Lineare Algebra I

Wintersemester 2017

Mohamed Barakat

Department Mathematik, Universität Siegen
mohamed.barakat@uni-siegen.de

Stand: 18. April 2018

Vorwort

Dies ist die geTEXte Version meiner Vorlesungsnotizen, die ich fortlaufend aktualisieren werde. Habt bitte Verständnis dafür, wenn Stand der Vorlesung und der Notizen nicht immer übereinstimmen werden. Daher gilt: Kommt zur Vorlesung und macht Eure eigenen Notizen. *Die sind sowieso besser als jedes Skript*. Die Form eines Skriptes erreichen diese Notizen vermutlich erst gegen Ende der Vorlesung, dies kann ich aber nicht garantieren. Die aktuelle Version ist unter der folgenden Adresse zu finden:

https://algebra.mathematik.uni-siegen.de/barakat/Lehre/WS17/LAI/Skript/LA_I.pdf

Als Vorlage benutz(t)e ich zum Teil das online-verfügbare Skript von Prof. Gabriele Nebe, das sie mir freundlicherweise zur Verfügung gestellt hat.

Für Korrektur- und Verbesserungsvorschläge bin ich stets dankbar
mohamed.barakat@uni-siegen.de

Inhaltsverzeichnis

Kapitel 0

Mathematische Grundlagen

0.1 Aussagen

Diese Sektion ist eine knappe und informelle Einführung in die mathematische Logik. Wir werden die nicht-konstruktive Logik kennenlernen, die mit der Mengenlehre eng verzahnt ist. Sie ist die sogenannte interne Logik der klassischen Mengenlehre.

0.1.a Definition und Beispiele

Definition 0.1.1. Eine **(mathematische) Aussage** ist ein sprachlicher Ausdruck, der einen eindeutigen **Wahrheitswert** besitzt, welcher entweder **wahr** oder **falsch** lauten kann[1]. Eine Aussage kann ggf. Formeln und Symbole enthalten.

Wir benutzen die Abkürzungen w und f für wahr und falsch. In der Literatur werden auch oft die Symbole $\top$ und $\bot$ für wahr und falsch benutzt.

Beispiel 0.1.2. Die Ausdrücke 'wahr' und 'falsch' sind selber Aussagen mit den jeweiligen Wahrheitswerten. Folgende sprachliche Ausdrücke sind mathematische Aussagen:

- '$1+1=2$' (w)
- '$1+1=3$' (f)
- 'Es gibt unendlich viele Primzahlen.' (w)
- 'Für jede reelle Zahl y gibt es eine reelle Zahl x mit $y=x^2$.' (f)
- 'Jede gerade ganze Zahl, die größer als 2 ist, ist die Summe aus zwei Primzahlen.' (unbekannt[2])

Aufgrund fehlender bzw. unvollständiger Spezifikation sind folgende Ausdrücke keine Aussagen:

- 'Aachen ist cool.'
- 'a+b=c'

[1] auch wenn der Wahrheitswert noch nicht bekannt ist

[2] Die Goldbachsche Vermutung vom Jahr 1742 besagt, dass diese Aussage wahr ist. Sie ist bereits für 18-stellige Zahlen mit dem Computer verifiziert worden.

0.1.b Zusammensetzung und Verneinung

Definition 0.1.3. Für beliebige Aussagen A und B definieren wir die Wahrheitswerte für folgende **zusammengesetzte Aussagen**:

(1) 'Nicht A': Die **Verneinung** (oder **Negation**) $\neg A$ ist genau dann wahr, wenn A falsch ist.

(2) 'A und B': Die **Konjunktion** $A \wedge B$ ist genau dann wahr, wenn A und B wahr sind.

(3) 'A oder B': Die **Disjunktion** $A \vee B$ ist genau dann wahr, wenn A oder B (oder beide) wahr sind.

(4) 'Entweder A oder B': Das **exklusive oder** $A \veebar B$ ist genau dann wahr, wenn entweder A oder B wahr ist (aber nicht beide wahr sind).
Andere Sprechweisen: 'A x-or B' bzw. 'A x-oder B'.

(5) 'A impliziert B': Die **Implikation** $A \implies B$ ist genau dann falsch, wenn A wahr ist und B nicht.
Andere Sprechweisen: 'aus A folgt B' bzw. 'wenn A dann B'.

(6) 'A ist äquivalent zu B': Die **Äquivalenz** $A \iff B$ ist genau dann wahr, wenn A und B den gleichen Wahrheitswert besitzen.
Andere Sprechweisen: 'A gilt genau dann, wenn B gilt'.

Die obigen Definitionen fassen wir in einer **Wahrheitstabelle** zusammen:

A	B	$\neg A$	$A \wedge B$	$A \vee B$	$A \veebar B$	$A \implies B$	$A \iff B$
w	w	f	w	w	f	w	w
w	f	f	f	w	w	f	f
f	w	w	f	w	w	w	f
f	f	w	f	f	f	w	w

Die Symbole $\neg$, $\wedge$, $\vee$, $\veebar$, $\implies$, $\iff$ werden **Junktoren** genannt, der erste ist 1-stellig und die restlichen 2-stellig. Die konstanten Aussagen w und f (bzw. $\top$ und $\bot$) können als 0-stellige Junktoren angesehen werden. Aufgrund der fehlenden Operatorrangfolge benutzen wir Klammern, um etwa $(A \wedge B) \vee C$ von $A \wedge (B \vee C)$ zu unterscheiden.

Beispiel 0.1.4.

(1) Die Verneinung von '$1+1=2$' ist: 'Es gilt nicht, dass $1+1=2$ ist'. Eine kürzere Form wäre '$1+1$ ist ungleich 2' oder rein symbolisch '$1+1 \neq 2$'. Diese negierte Aussage ist falsch.

(2) Die Negation von 'Das Glas ist voll' ist 'Das Glas ist nicht voll'. Beachte: 'Das Glas ist leer' ist nicht die Negation von 'Das Glas ist voll'.

(3) Die Verneinung von 'Alle Gläser sind voll' ist 'Nicht alle Gläser sind voll', oder gleichbedeutend 'Es gibt ein Glas, das nicht voll ist'.

(4) Folgende Aussage ist wahr: Wenn $1+1=3$, dann ist $1+1=2$.

(5) Folgende Aussage ist wahr: Wenn $1+1=3$, dann ist $1+1=5$.

0.1.c Aussageformen und Prädikatenlogik

Dies ist ein Miniabstecher in die Prädikatenlogik.

Definition 0.1.5. Eine **Aussageform** (oder **Prädikat**) ist ein sprachlicher Ausdruck, der endlich viele **ungebundene** (oder **freie**) **Individuenvariablen** enthält, und der für jede **Belegung** aller vorkommenden Individuenvariablen mit konkreten Objekten zu einer Aussage wird.

Beispiel 0.1.6.

- '$a+b=c$' ist eine Aussageform. Werden die Variablen a, b, c etwa mit (reellen) Zahlen belegt, so entsteht eine Aussage (mit eindeutigem Wahrheitswert).
- 'x ist an der RWTH Aachen eingeschrieben' ist eine Aussageform. Wird die Variable x etwa mit einem beliebigen Menschen belegt, so entsteht eine Aussage.
- 'Für jede reelle Zahl y gibt es eine reelle Zahl x mit $y = x^2$.' ist dagegen bereits eine (falsche) Aussage, da die Variablen x und y durch die **logischen Quantoren** 'für alle' bzw. 'für jeden' (Allquantor $\forall$) und 'es existiert' bzw. 'es gibt' (Existenzquantor $\exists$) bereits gebunden wurden.

Bemerkung 0.1.7. Eine Aussageform ist selbst *keine* Aussage. Die Zusammensetzung von Aussageformen mittels Junktoren ist wieder eine Aussageform.

Beispiel 0.1.8. Seien $A(t)$ die Aussageform 'Der Projektor im Hörsaal TEMP 2 ist zum Zeitpunkt t aus' und $B(t)$ die Aussageform 'Im Hörsaal TEMP 2 findet zum Zeitpunkt t keine Vorlesung statt'. Dann ist auch '$A(t) \implies B(t)$' eine Aussageform. Für jede Belegung der Variable t mit einem Zeitpunkt erhalten wir eine Aussage, deren Wahrheitswert von t abhängt. Wann ist sie falsch?

0.1.d Logische Äquivalenzen und Tautologien

Definition 0.1.9.

(1) Ein **logischer Term** ist ein Ausdruck bestehend aus endlich vielen *Aussagevariablen* $A, B, \ldots$, die mit Junktoren $w, f, \neg, \wedge, \ldots$ verknüpft sind. Durch die **Belegung** der Aussagevariablen mit Wahrheitswerten erhält der Term selbst einen Wahrheitswert.

(2) Zwei logische Terme S und T, definiert auf derselben Menge von Aussagevariablen, heißen **logisch äquivalent** (geschrieben $S \equiv T$), wenn S und T denselben Wahrheitswert für jede Belegung der Aussagevariablen mit Wahrheitswerten haben.

(3) Eine logischer Term T heißt **Tautologie**, falls $T \equiv w$.

(4) Eine logischer Term W heißt **Widerspruch**, falls $W \equiv f$.

Beispiel 0.1.10.

(1) $A \veebar B \equiv (A \wedge \neg B) \vee (\neg A \wedge B)$.
Wir sagen daher, das $\veebar$ durch $\neg, \wedge, \vee$ ausgedrückt werden kann.

(2) $A \implies B \equiv \neg(A \wedge \neg B) \equiv \neg A \vee B$.

(3) $A \iff B \equiv (A \implies B) \wedge (B \implies A) \equiv \neg(A \veebar B)$.

Beispiel 0.1.11. Wichtige Tautologien sind:

(1) Modus Ponens:
$$(A \wedge (A \implies B)) \implies B$$

(2) Tertium non datur (Gesetz des ausgeschlossenen Dritten):
$$A \vee \neg A$$

(3) de Morgan Gesetze:
$$\neg(A \wedge B) \iff (\neg A \vee \neg B)$$
$$\neg(A \vee B) \iff (\neg A \wedge \neg B)$$

(4) Kontraposition:
$$(A \implies B) \iff (\neg B \implies \neg A)$$

Bemerkung 0.1.12. Seien S, T logische Terme. Dann gilt $S \equiv T$ genau dann, wenn $S \iff T$ eine Tautologie ist.

Bemerkung 0.1.13. Tautologien können als Beweisstrategien benutzt werden. Möchte man etwa $A \implies B$ zeigen, so kann man nach der Kontraposition anstelle dessen $\neg B \implies \neg A$ zeigen.

0.1.e Sprachliche Konventionen

(1) Das Wort 'ein' bedeutet immer 'mindestens ein'. Wenn 'genau ein' gemeint ist, dann muss dies explizit gesagt werden.

(2) In einer Aufzählung von Objekten $x_1, \ldots, x_n$ heißen $x_1, \ldots, x_n$ **paarweise verschieden**, wenn keine zwei Objekte der Aufzählung gleich sind. Davon zu unterscheiden ist 'verschieden' im Sinne von 'nicht alle gleich'.

(3) Es gibt diverse Sprechweisen für das Wort impliziert:

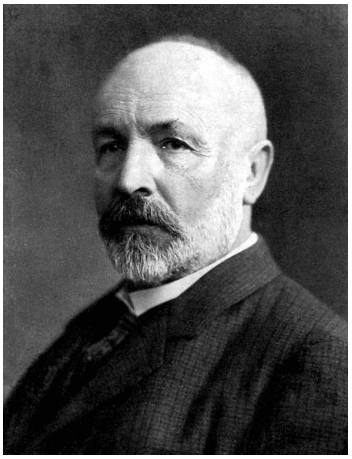
Term	Sprechweise
$A \implies B$	A impliziert B.
	Wenn A, dann B.
	Sei A. Dann B.
	Aus A folgt B.
	Sei A. Daraus folgt B.
	B dann, wenn A.
	B, falls A.
	Sei A. Insbesondere B.
	A vorausgesetzt, dann B.
	A ist eine hinreichende Bedingung für B.
	B ist eine notwendige Bedingung für A.
	A nur dann, wenn B.

Ein suggestives Beispiel solcher Aussagen wäre:
A = 'Er besteht die Klausur'
B = 'Er hat den Stoff verstanden'.

0.2 Mengen

0.2.a Definition und Beispiele

Georg Kantor gilt als Begründer der Mengenlehre.



> Unter einer "Menge" verstehen wir jede Zusammenfassung M von bestimmten wohlunterschiedenen Objekten m unserer Anschauung oder unseres Denkens (welche die "Elemente" von M genannt werden) zu einem Ganzen.
>
> — Georg Cantor, 1895

Schränkt man den Begriff der "Zusammenfassung" nicht weiter ein, so darf z.B. eine Menge sich selbst enthalten. Bertrand Russel hat 1901 eingesehen, dass wenn man wiederum alle Mengen, die sich nicht enthalten, zu einer Menge zusammenfasst, so einen offensichtlichen Widerspruch erhält. Dessen ungeachtet einigen wir uns auf folgende (naive) Definition.

Definition 0.2.1. Eine Menge M ist etwas, zu dem jedes beliebige Objekt x entweder **Element** der Menge ist (geschrieben $x \in M$) oder nicht (geschrieben $x \notin M$).

Mengen sind also dadurch gekennzeichnet, dass '$x \in M$' für jedes konkrete Objekt x eine Aussage ist, bzw. dass '$x \in M$' eine Aussageform ist. Umgekehrt ist für jede Aussageform $A(x)$ die Zusammenfassung aller x, für die $A(x)$ wahr ist, eine Menge (vgl. Schreibweise (c) unten).

Bemerkung 0.2.2. Zermelo und Fraenkel haben die obige naive Definition der Mengenlehre durch ein Axiomensystem (ZF) ersetzt, welches wir in dieser Vorlesung aus Zeitgründen auslassen müssen. Eine Menge im Sinne von ZF darf sich nicht enthalten.

Definition 0.2.3. Seien M, N zwei Mengen. Die Menge N heißt eine **Teilmenge** von M und M eine **Obermenge** von N (geschrieben $N \subset M$), wenn für alle $x \in N$ gilt: $x \in M$. Das Zeichen $\subset$ heißt **Inklusion**. Die Mengen M und N heißen **gleich** (geschrieben $M = N$), wenn $M \subset N$ und $N \subset M$.

Eine Menge M heißt **endlich**, wenn M nur endlich viele Elemente besitzt. Man schreibt in diesem Fall $|M|$ für die Anzahl der Elemente von M. Man nennt $|M|$ die Mächtigkeit von M. Andernfalls heißt M **unendlich** und man schreibt $|M| = \infty$.

Schreibweise.

(1) Aufzählen: Die Elemente werden zwischen geschweiften Klammern aufgelistet. Reihenfolge und Wiederholung spielen dabei keine Rolle, z.B. ist

$$\{1, 2, 3\} = \{2, 1, 3\} = \{1, 3, 2, 2, 3\}.$$

(2) Beschreiben: Mengen können durch Worte beschrieben werden, etwa:

$$\text{Menge der natürlichen Zahlen} = \{1, 2, 3, 4, 5, \ldots\}$$
$$\text{Menge der ganzen Zahlen} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$$

(3) Aussondern: Sei M eine Menge und $A(x)$ eine Aussageform, so bezeichnet

$$\{x \in M \mid A(x)\}$$

diejenige Teilmenge von M, die aus allen Elementen besteht, für die $A(x)$ wahr ist (gesprochen 'Die Menge aller x aus M mit $A(x)$').

Bezeichnen wir wie üblich die Menge der natürlichen mit $\mathbb{N}$, so ist

$$\{n \in \mathbb{N} \mid n \text{ ist ungerade}\}$$

die Menge der ungeraden natürlichen Zahlen, also $\{1, 3, 5, 7, \ldots\}$.

(4) Abbilden: Ist M eine Menge und $e(x)$ für jedes $x \in M$ ein Objekt, so bezeichnet

$$\{e(x) : x \in M\}$$

die Menge aller Objekte $e(x)$, wobei x alle Elemente der Menge M durchläuft.

Z.B. ist $\{n^2 : n \in \mathbb{N}\}$ die Menge aller natürlichen Quadratzahlen.

Abbilden und Aussondern können kombiniert werden: $\{n^2 : n \in \mathbb{N} \mid n \text{ ungerade}\}$ ist somit die Menge $\{1, 9, 25, 49, ...\}$ aller Quadrate von ungeraden natürlichen Zahlen.

Bemerkung 0.2.4. Eine Kombination von : und | wird häufig abgekürzt. Z.B. würde man für die obige Menge eher $\{n^2 \mid n \in \mathbb{N}, n \text{ ungerade}\}$ schreiben.

Beispiel 0.2.5. Häufig auftretende Mengen sind:

Symbol	Beschreibung	Definition
$\emptyset$	leere Menge	$\{\,\}$
$\underline{n}$	n-elementige Menge	{1,2,...,n}
$\mathbb{N}$	natürliche Zahlen	{1,2,3,...}
$\mathbb{N}_0$	natürliche Zahlen einschließlich 0	$\{0, 1, 2, 3, \ldots\}$
$\mathbb{P}$	Primzahlen	$\{2, 3, 5, 7, 11, 13, \ldots\}$
$\mathbb{Z}$	ganze Zahlen	$\{\ldots, -2, -1, 0, 1, 2, \ldots\}$
$\mathbb{Q}$	rationale Zahlen	$\{\frac{a}{b} : a \in \mathbb{Z}, b \in \mathbb{N}\}$
$\mathbb{R}$	reelle Zahlen	$\{\pm a_1 \ldots a_r, b_1 b_2 \ldots : a_i, b_i \in \{0, 1, \ldots, 9\}\}$
$\mathbb{R}_{>0}$	positive reelle Zahlen	$\{x \in \mathbb{R} \mid x > 0\}$
$\mathbb{R}_{\geq 0}$	nicht-negative reelle Zahlen	$\{x \in \mathbb{R} \mid x \geq 0\}$
$\mathbb{C}$	komplexe Zahlen	$\{a + bi \mid a, b \in \mathbb{R}\}$

Nur die ersten beiden Mengen der Tabelle sind endlich: $|\emptyset| = 0$ und $|\underline{n}| = n$ für alle $n \in \mathbb{N}_0$. Es gilt:

$$\emptyset = \underline{0} \subset \underline{1} \subset \underline{2} \subset \cdots \subset \mathbb{N} \subset \mathbb{N}_0 \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}.$$

Übung 0.2.1 (Das Russelsche Paradoxon)**.** Man betrachte die „Menge"

$$\mathcal{M} \coloneqq \{N \mid N \text{ Menge und } N \notin N\},$$

d.h. die Menge aller Mengen, die sich selbst als Element *nicht* enthalten. Gilt $\mathcal{M} \in \mathcal{M}$ oder $\mathcal{M} \notin \mathcal{M}$?

0.2.b Quantifizierte Aussagen

Sei $A(x)$ eine Aussageform. Setzt man in $A(x)$ für x ein konkretes Objekt ein, so sagt man, x wird **spezifiziert**; es entsteht eine Aussage. Wie wir in der ersten Vorlesung erwähnt haben, haben wir durch die Quantifizierung über x zwei weiteren Möglichkeiten aus $A(x)$ eine Aussage zu machen:

'Für alle $x \in M$ gilt $A(x)$' bzw. 'Es gibt ein $x \in M$, für das $A(x)$ gilt'.

In der ersten Vorlesung konnten wir nicht genauer darauf eingehen, da wir noch keinen Mengenbegriff hatten.

Beispiel 0.2.6.

(1) Sei $A(x)$ die Aussageform '$x > 1$'. Dann ist 'Es existiert ein $x \in \mathbb{N}$ mit $A(x)$' wahr. Dagegen ist 'Für alle $x \in \mathbb{N}$ gilt $A(x)$' falsch.

(2) Sei $A(t)$ die Aussageform 'Zum Zeitpunkt t gilt: Projektor ist aus $\implies$ Vorlesung findet nicht statt'. Die Aussage wäre falsch, sobald es einen Zeitpunkt gibt, wo eine Dozentin oder ein Dozent Vorlesung hat und sie oder er den Projektor nicht braucht und ihn deswegen oder aus Versehen ausschaltet. Dann wäre auch die Aussage 'Es gibt einen Zeitpunkt t mit $\neg A(t)$' wahr und 'Für alle Zeitpunkte t gilt $A(t)$' falsch.

(3) Die Verneinung von 'Für alle $x \in M$ gilt $A(x)$' lässt sich als 'Es existiert $x \in M$ mit $\neg A(x)$' bzw. 'Es existiert $x \in M$ für das $A(x)$ nicht gilt' formulieren.

(4) Die Verneinung von 'Es existiert ein $x \in M$ mit $A(x)$' lässt sich als 'Für alle $x \in M$ gilt $\neg A(x)$' formulieren.

Übung 0.2.2.

(1) Was ist die Verneinung von 'Es gibt eine Person im Hörsaal, die ihr Handy aus hat'?

(2) Wie lautet der Wahrheitswert der Aussagen 'Für alle $x \in \emptyset$ gilt $A(x)$' und 'Es gibt $x \in \emptyset$ mit $A(x)$'?

0.2.c Konstruktion von Mengen

Definition 0.2.7 (Mengenoperationen)**.** Seien M, N Mengen.

(1) $M \cap N \coloneqq \{x \mid x \in M \wedge x \in N\}$ heißt der **Durchschnitt** von M und N.

(2) $M \cup N \coloneqq \{x \mid x \in M \vee x \in N\}$ heißt die **Vereinigung** von M und N.

(3) $M - N \coloneqq \{x \mid x \in M \wedge x \notin N\}$ heißt die **Differenzmenge**, gesprochen "M ohne N".

(4) $M \times N \coloneqq \{(x, y) \mid x \in M \wedge y \in N\}$ heißt **kartesisches Produkt** von M und N. Hierbei ist (x, y) ein **geordnetes Paar**. Zwei geordnete Paare (x, y) und (x', y') sind genau dann gleich, wenn $x = x'$ und $y = y'$ ist.

(5) $M^n \coloneqq \{(x_1, \ldots, x_n) \mid x_1, \ldots, x_n \in M\}$ heißt **n-faches kartesisches Produkt** von M ($n \in \mathbb{N}$). Hierbei ist $(x_1, \ldots, x_n)$ ein **n-Tupel** über M. Zwei n-Tupel sind genau dann gleich, wenn ihre i-ten **Einträge** gleich sind für $i = 1, \ldots, n$.

(6) $\mathrm{Pot}(M) \coloneqq \{S \mid S \subset M\}$ heißt die **Potenzmenge** von M.

Bemerkung 0.2.8. Folgende Tabelle deutet auf den Zusammenhang zwischen der klassischen Logik und der Cantorschen Mengenlehre:

Stelligkeit	Junktoren	Mengenoperationen
0	w f	$\mathcal{U} \coloneqq$ „Menge“ von allem $\emptyset$
1	$\neg A$	$\complement M \coloneqq \mathcal{U} - M$
2	$A \wedge B$ $A \vee B$ $A \veebar B$ $A \wedge \neg B$ $A \implies B$ $A \iff B$	$M \cap N$ $M \cup N$ $M \triangle N$ $M - N = M \cap \complement N$ $M \subset N$ $M = N$
beliebig	$\bigwedge_{i\in I} A_i, \quad \forall i \in I : A(i)$ $\bigvee_{i\in I} A_i, \quad \exists i \in I : A(i)$	$\bigcap_{i\in I} M_i$ $\bigcup_{i\in I} M_i$

Beispiel 0.2.9.

(1) Die leere Menge ist Teilmenge jeder beliebigen Menge (auch von sich selbst).

(2) Für jede Menge M gilt $M^2 = M \times M$.

(3) Ein Element in $\mathbb{R}^5$ ist z.B. $\left(1, \frac{1}{3}, 0, -2, \sqrt{5}\right)$.

(4) Es gilt:

$$\begin{aligned} \mathrm{Pot}(\emptyset) &= \{\emptyset\}, \\ \mathrm{Pot}(\{\emptyset\}) &= \{\emptyset, \{\emptyset\}\}, \\ \mathrm{Pot}(\{\emptyset, \{\emptyset\}\}) &= \{\emptyset, \{\emptyset\}, \{\{\emptyset\}\}, \{\emptyset, \{\emptyset\}\}\}, \\ &\vdots \end{aligned}$$

Übung 0.2.3.

(1) Wie viele Elemente hat $\mathrm{Pot}(n)$ für $n \in \mathbb{N}_0$?

(2) Wie viele Elemente hat M^n für $n \in \mathbb{N}$?

0.2.d Indexmengen

Definition 0.2.10. Für $n \in \mathbb{N}$ seien $a_1, \ldots, a_n$ Zahlen, $A_1, \ldots, A_n$ Aussagen, und $M_1, \ldots, M_n$ Mengen. Wir definieren:

(1) $\sum_{i=1}^n a_i \coloneqq a_1 + \cdots + a_n$.

(2) $\prod_{i=1}^n a_i \coloneqq a_1 \cdot \ldots \cdot a_n$.

(3) $\bigvee_{i=1}^n A_i \coloneqq A_1 \vee \ldots \vee A_n$.

(4) $\bigwedge_{i=1}^n A_i \coloneqq A_1 \wedge \ldots \wedge A_n$.

(5) $\bigcup_{i=1}^n M_i \coloneqq M_1 \cup \ldots \cup M_n$.

(6) $\bigcap_{i=1}^n M_i \coloneqq M_1 \cap \ldots \cap M_n$.

Die leere Summe ist definiert als 0 und das leere Produkt als 1. Die leere Disjunktion ist definiert als f und die leere Konjunktion als w. Die leere Vereinigung ist definiert als die leere Menge und der leere Schnitt als “die Menge von allem”.

Diese Schreibweise der Aufzählung kann teilweise auf beliebige **Indexmengen** I verallgemeinert werden, die auch unendlich sein dürfen:

Definition 0.2.11. Für jedes $i \in I$ sei M_i eine Menge.

(1) Definiere $\bigcup_{i\in I} M_i$ durch

$$x \in \bigcup_{i\in I} M_i :\Longleftrightarrow \text{ es gibt ein } i \in I \text{ mit } x \in M_i.$$

(2) Definiere $\bigcap_{i\in I} M_i$ durch

$$x \in \bigcap_{i\in I} M_i :\Longleftrightarrow \text{ für alle } i \in I \text{ gilt } x \in M_i.$$

Übung 0.2.4. Was sind $\bigcup_{i\in\emptyset} M_i$ und $\bigcap_{i\in\emptyset} M_i$?

0.2.e Mengenpartitionen

Definition 0.2.12.

(1) Zwei Mengen A, B heißen **disjunkt**, wenn $A \cap B = \emptyset$.

(2) Mengen M_i, $i \in I$, heißen **paarweise disjunkt**, wenn für alle $i, j \in I$ mit $i \neq j$ gilt: $M_i \cap M_j = \emptyset$.

(3) Sei $\mathcal{M}$ eine Menge von Mengen ($\mathcal{M}$ darf hier unendlich sein). Die Elemente von $\mathcal{M}$ heißen **paarweise disjunkt**, wenn je zwei davon disjunkt sind, d.h. wenn für $M, M' \in \mathcal{M}$ mit $M \neq M'$ gilt: $M \cap M' = \emptyset$.

(4) Sei M eine Menge. Eine **Partition** von M ist eine Menge $\mathcal{P}$ nicht-leerer, paarweise disjunkter Teilmengen von M mit $M = \bigcup_{C\in\mathcal{P}} C$. Die Elemente $C \in \mathcal{P}$ heißen die **Teile** der Partition.

Bemerkung 0.2.13. Für jede Partition $\mathcal{P}$ von M ist $\mathcal{P} \subset \mathrm{Pot}(M) - \{\emptyset\}$.

Beispiel 0.2.14.

(1) $\mathcal{P} \coloneqq \{\{n \in \mathbb{N} \mid n \text{ gerade}\}, \{n \in \mathbb{N} \mid n \text{ ungerade}\}\}$ ist eine Partition von $\mathbb{N}$ in zwei Teile.

(2) $\mathcal{P} \coloneqq \{\{n \in \mathbb{N} \mid n \text{ hat } k \text{ Dezimalstellen}\} \mid k \in \mathbb{N}\}$ ist eine Partition von $\mathbb{N}$ mit unendlich vielen Teilen.

(3) Die leere Menge hat nach der obigen Definition nur die Partition $\mathcal{P} = \emptyset$.

0.3 Beweisprinzipien

0.3.a Direkter Beweis

Prinzip. Das Ziel ist, die Wahrheit der Aussage $A \implies B$ **direkt** zu beweisen. Dafür nehmen wir an, dass A wahr ist und folgern daraus (mittels logischer Schlüsse), dass daraufhin B auch wahr ist.

Beispiel 0.3.1. Für alle $n \in \mathbb{N}$ gilt: n ungerade $\implies$ n^2 ungerade.

Beweis. Sei $n \in \mathbb{N}$ ungerade. D.h. es existiert ein $k \in \mathbb{N}$ mit $n = 2k - 1$. Dann ist $n^2 = (2k-1)^2 = 4k^2 - 4k + 1 = 2(2k^2 - 2k) + 1$ eine ungerade Zahl. □

0.3.b Indirekter Beweis durch Kontraposition

Prinzip. Das Ziel ist, die Wahrheit der Aussage $A \implies B$ durch **Kontraposition** zu beweisen. D.h. die Wahrheit der Aussage $\neg B \implies \neg A$ direkt zu beweisen.

Beweis des Prinzips. Dies ist zulässig aufgrund der Tautologie in 0.1.d

$$(A \implies B) \iff (\neg B \implies \neg A).$$ □

Beispiel 0.3.2. Für alle $n \in \mathbb{N}$ gilt: n^2 gerade $\implies$ n gerade.

Beweis. Sei $n \in \mathbb{N}$ ungerade, so ist nach dem obigen Beweis n^2 ungerade. □

0.3.c Indirekter Beweis durch Widerspruch

Prinzip. Das Ziel ist, die Wahrheit der Aussage A durch **Widerspruch** zu beweisen. D.h. die Wahrheit der Aussage $(\neg A \implies \text{f})$ direkt zu beweisen. Dabei taucht f typischerweise in Form eines Widerspruchs $B \wedge \neg B$ auf.

Beweis des Prinzips. Dies ist zulässig aufgrund der Tautologie

$$A \iff (\neg A \implies \text{f}).$$ □

Beispiel 0.3.3. $\sqrt{2} \notin \mathbb{Q}$.

Beweis. Sei $\sqrt{2} \in \mathbb{Q}$ (dies entspricht der Aussage $\neg A$). Dann gibt es einen gekürzten Bruch $\sqrt{2} = \frac{m}{n}$ mit $n, m \in \mathbb{N}$. Insbesondere sind n und m nicht beide gerade (dies entspricht der Aussage B). Dies impliziert $2n^2 = m^2$, d.h. m^2 ist gerade. Nach Beispiel 0.3.b ist auch m gerade. Also gibt es ein $k \in \mathbb{N}$ mit $m = 2k$. Dann gilt $2n^2 = m^2 = 4k^2$, also $n^2 = 2k^2$. D.h. n^2 ist gerade. Und wieder nach Beispiel 0.3.b ist auch n gerade. Somit folgt, dass n und m beide gerade sind (dies entspricht der Aussage $\neg B$). Widerspruch. □

0.3.d Vollständige Induktion

Prinzip. Das Ziel ist, die Wahrheit der Aussage 'Für alle $n \in \mathbb{N} : A(n)$' **per Induktion** zu beweisen. Dies bedeutet:

- **Induktionsanfang**: Beweise die Wahrheit der Aussage $A(1)$.
- **Induktionsschritt**: Beweise die Wahrheit der Aussage 'Für alle $n \in \mathbb{N} : A(n) \implies A(n+1)$'.

Im Induktionsschritt nennt man die Aussage $A(n)$ die **Induktionsvoraussetzung**.

Beweis des Prinzips. Der Beweis beruht auf folgender Eigenschaft von $\mathbb{N}$, die wir als gegeben annehmen:

> Sei $N \subset \mathbb{N}$ mit den zwei Bedingungen:
>
> - $1 \in N$.
> - Für alle $n \in \mathbb{N}$: $n \in N \implies n+1 \in N$.
>
> Dann ist $N = \mathbb{N}$.

Bei der vollständigen Induktion zeigen wir, dass die Menge $N := \{n \in \mathbb{N} \mid A(n) \text{ ist wahr}\}$ die zwei Bedingungen erfüllt und somit gleich $\mathbb{N}$ ist. □

Bemerkung 0.3.4. Der Induktionsanfang ist für die vollständige Induktion unerlässlich. Würde sie fehlen, könnte man falsche Aussagen wie 'Für alle $n \in \mathbb{N}$ gilt: $n + 1 = n + 2$' scheinbar beweisen.

Beispiel 0.3.5 (Gaußsche Summenformel)**.** Für alle $n \in \mathbb{N}$ gilt: $\sum_{i=1}^{n} i = \frac{n(n+1)}{2}$.

Beweis. Sei $A(n)$ die Aussageform $\sum_{i=1}^{n} i = \frac{n(n+1)}{2}$.

- Induktionsanfang: Die Aussage $A(1)$ ist wahr, da $\sum_{i=1}^{1} i = 1 = \frac{1 \cdot 2}{2}$.
- Induktionsschritt: Sei $n \in \mathbb{N}$ beliebig, so dass die Aussage $A(n)$ wahr ist. Dann ist

$$\begin{aligned}\sum_{i=1}^{n+1} i &= \sum_{i=1}^{n} i + (n+1) \overset{A(n)}{=} \frac{n(n+1)}{2} + (n+1) = \frac{n(n+1) + 2(n+1)}{2} \\ &= \frac{(n+1)(n+2)}{2}.\end{aligned}$$

  Somit ist die Aussage $A(n+1)$ wahr. □

Bemerkung 0.3.6. Es gibt weitere Varianten der vollständigen Induktion:

- Der Induktionsanfang kann bei $n_0 \in \mathbb{N}_0$ statt bei 1 gemacht werden. Damit wird die Aussage für alle $n \geq n_0$ gezeigt. Z.B. ist die obige Aussageform bereits ab $n_0 = 0$ wahr und die Aussageform $2^n \geq n^2$ erst ab $n_0 = 4$ wahr.
- Als Induktionsvoraussetzung könnte die evtl. stärkere Aussage $A(1) \wedge \ldots \wedge A(n)$ anstelle von $A(n)$ notwendig sein. Hierbei braucht man die Induktionsvoraussetzung nicht zu verschärfen.
- Als Induktionsvoraussetzung könnte die evtl. stärkere Aussage $A(n-1) \wedge A(n)$ anstelle von $A(n)$ notwendig sein. Hierbei muss man die Induktionsvoraussetzung auf $A(1)$ und $A(2)$ ausdehnen.

Übung 0.3.1. Sei $a_1 = 1$, $a_2 = 8$ und $a_n = a_{n-1} + 2a_{n-2}$ für $n \geq 3$. Beweisen Sie mit vollständiger Induktion, dass $a_n = 3 \cdot 2^{n-1} + (-1)^n \cdot 2$ für alle $n \in \mathbb{N}$ gilt.

Übung 0.3.2. Beweisen Sie mit vollständiger Induktion, dass jede natürliche Zahl, die größer als 1 ist, Produkt von Primzahlen ist.
Hinweis: Benutze die verschärfte Induktionsvoraussetzung $A(1) \wedge \ldots \wedge A(n)$.

Übung 0.3.3. Beweisen Sie mit vollständiger Induktion, dass $\mathrm{Pot}(\underline{n}) = 2^n$ für all $n \in \mathbb{N}$ gilt.

0.4 Abbildungen

0.4.a Definition und Beispiele

Definition 0.4.1. Seien M, N Mengen. Eine **Abbildung f von N nach M** ist eine "Vorschrift" (z.B. eine Formel), die jedem $x \in N$ genau ein Element $f(x) \in M$ zuordnet. Wir schreiben

$$f : N \to M,\ x \mapsto f(x).$$

- N heißt der **Definitionsbereich** von f,
- M der **Ziel-** oder **Wertebereich** von f,

- $f(x)$ das **Bild von** x unter f.

Zwei Abbildungen $f : N \to M$ und $g : N' \to M'$ sind nur dann gleich, wenn $N = N'$, $M = M'$ und $f(x) = g(x)$ für alle $x \in N$.

Die Menge aller Abbildungen von N nach M bezeichnen wir mit $\mathrm{Abb}(N, M)$ oder mit M^N.

Beispiel 0.4.2.

(1) $f : \mathbb{N} \to \mathbb{Q},\ i \mapsto i^2$.

(2) Sei P eine Menge von Personen. Wir definieren die Abbildung

$$L : P \to \mathbb{R},\ x \mapsto \text{Länge in cm von } x.$$

(3) Sei P eine Menge von Personen. Wir definieren die Abbildung

$$J : P \to \mathbb{Z},\ x \mapsto \text{Geburtsjahr von } x.$$

(4) Die Addition in $\mathbb{Z}$ kann als die Abbildung

$$\mathbb{Z} \times \mathbb{Z} \to \mathbb{Z},\ (x, y) \mapsto x + y$$

aufgefasst werden.

(5) Für jede Menge M gibt es die **Identitätsabbildung**

$$\mathrm{id}_M : M \to M,\ x \mapsto x.$$

(6) Betrachte die Abbildung

$$\begin{aligned} f :&\mathbb{R} \to \mathbb{R},\ x \mapsto \sqrt{x^2}, \\ g :&\mathbb{R} \to \mathbb{R},\ x \mapsto |x|, \\ h :&\mathbb{R} \to \mathbb{R}_{\geq 0},\ x \mapsto |x|. \end{aligned}$$

Dann ist $f = g \neq h$.

(7) $\mathrm{Abb}(\mathbb{R}, \mathbb{R}) = \mathbb{R}^{\mathbb{R}} =$ Menge aller reellen Funktionen.

(8) Für jede Menge M existiert *genau eine* Abbildung $\emptyset \to M$.

(9) Für jede nicht-leere Menge N existiert *keine* Abbildung $N \to \emptyset$.

(10) Für jede Menge N existiert *genau eine* Abbildung $N \to \{\emptyset\}$.

Bemerkung 0.4.3.

(1) Eine Abbildung $a : \mathbb{N} \to M$ wird auch **Folge in** M genannt. Oft benutzt man für Folgen die Schreibweise $a_1, a_2, a_3, \ldots$ oder $(a_i)_{i \in \mathbb{N}}$, wobei a_i für das Bild $a(i) \in M$ steht. Die erste Abbildung aus dem letzten Beispiel würde als $1, 4, 9, 16, \ldots$ oder als $(i^2)_{i \in \mathbb{N}}$ geschrieben.

Die Menge aller Folgen in M wird daher mit $M^{\mathbb{N}}$ bezeichnet. Z.B. ist $2^{\mathbb{N}}$ die Menge aller Binärfolgen, $\mathbb{R}^{\mathbb{N}}$ die Menge aller reellen Folgen, etc.

(2) Ein n-Tupel $(x_1, \ldots, x_n)$ über M kann als die Abbildung $t : \underline{n} \to M,\ i \mapsto x_i$ aufgefasst werden.

Z.B. kann das 5-Tupel $\left(1, \frac{1}{3}, 0, -2, \sqrt{5}\right)$ über $\mathbb{R}$ als die Abbildung $t : \underline{5} \to \mathbb{R}$ mit $t(1) = 1, t(2) = \frac{1}{3}, t(3) = 0, t(4) = -2, t(5) = \sqrt{5}$ aufgefasst werden.

Übung 0.4.1. Seien M und N endliche Menge. Beweisen Sie $\left|M^N\right| = |M|^{|N|}$.

0.4.b Bild, Urbild und Faser

Definition 0.4.4. Sei $f : N \to M$ eine Abbildung.

(1) Für jede Teilmenge $X \subset N$ heißt

$$f(X) \coloneqq \{f(x) : x \in X\}$$

das **Bild von X unter f**.

(2) Das Bild $f(N)$ von N unter f wird das **Bild** oder die **Bildmenge** von f genannt.

(3) Für jede Teilmenge $Y \subset M$ heißt

$$f^{-1}(Y) \coloneqq \{x \in N \mid f(x) \in Y\}$$

das **Urbild von Y unter f**.

(4) Das Urbild $f^{-1}(\{y\})$ einer einpunktigen Menge $\{y\} \subset M$ heißt die **Faser von f über y**.

In dieser Schreibweise ist

$$f^{-1} : \mathrm{Pot}(M) \to \mathrm{Pot}(N).$$

Sie darf daher nicht mit der Umkehrabbildung, falls sie existiert, verwechselt werden.

Beispiel 0.4.5. Die Faser der Abbildung L über $180 \in \mathbb{R}$ aus dem vorigen Beispiel 0.4.a sind alle Personen in der Menge P, die 180 cm groß sind.

Bemerkung 0.4.6. Die nicht-leeren Fasern einer Abbildung bilden eine Partition des Definitionsbereichs.

0.4.c Injektive, surjektive und bijektive Abbildungen

Definition 0.4.7. Sei $f : N \to M$ eine Abbildung.

(1) f heißt **injektiv**, falls für alle $x, x' \in N$ gilt: $f(x) = f(x') \implies x = x'$.

(2) f heißt **surjektiv**, falls $f(N) = M$.

(3) f heißt **bijektiv**, falls f injektiv und surjektiv ist.

Bemerkung 0.4.8. Sei $f : N \to M$ eine Abbildung. Dann gilt:

(1) f ist injektiv $\iff$ jede Faser $f^{-1}(\{y\})$ hat *höchstens* ein Element.

(2) f ist surjektiv $\iff$ jede Faser $f^{-1}(\{y\})$ hat *mindestens* ein Element.

(3) f ist bijektiv $\iff$ jede Faser $f^{-1}(\{y\})$ hat *genau* ein Element.

Beispiel 0.4.9.

(1) $f : \mathbb{Z} \to \mathbb{Z},\ z \mapsto 2z$ ist injektiv, aber nicht surjektiv.

(2) $f : \mathbb{R} \to \mathbb{R},\ x \mapsto 2x$ ist bijektiv.

(3) $f : \mathbb{R} \to \mathbb{R},\ x \mapsto x^2$ ist weder injektiv ($f(1) = f(-1) = 1$) noch surjektiv ($f(\mathbb{R}) = \mathbb{R}_{\geq 0}$).

(4) Sei $J : P \to \mathbb{Z}$ die Geburtsjahr-Abbildung aus Beispiel 0.4.a. J ist genau dann injektiv, wenn keine zwei Personen in P im selben Jahr geboren wurden.

(5) Die Abbildung $\emptyset \to M$ ist injektiv. Sie ist genau dann surjektiv, wenn $M = \emptyset$.

(6) Hashfunktionen (auch "Prüfsummen" genannt), z.B.

$$\text{md5sum} : \{\text{Texte}\} \to \underline{2^{128}},$$

die einen 128 Bit-Hashwert produziert, sind offensichtlich nicht injektiv, aber idealerweise "kollisionsresistent" und surjektiv.

(7) Dagegen ist eine Verschlüsselungsfunktion

$$\text{crypt} : \underline{2^k} \to \underline{2^k}$$

injektiv, und somit surjektiv und bijektiv, damit eine eindeutige Entschlüsselung möglich ist.

Bemerkung 0.4.10. Für eine Abbildung $f : N \to M$ zwischen *endlichen* Mengen N, M gilt:

$$|f(N)| \leq |N|, |M|$$

Weiter gilt:

(1) f ist genau dann injektiv, wenn $|f(N)| = |N|$.

(2) f ist genau dann surjektiv, wenn $|f(N)| = |M|$.

Der Spezialfall $|N| = |M|$ (z.B. bei $M = N$) ist besonders wichtig:

$$|M| = |N| \implies (\text{injektiv} \iff \text{surjektiv} \iff \text{bijektiv}).$$

0.4.d Einschränkungen

Definition 0.4.11. Sei $f : N \to M$ eine Abbildung und $N' \subset N$. Dann heißt die Abbildung

$$f_{|N'} : N' \to M,\ x \mapsto f(x)$$

die **Einschränkung** von f auf N'

Bemerkung 0.4.12. Jede Abbildung kann durch Einschränkung auf eine geeignete Teilmenge des Definitionsbereiches injektiv gemacht werden. Z.B. sind für $f : \mathbb{R} \to \mathbb{R}, x \mapsto x^2$ die Einschränkungen $f_{|\mathbb{R}_{\geq 0}}$ und $f_{|\mathbb{R}_{\leq 0}}$ beide injektiv.

0.4.e Komposition von Abbildungen

Definition 0.4.13. Seien N, M, M', L Mengen. Weiter seien $f : N \to M$ und $g : M' \to L$ zwei Abbildungen mit $f(N) \subset M'$. Dann heißt die Abbildung

$$g \circ f : N \to L,\ x \mapsto (g \circ f)(x) \coloneqq g(f(x))$$

die **Komposition von f und g**. Häufig ist $M' = M$.

Beispiel 0.4.14. Für die Abbildungen

$$f : \mathbb{R} \to \mathbb{R},\ x \mapsto (x-3)^2,$$
$$g : \mathbb{R}_{\geq 0} \to \mathbb{R},\ x \mapsto \sqrt{x}$$

gibt es zwei mögliche Kompositionen:

$$g \circ f : \mathbb{R} \to \mathbb{R},\ x \mapsto \sqrt{(x-3)^2} = |x-3|,$$
$$f \circ g : \mathbb{R}_{\geq} 0 \to \mathbb{R},\ x \mapsto (\sqrt{x}-3)^2.$$

Bemerkung 0.4.15. Seien f, g, h Abbildungen.

(1) Ist die Komposition $h \circ (g \circ f)$ definiert, so braucht die Komposition $(h \circ g) \circ f$ nicht definiert zu sein. Aber falls doch, so stimmen sie überein. D.h. die Komposition von Abbildungen ist assoziativ, sprich die Klammern dürfen weggelassen werden: $h \circ g \circ f$.

(2) Ist $g \circ f$ definiert, so braucht $f \circ g$ im Allgemeinen nicht definiert zu sein.

0.4.f Umkehrabbildung

Definition 0.4.16. Seien $f : N \to M$ und $g : M \to N$ zwei Abbildungen (Kurzschreibweise: $N \underset{g}{\overset{f}{\rightleftarrows}} M$). Dann heißt g eine

- **linksseitige Umkehrabbildung von** f, wenn $g \circ f = \mathrm{id}_N$,
- **rechtsseitige Umkehrabbildung von** f, wenn $f \circ g = \mathrm{id}_M$,
- **Umkehrabbildung von** f, wenn sie sowohl links- als auch rechtsseitige Umkehrabbildung von f ist.

Entsprechend sagen wir, f ist linksinvertierbar, rechtsinvertierbar, bzw. invertierbar.

Satz 0.4.17. *Sei $f : N \to M$ eine Abbildung und $N \neq \emptyset$.*

(1) f ist genau dann linksinvertierbar, wenn f injektiv ist.

(2) f ist genau dann rechtsinvertierbar, wenn f surjektiv ist.

(3) f ist genau dann invertierbar, wenn f bijektiv ist. In diesem Fall existiert genau eine Umkehrabbildung, die mit $f^{-1} : M \to N$ bezeichnet wird.

Bemerkung 0.4.18. Aussage (2) ist eine äquivalente Formulierung des sogenannten Auswahlaxioms der Mengenlehre. Aus dem ZF (=Zermelo-Fraenkel Axiomensystem) kann man weder die Gültigkeit noch die Ungültigkeit von (2) in dieser Allgemeinheit beweisen[3]. Man sagt daher: (2) ist logisch unabhängig von ZF. Nimmt man die Gültigkeit von (2) an, so spricht man vom ZFC, d.h. ZF mit "axiom of choice".

Beweis des Satzes.

(1) ($\Longrightarrow$) Sei $g : M \to N$ eine linksseitige Umkehrabbildung von f. Dann gilt für alle $x, x' \in N$:

$$f(x) = f(x') \implies g(f(x)) = g(f(x')) \implies \underbrace{(g \circ f)}_{\mathrm{id}_N}(x) = \underbrace{(g \circ f)}_{\mathrm{id}_N}(x') \implies x = x'.$$

($\Longleftarrow$) Sei $f : N \to M$ injektiv, d.h. $\forall y \in M : |f^{-1}(\{y\})| \leq 1$. *Wähle* $x_0 \in N$ beliebig (möglich da $N \neq \emptyset$) und definiere $g : M \to N$ durch

$$g(y) := \begin{cases} x, & \text{falls } f^{-1}(\{y\}) = \{x\}, \qquad (\implies f(x) = y) \\ x_0, & \text{falls } f^{-1}(\{y\}) = \emptyset. \end{cases}$$

Damit gilt $(g \circ f)(x) = g(f(x)) = x$ für alle $x \in N$, d.h. $g \circ f = \mathrm{id}_N$, wie gewünscht.

[3] Man kann den Fall $|M| < \infty$ aus ZF folgern, aber nicht wenn M eine beliebige Kardinalität hat.

(2) ($\Longrightarrow$) Sei $g : M \to N$ eine rechtsseitige Umkehrabbildung von f. Dann ist

$$f(g(y)) = \underbrace{(f \circ g)}_{\mathrm{id}_M}(y) = y \quad \text{für alle } y \in M,$$

d.h. $g(y) \in N$ ist ein Urbild von y unter f. Mit anderen Worten, eine rechtsseitige Umkehrabbildung zeichnet faserweise Urbilder aus.
($\Longleftarrow$) Sei $f : N \to M$ surjektiv, d.h. $\forall y \in M : |f^{-1}(\{y\})| \geq 1$. Definiere $g : M \to N$ durch *Wahl*[4] eines Urbildes pro Faser

$$g(y) \in f^{-1}(\{y\}) \neq \emptyset.$$

Damit gilt $f(g(y)) = y$ für alle $y \in M$, d.h. $f \circ g = \mathrm{id}_M$, wie gewünscht.

(3) Die erste Aussage folgt aus (1) und (2). Nun zur Eindeutigkeit: Sei $g : M \to N$ eine linksseitige und $g' : M \to N$ eine rechtsseitige Umkehrabbildung. Dann ist

$$g = g \circ \mathrm{id}_M = g \circ f \circ g' = \mathrm{id}_N \circ g' = g'. \qquad \square$$

Beispiel 0.4.19.

(1) $f : \mathbb{R} \to \mathbb{R}, x \mapsto 2x$ ist bijektiv mit der Umkehrabbildung

$$f^{-1} : \mathbb{R} \to \mathbb{R},\ x \mapsto \frac{1}{2}x.$$

(2) $f : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 1},\ x \mapsto x^2 + 1$ ist bijektiv mit der Umkehrabbildung

$$f^{-1} : \mathbb{R}_{\geq 1} \to \mathbb{R}_{\geq 0}, x \mapsto \sqrt{x-1}.$$

(3) Sei $N \subset M$ eine Mengeninklusion. Dann ist $f : N \to M,\ x \mapsto x$ injektiv. Die Abbildung f ist genau dann surjektiv, wenn $N = M$ ist.

(4) $f : \mathbb{R} \to \mathbb{R}_{\geq 0},\ x \mapsto |x|$ ist surjektiv, aber nicht injektiv, und daher auch nicht linksinvertierbar. Die Abbildungen $g : \mathbb{R}_{\geq 0} \to \mathbb{R},\ x \mapsto x$ und $g' : \mathbb{R}_{\geq 0} \to \mathbb{R},\ x \mapsto -x$ sind zwei verschiedene rechtsseitige Umkehrabbildungen.

Satz 0.4.20. *Seien $f : N \to M$ und $g : M \to L$ zwei Abbildungen.*

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

(2) Sind f, g surjektiv so auch $g \circ f$.

(3) Sind f, g bijektiv so auch $g \circ f$. In diesem Fall gilt:

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

Beweis. Übung. $\square$

[4] Das ist das Auswahlaxiom!

0.4.g Mächtigkeit von Mengen

Definition 0.4.21. Zwei Mengen M und N heißen **gleichmächtig** oder **isomorph**, wenn eine bijektive Abbildung $f : M \to N$ existiert.

Übung 0.4.2. Beweisen Sie: $\mathbb{N}$, $\mathbb{Z}$ und $\mathbb{Q}$ sind gleichmächtig.

Satz 0.4.22 (Cantor). *Für jede Menge M sind M und $\mathrm{Pot}(M)$* nicht *gleichmächtig.*

Beweis. Sei $f : M \to \mathrm{Pot}(M)$ eine beliebige Abbildung. Definiere

$$A_f := \{x \in M \mid x \notin f(x)\} \in \mathrm{Pot}(M).$$

Angenommen es gibt ein $m \in M$ mit $f(m) = A_f$. Dann gilt: $m \in A_f \implies m \notin f(m) = A_f$ und $m \notin A_f \implies m \in f(m) = A_f$. Widerspruch. Also ist die Annahme falsch und f kann nicht surjektiv sein. □

Übung 0.4.3. Beweisen Sie:

(1) $\mathbb{N}$ und $\mathbb{R}$ sind nicht gleichmächtig.

(2) Die Zusammenfassung aller Mengen ist keine Menge. Man nennt sie daher eine **Klasse**.

0.4.h Selbstabbildung = Abbildung einer Menge in sich

Definition 0.4.23. Seien M eine Menge, $f : M \to M$ eine Abbildung und $n \in \mathbb{N}$. In diesem Fall spricht man von einer **Selbstabbildung**. Wir setzen

$$f^n := \underbrace{f \circ \cdots \circ f}_{n} \quad \text{und} \quad f^0 := \mathrm{id}_M .$$

Falls zusätzlich f bijektiv ist, so setzen wir $f^{-n} := (f^{-1})^n$.

Bemerkung 0.4.24. Für $a, b \in \mathbb{N}_0$ gelten die Potenzregeln:

$$f^a \circ f^b = f^{a+b} \quad \text{und} \quad (f^a)^b = f^{ab}.$$

Ist f bijektiv, so gilt dies sogar für alle $a, b \in \mathbb{Z}$.

0.5 Relationen

0.5.a Definition und Beispiele

Definition 0.5.1. Seien M und N Mengen.

(1) Eine Teilmenge $R \subset M \times N$ heißt **Relation zwischen M und N**. Für $(x, y) \in R$ schreiben wir auch xRy und sagen "x steht in Relation zu y bzgl. R".

(2) Ist $R \subset M \times M$, so sprechen wir von einer **Relation auf M**. Sie heißt

- (R) **reflexiv**, falls xRx für alle $x \in M$;
- (R') **antireflexiv**, falls nicht xRx für alle $x \in M$;
- (S) **symmetrisch**, falls $xRy \implies yRx$ für alle $x, y \in M$;
- (A) **antisymmetrisch**, falls $(xRy \wedge yRx) \implies x = y$ für alle $x, y \in M$;

(T) **transitiv**, falls $(xRy \wedge yRz) \implies xRz$ für alle $x, y, z \in M$.

(3) Eine Relation, die (R), (A) und (T) erfüllt heißt **(partielle) Ordnung**.

(4) Eine Ordnung heißt **Totalordnung**, falls $xRy \vee yRx$ für alle $x, y \in M$.

(5) Eine Relation, die (R), (S) und (T) erfüllt heißt **Äquivalenzrelation** (ÄR).

Beispiel 0.5.2.

(1) Jede Abbildung $f : N \to M$ kann als eine Relation aufgefasst werden. Nämlich die Abbildung f legt ihren **Graph**

$$\Gamma_f := \{(x, f(x)) : x \in N\} \subset N \times M$$

als Relation fest. Umgekehrt ist f durch ihren Graphen Γ_f eindeutig bestimmt.

(2) $M = \mathrm{Pot}(N)$ und $R =$ "$\subset$", d.h. $xRy :\iff x \subset y$. Da "$\subset$" (R), (A) und (T) ist, ist sie eine Ordnung auf M. Sie ist keine Totalordnung sobald $|N| \geq 2$ ist.

(3) $M = \mathbb{R}$ und $R =$ "$\leq$". Da "$\leq$" (R), (A) und (T) ist, ist sie eine Ordnung auf M. Sie ist sogar eine Totalordnung.

(4) $M = \mathbb{N}$ mit der Teilbarkeitsrelation $R =$ "$|$", d.h. $xRy :\iff x \mid y$. Sie ist eine Ordnung auf $\mathbb{N}$, aber keine Totalordnung.

(5) Die Teilbarkeitsrelation ist keine Ordnung auf $\mathbb{Z}$, da $-1 \mid 1$ und $1 \mid -1$ aber $1 \neq -1$.

(6) M eine Menge und $R =$ "$=$" die Gleichheit, d.h. $xRy :\iff x = y$. Sie ist (R), (S) und (T) und daher eine Äquivalenzrelation.

Bemerkung 0.5.3.

- Man kann auf einer endlichen Menge $M = \{m_1, \ldots, m_n\}$ eine Relation durch eine $n \times n$-Matrix definieren, die eine 1 an der Position (i, j) hat, falls $m_i R m_j$, und sonst 0. Die Eigenschaften (R), (S), (A) kann man dann sofort an der Matrix ablesen.

- Ist M eine Menge und $M' \subset M$ so ist $R' := R \cap (M \times M)$ eine Relation. Erfüllt R eine der Eigenschaften aus (2), so auch R'.

Übung 0.5.1.

(1) Welche Bedingungen muss eine Relation $R \subset N \times M$ erfüllen, damit sie gemäß Beispiel (1) als eine Abbildung von N nach M aufgefasst werden kann?

(2) Unter welchen Bedingungen ist diese Abbildung injektiv, surjektiv bzw. bijektiv?

(3) Welche Relation gehört im bijektiven Fall zur Umkehrabbildung?

0.5.b Partielle Ordnungen

Sei $\preceq$ eine Ordnung auf M.

Konvention. Wie üblich schreiben wir $m' \succeq m$ für $m \preceq m'$.

Definition 0.5.4. Ein Element $m \in M$ heißt

(1) **minimal** in M, falls $m' \preceq m \implies m' = m$, d.h. falls kein anderes Element "kleiner" als m ist.

(2) **Minimum** von M, falls für alle $m' \in M$ gilt $m \preceq m'$, d.h. falls alle anderen Element "größer" sind.

(3) **maximal** in M, falls $m' \succeq m \implies m' = m$.

(4) **Maximum** von M, falls für alle $m' \in M$ gilt $m \succeq m'$.

Satz 0.5.5. *Sei $\preceq$ eine partielle Ordnung auf M.*

(1) Jedes Minimum von M ist minimal in M.

(2) Existiert ein Minimum in M, so ist es das einzige minimale Element in M. Insbesondere ist das Minimum eindeutig.

(3) Bei einer Totalordnung sind die Begriffe "minimal" und "Minimum" gleichbedeutend.

Die dualen Aussagen für maximal und Maximum gelten entsprechend.

Beweis.

(1) Sei m ein Minimum und $m' \preceq m$, so folgt aus $m \preceq m'$ und (A), dass $m' = m$ ist.

(2) Seien m ein Minimum und m' minimal. Dann gilt $m \preceq m'$ (wegen m Minimum) und daher $m = m'$ (wegen m' minimal).

(3) Sei nun $\preceq$ eine Totalordnung auf M. Sei m minimal in M. So gilt für jedes $m' \in M$ (wegen Totalordnung) $m' \preceq m$ oder $m \preceq m'$ und daher (wegen m minimal) $m \preceq m'$. D.h. m ist das Minimum von M. □

Beispiel 0.5.6. Sei $|$ die Teilbarkeitsrelation auf $\mathbb{N}$ bzw. auf jeder Teilmenge $N \subset \mathbb{N}$.

(1) Die Teilmenge $\{2, 3, 4, 6, 12\}$ hat zwei minimale Elemente, nämlich 2 und 3 aber kein Minimum. Sie besitzt dagegen ein Maximum, nämlich 12.

(2) Das Minimum von $\mathbb{N}$ ist 1. Es gibt kein maximales Element und entsprechend auch kein Maximum.

Übung 0.5.2. Jede nicht-leere Teilmenge von $\mathbb{N}$ besitzt bzgl. der Ordnung $\leq$ ein Minimum.

Beweis. Man führe einen Widerspruchsbeweis mit vollständiger Induktion. □

0.5.c Äquivalenzrelationen

Definition 0.5.7. Sei $\sim$ eine Äquivalenzrelation auf M. Für $x \in M$ heißt

$$[x] \coloneqq [x]_\sim \coloneqq \{y \in M \mid x \sim y\}$$

die **Äquivalenzklasse von $\sim$ zu x** (oder **von x bezüglich $\sim$**).

Die Menge aller Äquivalenzklassen von $\sim$ wird mit

$$M/\sim \coloneqq \{[x]_\sim : x \in M\}$$

bezeichnet.

Übung 0.5.3. Die Isomorphie von Mengen (siehe Definition 0.4.21) ist eine Äquivalenzrelation auf der Klasse aller Mengen.

Bemerkung 0.5.8. Sei $\sim$ ein Äquivalenzrelation auf M. Dann gilt für $x, y \in M$:

(R) $x \in [x]_\sim$,

(S) $y \in [x]_\sim \implies x \in [y]_\sim$,

(S,T) $x \sim y \implies [x]_\sim = [y]_\sim$.

Wegen der letzten Eigenschaft bezeichnet man jedes Element einer Äquivalenzklasse als einen **Repräsentant** derselben.

Satz 0.5.9. *Sei M eine Menge.*

(1) Ist $\sim$ eine Äquivalenzrelation auf M, so ist $M/\sim$ eine Partition von M.

(2) Ist $\mathcal{P}$ eine Partition von M, so existiert eine Äquivalenzrelation $\sim$ auf M mit $M/\sim = \mathcal{P}$.

D.h. Äquivalenzrelationen auf M entsprechen Partitionen von M.

Beweis.

(1) Wegen $x \in [x]_\sim$ sind alle Äquivalenzklassen nicht leer und ihre Vereinigung ganz M. Also müssen wir nur noch zeigen, dass die Äquivalenzklassen paarweise disjunkt sind. Sind $[x]_\sim, [y]_\sim$ zwei solche, so wollen wir zeigen

$$[x]_\sim \neq [y]_\sim \implies [x]_\sim \cap [y]_\sim = \emptyset,$$

bzw. die Kontraposition

$$[x]_\sim \cap [y]_\sim \neq \emptyset \implies [x]_\sim = [y]_\sim.$$

Also sei $z \in [x]_\sim \cap [y]_\cap \neq \emptyset$, d.h. $z \in [x]_\sim$ und $z \in [y]_\sim$. Aus (S,T) der letzten Bemerkung folgt, dass dann $[x]_\sim = [z]_\sim = [y]_\sim$.

(2) Wir definieren $\sim$ durch die Vorschrift

$$x \sim y :\iff x \text{ und } y \text{ liegen in demselben Teil der Partition } \mathcal{P}.$$

Dies ist offensichtlich eine Äquivalenzrelation. Die Äquivalenzklassen sind per Definition von $\sim$ die Teile von $\mathcal{P}$.

□

Beispiel 0.5.10.

(1) Für die Gleichheitsrelation auf einer Menge M ist $[x]_= = \{x\}$ und $M/_= = \{\{x\} : x \in M\}$.

(2) Sei $f : N \to M$ eine Abbildung und R_f die Bildgleichheitsrelation auf N, d.h.

$$xR_fx' :\iff f(x) = f(x').$$

Sie ist offensichtlich eine Äquivalenzrelation. Ihre Äquivalenzklassen sind die verschieden nicht-leeren Fasern von f, d.h. für alle $x \in N$ ist

$$[x]_{R_f} = \{x' \in N \mid f(x) = f(x')\} = f^{-1}(\{f(x)\}).$$

Die Partition M/R_f ist demnach die Menge der nicht-leeren Fasern von f.

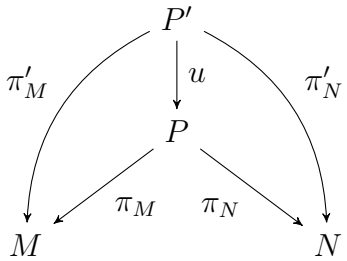
(3) Die Relation $\equiv_2$ definiert durch $x \equiv_2 y :\iff 2 \mid x - y$ ist eine Äquivalenzrelation. Ihre Äquivalenzklassen sind

$$[0]_{\equiv_2} = \{a \in \mathbb{Z} \mid a \text{ gerade}\},$$
$$[1]_{\equiv_2} = \{a \in \mathbb{Z} \mid a \text{ ungerade}\}.$$

0.6 Die Kategorie der Mengen

Die „Kategorie“ **Sets** aller Mengen (als Objektklasse) und Abbildungen (als Morphismenklasse) ist eine der wichtigsten Referenzkategorien in der Mathematik. Den Begriff der Kategorie werden wir hier nicht weiter präzisieren, da dies für die Bearbeitung folgender Übungsaufgaben nicht notwendig ist.

Übung 0.6.1. Zeigen Sie, dass die Kategorie **Sets** **binäre Produkte** besitzt. D.h. gegeben zwei Mengen M, N, dann existiert ein **Produkt** von M und N, sprich eine Menge P zusammen mit zwei Abbildungen $\pi_M : P \to M$ und $\pi_N : P \to N$, dass im folgenden Sinne universell ist: Für jede weitere Menge P' zusammen mit zwei Abbildungen $\pi'_M : P' \to M$ und $\pi'_N : P' \to N$ existiert eine *eindeutig bestimmte* Abbildung $u : P' \to P$, so dass das folgende Diagramm kommutiert[5]



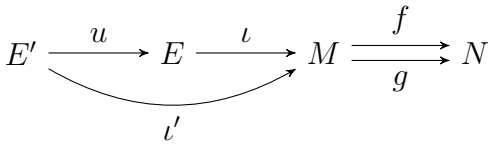
Hinweis: Betrachten Sie das kartesische Produkt $P = M \times N$.

Übung 0.6.2. Zeigen Sie, dass die Kategorie **Sets** ein **terminales Objekt** besitzt, d.h. eine Menge T, dass im folgenden Sinne universell ist: Für jede weitere Menge T' existiert eine *eindeutig bestimmte* Abbildung $u : T' \to T$.
Hinweis: Betrachten Sie für T eine einelementige Menge, egal welche.

Bemerkung 0.6.1. Ein terminales Objekt ist in einem gewissen Sinne ein **leeres Produkt**. Übungen 0.6.1 und 0.6.2 zeigen somit, dass **Sets** **endliche Produkte** besitzt.

[5]Sprich, alle Kompositionswege im Diagramm führen zur gleichen Abbildung.

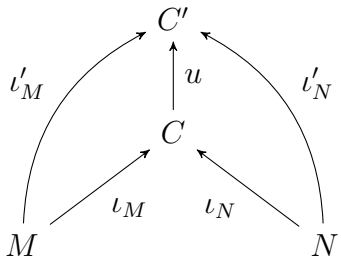
Übung 0.6.3. Zeigen Sie, dass die Kategorie **Sets** **binäre Egalisatoren** besitzt. D.h. gegeben zwei Abbildungen $f, g : M \to N$, dann existiert ein Egalisator von f und g, sprich eine Menge E zusammen mit einer Abbildung $\iota : E \to M$, die $f \circ \iota = g \circ \iota$ erfüllt, dass im folgenden Sinne universell ist: Für jede weitere Menge E' zusammen mit einer Abbildung $\iota' : E' \to M$, die $f \circ \iota' = g \circ \iota'$ erfüllt, existiert eine *eindeutig bestimmte* Abbildung $u : E' \to E$, so dass das folgende Diagramm kommutiert



Bemerkung 0.6.2. Eine Kategorie mit endlichen Produkten und Egalisatoren besitzt automatisch alle **endlichen Limiten**. Solche Kategorien nennt man **endlich vollständig**.

Dreht man alle Pfeile in Übungen 0.6.1, 0.6.2, 0.6.3 zeigen alle Pfeile um, so entsteht der duale Begriff des Koproduktes, des initialen Objektes, und des Koegalisators:

Übung 0.6.4. Zeigen Sie, dass die Kategorie **Sets** **binäre Koprodukte** besitzt. D.h. gegeben zwei Mengen M, N, dann existiert das **Koprodukt** von M und N, sprich eine Menge C zusammen mit zwei Abbildungen $\iota_M : M \to C$ und $\iota_N : N \to C$, dass im folgenden Sinne universell ist: Für jede weitere Menge C' zusammen mit zwei Abbildungen $\iota'_M : M \to C'$ und $\iota'_N : N \to C'$ existiert eine *eindeutig bestimmte* Abbildung $u : C \to C'$, so dass das folgende Diagramm kommutiert[6]

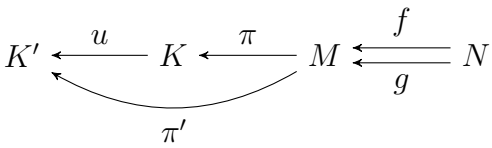


Hinweis: Betrachten Sie die disjunkte Vereinigung $C = (M \times \{0\}) \cup (N \times \{1\})$.

Übung 0.6.5. Zeigen Sie, dass die Kategorie **Sets** ein **initiales Objekt** besitzt, d.h. eine Menge I, dass im folgenden Sinne universell ist: Für jede weitere Menge I' existiert eine *eindeutig bestimmte* Abbildung $u : I \to I'$.
Hinweis: Betrachten Sie für I die leere Menge.

Bemerkung 0.6.3. Ein initiales Objekt ist in einem gewissen Sinne ein **leeres Koprodukt**. Übungen 0.6.4 und 0.6.5 zeigen somit, dass **Sets** **endliche Koprodukte** besitzt.

Übung 0.6.6. Zeigen Sie, dass die Kategorie **Sets** **binäre Koegalisatoren** besitzt. D.h. gegeben zwei Abbildungen $f, g : N \to M$, dann existiert ein Koegalisator von f und g, sprich eine Menge K zusammen mit einer Abbildung $\pi : M \to K$, die $\pi \circ f = \pi \circ g$ erfüllt, dass im folgenden Sinne universell ist: Für jede weitere Menge K' zusammen mit einer Abbildung $\pi' : M \to K'$, die $\pi' \circ f = \pi' \circ g$ erfüllt, existiert eine *eindeutig bestimmte* Abbildung $u : K \to K'$, so dass das folgende Diagramm kommutiert



[6]Sprich, alle Kompositionswege im Diagramm führen zur gleichen Abbildung.

Übung 0.6.7. Zeigen Sie, dass alle universellen Objekte in den obigen Übungen bis auf Isomorphie in **Sets** eindeutig bestimmt sind.

Bemerkung 0.6.4. Eine Kategorie mit endlichen Koprodukten und Koegalisatoren besitzt automatisch alle **endlichen Kolimiten**. Solche Kategorien nennt man **endlich kovollständig**.

Die Kategorie aller Mengen erfüllt eine Liste von Eigenschaften, die wir zum Teil in den obigen Übungsaufgaben kennenlernt haben. Eine Kategorie, die diese Liste von Eigenschaften erfüllt, nennt man einen **Topos**. Jeder Topos kommt mit seiner internen Logik. Die allermeisten Topoi haben die **konstruktive Logik** als interne Logik, eine Logik wo insbesondere tertium non datur und Widerspruchsbeweise (siehe Beispiel 0.1.11) nicht als Deduktionsregeln benutzt werden dürfen. Ebenfalls gilt in diesen Topoi nicht das Auswahlaxiom, da es tertium non datur implizieren würde.

Man kann Mathematik in jedem Topos als Ersatz für den Topos **Sets** betreiben. Nur Sätze, die mit konstruktiver Logik bewiesen werden, sind in jedem Topos gültig. Daher ist die konstruktive Mathematik die „lingua franca“ der Mathematik. Hier ein Beispiel aus der Analysis: Man kann in jedem Topos mit einem „natürliche Zahlen“-Objekt rationale und reelle Zahlen definieren und Analysis betreiben. Es gibt Topoi, in denen der Zwischenwertsatz in seiner klassischen Formulierung schlicht falsch ist. Der klassische Beweis benutzt in der Tat tertium non datur und ist daher nicht in jedem Topos mit reellen Zahlen gültig. Glücklicherweise gibt es eine Variante des Zwischenwertsatzes, die in jedem Topos mit reellen Zahlen gültig ist. Diese Variante ist sogar für die numerische Analysis ausreichend!

Kapitel 1

Lineare Gleichungssysteme

Was ist ein Gleichungssystem? Was bedeutet es, ein Gleichungssystem zu lösen?

1.1 Fasern einer Abbildung

Seien X, B Mengen. Eine Teilmenge

$$\Gamma_f \subseteq X \times B = \{(x, b) \mid x \in X, b \in B\}$$

definiert den Graph einer Abbildung $f : X \to B$, falls für alle $x \in X$ genau ein $b \in B$ existiert mit $(x, b) \in \Gamma_f$ (siehe Übung 0.5.1 und Beispiel 0.5.2.(1)), sprich

$$f(x) = b \iff (x, b) \in \Gamma_f.$$

Für $b \in B$ heißt die Menge

$$f^{-1}(\{b\}) \coloneqq \{x \in X \mid f(x) = b\}$$

das Urbild von b oder auch die **Faser** von f **über** b.

Aus Definition 0.4.7 und Bemerkung 0.4.8 wissen wir:

- Haben alle Fasern höchstens ein Element, so heißt α **injektiv**.
- Ist keine Faser leer, so heißt α **surjektiv**.
- Sind alle Fasern einelementig, so heißt α **bijektiv**.

Warum wiederholen wir diese Begriffe? Die zwei folgenden Bemerkungen klären den Zusammenhang zwischen diesen Begriffen und dem Lösen von Gleichungssystemen:

Bemerkung 1.1.1. Seien X, B Mengen und $f : X \to B$ eine Abbildung. Für jedes $b \in B$ ist das zu f und b gehörige **Gleichungssystem** gegeben durch

$$f(x) = b,$$

und seine **Lösungsmenge** ist gerade die Faser

$$f^{-1}(\{b\}) = \{x \in X \mid f(x) = b\}.$$

Das Lösen eines Gleichungssystems ist somit nichts anderes als die Bestimmung einer Faser einer Abbildung f.

Die Begriffe surjektiv, injektiv und bijektiv können wir somit folgendermaßen interpretieren:

Bemerkung 1.1.2. Das Gleichungssystem $f(x) = b$ für die Abbildung $f : X \to B$

- ist immer[1] lösbar genau dann, wenn f surjektiv ist.
- hat immer *höchstens* eine Lösung genau dann, wenn f injektiv ist.
- ist immer *eindeutig* lösbar genau dann, wenn f bijektiv ist.

Beispiel 1.1.3. Seien $X = B = \mathbb{R}^2 = \{(x, y) \mid x, y \in \mathbb{R}\}$ und

$$f : X \to B, (x, y) \mapsto (x^2 + y^2, x + 3y).$$

Gesucht ist $f^{-1}(\{(1, 1)\})$. Wir suchen also die Paare $(x, y) \in \mathbb{R}^2$ mit $x^2+y^2 = 1$ und $x+3y = 1$. Man rechnet leicht nach, dass dieses Gleichungssystem genau 2 Lösungen hat, die man als Schnittpunkte von einem Kreis und einer Gerade finden kann.

Beispiel 1.1.4. Als motivierendes Beispiel betrachten wir das folgende **lineare Gleichungssystem**:

$$\begin{array}{rrrrcl} 5x_1 & +7x_3 & +x_4 & -x_5 & = & 2 \\ & x_3 & +x_4 & +x_5 & = & 0 \\ x_1 & & & -x_5 & = & 1 \end{array} \qquad (\star)$$

Seine Lösungsmenge ist die Menge aller $(x_1, x_2, x_3, x_4, x_5) \in \mathbb{Q}^5$, die diese Gleichungen erfüllen. Dies ist die Faser der Abbildung f über $(2, 0, 1)$, wobei

$$f : \mathbb{Q}^5 \to \mathbb{Q}^2, (x_1, x_2, x_3, x_4, x_5) \mapsto (5x_1 + 7x_3 + x_4 - x_5, x_3 + x_4 + x_5, x_1 - x_5).$$

Die Lösungsmenge von ($\star$) kann man parametrisieren:

$$\mathcal{L}(\star) = \{(b + 1, a, (-1 - b)/2, (1 - b)/2, b) \mid a, b \in \mathbb{Q}\}.$$

1.2 Lineare Abbildungen und Matrizen

Definition 1.2.1. Seien $m, n \in \mathbb{N}$ natürliche Zahlen. Mit $\underline{n} \coloneqq \{1, \ldots, n\}$ bezeichnen wir die Menge aller natürlichen Zahlen $\leq n$. Eine rationale $m \times n$-**Matrix** ist eine Abbildung

$$A : \underline{m} \times \underline{n} \to \mathbb{Q}, (i, j) \mapsto A_{i,j}.$$

Notation:

$$A = \begin{pmatrix} A_{1,1} & A_{1,2} & \ldots & A_{1,n-1} & A_{1,n} \\ A_{2,1} & A_{2,2} & \ldots & A_{2,n-1} & A_{2,n} \\ \vdots & \vdots & \vdots & \vdots & \\ A_{m,1} & A_{m,2} & \ldots & A_{m,n-1} & A_{m,n} \end{pmatrix}.$$

Mit anderen Worten, die Positionen in der Matrix entsprechen dem Definitionsbereich, d.h. die Position „i-te Zeile, j-te Spalte“ entspricht $(i, j) \in \underline{m} \times \underline{n}$ und der Wert von A für (i, j) wird in die entsprechende Position eingetragen. Die Menge aller rationalen $m \times n$-Matrizen wird mit $\mathbb{Q}^{m \times n}$ bezeichnet.
Ist $m = 1$ bzw. $n = 1$ spricht man von **Zeilen** bzw. **Spalten** statt von Matrizen.[2]

[1] d.h. für jede rechte Seite $b \in B$

[2] Um leere Matrizen korrekt miteinzubeziehen muss man $\mathbb{N}$ durch $\mathbb{N}_0$ ersetzen und das Mengenpaar $(\underline{m}, \underline{n})$ als Teil des Datums des Definitionsbereiches betrachten.

Definition 1.2.2. Ein **lineares Gleichungssystem** (über $\mathbb{Q}$) mit m Gleichungen und n Unbekannten $x_1, \ldots, x_n$ ist gegeben durch

$$\begin{array}{rcl} A_{11}x_1 + A_{12}x_2 + \ldots + A_{1n}x_n & = & b_1 \\ A_{21}x_1 + A_{22}x_2 + \ldots + A_{2n}x_n & = & b_2 \\ \vdots & & \\ A_{m1}x_1 + A_{m2}x_2 + \ldots + A_{mn}x_n & = & b_m \end{array} \tag{$\star\star$}$$

wobei $A = (A_{i,j}) \in \mathbb{Q}^{m\times n}$ eine (fest vorgegebene) Matrix ist und $b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \in \mathbb{Q}^{m\times 1}$ eine (fest vorgegebene) Spalte ist.

- Die Matrix A heißt die **Matrix** vom Gleichungssystem ($\star\star$).
- Die Matrix $(A|b) \in \mathbb{Q}^{m\times(n+1)}$ definiert durch

  $$(A|b) : \underline{m} \times \underline{n+1} \to \mathbb{Q},\ (i,j) \mapsto \begin{cases} A_{ij} & j \leq n \\ b_i & j = n+1 \end{cases}$$

  heißt die **erweiterte Matrix** vom Gleichungssystem ($\star\star$).
- Unter einer **Lösung** von ($\star\star$) versteht man eine Spalte $\xi := \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} \in \mathbb{Q}^{n\times 1}$, derart, dass durch Einsetzen von ξ_i für x_i in ($\star\star$) für $i = 1, \ldots, n$ alle m Gleichungen von ($\star\star$) erfüllt sind.
- Die **Lösungsmenge** $\mathcal{L}(\star\star)$ ist die Menge aller solcher Lösungen.

Beispiel 1.2.3. In dem Eingangsbeispiel 1.1.4

$$\begin{array}{rrrrcl} 5x_1 & +7x_3 & +x_4 & -x_5 & = & 2 \\ & x_3 & +x_4 & +x_5 & = & 0 \\ x_1 & & & -x_5 & = & 1 \end{array} \tag{$\star$}$$

ist die Matrix von ($\star$)

$$\begin{pmatrix} 5 & 0 & 7 & 1 & -1 \\ 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & -1 \end{pmatrix} \in \mathbb{Q}^{3\times 5}$$

und die erweiterte Matrix

$$\left(\begin{array}{rrrrr|r} 5 & 0 & 7 & 1 & -1 & 2 \\ 0 & 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 & -1 & 1 \end{array}\right) \in \mathbb{Q}^{3\times 6}.$$

Ist das Gleichungssystem linear, so ist die dazugehörige Abbildung auch linear und umgekehrt. Auch diese Begriffe wollen wir sauber definieren:

Definition 1.2.4. Auf $\mathbb{Q}^{n\times 1} = \{ \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \mid a_1, \ldots, a_n \in \mathbb{Q}\}$ ist eine Verknüpfung

$$+ : \mathbb{Q}^{n\times 1} \times \mathbb{Q}^{n\times 1} \to \mathbb{Q}^{n\times 1},\ (a,b) \mapsto a + b$$

gegeben, die **Addition** heißt, definiert durch

$$\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} + \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} := \begin{pmatrix} a_1 + b_1 \\ \vdots \\ a_n + b_n \end{pmatrix}.$$

Weiter ist eine sogenannte Operation von $\mathbb{Q}$ auf $\mathbb{Q}^{n\times 1}$

$$\cdot : \mathbb{Q} \times \mathbb{Q}^{n\times 1} \to \mathbb{Q}^{n\times 1},\ (r, a) \mapsto r \cdot a$$

gegeben, genannt **Skalarmultiplikation**, definiert durch

$$r \cdot \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} := \begin{pmatrix} ra_1 \\ \vdots \\ ra_n \end{pmatrix}.$$

Schließlich kann man die beiden Verknüpfungen kombinieren: Sind $S_1, \ldots, S_k \in \mathbb{Q}^{n\times 1}$ und $r_1, \ldots, r_k \in \mathbb{Q}$ (oder kürzer $S \in (\mathbb{Q}^{n\times 1})^k, r \in \mathbb{Q}^k$), so heißt

$$r_1S_1 + \cdots + r_kS_k \qquad (\in \mathbb{Q}^{n\times 1})$$

die **Linearkombination** der S_i mit den Koeffizienten r_i.

Definition 1.2.5. Eine Abbildung $\alpha : \mathbb{Q}^{n\times 1} \to \mathbb{Q}^{m\times 1}$ heißt **linear** oder auch eine **lineare Abbildung**, falls für alle $r \in \mathbb{Q}$, $x, y \in \mathbb{Q}^{n\times 1}$ gilt:

$$\alpha(rx + y) = r\alpha(x) + \alpha(y).$$

Bemerkung 1.2.6. Eine Abbildung $\alpha : \mathbb{Q}^{n\times 1} \to \mathbb{Q}^{m\times 1}$ ist linear, genau dann wenn für alle $S \in (\mathbb{Q}^{n\times 1})^k, r \in \mathbb{Q}^k$,

$$\alpha(r_1S_1 + \cdots + r_kS_k) = r_1\alpha(S_1) + \cdots + r_k\alpha(S_k).$$

Da jede Spalte $\xi := \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} \in \mathbb{Q}^{n\times 1}$ trivialerweise die Linearkombination

$$\boxed{\xi = \xi_1e_1 + \cdots + \xi_ne_n}$$

der **Einheitsspalten**

$$e_i = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \leftarrow i\text{-te Zeile}$$

ist, ist das Bild von ξ unter der linearen Abbildung α die entsprechende Linearkombination der Bilder der Einheitsspalten:

$$\boxed{\alpha(\xi) = \xi_1\alpha(e_1) + \ldots + \xi_n\alpha(e_n).}$$

Also ist jede lineare Abbildung α durch die Bilder der Einheitsspalten eindeutig festgelegt.

Beispiel 1.2.7. Die Identitätsabbildung

$$\mathrm{id}_{\mathbb{Q}^{n\times 1}} : \mathbb{Q}^{n\times 1} \to \mathbb{Q}^{n\times 1},\ x \mapsto x$$

ist linear, trivialerweise.

Und nun zur Hauptquelle linearer Abbildungen:

Definition 1.2.8. Sei $A : \underline{m} \times \underline{n} \to \mathbb{Q},\ (i,j) \mapsto A_{ij}$ eine rationale $m \times n$-Matrix, kurz $A = (A_{ij}) \in \mathbb{Q}^{m\times n}$. Die von A **induzierte lineare Abbildung**

$$\widetilde{A} : \mathbb{Q}^{n\times 1} \to \mathbb{Q}^{m\times 1},\ \xi \mapsto A\xi$$

ist definiert durch

$$\underbrace{\widetilde{A}\left(\begin{pmatrix}\xi_1\\ \xi_2\\ \vdots\\ \xi_n\end{pmatrix}\right)}_{\widetilde{A}(\xi)} := \underbrace{A\begin{pmatrix}\xi_1\\ \xi_2\\ \vdots\\ \xi_n\end{pmatrix}}_{A\xi} := \begin{pmatrix} A_{11}\xi_1 + A_{12}\xi_2 + \ldots + A_{1n}\xi_n\\ A_{21}\xi_1 + A_{22}\xi_2 + \ldots + A_{2n}\xi_n\\ \vdots\\ A_{m1}\xi_1 + A_{m2}\xi_2 + \ldots + A_{mn}\xi_n\end{pmatrix}.$$

$A\xi$ heißt das **Produkt** der Matrix A mit der Spalte ξ.

Bemerkung 1.2.9. $\widetilde{A}$ ist in der Tat eine lineare Abbildung, d.h. für alle $r \in \mathbb{Q}$ und alle $x, y \in \mathbb{Q}^{n\times 1}$ gilt

$$\widetilde{A}(rx + y) = r\widetilde{A}(x) + \widetilde{A}(y).$$

Ende Vorl. 1 10.10

Beweis. Übung. □

Bemerkung 1.2.10. Unmittelbar aus Definition 1.2.8 folgt für die Einheitsspalten, dass

$$\widetilde{A}(e_j) = Ae_j = A_{-,j} \quad \text{für alle } j = 1, \ldots, n,$$

wobei

$$A_{-,j} := \begin{pmatrix} A_{1j}\\ A_{2j}\\ \vdots\\ A_{mj}\end{pmatrix}$$

die j-te Spalte von A bezeichnet. Sprich, *die Spalten der Matrix A sind die Bilder der Einheitsspalten unter $\widetilde{A}$.*

Bemerkung 1.2.11. Aus der Linearität von $\widetilde{A}$ (Bemerkung 1.2.9) und Bemerkung 1.2.6 (oder direkt aus der Definition des Produktes $A\xi$ in Definition 1.2.8) folgt

$$A\xi = \widetilde{A}(\xi) = \xi_1 A_{-,1} + \xi_2 A_{-,2} + \cdots + \xi_n A_{-,n}.$$

Sprich, $\widetilde{A}(\xi) = A\xi$ *ist Linearkombination der Spalten von A mit den Koeffizienten ξ_i, $i = 1, \ldots, n$.* Dies ist eine Verallgemeinerung von Bemerkung 1.2.10.

Folgerung 1.2.12. *Das Gleichungssystem ($\star\star$) genau dann lösbar, wenn man die Spalte b aus den Spalten von A linearkombinieren kann.*

Beispiel 1.2.13. Die Abbildung f aus Beispiel 1.1.4 erfüllt[3] $f = \widetilde{A}$ mit

$$A := \begin{pmatrix} 5 & 0 & 7 & 1 & -1\\ 0 & 0 & 1 & 1 & 1\\ 1 & 0 & 0 & 0 & -1\end{pmatrix} \in \mathbb{Q}^{3\times 5}.$$

[3] Genau genommen müsste man dafür die Menge der rationalen n-Tupel $\mathbb{Q}^n$ mit der Menge aller rationalen Spalten mit n Zeilen $\mathbb{Q}^{n\times 1}$ identifizieren.

Beispiel 1.2.14. Die Matrix

$$I_n \coloneqq \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix} : \underline{n} \times \underline{n} \to \mathbb{Q},\ (i,j) \mapsto \delta_{ij} \coloneqq \begin{cases} 1 & \text{falls } i = j \\ 0 & \text{falls } i \neq j \end{cases}$$

heißt die (rationale) **Einheitsmatrix** vom Grad n. Die Spalten von I_n sind nichts anderes als die Einheitsspalten $e_1, \ldots, e_n$. Daher gilt nach Bemerkung 1.2.6 und Bemerkung 1.2.11

$$\xi = \xi_1 e_1 + \cdots + \xi_n e_n = \xi_1 (I_n)_{-,1} + \cdots + \xi_n (I_n)_{-,n} = \widetilde{I_n}(\xi)$$

für alle $\xi \in \mathbb{Q}^{n\times 1}$. Also induziert I_n die Identität von $\mathbb{Q}^{n\times 1}$ als lineare Abbildung:

$$\widetilde{I_n} = \mathrm{id}_{\mathbb{Q}^{n\times 1}}\,.$$

Das n-Tupel $(e_1, \ldots e_n)$ heißt die **Standardbasis** von $\mathbb{Q}^{n\times 1}$, weil jede Spalte $\xi \in \mathbb{Q}^{n\times 1}$ in eindeutiger Weise aus den Einheitsspalten linearkombiniert werden kann.

Satz 1.2.15. *Zu jeder linearen Abbildung $\alpha : \mathbb{Q}^{n\times 1} \to \mathbb{Q}^{m\times 1}$ gibt es genau eine Matrix $A \in \mathbb{Q}^{m\times n}$ mit $\alpha = \widetilde{A}$.*

Beweis.
Eindeutigkeit. Dies folgt aus Bemerkung 1.2.10: $\widetilde{A}(e_j) = A_{-,j}$, die j-te Spalte von A. Sprich, man kann die Spalten der Matrix A aus den Bildern der Einheitsspalten unter $\widetilde{A}$ rekonstruieren.
Existenz. Sei α linear. Aus Bemerkung 1.2.6 wissen wir, dass α eindeutig bestimmt ist durch $\alpha(e_1) =: \begin{pmatrix} A_{11} \\ \vdots \\ A_{m1} \end{pmatrix}, \ldots, \alpha(e_n) =: \begin{pmatrix} A_{1n} \\ \vdots \\ A_{mn} \end{pmatrix}$. Setze also

$$A \coloneqq (\alpha(e_1), \ldots, \alpha(e_n)) = \begin{pmatrix} A_{1,1} & \ldots & A_{1,n} \\ \vdots & & \vdots \\ A_{m,1} & \ldots & A_{m,n} \end{pmatrix} \in \mathbb{Q}^{m\times n}.$$

Dann ist nach der Rechenregel für Linearkombinationen

$$\alpha(\xi) = \sum_j \xi_j \alpha(e_j) = \sum_j \xi_j A_{-,j} = \widetilde{A}(\xi),$$

wobei die letzte Gleichheit wieder Bemerkung 1.2.11 ist. □

1.3 Matrixmultiplikation und Komposition linearer Abbildungen

Lernziel: Matrixmultiplikation und Komposition von linearen Abbildungen, injektive, surjektive und bijektive lineare Abbildungen.

Erinnerung:
Sind $f : S \to T$ und $g : T \to U$ Abbildungen, so heißt die Abbildung

$$g \circ f \coloneqq gf : S \to U : s \mapsto g(f(s))$$

die **Komposition** oder **Hintereinanderausführung** von f mit g.

Wir werden jetzt zeigen:

Die Komposition linearer Abbildungen ist wieder linear.

Dies brauchen wir nur für Abbildungen der Form $\widetilde{A}$ zu zeigen, da nach Satz 1.2.15 jede lineare Abbildung von $\mathbb{Q}^{n\times 1} \to \mathbb{Q}^{m\times 1}$ von dieser Form ist:

Satz 1.3.1. *Seien $A \in \mathbb{Q}^{m\times n}, B \in \mathbb{Q}^{n\times o}$. Dann ist die Komposition $\widetilde{A} \circ \widetilde{B} : \mathbb{Q}^{o\times 1} \to \mathbb{Q}^{m\times 1}$ der linearen Abbildungen $\widetilde{A} : \mathbb{Q}^{n\times 1} \to \mathbb{Q}^{m\times 1}$ und $\widetilde{B} : \mathbb{Q}^{o\times 1} \to \mathbb{Q}^{n\times 1}$ wieder eine lineare Abbildung und es gilt*

$$\widetilde{A} \circ \widetilde{B} = \widetilde{AB}$$

wobei die Matrix $AB \in \mathbb{Q}^{m\times o}$ aus den Spalten $AB_{-,j}$ für $j = 1, \ldots, o$ besteht. Die Matrix AB heißt das **Matrixprodukt** *der Matrizen A und B.*

Beweis. (Spaltenphilosophie, in zwei Schritten)

(1) $\widetilde{A} \circ \widetilde{B}$ ist linear, denn:
Seien $r \in \mathbb{Q}, x, y \in \mathbb{Q}^{o\times 1}$. Dann gilt

$$\begin{aligned}(\widetilde{A} \circ \widetilde{B})(rx + y) &= \widetilde{A}\left(\widetilde{B}(rx + y)\right)\\ &= \widetilde{A}\left(r\widetilde{B}(x) + \widetilde{B}(y)\right)\\ &= r\widetilde{A}(\widetilde{B}(x)) + \widetilde{A}(\widetilde{B}(y))\\ &= r(\widetilde{A} \circ \widetilde{B})(x) + (\widetilde{A} \circ \widetilde{B})(y).\end{aligned}$$

(2) Ist C die nach Satz 1.2.15 existierende Matrix zu der nach (1) linearen Abbildung $\widetilde{A}\circ\widetilde{B}$, so gilt für die i-te Spalte von C, dass

$$\begin{aligned}C_{-,j} &= (\widetilde{A} \circ \widetilde{B})(e_j)\\ &= \widetilde{A}(\widetilde{B}(e_j))\\ &= \widetilde{A}(B_{-,j}) && \text{Bemerkung 1.2.10 (angewandt auf } \widetilde{B})\\ &= AB_{-,j}\end{aligned}$$

womit der Satz bewiesen ist. □

Übung 1.3.1. Führen Sie den Beweis von Satz 1.3.1 mit der Substitutionsphilosophie: $\widetilde{B}(x) = y$ bedeutet ausgeschrieben $y_i = \sum_j B_{ij}x_j$. Setzen Sie in diese Gleichung $x_j = \sum_k A_{jk}y_k$ ein.

Beispiel 1.3.2. Wir stellen uns ein System mit drei Zuständen vor, etwa eine Maus, die in einem der drei Zimmer eines Hauses ist. Eine Spalte

$$w := \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} \text{ mit } w_i \in \mathbb{Q}, \quad 0 \le w_i \le 1,\ w_1 + w_2 + w_3 = 1$$

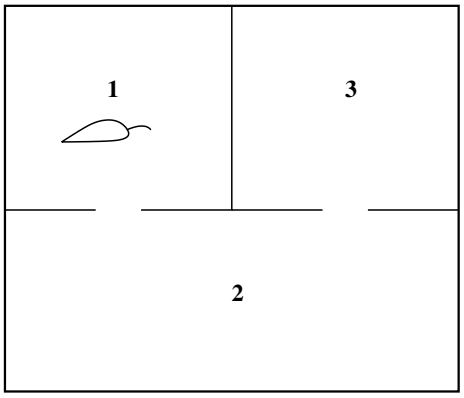
interpretieren wir als Wahrscheinlichkeitsverteilung, die uns sagt, dass das System mit Wahrscheinlichkeit w_i im Zustand i ist, etwa

$$w := \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

sagt uns, dass die Maus mit Sicherheit in Zimmer 1 ist. Wir betrachten die Matrix

$$A := \begin{pmatrix} 1/2 & 1/3 & 0 \\ 1/2 & 1/3 & 1/2 \\ 0 & 1/3 & 1/2 \end{pmatrix} \in \mathbb{Q}^{3\times 3}$$

wobei wir den Eintrag A_{ij} als Wahrscheinlichkeit dafür interpretieren, dass das System vom Zustand j in den Zustand i übergeht.



Man beachte: Für alle $j = 1, 2, 3$ gilt:

$$0 \leq A_{ij} \leq 1 \text{ für } i = 1, 2, 3 \text{ und } A_{1j} + A_{2j} + A_{3j} = 1.$$

Wir betrachten nun die Abbildung

$$\widetilde{A} : \mathbb{Q}^{3\times 1} \to \mathbb{Q}^{3\times 1} : w \mapsto Aw.$$

Wir stellen folgende Frage: Kann man mit Hilfe der Matrix A feststellen, wie sich die Wahrscheinlichkeitsverteilung beim Übergang in den neuen Zustand verändert? Für die Ausgangszustände

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

ist dies wegen Bemerkung 1.2.10 ja gerade die Definition. Welche Interpretation hat

$$\widetilde{A}^2(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}) = \widetilde{A}(\widetilde{A}(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}) = A \begin{pmatrix} 1/2 \\ 1/2 \\ 0 \end{pmatrix} = \begin{pmatrix} 5/12 \\ 5/12 \\ 1/6 \end{pmatrix}?$$

Dies ist die Wahrscheinlichkeitsverteilung nach zwei Zeittakten, wenn die Maus anfangs mit Sicherheit im ersten Zimmer war. Die Produktmatrix

$$A^2 := AA = \begin{pmatrix} 1/2 & 1/3 & 0 \\ 1/2 & 1/3 & 1/2 \\ 0 & 1/3 & 1/2 \end{pmatrix} \begin{pmatrix} 1/2 & 1/3 & 0 \\ 1/2 & 1/3 & 1/2 \\ 0 & 1/3 & 1/2 \end{pmatrix} = \begin{pmatrix} 5/12 & 5/18 & 1/6 \\ 5/12 & 4/9 & 5/12 \\ 1/6 & 5/18 & 5/12 \end{pmatrix}$$

gibt daher den Übergang innerhalb von zwei Zeittakten an. Wir werden später beweisen können, dass für sehr große n die Einträge von A^n sehr nahe an die der Matrix

$$\begin{pmatrix} 2/7 & 2/7 & 2/7 \\ 3/7 & 3/7 & 3/7 \\ 2/7 & 2/7 & 2/7 \end{pmatrix} = \frac{1}{7} \begin{pmatrix} 2 & 2 & 2 \\ 3 & 3 & 3 \\ 2 & 2 & 2 \end{pmatrix}$$

Bemerkung 1.3.3. Sei $A \in \mathbb{Q}^{m\times n}$. Dann gilt:

(1) $\widetilde{A}$ ist genau dann injektiv, wenn die Faser von $\widetilde{A}$ über der Nullspalte $0_m \in \mathbb{Q}^{m\times 1}$ nur aus der Nullspalte $0_n \in \mathbb{Q}^{n\times 1}$ besteht. Nach Bemerkung 1.2.11 heißt dies, dass 0_m sich nur trivial aus den Spalten von A linearkombinieren lässt, d.h. mit Nullen als Koeffizienten.

(2) $\widetilde{A}$ ist genau dann surjektiv, wenn sich alle Spalten aus $\mathbb{Q}^{m\times 1}$ aus den Spalten von A linearkombinieren lassen.

Ende Vorl. 2 12.10

Beweis.

(1) ($\Rightarrow$) Ist $\widetilde{A}$ injektiv, so haben alle Fasern unter $\widetilde{A}$ höchstens ein Element. Nun gilt $\widetilde{A}(0_n) = 0_m$, da $\widetilde{A}$ linear ist. Also ist $\widetilde{A}^{-1}(\{0_m\}) = \{0_n\}$.

($\Leftarrow$) Seien $x, y \in \mathbb{Q}^{n\times 1}$ mit $\widetilde{A}(x) = \widetilde{A}(y)$. Dann ist $\widetilde{A}(x-y) = 0_m$, also nach Voraussetzung $x - y = 0_n$, d.h. $x = y$.

(2) Es gilt $\widetilde{A}(x) = x_1 A_{-,1} + \ldots + x_n A_{-,n}$ (Bemerkung 1.2.11), also besteht das Bild von $\widetilde{A}$ aus allen Linearkombinationen der Spalten von A. □

An dieser Stelle ist eigentlich noch viel mehr zu sagen, etwa dass bei injektiven $\widetilde{A} : \mathbb{Q}^{n\times 1} \to \mathbb{Q}^{m\times 1}$ gilt: $n \leq m$ und ein Linksinverses kann als lineare Abbildung gewählt werden. Oder dual bei surjektivem $\widetilde{A}$ gilt: $m \leq n$ und ein Rechtsinverses kann linear gewählt werden.

Nun werden wir eine wichtige Folgerung aus der Assoziativität der Komposition von Abbildungen für die Assoziativität der Matrixmultiplikation ziehen:

Folgerung 1.3.4. *Die Matrixmultiplikation ist assoziativ, genauer: Sind $A \in \mathbb{Q}^{m\times n}, B \in \mathbb{Q}^{n\times o}, C \in \mathbb{Q}^{o\times p}$, so gilt:*

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

Beweis.

$$\begin{aligned}
\widetilde{(AB)C} &= \widetilde{(AB)} \circ \widetilde{C} && \text{Satz 1.3.1}\\
&= (\widetilde{A} \circ \widetilde{B}) \circ \widetilde{C} && \text{Satz 1.3.1}\\
&= \widetilde{A} \circ (\widetilde{B} \circ \widetilde{C}) && \text{(Komposition von Abbildungen ist assoziativ)}\\
&= \widetilde{A} \circ \widetilde{BC} && \text{Satz 1.3.1}\\
&= \widetilde{A(BC)} && \text{Satz 1.3.1}
\end{aligned}$$

Aus Satz 1.2.15 folgt nun die Behauptung. □

Hier ist eine kleine Anwendung der Assoziativität der Matrixmultiplikation in dem wahrscheinlichkeitstheoretischen Beispiel 1.3.2.

Beispiel 1.3.5. In Beispiel 1.3.2 war die Frage offengeblieben, ob $\widetilde{A}$ die Menge der Wahrscheinlichkeitsverteilungen

$$W := \{w \in \mathbb{Q}^{3\times 1} \mid 0 \leq w_i \leq 1 \text{ für alle } i = 1,2,3, w_1 + w_2 + w_3 = 1\},$$

in sich transformiert. Greifen wir uns die letzte Bedingung heraus. Man kann sie auch so schreiben:

$$\begin{pmatrix} 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} = \begin{pmatrix} 1 \end{pmatrix}.$$

Inhaltlich war die entscheidende Eigenschaft der Matrix A, dass die Einträge nicht negativ waren und die Spaltensummen alle gleich 1 sind. Letzteres bedeutet:

$$\begin{pmatrix}1 & 1 & 1\end{pmatrix} A = \begin{pmatrix}1 & 1 & 1\end{pmatrix}.$$

Somit bekommen wir aus der Assoziativität der Matrixmultiplikation

$$\begin{pmatrix}1 & 1 & 1\end{pmatrix}(Aw) = \left(\begin{pmatrix}1 & 1 & 1\end{pmatrix} A\right) w = \begin{pmatrix}1 & 1 & 1\end{pmatrix} w = \begin{pmatrix}1\end{pmatrix},$$

womit die letzte Bedingung für Aw überprüft ist. Die anderen Bedingungen zu überprüfen lasse ich als eine analoge Übung.

Erinnerung: Genau dann ist eine Abbildung $f : X \to Y$ bijektiv, falls eine Abbildung $g : Y \to X$ existiert mit

$$g \circ f = \mathrm{id}_X \text{ und } f \circ g = \mathrm{id}_Y.$$

Eine solche Abbildung ist notwendigerweise eindeutig bestimmt:

$$g' = g' \circ \mathrm{id}_Y = g' \circ f \circ g = \mathrm{id}_X \circ g = g.$$

Man nennt sie die zu f **inverse Abbildung** oder einfach das **Inverse** von f und schreibt $g =: f^{-1}$.

Bemerkung 1.3.6. Ist $\widetilde{A} : \mathbb{Q}^{n\times 1} \to \mathbb{Q}^{n\times 1}$ bijektiv, so ist $(\widetilde{A})^{-1} : \mathbb{Q}^{n\times 1} \to \mathbb{Q}^{n\times 1}$ linear.

Beweis. Seien $x, y \in \mathbb{Q}^{n\times 1}$ und $r \in \mathbb{Q}$. Wir wollen zeigen, dass

$$(\widetilde{A})^{-1}(rx + y) = r(\widetilde{A})^{-1}(x) + (\widetilde{A})^{-1}(y) \tag{$-$}$$

gilt. Dazu wenden wir die injektive Abbildung $\widetilde{A}$ auf beide Seiten an:

$$\begin{aligned}\widetilde{A}(r(\widetilde{A})^{-1}(x) + (\widetilde{A})^{-1}(y)) &= r\widetilde{A}((\widetilde{A})^{-1}(x)) + \widetilde{A}((\widetilde{A})^{-1}(y)) \\ &= rx + y \\ \widetilde{A}((\widetilde{A})^{-1}(rx + y)) &= rx + y\end{aligned}$$

Da $\widetilde{A}$ injektiv ist, folgt aus der Gleichheit der Bilder die Gleichung ($-$). □

Definition 1.3.7. Eine Matrix $A \in \mathbb{Q}^{n\times n}$ heißt **invertierbar**, falls $\widetilde{A} : \mathbb{Q}^{n\times 1} \to \mathbb{Q}^{n\times 1}$ bijektiv ist. In dem Fall heißt die eindeutig bestimmte Matrix $B \in \mathbb{Q}^{n\times n}$ mit $AB = BA = I_n$ die zu A **inverse Matrix** und wird mit A^{-1} bezeichnet, d.h.

$$(\widetilde{A})^{-1} = \widetilde{A^{-1}}.$$

Beim Verständnis der Matrixmultiplikation helfen uns Zeilen auch weiter. Wenn wir die beiden Beweise des Satzes 1.3.1 analysieren, so kommen wir zu dem Schluss, dass der erste Beweis spaltenorientiert war, der Substitutionsbeweis aber zeilenorientiert. Die folgende Bemerkung gibt ein ausgewogenes Bild.

Bemerkung 1.3.8. Sei $A \in \mathbb{Q}^{m\times n}, B \in \mathbb{Q}^{n\times o}$ und $C = AB \in \mathbb{Q}^{m\times o}$. Weiter bezeichne $A_{i,-} := (A_{i1}, \ldots, A_{in}) \in \mathbb{Q}^{1\times n}$ die i-te Zeile von A. Dann gilt:

(1) $C_{ij} = A_{i,-}B_{-,j}$ (Zeile mal Spalte).

(2) $C_{-,j} = AB_{-,j}$ (spaltenorientiert).

(3) $C_{i,-} = A_{i,-}B$ (zeilenorientiert).

(4) $C = A_{-,1}B_{1,-} + A_{-,2}B_{2,-} + \ldots + A_{-,n}B_{n,-}$ (Spalte mal Zeile).

Beweis. Übung. □

1.4 Der Gaußsche Algorithmus

Lernziel: Gaußsches Eliminationsverfahren mit Anwendungen auf Bestimmung von Lösungsmengen linearer Gleichungssysteme, Invertieren von Matrizen, Transponieren von Matrizen.

Wir wollen ein Verfahren kennenlernen (oder für die meisten wiederholen), welches die Faser über einem Punkt im Bildbereich unter der linearen Abbildung bestimmt. Es handelt sich um den Gaußschen Algorithmus, den Carl Friedrich Gauß vor etwa 200 Jahren für die Behandlung astronomischer Fragestellungen entwickelt hat. Es hat sich gezeigt, dass man dieses Verfahren schon vor 2000 Jahren in China kannte[4], ist dort aber wohl wieder in Vergessenheit geraten. Unsere Ausgangssituation ist das lineare Gleichungssystem

$$Ax = b \tag{*}$$

mit $A \in \mathbb{Q}^{m\times n}$ und $b \in \mathbb{Q}^{m\times 1}$ fest vorgegeben. Gesucht ist die Faser von $\widetilde{A}$ über b, also alle $x \in \mathbb{Q}^{n\times 1}$, die (*) erfüllen. Ausgeschrieben haben wir also

$$\begin{array}{rcl} A_{11}x_1 + A_{12}x_2 + \ldots + A_{1n}x_n & = & b_1 \\ A_{21}x_1 + A_{22}x_2 + \ldots + A_{2n}x_n & = & b_2 \\ \vdots & & \\ A_{m1}x_1 + A_{m2}x_2 + \ldots + A_{mn}x_n & = & b_m \end{array} \tag{**}$$

Statt dies immer auszuschreiben, arbeiten wir einfach mit der erweiterten Matrix $(A|b)$ des linearen Gleichungssystems. Der senkrechte Strich deutet an, wo die Matrix des Gleichungssystems aufhört und die rechte Seite anfängt. Es versteht sich von selbst, dass nicht alle linearen Probleme gleich in dieser Gestalt gegeben sind, sondern dass man manchmal etwas dafür arbeiten muss damit man diese Gestalt erhält, einige Beispiele dafür erhalten Sie in den Übungen.

Die erste Beobachtung besteht darin, dass es besonders einfache Situationen gibt, in denen man die Lösungen fast direkt ablesen kann.

Definition 1.4.1. Sei $M \in \mathbb{Q}^{s\times t}$ eine Matrix.

(1) Für $i \in \underline{s}$ ist der i-te **Stufenindex** $St_i(M)$ definiert als

$$St_i(M) \coloneqq min\{j \in \underline{t} \mid M_{ij} \neq 0\}$$

Falls $M_{i,-}$ die Nullzeile ist, setzen wir $St_i(M) \coloneqq t + i$.

(2) M ist in **Stufenform**, falls die Folge

$$St(M) \coloneqq (St_1(M), St_2(M), \ldots, St_s(M))$$

streng monoton steigend ist, d.h.

$$St_1(M) < St_2(M) < \cdots < St_s(M).$$

Falls zusätzlich noch für jeden Stufenindex $j \coloneqq St_i(M) \leq t$ gilt, $M_{-,j} = (I_s)_{-,j}$, so ist M in **strikter Stufenform** (bzw. **reduzierter Stufenform**).

[4] FANG-CHENG-ALGORITHMUS, vgl. P. Gabriel: Matrizen, Geometrie, lineare Algebra

Beispiel 1.4.2.

$$\begin{array}{rrrrcl} 1z_1+ & 2z_2+ & 3z_3+ & 21z_4 & = & 2 \\ 2z_1+ & 4z_2+ & 4z_3+ & 28z_4 & = & 3 \end{array}$$

hat zugeordnete Matrix

$$\left(\begin{array}{cccc|c} 1 & 2 & 3 & 21 & 2 \\ 2 & 4 & 4 & 28 & 3 \end{array}\right)$$

hat Stufenfolge $(1,1)$, ist also nicht in Stufenform. Hingegen

$$\left(\begin{array}{cccc|c} 1 & 2 & 3 & 21 & 2 \\ 0 & 0 & -2 & -14 & -1 \end{array}\right)$$

hat Stufenfolge $(1,3)$, ist also in Stufenform, jedoch nicht in strikter Stufenform. Letztere liegt bei

$$\left(\begin{array}{cccc|c} 1 & 2 & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 1 & 7 & \frac{1}{2} \end{array}\right)$$

vor, wo die Stufenspalten entsprechende Einheitsspalten sind. Das zugehörige lineare Gleichungssystem ist

$$\begin{array}{rrrrcl} z_1+ & 2z_2 & & & = & \frac{1}{2} \\ & & z_3+ & 7z_4 & = & \frac{1}{2} \end{array}$$

Dieses Gleichungssystem kann man nun sehr leicht lösen: Man fügt für jeden Nichtstufenindex einen Parameter p_i und eine neue Gleichung ein. Im vorliegenden Fall: $z_2 = p_1, z_4 = p_2$. Durch Addition geeigneter Vielfache dieser neuen Gleichungen bringt man das neue Gleichungssystem in die strikte Stufenform und kann alle Lösungen ablesen:

$$\begin{array}{rrrrcl} z_1+ & 2z_2 & & & = & \frac{1}{2} \\ & & z_3+ & 7z_4 & = & \frac{1}{2} \\ & z_2 & & & = & p_1 \\ & & & z_4 & = & p_2 \end{array} \qquad \left(\begin{array}{cccc|c} 1 & 2 & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 1 & 7 & \frac{1}{2} \\ 0 & 1 & 0 & 0 & p_1 \\ 0 & 0 & 0 & 1 & p_2 \end{array}\right)$$

Also

$$\left(\begin{array}{cccc|c} 1 & 0 & 0 & 0 & \frac{1}{2}-2p_1 \\ 0 & 1 & 0 & 0 & p_1 \\ 0 & 0 & 1 & 0 & \frac{1}{2}-7p_2 \\ 0 & 0 & 0 & 1 & p_2 \end{array}\right) \text{ also } \begin{array}{l} z_1 = \frac{1}{2}-2p_1 \\ z_2 = p_1 \\ z_3 = \frac{1}{2}-7p_2 \\ z_4 = p_2 \text{ mit } p_1, p_2 \in \mathbb{Q} \text{ beliebig.} \end{array}$$

Ende Vorl. 3 17.10

Bei einem linearen Gleichungssystem in Stufenform ist somit die Gesamtheit der Lösungen ablesbar. Leider ist nicht jedes Gleichungssystem in dieser Form gegeben. Hier ist eine ganz allgemeine Strategie, die über den Fall der linearen Gleichungssysteme hinausgeht, wie man ein solches Problem angehen kann.

Bemerkung 1.4.3. Sei $f : N \to M$ und $m \in M$. Für jede injektive Selbstabbildung $g : M \to M$ gilt für die Faser[5] von f über m:

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

Beweis.

$$\begin{array}{lll} n \in (g \circ f)^{-1}(\{g(m)\}) & \iff & (g \circ f)(n) = g(m) \\ & \iff & g(f(n)) = g(m) \\ & \iff & f(n) = m \qquad \text{(für „⇒“ brauchen wir } g \text{ injektiv)} \\ & & \qquad\qquad\quad \text{(und für „⇐“ nur } g \text{ Abbildung)} \\ & \iff & n \in f^{-1}(\{m\}). \end{array}$$

□

[5] Die Faser $f^{-1}(\{m\})$ von f über m ist die Lösungsmenge des Gleichungssystems $f(n) = m$.

Die sich ergebende Strategie ist somit, eine Folge von injektiven und daher fasererhaltenden Selbstabbildungen anzuwenden, bis man die Lösungen hoffentlich ablesen kann. Im vorliegenden Fall der linearen Gleichungssysteme wird man selbstverständlich die Kategorie der *linearen* Abbildungen nicht verlassen wollen.

Um die formale Definition dieser injektiven linearen Selbstabbildungen besser verstehen zu können, bemerken wir zuerst:

Bemerkung 1.4.4. Seien $m \in \mathbb{N}$ und $e_i := (I_m)_{-,i} \in \mathbb{Q}^{m\times 1}$ (die i-te Einheitsspalte).

(1) Ist $z \in Q^{1\times o}$ eine Zeile, so ist $e_i z$ die $m \times o$-Matrix, die z als i-te Zeile hat und sonst aus Nullen besteht.

(2) Sei $f_j := (I_m)_{j,-} \in \mathbb{Q}^{1\times m}$ die j-te **Einheitszeile**. Nach 1. hat die $m \times m$-Matrix $e_i f_j$ genau eine 1 an der Stelle (i, j) und sonst Nullen.

(3) Für $M \in \mathbb{Q}^{m\times n}$ ist $f_j M = M_{j,-}$, die j-te Zeile von M (vgl. Bemerkung 1.2.10).

(4) Insbesondere ist $e_i f_j M$ ebenfalls eine $m \times n$-Matrix, die die j-te Zeile von M als ihre i-te Zeile hat und sonst aus Nullen besteht.

Hier ist eine Auswahl von quadratischen Matrizen, die injektive, sogar bijektive[6] Selbstabbildungen induzieren. Diese Matrizen kann man also von links an die erweiterte Matrix eines linearen Gleichungssystems multiplizieren, ohne die Lösungsmenge zu verändern, wie wir aus Bemerkung 1.4.3 bereits wissen.

Definition 1.4.5. Seien n, e_i, f_j wie in der letzten Bemerkung. Jede Matrix, welche von einem der drei nachfolgenden Typen ist, heißt **elementare Umformungsmatrix**:

(1) Für $1 \le i, j \le m$ mit $i \neq j$ und $a \in \mathbb{Q}$ sei

$$\mathrm{Add}_m(i, j; a) := I_m + a e_i f_j : \underline{m} \times \underline{m} \to \mathbb{Q} : (p, q) \mapsto \begin{cases} 1 & \text{falls } p = q \\ a & \text{falls } (p, q) = (i, j) \\ 0 & \text{sonst} \end{cases}$$

(2) Für $1 \le i \le m$ und $a \in \mathbb{Q} - \{0\}$ sei

$$\mathrm{Mul}_m(i; a) := I_m + (a-1) e_i f_i : \underline{m} \times \underline{m} \to \mathbb{Q} : (p, q) \mapsto \begin{cases} 1 & \text{falls } p = q \neq i \\ a & \text{falls } (p, q) = (i, i) \\ 0 & \text{sonst} \end{cases}$$

(3) Für $1 \le i < j \le m$ sei

$$\mathrm{Ver}_m(i, j) : \underline{m} \times \underline{m} \to \mathbb{Q} : (p, q) \mapsto \begin{cases} 1 & \text{falls } p = q \notin \{i, j\} \\ 1 & \text{falls } (p, q) \in \{(i, j), (j, i)\} \\ 0 & \text{sonst} \end{cases}$$

Es gilt also

$$\begin{aligned} \mathrm{Ver}_m(i, j) &:= (e_1 \dots e_{i-1} e_j e_{i+1} \dots e_{j-1} e_i e_{j+1} \dots e_m) \\ &= (f_1 \dots f_{i-1} f_j f_{i+1} \dots f_{j-1} f_i f_{j+1} \dots f_m)\,. \end{aligned}$$

Auch im Hinblick auf das Lösen linearer Gleichungssysteme sind folgende Bemerkungen wichtig:

[6] Wir werden später beweisen, dass jede injektive lineare Selbstabbildung auf $M = \mathbb{Q}^{m\times 1}$ automatisch bijektiv ist.

Bemerkung 1.4.6. Sei $M \in \mathbb{Q}^{m\times n}$.

(1) Das Produkt $\mathrm{Add}_m(i,j;a)M$ unterscheidet sich von M nur in der i-ten Zeile, welche in dem Produkt gleich $M_{i,-} + aM_{j,-}$ ist. Insbesondere ist $\mathrm{Add}_m(i,j;a) \in \mathbb{Q}^{m\times m}$ invertierbar mit inverser Matrix $\mathrm{Add}_m(i,j;-a)$.

(2) Das Produkt $\mathrm{Mul}_m(i,a)M$ hat mit M alle Zeilen außer der i-ten gemeinsam, und diese ist gleich $aM_{i,-}$. Insbesondere ist $\mathrm{Mul}_m(i,a) \in \mathbb{Q}^{m\times m}$ invertierbar mit Inverser $\mathrm{Mul}_m(i,a^{-1})$.

(3) Im Produkt $\mathrm{Ver}_m(i,j)M$ sind gegenüber M die i-te und die j-te Zeile vertauscht. Insbesondere ist $\mathrm{Ver}_m(i,j) \in \mathbb{Q}^{m\times m}$ zu sich selbst invers.

Beweis. Übung. □

Algorithmus 1.4.7 (Gauß-Algorithmus)**.**
Gegeben: $M \in \mathbb{Q}^{m\times n}$.
Output: Eine $m \times n$-Matrix in Stufenform bzw. strikte Stufenform, die durch Linksmultiplikation mit elementaren Umformungsmatrizen aus M hervorgeht.
Algorithmus:

(1) Überführe M in eine Stufenform:

(a) Finde den kleinsten Spaltenindex j mit $M_{-,j} \neq 0$ (sprich finde die erste Spalte, die nicht Null ist).

(b) Finde den kleinsten Zeilenindex i mit $M_{i,j} \neq 0$.

(c) Falls $i \neq 1$, vertausche die erste und i-te Zeile, d.h. ersetze

$$M \rightsquigarrow \mathrm{Ver}_m(1,i)M$$

so, dass wir jetzt $i = 1$ haben.

(d) Falls $M_{1,j} \neq 1$, ersetze

$$M \rightsquigarrow \mathrm{Mul}_m(1, M_{1,j}^{-1})M$$

so, dass wir mit $M_{1,j} = 1$ weiterarbeiten können.

(e) Räume die j-te Spalte aus durch Subtraktion der $M_{i,j}$-Vielfachen der ersten Zeile von der i-ten Zeile, d.h. ersetze der Reihe nach

$$M \rightsquigarrow \mathrm{Add}_m(i,1;-M_{i,j})M$$

für $i = 2,\dots,m$.

(f) Wiederhole (1e) mit der Teilmatrix von M, die durch Streichen (bzw. Ignorierung) der ersten Zeile und der ersten j Spalten hervorgeht. Am Ende hat man eine Matrix M in Stufenform.

(2) Überführe M in eine *strikte* Stufenform:
Bringe die resultierende Matrix durch die Linksmultiplikationen mit elementare Umformungsmatrizen vom Typ $\mathrm{Add}_m(i,j;a)$ auf strikte Stufenform.

Beweis. Wir müssen nur noch zeigen, dass der Algorithmus nach endlich vielen Schritten terminiert. Dies sieht man wie folgt: Nach spätestens $n-1$ Übergängen zu Teilmatrizen wird die Stufenform erreicht. Hat eine dieser Teilmatrizen r Zeilen, so sind zu jedem Übergang maximal $1+1+(r-1)$ Zeilenumformungen notwendig. Die Anzahl der Zeilenumformungen im zweiten Teil kann man auch leicht abschätzen (Übung). □

Bemerkung 1.4.8. Falls man über $\mathbb{Q}$ arbeitet ergeben sich für den Schritt (1b) mehrere Alternativstrategien. Hier ist eine davon: Finde den kleinsten Zeilenindex i mit $M_{i,j} \neq 0$, so dass eine Division der i-ten Zeile durch $M_{i,j}$ eine ganzzahlige Zeile ergibt. Falls es keine solche Zeile gibt, verfahre wie in (1b).

Beispiel 1.4.9.

$$\begin{pmatrix} 3 & 0 & 6 & 3 & 3 \\ 2 & 0 & 4 & 3 & 5 \\ 0 & 3 & 0 & 2 & 3 \\ 1 & 2 & 2 & 3 & 5 \end{pmatrix} \underset{\rightsquigarrow}{\mathrm{Mul}_4(1;\frac{1}{3})} \begin{pmatrix} \mathbf{1} & \mathbf{0} & \mathbf{2} & \mathbf{1} & \mathbf{1} \\ 2 & 0 & 4 & 3 & 5 \\ 0 & 3 & 0 & 2 & 3 \\ 1 & 2 & 2 & 3 & 5 \end{pmatrix} \underset{\rightsquigarrow}{\mathrm{Add}_4(2,1;-2)} \begin{pmatrix} 1 & 0 & 2 & 1 & 1 \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{3} \\ 0 & 3 & 0 & 2 & 3 \\ 1 & 2 & 2 & 3 & 5 \end{pmatrix}$$

$$\underset{\rightsquigarrow}{\mathrm{Add}_4(4,1;-1)} \begin{pmatrix} 1 & 0 & 2 & 1 & 1 \\ 0 & 0 & 0 & 1 & 3 \\ 0 & 3 & 0 & 2 & 3 \\ \mathbf{0} & \mathbf{2} & \mathbf{0} & \mathbf{2} & \mathbf{4} \end{pmatrix} \underset{\rightsquigarrow}{\mathrm{Ver}_4(2,4)} \begin{pmatrix} 1 & 0 & 2 & 1 & 1 \\ \mathbf{0} & \mathbf{2} & \mathbf{0} & \mathbf{2} & \mathbf{4} \\ 0 & 3 & 0 & 2 & 3 \\ 0 & 0 & 0 & 1 & 3 \end{pmatrix}$$

$$\underset{\rightsquigarrow}{\mathrm{Mul}_4(2;\frac{1}{2})} \begin{pmatrix} 1 & 0 & 2 & 1 & 1 \\ \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{1} & \mathbf{2} \\ 0 & 3 & 0 & 2 & 3 \\ 0 & 0 & 0 & 1 & 3 \end{pmatrix} \underset{\rightsquigarrow}{\mathrm{Add}_4(3,2;-3)} \begin{pmatrix} 1 & 0 & 2 & 1 & 1 \\ 0 & 1 & 0 & 1 & 2 \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{-1} & \mathbf{-3} \\ 0 & 0 & 0 & 1 & 3 \end{pmatrix}$$

$$\underset{\rightsquigarrow}{\mathrm{Mul}_4(3;-1)} \begin{pmatrix} 1 & 0 & 2 & 1 & 1 \\ 0 & 1 & 0 & 1 & 2 \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{3} \\ 0 & 0 & 0 & 1 & 3 \end{pmatrix} \underset{\rightsquigarrow}{\mathrm{Add}_4(4;3,-1)} \begin{pmatrix} 1 & 0 & 2 & 1 & 1 \\ 0 & 1 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 & 3 \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix}$$

$$\underset{\rightsquigarrow}{\mathrm{Add}_4(1,3,-1)} \begin{pmatrix} \mathbf{1} & \mathbf{0} & \mathbf{2} & \mathbf{0} & \mathbf{-2} \\ 0 & 1 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \underset{\rightsquigarrow}{\mathrm{Add}_4(2,3,-1)} \begin{pmatrix} 1 & 0 & 2 & 0 & -2 \\ \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{0} & \mathbf{-1} \\ 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Ende Vorl. 4 19.10

Algorithmus 1.4.10 (Gauß-Algorithmus zum Lösen)**.**
Gegeben: $M \in \mathbb{Q}^{m\times(n+1)}$ erweiterte Matrix eines linearen Gleichungssystems.
Gesucht: Lösungsmenge des linearen Gleichungssystems.

(1) Wende den Gauß-Algorithmus auf M an und erhalte eine strikte Stufenform mit derselben Lösungsmenge. *Letztere ist genau dann leer, wenn $n+1$ ein Stufenindex ist.*

(2) Mache Lösungen explizit:
Falls $n+1$ kein Stufenindex ist, streiche die Nullzeilen von M und füge für jeden Nichtstufenindex i_ℓ mit $\ell = 1, \ldots, d$ der linken Seite eine neue Zeile $((I_n)_{i_\ell,-}|p_\ell)$ zu der erweiterten Matrix hinzu, wo p_ℓ paarweise verschiedene Parameter sind. Bringe die resultierende Matrix durch die Linksmultiplikationen mit elementaren Umformungsmatrizen vom Typ $\mathrm{Ver}_m(i,j)$ und $\mathrm{Add}_m(i,j,a)$ wieder auf strikte Stufenform. Diese ist gegeben durch (I_n, L), wo L eine Spalte ist, die die Lösungen in Abhängigkeit von den Parametern angibt.

Beweis. Wir müssen zeigen, dass wir am Ende wirklich sehen können, ob eine Lösung existiert und alle Lösungen ablesbar sind. Nach Bemerkung 1.4.3 und der Invertierbarkeit der Umformungsmatrizen aus Definition 1.4.5 ändert sich die Lösungsmenge nicht. Das Weglassen von Nullzeilen ist kein Informationsverlust und das Hinzufügen der Zeilen mit den Parametern nur eine Namensgebung. □

Mit dem Gaußschen Algorithmus sind wir im Besitz einer Schlüsseltechnologie. Mit seiner Hilfe können wir nicht nur Fasern von linearen Abbildungen bestimmen, sondern viele der Begriffe aus dem Abbildungsabschnitt algorithmisch beherrschen:

- Bestimmung des Bildes einer linearen Abbildung (vgl. Beispiel 1.4.11),
- Bestimmung von linearen Rechtsinversen (vgl. Beispiel 1.4.12) bzw. Linksinversen,
- Surjektivitätstests und Injektivitätstests.
- ...

Wir begnügen uns jeweils mit Beispielen.

Wir wollen jetzt mit Hilfe des Gaußschen Algorithmus das Bild einer linearen Abbildung bestimmen. Man kann natürlich sagen, dass das Bild einfach aus allen Linearkombinationen der Spalten der Matrix besteht. Dies ist Bemerkung 1.2.11. Diese **explizite** Beschreibung eignet sich in der Tat dafür, alle Bilder zu *parametrisieren*. Dagegen eignet sie sich diese explizite Beschreibung schlecht dafür, um schnell entscheiden zu können, ob eine vorgegebene Spalte aus dem Zielbereich im Bild ist oder nicht. Dafür wäre eine **implizite** Beschreibung des Bildes (etwa als Lösungsmenge einer Gleichung) wesentlich vorteilhafter.

Wir gehen hier so vor, dass wir die rechte Seite mit Unbestimmten vorbesetzen, den ersten Teil des Gaußschen Algorithmus Algorithmus 1.4.10 durchführen und dann durch die Lösbarkeitsbedingung ein lineares Gleichungssystem für die Lösungsmenge bekommen.

Beispiel 1.4.11.
Aufgabe: Beschreibe $\mathrm{Bild}\,\widetilde{A}$ implizit, wo A gegeben ist durch

$$A := \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 6 \end{pmatrix}$$

Lösung: Wir wenden den Gaußschen Algorithmus auf die Matrix

$$\left(\begin{array}{cccc|c} 1 & 2 & 3 & 4 & x \\ 2 & 3 & 4 & 5 & y \\ 3 & 4 & 5 & 6 & z \end{array}\right)$$

an. Also:

$$\left(\begin{array}{cccc|c} 1 & 2 & 3 & 4 & x \\ 0 & -1 & -2 & -3 & y-2x \\ 0 & -2 & -4 & -6 & z-3x \end{array}\right) \rightsquigarrow \left(\begin{array}{cccc|c} 1 & 2 & 3 & 4 & x \\ 0 & 1 & 2 & 3 & -y+2x \\ 0 & 0 & 0 & 0 & z+x-2y \end{array}\right)$$

Also $\begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathrm{Bild}(\widetilde{A})$ genau dann, wenn $z+x-2y=0$. Hieran kann man sofort entscheiden, ob eine vorgegebene Spalte im Bild von A liegt oder nicht.

Für den Abschluss dieses Kapitels brauchen wir noch eine kleine Erinnerung aus der Mengenlehre. So wie man bijektive Abbildungen durch die Existenz einer Inversen charakterisieren kann, kann man surjektive resp. injektive Abbildungen durch die Existenz von Rechts- resp. Linksinversen charakterisieren.

Erinnerung: Sei $f : M \to N$ eine Abbildung.

(1) Genau dann ist f surjektiv, wenn es eine **Rechtsinverse** von f gibt, also eine Abbildung $g : N \to M$ mit $f \circ g = \mathrm{id}_N$.

(2) Genau dann ist f injektiv, wenn es eine **Linkssinverse** von f gibt, also eine Abbildung $h : N \to M$ mit $h \circ f = \mathrm{id}_M$.

Als nächstes wollen wir an einem Beispiel diskutieren, wie man eine Rechtsinverse für surjektive lineare Abbildungen bestimmen kann. Als Vorübung eine kleine Aufgabe:

Übung 1.4.1. Stellt $A \in \mathbb{Q}^{m\times n}$ eine surjektive lineare Abbildung $\widetilde{A} : \mathbb{Q}^{n\times 1} \to \mathbb{Q}^{m\times 1}$ dar, so gibt es ein lineares Rechtsinverses von $\widetilde{A}$, d.h. eine Matrix $B \in \mathbb{Q}^{n\times m}$ mit $AB = I_m$.
Hinweis: Überlege, warum man nur die Bilder der Standardbasisvektoren von $\mathbb{Q}^{m\times 1}$ verfolgen muss.

Beispiel 1.4.12.
Aufgabe: Sei $A \in \mathbb{Q}^{2\times 3}$ gegeben durch

$$A := \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \end{pmatrix}.$$

Man überprüfe, ob $\widetilde{A} : \mathbb{Q}^{3\times 1} \to \mathbb{Q}^{2\times 1}$ ein Rechtsinverses hat und berechne gegebenenfalls alle linearen Rechtsinversen.
Lösung: Gesucht sind alle Matrizen $B \in \mathbb{Q}^{3\times 2}$ mit $AB = I_2$. Somit haben wir zwei lineare Gleichungssysteme zu lösen, nämlich

$$AB_{-,1} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \text{ und } AB_{-,2} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Da diese beiden Gleichungssysteme dieselbe linke Seite, nämlich A, haben, kann man sie simultan lösen, indem man beide rechten Seiten zu einer zweispaltigen rechten Seite I_2 zusammenfasst und dann den Gaußschen Algorithmus anwendet:

$$\left(\begin{array}{ccc|cc} 1 & 2 & 3 & 1 & 0 \\ 2 & 3 & 4 & 0 & 1 \end{array}\right) \rightsquigarrow \left(\begin{array}{ccc|cc} 1 & 2 & 3 & 1 & 0 \\ 0 & -1 & -2 & -2 & 1 \end{array}\right) \rightsquigarrow \left(\begin{array}{ccc|cc} 1 & 0 & -1 & -3 & 2 \\ 0 & 1 & 2 & 2 & -1 \\ 0 & 0 & 1 & a & b \end{array}\right) \rightsquigarrow$$

$$\left(\begin{array}{ccc|cc} 1 & 0 & 0 & -3+a & 2+b \\ 0 & 1 & 0 & 2-2a & -1-2b \\ 0 & 0 & 1 & a & b \end{array}\right).$$

Alle linearen Rechtsinversen sind somit durch die Matrizen

$$\begin{pmatrix} -3+a & 2+b \\ 2-2a & -1-2b \\ a & b \end{pmatrix}$$

gegeben mit $a, b \in \mathbb{Q}$ beliebig. Man nennt auch diese Matrizen die rechtsinversen Matrizen von A. Man beachte jedoch, dass unter den rechtsinversen Abbildungen von $\widetilde{A}$ auch nichtlineare Abbildungen gibt.

Übung 1.4.2. Das surjektive $\widetilde{A}$ ist genau dann bijektiv, wenn sein Rechtsinverses $\widetilde{B}$ eindeutig bestimmt ist. In diesem Fall ist es gleichzeitig Linksinverses und somit Inverses von $\widetilde{A}$.
Hinweis: Benutze den Hinweis von Übung 1.4.1.

Linksinverse für injektive lineare Abbildungen kann man durch einen kleinen Trick auf die Bestimmung von Rechtsinversen zurückführen.

Definition 1.4.13. Ist $A : \underline{m} \times \underline{n} \to \mathbb{Q} : (i, j) \mapsto A_{ij}$ eine Matrix, so heißt

$$A^{tr} : \underline{n} \times \underline{m} \to \mathbb{Q} : (i, j) \mapsto A_{ji}$$

die **transponierte Matrix** oder einfach die **Transponierte** von A.

Lemma 1.4.14. *Sind $A \in \mathbb{Q}^{m\times n}$ und $B \in \mathbb{Q}^{n\times o}$, so sind $B^{tr} \in \mathbb{Q}^{o\times n}$ und $A^{tr} \in \mathbb{Q}^{n\times m}$ und*

$$(AB)^{tr} = B^{tr}A^{tr}.$$

Ist $\widetilde{A}$ surjektiv mit Rechtsinversem $\widetilde{B}$, so ist $\widetilde{A^{tr}}$ injektiv mit Linksinversem $\widetilde{B^{tr}}$ und umgekehrt.

Beweis. Übung. □

Kapitel 2

Zahlen, Vektoren, Polynome

2.1 Gruppen, Ringe und Körper

Lernziel: Gruppenaxiome und ihre Bedeutung, Rechnen in Körpern im Sinne von Addieren, Subtrahieren, Multiplizieren und Dividieren. Beispiele von Körpern: $\mathbb{Q}, \mathbb{R}, \mathbb{C}, \mathbb{F}_p$. Rolle von $\mathbb{Z}$, EUKLIDischer Algorithmus und Teilbarkeitstheorie in $\mathbb{Z}$.

Jetzt ist der Zeitpunkt gekommen, wo wir erstmalig axiomatisch an unsere Probleme herangehen wollen. Es geht also darum, dass man sich fragt: Was ist der allgemeinste Rahmen für meine Schlüsse und Rechnungen? Kann ich aus einer Rechnung in einer konkreten Situation auf eine allgemeine Vermutung kommen und diese dann durch Übertragung der Schlüsse auch beweisen? Kann ich Analogien zwischen Situationen sehen, wo der Außenstehende keine Gemeinsamkeiten ahnt? Der erste Begriff, den wir kennenlernen wollen, ist der der Gruppe, welcher sich im Laufe des 19. Jahrhunderts herausgebildet hat. Zunächst werden wir ihn nur zur Definition von Zahlbereichen heranziehen, später werden wir sehen, dass er auch außerhalb der Zahlbereiche eine grundlegende Rolle spielt.

Definition 2.1.1. Sei G eine nicht leere Menge und $\cdot : G \times G \to G$ eine Verknüpfung auf G. Man nennt $(G, \cdot)$ **Gruppe**, falls $\cdot$ folgende drei Axiome erfüllt:

(1) (Assoziativgesetz) Für alle $x, y, z \in G$ gilt

$$x \cdot (y \cdot z) = (x \cdot y) \cdot z.$$

(2) (Einselement oder neutrales Element) Es existiert ein eindeutiges Element $1 \in G$, so dass für alle $g \in G$ gilt

$$1 \cdot g = g \cdot 1 = g.$$

(3) (Inverses Element) Zu jedem $g \in G$ existiert ein $g^{-1} \in G$ mit

$$g \cdot g^{-1} = g^{-1} \cdot g = 1.$$

Oft schreibt man gh statt $g \cdot h$.
Falls noch folgende Bedingung gilt, heißt G **Abelsche Gruppe** oder **kommutative Gruppe**:

(4) (Kommutativgesetz) Für alle $g, h \in G$ gilt

$$gh = hg.$$

Bei Abelschen Gruppen wird manchmal $+$ statt $\cdot$ als Verknüpfungssymbol genommen. Dann bezeichnet man das neutrale Element mit 0 und das Inverse von g mit $-g$. (Im Unterschied zu $\cdot$, lässt man $+$ nicht weg.)

Übung 2.1.1. Zeigen Sie folgende Aussagen:

(1) Beim zweiten Axiom (Existenz und Eindeutigkeit des Einselementes) kann eine schwächere Formulierung benutzt werden: Die Existenz impliziert bereits die Eindeutigkeit.

(2) Beim dritten Axiom ist das Inverse eines Elementes $g \in G$ eindeutig bestimmt.

Übung 2.1.2. Sei $(G, \cdot)$ Gruppe. Für $a, b \in G$ drücke man $(ab)^{-1}$ durch a^{-1} und b^{-1} aus.

Erinnerung: Sind M, N Mengen, so bezeichnet

$$M^N := \{f | f : N \to M\}$$

die Menge der Abbildungen von N nach M.

Bemerkung 2.1.2. Es gibt diverse Abschwächungen des Gruppenbegriffs: $(G, \cdot)$ heißt

- **Halbgruppe**, falls (1) erfüllt ist;
- **Monoid** oder **Halbgruppe mit Eins**, falls (1) und (2) erfüllt sind.

Beispiel 2.1.3.

(1) Für $(\mathbb{N}, +)$ gilt nur das Assoziativgesetz und das Kommutativgesetz. (Kommutative Halbgruppe).

Ende Vorl. 5 24.10

(2) Für $a \in \mathbb{Z}$ sei $Z_{\geq a} := \{r \in \mathbb{Z} | r \geq a\}$. Dann sind für $(\mathbb{Z}_{\geq 0}, +)$ das Assoziativgesetz, die Existenz des neutralen Elementes, nämlich 0 und das Kommutativgesetz erfüllt, (Kommutative Halbgruppe mit Eins, kommutatives Monoid).

(3) Für $(\mathbb{Z}_{\geq -1}, +)$ ist nicht einmal $+$ als Verknüpfung definiert.

(4) $(\mathbb{Z}, \cdot)$ ist kommutative Halbgruppe mit Eins, aber keine Gruppe: $0a = 0$ für alle $a \in \mathbb{Z}$.

(5) $\mathbb{Q}^* := (\mathbb{Q} - \{0\}, \cdot)$ ist eine kommutative Gruppe. Ebenso $\mathbb{R}^* := (\mathbb{R} - \{0\}, \cdot)$.

(6) Da die Komposition von Abbildungen assoziativ ist, ist $(M^M, \circ)$ eine Halbgruppe mit $1 = \mathrm{id}_M$ für jede Menge M. Diese ist nicht kommutativ, sobald M mehr als ein Element enthält. (Nicht kommutatives Monoid).

(7) $(S_M, \circ) := (\mathrm{Sym}(M), \circ) := (\{f \in M^M | f \text{ bijektiv}\}, \circ)$ ist eine Gruppe für jede Menge M. Diese ist nicht kommutativ, sobald M mehr als zwei Elemente enthält. In Falle $M := \underline{n}$ für $n \in \mathbb{N}$ schreibt man S_n statt $S_{\underline{n}}$. Man nennt S_M die **symmetrische Gruppe** auf M.

(8) Sei $\mathrm{GL}_n(\mathbb{Q})$ die Menge der invertierbaren Matrizen in $\mathbb{Q}^{n \times n}$. (Erinnerung: $A \in \mathbb{Q}^{n \times n}$ heißt invertierbar, falls ein $B \in \mathbb{Q}^{n \times n}$ existiert, mit $AB = I_n$. Übung: Man hat dann automatisch $BA = I_n$.) Dann ist $\mathrm{GL}_n(\mathbb{Q})$ zusammen mit der Matrixmultiplikation eine Gruppe mit der Einheitsmatrix I_n als Einselement. Die Gruppe $\mathrm{GL}_n(\mathbb{Q})$ heißt die **generelle lineare Gruppe** über $\mathbb{Q}$.

(9) Ist $M \neq \emptyset$ eine Menge, so ist $\mathbb{Q}^M$ eine kommutative Gruppe mit der werteweisen Addition: Für $f, g \in \mathbb{Q}^M$ definiert man:

$$(f + g)(m) := f(m) + g(m) \text{ für alle } m \in M$$

Man beachte, $\mathbb{Q}^{\mathbb{Q}}$, $\mathbb{Q}^{n \times 1} := \mathbb{Q}^{\underline{n} \times \{1\}}$ und $\mathbb{Q}^{m \times n} := \mathbb{Q}^{\underline{m} \times \underline{n}}$ sind Spezialfälle hiervon. Die Nullabbildung ist immer das 1-Element oder besser 0-Element, wie man bei der additiven Sprechweise sagen sollte.

(10) (Addition von Matrizen) Sind $A, B \in \mathbb{Q}^{m\times n}$ Matrizen, so ist nach 9. ihre Verknüpfung definiert als die Matrix $A + B$ mit Einträgen $(A + B)_{i,j} = A_{i,j} + B_{i,j}$.

Für uns haben Gruppen zweierlei Bedeutung: Sie treten in den nachfolgenden Definitionen von Ringen und Körpern, also Zahlbereichen im weitesten Sinne, wieder auf als Teile der Definitionen. Zweitens haben Gruppen ein Eigenleben, welches fast alle Teile der Mathematik beeinflusst. Auf dieses kommen wir später zurück.

Unser bisheriger Umgang mit rationalen Matrizen legt die Frage nahe, ob man Ähnliches mit Matrizen über anderen Zahlbereichen durchführen kann. Welche Zahlbereiche sind dies? Welche entscheidenden Eigenschaften brauchen wir, um die Theorie auf diese anderen Fälle zu übertragen. Indem man solche Fragen stellt, nimmt man den sogenannten axiomatischen Standpunkt ein. Es stellt sich heraus, dass man statt mit dem Bereich $\mathbb{Q}$ der rationalen Zahlen einen beliebigen Körper, wie wir ihn jetzt definieren, hernehmen kann. Natürlich hat der Körper der rationalen Zahlen noch viele Eigenschaften, die nicht aus den hier aufgelisteten folgen. Aber diese Eigenschaften sind für die Theorie, wie wir sie in dieser Vorlesung entwickeln wollen, nicht relevant.

Definition 2.1.4. Sei K eine nicht leere Menge mit zwei (inneren) Verknüpfungen $+$ (genannt Addition) und $\cdot$ (genannt Multiplikation). Das Tripel $(K, +, \cdot)$ heißt ein **Körper**, falls

(1) $(K, +)$ ist eine Abelsche Gruppe mit neutralem Element 0;

(2) für $(K, \cdot)$ gilt das Assoziativgesetz, die Existenz des 1-Elementes $1 \neq 0$ und das Kommutativgesetz. Weiter ist $(K^*, \cdot)$ mit $K^* := K - \{0\}$ eine Abelsche Gruppe mit neutralem Element 1, insbesondere haben alle Elemente $\neq 0$ ein multiplikatives Inverses;

(3) es gelten die beiden **Distributiv**gesetze:

$$a(b + c) = ab + ac \qquad \text{für alle } a, b, c \in K.$$

und

$$(b + c)a = ba + ca \qquad \text{für alle } a, b, c \in K.$$

Für (2) gibt es diverse Abschwächungen, die zu allgemeineren Strukturen führen:

- Fordert man lediglich, dass $(K^*, \cdot)$ ein kommutatives Monoid ist und verzichtet man auf die Forderung[1] $1 \neq 0$, so heißt $(K, +, \cdot)$ ein **kommutativer Ring mit** 1.
- Verzichtet man noch zusätzlich auf die Kommutativität der Multiplikation, spricht man von einem Ring mit Eins.

Bemerkung 2.1.5. Sei K ein Körper. Dann gilt:

(1) $0a = 0$ für alle $a \in K$.

(2) Statt $a \cdot b^{-1}$ schreibt man auch $\frac{a}{b}$ für alle $a \in K, b \in K^*$. Man hat für $a, c \in K, b, d \in K^*$:

$$\frac{a}{b} + \frac{c}{d} = \frac{ad + cb}{bd}.$$

Beweis.

[1] Im Falle $1 = 0$ gibt es jedoch keine weiteren Elemente in dem Ring; wir sprechen dann vom 0-Ring.

(1) $0a = (0+0)a = 0a + 0a$, also durch Subtraktion von $0a$ erhält man $0a = 0$.

(2) Übung. □

Beispiel 2.1.6.

(1) $(\mathbb{R}, +, \cdot)$, kurz $\mathbb{R}$, ist ein Körper.

(2) $\mathbb{Q}$ ist ein Körper (Teilkörper von $\mathbb{R}$).

(3) $\mathbb{Z}$ ist kein Körper, sondern nur ein kommutativer Ring mit 1. Der Körper $\mathbb{Q}$ der rationalen Zahlen geht aus dem Ring $\mathbb{Z}$ durch Bereichsvergrößerung hervor.

(4) $\mathbb{F}_2 := \mathbb{Z}/2\mathbb{Z} = \{[0] := 2\mathbb{Z}, [1] := 1 + 2\mathbb{Z}\}$ ist ein Körper, mit folgenden Verknüpfungen:

$$(i + 2\mathbb{Z}) + (j + 2\mathbb{Z}) := (i + j) + 2\mathbb{Z}, \qquad (i + 2\mathbb{Z})(j + 2\mathbb{Z}) := ij + 2\mathbb{Z}$$

für alle $i, j \in \mathbb{Z}$. Man muss hier zeigen, dass die Addition und Multiplikation wohldefiniert ist, also vertreterunabhängig, weil in der Definition mit Vertretern der Äquivalenzklassen gearbeitet wird. Wir lassen dies als Übung. Man hat also eine surjektive Abbildung:

$$^- : \mathbb{Z} \to \mathbb{Z}/2\mathbb{Z} : i \mapsto i + 2\mathbb{Z},$$

welche das Rechnen überträgt. Man kann also von einer Bereichsvergröberung von $\mathbb{Z}$ sprechen. Man beachte $\overline{0}$ ist das Nullelement und $\overline{1}$ das Einselement von $\mathbb{Z}/2\mathbb{Z}$. Deshalb liegt es nahe, 0 statt $\overline{0}$ und 1 statt $\overline{1}$ zu schreiben, wenn klar ist, dass wir in $\mathbb{F}_2 = \mathbb{Z}/2\mathbb{Z}$ arbeiten.

Bemerkung 2.1.7.

(1) In einem kommutativen Ring R mit Eins gilt das **allgemeine Assoziativgesetz** für die Addition, d.h. bei Summen von mehr als zwei Summanden kann man die Klammern weglassen. Entsprechendes gilt für die Multiplikation. Es gilt das **allgemeine Kommutativgesetz für Addition und Multiplikation**, d.h. auf die Reihenfolge von Summanden bzw. Faktoren braucht man auch nicht zu achten. Es gilt **das allgemeine Distributivgesetz**, z.B.

$$\prod_{i=1}^{m} \left(\sum_{j=1}^{n} A_{ij} \right) = \sum_{\varphi \in \underline{n}^{\underline{m}}} \prod_{i=1}^{m} A_{i\varphi(i)}$$

für jedes $A \in R^{m \times n}$.

(2) Für $i \in \mathbb{Z}_{\geq 0}$ und $a \in R$ sei $ia := \underbrace{a + \cdots + a}_{i}$ und für $i \in Z_{<0}$ sei $ia := -((-i)a)$. Man hat einen Ringhomomorphismus

$$^- : \mathbb{Z} \to R : i \mapsto i \cdot 1,$$

d.h. eine Abbildung mit

$$\overline{(i + j)} = \overline{i} + \overline{j}, \qquad \overline{(i \cdot j)} = \overline{i} \cdot \overline{j}$$

für alle $i, j \in \mathbb{Z}$. Die Abbildung in Beispiel 2.1.6.(4) ist ein Spezialfall hiervon.

Wegen des ersten Teils der Bemerkung können wir alle Betrachtungen, die wir für rationale Matrizen angestellt haben, übertragen auf K-wertige Matrizen für einen beliebigen Körper K, und zwar inklusive GAUSSalgorithmus, weil man durch Körperelemente (ungleich Null) auch durchdividieren kann. Wir formulieren dies als ganz wichtige Übungsaufgabe, mit der Sie sich geraume Zeit befassen sollten.

Übung 2.1.3. Übertragen Sie alle Ergebnisse aus dem vorherigen Kapitel über rationale Matrizen und lineare Gleichungssysteme auf Matrizen und lineare Gleichungssysteme über beliebigen Körpern. Man überzeuge sich insbesondere beim GAUSSschen Algorithmus, dass er über einem beliebigen Körper funktioniert.

Beispiel 2.1.8. Aufgabe: Invertiere $A \in (\mathbb{Z}/2\mathbb{Z})^{3\times 3} = \mathbb{F}_2^{3\times 3}$ mit

$$A := \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

Lösung mit GAUSSschem Algorithmus:

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

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 \end{array}\right) \rightsquigarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 \end{array}\right).$$

Also ist

$$A^{-1} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}.$$

Übung 2.1.4. Sei

$$A := \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

Zeigen Sie: Es existiert ein $m \in \mathbb{N}$ mit $A^m = I_3$. Finde das kleinste derartige m.

Wir wollen noch eine Folgerung aus der letzten Bemerkung ziehen.

Ende Vorl. 6 26.10

Definition 2.1.9. Sei $n \in \mathbb{Z}_{\geq 0}$.

- Für $i \in \mathbb{Z}_{\geq 0}$ heißt
$$\binom{n}{i} := |\operatorname{Pot}_i(\underline{n})|$$
Binomialkoeffizient (gesprochen: n über i), wobei $\operatorname{Pot}_i(\underline{n})$ die Menge der i-elementigen Teilmengen von $\underline{n}$ bezeichnet.

- Für $i_1, \ldots, i_k \in \mathbb{Z}_{\geq 0}$ mit $i_1 + \cdots + i_k = n$ heißt
$$\binom{n}{i_1, \ldots, i_k} := |\{\varphi \in \underline{k}^{\underline{n}} \mid |\varphi^{-1}(\{j\})| = i_j \text{ für } j = 1, \ldots k\}|$$
Multinomialkoeffizient.

Es ist klar, dass $\binom{n}{i} = \binom{n}{i,n-i}$ gilt. Wegen des allgemeinen Kommutativ- und Distributivgesetze (Bemerkung 2.1.7) bekommen wir aus der Formel von 2.1.7 den Multinomialsatz:

Bemerkung 2.1.10 (Multinomialsatz)**.** Sei R ein kommutativer Ring und $a_1, \ldots, a_n \in R$. Für $m \in \mathbb{N}$ gilt:

$$(a_1 + \cdots + a_n)^m = \sum_{i_1+\cdots+i_n=m} \binom{m}{i_1, \ldots, i_n} a_1^{i_1} \cdots a_n^{i_n}.$$

Ein ganz wichtiger Körper, der in vielen Situationen eine große Rolle spielt, ist der komplexe Zahlkörper $\mathbb{C}$.

Satz 2.1.11. *Die Menge*

$$\mathbb{C} := \left\{ \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \mid a, b \in \mathbb{R} \right\}$$

ist zusammen mit der Matrixaddition und Matrixmultiplikation ein Körper, genannt der Körper der **komplexen Zahlen**. *Man schreibt abkürzend $a + bi$ für $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$. Durch die Identifizierung von $a \in \mathbb{R}$ mit $a = a + 0i \in \mathbb{C}$, wird $\mathbb{R}$ zu einem Teilkörper von $\mathbb{C}$, d.h. $\mathbb{R} \subseteq \mathbb{C}$ und die Addition und Multiplikation komplexer Zahlen angewandt auf reelle Zahlen ist die bekannte Addition und Multiplikation in $\mathbb{R}$.*

Beweis. Zuerst müssen wir zeigen, dass die Addition und Multiplikation wohldefiniert sind in dem Sinne, dass für zwei Elemente aus $\mathbb{C}$ die Summe und das Produkt auch wieder in $\mathbb{C}$ liegen (und nicht irgendwelche Matrizen außerhalb von $\mathbb{C} \subset \mathbb{R}^{2\times 2}$ sind). Bei der Summe ist das klar, bei dem Produkt folgt es aus der leicht verifizierten Formel

$$(a + bi)(c + di) := (ac - bd) + (ad + bc)i.$$

Die meisten Axiome folgen aus den Eigenschaften der Matrixmultiplikation. Wir lassen ihre Verifikation als Übung. Die Kommutativität der Multiplikation liest man ab aus der obigen Formel. Die zu $a + bi \neq 0$ multiplikativ inverse komplexe Zahl ist $\frac{1}{a^2+b^2}(a - bi)$. Dass $\mathbb{R}$ Teilkörper von $\mathbb{C}$ ist, ist ebenfalls klar aus der Multiplikationsformel. □

Übung 2.1.5. Zeige $(a + bi)(c + di) = e + fi$ genau dann, wenn

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix} \begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} e \\ f \end{pmatrix}.$$

In welcher Beziehung stehen also Linearität und Distributivgesetz?

Man kann $\mathbb{C}$ als Menge mit der EUKLIDischen Ebene identifizieren und spricht von der GAUSSschen Zahlenebene. Die Addition und Multiplikation kann dann geometrisch interpretiert werden. Wir kommen später hierauf zurück. Das Körperelement $i \in \mathbb{C}$ heißt die imaginäre Einheit. Der Name kommt wohl daher, dass es schwer ist, sich eine Zahl vorzustellen, deren Quadrat -1 ist. Diese Vorurteile wurden nachhaltig durch die Einführung der GAUSSschen Zahlenebene überwunden. Wir kommen später auf diverse Eigenschaften der komplexen Zahlen zurück. Folgende Übung ist eine gute Vorbereitung:

Übung 2.1.6. Für $a + bi \in \mathbb{C}$ mit $a, b \in \mathbb{R}$ heißt $\overline{a + bi} := a - bi$ die **konjugiert** komplexe Zahl. Zeige:

(1) $z + \overline{z}, z\overline{z} \in \mathbb{R}$ für alle $z \in \mathbb{C}$.

(2) $z\overline{z} \geq 0$ mit Gleichheit genau dann, wenn $z = 0$.

Definiere $|z| := \sqrt{z\overline{z}}$ als den **Betrag** oder Absolutbetrag z. Dann gilt:

(3) $|z_1 z_2| = |z_1||z_2|$ für alle $z_i \in \mathbb{C}$.

Es gibt nun diverse Konstruktionen, wie man aus kommutativen Ringen mit 1 Körper machen kann und aus Körpern kommutative Ringe mit 1. Wir wollen einige hiervon kennenlernen, hauptsächlich um noch einige neue Körper kennenzulernen, die wichtig sind. Schauen wir uns zuerst den Übergang von $\mathbb{Z}$ nach $\mathbb{Z}/2\mathbb{Z}$ an. Dies ist gleichzeitig ein wichtiges Beispiel für Äquivalenzrelationen und Partitionen. Um es zu verallgemeinern, erinnern wir an die Division mit Rest, die zu einer interessante Teilbarkeitstheorie bei $\mathbb{Z}$ führt, welche man bei Körpern natürlich nicht haben kann. Ausgangspunkt für die Teilbarkeitstheorie ist die folgende einfache Aussage:

Bemerkung 2.1.12. Seien $a, b \in \mathbb{Z}$ mit $b \neq 0$. Wir schreiben

$$|a| := \begin{cases} a, & a \geq 0, \\ -a, & a < 0, \end{cases}$$

für den Absolutbetrag von a. Dann gibt es eindeutige $q, r \in \mathbb{Z}$ mit

$$a = qb + r \text{ und } 0 \leq r < |b|$$

r heißt auch der kleinste nicht negative **Rest** von a modulo b, abgekürzt: $r = a \pmod{b}$. Falls $r = 0$ gilt, sagt man b **teilt** a oder b ist ein **Teiler** von a, kurz $b|a$.

Satz 2.1.13. *Sei $p \in \mathbb{Z}$ fest vorgegeben und*

$$\sim_p := \{(a, b) \in \mathbb{Z} \times \mathbb{Z} |\, p|a - b\} \subset \mathbb{Z} \times \mathbb{Z}.$$

(1) Dann ist $\sim_p$ eine Äquivalenzrelation auf $\mathbb{Z}$ und wir schreiben statt $(a, b) \in \sim_p$ kurz $a \sim_p b$ (oder wie in der Literatur üblich $a \equiv b \pmod{p}$). Die Äquivalenzklassen

$$[a] := \{b \in \mathbb{Z}, b \sim_p a\}$$

für $a \in \mathbb{Z}$ heißen **Restklasse***n von $\mathbb{Z}$ nach p und bilden eine Partition von $\mathbb{Z}$ in $|p|$ verschiede Klassen.*

(2) Die Addition und Multiplikation in $\mathbb{Z}$ ist verträglich mit $\sim_p$, d.h. für[2] $[a], [b] \in \mathbb{Z}/\sim_p$ gibt es eindeutige Restklassen $[c], [d] \in \mathbb{Z}/\sim_p$ mit

$$\alpha \in [a], \beta \in [b] \text{ impliziert } \alpha + \beta \in [c], \alpha\beta \in [d],$$

d.h. auf $\mathbb{Z}/\sim_p$ ist eine induzierte Addition und Multiplikation definiert:

$$[a] + [b] := [a + b], \quad [a] \cdot [b] := [ab]$$

für alle $a, b \in \mathbb{Z}$. Man schreibt üblicherweise $\mathbb{Z}/p\mathbb{Z}$ für $(\mathbb{Z}/\sim_p, +, \cdot)$ und $a + p\mathbb{Z}$ statt $[a]$ für $a \in \mathbb{Z}$. Es gilt: $\mathbb{Z}/p\mathbb{Z}$ ist ein kommutativer Ring mit Eins.

(3) Ist p eine Primzahl (also eine natürliche Zahl $p \neq \pm 1$ mit der Eigenschaft: $p \mid nm \implies p|n \vee p|m$), so ist $\mathbb{Z}/p\mathbb{Z}$ ein Körper, der **Restklassenkörper** *von $\mathbb{Z}$ modulo p.*
Andere Bezeichnung: $\mathbb{F}_p$.

Beweis.

(1) Wir prüfen die definierenden Eigenschaften für Äquivalenzrelationen nach. Seien also $a, b, c \in \mathbb{Z}$.

- Reflexivität: $a \sim_p a$, denn $p|0 = a - a$ für alle $a \in \mathbb{Z}$.
- Symmetrie: $a \sim_p b$ heißt $p|(a - b)$, also auch $p|(b - a)$, d.h. $b \sim_p a$.
- Transitivität: $a \sim_p b$ und $b \sim_p c$ heißt $p|a-b$ und $p|b-c$, also auch $p|(a-b)+(b-c) = a - c$, d.h. $a \sim_p c$.

Wir haben genau $|p|$ Äquivalenzklassen, denn $0, 1, \ldots, |p| - 1$ sind paarweise inäquivalent, liegen also in verschiedenen Klassen. Andererseits läßt jede Zahl $a \in \mathbb{Z}$ einen Rest zwischen 0 und $|p| - 1$, so dass es keine weiteren Klassen gibt.

[2] Wie üblich, bezeichnet $\mathbb{Z}/\sim_p$ die Menge der $\sim_p$-Äquivalenzklassen.

(2) Seien $\alpha, \alpha' \in [a], \beta, \beta' \in [b]$. Dann gibt es (sogar eindeutig bestimmte) $t, u \in \mathbb{Z}$ mit $\alpha' = \alpha + pt$ und $\beta' = \beta + pu$, also $\alpha' + \beta' = \alpha + \beta + p(t + u)$, d.h. die Addition der Äquivalenzklassen ist wohldefiniert, und $\alpha'\beta' = \alpha\beta + p(\alpha u + t\beta + ptu)$, d.h. die vertreterweise Multiplikation von Äquivalenzklassen ist auch wohldefiniert. Dass $\mathbb{Z}/p\mathbb{Z}$ ein kommutativer Ring mit 1 ist, folgt aus der Tatsache, dass $\mathbb{Z}$ ein kommutativer Ring mit Eins ist und sich alle Eigenschaften von $\mathbb{Z}$ sich übertragen, z.B. die Assoziativität der Addition etc.. Nullelement ist $0 + p\mathbb{Z}$, Einselement ist $1 + p\mathbb{Z}$. Einzelheiten: Übung. Der Trivialfall $p = 1$ liefert den Nullring, also der mit $1 = 0$.

(3) Um im Fall p Primzahl zu zeigen, dass $\mathbb{Z}/p\mathbb{Z}$ ein Körper ist (und nur in diesem Fall), brauchen wir nur noch zu zeigen: Zu $a + p\mathbb{Z} \neq p\mathbb{Z}$ existiert ein $b + p\mathbb{Z}$ mit $(a + p\mathbb{Z})(b + p\mathbb{Z}) = 1 + p\mathbb{Z}$ oder anders ausgedrückt: zu $a \notin p\mathbb{Z}$ existieren $b, n \in \mathbb{Z}$ mit $ba + np = 1$. Dies folgt aus dem EUKLIDischen Algorithmus, den wir gleich kennenlernen werden. □

Folgende Definition dient zur Erinnerung:

Definition 2.1.14. Für $a, b \in \mathbb{Z}$ ist der größte gemeinsame Teiler t von a und b (auch ggT von a und b genannt, geschrieben $\mathrm{ggT}(a, b)$) charakterisiert durch:

$$t|a \wedge t|b \quad \text{und aus } d|a \wedge d|b \implies d|t.$$

Insbesondere gilt $\mathrm{ggT}(a, 0) = a$.

Übung 2.1.7. Zeigen Sie, dass der größte gemeinsame Teiler zweier ganzer Zahlen bis auf Vorzeichen eindeutig bestimmt ist.
Hinweis: Benutzen Sie nur die Definition.

Algorithmus 2.1.15. (EUKLIDischer Algorithmus, 1. Teil)
Gegeben: $a, b \in \mathbb{Z}, a \neq 0 \neq b$.
Gesucht: $\mathrm{ggT}(a, b)$.
Algorithmus:

(1) Setze $a_1 := a, a_2 := b$.

(2) Für $n \geq 3$ setze $a_n := a_{n-2} \bmod a_{n-1}$, falls $a_{n-1} \neq 0$.

Nach endlich vielen Schritten hat man das erste $k \in \mathbb{N}$ mit $a_{k+1} = 0$. Dann ist a_k der größte gemeinsame Teiler von a und b, kurz $a_k = \mathrm{ggT}(a, b)$.

Beweis. Wir müssen erstens zeigen, dass der Algorithmus nach endlich vielen Schritten terminiert. Dies ist klar, denn $a_n \in \mathbb{Z}_{\geq 0}$ für $n \geq 2$ und die Folge ist ab dem dritten Glied streng monoton fallend.
Nächste Behauptung: $a_k = \mathrm{ggT}(a, b)$. Zu diesem Zweck mache man sich klar: Für jedes n mit $3 \leq n \leq k + 1$ haben a_{n-2}, a_{n-1} dieselben gemeinsamen Teiler wie a_{n-1}, a_n: Dies ist klar, da

$$a_{n-2} = q_{n-2}a_{n-1} + a_n \text{ mit } q_{n-2} \in \mathbb{Z}.$$

Schließlich ist a_k der größte gemeinsame Teiler von $a_k, 0$, so dass die Behauptung folgt. □

Ende Vorl. 7 02.11

Beispiel 2.1.16. Bestimme $\mathrm{ggT}(1002, 912)$:
Wir erhalten die Folge $1002, 912, 90, 12, 6, 0$, also $\mathrm{ggT}(1002, 912) = 6$.

Bemerkung 2.1.17. Sei

$$\varepsilon : \mathbb{Z}^{2\times 1} \to \mathbb{Z}^{2\times 1} : \begin{pmatrix} a \\ b \end{pmatrix} \mapsto \begin{cases} \begin{pmatrix} b \\ a \bmod b \end{pmatrix} & \text{falls } b \neq 0, \\ \begin{pmatrix} a \\ 0 \end{pmatrix} & \text{falls } b = 0. \end{cases}$$

(1) Der Euklidische Algorithmus kann auch so formuliert werden:
Iteriere die Anwendung von ε, solange der zweite Fall „$b = 0$“ nicht eintritt, d.h. bis zu einem k mit

$$\varepsilon^{k-1}(\begin{pmatrix} a \\ b \end{pmatrix}) = \begin{pmatrix} t \\ 0 \end{pmatrix}.$$

Dann ist $t = \mathrm{ggT}(a, b)$.

(2) Es gilt im ersten Fall „$b \neq 0$“:

$$\varepsilon(\begin{pmatrix} a \\ b \end{pmatrix}) = \begin{pmatrix} 0 & 1 \\ 1 & -q \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} b \\ r \end{pmatrix}$$

wobei

$$a = qb + r \text{ mit } q \in \mathbb{Z}, 0 \leq r < |b|.$$

Mit dieser recht offensichtlichen Bemerkung erhält man den zweiten Teil des Euklidischen Algorithmus, der den $\mathrm{ggT}(a, b)$ als ganzzahlige Linearkombination von a und b darstellt, die sogenannte Bézout Identität:

Algorithmus 2.1.18. (EUKLIDischer Algorithmus mit Bézout Identität)
Gegeben: $a, b \in \mathbb{Z}, a \neq 0 \neq b$.
Gesucht: $t = \mathrm{ggT}(a, b)$ und $\alpha, \beta \in \mathbb{Z}$ mit $\alpha a + \beta b = t$ (die Bézout Identität).
Algorithmus:

(1) Wie oben $a_1 := a$, $a_2 := b$.

(2) Setze $a_n = q_n a_{n+1} + a_{n+2}$ mit $q_n \in \mathbb{Z}$ für $n \geq 1$.

Definiere

$$A_1 := \begin{pmatrix} 0 & 1 \\ 1 & -q_1 \end{pmatrix}$$

und

$$A_n := \begin{pmatrix} 0 & 1 \\ 1 & -q_n \end{pmatrix} A_{n-1} \text{ für } n \geq 2.$$

Dann liefert die erste Zeile von A_{k-1} das gewünschte Paar (α, β) und es gilt

$$A_{k-1} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} t \\ 0 \end{pmatrix}.$$

Man beachte, dass der Algorithmus so formuliert ist, dass man nur wenige Zwischenergebnisse abspeichern muss. Außerdem ist die erste Zeile von A_{k-1} gleich der zweiten Zeile von A_{k-2}.

Beispiel 2.1.19. Bestimme $(25 + 31\mathbb{Z})^{-1}$ in $\mathbb{Z}/31\mathbb{Z} = \mathbb{F}_{31}$. Dazu bestimme die Bézout Identität vom $\mathrm{ggT}(31, 25) = 1$:

$$\begin{array}{rcl} 31 & = & 1 \cdot 25 + 6 \\ 25 & = & 4 \cdot 6 + 1 \quad \text{also} \\ 6 & = & 6 \cdot 1 + 0 \end{array}$$

$$\begin{pmatrix} 0 & 1 \\ 1 & -6 \end{pmatrix} (\begin{pmatrix} 0 & 1 \\ 1 & -4 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix}) = \begin{pmatrix} 0 & 1 \\ 1 & -6 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ -4 & 5 \end{pmatrix} = \begin{pmatrix} -4 & 5 \\ * & * \end{pmatrix}$$

also

$$-4 \cdot 31 + 5 \cdot 25 = 1, \text{ insbesondere } 5 \cdot 25 \equiv 1 \,(\mathrm{mod} 31) \text{ und somit } (25 + 31\mathbb{Z})^{-1} = 5 + 31\mathbb{Z}.$$

Übung 2.1.8. Seien $a, b \in \mathbb{Z}$ mit $b \neq 0$ und $A \in \mathbb{Z}^{2\times 2}$ mit ganzzahligem Inversen, also $A^{-1} \in \mathbb{Z}^{2\times 2}$. Zeigen Sie:

(1) Ist $A \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} d \\ 0 \end{pmatrix}$, so ist $d = \mathrm{ggT}(a, b)$.

(2) Für $(\alpha, \beta) \in \mathbb{Z}^{1\times 2}$ gilt genau dann $\alpha a + \beta b = d$, wenn ein $z \in \mathbb{Z}$ existiert mit $(\alpha, \beta) = A_{1,-} + zA_{2,-}$.

Definition 2.1.20. Sei R ein Ring mit 1. Ein Element $u \in R$ heißt Einheit, falls es ein $v \in R$ gibt mit $uv = 1 = vu$. Die Menge der Einheiten $R^\times := \{u \in R \mid u \text{ Einheit}\}$ nennt man die Einheitengruppe von R.

Beispiel 2.1.21.

(1) Ein kommutative Ring R mit Eins ist genau dann ein Körper, wenn $R^\times = R - \{0\}$.

(2) Sei K ein Körper. Die Einheitengruppe von $K^{n\times n}$ ist die generell lineare Gruppe $\mathrm{GL}_n(K)$.

Übung 2.1.9. Sei R ein Ring mit 1.

(1) Zeigen Sie, dass die Einheitengruppe $R^\times$ in der Tat eine Gruppe ist.

(2) Bestimmen Sie alle Elemente der Einheitengruppe des Ringes $\mathbb{F}_2^{2\times 2}$.

Definition 2.1.22. Sei R ein kommutativer Ring mit 1. Ein Element $r \in R - \{0\}$ heißt Nullteiler, falls es ein $s \in R - \{0\}$ existiert mit $rs = 0$.

Bemerkung 2.1.23. Jeder Körper ist nullteilerfrei.

Beweis. Sei K ein Körper $r \in K - \{0\}$ ein Nullteiler, sprich es existiert ein $s \in K - \{0\}$ mit $rs = 0$. Wende s^{-1} auf beide Seiten an und erhalte $r = 0$. Widerspruch. □

Übung 2.1.10. Sei R ein *endlicher* kommutativer Ring mit Eins. Zeigen Sie:

(1) Jedes Element $a \in R - \{0\}$ ist entweder eine Einheit oder ein Nullteiler.

(2) $(\mathbb{Z}/m\mathbb{Z})^\times = \{a + m\mathbb{Z} \mid \mathrm{ggT}(a, m) = 1\}$.

(3) Bestimmen Sie die Nullteiler von $\mathbb{Z}/24\mathbb{Z}$ und die Einheitengruppe $(\mathbb{Z}/24\mathbb{Z})^\times$.

Hinweis zu (1): Betrachte Potenzen von a und beachte, dass der Ring nur endlich viele Elemente hat.

2.2 Gruppenoperationen

Lernziel: Operationen von Gruppen als Äquivalenzprinzip, Bahnen, Lösungsmenge eines linearen Gleichungssystems als Bahn, GAUSSalgorithmus als Vertreterbestimmung einer Bahn.

Die Gruppenaxiome kommen nicht von ungefähr. Erstens sind sie zu sehen im Kontext von Abbildungen: Wir wissen bereits, dass die Komposition von Abbildungen einer Menge in sich dem Assoziativgesetz genügt und dass wir in der Identitätsabbildung ein 1-Element haben. Inverse bekommt man nur wenn man bijektive Abbildungen nimmt. Zweitens sind sie zu sehen in Parallele zu den Axiomen einer Äquivalenzrelation: Das Assoziativgesetz entspricht der Transitivität, die Existenz der Eins der Reflexivität und die Existenz eines

Inversen der Symmetrie. Hierauf werden wir bald zurückkommen. Zunächst einige Beispiele.

Unsere Frage ist: Was macht die Gruppe?

Definition 2.2.1. Sei G eine Gruppe und M eine Menge.

(1) G **operiert** auf M (von links), falls eine Abbildung

$$\omega : G \times M \to M : (g, m) \mapsto gm$$

gegeben ist mit folgenden zwei Eigenschaften:

Op1. $1m = m$ für alle $m \in M$, wobei 1 das 1-Element von G ist.

Op2. $g(hm) = (gh)m$ für alle $g, h \in G$ und alle $m \in M$.

ω heißt dann auch **Operation** von G auf M.

(2) G operiere auf M und es sei $m \in M$. Dann heißt

$$Gm := \{gm \mid g \in G\} \subseteq M$$

die **Bahn** von m unter G.

Übung 2.2.1. Die Gruppe G operiere auf der Menge M. Definiere für $g \in G$

$$\overline{g} : M \to M : m \mapsto gm$$

Zeigen Sie: $\overline{g}$ ist bijektiv (sprich $\overline{g} \in S_M$) und $\overline{gh} = \overline{g} \circ \overline{h}$ für alle $g, h \in G$.

Zuerst ein Beispiel, welches die Herkunft des Namens Bahn bis zu einem gewissen Grad erklärt.

Beispiel 2.2.2. Wir nehmen $(\mathbb{R}, +)$ als Gruppe und $M = \mathbb{R}^2$ als Menge mit Operation von $(\mathbb{R}, +)$

$$\mathbb{R} \times \mathbb{R}^2 \to \mathbb{R}^2 : (t, (x, y)) \mapsto (x + t, y - 2t)$$

Man rechnet nach, dass eine Operation vorliegt und identifiziert die Bahnen als Schar von parallelen Geraden.

Jetzt einige grundsätzlichere Beispiele.

Beispiel 2.2.3. Sei M eine Menge.

(1) Die symmetrische Gruppe $S_M = \{f : M \to M \mid f \text{ ist bijektiv }\}$ operiert auf M durch Anwenden. Die Bahn eines jeden Elementes von $m \in M$ ist ganz M.

(2) Sei $M = \mathbb{Z}/6\mathbb{Z}$ und $f : M \to M, x \mapsto x + 1$ die zyklische Permutation der Elemente von M. Dann ist $G := \{\mathrm{id}_M, f, f^2, f^3, f^4, f^5\}$ eine Untergruppe von S_M. Wie die volle symmetrische Gruppe, operiert auch G auf M durch Anwenden. Wieder besteht M aus genau einer Bahn. Wir können G jedoch auch auf den 2-elementigen Teilmengen von M operieren lassen. Es gilt $|\mathrm{Pot}_2(M)| = 15$ und G hat auf $\mathrm{Pot}_2(M)$ genau 3 Bahnen:
$G \cdot \{0, 1\} = \{\{i, i + 1\} \mid i \in M\}$, der Länge 6,
$G \cdot \{0, 2\} = \{\{i, i + 2\} \mid i \in M\}$, der Länge 6,
sowie $G \cdot \{0, 3\} = \{\{0, 3\}, \{1, 4\}, \{2, 5\}\}$, der Länge 3.
Übung: Wie sehen die G-Bahnen auf den 3-, 4- und 5- elementigen Teilmengen aus?
Hinweis: $G \cdot A = (G \cdot A^c)^c$ wobei $A^c = M - A$ das Komplement der Teilmenge A von M bezeichnet.

Sei K ein Körper.

(3) $\mathrm{GL}_n(K)$ operiert auf $K^{n\times 1}$ durch Linksmultiplikation.
Übung: Zeigen Sie dies und zeigen, dass die Operation genau zwei Bahnen hat.
Hinweis: Gauß-Algorithmus!

(4) $\mathrm{GL}_m(K)$ operiert auf $K^{m\times n}$ durch Linksmultiplikation.
Übung: Der GAUSSsche Algorithmus, der die *strikte* Stufenform herstellt, liefert uns für jede Bahn einen *eindeutigen* Vertreter.

Ende Vorl. 8 07.11

Der entscheidende Satz ist nun der folgende, den sicher diejenigen von Ihnen schon geahnt haben, die eine Parallelität zwischen den Axiomen für Äquivalenzrelationen und für Gruppen gesehen haben: Reflexivität und Existenz des 1-Elementes, Symmetrie und Existenz des Inversen, Transitivität und Existenz eines Produktes.

Satz 2.2.4. *Die Gruppe G operiere auf der Menge M. Definiere die Relation $\sim_G$ auf M durch*

$$m \sim_G m' \iff \exists g \in G : gm = m'.$$

Dann ist $\sim_G$ eine Äquivalenzrelation und die Äquivalenzklassen sind die Bahnen von M unter G:

$$\boxed{M/\sim_G = \{Gm \mid m \in M\}}$$

Beweis.

- Da $1 \in G$ ist $\sim_G$ reflexiv,
- da Inverse existieren, ist $\sim_G$ symmetrisch,
- und schließlich die Transitivität etwas ausführlicher: Sei $m \sim_G m'$ und $m' \sim_G m''$. Dann existieren $g, h \in G$ mit $m' = gm$ und $m'' = hm'$, also $m'' = h(gm) = (gh)m$, d.h. $m \sim_G m''$.

Somit ist $\sim_G$ eine Äquivalenzrelation, die zugehörigen Äquivalenzklassen sind die Bahnen. □

Definition 2.2.5. Sei G eine Gruppe, die auf der Menge M operiert.

- Eine Abbildung $f : M \to N$ in eine Menge N heißt **Invariante** der Operation, falls f konstant auf Bahnen ist, sprich falls $f(gm) = f(m)$ für alle $m \in M, g \in G$.
  Anders ausgedrückt: $\sim_f \supseteq \sim_G$.
- Die Invariante heißt **trennend**, falls verschiedenen Bahnen verschiedene Werte zugeordnet werden, sprich falls $f(m) = f(m')$ für $m, m' \in M$ bereits $Gm = Gm'$ impliziert.
  Anders ausgedrückt: $\sim_f = \sim_G$.

Beispiel 2.2.6. S_n operiert auf $\mathrm{Pot}(\underline{n})$ durch

$$S_n \times \mathrm{Pot}(\underline{n}) \to \mathrm{Pot}(\underline{n}) : (g, T) \mapsto gT := \{gt \mid t \in T\}$$

Zwei Teilmengen von $\underline{n}$ liegen offensichtlich genau dann in derselben Bahn, wenn sie gleich viele Elemente haben, d.h.

$$|\cdot| : \mathrm{Pot}(\underline{n}) \to \mathbb{Z} : T \mapsto |T|$$

ist eine trennende Invariante und die Bahnen sind durch die Teilmengen $\mathrm{Pot}_k(\underline{n})$ mit $k = 0, \ldots, n$ gegeben.

Beispiel 2.2.7. Bei der Operation

$$\mathbb{R} \times \mathbb{R}^2 \to \mathbb{R}^2 : (t, (x, y)) \mapsto (e^t x, e^{-t} y)$$

ist

$$\varphi : \mathbb{R}^2 \to \mathbb{R} : (x, y) \mapsto xy$$

eine nicht trennende Invariante, wie man am Urbild $\varphi^{-1}(\{0\})$ der Null sieht, welches mehr als eine Bahn enthält.

Übung 2.2.2. Zeigen Sie, dass

$$\mathrm{GL}_2(\mathbb{Z}) := \{A \in \mathbb{Z}^{2\times 2} \mid A \in \mathrm{GL}_2(\mathbb{Q}), A^{-1} \in \mathbb{Z}^{2\times 2}\}$$

eine Gruppe ist, die auf $\mathbb{Z}^{2\times 1}$ operiert, so dass

$$\gamma : \mathbb{Z}^{2\times 1} \to \mathbb{Z}_{\geq 0} : \begin{pmatrix} a \\ b \end{pmatrix} \to \mathrm{ggT}(a, b)$$

eine trennende Invariante ist.

Hat man eine Äquivalenzrelation, so ist es immer eine interessante Frage, ob diese Äquivalenzrelation in natürlicher[3] Weise von einer Gruppenoperation herstammt, d.h. ob die Äquivalenzklassen Bahnen einer Gruppe sind. Hier ist ein Beispiel aus dem Bereich der linearen Abbildungen und linearen Gleichungen.

Beispiel 2.2.8. Sei K ein Körper und $A \in K^{m\times n}$ und $\widetilde{A} : K^{n\times 1} \to K^{m\times 1}$ die induzierte lineare Abbildung.
Frage: Ist die Bildgleichheitsäquivalenzrelation $\sim_{\widetilde{A}}$ bezüglich $\widetilde{A}$, also die Äquivalenzrelation auf $K^{n\times 1}$ definiert durch $x \sim_{\widetilde{A}} y$ genau dann, wenn $Ax = Ay$, in natürlicher Weise induziert durch eine Gruppenoperation? Anders ausgedrückt, ist $\sim_{\widetilde{A}} = \sim_G$ für eine in diesem Kontext natürliche Gruppenoperation G.
Antwort: Ja. Die Gruppe ist $G := \widetilde{A}^{-1}(\{0\})$, die Faser von $\widetilde{A}$ über $0 \in K^{m\times 1}$, die wir auch später den Kern von $\widetilde{A}$ nennen werden (vgl. Bemerkung 1.3.3.(1)). Dies ist nämlich eine nicht leere Menge, da die Nullspalte aus $K^{n\times 1}$ dazugehört. Mit $x, y \in \widetilde{A}^{-1}(\{0\})$ ist auch $x + y \in \widetilde{A}^{-1}(\{0\})$, denn $\widetilde{A}(x + y) = \widetilde{A}(x) + \widetilde{A}(y) = 0 + 0 = 0$, und mit $x \in \widetilde{A}^{-1}(\{0\})$ ist auch $-x \in \widetilde{A}^{-1}(\{0\})$, d.h. $\widetilde{A}^{-1}(\{0\})$ ist eine sogenannte Untergruppe von $K^{n\times 1}$.
Die Gruppe $\widetilde{A}^{-1}(\{0\})$ operiert auf $K^{n\times 1}$ vermöge

$$\widetilde{A}^{-1}(\{0\}) \times K^{n\times 1} \to K^{n\times 1} : (a, x) \mapsto a + x$$

Die Abbildung $\widetilde{A}$ ist konstant auf den Bahnen dieser Operation ist, da $\widetilde{A}$ linear ist: $\widetilde{A}(a + x) = \widetilde{A}(x)$ für alle $a \in \widetilde{A}^{-1}(\{0\})$ und $x \in K^{n\times 1}$. Sei umgekehrt $\widetilde{A}(x) = \widetilde{A}(y)$, dann ist $x - y \in \widetilde{A}^{-1}(\{0\})$ und $x = (x - y) + y$, d.h. die Bahnen sind genau die Äquivalenzklassen und $\widetilde{A}$ ist eine trennende Invariante dieser Operation.

Wir halten fest, Invarianten sind nützlich, um festzustellen, ob zwei Elemente in derselben Bahn liegen. Es gibt aber noch eine andere Methode, die wir beim GAUSSschen Algorithmus schon gesehen haben: Normalformen. Normalformen bilden einfach ein ausgezeichnetes Vertretersystem der Bahnen. Hat man einen Algorithmus, der die Normalform herstellt, hat man dann auch ein Verfahren, um die Zugehörigkeit zu derselben Bahn zu testen.

[3] Künstlich kriegt man das immer hin! Übung.

2.3 Vektorräume

Lernziel: Formale Definition von Vektorräumen und lineare Abbildungen motiviert an einigen Beispielen, Rechnen in einigen konkreten Vektorräumen, direkte Summen von Vektorräumen, Teilräume, Faktorräume, Homomorphiesatz, Interpretation für lineare Gleichungssysteme.

Beispiel 2.3.1 (Beispiele für natürlich auftretende lineare Abbildungen)**.** Folgende Abbildungen $\varphi : D \to W$ sind linear in dem Sinne, dass $\varphi(aX) = a\varphi(X)$ und $\varphi(X + Y) = \varphi(X) + \varphi(Y)$ für alle $a \in K$ und $X, Y \in D$, wobei K ein Körper ist:

(1) Sei $A \in K^{m\times n}$ eine Matrix über dem Körper K. Sei $D = K^{n\times 1}, W = K^{m\times 1}$ und $\varphi := \widetilde{A} : D \to W : s \mapsto As$ die durch A induzierte lineare Abbildung.

(2) Sei $K = \mathbb{R}$, D die Menge der stetig differenzierbaren Funktionen auf $\mathbb{R}$ und W die Menge der stetigen Funktionen und φ die Ableitung:

$$\varphi : f \mapsto f' := \text{ 1. Ableitung von } f.$$

(3) Sei $D = W$ die Menge der stetigen Funktionen auf $\mathbb{R}$ und φ das Integral:

$$\varphi : f \mapsto If \text{ mit } If(x) := \int_0^x f(u)du.$$

(4) Sei $D = W = K^{\mathbb{N}}$, K ein beliebiger Körper und φ einer der beiden Schiebeoperatoren

$$\varphi : (a_1, a_2, a_3, \ldots) \mapsto (a_2, a_3, \ldots)$$

oder

$$\varphi : (a_1, a_2, a_3, \ldots) \mapsto (0, a_1, a_2, \ldots).$$

(5) Sei $D = W = K^{\mathbb{N}}$, K ein beliebiger Körper und φ der Differenzenoperator

$$\varphi : (a_1, a_2, a_3, \ldots) \mapsto (a_1, a_2 - a_1, a_3 - a_2, \ldots)$$

oder sein Inverser, der Summenoperator

$$\varphi : (a_1, a_2, a_3, \ldots) \mapsto (a_1, a_1 + a_2, a_1 + a_2 + a_3, \ldots)$$

(welcher im Falle $K = \mathbb{R}$ wiederum verwandt ist mit dem Arithmetisches-Mittel-Operator

$$\varphi : (a_1, a_2, a_3, \ldots) \mapsto (a_1, \frac{a_1 + a_2}{2}, \frac{a_1 + a_2 + a_3}{3}, \ldots).$$

(6) Sei D die Menge der konvergenten reellen Folgen, $K = \mathbb{R}$ und $W = \mathbb{R}$ und φ ist der Grenzwertoperator

$$\varphi : (a_1, a_2, a_3, \ldots) \mapsto \lim_{n\to\infty} a_n.$$

Vektorräume sind die Antwort auf die Frage nach den Definitions- und Wertebereichen linearer Abbildungen.

Definition 2.3.2. Sei K ein Körper. Eine Abelsche Gruppe $(\mathcal{V}, +)$ zusammen mit einer äußeren Verknüpfung

$$K \times \mathcal{V} \to \mathcal{V} : (a, X) \mapsto aX$$

heißt **Vektorraum** über K oder K**-Vektorraum**, falls gilt

$$\begin{array}{lrcll} 1) & a(X+Y) & = & aX+aY & \text{für alle } a \in K \text{ und } X,Y \in \mathcal{V}; \\ 2) & (a+b)X & = & aX+bX & \text{für alle } a,b \in K \text{ und } X \in \mathcal{V}; \\ 3) & (ab)X & = & a(bX) & \text{für alle } a,b \in K \text{ und } X \in \mathcal{V}; \\ 4) & 1X & = & X & \text{für alle } X \in \mathcal{V}. \end{array}$$

Die Elemente von $\mathcal{V}$ heißen **Vektoren**.

Bemerkung 2.3.3. Ist $\mathcal{V}$ ein K-Vektorraum, so gilt:

(1) $0X = 0$ für alle $X \in \mathcal{V}$, wobei die linke Null das Nullelement von K ist und die rechte das Nullelement von $\mathcal{V}$.

(2) $(-1)X = -X$ für alle $X \in \mathcal{V}$.

Beweis.

(1) $0X = (0+0)X = 0X + 0X$, also durch Subtraktion von $0X$ folgt die Behauptung.

(2) Übung. □

Beispiel 2.3.4. Sei K ein Körper und M eine Menge. Dann ist K^M ein Vektorraum über K mit werteweiser Addition:

$$f+g : M \to K : m \mapsto f(m)+g(m) \qquad \text{für alle } f,g \in K^M$$

und werteweiser Multiplikation

$$af : M \to K : m \mapsto af(m) \qquad \text{für alle } f \in K^M \text{ und } a \in K.$$

Die Verifikation ist eine ganz wichtige Übung.
Man beachte, $K^n, K^{1\times n}, K^{m\times n}, K^K, K^{\mathbb{N}}$ sind Spezialfälle dieses Beispiels. Beachte weiter, $K \equiv K^1$ ist ebenfalls Vektorraum über K.

Definition 2.3.5. Seien $\mathcal{V}$ und $\mathcal{W}$ Vektorräume über demselben Körper K. Eine Abbildung $\varphi : \mathcal{V} \to \mathcal{W}$ heißt K**-linear** (oder ein K-**Homomorphismus**), falls für alle $X,Y \in \mathcal{V}$ und alle $a \in K$ gilt

$$\varphi(aX+Y) = a\varphi(X)+\varphi(Y).$$

Ist φ noch zusätzlich injektiv, surjektiv bzw. bijektiv, so heißt φ ein **Monomorphismus**, **Epimorphismus** bzw. **Isomorphismus**; ist $\mathcal{V} = \mathcal{W}$, so heißt φ auch **Endomorphismus**. Bijektive Endomorphismen heißen **Automorphismen**. K-Vektorräume $\mathcal{V}, \mathcal{W}$ zwischen denen ein Isomorphismus existiert, heißen **isomorph**, in Zeichen: $\mathcal{V} \cong \mathcal{W}$.

Übung 2.3.1. Ist $\varphi : \mathcal{V} \to \mathcal{W}$ ein Isomorphismus. Zeigen Sie: $\varphi^{-1} : \mathcal{W} \to \mathcal{V}$ ist auch ein Isomorphismus.

Beispiel 2.3.6. Sei K ein Körper.

(1) Die Abbildung ${}^{tr} : K^{n\times m} \to K^{m\times n}, A \mapsto A^{tr}$ ist ein Isomorphismus von K-Vektorräumen. Was ist das Inverse?

(2) Jede Abbildung $\varphi : M \to N$ liefert einen K-Vektorraum Homomorphismus

$$\varphi^* : K^N \to K^M, f \mapsto f \circ \varphi.$$

Dieser ist ein Isomorphismus, falls φ bijektiv ist. Was ist dann das Inverse?

Ende Vorl. 9 09.11

(3) Ist $\mathcal{V}$ ein K-Vektorraum mit $V_1, \ldots, V_n \in \mathcal{V}$, also $V \in \mathcal{V}^n$, dann ist

$$\lambda_V : K^n \to \mathcal{V} : a \mapsto a_1V_1 + \cdots + a_nV_n$$

eine lineare Abbildung. Der Wert $\lambda_V(a)$ heißt auch **Linearkombination** der Vektoren V_i mit Koeffizienten a_i.

Übung 2.3.2. Zeigen Sie: Sind $\alpha : \mathcal{V} \to \mathcal{W}$, $\alpha' : \mathcal{V} \to \mathcal{W}$ und $\beta : \mathcal{W} \to \mathcal{T}$ lineare Abbildungen, so sind

- $a\alpha : \mathcal{V} \to \mathcal{W}$ für alle $a \in K$;
- $\alpha + \alpha' : \mathcal{V} \to \mathcal{W}$;
- $\beta \circ \alpha : \mathcal{V} \to \mathcal{T}$

auch lineare Abbildungen.

Bemerkung 2.3.7. Sei

$$\mathrm{End}(\mathcal{V}) := \{\alpha : \mathcal{V} \to \mathcal{V} \mid \alpha \text{ linear } \}.$$

Dann ist $\mathrm{End}(\mathcal{V})$ ein Ring mit werteweiser Addition und Komposition von Abbildungen als Multiplikation, genannt der **Endomorphismenring** des Vektorraums $\mathcal{V}$. Seine Einheitengruppe ist $\mathrm{GL}(\mathcal{V}) := \{\alpha : \mathcal{V} \to \mathcal{V} \mid \alpha \text{ linear und bijektiv } \}$, genannt die **generelle lineare Gruppe** über $\mathcal{V}$.

Beweis. $\mathrm{End}(\mathcal{V})$ ist abgeschlossen unter Addition und Komposition. Addition ist assoziativ und kommutativ, mit der Nullabbildung als neutrales Element und $-\varphi$ als additiv Inverses von φ. Komposition ist assoziativ, die Identitätsabbildung ist linear und neutrales Element der Komposition. Zu den Distributivgesetzen: Seien $\alpha, \beta, \gamma \in \mathrm{End}(\mathcal{V})$. Dann gilt für $X \in \mathcal{V}$:

$$((\alpha + \beta) \circ \gamma)(X) = (\alpha + \beta)(\gamma(X)) = \alpha(\gamma(X)) + \beta(\gamma(X)) = (\alpha \circ \gamma + \beta \circ \gamma)(X),$$

also $(\alpha + \beta) \circ \gamma = \alpha \circ \gamma + \beta \circ \gamma$ wegen der Definition der Addition und ebenfalls

$$(\alpha \circ (\beta + \gamma))(X) = \alpha(\beta(X) + \gamma(X)) = \alpha(\beta(X)) + \alpha(\gamma(X)),$$

also $\alpha \circ (\beta + \gamma) = \alpha \circ \beta + \alpha \circ \gamma$ wegen der Linearität von α. □

Definition 2.3.8. Ist $\mathcal{V}$ ein K-Vektorraum, so heißt eine Teilmenge $\mathcal{T} \subseteq \mathcal{V}$ **Teil(vektor)-raum** oder **Unter(vektor)raum** von $\mathcal{V}$, falls

(1) $\mathcal{T} \neq \emptyset$,

(2) Für $X, Y \in \mathcal{T}$ und $a \in K$ ist $aX + Y \in \mathcal{T}$.

Schreibweise: $\mathcal{T} \leq \mathcal{V}$.

Übung 2.3.3. Teilräume sind Vektorräume.

Beispiel 2.3.9. Bei der Übertragung von Nachrichten benutzt man sogenannte **lineare Codes**, welche einfach $\mathbb{F}_2$-Teilräume von $\mathbb{F}_2^n$ sind. Dabei heißt n dann die Länge des Codes. Die Idee ist, die Codes so zu konstruieren, dass sie geringfügige Fehler, wo also nur wenige Bits verändert sind erkennen und möglichst auch korrigieren. Der sogenannte Hamming-Code

$$\{aA + bB + cC \mid a, b, c \in \mathbb{F}_2\}$$

ist ein linearer Code der Länge 7 ist, welcher 8 verschiedene Folgen übertragen kann: Die 0-Folge und 7 Folgen mit vier Einsen und drei Nullen. Dabei sind A, B, C die Zeilen von

$$\begin{pmatrix} \mathbf{1} & \mathbf{0} & 1 & \mathbf{0} & 1 & 0 & 1 \\ \mathbf{0} & \mathbf{1} & 1 & \mathbf{0} & 0 & 1 & 1 \\ \mathbf{0} & \mathbf{0} & 0 & \mathbf{1} & 1 & 1 & 1 \end{pmatrix}.$$

Rechne nach: Jede 01-Folge (Bit-Folge) der Länge 7 mit 4 Einsen und 3 Nullen, die nicht zum Code gehört, muss sich an mindestens 2 Stellen (bzw. Bits) von den Elementen des Codes unterscheiden. Zwei verschiedene Folgen im Code unterscheiden sich in mindestens 3 Bits. Dadurch kann man bis zu zwei Bit-Flips erkennen und einen Bit-Flip sogar korrigieren.

Mit linearen Abbildungen sind gleich zwei Teilräume verbunden, einer im Definitionsbereich und einer im Wertebereich.

Satz 2.3.10. *Sei $\varphi : \mathcal{V} \to \mathcal{W}$ eine lineare Abbildung von K-Vektorräumen. Dann gilt:*

(1) $\mathrm{Kern}(\varphi) := \varphi^{-1}(\{0\}) \leq \mathcal{V}$. $\mathrm{Kern}(\varphi)$ *heißt der* **Kern** *von* φ.

(2) $\mathrm{Bild}(\varphi) \leq \mathcal{W}$.

(3) Ist $\mathcal{S} \leq \mathcal{W}$, so ist $\varphi^{-1}(\mathcal{S}) := \{X \in \mathcal{V} \mid \varphi(X) \in \mathcal{S}\}$ ein Teilraum von $\mathcal{V}$.

(4) Ist $\mathcal{T} \leq \mathcal{V}$, so ist $\varphi(\mathcal{T}) \leq \mathcal{W}$.

Beweis.

(1) folgt aus (3) da $\{0\} \leq \mathcal{W}$.

(2) folgt aus (4) mit $\mathcal{T} = \mathcal{V} \leq \mathcal{V}$.

(3)
- $0 \in \varphi^{-1}(\mathcal{S})$, da $\varphi(0) = 0 \in \mathcal{S}$. Also ist $\varphi^{-1}(\mathcal{S}) \neq \emptyset$.
- Sind $X, Y \in \varphi^{-1}(\mathcal{S})$ und $a \in K$, dann gilt
$$\varphi(aX + Y) = a\varphi(X) + \varphi(Y) \in \mathcal{S},$$
da $\varphi(X) \in \mathcal{S}$ und $\varphi(Y) \in \mathcal{S}$, also $aX + Y \in \varphi^{-1}(\mathcal{S})$.

(4)
- $0 = \varphi(0) \in \varphi(\mathcal{T})$. Also ist $\varphi(\mathcal{T}) \neq \emptyset$.
- Sind $X, Y \in \varphi(\mathcal{T})$ und $a \in K$, dann existieren $X', Y' \in \mathcal{T}$ mit $X = \varphi(X'), Y = \varphi(Y')$. Da $\mathcal{T} \leq \mathcal{V}$ ist auch $aX' + bY' \in \mathcal{T}$ und es gilt
$$aX + Y = a\varphi(X') + \varphi(Y') = \varphi(aX' + Y') \in \varphi(\mathcal{T}).$$ □

Übung 2.3.4. Sei $\varphi : \mathcal{V} \to \mathcal{W}$ linear. Zeigen Sie: φ ist genau dann injektiv, wenn $\mathrm{Kern}(\varphi) = \{0\}$ gilt.

Definition 2.3.11.

(1) Seien $\mathcal{V}$ und $\mathcal{W}$ zwei K-Vektorräume. Auf dem kartesischen Produkt $\mathcal{V} \times \mathcal{W}$ definieren wir durch
$$+ : (\mathcal{V} \times \mathcal{W}) \times (\mathcal{V} \times \mathcal{W}) \to (\mathcal{V} \times \mathcal{W}) : ((V, W), (V', W')) \mapsto (V + V', W + W')$$
eine (komponentenweise) Addition und durch
$$K \times (\mathcal{V} \times \mathcal{W}) \to (\mathcal{V} \times \mathcal{W}) : (a, (V, W)) \mapsto (aV, aW)$$
eine (komponentenweise) Multiplikation mit Körperelementen. Man verifiziert nun leicht als Übung, dass $(\mathcal{V} \times \mathcal{W}, +, .)$ wieder ein Vektorraum über K ist. Er wird mit $\mathcal{V} \oplus \mathcal{W}$ bezeichnet und heißt die **(äußere) direkte Summe** von $\mathcal{V}$ und $\mathcal{W}$. Wenn wir betonen wollen, dass wir die äußere direkte Summe meinen, schreiben wir auch $\mathcal{V} \oplus_a \mathcal{W}$.

(2) Sei $\mathcal{V}$ ein K-Vektorraum mit zwei Teilräumen $\mathcal{T}_1, \mathcal{T}_2 \leq \mathcal{V}$. Falls jedes $V \in \mathcal{V}$ in *eindeutiger* Weise als $V = T_1 + T_2$ mit $T_1 \in \mathcal{T}_1, T_2 \in \mathcal{T}_2$ geschrieben werden kann, heißt $\mathcal{V}$ die **(innere) direkte Summe** von $\mathcal{T}_1$ und $\mathcal{T}_2$. Man schreibt $\mathcal{V} = \mathcal{T}_1 \oplus \mathcal{T}_2$ oder $\mathcal{V} = \mathcal{T}_1 \oplus_i \mathcal{T}_2$, wenn man betonen will, dass die innere direkte Summe gemeint ist.

Beispiel 2.3.12. Ist M die disjunkte Vereinigung $M = N \dot\cup T$ und $U := \{f \in K^M \mid f_{|N} = 0\}$, $V := \{f \in K^M \mid f_{|T} = 0\}$, so ist

$$K^M = V \oplus_i U \cong K^N \oplus_a K^T.$$

Ende Vorl. 10 14.11

Beispiel 2.3.13. Sei $\mathcal{V} = \mathcal{T}_1 \oplus \mathcal{T}_2$ die (innere) direkte Summe der beiden Teilräume $\mathcal{T}_i \leq \mathcal{V}$.

(1) $\pi_i : \mathcal{T}_1 \oplus \mathcal{T}_2 \to \mathcal{T}_i : T_1 + T_2 \mapsto T_i$ ist eine surjektive lineare Abbildung (Epimorphismus), welche man **Projektion** (auf $\mathcal{T}_i$ bezüglich der Zerlegung $\mathcal{V} = \mathcal{T}_1 \oplus \mathcal{T}_2$) nennt. Es gilt

$$\mathrm{Kern}(\pi_1) = \mathcal{T}_2 \qquad \text{und} \qquad \mathrm{Kern}(\pi_2) = \mathcal{T}_1.$$

(2) $\iota_i : \mathcal{T}_i \to \mathcal{V} : T_i \mapsto T_i$ ist eine injektive lineare Abbildung (Monomorphismus), die Einbettung von $\mathcal{T}_i$ in $\mathcal{V}$. Es gilt: $\mathrm{Bild}(\iota_i) = \mathcal{T}_i$ für $i = 1, 2$. Man beachte:

$$\pi_i \circ \iota_i = \mathrm{id}_{\mathcal{T}_i} \qquad \text{für } i = 1, 2$$

Weiter: $\pi_2 \circ \iota_1 : \mathcal{T}_1 \to \mathcal{T}_2$ und $\pi_1 \circ \iota_2 : \mathcal{T}_2 \to \mathcal{T}_1$ die Nullabbildungen.

Wir wollen unsere neuen Einsichten auf lineare Gleichungssysteme anwenden. Vorher verallgemeinern wir Beispiel 2.2.8. Auch der Beweis ist eine offensichtliche Verallgemeinerung des Beweises vom Beispiel 2.2.8.

Satz 2.3.14. *Sei $\varphi : \mathcal{V} \to \mathcal{W}$ eine lineare Abbildung. Die nicht leeren Fasern von φ sind gleichzeitig die Bahnen der Operation von $\mathrm{Kern}(\varphi)$ per Addition auf $\mathcal{V}$:*

$$\mathrm{Kern}(\varphi) \times \mathcal{V} \to \mathcal{V},\ (X, V) \mapsto X + V.$$

Beweis. Das dies eine Operation definiert ist trivial.
Zuerst beweisen wir, dass jede Bahn in einer Faser enthalten ist: Für $\varphi(V) = W$ und $X \in \mathrm{Kern}(\varphi)$ ist $\varphi(X + V) = \varphi(X) + \varphi(V) = 0 + W = W$.
Schließlich beweisen wir, dass je zwei Elemente einer nicht-leeren Faser in einer Bahn liegen: Für bildgleiche V, V', d.h. $\varphi(V) = \varphi(V')$ existiert ein $X \in \mathrm{Kern}(\varphi)$ mit $V' = X + V$, nämlich $X = V' - V$. (Es ist keine andere Wahl für X möglich.) □

Folgerung 2.3.15. *Sei K ein Körper, $A \in K^{m\times n}$ eine Matrix und $b \in K^{m\times 1}$ eine Spalte. Das lineare Gleichungssystem*

$$Ax = b \tag{$*$}$$

ist genau dann lösbar, wenn $b \in \mathrm{Bild}(\widetilde{A})$. Es gilt:

(1) Die Lösungsmenge des sogenannten **zugehörigen homogenen Systems**, *also*

$$Ax = 0 \tag{$*_0$}$$

ist bekanntlich $\mathrm{Kern}(\widetilde{A})$. Sie ist daher ein Teilraum von $K^{n\times 1}$ und insbesondere nie leer.

(2) Die nicht leeren Fasern von $\widetilde{A}$ sind gleichzeitig die Bahnen von $\mathrm{Kern}(\widetilde{A})$ auf $K^{n\times 1}$ unter der Operation

$$\mathrm{Kern}(\widetilde{A}) \times K^{n\times 1} \to K^{n\times 1} : (X, Y) \mapsto X + Y.$$

Ist insbesondere $Y \in K^{n\times 1}$ eine Lösung von ($*$), so durchläuft $X + Y$ mit $X \in \mathrm{Kern}(\widetilde{A})$ (d.h. X Lösung von ($*_0$)) alle Lösungen von ($*$).

Beweis.

(1) Die Lösungsmenge von ($*_0$) ist gleich $\widetilde{A}^{-1}(\{0\}) =: \mathrm{Kern}(\widetilde{A})$. Die restliche Behauptung folgt aus 2.3.10.

(2) Dies haben wir bereits in Beispiel 2.2.8 bewiesen. Außerdem ist die Aussage nun ein Spezialfall vom Satz 2.3.14 für $\varphi := \widetilde{A}$. □

Definition 2.3.16. Sei $\mathcal{V}$ ein K-Vektorraum. Eine Äquivalenzrelation $\sim$ auf $\mathcal{V}$ heißt **verträglich** mit der Vektorraumstruktur oder einfach **linear** oder **Kongruenz**, falls aus $X \sim X'$ und $Y \sim Y'$ für $X, X', Y, Y' \in \mathcal{V}$ und $a \in K$ folgt $aX + Y \sim aX' + Y'$.

Beispiel 2.3.17.

(1) Ist $\varphi : \mathcal{V} \to \mathcal{W}$ linear und $\sim_\varphi$= "Bildgleichheit bez. φ", so ist $\sim_\varphi$ eine Kongruenz.

(2) Ist $\mathcal{U} \leq \mathcal{V}$ ein Teilraum von $\mathcal{V}$, so ist $\sim^{\mathcal{U}}$ eine Kongruenz definiert durch $X \sim^{\mathcal{U}} Y$ dann und nur dann, wenn $X - Y \in \mathcal{U}$. Wir nennen sie die **Kongruenz nach** $\mathcal{U}$

Beweis. Übung. □

Lemma 2.3.18. *Ist $\sim$ eine Kongruenz auf dem K-Vektorraum $\mathcal{V}$, so gilt:*

(1) Die **Kongruenzklasse** $[0]$ *des Nullelementes ist ein Teilraum $\mathcal{U}$ von $\mathcal{V}$; $[0] =: \mathcal{U} \leq \mathcal{V}$.*

(2) $\sim = \sim^{\mathcal{U}}$ aus dem Beispiel oben.

(3) $\sim^{\mathcal{U}} = \sim_{\mathcal{U}}$, d.h. die Kongruenz $\sim^{\mathcal{U}}$ stimmt mit der Äquivalenzrelation $\sim_{\mathcal{U}}$ überein[4], die durch die Operation von $\mathcal{U}$ auf $\mathcal{V}$ durch Addition

$$\mathcal{U} \times \mathcal{V} \to \mathcal{V} : (U, V) \mapsto U + V$$

induziert wird.

(4) Die Kongruenzklasse $[X]$ von $X \in \mathcal{V}$ ist gegeben durch

$$\mathcal{U} + X := \{U + X \mid U \in \mathcal{U}\} = \{X + U \mid U \in \mathcal{U}\} = X + \mathcal{U},$$

also durch die Bahn von X unter der obigen Operation von $\mathcal{U}$ auf $\mathcal{V}$ durch Addition.

Beweis.

(1) $[0] \neq \emptyset$, da $0 \in [0]$. Sind $X, Y \in [0]$ und $a \in K$, dann ist $X \sim 0$ und $Y \sim 0$, also wegen der Verträglichkeit $aX + Y \sim a0 + 0 = 0$, d.h. $aX + Y \in [0]$.

(2) Sei $\mathcal{U} := [0]$. Behauptung: Für $X, Y \in \mathcal{V}$ sind äquivalent: $X \sim Y$ und $X - Y \in \mathcal{U}$. Dies ist klar, da wegen der Verträglichkeit $X \sim Y$ äquivalent zu $X - Y \sim 0$ ist.

(3) $Z \sim^{\mathcal{U}} X \iff Z - X \in \mathcal{U} \iff \exists U \in \mathcal{U} : Z - X = U \iff \exists U \in \mathcal{U} : Z = U + X \iff Z \sim_{\mathcal{U}} X$.

(4) Folgt mit (3) aus Satz 2.2.4. □

Ende Vorl. 11 16.11

Folgerung 2.3.19. *Sei $\varphi : \mathcal{V} \to \mathcal{W}$ eine lineare Abbildung. Dann ist*

$$\sim_\varphi = \sim_{\mathrm{Kern}(\varphi)} = \sim^{\mathrm{Kern}(\varphi)},$$

d.h. die Bildgleichheitsäquivalenzrelation von φ stimmt sowohl mit der durch die Operation von $\mathrm{Kern}(\varphi)$ auf $\mathcal{V}$ induzierten Äquivalenzrelation als auch mit der Kongruenz nach $\mathrm{Kern}(\varphi)$ überein.

[4] Vgl. Satz 2.2.4

Beweis. Die erste Gleichheit ist eine Umformulierung von Satz 2.3.14 und die zweite die Aussage von 2.3.18.(3). □

Satz 2.3.20. *Sei $\mathcal{U} \leq \mathcal{V}$ ein Unterraum des K-Vektorraumes $\mathcal{V}$ und $\sim^{\mathcal{U}}$ die zugehörige Kongruenz. Die Menge $\mathcal{V}/\sim^{\mathcal{U}} = \mathcal{V}/\sim_{\mathcal{U}}$ der Kongruenzklassen (bzw. der Bahnen von $\mathcal{U}$) wird mit $\mathcal{V}/\mathcal{U}$ (lies $\mathcal{V}$ modulo $\mathcal{U}$ oder $\mathcal{V}$ nach $\mathcal{U}$) bezeichnet. Die Elemente von $\mathcal{V}/\mathcal{U}$ heißen ebenfalls* **Restklassen** *nach $\mathcal{U}$.*

(1) $\mathcal{V}/\mathcal{U}$ wird mit der wohldefinierten Addition

$$(X+\mathcal{U})+(Y+\mathcal{U}) := (X+Y)+\mathcal{U} \text{ für alle } X,Y \in \mathcal{V}$$

und Multiplikation

$$a(X+\mathcal{U}) := aX+\mathcal{U} \text{ für alle } X \in \mathcal{V}, a \in K$$

zu einem K-Vektorraum, genannt **Faktorraum**, **Quotientenraum** *oder* **Restklassenraum** *von $\mathcal{V}$ nach $\mathcal{U}$.*

(2) Die Abbildung

$$\nu : \mathcal{V} \to \mathcal{V}/\mathcal{U},\ X \mapsto X+\mathcal{U}$$

ist eine lineare Abbildung, genannt der **natürliche Epimorphismus** *von $\mathcal{V}$ auf $\mathcal{V}/\mathcal{U}$. Es gilt $\operatorname{Kern}(\nu) = \mathcal{U}$.*
Insbesondere ist jeder Teilraum eines Vektorraumes Kern eines geeigneten Homomorphismus.

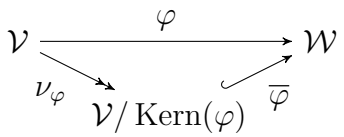
Beweis.

(1) Wir müssen zeigen, dass die Verknüpfungen wohldefiniert sind, d.h. **vertreterunabhängig** .
Wohldefiniertheit von $(X+\mathcal{U})+(Y+\mathcal{U})$:
Ist $X'+\mathcal{U} = X+\mathcal{U}$ und $Y'+\mathcal{U} = Y+\mathcal{U}$, so ist $(X'+Y')+\mathcal{U} = (X+Y)+\mathcal{U}$ zu zeigen. Per Definition existieren $U_1, U_2 \in \mathcal{U}$ mit $X' = X+U_1, Y' = Y+U_2$, also $(X'+Y')-(X+Y) = U_1+U_2 \in \mathcal{U}$, d.h. $(X'+Y')+\mathcal{U} = (X+Y)+\mathcal{U}$.
Wohldefiniertheit von $a(X+\mathcal{U})$:
Sei also $X'+\mathcal{U} = X+\mathcal{U}$, dann ist $X'-X \in \mathcal{U}$, also auch $aX'-aX = a(X'-X) \in \mathcal{U}$, also ist $aX+\mathcal{U} = aX'+\mathcal{U}$.
Jetzt müssen die Vektorraumaxiome überprüft werden, z.B. das Assoziativgesetz für $\mathcal{V}$ impliziert die Assoziativität der Addition von $\mathcal{V}/\mathcal{U}$ und $\mathcal{U} = 0+\mathcal{U}$ ist das Nullelement von $\mathcal{V}/\mathcal{U}$. Den Rest lassen wir als Übung.

(2) $\nu(aX+Y) = (aX+Y)+\mathcal{U} = a(X+\mathcal{U})+(Y+\mathcal{U}) = a\nu(X)+\nu(Y)$ für alle $a \in K$ und $X,Y \in \mathcal{V}$. Damit ist ν linear; dass ν surjektiv ist, ist klar und ebenso, dass $\operatorname{Kern}(\nu) = \mathcal{U}$. □

Jetzt ist alles für den Hauptsatz vorbereitet, den wir schon für Mengen und beliebige Abbildungen kennengelernt hatten.

Hauptsatz 2.3.21. (**Homomorphiesatz**) *Sei $\varphi : \mathcal{V} \to \mathcal{W}$ eine lineare Abbildung von K-Vektorräumen. Dann faktorisiert φ in die Komposition des natürlichen Epimorphismus $\nu = \nu_\varphi : \mathcal{V} \to \mathcal{V}/\operatorname{Kern}(\varphi)$ und des Monomorphismus $\overline{\varphi} : \mathcal{V}/\operatorname{Kern}(\varphi) \to \mathcal{W} : X+\operatorname{Kern}(\varphi) \mapsto \varphi(X)$, also $\varphi = \overline{\varphi} \circ \nu$. d.h. wir haben das* **kommutative Diagramm** *linearer Abbildungen*



Man beachte, dass der Monomorphismus $\overline{\varphi}$ einen Isomorphismus von $\mathcal{V}/\operatorname{Kern}(\varphi)$ auf $\operatorname{Bild}(\varphi)$ induziert, den wir ebenfalls mit $\overline{\varphi}$ bezeichnen werden. Insbesondere kennen wir bis auf Isomorphie alle epimorphen Bilder von $\mathcal{V}$, wenn wir alle Teilräume von $\mathcal{V}$ kennen:

$$\boxed{\varphi : \mathcal{V} \to \mathcal{W} \text{ linear } \implies \mathcal{V}/\operatorname{Kern}(\varphi) \cong \operatorname{Bild}(\varphi)}$$

Die exakte Analogie zum Homomorphiesatz für Mengen gilt wegen

$$\mathcal{V}/\operatorname{Kern}(\varphi) := \mathcal{V}/\sim_{\operatorname{Kern}(\varphi)} = \mathcal{V}/\sim_{\varphi} .$$

Die erste Gleichheit ist eine Definition und die zweite eine Umformulierung von Satz 2.3.14 bzw. die Aussage von Folgerung 2.3.19.

Beweis. $\nu := \nu_\varphi$ war bereits in 2.3.20 eingeführt, wobei wir als Teilraum $\mathcal{U} := \operatorname{Kern}(\varphi)$ wählen. Die beiden Bildgleichheitsäquivalenzrelationen $\sim_\varphi$ und $\sim_\nu$ sind gleich: Denn seien $X, Y \in \mathcal{V}$, dann gilt $\varphi(X) = \varphi(Y)$ genau dann, wenn $X - Y \in \operatorname{Kern}(\varphi) = \mathcal{U} = \operatorname{Kern}(\nu)$, was also äquivalent zu $\nu(X) = \nu(Y)$ ist.
Letzteres impliziert aber sowohl die Wohldefiniertheit von $\overline{\varphi}$ als auch die Injektivität von $\overline{\varphi}$. Die Linearität von ν hatten wir schon in 2.3.20 gesehen, die von $\overline{\varphi}$ folgt unmittelbar aus der von Linearität von φ. □

Beispiel 2.3.22. Was sagt uns der Homomorphiesatz über lineare Gleichungssysteme? Sei $A \in K^{m\times n}$ eine Matrix und $\varphi := \widetilde{A} : K^{n\times 1} \to K^{m\times 1}$, $x \mapsto Ax$ die induzierte lineare Abbildung. Dann sagt uns der Homomorphiesatz und die vorangegangenen Überlegungen:

(1) Diejenigen $b \in K^{m\times 1}$, für die $Ax = b$ lösbar ist, bilden einen Teilraum, nämlich $\operatorname{Bild}(\varphi)$ von $K^{m\times 1}$.

(2) Dieser Teilraum $\operatorname{Bild}(\varphi)$ ist isomorph zu $K^{n\times 1}/\operatorname{Kern}(\varphi)$.

(3) Die Lösungsmenge von $Ax = b$ für $b \in \operatorname{Bild}(\varphi)$ ist eine Restklasse nach $\operatorname{Kern}(\varphi)$ und wird unter dem Isomorphismus $\overline{\varphi} : K^{n\times 1}/\operatorname{Kern}(\varphi) \to \operatorname{Bild}(\varphi)$ auf b abgebildet. Insbesondere bildet die Gesamtheit aller (nicht leerer) Lösungsmengen (=Fasern) für variierendes $b \in \operatorname{Bild}(\varphi)$ einen Vektorraum, den Faktorraum nach $\operatorname{Kern}(\varphi)$.

(4) Je größer der Kern von φ ist, desto weniger rechte Seiten b gibt es, für die das Gleichungssystem lösbar ist. (Diese Tatsache werden wir später noch quantitativ untersuchen.)

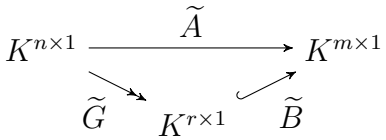
Beispiel 2.3.23. Was sagt uns der Homomorphiesatz über direkte Summen?
$\pi_1 : \mathcal{T}_1 \oplus \mathcal{T}_2 \to \mathcal{T}_1 : T_1 + T_2 \mapsto T_1$ ist eine surjektive lineare Abbildung (Epimorphismus) mit $\operatorname{Kern}(\pi_1) = \mathcal{T}_2$ und der Homomorphiesatz sagt

$$(\mathcal{T}_1 \oplus \mathcal{T}_2)/\mathcal{T}_2 \cong \mathcal{T}_1.$$

Insbesondere ist $\mathcal{T}_1$ eine **Transversale**, sprich ein Vertretersystem für die Restklassen von $\mathcal{T}_1 \oplus \mathcal{T}_2$ nach $\mathcal{T}_2$.

Satz 2.3.24. *Sei K ein Körper und $A \in K^{m\times n}$. Die induzierte lineare Abbildung $\widetilde{A} : K^{n\times 1} \to K^{m\times 1}$ kann in eine surjektive lineare Abbildung $\widetilde{G} : K^{n\times 1} \to K^{r\times 1}$ für ein $G \in K^{r\times n}$ mit $r \in \mathbb{N}$ und eine injektive lineare Abbildung $\widetilde{B} : K^{r\times 1} \to K^{m\times 1}$ mit $B \in K^{m\times r}$ faktorisiert werden:*

$$\widetilde{A} = \widetilde{B} \circ \widetilde{G} \qquad \textit{oder} \qquad A = BG:$$



Beweis. Wähle G als die Matrix, die aus der strikten Stufenform zu der zu A gehörigen Matrix durch Streichen der Nullzeilen hervorgeht. Die Zeilenzahl von G sei also r. Seien $s(i) := St_i(G)$ die Stufenindizes von G. Definiere $B \in K^{m\times r}$ dadurch, dass die j-te Spalte von B gleich der $s(j)$-ten Spalte von A ist: $B_{-,j} := A_{-,s(j)}$.
Es gilt $BG = A$, denn die Koeffizienten der i-ten Spalte von G sagen uns, mit welchen Koeffizienten die entsprechende Spalte $A_{-,i}$ aus den Stufenindexspalten $A_{-,s(i)}$ von A linearkombiniert werden, denn diese entsprechen den Stufenindexspalten von G, welche ja die Standardbasis von $K^{r\times 1}$ bilden.
Es gilt: $\widetilde{G}$ ist surjektiv. Dies ist klar, da die Standardbasisspalten von $K^{r\times 1}$ unter den Spalten von G vorkommen.
Es gilt: Bildgleichheit bezüglich $\widetilde{A}$ und $\widetilde{G}$ sind identisch, da nach Konstruktion des GAUSSschen Algorithmus gilt:

$$\mathrm{Kern}(\widetilde{A}) = \widetilde{A}^{-1}(\{0_m\}) = \widetilde{G}^{-1}(\{0_r\}) = \mathrm{Kern}(\widetilde{G}).$$

Es gilt: $\widetilde{B}$ ist injektiv. Dies folgt sofort aus dem letzten Schritt: Sei $X \in K^{r\times 1}$ mit $BX = 0_m$. Da $\widetilde{G}$ surjektiv ist, existiert ein $Y \in K^{n\times 1}$ mit $GY = X$. Man hat $AY = BGY = BX = 0_m$, also nach dem letzten Schritt $X = GY = 0_r$. □

Beispiel 2.3.25.

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 4 & 3 & 2 & 1 \\ 6 & 6 & 6 & 6 & 6 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 5 & 4 \\ 6 & 6 \end{pmatrix} \begin{pmatrix} 1 & 0 & -1 & -2 & -3 \\ 0 & 1 & 2 & 3 & 4 \end{pmatrix}.$$

Diese Darstellung ist gut, wenn man die Faser über der Nullspalte bestimmen will. Transponiert man die entsprechende Darstellung der transponierten Matrix, bekommt man eine Faktorisierung, die gut ist, um festzustellen, ob eine Spalte im Bild liegt:

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 4 & 3 & 2 & 1 \\ 6 & 6 & 6 & 6 & 6 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 4 & 3 & 2 & 1 \end{pmatrix}.$$

Ende Vorl. 12 21.11

2.4 Polynomringe

Lernziel: Formelle Einführung von Polynomen, Polynomdivision und EUKLIDischer Algorithmus, Körper der rationalen Funktionen, Restklassenkörper des Polynomrings, diverse Beispiele von Vektorräumen

Wir haben jetzt einerseits über kommutative Ringe und andererseits über Vektorräume gesprochen. Wir wollen nun über einen ganz wichtigen Ring sprechen, der gleichzeitig ein Vektorraum ist und der eine absolut grundlegende Rolle in der linearen Algebra spielt.

Erinnerung: $(\mathbb{Z}_{\geq 0}, +, 0)$ ist ein (abelsches) Monoid. Dies nutzen wir aus, um auf $K^{\mathbb{Z}_{\geq 0}}$ eine zweite Multiplikation zu definieren, die eine interessantere Ringstruktur liefert.

Definition 2.4.1. Sei K ein Körper.

(1) Auf dem K-Vektorraum $K^{\mathbb{Z}_{\geq 0}}$ definieren wir eine *kommutative* Multiplikation durch

$$(a_0, a_1, a_2, a_3, \ldots) \cdot (b_0, b_1, b_2, b_3, \ldots) := (c_0, c_1, c_2, c_3, \ldots)$$

mit

$$c_0 := a_0b_0, c_1 := a_0b_1 + a_1b_0, c_2 := a_0b_2 + a_1b_1 + a_2b_0, \ldots$$

allgemein für alle $n \geq 0$:

$$c_n := a_0b_n + a_1b_{n-1} + \ldots + a_nb_0$$

$(K^{\mathbb{Z}_{\geq 0}}, +, \cdot)$ zusammen mit der K-Vektorraumestruktur von $K^{\mathbb{Z}_{\geq 0}}$ wird auch mit $K[[x]]$ bezeichnet, dem **Potenzreihenring** über K in der **Unbestimmten** $x := (0, 1, 0, 0, \ldots)$, genauer, dem Ring der **formalen Potenzreihen** über K.
Statt $(a_0, a_1, a_2, a_3, \ldots)$ schreibt man auch $\sum_{i=0}^{\infty} a_i x^i$. Dadurch wird $(0, 1, 0, 0, \ldots)$ korrekterweise x zugeordnet.

(2) Eine Potenzreihe $a = (a_0, a_1, a_2, a_3, \ldots) \in K[[x]]$ heißt **Polynom**, falls ein $n \in \mathbb{Z}_{\geq 0}$ existiert mit $a_i = 0$ für alle $i > n$. Für $a \neq 0$ heißt das kleinste derartige n der **Grad** von a. Wir setzen $\mathrm{Grad}(0) := -\infty$. Die Menge aller Polynome bilden offensichtlich einen K-Teilraum von $K[[x]]$. Außerdem ist das Produkt von zwei Polynomen wieder ein Polynom.[5] Die Menge aller Polynome zusammen mit dieser Vektorraumestruktur und der von $K[[x]]$ ererbten Multiplikation heißt der **Polynomring** $K[x]$ über K.

Selbstverständlich kann man auch einen anderen Buchstaben als x für die Unbestimmte benutzen.

Bemerkung 2.4.2. Es gilt $x \cdot (a_0, a_1, a_2, \ldots) = (0, a_0, a_1, \ldots)$ insbesondere ist die Multiplikation mit x linear und injektiv.

Allgemeiner:

Übung 2.4.1. Sei $a \in K[[x]]$. Zeigen Sie (z.B. durch Induktion), dass

$$\mu_a : K[[x]] \to K[[x]] : b \mapsto ab$$

eine lineare Abbildung ist (insbesondere gelten in $K[[x]]$ die Distributivgesetze). Die Abbildung μ_a ist genau dann injektiv, wenn $a \neq 0$ gilt.

Die letzte Bemerkung 2.4.2 erlaubt uns mit Hilfe der in Übung 2.4.1 bewiesenen Distributivität das Produkt zweier Polynome leicht auszurechnen:

Beispiel 2.4.3 (Schriftliche Multiplikation ohne Übertrag)**.** In $\mathbb{Q}[x]$ berechnen wir ab mit $a := (1, 2, 0, 1, 0, 0, \ldots) = 1 + 2x + 1x^3$ und $b := (4, 3, 2, 1, 0, 0, \ldots) = 4 + 3x + 2x^2 + x^3$:

$$\begin{array}{ccccccccc}
4 & 3 & 2 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 8 & 6 & 4 & 2 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 4 & 3 & 2 & 1 & 0 & 0 \\
\hline
4 & 11 & 8 & 9 & 5 & 2 & 1 & 0 & 0
\end{array}$$

d.h. $ab = (4, 11, 8, 9, 5, 2, 1, 0, 0, \ldots) = 4 + 11x + 8x^2 + 9x^3 + 5x^4 + 2x^5 + x^6$.

Bemerkung 2.4.4.

(1) $1 := (1, 0, 0, \ldots) \in K[x]$ ist neutrales Element der Multiplikation in $K[[x]]$ und in $K[x]$.

[5] Für eine quantitative Präzisierung siehe die Grad-Formel in Bemerkung 2.4.4.(5).

(2) $xa = (0, a_0, a_1, \ldots)$ für alle $a \in K[[x]]$.

(3) Es gilt $x^i x^j = x^{i+j}$. Man setzt $x^0 := 1$. (Man sieht, wie die Multiplikation der Monome x^i der Addition der Exponenten entspricht. Man sagt: $K[x]$ ist die Halbgruppenalgebra von von $(\mathbb{Z}_{\geq 0}, +)$ über K.)

(4) Sei $a \in K[x]$ ein Polynom vom Grad n, dann gilt

$$a = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n.$$

Dies liefert eine offensichtliche Identifikation von K mit

$$K[x]_{\mathrm{Grad} \leq 0} \coloneqq \{a \in K[x] \mid \mathrm{Grad}(a) \leq 0\} = \{0\} \cup \{a \in K[x] \mid \mathrm{Grad}(a) = 0\}.$$

(5) $K[x]_{\mathrm{Grad} < n} \coloneqq \{a \in K[x] \mid \mathrm{Grad}(a) < n\}$ ist ein Teilvektorraum von $K[x]$.

(6) Für $a, b \in K[x]$ gilt die **Grad-Formel**

$$\mathrm{Grad}(ab) = \mathrm{Grad}(a) + \mathrm{Grad}(b).$$

Jede einzelne Aussage dieser Bemerkung ist sehr leicht zu beweisen und gleichzeitig sehr wichtig.

Definition 2.4.5. Sei R ein Ring mit Eins, der gleichzeitig ein K-Vektorraum für den Körper K ist. Man nennt dann R eine **assoziative K-Algebra** mit Eins oder kürzer **K-Algebra** , falls gilt:

$$k(ab) = (ka)b = a(kb)$$

für alle $a, b \in R$ und $k \in K$. Ist S eine weitere K-Algebra mit Eins, so heißt eine K-lineare Abbildung $\varphi : R \to S$ ein K-Algebrenhomomorphismus, falls

- $\varphi(ab) = \varphi(a)\varphi(b)$ für alle $a, b \in R$ ist und
- $\varphi(1_R) = 1_S$ gilt.

Satz 2.4.6. *Sei K ein Körper.*

(1) $K[x]$ ist ein kommutativer Ring, sogar eine kommutative K-Algebra.

(2) Polynomdivision als Division mit Rest: Für $a, b \in K[x]$ mit $b \neq 0$ existieren eindeutige $q, r \in K[x]$ mit

$$a = qb + r, \quad \mathrm{Grad}(r) < \mathrm{Grad}(b).$$

Beweis.

(1) Dass die Multiplikation kommutativ ist und wir ein Einselement haben, ergibt sich direkt aus der Definition der Multiplikation. Das Distributivgesetze folgen direkt aus der (Bi)linearität (und der Kommutativität) der Multiplikation, wie wir bereits in Übung 2.4.1 bemerkt haben. Es bleibt die Assoziativität der Multiplikation zu zeigen. Behauptung: $a(bc) = (ab)c$ für alle $a, b, c \in K[x]$.
Beweis durch Induktion nach $\mathrm{Grad}(c)$: Die Behauptung ist sicher richtig, wenn $\mathrm{Grad}(c) = 0$. Angenommen sie gilt für alle a, b, c mit $\mathrm{Grad}(c) \leq n$. Wir zeigen, dass sie dann

auch für $\mathrm{Grad}(c) = n+1$ gilt. Zu diesem Zweck schreiben wir $c = C + c_{n+1}x^{n+1}$ mit $\mathrm{Grad}(C) \leq n$ (schließt den Fall $C = 0$ ein). Dann gilt:

$$\begin{aligned} a(bc) &= a(b(C + c_{n+1}x^{n+1})) \\ &= a(bC + c_{n+1}bx^{n+1}) \\ &= a(bC) + c_{n+1}a(bx^{n+1}) \\ &= (ab)C + c_{n+1}(ab)x^{n+1} \\ &= (ab)(C + c_{n+1}x^{n+1}) \\ &= (ab)c, \end{aligned}$$

wobei im entscheidenden vierten Schritt einerseits die Induktionsvoraussetzung und andererseits eine einfache Beobachtung über Verschieben eingeht.

(2) Sie kennen das Verfahren von der schriftlichen Division her, nur dass hier die Situation einfacher ist, als bei ganzen Zahlen, da man keine Überträge hat. Sei $\mathrm{Grad}(a) = m$ und $\mathrm{Grad}(b) = n$. Falls $m < n$, sind wir bereits fertig mit $q = 0, r = a$. Falls $m \geq n$ ist, ersetzen wir a durch $a - \frac{a_m}{b_n}x^{m-n}b$ und belassen b. Da $\mathrm{Grad}\left(a - \frac{a_m}{b_n}x^{m-n}b\right) < \mathrm{Grad}(a)$ ist, ist nach endlich vielen Schritten der Grad des ersten Polynoms schließlich kleiner als $n = \mathrm{Grad}(b)$ und wir können $q = \frac{a_m}{b_n}x^{m-n} + \ldots$ sowie r als das letzte der Polynome aus der Folge der a's ablesen. Soweit die Existenz von q und r.
Zur Eindeutigkeit: Sei $a = q'b + r'$ mit $\mathrm{Grad}(r') < \mathrm{Grad}(b)$. Dann folgt

$$r - r' = (q' - q)b$$

Wäre $q' - q \neq 0$, dann folgt $\mathrm{Grad}(r - r') \geq \mathrm{Grad}(b)$, was ein Widerspruch ist. Also ist $q = q'$ und $r = r'$. □

Beispiel 2.4.7. $a := x^6 - x - 1, b = x^2 - x + 1 \in \mathbb{Q}[x]$. Wir suchen den Quotienten q und den Rest r. Wir haben also

Ende Vorl. 13 23.11

$$q = -1 - x + x^3 + x^4, r = -x.$$

Man vergleiche dieses Schema mit dem von der schriftlichen Division.

Übung 2.4.2. Zeige als Folgerung des letzten Satzes, dass der Potenzreihenring $K[[x]]$ ein kommutativer Ring mit 1, sogar kommutative K-Algebra ist. Man beachte, dass bei der üblichen Schreibweise für $a \in K[[x]]$ als Potenzreihe

$$a = \sum_{i=0}^{\infty} a_i x^i$$

es sich um eine formale Schreibweise handelt. Im algebraischen Sinne sind Summen mit unendlich vielen Summanden nicht definiert.

Bemerkung 2.4.8 (In der Vorlesung übersprungen)**.** Bei der Bestimmung von q und r handelt es sich bei der Polynomdivision um das Lösen eines linearen Gleichungssystems, welches bereits in Dreiecksgestalt gegeben ist.

0	0	0	0	1	−1	1	0	0	…	$1x^4 \cdot b$
0	0	0	1	−1	1	0	0	0	…	$1x^3 \cdot b$
0	0	0	0	0	0	0	0	0	…	$0x^2 \cdot b$
0	−1	1	−1	0	0	0	0	0	…	$-1x \cdot b$
−1	1	−1	0	0	0	0	0	0	…	$-1 \cdot b$
−1	−1	0	0	0	0	1	0	0	…	a

An den Spalten 3 ($= 1+\mathrm{Grad}(b)$) bis 7 ($= 1+\mathrm{Grad}(a)$) sieht man, welches Gleichungssystem man lösen muss, aus den ersten 2 ($= \mathrm{Grad}(b)$) kann man den Rest bestimmen und die Lösung des Gleichungssystems, also q kann man aus der letzten Spalte ablesen.

Folgerung 2.4.9. *Sei $0 \neq p \in K[x]$ von Grad n. Dann bilden die Vielfachen von p einen Teilraum $pK[x] \leq K[x]$ und*

$$K[x] = K[x]_{\mathrm{Grad}<n} \oplus pK[x].$$

Insbesondere hat der Faktorraum $K[x]/pK[x]$ den K-Teilraum $K[x]_{\mathrm{Grad}<n}$ als Vertretersystem.

Beweis. Jedes $f \in K[x]$ lässt sich mit Hilfe der Polynomdivision mit Rest eindeutig schreiben als $f = qp+r$ mit $r, q \in K[x]$, $\mathrm{Grad}(r) < n = \mathrm{Grad}(p)$. Also ist $K[x] = K[x]_{\mathrm{Grad}<n} \oplus pK[x]$. Nach dem Homomorphiesatz 2.3.21 bzw. Beispiel 2.3.23 ist somit $K[x]/pK[x] \cong K[x]_{\mathrm{Grad}<n}$, was man ganz konkret so verstehen kann, dass jede Restklasse $f+pK[x]$ einen **kanonischen Vertreter** $r = f - pq \in f + pK[x]$ hat, mit $\mathrm{Grad}(r) < n$. □

Bemerkung 2.4.10. Sei $p = a_0 + a_1x + \ldots + a_{n-1}x^{n-1} + x^n \in K[x]$. Die Multiplikation mit x induziert einen Endomorphismus von $K[x]/pK[x]$

$$x^i + pK[x] \mapsto x^{i+1} + pK[x] \text{ für } i = 0, \ldots, n-2,$$

und

$$x^{n-1} + pK[x] \mapsto x^n + pK[x] = -a_0 - a_1x - \ldots - a_{n-1}x^{n-1} + pK[x].$$

Beweis. Dass die Multiplikation mit x auf $K[x]$ linear ist, wissen wir schon. Wir müssen zeigen, dass sie eine wohldefinierte Selbstabbildung von $K[x]/pK[x]$ induziert, die dann natürlich automatisch linear ist. Sei also $r+pK[x] = s+pK[x]$ für zwei Elemente $r, s \in K[x]$. Behauptung: $xr+pK[x] = xs+pK[x]$. Beweis: $r-s \in pK[x]$, also $xr-xs = x(r-s) \in pK[x]$, also $xr + pK[x] = xs + pK[x]$. Der Rest ist klar. □

Schreibweise. Für die Kongruenz $a \sim^{pK[x]} b$ nach $pK[x]$ (d.h. $a + pK[x] = b + pK[x]$) schreiben wir oft $a \equiv b \pmod{p}$.

Übung 2.4.3. Zeigen Sie: $x^{100} \equiv x \pmod{x^2 + x + 1}$.
Hinweis: Rechne erst modulo $x^3 - 1$.

Übung 2.4.4. Zeigen Sie: Elemente $a = \sum_{i=0}^{\infty} a_i x^i \in K[[x]]$ mit $a_0 \neq 0$ sind invertierbar. Modifiziere das Schema des letzten Beispiels um ein Schema anzugeben, wie man die ersten n Glieder von a^{-1} ausrechnen kann. Gibt es Alternativen? Betrachte $(1-x)^{-1}$. Betrachte auch $(1 - x - x^2)^{-1} = \sum_{i=0}^{\infty} a_i x^i$. Zeige: $a_0 = a_1 = 1$ und $a_{i+2} = a_i + a_{i+1}$ für $i \geq 0$.

Inzwischen sollte die Parallelität vieler Eigenschaften von $\mathbb{Z}$ und $K[x]$ klar sein:

(1) In $\mathbb{Z}$ gilt $|ab| = |a||b|$, in $K[x]$ gilt $\mathrm{Grad}(ab) = \mathrm{Grad}(a) + \mathrm{Grad}(b)$, insbesondere hat $K[x]$ auch einen Quotientenkörper, den man mit $K(x)$ bezeichnet, der **Körper der rationalen Funktionen über** K (siehe Übung 2.4.5).

(2) In beiden Ringen haben wir Division mit Rest und damit hat $K[x]$ auch einen erweiterten EUKLIDischen Algorithmus (samt Bézout Identität), größte gemeinsame Teiler sind definiert und eindeutig bis Elemente in K^*, sprich bis auf Faktoren vom Grad 0. Weiter hat man auch das Analogon von Primzahlen in $K[x]$: **irreduzible** Polynome, also solche, die nicht als Produkt von zwei Polynomen echt kleineren Grades geschrieben werden können.

(3) In Analogie zu den Restklassenkörpern $\mathbb{F}_p := \mathbb{Z}/p\mathbb{Z}$ für Primzahlen p, hat man Restklassenkörper $K[x]/pK[x]$ für irreduzible Polynome $p = p(x) \in K[x]$ (siehe Satz 2.4.11).

Übung 2.4.5. Ausgehend von einem Körper K konstruiere aus dem Polynomring $K[x]$ einen Körper $K(x)$ in Analogie zu der Konstruktion von $\mathbb{Q}$ aus $\mathbb{Z}$.
Hinweis: $K(x) := (K[x] \times (K[x] - \{0\}))/\sim$ mit $(p, q) \sim (r, t)$ genau dann, wenn $pt = qr$. Man nennt $K(x)$ den Körper der rationalen Funktionen über K.

Satz 2.4.11.

(1) Sei $p \in K[x]$. Dann ist $K[x]/pK[x]$ durch vertreterweise Addition und Multiplikation ein Ring mit $\overline{1} := 1 + pK[x]$.

(2) Ist $p \in K[x]$ **irreduzibel**, *d.h. $\mathrm{Grad}(p) > 0$ und p hat keine Teiler in $K[x]$ von Grad g mit $0 < g < \mathrm{Grad}(p)$, so ist $K[x]/pK[x]$ ein Körper.*

Beweis.

(1) $K[x]$ ist assoziative, kommutative K-Algebra mit Eins, so dass sich die meisten Gesetze auf die Restklassenalgebra $K[x]/pK[x]$ sofort vererben. Es bleibt die Wohldefiniertheit der Multiplikation zu zeigen. Sind $a, a', b, b' \in K[x]$ mit $a = a' + pu, b = b' + pv$, so gilt
$$ab = (a' + pu)(b' + pv) = a'b' + p(ub' + a'v + puv) \in a'b' + pK[x].$$

(2) Ist nun $p \in K[x]$ irreduzibel, so ist zu zeigen, dass jedes Element $\neq 0$ in $K[x]/pK[x]$ invertierbar ist: Sei also $a \in K[x]$ mit $\overline{a} \neq 0$ in $K[x]/pK[x]$. Da p irreduzibel ist, liefert der erweiterte Euklidische Algorithmus $\alpha, \beta \in K[x]$ mit $\alpha a + \beta p = 1$. Es folgt $\overline{a}^{-1} = \overline{\alpha}$. □

Beispiel 2.4.12.

(1) Neue Konstruktion von $\mathbb{C}$: In $\mathbb{R}[x]$ ist $p = x^2 + 1$ irreduzibel. Bezeichne die Restklasse von x mit $\overline{x}$. Dann gilt somit $\overline{x}^2 = -1$. Die Element von $\mathbb{R}[x]/p\mathbb{R}[x]$ sind gegeben durch $a + b\overline{x}$ mit $a, b \in \mathbb{R}$. Es gilt
$$(a + b\overline{x})(c + d\overline{x}) = ac - bd + (ad + bc)\overline{x},$$
d.h. wir haben den komplexen Zahlkörper neu konstruiert.
Übung: Benutze den EUKLIDischen Algorithmus um $a + b\overline{x}$ zu invertieren.

(2) Körper mit vier Elementen: In $\mathbb{F}_2[x]$ ist $p = x^2 + x + 1$ irreduzibel. Damit ist $\mathbb{F}_2[x]/p\mathbb{F}_2[x]$ ein Körper mit vier Elementen: $0, 1, \overline{x}, 1 + \overline{x}$.
Übung: Man gebe die Additions- und Multiplikationstabellen an.

(3) Ein algebraisches Modell des Körpers $\mathbb{Q}[\sqrt[3]{2}]$: Das Polynom $x^3 - 2 \in \mathbb{Q}[x]$ ist sicher irreduzibel, da es sonst einen Teiler der Form $x - a$ mit $a \in \mathbb{Q}$ hätte. Indem man eine Primfaktorzerlegung für Zähler und Nenner ansetzt kommt man wegen $a^3 = 2$ schnell zu einem Widerspruch.
Also ist $\mathbb{Q}[x]/(x^3 - 2)\mathbb{Q}[x]$ ein Körper. Setze $(\sqrt[3]{2} :=)\overline{x} := x + (x^3 - 2)\mathbb{Q}[x]$.
Aufgabe: Bestimme $(\overline{x}^2 + \overline{x} + 1)^{-1}$, sprich ihre Normalform.
Lösung: Nach einem einschrittigen EUKLIDischen Algorithmus erhalten wir
$$x^3 - 2 = (x - 1)(x^2 + x + 1) - 1,$$
und somit
$$(\overline{x}^2 + \overline{x} + 1)^{-1} = \overline{x} - 1,$$
oder
$$\frac{1}{\sqrt[3]{2}^2 + \sqrt[3]{2} + 1} = -1 + \sqrt[3]{2}.$$
Mit dem Begriff der Irreduzibilität von Polynomen ist der der Wurzel eng verbunden.

Bemerkung 2.4.13. Sei K ein Körper und A eine assoziative K-Algebra (also z.B. $A = K$ oder $A = K^{n\times n}$).

(1) Für jedes $a \in A$ ist durch $x^i \mapsto a^i$ eine lineare Abbildung $\varepsilon_a : K[x] \to A : p \mapsto p(a)$ definiert (genannt der **Einsetzungshomomorphismus**), sogar ein K-Algebrenhomomorphismus.

(2) Ein Element $a \in K$ heißt **Wurzel** des Polynoms p, falls $p(a) = 0$, also $\varepsilon_a(p) = 0$.

(3) Für $a \in K$ ist $p(a)$ der Rest der Division von p durch $(x - a)$:

$$p \equiv p(a) \mod x - a$$

Ende Vorl. 14 28.11

(4) Für $a \in K$ ist $\mathrm{Kern}(\varepsilon_a) = (x - a)K[x]$.

(5) Ein Polynom vom Grad n hat höchstens n verschiedene Wurzeln in K.

(6) Eine Abbildung $f \in K^K$ heißt **Polynomfunktion**, falls ein $p \in K[x]$ existiert mit $f(a) = p(a)$ für alle $a \in K$. In diesem Fall heißt $f =: f_p$ die von p induzierte Polynomfunktion. Es ist

$$\varepsilon : K[x] \to K^K : p \mapsto f_p$$

ein K-Algebrenhomomorphismus. Das Bild bezeichnen wir mit $\mathrm{PolFu}(K)$.

- Die Abbildung ε ist genau dann injektiv, wenn K unendlich ist. Dann ist ihre Korestriktion $K[x] \to \mathrm{PolFu}(K)$ auf ihr Bild ein Isomorphismus.
- Die Abbildung ist genau dann surjektiv, wenn K endlich ist. In dem Fall ist sie aus Mächtigkeitsgründen nicht injektiv.

Beweis.

(1) Ist $f = \sum_{i=0}^n f_i x^i \in K[x]$, so ist $\varepsilon_a(f) = \sum_{i=0}^n f_i a^i \in A$. Die so gegebene Abbildung $\varepsilon_a : K[x] \to A$ ist wohldefiniert, da durch das Polynom $f = (f_0, f_1, \ldots, f_n, 0, \ldots) \in K^{\mathbb{Z}_{\geq 0}}$ seine Koeffizienten f_i eindeutig bestimmt sind und die Potenzen von a Elemente des K-Vektorraumes A sind.
ε_a ist linear: Seien $f = \sum_{i=0}^n f_i x^i, g = \sum_{j=0}^m g_j x_j \in K[x]$, $h \in K$, wobei wir nach Ergänzung von Nullen ohne Einschränkung annehmen dürfen, dass $m = n$ ist. Dann ist für $h \in K$

$$\varepsilon_a(hf + g) = \sum_{i=0}^n (hf_i + g_i)a^i = \sum_{i=0}^n (hf_i a^i + g_i a^i) = h\sum_{i=0}^n f_i a^i + \sum_{i=0}^n g_i a^i = h\varepsilon_a(f) + \varepsilon_a(g).$$

Dass es ein K-Algebrenhomomorphismus ist, ist nun völlig analog.

(2) ist eine Definition.

(3) Übung.

(4) Da $\varepsilon_a(p) = p(a)$ ist per Definition ist, ist (4) bloß ein Spezialfall von (3).

(5) Sind $a_1, \ldots, a_s$ paarweise verschiedene Wurzeln von p, so gilt nach (3) bzw. (4): $(x - a_i)$ teilt p für $i = 1, \ldots, s$. Da die $(x - a_i)$ paarweise verschiedene irreduzible Polynome sind, ist auch das Polynom $\prod_{i=1}^s (x - a_i)$ vom Grad s ein Teiler von p und somit $\mathrm{Grad}(p) \geq s$.

(6) Übung, etwa mit Hilfe der Lagrangeinterpolation (siehe Übung 2.4.7). □

Übung 2.4.6. Sei $p \in K[x]$ vom Grad 2 oder 3. Zeigen Sie: p ist genau dann irreduzibel, falls p keine Wurzeln (in K) hat. Was ist bei Polynomen vom Grad 4?

Übung 2.4.7 (Lagrangeinterpolation)**.** Seien $a_1, \ldots, a_n \in K$ beliebige Elemente und $s_1, \ldots, s_n \in K$ paarweise verschiedene Elemente. Man zeige: Es existiert genau ein $p \in K[x]_{\mathrm{Grad}<n}$ mit $p(s_i) = a_i$ für $i = 1, \ldots, n$.
Hinweis: Setzen Sie $q := (x - s_1) \ldots (x - s_n)$ und betrachten Sie $\frac{q}{x - s_i}$ für $i = 1, \ldots, n$.

Übung 2.4.8. Definieren Sie die Vielfachheit einer Wurzel.

Übung 2.4.9. Geben sie eine Normalformen der Elemente von $K(x)/K[x]$ an. Was versteht man unter Partialbruchzerlegung?

Definition 2.4.14. Ein Körper K heißt **algebraisch abgeschlossen**, falls jedes nicht konstante Polynom über K eine Wurzel (in K) hat, d.h. falls jedes nicht konstante Polynom in Linearfaktoren zerfällt.

Wir zitieren ohne Beweis:

Hauptsatz 2.4.15 (GAUSS , sogenannter Fundamentalsatz der Algebra)**.** *Der Körper $\mathbb{C}$ der komplexen Zahlen ist algebraisch abgeschlossen.*

Bemerkung 2.4.16.

(1) Sei R ein kommutativer Ring. $R^{\mathbb{Z}_{\geq 0}} =: R[[x]]$ mit der bekannten Multiplikation heißt der Potenzreihenring über R und entsprechend $R[x]$ der Polynomring über R. Beides sind kommutative Ringe.

(2) Nimmt man in (1) $R := K[t]$ den Polynomring in t über dem Körper K, so erhält man $K[t, x] := K[t][x]$, den Polynomring über K in t, x, eine kommutative K-Algebra. Es gilt:
$$K[t, x] \to K[x, t] : \sum_i (\sum_j a_{i,j} t^j) x^i \mapsto \sum_j (\sum_i a_{i,j} x^j) t^i$$
is ein K-Algebrenisomorphismus. Man nennt das maximale $i + j$ mit $a_{i,j} \neq 0$ auch den Gesamtgrad oder einfach Grad des Polynoms $\sum_{i,j} a_{i,j} t^i x^j$.

(3) Analog konstruiert man $K[[t, x]] := K[[x]][[t]]$ und sieht dass diese K-Algebra zu $K[[x, t]]$ isomorph ist.

Übung 2.4.10. Sind $K[[t]][x]$ und $K[t][[x]]$ auch isomorph, und zwar durch einen Isomorphismus, der Isomorphismus von 2.4.16.(2) fortsetzt?

Kapitel 3

Struktur endlich erzeugter Vektorräume

3.1 Erzeugen von Teilräumen

Lernziel: Erzeugnisse als Schnitte von Teilräumen, der kleinste umfassende Teilraum einer Menge, Erzeugnisse als Menge von Linearkombinationen, endliche Erzeugendensysteme, minimale Erzeugendensysteme.

Für Teilmengen hatten wir Durchschnitt und Vereinigung als Operationen, die aus Teilmengen neue Teilmengen machte. Gibt es etwas entsprechendes für Teilräume von Vektorräumen?

Beispiel 3.1.1. Sei $\mathcal{V}$ ein K-Vektorraum mit Teilräumen $\mathcal{T}_1, \mathcal{T}_2 \leq \mathcal{V}$. Es ist $\mathcal{T}_1 \cup \mathcal{T}_2 \leq \mathcal{V}$ genau dann, wenn $\mathcal{T}_1 \subseteq \mathcal{T}_2$ oder $\mathcal{T}_2 \subseteq \mathcal{T}_1$.

Beweis. Die Rückrichtung ist trivial. Nun zur Hinrichtung: Angenommen es gilt weder $\mathcal{T}_1 \subseteq \mathcal{T}_2$ noch $\mathcal{T}_2 \subseteq \mathcal{T}_1$. Dann wähle $x \in \mathcal{T}_1 \setminus \mathcal{T}_2$ und $y \in \mathcal{T}_2 \setminus \mathcal{T}_1$. Dann ist $x + y \notin \mathcal{T}_1 \cup \mathcal{T}_2$ denn $x + y \notin \mathcal{T}_1$, da sonst auch $y = (x + y) - x \in \mathcal{T}_1$, und $x + y \notin \mathcal{T}_2$, da sonst auch $x = (x + y) - y \in \mathcal{T}_2$. □

Beim Durchschnitt sieht es schon besser aus.

Satz 3.1.2. *Sei $M \neq \emptyset$ eine Menge von Teilräumen $\mathcal{W}$ von $\mathcal{V}$. Dann gilt:*

$$\bigcap_{\mathcal{W} \in M} \mathcal{W} \leq \mathcal{V}.$$

Beweis. Setze $\mathcal{T} := \bigcap_{\mathcal{W} \in M} \mathcal{W}$. Zeige $\mathcal{T} \leq \mathcal{V}$:

- $\mathcal{T} \neq \emptyset$, denn $0 \in \mathcal{W}$ für alle $\mathcal{W} \in M$ und daher $0 \in \bigcap_{\mathcal{W} \in M} \mathcal{W}$.
- Seien $X, Y \in \mathcal{T}$ und $a \in K$. Zu zeigen ist $aX+Y \in \mathcal{T}$: Es gilt aber $X, Y \in \mathcal{W}$ für alle $\mathcal{W} \in M$, also $aX + Y \in \mathcal{W}$ für alle $\mathcal{W} \in M$, also $aX + Y \in \bigcap_{\mathcal{W} \in M} \mathcal{W}$. □

Beispiel 3.1.3. Gegeben sei ein homogenes lineares Gleichungssystem mit m Gleichungen. Dann ist der Lösungsraum der Schnitt der Lösungsräume der einzelnen Gleichungen.

Was ist nun der angemessene Ersatz für die Vereinigung? Wir gehen diese Frage etwas allgemeiner an, indem wir fragen, wie kann man aus einer Teilmenge eines K-Vektorraumes einen Teilraum machen? Gibt es z.B. einen kleinsten Teilraum, der diese Menge enthält?

Definition 3.1.4. Sei $\mathcal{V}$ ein K-Vektorraum und $M \subset \mathcal{V}$.

(1) Das **Erzeugnis (Vektorraumerzeugnis)** $\langle M \rangle$ **von** M ist der Schnitt aller Teilräume von $\mathcal{V}$, die M enthalten:

$$\langle M \rangle := \bigcap_{\substack{\mathcal{W} \leq \mathcal{V} \\ M \subseteq \mathcal{W}}} \mathcal{W},$$

und somit der kleinste Teilraum von $\mathcal{V}$, der M enthält.

(2) Eine **Linearkombination von Elementen aus** M ist ein Vektor $V \in \mathcal{V}$, für den ein $n \in \mathbb{N}$, $a_1, a_2, \ldots, a_n \in K$, kurz $a \in K^n$, und $X_1, X_2, \ldots, X_n \in M$, kurz $X \in M^n$, existieren mit $V = a_1X_1 + \ldots + a_nX_n$. Ist $M = \emptyset$, so ist der Nullvektor 0 die einzige Linearkombination von Vektoren aus M. Die Menge aller Linearkombination von M bezeichnen wir mit

$$\mathcal{LK}(M) := \left\{ \sum_{i=1}^{n} a_i X_i \mid n \in \mathbb{N}_0, X_1, \ldots, X_n \in M, a_1, \ldots, a_n \in K \right\}.$$

Satz 3.1.5. *Sei $\mathcal{V}$ ein K-Vektorraum und $M \subseteq \mathcal{V}$ ein Teilmenge. Dann gilt*

$$\langle M \rangle = \mathcal{LK}(M).$$

Beweis.

0. Behauptung: $M \subseteq \mathcal{LK}(M)$. Trivial.

1. Behauptung: $\mathcal{LK}(M) \leq \mathcal{V}$.
Beweis: $0 \in \mathcal{LK}(M)$, also $\mathcal{LK}(M) \neq \emptyset$. Seien $V, W \in \mathcal{LK}(M)$ und $s \in K$. Dann existieren $m, n \in \mathbb{N}, a \in K^m, X \in M^m$ mit $V = a_1X_1 + \ldots + a_mX_m$ und $b \in K^n, Y \in M^n$ mit $W = b_1Y_1 + \ldots + b_nY_n$. Also ist auch

$$sV + W = sa_1X_1 + \cdots + sa_mX_m + b_1Y_1 + \cdots + b_nY_n \in \mathcal{LK}(M).$$

Aus 0. und 1. folgt, dass $\mathcal{LK}(M)$ ein Teilraum von $\mathcal{V}$ ist, welcher M enthält. Da $\langle M \rangle$ der kleinste solche Teilraum ist, gilt somit

$$\langle M \rangle \subseteq \mathcal{LK}(M).$$

Wir wollen nun die umgekehrte Inklusion zeigen. Dazu zeigen wir, dass jeder Teilraum von $\mathcal{V}$, welcher M enthält auch $\mathcal{LK}(M)$ enthält und somit $\mathcal{LK}(M)$ im Durchschnitt dieser Teilräume liegt:

2. Behauptung: Ist $\mathcal{W} \leq \mathcal{V}$ mit $M \subseteq \mathcal{W}$, dann gilt $\mathcal{LK}(M) \subseteq \mathcal{W}$.
Beweis: Sei $V \in \mathcal{LK}(M)$. Dann gibt es $X_1, \ldots, X_m \in M$ und $a_1, \ldots, a_m \in K$ mit $V = a_1X_1 + \cdots + a_mX_m$. Da $M \subseteq \mathcal{W}$ gilt $X_1, \ldots, X_m \in \mathcal{W}$. Da $\mathcal{W} \leq \mathcal{V}$ ein Teilraum ist, liegt somit auch $V = a_1X_1 + \ldots + a_mX_m$ in $\mathcal{W}$, wie behauptet.

Insgesamt gilt daher die behauptete Gleichheit $\langle M \rangle = \mathcal{LK}(M)$. □

Beispiel 3.1.6. Sei

$$M := \left\{ \begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \right\} \subseteq \mathbb{Q}^3.$$

Dann ist

$$\langle M\rangle = \mathcal{LK}(M) = \left\{ a_1 \begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix} + a_2 \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix} + a_3 \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \mid a_1, a_2, a_3 \in \mathbb{Q} \right\} \leq \mathbb{Q}^3.$$

Also ist $\langle M\rangle = \mathrm{Bild}(\widetilde{A})$ mit $A = \begin{pmatrix} 1 & 1 & -1 \\ 2 & 1 & 1 \\ -3 & -2 & 0 \end{pmatrix}$. Als Übung können Sie zeigen, dass

$$\langle M\rangle = \mathrm{Kern}(\widetilde{S}) = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{Q}^{3\times 1} \mid x_1 + x_2 + x_3 = 0 \right\} \text{ mit } S = \begin{pmatrix} 1 & 1 & 1 \end{pmatrix}.$$

Insbesondere folgt

$$\langle M\rangle = \left\langle \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \right\rangle.$$

Definition 3.1.7. Seien $\mathcal{T}_1, \mathcal{T}_2 \leq \mathcal{V}$. Dann definiert man als Ersatz für die Vereinigung die **Summe** der beiden Teilräume:

$$\mathcal{T}_1 + \mathcal{T}_2 := \langle \mathcal{T}_1 \cup \mathcal{T}_2 \rangle.$$

Oft schreiben wir auch $\langle \mathcal{T}_1, \mathcal{T}_2 \rangle$ statt $\langle \mathcal{T}_1 \cup \mathcal{T}_2 \rangle$.

Bemerkung 3.1.8. Seien $\mathcal{T}_1$ und $\mathcal{T}_2$ Teilräume des K-Vektorraumes $\mathcal{V}$.

(1) $\mathcal{T}_1 + \mathcal{T}_2 = \{X_1 + X_2 \mid X_1 \in \mathcal{T}_1, X_2 \in \mathcal{T}_2\}$.

(2) $\varphi : \mathcal{T}_1 \oplus_a \mathcal{T}_2 \to \mathcal{V} : (X_1, X_2) \mapsto X_1 + X_2$ ist eine lineare Abbildung mit $\mathrm{Bild}(\varphi) = \mathcal{T}_1 + \mathcal{T}_2$ und $\mathrm{Kern}(\varphi) = \{(T, -T) | T \in \mathcal{T}_1 \cap \mathcal{T}_2\}$. Insbesondere gilt die Gleichheit $\mathcal{T}_1 + \mathcal{T}_2 = \mathcal{T}_1 \oplus_i \mathcal{T}_2$ genau dann, wenn $\mathcal{T}_1 \cap \mathcal{T}_2 = \{0\}$.

(3) Ist $M = \{X\} \subseteq \mathcal{V}$. Dann ist $\langle X\rangle := \langle M\rangle = \{aX \mid a \in K\}$.

Beispiel 3.1.9. Im Beispiel 3.1.6 ist $\langle M\rangle = \mathcal{T}_1 + \mathcal{T}_2$ mit

Ende Vorl. 15 30.11

$$\mathcal{T}_1 = \langle \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \rangle \text{ und } \mathcal{T}_2 = \langle \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \rangle.$$

Hier ist $\mathcal{T}_1 \cap \mathcal{T}_2 = \{0\}$ und φ ein Isomorphismus. Es ist $\mathcal{T}_1 + \mathcal{T}_2 = \mathcal{T}_1 \oplus_i \mathcal{T}_2$.

Definition 3.1.10. Ein K-Vektorraum $\mathcal{V}$ heißt **endlich erzeugt**, falls eine endliche Teilmenge $M \subseteq \mathcal{V}$ mit $\mathcal{V} = \langle M\rangle$ existiert. Jedes solche M heißt ein endliches **Erzeugendensystem** von $\mathcal{V}$.

Bemerkung 3.1.11.

(1) $\mathcal{V} := K^{n\times 1}$ ist endlich erzeugt, denn die Spalten $e_i := (I_n)_{-,i}$ der Einheitsmatrix bilden ein Erzeugendensystem, d.h.

$$\mathcal{V} = \langle e_1, \ldots, e_n \rangle,$$

da nach Bemerkung 1.2.6

$$\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = a_1 e_1 + \cdots + a_n e_n.$$

(2) $K[x]$ ist als K-Vektorraum nicht endlich erzeugt, denn für jede endliche Teilmenge $M \subseteq K[x]$ sind die Grade der Polynome in $\langle M \rangle$ beschränkt durch das Maximum der Grade der Polynome in M.
Jedoch ist $\{1, x, x^2, \ldots\} = \{x^i \mid i \in \mathbb{N} \cup \{0\}\}$ ein unendliches Erzeugendensystem von $K[x]$, denn jedes Polynom ist (endliche) Linearkombination der Potenzen von x.

(3) Ist $\varphi : \mathcal{V} \to \mathcal{W}$ ein *Epimorphismus* und $M \subseteq \mathcal{V}$ ein Erzeugendensystem von $\mathcal{V}$. Dann ist $\varphi(M) \coloneqq \{\varphi(V) \mid V \in M\}$ ein Erzeugendensystem von $\mathcal{W}$. D.h. Erzeugendensysteme gehen bei Epimorphismen in Erzeugendensysteme über.

(4) Ist $\mathcal{V}$ endlich erzeugter K-Vektorraum und $\varphi : \mathcal{V} \to \mathcal{W}$ ein Epimorphismus, so ist auch $\mathcal{W}$ endlich erzeugt.

(5) Sind $\mathcal{V}, \mathcal{W}$ endlich erzeugte K-Vektorräume, so ist auch $\mathcal{V} \oplus \mathcal{W}$ endlich erzeugt.

Definition 3.1.12. Eine Teilmenge M eines K-Vektorraumes $\mathcal{V}$ heißt **minimales Erzeugendensystem** von $\mathcal{V}$, falls $\langle M \rangle = \mathcal{V}$ aber $\langle M \setminus \{X\} \rangle \neq \mathcal{V}$ für alle $X \in M$.

Bemerkung 3.1.13. Ist $M \subseteq \mathcal{V}$ ein minimales Erzeugendensystem, so gilt $X \notin \langle M - \{X\} \rangle$ für alle $X \in M$.

Bemerkung 3.1.14. Ein endlich erzeugter K-Vektorraum $\mathcal{V}$ ist epimorphes Bild von K^n, wobei n die Anzahl der Erzeuger ist.

Beweis. Sei $\{X_1, \ldots, X_n\} \subseteq \mathcal{V}$ Erzeugendensystem des K-Vektorraumes $\mathcal{V}$. Dann ist für das Tupel $X = (X_1, \ldots, X_n) \in \mathcal{V}^n$ (vgl. Beispiel 2.3.6.(3)) der Linearkombinationshomomorphismus bezüglich X

$$\lambda_X : K^n \to \mathcal{V},\ a \mapsto a_1 X_1 + \cdots + a_n X_n$$

ein Epimorphismus. □

Es wird sich zeigen, dass λ_X sogar ein Isomorphismus ist, wenn die X_i's ein minimales Erzeugendensystem bilden.

3.2 Lineare Unabhängigkeit

Lernziel: Eindeutigkeitsfragen bei Linearkombinationen, lineare Unabhängigkeit von Folgen von Vektoren, Reduktion eines Erzeugendensystems auf ein linear unabhängiges Erzeugendensystem, Charakterisierung von Basen.

Wie aus dem Beweis der letzten Bemerkung bereits ersichtlich, ist es für viele Zwecke günstiger, mit Folgen statt mit Mengen von Vektoren zu arbeiten, wenn man über Erzeugen, Linearkombinationen, etc. spricht. Alles was im letzten Abschnitt über endliche Erzeugendensysteme gemacht wurde, kann man auch mit endlichen Folgen von Vektoren machen, indem man die Begriffsbildungen auf das Bild der Folge bezieht.

Bemerkung 3.2.1. Sei $\mathcal{V}$ ein K-Vektorraum und $X = (X_1, \ldots, X_n) \in \mathcal{V}^n$. Folgende Aussagen sind äquivalent:

(1) Aus $a_1 X_1 + a_2 X_2 + \cdots + a_n X_n = 0$ mit $a_i \in K$ folgt $a_1 = a_2 = \ldots = a_n = 0$.

(2) Der Linearkombinationshomomorphismus bezüglich X

$$\lambda_X : K^n \to \mathcal{V},\ a \mapsto a_1 X_1 + a_2 X_2 + \cdots + a_n X_n$$

ist ein Monomorphismus, sprich $\mathrm{Kern}(\lambda_X) = \{0\}$.

(3) Für jedes $Y \in \langle X \rangle$ gibt es genau ein $a \in K^n$ mit $Y = a_1X_1 + a_2X_2 + \cdots + a_nX_n$.

Beweis.

(1) $\Rightarrow$ (2): Sei $(b_1, \ldots, b_n) \in \mathrm{Kern}(\lambda_X)$, d.h. $\sum_{i=1}^n b_iX_i = 0$. Dann gilt mit (1), dass auch $b_1 = \ldots = b_n = 0$ ist. Also ist $\mathrm{Kern}(\lambda_X) = \{0\}$.

(2) $\Rightarrow$ (3): Nach (2) ist λ_X eine injektive lineare Abbildung, d.h. für jedes $Y \in \mathrm{Bild}(\lambda_X) = \langle X \rangle := \langle X_1, \ldots, X_n \rangle$ gibt es genau ein $a \in K^n$ mit $Y = \lambda_X(a) = a_1X_1 + a_2X_2 + \cdots + a_nX_n$.

(3) $\Rightarrow$ (1): $0 = 0X_1 + \ldots + 0X_n$ hat nach (3) eine eindeutige Darstellung. □

Definition 3.2.2.

- $X \in \mathcal{V}^n$ heißt **linear unabhängig**, falls eine (und damit alle drei) der Aussagen von Bemerkung 3.2.1 zutreffen. Anderenfalls heißt X **linear abhängig**.
- Eine endliche Teilmenge $\{X_1, \ldots, X_n\}$ von n Elementen von $\mathcal{V}$ heißt linear unabhängig oder linear abhängig, wenn die Folge $(X_1, \ldots, X_n)$ die entsprechende Eigenschaft hat.
- Eine unendliche Teilmenge $X \subseteq \mathcal{V}$ heißt linear unabhängig, falls jede endliche Teilmenge von X linear unabhängig ist.

Der Begriff „lineare Unabhängigkeit“ von endlichen Mengen ist ein wohldefinierter Begriff, d.h. unabhängig von der Umordnungen der Vektoren (Übung). Daher sagt man anstelle von $(X_1, \ldots, X_n)$ ist linear unabhängig etwas lockerer: „$X_1, \ldots, X_n$ sind linear unabhängig“.

Man beachte, über die Größe von $\mathrm{Kern}(\lambda_X)$ liefert 3.2.1.(2) auch eine Vorstellung, wie sehr X linear abhängig ist. Man könnte die Elemente von $\mathrm{Kern}(\lambda_X)$ als „lineare Abhängigkeiten“ von X bezeichnen. Lineare Unabhängigkeit liegt vor, wenn man nur die triviale lineare Abhängigkeit hat.

Bemerkung 3.2.3. Ist $X \in \mathcal{V}^n$, $\varphi : \mathcal{V} \to \mathcal{W}$ linear und $\varphi \circ X = (\varphi(X_1), \ldots, \varphi(X_n)) \in \mathcal{W}^n$ linear unabhängig, so ist auch X linear unabhängig.

Beweis. Sei $a \in K^n$ mit $a_1X_1 + \cdots + a_nX_n = 0$, dann ist auch $a_1\varphi(X_1) + \cdots + a_n\varphi(X_n) = 0$, also $a_1 = \ldots = a_n = 0$. □

Beispiel 3.2.4.

(1) $(e_1, \ldots, e_n) \in (K^{n\times 1})^n$ aus Bemerkung 3.1.11.(1) ist linear unabhängig.

(2) $(\sin, \cos) \in (\mathbb{R}^{\mathbb{R}})^2$ ist linear unabhängig. Denn

$$\alpha : \langle \sin, \cos \rangle \to \mathbb{R}^2,\ f \mapsto (f(0), f(\pi/2))$$

ist linear und $(\alpha(\sin), \alpha(\cos)) = ((0,1),(1,0))$ ist linear unabhängig.

(3) Sei $\kappa_1; \mathbb{R} \to \mathbb{R} : x \mapsto 1$. Dann ist $(\sin^2, \cos^2, \kappa_1) \in (\mathbb{R}^{\mathbb{R}})^3$ linear abhängig, denn nach Pythagoras gilt

$$\sin(x)^2 + \cos(x)^2 = 1 \text{ für alle } x \in \mathbb{R}$$

Die Begriffe „erzeugen“ und „linear unabhängig“ verhalten sich unter bestimmten Aspekten dual zueinander. In diesem Sinne bearbeite man die folgende Aufgabe.

Übung 3.2.1. Sei $M \subseteq \mathcal{V}$ endlich. Zeigen Sie:

(1) Gilt $\langle M\rangle = \mathcal{V}$ und $M \subseteq N \subseteq \mathcal{V}$, so gilt $\langle N\rangle = \mathcal{V}$.

(2) Ist M linear unabhängig und $N \subseteq M$, so ist N linear unabhängig.

Ein weiteres Indiz für dieses duale Verhalten, war das Zusammenspiel mit linearen Abbildungen: Erzeugendensysteme werden durch lineare Abbildungen auf Erzeugendensysteme des Bildes abgebildet. Linear unabhängige Vektoren im Bild, kommen von linear unabhängigen Vektoren im Urbild, vgl. Bemerkung 3.2.3.

Bemerkung 3.2.5. Sei $X = (X_1, \ldots, X_n)$ linear unabhängig und $Y \in \mathcal{V}$. Dann ist $Y \in \langle X_1, \ldots, X_n\rangle$ genau dann, wenn $(X_1, \ldots, X_n, Y)$ linear abhängig ist.

Beweis.

($\Rightarrow$) Ist $Y \in \langle X_1, \cdots, X_n\rangle$, so gibt es $a_1, \ldots, a_n \in K$ mit $Y = a_1X_1 + \cdots + a_nX_n$ und somit ist $a_1X_1 + \cdots + a_nX_n + \underbrace{(-1)}_{\neq 0}Y = 0$ eine nichttriviale Linearkombination der 0.

($\Leftarrow$) Umgekehrt sei $a_1X_1 + \ldots + a_nX_n + bY = 0$ mit $\{a_1, \ldots, a_n, b\} \neq \{0\}$. Dann ist $b \neq 0$, denn sonst wäre schon $a_1X_1 + \ldots + a_nX_n = 0$ eine nichttriviale Linearkombination, im Widerspruch zur Voraussetzung, dass X linear unabhängig ist. Also ist $Y = -\frac{a_1}{b}X_1 - \ldots - \frac{a_n}{b}X_n \in \mathcal{LK}(X) = \langle X_1, \ldots, X_n\rangle$. □

Wie zu erwarten ist die Frage nach minimalen Erzeugendensystemen dual zu der nach maximal linear unabhängigen Systemen:

Satz 3.2.6. *Sei $\mathcal{V}$ ein K-Vektorraum und $X = (X_1, \ldots, X_n) \in \mathcal{V}^n$. Folgende Aussagen sind äquivalent:*

(1) X ist ein minimales Erzeugendensystem von $\mathcal{V}$.

(2) X ist maximal linear unabhängig in $\mathcal{V}$, d.h. X ist linear unabhängig und $(X_1, \ldots, X_n, Y)$ ist linear abhängig für jedes $Y \in \mathcal{V}$.

(3) X ist ein linear unabhängiges Erzeugendensystem.

Ein derartiges X heißt eine **Basis** *von $\mathcal{V}$.*

Beweis.

(2) $\Leftrightarrow$ (3): Folgt unmittelbar aus Bemerkung 3.2.5.

(1) $\Leftrightarrow$ (3): Dazu sei $X = (X_1, \ldots, X_n)$ ein Erzeugendensystem von $\mathcal{V}$. Wir zeigen die Kontraposition: X ist nicht minimal $\Leftrightarrow$ X ist linear abhängig.

($\Rightarrow$) Sei X nicht minimal. Dann existiert ein $i \in \{1, \ldots, n\}$ mit

$$\mathcal{V} = \langle X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n\rangle.$$

Da $X_i \in \mathcal{V}$ ist, heißt das, dass es $(a_1, \ldots a_{i-1}, a_{i+1}, \ldots, a_n) \in K^{n-1}$ gibt, so dass $X_i = a_1X_1 + \cdots + a_{i-1}X_{i-1} + a_{i+1}X_{i+1} + \cdots + a_nX_n$. Wie oben erhalten wir daraus eine nichttriviale lineare Abhängigkeit von X.

($\Leftarrow$) Sei X linear abhängig und $0 \neq (a_1, \ldots, a_n) \in K^n$ mit $\sum_{j=1}^n a_jX_j = 0$. Dann gibt es ein i mit $a_i \neq 0$ und also $X_i = \sum_{i\neq j} \frac{-a_j}{a_i}X_j \in \langle X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n\rangle$. Somit ist X nicht minimal. □

Folgerung 3.2.7. *$X = (X_1, \ldots, X_n) \in \mathcal{V}^n$ ist genau dann eine Basis von $\mathcal{V}$, wenn*

$$\lambda = \lambda_X : K^n \to \mathcal{V},\ a \mapsto a_1X_1 + a_2X_2 + \ldots + a_nX_n$$

ein Isomorphismus ist, d.h. falls für jedes $V \in \mathcal{V}$ ein eindeutiges $a \in K^n$ existiert mit

$$V = a_1X_1 + \cdots + a_nX_n.$$

Die dann eindeutig existierende Umkehrabbildung $\kappa_X^{tr} = \lambda_X^{-1}$ mit

$$\kappa_X^{tr} : \mathcal{V} \to K^n, a_1X_1 + a_2X_2 + \cdots + a_nX_n \mapsto (a_1, \ldots, a_n)$$

*heißt Zeilen-***Koordinatenabbildung** *bzgl. der Basis X.*

Alternativ kann man natürlich mit Spalten arbeiten, sprich mit

$$\kappa_X : \mathcal{V} \to K^{n\times 1}, a_1X_1 + a_2X_2 + \cdots + a_nX_n \mapsto \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$$

*Wir nennen κ_X die Spalten-***Koordinatenabbildung** *bzgl. der Basis X und $\kappa_X(V)$ die* **Koordinatenspalte** *von $V = \sum_i a_iX_i$ bezüglich der Basis X.*

Ende Vorl. 16 05.12

Merke: $B \in \mathcal{V}^n$ ist Basis von $\mathcal{V}$, genau dann wenn jedes $V \in \mathcal{V}$ eine *eindeutige* Linearkombination der Vektoren in B ist.

Existenz von Linearkombinationen	=	B ist Erzeugendensystem
Eindeutigkeit von Linearkombinationen	=	B ist linear unabhängig
Existenz+Eindeutigkeit von Linearkombinationen	=	B ist Basis.

Die Wahl einer Basis $B \in \mathcal{V}^n$ definiert einen *Isomorphismus* $\kappa_B : \mathcal{V} \to K^{n\times 1}$ zwischen $\mathcal{V}$ und dem Spaltenraum.

Zwar haben wir jetzt eine schöne Charakterisierung von Basen, aber es drängt sich eine Frage auf: Haben je zwei Basen (von endlich erzeugten Vektorräumen) dieselbe Anzahl von Vektoren? Dies wollen wir im nächsten Abschnitt beantworten.

3.3 Der STEINITZsche Austauschsatz

Lernziel: Beweis und Anwendungen des Austauschsatzes, Wohldefiniertheit der Dimension, diverse Dimensionsformeln.

Nach unserer Charakterisierung von Basen von endlich erzeugten Vektorräumen gibt es zwei mögliche Strategien, eine Basis zu konstruieren:

(1) Man beginnt mit einem endlichen Erzeugendensystem und läßt der Reihe nach Vektoren weg, die im Erzeugnis der übrigen liegen, also linear abhängig von den übrigen sind, bis man ein minimales Erzeugendensystem hat, welches dann auch automatisch linear unabhängig ist.

(2) Man beginnt mit einem linear unabhängigen System und fügt der Reihe nach Vektoren hinzu, so dass das erweiterte System wieder linear unabhängig ist. Wenn dieser Vorgang terminiert hat man ein maximal linear unabhängiges System.

Bei dem ersten Prozess ist klar, dass man eine Basis bekommt; allerdings weiß man nicht ob zwei Basen immer gleich viele Elemente enthalten. Bei dem zweiten Prozess ist nicht einmal klar, dass er terminiert. Aus diesem Dilemma führt der STEINITZsche Austauschsatz heraus, welcher der erste tiefere Struktursatz über endlich erzeugte Vektorräume ist, den wir in dieser Vorlesung kennenlernen.

Hauptsatz 3.3.1. *(STEINITZscher Austauschsatz) Sei $\mathcal{V}$ ein K-Vektorraum und $X \in \mathcal{V}^n$ ein Erzeugendensystem von $\mathcal{V}$, d.h. $\mathcal{V} = \langle X_1, X_2, \ldots, X_n\rangle$, und sei $Y = (Y_1, \ldots, Y_s) \in \mathcal{V}^s$ linear unabhängig. Dann gilt $s \leq n$ und nach geeigneter Umordnung der X_i's ist $(Y_1, \ldots, Y_s, X_{s+1}, \ldots, X_n)$ ein Erzeugendensystem von $\mathcal{V}$. Mit anderen Worten: Es existiert eine Permutation $\sigma \in S_n$, so dass $(Y_1, \ldots, Y_s, X_{\sigma(s+1)}, \ldots, X_{\sigma(n)})$ ein Erzeugendensystem von $\mathcal{V}$ ist.*

Beweis. Wir führen den Beweis durch vollständige Induktion über s, der Anzahl der linear unabhängigen Vektoren.
Induktionsanfang: Sei $s = 1$. Da $Y_1 \in \mathcal{V} = \langle X\rangle = \langle X_1, \ldots, X_n\rangle$, folgt: Es existiert ein $a \in K^n$ mit

$$Y_1 = a_1X_1 + \ldots + a_nX_n.$$

Da Y_1 linear unabhängig ist, d.h. $Y_1 \neq 0$ ist $a \neq 0$. Also existiert ein $i \in \underline{n}$ mit $a_i \neq 0$. Nach eventueller Umordnung der X_j können wir ohne Beschränkung der Allgemeinheit (oBdA) annehmen, dass $i = 1$ ist, also $a_1 \neq 0$. Also folgt

$$X_1 = \frac{1}{a_1}Y_1 - \frac{a_2}{a_1}X_2 - \cdots - \frac{a_n}{a_1}X_n.$$

Also $X_1 \in \langle Y_1, X_2, \ldots, X_n\rangle$, d.h. $\mathcal{V} = \langle Y_1, X_2, \ldots, X_n\rangle$.
Induktionsannahme: Sei die Behauptung gültig für $s - 1$, sprich sind $(Y_1, \ldots, Y_{s-1})$ linear unabhängig, so ist $s - 1 \leq n$ und nach Umordnung von X gilt $(Y_1, \ldots, Y_{s-1}, X_s, \ldots X_n)$ erzeugt $\mathcal{V}$.
Induktionsschritt: Sei $(Y_1, \ldots, Y_s)$ linear unabhängig. Dann ist auch $(Y_1, \ldots, Y_{s-1})$ linear unabhängig und wir wenden die Induktionsannahme darauf an. Insbesondere haben wir

$$Y_s \in \mathcal{V} = \langle Y_1, \ldots, Y_{s-1}, X_s, \ldots X_n\rangle$$

also existiert ein $a \in K^n$ mit

$$Y_s = a_1Y_1 + \cdots + a_{s-1}Y_{s-1} + a_sX_s + \cdots + a_nX_n.$$

Da Y linear unabhängig ist, ist $Y_s \neq 0$, also $a \neq 0$. Wäre $a_s = \cdots = a_n = 0$, so wäre Y linear abhängig. Also existiert ein i mit $s \leq i \leq n$ mit $a_i \neq 0$. Dies impliziert bereits $s \leq n$. Nach Umnummerierung der X_i können wir wieder oBdA[1] annehmen, dass $i = s$ gilt. Wie oben bekommen wir wieder $X_s \in \langle Y_1, \ldots, Y_s, X_{s+1}, \ldots X_n\rangle$ und folgern schließlich $\mathcal{V} = \langle Y_1, \ldots, Y_s, X_{s+1}, \ldots X_n\rangle$. □

Folgerung 3.3.2. *Sei $\mathcal{V}$ ein endlich erzeugter K-Vektorraum. Dann existiert ein eindeutiges $n \in \mathbb{Z}_{\geq 0}$, so dass jede Basis von $\mathcal{V}$ aus genau n Vektoren besteht. Dieses n nennt man die* **Dimension** *von $\mathcal{V}$ und schreibt $\operatorname{Dim}\mathcal{V} = n$.*

Falls $\mathcal{V}$ nicht endlich erzeugbar ist, also kein endliches Erzeugendensystem hat, schreiben wir pauschal $\operatorname{Dim}\mathcal{V} = \infty$.

Beweis. Der Fall des Nullvektorraumes ist klar, $\operatorname{Dim}(\{0\}) = 0$.
Sei also $\mathcal{V} \neq \{0\}$. Wegen der endlichen Erzeugbarkeit existieren ein $n \in \mathbb{N}$ und ein Erzeugendensystem $X \in \mathcal{V}^n$. Durch Weglassen von Vektoren können wir erreichen, dass X ein

[1] (Man beachte: Die beiden oBdA-Annahmen sind lockere Umschreibung der Tatsache, dass wir an diesen Stellen eine Permutation der X_i vornehmen müssen.)

linear unabhängiges Erzeugendensystem ist, also oBdA X linear unabhängig. Ist $Y \in \mathcal{V}^s$ linear unabhängig, so folgt aus dem Satz von STEINITZ $s \leq n$. Ist Y zudem ein Erzeugendensystem, also eine Basis, so folgt wiederum aus dem Satz von STEINITZ 3.3.1 $n \leq s$, also $s = n$. □

Beispiel 3.3.3.

(1) $\operatorname{Dim}\{0\} = 0$, denn $\{0\} = \langle\emptyset\rangle$.

(2) $\operatorname{Dim} K^n = n$, denn die Standardbasis

$$\mathcal{S} := ((1, 0, \ldots, 0), (0, 1, 0, \ldots, 0), \ldots, (0, \ldots, 0, 1))$$

ist offensichtlich eine Basis aus n Elementen. Entsprechend gilt

$$\operatorname{Dim} K^{n\times 1} = \operatorname{Dim} K^{1\times n} = n.$$

(3) Allgemeiner haben wir für beliebige (insbesondere endliche) Mengen M

$$\operatorname{Dim} K^M = |M|$$

denn im endlichen Fall bilden die charakteristischen Funktionen der einelementigen Teilmengen von M (in irgendeiner Anordnung) eine Basis. Für den Fall $M = \emptyset$ ist K^M ein Nullvektorraum[2].

Ist M unendlich, so sind die charakteristischen Funktionen der einelementigen Teilmengen von M immer noch linear unabhängig (aber sicher keine Basis mehr), weshalb man dann $\operatorname{Dim} K^M = \infty$ hat.

(4) $\operatorname{Dim}(K^{n\times m}) = |\underline{n} \times \underline{m}| = nm$.

(5) $\operatorname{Dim}(K[X]) = \infty$. Eine Basis von $K[X]$ ist $(1, X, X^2, \ldots)$.

(6) $\operatorname{Dim}(K[X]_{\mathrm{Grad}<n}) = n$. Eine Basis ist $(1, x, \ldots, x^{n-1})$.

Bemerkung 3.3.4. Seien $\mathcal{V}, \mathcal{W}$ zwei K-Vektorräume, $\varphi : \mathcal{V} \to \mathcal{W}$ linear.

(1) Ist $(\varphi(X_1), \ldots, \varphi(X_n)) \in \mathcal{W}^n$ linear unabhängig, so ist auch $(X_1, \ldots, X_n) \in \mathcal{V}^n$ linear unabhängig.

(2) Ist $\langle Y_1, \ldots, Y_m\rangle = \mathcal{V}$, so ist $\langle\varphi(Y_1), \ldots, \varphi(Y_m)\rangle = \operatorname{Bild}(\varphi) \leq \mathcal{W}$.

(3) Ist φ ein Isomorphismus, so bildet φ jede Basis von $\mathcal{V}$ auf eine Basis von $\mathcal{W}$ ab. Insbesondere gilt dann $\operatorname{Dim}\mathcal{V} = \operatorname{Dim}\mathcal{W}$.

Folgerung 3.3.5. *Sei $\mathcal{V}$ ein endlich erzeugter K-Vektorraum.*

(1) (Basisergänzungssatz) Ist $X \in \mathcal{V}^s$ linear unabhängig, so kann X zu einer Basis von $\mathcal{V}$ ergänzt werden, d.h. es gilt $s \leq \operatorname{Dim}\mathcal{V} =: n$ und es existiert eine Basis $Y \in \mathcal{V}^n$ mit $Y_i = X_i$ für $1 \leq i \leq s$.

(2) Ist $\mathcal{T} \leq \mathcal{V}$, so gilt $\operatorname{Dim}\mathcal{T} \leq \operatorname{Dim}\mathcal{V}$ mit Gleichheit genau dann, wenn $\mathcal{T} = \mathcal{V}$.

(3) Ist $\mathcal{T} \leq \mathcal{V}$ mit Basis $(X_1, \ldots, X_m)$ und $(X_1, \ldots, X_m, Y_1, \ldots, Y_k)$ eine Basis von $\mathcal{V}$, so ist $(Y_1 + \mathcal{T}, \ldots, Y_k + \mathcal{T})$ eine Basis von $\mathcal{V}/\mathcal{T}$. Insbesondere gilt:

$$\boxed{\operatorname{Dim}\mathcal{V}/\mathcal{T} = \operatorname{Dim}\mathcal{V} - \operatorname{Dim}\mathcal{T}.}$$

[2] sprich, ein Vektorraum, der nur aus dem Nullvektor besteht

Beweis.

(1) Direkt aus dem Satz von STEINITZ: Nehme ein endliches Erzeugendensystem, ersetze davon s Vektoren durch $X_1, \ldots, X_s$ (es bleibt ein Erzeugendensystem). Entferne weitere Vektoren, die linear abhängig sind, bis dies nicht mehr geht, also bis nach endlich vielen Schritten ein linear unabhängiges Erzeugendensystem, sprich eine Basis übrig bleibt.

(2) Linear unabhängige Vektoren aus $\mathcal{T}$ sind auch linear unabhängig in $\mathcal{V}$. Also ist die Maximalzahl linear unabhängiger Vektoren aus $\mathcal{T}$ beschränkt durch $\operatorname{Dim}\mathcal{V}$. Diese Maximalzahl ist aber $\operatorname{Dim}\mathcal{T}$.

(3) Dass $(Y_1+\mathcal{T}, \ldots, Y_k+\mathcal{T})$ ein Erzeugendensystem von $\mathcal{V}/\mathcal{T}$ ist, ist klar. Wir zeigen die lineare Unabhängigkeit: Sei $a \in K^k$ mit

$$a_1(Y_1+\mathcal{T}) + \ldots + a_k(Y_k+\mathcal{T}) = 0+\mathcal{T}$$

d.h. $a_1Y_1 + \ldots + a_kY_k \in \mathcal{T}$. Also existiert $b \in K^m$ mit

$$a_1Y_1 + \ldots + a_kY_k = b_1X_1 + \ldots + b_mX_m$$

d.h.

$$a_1Y_1 + \ldots + a_kY_k - b_1X_1 - \ldots - b_mX_m = 0,$$

woraus wir mit der linearen Unabhängigkeit von $(X_1, \ldots, X_m, Y_1, \ldots, Y_k)$ schließen, dass $a = 0$ (und $b = 0$) gilt, d.h. $(Y_1+\mathcal{T}, \ldots, Y_k+\mathcal{T})$ ist linear unabhängig. □

Des Weiteren können wir den Homomorphiesatz um eine quantitative Aussage erweitern.

Folgerung 3.3.6. *Sei $\mathcal{V}$ ein endlich erzeugter K-Vektorraum und $\alpha : \mathcal{V} \to \mathcal{W}$ eine K-lineare Abbildung. Dann gilt:*

$$\boxed{\operatorname{Dim}\operatorname{Bild}(\alpha) + \operatorname{Dim}\operatorname{Kern}(\alpha) = \operatorname{Dim}\mathcal{V}.}$$

Beweis. Da $\operatorname{Bild}(\alpha)$ und $\mathcal{V}/\operatorname{Kern}(\alpha)$ nach dem Homomorphiesatz isomorph sind, haben beide gleiche Dimension und die Behauptung folgt aus Folgerung 3.3.5.(3). □

Beispiel 3.3.7. Sei $\mathcal{W} := \{p \in K[X] \mid \operatorname{Grad}(p) < 6, p(1) = p(2) = 0\}$. Wir zeigen, dass $\mathcal{W}$ ein Teilraum von $K[X]$ ist und bestimmen Sie seine Dimension.
Dazu definieren wir

$$\alpha : \mathcal{V} := K[X]_{\leq 5} \to K^2, p \mapsto (p(1), p(2)).$$

Dann ist α eine lineare Abbildung, $\operatorname{Kern}(\alpha) = \mathcal{W}$, also insbesondere $\mathcal{W} \leq \mathcal{V}$. Weiter ist α surjektiv, da $\alpha(1)$ und $\alpha(X)$ den Raum K^2 erzeugen. Also ist

$$\dim(\mathcal{W}) = \dim(\mathcal{V}) - \dim(\operatorname{Bild}(\alpha)) = 6 - 2 = 4.$$

Wir schließen diesen Abschnitt mit der berühmten GRASSMANNidentität ab.

Ende Vorl. 17 07.12

Folgerung 3.3.8. *Seien $\mathcal{V}_1, \mathcal{V}_2$ endlich erzeugte K-Vektorräume. Dann gilt:*

$$\operatorname{Dim}(\mathcal{V}_1 \oplus_a \mathcal{V}_2) = \operatorname{Dim}\mathcal{V}_1 + \operatorname{Dim}\mathcal{V}_2.$$

Beweis. Mit der Ausführung im Beispiel 2.3.23 ist $\mathcal{V}_1 \oplus_a \mathcal{V}_2/\mathcal{V}_1 \cong \mathcal{V}_2$. Der Rest folgt aus Folgerung 3.3.5.(3). □

Übung 3.3.1. Geben Sie einen direkten Beweis für Folgerung 3.3.8 an.

Folgerung 3.3.9. *Sei $\mathcal{V}$ ein K-Vektorraum mit endlich erzeugten Teilräumen $\mathcal{T}_1, \mathcal{T}_2 \leq \mathcal{V}$. Dann gilt:*

$$\mathrm{Dim}(\mathcal{T}_1 + \mathcal{T}_2) + \mathrm{Dim}(\mathcal{T}_1 \cap \mathcal{T}_2) = \mathrm{Dim}(\mathcal{T}_1 \oplus_a \mathcal{T}_2) = \mathrm{Dim}\,\mathcal{T}_1 + \mathrm{Dim}\,\mathcal{T}_2.$$

Beweis. Offenbar ist

$$\alpha : \mathcal{T}_1 \oplus_a \mathcal{T}_2 \to \mathcal{V} : (T_1, T_2) \mapsto T_1 + T_2$$

linear mit $\mathrm{Bild}(\alpha) = \mathcal{T}_1 + \mathcal{T}_2$ und $\mathrm{Kern}(\alpha)$ isomorph zu $\mathcal{T}_1 \cap \mathcal{T}_2$. Die Behauptung folgt jetzt aus 3.3.6. □

Anhang: Jeder Vektorraum hat eine Basis.

In diesem Abschnitt wollen wir (nicht konstruktiv) zeigen, dass jeder Vektorraum eine Basis besitzt. Dazu müssen wir ein Axiom der Mengenlehre voraussetzen, das sogenannte Zorn'sche Lemma.

Satz 3.3.10. *Folgende Aussagen sind äquivalent.*

(A) (Auswahlaxiom) Sei $\Lambda \neq \emptyset$ und für jedes $\lambda \in \Lambda$ eine nichtleere Menge X_λ gegeben. Dann ist das kartesische Produkt

$$\prod_{\lambda \in \Lambda} X_\lambda = \{(x_\lambda)_{\lambda \in \Lambda} \mid x_\lambda \in X_\lambda\}$$

nicht leer.
(Man kann also simultan für jedes $\lambda \in \Lambda$ ein $x_\lambda \in X_\lambda$ auswählen.)

(Z) (Lemma von Zorn) Sei $(M, \leq)$ eine nicht leere geordnete Menge. Hat jede Kette in M eine obere Schranke in M, so gibt es ein $x \in M$ mit $m \geq x \Rightarrow m = x$ für jedes $m \in M$.

(A) und (Z) sind äquivalente Axiome (vgl. z.B. Halmos, Naive Mengenlehre oder meine Algebra Vorlesung), sie folgen nicht aus den Grundaxiomen der Mengenlehre, man muss eines von ihnen zusätzlich fordern.
Die Aussage (A) haben wir implizit schon einmal benutzt, als wir zeigten, dass jede surjektive Funktion eine Rechtsinverse besitzt, da wir aus jeder Faser ein Element auswählen mussten, um die Rechtsinverse zu „konstruieren".
Wir werden jetzt (Z) benutzen, um die Existenz von Basen zu beweisen.
Zunächst müssen wir die Begriffe „Erzeugendensystem" und „lineare Unabhängigkeit" auch für unendliche Teilmengen eines Vektorraumes präzisieren.

Definition 3.3.11. Sei $\mathcal{V}$ ein K-Vektorraum und $X \subseteq \mathcal{V}$ eine Teilmenge.
(a) X heißt **Erzeugendensystem** von $\mathcal{V}$, falls jedes Element von $\mathcal{V}$ eine endliche Linearkombination der Elemente von X ist.
(b) X heißt **linear unabhängig**, falls alle endlichen Teilmengen von X linear unabhängig sind, d.h. für alle $n \in \mathbb{N}$, $X_1, \ldots, X_n \in X$ mit $X_i \neq X_j$ für $i \neq j$ gilt

$$\sum_{i=1}^{n} a_i X_i = 0 \text{ mit } a_1, \ldots, a_n \in K \Leftrightarrow a_1 = \ldots = a_n = 0.$$

Beachten Sie: $(1, x, x^2, x^3, \ldots) = (x^n | n \in \mathbb{N}_0)$ ist eine Basis von $K[x]$, jedoch kein Erzeugendensystem von $K[[x]]$, da sich Potenzreihen i.a. nur als "unendliche"[3] Linearkombinationen der x^n darstellen lassen. Können Sie eine Basis von $K[[x]]$ angeben?

Satz 3.3.12. *Jeder Vektorraum besitzt eine Basis.*

[3] Solche haben wir explizit ausgeschlossen

Beweis. Sei $\mathcal{V}$ ein Vektorraum und $\mathcal{B} := \{B \subset \mathcal{V} \mid B \text{ linear unabhängig }\}$. Dann ist $\mathcal{B}$ geordnet durch $B_1 \leq B_2 \Leftrightarrow B_1 \subset B_2$. Ist nun $\mathcal{K} \subset \mathcal{B}$ eine Kette, also eine total geordnete Menge, so ist die Vereinigung

$$K := \bigcup_{B \in \mathcal{K}} B$$

eine linear unabhängige Teilmenge von $\mathcal{V}$ (beachten Sie, linear unabhängig heißt dass jede endliche Linearkombination der 0 trivial ist) und somit ein Element von $\mathcal{B}$. Nach dem Zorn'schen Lemma hat also $\mathcal{B}$ maximale Elemente, also maximal linear unabhängige Teilmengen $X \subset \mathcal{V}$. Jedes solche X ist ein Erzeugendensystem, denn für $v \in \mathcal{V} \setminus \langle X \rangle$ ist $X \cup \{v\}$ linear unabhängig. □

Beachten Sie, dass man nicht unbedingt eine Basis von $\mathcal{V}$ angeben kann. Für $\mathcal{V} = K[[x]]$, $\mathcal{V} = \mathbb{R}^{[0,1]}$ oder $\mathcal{V} = \{f : [0,1] \to \mathbb{R} \mid f \text{ stetig }\}$ kann man keine solche Basis angeben, glaubt man an das Auswahlaxiom, was durchaus sinnvoll erscheint, so gibt es aber eine Basis.

Kapitel 4

Konstruktive Aspekte

4.1 Die Matrix einer linearen Abbildung

Lernziel: Matrix einer linearen Abbildung, typische Beispiele, Basistransformationen.

Wir wollen zuerst das Zusammenspiel der Begriffe "linear unabhängig" und "erzeugen" mit linearen Abbildungen näher beleuchten.

Satz 4.1.1. *Seien $\mathcal{V}$ und $\mathcal{W}$ K-Vektorräume und $X = (X_1, \ldots, X_n) \in \mathcal{V}^n, Y = (Y_1, \ldots, Y_n) \in \mathcal{W}^n$.*

(1) Ist X ein Erzeugendensystem von $\mathcal{V}$, so gibt es höchstens[1] *eine lineare Abbildung $\varphi : \mathcal{V} \to \mathcal{W}$ mit $\varphi \circ X = Y$, d.h. mit $\varphi(X_i) = Y_i$ für $i = 1, \ldots n$.*

(2) Ist X linear unabhängig (und $\mathcal{V}$ endlich erzeugt), so gibt es eine[2] *lineare Abbildung $\varphi : \mathcal{V} \to \mathcal{W}$ mit $\varphi \circ X = Y$.*

(3) Ist X eine Basis von $\mathcal{V}$, so gibt es genau eine lineare Abbildung $\varphi : \mathcal{V} \to \mathcal{W}$ mit $\varphi \circ X = Y$.

Beweis.

(1) Sei $V \in \mathcal{V}$. Da X ein Erzeugendensystem von $\mathcal{V}$ ist, gibt es $a \in K^n$ mit $V = a_1X_1 + \ldots + a_nX_n$. Für ein solches lineares φ gilt $\varphi(V) = a_1\varphi(X_1) + \ldots + a_n\varphi(X_n)$, also legen die Bilder$\varphi(X_i)$ wegen der Linearität die Abbildung φ eindeutig fest.

(2) Wir können X zu einer Basis ergänzen und nehmen daher oBdA an, dass X schon eine Basis ist. In diesem Fall ist

$$\varphi : \mathcal{V} \to \mathcal{W} : a_1X_1 + \ldots + a_nX_n \mapsto a_1Y_1 + \ldots + a_nY_n$$

eine wohldefinierte lineare Abbildung, denn die Darstellung der Elemente von $\mathcal{V}$ als Linearkombinationen von X ist eindeutig und die Linearität ist klar.
(Beachte, die ergänzten Vektoren zeigen, wieviel Freiheit man für die Festlegung von φ noch hat: Ist X keine Basis sondern nur linear unabhängig, so stehen die möglichen linearen Abbildungen φ in Bijektion mit $\mathcal{W}^k$, $k = \mathrm{Dim}(\mathcal{V}) - n$.)

(3) folgt sofort aus (1) und (2). □

Beispiel 4.1.2. Sei $\mathcal{V} = \mathbb{F}_2^2$, $\mathcal{W} = \mathbb{F}_2^3$,

$$X_1 := \begin{pmatrix} 1 \\ 0 \end{pmatrix}, X_2 := \begin{pmatrix} 0 \\ 1 \end{pmatrix}, X_3 := \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

[1] **Achtung**: So ein φ muss nicht unbedingt existieren.
[2] wie immer, im Sinne von „mindestens eine"

und $Y_i = e_i$ die Einheitsspalten. Dann ist $X = (X_1, X_2, X_3)$ ein Erzeugendensystem von $\mathcal{V}$, jedoch gibt es keine lineare Abbildung $\varphi : \mathcal{V} \to \mathcal{W}$ mit $\varphi(X_i) = e_i$, da jedes solche φ die Gleichung

$$0 = \varphi(0) = \varphi(X_1 + X_2 + X_3) = \varphi(X_1) + \varphi(X_2) + \varphi(X_3) = e_1 + e_2 + e_3 \neq 0$$

erfüllen muss. Hingegen kann man das Bild von je zwei der drei X_i's beliebig vorschreiben und erhält eine eindeutige lineare Abbildung φ. Z.B. $\varphi : \mathcal{V} \to \mathcal{W}, \varphi(X_1) = 0, \varphi(X_3) = e_1 + e_2 + e_3$, definiert eine eindeutige lineare Abbildung. Es gilt dann $\varphi(X_2) = e_1 + e_2 + e_3$.

Wir haben uns früher davon überzeugt, dass jede Matrix $A \in K^{m\times n}$ eine lineare Abbildung $\widetilde{A} : K^{n\times 1} \to K^{m\times 1}, X \mapsto AX$ induziert und, dass jede lineare Abbildung von $K^{n\times 1}$ nach $K^{m\times 1}$ auf diese Art mit einer eindeutig bestimmten Matrix $A \in K^{m\times n}$ zustande kommt. Dies können wir mit dem obigen Satz jetzt gut verstehen:
Die Spalten der Matrix A sind gerade die Bilder der Basisvektoren aus der Standardbasis von $K^{n\times 1}$ in $K^{m\times 1}$. Diesen Sachverhalt wollen wir zuerst einmal etwas formaler festhalten.

Definition 4.1.3. Seien $\mathcal{V}$ und $\mathcal{W}$ Vektorräume über dem Körper K. Dann bezeichnet

$$\mathrm{Hom}(\mathcal{V}, \mathcal{W}) := \{\varphi \mid \varphi : \mathcal{V} \to \mathcal{W}, \varphi \text{ linear } \} \leq \mathcal{W}^{\mathcal{V}}$$

den Vektorraum der linearen Abbildungen von $\mathcal{V}$ nach $\mathcal{W}$.

Man erinnert sich, dass $\mathcal{W}^{\mathcal{V}}$ mit der werteweisen Addition und entsprechenden Multiplikation mit Körperelementen ein Vektorraum ist, da der Wertebereich $\mathcal{W}$ diese Operationen zulässt und ein K-Vektorraum ist. Unsere obige Erinnerung lässt sich etwas verschärfen zu der Aussage:

Bemerkung 4.1.4.

$$K^{m\times n} \to \mathrm{Hom}(K^{n\times 1}, K^{m\times 1}) : A \mapsto \widetilde{A}$$

ist ein Isomorphismus von K-Vektorräumen.

Beweis. Dies war eine Übungsaufgabe! □

Die Umkehrabbildung der Linearkombinationsabbildung $\lambda_B : K^n \to \mathcal{V}$ für eine Basis $B \in \mathcal{V}^n$ nannten wir Koordinatenabbildung. Die Spalten-Variante dieser Umkehrabbildung halten wir nochmal in einer Definition fest:

Definition 4.1.5. Sei $\mathcal{V}$ ein n-dimensionaler K-Vektorraum mit Basis $B = (B_1, \ldots, B_n) \in \mathcal{V}^n$. Dann heißt der Isomorphismus

$$\kappa_B : \mathcal{V} \to K^{n\times 1} : X = a_1B_1 + \ldots + a_nB_n \mapsto {}^BX := \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$$

die **Koordinatenabbildung** (genauer Spaltenkoordinatenabbildung) von $\mathcal{V}$ bezüglich B.

Beispiel 4.1.6. Sei $\mathcal{V} = \mathbb{Q}[x]_{\mathrm{Grad}<4}$, $B = (1, x, x^2, x^3)$. Die Verschiebeabbildung

$$\varphi : \mathcal{V} \to \mathcal{V},\ p(x) \mapsto p(x+1)$$

ist linear. $\varphi(B_1) = 1 = B_1$, $\varphi(B_2) = B_1 + B_2$, $\varphi(B_3) = B_1 + 2B_2 + B_3$, $\varphi(B_4) = B_1 + 3B_2 + 3B_3 + B_4$. Bestimme $\varphi(p)$ für $p = x^2 + x + 1$. Dazu

$${}^B\varphi^B = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{pmatrix} =: A,\ {}^Bp = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}$$

Man berechnet ${}^B\varphi(p) = {}^B\varphi^B\, {}^Bp = (3,3,1,0)^{tr}$ und liest davon ab, dass $\varphi(p) = 3 + 3x + x^2$ ist.

Satz 4.1.7. *Sei $\mathcal{V}$ ein K-Vektorraum mit Basis $B = (B_1, \ldots, B_n) \in \mathcal{V}^n$ und $\mathcal{W}$ ein K-Vektorraum mit Basis $C = (C_1, \ldots, C_m) \in \mathcal{W}^m$. Es gilt:*

(1) Die Abbildung

$$ {}_C\Lambda_B : K^{m\times n} \to \mathrm{Hom}(\mathcal{V}, \mathcal{W}) : A \mapsto \kappa_C^{-1} \circ \widetilde{A} \circ \kappa_B $$

ist ein Isomorphismus von K-Vektorräumen.

(2) Sei die Umkehrabbildung von ${}_C\Lambda_B$ gegeben durch

$$ {}_C\Lambda_B^{-1} : \mathrm{Hom}(\mathcal{V}, \mathcal{W}) \to K^{m\times n} : \varphi \mapsto {}^C\varphi^B. $$

Man nennt ${}^C\varphi^B$ die **Matrix** *von φ bezüglich der Basen B und C. Per Definition gilt also*[3]

$$ \boxed{\varphi = \kappa_C^{-1} \circ \widetilde{{}^C\varphi^B} \circ \kappa_B} $$

oder äquivalent

$$ \boxed{\kappa_C \circ \varphi = \widetilde{{}^C\varphi^B} \circ \kappa_B} $$

und wir haben das kommutative Diagramm

$$ \begin{array}{ccc} \mathcal{V} & \xrightarrow{\varphi} & \mathcal{W} \\ \kappa_B \downarrow & & \downarrow \kappa_C \\ K^{n\times 1} & \xrightarrow[\widetilde{{}^C\varphi^B}]{} & K^{m\times 1} \end{array} $$

Insbesondere gilt für alle $V \in \mathcal{V}$:

$$ \boxed{{}^C(\varphi(V)) = {}^C\varphi^B \cdot {}^BV} $$

Insbesondere steht in der i-ten Spalte von ${}^C\varphi^B$ die Koordinatenspalte von $\varphi(B_i)$ bezüglich der Basis C.

Beweis.

(1) folgt aus Bemerkung 4.1.4, da die Komposition von Isomorphismen wieder ein Isomorphismus ist:

$$ \begin{array}{ccccccc} K^{m\times n} & \to & \mathrm{Hom}(K^{n\times 1}, K^{m\times 1}) & \xrightarrow{\kappa_C^{-1}\circ(-)} & \mathrm{Hom}(K^{n\times 1}, \mathcal{W}) & \xrightarrow{(-)\circ\kappa_B} & \mathrm{Hom}(\mathcal{V}, \mathcal{W}), \\ A & \mapsto & \widetilde{A} & \mapsto & \kappa_C^{-1} \circ \widetilde{A} & \mapsto & \kappa_C^{-1} \circ \widetilde{A} \circ \kappa_B \end{array} $$

(2) Die ersten beiden Formeln folgen aus (1) und der Definition von ${}^C\varphi^B$. Für die dritte Gleichheit benutze die Formel $\widetilde{A}(X) = AX$ für $X = \kappa_B(V) = {}^BV$. □

Beispiel 4.1.8. Die Abbildung $\mu : \mathbb{Q}[x]_{\mathrm{Grad}<4} \to \mathbb{Q}[x]/p\mathbb{Q}[x] : q \mapsto \overline{q} := q + p\mathbb{Q}[x]$ ist eine $\mathbb{Q}$-lineare Abbildung für jedes $p \in \mathbb{Q}[x]$. Wir wählen als Beispiel $p = 1 + x + x^2$. Ende Vorl. 18 12.12

Aufgabe: Bestimme die Matrix von μ bezüglich der Basen

$$ B := (1, x, x^2, x^3) \text{von } \mathbb{Q}[x]_{\mathrm{Grad}<4} \text{ und} $$
$$ C = ([1], [x]) := (1 + p\mathbb{Q}[x], x + p\mathbb{Q}[x]) \text{von } \mathbb{Q}[x]/p\mathbb{Q}[x]. $$

Diese Abbildung kodiert somit die Restbestimmung aller Polynome bis zum Grad drei (einschließlich) modulo p.

Lösung:

[3] die rechte Seite ist bloß nur ${}_C\Lambda_B({}_C\Lambda_B^{-1}(\varphi))$

- $\mu(1) = \overline{1} = 1 \cdot \overline{1} + 0 \cdot \overline{x}$, womit wir die erste Spalte haben.
- $\mu(x) = \overline{x} = 0 \cdot \overline{1} + 1 \cdot \overline{x}$, womit wir die zweite Spalte haben.
- $\mu(x^2) = \overline{x^2} = -1 \cdot \overline{1} - 1 \cdot \overline{x}$, womit wir die dritte Spalte haben.
- $\mu(x^3) = \overline{x^3} = -1 \cdot \overline{x} - 1 \cdot \overline{x^2} = -1 \cdot \overline{x} - (-1 \cdot \overline{1} - 1 \cdot \overline{x}) = 1 \cdot \overline{1} + 0 \cdot \overline{x}$, was die vierte Spalte ergibt.

Insgesamt haben wir also

$$ {}^C\mu^B = \begin{pmatrix} 1 & 0 & -1 & 1 \\ 0 & 1 & -1 & 0 \end{pmatrix} $$

Diese Matrix kodiert somit die Restbestimmung aller Polynome bis zum Grad drei (einschließlich) modulo p. Z.B. $7x^3 + 3x^2 + 2x + 10$ modulo p berechnet man nun wegen

$$ \begin{pmatrix} 1 & 0 & -1 & 1 \\ 0 & 1 & -1 & 0 \end{pmatrix} \begin{pmatrix} 10 \\ 2 \\ 3 \\ 7 \end{pmatrix} = \begin{pmatrix} 14 \\ -1 \end{pmatrix} $$

zu $14 - x$.

Man kann übrigens die Idee des letzten Beispiels benutzen, um bequem im Restklassenring $K[x]/pK[x]$ zu rechnen, was wir aber jetzt nicht verfolgen wollen.

Hier ist eine sehr wichtige Folgerung aus unserem Satz. (Vgl. auch 1.3.1 aus dem ersten Kapitel.)

Satz 4.1.9. *Seien $\varphi : \mathcal{V} \to \mathcal{W}$ und $\psi : \mathcal{W} \to \mathcal{U}$ K-lineare Abbildungen und $\mathcal{V}, \mathcal{W}, \mathcal{U}$ sollen die Basen $B \in \mathcal{V}^n, C \in \mathcal{W}^m, D \in \mathcal{U}^p$ haben. Dann gilt:*

$$ \underbrace{{}^D(\psi \circ \varphi)^B}_{p \times n} = \underbrace{{}^D\psi^C}_{p \times m} \cdot \underbrace{{}^C\varphi^B}_{m \times n}. $$

Beweis. Sei $V \in \mathcal{V}$ beliebig. Dann gilt nach dem letzten Satz 4.1.7 einerseits:

$$ {}^D((\psi \circ \varphi)(V)) = {}^D(\psi \circ \varphi)^B \cdot {}^BV $$

und andererseits

$$ {}^D((\psi \circ \varphi)(V)) = {}^D(\psi(\varphi(V))) = {}^D\psi^C \cdot {}^C(\varphi(V)) = {}^D\psi^C \cdot {}^C\varphi^B \cdot {}^BV $$

Lassen wir nun V die Basis B durchlaufen, so folgt

$$ {}^D(\psi \circ \varphi)^B = {}^D\psi^C \cdot {}^C\varphi^B. $$

□

Einen Spezialfall sollte man besonders herausstellen: den Basiswechsel.

Folgerung 4.1.10.

(1) Sind $B, B' \in \mathcal{V}^n$ Basen von $\mathcal{V}$, so heißt ${}^B\,\mathrm{id}_{\mathcal{V}}^{B'}$ die Matrix der **Basistransformation** *und es gilt*

$$ {}^{B'}\,\mathrm{id}_{\mathcal{V}}^{B} = ({}^B\,\mathrm{id}_{\mathcal{V}}^{B'})^{-1}. $$

(2) Sind $C, C' \in \mathcal{W}^n$ Basen von $\mathcal{W}$ und ist $\varphi : \mathcal{V} \to \mathcal{W}$ linear, so gilt

$$ {}^{C'}\varphi^{B'} = {}^{C'}\,\mathrm{id}_{\mathcal{W}}^{C} \cdot {}^C\varphi^B \cdot {}^B\,\mathrm{id}_{\mathcal{V}}^{B'} $$

Beweis.

(1) Aus Satz 4.1.9 folgt
$$ {}^{B'}\mathrm{id}_{\mathcal{V}}^{B}\,{}^{B}\mathrm{id}_{\mathcal{V}}^{B'} = {}^{B'}\mathrm{id}_{\mathcal{V}}^{B'} = I_n. $$

(2) Wende Satz 4.1.9 zweimal an. □

Bemerkung 4.1.11. Sind $B = (B_1, \ldots, B_n)$ und $B' = (B'_1, \ldots, B'_n)$ Basen des K-Vektorraumes $\mathcal{V}$, so gilt:
$$ {}^{B}\mathrm{id}_{\mathcal{V}}^{B'} = {}^{B}\tau^{B}, $$
wobei $\tau : \mathcal{V} \to \mathcal{V}$, $B_i \mapsto B'_i$ die Abbildung ist, die B auf B' abbildet. Die linke Seite nennt man die passive Interpretation des Basiswechsels un die rechte Seite die aktive.

Beispiel 4.1.12. Der Vektorraum $\mathcal{V} := \mathbb{Q}[x]_{\mathrm{Grad}<5}$ hat die Basen $B = (1, x, x^2, x^3, x^4)$ und $B' = (1, x+1, (x+1)^2, (x+1)^3, (x+1)^4)$. Der binomische Lehrsatz liefert sofort das transponierte PASCALsche Dreieck:
$$ {}^{B}\mathrm{id}_{\mathcal{V}}^{B'} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 & 4 \\ 0 & 0 & 1 & 3 & 6 \\ 0 & 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} = {}^{B}\tau^{B} $$
für die lineare Abbildung
$$ \tau : \mathcal{V} \to \mathcal{V} : p(x) \mapsto p(x+1). $$
Für die inverse Matrix erhält man
$$ {}^{B'}\mathrm{id}_{\mathcal{V}}^{B} = \begin{pmatrix} 1 & -1 & 1 & -1 & 1 \\ 0 & 1 & -2 & 3 & -4 \\ 0 & 0 & 1 & -3 & 6 \\ 0 & 0 & 0 & 1 & -4 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} = ({}^{B}\mathrm{id}_{\mathcal{V}}^{B'})^{-1} = {}^{B}(\tau^{-1})^{B} $$
wobei $\tau^{-1}(p) = p(x-1)$ ist.

4.2 Matrizen und Teilräume: Zeilenraum und Spaltenraum

Lernziel: Algorithmen zur Herstellung von Basen, Rang einer Matrix, genaue Analyse des GAUSSalgorithmus, Algorithmen zur Berechnung von Schnitten von Teilräumen.

Der Satz von STEINITZ ist konstruktiv und hat eine innere Verwandtschaft mit dem GAUSSschen Algorithmus. Wir wollen jetzt die abstrakten Dimensions- und Existenzaussagen des letzten Kapitels in konkrete Algorithmen umwandeln. Wesentlich ist, dass wir explizit in unserem Vektorraum rechnen können. Insbesondere wollen wir jetzt eine konstruktive Behandlung von endlichen Erzeugendensystemen vornehmen.

Zur Erinnerung:
$$ \mathrm{GL}_m(K) := \{g \in K^{m\times m} \mid \text{ es existiert ein } h \in K^{m\times m} \text{ mit } gh = I_m\} $$
operiert auf $K^{m\times n}$ durch Linksmultiplikation. Früher hatten wir gesehen, dass die Linksmultiplikation mit sogenannten elementaren Matrizen aus $\mathrm{GL}_m(K)$ den elementaren Zeilenumformungen aus dem GAUSSalgorithmus entsprechen. Neue Interpretation dieser Umformungen: Wechsel des Erzeugendensystems des Zeilenraumes der Matrix.

Definition 4.2.1. Sei $A \in K^{m\times n}$ und $\mathcal{V}$ ein K-Vektorraum.

(1) Es bezeichnet
$$\mathrm{TR}(\mathcal{V}) := \{\mathcal{T} \mid \mathcal{T} \leq \mathcal{V}\}$$
die Menge aller Teilräume von $\mathcal{V}$.

(2) $Z(A) := \langle A_{1,-}, A_{2,-}, \ldots, A_{m,-}\rangle \leq K^{1\times n}$ heißt der **Zeilenraum** von A und seine Dimension der **Zeilenrang** von A.

(3) $S(A) := \langle A_{-,1}, A_{-,2}, \ldots, A_{-,n}\rangle \leq K^{m\times 1}$ heißt der **Spaltenraum** von A und seine Dimension der **Spaltenrang** von A.

Klar: Z ist eine Abbildung $K^{m\times n} \to \mathrm{TR}(K^{1\times n})$, und der Zeilenrang die Komposition von Z mit $\mathrm{Dim} : \mathrm{TR}(K^{1\times n}) \to \mathbb{Z}_{\geq 0}$. Entsprechend für Spaltenraum und Spaltenrang mit dem Zusatz $S(A) = \mathrm{Bild}(\widetilde{A})$.

Bemerkung 4.2.2. Sei $A \in K^{m\times n}$, $g \in \mathrm{GL}_m(K)$ und $h \in \mathrm{GL}_n(K)$.

(1) Dann haben A und gA denselben Zeilenraum und somit auch denselben Zeilenrang. Insbesondere produziert der GAUSSsche Algorithmus eine Basis des Zeilenraumes.

(2) A und gA haben denselben Spaltenrang.

(3) A und Ah haben denselben Spaltenraum und somit denselben Spaltenrang.

(4) A und Ah haben denselben Zeilenrang.

Beweis.

(1) Klar.

(2) $\widetilde{g} : K^{m\times 1} \to K^{m\times 1}$ ist ein Automorphismus (d.h. Endomorphismus + Isomorphismus) und seine Einschränkung auf den Spaltenraum von A liefert einen Isomorphismus zwischen $S(A)$ und $S(gA)$, also haben beide dieselbe Dimension.

(3) Klar.

(4) so wie (2). □

Folgerung 4.2.3. *Für jede Matrix $A \in K^{m\times n}$ über einem Körper K sind Zeilen- und Spaltenrang gleich. Sie werden mit* $\mathrm{Rang}(A) =$ **Rang** *von A bezeichnet.*

Beweis. Da $Z(A) = Z(gA)$ für alle $g \in \mathrm{GL}_n(K)$, kann man eine Basis des Zeilenraumes nach Anwenden des GAUSSalgorithmus aus der strikten Zeilenstufenform $B = gA$ von A direkt ablesen, nämlich die Nicht-Null-Zeilen von B. Die Dimension von $Z(B)$ ist somit die Anzahl der Nicht-Null-Zeilen von B, also die Anzahl der Stufenindizes von B. Aber diese Anzahl ist ebenfalls die Dimension des Spaltenraums von B, da die Spalten an den Stufenindizes offensichtlich eine Basis des Spaltenraumes von B bilden. Diese Dimension ist aber auch der Spaltenrang von A nach Bemerkung 4.2.2.(2). □

Beispiel 4.2.4.

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ -1 & -3 & -4 & 5 \\ 0 & -1 & -1 & 9 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 1 & 22 \\ 0 & 1 & 1 & -9 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Also ist $\mathrm{Rang}(A) = 2$.

$$\mathcal{L}(Ax = 0) = \mathrm{Kern}(\widetilde{A}) = \langle \begin{pmatrix} -1 \\ -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -22 \\ 9 \\ 0 \\ 1 \end{pmatrix} \rangle$$

$\mathrm{Dim}(\mathrm{Kern}(\widetilde{A})) = 2 = 4 - \mathrm{Rang}(A)$.

Folgerung 4.2.5. *Für die Matrix $A \in K^{m\times n}$ ist $\mathrm{Kern}(\widetilde{A})$ der Lösungsraum des linearen Gleichungssystems $Ax = 0, x \in K^{n\times 1}$. Es gilt:*

$$\boxed{\mathrm{Dim}(\textit{Lösungsraum}) = n - \mathrm{Rang}(A).}$$

Beweis. Es gilt wegen $\mathrm{Bild}(\widetilde{A}) = S(A)$

$$\begin{aligned} \mathrm{Dim}(\text{Lösungsraum}) &= \mathrm{Dim}(\mathrm{Kern}(\widetilde{A})) \\ &= n - \mathrm{Dim}(\mathrm{Bild}(\widetilde{A})) \\ &= n - \mathrm{Rang}(A). \end{aligned}$$

□

Wir wollen jetzt den GAUSSalgorithmus zur Herstellung der strikten Stufengestalt einer Matrix als Wechsel des Erzeugendensystems des Zeilenraumes interpretieren und die nichtverschwindenden Zeilen der Matrix in strikter Stufengestalt als eine normierte Basis. Ende Vorl. 19 14.12

Satz 4.2.6.

(1) $\mathrm{GL}_m(K)$ *operiert auf* $K^{m\times n}$ *durch Linksmultiplikation:*

$$\mathrm{GL}_m(K) \times K^{m\times n} \to K^{m\times n} : (g, A) \mapsto gA.$$

(2) Der Zeilenraum $Z : K^{m\times n} \to \mathrm{TR}(K^{1\times n})$ *ist eine trennende Invariante, d.h.* $A, B \in K^{m\times n}$ *sind genau dann in derselben Bahn, wenn* $Z(A) = Z(B)$ *gilt.*

(3) Jede Bahn enthält genau eine Matrix in strikter Stufengestalt.

Beweis.

(1) & (3) kennen wir bereits.

(2) Die Gleichheit $Z(A) = Z(gA)$ für $A \in K^{m\times n}$ und $g \in \mathrm{GL}_m(K)$ haben wir schon in Bemerkung 4.2.2 gesehen, d.h. der Zeilenraum ist eine Invariante, ebenso wie der Zeilenrang.
Wir zeigen zunächst, dass die Stufenindizes (je)der zu $A \in K^{m\times n}$ gehörigen Matrix in strikter Stufengestalt bereits aus $Z(A)$ abgelesen werden können, also strukturell bestimmt sind: Definiere für $1 \le d \le n$

$$\pi_d : Z(A) \to K^{1\times d} : (x_1, \ldots, x_n) \mapsto (x_1, \ldots, x_d)$$

und π_0 als die Nullabbildung auf $Z(A)$. Dann ist sofort klar: d ist Stufenindex von gA genau dann, wenn

$$\mathrm{Dim}(\mathrm{Bild}(\pi_d)) > \mathrm{Dim}(\mathrm{Bild}(\pi_{d-1})).$$

Seien nun $1 \le s(1) < \ldots < s(r) \le n$ die Stufenindizes die durch $Z(A)$ festgelegt sind. Für den Rest des Beweises nehmen wir Œ an, dass wir es nur mit Matrizen vom Zeilenrang m zu tun haben, also $r = m$. Eine zu A gehörige Matrix in strikter Stufenform ist dann gegeben durch $G^{-1}A$, wobei $G \in K^{m\times m}$ die Stufenspalten von

A als Spalten hat: $G_{-,i} := A_{-,s(i)}$. (Übung: Warum ist G invertierbar?) Jede andere Matrix $B \in K^{m\times n}$ mit $Z(A) = Z(B)$ ist dann gegeben durch $HG^{-1}A$, wo $H \in K^{m\times m}$ mit Spalten $H_{-,i} := B_{-,s(i)}$ gegeben ist. Damit folgt, dass $G^{-1}A$ die einzige Matrix in strikter Stufengestalt unter den Matrizen mit Zeilenraum $Z(A)$ ist und auch, dass Z eine trennende Invariante der $\mathrm{GL}_m(K)$-Operation auf $K^{m\times n}$ ist. □

Folgerung 4.2.7. *Jeder k-dimensionale Teilraum von $K^{1\times n}$ mit $k > 0$ hat eine eindeutige* **Standardbasis** *$(Z_1, \ldots, Z_k)$, die dadurch gekennzeichnet ist, dass die Matrix $Z \in K^{k\times n}$ mit $Z_{i,-} = Z_i$ in strikter Stufengestalt ist.*

Beispiel 4.2.8. Die 2-dimensionalen Teilräume von $K^{1\times 3}$ sind genau die Zeilenräume von:

$$\begin{pmatrix} 1 & 0 & a \\ 0 & 1 & b \end{pmatrix}, \begin{pmatrix} 1 & c & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

mit $a, b, c \in K$.

Entsprechend sind die verschiedenen 2-dimensionalen Teilräume eines Raums $\mathcal{V}$ mit Basis (B_1, B_2, B_3) gegeben durch

$$\langle B_1 + aB_3, B_2 + bB_3\rangle, \langle B_1 + cB_2, B_3\rangle, \langle B_2, B_3\rangle.$$

wobei a, b, c die Elemente von K durchläuft. Hat K also q Elemente, so besitzt $\mathcal{V}$ genau q^2+q+1 2-dimensionale Teilräume. Dies ist auch die Anzahl der 1-dimensionalen Teilräume von $\mathcal{V}$. Warum?

Definition 4.2.9. Sei $\varphi : \mathcal{V} \to \mathcal{W}$ eine lineare Abbildung. Definiere

$$\mathrm{Rang}(\varphi) \coloneqq \mathrm{Dim}(\mathrm{Bild}(\varphi)).$$

Folgerung 4.2.10. *Sei $\varphi : \mathcal{V} \to \mathcal{W}$ eine lineare Abbildung von endlich dimensionalen (=endlich erzeugte) K-Vektorräumen, B eine Basis von $\mathcal{V}$ und C eine Basis von $\mathcal{W}$. Dann ist*

$$\mathrm{Rang}(\varphi) = \mathrm{Rang}({}^C\varphi^B).$$

Beweis. Übung. □

Folgerung 4.2.11. *Sei $\varphi : \mathcal{V} \to \mathcal{W}$ eine lineare Abbildung von endlich dimensionalen (=endlich erzeugte) K-Vektorräumen. Dann existiert eine Basis B von $\mathcal{V}$ und eine Basis C von $\mathcal{W}$, derart dass*

$${}^C\varphi^B = \begin{pmatrix} 1 & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ 0 & 1 & 0 & \cdots & \cdots & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \cdots & \cdots & 0 \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \end{pmatrix},$$

mit genau $\mathrm{Rang}(\varphi)$ Einsen auf der Diagonalen. Sprich, es gilt[4]

$$\begin{aligned} \varphi(B_i) &= C_i && \text{für } i = 1, \ldots \mathrm{Rang}(\varphi), \\ \varphi(B_i) &= 0 && \text{für } i = \mathrm{Rang}(\varphi) + 1, \ldots, \mathrm{Dim}(\mathcal{V}). \end{aligned}$$

[4] Etwas derartig einfaches gibt es bei (endlichen) Mengen und Abbildungen dazwischen nicht!

Beweis. Starte mit zwei beliebigen Basen $B' \in \mathcal{V}^n$ und $C' \in \mathcal{W}^m$. Berechne die Matrix ${}^{C'}\varphi^{B'}$ und bringe sie mit Hilfe des üblichen zeilenorientierten GAUSS-Algorithmus auf strikte Zeilenstufenform und mit dem völlig analogen spaltenorientierten Gauss-Algorithmus[5] schließlich auf die gewünschte Form, nennen wir sie A. Durch die Buchführung bei den GAUSS-Algorithmen bestimmen wir also Matrizen $g \in \mathrm{GL}_n(K)$ und $h \in \mathrm{GL}_m(K)$ mit $g\left({}^{C'}\varphi^{B'}\right)h = A$. Nun interpretiere g als die Basiswechselmatrix ${}^{B}\mathrm{id}_{\mathcal{V}}^{B'}$ und h als die Basiswechselmatrix ${}^{C'}\mathrm{id}_{\mathcal{W}}^{C}$ und lese daraus B und C ab. □

Folgerung 4.2.12. *Die Gruppe* $\mathrm{GL}_m(K) \times \mathrm{GL}_n(K)$ *operiert auf* $K^{m\times n}$ *durch*

$$(\mathrm{GL}_m(K) \times \mathrm{GL}_n(K)) \times K^{m\times n} \to K^{m\times n},\ ((g,h),A) \mapsto gAh^{-1}.$$

Matrizen in derselben Bahn heißen **äquivalent**. *Der Rang*[6] *ist eine trennende Invariante dieser Operation, d.h. Matrizen in* $K^{m\times n}$ *sind genau dann äquivalent, wenn sie den gleichen Rang haben.*

[5] Z.B. transponiere die Eingabematrix, wende den zeilenorientierten GAUSS-Algorithmus an und transponiere anschließend das Ergebnis zurück.

[6] Genauer: $\mathrm{Rang} : K^{m\times n} \to \mathbb{Z}_{\geq 0}$

Kapitel 5

Endomorphismen

5.1 Der Endomorphismenring

Lernziel: Matrizen von Endomorphismen, Ähnlichkeit von Matrizen, Einsetzungshomomorphismus, Minimalpolynom und seine Berechnung, Begleitmatrizen.

Definition 5.1.1. Sei $\mathcal{V}$ ein K-Vektorraum . Dann heißt

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

zusammen mit der Addition von linearen Abbildungen und der Komposition der **Endomorphismenring** von $\mathcal{V}$.

Die Endomorphismen von $\mathcal{V}$ bilden einen Ring mit $\mathrm{id}_\mathcal{V}$ als Eins, sogar eine K-Algebra, die für $\mathrm{Dim}\,\mathcal{V} > 1$ nicht-kommutativ ist. Die invertierbaren Elemente bilden die Einheitengruppe des Ringes, die wir bereits früher als generelle lineare Gruppe $\mathrm{GL}(\mathcal{V})$ bezeichnet haben. Was bewirkt die Festlegung einer Basis?

Bemerkung 5.1.2. Sei $B \in \mathcal{V}^n$ eine Basis von $\mathcal{V}$. Dann ist

$$\mathrm{End}(\mathcal{V}) \to K^{n\times n},\ \alpha \mapsto {}^B\alpha^B$$

ein K-Algebrenisomorphismus.

Unser Thema ist weniger die algebraische Struktur von $\mathrm{End}(\mathcal{V})$ als Ring sondern vielmehr, wie wir vorgegebene Endomorphismen besser verstehen können. In Folgerung 4.2.11 haben wir uns dafür erlaubt zwei Basen zu wählen, eine im Definitionsbereich und eine im Wertebereich. Da es sich hier um Endomorphismen handelt, möchten wir gerne nur mit einer Basis B auskommen, und dass weiterhin die Abbildungsmatrix eine besonders einfache Gestalt annimmt. Am einfachsten ist die Diagonalgestalt. Zuerst rekapitulieren wir, wie ein Basiswechsel sich auswirkt:

Bemerkung 5.1.3.

(1) $\alpha \in \mathrm{End}(\mathcal{V})$ und $B, B' \in \mathcal{V}^n$ Basen von $\mathcal{V}$. Es gilt

$${}^{B'}\alpha^{B'} = ({}^B\,\mathrm{id}_\mathcal{V}^{B'})^{-1} \cdot {}^B\alpha^B \cdot {}^B\,\mathrm{id}_\mathcal{V}^{B'} = {}^{B'}\,\mathrm{id}_\mathcal{V}^B \cdot {}^B\alpha^B \cdot ({}^{B'}\,\mathrm{id}_\mathcal{V}^B)^{-1}.$$

(2) Die Gruppe $\mathrm{GL}_n(K)$ operiert auf $K^{n\times n}$ durch

$$\mathrm{GL}_n(K) \times K^{n\times n} \to K^{n\times n},\ (g, A) \mapsto gAg^{-1}.$$

Matrizen in derselben Bahn heißen **ähnlich**. Die Operation heißt auch **Konjugationsoperation** von $\mathrm{GL}_n(K)$ auf $K^{n\times n}$.

Zwei Matrizen $A, A' \in K^{n\times n}$ sind also genau dann **ähnlich**, wenn es eine Matrix $C \in \mathrm{GL}_n(K)$ existiert, mit $A' = C^{-1}AC$. Die Äquivalenzklassen heißen **Ähnlichkeitsklassen**. Insbesondere sind ${}^B\alpha^B$ und ${}^{B'}\alpha^{B'}$ ähnlich.

Es ist recht schwierig, trennende Invarianten oder Normalformen für die Ähnlichkeitsklassen anzugeben. Dies wird für algebraisch abgeschlossene Körper in einem späteren Kapitel geschehen und für allgemeine Körper wahrscheinlich erst in der Algebravorlesung. Aber wir können jetzt schon einige erste Schritte unternehmen. Klar ist, dass ähnliche Matrizen denselben Rang haben. Eine zweite einfache Invariante ist die Spur.

Definition 5.1.4.

(1) Für $A \in K^{n\times n}$ heißt $\mathrm{Spur}(A) := \sum_{i=1}^{n} A_{ii}$ die **Spur** der Matrix A.

(2) Sei $\mathcal{V}$ ein endlich dimensionaler K-Vektorraum. Für $\alpha \in \mathrm{End}(\mathcal{V})$ definiert man die **Spur** von α als
$$\mathrm{Spur}(\alpha) := \mathrm{Spur}({}^B\alpha^B)$$
für irgendeine Basis $B \in \mathcal{V}^n$ von $\mathcal{V}$.

Bemerkung 5.1.5.

(1) Für $A \in K^{m\times n}$ und $B \in K^{n\times m}$ gilt:
$$\mathrm{Spur}(AB) = \sum_{i,j} A_{ij}B_{ji} = \mathrm{Spur}(BA).$$

(2) Für $A \in K^{n\times n}$ und $g \in \mathrm{GL}_n(K)$ gilt
$$\mathrm{Spur}(g^{-1}Ag) = \mathrm{Spur}(A).$$

(3) $\mathrm{Spur}(\alpha)$ für $\alpha \in \mathrm{End}(\mathcal{V})$ ist wohldefiniert.

Beweis.

(1) Klar.

(2) Aus (1): $\mathrm{Spur}(g^{-1}Ag) = \mathrm{Spur}(gg^{-1}A) = \mathrm{Spur}(A)$.

(3) Aus (2) und 5.1.3. □

5.2 Das Minimalpolynom

Hier kommt eine brauchbarere Invariante, die auf einem Test der linearen Abhängigkeiten der Potenzen einer linearen Abbildung beruht.

Beispiel 5.2.1. Wir halten folgende Spezialfälle von Einsetzhomomorphismen fest:

(1) Für $A \in K^{n\times n}$ und $p = p(x) = a_0 + a_1x + \ldots + a_dx^d \in K[x]$ sei $p(A)$ definiert als $p(A) := a_0I_n + a_1A + \ldots + a_dA^d$. Weiter heißt
$$\varepsilon_A : K[x] \to K^{n\times n} : p \to p(A)$$
der **Einsetzungshomomorphismus** zu A.

(2) Für $\alpha \in \mathrm{End}(\mathcal{V})$ und $p = p(x) = a_0 + a_1 x + \ldots + a_d x^d \in K[x]$ sei $p(\alpha)$ definiert als $p(\alpha) := a_0 \,\mathrm{id}_{\mathcal{V}} + a_1\alpha + \ldots + a_d\alpha^d$. Weiter heißt

$$\varepsilon_\alpha : K[x] \to \mathrm{End}(\mathcal{V}) : p \to p(\alpha)$$

der **Einsetzungshomomorphismus** zu α.

Ende Vorl. 20 19.12

Erinnerung: Ist $A \in K^{n\times n}$ so ist ε_A ein K-Algebrenhomomorphismus, d.h. für $p, q \in K[x]$ und $a, b \in K$ ist

$$\varepsilon_A(ap + bq) = a\varepsilon_A(p) + b\varepsilon_A(q),\ \varepsilon_A(pq) = \varepsilon_A(p)\varepsilon_A(q).$$

Ist $\alpha \in \mathrm{End}(\mathcal{V})$ so ist ε_α ein K-Algebrenhomomorphismus.

Lemma 5.2.2.

(1) Ist $A \in K^{n\times n}$, so gibt es genau ein normiertes Polynom $\mu_A \in K[x]$ mit $\mathrm{Kern}(\varepsilon_A) = \mu_A K[x]$. Dieses Polynom heißt das **Minimalpolynom** *von A.*

(2) Ist $\mathcal{V}$ ein endlich dimensionaler K-Vektorraum und $\alpha \in \mathrm{End}(\mathcal{V})$, so gibt es genau ein normiertes Polynom $\mu_\alpha \in K[x]$ mit $\mathrm{Kern}(\varepsilon_\alpha) = \mu_\alpha K[x]$. Dieses Polynom heißt das **Minimalpolynom** *von α.*

Beweis. (2) geht genauso wie (1)

(1) Sei $s \in \mathbb{N}$ minimal, so dass $(I_n, A, \ldots, A^s)$ linear abhängig im K-Vektorraum $K^{n\times n}$ ist und $a_0, \ldots, a_{s-1} \in K$ mit $A^s + a_{s-1}A^{s-1} + \ldots + a_0 I_n = 0$. Setze $a_s = 1$ und $\mu_A = \sum_{i=0}^{s} a_i x^i \in K[x]$.
Behauptung: $\mu_A K[x] = \mathrm{Kern}(\varepsilon_A)$.

$\subseteq$: klar, da $\mu_A(A) = 0$.

$\supseteq$: Sei $0 \neq q \in \mathrm{Kern}(\varepsilon_A)$. Dann gibt es $a, b \in K[x]$ mit $q = a\mu_A + b$ mit $\mathrm{Grad}(b) < \mathrm{Grad}(\mu_A) = s$. Es gilt jedoch

$$0 = q(A) = a(A)\mu_A(A) + b(A) = 0 + b(A) \text{ also auch } b(A) = 0.$$

Wegen der Minimalität von $s = \mathrm{Grad}(\mu_A)$ folgt also $b = 0$. Somit ist $q \in \mu_A K[x]$. □

Übung 5.2.1. Zeigen Sie: Für $A \in K^{n\times n}$, $g \in \mathrm{GL}_n(K)$ und $p \in K[x]$ ist

$$p(g^{-1}Ag) = g^{-1}p(A)g.$$

Bemerkung 5.2.3.

(1) Der Grad des Minimalpolynoms von A ist das kleinste $s \in \mathbb{N}$ mit $(I_n, A, \ldots, A^s)$ linear abhängig. Insbesondere ist das Minimalpolynom wohldefiniert.

(2) Das Minimalpolynom, genauer

$$\mu : K^{n\times n} \to K[x] : A \mapsto \mu_A(x)$$

ist eine Invariante der Ähnlichkeitsklassen in $K^{n\times n}$, sprich zwei ähnliche Matrizen haben das gleiche Minimalpolynom.

(3) Sei $\alpha \in \mathrm{End}(\mathcal{V})$, $B \in \mathcal{V}^n$ eine Basis von $\mathcal{V}$ und $A = {}^B\alpha^B \in K^{n\times n}$. Dann sind die Minimalpolynome von α und A gleich:

$$\mu_\alpha(x) = \mu_A(x).$$

Beweis. Übung. □

Beispiel 5.2.4.

(1) Sei $A = 0 \in K^{n\times n}$ die Nullmatrix. Dann gilt $\mu_0(x) = x$. Entsprechend $\mu_{0_\mathcal{V}} = x$, wobei wir mit $0_\mathcal{V}$ den Nullendomorphismus bezeichnet.

(2) Sei $A = I_n \in K^{n\times n}$. Dann gilt $\mu_{I_n}(x) = x - 1$ für $n > 0$. Entsprechend $\mu_{\mathrm{id}_\mathcal{V}} = x - 1$ falls $\mathcal{V} \neq \{0\}$.

(3) Sei $A = \begin{pmatrix} 2 & 0 \\ 0 & -5 \end{pmatrix} \in \mathbb{Q}^{2\times 2}$. Dann gilt $\mu_A = (x-2)(x+5)$.

(4) Sei $\mathcal{V} = \langle \sin, \cos\rangle \leq \mathbb{R}^\mathbb{R}$ der von Sinus und Cosinus erzeugte Teilraum von $\mathbb{R}^\mathbb{R}$ und $\partial \in \mathrm{End}(\mathcal{V})$ die Ableitung. Dann gilt $\partial(\sin) = \sin' = \cos$ linear unabhängig von $\sin$, und $\partial^2(\sin) = -\sin$ und $\partial^2(\cos) = -\cos$, also $\mu_\partial(x) = x^2+1$. Man beachte, wir erhalten dadurch eine neue Realisierung der komplexen Zahlen $\mathbb{C} \cong \mathbb{R}[x]/(x^2+1)\mathbb{R}[x]$ als den Teilring von $\mathrm{End}(\mathcal{V}) \cong \mathbb{R}^{2\times 2}$, der als $\mathbb{R}$-Vektorraum von $\mathrm{id}_\mathcal{V}$ und ∂ erzeugt wird.

(5) Im Allgemeinen gilt für $A \in K^{n\times n}$:

$$K^{n\times n} \geq K[A] := \langle I_n, A, A^2, \ldots, A^{s-1}\rangle = \mathrm{Bild}(\varepsilon_A) \cong K[x]/\operatorname{Kern}(\varepsilon_A) = K[x]/\mu_A K[x]$$

als K-Algebra.

Lemma 5.2.5. *Sei $\alpha \in \mathrm{End}(\mathcal{V})$ und $\mathcal{U} \leq \mathcal{V}$ ein α-invarianter Teilraum, d.h. $\alpha(\mathcal{U}) \subseteq \mathcal{U}$. Dann definiert α zwei lineare Abbildungen:*

$$\beta := \alpha_{|\mathcal{U}} \in \mathrm{End}(\mathcal{U}) \text{ und } \gamma \in \mathrm{End}(\mathcal{V}/\mathcal{U}), \gamma(X+\mathcal{U}) := \alpha(X)+\mathcal{U}.$$

Es gilt

$$kgV(\mu_\beta, \mu_\gamma) \mid \mu_\alpha \mid \mu_\beta\mu_\gamma,$$

d.h. $kgV(\mu_\beta, \mu_\gamma)$ teilt μ_α und μ_α teilt $\mu_\beta\mu_\gamma$.

Beweis. Es ist $\mu_\beta K[X] = \operatorname{Kern}(\varepsilon_\beta)$, insbesondere ist jedes Polynom $f \in K[X]$ mit $f(\beta) = 0$ durch μ_β teilbar. Jetzt genügt es zu beobachten, dass $\mu_\alpha(\beta) = 0$ da

$$\mu_\alpha(\beta) = \mu_\alpha(\alpha)_{|\mathcal{U}}.$$

Also gilt μ_β teilt μ_α. Es ist leicht einzusehen, dass γ wohldefiniert ist und $\mu_\alpha(\gamma) = 0$ gilt. Daraus ergibt sich die erste Teilbarkeitsrelation.
Für die zweite Teilbarkeit sei $X \in \mathcal{V}$. Dann gilt

$$(\mu_\beta\mu_\gamma(\alpha))(X) = \mu_\beta(\alpha)(\mu_\gamma(\alpha)(X)) = \mu_\beta(\alpha)(Y) = 0$$

wobei $Y = \mu_\gamma(\alpha)(X) \in \mathcal{U}$ ist und daher $\mu_\beta(\alpha)(Y) = \mu_\beta(\beta)(Y) = 0$. □

Bemerkung 5.2.6. Lemma 5.2.5 liest sich für Matrizen wie folgt: Sei $A = \begin{pmatrix} B & \star \\ 0 & C \end{pmatrix} \in K^{n\times n}$ mit quadratischen Matrizen B und C. Dann gilt

$$\mathrm{kgV}(\mu_C, \mu_B) \mid \mu_A \mid \mu_B\mu_C,$$

d.h. $\mathrm{kgV}(\mu_C, \mu_B)$ teilt μ_A und μ_A teilt $\mu_B\mu_C$.

Beispiel 5.2.7. Sei $\mathcal{V}$ endlich erzeugter K-Vektorraum und $\alpha \in \mathrm{End}(\mathcal{V})$. Wähle $0 \neq V \in \mathcal{V}$ und schreibe die erste lineare Abhängigkeit von $(V, \alpha(V), \alpha^2(V), \ldots, \alpha^{k-1}(V), \alpha^k(V)) \in \mathcal{V}^{k+1}$ mit k minimal als Polynom

$$a_0 V + a_1\alpha(V) + \ldots + a_{k-1}\alpha^{k-1}(V) + 1\alpha^k(V) = 0$$

und definiere $q := a_0 + a_1 x + \ldots + a_{k-1}x^{k-1} + x^k \in K[x]$, so gilt nach der gleichen Beweisführung wie oben, dass $q|\mu_\alpha(x)$ und, falls $\langle V, \alpha(V), \alpha^2(V), \ldots, \alpha^{k-1}(V)\rangle = \mathcal{V}$, so gilt sogar $q = \mu_\alpha(x)$ (siehe Bemerkung 5.2.8).
Z.B.

$$A := \begin{pmatrix} 1 & -2 & 3 \\ -4 & 0 & -4 \\ 3 & -2 & 1 \end{pmatrix} \in \mathbb{Q}^{3\times 3}$$

liefert mit dem Vektor $V := \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \in \mathbb{Q}^{3\times 1}$ die Folge

$$(V, \widetilde{A}(V), \widetilde{A}^2(V), \widetilde{A}^3(V)) = (\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ -4 \\ 3 \end{pmatrix}, \begin{pmatrix} 18 \\ -16 \\ 14 \end{pmatrix}, \begin{pmatrix} 92 \\ -128 \\ 100 \end{pmatrix})$$

Man setzt ein lineares Gleichungssystem an und bekommt die eindeutige Lösung

$$32V + 24\widetilde{A}(V) + 2\widetilde{A}^2(V) = \widetilde{A}^3(V),$$

da die ersten drei Vektoren noch linear unabhängig sind, und somit $x^3 - (32 + 24x + 2x^2)$ als Minimalpolynom von A. Man beachte, $B := (V, \widetilde{A}(V), \widetilde{A}^2(V))$ ist eine Basis von $\mathbb{Q}^{3\times 1}$ und

$${}^B\widetilde{A}^B = \begin{pmatrix} 0 & 0 & 32 \\ 1 & 0 & 24 \\ 0 & 1 & 2 \end{pmatrix}$$

Nicht immer ist die Bestimmung des Minimalpolynoms so schmerzfrei wie bei den obigen Beispielen. Wir geben daher einen Algorithmus an, der die Bestimmung des Minimalpolynoms eines Endomorphismus α auf die (leichtere) Bestimmung hinreichend vieler Minimalpolynome von Vektoren von $\mathcal{V}$ reduziert: Ende Vorl. 21 21.12

Bemerkung 5.2.8. Sei $\mathcal{V}$ endlich erzeugter K-Vektorraum, $\alpha \in \mathrm{End}(\mathcal{V})$ und $0 \neq V \in \mathcal{V}$. Dann gibt es ein kleinstes $k \leq \mathrm{Dim}(\mathcal{V})$, so dass

$$(V, \alpha(V), \alpha^2(V), \ldots, \alpha^k(V)) \in \mathcal{V}^{k+1}$$

linear abhängig ist und eine eindeutige lineare Abhängigkeit $(a_0, a_1, \ldots, a_{k-1}, 1) \in K^{k+1}$mit

$$a_0 V + a_1\alpha(V) + \cdots + a_{k-1}\alpha^{k-1}(V) + \alpha^k(V) = 0.$$

Dann heißt

$$\mu_{\alpha,V}(x) = a_0 + a_1 x + \cdots + a_{k-1}x^{k-1} + x^k$$

das **Minimalpolynom** des Vektors V bezüglich α. Der k-dimensionale Teilraum

$$\mathcal{W} := K[\alpha]V := \{p(\alpha)(V) \mid p \in K[x]\} = \langle V, \alpha(V), \alpha^2(V), \ldots, \alpha^{k-1}(V)\rangle$$

ist invariant unter α, d.h. $\alpha(\mathcal{W}) \subseteq \mathcal{W}$ und $\mu_{\alpha,V}(x)$ ist das Minimalpolynom der Einschränkung

$$\beta : \mathcal{W} \to \mathcal{W} : W \mapsto \alpha(W).$$

Mit Lemma 5.2.5 gilt $\mu_{\alpha,V}(x)$ teilt $\mu_\alpha(x)$.

Beweis. Übungsaufgabe. □

Bemerkung 5.2.9. Sei $p = x^d + a_{d-1}x^{d-1} + \ldots + a_0 \in K[x]$ normiert vom Grad d. Die Multiplikation mit x induziert eine lineare Abbildung m_p auf $K[x]/pK[x]$, die bezüglich der Basis

$$B = (\overline{1}, \overline{x}, \overline{x}^2, \ldots, \overline{x}^{d-1}) \in (K[x]/pK[x])^d \quad \text{mit } \overline{x} := x + pK[x]$$

die Matrix

$${}^B m_p^B =: M_p = \begin{pmatrix} 0 & 0 & 0 & \ldots & 0 & 0 & -a_0 \\ 1 & 0 & 0 & \ldots & 0 & 0 & -a_1 \\ 0 & 1 & 0 & \ldots & 0 & 0 & -a_2 \\ 0 & 0 & 1 & \ldots & 0 & 0 & -a_3 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 1 & 0 & -a_{d-2} \\ 0 & 0 & 0 & \ldots & 0 & 1 & -a_{d-1} \end{pmatrix} \in K^{d\times d}$$

hat. Diese Matrix heißt die **Begleitmatrix** von p. Nach Bemerkung 5.2.8 gilt

$$\mu_{M_p} = p.$$

Algorithmus 5.2.10.
Gegeben: $\alpha \in \mathrm{End}(\mathcal{V})$, $\mathcal{V}$ endlich erzeugter K-Vektorraum.
Gesucht: Das Minimalpolynom $\mu_\alpha(x)$.
Algorithmus:

(1) Setze $\mathcal{W} := \{0\} \leq \mathcal{V}$ und $\mu := 1 \in K[x]$.

(2) Falls $\mathcal{W} = \mathcal{V}$ gebe $\mu_\alpha(x) = \mu(x)$ zurück.

(3) Wähle $V \in \mathcal{V} \setminus \mathcal{W}$ und bestimme das Minimalpolynom $\mu_{\alpha,V}(x)$ von α bezüglich V.

(4) Ersetze

$$\mu(x) \qquad \text{durch} \qquad \mathrm{kgV}(\mu(x), \mu_{\alpha,V}(x)) = \frac{\mu(x)\mu_{\alpha,V}(x)}{\mathrm{ggT}(\mu(x), \mu_{\alpha,V}(x))}$$

und

$$\mathcal{W} \qquad \text{durch} \qquad \mathcal{W} + K[\alpha]V := \langle \mathcal{W}, K[\alpha]V \rangle.$$

Springe zu Schritt (2).

Beweis. Der Algorithmus terminiert nach spätestens $\mathrm{Dim}(\mathcal{V})$ Schritten. Wir zeigen durch Induktion nach der Anzahl der Schritte, dass $\mathcal{W}$ in jedem Schritt invariant unter α und $\mu(x)$ das Minimalpolynom der Einschränkung von α auf $\mathcal{W}$ ist. Der Induktionsanfang ist gerade Bemerkung 5.2.8.
Induktionsannahme: Für das letzte $\mathcal{W}$ gilt: $\mathcal{W}$ ist invariant unter α und $\mu(x)$ ist das Minimalpolynom der Einschränkung von α auf $\mathcal{W}$.
Induktionsschritt: Sei nun $V \in \mathcal{V} \setminus \mathcal{W}$. Da $\mathcal{W}$ und $K[\alpha]V$ invariant unter α sind, gilt dies auch für das Erzeugnis $\widetilde{\mathcal{W}} := \mathcal{W} + K[\alpha]V$. Da $\widetilde{\mu}(x) := \mathrm{kgV}(\mu(x), \mu_{\alpha,V}(x))$ sowohl Vielfaches von $\mu(x)$ als auch von $\mu_{\alpha,V}(x)$ ist, gilt $\widetilde{\mu}(\alpha)(U) = 0$ sowohl für alle $U \in \mathcal{W}$ als auch für alle $U \in K[\alpha]V$, somit auch für alle U in dem Erzeugnis der beiden. Ein Polynom $r(x)$ mit $r(\alpha)(\mathcal{W} + K[\alpha]V) = \{0\}$ muss sowohl ein Vielfaches von $\mu(x)$ als auch von $\mu_{\alpha,V}(x)$ sein, also ein Vielfaches des kleinsten gemeinsamen Vielfaches $\widetilde{\mu}(x)$. Somit ist $\widetilde{\mu}(x)$ das Minimalpolynom der Einschränkung von α auf $\widetilde{W}$. □

Übung 5.2.2. Zeigen Sie, dass $\mathrm{Grad}(\mu_\alpha(x)) \leq \mathrm{Dim}(\mathcal{V})$.
Hinweis: Benutze die Beweisidee des Algorithmus. Das Minimalpolynom der Einschränkung von α auf $\mathcal{W} \cap K[\alpha]V$ teilt $\mu(x)$ und $\mu_{\alpha,V}(x)$.

Man beachte, dass der Algorithmus, wenn er nicht schon nach einem Schritt terminiert, wie im Beispiel 5.2.7 der Fall war, dann eine Faktorisierung des Minimalpolynoms gleichzeitig mitliefert. Dies wird sich als vorteilhaft erweisen.

5.3 Eigenvektoren und Diagonalisierbarkeit

Lernziel: Eigenwerte und Eigenvektoren, Beispiele von Eigenvektorbasen, diagonalisierbare Matrizen.

Definition 5.3.1. Sei $\alpha : \mathcal{V} \to \mathcal{V}$ ein Endomorphismus des K-Vektorraumes $\mathcal{V}$.

(1) Ein $a \in K$ heißt **Eigenwert** von α, falls ein Vektor $V \in \mathcal{V}$ existiert mit

$$\alpha(V) = aV \text{ und } V \neq 0.$$

In diesem Fall heißt V **Eigenvektor**[1] von α zum Eigenwert a. Allgemeiner heißt für $b \in K$ der Teilraum

$$E_\alpha(b) = E(b) := \mathrm{Kern}(\alpha - b\,\mathrm{id}_\mathcal{V}) \leq \mathcal{V}$$

der **Eigenraum** von α zu b. Eine Zahl $a \in K$ ist also genau dann Eigenwert von α, wenn $E_a(A) \neq \{0\}$, also genau dann, wenn es einen Eigenvektor von α zu a gibt.

(2) Ist E eine Basis von $\mathcal{V}$ bestehend aus Eigenvektoren von α, so heißt E eine **Eigenvektorbasis** von $\mathcal{V}$ bezüglich α.

(3) Wir nennen α **diagonalisierbar**, falls eine Eigenvektorbasis für α existiert.

(4) Vermöge des K-Algebrenisomorphismus

$$\widetilde{\cdot} : K^{n\times n} \to \mathrm{End}(K^{n\times 1}), A \mapsto \widetilde{A}$$

lassen sich die Begriffe auf quadratische Matrizen übertragen: Für $A \in K^{n\times n}$ heißt ein Vektor $X \in K^{n\times 1} \setminus \{0\}$ **Eigenvektor** von A zu $a \in K$, falls $AX = aX$ gilt und $E_a(A) = \{X \in K^{n\times 1} \mid AX = aX\}$ der **Eigenraum** von A zu a, etc.

Übung 5.3.1. Sei $A \in K^{n\times n}$. Zeigen Sie:

(1) Für $g \in \mathrm{GL}_n(K)$ ist $g^{-1}Ag$ genau dann eine Diagonalmatrix, wenn die Spalten von g eine Eigenvektorbasis von A bilden.

(2) Der Endomorphismus $\widetilde{A}$ (bzw. die Matrix A) ist genau dann diagonalisierbar, wenn eine Matrix $g \in \mathrm{GL}_n(K)$ existiert mit $g^{-1}Ag$ eine Diagonalmatrix.

Beispiel 5.3.2. Seien $s_1, \ldots, s_d \in K$ genau d verschiedene Elemente von K und $p := (x - s_1)\cdots(x - s_d) \in K[x]$. Dann ist

$$(q_1, \ldots, q_d) \text{ mit } q_i := p/(x - s_i) \in K[x]$$

eine Basis von $K[x]_{\mathrm{Grad}<d}$ und somit $E := (\overline{q_1}, \ldots, \overline{q_d}) \in (K[x]/pK[x])^d$ eine Basis von $K[x]/pK[x]$. Wegen $(x - s_i)q_i = p$ sieht man sofort $\overline{x} \cdot \overline{q_i} = s_i\overline{q_i}$, d.h. die Matrix von m_p bezüglich der Basis E hat Diagonalgestalt:

$${}^E m_p^E = \mathrm{Diag}(s_1, \ldots, s_d) := \begin{pmatrix} s_1 & 0 & 0 & \ldots & 0 & 0 & 0 \\ 0 & s_2 & 0 & \ldots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 0 & s_{d-1} & 0 \\ 0 & 0 & 0 & \ldots & 0 & 0 & s_d \end{pmatrix}$$

[1] Der Nullvektor erfüllt die erste Bedingung für alle $a \in K$, darum müssen wir ihn explizit ausschließen.

Jedes $\overline{q_i}$ ist also Eigenvektor von m_p zum Eigenwert s_i.

Satz 5.3.3. *Sei $\alpha : \mathcal{V} \to \mathcal{V}$ Endomorphismus des endlich erzeugten K-Vektorraumes $\mathcal{V}$. Genau dann ist $a \in K$ Eigenwert von α, wenn a Wurzel des Minimalpolynoms ist (d.h. $\mu_\alpha(a) = 0$).*

Beweis. Sei a Eigenwert von α, d.h. $E_\alpha(a) = \mathrm{Kern}(\alpha - a\,\mathrm{id}_\mathcal{V}) \neq \{0\}$. Dann induziert α auf $E_\alpha(a)$ die lineare Abbildung $\beta =$ Multiplikation mit a. Diese hat $x - a$ als Minimalpolynom. Wegen Lemma 5.2.5 ist also a Wurzel von $\mu_\alpha(x)$.
Sei umgekehrt $a \in K$ Wurzel von $\mu_\alpha(x)$, also $\mu_\alpha(x) = (x - a)q$ für ein $q \in K[x]$. Angenommen $E_\alpha(a) = \{0\}$. Dann ist der Kern von $\alpha - a\,\mathrm{id}_\mathcal{V}$ gleich 0, also $\alpha - a\,\mathrm{id}_\mathcal{V}$ bijektiv. Insbesondere

$$\mu_\alpha(\alpha) = 0 \text{ genau dann, wenn } q(\alpha) = 0,$$

was der Definition von $\mu_\alpha(x)$ als Minimalpolynom widerspricht. □

Beispiel 5.3.4. Sei $\mathcal{V}$ ein 2-dimensionaler $\mathbb{Q}$-Vektorraum mit Basis $B \in \mathcal{V}^2$ und Endomorphismus $\alpha \in \mathrm{End}(\mathcal{V})$, so dass

$${}^B\alpha^B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Man sieht sofort, dass $\mu_\alpha(x) = x^2 - 1$ das Minimalpolynom ist, also 1 und -1 die Eigenwerte sind. Die Koordinatenspalten ${}^BV$ der Eigenvektoren zum Eigenwert 1 bzw. -1 bekommen wir durch Lösen des linearen Gleichungssystems

$$({}^B\alpha^B - I_2)X = 0 \text{ bzw. } ({}^B\alpha^B + I_2)X = 0$$

und erhalten

$$E_\alpha(1) = \{V \mid {}^BV = \begin{pmatrix} a \\ a \end{pmatrix}, a \in \mathbb{Q}\}$$

und

$$E_\alpha(-1) = \{V \mid {}^BV = \begin{pmatrix} a \\ -a \end{pmatrix}, a \in \mathbb{Q}\}.$$

Also die Koordinatendarstellung einer möglichen Eigenvektorbasis E bzgl. B ist gegeben durch die Spalten der Matrix

$${}^B\,\mathrm{id}_\mathcal{V}^E = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix},$$

wobei man aber auch jede Spalte durch ein Vielfaches $\neq 0$ ersetzen kann. Jedenfalls liefert diese oder eine in dieser Weise modifizierte Eigenvektorbasis die Matrix

$${}^E\alpha^E = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

für α, und zwar ohne jede weitere Rechnung. Insbesondere braucht die Inverse ${}^E\,\mathrm{id}_\mathcal{V}^B = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ von ${}^B\,\mathrm{id}_\mathcal{V}^E$ nicht berechnet zu werden.

Ende Vorl. 22 09.01

Beispiel 5.3.5 (Projektionen)**.** Eine **Projektion** ist eine Abbildung $\pi \in \mathrm{End}(\mathcal{V})$ mit $\pi^2 = \pi$. Sieht man von den Grenzfällen $\pi = \mathrm{id}_\mathcal{V}$ und $\pi = 0_\mathcal{V}$ ab, so heißt dies, dass $x^2 - x = x(x - 1)$ das Minimalpolynom von π ist. Somit sind 0 und 1 die Eigenwerte von π und aus $\mathrm{id}_\mathcal{V} = \pi + (\mathrm{id}_\mathcal{V} - \pi)$ folgt, dass

$$\mathcal{V} = E_\pi(1) \oplus E_\pi(0)$$

gilt mit

$$E_\pi(0) := \mathrm{Kern}(\pi) = \mathrm{Bild}(\mathrm{id}_\mathcal{V} - \pi),$$
$$E_\pi(1) := \mathrm{Kern}(\pi - \mathrm{id}_\mathcal{V}) = \mathrm{Bild}(\pi).$$

Z.B. sieht man die letzte Gleichheit so: $(\pi - \mathrm{id}_\mathcal{V}) \circ \pi = 0$ besagt, dass $\mathrm{Bild}(\pi) \leq \mathrm{Kern}(\pi - \mathrm{id}_\mathcal{V})$. Umgekehrt ist $X \in \mathrm{Kern}(\pi - \mathrm{id}_\mathcal{V}) \iff \pi(X) - X = 0 \iff X = \pi(X)$, also $X \in \mathrm{Bild}(\pi)$. Die Trivialität des Durchschnittes $E_\pi(0) \cap E_\pi(1)$ folgt daraus, dass kein Vektor (ungleich Null) Eigenvektor zu zwei verschiedenen Eigenwerten sein kann.

Insbesondere hat man eine Eigenvektorbasis E für π mit ${}^E\pi^E = \mathrm{Diag}(\underbrace{1, \ldots, 1}_{\mathrm{Dim}\, E(1)}, \underbrace{0, \ldots, 0}_{\mathrm{Dim}\, E(0)})$.

Lemma 5.3.6. *Sei $\mathcal{V}$ ein endlich erzeugter K-Vektorraum und $\alpha \in \mathrm{End}(\mathcal{V})$. Sei $\pi \in \mathrm{End}(\mathcal{V})$ eine mit α vertauschbare Projektion, d.h. $\pi^2 = \pi$ und $\alpha \circ \pi = \pi \circ \alpha$. Dann sind $\mathrm{Kern}(\pi)$ und $\mathrm{Bild}(\pi) = \mathrm{Kern}(\pi - \mathrm{id}_\mathcal{V})$ beides α-invariante Teilräume von $\mathcal{V}$ und $\mathcal{V} = \mathrm{Kern}(\pi) \oplus_i \mathrm{Bild}(\pi)$.*

Beweis. Übung. □

Übung 5.3.2. Seien $\mathcal{V}$ ein endlich erzeugter K-Vektorraum, $\alpha \in \mathrm{End}(\mathcal{V})$ und $\mathcal{V} = \mathcal{T}_1 \oplus_i \mathcal{T}_2$ eine α-invariante direkte Summenzerlegung mit $\alpha_i := \alpha_{|\mathcal{T}_i} : \mathcal{T}_i \to \mathcal{T}_i$. Dann ist $\mu_\alpha = \mathrm{kgV}(\mu_{\alpha_1}, \mu_{\alpha_2})$. Vergleiche diese Aussage mit Lemma 5.2.5. Formuliere analog zu Bemerkung 5.2.6 die „Matrixversion“ dieser Aussage aus.

Bemerkung 5.3.7. Sei $\mathcal{V}$ ein endlich erzeugter K-Vektorraum und $\alpha \in \mathrm{End}(\mathcal{V})$.

(1) Sei $V \in \mathcal{V} \setminus \{0\}$ und $\mu_{\alpha,V}$ das Minimalpolynom von Grad d von α bezüglich V. Ferner sei B die Ergänzung von $(V, \alpha(V), \ldots, \alpha^{d-1}(V))$ zu einer Basis von $\mathcal{V}$. Dann gilt:

$${}^B\alpha^B = \begin{pmatrix} M_{\mu_{\alpha,V}} & * \\ 0 & * \end{pmatrix}.$$

(2) Ist $\mu_\alpha = \mu_{\alpha,V}$ für eine $V \in \mathcal{V}$, so gibt es eine Basis B von $\mathcal{V}$ mit

$${}^B\alpha^B = \begin{pmatrix} M_{\mu_\alpha} & * \\ 0 & * \end{pmatrix}.$$

(3) Wir werden später einsehen, dass ein solches V immer existiert.

Satz 5.3.8. *Sei $\mathcal{V}$ ein endlich erzeugter K-Vektorraum und $\alpha \in \mathrm{End}(\mathcal{V})$ mit Minimalpolynom $\mu_\alpha(x) \in K[x]$.*

(1) Ist $\mu_\alpha(x) = p_1(x)p_2(x)$ mit $p_1, p_2 \in K[x]$ teilerfremd von positiven Graden und normiert, d.h. $\mathrm{ggT}(p_1, p_2) = 1$, dann gibt es eine mit α verträgliche (sprich α-invariante) direkte Summenzerlegung $\mathcal{V} = \mathcal{T}_1 \oplus \mathcal{T}_2$, so dass $\alpha_i : \mathcal{T}_i \to \mathcal{T}_i : T \mapsto \alpha(T)$ Minimalpolynom p_i für $i = 1, 2$ hat. Insbesondere hat ${}^B\alpha^B$ Blockdiagonalgestalt für angepaßte Basen B von $\mathcal{V}$.

(2) Ist $\mu_\alpha(x) = \prod_{i=1}^d p_i$ mit $p_i \in K[x]$ paarweise teilerfremd und normiert, dann gibt es eine mit α verträgliche direkte Summenzerlegung $\mathcal{V} = \bigoplus_{i=1}^d \mathcal{T}_i$, so dass $\alpha_i := \alpha_{|\mathcal{T}_i} : \mathcal{T}_i \to \mathcal{T}_i$ Minimalpolynom p_i hat.

Bemerkung 5.3.9. $\mathcal{V} = \bigoplus_{i=1}^d \mathcal{T}_i$ bedeutet, dass sich jedes $X \in \mathcal{V}$ eindeutig schreiben läßt als $X = \sum_{i=1}^d X_i$ mit $X_i \in \mathcal{T}_i$. Aus einer Übungaufgabe wissen wir, dass folgende Aussagen äquivalent sind:

(1) $\mathcal{V} = \bigoplus_{i=1}^d \mathcal{T}_i$

(2) Sind B_i Basen von $\mathcal{T}_i$ ($i = 1, \ldots, d$), so ist $B = \cup_{i=1}^{d} B_i$ eine Basis von $\mathcal{V}$ (eine solche Basis heißt **eine an die Zerlegung angepasste Basis**).

(3) $\mathcal{V} = \langle \mathcal{T}_1, \mathcal{T}_2, \ldots, \mathcal{T}_d \rangle =: \mathcal{T}_1 + \mathcal{T}_2 + \ldots + \mathcal{T}_d =: \sum_{i=1}^{d} \mathcal{T}_i$ (d.h. der Vektorraum $\mathcal{V}$ wird also erzeugt von den $\mathcal{T}_i$'s) und für jedes $j \in \{1, \ldots, d\}$ gilt $\mathcal{T}_j \cap (\sum_{i \neq j} \mathcal{T}_i) = \{0\}$.

Beweis von Satz 5.3.8. Wir übertragen eine Polynomrechnung in eine Rechnung mit Endomorphismen vermöge des Einsetzhomomorphismus $K[x] \to \mathrm{End}(\mathcal{V}) : p(x) \mapsto p(\alpha)$, welcher nach dem Homomorphiesatz einen Monomorphismus $K[x]/\mu_\alpha K[x] \to \mathrm{End}(\mathcal{V})$ induziert.

(1) Wegen der Teilerfremdheit liefert der EUKLIDische Algorithmus Polynome $q_1, q_2 \in K[x]$ mit $1 = q_1 p_1 + q_2 p_2$. Setze $\pi_1 := (q_2 p_2)(\alpha) = q_2(\alpha) \circ p_2(\alpha)$ und $\pi_2 := q_1(\alpha) \circ p_1(\alpha)$. Dann gilt für $i = 1, 2$:
(a) $\pi_1 + \pi_2 = \mathrm{id}_\mathcal{V}$, denn $1 = q_1 p_1 + q_2 p_2$.
(b) $\pi_i \circ \alpha = \alpha \circ \pi_i$, denn $q_i p_i x = x q_i p_i$; insbesondere sind $\mathrm{Kern}(\pi_1) = \mathrm{Bild}(\pi_2)$ und $\mathrm{Kern}(\pi_2) = \mathrm{Bild}(\pi_1)$ beides α-invariante Teilräume.
(c) $\pi_1 \circ \pi_2 = \pi_2 \circ \pi_1 = 0$, da $p_1(\alpha) p_2(\alpha)$ ein Faktor von beiden ist.
(d) $\pi_i^2 = \pi_i$, denn
$$\pi_1 \circ \pi_1 = \pi_1 \circ (1 - \pi_2) = \pi_1 - \pi_1 \circ \pi_2 = \pi_1.$$
Also ist π_1 eine mit α vertauschbare Projektion und man erhält mit Lemma 5.3.6 die α-invariante direkte Summenzerlegung
$$\mathcal{V} = \mathrm{Bild}(\pi_1) \oplus_i \mathrm{Bild}(\pi_2) = \mathcal{T}_1 \oplus \mathcal{T}_2$$
wobei $\mathcal{T}_i := \pi_i(\mathcal{V})$. Wegen $\mathcal{T}_1 = \mathrm{Kern}(\pi_2)$ und $\mathcal{T}_2 = \mathrm{Kern}(\pi_1)$ folgt (leichte Übung), dass p_i das Minimalpolynom von $\alpha_i := \alpha_{|\mathcal{T}_i} : \mathcal{T}_i \to \mathcal{T}_i$ ist.

(2) Aus (1) durch Iteration. □

Beispiel 5.3.10. Sei $A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix} \in \mathbb{F}_2^{3\times 3}$. Dann ist
$$(E_1, AE_1, A^2E_1, A^3E_1) = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \end{pmatrix}$$
also $\mu_{A,E_1} = x^3 + x^2 + x = \underbrace{x}_{p_1} \cdot \underbrace{(x^2 + x + 1)}_{p_2} = \mu_A$,
$$\mathrm{ggT}(x, x^2 + x + 1) = 1 = (x + 1) \cdot x + 1 \cdot (x^2 + x + 1).$$
Also ist $\pi_1 = A^2 + A + 1$ und $\pi_2 = A^2 + A = I_3 - \pi_1$. Bezüglich geeigneter Basen von $\mathrm{Bild}(\pi_1)$ und $\mathrm{Kern}(\pi_1)$ hat $\widetilde{A}$ die Gestalt
$$\mathrm{Diag}(M_x, M_{x^2+x+1}) = \mathrm{Diag}\left((0), \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}\right).$$
Dies gilt da $\mathrm{Dim}\,\mathrm{Bild}(\pi_1) = \mathrm{Grad}\, p_1$ und $\mathrm{Dim}\,\mathrm{Bild}(\pi_2) = \mathrm{Grad}\, p_2$ ist (vgl. Bemerkung 5.3.7).

Es ist $\mathrm{Bild}(\pi_1) = \mathrm{Kern}(\pi_2)$ eindimensional, $\mathrm{Bild}(\pi_1) = \langle E_1 + AE_1 + A^2E_1 = (1, 1, 1)^{tr} \rangle$. Eine geeignete Basis von $\mathrm{Bild}(\pi_2)$ erhält man als $(A^2E_1 + AE_1 = (0, 1, 1)^{tr}, A(A^2E_1 + AE_1) = A^3E_1 + A^2E_1 = (1, 0, 1)^{tr}$, so dass $g^{-1}Ag = \mathrm{Diag}((0), \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix})$ mit $g = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \in \mathrm{GL}_3(\mathbb{F}_2)$. Wieso gilt $\mathrm{Bild}(\pi_2) = \mathrm{Bild}(\widetilde{A})$?

Ende Vorl. 23 11.01

Beispiel 5.3.11. Sei $\mathcal{V}$ ein 6-dimensionaler $\mathbb{F}_2$-Vektorraum mit Basis B und $\alpha \in \mathrm{End}(\mathcal{V})$ mit

$$ {}^B\alpha^B = \begin{pmatrix} \cdot & 1 & \cdot & 1 & \cdot & \cdot \\ 1 & 1 & 1 & \cdot & \cdot & 1 \\ 1 & \cdot & 1 & 1 & 1 & 1 \\ \cdot & 1 & 1 & \cdot & 1 & \cdot \\ \cdot & 1 & \cdot & 1 & \cdot & 1 \\ \cdot & 1 & 1 & \cdot & \cdot & 1 \end{pmatrix} $$

Wenn wir α und seine Potenzen auf B_1 anwenden, sind die Koordinatenspalten der resultierenden Vektoren die Spalten der folgenden Matrix:

$$ \begin{pmatrix} 1 & \cdot & 1 & \cdot & \cdot & 1 \\ \cdot & 1 & \cdot & \cdot & \cdot & \cdot \\ \cdot & 1 & 1 & 1 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & 1 & 1 \\ \cdot & \cdot & 1 & \cdot & 1 & 1 \\ \cdot & \cdot & \cdot & 1 & \cdot & \cdot \end{pmatrix} $$

Die ersten 5 Spalten sind noch linear unabhängig, die letzte ist abhängig von den ersten 5 und wir erhalten $\mu_{\alpha,B_1} = 1 + x^4 + x^5$ als Teiler des Minimalpolynoms μ_α. Das Polynom $1 + x^4 + x^5$ hat keine Wurzeln in $\mathbb{F}_2$, aber $1 + x + x^2$ als irreduziblen Teiler, so dass wir

$$ 1 + x^4 + x^5 = (1 + x + x^2)(1 + x + x^3) $$

bekommen. Wenn wir geeignete Basen gefunden haben, so dass die Matrix von α Blockdiagonalgestalt hat, wird $p_1 := 1 + x + x^2$ den Diagonalblock $M_{p_1} := \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$ beitragen und $p_2 := 1 + x + x^3$ den Diagonalblock $M_{p_2} := \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$. Es kann höchstens noch ein Diagonalblock vom Grad 1 dazukommen. Wenn wir Spuren vergleichen, kommen wir zu dem Schluss, dass es (0) sein muss. Insbesondere sollte 0 Eigenwert sein. Man überzeugt sich davon, dass $x(1 + x + x^2)(1 + x + x^3)$ nach Übung 5.3.2 das Minimalpolynom von α ist und weiß, dass es eine Basis C gibt mit

$$ {}^C\alpha^C = \mathrm{Diag}((0), \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}). $$

Wie bestimmt man nun ${}^B\,\mathrm{id}_{\mathcal{V}}^C$? Der Weg, die Projektionen in die Komponenten mit Hilfe des EUKLIDischen Algorithmus auszurechnen, ist langwierig, weil man ja die Matrix einsetzen muss, aber möglich:

$$ 1 = (1 + x)x + (1 + x + x^2) $$

wird mit $1 + x + x^3$ multipliziert und in

$$ 1 = xx(1 + x + x^2) + (1 + x)(1 + x + x^3) $$

eingesetzt ergibt

$$ 1 = xx(1 + x + x^2) + (1 + x)^2x(1 + x + x^3) + (1 + x)(1 + x + x^2)(1 + x + x^3) $$

Statt nun α in jeden der drei Summanden einzusetzen, um die Projektionen zu bekommen kann man sich damit begnügen, nur jeweils einen Vektor in $\mathrm{Bild}(\alpha \circ (1 + \alpha + \alpha^2)), \mathrm{Bild}(\alpha \circ$

$(1+\alpha+\alpha^3)), \mathrm{Bild}((1+\alpha+\alpha^2)\circ(1+\alpha+\alpha^3))$ zu bestimmen. Dies bekommt man mit einer sehr schmerzfreien Rechnung, weil die $\alpha^i(B_1)$ schon bekannt sind. In den ersten beiden Fällen bekommt man die Vektoren mit den Komponenten

$$\begin{pmatrix} 1 \\ 1 \\ 1 \\ \cdot \\ 1 \\ 1 \end{pmatrix} \text{ bzw. } \begin{pmatrix} 1 \\ 1 \\ \cdot \\ 1 \\ \cdot \\ \cdot \end{pmatrix}$$

jedoch beim dritten Fall leider Null. Also muss man einen anderen Vektor als B_1 iterieren. Im vorliegenden Fall kann man sich auch noch anders helfen: Man berechnet $\mathrm{Kern}(\alpha) = E_\alpha(0)$. Mit diesen Vektoren erhalten wir nach Umstellung der Komponenten entsprechend unserer angestrebten Blockdiagonalmatrix für α:

$${}^B\mathrm{id}_{\mathcal{V}}^C = \left(\begin{array}{c|cc|ccc} \cdot & 1 & \cdot & 1 & 1 & 1 \\ 1 & 1 & \cdot & 1 & \cdot & \cdot \\ 1 & \cdot & \cdot & 1 & \cdot & 1 \\ 1 & 1 & 1 & \cdot & 1 & \cdot \\ \cdot & \cdot & \cdot & 1 & \cdot & \cdot \\ \cdot & \cdot & 1 & 1 & 1 & 1 \end{array}\right)$$

Man würde jetzt nicht auf die Idee kommen, nach der Formel ${}^C\alpha^C = ({}^B\mathrm{id}_{\mathcal{V}}^C)^{-1}{}^B\alpha^B\,{}^B\mathrm{id}_{\mathcal{V}}^C$ nachzurechnen, ob wirklich die gewünschte Blockdiagonalmatrix herauskommt, sondern nur überprüfen, wie sich die Bilder der Spalten unmittelbar vor dem | jeweils aus den vorausgehenden Spalten (eine, zwei oder drei) linearkombinieren.

$$\left.\begin{array}{cc} \cdot & \cdot \\ 1 & \cdot \\ 1 & \cdot \\ 1 & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \end{array}\right| \left.\begin{array}{ccc} 1 & \cdot & 1 \\ 1 & \cdot & 1 \\ \cdot & \cdot & \cdot \\ 1 & 1 & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & 1 & 1 \end{array}\right| \begin{array}{cccc} 1 & 1 & 1 & \cdot \\ 1 & \cdot & \cdot & 1 \\ 1 & \cdot & 1 & 1 \\ \cdot & 1 & \cdot & 1 \\ 1 & \cdot & \cdot & 1 \\ 1 & 1 & 1 & \cdot \end{array}$$

Folgerung 5.3.12. *Sei $\mathcal{V}$ endlich erzeugter K-Vektorraum und $\alpha \in \mathrm{End}(\mathcal{V})$ mit Minimalpolynom vom Grad d. Genau dann existiert eine Eigenvektorbasis für α, wenn $\mu_\alpha(x)$ genau d verschiedene Wurzeln $s_1, \ldots, s_d$ in K hat. (Also $\mu_\alpha(x) = \prod_{i=1}^d (x - s_i)$ für paarweise verschiedene $s_i \in K$.)*

Beweis. Sei E eine Eigenvektorbasis. Aus der Matrix

$${}^E\alpha^E = \mathrm{Diag}(a_1, \ldots, a_n)$$

lesen wir sofort das Minimalpolynom als $\prod_{i=1}^d (x - s_i)$ ab, wo s_i die *verschiedenen* Eigenwerte a_j durchläuft.
Sei umgekehrt $\mu_\alpha(x) = \prod_{i=1}^d (x - s_i)$ mit $s_i \in K$ paarweise verschieden. Wir wenden Satz 5.3.8 mit den teilerfremden Polynomen $p_i := x - s_i$ $(i = 1, \ldots, d)$ an und erhalten eine Zerlegung $\mathcal{V} = \bigoplus_{i=1}^d \mathcal{T}_i$ von $\mathcal{V}$ in α-invariante Teilräume $\mathcal{T}_i$ mit $\alpha_{|\mathcal{T}_i} = s_i\,\mathrm{id}_{\mathcal{T}_i}$. Es ist also $\mathcal{T}_i = E_\alpha(s_i)$. Eine Eigenvektorbasis von $\mathcal{V}$ erhält man durch Zusammenfügen beliebiger Basen der Teilräume $\mathcal{T}_i$. □

Beispiel 5.3.13. Seien $a_1, \ldots, a_n \in K$ paarweise verschieden und $A \in K^{n\times n}$ mit $A_{i,i} = a_i$ und $A_{i,j} = 0$ für $i > j$, $1 \le i, j \le n$ (also eine obere Dreiecksmatrix). Dann ist $\mu_A = \prod_{i=1}^n (x - a_i)$ (etwa mit Bemerkung 5.2.6) und A ist diagonalisierbar also ähnlich zu $\mathrm{Diag}(a_1, \ldots, a_n)$.

5.4 Determinanten

5.4.a Unsere Wunschliste

Lernziel: Definition und Beispiele von Multilinearformen, Determinante, Berechnungsverfahren und Anwendungen

Wir wollen in diesem Abschnitt Determinanten einführen. Sie sind wichtige Werkzeuge der Theorie und haben viele Anwendungen auch außerhalb der linearen Algebra.

Definition 5.4.1. Sei $\mathcal{V}$ ein K-Vektorraum der Dimension n und $B \in \mathcal{V}^n$ eine Basis von $\mathcal{V}$. Die (bezüglich B normierte) **Determinante** von $\mathcal{V}$ ist eine Abbildung

$$\det_B : \mathcal{V}^n \to K : (V_1, \ldots, V_n) \mapsto \det_B(V_1, \ldots, V_n)$$

mit folgenden drei Eigenschaften:

(1) $\det_B$ ist **multilinear**, d. h.

$$\begin{aligned}\det_B(X_1, \ldots, X_{i-1}, aX_i + bX_i', X_{i+1}, \ldots, X_n) = \\ a \det_B(X_1, \ldots, X_{i-1}, X_i, X_{i+1}, \ldots, X_n) + \\ b \det_B(X_1, \ldots, X_{i-1}, X_i', X_{i+1}, \ldots, X_n)\end{aligned}$$

für alle $X \in \mathcal{V}^n, i \in \underline{n}, X_i' \in \mathcal{V}$ und $a, b \in K$.

(2) $\det_B$ ist **alternierend**, d.h. $\det_B(X) = 0$ für alle $X \in \mathcal{V}^n$, für die $i, j \in \underline{n}, i \neq j$ existieren mit[2] $X_i = X_j$.

(3) $\det$ ist normiert, d.h. $\det_B(B) = 1$.

Zusammenfassend ist die (bezüglich B normierte) Determinante eine alternierende Multilinearform mit $\det_B(B) = 1$.

Beispiel 5.4.2.

(1) Für $\mathcal{V} = K$ mit Basis $B = (1)$ ist $\det = \mathrm{id}_K$.

(2) Für $\mathcal{V} = K^{2\times 1}$ mit der Standardbasis $E = (e_1, e_2)$ als Basis B ist

$$\det_E(\begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} c \\ d \end{pmatrix}) = ad - bc.$$

Wir haben jetzt zwei Aufgaben: Nachweis der Eindeutigkeit und Nachweis der Existenz. Zum Nachweis der Eindeutigkeit, der gleichzeitig eine Idee vermittelt, wie man eine Determinante ausrechnet, brauchen wir etwas Vorbereitung aus der Gruppentheorie.

5.4.b Exkurs in die Gruppentheorie der symmetrischen Gruppe

Definition 5.4.3.

(1) Eine **Permutation** der Menge M ist ein Element der symmetrischen Gruppe S_M. Für $a, b \in M, a \neq b$ bezeichnet $\tau_{a,b}$ die Permutation

$$\tau_{a,b} : M \to M,\ m \mapsto \begin{cases} m, & m \neq a, b \\ a, & m = b \\ b, & m = a \end{cases}$$

Permutationen der Form $\tau_{a,b}$ heißen **Transpositionen**.

[2] also nicht injektive $X \in \mathcal{V}^n$.

(2) Für $\pi \in S_n$ sei $a(\pi) := |\{(i,j) \in \underline{n} \times \underline{n} \mid i < j, \pi(i) > \pi(j)\}|$ und

$$\operatorname{sign}(\pi) := (-1)^{a(\pi)}$$

heißt das **Signum** von π.

Lemma 5.4.4.

(1) Für $\pi \in S_n$ gilt

$$\operatorname{sign}(\pi) = \prod_{i<j} \frac{\pi(j) - \pi(i)}{j - i}$$

und

$$\operatorname{sign} : S_n \to \{1, -1\},\ \pi \mapsto \operatorname{sign}(\pi)$$

ist ein Homomorphismus mit $\operatorname{sign}(\tau_{a,b}) = -1$ für alle $a, b \in \underline{n}, a \neq b$.

(2) Jede Permutation $\pi \in S_n$ kann als Produkt (= Komposition) von Transpositionen geschrieben werden.

Beweis.

(1) Jeder Faktor von $\prod_{i<j}(j-i)$ taucht bis aufs Vorzeichen auch als Faktor von $\prod_{i<j}(\pi(j)-\pi(i))$ auf und umgekehrt, wobei genau $a(\pi)$ Vorzeichenwechsel auftreten. Damit folgt der erste Teil der Behauptung.
Seien nun $\pi, \sigma \in S_n$. Dann gilt

$$\begin{aligned} \operatorname{sign}(\pi \circ \sigma) &= \prod_{i<j} \tfrac{\pi(\sigma(j))-\pi(\sigma(i))}{j-i} \\ &= \prod_{i<j} \tfrac{\pi(\sigma(j))-\pi(\sigma(i))}{\sigma(j)-\sigma(i)} \tfrac{\sigma(j)-\sigma(i)}{j-i} \\ &= \prod_{i<j} \tfrac{\pi(j)-\pi(i)}{j-i} \prod_{i<j} \tfrac{\sigma(j)-\sigma(i)}{j-i} \\ &= \operatorname{sign}(\pi)\operatorname{sign}(\sigma). \end{aligned}$$

(2) Induktion über n: Für $n = 1$ ist alles richtig, da die Identität ein leeres Produkt von Transpositionen ist. Angenommen die Behauptung gilt für alle $\pi \in S_n$.
Sei nun $\pi \in S_{n+1}$. Falls $\pi(n+1) = n+1$, folgt die Behauptung wegen der Induktionsannahme. Falls $\pi(n+1) = i \neq n+1$, dann gilt $(\tau_{i,n+1} \circ \pi)(n+1) = n+1$ und wir können die Induktionsannahme auf $\tau_{i,n+1} \circ \pi$ anwenden und die Behauptung folgt wegen $\tau_{i,n+1}^2 = \mathrm{id}_{\underline{n}}$. □

Ende Vorl. 24 16.01

Offensichtlich hat S_n nur endlich viele Elemente. Die erste Frage, die man stellt, wenn man es mit einer endlichen Gruppe G zu tun hat, ist die nach der **Ordnung** $|G|$ der Gruppe, also nach der Anzahl der Elemente von G.

Bemerkung 5.4.5.

$$|S_n| = n!$$

Beweis. Volkstümlich formuliert man den Beweis so: Sei $\pi \in S_n$. Für $\pi(1)$ hat man n Möglichkeiten. Nachdem $\pi(1)$ festgelegt ist, hat man für $\pi(2)$ nur noch $n-1$ Möglichkeiten, also für $(\pi(1), \pi(2))$ insgesamt $n(n-1)$ Möglichkeiten, etc..

Hier ist ein anderer Beweis, der sich auf andere Situationen besser übertragen lässt. Wir fangen bei $\pi(n)$ statt bei $\pi(1)$ an:
Betrachte die Abbildung

$$\Lambda : S_n \to \underline{n},\ \pi \mapsto \pi(n).$$

Diese Abbildung ist surjektiv. Wenn wir noch zeigen, dass jede Faser $|S_{n-1}|$ Elemente hat, dann folgt die Behauptung durch Induktion, denn offensichtlich ist $|S_1| = 1$.

Die Faser $\Lambda^{-1}(\{n\})$ steht in offensichtlicher Bijektion mit S_{n-1}, kurz $\Lambda^{-1}(\{n\}) = S_{n-1}$. Weiter ist klar: $\tau_{i,n}\Lambda^{-1}(\{i\}) = \{\tau_{i,n} \circ \sigma \mid \sigma \in \Lambda^{-1}(\{i\})\} = \Lambda^{-1}(\{n\}) = S_{n-1}$, oder äquivalent $\Lambda^{-1}(\{i\}) = \tau_{i,n}S_{n-1} := \{\tau_{i,n} \circ \pi \mid \pi \in S_{n-1}\}$. Genauer:

$$S_{n-1} \to \Lambda^{-1}(\{i\}),\ \pi \mapsto \tau_{i,n} \circ \pi$$

ist eine Bijektion, da $\tau_{i,n}$ als Element der Gruppe S_n invertierbar ist. Damit folgt die Behauptung. □

Übung 5.4.1. Sei $\mathcal{V}$ ein Vektorraum der Dimension n über einem Körper K von $q < \infty$ Elementen. Zeige:

$$|\operatorname{GL}(\mathcal{V})| = (q^n - 1)(q^n - q)\cdots(q^n - q^{n-1}).$$

Folgerung 5.4.6.

$$S_n = \bigcup_{i=1}^{n} \tau_{i,n}S_{n-1}$$

wobei $\tau_{i,n}S_{n-1} \cap \tau_{j,n}S_{n-1} = \emptyset$ *für* $i \neq j$.

Später werden wir lernen, dass S_{n-1} eine Untergruppe von S_n ist und die $\tau_{i,n}S_{n-1}$ Restklassen nach dieser Untergruppe sind. Wir haben jetzt aber genügend Hilfsmittel aus der Gruppentheorie, um mit der Determinante fortfahren zu können.

5.4.c Eindeutigkeit und Existenz der Determinante

Wir können jetzt erste Eigenschaften der Determinante beweisen, die zu einem Eindeutigkeitsbeweis führen.

Lemma 5.4.7. *Ist* $X \in \mathcal{V}^n$ *und* $\pi \in S_n$, *so gilt*

$$\det_B(X \circ \pi) = \det_B(X)\operatorname{sign}(\pi).$$

Beweis. Für $\pi \in S_n$ sei

$$\ell(\pi) := \min\{k \mid \text{ Es existieren Transpositionen } \tau_1, \ldots, \tau_k \in S_n \text{ mit } \pi = \tau_1 \circ \ldots \circ \tau_k\}.$$

Wir führen den Beweis durch Induktion über die Länge $\ell(\pi)$. Die Behauptung ist klar, falls $\ell(\pi) = 0$, also $\pi = \mathrm{id}$. Im Falle $\ell(\pi) = 1$ ist π Transposition, sagen wir $\pi = \tau_{i,j}$. Setze

$$\widetilde{X} : \underline{n} \to \mathcal{V} : k \mapsto \begin{cases} X_k & k \neq i, j \\ X_i + X_j & k \in \{i, j\} \end{cases}$$

Dann ist

$$\begin{aligned} 0 &= \det_B(\widetilde{X}) \\ &= \det_B(\ldots, X_i + X_j, \ldots, X_i + X_j, \ldots) \\ &= \det_B(\ldots, X_i, \ldots, X_i, \ldots) + \det_B(\ldots, X_j, \ldots, X_j, \ldots) \\ &\quad + \det_B(\ldots, X_j, \ldots, X_i, \ldots) + \det_B(\ldots, X_i, \ldots, X_j, \ldots) \\ &= 0 + 0 + \det_B(X) + \det_B(X \circ \pi) \end{aligned}$$

und die Behauptung folgt für $\ell(\pi) = 1$.
Angenommen, die Behauptung gilt für alle $X \in \mathcal{V}^n$ und $\ell(\pi) = k$. Sei nun $\pi \in S_n$ mit $\ell(\pi) = k + 1$, also $\pi = \tau \circ \pi'$ mit $\ell(\tau) = 1, \ell(\pi') = k$. Dann gilt

$$\begin{aligned} \det_B(X \circ \pi) &= \det_B((X \circ \tau) \circ \pi') \\ &= \det_B(X \circ \tau)\operatorname{sign}(\pi') \\ &= \det_B(X)\operatorname{sign}(\tau)\operatorname{sign}(\pi') \\ &= \det_B(X)\operatorname{sign}(\pi) \end{aligned}$$

□

Satz 5.4.8. *Falls eine Determinante auf $\mathcal{V}$ mit Normierung bezüglich der Basis $B \in \mathcal{V}^n$ existiert, ist sie eindeutig bestimmt.*

Beweis. Sei $X \in \mathcal{V}^n$. Dann existiert eine eindeutige Matrix $A \in K^{n\times n}$ mit $X = BA$, nämlich $A = ({}^BX_1, \ldots, {}^BX_n)$. Es gilt:

$$\begin{aligned}
\det_B(X) &= \det_B(\textstyle\sum_{i_1} A_{i_1,1}B_{i_1}, \ldots, \sum_{i_n} A_{i_n,n}B_{i_n}) \\
&= \textstyle\sum_{i_1} A_{i_1,1} \det_B(B_{i_1}, \sum_{i_2} A_{i_2,2}B_{i_2}, \ldots, \sum_{i_n} A_{i_n,n}B_{i_n}) \\
&= \textstyle\sum_{i_1,i_2,\ldots,i_n} A_{i_1,1}A_{i_2,2}\cdots A_{i_n,n} \det_B(B_{i_1}, \ldots, B_{i_n}) \\
&= \textstyle\sum_{\pi\in S_n} A_{\pi(1),1}A_{\pi(2),2}\cdots A_{\pi(n),n} \det_B(B\circ\pi) \\
&= \textstyle\sum_{\pi\in S_n} A_{\pi(1),1}A_{\pi(2),2}\cdots A_{\pi(n),n} \det_B(B) \operatorname{sign}(\pi) \\
&= \textstyle\sum_{\pi\in S_n} A_{\pi(1),1}A_{\pi(2),2}\cdots A_{\pi(n),n} \operatorname{sign}(\pi)
\end{aligned}$$

Damit ist die Eindeutigkeit nachgewiesen. □

Der Beweis des Eindeutigkeitssatzes gibt uns auch einen Hinweis für den Existenzsatz. Wir können den Beweis nämlich so lesen: Wenn es überhaupt eine (bezüglich der Basis B normierte) Determinante gibt, dann ist sie durch

$$\boxed{\det_B(X_1, \ldots, X_n) = \sum_{\pi\in S_n} \operatorname{sign}(\pi)({}^BX_1)_{\pi(1)}({}^BX_2)_{\pi(2)}\cdots({}^BX_n)_{\pi(n)}}$$

gegeben.

Satz 5.4.9. *Die bezüglich der Basis $B \in \mathcal{V}^n$ normierte Determinante auf $\mathcal{V}$ existiert.*

Beweis. Wir müssen zeigen, dass die durch die obige Formel definierte Funktion $\det_B$ auf $\mathcal{V}^n$ die drei definierenden Eigenschaften der Determinante hat. Sofort klar sind die Multilinearität und die Normierungsbedingung, denn B eingesetzt liefert genau einen Summanden 1 und alle anderen Summanden gleich 0. Behauptung: $\det_B$ ist alternierend. Sei also $X \in \mathcal{V}^n$ mit $X_i = X_j$ für ein Paar $i, j \in \underline{n}, i \neq j$. Wir müssen zeigen, dass $\det_B(X) = 0$ gilt. In der Darstellung

$$\det_B(X) = \sum_{\pi\in S_n} \operatorname{sign}(\pi)({}^BX_1)_{\pi(1)}\cdots\ ({}^BX_n)_{\pi(n)}$$

vergleichen wir die beiden Summanden für $\pi \in S_n$ und $\pi \circ \tau_{i,j} \in S_n$. Beachte, dass diese beiden Permutationen verschieden sind, also wirklich zwei Summanden vorliegen. Beachte weiter, dass $X = X \circ \tau_{i,j}$ gilt, also

$$\prod_k ({}^BX_k)_{\pi(k)} = \prod_k ({}^BX_{\tau_{i,j}(k)})_{\pi(k)} = \prod_l ({}^BX_l)_{\pi(\tau_{i,j}(l))}$$

Andererseits aber gilt $\operatorname{sign}(\pi) = -\operatorname{sign}(\pi \circ \tau_{i,j})$. Also heben sich die beiden Summanden gegenseitig auf und die Gesamtsumme ist Null. □

Bemerkung 5.4.10. Sind B und B' Basen von $\mathcal{V}$, so gilt

$$\det_{B'} = \det_B(B')^{-1} \det_B$$

Beweis. $\det_{B'}$ und $\det_B$ sind beides alternierende Multilinearformen. Im Beweis der Eindeutigkeit haben wir die Normierung erst ganz am Ende benutzt. Daher sieht man, dass es ein $a \in K$ gibt mit $\det_{B'} = a \det_B$. Das a bestimmt aus

$$1 = \det_{B'}(B') = a \det_B(B').$$

□

Satz 5.4.11.

(1) *Seien* $X \in \mathcal{V}^n$, $1 \leq k \leq n$ *und*

$$Y : \underline{n} \to \mathcal{V} : i \mapsto \begin{cases} X_i & i \neq k \\ X_k + Z & i = k \end{cases}$$

für ein $Z \in \langle X_1, \ldots, X_{k-1}, X_{k+1}, \ldots X_n \rangle$. *Dann gilt* $\det_B(X) = \det_B(Y)$.

(2) $X \in \mathcal{V}^n$ *ist genau dann linear abhängig, wenn* $\det_B(X) = 0$ *gilt.*

Beweis.

(1) Sei $Z = Xa := \sum_i a_i X_i$ mit $a \in K^{n\times 1}$, $a_k = 0$. Dann gilt

$$\det_B(Y) = \det_B X + \textstyle\sum_{i\neq k} a_i 0$$

wegen der Linearität in der k-ten Komponente und weil $\det_B$ alternierend ist.

(2) Ist X linear abhängig, so folgt $\det_B(X) = 0$ aus (1). Ist X linear unabhängig, so ist X eine Basis und nach Bemerkung 5.4.10 $\det_B(X) \neq 0$. □

5.4.d Die Determinante einer Matrix

Definition 5.4.12. Sei $A \in K^{n\times n}$. Dann setzt man

$$\det(A) := \det_E(A_{-,1}, \ldots, A_{-,n}),$$

wobei die Determinante auf der rechten Seite die bezüglich der Standardbasis $E = (e_1, \ldots, e_n)$ normierte Determinante des $K^{n\times 1}$ ist.

Aus der oben entwickelten Formel für Determinanten erhalten wir die Leibniz Regel:

Bemerkung 5.4.13 (Leibniz Regel)**.** Für $A \in K^{n\times n}$ ist

$$\det(A) = \sum_{\pi\in S_n} \operatorname{sign}(\pi) \prod_{i=1}^{n} A_{\pi(i),i}$$

Beispiel 5.4.14.

- $\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc.$
- $\det \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix} = a_1b_2c_3 + a_2b_3c_1 + a_3b_1c_2 - a_1b_3c_2 - a_2b_1c_3 - a_3b_2c_1.$

Hier sind wichtige Eigenschaften von Determinanten von Matrizen.

Ende Vorl. 25 18.01

Satz 5.4.15.

(1) *Für* $A \in K^{n\times n}$ *gilt*

$$\det(A) = \det(A^{tr}).$$

(2) *Für* $A \in K^{n\times n}$ *gilt*

$$A \text{ invertierbar} \Leftrightarrow \det(A) \neq 0.$$

(3) *Für* $A_1, A_2 \in K^{n\times n}$ *gilt*

$$\det(A_1A_2) = \det(A_1)\det(A_2).$$

(4) *Für $A \in K^{n\times n}$ und $g \in \mathrm{GL}_n(K)$ ist $\det(g^{-1}Ag) = \det(A)$. Ähnliche Matrizen haben also die gleiche Determinante.*

(5) *Sei $\alpha \in \mathrm{End}(\mathcal{V})$ ein Endomorphismus eines endlich dimensionalen K Vektorraums. Ist B eine Basis von $\mathcal{V}$ so setzen wir*

$$\det(\alpha) := \det({}^B\alpha^B), \mathrm{Spur}(\alpha) = \mathrm{Spur}({}^B\alpha^B).$$

Beweis.

(1) Man hat

$$\begin{aligned}\det(A) &= \sum_{\pi\in S_n} \mathrm{sign}(\pi) \prod_i A_{\pi(i),i} \\ &= \sum_{\pi\in S_n} \mathrm{sign}(\pi) \prod_\ell A_{\ell,\pi^{-1}(\ell)} \\ &= \sum_{\pi\in S_n} \mathrm{sign}(\pi) \prod_\ell A^{tr}_{\pi^{-1}(\ell),\ell} \\ &= \sum_{\pi\in S_n} \mathrm{sign}(\pi) \prod_\ell A^{tr}_{\pi(\ell),\ell} \\ &= \det(A^{tr}).\end{aligned}$$

(2) Es gilt: A ist invertierbar, genau dann wenn die Spalten von A eine Basis von $K^{n\times 1}$ bilden. Daher folgt die Behauptung aus Satz 5.4.11.(2).

(3) Ist $\det(A_1) = 0$, so ist die Behauptung klar, da dann auch die Spalten von A_1A_2 linear abhängig sind. Sei also $\det(A_1) \neq 0$. Dann ist neben $\det_E$ auch

$$\Delta : (K^{n\times 1})^n \to K : X \mapsto \det_E(A_1X)$$

eine alternierende Multilinearform, und somit ein Vielfaches von $\det_E$. Durch Einsetzen der Standardbasis erhält man $\Delta(E) = \det_E(A_1)$, also $\Delta = \det(A_1)\det_E$.

(4) folgt aus (3), denn es ist $\det(g^{-1}) = \det(g)^{-1}$, da $\det(I_n) = 1$.

(5) Die Definition von Determinante und Spur eines Endomorphismus ist unabhängig von der Basiswahl nach (4). □

Folgerung 5.4.16. *Bezeichnen wir die Einschränkung der Determinante von $K^{n\times n}$ auf $\mathrm{GL}_n(K)$ wieder mit $\det$, so gilt:*

$$\det : \mathrm{GL}_n(K) \to K^* : A \to \det(A)$$

ist ein Gruppenhomomorphismus auf die multiplikative Gruppe des Körpers K. Der Kern dieses Homomorphismus, also das volle Urbild der 1 wird mit $\mathrm{SL}(n, K)$ bezeichnet und heißt die **spezielle lineare Gruppe**.

Bemerkung 5.4.17.

(1) Die Determinante von A ändert sich nicht, wenn man ein Vielfaches einer Spalte (Satz 5.4.11.(1)) oder Zeile (Satz 5.4.15.(1)) zu einer anderen hinzuaddiert.

(2) Es gilt

$$\det\left(\begin{pmatrix} a_1 & * & \dots & * \\ 0 & a_2 & * & \vdots \\ \vdots & \ddots & & \vdots \\ 0 & \dots & 0 & a_n \end{pmatrix}\right) = a_1a_2\dots a_n$$

(3) Die Determinante von A multipliziert sich mit (-1), wenn man zwei Spalten oder Zeilen vertauscht.

(4) Entsteht A' aus A durch Multiplikation einer Spalte oder Zeile mit $a \in K$, so ist $\det(A') = a\det(A)$.

(5) Zur Berechnung der Determinante einer Matrix bringt man sie mit dem Gauß-Algorithmus unter Beachtung von (1), (3) und (4) auf eine obere Dreiecksmatrix bringt und dann (2) benutzt.

Beispiel 5.4.18.

$$\det(\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 1 & 3 & 6 \end{pmatrix}) = \det(\begin{pmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & 1 & 3 \end{pmatrix}) = \det(\begin{pmatrix} 1 & 2 & 3 \\ 0 & 0 & 3 \\ 0 & 1 & 3 \end{pmatrix}) = -3$$

Übung 5.4.2. Sei $k + l = n$, $A_1 \in K^{k\times k}, A_2 \in K^{l\times l}, A_3 \in K^{l\times k}$. Man zeige

$$\det(\begin{pmatrix} A_1 & 0 \\ A_3 & A_2 \end{pmatrix}) = \det(A_1)\det(A_2).$$

Satz 5.4.19. *(*LAPLACE**scher Entwicklungssatz** *(Entwicklung nach einer Spalte)) Für $A \in K^{n\times n}$, $i, j \in \underline{n}$ sei $A^{(i,k)} \in K^{(n-1)\times(n-1)}$ die Matrix, die aus A durch Streichen der i-ten Zeile und k-ten Spalte entsteht. Dann gilt für $k \in \{1, \ldots, n\}$*

$$\det(A) = \sum_{i=1}^{n} A_{i,k}(-1)^{k+i}\det(A^{(i,k)})$$

Beweis. Es ist

$$\det(A) = \det_E(A_{-,1}, \ldots, A_{-,n}) = \sum_{i=1}^{n} A_{i,k}\det_E(A_{-,1}, \ldots, e_i, \ldots, A_{-,n}),$$

wobei die i-te Einheitsspalte E_i gerade an der k-ten Stelle steht. Die Idee ist jetzt, durch sukzessives Vertauschen von nebeneinanderliegenden Spalten und Zeilen die Matrix im i-en Summanden auf Block-Dreiecksgestalt $\begin{pmatrix} 1 & * \\ 0 & A^{(i,k)} \end{pmatrix}$ zu bringen. Um die k-te Spalte nach vorne zu bringen muss man $k-1$ mal benachbarte Spalten vertauschen, also ist

$$\det_E(A_{-,1}, \ldots, e_i, \ldots, A_{-,n}) = (-1)^{k-1}\det_E(e_i, A_{-,1}, \ldots, A_{-,n}).$$

Um die i-te Zeile nach oben zu bringen muss man $i-1$ mal benachbarte Zeilen vertauschen, also ist

$$\det_E(A_{-,1}, \ldots, e_i, \ldots, A_{-,n}) = (-1)^{(k-1)+(i-1)}\det(\begin{pmatrix} 1 & * \\ 0 & A^{(i,k)} \end{pmatrix}) = (-1)^{k+i}\det(A^{(i,k)}).$$

□

Übung 5.4.3. Man formuliere die Entwicklung der Determinante nach einer Zeile.

Beispiel 5.4.20. Sei

$$A := \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 2 & 3 & 4 \\ 1 & 2 & -1 & -2 \\ a & 1 & 1 & 2 \end{pmatrix} \in \mathbb{Q}^{4\times 4}.$$

Aufgabe: Berechne $\det(A)$.
1. Lösung (sehr schlecht): Entwicklung nach der ersten Zeile:

$$\det(A) = 1\det\begin{pmatrix} 2 & 3 & 4 \\ 2 & -1 & -2 \\ 1 & 1 & 2 \end{pmatrix} - 2\det\begin{pmatrix} 0 & 3 & 4 \\ 1 & -1 & -2 \\ a & 1 & 2 \end{pmatrix}$$

$$+3\det\begin{pmatrix} 0 & 2 & 4 \\ 1 & 2 & -2 \\ a & 1 & 2 \end{pmatrix} - 4\det\begin{pmatrix} 0 & 2 & 3 \\ 1 & 2 & -1 \\ a & 1 & 1 \end{pmatrix} = \dots$$

Etwas besser wäre die Entwicklung nach der ersten Spalte (Vorzeichen $+ - +-$).
2. Lösung (mit Zeilen und Spaltenumformungen):

$$\begin{aligned}\det(A) &= \det\begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 2 & 3 & 4 \\ 1 & 2 & -1 & -2 \\ a & 1 & 1 & 2 \end{pmatrix} = \det\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 3 & 4 \\ 1 & 2 & -1 & -2 \\ a & 1 & 1 & 2 \end{pmatrix} \\ &= \det\begin{pmatrix} 2 & 3 & 4 \\ 2 & -1 & -2 \\ 1 & 1 & 2 \end{pmatrix} = \det\begin{pmatrix} 0 & 1 & 0 \\ 2 & -1 & -2 \\ 1 & 1 & 2 \end{pmatrix} \\ &= -\det\begin{pmatrix} 2 & -2 \\ 1 & 2 \end{pmatrix} = -(2^2+2) = -6.\end{aligned}$$

Wie kann man übrigens sofort sehen, dass der Wert der Determinante unabhängig von a ist?

Der folgende Satz ist mehr von theoretischem als praktischem Interesse.

Satz 5.4.21. *(CRAMERsche Regel) Ist $A \in K^{n\times n}$ von Höchstrang, also $\det(A) \neq 0$, und $b \in K^{n\times 1}$, dann ist die eindeutige Lösung des Gleichungssystems $AX = b$, $X \in K^{n\times 1}$ gegeben durch*

$$X_{i,1} = \frac{\det(A^{i,b})}{\det(A)}$$

für $i = 1, \dots, n$, wobei $A^{i,b}$ dieselben Spalten wie A hat, außer dass die i-te Spalte durch b ersetzt ist.

Beweis. Dass die Gleichung eindeutig lösbar ist, wissen wir bereits. Sei also $X \in K^{n\times 1}$ die eindeutige Lösung. Dann haben wir $b = \sum X_{i,1}A_{-,i}$, also durch Einsetzen

$$\det(A^{i,b}) = 0 + \dots + 0 + X_{i,1}\det(A) + 0 + \dots + 0. \qquad \square$$

Ende Vorl. 26 23.01

Als Folgerung aus der CRAMERschen Regel bekommt man eine Formel für die Inverse einer Matrix, die aber nur von theoretischem Wert ist.

Folgerung 5.4.22. *Sei $A \in \mathrm{GL}_n(K)$. Die Einträge der Inversen A^{-1} sind gegeben durch*

$$(A^{-1})_{ij} = \frac{(-1)^{i+j}\det(A^{(j,i)})}{\det(A)}$$

wo $A^{(j,i)} \in K^{(n-1)\times(n-1)}$ durch Streichen der j-ten Zeile und i-ten Spalte entsteht. Die Matrix $((-1)^{i+j}\det(A^{(j,i)}))_{i,j}$ heißt auch Matrix der **Kofaktoren**.

Beweis. Die j-te Spalte $(A^{-1})_{-,j}$ von A^{-1} ist die Lösung des Gleichungssystems $AX = I_{-,j}$, berechnet sich also nach der CRAMERschen Regel:

$$(A^{-1})_{i,j} = \frac{\det(A^{i,I_{-,j}})}{\det(A)}.$$

Es bleibt zu zeigen, dass $(-1)^{i+j}\det(A^{(j,i)}) = \det(A^{i,I_{-,j}})$ gilt. Dies folgt einfach durch Entwickelung von $\det(A^{i,I_{-,j}})$ nach der i-ten Spalte. □

5.5 Das charakteristische Polynom

5.5.a Das charakteristische Polynom eines Endomorphismus

Definition 5.5.1.

(1) Sei $A \in K^{n\times n}$. Das **charakteristische Polynom** von A ist $\chi_A(x) = \det(xI_n - A) \in K[x]$.

(2) Sei $\mathcal{V}$ ein endlich erzeugter K-Vektorraum und $\alpha \in \mathrm{End}(\mathcal{V})$. Dann heißt

$$\chi_\alpha(x) := \det(xI_n - {}^B\alpha^B)$$

das **charakteristische Polynom** von α, wobei $B \in \mathcal{V}^n$ eine Basis von $\mathcal{V}$ ist.

Lemma 5.5.2. *Das charakteristische Polynom $\chi_\alpha(x)$ ist wohldefiniert und hängt insbesondere nicht von der Wahl der Basis B ab.*

Beweis. Man beachte zuerst, $xI_n - {}^B\alpha^B \in K(x)^{n\times n}$ und auf $K(x)^{n\times n}$ haben wir eine Determinante (siehe Übung 2.4.5). Weiter ist das Ergebnis ein Polynom, also in $K[x]$ nach der Leibniz Regel für Determinanten. Schließlich sei $C \in \mathcal{V}^n$ eine weitere Basis von $\mathcal{V}$. Dann gilt mit $T = {}^B\mathrm{id}_{\mathcal{V}}^C$:

$$\begin{aligned}\det(xI_n - {}^C\alpha^C) &= \det(xI_n - T^{-1}({}^B\alpha^B)T)\\ &= \det(T^{-1}(xI_n - {}^B\alpha^B)T)\\ &= \det(T)^{-1}\det(xI_n - {}^B\alpha^B)\det(T)\\ &= \det(xI_n - {}^B\alpha^B).\end{aligned}$$

□

Beispiel 5.5.3.

(1) Ist

$$A = \begin{pmatrix} a_1 & * & \dots & * \\ 0 & a_2 & * & \vdots \\ \vdots & \ddots & & * \\ 0 & \dots & 0 & a_n \end{pmatrix}$$

so ist $\chi_A = (x-a_1)(x-a_2)\dots(x-a_n)$.

(2) Ist $A = \begin{pmatrix} A_1 & * \\ 0 & A_2 \end{pmatrix}$ mit A_1, A_2 quadratisch, so ist $\chi_A = \chi_{A_1}\chi_{A_2}$.

(3) Ähnliche Matrizen haben das gleiche charakteristische Polynom.

(4) Ist also A diagonalisierbar, so ist

$$\chi_A = \prod_{a \in EW(A)} (x-a)^{\dim E_A(a)},$$

wo $EW(A)$ die Menge der Eigenwerte von A bezeichnet.

Übung 5.5.1. Ist $A = aI_n + bJ_n$ mit $b \neq 0$ und

$$J_n = \begin{pmatrix} 1 & \dots & 1 \\ \vdots & \dots & \vdots \\ 1 & \dots & 1 \end{pmatrix} \in K^{n\times n}.$$

Zeige: J_n und somit A ist genau dann diagonalisierbar wenn $n1_K \neq 0$. In diesem Fall ist $EW(A) = \{a+nb, a\}$ und die Dimension der Eigenräume ist 1 bzw. $n-1$. Also ist $\chi_A = (x-(a+nb))(x-a)^{n-1}$, $\mu_A = (x-(a+nb))(x-a)$.

Hier sind die wichtigsten Eigenschaften des charakteristischen Polynoms.

Satz 5.5.4. *Sei $\mathcal{V}$ ein endlich erzeugter K-Vektorraum und $\alpha \in \mathrm{End}(\mathcal{V})$. Dann gilt:*

(1) $\chi_\alpha(x) \in K[x]$ ist normiert vom Grad $n = \mathrm{Dim}(\mathcal{V})$. Der Koeffizient von x^{n-1} ist gleich $-\mathrm{Spur}(\alpha)$ und der Koeffizient von x^0 ist $(-1)^n \det(\alpha)$.

(2) $a \in K$ ist Eigenwert von α genau dann, wenn $\chi_\alpha(a) = 0$, d.h. falls a eine Wurzel von $\chi_\alpha(x)$. In anderen Worten, das Minimalpolynom und das charakteristische Polynom haben dieselben Nullstellen[3].

(3) (HAMILTON-CAYLEY) $\chi_\alpha(\alpha) = 0$, d.h. das Minimalpolynom teilt das charakteristische Polynom: $\mu_\alpha(x) | \chi_\alpha(x)$.

Beweis.

(1) Setze $A := {}^B\alpha^B$ und $M := xI_n - A$ und für jede Teilmenge $T \subseteq \underline{n} := \{1, \dots, n\}$ sei $M_T \in K(x)^{n\times n}$ gegeben durch die Spalten

$$(M_T)_{-,i} := \begin{cases} x(I_n)_{-,i} & i \notin T \\ -A_{-,i} & i \in T \end{cases}$$

Dann ist wegen der Multilinearität der Determinante und dem Laplaceschen Entwicklungssatz

$$\begin{aligned} \chi_\alpha(x) &= \det(M) \\ &= \textstyle\sum_{T\subseteq\underline{n}} \det(M_T) \\ &= \textstyle\sum_{T\subseteq\underline{n}} x^{n-|T|} \det(-A_{|T\times T}), \end{aligned}$$

wobei $\det(-A_{|T\times T})$ für $T = \emptyset$ als 1 zu interpretieren ist. Die erste Behauptung folgt nun leicht.

(2) $\mathrm{Kern}(\alpha - a\,\mathrm{id}_\mathcal{V}) \neq \{0\}$ ist äquivalent mit $\det(\alpha - a\,\mathrm{id}_\mathcal{V}) = 0$ und somit zu $\chi_\alpha(a) = 0$.

(3) Da $\chi_\alpha(x) \neq 0$, ist $xI - A \in K(x)^{n\times n}$ invertierbar. Sei die Matrix der Kofaktoren[4] von $xI_n - A$ gleich $M \in K(x)^{n\times n}$. Nach der Cramerschen Regel ist klar, dass die Einträge von M Polynome vom Grad $\leq n-1$ sind, so dass wir schreiben können:

$$M = M^{(0)} + xM^{(1)} + x^2M^{(2)} + \ldots + x^{n-1}M^{(n-1)}$$

[3] ZUNÄCHST NOCH IN K, DAHER IST HAMILTON-CAYLEY ist schärfer.

[4] Die Matrix der Kofaktoren von $A \in K^{n\times n}$ ist die nach der Cramerschen Regel eindeutig bestimmte Matrix B mit $(\det A)I_n = AB$.

mit $M^{(i)} \in K^{n\times n}$. Wir wissen wegen der Darstellung der Inversen nach der CRAMERschen Regel:

$$\begin{aligned}\chi_\alpha(x)I_n &= (xI_n - A)M \\ &= (xI_n - A)(M^{(0)} + xM^{(1)} + x^2M^{(2)} + \ldots + x^{n-1}M^{(n-1)}) \\ &= -AM^{(0)} + x(M^{(0)} - AM^{(1)}) + x^2(M^{(1)} - AM^{(2)}) + \cdots \\ &\quad + x^{n-1}(M^{(n-2)} - AM^{(n-1)}) + x^nM^{(n-1)}\end{aligned}$$

Ist nun $\chi_\alpha(x) = a_0 + a_1x + \cdots + a_{n-1}x^{n-1} + x^n$, so bekommen wir durch Vergleich der Matrixkoeffizienten der x^i aus der letzten Formel

$$a_0I_n = -AM^{(0)}, a_1I_n = M^{(0)} - AM^{(1)}, \ldots$$

$$a_{n-1}I_n = M^{(n-2)} - AM^{(n-1)}, I_n = M^{(n-1)}.$$

Multiplizieren wir jedes a_iI_n mit A^i und summieren auf, bekommen wir eine Teleskopreihe, d.h. $\chi_\alpha(A) = 0$. □

Folgerung 5.5.5. *Ist $p \in K[x]$ normiert vom Grad n und $A = M_p$ die Begleitmatrix von p, dann gilt $\chi_A = \mu_A = p$. Allgemeiner gilt wegen $\mu_A \mid \chi_A$, dass $\chi_A = \mu_A$, falls $\mathrm{Grad}(\mu_A) = n$.*

Ende Vorl. 27 25.01

5.5.b Die Zerlegung in Haupträume

Die folgende Bemerkung ist bloß eine Umformulierung von Satz 5.3.8.

Bemerkung 5.5.6 (Zerlegung in Haupträume)**.** Sei $\mathcal{V}$ ein K-Vektorraum endlicher Dimension und $\alpha \in \mathrm{End}(\mathcal{V})$. Schreibe das Minimalpolynom $\mu_\alpha = \prod_{i=1}^{\ell} p_i^{m_i}$ mit p_i irreduzibel, normiert und paarweise verschieden. Setzt man $q_i := \prod_{j\neq i} p_j^{m_j}$, so ist $\mathrm{ggT}(q_1, \ldots, q_\ell) = 1$. Schreibt man

$$1 = \sum_{i=1}^{\ell} a_iq_i \in K[x],$$

so sind die $\pi_i = a_i(\alpha)q_i(\alpha)$ mit α vertauschbare Projektionen, die folgendes erfüllen:

$$\pi_i \circ \pi_j = \delta_{ij}\pi_i,\ \mathrm{id}_\mathcal{V} = \pi_1 + \ldots + \pi_\ell.$$

Die Teilräume

$$\mathcal{U}_i := \mathrm{Bild}(\pi_i)$$

sind α-invariante Teilräume von $\mathcal{V}$, die wir auch **Haupträume** nennen wollen, genauer, wir nennen $\mathcal{U}_i$ den Hauptraum zum Faktor p_i. Es gilt

$$\mathcal{V} = \bigoplus_i \mathcal{U}_i$$

und für $\alpha_i := \alpha_{|\mathcal{U}_i}$ ist $\mu_{\alpha_i} = p_i^{m_i}$: Das Minimalpolynom μ_{α_i} von α_i teilt sicherlich $p_i^{m_i}$, da $p_i^{m_i}(\alpha)_{|\mathcal{U}_i} = 0$. Andererseits teilt μ_α das Produkt $\prod_{i=1}^{\ell} \mu_{\alpha_i}$ und somit muss $\mu_{\alpha_i} = p_i^{m_i}$ gelten. Daher gilt (leichte Übung)

$$\mathcal{U}_i = \mathrm{Kern}(p_i(\alpha)^{m_i}) = \mathrm{Bild}(q_i(\alpha)).$$

Ist B_i eine Basis von $\mathcal{U}_i$ ($1 \leq i \leq \ell$), so ist $B := (B_1, \ldots, B_\ell)$ eine Basis von $\mathcal{V}$ und

$${}^B\alpha^B = \mathrm{Diag}({}^{B_1}\alpha_1^{B_1}, \ldots, {}^{B_\ell}\alpha_\ell^{B_\ell}).$$

Satz 5.5.7. *Sei $\alpha \in \mathrm{End}(\mathcal{V})$ mit $\mu_\alpha = p^m$ für ein irreduzibles normiertes Polynom $p \in K[x]$. Dann gibt es $1 \leq m_1, m_2, \ldots, m_s \leq m$ und eine Basis B von $\mathcal{V}$, so dass*

$$ {}^B\alpha^B = \begin{pmatrix} M_{p^{m_1}} & * & \ldots & * \\ 0 & M_{p^{m_2}} & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \ldots & 0 & M_{p^{m_s}} \end{pmatrix} $$

Insbesondere gilt $d := \mathrm{Grad}(p) \mid \mathrm{Dim}(\mathcal{V}) =: n$ und $\chi_\alpha = p^c$ mit $c := m_1 + \cdots + m_s = \frac{n}{d} \geq m$. Schließlich gilt für mindestens ein $i \in \{1, \ldots, s\}$, dass $m_i = m$ ist.

Beweis. Eine solche Basis B erhält man, indem man zunächst ein $0 \neq X_1 \in \mathcal{V}$ wählt. Dieses X_1 erzeugt einen α-invarianten Teilraum $\mathcal{V}_1 \leq \mathcal{V}$, $\mathcal{V}_1 = \langle X_1, \alpha(X_1), \ldots, \alpha^{dm_1-1}(X_1)\rangle$ der Dimension dm_1 (Übung: Warum ist die Dimension ein Vielfaches von d?). Auf dem Faktorraum $\mathcal{V}/\mathcal{V}_1$ induziert α einen Endomorphismus dessen Minimalpolynom ein Teiler von μ_α ist. Dort wählt man wieder ein $0 \neq X_2 + \mathcal{V}_1$, bildet den von $X_2 + \mathcal{V}_1$ erzeugten α-invarianten Teilraum der Dimension dm_2 und setzt $\mathcal{V}_2 = \langle \alpha^t(X_1), \alpha^t(X_2) | t \in \mathbb{N}_0\rangle$, usw. Die Basis B ergibt sich dann als

$$ B = (X_1, \alpha(X_1), \ldots, \alpha^{dm_1-1}(X_1), X_2, \alpha(X_2), \ldots, \alpha^{dm_2-1}(X_2), \ldots, \alpha^{dm_s-1}(X_s)). \quad \square $$

Folgerung 5.5.8. *Sei $\mu_\alpha(x) = \prod_{i=1}^{\ell} p_i^{m_i}$ eine Zerlegung des Minimalpolynoms in normierte, irreduzible und paarweise verschiedene Polynome p_i. Dann gilt $\chi_\alpha(x) = \prod_{i=1}^{\ell} p_i^{c_i}$ mit $c_i \geq m_i$. Weiter gilt für die Dimension des p_i-***Hauptraumes*** $\mathcal{U}_i := \mathrm{Kern}(p_i^{m_i}(\alpha))$*

$$ \mathrm{Dim}(\mathcal{U}_i) = \mathrm{Dim}(\mathrm{Kern}(p_i^{m_i}(\alpha))) = c_i \,\mathrm{Grad}(p_i). $$

Insbesondere ist $\sum_i c_i \,\mathrm{Grad}(p_i) = \mathrm{Dim}\,\mathcal{V}$.

Übung 5.5.2. Zeige $\mathrm{Kern}(p_i^{m_i}(\alpha)) = \mathrm{Kern}(p_i^{c_i}(\alpha)) = \mathrm{Bild}(q_i(\alpha)) = \mathrm{Bild}(r_i(\alpha))$ wobei $r_i := \prod_{j \neq i} p_j^{c_j}$. Man hätte also die Zerlegung auch mit dem charakteristischen Polynom bekommen. Man zähle die Vorteile des Minimalpolynoms auf und entscheide, ob diese durch die explizite Formel für das charakteristische Polynom sowie durch die Dimensionsformel für die Haupträume aufgehoben werden.

Beispiel 5.5.9. Sei

$$ A := \begin{pmatrix} 1 & 1 & 1 & \cdot & \cdot & 1 \\ 1 & \cdot & \cdot & 1 & 1 & 1 \\ \cdot & \cdot & 1 & \cdot & 1 & \cdot \\ \cdot & \cdot & 1 & \cdot & 1 & 1 \\ \cdot & 1 & 1 & 1 & \cdot & \cdot \\ 1 & \cdot & \cdot & \cdot & \cdot & 1 \end{pmatrix} \in \mathbb{F}_2^{6\times 6} $$

Die Vektoren $E_1, AE_1, A^2E_1, A^3E_1, A^4E_1$ bilden die Spalten der Matrix

$$ M := \begin{pmatrix} 1 & 1 & 1 & 1 & \cdot \\ \cdot & 1 & \cdot & 1 & \cdot \\ \cdot & \cdot & \cdot & 1 & \cdot \\ \cdot & \cdot & 1 & 1 & 1 \\ \cdot & \cdot & 1 & 1 & 1 \\ \cdot & 1 & \cdot & 1 & \cdot \end{pmatrix} $$

Man liest ab $\mu_{A,E_1} = x^4 + x^2 + 1 = (x^2 + x + 1)^2$. Der Raum $\mathcal{V}_1 := \langle E_1, AE_1, A^2E_1, A^3E_1\rangle$ ist 4-dimensional und enthält nicht E_2. Die Vektoren E_2, $AE_2 = (1, 0, 0, 0, 1, 0)^{tr}$, $A^2E_2 =$

$(1,0,1,1,0,1)^{tr}$ erfüllen $E_2 + AE_2 + A^2E_2 = (0,1,1,1,1,1)^{tr} \in \mathcal{V}_1$. Setzt man also

$$T := \begin{pmatrix} 1 & 1 & 1 & 1 & \cdot & 1 \\ \cdot & 1 & \cdot & 1 & 1 & \cdot \\ \cdot & \cdot & \cdot & 1 & \cdot & \cdot \\ \cdot & \cdot & 1 & 1 & \cdot & \cdot \\ \cdot & \cdot & 1 & 1 & \cdot & 1 \\ \cdot & 1 & \cdot & 1 & \cdot & \cdot \end{pmatrix}$$

so erhält man

$$T^{-1}AT = \begin{pmatrix} \cdot & \cdot & \cdot & 1 & \cdot & 1 \\ 1 & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & 1 & \cdot & 1 & \cdot & \cdot \\ \cdot & \cdot & 1 & \cdot & \cdot & 1 \\ \cdot & \cdot & \cdot & \cdot & \cdot & 1 \\ \cdot & \cdot & \cdot & \cdot & 1 & 1 \end{pmatrix}.$$

Es gilt $\mu_A = p^2$, $\chi_A = p^3$ wobei $p = x^2 + x + 1 \in \mathbb{F}_2[x]$. Was muss man machen um eine Matrix T_1 zu finden mit $T_1^{-1}AT_1 = \mathrm{Diag}(M_{p^2}, M_p)$?

Ende Vorl. 28 30.01

Literaturverzeichnis

Index