Skript zu

Lineare Algebra für Informatiker und Statistiker

Inhaltsverzeichnis

1 Grundlagen

1.1 Logik

1.1.1 Was es bedeutet

Eine Aussage ist immer entweder wahr oder falsch.

Beispiel 1.1. 1) "Es gibt keine rote Kreide im Karton". Nach dem öffnen des Kartons kann dies eindeutig bewertet werden. Also ist dies eine Aussage, auch wenn wir nicht immer wissen, welchen Wahrheitswert sie besitzt. Entweder die Aussage oder ihre Negation ist wahr.

2) "Vollkornbrot schmeckt gut" ist keine Aussage, denn ihre Negation ist ebenso unentschieden, wie der Satz selbst. Es geht hierbei um eine Meinung, nicht um eine Aussage.

3) "n ist größer oder gleich Null" ($n \geq 0$) ist keine Aussage, denn es hängt von dem Wert von n ab, ob der Wahrheitswert wahr (w) oder falsch (f) ist. (Natürlich müssten zunächst auch alle anderen Begriffe wie "größer oder gleich" oder "Null" erläutert werden.)
Aber:

3a) "Für jede natürliche Zahl n gilt $n \geq 0$." ist eine wahre Aussage.

3b) "Für jede reelle Zahl n gilt $n \geq 0$" ist eine falsche Aussage, da -1 ebenfalls eine reelle Zahl ist, jedoch nicht ≥ 0.

3c) "Es gibt eine reelle Zahl n so dass $n \geq 0$" ist eine wahre Aussage, da 0 eine reelle Zahl ist.

Wir sehen also, dass eine Aussage wahr oder falsch unabhängig von äußeren Umständen sein muss.
Ein Prädikat hingegen kann Variablen (Unbekannte) enthalten, durch einsetzen von Werten für diese Variablen wird das Prädikat zu einer Aussage. Wir verwenden dafür die folgenden *Quantoren*

- $\forall$ ("für alle")
- $\exists$ ("es existiert")
- $\nexists$ ("es existiert nicht/kein")

Beispiel 1.2.

1. "$\forall x$ reell gilt $x^2 \geq 0$" ist wahr.

2. "$\exists x$ ganze Zahl so dass $x^2 = 5$ ist falsch.

3. "$\nexists$ blaue Tomaten" oder "$\forall$ Tomaten x ist x nicht blau" (hier muss man prinzipiell "blau" und "Tomate" noch genau definieren). Diese Aussage ist vermutlich war, jedoch können wir es derzeit nicht Beweisen.

Bemerkung. $\exists x P(x)$ kann "gezeigt"/bewiesen werden, indem wir ein x angeben, das die Eigenschaft $P(x)$ besitzt. Dies bezeichnet man dann als einen konstruktiven Beweis. Es gibt aber auch Fälle, wo andere Beweise existieren. Die Aussage kann jedoch allein durch Angabe von (beliebig vielen) Beispielen *nicht widerlegt* werden.

Bemerkung. Analog kann $\forall x P(x)$ nicht allein durch Prüfen von Beispielen für x bewiesen werden (ausser die Zahl der möglichen x ist endlich *und klein genug*, damit dies auch praktisch durchführbar ist). Sie kann aber durch Angabe eines *Gegenbeispieles* widerlegt werden.

Beispiel 1.3. $\forall x, y$ ganzzahlig gilt $x + y = y + x$. Wenn wir für 1000000000 Werte von x und y überprüfen, ob diese Aussage für diese Werte von x und y stimmt, so ist dies noch immer kein Beweis, auch wenn wir intuitiv dadurch überzeugt sind, dass die Aussage wahr ist. Für einen Beweis ist es notwendig, dass wir "verstehen", was eine ganze Zahl ist und wie die Addition $+$ auf ganzen Zahlen definiert wird, also was $x + y$ formal bedeutet.

Beispiel 1.4. "$\forall x$ reell positiv gilt $x \log x < 1$ ist falsch, denn wir können als Gegenbeispiel $x = e$ betrachten, da gilt $\log e = 1$ und $e > 1$, also damit $e \log e > 1$.
Aber: " $\forall x$ reell positiv gilt $x \log x \neq 2$ kann zwar auch durch Angabe einer Lösung der Gleichung $x \log x = 2$ widerlegt werden, jedoch kann man dieses x nicht explizit angeben. Man kann jedoch beweisen, dass es sie gibt. Dies gehört aber zum Gebiet der Analysis.

Fazit: Der Wahrheitswert einer Aussage hängt nicht davon ab, wie einfach oder schwierig sie zu beweisen ist. Es gibt Aussagen, deren Wahrheitswert unbekannt ist, für die es also weder einen Beweis, noch einen Beweis ihrer Negation bekannt ist. Dies führt uns zu den Begriffen

- "Vermutung", falls man Behauptet, ein Beweis sollte irgendwann zu finden sein.
- "Axiome", falls man *annimmt*, dass sie wahr ist, und man schätzt, dass sie aus den bestehenden Aussagen nicht *bewiesen werden kann*. Die Axiome sind Grundbausteine der Logik/Mathematik.

Beweis: Es gibt grundsätzlich zwei Möglichkeiten, eine Aussage zu beweisen.

- Logik: Beweis durch strenge Regeln, die auf bestimmte Zeichenreihen bestehend aus einer gewissen Menge an Symbolen angewandt werden können. Dieses Verfahren ist sehr gut für Computer geeignet.

- Mathematik (außerhalb der Logik): Man überzeugt sich davon, dass ein Beweis im Sinne der strengen Logik *möglich* ist und schreibt eine für Menschen verständliche Fassung davon auf.

Wir werden später noch Beispiele dazu angeben, aber hierfür benötigen wir vorerst einen Formalismus, um das zu ermöglichen.

1.1.2 Regeln der Logik

Sei A eine Aussage. Dann ist $\neg A$ ("non A" = "nicht A" = "Negation von A") wahr genau dann, wenn A falsch ist und falsch genau dann, wenn A wahr ist.
$A \vee B$ ("A oder B") ist genau dann wahr, wenn A wahr ist, oder B wahr ist (oder beide). Man bezeichnet $A \vee B$ auch als die Disjunktion von A und B.
$A \wedge B$ ("A und B") ist genau dann wahr, wenn sowohl A als auch B wahr sind. Man nennt $A \wedge B$ auch die Konjunktion von A und B.
Wir kommen nun zum Begriff der Wahrheitstafel: Wir stellen sie wie folgt dar:

A \ B	w	f
w		
f		

wobei man in die frei gebliebenen Felder die Wahrheitswerte von $A \vee B$, $A \wedge B$ etc. eintragen kann, zum Beispiel für $A \wedge B$:

A \ B	w	f
w	w	f
f	f	f

Häufiger stellt man sie aber in der folgenden Form dar:

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

wobei wir $A \Rightarrow B$ ("Implikation") als "A impliziert B" oder als "aus A folgt B" verstehen wollen und $A \Leftrightarrow B$ ("Äquivalenz") als "A gilt genau dann, wenn B gilt".

Übung. $A \Leftrightarrow B$ hat die gleichen Wahrheitswerte wie $(A \Rightarrow B) \wedge (B \Rightarrow A)$.

Bedeutung: Wenn man $A \Leftrightarrow B$ beweisen will, so beweist man oft $A \Rightarrow B$ und $B \Rightarrow A$. In nur wenigen (einfachen) Fällen ist es möglich, die Äquivalenz zweier Aussagen ohne diese Trennung der Implikationsrichtungen zu beweisen.

Bemerkung. $A \Rightarrow B$ ist äquivalent zu $\neg A \vee B$ (Übung), d.h. wenn A falsch ist, so stimmt die Implikation $A \Rightarrow B$ *immer*, d.h. etwas Falsches kann *alles* implizieren. $A \Rightarrow B$ kann also wahr sein, ohne dass A oder B wahr sind.

Bemerkung. Der Schwerpunkt bei einer Aussage vom Typ $A \Rightarrow B$ liegt in dem "Pfeil" (der eine asymmetrische logische Verbindung zwischen A und B darstellt:"falls A, dann B"), und nicht in dem tatsaechlichen Wahrheitswert der Aussagen A und B. Die Implikationen sind die Grundbausteine eines Beweises (siehe insbesondere Satz 1.2 unten). Die Äquivalenzen ebenso (eine Äquivalenz lässt sich aber in einem Paar von Implikationen zerlegen - Satz 1.2 *a).*)

Satz 1.1.

(a) $A \vee B \Leftrightarrow B \vee A$

(b) $A \wedge B \Leftrightarrow B \wedge A$

(c) $(A \vee B) \vee C \Leftrightarrow A \vee (B \vee C)$

(d) $(A \wedge B) \wedge C \Leftrightarrow A \wedge (B \wedge C)$

(e) $(A \vee B) \wedge C \Leftrightarrow (A \wedge C) \vee (B \wedge C)$

(f) $(A \wedge B) \vee C \Leftrightarrow (A \vee C) \wedge (B \vee C)$

(g) $\neg(A \wedge B) \Leftrightarrow \neg A \vee \neg B$

(h) $\neg(A \vee B) \Leftrightarrow \neg A \wedge \neg B$

(i) $\neg(\neg A) \Leftrightarrow A$

Beweis. Wir beweisen hier nur exemplarisch (e):

A	B	$A \vee B$	C	$(\mathbf{A} \vee \mathbf{B}) \wedge \mathbf{C}$	$(A \wedge C)$	$(B \wedge C)$	$(\mathbf{A} \wedge \mathbf{C}) \vee (\mathbf{B} \wedge \mathbf{C})$
w	w	w	w	w	w	w	w
w	w	w	f	f	f	f	f
w	f	w	w	w	w	f	w
w	f	w	f	f	f	f	f
f	w	w	w	w	f	w	w
f	w	w	f	f	f	f	f
f	f	f	w	f	f	f	f
f	f	f	f	f	f	f	f

Man kann also an der fünften und letzten Spalte ablesen, dass sie für alle Belegungen von A, B, C den gleichen Wahrheitswert liefern, also sind sie äquivalent. Der Rest der Regeln bleibt als Übung zu beweisen. □

Es gelten die folgenden Regeln:

$$\forall x\colon A(x) \wedge B(x) \Leftrightarrow \forall x\colon A(x) \wedge \forall x\colon B(x)$$
$$\exists x\colon A(x) \vee B(x) \Leftrightarrow \exists x\colon A(x) \vee \exists x\colon B(x)$$

Jedoch lassen sich die folgenden nicht zerlegen:

$$\forall x\colon A(x) \vee B(x)$$
$$\exists x\colon A(x) \wedge B(x)$$

Beispiel 1.5. Wir defineren $A(x) :\Leftrightarrow x = 1$ und $B(x) :\Leftrightarrow x \neq 1$. Dann gilt $\forall x\colon A(x) \vee B(x) \Leftrightarrow w$. Es gilt jedoch $\forall x\colon A(x) \Leftrightarrow f$ und $\forall x\colon B(x) \Leftrightarrow f$.

Vorrangsregeln: Um die Schreibweise zu erleichtern und die Anzahl der Klammern zu reduzieren, wir legen in logischen Konstruktionen die folgende Vorrangsregel fest: $\neg$ wird zuerst angewandt, dann $\wedge$ und $\vee$, und danach $\Rightarrow$ und $\Leftrightarrow$.

Satz 1.2.

a) $(A \Rightarrow B) \wedge (B \Rightarrow A) \Leftrightarrow (A \Leftrightarrow B)$

b) $(A \Leftrightarrow B) \wedge (B \Leftrightarrow C) \Rightarrow (A \Leftrightarrow C)$

c) $(A \Rightarrow B) \wedge (B \Rightarrow C) \Rightarrow (A \Rightarrow C)$

Dies lässt sich auch verallgemeinern zu

b_n) $(A_1 \Leftrightarrow A_2) \wedge (A_2 \Leftrightarrow A_3) \wedge ... \wedge (A_{n-1} \Leftrightarrow A_n) \Rightarrow (A_1 \Leftrightarrow A_n)$

c_n) $(A_1 \Rightarrow A_2) \wedge (A_2 \Rightarrow A_3) \wedge ... \wedge (A_{n-1} \Rightarrow A_n) \Rightarrow (A_1 \Rightarrow A_n)$

Die Aussagen (a)-(c) lassen sich über Wahrheitstafel prüfen (Übung), (b_n) und (c_n) benötigen *Induktion* (siehe unten).

Bedeutung: *c).* und seine allgemeinere Form *c_n).* stellt eine generelle Strategie für (komplizierte) Beweise dar: durch Folgen von einfachen Implikationen können also andere bewiesen werden, also die Wahrheitswerte einer Aussage (zum Beispiel A_n aus c_n) aus den Wahrheitswerten von A_1 können "stufenweise" hergeleitet werden. Ähnliches gilt für *b).*, bzw. *b_n).*, wobei die Äquivalenzen laut *a).* als "doppelte Implikationen" gesehen werden können.

Satz 1.3.

a) $(A \Rightarrow B) \Leftrightarrow (\neg A \vee B)$

b) $(A \Rightarrow B) \Leftrightarrow (\neg B \Rightarrow \neg A)$

Verneinung (Negation) von Implikationen: $\neg(A \Rightarrow B)$ ist keine Implikation, denn man kriegt in der Wahrheitstafel $1\times$ den Wert wahr, jedoch $3\times$ den Wert falsch, was nicht der Wahrheitstafel einer Implikation entspricht.

Beweistechnik für $A \Rightarrow B$:

1. Direkter Beweis: Man nimmt an, A sei wahr und benutzt dann andere Axiome oder bekannte Sätze und beweist damit, B sei wahr. Insbesondere mithilfe von b_n und c_n aus dem Satz oben.

2. Indirekter Beweis: Man nimmt an, B sei falsch und beweist, dass dann auch A falsch sein muss. Dies entspricht genau der Aussage des obigen Satzes.

3. Widerspruchsbeweis: Man nimmt an, A sei wahr und B sei falsch. Man leitet dann aus diesen Annahmen einen Widerspruch her, woraus man schließen kann, dass die Annahme, dass B nicht gelte, falsch wahr. Das heißt man zeigt $A \wedge \neg B$ ist falsch, d.h. $\neg(A \wedge \neg B) \Leftrightarrow \neg A \vee B \Leftrightarrow (A \wedge B)$ ist also wahr.

1.2 Mengenlehre

Eine *Menge* A besteht aus ihren *Elementen* p. Wir schreiben $p \in A$ ("p ist Element von A"), falls p zu A gehört, oder auch $A \ni p$ ("A enthält p"). Falls p nicht zu A gehört, so schreibt man $p \notin A$. Beispiele für Mengen sind die folgenden:

$$\begin{aligned}
\mathbb{N} &= \{0, 1, 2, ...\} \text{ Menge der natürlichen Zahlen} \\
\mathbb{N}^* &:= \mathbb{N} \setminus \{0\} = \{n \in \mathbb{N} \mid n \neq 0\} \text{ Menge der nichtnull natürlichen Zahlen} \\
\mathbb{Z} &= \{..., -2, -1, 0, 1, 2, ...\} \text{ Menge der ganzen Zahlen} \\
\mathbb{Q} &= \left\{ \frac{p}{q} \mid p \in \mathbb{Z}, q \in \mathbb{N}^* \right\} \text{ Menge der rationalen Zahlen} \\
\mathbb{R} &= \text{ Menge der reellen Zahlen (siehe Analysis)}
\end{aligned}$$

Es gilt:

$$\mathbb{N} \subsetneq \mathbb{Z} \subsetneq \mathbb{Q} \subsetneq \mathbb{R}$$

Ein weiteres Beispiel für eine Menge ist $\varnothing$ = leere Menge. Sie enthält kein Element.

Bemerkung. Die Reihenfolge oder das mehrfache Auftreten eines Elementen in einer Menge ist irrelevant. Das heißt es gilt, zum Beispiel:

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

Definition 1.1.

(a) $A = B \ :\Leftrightarrow \ \forall x\colon \ (x \in A \Leftrightarrow x \in B)$.

(b) $A \subseteq B \;:\Leftrightarrow\; \forall x\colon (x \in A \Rightarrow x \in B)$. Es heißt dann A *Teilmenge* von B, und $A \subseteq B$ ist die *Inklusion* der Mengen A in B.

(c) $\mathcal{P}(A) \coloneqq \{M \mid M \subseteq A\}$. Es heißt $\mathcal{P}(A)$ die Potenzmenge von A.

Die Symbole $:\Leftrightarrow$ und $\coloneqq$ zeigen, dass die entsprechende Äquivalenz, bzw. Gleichung eine Definition darstellt. Bei manchen Authoren (und in dieser Vorlesung) wird $\subseteq$ auch $\subset$ geschrieben. Andere Authoren dagegen benutzen das Symbol $\subset$ nur bei Inklusionen ohne Gleichheit.

Beispiel 1.6.

$$\mathcal{P}(\varnothing) = \{\varnothing\}$$
$$\mathcal{P}(\{\varnothing\}) = \{\varnothing, \{\varnothing\}\}$$

Satz 1.4. *Es gilt:*

(a) $A \subseteq B \wedge B \subseteq A \Leftrightarrow A = B$

(b) $A \subseteq B \wedge B \subseteq C \Rightarrow A \subseteq C$

Definition 1.2. $A \cap B \coloneqq \{x \mid x \in A \wedge x \in B\}$ heißt der Schnitt (oder Durchschnitt, Schnittmenge) von A und B.
$A \cup B \coloneqq \{x \mid x \in A \;\vee\; x \in B\}$ heißt die Vereinigung von A und B.
$A \setminus B \coloneqq \{x \mid x \in A \;\wedge\; x \notin B\}$ heißt (Mengen-)Differenz von A und B, oder auch "A ohne B".

Satz 1.5. *Seien A, B, C Mengen. Dann gelten folgende Regeln:*

(a) $A \cup B = B \cup A$

(b) $A \cap B = B \cap A$
(*Kommutativität der Vereinigung (a), bzw. des Durchschnitts (b)*)

(c) $(A \cup B) \cup C = A \cup (B \cup C)$

(d) $(A \cap B) \cap C = A \cap (B \cap C)$
(*Assoziativität der Vereinigung (c), bzw. des Durchschnitts (d)*)

(e) $(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$

(f) $(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$
(*Distributivität der Vereinigung bzgl. des Durchschnitts (e), bzw. des Durchschnitts bzgl. der Vereinigung (f)*)

Bemerkung. Es gilt: $A \cup A = A,\; A \cap A = A,\; A \cup \varnothing = A,\; A \cap \varnothing = \varnothing$.

Komplemente: Seien A, M Mengen und $A \subseteq M$. Wir bezeichnen durch $C_A := M \setminus A$.

Satz 1.6. *Seien A, B Teilmengen von M. Dann gilt:*

(a) $C_{(A \cup B)} = C_A \cap C_B$

(b) $C_{(A \cap B)} = C_A \cup C_B$

(c) $C_{C_A} = A$

1.2.1 Quantoren über Mengen

$$\forall x \in M \colon P(x) :\Leftrightarrow \forall x \colon (x \in M \Rightarrow P(x))$$
$$\exists x \in M \colon P(x) :\Leftrightarrow \exists x \colon (x \in M \wedge P(x))$$

Negation von Quantoren:

$$\neg(\exists x \colon P(x)) \;\Leftrightarrow\; \forall x \colon \neg P(x)$$
$$\neg(\forall x \colon P(x)) \;\Leftrightarrow\; \exists x \colon \neg P(x)$$

Für Quantoren über Mengen:

$$\neg(\forall x \in M \colon P(x)) \;\Leftrightarrow\; \exists x \in M \colon \neg P(x)$$
$$\neg(\exists x \in M \colon P(x)) \;\Leftrightarrow\; \forall x \in M \colon \neg P(x)$$

Beispiel 1.7. Seien $k \in \mathbb{N} \setminus \{0\}$ und $n \in \mathbb{Z}$. Man sagt, dass "k teilt n" oder "n teilbar durch k", oder - in Zeichen -

$$k \mid n$$

genau dann, wenn gilt:

$$\exists p \in \mathbb{Z} \colon n = kp.$$

Anders gesagt $n \in k\mathbb{Z}$. Die Negation hiervon ist

$$\neg(\exists p \in \mathbb{Z} \colon n = kp) \Leftrightarrow \forall p \in \mathbb{Z} \colon n \neq kp.$$

Alternativ natürlich $n \notin k\mathbb{Z}$. Man sagt dann, dass n nicht durch k teilbar sei. Gerade Zahlen sind solche, die durch 2 teilbar sind, ungerade Zahlen die, die nicht durch 2 teilbar sind.

1.3 Das Prinzip der vollständigen Induktion

Es gibt das sogenannte *Wohlordnungszprinzip* für $\mathbb{N}$. Es besagt: $\forall A \subseteq \mathbb{N}$ nicht leer besitzt ein minimales Element. Mit Quantoren:

$$\forall A \subseteq \mathbb{N}, A \neq \varnothing \colon \exists n \in A \colon \forall k \in A \colon n \leq k$$

Man schreibt dann $n \coloneqq \min A$.

Bemerkung. Das kleinste Element ist eindeutig bestimmt.

Man kann manchmal auch für $A \subseteq \mathbb{Z}, \mathbb{Q}, \mathbb{R}$ ein kleinstes Element $\min A$ und ein größtes Element $\max A$ definieren. Diese müssen jedoch nicht existieren. Solche Beispiele werden oft in der Analysis gesehen.

Satz 1.7 ((Vollständige) Induktion)**.** *Sei $P(\cdot)$ ein Prädikat, das für jedes $n \in \mathbb{N}$ eine Aussage $P(n)$ bildet. Wenn gilt:*

1) *$P(0) \Leftrightarrow w$ (Induktionsanfang)*

2) *$\forall n \in \mathbb{N} \colon P(n) \Rightarrow P(n+1)$ (Induktionsschritt).*

so ist $P(n)$ wahr $\forall n \in \mathbb{N}$. (Induktionsschluss)

Beweis. Sei $M \coloneqq \{k \in \mathbb{N} \mid P(k) \text{ falsch}\}$. Falls $M \neq \varnothing$, sei $n_0 \coloneqq \min M$. $n_0 = 0$ kann nicht sein, da dies 1) widerspricht. Falls $n_0 > 0$, so ist $n_0 - 1 \in \mathbb{N}$ und es gilt damit $P(n_0 - 1)$, da n_0 das kleinste Element war, für das $P(n)$ nicht gilt. Mit 2) folgt dann, dass $P(n_0)$ wahr ist. Dies ist ein Widerspruch, also muss die Annahme $M \neq \varnothing$ falsch gewesen sein. □

Bemerkung (Variante des Induktionsschemas)**.** Sei $n_0 \in \mathbb{Z}$ fest. Es gelte:

1) $P(n_0)$ ist wahr.

2) $\forall n \in n_0 + \mathbb{N} \colon P(n_0) \wedge P(n_0 + 1) \wedge .... \wedge P(n) \Rightarrow P(n+1)$.

Dann gilt $(\forall n \in n_0 + \mathbb{N} \colon P(n))$ ist wahr.

Beispiel 1.8. Wir wollen zeigen: $1 + 2 + ... + (n-1) = \frac{n(n-1)}{2}$. Dies ist unser $(P(n))$. Wir wenden Induktion an:

$P(1)$: $0 = 0$ ist wahr

$P(n) \Rightarrow P(n+1)$: $1 + 2 + 3 + ... : (n-1) + n \overset{P(n)}{=} \frac{n(n-1)}{2} + n = \frac{n(n+1)}{2}$. Also ist $P(n+1)$ wahr.

Bemerkung. Im Induktionsschritt $P(n) \Rightarrow P(n+1)$ wird oft $P(n)$ *Induktionssvoraussetzung* und $P(n+1)$ *Induktionsbehauptung* genannt.

Definition 1.3. Das Symbol fuer die mehrfache Summe $\sum_{k=0}^{n} a_k$ wird wie folgt (rekursiv) definiert:

1) $\sum_{k=0}^{-1} a_k \coloneqq 0$

2) $\forall n \in \mathbb{N}\colon\ \sum_{k=0}^{n} a_k \coloneqq (\sum_{k=0}^{n-1} a_k) + a_n$

Man schreibt dann also $1 + 2 + 3 + ... + n$ genauer als $\sum_{k=1}^{n} k$.

Definition 1.4. Analog zu $\sum_{k=0}^{n} a_k$ definieren wir das Symbol fuer mehrfache Produkte $\prod_{k=0}^{n}$ rekursiv wie folgt:

1) $\prod_{k=0}^{-1} a_k = 1$

2) $\prod_{k=0}^{n} a_k = (\prod_{k=0}^{n-1} a_k) \cdot a_n$

Hier, wie bei der Summen, braucht der Index nicht unbedingt die Werte von 0 nach n zu nehmen, sondern kann von n_0 nach $m_0 \geq n_0$, mit $n_o, m_o \in \mathbb{Z}$:

$$\sum_{k=n_0}^{m_0} a_k \coloneqq a_{n_0} + a_{n_0+1} + \cdots + a_{m_0},\ \text{bzw.}\ \prod_{k=n_0}^{m_0} a_k \coloneqq a_{n_0} \cdot a_{n_0+1} \cdot \cdots \cdot a_{m_0}.$$

Häufig wird $n_0 = 1$ getroffen. Man kann auch $m_0 < n_0$ zulassen, dann wird die obige Summe als Null definiert, bzw. das obige Produkt gleich 1 per Definition gesetzt. Im Allgemeinen sind die Symbole, die für die Indizen verwendet werden, unwichtig:

$$\sum_{k=1}^{n} a_k = \sum_{\mathbf{i}=1}^{n} a_{\mathbf{i}},$$

wichtig ist, dass der *laufende Index* im Summenzeichen (also k, bzw. $\mathbf{i}$ unter der Summenzeichen) ist derjenige, der in der Beschreibung der Summenterme verwendet wird (Analog für Produkte). Man soll also die (ungenauere, aber) anschaulichere Darstellung

$$a_1 + a_2 + \cdots + a_n,\ \text{bzw.}\ b_{-1} \cdot b_0 \cdot \cdots \cdot b_{n+2}$$

durch Auswertung der laufende Indizen in dem vorgeschriebenen Wertebereich hergeleitet werden:

$$\sum_{k=1}^{n} a_k,\ \text{bzw.}\ \prod_{i=-1}^{n+2} b_i.$$

Daher sind z.B. die folgende Summen gleich:

$$\sum_{k=1}^{n} a_k = \sum_{j=0}^{n-1} a_j,$$

dagegen ist $\sum_{k=1}^{n-1} a_k$ von den obigen verschieden (es sei denn, die Differenz, die aus dem "weggelassenen" Term a_n besteht, Null ist). Beim Vergleich zweier Summen ist also nicht die Erscheinung der Terme (also a_k, oder a_k, oder b_l) entscheidend, sondern die Werte, die diese Terme nehmen, wenn man die laufende Indizen jeweils durch die vorgeschriebene (d.h. durch die im Summenzeichen angegebene Wertgrenzen "von (untere Grenze) nach (obere Grenze)") Wertebereiche ersetzt. (Analog für Produkte)

Beispiel 1.9 (rekursive Definition der Fibonacci-Zahlen)**.** Wir definieren: $F_0 := 1, F_1 := 1$ und $F_{n+2} := F_n + F_{n+1}$

Übung. Man beweise: $\exists n_0 \in \mathbb{N}\colon \forall n \geq n_0 \colon F_n > n^2$.

1.3.1 Teilbarkeit und Division mit Rest

Definition 1.5. Sei $x \in \mathbb{R}$. Der *Betrag* (Absolutwert) von x ist die folgende nichtnegative reele Zahl:

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

Intuitiv ist der Betrag "die Zahl ohne Vorzeichen" oder "die Entfernung (der Abstand) zu 0".

Satz 1.8 (Eigenschaften der Teilbarkeit)**.** *Für $a, b, m, l, n \in \mathbb{Z}$ gilt:*

(a) $m \mid a \ \wedge \ a \neq 0 \ \Rightarrow \ m \leq |a|$

(b) $m \mid n \ \wedge \ n \mid m \ \Rightarrow \ m = n$

(c) $l \mid n \ \wedge \ n \mid m \ \Rightarrow \ l \mid m$

(d) $m \mid a \ \wedge \ m \mid b \ \Rightarrow \ m \mid \alpha a + \beta b$ *für* $\alpha, \beta \in \mathbb{Z}$

(e) $m \mid a \ \wedge \ n \mid b \ \Rightarrow \ m \cdot n \mid a \cdot b$

Satz 1.9 (Division mit Rest)**.** *Es sei $a \in \mathbb{Z}, b \in \mathbb{N}^*$. Dann*

$$\exists!\, q \in \mathbb{Z}, \ \exists!\, r \in \{0, 1, ..., b-1\} : a = bq + r$$

Bemerkung. Das Symbol $\exists!$ bedeutet "es gibt genau ein", oder "es gibt ein [...] und es ist eindeutig". Genauer gesagt,

$$(\exists!\, x\colon P(x)) :\Leftrightarrow (\exists x\colon P(x)) \wedge \big(\forall x_1, x_2\colon (P(x_1) \wedge P(x_2) \Rightarrow (x_1 = x_2))\big)$$

Beweis.

Existenz: Wir definieren $M := \{a - bq \mid q \in \mathbb{Z} \wedge a - bq \geq 0\}$. Es gilt $M \subseteq \mathbb{N}$ und $M \neq \varnothing$, denn $a - (-|a|\, b) \in M$. Sei $r := \min M$. Dann gilt $r \geq 0$. Falls $r \geq b$, so gilt $r \geq r - b \geq 0$ und $r - b \in M$. Dies kann nicht sein. Also gilt $r \in \{0, 1, ..., b-1\}$. Nach Definition von r gibt es dann ein $q \in \mathbb{Z}$ so dass $a = qb + r$.

Eindeutigkeit: Falls $a = bq + r = bq' + r'$, so gilt $b \mid (r - r')$. Also $|r - r'| \geq b$ oder $r - r' = 0$. Aber es gilt $r, r' \in \{0, 1, ..., b-1\}$. Also $|r - r'| \leq b - 1$ und damit $r = r'$. Dies impliziert dann ebenfalls $q_1 = q_2$, weil $b(q_1 - q_2) = 0$ und $b \neq 0$.

□

Definition 1.6. $p \in \mathbb{N}$ heißt Primzahl, falls $p \geq 2$ und $\forall n \in \mathbb{N}$: $(n \mid p \Rightarrow n = 1 \vee n = p$.

Satz 1.10. *$\forall n \in \mathbb{N}, n \geq 2$ gilt: n ist das Produkt von Primzahlen.*

Beweis. Durch Induktion über n. Wir definieren hierzu $P(n) :\Leftrightarrow n = p_1 \cdot .... \cdot p_k$ für irgendwelche Primzahlen p_i.

$n = 2$**:** Es ist $P(2)$ offenbar wahr.

$P(2) \wedge P(3) \wedge .... \wedge P(n-1) \Rightarrow P(n)$**:** Falls $n > 2$ Primzahl ist, so $n = n$ und damit $P(n)$ wahr. Falls n keine Primzahl ist, so schreibe $n = n_1 \cdot n_2$ mit $1 < n_1, n_2$. Wir wissen dann $P(n_1) \wedge P(n_2)$, also

$$n_1 = \prod_{j=1}^{k_1} p_j \quad \wedge \quad n_2 = \prod_{i=1}^{k_2} q_i \Rightarrow n = n_1 n_2 = \prod_{j=1}^{k_1} p_j \cdot \prod_{i=1}^{k_2} q_i$$

ist ein Produkt von Primzahlen $p_1 \cdot ... \cdot p_{k_1} \cdot q_1 \cdot ... \cdot q_{k_2}$.

□

Satz 1.11 (Euklid)**.** *Es gibt unendlich viele Primzahlen.*

Beweis. Sei $P = \{p \in \mathbb{N} \mid p \text{ Primzahl}\}$. Falls P endlich ist, so ist $P = \{p_1, ..., p_k\}$ mit $k := \#P = |P|$ die Anzahl an Elementen in P. Sei nun $N = (\prod_{j=1}^{k} p_j) + 1$. Dann ist N entweder selbst eine Primzahl (Widerspruch), oder N ist ein Produkt von Primzahlen. Aber es gilt $p_j \nmid N \quad \forall j \in \{1, , , k\}$, da sonst $p_j \mid N$ und $p_j \mid \prod_{i=1}^{k} p_i$ gelten müsste. Dann würde aber auch mit obigen Regeln über die Teilbarkeit folgen: $p_j \mid N - \prod_{i=1}^{k} p_i$, also $p_j \mid 1$. Dies kann aber nicht sein. □

Definition 1.7 (Größter gemeinsamer Teiler, Teilerfremdheit)**.**

a) Seien $a, b \in \mathbb{Z}, a \neq 0 \vee b \neq 0$. Dann definiert man

$$\text{ggT}(a, b) := \max \{m \in \mathbb{N} \mid \ m \mid a \ \wedge \ m \mid b \ \}$$

b) Falls $\mathrm{ggT}(a,b) = 1$, so heißen a, b teilerfremd.

Beispiel 1.10. Es gilt $\mathrm{ggT}(6,16) = 2$, $\mathrm{ggT}(2,5) = 1$, also sind 2 und 5 teilerfremd. Seien allgemein $p \neq q$ zwei Primzahlen. Dann gilt p^{k_1}, q^{k_2} sind teilerfremd.

Satz 1.12 (Bézout)**.** *Seien $a, b \in \mathbb{Z}$ und $a \neq 0 \vee b \neq 0$. Dann*

$$\mathrm{ggT}(a,b) = \min\left\{ax + by \mid x \in \mathbb{Z} \ \wedge \ y \in \mathbb{Z} \ \wedge \ ax + by \in \mathbb{N}^*\right\}$$

Beweis. Wir setzen $M = \left\{ax + by \mid x \in \mathbb{Z} \ \wedge \ y \in \mathbb{Z} \ \wedge \ ax + by \in \mathbb{N}^*\right\}$. Dann gilt $\varnothing \neq M \subseteq \mathbb{N} \Rightarrow \exists m \coloneqq \min M,\ m = ax_0 + by_0$.

1. Wir zeigen: $\mathrm{ggT}(a,b) \leq m$. Das folgt aus dem Satz über die Teilbarkeit (e) und (a).

2. Wir zeigen: $m \leq \mathrm{ggT}(a,b)$. Division mit Rest liefert: $a = mq + r, q \in \mathbb{Z}, 0 \leq r < m$. Also $r = a - m(ax_0 + by_0) \in \mathbb{N}^*$, also falls $r \neq 0$, $r \in M$ und $r < m$. Widerspruch. Also $r = 0$ und $m \mid a$. Analog $m \mid b$. Dann gilt $m \leq \mathrm{ggT}(a,b)$, da $\mathrm{ggT}(a,b)$ die größte Zahl ist, die sowohl a als auch b teilt. Also folgt zusammen $m = \mathrm{ggT}(a,b)$.

□

Korollar 1.13. $k \mid a \ \wedge \ k \mid b \ \Rightarrow \ k \mid \mathrm{ggT}(a,b)$.

1.4 Kartesische Produkte

Seien A, B Mengen. Ein *Paar* (a,b) mit $a \in A, b \in B$ ist charakterisiert durch die Eigenschaft

$$(a,b) = (a',b') :\Leftrightarrow a = a' \ \wedge \ b = b'$$

Zum Beispiel kann das Paar (a,b) wie folgt definiert werden: $(a,b) \coloneqq \left\{\{a\}, \{a,b\}\right\} \subseteq \mathcal{P}(A \cup B)$ bzw $\in \mathcal{P}(\mathcal{P}(A \cup B))$.

Wir können damit definieren $A \times B \coloneqq \left\{(a,b) \mid a \in A, b \in B\right\}$ heißt *das kartesische Produkt* der Mengen A und B.

Definition 1.8. Seien $A_1, ..., A_k$ Mengen für $k \geq 2$. Wir setzen

$$A_1 \times ... \times A_k \coloneqq \left\{(a_1, ..., a_k) \mid \forall i \in \{1, ..., k\} : a_i \in A_i\right\}$$

heißt das kartesische Produkt der Mengen $A_1, .., A_k$. Die "k-Tupel" $(a_1, .., a_k)$ und $(b_1, ..., b_k)$ heißen gleich $:\Leftrightarrow a_1 = b_1 \ \wedge \ ... \ \wedge \ a_k = b_k$. Falls $A_i = A$ für $i = 1, .., k$, so schreibt man $A^k \coloneqq A \times A \times ... \times A$ (k-Mal). Es ist dann $A^1 \coloneqq A$.

1.5 Relationen

Definition 1.9. Seien A, B Mengen und $R \subseteq A \times B$. Dann heißt $\mathcal{R} \coloneqq (A, B, R)$ eine Relation zwischen (Elementen von) A und (Elementen von) B. Ist $A = B$, so heißt $\mathcal{R}$ dann Relation auf A. Man schreibt $x \sim y :\Leftrightarrow (x, y) \in R$, oder auch $x \sim_R y$. Man schreibt auch $(A, \sim_R)$ oder $(A, \sim)$ anstatt (A, A, R).

Beispiel 1.11. Man kann die Relationen $\underbrace{(\mathbb{R}, \leq)}_{a}, \underbrace{(\mathbb{R}, <)}_{b}, \underbrace{(\mathbb{R}, \geq)}_{c}, \underbrace{(\mathbb{R}, =)}_{d}$ betrachten. Es gilt dann beispielweise:

$$R_a = \Big\{(x, y) \in \mathbb{R}^2 \mid x \leq y\Big\}$$
$$R_d = \big\{(x, x) \mid x \in \mathbb{R}\big\}$$

1.5.1 Eigenschaften von Relationen

Definition 1.10.

a) $(X, \sim)$ heißt reflexiv, falls $\forall x \in X : x \sim x$.

b) $(X, \sim)$ heißt symmetrisch, falls $\forall x, y \in X : (x \sim y \Rightarrow y \sim x)$

c) $(X, \sim)$ heißt transitiv, falls $\forall x, y, z \in X : x \sim y \ \wedge \ y \sim z \ \Rightarrow \ x \sim z$.

d) $(X, \sim)$ heißt antisymmetrisch, falls $\forall x, y \in X : x \sim y \ \wedge \ y \sim x \ \Rightarrow \ x = y$.

e) $(X, \sim)$ heißt Äquivalenzrelation, falls sie reflexiv, symmetrisch und transitiv ist.

f) $(X, \sim)$ heißt Ordnungsrelation, falls sie reflexiv, transitiv und antisymmetrisch ist.

Beispiel 1.12. 1) Für eine beliebige Menge X ist $(X, =)$ eine Äquivalenzrelation.

2) $(\mathbb{R}, \leq)$ ist eine Ordungsrelation.

3) $(\mathcal{P}(A), \subseteq)$ ist ebenfalls eine Ordnungsrelation.

4) Sei $m \in \mathbb{N} \setminus \{0\}$. Eine Relation auf $\mathbb{Z}$ kriegen wir so: Wir definieren $a \sim b :\Leftrightarrow m \mid (a - b)$ Man schreibt dann auch $a \equiv b \pmod{m}$ und sagt "a ist kongruent zu b modulo m. Dann ist $(\mathbb{Z}, \sim)$ eine Äquivalenzrelation.

Beweis. Gilt wegen der Teilbarkeitseigenschaft $m \mid p \wedge m \mid q \Rightarrow m \mid \alpha p + \beta q$ für $\alpha, \beta \in \mathbb{Z}$. □

Definition 1.11 (Äquivalenzklasse)**.** Sei $(X, \sim)$ eine Äquivalenzrelation. Für $a \in X$ definieren wir $[a] \coloneqq \big\{x \in X \mid x \sim a\big\}$. Es heißt dann $[a]$ Äquivalenzklasse zu a.

Bemerkung. Es gilt immer $[a] = [b]$ falls $a \sim b$ oder $[a] \cap [b] = \emptyset$, falls $a \nsim b$.

Beispiel 1.13. Man kriegt in unserem Beispiel der Kongruenz modulo m für $m \in \mathbb{N} \setminus \{0\}$ die m verschiedenen Äquivalenzklassen aus $\mathbb{Z}_m \coloneqq \big\{[0], [1], ..., [m-1]\big\}$. Man schreibt dann auch $[a] = a + m\mathbb{Z}$ und bezeichnet $[a]$ dann auch als Kongruenzklasse von a modulo m und $\mathbb{Z}_m$ als die Menge der Restklassen modulo m.

Bemerkung. Man kann sich für eine Äquivalenzrelation $(X, \sim)$ die Menge X als eine Menge von Gegenständen die Relation $x \sim y$ als "x hat die gleiche Farbe wie y" vorstellen.

Beispiel 1.14. Es sei $X = \{1, 2, 3\}, Y = \{1, 2\}$ und $R \subseteq X \times Y$ gegeben durch $R = \big\{(1, 2), (3, 1), (3, 2)\big\}$. Man kann sich diese Relation leicht an einem Bild veranschaulichen.

Umkehrrelation Ist (X, Y, R) eine Relation, so bezeichnet man (Y, X, R^{-1}) mit $R^{-1} = \big\{(y, x) \in Y \times X \mid (x, y) \in R\big\}$ als die Umkehrrelation von R. Es gilt dann $(R^{-1})^{-1} = R$.

1.6 Funktionen, Abbildungen

Definition 1.12. Eine Relation (X, Y, R) mit der Eigenschaft $\forall x \in X,\ \exists! y \in Y \colon (x, y) \in R$ heißt Funktion (Abbildung) von X nach Y. Man bezeichnet dann y als "Funktionswert von $f \colon X \to Y$ an der Stelle x": $y = f(x)$ oder $x \overset{f}{\mapsto} y$. Man nennt dann X den Definitionsbereich und Y den Wertebereich von f. R ist dann *der Graph* von f, $\operatorname{graph}(f) \coloneqq \big\{(x, f(x)) \mid x \in X\big\} \subseteq X \times Y$.

Beispiel 1.15. **a)** $f \colon \{1, 2, 3\} \to \{1, 2\}$ mit $f(1) = 2, f(2) = 1, f(3) = 1$ ist eine Abbildung.

b) $f \colon [0, 1] \to \mathbb{R}, f(x) = x^2$ ist eine Abbildung.

c) Sei $(X, \sim)$ eine Äquivalenzrelation. Dann ist $f \colon X \to \mathcal{P}(X), a \mapsto [a]$ eine Funktion.

d) Eine Abbildung $f \colon \{1, ..., n\} \to X_1 \cup ... \cup X_n, k \mapsto x_k \in X_k$ ist ein n-Tupel

$$(x_1, \ldots, x_n) \in X_1 \times \cdots \times X_n.$$

Allgemeiner sei I eine (Index)Menge und M eine Menge. Dann heißt $f \colon I \to M$ ein I-Tupel. Dieses bezeichnet man dann auch als Familie von Elementen aus M. Man bezeichnet mit M^I die Menge $\{f \colon I \to M\}$. Dies ist eine Verallgemeinerung von $A^n \coloneqq \underbrace{A \times \cdots \times A}_{n-\text{mal}} = \big\{f \colon \{1, .., n\} \to A\big\}$.

e) $\exists! f\colon \emptyset \to A$ für $A \neq \emptyset$. Dies ist die leere Funktion. Aber: $\nexists f\colon A \to \emptyset$, falls $A \neq \emptyset$. Insbesondere gibt es genau ein 0-Tupel (oder $\emptyset$-Tupel) von Elementen aus A.

f) Für eine Funktion $f\colon \mathbb{N} \to X$ mit $f(n) = x_n$ bezeichnet man $(x_n)_{n\in\mathbb{N}}$ als eine Folge in X.

Mit der Indexnotation werden die übliche Operationen (Vereinigung, Durchschnitt, kartesisches Produkt) für beliebig viele Mengen definiert:

Definition 1.13. Sei $\mathcal{M}$ eine Menge von Mengen, I eine nicht leere (Index)menge und $f : I \to \mathcal{M}$ eine I-Familie von Mengen. Wir bezeichnen

$$\forall i \in I \ : \ A_i := f(i) \in \mathcal{M}$$

und wir schreiben $(A_i)_{i\in I}$ die entsprechende I-Familie (oder I-Tupel von Mengen).

a) Die Vereinigung der Mengen $(A_i)_{i\in I}$ ist definiert durch

$$\bigcup_{i\in I} A_i := \{x \mid \exists i \in I \ : \ x \in A_i\}.$$

b) Der Durchschnitt der Mengen $(A_i)_{i\in I}$ ist definiert durch

$$\bigcap_{i\in I} A_i := \{x \mid \forall i \in I \ : \ x \in A_i\}.$$

c) Das kartesische Produkt der Mengen $(A_i)_{i\in I}$ ist definiert durch

$$\prod_{i\in I} A_i := \left\{g : I \to \cup_{i\in I} A_i \ \mid \ \forall i \in I \ : \ g(i) \in A_i\right\}.$$

Ein Element aus diesem Produkt ist also ein I-Tupel:

$$(x_i)_{i\in I} \in \prod_{i\in I} A_i \ \Leftrightarrow \forall i \in I \ : \ x_i \in A_i.$$

Für $I := \{1, 2\}$ die obige Definitionen a) und b) stimmen mit den bekannten Operationen $A_1 \cup A_2$, bzw. $A_1 \cap A_2$ überein (für (a) und (b) (Übung); für das kartesische Produkt wird in dieser Vorlesung ohne Beweis akzeptiert, dass (c) eine äquivalente[1] Definition zu 1.4 ist.

Für $I := \{1, 2, \ldots, n\}$ schreibt man auch

$$\bigcup_{i\in\{1,\ldots,n\}} A_i = \bigcup_{i=1}^{n} A_i = A_1 \cup \cdots \cup A_n,$$

bzw. analog für den Durchschnitt, bzw. das kartesische Produkt.

[1] Wichtig dabei ist nicht die konkrete Realisierung eines Paaren, als $\{\{a\}, \{a, b\}\}$ oder als Funktion f definiert auf $\{1, 2\}$, mit Werten $f(1) = a$ und $f(b) = b$, sondern die Grundeigenschaften der Paaren und des kartesischen Produktes

Bemerkung. 1. Die Notationen $A_1 \cup \cdots \cup A_n$ oder $A_1 \cap \cdots \cap A_n$ sind durch die Assoziativität der Vereinigung, bzw. Durchschnitt berechtigt. Für das kartesische Produkt $A_1 \times \cdots \times A_n$ wird ebenfalls die Assoziativität ohne Beweis akzeptiert.

2. Für endliche oder *abzählbare* Mengen I (z.B. $I := \{1, \ldots, n\}$, bzw. $I := \mathbb{N}$) (siehe den Abschnitt über Mächtigkeit unten) ist die Indexnotation

$$\bigcup_{i \in I} A_i = \bigcup_{i=1}^{n} A_i, \text{ bzw. } \bigcup_{i \in \mathbb{N}} A_i = \bigcup i = 0^{\infty} A_i$$

eine kompakte (und eindeutige) schreibweise für die intuitive Notation $A_1 \cup \cdots \cup A_n$, bzw. $A_0 \cup \cdots \cup A_n \cup \ldots$. Für I unendlich, insbesondere *überabzählbar* (siehe unten) gibt es keine andere Notation für die Vereinigung, Durchschnitt und kartesisches Produkt einer I-Familie von Mengen.

Per Induktion kann man zeigen, dass das kartesische Produkt $A_1 \times \cdots \times A_n$ nicht leer ist, wenn alle Mengen $(A_i)_{i \in I}$ nicht leer sind. Für unendliche Indexmengen kann eine solche Aussage weder bewiesen, noch widerlegt werden; sie wird meistens als *Axiom* angenommen:

Das Auswahlsaxiom Sei $(A_i)_{i \in I}$ eine Familie von nicht leeren Mengen. Dann ist das kartesische Produkt $\prod_{i \in I} A_i$ nicht leer.

Definition 1.14 (Bild, Urbild)**.** Sei $f : X \to Y$ eine Funktion.

a) Sei $A \subset X$. Die Teilmenge

$$f_P(A) := \{f(x) \ : \ x \in X\} \subset Y,$$

oft einfach $f(A)$ geschrieben, ist *das Bild von A durch f*

b) Sei $B \subset Y$. Die Teilmenge

$$f_P^{-1}(B) := \{x \in X \ : \ f(x) \in B\} \subset X,$$

oft $f^{-1}(B)$ geschrieben, heißt *das Urbild durch f der Menge B.*

f_P, bzw. f_P^{-1} sind Abbildungen von $\mathcal{P}(X)$ nach $\mathcal{P}(Y)$ und sind von der Abbildung $f : X \to Y$, bzw. von seiner *Umkehrabbildung* (siehe unten) stets zu unterscheiden!

Definition 1.15 (Injektivität, Surjektivität, Bijektivität)**.** Es sei $f \colon X \to Y$ eine Abbildung.

a) Wir sagen f ist *injektiv*, falls gilt:

$$\forall\, x, y \in X \colon f(x) = f(y) \Rightarrow x = y$$

b) Wir sagen f ist *surjektiv*, falls gilt:

$$\forall\, y \in Y \, \exists\, x \in X \colon f(x) = y$$

c) Wir sagen f ist bijektiv, wenn f sowohl injektiv als auch surjektiv ist.

1.6.1 Verkettung von Funktionen

Seien $f\colon X \to Y$ und $g\colon Y \to Z$ Funktionen. Die Verkettung (Komposition) von g mit f ist die Funktion $g \circ f : X \to Z$, definiert als

$$\forall x \in X \ : \ (g \circ f)(x) \coloneqq g(f(x)).$$

Bemerkung. 1. Erst wird f berechnet in x, dann wird g berechnet in $f(x)$. Man schreibt aber erst g, dann f in $g \circ f$.

2. Für die Komposition von Abbildungen ist es wichtig, dass der Definitionsbereich von g mit dem Wertebereich von f übereinstimmt! Manchmal lässt man jedoch zu, dass $f\colon X \to Y \wedge g\colon Y' \to Z$ mit $Y \subseteq Y'$, um die Verkettung $g \circ f$ zu definieren. Korrekt wäre, $g\,|_Y \circ f$ zu schreiben, wobei $g\,|_Y$ die Einschränkung von g auf Y ist.

Ein Spezialfall der Verkettung von Funktionen ist $X = Z, f\colon X \to Y, g\colon Y \to X$. Es können dann also $f \circ g$ und $g \circ f$ definiert werden. Diese sind aber im Allgemeinen nicht die gleichen Funktionen, auch nicht wenn $X = Y$.

Beispiel 1.16. $f\colon \mathbb{Z} \to \mathbb{Z}, f(x) \coloneqq x + 1$ und $g\colon \mathbb{Z} \to \mathbb{Z}, g(x) \coloneqq 2x$. Dann gilt:

$$(f \circ g)(x) = f(g(x)) = f(2x) = 2x + 1$$
$$(g \circ f)(x) = g(f(x)) = g(x + 1) = 2x + 2$$

Offenbar sind diese zwei Funktionen nicht gleich.

1.6.2 Inverse (Umkehr-)Abbildung

Definition 1.16. Eine Funktion $f\colon X \to Y$ heißt *invertierbar* genau dann, wenn es existiert eine Abbildung $g\colon Y \to X \Leftrightarrow f \circ g = \mathrm{id}_Y \ \wedge \ g \circ f = \mathrm{id}_X$. g heißt dann eine Umkehrfunktion zu f, und man sagt, f besitzt eine Umkehrfunktion.

Beispiel 1.17. 1) Eine Umkehrabbildung zur Identitätsabbildung $\mathrm{id}_X : X \to X$ ist id_X selbst;

2) $f\colon X \to Y, f(x) = y$, dann falls g eine Umkehrfunktion zu f ist, so gilt $g(y) = x$. $g(y)$ ist also die Lösung der Gleichung $f(x) = y$. Beispielsweise besitzt $f\colon \mathbb{Z} \to \mathbb{Z}, f(x) = 2x$ keine Umkehrabbildung, weil es sonst x mit $g(1) = x$ geben müsste, also $f(x) = 1$. Ein solches x existiert aber in $\mathbb{Z}$ nicht.
Aber: $f\colon \mathbb{Z} \to 2\mathbb{Z}, f(x) \coloneqq 2x$ besitzt die Abbildung $g\colon 2\mathbb{Z} \to \mathbb{Z}, g(y) \coloneqq \frac{1}{2}y \in \mathbb{Z}$. Wir sehen also, dass Definitions- und Wertebereich von Funktionen entscheidend sein können.

Satz 1.14. *$f\colon X \to Y$ besitzt eine Umkehrabbildung $\Leftrightarrow$ f ist bijektiv. Dann ist die Umkehrabbildung von f eindeutig bestimmt.*

Wir beweisen dazu zunächst das folgende

Lemma 1.15. *Seien $f\colon X \to Y, g\colon Y \to Z$ Funktionen.*

1) *Falls $g \circ f$ injektiv ist, so ist f injektiv.*

2) *falls $g \circ f$ surjektiv ist, so ist g surjektiv.*

Beweis. **1)** Seien $x_1, x_2 \in X, f(x_1) = f(x_2)$. Dann ist $g(f(x_1)) = g(f(x_2)) \Leftrightarrow (g \circ f)(x_1) = (g \circ f)(x_2)$. Also $x_1 = x_2$, weil $g \circ f$ injektiv ist.

2) $\forall\, z \in Z\, \exists\, x \in X\colon (g \circ f)(x) = z$. Also $y \coloneqq f(x) \in Y$ erfüllt $g(y) = z$. Also ist g surjektiv.

□

Übung. Man mache sich klar, dass auch die folgenden Aussagen gelten:

3) f injektiv $\wedge$ g injektiv $\Rightarrow$ $g \circ f$ injektiv

4) f surjektiv $\wedge$ g surjektiv $\Rightarrow$ $g \circ f$ surjektiv

Korollar 1.16. ***1)*** *Falls $g \circ f$ bijektiv ist, so ist g surjektiv und f injektiv.*

2) *Falls f bijektiv, g bijektiv, so auch $g \circ f$ bijektiv.*

Beweis des Satzes. Sei $g\colon Y \to X$ Umkehrabbildung zu f. Dann sind $g \circ f$ und $f \circ g$ bijektiv, also f und g sind bijektiv nach dem obigen Korollar.
Sei nun umgekehrt f bijektiv. Dann $\forall\, y \in Y\, \exists\, !x \in X\colon f(x) = y$. Wir definieren $g(y) \coloneqq x$. Als Relation gilt dann

$$G_g \coloneqq \big\{(y, x) \in Y \times X \mid (x, y) \in G_f\big\}$$

ist der Graph einer Funktion, da gilt:

$$\forall\, y \in Y\, \exists\, !x \in X\colon (y, x) \in G_{f^{-1}}$$

Dies ist nämlich gerade die Bedingung, dass f bijektiv sein sollte. Die obige Konstruktion von g zeigt insbesondere, dass die Umkehrfunktion eindeutig bestimmt ist. □

Die Umkehrabbildung zu f wird meistens f^{-1} notiert.

Bemerkung. der Graph von f^{-1} ist, als Relation zwischem Y und X, genau die Umkehrrelation zu $G_f \subset X \times Y$. Dabei ist zu beachten, dass die Umkehrrelation zu G_f *immer existiert*, aber die Umkehrfunktion zu f nur für bijektive Funktionen gibt. Die Umkehrrelation zu einer nicht-bijektiven Funktion ist also *kein Graph.*

1.6.3 Endliche Mengen, Mächtigkeit

Definition 1.17. 1) Seien A, B Mengen. A und B heißen gleichmächtig, falls $\exists\, f\colon A \to B$ und f ist bijektiv. Gleichmächtigkeit definiert eine Äquivalenzrelation auf jeder Menge von Mengen.

2) Sei $\mathcal{M}$ eine Menge von Mengen. Die Äquivalenzklassen für die Gleichmächtigkeitsrelation auf $\mathcal{M}$ heißen Kardinalzahlen. $|A|$ heißt die Kardinalzahl zu A.

3) Eine Menge A ist endlich, falls $\exists\, n \in \mathbb{N}$ so dass $|A| = \big|\{1, .., n\}\big|$. Man schreibt dann $|A| = n$ (oder $\#A = n$) und sagt dann "A hat n Elemente".

4) Eine Menge A heißt abzählbar, falls $|A| = |\mathbb{N}|$.

Bemerkung. $\emptyset$ hat 0 Elemente, $\big|\emptyset\big| = 0$.

Beispiel 1.18. $\mathbb{N} \times \mathbb{N}, \mathbb{Q}, \mathbb{Z}^k$ sind abzählbar.

1.6.4 "Weniger mächtig"-Relation

Definition 1.18. Seien A, B Mengen. A heißt weniger mächtig als $B \Leftrightarrow \exists\, f\colon A \to B$ so dass f injektiv ist. Wir schreiben dann $|A| \leq |B|$.

Beispiel 1.19. Jede Teilmenge einer Menge ist weniger mächtig als die Menge selbst.

Bemerkung. Für endliche Mengen $|A| \leq |B|$ ist äquivalent zu der Ungleichung zwischen den natürlichen Zahlen $|A|, |B|$ (siehe den folgenden Beweis).

Satz 1.17. *Seien A, B zwei endliche Mengen. Die folgende Aussagen sind dann gültig:*

1) $\exists\, f\colon A \to B$ *injektiv* $\quad \vee \quad \exists\, g\colon B \to A$ *injektiv.*

2) $\exists\, f\colon A \to B$ *injektiv* $\Leftrightarrow \exists\, g\colon B \to A$ *surjektiv für* $A, B \neq \emptyset$.

3) $\exists\, f\colon A \to B$ *injektiv und* $\exists\, g\colon B \to A$ *injektiv, so* $|A| = |B|$

4) *Falls* $|A| = |B|$ *und* $f\colon A \to B$*, so sind äquivalent:*

- ***a)*** *f ist injektiv.*
- ***b)*** *f ist surjektiv.*
- ***c)*** *f ist bijektiv.*

Beweis. Sei $n_1 := |A|$ und $n_2 := |B|$. Es gibt Abbildungen $f_1\colon \{1, .., n_1\} \to A$ bijektiv und $f_2\colon \{1, .., n_2\} \to B$ bijektiv.

1) Sei $n_1 \leq n_2$. Dann ist $i\colon \{1, .., n_1\} \to \{1, .., n_2\}, k \mapsto k$ injektiv. Also ist $f_2 \circ i \circ f_2^{-1}\colon A \to B$ injektiv. Analog falls $n_2 \leq n_1$, so gibt es $g\colon B \to A$ injektiv.

2) Sei $f\colon A \to B$ injektiv und $A \neq \emptyset$. Dann $\exists\, x_0 \colon x_0 \in A$. Sei $g\colon B \to A$ definiert durch:

$$g(y) \coloneqq \begin{cases} x, & \text{falls } y = f(x) \\ x_0, & \text{falls } y \notin f(A) \end{cases}$$

g ist surjektiv, weil $g \circ f = \mathrm{id}_A$ ist. Sei nun umgekert $g\colon B \to A$ surjektiv. Wir definieren $f\colon A \to B$ durch Auswahl je eines Elementes $f(x)$ in $g^{-1}(\{x\}) \neq \emptyset$. Dann ist f injektiv, da $g \circ f = \mathrm{id}_A$.

Bemerkung. Wollten wir dies für unendliche Mengen tun, so bräuchten wir das Auswahlaxiom.

3) Folgt unmittelbar aus dem Beweis der ersten Aussage: Es folgt $n_1 = n_2$.

4) Sei $n \coloneqq |A| = |B|$. Wir beweisen per Induktion, dass a) $\Leftrightarrow$ b). Dies genügt.

I.A. $n = 0$: Trivial.

I.S. $n > 0$: Sei $x_0 \in A$ und $f\colon A \to B$ injektiv. Dann ist $\hat{f}\colon A \setminus \{x_0\} \to B \setminus \{f(x_0)\}$ wohldefiniert und injektiv. Nach Induktionsvoraussetzung ist $\hat{f}$ bijektiv und damit auch f.
Sei nun umgekehrt $f\colon A \to B$ surjektiv. Dann gibt es nach 2. ein $g\colon B \to A$ injektiv und aus der Induktion von oben folgt g ist bijektiv.

□

Bemerkung. 1), 2), 3) aus dem Satz sind auch für unendliche Mengen richtig. Der Beweis ist in diesem Fall jedoch viel schwieriger.

Korollar 1.18. *Seien A, B Mengen. Dann:*

1) $|A| \leq |B| \ \vee\ |B| \leq |A|$

2) $|A| \leq |B| \ \wedge\ |B| \leq |A| \ \Leftrightarrow\ |A| = |B|$.

Eine Ordnungsrelation $(X, \preceq)$ heißt total*, falls*

$$\forall x, y \in X \ \Rightarrow (x \preceq y \ \vee\ y \preceq x).$$

Die "weniger mächtig" Relation ist also eine totale Ordnungsrelation auf der Menge der Kardinalzahlen. Man schreibt $|A| < |B|$ für $|A| \leq |B| \ \wedge\ |A| \neq |B|$.

Satz 1.19 (Cantor)**.** *Sei A eine Menge. Dann gilt $|A| < \big|\mathcal{P}(A)\big|$.*

Beweis. Beweis durch Widerspruch. Sei also $f\colon A \to \mathcal{P}(A)$ surjektiv. Sei $B \coloneqq \big\{x \in A \mid x \notin f(x)\big\}$. Da f surjektiv ist $\exists\, x_0 \in A\colon f(x_0) = B$. Es gilt $x_0 \in B \vee x_0 \notin B$. Aber $x_0 \in B \Rightarrow x_0 \notin f(x_0) = B$ ist ein Widerspruch und $x_0 \notin B \Rightarrow x_0 \notin f(x_0)$, also $x_0 \in B$ ist ebenfalls ein Widerspruch. Es kann also keine solche surjektive Abbildung f geben. □

Definition 1.19. Eine Menge A so dass $|A| > |\mathbb{N}|$ heißt überabzählbar.

Korollar 1.20. *$\mathcal{P}(\mathbb{N})$ ist überabzählbar.*

In der Analysis zeigt man $|\mathbb{R}| = \left|\mathcal{P}(\mathbb{N})\right|$, also ist $\mathbb{R}$ überabzählbar.

Dagegen ist $\mathbb{Q}$ abzählbar:

Dies folgt aus den Aufgaben 4 in den Übungsblatt 3, bzw. Tutoriumsblatt 3: zuerst wird gezeigt, $\mathbb{Z} = \mathbb{N} \cup (-\mathbb{N})$ ist abzählbar, dann $M := \mathbb{Z} \times (\mathbb{N} \setminus \{0\})$ ist abzählbar, und man definiert

$$f : M \to \mathbb{Q},\ f(p,q) := \frac{p}{q}.$$

f ist dann surjektiv und also $|\mathbb{N}| = |M| \geq |\mathbb{Q}|$. Da die Kardinalzahl von $\mathbb{Q}$ unendlich ist, folgt $|\mathbb{Q}| = |\mathbb{N}|$.

In der Analysis ist die Abzählbarkeit ein extrem wichtiges Begriff, denn die Summe von unendlich, abzählbar vielen Termen kann unter Umständen wolhdefiniert werden.

Oft können durch solchen Summen verschiedene Zahlen oder andere Objekte approximiert werden. Der wohlbekannteste Beispiel ist die Dezimaldarstellung der reellen Zahlen:

$$a = a_n a_{n-1} \dots a_0, a_{-1} a_{-2} \cdots \in \mathbb{R},$$

mit $a_k \in \{0, \dots, 9\},\ \forall k \in \mathbb{Z},\ k \leq n,\ a_n \neq 0$, die eigentliche Bedeutung dieser Darstellung ist

$$a = \sum_{k \in \mathbb{Z}, k \leq n} a_k \cdot 10^k.$$

Fazit: viele der Quantitäten, Mengen, Objekte, die wir berechnen müssen, können durch Algorithmen mit abzählbar vielen Schritten bestimmt werden. Falls der Schritt $n \to n+1$ automatisch funktioniert (wie in einem Induktionsbeweis), kann mann gewisse Eigenschaften über das Endergebnis theoretisch daraus schliessen. In der Praxis sucht man aber dies Endergebnis selbst und beschränkt man sich auf eine endliche (grosse) Anzahl von solchen Schritten, wodurch aber ein Fehlerterm entsteht (möglicherweise klein und/oder unwichtig für die Aufgabe).

In allen diesen methoden ist der Begriff der Abzählbarkeit wesentlich.

2 Algebraische Strukturen

Definition 2.1 (Verknüpfung)**.** Sei M eine Menge. Eine Verknüpfung auf M ist eine Abbildung $\circ\colon M \times M \to M, (a,b) \mapsto \circ(a,b) =: a \circ b$. Schreibweise $(M, \circ)$, "M mit Verknüpfung $\circ$."

Beispiel 2.1. $(\mathbb{N}, +), (\mathbb{Z}, -), (\mathcal{P}(A), \cup), (\mathcal{P}(A), \cap)$ sind Mengen mit Verknüpfungen, jedoch ist $(\mathbb{N}, -)$ keine Menge mit Verknüpfung.

Definition 2.2 (Gruppe)**.** $(G, \circ)$ heißt Gruppe, falls gilt:

(G1) $\forall\, a, b, c \in G\colon (a \circ b) \circ c = a \circ (b \circ c)$. $\circ$ soll also *assoziativ* sein.

(G2) $\exists\, e \in G\ \forall\, a \in G\colon a \circ e = e \circ a = a$. Es gibt in Gruppen also ein *neutrales Element* bezüglich $\circ$.

G3) $\forall\, a \in G\ \exists\, b \in G\colon a \circ b = b \circ a = e$. Es gibt *inverse Elemente* bezüglich $\circ$.

Die Bedingungen (G1),(G2),(G3) (und eventuell die Bedingung (Ab) weiter unten) heißen *Gruppenaxiome*. Eine Menge mit Verknüpfung ist also eine Gruppe, falls die Gruppenaxiome gelten (oder: die Verknüpfung erfüllt die Gruppenaxiome).

Satz 2.1. ***a)*** *Das neutrale Element e ist eindeutig bestimmt.*

b) *Inverse Elemente sind eindeutig bestimmt.*

Beweis. **a)** Seien e, e' zwei neutrale Elemente. Dann gilt:
$e \circ e' = e$, weil e' ist ein neutrales Element (siehe (G2)).
$e \circ e' = e'$, weil e ist ein neutrales Element (siehe (G2)).Also $e = e'$.

b) Seien b, b' inverse Elemente von a. Dann gilt:

$$\left.\begin{array}{lll} a \circ b & = e & \overset{(G2)}{\Rightarrow} b' \circ (a \circ b) = b' \\ b' \circ a & = e & \overset{(G2)}{\Rightarrow} (b' \circ a) \circ b = b \end{array}\right\} \overset{(G1)}{\Rightarrow} b = b'$$

□

Beispiel 2.2. • $(\mathbb{Z}, +)$ ist eine Gruppe. Das neutrale Element ist 0 und für jedes x ist $-x$ das inverse Element.

• $(\mathbb{N}, +)$ ist keine Gruppe!

• $(\mathbb{Q} \setminus \{0\}, \cdot)$ ist eine Gruppe mit neutralem Element 1 und für jedes x ist $x^{-1} = \frac{1}{x}$ das inverse Element.

• $(\mathbb{Z}, \cdot)$ ist keine Gruppe!

- $(\mathbb{R}^n, +)$ ist eine Gruppe mit $(x_1, .., x_n) + (y_1, ..., y_n) \coloneqq (x_1 + y_1, ..., x_n + y_n)$. Das neutrale Element ist dann $0 = (0, ..., 0)$ und $-(x_1, .., x_n) \coloneqq (-x_1, ..., -x_n)$ ist das inverse Element zu $(x_1, ..., x_n)$.

Notation 2.3. Wir schreiben für Multiplikative Gruppen $(G, \cdot), (G, *)$ und bezeichnen das neutrale Element mit e oder 1 und das Inverse zu x mit x^{-1}.

Definition 2.4. Eine Gruppe $(G, \circ)$ heißt abelsch oder kommutativ, wenn gilt:

$$\forall\, a, b \in G\colon a \circ b = b \circ a. \tag{Ab}$$

Notation 2.5. Manchmal wird die Verknüpfung einer abelschen Gruppe additiv geschrieben. Das neutrale Element bezeichnet man dann mit 0 und das Inverse zu x mit $-x$.

Beispiel 2.3. Alle obigen Beispiele sind abelsch. Wir möchten nun also ein Beispiel einer nicht abelschen Gruppe angeben. Sei hierzu A eine Menge. Definiere $S(A) \coloneqq \{ f\colon A \to A \mid f \text{ bijektiv } \}$ mit Verknüpfung $\circ \coloneqq$ Verkettung von Abbildungen. Dann ist $(S(A), \circ)$ eine Gruppe. Sie ist im Allgemeinen nicht abelsch. Genauer gesagt ist sie nur für $|A| \leq 2$ abelsch (Übung). Man nennt $(S(A), \circ)$ die symmetrische Gruppe auf A. Für $A = \{1, .., n\}$ setzen wir $S(n) \coloneqq S(A)$. Sie ist die Gruppe der Permutationen auf n Elementen.

Notation 2.6. Für $f\colon \{1, .., n\} \to \{1, .., n\}$ schreiben wir auch: $\begin{pmatrix} 1 & 2 & ... & n \\ f(1) & f(2) & ... & f(n) \end{pmatrix}$

Bemerkung. Es gilt für $S_n \coloneqq S(n)\colon |S_n| = n!$.

Beweis. Durch Induktion über n:

I.A. $n = 0$: $S_0 = \{ f\colon \emptyset \to \emptyset \mid f \text{ bijektiv } \}$. $|S_0| = 1 = 0!$

I.S. $n \geq 0$: Wir setzen $G\colon S_n \times \{1, ..., n+1\} \to S_{n+1}$ mit $G(\phi, m) \coloneqq \psi$ die folgende Abbildung:

$$\psi(p) \coloneqq \begin{cases} \phi(p), & \text{falls } p \leq n \ \wedge\ \phi(p) < m \\ \phi(p) + 1, & \text{falls } p \leq n \ \wedge\ \phi(p) \geq m \\ m, & \text{falls } p = n + 1 \end{cases}$$

Wir zeigen zuerst, ψ ist injektiv, also ist G wohldefiniert: Seien $p_1, p_2 \in \{1, \ldots, n+1\}$, $\psi(p_1) = \psi(p_2)$. Es gilt erst $\psi(p_1) = m\, (= \psi(p_2))$ ist äquivalent zu $p_1 = p_2 = n + 1$. Weiterhin gilt:

$$\begin{array}{llllll} \forall p \in \{1, \ldots, n\}, & \psi(p) < m & \Leftrightarrow & \phi(p) < m & \Leftrightarrow & \psi(p) = \phi(p), \text{ und} \\ \forall p \in \{1, \ldots, n\}, & \psi(p) > m & \Leftrightarrow & \phi(p) \geq m & \Leftrightarrow & \psi(p) = \phi(p) + 1. \end{array} \tag{1}$$

Es folgt also, für $p_1, p_2 \leq n$, dass $\psi(p_1) = \psi(p_2) \Leftrightarrow \phi(p_1) = \phi(p_2)$, und dass ist äquivalent zu $p_1 = p_2$, weil ϕ injektiv ist.

Wir zeigen nun, dass G bijektiv ist. Sei $\psi \in S_{n+1}$. Sei $m \coloneqq \psi(n+1)$ und sei $\phi\colon \{1,..,n\} \to \{1,..,n\}, \phi(p) \coloneqq \begin{cases} \psi(p), & \psi(p) < m \\ \psi(p) - 1, & \psi(p) > m \end{cases}$. Wie oben folgt aus (1), dass $\phi : \{1,\ldots,n\} \to \{1,\ldots,n\}$ injektiv (also bijektiv) ist. Es gilt $G(\phi, m) = \psi$. Also ist G surjektiv. Wir müssen noch zeigen, dass G injektiv ist:

Seien $\phi_1, \phi_2 \in S_n$ und $m_1, m_2 \in \{1,..,n+1\}$ so dass $G(\phi_1, m_1) = G(\phi_2, m_2) = \psi$. Dann gilt $\psi(n+1) = m_1 = m_2 = m$. Wie in (1) gilt es

$$\phi_1(p) < m \Leftrightarrow \psi(p) < m \Leftrightarrow \phi_2(p) < m \Leftrightarrow \phi_1(p) = \psi(p) = \phi_2(p).$$

Ebenfalls,

$$\phi_1(p) \geq m \Leftrightarrow \psi(p) > m \Leftrightarrow \phi_2(p) \geq m \Leftrightarrow \phi_1(p) = \psi(p) - 1 = \phi_2(p).$$

Es folgt also $\forall p \in \{1,\ldots,n\},\ \phi_1(p) = \phi_2(p)$, also $\phi_1 = \phi_2$.

Die Abbildung $G : S_n \times \{1,\ldots,n+1\} \to S_{n+1}$ ist also bijektiv und $|S_n| \cdot (n+1) = |S_{n+1}|$. Da $|S_n| = n!$ aus der Induktionsvoraussetzung, folgt $|S_{n+1}| = (n+1)!$.

□

Satz 2.2. ***(a)*** *Sei $m \in \mathbb{N}$. Dann ist $(\mathbb{Z}_m, b)$ mit der Verknüpfung $[a]_m + [b]_m \coloneqq [a+b]_m$ eine abelsche Gruppe mit m Elementen.*

(b) *Sei $p \in \mathbb{N}$ eine Primzahl. Dann ist $(\mathbb{Z}_p \setminus \{[0]_p\}, \cdot)$ (die Verknüpfung "$\cdot$" wird durch: $[a]_p \cdot [b]_p \coloneqq [a \cdot b]_p$ definiert) eine abelsche Gruppe mit $p-1$ Elementen.*

Beweis. **(a)** $+$ ist wohldefiniert, denn $\forall k \in \mathbb{Z}\colon [a]_m = [a+km]_m$ und $\forall l \in \mathbb{Z}\colon [b]_m = [b+lm]_m$ aber $a + km + b + lm = a + b + (k+l)m \equiv a + b \pmod{m}$. Also $[a+b]_m = [(a+km) + (b+lm)]_m$. Die Gruppenaxiome folgen direkt aus den Eigenschaften von $(\mathbb{Z}, +)$.

(b) Analog zu (a), hängt die Restklasse von $a \cdot b \pmod{p}$ nur von den Restklassen von a und $b \pmod{p}$. Weiter gilt $a \not\equiv 0 \pmod{p} \wedge b \not\equiv 0 \pmod{p} \Rightarrow p \nmid ab$. Also ist die Multiplikation $\cdot$ eine Verknüpfung auf $\mathbb{Z}_p \setminus \{[0]_p\}$. Die Assoziativität und Kommuntativität folgen direkt aus den entsprechenden Eigenschaften der Multiplikation aus $\mathbb{Z}$. Das neutrale Element ist $[1]_p$. Zum Inversen: Sei $a \in \mathbb{Z}$ und $p \nmid a$. Dann $\mathrm{ggT}(p, a) = 1 \overset{\text{Bezout}}{\Rightarrow} \exists l, m \in \mathbb{Z}\colon pl + am = 1 \Rightarrow [m]_p \cdot [a]_p = [1]_p$.
□

Beispiel 2.4. $\mathbb{Z}_2$ Verknüpfungstafel. Es ist $\mathbb{Z}_2 = \big\{[0],[1]\big\}$ wobei [0] die Klasse der geraden Zahlen und [1] die Klasse der ungeraden Zahlen ist.

Es gilt in $\mathbb{Z}_2$:

$+$	$[0]$	$[1]$
$[0]$	$[0]$	$[1]$
$[1]$	$[1]$	$[0]$

Allgemeiner gilt $\mathbb{Z}_k = \big\{[0],...,[k-1]\big\}$ und $[p]+[q] = [p+q]$. Man bezeichnet $[p]$ auch mit $\overline{p}$.

Damit gilt dann

$+$	$\overline{0}$	$\overline{1}$	$\overline{2}$
$\overline{0}$	$\overline{0}$	$\overline{1}$	$\overline{2}$
$\overline{1}$	$\overline{1}$	$\overline{2}$	$\overline{0}$
$\overline{2}$	$\overline{2}$	$\overline{0}$	$\overline{1}$

in $\mathbb{Z}_3$ und für $\cdot$ gilt in $\mathbb{Z}_3$:

$\cdot$	$\overline{1}$	$\overline{2}$
$\overline{1}$	$\overline{1}$	$\overline{2}$
$\overline{2}$	$\overline{2}$	$\overline{1}$

und in $\mathbb{Z}_5$:

$\cdot$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{4}$
$\overline{1}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{4}$
$\overline{2}$	$\overline{2}$	$\overline{4}$	$\overline{1}$	$\overline{3}$
$\overline{3}$	$\overline{3}$	$\overline{1}$	$\overline{4}$	$\overline{2}$
$\overline{4}$	$\overline{4}$	$\overline{3}$	$\overline{2}$	$\overline{1}$

$\Leftrightarrow$

$\cdot$	$\overline{1}$	$\overline{2}$	$\overline{4}$	$\overline{3}$
$\overline{1}$	$\overline{1}$	$\overline{2}$	$\overline{4}$	$\overline{3}$
$\overline{2}$	$\overline{2}$	$\overline{4}$	$\overline{3}$	$\overline{1}$
$\overline{4}$	$\overline{4}$	$\overline{3}$	$\overline{1}$	$\overline{2}$
$\overline{3}$	$\overline{3}$	$\overline{1}$	$\overline{2}$	$\overline{4}$

. Diese zweite (äquivalente) Verknüpfungstafel zu $(\mathbb{Z}_5 \setminus \{\overline{0}\}, \cdot)$ ist "ähnlich" wie die Verknüpfungstafel zu $(\mathbb{Z}_4, +)$:

$\cdot$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$
$\overline{0}$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$
$\overline{1}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{0}$
$\overline{2}$	$\overline{2}$	$\overline{3}$	$\overline{0}$	$\overline{1}$
$\overline{3}$	$\overline{3}$	$\overline{0}$	$\overline{1}$	$\overline{2}$

. Wir werden weiter sehen (Satz von Fermat) dass dies kein Zufall ist. Aber zuerst müssen wir das Begriff der "Ähnlichkeit" präzisieren.

Definition 2.7. Seien $(G,\circ),(H,*)$ Gruppen und $f\colon G \to H$ eine Abbildung.

(a) f heißt Gruppenhomomorphismus, falls $\forall a,b \in G\colon f(a\circ b) = f(a) * f(b)$.

(b) f heißt Gruppenisomorphismus $\Leftrightarrow$ f ist ein bijektiver Grummenhomomorphismus.

Bemerkung. Die Verkettung zweier Gruppenhomomorphismen ist wieder ein Gruppenhomomorphismus. Wenn f Gruppenismomorphismus, so ist f^{-1} ebenfalls ein Gruppenhomomorphismus (Übung).

Falls solch ein Isomorphismus $f\colon G \to H$ existiert, so heißen $(G,\circ)$ und $(H,*)$ isomorph. Man schreibt dann auch $G \simeq H$. Im Spezialfall $(H,*) = (G,\circ)$ heißt ein Gruppenisomorphismus ein *Automorphismus* der Gruppe $(G,\circ)$. Sei $\mathrm{Hom}(G,H)$ die Menge der Gruppenhomomorphismen von G nach H, bzw. $\mathrm{Aut}(G) := \mathrm{Hom}(G,G)$. Aus der vorigen Bemerkung folgt dann, $(\mathrm{Aut}(G),\circ)$ ist eine Gruppe (mit der Verkettung $\circ$ der Abbildungen).

Beispiel 2.5. a) $(\mathbb{R},+) \simeq (\mathbb{R}_+^*,\cdot)$ wobei $\mathbb{R}_+^* := \big\{x \in \mathbb{R} \mid x > 0\big\}$. Definiere $\phi\colon \mathbb{R} \to \mathbb{R}_+^*$, $\phi(x) := 2^x$. Es gilt $\phi(x+y) = 2^{x+y} = 2^x \cdot 2^y = \phi(x)\phi(y)$. Weiter ist

ϕ bijektiv mit Umkehrabbildung $\log_2 \colon \mathbb{R}_+^* \to \mathbb{R}$. (Man kann auch eine andere reele Zahl $a > 1$ an der Stelle von 2 nehmen; üblich ist die Eulersche Zahl $e = 2,71...$ zu betrachten, dann ist $\phi(x) := e^x = exp(x)$ die *(natürliche) Exponentialfunktion* und ihre Umkehrfunktion ist $\log_e = \ln$, die *natürliche Logarithmus*-Funktion. Beachten Sie, dass e in diesem Kontext kein neutrales Element in einer Gruppe bezeichnet, sondern eine gewisse reele Zahl!.)

(b) Wir definieren $\phi \colon (\mathbb{Z}_4, +) \to (\mathbb{Z}_5 \setminus \{\overline{0}\}, \cdot)$ durch $\phi([k]_4) \coloneqq [2^k]_5$. Dann ist ϕ ein Gruppenisomorphismus.

Beweis. **1)** Es gilt $2^4 = 16 \equiv 1 \pmod 5$. Also $2^{4p} \equiv 1 \pmod 5$ und $\phi(4p + k) = 2^k \cdot 2^{4p} = 2^k \cdot (2^4)^p \equiv 2^k \cdot 1^p \equiv 2^k = \phi(k) \pmod 5$. Also ist $\phi : \mathbb{Z}_4 \to \mathbb{Z}_5 \setminus \{\overline{0}\}$ wohldefiniert. Es gilt:

$[k]_4$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$
$\phi([k]_4)$	$[1]_5$	$[2]_5$	$[4]_5$	$[3]_5$

womit ϕ bijektiv ist. Es gilt $\phi([k + l]_4) = [2^{k+l}]_5 = [2^k \cdot 2^l]_5 = \phi([k]_4) \cdot \phi([l]_4)$. Also ist ϕ auch ein Gruppenhomomorphismus und mit der Bijektivität also auch bereits Isomorphismus.

□

Lemma 2.3. *Sei $(G, \circ)$ eine Gruppe mit multiplikativer Notation. Dann gilt:*

a) $\forall a, b \in G \colon (a \circ b)^{-1} = b^{-1} \circ a^{-1}$

(b) $\forall a \in G \colon (a^{-1})^{-1} = a$

(c) $\forall a, b \in G \colon \exists! x \in G \ \wedge \ \exists! y \in G \colon a \circ x = b, \ y \circ a = b.$

(d) $\forall a \in G$ *sind die Abbildungen* $l_a, r_a \colon G \to G, l_a(x) = a \circ x, \ r_a(x) = x \circ a$ *bijektiv.*

Beweis. **c) ⇔ d)** x, y sind die Lösungen zu $l_a(x) = b, r_a(y) = b$.

a),b),d) Übung.

□

Die Abbildungen l_a und r_a heißen *die linke (bzw. rechte) Verschiebung um a* oder, falls die Gruppe eine Multiplikative Notation hat, *die linke (bzw. rechte) Multiplikation mit a.*

Bemerkung. x, y müssen in (c) nicht unbedingt die gleichen sein.

Notation 2.8. Wir definieren

$$\begin{aligned} a^n &:= \underbrace{a \circ a \circ .... \circ a}_{n \text{ mal}} \\ a^0 &:= e \\ a^{-n} &:= \underbrace{a^{-1} \circ a^{-1} \circ ... \circ a^{-1}}_{n \text{ mal}} \end{aligned}$$

Übung. Es gilt: $\forall a \in G$,

$$\begin{aligned} a^{m+n} &= a^m \circ a^n \\ (a^m)^n &= a^{m \cdot n} \end{aligned}$$

Notation 2.9. Ist $(G, +)$ eine abelsche Gruppe mit additiver Notation, so ist 0 das neutrale Element, $-x$ das Inverse zu x und man schreibt für ein $n > 0$:

$$\begin{aligned} nx &:= \underbrace{x + ... + x}_{n \text{ mal}} \\ 0x &:= 0 \\ (-n)x &:= \underbrace{(-x) + ... + (-x)}_{n \text{ mal}} \end{aligned}$$

Bemerkung. In der obigen Notation heißt nx nicht, dass $n \in \mathbb{Z}$ mit der "Zahl" x multipliziert wird; in der Tat wird die Notation oft für Gruppen G benutzt, deren Elemente keine Zahlen sind. Es bedeutet ebenfalls auch nicht, dass n ein Element in G sein muss, damit nx das Ergebnis einer Verknüpfung sein soll. Das fehlende Verknüpfungszeichen $+$ oder $\cdot$ ist beabsichtigt, um nx, mit $n \in \mathbb{N}$, $x \in G$ von $a \cdot x$, mit $a, x \in G$ zu unterscheiden.

Übung. Es gilt: $\forall x \in G$,

$$\begin{aligned} (m + n)x &= mx + nx \\ (mn)x &= m(nx) \end{aligned}$$

Natürlich ist diese Übung inhaltlich identisch mit der vorigen (für die Multiplikativ notierte Verknüpfung).

Definition 2.10 (Untergruppe)**.** Es sei $(G, \circ)$ eine Gruppe. Eine nicht-leere Teilmenge $\neq U \subseteq G$ heißt Untergruppe in (von) G, falls gilt:

(UG1) $\forall a, b \in U : a \circ b \in U$.

(UG2) $\forall a \in U : a^{-1} \in U$.

Bemerkung. 1) $(U, \circ)$ ist dann eine Gruppe. (die Einschränkung $\circ_U : U \times U \to U$ der Verknüpfung $\circ$ auf G ist wegen der Bedingung (UG1) in der obigen Definition wohldefiniert; sie wird ebenfalls mit $\circ$ bezeichnet. Die Gruppenaxiome für $(U, \circ)$ folgen aus denen für $(G, \circ)$)

2) Es gilt insbesondere $e \in U$, da $\exists a \in U \overset{(UG2)}{\Rightarrow} a^{-1} \in U \overset{(UG1)}{\Rightarrow} a \circ a^{-1} = e \in U$.

Beispiel 2.6. 1) Es ist $(n\mathbb{Z}, +) \subseteq (\mathbb{Z}, +)$ eine Untergruppe.

2) Es sind $(\mathbb{Z}, +), (\mathbb{Q}, +)$ Untergruppen von $(\mathbb{R}, +)$.

3) Sei $(G, *)$ eine Gruppe. Dann ist die symmetrische Gruppe $S(G)$ von G eine Gruppe bezüglich der Verkettung von Abbildungen. Dann ist $\mathrm{Aut}(G) := \{f \in S(G) \mid f \text{ Gruppenisomorphismus}\}$ eine Untergruppe von $S(G)$: $\forall f, g \in \mathrm{Aut}(G),\ \forall x, y \in G\ : (g \circ f)(x * y) = g(f(x * y)) - g(f(x) * f(y)) = g(f(x)) * g(f(y)) = (g \circ f)(x) * (g \circ f)(y)$, also $g \circ f$ ist auch ein Automorphismus von $(G, *)$. Man nennt $\mathrm{Aut}(G)$ die Automorphismengruppe von G; die Elemente $f \in \mathrm{Aut}(G)$ sind die Automorphismen von G.

4) Sei $f \colon (G, \circ) \to (H, *)$ ein Gruppenhomomorphismus. Dann ist

a) das *Bild* von f, $f(G)$ eine Untergruppe von H.

b) der *Kern* von f, $\ker f := \{x \in G \mid f(x) = e_H\}$ eine Untergruppe in G.

(Übung)

Definition 2.11. Sei $(G, \circ)$ eine Gruppe und $a \in G$. Setze $\langle a \rangle := \{a^k \mid k \in \mathbb{Z}\}$. $\langle a \rangle$ heißt die von a erzeugte *zyklische Untergruppe* von G. Falls $G = \langle a \rangle$ für ein $a \in G$, so heißt G *zyklisch*.

Übung. Man mache sich klar, dass $\langle a \rangle$ tatsächlich eine Untergruppe von G ist.

Beispiel 2.7. 1) $\forall k \in \mathbb{N} \setminus \{0\}$ ist $k\mathbb{Z} := \{kn \ : \ n \in \mathbb{Z}\} \subset \mathbb{Z}$ eine von k erzeugte (zyklische) Untergruppe. Insbesondere ist $(\mathbb{Z}, +)$ zyklisch; sie wird von 1 erzeugt.

2) $(\mathbb{Z}_m, +)$ wird von $[1]$ erzeugt.

3) $(\mathbb{Q}, +)$ ist nicht zyklisch, denn $\langle \frac{p}{q} \rangle = \frac{p}{q}\mathbb{Z}$ enthält nur solche unkürzbare Brüche, für die der Nenner q oder ein Teiler von q ist. Zum Beispiel $\frac{1}{2q} \notin \frac{p}{q}\mathbb{Z}$. Also $\langle a \rangle = a\mathbb{Z} \neq \mathbb{Q}$ für beliebiges $a \in \mathbb{Q}$.

Es folgt aus dem letzten Beispiel auch sofort, dass $(\mathbb{Z}, +)$ und $(\mathbb{Q}, +)$ nicht Isomorph sind, denn sonst wäre $\mathbb{Q}$ von $f(1)$ erzeugt für einen Gruppenisomorphismus $f \colon \mathbb{Z} \to \mathbb{Q}$.

Satz 2.4 (Lagrange)**.** *Sei $(G, \circ)$ eine endliche Gruppe. Sei $U \subseteq G$ eine Untergruppe. Dann ist $|U|$ ein Teiler von $|G|$.*

Beweis. Sei $a \in G$. Die Einschränkung $l_a|_U \colon U \to a \circ U$ der linken Verschiebung um a (definiert durch $l_a(x) := a \circ x$) ist injektiv und damit bijektiv (weil der Zielbereich $a \circ U$ so definiert wurde, damit $l_a|_U$ surjektiv ist). Also $|a \circ U| = \left|l_a(U)\right| = |U|$.

Falls $c \in b \circ U \cap a \circ U$, so $c = a \circ u = b \circ v$ für $u, v \in U$. Dann ist $b^{-1} \circ a \in U$ und $a \circ U = b \circ (b^{-1} \circ a) \circ U = b \circ U$. Also $a \circ U = b \circ U$ oder $a \circ U \cap b \circ U = \emptyset$. Damit ist

$$G = \bigcup_{a \in G} (a \circ U)$$

eine endliche Vereinigung von Mengen, die paarweise teilfremd sind. Sei n ihre Anzahl. Dann $G = a_1 \circ U \cup ... \cup a_n \circ U$ und damit

$$|G| = |a_1 \circ U| + ... + |a_n \circ U| = n|U|$$

da die $a_i \circ U$ paarweise disjunkt sind und $|a_1 \circ U| = \cdots = |a_n \circ U| = |U|$. □

Beispiel 2.8. 1) Sei p eine Primzahl. Dann hat $\mathbb{Z}_p$ keine Untergruppen außer $\left\{\overline{0}\right\}$ und $\mathbb{Z}_p$.

2) Falls $f \colon (\mathbb{Z}_p, +) \to (\mathbb{Z}_q, +)$ ein Gruppenhomomorphismus und $p \neq q$ Primzahlen sind, so ist f konstant.

Beweis: Da $f(\mathbb{Z}_p)$ eine Untergruppe von $\mathbb{Z}_q$ ist, gilt entweder $f(\mathbb{Z}_p) = \{\overline{0}\}$ (dann sind wir fertig) oder f ist surjektiv. Da $\ker(f)$ eine Untergruppe von $(\mathbb{Z}_p)$ ist, gilt entweder $\ker(f) = \mathbb{Z}_p$ ($\Leftrightarrow f \equiv 0$) oder $\ker(f) = \{\overline{0}\} \subset \mathbb{Z}_p$. In diesem Fall ist aber f injektiv, denn $f(\overline{a}) = f(\overline{b}) \Leftrightarrow f(\overline{a} - \overline{b}) = \overline{0} \Leftrightarrow \overline{a - b} \in \ker(f) \Leftrightarrow \overline{a} = \overline{b}$. Dann ist f bijektiv, also $\left|\mathbb{Z}_p\right| = \left|\mathbb{Z}_q\right| \Leftrightarrow p = q$, Widerspruch.

Satz 2.5 (Fermat für Gruppen)**.** *Sei $(G, \circ)$ eine endliche Gruppe mit n Elementen. Dann gilt $\forall a \in G \colon a^n = e$.*

Beweis. $\forall a \in G \colon \langle a \rangle$ ist eine abelsche Untergruppe von G mit k Elementen. Aus dem Satz von Lagrange folgt $n = k \cdot l$ für ein $l \in \mathbb{N}$. Sei $\langle a \rangle = \{x_1, ..., x_k\}$. Dann ist $\{a \circ x_1, ..., a \circ x_k\} = \{x_1, ..., x_k\}$, da $l_a : \langle a \rangle \to \langle a \rangle$ bijektiv ist. Also $a^k \circ x_1 \circ ... \circ x_k = (a \circ x_1) \circ ... \circ (a \circ x_k) = x_1 \circ ... \circ x_k$. Also $a^k = e$. Es folgt $a^n = a^{k \cdot l} = (a^k)^l = e$. □

Satz 2.6 (Sätze von Fermat und Euler)**.** ***(der kleine Satz von) Fermat:*** *Sei p eine Primzahl und $a \in \mathbb{Z}$ so dass $p \nmid a$. Dann $a^{p-1} \equiv 1 \pmod{p}$.*

Satz von Euler: *Sei $n \in \mathbb{N}$. Sei $\varphi(n) := \left|\left\{k \in \{1, .., n\} \mid \mathrm{ggT}(k, n) = 1\right\}\right|$ die Eulersche Phifunktion. $\forall a \in \mathbb{Z} \colon \mathrm{ggT}(a, n) = 1 \Rightarrow a^{\varphi(n)} \equiv 1 \pmod{n}$.*

Bemerkung. $\varphi(n)$ ist die Anzahl der zu n Teilerfremden Zahlen zwischen 1 und n. Der kleine Satz von Fermat ist also ein Spezialfall (für n eine Primzahl) des Satzes von Euler.

Beispiel 2.9. $\varphi(p) = p-1 \Leftrightarrow p$ ist prim. $\varphi(pq) = (p-1)(q-1)$ falls $p \neq q$ Primzahlen sind.

Beweis. Sei $a \in \{1, ..., pq-1\}$ so dass $\mathrm{ggT}(a, pq) > 1$. Dann ist $\mathrm{ggT}(a, pq) = p$ oder q. Es gibt genau $q-1$ Zahlen $a \in \{1, .., pq-1\}$, die durch q teilbar sind: $q, 2q, ..., (p-1)q$ und genau $p-1$ Zahlen $a \in \{1, .., pq-1\}$, die durch p teilbar sind: $p, 2p, .., (q-1)p$. Diese insgesamt $(p-1)+(q-1)$ Zahlen sind paarweise verschieden, also $\varphi(pq) = pq - (p-1) - (q-1) - 1 = pq - p - q + 1 = (p-1)(q-1)$. □

Beweis des Satzes von Euler. Sei $G := \left\{[k]_n \colon \mathrm{ggT}(k, n) = 1\right\} \subseteq \mathbb{Z}_n$. Es gilt $|G| = \varphi(n)$. Falls a, b teilerfremd zu n sind, so ist auch ab. Also induziert die Multiplikation in $\mathbb{Z}_n$ eine Verknüpfung auf G, die assoziativ und kommutativ ist und ein neutrales Element $[1]_n$ besitzt. Nach dem Satz von Bézout hat jedes Element $[k]_n \in G$ ein inverses Element $[l]_n \in G$, da es $l, m \in \mathbb{Z}$ gibt so dass $kl + nm = 1$ ist. Also ist $(G, \cdot)$ eine Gruppe mit $\varphi(n)$ Elementen und der Satz von Fermat für Gruppen impliziert, $\forall [a]_n \in G \colon [a]_n^{\varphi(n)} = [a^{\varphi(n)}]_n = [1]_n$, also $a^{\varphi(n)} \equiv 1 \pmod{n}$. □

Bemerkung. Die RSA-Verschlüsselung ist eine Anwendung des Satzes von Euler (Fermat) für $n = pq$ für p, q sehr große Primzahlen. Das Grundprinzip (in der Praxis werden zusätzliche Sicherheitsverfahren verwendet, die die Verschlüsselung von Angriffen schützen) wird hier erklärt:

Der Nutzer A wählt $p, q > N_0$ Primzahlen (N_0 groß genug, siehe unten), und rechnet $n := pq$ und $\varphi(n) = (p-1)(q-1)$ aus. Er wählt dazu eine Primzahl l, die $\varphi(n)$ nicht teilt. Nach dem Satz von Bézout gibt es also $s \in \mathbb{N}$, so dass

$$\varphi(n) \mid s \cdot l - 1.$$

Der Satz von Euler impliziert, dass $(m^l)^s \equiv m \pmod{n}$, $\forall m \in \mathbb{Z}$ für alle m teilerfremd zu n, zum Beispiel für alle $m \in \{1, ..., N_0 - 1\}$.

Jetzt veröffentlicht Nutzer A die Zahlen n und l; ein Nutzer B, der A eine Nachricht (in der Form einer Zahl m) schicken will, muss $k := m^l \pmod{n}$ ausrechnen, und das Ergebnis A schicken.

A braucht nun $k^s \pmod{n}$ auszurechnen, und bekommt m wieder.

Der Erfolg dieser Verschlüsselungsmethode liegt in der Schwierigkeit, eine Zahl n in Primfaktoren zu zerlegen, um $\varphi(n)$ auszurechnen (und so s zu bekommen). Im Fall einer hundertstelligen Zahl N_0 gibt es relativ wirksame Algorithmen, die Primzahlen p und q mit $p, q > N_0$ finden. Dagegen ist die Primfaktorzerlegung von $n = pq$ so aufwendig, dass ein leistungsstarken Computer mehrere Jahre dafür brauchen könnte.

Die Rechnungen, die A und B machen müssen, sind einfach, und für N_0 groß genug, ist es kein Problem eine Nachricht durch eine Zahl $m < N_0 - 1$ darzustellen (diese

Bedingung garantiert, dass $\text{ggT}(m, \varphi(n)) = 1$). Ein eventueller Angreifer kann in der Praxis die Zahl s aus den veröffentlichten Daten wegen dem hohen Rechenaufwand nicht herleiten.

2.1 Ringe

Definition 2.12. Eine Menge R mit zwei Verknüpfungen

$$+\colon R \times R \to R, (a,b) \mapsto a+b \quad (\text{"Addition"})$$
$$\cdot\colon R \times R \to R, (a,b) \mapsto a \cdot b \quad (\text{"Multiplikation"})$$

heißt Ring, wenn gilt:

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

(b) Die Multiplikation $\cdot$ ist assoziativ. Man sagt dann auch, dass $(R,\cdot)$ eine Halbgruppe ist.

(c) $\forall a,b,c \in R\colon (a+b)\cdot c = a\cdot c + b\cdot c \quad \wedge \quad c\cdot(a+b) = c\cdot a + c\cdot b$,
d.h. $+,\cdot$ verhalten sich distributiv.

Das neutrale Element der Addition wird mit 0 bezeichnet und heißt Nullelement. Das inverse Element zu $a \in R$ bezüglich der Addition wird mit $-a$ bezeichnet. Wir definieren $a - b \coloneqq a + (-b)$.

Definition 2.13 (Kommutativer Ring, Einselement, nullteilerfrei)**.** Sei $(R,+,\cdot)$ ein Ring.

(a) R heißt kommutativ, wenn gilt:

$$\forall a,b \in R\colon a\cdot b = b\cdot a$$

(b) Ein Element 1 heißt Einselement, wenn gilt:

$$\forall a \in R\colon a\cdot 1 = 1\cdot a = a$$

(c) Ein Element $a \in R \setminus \{0\}$ heißt Nullteiler, wenn $\exists b \in R\setminus\{0\}$, so dass $a\cdot b = 0 \ \vee \ b\cdot a = 0$. Der Ring R heißt nullteilerfrei, falls es keine Nullteiler in R gibt, d.h., es gilt:

$$\forall a,b \in R\colon a\cdot b = 0 \Rightarrow a = 0 \vee b = 0$$

Falls

Bemerkung. Das Einselement ist eindeutig bestimmt, falls es existiert.

Beispiel 2.10. 1. $(\mathbb{Z},+,\cdot)$ ist ein kommutativer nullteilerfreier Ring mit Einselement 1.

2. $(m\mathbb{Z},+,\cdot)$ ist für $2 \leq m \in \mathbb{N}$ ein kommutativer nullteilerfreier Ring ohne Einselement.

3. $(\mathbb{Z}_m, +, \cdot)$ ist für $m \geq 2$ ein kommutativer Ring mit Einselement. Er ist nullteilerfrei genau dann, wenn m eine Primzahl ist. Dass $(\mathbb{Z}_m, +)$ eine abelsche Gruppe ist, wissen wir bereits. Mit $[a]_m \cdot [b]_m := [a \cdot b]_m$ folgt die Assoziativität von $\cdot$ direkt mit der aus $\mathbb{Z}$ und ebenso die Distributivität. Weiter ist $[1]_m$ dann das Einselement.
Ist $m \geq 2$ keine Primzahl, so gibt es $a, b \in \mathbb{N}, a, b \geq 2$: $m = a \cdot b$. Dann aber $[a]_m \cdot [b]_m = [a \cdot b]_m = [m]_m = [0]_m$, aber $[a]_m \neq [0]_m, [b]_m \neq [0]_m$, also gibt es Nullteiler.
Ist m eine Primzahl, so $[a]_m \cdot [b]_m = [0]_m \Leftrightarrow [a \cdot b]_m = [0]_m \Leftrightarrow m \mid ab \Rightarrow m \mid a \vee m \mid b \Leftrightarrow [a]_m = [0]_m \vee [b]_m = [0]_m$.

Lemma 2.7 (Rechenregeln in Ringen)**.** *Sei $(R, +, \cdot)$ ein Ring und $a, b, c \in R$. Dann gilt*

(a) $0 \cdot a = a \cdot 0 = 0$

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

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

(d) *R hat Einselement* $\Rightarrow -a = (-1) \cdot a = a \cdot (-1)$

(e) *R nullteilerfrei* $\Rightarrow \begin{cases} (c \neq 0 \wedge a \cdot c = b \cdot c) \Rightarrow a = b \\ (c \neq 0 \wedge c \cdot a = c \cdot b) \Rightarrow a = b \end{cases}$ *("Kürzungsregel")*

Beweis. **(a)** $0 \cdot a + 0 \cdot a \overset{\text{Distr.}}{=} (0 + 0) \cdot a = 0 \cdot a \Rightarrow 0 \cdot a + 0 \cdot a = 0 \cdot a \Rightarrow 0 \cdot a = 0$, da man in Gruppen kürzen darf. Analog folgt $a \cdot 0 = 0$.

(b) $a \cdot b + (-a) \cdot b \overset{\text{Distr.}}{=} (a + (-a)) \cdot b = 0 \cdot b = 0$ und $a \cdot b + -(a \cdot b) = 0 \Rightarrow -(a \cdot b) = (-a) \cdot b$ wie oben. Die zweite Gleichung zeigt man analog.

(c) $(-a) \cdot (-b) \overset{\text{(b)}}{=} -(a \cdot (-b)) = -(-(a \cdot b)) = a \cdot b$

(d) $(-1) \cdot a = -(1 \cdot a) = -a$ und zweite Gleichung analog.

(e) Sei R nullteilerfrei und $c \neq 0$:

$$a \cdot c = b \cdot c \Rightarrow a \cdot c + (-(b \cdot c)) = 0 \Rightarrow a \cdot c + (-b) \cdot c = 0$$
$$\overset{\text{Distr.}}{\Rightarrow} (a + (-b)) \cdot c = 0 \overset{\text{R nullteilerfrei}}{\Rightarrow} a + (-b) = 0 \Rightarrow a = -(-b) \Leftrightarrow a = b$$

□

2.2 Körper

Definition 2.14. $(K, +, \cdot)$ heißt Körper, wenn gilt:

(a) $(K, +, \cdot)$ ist ein Ring.

(b) $(K \setminus \{0\}, \cdot)$ ist eine abelsche Gruppe.

Mit 1 bezeichnen wir das neutrale Element von $K \setminus \{0\}$. Wir definieren $\frac{a}{b} \coloneqq a \cdot b^{-1}$ ("Division")

Bemerkung. 1. Jeder Körper ist ein kommutativer nullteilerfreier Ring mit 1 als Einselement und $1 \neq 0$:

$$\left.\begin{array}{l} \forall a, b \in K \setminus \{0\} : a \cdot b = b \cdot a \text{ (wegen (b))} \\ \forall a \in K : a \cdot 0 = 0 \cdot a \text{ (Lemma von oben)} \end{array}\right\} \Rightarrow K \text{ kommutativer Ring.}$$

$$\left.\begin{array}{l} \forall a \in K \setminus \{0\} : a \cdot 1 = 1 \cdot a \text{ (wegen (b))} \\ 0 \cdot 1 = 1 \cdot 0 \text{ (wegen obigem Lemma)} \end{array}\right\} \Rightarrow K \text{ Ring mit 1 als Einselement. Es}$$

gilt $1 \neq 0$, da K (wegen (b)) mindestens 2 Elemente enthalten muss, jedoch $1 = 0$ nur im Fall eines einelementigen Ringes gilt.
Sei nun $a \cdot b = 0$. Zeige $a = 0 \ \vee \ b = 0$. Angenommen $a \neq 0$. Dann ist $a^{-1} \cdot (a \cdot b) = 0$, also $b = 0$.

2. Endliche nullteilerfreie kommutative Ringe mit $1 \neq 0$ sind Körper. (evtl. Übung)

Beispiel 2.11. 1. $(\mathbb{Q}, +, \cdot)$ und $(\mathbb{R}, +, \cdot)$ sind Körper. $(\mathbb{Z}, +, \cdot)$ ist kein Körper.

2. Für p Primzahl ist $(\mathbb{Z}_p, +, \cdot)$ ein Körper. Für $m \in \mathbb{N}, m$ keine Primzahl, ist $(\mathbb{Z}_m, +, \cdot)$ *kein* Körper, weil $(\mathbb{Z}_m, +, \cdot)$ nicht nullteilerfrei ist.

3 Polynome

Definition 3.1. Sei $(R, +, \cdot)$ ein kommutativer Ring. Ein Polynom mit Koeffizienten in R ist eine endliche Folge $\varphi_f \colon \mathbb{N} \to R,\ f_n \coloneqq \varphi_f(n)$, so dass $\exists N \in \mathbb{N}\ \forall n > N \colon f_n = 0 \in R$. Die Folge hat also nur endlich viele Einträge $\neq 0$.

Man schreibt $f(X) = f_0X^0 + f_1X^1 + ... + f_NX^N$. Definiere weiter $\operatorname{grad}(f) \coloneqq \max(\{n \in \mathbb{N} \mid f_n \neq 0\} \cup \{-\infty\})$.

Das Nullpolynom 0 ist die Folge $\varphi_0 \coloneqq (0, 0, ...)$. Es gilt $\operatorname{grad}(0) = -\infty$ und $\operatorname{grad}(f) \in \mathbb{N}\ \forall f \neq 0$ Polynom.

$f_0, ..., f_N$ heißen die Koeffizienten von f. $N = \operatorname{grad}(f) \Leftrightarrow f_N \neq 0\ \wedge\ \forall n > N\ : f_n = 0$.

Falls R ein Einselement 1_R (weiter kurz 1 genannt) hat, wird 1_R als Koeffizient meistens nicht geschrieben und, falls -1_R ein Koeffizient ist, so wird nur das Vorzeichen geschrieben, wie in $X^2 - X \coloneqq 1_R \cdot X^2 + (-1_R) \cdot X \in R[X]$.

Beispiel 3.1. $f(X) = 1 + X - X^2 + 2X^3 \in \mathbb{Z}[X]$, also $\varphi_f = (1, 1, -1, 2, 0, 0, ...) \in \{f \colon \mathbb{N} \to \mathbb{Z}\}$. Es ist $\operatorname{grad}(f) = 3$.

Definition 3.2. Die Menge aller Polynome mit Koeffizienten in R heißt $R[X]$, wobei X die Variable ist und $f_0, ..., f_N$ die Koeffizienten eines Polynomes f.

Wir führen auf $R[X]$ die folgenden Verknüpfungen ein:

Addition: $\forall f, g \in R[X]$ entspricht $f + g$ der Funktion $\varphi_f + \varphi_g \colon \mathbb{N} \to R,\ (\varphi_f + \varphi_g)(n) \coloneqq \varphi_f(n) + \varphi_g(n) = f_n + g_n \in R$. Die Addition ist assoziativ und kommutativ, hat 0 (das Nullpolynom: alle Koeffizienen sind Null) als neutrales Element und $\forall f \in R[X]$ ist $-f$ gegeben durch $-(\varphi_f)$, d.h. $\forall n \in \mathbb{N} \colon \varphi_{-f}(n) \coloneqq -\varphi_f(n)$, das Inverse zu f.

Beispiel 3.2. $f(X) = 1 - X^2 \Rightarrow (-f)(X) = -1 + X^2$

Mit der hier definierten addition wird $(R[X], +)$ zu einer abelschen Gruppe.

Multiplikation: Wir setzen $\varphi_{fg}(n) \coloneqq \sum_{k=0}^{n} f_k \cdot g_{n-k}$. Es ist $\varphi_{fg}(n) = 0$, falls $n > \operatorname{grad}(f) + \operatorname{grad}(g)$, weil in jedem Term $f_k \cdot g_{n-k}$ ist $k > \operatorname{grad}(f)$ oder $n - k > \operatorname{grad}(g)$, also das Produkt ist 0.

Beispiel 3.3. $f(X) = 1 - X^2$, $g(X) = 1 + 2X + X^2$. Die Koeffizienten sind dann $f_0 = 1, f_1 = 0, f_2 = -1, g_0 = 1, g_1 = 2, g_2 = 1$ und $\operatorname{grad}(f) = 2, \operatorname{grad}(g) = 2$. Die Methode zur Multiplikation ist

g \ f	1	0	−1
1	1	0	−1
2	2	0	−2
1	1	0	−1

woran man ablesen kann: $(fg)_0 = 1, (fg)_1 = 2+0, (fg)_2 = 1+0+-1, (fg)_3 = 0+(-2), (fg)_4 = -1$.
Also $(1-X^2)(1+2X+X^2) = 1+2X-2X^3-X^4$.

Die Variable X kommutiert mit allen Koeffizienten.

Satz 3.1. *$(R[X], +, \cdot)$ ist ein kommutativer Ring. Falls R nullteilerfrei ist, so ist $R[X]$ nullteilerfrei. Falls $1 \in R$ Einselement ist, so ist das Polynom 1 ($\varphi_1 = (1,0,0,....)$) ein Einselement in $R[X]$. $R[X]$ ist kein Körper.*

Beweis. **Kommutativität:** Die Multiplikation ist kommutativ, da

$$\forall f, g \in R[X]\colon (fg)_n = \sum_{k=0}^{n} f_n \cdot g_{n-k}$$

$$= f_0 \cdot g_n + f_1 \cdot g_{n-1} + ... + f_{n-1} \cdot g_1 + f_n \cdot g_0 = \sum_{l=0}^{n} g_l \cdot f_{n-l} = (g \cdot f)_n$$

Assoziativität: Seien $f, g, h \in R[X]$. Wir bezeichnen die Koeffizienten mit $f_n, g_n, h_n, (fg)_n$ etc. Dann gilt für alle $n \in \mathbb{N}$:

$$((f \cdot g) \cdot h)_n = \sum_{k=0}^{n} (f \cdot g)_k \cdot h_{n-k} = \sum_{k=0}^{n} \sum_{l=0}^{k} f_l \cdot g_{k-l} \cdot h_{n-k} \overset{\substack{Distr.\\ Komm.}}{=} \sum_{0 \le l \le k \le n} f_l \cdot g_{k-l} \cdot h_{n-k}$$

$$= \sum_{l=0}^{n} f_l \cdot \sum_{k=l}^{n} g_{k-l} \cdot h_{n-k} = \sum_{l=0}^{n} f_l \cdot \sum_{j=0}^{n-l} g_j \cdot h_{(n-l)-j} = \sum_{l=0}^{n} f_l \cdot (g \cdot h)_{n-l} = (f \cdot (gh))_n$$

Distributivität: Seien $f, g, h \in R[X]$. Für alle $n \in \mathbb{N}$ gilt:

$$((f+g) \cdot h)_n = \sum_{k=0}^{n} (f+g)_k \cdot h_{n-k} = \sum_{k=0}^{n} (f_k + g_k) \cdot h_{n-k} = \sum_{k=0}^{n} (f_k \cdot h_{n-k} + g_k \cdot h_{n-k})$$

$$= \sum_{k=0}^{n} f_k \cdot h_{n-k} + \sum_{k=0}^{n} g_k \cdot h_{n-k} = (f \cdot h)_n + (g \cdot h)_n$$

Also ist $R[X]$ ein kommutativer Ring.
Sei nun R nullteilerfrei. Seien $f, g \in R[X]$. Sei $f \cdot g = 0$ und $f, g \neq 0$. Also $N = \text{grad}(f) \in \mathbb{N}$. Falls $g \neq 0$, so $M = \text{grad}(g) \in \mathbb{N}$ und $(f \cdot g)_{M+N}$ hat genau einen Koeffiezient, der nicht offensichtlich 0 ist: $(f \cdot g)_{M+N} = f_N \cdot g_M$. Dieses Koeffizient ist dann nicht Null, da $f_N \neq 0, g_M \neq 0$ und R nullteilerfrei ist. Also ist $R[X]$ nullteilerfrei.
Sei nun $1 \in R$. Das Polynom f mit $f_0 = 1$, $f_n = 0\ \forall n \in \mathbb{N} \setminus \{0\}$ ist ein neutrales

Element für die Multiplikation der Polynome, denn in allen Summen, die die Koeffizienten vom Polynom f mit einem Polynom $g \in R[X]$ bestimmen, nur ein einziges Term nicht Null sein kann:

$$(g \cdot f)_n = \sum_{k=0}^{n} g_k \cdot f_{n-k} = g_n \cdot 1 = g_n.$$

Weiter ist $R[X]$ kein Körper, da X kein multiplikatives Inverses besitzt. □

Notation 3.3. $X, X^2, ...$ heißen *Monome*. X^k ist das Polynom gegeben durch die Folge

$$(0, 0, ..., \underbrace{1}_{k\text{-te Stelle}}, 0, 0, ...)$$

$-X^k \coloneqq (-1) \cdot X^k$, also zum Beispiel $X - X^2 = 1 \cdot X + (-1) \cdot X^2$.
$f_n X^n$ heißen *Terme*, f_0 heißt *konstanter Term*. f_N mit $\text{grad}(f) = N \in \mathbb{N}$ heißt *Leitkoeffizient* von f. In diesem Fall heißt $f_N X^N$ *Leitterm* (*führender Term*) von f.

Lemma 3.2. *Sei R ein Ring und seien $f, g \in R[X]$. Dann gilt*

1. $\text{grad}(f + g) \leq \max\{\text{grad}(f), \text{grad}(g)\}$ *mit Gleichheit, falls* $\text{grad}(f) \neq \text{grad}(g)$.

2. $\text{grad}(f \cdot g) \leq \text{grad}(f) + \text{grad}(g)$. *Ist R sogar nullteilerfrei, so gilt stets* $\text{grad}(f \cdot g) = \text{grad}(f) + \text{grad}(g)$.

Konvention:

$$-\infty + n = -\infty,\ -\infty + (-\infty) = -\infty,\ \forall n \in \mathbb{N}\colon\ -\infty < n$$

Mit dieser Konvention können wir alle Fälle gleichzeitig betrachten, ohne das, das Nullpolynom eine Ausnahme zu den Formeln im obigen Satz bildet.

Sei K ein Körper. Dann ist $K[X]$ ein kommutativer, nullteilerfreier Ring mit Einselement.

3.1 Teilbarkeit im Polynomring

Definition 3.4 (Teilbarkeit)**.** Seien $f, g \in K[X]$, $f \neq 0$. g ist durch f teilbar $:\Leftrightarrow$ $\exists h \in K[X]\colon g = f \cdot h$.

Eigenschaften der Teilbarkeit:

(a) $f \mid g\ \wedge g\ \neq 0 \Rightarrow \text{grad}(f) \leq \text{grad}(g)$

(b) $f \mid g\ \wedge g\ \mid f \Rightarrow \exists a \in K \setminus \{0\} : f = a \cdot g,\ g = a^{-1} \cdot f$

(c) $f \mid g\ \wedge\ g \mid h \Rightarrow f \mid h$

(d) $f \mid g \wedge f \mid h \Rightarrow f \mid \alpha \cdot g + \beta \cdot h$ für $\alpha, \beta \in K[X]$

(e) $f \mid \alpha \ \wedge \ g \mid \beta \Rightarrow f \cdot g \mid \alpha \cdot \beta$

Satz 3.3 (Division mit Rest (Polynomdivision))**.** *Sei* $f \in K[X] \setminus \{0\}, g \in K[X]$. *Dann* $\exists! q, r \in K[X]\colon \operatorname{grad}(r) < \operatorname{grad}(f) \ \wedge \ g = f \cdot q + r$

Beweis. Sei $N \coloneqq \operatorname{grad}(f) \in \mathbb{N}$ und $n = \operatorname{grad}(g)$. falls $n < N$, so setze $q = 0, r = g$. Induktion nach $n \geq N$:

IA: Sei $n = N$ und sei $q_0 \coloneqq f_N \cdot g_N^{-1}$. Weiter $q_i \coloneqq 0 \ \forall i > 0$. Also q konstant. $q(X) = q_0, f(X) = f_0 + ... + f_N X^N, g(X) = g_0 + ... + g_n X^n$. Sei $r \coloneqq g - f \cdot q$. Dann $\operatorname{grad}(r) \leq N$, aber $r_N = f_N - q \cdot g_N = 0$, also $\operatorname{grad}(r) < N$ und $g = f \cdot q + r$.

IS: Sei $n > N$ und sei $p_{n-N} \coloneqq f_N \cdot g_n^{-1}$ und $p_i = 0 \ \forall i \in \mathbb{N} \setminus \{n - N\}$. Dann $g' \coloneqq g - p \cdot f$ hat $\operatorname{grad}(g') < \operatorname{grad}(g)$ (folgt analog wie im IA). Also $g' = q' \cdot f + r$ nach Induktionsvoraussetzung. Damit ist dann $g = (p + q') \cdot f + r$.

Die Eindeutigkeit von q, r rechnet man leicht nach. □

Definition 3.5. Ein Polynom $f \in K[X] \setminus \{0\}$ heißt irreduzibel, falls $\nexists g, h \in K[X]$ mit $f = g \cdot h$ und $0 < \operatorname{grad}(g), \operatorname{grad}(h) < \operatorname{grad}(h) < f$.

Bemerkung. $\operatorname{grad}(f) \in \{0, 1\} \Rightarrow f$ irreduzibel.

Beispiel 3.4. $f(X) \coloneqq X^2 + 1 \in \mathbb{Q}[X]$. Falls $\exists g, h \in \mathbb{Q}[X]$ mit $\operatorname{grad}(g), \operatorname{grad}(h) < \operatorname{grad}(f) = 2$, so $\operatorname{grad}(g) = \operatorname{grad}(h) = 1$, also $g(X) = aX + b, h(X) = cX + d$ mit $a, b, c, d \in \mathbb{Q}$. Dann

$$(g \cdot h)(X) = (aX + b)(cX + d) = acX + (ad + bc)X + bd$$

Also $gh = f \Leftrightarrow \left\{ \begin{matrix} ac = 1 \Leftrightarrow c = a^{-1} \\ ad + bc = 0 \\ bd = 1 \Leftrightarrow d = b^{-1} \end{matrix} \right\} \Leftrightarrow ab^{-1} + ba^{-1} = 0 \Leftrightarrow a^2 + b^2 = 0 \overset{a,b \in \mathbb{Q}}{\Rightarrow} a = b = 0$. Widerspruch, also ist $X^2 + 1 \in \mathbb{Q}[X]$ irreduzibel.

Bmerkung: Wir können nun, dass wir K mit den konstanten Polynomen in $K[X]$ identifiziert haben, die Schreibweise

$$\sum_{k=0}^{n} f_n \cdot X^n$$

als eine Summe von den Produkten $f_k \cdot X^k$ von den konstanten Polynomen $f_k \in K \subset K[X]$ mit den Polynomen (Monomen) X^k, $k \in \{0, \ldots, n\}$. Das Ergebnis dieser Summe ist das Polynom

$$\sum_{k=0}^{n} f_k X^k,$$

das also ebenfalls mit "$\cdot$" zwischen Koeffizienten und Monomen geschrieben werden kann.

3.2 Ring- und Körperhomomorphismen

Wir möchten nun noch einige Begriffe für allgemeine Ringe einführen, die beim Studium von Polynomen relevant sind:

Definition 3.6. Seien $(R, +, \cdot)$ und $(S, \oplus, \odot)$ zwei Ringe. Eine Funktion $f\colon R \to S$ heißt Ringhomomorphismus, falls gilt:

1) $\forall a, b \in R\colon f(a + b) = f(a) \oplus f(b)$

2) $\forall a, b \in R\colon f(a \cdot b) = f(a) \odot f(b)$

3) Falls R, S Einselemente $1_R \in R$, $1_S \in S$ besitzen, so wird zusätzlich $f(1_R) = 1_S$ verlangt.

Sind R, S Körper, so heißt f ein Körperhomomorphismus.
Ist f bijektiv, so heißt f ein Ringisomorphismus.

Bemerkung. Falls $(K, +, \cdot)$ ein Körper und $(R, \oplus, \odot)$ ein Ring mit Einselement $(1 \neq 0$ ist, so ist jeder Ringhomomorphismus $f\colon K \to R$ injektiv.

Beweis. $f(a) = f(b) \Leftrightarrow f(a - b) = 0_R$. Falls $a - b \neq 0_K$, so existiert $c = (a - b)^{-1}$. Dann

$$f(c \cdot (a - b)) = f(c) \odot f(a - b) = 0_R$$

aber es gilt $c \cdot (a - b) = 1_K$, also $f(c \cdot (a - b)) = f(1_K) = 1_R \neq 0_R$. Widerspruch. Also $a = b$ und f ist injektiv. □

Insbesondere ist jeder Körperhomomorphismus injektiv.

Beispiel 3.5. $K \xrightarrow{\psi} K[X]$, $\varphi(a) \coloneqq (a, 0, 0, ...)$ oder $\psi(a)(X) \coloneqq a(+0 \cdot X) + ...$ besteht also aus dem konstanten Term a. ψ ist ein Ringhomomorphismus. Wir bezeichnen auch mit $a \in K$ das Polynom $\psi(a) \in K[X]$. Da ψ injektiv ist, können wir somit K mit seinem Bild in $K[X]$ identifizieren.

Bemerkung. $K[X]$ ist konstruiert wie folgt: Zu K geben wir die Variable X an und noch alle Kombinationen (endliche Summen und Produkte) aus Elementen aus $K \cup \{X\}$, so dass wir ein geschlossenes System (unter Addition und Multiplikation) erhalten.

Definition 3.7. Sei $(K, +, \cdot)$ ein Körper und $\mathrm{Abb}(K, K) : \{f \ : \ K \to K$ Abbildung$\}$. Der Einsetzhomomorphismus (Einsetzungshomomorphismus) $F\colon K[X] \to \mathrm{Abb}(K, K)$ ist definiert durch:

$$\forall f(X) = \sum_{k=0}^{n} f_k \cdot X^k \in K[X] \ ; \ F(f) \coloneqq \overline{f} : K \to K \ : \ \forall a \in K \ : \ \overline{f}(a) \coloneqq \sum_{k=0}^{n} f_k a^k.$$

Für ein Polynom $f \in K[X]$ heißt $\overline{f}$ die Polynomialfunktion zu f.

Sei M eine nicht leere Menge und $(R, +, \cdot)$ ein Ring mit Einselement. Wir definieren nun eine Ringstruktur auf $\mathrm{Abb}(M, R)$:

$$\forall f, g : M \to R,\ (f+g) : M \to R\ :\ \forall a \in M\ :\ (f+g)(a) := f(a) + g(a),$$

$$\forall f, g : M \to R,\ (f \cdot g) : M \to R\ :\ \forall a \in M\ :\ (f \cdot g)(a) := f(a) \cdot g(a).$$

$\mathrm{Abb}(M, R)$ ist dann ein Ring (Übung). Insbesondere ist $\mathrm{Abb}(K, K)$ ein Ring.

Satz 3.4. *F ist ein Ringhomomorphismus.*

(evtl. Übung)

3.3 Nullstellen

Definition 3.8. Sei $f \in K[X]$. Ein $a \in K$ heißt *Nullstelle* von f in K, falls $\overline{f}(a) = 0$.

Beispiel 3.6. 1) $f(X) = X - a$ für $a \in K$ hat genau die Nullstelle a.

2) Sei p eine Primzahl. $f(X) := X^p - X \in \mathbb{Z}_p[X]$ hat p Nullstellen in $\mathbb{Z}_p$. Nach den Satz von Fermat ist $a^{p-1} = [1]_p$, also $a^p = a$ für alle $a \in \mathbb{Z}_p \setminus \{[0]_p\}$. Diese letzte Gleichheit gilt aber offenbar auch für $a = [0]_p$. Damit also

$$\forall a \in \mathbb{Z}_p \colon \overline{f}(a) = a^p - a = [0]_p$$

Dies ist ein Beispiel eines Polynoms $\neq 0$, dessen Polynomfunktion $\overline{f} \colon K \to K$ jedoch die Nullabbildung ist. Für unendliche Körper ist dies hingegen nicht möglich (siehe später).

Bemerkung: In diesem Beispiel haben wir unsere Standardnotation von $[1]_p, [0]_p$ fuer das Eins-, bzw. Nullelement in $\mathbb{Z}_p$ verwendet. Diese Elemente könnten eigentlich ebenfalls mit 1, bzw. 0 bezeichnet werden (unsere einheitliche Notation für das Eins-, bzw. Nullelement in einem Ring). Leider ist in der Mathematik nur selten möglich, eine perfekt eindeutige Notation zu verwenden, ohne dass der Text schnell überladen wird mit Indizen, Klammern, usw. Der Leser muss also stets *den Kontext* im Auge behalten, in welchem eine "0", eine "1", ein "+", ein "$\cdot$", usw. eigentlich gemeint sind (also in welchem Ring sie zu verstehen sind). Für Elemente in $\mathbb{Z}_p$ werden wir trotzdem eine Notation verwenden, die deutlich zeigt, dass es *nicht um ganze Zahlen geht*, denn die Menge $\mathbb{Z}_p$ der Restklassen modulo p liegt der Menge $\mathbb{Z}$ nahe genug, dass eine Verwechselung leicht passieren kann.

Eine etwas weniger beladene Notation für die Elemente aus $\mathbb{Z}_m$, $m \in \mathbb{N} \setminus \{0, 1\}$ ist $\bar{k} := [k]_m$. Für Rechnungen ist sie angebracht. Dies enthält keine Information über m, erfüllt jedoch den Zweck, deutlich zu machen, dass $\bar{k}$ keine ganze Zahl ist, sondern eine Restklasse.

Satz 3.5 (Teilbarkeitskriterium für lineare Faktoren)**.** *Sei $f \in K[X]$ und $a \in K$. Dann gilt $(X - a) \mid f \Leftrightarrow \overline{f}(a) = 0$.*

Beweis. "⇒" Falls $f(X) = g(X) \cdot (X - a)$ so wenden wir den Einsetzhomomorphismus an:

$$\overline{f}(a) = \overline{g}(a) \cdot (a - a) = 0$$

"⇐" Sei $f(X) = g(X) \cdot (X - a) + b$ nach Division mit Rest durch $(X - a)$, also $b \in K$. Dann

$$0 = \overline{f}(a) = \overline{g}(a) \cdot (a - a) + b = 0 + b$$

Also $b = 0$ und $(X - a) \mid f(X)$.

□

Beispiel 3.7. Sei $f(X) \coloneqq X^p - X + \overline{1} \in \mathbb{Z}_p[X]$ für p Primzahl. f ist durch kein Polynom vom Grad 1 teilbar, denn so ein Polynom würde eine Nullstelle haben, die auch eine von f sein müsste, aber $\overline{f}(a) = \overline{1} \neq \overline{0}$ für alle $a \in \mathbb{Z}_p$ nach obigem Beispiel. (Hier haben wir die Notation $\bar{k} \coloneqq [k]_p$ verwendet.)

Definition 3.9. Ein Polynom $f \in K[X] \setminus \{0\}$ heißt *normiert*, falls der Leitkoeffizient $f_{\text{grad}(f)} = 1 \in K$.

Satz 3.6. *Sei $f \in \mathbb{Q}[X]$ ein normierter Polynom mit Koeffizienten in $\mathbb{Z}$. Dann ist jede Nullstelle $a \in \mathbb{Q}$ von f eine ganze Zahl.*

Beweis. Sei $f(X) = \sum_{i=0}^{n} a_i X^i$, mit $a_i \in \mathbb{Z}$, $\forall i \in \{0, ..., n\}$ und sei $a \coloneqq \frac{p}{q} \in \mathbb{Q}$ eine Nullstelle von f, mit $p, q \in \mathbb{Z}$, $q > 0$, $\text{ggT}(p, q) = 1$. Dann gilt $f(a) = 0 \Leftrightarrow$

$$\sum_{i=0}^{n} a_i \left(\frac{p}{q}\right)^i = 0 \;\Leftrightarrow\; \sum_{i=0}^{n} a_i p^i q^{n-i} = 0.$$

Insbesondere sind alle Terme $a_i p^i q^{n-i}$ (für $i \in \{0, ..., n-1\}$) durch q teilbar; demzufolge ist auch der Term $a_n p^n = p^n$ (denn f ist normiert) durch q teilbar. Falls $q > 1$, jeder Primfaktor von q muss also p teilen, was ein Widersprch zur Annahme $\text{ggT}(p, q) = 1$ darstellt. Also $q = 1$ und $a \in \mathbb{Z}$. □

Bemerkung. Dieser Beweis ist eine Verallgemeinerung des Beweises, dass $\sqrt{2}$ keine rationale Zahl ist, denn $\sqrt{2}$ ist eine Nullstelle des Polynoms $X^2 - 2 \in \mathbb{Q}[X]$. Sehr oft ist es einfach, ganzzahlige Nullstellen für ein Polynom mit ganzzahligen Koeffizienten auszuschließen (im Fall $f(X) \coloneqq X^2 - 2$ wird z.B. gezeigt, $f(a) > 0$, $\forall a \in \mathbb{Z} \setminus \{0, 1, -1\}$, und dann bleibt es lediglich, die Werte von f in diesen 3 (im allgemeinen Fall: endlich vielen) Zahlen $0, 1$ und -1 auszurechnen, um zu zeigen, $f(a) \neq 0$, $\forall a \in \mathbb{Z}$); durch Anwendung dieses obigen Satzes kann man also zeigen, ein solches Polynom (falls normiert!) hat keine rationale Nullstellen.

Definition 3.10. $f, g \in K[X] \setminus \{0\}$ heißen teilerfremd, falls $\forall h \in K[X] \setminus \{0\}$ gilt: $h \mid f \ \wedge \ h \mid g \Rightarrow \text{grad}(h) = 0$.

Satz 3.7 (Bézout für Polynome)**.** *Seien $f, g \in K[X] \setminus \{0\}$ teilerfremd. Dann $\exists p, q \in K[X]$: $f \cdot p + g \cdot q = 1$.*

Beweis. Wir können $n_1 \coloneqq \text{grad}(f) \geq \text{grad}(g) =: n_2$ annehmen, sonst können wir die Rollen von f und g vertauschen. Der Beweis ist per Induktion über n_2 (also über den niedriegeren Grad).

IA: Sei $n_2 = 0$. Dann $g(X) = a \in K \setminus \{0\}$. Sei $p = -1 \in K$ und $q = a^{-1} \cdot (f + 1)$. Dann gilt

$$f \cdot (-1) + a \cdot a^{-1} \cdot (f + 1) = 1$$

IS: Sei $n_2 > 0$. Sei $f = q' \cdot g + r$ nach Polynomdivision, $\text{grad}(r) < \text{grad}(g)$. $r \neq 0$, sonst $g \mid f$ und $\text{grad}(g) > 0$. Also $r \in K[X] \setminus \{0\}$. Sei $f' \coloneqq g, g' \coloneqq r$, also $\text{grad}(f') = n_1' = n_2$ und $\text{grad}(g') = \text{grad}(r) =: n_2' < n_2 = n_1'$. Also folgt mit Induktionsvoraussetzung: $\exists p'', q'' \in K[X]$:

$$f' \cdot p'' + g' \cdot q'' = 1 \Leftrightarrow g \cdot p'' + (f - g \cdot q') \cdot q'' = 1 \Leftrightarrow f \cdot q'' + g \cdot (p'' - q' \cdot q'') = 1$$

□

3.4 Restklassenringe. Körpererweiterungen

Satz und Definition 3.8. *Sei $p \in K[X] \setminus \{0\}$ mit $\text{grad}(p) > 0$. Auf $K[X]$ führen wir die folgende Äquivalenzrelation ein:*

$$\forall f, g \in K[X]\colon f \sim_p g \coloneqq p \mid f - g$$

Sei $[f]_p$ die Äquivalenzklasse zu f und $K[X]/(p)$ die Menge der Äquivalenzklassen. Wir definieren für $[f]_p, [g]_p \in K[X]/(p)$:

$$[f]_p + [g]_p \coloneqq [f + g]_p$$
$$[f]_p \cdot [g]_p \coloneqq [f \cdot g]_p$$

Dann ist $(K[X]/(p), +, \cdot)$ ein kommutativer Ring mit Einselement $[1]_p$, der K enthält (als die Klassen der konstanten Polynomen).

Die Abbildung $\pi\colon K[X] \to K[X]/(p)$, $\pi(f) \coloneqq [f]_p$ ist ein Ringhomomorphismus. Weiter ist $(K[X]/(p), +, \cdot)$ nullteilerfrei $\Leftrightarrow$ p ist irreduzibel $\Leftrightarrow$ $(K[X]/(p), +, \cdot)$ ist ein Körper.

Beweis. Wir zeigen nur den letzten Teil, also die 3 Äquivalenzen (für den ersten Teil wird der Beweis für die entsprechende Aussagen über Restklassen in $\mathbb{Z}$ fast ohne Änderung übertragen). Genauer gesagt, beweisen wir (a)⇒(b)⇒(c)⇒(a), wobei die letzte Implikation offensichtlich ist (ein Körper ist nullteilerfrei).

"(a)⇒(b)": Sei $p = q \cdot r$ mit $\text{grad}(q), \text{grad}(r) > 0$. Dann $[q]_p, [r]_p \neq [0]_p$ weil q, r nicht durch p teilbar sein können. Aber $[q]_p \cdot [r]_p = [q \cdot r]_p = [p]_p = [0]_p$. Also $K[X]/(p)$ nullteilerfrei $\Rightarrow$ p irreduzibel.

"(b)⇒(c)": Falls p irreduzibel ist und $q \in K[X]$ mit $p \nmid q$, so sind p, q teilerfremd, also $\exists f, g \in K[X]$ so dass $p \cdot f + q \cdot g = 1$. Also ist dann $[g]_p$ das inverse zu $[q]_p$ in $K[X]/(p)$. Also hat jedes Element in $K[X]/(p) \setminus \left\{[0]_p\right\}$ ein Inverses Element und damit ist $(K[X]/(p), +, \cdot)$ ein Körper.

"(c)⇒(a)" ist offensichtlich. □

Bemerkung. Die Abbildung $K \ni a \mapsto \pi(a) \in K[X]/(p)$ wobei $K \subseteq K[X]$ ist ein Ringhomomorphismus und da K Körper ist also injektiv. Das heißt $(K[X]/(p), +, \cdot)$ ist "größer" als K.

Satz 3.9. *Falls* $\text{grad}(p) = n \in \mathbb{N} \setminus \{0\}$*, so gibt es eine bijektive Abbildung* $\psi \colon K^n \xrightarrow{\sim} K[X]/(p)$.

Beweis. Betrachte $\psi(a_0, ..., a_{n-1}) \coloneqq [f]_p$ wobei $f(X) \coloneqq a_0 + a_1 X + ... + a_{n-1} X^{n-1}$. Die Abbildung ψ ist surjektiv (zeigt man durch Polynomdivision) und injektiv. □

Definition 3.11 (Charakteristik)**.** Sei $(K, +, \cdot)$ ein Körper. Die kleinste natürliche Zahl p so dass $\underbrace{1 + ... + 1}_{p-\text{mal}} = 0 \in K$ heißt die Charakteristik $\text{Char}(K)$ von K. Existiert kein solches p, so setzt man $\text{Char}(K) = 0$.

Satz 3.10. *Sei* $(K, +, \cdot)$ *ein Körper mit* $\text{Char}(K) \neq 0$*. Dann ist* $\text{Char}(K) = p$ *eine Primzahl.*

Beweis. Übung. □

Satz 3.11. ***a)*** *Sei* $(K, +, \cdot)$ *ein endlicher Körper. Dann* $\exists n \in \mathbb{N}$ *so dass* K *genau* p^n *Elemente besitzt.*

b) *Jeder endlicher Körper ist zu* $\mathbb{Z}_p[X]/(f)$ *isomorph für ein irreduzibles Polynom mit* $\text{grad}(f) = n$*. Für jedes* $n \in \mathbb{N} \setminus \{0\}$ *gibt es ein irreduzibles Polynom* $f \in \mathbb{Z}_p[X]$ *vom Grad* n*; falls* $f, g \in \mathbb{Z}_p[X]$ *irreduzibel und* $\text{grad}(f) = \text{grad}(g)$*, so sind die Körper* $\mathbb{Z}_p[X]/(f)$ *und* $\mathbb{Z}_p[X]/(g)$ *isomorph.*

c) *Für einen endlichen Körper ist* $(K \setminus \{0\}, \cdot)$ *eine zyklische abelsche Gruppe.*

Beweis. (a) folgt aus (b) und der bijektiven Abbildung von oben. (b) und (c) sind schwierig und werden hier deswegen nicht bewiesen. □

Bemerkung. Aus (a) und (b) folgt, dass alle endliche Körper, die die gleiche Anzahl von Elementen haben, zueinander isomorph sind. Diese Anzahl ist immer eine Potenz einer Primzahl, und jede solche Potenz entspricht eines endliches Körpers.

Folgerung: Die endlichen Körper haben wichtige Anwendungen in Verschlüsselung und Fehlerkorrekturen von Codes. Insbesondere $\mathbb{Z}_2[X]/(f)$ mit $\text{grad}(f) = 8$ und f irreduzibel. Die Eigenschaft (c) aus dem obigen Satz ist dabei besonders wichtig.

Die Restklassenkörper (für irreduzible Polynome in $K[X]$) sind also Erweiterungen des Körpers K. Solche Erweiterungen "produzieren Nullstellen von f" (welches eigentlich keine Nullstellen in K hat, denn es ist irreduzibel):

Bemerkung. Falls $K \subseteq K'$ eine Körpererweierung ist, ist $K[X]$ identifiziert mit der Teilmenge der Polynomen in $K'[X]$, deren Koeffizienten in $K \subseteq K'$ sind. Wir schriben $K[X] \subseteq K'[X]$.

Beispiel 3.8. $f(X) := X^2 - 2$ kann sowohl als Polynom in $\mathbb{Q}[X]$ (dort ist f irreduzibel) gesehen werden, als in $\mathbb{R}[X]$ (dort ist f aber reduzibel, denn es hat $\sqrt{2}, -\sqrt{2} \in \mathbb{R} \setminus \mathbb{Q}$ als Nullstellen).

Satz 3.12. *Sei K ein Körper und $f \in K[X]$, $\text{grad}(f) > 1$ irreduzibel, und sei $K' := K[X]/(f)$. Dann ist $f \in K'[X]$* reduzibel*, er hat in K' die Nullstelle $[X]_f \in K[X]/(f)$.*

Beweis. Sei $f(X) := \sum_{k=0}^{n} f_k X^k \in K[X] \subset K'[X]$. Dann gilt

$$\overline{f}([X]_f) = \sum_{k=0}^{n} f_k \cdot ([X]_f)^k = \sum_{k=0}^{n} \left[f_k \cdot X^k \right]_f = [f(X)]_f = [0]_f = 0 \in K'.$$

Somit ist $[X]_f \in K'$ eine Nullstelle von f in K'. □

Bemerkung. 1. Die Körpererweiterungen der Form $K \subset K[X]/(f)$, mit $f \in K[X]$ irreduzibel, "ergänzen" K zu einem "größerem" Körper K', der Nullstellen für f enthält.

2. Wenn wir sagen, dass ein Polynom f keine Nullstelle hat, müssen wir stets präzisieren, *in welchem Körper* wir die Nullstellen von f suchen, oder mindestens deutlich machen, dass die Koeffizienten von f in einem bestimmten Körper sind (Die Aussage "$f \in K[X]$ ist nullstellenfrei" wird dann der deutlichere Aussage "$f \in K[X]$ hat keine Nullstelle in K" gleichgesetzt).

Beispiel 3.9. Beispiele von Körpererweiterungen $K \subseteq K[X]/(p)$.

1) $\text{grad}(p) = 1 \Leftrightarrow K \cong K[X]/(p)$

2) $\mathbb{Q}[X]/(X^2-2)$ ist ein Körper, der Nullstellen von X^2-2 enthält (dies Polynom ist in $\mathbb{Q}[X]$ nullstellenfrei). Man kann $\mathbb{Q}$ mit der Menge $\mathbb{Q}[\sqrt{2}] := \{a+b\sqrt{2} : a,b \in \mathbb{Q}\} \subset \mathbb{R}$ identifizieren, wobei die Addition und Multiplikation die von $\mathbb{R}$ sind (evtl. Übung).

3) $\mathbb{C} := \mathbb{R}[X]/(X^2+1)$ ist der Körper der komplexen Zahlen. Wir notieren $i := [X]_{(X^2+1)}$, also wir schreiben $[bX+a] =: (a+bi)$ mit $a,b \in \mathbb{R}$. Es ist $i^2 = 1$, da $i \in \mathbb{C}$ eine Nullstelle des Polynoms X^2+1 ist (die andere Nullstelle ist $-i \in \mathbb{C}$).

Definition 3.12. Eine Nullstelle von $f \in K[X]$ hat Vielfachheit $k \in \mathbb{N} :\Leftrightarrow (X-a)^k \mid$ und $(X-a)^{k+1} \nmid f$.

Beispiel 3.10. $-1 \in \mathbb{Q}$ ist eine zweifache (doppelte) Nullstelle von $X^2+2X+1 \in \mathbb{Q}[X]$.

Satz 3.13. *Ein Polynom $f \in K[X]$ mit $\operatorname{grad}(f) = n > 0$ hat höchstens n Nullstellen in K, gerechnet mit Vielfachheiten.*
Sind $a_1, .., a_l$ Nullstellen mit Vielfachheiten $n_1, .., n_l$ so gilt $n_1 + ... + n_l \leq n$

Beweis. Übung. □

Korollar 3.14. *Der Einsetzhomomorphismus $F\colon K[X] \to \operatorname{Abb}(K,K)$, $F(f) := \overline{f} : K \to K$ ist injektiv $\Leftrightarrow$ K ist unendlich.*

Beweis. F ist injektiv $\Leftrightarrow (\forall f,g \in K[X]\colon \overline{f} = \overline{g} \Rightarrow f = g) \Leftrightarrow (\forall a \in K\colon \overline{f-g}(a) = 0 \Rightarrow f-g = 0)$. Sei also $f \in K[X]$ so dass $F(f) = 0 \in \operatorname{Abb}(K,K)$. $0 = \overline{f}$ ist also die Polynomialfunktion, die zu f assoziiert ist.

Falls $\operatorname{grad}(f) = n \in \mathbb{N}$, so hat f höchstens n verschiedenen Nullstellen. Da aber $\overline{f}(a) = 0\ \forall a \in K$ und K unendlich ist, folgt $f = 0$ in $K[X]$.

Falls K endlich ist, $|K| = p^n$ mit $n \in \mathbb{N}$ und p Primzahl, so die Polynomialfunktion zu $X^{p^n} - X$ verschwindet auf ganz K (d.h. jedes Element in K ist Nullstelle für das obige Polynom). Für $a = 0$ ist dies klar.

Falls $a \in K \setminus \{0\}$, so folgt aus dem Satz von Fermat für Gruppen, dass $a^{|K\setminus\{0\}|} = 1$, also $a^{p^n-1} = 1$ für alle $a \in K \setminus \{0\}$. Also $X^{p^n} - X \neq 0$ hat als assoziierte Funktion die Nullfunktion. Damit ist F nicht injektiv. □

Definition 3.13. Ein Körper $(K,+,\cdot)$ heißt algebraisch abgeschlossen, falls $\forall f \in K[X]$ gilt, dass f eine Nullstelle in K besitzt.

Bemerkung. Falls K algebraisch abgeschlossen ist und $\operatorname{grad}(f) = n$, so besitzt f genau $\operatorname{grad}(f)$ Nullstellen mit Vielfachheiten. Dies bedeutet genau, dass alle irreduziblen Polynome Grad 1 besitzen. (Übung)

4 Komplexe Zahlen

Es ist $\mathbb{C} := \mathbb{R}[X]/(X^2+1)$. Die Restklassen modulo (X^2+1) kann man durch den Rest der Polynomdivision charakterisieren: Als Rest sind die Polynome $a+bX$ möglich mit $a, b \in \mathbb{R}$. Also

$$\mathbb{C} = \left\{ [a+bX]_{(X^2+1)} \mid a, b \in \mathbb{R} \right\}$$

Sei $i := [X]_{(X^2+1)}$. Dann gilt $i^2 + [1] \equiv 0 \pmod{X^2+1}$, also $i^2 + 1 = 0$ in $\mathbb{C}$ (wobei hier $1, 0$ die Notationen für Null- bzw. Einselement in $\mathbb{C}$ seien. Wir identifizieren $\mathbb{R}$ mit den Klassen $[a]_{(X^2+1)}$ für $a \in \mathbb{R}$, d.h. mit den Restklassen der konstanten Polynome. Wir haben dann $\mathbb{R} \subseteq \mathbb{C}$ und $\mathbb{C} = \left\{ a + ib \mid a, b \in \mathbb{R} \right\}$ mit $i^2 = -1$ (die Gleichung $i = \sqrt{-1}$ ist zu vermeiden, denn es gilt ebenfalls $(-i)^2 = -1$).

Komplexkonjugation Sei $z = a + ib \in \mathbb{C}$, $a, b \in \mathbb{R}$. Dann ist $\overline{z} := a - ib$ das komplex konjugierte zu z und die Abbildung

$$C \colon \mathbb{C} \to \mathbb{C},\ z \mapsto \overline{z}$$

heißt komplexe Konjugation. Weiter schreibt man für $z = a + ib \in \mathbb{C}$:

$$\begin{aligned} a &:= \mathrm{Re}(z) \text{ und nennt ihn den Realteil von } z \\ b &:= \mathrm{Im}(z) \text{ und nennt ihn den Imaginärteil von } z \\ |z| &:= \sqrt{a^2+b^2} \in \mathbb{R}_+ \text{ und nennt ihn den Absolutbetrag von } z \end{aligned}$$

Es gilt weiter $(a+ib)(c+id) = (ac-bd) + i(bc+ad)$

Satz 4.1 (Rechenregeln in $\mathbb{C}$)**.** *Für alle $z \in \mathbb{C}$ gilt:*

a) $z + \overline{z} = 2Re(z)$ *und* $z - \overline{z} = 2i\,\mathrm{Im}(z)$*, sowie* $z \cdot \overline{z} = |z|^2$ *und* $\overline{\overline{z}} = z$*. Weiter gilt* $\overline{z} = z \Leftrightarrow z \in \mathbb{R}$*.*

(b) $\forall z, w \in \mathbb{C}$:

$$\left.\begin{aligned} \overline{zw} &= \overline{z} \cdot \overline{w} \\ \overline{z+w} &= \overline{z} + \overline{w} \end{aligned}\right\} \; C \colon \mathbb{C} \to \mathbb{C} \textit{ ist ein Körperisomorphismus}$$

Damit folgt insbesondere:

$$|zw| = |z||w|$$

c) *Es gilt*

$$|z+w| \leq |z| + |w|$$

Beweis. **a,b)** Übung.

c) Sei $z = a + ib, w = c + id$. Dann $z + w = (a + c) + i(b + d)$. Es gilt

$$|z + w|^2 = (a + c)^2 + (b + d)^2 = a^2 + b^2 + c^2 + d^2 + 2ac + 2bd$$

Und

$$(|z| + |w|)^2 = |z|^2 + 2|zw| + |w|^2 = a^2 + b^2 + c^2 + d^2 + 2(\sqrt{a^2 + b^2}\sqrt{c^2 + d^2})$$

Also

$$|z + w| \leq |z| + |w| \Leftrightarrow |z + w|^2 \leq (|z| + |w|)^2 \Leftrightarrow ac + bd \overset{(CS)}{\leq} \sqrt{a^2 + b^2}\sqrt{c^2 + d^2}$$

für $a, b, c, d \in \mathbb{R}$. Diese Ungleichung folgt aus dem folgenden Spezialfall der Ungleichung von Cauchy-Schwarz:

$$\forall a, b, c, d \in \mathbb{R},\ (ac + bd)^2 \leq (a^2 + b^2)(c^2 + d^2). \tag{CS}$$

Beweis. Seien $a, b \neq 0$. Seien $f_1(x) := a^2x^2 - 2acx + c^2 = (ax - c)^2 \geq 0\ \forall x \in \mathbb{R}$ und $f_2(x) := b^2x^2 - 2bdx + d^2 = (bx - d)^2 \geq 0\ \forall x \in \mathbb{R}$. Setze nun $f(x) := f_1(x) + f_2(x)$. Dann also $f(x) \geq 0\ \forall x \in \mathbb{R}$. Entweder haben f_1 und f_2 eine gemeinsame Nullstelle $\frac{c}{a} = \frac{d}{b}$, also $b + id = q(a + ic)$ mit $q \in \mathbb{R}$, oder aber $f(x)$ hat keine reelle Nullstelle. Das bedeutet dann, dass $f(x) = \alpha x^2 + \beta x + \gamma$ mit $\alpha, \beta, \gamma \in \mathbb{R}$ und

$$\frac{\beta^2}{4} < \alpha\gamma \tag{$**$}$$

In diesem Fall ist $\alpha = a^2 + b^2, \gamma = c^2 + d^2$ und $\beta = -2ac - 2bd$ und damit folgt (CS) aus $(**)$. Falls $a, b = 0$, so ist die Ungleichung in (CS) trivial. Falls $a = 0, b \neq 0$, oder $a \neq 0, b = 0$ so ist f eine quadratische Polynomfunktion. Falls $ac + bd < 0$, so folgt (CS) automatisch. Sonst gilt:

$$\begin{aligned}
&(\text{CS}) \Leftrightarrow (ac + bd)^2 \leq (a^2 + b^2)(c^2 + d^2) \\
&\Leftrightarrow a^2c^2 + b^2d^2 + 2abcd \leq a^2c^2 + a^2d^2 + b^2c^2 + b^2d^2 \\
&\Leftrightarrow 2abcd \leq a^2d^2 + b^2c^2 \Leftrightarrow (ad - bc)^2 \geq 0
\end{aligned}$$

also die Behauptung. Gleichheit gilt, falls $ad - bc = 0$, also auch, falls $a + ib \neq 0$ und $c + id = q(a + ib)$ für ein $q \in \mathbb{R}$. □

□

Bemerkung. Die Ungleichung aus c) zeigt, dass $|\cdot| : \mathbb{C} \to \mathbb{R}, z \mapsto |z|$ kein Körperhomomorphismus ist. Dies ist natürlich nicht überraschend, da $|\cdot|$ nicht injektiv ist. Es ist aber $|\cdot| : \mathbb{C} \setminus \{0\} \to \mathbb{R}_+^*$ ein Gruppenhomomorphismus der multiplikativen Gruppen.

4.1 Polarkoordinaten

Sei $z = a + ib \in \mathbb{C} \setminus \{0\}$ und $r := |z| \in \mathbb{R}_+^*$, da $z \neq 0$. Dann $\exists! \phi \in [0, 2\pi)$ mit $z = r(\cos\phi + i\sin\phi)$. Man nennt r den Betrag von z und ϕ das Argument von z und schreibt dafür auch $\arg(z)$.

Bemerkung. $\{\phi \in \mathbb{R} \mid r(\cos\phi + i\sin\phi) = z\} = \{\arg(z) + 2k\pi \mid k \in \mathbb{Z}\}$, da $\sin, \cos$ periodisch sind.

Notation 4.1. $e^{i\phi} := \exp(i\phi) := \cos\phi + i\sin\phi$. Allgemein wird exp definiert durch:

$$\exp\colon \mathbb{C} \to \mathbb{C} \setminus \{0\}$$
$$\exp(a + ib) = e^a(\cos b + i\sin b) =: e^{a+ib} \quad \forall a, b \in \mathbb{R}$$

Diese Definition wird durch die Reihendarstellung der Exponentialfunktion gerechtfertigt.

Satz 4.2. *Seien $z, w \in \mathbb{C} \setminus \{0\}$, $z = |z|(\cos\phi + i\sin\phi), w = |w|(\cos\psi + i\sin\psi)$ mit $\phi, \psi \in \mathbb{R}$. Dann gilt*

a) $zw = |zw|(\cos(\phi + \psi) + i\sin(\phi + \psi))$

b) $z^{-1} = \frac{1}{z} = \frac{1}{|z|}(\cos\phi - i\sin\phi) = \frac{1}{|z|}(\cos\phi + i\sin(-\phi))$.

c) *Falls $|z| = 1$, so $\overline{z} = z^{-1}$.*

d) $z^n = |z|^n(\cos(n\phi) + i\sin(n\phi))$.

Beweis. Übung. Man verwende bei a) $\cos(\phi + \psi) = \cos\phi\cos\psi - \sin\phi\sin\psi$ und $\sin(\phi + \psi) = \sin\phi\cos\psi + \sin\psi\cos\phi$. □

Beispiel 4.1 (Bestimmen von Nullstellen in $\mathbb{C}$)**. 1)** Sei $f(X) := X^n - a, a \in \mathbb{C}$. $z \in \mathbb{C}$ ist eine Nullstelle von f, falls

$$z^n = a \Leftrightarrow |z|^n(\cos(n\phi) + i\sin(n\phi)) = |a|(\cos\psi + i\sin\psi)$$

für $\psi = \arg(a)$ und $|a| \neq 0$. Also $z = |z|(\cos\phi + i\sin\phi)$ ist eine Nullstelle von f

$$\Leftrightarrow \begin{cases} |z| = |a|^{\frac{1}{n}} \in \mathbb{R}_+^* \\ \begin{cases} \cos(n\phi) = \cos(\psi) \\ \sin(n\phi) = \sin(\psi) \end{cases} \Leftrightarrow n\phi = \psi + 2k\pi, k \in \mathbb{Z} \end{cases}$$

$\Leftrightarrow \phi = \frac{\psi}{n} + \frac{2k\pi}{n}$ für $k \in \mathbb{Z}$. Es besitzt f also n *verschiedene* Nullstellen

$$\left\{|a|^{\frac{1}{n}}(\cos(\frac{\psi}{n} + \frac{2k\pi}{n}) + i\sin(\frac{\psi}{n} + \frac{2k\pi}{n})) \mid k \in \{0, ..., n-1\}\right\}$$

Alle anderen Werte von k liefern Elemente aus dieser Menge.

2) Seien $a, b \in \mathbb{C}$ und $a \neq 0$. Betrachte $f(X) := aX^2 + bX + c$. Wir möchten Nullstellen von f bestimmen. Wir verfahren wie bei reellen Polynome vom Grad 2. Setze also $\Delta := b^2 - 4ac$. Sei δ eine der 2 Lösungen der Gleichung $\delta^2 = \Delta \neq 0 \in \mathbb{C}$. Falls $\Delta = 0$, so hat die Gleichung offenbar die eindeutige Lösung $\delta = 0$. Also gilt für $\Delta = 0$:

$$f(X) = a(X + \frac{b}{2a})^2$$

mit Nullstelle $-\frac{b}{2a}$ und im Fall $\Delta \neq 0$ erhalten wir:

$$f(X) = a(X + \frac{b+\delta}{2a})(X + \frac{b-\delta}{2a})$$

mit Nullstellen $\frac{b+\delta}{2a}, \frac{b-\delta}{2a} \in \mathbb{C}$, wie man durch direkte Rechnung verifiziert.

Satz 4.3. *Sei $f \in \mathbb{R}[X]$ mit* $\text{grad}(f) \geq 1$. *Dann gilt f irreduzibel* $\Rightarrow \text{grad}(f) = 1$ *oder* $\text{grad}(f) = 2$.

Beweis. Sei $f \in \mathbb{R}[X]$ mit $\text{grad}(f) = n > 1$. Dann hat f in $\mathbb{C}$ genau n Nullstellen (mit Vielfachheiten) Sei also

$$f(X) = a(X - a_1)...(X - a_n), \quad a_1, ..., a_n \in \mathbb{C}, a \in \mathbb{R} \setminus \{0\}$$

Sei $f(X) = \sum_{k=0}^{n} f_k X^k$. Dann

$$f(\overline{z}) = \sum_{k=0}^{n} f_k (\overline{z})^k = \sum_{k=0}^{n} f_k \overline{z^k} = \sum_{k=0}^{n} \overline{f_k z^k} = \overline{f(z)}, \quad \forall z \in \mathbb{C}$$

weil die Koeffizienten reell sind. Ist nun f irreduzibel in $\mathbb{R}[X]$, so folgt aus $\text{grad}(f) > 1$, dass f keine reelle Nullstelle besitzen kann. Also sind alle Nullstellen von f aus $\mathbb{C} \setminus \mathbb{R}$ und können als $b_1, ..., b_l, \overline{b_1}, ..., \overline{b_l} \in \mathbb{C} \setminus \mathbb{R}$ umbenannt werden. Inbesondere ist also $n = 2l$. Es gilt also

$$f(X) = a(X - b_1)(X - \overline{b_1})...(X - b_l)(X - \overline{b_l})$$

Aber es gilt $g_1(X) := (X - b_1)(X - \overline{b_1}) = X^2 - (b_1 + \overline{b_1})X + b_1\overline{b_1}$ hat reelle Koeffizienten, da $b_1 + \overline{b_1} = 2\,\text{Re}(b_1) \in \mathbb{R}$ und $b_1\overline{b_1} = |b_1| \in \mathbb{R}$. Also $g_1 \in \mathbb{R}[X]$ und $g_1 \mid f$. Falls $\text{grad}(f) > 2$, so ist f also reduzibel. □

Bemerkung. Ob ein reelles Polynom vom Grad 2 irreduzibel ist, lässt sich aus des bekannten Theorie der quadratischen Funktionen herleiten. Sei $f(X) = aX^2 + bX + c$ für $a, b, c \in \mathbb{R}, a \neq 0$. Sei $\Delta := b^2 - 4ac$. Dann

- f hat 2 einfache reelle Nullstellen $\Leftrightarrow \Delta > 0$
- f hat eine doppelte Nullstelle $\Leftrightarrow \Delta = 0$
- f hat keine *reelle* Nullstelle $\Leftrightarrow \Delta < 0$. In diesem Fall ist $f \in \mathbb{R}[X]$ irreduzibel, natürlich aber nicht in $\mathbb{C}[X]$, da es dort Nullstellen besitzt.

5 Vektorräume

Definition 5.1. Sei K ein Körper. Eine abelsche Gruppe $(V, +)$ heißt K-Vektorraum, falls eine Abbildung

$$\cdot \colon K \times V \to V$$

existiert, welche die folgenden Eigenschaften erfüllt

(V1) $\forall a \in K \; \forall v_1, v_2 \in V \colon a \cdot (v_1 + v_2) = a \cdot v_1 + a \cdot v_2$

(V2) $\forall a_1, a_2 \in K \; \forall v \in V \colon (a_1 + a_2) \cdot v = a_1 \cdot v + a_2 \cdot v$

(V3) $\forall a_1, a_2 \in K \; \forall v \in V \colon (a_1 \cdot a_2) \cdot v = a_1 \cdot (a_2 \cdot v)$

(V4) $\forall v \in V \colon 1_K \cdot v = v$

Die Elemente aus V heißen Vektoren, die Elemente von K bezeichnet man als Skalare.

Beispiel 5.1. 1) Falls $V = K$ und $\cdot \colon K \times K \to K$ die übliche Multiplikation in K ist , so wird K zu einem K-Vektorraum.

2) Sei $K \xrightarrow{\alpha} R$ ein Ringhomomorphismus. Dann ist die skalare Multiplikation von K auf R durch $\forall a \in K, \; \forall v \in R \colon a \cdot v \coloneqq \alpha(a) \cdot v$ definiert, wobei die zweite Multiplikation in R zu verstehen ist. Damit wird R zu einem K-Vektorraum. Insbesondere sind damit $\mathbb{R}, \mathbb{C}$ $\mathbb{Q}$-Vektorräume und $\mathbb{C}$ ein $\mathbb{R}$-Vektorraum. Weiter ist für alle Körper K dann $K[X]/(p)$ ein K-Vektorraum.

3) Sei K ein Körper und M eine Menge sowie V ein K-Vektorraum. Sei $V^M \coloneqq \mathrm{Abb}(M, V)$. Wir definieren $+ \colon V^M \times V^M \to V^M$ durch

$$\forall f, g \in V^M \colon f + g \colon M \to V \colon (f + g)(a) \coloneqq f(a) + g(a) \in V, \quad \forall a \in M$$

Dann ist $(V^M, +)$ eine abelsche Gruppe. Wir definieren nun $\cdot \colon K \times V^M \to V^M$ durch

$$\forall \alpha \in K \; \forall f \in V^M \colon \alpha \cdot f \colon M \to V$$
$$\forall a \in M \colon (\alpha \cdot f)(a) \coloneqq \alpha \cdot f(a) \in V$$

Dann ist V^M ein K-Vektorraum.

Beweis. **(V1)** $\forall \alpha \in K \; \forall f, g \in V^M \colon$

$$\begin{aligned} \forall a \in M \colon (\alpha \cdot (f_1 + f_2))(a) &\overset{\text{(V1) für } V}{=} \alpha \cdot f_1(a) + \alpha \cdot f_2(a) \\ &= (\alpha \cdot f_1)(a) + (\alpha \cdot f_2)(a) = (\alpha \cdot f_1 + \alpha \cdot f_2)(a) \end{aligned}$$

(V2) $\forall \alpha_1, \alpha_2 \in K \; \forall f \in V^M$:

$$\forall a \in M\colon ((\alpha_1 + \alpha_2) \cdot f)(a) \overset{\text{Def.}}{=} (\alpha_1 + \alpha_2) \cdot f(a) \overset{\text{(V2) für } V}{=} \alpha_1 \cdot f(a) + \alpha_2 f(a)$$
$$\overset{\text{Def.}}{=} (\alpha_1 \cdot f)(a) + (\alpha_2 \cdot f)(a) \overset{\text{Def.}}{=} (\alpha_1 \cdot f + \alpha_2 \cdot f)(a)$$

(V3) $\forall \alpha_1, \alpha_2 \in K \; \forall f \in V^M$:

$$\forall a \in M\colon ((\alpha_1 \cdot \alpha_2) \cdot f)(a) \overset{\text{Def.}}{=} (\alpha_1 \cdot \alpha_2) \cdot f(a) \overset{\text{(V3) für } V}{=} \alpha_1 \cdot (\alpha_2 \cdot f(a))$$
$$\overset{\text{Def.}}{=} \alpha_1 \cdot (\alpha_2 \cdot f)(a) \overset{\text{Def.}}{=} (\alpha_1 \cdot (\alpha_2 \cdot f))(a)$$

(V4) $\forall f \in V^M$

$$\forall a \in M\colon (1 \cdot f)(a) \overset{\text{Def.}}{=} 1 \cdot f(a) \overset{\text{(V4) für } V}{=} f(a)$$

□

Spezialfälle: 1) $K = V = \mathbb{R}$ oder $\mathbb{C}$ und $\mathrm{Abb}(M, \mathbb{R})$ bzw $\mathrm{Abb}(M, \mathbb{C})$ (Mengen der reell- beziehungsweise komplexwertigen Funktionen) sind $\mathbb{R}$- beziehungsweise $\mathbb{C}$-Vektorräume. Für alle Mengen M (oft hat man $M \subseteq \mathbb{R}$ oder $M \subseteq \mathbb{C}$)

1a) Für $K = V = \mathbb{C}$ und $M = \mathbb{N}$ erhält man $\mathrm{Abb}(\mathbb{N}, \mathbb{C})$ als Menge der Folgen aus $\mathbb{C}$. Statt $a\colon \mathbb{N} \to \mathbb{C}$ schreibt man stattdessen $(a_0, a_1, ...) = (a_k)_{k \in \mathbb{N}}$.

1b) Ist $M = \{1, ..., n\}$ mit $n \in \mathbb{N} \setminus \{0\}$ und $V = K$. Dann ist $V^M = K^{\{1,..,n\}} =: K^n$ die Menge der n-Tupel aus K. Wir schreiben statt $a\colon \{1, .., n\} \to K$ zumeist $(a_1, .., a_n) \in K^n$. Man kann K^n also als n-faches kartesisches Produkt $K \times ... \times K$ verstehen. Also ist K^n ein K-Vektoraum. Falls $n = 1$, so $K^1 = K$ und falls $n = 0$, so $K^0 \coloneqq \{0\}$ der triviale Vektorraum.

1c) Sei $M \coloneqq \{1, .., n\} \times \{1, .., m\}$ mit $m, n \in \mathbb{N} \setminus \{0\}$ und $K = V$. Dann ist $V^M = K^{\{1,..,n\} \times \{1,..,m\}} =: \mathrm{Mat}(n \times m, K)$ der Vektorraum der $n \times m$-Matrizen über K. Statt $a\colon \{1, .., n\} \times \{1, .., m\} \to K$ schreiben wir stattdessen die Werte $a_{ij} \coloneqq a(i, j)$ in eine Tabellen

$$\begin{pmatrix} a_{11} & ... & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & ... & a_{nm} \end{pmatrix}$$

mit n Zeilen und m Spalten. Diese Tabelle heißt Matrix mit n Zeilen und m Spalten. Eine alternative, kompakte Schreibweise hierfür ist $(a_{ij})_{\substack{i \in \{1,..,n\} \\ j \in \{1,..,m\}}}$.

Bemerkung. Es ist $(a_{ij})_{\substack{i\in\{1,..,n\}\\ j\in\{1,..,m\}}} = (a_{kl})_{\substack{k\in\{1,..,n\}\\ l\in\{1,..,m\}}} = (a_{ji})_{\substack{j\in\{1,..,n\}\\ i\in\{1,..,m\}}} = (a_{ij})_{\substack{j\in\{1,..,m\}\\ i\in\{1,..,n\}}}$ und so weiter, aber es es ist $(a_{ij})_{\substack{i\in\{1,..,n\}\\ j\in\{1,..,m\}}} \neq (a_{ij})_{\substack{i\in\{1,..,m\}\\ j\in\{1,..,n\}}}$ für $m \neq n$.

Merkhilfe: Man denke am bestem immer an die tabellarische Darstellung von Matrizen. Dann ist der Zeilenindex der zweite Index und der Spaltenindex der erste.
Zu 1b): Ist $\lambda \in K, a = (a_1, .., a_n) \in K^n$, (oft schreibt man a als Spaltenvektor $\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$). Dann ist $\lambda \cdot a = (\lambda a_1, .., \lambda a_n)$ oder auch $\begin{pmatrix} \lambda a_1 \\ \lambda a_2 \\ \vdots \\ \lambda a_n \end{pmatrix}$ und $a + b = (a_1 + b_1, ..., a_n + b_n)$ oder auch $a + b = \begin{pmatrix} a_1 + b_1 \\ \vdots \\ a_n + b_n \end{pmatrix}$.

Satz 5.1 (Rechenregeln in einem K-Vektorraum)**.** *Sei K Körper und V ein K-Vektorraum. Dann gelten:*

(a) $0_K \cdot v = 0_V \ \forall v \in V \ \wedge \ a \cdot 0_V = 0_V \ \forall a \in K$

(b) $\alpha \cdot v = 0_V \Rightarrow \alpha = 0_K \ \vee \ v = 0_V$

(c) $(-1_K) \cdot v = -v \ \forall v \in V$

Beweis. **(a)** $0_K \cdot v = (0_K + 0_K) \cdot v = 0_K \cdot v + 0_K \cdot v$, also ist $0_K \cdot v$ ein neutrales Element für $(V, +)$, also gleich 0_V. Auch gilt $\alpha \cdot 0_V + \alpha \cdot 0_V = \alpha \cdot (0_V + 0_V) = \alpha \cdot 0_V$, also ist auch $\alpha \cdot 0_V$ ein neutrales Element in $(V, +)$, also $\alpha \cdot 0_V = 0_V$.

(b) Sei $\alpha \neq 0_K$. Dann $\alpha^{-1} \in K$ und $\alpha^{-1}(\alpha \cdot v) = \alpha^{-1} \cdot 0_V = 0_V$ aus (a), aber auch $\alpha^{-1} \cdot (\alpha \cdot v) = (\alpha^{-1} \cdot \alpha) \cdot v \overset{(V3)}{=} 1_K \cdot v \overset{(V4)}{=} v$. Also $v = 0_V$.

(c) Es ist $1_K - 1_K = 0_K$, also

$$v + (-1)_K \cdot v = 1_K \cdot v + (-1)_K \cdot v = (1_K + (-1)_K) \cdot v = 0_V$$

und damit ist $(-1)_K \cdot v$ das (eindeutig bestimmte) inverse Element zu v, also $(-1)_K \cdot v = -v$.

□

Definition 5.2 (Untervektorraum)**.** Sei K ein Körper und V ein K-Vektorraum. $U \subseteq V$ heißt Untervektorraum, falls gilt

(U1) $\emptyset \neq U$

(U2) $\forall \alpha \in K \, \forall v \in U \colon \alpha \cdot v \in U$

(U3) $\forall v_1, v_2 \in U \colon v_1 + v_2 \in U$

Bemerkung. 1. (U1) $\wedge$ (U3) (zusammen mit (U2) für $\alpha = -1_K$) sagen genau aus, dass $U \subseteq V$ eine Untergruppe bezüglich der Addition ist.

2. U wird mit der Addition, und der von V induzierten Skalarmultiplikation $\cdot \colon K \times U \to U$ wieder zu einem K-Vektorraum.

Beispiel 5.2. 1. $\mathbb{R} \subseteq \mathbb{C}$ ist ein $\mathbb{Q}$-oder $\mathbb{R}$-Untervektorraum, jedoch *kein* $\mathbb{C}$-Untervektorraum.

Notation 5.3. Sei V ein K-Vektorraum. Dann $V^{\emptyset} \coloneqq \{0\}$ und $V^0 \coloneqq \{0\}$ also 'nullte Potenz' von V sind triviale Gruppen (bzw. Vektorräume).

2. Sei $U_m \coloneqq \big\{(a_1, .., a_n) \in K^n \mid a_{m+1} = ... = a_n = 0\big\}$. Dann ist $U_m \subseteq K^n$ ein K-Untervektorraum.
Allgemeiner gilt: Sei V ein K-Vektorraum, M Menge und $N \subseteq M$ eine Teilmenge. Dann ist $U_N \coloneqq \big\{f \colon M \to V \mid f(x) = 0_V \, \forall x \in M \setminus N\big\}$ ein Untervektorraum.

Beweis. **(U1)** Gilt immer, denn U_N enthält offenbar die Nullfunktion $f_0 \colon M \to V$, $f_0(x) = 0_V \, \forall x \in M$.

(U2) Sei $\alpha \in K$ und $f \in U_N$. Dann ist $(\alpha \cdot f) \in V^M$ definiert durch $(\alpha \cdot f)(x) \coloneqq \alpha \cdot f(x) \, \forall x \in M$. Insbesondere $\alpha \cdot f(x) = \alpha \cdot 0_V = 0_V \, \forall x \in M \setminus N$, also $(U2)$ gilt für $U_N \subseteq V^M$.

(U3) Falls $f, g \in U_N$, so $(f + g)(x) = f(x) + g(x) = 0_V \, \forall x \in M \setminus N$, also (U3) für $U_N \subseteq V^M$.

□

Satz 5.2. *Eine nichtleere Teilmenge U in einem K-Vektorraum V ist ein Untervektorraum $\Leftrightarrow \forall k \in \mathbb{N} \setminus \{0\} \, \forall \alpha_1, .., \alpha_k \, \forall v_1, .., v_k \in U$ gilt $\alpha_1 v_1 + ... + \alpha_k v_k \in U$.*

Beweis. "$\Rightarrow$" Induktion nach $k \in \mathbb{N}$:

$k = 1$: $\forall \alpha_1 \in K \, \forall v_1 \in U \colon \alpha_1 v_1 \in U$ nach (U2).

$k \to k + 1$: $\forall \alpha_1, ..., \alpha_{k+1} \in K \, \forall v_1, ..., v_{k+1} \in U \colon$

$$(\alpha_1 \cdot v_1 + ... + \alpha_k v_k) + \alpha_{k+1} v_{k+1} \in U$$

nach Induktionsvoraussetzung und (U2), da beide Terme in U liegen und damit auch ihre Summe.

"$\Leftarrow$" (U1) folgt mit $k = 1$ und (U3) folgt aus $k = 2$ mit $\alpha_1 = \alpha_2 = 1_K$.

□

Satz 5.3. *Seien U_1, U_2 Untervektorräume in V. Dann sind $U_1 \cap U_2$ und*

$$U_1 + U_2 \coloneqq \{v_1 + v_2 \mid v_i \in U_1, v_2 \in U_2\}$$

Untervektorräume.

Beweis. Es gilt $0_V \in U_1 \cap U_2$ und $0_V = 0_V + 0_V \in U_1 + U_2$. Sei nun $\alpha \in K$. Dann ist $\alpha \cdot v \in U_1 \wedge \alpha \cdot v \in U_2$, also (U2) für $U_1 \cap U_2$. Sei nun $v = v_1 + v_2 \in U_1 + U_2$ und $\alpha \in K$. Dann $\alpha \cdot v = (\alpha \cdot v_1) + (\alpha \cdot v_2) \in U_1 + U_2$, also (U2) für $U_1 + U_2$. Seien nun $v, v' \in U_1 \cap U_2$. Dann ist $v + v' \in U_1$ und $v + v' \in U_2$, also (U3) gilt in $U_1 \cap U_2$. Sei weiter $v = v_1 + v_2, v' = v_1' + v_2' \in U_1 + U_2$. Dann gilt $v + v' = (v_1 + v_1') + (v_2 + v_2') \in U_1 + U_2$, also (U3) in $U_1 + U_2$. Also sind $U_1 \cap U_2$ und $U_1 + U_2$ Untervektorräume in V. □

Allgemeiner gilt:

a) Sei $(U_i)_{i \in I}$ eine Familie von K-Untervektorräumen in V wobei I eine beliebige Menge sei. Dann ist $\bigcap_{i \in I} U_i \subseteq V$ ein Untervektorraum.

b) Sei $k \in \mathbb{N} \setminus \{0\}$ und $U_1, .., U_k \subseteq V$ Untervektorräume. Dann ist $U_1 + ... + U_k \coloneqq \{v_1 + ... + v_k \mid v_i \in U_i\}$ ein Untervektorraum.

Bemerkung. 1) $U_\alpha \coloneqq \{(x, y) \in \mathbb{R}^2 \mid y = \alpha\}$ ist ein $\mathbb{R}$-Untervektorraum in $\mathbb{R}^2 \Leftrightarrow \alpha = 0$.

2) Sei $U_1 \coloneqq \{(x, 0) \mid x \in \mathbb{R}\} \subseteq \mathbb{R}^2$ und $U_2 \coloneqq \{(0, y) \mid y \in \mathbb{R}\} \subseteq \mathbb{R}^2$. Dann sind U_1, U_2 Untervektorräume, aber $U_1 \cup U_2$ ist kein Untervektorraum, denn $(1, 0) \in U_1, (0, 1) \in U_2$, aber $(1, 1) = (1, 0) + (0, 1) \notin U_1 \cup U_2$.

5.1 Lineare Abbildungen. Matrizen

Definition 5.4. Sei K Körper und seien V, W K-Vektorräume. Eine Abbildung $f \colon V \to W$ heißt K-linear, oder auch Vektorraumhomomorphismus, falls gilt

(L1) $\forall v_1, v_2 \in V \colon f(v_1 + v_2) = f(v_1) + f(v_2)$

(L2) $\forall \alpha \in K \, \forall v \in V \colon f(\alpha \cdot v) = \alpha \cdot f(v)$

Bemerkung. (L1) $\Leftrightarrow f \colon (V, +) \to (W, +)$ ist ein Gruppenhomomorphismus.

Satz 5.4. *Sei $f \colon V \to W$ eine lineare Abbildung zwischen K-Vektorräumen. Dann sind* $\ker(f) \coloneqq \{v \in V \mid f(v) = 0_W\} \subseteq V$ *und* $\mathrm{Im}(f) \coloneqq f(V) = \{w \in W \colon \exists \in V \colon F(v) = w\} \subseteq W$ *Untervektorräume.*

Beweis. Es ist bereits bekannt, dass $\ker(f), \mathrm{Im}(f)$ Untergruppen in V beziehungsweise W sind, weil f ein Gruppenhomomorphismus ist. Es ist also nur noch (U3) zu prüfen. Sei $v \in \ker(f), \alpha \in K$. Dann $f(\alpha \cdot v) = \alpha \cdot f(v) = 0_W$, also $\alpha \cdot v \in \ker(f)$. Ist nun $w = f(v) \in \mathrm{Im}(f)$ und $\alpha \in K$, so ist $f(\alpha \cdot v) = \alpha \cdot w$, also $\alpha \cdot w \in \mathrm{Im}(f)$. □

Beispiel 5.3. $f\colon \mathbb{R}^2 \to \mathbb{R}^2,\ f(x,y) \coloneqq (ax+by, cx+dy)$ mit $a,b,c,d \in \mathbb{R}$ ist eine $\mathbb{R}$-lineare Albbildung. Dagegen ist $g(x,y) \coloneqq (ax+b, cy+d)$ für $b,d \neq 0$ nicht linear, denn

$$g(\alpha \cdot (x,y)) = (a \cdot \alpha \cdot x + b, c \cdot \alpha \cdot y + d) \neq \alpha(ax+b, cy+d)$$

falls $\alpha \neq 1$.

Bemerkung. $f\colon V \to W$ linear $\Rightarrow f(0_V) = 0_W$.

Definition 5.5. Sei K Körper und V, W K-Vektorräume. Eine lineare Abbildung $f\colon V \to W$ heißt Isomorphismus von Vektorräumen, falls sie bijektiv ist.

Bemerkung. Dann ist $f^{-1}\colon W \to V$ ebenfalls K-linear, also ein Isomorphismus von K-Vektorräumen.

Beispiel 5.4. $C\colon \mathbb{C} \to \mathbb{C}, C(z) \coloneqq \overline{z}$ ist ein $\mathbb{R}$-linearer Isomorphismus, jedoch nicht $\mathbb{C}$-linear.

Definition 5.6 (Lineare Hülle, erzeugter Vektorraum)**.** Sei K Körper, V ein K-Vektorraum und $M \subseteq V$ eine Menge (kann auch leer sein). Die lineare Hülle von M in V, oder der von M erzeugte Untervektorraum $\langle M \rangle$ ist definiert durch

$$\langle M \rangle \coloneqq \left\{ \sum_{i=1}^{k} \alpha_i v_i \mid k \in \mathbb{N}, v_i \in M, \alpha_i \in K \right\}$$

Notation 5.7. $\sum_{i=1}^{0} \alpha_i v_i \coloneqq 0_V$.

Satz 5.5. *$\langle M \rangle \subseteq V$ ist ein Untervektorraum.*

Beweis. $\langle M \rangle \ni 0_V$. Sei $\alpha \in K, v = \sum_{i=1}^{k} \alpha_i v_i \in \langle M \rangle$. Dann ist $\alpha \cdot v = \sum_{i=1}^{k} (\alpha \cdot \alpha_i) \cdot v_i \in \langle M \rangle$. Sei $v = \sum_{i=1}^{k} \alpha_i \cdot v_i, v' = \sum_{i=1}^{k} \alpha_i' v_i' \in \langle M \rangle$. Dann

$$v + v' = \sum_{i=1}^{k} \alpha_i v_i + \sum_{i=1}^{k'} \alpha_i' v_i' = \sum_{i=1}^{k+k'} \beta_i \cdot w_i \in \langle M \rangle$$

wobei $\forall i \in \{1, .., k\} : \beta_i \coloneqq \alpha_i, w_i \coloneqq v_i$ und $\forall i \in \left\{k+1, ..., k+k'\right\} \beta_i \coloneqq \alpha_{i-k}', w_i \coloneqq v_{i-k}'$. □

Satz 5.6. *Sei $U \subseteq V$ ein Untervektorraum und $M \subseteq U$ eine Teilmenge. Dann ist $\langle M \rangle \subseteq U$ und ist die lineare Hülle in U von M. (U ist ebenfalls K-Vektorraum.)*

Beispiel 5.5. 1. Sei $v \coloneqq (a,b) \in \mathbb{R}^2 \backslash \left\{(0,0)\right\}$. Dann ist $\langle v \rangle = \mathbb{R} \cdot v = \left\{(\alpha a, \alpha b) \in \mathbb{R}^2 \mid \alpha \in \mathbb{R}\right\}$ die Gerade durch $(0,0$ und $(a,b) \in \mathbb{R}^2$.

2. Sei $v \coloneqq (1,0,0), w \coloneqq (0,1,0) \in \mathbb{R}^3$. Dann

$$\langle \{v,w\} \rangle = \big\{(a,b,0) \mid a,b \in \mathbb{R}\big\} \subseteq \mathbb{R}^3$$

ist die x-y-Ebene im $\mathbb{R}^3$.

Definition 5.8 (Erzeugendensystem)**.** Eine Menge $M \subseteq V$ ist ein Erzeugendensystem von V, falls $\langle M \rangle = V$.

Definition 5.9 (Linear unabhängiges System)**.** Eine Teilmenge $M \subseteq V$ heißt linear unabhängig, falls $\forall k \in \mathbb{N}\, \forall v_1, .., v_k \in M\, \forall \alpha_1, ..., \alpha_k \in K$:

$$\sum_{i=1}^{k} \alpha_i v_i = 0_V \Rightarrow \alpha_1 = ... = \alpha_k = 0_K$$

Beispiel 5.6. $\big\{(1,0,1),(1,1,0),(0,1,0),(1,0,0)\big\} \subseteq \mathbb{R}^3$ ist ein Erzeugendensystem, denn $\forall (a,b,c) \in \mathbb{R}^3$ kann man dies als $c \cdot (1,0,1) + b \cdot (0,1,0) + (a-c) \cdot (1,0,0)$ darstellen. Dagegen ist M nicht linear unabhängig, denn $(1,1,0) = (0,1,0) + (1,0,0)$.

Satz 5.7. *Sei $M = \{v_1, ..., v_n\} \subseteq V$ eine endliche Menge. Dann gibt es genau eine lineare Abbildung*

$$f_M \colon K^n \to V, \text{ so dass}$$
$$f_M((0, ..., \underbrace{1}_{k}, ..., 0)) \coloneqq v_k$$

f_M heißt die zu M assoziierte lineare Abbildung.

Beweis. Sei $f_M(a_1, .., a_k) \coloneqq \sum_{k=1}^{n} a_k v_k\, \forall (a_1, ..., a_n) \in K^n$. Dann gilt

$$\begin{aligned} &f_M((a_1, ..., a_n)) + f_M((b_1, .., b_n)) \\ &= \sum_{k=1}^{n} a_k v_k + \sum_{k=1}^{n} b_k v_k = \sum_{k=1}^{n} (a_k + b_k) v_k \\ &= f_M((a_1, ..., a_n) + (b_1, ..., b_n)) \end{aligned}$$

und $\forall \alpha \in K \colon f_M(\alpha \cdot (a_1, .., a_k)) = \sum_{k=1}^{n} \alpha a_k v_k = \alpha \sum_{k=1}^{n} a_k v_k = \alpha \cdot f(a_1, .., a_n)$ □

Satz 5.8. *Eine endliche Teilmenge $M \subseteq V$ ist ein Erzeugendensystem, falls f_M surjektiv ist. Es ist linear unabhängig, falls f_M injektiv ist.*

Bemerkung. $f \colon V \to W$ ist surjektiv $\Leftrightarrow f(V) = W$ und es ist injektiv $\Leftrightarrow \ker(f) = \{0\}$.

Beweis. Es gilt $f_M(K^n) = \langle M \rangle$, also f_M surjektiv $\Leftrightarrow M$ ist Erzeugendensystem. $\ker(f_M) = \big\{(a_1, .., a_k) \in K^n \mid \sum_{i=1}^{n} a_i v_i = 0\big\}$. $\ker(f_M) = \big\{(0, .., 0)\big\} \in K^n \Leftrightarrow M$ ist linear unabhängig. □

Definition 5.10. Sei K Körper und seien V, W K-Vektorräume. Definiere

$$\begin{aligned}\mathrm{Hom}_K(V,W) &:= \{f\colon V \to W \mid f \ K\text{-linear}\}\\ \mathrm{End}_K(V) &:= \mathrm{Hom}_k(V,V)\\ \mathrm{Gl}_K(V) &:= \{f \in \mathrm{End}_K(V) \mid \ f \text{ bijektiv}\}\end{aligned}$$

Satz 5.9. ***a)*** $\mathrm{Hom}_K(V,W)$ *ist ein K-Vektorraum.*

b) $(\mathrm{End}_K(V), +, \circ)$ *mit $\circ$ = 'Verkettung von Funktionen' ist ein Ring, welcher im allgemeinen nicht kommutativ ist und Nullteiler besitzt.*

c) $(\mathrm{Gl}_K(V), \circ)$ *ist eine (nicht abelsche) Gruppe.*

Satz 5.10 (Darstellung einer linearen Abbildung)**.** *Sei $V = K^n, W = K^m$, $m, n \in \mathbb{N} \setminus \{0\}$. Dann existiert ein Isomorphismus von Vektorräumen:*

$$\mathcal{F}\colon \mathrm{Mat}(m \times n, K) \xrightarrow{\cong} \mathrm{Hom}_K(K^n, K^m)$$

mit

$$\mathcal{F}((a_{ij})_{\substack{i\in\{1,..,m\}\\ j\in\{1,..,m\}}})(x_1,...,x_n) = (\sum_{j=1}^{n} a_{ij}x_j, ..., \sum_{j=1}^{n} a_{mj}x_j) \in K^m$$

Beweis. $\mathcal{F}$ **ist linear:** Wir setzen zur Abkürzung $(a_{ij}) := (a_{ij})_{\substack{i\in\{1,..,m\}\\ j\in\{1,..,n\}}}$ für alle Matrizen. Es gilt $\forall \alpha \in K, (a_{ij}) \in \mathrm{Mat}(m \times n, K)$:

$$\alpha \cdot (a_{ij}) = (\alpha a_{ij})$$

und

$$\mathcal{F}((\alpha \cdot a_{ij}))(x_1,..,x_n) = \alpha \cdot \mathcal{F}((a_{ij}))(x_1,..,x_n)$$

Weiter gilt $\forall (a_{ij}), (b_{ij}) \in \mathrm{Mat}(m \times n, K)$:

$$\mathcal{F}((a_{ij}) + (b_{ij})) = \mathcal{F}((a_{ij})) + \mathcal{F}((b_{ij}))$$

$\mathcal{F}$ **ist injektiv:** ($\Leftrightarrow \ker \mathcal{F} = (0_{ij}) \in \mathrm{Mat}(m \times n, K)$). Es gilt

$$\mathcal{F}((a_{ij}))(x_1,..,x_n) = 0 \in K^m \Rightarrow \mathcal{F}((a_{ij}))(0,...,\underbrace{1}_{k},...,0) = 0 \in K^m$$

Also $\forall l \in \{1,..,m\}$:

$$\sum_{j=1}^{n} a_{lj} \cdot x_j = 0$$

wobei $x_j = 0$ für $j \neq k$ und $x_j = 1$, falls $j = k$. Also $a_{lk} = 0$. Da l und k beliebig gewählt waren, folgt also $(a_{ij}) = (0_{ij}) \in \mathrm{Mat}(m \times n, K)$.

$\mathcal{F}$ **ist surjektiv:** Sei $f\colon K^n \to K^m$ linear. Sei $f((0,...,\underbrace{1}_{j},...,0)) = (b_1,...,b_m)$. Sei $a_{ij} \coloneqq b_i$ aus der obigen Gleichung. Dann gilt: $f((0,...,1,...,0)) = (a_{1j},...,a_{mj})$ und

$$f((x_1,..,x_n)) = f(\sum_{j=1}^{n} x_j(0,..,1,..,0)) = \sum_{j=1}^{n} x_j \cdot (a_{1j},..,a_{mj}) = (\sum_{j=1}^{n} a_{1j}x_j,...,\sum_{j=1}^{n} a_{mj}xj)$$

also $f = \mathcal{F}((a_{ij}))$.

□

Bemerkung. Dieser Isomorphismus ist der Grundbaustein der Matrizenrechnung: Die Matrizenmultiplikation (welche noch zu definieren bleibt) entspricht der Verkettung von linearen Abbildungen.

Satz 5.11 (Basisauswahlsatz)**.** *Sei V ein endlich erzeugter K-Vektorraum, d.h. $\exists A \subseteq V$ endlich mit $V = \langle A\rangle$. Dann existiert $A' \subseteq A$ Basis von V.*

Beweis. Durch Induktion nach $n \coloneqq |A|$.

I.A. $n = 0$**:** $V = \langle\emptyset\rangle = \{0\}$ und $\emptyset$ ist Basis von V.

I.S. $n \geq 1$**:** Falls A linear unabhängig ist, so ist A bereits Basis von V. Falls nicht, so sei $A = \{v_1,...,v_n\}$ und $\exists a_1,..,a_n \in K$ so dass

$$a_1v_1 + ... + a_nv_n = 0 \wedge \exists i \in \{1,..,n\} : a_i \neq 0$$

Nach einer eventuellen Umbenennung von $v_1,..,v_n$ können wir annehmen, dass $a_n \neq 0$. Dann ist

$$v_n = -a_n^{-1} \cdot \sum_{k=1}^{n-1} a_k v_k \in \langle\{v_1,..,v_{n-1}\}\rangle.$$

Also $A' \coloneqq A\backslash\{v_n\}$ erzeugt ebenfalls V. Nach Induktionsvoraussetzung existiert nun $A'' \subseteq A' \subseteq A$ Basis von V.

□

Lemma 5.12. *Sei $M \coloneqq \{f_1,..,f_m\}$ Basis eines K-Vektorraums $V \neq \{0\}$, $m \in \mathbb{N}\setminus\{0\}$ und sei $v \coloneqq \sum_{i=1}^{m} a_i f_i$ mit $\forall i \in \{1,..,m\} : a_i \in K, a_i \neq 0$. Dann ist $M' \coloneqq \{v, f_2,..,f_m\}$ ebenfalls eine Basis von V.*

Beweis. **Lineare Unabhängigkeit:** Seien $a, b_2,..,b_m \in K\colon a\cdot v + b_2f_2 + ... + b_mf_m = 0$. Dann $aa_1f_1 + (aa_2 + b_2)f_2 + ... + (aa_m + b_m)f_m = 0$. Da M Basis ist folgt $aa_1 = 0 \wedge \ \forall i \in \{2,..,m\} : aa_i + b_i = 0$, also $a = 0$, da $a_1 \neq 0$ und damit $\forall i \in \{2,..,m\} : b_i = 0$. Also ist M' linear unabhängig.

Erzeugendensystem: Dafür genügt es $M \subseteq \langle M' \rangle$ zu zeigen, denn dann $V \subseteq \langle M \rangle \subseteq \langle M' \rangle$. Es gilt $f_2, .., f_m \in M' \subseteq \langle M' \rangle$. Weiter ist $f_1 = a_1^{-1}(v - \sum_{i=2}^{m} a_i f_i)$, also $f_1 \in \langle M' \rangle$ und M' ist Basis von V.

□

Lemma 5.13. *Falls $M \coloneqq \{e_1, .., e_n\}$ eine Basis eines K-Vektorraum V ist, so ist $\forall a_2, .., a_n \in K$ $M' \coloneqq \{e_1, e_2 + a_2 e_1, .., e_n + a_n e_1\}$ ebenfalls eine Basis von V.*

Beweis. **Lineare Unabhängigkeit:** Seien $b_1, .., b_n \in K$: $b_1 e_1 + \sum_{i=2}^{n} b_i \cdot (e_i + a_i e_1) = 0$. Dann $(b_1 + \sum_{i=2}^{n} b_i a_i) e_i + \sum_{i=2}^{n} b_i e_i = 0$. Da M Basis ist, also $\forall i = 2, .., n$: $b_i = 0 \wedge b_1 + \sum_{i=2}^{n} b_i a_i = 0 \Rightarrow b_1 = 0$. Also ist M' linear unabhängig.

Erzeugendensystem: Wir zeigen $M \subseteq \langle M' \rangle$. Es ist $e_1 \in M'$ und $\forall i \in \{2, .., n\} : e_i = (e_i + a_i e_1) - a_i e_1 \in \langle M' \rangle$. Also erzeugt $M'V$ und ist somit eine Basis.

□

Satz 5.14. *Seien $M_1 \coloneqq \{e_1, .., e_n\}$ und $M_2 \coloneqq \{f_1, .., f_m\}$ Basen eines K-Vektorraum V, $m, n \in \mathbb{N}$. Dann gilt $m = n$. Insbesondere ist die Dimension $\dim_K V \coloneqq n$ wohldefiniert.*

Bemerkung. Falls V nicht endlich erzeugt ist, so ist seine Dimension per Definition unendlich: $\dim_K V = \infty$.

Beweis. Wir zeigen per Induktion die folgende Aussage für $n \in \mathbb{N}$:

$P(n)$: Jeder K-Vektorraum V, der eine n-elementige Basis besitzt besitzt nur n-elementige Basen

I.A. $n = 0$: Dann $V = \{0\}$ und er besitzt nur $\emptyset$ als Basis, denn 0 kann zu keinem linear unabhängigen System von Vektoren gehören.

I.S. $n \geq 1$: Sei V ein K-Vektorraum und $M \coloneqq \{e_1, .., e_n\}$ eine Basis und $M_1 \coloneqq \{f_1, .., f_m\}, m \in \mathbb{N} \setminus \{0\}$ eine andere Basis. Zu zeigen ist $m = n$. Nach eventueller Umbenennung der Elemente in M, M_1 können wir annehmen, dass $\exists a_1, .., a_m \in K$: $e_1 = \sum_{i=1}^{m} a_i f_i$, $a_1 \neq 0$. Lemma 5.12 zeigt, dass $M_2 \coloneqq \{e_1, f_2, ..., f_m\}$ eine Basis ist. Dann gilt $\forall j \in \{2, .., n\} : \exists a_{1j}, .., a_{mj} \in K$:

$$e_j = a_{1j} e_1 + \sum_{k=2}^{m} a_{kj} \cdot f_k \Leftrightarrow e'_j \coloneqq e_j - a_{1j} e_1 = \sum_{k=2}^{m} a_{kj} \cdot f_k \qquad (*)$$

Mit Lemma 5.13 folgt, dass $M_3 \coloneqq \{e_1, e'_2, .., e'_n\}$ eine Basis von V ist. Dann ist jedes Element $f_2, .., f_m$ als Linearkombination von $e_1, e'_2, .., e'_n$ darstellbar. Wir zeigen nun: $\forall i \in \{2, .., m\} : f_i \in \langle \{e'_2, .., e'_n\} \rangle$. Seien $b_1, .., b_n \in K$ so dass

$$f_i = b_1 e_1 + \sum_{j=2}^{n} b_j e'_j \qquad (**)$$

Diese b_i existieren, da M_3 eine Basis ist. Falls $b_1 \neq 0$ impliziert 5.12, dass $M_4 \coloneqq \{f_i, e'_2, .., e'_n\}$ eine Basis von V ist. Aus $(*)$ folgt aber, dass $e'_2, .., e'_n \in \langle\{f_2, .., f_m\}\rangle$ ebenfalls $f_i \in \langle\{f_2, .., f_m\}\rangle$ für $i \in \{2, .., m\}$, also $V = \langle M_4\rangle = \langle\{f_2, .., f_m\}\rangle$, also existieren $c_2, .., c_m \in K\colon f_1 = c_2 f_2 + ... + c_m f_m$, was der linearen Unabhängigkeit von $\{f_1, .., f_m\}$ widerspricht. Also $b_1 = 0$ in $(**)$, was zeigt, dass $f_i \in \langle\{e'_2, .., e'_n\}\rangle$ für $i \in \{2, .., m\}$. Sei $U \coloneqq \langle\{e'_2, .., e'_n\}\rangle$. Es gilt $f_2, .., f_m \in U$, aber auch $e'_2, .., e'_n \in \langle\{f_2, .., f_m\}\rangle$. Es folgt $U = \langle\{f_2, .., f_m\}\rangle = \langle\{e'_2, .., e'_n\}\rangle$ und $\{f_2, .., f_m\}, \{e'_2, .., e'_n\}$ sind beides linear unabhängige Systeme. Also hat U eine $n-1$-elementige Basis und nach Induktionsvoraussetzung gilt damit $m - 1 = n - 1$. Also $m = n$.

□

Beispiel 5.7. 1) $\dim_K K^n = n$

2) $\dim_{\mathbb{R}} \mathbb{C} = 2$

3) $\mathbb{R}$ ist nicht endlich erzeugt als $\mathbb{Q}$-Vektorraum, also $\dim_{\mathbb{Q}} \mathbb{R} = \infty$. Der Beweis wird hier nicht ausgeführt.

4) $\dim_K \{0\} = 0$ und $\dim_K V > 0$ falls $V \neq \{0\}$.

Satz 5.15. *Eine lineare Abbildung $f\colon K^m \to K^n$ kann nur im Fall $m = n$ bijektiv sein. Allgemeiner kann eine lineare Abbildung $f\colon V \to W$ zwischen K-Vektorräumen nur im Fall $\dim_K V = \dim_K W$ bijektiv sein.*

Beweis. Sei $f\colon V \to W$ ein Isomorphismus von K-Vektorräumen. Nehmen wir an, V ist endlich erzeugt. Dann $V = \langle M\rangle$, $M = \{e_1, .., e_n\}$ eine Basis von V. Dann ist $f(M)$ eine Basis von W, also $\dim_K W = \dim_K V = n$. □

Bemerkung. Der Basisauswahlsatz und der Dimensionssatz gelten auch für unendlich erzeugte Vektorräume. Zum Beispiel, falls M eine unendliche Menge ist, ist K^M ein Vektorraum und $K^{(M)} \coloneqq \{f\colon M \to K \mid f \text{ Abbildung mit endlichem Träger}\}$ ebenfalls ein K-Vektorraum. Hierbei ist der Träger $\operatorname{supp}(f)$ definiert durch:

$$\operatorname{supp}(f) \coloneqq \{x \in M \mid f(x) \neq 0\}$$

Es gilt $\operatorname{supp}(f) = 0 \Leftrightarrow f$ ist die triviale Abbildung. Dann ist $M_0 \coloneqq \{f_x \mid x \in M\} \subseteq K^{(M)} \subseteq K^M$ eine Basis von $K^{(M)}$, wobei

$$f_x\colon M \to K, \quad f_x(y) = \begin{cases} 1, & x = y \\ 0, & \text{sonst} \end{cases}$$

Es kann beispielsweise $g\colon M \to K, \quad g(x) = 1\ \forall x \in M$ nicht als Linearkombination von Elementen aus M_0 dargestellt werden, da diese Linearkombination per Definition endlich sein müsste.

Satz 5.16. *Sei $M = \{e_1, .., e_n\}$ linear unabhängig in K-Vektorraum $V = \langle \widetilde{M} \rangle$ mit $\widetilde{M} = \{f_1, .., f_m\}$. Dann existiert $\overline{M} \subseteq M \cup \widetilde{M}$ Basis und $M \subseteq \overline{M}$. Ein linear unabhängiges System kann also durch hinzufügen von Vektoren aus einen vorgegebenem Erzeugendensystem zu einer Basis erweitert werden.*

Beweis. Falls $U \coloneqq \langle M \rangle \subsetneq V$, so ist $\widetilde{M} \not\subseteq U$. Also $\exists i \in \{1, .., m\} : f_i \notin U$. Nach eventueller Umbenennung von $f_1, .., f_m$ können wir annehmen, dass $f_1 \notin U$. Dann ist $\{f_1\} \cup M$ ebenfalls linear unabhängig:

$$af_1 + \sum_{i=1}^{n} a_i e_i = 0 \Rightarrow \begin{cases} a \neq 0 \Rightarrow f_1 = -a^{-1} \cdot \sum_{i=1}^{n} a_i e_i \in \langle M \rangle : \text{ Widerspruch!} \\ a = 0 \Rightarrow \sum_{i=1}^{n} a_1 e_1 = 0 \Rightarrow a_1 = ... = a_n = 0 \end{cases}$$

Sei $M_1 \coloneqq M \cup \{f_1\}$ linear unabhängig. Wir wiederholen das Verfahren und ergänzun M_1 zu $M_2 \coloneqq M_1 \cup \{f_2\}$ bis wir $\langle M_k \rangle = V$ erreichen, also M_k Basis von V ist. □

Korollar 5.17. *Sei U ein Untervektorraum im endlich erzeugten Vektorraum V. Dann ist $\dim_K U \leq \dim_K V$. Die Gleichheit gilt $\Leftrightarrow U = V$.*

Satz 5.18. ***1)*** *Sei V endlich erzeugter K-Vektorraum. Dann ist V isomorph zu K^n wobei $n = \dim_K V$.*

2) *Sei $f \colon V \to W$ linear wobei V endlich erzeugt sei. Dann gilt $\dim_K(\ker f) + \dim_K(\operatorname{Im} f) = \dim_K V$.*

3) *Sei $f \colon V \to V$ eine lineare Abbildung und V endlich erzeugt. Dann sind äquivalent:*

- ***a)*** *f injektiv.*
- ***b)*** *f surjektiv.*
- ***c)*** *f bijektiv.*

Beweis. **1)** $\forall \{e_1, .., e_n\} \subseteq V$ Basis existiert $f \colon K^n \to V$ linear mit $f(e_i^0) = e_i \; \forall i \in \{1, .., n\}$ wobei $e_i^0 = (0, .., \underbrace{1}_{i}, 0, ..., 0) \in K^n$. Dann $\{e_1, ., e_n\}$ ist Basis $\Rightarrow f$ ist bijektiv $\Leftrightarrow f$ ist Isomorphismus, also $V \cong K^n$.

2) Sei $\{f_1, .., f_m\}$ Basis von $\operatorname{Im} f = f(V) \subseteq W$ und seien $e_1, .., e_m \in V$ so dass $f(e_i) = f_i$ für $i = 1, .., m$. Sei $\{v_1, .., v_k\} \subseteq \ker f$ Basis von $\ker f$. Wir zeigen $\{v_1, .., v_k, e_1, .., e_m\}$ ist eine Basis von V.

Lineare Unabhängigkeit: Seien $a_1, .., a_k, b_1, .., b_m \in K$ mit

$$\sum_{j=1}^{k} a_j v_j + \sum_{i=1}^{m} b_i e_i = 0$$

Wir wenden f an:

$$\sum_{i=1}^{m} b_i f(e_i) = 0 \Rightarrow b_1 = ... = b_m = 0$$

da $\{f_1, .., f_m\}$ linear unabhängig ist. Also $\sum_{j=1}^{k} a_j v_j = 0 \Rightarrow a_1 = ... = a_k = 0$.

Erzeugendensystem: Sei $x \in V$. Dann $\exists b_1, .., b_m \in K$:

$$f(x) = \sum_{i=1}^{m} b_i f_i$$

da $\{f_1, .., f_m\}$ eine Basis von $\operatorname{Im} f$ ist und $f(x) \in \operatorname{Im} f$. Also $f(x - \sum_{i=1}^{m} b_i e_i) = 0$, das heißt $x - \sum_{i=1}^{m} b_i e_i \in \ker f \Rightarrow \exists a_1, .., a_k \in K\colon x - \sum_{i=1}^{m} b_i e_i = \sum_{j=1}^{k} a_j v_j$, also $x = \sum_{i=1}^{m} b_i e_i + \sum_{j=1}^{k} a_j v_j \in \langle \{v_1, .., v_k, e_1, .., e_m\} \rangle$.

3) Folgt aus 2): f injektiv $\Leftrightarrow \dim \ker f = 0 \Leftrightarrow \dim \operatorname{Im} f = \dim V \Leftrightarrow f$ surjektiv.

□

Erinnerung: $A \in \operatorname{Mat}(m \times n, K) \rightsquigarrow f_A\colon K^n \to K^m$ lineare Abbildung. $A = (a_{ij})_{\substack{i \in \{1,..m\} \\ j \in \{1,..,n\}}} \rightsquigarrow f_A(x_1, .., x_n) = (\sum_{j=1}^{n} a_{1j} x_j, ..., \sum_{j=1}^{n} a_{mj} x_j)$.

Definition 5.11 (Matrixmultiplikation)**.** Für $A \in \operatorname{Mat}(m \times n, K), B \in \operatorname{Mat}(n \times p, K)$ und die zugehörigen Abbildungen $f_A\colon K^n \to K^m$, $f_B\colon K^p \to K^n$ sei C die Matrix, die zur linearen Abbildung $f_C \coloneqq f_A \circ f_B\colon K^p \to K^m$ gehört. Falls $C = (c_{ik})_{\substack{i \in \{1,..,m\} \\ k \in \{1,..,p\}}}, A = (a_{ij})_{\substack{i \in \{1,..,m\} \\ j \in \{1,..,n\}}}, B = (b_{jk})_{\substack{j \in \{1,..,n\} \\ k \in \{1,..,p\}}}$, so gilt

$$c_{ik} = \sum_{j=1}^{n} a_{ij} b_{jk} \quad \forall i \in \{1, .., m\}, k \in \{1, .., p\}$$

Man kann auch sagen, dass die p Spalten der Matrix $A \cdot B$ genau die Vektoren $A \cdot b_i$ sind, wobei $b_1, ..., b_p$ die Spalten(vektoren) der Matrix B sind.

Beispiel 5.8. 1)

$$\begin{pmatrix} 1 & 1 \\ 2 & 1 \\ 0 & 3 \end{pmatrix} \cdot \begin{pmatrix} 1 & 4 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} 1 \cdot 1 + 1 \cdot (-1) & 1 \cdot 4 + 1 \cdot 0 \\ 2 \cdot 1 + 1 \cdot (-1) & 2 \cdot 4 + 1 \cdot 0 \\ 0 \cdot 1 + 3 \cdot (-1) & 0 \cdot 4 + 3 \cdot 0 \end{pmatrix} = \begin{pmatrix} 0 & 4 \\ 1 & 8 \\ -3 & 0 \end{pmatrix} \in \operatorname{Mat}(3 \times 2, \mathbb{Q})$$

2)

$$\begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix} \cdot \begin{pmatrix} b_1 ... b_n \end{pmatrix} = \begin{pmatrix} a_1 b_1 & \dots & a_1 b_n \\ \vdots & & \vdots \\ a_m b_1 & \dots & a_m b_n \end{pmatrix} \in \operatorname{Mat}(m \times n, K)$$

3)

$$\Big(b_1, ..., b_n\Big) \cdot \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = a_1 b_1 + ... + a_n b_n \in \mathrm{Mat}(1 \times 1, K) = K$$

Bemerkung. Die Multiplikation der Matrizen $A \cdot B$ ist nur dann definiert, falls Spaltenzahl von A = Zeilenzahl von B. Insbesondere ist $A \cdot B$ und $B \cdot A$ definiert, falls $A \in \mathrm{Mat}(m \times n, K)$ und $B \in \mathrm{Mat}(n \times m, K)$. Es bildet $(\mathrm{Mat}(n \times n, K), +, \cdot)$ einen Ring, welcher isomorph zu $\mathrm{End}(K^n) = \mathrm{Hom}(K^n, K^n) = \big\{f \colon K^n \to K^n \mid f \text{ linear}\big\}$ ist. Der Isomorphismus ist $A \mapsto f_A$ wie oben.

Bemerkung. $A \in \mathrm{Mat}(m \times n, K) \rightsquigarrow f_A \colon K^n \to K^m$ linear. Sei $\big\{e_1^0, ..., e_n^0\big\}$ die Standardbasis von K^n wobei e_i^0 der Spaltenvektor $\begin{pmatrix} 0 \\ \vdots \\ \mathrm{i} \\ \vdots \\ 0 \end{pmatrix} \in K^n$ mit 1 an der i-ten Stelle ist. Die i-te Spalte von $A = (a_{ij})_{i \in \{1,..,m\}, j \in \{1,..,n\}}$ ist

$$\begin{pmatrix} a_{1i} \\ \vdots \\ a_{mi} \end{pmatrix} = \begin{pmatrix} a_{11} & ... & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & ... & a_{mn} \end{pmatrix} \cdot e_i^0$$

Also sind die Spalten von A genau die Bilder der Vektoren e_i^0 der Standardbasis von K^n unter f_A.

Definition 5.12. Der Rang einer $m \times n$-Matrix A ist die Dimension des Vektorraums $\mathrm{Im}(f_A) = f_A(K^n) \subseteq K^m$. Insbesondere ist A in $\mathrm{Mat}(n \times n, K)$ invertierbar $\Leftrightarrow$ $\mathrm{rang}(A) = n$.

5.2 Lineare Gleichungssysteme

Seien $A = (a_{ij})_{i \in \{1,..,m\}, j \in \{1,..,n\}} \in \mathrm{Mat}(m \times n, K), b = (b_i)_{i \in \{1,..,m\}} \in K^m$. Es ist ein Vektor $x = (x_j)_{j=1,..,n} \in K^n$ gesucht, so dass gilt:

$$Ax = b, \tag{L}$$

das heißt

$$\begin{pmatrix} a_{11} & ... & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & ... & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$$

Oder noch ausführlicher:

$$\begin{aligned} a_{11} \cdot x_1 + ... + a_{1n} \cdot x_n &= b_1 \\ &\vdots \\ a_{m1} \cdot x_1 + ... + a_{mn} \cdot x_n &= b_m \end{aligned}$$

Das lineare Gleichungssystem zu lösen ist dazu äquivalent, die Lösungsmenge

$$S(A, b) = \{x \in K^n \mid Ax = b\}$$

zu bestimmen. Im Fall $b = 0 \in K^m$ heißt (L) *homogen*, sonst *inhomogen*. Die Lösungsmenge eines homogenen linearen Gleichungssystems, von $A \in \mathrm{Mat}(m \times n, K)$ bestimmt, ist der kern der zugehörigen Abbildung $f_A \colon K^n \to K^m$, also

$$S(A, 0) = \ker(f_A)$$

Es folgt, dass $S(A, 0)$ ein Untervekottaum in K^n ist, der die Dimension $n - \mathrm{rang}(A)$ besitzt.

Bemerkung. Die Inverse Matrix A^{-1} zu A entspricht der Umkehrabbildung $f_A^{-1} = f_{A^{-1}} \colon K^m \to K^n$ für $f_A \colon K^n \to K^m$. Diese existiert also nur für bijektive lineare Abbildungen, also muss insbesondere $m = n$ gelten. Falls $A \in \mathrm{Mat}(n \times n, K)$ invertierbar ist, so $Ax = b \Leftrightarrow A^{-1}Ax = A^{-1}b \Leftrightarrow x = A^{-1}b$, und $A^{-1}b$ ist die eindeutige Lösung der Gleichung.

Kriterium Das inhomogene lineare Gleichungssystem $Ax = b$ hat eine Lösung $\Leftrightarrow \mathrm{rang}(A) = \mathrm{rang}(A \mid b)$ wobei

$$(A \mid b) = \left(\begin{array}{ccc|c} a_{11} & ... & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{m1} & ... & a_{mn} & b_m \end{array}\right)$$

Beweis. Es ist $\mathrm{rang}(A) = \dim(f_A(K^n)) = \dim(\langle \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix}, ..., \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix} \rangle)$ Bezeichnen wir mit $A_{(j)}$ die j-te Spalte der Matrix A. Dann also $\mathrm{rang}(A) = \dim\langle \{A_{(1)}, .., A_{(n)}\} \rangle$ und $\mathrm{rang}(A \mid b) = \dim\langle \{A_{(1)}, .., A_{(n)}, b\} \rangle$. Es gilt also $\mathrm{rang}(A) \leq \mathrm{rang}(A \mid b)$. Die Gleichheit ist äquivalent zu $b \in \langle \{A_{(1)}, .., A_{(n)}\} \rangle = f_A(K^n) \Leftrightarrow$ (L) hat eine Lösung. $\square$

5.3 Methode zum lösen von linearen Gleichungssystemen und invertieren von Matrizen

Definition 5.13. Die folgenden *Elementaren Zeilenumformungen* (EZU) von $A \in \mathrm{Mat}(m \times n, K)$ entstehen durch Multiplikation von links mit den folgenden $m \times m$ Matrizen, $A = (a_{ij})_{i \in \{1,..,m\}, j \in \{1,..,n\}}$.

Typ I, $(S_i(\lambda))$ für $i \in \{1, .., m\}, \lambda \in K^\times$.

$$S_i(\lambda) = \begin{pmatrix} 1 & 0 & ... & 0 & & \\ 0 & 1 & ... & & 0 & \\ \vdots & \vdots & \ddots & & & \\ 0 & & 0 & \lambda & 0 & \\ & & & 0 & 1 & \\ & & & & & \ddots \end{pmatrix}$$

Die Einträge von $S_i(\lambda)$ außerhalb der Diagonalen sind 0 und auf der Diagonalen stehen überall $1 \in K$ außer an der (i,i)-ten Position, wo ein λ steht. Dann gilt:

$$S_i(\lambda) \cdot A = \begin{pmatrix} a_{11} & ... & a_{1n} \\ \vdots & & \vdots \\ \lambda a_{i1} & ... & \lambda a_{in} \\ \vdots & & \vdots \\ a_{m1} & ... & a_{mn} \end{pmatrix}$$

Also die i-te Zeile von A wird mit λ multipliziert, alle anderen Einträge bleiben unverändert.

Typ II, $T_{ij}(\lambda)$ $i, j \in \{1, .., m\}, i \neq j, \lambda \in K$.

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

Die Diagonaleinträge von $T_{ij}(\lambda)$ sind 1. Alle anderen Einträge sind 0 außer der

(i,j)-te Eintrag (i-te Zeile, j-te Spalte), welcher gleich λ ist. Dann gilt:

$$T_{ij}(\lambda)\cdot A = \begin{pmatrix} a_{11} & ... & a_{1n} \\ \vdots & & \vdots \\ a_{i1}+\lambda a_{j1} & ... & a_{in}+\lambda a_{jn} \\ \vdots & & \vdots \\ a_{m1} & ... & a_{mn} \end{pmatrix}$$

Auf die i-te Zeile von A wird die j-te Zeile λ-mal draufaddiert, die anderen Einträge bleiben alle unverändert.

Typ III, W_{ij} , $i,j \in \{1,..,m\}, i<j$.

$$W_{ij} = \begin{pmatrix} 1 & & & & & & & \\ & \ddots & & & & & & \\ & & 1 & & & & & \\ & & & 0 & ... & 1 & & \\ & & & \vdots & \ddots & \vdots & & \\ & & & 1 & ... & 0 & & \\ & & & & & & \ddots & \\ & & & & & & & 1 \end{pmatrix}$$

Wir erhalten W_{ij} aus Vertauschung der i-ten und j-ten Zeile der Einheitsmatrix

$$I_m = \begin{pmatrix} 1 & \dots & 0 \\ & \ddots & \\ 0 & \dots & 1 \end{pmatrix}$$

Bemerkung. Man kann $I_m \in \mathrm{Mat}(m\times m, K)$ auch durch die Kroneckersymbole $\delta_{ij} = \begin{cases} 0, i \neq j \\ 1, i = j \end{cases}$ schreiben: $I_m = (\delta_{ij})_{i,j\in\{1,..,m\}}$.

Man erhält dann:

$$W_{ij}\cdot A = \begin{pmatrix} a_{11} & ... & a_{1n} \\ \vdots & & \vdots \\ a_{j1} & ... & a_{jn} \\ \vdots & & \vdots \\ a_{i1} & ... & a_{in} \\ \vdots & & \vdots \\ a_{m1} & ... & a_{mn} \end{pmatrix}$$

Die i-te und j-te Zeile von A werden also miteinander vertauscht, während alle anderen Zeilen von A unverändert bleiben.

Satz 5.19. *Die Matrizen $S_i(\lambda), i \in \{1,..,m\}, \lambda \in K^\times$, $T_{ij}(\lambda), i,j \in \{1,..,m\}, i \neq j, \lambda \in K$ und $W_{ij}, i,j \in \{1,..,m\}, i < j$ aus $\mathrm{Mat}(m \times m, K)$ sind alle invertierbar und $S_i(\lambda)^{-1} = S_i(\lambda^{-1}), T_{ij}(\lambda)^{-1} = T_{ij}(-\lambda), W_{ij}^{-1} = W_{ij}$.*

Beweis. Übung. □

Definition 5.14 (Zeilenstufenform - ZSF)**.** Eine Matrix $A \in \mathrm{Mat}(m \times n, K)$ ist in Zeilenstufenform, falls $\exists r \in \{0,..,m\}$ und $\forall i \in \{1,..,r\}\ \exists j_i \in \{1,..,n\}$, so dass $j_1 < j_2 < ... < j_r$ und $\forall i \in \{1,..,r\}\ \forall j \in \{1,..,j_i - 1\} : a_{ij} = 0, a_{ij_i} \neq 0$ und $\forall i \in \{r+1,..,m\}\ \forall j \in \{1,..,n\} : a_{ij} = 0$. Anschaulich:

$$\left(\begin{array}{ccccccccccc}
0 & \cdots & 0 & * & \cdots & & & & & & \\
0 & \cdots & & 0 & \cdots & 0 & * & \cdots & & & \\
0 & \cdots & & & & \cdots & 0 & * & \cdots & & \\
0 & \cdots & & & & & & \cdots & 0 & &
\end{array}\right)$$

Im unteren Bereich sind alle Einträge 0, die Einträge $*$ (am Anfang jeder "Stufe") sind nicht Null.

Satz 5.20. *Jedes $A \in \mathrm{Mat}(m \times n, K)$ kann durch EZUs vom Typ II, III in eine Matrix $A' \in \mathrm{Mat}(m \times n, K)$ transformiert werden, welche in Zeilenstufenform ist. Die Anzahl an Zeilen von A', welche nicht trivial sind, ist $r = \mathrm{rang}(A)$.*

Beweisalgorithmus: Es gilt $\mathrm{rang}(A) = 0 \Leftrightarrow A = 0 \Leftrightarrow A' = 0$. In diesem Fall ist A bereits in Zeilenstufenform. Falls $A \neq 0$ sei j_i die Nummer der ersten Spalte von A, dich nicht nur aus Nullen besteht. Es wird nach eventueller Vertauschung der Zeilen (Typ III) die erste Zeile von A so festgelegt, dass der Eintrag $a'_{1j_1} \neq 0$ ist. Es werden nun alle anderen $m-1$ Zeilen betrachtet. Von der $k > 1$-ten Zeile wird die erste Zeile $(a'_{1j_1})^{-1} \cdot a'_{kj_1}$ mal subtrahiert (Typ II). Die entstandene Matrix A'' hat die Einträge $a'_{1j} = a'_{1j}\ \forall j \in \{1,..,n\}$ und $a''_{ij} = 0\ \forall j \leq j_1\ \forall i \in \{2,..,m\}$. Es wird das obige Verfahren für die so entstandene Matrix $(a''_{ij})_{i \in \{2,..,m\}, j \in \{1,..,n\}}$ wiederholt, so wird die 2. Zeile festgelegt. Dies wiederholen wir nun für alle Zeilen. Die so erhaltene Matrix A' ist in Zeilenstufenform. Da bei jedem Schritt die Matrix A mit einer invertierbaren $m \times m$-Matrix multipliziert wird besitzt $A' = M \cdot A$ für $M \in \mathrm{Mat}(m \times m, K)$ invertierbar $\mathrm{rang}(A') = \mathrm{rang}(A)$, da $f_M \colon K^m \to K^m$ dann bijektiv ist, also einen Isomorphismus zwischen $f_A(K^n)$ und $f_{A'}(K^n)$ induziert. Die Spalten $j_1,..,j_r$ sind linear unabhängig und erzeugen $f_{A'}(K^n)$. Also $r = \mathrm{rang}(A')$. Wir zeigen hierfür die lineare Unabhängigkeit: Seien $b_1,..,b_n \in K, b_j = 0$ falls $j \notin \{j_1,..,j_r\}$ (also werden eigentlich nur r Elemente $b_{j_1},..,b_{j_r}$ gesucht), so dass $(\sum_{j=1}^n a'_{ij} b_j)_{i \in \{1,..,m\}} \in K^m$ ist

eine Linearkombination der Spalten $j_1, .., j_r$. Dann $\sum_{j=1}^{n} a'_{ij} b_j = 0 \; \forall i \in \{1, .., r\}$ (für $i > r$ sind alle Einträge $a'_{ij} = 0$) $\Leftrightarrow$

$$\begin{aligned} a'_{1j_1} b_{j1} + a_{1j_2} b_{j_2} + ... + a_{1j_r} b_{j_r} &= 0 \\ &\vdots \\ a'_{rj_1} b_{j_1} + a'_{rj_2} b_{j_2} + ... + a_{rj_r} b_{j_r} &= 0 L') \end{aligned} \tag{(()}$$

Dieses System ist in *Dreiecksform*, das heißt in der letzten Zeile sind alle Koeffizienten gleich Null bis a'_{rj_r}, in der vorletzten alle bis $a'_{r-1,j_{r-1}}$ und so weiter, bis zur ersten Zeile. Wir lösen (L') von unten nach oben. $a'_{rj_r} \neq 0 \Rightarrow b_{j_r} = 0$. Dann ist die vorletzte Zeile $a'_{r-1,j_{r-1}} b_{j_{r-1}} + a_{r-1,j_r} b_{j_r} = 0$ aber $b_{j_r} = 0$ und $a_{r-1,j_{r-1}} \neq 0 \Rightarrow b_{j_{r-1}} = 0$. Es folgt also Schrittweise $b_{j_i} = 0$ für $i = 1, .., r$. Also sind die Spalten $A'_{(j_1)}, .., A'_{(j_r)}$ linear unabhängig. Wir zeigen nun, dass $A'_{(j)} \in W \coloneqq \langle \left\{ A'_{(j_1)}, .., A'_{(j_r)} \right\} \rangle \; \forall j \in \{1, .., n\}$.

1) Falls $j < j_1$, so $A'_{(j)} = 0 \in K^m$.

2) Falls $\exists i \in \{1, .., r\} : j = j_i \Rightarrow A'_{(j)} = A_{(j_i)} \in W$.

3) Falls $j_i < j < j_{i+1}, i \in \{1, .., r\}$ (im Fall $i = r$ definieren wir $j_{r+1} \coloneqq (n+1)$) zeigen wir $A'_{(j)} \in \langle \left\{ A'_{(j_1)}, ..., A'_{(j_i)} \right\} \rangle$ durch lösen des linearen Gleichungssystems

$$\begin{aligned} a'_{1j_1} x_1 + a'_{1j_2} x_2 + ... + a'_{1j_i} x_i &= a'_{ij} \\ a'_{2j_2} x_2 + ... + a'_{2j_i} x_i &= a'_{2j} \\ &\vdots \\ a'_{ij_i} x_i &= a_{ij} \end{aligned} \tag{L_i}$$

Das System wird von unten nach oben gelöst (also zuerst wird x_i aus der letzten Gleichung bestimmt, dann x_{i-1} aus der vorletzten und am Ende x_1 aus der ersten Gleichung und den bereits bestimmten $x_2, .., x_i$ berechnet. Es gilt dann $A'_{(j)} = x_1 A'_{(j_1)} + ... + x_i A'_{(j_i)}$, also $A'_{(j)} \in \langle \left\{ A'_{(j_1)}, ..., A'_{(j_r)} \right\} \rangle$ und somit ist $\left\{ A'_{(j_1)}, ..., A_{(j_r)} \right\}'$ eine basis von $f_{A'}(K^n)$, also $r = \text{rang}(A') = \text{rang}(A)$.

□

Lösen eines linearen Gleichungssystems Für $A \in \text{Mat}(m \times n, K), b \in K^m$ werde ein $x \in K^n$ gesucht. Es wird A in Zeilenstufenform gebracht, also $A' = MA$ mit $M \in \text{Mat}(m \times m, K)$ invertierbar und $A' \in \text{Mat}(m \times n, K)$ in Zeilenstufenform. Dann $Ax = b \Leftrightarrow A'x = b'$ mit $b' = Mb \in K^m$. Das lineare Gleichungssystem $A'x = b'$

wird von unten nach oben gelöst, in dem alle Komponenten von x mit Ausnahme von $x_{j_1}, .., x_{j_r}$ beliebige Werte vorgegeben werden. Dieses Sytem ist in der Form

$$\begin{aligned} a'_{1j_1}x_{j_1} + a'_{1j_2}x_{j_2} + ...+a'_{1j_r}x_{j_r} &= c_1 \\ a'_{2j_2}x_{j_2} + ...+a'_{2j_r}x_{j_r} &= c_2 \\ &\vdots \\ a'_{r-1,j_{r-1}}x_{j_{r-1}}+a'_{r-1j_r}x_{j_r} &= c_{r-1} \\ a'_{rj_r}x_{j_r} &= c_r \end{aligned} \tag{L'}$$

wobei

$$c_i \coloneqq b'_i - \sum_{\substack{j=1 \\ j\notin\{j_1,...,j_r\}}}^{n} a_{ij} \cdot x_j \in K$$

Also die Komponenten x_j mit $j \notin \{j_1, .., j_r\}$ werden beliebig gewählt und $x_{j_1}, ..., x_{j_r}$ werdem aus dem System (L') von unten nach oben bestimmt. Da die Koeffizienten $c_1, .., c_r$ von den Werten von $(x_j)_{j\in\{1,..,n\}\setminus\{j_1,..,j_r\}}$ abhängen, hängen auch die bestimmten Werte von $x_{j_1}, .., x_{j_r}$ davon ab.

Eine $n \times n$-Matrix invertieren

Sei $A \in \mathrm{Mat}(n \times n, K)$ sei $A' = MA$ die Matrix, die durch Elementare Zeilenumformungen vom Typ II und III in Zeilenstufenform gebracht wird. Es gilt $\mathrm{rang}(A) = n \Leftrightarrow$ die Zeilenstufenform A' ist in Dreiecksform, also $j_i = i \in \{1, .., n\}$ und die Diagonaleintrage $a'_{ii} \neq 0 \; \forall i \in \{1, .., n\}$.

Es werden nun Elementare Zeilenumformungen angewandt, und zwar $S \coloneqq \prod_{i=1}^{n} S_i(a'^{-1}_{ii}) = S_1(a'^{-1}_{ii}) \cdots S_n(a'^{-1}_{nn})$. Es folgt $A'' \coloneqq SA'$ ist auch in Dreiecksform, aber alle Diagonaleinträge sind 1. Es werden nun wieder Elementare Zeilenumformungen vom Typ II durchgeführt, und zwar zuerst $T_n \coloneqq \prod_{i=1}^{n} T_{in}(-a''_{in})$. Dadurch werden alle Einträge in der letzten Spalte (außer $(a''_{nn} = 1)$ gleich 0 gestellt. Danach wird $T_{n-1} \coloneqq \prod_{i=1}^{n-1} T_{in-1}(-a'''_{in-1})$ wobei die a'''_{ij} zur Matrix A''' gehören, welche aus dem vorigen Schritt entstanden ist. Diese Matrix ist ebenfalls wieder in Dreiecksform mit den letzten beiden Spalten 0 außer den Diagonaleinträgen. Dieses Verfahren wird nun wiederholt, bis alle Einträge der entstehenden Matrix den Einträgen der Einheitsmatrix I_n entsprechen. Also $I_n = T_1 \cdot T_2 \cdots T_n \cdot S \cdot M \cdot A$. Wir erhalten also:

$$A^{-1} = T_1 \cdots T_n \cdot S \cdot M$$

Diese Matrix wird durch die gleichzeitige Anwendung (der gleichen Elementaren Zeilenumformungen wie für A) auf I_n erhalten.

Beispiel 5.9. Sei $A \coloneqq \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$. In der folgenden Beschreibung sind die römischen Zahlen die betroffenen Zeilennummern; I + 2II bedeutet Multiplikation mit

$T_{12}(2)$ von links, II $\leftrightarrow$ III bedeutet, dass Zeilen II und III vertauscht werden und $\lambda\cdot$ II bedeutet, dass $S_2(\lambda)$ verwendet wird. Man beachte: I + II $\neq$ II + I in diesem Zusammenhang. Wir betrachten nun die 3×6 Matrix $(A \mid I_3)$ und führen die notwendigen Elementaren Zeilenumformungen durch, damit A in I_3 übergeht. Das Anwenden der gleichen EZU auf I_3 liefert damit A^{-1}:

$$\left(\begin{array}{ccc|ccc} 2 & 1 & 1 & 1 & 0 & 0 \\ 1 & 2 & 1 & 0 & 1 & 0 \\ 1 & 1 & 2 & 0 & 0 & 1 \end{array}\right) \overset{I-II}{\to} \left(\begin{array}{ccc|ccc} 1 & -1 & 0 & 1 & -1 & 0 \\ 1 & 2 & 1 & 0 & 1 & 0 \\ 1 & 1 & 2 & 0 & 0 & 1 \end{array}\right) \overset{II-I,III-I}{\to}$$

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

$$\left(\begin{array}{ccc|ccc} 1 & -1 & 0 & 1 & -1 & 0 \\ 0 & 1 & -1 & 0 & 1 & -1 \\ 0 & 0 & 4 & -1 & -1 & 3 \end{array}\right) \overset{\frac{1}{4}III}{\to} \left(\begin{array}{ccc|ccc} 1 & -1 & 0 & 1 & -1 & 0 \\ 0 & 1 & -1 & 0 & 1 & -1 \\ 0 & 0 & 1 & -\frac{1}{4} & \frac{-1}{4} & \frac{3}{4} \end{array}\right) \overset{II+III}{\to}$$

$$\left(\begin{array}{ccc|ccc} 1 & -1 & 0 & 1 & -1 & 0 \\ 0 & 1 & 0 & -\frac{1}{4} & \frac{3}{4} & -\frac{1}{4} \\ 0 & 0 & 1 & -\frac{1}{4} & -\frac{1}{4} & \frac{3}{4} \end{array}\right) \overset{I+II}{\to} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{3}{4} & -\frac{1}{4} & \frac{-1}{4} \\ 0 & 1 & 0 & -\frac{1}{4} & \frac{3}{4} & -\frac{1}{4} \\ 0 & 0 & 1 & -\frac{1}{4} & -\frac{1}{4} & \frac{3}{4} \end{array}\right)$$

Also $A^{-1} = \frac{1}{4}\begin{pmatrix} 3 & -1 & -1 \\ -1 & 3 & -1 \\ -1 & -1 & 3 \end{pmatrix}$ Überprufen (nicht notwendig, aber sehr zu empfehlen):

$$A \cdot A^{-1} = \frac{1}{4}\begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix} \cdot \begin{pmatrix} 3 & -1 & -1 \\ -1 & 3 & -1 \\ -1 & -1 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Lineare Gleichungssysteme $A \in \mathrm{Mat}(m \times n, K), b \in K^n$ gegeben. Gesucht wird $x \in K^n$ so dass

$$Ax = b \qquad (*)$$

Dies gilt genau dann, wenn $f_A(x) = b$ mit $f_A \colon K^n \to K^m$ gegeben durch $f_A(x_1, .., x_n) = (\sum_{j=1}^n a_{1j}x_j, ..., \sum_{j=1}^n a_{mj}x_j)$ für $A = (a_{ij})_{\substack{i\in\{1,..,m\} \\ j\in\{1,..,n\}}}$. Zum lösen von $(*)$ geht man wie folgt vor:

1) rang A bestimmen. Dann $\ker f_A = n - \mathrm{rang}\, A$. Dies ist der Lösungsraum für das homogene System $Ax = 0$.

2) Existenz einer Lösung von $(*) \Leftrightarrow \operatorname{rang} A = \operatorname{rang}(A \mid b) \Leftrightarrow$ die Elementaren Zeilenumformungen, die A in Zeilenstufenform A' bringen, ändern b zu b' so dass b' Nulleinträge dort hat, wo die Zeilen von A' identisch Null sind.

3) Falls $m = n = \operatorname{rang} A$, so ist A invertierbar und die Lösung von $(*)$ ist eindeutig bestimmt durch $x = A^{-1}b$. Es gibt 2 Wege, diese Lösung zu bestimmen:

- **a)** A^{-1} ausrechnen, $(A \mid I_n) \overset{\text{EZU}}{\to} (I_n \mid A^{-1})$
- **b)** Direkt: $(A \mid b) \overset{\text{EZU}}{\to} (I_n \mid A^{-1}b)$

4) Das System $(*)$ für $m \neq n$ oder $\operatorname{rang} A < m$ lösen, also die Lösungsmenge L bestimmen.

- **a)** Das homogene System $Ax = 0$ lösen, d.h. $\ker f_A$ bestimmen.
- **b)** Eine Lösung x_0 von $(*)$ finden (durch Lösen "von unten nach oben" des Systems in Zeilenstufenform)
- **c)** $L = \{x_0 + x' \mid x' \in \ker f_A\}$ ist die Lösungsmenge zu $(*)$.

A) Wieso dies funktioniert:

Satz 5.21. *Sei A' eine Zeilenstufenform zu A. Dann ist* $\operatorname{rang} A' = \operatorname{rang} A$.

Beweis. Eine Elementare Zeilenumformung entspricht der Multiplikation von A mit einer invertierbaren Matrix
$M \in \{S_i(\lambda) \mid i \in \{1,..,m\}, \lambda \in K \setminus \{0\}\} \cup \{T_{ij}(\lambda) \mid i \neq j, i,j \in \{1,..,m\}, \lambda \in K\} \cup \{W_{ij} \mid 1 \leq i < j \leq m\}$. Also $\operatorname{rang} MA = \dim(f_M \circ f_A(K^n)) = \dim f_A(K^n) = \operatorname{rang} A$, da $f_M \colon K^m \to K^m$ Isomorphismus, also $\dim f_M(U) = \dim U$ für alle Untervektorräume $U \subseteq K^m$. Nach endlich vielen Elementaren Zeilenumformungen erhält man $A' = M_k \cdots M_1 \cdot A$ und $\operatorname{rang} A' = \operatorname{rang} M_{k-1} \cdots A = ... = \operatorname{rang} A$. □

Satz 5.22. $A \in \operatorname{Mat}(n \times n, K)$ *invertierbar* $\Rightarrow$

a) $(A \mid I_n) \overset{EZU}{\to} (I_n \mid B) \Leftrightarrow B = A^{-1}$

b) $(A \mid b) \overset{EZU}{\to} (I_n \mid b') \Leftrightarrow b' = A^{-1}b$.

Beweis. A invertierbar $\Leftrightarrow \operatorname{rang} A = n \Leftrightarrow$ Eine Zeilenstufenform von A ist in Dreiecksform

$$A' = \begin{pmatrix} * & * & ... & * \\ 0 & * & \ddots & \\ \vdots & 0 & \ddots & \\ 0 & & & * \end{pmatrix}$$

mit allen Einträgen auf der Diagonalen ungleich Null. Dann $A' \overset{\text{EZU}}{\to} I_n$. Das Produkt der invertierbaren Matrizen, ist also eine invertierbare Matrix $M \in \text{Mat}(n \times n, K)$, so dass $MA = I_n$. Also $M = A^{-1}$. Falls dieselben Elementaren Zeilenumformungen an I_n beziehungsweise b angewandt werden, so werden I_n beziehungsweise b mit A^{-1} multipliziert. □

Satz 5.23. *(∗) hat eine Lösung* $\Leftrightarrow \operatorname{rang} A = \operatorname{rang}(A \mid b)$.

Beweis. Sei $A' = MA, b' = Mb$ mit A' in Zeilenstufenform und $M \in \text{Mat}(m \times m, K)$ invertierbar.

$$A'x = b' \Leftrightarrow Ax = b$$

Also hat $(*)$ eine Lösung $x \in K^n \Leftrightarrow A'x = b'$ hat eben jene Lösung $x \in K^n$. Falls $A'x = b$ eine Lösung hat, und die letzten $m - \operatorname{rang} A$ Zeilen von A' sind Null, so müssen auch die letzten Einträge von b' Null sein $\Leftrightarrow \operatorname{rang} A' = \operatorname{rang}(A' \mid b')$. Es bleibt noch zu zeigen, dass falls diese Bedingung erfüllt ist, dann existiert eine Lösung zu $A'x = b'$. Dies zeigen wir weiter unten. □

B) Wie: Lösen eines Gleichungssystems in Zeilenstufenform: Man ordnet einer Matrix A' in Zeilenstufenform

$$A' = \begin{pmatrix} 0 & \dots & 0 & * & & & & \\ 0 & & & 0 & 0 & * & & \\ 0 & & & & & 0 & * & \\ & & & & & & & \\ & & & & & & & \end{pmatrix}$$

eine $r \times r$-Dreiecksmatrix ($r = \operatorname{rang} A$):

$$A'_{red} := \begin{pmatrix} * & * & \dots & * \\ 0 & * & \ddots & \vdots \\ 0 & 0 & \ddots & * \\ 0 & 0 & \dots & * \end{pmatrix}$$

deren Spalten diejenige Spalten von A' (gekürzt um die letzte $m - r$ Einträge, die so woe so Null sind) sind, die ein "Zeilenstufenanfang" enthalten, und das folgende System lösen:

$$A'_{red} \cdot x_{red} = c_{red},$$

wobei $c_{red} := b' - A'_{rest} \cdot x_{rest}$, wobei A'_{rest} die $r \times (n - r)$-Matrix, die aus den übrig gebliebenen Splten von A' entsteht (auch hier werden die letzte $m - r$ Nulleinträge aus den Spalten von A' weggelassen) und $x_{rest} \in K^{n-r}$ beliebig gewählt ist.

Beispiel 5.10.

$$\begin{pmatrix} 0 & 1 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$$

Die Matrix A ist in Zeilenstufenform, und sie hat 3 Stufen, deren Anfänge sich jeweils in der Spalte 2, 4 und 5 befinden.

Wir bestimmen A_{red}, indem wir diese 3 Spalten nehmen, in denen die ersten nicht-null Elemente der jeweiligen Zeilen stehen, also

$$A_{red} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} x_{red} = \begin{pmatrix} x_2 \\ x_4 \\ x_5 \end{pmatrix} x_{rest} = \begin{pmatrix} x_1 \\ x_3 \\ x_6 \end{pmatrix} A_{rest} = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$c = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} - \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_3 \\ x_6 \end{pmatrix} = \begin{pmatrix} 1 - x_3 - x_6 \\ 2 \\ 1 - x_6 \end{pmatrix}$$

(In diesem Beispiel brauchen wir nicht, die Spalten zu kürzen.)

Aus der Aufgabe 1, Blatt 9, folgt $A_{red}^{-1} = \begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}$, also

$$\begin{pmatrix} x_2 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 - x_3 - x_6 \\ 2 \\ 1 - x_6 \end{pmatrix} = \begin{pmatrix} -x_3 - 2x_6 \\ 1 + x_6 \\ 1 - x_6 \end{pmatrix}$$

Führt man nun die Probe durch, so sieht man

$$\begin{pmatrix} 0 & 1 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ -x_3 - 2x_6 \\ x_3 \\ 1 + x_6 \\ 1 - x_6 \\ x_6 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$$

Wir erhalten damit den Lösungsraum $L = \{(x_1, -x_3 - 2x_6, x_3, 1 + x_6, 1 - x_6, x_6) \mid x_1, x_3, x_6 \in K\} \subseteq K^6$. falls $b = 0 \in K^m$, so wird $\ker f_A$ bestimmt, also x_{rest} hat beliebige Werte und die Komponenten von x_{red} werden durch $A_{red} x_{red} = c$ bestimmt.

Bemerkung. Die Spalten in A_{rest} können durch Elementare Spaltenumformungen vom Typ II eliminiert werden: Die Spalte $A_{(j)}$ kann als Linearkombination der spalten $A_{(j_1)}, ..., A_{(j_l)}$ mit $j_l < j < j_{l+1}$ geschrieben werden.

Die Elementaren Spaltenumformungen werden analog zu den Elementaren Zeilenumformungen definiert und sie entsprechen der Multiplikation *von rechts* mit

$$\left\{ \begin{array}{ll} S_i(\lambda), & i \in \{1,..,n\}, \lambda \in K \setminus \{0\} \\ T_{ij}(\lambda), & i \neq j, i, j \in \{1,..,n\}, \lambda \in K \\ W_{ij}, & i < j, i, j \in \{1,..,n\} \end{array} \right\} \subseteq \mathrm{Mat}(n \times n, K)$$

Also $A \overset{\text{EZU, Typ II, III}}{\to} A''$ in Zeilenstufenform

$$M' \cdot A = A' \overset{\text{EZU, Typ I, II}}{\to} A''$$

in reduzierter Zeilenstufenform ($A''_{red} = I_r$)

$$M'' \cdot A = A'' \overset{\text{ESU, Typ II, III}}{\to} \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} = M''A \cdot N$$

mit $M'' \in \mathrm{Mat}(m \times m, K), N \in \mathrm{Mat}(n \times n, K)$ invertierbar.

Definition 5.15. Sei $A = (a_{ij})_{\substack{i \in \{1,..,m\} \\ j \in \{1,..,n\}}} \in \mathrm{Mat}(m \times n, K)$. Die *transponierte Matrix* ist dann $A^T \coloneqq (a_{ji})_{\substack{i \in \{1,..,m\} \\ j \in \{1,..,n\}}} \in \mathrm{Mat}(n \times m, K)$

Beispiel 5.11.

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

Satz 5.24. ***a)*** $(A + B)^T = A^T + B^T$ *für alle* $A, B \in \mathrm{Mat}(m \times n, K)$.

b) $\forall A \in \mathrm{Mat}(m \times n, K),\ \forall B \in \mathrm{Mat}(n \times p, K)$ *gilt* $(A \cdot B)^T = B^T \cdot A^T$.

Beweis. **a)** Klar.

b) Sei $C = AB$ mit $A = (a_{ij}), B = (b_{jk})$. Dann

$$c_{ik} = \sum_{j=1}^{n} a_{ij} b_{jk} \quad \forall i \in \{1,..,m\}, k \in \{1,..,p\}$$

Falls $D = B^T A^T$, so $D = (d_{ki})_{\substack{k \in \{1,..,p\} \\ i \in \{1,..,m\}}}$ so dass $d_{ki} = \sum_{j=1}^{n} b_{kj}^T a_{ji}^T$ wobei $a_{ji}^T = a_{ij}, b_{kj}^T = b_{jk}$, also $d_{ki} = c_{ik}$ und $D = C^T$.

□

Korollar 5.25. ***1)*** *Sei $A \in \mathrm{Mat}(n \times n, K)$ invertierbar. Dann ist auch A^T invertierbar.*

2) $A \overset{EZU}{\to} A' \Leftrightarrow A^T \overset{ESU}{\to} (A')^T$

Beweis. **1)** $AB = I_n \Rightarrow B^T A^T = I_n \Rightarrow A^T$ ist invertierbar und $(A^T)^{-1} = (A^{-1})^T$.

2) $S_i(\lambda)^T = S_i(\lambda), T_{ij}(\lambda)^T = T_{ji}(\lambda), W_{ij}^T = W_{ij}$ für alle i, j, λ für die diese Matrizen definiert sind. Also $(MA)^T = A^T M^T$ und falls M eier Elementaren Zeilenumformung $A' = MA$ entspricht, so entspricht M^T einer Elementaren Spaltenumformung $(A')^T = A^T M^T$. und umgekehrt.

□

Korollar 5.26. $\operatorname{rang} A = \operatorname{rang} A^T$.

Beweis. Sei $A \in \mathrm{Mat}(m \times n, K)$ und $r \coloneqq \operatorname{rang} A$. Dann existiert eine Zeilenumformung und Spaltenumformung so dass $MAN = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$ mit $M \in \mathrm{Mat}(m \times m, K), N \in \mathrm{Mat}(n \times n, K)$ invertierbar. Also $N^T A^T M^T = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$, da offenbar $I_r^T = I_r$. und N^T, M^T entsprechen Elementaren Zeilenumoformungen beziehungsweise Spaltenumoformungen für A^T. Also $\operatorname{rang} A^T = \operatorname{rang} \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} = r = \operatorname{rang} A$. □

Matrix einer linearen Abbildung Seien V, W endlich erzeugte K-Vektorräume. Dann $\exists m, n \in \mathbb{N}\ (m = \dim_K W, n = \dim_K V)$ und Isomorphismen $\psi\colon K^m \to W, \varphi\colon K^n \to V$. Falls $B_n \coloneqq \{e_1^0, ..., e_n^0\} \subseteq K^n$ und $B_m \coloneqq \{e_1^0, .., e_m^0\} \subseteq K^m$ die Standardbasen sind, so sind $B_V \coloneqq \left\{x_j \coloneqq \varphi(e_j^0) \mid j \in \{1, .., n\}\right\}$ und $B_W \coloneqq \{y_i \coloneqq \psi(e_i^0) \mid i \in \{1, .., m\}\}$ Basen von V beziehungsweise W. Umgekehrt existieren für gegebene Basen $B_V = \{x_1, .., x_n\} \subseteq V$ und $B_W = \{y_1, .., y_m\} \subseteq W$ (eindeutig bestimmte) Isomorphismen $\varphi\colon K^n \to V$ und $\psi\colon K^m \to W$ von K-Vektorräumen, so dass $\varphi(e_j^0) = x_j\ \forall j \in \{1, .., n\}$ und $\psi(e_i^0) = y_i\ \forall i \in \{1, .., m\}$. Jeder linearen Abbildung $F\colon V \to W$ und Basiswahl B_V, B_W mit zugehörigen Abbildungen φ, ψ wie oben entspricht dann eine Lineare Abbildung:

$$\psi^{-1} \circ F \circ \varphi\colon K^n \to K^m$$

und somit eine Matrix $A_{F,\varphi,\psi} \in \mathrm{Mat}(m \times n, K)$ mit $f_{A_{F,\varphi,\psi}} = \psi^{-1} \circ F \circ \varphi$. $A_{F,\varphi,\psi}$ hängt von F, aber auch von φ und ψ ab. Wir beschreiben die Abhängigkeit dieser Matrizen von den basen $\{x_1, .., x_n\} \subseteq V$ und $\{y_1, .., y_n\} \subseteq W$ (oder äquivalent von den Isomorphismen φ, ψ) für festgelegtes $F\colon V \to W$. Seien $\varphi'\colon K^n \to V, \psi'\colon K^m \to W$ Isomorphismen. Dann

$$\varphi^{-1} \circ \varphi'\colon K^n \to K^n, \quad \psi^{-1} \circ \psi'\colon K^m \to K^m$$

sind Isomorphismen, und seien $M_1 \in \mathrm{Mat}(n \times n, K)$ beziehungsweise $M_2 \in \mathrm{Mat}(m \times m, K)$ die zugehörigen invertierbaren Matrizen.

Satz 5.27. *In der Situation wie oben gilt*

$$A_{F,\varphi',\psi'} = M_2^{-1} A_{F,\varphi,\psi} M_1 \qquad (**)$$

Beweis. Wir betrachten die zu $A_{F,\varphi',\psi'}$ gehörige lineare Abbildung $f_{A_{F,\varphi',\psi'}} = (\psi')^{-1} \circ F \circ \varphi' = (\psi'^{-1} \circ \psi) \circ (\psi^{-1} \circ F \circ \varphi) \circ (\varphi^{-1} \circ \varphi') = f_{M_2}^{-1} \circ f_{A_{F,\varphi,\psi}} \circ f_{M_1}$. Hieraus folgt die Behauptung $(**)$. □

Korollar 5.28. *Seien $V \coloneqq K^n, W \coloneqq K^m$ mit Basen $\{x_1, .., x_n\} \subseteq K^n$ und $\{y_1, .., y_m\} \subseteq K^m$. Sei $M_1 \coloneqq (x_1, .., x_n)$ die $n \times n$-Matrix, deren Spalten die Vektoren $x_1, .., x_n$ sind und $M_2 \coloneqq (y_1, .., y_n)$ die $m \times m$-Matrix, die aus dem Zusammensetzen der Spaltenvektoren $y_1, .., y_m$ entsteht. Sei $f_A \colon K^n \to K^m$ die zu $A \in \mathrm{Mat}(m \times n, K)$ gehörige lineare Abbildung. Die zu $F \coloneqq f_A$ und den Basen $\{x_1, .., x_n\}, \{y_1, .., y_m\}$ assoziierte Matrix ist dann $M_2^{-1} \circ A \circ M_1$.*

Korollar 5.29. *Sei $V = W$ und $\varphi = \psi, \varphi' = \psi'$ und so weiter (also die Basis in V ist gleichzeitig die ausgewählte Basis im Zielbereich $W = V$). Sei $F \colon V \to V$ und $A, A' \in \mathrm{Mat}(n \times n, K)$ die zu F und Basis φ bzw zu F und Basis φ' assoziierten Matrizen. Sei $M \in \mathrm{Mat}(n \times n, K)$ die zu $\varphi^{-1} \circ \varphi'$ assoziierte invertierbare Matrix. Es gilt: $A' = M^{-1} \circ A \circ M$.*

Bemerkung. Die obige Matrix M hat $\varphi^{-1} \circ \varphi'(e_1^0), ..., \varphi^{-1} \circ \varphi'(e_n^0)$ als Spalten (im Fall $\varphi = \mathrm{id}_{K^n}$, also falls $V = K^n$, sind die Spalten von M gleich $\varphi'(e_1^0), ..., \varphi'(e_n^0)$). Diese Spalten bilden eine Basis in K^n, daher entspricht die invertierbare Matrix M eines *Basiswechsels* von $\{e_1^0, ..., e_n^0\}$ nach $\{\varphi^{-1} \circ \varphi'(e_1^0), ..., \varphi^{-1} \circ \varphi'(e_n^0)\}$ oder auch von $\{\varphi(e_1^0), ..., \varphi(e_n^0)\} = \{x_1, .., x_n\}$ nach $\{x_1', ..., x_n'\}$ wobei $x_i' \coloneqq \varphi'(e_i^0)\ \forall i \in \{1, .., n\}$. Die Matrix $M^{-1}AM$ entspricht also derselben linearen Abbildung $F \colon V \to V$ wie A bezuglich des Basiswechsels zu M. Eine wichtige Aufgabe der linearen Algebra ist, eine möglichst "günstige" Basis zu finden, in der der linearen Abbildung F eine möglichst einfache Matrix entspricht. (Dies ist insbesondere zum Studium von Eigenwerten und Eigenvektoren wichtig.)

Bemerkung. Falls die Basen in V und W unabhängig von einander gewählt werden dürfen, also es werden nicht nur nach einer, sondern nach zwei Basiswechseln in V und in W gesucht, dann ist die einfachste Form der Matrix $A' = M_2^{-1} A M_1$ bereits gefunden und zwar $A' = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$ mit $r \coloneqq \operatorname{rang} A$. Die Matrix M_2^{-1} entsteht aus Elementaren Zeilenumformungen und M_1 aus Elementaren Spaltenumformungen.

Bemerkung. Um ein lineares Gleichungssystem $Ax = b$ zu lösen, sind die Elementaren Zeilenumformungen bevorzugt einzusetzen, denn $MAx = Mb$ hat die gleiche

Lösungsmenge wie $Ax = b$. Dadurch kann die Matrix $A_{red} = I_r$ gesetzt werden, die Matrix A_{rest}, die aus den Spalten besteht, die kein "Stufenanfang" a_{ij_i} enthalten werden nicht geändert. Durch Elementare Spaltenumformungen kann die Matrix MA in ihre einfachste Form $MAN = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$ gebracht werden, aber die Lösungsmenge zu $Ax = b$ ist $\left\{N^{-1}y \mid (MAN)y = Mb\right\}$. Es wird also ein Basiswechsel in K^n benötigt. Hier, wie sonst so oft in der linearen Algebra, ist es wichtig, ein lineares Gleichungssystem $Ax = b$ als eine Gleichung

$$f_A(x) = b$$

zu betrachten. Falls $A' = MAN$ (mit M, N invertierbar, so $f_{A'} = f_M \circ f_A \circ f_N$ und die Korrespondenz zwischen den Lösungsmengen zu $Ax = b$ und zu $A'y = b'$ folgt direkt aus dieser Gleichung.

Bemerkung. Sei $V = W$. Für jeden Isomorphismus $\varphi\colon K^n \to V$ in K-Vektorräumen entsteht ein Isomorphismus

$$\Phi\colon \operatorname{Mat}(n \times n, K) \to \operatorname{End}_K(V) \coloneqq \hom_K(V, V), \quad \Phi(A) \coloneqq \varphi \circ f_A \circ \varphi^{-1}\colon V \to V$$

Falls eine Abbildung $\alpha\colon \operatorname{Mat}(n \times n, K) \to S$ für eine Menge S die Bedingung

$$\alpha(MAM^{-1}) = \alpha(A) \quad \forall A, M \in \operatorname{Mat}(n \times n, K)$$

mit M invertierbar erfüllt, so induziert α eine Abbildung

$$\alpha_V\colon \operatorname{End}_K(V) \to S, \quad \alpha_V(F) \coloneqq \alpha(\Phi^{-1}(F))$$

(unabhängig vom Isomorphismus $\varphi\colon K^n \to V$).

5.4 Spur einer quadratischen Matrix

Sei $A \in \operatorname{Mat}(n \times n, K), A = (a_{ij})_{i,j\in\{1,..,n\}}$. Die *Spur* der Matrix A ist

$$\operatorname{tr} A \coloneqq \sum_{i=1}^{n} a_{ii}$$

die Summe der Diagonaleinträge von A.

Satz 5.30. $\forall A \in \operatorname{Mat}(m \times n, K)\ \forall B \in \operatorname{Mat}(n \times m, K)$ *gilt* $\operatorname{tr} AB = \operatorname{tr} BA$.

Beweis. Wir bemerken zunächst, dass AB, BA in der Tat quadratische Matrizen sind, wir also die Spur für diese Produkte definiert haben. Sei $A = (a_{ij})_{\substack{i\in\{1,..,m\}\\ j\in\{1,..,n\}}}, B = (b_{jk})_{\substack{j\in\{1,..,n\}\\ k\in\{1,..,m\}}}$. Die $m \times m$ Einträge von AB sind

$$c_{ik} \coloneqq \sum_{j=1}^{n} a_{ij}b_{jk}, \quad \forall i, k \in \{1, .., m\}$$

und die $n \times n$ Einträge von BA sind

$$d_{jl} \coloneqq \sum_{k=1}^{m} b_{jk} a_{kl}, \quad \forall j, l \in \{1, .., n\}$$

Die Spuren sind

$$\operatorname{tr} AB = \sum_{i=1}^{m} c_{ii} = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} b_{ji} = \sum_{j=1}^{n} \sum_{k=1}^{m} b_{jk} a_{kj} = \sum_{j=1}^{n} d_{jj} = \operatorname{tr} BA$$

da das Vertauschen der Summenzeichen und Umbenennen des Summationsindexes erlaubt ist. □

Satz 5.31. *Seien $A, M \in \operatorname{Mat}(n \times n, K)$ und M invertierbar. Dann*

$$\operatorname{tr} A = \operatorname{tr}(MAM^{-1})$$

Beweis.

$$\operatorname{tr} A = \operatorname{tr}(M^{-1}MA) = \operatorname{tr}(MAM^{-1})$$

□

Bemerkung. $\operatorname{tr}(ABC) \neq \operatorname{tr}(BAC)$ und $\operatorname{tr}\colon \operatorname{Mat}(n \times n, K) \to K$ ist eine lineare Abbildung.

6 Determinanten

6.1 Kreuzprodukt, Flächeninhalte und Volumen

Seien $X, Y \in \mathbb{R}^2$ Vektoren. Der Flächeninhalt des Parallelogramms, welches von X und Y erzeugt wird ist linear in X und linear in Y und verschwindet, falls $X = Y$. Als Vektoren im $\mathbb{R}^3$ (wenn wir $\mathbb{R}^2 \subseteq \mathbb{R}^3$ als die Basisebene betrachter) ist $X \times Y$ ein Vertikaler Vektor im $\mathbb{R}^3$, dessen Länge gleich dem Flächeninhalt des obigen Parallelogramms ist (wobei jedoch die Richtung von der Reihenfolge von X, Y abhängt). Falls $X = (x_1, x_2), Y = (y_1, y_2)$ so ist dieser Flächeninhalt genau

$$x_1 y_2 - y_1 x_2 = \det(X, Y) = \det \begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \end{pmatrix}$$

6.2 Definition und Leibniz-Formel

Definition 6.1. Eine Determinante auf $\mathrm{Mat}(n \times n, K) (\cong K^n)^n$ ist eine Abbildung $\det\colon (K^n)^n \to K$ mit den folgenden Eigenschaften:

D1) det ist multilinear: $\forall x_1, .., x_n \in K^n, x_i' \in K^n, \lambda \in K$ gilt:

$$\det(x_1, ..., x_i + \lambda x_i', .., x_n) = \det(x_1, .., x_n) + \lambda \det(x_1, .., x_i', .., x_n)$$

D2) det ist alternierend: $\forall x_1, .., x_n \in K^n$ mit

$$\exists i \neq j\colon x_i = x_j \Rightarrow \det(x_1, .., x_n) = 0$$

D3) det ist normiert: Für die Standardbasis $\{e_1, .., e_n\}$ von K^n gilt:

$$\det(e_1, .., e_n) = 1$$

Bemerkung. Da $(K^n)^n \cong \mathrm{Mat}(n \times n, K)$ ist die Determinante ebenfalls eine Abbildung $\det\colon \mathrm{Mat}(n \times n, K) \to K$.

Satz 6.1. $\forall n \in \mathbb{N}\ \exists!\, \det\colon (K^n)^n \to K$, *die die Eigenschaft D1)-D3) erfüllt.*

Der Beweis der Existenz ist schwierig: Sie folgt entweder aus der Leibniz-Formel:

$$\det(a_{ij}) \coloneqq \sum_{\sigma \in S_n} \varepsilon(\sigma) \prod_{i=1}^{n} a_{i\sigma(i)}$$

wobei $\sigma \in S_n$ eine Permutation der Menge $\{1, .., n\}$ und $\varepsilon\colon S_n \to \{-1, 1\}$ ein Gruppenhomomorphismus ist, der noch zu definieren ist, oder sie folgt aus der Eindeutigkeit der Determinante und einer Induktion nach n. Wir skizzieren zuerst die Definition (und den zugehörigen Satz, dass ε wohldefiniert ist) der Determinante durch die Leibniz-Formel:

Definition 6.2. $\tau \in S_n$ heißt Transposition, falls sie aus der Vertauschung von $i, j \in \{1, .., n\}, i < j$ entsteht. Für alle anderen $l \in \{1, .., n\} \setminus \{i, j\}$ gilt also $\tau(l) = l$. Wir schreiben:

$$\tau_{ij}(l) = \begin{cases} i, & l = j \\ j, & l = i \\ l, & \text{sonst} \end{cases}$$

Wir schreiben weiter $\mathcal{T}_n = \{\tau_{ij} \mid i < j, i, j \in \{1, .., n\}\}$.

Satz 6.2. *Sei $\sigma \in S_n$. Dann gibt es Transpositionen $\tau_1, .., \tau_n \in \mathcal{T}_n$ so dass $\sigma = \tau_1 \circ ... \circ \tau_k$.*

Beweis. Durch Induktion nach n:

I.A. $n = 1$: Dann $\sigma = \mathrm{id}_{\{1\}}$, also ist σ Produkt von 0 Transpositionen.

I.S. $n > 1$: Falls $\sigma(n) = n$, so sei $\sigma'(l) \coloneqq \sigma(l)$ für alle $l \in \{1, .., n-1\}$. Nach Induktionsvoraussetzung gibt es dann $\tau'_1, .., \tau'_k \in \mathcal{T}_{n-1}$ mit $\sigma' = \tau'_1 \circ .. \circ \tau'_k$. Aber falls $\tau'_l = \tau_{ij} \in \mathcal{T}_{n-1}$, so sei $\tau_l \coloneqq \tau_{ij} \in \mathcal{T}_n$ und dann gilt $\sigma = \tau_1 \circ ... \circ \tau_k$. Falls $\sigma(n) \neq n$, so sei $\tau_0 \coloneqq \tau_{\sigma(n)n} \in \mathcal{T}_n$. Dann folgt $\overline{\sigma} \coloneqq \tau_0 \circ \sigma$ erfüllt $\overline{\sigma}(n) = n$, also $\overline{\sigma} = \tau_1, .., \tau_k$ wie oben, also

$$\sigma = \tau_0^{-1} \circ \tau_1 \circ ... \circ \tau_k = \tau_0 \circ ... \circ \tau_k$$

□

Wir zeigen nun, dass die Parität der Anzahl von Transpositionen, deren Produkt $\sigma \in S_n$ ist nur von σ abhängt. Nun gilt:

$$\sigma = \tau_1 \circ .. \circ \tau_k = \tau'_1 \circ .. \circ \tau'_l \Leftrightarrow \tau_1 \circ .. \circ \tau_k \circ \tau'_l \circ .. \circ \tau'_1 = \mathrm{id}_{\{1,..,n\}}$$

Also genügt es zu zeigen, dass $\tau_1 \circ ... \circ \tau_k = \mathrm{id}_{\{1,..,n\}}$ und $\tau_1, .., \tau_k \in \mathcal{T}_n \Rightarrow k$ gerade. Um dies zu zeigen, zerlegen wir zuerst jede Transposition τ_{ij} in elementare Transpositionen aus der Menge $\mathcal{T}_n^0 \coloneqq \{\tau_{i,i+1} \mid i \in \{1, .., n\}\}$. Tatsächlich gilt für alle $i < j, i, j \in \{1, .., n\}$:

$$\tau_{ij} = \tau_{i,i+1} \circ \tau_{i+1,i+2} \circ ... \circ \tau_{j-1,j} \circ \tau_{j-2,j-1} \circ ... \circ \tau_{i,i+1} \qquad (*)$$

Dies sind $2(j-i)+1$ elementare Transpositionen. Zur Überprüfung von $(*)$ merken wir, dass $\forall l.i \{i+1, .., j-1\}$ gilt: $\tau_{j-1,j} \circ .. \circ \tau_{i,i+1}(l) = l - 1$ (der erste Faktor, der l von seiner Stelle bewegt, ist $\tau_{l-1,l}$ und die nächsten bewegen $l-1$ nicht mehr und $\tau_{i,i+1} \circ .. \circ \tau_{j-2,j-1}(l-1) = l$ (der erste Faktor, der $l-1$ bewegt ist $\tau_{l-1,l}$, danach wird l nicht mehr bewegt). Dagegen wird jeder Faktor in $\tau_{j-1} \circ .. \circ \tau_{i,i+1}$ das Element

i bewegen so dass $\tau_{j-1,j} \circ .. \circ \tau_{i,i+1}(i) = j$. Analog wird j erstmal von $\tau_{j-1,j}$ bewegt, dann weiter von allen anderer Faktoren in $\tau_{i,i+1} \circ .. \circ \tau_{j-2,j-1}$, so dass

$$\tau_{i,i+1} \circ .. \circ \tau_{j-2,j-1} \circ \tau_{j-1,j} \circ \tau_{j-2,j-1} \circ .. \circ \tau_{i,i+1}(l) = \begin{cases} j, & l = i \\ i, & l = j \\ l, & \text{sonst} \end{cases}$$

was $(*)$ beweist. Wir müssen also zeigen, falls $\mathrm{id} = \tau_1, .., \tau_k$ ein Produkt von elementaren Transpositionen ist, so ist k gerade. Dazu zählen wir ab, wie oft eine Transposition $\tau_{i,i+1}$ im obigen Produkt vorkommt:
Seien also $\tau_1, .., \tau_k \in \mathcal{T}_n^0$ mit $\tau_k ... \tau_1 = \mathrm{id}_{\{1,..,n\}}$. Wir nennen die k-Tupel $Q_i = (i, \tau_1(i), \tau_2(\tau_1(i)), ..., \tau_k(...(\tau_1(i))))$ eine *Kette* für $i \in \{1, .., n\}$. Wir sagen 2 Ketten Q_i und Q_j überkreuzen sich an der Stelle l, falls

$$(\tau_{l-1} ... \tau_0)(i) = (\tau_l ... \tau_0)(j) \ \wedge \ (\tau_{l-1} ... \tau_0)(j) = (\tau_l ... \tau_0)(j)$$

wir definieren hier $\tau_0 = \mathrm{id}_{\{1,..,n\}}$. Dies passiert genau dann, wenn $\tau_l = \tau_{ab}$ wobei $a = (\tau_{l-1} ... \tau_0)(i)$ und $b = (\tau_{l-1} ... \tau_0)(j)$. An jeder Stelle zwischen 1 und k passiert genau eine Überkreuzung von Ketten. k ist also die Anzahl der Überkreuzungen. Wir zählen die Überkreuzungen pro kettenpaar, dann werden alle Überkreuzungen genau einmal gezählt. Für alle Paare $i, j \in \{1, .., n\}, i < j$ gibt es aber eine Gerade Anzahl an Überkreuzungen zwischen Q_i und Q_j, weil bei jeder Überkreuzung ändert sich das Vorzeichen von $(\tau_l ... \tau_0)(i) - (\tau_l - ... \tau_0)(j)$, und am Ende ($l = k$ wie am Anfang ($l = 0$)) ist diese Differenz positiv. Die Gesamtanzahl alle Überkreuzungen also k muss also gerade sein.

Beispiel 6.1. Sei $\sigma \in S_4$ gegeben durch

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 3 & 1 \end{pmatrix}$$

Dann gilt:

$$\tau_{14} \circ \sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 3 & 4 \end{pmatrix} = \tau_{12}$$

Also $\sigma = \tau_{14} \circ \tau_{12}$ und damit $\varepsilon(\sigma) = 1$.

Satz 6.3. *Die Abbildung* $\det^L \colon \mathrm{Mat}(n \times n, K) \to K$ *definiert durch*

$$\det{}^L A \coloneqq \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)}$$

für $A = (a_{ij})_{i,j \in \{1,..,n\}}$. *erfüllt die Bedingungen D1)-D3), ist also eine Determinante.*

Beweis. **D1)** Seien $a'_1, .., a'_k \in K$ und $A'_{(i)} = \begin{pmatrix} 0 & a'_1 & 0 \\ \vdots & \vdots & \vdots \\ 0 & a'_n & 0 \end{pmatrix}$ und $A_{(i)} \coloneqq \begin{pmatrix} 0 & a_1 & 0 \\ \vdots & \vdots & \vdots \\ 0 & a_n & 0 \end{pmatrix}$.

Es gilt dann

$$\text{D1)} \Leftrightarrow \det(A + \lambda A'_{(i)}) = \det A + \lambda \det(A - A_{(i)} + A'_{(i)})$$

Wir überprufen nun D1) an $\det^L$. Es gilt:

$$\det{}^L(A + \lambda A'_{(i)}) = \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{1\sigma(1)}..(a_{i\sigma(i)} + \lambda a'_{\sigma(i)})..a_{n\sigma(n)}$$

$$= \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{1\sigma(1)}..a_{n\sigma(n)} + \lambda \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{1\sigma(1)}..a'_{\sigma(i)}..a_{n\sigma(n)}$$
$$= \det{}^L A + \lambda \det{}^L(A - A_{(i)} + A'_{(i)})$$

D2) Sei $i < j$ und $(a_{1i}, ..., a_{ni})^t = (a_{1j}, .., a_{nj})$. Dann $a_{1\sigma(1)}...a_{n\sigma(n)} = a_{1\sigma'(1)}...a_{n\sigma'(n)}$ wobei $\sigma' \coloneqq \tau_{ij} \circ \sigma$. Sei $\varphi \colon S^n \to S^n$, $\varphi(\sigma) = \tau_{ij} \circ \sigma$. φ ist eine bijektive Abbildung und es gilt:

$$\varepsilon(\varphi(\sigma)) = -\varepsilon(\sigma)$$

Es gilt also

$$a_{1\sigma(1)}..a_{n\sigma(n)} = a_{1\varphi(\sigma)(1)}..a_{n\varphi(\sigma)(n)} \qquad (**)$$

Sei $S_n^+ \coloneqq \{\sigma \in S_n \mid \varepsilon(\sigma) = 1\}$ und $S_n^- \coloneqq \{\sigma \in S_n \mid \varepsilon(\sigma) = -1\}$. Dann ist $\varphi_+ \colon S_n^+ \to S_n^-$ wobei $\varphi_+(\sigma) = \varphi(\sigma) \forall \sigma \in S_n^+$ also bijektiv. Also:

$$\sum_{\sigma \in S_n} \varepsilon(\sigma) a_{1\sigma(1)}..a_{n\sigma(n)} = \sum_{\sigma \in S_n^+} \varepsilon(\sigma) a_{1\sigma(1)}..a_{n\sigma(n)} + \sum_{\sigma \in S_n^+} \varepsilon(\varphi_+(\sigma)) a_{1\varphi(\sigma)(1)}..a_{n\varphi(\sigma)(n)}$$
$$\overset{(**)}{=} \sum_{\sigma \in S_n^+} \varepsilon(\sigma) a_{1\sigma(1)}..a_{n\sigma(n)} - \sum_{\sigma \in S_n^+} \varepsilon(\sigma) a1\sigma(1)..a_{n\sigma(n)} = 0$$

D3) Falls $A = I_n$, so sind die einzigen Einträge von A, die nicht Null sind auf der Diagonale. Also ist das Produkt $a_{1\sigma(1)}..a_{n\sigma(n)} \neq 0 \Leftrightarrow \sigma = \mathrm{id}_{\{1,..,n\}}$. Es folgt:

$$\det{}^L(I_n) = 1 \cdots 1 = 1$$

□

Dieser Satz beweist die Existenz einer Determinante. Um die Eindeutigkeit zu beweisen, werden wir die Eigenschaften der Determinante untersuchen. Diese Eigenschaften werden künftig auch die konkreten Berechnungsmethoden für Determinanten implizieren.

6.3 Eigenschaften und Eindeutigkeit der Determinante

Satz 6.4 (S1)**.** *Seien* $\det\colon (K^n)^n \to K$ *eine Determinante (eine Abbildung die D1)-D3) erfüllt). Dann gilt:*

a) $a_1, .., a_n$ *linear abhängig* $\Rightarrow \det(a_1, .., a_n) = 0$.

b) $\det(.., a_i, .., a_j, ..) = -\det(.., a_j, .., a_i, ..)$. *Durch Vertauschen zweier Vektoren ändert sich also das Vorzeichen der Determinante.*

Beweis. **a)** Sei $a_n = \sum_{j=1}^{n} \lambda_j a_j$. Dann gilt:

$$\det(a_1, .., a_n) \overset{\text{D1)}}{=} \sum_{j=1}^{n-1} \lambda_j (a_1, .., a_{n-1}, \lambda_j a_j) \overset{\text{D2)}}{=} 0$$

b) Wir wissen $\det(..a_i + a_j, .., a_j + a_i..) = 0$. Aus D1) folgt:

$$\begin{aligned} 0 &= \det(...a_i + a_j, .., a_j + a_i..) = \det(..a_i, .., a_j + a_i..) + \det(..a_j, .., a_j + a_i..) \\ &= \det(..a_i, .., a_j..) + \det(..a_i, .., a_j..) + \det(..a_j, .., a_j..) + \det(..a_j, .., a_i..) \end{aligned}$$

Also $\det(..a_i, .., a_j..) = -\det(..a_j, .., a_i..)$.

□

Satz 6.5 (S)**.** *Seien* $S_j(\lambda), T_{ij}(\lambda), W_{ij}$ *die Elementarmatrizen, die Elementare Spalten-/Zeilenumformungen für* $n \times n$*-Matrizen definieren. Dann gilt:*

a) $\det(A \cdot S_j(\lambda)) = \lambda \det A$

b) $det(A \cdot T_{ij}(\lambda)) = \det A$

c) $\det(A \cdot W_{ij}) = -\det A$

Beweis. Seien $A_1, .., A_n$ die Spalten der Matrix A. Es gilt $A \cdot S_j(\lambda) = (A_1, .., \lambda A_j, .., A_n)$, also

$$\det(A \cdot S_j(\lambda)) \overset{\text{D1)}}{=} \lambda \det A \Rightarrow a)$$

Weiter gilt für $i < j$: $A \cdot T_{ij}(\lambda) = (A_1, .., A_i + \lambda A_j, .., A_n)$, also

$$\det(A \cdot T_{ij}) \overset{\text{D1)}}{=} \det A + \lambda \det(A_1, .., A_j, .., A_j, .., A_n) \overset{\text{D2)}}{=} \det A \Rightarrow b)$$

Für $i < j$ gilt weiter:

$$\det(A \cdot W_{ij}) = \det(A_1, .., A_j, .., A_i, .., A_n) = -\det(A_1, .., A_i, .., A_j, .., A_n) = -\det A \Rightarrow c)$$

□

Insbesondere folgt für $A = I_n$:

a') $\det(S_j(\lambda)) = \lambda$ für $\lambda \neq 0$.

b') $\det(T_{ij}(\lambda)) = 1$ für $i \neq j$

c') $\det(W_{ij}) = -1$ für $i < j$

Lemma 6.6. *Sei $A \in \mathrm{Mat}(n \times n, K)$. Dann gilt:*

a) *A invertierbar $\Leftrightarrow \operatorname{rang} A = n \Leftrightarrow A = (A_1, .., A_n)$ mit $A_1, .., A_n \in K^n$ linear unabhängig.*

b) *A invertierbar $\Leftrightarrow$ A ist ein Produkt von Elementarmatrizen.*

Beweis. **a)** Bekannt: Die Spalten von A sind die Vektoren $f_A(e_1^0), .., f_A(e_n^0)$, also Bilder der Standardbasis von K^n unter f_A.

b) Folgt aus den elementaren Zeilenumformungen, die A in I_n überführen. $A^{-1} = E_1 \cdots E_k \Leftrightarrow A = E_k^{-1} \cdots E_1^{-1}$ und die Inversen der Elementarmatrizen sind selbst wieder Elementarmatrizen.

□

Satz 6.7. *Es gibt genau eine Determinante* $\det\colon (K^n)^n \to K$.

Beweis. Seien det und $\widetilde{\det}$ Determinanten. Für beide folgt $\det(A) = \widetilde{\det}(A)$ für A Elementarmatrix. Aus dem Satz (S) folgt $\det(AE) = \widetilde{\det}(AE)$ für alle $A \in \mathrm{Mat}(n \times n, K)$ und E Elementarmatrix. Sei A invertierbar in $\mathrm{Mat}(n \times n, K)$. Dann ist $A = E_1 \cdots E_k$ mit $E_1, .., E_k$ Elementarmatrizen. Also:

$$\begin{aligned}&\det(E_1 \cdots E_k) \overset{(S)}{=} \det(E_1 \cdots E_{k-1}) \det(E_k) \overset{(S)}{=} \ldots \overset{(S)}{=} \det(E_1) \cdots \det(E_k)\\ &= \widetilde{\det}(E_1) \cdots \widetilde{\det}(E_k) \overset{\text{analog}}{=} \widetilde{\det}(E_1 \cdots E_k)\end{aligned}$$

Falls A nicht invertierbar ist, so ist $\operatorname{rang} A < n$ und die Spalten von A sind linear abhängig. Aus Satz (S1)(a) folgt $\det(A) = \widetilde{\det}(A) = 0$. □

Der obige Satz beweist die Eindeutigkeit der Determinante. Der Beweis der Existenz (ohne die angeführte Darstellung nach Leibniz) folgt aus den Eigenschaften, die gleichzeitig Methoden zur Berechnung der Determinante bestimmen:

Satz 6.8 (Rechenregeln)**.** *Sei K ein Körper und* $\det\colon (K^n)^n \to K$ *eine normierte, alternierende Multilinearform auf K^n. Seien $A, B \in \mathrm{Mat}(n \times n, K)$. Dann gilt:*

a) $\det(\lambda A) = \lambda^n \det(A)$

b) $\det(AB) = \det A \det B$

c) *A invertierbar* $\Leftrightarrow \det A \neq 0$

d) $\det A^T = \det A$

Bemerkung. Es gilt *nicht*: $\det(\lambda A) = \lambda \det(A)$ oder $\det(A+B) = \det(A) + \det(B)$ für $n > 1$.

Beweis. Sei $A = (A_1, .., A_n)$.

a) $\det(\lambda A) = \det(\lambda A_1, .., \lambda A_n) \overset{\text{D1)}}{=} \lambda \det(A_1, \lambda A_2, .., \lambda A_n) \overset{\text{D1)}}{=} .. \overset{\text{D1)}}{=} \lambda^n \det(A_1, .., A_n) = \lambda^n \det(A)$

b) Folgt aus dem Beweis der Eindeutigkeit für A, B invertierbar. Falls A oder B nicht invertierbar, so auch AB nicht invertierbar, also $\det(AB) = 0 = \det A \cdot \det B$.

c) Wie oben.

d) Wird erstmal für Elementarmatrizen bewiesen: $S_j(\lambda)^T = S_j(\lambda), T_{ij}(\lambda)^T = T_{ji}(\lambda), W_{ij}^T = W_{ij}$, und es gilt $\det E = \det E^T$ für E Elementarmatrix. Aus b) folgt also

$$\begin{aligned}&\det(E_1 \cdots E_k) = \det(E_1) \cdots \det(E_k) = \det(E_k^T) \cdots \det(E_1^T)\\ &= \det(E_k^T \cdots E_1^T) = \det((E_1 \cdots E_k)^T)\end{aligned}$$

da $(E_1 \cdots E_k)^T = E_k^T \cdots E_1^T$. Falls A nicht invertierbar ist, so auch A^T nicht, da $\operatorname{rang} A = \operatorname{rang} A^T$ und damit $\det A = 0 = \det A^T$.

□

Korollar 6.9. *Seien $A, B \in \mathrm{Mat}(n \times n, K)$ und B invertierbar ($B \in \mathrm{Gl}(n, K)$). Dann gilt:*

a) $\det(B^{-1}) = (\det(B))^{-1}$

b) $\det(BAB^{-1}) = \det A$

Beweis. $I_n = BB^{-1} \Rightarrow 1 = \det B \cdot \det B^{-1} \Rightarrow$ a). b) folgt direkt aus a) mit der Produktregel. □

Satz 6.10. *Sei $f \in \mathrm{End}(V)$ und V ein endlich erzeugter K-Vektorraum. dann $\det(A_{f,\varphi})$ ist von der Basis $\varphi\colon K^n \to V$ ($n = \dim_K V$) unabhängig.*

Beweis. Für zwei Basen $\varphi, \varphi'\colon K^n \to V$ ist $\varphi^{-1} \circ \varphi' = f_M$ mit $M \in \mathrm{Gl}(n, K)$ und es gilt $A_{f,\varphi'} = M^{-1} \circ A_{f,\varphi} \circ M$ und der Satz folgt mit b) aus dem Korollar. □

6.4 Entwicklungssatz von Laplace. Komplementärmatrix

Definition 6.3. Sei $A \in \mathrm{Mat}(n \times n, K)$, $A = (a_{ij})_{i,j\in\{1,..,n\}}$ eine Matrix. Die $(n-1) \times (n-1)$-Matrix $A_{(ij)}$ entsteht aus A durch Streichen der i-ten Zeile und j-ten Spalte.

Lemma 6.11. *Sei $A = (a_{ij})_{i,j\in\{1,..,n\}} \lambda \mathrm{Mat}(n \times n, K)$, so dass für $j-$te Spalte A_j gilt: $A_j = e_i^0$. Dann gilt:*

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

Beweis. Sei zunächst $i = j = 1$. Dann gilt:

$$A = \begin{pmatrix} 1 & a_{12} & ... & a_{1n} \\ 0 & & & \\ & & A_{(11)} & \\ 0 & & & \end{pmatrix}$$

Seien $B_2, .., B_n \in K^{n-1}$ die Spalten von $A_{(11)}$. Dann $\det(A) = \det \begin{pmatrix} 1 & 0 & ... & 0 \\ 0 & & & \\ \vdots & b_2 & & b_n \\ 0 & & & \end{pmatrix}$ ist eine Abbildung $\widetilde{\det}(b_2, .., b_n) \coloneqq \det \begin{pmatrix} 1 & 0 & ... & 0 \\ 0 & & & \\ \vdots & b_2 & & b_n \\ 0 & & & \end{pmatrix}$, $\widetilde{\det}\colon (K^{n-1})^{n-1} \to K$, die die Bedingung D1)-D3) erfüllt, also $\det A = \det A_{(11)}$. Seien nun $i, j \in \{1, .., n\}$ beliebig. Dann gilt: $A = \begin{pmatrix} a_{11} & 0 & a_{1n} \\ & \vdots & \\ & 1 & \\ & \vdots & \\ a_{n1} & 0 & a_{nn} \end{pmatrix}$ wobei die ausgezeichnete 1 an (i, j)-ter Stelle sei. Dann gilt:

$$\det A = \det \begin{pmatrix} 0 & a_{11} & a_{1n} \\ \vdots & & \\ 1 & & \\ \vdots & & \\ 0 & a_{n1} & a_{nn} \end{pmatrix} \cdot (-1)^{j-1}$$

da hierfür $j-1$ Spaltenvertauschungen $W_{12}\cdots W_{j-1j}$ benötigt werden. Analog folgt:

$$\det\begin{pmatrix} 0 & a_{11} & & a_{1n} \\ \vdots & & & \\ 1 & & & \\ \vdots & & & \\ 0 & a_{n1} & & a_{nn} \end{pmatrix} = (-1)^{i-1}\det\begin{pmatrix} 1 & a_{i1} & & a_{in} \\ 0 & a_{11} & & a_{1n} \\ \vdots & & & \\ 0 & a_{n1} & & a_{nn} \end{pmatrix}$$

Es werden hierfür nämlich $i-1$ Zeilenvertauschungen benötigt. Es folgt:

$$\det A = (-1)^{i-1+j-1}\det\begin{pmatrix} 1 & a_{i1} & & a_{in} \\ 0 & & & \\ \vdots & & A_{(ij)} & \\ 0 & & & \end{pmatrix} = (-1)^{i+j}\det A_{(ij)}$$

□

Satz 6.12 (Laplace-Entwicklung)**.** *Sei* $A = (a_{ij})_{i,j\in\{1,..,n\}} \in \mathrm{Mat}(n\times n, K)$*. Sei* $j \in \{1,..,n\}$*. Dann gilt:*

a)

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

Dies ist die Entwicklung nach der j*-ten Spalte.*

b)

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

Dies ist die Entwicklung nach der j*-ten Zeile.*

Beweis. **a)** Folgt aus dem Lemma, D1) und $A_j = \sum_{i=1}^{n} a_{ij}e_i^0$

b) Folgt aus a), $\det A^T = \det A$, $A_{(ji)} = (A^T_{(ij)})^T$ wobei a_{ji} der (i,j)-te Eintrag von A^T ist.

□

Korollar 6.13. *Sei* $A = (a_{ij})_{i,j\in\{1,..,n\}} \in \mathrm{Mat}(n\times n, K)$ *und* $a_{ij} \neq 0 \Rightarrow j \geq i$ *(A obere Dreieckmatrix). Dann* $\det A = a_{11}\cdots a_{nn}$ *ist das Produkt der Diagonaleinträge.*

Definition 6.4 (Komplementärmatrix)**.** Sei $A = (a_{ij})_{i,j\in\{1,..,n\}}$. Dann können wir definieren $\widetilde{A} \coloneqq ((-1)^{i+j}\det(A_{(ji)}))_{i,j\in\{1,..,n\}} \in \mathrm{Mat}(n\times n, K)$. Also $\widetilde{a}_{ij} = (-1)^{i+j}\det(A_{(ji)}))$ (man beachte, dass i, j hier vertauscht werden). Dann heißt $\widetilde{A}$ die Komplementärmatrix zu A.

Satz 6.14. $A \cdot \widetilde{A} = \widetilde{A} \cdot A = \det A \cdot I_n$

Beweis. $A \cdot \widetilde{A} = (b_{ik})_{i,k\in\{1,..n\}}$ wobei

$$b_{ik} = \sum_{j=1}^{n} a_{ij}\widetilde{a_{jk}} = \sum_{j=1}^{n} a_{ij}(-1)^{i+k}\det A_{(kj)}$$

Falls $i = k$ so $b_{ii} = \det A$ nach Laplace. Falls $i \neq k$, so

$$b_{ik} = \det \begin{pmatrix} a_{11} & ... & ... & a_{1n} \\ & & & \\ a_{i1} & & & a_{in} \\ & & & \\ a_{i1} & & & a_{in} \\ & & & \\ a_{n1} & & & a_{nn} \end{pmatrix} = 0$$

Die k-te Zeile von A wird also durch die i-te Zeile ersetzt. Es folgt $A \cdot \widetilde{A} = \det A \cdot I_n$. Die andere Gleichung lässt sich analog durch Spaltenentwicklung zeigen. □

Der obige Satz liefert für invertierbare Matrizen A eine direkte Formel $A^{-1} = \frac{1}{\det A}\widetilde{A}$.
Die Determinante liefert ebenfalls eine Formel für die eindeutige Lösung eines linearen Gleichungssystems $Ax = b$ mit A invertierbar.

Satz 6.15. *Sei $A = (A_1, .., A_n) \in \mathrm{Mat}(n \times n, K)$ invertierbar und $b \in K^n$. Dann ist $x = (x_1, .., x_n)^T$ gegeben durch*

$$x_i \coloneqq \frac{\det(A_1, .., \overbrace{b}^{i}, .., A_n}{\det A}$$

die eindeutige Lösung zu $Ax = b$.

Beweis. Mit Laplace angewandt auf die i-te Spalte folgt: $\det(A_1, .., b, .., A_n) = \sum_{j=1}^{n}(-1)^{i+j}b_j A_{(ji)} = \sum_{j=1}^{n} b_j \cdot \widetilde{a}_{ij} = \det A \cdot (A^{-1}b)_i$ ist die i-te Komponente des Vektors $\det A \cdot (A^{-1}b)$. Es folgt:

$$Ax = \frac{1}{\det A} \cdot A((A^{-1}b) \cdot \det A) = b$$

□

7 Eigenwerte

Definition 7.1. Sei $A \in \mathrm{Mat}(n \times n, K)$. Ein Element $\lambda \in K$ heißt *Eigenwert* von A, falls $\exists x \in K^n \setminus \{0\} : Ax = \lambda x$. x heißt dann *Eigenvektor* zum Eigenwert λ.

Bemerkung. Eigenwerte können 0 sein, Eigenvektoren jedoch nicht.

7.1 Charakteristisches Polynom

Satz 7.1. *$\lambda \in K$ ist Eigenwert zu $A \in \mathrm{Mat}(n \times n, K) \Leftrightarrow \det(A - \lambda I_n) = 0$.*

Beweis. $\exists x \neq 0 \colon Ax = \lambda x \Leftrightarrow \exists x \neq 0 \colon (A - \lambda I_n)x = 0 \Leftrightarrow A - \lambda I_n$ nicht invertierbar $\Leftrightarrow \det(A - \lambda I_n) = 0$. □

Satz 7.2. *Es gibt ein Polynom $\chi_A \in K[X]$, so dass $\chi_A(\lambda) = \det(A - \lambda I_n)\ \forall \lambda \in K$. χ_A heißt* Charakteristisches Polynom *von A.*

Beweis. Sei $\delta_{ij} \coloneqq 1$ falls $i = j$ und $\delta_{ij} = 0$ sonst. Dann $I_n = (\delta_{ij})_{i,j \in \{1,..,n\}}$. Die Formel von Leibniz liefert:

$$\det(A - \lambda I_n) = \det((a_{ij}) - \lambda(\delta_{ij})) = \sum_{\sigma \in S_n} \varepsilon(\sigma)(a_{1\sigma(1)} - \lambda\delta_{1\sigma(1)}) \cdots (a_{n\sigma(n)} - \lambda\delta_{n\sigma(n)})$$

Dies ist eine Summe und jeder Term ist ein Produkt von n Faktoren vom Typ $a' - Xb'$. Setze

$$\chi_A(X) \coloneqq \sum_{\sigma \in S_n} \varepsilon(\sigma)(a_{1\sigma(1)} - X\delta_{1\sigma(1)}) \cdots (a_{n\sigma(n)} - X\delta_{n\sigma(n)}) \qquad (*)$$

Dann $\chi_A(X) \in K[X]$ mit $\mathrm{grad}(\chi_A(X)) \leq n$ und $\chi_A(\lambda) = \det(A - \lambda I_n)$. □

Satz 7.3. *Der Leitkoeffizient von χ_A ist $(-1)^n$, der Koeffizient von X^{n-1} ist $(-1)^{n-1} \operatorname{tr} A$ und der konstante Term von χ_A ist $\det A$.*

Beweis. Ein Term in X^n kann nur in den Termen in $(*)$ vorkommen, wo alle $\delta_{1\sigma(1)} \cdots \delta_{n\sigma(n)}$ ungleich Null sind. Dies ist für genau für die Permutation $\sigma = \mathrm{id}_{\{1,..,n\}}$ der Fall. Es ist $\varepsilon(\mathrm{id}) = 1$ und der entsprechende Term ist:

$$(a_{11} - X) \cdots (a_{nn} - X) \qquad (**)$$

und der Koeffizient von X^n ist $(-1)^n$. Ein Term in X^{n-1} kann nur in einem Term $\varepsilon(\sigma)(a_{1\sigma(1)} - X\delta_{1\sigma(1)}) \cdots (a_{n\sigma(n)} - X\delta_{n\sigma(n)})$ in $(*)$ vorkommen, falls mindestens $n-1$ der Koeffizienten von X in den obigen Faktoren nicht Null sind. Also $\sigma(i) = i$ für mindestens $n - 1$ Elemente $i \in \{1, .., n\}$. Weil σ bijektiv ist, folgt $\sigma = \mathrm{id}_{\{1,..,n\}}$. Die Terme in X^{n-1} kommen also alle vom Term $(**)$, und sind die Produkte von einem

konstanten Term aus einem Faktor $a_{ii} - X$ und $(n-1)$-mal $-X$ (aus den restlichen Faktoren $a_{jj} - X, j \neq i$). Es folgt, dass die Summe dieser Terme

$$(-X)^{n-1}(a_{11} + .. + a_{nn}) = (-1)^{n-1} \operatorname{tr}(A) X^{n-1}$$

ist. Der konstante Term von $\chi_A(X)$ ist $\chi_A(0) = \det A$. □

Satz 7.4. *Ein Element $\lambda \in K$ ist ein Eigenwert der $n \times n$ Matrix $A \Leftrightarrow \lambda$ ist Nullstelle des charakteristischen Polynoms χ_A.*

Beweis. Klar. □

$\chi_A \in K[X]$ lässt sich in irreduzible Faktoren zerlegen. Zwei solche Faktoren, die nicht teilerfremnd sind unterscheiden sich um einen Faktor in $K \setminus \{0\}$. O. B. d. A. können wir schreiben: $\chi_A = a f_1^{k_1} \cdots f_m^{k_m}$ mit $m, k_1, .., k_m \in \mathbb{N} \setminus \{0\}$ und $f_1, .., f_m \in K[X]$ irreduzibel, $f_i, .., f_j$ teilerfremnd und $a \in K \setminus \{0\}$. Falls λ Eigenwert von A ist, so $X - \lambda \mid \chi_A(X)$. Also können wir o. B. d. A. annehmen, dass $\exists i \in \{1, ., m\} : f_i = X - \lambda$.

7.2 Vielfachkeiten eines Eigenwerts

Definition 7.2. Die algebraische Vielfachheit des Eigenwertes $\lambda \in K$ von A ist $a_\lambda \in \mathbb{N} \setminus \{0\}$ so dass $(X - \lambda)^{a_\lambda} \mid \chi_A(X)$ und $(X - \lambda)^{a_\lambda + 1} \nmid \chi_A(X)$. Falls $f_i = X - \lambda$, so ist $\alpha_\lambda = k_i$ die algebraische Vielfachheit von λ. Falls $\alpha_\lambda = 1$, so heißt λ einfacher Eigenwert von A.

Satz 7.5 (Fundamentalsatz der Algebra)**.** *Jedes Polynom $f \in \mathbb{C}[X]$ hat genau $n = \operatorname{grad} f$ Nullstellen in $\mathbb{C}$ (gerechnet mit Vielfachheiten). Das heißt $\forall f \in \mathbb{C}[X]$ ist $f = a f_1^{k_1} \cdots f_m^{k_m}$ und $f_i(X) = X - \lambda_i$ mit $\lambda_i \in \mathbb{C}$ sowie $k_1 + .. + k_m = n$.*

Satz 7.6. ***a)*** *Sei $A \in \operatorname{Mat}(n \times n, K)$. Die Summe der algebraischen Vielfachheiten der Eigenwerte ist $\leq n$.*

b) *Ist $K = \mathbb{C}$, so entspricht die Summe der algebraischen Vielfachheiten genau n.*

Beispiel 7.1. Die Matrix $A = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \in \operatorname{Mat}(2 \times 2, \mathbb{R})$ hat keine *reellen* Eigenwerte, denn $\chi_A(\lambda) = \det \begin{pmatrix} 1-\lambda & 1 \\ -1 & 1-\lambda \end{pmatrix} = (1-\lambda)^2 + 1 = \lambda^2 - 2\lambda + 2 > 0 \; \forall \lambda \in \mathbb{R}$. Betrachtet man aber A also Matrix in $\operatorname{Mat}(2 \times 2, \mathbb{C})$, so hat sie die Eigenwerte $1 + i$ und $1 - i$, denn

$$\chi_A(1 \pm i) = (1 - (1 \pm i))^2 + 1 = (\pm i)^2 + 1 = -1 + 1 = 0$$

Beide Eigenwerte sind hier einfach.

Definition 7.3. Sei λ Eigenwert von $A \in \mathrm{Mat}(n \times n, K)$. Die geometrische Vielfachheit g_λ von λ ist

$$g_\lambda \coloneqq \dim_K(\ker(A - \lambda I_n))$$

Der Untervektorraum $\ker(A - \lambda I_n)$ heißt *Eigenraum* zum Eigenwert λ.

Satz 7.7. $g_\lambda \leq a_\lambda$

Beweis. Zu zeigen ist $(X - \lambda)^{g_\lambda} \mid \chi_A(X)$. Sei $k \coloneqq g_\lambda$ und $b_1, .., b_k$ eine Basis von $\ker(A - \lambda I_n)$. Wir ergänzen $b_1, .., b_k$ zu einer Basis $b_1, .., b_n$. Es gilt $f_A(b_1) = \lambda b_1, .., f_A(b_k) = \lambda b_k$. Das heißt falls $M \coloneqq (b_1, .., b_n) \in \mathrm{Gl}(n, K)$, so gilt

$$M^{-1}AM = \left(\begin{array}{ccc|c} \lambda & & & \\ & \lambda & & B \\ & & \ddots & \\ & & & \lambda \\ \hline & 0 & & C \\ & & & \end{array}\right)$$

Mit $C \in \mathrm{Mat}(n - k \times n - k, K)$ und $B \in \mathrm{Mat}(k \times n - k, K)$. Dann

$$\det(A - \mu I_n) = \det(M^{-1}AM - \mu I_n) \overset{\text{Üb } 11.3}{=} \det((\lambda - \mu) I_k) \cdot \det(C - \mu I_{n-k})$$

Es folgt $\chi_A(X) = \chi_{\lambda I_k}(X) \cdot \chi_C(X) = (\lambda - X)^k \cdot \chi_C(X)$, also $k \leq a_\lambda$. □

Bemerkung. Das Konzept der Eigenwerte und Eigenvektoren bezieht sich auf die assoziierte lineare Abbildung f_A zu A. Man kann ebenso gut Konzepte der Eigenwerte und Eigenvektoren für Endomorphismen für (endlich erzeugte) Vektorräume definieren. Sei $f \colon V \to V$ linear. $\lambda \in K$ heißt Eigenwert von f, falls $\exists x \in V \backslash \{0\} : f(v) = \lambda v$. Das charakteristische Polynom von f kann durch $\chi_f(X) \coloneqq \chi_{A_{f,\varphi}}(X)$ definiert werden, wobei $A_{f,\varphi} \in \mathrm{Mat}(n \times n, K)$ die zu f und zur Basis $\varphi \colon K^n \overset{\cong}{\to} V$ assoziietre Matrix, wobei $f_{A_{f,\varphi}} = \varphi^{-1} \circ f_A \circ \varphi$. Es gilt $\chi_f(\lambda) = \det(f - \lambda \,\mathrm{id}_V)$, wobei die Determinante des Endomorphismus $f - \lambda \,\mathrm{id}_V$ in jeder Basis von V definiert werden kann, denn sie hängt nicht von der gewählten Basis ab. Ebenfalls ist $g_\lambda = \dim_K(\ker(f - \lambda \,\mathrm{id})$ und a_λ wird wie für Matrizen definiert.

Fazit: Eigenwerte, g_λ, a_λ und das Charakteristische Polynom einer Linearen Abbildung hängen von keiner Basis ab. Für eine Matrix $A \in \mathrm{Mat}(n \times n, K)$ sind ihre Eigenwerte, g_λ, a_λ gleich deren der Matrix $M^{-1}AM$ mit $M \in \mathrm{Gl}(n, K)$.
Wir untersuchen weiter den Fall, wo die Matrix eine Basis aus Eigenvektoren besitzt.

7.3 Diagonalisierbarkeit

Definition 7.4. a) $A \in \mathrm{Mat}(n \times n, K)$ heißt *diagonalisierbar* $\Leftrightarrow \exists M \in \mathrm{Gl}(n, K)\colon M^{-1}AM = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$

b) $A \in \mathrm{Mat}(n \times n, K)$ heißt *trigonalisierbar* $\Leftrightarrow \exists M \in \mathrm{Gl}(n, K)\colon M^{-1}AM = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$

Bemerkung. In beiden Fällen sind $\lambda_1, .., \lambda_n$ die Eigenwerte von A, wobei die $\lambda_1, .., \lambda_n$ nicht paarweise verschieden sein müssen.

Satz 7.8. *$A \in \mathrm{Mat}(n \times n, K)$ hat eine Basis aus Eigenvektoren $\Leftrightarrow$ A ist diagonalisierbar.*

Beweis. "$\Rightarrow$" Sei $\{b_1, .., b_n\}$ eine Basis aus Eigenvektoren zu den Eigenwerten $\lambda_1, .., \lambda_n$
Dann $Ab_i = \lambda_i b_i \Leftrightarrow$

$$M^{-1}AM = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

mit $M = (b_1, .., b_n)$.

"$\Leftarrow$" Falls $M^{-1}AM = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$, so sind die Spalten von M Eigenvektoren von A. $AM = M \cdot \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$, also falls $M = (b_1, .., b_n)$, so $Ab_i = \lambda_i b_i$ und $\{b_1, .., b_n\}$ bildet eine Basis, da M invertierbar ist.

□

Satz 7.9. *$A \in \mathrm{Mat}(n \times n, K)$ ist diagonalisierbar $\Leftrightarrow$ die Summe der geometrischen Vielfachheiten aller Eigenwerte ist gleich n. Dann gilt insbesondere: χ_A zerfällt in Linearfaktoren und $g_\lambda = a_\lambda \; \forall \lambda$ Eigenwerte von A.*

Beweis. "⇒" Die Diagonalform $M^{-1}AM = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$ können wir o. B. d. A. als

$$\begin{pmatrix} \lambda_1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & \lambda_1 & & & & & & \\ & & & \lambda_2 & & & & & \\ & & & & \ddots & & & & \\ & & & & & \lambda_2 & & & \\ & & & & & & \ddots & & \\ & & & & & & & \lambda_l & \\ & & & & & & & & \ddots \\ & & & & & & & & & \lambda_l \end{pmatrix}$$

umschreiben mit den paarweise verschiedenen Eigenwerte $\lambda_1, .., \lambda_l$ von A. Es folgt $k_i = g_{\lambda_i} = a_{\lambda_i}\ \forall i \in \{1, .., l\}$ und $\chi_A(X) = \prod_{i=1}^{l}(\lambda_i - X)^{k_i}$.

"⇐" Seien $\lambda_1, .., \lambda_l$ die paarweise verschiedenen Eigenwerte von A und nehmen wir $g_{\lambda_1} + .. + g_{\lambda_l} = n$ an. Da $a_{\lambda_i} \geq g_{\lambda_i}$ für alle $i \in \{1, .., l\}$ gilt $a_{\lambda_1} = g_{\lambda_1}\ \forall i \in \{1, .., l\}$ und $a_{\lambda_1} + .. + a_{\lambda_l} = n$. Dies impliziert $\chi_A(X) = \prod_{i=1}^{l}(\lambda_i - X)^{a_{\lambda_i}}$. Sei nun $k_i \coloneqq a_{\lambda_i}$ und sei $b_1, .., b_{k_1}$ eine Basis von $\ker(A - \lambda_1 I_n)$, $b_{k_1+1}, .., b_{k_1+k_2}$ eine Basis von $\ker(A - \lambda_2 A_n)$ und so weiter. Wir zeigen $b_1, .., b_n$ sind linear unabhängig. Seien $\alpha_1, .., \alpha_n \in K$ mit $\sum_{i=1}^{n} \alpha_i b_i = 0$. Wir setzen nun alle Vektoren der einzelnen Eigenräume zusammen:

$$\begin{aligned} v_1 &\coloneqq \sum_{i=1}^{k_1} \alpha_i b_i \in \mathrm{ER}(\lambda_1) \\ v_2 &\coloneqq \sum_{i=k_1+1}^{k_1+k_2} \alpha_i b_i \in \mathrm{ER}(\lambda_2) \\ &\vdots \\ v_l &\coloneqq \sum_{i=k_1+..+k_{l-1}+1}^{n} \alpha_i b_i \in \mathrm{ER}(\lambda_l) \end{aligned}$$

Wir haben $v_1 + ... + v_l = 0$. Wir wenden $A - \lambda_l I_n$ an und erhalten

$$(\lambda_1 - \lambda_l)v_1 + .. + (\lambda_{l-1} - \lambda_l)v_{l-1} = 0$$

und ersetzen $w_i := (\lambda_i - \lambda_l)v_i \in \mathrm{ER}(\lambda_i)\ \forall i \in \{1,..,l-1\}$. Es gitl $v_i = 0 \Leftrightarrow w_i = 0$ weil $\lambda_i \neq \lambda_l$ für $i \in \{1,..,l-1\}$. Wenden wir nun $A - \lambda_{l-1}I_n$ an, so erhalten wir:

$$(\lambda_1 - \lambda_{l-1})w_1 + .. + (\lambda_{l-2} - \lambda_{l-1})v_{l-2} = 0$$

Dies führen wir für alle $i \in \{1,..,l\}$ durch und kriegen damit $v_1 = 0$, dann $v_2 = 0, ..., v_l = 0$. Aber v_i ist eine Linearkombination von $b_{k_1+..k_{i-1}+1}, .., b_{k_1+..+k_i}$ von $\mathrm{ER}(\lambda_i)$. Dann $v_i = 0\ \forall i \in \{1,..,l\} \Rightarrow \alpha_j = 0\ \forall j \in \{1,..,n\}$. Also ist $b_1, ., b_n$ eine Basis von K^n, die aus Eigenvektoren von A besteht. A ist also diagonalisierbar.

□

Eine notwendige Bedingen, damit eine $n \times n$ Matrix A über K diagonalisierbar ist, ist dass ihr charakteristisches Polynom genau n Nullstellen inklusive Vielfachheiten über K besitzt. Dies ist immer der Fall, wenn K algebraisch abgeschlossen ist, zum Beispiel für $K = \mathbb{C}$. Eine zweite Bedingung ist, dass die geometrische und algebraische Vielfachheiten übereinstimmen. Dies ist aber automatisch der Fall, falls alle Eigenwerte einfach sind: $a_\lambda = 1\ \forall \lambda$ Eigenwert von A.

Satz 7.10. *Falls $\chi_A(X)$ genau n verschiedene Nullstellen besitzt, so ist $A \in \mathrm{Mat}(n \times n, K)$ diagonalisierbar. In diesem Fall sind die Eigenvektoren bis auf einen Faktor aus $K \setminus \{0\}$ eindeutig bestimmt.*

Beispiel 7.2. $A = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$ besitzt die Eigenwerte 2 und 1 (obere Dreiecksmatrix). $a_2 = 1, a_1 = 2, g_2 = 1$, aber $g_1 = \dim(\ker(A - I_3)$ mit $A - I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$, also $(x_1, x_2, x_3)^T \in \ker(A - I_3) \Leftrightarrow x_1 = 0\ \wedge\ x_3 = 0$, also $\ker(A - I_3) = \{(0, 0, x) \in K^3\}$ und $g_1 = 1 < a_1$. Also ist A nicht diagonalisierbar. Wir sehen also, dass selbst dann $g_\lambda < a_\lambda$ möglich ist, wenn χ_A in Linearfaktoren zerfällt.

Satz 7.11. *Jede symmetrische Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{R})$ ist diagonalisierbar (über $\mathbb{R}$).*

Beweis. **Teil 1:** χ_A zerfällt über $\mathbb{R}[X]$ in Linearfaktoren. Sei $\chi_A(X) \in \mathbb{R}[X] \subseteq \mathbb{C}[X]$ und seien $\lambda_1, .., \lambda_n \in \mathbb{C}$ die komplexen Nullstellen von χ_A. Wir zeigen $\chi_A(\lambda) = 0 \Rightarrow \lambda \in \mathbb{R}$. Sei λ komplexer Eigenwert von A. Dann $\exists z \in \mathbb{C}^n \setminus \{0\}$ mit (1) $Az = \lambda z$, also $\overline{Az} = \overline{\lambda}\overline{z} \Leftrightarrow$ (2) $A\overline{z} = \overline{\lambda}\overline{z}$, also ist $\overline{z}$ Eigenvektor zum Eigenwert $\overline{\lambda}$. Sei $z = (z_1, .., z_n) \in \mathbb{C}^n \setminus \{0\}$. Dann ist $\overline{z}^T z = \sum_{i=1}^n \overline{z_i} z_i = \sum_{i=1}^n |z_i|^2 > 0$ (3) weil $z \neq 0$. Es gilt also

(1) $\Rightarrow \overline{z}^T A z = \lambda \overline{z}^T z = \lambda \sum_{i=1}^n |z_i|^2$.
(2) $\Rightarrow z^T A \overline{z} = \overline{\lambda} z^T \overline{z} = \lambda \sum_{i=1}^n |z_i|^2$.
Wegen $A = A^T$ gilt also $(\overline{z}^T A z)^T = z^T A^T \overline{z} = z^T A \overline{z}$, also

$$\lambda \sum_{i=1}^n |z_i|^2 = \overline{\lambda} \sum_{i=1}^n |z_i|^2$$

also wegen (3) $\lambda = \overline{\lambda}$, also $\lambda \in \mathbb{R}$. Also sind alle Eigenwerte von A reell. Das heißt $\chi_A(X)$ zerfällt über $\mathbb{R}[X]$ in Linearfaktoren.

2. Teil: A ist diagonalisierbar. Dies werden wir im nächsten Kapitel sehen.

□

8 $\mathbb{R}^n$ als Euklidischer Raum. Orthogonale Matrizen und Basen

In $\mathbb{R}^n$ können wir den Abstand zwischen $x = (x_1, .., x_n)$ und $y = (y_1, .., y_n)$ mit $x, y \in \mathbb{R}^n$ durch

$$d(x, y) := \|x - y\| := \sqrt{(x_1 - y_2)^2 + .. + (x_n - y_n)^2} \in \mathbb{R}_+$$

bestimmen, wobei die *Norm* $\|x - y\|$ des Vektors $x - y \in \mathbb{R}^n$ stets ≥ 0 ist. Gleichheit gilt $\Leftrightarrow x - y = 0 \Leftrightarrow x = y$.

Definition 8.1. Sei M eine menge und $d\colon M \times M \to \mathbb{R}_+$ eine Funktion. d heißt Abstandsfunktion (Metrik), falls gilt:

M1) $\forall x, y \in M\colon d(x, y) = d(y, x)$

M2) $\forall x, y \in M\colon d(x, y) = 0 \Leftrightarrow x = y$

M3) $\forall x, y, z \in M\colon d(x, y) + d(y, z) \geq d(x, z)$

M3) heißt Dreiecksungleichung. Man nennt dann (M, d) einen *Metrischen Raum*.

Definition 8.2. Sei V ein $\mathbb{R}$-Vektorraum. Eine Abbildung $\nu\colon V \to \mathbb{R}_+$ heißt *Norm*, falls gilt:

N1) $\forall x \in V\ \forall \lambda \in \mathbb{R}\colon \nu(\lambda x) = |\lambda|\, \nu(x)$

N2) $\forall x \in V\colon \nu(x) = 0 \Leftrightarrow x = 0$

N3) $\forall x, y \in V\colon \nu(x + y) \leq \nu(x) + \nu(y)$

(V, ν) heißt dann Normierter Raum.

Satz 8.1. *Sei (V, ν) ein normierter Vektorraum. Dann ist $d\colon V \times V \to \mathbb{R}, d(x, y) := \nu(x - y)$ eine Metrik.*

Beweis. Vielleicht Übung. □

Beispiel 8.1. $\nu\colon \mathbb{R}^n \to \mathbb{R}$, $\nu(x_1, .., x_n) := \max_{i \in \{1,..,n\}} |x_i|$ ist eine Norm.

Bemerkung. In der Definition der Norm kann auch V ein $\mathbb{C}$-Vektorraum sein und $\lambda \in \mathbb{C}$ in N1). Man sagt dann, die Norm ist kompatibel mit der Struktur von V als $\mathbb{C}$-Vektorraum.

Definition 8.3. Sei V ein $\mathbb{R}$-Vektorraum. Eine Abbildung $g\colon V \times V \to \mathbb{R}$ heißt *Skalarprodukt*, falls gilt:

S1) g ist symmetrisch: $\forall x, y \in V\colon g(x, y) = g(y, x)$

S2) g ist bilinear: $g(x+\lambda x', y) = g(x,y) + \lambda g(x', y)$ und für alle $x, y, x' \in V, \lambda \in \mathbb{R}$.

S3) g ist positiv definit: $\forall x \in V\colon g(x,x) \geq 0$ und $g(x,x) = 0 \Leftrightarrow x = 0$.

Bemerkung. Man kann für $\mathbb{C}$-Vektorräume *hermitesche* Skalarprodukte $g\colon V \times V \to \mathbb{C}$ definieren, wobei S1) $\to$ H1): $g(x,y) = \overline{g(y,x)}$ $\forall x, y \in V$ sowie S2) $\to$ H2): $g(x+\lambda x', y) = g(x,y) + \lambda g(x', y)$ $\forall x, y, x' \in V$, $\forall \lambda \in \mathbb{C}$ und H3) = S3)

Bemerkung. S1) $\wedge$ S2) $\Rightarrow g(x, y+\lambda y') = g(x,y) + \lambda g(x, y')$ $\forall \lambda \in \mathbb{R}$ $\forall x, y, y' \in V$.

Satz 8.2. *Sei g ein Skalarprodukt. Dann ist $\nu\colon V \to \mathbb{R}$, $\nu(x) \coloneqq \sqrt{g(x,x)}$ eine Norm.*

Beweis. N1) und N2) folgen direkt aus der Defintion. N3) ist die Ungleichung von Cauchy-Schwarz:

$$\sqrt{g(x,x)} + \sqrt{g(y,y)} \geq \sqrt{g(x+y, x+y)} \Leftrightarrow \sqrt{g(x,x)g(y,y)} \geq g(x,y) \qquad (4)$$

Es gibt 2 Fälle:

1) Sind x, y linear abhängig über $\mathbb{R}$, so sei o. B. d. A. $y = \lambda x$ für ein $\lambda \in \mathbb{R}$. Dann $g(y,y) = \lambda^2 g(x,x)$ und $g(x,y) = \lambda g(x,x)$ und (4) ist automatisch erfüllt, denn $|\lambda| \geq \lambda$.

2) Sind x, y linear unabhängig, so dass $\forall \lambda \in \mathbb{R}\colon x + \lambda y \neq 0$. Dann $g(x+\lambda y, x+\lambda y) > 0$. Dann $P(\lambda) \coloneqq g(x+\lambda y, x+\lambda y) = \lambda^2 g(y,y) + 2\lambda g(x,y) + g(x,x)$ ist ein reelles Polynom von Grad 2, das nur positive Werte auf $\mathbb{R}$ annimmt. Seine Diskriminante $g(x,y)^2 = g(y,y)g(x,x)$ ist daher negativ: Also $\big|g(x,y)\big| < \sqrt{g(x,x)g(y,y)} \Rightarrow$ (4). Es folgt N3) mit der Zusatzeigenschaft $\nu(x) + \nu(y) = \nu(x+y) \Leftrightarrow y = \lambda x$ für $\lambda \geq 0$ oder $x = 0$.

□

Das euklidische Skalarprodukt $g_0(x,y) \coloneqq x_1 y_1 + ... + x_n y_n$ auf $\mathbb{R}^n$ induziert damit die Norm $\nu(x) \coloneqq \|x\| \coloneqq \sqrt{x_1^2 + .. + x_n^2}$.

Bemerkung. Das Standard hermitesche Skalarprodukt auf $\mathbb{C}^n$ ist $h_0(x,y) \coloneqq \sum_{i=1}^n x_i \overline{y}_i$ und die zugehörige norm ist $\|x\| \coloneqq \sqrt{|x_1|^2 + .. + |x_n|^2}$. Sie stimmt mit der euklidischen Norm auf $\mathbb{R}^{2n} \cong \mathbb{C}^n$ überein, $(x,y) \mapsto x + iy \in \mathbb{C}^n$ $\forall x, y \in \mathbb{R}^n$, denn $z \coloneqq x + iy$ hat die Komponenten $z_j = x_j + iy_j$, also $\|z\| = \sqrt{\sum_{j=1}^n |z_j|^2} = \sqrt{\sum_{j=1}^n x_j^2 + y_j^2} = \sqrt{\sum_{j=1}^n x_j^2 + \sum_{j=1}^n y_j^2}$.

Definition 8.4. Sei (V, g) ein euklidischer Vektorraum (zum Beispiel $V = \mathbb{R}^n, g = g_0$). Eine Menge $M \subseteq V$ heißt *Orthonormalsystem*, falls gilt:

1) $\forall v, w \in M, v \neq w\colon g(v,w) = 0$. v und w heißen dann auch orthogonal oder senkrecht zu einander und man schreibt auch $v \perp w$.

2) $\forall v \in M\colon g(v,v) = 1$. Alle Vektor besitzen also Norm 1.

Satz 8.3. *Ein Orthonormalsystem ist linear unabhängig.*

Beweis. Seien $v_1, .., v_k \in M$ mit $v_i \perp v_j$ für alle $i \neq j, i, j \in \{1, .., k\}$ und $\|v_i\| = 1$ für alle $i \in \{1, .., k\}$. Seien $\alpha_1, .., \alpha_k \in \mathbb{R}\colon \sum_{i=1}^{k} \alpha_i v_i = 0$. Wir nehmen das Skalarprodukt mit v_j:

$$0 = g(0, v_j) = g(\sum_{i=1}^{k} \alpha_i v_i, v_j) = \sum_{i=1}^{k} \alpha_i g(v_i, v_j) = \alpha_i g(v_j, v_j) = \alpha_i$$

Dies gilt $\forall j \in \{1, .., k\}$, also sind alle $\alpha_j = 0$. □

Satz 8.4 (Orthonormalisierungsverfahren)**.** *Sei $\{v_1, .., v_k\}$ ein linear unabhängiges System in einem euklidischen Vektorraum (V, g). Dann existiert ein Orthonormalsystem $\{w_1, .., w_k\}$ so dass $\langle\{w_1, .., w_i\}\rangle = \langle\{v_1, .., v_i\}\rangle \; \forall i \in \{1, .., k\}$.*

Beweis. Induktion nach $k \in \mathbb{N} \setminus \{0\}$:

I.A. $k = 1$**:** Sei $w_1 \coloneqq \frac{1}{\|v_1\|} v_1$. Dies ist wohldefiniert, da $v_1 \neq 0$, also $\|v_1\| > 0$ und $\|w_1\| = 1$, also ist $\{w_1\}$ ein Orthonormalsystem, das $\langle w_1\rangle = \langle v_1\rangle$ erfüllt.

I.S. $k \geq 2$**:** Sei $\{w_1, .., w_{k-1}\}$ ein Orthonormalsystem so dass $\langle\{v_1, .., v_v\}\rangle = \langle\{w_1, .., w_i\}\rangle \; \forall i \in \{1, .., k-1\}$ und $v_k \in V \setminus \langle\{v_1, .., v_{k-1}\}\rangle$, denn $v_1, .., v_k$ sind linear unabhängig. Seien $\alpha_i \coloneqq g(v_k, w_i) \; \forall i \in \{1, .., k-1\}$. Dann ist $v \coloneqq \sum_{i=1}^{k-1} \alpha_i w_i \in \langle\{v_1, .., v_{k-1}\}\rangle$ und es gilt $v_k - v \perp w_i \; \forall i \in \{1, .., k-1\}$:

$$g(v_k - v, w_i) = \alpha_i - \sum_{j=1}^{k-1} \alpha_j g(w_j, w_i) = \alpha_i - \alpha_i = 0$$

weil $g(w_i, w_j) = \delta_{ij}$. Weiter ist $v_k - v \neq 0$ weil $v_k \notin \langle\{v_1, .., v_{k-1}\}\rangle$. Wir definieren also $w_k \coloneqq \frac{1}{\|v_k - v\|}(v_k - v)$. Dann gilt $\|w_k\| = 1$ und $g(w_k, w_i) = 0 \; \forall i \in \{1, .., k\}$, also ist $\{w_1, ., w_k\}$ ein Orthonormalsystem. Zu zeigen ist also $\langle\{w_1, .., w_k\}\rangle = \langle\{v_1, .., v_k\}\rangle$. Es gilt $w_k = \frac{1}{\|v_k - v\|} v_k - \frac{1}{\|v_k - v\|} \sum_{i=1}^{k-1} \alpha_i w_i$, also $w_k \in \langle\{w_1, .., w_{k-1}, v_k\}\rangle = \langle\{v_1, .., v_k\}\rangle$. Also ist $\{w_1, .., w_k\}$ ein linear unabhängiges System im k-Dimensionalen Vektorraum $\langle\{v_1, .., v_k\}\rangle$, also eine Basis des Vektorraums. Insbesondere also $\langle\{v_1, .., v_k\}\rangle = \langle\{w_1, .., w_k\}\rangle$.

□

Korollar 8.5. *Seien $\{v_1, .., v_k\}$ und $\{w_1, .., w_k\}$ Orthonormalsysteme in $\mathbb{R}^n$. Dann $\exists A \in O(n) \coloneqq \left\{A \in \mathrm{Gl}(n, \mathbb{R}) \mid A^{-1} = A^\top\right\}$ so dass $Av_i = w_i \; \forall i \in \{1, .., k\}$.*

Beweis. Wir vervollständigen $\{v_1, .., v_k\}$ zu einer orthonormalen Basis $\{v_1, .., v_n\}$ *von*$\mathbb{R}^n$ indem wir zunächst zu einer Basis $\left\{v_1, .., v_k, v'_{k+1}, .., v'_n\right\}$ ergänzen und dann das Orthonormalisierungsverfahren anwenden. Ebenfalls sei $\{w_1, .., w_n\}$ eine Orthonormalbasis von $\mathbb{R}^n$. Seien $M \coloneqq (v_1, .., v_n)$ und $N \coloneqq (w_1, .., w_n)$ die aus den Vektoren $v_1, .., v_n$ beziehungsweise $w_1, .., w_n$ als Spalten zusammengesetzten $n \times n$ Matrizen. M und N sind invertierbar.

Lemma 8.6. *Eine Matrix $B \in \mathrm{Mat}(n \times n, \mathbb{R})$ ist genau dann orthogonal ($B^\top B = BB^\top = I_n$), wenn ihre Spalten eine Orthonormalbasis von $\mathbb{R}^n$ bilden.*

Beweis des Lemmas. Sei $B \coloneqq (b_1, .., b_n)$. Dann $g(b_i, b_j) = b_i^\perp b_j = (B^\top B)_{ij}$. Es gilt $\{b_1, .., b_n\}$ ist eine Orthonormalbasis $\Leftrightarrow g(b_i, b_j) = \delta_{ij}\ \forall i, j \in \{1, ., n\} \Leftrightarrow B^\top B = I_n \Leftrightarrow B^\top = B^{-1}$. □

Aus dem Lemma folgt, dass M, N invertierbare Orthogonalmatrizen sind.

Lemma 8.7. *$O(n)$ ist eine Untergruppe von* $\mathrm{Gl}(n, \mathbb{R})$

Beweis des Lemmas. A orthogonal $\Leftrightarrow A^\top = A^{-1} \Leftrightarrow AA^\top = A^\top A = I_n \Leftrightarrow (A^\top)^{-1} = (A^\top)^\top \Leftrightarrow A^\top = A^{-1}$ orthogonal. Seien weiter $A, B \in O(n)$. Dann $(AB)^\top = B^\top A^\top = B^{-1}A^{-1} = (AB)^{-1}$, also $AB \in O(n)$. Weiter ist $O(n)$ offenbar $\neq \emptyset$. □

Sei nun $A \coloneqq NM^{-1} \in O(n)$. Dann gilt $AM = N$, also $A \cdot (v_1, .., v_n) = (w_1, .., w_n) \Leftrightarrow Av_i = w_i\ \forall i \in \{1, .., n\}$. □

8.1 Haupachsentransformation

Satz 8.8. *Sei $S \in \mathrm{Mat}(n \times n, \mathbb{R})$ eine symmetrische Matrix. Dann existiert eine Orthonormalbasis $\{v_1, .., v_n\} \subseteq \mathbb{R}^n$ aus Eigenvektoren von S*

Beweis. Wir haben bereits gezeigt, dass $\chi_A(x) = (\lambda_1 - X) \cdots (\lambda_n - X)$ mit $\lambda_1, .., \lambda_n \in \mathbb{R}$. Wir beweisen nun die Behautung per Induktion nach $n \in \mathbb{N} \setminus \{0\}$.

I.A. $n = 1$: $\mathrm{Mat}(1 \times 1, \mathbb{R}) \cong \mathbb{R}$ und jede Matrix ist Diagonal.

I.S. $n \geq 2$: Sei v_n ein Eigenvektor zum Eigenwert λ_n. O. B. d. A. können wir $\|v_n\| = 1$ annehmen. Sei $M_n \in O(n)$ so dass $M_n e_n^0 = v_n$. Die Existenz dieser Matrix folgt aus obigen Korollar. Sei $V \coloneqq \langle \{e_1^0, .., e_{n-1}^0\} \rangle \cong \mathbb{R}^{n-1}$. Wir zeigen $M_n^{-1} A M_n v \in V\ \forall v \in V$. Es gilt:

$$g(e_n^0, M_n^{-1} A M_n v) = (e_n^0)^\top M_n^{-1} A M_n v = (e_n^0)^\top M_n^\top A M_n v = v_n^\top A M_n v$$

wegen $M_n e_n^0 = v_n \Rightarrow v_n^\top = (e_n^0)^\top M_n^\top$. Nun gilt $Av_n = \lambda_n v_n \Leftrightarrow v_n^\top A^\top = \lambda_n v_n^\top$, also

$$\begin{aligned} g_0(e_n^0, M_n^{-1} A M_n v) &= \lambda_n v_n^\top M_n v = \lambda_n v_n^\top (M_n^{-1})^\top v \\ &= \lambda_n (M_n^{-1} v_n)^\top v = \lambda_n (e_n^0)^\top v = \lambda_n g(e_n^0, v) = 0 \end{aligned}$$

Also $M_n^{-1}AM_nv \perp e_n^0 \Leftrightarrow M_n^{-1}AM_nv \in V \ \forall v \in V$. $M_n^{-1}AM_n$ ist also eine Matrix der Form $\left(\begin{array}{c|c} A' & 0 \\ \hline 0 & \lambda_n \end{array}\right)$ mit $A' \in \mathrm{Mat}((n-1)\times(n-1), \mathbb{R})$ symmetrisch. Nach Induktionsvoraussetzung $\exists M' \in O(n-1)$ so dass $(M')^{-1}A'M' = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_{n-1} \end{pmatrix}$.

Sei $M \coloneqq \left(\begin{array}{c|c} M' & 0 \\ \hline 0 & 1 \end{array}\right) M_n \in O(n)$. Dann

$$M^{-1}AM = \left(\begin{array}{c|c} M'^{-1} & 0 \\ \hline 0 & 1 \end{array}\right)\left(\begin{array}{c|c} A' & 0 \\ \hline 0 & \lambda_n \end{array}\right)\left(\begin{array}{c|c} M' & 0 \\ \hline 0 & 1 \end{array}\right)$$

$$= \left(\begin{array}{c|c} M'^{-1}A'M' & 0 \\ \hline 0 & \lambda_n \end{array}\right) = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

□

Das Verfahren, in dem eine Symmetrische Matrix diagonalisiert wird heißt Hauptachsentransformation. (Es werden die "Hauptachsen" $v_1, .., v_n$ in $\mathbb{R}^n$ bestimmt, für die die Lineare Abbildung $f_S \colon \mathbb{R}^n \to \mathbb{R}^n$ die einfache Form $f_S(v_i) = \lambda_i v_i \ \forall i \in \{1, .., n\}$ hat.)

Bemerkung. Falls $A \in \mathrm{Mat}(n \times n, \mathbb{R})$ eine Matrix ist, die mit Hilfe einer Orthogonalmatrix M diagonalisierbar ist, so ist A symmetrisch.

Beweis. Eine Diagonalmatrix ist symmetrisch. Falls A wie oben ist, so ist $M^{-1}AM$ symmetrisch, also $M(M^{-1}AM)M^\top = A$ ist ebenfalls symmetrisch, denn $A^\top = (M(M^{-1}AM)M^\top)^\top = M(M^{-1}AM)^\top M^\top = M(M^{-1}AM)M^\top = A$. □

Der Satz über die Hauptachsentransformation zeigt also dass umgekehrt jede symmetrische Matrix eine Orthonormalbasis von Eigenvektoren besitzt.
Zusammenfassung: Es gibt 3 Fälle, in der die Diagonalisierbarkeit einer reellen Matrix A garantiert werden kann:

1) χ_A zerfällt in Linearfaktoren in $\mathbb{R}[X]$ und $a_\lambda = g_\lambda \ \forall \lambda$ Nullstelle von χ_A.

2) χ_A zerfällt in Linearfaktoren in $\mathbb{R}[X]$ mit paarweise verschiedenen Nullstellen.

3) A ist symmetrisch. Dann kann eine Orthonormalbasis von Eigenvektoren gefunden werden.