Skript zur Vorlesung

Lineare Algebra I

gehalten von Prof. Dr. Markus Schweighofer
im Wintersemester 2021/2022
an der Universität Konstanz

entstanden aus einer

früheren elektronischen Mitschrift

von Michael Strecke,

erweitert um die lineare Algebra II von David Jannack

und editiert von Tom Folgmann

Fassung vom
14. März 2023, 21:23 Uhr

Vorwort des Originalskripts zur linearen Algebra 1

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.

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. 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 schöne, jedoch von mir nicht überprüfte elektronische Mitschriften zu meiner Vorlesung.

Im Wintersemester 2017/2018 las ich die Vorlesung mit nur geringfügigen Änderungen zum dritten Mal. Ich habe bei dieser Gelegenheit das Skript von Michael Strecke zu einem von mir autorisierten Skript umgebaut, um darauf verweisen zu können.

Im Wintersemester 2021/2022 lese ich die Vorlesung nun zum vierten Mal. Aufgrund der COVID-19-Pandemie wird es dieses Mal zusätzlich zum Skript sogar eine Echtzeitübertragung und eine Aufzeichnung der Vorlesung geben, obwohl man zumindest zur Zeit den Hörsaal wieder eingeschränkt betreten kann. Die Vorlesungszeit des Wintersemesters wurde mittlerweile um eine Woche verkürzt, so dass ich zusätzlich einige Teile in zusätzlichen Lernvideos erklären werde. All diese Videos werden bereitgestellt auf einer YouTube-Playlist:

`https://youtube.com/playlist?list=PLbQ93L5pV-a_iL7cSoU9ZZrj3KH97NObn`

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 mache ich verfügbar unter:

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

Skript zur linearen Algebra I & II

Inhaltsverzeichnis

Skript zur linearen Algebra I & II

§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 **??** 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 **??** 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 **??**. 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 $f\colon A \to B$.

Bemerkung 1.1.17. Sind $f\colon A \to B$ 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.

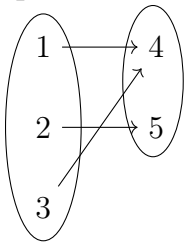
Für $f\colon \{0,1\} \to \{0,1\}, x \mapsto x$
$g : \{0,1\} \to \{0,1\}, x \mapsto x^2$ und
$h : \{0,1\} \to \{0,1\}, 0 \mapsto 0, 1 \mapsto 1$ gilt
$f = g = h$, aber $f \neq p$ für $p : \{0,1\} \to \{0,1,2\}, 0 \mapsto 0, 1 \mapsto 1$

Definition 1.1.19. Eine Abbildung $f\colon A \to B$ 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:

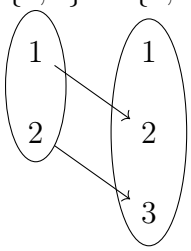
$$\begin{aligned} f \text{ injektiv} &\iff \forall a_1, a_2 \in A : (f(a_1) = f(a_2) \Longrightarrow a_1 = a_2) \\ f \text{ surjektiv} &\iff \forall b \in B : \exists a \in A : f(a) = b \text{ und} \\ f \text{ bijektiv} &\iff (f \text{ injektiv } \& \ f \text{ surjektiv}) \end{aligned}$$

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



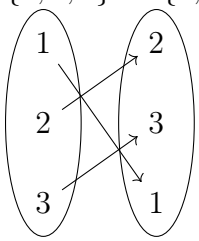
nicht injektiv, surjektiv, nicht bijektiv

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



injektiv, nicht surjektiv, nicht bijektiv

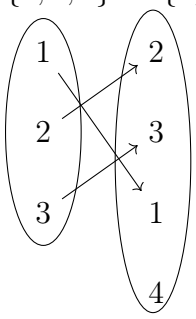
(c) $\{1,2,3\} \to \{1,2,3\}, x \mapsto x$



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

Bis hierher sollten wir am 25. Oktober kommen.

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 $f\colon \mathbb{N} \to A$ 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 $f\colon A \to \mathscr{P}(A)$ 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 $f(a) = B$. $\lightning$ „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 $A = \mathbb{N}$:

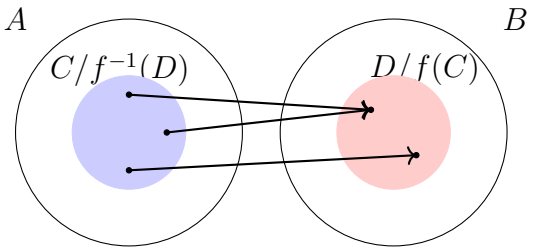
$f(1) = \{\ \underline{1},\ 3,\ 4,\ 7,\ 9,\ 10,\ \ldots\}$
$f(2) = \{\ \underline{2},\ 4,\ 5,\ 8,\ 10,\ \ldots\}$
$f(3) = \{\ 2,\ \_,\ 4,\ 6,\ 7,\ 8,\ 9,\ \ldots\}$
$\_ = \mathbb{N} \setminus B$ „Cantors Diagonalargument“

Bemerkung 1.1.23. Für endliche Mengen A folgt **??** 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 **??** eine bijektive Abbildung $\mathbb{N} \to \mathscr{P}(\mathbb{N})$. Dies ist nach dem Satz von Cantor **??** 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	Mengeneinbettung/ 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 **??** 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:

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

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 [→ **??**] und auch das Bild [→ **??** (d)], nicht aber die Zielmenge von f zurückgewinnen.

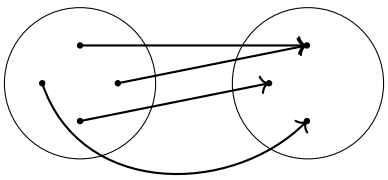
$$A = \{a \mid \exists b : (a, b) \in \Gamma_f\}$$

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

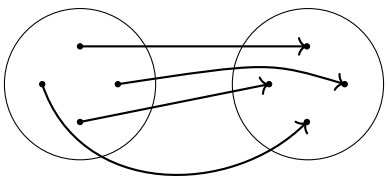
$$f(A) = \{b \mid \exists a : (a, b) \in \Gamma_f\} \subseteq B$$

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 [→ **??**].
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 **??** 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.

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* $\left\{ \begin{matrix} \textit{schaltung} \\ \textit{ausführung} \end{matrix} \right\}$ (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

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

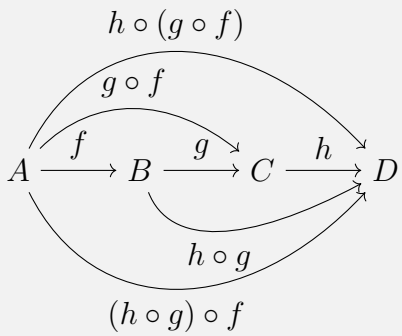
(„$\circ$ ist assoziativ“).

(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 $f \circ f^{-1} = \mathrm{id}_B$.

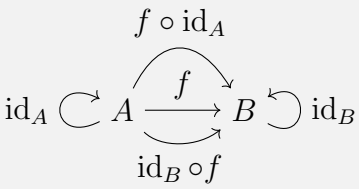
(d) Für bijektive $f : A \to B$ ist auch f^{-1} bijektiv und es gilt $(f^{-1})^{-1} = f$.

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



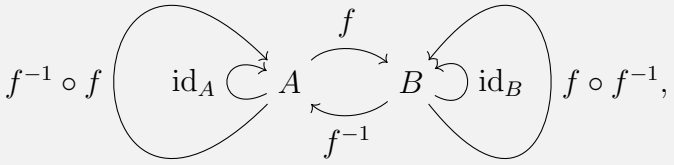
und$(h\circ(g\circ f))(a) = h((g\circ f)(a)) = h(g(f(a))) = (h\circ g)(f(a)) = ((h\circ g)\circ f)(a)$ für alle $a \in A$. Nach **??** gilt also $h \circ (g \circ f) = (h \circ g) \circ f$.

(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 **??** gilt also $f \circ \mathrm{id}_A = f = \mathrm{id}_B \circ f$.

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



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

Nach **??** 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 $\{b \in B \mid f^{-1}(b) = a\} \overset{??}{=} \{b \in B \mid f(a) = b\} = \{f(a)\}$, d.h. jedes Element von A hat genau ein Urbild unter f^{-1}. Weiter gilt $(f^{-1})^{-1} : A \to B$ und

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

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

Bis hierher sollten wir am 28. Oktober kommen.

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

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

Analog folgt $f = g^{-1}$.

Sprechweise und Notation 1.2.7. Die Situation von **??** 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. [→??] 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 $\bigcup(A/\sim) \overset{??}{=} \{a \mid \exists B \in A/\sim : a \in B\} \overset{??(a)}{=} \{a \mid \exists b \in A : a \in \widetilde{b}\}$. 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 **??** 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 **??** entsprechen sich die Äquivalenzrelation $\sim$ auf $\mathbb{Z}$ definiert durch

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

und die Zerlegung

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

Bis hierher sollten wir am 4. November kommen.

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 **??**(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 **??** gilt.

[Zeichne Bild!]

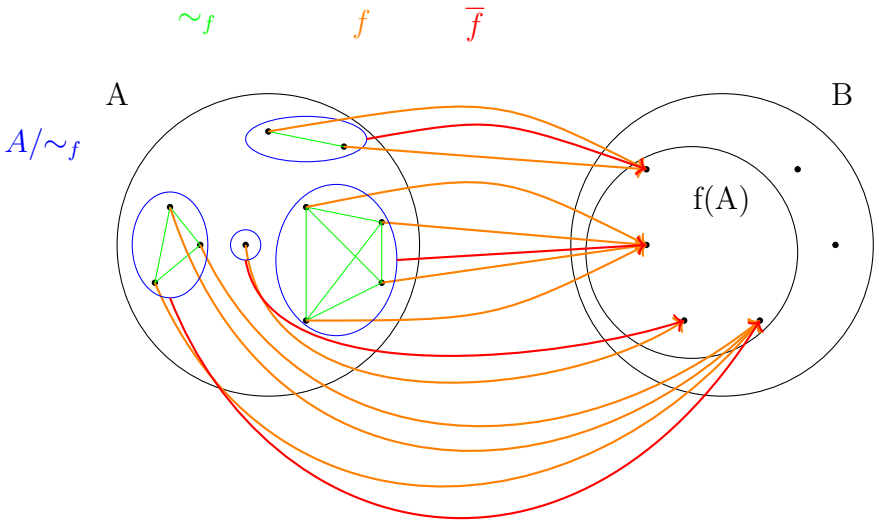
Definition und Proposition 1.3.9. Sei $f\colon A \to B$ eine Abbildung. Dann wird durch

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

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.



Skript zur linearen Algebra I & II

§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 $a \in G$. 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 $\left\{\begin{array}{c} ab = a \cdot b \\ 1 \\ a^{-1} \end{array}\right\}$ statt $\left\{\begin{array}{c} a + b \\ 0 \\ -a \end{array}\right\}$, 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“.

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:

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

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

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

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 [→ **??** (d)]) abelsche Gruppe, nicht jedoch $\{0,1\}, \mathbb{Q}$ oder $\mathbb{R}$.

Proposition 2.1.4. Sei G eine abelsche Gruppe [→ **??** (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 [→ **??** (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

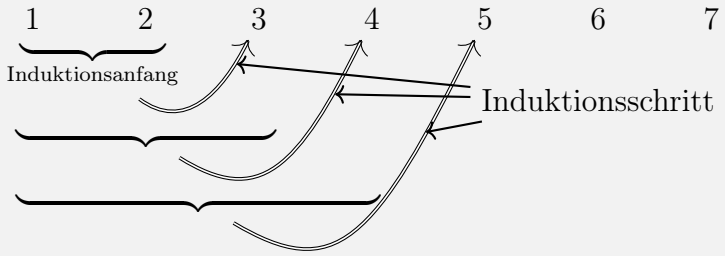
$$\begin{aligned}(a+b)+((-a)+(-b)) &\overset{(A)}{=} a+(b+((-a)+(-b))) \overset{(K)}{=} a+(b+((-b)+(-a)))\\ &\overset{(A)}{=} a+((b+(-b))+(-a)) \overset{(I)}{=} a+(0+(-a))\\ &\overset{(K)}{=} a+((-a)+0) \overset{(N)}{=} a+(-a) \overset{(I)}{=} 0.\end{aligned}$$

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

Bis hierher sollten wir am 8. November kommen.

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).



$n \in \{1, 2\}$ klar.
$1, 2, \ldots, n-1 \to n \ (n \geq 3)$: 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

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

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

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

Satz 2.1.9 (über Umordnung)**.** Sei A eine Menge, $\varrho : A \times A \to A$ assoziativ [→ **??**] 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 $n \in \mathbb{N}, a_1, \dots, a_n \in A$ und $\sigma \in S_n : a_1 \varrho \ \dots \ \varrho\, a_n = a_{\sigma(1)} \varrho \ \dots \ \varrho\, a_{\sigma(n)}$.

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

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

Notation 2.1.10. Sei G eine abelsche Gruppe [→ **??** (e)]. Ist $(a_i)_{i \in I}$ eine Familie in G (d.h. $I \to G, \ i \mapsto a_i$ eine Abbildung [→ **??** (c)]) und $n := \#I \in \mathbb{N}$, so gilt

$$a_{\sigma(1)} + \dots + a_{\sigma(n)} = a_{\tau(1)} + \dots + a_{\tau(n)}$$

für alle Bijektionen $\sigma, \tau : \{1, \dots, n\} \to I$ [→ **??**, **??**] (denn $\sigma^{-1} \circ \tau \in S_n$ und daher ist $b_1 + \dots + b_n = b_{\sigma^{-1}(\tau(1))} + \dots + b_{\sigma^{-1}(\tau(n))}$ für $b_1 = a_{\sigma(1)}, \dots, b_n = a_{\sigma(n)}$) und wir notieren dieses Element von G mit

$$\sum_{i \in I} a_i.$$

Wir setzen $\sum_{i \in \emptyset} a_i := 0$. Statt $\sum_{i \in \{m, \dots, n\}} a_i$ schreibt man auch $\sum_{i=m}^{n} a_i$. Beachte $\sum_{i=1}^{0} a_i \overset{??}{=} \sum_{i \in \emptyset} = 0$.

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

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

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“).

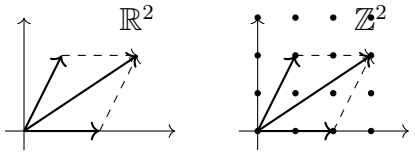
Beweis. Übung.

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

$$(a_1, \ldots, a_n) + (b_1, \ldots, b_n) := (a_1 +_1 b_1, \ldots, a_n +_n b_n)$$

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 $0 = (0_1, \ldots, 0_n)$ und $-(a_1, \ldots, a_n) = (-_1 a_1, \ldots, -_n a_n)$ 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') $\forall a \in H : -_H a = -_G a$

Beweis. Wir zeigen zunächst:

$(*)$ $((a)\&(b)\&(c)\&(d)) \implies H$ ist Trägermenge einer Untergruppe von $(G, +_G)$.

Gelte hierzu (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). Man sieht sofort, dass $(H, +_H)$ wegen (a) die Eigenschaften (K) und (A) aus **??** „erbt“. Gleiches gilt für (N) und (I) wegen (b) und (d). Damit ist $(*)$ gezeigt.
Für den Rest des Beweises sei H Trägermenge einer Untergruppe von $(G, +_G)$. Dann gibt es eine Gruppenaddition $+_H : H \times H \to H$ derart, dass $(H, +_H)$ eine Untergruppe von $(G, +_G)$ ist. Nach **??** gelten (a) und (c'). Aus (c') folgt weiter, dass es *genau* eine solche Gruppenaddition gibt. Es bleiben nur noch (b') und (d') zu zeigen, denn (b') $\implies$ (b), (c') $\implies$ (c) und (d') $\implies$ (d).
Zu (b'). Es gilt: $0_H +_G 0_H \overset{\text{(c')}}{=} 0_H +_H 0_H \overset{\text{(N) für } H}{=} 0_H$.
Daraus folgt

$$0_H \underset{\text{für } G}{\overset{\text{(N)}}{=}} 0_H +_G 0_G \underset{\text{für } G}{\overset{\text{(I)}}{=}} 0_H +_G (0_H +_G (-_G 0_H)) \underset{\text{für } G}{\overset{\text{(A)}}{=}} (0_H +_G 0_H) +_G (-_G 0_H) = 0_H +_G (-_G 0_H) \underset{\text{für } G}{\overset{\text{(I)}}{=}} 0_G.$$

Zu (d'). Sei $a \in H$. Dann ist $a +_G (-_H a) \overset{\text{(c')}}{=} a +_H (-_H a) \underset{\text{für } H}{\overset{\text{(I)}}{=}} 0_H \overset{\text{(b')}}{=} 0_G$ und daher $-_H a = -_G a$ nach der Definition von $-_G a$ in **??**(b).

Bemerkung 2.2.3. Ist $(H, +_H)$ eine Untergruppe der abelschen Gruppe $(G, +)$, so schreibt man wegen (b'), (c') und (d') fast immer $+, -, 0$ statt $+_H, -_H, 0_H$. Daher erwähnt man die Gruppenaddition oft nicht mehr explizit und spricht zum Beispiel von einer „Untergruppe H der (additiv geschriebenen) 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. Dies ist klar für $\#G \leq 2$, für $\#G \leq 3$ betrachte man die Additionstabelle aus **??**(d).

(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. Hierzu ist zu zeigen: $H \subseteq \mathscr{P}(A)$, $0 \in H$, $\forall B, C \in H : B + C \in H$ und

$\forall B \in H : -B \in H$. Wegen $0 = \emptyset$ und $-B = B$ für $B \in H$ ist nur $(B \Delta C) \cap X = \emptyset$ für alle $B, C \in H$ zu zeigen, was einfach ist.

(c) Folgende Inklusionen sind Untergruppenbeziehungen:

$$\{10n \mid n \in \mathbb{Z}\} \subseteq \left\{ \begin{matrix} \{5n \mid n \in \mathbb{Z}\} \\ \{2n \mid n \in \mathbb{Z}\} \end{matrix} \right\} \subseteq \mathbb{Z} \subseteq \left\{ \begin{matrix} \left\{\frac{n}{3} \mid n \in \mathbb{Z}\right\} \\ \left\{\frac{n}{2} \mid n \in \mathbb{Z}\right\} \end{matrix} \right\} \subseteq \left\{\frac{n}{6} \mid n \in \mathbb{Z}\right\} \subseteq \mathbb{Q},$$

$$\mathbb{Q} \subseteq \left\{p + q\sqrt{2} \mid p, q \in \mathbb{Q}\right\} \subseteq \mathbb{R},$$

$$\{(0,0)\} \subseteq \{(m,0) \mid m \in \mathbb{Z}\} \subseteq \{(m,2n) \mid m, n \in \mathbb{Z}\} \subseteq \mathbb{Z} \times \mathbb{Z},$$

$$\mathbb{Z} \times \mathbb{Z} \subseteq \left\{ \begin{matrix} \mathbb{Q} \times \mathbb{Z} \\ \mathbb{Z} \times \mathbb{Q} \end{matrix} \right\} \subseteq \mathbb{Q} \times \mathbb{Q} \subseteq \mathbb{Q} \times \mathbb{R} \subseteq \mathbb{R} \times \mathbb{R}.$$

(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 (mit $\bigcap \emptyset := G$) [→ **??**].

Beweis. Wir verwenden im Folgenden, dass

$$\bigcap M = \{a \in G \mid \forall I \in M : a \in I\},$$

was man sofort durch Fallunterscheidung nach den Fällen $M = \emptyset$ und $M \neq \emptyset$ zeigt. Zu zeigen:

(a) $\bigcap M \subseteq G$

(b) $0 \in \bigcap M$

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

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

Zu (a). Œ $M \neq \emptyset$. Sei $a \in \bigcap M$. Wähle $H \in M$ (beachte $M \neq \emptyset$). Dann $a \in H \subseteq G$, also $a \in G$.
Zu (b). Sei $H \in M$. Zu zeigen: $0 \in H$. Klar, da H Untergruppe.
Zu (c). Seien $a, b \in \bigcap M$. Zu zeigen: $a+b \in \bigcap M$. Sei $H \in M$. Zu zeigen: $a+b \in H$. Wissen $a, b \in H$. Da H Untergruppe, folgt $a + b \in H$.
Zu (d). Sei $a \in \bigcap M$. Zu zeigen: $-a \in \bigcap M$. Sei $H \in M$. Zu zeigen: $-a \in H$. Wissen $a \in H$. Da H Untergruppe, folgt $-a \in H$.

Satz und Definition 2.2.6. Sei G eine abelsche Gruppe und $E \subseteq G$. Dann existiert die kleinste Untergruppe H von G mit $E \subseteq H$. (d.h. H ist eine Untergruppe von G mit $E \subseteq H$ und wenn H' auch eine solche ist, so gilt $H \subseteq H'$). Man nennt sie die von E in G *erzeugte* Untergruppe von G und notiert sie mit $\langle E \rangle_G$.

Beweis. Für $M := \{I \mid I \text{ Untergruppe von } G, E \subseteq I\}$ ist $H := \bigcap M$ ist eine Untergruppe von G nach **??**. Offensichtlich gilt $E \subseteq H$. Schließlich ist $H \subseteq H'$ zu zeigen für alle $H' \in M$. Dies ist aber trivial wegen $H = \bigcap M$.

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. (gleichzeitig alternativer Beweis für Satz **??**!) Man überlegt sich leicht, dass die rechtsstehende Menge eine Untergruppe von G ist, die E enthält. Sei H eine weitere Untergruppe von G mit $E \subseteq H$. Zu zeigen ist, dass die rechtsstehende Menge in H enthalten ist, was aber trivial ist.

Beispiel 2.2.8. $\langle \{3, 2\} \rangle_{\mathbb{Z}} = \mathbb{Z}$, $\langle \{12, 16\} \rangle_{\mathbb{Z}} = \{4n \mid n \in \mathbb{Z}\}$

Definition 2.2.9. Seien $(G, +_G)$ und $(H, +_H)$ abelsche Gruppen. Dann heißt f ein (Gruppen-)Homomorphismus von $(G, +_G)$ nach $(H, +_H)$, wenn $f \colon G \to H$ eine Abbildung ist mit

$$\forall a, b \in G : f(a +_G b) = f(a) +_H f(b).$$

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

Proposition 2.2.11. Seien $(G, +_G)$ und $(H, +_H)$ abelsche Gruppen und sei f ein Homomorphismus von $(G, +_G)$ nach $(H, +_H)$. Dann gilt $f(0_G) = 0_H$ und $f(-_G a) = -_H f(a)$ für alle $a \in G$.

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, ist $f(a) +_H f(-_G a) = 0_H$ zu zeigen. Nun gilt aber $f(a) +_H f(-_G a) = f(a +_G (-_G a)) = f(0_G) = 0_H$.

Notation 2.2.12. [→ **??**(c)] Ein Gruppenhomomorphismus $f : G \to H$ heißt (Gruppen-)$\left\{\begin{array}{c}\textit{Einbettung oder Mono-}\\ \textit{Epi-}\\ \textit{Iso-}\end{array}\right\}$*morphismus*, wenn f $\left\{\begin{array}{c}\text{injektiv}\\ \text{surjektiv}\\ \text{bijektiv}\end{array}\right\}$ ist. Einen Gruppenhomomorphismus $f : G \to G$ nennt man auch einen (Gruppen-)*Endomorphismus* von G und, falls er bijektiv ist, (Gruppen-)*Automorphismus* von G.

Beispiel 2.2.13.

	Hom.	Einb.	Epi.	Iso.	Endo.	Auto.
$\mathbb{Z} \to \mathbb{Z}, a \mapsto 2a$	✓	✓	—	—	✓	—
$\mathbb{Z} \to \mathbb{Q}, a \mapsto 2a$	✓	✓	—	—	—	—
$\mathbb{Q} \to \mathbb{Q}, a \mapsto 2a$	✓	✓	✓	✓	✓	✓
$\mathbb{Z} \to \mathbb{Z}, a \mapsto -a$	✓	✓	✓	✓	✓	✓
$\mathbb{Z} \times \mathbb{Z} \to \mathbb{Z}, (a,b) \mapsto a$	✓	—	✓	—	—	—
$\mathbb{Z} \times \mathbb{Z} \to \mathbb{Z}, (a,b) \mapsto a+b$	✓	—	✓	—	—	—
$\mathbb{Z} \to \mathbb{Z}, a \mapsto a+1$	—	—	—	—	—	—
$\mathbb{Z} \to \mathbb{Z} \times \mathbb{Z}, a \mapsto (a,a)$	✓	✓	—	—	—	—
$\mathbb{Z} \to \mathbb{Z} \times \mathbb{Z}, a \mapsto (a,0)$	✓	✓	—	—	—	—
$\mathbb{R} \setminus \{0\} \to \mathbb{R}_{>0}, \quad a \mapsto \lvert a \rvert$ (jeweils mit Multiplikation)	✓	—	✓	—	—	—

Proposition 2.2.14. (a) Seien G, H, I abelsche Gruppen. Sind $G \xrightarrow{f} H \xrightarrow{g} I$ Gruppenhomomorphismen, so auch $g \circ f$.

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

Beweis. (a) Seien $G \xrightarrow{f} H \xrightarrow{g} I$ Gruppenhomomorphismen und $a, b \in G$. Dann

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

(b) Sei $f : G \to H$ ein Gruppenisomorphismus. Nach **??**(d) ist f^{-1} bijektiv. Es ist noch zu zeigen, dass f^{-1} ein Homomorphismus ist. Seien hierzu $a, b \in H$. Da f injektiv ist, reicht es $f(f^{-1}(a+b)) = f(f^{-1}(a) + f^{-1}(b))$ zu zeigen. Es gilt aber $f(f^{-1}(a+b)) = (f \circ f^{-1})(a+b) \overset{??(c)}{=} \mathrm{id}_H(a+b) = a+b$ und $f(f^{-1}(a)+f^{-1}(b)) \overset{f \text{ Hom.}}{=} f(f^{-1}(a)) + f(f^{-1}(b)) = (f \circ f^{-1})(a) + (f \circ f^{-1})(b) = \mathrm{id}_H(a) + \mathrm{id}_H(b) = a + b$.

Bis hierher sollten wir am 11. November kommen.

Definition und Bemerkung 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:

$+$	$\cdots$	b	$\cdots$
$\vdots$		$\vdots$	
a	$\cdots$	$a+b$	$\cdots$
$\vdots$		$\vdots$	

$\rightsquigarrow$

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

Er tauscht die Elemente aus, ohne die „Gruppenstruktur“ zu verändern (das heißt die

über die Addition geregelten Beziehungen der Elemente untereinander). Alle „strukturellen“ (das heißt nicht auf die Natur der Elemente bezogenen) Eigenschaften einer abelschen Gruppe übertragen sich daher unter Isomorphismen. Dass zwei abelsche Gruppen isomorph sind, heißt, dass ihre Additionstabellen dieselbe Form haben.

Beispiel 2.2.16. (a) Seien $(G, +_G)$ und $(H, +_H)$ abelsche Gruppen mit $\#G = 3 = \#H$. Nach **??**(d) kann man dann a, b, c, d mit $G = \{0_G, a, b\}$ und $H = \{0_H, c, d\}$ finden so, dass die Additionstabellen von $(G, +_G)$ und $(H, +_H)$ wie folgt lauten:

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

und

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

Daher ist $f : G \to H$, $\begin{array}{c} 0_G \mapsto 0_H \\ a \mapsto c \\ b \mapsto d \end{array}$ ein Isomorphismus von $(G, +_G)$ nach $(H, +_H)$, also $(G, +_G) \cong (H, +_H)$. Man sagt, dass es „*bis auf Isomorphie*“ genau eine dreielementige abelsche Gruppe gibt. Dasselbe gilt nach **??** für ein- und zweielementige abelsche Gruppen.

(b) Sei G eine abelsche Gruppe mit $\#G \leq 3$. Ist $\#G \leq 2$, so ist id_G der einzige Automorphismus von G (für $\#G = 2$ folgt dies aus **??**). Ist $\#G = 3$, so gibt es genau zwei Automorphismen von G, wie man an der Additionstabelle in **??**(d) leicht erkennt: id_G und die Permutation von G, die die zwei nichtneutralen Elemente vertauscht.

Sprechweise 2.2.17. „Isomorphismus = Umbenennung der Elemente“

2.3 Quotientengruppen[→ §??]

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

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

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

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 $G/\equiv$ 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 **??** geklärt. Wir prüfen die Axiome (K), (A), (N) und (I) aus **??** 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 **??** 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. [→ **??**(b)] Sei G eine abelsche Gruppe. Zu jeder Untergruppe H von G definieren wir eine Relation $\equiv_H$ auf G durch

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

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

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

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 **??** gezeigt.

Zu (b). Sei H eine Untergruppe von G. Zu zeigen ist, dass $\equiv_H$ eine Äquivalenzrelation ist [→ **??**] und dass $(*)$ aus **??** gilt. Also ist zu zeigen:

(1) $\forall a \in G : a \equiv_H a$

(2) $\forall a, b \in G : (a \equiv_H b \implies b \equiv_H a)$

(3) $\forall a, b, c \in G : ((a \equiv_H b\ \&\ b \equiv_H c) \implies a \equiv_H c)$

(4) $\forall a, a', b, b' \in G : ((a \equiv_H a'\ \&\ b \equiv_H b') \implies a + b \equiv_H a' + b')$.

Zu (1). Ist $a \in G$, dann $a - a = 0 \in H$ und damit $a \equiv_H a$.
Zu (2). Seien $a, b \in G$ mit $a \equiv_H b$. Dann $a - b \in H$ und daher $b - a = -(a - b) \in H$.
Also $b \equiv_H a$.

Zu (3). Seien $a, b, c \in G$ mit $a \equiv_H b$ und $b \equiv_H c$. Dann $a - b \in H$ und $b - c \in H$. Also $a - c = (a - b) + (b - c) \in H$.
Zu (4). Seien $a \equiv_H a'$ und $b \equiv_H b'$. Dann $a - a' \in H$ und $b - b' \in H$. Also $a + b - (a' + b') = (a - a') + (b - b') \in H$, das heißt $a + b \equiv_H a' + b'$.

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:

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

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

$$\overset{-H}{0} \overset{??(b)}{=} \{a \in G \mid a \equiv_H 0\} \overset{??}{=} \{a \in G \mid a - 0 \in H\} = H.$$

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 $\overset{-H}{a}$ $(a \in G)$ [→ **??**] von $\equiv_H$ nennen wir auch die *Nebenklassen* von H (in G). Wegen **??**(**??**) haben alle Nebenklassen von H dieselbe Mächtigkeit [→ **??**] und da sie eine Zerlegung [→ **??**] 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

$$\mathbb{Z}/\langle n \rangle = \left\{ \overset{-\langle n \rangle}{a} \mid a \in \mathbb{Z} \right\}, \quad \text{wobei}$$

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

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

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

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 $\equiv_f := \sim_f$ [→??] eine Kongruenzrelation auf G.

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

(c) Unter der Bijektion aus Satz **??** 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 **??** 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

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

(b) Nach **??** 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 **??**. 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 $f(-a) \overset{??}{=} -f(a) = -0 \overset{0+0=0}{=} 0$ und daher $-a \in \ker f$.

(c) Wegen Satz **??** sind die beiden Gleichheiten äquivalent. Es reicht daher eine der beiden zu zeigen. Wir zeigen die erste. Diese ergibt sich wie folgt:

$$\ker f \overset{(b)}{=} f^{-1}(\{0\}) \overset{??}{=} \{a \in G \mid f(a) = 0\} \overset{??}{=} \{a \in G \mid f(a) = f(0)\} \overset{??}{=} \{a \in G \mid a \equiv_f 0\} \overset{??(b)}{=} \overline{0}^{f}.$$

Bis hierher sollten wir am 15. November kommen.

Satz 2.3.11 (Homomorphiesatz für abelsche Gruppen)**.** [→??] 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 $\overline{f}(\overline{a}^{I}) = f(a)$ 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 **??** (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 **??** und damit $f(a) - f(b) \overset{f \text{ Hom.}}{=} f(a-b) = 0$. Unter Beachtung von $G/I \overset{??}{=} G/\equiv_I$ erhält man die in (a) behauptete Existenz und Eindeutigkeit der Abbildung $\overline{f}$ daher aus dem Homomorphiesatz für Mengen **??**. Dass $\overline{f}$ ein Homomorphismus ist, rechnet man sofort nach:

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

für alle $a, b \in G$. Damit ist (a) gezeigt. Die Aussage (c) folgt direkt aus **??**(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 **??**(b) ist.

Bemerkung 2.3.12. [→**??**] 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* [→**??**]

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

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 $\operatorname{im} f := f(G) \overset{??}{=} \{f(a) \mid a \in G\}$ für das in **??**(d) eingeführte Bild von f. Es ist $\operatorname{im} f$ eine Untergruppe von H.

Beweis. Zu zeigen sind gemäß **??**:

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

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

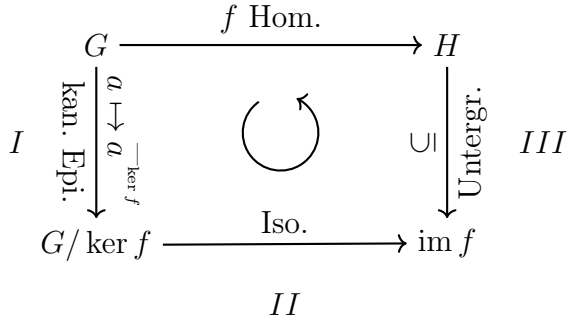
(c) $\forall a, b \in \operatorname{im} f : a + b \in \operatorname{im} f$

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

(a) ist trivial und (b) folgt aus $f(0) \overset{??}{=} 0$. 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 $f(-c) \overset{??}{=} -f(c) = -a$ 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)**.** [→**??**] 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 $\overline{f}(\overline{a}^{\ker f}) = f(a)$ für $a \in G$ ein Isomorphismus [→**??**]. Insbesondere $G/\ker f \cong \operatorname{im} f$ [→**??**].

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



Drei Phasen:

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

Skript zur linearen Algebra I & II

§3 Kommutative Ringe

[Julius Wilhelm 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 ($\dot{\text{N}}$) 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 **??** 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 $\left\{\begin{matrix}A \\ (A, +)\end{matrix}\right\}$ die *zugrundeliegende* $\left\{\begin{matrix}\textit{Menge} \\ \textit{abelsche Gruppe}\end{matrix}\right\}$ oder die $\left\{\begin{matrix}\textit{Trägermenge} \\ \textit{additive Gruppe}\end{matrix}\right\}$ und $\left\{\begin{matrix}+ \text{ die Addition} \\ \cdot \text{ die Multiplikation}\end{matrix}\right\}$ 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“ [→ **??** (e)].

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

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

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

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 $a \in A$

$$a \overset{(\dot{\mathrm{N}})}{=} a \cdot 1_A = a \cdot 0_A \overset{\text{?? (g)}}{=} 0_A,$$

also $A = \{0_A\} = \{1_A\}$.

Beispiel 3.1.4. [→ **??**]

(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. Es gilt $1 = A$.

(d) Genauso wie man in **??** 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 **??** schon, dass $(A^{\mathbb{N}_0}, +)$ eine abelsche Gruppe ist. Man rechnet nun als Übung (K̇), (Ȧ),(Ṅ) und (D) nach.

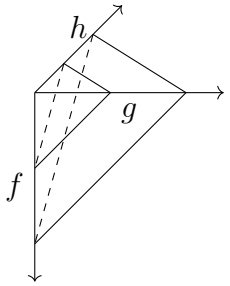
	$g(0)$	$g(1)$	$g(2)$	$\cdots$
$f(0)$	$f(0)g(0)$	$f(0)g(1)$	$f(0)g(2)$	$\cdots$
$f(1)$	$f(1)g(0)$	$f(1)g(1)$	$\cdots$	$\cdots$
$f(2)$	$f(2)g(0)$	$\cdots$	$\cdots$	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

Faltung bildet Summen über die Diagonalen.

(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 bildet Summen über die Raumdiagonalen.

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

3.2 Unterringe, Ringhomomorphismen und Polynome

Definition 3.2.1. [→ **??**] 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 $B \subseteq A, \underline{\underline{1_A \in B}}, \forall a, b \in B : a +_B b = a +_A b$ und $\forall a, b \in B : a \cdot_B b = a \cdot_A b$.

Proposition 3.2.2. [→ **??**] Sei $(A, +_A, \cdot_A)$ ein kommutativer Ring und B eine Menge. Genau dann ist B Trägermenge [→ **??** (c)] eines Unterrings von $(A, +_A, \cdot_A)$, wenn B Trägermenge einer Untergruppe von $(A, +_A)$ ist [→ **??**] 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 **??**.

Bis hierher sollten wir am 18. November kommen.

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 $B \cup \{x\} \subseteq B[x]$.

(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) $B \cup \{x\} \subseteq B[x] \subseteq A$

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

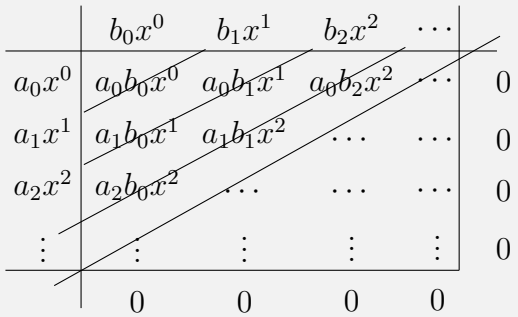
(3) $\forall a \in B[x] : -a \in B[x]$

Zu (1). Es gilt $a = a \cdot x^0 \in B[x]$ für alle $a \in B$ und $x = 1 \cdot x^1 \in B[x]$. Klar ist auch $B[x] \subseteq A$.

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 $\max\{m, n\} :=$ größtes Element der Menge $\{m, n\}$:

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

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



(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 [→ **??** (b)], der $\mathbb{Q} \cup \left\{\sqrt{2}\right\}$ enthält und $\mathbb{Q}[\sqrt{2}]$ der kleinste solche ist [→ **??**].

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* in 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 kommutativer Ring [→ **??**, **??** (a)], so ist A ein Polynomring über sich selber in 0.

Definition 3.2.8. [→ **??**] 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 $\underline{\underline{f(1_A) = 1_B}}$ und $\forall a, b \in A : f(a \cdot_A b) = f(a) \cdot_B f(b)$.

Beispiel 3.2.9. (a) $f : \mathbb{R} \to \mathbb{R}, x \mapsto 0$ 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 **??** (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. [→ **??**] Ein Ringhomomorphismus $f : A \to B$ heißt (Ring-) $\left\{\begin{array}{r} \textit{Einbettung oder Mono} \\ \textit{Epi-} \\ \textit{Iso-} \end{array}\right.$ wenn $f \left\{\begin{array}{c} \text{injektiv} \\ \text{surjektiv} \\ \text{bijektiv} \end{array}\right\}$ ist. Einen Ringhomomorphismus $f : A \to A$ nennt man auch einen *(Ring-)Endomorphismus* von A und, falls er bijektiv ist, *(Ring-)Automorphismus* von A.

Proposition 3.2.11. [→**??**] 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 **??** wissen wir schon, dass $\operatorname{im} f$ eine Untergruppe der additiven Gruppe von B ist. Zu zeigen sind dann gemäß **??** 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 **??** eines Ringhomomorphismus gilt $f(1) = 1$. Um (b) zu zeigen, seien $a, b \in \operatorname{im} f$. Wähle dann $c, d \in A$ mit $f(c) = a$ und $f(d) = b$. Es folgt $f(cd) \overset{f \text{ Hom.}}{=} f(c)f(d) = ab$ und damit $ab \in \operatorname{im} f$.

Satz 3.2.12. 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 **??**. Es ist $\varphi : A \to A^{\mathbb{N}_0}, a \mapsto (a, 0, 0, 0, \ldots)$ [→ **??** (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 $A^{\mathbb{N}_0}$ gemäß **??**. Da die Behauptung eine *strukturelle* Aussage über A macht (**??** 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 **??** 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 $A'[x]$ [→ **??**] von $A^{\mathbb{N}_0}$ ein Polynomring über A' in x ist [→ **??**]. Seien hierzu $n \in \mathbb{N}_0$ und $a_0, \ldots, a_n \in A$ mit

$$\sum_{k=0}^{n} \underbrace{\widehat{\varphi}(a_k)x^k}_{\widehat{\varphi}(a_k)*x*\ldots*x} = 0.$$

Zu zeigen: $\widehat{\varphi}(a_0) = \ldots = \widehat{\varphi}(a_n) = 0$. Durch Induktion nach $k \in \mathbb{N}_0$ zeigt man

$$x^k = (\underbrace{0, 0, \ldots, 0}_{k \text{ Nullen}}, 1, 0, 0, \ldots).$$

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 $a_0 = \ldots = a_n = 0$ und daher $\widehat{\varphi}(a_0) = \ldots = \widehat{\varphi}(a_n) = 0$.

Bemerkung 3.2.13. Schreibt man $A[X]$ mit großem X, so meint man meist stillschweigend, dass $A[X]$ ein Polynomring über A in X ist.

Satz 3.2.14. 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. Dass man ψ so definieren kann, liegt an der Eindeutigkeit der Koeffizienten eines Polynoms [→ **??**]. Es bleibt für das so definierte ψ 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 (d):

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

Zu (e):

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

Korollar 3.2.15. 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 3.2.16. In der Situation von **??** schreibt man auch $p(x)$ („p ausgewertet in 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 3.2.17. 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 **??** 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 $\psi\left(\sum_k a_k X^k\right) = 0$ ($a_k \in A$). 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 3.2.18. Wegen **??** spricht man oft auch von <u>*dem*</u> Polynomring $A[X]$ über A in X.

3.3 Ideale und Quotientenringe [→ §??, §??]

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 [→**??**], 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 **??** 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 $A/\equiv$ [$\to$**??**] 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 **??** wissen wir schon, dass A bezüglich der Addition eine abelsche Gruppe bildet. Es sind daher nur noch ($\dot{\text{K}}$), ($\dot{\text{A}}$), ($\dot{\text{N}}$) und (D) aus Definition **??** nachzurechnen:

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

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

($\dot{\text{N}}$) 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 [$\to$**??**(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 **??** und **??**, denn wenn (**??**) gilt, so ist Bedingung **??**(d) automatisch erfüllt, da dann

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

für alle $a \in I$.

Satz 3.3.6. [→??] 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 **??** (oder **??**) schon, dass $\overline{0}$ eine Untergruppe von A ist. Gemäß Definition **??** 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 **??**$(*)$ folgt

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

das heißt $ab \in \overline{0}$.

Zu (b). Sei I eine Ideal von A. Wir wissen aus **??** schon, dass $\equiv_I$ ein Kongruenzrelation der additiven Gruppe von A ist. Es bleibt **??**$(*)$ 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 **??**.

Definition 3.3.7. [→??] 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 $\overline{a}^{I}$ $(a \in A)$ [→??] von $\equiv_I$ nennt man manchmal auch die *Restklassen* von I (in A).

Bis hierher sollten wir am 22. November kommen.

Proposition 3.3.8. [→??] 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 **??** wissen wir schon, dass $\bigcap M$ eine Untergruppe der additiven Gruppe von A ist. Nach Definition **??** 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. [→**??**] 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 **??** (benutze **??** statt **??**).

Satz 3.3.10. [→**??**] 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 **??** und stellt gleichzeitig einen neuen Beweis für **??** dar!

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

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

mit $n \in \mathbb{N}_0$ und $a_1, \ldots, a_n \in A$ *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

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

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 $\overset{-(9)}{a} = \overset{\equiv(9)}{a}$. 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. [→??] 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 **??** unter der Bijektion aus Satz **??** 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 **??**(b) gezeigt. Nach Definition **??** 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

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

also $ab \in \ker f$.

Satz 3.3.16 (Homomorphiesatz für kommutative Ringe)**.** [→??] 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 **??**. Dass $\overline{f}$ ein Gruppenhomomorphismus ist, wissen wir ebenfalls aus **??**. Wir rechnen nach, dass $\overline{f}$ ein Ringhomomorphismus [→??]

ist, wobei wir $\equiv := \equiv_I$ und damit $\overline{a} = \overline{a}^I = \overline{\overline{a}}^{\equiv_I}$ für alle $a \in A$ schreiben:

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

und

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

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

Bemerkung 3.3.17. [→**??**] 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 **??**.

Korollar 3.3.18 (Isomorphiesatz für kommutative Ringe)**.** [→**??**] 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 [→**??**]. Insbesondere $A/\ker f \cong \operatorname{im} f$ (in der zu **??** analogen Bedeutung).

Skript zur linearen Algebra I & II

§4 Körper

4.1 Definition und Beispiele von Körpern

Definition und Proposition 4.1.1. Sei $(A, +, \cdot)$ ein kommutativer Ring [→ **??**]. 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 $(ab)(a'b') \overset{??\,(e)}{=} (aa')(bb') = 1 \cdot 1 \overset{(\dot{N})}{=} 1$, also $ab \in A^\times$. Die Axiome (K), (A), (N) für $(A^\times, \cdot)$ [→ **??**] folgen aus den Axiomen $(\dot{K})$, $(\dot{A})$, $(\dot{N})$ für $(A, +, \cdot)$ [→ **??**], 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 $ba \overset{(\dot{K})}{=} ab = 1$, 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 $\overline{2}, \overline{5} \notin (\mathbb{Z}/(10))^\times$ (denn wäre etwa $\overline{2} \in (\mathbb{Z}/(10))^\times$, so wäre $\overline{5} = \overline{2}^{-1} \cdot \overline{2} \cdot \overline{5} = \overline{2}^{-1} \cdot 0 = 0$ in $\mathbb{Z}/(10)$)
$\overline{1} \cdot \overline{1} = 1$ in $\mathbb{Z}/(10)$, also $\overline{1} \in (\mathbb{Z}/(10))^\times$
$\overline{4} \cdot \overline{5} = \overline{6} \cdot \overline{5} = \overline{8} \cdot \overline{5} = 0$ in $\mathbb{Z}/(10)$, 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 kommutativer Ring $A = \{0\} = \{1\}$ [→ **??**, **??** (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: $n = 0$

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

Da $\mathbb{Z}$ kein Körper ist und $\mathbb{Z}/(n) \cong \mathbb{Z}$, ist $\mathbb{Z}/(n)$ auch kein Körper (vgl. **??**). 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. **??** (a)) ist kein Körper [$\to$ **??** (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 $s = 1 \iff t = n$ sowie $t = 1 \iff s = n$. Wir zeigen also: $t = n$ oder $s = n$.

Annahme: Weder $t = n$ noch $s = n$.
Dann $1 \leq t \leq n - 1$ und $1 \leq s \leq n - 1$.
Also $\overline{s}, \overline{t} \in \left\{\overline{1}, \ldots, \overline{n-1}\right\} = (\mathbb{Z}/(n))^\times$.
Dann $0 = \overline{n} = \overline{st} = \overline{s}\,\overline{t} \in (\mathbb{Z}/(n))^\times$, 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 **??**

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

woraus $\#(a) = n$ folgt wie gewünscht.

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

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

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 $s = t = 0$ ↯.

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

Bemerkung 4.1.9. Mit dem „Wissen“ über „Primfaktorzerlegungen“ aus der Schule ist **??** auch klar, aber wurde dort dieses „Wissen“ begründet? Mit **??** kann man es begründen! Wir werden dies aber in der Linearen Algebra II 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 $b \in A^\times$.

$$\frac{a}{b} := ab^{-1}$$

„a durch b“ $\qquad [\to$ **??** (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 $[\to$**??**$]$

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

(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 [→**??**]. In der Tat: φ ist ein Ringhomomorphismus [→ **??**] (nämlich die Einschränkung [→ **??**] des kanonischen Epimorphismus [→ **??**] 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 [→**??**] (X^2+1) einen Grad [→**??**] ≥ 2. Nun ist $\hat{\varphi}\colon K \to \operatorname{im}\varphi$ ein Ringisomorphismus und $K' := \operatorname{im}\varphi$ ein Unterring von $K[X]/(X^2+1)$ [→**??**]. Es reicht, die Behauptung für K' statt K zu zeigen, denn $K \cong K'$ (vergleiche **??** und Beweis von **??**). 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 **??** 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 damit natürlich $a = 0$.

(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: 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 [→**??**], denn

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

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

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.

Bis hierher sollten wir am 25. November kommen.

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 [→**??**(a)]. Wegen **??**(c) ist $K[\mathring{i}]$ im Wesentlichen 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 **??**(b) 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 $\mathring{\jmath} := -\mathring{i}$, denn $\mathring{\jmath}^2 = \mathring{\jmath}\mathring{\jmath} = (-\mathring{i})(-\mathring{i}) = -\mathring{i}(-\mathring{i}) = -(-\mathring{i}\mathring{i}) = \mathring{i}\mathring{i} = \mathring{i}^2 = -1$. Ist nun K ein Körper ohne imaginäre Einheit, so ist $K[\mathring{i}] = K[-\mathring{i}]$ und nach **??**(c) gibt es genau einen Ringautomorphismus [→**??**] φ von $K[\mathring{i}]$ 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 **??**(b) können wir $z = a + b\mathring{i}$ 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

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

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 $p(a) = 0$ [→**??**].

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 [→**??**] $(X - a)$ 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 $\overline{p} \overset{??}{=} \sum_{k=0}^{n} \overline{a_k}\overline{X}^k \overset{\overline{X}=\overline{a}}{=} \sum_{k=0}^{n} \overline{a_k}\,\overline{a}^k \overset{??}{=} \overline{p(a)} = \overline{0} = 0$.

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

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

(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 2.17 im Buch von Basu, Pollack und Roy: Algorithms in Real Algebraic Geometry, `https://perso.univ-rennes1.fr/marie-francoise.roy/bpr-ed2-posted3.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 **??** *(nicht klausurrelevant).*
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

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

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 $\mathbb{F}_3 \overset{??}{=} \mathbb{Z}/(3) = \{\overline{0}, \overline{1}, \overline{2}\}$ 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 **??**(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. [→ **??**]

5.1 Matrizen in Stufenform

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

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

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 ist 0
„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 $(1+\mathring{i})(1-3\mathring{i}) - 2 \cdot 2 + 2\mathring{i} = 1 - 3\mathring{i} + \mathring{i} + 3 - 4 + 2\mathring{i} = 0$ 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 **??** $(*)$ (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 $x + y \overset{??}{=} \begin{pmatrix} x_1+y_1 \\ \vdots \\ x_n+y_n \end{pmatrix} \in U$.

(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 **??** (g)

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

$$a_{i1}(x_1 + y_1) + \ldots + a_{in}(x_n + y_n) \overset{\substack{??\ (D) \\ ??}}{=} (a_{i1}x_1 + \ldots + a_{in}x_n) + (a_{i1}y_1 + \ldots + a_{in}y_n) = 0 + 0 = 0.$$

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

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

Bemerkung 5.1.4. In der Situation von **??** 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) $\overset{??}{\leftrightsquigarrow} r = 0$, (c) $\leftrightsquigarrow r = 1$, (b) $\leftrightsquigarrow r = 2$ & $\lambda_1 = \lambda_2 = 1$, (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 deren 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

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

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 §**??** beschriebene *Gauß-Verfahren* (lange vor Gauß bekannt, z.B. um -100 in China) wird dies leisten.

Bis hierher sollten wir am 29. November kommen.

Erinnerung und Definition 5.1.8. [→ **??** (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

$$(A(i,j))_{1 \le i \le m, 1 \le j \le n} := (A(i,j))_{(i,j) \in \{1,\ldots,m\} \times \{1,\ldots,n\}} \overset{??(c)}{=} \begin{pmatrix} A(1,1) & \ldots & A(1,n) \\ \vdots & & \vdots \\ A(m,1) & \ldots & A(m,n) \end{pmatrix}$$

schreibt. Die Menge aller $m \times n$-Matrizen über Z bezeichnet man mit $Z^{m \times n}$.

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

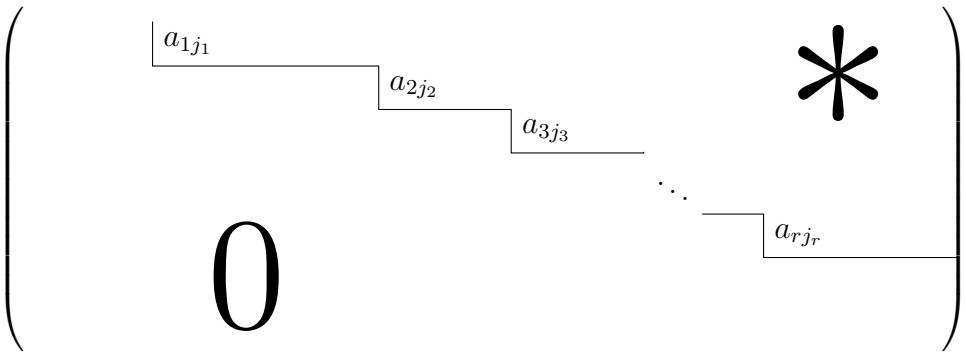
$$Ax := A \cdot x := \begin{pmatrix} a_{11}x_1 & + & \dots & + & a_{1n}x_n \\ \vdots & & & & \vdots \\ a_{m1}x_1 & + & \dots & + & a_{mn}x_n \end{pmatrix} \in K^n.$$

Damit können wir **??** $(*)$ 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



ist mit $r \in \{0, \dots, m\}, j_1, \dots, j_r \in \{1, \dots, n\}, j_1 < j_2 < \dots < j_r$ und $a_{ij_i} \neq 0$ für alle $i \in \{1, \dots, r\}$, wobei die „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 (das heißt 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) $\begin{pmatrix} 1 \end{pmatrix}$ ist eine 1×1-Matrix in reduzierter Stufenform mit 1 Stufe.

(d) $\begin{pmatrix} 1 & 0 \end{pmatrix}$; ist 1×2-Matrix in reduzierter Stufenform mit 1 Stufe und Stufenposition 1.

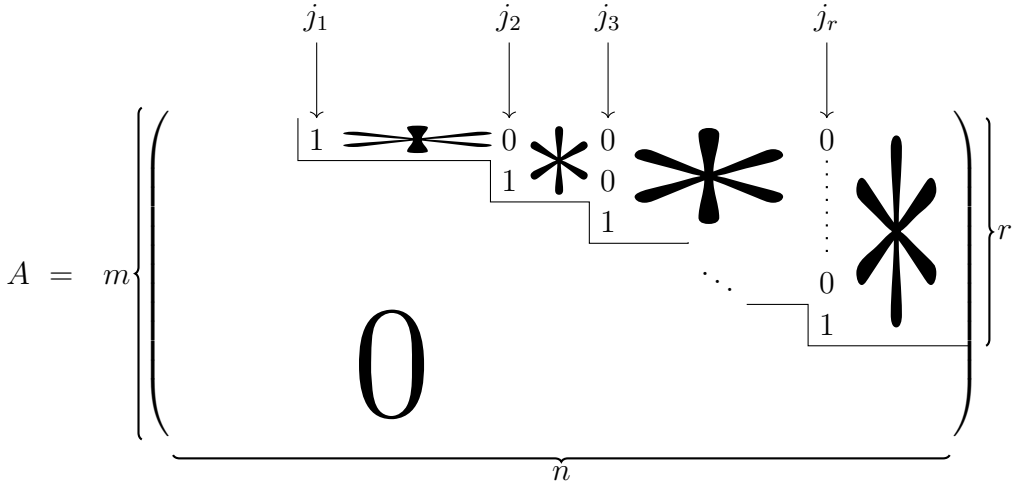
(e) $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ ist eine 2×1-Matrix *nicht in Stufenform*.

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

Sprechweise und Bemerkung 5.1.12. Ist ein lineares Gleichungssystem [→ **??**]

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

mit $A \in K^{m \times n}$ in Stufenform gegeben, so nennen wir für $j \in \{1, \ldots, n\}$ die Unbekannte x_j $\left\{ \begin{matrix} \textit{abhängig} \\ \textit{frei} \end{matrix} \right\}$ in $(*)$, wenn j $\left\{ \begin{matrix} \text{eine} \\ \text{keine} \end{matrix} \right\}$ Stufenposition [→ **??**] 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 [→ **??**], also



so kann man offensichtlich $(*)$ als ein System von r homogenen linearen Gleichungen

$$\begin{aligned} x_{j_1} &= \ldots \\ &\vdots \qquad\qquad (x \in K^n) \\ x_{j_r} &= \ldots \end{aligned}$$

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 $x^{(1)}, \ldots, x^{(n-r)} \in K^n$ bestimmen mit

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

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

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

kann man schreiben als

$$(*) \quad Ax = 0 \quad (x \in \mathbb{Q}^7)$$

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 **??** und in $(*)$ sind x_1, x_2, x_6 abhängig und x_3, x_4, x_5, x_7 frei gemäß **??**. Die Lösungsmenge von $(*)$ ist

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

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 $x_j \neq 0$ (etwa $x_j = 1$) und eine solche Lösung bekommen wir nicht, ohne die zu x_j gehörige Basislösung zu benutzen, denn die j-ten Komponenten der anderen Basislösungen sind null.

Bemerkung 5.1.14. (a) Da man stets wie in **??** 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 *elementaren Zeilenoperationen* [→ **??** (b) (1),(2)]:

- $\underbrace{Z_i}_{\text{„Zeile } i\text{“}} \underbrace{\leftarrow}_{\text{„wird“}} Z_i + \lambda Z_j \qquad (i,j \in \{1,\dots,m\}, i \neq j, \lambda \in K)$
  (Addieren des λ-fachen einer Zeile zu einer anderen)
- $Z_i \leftarrow \lambda Z_i \qquad (i \in \{1,\dots,m\}, \lambda \in K^\times)$
  (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, zum Beispiel die folgenden:

- $Z_i \underbrace{\leftrightarrow}_{\text{„wird vertauscht mit“}} Z_j \qquad (i,j \in \{1,\dots,m\})$
  (Vertauschen zweier Zeilen)
  $$\left[\begin{array}{l} \text{Simulation durch} \begin{cases} \text{Nichtstun falls } i=j \\ \left.\begin{array}{l} Z_i \leftarrow Z_i + Z_j \\ Z_j \leftarrow Z_j - Z_i \\ Z_i \leftarrow Z_i + Z_j \\ Z_j \leftarrow -Z_j \end{array}\right\} \text{falls } i \neq j \end{cases} \\ \binom{a}{b} \overset{Z_1 \leftarrow Z_1+Z_2}{\sim} \binom{a+b}{b} \overset{Z_2 \leftarrow Z_2-Z_1}{\sim} \binom{a+b}{-a} \overset{Z_1 \leftarrow Z_1+Z_2}{\sim} \binom{b}{-a} \overset{Z_2 \leftarrow -Z_2}{\sim} \binom{b}{a} \end{array}\right]$$
- $Z_i \leftarrow \sum_{j=1}^m \lambda_j Z_j \qquad (i \in \{1,\dots,m\}, \lambda_1,\dots,\lambda_m \in K \text{ mit } \lambda_i \neq 0)$
  $$\left[\begin{array}{lll} \text{Simulation durch} & Z_i \leftarrow \lambda_i Z_i & \\ & \left.\begin{array}{l} Z_i \leftarrow Z_i + \dots \\ \vdots \\ Z_i \leftarrow Z_i + \dots \end{array}\right\} & m-1 \text{ mal} \end{array}\right]$$

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. Geht B aus A durch Zeilenoperationen hervor, so auch A aus B (vgl. **??** (b)).
Durch

$$A \sim B :\iff B \text{ geht aus } A \text{ durch Zeilenoperation hervor} \qquad (A, B \in K^{m\times n})$$

wird also eine Äquivalenzrelation auf $K^{m\times n}$ definiert [→ **??** (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:

$$A = \begin{pmatrix} 0 & \overset{a_{ij}\neq 0}{\overset{\text{erste Spalte} \neq 0}{\overset{\downarrow}{*}}} \\ & a_{ij} \\ & * \end{pmatrix} \overset{Z_i \leftrightarrow Z_1}{\sim} \begin{pmatrix} 0 & \overset{a_{1j}\neq 0}{a_{1j}} & * \\ & * & * \end{pmatrix} \overset{Z_1 \leftarrow \frac{1}{a_{1j}} Z_1}{\sim} \begin{pmatrix} 0 & 1 \\ & a_{i2} \\ & \vdots \\ & a_{im} \end{pmatrix}$$

$$\overset{\substack{Z_2 \leftarrow Z_2 - a_{i2}Z_1 \\ Z_3 \leftarrow Z_3 - a_{i3}Z_1 \\ \vdots \\ Z_m \leftarrow Z_m - a_{im}Z_1}}{\sim} \begin{pmatrix} 0 & 1 & \overset{\text{WENDE VERFAHREN REKURSIV AN}}{\boxed{*}} \\ & 0 \\ & \vdots \\ & 0 \end{pmatrix} \sim \ldots \sim B$$

Bis hierher müssen wir am 2. Dezember kommen.

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:

$$
\begin{array}{l}
\xrightarrow[(i\in\{1,\dots,r\})]{Z_i \leftarrow \frac{1}{b_{ij_i}} Z_i} B \cdot \\
\xrightarrow[(i\in\{1,\dots,r-1\})]{Z_i \leftarrow Z_i - b_{ij_r} Z_r} \\
\xrightarrow[(i\in\{1,\dots,r-2\})]{Z_i \leftarrow Z_i - b_{ij_{r-1}} Z_{r-1}} \\
\qquad \vdots \\
\xrightarrow[(i\in\{1,\dots,1\})]{Z_i \leftarrow Z_i - b_{ij_2} Z_2}
\end{array}
\begin{pmatrix}
1 & b_{1j_2} & b_{1j_3} & \cdots & b_{1j_r} \\
 & 1 & b_{2j_3} & \cdots & b_{2j_r} \\
 & & 1 & \ddots & \vdots \\
 & & & & b_{(r-1)j_r} \\
 & & & & 1 \\
 & & 0 & & \\
 & & & &
\end{pmatrix}
\begin{pmatrix}
1 & b_{1j_2} & b_{1j_3} & \cdots & 0 \\
 & 1 & b_{2j_3} & \cdots & \vdots \\
 & & 1 & \ddots & 0 \\
 & & & & 1 \\
 & & 0 & & \\
 & & & &
\end{pmatrix}
\cdots
\begin{pmatrix}
1 & 0 & 0 & \cdots & 0 \\
 & 1 & 0 & \cdots & \vdots \\
 & & 1 & \ddots & 0 \\
 & & & & 1 \\
 & & 0 & & \\
 & & & &
\end{pmatrix}
= C
$$

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

(a) Bringe Koeffizientenmatrix auf Stufenform [→ **??**]

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

(c) Schreibe die Lösungsmenge als Spann [→ **??**]

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

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

kann man schreiben als

$$(*) \qquad Ax = 0 \qquad (x \in {\mathbb{F}_5}^4)$$

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

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

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 **??**

und **??** kann man seine Koeffizientenmatrix in reduzierter Stufenform annehmen und da diese Matrix breiter als hoch ist, hat das Gleichungssystem dann eine freie Unbekannte [→ **??**].

Bemerkung 5.2.8. Beim Lösen von homogenen linearen Gleichungssystemen kann es manchmal sinnvoll sein, die Unbekannten $x_1, \dots, x_n$ 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 $A = \begin{pmatrix} 1 & 0 & 0 & 3 & 1 \\ 2 & 0 & 1 & 3 & 1 \\ 3 & 1 & 0 & 3 & 1 \end{pmatrix}$.

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

$$\begin{bmatrix} A \overset{\text{VORSICHT}}{\underset{\text{wahrscheinlich}}{\not\sim}} \end{bmatrix} \begin{pmatrix} 1 & 3 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 2 \end{pmatrix} \begin{matrix} x_3 \\ \\ \end{matrix}$$

reduzierte Stufenform
3 Stufen
Stufenpositionen 1, 3, 4
abhängig: x_5, x_3, x_2
frei: x_4, x_1

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

Bis hierher sollten wir am 6. Dezember kommen.

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 $\ker A := \left\{ x \in K^n \;\middle|\; \underbrace{Ax}_{[\to \text{??}]} = 0 \right\}$ 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 $A = () = B$.

$\underline{n - 1 \to n \quad (n \in \mathbb{N})}$ 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: $A = \begin{pmatrix} 0 \\ \vdots & * \\ 0 \end{pmatrix}$.

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

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

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

Fall 2: $A = \begin{pmatrix} 1 \\ 0 \\ \vdots & * \\ 0 \end{pmatrix}$

Dann $\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \notin \ker A = \ker B$ und daher $B = \begin{pmatrix} 1 \\ 0 \\ \vdots & * \\ 0 \end{pmatrix}$.

Schreibe $A = \begin{pmatrix} 1 & \times\!\times\!\times \\ 0 & \\ \vdots & A' \\ 0 & \end{pmatrix}$ und $B = \begin{pmatrix} 1 & \times\!\times\!\times \\ 0 & \\ \vdots & B' \\ 0 & \end{pmatrix}$ mit $A', B' \in K^{(m-1)\times(n-1)}$. Dann sind A' und B' in reduzierter Stufenfom und es gilt

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

und daher $A' = B'$ nach IV. Es bleibt nur noch zu zeigen, dass die jeweils ersten Zeilen von A und B übereinstimmen.

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 $A \iff j-1$ Stufenposition 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

$$\ker A = \{x \in K^n \mid \forall a \in \operatorname{row} A : a_1 x_1 + \cdots + a_n x_n = 0\} \;\; [\to \textbf{??}, \textbf{??}].$$

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 aus $K^{m\times n}$ ändern weder ihre Äquivalenzklasse [→ **??**], noch ihren Kern [→ **??** (b)] noch ihren Zeilenraum (sieht man leicht). Also können wir nach **??** und **??** A und B in reduzierter Stufenform

annehmen. Dann

$$\ker A = \ker B \overset{??}{\Longrightarrow} A = B \Longrightarrow A \sim B \Longrightarrow \operatorname{row} A = \operatorname{row} B \overset{??}{\Longrightarrow} \ker A = \ker B.$$

Korollar 5.3.5. [→ **??**] 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. **??**.

„$\supseteq$“ Sei $a \in K^n$ mit $\forall x \in \ker A : a_1x_1 + \ldots + a_nx_n = 0$.

$$\text{Dann } \ker A = \ker \begin{pmatrix} \boldsymbol{A} \\ a_1 \quad \ldots \quad a_n \end{pmatrix} = \ker \begin{pmatrix} \boldsymbol{A} \\ 0 \quad \ldots \quad 0 \end{pmatrix} \text{ und daher}$$

$$\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \in \operatorname{row} \begin{pmatrix} \boldsymbol{A} \\ a_1 \quad \ldots \quad a_n \end{pmatrix} \overset{??}{=} \operatorname{row} \begin{pmatrix} \boldsymbol{A} \\ 0 \quad \ldots \quad 0 \end{pmatrix} = \operatorname{row} A.$$

Korollar 5.3.6. Ist $A \in K^{m\times n}$ und $X \in K^{\ell\times n}$, so

$$\ker A = \operatorname{row} X \iff \ker X = \operatorname{row} A.$$

Beweis. Aus Symmetriegründen reicht es „ $\Longrightarrow$ “ zu zeigen. Gelte also $\ker A = \operatorname{row} X$. Aus $\ker A \supseteq \operatorname{row} X$ folgt leicht $\ker X \supseteq \operatorname{row} A$. Es bleibt $\ker X \subseteq \operatorname{row} A$ zu zeigen. Sei hierzu $a \in \ker X$. Zu zeigen ist $a \in \operatorname{row} A$. Nach Korollar **??** reicht es hierzu zu zeigen, dass

$$a_1x_1 + \ldots + a_nx_n = 0$$

für alle $x \in \ker A$. Diese Gleichung ist für alle Zeilen x von X klar wegen $a \in \ker X$ und der Kommutativität der Multiplikation in K. Damit ist sie aber auch klar für alle x aus dem Zeilenraum $\operatorname{row} X$ von X. Dieser ist aber nach Voraussetzung gleich $\ker A$.

Bemerkung 5.3.7. Gegeben seien $x^{(1)}, \ldots, x^{(\ell)} \in K^n$. Wir wollen ein homogenes lineares Gleichungssystem ohne redundante Gleichungen finden, dessen Lösungsmenge genau $\operatorname{span}\left(x^{(1)}, \ldots, x^{(\ell)}\right)$ ist. Diese sogenannte *duale Aufgabe* kann man wegen **??** wie folgt lösen:

Schreibe $x^{(1)}, \ldots, x^{(\ell)}$ als Zeilen der Matrix $X \in K^{\ell\times n}$. Löse das Gleichungssystem

$$X \cdot a = 0 \qquad (a \in K^n)$$

und fasse die gefundenen Basislösungen $a^{(1)}, \ldots, a^{(m)} \in K^n$ als Zeilen einer Matrix $A \in K^{m \times n}$ auf. Es ist dann

$$A \cdot x = 0 \qquad (x \in K^n)$$

ein Gleichungssystem wie gewünscht.

Beispiel 5.3.8. 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.

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

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

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

Skript zur linearen Algebra I & II

§6 Vektorräume

Skript zur linearen Algebra I & II

§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 [→ **??**(c), **??**] $\underline{v} = (v_1, \ldots, v_n)$ 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 [→ **??**(e)]

$$(*) \qquad f = \mathrm{vec}_{\underline{w}} \circ f_A \circ \mathrm{coord}_{\underline{v}}$$

$$\begin{array}{ccccc} \sum_{j=1}^n \lambda_j v_j & & V & \xrightarrow{\quad f \quad} & W & & \sum_{i=1}^m \mu_i w_i \\ \downarrow & & \mathrm{coord}_{\underline{v}} \downarrow \cong & & \cong \uparrow \mathrm{vec}_{\underline{w}} & & \uparrow \\ \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} & & K^n & \xrightarrow[x \mapsto Ax]{} & K^m & & \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_m \end{pmatrix} \end{array}$$

Bemerkung 7.1.2. In der Situation von **??** gilt

$$\begin{aligned} (*) &\overset{\mathrm{vec}_{\underline{w}}^{-1} = \mathrm{coord}_{\underline{w}}}{\iff} \mathrm{coord}_{\underline{w}} \circ f = f_A \circ \mathrm{coord}_{\underline{v}} \\ &\overset{??}{\iff} \forall j \in \{1, \ldots, n\} : \mathrm{coord}_{\underline{w}}(f(v_j)) = \underbrace{f_A(\underbrace{\mathrm{coord}_{\underline{v}}(v_j)}_{=e_j})}_{=Ae_j} \\ &\iff \forall j \in \{1, \ldots, n\} : Ae_j = \mathrm{coord}_{\underline{w}}(f(v_j)) \end{aligned}$$

„In den Spalten stehen die **Koordinaten der** Bilder der Basisvektoren."

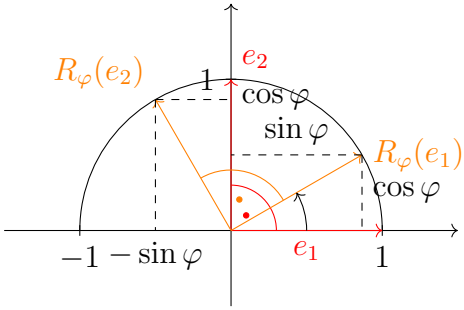
Zu jeder linearen Abbildung zwischen endlichdimensionalen Vektorräumen gibt es also bezüglich gegebener Basen jeweils *genau eine* Darstellungsmatrix.

Notation 7.1.3. $M(f, \underline{v}, \underline{w})$ steht für das eindeutig bestimmte A aus Definition **??**.

Beispiel 7.1.4. [→ **??**] $\underline{e} = (e_1, e_2) = (\binom{1}{0}, \binom{0}{1})$ Standardbasis des $\mathbb{R}^2$ [→ **??**].

(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 auch 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 [→ **??**] 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 $K[X]_d$ [→ **??**(d)]

$$D^{(d)}(X^k) = \begin{cases} 0 & \text{falls } k = 0 \\ kX^{k-1} & \text{falls } k \in \{1, \ldots, d\}. \end{cases}$$

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 $\mathbb{C}$ [→ **??**(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 $\mathbb{C}$.

$$C(1 + \mathring{i}) = 1 - \mathring{i}, \quad C(1 - \mathring{i}) = 1 + \mathring{i}$$

$$M(C, \underline{w}, \underline{w}) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

$$C(1) = 1 = \underline{\frac{1}{2}}(1 + \mathring{i}) + \underline{\frac{1}{2}}(1 - \mathring{i}), \quad C(\mathring{i}) = -\mathring{i} = \underline{\left(-\frac{1}{2}\right)}(1 + \mathring{i}) + \underline{\frac{1}{2}}(1 - \mathring{i})$$

$$M(C, \underline{v}, \underline{w}) = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

$$C(1 + \mathring{i}) = \underline{1} \cdot 1 + \underline{(-1)}\mathring{i}, \quad C(1 - \mathring{i}) = \underline{1} \cdot 1 + \underline{1} \cdot \mathring{i}$$

$$M(C, \underline{w}, \underline{v}) = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$$

Erinnerung 7.1.5 (Spezialfall von **??**)**.** Ist I eine Menge und V ein K-Vektorraum, so ist auch $V^I \overset{??}{=} \{f \mid f : I \to V\}$ 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 $I = \{1, \dots, m\} \times \{1, \dots, n\}$ [→ **??**] 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 **??** ist zu zeigen:

(a) $0 \in \mathrm{Hom}(V,W)$ (wobei $0 : V \to W, v \mapsto 0$)

(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 $v_1, v_2 \in K$
$0(\lambda v) = 0_W = \lambda 0_W = \lambda 0(v)$ für $v \in V$ und $\lambda \in K$

Zu (b). Seien $f, g : V \to W$ linear. Zu zeigen: $f + g$ linear.

$$\begin{aligned}(f+g)(v_1+v_2) &= f(v_1+v_2) + g(v_1+v_2) = f(v_1) + f(v_2) + g(v_1) + g(v_2)\\ &= f(v_1) + g(v_1) + f(v_2) + g(v_2) = (f+g)(v_1) + (f+g)(v_2)\end{aligned}$$

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 $v_1, v_2 \in V$

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

für alle $v \in V$ und $\lambda \in K$.

Übung 7.1.7. (a) Sind V und W K-Vektorräume, $U \overset{f}{\underset{g}{\rightrightarrows}} V \xrightarrow{h} W$ Abbildungen, 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 **??** und **??**(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 $(\Phi \circ \Psi)(A) = \Phi(\Psi(A)) = M(\mathrm{vec}_{\underline{w}} \circ f_A \circ \mathrm{coord}_{\underline{v}}, v, w) \overset{\substack{??\\??}}{=} A$.

Zu (b). Für $f \in \mathrm{Hom}(V,W)$ gilt $(\Psi \circ \Phi)(f) = \mathrm{vec}_{\underline{w}} \circ f_{M(f,\underline{v},\underline{w})} \circ \mathrm{coord}_{\underline{v}} \overset{\substack{??\\??}}{=} f$.

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 **??** 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 **??** 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 $\underline{v}$.

Bis hierher sollten wir am 23. Dezember kommen.

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

$3 \times \underline{2}$ $\qquad$ $\underline{2} \times 4$ $\qquad$ 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 [→ **??** (e)]

$$K^r \xrightarrow{f_B} K^n \xrightarrow{f_A} K^m, \quad f_{AB} : K^r \to K^m$$

so dass Definitions- und Zielmengen von f_{AB} und $f_A \circ f_B$ übereinstimmen. Da beide Abbildungen linear sind, reicht es nach **??**, die Gleichheit auf den Standardvektoren [→ **??**] $e_k \in K^r$ 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 **??**(**??**) 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 $f \circ g = \mathrm{vec}_{\underline{w}} \circ f_{AB} \circ \mathrm{coord}_{\underline{u}}$ [→**??**]. Nun gilt aber:

$$\begin{aligned} f \circ g &= (\mathrm{vec}_{\underline{w}} \circ f_A \circ \mathrm{coord}_{\underline{v}}) \circ (\mathrm{vec}_{\underline{v}} \circ f_B \circ \mathrm{coord}_{\underline{u}}) \\ &= \mathrm{vec}_{\underline{w}} \circ f_A \circ \underbrace{(\mathrm{coord}_{\underline{v}} \circ \mathrm{vec}_{\underline{v}})}_{=\mathrm{id}_{K^n}} \circ f_B \circ \mathrm{coord}_{\underline{u}} \\ &= \mathrm{vec}_{\underline{w}} \circ (f_A \circ f_B) \circ \mathrm{coord}_{\underline{u}} \overset{??}{=} \mathrm{vec}_{\underline{w}} \circ f_{AB} \circ \mathrm{coord}_{\underline{u}} \end{aligned}$$

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

$$\begin{aligned}(AB)C &= (M(f_A, \underline{e}^{(n)}, \underline{e}^{(m)})M(f_B, \underline{e}^{(r)}, \underline{e}^{(n)}))M(f_C, \underline{e}^{(s)}, \underline{e}^{(r)}) \\ &= M(f_A \circ f_B, \underline{e}^{(r)}, \underline{e}^{(m)})M(f_C, \underline{e}^{(s)}, \underline{e}^{(r)}) \\ &= M((f_A \circ f_B) \circ f_C, \underline{e}^{(s)}, \underline{e}^{(m)}) \overset{??(a)}{=} M(f_A \circ (f_B \circ f_C), \underline{e}^{(s)}, \underline{e}^{(m)}) \\ &= M(f_A, \underline{e}^{(n)}, \underline{e}^{(m)})M(f_B \circ f_C, \underline{e}^{(s)}, \underline{e}^{(n)}) \\ &= M(f_A, \underline{e}^{(n)}, \underline{e}^{(m)})(M(f_B, \underline{e}^{(r)}, \underline{e}^{(n)})M(f_C, \underline{e}^{(s)}, \underline{e}^{(r)})) = A(BC)\end{aligned}$$

Bemerkung 7.2.7. (a) Man kann für Korollar **??** 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 **??** und **??** folgt ähnlich wie im Beweis von **??**, 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 **??** „eintragweise“ definiert (und $A+B$ für $A, B \in K^{m\times n}$ schon durch **??** genauso wie in **??** „eintragweise“ definiert ist).

(c) Wegen **??** können wir beim Multiplizieren von mehreren Matrizen auf Klammern verzichten [→**??**].

Beispiel 7.2.8. Seien $\varphi, \psi \in \mathbb{R}$. Dann $R_{\varphi+\psi} = R_\varphi \circ R_\psi$ aus geometrischen Gründen und mit **??** daher $M(R_{\varphi+\psi}, \underline{e}) = M(R_\varphi, \underline{e}) M(R_\psi, \underline{e})$, was mit **??**(a) heißt

$$\begin{aligned}\begin{pmatrix} \cos(\varphi+\psi) & -\sin(\varphi+\psi) \\ \sin(\varphi+\psi) & \cos(\varphi+\psi) \end{pmatrix} &= \begin{pmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{pmatrix} \begin{pmatrix} \cos(\psi) & -\sin(\psi) \\ \sin(\psi) & \cos(\psi) \end{pmatrix} \\ &\overset{??}{=} \begin{pmatrix} (\cos\varphi)(\cos\psi) - (\sin\varphi)(\sin\psi) & -(\cos\varphi)(\sin\psi) - (\sin\varphi)(\cos\psi) \\ (\sin\varphi)(\cos\psi) + (\cos\varphi)(\sin\psi) & -(\sin\varphi)(\sin\psi) + (\cos\varphi)(\cos\psi) \end{pmatrix}.\end{aligned}$$

Es folgen die *Additionstheoreme*

$$\begin{aligned}\cos(\varphi+\psi) &= (\cos\varphi)(\cos\psi) - (\sin\varphi)(\sin\psi) \text{ und} \\ \sin(\varphi+\psi) &= (\sin\varphi)(\cos\psi) + (\cos\varphi)(\sin\psi).\end{aligned}$$

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 $B' = B'I_n = B'AB = I_n B = B$) 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. $M(f, \underline{v}) = I_n \overset{??}{\iff} f = \mathrm{vec}_{\underline{v}} \circ \underbrace{f_{I_n}}_{=\mathrm{id}_{K^n}} \circ \mathrm{coord}_{\underline{v}} \iff f = \underbrace{\mathrm{vec}_{\underline{v}} \circ \mathrm{coord}_{\underline{v}}}_{\mathrm{id}_V}$

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 **??**(c), also

$$I_n \overset{??}{=} M(f \circ f^{-1}, \underline{w}, \underline{w}) \overset{??}{=} M(f, \underline{v}, \underline{w}) M(f^{-1}, \underline{w}, \underline{v}) \text{ und}$$
$$I_n \overset{??}{=} M(f^{-1} \circ f, \underline{v}, \underline{v}) \overset{??}{=} M(f^{-1}, \underline{w}, \underline{v}) M(f, \underline{v}, \underline{w}),$$

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 $f \overset{??}{=} \mathrm{vec}_{\underline{w}} \circ f_A \circ \mathrm{coord}_{\underline{v}}$:

$$g \circ f = \mathrm{vec}_{\underline{v}} \circ f_B \circ f_A \circ \mathrm{coord}_{\underline{v}} = \mathrm{id}_V \text{ und } f \circ g = \mathrm{vec}_{\underline{w}} \circ f_A \circ f_B \circ \mathrm{coord}_{\underline{w}} = \mathrm{id}_W,$$

da $f_B \circ f_A \overset{??}{=} f_{BA} = f_{I_n} \overset{??}{=} \mathrm{id}_{K^n}$ und $f_A \circ f_B \overset{??}{=} f_{AB} = f_{I_n} \overset{??}{=} \mathrm{id}_{K^n}$. Aus **??** folgt, dass dann f bijektiv ist.

Proposition 7.2.12. Seien V und W endlichdimensionale K-Vektorräume derselben Dimension *[→**??**]* und $f \colon V \to W$ linear. Dann gilt

$$f \text{ injektiv} \iff f \text{ bijektiv} \iff f \text{ surjektiv}.$$

Beweis. Wähle mit **??** eine Basis $\underline{v} = (v_1, \ldots, v_n)$ von V. Nach **??** 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 **??** 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 $f_A \circ f_B \overset{??}{=} f_{AB} = f_{I_n} \overset{??}{=} \mathrm{id}_{K^n}$, woraus folgt, dass f_B injektiv ist. Aus **??** folgt, dass f_B bijektiv ist, woraus man mit **??** 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 **??** und **??**(**??**).

Zur Berechnung von Matrixinversen, muss man also sogenannte *inhomogene* lineare Gleichungssysteme lösen, was Gegenstand des nächsten Abschnitts ist.

Bis hierher sollten wir am 10. Januar kommen.

7.3 Inhomogene lineare Gleichungssysteme [→ §??]

Sprechweise und Bemerkung 7.3.1. [→ **??**, **??**] Ein *lineares Gleichungssystem* über K ist (ggf. nach Umstellen) von der Form

$$(*) \qquad Ax = b \quad (x \in K^n),$$

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 §**??** behandelt. Ist A in Stufenform [→ **??**], so nennen wir wieder für $j \in \{1, \dots, n\}$ die Unbekannte x_j $\left\{\begin{matrix} \textit{abhängig} \\ \textit{frei} \end{matrix}\right\}$ in $(*)$, wenn j $\left\{\begin{matrix} \text{eine} \\ \text{keine} \end{matrix}\right\}$ Stufenposition [→ **??**] von A ist. Ist A sogar in *reduzierter Stufenform* mit r Stufen und Stufenpositionen $j_1, \dots, j_r$, so kann man offensichtlich $(*)$ als ein System von m linearen Gleichungen

$$\begin{aligned} x_{j_1} &= b_1 + \dots \\ &\vdots \\ x_{j_r} &= b_r + \dots \\ 0 &= b_{r+1} \\ &\vdots \\ 0 &= b_m \end{aligned} \qquad (x \in K^n)$$

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. [→ ??] $K = \mathbb{Q}$

$$(*) \qquad \begin{array}{rrrrrl} x_1 & -3x_4 & +x_5 & +x_7 & = 1 & \\ & x_2 & -x_3 & +x_5 & = 2 & \qquad (x_1, \ldots, x_7 \in \mathbb{Q}) \\ & & x_6 & -x_7 & = -1 & \end{array}$$

x_1, x_2, x_6 abhängig, x_3, x_4, x_5, x_7 frei.

$$\left\{ x \in \mathbb{Q}^7 \;\middle|\; \begin{array}{rrrrrl} x_1 = & 1 & +3x_4 & -x_5 & -x_7 & \\ x_2 = & 2 & +x_3 & -x_5 & & , x_3, x_4, x_5, x_7 \in \mathbb{Q} \\ x_6 = & -1 & +x_7 & & & \end{array} \right\}$$

$$= \left\{ \begin{pmatrix} 1+3x_4-x_5-x_7 \\ 2+x_3-x_5 \\ x_3 \\ x_4 \\ x_5 \\ -1+x_7 \\ x_7 \end{pmatrix} \;\middle|\; x_3, x_4, x_5, x_7 \in \mathbb{Q} \right\}$$

$$= \left\{ \begin{pmatrix} 1 \\ 2 \\ 0 \\ 0 \\ 0 \\ -1 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} 3 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + x_5 \begin{pmatrix} -1 \\ -1 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} + x_7 \begin{pmatrix} -1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} \;\middle|\; x_3, x_4, x_5, x_7 \in \mathbb{Q} \right\}$$

$$= \left\{ \underbrace{\begin{pmatrix} 1 \\ 2 \\ 0 \\ 0 \\ 0 \\ -1 \\ 0 \end{pmatrix}}_{x_{(0)}} + \xi \;\middle|\; \xi \in \operatorname{span} \left(\underbrace{\begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}}_{\xi^{(1)}}, \underbrace{\begin{pmatrix} 3 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}}_{\xi^{(2)}}, \underbrace{\begin{pmatrix} -1 \\ -1 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}}_{\xi^{(3)}}, \underbrace{\begin{pmatrix} -1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}}_{\xi^{(4)}} \right) \right\}$$

Keines der so berechneten $\xi^{(\ell)}$ ist überflüssig, denn ist x_j eine freie Unbekannte, so gibt es im Fall 2 von **??** eine Lösung von $(*)$ mit $x_j \neq 0$ (etwa $x_j = 1$) und eine solche Lösung bekommen wir nicht, ohne das zu x_j gehörige $\xi^{(\ell)}$ zu benutzen, denn die j-te Komponente von $x^{(0)}$ und von den anderen $\xi^{(k)}$ ist jeweils null.

Bemerkung 7.3.3. [→ ??]

(a) Da man stets wie in **??** vorgehen kann, ist geklärt, wie man lineare Gleichungssysteme

$$Ax = b \quad (x \in K^n)$$

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* $(A\ b) \in K^{m \times (n+1)}$ und führt sie durch erlaubte Zeilenoperationen [→ §**??**] in eine Matrix $(A'\ b') \in K^{m \times (n+1)}$ mit A' in Stufenform über. Es gilt dann

$$\{x \in K^n \mid Ax = b\} = \{x \in K^n \mid \left(\begin{smallmatrix} x \\ -1 \end{smallmatrix}\right) \in \ker(A\ b)\}$$

$$\overset{(A\ b) \sim (A'\ b')}{\underset{??}{=}} \qquad \{x \in K^n \mid \left(\begin{smallmatrix} x \\ -1 \end{smallmatrix}\right) \in \ker(A'\ b')\} = \{x \in K^n \mid A'x = b'\}\,.$$

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. [→ **??**] $K = \mathbb{F}_5, 2 := 1+1, 3 := 1+1+1$ 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 $(A\ b_1\ b_2) \sim (A'\ b_1'\ b_2')$.

$$\left(\begin{array}{cccc|cc} 0 & 4 & 1 & 3 & 1 & 1 \\ 2 & 3 & 2 & 1 & 1 & 0 \\ 1 & 2 & 3 & 4 & 0 & 0 \\ 2 & 4 & 1 & 3 & 0 & 0 \end{array}\right) \overset{\substack{Z_1 \leftrightarrow Z_3 \\ Z_2 \leftarrow Z_2 - 2Z_1 \\ Z_4 \leftarrow Z_4 - 2Z_1}}{\sim} \left(\begin{array}{cccc|cc} 1 & 2 & 3 & 4 & 0 & 0 \\ 0 & 4 & 1 & 3 & 1 & 0 \\ 0 & 4 & 1 & 3 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

$$\overset{Z_3 \leftarrow Z_3 - Z_2}{\sim} \left(\begin{array}{cccc|cc} 1 & 2 & 3 & 4 & 0 & 0 \\ 0 & 4 & 1 & 3 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right) \overset{Z_2 \leftarrow \frac{1}{4} Z_2}{\sim} \left(\begin{array}{cccc|cc} 1 & 2 & 3 & 4 & 0 & 0 \\ 0 & 1 & 4 & 2 & 4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

$$\overset{Z_1 \leftarrow Z_1 - 2Z_2}{\sim} \underbrace{\left(\begin{array}{cccc|cc} 1 & 0 & 0 & 0 & 2 & 0 \\ 0 & 1 & 4 & 2 & 4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right)}_{\text{Stufenform, 2 Stufen, Stufenpositionen } 1,2 \text{ abhängig: } x_1, x_2, \text{ frei: } x_3, x_4}$$

$$\left\{x \in \mathbb{F}_5^4 \mid Ax = b_2\right\} = \emptyset$$

$$\begin{aligned} \left\{x \in \mathbb{F}_5^4 \mid Ax = b_1\right\} &= \left\{ \begin{pmatrix} 2 \\ 4-4x_3-2x_4 \\ x_3 \\ x_4 \end{pmatrix} \;\middle|\; x_3, x_4 \in \mathbb{F}_5 \right\} \\ &= \left\{ \begin{pmatrix} 2 \\ 4 \\ 0 \\ 0 \end{pmatrix} + \xi \;\middle|\; \xi \in \operatorname{span}\left(\begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 3 \\ 0 \\ 1 \end{pmatrix} \right) \right\} \end{aligned}$$

Mit **??** 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 [→ **??**] 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 $\ \& A^{-1} = B)$

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 [→ **??**]. Daher gibt es *kein* oder *mehrere* $x^{(1)} \in K^n$ mit $Ax^{(1)} = e_1$, je nachdem ob in **??** Fall 1 oder 2 eintritt. Wäre A invertierbar, so stünde dies im Widerspruch zu **??**.

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 **??** 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 [→ §??, §??, §??]

Definition 8.1.1. [→ **??**, **??**] 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. [→ **??**, **??**] Definition **??** wurde gerade so gemacht, 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. [→ **??**, **??**] 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 **??**.

Satz 8.1.4. [→ **??**, **??**] Sei V ein Vektorraum. Die Zuordnungen

$$\begin{aligned} \equiv &\mapsto \overline{0} \\ \equiv_U &\leftarrow\!\shortmid U \end{aligned}$$

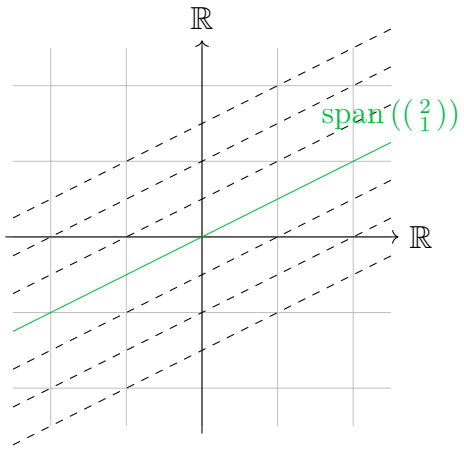
vermitteln eine Bijektion zwischen der Menge der Kongruenzrelationen auf V und der Menge der Unterräume auf V.

Beweis. Übung (vgl. **??**).

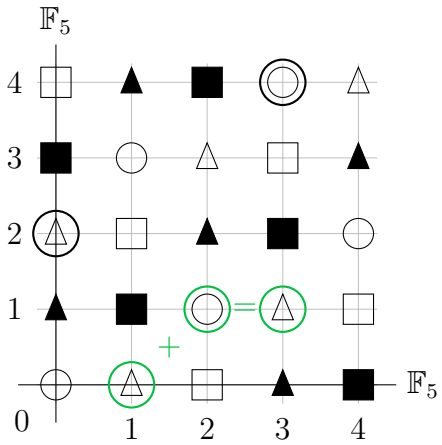
Notation 8.1.5. [→ ??, ??] Sei V ein Vektorraum und U ein Unterraum von V.

$$V/U := V/\equiv_U \qquad \text{„}V \text{ modulo } U\text{“}$$

Beispiel 8.1.6. (a) $\mathbb{R}^2/\operatorname{span}\left(\left(\begin{smallmatrix}2\\1\end{smallmatrix}\right)\right)$ besteht aus allen Geraden in der Ebene mit Steigung $\frac{1}{2}$.



(b) $\mathbb{F}_5^2/\operatorname{span}\left(\left(\begin{smallmatrix}2\\1\end{smallmatrix}\right)\right) = \left\{\underbrace{\overline{0}}_{\bigcirc}, \underbrace{\overline{\left(\begin{smallmatrix}1\\0\end{smallmatrix}\right)}}_{\triangle}, \underbrace{\overline{\left(\begin{smallmatrix}2\\0\end{smallmatrix}\right)}}_{\square}, \underbrace{\overline{\left(\begin{smallmatrix}3\\0\end{smallmatrix}\right)}}_{\blacktriangle}, \underbrace{\overline{\left(\begin{smallmatrix}4\\0\end{smallmatrix}\right)}}_{\blacksquare}\right\}$



$$\overline{\left(\begin{smallmatrix}0\\2\end{smallmatrix}\right)} + \overline{\left(\begin{smallmatrix}3\\4\end{smallmatrix}\right)} = \overline{\left(\begin{smallmatrix}3\\1\end{smallmatrix}\right)}$$

Proposition 8.1.7. [→ ??, ??, ??, ??] Seien V und W K-Vektorräume und $f\colon V \to W$ 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.

Bis hierher sollten wir am 13. Januar kommen.

Satz 8.1.9 (Homomorphiesatz für Vektorräume)**.** [→ **??**, **??**] Seien V und W K-Vektorräume, U ein Unterraum von V und $f\colon V \to W$ 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 **??**. Nur noch zu zeigen [→ **??**]:

$$\forall v \in V : \forall \lambda \in K : \overline{f}(\lambda\overline{v}) = \lambda\overline{f}(\overline{v}).$$

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)**.** [→ **??**, **??**] Seien V und W K-Vektorräume und $f\colon V \to W$ 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 $(u_1, \dots, u_m)$. Seien $v_1, \dots, v_n \in V$. Dann

$$(\overline{v_1}, \dots, \overline{v_n}) \text{ Basis von } V/U \iff (u_1, \dots, u_m, v_1, \dots, v_n) \text{ Basis von } V.$$

Insbesondere $\dim(U) + \dim(V/U) = \dim(V)$ falls $\dim V < \infty$.

Beweis. $K :=$ Grundkörper von V [→ **??**(c)].

„$\Longrightarrow$“ Sei $(\overline{v_1}, \dots, \overline{v_n})$ Basis von V/U. Zu zeigen:

(a) $u_1, \dots, u_m, v_1, \dots, v_n$ linear unabhängig in V

(b) $V = \operatorname{span}(u_1, \dots, u_m, v_1, \dots, v_n)$

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 $\sum_{j=1}^m \mu_j \overline{v_j} = \overline{\sum_{i=1}^m \lambda_i u_i + \sum_{j=1}^n \mu_j v_j} = 0$ folgt $\mu_j = 0$ für alle j, da $\overline{v_1}, \dots, \overline{v_n}$ linear unabhängig.

Da $u_1, \dots, u_m$ 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 $\overline{v_1}, \dots, \overline{v_n}$ 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. $\dim \ker f + \dim V/\ker f \overset{??}{=} \dim V$ und $V/\ker f \overset{??}{\cong} \operatorname{im} f$

Korollar 8.1.13. Zeilen- und Spaltenraum [→ **??**, **??**(e)] einer Matrix über einem Körper haben dieselbe Dimension.

Beweis. Sei $A \in K^{m \times n}$. Die Dimensionsformel **??** für $f_A \colon K^n \to K^m$ 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 $A \sim B$ [→ **??**]. Wegen $\ker B = \ker A$ und $\operatorname{row} B = \operatorname{row} A$ [→ **??**] reicht es zu zeigen, dass $\dim \ker B + \dim \operatorname{row} B = n$. Benutze nun **??**.

Definition 8.1.14. Sei $A \in K^{m \times n}$. Man nennt

$$\operatorname{rank} A := \dim \operatorname{im} A \overset{??}{=} \dim \operatorname{row} A$$

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 $f \colon V \to W$ linear. Dann gilt

$$\operatorname{rank} M(f, \underline{v}, \underline{w}) = \dim \operatorname{im} f.$$

Beweis. $A := M(f, \underline{v}, \underline{w}), f \overset{??}{=} \mathrm{vec}_{\underline{w}} \circ f_A \circ \mathrm{coord}_{\underline{v}}$

$$\dim \mathrm{im}\, f = \dim \mathrm{vec}_{\underline{w}}(\mathrm{im}(f_A \circ \mathrm{coord}_{\underline{v}})) \overset{\mathrm{vec}_{\underline{w}} \text{ Iso.}}{\underset{??}{=}} \dim \mathrm{im}(f_A \circ \mathrm{coord}_{\underline{v}})$$
$$\overset{\mathrm{coord}_{\underline{v}} \text{ Iso.}}{\underset{??}{=}} \dim \mathrm{im}\, f_A \overset{??(e)}{=} \dim \mathrm{im}\, A \overset{??}{=} \mathrm{rank}\, A$$

Proposition 8.1.16. Sei $A \in K^{n \times n}$. Dann

$$A \text{ invertierbar } [\to ??] \iff \mathrm{rank}\, A = n.$$

Beweis.

$$A \text{ invertierbar} \overset{??}{\iff} f_A \text{ bijektiv} \overset{??}{\iff} f_A \text{ surjektiv} \iff \mathrm{im}\, f_A = K^n$$
$$\overset{??}{\iff} \dim \mathrm{im}\, f_A = n \overset{??}{\iff} \mathrm{rank}\, A = n$$

Proposition 8.1.17. Seien $A \in K^{m \times n}$ und $B \in K^{n \times r}$. Dann gilt

$$\mathrm{rank}(AB) \leq \mathrm{rank}\, A \qquad \text{und} \qquad \mathrm{rank}(AB) \leq \mathrm{rank}(B).$$

Beweis. $\mathrm{im}(AB) = \{ABx \mid x \in K^r\} \subseteq \{Ay \mid y \in K^n\} = \mathrm{im}\, A$ und daher $\mathrm{rank}(AB) = \dim \mathrm{im}(AB) \overset{??}{\leq} \dim \mathrm{im}\, A = \mathrm{rank}\, A$
$\mathrm{row}(AB) = \{(x_1 \dots x_m)AB \mid x \in K^m\} \subseteq \{(y_1 \dots y_n)B \mid y \in K^n\} = \mathrm{row}\, B$ und daher $\mathrm{rank}(AB) = \dim \mathrm{row}(AB) \overset{??}{\leq} \dim \mathrm{row}\, B = \mathrm{rank}\, B$

8.2 Direkte Summen

Definition 8.2.1. Seien $n \in \mathbb{N}_0$ und $U_1, \dots, U_n$ Unterräume des Vektorraums V. Betrachte die lineare Abbildung

$$f \colon U_1 \times \dots \times U_n \to V, (u_1, \dots, u_n) \mapsto u_1 + \dots + u_n$$

$[\to ??]$. Der Unterraum

$$\sum_{i=1}^{n} U_i := U_1 + \dots + U_n := \mathrm{im}\, f = \{u_1 + \dots + u_n \mid u_1 \in U_1, \dots, u_n \in U_n\}$$

heißt *Summe* von $U_1, \dots, U_n$. Falls f injektiv ist, so sagt man, diese Summe ist *direkt* und schreibt dann auch

$$\bigoplus_{i=1}^{n} U_i := U_1 \oplus \dots \oplus U_n := \mathrm{im}\, f.$$

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

$$V = U_1 \oplus \ldots \oplus U_n$$

[→ **??**]. 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.

$$V = U_1 + \ldots + U_n \overset{??}{\iff} \dim V = \dim\left(\sum_{i=1}^{n} U_i\right)$$

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

$$\sum_{i=1}^{n} \dim U_i \overset{??}{=} \dim(U_1 \times \ldots \times U_n) = \dim(U_1 + \ldots + U_n).$$

„$\Longleftarrow$“ Ist $\sum_{i=1}^{n} \dim U_i = \dim\left(\sum_{i=1}^{n} U_i\right)$, so haben nach **??** Definitions- und Zielvektorraum von f dieselbe Dimension und nach **??** 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. $f\colon \begin{cases} U \times W & \to U + W \\ (u, w) & \mapsto u + w \end{cases}$ ist Vektorraumepimorphismus [$\to$ **??**]

$$\begin{aligned} \ker f &= \{(u, w) \in U \times W \mid u + w = 0\} \\ &= \{(u, w) \in U \times W \mid w = -u\} \\ &= \{(u, -u) \mid u \in U, -u \in W\} \\ &= \{(u, -u) \mid u \in U, u \in W\} \\ &= \{(u, -u) \mid u \in U \cap W\} \end{aligned}$$

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 [$\to$ **??**] liefert $\dim \ker f + \dim \operatorname{im} f = \dim(U \times W)$. Daraus folgt mit **??** die Behauptung.

Skript zur linearen Algebra I & II

§9 Determinanten

In diesem Kapitel sei stets K ein kommutativer Ring.

9.1 Definition und Eigenschaften von Determinanten

Definition 9.1.1. Sei $\sigma \in S_n$ [→??]. 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

$$(\tau(1),\tau(2)),(\tau(1),\tau(3)),(\tau(1),\tau(4)),\dots,(\tau(1),\tau(n)),\quad\dots\quad,(\tau(n-1),\tau(n))$$

für jedes $(i,j)\in\{1,\dots,n\}$ 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 **??** 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.

Bis hierher sollten wir am 17. Januar kommen.

Lemma 9.1.6. Jede Permutation $\sigma\in S_n$ ist Hintereinanderschaltung *[→??]* 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 $i\in\{1,\dots,n\}$ und Zeilen $a_1,\dots,a_{i-1},b,c,a_{i+1},\dots,a_n\in K^n$ 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 $i\in\{1,\dots,n\}$, $a_1,\dots,a_n\in K^n$ 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 $e_j := (0,\dots,0,\overbrace{1}^{j\text{-te Stelle}},0\dots,0) \in K^n$ für $j \in \{1,\dots,n\}$. 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 **??**, **??** und **??** 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

$$\begin{aligned} \delta(A) &= \delta \begin{pmatrix} \sum_{j=1}^n a_{1j}e_j \\ \vdots \\ \sum_{j=1}^n a_{nj}e_j \end{pmatrix} \overset{(a)}{\underset{(b)}{=}} \sum_{j_1=1}^n \cdots \sum_{j_n=1}^n a_{1j_1}\cdots a_{nj_n} \delta \begin{pmatrix} e_{j_1} \\ \vdots \\ e_{j_n} \end{pmatrix} \\ &\overset{(c)}{=} \sum_{\sigma\in S_n} a_{1\sigma(1)}\cdots a_{n\sigma(n)} \delta \begin{pmatrix} e_{\sigma(1)} \\ \vdots \\ e_{\sigma(n)} \end{pmatrix} = e \sum_{\sigma\in S_n} (\operatorname{sgn}\sigma) a_{1\sigma(1)}\cdots a_{n\sigma(n)}. \end{aligned}$$

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 ($k \neq \ell$) von A übereinstimmen. Definiere $A_n := \{\sigma \in S_n \mid \operatorname{sgn}\sigma = 1\}$ und beachte, dass [$\to$**??**, **??**]

$$\Phi\colon A_n \to S_n \setminus A_n,\ \sigma \mapsto \sigma \circ \tau_{k\ell}$$

eine Bijektion ist. Dann

$$\begin{aligned}\delta(A) &\overset{(*)}{=} e \sum_{\sigma\in A_n} \left(a_{1\sigma(1)}\cdots a_{n\sigma(n)} - a_{1(\sigma\circ\tau_{k\ell})(1)}\cdots a_{n(\sigma\circ\tau_{k\ell})(n)}\right)\\ &= e\sum_{\sigma\in A_n}\left(\prod_{i\in\{1,\ldots,n\}\setminus\{k,\ell\}} a_{i\sigma(i)}\right)\underbrace{\left(a_{k\sigma(k)}a_{\ell\sigma(\ell)} - \underbrace{a_{k\sigma(\ell)}}_{=a_{\ell\sigma(\ell)}}\underbrace{a_{\ell\sigma(k)}}_{=a_{k\sigma(k)}}\right)}_{=0} = 0.\end{aligned}$$

Bemerkung 9.1.8. (a) In **??** 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) **??**(**??**) 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

$$\det(A) := \delta_1^{(n)}(A) \overset{??(??)}{=} \sum_{\sigma\in S_n} (\operatorname{sgn}\sigma) a_{1\sigma(1)}\cdots a_{n\sigma(n)}$$

die *Determinante* von A.

Beispiel 9.1.10. (a) $\det() = \delta_1^{(0)}() \overset{()=I_0}{=} \delta_1^{(0)}(I_0) \overset{??(d)}{=} 1$

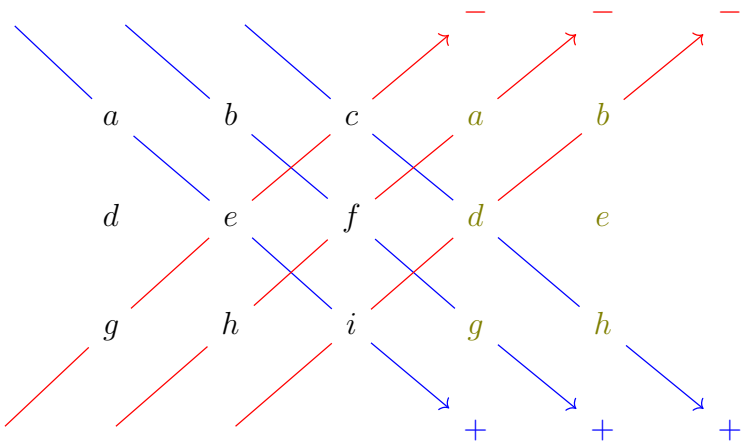
(b) Die Determinante einer Matrix, die eine Nullzeile enthält, ist null. Dies folgt zum Beispiel aus **??**(b). Insbesondere gilt für die Nullmatrix $0 \in K^{n\times n}$ im Fall $n \ge 1$, dass $\det(0) = 0$ (nicht allerdings, wenn $n = 0$ und $0 \neq 1$ in K, wie in (a) gesehen).

(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

$$A = \begin{pmatrix} \boxed{A_1} & & * \\ & \boxed{A_2} & \\ 0 & & \ddots \\ & & \boxed{A_m} \end{pmatrix}$$ 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 **??**(**??**) 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 [→**??**] 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

$$\det A = \left(\sum_{\varrho\in S_r} (\operatorname{sgn}\varrho) b_{1\varrho(1)}\cdots b_{r\varrho(r)}\right)\left(\sum_{\tau\in S_t} (\operatorname{sgn}\tau) c_{1\tau(1)}\cdots c_{t\tau(t)}\right) \overset{??(??)}{=} (\det B)(\det C).$$

Bemerkung 9.1.12. Sei K ein Körper. Zur Berechnung von Determinanten ist es oft am effizientesten, die Matrix durch Zeilenoperationen [→§??] auf obere Dreiecksgestalt (zum Beispiel Stufenform [→??]) zu bringen und den Effekt auf die Determinante dabei mitzuprotokollieren. Für die beiden *elementaren* Zeilenoperationen [→??] 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 $\det A = \det B$ [→??(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 $\det A = \frac{1}{\lambda}\det B$ [→??(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 ?? aus (a) und (b) oben oder wie im Beweis von ??].

(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 ?? aus (a) und (b) oben].

Beispiel 9.1.13. Sei K ein Körper, schreibe $2 := 1+1 \in K$ und so weiter. Gelte $3 \in K^\times$.

$$A := \begin{pmatrix} 1 & 2 & -1 & 0 \\ 2 & 3 & 0 & 1 \\ -1 & 1 & 1 & 0 \\ 0 & 9 & 3 & -3 \end{pmatrix} \overset{\substack{Z_2\leftarrow Z_2-2Z_1 \\ Z_3\leftarrow Z_3+Z_1}}{\sim} \begin{pmatrix} 1 & 2 & -1 & 0 \\ 0 & -1 & 2 & 1 \\ 0 & 3 & 0 & 0 \\ 0 & 9 & 3 & -3 \end{pmatrix}$$

$$\overset{\substack{Z_4\leftarrow \frac{1}{3}Z_4 \\ Z_4\leftarrow Z_4-Z_3 \\ Z_3\leftarrow Z_3+3Z_2}}{\sim} \begin{pmatrix} 1 & 2 & -1 & 0 \\ 0 & -1 & 2 & 1 \\ 0 & 0 & 6 & 3 \\ 0 & 0 & 1 & -1 \end{pmatrix} \overset{Z_3\leftrightarrow Z_4}{\underset{Z_4\leftarrow Z_4-6Z_3}{\sim}} \begin{pmatrix} 1 & 2 & -1 & 0 \\ 0 & -1 & 2 & 1 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 9 \end{pmatrix} =: B$$

Also $\det A \overset{??(b)(c)}{=} \frac{1}{3}\cdot(-1)\det B = (-3)(-9) = 27$.

Bemerkung 9.1.14. Sei K ein Körper und $A \in K^{n\times n}$. Ist A in Stufenform [→??], so ist der Rang von A [→??] gleich der Anzahl der Stufen von A und die Determinante von A [→??] gleich dem Produkt der Diagonaleinträge von A [→??]. 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 [→**??**(c)] und die Determinante nur um einen Faktor $\neq 0$ [→**??**]. Wegen A invertierbar $\overset{??}{\iff}$ $\operatorname{rank} A = n$ 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

$$f\colon K^{n\times n} \to K,\ A \mapsto \det(AB) \qquad \text{sowie}$$
$$g\colon K^{n\times n} \to K,\ A \mapsto (\det A)(\det B).$$

Zu zeigen ist $f = g$. Wir zeigen $f = \delta^{(n)}_{\det B} = g$, indem wir die Eigenschaften **??**(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):

$$f\begin{pmatrix} a_1 \\ \vdots \\ a_{i-1} \\ b+c \\ a_{i+1} \\ \vdots \\ a_n \end{pmatrix} = \det\left(\begin{pmatrix} a_1 \\ \vdots \\ a_{i-1} \\ b+c \\ a_{i+1} \\ \vdots \\ a_n \end{pmatrix} B\right) = \det\begin{pmatrix} a_1B \\ \vdots \\ a_{i-1}B \\ (b+c)B \\ a_{i+1}B \\ \vdots \\ a_nB \end{pmatrix} = \det\begin{pmatrix} a_1B \\ \vdots \\ a_{i-1}B \\ bB+cB \\ a_{i+1}B \\ \vdots \\ a_nB \end{pmatrix}$$
$$\overset{??(a)}{=} \det\begin{pmatrix} a_1B \\ \vdots \\ a_{i-1}B \\ bB \\ a_{i+1}B \\ \vdots \\ a_nB \end{pmatrix} + \det\begin{pmatrix} a_1B \\ \vdots \\ a_{i-1}B \\ cB \\ a_{i+1}B \\ \vdots \\ a_nB \end{pmatrix} = f\begin{pmatrix} a_1 \\ \vdots \\ a_{i-1} \\ b \\ a_{i+1} \\ \vdots \\ a_n \end{pmatrix} + f\begin{pmatrix} a_1 \\ \vdots \\ a_{i-1} \\ c \\ a_{i+1} \\ \vdots \\ a_n \end{pmatrix}$$

für alle $a_i, b, c \in K^n$.

Bis hierher sollten wir am 20. Januar kommen.

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 **??**(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 $A \in K^{n\times n}$ [→**??**]

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 $\det A \overset{??}{=} (\det(P^{-1}))(\det B)(\det P)$ und

$$1 \overset{??}{=} \det I_n \overset{??}{=} \det(P^{-1}P) \overset{??}{=} (\det(P^{-1}))(\det P).$$

Proposition 9.1.19. Sei f ein Endomorphismus [→**??**] eines endlichdimensionalen Vektorraums mit geordneten Basen $\underline{v}$ und $\underline{w}$. Dann $M(f, \underline{v}) \approx M(f, \underline{w})$ [→**??**].

Beweis. Es gilt $M(f, \underline{v}) \overset{??}{=} M(\underline{w}, \underline{v})M(f, \underline{w})M(\underline{v}, \underline{w})$ und

$$I_n \underset{??}{\overset{??}{=}} M(\underline{v}, \underline{v}) \overset{??}{=} M(\underline{w}, \underline{v})M(\underline{v}, \underline{w}),$$

das heißt $M(\underline{w}, \underline{v}) = M(\underline{v}, \underline{w})^{-1}$ [→**??**]. Mit $P := M(\underline{v}, \underline{w})$ gilt also

$$M(f, \underline{v}) \overset{??}{=} P^{-1}M(f, \underline{w})P.$$

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 [→**??**, **??**]

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

$$A^T := (a_{ij})_{1\le j\le n, 1\le i\le m} \in K^{n\times m}$$

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

$$\begin{aligned}
\det(A^T) &\overset{??(??)}{=} \sum_{\sigma\in S_n} (\operatorname{sgn}\sigma) a_{\sigma(1)1}\cdots a_{\sigma(n)n} \\
&\overset{\Phi \text{ bijektiv}}{=} \sum_{\sigma\in S_n} (\operatorname{sgn}(\sigma^{-1})) a_{\sigma^{-1}(1)1}\cdots a_{\sigma^{-1}(n)n} \\
&\overset{\substack{\sigma \text{ bijektiv} \\ \text{für } \sigma\in S_n}}{=} \sum_{\sigma\in S_n} (\operatorname{sgn}(\sigma^{-1})) a_{\sigma^{-1}(\sigma(1))\sigma(1)}\cdots a_{\sigma^{-1}(\sigma(n))\sigma(n)} \\
&\overset{\substack{\operatorname{sgn}(\sigma^{-1})=\operatorname{sgn}\sigma \\ \text{da } (\operatorname{sgn}(\sigma^{-1}))(\operatorname{sgn}\sigma)=\operatorname{sgn}(\mathrm{id})=1}}{=} \sum_{\sigma\in S_n} (\operatorname{sgn}\sigma) a_{1\sigma(1)}\cdots a_{n\sigma(n)} \overset{??(??)}{=} \det A
\end{aligned}$$

Bemerkung 9.1.24. Proposition **??** hat offensichtliche Konsequenzen: Die Determinantenfunktion ist auch *linear in den Spalten* [→**??**(a)], zur Berechnung von Determinanten kann man auch (analog zu **??** definierte) *Spaltenoperationen* heranziehen [→**??**], die Determinante einer Matrix in *unterer* Dreiecksgestalt [→**??**] 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 $i, j \in \{1, \ldots, n\}$ die Matrix, die aus A durch Streichen der i-ten Zeile und j-ten Spalte entsteht.

(a) Für alle $i \in \{1, \ldots, n\}$ gilt

$$\det A = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det A_{ij}$$

„Entwicklung nach der i-ten Zeile“

(b) Für alle $j \in \{1, \ldots, n\}$ gilt

$$\det A = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det A_{ij}$$

„Entwicklung nach der j-ten Spalte“

Beweis. Wegen **??** reicht es, (a) zu zeigen. Hierzu rechnet man

$$\det A \overset{\text{??(a)(b)}}{=} \sum_{j=1}^{n} a_{ij} \det \left(\begin{array}{c} \text{„wie in } A\text{“} \\ \hline 0 \ldots 0 \; 1 \; 0 \ldots 0 \end{array} \right) \leftarrow i$$

$$\uparrow j$$

$$\overset{\text{??(a)}}{=} \sum_{j=1}^{n} a_{ij} \det \left(\begin{array}{ccc} \text{„wie in } A\text{“} & 0 & \text{„wie in } A\text{“} \\ & \vdots & \\ & 0 & \\ \hline 0 \ldots 0 & 1 & 0 \ldots 0 \\ \hline \text{„wie in } A\text{“} & 0 & \text{„wie in } A\text{“} \\ & \vdots & \\ & 0 & \end{array} \right) \leftarrow i$$

$$\uparrow j$$

$$\overset{\text{??(c)}}{=} \sum_{j=1}^{n} a_{ij} \underbrace{(-1)^{(n-i)+(n-j)}}_{(-1)^{i+j}} \underbrace{\det \left(\begin{array}{c|c} A_{ij} & \begin{matrix} 0 \\ \vdots \\ 0 \end{matrix} \\ \hline 0 \ldots 0 & 1 \end{array} \right)}_{\overset{\text{??}}{=} (\det A_{ij}) \underbrace{\det(1)}_{=1}}$$

Beispiel 9.2.2.

$$\det \begin{pmatrix} 2 & 0 & 4 & 1 \\ 1 & 9 & -1 & 0 \\ 0 & 1 & 1 & -3 \\ -1 & 1 & 2 & 0 \end{pmatrix} \qquad \begin{pmatrix} + & - & + & - \\ - & + & - & + \\ + & - & + & - \\ - & + & - & + \end{pmatrix}$$

$$\overset{\text{Entwicklung nach letzter Spalte}}{=} -1 \underbrace{\det \begin{pmatrix} 1 & 9 & -1 \\ 0 & 1 & 1 \\ -1 & 1 & 2 \end{pmatrix}}_{\overset{\text{Entwicklung nach erster Spalte}}{=} \det\left(\begin{smallmatrix} 1 & 1 \\ 1 & 2 \end{smallmatrix}\right) - \det\left(\begin{smallmatrix} 9 & -1 \\ 1 & 1 \end{smallmatrix}\right) = (2-1)-(9+1) = -9} - (-3) \underbrace{\det \begin{pmatrix} 2 & 0 & 4 \\ 1 & 9 & -1 \\ -1 & 1 & 2 \end{pmatrix}}_{\overset{\text{Entwicklung nach erster Zeile}}{=} 2\det\left(\begin{smallmatrix} 9 & -1 \\ 1 & 2 \end{smallmatrix}\right) + 4\det\left(\begin{smallmatrix} 1 & 9 \\ -1 & 1 \end{smallmatrix}\right) = 2 \cdot 19 + 4 \cdot 10 = 78} = 9 + 3 \cdot 78 = 243$$

Definition 9.2.3. Ist $A = (a_{ji})_{1 \le i,j \le n} \in K^{n \times n}$ 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

$$\operatorname{com} A := \left((-1)^{i+j} \det A_{ij}\right)_{1 \le i,j \le n}$$

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 $(\operatorname{com} A)^T = ((-1)^{i+j} \det A_{ji})_{1 \le i,j \le n}$ [→ **??**, **??**]. Schreibe $I_n = (\delta_{ij})_{1 \le i,j \le n}$. Seien $i, k \in \{1 \dots, n\}$. Nach **??** 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 $\left\{\begin{matrix} \text{links} \\ \text{rechts} \end{matrix}\right\}$ die Entwicklung der Determinante von A nach der k-ten $\left\{\begin{matrix} \text{Zeile} \\ \text{Spalte} \end{matrix}\right\}$. Für $i \neq k$ steht $\left\{\begin{matrix} \text{links} \\ \text{rechts} \end{matrix}\right\}$ die Entwicklung der Determinante einer Matrix, deren i-te und k-te $\left\{\begin{matrix} \text{Zeile} \\ \text{Spalte} \end{matrix}\right\}$ übereinstimmen nach der $\left\{\begin{matrix} k\text{-ten Zeile} \\ i\text{-ten Spalte} \end{matrix}\right\}$.

Korollar 9.2.5 (in **??** 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 [→ **??**], so $1 = \det I_n = \det(AA^{-1}) \overset{??}{=} (\det A)(\det(A^{-1}))$, also $\det A \in K^\times$. Ist $\det A \in K^\times$, so $A\left(\frac{1}{\det A}(\operatorname{com} A)^T\right) \overset{??}{=} I_n \overset{??}{=} \left(\left(\frac{1}{\det A}\right)(\operatorname{com} A)^T\right) A$.

Korollar 9.2.6 (in **??** 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 **??**. Also $\det B \in K^\times$ und nach **??** ist B invertierbar. Dann aber

$$BA = BA(BB^{-1}) = B(AB)B^{-1} = BB^{-1} = I_n$$

wie im Beweis von **??**.

Satz 9.2.7 (Cramersche Regel [Gabriel Cramer *1704, †1752])**.** Seien $A \in K^{n\times n}$ und $x, b \in K^n$ mit $Ax = b$ [→ §**??**]. 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 $X_i := \left(\begin{array}{ccc|c|ccc} 1 & & & x_1 & & & \\ & \ddots & & \vdots & & \text{„wie in } I_n\text{“} & \\ & & & x_i & & & \\ & \text{„wie in } I_n\text{“} & & \vdots & & \ddots & \\ & & & x_n & & & 1 \end{array}\right)$ gilt

$$\det X_i \overset{\substack{??(a)\\??}}{=} \det \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & x_i & & \\ & & & & 1 & \\ & & & & & \ddots \\ & & & & & & 1 \end{pmatrix} \overset{??}{=} x_i$$

sowie $AX_i \overset{\substack{Ax=b\\??(c)}}{=} A_i$ und daher

$$(\det A)x_i = (\det A)(\det X_i) = \det(AX_i) = \det(A_i)$$

für $i \in \{1, \dots, n\}$.

Bis hierher sollten wir am 24. Januar kommen.

§10 Eigenvektoren

In diesem Kapitel sei stets K ein Körper.

10.1 Charakteristisches Polynom und Eigenwerte

Notation und Wiederholung 10.1.1. [→ ??, ??] Ist V ein K-Vektorraum, so ist

$$\begin{aligned}\operatorname{End}(V) := \operatorname{Hom}(V,V) &= \{f \mid f \text{ Endomorphismus der Vektorraums } V\} \\ &= \{f \mid f\colon V \to V \text{ linear}\}\end{aligned}$$

ein Unterraum des K-Vektorraums V^V.

Definition 10.1.2. Sei V ein K-Vektorraum und $f \in \operatorname{End}(V)$. 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 ein *Eigenvektor* (EV) von f zum Eigenwert λ. Für jeden Eigenwert λ von f nennt man den Unterraum

$$\ker(f - \lambda \operatorname{id}_V) = \{v \in V \mid f(v) = \lambda v\} \subseteq V,$$

welcher aus dem Nullvektor und den Eigenvektoren zum Eigenwert λ besteht, den *Eigenraum* von f zum Eigenwert λ.

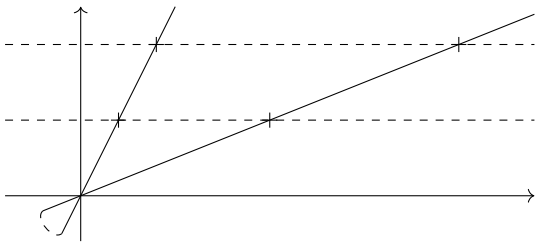
Beispiel 10.1.3. [→ ??, ??]

(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 $\operatorname{span}\left(\begin{smallmatrix}0\\1\end{smallmatrix}\right)$ der Eigenraum zum Eigenwert 1 und $\operatorname{span}\left(\begin{smallmatrix}1\\0\end{smallmatrix}\right)$ der Eigenraum zum Eigenwert -1.

(c) Die Projektion P hat die Eigenwerte 1 und 0. Es ist $\operatorname{span}\left(\begin{smallmatrix}1\\0\end{smallmatrix}\right)$ der Eigenraum zum Eigenwert 1 und $\operatorname{span}\left(\begin{smallmatrix}0\\1\end{smallmatrix}\right)$ 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 $\operatorname{span}\left(\begin{smallmatrix}1\\0\end{smallmatrix}\right)$.



(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 $D\colon K[X] \to K[X]$ 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 $n := \dim V < \infty$ und geordneten Basen $\underline{v}$ und $\underline{w}$. Sei $f \in \mathrm{End}(V)$. Dann sind die beiden Matrizen

$$M(f,\underline{v}) - XI_n, M(f,\underline{w}) - \underbrace{XI_n}_{} \in K[X]^{n\times n}$$
$$= \begin{pmatrix} X & & 0 \\ & \ddots & \\ 0 & & X \end{pmatrix}$$

ähnlich [→ **??**] und haben daher dieselbe Determinante [→ **??**].

Beweis. Nach **??** 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

$$\begin{aligned} P^{-1}(M(f,\underline{w}) - XI_n)P &= P^{-1}M(f,\underline{w})P - P^{-1}(XI_n)P \\ &\overset{??(b)}{=} M(f,\underline{v}) - XP^{-1}P = M(f,\underline{v}) - XI_n \end{aligned}$$

und daher $M(f,\underline{v}) - XI_n \approx M(f,\underline{w}) - XI_n$.

Definition 10.1.5. Sei V ein endlichdimensionaler Vektorraum und $f \in \mathrm{End}(V)$. Dann ist das *charakteristische Polynom* von f definiert als

$$\chi_f := \det(M(f, \underline{v}) - XI_n) \in K[X],$$

wobei $\underline{v}$ eine beliebig gewählte Basis von V ist [→ **??**].

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 $\chi_f = (-1)^n X^n + a_{n-1}X^{n-1} + \ldots + a_1 X_1 + a_0$, wobei $a_0 = \det f$. Insbesondere hat χ_f den Grad n. Dies folgt leicht aus der Leibniz-Formel **??**(**??**), weil $M(f, \underline{v}) - XI_n)$ von der Form

$$\begin{pmatrix} b_{11} - X & & \text{„aus } K\text{“} \\ & \ddots & \\ \text{„aus } K\text{“} & & b_{nn} - X \end{pmatrix}$$

ist.

Proposition 10.1.7. Sei V ein K-Vektorraum, $n := \dim V < \infty$, $f \in \mathrm{End}(V)$ und $\lambda \in K$. Dann

$$\lambda \text{ Eigenwert von } f \text{ [→ ??]} \iff \chi_f(\lambda) = 0 \text{ [→ ??]}.$$

„Die Eigenwerte sind die Nullstellen [→ **??**] des charakteristischen Polynoms.“

Beweis. Wähle eine geordnete Basis $\underline{v}$ von V.

$$\begin{aligned}
\lambda \text{ Eigenwert von } f &\overset{??}{\iff} \exists v \in V \setminus \{0\} : f(v) = \lambda v \\
&\iff \ker(f - \lambda \mathrm{id}_V) \neq \{0\} \\
&\overset{??(\mathrm{b})}{\iff} \dim \ker(f - \lambda \mathrm{id}_V) \neq 0 \\
&\overset{??}{\iff} \dim \mathrm{im}(f - \lambda \mathrm{id}_V) \neq n \\
&\overset{??}{\iff} \mathrm{rank}(M(f - \lambda \mathrm{id}_V, \underline{v})) \neq n \\
&\overset{??}{\iff} M(f - \lambda \mathrm{id}_V, \underline{v}) \text{ nicht invertierbar} \\
&\overset{??}{\iff} \det \underbrace{M(f - \lambda \mathrm{id}_V, \underline{v})}_{\overset{??}{=} M(f,\underline{v}) - \lambda \underbrace{M(\mathrm{id}_V, \underline{v})}_{\overset{??}{=} I_n}} = 0 \\
&\iff \det(M(f, \underline{v}) - \lambda I_n) = 0 \\
&\underset{??}{\overset{??(??)}{\iff}} \chi_f(\lambda) = 0
\end{aligned}$$

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. **??**, **??**, **??**

Beispiel 10.1.9. [→ **??**, **??**, **??**]

(a) Sei $\varphi \in \mathbb{R}$.

$$2\chi_{R_\varphi} = \det \begin{pmatrix} (\cos\varphi) - X & -\sin\varphi \\ \sin\varphi & (\cos\varphi) - X \end{pmatrix} \overset{\text{?? (d)}}{=} ((\cos\varphi) - X)^2 + (\sin\varphi)^2$$

Für $\lambda \in \mathbb{R}$ gilt also

$$\begin{aligned} \chi_{R_\varphi}(\lambda) = 0 &\iff (\lambda = \cos\varphi \,\&\, \sin\varphi = 0) \\ &\iff ((\exists n \in \mathbb{Z} \text{ ungerade} : \varphi = n\pi) \,\&\, \lambda = -1) \text{ oder} \\ &\qquad (\exists n \in \mathbb{Z} \text{ gerade} : \varphi = n\pi) \,\&\, \lambda = 1)). \end{aligned}$$

(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) $\chi_{T_a} = \det \begin{pmatrix} 1 - X & a \\ 0 & 1 - X \end{pmatrix} \overset{??}{=} (X - 1)^2$ 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 $\lambda_1, \ldots, \lambda_m$ paarweise verschiedene Eigenwerte von f und v_i ein Eigenvektor zum Eigenwert λ_i für jedes $i \in \{1, \ldots, m\}$ [→ **??**]. Dann sind $v_1, \ldots, v_m$ linear unabhängig in V.

„Eigenvektoren zu verschiedenen Eigenwerten sind linear unabhängig."

Beweis. Induktion nach $m \in \mathbb{N}_0$.

$\underline{m = 0}$ $\quad \emptyset$ ist linear unabhängig ✓

$\underline{m - 1 \to m \ \ (m \in \mathbb{N})}$ 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. **??** folgt auch aus **??** und **??**.

Korollar 10.1.12. Sei f ein Endomorphismus eines endlichdimensionalen Vektorraums V. Dann ist die Summe der Eigenräume [→ **??**] von f direkt [→ **??**], 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 **??** folgt sofort, dass diese lineare 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 [→ **??**] von p (es gilt $m \leq n$

[→ **??**]). Dann gibt es eindeutig bestimmte $\alpha_1, \ldots, \alpha_m \in \mathbb{N}$ und $r \in K[X]$ derart, dass

$$p = (X - \lambda_1)^{\alpha_1} \cdots (X - \lambda_m)^{\alpha_m} r$$

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 **??**, Eindeutigkeit leicht zu sehen.

Definition 10.1.14. Sei V ein endlichdimensionaler Vektorraum und $f \in \mathrm{End}(V)$. 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 ein Eigenwert von f. Wähle eine Basis $(v_1, \ldots, v_m)$ von $\ker(f - \lambda \mathrm{id}_V)$ [→ **??**] und ergänze sie zu einer Basis $\underline{v} = (v_1, \ldots, v_n)$ von V [→ **??**]. 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 $i \in \{1, \ldots, m\}$ [→ **??**]. Daher

$$\begin{aligned} \chi_f &\overset{??}{=} \det\left(\begin{array}{c|c} (\lambda - X) I_m & A \\ \hline 0 & B - X I_{n-m} \end{array}\right) \\ &\overset{??}{=} (\det((\lambda - X) I_m)) \det(B - X I_{n-m}) \overset{??}{=} (\lambda - X)^m r \end{aligned}$$

für $r := \det(B - X I_{n-m}) \in K[X]$.

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. [→??] 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 **??** $q, r \in K[X]$ 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 **??** 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 $p = q$.

Bis hierher sollten wir am 27. Januar kommen.

Proposition 10.2.5. Sei $p \in K[X]$ ein Polynom vom Grad $n \in \mathbb{N}_0$ [→??]. Dann ist das Ideal (p) [→??] des kommutativen Ringes $K[X]$ [→??] ein Unterraum des Vektorraums $K[X]$ [→??(b)] und $(\overline{1}, \overline{X}, \ldots, \overline{X^{n-1}})$ eine Basis des Quotientenvektorraums $K[X]/(p)$ [→??].

Beweis. Nach der Definition einer Basis **??**(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 **??**. 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 **??** und **??** 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 **??** 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 $p = X^n + a_{n-1}X^{n-1} + \ldots + a_1 X + a_0 \in K[X]$ 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

$$C_p = \begin{pmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & & \vdots & -a_1 \\ 0 & 1 & \ddots & \vdots & -a_2 \\ 0 & 0 & \ddots & 0 & \vdots \\ \vdots & \vdots & & \vdots & \\ 0 & 0 & \cdots & 1 & -a_{n-1} \end{pmatrix}.$$

Beweis. Da im Quotientenring $K[X]/(p)$ gilt $\overline{Xq} \overset{??}{=} \overline{X} \cdot \overline{q}$ 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 **??** ist $\underline{v}$ eine Basis des Quotientenvektorraums $K[X]/(p)$. Die behauptete Gleichheit für C_p ergibt sich aus **??** wegen $f(\overline{X^k}) = \overline{X^{k+1}}$ für alle $k \in \{0, \ldots, n-2\}$ und $f(\overline{X^{n-1}}) = \overline{X^n} = -a_0 1 - a_1 \overline{X} - \ldots - a_{n-1} \overline{X^{n-1}}$.

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 $p = 1 \overset{??(a)}{=} (-1)^0 \det() = (-1)^0 \chi_{C_p}$. Sei also Œ $n \geq 1$. Benutze wieder die Basis $\underline{v} := (\overline{1}, \overline{X}, \ldots, \overline{X^{n-1}})$ des Quotientenvektorraums $K[X]/(p)$ [→**??**] und schreibe $p = X^n + a_{n-1}X^{n-1} + \ldots + a_1 X + a_0 \in K[X]$ mit $a_0, \ldots, a_{n-1} \in K$. Es gilt:

$$\begin{aligned} \chi_{C_p} &\overset{??(e)}{=} \det(C_p - XI_n) \\ &= \det \begin{pmatrix} -X & 0 & \cdots & 0 & -a_0 \\ 1 & -X & & & -a_1 \\ 0 & 1 & \ddots & & -a_2 \\ 0 & 0 & \ddots & & \vdots \\ \vdots & \vdots & & -X & -a_{n-2} \\ 0 & 0 & \cdots\ 0 & 1 & -a_{n-1} - X \end{pmatrix}. \end{aligned}$$

Addiert man nun in der großen Matrix nacheinander von unten beginnend jeweils das X-fache einer Zeile zur vorherigen, dann ändert sich gemäß **??**(a) die Determinante nicht und man erhält

$$\chi_{C_p} = \det \begin{pmatrix} 0 & 0 & \cdots & 0 & -a_0 - a_1X - \ldots - a_{n-1}X^{n-1} - X^n \\ 1 & 0 & & & -a_1 - a_2X - \ldots - a_{n-1}X^{n-2} - X^{n-1} \\ 0 & 1 & \ddots & & -a_2 - a_3X - \ldots - a_{n-1}X^{n-3} - X^{n-2} \\ 0 & 0 & \ddots & & \vdots \\ \vdots & \vdots & & 0 & -a_{n-2} - a_{n-1}X - X^2 \\ 0 & 0 & \cdots\ 0 & 1 & -a_{n-1} - X \end{pmatrix}.$$

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 **??**(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 $R_\varphi \in \mathrm{End}(\mathbb{R}^2)$ [→**??**(a)]

$$(R_\varphi)^k = R_{k\varphi}.$$

Erinnerung 10.2.10. [→**??**,**??**] 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. [→**??**]

(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 **??** nachzurechnen. Mit Erinnerung **??** geht dies analog zum Beweis von **??**.

(b) geht genauso unter Benutzung von **??** und **??**(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 **??** 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

$$\psi\left(\sum_{k=0}^{n} a_k X^k\right) = \sum_{k=0}^{n} \varphi(a_k) f^k = \sum_{k=0}^{n} \underbrace{(a_k\,\mathrm{id}_V) \circ f^k}_{\overset{??}{=} a_k(\mathrm{id}_V \circ f^k) = a_k f^k} .$$

(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 **??**, so ist dies wegen **??** und **??** äquivalent zu $\chi_{C_p}(f) = 0$ und damit wegen **??** 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 **??** 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 $f^m(v) + a_{m-1}f^{m-1}(v) + \ldots + a_1 f(v) + a_0 v = 0$. Da $v, f(v), \ldots, f^{m-1}(v)$ linear unabhängig sind, findet man mit **??** eine Basis $\underline{v} = (v, f(v), \ldots, f^{m-1}(v), v_{m+1}, \ldots, v_n)$ von V. Setzt man nun

$$p := X^m + a_{m-1}X^{m-1} + \ldots + a_1 X + a_0,$$

so ist nach **??**

$$M(f, \underline{v}) = \left(\begin{array}{c|c} C_p & * \\ \hline 0 & A \end{array}\right)$$

für ein $A \in K^{(n-m)\times(n-m)}$. Es ergibt sich

$$\begin{aligned}\chi_f &\overset{??}{=} \det(M(f,\underline{v}) - XI_n) = \det\left(\begin{array}{c|c} C_p - XI_m & * \\ \hline 0 & A - XI_{n-m} \end{array}\right) \\ &\overset{??}{=} \det(C_p - XI_m)\det(A - XI_{n-m}) \overset{??}{=} \chi_{C_p}\chi_A \overset{??}{=} pq = qp\end{aligned}$$

mit $q := (-1)^m\chi_A \in K[X]$. Nun gilt

$$\chi_f(f) = (qp)(f) \overset{??(a)}{=} q(f) \circ p(f)$$

und daher

$$(\chi_f(f))(v) = (q(f))((p(f))(v)) = (q(f))(0) = 0,$$

da

$$\begin{aligned}(p(f))(v) &= (f^m + a_{m-1}f^{m-1} + \ldots + a_1 f + a_0 \,\mathrm{id}_V)(v) \\ &= f^m(v) + a_{m-1}f^{m-1}(v) + \ldots + a_1 f(v) + a_0 v = 0.\end{aligned}$$

(b) Für $A \in K^{n\times n}$ ist $f_A \in \mathrm{End}(K^n)$ [→**??**(e)] und wir haben

$$\begin{aligned}\chi_A(A) &\overset{??(e)}{=} \chi_{f_A}(A) = \chi_{f_A}(M(f_A,\underline{e})) = \\ &\overset{??}{=} M(\chi_{f_A}(f_A),\underline{e}) \overset{(a)}{=} M(0,\underline{e}) = 0.\end{aligned}$$

Bemerkung 10.2.15. (a) Im Beweis des Satzes von Cayley-Hamilton **??** 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 **??** und **??** 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 [→**??**(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 [→**??**(a)]. Wo liegt aber genau der Fehler? Da das Einsetzen von A für X ein Ringhomomorphismus von $K[X]$ nach $K[A]$ ist [→**??**(b)], kann man wegen der Leibniz-Formel **??**(**??**) 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 [→**??**(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 **??**. 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 $\chi_A = a_nX^n + \ldots + a_1X + a_0$ mit $a_0, \ldots, a_n \in K$ und $(\mathrm{com}(A - XI_n))^T = X^{n-1}B_{n-1} + \ldots + XB_1 + B_0$ mit $B_0, \ldots, B_{n-1} \in K^{n\times n}$. Dann gilt $(A - XI_n)(X^{n-1}B_{n-1} + \ldots + XB_1 + B_0) = (A - XI_n)(\mathrm{com}(A - XI_n))^T \overset{??}{=} (\det(A - XI_n))I_n = X^na_nI_n + \ldots + Xa_1I_n + a_0I_n$, 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

$$\begin{aligned} -A^nB_{n-1} &= a_nA^n \\ A^nB_{n-1} - A^{n-1}B_{n-2} &= a_{n-1}A^{n-1} \\ &\vdots \\ A^2B_1 - AB_0 &= a_1A \\ AB_0 &= a_0I_n. \end{aligned}$$

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`.

Bis hierher sollten wir am 31. Januar kommen.

Definition 10.2.16. [→??, ??]

(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 **??** 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. [→??]

(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. [→??]

(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. [→??, ??, ??, ??] 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 **??** gilt

$$\begin{aligned} \underbrace{(R_\varphi)^2}_{\overset{??}{=}R_{2\varphi}} - 2(\cos\varphi)R_\varphi + \mathrm{id}_{\mathbb{R}^2} &= 0, \quad \text{also} \\ R_{2\varphi}(v) - 2(\cos\varphi)R_\varphi(v) + v &= 0 \quad \text{für alle } v \in \mathbb{R}^2. \end{aligned}$$

(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 **??** und **??**.

(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 [→ **??**] mit ebensoviel Zeilen wie Spalten nennt man *quadratisch* [→ **??**(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 $\left\{\begin{array}{c} \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \\ \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \\ \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ * & & \lambda_n \end{pmatrix} \end{array}\right\}$ ist.

Definition 10.3.2. Sei V ein Vektorraum und $f \in \mathrm{End}(V)$.

(a) f heißt $\left\{\begin{matrix}\textit{diagonalisierbar}\\ \textit{trigonalisierbar}\end{matrix}\right\}$, wenn es eine geordnete Basis $\underline{v}$ von V gibt derart, dass $M(f,\underline{v})$ [→ **??**] $\left\{\begin{matrix}\text{Diagonal-}\\ \text{obere Dreiecks-}\end{matrix}\right\}$gestalt hat.

(b) Eine (geordnete) Basis von V heißt (geordnete) *Eigenbasis* für f, wenn sie aus Eigenvektoren [→**??**] von f besteht.

Satz 10.3.3. Sei f ein Endomorphismus eines endlichdimensionalen Vektorraums V. Dann sind folgende Aussagen äquivalent:

(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 [→ **??**, **??**].

Beweis. Sei V ein K-Vektorraum, $n := \dim V < \infty$ und $f \in \mathrm{End}(V)$.
(a) $\iff$ (b) folgt aus der folgenden Tatsache:
(∗) Sei $\underline{v} = (v_1, \ldots, v_n)$ Basis von V. Dann gilt für $\lambda_1, \ldots, \lambda_n \in K$:

$$M(f,\underline{v}) = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \overset{??}{\iff} \forall i \in \{1,\ldots,n\} : \mathrm{coord}_{\underline{v}}(v_i) = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \lambda_i \leftarrow i\text{-te Stelle} \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

$$\overset{??}{\iff} \forall i \in \{1,\ldots,n\} : f(v_i) = \underbrace{0v_1 + \ldots + 0v_{i-1} + \lambda_i v_i + 0v_{i+1} + \ldots + 0v_n}_{\lambda_i v_i}$$

$$\iff \forall i \in \{1,\ldots,n\} : v_i \text{ ist Eigenvektor von } f \text{ zum Eigenwert } \lambda_i$$

(a) $\implies$ (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

$$\chi_f \overset{??}{=} \det(M(f,\underline{v}) - XI_n) = \det \begin{pmatrix} \lambda_1 - X & & 0 \\ & \ddots & \\ 0 & & \lambda_n - X \end{pmatrix}$$

$$\overset{??}{=} \prod_{i=1}^{n} (\lambda_i - X) = (-1)^n \prod_{i=1}^{n} (X - \lambda_i).$$

Es zerfällt χ_f also [→ **??**]. Sei nun λ ein Eigenwert von f. Nach Umnummerieren

der λ_i können wir $\lambda = \lambda_1 = \ldots = \lambda_m \neq \lambda_i$ für $i \in \{m+1, \ldots, n\}$ annehmen. Zu zeigen: $\dim \ker(f - \lambda \operatorname{id}_V) = m$ [→ **??**].
Aus **??** wissen wird schon „≤“.
„≥“ 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) $\implies$ (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 $\alpha_1 + \cdots + \alpha_m = \deg \chi_f \overset{??}{=} n$. Wegen $\dim \ker(f - \lambda_i \operatorname{id}_V) \overset{(c)}{=} \alpha_i$ folgt [→ **??**]

$$\dim \left(\bigoplus_{i=1}^{n} \ker(f - \lambda_i \operatorname{id}_V) \right) \overset{??}{=} \sum_{i=1}^{m} \alpha_i = n,$$

also $\bigoplus_{i=1}^{m} \ker(f - \lambda_i \operatorname{id}_V) = V$ nach **??**, 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 **??**.

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 [→ **??**(e)] $\left\{ \begin{matrix} \text{diagonalisierbar} \\ \text{trigonalisierbar} \end{matrix} \right\}$ [→ **??**] 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) $\iff$ (c) $\iff$ (d) ist klar, da dasselbe für f_A gilt [→ **??**].
(a) $\implies$ (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

$$A \overset{??(e)}{=} M(f_A, \underline{e}) \overset{??}{\approx} M(f_A, \underline{v}) = D.$$

(b) $\implies$ (c) 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

$$f_A(v_i) = Av_i = P^{-1}DPP^{-1}e_i = P^{-1}De_i = P^{-1}\lambda_i e_i = \lambda_i P^{-1}e_i = \lambda_i v_i.$$

Proposition 10.3.6. Sei f ein Endomorphismus des Vektorraums V. Ein Unterraum U von V heißt *f-invariant*, wenn $f(U) \subseteq U$. Ist U f-invariant, so sind

$$f|_U \colon U \to U, v \mapsto f(v) \qquad \text{und}$$
$$\overline{f}^U \colon V/U \to V/U, \overline{v}^U \mapsto \overline{f(v)}^U$$

wohldefinierte Endomorphismen.

Beweis. klar für $f|_U$
für $\overline{f}^U$ $\quad V \xrightarrow{f} V \xrightarrow{v \mapsto \overline{v}^U} V/U$ sind linear [→ **??**], also nach **??** auch

$$g \colon V \to V/U,\ v \mapsto \overline{f(v)}^U.$$

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 **??** ist $\overline{f}^U = \overline{g}$ wohldefiniert und linear.

Lemma 10.3.7. [→ **??**] 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 ist. Dann $\underline{v} := (v_1, \ldots, v_n)$ eine Basis von V und $M(f, \underline{v})$ von der Gestalt

$$M(f, \underline{v}) = \left(\begin{array}{c|c} M(f|_U, \underline{u}) & * \\ \hline 0 & M\left(\overline{f}^U, \underline{w}\right) \end{array}\right).$$

Beweis. $\underline{v}$ ist eine Basis von V nach **??**. 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 $f(v_j) = \sum_{i=1}^n \lambda_i v_i$ [→ **??**]. Ist $j \in \{1, \ldots, m\}$, so ist $v_j \in U$, also $f(v_j) \in U$ und daher $\lambda_{m+1} = \ldots = \lambda_n = 0$. Ist $j \in \{m+1, \ldots, n\}$, so ist

$$\overline{f}^U(v_j) \overset{??}{=} \overline{f(v_j)}^U = \overline{\sum_{i=1}^n \lambda_i v_i}^U = \sum_{i=1}^n \lambda_i \overline{v_i}^U = \sum_{i=m+1}^n \lambda_i \overline{v_i}^U.$$

Korollar 10.3.8. [→ **??**] 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. [→ **??**] Sei f ein Endomorphismus eines endlichdimensionalen Vektorraums V. Dann sind äquivalient:

(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 $M(f,\underline{v}) = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$, so

$$\chi_f \overset{??}{=} \prod_{i=1}^{n}(\lambda_i - X) = (-1)^n \prod_{i=1}^{n}(X - \lambda_i).$$

(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 $\Longleftrightarrow$ $M(f,\underline{w})$ hat untere Dreiecksgestalt.

(b) $\Longrightarrow$ (a) Induktion nach $n := \dim V$.

$n = 0$ Die 0×0-Matrix ist eine obere Dreiecksmatrix.

$n-1 \to n$ $(n \in \mathbb{N})$ Sei V ein K-Vektorraum mit $\dim V = n$ und $f \in \mathrm{End}(V)$. Es zerfalle χ_f. Wegen $n \geq 1$ hat dann χ_f eine Nullstelle $\lambda \in K$. Wähle einen Eigenvektor v_1 von f zum Eigenwert λ [$\to$ **??**]. Dann ist $U := \mathrm{span}\,(v_1)$ 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 **??** $\chi_f = \chi_{f|_U}\chi_{\overline{f}^U} = (\lambda - X)\chi_{\overline{f}^U}$, weshalb auch $\chi_{\overline{f}^U}$ zerfällt. Wegen $\dim(V/U) \overset{??}{=} n-1$ 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 **??** 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. [$\to$ **??**] 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 **??** (mit **??** statt **??**).

Bis hierher sollten wir am 3. Februar kommen.

Skript zur linearen Algebra I & II

§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 [→**??**] (falls $\mathbb{K} = \mathbb{R}$, so gilt natürlich $a^* = a$). Allgemeiner schreiben wir A^* für die komplex-konjugierte transponierte Matrix $(a_{ij}^*)_{1\le j\le n, 1\le i\le m} \in \mathbb{K}^{n\times m}$ 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 $A^* = A^T$ [→**??**]). 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) $v \neq 0 \implies \langle v, v\rangle > 0$

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 **??**(4),(5) erfüllt. Dann gilt $\langle v, 0\rangle = \langle v, 0_{\mathbb{K}} \cdot 0\rangle = 0_{\mathbb{K}} \langle v, 0\rangle = 0_{\mathbb{K}}$ 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 **??**(6) äquivalent dazu, dass für alle $v \in V$ gilt

(6') $\langle v, v\rangle \geq 0$ und $(\langle v, v\rangle = 0 \implies v = 0)$.

[1] „sesqui“ kommt aus dem Lateinischen und bedeutet „anderthalb“.

(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) $\mathbb{K}^n \times \mathbb{K}^n \to \mathbb{K},\ (x,y) \mapsto \langle x,y\rangle$ 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 **??**(6) folgt aus $\langle x,x\rangle = \sum_{i=1}^n |x_i|^2$ für alle $x \in \mathbb{K}^n$ [→**??**].

(b) $\mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R},\ (x,y) \mapsto \langle x,y\rangle$ definiert durch

$$\left\langle \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \right\rangle := x_1y_1 + 5x_1y_2 + 5x_2y_1 + 26x_2y_2 \qquad \text{für } x_1,x_2,y_1,y_2 \in \mathbb{R}$$

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 $\binom{x_1}{x_2} \in \mathbb{R}^2 \setminus \{0\}$.

(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 $\mathbb{K}^{[0,1]}$ [→**??**] 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

$$\begin{array}{rcll} \|v+w\| & \leq & \|v\| + \|w\| & \text{„Dreiecksungleichung“} \\ \|\lambda v\| & = & |\lambda|\|v\| & \text{„absolute Homogenität“} \\ v \neq 0 & \Longrightarrow & \|v\| > 0 & [\text{beachte auch } \|0\| = \|0_{\mathbb{K}}0\| = |0_{\mathbb{K}}|\|0\| = 0_{\mathbb{K}}] \end{array}$$

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 [→**??**]. So wie wir in einer abelschen Gruppe die Addition fast immer mit + notieren [→**??**(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. $\langle v+w, v+w \rangle = \langle v, v+w \rangle + \langle w, v+w \rangle = \langle v, v \rangle + \langle v, w \rangle + \langle w, v \rangle + \langle w, w \rangle$
$= \langle v, v \rangle + \langle v, w \rangle + \langle v, w \rangle^* + \langle w, w \rangle = \langle v, v \rangle + 2\,\mathrm{Re}(\langle v, w \rangle) + \langle w, w \rangle$ [→**??**]

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 **??**

$$0 \leq \langle v + x\zeta w, v + x\zeta w \rangle = \langle v, v \rangle + 2x\,\mathrm{Re}(\zeta \langle v, w \rangle) + x^2 \langle w, w \rangle.$$

Wähle $\zeta \in S$ mit $\mathrm{Re}(\zeta \langle v, w \rangle) = |\langle v, w \rangle|$ (nehme $\zeta := \frac{\langle v,w \rangle^*}{|\langle v,w \rangle|} \in S$ falls $\langle v, w \rangle \neq 0$). Dann gilt für alle $x \in \mathbb{R}$

$$0 \leq \langle v + x\zeta w, v + x\zeta w \rangle = \langle v, v \rangle + 2x|\langle v, w \rangle| + x^2 \langle w, w \rangle.$$

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

$$\begin{aligned}
\|v+w\|^2 &= \langle v+w, v+w\rangle \\
&\overset{??}{=} \langle v,v\rangle + 2\,\mathrm{Re}(\langle v,w\rangle) + \langle w,w\rangle \\
&\overset{??}{\leq} \langle v,v\rangle + 2|\langle v,w\rangle| + \langle w,w\rangle \\
&\overset{??}{\leq} \langle v,v\rangle + 2\|v\|\|w\| + \langle w,w\rangle \\
&= (\|v\| + \|w\|)^2.
\end{aligned}$$

Satz 11.1.10 (Polarisationsformel)**.** Sei V ein $\mathbb{K}$-Vektorraum mit Skalarprodukt. Dann gilt für alle $v, w \in V$:

$$4\langle v,w\rangle = \|v+w\|^2 - \|v-w\|^2 \text{ falls } \mathbb{K} = \mathbb{R} \text{ und}$$
$$4\langle v,w\rangle = \|v+w\|^2 - \|v-w\|^2 - \mathring{\imath}\|v+\mathring{\imath}w\|^2 + \mathring{\imath}\|v-\mathring{\imath}w\|^2 \text{ falls } \mathbb{K} = \mathbb{C}.$$

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 $-\mathring{\imath}(\langle v+\mathring{\imath}w, v+\mathring{\imath}w\rangle - \langle v-\mathring{\imath}w, v-\mathring{\imath}w\rangle) = -2\mathring{\imath}(\langle v, \mathring{\imath}w\rangle + \langle \mathring{\imath}w, v\rangle) = 2(\langle v,w\rangle - \langle w,v\rangle)$ 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

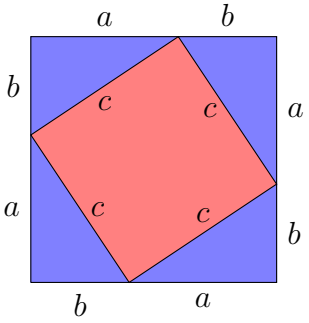
$$\cos\alpha = \frac{\langle v,w\rangle}{\|v\|\|w\|}.$$

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 **??** gilt

$$-1 \leq \frac{\langle v,w\rangle}{\|v\|\|w\|} \leq 1.$$

Bemerkung 11.1.12. (a) Die durch das Standardskalarprodukt auf dem $\mathbb{R}^2$ [→**??**(a)] induzierte Norm $\mathbb{R}^2 \to \mathbb{R},\ \left(\begin{smallmatrix} a \\ b \end{smallmatrix}\right) \mapsto \sqrt{a^2+b^2}$ [→**??**] 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 **??** 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 ($\varphi \in \mathbb{R}$) [$\to$**??**(a)] die Norm eines Vektors nicht, woraus mit der Polarisationsformel **??** folgt, dass R_φ auch das Skalarprodukt und damit Winkel im Sinne von Definition **??** nicht verändert. Es reicht also $\measuredangle(v,w)$ für den Fall zu betrachten, dass v auf der Halbachse $\mathbb{R}_{>0} \times \{0\}$ und $w \neq 0$ in der oberen Halbebene $\mathbb{R} \times \mathbb{R}_{\geq 0}$ liegt (sonst drehe geeignet). Weiter kann man $\|v\| = \|w\| = 1$ annehmen, also $v = \left(\begin{smallmatrix}1\\0\end{smallmatrix}\right)$ und $w = \left(\begin{smallmatrix}\cos\beta\\ \sin\beta\end{smallmatrix}\right)$ mit $\beta \in [0,\pi]$. Dann gilt für $\alpha := \measuredangle(v,w)$, dass $\cos\alpha = \langle \left(\begin{smallmatrix}1\\0\end{smallmatrix}\right), \left(\begin{smallmatrix}\cos\beta\\ \sin\beta\end{smallmatrix}\right)\rangle = \cos\beta$. Also gilt $\alpha = \beta$, das heißt α ist der anschauliche Winkel zwischen v und w.

(c) Die durch das Standardskalarprodukt auf dem $\mathbb{R}^3$ [$\to$**??**(a)] induzierte Norm $\mathbb{R}^3 \to \mathbb{R}$, $\left(\begin{smallmatrix}a\\b\\c\end{smallmatrix}\right) \mapsto \sqrt{a^2+b^2+c^2}$ [$\to$**??**] gibt gerade die anschauliche Länge eines Vektors im $\mathbb{R}^2$ wieder, denn nach Pythagoras aus Teil (a) ist die Länge von $\left(\begin{smallmatrix}a\\b\\0\end{smallmatrix}\right)$ gleich $\sqrt{a^2+b^2}$ und wieder nach Pythagoras ist die Länge von $\left(\begin{smallmatrix}a\\b\\c\end{smallmatrix}\right)$ 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 **??** eingeführte Winkelbegriff mit dem üblichen übereinstimmt.

(e) Sind $a_1, \ldots, a_n, b_1, \ldots, b_n \in \mathbb{R}$, so gilt im $\mathbb{C}^n$

$$\left\| \begin{pmatrix} a_1 + \mathring{\imath} b_1 \\ \vdots \\ a_n + \mathring{\imath} b_n \end{pmatrix} \right\| = \sqrt{(a_1 - \mathring{\imath} b_1)(a_1 + \mathring{\imath} b_1) + \cdots + (a_n - \mathring{\imath} b_n)(a_n + \mathring{\imath} b_n)}$$
$$= \sqrt{a_1^2 + b_1^2 + \cdots + a_n^2 + b_n^2}.$$

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 $v, w \in V \setminus \{0\}$. Dann $v \perp w \iff \measuredangle(v, w) = \frac{\pi}{2}$ nach der Definition von Winkeln **??**, denn $\cos \frac{\pi}{2} = 0$. Insbesondere stimmt unsere Definition von Senkrechtstehen in $\mathbb{R}^2$ und $\mathbb{R}^3$ nach **??** 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 $\|v\|^2 + \|w\|^2 = \|v + w\|^2$ [→**??**].

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 **??** gerechtfertigt). Man könnte daher sagen, dass der *eigentliche* Beweis des Satzes von Pythagoras in **??**(a) steht. Es stellt aber **??**(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. [→**??**] 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 $\mathbb{K}^n$ [→**??**] ist eine ONB des $\mathbb{K}^n$ (versehen mit dem Standardskalarprodukt [→**??**(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 $j \in \{1, \dots, m\}$

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 $v \in V$. 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 [→**??**, **??**] 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 [→**??**(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).

Bis hierher sollten wir am 7. Februar kommen.

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

$$\operatorname{coord}_{\underline{v}}(v) = \begin{pmatrix} \langle v_1, v\rangle \\ \vdots \\ \langle v_n, w\rangle \end{pmatrix} \qquad \text{für jedes } v \in V.$$

Beweis. Die orthogonale Projektion eines $v \in V$ auf V ist v selber [→**??**(a)]. Also gilt nach **??** $v = \sum_{i=1}^{n} \langle v_i, v\rangle v_i$ 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 [→**??**]. 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. [→**??**] Jeder endlichdimensionale Vektorraum mit Skalarprodukt besitzt eine ONB.

Beweis. Induktion nach der Dimension, wobei der Induktionsschritt mit Gram-Schmidt **??** 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 **??** eine ONB $(v_1,\dots,v_m)$ von U. Nach **??** 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 **??**(3),(4) rechnet man nun sofort die Linearität von P_U nach.

Weiter gilt

$$\ker P_U = \left\{ v \in V \mid \sum_{i=1}^{m} \langle v_i, v\rangle v_i = 0 \right\} \overset{??}{=} \{v \in V \mid \langle v_1, v\rangle = \ldots = \langle v_m, v\rangle = 0\}$$
$$\overset{??}{=} (\operatorname{span}(v_1, \ldots, v_m))^\perp = U^\perp,$$

$U \subseteq \operatorname{im} P_U$ nach **??**(a) und selbstverständlich $\operatorname{im} P_U \subseteq U$ nach **??**.

Proposition 11.2.17. Sei U ein endlichdimensionaler Unterraum des Vektorraums mit Skalarprodukt V. Dann gilt $V = U \oplus U^\perp$ [→**??**]. Für endlichdimensionales V gilt insbesondere $\dim(U) + \dim(U^\perp) = \dim(V)$ [→**??**].

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 $u = -v \in U \cap U^\perp \overset{??}{=} \{0\}$.

Korollar 11.2.18. Sei U ein Unterraum des endlichdimensionalen $\mathbb{K}$-Vektorraums mit Skalarprodukt V. Dann $U = (U^\perp)^\perp$.

Beweis. Nach **??** gilt $U \subseteq (U^\perp)^\perp$ und mit **??** angewandt auf U und auf $U^\perp$ haben wir $\dim(U) = \dim(V) - \dim(U^\perp) = \dim((U^\perp)^\perp)$. Benutze nun **??**.

Definition 11.2.19. Seien V und W $\mathbb{K}$-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 **??** 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 **??** bereits bemerkt ist $R_\varphi \in \operatorname{End}(\mathbb{R}^2)$ 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 $(f(v_1), \dots, f(v_n))$ eine ONB von W. Nach **??** 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 Orthonormalbasen $(v_1, \dots, v_n)$ von V und $(w_1, \dots, w_n)$ von W. Dann ist die lineare Abbildung $f\colon V \to W$ mit $f(v_i) = w_i$ für $i \in \{1, \dots, n\}$ [→**??**] ein Isomorphismus von Vektorräumen mit Skalarprodukt.

Korollar 11.2.24. Sei V ein Vektorraum mit Skalarprodukt und $\underline{v} = (v_1, \dots, v_n)$ 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 Orthonormalbasen $\underline{v} = (v_1, \dots, v_n)$ und $\underline{w} = (w_1, \dots, w_n)$. 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 **??** 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 **??** 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 **??**, **??** und **??** 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 **??** folgen (b) $\iff$ (d) und (c) $\iff$ (e). Schließlich gilt (d) $\iff$ (e) $\iff$ (f) wegen **??**.

Beispiel 11.2.28. [$\to$**??**(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 [$\to$**??**] ist selbstadjungiert. In der Tat: Wegen $V = U + U^\perp$ [$\to$**??**] 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 $f = \mathrm{vec}_{\underline{v}} \circ f_{M(f,\underline{v})} \circ \mathrm{coord}_{\underline{v}}$ [$\to$**??**] 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 **??** 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

$$(AB)^T \overset{??}{\underset{??}{=}} \left(\sum_{j=1}^{n} a_{ij}b_{jk}\right)_{1\le k\le r, 1\le i\le m} = \left(\sum_{j=1}^{n} b_{jk}a_{ij}\right)_{1\le k\le r, 1\le i\le m} \overset{??}{\underset{??}{=}} B^T A^T.$$

(b) Ist $K = \mathbb{K}$, so gilt

$$(AB)^* \overset{??}{\underset{??}{=}} \left(\sum_{j=1}^{n} (a_{ij}b_{jk})^*\right)_{1\le k\le r, 1\le i\le m} \overset{??}{=} \left(\sum_{j=1}^{n} b_{jk}^* a_{ij}^*\right)_{1\le k\le r, 1\le i\le m} \overset{??}{\underset{??}{=}} B^* A^*.$$

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. $\langle A^* x, y\rangle \overset{??(a)}{=} (A^* x)^* y \overset{??}{=} x^*(A^*)^* y = x^* A y \overset{??(a)}{=} \langle x, Ay\rangle$

Proposition 11.3.7. Für $A \in \mathbb{K}^{n\times n}$ gilt

$$A \text{ selbstadjungiert} \iff A^* = A.$$

Beweis.

$$\begin{aligned}
A \text{ selbstadjungiert} &\overset{??}{\underset{??}{\iff}} \forall x, y \in \mathbb{K}^n\colon \langle Ax, y\rangle = \langle x, Ay\rangle\\
&\overset{??}{\iff} \forall x, y \in \mathbb{K}^n\colon \langle Ax, y\rangle = \langle A^* x, y\rangle\\
&\iff A = A^*
\end{aligned}$$

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 [→**??**, **??**(e)], das heißt es gibt $x \in \mathbb{C}^n \setminus \{0\}$ mit $Ax = \lambda x$. Es folgt

$$\lambda\langle x, x\rangle \overset{\text{??}(4)}{=} \langle x, \lambda x\rangle = \langle x, Ax\rangle \underset{\text{??}}{\overset{\text{??}}{=}} \langle Ax, x\rangle = \langle \lambda x, x\rangle \overset{\text{??}(2)}{=} \lambda^*\langle x, x\rangle$$

und daher $\lambda = \lambda^*$ nach **??**(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 [→**??**(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 **??**, dass f einen *reellen* Eigenwert λ besitzt: Wählt man nämlich eine ONB $\underline{w}$ von V [→**??**] und setzt $A := M(f, \underline{w})$, so ist A nach **??** selbstadjungiert und daher jedes $\lambda \in \mathbb{C}$ mit $\chi_A(\lambda) = 0$ nach Lemma **??** sogar aus $\mathbb{R}$. Nun gilt aber $\chi_f = \chi_A$ [→**??**, **??**(e)] und es gibt ein solches λ wegen $\deg \chi_f \overset{\text{??}}{=} n \geq 1$ nach dem Fundamentalsatz der Algebra **??**.
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 $1 + \dim(U^\perp) = \dim(U) + \dim(U^\perp) \overset{\text{??}}{=} \dim(V) = n$, 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 **??** eine ONB $\underline{v}$ von $\mathbb{K}^n$, die aus Eigenvektoren von A zu reellen Eigenwerten besteht. Nach Satz **??** ist $P := M(\underline{e}, \underline{v})$ [→**??**] dann orthogonal und daher

$$P^* \overset{\text{??}(f)}{=} P^{-1} \overset{\text{??}}{=} M(\underline{v}, \underline{e}).$$

Also

$$A \overset{\text{??}(e)}{=} M(f_A, \underline{e}) \overset{\text{??}}{=} M(\underline{v}, \underline{e})M(f_A, \underline{v})M(\underline{e}, \underline{v}) = P^*DP,$$

[2] Im Beweis dieses Satzes benutzen wir den Fundamentalsatz der Algebra **??** [→**??**(g)].
[3] Im Beweis dieses Korollars benutzen wir den Fundamentalsatz der Algebra **??** [→**??**(g)].

wobei $D := M(f_A, \underline{v})$ eine reelle Diagonalmatrix ist.

Bis hierher sollten wir am 10. Februar kommen.

§12 Halbordnungen und Ordnungen

12.1 Infima und Suprema, Minima und Maxima

Definition 12.1.1. Sei A eine Menge. Eine *Halbordnung* (auch: *partielle Ordnung*) ist eine Relation [→ **??** (e)] $\preceq$ auf A, für die gilt:

- $\forall a \in A : a \preceq a$ - "reflexiv",
- $\forall a, b \in A : (a \preceq b) \wedge (b \preceq b) \implies a = b$ - "antisymmetrisch",
- $\forall a, b, c \in A : (a \preceq b) \wedge (b \preceq c) \implies a \preceq c$ - "transitiv".

Gilt zusätzlich

- $\forall a, b \in A : (a \preceq b) \vee (b \preceq a)$ - "linear",

so heißt $\preceq$ sogar eine *Ordnung* (auch: *lineare Ordnung*, *Totalordnung).*

Beispiel 12.1.2. (a) Sei A eine Menge von Mengen. Dann wird durch

$$\forall N, M \in A : N \preceq M :\iff N \subseteq M$$

per Mengeninklusion eine Halbordnung auf A definiert.

(b) Die Teilsbarkeitsbeziehung auf $\mathbb{N}$ definiert durch

$$a \mid b :\iff \exists c \in \mathbb{N} : b = ca.$$

ist eine Halbordnung auf den natürlichen Zahlen.

(c) Die *natürliche Ordnung* $\leq$ auf $\mathbb{N}$ ist eine Ordnung.

(d) Die rationale Ordnung $\leq$ auf $\{\infty\} \cup \mathbb{R} \cup \{-\infty\}$ ist eine Ordnung.

Bemerkung 12.1.3. (a) Ist $\preceq$ eine (Halb-)Ordnung auf A, dann auch $\succeq$ vermöge der Relation $\forall a, b \in A : a \succeq b :\iff b \preceq a$. Es ist $\succeq$ die zu $\preceq$ *inverse* (Halb-)Ordnung.

(b) Ist $\preceq$ eine (Halb-)Ordnung auf A, so finden wir auch eine Halbordnung für ein $B \subseteq A$ über $\preceq' := \preceq \cap (B \times B)$. Es ist $\preceq'$ die *Einschränkung* von $\preceq$ auf B.

Definition 12.1.4. Ein geordnetes Paar $(A, \preceq)$ bestehend aus einer Menge A und einer (Halb-)Ordnung $\preceq$ auf A heißt (halb-)geordnete Menge. Wir nennen A die zugrundeliegende Menge (auch: Trägermenge/ Universum) und $\preceq$ die (Halb-)Ordnung von $(A, \preceq)$.

Bemerkung 12.1.5. Wie immer pflegen wir einen "schlamigen" Sprachgebrauch, z.B. "Sei A eine geordnete Menge und seien $a, b \in A$ mit $a \preceq b$".

Definition 12.1.6. Sei $(A, \preceq)$ eine halbgeordnete Menge und $B \subseteq A$. Ein Elemente $a \in A$ heißt $\left\{\begin{matrix} \textit{obere} \\ \textit{untere} \end{matrix}\right\}$ Schranke von B, wenn $\forall b \in B : \left\{\begin{matrix} a \preceq b \\ b \preceq a \end{matrix}\right\}$ gilt. Wir notieren

- $\mathrm{ub}(B) := \{a \in A \mid a \text{ obere Schranke}\}$
- $\mathrm{lb}(B) := \{a \in A \mid a \text{ untere Schranke}\}$.

Ein Element $a \in A$ heißt $\left\{\begin{matrix} \textit{Minimum} \text{ (auch: kleinstes Element)} \\ \textit{Maximum} \text{ (auch: größtes Element)} \end{matrix}\right\}$, wenn $\left\{\begin{matrix} a \in B \cap \mathrm{lb}(B) \\ a \in B \cap \mathrm{ub}(B) \end{matrix}\right\}$ gilt. [Wenn existent, ist ein solches Element wegen der Antisymmetrie von $\preceq$ eindeutig und wir notieren dieses mit $\left\{\begin{matrix} \min B \\ \max B \end{matrix}\right\}$.]

Ein Element $a \in A$ heißt $\left\{\begin{matrix} \textit{Infimum} \text{ (auch: größte untere Schranke)} \\ \textit{Supremum} \text{ (auch: kleinste obere Schranke)} \end{matrix}\right\}$ von B, wenn $\left\{\begin{matrix} a = \max \mathrm{lb}(B) \\ a = \min \mathrm{ub}(B) \end{matrix}\right\}$ gilt. Wenn existent, notieren wir ein solches Element mit $\left\{\begin{matrix} \inf B \\ \sup B \end{matrix}\right\}$.

Proposition 12.1.7. Sei A eine halbgeordnete Menge und $B \subseteq A$.

(a) Besitzt B ein $\left\{\begin{matrix} \text{Minimum} \\ \text{Maximum} \end{matrix}\right\}$, so auch ein $\left\{\begin{matrix} \text{Infimum} \\ \text{Supremum} \end{matrix}\right\}$ und es gilt $\left\{\begin{matrix} \inf B = \min B \\ \sup B = \max B \end{matrix}\right\}$.

(b) Besitzt B ein $\left\{\begin{matrix} \text{Infimum} \\ \text{Supremum} \end{matrix}\right\}$, so ist dieses genau dann das $\left\{\begin{matrix} \text{Minimum} \\ \text{Maximum} \end{matrix}\right\}$, wenn dieses in B liegt.

(c) Sei $a \in A$. Dann ist a genau dann $\left\{\begin{matrix} \text{Infimum} \\ \text{Supremum} \end{matrix}\right\}$ von B, wenn gilt $\forall c \in A : \left\{\begin{matrix} c \in \mathrm{lb}(B) \iff c \preceq a \\ c \in \mathrm{ub}(B) \iff a \preceq c \end{matrix}\right\}$.

Beweis. Wir zeigen die Aussagen je für das Infimum/ Minimum. Der entsprechende Beweis für das Supremum/ Maximum folgt dann über die Halbordnung $\succeq$.

(a) Sei $m := \min B$. Dann gilt $\forall b \in B : m \preceq b$ und wegen $m \in B$ auch $\forall l \in \mathrm{lb}(B) : l \preceq b$. Insgesamt folgt $b = \max \mathrm{lb}(B) = \inf B$.

(b) Sei $i := \inf B$. Ist i das Minimum von B, liegt dieses per Definition bereits darin. Liegt i in B, dann folgt aus $i \in \mathrm{lb}(B)$ die gewünschte Minimaleigenschaft.

(c) Sei $a \in A$. Es gilt

$$\begin{aligned}(a = \inf A) &\iff (a \in \max \mathrm{lb}(B)) \\ &\iff (\forall c \in A : c \in \mathrm{lb}(B) \Rightarrow c \preceq b) \wedge a \in \mathrm{lb}(B).\end{aligned}$$

Die letzte Aussage ist erkennbar äquivalent zu der Behauptung, da $a \preceq a$.

Beispiel 12.1.8. (a) Sei M eine Menge und $A := \mathscr{P}(M)$ durch Inklusion halbgeordnet. Dann gilt für $B \subseteq A$:

- $\inf B := \begin{cases} M & \text{wenn } B = \emptyset \\ \bigcap B & \text{sonst} \end{cases}$
- $\sup B := \bigcup B$

(b) Sei V ein K-Vektorraum und

$$A := \{U \mid U \text{ Unterraum von } V\}$$

durch Inklusion halbgeordnet. Dann gilt für $B \subseteq A$:

- $\inf B := \begin{cases} V & \text{wenn } B = \emptyset \\ \bigcap B & \text{sonst} \end{cases}$.
- $\sup B := \operatorname{span}(\bigcup B)$

(c) Sei

$$A := \{\underbrace{\{1\}}_{=a}, \underbrace{\{2\}}_{=b}, \underbrace{\{3\}}_{=c}, \underbrace{\{1,2\}}_{=d}, \underbrace{\{2,3\}}_{=e}\}$$

durch Inklusion halbgeordnet. Es gilt:

B	$\mathrm{lb}(B)$	$\mathrm{ub}(B)$	$\inf(B)$	$\min(B)$	$\sup(B)$	$\max(B)$
$\emptyset$	A	A	–	–	–	–
$\{a\}$	$\{a\}$	$\{a,d\}$	a	a	a	a
$\{b\}$	$\{b\}$	$\{b,d,e\}$	b	b	b	b
$\{d\}$	$\{a,b,d\}$	$\{d\}$	e	e	e	e
$\{a,b\}$	–	$\{d\}$	–	–	d	–
$\{d,e\}$	–	$\{b\}$	b	–	–	–

(d) In $(\mathbb{N}, |)$ gilt für $B \subseteq \mathbb{N}$:

- $\mathrm{lb}(B)$ ist die Menge der gemeinsamen Teiler B.
- $\mathrm{ub}(B)$ ist die Menge der gemeinsamen Vielfachen von V.
- $\inf(\emptyset)$ existiert nicht wegen der Unbeschränktheit von $\mathrm{lb}(\emptyset) = \mathbb{N}$.

Für $B \neq \emptyset$ gilt weiter:

- $\inf(B) = \underbrace{\text{größter}}_{\text{in Teilerrelation}}$ gemeinsamer Teiler aller Zahlen in B
- $\sup(B) = \underbrace{\text{kleinste}}_{\text{in Teilerrelation}}$ gemeinsame Vielfache aller Zahlen in B, wenn existent!

(e) In $\mathbb{N}$ mit der natürlichen Ordnung gilt:

- $\sup \emptyset = 1$, aber das Maximum von $\emptyset$ existiert nicht. Ebenso existieren das Infimum und Minimum von $\emptyset$ nicht.
- ist $B \neq \emptyset$, dann hat B immer ein Minimum und somit Infimum. Die Begriffe von Maximum und Supremum gleichen sich hier. Also

$$B \text{ hat Maximum} \iff B \text{ hat Minimum} \iff B \text{ ist endlich.}$$

(f) In der auf natürliche Weise geordneten Menge $\mathbb{R}_\infty := \{-\infty\} \cup \mathbb{R} \cup \{\infty\}$ gilt: per Supremumsaxiom in $\mathbb{R}$ und hinzugefügten Elementen zu $\mathbb{R}_\infty$ hat jede Teilmenge $B \subseteq \mathbb{R}_\infty$ ein Infimum und Supremum. Insbesondere gilt für das Infimum:

- $\inf B = -\infty \iff B$ hat keine reelle untere Schranke
- $\inf B = \infty \iff B \in \{\emptyset, \{\infty\}\}$

Anaolog gilt für das Supremum:

- $\sup B = \infty \iff B$ hat keine reelle obere Schranke
- $\sup B = -\infty \iff B \in \{\emptyset, \{-\infty\}\}$

12.2 Das Zornsche Lemma

Notation 12.2.1. Wird eine Halbordnung R auf A durch $\preceq$, $\leq$, $\succeq$ oder $\geq$ notiert, so bezeichnen man mit $\prec$, $<$, $\succ$ oder $>$ die Relation R' auf A definiert durch

$$\forall a, b \in A : aR'b :\iff (aRb \wedge a \neq b)$$

Definition 12.2.2. Sei $(A \preceq)$ eine halbgeordnete Menge und $B \subseteq A$. Mann nennt b ein $\left\{\begin{matrix} \textit{maximales} \\ \textit{minimales} \end{matrix}\right\}$ Element von B, wenn $b \in B$ und $\left\{\begin{matrix} \nexists c \in B : b \prec c \\ \nexists c \in B : c \prec b \end{matrix}\right\}$ gilt.

Bemerkung 12.2.3. Dies ist nicht zu verwechseln mit den in **??** eingeführten Begriffen. Alle Maxima und Minima sind zwar auch maximale bzw. minimale Elemente, die Umkehrung dieser Aussage gilt jedoch im Allgemeinen nicht. In z.B. $(\mathbb{N}, |)$ ist 3 ein maximales Element von $\{3, 28\}$, aber kein Maximum. Diese Menge besitzt kein Maximum.

Definition 12.2.4. Sei $(A, \preceq)$ eine halbgeordnete Menge und $C \subseteq A$. Dann heißt C eine *Kette* (in $(A, \preceq)$), wenn die Einschränkung von $\preceq$ auf C eine Ordnung ist, d.h. zusätzlich gilt:

- $\forall b, c \in C : (b \preceq c) \vee (c \preceq b)$ - "Totalität"

Anschaulich wird C durch $\preceq|_{C \times C}$ linear zu einer "Kette".

Beispiel 12.2.5. (a) Es ist $B := \{\emptyset, \{2\}, \{1, 2, 4\}\}$ eine Kette in $(\mathscr{P}(\{1, 2, 3, 4\}), \subseteq)$.

(b) Die Menge $\{2^n \mid n \in \mathbb{N}\}$ ist eine Kette in $(\mathbb{N}, |)$, aber nicht $\{2, 3\}$ denn $2 \nmid 3$ und $3 \nmid 2$.

Lemma 12.2.6. Sei M eine Menge und $\mathscr{P}(M)$ durch Inklusion halbgeordnet. Sei $B \in \mathscr{P}(M)$ derart, dass für jede Kette $C \subseteq B$ gilt $\bigcup C \in B$. Es ist B gewissermaßen abgeschlossen unter Supremumsbildung. Ferner sei $f : B \to B$ eine Abbildung mit $\forall N \in B : N \subseteq f(N)$. Dann gibt es $N_0 \in B$ mit $N_0 = f(N_0)$.

Beweis. Wir nennen $B' \subseteq B$ als f-induktiv, wenn $\forall N \in B' : f(N) \in B'$ und $\bigcup C \in B'$ für jede Kette $C \subseteq B'$ gilt. Somit ist B per Voraussetzung f-induktiv.

Schritt 1: Man darf annehmen, dass keine echte Teilmenge von B f-induktiv ist.
Begründung: Wir zeigen, dass eine kleinste f-induktive Teilmenge von B existiert ("minimalste" würde aber reichen!). Setze

$$B'' := \bigcap \{B' \subseteq B \mid B' \text{ ist} f\text{-induktiv}\}.$$

Klarerweise ist B'' dann f-induktiv. Wir könnten dann mit B'' statt B arbeiten - da aber $B'' \subseteq B$, folgt die Aussage somit auch für B. Œ sei also $B = B''$.

Wir definieren die Menge der Fixpunkte als solche, für welche alle Teilmengen auch dessen f-Vergrößerung kleiner als diese ist:

$$B' := \{N' \in B \mid \forall N \in B : (N \subset N' \implies f(N) \subseteq N')\}.$$

Für alle $N' \in B'$ definieren wir zudem

$$B'_N := \{N \in B \mid (N \subseteq N') \vee (f(N') \subseteq N)\}.$$

Dies ist also die Menge aller Punkte unter N' (in der Halbordnung auf $\mathscr{P}(M)$) und alle Punkte, welcher größer als die f-Vergrößerung von N' ist. Es fehlen alle Punkte echt zwischen N' und $f(N')$.

Schritt 2: Für jedes $N' \in B'$ ist B'_N f-induktiv.
Begründung: Sei $N' \in B'$. Zu zeigen ist

(b) $\forall N \in B'_N : f(N) \in B'_N$,

(b) Für jede Kette $C \subseteq B'_N$ gilt $\bigcup C \in B'_N$.

(a) Sei $N \in B'_N$. Dann $N \in B$ und $N \subseteq N'$ oder $f(N') \subseteq N$. Zu zeigen ist $f(N) \in B'_N$. $f(N) \in B$ gilt immer. Ist $N \subseteq N'$, folgt entweder $N = N'$ oder

$N \subset N'$. Ersterer ist klar wegen $N' \subseteq N'$. Aus letzterem folgt $f(N) \subseteq N'$, da N' ein Fixpunkt ist. Ist $f(N') \subseteq N$, dann folgt wegen $N \subseteq f(N)$ auch $f(N') \subseteq f(N)$.

(b) Sei $C \subseteq B'_N$ eine Kette. Zu zeigen ist $\bigcup C \in B'_N$. Es gilt immer $\bigcup C \in B$. Zu zeigen ist $\bigcup C \subseteq N'$ oder $f(N') \subseteq \bigcup C$. Gelte ersteres nicht, zu zeigen ist dann letzteres. Wir können $N \in C$ mit $N \subsetneq N'$ wählen. Mit $N \in B'_N$ folgt $f(N') \subseteq N \subseteq \bigcup C$.

Insgesamt folgt für alle $N' \in B'$, dass $B'_N = N$ gilt. Also ist

$$(*)\ \forall N' \in B' : \forall N \in B : (N \subseteq N' \vee f(N') \subseteq N').$$

Schritt 3: B' ist f-induktiv.
Begründung: Zu zeigen ist

(a) $\forall N \in B' : f(N) \in B'$,

(b) Für jede Kette $C \subseteq B'$ gilt $\bigcup C \in B'$.

(a) Sei $N' \subseteq B'$. Zu zeigen ist $f(N') \in B$ und $\forall N \in B : (N \subset f(N') \implies f(N) \subseteq f(N'))$. Ersteres ist klar. Sei nun $N \in B$ mit $N \subset f(N')$. Zu zeigen ist $f(N) \subseteq f(N')$. Wegen (*) folgt aber $N \subseteq N'$. Sei $N \neq N'$. Dann $f(N) \subseteq N' \subseteq f(N')$, da N' ein Fixpunkt ist.

(b) Sei $C \subseteq B'$ eine Kette. Zu zeigen ist $\bigcup C \in B$ und $\forall N \in B : (N \subset \bigcup C) \implies f(N) \subseteq \bigcup C)$. Ersteres ist klar. Sei nun $N \in B$ mit $N \subset \bigcup C$. Zu zeigen ist $f(N) \subseteq \bigcup C$. Wegen $\bigcup C \subsetneq N$ gibt es $N' \in C \subseteq B'$ mit $N' \subsetneq N$. Insbesondere ist $f(N') \subseteq N$. Wegen (*) folgt $N \subseteq N'$ und hier sogar $N \subset N'$. Also $f(N) \subseteq N' \subseteq \bigcup C$ und wir sind fertig.

Wegen *Schritt 1* ist $B' = B$, weshalb aus (*) folgt, dass B eine Kette ist. Da B f-induktiv ist, folgt $N_0 := \bigcup B \in B$ und $f(N_0) \in B$. Daraus folgt $N_0 \subseteq f(N_0) \subseteq \bigcup B = N_0$, also $N_0 = f(N_0)$.

Lemma 12.2.7. Sei A eine halbgeordnete Menge. Dann gibt es eine maximale Kette C in A.

Beweis. Bezeichne $\mathscr{C}$ die durch Inklusion halbgeordnete Menge aller Ketten in A:

$$\mathscr{C} := \{C \subseteq A \mid C \text{ ist Kette}\}.$$

Nehmen wir an, es gäbe kein maximales Element in $\mathscr{C}$, also zu jedem $C \in \mathscr{C}$ gibt es ein $f(C) \in \mathscr{C}$ mit $C \subset f(C)$. Dies definiert eine Abbildung $f : \mathscr{C} \to \mathscr{C}$. Wir zeigen nun, für jede Kette $K \subseteq \mathscr{C}$ gilt $\bigcup K \in \mathscr{C}$. Sei also $K \subseteq \mathscr{C}$ eine Kette und seien $a, b \in \bigcup K$. Zu zeigen ist $a \preceq b \vee b \preceq a$. Wählen wir nun $C_1, C_2 \in K$ so,

dass $a \in C_1$ und $b \in C_2$. Da K eine Kette ist, folgt $(C_1 \subseteq C_2) \vee (C_2 \subseteq C_1)$. gelte ersteres. Dann $a, b \in C_2$ und da C_2 eine Kette ist, folgt die Behauptung.

Insgesamt ist dann $\mathscr{C}$ f-induktiv. Mit Lemma (**??**) finden wir dann aber eine Kette $C_0 \in \mathscr{C}$ mit $C_0 = f(C_0)$, was unserer Annahme widerspricht. Per Widerspruch sind wir somit fertig.

Satz 12.2.8. Als *Lemma von Zorn* $\begin{bmatrix}\text{Max August Zorn}\\ *1906,\dagger1993\end{bmatrix}$ wird der folgende Satz von Kuratowski $\begin{bmatrix}\text{Kazimierz, Kuratowski}\\ *1896,\dagger1980\end{bmatrix}$ bezeichnet.
Sei A eine halbgeordnete Menge derat, dass jede Kette in A eine obere Schranke in A besitzt. Dann besitzt A ein maximales Element.

Beweis. Wähle nach Lemma (**??**) eine maximale Kette C in A. Wähle $a \in \mathrm{ub}(C)$ gemäß Voraussetzung. Wir behaupten, dass a ein maximales Element von A ist. Sei hierfür $c \in A$ mit $a \preceq c$. Zu zeigen ist $a = c$. Es ist wohl dann $C \cup \{a, c\}$ eine Kette in A. Da C maximal ist, folgt aber $a, c \in C$. Da $a \in \mathrm{ub}(C)$, muss $c \preceq a$ gelten und per Antisymmetrie also $a = c$.

Bemerkung 12.2.9. Dass die leere Kette in einer halbgeordneten Menge A eine obere Schranke besitzt, führt dazu, dass A nichtleer sein darf. Deshalb formulieren manche Autoren: "Sei A eine nichtleere halgeordnete Menge derart, dass jede nichtleere Kette...".

Korollar 12.2.10. Sei A eine halbgeordnete Menge derart, dass jede nichtleere Kette in A eine obere Schranke in A besitzt. Dann gibt es zu jedem $a \in A$ ein maximales Element $b \in A$ mit $a \preceq b$.

Beweis. Betrachte $A' := \mathrm{ub}(\{a\})$ mit der auf A' eingeschränkten Halbordnung. Wende das Zornsche Lemma auf A' an und erhalte ein maximales Element b von A' zu erhalten. Es ist b auch ein maximales Element von A, denn ist $c \in A$ mit $b \preceq c$, so gilt auch $c \in A'$. Also folgt $b = c$.

Korollar 12.2.11. Sei A eine durch Inklusion halbgeordnete Menge von Mengen derart, dass für jede nichtleere Kette $C \subseteq A$ gilt $\begin{Bmatrix}\bigcup C \in A\\ \bigcap C \in A\end{Bmatrix}$. Dann gibt es zu jedem $M \in A$ ein $\begin{Bmatrix}\text{maximales}\\ \text{minimales}\end{Bmatrix}$ $N \in A$ mit $\begin{Bmatrix}M \subseteq N\\ N \subseteq M\end{Bmatrix}$.

Bemerkung 12.2.12. Wie in **??** angedeutet, gibt es bei der Mengenbildung Spielregeln einzuhalten. Dies wird auch durch obiges Korollar deutlich, welches im Widerspruch zur Existenz der Menge aller Mengen steht. Denn diese hat keine maximalen Elemente.

Anschauung. Das Zorn'sche Lemma besagt, dass Mengen nicht furchtbar groß sein können, was folgender Gedanke verdeutlicht: Sei A eine Menge, die alle Voraussetzung für das Lemma von Zorn erfüllt. Wie in Bemerkung **??** erläutert, muss es ein Element $a_1 \in A$ geben. Ist dieses nicht maximal, finden wir $a_2, a_3...$, die größer als a_1 sind und erhalten die Kette

$$a_1, a_2, a_3, ...$$

Im schlimmsten Fall ist diese unendlich. Da aber jede Kette nach Voraussetzung ein maximales Element hat, gibt es ein b_1 in A, das größer als $a_1, a_2, a_3, ...$ ist und weiter eine Kette $b_1, b_2, b_3, ...$ Irgendwann erhält man so im schlimmsten Fall unendlich viele Ketten

$$\begin{aligned} &a_1, a_2, a_3, ... \\ &b_1, b_2, b_3, ... \\ &c_1, c_2, c_3, ... \\ &..., \end{aligned}$$

wobei sichtbar $a_1, b_1, c_1, ...$ wieder eine Kette ist, mit welcher man gleiche Überlegungen wie zu der Kette $a_1, a_2, a_3, ...$ anstellen kann. Nach dem Zorn'schen Lemma muss aber irgendwann damit Schluss sein - es muss ein maximales Element geben.

12.3 Existenz von Basen in beliebigen Vektorräumen

Satz 12.3.1. [→ **??**] Sei V ein Vektorraum, $F \subseteq G \subseteq V$. Sei ferner F linear unabhängig in V und G ein Erzeugendensystem von V. Betrachte die durch Inklusion halbgeordnete Menge

$$\mathscr{A} := \{C \mid F \subseteq C \subseteq G\}.$$

Dann sind äquivalent folgende Aussagen für alle $F \subseteq B \subseteq G$:

(a) B ist eine Basis von V,

(b) B ist ein minimales Element von

$$\{C \in \mathscr{A} \mid C \in \text{ Erzeugendensystem von } V\},$$

(c) B ist ein maximales Element von

$$\{C \in \mathscr{A} \mid C \in \text{ linear unabhängig in } V\}.$$

Beweis. Schon gezeigt.

Korollar 12.3.2. Sei V ein Vektorraum, $F \subseteq G \subseteq V$. Sei ferner F linear unabhängig in V und G ein Erzeugendensystem von V. Dann gibt es eine Basis B von V mit $F \subseteq B \subseteq G$.

Beweis. Betrachte die durch Inklusion halbgeordnete Menge

$$\mathscr{B} := \{C \mid F \subseteq C \subseteq G, C \text{ linear unabhängig in } V\}.$$

Gemäß **??**(c) ist zu zeigen, dass $\mathscr{B}$ ein maximales Element besitzt. Nach dem Zornschen Lemma reicht es hierfür zu zeigen, dass jede Kette in $\mathscr{B}$ eine obere Schranke besitzt. Sei also $K \subseteq \mathscr{B}$ eine Kette. Falls $K = \emptyset$, so ist F eine obere Schranke von K. Sei also $K \neq \emptyset$. Wir behaupten, dass

$$C := \bigcup K$$

eine obere Schranke von K ist. Es ist $F \subseteq C \subseteq G$ klar. Zu zeigen bleibt, dass C linear unabhängig ist. Hierfür zeigen wir, dass jede endliche Teilmenge von C linear unabhängig ist. Sei also $C' \subseteq C$ endlich. Da $C' \subseteq K$, gibt es eine endliche Teilkette $K' \subseteq K$ mit $C' \subseteq \bigcup K'$. gelte $K' \neq \emptyset$. Da K' eine endliche nichtleere Kette ist, hat sie ein maximales Element $m_k := \bigcup K'$. Daher gilt gilt $\bigcup K' \in K' \subseteq K \subseteq \mathscr{B}$. Insbesondere ist K' linear unabhängig und damit auch C'.

Skript zur linearen Algebra I & II

§13 Linearformen, Bilinearformen und quadratische Formen

In diesem Abschnitt sei K stets ein Körper.

13.1 Algebraischer Dualraum

Definition 13.1.1. Sei V ein K-Vektorraum. Eine *Linearform* auf V ist eine lineare Funktion $V \to K$. Der K-Vektorraum

$$V^* := \mathrm{Hom}(V, K)$$

aller Linearformen auf V heißt der *(algebraische) Dualraum* von V.

Erinnerung 13.1.2. Die K-Vektorraumstruktur auf V^* ist gegeben durch $(l_l + l_2)(v) = l_1(v) + l_2(v)$, $(\lambda l)(v) = \lambda\, l(v)$ für alle $l, l_1, l_2 \in V^*, \lambda \in K$ und $v \in V$.

Bemerkung 13.1.3. Sei V ein K-Vektorraum mit Basis $\underline{v} = (v_1, ...v_n)$. Bezeichne mit $\underline{e} = (1)$ die Standardbasis des K-Vektorraums K. Dann ist für $l \in V^*$

$$M(l, \underline{v}, \underline{e}) = \begin{pmatrix} l(v_1) \\ \vdots \\ l(v_n) \end{pmatrix}^T$$

ein Zeilenvektor. Folgender Satz ist daher ein Speziallfall von **??**.

Satz 13.1.4. Sei V ein K-Vektorraum mit Basis $\underline{v} = (v_1, ...v_n)$. Dann sind $\Phi : \begin{cases} V^* \to K^n \\ l \mapsto \begin{pmatrix} l(v_1) \\ \vdots \\ l(v_n) \end{pmatrix} \end{cases}$

und $\Psi : \begin{cases} K^n \to V^* \\ a \mapsto \begin{pmatrix} V \to K \\ \sum_{i=1}^n \lambda_i v_i \mapsto a^T \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} \end{pmatrix} \end{cases}$ zueinander inverse K-Vektorraumisomorphismen

Beweis. Schon gezeigt.

Korollar und Definition 13.1.5. Sei V ein K-Vektorraum mit Basis $\underline{e} = (v_1, ...v_n)$. Definiere $v_1^*, ..., v_n^* \in V^*$ durch

$$v_i^*(v_j) := \begin{cases} 1 & \text{falls } i = j \\ 0 & \text{sonst} \end{cases} \quad (\forall i, j \in \{1...n\}).$$

Dann ist $\underline{v}^* := (v_1^*, ...v_n^*)$ eine Basis von V^*, genannt die zu $\underline{v}$ *duale Basis*.

Beweis. Es ist $\underline{v}^* = (\Psi(e_1), ..., \Psi(e_n))$ eine Basis, da Ψ ein Vektorraumisomorphismus ist.

Korollar 13.1.6. Sei $n \in \mathbb{N}_0$. Dann sind $\Phi : \begin{cases} (K^n)^* \to K^n \\ l \mapsto \begin{pmatrix} l(e_1) \\ \vdots \\ l(e_n) \end{pmatrix} \end{cases}$ und $\Psi : \begin{cases} K^n \to (K^n)^* \\ a \mapsto \begin{pmatrix} K^n \to K \\ x \mapsto a^T x \end{pmatrix} \end{cases}$ zueinander inverse K-Vektorraumisomorphismen.

Korollar 13.1.7. Sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ und V ein endlichdimensionaler $\mathbb{K}$-Vektorraum mit Skalarprodukt. Dann ist die Abbildung $\Psi : V \to V^*, v \mapsto \begin{pmatrix} V \to \mathbb{K} \\ w \mapsto \langle v, w \rangle \end{pmatrix}$ bijektiv und im Fall $\mathbb{K} = \mathbb{R}$ auch linear.

Beweis. In **??** wurde gezeigt, dass je zwei n-dimensionale $\mathbb{K}$-Vektorräume mit Skalarprodukt als solche isomorph sind. Da die Aussage nur Bezug nimmt auf die Struktur als Vektorraum mit Skalarprodukt, können wir in der Behauptung V durch $\mathbb{K}^n$ mit Standardskalarprodukt ersetzen. Dann ist die Behauptung, dass $Psi : \mathbb{K}^n \to (\mathbb{K}^n)^*, x \mapsto \begin{pmatrix} \mathbb{K}^n \to \mathbb{K} \\ y \mapsto x^* \end{pmatrix}$ bijektiv und falls $\mathbb{K} = \mathbb{R}$ linear ist. Der Fall $\mathbb{K} = \mathbb{R}$ ergibt sich dann genau aus Korollar **??**. Sei also $\mathbb{K} = \mathbb{C}$. Dann gilt für $x := \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{C}^n$, dass $x^* = (x_1^*, ...x_n^*) = \begin{pmatrix} x_1^* \\ \vdots \\ x_n^* \end{pmatrix}^T = c(x)^T$, wobei $c : \mathbb{C}^n \to \mathbb{C}^n, \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto \begin{pmatrix} x_1^* \\ \vdots \\ x_n^* \end{pmatrix}$ bijektiv ist. Es gilt nun:

$$\begin{aligned} \Psi \text{ bijektiv} &\iff \Psi \circ c \text{ bijektiv} \\ &\iff \begin{pmatrix} \mathbb{K}^n \to (\mathbb{K}^n)^* \\ x \mapsto \begin{pmatrix} \mathbb{K}^n \to \mathbb{K} \\ y \mapsto x^T y \end{pmatrix} \end{pmatrix} \text{ bijektiv} \end{aligned}$$

Letzteres folgt aus Korollar **??** für $\mathbb{K} = \mathbb{C}$.

Proposition und Definition 13.1.8. Seien V, W je K-Vektorräume und sei $f : V \to W$ linear. Dann ist die Abbildung $f^* : W^* \to V^*, l \mapsto l \circ f$ linear und heißt die zu f *duale lineare Abbildung*.

Beweis. f^* ist wohldefiniert, da für $l \in W^*$ gilt $l \circ f \in V^*$. Denn ist $l : W \to K$ linear, so auch $l \circ f : V \to K$. Zu zeigen verbleibt die Linearität von f^*. Sei hierfür $l_1, l_2 \in W^*$ und $v \in V$. Dann gilt

$$\begin{aligned}(f^*(l_1 + l_2))(v) &= ((l_1 + l_2) \circ f)(v) = (l_1 + l_2)(f(v)) = l_1(f(v)) + l_2(f(v)) \\ &= (l_1 \circ f)(v) + (l_2 \circ f)(v) = (f^*(l_1))(v) + (f^*(l_2))(v),\end{aligned}$$

also $f^*(l_1 + l_2) = f^*(l_1) + f^*(l_2)$. Sei nun $l \in W^*$, $v \in V$ und $\lambda \in K$. Es folgt

$$(f^*(\lambda l))(v) = ((\lambda l) \circ f)(v) = (\lambda l)(f(v)) = \lambda(l(f(v))) = \lambda((f^*(l)(v)),$$

Proposition 13.1.9. Seien U, V, W je K-Vektorräume und $U \xrightarrow{g} V \xrightarrow{f} W$ linear. Dann gilt $(f \circ g)^* = g^* \circ f^*$.

Beweis. Es ist $f \circ g : U \to W$, also $(f \circ g)^* : W \to U$. Auch ist $g^* : W \to V$ und $f^* : V \to U$, also $g^* \circ f^* : W \to U$. Die Definitions und Zielbereiche stimmen folglich überein. Zum Beweis der punktweisen Gleichheit sei $l \in W^*$. Zu zeigen ist $(f \circ g)^*(l) = g^*(f^*(l))$. Wir tun dies über die Gleichungskette

$$(f \circ g)^*(l) = l \circ (f \circ g) = (l \circ f) \circ g = (f^*(l)) \circ g = g^*(f^*(l)).$$

Proposition 13.1.10. Die Abbildung $\mathrm{Hom}(V, W) \to \mathrm{Hom}(W^*, V^*), f \mapsto f^*$ ist linear.

Beweis. Sei $f, g \in \mathrm{Hom}(V, W)$ und $\lambda \in K$. Zu zeigen ist:

(a) $(f + g)^* = f^* + g^*$

(b) $(\lambda f)^* = \lambda f^*$

(a) Für alle $l \in W^*$ und $v \in V$ gilt

$$\begin{aligned}((f + g^*)(l))(v) &= (l \circ (f + g))(v) = (l \circ f + l \circ g)(v) = (l \circ f)(v) + (l \circ g)(v) \\ &= (f^*(l))(v) + (g^*(l))(v)),\end{aligned}$$

wobei wir genutzt haben, dass l linear ist.

(b) Für alle $l \in W^*$, $v \in V$ gilt

$$((\lambda f)^*(l))(v) = (l \circ (\lambda f))(v) = \lambda((l \circ f)(v)) = \lambda(f^*(v)),$$

wieder per Linearität von l.

Proposition 13.1.11. Seien V, W endlichdimensionale K-Vektorräume mit geordneten Basen $\underline{v}$ und $\underline{w}$. Sei $f : V \to W$ linear. Dann gilt

$$M(f^*, \underline{w}^*, \underline{v}^*) = M(f, \underline{v}, \underline{w})^T.$$

"Transponieren heißt Komponenten dualisieren."

Beweis. Schreibe $\underline{v} = (v_1, ..., v_n)$, $\underline{w} = (w_1, ...w_m)$, $\underline{v}^* = (v_1^*, ...v_n^*)$ und $\underline{w}^* = (w_1^*, ..., w_m^*)$. Schreibe $M(f, \underline{v}, \underline{w}) = (a_{i,j})_{(i,j)\in\{1...m\}\times\{1...n\}}$. Dann ist

$$f(v_j) = \sum_{i=1}^{m} a_{i,j} w_i \text{ für } j \in \{1...n\}.$$

Daher gilt

$$\begin{aligned}(f^*(w_i^*))(v_j) = (w_i^* \circ f)(v_j) = w^*(f(v_j)) = w_i^* \left(\sum_{i=1}^{m} a_{i,j} w_i\right) \\ = \sum_{k=1}^{m} a_{k,j} w_i^*(w_k) = a_{i,j} = \sum_{k=1}^{m} a_{i,k} v_k^*(v_j) = \left(\sum_{k=1}^{m} a_{i,k} v_k^*\right)(v_j)\end{aligned}$$

für $i \in \{1...n\}$ und $j \in \{1, ..., n\}$. Es folgt

$$f^*(w_i^*) = \sum_{k=1}^{m} a_{i,k} v_k^* \text{ für } i \in \{1...n\}.$$

Dies bedeutet $M(f^*, \underline{v}^*, \underline{w}^*) = (a_{i,j})_{(j,i)\in\{1...n\}\times\{1...m\}} = M(f, \underline{v}, \underline{w})^T$.

Proposition 13.1.12. Seien V, W eindlichdimensionale K-Vektorräume und $f : V \to W$ linear. Dann gilt:

(a) $\dim \ker f + \dim \operatorname{im} f^* = \dim V$ und $\dim \operatorname{im} f + \dim \ker f^* = \dim W$,

(b) f injektiv $\iff$ f^* surjektiv,

(c) f surjektiv $\iff$ f^* injektiv,

(d) f bijektiv $\iff$ f^* bijektiv.

Beweis. (a) Nach der Dimensionsformel für die lineare Abbildung f und f^* [→**??**] genügt es $\dim \operatorname{im} f = \dim \operatorname{im} f^*$ zu zeigen. Wähle Basen $\underline{v}$ von V und $\underline{w}$ von W. Es gilt

$$\begin{aligned}\dim \operatorname{im} f^* &\overset{??}{=} \operatorname{rank} M(f^*, \underline{v}^*, \underline{w}^*) = \dim \operatorname{im} M(f^*, \underline{v}^*, \underline{w}^*) \\ &= \dim \operatorname{row} M(f^*, \underline{v}^*, \underline{w}^*)^T = \operatorname{rank} M(f^*, \underline{v}^*, \underline{w}^*)^T \\ &\overset{??}{=} \operatorname{rank} M(f, \underline{v}, \underline{w}) = \dim \operatorname{im} f.\end{aligned}$$

(b) und (c) folgen sofort aus (a). (d) folgt aus (b) und (c).

Definition 13.1.13. Ist V ein K-Vektorraum, so heißt $V^{**} := (V^*)^*$ das *Doppeldual* (auch: Bidual) von V.

Proposition 13.1.14. Sei V ein K-Vektorraum. Die kanonische Abbildung *gnore* : $V \to V^{**}, v \mapsto \begin{pmatrix} V^* \to K \\ l \mapsto l(v) \end{pmatrix}$ ist linear und injektiv. Ist V endlichdimensional, so ist sie auch surjektiv.

Beweis. *gnore* wohldefiniert. Sei $v \in V$. Zu zeigen ist, $V^* \to K, l \mapsto l(v)$ ist linear. Dies ist klar nach Linearität der Elemente in V^*.

gnoreall linear. Seien $v_1, v_2, v \in V$ und $\lambda \in K$. Es gilt für alle $l \in V^*$

$$((v_1 + v_2))(l) = l(v_1 + v_2) = l(v_1) + l(v_2) = ((v_1))(l) + ((v_2))(l)$$

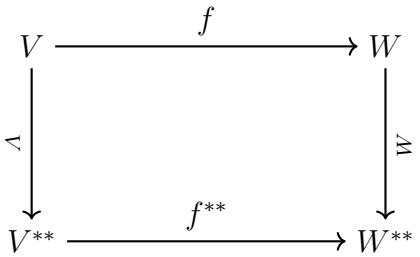
sowie

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

ε injektiv. Sei $v \in V \setminus \{0\}$. Zu zeigen ist $\varepsilon \neq 0$. Nach **??** gibt es dann eine Basis B von V mit $v \in B$. Definiere eine linear Abbildung $l : V \to K$ mit $l(b) = 1$ für alle $b \in B$. Dann insbesondere $l(v) = 1$, also $(\varepsilon(v))(l) = 1$ und daher $\varepsilon(v) \neq 0$.

ε surjektiv. Ist V endlichdimensional, so gilt wegen (**??**) $\dim V^{**} = \dim V^* = \dim V$, weswegen ε dann surjektiv ist.

Bemerkung 13.1.15. Seien V, W je K-Vektorräume und $f : V \to W$ eine lineare Abbildung. Seine $\varepsilon_V : V \to V^{**}$ und $\varepsilon_W : W \to W^{**}$ die kanonischen Abbildungen wie in **??**. Dann kommutiert das folgende Diagramm:



Also $\varepsilon_W \circ f = f^{**} \circ \varepsilon_V$. Ist nämlich $v \in V$ und $l \in W^*$, so gilt

$$\begin{aligned}((f^{**} \circ \varepsilon_V)(v))(l) &= (f^{**}(\varepsilon_V(v)))(l) = (\varepsilon_V(v) \circ f^*)(l) = (\varepsilon_V(v))(f^*(l)) \\ &= (\varepsilon_V(v))(l \circ f) = (l \circ f)(v) = l(f(v)) \\ &= (\varepsilon_W(f(v)))(l) = ((\varepsilon_W \circ f)(v))(l)\end{aligned}$$

13.2 Bilineare Abbildungen

Definition 13.2.1. Seien V_1, V_2 und W Mengen und $b : V_1 \times V_2 \to W$ eine Abbildung. Dann bezeichen wir für $v_1 \in V_1$ mit $b(v_1, \cdot)$ die Abbildung $V_2 \to W, v_2 \mapsto b(v_1, v_2)$ und für $v_2 \in V_2$ mit $b(\cdot, v_2)$ die Abbildung $V_1 \to W, v_1 \mapsto b(v_1, v_2)$.

Definition 13.2.2. Seien V_1, V_2 und W je K-Vektorräume. Eine Abbildung $b : V_1 \times V_2 \to W$ heißt bilinear *bilinear* , wenn $b(\cdot, v_2)$ für alle $v_2 \in V_2$ und $b(v_1, \cdot)$ für alle $v_1 \in V_1$ linear sind.

Beispiel 13.2.3. (a) Seien U, V, W je K-Vektorräume. In **??** wurde gezeigt, dass die Hintereinanderschaltung von linearer Abbildungen

$$\begin{aligned} \mathrm{Hom}(V, W) \times \mathrm{Hom}(U, V) &\to \mathrm{Hom}(U, W) \\ (g, f) &\mapsto g \circ f, \end{aligned}$$

eine bilineare Abbildung ist.

(b) Seien $m, n, r \in N_0$. In **??**(b) wurde gezeigt, dass die Matrizenmultiplikation,

$$\begin{aligned} K^{m\times n} \times K^{n\times r} &\to K^{m\times r} \\ (A, B) &\mapsto AB, \end{aligned}$$

bilinear ist.

Bemerkung 13.2.4. Seien V_1, V_2 und W je K-Vektorräume.

(a) Es ist $\{b \mid b : V_1 \times V_2 \to W \text{ bilinear}\}$ ein Unterraum des K-Vektorraums $W^{V_1\times V_2}$.

(b) ist $f : V_1 \times V_2 \to W$ gleichzeitig linear und bilinear, so gilt $f = 0$. Denn ist $(v_1, v_2) \in V_1 \times V_2$, so gilt

$$f((v_1, v_2)) = f((v_1, 0)) + f((0, v_2)) = 0 + 0 = 0.$$

(c) Ist $b : V_1 \times V_2 \to W$ bilinear, so auch $V_2 \times V_1 \to W, (v_2, v_1) \mapsto b(v_1, v_2)$. Daher kann man z.B. im nächsten Satz die Rolle von erstem und zweitem Argument vertauschen.

(d) Das Studium bilinearer Abbildungen kann teilweise auf das Studium linearer Abbildung zurückgeführt werden. Siehe nächster Satz.

Lemma 13.2.5. Seien V_1, V_2 und W Mengen. Dann sind die Abbildungen

$$\Phi : \begin{cases} W^{V_1\times V_2} \to (W^{V_2})^{V_1} \\ b \mapsto \begin{pmatrix} V_1 \to W^{V_2} \\ v_1 \mapsto b(v_1, \cdot) \end{pmatrix} \end{cases} \quad \text{und} \quad \Psi : \begin{cases} (W^{V_2})^{V_1} \to W^{V_1\times V_2} \\ f \mapsto \begin{pmatrix} V_1 \times V_2 \to W \\ (v_1, v_2) \mapsto (f(v_1))(v_2) \end{pmatrix} \end{cases}$$

zueinander invers und damit insbesondere bijektiv. Ist W ein K-Vektorraum, so sind Φ und Ψ sogar K-Vektorraumisomorphismen.

Beweis. Übung.

Satz 13.2.6. Seien V_1, V_2 und W je K-Vektorräume. Dann sind die Abbildungen

$$\Phi : \begin{cases} \{b \mid b : V_1 \times V_2 \to W \text{ bilinear}\} \to \mathrm{Hom}(V_1, \mathrm{Hom}(V_2, W)) \\ b \mapsto \begin{pmatrix} V_1 \to \mathrm{Hom}(V_2, W) \\ v_1 \mapsto b(v_1, \cdot) \end{pmatrix} \end{cases} \quad \text{und}$$

$$\Psi : \begin{cases} \mathrm{Hom}(V_1, \mathrm{Hom}(V_2, W)) \to \{b \mid b : V_1 \times V_2 \to W \text{ bilinear}\} \\ f \mapsto \begin{pmatrix} V_1 \times V_2 \to W \\ (v_1, v_2) \mapsto (f(v_1))(v_2) \end{pmatrix} \end{cases} \quad \text{zueinander inverse } K\text{-Vektorraumisomorphismen.}$$

Beweis. Es reicht zu zeigen, dass für Φ definiert wie in Lemma **??** und jede Abbildung $b : V_1 \times V_2 \to W$ gilt b bilinear $\iff \Phi(b) \in \mathrm{Hom}(V_1, \mathrm{Hom}(V_2, W))$.

$\Longrightarrow$ Sei b bilinear. Dann ist ist $(\Phi(b))(v_1) = b(v_1, \cdot)$ linear für also $v_1 \in V$, also $\Phi(b) : V_1 \to \mathrm{Hom}(V_2, W)$ eine Abbildung. Zu zeigen ist, $\Phi(b)$ ist linear. Seien hierzu $v_1, v_1' \in V_1$ und $\lambda \in K$. Nach Definition von Φ ist zu zeigen $b(v_1 + v_1', \cdot) = b(v_1, \cdot) + b(v_1', \cdot)$ und $b(\lambda v_1, \cdot) = \lambda b(v_1, \cdot)$. Sei $v_2 \in V_2$. Dann gilt

$$\begin{aligned}(b(v_1 + v_1', \cdot))(v_2) &= b(v_1 + v_1', v_2) = b(v_1, v_2) + b(v_1', v_2) \\ &= (b(v_1, \cdot))(v_2) + (b(v_1', \cdot))(v_2)\end{aligned}$$

und

$$(b(\lambda v_1, \cdot))(v_2) = b(\lambda v_1, v_2) = \lambda b(v_1, v_2) = (\lambda b(v_1, \cdot))(v_2).$$

$\Longleftarrow$ Sei $\Phi(b) : V_1 \to \mathrm{Hom}(V_2, W)$ linear. Zu zeigen ist $\forall v_1 \in V_1 : b(v_1, \cdot)$ linear und $\forall v_2 \in V_2 : b(\cdot, v_2)$ linear. Ersteres ist klar, denn $\forall v_1 \in V_2 : (\Phi(b))(v_1) = b(v_1, \cdot))$. Für Zweiteres sei $v_2 \in V_2$. Dann ist $: \mathrm{Hom}(V_2, W) \to W, g \mapsto g(v_2)$ linear und $b(\cdot, v_2) = \circ\Phi(b)$ ebenso.

13.3 Bilinearformen

Definition 13.3.1. Sei V ein K-Vektorraum. Eine *bilinear* auf V ist eine bilineare Funktion $f : V \times V \to K$. Wir bezeichnen den K-Vektorraum aller Bilinearformen auf V mit $\mathrm{Bil}(V)$.

Beispiel 13.3.2. (a) Ist $n \in N_0$ und $A = (a_{i,j})_{1 \le i,j \le n} \in K^{n \times n}$, so ist $b_A : K^n \times K^n \to K, (x, y) \mapsto x^T A y = \sum_{i,j=1}^n a_{i,j} x_i y_j$ bilinear.

(b) Jedes Skalarprodukt auf einem reellen Vektorraum V ist eine Bilinearform auf V.

Definition 13.3.3. [→??] Sei V ein K-Vektorraum mit Basis $\underline{v} = (v_1, ..., v_n)$ und b eine Bilinearform auf V. Eine Matrix $A \in K^{n\times n}$ heißt *Darstellungsmatrix* von b bezüglich der Basis $\underline{v}$, falls für alle $v, w \in V$ gilt

$$(*)\ b(v,w) = b_A(\mathrm{coord}_{\underline{v}}(v), \mathrm{coord}_{\underline{v}}(w))$$

Bemerkung 13.3.4. Man beachte

$$\begin{aligned}(*) &\overset{\underline{v}\text{ Basis}}{\iff} \forall i,j \in \{1...n\} : b(v_i, v_j) = b_A(\mathrm{coord}_{\underline{v}}(v_i), \mathrm{coord}_{\underline{v_j}}(w)) \\ &\iff \forall i,j \in \{1...n\} : b(v_i, v_j) = b_A(e_i, e_j) \\ &\iff \forall i,j \in \{1...n\} : b(v_i, v_j) = e_i^T A e_j \\ &\iff A = (b(v_i, v_j))_{1\le i,j\le n}.\end{aligned}$$

Zu jeder Bilinearform auf einem endlichdimensionalen Vektorraum gibt es also eine bezüglich einer gegeben Basis jeweils genau eine Darstellungsmatrix.

Notation 13.3.5. $M(b, \underline{v})$ steht für das eindeutig bestimmte $A \in K^{n\times n}$ aus Definition **??**.

Beispiel 13.3.6. Sei $d \in \mathbb{N}_0$. Betrachte den $\mathbb{R}$-Vektorraum $\mathbb{R}[X]_d$ aller reellen Polynome vom Grad $\le d$ mit dem Skalarprodukt

$$b : \mathbb{R}[X]_d \times \mathbb{R}[X]_d \to \mathbb{R}, (f,g) \mapsto \int_0^1 f(x)g(x)\,\mathrm{d}x$$

und der Basis $\underline{v} = (1, X, ...X^d)$. Dann ist

$$M(b, \underline{v}) = \begin{pmatrix} 1 & 1/2 & 1/3 & ... & 1/(d+1) \\ 1/2 & 1/3 & 1/4 & ... & 1(d+2) \\ 1/3 & 1/4 & 1/5 & ... & 1/(d+3) \\ ... & ... & ... & ... & ... \\ 1/(d+1) & 1/3 & 1/4 & ... & 1(d+d) \end{pmatrix} \in K^{(d+1)\times(d+1)}.$$

Definition 13.3.7. Sei V ein K-Vektorraum. Dann definieren wir die beiden Vektorraumisomorphismen [→??] $\leftarrow: \begin{cases} \mathrm{Bil}(V) & \to \mathrm{Hom}(V, V^*) \\ b & \mapsto \overleftarrow{b} := \begin{pmatrix} V \to V^* \\ v \mapsto b(\cdot, v) \end{pmatrix} \end{cases}$ und

$\rightarrow: \begin{cases} \mathrm{Bil}(V) & \to \mathrm{Hom}(V, V^*) \\ b & \mapsto \overrightarrow{b} := \begin{pmatrix} V \to V^* \\ v \mapsto b(v, \cdot) \end{pmatrix} \end{cases}$.

Proposition 13.3.8. Sei V ein K-Vektorraum mit geordneter Basis $\underline{v}$ und $b \in \mathrm{Bil}(V)$. Dann gilt $M(b, \underline{v}) = M(\overleftarrow{b}, \underline{v}, \underline{v}^*) = M(\overrightarrow{b}, \underline{v}, \underline{v}^*)^T$.

Beweis. Zu zeigen ist:

(a) $\overset{\leftarrow}{b} = \mathrm{vec}_{\underline{v}^*} \circ f_{M(b,\underline{v})} \circ \mathrm{coord}_{\underline{v}}$

(b) $\vec{b} = \mathrm{vec}_{\underline{v}^*} \circ f_{M(b,\underline{v})^T} \circ \mathrm{coord}_{\underline{v}}$

Wir zeigen nur (a), denn (b) geht fast genauso. Es gilt

$$(a) \overset{\underline{v}\text{Basis}}{\iff} \forall j \in \{1,...,n\} : \overset{\leftarrow}{b}(v_j) = \mathrm{vec}_{\underline{v}^*}(M(b,\underline{v}) \underbrace{\mathrm{coord}_{\underline{v}}(v_j)}_{=e_j})$$

$$\iff \forall j \in \{1,...,n\} : b(\cdot, v_j) = \sum_{k=1}^{n} \left(e_k^T M(b,\underline{v}) e_j\right) v_k^*$$

$$\overset{\underline{v}\text{Basis}}{\iff} \forall i,j \in \{1,...,n\} : b(v_i,v_j) = e_i^T M(b,\underline{v}) e_j = b(v_i,v_j).$$

Satz 13.3.9. Sei V ein K-Vektorraum mit Basis $\underline{v} := (v_1, ...v_n)$. Dann sind

$$\Phi : \begin{cases} \mathrm{Bil}(V) \to \mathbb{K}^{n\times n} \\ b \mapsto M(b,\underline{v}) \end{cases} \quad \text{und} \quad \Psi : \begin{cases} K^{n\times n} \to \mathrm{Bil}(V) \\ A \mapsto \begin{pmatrix} V \times V \to K \\ (v,w) \mapsto \mathrm{coord}(v)_{\underline{v}}^T A\, \mathrm{coord}_{\underline{v}}(w) \end{pmatrix} \end{cases}$$ zueinander inverse K-Vektorraumisomorphismen.

Beweis. Φ ist als Hintereinanderschaltung zweier Vektorraumisomorphismen ebenfalls ein Vektorraumisomorphismen:

$$\mathrm{Bil}(V) \underset{[\to ??]}{\overset{\cong}{\longrightarrow}} \mathrm{Hom}(V,V^*) \underset{[\to ??]}{\overset{\cong}{\longrightarrow}} K^{n\times n}$$
$$b \mapsto \overset{\leftarrow}{b} \mapsto M(\overset{\leftarrow}{b}, \underline{v}, \underline{v}^*) \overset{??}{=} M(b,\underline{v}).$$

Ist $b \in \mathrm{Bil}(V)$, so gilt für alle $i,j \in \{1,...,n\}$

$$(\Psi(\Phi(b)))(v_i,v_j) = e_i^T \Phi(b) e_j = e_i^T M(b,\underline{v}) e_i = b(v_i,v_j).$$

Daraus folgt für $A \in K^{n\times n}$

$$\Psi(A) = \Psi(\Phi(\Phi^{-1}(A))) = \Phi^{-1}(A).$$

Daher ist $\Phi = \Psi^{-1}$ auch ein Vektorraumisomorphismus.

Satz 13.3.10. Sei V ein K-Vektorraum mit geordneten Basen $\underline{v}$, $\underline{w}$ und $b \in \mathrm{Bil}(V)$. Dann gilt $M(b,\underline{v}) = M(\underline{v},\underline{w})^T M(b,\underline{w}) M(\underline{v},\underline{w})$.

Beweis. Es gilt

$$\begin{aligned} M(b,\underline{v}) &\overset{??}{=} M(\overset{\leftarrow}{b},\underline{v},\underline{v}^*) \overset{??}{=} M(\underline{w}^*,v^*)M(\overset{\leftarrow}{b},\underline{w},\underline{w}^*)M(\underline{v},\underline{w}) \\ &\overset{??}{=} M(\underline{v},\underline{w})^T M(\overset{\leftarrow}{b},\underline{w},\underline{w}^*)M(\underline{v},\underline{w}) \\ &\overset{13.3.8}{=} M(\underline{v},\underline{w})^T M(b,\underline{w})M(\underline{v},\underline{w}) \end{aligned}$$

Definition und Proposition 13.3.11. Sei V ein K-Vektorraum mit Basis $\underline{v} = (v_1, ..., v_n)$. Sei $b \in \mathrm{Bil}(V)$. Man nennt b *nicht ausgeartet*, wenn folgende äquivalente Bedingungen gelten:

(a) $\overset{\leftarrow}{b}$ ist injektiv, d.h. $\forall w \in V : b(\cdot, w) = 0 \implies w = 0$),

(b) $\overset{\leftarrow}{b}$ ist surjektiv, d.h. $\forall \ell \in V^* : \exists w \in V : \ell = b(\cdot, w)$,

(c) $\overset{\leftarrow}{b}$ ist ein Vektorraumisomorphismus,

(d) $M(b, \underline{v})$ ist invertierbar,

(e) $\vec{b}$ ist injektiv, d.h. $\forall v \in V : b(v, \cdot) = 0 \implies v = 0$),

(f) $\vec{b}$ ist surjektiv, d.h. $\forall \ell \in V^* : \exists v \in V : \ell = b(v, \cdot)$,

(g) $\vec{b}$ ist ein Vektorraumisomorphismus.

Beweis. $(a) \iff (b) \iff (c) \iff (d)$ ist klar, da $M(b, \underline{v}) = M(\overset{\leftarrow}{b}, \underline{v}, \underline{v}^*)$. Ebenso ist $(d) \iff (e) \iff (f) \iff (g)$, da $M^T(b, \underline{v}) = M(\vec{b}, \underline{v}, \underline{v}^*)$ und $M(b, \underline{v})$ invertierbar $\iff M^T(b, \underline{v})$ invertierbar.

13.4 Symmetrische Bilinearformen und quadratische Formen

Definition 13.4.1. Eine Bilinearform b auf einen K-Vektorraum V heißt symmetrisch *symmetrisch*, wenn $b(v, w) = b(w, v)$ für alle $v, w \in V$ gilt. Wir bezeichnen den K-Vektorraum aller symmetrischen Bilinearformen auf einen K-Vektorraum V mit $\mathrm{SBil}(V)$.

Definition 13.4.2. [→??] Eine Matrix $A \in K^{n\times n}$ heißt *symmetrisch*, wenn $A = A^T$. Wir bezeichnen den K-Vektorraum aller symmetrischen $n \times n$-Matrizen über K mit $\mathrm{SK}^{n\times n}$.

Proposition 13.4.3. Sei V ein K-Vektorraum mit Basis $\underline{v} = (v_1, ..., v_n)$. Sei $b \in \mathrm{Bil}(V)$. Dann gilt $b \in \mathrm{SBil}(V) \iff M(b, \underline{v}) \in \mathrm{SK}^{n\times n}$.

Beweis. Definiere $\widetilde{b}\colon V \times V \to K, (v,w) \mapsto b(w,v)$. Es gilt

$$\begin{aligned} b \in \mathrm{SBil}(V) &\iff b = \widetilde{b} \iff \forall i,j \in \{1...n\} : b(v_i,v_j) = \widetilde{b}\ (v_j,v_i) \\ &\iff (b(v_i,v_j))_{i,j\in\{1...n\}} = (\widetilde{b}\ (v_j,v_i))_{i,j\in\{1...n\}} \\ &\iff M(b,\underline{v}) = M(b,\underline{v})^T \iff M(b,\underline{v}) \in \mathrm{SK}^{n\times n}. \end{aligned}$$

Beispiel 13.4.4. Jedes Skalarprodukt auf einem reellen Vektorraum V ist eine symmetrische Biliniearform auf V.

Definition 13.4.5. Sei V ein K-Vektorraum. Eine Funktion $q : V \to K$ heißt *quadratische Form* auf V, wenn es $b \in \mathrm{Bil}(V)$ gibt mit $q(v) = b(v,v)$ für alle $v \in V$. Wir bezeichnen den K-Vektorraum aller quadratischen Formen auf einem K-Vektorraum V mit $\mathrm{Q}(V)$.

Notation und Proposition 13.4.6. Gelte $1+1 \neq 0$ in K. Dann sind die Abbildungen $\mathrm{SBil}(V) \to \mathrm{Q}(V), b \mapsto q_b$ mit $q_b : \begin{cases} V \to K \\ v \mapsto b(v,v) \end{cases}$ und $\mathrm{Q}(V) \to \mathrm{SBil}(V), q \mapsto b_q$ mit $b_q : \begin{cases} V \times V \to K \\ (v,w) \mapsto \frac{1}{2}(q(v+w) - q(v) - q(w)) \end{cases}$ zueinander inverse K-Vektorraumisomorphismen.

Beweis. Sei $q \in \mathrm{Q}(V)$. Zu zeigen ist $b_q \in \mathrm{SBil}(V)$. Wähle $b \in \mathrm{SBil}(V)$ mit $q(v) = b(v,v)$ für alle $v \in V$. Dann ist

$$\begin{aligned} b_q(v,w) &= \frac{1}{2}(b(v+w,v+w) - b(v,v) - b(w,w)) \\ &= \frac{1}{2}(b(v,v) + b(w,w) + b(v,w) + b(w,v) - b(v,v) - b(w,w)) \\ &= \frac{1}{2}(b(v,w) + b(v,w)) \\ &= b(v,w). \end{aligned}$$

Genauso zeigt man, dass $b_{q_b} = b$ für alle $b \in \mathrm{SBil}(V)$. Sei $q \in \mathrm{Q}(V)$, etwa $q(v) = b(v,v)$ für alle $v \in V$ für ein $b \in \mathrm{SBil}(V)$. Dann gilt $q_{b_q}(v) = b_q(v,v) = b(v,v) = q(v)$. Daher sind die zwei Abbildungen invers zueinander. Die Linearität ist klar.

Proposition 13.4.7. Die folgende Proposition motiviert die Bezeichnung "quadratische Form". Sei V ein K-Vektorraum mit Basis $\underline{v} = (v_1, ..., v_n)$ und $q : V \to K$ eine Funktion. Dann ist $q \in \mathrm{Q}(V)$ genau dann, wenn es eindeutig bestimmte $a_{i,j} \in K\ (1 \leq i \leq j \leq n)$ gibt mit $\forall x \in K^n : q\left(\sum_{i=1}^n x_i v_i\right) = \sum_{i\leq j} a_{i,j} x_i x_j$.

Beweis. $\Longrightarrow$ Sei $q \in \mathrm{Q}(V)$. Wähle $b \in \mathrm{Bil}(V)$ mit $\forall v \in V : q(v) = b(v,v)$. Dann gilt für $a_{i,j} \in K\,(1 \le i \le j \le n)$:

$$\begin{aligned}
&\forall x \in K^n : b\left(\sum_{i=1}^n x_i v_i, q\sum_{i=1}^n x_i v_i\right) = \sum_{i\le j}^n a_{i,j}x_i x_j \\
\Longleftrightarrow &\forall x \in K^n : \sum_{i,j=1}^n b(v_i,v_j)x_i x_j = \sum_{i\le j}^n a_{i,j}x_i x_j \\
\Longleftrightarrow &\forall x \in K^n (i \in \{1...n\} : a_{i,i} = b(v_i,v_i) \\
&\quad \wedge \forall i,j \in \{1...n\} : (i<j \implies a_{i,j} = b(v_i,v_j) + b(v_j,v_i)).
\end{aligned}$$

Daher gibt es eindeutig bestimmte $a_{i,j} \in K\,(1 \le i \le j \le n)$ mit $\forall x \in K^n : q\left(\sum_{i=1}^n x_i v_i\right) = \sum_{i\le j} a_{i,j}x_i x_j$.

$\Longleftarrow$ Gebe es nun umgekehrt $a_{i,j} \in K\,(1 \le i \le j \le n)$ mit $\forall x \in K^n : q\left(\sum_{i=1}^n x_i v_i\right) = \sum_{i\le j} a_{i,j}x_i x_j$. Definiere dann $b \in \mathrm{Bil}(V)$ durch

$$b(v_i,v_j) := \begin{cases} a_{i,j} & \text{falls } i \le j \\ 0 & \text{sonst} \end{cases}.$$

Dann gilt

$$q\left(\sum_{i=1}^n x_i v_i\right) = \sum_{i\le j} a_{i,j}x_i x_j = \sum_{i\le j} x_i x_j b(v_i,v_j) = b\left(\sum_{i=1}^n x_i v_i, \sum_{i=1}^n x_i v_i\right).$$

Bemerkung 13.4.8. Im Fall $2 \neq 0$ in K kann man in der Definition **??** "wenn es $b \in \mathrm{Bil}(V)$ gibt" ersetzen durch "wenn es (ein eindeutig bestimmtes) $b \in \mathrm{SBil}(V)$ gibt". Dieses $b \in \mathrm{SBil}(V)$ ist natürlich b_q und man identifiziert dann oft q mit b_q. Z.B. kann man von einer "Darstellungsmatrix von q" reden und damit eine Darstellungsmatrix von b_q meinen.

Im Fall $2 = 0$ in K, der uns hier nicht weiter interessiert, ist all dies nicht möglich, wie die folgenden beiden Beispiele zeigen.

Beispiel 13.4.9. (a) $q : \mathbb{F}_2^2 \to \mathbb{F}_2, \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto x_1 x_2$ ist eine quadratische Form auf $\mathbb{F}_2^2$, aber es gibt kein $b \in \mathrm{SBil}(V)$ mit $q(x) = b(x,x)$ für alle $x \in \mathbb{F}_2^2$. Denn sonst würde gelten

$$\begin{aligned}
q(x) = b(x,x) &= b(x_1e_1 + x_2e_2, x_1e_1 + x_2e_2) = x^2 b(e_1,e_1) + 2x_1x_2b(e_1,e_2) + x_2^2 b(e_2,e_2) \\
&= x^2 b(e_1,e_1) + x_2^2 b(e_2,e_2)
\end{aligned}$$

für alle $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{F}_2^2$ und daher $1 = q\left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right) = b(e_1,e_1) + b(e_2,e_2) = q(e_1) +$

$q(e_2) = 0$.

(b) Für die Nullform $0 \in \mathrm{Q}(\mathbb{F}_2^2)$ gibt es außer $0 \in \mathrm{SBil}(\mathbb{F}_2^2)$ noch eine andere symmetrische Bilinearform $b : \mathbb{F}_2^2 \to \mathbb{F}_2$ mit $q_b = 0$, nämlich die durch $M(b, \underline{e}) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ gegebene.

(c) $q : \mathbb{R}^3 \to \mathbb{R}, \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mapsto 2x_1^2 - 3x_1x_2 + x_2^2 - x_1x_3 - 2x_3^2$ ist eine quadratische Form auf $\mathbb{R}^3$ mit zugehöriger Bilinearform

$$b : \begin{cases} \mathbb{R}^3 \times \mathbb{R}^3 & \to \mathbb{R} \\ \left(\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \right) & \mapsto 2x_1y_1 - \frac{3}{2}x_1y_2 - \frac{3}{2}x_2y_1 + x_2y_2 - \frac{1}{2}x_1y_3 - \frac{1}{2}x_3y_1 - 2x_3y_3 \end{cases}.$$

Es gilt

$$M(q, \underline{e}) = \begin{pmatrix} 2 & -3/2 & -1/2 \\ -3/2 & 1 & 0 \\ -1/2 & 0 & -2 \end{pmatrix}.$$

13.5 Eine verallgemeinerte Cholesky-Zerlegung

Lemma 13.5.1. Gelte $2 \neq 0$ in K. Sei V ein K-Vektorraum mit Basis $\underline{v} = (v_1, ..., v_n)$, $b \in \mathrm{SBil}(V)$ und $q := q_b \in \mathrm{Q}(V)$. Seien $l_1, ..., l_m \in V^*$ und $\lambda_1, ..., \lambda_m \in K$. Dann sind äquivalent:

(a) $q(v) = \sum_{k=1}^m \lambda_k l_k^2(v)$ für alle $v \in V$

(b) $b(v, w) = \sum_{k=1}^m \lambda_k l_k(v) l_k(w)$ für alle $v, w \in V$

(c) $M(b, \underline{v}) = P^T D P$ mit $D := \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_m \end{pmatrix} \in K^{m \times m}$ und $P := (l_i(v_j))_{1 \leq i \leq m, 1 \leq j \leq n} \in K^{m \times n}$

Beweis. $(a) \impliedby (b)$. Ist klar.

$(a) \implies (b)$. Gelte (a). Es ist $b_0 : V \times V \to K, (v, w) \mapsto \sum_{k=1}^m \lambda_k l_k(v) l_k(w)$ eine symmetrische Bilinearform, für die gilt $q_{b_0} = q = q_b$ und daher $b_0 = b$.

$(b) \iff (c)$. Es gilt:

- $(b) \iff \forall i, j \in \{1, ..., n\} : b(v_i, v_j) = \sum_{k=1}^m \lambda_k l_k(v_i) l_k(v_j)$

- $(c) \iff \forall i,j \in \{1,...,m\}. e_i^T M(b,\underline{v}) e_j = (Pe_i)^T D P e_j$

Seien nun $i,j \in \{1,...,n\}$. Wegen $b(v_i,v_j) = e_i^T M(b,\underline{v}) e_j$ reicht es zu zeigen, dass $\sum_{k=1}^m \lambda_k l_k(v_i) l_k(v_j) = (Pe_i)^T D P e_j$. Sei hierzu $k \in \{1,...,m\}$. Wir zeigen

$$l_k(v_i) l_k(v_j) = \left(\begin{pmatrix} 0 & & & & \\ & \ddots & & & \\ & & 0 & & \\ & & & 1 & \\ & & & & \ddots \\ & & & & & 0 \end{pmatrix} Pe_i \right)^T \quad \left(\begin{pmatrix} 0 & & & & \\ & \ddots & & & \\ & & 0 & & \\ & & & 1 & \\ & & & & \ddots \\ & & & & & 0 \end{pmatrix} Pe_j \right)$$

$$= \begin{pmatrix} 0 \\ \vdots \\ l(e_i) \\ \vdots \\ 0 \end{pmatrix}^T \qquad = \begin{pmatrix} 0 \\ \vdots \\ l(e_j) \\ \vdots \\ 0 \end{pmatrix} \longleftarrow \textbf{k-te Stelle}$$

k-te Stelle

Multipliziert man auf beiden Seiten mit λ_k und summiert man auf, so erhält man

$$\sum_{k=1}^m \lambda_k l_k(v_i) l_k(v_j) = \sum_{k=1}^m (Pe_i)^T \begin{pmatrix} 0 & & & & \\ & \ddots & & & \\ & & 0 & & \\ & & & 1 & \\ & & & & \ddots \\ & & & & & 0 \end{pmatrix}^2 Pe_j$$

$$= (Pe_i)^T \left(\sum_{k=1}^m \begin{pmatrix} 0 & & & & \\ & \ddots & & & \\ & & 0 & & \\ & & & 1 & \\ & & & & \ddots \\ & & & & & 0 \end{pmatrix} \right) Pe_j$$

$$= (Pe_i)^T D P e_j.$$

k-te Stelle

Satz 13.5.2. Gelte $2 \neq 0$ in K. Sei V ein K-Vektorraum mit Basis $\underline{v} = (v_1,...,v_n)$ und $q \in \mathrm{Q}(V)$. Seien $l_1,...,l_n \in V^*$ und $\lambda_1,...,\lambda_n \in K^\times$. Dann sind äquivalent:

(a) $l_1,...l_n$ sind linear unabhängig in V^* und $q(v) = \sum_{k=1}^n \lambda_k l_k^2(v)$ für alle $v \in V$.

(b) $P := (l_i(v_j))_{1 \leq i,j \leq n} \in K^{n \times n}$ ist invertierbar und $M(q,\underline{v}) = P^T D P$ mit $D := \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \in K^{n \times n}$.

Beweis. Nach Lemma **??** reicht es zu zeigen $l_1,...,l_n$ lin. unabhängig in $V^* \iff$

P invertierbar. Da $\underline{v}$ eine Basis von V ist, ist $V^* \to K^n, l \mapsto \begin{pmatrix} l(v_1) \\ \vdots \\ l(v_n) \end{pmatrix}$ ein K-Vektorraum Isomorhismus. Daher gilt

$$l_1, ..., l_n \text{ lin. unabhängig in } V^* \iff \begin{pmatrix} l_1(v_1) \\ \vdots \\ l_1(v_n) \end{pmatrix}, ..., \begin{pmatrix} l_n(v_1) \\ \vdots \\ l_n(v_n) \end{pmatrix} \text{ lin. unabhängig in } K$$
$$\iff \operatorname{rank} P = n \iff P \text{ invertierbar.}$$

Beispiel 13.5.3. Gegeben ist die Matrix $A := \begin{pmatrix} 2 & -2 & 0 \\ -2 & -1 & -4 \\ 0 & -4 & 0 \end{pmatrix} \in \mathrm{SQ}^{3\times 3}$. Gesucht ist eine invertierbare Matrix $P \in \mathrm{SQ}^{3\times 3}$ und eine Diagonalmatrix $D \in \mathbb{Q}^{3\times 3}$ mit $A = P^T D P$. Betrachte die quadratische Form q auf $\mathbb{Q}^3$ mit $M(q, \underline{e}) = A$. Es gilt für $x, y, z \in \mathbb{Q}$:

$$q(x,y,z) = 2x^2 - 4xy - y^2 - 8yz = (\underbrace{2}_{\lambda_1} \underbrace{(x-y)^2}_{l_1(x,y,z)} \underbrace{-2(-y)^2) - y^2 - 8yz}_{q_1(y,z)}$$
$$q_1(y,z) = -3y^3 - 8zy = \underbrace{-3}_{\lambda_2} \underbrace{\left(y - \frac{4}{3}z\right)^2}_{l_2(y,z)} \underbrace{+\frac{16}{9}}_{\lambda_3} \underbrace{z^2}_{l_3(z)}.$$

Setzt man $P := \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 4/3 \\ 0 & 0 & 1 \end{pmatrix}$ und $D := \begin{pmatrix} 2 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 16/3 \end{pmatrix}$, so gilt nach Satz **??** $A = P^T D P$.

Beispiel 13.5.4. Gegeben ist die Matrix $A := \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \in \mathrm{SQ}^{4\times 4}$. Gesucht ist eine invertierbare Matrix $P \in \mathbb{Q}^{4\times 4}$ und eine Diagonalmatrix $D \in \mathbb{Q}^{4\times 4}$ mit $A = P^T D P$.

Betrachte $q \in \mathrm{Q}(\mathbb{Q}^4)$ mit $M(q, \underline{e}) = A$. Für $x \in \mathbb{Q}^4$ gilt

$$\begin{aligned} q(x) &= 2x_1x_2 + 2x_1x_3 + 2x_2x_3 + 2x_3x_4 \\ &= (2\underbrace{(x_1 + x_3)}_{h_1(x_1,x_3)}\underbrace{(x_2 + x_3)}_{h_2(x_2,x_3)}\underbrace{-2x_3^3}_{q_1(x_3,x_4)}) + 2x_3x_4 \\ &= \underbrace{\frac{1}{2}}_{\lambda_1=-\lambda_2}\left(\underbrace{(h_1 + h_2)^2}_{l_1(x_1,x_2,x_3)} - \underbrace{(h_1 - h_2)^2}_{l_2(x_1,x_2,x_3)}\right) + q_1 \end{aligned}$$

$$q_1(x_3, x_4) = -2x_3^2 + 2x_3x_4 = \underbrace{-2}_{\lambda_3}\left(\underbrace{x_3 - \frac{1}{2}x_4}_{l_3(x_3,x_4}\right)^2 \underbrace{-\frac{1}{2}}_{\lambda_4}\underbrace{x_4^2}_{l_4(x_4)}.$$

Wir rechnen noch l_1, l_2 aus: $l_1 = h_1 + h_2 = x_1 + x_2 + 2x_3$ und $l_2 = h_1 - h_2 = x_1 - x_2$. Setzt man $P := \begin{pmatrix} 1 & 1 & 2 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1/2 \\ 0 & 0 & 0 & 1 \end{pmatrix}$ und $D := \begin{pmatrix} 1/2 & 0 & 0 & 0 \\ 0 & -1/2 & 0 & 0 \\ 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & -1/2 \end{pmatrix}$, so gilt $A = P^T DP$. Man beachte hierbei, dass die Matrix $P = (l_i(e_j))_{1 \le i,h \le 4}$ invertierbar ist, da l_1, l_2, l_3, l_4 eine Basis von $(\mathbb{Q}^4)^*$ bilden, denn $\mathrm{span}(l_1, l_2, l_3, l_4) = \mathrm{span}(h_1, h_2, l_3, l_4) = (\mathbb{Q}^4)^*$.

Definition 13.5.5. Bezeichne $\underline{e} = (e_1, ..., e_n)$ wie immer die Standardbasis des K^n. Eine Permutationsmatrix *Permutationsmatrix* ist eine Matrix $S \in K^{n \times n}$ der Form $S = (e_{\sigma(1)}, ..., e_{\sigma(n)})$ für eine Permutation $\sigma \in S_n$.

Bemerkung 13.5.6. (a) Sei V ein Vektorraum mit Basis $\underline{v} = (v_1, ..., v_n)$ und $\sigma \in S_n$. Dann ist $\underline{v_\sigma} := (v_{\sigma(1)}, ..., v_{\sigma(n)})$ auch eine Basis von V und die Basiswechselmatrizen $M(\underline{v}, \underline{v_\sigma}) = (e_{\sigma^{-1}(1)}, ..., e_{\sigma^{-1}(n)})$ und $M(\underline{v_\sigma}, \underline{v}) = (e_{\sigma(1)}, ..., e_{\sigma(n)})$ Permutationsmatrizen.

(b) Ist $\sigma \in S_n$, so gilt $\begin{pmatrix} e_{\sigma(1)}^T \\ \vdots \\ e_{\sigma(n)}^T \end{pmatrix}(e_{\sigma(1)}, ..., e_{\sigma(n)}) = I_n$. Für jede Permutationsmatrix S gilt daher $S^{-1} = S^T$.

Definition 13.5.7. Sei $A \in \mathrm{SK}^{n \times n}$. Unter einer *verallgemeinerten Cholesky-Zerlegung* von A verstehen wir ein Tripel (S, P, D) von Matrizen $S, P, D \in K^{n \times n}$ mit $A = S^T P^T DPS$, wobei S eine Permtuationsmatrix und P wie D von der Gestalt

$$P = \begin{pmatrix} \boxed{B_1} & & * \\ & \boxed{B_2} & \\ 0 & & \boxed{B_k} \end{pmatrix} \text{ und } P = \begin{pmatrix} \boxed{C_1} & & 0 \\ & \boxed{C_2} & \\ 0 & & \boxed{C_k} \end{pmatrix}$$

sind, wobei jedes $i \in \{1...k\}$ das Paar (B_i, C_i)

- <u>entweder</u> ein Paar von 2×2-Matrizen ist, nämlich $B_i := \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ und $C_i = \begin{pmatrix} \lambda & 0 \\ 0 & -\lambda \end{pmatrix}$ für ein $\lambda \in K^\times$,
- <u>oder</u> ein Paar von $1\times$-Matrizen ist, nämlich $B_i = (1)$ und $C_i = (\lambda)$ für ein $\lambda \in K$.

Bemerkung 13.5.8. Ist (S, P, D) eine verallgemeinerte Cholesky-Zerlegung einer Matrix $A \in \mathrm{SK}^{n\times n}$, so gilt

$$\begin{aligned}\det(A) &= \det(S^T P^T D P S) = \det(S^T)\det(P^T)\det(D)\det(S)\det(P) \\ &= \underbrace{(\det(S))}_{=\pm 1}{}^2 \det(P)^2 \det(D)\end{aligned}$$

und $\det(P) = \prod_{k\in\{1...n\}} B_i = \left(\det \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\right)^{\{\#i|P_{i,i}=-1\}}$. Also $\det A = 4^{\{\#i|P_{i,i}=-1\}} \det D$.

Satz 13.5.9. Jede symmetrische Matrix über einen Körper K mit $2 \neq 0$ besitzt eine verallgemeinerte Cholesky-Zerlegung.

Beweis. Gelte $2 \neq 0$ in K. Sei $n \in \mathbb{N}_0$ und $B \in \mathrm{SK}^{n\times n}$. Betrachte eine quadratische Form $q \in \mathrm{Q}(K^n)$ mit $M(q, e) = B$. Schreibe $q(x) = \sum_{i\leq j} a_{i,j} x_i x_j$ für $x \in K^n$ mit $a_{i,j} \in K$ $(1 \leq i \leq j \neq n)$.

<u>Fall 1</u>. $a_{1,1} = a_{1,2} = ... = a_{1,n} = 0$. Definiere $\lambda_1 := 0$ und $l_1 \in (K^n)^*$ durch $l(x) = x_1$ für $x \in K^n$. Betrachte sodann $q_1 \in \mathrm{Q}(K^{n-1})$ definiert durch $q_1 \begin{pmatrix} x_2 \\ \vdots \\ x_n \end{pmatrix} = \sum_{2\leq i\leq j} a_{i,j} x_i x_j$ für $q(x) = \lambda_1 l_1^2(x) + q_1 \begin{pmatrix} x_2 \\ \vdots \\ x_n \end{pmatrix}$ für $x \in K^n$.

<u>Fall 2</u>. $a_{1,1} \neq 0$. Definiere $\lambda_1 := a_{1,1}$ und $l_1 \in (K^n)^*$ durch $l_1(x) = x_1 + \sum_{j=2}^n \frac{a_{1,j}}{2a_{1,1}} x_j$ für $x \in K^n$. Dann gilt $\lambda_1 l_1^2(x) = a_{1,1}x^2 + \sum_{j=2}^n a_{1,j} x_1 x_j + \sum_{2\leq i\leq j} b_{i,j} x_i x_j$ für alle $x \in K^n$ und gewisse $b_{i,j} \in K$. Betrachte sodann $q_1 \in \mathrm{Q}(K^{n-1})$ definiert durch $q_1 \begin{pmatrix} x_2 \\ \vdots \\ x_n \end{pmatrix} = \sum_{2\leq i\leq j} (a_{i,j} - b_{i,j})_i^x 1 x_j$ für $\begin{pmatrix} x_2 \\ \vdots \\ x_n \end{pmatrix} \in K^{n-1}$. Es gilt $q(x) = \lambda_1 l_1^2(x) + q_1 \begin{pmatrix} x_2 \\ \vdots \\ x_n \end{pmatrix}$ für $x \in K^n$.

<u>Fall 3</u> $a_{1,1} = 0$ aber $a_{1,j} \neq 0$ für ein j.

> <u>Fall 3.1</u> $a_{j,j} \neq 0$. Vertausche Indizes 1 und j auf Kosten der Permutationsmatrix und nehme daher Œ an, dass wir im Fall 1 sind.
>
> <u>Fall 3.2</u> $a_{j,j} = 0$. Wieder durch Permutation können wir Œ $j = 2$ annehmen. Definiere $\lambda_1 := \frac{a_{1,2}}{4}, \lambda_2 := \frac{-a_{1,2}}{4}$ und $h_1, h_2 \in (K^n)^*$ durch

$h_1(x) = x_1 + \sum_{j=3}^n \frac{a_{2,j}}{a_{1,2}} x_j$ und $h_2(x) = x_2 + \sum_{j=3}^n \frac{a_{1,j}}{a_{1,2}} x_j$ für $x \in K^n$. Setze ferner $l_1 := h_1 + h_2$ und $l_2 := h_1 - h_2$. Dann gilt $\lambda_1 l_1^2(x) + \lambda_2 l_2^2(x) = a_{1,2} h_1(x) h_2(x) = a_{1,2} x_1 x_2 + \underbrace{\sum_{j=3}^n a_{1,j} x_j + \sum_{j=3}^n a_{2,j} x_j}_{= \sum_{i=1}^2 \sum_{j=1}^n a_{i,j} x_i x_j} + \sum_{3 \le i \le j} b_{i,j} x_i x_j$ für alle $x \in K^n$ und gewisse $b_{i,j} \in K$. Betrachte $q_1 \in \mathrm{Q}(K^{n-2})$ definiert durch $q_1 \begin{pmatrix} x_3 \\ \vdots \\ x_n \end{pmatrix} = \sum_{2 \le i \le j} (a_{i,j} - b_{i,j}) x_i x_j$ für $\begin{pmatrix} x_3 \\ \vdots \\ x_n \end{pmatrix} \in K^{n-2}$. Es gilt $q(x) = \lambda_1 l_1^2(x) + \lambda_2 l_2^2(x) + q_1 \begin{pmatrix} x_2 \\ \vdots \\ x_n \end{pmatrix}$ für $x \in K^n$.

Wendet man nun dasselbe Verfahren rekursiv auf $q_1 \in \mathrm{Q}(K^{n-1})$ bzw. im Fall 3.2 auf $q_1 \in \mathrm{Q}(K^{n-2})$ an usw., so erhält man eine Darstellung

$$q(x) = \sum_{i=1}^n \lambda_i l_i^2(x) \ (x \in K^n)$$

mit $\lambda_1, ..., \lambda_n \in K$ und $l_1, ..., l_n \in (K_n)^*$. Nach Lemma **??** gilt $M(q, \underline{e}) = P^T D P$ mit $D := \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \in K^{n \times n}$ und $P := (l_i(v_j))_{1 \le i \le m, 1 \le j \le n} \in K^{n \times n}$ und $P := (l_i(e_j))_{1 \le i,j \le n} \in K^{n \times n}$ Man erkennt, dass P die gewünschte Gestalt hat. In der Tat: in den Fällen 1 und 2 gilt (mit B_i und C_i wie in der Definition der verallgemeinerten Cholesky-Zerlegung) $B_1 = (l_1(e_1)) \in K^{1 \times 1}$ und $C_1 = (\lambda_1) \in K^{1 \times 1}$. Im Fall 3.2 gilt

$$B_1 = \begin{pmatrix} l_1(e_1) & l_1(e_2) \\ l_2(e_1) & l_2(e_2) \end{pmatrix} = \begin{pmatrix} (h_1 + h_2)(e_1) & (h_1 + h_2)(e_2) \\ (h_1 - h_2)(e_1) & (h_1 - h_2)(e_2) \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \in K^{2 \times 2}$$

und $C_1 = \begin{pmatrix} \lambda_1 & 0 \\ 0 & -\lambda_1 \end{pmatrix}$ mit $\lambda_1 = \frac{a_{1,2}}{4} \neq 0$.

Bemerkung 13.5.10. (a) Der Beweis von **??** zeigt, wie man im Fall $2 \neq 0$ in K verallgemeinerte Cholesky-Zerlegungen berechnen kann. Auf diese Weise wurde in den letzten beiden Beispielen bereits eine solche Zerlegung berechnet (mit der Einheitsmatrix als Permutationsmatrix).

(b) Es ist natürlich erstrebenswert in einer verallgemeinerten Cholesky-Zerlegung (S, P, D) einer Matrix $A \in \mathrm{SK}^{n \times n}$ zu erreichen, dass P in oberer Dreiecksgestalt vorliegt (vermeide Fall 3.2 im Beweis des Satzes) und $S = I_n$ gilt (vermeide Fall 3.1 , also unnötige Variablenvertauschungen).

(c) Ist $K = \mathbb{R}$, gilt $q(x) \ge 0$ für alle $x \in \mathbb{R}^n$, so sieht man leicht, dass Fall 3 in obigem Beweis im ersten Schritt nicht eintreten kann. In den Fällen 1 und 2 gibt es aber für

alle $x_2, ...x_n \in \mathbb{R}$ ein $x_1 \in \mathbb{R}$ mit $l_1(x_1) = 0$, woraus $q_1 \begin{pmatrix} x_2 \\ \vdots \\ x_n \end{pmatrix} = q(x) \geq 0$ folgt. Daher kann dann Fall 3 auch in den folgenden Schritten automatisch nicht eintreten.

Korollar 13.5.11. Gelte $0 \neq 2$ in K. Sei $n \in \mathbb{N}_0$ und $A \in \mathrm{SK}^{n\times n}$. Dann gibt es eine invertierbare Matrix $P \in K^{n\times n}$ und eine Diagonalmatrix $D \in K^{n\times n}$ mit $A = P^T DP$.

Korollar 13.5.12. (Diagonalisierung quadratischer Formen) Gelte $0 \neq 2$ in K. Sei V ein endlichdimensionaler Vektorraum und $b \in \mathrm{SBil}(V)$. Dann gibt es eine geordnete Basis $\underline{v}$ und V derart, dass $M(b, \underline{v})$ Diagonalgestalt hat.

Beweis. Wähle eine Basis $\underline{w} = (w_1, ..., w_n)$ von V. Wähle eine invertierbare Matrix $P \in K^{n\times n}$ und eine Diagonalmatrix $D \in K^{n\times n}$ mit $M(b, \underline{w}) = P^T DP$. Setze $v_i := \mathrm{vec}_{\underline{w}} P^{-1} e_i$. Dann ist $\underline{v} = (v_1, ..., v_n)$ eine Basis von V und es gilt

$$\begin{aligned} b(v_i, v_j) &= \mathrm{coord}_{\underline{w}}(v_i)^T M(b, \underline{w}) \,\mathrm{coord}_{\underline{w}}(v_j) = (P^{-1}e_i)^T P^T DP(P^{-1}e_j) \\ &= (PP^{-1}e_i)^T D(PP^{-1}e_j) = e_i^T De_j \end{aligned}$$

für $i, j \in \{1...n\}$, also $M(b, \underline{v}) = D$.

Skript zur linearen Algebra I & II

§14 Reelle quadratische Formen

14.1 Der Trägheitssatz von Sylvester

[James Joses Sylvester *1814 †1897]

Satz und Definition 14.1.1. (Trägheitssatz) Sei V ein endlichdimensionaler $\mathbb{R}$-Vektorraum und $q \in \mathrm{Q}(V)$. Dann gibt es genau ein Paar $(r,s) \in \mathbb{N}_0^2$, genannt Sylvester-Signatur mit a derart, dass es eine geordnete Basis $\underline{v}$ von V gibt mit

$$M(q,\underline{v}) = \begin{pmatrix} \underbrace{\begin{matrix} 1 & & \\ & \ddots & \\ & & 1 \end{matrix}}_{r} & & \\ & \underbrace{\begin{matrix} -1 & & \\ & \ddots & \\ & & -1 \end{matrix}}_{s} & \\ & & \mathbf{0} \end{pmatrix}$$

Beweis. Existenz. Nach Korollar **??** gibt es eine Basis $\underline{w} = (w_1, ..., w_n)$ von V derart, dass $M(q,\underline{w})$ Diagonalgestalt hat. Œ gibt es $r,s \in \mathbb{N}_0$ mit $q(w_1) > 0, ..., q(w_r) > 0, q(w_{r+1}) < 0, ..., q(w_{r+s}) < 0, q(w_{r+s+1}) = ... = q(w_n) = 0$. Setze nun

$$\underline{v} := \left(\frac{w_1}{\sqrt{q(w_1)}}, ..., \frac{w_r}{\sqrt{q(w_r)}}, \frac{w_{r+1}}{\sqrt{-q(w_{r+1})}}...., \frac{w_{r+s}}{\sqrt{-q(w_{r+s})}}, w_{r+s+1}, ..., w_n \right).$$

Dann ist $M(q,\underline{v})$ von der gewünschten Gestalt.

Eindeutigkeit. Seien $(r,s),(t,u) \in \mathbb{N}_0^2$ und $\underline{v} = (v_1, ..., v_n), \underline{w} = (w_1, ..., w_n)$ Basen

von V mit

$$M(q,\underline{v}) = \begin{pmatrix} \underbrace{\begin{matrix} 1 & & \\ & \ddots & \\ & & 1 \end{matrix}}_{r} & & & \\ & \underbrace{\begin{matrix} -1 & & \\ & \ddots & \\ & & -1 \end{matrix}}_{s} & & \\ & & \underbrace{\begin{matrix} 0 & & \\ & \ddots & \\ & & 0 \end{matrix}}_{} & \end{pmatrix} \text{ und}$$

$$M(q,\underline{w}) = \begin{pmatrix} \underbrace{\begin{matrix} 1 & & \\ & \ddots & \\ & & 1 \end{matrix}}_{t} & & & \\ & \underbrace{\begin{matrix} -1 & & \\ & \ddots & \\ & & -1 \end{matrix}}_{u} & & \\ & & \underbrace{\begin{matrix} 0 & & \\ & \ddots & \\ & & 0 \end{matrix}}_{} & \end{pmatrix}$$

Zu zeigen ist $(r,s) = (t,u)$. Setze $b := b_q$. Dann gilt für $\lambda_1, ..., \lambda_n \in K$:

$$\begin{aligned} \sum_{i=1}^{n} \lambda_i v_i \in \ker \overrightarrow{b} &\iff \overrightarrow{b}\left(\sum_{i=1}^{n} \lambda_i v_i\right) = 0 \\ &\iff b\left(\sum_{i=1}^{n} \lambda_i v_i, \cdot\right) = 0 \\ &\iff \forall j \in \{1,...,n\} : \underbrace{b\left(\sum_{i=1}^{n} \lambda_i v_i, v_j\right)}_{=\lambda_j q(v_j)} = 0 \\ &\iff \forall j \in \{1,...,r+s\} : \lambda_j = 0. \end{aligned}$$

Also $\operatorname{span}(v_{r+s+1}, ...v_n) = \ker \overrightarrow{b}$ und ebenso $\operatorname{span}(w_{r+s+1}, ...w_n) = \ker \overrightarrow{b}$. Es folgt $r+s = t+u$. Betrachte nun die Untervektorräume

- $U := \operatorname{span}(v_1, ...v_r, v_{r+s+1}, ..., v_n)$ und
- $W := \operatorname{span}(w_{t+1}, ..., w_{t+u})$.

Es gilt $q(v) \geq 0$ für $v \in U$ und $q(v) < 0$ für $W \backslash \{0\}$. Daher gilt $U \cap W = \{0\}$ und mit der Dimensionsformel [→**??**] für Untervektorräume $n \geq \dim U + \dim W = n - s + u$. Also ist $u \leq s$. Genauso zeigt man $s \leq u$. Es folgt $s = u$ und daher auch $r = t$. Somit $(r,s) = (t,u)$.

Satz 14.1.2. Es gelte der Fundamentalsatz der Algebra. Sei V ein endlichdimensionaler $\mathbb{R}$-Vektorraum und $q \in \mathrm{Q}(V)$ mit Sylvestersignatur $(r,s) \in \mathbb{N}_0$.

(a) Ist $\underline{v} = (v_1, ..., v_n)$ eine Basis und sind $\lambda_1, ...\lambda_n$ die Eigenwerte von $M(q,\underline{v})$ wobei

jedes λ_i seiner algebraischen Vielfachheit entsprechend oft aufgeführt ist [→**??**,**??**], das heißt $\chi_{M(q,\underline{v})} = \det(M(q,\underline{v}) - XI_n) = (-1)^n \prod_{i=1}^n (X - \lambda_i)$, so gilt

- $r = \#\{i \in \{1,...,n\} \mid \lambda_i > 0\}$ und
- $s = \#\{i \in \{1,...,n\} \mid \lambda_i < 0\}$.

(b) Ist $\underline{v} = (v_1,...,v_n)$ eine Basis von V, $D := \begin{pmatrix} d_1 & & 0 \\ & \ddots & \\ 0 & & d_n \end{pmatrix} \in \mathbb{R}^{n\times n}$ eine Diagonalmatrix und $P \in \mathbb{R}^{n\times n}$ invertierbar mit $M(q,\underline{v}) = P^T DP$ oder $M(q,\underline{v}) = P^{-1}DP$, so gilt

- $r = \#\{i \in \{1,...,n\} \mid d_i > 0\}$ und
- $s = \#\{i \in \{1,...,n\} \mid d_i < 0\}$.

(c) Sind $l_1,...,l_m \in V^*$ linear unabhängig und $\lambda_1,...,\lambda_n \in \mathbb{R}$ mit $\forall v \in V : q(v) = \sum_{i=1}^m \lambda_i l_i^2(v)$, so gilt

- $r = \#\{i \in \{1,...,n\} \mid \lambda_i > 0\}$ und
- $s = \#\{i \in \{1,...,n\} \mid \lambda_i < 0\}$.

Beweis. Wir zeigen zunächst den Fall $M(q,\underline{v}) = P^T DP$ von (b), dann (a), dann den Fall $M(q,\underline{v}) = P^{-1}DP$ von (b) und schließlich (c).

(b) Fall $M(q,\underline{v}) = P^T DP$.

Schritt 1. Œ gelte $P = I_n$.
Begründung. Setze $w_i := \mathrm{vec}_{\underline{v}}(P^{-1}e_i)$ für alle $i \in \{1,...,n\}$. Dann ist $\underline{w} = (w_1,...,w_n)$ eine Basis von V [→ **??**,**??**] und es gilt $P^{-1} = M(\underline{w},\underline{v})$ und somit

$$\begin{aligned} M(q,\underline{w}) &\overset{??}{=} M(\underline{w},\underline{v})^T M(q,\underline{w}) M(\underbrace{\ w\ }, \underline{v}) \\ &= (P^{-1})^T P^T DPP^{-1})(PP^{-1})^T D(PP^{-1}) = D. \end{aligned}$$

Schritt 2. Œ sei D von der Gestalt $\begin{pmatrix} 1 & & & & & & 0 \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & -1 & & & \\ & & & & \ddots & & \\ & & & & & -1 & \\ & & & & & & 0 \\ & & & & & & & \ddots \\ 0 & & & & & & & & 0 \end{pmatrix}$.

Begründung. Vertauschen der Basiselemente und Skalieren ändert die Anzahl der positiven bzw. negativen Diagonaleinträge von $M(q,\underline{v})$ nicht. $\square$

(a) Da $M(q,v)$ eine reelle symmetrische Matrix ist, gibt es nach **??** eine orthogonale Matrix $P \in \mathbb{R}^{n\times n}$ und einer Diagonalmatrix $D = \begin{pmatrix} d_1 & & 0 \\ & \ddots & \\ 0 & & d_n \end{pmatrix} \in \mathbb{R}^{n\times n}$ mit

$M(q,\underline{v}) = P^TDP = P^{-1}DP$. Da $M(q,\underline{v})$ und D ähnlich sind, haben sie dasselbe charakteristische Polynom [→**??**], das heißt $\prod_{i=1}^n(X-\lambda_i) = \prod_{i=1}^n(X-d_i)$. Dann gibt es nach **??** eine Permutation $\sigma \in S_n$ mit $(\lambda_1,...\lambda_n) = (d_{\sigma(1)},...d_{\sigma(n)})$, weshalb in der Behauptung die λ_i durch die d_i ersetzt werden können. Wegen $M(q,\underline{v}) = P^TDP$ folgt dann die Behauptung nach dem bereits bewiesenen Teil von (b).

(b) Fall $M(q,\underline{v}) = P^{-1}DP$. Setze $\lambda_i := d_i$ für $i \in \{1,...,n\}$. Dann gilt $\chi_{M(q,\underline{v})} \overset{??}{=} \chi_D = (-1)^n\prod_{i=1}^n(X-\lambda_i)$ und die Behauptung folgt nun aus (a).

(c) Ergänze $l_1,...,l_m$ zu einer Basis $l_1,...l_m,l_{m+1},...,l_n$ von V^* und setze $\lambda_{m+1} = ... = \lambda_n = 0$. Wähle eine Basis $\underline{v} = (v_1,...,v_n)$ von V. Nach **??** ist dann $P := (l_i(v_j))_{1\le i,j\le n} \in \mathbb{R}^{n\times n}$ invertierbar und mit $D := \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \in \mathbb{R}^{n\times n}$ gilt $M(q,\underline{v}) = P^TDP$. Nun folgt die Behauptung aus (b).

Definition 14.1.3. Sei $n \in \mathbb{N}_0$ und $A \in \mathrm{S}\mathbb{R}^{n\times n}$. Dann definiert man die Sylvester-Signatur der von A als die Sylvester-Signatur der zu A gehörigen quadratischen Form $q_A : \mathbb{R}^n \to \mathbb{R}, x \mapsto x^TAx$.

Bemerkung 14.1.4. Sei V ein $\mathbb{R}$-Vektorraum mit Basis $\underline{v} = (v_1,...,v_n)$ und $q \in \mathrm{Q}(V)$. Dann stimmen die Sylvester-Signaturen von q und $M(q,\underline{v})$ natürlich überein, denn setzt man $A := M(q,\underline{v})$, so gilt $M(q_A,\underline{e}) = A = M(q,\underline{v})$ und es liefert zum Beispiel (a) des obigen Satzes das Gewünschte.

Korollar 14.1.5. Sei $n \in \mathbb{N}_0$ und $A \in \mathrm{S}\mathbb{R}^{n\times n}$ mit Sylvester-Signatur (r,s).

(a) Gilt $\chi_A = (-1)^n\prod_{i=1}^n(X-\lambda_i)$ mit $\lambda_1,...\lambda_n \in \mathbb{R}$, so gilt
- $r = \#\{i \in \{1,...,n\} \mid \lambda_i > 0\}$ und
- $s = \#\{i \in \{1,...,n\} \mid \lambda_i < 0\}$.

(b) $D := \begin{pmatrix} d_1 & & 0 \\ & \ddots & \\ 0 & & d_n \end{pmatrix} \in \mathbb{R}^{n\times n}$ eine Diagonalmatrix und $P \in \mathbb{R}^{n\times n}$ invertierbar mit $M(q_A,\underline{v}) = P^TDP$ oder $M(q_A,\underline{v}) = P^{-1}DP$, so gilt
- $r = \#\{i \in \{1,...,n\} \mid d_i > 0\}$ und
- $s = \#\{i \in \{1,...,n\} \mid d_i < 0\}$.

(c) Sind $l_1,...l_m \in (\mathbb{R}^n)^*$ linear unabhängig und $\lambda_1,...,\lambda_n \in \mathbb{R}$ mit $\forall x \in \mathbb{R}^n : x^TAx = \sum_{i=1}^m \lambda_i l_i^2(v)$, so gilt
- $r = \#\{i \in \{1,...,n\} \mid \lambda_i > 0\}$ und
- $s = \#\{i \in \{1,...,n\} \mid \lambda_i < 0\}$.

14.2 Positiv semidefinite Matrizen

Definition 14.2.1. Sei V ein $\mathbb{R}$-Vektorraum. Mann nennt $q \in \mathrm{Q}(V)$ $\left\{\begin{matrix}\text{positiv}\\\text{negativ}\end{matrix}\right\}$ semidefinit (auch: $\left\{\begin{matrix}\text{psd}\\\text{nsd}\end{matrix}\right\}$), wenn $\forall v \in V : q(v)\left\{\begin{matrix}\geq\\\leq\end{matrix}\right\}0$. Gilt zusätzlich $\forall v \in V : (q(v) = 0 \implies v = 0)$, so nennt man q $\left\{\begin{matrix}\text{positiv}\\\text{negativ}\end{matrix}\right\}$ definit (auch: $\left\{\begin{matrix}\text{pd}\\\text{nd}\end{matrix}\right\}$). Man nenn $b \in \mathrm{SBil}(V)$ psd/nsd/pd/ng, wenn b symmetrisch und q_b psd/nsd/pd/nd ist.

Beispiel 14.2.2. Ein Skalarprodukt auf einem reellen Vektorraum ist per Definition nichts anderes als eine pd Bilinearform [→**??**].

Definition 14.2.3. Sei $n \in \mathbb{N}_0$ und $A \in \mathbb{R}^{n\times n}$. Es heißt A psd/nsd/pd/nd, wenn die zu gehörige Bilinearform $b_A : \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times n} \to \mathbb{R}, (x, y) \mapsto x^T Ay$ [→**??**(a)] psd/nsd/pd/nd ist

Bemerkung 14.2.4. Sei $n \in \mathbb{N}_0$ und $A \in \mathbb{R}^{n\times n}$. Dann ist A psd/nsd/pd/nd genau dann, wenn A symmetrisch ist [→**??**,**??**] und wenn $q_A : \mathbb{R}^n \to \mathbb{R}, x \mapsto x^T Ax$ psd/nsd/pd/nd ist.

Bemerkung 14.2.5. Sei V ein $\mathbb{R}$-Vektorraum mit Basis $\underline{v} = (v_1, ..., v_n)$ und $q \in \mathrm{Q}(V)$. Dann ist q psd / nsd / pd / nd $\iff M(q, \underline{v})$ psd / nsd / pd / nd.

Bemerkung 14.2.6. Für reelle quadratische Formen q gilt natürlich q nsd $\iff -q$ psd und q nd $\iff -q$ pd. Analoges gilt für reelle Bilinearformen und Matizen. Daher betrachten wir im Folgenden nur noch die Begriffe psd/pd.

Satz 14.2.7. Sei V ein $\mathbb{R}$-Vektorraum mit $n = \dim V < \infty$ und $q \in \mathrm{Q}(V)$. Dann sind äquivalent:

(a) q ist psd.

(b) Die Sylvester-Signatur ist $(r, 0)$ for ein $r \in \mathbb{N}_0$,

(c) $\exists l_1, ..., l_n \in V^* : \forall v \in V : q(v) = \sum_{i=1}^n l_i^2(v)$.

(d) $\exists m \in \mathbb{N}_0 : \exists l_1, ..., l_m \in V^* : \forall v \in V : q(v) = \sum_{i=1}^m l_i^2(v)$.

Beweis. **(*a*)** $\implies$ **(*b*)**. Dies folgt direkt aus der Definition der Sylvester-Signatur **??**.

(*b*) $\implies$ **(*c*)**. Dies folgt ebenfalls aus der Definition der Sylvester-Signatur mit Lemma **??**.

(*c*) $\implies$ **(*d*)** $\implies$ **(*a*)**. Diese sind trivial.

Definition 14.2.8. [→??] Sei $A \in \mathrm{S}\mathbb{R}$. Unter einer Cholesky-Zerlegung [André Louis Cholesky *1875 †1918] von A verstehen wir ein Paar (P, D) von Matrizen $P, D \in \mathbb{R}^{n\times n}$ mit $A = P^T DP$, wobei P von oberere Dreiecksgestalt ist mit lauter Einsen auf der Diagonale und D von Diagonalgestalt ohne negative Einträge.

Definition 14.2.9. Sei K ein kommutativer Ring, $n \in \mathbb{N}_0$ und $A \in K^{n\times n}$. Für jedes $I \subseteq \{1, ..., n\}$ bezeichne $A_I \in K^{(\#I)\times(\#I)}$ die Matrix, die aus A durch Streichen aller Zeichen i und Spalten j mit $i, j \notin I$ entsteht. Wir bezeichnen die Determinanten der n Matrizen $A_{\{1\}}$,$A_{\{1,2\}}$,$A_{\{1,2,3\}}$,...,$A_{\{1,...,n\}}$ als die Leithauptminoren (auch: führende Hauptminoren) von A und die Determinanten der 2^n-1 Matrizen A_I ($\emptyset \neq I \subseteq \{1, ..., n\}$) als die Hauptminoren von A. [Vorsicht: manche deutschsprachige Autoren bezeichnen nur die Leithauptminoren als Hauptminoren und haben keine Bezeichnung für unsere Hauptminoren].

Beispiel 14.2.10. Die Leithauptminoren von $A := \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$ sind $1, 0, 0$ und ihre Hauptminoren sind die Diagonaleinträge $1, 1, -1$, die Determinante $\det \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = 0$, $\det \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = -1$, $\det \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = -1$ und $\det(A) = 0$.

Satz 14.2.11. Es gelte der Fundamentalsatz der Algebra. Sei $A \in \mathrm{S}\mathbb{R}^{n\times n}$. Dann sind äquivalent:

(a) A ist psd.

(b) $\forall x \in \mathbb{R}^n : x^T Ax \geq 0$.

(c) Die Sylvester-Signatur von A ist $(r, 0)$ für ein $r \in \mathbb{N}_0$.

(d) Alle Eigenwerte von A sind ≥ 0.

(e) Alle Koeffizienten von $\det(A + XI_n) = \chi_A(-X) \in \mathbb{R}[X]$ sind ≥ 0.

(f) Alle Hauptminoren von A sind ≥ 0.

(g) A besitzt eine Cholesky-Zerlegung.

(h) Es gibt eine obere Dreiecksmatrix [→??] $B \in \mathbb{R}^{n\times n}$ mit $A = B^T B$.

(i) $\exists B \in \mathbb{R}^{n\times n} : A = B^T B$.

(j) $\exists m \in \mathbb{N}_0 : \exists B \in \mathbb{R}^{m\times n} : A = B^T B$.

(k) $\exists v_1, ..., v_n \in \mathbb{R}^n : A = \begin{pmatrix} \langle v_1, v_1\rangle & \dots & \langle v_1, v_n\rangle \\ \vdots & \ddots & \vdots \\ \langle v_n, v_1\rangle & \dots & \langle v_n, v_n\rangle \end{pmatrix}$.

(l) $\exists m \in \mathbb{N}_0 : \exists v_1, ..., v_n \in \mathbb{R}^m : A = \begin{pmatrix} \langle v_1, v_1 \rangle & \dots & \langle v_1, v_n \rangle \\ \vdots & \ddots & \vdots \\ \langle v_n, v_1 \rangle & \dots & \langle v_n, v_n \rangle \end{pmatrix}$.

(m) $\exists v_1, ...v_n \in \mathbb{R}^n : A = \sum_{i=1}^n v_i v_i^T$.

(n) $\exists m \in \mathbb{N}_0 : \exists v_1, ...v_m \in \mathbb{R}^n : A = \sum_{i=1}^m v_i v_i^T$.

Beweis. $(a) \iff (b)$ ist trivial, da A symmetrisch ist.

$(b) \iff (c)$ ist klar nach Definition **??** der Sylvester-Signatur.

$(c) \iff (d)$ folgt aus **??**(a).

$(d) \iff (e)$

" $\implies$ " Sind $\lambda_1, ..., \lambda_n \in \mathbb{R}$ die Eigenwerte von A gezählt mit algebraischer Vielfachheit [→**??**,**??**], so gilt $\chi_A = \prod_{i=1}^n (\lambda_i - X)$ und daher $\chi_A = \prod_{i=1}^n (\lambda_i + X)$. Die Koeffizienten von $\chi_A(-X)$ sind daher Summen von Produkten der λ_i.

" $\impliedby$ " Gelte (e). Sei $\lambda \in \mathbb{R}$ ein Eigenwert von A. Dann $\det(A - \lambda I_n) = \chi_A(\lambda) = 0$. Setzt man also $-\lambda$ anstelle von X in das Polynom $0 \neq \det(A + XI_n)$ ein, so erhält man 0. Hat dieses Polynom nur nichtnegativen Koeffizienten, so folgt $-\lambda \leq 0$.

$(e) \iff (f)$

" $\implies$ " Gelte (e). Sei $\emptyset \neq I \subseteq \{1, ..., n\}$. Zu zeigen ist $\det(A_I) \geq 0$. Wegen $(b) \iff (e)$ gilt auch (b). Insbesondere $\forall x \in \mathbb{R}^{\#I} : x^T A_I x \geq 0$. Wieder wegen $(b) \iff (e)$ hat $\det(A_I + XI_{\#I}) \in \mathbb{R}[X]$ keine negativen Koeffizienten. Insbesondere ist der konstante Koeffizient $\det(A_I)$ dieses Polynoms ≥ 0.

" $\impliedby$ " Schreibt man $\det(A + XI_n) = X^n + a_{n-1}X^{n-1} + ... + a_0$ mit $a_0, ..., a_{n-1} \in K$, so sieht man mit scharfem Auge direkt an der Definition einer Determinante **??**, dass $a_i = \sum_{\substack{I \subseteq \{1,...,n\} \\ \#(\{1,...,n\} \setminus I) = i}} \det(A_I)$ für $i \in \{0, ..., n-1\}$.

$(b) \implies (g)$ folgt wie in Bemerkung **??** (c) angekündigt durch Inspektion des Beweises von Satz **??**.

$(g) \implies (h)$. Ist $A = P^T D P$ mit $P = \begin{pmatrix} 1 & & * \\ & \ddots & \\ 0 & & 1 \end{pmatrix} \in \mathbb{R}^{n \times n}$ und $D = \begin{pmatrix} d_1 & & 0 \\ & \ddots & \\ 0 & & d_n \end{pmatrix} \in \mathbb{R}^{n \times n}$ mit $d_i \geq 0$, so ist $B := \begin{pmatrix} \sqrt{d_1} & & 0 \\ & \ddots & \\ 0 & & \sqrt{d_n} \end{pmatrix} P \in R^{n \times n}$ und $A = B^T B$.

$(h) \implies (i) \implies (j)$ ist trivial.

$\underline{(j) \implies (b)}$. Ist $A = B^T B$ mit $B \in \mathbb{R}^{m\times n}$, so gilt $x^T AX = x^T B^T Bx = \overline{(Bx)^T Bx = \langle Bx, Bx\rangle \geq 0}$ für alle $x \in \mathbb{R}^n$.

Es ist nun die Äquivalenz der Aussagen (a)-(j) gezeigt. Die Äquivalenzen $\underline{(i) \iff (k)}$ und $\underline{(j) \iff (l)}$ ergeben sich sofort, indem man die v_i als die Spalten von B auffasst, denn für $v_1, ..., v_n \in \mathbb{R}^m$ gilt $\begin{pmatrix} v_1^T \\ \vdots \\ v_n^T \end{pmatrix} \begin{pmatrix} v_1 \dots v_n \end{pmatrix} = \begin{pmatrix} \langle v_1,v_1\rangle & \dots & \langle v_1,v_n\rangle \\ \vdots & \ddots & \vdots \\ \langle v_n,v_1\rangle & \dots & \langle v_n,v_n\rangle \end{pmatrix}$.
Die Äquivalenzen $\underline{(i) \iff (m)}$ und $\underline{(j) \iff (n)}$ ergeben sich, indem man die v_i^T als die Zeilen von B auffasst, denn für $v_1, ..., v_m \in \mathbb{R}^n$ gilt

$$(v_1 \dots v_n) \begin{pmatrix} v_1^T \\ \vdots \\ v_n^T \end{pmatrix} = \sum_{i=1}^{m} (0 \dots 0 \underbrace{v_i}_{\uparrow\, i-\text{te Spalte}} 0 \dots 0) \begin{pmatrix} v_1^T \\ \vdots \\ v_n^T \end{pmatrix}$$

$$= \sum_{i=1}^{m} (0 \dots 0 \underbrace{v_i}_{\uparrow\, i-\text{te Spalte}} 0 \dots 0) \begin{pmatrix} 0 \\ \vdots \\ 0 \\ v_i^T \leftarrow i\text{-te Zeile} \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \sum_{i=1}^{m} v_i v_i^T.$$

Satz 14.2.12. [→**??**] Sei V ein endlichdimensionaler $\mathbb{R}$-Vektorraum, $n := \dim V$ und $q \in \mathbb{Q}(V)$. Es sind äquivalent:

(a) q ist pd

(b) Die Sylvester-Signatur von q ist $(n, 0)$

(c) Es gibt eines Basis $l_1, ..., l_n$ von V^* mit $\forall v \in V : q(v) = \sum_{i=1}^{n} l_i^2(v)$

Beweis. $\underline{(a) \implies (b)}$ folgt direkt aus der Definition der Sylvester-Signatur **??**.

$\underline{(b) \implies (c)}$ folgt ebenfalls aus dieser Definition zusammen mit Lemma **??**

$\underline{(c) \implies (a)}$ Gelte (c) und sei $0 \neq v \in V$. Zu zeigen ist $q(v) > 0$. Da die kanonische Auswertung $V \to V^{**}$ [→**??**] injektiv ist, gibt es $l \in V^*$ mit $l(v) \neq 0$. Wegen $l \in \operatorname{span}(l_1, ..., l_n)$ gibt es $i \in \{1, ..., n\}$ mit $l_i(v) \neq 0$. Daraus folgt $q(v) \geq l_i^2(v) > 0$.

Satz 14.2.13. Es gelte der Fundamentalsatz der Algebra. Sei $A \in \mathrm{S}\mathbb{R}^{n\times n}$. Dann sind äquivalent:

(a) A ist pd.

(b) $\forall x \in \mathbb{R}^n : x^T A x > 0$.

(c) Die Sylvester-Signatur von A ist $(n, 0)$ für ein $r \in \mathbb{N}_0$.

(d) Alle Eigenwerte von A sind > 0.

(e) Die Koeffizienten zu den Monomen $1, X, ..., X^{n-1}$ von $\det(A + XI_n) = \chi_A(-X) \in \mathbb{R}[X]$ sind > 0.

(f) Alle Leithauptminoren von A sind > 0

(g) Alle Hauptminoren von A sind > 0.

(h) A besitzt eine Cholesky-Zerlegung (P, D) derart, dass alle Diagonaleinträge von D positiv sind.

Beweis. Die Äquivalent aller Aussagen mit Ausnahme von (f) zeigt man analog zum Beweis von Satz **??**. Es ist $\underline{(g) \implies (f)}$ trivial. Wir zeigen schließlich $\underline{(f) \implies (b)}$ durch Induktion nach $n \in \mathbb{N}_0$.

$\underline{n = 0}$ Hier ist (b) die leer Aussage, da $\mathbb{R}^n = \mathbb{R}^0 = \{0\}$.
$\underline{n \to n+1 \ (n \in \mathbb{N}_0)}$ Seien die Leithauptminoren der Matrix $A \in \mathbb{R}^{(n+1)\times(n+1)}$ positiv. Schreibt man $A = \left(\begin{array}{c|c} B & \begin{matrix} a_1 \\ \vdots \\ a_n \end{matrix} \\ \hline a_1 \dots a_n & c \end{array}\right)$ mit $B \in \mathrm{S}\mathbb{R}^{n\times n}$ und $a_1, ...,_n, c \in \mathbb{R}$, so $x^T B x > 0$ für alle $x \in \mathbb{R}^n \setminus \{0\}$ (denn insbesondere sind alle Leithauptminoren von B positiv). Wähle nun $0 \neq v \in \ker \underbrace{\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}}_{\in \mathbb{R}^{n\times(n+1)}} B$. Wegen $\ker B = \{0\}$ kann nicht $v \in \mathrm{span}(e_1, ..., e_n)$ gelten, wobei $\underline{e} = (e_1, ..., e_{n+1})$ die Standardbasis des $\mathbb{R}^{n+1}$ bezeichnet. Dann ist $\underline{v} := (e_1, ..., e_n, v)$ und es gilt $e_i^T A v = 0$ f+r $i \in \{0, ..., n\}$. Es folgt, dass die Darstellungsmatrix $M(q_A, \underline{v})$ von der Form $M(q_A, \underline{v}) = \left(\begin{array}{c|c} B & \begin{matrix} 0 \\ \vdots \\ 0 \end{matrix} \\ \hline 0 \dots 0 & d \end{array}\right)$ mit $d \in \mathbb{R}$ ist. Wegen $A = M(q_A, \underline{e}) \overset{??}{=} M(\underline{e}, \underline{v})^T M(q_A, \underline{v}) M(\underline{e}, \underline{v})$ gilt

$$0 < \det A \overset{??}{\underset{??}{=}} (\det M(\underline{e}, \underline{v}))^2 \det(M(q_A, \underline{v})) = \underbrace{\det M(\underline{e}, \underline{v})}_{>0} \underbrace{(\det B)}_{>0} d$$

und daher $d > 0$. Nun gilt für alle $x \in \mathbb{R}^n$ und $y \in \mathbb{R}$, dass $\binom{x}{y}^T M(q_A, \underline{v}) \binom{x}{y} = x^T B x + d y^2 > 0$ falls $\binom{x}{y} \neq 0$ und damit q_A positiv definit, das heißt A ist positiv

definit.

Bemerkung 14.2.14. Wie man eine Cholesky-Zerlegung einer psd Matrix berechnet, ist aus dem Beweis von **??** wegen Bemerkung **??**(c) klar. Ist die Matrix sogar positiv definit, so kann auch der dortige Fall 1 nicht auftreten. Da im dortigen Fall 2 die Wahl der Linearform l_1 zwingend (d.h. eindeutig) ist, sieht man mit Hilfe von **??** leicht, dass die Cholesky-Zerlegung einer pd eindeutig ist.

Die in **??** bewiesene Diagonalisierung quadratischer Formen über beliebigen Körper mit $0 \neq 2$ kann über dem Körper der reellen Zahlen zu folgender in §11.3 in einer anderen Sprache formulierten Aussage verschärft werden (ßimultane Diagonalisierung").

Satz 14.2.15. Es gelte der Fundamentalsatz de Algebra. Sei V ein endlichdimensionaler $\mathbb{R}$-Vektorraum und $q_1, q_2 \in \mathrm{Q}(V)$. Ist q_1 pd oder nd, so gibt es eine geordnete Basis $\underline{v}$ von V derart, dass $M(q_1, \underline{v})$ und $M(q_2, \underline{v})$ beide Diagonalgestalt haben.

Beweis. Œ sei q_1 pd. Es ist b_{q_1} [→**??**] ein Skalarprodukt auf V vermöge dessen V zu einem Vektorraum mit Skalarprodukt wird [→**??**]. Wähle eine ONB $\underline{w}$ von V [→**??**,**??**]. Wähle $f \in \mathrm{End}(V)$ mit $M(f, \underline{w}) = M(q_2, \underline{w})$ (nämlich $f : \mathrm{vec}_{\underline{w}} \circ f_{M(q_2,\underline{w})} \circ \mathrm{coord}_{\underline{v}}$). Da $M(f, \underline{w})$ symmetrisch (also selbstadjungiert [→**??**]) und $\underline{w}$ eine ONB ist, ist f nach **??** selbstadjungiert. Nach Satz **??** gibt es eine ONB $\underline{v}$ von V, die aus Eigenvektoren von f besteht. Dann ist $M(q_1, \underline{v}) = M(b_{q_1}, \underline{v})$ die Einheitsmatrix (da $\underline{v}$ eine ONB ist) und

$$\begin{aligned} M(q_2, \underline{v}) &= (\underline{v}, \underline{w})^T M(q_2, \underline{w}) M(\underline{v}, \underline{w}) = (\underline{v}, \underline{w})^T M(f, \underline{w}) M(\underline{v}, \underline{w}) \\ &\overset{\underline{v},\underline{w}\ \mathrm{ONB}}{\underset{??}{=}} (\underline{v}, \underline{w})^{-1} M(f, \underline{w}) M(\underline{v}, \underline{w}) \overset{??}{=} M(\underline{w}, \underline{v}) M(f, \underline{w}) M(\underline{v}, \underline{w}) \\ &\overset{??}{=} M(f, \underline{v}) \end{aligned}$$

von Diagonalgestalt.

§15 Skalarprodukte

Dieses Kapitel ist eine Fortsetzung von §11. Wie dort, so sei auch hier stets $\mathbb{K} = \{\mathbb{R}, \mathbb{C}\}$.

15.1 Die adjungierte Abbildung

Definition und Proposition 15.1.1. Seien V, W $\mathbb{K}$-Vektorräume mit Skalarprodukt und $f : V \to W$ linear. Dann gibt es zu jedem $w \in W$ höchstens ein $v' \in V$ mit $\forall v \in V : \langle f(v), w\rangle = \langle v, v'\rangle$. Setzt man

$$W' = \{w \in W \mid \exists v' \in V : \forall v \in V : \langle f(v), w\rangle = \langle v, v'\rangle\},$$

so gibt es also genau eine Abbildung $f^* : W' \to V$ mit $\forall v \in V : \langle f(v), w\rangle = \langle v, f^*(w)\rangle$. Man nennt f^* die zu f adjungierte Abbildung. Es ist W' ein Unterverktorraum von W und f^* linear.

Beweis. Zu zeigen ist:

(a) die Eindeutigkeit von f^*

(b) $\forall w_1, w_2 \in W' : [w_1 + w_2 \in W'] \wedge [f^*(w_1 + w_2) = f^*(w_1) + f^*(w_2)]$

(c) $\forall w \in W', \lambda \in \mathbb{K}[\lambda w \in W'] \wedge [f^*(\lambda w) = \lambda f^*(w)]$

Zu (a). Sei $w \in W$. Sind $v_1', v_2' \in V$ mit $\forall v \in V : \langle v, v_1'\rangle = \langle f(v), w\rangle = \langle v, v_2'\rangle$ So folgt $\forall v \in V : \langle v, v_1' - v_2'\rangle = 0$, insbesondere $\langle v_1' - v_2', v_1' - v_2'\rangle = 0$, also $v_1' = v_2'$.

Zu (b). Seien $w_1, w_2 \in W'$. Dann gilt für alle $v \in V$

$$\begin{aligned}\langle f(v), w_1 + w_2\rangle &= \langle f(v), w_1\rangle + \langle f(v), w_2\rangle = \langle v, f^*(w_1)\rangle + \langle v, f^*(w_2)\rangle \\ &= \langle v, f^*(w_1) + f^*(w_2)\rangle\end{aligned}$$

Zu (c). Sei $w \in W'$ und $\lambda \in \mathbb{K}$. Dann gilt für alle $v \in V$

$$\langle f(v), \lambda w\rangle = \lambda\langle f(v), w\rangle = \lambda\langle v, f^*(w)\rangle = \langle v, \lambda f^*(w)\rangle.$$

Beispiel 15.1.2. Sei $V := C^\infty([0,1],\mathbb{R})$ die der $\mathbb{R}$-Vektorraum der unendlich oft differenzierbaren reellen Funktionen definiert auf $[0,1]$ mit dem durch $\langle f,g\rangle := \int_0^1 fg$ für alle $f,g \in V$ definierten Skalarprodukt. Die Ableitung sei geschrieben als $D : V \to V, f \mapsto f'$. Ist g im Definitionsbereich von D^*, so gilt

$$\forall f \in V : \int_0^1 f'g = \int_0^1 fD^*(g).$$

Andererseits gilt gemäß partieller Integration auch

$$\forall f \in V : \int_0^1 f'g = [fg]_0^1 - \int_0^1 fg'$$

und daher

$$\forall f \in V : \int_0^1 f(g + D^*(g)) = [fg]_0^1 .$$

Durch Einsetzen von "Nadelfunktionen" (siehe etwas Wikipedia-Eintrag zu "bumpfunction") für f sieht man nun $D^*(g) = -g'$. Daher ist der Definitionsbereich von D^* gleich

$$\begin{aligned}\left\{g \in V \mid \forall v \in V : \int_0^1 fg' = \int_0^f (-g')\right\} &= \{g \in V \mid \forall v \in V : [fg]_0^1 = 0\} \\ &= \{g \in V \mid g(0) = g(1) = 0\}.\end{aligned}$$

Es gilt also $D^* : \begin{cases}\{g \in V \mid g(0) = g(1) = 0\} \to V \\ g \mapsto -g'\end{cases}$.

Erinnerung 15.1.3. [→??] Sei V ein $\mathbb{K}$-Vektorraum mit Skalarprodukt und $f : V \to V$ linear. Dann heißt f selbstadjungiert, wenn $\forall v, w \in V : \langle f(v), w\rangle = \langle v, f(w)\rangle$.

Proposition 15.1.4. Sei V ein $\mathbb{K}$-Vektorraum mit Skalarprodukt und $f : V \to V$ linear. Dann ist f selbstadjungiert genau dann, wenn $f = f^*$ (was natürlich beinhaltet, dass f^* auf V definiert ist).

Beispiel 15.1.5. Sei $V := \{f \in C^\infty([0,1],\mathbb{C}) \mid f(0) = f(1)\}$ ein $\mathbb{C}$-Vektorraum mit dem durch $\langle f,g\rangle := \int_0^1 f^*g$ für alle $f,g \in V$ definierten Skalarpodukt, wobei für $f,g \in V$ jeweils f^* die zu f punktweise komplex-konjugierte Funktion ist. Betrachte $T : V \to V, f \mapsto \mathring{\imath} f'$. Es gilt für alle $f,g \in V$

$$\begin{aligned}\langle T(f), g\rangle = \int_0^1 (\mathring{\imath} f')^* g = -\mathring{\imath} \int_0^1 f^* g \overset{\text{part. Int.}}{=} -\mathring{\imath}\left(\underbrace{[f^*g]_0^1}_{\substack{=0 \\ \text{da } f,g\in V}} - \int_0^1 f^* g'\right) = \int_0^1 f^* T(g) \\ = \langle f, T(g)\rangle.\end{aligned}$$

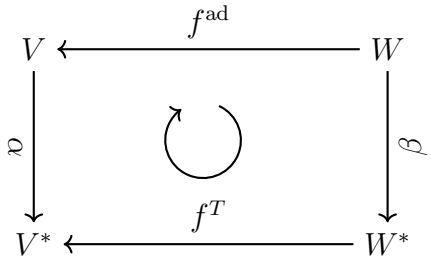
Daher ist $T = T^*$.

Satz 15.1.6. Seien V, W je $\mathbb{K}$-Vektorräume mit Skalarprodukt. Sei $f : V \to W$ linear und $\dim V < \infty$. Dann ist f^* auf ganz W definiert.

Beweis. Wähle mit **??** eine ONB $\underline{v} = (v_1, ..., v_n)$ von V. Dann gilt für jede lineare Abbildung $g : W \to V$:

$$\begin{aligned} g = f^* &\iff \forall v \in V : \forall w \in W : \langle f(v), w\rangle = \langle v, g(w)\rangle \\ &\iff \forall i \in \{1, ..., n\} : \forall w \in W : \langle f(v_i), w\rangle = \langle v_i, g(w)\rangle \\ &\iff \forall w \in W : \sum_{i=1}^{n} \langle f(v_i), w\rangle v_i = \sum_{i=1}^{n} \langle v_i, g(w)\rangle v_i \overset{??}{=} g(w). \end{aligned}$$

Bemerkung 15.1.7. Die Notation f^* hatten wir in **??** anders verwendet, nämlich für die zu einer linearen Abbildung gehörige lineare Abbildung $f^* : W^* \to V^*$. Man verwendet fast nie die adjungierte und die duale Abbildung gleichzeitig und aus dem Kontext ist fast immer klar, welche gemeint ist.
Wenn V, W endlichdimensionale $\mathbb{R}$-Vektorräume mit Skalarprodukt sind und $f : V \to W$ linear ist, dann sind außerdem die zu f duale Abbildung $f^T := f^* : W^* \to V^*$ und die zu f adjungierte Abbildung $f^{\text{ad}} := f^* : W \to V$ "im Prinzip dieselben", denn das Diagramm



mit den kanoninischen Isomorphismen

$$\begin{aligned} \alpha : V \to V^*, v_1 \mapsto (v_2 \mapsto \langle v_1, v_2\rangle) \text{ und} \\ \beta : W \to W^*, w_1 \mapsto (w_2 \mapsto \langle w_1, w_2\rangle) \end{aligned}$$

kommutiert. In der Tat: es gilt für alle $w \in W$ und $v \in V$

$$\begin{aligned} (\alpha(f^{ad}(w)))(v) &= \langle v, f^{ad}(w)\rangle = \langle f(v), w\rangle = \langle w, f(v)\rangle \\ &= (\beta(w))(f(v)) = (\beta(w) \circ f)(v) = (f^T(\beta(w)))(v). \end{aligned}$$

Lemma 15.1.8. Sei $A \in \mathbb{K}^{m\times n}$, $x \in \mathbb{K}^n$ und $y \in \mathbb{K}^m$. Dann ist

$$\langle Ax, y\rangle = (Ax)^* y = x^* A^* y = \langle x, A^* y\rangle.$$

Proposition 15.1.9. Seien V und W endlichdimensionale $\mathbb{K}$-Vektorräume mit Skalaprodukt und $f : V \to W$ linear. Sei $\underline{v}$ eine ONB von V und $\underline{w}$ eine ONB von W. Dann gilt $M(f^*, \underline{w}, \underline{v}) = M(f, \underline{v}, \underline{w})^*$.

Beweis. Es ist $(*)$ $f^* = \mathrm{vec}_{\underline{v}} \circ f_{M(f,\underline{v},\underline{w})^*} \circ \mathrm{coord}_{\underline{w}}$ zu zeigen. Es gilt

$$\begin{aligned}(*) &\iff \forall v \in V : \forall w \in W : \langle f(v), w\rangle = \langle v, \mathrm{vec}_{\underline{v}}(M(f,\underline{v},\underline{w})^* \,\mathrm{coord}_{\underline{w}}(w))\rangle \\ &\overset{??}{\iff} \forall v \in V : \forall w \in W : \langle \mathrm{coord}_{\underline{w}}(f(v)), \mathrm{coord}_{\underline{w}}(w)\rangle \\ &= \langle \mathrm{coord}_{\underline{v}}(v), M(f,\underline{v},\underline{w})^* \,\mathrm{coord}_{\underline{w}}(w)\rangle.\end{aligned}$$

Proposition 15.1.10. Seien U, V und W je $\mathbb{K}$-Vektorräume mit Skalarprodukt, und U und V endlichdimensional. Sei $\lambda \in \mathbb{K}$ und $f, f_1, f_2 : U \to V$ sowie $g : V \to W$ linear. Dann gilt $(f_1 + f_2)^* = f_1^* + f_2^*$, $(\lambda f)^* = \lambda^* f^*$, $(g \circ f)^* = f^* \circ g^*$, $id_V^* = id_V$ und $f^{**} = f$.

Beweis. Zu zeigen ist:

$\forall u \in U : \forall v \in V : \langle (f_1 + f_2)(u), v)\rangle = \langle u, (f_1^* + f_2^*)(v)\rangle$,

(b) $\forall u \in U : \forall v \in V : \langle (\lambda f)(u), v\rangle = \langle u, (\lambda^* f^*)(v)\rangle$,

(c) $\forall u \in U : \forall v \in W : \langle (g \circ f)(u), w\rangle = \langle u, (f^* \circ g^*)(w)\rangle$,

(d) $\forall u \in U : \forall v \in V : \langle (id_V(u), v)\rangle = \langle u, id_V(v)\rangle$,

(e) $\forall u \in U : \forall v \in V : \langle (f^*(u), v)\rangle = \langle u, f(v)\rangle$.

Zu (a). Seien $u \in U$ und $v \in V$. Dann

$$\begin{aligned}\langle (f_1 + f_2)(u), v\rangle &= \langle f_1(u), v\rangle + \langle f_2(u), v\rangle = \langle u, f_1^*(v)\rangle + \langle u, f_2^*(v)\rangle \\ &= \langle u, f_1^*(v) + f_2^*(v)\rangle = \langle u, (f_1^* + f_2^*)(v)\rangle.\end{aligned}$$

Zu (b). Seien $u \in V$ und $v \in V$. Dann

$$\begin{aligned}\langle (f)(u), v\rangle &= \langle \lambda(f(u)), v\rangle = \lambda^* \langle f(u), v\rangle = \lambda^* \langle u, f^*(v)\rangle \\ &= \langle u, \lambda^*(f^*(v))\rangle = \langle u, (\lambda^* f^*)(v)\rangle.\end{aligned}$$

Zu (c). Seien $u \in V$ und $w \in W$. Dann gilt

$$\langle (g \circ f)(u), w\rangle = \langle f(u), g^*(w)\rangle = \langle u, (f^* \circ g^*)(w)\rangle.$$

(d) ist trivial.

Zu (d). Seien $u \in U$ und $v \in V$. Dann gilt

$$\langle f^*(v), u\rangle = \langle u, f^*(v)\rangle^* = \langle f(u), v\rangle^* = \langle v, f(u)\rangle.$$

Falls W auch endlichdimensional ist, kann man den Beweis von Proposition **??** durch Rückführung auf die entsprechenden Tatsachen für Matrizen führen.

Proposition 15.1.11. Seien V und W je $\mathbb{K}$-Vektorräume mit Skalarprodukt und $f : V \to W$ linear. Dann gilt $\ker(f^*) = (\mathrm{im}\, f)^\perp$.

Beweis. Für $w \in W$ gilt gemäß Definition **??** der adjungierten Abbildung

$$w \in \ker(f^*) \iff \forall v \in V : \langle f(v), w\rangle = \langle v, 0\rangle.$$

Satz 15.1.12. Seien V und W $\mathbb{K}$-Vektorräume mit Skalarprodukt und $f : V \to W$ linear. Dann sind äquivalent:

(a) f ist ein Isomorphismus von Vektorräumen mit Skalarprodukt.

(b) f^* ist auf ganz W definiert und es gilt sowohl $f^* \circ f = id_V$ als auch $f \circ f^* = id_W$.

(c) f^* ist eine Bijektion von $W \to V$, deren Umkehrabbildung f ist.

Beweis. (c) ist nur eine Umformulierung von (b).

$(a) \implies (b)$. Gelte (a). Sei $v' \in V$. Setze $w := f(v')$. Wegen $\langle f(v), w\rangle \overset{(a)}{=} \langle v, v'\rangle$ für alle $v \in V$. ist f^* in ganz W definiert und es gilt $f^*(w) = v'$, das heißt $f^*(f(v')) = v'$. Da $v' \in V$ beliebig war, ist f^* auf $\operatorname{im} f \overset{(a)}{=} W$ definiert und es gilt $f^* \circ f = id_V$ Weiter gilt $f^* = f^* \circ (f \circ f^{-1}) = (f^* \circ f) \circ f^{-1} = f^{-1}$ und daher auch $f \circ f^* = id_W$.

$(b) \implies (a)$. Gelte (b). Dann hat f eine Umkehrabbildung und ist bijektiv. Da f auch linear ist, ist f ein Isomorphismus mit Vektorräumen. Es bleibt zu zeigen, dass für alle $v, v' \in V$ gilt $\langle f(v), f(v')\rangle = \langle v, v'\rangle$. Seien also $v, v' \in V$. Dann $\langle f(v), f(v')\rangle = \langle v, f^*(f(v'))\rangle = \langle v, (f^* \circ f)(v')\rangle = \langle v, v'\rangle$.

15.2 Normale Abbildungen

Definition 15.2.1. Sei V ein Vektorraum mit Skalarpodukt und $f : V \to V$ linear. Dann heißt f normal, wenn f^* auf ganz V definiert ist [→**??**] und $f \circ f^* = f^* \circ f$.

Beispiel 15.2.2. Sei V ein Vektorraum mit Skalarprodukt und $f : V \to V$ linear.

(a) Ist f selbstadjungiert (das heißt $f = f^*$), so ist f normal.

(b) Ist f ein Automorphismus von Vektorräumen mit Skalarprodukt [→**??**], so ist f ebenfalls normal, denn es ist $f \circ f^* = id_V = f^* \circ f$.

Lemma 15.2.3. Sei V ein endlichdimensionaler $\mathbb{K}$-Vektorraum mit Skalarprodukt, $\lambda \in \mathbb{K}$ und $f : V \to V$ normal. Dann gilt $\ker(f - \lambda \operatorname{id}_V) = \ker(f^* - \lambda^* \operatorname{id}_V)$.

Beweis. Wegen

$$\begin{aligned}(f-\lambda id_V)\circ(f-\lambda id_V)^* &\overset{??}{=} (f-\lambda id_V)\circ(f^*-\lambda^* id_V)\\ &= f\circ f^* - \lambda f^* - \lambda^* f + \lambda\lambda^* id_V\\ &= f^*\circ f - \lambda f^* - \lambda^* f + \lambda^*\lambda id_V\\ &= (f^*-\lambda^* id_V)\circ(f-\lambda id_V)\end{aligned}$$

ist auch $f-\lambda id_V$ normal. Daher reicht es $\ker f = \ker f^*$ zu zeigen. Dies folgt aus

$$\begin{aligned}\|f(v)\|^2 &= \langle f(v), f(v)\rangle = \langle v, f^*(f(v))\rangle\\ &= \langle v, f(f^*(v))\rangle = \langle f^*(v), f^*(v)\rangle = \|f^*(v)\|^2\end{aligned}$$

für alle $v\in V$.

Satz 15.2.4. [→**??**] Es gelte der Fundamentalsatz der Algebra. Sei V ein endlichdimensionaler $\mathbb{C}$-Vektorraum mit Skalarprodukt und $f: V\to V$ linear. Es sind äquivalent:

(a) f ist normal,

(b) V hat eine ONB, die aus Eigenvektoren von f besteht.

(c) Es gibt eine ONB $\underline{v}$ von V derart, dass $M(f,\underline{v})$ Diagonalgestalt hat.

Beweis. $(\boldsymbol{b}) \implies (\boldsymbol{c})$ ist klar.

$(\boldsymbol{c}) \implies (\boldsymbol{a})$. Sei $\underline{v}$ eine ONB von V mit $M(f,\underline{v})$ in Diagonalgestalt. Nach **??**, **??** reicht es $M(f\circ f^*,\underline{v}) = M(f^*\circ f,\underline{v})$ zu zeigen. Es gilt aber

$$\begin{aligned}M(f\circ f^*,\underline{v}) &= M(f,\underline{v})M(f^*,\underline{v}) \overset{\underline{v}\ \text{ONB}}{\underset{??}{=}} M(f,\underline{v})M(f,\underline{v})^*\\ &\overset{Diagonagestalt}{=} M(f,\underline{v})^*M(f,\underline{v}) \overset{\underline{v}\ \text{ONB}}{\underset{??}{=}} M(f^*,\underline{v})M(f,\underline{v})\\ &= M(f^*\circ f,\underline{v}).\end{aligned}$$

$(\boldsymbol{a}) \implies (\boldsymbol{c})$. Wir zeigen die Implikation per Induktion nach $n := \dim V \in \mathbb{N}_0$:
$\underline{n=0}$ Es ist nichts zu zeigen.
$\underline{n-1\to n\ (n\in\mathbb{N})}$ Sei $f: V\to V$ normal. Wegen $\deg\chi_f = n \geq 1$ gibt es nach dem Fundmantalsatz der Algebra einen Eigenwert λ von f. wähle dazu einen Eigenvektor $u\in V$, das heißt $u\neq 0$ und $f(u)=\lambda u$. Setze $U := \operatorname{span}(u)$ Es gilt $f(U^\perp)\subseteq U^\perp$ und $f^*(U^\perp)\subseteq U^\perp$, denn ist $v\in U^\perp$ und $u\in U$, so gilt

$$\begin{aligned}\langle f(v),u\rangle = \langle v, f^*(u)\rangle &\overset{??}{=} \langle v,\lambda^* u\rangle = \lambda^*\langle v,u\rangle = 0 \text{ und}\\ \langle f^*(v),u\rangle = \langle v, f(u)\rangle &= \langle v,\lambda u\rangle = \lambda\langle v,u\rangle = 0.\end{aligned}$$

Betrachte nun $f|_{U^\perp}: U^\perp\to U^\perp$ und $f^*|_{U^\perp}: U^\perp\to U^\perp$. Anhand von **??** sieht man

leicht, dass $f^*|_{U^\perp} = f|_{U^\perp}^*$. Daher hat man

$$f|_{U^\perp}^* \circ f|_{U^\perp} = f^*|_{U^\perp} \circ f|_{U^\perp} = (f^* \circ f)|_{U^\perp} = (f \circ f^*)|_{U^\perp} = f|_{U^\perp} \circ f^*|_{U^\perp} = f|_{U^\perp} \circ f|_{U^\perp}^*,$$

weswegen $f|_{U^\perp}$ ebenfalls normal ist. Wegen $\dim(U^\perp) \overset{??}{=} n-1$ gibt es nach IV eine ONB $(v_2, ..., v_n)$ von $U^\perp$, die aus Eigenwerten von f besteht. Setze $v_1 := u/\|u\|$, so erhält man eine ONB $(v_1, ..., v_n)$ von V, die aus Eigenwerten von f besteht.

Korollar 15.2.5. Es gelte der Fundamentalsatz der Algebra. Sei V ein endlichdimensionaler $\mathbb{C}$-Vektorraum mit Skalarprodukt und $f : V \to V$ normal. Dann ist f diagonalisierbar [→**??**(a)].

Korollar 15.2.6. Es gelte der Fundamentalsatz der Algebra. Sei V ein endlichdimensionaler $\mathbb{C}$-Vektorraum mit Skalarprodukt und $f : V \to V$ linear. Es sind äquivalent:

(a) f ist orthogonal (auch: unitär [→**??**]), das heißt $\forall v, w \in V : \langle f(v), f(w)\rangle = \langle v, w\rangle$.

(b) f ist ein Isomorphismus von Vektorräumen mit Skalarprodukt.

(c) V besitzt eine ONB die aus Eigenvektoren von f zu Eigenwerten vom Betrag Eins besteht.

(d) Es gibt eine ONB $\underline{v}$ von V derart, dass $M(f, \underline{v})$ Diagonalgestalt mit Diagonaleinträgen vom Betrag 1 hat.

Beweis. **(*a*)** $\implies$ **(*b*)** ist klar, da f injektiv, und weil V endlichdimensional ist, auch surjektiv ist [→**??**].

(*b*) $\implies$ **(*c*)** folgt sofort aus Satz **??**, denn wenn (b) gilt, so ist f normal und alle Eigenwerte von f haben den Absolutbetrag 1: Sei $\lambda \in \mathbb{C}$ und $v \in V \setminus \{0\}$ mit $f(v) = \lambda v$). Dann ist $|\lambda|\|v\| = \|\lambda v\| = \|f(v)\| = \|v\|$ und daher $|\lambda| = 1$.

(*c*) $\implies$ **(*d*)** ist klar.

(*d*) $\implies$ **(*a*)**. Sei $\underline{v} = (v_1, ..., v_n)$ eine ONB von V mit $M(f, \underline{v}) = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$ und $|\lambda_i| = 1$ für $i \in \{1, ..., n\}$. Dann gilt

$$\begin{aligned} M(f \circ f^*, \underline{v}) &= M(f, \underline{v})M(f^*, \underline{v}) \overset{\underline{v}\ \text{ONB}}{=} M(f, \underline{v})M(f, \underline{v})^* \\ &= \begin{pmatrix} \lambda_1\lambda_1{}^* & & 0 \\ & \ddots & \\ 0 & & \lambda_n\lambda_n{}^* \end{pmatrix} = \begin{pmatrix} |\lambda_1|^2 & & 0 \\ & \ddots & \\ 0 & & |\lambda_n|^2 \end{pmatrix} = I_n \end{aligned}$$

und daher $f \circ f^* = id_V$. Analolg folgt $f^* \circ f = id_V$ [alternativ: aus $f \circ f^* = id_V$ folgt, f^* ist injektiv und damit f^* auch bijektiv, wodurch folgt $f^* \circ f = f^* \circ f \circ f^* \circ (f^*)^{-1} = f^* \circ id_V \circ (f^*)^{-1} = \mathrm{id}_V$].

Um dem Leser zu helfen, die Resultate einzuordnen, formulieren wir Satz **??** noch einmal leicht anders. Er wurde damals durch eine kleine Variante des Beweises von **??** gezeigt, aber zumindest für $\mathbb{K} = \mathbb{C}$ erhält man ihn auch leicht als Korollar aus dem obigen Satz **??**.

Satz 15.2.7. Es gelte der Fundamentalsatz der Algebra. Sei V ein endlichdimensionaler $\mathbb{C}$-Vektorraum mit Skalarprodukt und $f : V \to V$ linear. Dann sind äquivalent:

(a) f ist selbstadjungiert (für $\mathbb{K} = \mathbb{C}$ auch hermitesch genannt [→**??**]).

(b) V hat eine ONB die aus Eigenvektoren zu reellen Eigenwerten von f besteht.

(c) Es gibt eine ONB $\underline{v}$ von V derart, dass $M(f, \underline{v})$ eine reelle Diagonalmatrix ist.

Anschauung. Normale, orthogonale und selbstadjungierte Endomorphismen können bezüglich gewisser ONB anschaulich in der komplexen Zahlebene dargestellt werden.

Endomorphismusart	**Charakteristik der ONB**
normal	ist existent
orthogonal	Eigenwerte liegen auf Einheitskreis
selbstadjungiert	Eigenwerte sind reell

Sei V ein endlichdimensionaler $\mathbb{K}$-Vektorraum. Dann kann man für $f \in \mathrm{End}(V)$ und eine bestimmte ONB davon mit Eigenwerten $\lambda_1, ...\lambda_n \in \mathbb{K}$ ($n := \dim V$) obige Tabelle auch formal darstellen.

Bedingung an f	**Bedingung an alle λ_i**
$f \circ f^* = f^* \circ f$	$\lambda_i \lambda_i^* = \lambda_i^* \lambda_i$
$\forall v \in V : \|f(v)\| = \|v\|$ [→**??**]	$\|\lambda_i\| = 1$
$f = f^*$	$\lambda_i = \lambda_i^*$

Es bleibt anzumerken, dass die Bedingung der Eigenwerte zu der ONB einer normalen Abbildung nur die immer gegebene Kommutativität der Körpermultiplikation darstellt, d.h. immer gegeben ist.

Notation 15.2.8. Für den Rest des Abschnitts notieren wir $\overline{x} := \begin{pmatrix} x_1^* \\ \vdots \\ x_n^* \end{pmatrix} = (x^*)^T$ für $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{C}^n$.

Lemma 15.2.9. Seien $x, y \in \mathbb{R}^n$ und $w, \in \mathbb{C}^n$ mit $\sqrt{2}w = x + iy$. Ist dann $(w, \overline{w})$ ein ONS in $\mathbb{C}^n$, so ist (x, y) eine ONB von $\mathrm{span}(w, \overline{w})$ in $\mathbb{C}^n$.

Beweis. Sei $(w, \overline{w})$ ein ONS in $\mathbb{C}^n$, das heißt $\langle w, w\rangle = \langle \overline{w}, \overline{w}\rangle = 1$ und $\langle \overline{w}, w\rangle = 0$.

Dann ist

$$1 = \langle w, w\rangle = \frac{1}{2}\langle x+iy, x+iy\rangle = \frac{1}{2}(\langle x,x\rangle + i\langle x,y\rangle - i\langle y,x\rangle - i^2\langle y,y\rangle) =$$
$$\frac{1}{2}(\langle x,x\rangle + \langle y,y\rangle + \underbrace{i(\underbrace{\langle x,y\rangle}_{\in\mathbb{R}} - \langle x,y\rangle^*)}_{=0}) = \frac{1}{2}(\langle x,x\rangle + \langle y,y\rangle)$$

und

$$0 = \langle w, \overline{w}\rangle = \frac{1}{2}\langle x+iy, x-iy\rangle = \frac{1}{2}(\langle x,x\rangle - i\langle x,y\rangle - i\langle x,y\rangle - \langle y,y\rangle)$$
$$= \frac{1}{2}(\langle x,x\rangle - \langle y,y\rangle - 2i\langle x,y\rangle),$$

woraus folgt $0 = \frac{1}{2}(\langle x,x\rangle - \langle y,y\rangle)$ und $\langle x,y\rangle = 0$. Insgesamt folgt $\langle x,x\rangle = \langle y,y\rangle = 1$ und $0 = \langle x,y\rangle$, wehsalb (x,y) ein ONS in $\mathbb{C}^n$ ist mit $w, \overline{w} \in \operatorname{span}_{\mathbb{C}}(x,y)$. Wegen $\operatorname{span}_{\mathbb{C}}(w,\overline{w}) \subseteq \operatorname{span}_{\mathbb{C}}(x,y)$ und $\dim(\operatorname{span}_{\mathbb{C}}(w,\overline{w})) = 2$ folgt $\operatorname{span}_{\mathbb{C}}(w,\overline{w}) = \operatorname{span}_{\mathbb{C}}(x,y)$.

Satz 15.2.10. [→**??**] Es gelte der Fundamentalsatz der Algebra. Sei V ein endlichdimensionaler $\mathbb{R}$-Vektorraum mit Skalarprodukt und $f : V \to V$ linear. Es sind äquivalent:

(a) f ist normal [→**??**].

(a) Es gibt eine ONB $\underline{v}$ von V derart, dass $M(f,\underline{v})$ von der Gestalt

$$\begin{pmatrix} \lambda_1 & & & & \\ & \ddots & & & \\ & & \lambda_k & & \\ & & & \boxed{\begin{matrix} a_1 & b_1 \\ -b_1 & a_1 \end{matrix}} & \\ & & & & \ddots \\ & & & & \boxed{\begin{matrix} a_l & b_l \\ -b_l & a_l \end{matrix}} \end{pmatrix}$$ ist.

Beweis. **(*b*)** $\Longrightarrow$ **(*a*)**. Für $(a,b) \in \mathbb{R}$ gilt

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix}^* \begin{pmatrix} a & b \\ -b & a \end{pmatrix} = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}\begin{pmatrix} a & b \\ -b & a \end{pmatrix} = \begin{pmatrix} a^2+b^2 & 0 \\ 0 & a^2+b^2 \end{pmatrix}$$
$$= \begin{pmatrix} a & b \\ -b & a \end{pmatrix}\begin{pmatrix} a & -b \\ b & a \end{pmatrix} = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}\begin{pmatrix} a & b \\ -b & a \end{pmatrix}^*$$

und der Beweis geht daher wie der Beweis von $(c) \Longrightarrow (a)$ in Satz **??**

(*a*) $\Longrightarrow$ **(*b*)**. Nach **??** ist V als Vektorraum mit Standardskalarprodukt isomorph zu $\mathbb{R}^n$ ($n = \dim V$) mit dem Standardskalarprodukt. Daher sei Œ $V = \mathbb{R}^n$ und

$f = f_A$ mit $A := M(f, \underline{e}) \in \mathbb{R}^{n\times n}$. Betrachte $g : \mathbb{C}^n \to \mathbb{C}^n, x \mapsto Ax$. Offensichtlich gilt $A = M(g, \underline{e})$ und wegen $A^*A = AA^*$ ist g normal. Daher gibt es eine ONB $\underline{w} = (w_1, ..., w_n)$ von $\mathbb{C}^n$, die aus Eigenwerten von g besteht. Bezeichne $\lambda_i \in \mathbb{C}$ den zu w_i gehörigen Eigenwert von g ($i \in \{1, ..., n\}$). Dann gilt $\chi_g = \det(A - XI_n) = (-1)^n \prod_{i=1}^n (X - \lambda_i)$ und wegen $\chi_g \in \mathbb{R}[X]$ (weil $A \in \mathbb{R}^{n\times n}$) auch $\chi_g = (-1)^n \prod_{i=1}^n (X - \lambda_i^*)$. Es folgt [→**??**], dass die Tupel $(\lambda_1, ...\lambda_n)$ und $(\lambda_1^*, ..., \lambda_n^*)$ bis auf Permutationen der Einträge dieselben sind. Nach allfälligem Umnummerieren der λ_i und w_i können wir daher davon ausgehen, dass

- $\lambda_1, ..., \lambda_k$ reell sind,
- $\lambda_{k+1}, \lambda_{k+3}, ...$ einen positiven Imaginärteil haben und
- $\lambda_{k+1}^* = \lambda_{k+2}$, $\lambda_{k+3} = \lambda_{k+4}^*, ...$

Wir ersetzen nun Œ in der ONB $\underline{w}$ mehrmals jeweils endlich viele $w_{i_1}, ...w_{i_r}$ ($1 \le i_1 \le i_2 \le ... \le i_r$) durch eine ONB $(u_1, ..., u_r)$ von $\operatorname{span}(w_{i_1}, ...w_{i_r})$ mit $f(u_j) = \lambda_{i_j} u_j$ für $j \in \{1, ..., r\}$. Wir behaupten, dass wir so $w_1, ..., w_k \in \mathbb{R}^n$ und $\overline{w}_{k+1} = w_{k+2}, \overline{w}_{k+3} = w_{k+4}, ...$ erreichen können [→**??**].

Schritt 1. Œ $w_1, ..., w_k \in \mathbb{R}^n$.
Begründung. Für jeden reellen Eigenwert λ von g seien $r \in \mathbb{N}$ und $1 \le i_1 \le ... \le i_r \le n$ derart, dass $\{i_1, ..., i_r\} = \{i \mid \lambda_i = \lambda\} \subseteq \{1, ..., k\}$. Dann gilt $\operatorname{span}(w_{i_1}, ..., w_{i_r}) = \ker(g - \lambda \mathrm{id}_{\mathbb{C}^n})$, denn " $\subseteq$ " ist trivial und $r \ge \dim \ker(g - \lambda \mathrm{id}_{\mathbb{C}^n})$ nach **??** ("geometrische Vielfachheit ist kleiner/gleich algebraischer").
Nun gilt $\ker(g - \lambda \operatorname{id}_{\mathbb{C}^n}) = \ker_{\mathbb{C}^n}(A - \lambda I_n) \supseteq \ker_{\mathbb{R}^n}(A - \lambda I_n)$. Da $A - \lambda I_n$ als reelle Matrix denselben Rang hat wie als komplexe Matrix (der Rang kann durch Überführung in Stufenform über $\mathbb{R}$ ermittelt werden [→,**??**], dieselben Zeilenoperationen sind erst recht über $\mathbb{C}$ durchführbar), gilt $r = \dim \ker_{\mathbb{C}^n}(A - \lambda I_n) = \dim \ker_{\mathbb{R}^n}(A - \lambda I_n)$. Wähle nun eine ONB $(u_1, ..., u_r)$ des $\mathbb{R}$-Vektorraums $\ker_{\mathbb{R}^n}(A - \lambda I_n)$ (mit Standardskalarprodukt). Dann ist $(u_1, ..., u_r)$ auch eine ONB des $\mathbb{C}$-Vektorraums $\ker_{\mathbb{C}^n}(A - \lambda I_n) = \operatorname{span}(w_{i_1}, ..., w_{i_r})$.

Schritt 2. Œ $\overline{w_{k+1}} = w_{k+2}, \overline{w_{k+3}} = w_{k+4}, ...$.
Begründung. Für jeden Eigenwert λ von g mit negativem Imaginärteil seien $r \in \mathbb{N}$ und $1 \le i_1 \le ... \le i_r \le n$ derart, dass $\{i_1, ..., i_r\} = \{i \mid \lambda_i = \lambda\} \subseteq \{k+2, k+4, ..., n\}$. Dann gilt wieder $\operatorname{span}(w_{i_1}, ...w_{i_r}) = \ker(g - \lambda \operatorname{id}_{\mathbb{C}^n})$. Nun gilt

$$\begin{aligned}
\ker(g - \lambda \operatorname{id}_{\mathbb{C}^n}) &= \ker(A - \lambda I_n) = \{x \in \mathbb{C}^n \mid Ax = \lambda x\} \\
&= \{\overline{x} \in \mathbb{C}^n \mid A\overline{x} = \lambda^* \overline{x}\} \\
&\overset{A\in\mathbb{R}^{n\times n}}{=} \{\overline{x} \in \mathbb{C}^n \mid Ax = \lambda^* x\} \\
&= \{\overline{x} \in \mathbb{C}^n \mid x \in \ker(A - \lambda^* \operatorname{id}_{\mathbb{C}^n})\}.
\end{aligned}$$

Wegen $\{i \mid \lambda_i = \lambda^*\} = \{i_1 - 1, ...i_r - 1\} \subseteq \{k+1, k+3, ..., n-1\}$ ist $\operatorname{span}(w_{i_1-1}, ...w_{i_r-1}) = \ker(g - \lambda^* \operatorname{id}_{\mathbb{C}^n})$ und mit $(u_1, ..., u_r) := (\overline{w_{i_1-1}}, ...\overline{w_{i_r-1}})$ daher $\operatorname{span}(u_1, ...u_r) = \operatorname{span}(g - \lambda \operatorname{id}_{\mathbb{C}^n}) = \operatorname{span}(w_{i_1}, ..., w_{i_r})$. Es ist $(u_1, ..., u_r)$ auch ein ONS und damit eine ONB von $\operatorname{span}(w_{i_1}, ..., w_{i_r})$.

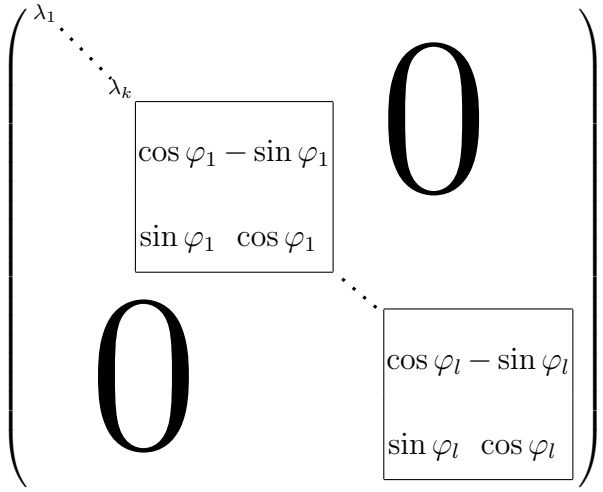
Schritt 3. Für $v_1, ..., v_n \in \mathbb{R}^n$ definiert durch $v_1 = w_1, ...v_k = w_k, \sqrt{2}w_{k+1} = v_{k+1} + \mathring{i}v_{k+2}, \sqrt{2}w_{k+3} = v_{k+3} + \mathring{i}v_{k+4}, ..., \sqrt{2}w_{n-1} = v_{n-1} + \mathring{i}v_n$ ist $\underline{v} := (v_1, ..., v_n)$ eines ONB des $\mathbb{R}^n$ mit $M(f, \underline{v})$ von der gewünschten Gestalt.
Begründung. Mit Lemma **??** sieht man leicht, dass $\underline{v}$ eine ONB des $\mathbb{C}^n$ und damit auch des $\mathbb{R}^n$ ist. Dass $M(f, \underline{v})$ von der gewünschten Gestalt ist, folgt leicht wie folgt. Sei $j \in \{k+1, k+3, ...\}$ und schreibe $\lambda_j = a_j + \mathring{i}b_j$ mit $a_j, b_j \in \mathbb{R}$. Dann gilt

$$\begin{aligned} f(v_j) + \mathring{i}f(v_{j+1}) &= g(v_j) + \mathring{i}g(v_{j+1}) = g(v_j + \mathring{i}v_{j+1}) \\ &= g(\sqrt{2}w_j) = \sqrt{2}\lambda_j w_j = \lambda_j(v_j + \mathring{i}v_{j+1}) \\ &= (a + \mathring{i}b)(v_j + \mathring{i}v_{j+1}) = av_j - b_j v_{j+1} + \mathring{i}(a_j v_{j+1} + b_j v_j). \end{aligned}$$

Satz 15.2.11. [→**??**] Es gelte der Fundamentalsatz der Algebra. Sei V ein endlichdimensionaler $\mathbb{R}$-Vektorraum mit Skalarprodukt und $f : V \to V$ linear. Es sind äquivalent:

(a) f ist orthogonal [→**??**].

(b) Es gibt eine ONB $\underline{v}$ von V derart, dass $M(f, \underline{v})$ von der Gestalt



Drehmatrizen
[→**??**(a),**??**]

mit $\lambda_1, ...\lambda_k \in \{-1, 1\}, \varphi_1, ...\varphi_l \in \mathbb{R}$

ist.

Beweis. **(b)** $\Longrightarrow$ **(a)** ist einfach.

(a) $\Longrightarrow$ **(b)**. Es reicht zu zeigen, dass es für $a, b \in \mathbb{R}$ mit $\left(\begin{smallmatrix} a & b \\ -b & a \end{smallmatrix}\right)^* \left(\begin{smallmatrix} a & b \\ -b & a \end{smallmatrix}\right) = I_2$ ein $\varphi \in \mathbb{R}$ gibt mit $\left(\begin{smallmatrix} a & b \\ -b & a \end{smallmatrix}\right) = \left(\begin{smallmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{smallmatrix}\right)$. Seien also $a, b \in \mathbb{R}$ derart, dass sie vorige

Bedingung erfüülen. Das heißt $a^2 + (-b)^2 = a^2 + b^2 = 1$. Dann gibt es $\varphi \in \mathbb{R}$ mit $\binom{a}{-b} = \binom{\cos\varphi}{\sin\varphi}$ und somit $\begin{pmatrix} a & b \\ -b & a \end{pmatrix} = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}$.

Definition 15.2.12. [→??, ??] Eine Matrix $A \in \mathbb{K}^{n\times n}$ heißt normal, wenn f_A normal ist.

Proposition 15.2.13. [→??, ??] Sei $A \in \mathbb{K}^{n\times n}$. Dann gilt

$$A \text{ normal} \iff A^*A = AA^*.$$

Beweis. Es gilt

$$\begin{aligned} A \text{ normal} &\iff f_A \text{ normal} \\ &\iff f_A^* \circ f_A = f_A \circ f_A^* \\ &\iff M(f_A^* \circ f_A, \underline{e}) = M(f_A \circ f_A^*, \underline{e}) \\ &\iff M(f_A^*, \underline{e})M(f_A, \underline{e}) = M(f_A, \underline{e})M(f_A^*, \underline{e}) \\ &\overset{??}{\iff} M(f_A, \underline{e})^*M(f_A, \underline{e}) = M(f_A, \underline{e})M(f_A, \underline{e})^* \\ &\iff A^*A = AA^*. \end{aligned}$$

Proposition 15.2.14. [→??, ??] Seien V ein $\mathbb{K}$-VR mit Skalarprodukt und $\underline{v} = (v_1, ..., v_n)$ eine ONB. Sei $f : V \to V$ linear. Dann gilt

$$f \text{ ist normal} \iff M(f, \underline{v}) \text{ ist normal}$$

Beweis. Es gilt

$$\begin{aligned} f \text{ normal} \iff f \circ f^* = f^* \circ &\iff M(f \circ f^*, \underline{v}) = M(f^* \circ f, \underline{v}) \\ f &\iff M(f, \underline{v})M(f^*, \underline{v}) = M(f^*, \underline{v})M(f, \underline{v}) \\ &\overset{??}{\iff} M(f, \underline{v})M(f, \underline{v})^* = M(f, \underline{v})^*M(f, \underline{v}) \\ &\overset{??}{\iff} M(f_A, \underline{v}) \text{ ist normal.} \end{aligned}$$

Ähnlich wie man z.B. **??** in **??** übersetzen kann, kann man die Sätze in diesem Abschnitt matrizentheoretisch formulieren. Wir überlassen dies dem Leser.

§16 Teilbarkeit in kommutativen Ringen

In diesem Kapitel sei stets A ein kommutativer Ring [→**??**].

16.1 Teilerbeziehung und Ideale

Definition 16.1.1. Wir führen auf A die Relationen $\mid$ und $\mathrel{\widehat{=}}$ ein für durch

$$a \mid b :\Longleftrightarrow \exists c \in A : ac = b,\ (a, b \in A),$$

gesprochen

- $''a$ *teilt* b (in $A)''$
- $''a$ ist Teiler von b (in $A)''$
- $''b$ ist Vielfaches von a (in $A)''$,

sowie

$$a \mathrel{\widehat{=}} b :\Longleftrightarrow (a \mid b \wedge b \mid a),\ (a, b \in A),$$

gesprochen

- $''a$ ist assoziativ zu b (in $A)''$
- $''a$ und b sind assoziativ (zueinander) (in$A)''$.

Erinnerung 16.1.2. [→ **??**] Eine Untergruppe I der additiven Gruppe von A heißt Ideal von A, wenn $\forall a \in A : \forall b \in I : ab \in I$ [→**??**]. Für $a \in A$ ist $(a) := \{ca \in A\}$ das kleinste Ideal I von A mit $a \in I$, genannt das von a erzeugte (Haupt-)Ideal. Man nennt ein Ideal I von A ein Hauptideal, wenn es $a \in A$ gibt mit $I = (a)$ [→**??**].

Bemerkung 16.1.3. Seien $a, b \in A$. Dann gilt $a \mid b \Longleftrightarrow b \in (a) \Longleftrightarrow (b) \subseteq (a)$ und $a \mathrel{\widehat{=}} b \Longleftrightarrow (a) = (b)$ Insbesondere ist $\mathrel{\widehat{=}}$ eine Äquivalenzrelation [→**??**(b)] auf A und auf

der Quotientenmenge [→**??**(a)] $A/\widehat{=}$ können wir eine Halbordnung [→**??**] $\preceq$ definieren durch

$$\widehat{a} \preceq \widehat{b} :\Longleftrightarrow a \mid b,\ (a, b \in A),$$

wobei für jedes $a \in A$ mit $\widehat{a} := \widehat{\widehat{a}}$ die Äquivalenzklasse [→**??**(b),**??**] von a bezeichnet wird. In der Tat: Sind $a, a', b, b' \in A$ mit $\widehat{a} = \widehat{a'}$ und $\widehat{b} = \widehat{b'}$, so gilt, dass $(a) = (a')$ und $(b) = (b')$ und daher

$$a \mid b \Longleftrightarrow (b) \subseteq (a) \Longleftrightarrow (b') \subseteq (a') \Longleftrightarrow a' \mid b'.$$

Definition 16.1.4. Sei $B \subseteq A$ und $a \in A$. Es heißt a ein $\left\{\begin{array}{c} \textit{gemeinsamer Teiler (gT)} \\ \textit{gemeinsames Vielfaches (gV)} \end{array}\right\}$ der Elemente von B (in A), wenn $\widehat{a}$ eine $\left\{\begin{array}{c}\text{obere}\\ \text{untere}\end{array}\right\}$ Schranke von $\{\widehat{b} \mid b \in B\}$ in $(A/\widehat{=}, \preceq)$ ist. Es heißt a ein $\left\{\begin{array}{c} \textit{größten gemeinsamen Teiler (ggT)} \\ \textit{kleinsten gemeinsamen Vielfachen (kgV)} \end{array}\right\}$ der Elemente von B, wenn a $\left\{\begin{array}{c}\text{Supremum}\\ \text{Infimum}\end{array}\right\}$ von $\{\widehat{b} \mid b \in B\}$ in $(A/\widehat{=}, \preceq)$ ist.

Definition 16.1.5. Seien $(A_1, \preceq_2)$ und $(A_2, \preceq_2)$ halbgeordnete Mengen. Eine Abbildung $f : A_1 \to A_2$ heißt Isomorphismus halbgeordneter Mengen, wenn f bijektiv ist und

$$\forall a, b \in A_1 : (a \preceq_1 b \Longleftrightarrow f(a) \preceq_2 f(b)).$$

Bemerkung 16.1.6. Sei $B \subseteq A$ und $a \in A$.

(a) Es ist klar, wie man die Definition von gT, gV, ggT, kgV ohne die Verwendung von $\widehat{=}$ lesen kann:

- a gT der El. von B $\Longleftrightarrow \forall b \in B : a \mid b$
- a ggT der El. von $B \Longleftrightarrow \begin{pmatrix}\text{a gT der El. von B}\wedge \\ \forall b \in A : (\text{b gT der El. von B} \implies b \mid a)\end{pmatrix}$

(b) Wegen der Eindeutigkeit von Infima und Suprema in halbgeordneten Mengen [→**??**] sind ggT und kgV in kommutativen Ringen genau bis auf Assoziiertheit eindeutig, sofern sie existieren.

(c) Betrachte die durch Obermengeninklusion (umgekehrte Inklusion) halbgeordnete Menge aller Hauptideale $H := \{(a) \mid a \in A\}$. Nach Definition von $\widehat{=}$ ist die Abbildung $A/\widehat{=} \to H, \hat{a} \mapsto (a)$ wohldefiniert und injektiv. Offenbar ist sie auch surjektiv. Nach Definition von $\preceq$ ist sie ein Isomorphismus halbgeordneter Mengen. In der Definition von gT/ gV/ ggT/ kgV [→**??**] könnte man daher genauso (a) statt $\widehat{a}$, (b) statt $\widehat{b}$ und $(H, \supseteq)$ statt $(A/\widehat{=}, \preceq)$ schreiben.

Erinnerung 16.1.7. (a) Die Schnittmenge einer Menge von Idealen in A ist wieder ein Ideal von A [→**??**].

(b) Für jede Teilmenge $E \subseteq A$ gibt es ein kleinstes Ideal, welches E enthält [→**??**]. Man nennt es das von E erzeugte Ideal und notiert es mit $(E) = (E)_A$. Es besteht gerade aus allen Summen von Vielfachen von Elementen von E [→**??**].

Proposition 16.1.8. Sei $B \subseteq A$ und $a \in A$.

(a) a gT der El. von $B \iff B \subseteq (a) \iff (B) \subseteq (a)$

(b) a gV der El. von $B \iff a \in \bigcap\{(b) \mid b \in B\} \iff (a) \subseteq \bigcap\{(b) \mid b \in B\}$

(c) a ggT. der El. von $B \impliedby (a) = (B)$
und wenn (B) ein Hauptideal ist, gilt auch " $\implies$ ".

(d) a kgV. der El. von $B \iff (a) = \bigcap\{(b) \mid b \in B\}$

Beweis. (a) und (b) sind klar.

Zu (c).

$\impliedby$ Gelte $(a) = (B)$. Nach (a) ist a ein gT der Elemente von B. Sei b ein gT der Elemente von B. Zu zeigen ist $b \mid a$. Dann $a \in (a) = (B) \overset{(a)}{\subseteq} (b)$ und daher $b \mid a$.

$\implies$ Sei a ein ggT. der Elemente von B und $b \in A$ mit $(b) = (B)$. Nach (a) gilt $(B) \subseteq (a)$. Noch zu zeigen ist $(a) \subseteq (B)$. Da b ein gT. und a ein ggT. der Elemente von B ist, gilt $b \mid a$, also $(a) \subseteq (b) = (B)$.

Zu (d). Wegen (b) genügt es zu zeigen

$$(\forall c \in A : (\underbrace{c \text{ gV der El. von } B}_{\iff c \in \bigcap\{(b) \mid b \in B\}} \implies a \mid c)) \iff (a) \supseteq \bigcap\{(b) \mid b \in B\}.$$

Dies ist klar.

Bemerkung 16.1.9. In **??** haben wir gezeigt, dass in $\mathbb{Z}$ jedes Ideal ein Hauptideal ist. In **??** haben wir dasselbe für den Polynomring $K[X]$ über einem Körper K gezeigt. Nach **??**(c) läuft also in diesen Ringen das Bestimmen eines ggT auf die Berechnung eines Erzeugers eines Ideals hinaus. Im Polynomring $\mathbb{Z}[X]$ zum Beispiel ist diese jedoch nicht so, wie Beispiel **??** unten zeigt.

Sprechweise 16.1.10. Seien $b_1, ..., b_n \in A$. Wenn wir von einem $gT/gV/ggT/kgV$ von $b_1, ..., b_n$ shreiben, so meinen wir einen $gT/gV/ggT/kgV$ von $\{b_1, ..., b_n\}$. Wir nennen wir

$$(b_1, ..., b_n) := (\{b_1, ..., b_n\}) = \left\{ \sum_{i=1}^{n} a_i b_i \mid a_1, ...a_n \in A \right\} \ [\rightarrow \textbf{??}]$$

das von $b_1, ..., b_n$ erzeugte Ideal.

Beispiel 16.1.11. Die Teiler von 2 in $\mathbb{Z}[X]$ sind $-2, -1, 1, 2$. Die Teiler von X in $\mathbb{Z}[X]$ sind $-X, -1, 1, X$. Die gemeinsamen Teiler von 2 und X sind daher $-1, 1$. Die größten ggT von 2 und X sind ebenso $1, -1$. Allerdings gilt

$$(2, X) = \{2p + Xq \mid q, p \in \mathbb{Z}[X]\} = \{2a + Xp \mid a \in \mathbb{Z}, p \in \mathbb{Z}[X]\} \neq (1) = \mathbb{Z}[X].$$

Beispiel 16.1.12. Sei A ein kommutativer Ring.

(a) $\forall a \in A : (1 \mid a \wedge a \mid 0)$

(b) $\forall a \in A : (a \mid 1 \iff a \in A^\times)$,
wobei $A^\times = \{a \in A : \exists b \in A : ab = 1\}$ die Einheitengruppe von A ist [→**??**].

(c) $\forall a \in A : (0 \mid a \iff a = 0)$

(d) $\widehat{1} = A^\times$, $\widehat{0} = \{0\}$

(e) Wegen $(\emptyset) = (0)$ ist 0 ein ggT der Elemente von $\emptyset$. Wegen $\widehat{0} = \{0\}$ ist es der einzige.

(f) Wegen $(A) = A = (1)$ ist 1 ein ggT der Elemente von A. Wegen $\widehat{1} = A^\times$ sind diese ggT genau die Einheiten.

16.2 Integritäts- und Hauptidealringe

Definition 16.2.1. Ein kommutativer Ring A heißt Integritätsring (auch: Integritätsbereich), wenn $1 \neq 0$ in A und $\forall a, b \in A : (ab = 0 \implies (a = 0 \vee b = 0))$. Ein Integritätsring A heißt Hauptidealring (auch: Hauptidealbereich), wenn jedes Ideal von A ein Hauptideal ist.

Beispiel 16.2.2. (a) Jeder Unterring [→**??**,**??**] eines Körpers ist ein Integritätsring.

(b) $\mathbb{Z}$ ist ein Hauptidealring [→**??**].

(c) Für jeden Körper K ist $K[X]$ ein Hauptidealring [→**??**].

Proposition 16.2.3. (a) Sei A ein Integritätsring. Dann gilt die Kürzungsregel

$$\forall a, b, c \in A : ((ac = ab \wedge c \neq 0) \implies a = b).$$

(b) Sei A ein Hauptidealring und $B \subseteq A$. Dann gibt es einen ggT und ein kgV der Elemente von B.

Beweis. **Zu (a).** Seien $a, b, c \in A$ mit $ac = ab$ mit $c \neq 0$. Dann gilt $(a-b)c = ac-bc0$ und daher $a - b = 0$ oder $c = 0$. Wegen $c \neq 0$ folgt $a = b$.

(b) folgt direkt aus **??**(c), da die Ideale (B) und $\bigcap\{(b) \mid b \in B\}$ Hauptideale sind.

Proposition 16.2.4. Sei A ein Integritätsring und $a, b \in A$. Dann gilt

$$a \mathrel{\widehat{=}} b \iff \exists c \in A^\times : a = bc.$$

Beweis. $\Longrightarrow$ Gelte $a \mathrel{\widehat{=}} b$, also $(a) = (b)$. Dann gibt es $c, d \in A$ mit $a = bc$ und $b = ad$. Es folgt $b = bcd$. Ist $b = 0$, so $(a) = (0)$ und daher $a = 0 = 0 \cdot 1 = b \cdot 1$ (und $1 \in A^\times$). Sei also $b \neq 0$. Dann $1 = cd$ wegen **??**(a) und daher $c \in A^\times$.

$\Longleftarrow$ Sei $c \in A^\times$ und $a = bc$. Dann $b \mid a$. Wegen $c^{-1}a = b$ aber auch $a \mid b$. Daher ist $a \mathrel{\widehat{=}} b$.

Beispiel 16.2.5. (a) $\mathbb{N}_0 \to \mathbb{Z}/\mathrel{\widehat{=}}, n \mapsto \widehat{n}$ ist eine Bij, denn $\mathbb{Z}^\times = \{-1, 1\}$ und $\mathbb{N}_0 \geq 0$.

(b) Sei K ein Körper. Dann ist $\{p \in K[X] : p \text{ normiert oder Null}\} \to K[X]/\mathrel{\widehat{=}}, p \mapsto \widehat{p}$ ist eine Bijektion, denn $K[X]^\times = K^\times$ [→**??**] und der Leitkoeffizient ist immer eindeutig bestimmt.

Bemerkung 16.2.6. Aus **??**(b)(c) folgt nun:

Zu jeder Menge $B \subseteq \mathbb{Z}$ gibt es genau $\left\{\begin{matrix}\text{einen ggT}\\ \text{ein kgV}\end{matrix}\right\}$ $a \in \mathbb{N}_0$ der Elemente von B in $\mathbb{Z}$, oft notiert mit $\left\{\begin{matrix}\gcd(B)\\ \operatorname{lcm}(B)\end{matrix}\right\}$.

Sei K ein Körper. ZU jeder Menge $B \subseteq K[X]$ gibt es genau $\left\{\begin{matrix}\text{einen ggT}\\ \text{ein kgV}\end{matrix}\right\}$ $p \in K[X]$ der Elemente von B in $K[X]$ mit $\deg p < 0$ oder p normiert, oft notiert mit $\left\{\begin{matrix}\gcd(B)\\ \operatorname{lcm}(B)\end{matrix}\right\}$.

Beispiel 16.2.7. (a) $\operatorname{lcm}(\emptyset) = 1$ und $\gcd(\emptyset) = 0$ sowohl in $\mathbb{Z}$ als auch in $K[X]$ für einen Körper K.

(b) $\operatorname{lcm}(\{-5\}) = 5$ in $\mathbb{Z}$ und $\operatorname{lcm}(\{-5\}) = 1$ in $\mathbb{R}[X]$.

(c) $\operatorname{lcm}(\{5, 3, 2, 6\}) = 30$ in $\mathbb{Z}$, denn $(5) \cap (3) \cap (2) \cap (6) = 30$.

(d) $\operatorname{lcm}(\{2X^2 - 2, X^2 - 2X + 1\}) = (X - 1)^2(X + 1)$ in $\mathbb{R}[X]$, denn

$$\begin{aligned}(2X^2 - 2) \cap (X^2 - 2X + 1) &= (X^2 - 1) \cap ((X - 1)^2)\\ &= ((X - 1(X + 1)) \cap ((X - 1)^2)\\ &\overset{??}{=} ((X + 1)(X - 1)^2).\end{aligned}$$

(e) $\gcd(\{153, 204, -357, 0\}) = 51$ in $\mathbb{Z}$, denn

$$\begin{aligned}(153, 204, -357, 0) &= (153, 204, 357) = (153, 204, 357 - 153)\\ &= (153, 204, 204) = (153, 204 - 153) = (153, 51)\\ &= (153 - 51, 51) = (102, 51) = (51).\end{aligned}$$

(f) $\gcd(\{2X^2 - 2, X^2 - 2X + 1\}) = X - 1$ in $\mathbb{Q}[X]$, denn

$$\begin{aligned}(2X^2 - 2, X^2 - 2X + 1) &= (X^2 - 1, X^2 - 2X + 1 - 2X + 2) = (X^2 - 1, X - 1)\\ &= (X - 1).\end{aligned}$$

(g) $\operatorname{lcm}(\{2^m \mid m \in \mathbb{N}^2\}) = 0$ in $\mathbb{Z}$.

16.3 Zur Berechnung größter gemeinsamer Teiler

Gegeben seien $b_1, ...b_n \in A$. Um die gT von $b_1, ..., b_n$ zu bestimmen oder sogar einen ggT von $b_1, ..., b_n$ zu bestimmen, falls er existiert, ist es generell eine gute Strategie "einfachere" $c_1, ...c_m \in A$ zu suchen mit $(b_1, ..., b_n) = (c_1, ..., c_m)$. Gilt nämlich diese Gleichung, so haben $b_1, ..., b_n$ und $c_1, ..., c_m$ jeweils dieselben gT nach **??**(a). Existiert sogar ein $c \in A$ mit $(b_1, ..., b_n) = (c)$, was in Hauptidealringen immer der Fall ist, so ist c nach **??**(a) ein ggT von $b_1, ..., b_n$. Wir werden sehen, wie man in $\mathbb{Z}$ und in $K(X]$ (modulo dem Problem, die Rechenoperationen im Körper K durchzuführen). Ein solches c und damit $\gcd\{b_1, ..., b_n\}$ berechnen kann. Weiter werden wir sehen, wie man in $\mathbb{Z}$ und $K[X]$, allerdings mit erheblich mehr Aufwand, sogar $a_1, ..., a_n$ berechnen kann mit $a_1b_1 + ... + a_nb_n = \gcd\{b_1, ..., b_n\}$.

Satz 16.3.1. Seien $b_1, ..., b_n \in A$, $i \in \{1, ..., i\}$ und sei $d \in A$ derart, dass $\overline{b}_i \,\widehat{=}\, \overline{d}$ in $A/(b_1, ...b_{i-1}, b_{i+1}, ..., b_n)$. Dann gilt $(b_1, ..., b_n) = (b_1, ...b_{i-1}, d, b_{i+1}, ..., b_n)$.

Beweis. Da man in der Behauptung b_i und d vertauschen kann, reicht es "$\subseteq$" zu zeigen. Dafür reicht es, $b_i \in (b_1, ...b_{i-1}, d, b_{i+1}, ..., b_n)$ zu zeigen. Wegen $(\overline{b_i}) = (\overline{d})$ gibt es ein $e \in A$ mit $\overline{b_i} = \overline{e}\overline{d}$ in $A/(b_1, ...b_{i-1}, b_{i+1}, ..., b_n)$. Daher gilt $b_1 - ed \in (b_1, ...b_{i-1}, b_{i+1}, ..., b_n)$ und somit $b_i \in (b_1, ...b_{i-1}, d, b_{i+1}, ..., b_n)$.

Beispiel 16.3.2. (a) In $\mathbb{Z}$ gilt

$$\begin{aligned}(38\,321\,783\,292, 27, 45) &= (38\,321\,783\,292, 27, 18) = (38\,321\,783\,292, 9, 18)\\ &= (38\,321\,783\,292, 9)\\ &= (3 + 8 + 3 + 2 + 1 + 7 + 8 + 3 + 2 + 9 + 2, 9)\\ &= (3 + 8 + 3 + 2 + 1 + 7 + 8 + 3 + 2 + 9 + 2, 9)\\ &= (12, 9) = (3)\end{aligned}$$

und daher $\gcd\{(38\,321\,783\,292, 27, 45)\} = 3$.

(b) In $\mathbb{Z}$ gilt

$$\begin{aligned}(43\,733, 17\,473, 27\,977) &= (43\,733, 17\,473, 10\,504) = (43\,733, 303, 10\,504)\\ &= (1\,717, 303, 1\,414) = (303, 303, 1\,414)\\ &= (303, 1\,414) = (303, 202) = (101)\end{aligned}$$

und daher $\gcd\{(43\,733, 17\,473, 27\,977)\} = 101$.

(c) In $\mathbb{Z}$ gilt

$$\begin{aligned}
&(\underbrace{\underbrace{34}_{\equiv_{(24)}10}}_{\equiv_{(54)}000}\ 8\ \underbrace{39}_{\equiv_{(24)}15}\ 0\,4\ \underbrace{58}_{\equiv_{(54)}04}\ \underbrace{50}_{\equiv_{(24)}02}\ 2, 24, 54) = (\underbrace{15}_{\equiv_{6)}03}\,0\ \underbrace{40}_{\equiv_{(24,6)}04}\ \underbrace{4\,0}_{\equiv_{(24,6)}04}\ \underbrace{22}_{\equiv_{(4)}04}, 24, 6)\\
=&(\underbrace{30}_{\equiv_{(24,6)00}}\ 0\ \underbrace{40}_{\equiv_{(24)}04}\ \underbrace{40}_{\equiv_{(24)}04}\ 4, 24, 6) = (\underbrace{40}_{\equiv_{(24)04}}\ \underbrace{44}_{\equiv_{(24)}8}, 24, 6)\\
=&(\underbrace{408}_{\equiv_{(24)}04}, 24, 6) = (48, 24, 6) = (24, 6) = (6)
\end{aligned}$$

und daher $\gcd\{348\,390\,458\,502, 24, 54\} = 6$.

(d) In $\mathbb{Z}$ gilt

$$\begin{aligned}
(3^{30}-1, 3^{45}-1) &= (3^{30}-1, 3^{15}-1) = (1-1, 3^{15}-1)\\
&= (3^{15}-1)
\end{aligned}$$

und daher $\gcd\{3^{30}-1, 3^{45}-1\} = 3^{15}-1$.

(e) In $\mathbb{Q}[X]$ gilt

$$\begin{aligned}
&(-X^7+X^5-X^3+X, X^8-1) \overset{\overline{X}\in(\mathbb{Q}[X]/(X^8-1))^\times}{=} (-X^6+X^4-X^2+1, X^8-1)\\
=&(-X^6+X^4-X^2+1, -X^6+X^4-X^2+1) = (-X^6+X^4-X^2+1)
\end{aligned}$$

und daher $\gcd = \{-X^7+X^5-X^3+X, X^8-1\} = X^6-X^4+X^2-1$.

(f) In $\mathbb{Q}(X]$ gilt

$$\begin{aligned}
&(X^4-2X^3+2X^2-2X+1, \overbrace{4X^3-6X^2+4X-2}^{=p}, \overbrace{X^3+3X^2-9X+5}^{=q})\\
=&(-5X^3+11X^2-7X+1, p, q) = (8X^2-12X+4, -18X^2+40X-22, q)\\
=&(2X^2-3X+1, 13X-13, q) = (2-3+1, X-1, 1+3-9+5)\\
=&(X-1)
\end{aligned}$$

und daher $\gcd\{X^4-2X^3+2X^2-2X+1, p, q\} = X-1$.

Es sollte nun klar sein, dass man in $\mathbb{Z}$ und in $K[X]$ (sofern K ein Körper ist, in dem man zu rechnen weiß) durch mehrfache Anwendung von Satz **??** zu gegebenen Elemente $b_1, ...b_n$ stets $\gcd\{b_1, ..., b_n\}$ berechnen kann. Solange man nämlich nicht $(b_1, ..., b_n)$ noch nicht als Hauptideal geschrieben hat (also mindestens zwei Erzeuger $\neq 0$ hat), kann man die Summe der Absolutbeträge (für $\mathbb{Z}$) bzw. die Summe der Gerade (für $K[X]$) der Erzeuger $\neq 0$ in jedem Schritt verringern.

Will man zu gegebenen $b_1, ..., b_n \in A$ nicht nur, falls es existiert, ein $b \in A$ finden mit $(b_1, ..., b_n) = (b)$, sondern will man auch noch $a_1, ..., a_n$ mit $b = a_1 b_1 + ... + a_n b_n$ finden, so kann man das wie folgt versuchen: Man bildet die Matrix

$$M(q, \underline{v}) = \begin{pmatrix} b_1 & & \\ \vdots & & \\ b_n & & \end{pmatrix} = \begin{pmatrix} b_1 & 1 & & 0 \\ \vdots & & \ddots & \\ b_n & 0 & & 1 \end{pmatrix} \in A^{m \times (n+1)}$$

Die i-te Zeile dieser Matrix kann man als die (trivialeriwese) gültige Gleichung $b_i = 0 \cdot b_1 + 0 \cdot b_{i-1} + b_i + 0 \cdot b_{i+1} + ... + 0 \cdot b_{n+1}$ interpretieren. Nun überträgt man die vom Gauß-Verfahren aus §5.2 bekannten elementaren Zeilenoperationen von Matrizen über Körpern auf Matrizen über kommutativen Ringen:

- $\underbrace{Z_i}_{\text{„Zeile } i\text{“}} \underbrace{\leftarrow}_{\text{„wird“}} Z_i + \lambda Z_j \qquad (i, j \in \{1, \ldots, m\}, i \neq j, \lambda \in A)$
  (Addieren des λ-fachen einer Zeile zu einer anderen)

- $Z_i \leftarrow \lambda Z_i \qquad (i \in \{1, \ldots, m\}, \lambda \in A^\times)$
  (Multiplizieren einer Zeile mit einem $\lambda \neq 0$).

Nach Satz **??** ändert sich das von den Einträgen der ersten Spalte aufgespannte Ideal dabei nicht. Auch kann man die dabei entstehenden Zeilen jeweils als eine gültige Gleichung interpretieren.

Sowohl für $A = \mathbb{Z}$ als auch für $A = K[X]$ (K sei ein Körper) kann man durch endlich viele dieser Zeilenoperationen die Matrix $\begin{pmatrix} b_1 & \\ \vdots & I_n \\ b_n & \end{pmatrix}$ überführen in eine Matrix $\begin{pmatrix} \gcd\{b_1, ..., b_n\} & a_1 & \ldots & a_n \\ 0 & & & \\ \vdots & & * & \\ 0 & & & \end{pmatrix}$, deren erste Zeile als Gleichung gelesen besagt, dass $\gcd\{b_1, ..., b_n\} = a_1 b_1 + ... + a_n b_n$.

Beispiel 16.3.3. Durch Anwenden elementarer Zeilenoperationen über $\mathbb{Z}$ erhält man

$$\begin{pmatrix} 43\,733 & 1 & 0 & 0 \\ 17\,473 & 0 & 1 & 0 \\ 27\,977 & 0 & 0 & 1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 43\,733 & 1 & 0 & 0 \\ 17\,473 & 0 & 1 & 0 \\ 10\,504 & 0 & -1 & 1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1\,717 & 1 & 4 & -4 \\ 17\,473 & 0 & 1 & 0 \\ 10\,504 & 0 & -1 & 1 \end{pmatrix}$$
$$\rightsquigarrow \begin{pmatrix} 1\,717 & 1 & 4 & -4 \\ 303 & -10 & -39 & 40 \\ 10\,504 & 0 & -1 & 1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1\,717 & 1 & 4 & -4 \\ 303 & -10 & -39 & 40 \\ 1\,414 & 300 & 1\,169 & -1\,199 \end{pmatrix}$$
$$\rightsquigarrow \begin{pmatrix} 303 & * & * & * \\ 303 & -10 & -39 & 40 \\ 1\,414 & 300 & 1\,169 & -1\,199 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 303 & * & * & * \\ 303 & -10 & -39 & 40 \\ 202 & 340 & 1\,325 & -1\,359 \end{pmatrix}$$
$$\rightsquigarrow \begin{pmatrix} 303 & * & * & * \\ 101 & -350 & -1\,364 & 1\,399 \\ 202 & * & * & * \end{pmatrix} \rightsquigarrow \begin{pmatrix} 0 & * & * & * \\ 101 & -350 & -1\,364 & 1\,399 \\ 0 & * & * & * \end{pmatrix}.$$

Also ist $\gcd = \{43\,733, 17\,473, 27\,977\} = 101 = (-350) \cdot 43\,733 + (-1\,364) \cdot 17\,473 + 1\,399 \cdot 27\,977$.

16.4 Faktorielle Ringe

Definition 16.4.1. Sei $p \in A$. Es heißt p irreduzibel (in A), wenn gilt $p \notin A^\times \wedge (\forall a, b \in A : (p = ab \implies (a \in A^\times \vee b \in A^\times))$ und prim (in A) (auch: Primeleent von A), wenn $p \notin A^\times \wedge \forall a, b \in A : (p \mid ab \implies (p \mid a \vee p \mid b))$

Bemerkung 16.4.2. (a) 0 ist niemals irreduzibel, denn sonst erhielte man aus $0 = 0 \cdot 0$, dass $0 \in A^\times$ im Widerspruch zur angenommen Irreduzibilität.

(b) 0 ist prim in $A \iff A$ ist Integritätsring [→??]

Proposition 16.4.3. Sei A ein Integritätsring. Dann ist jedes Primelement $\neq 0$ von A irreduzibel.

Beweis. Sei $p \neq 0$ prim in A. Zu zeigen sind (a) $p \notin A^\times$ und (b) $\forall a, b \in A : (p = ab \implies (a \in A^\times \vee b \in A^\times)$.

(a) ist Teil der Definition eines Primelements.

Zu (b). Seien $a, b \in A$ mit $p = ab$. Wegen $p \mid ab$ gilt dann $p \mid a$ oder $p \mid b$. Œ gelte $p \mid a$, etwa $a = a'p$ mit $a' \in A$. Es folgt $p = ab = a'pb$ und daher $1 = a'b$ wegen $p \neq 0$. Somit ist $b \in A^\times$.

Die Voraussetzung, dass A Integritätsring ist, ist nicht überflüssig:

Beispiel 16.4.4. $\overline{2} = \overline{2} \cdot \overline{4}$ in $\mathbb{Z}/(6)$ und $\overline{2}, \overline{4} \notin (\mathbb{Z}/(6))^\times$. Daher ist $\overline{2}$ nicht irreduzibel in $\mathbb{Z}/(6)$. Es ist aber $\overline{2}$ ein Primelement $\neq 0$ in $\mathbb{Z}/(6)$, denn $\overline{2} \notin (\mathbb{Z}/(6))^\times$ und sind $a, b \in \mathbb{Z}$

mit $\overline{2} \mid \overline{a}\overline{b}$ in $\mathbb{Z}/(6)$, so ist a oder b eine gerade ganze Zahl und daher $\overline{2} \mid \overline{a}$ oder $\overline{2} \mid \overline{b}$ in $\mathbb{Z}/(6)$.

Erinnerung 16.4.5. [→??] Wir nennen $n \in \mathbb{N}$ mit $n \geq 2$ eine Primzahl, wenn es nicht $s, t \in \mathbb{N}$ mit $s, t \geq 2$ und $n = st$ gibt. Wir schreiben $\mathbb{P} := \{2, 3, 5, 7, 11, ...\}$ für die Menge der Primzahlen.

Proposition 16.4.6. Für $p \in \mathbb{Z}$ sind äquivalent:

(a) $p \in \mathbb{P}$

(b) p ist irreduzibel in $\mathbb{Z}$ und $p \geq 0$

(c) p ist prim in $\mathbb{Z}$ und $p > 0$

Beweis. $\underline{(a) \iff (b)}$ ist sehr einfach.

$\underline{(c) \implies (b)}$ folgt aus Proposition **??**

$\underline{(b) \implies (c)}$. Gelte $p \in \mathbb{P}$. Dann $p \notin \{1, -1\} \in \mathbb{Z}^\times$. Noch zu zeigen ist $\forall a, b \in \mathbb{Z} : (p = ab \implies (a \in A^\times \vee b \in A^\times)$. Mit HIlfe von $\mathbb{Z}/(p)$ kann man das anders schreiben als: $\forall a, b \in Z/(p) = (ab = 0 \implies (a = 0 \vee b =))$. Wegen $p \in \mathbb{P}$ ist aber nach **??** der kommutative Ring $\mathbb{Z}/(p)$ ein Körper, woraus dies sofort folgt.

Beispiel 16.4.7. Im Unterring $\mathbb{Z}[2\mathring{i}] = \{a + b2\mathring{i} \mid a, b \in \mathbb{Z}\}$ von $\mathbb{C}$ [→??,??] ist 2 irreduzibel aber nicht prim. Wegen $(2\mathring{i})(2\mathring{i}) = -4 = (-2)2$ gilt nämlich $2 \mid (2\mathring{i})(2\mathring{i})$ in $\mathbb{Z}[2\mathring{i}]$, aber offensichtlich gilt nicht $2 \mid 2\mathring{i}$ in $\mathbb{Z}[2\mathring{i}]$. Daher ist 2 nicht prim in $\mathbb{Z}[2\mathring{i}]$. Offensichtlich ist 2 keine Einheit in $\mathbb{Z}[2\mathring{i}]$. Um zu zeigen, dass 2 irreduzibel in $\mathbb{Z}[2\mathring{i}]$, reicht es schließlich $a, b, c, d \in \mathbb{Z}$ zu betrachten mit $2 = (a + 2b\mathring{i})(c + 2d\mathring{i})$. Zu zeigen ist $a + 2b\mathring{i} \in \mathbb{Z}[2\mathring{i}]^\times$ oder $c + 2d\mathring{i} \in \mathbb{Z}[2\mathring{i}]^\times$. Es gilt

$$\begin{aligned} 4 = 2 \cdot 2^* &= (a + 2b\mathring{i})(c + 2d\mathring{i}) \cdot (a - 2b\mathring{i})(c - 2d\mathring{i}) \\ &= (a^2 + 4b^2)(c^2 + 4d^2). \end{aligned}$$

Da $a^2 + 4b^2 \neq 2$ und $c^2 + 4d^2 \neq 2$, folgt $a^2 + 4b^2 = 1$ oder $c^2 + 4d^2 = 1$, also $(a \in \{1, -1\} \wedge b = 0)$ oder $(c \in \{1, -1\} \wedge d = 0)$.

Notation 16.4.8. Im Folgenden fixieren wir eine Menge $\mathbb{P}_A$ von Primelementen $\neq 0$ in A derart, dass jedes Primelement $\neq 0$ in A zu genau einem Element von $\mathbb{P}_A$ assoziiert ist. Es enthält $\mathbb{P}_A$ also aus jedem $\widehat{p}$ mit $p \in A \setminus \{0\}$ mit p prim genau einen Vertreter.

Beispiel 16.4.9. (a) In der Regel wird man $\mathbb{P}_\mathbb{Z} = \mathbb{P}$ nehmen. Man könnte aber auch $\mathbb{P}_\mathbb{Z} := \{2, -3, 5, 7, -11, 13, ...\}$.

(b) Ist K ein Körper, so nimmt man in der Regel $\mathbb{P}_{K[X]} := \{p \in K[X] \mid p \neq 0, p \text{ normiert}, p \text{ prim in } K[X]\}$.

Definition 16.4.10. Es bezeichne $\mathbb{N}_0^{(\mathbb{P}_A)}$ die Menge der Funktionen $\alpha : \mathbb{P}_A \to \mathbb{N}_0$ mit endlichem Träger $\operatorname{supp}(\alpha) := \{p \in \mathbb{P}_A \mid \alpha(p) \neq 0\}$. Für jedes $\alpha \in \mathbb{N}_0^{(\mathbb{P}_A)}$ setzen wir $\mathbb{P}_A^\alpha := \prod_{p \in \operatorname{supp}(\alpha)} p^{\alpha(p)}$. Wir nennen $(c, \alpha) \in A^\times \times \mathbb{N}_0^{(\mathbb{P}_A)}$ eine Primfaktorzerlegung von $a \in A$, wenn $a = c\mathbb{P}_A^\alpha$.

Beispiel 16.4.11. Mit $\mathbb{P}_Z = \mathbb{P}$ ist $(-1, \alpha)$ eine Primfaktorzerlegung von -20 in $\mathbb{Z}$, wenn man $\alpha \in \mathbb{N}_0^{(\mathbb{P}_A)}$ definiert durch $\operatorname{supp}(\alpha) = \{2, 5\}$, $\alpha(2) = 2$ und $\alpha(5) = 1$. Informell würde man sagen, $-20 = (-1)2^2 5$ eine Primfaktorzerlegung von 20.

Lemma 16.4.12. Sei A ein Integritätsring und $p, q \in \mathbb{P}_A$ mit $p \mid q$. Dann gilt $p = q$.

Beweis. Schreibe $p = qa$ mit $a \in A$. Wegen $q \mid pa$ gilt $q \mid p$ oder $q \mid a$. Falls $q \mid p$, so gilt $p \mathrel{\widehat{=}} q$ und $p = q$. Es also, die Annahme $q \mid a$ zum Widerspruch zu führen. Wäre aber $b \in A$ mit $a = qb$, so folgt $q = pa = pqb$ und daher $1 = pb$ wegen $q \neq 0$. Dann wäre aber $p \in A^\times$ und damit p nicht prim.

Proposition 16.4.13. In Integritätsringen sind Primfaktorzerlegungen eindeutig, das heißt ist A ein Integritätsring und sind $(c, \alpha), (d, \beta) \in A^\times \times \mathbb{N}_0^{(\mathbb{P}_A)}$ mit $c\mathbb{P}_A^\alpha = d\mathbb{P}_A^\beta$. So folgt $(c, \alpha) = (d, \beta)$.

Beweis. Der Beweis erfolgt durch Induktion nach der Anzahl der Primfaktoren in der ersten Primfaktorzerlegung, das heißt wir zeigen:

$$\forall (c, \alpha), (d, \beta) \in A^\times \times \mathbb{N}_0^{(\mathbb{P}_A)} : \left(\left(\sum_{p \in \operatorname{supp}(\alpha)} \alpha(p) = n \right) \wedge c\mathbb{P}_A^\alpha = d\mathbb{P}_A^\beta \right) \\ \implies (c, \alpha) = (d, \beta)$$

durch Induktion nach $n \in \mathbb{N}_0$.

$\underline{n = 0}$ Seien $(c, \alpha), (d, \beta) \in A^\times \times \mathbb{N}_0^{(\mathbb{P}_A)}$ mit $\alpha = 0$ und $c\mathbb{P}_A^\alpha = d\mathbb{P}_A^\beta$. Dann $\mathbb{P}_A^\beta = cd^{-1} \in A^\times$. Da kein Primelement eine Einheit ist, folgt $\beta = 0$ und $c = d$.

$\underline{n - 1 \to n \ (n \in \mathbb{N})}$ Seien $(c, \alpha), (d, \beta) \in A^\times \times \mathbb{N}_0^{(\mathbb{P}_A)}$ MIT $\sum_{p \in \operatorname{supp}(\alpha)} \alpha(p) = n$ und $c\mathbb{P}_A^\alpha = d\mathbb{P}_A^\beta$. Wähle $p \in \operatorname{supp}(\alpha)$. Wegen $p \mid \mathbb{P}_A^\beta$ gibt es $q \in \operatorname{supp}(\beta)$ mit $p \mid q$, denn p ist prim. Gemäß Lemma **??** gilt $p = q$. Definiere $\alpha', \beta' \in \mathbb{N}_0^{(\mathbb{P}_A)}$ durch $\alpha'|_{\mathbb{P}_A \setminus \{p\}} = \alpha|_{\mathbb{P}_A \setminus \{p\}}, \beta'|_{\mathbb{P}_A \setminus \{p\}} = \beta|_{\mathbb{P}_A \setminus \{p\}}$, $\alpha'(p) = \alpha(p) - 1$ und $\beta'(p) = \beta(p) - 1$. Es folgt $c\mathbb{P}_A^{\alpha'} = d\mathbb{P}_A^{\beta'}$ und nach der Induktionsvoraussetzung daher $(c, \alpha') = (d, \beta')$. Somit gilt auch $(c, \alpha) = (d, \beta)$.

Definition 16.4.14. Ein Integritätsring A heißt faktoriell, wenn jedes $a \in A \setminus \{0\}$ eine Primfaktorzerlegung in A besitzt, das heißt es gibt $(c, \alpha) \in A^\times \times \mathbb{N}_0^{(\mathbb{P}_A)}$ mit $a = c\mathbb{P}_A^\alpha$.

Bemerkung 16.4.15. Sei A ein Integritätsring.

(a) Definition **??** ist wegen Proposition **??** unabhängig von der Wahl von $\mathbb{P}_A$.

(b) Proposition **??** besagt, dass $A^\times \times \mathbb{N}_0^{(\mathbb{P}_A)} \to A \setminus \{0\}, (c, \alpha) \mapsto c\mathbb{P}_A^\alpha$ injektiv ist und gemäß Definition **??** ist A genau dann faktoriell, wenn diese Abbildung auch surjektiv und damit bijektiv ist.

Notation 16.4.16. Wir führen auf $\mathbb{N}_0^{(\mathbb{P}_A)}$ die Halbordnung $\preceq$ definiert durch

$$\alpha \preceq \beta :\Longleftrightarrow (\forall p \in \mathbb{P}_A : \alpha(p) \leq \beta(p)) \ (\alpha, \beta \in \mathbb{N}_0^{(\mathbb{P}_A)})$$

ein.

Proposition 16.4.17. Sei A ein faktorieller Ring. Dann gilt für alle $(c, \alpha), (d, \beta) \in A^\times \times \mathbb{N}_0^{(\mathbb{P}_A)}$, dass $c\mathbb{P}_A^\alpha \mid d\mathbb{P}_A^\beta \Longleftrightarrow \alpha \preceq \beta$.

Beweis. $''\Longleftarrow''$ Ist $\alpha \preceq \beta$, so ist $\gamma := \beta - \alpha \in \mathbb{N}_0^{(\mathbb{P}_A)}$ und $d\mathbb{P}_A^\beta = (\frac{d}{c}\mathbb{P}_A^\gamma)(c\mathbb{P}_A^\alpha)$.

$''\Longrightarrow''$ Gelte $c\mathbb{P}_A^\alpha \cdot a = d\mathbb{P}_A^\beta$ für ein $a \in A$. Zu zeigen ist $\alpha \preceq \beta$. Da A faktoriell und $a \neq 0$ ist, gibt es $(e, \gamma) \in A^\times \times \mathbb{N}_0^{(\mathbb{P}_A)}$ mit $a = e\mathbb{P}_A^\gamma$. Es folgt $ce\mathbb{P}_A^{\alpha+\gamma} = d\mathbb{P}_A^\beta$ und wegen **??** daher $\alpha + \gamma = \beta$ (und $ce = d$).

Satz 16.4.18. Sei A ein Integritätsring. Dann ist A faktoriell genau dann, wenn in A jeder irreduzible Element prim ist und die folgende "Teilerkettenbedingung": Ist $(a_n)_{n\in\mathbb{N}}$ eine Folge der Elemente $a_n \in A$ mit $(a_1) \subseteq (a_2) \subseteq (a_3) \subseteq ...$, so gibt es $k \in \mathbb{N}$ mit $(a_k) = (a_{k+1}) = (a_{k+2}) = ...$

Beweis. Sei zunächst A faktoriell. Da ein irreduzibles Element per Definition keine Einheit ist, muss in seiner Primfaktorzerlegung mindestens ein Primfaktor auftauchen. Da Primelemente keine Einheiten sind, kann aber dort auch höchstens ein Primfaktor auftreten. Sei nun $(a_n)_{n\in\mathbb{N}}$ eine Folge mit $(a_1) \subseteq (a_2) \subseteq (a_3) \subseteq ....$ Œ $a_1 \neq 0$ und damit $\forall n \in \mathbb{N} : a_n \neq 0$. Schreibe $a_n = c_n\mathbb{P}_A^{\alpha_n}$ mit $(c_n, \alpha_n), \in A^\times \times \mathbb{N}_0^{(\mathbb{P}_A)}$ für alle $n \in \mathbb{N}$. Mit Proposition **??** folgt $\alpha_1 \succeq \alpha_2 \succeq \alpha_3 \succeq ....$ Weil $\mathrm{supp}(\alpha_1)$ endlich ist, folgt hieraus, dass es $k \in \mathbb{N}$ gibt mit $\alpha_k = \alpha_{k+1} = ...$ Dann ist $(a_k) = (a_{k+1}) = ...$

Sei nun umgekehrt jede irreduzible Element prim und gelte die Teilerkettenbedingung. Es reicht dann zu zeigen, dass jedes $a \in A \setminus \{a\}$ ein Produkt von irreduziblen Elementen und einer Einheit ist. Sei M die durch Inklusion halbgeordnete Menge von Hauptidealen (a) derart, dass $a \in A \setminus \{a\}$ kein Produkt von irreduziblen Elementen und keiner Einheit ist. Es reicht, die Annahme $M \neq \emptyset$ zum Widerspruch zu führen. Wegen der Teilerkettenbedingung besitzt M mindestens ein maximales Element [→**??**](a) mit $a \in A \setminus \{a\}$, welches kein Produkt der gewünschten Form ist. Insbesondere ist a weder irreduzibel noch eine Einheit, weswegen es $b, c \in A \setminus A^\times$ mit $a = bc$. Es folgt $a \subset (b)$ und $(a) \subset (b)$ (wäre etwa $(a) = (b)$; das heißt $a \mathrel{\widehat{=}} b$,

so gäbe es nach Proposition **??** ein $c \in A^\times$ mit $a = bc'$ und es folgt $c = c' \in A^\times$ wegen $b \neq 0$). Wegen der Maximalität von (a) sind sowohl b als auch c Produkte von irreduziblen Elementen und einer Einheit, dann aber auch a. Dies ist ein Widerspruch.

Korollar 16.4.19. Jeder Hauptidealring ist faktoriell.

Beweis. Sei A ein Hauptideal. Zu zeigen ist (a) ist jedes irreduzible Element prim und (b) die Teilerkettenbedingung.

Zu (a). Sei $p \in A$ irreduzibel. Zu zeigen ist, dass p prim ist. Per Definition ist $p \neq A\times$. Seien $a, b \in A$ mit $p \mid ab$. Zu zeigen ist $p \mid a$ oder $p \mid b$. Wähle $c \in A$ mit $(c) = (p, a)$. Wegen $p \in (c)$ und der Irreduzibilität von p folgt $c \in A^\times$ oder $p \mathrel{\widehat{=}} c$. Fall $p \mathrel{\widehat{=}} c$, so gilt $p \mid c$ und $c \mid a$, also $p \mid a$. Gelte also $c \in A^\times$. Dann $(1) = (p, a)$ und es gibt $s, t \in A$ mit $1 = sp + ta$. Dann $b = sbp + tab$ und mit $p \mid (sbp + tab)$ (wegen $p \mid ab$) folgt $p \mid b$.

Zu (b). Sei $(a_n)_{n\in\mathbb{N}}$ eine Folge in A mit $(a_1) \subseteq (a_2) \subseteq (a_3) \subseteq ....$ Dann ist $I := \bigcup\{(a_n) \mid n \in \mathbb{N}\}$ ein Ideal von A, wie man sich leicht überlegt. Wähle $a \in A$ mit $I = (a)$. Wähle $k \in \mathbb{N}$ mit $a \in (a_k)$. Dann $(a_k) \subseteq (a_{k+1}) \subseteq ...I \subseteq (a_k)$, also $(a_k) = (a_{k+1}) = ...$

Beispiel 16.4.20. (a) $\mathbb{Z}$ ist faktoriell: Für alle $a \in \mathbb{Z} \setminus \{0\}$ gibt es genau ein $(c, \alpha) \in \{-1, 1\} \times \mathbb{N}_0^{(\mathbb{P}_A)}$ mit $a = c\mathbb{P}^\alpha$.

(b) Sei K ein Körper. Dann ist $K[X]$ faktoriell. Insbesondere sind die Primelemente $p \in K[X] \setminus \{0\}$ genau die irreduziblen Polynome in $K[X]$. Mann setzt in Übereinstimmung mit **??**(b)

$$\mathbb{P}_{K[X]} := \{p \in K[X] \mid p \text{ normiert und irreduzibel}\},$$

sofern nichts anderes ewähnt wird. Es gelte der Fundamentalsatz der Algebra. Dann gilt $\mathbb{P}_{\mathbb{C}[X]} = \{X - z \mid z \in \mathbb{C}\}$ und $\mathbb{P}_{\mathbb{R}[X)} = \{Xa \mid a \in \mathbb{R}\} \cup \{(X + a)^2 + c \mid a, b \in R, c > 0\}$. Letzteres folgt, da in der Primfaktorzerlegung eines $p \in \mathbb{R}[X]$ in $\mathbb{C}[X]$ mit jedem $X + a + \mathring{i}b \in \mathbb{P}_{\mathbb{C}[X]}$ $(a, b \in \mathbb{R}, b \neq 0)$ auch $X + a - \mathring{i}b \in \mathbb{P}_{\mathbb{C}[X]}$ auftaucht und mit $c := b^2 > 0$ gilt $(X + a + \mathring{i}b)(X + a - \mathring{i}b) = (X + a)^2 + c$ (vergleiche auch **??** und den Beweis von **??**).

(c) Sei K ein Körper und $p \in K[X] \setminus \{0\}$ mit Primfaktorzerlegung $c\mathbb{P}^\alpha_{K[X]}$ $(c \in A^\times, \alpha \in \mathbb{N}_0^{(\mathbb{P}_A)})$. Ist $\lambda \in K$, so ist λ eine Nullstelle von p, genau dann, wenn $\alpha(X - \lambda) \geq 1$. In diesem Fall ist $\alpha(X - \lambda)$ die in **??** definierte Vielfachheit der Nullstelle λ von p.

Notation 16.4.21. Wir führen zu $\mathbb{N}_0^{(\mathbb{P}_A)}$ ein neues Elemente ∞ hinzu und erweiterten die Halbordnung $\preceq$ aus **??** auf $\mathbb{N}_0^{(\mathbb{P}_A)} \cup \{\infty\}$, indem wir festlegen, das ∞ das größte Element der Halbordnung wird:

$$\alpha \preceq \beta :\Longleftrightarrow (\beta = \infty \vee (\alpha \neq \infty \neq \beta \wedge \forall p \in \mathbb{P}_A : \alpha(p) \leq \beta(p))) \ (\alpha, \beta \in \mathbb{N}_0^{(\mathbb{P}_A)} \cup \{\infty\})$$

Proposition 16.4.22. Sei A faktoriell. Dann ist die Abbildung

$$\begin{aligned} \mathbb{N}_0^{(\mathbb{P}_A)} \cup \{\infty\} \to & \qquad A/\widehat{=} \\ \alpha \mapsto & \begin{cases} \widehat{0} & \text{falls } \alpha = \infty \\ \widehat{\mathbb{P}_A^\alpha} & \text{sonst} \end{cases} \end{aligned}$$

ein Isomorphismus halbgeordneter Mengen [→**??**,**??**,**??**].

Beweis. **surjektiv**. Sei $a \in A$. Ist $a = 0$, so ist $\widehat{a}$ Bild von ∞. Sei also $a \neq 0$. Schreibe $a = c\mathbb{P}_a^\alpha$ mit $c \in A^\times$ und $\alpha \mathbb{N}_0^{(\mathbb{P}_A)}$. Dann $a \mathrel{\widehat{=}} \mathbb{P}_a^\alpha$, also ist $\widehat{a} \mathrel{\widehat{=}} \widehat{\mathbb{P}_a^\alpha}$ bild von α.

injektiv. $\widehat{0}$ hat nur ∞ als Urbild. Sin $\alpha, \beta \in \mathbb{N}_0^{(\mathbb{P}_A)}$ mit $\widehat{\mathbb{P}_A^\alpha} = \widehat{\mathbb{P}_A^\beta}$, so $\widehat{\mathbb{P}_A^\alpha} \mathrel{\widehat{=}} \widehat{\mathbb{P}_A^\beta}$ und es gibt $c \in A^\times$ mit $c\widehat{\mathbb{P}_A^\alpha} = \widehat{\mathbb{P}_A^\beta}$. Aus **??** ergibt sich $\alpha = \beta$ (und $c = 1$).

Isomorphismus. ∞ ist das größte Element von $\mathbb{N}_0^{(\mathbb{P}_A)}$ und sein Bild $\widehat{0}$ das größte Element von $A/\widehat{=}$. Für $\alpha, \beta \in \mathbb{N}_0^{(\mathbb{P}_A)}$ gilt $\alpha \preceq \beta \Longleftrightarrow \mathbb{P}_A^\alpha \mid \mathbb{P}_A^\beta \Longleftrightarrow \widehat{\mathbb{P}_A^\alpha} \preceq \widehat{\mathbb{P}_A^\beta}$.

In Verallgemeinerung [→**??**] von Proposition **??**(b) halten wir fest:

Korollar 16.4.23. Ist A ein faktorieller Ring und $B \subseteq A$, so gibt es einen ggT und ein kgV der Elemente von B.

Beweis. Luat Definition **??** ist zu zeigen, dass in der halbgeordneten Menge $(A/\widehat{=}, \preceq)$ ein Infimum und ein Supremum besitzt. Nach der letzten Proposition reicht es, dasselbe für die halbgeordnete Menge $(\mathbb{N}_0^{(\mathbb{P}_A)} \cup \{\infty\}, \preceq)$ zu zeigen. Sei also $B \subseteq \mathbb{N}_0^{(\mathbb{P}_A)} \cup \{\infty\}$. Setze

$$\alpha_{ggT} := \begin{cases} \infty & \text{falls } B \subseteq \{\infty\} \\ \begin{cases} \mathbb{P}_A \to \mathbb{N}_0 \\ p \mapsto \min\{\beta(p) \mid \beta \in B \setminus \{\infty\}\} \end{cases} & \text{falls } B \subsetneq \{\infty\} \end{cases}.$$ Dann ist α_{ggT} nach **??**(c) das Infimum von B, denn es gilt

$$\forall \beta \in \mathbb{N}_0^{(\mathbb{P}_A)} \cup \{\infty\} : (\beta \text{ untere Schranke von } B \Longleftrightarrow \beta \preceq \alpha_{ggT}).$$

Setze weiter

$$\alpha_{kgV} := \begin{cases} 0 & \text{falls } B = \emptyset \\ \infty & \text{falls } B \text{ keine obere Schranke besitzt} \\ \begin{cases} \mathbb{P}_A \to \mathbb{N}_0 \\ p \mapsto \max\{\beta(p) \mid \beta \in B \setminus \{\infty\}\} \end{cases} & \text{sonst} \end{cases}.$$

Wieder nach **??**(c) ist α_{kgV} das Supremum von B, denn es gilt

$$\forall \beta \in \mathbb{N}_0^{(\mathbb{P}_A)} \cup \{\infty\} : (\beta \text{ obere Schranke von } B \iff \alpha_{kgV} \preceq \beta).$$

Beispiel 16.4.24. $\gcd\{3^{17}, 5^{13}, 3^{14}, 5^9, 7\} = 3^{14}5^9$ und $\operatorname{lcm}\{3^{17}, 5^{13}, 3^{14}, 5^9, 7\} = 3^{17}5^{13}$ in $\mathbb{Z}$.

Skript zur linearen Algebra I & II

§17 Normalformen von Matrizen

17.1 Existenz der Normalform von Smith

In diesem Abschnitt sei stets A ein Hauptidealring (z.B. $A = \mathbb{Z}$ oder $A = K[X]$ mit K als Körper). Wie in **??** fixieren wir wieder eine Menge $\mathbb{P}_A$ von Primelementen von A derart, dass $\mathbb{P}_A \to \{\widehat{p} \mid p \in A, p \text{ prim}, p \neq 0\}$ eine Bijektion ist (z.B. $\mathbb{P}_Z = \mathbb{P}$ oder $\mathbb{P}_{K[X]} = \{f \in K[X] \mid f \text{ normiert und irreduzibel}\}$). Wie in §16.3 (vergleiche aus §5.2) betrachten wir wieder die folgenden elementaren Zeilenoperationen auf Matrizen über A:

- $Z_i \leftarrow Z_i + \lambda Z_j \qquad (i \neq j, \lambda \in A)$
- $Z_i \leftarrow \lambda Z_i \qquad (\lambda \in A^\times)$

Für theoretische Zwecke erlauben wir erstmals noch eine dritte elementare Zeilenoperation:

- $\begin{pmatrix} Z_i \\ Z_j \end{pmatrix} \underbrace{\leftarrow}_{\text{„simultan“}} \begin{pmatrix} aZ_i + bZ_j \\ cZ_i + dZ_j \end{pmatrix} \qquad (i \neq j, a, b, c, d \in A \text{ mit } ad - bc = 1)$

$$\left[\begin{array}{c} \text{umkehrbar: Sind } a, b, c, d \in A \text{ mit } ab - cd = 1, \text{ so } da - (-b)(-c) = 1 \text{ und z.B.} \\ \begin{pmatrix} L_1 \\ L_2 \end{pmatrix} \overset{\binom{Z_1}{Z_2} \leftarrow \binom{aZ_1 + bZ_2}{cZ_1 + dZ_2}}{\sim} \begin{pmatrix} aL_1 + bL_2 \\ cL_1 + dL_2 \end{pmatrix} \overset{\binom{Z_1}{Z_2} \leftarrow \binom{dZ_1 - bZ_2}{-cZ_1 + aZ_2}}{\sim} \begin{pmatrix} L_1 \\ L_2 \end{pmatrix} \end{array} \right]$$

Man sieht aber gleich, dass diese im Fall $A = \mathbb{Z}$, $A = K[X]$ und $A = K$ (K sei ein Körper) durch die anderen beiden simuliert werden kann und somit in diesem Fällen nichts beisteuert. [Man kann nämlich in allen drei Fällen die Matrix $I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \in A^{2\times 2}$ durch Anwendung der ersten beiden Operationen in jede Matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in A^{2\times 2}$ mit $\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = 1$ überführen $\left(\text{und dementsprechend } \begin{pmatrix} L_1 \\ L_2 \end{pmatrix} = \begin{pmatrix} 1L_1 + 0L_2 \\ 0L_1 + 1L_2 \end{pmatrix} \text{ in } \begin{pmatrix} aL_1 + bL_2 \\ cL_1 + dL_2 \end{pmatrix}\right)$. Da alle Operationen umkehrbar sind, reicht es, sich zu überlegen, dass man $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ mit der Determinante 1 in I_2 durch die ersten beiden Operationen überführen kann. Im Fall

$A = K$ ist dies klar, da $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ invertierbar ist. In den Fällen $A = \mathbb{Z}$ und $A = K[X]$ kann man mit den ersten beiden Operationen $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ auf die Form $\begin{pmatrix} e_1 & h \\ 0 & e_2 \end{pmatrix}$ bringen (vergleiche §16.3). Da sich die Determinante dabei nur um eine Einheit ändert (siehe §9.1), bleibt die Determinante eine Einheit. Mit $e_1 e_2 = \det \begin{pmatrix} e_1 & h \\ 0 & e_2 \end{pmatrix} \in A^\times$ gilt aber auch $e_1, e_2 \in A^\times$ un daher Œ $e_1 = e_2 = 1$ und dann auch Œ $h = 0$, also $\begin{pmatrix} e_1 & h \\ 0 & e_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.]
Wie in §5.2 kann man natürlich auch wieder Zeilenoperationen erlauben, die man durch endlich viele obiger Zeilenoperationen simulieren kann, z.B. $Z_i \longleftrightarrow Z_j$. Wir erlauben nun zusätzlich auch noch die entsprechenden Spaltenoperationen:

- $S_i \leftarrow S_i + \lambda S_j \qquad (i \neq j, \lambda \in A)$
- $S_i \leftarrow \lambda S_i \qquad (\lambda \in A^\times)$
- $\begin{pmatrix} S_i \\ S_j \end{pmatrix} \underbrace{\leftarrow}_{\text{„simultan“}} \begin{pmatrix} aS_i + bS_j \\ cS_i + dS_j \end{pmatrix} \qquad (i \neq j, a, b, c, d \in A \text{ mit } ad - bc = 1)$

Letztere ist für $A = \mathbb{Z}$, $A = K[X]$ und $A = K$ (K sei ein Körper) wiederum überflüssig. Offenbar sind alle diese Operationen umkehrbar und verändern die Determinante nur um einen Faktor, der eine Einheit ist.
Wir führen eine Äquivalenzrelation $\sim$ auf $A^{m \times n}$ ein durch

$$M \sim N :\Longleftrightarrow N \text{ geht aus } M \text{ durch obige Zeilen- und Spaltenoperationen hervor}$$

für alle $M, N \in A^{m \times n}$.

Warnung 17.1.1. Dies ist eine andere Äquivalenzrelation als die in §5.2 betrachtete, welche man auch "Zeilenäquivalenz nennt". Es handelt sich hier um "Zeilenspaltenäquivalenz".

Definition 17.1.2. Seien $m, n \in \mathbb{N}_0$. Eine Matrix $S \in A^{m \times n}$ findet sich in Smithscher Normalform, wenn mit $l := \min\{m, n\}$ es $\alpha_1, ..., \alpha_l \in \mathbb{N}_0^{(\mathbb{P}_A)} \cup \{\infty\}$. mit $\alpha_1 \preceq \alpha_2 \preceq ... \preceq \alpha_l$ [→**??**] und

$$S = \left(\begin{array}{ccc} \mathbb{P}_A^{\alpha_1} & & 0 \\ & \ddots & \\ 0 & & \mathbb{P}_A^{\alpha_l} \\ \hline & & \end{array}\right) \qquad \text{falls } m \geq n$$

$$\text{bzw.} \qquad S = \left(\begin{array}{ccc|c} \mathbb{P}_A^{\alpha_1} & & 0 & \\ & \ddots & & \\ 0 & & \mathbb{P}_A^{\alpha_l} & \end{array}\right) \qquad \text{falls } m \leq n, \text{ wobei } \mathbb{P}_A^\infty := 0.$$

Satz 17.1.3. Zu jeder Matrix $M \in A^{m\times n}$ gibt es eine Matrix $S \in A^{m\times n}$ in Smithscher Normalform mit $M \sim S$.

Beweis. Kommt später.

Beispiel 17.1.4. (a) Sei $M := \begin{pmatrix} 2 & 1 & -3 & -1 \\ 1 & -1 & -3 & 1 \\ 4 & -4 & 0 & 16 \end{pmatrix} \in \mathbb{Z}^{3\times 4}$. Es gilt:

$$
\begin{aligned}
M &\overset{Z_1 \leftrightarrow Z_2}{\sim} \begin{pmatrix} 1 & -1 & -3 & 1 \\ 2 & 1 & -3 & -1 \\ 4 & -4 & 0 & 16 \end{pmatrix} \overset{\substack{Z_1 \leftarrow Z_2 - Z_1 \\ Z_3 \leftarrow Z_1 - 4Z_1}}{\sim} \begin{pmatrix} 1 & -1 & -3 & 1 \\ 0 & 3 & 3 & -3 \\ 0 & 0 & 12 & 12 \end{pmatrix} \\
&\overset{\substack{S_2 \leftarrow S_2 + S_1 \\ S_3 \leftarrow S_3 + 3S_1 \\ S_4 \leftarrow S_4 - S_1}}{\sim} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 3 & 3 & -3 \\ 0 & 0 & 12 & 12 \end{pmatrix} \overset{\substack{S_3 \leftarrow S_3 - S_1 \\ S_4 \leftarrow S_4 + S_3}}{\sim} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 12 & 12 \end{pmatrix} \\
&\overset{S_4 \leftarrow S_4 - S_3}{\sim} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 12 & 0 \end{pmatrix} := S
\end{aligned}
$$

Also $M \sim S$ und S ist in Smithscher Normalform.

(b) Sei $M := \begin{pmatrix} 3 & 3 & 7 \\ -2 & 5 & 7 \\ 2 & 0 & -2 \end{pmatrix} \in \mathbb{Z}^{3\times 3}$. Es gilt:

$$
\begin{aligned}
M &\overset{\substack{Z_1 \leftarrow Z_1 - Z_3 \\ Z_2 \leftarrow Z_2 + Z_3 \\ Z_3 \leftarrow Z_3 - 2Z_1}}{\sim} \begin{pmatrix} 1 & 3 & 9 \\ 0 & 5 & 0 \\ 0 & -6 & -20 \end{pmatrix} \overset{\substack{S_2 \leftarrow S_1 - 3S_1 \\ S_3 \leftarrow S_3 - 9S_1 \\ Z_2 \leftarrow Z_2 + Z_3 \\ Z_3 \leftarrow -Z_3}}{\sim} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & -20 \\ 0 & 6 & 20 \end{pmatrix} \\
&\overset{\substack{Z_3 \leftarrow Z_3 + 6Z_2 \\ Z_2 \leftarrow -Z_2}}{\sim} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 20 \\ 0 & 0 & -100 \end{pmatrix} \overset{\substack{S_3 \leftarrow S_3 - 20S_2 \\ Z_3 \leftarrow -Z_3}}{\sim} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 100 \end{pmatrix} := S
\end{aligned}
$$

(c) Sei $M := \begin{pmatrix} 1-X^2 & 1+X \\ 1-X & 1+X+X^2+X^3 \\ 1-X & 1+X \end{pmatrix} \in K[X]$ mit K als Körper mit $2 \neq 0$. Es

gilt:

$$M \overset{\substack{S_1\leftrightarrow S_2\\ Z_3\leftarrow Z_3-Z_1\\ Z_2\leftarrow Z_2-(X^2+1)Z_1}}{\sim} \begin{pmatrix} 1+X & 1-X^2 \\ 0 & X^4-X \\ 0 & X^2-X \end{pmatrix} \overset{Z_2\leftarrow Z_2-(X^2+X+1)Z_3}{\underset{Z_2\leftrightarrow Z_3}{\sim}} \begin{pmatrix} 1+X & 1-X^2 \\ 0 & X^2-X \\ 0 & 0 \end{pmatrix}$$
$$\overset{Z_1\leftarrow Z_1+Z_2}{\sim} \begin{pmatrix} 1+X & 1-X \\ 0 & X^2-X \\ 0 & 0 \end{pmatrix} \overset{S_1\leftarrow S_1+S_2}{\sim} \begin{pmatrix} 2 & 1-X \\ X^2-X & X^2-X \\ 0 & 0 \end{pmatrix}$$
$$\overset{\substack{Z_2\leftarrow 2Z_2\\ Z_2\leftarrow Z_2-(X^2-X)Z_1}}{\sim} \begin{pmatrix} 2 & 1-X \\ 0 & X^3-X \\ 0 & 0 \end{pmatrix} := S$$

Lemma 17.1.5. Seien $a, b, c, d, g \in A$ mit $(a, b) = (g)$. Dann gibt es $h, j, i \in A$ mit $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \sim \begin{pmatrix} g & h \\ i & j \end{pmatrix}$ und $h \in (a, b)$ sowie $i, j \in (c, d)$.

Beweis. Œ $g \neq 0$. Schreibe $g = sa + tb$ mit $s, t \in A$ und $a = a'g$ sowie $b = b'g$ mit $a', b' \in A$. Dann $1 = sa' + tb' = sa - t(-b')$ und $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \overset{\binom{S_1}{S_2}\leftarrow\binom{sS_1+tS_2}{-b'S_1+a'S_2}}{\sim} \begin{pmatrix} g & h \\ i & j \end{pmatrix}$ für gewisse $h, j, j \in A$.

Bemerkung 17.1.6. Im Fall $A = \mathbb{Z}$ oder $A = K[X]$ (K sei ein Körper) kann man g aus der Aussage und $s, t, a', b' \in A$ im Beweis explizit berechnen, siehe §16.3 (nehme $g = \gcd = \{a, b\}$). Man kann also dann h, i, j explizit berechnen.

Lemma 17.1.7. Seien $a, b \in A$. Dann gibt es $c, d \in A$ mit $\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \sim \begin{pmatrix} c & 0 \\ 0 & d \end{pmatrix}$ mit $(c) = (a, b)$ und $c \mid d$.

Beweis. Wähle $c \in A$ mit $(a, b) = (c)$. Dann gibt es nach **??** ein $h \in (a, b)$ und $i, j \in (b)$ mit $\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \sim \begin{pmatrix} a & b \\ 0 & b \end{pmatrix} \sim \begin{pmatrix} c & h \\ i & j \end{pmatrix}$. Wegen $h \in (a, b) = (c)$ gilt $\begin{pmatrix} c & h \\ i & j \end{pmatrix} \sim \begin{pmatrix} c & 0 \\ i & d \end{pmatrix}$ für ein $d \in (i, j) \subseteq (b) \subseteq (c)$. Wegen $i \in (b) \subseteq (c)$ gilt $\begin{pmatrix} c & 0 \\ i & d \end{pmatrix} \sim \begin{pmatrix} c & 0 \\ 0 & d \end{pmatrix}$.

Lemma **??** brauchen wir später noch einmal in folgender Variante:

Lemma 17.1.8. Seien $a, b, g \in A$ mit $(a, b) = (g)$. Dann gibt es $h \in A$ mit $\begin{pmatrix} a & b \end{pmatrix} \sim \begin{pmatrix} g & h \end{pmatrix}$ und $\begin{pmatrix} a \\ b \end{pmatrix} \sim \begin{pmatrix} g \\ h \end{pmatrix}$.

Beweis. $\begin{pmatrix} a & b \end{pmatrix} \sim \begin{pmatrix} g & h \end{pmatrix}$ folgt durch Inspektion des Beweises von **??**, da durch nur eine Spaltenoperation durchgeführt wird.

$\begin{pmatrix} a \\ b \end{pmatrix} \sim \begin{pmatrix} g \\ h \end{pmatrix}$ folgt mir der entsprechenden Zeilenoperation.

Lemma 17.1.9. Seien $m, n \in \mathbb{N}$ und $M \in A^{m \times n}$. Dann gibt es $a \in A$ mit

$$M \sim \begin{pmatrix} a & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & * & \\ 0 & & & \end{pmatrix}.$$

Beweis. Bezeichne G die Menge der Matrizen $N \in A^{m \times n}$, deren linker obere Eintrag alle Einträge der ersten Zeile und ersten Spalte von N teilt. Es reicht, ein $N \in G$ mit $M \sim N$ zu finden.

Behauptung. Ist $M' \in A^{m \times n} \setminus G$ mit linkem oberen Eintrag a', so gibt es $M'' \in A^{m \times n}$ mit linkem oberen Eintrag a'' derart, dass $M' \sim M''$ und $(a') \subset (a'')$.

Begründung. Sei $M' \in A^{m \times n} \setminus G$, etwa $M' = \begin{pmatrix} a' & b' \\ * & * \end{pmatrix}$ mit $a' \nmid b'$. Wähle $a'' \in A$ mit $(a', b') = (a'')$. Nach **??** gibt es dann $h \in A$ und $M'' = \begin{pmatrix} a'' & h \\ * & * \end{pmatrix}$ mit $M' \sim M''$. Es gilt $(a') \subseteq (a'')$, aber nicht $(a') = (a'')$, denn sonst $b' \in (a'') = (a')$ und daher $a' \mid b'$.

Ist $M_1 := M \in G$, so setzen wir $N := M_1$ und sind fertig. Sonst gibt es $M_2 \in A^{m \times n}$ mit $M_1 \sim M_2$ und $(a_1) \subset (a_2)$ (a_i bezeichne den linken oberen Eintrag von M_i). Ist $M_2 \in G$, so setze $N := M_2$. Ist $M_2 \notin G$, so findent wir M_3 mit $M_1 \sim M_2 \sim M_3$ und $(a_1) \subset (a_2) \subset (a_3)$. Da A ein Hauptidealring und damit faktoriell ist, gilt die Teilerkettenbedingung aus **??**, das heißt die so weitergeführte Konstruktion muss irgendwann abbrechen (andernsfalls folgt $(a_1) \subset (a_2) \subset (a_3) \subset (a_4) \subset ...$) und wir finden $M_1, ..., M_k \in A^{m \times n}$ mit $M = M_1 \sim M_2 \sim M_3 \sim ... \sim M_k \in G$.

Lemma 17.1.10. Sei $D \in A^{n \times n}$ eine Diagonalmatrix. Dann gibt es $c_1, ..., c_n \in A$ mit $c_1 \mid c_2 \mid c_3 \mid ... \mid c_n$ und $D \sim \begin{pmatrix} c_1 & & 0 \\ & \ddots & \\ 0 & & c_n \end{pmatrix}$.

Beweis. Man kann die Nullen auf der Diagonalen von D alle nach rechts unten wandern lassen, denn $\begin{pmatrix} 0 & 0 \\ 0 & a \end{pmatrix} \overset{S_1 \leftrightarrow S_2}{\sim} \begin{pmatrix} 0 & 0 \\ a & 0 \end{pmatrix} \overset{Z_1 \leftrightarrow Z_2}{\sim} \begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix}$. Wegen $a \mid 0$

für alle $a \in A$ kann man dann gleich voraussetzen, dass $D = \begin{pmatrix} d_1 & & 0 \\ & \ddots & \\ 0 & & d_n \end{pmatrix}$ mit $d_1, ..., d_n \in A \setminus \{0\}$. Für jedes $d \in A \setminus \{0\}$ bezeichne $l(d)$ die Summe der Exponenten in der Primfaktorzerlegung von d, das heißt ist $(c, \alpha) \in A^\times \times \mathbb{N}_0^{(\mathbb{P}_A)}$ mit $d = c\mathbb{P}_A^\alpha$, so gilt $l(d) := \sum_{p \in \mathrm{supp}(\alpha)} \alpha(p)$. Setze $\delta : \begin{cases} (A \setminus \{0\})^n \to \mathbb{N}_0 \\ (c_1, ..., c_n) \mapsto \sum_{i=1}^n (n-i) l(c_i) \end{cases}$ und $G := \{(c_1, ..., c_n) \in (A \setminus \{0\})^n : c_1 \mid c_2 \mid ... \mid c_n\}$.

Behauptung. Ist $(d_1', ..., d_n') \in (A \setminus \{0\})^n \setminus G$, so gibt es $(d_1', ..., d_n') \in (A \setminus \{0\})^n$ mit $\begin{pmatrix} d_1' & & 0 \\ & \ddots & \\ 0 & & d_n' \end{pmatrix} \sim \begin{pmatrix} d_1'' & & 0 \\ & \ddots & \\ 0 & & d_n'' \end{pmatrix}$ und $\delta(d_1', ...d_n') > \delta(d_1'', ...d_n'')$.

Begründung. Sei $(d_1', ..., d_n') \in (A \setminus \{0\})^n \setminus G$. Wähle $i \in \{1, ..., n-1\}$ mit $d_i' \nmid d_{i+1}'$. Nach Lemma **??** gibt es $c, d \in A$ mit $(c) = (d_i', d_{i+1}')$ und $\begin{pmatrix} d_i' & 0 \\ 0 & d_{i+1}' \end{pmatrix} \sim \begin{pmatrix} c & 0 \\ 0 & d \end{pmatrix}$ (und $c \mid d$, was wir aber nicht brauchen). Insbesondere $d_i' d_{i+1}' = \det \begin{pmatrix} d_i' & 0 \\ 0 & d_{i+1}' \end{pmatrix} \mathrel{\widehat{=}} \begin{pmatrix} c & 0 \\ 0 & d \end{pmatrix} = cd$ und daher $c, d \in A \setminus \{0\}$ sowie

$$(*) \qquad l(d_i') + l(d_{i+1}') = l(d_i' d_{i+1}') = l(cd) = l(c) + l(d).$$

Setze $(d_1'', ..., d_n'') := (d_1', ..d_{i-1}, c, d, d_{i+2}', ..., d_n')$. Noch zu zeigen ist: $\delta(d_1', ..., d_n') > \delta(d_1'', ..., d_n'')$. Wegen (*) reicht es zu zeigen, dass $l(d_i') > l(c)$. Dies folgt aus $c \nmid d_i'$ (sonst ist $d_{i+1} \in (c) = (d_i')$ und daher $d_i \mid d_{i+1}$).

Nun folgt alles durch mehrfache Anwendung der Hilfsbehauptung.

Beweis von Satz **??**. Sei $M \in A^{m \times n}$. Zu zeigen ist, dass es $S \in A^{m \times n}$ gibt in Smithscher Normalform mit $M \sim S$. Setzte $l := \min\{m, n\}$. Durch l-fache Anwendung von Lemma **??** erhält man $D \in A^{l \times l}$ diagonal mit $M \sim \begin{pmatrix} D \\ 0 \end{pmatrix}$ falls $m \geq n$ bzw. $M \sim \begin{pmatrix} D & 0 \end{pmatrix}$ falls $m \leq n$. Wende nun letztes Lemma auf D an, um $c_1, ..., c_l \in A$ mit $c_1 \mid c_2 \mid ... \mid c_l$ zu finden mit $D \sim \begin{pmatrix} c_1 & & 0 \\ & \ddots & \\ 0 & & c_n \end{pmatrix}$. Nach Operationen von Typ $Z_i \leftarrow \lambda Z_i \qquad (\lambda \in A^\times)$ gilt Œ $c_i = \mathbb{P}_A^{\alpha_i}$ mit $\alpha_I \in \mathbb{N}_0^{(\mathbb{P}_A)} \cup \{\infty\}$. Wegen $c_1 \mid c_2 \mid ... \mid c_l$ gilt $\widehat{c_1} \preceq \widehat{c_2} \preceq ... \preceq \widehat{c_l}$ und daher $\alpha_1 \preceq \alpha_2 \preceq ... \preceq \alpha_l$ [→**??**,**??**].

Alle betrachteten Zeilenoperation auf $(m \times n)$-Matrizen sind von der Form

$$\begin{pmatrix} Z_1 \\ \vdots \\ Z_m \end{pmatrix} \leftarrow \begin{pmatrix} a_{1,1}Z_1 + & \dots & +a_{1,m}Z_m \\ \vdots & & \vdots \\ a_{m,1}Z_1 + & \dots & +a_{m,m}Z_m \end{pmatrix}$$

mit $a_{i,j} \in A$. Setzt man $P := (a_{i,j})_{1 \le i,j \le m} \in A^{m \times m}$, so erhält man aus $M \in A^{m \times n}$ durch Anwendung dieser Operation offenbar PM. Analog sind die Spaltenoperation auf $(m \times n)$-Matrizen sind von der Form

$$\begin{pmatrix} S_1 \\ \vdots \\ S_m \end{pmatrix} \leftarrow \begin{pmatrix} b_{1,1}S_1 + & \dots & +b_{1,n}S_n \\ \vdots & & \vdots \\ b_{n,1}S_1 + & \dots & +b_{n,n}S_n \end{pmatrix}$$

mit $b_{j,i} \in A$. Setzt man $Q := (b_{i,j})_{1 \le i,j \le n} \in A^{n \times n}$, so überführt diese Operation $M \in A^{m \times n}$ in MQ.

17.2 Die Formel von Cauchy-Binet

[Augustin Louis-Cauchy *1789 †1857; Jacques Philippe Marie Binet *1786, †1856]

In diesem Unterabschnitt sei stets K ein kommutativer Ring.

Definition 17.2.1. Seien $m, n \in \mathbb{N}_0$ und $A \in K^{m \times n}$. Für alle $I \subseteq \{1, ..., m\}$ und $J \subseteq \{1, ..., n\}$ bezeichnet $A_{I,J} \in K^{(\#I) \times (\#J)}$ die Matrix, durch aus A durch Streichen aller Zeilen i mit $i \notin I$ und Spalten j mi $j \notin J$ entsteht.

Bemerkung 17.2.2. Für $A \in K^{n \times n}$ gilt also mit der Notation aus **??** für $I \subseteq \{1, ..., n\}$: $A_I = A_{I,I}$

Definition 17.2.3. [→**??**] Seien $m, n, k \in \mathbb{N}_0$ und $A \in K^{m \times n}$. Die Determinanten der $\binom{m}{k} \cdot \binom{n}{k}$ Matrizen $A_{I,J}$ ($I \subseteq \{1, ..., n\}$,$J \subseteq \{1, ..., m\}$, $\#I = k = \#J$) nennt man die Minoren der Ordnung k von A. Ein Minor von A ist ein Minor irgendeiner Ordnung von A.

Bemerkung 17.2.4. Ein Minor einer Matrix A ist also die Determinante einer quadratischen Matrix, die durch eventuelles Streichen von Zeilen und Spalten aus A hervorgeht.

Satz 17.2.5. (Formel von Cauchy-Binet) Seien $m, n \in \mathbb{N}_0$, $A \in K^{m \times n}$ und $B \in K^{n \times m}$. Mit $I := \{1, ..., m\}$ gilt dann

$$\det(AB) = \sum_{\substack{J \subseteq \{1,...,n\} \\ \#J = m}} (\det A_{I,J})(\det B_{J,I}).$$

Beweis. Setze $J_0 := \{1, ..., n\}$. Da Matrixenmultiplikation "simultanes Multiplizieren mit Spaltenvektoren" ist [→**??**(c)], gilt

$$AB = A(B_{J_0,\{0\}} \dots B_{J_0,\{m\}}) = \left(\sum_{j=1}^{n} B_{\{j\},\{1\}} A_{I,\{j\}} \dots \sum_{j=1}^{n} B_{\{j\},\{m\}} A_{I,\{j\}} \right).$$

Da die Determinante "linear in den Spalten" ist, folgt

$$\det(AB) = \sum_{j_1=1}^{n} \dots \sum_{j_m=1}^{n} B_{\{j_1\},\{1\}} \dots B_{\{j_m\},\{m\}} \det(A_{I,\{j_1\}} \dots A_{I,\{j_m\}}).$$

Weil die Determinante einer Matrix mit zwei identischen Spalten gleich Null ist, folgt

$$\det(AB) = \sum_{1 \le j_1 < \dots < j_m \le n} \sum_{\sigma \in S_n} B_{\{j_{\sigma(1)}\},\{1\}} \dots B_{\{j_{\sigma(m)}\},\{m\}} \det(A_{I,\{j_{\sigma(1)}\}} \dots A_{I,\{j_{\sigma(m)}\}}).$$

Da sich beim Vertauschen zweier Spalten das Vorzeichen der Determinante ändert und jede Permutation einer Hintereinanderschaltung von Transpositionen ist, ist die hier rechts auftretende Determinante gleich $(\operatorname{sgn}\sigma)\det(A_{I,\{j_1\}} \dots A_{I,\{j_m\}}) = (\operatorname{sgn}\sigma)\det(A_{I,\{j_1,\dots,j_m\}})$ und es folgt

$$\begin{aligned} \det(AB) &= \sum_{1 \le j_1 < \dots < j_m \le n} \det(A_{I,\{j_1,\dots,j_m\}}) \sum_{\sigma \in S_n} (\operatorname{sgn}\sigma) B_{\{j_{\sigma(1)}\},\{1\}} \dots B_{\{j_{\sigma(m)}\},\{m\}} \\ &= \sum_{\substack{J \subseteq \{1,\dots,n\} \\ \#J = m}} \det(A_{I,J}) \det(B_{J,I}). \end{aligned}$$

Bemerkung 17.2.6. (a) Im Fall $m = n$ gibt es in der Formel von Cauchy-Binet nur den Summanden mit $J = \{1, .., n\}$. Dabei handelt es sich in diesem Fall um einen neuen Beweis des Determinantenproduktsatzes **??**.

(b) Falls $m > n$, gibt es in der Formel von Cauchy-Binet keinen Summanden. Die Formel besagt dann, dass $\det(AB) = 0$, was im Falle mit K als Körper schon bekannt ist, da die Abbildung $K^m \to K^m, x \mapsto ABx$ dann aus Dimensionsgründen nicht surjektiv (oder gleichbedeutend injektiv) sein kann. EVENTUELL DIAGRAMM HINZUFÜGEN.

Beispiel 17.2.7. Setze $A := \begin{pmatrix} 1 & 1 & -2 \\ 3 & 1 & -1 \end{pmatrix} \in Z^{2\times 3}$, $B := \begin{pmatrix} 1 & 1 \\ 3 & 1 \\ 0 & 2 \end{pmatrix} \in \mathbb{Z}^{3\times 2}$ und $I :=$

$\{1,2\}$. Dann

$$\begin{aligned}\det(AB) &= \det(A_{I,\{1,2\}})\det(B_{\{1,2\},I}) + \det(A_{I,\{1,3\}})\det(B_{\{1,3\},I}) + \det(A_{I,\{2,3\}})\det(B_{\{2,4\},I}) \\ &= \det\begin{pmatrix}1 & 1\\ 3 & 1\end{pmatrix}\det\begin{pmatrix}1 & 1\\ 3 & 1\end{pmatrix} + \det\begin{pmatrix}1 & 2\\ 3 & -1\end{pmatrix}\det\begin{pmatrix}1 & 1\\ 0 & 2\end{pmatrix} + \det\begin{pmatrix}1 & 2\\ 1 & -1\end{pmatrix}\det\begin{pmatrix}3 & 1\\ 0 & 2\end{pmatrix} \\ &= (-2)(-2) + (-7)2 + (-3)6 = 4 - 18 - 18 = -28.\end{aligned}$$

Korollar 17.2.8. (Minorenformel) Seien $m, n, r \in \mathbb{N}_0$, $A \in A^{m\times n}$ und $B \in K^{n\times r}$. Seien $I \subseteq \{1, ..., m\}$ und $L \subseteq \{1, ..., r\}$ mit $\#I = \#L =: k$. Dann gilt
$\det((AB)_{I,L}) = \sum_{\substack{J\subseteq\{1,...,n\}\\ \#J=m}} (\det A_{I,J})(\det B_{J,L})$.

Beweis. Beachte $(AB)_{I,L} = A_{I,\{1,...,n\}}B_{\{1,...,n\},L}$ und nutze Cauchy-Binet.

Bemerkung 17.2.9. (a) Im Fall $m = r = k$ wird die Minorenformel zur Formel von Cauchy-Binet.

(b) Im Fall $k = 1$ wird die Minorenformel zur Formel der Matrixmultiplikation.

Definition 17.2.10. Seien $m, n, k \in \mathbb{N}_0$ und $A \in K^{m\times n}$. Das k-te Minorenideal von A, geschrieben $\mathrm{minor}_k(A)$, ist das Ideal von K, welches von allen Minoren der Ordnung k von A erzeugt wird.

Beispiel 17.2.11. Sei $A := \begin{pmatrix}2 & 4\\ -2 & 2\\ 0 & 6\end{pmatrix} \in \mathbb{Z}^{3\times 2}$. Dann ist $\mathrm{minor}_0(A) = (1) = \mathbb{Z}$, $\mathrm{minor}_1(A) = (\gcd\{A\}) = (2)$, $\mathrm{minor}_2(A) = (12, 12, -12) = (12)$, $\mathrm{minor}_3(A) = \mathrm{minor}_4(A) = ... = (0)$.

Korollar 17.2.12. Seien $m, n, k \in \mathbb{N}_0$, $M \in K^{m\times n}$ und seien $P \in K^{m\times m}$ und $Q \in K^{n\times n}$ invertierbar. Dann stimmen die k-ten Minorenideale von M und von PMQ überein.

Beweis. Nach der Minorenformel ist das k-te Minorenideal von PMQ in dem von M enthalten. Genauso ist das k-te Minorenideal von $M = P^{-1}(PQM)Q^{-1}$ in dem von PMQ enthalten.

Proposition 17.2.13. Seien $m, n, k \in \mathbb{N}_0$ und $A \in K^{m\times n}$. Dann $\mathrm{minor}_0(A) = K$, $\mathrm{minor}_k(A) = \{0\}$ falls $k > \min\{m, n\}$ und $\mathrm{minor}_0(A) \geq \mathrm{minor}_1(A) \geq \mathrm{minor}_2(A) \geq ....$

Beweis. Entwickelt man einen Minor der Ordnung $k+1$ nach einer Zeile oder einer Spalte [→**??**], so sieht man, dass er im Ideal $\mathrm{minor}_k(A)$ liegt.

17.3 Eindeutigkeit der Normalform von Smith

Wir springen zurück in die Notation und in die Generalvoraussetzungen von Paragraph §17.1. Insbesondere sei A ein Hauptidealring. Daher kann man die gesamte Information über die Minorenideale einer Matrix $M \in A^{m\times n}$ durch die sogenannten Determinantenteiler von M erfassen.

Definition 17.3.1. Seien $m, n \in \mathbb{N}_0$, $M \in A^{m\times n}$ und $l := \min\{m, n\}$. Dann heißt für $k \in \{1, ..., l\}$ das eindeutig bestimmte $\mathbb{P}_A^\alpha$ ($\alpha \in \mathbb{N}_0^{(\mathbb{P}_A)} \cup \{\infty\}$, hier wieder $\mathbb{P}_A^\infty := 0$) mit $\text{minor}_k(M) = (\mathbb{P}_A^\alpha)$ der k-te Determinantenteiler $d_k(M)$ von M. Wir setzen $d(M) := (d_1(M), ..., d_l(M))$.

Beispiel 17.3.2. Sei $M := \begin{pmatrix} 2 & 4 \\ -2 & 2 \\ 0 & 6 \end{pmatrix} \in \mathbb{Z}^{3\times 2}$ [→**??**]. Setzt man wie üblich $\mathbb{P}_\mathbb{Z} = \mathbb{P}$, so gilt $d(M) = (2, 12)$.

Bemerkung 17.3.3. (a) $d_k(M)$ ist nach **??**(c) ein ggT der Minoren der Ordnung k von M.

(b) Wegen $\text{minor}_1(M) \supseteq ... \supseteq \text{minor}_l(M)$ gilt, dass $d_1(M) \mid ... \mid d_l(M)$.

(c) Sind $P \in A^{m\times m}$ und $Q \in A^{n\times n}$ invertierbar, so $d(M) = d(PMQ)$ [→**??**].

Lemma 17.3.4. Sei $S \in A^{m\times n}$ in Smithscher Normalform, $l := \min\{n, m\}$ und $c_1, ..., c_l \in A$ mit

$$S = \left(\begin{array}{ccc} c_1 & & 0 \\ & \ddots & \\ 0 & & c_l \\ \hline & & \end{array}\right) \quad \text{falls } m \geq n$$

$$\text{bzw.} \qquad S = \left(\begin{array}{ccc|c} c_1 & & 0 & \\ & \ddots & & 0 \\ 0 & & c_l & \end{array}\right) \quad \text{falls } m \leq n.$$

Dann gilt $d_k(M) = c_1...c_k$ für $k \in \{1, ..., l\}$.

Beweis. Sei $k \in \{1, ...l\}$. Dann $(c_1...c_k) \subseteq \text{minor}_k(M) \overset{??}{\subseteq} (\{c_{i_1}...c_{i_k} \mid 1 \leq i_1 \leq ... \leq i_k \leq l\}) \overset{c_1|...|c_l}{\subseteq} (c_1...c_k)$, also $(d_k(M)) = (c_1...c_k)$ und daher $d_k(M) \widehat{=} c_1...c_k$ und nach Definition **??** und **??** sogar $d_k(M) = c_1...c_k$.

Satz 17.3.5. (Elementarteilersatz, auch Invariantenteilersatz genannt)

(a) Zu jeder Matrix $M \in A^{m\times n}$ gibt es genau eine Matrix $S \in A^{m\times n}$ in Smithscher Normalform mit $M \sim S$.

(b) Sind $M, N \in A^{m\times n}$, so gilt $M \sim N$ genau dann, wenn es invertierbare Matrizen $P \in A^{m\times m}$ und $Q \in A^{n\times n}$ gibt mit $M = PNQ$.

Beweis. **Zu (a)**. Die Existenz ist in **??** gezeigt. Zur Eindeutigkeit: Seien $S, T \in A^{m\times n}$ in Smithscher Normalform mit $S \sim T$. Zu zeigen ist $S = T$. Nach §17.1 gibt es invertierbare $\mathbb{P} \in A^{m\times n}$ und $Q \in A^{n\times n}$ mit $S = PTQ$. Nach **??** gilt dann $\text{minor}_k(S) = \text{minor}_k(T)$ für alle $k \in \mathbb{N}_0$ und daher $d(S) = d(T)$. Mit **??** folgert man leicht $S = T$.

Zu (b). Seien $P \in A^{m\times m}$ und $Q \in A^{n\times n}$ invertierbar mit $M = PNQ$. Zu zeigen ist $M \sim N$. Seien S und T die Smithschen Normalformen von M und N. Wie im Beweis von (a) gerade gezeigt, gilt dann $S = T$, also $M \sim S = T \sim N$ und daher $M \sim N$.

Definition 17.3.6. Sei $M \in A^{m\times n}$. Die eindeutig bestimmte Matrix $S \in A^{m\times n}$ in Smithscher Normalform mit $M \sim S$ heißt die Smithsche Normalform von M. Für $k \in \{1, ..., l\}$ mit $l := \min\{n, m\}$ heißt der Eintrag in der k-ten Zeile und k-Spalte von S der k-te Elementarteiler von M, geschrieben $c_k(M)$. Wir setzen $c(M) := (c_1(M), ..., c_l(M))$.

Korollar 17.3.7. (a) $d_k(M) = c_1(M)...c_k(M)$ für alle $M \in A^{m\times n}$ und $k \in \{1, ..., l\}$ mit $l := \min\{n, m\}$.

(b) $M \sim N \iff d(M) = d(N) \iff c(M) = c(N)$ für alle $M, N \in A^{m\times n}$.

Beispiel 17.3.8. Sei wieder $M := \begin{pmatrix} 2 & 4 \\ -2 & 2 \\ 0 & 6 \end{pmatrix} \in \mathbb{Z}^{3\times 2}$ [→**??**,**??**]. Wegen $d(M) = (2, 12)$ gilt $c(M) = (2, 6)$. Die Smithsche Normalform von M ist also $S = \begin{pmatrix} 2 & 0 \\ 0 & 6 \\ 0 & 0 \end{pmatrix}$.

17.4 Charakterisierung der Ähnlichkeit und nochmals Cayley-Hamilton

In diesem Abschnitt sei stets K ein kommutativer Ring mit $1 \neq 0$. Nach **??**(a) sind zwei Matrizen über einen Hauptidealring (zeilenspalten-)äquivalent genau dann, wenn sie dieselbe Smithsche Normalform haben. Oftmals interessiert man sich aber nicht für Äquivalenz, sondern für Ähnlichkeit von Matrizen.

Erinnerung 17.4.1. [→**??**, **??**] $A, B \in K^{n\times n}$ heißen ähnlich, in Zeichen $A \approx B$, wenn es eine invertierbare Matrix $P \in K^{n\times n}$ gibt mit $A = P^{-1}BP$. Ähnlichkeit ist eine Äquivalenzrelation.

Für $A, B \in K^{n\times n}$ gilt

$$\begin{aligned} A \approx B &\iff \text{es gibt ein invertierbares } P \in K^{n\times n} \text{ mit } A = PBP^{-1} \\ &\overset{??}{\iff} \text{es gibt (invertierbare) } P, Q \in K^{n\times n} \text{ mit } PAQ = B \text{ und } PQ = I_n \\ &\overset{Koeff.vergleich}{\iff} \text{es gibt (invertierbare) } P, Q \in K^{n\times n} \text{ mit } P(A - XI_n)Q = B - XI_n \end{aligned}$$

und andererseits

$$A - XI_n \sim B - XI_n \overset{??}{\iff} \text{es gibt invertierbare } P, Q \in K[X]^{n\times n} \text{ mit}$$
$$(*) \qquad P(A - XI_n)Q = B - XI_n.$$

Definition 17.4.2. Ist $A \in K^{n\times n}$, so heißt $A - XI_n \in K[X]^{n\times n}$ die charakteristische Matrix von A.

Bemerkung 17.4.3. Sei K ein Körper.

(a) Die Determinante der charakteristischen Matrix ist das charakteristische Polynom [→**??**(e)].

(b) Obige Beobachtung lässt die Idee aufkommen, dass man die Ähnlichkeit von Matrizen im Zusammenhang mit der Äquivalenz mit der Äquivalenz ihrer charakteristischen Matrizen bringen könnte. Das Problem scheint dabei, dass die Übergangsmatrizen P und Q einmal aus $K^{n\times n}$ und einmal aus $K[X]^{n\times n}$ sein sollen. Trotzdem wird sich auf überraschende Weise herausstellen, dass die Ähnlichkeit von Matrizen gleichbedeutend zur Äquivalenz ihrer charakteristischen Matrizen ist.

(c) Da man die Elemente von $K[X]^{n\times n}$ in der Form $X^dP_d + X^{d-1}P_{d-1} + ... + P_0$ mit eindeutig bestimmten $P_i \in K^{n\times n}$ schreiben kann, nennt man sie auch Matrixpolynome.

(d) Im Hinblick auf (b) untersuchen wir Möglichkeiten durch Einsetzen von etwas für X aus einem Matrixpolynom eine Matrix zu machen.
Einsetzen eines Skalars $\lambda \in K$ für X scheint leider nichts zu bringen, denn aus (*) wird dann $P(\lambda)(A - \lambda I_n)Q(\lambda) = B - \lambda I_n$ mit invertierbaren $P(\lambda), Q(\lambda) \in K^{n\times n}$, aber da die Unbestimmte verschwunden ist, kann man dies nicht aufspalten in $P(\lambda)AQ(\lambda) = B(\lambda)$ und $P(\lambda)Q(\lambda) = I_n$.
Gewöhnliches Einsetzen einer Matrix $M \in K^{n\times n}$ für X würde (*) zu einer Gleichung von Elementen aus $K[M]^{n\times n}$ [→**??**(b)] machen und die Sache eher noch verkomplizieren. Wir definieren daher die Linkseinsetzung einer Matrix in ein Matrixpolynom. Sie hat (genauso wie die Rechtseinsetzung, die wir alternativ verwenden könnten) viel schlechtere Eigenschaften als etwa die in **??**(b) definierte Einsetzung einer Matrix in ein Polynom. Sie wird uns aber in erstaunlicher Weise helfen.

Definition 17.4.4. Ist $P = X^dP_d + X^{d-1}P_{d-1} + ... + P_0 \in K[X]^{n\times n}$ mit $P_i \in K^{n\times n}$, dann nennen wir $P_A := A^dP_d + ... + P_0$ für $A \in K^{n\times n}$ die Linkseinsetzung von der Matrix A in (das Matrixpolynom) P

Beispiel 17.4.5. Ist $P := \begin{pmatrix} 2-X^2 & X \\ 1 & 1+X \end{pmatrix} = X^2 \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix} + X \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} + \begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix}$ und $A := \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix}$, so ist

$$\begin{aligned} P_A &= \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix}^2 \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} + \begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 3 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 3 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 3 \\ 1 & 1 \end{pmatrix}. \end{aligned}$$

Lemma 17.4.6. Seien $P, Q \in K[X]^{n\times n}$ und $A \in K^{n\times n}$. Sei weiter $P = X^d P_d + ... + P_0$ mit $P_i \in K^{n\times n}$ derart, dass $P_i A = A P_i$ für alle $i \in \{1, ..., d\}$. Dann gilt $(PQ)_A = P_A Q_A$.

Beweis. Schreibe $Q = X^e Q_e + ... + Q_0$ mit $Q_j \in K^{n\times n}$. Dann ist

$$\begin{aligned} (PQ)_A &= \left(\sum_{k=0}^{d+e} X^k \sum_{i+j=k} P_i Q_j \right)_A \begin{bmatrix} \text{setze } P_i := 0 \text{ für i>d} \\ \text{und } Q_j := 0 \text{ für j>e} \end{bmatrix} \\ &= \sum_{k=0}^{d+e} A^k \sum_{i+j=k} P_i Q_j \\ &\overset{P_iA=AP_i}{=} \sum_{k=0}^{d+e} \sum_{i+j=k} (A^i P_i)(A^j Q_j) \\ &= \sum_{i=0}^{d} \sum_{j=0}^{e} (A^i P_i)(A^j Q_j) \\ &= \left(\sum_{i=0}^{d} A^i P_i \right) \left(\sum_{j=0}^{e} A^j Q_j \right) = P_A Q_A. \end{aligned}$$

Lemma 17.4.7. [→??] Sei $P \in K[X]^{n\times n}$ und $A \in K^{n\times n}$. Dann $P_A = 0 \iff \exists Q \in K[X]^{n\times n} : P = (XI_n - A)Q$.

Beweis. " $\Longleftarrow$ " Sei $Q \in K[X]^{n\times n}$ mit $P = (XI_n - A)Q$. Dann gilt wegen $I_n A = AI_n$ und $AA = AA$ nach **??** $P_A = (XI_n - A)_A Q_A = \underbrace{(AI_n - A)}_{=0} Q_A = 0$.

" $\Longrightarrow$ " Gelte $P_A = 0$. Schreibe $P = X^d P_d + ... + P_0$ mit $P_i \in K^{n\times n}$. Dann

$$\begin{aligned} P = P - P_A &= (X^d I_n - A^d)P_d + ... + (XI_n - A)P_1 \\ &= (XI_n - A)(X^{d-1}I_n + X^{d-2}A + X^{d-3}A^2 + ... + A^{d-1}) \\ &\quad + ... + (XI_n - A)P_1. \end{aligned}$$

Lemma 17.4.8. Seien $A, B \in K^{n\times n}$ und $P, Q \in K[X]^{n\times n}$ mit $P(B - XI_n) = (A - XI_n)Q$. Dann gilt $AP_A = P_A B$.

Beweis. Es gilt

$$\begin{aligned} 0 &= (A - AI_n)Q \overset{??}{=} ((A - XI_n)Q)_A = (P(B - XI_n))_A = (PB - XP)_A \\ &= (PB_A) - (XP)_A = P_A B - AP_A. \end{aligned}$$

Satz 17.4.9. Seien $P, Q \in K[X]^{n\times n}$ invertierbar und $A, B \in K^{n\times n}$ mit $P(B - XI_n)Q = A - XI_n$. Dann ist $P_A \in K^{n\times n}$ invertierbar mit $P_A^{-1} A P_A = B$.

Beweis. Wegen $P(B - XI_n) = (A - XI_n)Q^{-1}$ folgt nach dem letzten Lemma $AP_A = P_A B$. Zu zeigen ist, P_A ist invertierbar. Setze $R := P^{-1} \in K[X]^{n\times n}$, das heißt $PR = I_n$. Nach Lemma **??** gibt es $F, G \in K[X]^{n\times n}$ mit $P - P_A = (XI_n - A)F$ und $R - R_B = (XI_n - B)G$. Es folgt

$$\begin{aligned} I_n = PR &= P(XI_n - B)G + PR_B = (XI_n - A)Q^{-1}G + PR_B \\ &= (XI_n - A)Q^{-1}G + P_A R_B + (XI_n - A)FR_B \end{aligned}$$

und daher $I_n - P_A R_B = (XI_n - A)(Q^{-1}G + FR_B)$. Es folgt $I_n - P_A R_B = (I_n - P_A R_B)_A \overset{??}{=} 0$, also $P_A R_B = I_n$ und nach **??** ist P_A invertierbar.

Korollar 17.4.10. Sei K ein Körper und $A, B \in K^{n\times n}$. Dann äquivalent:

(a) $A \approx B$

(b) $A - XI_n \sim B - XI_n$

(c) $c(A - XI_n) = c(B - XI_n)$

(d) $d(A - XI_n) = d(B - XI_n)$

(e) $A - XI_n$ und $B - XI_n$ haben dieselbe Smithsche Normalform.

Beweis. Beachte, dass $K[X]$ ein Hauptidealring ist.
(a) $\Longrightarrow$ ***(b)*** folgt aus **??**.

$(b) \iff (c) \iff (d) \iff (e)$ folgt aus **??** und **??**.

$(b) \implies (a)$ folgt aus **??**.

Algorithmus 17.4.11. Sei K ein Körper und $A, B \in K^{n\times n}$ gegeben. Um zu entscheiden, ob $A \approx B$ gilt, und um gegebenenfalls $R \in K^{n\times n}$ mit $R^{-1}AR = B$ zu berechnen, kann man wie folgt vorgehen: Berechne wie in §17.1 durch Zeilen- und Spaltenoperatonen der Form

- $Z_i \leftarrow Z_i + \lambda Z_j \qquad (i \neq j, \lambda \in K[X])$
- $Z_i \lambda Z_i \qquad (\lambda \in K^\times)$
- $S_i \leftarrow S_i + \lambda S_j \qquad (i \neq j,\ \lambda \in K[X], i \leq n, j \leq n)$
- $S_i \lambda Z_i \qquad (\lambda \in K^\times, i \leq n)$

$S, T \in K[X]^{n\times n}$ in Smithscher Normalform und $L \in K[X]^{n\times n}$ mit $(A - XI_n \mid I_n) \sim (S \mid L)$ und $B - XI_n \sim T$. Notiere dabei die Zeilenoperationen, die zu $B - XI_n \sim T$ führen (aus Effizienzgründen sollte man im Zweifelsfall Spaltenoperationen vorziehen!). Gilt $S \neq T$, so $A \not\approx B$.
Gelte also nun $S = T$. Wende dann die zu den notieren Zeilenoperationen inversen Zeilenoperationen in umgekehrter Reihenfolge auf L an und nenne die so erhaltene Matrix $P \in K[X]^{n\times n}$. Setze nun $R := P_B$. Dann ist R invertierbar und es gilt $R^{-1}BR = A$.

Beweis. $L \in K[X]^{n\times n}$ ist invertierbar und es gibt ein invertierbares $M \in K[X]^{n\times n}$ mit $L(A - XI_n)M = S$ (vergleiche **??**, im Gegensatz zu dort haben wir M nicht berechnet, da wir es nicht brauchen). Weiter gibt es invertierbare $P', Q' \in K[X]^{n\times n}$ mit $P'(B - XI_n)Q' = T$, wobei Multiplikation von links mit P' der Anwendung der notierten Zeilenoperationen entspricht (siehe §17.1). Gilt $S \neq T$, so ist $A \not\approx B$ nach **??**.
Gelte also $S = T$. Dann ist $L(A-XI_n)M = P'(B-XI_n)Q'$. Nun gilt $P = (P')^{-1}L$, also $P(A - XI_n)(M(Q')^{-1}) = B - XI_n$, weswegen nach **??** gilt, dass $R = P_B \in K^{n\times n}$ invertierbar ist mit $R^{-1}BR = P_B^{-1}BP_B = A$.

Beispiel 17.4.12. Betrachte $A, B \in \mathbb{Q}^{3\times 3}$ gegeben durch $A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}$ und $B =$

$\begin{pmatrix} 1 & 1 & \frac{1}{2} \\ 1 & 1 & -\frac{1}{2} \\ 0 & 2 & 1 \end{pmatrix}$. Es gilt

$$
\begin{aligned}
(A - XI_3 \mid I_3) &= \left(\begin{array}{ccc|ccc} 1-X & 1 & 0 & 1 & 0 & 0 \\ 0 & 1-X & 1 & 0 & 1 & 0 \\ 1 & 0 & 1-X & 0 & 0 & 1 \end{array}\right) \\
&\sim \left(\begin{array}{ccc|ccc} 0 & 1 & -(1-X)^2 & 1 & 0 & -(1-X) \\ 0 & 0 & 1+(1-X)^3 & -(1-X) & 1 & (1-X)^2 \\ 1 & 0 & 0 & 0 & 0 & 1 \end{array}\right) \\
&\sim \left(\begin{array}{ccc|ccc} 0 & 1 & 0 & 1 & 0 & X-1 \\ 0 & 0 & 1+(1-X)^3 & -X-1 & 1 & (X-1)^2 \\ 1 & 0 & 0 & 0 & 0 & 1 \end{array}\right) \\
&\sim \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & X-1 \\ 0 & 0 & (X-1)^3-1 & X-1 & 1 & (X-1)^2 \end{array}\right) := (S \mid L).
\end{aligned}
$$

Notiere im Folgenden nur die Zeilenoperationen!

$$
\begin{aligned}
B - XI_3 = \begin{pmatrix} 1-X & 1 & \frac{1}{2} \\ 1 & 1-X & -\frac{1}{2} \\ 0 & 2 & 1-X \end{pmatrix} &\overset{\substack{Z_1 \leftarrow Z_1 - (1-X)Z_2 \\ Z_3 \leftarrow \frac{1}{2}Z_3}}{\sim} \begin{pmatrix} 0 & 1-(1-X)^2 & \frac{1}{2}+\frac{1}{2}(1-X) \\ 1 & 0 & 0 \\ 0 & 1 & \frac{1}{2}(1-X) \end{pmatrix} \\
&\overset{Z_1 \leftarrow Z_1 - Z_3}{\sim} \begin{pmatrix} 0 & -(1-X)^2 & \frac{1}{2} \\ 1 & 0 & 0 \\ 0 & 1 & \frac{1}{2}(1-X) \end{pmatrix} \\
&\overset{Z_1 \leftarrow Z_1 + (1-X)^2 Z_3}{\sim} \begin{pmatrix} 0 & 0 & \frac{1}{2} \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \overset{Z_1 \leftrightarrow Z_3}{\sim} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & (X-1)^3-1 \end{pmatrix} =: T.
\end{aligned}
$$

S und T sind in Smithscher Normalform und es gilt $S = T$, also $A \approx B$. Führe nun die zu den notieren Zeilenoperationen inversen Operationen in umgekehrter Reihenfolge auf L aus.

$$
\begin{aligned}
L = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & X-1 \\ X-1 & 1 & (X-1)^2 \end{pmatrix} &\overset{Z_1 \leftrightarrow Z_3}{\sim} \begin{pmatrix} X-1 & 1 & (X-1)^2 \\ 1 & 0 & X-1 \\ 0 & 0 & 1 \end{pmatrix} \\
&\overset{Z_1 \leftarrow Z_1 - (1-X)^2 Z_3}{\sim} \begin{pmatrix} X-1 & 1 & 0 \\ 1 & 0 & X-1 \\ 0 & 0 & 1 \end{pmatrix} \overset{Z_1 \leftarrow Z_1 - (1-X)^2 Z_3}{\sim} \begin{pmatrix} X-1 & 1 & 0 \\ 1 & 0 & X-1 \\ 0 & 0 & 1 \end{pmatrix} \\
&\overset{Z_1 \leftarrow Z_1 + Z_3}{\sim} \begin{pmatrix} X-1 & 1 & 1 \\ 1 & 0 & X-1 \\ 0 & 0 & 1 \end{pmatrix} \overset{Z_3 \leftarrow 2Z_3}{\sim} \begin{pmatrix} X-1 & 1 & 1 \\ 1 & 0 & X-1 \\ 0 & 0 & 2 \end{pmatrix} \\
&\overset{Z_1 \leftarrow Z_1 + 2(1-X)_2}{\sim} \begin{pmatrix} 0 & 1 & 1-(X-1)^2 \\ 1 & 0 & X-1 \\ 0 & 0 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 1 & -X^2+2X \\ 1 & 0 & X-1 \\ 0 & 0 & 2 \end{pmatrix}.
\end{aligned}
$$

Somit ist

$$P := X^2 \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + X \begin{pmatrix} 0 & 0 & 2 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & 0 & 2 \end{pmatrix}.$$

Mit

$$\begin{aligned} R := P_B &= \begin{pmatrix} 1 & 1 & \frac{1}{2} \\ 1 & 1 & -\frac{1}{2} \\ 0 & 2 & 1 \end{pmatrix}^2 \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 1 & 1 & \frac{1}{2} \\ 1 & 1 & -\frac{1}{2} \\ 0 & 2 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 2 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & 0 & 2 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 1 & \frac{1}{2} \\ 1 & 1 & -\frac{1}{2} \\ 0 & 2 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 3 \\ 0 & 0 & 3 \\ 0 & 0 & 2 \end{pmatrix} + \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & 0 & 2 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 & -2 \\ 0 & 0 & -2 \\ 0 & 0 & -2 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 3 \\ 0 & 0 & 3 \\ 0 & 0 & 2 \end{pmatrix} + \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & 0 & 2 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix} \end{aligned}$$

gilt $R^{-1}BR = A$.

Proposition 17.4.13. Sei L ein Körper und K ein Unterkörper von L. Seien $A, B \in K^{n\times n}$. Dann sind A und B ähnlich über K (das heißt aufgefasst als Matrizen in $K^{n\times n}$) genau dann wenn, wenn A und B ähnlich über L sind.

Beweis. Da $K[X]$ ein Hauptidealring ist, gibt es $S, T \in K[X]^{n\times n}$ in Smithscher Normalform mit $A - XI_n \sim S$ über $K[X]$ und $B - XI_n \sim T$ über $K[X]$. Es sind dann S und T auch im Smithscher Normalform über $L[X]$ (das heißt als Matrizen aus $L[X]^{n\times n}$) und auch über L gilt $A - XI_n \sim S$ und $B - XI_n \sim T$. Daher gilt

$$A \approx B \text{ über } K \overset{??}{\iff} S = T \overset{??}{\iff} A \approx B \text{ über } L.$$

Mit den in diesem Abschnitt entwickelten Werkzeugen können wir den Satz von Caley-Hamilton **??**(b) jetzt mühelos noch einmal beweisen und zwar sogar in einer allgemeineren Form.

Satz 17.4.14. (Cayley-Hamilton) Sei $A \in K^{n\times n}$ und $\chi_A := \det(A - XI_n)$. Dann ist $\chi_A(A) = 0$.

Beweis. Sei $P := (A - XI_n)(\mathrm{com}(A - XI_n))^T \overset{??}{=} (\det(A - XI_n))I_n = \chi_A I_n$. Dann gilt $0 \overset{??}{=} P_A = \chi_A(A)I_n = \chi_A(A)$.

17.5 Die Normalform von Frobenius

[Ferdinand Georg Frobenius *1849 †1917]

In diesem Abschnitt sei K stets ein Körper.

Erinnerung 17.5.1. Sei $p = X^n + a_{n-1}X^{n-1} + ... + a_1X + a_0 \in K[X]$ mit $a_0, ..., a_n \in K$ ein normiertes Polynom. Dann ist das Ideal (p) des kommutativen Ringes $K[X]$ ein Unterraum des K-Vektorraums $K[X]$ und $\underline{v} := (\overline{1}, \overline{X}, ...\overline{X^{n-1}})$ eine Basis des Quotientenvektorraums $K[X]/(p)$ [→**??**]. Die Abbildung

$$f : K[X]/(p) \to K[X]/(p), \overline{q} \mapsto \overline{Xq} \qquad (q \in K[X])$$

ist wohldefiniert und linear und

$$C(p) := M(f, \underline{v}) = \begin{pmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \ddots & \vdots & -a_1 \\ 0 & 1 & \ddots & \vdots & -a_2 \\ 0 & 0 & \ddots & 0 & \vdots \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -a_{n-1} \end{pmatrix} \in K^{n \times n}$$

heißt die Begleitmatrix von p [→**??**]. Man sieht leicht $p(f) = 0$ und $r(f) \neq 0$ für alle $r \in K[X] \setminus \{0\}$ mit $\deg(r) < n$ [→**??**]. Es folgt $\mu_{C(p)} = \mu_f = p$ und $\chi_{C(p)} = \chi_f = (-1)^n p$ [→**??**,**??**].

Lemma 17.5.2. Sei $p \in K[X]$ normiert vom Grad ≥ 1. Dann ist die Smithsche Normalform der charakteristischen Matrix $C(p) - XI_n \in K[X]^{n \times n}$ gleich $\begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \\ & & & p \end{pmatrix}$.

Beweis. Sei $n := \deg p \geq 1$. Streicht man in $C(p) - XI_n = \begin{pmatrix} -X & & 0 & -a_0 \\ 1 & \ddots & & -a_1 \\ & \ddots & -X & \vdots \\ 0 & & 1 & -a_{n-2} \\ & & & -a_{n-1} - X \end{pmatrix}$ die erste Zeile und letzte Spalte, so erhält man eine obere Dreiecksmatrix mit Determinante 1. Es folgt $d(C(p) - XI_n) = (1, ..., 1, p) = d \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \\ & & & p \end{pmatrix}$.

Lemma 17.5.3. Sei $m \in \mathbb{N}_0$ und seien $p_1, ..., p_m \in K[X]$ normiert vom Grad ≥ 1 mit $p_1 \mid ... \mid p_m$. Dann ist die Smithsche Normalform von

$$A = \begin{pmatrix} \boxed{C(p_1)} & & 0 \\ & \ddots & \\ 0 & & \boxed{C(p_m)} \end{pmatrix} - XI_n \in K[X]^{n \times n} \text{ (mit } n := \textstyle\sum_{i=1}^{m} \deg(p_i))$$

gleich $\begin{pmatrix} 1 & & & & 0 \\ & \ddots & & & \\ & & 1 & & \\ & & & \ddots & \\ 0 & & & & p_m \end{pmatrix}$.

Beweis. Setze $n_i := \deg(p_i)$. Für $i \in \{1, ..., n\}$ gilt

$$\begin{pmatrix} \boxed{C(p_1)} & & 0 \\ & \ddots & \\ 0 & & \boxed{C(p_m)} \end{pmatrix} - XI_n \overset{??}{\sim} \begin{pmatrix} \boxed{\begin{smallmatrix} 1 & & \\ & \ddots & \\ & & 1 \\ & & & p_1 \end{smallmatrix}} & & 0 \\ & \ddots & \\ 0 & & \boxed{\begin{smallmatrix} 1 & & \\ & \ddots & \\ & & 1 \\ & & & p_m \end{smallmatrix}} \end{pmatrix}$$

$$\begin{pmatrix} 1 & & & & & & 0 \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & p_1 & & & \\ & & & & \ddots & & \\ & & & & & p_m & \\ 0 & & & & & & 0 \end{pmatrix}.$$

Satz und Definition 17.5.4. Sei $A \in K^{n \times n}$. Dann gibt es genau ein Tupel $(p_1, ..., p_m)$

von normierten Polynomen $p_i \in K[X]$ vom Grad ≥ 1 mit $p_1 \mid ...p_m$ und

$$A \approx \begin{pmatrix} \boxed{C(p_1)} & & 0 \\ & \ddots & \\ 0 & & \boxed{C(p_m)} \end{pmatrix}.$$

Die rechte Matrix heißt die Frobenius'sche Normalform von A. Es gilt $c(A - XI_n) = (1, ..., 1, p_1, ..., p_m)$, $\mu_A = p_m$ und $\chi_A = (-1)^n p_1 \dots p_m$.

Beweis. Seien $p_1, ..., p_m$ normierte Polynome vom Grad ≥ 1 mit $p_1 \mid ... \mid p_m$. Dann gilt

$$A \approx \begin{pmatrix} \boxed{C(p_1)} & & 0 \\ & \ddots & \\ 0 & & \boxed{C(p_m)} \end{pmatrix}$$

$\overset{??}{\iff}$ die Smitsche Normalform von $A - XI_n$ ist $\begin{pmatrix} 1 & & & & 0 \\ & \ddots & & & \\ & & 1 & & \\ & & & \ddots & \\ 0 & & & & p_m \end{pmatrix}$.

Hieraus folgt alles bis auf die Behauptung, dass in diesem Fall $\mu_A = p_m$ ist. Um das auch noch zu zeigen, betrachten wir das in **??** eingeführte Ideal $I_A = \{q \in$

$K[X] \mid q(A) = 0\}$ der algebraischen Identitäten von A. Es gilt

$$\begin{aligned}
I_A = \{q \in K[X] \mid q(A) = 0\} &= \left\{ q \in K[X] \mid q \begin{pmatrix} \boxed{C(p_1)} & & 0 \\ & \ddots & \\ 0 & & \boxed{C(p_m)} \end{pmatrix} = 0 \right\} \\
&= \left\{ q \in K[X] \mid \begin{pmatrix} \boxed{q(C(p_1))} & & 0 \\ & \ddots & \\ 0 & & \boxed{q(C(p_m))} \end{pmatrix} = 0 \right\} \\
&= \{q \in K[X] \mid q(C(p_1)) = 0, ..., q(C(p_m)) = 0\} \\
&= I_{C(p_1)} \cap ... \cap I_{C(p_m)} \\
&= (\mu_{C(p_1)}) \cap ... \cap (\mu_{C(p_m)}) \\
&= (p_1) \cap ... \cap (p_m) \overset{p_1|...|p_m}{=} (p_m)
\end{aligned}$$

und daher $\mu_A = p_m$.

Bemerkung 17.5.5. Wir haben insbesondere $\chi_A = (-1)^n p_1 \dots p_m \in (p_m) = I_A$ und daher einen weiteren Beweis des Satzes von Cayley-Hamilton **??** [→**??**]

Beispiel 17.5.6. Sei $K := \mathbb{Q}$, $A := \begin{pmatrix} 2 & 1 & 0 & 1 \\ 0 & 3 & 0 & 0 \\ -1 & 1 & 3 & 1 \\ -1 & 1 & 0 & 4 \end{pmatrix}$.

$$\begin{aligned}
&A - XI_4 \\
&= \begin{pmatrix} 2-X & 1 & 0 & 1 \\ 0 & 3-X & 0 & 0 \\ -1 & 1 & 3-X & 1 \\ -1 & 1 & 0 & 4-X \end{pmatrix} \sim \begin{pmatrix} -1 & 1 & 0 & 4-X \\ 0 & 1+(2-X) & 0 & 1+(2-X)(4-X) \\ 0 & 3-X & 0 & 0 \\ 0 & 0 & 3-X & 1-(4-X) \end{pmatrix} \\
&\sim \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & X-3 & 0 & -X^2+6X-9 \\ 0 & 0 & 0 & -X^2+6X-9 \\ 0 & 0 & X-3 & -X+3 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & X-3 & 0 & 0 \\ 0 & 0 & X-3 & 0 \\ 0 & 0 & 0 & X^2-6X+9 \end{pmatrix} := S
\end{aligned}$$

Wegen $3^2 - 6 \cdot 3 + 9 = 0$ gilt $(X-3) \mid X^2 - 6X + 9$ und S ist in Smithscher Normalform. Es gilt also $c(A - XI_4) = (1, X-3, X-3, X^2-6X+9)$ und daher lautet die Frobenuis'sche

Normalform von $A \begin{pmatrix} \boxed{3} & 0 & 0 & 0 \\ 0 & \boxed{3} & 0 & 0 \\ 0 & 0 & \boxed{\begin{matrix} 0 & -9 \\ 1 & 6 \end{matrix}} \end{pmatrix}$ und es gilt $\mu_A = X^2 - 6X + 9 = (X-3)^2$ und $\chi_A = (X-3)(X-3)(X^2-6X+9) = (X-3)^4$.

17.6 Die Normalform von Weierstraß

[Karl Theodor Wilhelm Weierstraß *1815 †1897]

Wieder sei stets K ein Körper.

Lemma 17.6.1. Sei $m \in \mathbb{N}_0$ und $p_1, ..., p_m \in K[X]$ normiert mit $\gcd(p_i, p_j) = 1$ für alle $i, j \in \{1, ..., m\}$ bei $i \neq j$. Dann ist

$$C(p_1 \dots p_m) \approx \begin{pmatrix} \boxed{C(p_1)} & & 0 \\ & \ddots & \\ 0 & & \boxed{C(p_m)} \end{pmatrix}.$$

Beweis. Setze $n := \deg(p_1 \dots p_m)$. Nach **??** ist die Behauptung äquivalent zu $C(p_1 \dots p_m) - XI_n \sim \begin{pmatrix} \boxed{C(p_1)} & & 0 \\ & \ddots & \\ 0 & & \boxed{C(p_m)} \end{pmatrix} - XI_n$, was nach **??** gleichbedeutend ist mit

$$\begin{pmatrix} 1 & & & 0 \\ & \ddots & & \\ & & 1 & \\ 0 & & & p_1 \dots p_m \end{pmatrix} \sim \begin{pmatrix} \boxed{\begin{smallmatrix} 1 & & \\ & \ddots & \\ & & p_1 \end{smallmatrix}} & & 0 \\ & \ddots & \\ 0 & & \boxed{\begin{smallmatrix} 1 & & \\ & \ddots & \\ & & p_m \end{smallmatrix}} \end{pmatrix}$$, das heißt mit

$$S := \begin{pmatrix} 1 & & & 0 \\ & \ddots & & \\ & & 1 & \\ 0 & & & p_1 \dots p_m \end{pmatrix} \sim \begin{pmatrix} 1 & & & & 0 \\ & \ddots & & & \\ & & 1 & & \\ & & & p_1 & \\ & & & & \ddots \\ 0 & & & & p_m \end{pmatrix} := T.$$

Œ sei $n \geq 2$. Es gilt $d_n(S) = \det(S) = p_1 \dots p_m = \det(T) = d_n(T)$ und $d_{n-1}(S) = 1 = d_{n-1}(T)$, denn $d_{n_1}(T)$ teilt $p_1 \dots p_{i-1} p_{i+1} \dots p_m$ für alle $i \in \{1, ..., m\}$. Wäre nämlich $d_{n-1}(T) \neq 1$, so gäbe es ein Primpolynom $g \in K[X]$ mit $g \mid d_{n-1}(T)$ und daher gilt $g \mid p_i$ für ein $i \in \{2, ..., m\}$, sowie $g \mid p_j$ für ein $j \in \{1, ..., m\} \setminus \{i\}$. Es folgt $g \mid \gcd(p_i, p_j) = 1$, was ein Widerspruch ist. Daher gilt $d(S) = (1, ..., 1, p_1 \dots p_m) = d(T)$ und somit $S \sim T$.

Satz und Definition 17.6.2. Sei $A \in K^{n \times n}$. Dann gibt es ein bis auf Reihenfolge der Einträge eindeutig bestimmtes Tupel $(p_1, ..., p_m)$ von normierten Polynomen $p_i \in K[X]$ vom Grad ≥ 1, die Potenzen von Primpolyonomen sind mit

$$A \approx \begin{pmatrix} \boxed{C(p_1)} & & 0 \\ & \ddots & \\ 0 & & \boxed{C(p_m)} \end{pmatrix}.$$

Die bis auf Reihenfolge der Kästchen eindeutig bestimmte Matrix heißt Weierstraß'sche Normalform von A.

Beweis. **Existenz**. Diese ist klar mit der Frobenius'schen Normalform **??**, Primfaktorzerlegung und Lemma **??**

Eindeutigkeit.

Behauptung. Sei $(p_1, ..., p_m)$ ein Tupel normierter $p_i \in K[X]$ vom Grad ≥ 1, die Potenzen von Primpolynomen sind. Dann gibt es eine Zerlegung $Z = \{I_1, ..., I_s\}$ von $\{1, ..., m\}$ mit $\#Z = s$ derart, dass mit $f_1, ..., f_s$ definiert durch $f_k := \prod_{i \in I_k} p_i$ $(k \in \{1, ..., s\})$ gilt:

$(*)_k$ $\gcd(p_i p_j) = 1$ für alle $i, j \in I_k$ (für alle $k \in \{1, ..., s\}$

$(**)_k$ $f_{k+1} \mid f_k$ (für alle $k \in \{1, ..., s-1\}$)

Begründung. Ist $m = 0$, so ist nichts zu tun. Sei also $m > 0$. Da jedes p_i eine Potenz von Primpolynomen ist gibt es dann $\emptyset \neq I_1 \subseteq \{1, ..., m\}$ derart, dass $f_1 := \prod_{i \in I_1} p_i = \operatorname{lcm}(p_1, ..., p_m)$ [wäre z.B. $K = \mathbb{C}$, $m = 5$, $p_1 = (X-1)^3, p_2 = X - 1, p_3 = X^5, p_4 = (X-1)^3, p_5 = X^7$, so wären $I_1 = \{1, 5\}$ und $I_1 = \{4, 5\}$ die beiden möglichen Wahlen für I_1]. Es gilt dann $(*_1)$. Ist $I_1 = \{1, ..., m\}$, so setzen wir $s := 1$ und sind fertig. Sonst gibt es $\emptyset \neq I_2 \subseteq \{1, ..., m\}$ derart, dass $f_2 := \prod_{i \in I_2} p_i = \operatorname{lcm}(p_i \mid i \in \{1, ..., m\} \setminus I_1)$. Es gilt dann $(**_1)$ und $(*_2)$. Ist $I_1 \cup I_2 = \{1, ...m\}$, so setze $s := 2$ und wir sind fertig. Sonst mache so weiter...

Seien nun $(p_1, ..., p_m)$ und $(q_1, ..., q_l)$ Tupel normierter Polynome $p_i, q_j \in K[X]$ vom Grad ≥ 1, die Potenzen von Primpolynomen sind und für welche die Behauptung gilt. Zu zeigen ist, nach Umnummerierung gilt $(p_1, ..., p_m) = (q_1, ..., q_l)$. Wähle gemäß Hilfsbehauptung Zerlegungen $Z = \{I_1, ..., I_s\}$ von $\{1, ..., m\}$ und $Z' = \{J_1, ..., J_t\}$ von $\{1, ..., l\}$ mit $\#Z = s$ und $\#Z' = t$ derart, dass mit $f_1, ...f_s, g_1, ..., g_t$ definiert durch $f_k := \prod_{i \in I_k} p_i$ $(k \in \{1, ..., s\})$ und $g_k := \prod_{j \in J_k} q_j$ $(k \in \{1, ..., t\})$ gilt:

- $\forall k \in \{1, ..., s\} : \forall i, j \in I_k : (i \neq j \implies \gcd(p_i, p_j) = 1$ und $f_s \mid ... \mid f_1$
- $\forall k \in \{1, ..., t\} : \forall i, j \in J_k : (i \neq j \implies \gcd(q_i, q_j) = 1$ und $g_t \mid ... \mid g_1$.

Nach Lemma **??** gilt nun

$$\begin{pmatrix} \boxed{C(f_s)} & & 0 \\ & \ddots & \\ 0 & & \boxed{C(f_1)} \end{pmatrix} \approx \begin{pmatrix} \boxed{C(p_1)} & & 0 \\ & \ddots & \\ 0 & & \boxed{C(p_m)} \end{pmatrix}$$

$$\approx \begin{pmatrix} \boxed{C(q_1)} & & 0 \\ & \ddots & \\ 0 & & \boxed{C(q_l)} \end{pmatrix} \approx \begin{pmatrix} \boxed{C(g_t)} & & 0 \\ & \ddots & \\ 0 & & \boxed{C(g_t)} \end{pmatrix}.$$

Wegen der Eindeutigkeit der Frobenius'schen Normalform aus **??** folgt $(f_s, ..., f_1) = (g_t, ..., g_1)$, insbesondere $s = t$. Wegen der Eindeutigkeit der Primfaktorzerlegung in $K[X]$ folgt, dass nach Umnummerieren gilt $(p_1, ..., p_m) = (q_1, ..., q_l)$.

Beispiel 17.6.3. (a) Für die in **??** betrachtete Matrix $A \in \mathbb{Q}^{4\times 4}$ galt $c(A - XI_4) = (1, X-3, X-3, X^2-6X+9)$. Da $X^2-6X+9 = (X-3)^2$ eine Potenz eines Primpolynoms ist, stimmt für A die Frobenius'sche Normalform mit der Weierstraß'schen Normalform überein.

(b) Die Matrix $A := \left(\begin{array}{c|ccc} 1 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 1 \end{array}\right) \in \mathbb{R}^{4\times 4}$ ist nach **??** in Frobenius'scher Normalform und $c(A - XI_4) = (1, 1, X-1, X^3 - X^2 + X - 1)$, wobei $X^3 - X^2 + X - 1 = (X-1)(X^2+1)$. Da $X-1$ und X^2+1 Primpolynome sind, ist die Weierstraß'sche Normalform von A gegeben durch $A \approx \left(\begin{array}{c|c|cc} 1 & 0 & 0 & 0 \\ \hline 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{array}\right)$.

17.7 Die Normalform von Jordan

[Marie Ennemand Camille Jordan *1838 †1922]

Wieder sei stets K ein Körper.

Definition 17.7.1. Sei $\lambda \in K$ und $n \in \mathbb{N}$. Dann nennt man

$J(\lambda, n) := \begin{pmatrix} \lambda & & & \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ & & 1 & \lambda \end{pmatrix} \in K^{n\times n}$ das Jordankästchen zum Eigenwert λ der Größe n [insbesondere $J(\lambda, 1) = (\lambda)$ falls $n = 1$, und beachte, dass $J(\lambda, n)$ den Eigenwert λ hat, denn $\chi_{J(\lambda,n)} = (\lambda - X)^n$].

Bemerkung 17.7.2. (vergleiche §17.5) Sei $n \in \mathbb{N}$ und $\lambda \in K$. Setze $p := (X-\lambda)^n = \sum_{k=0}^{n} \binom{n}{k} (-k)^{n-k} X^k$. Wie gehabt ist $\underline{v} := (\overline{1}, \overline{X}, ..., \overline{X^{n-1}})$ eine Basis von $K[X]/(p)$ und die lineare Abbildung $f : \begin{cases} K[X]/(p) \to K[X]/(p) \\ \overline{q} \mapsto \overline{Xq} \end{cases}$ $(q \in K[X])$ wohldefiniert. Die Begleitmatrix $C(p) = M(f, \underline{v})$ von p ist leider etwas kompliziert, da in der letzten Spalte die Koeffizienten von p eingehen. Daher betrachtete man die Basis $\underline{w} = (\overline{1}, \overline{X-\lambda}, ..., \overline{(X-\lambda)^{n-1}})$ von $K[X]/(p)$ bezüglich derer gilt $M(f, \underline{v}) = J(\lambda, n)$. Wegen

$$\begin{aligned} C(p) \overset{??}{=} M(f, \underline{v}) &\overset{??}{=} M(\underline{w}, \underline{v}) M(f, \underline{w}) M(\underline{v}, \underline{w}) \\ &= M(\underline{v}, \underline{w})^{-1} M(f, \underline{w}) M(\underline{v}, \underline{w}) \approx M(f, \underline{w}) = J(\lambda, n) \end{aligned}$$

gilt $C(p) \approx J(\lambda, n)$. Es ist also $J(\lambda, n)$ ein schöner Ersatz für $C(p)$ im Fall $p = (X-\lambda)^n$. Wie in **??** gilt $\mu_{J(\lambda,n))} = \mu_f = (X-\lambda)^n$ und $\chi_f = \chi_{J(\lambda,n)} = (-1)^n(X-\lambda)^n$.

Satz und Definition 17.7.3. Sei $A \in K^{n\times n}$ derart, dass χ_A zerfällt [→**??**,**??**]. Dann gibt es ein bis auf Reihenfolge der Einträge eindeutig bestimmtes Tupel $((\lambda_1, k_1), ..., (\lambda_m, k_m))$ von Paaren $(\lambda_i, k_i) \in K \times \mathbb{N}$ mit

$$A \approx \begin{pmatrix} \boxed{J(\lambda_1, k_1)} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \boxed{J(\lambda_m, k_m)} \end{pmatrix}.$$

Die bis auf Reihenfolge der Kästchen eindeutig bestimmte rechte obige Matrix heißt Jordansche Normalform von A.

Beweis. Beachte, dass

$$\begin{pmatrix} \boxed{J(\lambda_1, k_1)} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \boxed{J(\lambda_1, k_1)} \end{pmatrix} \approx \begin{pmatrix} \boxed{C((X-\lambda_1)^{k_1})} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \boxed{C((X-\lambda_m)^{k_m})} \end{pmatrix}.$$

Existenz. Bringe A auf die Weierstraß'sche Normalform aus §17.6, d.h. wähle ein Tupel $(p_1, ..., p_m)$ von normierten Polynomen $p_i \in K[X]$ vom Grad ≥ 1, die Potenzen von Primpolynomen sind mit $A \approx \begin{pmatrix} \boxed{C(p_1)} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \boxed{C(p_m)} \end{pmatrix}$. Mit $n_i := \deg p_i$ gilt dann

$$(-1)^{n_1+...+n_m} p_1 \dots p_m \overset{??}{=} \prod_{i=1}^{m} \deg(C(p_i) - XI_n) = \deg(A - XI_n) = \chi_A.$$

Mit χ_A zerfällt daher auch jedes p_i. Daher gibt es zu jedem $i \in \{1, ..., m\}$ ein Paar

$(\lambda_i, k_i) \in K \times \mathbb{N}$ mit $p_i = (X - \lambda_i)^{k_i}$.

Eindeutigkeit. Diese wird sofort klar aus der Eindeutigkeit der Weierstraß'schen Normalform.

Beispiel 17.7.4. (a) Für die in **??** und **??**(a) betrachtete Matrix $A \in \mathbb{Q}^{4\times 4}$ galt $c(A - XI_4) = (1, X-3, X-3, (X-3)^2)$. Daher ist $\left(\begin{array}{c|c|cc} 3 & 0 & 0 & 0 \\ \hline 0 & 3 & 0 & 0 \\ \hline 0 & 0 & 3 & 0 \\ 0 & 0 & 1 & 3 \end{array}\right) \approx A$ die Jordansche Normalform von A.

(b) Ist $A \in \mathbb{R}^{8\times 8}$ mit $c(A-XI_8) = (1,1,1,1,1,1,(X-1)^2(X-2),(X-1)^2(X-2)(X-3))$, so ist $\left(\begin{array}{cc|c|ccc|c|c} 1 & 0 & & & & & & \\ 1 & 1 & & & & & & \\ \hline & & 2 & & & & & \\ \hline & & & 1 & 0 & 0 & & \\ & & & 1 & 1 & 0 & & \\ & & & 0 & 1 & 1 & & \\ \hline & & & & & & 2 & \\ \hline & & & & & & & 3 \end{array}\right) \approx A$ die Jordansche Normalform von A, $\mu_A = (X-1)^3(X-2)(X-3)$ und $\chi_A = (X-1)^5(X-2)^2(X-3)$.

Skript zur linearen Algebra I & II

Index