LINEARE ALGEBRA I

OLIVER C. SCHNÜRER

Zusammenfassung. Bei diesem Manuskript handelt es sich um Notizen zu einer Vorlesung Lineare Algebra I (B1) an der Universität Konstanz.

Inhaltsverzeichnis

0. Vorbemerkungen

0.1. **Lineare Algebra in der Mathematik.** Lineare Algebra I ist eine Grundvorlesung, deren sämtliche Inhalte später in allen mathematischen Gebieten ständig auftauchen werden und daher ein guter Einstieg in die Mathematik.

Die lineare Algebra beschäftigt sich mit Objekten, die linear, d. h. gerade aussehen, also wie eine Gerade oder eine Ebene. Die Realität ist aber in der Regel nicht linear. Jedoch sehen Dinge aus der Nähe häufig linear aus.

0.2. **Was kann man mit Mathematik anfangen?**
Nach `http://www.maths.monash.edu.au/`

- Understanding our world (Bewegung der Sterne, Sonnenprotuberanzen, schwarze Löcher, Wasserwellen, Windhose, Buschbrände)
- Modelling and improving systems (Verkehrsleitsysteme, Logistik für Containerschiffe, Börse, Produktion, Medizin)
- Studying the beauty of perfection (Seifenblasen, Symmetrien in Sonnenblumen oder geometrischen Mustern, Fraktale, Wassertropfen)

0.3. **Literatur.** Die meisten Bücher über lineare Algebra sind geeignet, z. B. die Bücher von U. Stammbach [6] (häufig die Grundlage dieses Skriptes), G. Fischer [1] (auch häufiger benutzt), S. Lang [3] oder F. Lorenz [4], andere Lineare Algebra-Bücher und (mit Vorsicht; auch im Skript verwandt) für Definitionen Wikipedia (`http://www.wikipedia.org/`).

Date: 11. Oktober 2018.
1991 *Mathematics Subject Classification.* 15-01.
Vielen Dank an Dieter Hoffmann, Matthias Makowski, Reinhard Racke, Olaf Schnürer und an die Hörer der Vorlesung für Korrekturen. In den Wintersemestern 2010/11 und 2018/19 benutzt.

1. Grundlagen: Logik und Mengenlehre

Über Logik bzw. Mengenlehre gibt es ganze Vorlesungen. Wir vermitteln hier nur einige Grundlagen und folgen einem naiven Zugang.

1.1. **Logische Grundlagen.** Eine Definition führt eine Bezeichnung ein.

Definition 1.1.1 (Naive Definition einer Aussage)**.**

(i) Eine **Aussage** ist etwas, dem entweder der Wahrheitswert „wahr“ oder der Wahrheitswert „falsch“ zugeordnet ist. Wir müssen den Wahrheitswert einer Aussage nicht kennen.

(ii) Eine **Aussageform** ist eine Aussage, die eine noch nicht bestimmte oder **freie** Variable enthält. Ihr Wahrheitswert ist möglicherweise nicht zu bestimmen, solange die Variable unbekannt ist.

Für die folgenden Beispiele greifen wir auf Schulwissen zurück.

Beispiele 1.1.2.

(a) Beispiele für Aussagen sind:
  (i) 6 ist durch 3 teilbar.
  (ii) Es regnet.
  (iii) Es gibt unendlich viele Primzahlzwillinge.
  (Ich kenne den Wahrheitswert dieser Aussage momentan nicht.)

(b) Beispiele für Aussageformen sind:
  (i) x ist durch 3 teilbar. (unbekannter Wahrheitswert)
  (ii) Es gilt stets $x = y$ oder $x \neq y$. (ist stets wahr)

Definition 1.1.3 (Negation, Verneinung)**.** Sei p eine Aussage, so bezeichnen wir mit $\neg p$ die Negation dieser Aussage, als Wahrheitstabelle:

p	$\neg p$
w	f
f	w

Dabei steht „w“ für „wahr“ und „f“ steht für „falsch“.

Wir sagen für $\neg p$: „Es stimmt **nicht,** dass p gilt.“oder „p ist falsch.“

Bemerkung 1.1.4.

(i) Sei p die Aussage „91 ist eine Primzahl“, so ist $\neg p$ die Aussage „Es stimmt nicht, dass 91 eine Primzahl ist“ oder kurz: „91 ist keine Primzahl“.

(ii) Anhand der Wahrheitstabelle überlegt man sich leicht, dass p und $\neg\neg p$ Aussagen mit den gleichen Wahrheitswerten sind.

Warnung: Die doppelte Negation wird im Alltag aber gelegentlich anders aufgefasst. Beispiel: „Magst du nicht den Müll runter bringen?“ – “Nein.“

Ist der Antwortende ein Mathematiker oder Logiker, so bringt er den Müll gerne nach unten.

Wir verwenden die folgenden Verknüpfungen von zwei Aussagen zu einer neuen Aussage.

Definition 1.1.5. Seien p, q zwei Aussagen. Dann schreiben wir

- $p \wedge q$ für die Konjunktion, in Worten „p **und** q“,
- $p \vee q$ für die Disjunktion, in Worten „p **oder** q“,
- $p \dot{\vee} q$ (kaum verwendetes Symbol) für die Kontravalenz oder aussschließende Disjunktion, in Worten „**entweder** p **oder** q“,
- $p \Longrightarrow q$ für die Implikation, in Worten „wenn p, dann q“ (oder „p **impliziert** q“), und

- $p \iff q$ für die Äquivalenz, in Worten „p und q sind **äquivalent**“ oder „Genau dann, wenn p gilt, gilt auch q“,

und definieren die Wahrheitswerte durch die folgende Tabelle:

p	q	$p \wedge q$	$p \vee q$	$p \dot{\vee} q$	$p \implies q$	$p \iff q$
w	w	w	w	f	w	w
w	f	f	w	w	f	f
f	w	f	w	w	w	f
f	f	f	f	f	w	w

Bemerkung 1.1.6.

(i) Die beiden Aussagen in einer Implikation müssen nicht kausal zusammenhängen.
(ii) In der Aussage $p \implies q$ heißt p Voraussetzung oder Prämisse und q Behauptung oder Konklusion.
(iii) Gilt $p \implies q$, so heißt p eine hinreichende Bedingung für q und q eine notwendige Bedingung für p.
(iv) Wir verwenden $p \implies q$ und $q \impliedby p$ gleichbedeutend.
(v) Wir sagen, dass die Aussagen $p_1, p_2, \ldots, p_n$ (oder sogar einer ganzen Familie von Aussagen, was später erklärt wird) äquivalent sind, wenn für je zwei Aussagen p und q davon $p \iff q$ gilt.
(vi) „Wenn 6 eine Primzahl ist, dann gilt $0 = 1$“ und „Wenn 6 eine Primzahl ist, dann ist 2 eine Primzahl“ sind zwei wahre Aussagen.

Die logischen Elementarverknüpfungen $\neg$, $\vee$, $\wedge$, $\implies$ und $\iff$ erfüllen

Proposition 1.1.7. *Seien p, q, r Aussagen. Sei t (true) eine stets wahre Aussage. Dann gelten*

(i) $p \wedge t \iff p$, *(Wahres mit und)*
(ii) $p \vee t \iff t$, *(Wahres mit oder)*
(iii) $\neg\neg p \iff p$, *(Doppelte Verneinung)*
(iv) $p \vee \neg p$, ... *(tertium non datur)*
(v) $p \wedge q \iff q \wedge p$, *(Symmetrie)*
(vi) $p \vee q \iff q \vee p$, *(Symmetrie)*
(vii) $(p \iff q) \iff (q \iff p)$, *(Symmetrie)*
(viii) $p \wedge p \iff p$, ... *(Idempotenz)*
(ix) $p \vee p \iff p$, *(Idempotenz)*
(x) $p \wedge q \implies p$, *(Weglassen)*
(xi) $p \implies p \vee q$, *(Hinzufügen)*
(xii) $(p \iff q) \implies ((p \vee r) \iff (q \vee r))$, ... *(beidseitiges Hinzufügen)*
(xiii) $(p \iff q) \implies ((p \wedge r) \iff (q \wedge r))$, *(beidseitiges Hinzufügen)*
(xiv) $(p \iff q) \implies ((p \iff r) \iff (q \iff r))$, *(beidseitiges Hinzufügen)*
(xv) $(p \wedge q) \wedge r \iff p \wedge (q \wedge r)$, *(Assoziativität)*
(xvi) $p \vee (q \vee r) \iff (p \vee q) \vee r$, ... *(Assoziativität)*
(xvii) $p \vee (q \wedge r) \iff (p \vee q) \wedge (p \vee r)$, *(Distributivität)*
(xviii) $p \wedge (q \vee r) \iff (p \wedge q) \vee (p \wedge r)$, *(Distributivität)*
(xix) $\neg(p \wedge q) \iff \neg p \vee \neg q$, *(De Morgan)*
(xx) $\neg(p \vee q) \iff \neg p \wedge \neg q$, ... *(De Morgan)*
(xxi) $(p \iff q) \iff ((p \implies q) \wedge (q \implies p))$, *(beide Richtungen zeigen)*
(xxii) $((p \iff q) \wedge (q \iff r)) \implies (p \iff r)$, *(Zwischenschritt)*
(xxiii) $((p \implies q) \wedge (q \implies r)) \implies (p \implies r)$, *(Zwischenschritt)*
(xxiv) $(p \implies q) \iff (\neg p \vee q)$, ... *(Implikation auflösen)*
(xxv) $(p \implies q) \iff (\neg q \implies \neg p)$, *(Kontraposition)*
(xxvi) $p \iff ((p \wedge r) \vee (p \wedge \neg r))$. *(Fallunterscheidung)*

Genauer sollte man $(p \wedge p) \iff p$ bzw. $((\neg p) \vee q)$ schreiben, d. h. Klammern verwenden um die Reihenfolge anzudeuten.

Aus der Assoziativität folgt, dass wir hier Klammern weglassen dürfen, also $p \wedge q \wedge r$ statt $(p \wedge q) \wedge r$ schreiben dürfen. Für mehr als drei Aussage ist noch nachzuweisen, dass wir die Klammern weglassen dürfen.

Wir schreiben $p \iff q \iff r$ für $(p \iff q) \wedge (q \iff r)$.

Beweis. Die ersten Behauptungen erhält man mit Hilfe von Wahrheitstafeln, man zeigt also beispielsweise, dass in der Wahrheitstafel für $p \implies (p \vee p)$ unter dieser Aussage stets ein „wahr“ steht. Später kann man die Behauptungen auch durch Kombinationen erhalten. Wir zeigen nur, wie dies im Fall (xxv) funktioniert und lassen den Rest als Übung:

Es gilt

$$\begin{aligned}
(p \implies q) &\iff \neg p \vee q && \text{(xxiv)} \\
&\iff \neg p \vee \neg\neg q && \text{(iii), (vi) und (xii)} \\
&\iff \neg\neg q \vee \neg p && \text{(vi)} \\
&\iff \neg q \implies \neg p && \text{(xxiv).}
\end{aligned}$$

Somit ist $(p \implies q) \iff (\neg q \implies \neg p)$ aufgrund mehrfacher Anwendung von (xxii). □

Bemerkung 1.1.8. ⋆ Es gibt noch zahlreiche weitere interessante Verknüfungen von logischen Aussagen. Einige Beispiele sind:

(i) $((p \implies q) \wedge p) \implies q$, (modus ponens)
(ii) $(p \iff q) \implies (p \implies q)$, (Abschwächung)
(iii) $((p \implies r) \wedge (q \implies r) \wedge (p \vee q)) \implies r$, (beidseitige Abschwächung)
(iv) $((p \vee q) \wedge (\neg p)) \implies q$, ... (ausgeschlossene Variante)
(v) $((p \implies q) \wedge \neg q) \implies \neg p$, (modus tollens)
(vi) $(\neg(p \wedge q) \wedge p) \implies \neg q$, (ausgeschlossene Variante)
(vii) $((p \implies q) \wedge (r \implies s) \wedge (\neg q \vee \neg s)) \implies (\neg p \vee \neg r)$, (falsche Schlussfolgerungen)
(viii) $(p \implies q) \iff (p \implies (p \wedge q))$. ..... (Wiederholen der Prämisse)

Die Benennung der Aussagen hier und in Proposition 1.1.7 ist teilweise keine Standardbenennung.

Hinweis: Beim Beweis kann es sinnvoll sein, eine stets falsche Aussage f einzuführen, beispielsweise $\neg t$.

Beweis. Übung. □

1.2. **Naive Mengenlehre.** Wir benutzen hier auch [2].

Definition 1.2.1 (Naive Definition einer Menge)**.** Eine **Menge** ist eine Zusammenfassung von Objekten, den sogenannten **Elementen** dieser Menge. Ist A eine Menge und x ein Objekt, so schreiben wir die Aussage $x \in A$, falls x ein Element dieser Menge ist und $x \notin A$ für $\neg(x \in A)$.

Ist A eine Menge, die aus den Elementen a, b, c besteht, so schreiben wir $A = \{a, b, c\}$. Dabei ist die Reihenfolge der aufgeführten Elemente, oder ob wir Elemente mehrfach aufführen, unerheblich. Eine analoge Schreibweise verwenden wir auch für Mengen mit einer anderen Anzahl (noch nicht definiert) von Elementen. Wir schreiben $A = \{x_1, x_2, \ldots, x_{20}\}$ oder $A = \{x_1, x_2, \ldots\}$, falls klar ist, welche Elemente gemeint sind.

Die obige Definition erklärt nicht wirklich, was Mengen sind. Wir wollen Mengen später von einer Menge, der leeren Menge, ausgehend, beschreiben.

Definition 1.2.2 (Teilmenge, Gleichheit)**.** Seien A, B zwei Mengen.

(i) Dann ist A eine **Teilmenge** von B, notiert als $A \subset B$ oder $A \subseteq B$, falls für alle $x \in A$ auch $x \in B$ gilt.
(ii) Dann heißen A und B **gleich,** abgekürzt mit $A = B$, wenn $A \subset B$ und $B \subset A$ gelten. (Extensionalitätsaxiom)
Gilt für zwei Mengen nicht $A = B$ so schreiben wir $A \neq B$.
(iii) Gelten $A \subset B$ und $A \neq B$ so kürzen wir das mit $A \subsetneq B$ ab.

Bemerkung 1.2.3.
(i) Wir folgen einer weit verbreiteten Konvention, dass wir in Definitionen wie bei der Teilmenge nur „falls“ schreiben, obwohl wir hier eine Äquivalenz im Sinn haben.
(ii) Die Teilmengendefinition ist umgangssprachlich. Wir werden später stets die Variante aus Bemerkung 1.3.3 mit Quantoren benutzen.

Lemma 1.2.4. *Seien A, B, C Mengen. Dann gelten*
(i) $A \subset A$, *(Reflexivität)*
(ii) $x \in A$ *und* $A \subset B$ *impliziert* $x \in B$.
(iii) $A \subset B \subset C$ *impliziert* $A \subset C$, *(Transitivität)*

Beweis.
(i) Zu zeigen ist, dass für alle $x \in A$ auch $x \in A$ gilt. Das ist klar.
(ii) Gilt nach Definition von $A \subset B$.
(iii) Sei $x \in A$. Dann folgt $x \in B$ nach (ii). Nochmals nach (ii) folgt auch $x \in C$. Also gilt $A \subset C$. □

Axiom 1.2.5 (Aussonderungsaxiom)**.** Sei A eine Menge und $a(x)$ eine Aussageform. Dann gibt es eine Menge B, deren Elemente **genau** die Elemente x aus A sind, die $a(x)$ erfüllen (Aussonderungsaxiom). Wir schreiben

$$B = \{x \in A \colon a(x)\}.$$

Bemerkung 1.2.6.
(i) Mit „genau“ in dieser Definition beschreiben wir, dass B alle Elemente $x \in A$ enthält, die $a(x)$ erfüllen, aber keine, für die $\neg a(x)$ gilt.
(ii) Da wir die Menge B hier definieren, schreiben wir auch $B := \{x \in A \colon a(x)\}$ und sagen, dass die Menge B als die Menge auf der rechten Seite definiert ist.
(iii) Es ist klar, dass $B \subset A$ gilt.

Bemerkung 1.2.7. Zu jeder Menge A gibt es eine Menge B und eine Aussageform $p(x)$, so dass

$$A = \{x \in B \colon p(x)\}$$

gilt: Man nehme $B = A$ und für $p(x)$ die Aussageform $x \in B$.

Diese Bemerkung erlaubt es, mit Aussage(forme)n statt mit Mengen zu arbeiten.

Bemerkung 1.2.8 (Russelsche Antinomie)**.** ⋆ Naiverweise könnte man sich im Aussonderungsaxiom die Angabe sparen, aus welcher Menge man die Elemente aussondert, die eine gegebene Aussage erfüllen, würde also die Existenz einer „Allmenge“ annehmen. Dann hätte man zu jeder Aussage eine Menge, in der alle Objekte sind, die diese Aussage erfüllen. Dies führt jedoch zu einem Widerspruch:

Angenommen, es gäbe eine Allmenge A, also eine Menge, die alle Elemente enthält. Dann definieren wir

$$B := \{x \in A \colon x \notin x\}.$$

Nach Definition ist nun $y \in B \iff (y \in A \wedge y \notin y)$. Da wir angenommen haben, dass A eine Allmenge ist, ist $y \in A$ stets erfüllt und wir erhalten $y \in B \iff y \notin y$. Wenn wir nun überprüfen wollen, ob $B \in B$ gilt, können wir diese Bedingung mit $y = B$ anwenden und erhalten $B \in B \iff B \notin B$. Dies ist ein Widerspruch.

Daher werden wir zukünftig Allmengen vermeiden.

Es ist üblich, in einem Widerspruchsbeweis die Annahme, die zu einem Widerspruch geführt werden soll, im Irrealis zu schreiben, bei den weiteren Folgerungen daraus aber wieder den Realis zu verwenden.

Aus der Existenz irgendeiner Menge, was wir annehmen wollen, folgt auch die Existenz der leeren Menge.

Lemma 1.2.9 (Existenz der leeren Menge)**.** *Es gibt eine Menge $\emptyset$, die leere Menge, die kein Element enthält. Sie erfüllt*

(i) $\emptyset \subset A$ für alle Mengen A und
(ii) $\emptyset$ ist eindeutig bestimmt.

Beweis. Sei B eine beliebige Menge. Definiere

$$\emptyset := \{x \in B \colon x \neq x\}.$$

Aufgrund des Aussonderungsaxioms ist $\emptyset$ ein Menge.

(i) Sei A eine beliebige Menge. Wir müssen nachweisen, dass $\emptyset \subset A$ gilt, dass also $x \in \emptyset \implies x \in A$ gilt. Dazu zeigen wir, dass die Aussage $x \in \emptyset$ stets falsch ist. Sei $x \in \emptyset$ ein beliebiges Element, x ist also so gewählt, dass die Aussage $x \in \emptyset$ gilt. Aus der Aussage $x \in \emptyset$ folgt $x \neq x$. Dies ist stets falsch. Somit muss auch die Voraussetzung $x \in \emptyset$, unabhängig von x, falsch gewesen sein. Dies wollten wir zeigen.
(ii) Seien $\emptyset_1$ und $\emptyset_2$ zwei leere Mengen. Dann folgen nach (i) $\emptyset_1 \subset \emptyset_2$ und $\emptyset_2 \subset \emptyset_1$. Somit gilt nach Definition $\emptyset_1 = \emptyset_2$. □

1.3. **Quantoren.**

Definition 1.3.1 (Quantoren)**.** Quantoren verwenden wir bei freien Variablen in Aussageformen. Sie beziehen sich stets auf Elemente einer Menge:

Sei A eine Menge und $a(x)$ eine Aussageform mit der freien Variablen x.

(i) **Existenzquantor:** Wir schreiben

$$\exists\, x \in A \colon a(x) \quad \text{oder} \quad \exists_{x \in A}\, a(x)$$

für die folgende Aussage: „**Es gibt** ein x in der Menge A, so dass für dieses x die Aussage $a(x)$ gilt."

Wir schreiben

$$\exists!\, x \in A \colon a(x),$$

falls es ein x in der Menge A gibt, so dass für dieses x und kein anderes Element aus A die Aussage $a(x)$ gilt. Man sagt: Es gibt **genau** ein x aus A mit $a(x)$.

(ii) **Allquantor:** Wir schreiben

$$\forall\, x \in A \colon a(x) \quad \text{oder} \quad \forall_{x \in A}\, a(x)$$

oder auch $a(x)\ \forall x \in A$ für die Aussage: „**Für alle** x in der Menge A gilt die Aussage $a(x)$."

⋆ Es gibt auch die Bezeichnungen

(i) $\bigvee$ für den Existenzquantor und
(ii) $\bigwedge$ für den Allquantor.

Beispiele 1.3.2. Wir verwenden für diese Beispiele nochmals Schulwissen.

(i) Für alle reellen Zahlen x gilt $x^2 \geq 0$.
(ii) Es gibt eine reelle Zahl x mit $x^2 = 2$.
(iii) Falsch ist aber: $\exists!\, x \in \mathbb{R} \colon x^2 = 2$.

Bemerkung 1.3.3.

(i) Mit Quantoren können wir die Teilmengendefinition formaler angeben:

$$A \subset B \quad :\Longleftrightarrow \quad \forall\, x \in A : x \in B.$$

⋆ Dabei schreiben wir $:\Longleftrightarrow$ um anzudeuten, dass der Ausdruck auf der linken Seite eine Abkürzung für die Aussage auf der rechten Seite ist.

(ii) In einem Beweis zeigen wir die Aussage $\forall\, x \in A\colon a(x)$ wie folgt:
Sei $x \in A$ beliebig. ... Damit folgt $a(x)$.

(iii) Die Aussage $\exists\, x \in A\colon a(x)$ zeigt man beispielsweise, indem man ein x konkret hinschreibt und nachweist, dass $a(x)$ gilt.
Alternativ zeigt man $\{x \in A\colon a(x)\} \neq \emptyset$.

(iv) Die Aussage $\exists!\, x \in A\colon a(x)$ zeigt man, indem man $\exists\, x \in A\colon a(x)$ zeigt und dass für alle $x, y \in \{z \in A\colon a(z)\}$, wobei wir $x, y \in \{\ldots\}$ für $x \in \{\ldots\}$ und $y \in \{\ldots\}$ schreiben, die Aussage $x = y$ folgt.

Wir wollen die folgenden Quantorenregeln verwenden.

Axiom 1.3.4. Seien A, B Mengen und $p(x)$ bzw. $p(x, y)$ Aussageformen. Dann gelten

$$\text{(1.1)} \qquad \underset{x\in A}{\forall}\, \underset{y\in B}{\forall} p(x,y) \quad \Longleftrightarrow \quad \underset{y\in B}{\forall}\, \underset{x\in A}{\forall} p(x,y),$$

$$\text{(1.2)} \qquad \underset{x\in A}{\exists}\, \underset{y\in B}{\exists} p(x,y) \quad \Longleftrightarrow \quad \underset{y\in B}{\exists}\, \underset{x\in A}{\exists} p(x,y),$$

$$\text{(1.3)} \qquad \underset{x\in A}{\exists}\, \underset{y\in B}{\forall} p(x,y) \quad \Longrightarrow \quad \underset{y\in B}{\forall}\, \underset{x\in A}{\exists} p(x,y),$$

$$\text{(1.4)} \qquad \neg\left(\underset{x\in A}{\forall} p(x)\right) \quad \Longleftrightarrow \quad \underset{x\in A}{\exists} \neg p(x),$$

$$\text{(1.5)} \qquad \neg\left(\underset{x\in A}{\exists} p(x)\right) \quad \Longleftrightarrow \quad \underset{x\in A}{\forall} \neg p(x).$$

Bemerkung 1.3.5.

(i) Die beiden letzten Äquivalenzen heißen De Morgansche Regeln.

(ii) In (1.3) in Axiom 1.3.4 handelt es sich nicht um eine Äquivalenz: Ist $A = \mathbb{R}$ (die hier noch nicht definierten reellen Zahlen) und $p(x, y)$ die Aussage $x + y = 0$, so ist die rechte Seite wahr, die linke aber nicht.

Wir beenden diesen Abschnitt mit einem Lemma. Dabei geben wir nur einen knappen Beweis für einen Teil der Aussage.

- Ergänzen Sie die fehlenden Details!
- Welche Aussagen mit Quantoren werden hier eigentlich gezeigt?
- Warum folgt daraus die Behauptung?
- Beweisen Sie den noch fehlenden Rest!
- Die Einführung in das mathematische Arbeiten (EmA) eignet sich vorzüglich, um zu erlernen, wie man selbst solche Beweise eigenständig findet und formal korrekt aufschreibt.

Lemma 1.3.6. *Seien A, B Mengen. Dann sind die folgenden Aussagen äquivalent:*

(i) $A \subset B$,
(ii) $A \cup B = B$ *und*
(iii) $A \cap B = A$.

Beweis. Wir zeigen nur (ii) $\Longleftrightarrow$ (iii).

- „(ii)$\Longrightarrow$(iii)“: Es gilt stets $A \cap B \subset A$. Daher zeigen wir nur $A \subset A \cap B$: Sei $x \in A$ beliebig. Es folgt $x \in A \cup B$, also nach (ii) auch $x \in B$. Daher ist $x \in A \cap B$ wie behauptet.

- „(iii)$\Longrightarrow$(ii)“: Wiederum gilt $B \subset A \cup B$ stets. Es genügt also, $A \cup B \subset B$ zu zeigen: Sei $x \in A \cup B$ beliebig. Ist $x \in B$, so sind wir fertig. Sonst gilt $x \in A$. Nach (iii) folgt daraus $x \in A \cap B$. Somit gilt $x \in B$. Die Behauptung folgt. □

1.4. **Weitere Mengenlehre.**

Axiom 1.4.1 (Axiom der Existenz einer Obermenge)**.** Sei $\mathcal{M}$ eine Menge von Mengen. Dann gibt es eine Menge M mit der Eigenschaft

$$A \in \mathcal{M} \implies A \subset M.$$

Sie heißt Obermenge.

⋆ Die Obermenge ist nicht eindeutig bestimmt. Nehmen wir ein weiteres Element hinzu, so erhalten wir eine weitere Obermenge.

Definition 1.4.2 (Vereinigung und Durchschnitt)**.** Seien A, B Mengen und X eine beliebige Obermenge von A und B. Wir definieren

(i) die **Vereinigung** von A und B, mit $A \cup B$ bezeichnet, durch

$$A \cup B := \{x \in X \colon x \in A \vee x \in B\}.$$

(ii) den **Durchschnitt** (oder Schnitt) von A und B, in Zeichen $A \cap B$, durch

$$A \cap B := \{x \in X \colon x \in A \wedge x \in B\}.$$

Sei nun $\mathcal{M}$ eine Menge von Mengen und sei X eine gemeinsame Obermenge, gelte also $A \subset X$ für alle $A \in \mathcal{M}$. Dann definieren wir

(i) die Vereinigung der Mengen aus $\mathcal{M}$ durch

$$\bigcup_{A \in \mathcal{M}} A := \{x \in X \colon (\exists\, A \in \mathcal{M} \colon x \in A)\}.$$

(ii) den Schnitt (Durchschnitt oder Schnittmenge) der Mengen aus $\mathcal{M}$ durch

$$\bigcap_{A \in \mathcal{M}} A := \{x \in X \colon (\forall\, A \in \mathcal{M} \colon x \in A)\}.$$

Bemerkung 1.4.3.

(i) Enthält $\mathcal{M}$ genau zwei Mengen, so stimmen die beiden Definitionen überein.

(ii) Enthält $\mathcal{M}$ keine Menge, so gilt nach Definition der Quantoren $\bigcup\limits_{A \in \mathcal{M}} A = \emptyset$ sowie $\bigcap\limits_{A \in \mathcal{M}} A = X$, falls klar ist, was die Obermenge X ist.

(iii) Wir werden sehen, dass es bei der Vereinigung oder beim Schnitt von mehr als zwei Mengen nicht auf die Reihenfolge ankommt. Daher ist die Definition der Vereinigung und des Schnittes eine sinnvolle Verallgemeinerung des Falles von zwei Mengen.

Definition 1.4.4 (Disjunkte Mengen)**.** Seien A, B zwei Mengen.

(i) Dann heißen A und B **disjunkt,** falls $A \cap B = \emptyset$ gilt. In diesem Fall schreiben wir auch $A \dot\cup B$ statt $A \cup B$.

(ii) Sei $\mathcal{M}$ eine Menge von Mengen. Dann heißt diese disjunkt, falls für alle $A, B \in \mathcal{M}$ mit $A \neq B$ stets $A \cap B = \emptyset$ gilt. In diesem Falle schreiben wir auch $\dot\bigcup\limits_{A \in \mathcal{M}} A$ statt $\bigcup\limits_{A \in \mathcal{M}} A$.

Definition 1.4.5 (Komplement)**.** Seien A und B Mengen mit fester Obermenge X. Dann definieren wir

(i) das **Komplement** von A in B durch

$$B \setminus A := \{x \in B \colon x \notin A\}.$$

(ii) das **Komplement** von A durch

$$\complement A := \{x \in X \colon x \notin A\}.$$

Proposition 1.4.6. *Seien A, B und C drei Mengen und sei X eine feste Obermenge. Dann gelten*

(i)	$A \cup B = B \cup A$	*(Kommutativität)*
(ii)	$A \cap B = B \cap A$	*(Kommutativität)*
(iii)	$(A \cup B) \cup C = A \cup (B \cup C)$	*(Assoziativität)*
(iv)	$(A \cap B) \cap C = A \cap (B \cap C)$	*(Assoziativität)*
(v)	$(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$	*(Distributivität)*
(vi)	$(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$	*(Distributivität)*
(vii)	$\complement(A \cup B) = \complement A \cap \complement B$	*(De Morgansche Regel)*
(viii)	$\complement(A \cap B) = \complement A \cup \complement B$	*(De Morgansche Regel)*
(ix)	$\complement\complement A = A$	
(x)	$A \cup \complement A = X$	
(xi)	$A \setminus B = A \cap \complement B$	

Beweis. Wir betrachten nur zwei der Aussagen und lassen den Rest als Übung.

Erster Teil der Distributivität (v):

„$\subset$“: Sei $x \in (A \cap B) \cup C$ beliebig. Wir erhalten

$$\begin{aligned}
& x \in (A \cap B) \vee x \in C \\
\implies & (x \in A \wedge x \in B) \vee x \in C \\
\implies & (x \in A \vee x \in C) \wedge (x \in B \vee x \in C) && \text{(Proposition 1.1.7 (xvii))} \\
\implies & (x \in A \cup C) \wedge (x \in B \cup C) \\
\implies & x \in (A \cup C) \cap (B \cup C).
\end{aligned}$$

„$\supset$“: Folgt analog.

Erste De Morgansche Regel (vii):

„$\subset$“: Sei $x \in \complement(A \cup B)$. Es folgt

$$\begin{aligned}
& x \notin (A \cup B) \\
\implies & \neg(x \in (A \cup B)) \\
\implies & \neg(x \in A \vee x \in B) \\
\implies & \neg(x \in A) \wedge \neg(x \in B) && \text{(Proposition 1.1.7 (xx))} \\
\implies & x \notin A \wedge x \notin B \\
\implies & x \in \complement A \wedge x \in \complement B \\
\implies & x \in \complement A \cap \complement B.
\end{aligned}$$

„$\supset$“: Folgt wieder analog. □

Axiom 1.4.7 (Potenzmenge)**.** Sei A eine beliebige Menge. Dann gibt es die Menge $\mathcal{P}(A)$ (oder 2^A), die **Potenzmenge** von A. Ihre Elemente sind genau die Teilmengen von A.

Beispiel 1.4.8.

(i) Sei $A = \{a, b, c\}$, so ist

$$\mathcal{P}(A) = \{\emptyset, \{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}, \{b, c\}, \{a, b, c\}\}.$$

(ii) Ist $A = \emptyset$, so ist $\mathcal{P}(A) = \{\emptyset\}$. Beachte, dass (insbesondere in diesem Fall) $\mathcal{P}(A) \neq A$ ist.

(iii) Die Aussagen $M \subset A$ und $M \in \mathcal{P}(A)$ sind nach Definition der Potenzmenge äquivalent.

Axiom 1.4.9 (Kartesisches Produkt)**.** Seien A und B zwei Mengen. Dann gibt es eine Menge, das **kartesische Produkt** von A und B, $A \times B$, das aus allen geordneten Paaren (a, b) mit $a \in A$ und $b \in B$ besteht. a heißt die erste und b die zweite Komponente dieses Paares. Es gilt

$$A \times B = \{(a, b) \colon a \in A \wedge b \in B\}.$$

Bemerkung 1.4.10. $\star$ Mengentheoretisch kann man solche Paare als gewisse zweielementige Mengen darstellen. Dabei wird $(a, b) \in A \times B$ durch $\{a, \{a, b\}\} \subset A \cup \mathcal{P}(A \cup B)$ oder $\{a, \{a, b\}\} \in \mathcal{P}(A \cup \mathcal{P}(A \cup B))$ dargestellt.

Damit können wir definieren, was eine Funktion oder Abbildung ist:

Definition 1.4.11 (Funktion, Abbildung)**.** Seien A, B zwei Mengen.

(i) Eine **Funktion** oder **Abbildung** f von A nach B, $f \colon A \to B$, ist eine Teilmenge von $A \times B$ mit der Eigenschaft, dass es zu jedem $a \in A$ genau ein $b \in B$ mit $(a, b) \in f$ gibt:

$$\forall\, a \in A\, \exists\, !\, b \in B \colon (a, b) \in f.$$

Wir schreiben dafür auch $b = f(a)$ oder $a \mapsto b$.

Die Beschreibung einer Funktion als Teilmenge des kartesischen Produktes wird später kaum noch verwendet. Dafür definiert man dann den **Graphen** von f durch

$$\operatorname{graph} f := \{(x, f(x)) \in A \times B \colon x \in A\}.$$

Der Graph einer Funktion f ist gerade die oben betrachtete Teilmenge $f \subset A \times B$.

(ii) A heißt **Definitionsbereich** der Funktion f, bezeichnet mit $D(f)$, und

$$f(A) := \{f(x) \colon x \in A\} \equiv \{y \in B \colon (\exists\, x \in A \colon f(x) = y)\}$$

heißt **Bild** oder **Wertebereich** von f. Wir schreiben $\operatorname{im} f$ oder $R(f)$ für $f(A)$ (für image bzw. range).

(iii) Allgemeiner definieren wir für beliebige $M \subset A$ eine Menge $f(M)$, das **Bild** von M unter f, durch

$$f(M) := \{y \in B \colon (\exists x \in M \colon y = f(x))\} \equiv \{f(x) \colon x \in M\}.$$

Somit induziert f eine Abbildung $\mathcal{P}(A) \to \mathcal{P}(B)$, die wir wieder mit f bezeichnen.

(iv) Zu einer beliebigen Abbildung $f \colon A \to B$ definieren wir die **Urbildabbildung**

$$f^{-1} \colon \mathcal{P}(B) \to \mathcal{P}(A)$$

durch

$$f^{-1}(M) = \{x \in A \colon f(x) \in M\}$$

für beliebige $M \subset B$. Die Menge $f^{-1}(M) \subset A$ heißt das Urbild von M unter der Abbildung f.

Bemerkung 1.4.12. Zwei Funktionen $f \colon A \to B$ und $g \colon C \to D$ sind gleich, wenn f und g als Teilmengen von $A \times B$ bzw. $C \times D$ gleich sind. Insbesondere kann $f = g$ nur gelten, falls auch $A = C$ und $B = D$ gelten.

Wir definieren einige fundamentale Eigenschaften von Funktionen.

Definition 1.4.13. Sei $f \colon A \to B$ eine Funktion.

(i) Dann heißt f **injektiv,** falls aus $f(x) = f(y)$ bereits $x = y$ folgt.

Im Englischen sagt man, dass eine Abbildung "one to one" sei und in älteren Texten heißen solche Abbildungen eineindeutig.

(ii) f heißt **surjektiv,** falls $f(A) = B$ gilt.
Wir sagen, dass f die Menge A **auf** die Menge B abbildet. Ist f nicht notwendigerweise surjektiv, so heißt f im Gegensatz dazu eine Abbildung von A **nach** B oder von der Menge A **in** die Menge B.

(iii) f heißt **bijektiv,** falls f injektiv und surjektiv ist. Eine bijektive Abbildung heißt Bijektion.

(iv) Ist f injektiv, so definieren wir die **Inverse** oder **Umkehrabbildung** von f durch

$$\begin{aligned} f^{-1}\colon \operatorname{im} f &\to A, \\ f(x) &\mapsto x, \end{aligned}$$

d. h. wir definieren $f^{-1}(f(x)) := x$.

Bemerkung 1.4.14.

(i) Wir verwenden f^{-1} für die Inverse und für die Urbildabbildung. Die sollte nicht zu Verwirrungen führen, es gilt nämlich, falls beide definiert sind, für die Inverse I, die Urbildabbildung U und alle $x \in A$

$$\{I(f(x))\} = U(\{f(x)\}).$$

(ii) Ist $f\colon A \to B$ eine Abbildung, die $g\colon \mathcal{P}(A) \to \mathcal{P}(B)$ induziert, so gilt

$$\{f(x)\} = g(\{x\})$$

für alle $x \in A$.

Definition 1.4.15 (Komposition von Abbildungen)**.** Seien $f\colon A \to B$ und $g\colon B \to C$ zwei Abbildungen. Dann definieren wir eine Abbildung $g \circ f\colon A \to C$ durch $x \mapsto g(f(x))$. $g \circ f$ heißt **Komposition** der Abbildungen f und g.

Beispiele 1.4.16.

(i) Sei A eine Menge. Dann heißt

$$\Delta(A) = \{(x, x) \in A \times A\colon x \in A\}$$

die **Diagonale** von $A \times A$.

(ii) Die Abbildung $\Delta(A) \subset A \times A$ heißt **Identität** auf A, id_A oder id. Es gilt $\mathrm{id}_A(x) = x$ für alle $x \in A$.

(iii) Ist $f\colon A \to B$ bijektiv, so gelten $f^{-1} \circ f = \mathrm{id}_A$ und $f \circ f^{-1} = \mathrm{id}_B$.

(iv) Sei A eine Menge und gelte $C \subset A$. Dann heißt die Abbildung $i\colon C \to A$ mit $x \mapsto x$ **Inklusionsabbildung.** i ist injektiv. Wir schreiben dafür manchmal $i\colon C \hookrightarrow A$.

(v) Sei $f\colon A \to B$ eine Abbildung. Sei $C \subset A$ eine Menge. Dann definieren wir die **Einschränkung** oder **Restriktion** von f auf C durch

$$\begin{aligned} f|_C\colon C &\to B, \\ x &\mapsto f(x). \end{aligned}$$

Ist $i\colon C \hookrightarrow A$ die Inklusionsabbildung, so gilt $f|_C = f \circ i$.

Bemerkung 1.4.17. Seien $f\colon A \to B$, $g\colon B \to C$ und $h\colon C \to D$ drei Abbildungen. Dann gelten

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

und

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

wobei die zweite Gleichheit sowohl für die Urbildabbildungen als auch (falls definiert) für die Inversen gilt.

1.5. **Natürliche Zahlen und vollständige Induktion.** Wir benutzen die reellen Zahlen samt Dezimalschreibweise wie sie aus der Schule und insbesondere der Vorlesung Analysis I bekannt sind.

Definition 1.5.1 (Natürliche Zahlen)**.** Die **natürlichen Zahlen** $\mathbb{N}$ sind die kleinste Teilmenge $A \subset \mathbb{R}$ mit

(N1) $0 \in A$ und
(N2) $a + 1 \in A$ für alle $a \in A$.

$\mathbb{N}$ ist dabei die kleinste solche Menge in dem Sinn, dass jede weitere Menge $\mathcal{N} \subset \mathbb{R}$, die (N1) und (N2) erfüllt, auch $\mathbb{N} \subset \mathcal{N}$ erfüllt.

Lemma 1.5.2. *Es gibt die natürlichen Zahlen. Sie sind eindeutig bestimmt.*

Beweis.

(i) **Existenz:** Sei $\mathcal{M}$ die Menge aller Teilmengen A von $\mathbb{R}$, so dass $0 \in A$ und $a + 1 \in A$ für alle $a \in A$ gelten. Dann ist $\mathcal{M} \neq \emptyset$, da $\mathbb{R} \in \mathcal{M}$ ist. Definiere

$$\mathbb{N} := \bigcap_{A \in \mathcal{M}} A.$$

Da $0 \in A$ für alle $A \in \mathcal{M}$ gilt, folgt auch $0 \in \mathbb{N}$. Sei $a \in \mathbb{N}$. Dann folgt $a \in A$ für alle $A \in \mathcal{M}$. Somit ist auch $a + 1 \in A$ für alle diese Mengen A und wir erhalten $a + 1 \in \mathbb{N}$.

$\mathbb{N}$ ist die kleinste Teilmenge von $\mathbb{R}$, die (N1) und (N2) erfüllt, da wir $\mathbb{N}$ als Schnitt aller Teilmengen von $\mathbb{R}$ mit diesen beiden Eigenschaften definiert haben: Sei $B \subset \mathbb{R}$ beliebig mit (N1) und (N2). Dann gilt nach Definition von $\mathbb{N}$ als Schnitt solcher Mengen bereits $\mathbb{N} \subset B$.

(ii) **Eindeutigkeit:** Augfrund der Minimalität folgt für zwei solche Mengen $\mathbb{N}_1 \subset \mathbb{N}_2$ und $\mathbb{N}_2 \subset \mathbb{N}_1$. □

Theorem 1.5.3 (Vollständige Induktion)**.** *Erfülle $M \subset \mathbb{N}$ die Bedingungen*

(i) $0 \in M$ (Induktionsanfang) und
(ii) folgt aus $n \in M$ auch $n + 1 \in M$ (Induktionsschritt),

so gilt $M = \mathbb{N}$.

Beweis. Die Menge M erfüllt (N1) und (N2) und ist daher eine der Mengen, die im Schnitt in Lemma 1.5.2 auftaucht. Somit gilt $\mathbb{N} \subset M$ und damit $M = N$. □

Theorem 1.5.4. ⋆ *Sei p eine Aussageform auf $\mathbb{N}$. Gelten*

(i) $p(0)$ und
(ii) $p(n) \Longrightarrow p(n + 1)$ für alle $n \in \mathbb{N}$,

so gilt $p(n)$ für alle $n \in \mathbb{N}$.

Beweis. Definiere $M := \{n \in \mathbb{N} \colon p(n) \text{ ist wahr}\}$. Dann folgt die Behauptung aus Theorem 1.5.3. □

Bemerkung 1.5.5. ⋆

(i) Wir können vollständige Induktion auch benutzen, wenn $p(0)$ gilt und wenn für alle $n \in \mathbb{N}$ aus $p(0) \wedge p(1) \wedge \ldots \wedge p(n)$ die Aussage $p(n + 1)$ folgt: Wir benutzen dann die Aussageform

$$q(n) = (p(0) \wedge p(1) \wedge \ldots \wedge p(n)) = \bigwedge_{0 \leq k \leq n} p(k).$$

Wiederum folgt $p(n)$ für alle $n \in \mathbb{N}$.

(ii) Gilt $p(n_0)$ für $n_0 \in \mathbb{N}$ und folgt aus $p(n)$ die Aussage $p(n+1)$ für alle $n \geq n_0$, so gilt $p(n)$ für alle $n \in \mathbb{N}$ mit $n \geq n_0$.

Bemerkung 1.5.6. ⋆ Genauso wie wir Aussagen induktiv beweisen können, kann man auch entsprechend definieren, siehe [5] für Details.

Beispiel 1.5.7 (Summenschreibweise)**.** Seien $k \leq l$ mit $k, l \in \mathbb{N}$. Seien a_n mit $k \leq n \leq l$ und $n \in \mathbb{N}$ (letzteres wird häufig nicht hingeschrieben, wenn dies selbstverständlich sein sollte) reelle Zahlen (oder andere Elemente, die wir aufsummieren können). Dann definieren wir

$$\sum_{n=k}^{l} a_n$$

induktiv durch

$$\sum_{n=k}^{k} a_n := a_k \quad \text{und} \quad \sum_{n=k}^{m+1} a_n := \sum_{n=k}^{m} a_n + a_{m+1}$$

für $k \leq m < l$.

Beachte, dass wir die Ausdrücke auch auf diese Art und Weise definieren können, obwohl k nicht beliebig groß werden darf, da $k \leq l$ stets erfüllt bleiben muss.

Eine entsprechende Notation werden wir später auch verwenden, wenn der Summationsbereich zwischen zwei ganzen Zahlen, die wir noch definieren werden, liegt.

Wir behaupten, dass

$$\sum_{i=0}^{n} i = \frac{n(n+1)}{2}$$

gilt.

Beweis. Wir beweisen dies per Induktion:

(i) **Induktionsanfang:** Für $n = 0$ müssen wir nachweisen, dass $\sum_{i=0}^{0} i = 0$ gilt. Dies ist offensichtlich richtig.

(ii) **Induktionsschritt:** Wir nehmen nun an, dass die Aussage für ein festes $n \in \mathbb{N}$ bereits gezeigt sei (Induktionsannahme) und wollen nachweisen, dass sie auch für $n+1$ gilt. Es gilt

$$\begin{aligned}\sum_{i=0}^{n+1} i &= \sum_{i=0}^{n} i + (n+1) \overset{\text{I.A.}}{=} \frac{n(n+1)}{2} + (n+1) = \frac{n(n+1)}{2} + \frac{2(n+1)}{2} \\ &= \frac{(n+2)(n+1)}{2} = \frac{(n+1)((n+1)+1)}{2}.\end{aligned}$$

Damit ist der Induktionsschritt gezeigt.

Aufgrund des Prinzips der vollständigen Induktion haben wir damit die Aussage für alle $n \in \mathbb{N}$ gezeigt. □

Beispiele 1.5.8 (Rekursive Definition)**.**

(i) Fakultät: Wir setzen $0! := 1$ und für $n \in \mathbb{N}$ definieren wir $(n+1)! := n! \cdot (n+1)$.

Somit folgt $0! = 1$, $1! = 1$, $2! = 2$, $3! = 6$, $4! = 24$, $5! = 120$, $6! = 720$, $8! = 5760$, ...

(ii) Fibonaccizahlen: Wir definieren eine Funktion $f\colon \mathbb{N} \to \mathbb{N}$ durch $f(0) := 0$, $f(1) := 1$ und $f(n+2) := f(n+1) + f(n)$ für $n \in \mathbb{N}$.

Wir erhalten

n	0	1	2	3	4	5	6	7	8	9	10
f(n)	0	1	1	2	3	5	8	13	21	34	55

Dieses Beispiel benutzt die Variante der rekursiven Definition, bei der auf mehrere Vorgängerwerte zurückgegriffen wird.

Mit Hilfe von Eigenwerten kann man eine geschlossene Formel für die Fibonaccizahlen herleiten.

1.6. **Relationen.**

Definition 1.6.1 (Relationen)**.** Seien A, B Mengen.

(i) Eine Teilmenge $R \subset A \times B$ heißt **Relation**. Statt $(x, y) \in R$ sagen wir auch, dass $R(x, y)$ gilt.
(ii) Eine Relation $R \subset A \times A$ heißt
  (a) **reflexiv,** falls $R(x, x)$ für alle $x \in A$ gilt,
  (b) **symmetrisch,** falls für alle $x, y \in A$ aus $R(x, y)$ auch $R(y, x)$ folgt,
  (c) **antisymmetrisch,** falls für alle $x, y \in A$ aus $R(x, y)$ und $R(y, x)$ bereits $x = y$ folgt,
  (d) **transitiv,** falls für alle $x, y, z \in A$ aus $R(x, y)$ und $R(y, z)$ auch $R(x, z)$ folgt.
(iii) Sei $R \subset A \times A$ eine Relation. Dann heißt R
  (a) **Äquivalenzrelation,** falls R reflexiv, symmetrisch und transitiv ist. Häufig schreiben wir $x \sim y$ statt $R(x, y)$ für Äquivalenzrelationen.
  (b) **Ordnung** oder **Totalordnung,** falls R reflexiv, transitiv, antisymmetrisch und total/linear ist.
  (c) **Halbordnung,** falls R reflexiv, transitiv und antisymmetrisch ist.

Beispiele 1.6.2. (teils $\star$)

(i) Auf den reellen Zahlen gibt es folgende wichtige Relationen:
  (a) „$<$“ ($a < b$ ist genau dann wahr, falls a strikt kleiner als b ist),
  (b) „$=$“ ($a = b$ ist genau dann wahr, falls a und b gleich sind),
  (c) „$>$“ ($a > b$ ist genau dann wahr, falls a strikt größer als b ist),
  (d) „$\leq$“ ($a \leq b$ ist genau dann wahr, falls $a <$ oder $a = b$ wahr ist),
  (e) „$\geq$“ ($a \geq b$ ist genau dann wahr, falls $a > b$ oder $a = b$ wahr ist) und
  (f) „$\neq$“ ($a \neq b$ ist genau dann wahr, falls a und b verschiedene reelle Zahlen sind).

  (Genauer werden diese Relationen auf den reellen Zahlen in der Analysis im Zusammenhang mit geordneten Körpern besprochen.)

  $=$ ist eine Äquivalenzrelation, $\leq$ und $\geq$ sind Ordnungen.
(ii) Sei $0 < q \in \mathbb{Z}$. Auf $\mathbb{Z}$ definieren wir eine Äquivalenzrelation $\sim$ durch $n \sim m$ für $n, m \in \mathbb{Z}$, wenn $n - m$ durch q teilbar ist, d. h. wenn es ein $r \in \mathbb{Z}$ gibt, so dass $n - m = qr$ ist. Wir schreiben auch

$$n \equiv m \pmod{q}.$$

  (Rechnen modulo 24 verwenden wir bei der Uhr: 2 Stunden nach 23 Uhr ist es 1 Uhr.)
(iii) Sei A eine Menge. Dann ist „$\subset$“ eine Halbordnung auf $\mathcal{P}(A)$.

Definition 1.6.3. Sei $R \subset A \times A$ eine Äquivalenzrelation. Sei $x \in A$. Dann heißt

$$[x] := \{y \in A \colon R(x, y)\}$$

die **Äquivalenzklasse** oder **Restklasse** von x. Statt $y \in [x]$ schreiben wir auch $y \equiv x \mod R$.

Mit $A/R := \{[x] \colon x \in A\}$ bezeichnen wir die Menge aller Äquivalenzklassen dieser Relation.

Beispiel 1.6.4. Die folgenden Beispiele verwenden Schulwissen.

(i) Die Relation „wohnt in der gleichen Stadt“ ist eine Äquivalenzrelation. Die Äquivalenzklassen bestehen jeweils aus allen Bürgern von Konstanz, Kreuzlingen, ...
(ii) Auf der Menge $A := \mathbb{Z} \times (\mathbb{Z} \setminus \{0\})$ definieren wir eine Äquivalenzrelation $\sim$ durch die Festlegung, dass $(a, b) \sim (c, d)$ genau dann gilt, falls $ad = bc$ richtig

ist. Fassen wir nun (a, b) als den Bruch $\frac{a}{b}$ auf, so bestehen die Äquivalenzklassen $A/\sim$ gerade aus den Brüchen, die dieselbe reelle Zahl ergeben.

(iii) Sei $n \in \mathbb{N} \setminus \{0\}$. Auf $\mathbb{Z}$ definieren wir eine Äquivalenzrelation $\sim$ durch:

$$a \sim b :\Longleftrightarrow \exists\, k \in \mathbb{Z} \colon a - b = kn.$$

Die Äquivalenzklassen $\mathbb{Z}/\sim$ bestehen dann gerade aus den ganzen Zahlen, die bei Division durch n denselben Rest lassen. Statt $a \sim b$ schreiben wir hier auch $a \equiv b \pmod{n}$.

Definition 1.6.5 (Partition)**.** Sei A eine Menge. Eine **Partition** von A ist eine Überdeckung $(A_i)_{i\in I}$ von A mit $A_i \subset A$ für alle $i \in I$ durch paarweise disjunkte Mengen, d. h. es gilt $A_i \cap A_j = \emptyset$ für alle $i \neq j \in I$. Wir schreiben

$$A = \dot{\bigcup_{i\in I}} A_i.$$

Proposition 1.6.6.

(i) Sei $\sim$ eine Äquivalenzrelation auf A. Dann bilden die Restklassen von $\sim$ eine Partition von A.

(ii) Sei $(A_i)_{i\in I}$ eine Partition von A. Dann ist $\sim$ mit

$$x \sim y \quad :\Longleftrightarrow \quad \exists\, i \in I \colon x, y \in A_i$$

eine Äquivalenzrelation auf A.

Beweis.

(i) Sei $\sim$ eine Äquivalenzrelation auf A. Betrachte die Familie $([x])_{[x]\in A/\sim}$.
  (a) Wegen $x \in [x]$ wird ganz A überdeckt und es gilt $[x] \subset A$ für alle $x \in A$.
  (b) Disjunktheit: Seien $x, y \in A$ beliebig. Sind $[x], [y]$ nicht disjunkt, so gibt es ein $z \in [x] \cap [y]$. Wir behaupten, dass dann auch $[x] = [y]$ gilt, dass es sich also um dieselben Äquivalenzklassen handelt. Seien $\bar{x} \in [x]$ und $\bar{y} \in [y]$. Dann gilt $\bar{x} \sim x \sim z \sim y$, also folgt $\bar{x} \in [y]$. Ebenso gilt $\bar{y} \in [x]$ und $[x] = [y]$ folgt.

(ii) (a) reflexiv: $x \sim x$ ist klar.
  (b) symmetrisch: $x \sim y \Longrightarrow y \sim x$ ist ebenso klar.
  (c) transitiv: Gelte $x \sim y$ und $y \sim z$. Somit gibt es $k, l \in I$ mit $x, y \in A_k$ und $y, z \in A_l$. Weil die Familie $(A_i)_{i\in I}$ aus paarweise disjunkten Mengen besteht und $y \in A_k \cap A_l$ gilt, folgt $k = l$. Somit gilt auch $x, z \in A_k = A_l$ und nach Definition folgt $x \sim z$ wie behauptet. □

1.7. Familien und Auswahlaxiom.

Definition 1.7.1 (Familie)**.**

(i) Seien I, X Mengen und $f \colon I \to X$ eine Abbildung. Dann heißt f auch **Familie.** Wir schreiben $(x_i)_{i\in I}$. Dabei ist $x_i = f(i)$ für $i \in I$. I heißt Indexmenge.

(ii) Sei $(x_i)_{i\in I}$ eine Familie. Ist $I = \{1, 2, \ldots, n\}$ so schreiben wir auch $(x_i)_{1\leq i\leq n}$. Im Falle
  (a) $n = 2$ heißt die Familie ein **Paar:** (x_1, x_2),
  (b) $n = 3$ heißt die Familie ein **Tripel:** (x_1, x_2, x_3),
  (c) n heißt die Familie ein n-**Tupel:** $(x_1, x_2, \ldots, x_n)$.

Analog zu Definition 1.4.2 haben wir:

Definition 1.7.2 (Schnitt und Vereinigung von Familien)**.** Sei $(A_i)_{i\in I}$ eine Familie von Mengen mit Obermenge X, so definieren wir

(i) $\bigcup\limits_{i\in I} A_i := \{x \in X \colon (\exists\, i \in I \colon x \in A_i)\}$.

(ii) $\bigcap\limits_{i\in I} A_i := \{x \in X \colon (\forall\, i \in I \colon x \in A_i)\}$.

(iii) Im Falle $I = \{1, 2, \dots, n\}$ schreiben wir $\bigcup_{i=1}^{n} A_i \equiv \bigcup_{i \in I} A_i$ bzw. $\bigcap_{i=1}^{n} A_i \equiv \bigcap_{i \in I} A_i$.

(iv) $\star$ Eine Menge I heißt endlich, wenn es ein $n \in \mathbb{N}$ und eine bijektive Abbildung $\varphi\colon \{1, 2, \dots, n\} \to I$ gibt und sonst unendlich. Ist I endlich, so heißt $\bigcup_{i \in I} A_i$ eine endliche Vereinigung.

Wir verallgemeinern nun das kartesische Produkt auf mehr als zwei Faktoren. Den Fall von zwei Faktoren erhalten wir dabei, wenn wir $I = \{1, 2\}$ verwenden.

Definition 1.7.3 (Kartesisches Produkt)**.**

(i) Sei $I \neq \emptyset$ und sei $(A_i)_{i \in I}$ eine Familie von Mengen. Dann definieren wir das **kartesische Produkt** durch

$$\prod_{i \in I} A_i = \{(x_i)_{i \in I} \colon \text{für alle } i \in I \text{ gilt } x_i \in A_i\}.$$

(ii) Zu beliebigem $j \in I$ definieren wir die j-te **Projektion(sabbildung)**

$$\pi_j \colon \prod_{i \in I} A_i \to A_j \quad \text{durch} \quad \pi_j((x_i)_{i \in I}) := x_j.$$

Beispiel 1.7.4. Sei F eine Menge. Ist $I = \{1, \dots, n\}$ und $A_i = F$ für alle $i \in I$, so ist

$$\prod_{i=1}^{n} A_i \equiv \prod_{i \in I} A_i \equiv F^n = \left\{ \left(x^1, \dots, x^n\right) : x^i \in F \text{ für alle } 1 \leq i \leq n \right\}.$$

$\pi_i\left(x^1, \dots, x^n\right) = x^i$ heißt die i-te Komponente des n-Tupels $\left(x^1, \dots, x^n\right)$.

Aus Kovarianzgründen schreiben wir hier die Indices nach oben. Dies sind keine Exponenten. Eine genauere Begründung dafür werden wir erst deutlich später kennen lernen.

Beim ersten Lesen ist das folgende Axiom sicher schwer zugänglich. Ist I eine endliche Menge oder (allgemeiner) nicht mächtiger (noch zu definieren; siehe Analysis) als $\mathbb{N}$, so können wir die Aussage aus dem Bisherigen beweisen, i. a. klappt dies jedoch nicht.

Axiom 1.7.5 (Auswahlaxiom)**.** Sei $(A_i)_{i \in I}$ eine Familie von Mengen A_i mit $A_i \neq \emptyset$ für alle $i \in I$. Dann gilt $\prod_{i \in I} A_i \neq \emptyset$, d. h. es existiert eine Familie $(x_i)_{i \in I}$ mit $x_i \in A_i$ für alle $i \in I$.

Proposition 1.7.6. $\star$ *Seien $I \neq \emptyset$ und $(A_i)_{i \in I}$ eine Familie von Mengen. Dann gilt*

$$\prod_{i \in I} A_i = \emptyset \quad \Longleftrightarrow \quad \textit{Es gibt ein } i \in I \textit{ mit } A_i = \emptyset.$$

Beweis.

„$\Longleftarrow$“: Ist $i_0 \in I$ und $A_{i_0} = \emptyset$, so ist klar, dass es keine solche Familie $(x_i)_{i \in I}$ geben kann.

„$\Longrightarrow$“: Gilt umgekehrt $A_i \neq \emptyset$ für alle $i \in I$, so folgt $\prod_{i \in I} A_i \neq \emptyset$ gerade aus dem Auswahlaxiom. □

Eine Folgerung des Auswahlaxioms ist das Zornsche Lemma. Wir werden es benutzen um zu zeigen, dass jeder Vektorraum eine Basis besitzt.

Lemma 1.7.7 (Zornsches Lemma)**.** *Sei M eine nichtleere Menge mit einer Teilordnung $\leq$. Nehme an, dass jede total geordnete Teilmenge $\Lambda \subset M$ (= Kette) eine obere Schranke $b \in M$ besitzt, d. h. dass $x \leq b$ für alle $x \in \Lambda$ gilt. Dann enthält M ein maximales Element x_0, d. h. ein Element $x_0 \in M$, so dass aus $x \geq x_0$ bereits $x = x_0$ folgt. (Beachte, dass x_0 i. a. nicht eindeutig bestimmt ist.)*

Beweis. Logikvorlesung. □

Proposition 1.7.8. *Seien A, B, C Mengen. Seien $\varphi\colon A \to B$ und $\psi\colon B \to C$ Abbildungen. Sei $f\colon A \to B$ eine weitere Abbildung. Dann gelten die folgenden Aussagen:*

(i) Ist $\psi \circ \varphi$ injektiv, so ist φ injektiv.
(ii) Ist $\psi \circ \varphi$ surjektiv, so ist ψ surjektiv.
(iii) f ist genau dann surjektiv, wenn es eine Abbildung $g\colon B \to A$ mit $f \circ g = \mathrm{id}_B$ gibt.
(iv) f ist genau dann injektiv, wenn es eine Abbildung $g\colon B \to A$ mit $g \circ f = \mathrm{id}_A$ gibt.

Beweis.

(i) Falls nicht, so gäbe es $x, y \in A$ mit $x \neq y$ und $\varphi(x) = \varphi(y)$. Daraus folgt auch $\psi(\varphi(x)) = \psi(\varphi(y))$. Dies widerspricht der angenommenen Injektivität von $\psi \circ \varphi$. Somit folgt die Behauptung.
(ii) Übung.
(iii)
„$\Longleftarrow$“: Folgt aus (ii).
„$\Longrightarrow$“: Wir definieren **eine** solche Abbildung g wie folgt: Sei $y \in B$ beliebig. Da f surjektiv ist, ist die Menge $f^{-1}(\{y\}) \neq \emptyset$. Aus dieser Menge wählen wir ein x aus und definieren $g(y) := x$. Da $y \in B$ beliebig war, definiert dies eine Abbildung $g\colon B \to A$. Es folgt $f(g(y)) = f(x) = y = \mathrm{id}_B(y)$ wie gewünscht.
(Erläuterung zum Sprachgebrauch: „Wir definieren **die** Abbildung g wie folgt“ würde bedeuten, dass wir auch behaupten, dass es nur eine solche Abbildung g gibt, was hier im Allgemeinen nicht zutrifft.)
(iv) Übung. □

1.8. Weitere Zahlen ⋆.

Definition 1.8.1.

(i) Die Menge der reellen Zahlen $x \in \mathbb{R}$, für die es $m, n \in \mathbb{N}$ gibt mit $m - n = x$, heißt die Menge der **ganzen Zahlen:**

$$\mathbb{Z} := \{m - n\colon m, n \in \mathbb{N}\}.$$

(ii) Die **rationalen Zahlen** sind als Quotienten von ganzen Zahlen (ohne Null im Nenner) definiert:

$$\mathbb{Q} := \left\{ \frac{p}{q}\colon p, q \in \mathbb{Z}, q \neq 0 \right\}.$$

(iii) Die Menge $\mathbb{I} := \mathbb{R} \setminus \mathbb{Q}$ heißt die Menge der **irrationalen Zahlen.**
(iv) Die **komplexen Zahlen** sind Paare von reellen Zahlen:

$$\mathbb{C} := \{(a, b)\colon a, b \in \mathbb{R}\}.$$

Für $(a, b), (c, d) \in \mathbb{C}$ definieren wir eine Addition durch $(a, b) + (c, d) := (a + c, b + d)$ sowie eine Multiplikation durch $(a, b) \cdot (c, d) := (ac - bd, ad + bc)$. Statt (a, b) schreiben wir auch $a + ib$. Somit gilt insbesondere $i^2 = -1$. a heißt Realteil der komplexen Zahl $z = a + ib$ mit $a, b \in \mathbb{R}$, $a = \operatorname{Re} z$, und b heißt Imaginärteil der komplexen Zahl z, $b = \operatorname{Im} z$. Wir definieren $\overline{a + ib} := a - ib$, die zu $a + ib$ komplex konjugierte komplexe Zahl, und $|a + ib| := \sqrt{a^2 + b^2}$, den Betrag der komplexen Zahl $a + ib$. Es gelten für $a, b \in \mathbb{R}$ und $z, w \in \mathbb{C}$

$$|a + ib|^2 = (a + ib)\overline{(a + ib)},$$
$$\overline{z + w} = \overline{z} + \overline{w},$$

$$\begin{aligned}\overline{zw} &= \overline{z} \cdot \overline{w},\\ |z|^2 &= |\operatorname{Re} z|^2 + |\operatorname{Im} z|^2,\\ |z| &= |\overline{z}|.\end{aligned}$$

Wir betrachten $\mathbb{R}$ vermöge $x \mapsto (x, 0)$ als Teilmenge der komplexen Zahlen und erhalten $\overline{x} = x$ für $x \in \mathbb{R}$.

Es gilt

$$(a + ib)(a - ib) = a^2 + b^2.$$

Sei $(a, b) \neq (0, 0)$, so ist daher $\frac{a}{a^2+b^2} + \frac{-b}{a^2+b^2} i =: z \in \mathbb{C}$ mit der Eigenschaft $(a + ib) \cdot z = 1$.

2. Körper und Vektorräume

2.1. **Körper.** Bereits bekannte Beispiele von Körpern sind die reellen Zahlen $\mathbb{R}$, die komplexen Zahlen $\mathbb{C}$ oder die rationalen Zahlen $\mathbb{Q}$. Allgemein definieren wir

Definition 2.1.1 (Körper)**.** Eine Menge F (engl.: field) mit einer Addition „+“ und einer Multiplikation „·“ heißt Körper, wenn für alle $a, b, c \in F$ die folgenden Axiome erfüllt sind. (Die Verknüpfungen Addition und Multiplikation sind Abbildungen $F \times F \to F$, die wir in der Form $a + b$ oder $a \cdot b$ schreiben.)

(F1) $a + b = b + a$ (für alle $a, b \in F$) (Kommutativgesetz der Addition)
(F2) $a + (b + c) = (a + b) + c$ (Assoziativgesetz der Addition)
(F3) Es gibt ein **neutrales Element** $0 \in F$ der Addition mit $a + 0 = a$ für alle $a \in F$.
(F4) Zu jedem $a \in F$ gibt es ein (eindeutig bestimmtes) Element $-a \in F$, so dass $a + (-a) = 0$ gilt.
(F5) $a \cdot b = b \cdot a$ (Kommutativgesetz der Multiplikation)
(F6) $a \cdot (b \cdot c) = (a \cdot b) \cdot c$ (Assoziativgesetz der Multiplikation)
(F7) Es gibt ein neutrales Element $1 \in F$ der Multiplikation, $0 \neq 1$, mit $a \cdot 1 = a$ für alle $a \in F$.
(F8) Zu jedem $a \in F$ mit $a \neq 0$ gibt es ein (eindeutig bestimmtes) Element $a^{-1} \in F$ mit $a \cdot a^{-1} = 1$.
(F9) $a \cdot (b + c) = (a \cdot b) + (a \cdot c)$ (Distributivgesetz)

Bemerkung 2.1.2.
(1) Die Elemente aus (4) und (8) heißen **inverse** Elemente der Addition bzw. Multiplikation.
(2) Die eingeklammerten Eindeutigkeitsforderungen kann man auch beweisen.

Beispiele 2.1.3.
(i) Die Mengen $\mathbb{R}, \mathbb{Q}, \mathbb{C}$ mit der üblichen Addition und Multiplikation bilden jeweils einen Körper.
(ii) $F := \left\{a + b\sqrt{3} \equiv a + \left(b\sqrt{3}\right) : a, b \in \mathbb{Q}\right\} \equiv \mathbb{Q}\left[\sqrt{3}\right] \subset \mathbb{R}$ mit der üblichen Addition und Multiplikation von reellen Zahlen ist ein Körper.
(iii) Sei p eine Primzahl. Definiere die folgenden Teilmengen von $\mathbb{Z}$

$$\begin{aligned}\bar{0} &:= \{\ldots, -3p, -2p, -p, 0, p, 2p, 3p, \ldots\},\\ \bar{1} &:= \{\ldots, 1 - 3p, 1 - 2p, 1 - p, 1, 1 + p, 1 + 2p, 1 + 3p, \ldots\},\\ \bar{2} &:= \{\ldots, 2 - 3p, 2 - 2p, 2 - p, 2, 2 + p, 2 + 2p, 2 + 3p, \ldots\},\\ \bar{a} &:= \{\ldots, a - 3p, a - 2p, a - p, a, a + p, a + 2p, a + 3p, \ldots\}, \quad a \in \mathbb{Z}.\end{aligned}$$

Wir benutzen hier die Konvention $kp + q := (k \cdot p) + q$. Setze

$$\mathbb{Z}/p\mathbb{Z} := \left\{\bar{0}, \bar{1}, \bar{2}, \ldots, \overline{p-1}\right\}.$$

Auf $\mathbb{Z}/p\mathbb{Z}$ definieren wir Addition und Multiplikation durch

$$\bar{a} + \bar{b} := \overline{a+b} \quad \text{und} \quad \bar{a} \cdot \bar{b} := \overline{a \cdot b}.$$

(Hier ist noch die Wohldefiniertheit nachzuweisen, d. h. wir müssen zeigen, dass $\overline{a+kp} + \overline{b+lp} = \bar{a} + \bar{b}$ und $\overline{a+kp} \cdot \overline{b+lp} = \bar{a} \cdot \bar{b}$ für alle $k, l \in \mathbb{Z}$ gelten.) Wir multiplizieren also jeweils vertreterweise. Dann ist

$$\mathbb{F}_p := \left\{\bar{0}, \bar{1}, \bar{2}, \ldots, \overline{p-1}\right\}$$

mit der oben definierten Addition und Multiplikation ein Körper. (Wir schreiben die Elemente auch wieder als $0, 1, 2, \ldots, p-1$, falls keine Verwechslungsgefahr besteht.)

Beweis. Wir zeigen einige wichtige Eigenschaften, deuten einige Dinge an und lassen den Rest als Übung.

(i) Übung.

(ii) (a) $0 + 0\sqrt{3}$ ist das neutrale Element der Addition.

(b) $1 + 0\sqrt{3}$ ist das neutrale Element der Multiplikation.

(c) Sind $a + b\sqrt{3}, c + d\sqrt{3} \in F$, so folgt auch

$$\left(a + b\sqrt{3}\right) \cdot \left(c + d\sqrt{3}\right) = (ac + 3bd) + (ad + bc)\sqrt{3} \in F.$$

(d) Ist $a + b\sqrt{3} \in F$ mit $(a, b) \neq (0, 0)$, so gilt

$$\begin{aligned}\left(a + b\sqrt{3}\right)^{-1} \equiv \frac{1}{a + b\sqrt{3}} &= \frac{a - b\sqrt{3}}{\left(a + b\sqrt{3}\right)\left(a - b\sqrt{3}\right)} = \frac{a - b\sqrt{3}}{a^2 - 3b^2} \\ &= \frac{a}{a^2 - 3b^2} - \frac{b}{a^2 - 3b^2}\sqrt{3}\end{aligned}$$

und man mache sich klar, dass der Nenner nicht Null ist.

(iii) (a) Wohldefiniertheit der vertreterweise definierten Multiplikation: Seien a, b, k, $l \in \mathbb{Z}$. Aus

$$(a + kp)(b + lp) = ab + (al + bk + klp)p$$

sieht man, dass das Abhändern von a und b um ganzzahlige Vielfache von p auch das Ergebnis nur um ganzzahlige Vielfache von p ändert.

(b) Existenz der Inversen: Wir verwenden Schulwissen. Sei $a \in \mathbb{Z}$ mit $\bar{a} \notin \bar{0}$. Betrachte die Abbildung

$$\begin{aligned} m \colon (\mathbb{Z}/p\mathbb{Z}) \setminus \overline{0} &\to (\mathbb{Z}/p\mathbb{Z}) \setminus \overline{0}, \\ \overline{b} &\mapsto \overline{ab}. \end{aligned}$$

Die Abbildung ist wohldefiniert: Zunächst ist $\overline{ab} \neq \overline{0}$, denn sonst wäre ab durch die Primzahl p teilbar. Demnach wäre auch a durch p teilbar oder b durch p teilbar. Daraus folgte dann aber $\overline{a} = \overline{0}$ oder $\overline{b} = \overline{0}$. Widerspruch. Wir behaupten nun, dass $\overline{1} \in \operatorname{im} m$ gilt. Falls nicht, so ist m eine Abbildung von einer Menge mit $p-1$ Elementen in eine Menge mit $p-1$ Elementen, die nicht surjektiv ist. Somit gibt es mindestens ein Element, das mindestens doppelt getroffen wird, d. h. es gibt $b_1, b_2 \in \mathbb{Z}$ mit $\overline{b_1} \neq \overline{b_2}$ aber $m\left(\overline{b_1}\right) = m\left(\overline{b_2}\right)$. Somit ist $\overline{ab_1} = \overline{ab_2}$. Also gilt auch $\overline{a} \cdot \overline{b_1 - b_2} = \overline{0}$. Somit muss a oder $b_1 - b_2$ durch p teilbar sein. Widerspruch. Daher gibt es doch ein $b \in \mathbb{Z}$ mit $m\left(\overline{b}\right) = \overline{1}$, also $\overline{a}\overline{b} = \overline{ab} = \overline{1}$, wie gewünscht. □

Theorem 2.1.4. *Sei F ein Körper. Dann gilt für alle $a, b \in F$*

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

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

(iii) $-(-a) = a$,

(iv) $(a^{-1})^{-1} = a$ *für* $a \neq 0$,
(v) $(-1) \cdot a = -a$,
(vi) $(-a) \cdot (-b) = a \cdot b$,
(vii) $(-a)^{-1} = -\left(a^{-1}\right)$ *für* $a \neq 0$.

Beweis.

(i) Es gilt $0 \cdot a \overset{(F3)}{=} (0+0) \cdot a \overset{(F5),(F9)}{=} 0 \cdot a + 0 \cdot a$. Die rechte Seite lesen wir als $(0 \cdot a) + (0 \cdot a)$ nach der Regel „Punkt vor Strich“. Addiere nun auf beiden Seiten $-(0 \cdot a)$. Es folgt $0 = 0 \cdot a$. Nun folgt $a \cdot 0 = 0$ aus der Kommutativität der Multiplikation.

(ii) Es gilt $0 = a \cdot (b + (-b)) = a \cdot b + a \cdot (-b)$. Addiere nun auf beiden Seiten $-(a \cdot b)$ und erhalte $a \cdot (-b) = -(a \cdot b)$. Die andere Gleichheit folgt analog.

(iii) Es gelten $0 = a + (-a)$ und $0 = (-a) + (-(-a))$, also auch $(-a) + (-(-a)) = a + (-a)$. Durch Addition von a dürfen wir auf beiden Seiten den Term $(-a)$ streichen. Die Behauptung folgt.

(iv) Verfahre wie beim Beweis von (iii) und verwende Körperaxiom (F8) statt (F4) und Multiplikation statt Addition.

(v) Es gilt $a + (-1) \cdot a = 1 \cdot a + (-1) \cdot a = (1 + (-1)) \cdot a = 0 \cdot a = 0$ nach (i) für die letzte Gleichheit. Addiere nun auf beiden Seiten $(-a)$ und erhalte $(-1) \cdot a = -a$.

(vi) Wir multiplizieren $a + (-a) = 0$ mit $(-b)$ und erhalten

$$a \cdot (-b) + (-a) \cdot (-b) = 0 \cdot (-b) = 0.$$

Andererseits folgt aus $(-b) + b = 0$ durch Multiplikation mit a

$$a \cdot (-b) + a \cdot b = a \cdot 0 = 0.$$

Gleichsetzen und auf beiden Seiten $a \cdot (-b)$ abziehen liefert $(-a) \cdot (-b) = a \cdot b$.

(vii) Wegen $a \neq 0$ und somit $-a \neq 0$ gibt es $(-a)^{-1}$. Es gilt

$$1 = a \cdot a^{-1} \overset{\text{(vi)}}{=} (-a) \cdot \left(-\left(a^{-1}\right)\right).$$

Multipliziere nun beide Seiten von links mit $(-a)^{-1}$. Die Behauptung folgt.

□

2.2. **Vektorräume.**

Definition 2.2.1. Sei F ein Körper. Ein Vektorraum V über F (oder ein F-Vektorraum) ist eine nichtleere Menge zusammen mit einer Addition „+“ und einer Multiplikation „$\cdot$“ mit den folgenden Eigenschaften:

- Die Addition $+\colon V \times V \to V$ erfüllt die folgenden Axiome
  - (V1) $a + b = b + a$ für alle $a, b \in V$, (Kommutativität)
  - (V2) $a + (b + c) = (a + b) + c$ für alle $a, b, c \in V$ (Assoziativität)
  - (V3) Es gibt ein (eindeutig bestimmtes) Element, den Nullvektor 0, mit $a + 0 = a$ für alle $a \in V$. (neutrales Element)
  - (V4) Zu jedem $a \in V$ gibt es ein (eindeutig bestimmtes) Inverses $(-a) \in V$, so dass $a + (-a) = 0$ gilt. (additives Inverses)
- Die skalare Multiplikation $\cdot\colon F \times V \to V$ erfüllt
  - (V5) $\lambda \cdot (a + b) = (\lambda \cdot a) + (\lambda \cdot b)$ für alle $\lambda \in F$, $a, b \in V$
  - (V6) $(\lambda + \mu) \cdot a = (\lambda \cdot a) + (\mu \cdot a)$ für alle $\lambda, \mu \in F$, $a \in V$ (Distributivgesetze)
  - (V7) $(\lambda \cdot \mu) \cdot v = \lambda \cdot (\mu \cdot a)$ für alle $\lambda, \mu \in F$, $a \in V$ (Assoziativität)
  - (V8) $1 \cdot a = a$ für alle $a \in V$.

Die Elemente eines Vektorraumes heißen Vektoren.

Bemerkung 2.2.2.

(i) Ohne (V8) könnte man auch $\lambda \cdot a := 0$ für alle $\lambda \in F$ und alle $a \in V$ setzen.
(ii) Üblicherweise schreibt man (nicht nur hier) λa statt $\lambda \cdot a$.
(iii) Die eingeklammerten Eindeutigkeitsforderungen kann man auch hier wieder beweisen.

Beispiele 2.2.3.

(i) $\mathbb{R}^n$ mit der üblichen Addition und Skalarmultiplikation, in Koordinaten

$$(\xi^1, \ldots, \xi^n) + (\eta^1, \ldots, \eta^n) := (\xi^1 + \eta^1, \ldots \xi^n + \eta^n),$$
$$\lambda \cdot (\xi^1, \ldots, \xi^n) := (\lambda\xi^1, \ldots, \lambda\xi^n),$$

ist ein $\mathbb{R}$-Vektorraum. (Beachte, dass die oberen Indices Komponenten und keine Exponenten darstellen.)

(ii) $\mathbb{C}^n$, der Raum der komplexen n-Tupel mit ebensolchen Verknüpfungen ist ein $\mathbb{C}$-Vektorraum.

(iii) $\mathbb{Q}^n$ ist ein $\mathbb{Q}$-Vektorraum.

(iv) Sei F ein Körper. Dann ist F^n mit entsprechend definierten Verknüpfungen wie für $\mathbb{R}^n$ ein F-Vektorraum.

(v) Sei F ein Körper. Dann ist die Menge der Folgen reeller Zahlen

$$F^{\mathbb{N}} = \left\{ \left(\xi^0, \xi^1, \xi^2, \ldots\right) : \xi^i \in F \text{ für alle } i \in \mathbb{N} \right\}$$

mit den Verknüpfungen

$$\left(\xi^0, \xi^1, \xi^2, \ldots\right) + \left(\eta^0, \eta^1, \eta^2, \ldots\right) := \left(\xi^0 + \eta^0, \xi^1 + \eta^1, \xi^2 + \eta^2, \ldots\right),$$
$$\lambda \cdot \left(\xi^0, \xi^1, \xi^2, \ldots\right) := \left(\lambda\xi^0, \lambda\xi^1, \lambda\xi^2, \ldots\right)$$

ein $\mathbb{R}$-Vektorraum. Sei nun $F = \mathbb{R}$. Dann bilden die Teilmengen der konvergenten Folgen oder der Nullfolgen ebenfalls einen Vektorraum.

(vi) $\mathbb{C}^n$ ist ein $\mathbb{R}$-Vektorraum mit der üblichen Addition und der Skalarmultiplikation

$$\lambda \left(\xi^1, \ldots, \xi^n\right) := \left(\lambda\xi^1, \ldots, \lambda\xi^n\right).$$

(vii) Sei A eine Menge. Die Menge der Funktionen $f\colon A \to F$ bildet einen F-Vektorraum. Die Verknüpfungen sind

$$(f + g)(x) := f(x) + g(x) \quad \text{für } x \in A,$$
$$(\lambda f)(x) := \lambda(f(x)) \quad \text{für } x \in A.$$

(viii) $\star$ Die Menge der stetigen Funktionen $f\colon [0,1] \to \mathbb{R}$ bildet einen $\mathbb{R}$-Vektorraum $C^0([0,1])$ mit wie für beliebige Funktionen definierter Addition und Skalarmultiplikation. (Details: Analysis.)

(ix) $\star$

$$\{f \in C^0([0,1])\colon f(0) = 0\}$$

bildet mit der üblichen Addition und Multiplikation einen Vektorraum. (Details: Analysis.)

(x) Polynome: Sei F ein Körper. Betrachte in $F^{\mathbb{N}}$ die Menge aller abbrechenden Folgen, die also die Form $(\xi_0, \xi_1, \ldots, \xi_n, 0, 0, \ldots)$ für ein $n \in \mathbb{N}$ haben. Dies ist mit der Addition und der Skalarmultiplikation wie bei Folgen in F ein Vektorraum. Wir schreiben auch mit Hilfe einer Variablen X

$$\xi_0 + \xi_1 X + \xi_2 X^2 + \ldots + \xi_n X^n$$

und bezeichnen den Vektorraum als $F[X]$. Polynome können wir auch multiplizieren:

$$(\xi_0, \xi_1, \ldots) \cdot (\eta_0, \eta_1, \ldots) := (\xi_0 \cdot \eta_0, \xi_0 \cdot \eta_1 + \xi_1 \cdot \eta_0, \xi_0 \cdot \eta_2 + \xi_1 \cdot \eta_1 + \xi_2 \cdot \eta_0, \ldots),$$

wobei allgemein an Stelle k der Eintrag $\sum_{i=0}^{k} \xi_i \eta_{k-i}$ steht. In vertrauterer Notation mit der Variablen X schreiben wir

$$\begin{aligned}&(\xi_0 + \xi_1 X + \xi_2 X^2 + \ldots + \xi_m X^m) \cdot (\eta_0 + \eta_1 X + \eta_2 X^2 + \ldots + \eta_n X^n)\\ =\,&\xi_0\eta_0 + (\xi_0\eta_1 + \xi_1\eta_0)X + \ldots + \left(\sum_{i=0}^{k} \xi_i \eta_{k-i}\right) X^k + \ldots + \xi_m \eta_n X^{m+n}.\end{aligned}$$

Die Multiplikation von Polynomen werden eigentlich erst später im Zusammenhang mit „Ringen“ untersuchen. Wir schreiben z. B. $P(X)$ für ein Polynom und $P(a)$, wenn wir statt der Variablen X überall $a \in F$ einsetzen.

Theorem 2.2.4. *Sei V ein F-Vektorraum. Dann gilt für alle $a, b \in V$ und $\lambda, \mu \in F$*

(i) $0 \cdot a = 0$
(ii) $\lambda \cdot 0 = 0$
(iii) *Aus $\lambda \cdot a = 0$ folgt $a = 0$ oder $\lambda = 0$.*
(iv) $(-\lambda) \cdot a = \lambda \cdot (-a) = -(\lambda \cdot a)$
(v) *Definiere eine Subtraktion durch $a - b := a + (-b)$ und $\lambda - \mu := \lambda + (-\mu)$. Dann gelten*

$$\begin{aligned}\lambda \cdot (a - b) &= \lambda \cdot a - \lambda \cdot b,\\ (\lambda - \mu) \cdot a &= \lambda \cdot a - \mu \cdot a.\end{aligned}$$

Beweis.

(i) Nach (V6) gilt

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

Wir wenden (V4) auf $0 \cdot a$ an und erhalten

$$\begin{aligned}0 &= 0 \cdot a + (-(0 \cdot a))\\ &= (0 \cdot a + 0 \cdot a) + (-(0 \cdot a))\\ &= 0 \cdot a + (0 \cdot a + (-(0 \cdot a)) && \text{nach (V2)}\\ &= 0 \cdot a + 0\\ &= 0 \cdot a && \text{nach (V3).}\end{aligned}$$

(ii) Nach (V5) gilt

$$\lambda \cdot 0 = \lambda \cdot (0 + 0) = \lambda \cdot 0 + \lambda \cdot 0.$$

Wir addieren den zu $\lambda \cdot 0$ inversen Vektor auf beiden Seiten und erhalten

$$\begin{aligned}0 &= \lambda \cdot 0 + (-(\lambda \cdot 0))\\ &= (\lambda \cdot 0 + \lambda \cdot 0) + (-(\lambda \cdot 0)) = \lambda \cdot 0 + (\lambda \cdot 0 + (-(\lambda \cdot 0)))\\ &= \lambda \cdot 0 + 0 && \text{nach (V2)}\\ &= \lambda \cdot 0 && \text{nach (V3).}\end{aligned}$$

(iii) Sei $\lambda \cdot a = 0$ und gelte $\lambda \neq 0$. Zeige also, dass $a = 0$ gilt. Da $\lambda \neq 0$ ist, existiert $\frac{1}{\lambda} \equiv \lambda^{-1}$. Es gilt

$$\begin{aligned}0 &= \frac{1}{\lambda} \cdot 0 && \text{nach (ii)}\\ &= \frac{1}{\lambda}(\lambda \cdot a)\\ &= \left(\frac{1}{\lambda} \cdot \lambda\right) \cdot a && \text{nach (V7)}\\ &= 1 \cdot a\\ &= a && \text{nach (V8).}\end{aligned}$$

(iv) Es gilt

$$\lambda \cdot a + (-\lambda) \cdot a = (\lambda + (-\lambda)) \cdot a = 0 \cdot a = 0 = \lambda \cdot a + (-(\lambda \cdot a)).$$

Wir addieren auf beiden Seiten das Inverse von $\lambda \cdot a$ und erhalten

$$(-\lambda) \cdot a = -(\lambda \cdot a).$$

Ebenso addieren wir zu beiden Seiten von

$$\lambda \cdot a + \lambda \cdot (-a) = \lambda(a + (-a)) = 0 = \lambda \cdot a + (-(\lambda \cdot a))$$

auf beiden Seiten das Inverse von $\lambda \cdot a$ und erhalten $\lambda \cdot (-a) = -(\lambda \cdot a)$. Die Behauptung folgt.

(v) Unter Benutzung von (iv) erhalten wir

$$\lambda \cdot (a - b) = \lambda \cdot (a + (-b)) = \lambda \cdot a + \lambda \cdot (-b) = \lambda \cdot a + (-(\lambda \cdot b)) = \lambda \cdot a - \lambda \cdot b.$$

Ebenso folgt

$$(\lambda - \mu) \cdot a = (\lambda + (-\mu)) \cdot a = \lambda \cdot a + (-\mu) \cdot a = \lambda \cdot a + (-\mu \cdot a) = \lambda \cdot a - \mu \cdot a.$$

□

Bemerkung 2.2.5. Das Assoziativgesetz (V2) gilt auch für mehr als drei Summanden, ebenso gilt das Kommutativgesetz (V1) für mehr als zwei Summanden. Der Beweis ist nicht kompliziert, nur umständlich aufzuschreiben. Man benutzt Induktion nach der Anzahl der Summanden. Beispiele sind:

$$\begin{aligned}(a + (b + c)) + d &= ((a + b) + c) + d = (a + b) + (c + d)\\ &= a + (b + (c + d)) = a + ((b + c) + d),\\ a + b + c &= a + c + b = c + a + b = c + b + a = b + c + a = b + a + c.\end{aligned}$$

Daher wollen wir zukünftig Klammern weglassen und benutzen das Summenzeichen bei assoziativer und kommutativer Summation

$$\sum_{i=1}^{n} a_i = a_1 + a_2 + \ldots + a_n.$$

Dann gelten die üblichen Rechenregeln

(i) $\lambda \cdot \sum_{i=1}^{n} a_i = \sum_{i=1}^{n} \lambda \cdot a_i$ für alle $\lambda \in F$ und $a_1, \ldots, a_n \in V$,

(ii) $\left(\sum_{i=1}^{n} \lambda_i\right) \cdot a = \sum_{i=1}^{n} \lambda_i \cdot a$ für alle $\lambda_1, \ldots, \lambda_n \in F$ und $a \in V$,

(iii) $\sum_{i=1}^{n} \lambda^i a_i + \sum_{i=1}^{n} \mu^i a_i = \sum_{i=1}^{n} \left(\lambda^i + \mu^i\right) a_i$ für alle $\lambda^1, \ldots, \lambda^n, \mu^1, \ldots, \mu^n \in F$ und $a_1, \ldots, a_n \in V$.

Beweis. Übung. □

2.3. **Linearkombinationen.**

Definition 2.3.1 (Linearkombination)**.** Sei V ein F-Vektorraum. Sei $S \subset V$ nichtleer. Gilt

$$a = \sum_{i=1}^{n} \lambda^i a_i \quad \text{für } \lambda^1, \ldots, \lambda^n \in F \quad \text{und } a_1, \ldots, a_n \in S \quad \text{und ein } n \in \mathbb{N},$$

so sagen wir, dass a als **Linearkombination** von Vektoren aus S dargestellt ist. (Beachte, dass die Summen in Linearkombinationen stets endlich sind.)

Die Menge aller Vektoren, die sich als Linearkombination von Vektoren in S darstellen lässt, heißt lineare Hülle von S; wir schreiben $\langle S \rangle$. Ist S eine endliche Menge, so schreiben wir auch $\langle a_1, \ldots, a_m \rangle$ statt $\langle \{a_1, \ldots, a_m\} \rangle$. Wir sagen dann auch, dass $\langle a_1, \ldots, a_m \rangle$ aus den Linearkombinationen von $a_1, \ldots, a_m$ besteht.

Wir vereinbaren weiterhin $\langle\emptyset\rangle := \{0\}$.

Beispiele 2.3.2.

(i) Ist V ein F-Vektorraum und $S = \{a_1, \ldots, a_p\} \subset V$. Dann ist
$$\langle S\rangle = \left\{\lambda^1 a_1 + \ldots \lambda^p a_p \colon \lambda^1, \ldots, \lambda^p \in F\right\}.$$

(ii) Im Vektorraum $V = \mathbb{R}[X]$, der reellen Polynome in X sei
$$S = \{1, X^2, X^4, \ldots, X^{2n}, \ldots\}.$$
Dann ist $\langle S\rangle = \{P(X) \in \mathbb{R}[X] \colon P(a) = P(-a) \text{ für alle } a \in \mathbb{R}\}$.

(iii) Sei $V = \mathbb{R}^3$ und $S = \{(1,1,0), (1,0,1)\}$. Es folgt
$$\begin{aligned}\langle S\rangle &= \{a(1,1,0) + b(1,0,1) \colon a, b \in \mathbb{R}\}\\ &= \{(a+b, a, b) \colon a, b \in \mathbb{R}\}\\ &= \{(x,y,z) \colon x - y - z = 0,\ x, y, z \in \mathbb{R}\}.\end{aligned}$$
Ist $S' = \{(1,1,0), (2,1,1)\}$, so gilt $\langle S\rangle = \langle S'\rangle$.

2.4. Unterräume.

Definition 2.4.1. Sei V ein Vektorraum über F. Eine nichtleere Teilmenge $W \subset V$, so dass für alle $a, b \in W$ und alle $\lambda, \mu \in F$ auch $\lambda a + \mu b \in W$ gilt, heißt **Unterraum** von V.

Bemerkung 2.4.2. Ein Unterraum eines F-Vektorraumes V ist mit der Addition und Skalarmultiplikation von V ein F-Vektorraum.

Beispiele 2.4.3.

(i) Sei V ein Vektorraum. Dann sind V und $\{0\}$ (triviale) Unterräume von V.

(ii) Sei $V = F[X]$. Der Grad eines Polynomes $P(X) = a_0 + a_1 X + \ldots + a_m X^m$ ist das kleinste n, so dass für alle $i > n$ für die Koeffizienten $a_i = 0$ gilt. Wir schreiben $\deg P = n$ und setzen $\deg 0 \equiv -\infty$. Sei $m \leq n$. Der Vektorraum der Polynome vom Grad kleiner oder gleich m ist ein Unterraum der Polynome vom Grad kleiner oder gleich n.

(iii) Sei $c \in [0,1]$. Dann ist $\{f \in C([0,1]) \colon f(c) = 0\}$ ein Unterraum von $C([0,1])$.

(iv) Die Menge der Lösungen eines linearen homogenen Gleichungssystems in n Variablen sind ein Unterraum von F^n.

Hier und im folgenden betrachten wir Vektorräume über einem beliebigen aber festen Körper F, falls dies nicht ausdrücklich anders angegeben ist.

Theorem 2.4.4. *Sei V ein Vektorraum, $S \subset V$. Dann ist $\langle S\rangle$ der kleinste Unterraum von V, der S enthält.*

Beweis. Wegen $\sum_{\emptyset} = 0$ ist $\langle S\rangle \neq \emptyset$. $\langle S\rangle$ ist ein Unterraum, denn für $a, b \in \langle S\rangle$ mit $a = \sum_{i\in I} a^i s_i$ und $b = \sum_{j\in J} b^j s_j$ für endliche Mengen I, J und $\lambda, \mu \in F$ gilt auch
$$\lambda a + \mu b = \sum_{i\in I\cap J} \left(\lambda a^i + \mu b^i\right) s_i + \sum_{i\in I\setminus J} \lambda a^i s_i + \sum_{j\in J\setminus I} \mu b^j s_j \in \langle S\rangle.$$
Somit ist $\langle S\rangle$ ein Unterraum von V. $\langle S\rangle$ enthält insbesondere die Vektoren von S und daher auch alle Linearkombinationen von Vektoren aus S.

Sei $W \subset V$ ein beliebiger Unterraum von V mit $S \subset W$. Dann enthält W nach Definition auch alle Linearkombinationen von Vektoren aus S und somit $\langle S\rangle$. Die Behauptung folgt. □

Theorem 2.4.5. *Sei V ein Vektorraum, $S \subset V$, $b \in V$. Sei b eine Linearkombination von Vektoren aus S. Dann gilt $\langle S \cup \{b\}\rangle = \langle S\rangle$.*

Beweis. Es gilt $S \cup \{b\} \subset \langle S \rangle$. $\langle S \rangle$ ist ein Vektorraum. Somit folgt nach Theorem 2.4.4 $\langle S \cup \{b\} \rangle \subset \langle S \rangle$.
$\langle S \rangle \subset \langle S \cup \{b\} \rangle$ ist klar. □

Definition 2.4.6. Wir sagen, dass $\langle S \rangle$ von S erzeugt ist. Im Fall $\langle S \rangle = V$ heißt S ein **Erzeugendensystem** von V. Besitzt ein Vektorraum V ein endliches Erzeugendensystem, so heißt V **endlich erzeugt** bzw. endlich erzeugbar.

Beispiele 2.4.7.

(i) Sei $A \in F^{m \times n}$. Fassen wir die Spalten von A als Vektoren in F^m auf, so nennen wir den von ihnen erzeugten Unterraum den Spaltenraum von A. Ebenso ist der Zeilenraum von A der von den Zeilen von A in F^n erzeugte Unterraum.

(ii) Sei $V = \mathbb{R}^n$, $S = \{e_1, \ldots, e_n\}$ mit

$$\begin{aligned} e_1 &= (1, 0, 0, \ldots, 0), \\ e_2 &= (0, 1, 0, \ldots, 0), \\ e_3 &= (0, 0, 1, \ldots, 0), \\ &\vdots \\ e_n &= (0, 0, 0, \ldots, 1). \end{aligned}$$

(Für später eigentlich besser als Spaltenvektoren geschrieben.) Dann ist $\langle S \rangle = V$ und somit ist S ein Erzeugendensystem von V. Für $\xi = (\xi^1, \ldots, \xi^n)$ gilt $\xi = \sum_{i=1}^{n} \xi^i e_i$. Somit ist $\mathbb{R}^n$ endlich erzeugbar.

(iii) Sei $V = \mathbb{R}[X]$. Dann ist $S = \{1, X, X^2, \ldots\}$ ein Erzeugendensystem, da sich jedes Polynom als Linearkombination von Termen der Form X^i schreiben lässt. $\mathbb{R}[X]$ ist aber nicht endlich erzeugbar, da ein endliches Erzeugendensystem nicht Polynome beliebigen Grades erzeugen kann.

Theorem 2.4.8. *Sei $(W_i)_{i \in I}$ eine beliebige Familie von Unterräumen von V. Dann ist $W := \bigcap_{i \in I} W_i$ ein Unterraum von V.*

Beweis. Der Schnitt ist nicht leer, da jeder Unterraum die Null enthält. Seien $a, b \in W$, also $a, b \in W_i$ für alle $i \in I$. Seien $\lambda, \mu \in F$. Dann gilt $\lambda a + \mu b \in W_i$ für alle $i \in I$ und somit ist auch $\lambda a + \mu b \in W$. □

Im Allgemeinen ist die Vereinigung von Unterräumen eines Vektorraumes kein Unterraum mehr. Die folgende Konstruktion macht aus mehreren Unterräumen einen Unterraum, der die Vereinigung aller dieser Unterräume enthält. Dies funktioniert auch für beliebig viele Unterräume; dann erklärt man die Summe als die Menge beliebiger endlicher Summen von Vektoren aus den Unterräumen.

Definition 2.4.9. Seien $W_1, W_2, \ldots, W_p$ Unterräume eines F-Vektorraumes V. Definiere deren Summe durch

$$W_1 + W_2 + \ldots + W_p := \{a_1 + a_2 + \ldots + a_p \colon a_i \in W_i,\ i = 1, \ldots p\}.$$

Theorem 2.4.10. *Sei V ein F-Vektorraum. Seien $W_1, \ldots, W_p$ Unterräume von V. Dann ist $W := W_1 + \ldots + W_p$ der kleinste Unterraum von V, der jedes W_i enthält.*

Beweis. Zeige zunächst, dass W ein Unterraum von V ist: Seien $a = a_1 + \ldots + a_p$, $b = b_1 + \ldots + b_p \in W$ mit $a_i, b_i \in W_i$. Seien $\lambda, \mu \in F$. Dann ist

$$\lambda a + \mu b = (\lambda a_1 + \mu b_1) + (\lambda a_2 + \mu b_2) + \ldots + (\lambda a_p + \mu b_p).$$

Somit ist $\lambda a + \mu b \in W$ und W ist ein (nichtleerer) Unterraum.

Wegen $0 + \ldots + 0 + a_i + 0 + \ldots + 0 \in W$ für $a_i \in W_i$ gilt $W_i \subset W$. Sei W' ein weiterer Unterraum von V mit $W_i \subset W'$ für alle i. Zeige, dass $W \subset W'$ gilt: Sei $a_1 + \ldots + a_p \in W$ beliebig. Da $a_i \in W_i \subset W'$ ist, folgt auch $a_1 + \ldots + a_n \in W'$, denn W' ist ein Unterraum. Wir erhalten $W \subset W'$. □

Für das nächste Kapitel über lineare Gleichungssysteme benötigen wir

Definition 2.4.11. Sei V ein F-Vektorraum. Sei $A \subset V$. Dann heißt A ein **affiner Unterraum** von V, falls es einen Unterraum $U \subset V$ und einen Vektor $x \in V$ mit

$$A = x + U \equiv \{x + y \colon y \in U\}$$

gibt.

3. Lineare Gleichungssysteme

3.1. Lineare Gleichungssysteme.

Beispiel 3.1.1. ⋆ Wir wollen das folgende Gleichungssystem reeller Zahlen lösen:

$$\begin{array}{rcrcrcr} 2x & + & 6y & - & 4z & = & 10, \\ -x & + & 5y & + & 2z & = & 11, \\ 3x & + & 7y & + & z & = & -3. \end{array}$$

Dazu multiplizieren wir die erste Zeile mit $\frac{1}{2}$.

$$\begin{array}{rcrcrcr} x & + & 3y & - & 2z & = & 5, \\ -x & + & 5y & + & 2z & = & 11, \\ 3x & + & 7y & + & z & = & -3. \end{array}$$

Wir addieren die erste Zeile zur zweiten Zeile und subtrahieren sie, mit 3 multipliziert, von der dritten Zeile

$$\begin{array}{rcrcrcr} x & + & 3y & - & 2z & = & 5, \\ & & 8y & & & = & 16, \\ & - & 2y & + & 7z & = & -18. \end{array}$$

Multipliziere die zweite Zeile mit $\frac{1}{8}$ und addiere das Resultat, mit 2 multipliziert, zur dritten Zeile

$$\begin{array}{rcrcrcr} x & + & 3y & - & 2z & = & 5, \\ & & y & & & = & 2, \\ & & & & 7z & = & -14. \end{array}$$

Multipliziere die dritte Zeile mit $\frac{1}{7}$ und addiere das Resultat, mit 2 multipliziert, zur ersten Zeile. Addiere die zweite Zeile, mit -3 multipliziert, zur ersten Zeile und erhalte

$$\begin{array}{rcrcr} x & & & = & -5, \\ & & y & = & 2, \\ & & z & = & -2. \end{array}$$

Wir erhalten die Lösung $(x, y, z) = (-5, 2, -2)$. D. h. für $x = -5$, $y = 2$ und $z = -2$ sind sämtliche Gleichungen des obigen Gleichungssystems erfüllt.

Wir wollen jedoch nicht in erster Linie spezielle Probleme lösen, sondern Aussagen über allgemeine Probleme machen.

Definition 3.1.2 (Lineares Gleichungssystem)**.**

(i) Sei F ein Körper. Ein **lineares Gleichungssystem** in n Unbekannten x^1, …, x^n und m Gleichungen ist ein System von Gleichungen der Form

$$\begin{array}{rcrcrcrcr} a_1^1 x^1 & + & a_2^1 x^2 & + & \ldots & + & a_n^1 x^n & = & b^1 \\ a_1^2 x^1 & + & a_2^2 x^2 & + & \ldots & + & a_n^2 x^n & = & b^2 \\ \vdots & & & & & & \vdots & & \vdots \\ a_1^m x^1 & + & a_2^m x^2 & + & \ldots & + & a_n^m x^n & = & b^m \end{array} \tag{3.1}$$

Dabei sind a^i_j und b^i, $1 \leq i \leq m$, $1 \leq j \leq n$, Elemente von F, also $a^i_j, b^i \in F$.

(ii) Das System heißt **homogen,** falls $b^1 = \ldots = b^m = 0$ gilt und sonst **inhomogen.**

(iii) Ein n-Tupel $(\xi^1, \ldots, \xi^n) \in F^n$ heißt **Lösung** des linearen Gleichungssystems, wenn alle Gleichungen des Gleichungssystems erfüllt sind, wenn wir die Zahlen ξ^i statt x^i für $1 \leq i \leq n$ einsetzen.

Bemerkung 3.1.3. Anders als in vielen Büchern über lineare Algebra schreiben wir einen Index der Koeffizienten a^i_j nach oben. (a^i_j hier entspricht a_{ij} in solchen Büchern.) Dies bedeutet keine Potenzierung und ist eine in der Differentialgeometrie und Physik (Einsteinsche Summenkonvention) übliche Schreibweise.

Theorem 3.1.4. *Sei (H) ein homogenes lineares Gleichungssystem wie in* (3.1)*. Dann bildet die Menge der Lösungen von (H) einen Unterraum von F^n.*

Beweis.

(i) Es ist klar, dass $0 = (0, \ldots, 0) \in F^n$ eine Lösung von (H) ist.

(ii) Seien $(\xi^1, \ldots, \xi^n)$ und $(\eta^1, \ldots, \eta^n)$ Lösungen von (H). Sei $\lambda \in F$. Nach Voraussetzung gilt für alle $1 \leq i \leq m$

$$a^i_1\xi^1 + a^i_2\xi^2 + \ldots + a^i_n\xi^n = 0, \tag{3.2}$$

$$a^i_1\eta^1 + a^i_2\eta^2 + \ldots + a^i_n\eta^n = 0. \tag{3.3}$$

Wir addieren (3.2) und (3.3) und erhalten

$$a^i_1(\xi^1 + \eta^1) + a^i_2(\xi^2 + \eta^2) + \ldots + a^i_n(\xi^n + \eta^n) = 0.$$

Somit ist $(\xi^1 + \eta^1, \ldots, \xi^n + \eta^n)$ ebenfalls eine Lösung von (H).

(iii) Multipliziere (3.2) mit λ und erhalte

$$a^i_1(\lambda\xi^1) + a^i_2(\lambda\xi^2) + \ldots + a^i_n(\lambda\xi^n) = 0.$$

Daher ist $(\lambda\xi^1, \ldots, \lambda\xi^n)$ auch eine Lösung von (H). □

Bemerkung 3.1.5. Sei (L) ein lineares Gleichungssystem wie in (3.1). Dann bezeichnen wir das homogene lineare Gleichungssystem, das aus (3.1) entsteht, wenn wir die Koeffizienten b^i durch 0 ersetzen als das zugehörige homogene System (HL).

Bemerkung 3.1.6. Für n-Tupel verwenden wir die Abkürzungen

$$\eta = (\eta^1, \eta^2, \ldots, \eta^n),$$
$$\zeta = (\zeta^1, \zeta^2, \ldots, \zeta^n), \ldots,$$

und definieren

$$\eta + \zeta = (\eta^1 + \zeta^1, \eta^2 + \zeta^2, \ldots, \eta^n + \zeta^n)$$

sowie für $\lambda \in F$

$$\lambda\eta = (\lambda\eta^1, \lambda\eta^2, \ldots, \lambda\eta^n).$$

Theorem 3.1.7. *Sei (L) ein lineares Gleichungssystem wie in* (3.1) *und bezeichne (HL) das zugehörige homogene lineare Gleichungssystem. Dann ist die Menge M aller Lösungen von (L) entweder die leere Menge oder ein affiner Unterraum des F^n der Form $M = \zeta + M_0$, wobei M_0 die Lösungsmenge von (HL) und ζ eine beliebige Lösung von (L) sind.*

Beweis. Sei $M \neq \emptyset$. Wähle $\xi \in M$ beliebig.

(i) Sei $\xi \in M_0$ beliebig. Addiere

$$a_1^i\zeta^1 + a_2^i\zeta^2 + \ldots + a_n^i\zeta^n = b^i$$

und

$$a_1^i\xi^1 + a_2^i\xi^2 + \ldots + a_n^i\xi^n = 0$$

und erhalte

$$a_1^i(\zeta^1 + \xi^1) + a_2^i(\zeta^2 + \xi^2) + \ldots + a_n^i(\zeta^n + \xi^n) = b^i.$$

Somit folgt $\zeta + M_0 \subset M$.

(ii) Umgekehrt ist zu zeigen, dass für jede Lösung $x \in M$ der Vektor $\eta - \zeta = \eta + (-1)\zeta$ ein Element von M_0 ist. Subtrahiere dazu

$$a_1^i\zeta^1 + a_2^i\zeta^2 + \ldots + a_n^i\zeta^n = b^i$$

von

$$a_1^i\eta^1 + a_2^i\eta^2 + \ldots + a_n^i\eta^n = b^i$$

und erhalte

$$a_1^i(\eta^1 - \zeta^1) + a_2^i(\eta^2 - \zeta^2) + \ldots + a_n^i(\eta^n - \zeta^n) = 0.$$

Wir erhalten die zweite Behauptung. □

Lemma 3.1.8. *Sei (L) ein lineares Gleichungssystem. Dann ist die Lösungsmenge unter den folgenden Operationen invariant, d. h. unverändert.*

(i) Vertauschen von Gleichungen.
(ii) Multiplikation einer Gleichung mit $\lambda \in F$, $\lambda \neq 0$.
(iii) Wir addieren das λ-fache der j-ten Gleichung zur i-ten Gleichung.

Beweis. Sei L_1 die Lösungsmenge von (L) und sei L_2 die Lösungsmenge des Systems, das wir erhalten, wenn wir die jeweilige Operation anwenden.

(i) Klar.
(ii) Für jede Zahl $\lambda \in F$ gilt $L_1 \subset L_2$. Ist $\lambda \neq 0$, so erhalten wir durch Multiplikation derselben Gleichung mit $\frac{1}{\lambda}$ gerade wieder (L). Somit gilt auch $L_2 \subset L_1$, insgesamt also $L_1 = L_2$.
(iii) Auch hier ist $L_1 \subset L_2$ wieder klar. Addition des $(-\lambda)$-fachen der j-ten Gleichung zur i-ten Gleichung liefert wieder (L), also gilt auch $L_2 \subset L_1$ und daher ist auch hier wieder $L_1 = L_2$. □

3.2. **Gaußsches Eliminationsverfahren.** Das Gaußsche Eliminationsverfahren erlaubt es, allgemeine lineare Gleichungssysteme zu lösen oder nachzuweisen, dass sie keine Lösung besitzen.

Definition 3.2.1 (Zeilenstufenform)**.** Sei (L) ein lineares Gleichungssystem in der Form (3.1). Dann ist (L) in Zeilenstufenform, falls „in jeder Zeile links mehr Nullen als in der Vorgängerzeile stehen“, d. h. setzen wir

$$A_j^i := \begin{cases} a_j^i, & 1 \leq j \leq n,\ 1 \leq i \leq m, \\ b^i, & j = n+1,\ 1 \leq i \leq m, \end{cases}$$

so ist die Anzahl der „links stehenden Nullen“ in Zeile i durch

$$N(i) := \max\left\{k \in \{0, 1, \ldots, n+1\} \colon A_j^i = 0 \text{ für } 1 \leq j \leq k\right\}$$

definiert und wir fordern, dass es ein $i_0 \in \{0, \ldots, m-1\}$ gibt, so dass $N(i) < N(i+1)$ für alle $1 \leq i \leq i_0$ und $N(i) = n+1$ für $i_0 < i \leq m$ gelten.

Wir verwenden

$$\mathbb{N} = \{0, 1, 2, \ldots\} \quad \text{und} \quad \mathbb{N}^+ \equiv \mathbb{N}_{>0} \equiv \mathbb{N} \setminus \{0\}.$$

Theorem 3.2.2 (Gaußsches Eliminationsverfahren)**.**
Sei (L) ein lineares Gleichungssystem wie in (3.1)*. Dann lässt sich (L) mit endlich vielen Operationen aus Lemma 3.1.8 in ein lineares Gleichungssystem (G) der Form* (3.1) *umformen, so dass (G) in Zeilenstufenform ist. (L) und (G) haben dieselbe Lösungsmenge.*

Beweis. Die Gleichheit der Lösungsmengen folgt direkt nach Lemma 3.1.8.

Wir zeigen nun per Induktion nach der Anzahl der Spalten bzw. der Variablen, dass dies stets möglich ist.

Wir betrachten zuerst den Fall einer Variablen. Dann haben wir ein Gleichungssystem der Form

$$\begin{aligned} a_1^1 x^1 &= b^1, \\ \vdots \quad & \quad \vdots, \\ a_1^m x^1 &= b^m. \end{aligned}$$

Ist ein $a_1^j \neq 0$, so dürfen wir nach Vertauschung der Zeilen annehmen, dass $a_1^1 \neq 0$ gilt. Addiere nun $-\frac{a_1^2}{a_1^1}$-mal die erste Zeile zur zweiten, $-\frac{a_1^i}{a_1^1}$-mal die erste Zeile zur i-ten. Damit haben wir das obige System auf eine Form mit $a_1^i = 0$ für $2 \leq i \leq m$ gebracht. Wir behalten die Bezeichnungen bei. Ist ein $b^i \neq 0$, so dürfen wir nach Zeilenvertauschung ohne Einschränkung annehmen, dass dies b^2 ist. Falls für alle $2 \leq i \leq m$ auch $b^i = 0$ gilt, so sind wir fertig. Nehme daher an, dass dieser Fall nicht eintritt. Wir addieren nun $-\frac{b^3}{b^2}$-mal die zweite Zeile zur dritten und $-\frac{b^i}{b^2}$-mal die zweite Zeile zur i-ten. Das Ergebnis ist ein Gleichungssystem in Zeilenstufenform.

Sei nun die Anzahl der Variablen n in (3.1) größer als eins. Nach Umordnen der Gleichungen dürfen wir ohne Einschränkung annehmen, dass $a_1^1 \neq 0$ gilt. (Gibt es nämlich keinen Koeffizienten a_1^i in der ersten Spalte, der nicht verschwindet, so folgt die Behauptung per Induktion.) Addiere nun $-\frac{a_1^i}{a_1^1}$-mal die erste Zeile zur i-ten. Wir bezeichnen die Koeffizienten im neuen linearen Gleichungssystem wieder mit a_j^i. Dann gilt $a_1^i = 0$ für $2 \leq i \leq m$. Nach Induktionsvoraussetzung können wir den Block ab der zweiten Gleichung in einem System der Form

$$\begin{array}{ccccccccc} a_1^1 x^1 & + & a_2^1 x^2 & + & \ldots & + & a_n^1 x^n & = & b^1 \\ & & \tilde{a}_2^2 x^2 & + & \ldots & + & \tilde{a}_n^2 x^n & = & \tilde{b}^2 \\ & & \vdots & & & & \vdots & & \vdots \\ & & \tilde{a}_2^m x^2 & + & \ldots & + & \tilde{a}_n^m x^n & = & \tilde{b}^m. \end{array}$$

in Zeilenstufenform bringen. Die Behauptung folgt daher.

Nach Multiplikation der i-ten Zeile mit $\frac{1}{A^i_{N(i)+1}}$ wie in Definition 3.2.1 dürfen wir zusätzlich ohne Einschränkung annehmen, dass das Gleichungssystem in Zeilenstufenform $A^i_{N(i)+1} = 1$ erfüllt (falls $N(i) \leq n$ ist), dass also der erste von Null verschiedene Koeffizient gleich 1 ist, falls die Gleichung nicht $0 = 0$ lautet. □

Lemma 3.2.3. *Sei (L) ein lineares Gleichungssystem der Form* (3.1) *in Zeilenstufenform.*

(i) Dann besitzt (L) genau dann eine Lösung, wenn (L) keine Zeile der Form $0 = b^i$ mit $b^i \neq 0$ enthält.

(ii) Die Lösung ist eindeutig, falls eine Lösung existiert, $N(i) = i-1$ für $1 \leq i \leq m$ mit $N(i)$ wie in Definition 3.2.1 und $n = m$ gelten.

Beweis.

(i) Damit (L) eine Lösung besitzt, darf keine Zeile der Form $0 \neq 0$ auftreten. Nehme dies an. Dann hat die letzte Zeile, die nicht die Form $0 = 0$ hat, die Form

$$a_k^i x^k + \ldots + a_n^i x^n = b^i$$

mit $a_k^i \neq 0$. Wähle $\xi^{k+1}, \ldots, \xi^n$ beliebig und setze

$$\xi^k := \frac{1}{a_k^i} \left(b^i - \left(a_{k+1}^i \xi^{k+1} + \ldots + a_n^i \xi^n \right) \right).$$

$(\xi^k, \ldots, \xi^n)$ löst diese letzte Gleichung. Ersetze nun $x^k, \ldots, x^n$ im restlichen Gleichungssystem durch $\xi^k, \ldots, \xi^n$. Dieses Gleichungssystem hat nun eine Gleichung weniger als das bisher betrachtete. Per Induktion finden wir daher eine Lösung.

(ii) Gilt $N(1) = 0$, $N(2) = 1, \ldots$, so sind sämtliche ξ^i beim Lösen des Gleichungssystems von unten eindeutig bestimmt. Daher ist die Lösung eindeutig. □

Korollar 3.2.4. *Sei (H) ein homogenes lineares Gleichungssystem der Form* (3.1) *mit $m < n$, also mit weniger Zeilen als Variablen. Dann besitzt (H) eine nichttriviale Lösung, d. h. eine Lösung $(\xi^1, \ldots, \xi^n) \neq (0, \ldots, 0)$.*

Beweis. Wir bringen das Gleichungssystem in Zeilenstufenform. Wegen $m < n$ gibt es ein $i < m$ mit $N(i) + 2 \leq N(i+1)$ und $N(i) + 2 \leq n$ oder es ist $N(m) \leq n - 2$. Lösen wir das Gleichungssystem also wie in Lemma 3.2.3, so ist $\xi^{N(i)+2}$ bzw. ξ^n frei wählbar. □

Korollar 3.2.5. *Sei (L) ein beliebiges lineares Gleichungssystem wie in* (3.1) *mit $m = n$, also mit genausovielen Gleichungen wie Variablen. Dann gilt*

(i) Besitzt (HL) nur die triviale Lösung, so besitzt (L) genau eine Lösung.

(ii) Besitzt (HL) eine nichttriviale Lösung, so besitzt (L) entweder keine oder mindestens so viele Lösungen wie der Körper Elemente besitzt.

(Dieser Teil gilt mit unverändertem Beweis auch für $m \neq n$.)

Äquivalent dazu ist, dass (L) entweder eindeutig lösbar ist, oder (HL) eine nichttriviale Lösung besitzt.

Beweis. Verfahren wir wie in Theorem 3.2.2, so erhalten wir dasselbe homogene lineare Gleichungssystem, egal ob wir zunächst (L) auf Zeilenstufenform bringen und dann das homogene lineare Gleichungssystem dazu betrachten oder ob wir (HL) auf Zeilenstufenform bringen. Daher dürfen wir ohne Einschränkung annehmen, dass (L) bereits in Zeilenstufenform ist.

(i) Besitzt (HL) nur die triviale Lösung, so hat (L) die Gestalt

$$\begin{array}{ccccccccc} a_1^1 x^1 & + & a_2^1 x^2 & + & \ldots & + & a_n^1 x^n & = & b^1 \\ & & a_2^2 x^2 & + & \ldots & + & a_n^2 x^n & = & b^2 \\ & & & & \ddots & & \vdots & & \vdots \\ & & & & & & a_n^n x^n & = & b^n. \end{array}$$

mit $a_i^i \neq 0$ für $1 \leq i \leq n$, hat also „obere Dreiecksgestalt“. Nach Lemma 3.2.3 besitzt (L) daher eine eindeutige Lösung.

(ii) Sei $\eta = (\eta^1, \ldots, \eta^n)$ eine nichttriviale Lösung des homogenen linearen Gleichungssystems (HL), also eine Lösung mit $(\eta^1, \ldots, \eta^n) \neq (0, \ldots, 0)$. Nehme an, dass (L) eine Lösung $\xi = (\xi^1, \ldots, \xi^n)$ besitzt. Dann sind $\xi + \lambda\eta$ für alle $\lambda \in F$ Lösungen von (L). Die Behauptung folgt. □

3.3. **Matrizen.**

Definition 3.3.1. Eine $(m \times n)$-**Matrix** A ist ein Element in $F^{m \cdot n}$. Wir schreiben $A \in F^{m \times n}$. A hat $m \cdot n$ reelle **Komponenten,** die wir mit a^i_j, $1 \leq i \leq m$, $1 \leq j \leq n$, bezeichnen: $A = (a^i_j)$ oder genauer $A = (a^i_j)_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$. Wir stellen Matrizen wie folgt graphisch dar:

$$A = \begin{pmatrix} a^1_1 & a^1_2 & a^1_3 & \dots & a^1_n \\ a^2_1 & a^2_2 & a^2_3 & \dots & a^2_n \\ \vdots & \vdots & \vdots & & \vdots \\ a^m_1 & a^m_2 & a^m_3 & \dots & a^m_n \end{pmatrix}.$$

Der obere Index bezieht sich auf die Zeilen, der untere auf die Spalten der Matrix. a^i_j steht daher in Zeile i und Spalte j.

(i) Seien A, B zwei $(m \times n)$-Matrizen mit Komponenten a^i_j bzw. b^i_j. Wir definieren die Summe durch

$$A + B := \left(a^i_j + b^i_j\right),$$

also durch komponentenweise Addition.

(ii) Sei $\lambda \in F$. Wir definieren das Produkt einer Matrix mit einer reellen Zahl komponentenweise durch

$$\lambda A := \left(\lambda a^i_j\right).$$

Bemerkung 3.3.2. Wir können n-Tupel $\xi \in F^n$ als $(1 \times n)$-Matrizen auffassen. Dann stimmen die Operationen Summieren und Produktbildung mit einer reellen Zahl mit den für n-Tupel bereits definierten Operationen überein.

Definition 3.3.3. Eine $(n \times n)$-Matrix heißt **quadratische Matrix.** Die $(n \times n)$-Matrix

$$I = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

heißt **Einheitsmatrix** der Ordnung n. Man schreibt auch $I = \mathbf{1}$. Wir definieren das **Kroneckersymbol** durch

$$\delta^i_j := \begin{cases} 1, & i = j, \\ 0, & \text{sonst.} \end{cases}$$

Dann gilt $I = \left(\delta^i_j\right)$.

Eine quadratische Matrix heißt **Diagonalmatrix,** falls $a^i_j = 0$ für $i \neq j$ gilt. Die Elemente a^i_i heißen **Diagonalelemente.**

Definition 3.3.4 (Matrixmultiplikation)**.** Sei A eine $(m \times n)$-Matrix und B eine $(n \times p)$-Matrix, wie bisher mit Einträgen (a^i_j) und (b^j_k). Wenn die Spaltenzahl von A mit der Zeilenzahl von B wie angegeben übereinstimmt, so definieren wir das Produkt $C = AB$, eine $(m \times p)$-Matrix (c^i_k), für $1 \leq i \leq m$ und $1 \leq k \leq p$ durch

$$c^i_k := \sum_{j=1}^{n} a^i_j b^j_k \equiv a^i_1 b^1_k + a^i_2 b^2_k + \ldots + a^i_n b^n_k.$$

(Physiker benutzen die Einsteinsche Summenkonvention und schreiben $c^i_k = a^i_j b^j_k$.)

Beispiele 3.3.5. Es gilt

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

$$\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ -1 & 0 \\ 2 & 1 \end{pmatrix} \quad \text{ist nicht definiert,}$$

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

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

Wie das letzte Beispiel zeigt, gilt im allgemeinen nicht $AB = BA$, selbst, wenn beide Verknüpfungen definiert sind. Gilt $AB = BA$, so sagen wir, dass die beiden Matrizen vertauschen oder **kommutieren.**

Lemma 3.3.6. *Sei A eine $(m \times n)$-Matrix, B, C seien $(n \times p)$-Matrizen, D eine $(p \times q)$-Matrix und $\lambda \in F$. Dann gelten die Rechenregeln*

$$\begin{aligned} A(B+C) &= AB + AC, \\ (\lambda A)B &= \lambda(AB) = A(\lambda B), \\ A(BD) &= (AB)D. \end{aligned}$$

Beweis. Es gilt

$$\begin{aligned} A(B+C) &= \left(\sum_{j=1}^{n} a_j^i (b_k^j + c_k^j) \right)_{\substack{1 \le i \le m \\ 1 \le k \le p}} = \left(\sum_{j=1}^{n} a_j^i b_k^j \right)_{\substack{1 \le i \le m \\ 1 \le k \le p}} + \left(\sum_{j=1}^{n} a_j^i c_k^j \right)_{\substack{1 \le i \le m \\ 1 \le k \le p}} \\ &= AB + AC, \\ (\lambda A)B &= \left(\sum_{j=1}^{n} (\lambda a_j^i) b_k^j \right)_{\substack{1 \le i \le m \\ 1 \le k \le p}} = \underbrace{\left(\sum_{j=1}^{n} a_j^i (\lambda b_k^j) \right)_{\substack{1 \le i \le m \\ 1 \le k \le p}}}_{=A(\lambda B)} = \underbrace{\lambda \left(\sum_{j=1}^{n} a_j^i b_k^j \right)_{\substack{1 \le i \le m \\ 1 \le k \le p}}}_{=\lambda(AB)}, \\ A(BD) &= \left(\sum_{j=1}^{n} a_j^i \left(\sum_{k=1}^{p} b_k^j d_l^k \right) \right)_{\substack{1 \le i \le m \\ 1 \le l \le q}} = \left(\sum_{k=1}^{p} \left(\sum_{j=1}^{n} a_j^i b_k^j \right) d_l^k \right)_{\substack{1 \le i \le m \\ 1 \le l \le q}} \\ &= (AB)D. \end{aligned}$$

□

Bemerkung 3.3.7. Sei

$$\sum_{k=1}^{n} a_k^i x^k = b^i, \quad 1 \le i \le m,$$

ein lineares Gleichungssystem. Wir definieren die Koeffizientenmatrix A und die augmentierte Koeffizientenmatrix A' zu diesem linearen Gleichungssystem durch

$$A = \begin{pmatrix} a_1^1 & a_2^1 & \dots & a_n^1 \\ a_1^2 & a_2^2 & \dots & a_n^2 \\ \vdots & \vdots & & \vdots \\ a_1^m & a_2^m & \dots & a_n^m \end{pmatrix} \quad \text{und} \quad A' = \begin{pmatrix} a_1^1 & a_2^1 & \dots & a_n^1 & b^1 \\ a_1^2 & a_2^2 & \dots & a_n^2 & b^2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_1^m & a_2^m & \dots & a_n^m & b^m \end{pmatrix}.$$

(In A' verwendet man gelegentlich auch einen senkrechten Strich vor der letzten Spalte.) Definiere weiterhin

$$b = \begin{pmatrix} b^1 \\ b^2 \\ \vdots \\ b^m \end{pmatrix} \quad \text{und} \quad x = \begin{pmatrix} x^1 \\ x^2 \\ \vdots \\ x^n \end{pmatrix}.$$

Dann können wir das Gleichungssystem in der Form

$$Ax = b$$

schreiben.

4. Struktur von Vektorräumen

4.1. Lineare Unabhängigkeit.

Definition 4.1.1.

(i) Sei $\{a_1, \ldots, a_n\}$ eine endliche Familie von Vektoren in einem F-Vektorraum. Dann heißt $\{a_1, \ldots, a_n\}$ **linear unabhängig,** wenn aus

$$\sum_{i=1}^{n} \lambda^i a_i = 0$$

für $\lambda^i \in F$ bereits $\lambda^1 = \ldots = \lambda^n = 0$ folgt. Andernfalls heißt diese Familie **linear abhängig.**

Wir sagen auch, dass die Vektoren $a_1, \ldots, a_n$ linear unabhängig sind.

(ii) Eine beliebige Familie heißt linear unabhängig, wenn dies für jede endliche Teilfamilie $\{a_1, \ldots, a_n\}$ gilt. Andernfalls heißt sie linear abhängig.

Beispiele 4.1.2.

(i) $\{a\}$ ist genau dann linear unabhängig, wenn $a \neq 0$ gilt.
(ii) Ist S linear abhängig und $S \subset T$, so ist T linear abhängig.
(iii) $\{a + tb \colon t \in \mathbb{R}\}$ beschreibt genau dann eine Gerade, die den Ursprung 0 nicht enthält, wenn a, b linear unabhängig sind.
(iv) Die Menge $\{(2,-1,1),(1,0,0),(0,2,1)\} \equiv \{a_1, a_2, a_3\}$ ist linear unabhängig, denn für eine Linearkombination der Null gilt $0 = \lambda^1 a_1 + \lambda^2 a_2 + \lambda^3 a_3$, also

$$\begin{array}{rcrcrcl} 2\lambda^1 & + & \lambda^2 & + & & = & 0, \\ -\lambda^1 & & & + & 2\lambda^3 & = & 0, \\ \lambda^1 & & & + & \lambda^3 & = & 0 \end{array}$$

und diese lineare Gleichungssystem besitzt nur die Nulllösung.
(v) Sei $e_i = (0, \ldots, 0, 1, 0, \ldots, 0)$ mit Eintrag 1 an der Stelle i. Dann sind die Vektoren $\{e_1, \ldots, e_n\}$ in F^n linear unabhängig.
(vi) $\{1, X, X^2, X^3, \ldots\}$ sind in $\mathbb{R}[X]$ linear unabhängig.

Theorem 4.1.3. *Eine Familie von Vektoren $\{a_1, \ldots, a_n\}$ ist genau dann linear unabhängig, wenn sich keiner der Vektoren als Linearkombination der anderen schreiben lässt.*

Beweis.

(i) Sei zunächst ohne Einschränkung

$$a_1 = \sum_{i=2}^{n} \lambda^i a_i$$

für $\lambda^i \in F$, $i = 2, \ldots, n$. Setze $\lambda^1 := -1$. Dann gilt

$$\sum_{i=1}^{n} \lambda^i a_i = 0$$

mit $\lambda^1 = -1 \neq 0$. Somit ist die Familie $\{a_1, \ldots, a_n\}$ linear abhängig.

(ii) Gelte

$$\sum_{i=1}^{n} \lambda^i a_i = 0$$

und sei ohne Einschränkung $\lambda^1 \neq 0$. Dann folgt

$$a_1 = -\frac{1}{\lambda^1} \sum_{i=2}^{n} \lambda^i a_i.$$

Somit haben wir a_1 als Linearkombination der übrigen Vektoren dargestellt.

□

Theorem 4.1.4. *Eine Familie S von Vektoren ist genau dann linear unabhängig, wenn jedes $a \in \langle S \rangle$ nur auf genau eine Art und Weise linear aus den Vektoren aus S kombiniert werden kann.*

Beweis.

(i) Insbesondere ist $0 \in \langle S \rangle$. Der Nullvektor lässt sich als Linearkombination mit Koeffizienten 0 schreiben. Da dies nach Voraussetzung die einzige Möglichkeit ist, sind die Vektoren in S linear unabhängig.

(ii) Sei ein Vektor auf zwei verschiedene Arten dargestellt, gelte also

$$\sum_{i=1}^{n} \lambda^i a_i = \sum_{i=1}^{n} \mu^i a_i$$

mit $(\lambda^1, \ldots, \lambda^n) \neq (\mu^1, \ldots, \mu^n)$. Dann ist

$$\sum_{i=1}^{n} \left(\lambda^i - \mu^i\right) a_i = 0$$

eine nichttriviale Linearkombination der Null. Die Vektoren sind also linear abhängig. □

4.2. **Basen.**

Definition 4.2.1. Sei V ein Vektorraum. Dann ist $S \subset V$ eine **Basis** von V, wenn S linear unabhängig ist und $\langle S \rangle = V$ gilt.

Beispiele 4.2.2.

(i) $\{e_1, \ldots, e_n\}$ ist eine Basis des F^n, die **Standardbasis** von F^n.

(ii) $\{1, X, X^2, \ldots\}$ ist eine Basis von $F[X]$.

Als Korollar zu Theorem 4.1.4 erhalten wir

Korollar 4.2.3. *$S \subset V$ ist genau dann eine Basis von V, wenn sich jeder Vektor in eindeutiger Weise als Linearkombination von Vektoren aus S schreiben lässt.*

Theorem 4.2.4. *Sei $S \subset V$ linear unabhängig. Sei $b \in V$. Ist $b \notin \langle S \rangle$, so ist $S \cup \{b\}$ linear unabhängig.*

Beweis. Sei $\{a_1, \ldots, a_n\}$ eine endliche Teilmenge von S. Zeige, dass $\{a_1, \ldots, a_n, b\}$ linear unabhängig ist: Gelte

$$\mu b + \sum_{i=1}^{n} \lambda^i a_i = 0.$$

Dann folgt zunächst $\mu = 0$, denn sonst wäre b als Linearkombination der $\{a_i\}$ darstellbar. Aufgrund der linearen Unabhängigkeit von S folgt dann aber auch $\lambda^i = 0$ für alle i. Somit ist $S \cup \{b\}$ linear unabhängig. □

Mit Hilfe dieses Satzes kann man linear unabhängige Teilmengen sukzessive vergrößern.

Definition 4.2.5. Sei $S \subset V$ linear unabhängig. Dann heißt S **maximal** oder maximal linear unabhängig, wenn $S \cup \{b\}$ für jedes $b \in V \setminus S$ linear abhängig ist.

Theorem 4.2.6. *Sei $S \subset V$. Dann ist S genau dann eine Basis von V, wenn S linear unabhängig und maximal ist.*

Beweis.

(i) Sei S maximal und linear unabhängig. Gäbe es $b \in V \setminus \langle S \rangle$, so wäre $S \cup \{b\}$ nach Theorem 4.2.4 auch linear unabhängig. Dies widerspricht der Maximalität. Somit gibt es kein solches b und wir erhalten $\langle S \rangle = V$. Da S nach Voraussetzung linear unabhängig ist, handelt es sich um eine Basis.

(ii) Ist S eine Basis, so ist S nach Definition linear unabhängig. Da jedes $b \in V$ eine Linearkombination von Vektoren aus S ist, ist $S \cup \{b\}$ für beliebiges $b \in V$ nach Theorem 4.1.3 linear abhängig. Somit ist S bereits maximal. □

Wir zeigen der Übersichtlichkeit halber zunächst, dass jeder endlich erzeugte Vektorraum eine Basis besitzt.

Theorem 4.2.7. *Sei V ein endlich erzeugter Vektorraum. Dann besitzt V eine Basis.*

Beweis. Sei $S = \{a_1, \ldots, a_n\} \subset V$ ein endliches Erzeugendensystem. Ist S linear unabhängig, so sind wir fertig. Sonst gibt es ein $i_0 \in \{1, \ldots, n\}$ und $\lambda^i \in F$, $1 \leq i \leq n$, mit $\lambda^{i_0} \neq 0$ und $0 = \sum_{i=1}^{n} \lambda^i a_i$. Somit ist a_{i_0} eine Linearkombination von $S_1 := \{a_1, \ldots, a_{i_0-1}, a_{i_0+1}, \ldots, a_n\}$. Nach Theorem 2.4.5 folgt $\langle S \rangle = \langle S_1 \rangle$. Induktiv erhalten wir so, wie wir aus S die Menge S_1 gewonnen haben, aus S_1 eine Menge S_2, dann aus S_2 eine Menge S_3, ... Da S endlich ist, bricht dies irgendwann einmal ab und wir erhalten ein k, so dass S_k linear unabhängig ist. Wegen $V = \langle S \rangle = \langle S_1 \rangle = \ldots = \langle S_k \rangle$ ist S_k eine Basis von V. □

Das Zornsche Lemma erlaubt uns nun, zu zeigen, dass jeder Vektorraum eine Basis besitzt. Das Zornsche Lemma ist äquivalent zum Auswahlaxiom. Dies wird in einer Vorlesung über Logik oder Mengenlehre bewiesen.

Lemma 4.2.8 (Zornsches Lemma)**.** $\star$ *Sei A nichtleer und $\leq$ eine Halbordnung auf A. Besitzt jede total geordnete Teilmenge B eine obere Schranke in A, also ein $\beta \in A$ mit $b \leq \beta$ für alle $b \in B$, so besitzt A mindestens ein maximales Element m. (m ist maximales Element von A, falls aus $m \leq a$, $a \in A$ stets $a = m$ folgt.)*

Theorem 4.2.9. *Sei V ein Vektorraum. Dann besitzt V eine Basis.*

Nach Konvention ist $\emptyset$ die Basis von $\{0\}$, da $\sum_{\emptyset} = 0$ gilt.

Beweis. Die Relation $\subset$ ist eine Halbordnung auf $\mathcal{P}(V)$ und daher auch auf der Menge aller linear unabhängigen Teilmengen $\mathcal{U}$ von V. Sei also $\{A_i\}_{i \in I}$ eine total geordnete Teilmenge von $\mathcal{U}$, gelte also stets $A_i \subset A_j$ oder $A_j \subset A_i$. Setze $A := \bigcup_{i \in I} A_i$. Wir behaupten, dass A eine obere Schranke für $\{A_i\}_{i \in I}$ ist. $A_i \subset A$ ist klar.

Zur linearen Unabhängigkeit: Seien $a_1, \ldots, a_n \in A$ und gelte $\sum_{i=1}^{n} \lambda^i a_i = 0$. Dann gibt es Mengen $A_{j(i)}$ mit $a_i \in A_{j(i)}$. Es gilt $A_{j(1)} \subset A_{j(2)}$ oder $A_{j(2)} \subset A_{j(1)}$. Setze

$k(2) := j(2)$, falls $A_{j(1)} \subset A_{j(2)}$ gilt und sonst $k(2) := j(1)$. Dann gilt $A_{j(i)} \subset A_{k(2)}$ für alle $i = 1, 2$. Gelte per Induktion bereits $A_{j(i)} \subset A_{k(p)}$ für $1 \leq i \leq p$ und für ein $k(p) \in \{j(1), \ldots, j(p)\}$. Gilt auch $A_{j(p+1)} \subset A_{k(p)}$, so setze $k(p+1) := k(p)$, sonst $k(p+1) := j(p+1)$. Es folgt $A_{j(i)} \subset A_{k(p)}$ für $1 \leq i \leq p+1$. Somit gibt es einen Index $k(n) \in \{j(1), \ldots, j(n)\}$ mit $a_i \in A_{k(n)}$ für alle $1 \leq i \leq n$. Damit sind die Vektoren $\{a_i\}_{1 \leq i \leq n}$ linear unabhängig, A ist also eine obere Schranke der $(A_i)_{i \in I}$.

Das Zornsche Lemma impliziert nun, dass es ein maximales $B \in \mathcal{U}$ gibt. Nach Theorem 4.2.6 folgt $\langle B \rangle = V$. Damit ist B ein maximales linear unabhängiges Erzeugendensystem von V und somit eine Basis von V. □

Definition 4.2.10. Sei S ein Erzeugendensystem von V. Dann heißt S **minimal,** falls es keine echte Teilmenge von S gibt, die ebenfalls V erzeugt.

Theorem 4.2.11. *Eine Familie S von Vektoren ist genau dann eine Basis von V, wenn S ein minimales Erzeugendensystem von V ist.*

Beweis.

(i) Ist S nicht minimal, so gibt es $b \in S$, so dass $S \setminus \{b\}$ ebenfalls ein Erzeugendensystem von V ist. Insbesondere gilt dann aber auch

$$b = \sum_{i=1}^{n} \lambda^i a_i$$

für spezielle $\lambda^1, \ldots \lambda^n \in F$, $a_1, \ldots, a_n \in S \setminus \{b\}$, $n \in \mathbb{N}$. Bringen wir b auf die andere Seite, so widerspricht dies der linearen Unabhängigkeit von S. Somit ist eine Basis minimal.

(ii) Sei S ein minimales Erzeugendensystem. Ist S nicht linear unabhängig, so gibt es $a_1, \ldots, a_n \in S$ und $\lambda^1, \ldots, \lambda^n \in F$, $n \in \mathbb{N}$, mit

$$\sum_{i=1}^{n} \lambda^i a_i = 0$$

und $(\lambda^1, \ldots, \lambda^n) \neq (0, \ldots, 0)$. Nehme ohne Einschränkung an, dass $\lambda_1 \neq 0$ ist. Dann lässt sich a_1 als Linearkombination von $a_2, \ldots, a_n$ schreiben. Nach Theorem 2.4.5 gilt $\langle S \rangle = \langle S \setminus \{a_1\} \rangle$. Somit ist S nicht minimal. Widerspruch. Die Behauptung folgt. □

Korollar 4.2.12. *Sei V Vektorraum, $S \subset V$. Nach den Theoremen 4.2.6 und 4.2.11 sind die folgenden Aussagen äquivalent:*

(i) S ist eine Basis von V,
(ii) S ist eine maximale linear unabhängige Teilmenge von V,
(iii) S ist ein minimales Erzeugendensystem von V.

Im endlich erzeugten Fall können wir auch absteigende Folgen von Erzeugendensystemen betrachten um eine Basis zu bekommen.

Theorem 4.2.13. *Sei S ein endliches Erzeugendensystem von V. Dann gibt es eine Basis $S' \subset S$ von V.*

Beweis. Ist S keine Basis, so ist S nach Korollar 4.2.12 nicht minimal. Somit gibt es $S_1 \subsetneq S$, das ebenfalls ein Erzeugendensystem ist. Ist S_1 keine Basis, so wiederholen wir diesen Schritt und erhalten $S_2 \subsetneq S_1$, ... Da S endlich ist, bricht die strikt absteigende Folge

$$S \supsetneq S_1 \supsetneq S_2 \supsetneq \ldots \supsetneq S_k$$

bei einem $k \in \mathbb{N}$ ab. S_k ist minimal und somit nach Korollar 4.2.12 eine Basis. □

In den Übungen werden wir sehen, dass dieses Vorgehen ohne die Endlichkeit des Erzeugendensystems nicht funktioniert.

Es gilt aber auch allgemein

Theorem 4.2.14. *Sei S ein Erzeugendensystem von V. Dann gibt es $S' \subset S$, so dass S' eine Basis von V ist.*

Beweis. Argumentiere wie im Beweis von Theorem 4.2.9, benutze aber linear unabhängige Teilmengen von S statt von V. □

Theorem 4.2.15. *Sei $S \subset V$ linear unabhängig. Dann besitzt V eine Basis B mit $S \subset B$.*

Beweis. Gehe wie beim Beweis von Theorem 4.2.9 vor, benutze aber linear unabhängige Teilmengen, die S enthalten statt linearen Teilmengen von V. □

4.3. **Dimension.**

Definition 4.3.1. Sei V ein Vektorraum mit Basis S. Ist S endlich und besitzt n Elemente, $n \in \mathbb{N}$, so hat V die **Dimension** n, $\dim V = n$.

Zunächst einmal ist nicht klar, ob die Dimension wohldefiniert ist, da es verschiedene Basen mit unterschiedlich vielen Elementen geben könnte. Wir werden aber in Theorem 4.3.3 noch zeigen, dass dies nicht der Fall ist.

Lemma 4.3.2 (Austauschsatz von Steinitz)**.** *Sei V ein Vektorraum. Ist die Familie $\{a_1, \dots, a_p\}$ linear unabhängig und sei $\{b_1, \dots, b_n\}$ ein Erzeugendensystem von V. Dann gilt $p \leq n$ und nach einer Umnummerierung der b_i's ist auch die Familie $\{a_1, \dots, a_p, b_{p+1}, \dots, b_n\}$ ein Erzeugendensystem von V.*

Beweis. Wir beweisen per Induktion nach k, dass $\{a_1, \dots, a_k, b_{k+1}, \dots, b_n\}$ für $0 \leq k \leq n$ nach Umnummerierung der b_i's ebenfalls ein Erzeugendensystem von V ist. (Wäre $p > n$, so folgte aus diesem Beweis, dass insbesondere $\{a_1, \dots, a_n\}$ ein Erzeugendensystem von V ist. Dann wäre aber a_{n+1} aus $a_1, \dots, a_n$ linear kombinierbar. Widerspruch.) Für $k = 0$ ist die Aussage klar. Sei also $\{a_1, \dots, a_k, b_{k+1}, \dots, b_n\}$ ein Erzeugendensystem von V. Es gibt also eine Linearkombination davon, die a_{k+1} darstellt

$$a_{k+1} = \sum_{i=1}^{k} \lambda^i a_i + \sum_{j=k+1}^{n} \mu^j b_j.$$

Da $a_{k+1} \neq 0$ ist und da die a_i's linear unabhängig sind, ist ein $\mu^j \neq 0$, ohne Einschränkung $\mu^{k+1} \neq 0$. Daher ist $b_{k+1} \in \langle a_1, \dots, a_{k+1}, b_{k+2}, \dots, b_n \rangle$. Setze $S := \{a_1, \dots, a_k, b_{k+2}, \dots, b_n\}$. Nach Theorem 2.4.5 erhalten wir

$$\langle S \cup \{b_{k+1}\} \rangle = \langle S \cup \{a_{k+1}, b_{k+1}\} \rangle = \langle S \cup \{a_{k+1}\} \rangle.$$

Die Behauptung folgt nun per Induktion. □

Aufgrund des folgenden Theorems ist die Dimension wohldefiniert.

Theorem 4.3.3. *Sei V ein Vektorraum mit Basen S und T. Habe S endlich viele Elemente n, $n \in \mathbb{N}$. Dann besitzt T ebenfalls n Elemente.*

Beweis. Seien $|T|$ und $|S|$ die Anzahlen der Elemente von T bzw. S. Dann gilt nach Lemma 4.3.2 $|T| \leq |S|$ und $|S| \leq |T|$. Die Behauptung folgt. □

Korollar 4.3.4. *Sei V ein n-dimensionaler Vektorraum. Dann gelten folgende Aussagen.*

(i) Jedes Erzeugendensystem von V besitzt wenigstens n Vektoren.
(ii) Jede Familie von strikt mehr als n Vektoren ist linear abhängig.
(iii) Jedes Erzeugendensystem aus n Vektoren ist eine Basis von V.

(iv) Jede Familie von n linear unabhängigen Vektoren ist eine Basis von V.

Beweis.
(i) Nach Theorem 4.2.14 gäbe es zu einem Erzeugendensystem mit weniger als n Elementen auch eine Basis mit weniger als n Elementen. Widerspruch zu Theorem 4.3.3.
(ii) Nach Theorem 4.2.15 könnte man sonst diese Menge zu einer Basis mit mehr als n Elementen ergänzen. Dies widerspricht wiederum Theorem 4.3.3.
(iii) Ein minimales Erzeugendensystem ist eine Basis.
(iv) Eine maximale linear unabhängige Familie ist eine Basis. □

Korollar 4.3.5. *Sei W ein Unterraum eines Vektorraumes V mit $\dim V = n$. Dann gilt $\dim W \leq \dim V$ und Gleichheit gilt genau dann, wenn $V = W$ ist.*

Beweis. Sei S eine Basis von W und T mit $S \subset T$ eine Basis von V. Dann gilt $|S| \leq |T|$. Bei Gleichheit gilt $S = T$ und die Erzeugnisse stimmen überein. Die Rückrichtung ist trivial. □

Theorem 4.3.6 (Dimensionsformel)**.** *Seien $U, V \subset W$ endlichdimensionale Unterräume eines Vektorraumes W. Dann gilt*

$$\dim U + V = \dim U + \dim V - \dim(U \cap V).$$

Beweis. Sei $w_1, \ldots, w_{\dim U \cap V}$ eine Basis von $U \cap V$. Seien $u_1, \ldots, u_{\dim U - \dim U \cap V}$ so gewählt, dass sie zusammen mit den w_i's eine Basis von U bilden und seien $v_1, \ldots, v_{\dim V - \dim U \cap V}$ so gewählt, dass sie zusammen mit den w_i's eine Basis von V bilden. Wir behaupten nun, dass alle diese Vektoren zusammen eine Basis von $U + V$ bilden. Es ist klar, dass sie $U + V$ erzeugen. Zum Nachweis der linearen Unabhängigkeit sei

$$\underbrace{\sum_i c^i w_i}_{=:w \in U \cap V} + \underbrace{\sum_j a^j u_j}_{=:u \in U} + \underbrace{\sum_k b^k v_k}_{=:v \in V} = 0.$$

Dann ist $v \in V$ und $v = -w - u \in U$, also $v \in U \cap V$. Die w_i's bilden zusammen mit den v_i's eine Basis von V, Vektoren in $U \cap V$ werden aber bereits eindeutig als Linearkombination von w_i's dargestellt. Die obige Darstellung von v enhält keine w_i's. Somit ist $v = 0$. Ebenso ist $u = 0$. Die lineare Unabhängigkeit folgt. □

4.3.1. *Koordinaten.*

Definition 4.3.7 (Koordinaten)**.** Sei V ein F-Vektorraum mit einer (geordneten) Basis $\{a_1, \ldots, a_n\}$. Wir schreiben auch $(a_1, \ldots, a_n)$. Dann lässt sich nach Korollar 4.2.3 jedes $a \in V$ in eindeutiger Weise aus den a_i's linear kombinieren

$$a = \sum_{i=1}^{n} \lambda^i a_i,$$

$\lambda^i \in F$. Die Koeffizienten $\lambda^1, \ldots, \lambda^n$ heißen **Koordinaten** von a bezüglich der Basis $\{a_1, \ldots, a_n\}$.

4.4. **Lineare Abbildungen.**

Definition 4.4.1 (Lineare Abbildung)**.** Seien V, W zwei Vektorräume über einem Körper F. Eine Abbildung $f\colon V \to W$ heißt **linear,** falls

$$f(a + b) = f(a) + f(b) \qquad \text{für alle } a, b \in V \text{ und} \tag{4.1}$$

$$f(\lambda a) = \lambda f(a) \qquad \text{für alle } a \in V, \lambda \in F \text{ gelten.} \tag{4.2}$$

Lemma 4.4.2. *Sei $f\colon V \to W$ linear. Dann gelten die folgenden Aussagen*
(i) $f(0) = 0$,

(ii) $f\left(\sum_{i=1}^{n} \lambda^i a_i\right) = \sum_{i=1}^{n} \lambda^i f(a_i)$ *für alle* $\lambda^i \in F$, $a_i \in V$, *insbesondere ist*

$$f(\lambda a + \mu b) = \lambda f(a) + \mu f(b)$$

für alle $\lambda, \mu \in F$ *und* $a, b \in V$. *Diese letzte Gleichheit ist äquivalent zu* (4.1) *und* (4.2), *auch wenn wir* $\mu = 1$ *wählen.*

Beweis.

(i) Benutze (4.2) mit $\lambda = 0$.

(ii) Per Induktion erhalten wir

$$\begin{aligned} f\left(\sum_{i=1}^{n} \lambda^i a_i\right) &= f\left(\lambda^1 a_1 + \sum_{i=2}^{n} \lambda^i a_i\right) \\ &\overset{(4.1)}{=} f(\lambda^1 a_1) + f\left(\sum_{i=2}^{n} \lambda^i a_i\right) \\ &\overset{(4.2)}{=} \lambda^1 f(a_1) + \sum_{i=2}^{n} \lambda^i f(a_i) \qquad \text{nach Induktionsannahme} \\ &= \sum_{i=1}^{n} \lambda^i f(a_i). \end{aligned}$$

Die Gleichung (4.1) erhält man mit $\lambda = \mu = 1$, die Gleichung (4.2) mit $b = 0$. □

Beispiele 4.4.3.

(i) Die Identität $\mathbb{1} = \mathbb{1}_V \colon V \to V$, $\mathbb{1}(a) := a$ ist linear.

(ii) Seien $a^i_j \in F$, $1 \le i \le m$, $1 \le j \le n$. Dann ist $f \colon F^n \to F^m$ mit

$$f(\xi) \equiv f(\xi^1, \dots, \xi^n) \equiv (f^1(\xi), \dots, f^m(\xi)), \quad f^i(\xi) := \sum_{j=1}^{n} a^i_j \xi^j$$

linear.

(iii) Seien $f \colon V \to W$ und $g \colon W \to X$ linear. Dann ist $g \circ f \colon V \to X$ ebenfalls linear.

(iv) Sei $P \in F[X]$. Dann ist die Abbildung D (Differenzieren)

$$D : P[X] = \sum_{i=0}^{n} a_i X^i \mapsto \sum_{i=1}^{n} i a_i X^{i-1}$$

eine lineare Abbildung.

(v) Die Abbildung

$$C^0([0,1]) \ni f \mapsto \left(x \mapsto \int_0^x f(t)\, dt\right) \in C^0([0,1])$$

ist linear.

Definition 4.4.4.

(i) Eine lineare Abbildung $f \colon V \to W$ heißt **Isomorphismus,** wenn es eine Abbildung $g \colon W \to V$ mit $g \circ f = \mathbb{1}_V$ und $f \circ g = \mathbb{1}_W$ gibt. g heißt die zu f inverse Abbildung. Wir schreiben $g = f^{-1}$.

(ii) Zwei Vektorräume V und W heißen **isomorph,** wenn es einen Isomorphismus $f \colon V \to W$ gibt.

Bemerkung 4.4.5.

(i) Wir werden in Lemma 4.4.6 sehen, dass g wohldefiniert ist, d. h. insbesondere linear.
(ii) Isomorph zu sein ist eine Äquivalenzrelation auf einer Kollektion von F-Vektorräumen.

Lemma 4.4.6. *Sei $f\colon V \to W$ ein Isomorphismus. Dann ist die Inverse $g\colon W \to V$ von f eindeutig bestimmt und ebenfalls linear.*

Beweis. Sei $h\colon W \to V$ eine weitere Abbildung mit $h \circ f = \mathbb{1}_V$ und $f \circ h = \mathbb{1}_W$. Dann gilt für alle $b \in W$

$$g(b) = \mathbb{1}_V \circ g(b) = (h \circ f) \circ g(b) = h \circ (f \circ g)(b) = h \circ \mathbb{1}_W(b) = h(b).$$

Somit gilt $h = g$.
g ist linear, denn es gilt

$$\begin{aligned} g(a) + g(b) &= \mathbb{1}_V(g(a) + g(b)) = g \circ f(g(a) + g(b)) \\ &= g(f \circ g(a) + f \circ g(b)) && \text{da } f \text{ linear ist} \\ &= g(a + b); \\ \lambda g(a) &= g \circ f(\lambda g(a)) = g(\lambda f \circ g(a)) && \text{da } f \text{ linear ist} \\ &= g(\lambda a). \end{aligned}$$

□

Lemma 4.4.7 (Koordinatenabbildung)**.** *Sei V ein F-Vektorraum mit geordneter Basis $(a_1, \dots, a_n)$. Dann ist die Abbildung $f\colon V \to F^n$,*

$$V \ni a = \sum_{i=1}^{n} \lambda^i a_i \mapsto (\lambda^1, \dots \lambda^n) \in F^n,$$

ein Isomorphismus.

Beweis. Da die Darstellung von a eindeutig ist, ist f wohldefiniert. Die Linearität von f ist klar. Die Umkehrabbildung $g\colon F^n \to V$,

$$g(\lambda^1, \dots, \lambda^n) := \sum_{i=1}^{n} \lambda^i a_i,$$

ist ebenfalls offensichtlicherweise linear und zu f invers. □

Korollar 4.4.8. *Sei V ein n-dimensionaler F-Vektorraum. Dann ist V isomorph zu F^n.*

Theorem 4.4.9. *Sei $f\colon V \to W$ linear. Sei $\{a_1, \dots, a_n\} \subset V$ linear abhängig. Dann ist $\{f(a_1), \dots, f(a_n)\} \subset W$ ebenfalls linear abhängig.*

Beweis. Aus $\sum_{i=1}^{n} \lambda^i a_i = 0$ folgt mit Hilfe der Linearität $\sum_{i=1}^{n} \lambda^i f(a_i) = 0$. □

Korollar 4.4.10. *Sei $f\colon V \to W$ ein Isomorphismus. Dann ist $\{a_1, \dots, a_n\}$ genau dann linear abhängig, wenn $\{f(a_1), \dots, f(a_n)\}$ linear abhängig ist.*

Theorem 4.4.11. *Seien V, W über F endlich erzeugbare Vektorräume. Dann sind V und W genau dann isomorph, falls $\dim V = \dim W$ gilt.*

Beweis.
(i) Gelte zunächst $\dim V = \dim W$. Dann gibt es nach Lemma 4.4.7 Isomorphismen $\varphi\colon V \to F^n$ und $\psi\colon W \to F^n$. Somit ist $\psi^{-1} \circ \varphi\colon V \to W$ der gesuchte Isomorphismus.

(ii) Sei $\{a_1, \dots, a_n\}$ eine Basis von V. Sei $f\colon V \to W$ ein Isomorphismus. Dann ist $\{f(a_i)\}$ linear unabhängig. Es ist auch eine Basis, denn sei $b \in W$ beliebig, dann gilt

$$V \ni f^{-1}(b) = \sum_{i=1}^{n} \lambda^i a_i$$

für geeignete $\lambda_i \in F$. Somit ist $b = f\left(f^{-1}(b)\right) = \sum\limits_{i=1}^{n} \lambda^i f(a_i)$, $\{f(a_i)\}$ ist also auch ein Erzeugendensystem von W und damit eine Basis. Da die Anzahl von Basiselementen in beiden Vektorräumen übereinstimmt, gilt $\dim V = \dim W$.

□

4.5. **Rechenmethoden.** Sei V ein endlichdimensionaler Vektorraum, $S \subset V$ endlich. Wir wollen folgende Fragen konkret behandeln können:

(i) Ist S linear unabhängig?
(ii) Wie findet man eine Basis von $\langle S\rangle$, falls S linear abhängig ist?
(iii) Wie findet man eine Basis von $\langle S\rangle$, die aus Elementen von S besteht?
(iv) Gilt $a \in \langle S\rangle$, d. h. ist a aus Vektoren in S linear kombinierbar?

Wegen Lemma 4.4.7 und da diese Eigenschaften sich unter Isomorphismen sich nicht ändern, können wir ohne Einschränkung $V = F^n$ betrachten.

Notation 4.5.1. Sei $S = (a_1, \dots, a_m)$, $a_i \in F^n$.

(i) $\langle S\rangle := \langle\{a_1, \dots, a_m\}\rangle$ (Ebenso werden wir die Anordnung der Vektoren vergessen, wenn es um Fragen wie z. B. lineare Unabhängigkeit geht.)
(ii) $S_{ij} := (a_1, \dots, a_{i-1}, a_j, a_{i+1}, \dots, a_{j-1}, a_i, a_{j+1}, \dots, a_m)$; S_{ij} entsteht aus S durch Vertauschen des i-ten mit dem j-ten Vektor.
(iii) $S_i(\lambda) := (a_1, \dots, a_{i-1}, \lambda a_i, a_{i+1}, \dots, a_m)$, $0 \neq \lambda \in F$; $S_i(\lambda)$ entsteht aus S durch Multiplikation des i-ten Vektors mit λ.
(iv) $S_{ij}(\lambda) := (a_1, \dots, a_{i-1}, a_i + \lambda a_j, a_{i+1}, \dots, a_m)$, $\lambda \in F$, $i \neq j$; $S_{ij}(\lambda)$ entsteht aus S durch Addition des λ-fachen des Vektors a_j zu a_i.

Theorem 4.5.2.

(i) Es gilt $\langle S\rangle = \langle S_{ij}\rangle = \langle S_i(\lambda)\rangle = \langle S_{ij}(\lambda)\rangle$ für alle λ, i, j wie oben.
(ii) Ist eine der Familien S, S_{ij}, $S_i(\lambda)$, $S_{ij}(\lambda)$ linear unabhängig, so gilt dies auch für alle anderen dieser Familien.

Beweis.

(i) (a) $\langle S\rangle = \langle S_{ij}\rangle$ ist klar.
(b) Sei $a = \sum\limits_{k=1}^{m} \mu^k a_k \in \langle S\rangle$. Dann ist $\sum\limits_{k\neq i} \mu^k a_k + \left(\frac{1}{\lambda}\mu^i\right)(\lambda a_i) = a \in \langle S_i(\lambda)\rangle$. $\langle S_i(\lambda) \subset \langle S\rangle$ folgt analog.
(c) Benutze nun, dass sich die Linearkombination einer Familie von Vektoren nicht ändert, wenn man einen linear kombinierbaren Vektor hinzufügt, vergleiche Theorem 2.4.5.

$$\begin{aligned}\langle S\rangle &= \langle a_1, \dots, a_i, \dots, a_m\rangle \\ &= \langle a_1, \dots, a_i, \dots, a_m, a_i + \lambda a_j\rangle \\ &= \langle a_1, \dots, a_i + \lambda a_j, \dots, a_m\rangle, \quad \text{da } a_i = (a_i + \lambda a_j) - \lambda a_j \text{ gilt,} \\ &= \langle S_{ij}(\lambda)\rangle.\end{aligned}$$

(ii) Sei S linear abhängig. Man überlegt sich leicht, dass dann auch die anderen drei Familien von Vektoren linear abhängig sind. Insbesondere folgt aus $\mu^i a_i + \mu^j a_j = 0$ auch $\mu^i(a_i + \lambda a_j) + \left(\mu^j - \mu^i\lambda\right) a_j = 0$.

Sei umgekehrt eine der Familien S_{ij}, $S_i(\lambda)$, $S_{ij}(\lambda)$ linear abhängig. Die Operationen, die von S zu dieser Familie geführt haben, sind mit Hilfe derselben Operationen invertierbar: Ist $T = S_{ij}$, so ist $T_{ij} = S$; ist $T = S_i(\lambda)$, so ist $T_i\left(\frac{1}{\lambda}\right) = S$; ist $T = S_{ij}(\lambda)$, so ist $T_{ij}(-\lambda) = S$. Somit ist in diesem Fall auch S linear abhängig. Die Behauptung folgt. □

Definition 4.5.3. Seien $a^i = (a^i_1, \ldots, a^i_n)$, $1 \leq i \leq m$, Zeilenvektoren und

$$A = \begin{pmatrix} a^1_1 & a^1_2 & \ldots & a^1_n \\ a^2_1 & a^2_2 & \ldots & a^2_n \\ \vdots & \vdots & & \vdots \\ a^m_1 & a^m_2 & \ldots & a^m_n \end{pmatrix}$$

die aus diesen Zeilenvektoren bestehende Matrix. Dann heißt der von den Zeilen $(a^i_1, \ldots, a^i_n)$ in F^n erzeugte Unterraum **Zeilenraum** der Matrix A. (Analog heißt der von den Spalten in F^m erzeugte Unterraum Spaltenraum der Matrix A.)

In Analogie zu den Operatoren auf Erzeugendensystemen definieren wir die folgenden **elementaren Zeilenoperationen:**

(i) Vertausche die Zeilen i und j.
(ii) Ersetze die i-te Zeile durch das λ-fache der i-ten Zeile für $0 \neq \lambda \in F$.
(iii) Ersetze die i-te Zeile a^i durch $a^i + \lambda a^j$ für $j \neq i$ und $\lambda \in F$.

Wie bei linearen Gleichungssystemen sagen wir, dass eine Matrix in Zeilenstufenform ist, wenn die Anzahl der links stehenden Nullen von Zeile zu Zeile strikt wächst oder eine Zeile und alle folgenden Zeilen nur Nullen enthalten.

Theorem 4.5.4. *Eine $(m \times n)$-Matrix A lässt sich auf Zeilenstufenform bringen.*

Beweis. Verfahre so wie beim Beweis, dass sich ein lineares Gleichungssystem auf Zeilenstufenform bringen lässt. □

Lemma 4.5.5. *Sei A eine $(m \times n)$-Matrix. Bringt man A auf Zeilenstufenform, so bilden die Nicht-Null-Zeilen eine Basis des Zeilenraumes von A.*

Beweis. Nach Theorem 4.5.2 ändern elementare Zeilenumformungen den Zeilenraum einer Matrix nicht. Da in Zeilenstufenform die Nicht-Null-Zeilen eine Basis des Zeilenraumes bilden, folgt die Behauptung. □

Beispiel 4.5.6. Mit Hilfe elementarer Zeilenumformungen wollen wir den Zeilenraum bestimmen. Wir erhalten

$$\begin{pmatrix} 2 & 4 & -1 & 7 \\ -1 & -2 & 1 & -4 \\ 1 & 2 & 2 & 1 \\ 3 & 6 & 3 & 6 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 2 & 1 \\ -1 & -2 & 1 & -4 \\ 2 & 4 & -1 & 7 \\ 3 & 6 & 3 & 6 \end{pmatrix},$$

$$\begin{pmatrix} 1 & 2 & 2 & 1 \\ 0 & 0 & 3 & -3 \\ 0 & 0 & -5 & 5 \\ 0 & 0 & -3 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 2 & 1 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

4.6. **Anwendungen auf lineare Gleichungssysteme.** Gleichungen der Form

$$a_1 x^1 + a_2 x^2 + \ldots + a_n x^n = b$$

bilden unter den Verknüpfungen

$$\begin{aligned} &(a_1 x^1 + \ldots + a_n x^n = b) + \\ &(a'_1 x^1 + \ldots + a'_n x^n = b') := ((a_1 + a'_1)x^1 + \ldots + (a_n + a'_n)x^n = (b + b')), \\ &\lambda(a_1 x^1 + \ldots + a_n x^n = b) := ((\lambda a_1)x^1 + \ldots + (\lambda a_n)x^n = (\lambda b)) \end{aligned}$$

einen Vektorraum V. $f\colon V \to F^{n+1}$ mit

$$(a_1x^1 + \ldots + a_nx^n = b) \mapsto (a_1, \ldots, a_n, b)$$

ist ein Isomorphismus.

Sei nun ein lineares Gleichungssystem gegeben. Eine Lösung des Gleichungssystems löst auch alle Gleichungen im von den linearen Gleichungen erzeugten Unterraum U (und eine Lösung aller Gleichungen des Unterraumes ist auch eine Lösung des ursprünglichen linearen Gleichungssystems). Wenden wir dies zweimal an, so sehen wir: Sei S eine Basis von U. Dann ist x genau dann eine Lösung des ursprünglichen Gleichungssystems, wenn x alle Gleichungen in S erfüllt.

Technisches Vorgehen: Um

$$\sum_{k=1}^{n} a_k^i b^k = b^i$$

für alle $1 \leq i \leq m$ zu lösen, können wir alle Lösungen der Matrix bestimmen, die wir erhalten, wenn wir

$$\begin{pmatrix} a_1^1 & a_2^1 & \ldots & a_n^1 & b^1 \\ a_1^2 & a_2^2 & \ldots & a_n^2 & b^2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_1^m & a_2^m & \ldots & a_n^m & b^m \end{pmatrix}$$

auf Zeilenstufenform gebracht haben. Dies ist einfach, wenn wir die Gleichungen von unten nach oben betrachten und das bisherige Ergebnis jeweils in die weiter oben stehenden Gleichungen einsetzen.

4.7. **Der Rang.**

Definition 4.7.1. Sei A eine $(m \times n)$-Matrix

$$A = \left(a_j^i\right)_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} = \begin{pmatrix} a_1^1 & a_2^1 & \ldots & a_n^1 \\ a_1^2 & a_2^2 & \ldots & a_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ a_1^m & a_2^m & \ldots & a_n^m \end{pmatrix}$$

über F.

(i) Dann ist der **Zeilenrang** die Dimension des von den Zeilen $b^1, \ldots, b^m$, $b^i = (a_1^i, \ldots, a_n^i)$, in F^n erzeugten Vektorraumes.

(ii) Der **Spaltenrang** ist die Dimension des von den Spalten $d_1, \ldots, d_n$,

$$d_j = \left(a_j^i\right)_{1 \leq i \leq n} = \begin{pmatrix} a_j^1 \\ \vdots \\ a_j^n \end{pmatrix}$$

erzeugten Vektorraumes in F^m. Der von den Spalten erzeugte Vektorraum heißt Spaltenraum von A.

(iii) Der **Rang** einer Matrix A ist der Zeilenrang von A, $\operatorname{rang} A$.

Die Begriffsbildung bei der letzten Definition ist sinnvoll, denn es gilt

Theorem 4.7.2. *Sei A eine Matrix. Dann ist der Zeilenrang von A gleich dem Spaltenrang von A.*

Beweis. Wir wollen die Matrix auf Zeilenstufenform bringen. Für Matrizen in Zeilenstufenform gilt die Behauptung offensichtlicherweise. Zeige also noch, dass sich weder Zeilenrang noch Spaltenrang unter elementaren Zeilenumformungen ändern.

Da sich sogar der Zeilenraum einer Matrix unter elementaren Zeilenumformungen nicht ändert, bleibt auch der Zeilenrang konstant.

Unter elementaren Zeilenumformungen kann sich der Spaltenraum ändern. Sei

$$\sum_{i=1}^{n} \lambda^i d_i = 0$$

eine lineare Abhängigkeit der Spalten. Dann bleibt diese unter elementaren Zeilenumformungen erhalten. Der Spaltenrang wird unter elementaren Zeilenumformungen also höchstens kleiner. Da sich elementare Zeilenumformungen aber invertieren lassen, bleibt er gleich.

Die Behauptung folgt. □

Definition 4.7.3. Sei A eine $(n \times n)$-Matrix. Dann heißt A

(i) **regulär,** falls $\operatorname{rang} A = n$ ist,
(ii) sonst heißt sie **singulär.**

Theorem 4.7.4. *Sei A eine $(n \times n)$-Matrix. Dann sind die folgenden Aussagen äquivalent:*

(i) A *ist regulär,*
(ii) $\operatorname{rang} A = n$,
(iii) die Zeilen sind linear unabhängig,
(iv) die Spalten sind linear unabhängig,
(v) das lineare Gleichungssystem $Ax = 0$ *hat nur die triviale Lösung.*

Beweis. Nach Definition und Theorem 4.7.2 ist nur noch die Äquivalenz zur ausschließlich trivialen Lösbarkeit zu zeigen.

$Ax = 0$ besitzt genau dann eine nichttriviale Lösung, wenn die Spalten von A linear abhängig sind. □

Theorem 4.7.5. *Sei $Ax = 0$ ein lineares homogenes Gleichungssystem mit n Unbekannten. Dann hat der Raum der Lösungen die Dimension $n - \operatorname{rang} A$.*

Beweis. Bringe das Gleichungssystem auf Zeilenstufenform. Dann kann man die Aussage direkt am Gleichungssystem ablesen. □

Theorem 4.7.6. *Sei $Ax = b$ ein inhomogenes lineares Gleichungssystem. Sei $(A|b)$ die augmentierte Matrix. Dann ist das Gleichungssystem genau dann lösbar, wenn* $\operatorname{rang} A = \operatorname{rang}(A|b)$ *gilt.*

Beweis. Ist $Ax = b$ lösbar, so ist b eine Linearkombination der Spalten von A. Somit ändert sich der Spaltenrang durch Hinzunahme der Spalte b nicht.

Der Spaltenraum von A ist im Spaltenraum von $(A|b)$ enthalten. Da die Dimensionen nach Voraussetzung übereinstimmen, ist b im Spaltenraum von A enthalten. Die Linearkombination von b aus den Spalten von A liefert gerade die Lösung x. □

4.8. **Basiswechsel.**

Herleitung 4.8.1. Sei V ein n-dimensionaler Vektorraum über F mit geordneten Basen $S = (a_1, \ldots, a_n)$ und $T = (b_1, \ldots, b_n)$. Sei $\xi \in V$ und gelte

$$\xi = \sum_{i=1}^{n} \lambda^i a_i = \sum_{k=1}^{n} \mu^k b_k.$$

Aufgrund der Basiseigenschaft gibt es Zahlen $\left(d_j^i\right)_{1 \le i,j \le n}$ mit

$$a_i = \sum_{k=1}^{n} d_i^k b_k.$$

Wir wollen nun herleiten, wie sich die Koeffizienten (λ^i) aus den Koeffizienten (μ^k) bestimmen lassen. Es gilt

$$\xi = \sum_{i=1}^{n} \lambda^i a_i = \sum_{i,k=1}^{n} \lambda^i d_i^k b_k = \sum_{k=1}^{n} \left(\sum_{i=1}^{n} \lambda^i d_i^k \right) b_k = \sum_{k=1}^{n} \mu^k b_k.$$

Aufgrund der eindeutigen Darstellbarkeit eines Vektors mit Hilfe einer Basis folgt

$$\mu^k = \sum_{i=1}^{n} \lambda^i d_i^k.$$

Herleitung 4.8.2. Seien $(\alpha_i), (\beta_i)$ und (γ_i) geordnete Basen eines n-dimensionalen F-Vektorraumes V. Gelte

$$\beta_j = \sum_{k=1}^{n} a_j^k \alpha_k,$$

$$\gamma_i = \sum_{j=1}^{n} b_i^j \beta_j$$

und

$$\gamma_i = \sum_{k=1}^{n} c_i^k \alpha_k.$$

Dann folgt

$$\gamma_i = \sum_{j=1}^{n} b_i^j \beta_j = \sum_{j,k=1}^{n} b_i^j a_j^k \alpha_k.$$

Aufgrund der eindeutigen Darstellbarkeit in einer Basis erhalten wir

$$c_i^k = \sum_{j=1}^{n} b_i^j a_j^k.$$

In Matrizenform mit $A = \left(a_j^k\right)_{1 \le k,j \le n}$, $B = \left(b_i^j\right)$ und $C = \left(c_i^k\right)$ erhalten wir

$$C = AB.$$

Sei nun insbesondere $\alpha_i = \gamma_i$ für alle $1 \le i \le n$. Dann gilt $c_j^i = \delta_j^i$ und

$$\delta_i^k = \sum_{j=1}^{n} b_i^j a_j^k \quad \text{oder} \quad \mathbb{1} = AB.$$

Definition 4.8.3. Sei A eine $(n \times n)$-Matrix über F. Eine $(n \times n)$-Matrix B mit $AB = BA = \mathbb{1}$ heißt zu A **inverse Matrix,** $B = A^{-1}$.

Bemerkung 4.8.4. Besitzt A eine Inverse, so ist diese eindeutig bestimmt.

Beweis. Seien B und C Inverse. Aus $AC = I$ folgt $B = BI = B(AC) = (BA)C = IC = C$. Die Behauptung folgt. □

Theorem 4.8.5. *Sei A eine $(n \times n)$-Matrix über F. Dann sind die folgenden Aussagen äquivalent:*

(i) A ist regulär
(ii) A besitzt eine eindeutig bestimmte Inverse, d. h. eine $(n \times n)$-Matrix mit $AB = BA = \mathbb{1}$.
(iii) A ist die Matrix eines Basiswechsels in einem n-dimensionalen F-Vektorraum.

Beweis. Sei A die Matrix eines Basiswechsels. Da auch ein Basiswechsel in der anderen Richtung möglich ist, gibt es nach Herleitung 4.8.2 eine Matrix B mit $I = BA$. Vertauschen wir die Rollen der beiden Basen finden wir eine Matrix C mit $I = AC$. Es gilt $B = BI = B(AC) = (BA)C = IC = C$. B ist nach Bemerkung 4.8.4 eindeutig bestimmt.

Eine invertierbare Matrix ist regulär: Nach Theorem 4.7.4 ist A genau dann invertierbar, wenn das lineare Gleichungssystem $\sum_{i=1}^{n} a_i^j x^i = 0$, $1 \leq j \leq n$, nur die triviale Lösung besitzt. Sei B die Inverse zu A. Somit gilt $\sum_{j=1}^{n} b_j^k a_i^j = \delta_i^k$ für alle $1 \leq i, k \leq n$. Wir erhalten

$$0 = \sum_{i,j=1}^{n} b_j^k a_i^j x^i = \sum_{i=1}^{n} \delta_i^k x^i = x^k$$

für $1 \leq k \leq n$. Somit ist jede Lösung trivial.

Sei schließlich A regulär. $(e_1, \ldots, e_n)$ ist eine geordnete Basis von F^n. Setze

$$a_k := \sum_{i=1}^{k} a_k^i e_i.$$

Die a_k's sind also gerade die Spalten der Matrix A und sind damit linear unabhängig und somit eine Basis von F^n. Also ist A eine Basiswechselmatrix. □

4.9. **Gruppen.** Reguläre Matrizen sind Beispiele für die algebraische Struktur einer Gruppe:

Definition 4.9.1 (Gruppe)**.** ..... Eine **Gruppe** ist eine Menge G zusammen mit

(i) einer Verknüpfung „$\cdot$“: $G \times G \to G$,
(ii) einer Abbildung $G \to G$, die jedem Element $a \in G$ ein $a^{-1} \in G$ zuordnet und
(iii) einem ausgezeichneten Element $e \in G$,

so dass die folgenden Axiome erfüllt sind:

(G1) $a \cdot (b \cdot c) = (a \cdot b) \cdot c$ für alle $a, b, c \in G$ (Assoziativgesetz)
(G2) $a \cdot e = a = e \cdot a$ für alle $a \in G$ (Neutrales Element, Einselement)
(G3) $a \cdot a^{-1} = a^{-1} \cdot a = e$ für alle $a \in G$, d. h. a^{-1} ist die Inverse von a.
(G4) Gilt $a \cdot b = b \cdot a$ für alle $a, b \in G$, so heißt die Gruppe G **kommutativ** oder **abelsch.**

Beispiele 4.9.2.

(i) Die regulären $(n \times n)$-Matrizen über einem Körper F mit der Matrizenmultiplikation bilden eine Gruppe, $GL_n(F)$ oder $GL(n, F)$.
(ii) Die komplexen Zahlen mit Norm 1, also $z \cdot \bar{z} = 1$, bilden mit der Multiplikation eine abelsche Gruppe.
(iii) Sei F ein Körper. Dann bilden die von Null verschiedenen Elemente mit der Multiplikation eine abelsche Gruppe, $F^\times = F \setminus \{0\}$.
(iv) Sei F ein Körper. Dann ist $(F, +)$ eine additive Gruppe. (Die Bezeichnung additiv betont, dass die übliche additive Verknüpfung hier die Gruppenoperation ist.) Diese Gruppe ist abelsch.
(v) $(\mathbb{Z}, +)$ ist eine additive abelsche Gruppe.

Definition 4.9.3 (Untergruppe)**.** Sei G eine Gruppe. Dann ist $U \subset G$ eine **Untergruppe** von G, falls U nicht leer ist und mit $a, b \in U$ auch $a \cdot b \in U$ und $a^{-1} \in U$ folgen.

Beispiele 4.9.4.

(i) $\{1, i, -1, -i\}$ bilden eine bezüglich der Multiplikation eine Untergruppe der multiplikativen Gruppe $\mathbb{C} \setminus \{0\}$ der komplexen Zahlen.
(ii) Die geraden Zahlen bilden eine additive Untergruppe der ganzen Zahlen $\mathbb{Z}$.

4.10. **Komplemente und direkte Summen.**

Definition 4.10.1. Sei U ein Unterraum des Vektorraumes V. Dann ist W ein **Komplement** von U, wenn

(i) $U \cap W = \{0\}$ und
(ii) $U + W = V$ gelten.

Ist W ein Komplement von U in V, so heißt V die direkte Summe der Unterräume U und W, $V = U \oplus W$.

Bemerkung 4.10.2. Die Summe ist hier wie in Definition 2.4.9 erklärt. Im allgemeinen gibt es mehrere Komplemente eines Unterraumes.

Beispiele 4.10.3.

(i) Sei $V = \mathbb{R}^n$. Sei $\{a_1, \ldots, a_n\}$ eine Basis von V. Dann ist $\langle a_{k+1}, \ldots, a_n\rangle$ ein Komplement von $\langle a_1, \ldots, a_k\rangle$. Ein anderes ist $\langle a_{k+1} + a_1, a_{k+2}, \ldots, a_n\rangle$. Also gilt
$$V = \langle a_1, \ldots, a_k\rangle \oplus \langle a_{k+1}, \ldots, a_n\rangle.$$
(ii) Sei $V = F[X]$ und $U = \{P(X) \in V \colon P(0) = 0\}$. Dann ist der Unterraum W, der alle konstanten Polynome enthält, ein Komplement von U in V.

Theorem 4.10.4. *Seien U, W Unterräume von V. Dann ist W genau dann ein Komplement von U in V, wenn jedes $v \in V$ sich in eindeutiger Weise als $u + w$ mit $u \in U$ und $w \in W$ schreiben lässt.*

Beweis. Gelte $U \oplus W = V$. Sei $v \in V$ beliebig. Dann gibt es wegen $U + W = V$ Vektoren $u \in U$ und $w \in W$ mit $u + w = v$. Seien $u' \in U$ und $w' \in W$ weitere Vektoren mit $u' + w' = v$. Dann folgt $u - u' = w' - w \in U \cap W = \{0\}$. Es folgen $u = u'$ und $w = w'$.

Aus der Darstellbarkeit folgt $U + W = V$. Wäre $U \cap W$ strikt größer als $\{0\}$, $0 \neq v \in U \cap W$, so könnten wir 0 nichttrivial als $v - v \in U + W$ kombinieren. Widerspruch. □

Theorem 4.10.5. *Sei V ein Vektorraum und U ein Unterraum. Dann besitzt U ein Komplement W, $U \oplus W = V$.*

Beweis. Sei $\{a_i\}_{i\in I}$ eine Basis von U. Dann können wir diese Basis durch eine Familie $\{b_j\}_{j\in J}$ zu einer Basis von V ergänzen. Setze $\langle\{b_j\}_{j\in J}\rangle =: W$. Dann gilt $U + W = V$.

Zu $U \cap W = \{0\}$: Gelte nach einer Umbenennung der Vektoren a_i und b_i ohne Einschränkung
$$v = \sum_{i=1}^{n} \lambda^i a_i = \sum_{j=1}^{m} \mu^j b_j.$$
Dann folgt $0 = \sum_{i=1}^{n} \lambda^i a_i - \sum_{j=1}^{m} \mu^j b_j$. Da $\{a_i, b_j\}$ eine Basis von V ist, folgt $0 = \lambda^i = \mu^j$ für alle $1 \leq i \leq n$ und $1 \leq j \leq m$. Somit ist $v = 0$. □

5. Lineare Abbildungen

5.1. **Dimensionsformel.** Erinnerung: $f \colon V \to W$ heißt linear, falls
$$f(\lambda a + \mu b) = \lambda f(a) + \mu f(b)$$
für alle $\lambda, \mu \in F$ und $a, b \in V$ gilt.

Beispiel 5.1.1. Sei $V = U \oplus W$. Definiere $p_U \colon V \to V$, die Projektion auf U, wie folgt: Sei $v \in V$. Dann gibt es $u \in U$ und $w \in W$ mit $v = u + w$. Setze $p_U(v) := u$. p_U ist eine lineare Abbildung.

Definition 5.1.2. Sei $f \colon V \to W$ linear. Dann definieren wir

(i) den **Kern** von f als
$$\ker f := \{v \in V \colon f(v) = 0\} = f^{-1}(\{0\})$$
(ii) und das **Bild** von f als
$$\operatorname{im} f := \{f(v) \colon v \in V\} = f(V) = \{w \in W \colon (\exists\, v \in V \colon f(v) = w)\}.$$

Theorem 5.1.3. *Sei $f \colon V \to W$ linear. Dann ist* $\ker f$ *ein Unterraum von V und* $\operatorname{im} f$ *ein Unterraum von W.*

Beweis.

(i) Seien $v, w \in \ker f$, $\lambda, \mu \in F$. Dann gilt
$$f(\lambda v + \mu w) = \lambda f(v) + \mu f(w) = \lambda \cdot 0 + \mu \cdot 0 = 0.$$
Wegen $f(0) = 0$ ist $0 \in \ker f$ und $\ker f \neq \emptyset$.

(ii) Seien $a, b \in \operatorname{im} f$ und $u, v \in V$ mit $f(u) = a$ sowie $f(v) = b$. Dann folgt für $\lambda, \mu \in F$
$$\lambda a + \mu b = \lambda f(u) + \mu f(v) = f(\lambda u + \mu v).$$
Schließlich ist wiederum wegen $f(0) = 0$ auch $0 \in \operatorname{im} f$. □

Wir erinnern an

Definition 5.1.4. ⋆ Sei $f \colon V \to W$ eine (lineare) Abbildung. Dann heißt f

(i) **injektiv,** falls aus $f(a) = f(b)$ stets $a = b$ folgt;
(ii) **surjektiv,** falls $\operatorname{im} f = W$ gilt (d. h. für jedes $w \in W$ existiert ein $a \in V$ mit $f(a) = w$).

Theorem 5.1.5. *Sei $f \colon V \to W$ linear. Dann ist f genau dann injektiv, wenn* $\ker f = \{0\}$ *gilt.*

Beweis.

(i) Ist f injektiv, so folgt aus $f(a) = 0 = f(0)$, dass $a = 0$ gilt. Somit ist $\ker f = \{0\}$.
(ii) Ist f nicht injektiv, so finden wir $a \neq b \in V$ mit $f(a) = f(b)$. Somit ist $f(a - b) = 0$ und $0 \neq a - b \in \ker f$. □

Theorem 5.1.6. *Sei $f \colon V \to W$ linear. Dann ist f genau dann ein Isomorphismus, wenn f injektiv und surjektiv ist.*

Beweis.

(i) Sei f ein Isomorphismus. Dann existiert eine lineare Abbildung $g \colon W \to V$ mit $g \circ f = \mathbb{1}_V$ und $f \circ g = \mathbb{1}_W$. Sei $f(a) = f(b)$. Nach beidseitiger Anwendung von g erhalten wir
$$a = g \circ f(a) = g \circ f(b) = b.$$
Somit ist f injektiv. Sei $w \in W$ beliebig, $a := g(w)$. Dann folgt $f(a) = f \circ g(w) = w$. Somit ist f surjektiv.
(ii) Sei f injektiv und surjektiv, so gibt es (Surjektivität) zu jedem $w \in W$ ein eindeutig (Injektivität) bestimmtes $v \in V$ mit $f(v) = w$. Definiere $g(w) := v$. $g \colon W \to V$ ist eine Funktion, d. h. auf ganz W eindeutig definiert. Dann gelten nach Definition $g \circ f = \mathbb{1}_V$ und $f \circ g = \mathbb{1}_W$. (Die Linearität von g hatten wir in der Definition 4.4.4 nicht gefordert, sie folgt aber aus Lemma 4.4.6.) □

Lemma 5.1.7. *Sei $\{a_i\}_{i\in I}$ eine Basis von V und seien $b_i \in W$, $i \in I$, beliebig. Dann gibt es genau eine lineare Abbildung $f\colon V \to W$ mit $f(a_i) = b_i$ für alle $i \in I$.*

Beweis. Sei $J \subset I$ eine beliebige endliche Teilmenge. Sei $a = \sum_{i\in J} \lambda^i a_i$ eine beliebige Linearkombination der Vektoren a_i, also ein beliebiger Vektor in V. Definiere $f(a) := \sum_{i\in J} \lambda^i f(a_i) = \sum_{i\in J} \lambda^i b_i$.

Diese Abbildung f ist linear: Seien nämlich $a, b \in V$ und $\lambda, \mu \in F$ beliebig. Ohne Einschränkung wollen wir $a = \sum_{i=1}^{n} \alpha^i a_i$ und $b = \sum_{i=1}^{n} \beta^i a_i$ annehmen, dass sich also beide Vektoren aus denselben Basisvektoren linear kombinieren lassen und dass diese die angegebenen Indices tragen. Es folgt

$$\begin{aligned} f(\lambda a + \mu b) &= f\left(\lambda \sum_{i=1}^{n} \alpha^i a_i + \mu \sum_{i=1}^{n} \beta^i a_i\right) = f\left(\sum_{i=1}^{n} \left(\lambda\alpha^i + \mu\beta^i\right) a_i\right) \\ &= \sum_{i=1}^{n} \left(\lambda\alpha^i + \mu\beta^i\right) b_i = \lambda \sum_{i=1}^{n} \alpha^i b_i + \mu \sum_{i=1}^{n} \beta^i b_i = \lambda f(a) + \mu f(b). \end{aligned}$$

Zur Eindeutigkeit: Sei $g\colon V \to W$ eine weitere solche Abbildung. Dann folgt für einen beliebigen Vektor $a = \sum_{j\in J} \lambda^j a_j$, $J \subset I$ endlich, ohne Einschränkung $J = \{1, \dots, n\}$,

$$g\left(\sum_{i=1}^{n} \lambda^i a_i\right) = \sum_{i=1}^{n} \lambda^i g(a_i) = \sum_{i=1}^{n} \lambda^i f(a_i) = f\left(\sum_{i=1}^{n} \lambda^i a_i\right). \qquad \square$$

Korollar 5.1.8. *Sei $\{a_i\}_{i\in I}$ eine linear unabhängige Teilmenge von V und seien $b_i \in W$, $i \in I$, beliebig. Dann gibt es eine lineare Abbildung $f\colon V \to W$ mit $f(a_i) = b_i$ für alle $i \in I$.*

Beweis. Ergänze $\{a_i\}_{i\in I}$ zu einer Basis von V, wähle die zugehörigen b_i's beliebig und wende Lemma 5.1.7 an. $\square$

Direkt aus dem Beweis von Lemma 5.1.7 folgt auch

Korollar 5.1.9. *Sei $\{a_i\}_{i\in I}$ eine Basis von V und seien $b_i \in W$, $i \in I$, beliebig. Sei f die eindeutig bestimmte lineare Abbildung $f\colon V \to W$ mit $f(a_i) = b_i$. Dann gilt* $\operatorname{im} f = \langle \{b_i\}_{i\in I} \rangle$.

Theorem 5.1.10 (Dimensionsformel)**.** *Sei $f\colon V \to W$ eine lineare Abbildung. Sei* $\dim V = n$. *Dann gilt*

$$\dim(\ker f) + \dim(\operatorname{im} f) = n = \dim V.$$

Beweis. Da $\ker f$ ein Unterraum von V ist, gibt es eine Basis $\{a_1, \dots, a_p\}$ von $\ker f$. Ergänze diese zu einer Basis $\{a_1, \dots, a_n\}$ von V. Nach Korollar 5.1.9 ist $\operatorname{im} f$ von

$$f(a_1) = 0, \dots, f(a_p) = 0, f(a_{p+1}), \dots, f(a_n)$$

erzeugt. Daher genügt es zu zeigen, dass die Vektoren $f(a_{p+1}), \dots, f(a_n)$ linear unabhängig sind. Sonst hätten wir λ^i's, nicht alle gleich Null, mit $\sum_{i=p+1}^{n} \lambda^i f(a_i) = 0$, also $0 = f\left(\sum_{i=p+1}^{n} \lambda^i a_i\right)$. Widerspruch, denn der Nicht-Nullvektor $\sum_{i=p+1}^{n} \lambda^i a_i$ läge dann ebenfalls in $\ker f$. Somit gilt

$$\dim \ker f + \dim \operatorname{im} f = p + (n - p) = n = \dim V. \qquad \square$$

Bemerkung 5.1.11. Definiert man im Beweis der Dimensionsformel den Unterraum U durch $U := \langle a_{p+1}, \ldots, a_n \rangle$, so ist $V = \ker f \oplus U$. Dann ist $f|_U \colon U \to W$, die Einschränkung von f auf den Unterraum $U \subset V$ ebenfalls linear und der obige Beweis zeigt, dass $f|_U$ injektiv ist.

Definition 5.1.12. Sei G eine Gruppe, $U \subset G$ eine Untergruppe. Wir definieren auf G eine Relation $\sim_U$ durch

$$a \sim_U b \quad \Longleftrightarrow \quad b^{-1}a \in U.$$

Lemma 5.1.13. *Die Relation $\sim_U$ ist eine Äquivalenzrelation.*

Beweis.
(i) $\sim_U$ ist reflexiv, denn es gilt $a^{-1}a = e \in U$.
(ii) $\sim_U$ ist symmetrisch, denn aus $a \sim_U b$ folgt $b^{-1}a \in U$, also auch $\left(b^{-1}a\right)^{-1} = a^{-1}b \in U$ und somit ist $b \sim_U a$.
(iii) $\sim_U$ ist transitiv, denn aus $a \sim_U b$ und $b \sim_U c$ folgen $b^{-1}a, c^{-1}b \in U$, also auch $c^{-1}bb^{-1}a = c^{-1}a \in U$ und somit $a \sim_U c$. □

Die folgende Definition benötigen wir später nicht, wenn wir uns nur für Quotientenräume von abelschen Gruppen oder Vektorräumen interessieren.

Definition 5.1.14 (Normalteiler)**.** Ein **Normalteiler** N einer Gruppe G ist eine Untergruppe von G, so dass für alle $g \in G$ die Relation $gN = Ng$ gilt, wobei $gN = \{ga \colon a \in N\}$ ist (Linksnebenklasse) und Ng analog definiert ist.

Beispiele 5.1.15.
(i) Jede Untergruppe einer abelschen Gruppe ist ein Normalteiler.
(ii) In der Gruppe der regulären $(n \times n)$-Matrizen über einem Körper bilden die Vielfachen der Einheitsmatrix $\lambda \mathbb{1}$, $\lambda \in F^\times$, einen Normalteiler.

Lemma 5.1.16. *N ist genau dann ein Normalteiler von G, wenn für alle $x \in N$ und für alle $g \in G$ auch $g^{-1}xg \in N$ gilt.*

Beweis. $gN = Ng$ ist äquivalent zu $gNg^{-1} = N$ oder zu den beiden Aussagen

$$\forall\, x \in N\, \forall g \in G\, \exists y \in N \colon gxg^{-1} = y$$

und

$$\forall\, y \in N\, \forall\, g \in G\, \exists\, x \in N \colon gxg^{-1} = y.$$

Die erste Aussage ist gerade die Behauptung und die zweite ist äquivalent dazu (man multipliziere mit g von rechts und mit g^{-1} von links und setze $h = g^{-1}$). □

Theorem 5.1.17. *Sei G eine Gruppe mit Normalteiler N. Dann ist*

$$G/N := \{gN \colon g \in G\}$$

mit vertreterweise definierter Verknüpfung eine Gruppe.

Beweis.
(i) Setze $aN \cdot bN := (a \cdot b)N$. Dies ist wohldefiniert, denn seien $a' = an$ und $b' = bm$ mit $n, m \in N$, so folgt $a'N \cdot b'N = anbmN = anbN = anNb = aNb = abN$, wobei wir $nN = N$ verwendet haben. Zeige dies: Es gilt nach Definition einer Untergruppe $nN \subset N$ und Multiplikation mit n^{-1} von links liefert $N \subset n^{-1}N$; da $n \in N$ beliebig ist, folgt $nN = N$.
(ii) Setze $(aN)^{-1} := a^{-1}N$.
(iii) Das neutrale Element ist eN mit dem neutralen Element $e \in G$.
(iv) Die Assoziativität $(aN \cdot bN) \cdot cN = aN \cdot (bN \cdot cN)$ folgt nach Definition der Multiplikation aus der Assoziativität von G.

(v) Ebenso folgen $aN \cdot eN = aN = eN \cdot aN$
(vi) und $aN \cdot a^{-1}N = eN = a^{-1}N \cdot aN$. □

Theorem 5.1.18. *Sei V ein F-Vektorraum. Sei $U \subset V$ ein Unterraum. Definiere $\sim_U$ durch $a \sim_U b$ genau dann, wenn $a-b \in U$. Dann ist $\sim_U$ eine Äquivalenzrelation. Der Faktorraum*

$$V/U := \{v + U \colon v \in V\}$$

mit repräsentantenweise definierter Verknüpfung ist wiederum ein Vektorraum.

Beweis. $(V, +)$ ist eine abelsche Gruppe mit Untergruppe (und damit auch Normalteiler) U. Daher übertragen sich die Kommutativität der Addition (V1), die Assoziativität der Addition (V2), die Eigenschaft des Nullelementes (V3) und des Inversen (V4) direkt.

Die Skalarmultiplikation ist durch $\lambda \cdot (a + U) := \lambda \cdot a + U$ definiert. Sie ist wohldefiniert, denn aus $a + U = b + U$, also $a - b \in U$ folgt $\lambda a - \lambda b = \lambda(a - b) \in U$. Die vertreterweise definierten Distributiv- und Assoziativgesetze der Skalarmultiplikation ((V5), (V6) und (V7)) übertragen sich direkt, ebenso gilt $1 \cdot (a + U) = (1 \cdot a) + U = a + U$ für $1 \in F$. □

Bemerkung 5.1.19. Ein Diagramm aus Räumen und Pfeilen, die Abbildungen zwischen diesen Räumen entsprechen, heißt kommutativ, wenn folgendes gilt: Starte in einem Raum und folge den Pfeilen. Dann ist das Ergebnis in einem anderen Raum unabhängig von der speziellen Wahl der Pfeile. Beispiel:

$$\begin{array}{ccc} A & \xrightarrow{f} & B \\ {\scriptstyle\varphi}\downarrow & & \downarrow{\scriptstyle\psi} \\ C & \xrightarrow[g]{} & D \end{array}$$

ist genau dann kommutativ, wenn $\psi(f(a)) = g(\varphi(a))$ für alle $a \in A$ gilt. Wir schreiben auch

$$\begin{array}{ccc} A & \xrightarrow{f} & B \\ {\scriptstyle\varphi}\downarrow & /\!/\!/ & \downarrow{\scriptstyle\psi} \\ C & \xrightarrow[g]{} & D \end{array}$$

um die Kommutativität anzudeuten.

Weiterhin verwenden wir die Pfeile $A \xhookrightarrow{i} B$ für eine injektive und $A \xrightarrow{p} \!\!\!\!\rightarrow B$ für eine surjektive Abbildung. „$\cong$“ an einem Pfeil deutet an, dass es sich um einen Isomorphismus handelt.

Versionen des folgenden Theorems gibt es auch für andere Homomorphismen.

Theorem 5.1.20 (Homomorphiesatz)**.** *Sei $f \colon V \to W$ eine lineare Abbildung. Dann gibt es eine Zerlegung von f*

$$\begin{array}{ccccc} V & & \xrightarrow{\quad f \quad} & & W, \\ & {\scriptstyle\pi}\searrow & & \nearrow{\scriptstyle i} & \\ & V/\ker f & \xrightarrow[\cong]{\bar{f}} & \operatorname{im} f & \end{array}$$

wobei $\pi \colon V \to V/\ker f$, $v \mapsto v + \ker f \equiv [v]$, die Projektionsabbildung, $i \colon \operatorname{im} f \to W$ die Inklusionsabbildung $w \mapsto w$ und $\bar{f} \colon V/\ker f \to \operatorname{im} W$ die von f induzierte Abbildung $\bar{f}([a]) := f(a)$ ist. Das obige Diagramm ist kommutativ. $\bar{f}$ ist ein Isomorphismus.

Beweis. Sei $a \in V$. Dann gilt $\pi(a) = a + \ker f$, $\bar{f}(\pi(a)) = f(a)$ und $i\left(\bar{f}(\pi(a))\right) = f(a)$. Die Abbildung $\bar{f}$ ist wohldefiniert, da für $[a] = [b]$ ein $c \in \ker f$ existiert, so dass $a = b + c$ ist. Somit folgt $f(a) = f(b) + 0$.

$\bar{f}$ ist surjektiv, da $\operatorname{im} f$ aus Punkten der Form $f(a)$ besteht und $\bar{f}([a]) = f(a)$ ist. $\bar{f}$ ist injektiv, da aus $\bar{f}([a]) = \bar{f}([b])$ auch $f(a) = f(b)$ oder $f(a - b) = 0$ folgt. Somit ist $a - b \in \ker f$ und daher gilt $[a] = [b]$. □

5.2. **Matrizen.** Wir wollen lineare Abbildungen mit Hilfe von Matrizen darstellen.

Sei $f\colon V \to W$ eine lineare Abbildung. Seien $\{a_1, \dots, a_n\}$ und $\{b_1, \dots, b_m\}$ Basen von V bzw. W. Nach Lemma 5.1.7 definieren die Bilder $f(a_j)$, $1 \le j \le n$, die Abbildung f eindeutig. Es gibt $(a_j^i)_{\substack{1 \le i \le m \\ 1 \le j \le n}}$, so dass

$$f(a_j) = \sum_{i=1}^{m} a_j^i b_i$$

für alle $1 \le j \le n$ gilt. Sei $a = \sum\limits_{j=1}^{n} x^j a_j$ ein beliebiger Vektor aus V mit Koordinaten $(x^1, \dots, x^n)$. Mit Hilfe von a_j^i wollen wir angeben, welche Koordinaten $(y^1, \dots, y^m)$ sein Bild $f(a) = \sum\limits_{i=1}^{m} y^i b_i$ hat. Es gilt

$$f(a) = f\left(\sum_{j=1}^{n} x^j a_j\right) = \sum_{j=1}^{n} x^j f(a_j) = \sum_{j=1}^{n}\sum_{i=1}^{m} x^j a_j^i b_i = \sum_{i=1}^{m}\left(\sum_{j=1}^{n} x^j a_j^i\right) b_i. \tag{5.1}$$

Es folgt

$$y^i = \sum_{j=1}^{n} x^j a_j^i.$$

Wir erhalten

Theorem 5.2.1. *Seien V, W endlichdimensionale F-Vektorräume. Sei $(a_1, \dots, a_n)$ eine Basis von V und sei $(b_1, \dots, b_m)$ eine Basis von W. Dann wird eine lineare Abbildung $f\colon V \to W$ vollständig durch eine $(m \times n)$-Matrix A beschrieben. Es ist*

$$A = \begin{pmatrix} a_1^1 & a_2^1 & \dots & a_n^1 \\ a_1^2 & a_2^2 & \dots & a_n^2 \\ \vdots & \vdots & & \vdots \\ a_1^m & a_2^m & \dots & a_n^m \end{pmatrix}.$$

In der k-ten Spalte von A stehen dabei gerade die Koordinaten des Vektors $f(a_k)$ bezüglich der Basis $(b_1, \dots, b_m)$.

Umgekehrt definiert eine $(m \times n)$-Matrix A vermöge (5.1) eine eindeutig bestimmte lineare Abbildung $f\colon V \to W$, so dass die k-te Spalte von A gerade die Koordinaten von $f(a_k)$ bezüglich der Basis $(b_1, \dots, b_m)$ enthält.

A heißt die zur linearen Abbildung f gehörige Matrix bezüglich der beiden Basen $(a_1, \dots, a_n)$ und $(b_1, \dots, b_m)$ oder ***darstellende Matrix.***

Weiterhin folgt aus den obigen Überlegungen

Theorem 5.2.2. *Sei f wie oben und seien die Basen wiederum fixiert. Dann erhalten wir aus f eine Abbildung für die Koordinatenvektoren*

$$x = \begin{pmatrix} x^1 \\ x^2 \\ \vdots \\ x^n \end{pmatrix} \quad \textit{und} \quad \begin{pmatrix} y^1 \\ y^2 \\ \vdots \\ y^m \end{pmatrix}.$$

Ist x der Koordinatenvektor von $a \in V$, so erfüllt der Koordinatenvektor y von $f(a) \in W$

$$y = Ax.$$

Theorem 5.2.3. *Seien $f\colon U \to V$ und $g\colon V \to W$ lineare Abbildungen mit zugehörigen Matrizen A bzw. B. Dann ist $D = BA$ die zu $g \circ f$ gehörige Matrix.*

Beweis. Seien $(\alpha_i)_{1\le i\le l}$, $(\beta_j)_{1\le j\le m}$ und $(\gamma_k)_{1\le k\le n}$ Basen von U, V bzw. W. Die zu f gehörige Matrix sei $\left(a_j^i\right)_{\substack{1\le i\le m\\ 1\le j\le l}}$ und die zu g gehörige Matrix sei $\left(b_j^i\right)_{\substack{1\le i\le n\\ 1\le j\le m}}$. Es ist also zu zeigen, dass die zu $g \circ f\colon U \to W$ gehörige Matrix durch $\left(d_j^i\right)_{\substack{1\le i\le n\\ 1\le j\le l}}$ mit $d_j^i = \sum_{k=1}^{m} b_k^i a_j^k$ gegeben ist.

Nach Voraussetzung gelten

$$f(\alpha_j) = \sum_{i=1}^{m} a_j^i \beta_i \quad \text{und} \quad g(\beta_i) = \sum_{k=1}^{n} b_i^k \gamma_k.$$

Somit ist

$$g \circ f(\alpha_j) = \sum_{i=1}^{m} \sum_{k=1}^{n} a_j^i b_i^k \gamma_k$$

und die Behauptung folgt. □

Bemerkung 5.2.4. Da die Komposition von Abbildungen assoziativ ist, ist auch die Matrizenmultiplikation assoziativ.

5.3. **Rang einer linearen Abbildung.**

Definition 5.3.1. Sei $f\colon V \to W$ eine lineare Abbildung zwischen endlichdimensionalen Vektorräumen. Dann ist der **Rang** von f, $\operatorname{rang} f$, durch

$$\operatorname{rang} f := \dim \operatorname{im} f$$

definiert.

Theorem 5.3.2. *Sei $f\colon V \to W$ eine lineare Abbildung zwischen endlichdimensionalen Vektorräumen. Zu fixierten Basen $S = \{a_i\}$ von V und $T = \{b_j\}$ von W sei A die zu f gehörige Matrix. Dann gilt*

$$\operatorname{rang} f = \operatorname{rang} A.$$

Beweis. Der Raum $\operatorname{im} f$ ist durch die Vektoren $f(a_i)$ erzeugt. Bezüglich der Basis T sind diese Vektoren aber gerade die Spalten von A. Daher ist $\operatorname{rang} f$ gleich dem Spaltenrang von A. Die Behauptung folgt. □

Korollar 5.3.3. *Seien A und B zwei Matrizen, die bezüglich verschiedener Basen zu einer linearen Abbildung $f\colon V \to W$ gehören. Dann gilt* $\operatorname{rang} A = \operatorname{rang} B$.

Theorem 5.3.4. *Sei $f\colon V \to W$ eine lineare Abbildung zwischen endlichdimensionalen Vektorräumen. Dann gilt*

$$\operatorname{rang} f + \dim \ker f = \dim V.$$

Beweis. Mit Definition 5.3.1 folgt dies aus der Dimensionsformel, Theorem 5.1.10. □

Bemerkung 5.3.5. Wir können die Matrix $A = \left(a_j^i\right)$ eines linearen Gleichungssystems

$$\sum_{k=1}^{n} a_k^i x^k = b^i, \quad 1 \le i \le m,$$

als Matrix einer linearen Abbildung $f\colon F^n \to F^m$ bezüglich der Standardbasen betrachten. Dann erhalten wir:

(i) Das homogene lineare Gleichungssystem hat genau $\ker f$ als Lösungsraum und die Dimension des Lösungsraumes ist nach Theorem 5.3.4 gleich $n - \operatorname{rang} A$.
(ii) Das lineare Gleichungssystem ist genau dann lösbar, wenn $(b^1, \dots, b^n) \in \operatorname{im} f$. $\operatorname{im} f$ und der Spaltenraum von A stimmen überein.
(iii) Im Falle $\ker f = \{0\}$ hat das lineare Gleichungssystem höchstens eine Lösung. Nach Theorem 5.3.4 gilt $\ker f = \{0\}$ genau dann, wenn $\operatorname{rang} A = n$ gilt.

Theorem 5.3.6. *Sei $f\colon V \to W$ eine lineare Abbildung. Sei $\dim V = \dim W = n$. Dann sind die folgenden Aussagen äquivalent:*

(i) f ist ein Isomorphismus,
(ii) f ist injektiv,
(iii) f ist surjektiv.

Beweis. Dies folgt aus Theorem 5.1.6 und der Dimensionsformel

$$\dim \ker f + \dim \operatorname{im} f = \dim V. \qquad \square$$

Somit korrespondieren Isomorphismen genau zu regulären Matrizen.

5.4. **Basiswechsel.** Sei $f\colon V \to W$ eine lineare Abbildung. Sei

$$\{a_1, \dots, a_p, a_{p+1}, \dots, a_n\}$$

eine Basis von V, so dass $\{a_{p+1}, \dots, a_n\}$ eine Basis von $\ker f$ ist. Setze $f(a_i) =: b_i$ für $1 \le i \le p$. Dann ist die Familie $\{b_1, \dots, b_p\}$ linear unabhängig. Ergänzen wir sie zu einer Basis $\{b_1, \dots, b_p, b_{p+1}, \dots, b_m\}$ so hat die zu f gehörige Matrix die Form

$$\begin{pmatrix} 1 & 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & 0 & 0 & \dots & 0 \end{pmatrix}. \tag{5.2}$$

Diese Matrix hat p Einsen, m Zeilen und n Spalten.

Komplizierter wird es, wenn man eine Basis finden will, so dass $f\colon V \to V$ eine einfache Gestalt hat.

Bemerkung 5.4.1 (Transformationsregeln)**.** Sei $f\colon V \to W$ linear. Wir verwenden die folgenden Bezeichnungen

	V	W
Alte Basen	$S = \{a_1, \dots, a_n\}$	$T = \{b_1, \dots, b_m\}$
Alte Koordinaten	$x^1, \dots, x^n$	$y^1, \dots, y^m$
Transformationsmatrizen	$D \in F^{n\times n}$	$E \in F^{m\times m}$
Neue Basen	$S' = \{a'_1, \dots, a'_n\}$	$T' = \{b'_1, \dots, b'_m\}$
Neue Koordinaten	$(x')^1, \dots, (x')^n$	$(y')^1, \dots, (y')^m$

Nach Wahl von D und E gelten

$$x' = Dx \quad \text{und} \quad y' = Ey.$$

Sei f bezüglich S, T durch A und bezüglich S', T' durch A' dargestellt. Es gelten daher

$$y = Ax \quad \text{und} \quad y' = A'x'.$$

Daraus erhalten wir $Ey = A'Dx$ und daraus $y = E^{-1}A'Dx$ sowie durch Vergleich mit $y = Ax$ die Relation

$$A = E^{-1}A'D \quad \text{sowie} \quad A' = EAD^{-1}.$$

Korollar 5.4.2. *Seien S und S' zwei geordnete Basen von V. Sei D die Transformationsmatrix, die die Vektoren von S als die Linearkombinationen von Vektoren aus S' darstellt. Sei $f\colon V \to V$ bezüglich S durch A und bezüglich S' durch A' dargestellt. Dann gilt*

$$A' = DAD^{-1}.$$

Bei Abbildungen $f\colon V \to V$ wollen wir stets dieselbe Basis für den Definitions- und Zielbereich verwenden.

Kombinieren wir nun die spezielle Darstellung einer linearen Abbildung aus (5.2) mit den Transformationsregeln, so erhalten wir

Theorem 5.4.3. *Sei A eine beliebige $(m \times n)$-Matrix. Dann gibt es reguläre Matrizen $D \in F^{n\times n}$ und $E \in F^{m\times m}$, so dass die Matrix EAD^{-1} von der Form* (5.2) *ist.*

5.5. **Der Vektorraum** $\mathrm{Hom}(V,W)$**.** Seien V, W Vektorräume über F. Seien die Abbildungen $f, g\colon V \to W$ linear. Seien $\lambda, \mu \in F$. Definiere

$$(\lambda f + \mu g)(a) := \lambda f(a) + \mu g(a).$$

(Man rechnet nach, dass auch $\lambda f + \mu g$ wieder linear ist.) Wir erhalten

Theorem 5.5.1. *Seien V, W Vektorräume über F. Die Menge der linearen Abbildungen $V \to W$ mit der (wie oben) punktweise erklärten Addition und Skalarmultiplikation ist ein F-Vektorraum:* $\mathrm{Hom}(V,W)$.

Beweis. Übung. □

Theorem 5.5.2. *Seien V und W Vektorräume der Dimension n bzw. m über F. Die Zuordnung*

$$\mathrm{Hom}(V,W) \ni f \mapsto A \in F^{m\times n}$$

aus Theorem 5.2.1 ist ein Isomorphismus.

Beweis. Die Linearität ist klar. Die Zuordnung und ihre Umkehrung sind vor und in Theorem 5.2.1 beschrieben. □

Korollar 5.5.3. *Seien V, W endlichdimensionale Vektorräume mit Basen (a_i) bzw. (b_i). Dann bilden die durch $f_i^j(a_k) := b_i\delta_k^j$ definierten Abbildungen eine Basis von* $\mathrm{Hom}(V,W)$. *Insbesondere ist* $\dim \mathrm{Hom}(V,W) = \dim V \cdot \dim W$.

Die Aussage über die Basen gilt auch für beliebige Vektorräume.

Beweis. Diese Abbildungen entsprechen gerade den Matrizen, die in Zeile i und Spalte j eine 1 und sonst Nullen haben. □

Definition 5.5.4. Abbildungen $f \in \mathrm{Hom}(V,F)$ heißen (lineare) **Funktionale** auf V. Wir schreiben $V^* := \mathrm{Hom}(V,F)$. V^* heißt der zu V **duale Raum.**

Bemerkung 5.5.5.

(i) Nach Wahl einer Basis $(a_1, \ldots, a_n)$ von V und bezüglich der Basis 1 von F hat ein lineares Funktional die Gestalt

$$f\left(\sum_{i=1}^n x^i a_i\right) = \sum_{i=1}^n x^i f(a_i).$$

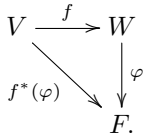
Wir bezeichnen die Abbildung $F^n \ni (x^1, \ldots, x^n) \mapsto \sum_{i=1}^n x^i f(a_i)$ auch als Linearform.

(ii) Für $0 \neq f \in V^* = \operatorname{Hom}(V, F)$ mit $\dim V = n$ gilt $\dim \ker f = n - 1$, da $\dim \operatorname{im} f = 1$ und $\dim \operatorname{im} f + \dim \ker f = n$ gelten.

(iii) Habe V die Basis $(a_1, \ldots, a_n)$. Dann bilden die Vektoren $f^i \in V^*$ mit $f^i(a_k) = \delta^i_k$ eine Basis von V^*. Wir schreiben auch $(a^*)^i = f^i$. Für $b = \sum_{i=1}^{n} x^i a_i$ gilt

$$(a^*)^j(b) = \sum_{i=1}^{n} x^i (a^*)^j(a_i) = \sum_{i=1}^{n} x^i \delta^j_i = x^j.$$

Definition 5.5.6 (Duale Abbildung)**.** Sei $f\colon V \to W$ linear. Dann definieren wir die **duale Abbildung** $f^*\colon W^* \to V^*$ zu f durch $f^*(\varphi) := \psi$ mit $\psi(\xi) = \varphi(f(\xi))$ für $\varphi \in W^*$ und $\xi \in V$.



(Es folgt aus der Definition, dass $f^*(\varphi) \in V^*$ gilt und dass f^* linear ist.)

Lemma 5.5.7. *Sei $f\colon V \to W$ durch die $(m \times n)$-Matrix $A = \left(a^i_j\right)_{\substack{1 \le i \le m \\ 1 \le j \le n}}$ bezüglich Basen $(a_i)_{1 \le i \le n}$ und $(b_j)_{1 \le j \le m}$ von V bzw. W dargestellt. Dann ist $f^*\colon W^* \to V^*$ durch die $(m \times n)$-Matrix $\left(b^i_j\right)_{\substack{1 \le i \le m \\ 1 \le j \le n}}$ mit $b^i_j = a^i_j$ bezüglich der Basen $\left((b^*)^j\right)_{1 \le j \le m}$ und $\left((a^*)^i\right)_{1 \le i \le n}$ in dem Sinne dargestellt, dass $f^*\left((b^*)^j\right) = \sum_{i=1}^{n} b^j_i (a^*)^i$ für alle $1 \le j \le m$ gilt.*

Für die Koordinaten erhalten wir die Abbildung

$$(\xi_1, \ldots, \xi_m) \mapsto \left(\sum_{j=1}^{m} \xi_j b^j_i\right)_{1 \le i \le n}.$$

Beweis. Nach Definition gilt

$$\begin{aligned}\left(f^*\left((b^*)^j\right)\right)(a_k) = \left((b^*)^j \circ f\right)(a_k) &= (b^*)^j\left(\sum_{i=1}^{m} a^i_k b_i\right) \\ &= \sum_{i=1}^{m} a^i_k (b^*)^j b_i = \sum_{i=1}^{m} a^i_k \delta^j_i = a^j_k.\end{aligned}$$

Aus $f^*\left((b^*)^j\right) = \sum_{l=1}^{n} b^j_l (a^*)^l$ folgt andererseits

$$\left(f^*\left((b^*)^j\right)\right)(a_k) = \sum_{l=1}^{n} b^j_l (a^*)^l(a_k) = \sum_{l=1}^{n} b^j_l \delta^l_k = b^j_k.$$

Die Behauptung für die Koordinaten ergibt sich direkt aus der Linearität. □

Bemerkung 5.5.8. Achtung, in der Darstellung als Matrix haben wir soeben über andere Indices als für Abbildungen $f\colon V \to W$ summiert. Die Koordinaten verändern sich folglich nach der Regel

$$(\xi_1, \ldots, \xi_m) \mapsto (\xi_1, \ldots, \xi_m) \begin{pmatrix} b^1_1 & \ldots & b^1_n \\ \vdots & & \vdots \\ b^m_1 & \ldots & b^m_n \end{pmatrix}.$$

Beachte dazu, dass wir die Basisvektoren von V^* anders als die von V oben indiziert haben. Die Koordinaten haben wir folglich unten indiziert. Somit handelt es sich um Zeilenvektoren.

In der Literatur weit verbreitet ist aber die (aus Kovarianzgründen nicht so saubere) Variante, auch diese Koordinaten als Spaltenvektoren zu schreiben, Indices generell unten anzubringen und den oberen Index einer Matrix an die erste Stelle zu senken, also a_{ij} statt a^i_j zu verwenden. Dann gilt für die Komponenten unter f^* die Abbildungsregel

$$\begin{pmatrix}\xi_1\\ \vdots\\ \xi_m\end{pmatrix} \mapsto \begin{pmatrix}b^1_1 & \dots & b^m_1\\ \vdots & & \vdots\\ b^1_n & \dots & b^m_n\end{pmatrix}\begin{pmatrix}\xi_1\\ \vdots\\ \xi_m\end{pmatrix}.$$

Wie man sich leicht überzeugt sind die Einträge im so erhaltenen Spaltenvektor dieselben wie beim Zeilenvektor. Hier ist nun die Transformationsmatrix die Matrix A^T, die transponierte Matrix, die für $A = (a^i_j)_{\substack{1\le i\le m\\ 1\le j\le n}}$ durch $A^T := (b^j_i)_{\substack{1\le j\le n\\ 1\le i\le m}}$ mit $b^j_i := a^i_j$ definiert ist. Graphisch erhält man

$$A^T = \begin{pmatrix}a^1_1 & a^2_1 & a^3_1 & \dots & a^m_1\\ a^1_2 & a^2_2 & a^3_2 & \dots & a^m_2\\ a^1_3 & a^2_3 & a^3_3 & \dots & a^m_3\\ \vdots & \vdots & \vdots & & \vdots\\ a^1_n & a^2_n & a^3_n & \dots & a^m_n\end{pmatrix} \quad \text{aus} \quad A = \begin{pmatrix}a^1_1 & a^1_2 & a^1_3 & \dots & a^1_n\\ a^2_1 & a^2_2 & a^2_3 & \dots & a^2_n\\ a^3_1 & a^3_2 & a^3_3 & \dots & a^3_n\\ \vdots & \vdots & \vdots & & \vdots\\ a^m_1 & a^m_2 & a^m_3 & \dots & a^m_n\end{pmatrix}.$$

5.6. **Endomorphismen.**

Definition 5.6.1 (Ring)**.** Ein **Ring** R ist eine additive abelsche Gruppe, d. h. wir schreiben die Verknüpfung als „+“, mit einer Multiplikation „·“, so dass folgende Axiome für alle $a, b, c \in R$ gelten

(R1) $a \cdot (b \cdot c) = (a \cdot b) \cdot c$ (Assoziativität)
(R2) $a \cdot (b + c) = a \cdot b + a \cdot c$ (Distributivität)
(R3) $(a + b) \cdot c = a \cdot c + b \cdot c$ (Distributivität)

Hierbei verwenden wir eine Klammersetzung nach der Regel „Punkt vor Strich“.

(R4) Ein Ring mit Einselement, also einem Element $1 = 1_R \in R$, das $a \cdot 1 = a = 1 \cdot a$ für alle $a \in R$ erfüllt, heißt Ring mit 1. Häufig wird die Existenz eines Einselementes bereits in der Definition eines Ringes verlangt.
(R5) Ein Ring heißt kommutativ, falls $a \cdot b = b \cdot a$ für alle $a, b \in R$ gilt.

Beispiele 5.6.2.

(i) $\mathbb{Z}$ ist ein kommutativer Ring mit Eins.
(ii) Die geraden Zahlen sind ein Ring ohne Eins.
(iii) Ein Körper ist ein kommutativer Ring mit Eins.
(iv) $F^{n\times n}$ ist ein Ring. Für $n \ge 2$ ist er nicht kommutativ.

Definition 5.6.3 (Ringhomomorphismus)**.** Seien R, S Ringe. Dann heißt eine Abbildung $\Phi\colon R \to S$ **Ringhomomorphismus,** wenn $\Phi(a + b) = \Phi(a) + \Phi(b)$ und $\Phi(a \cdot b) = \Phi(a) \cdot \Phi(b)$ für alle $a, b \in R$ gelten. Bei Ringen mit Eins verlangt man zusätzlich $\Phi(1) = 1$.

Der Deutlichkeit halber könnte man auch $\Phi(a +_R b) = \Phi(a) +_S \Phi(b)$ oder $\Phi(1_R) = 1_S$ schreiben.

Definition 5.6.4 (Algebra)**.** Eine **Algebra** A über einem Körper F ist ein Vektorraum mit einer Multiplikation „·“, so dass folgende Axiome für alle $a, b, c \in A$ und $\lambda \in F$ erfüllt sind:

(A1) $a \cdot (b \cdot c) = (a \cdot b) \cdot c$ (Assoziativität)
(A2) $a \cdot (b + c) = a \cdot b + a \cdot c$ (Distributivität)
(A3) $(a + b) \cdot c = a \cdot c + b \cdot c$ (Distributivität)
(A4) $(\lambda a) \cdot b = \lambda(a \cdot b) = a \cdot (\lambda b)$

(A5) Es existiert ein Einselement $1 = 1_A$ mit $a \cdot 1 = 1 \cdot a = a$ für alle $a \in A$.

Beispiele 5.6.5.

(i) $F^{n\times n}$ mit der Matrixmultiplikation ist eine Algebra.
(ii) Sei V ein F-Vektorraum. Dann ist $\mathrm{Hom}(V,V)$ eine Algebra. Lineare Selbstabbildungen eines Vektorraumes heißen auch Endomorphismen, $\mathrm{End}(V)$.
(iii) $F[X]$ ist eine Algebra über F und ein Ring, der Polynomring in der Variablen X über F.
(iv) $\mathbb{C}$ ist eine Algebra über $\mathbb{R}$.

Bemerkung 5.6.6. Sei V ein Vektorraum. Nach Wahl einer Basis ist jedem $f \in \mathrm{Hom}(V,V)$ eine Matrix A zugeordnet. Nach Theorem 5.5.2 ist diese Abbildung

$$\Phi\colon \mathrm{Hom}(V,V) \to F^{n\times n}$$

ein Isomorphismus von Vektorräumen. Dann gelten

$$\Phi(f\circ g) = \Phi(f)\cdot\Phi(g), \quad \Phi(f+g) = \Phi(f)+\Phi(g) \quad \text{sowie} \quad \Phi(\mathrm{id}) = \mathbb{1},$$

siehe insbesondere Theorem 5.2.3. Entsprechendes gilt für die Umkehrabbildung von Φ. Wir sagen daher, dass die Algebren $\mathrm{Hom}(V,V)$ und $F^{n\times n}$ isomorph sind.

Korollar 5.4.2 motiviert die folgende

Definition 5.6.7. Seien $A, A' \in F^{n\times n}$. Dann heißen A und A' **ähnlich,** wenn es eine reguläre Matrix D mit $A' = DAD^{-1}$ gibt.

Bemerkung 5.6.8. Äquivalent dazu hätte man auch zwei Matrizen ähnlich nennen können, wenn es eine lineare Abbildung und zugehörige Basen gibt, so dass diese Abbildung bezüglich der Basen durch diese Matrizen beschrieben wird.

Ziel: Im weiteren Verlauf der Vorlesung wollen wir uns intensiver mit der Frage beschäftigen, wie wir es erreichen können, dass die einer Abbildung $f\colon V \to V$ vermöge einer Basis zugeordnete Matrix möglichst einfach aussieht.

Als Vorbereitung definieren wir dazu

Definition 5.6.9. Ein Unterraum U von V heißt bezüglich einer linearen Abbildung $f\colon V \to V$ **invariant,** falls $f(U) \subset U$ gilt.

Beispiele 5.6.10.

(i) Betrachte $F[X]$ mit der linearen Abbildung $f(p(X)) := p'(X)$, der Ableitung von p,
$$f\left(\sum_{i=0}^{n} a_i X^i\right) := \sum_{i=1}^{n} i a_i X^{i-1}.$$
Die Unterräume der Polynome vom Grad $\le m$ sind beim Ableiten invariant. $\{p(X)\colon p(0)=0\}$ ist unter f nicht invariant.
(ii) Sei $f\colon V \to V$ eine beliebige lineare Abbildung. Dann sind $\ker f$ und $\mathrm{im}\, f$ invariante Unterräume.

Bemerkung 5.6.11.

(i) Sei U unter $f\colon V \to V$ invariant, $\dim V < \infty$. Sei $\{a_1,\dots,a_p\}$ eine Basis von U. Ergänze diese zu einer Basis $\{a_1,\dots,a_p,a_{p+1},\dots,a_n\}$ von V. Dann ist f bezüglich dieser Basis durch
$$\begin{pmatrix} a_1^1 & \dots & a_p^1 & a_{p+1}^1 & \dots & a_n^1 \\ \vdots & \ddots & \vdots & \vdots & & \vdots \\ a_1^p & \dots & a_p^p & a_{p+1}^p & \dots & a_n^p \\ 0 & \dots & 0 & a_{p+1}^{p+1} & \dots & a_n^{p+1} \\ \vdots & & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & a_{p+1}^n & \dots & a_n^n \end{pmatrix} \equiv \begin{pmatrix} A & B \\ 0 & C \end{pmatrix}$$

dargestellt, wobei A für einen $p \times p$-Block, B für einen $p \times (n-p)$-Block, 0 für einen $(n-p) \times p$-Block und C für einen $(n-p) \times (n-p)$-Block steht; hierbei ist ein Block eine rechteckige Anordnung von Zahlen wie aus der Gleichung ersichtlich.

(ii) Kann man gegebenenfalls auch noch $B = 0$ erreichen? Gelte $V = U_1 \oplus \ldots \oplus U_k$ für f-invariante Unterräume U_i. Dann hat die f darstellende Matrix bezüglich einer Basis, die aus den Basen der Unterräume U_i (in geeigneter Anordnung) besteht, die Blockgestalt

$$\begin{pmatrix} A_1 & 0 & \ldots & 0 \\ 0 & A_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \ldots & \ldots & A_k \end{pmatrix}.$$

Dabei ist A_i die Matrix, die $f|_{U_i}$ bezüglich einer Basis von U_i darstellt, $1 \leq i \leq k$.

Da eine solche Darstellung besonders einfach wird, wenn $\dim U_i = 1$ ist, definiert man Eigenvektoren und Eigenräume:

5.7. Eigenvektoren und Eigenräume.

Definition 5.7.1. Ein Vektor $a \in V$ mit $a \neq 0$ heißt **Eigenvektor** von $f\colon V \to V$ zum **Eigenwert** $\lambda \in F$, falls

$$f(a) = \lambda a$$

gilt. Der Unterraum

$$E_\lambda := \{b \in V\colon f(b) = \lambda b\}$$

heißt der zu λ gehörige **Eigenraum.** Die Dimension $\dim E_\lambda$ heißt **geometrische Vielfachheit** von λ.

Ebenso spricht man bei einer Matrix $A \in F^{n\times n}$ vom Eigenvektor x zum Eigenwert λ, falls

$$Ax = \lambda x$$

gilt. Eigenräume und geometrische Vielfachheit definiert man entsprechend.

Ein $\lambda \in F$ heißt Eigenwert von f bzw. A, wenn es einen zugehörigen Eigenvektor gibt.

Bemerkung 5.7.2.

(i) λ ist genau dann ein Eigenwert von f, wenn λ ein Eigenwert der f zugeordneten Matrix A ist.
(ii) $\ker f$ ist der Eigenraum zum Eigenwert 0. Ebenso gilt $\ker(f - \lambda\mathbb{1}) = E_\lambda$.
(iii) λ ist genau dann ein Eigenwert von f, wenn $f - \lambda\mathbb{1}$ singulär ist.
(iv) Die Eigenräume E_λ von f sind unter f invariant.
(v) Für $a \in E_\lambda$ gilt $f(a) = \lambda a$. Daher ist $f|_{E_\lambda}\colon E_\lambda \to E_\lambda$ bezüglich einer beliebigen Basis von der Form

$$\begin{pmatrix} \lambda & 0 & \ldots & 0 \\ 0 & \lambda & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda \end{pmatrix}.$$

Theorem 5.7.3. *Sein λ_i, $1 \leq i \leq m$, paarweise verschiedene Eigenwerte von f mit zugehörigen Eigenvektoren a_i, so sind die Vektoren a_i linear unabhängig.*

Beweis. Falls nicht, so gibt es eine nichttriviale Darstellung der Null. Wähle eine kürzeste nichttriviale Darstellung der Null, also eine solche, die möglichst wenige der Vektoren a_i verwendet. Ohne Einschränkung verwendet die kürzeste Darstellung $a_1, \dots, a_n$. Es gilt

$$0 = \sum_{i=1}^{n} \mu^i a_i$$

für geeignete $\mu^i \in F$. Wir erhalten einerseits

$$0 = f\left(\sum_{i=1}^{n} \mu^i a_i\right) = \sum_{i=1}^{n} \mu^i f(a_i) = \sum_{i=1}^{n} \mu^i \lambda_i a_i$$

und andererseits

$$0 = \lambda_1 \sum_{i=1}^{n} \mu^i a_i = \sum_{i=1}^{n} \lambda_1 \mu^i a_i.$$

Als Differenz erhalten wir

$$0 = \sum_{i=2}^{n} (\lambda_i - \lambda_1) \mu^i a_i.$$

Da dies eine kürzere Darstellung der Null als die ursprüngliche ist, muss sie trivial sein. Also gilt $(\lambda_i - \lambda_1)\mu^i a_i = 0$ für alle $2 \leq i \leq n$. Da die Eigenwerte paarweise verschieden sind, folgt bereits $\mu^i a_i = 0$ und somit auch $\mu^i = 0$ für $2 \leq i \leq n$. Also ist auch $\mu^1 a_1 = 0$. Wegen $a_1 \neq 0$, da es sich um einen Eigenvektor handelt, ist $\mu^1 = 0$, die Linearkombination der Null also doch trivial. Widerspruch. □

Korollar 5.7.4. *Sei V ein n-dimensionaler Vektorraum. Sei $f\colon V \to V$ ein Endomorphismus. Dann besitzt f höchstens n verschiedene Eigenwerte. Besitzt f genau n verschiedene Eigenwerte $\lambda_1, \dots, \lambda_n$, so bildet eine Auswahl zugehöriger Eigenvektoren $a_1, \dots, a_n$ eine Basis von V. Bezüglich dieser Basis hat f die Gestalt*

$$\begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}.$$

Bemerkung 5.7.5.

(i) Eine Matrix eines Endomorphismusses heißt diagonalisierbar, wenn es eine Basis gibt, so dass $A = \left(a^i_j\right)_{1 \leq i,j \leq n}$ diagonal ist, also $a^i_j = 0$ für $i \neq j$ erfüllt.

Eine solche Matrix ist genau dann diagonalisierbar, wenn es eine Basis aus Eigenvektoren gibt.

(ii) Die Matrix $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ ist nicht diagonalisierbar, da $\left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)$ (bis auf skalare Vielfache) der einzige Eigenvektor ist.

6. Determinanten

6.1. Polynome.

Bemerkung 6.1.1. Seien $p, q \in F[X]$. Dann gelten

(i) $\deg(p + q) \leq \max\{\deg p, \deg q\}$
(ii) $\det(p \cdot q) = \deg p + \deg q$

$a \in F$ heißt **Nullstelle** von p, wenn $p(a) = 0$ gilt, wobei $p(a)$ der Ausdruck ist, den wir erhalten, wenn wir in $p(X)$ überall die Variable X durch a ersetzen.

Theorem 6.1.2 (Polynomdivision)**.** *Seien $p, q \in F[X]$, $\deg q > 0$. Dann gibt es eindeutig bestimmte Polynome $s, r \in F[X]$ mit*

$$p(X) = s(X)q(X) + r(X)$$

und $\deg r < \deg q$.

Beweis. Zur Eindeutigkeit: Seien s, r wie behauptet und $s', r' \in F[X]$, so dass

$$p(X) = s'(X)q(X) + r'(X)$$

und $\deg r' < \deg q$ gelten. Wir erhalten

$$0 = (s(X) - s'(X))q(X) + (r(X) - r'(X)).$$

Ist $s \neq s'$, so ist $s(X) - s'(X)$ ein Polynom vom Grad ≥ 0. Also ist $0 + \deg q(X) \leq \deg((s(X) - s'(X))q(X)) = \deg(r'(X) - r(X)) < \deg q(X)$. Widerspruch.

Zur Existenz: Wir beweisen die Aussage per Induktion nach $\deg p$. Für $\deg p < \deg q$ können wir $s = 0$ und $r = p$ wählen. Sei also $\deg p \geq \deg q$. Nehme an, dass

$$p(X) = a_n X^n + \ldots + a_0 \quad \text{mit } a_n \neq 0$$

und

$$q(X) = b_m X^m + \ldots + b_0 \quad \text{mit } b_m \neq 0 \text{ und } m \leq n$$

gelten. Es folgt

$$\begin{aligned}
a_n X^n + \ldots + a_0 &= \frac{a_n}{b_m} X^{n-m} (b_m X^m + \ldots + b_0) + a_n X^n + a_{n-1} X^{n-1} + \ldots + a_0 \\
&\quad - a_n X^n - \frac{a_n b_{m-1}}{b_m} X^{n-1} - \frac{a_n b_{m-2}}{b_m} X^{n-2} - \ldots - \frac{a_n b_0}{b_m} X^{n-m} \\
&= \frac{a_n}{b_m} X^{n-m} q(X) + a_{n-1} X^{n-1} + \ldots + a_0 \\
&\quad - \frac{a_n b_{m-1}}{b_m} X^{n-1} - \frac{a_n b_{m-2}}{b_m} X^{n-2} - \ldots - \frac{a_n b_0}{b_m} X^{n-m} \\
&= \frac{a_n}{b_m} X^{n-m} q(X) + R(X),
\end{aligned}$$

wobei $R(X)$ ein Polynom vom Grad $\leq (n-1)$ ist. Nach Induktionsvoraussetzung existieren daher $A, B \in F[X]$ mit $\deg B < \deg q$, so dass

$$R(X) = A(X)q(X) + B(X)$$

gilt. Es folgt

$$p(X) = \left(\frac{a_n}{b_m} X^{n-m} + A(X) \right) q(X) + B(X)$$

wie behauptet. □

Korollar 6.1.3. *Sei $a \in F$ Nullstelle eines Polynomes $p \in F[X]$. Dann gibt es ein eindeutig bestimmtes Polynom $q(X)$, so dass*

$$p(X) = q(X)(X - a)$$

mit $\deg q = (\deg p) - 1$ *gilt.*

Beweis. Nach Bemerkung 6.1.1 ist die Aussage über den Grad klar.

Theorem 6.1.2 liefert eine Darstellung wie gefordert, $p(X) = q(X)(X-a)+r(X)$, aber möglicherweise mit einem zusätzlichen Polynom $r(X)$, das $\deg r < 1$ erfüllt und somit konstant ist. Auswerten an der Stelle $X = a$ liefert $r = 0$. □

Per Induktion erhält man hieraus

Korollar 6.1.4. *Ein Polynom $p(X)$ vom Grad $n \geq 1$ besitzt höchstens n Nullstellen.*

Definition 6.1.5. Sei $a \in F$ eine Nullstelle des Polynomes $p \in F[X]$. Dann besitzt p nach Korollar 6.1.3 (gegebenenfalls mehrfach angewandt) eine Darstellung der Form

$$p(X) = (X - a)^k q(X)$$

mit $k \in \mathbb{N}$ und $q \in F[X]$, so dass $q(a) \neq 0$ gilt. k heißt dann die **Vielfachheit der Nullstelle** a.

Wie das Beispiel $X^2 + 1 \in \mathbb{R}[X]$ zeigt, braucht ein Polynom keine Nullstelle zu besitzen. Es gilt aber

Theorem 6.1.6 (Fundamentalsatz der Algebra)**.** *Sei $p \in \mathbb{C}[X]$ mit $\deg p \geq 1$. Dann besitzt p eine Nullstelle.*

Beweis. Siehe Vorlesung „Funktionentheorie“. □

Korollar 6.1.7. *Sei $p \in \mathbb{C}[X]$ ein Polynom mit $\deg p = n > 0$. Dann gibt es komplexe Zahlen $d, \lambda_1, \ldots, \lambda_n$, so dass*

$$p(X) = d(X - \lambda_1)(X - \lambda_2) \cdots (X - \lambda_n)$$

ist. λ_i sind gerade die Nullstellen von p. Sie kommen entsprechend ihrer Vielfachheit ggf. mehrfach vor.

Beweis. Induktion, Übung. □

Lemma 6.1.8. *Sei $p \in \mathbb{C}[X]$ ein Polynom mit reellen Koeffizienten. Sei $a + ib$, $a, b \in \mathbb{R}$, eine Nullstelle von p, so ist auch die komplex konjugierte Zahl $\overline{a + ib} := a - ib$ eine Nullstelle von p.*

Beweis. Sei $p(x) = a_n X^n + \ldots a_1 X + a_0$ mit $a_i \in \mathbb{R}$. Setze $c := a + ib$. Es gilt

$$0 = \overline{p(c)} = \overline{a_n c^n + \ldots + a_1 c + a_0} = a_n \bar{c}^n + \ldots + a_1 \bar{c} + a_0 = p(\bar{c}).$$

Somit ist auch $\bar{c}$ eine Nullstelle von p. □

6.2. **Permutationen.**

Definition 6.2.1. Eine **Permutation** der Zahlen $1, \ldots, n$ ist eine bijektive Abbildung $\sigma \colon \{1, \ldots, n\} \to \{1, \ldots, n\}$.

Schematisch können wir eine Permutation durch

$$\begin{pmatrix} 1 & 2 & 3 & \ldots & n \\ \sigma(1) & \sigma(2) & \sigma(3) & \ldots & \sigma(n) \end{pmatrix}$$

oder kürzer durch

$$\begin{pmatrix} \sigma(1) & \sigma(2) & \sigma(3) & \ldots & \sigma(n) \end{pmatrix}$$

darstellen, im zweiten Fall auch mit Kommata getrennt. (Aufgrund der Bijektivität von σ kommt jede der Zahlen $1, 2, \ldots, n$ in jeder dieser Zeilen genau einmal vor.)

Bemerkung 6.2.2.

(i) Es gibt genau $n! := 1 \cdot 2 \cdots n$ Permutationen von $1, \ldots, n$.
(ii) Die Permutationen mit $(\sigma \circ \tau)(i) := \sigma(\tau(i))$ bilden eine Gruppe: S_n.
(iii) S_n ist genau für $n = 1, 2$ abelsch.

Beweis. Übung. □

Definition 6.2.3. Sei σ eine Permutation der Zahlen $1, 2, \ldots, n$. Sei $s = s(\sigma)$ die Anzahl der Paare (i, j) mit $1 \leq i < j \leq n$ und $\sigma(i) > \sigma(j)$. Dann definieren wir das **Signum** der Permutation σ durch

$$\operatorname{sign} \sigma := (-1)^{s(\sigma)}.$$

Eine Permutation σ mit $\operatorname{sign}(\sigma) = 1$ heißt **gerade,**
eine Permutation σ mit $\operatorname{sign}(\sigma) = -1$ heißt **ungerade.**

Definition 6.2.4 (Transposition)**.** Setze für $1 \leq i, j \leq n$ mit $i \neq j$

$$\tau_{ij}(k) := \begin{cases} j & k = i, \\ i & k = j, \\ k & \textit{sonst.} \end{cases}$$

Die Permutation vertauscht gerade die Zahlen i und j. Es gilt $\tau^{-1} = \tau$.

Theorem 6.2.5. *Sei σ eine Permutation der Zahlen $1, 2, \ldots, n$. Sei τ eine Transposition. Dann gilt*

$$\operatorname{sign}(\tau \circ \sigma) = -\operatorname{sign}(\sigma).$$

Beweis. Seien σ und σ' zwei Permutationen,

$$(\sigma(1), \sigma(2), \ldots, \sigma(n)) \quad \text{und} \quad (\sigma'(1), \sigma'(2), \ldots, \sigma'(n)),$$

die sich nur dadurch unterscheiden, dass in dieser Darstellung zwei benachbarte Einträge vertauscht wurden. Dann gilt nach Definition

$$\operatorname{sign} \sigma = -\operatorname{sign} \sigma'.$$

Sei $(\tau\sigma(1), \ldots, \tau\sigma(n))$ aus $(\sigma(1), \ldots, \sigma(n))$ durch Vertauschen der Zahlen i und j entstanden, die k weitere Zahlen in dieser Darstellung zwischen sich haben. Durch $2k + 1$ Vertauschungen nebeneinanderliegender Elemente kann man diese beiden Permutationen ineinander überführen. Daher gilt

$$\operatorname{sign}(\tau \circ \sigma) = (-1)^{2k+1} \operatorname{sign}(\sigma) = -\operatorname{sign}(\sigma). \qquad \square$$

Theorem 6.2.6. *Jede Permutation der Zahlen $1, 2, \ldots, n$ lässt sich als Produkt von Transpositionen schreiben.*

Beweis. Für die Identität und eine beliebige Permutation τ gilt $\tau \circ \tau = \mathbb{1}$. Sonst kann man durch Permutationen sukzessive erreichen, dass $(1, 2, \ldots, n)$ in eine Permutation überführt wird, bei der in dieser Darstellung das Anfangsstück, das mit $(\sigma(1), \sigma(2), \ldots, \sigma(n))$ übereinstimmt um mindestens eine Position länger wird. Die Aussage folgt nun per Induktion. $\square$

Theorem 6.2.7. *Sei σ eine Permutation, die sich als Produkt von r Transpositionen schreiben lässt. Dann gilt* $\operatorname{sign} \sigma = (-1)^r$.

Beweis. Benutze Theorem 6.2.5. $\square$

Da $\operatorname{sign} \sigma$ unabhängig von Transpositionen definiert ist, gilt

Korollar 6.2.8. *Lässt sich eine Permutation als Produkt von k und l Transpositionen schreiben, dann gilt*

$$k \equiv l \pmod 2,$$

d. h. k und l sind entweder beide gerade oder beide ungerade.

Theorem 6.2.9. *Sei $n \geq 2$. Dann gibt es genau $\frac{1}{2}n!$ gerade und $\frac{1}{2}n!$ ungerade Permutationen der Zahlen $1, 2, \ldots, n$.*

Beweis. Sei τ eine beliebige Transposition. Da die Permutationen eine Gruppe bilden, ist $\sigma \mapsto \tau\sigma$ bijektiv. Die geraden Permutationen werden dabei gerade auf die ungeraden und umgekehrt abgebildet. Folglich gibt es von jeder Sorte gleich viele. □

Bemerkung 6.2.10. Eine Untergruppe der symmetrischen Gruppe S_n ist die alternierende Gruppe. Sie besteht aus den Elementen σ der symmetrischen Gruppe mit $\operatorname{sign}\sigma = 1$.

Definition 6.2.11. Seien G, H Gruppen. Dann ist $\varphi\colon G \to H$ ein **Gruppenhomomorphismus,** wenn für alle $a, b \in G$

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

gilt.

Beispiele 6.2.12.

(i) Lineare Abbildungen sind Gruppenhomomorphismen, wenn wir einen Vektorraum als additive Gruppe betrachten.
(ii) Die Koordinatenabbildung eines Vektorraumes ist ein Gruppenhomomorphismus bezüglich der additiven Gruppe eines Vektorraumes.
(iii) Die Abbildung $\Phi\colon \operatorname{Hom}(V,V) \to F^{n\times n}$ wie in Bemerkung 5.6.6 ist ein Gruppenhomomorphismus.

Ein weiteres Beispiel für einen Gruppenhomomorphismus ergibt sich aus

Theorem 6.2.13. *Seien σ, τ zwei Permutationen. Dann gilt*

$$\operatorname{sign}(\sigma \circ \tau) = \operatorname{sign}(\sigma) \cdot \operatorname{sign}(\tau).$$

Beweis. Stelle die Permutationen als Verkettung von Transpositionen dar und zähle deren Anzahl. □

Korollar 6.2.14. *Sei σ eine Permutation. Dann gilt* $\operatorname{sign}\left(\sigma^{-1}\right) = \operatorname{sign}\sigma$.

Beweis. Es gilt $\operatorname{sign}(\sigma) \cdot \operatorname{sign}\left(\sigma^{-1}\right) = \operatorname{sign}\left(\sigma \circ \sigma^{-1}\right) = \operatorname{sign}\mathbb{1} = 1$. □

6.3. **Determinanten.** In diesem Kapitel werden wir häufiger zwischen einer Matrix $A = \left(a^i_j\right)_{1\le i,j\le n}$ und deren Spalten $a_j = \left(a^i_j\right)_{1\le i\le n} = \sum_{i=1}^{n} a^i_j e_i$ mit $e_i = (0,\dots,0,1,0,\dots,0)^T$ hin- und herwechseln. Ebenso werden wir zwischen $\det A$ und $\det(a_1,\dots,a_n)$ hin- und herwechseln.

Definition 6.3.1 (Determinante)**.** Eine Funktion $\det\colon F^{n\times n} \to F$, die n Vektoren in F^n ein Element in F zuordnet, $\det(a_1,\dots,a_n)$, heißt n-dimensionale Determinantenfunktion oder **Determinante**, wenn sie folgende Axiome erfüllt

(i) $\det(a_1,\dots,a_{i-1},\lambda a_i,a_{i+1},\dots,a_n) = \lambda\det(a_1,\dots,a_{i-1},a_i,a_{i+1},\dots,a_n)$ für alle $\lambda \in F$ und $1 \le i \le n$
(ii) $\det(a_1,\dots,a_{i-1},a_i+b_i,a_{i+1},\dots,a_n) = \det(a_1,\dots,a_{i-1},a_i,a_{i+1},\dots,a_n) + \det(a_1,\dots,a_{i-1},b_i,a_{i+1},\dots,a_n)$ für alle $1 \le i \le n$ (Linearität)
(iii) $\det(a_1,\dots,a_i,\dots,a_j,\dots,a_n) = 0$, falls $a_i = a_j$ für $i \ne j$
(iv) $\det(e_1,\dots,e_n) = 1$ (Normierung)

Da eine Determinantenfunktion in jedem Argument linear ist, heißt sie multilinear.

Bemerkung 6.3.2. ⋆ Die Determinante ist ein Polynom in den Einträgen. Wir werden später auch Matrizen statt Körperelementen einsetzen. Daher empfehlen wir, sich beim zweiten Lesen zu überlegen, dass sich die Resultate über Determinanten auf diesen Fall anpassen lassen.

Wir werden erst später zeigen, dass es eine eindeutig bestimmte Determinantenfunktion gibt.

Lemma 6.3.3. *Sei* det *eine Determinantenfunktion. Dann gilt*

(i) $$\det\left(a_1,\dots,a_{i-1},\sum_{k=1}^{m}\lambda^k b_k,a_{i+1},\dots,a_n\right) = \sum_{k=1}^{m}\lambda^k\det(a_1,\dots,a_{i-1},b_k,a_{i+1},\dots,a_n)$$

(ii) $$\det\left(a_1,\dots,a_{i-1},a_i+\sum_{\substack{k=1\\k\neq i}}^{n}\lambda^k a_k,a_{i+1},\dots,a_n\right) = \det(a_1,\dots,a_{i-1},a_i,a_{i+1},\dots,a_n)$$

(iii) $\det(a_1,\dots,a_i,\dots,a_j,\dots,a_n) = -\det(a_1,\dots,a_j,\dots,a_i,\dots,a_n)$ *für* $i\neq j$ *(Antisymmetrie)*

Beweis.

(i) Dies folgt direkt aus der Multilinearität.

(ii) Ist eine Folgerung aus der Multilinearität und aus (iii) in der Definition der Determinante.

(iii) Es gilt

$$\begin{aligned}\det(a_1,\dots,a_i,\dots,a_j,\dots,a_n) &= \det(a_1,\dots,a_i+a_j,\dots,a_j,\dots,a_n)\\ &= \det(a_1,\dots,a_i+a_j,\dots,a_j-a_i-a_j,\dots,a_n)\\ &= -\det(a_1,\dots,a_i+a_j,\dots,a_i,\dots,a_n)\\ &= -\det(a_1,\dots,a_j,\dots,a_i,\dots,a_n).\end{aligned}$$ □

Theorem 6.3.4. *Sei* det *eine Determinantenfunktion. Ist die Familie der Vektoren* $\{a_1,\dots,a_n\}$ *linear abhängig, so gilt*

$$\det(a_1,\dots,a_n)=0.$$

Beweis. Sei $a_i=\sum_{\substack{k=1\\k\neq i}}^{n}\lambda^k a_k$. Dann folgt

$$\det(a_1,\dots,a_n)=\sum_{\substack{k=1\\k\neq i}}^{n}\lambda^k\det(a_1,\dots,a_{i-1},a_k,a_{i+1},\dots,a_n)=0,$$

da in den Determinanten auf der rechten Seite stets zwei Vektoren doppelt vorkommen. □

Der Eindeutigkeitssatz für die Determinantenfunktion lautet

Theorem 6.3.5. *Sei* $n\geq 1$. *Dann gibt es höchstens eine* n*-dimensionale Determinantenfunktion.*

Beweis. Seien d, D zwei Determinantenfunktionen. Dann gilt

$$1=d(e_1,\dots,e_n)=D(e_1,\dots,e_n).$$

Sei $(i_1,\dots,i_n)$ eine Permutation von $1,\dots,n$. Da wir eine Permutation als Verkettung von Transpositionen schreiben können erhalten wir aus der Antisymmetrie

$$\begin{aligned}d(e_{i_1},\dots,e_{i_n}) &= \operatorname{sign}(i_1,\dots,i_n)d(e_1,\dots,e_n)\\ &= \operatorname{sign}(i_1,\dots,i_n)D(e_1,\dots,e_n)\\ &= D(e_{i_1},\dots,e_{i_n}).\end{aligned}$$

Sei nun $a_j = \sum_{i=1}^{n} a_j^i e_i$. Wir erhalten

$$\begin{aligned} d(a_1, \dots, a_n) &= \sum_{i_1=1}^{n} a_1^{i_1} \sum_{i_2=1}^{n} a_2^{i_2} \dots \sum_{i_n=1}^{n} a_n^{i_n} d(e_{i_1}, \dots, e_{i_n}) \\ &= \sum_{i_1=1}^{n} a_1^{i_1} \sum_{i_2=1}^{n} a_2^{i_2} \dots \sum_{i_n=1}^{n} a_n^{i_n} D(e_{i_1}, \dots, e_{i_n}) \\ &= D(a_1, \dots, a_n). \end{aligned}$$

Dabei haben wir das mittlere Gleichheitszeichen für den Fall, dass $(i_1, \dots, i_n)$ eine Permutation von $(1, \dots, n)$ ist, bereits nachgerechnet. Sonst stimmen jedoch mindestens zwei Indices überein und beide Terme verschwinden. Die Behauptung folgt. □

Der Existenzsatz für die Determinantenfunktion ist

Theorem 6.3.6. *Für $n \geq 1$ existiert eine Determinantenfunktion. Sie ist durch*

$$\det(a_1, \dots, a_n) \equiv \det A := \sum_{\sigma \in S_n} \operatorname{sign} \sigma \cdot a_1^{\sigma(1)} \cdot a_2^{\sigma(2)} \cdots a_n^{\sigma(n)}$$

gegeben. (Leibnizformel)

Beweis. Bereits aus dem Beweis des Eindeutigkeitssatzes folgt, dass es höchstens diese Determinantenfunktion geben kann.

Bei der angegebenen Funktion handelt es sich auch um eine Determinantenfunktion, denn sie ist offensichtlicherweise multilinear und normiert und wir erhalten für $a_j^i = a_k^i$ für alle i und $j \neq k$

$$\begin{aligned} \det A &= \det(a_1, \dots, a_j, \dots, \underbrace{a_k}_{=a_j}, \dots, a_n) \\ &= \sum_{\sigma \in S_n} \operatorname{sign} \sigma \cdot a_1^{\sigma(1)} \cdots a_j^{\sigma(j)} \cdots a_k^{\sigma(k)} \cdots a_n^{\sigma(n)} \\ &= \sum_{\substack{\sigma \in S_n \\ \sigma(j) < \sigma(k)}} \operatorname{sign} \sigma \cdot a_1^{\sigma(1)} \cdots a_j^{\sigma(j)} \cdots a_j^{\sigma(k)} \cdots a_n^{\sigma(n)} \\ &\quad + \sum_{\substack{\sigma \in S_n \\ \sigma(j) > \sigma(k)}} \operatorname{sign} \sigma \cdot a_1^{\sigma(1)} \cdots a_j^{\sigma(j)} \cdots a_j^{\sigma(k)} \cdots a_n^{\sigma(n)} = 0, \end{aligned}$$

denn für jeden Term in der obigen Summe gibt es einen Term in der unteren Summe, der zur gleichen Permutation gehört, nur dass $\sigma(j)$ und $\sigma(k)$ vertauscht sind. Da diese Permutationen sich also um genau eine Transposition unterscheiden, heben sich diese Terme gerade paarweise auf und die Behauptung folgt. □

Korollar 6.3.7. *Für quadratische Matrizen $A \in F^{n \times n}$, $C \in F^{m \times m}$ und $B \in F^{n \times m}$ gilt*

$$\det \begin{pmatrix} A & B \\ 0 & C \end{pmatrix} = \det A \cdot \det C.$$

Beweis. Sei $\left(a_j^i\right)_{1 \leq i,j \leq n+m} = \left(\begin{smallmatrix} A & B \\ 0 & C \end{smallmatrix}\right)$. In der Leibnizformel

$$\det \begin{pmatrix} A & B \\ 0 & C \end{pmatrix} = \sum_{\sigma \in S_{n+m}} \operatorname{sign} \sigma \cdot a_1^{\sigma(1)} \cdot a_2^{\sigma(2)} \cdots a_{n+m}^{\sigma(n+m)}$$

verschwinden alle Terme, für die $\sigma(i) > n$ für mindestens ein i mit $1 \leq i \leq n$ gilt. Zu $\sigma \in S_{n+m}$ mit $\sigma(i) \leq n$ für $1 \leq i \leq n$ gibt es Permutationen $\tau \in S_n$ und $\rho \in S_m$,

so dass $\sigma(i) = \tau(i)$ für alle $1 \leq i \leq n$ und $\sigma(n+i) = n+\rho(i)$ für alle $1 \leq i \leq n$ gilt. Stellt man τ und ρ als Verkettungen von Transpositionen dar, so kann man diese auch auf $1, \ldots, n$ bzw. $n+1, \ldots, m$ wirken lassen und erhält $\operatorname{sign}\sigma = \operatorname{sign}\tau \cdot \operatorname{sign}\rho$. Es folgt somit

$$\begin{aligned}
\det \begin{pmatrix} A & B \\ 0 & C \end{pmatrix} &= \sum_{\tau \in S_n} \sum_{\rho \in S_m} \operatorname{sign}\tau \cdot \operatorname{sign}\rho \cdot a_1^{\tau(1)} \cdots a_n^{\tau(n)} \cdot a_{n+1}^{n+\rho(1)} \cdots a_{n+m}^{n+\rho(m)} \\
&= \sum_{\tau \in S_n} \operatorname{sign}\tau \cdot a_1^{\tau(1)} \cdots a_n^{\tau(n)} \cdot \sum_{\rho \in S_m} \operatorname{sign}\rho \cdot a_{n+1}^{n+\rho(1)} \cdots a_{n+m}^{n+\rho(m)} \\
&= \det A \cdot \det C. \qquad \square
\end{aligned}$$

Per Induktion folgt

Korollar 6.3.8. *Sei $A = \left(a_j^i\right)_{1 \leq i,j \leq n}$ eine obere Dreiecksmatrix, d. h. gelte $a_j^i = 0$ für $i < j$. Dann gilt*

$$\det A = a_1^1 \cdots a_n^n = \prod_{i=1}^{n} a_i^i.$$

Theorem 6.3.9. *Sei A eine $(n \times n)$-Matrix. Dann gilt*

$$\det A = \det A^T.$$

Beweis. Sei $A = \left(a_j^i\right)$. Es folgt

$$\begin{aligned}
\det A &= \sum_{\sigma \in S_n} \operatorname{sign}\sigma \cdot a_1^{\sigma(1)} a_2^{\sigma(2)} \cdots a_n^{\sigma(n)} \\
&= \sum_{\sigma \in S_n} \underbrace{\operatorname{sign}\sigma}_{=\operatorname{sign}(\sigma^{-1})} \cdot a^1_{\sigma^{-1}(1)} a^2_{\sigma^{-1}(2)} \cdots a^n_{\sigma^{-1}(n)}
\end{aligned}$$

(nach Umsortieren der Terme nach dem oberen Index)

$$\begin{aligned}
&= \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) \cdot a^1_{\sigma(1)} a^2_{\sigma(2)} \cdots a^n_{\sigma(n)} \\
&= \det A^T. \qquad \square
\end{aligned}$$

Korollar 6.3.10. *Die Axiome, die wir für die Spalten einer Matrix in der Definition einer Determinantenfunktion gefordert haben, gelten somit auch für die Zeilen $a^i \equiv (a^i_j)_{1 \leq j \leq n}$, also etwa*

$$\det A = -\det \begin{pmatrix} a^2 \\ a^1 \\ a^3 \\ \vdots \\ a^n \end{pmatrix}.$$

Wir können die Berechnung von Determinanten von $(n \times n)$-Matrizen auf die von $((n-1) \times (n-1))$-Matrizen zurückführen.

Theorem 6.3.11 (Laplace). *Sei A eine $(n\times n)$-Matrix. Sei A^i_j die $((n-1)\times(n-1))$-Matrix, die entsteht, wenn wir die i-te Zeile und die j-te Spalte streichen.*

$$\begin{pmatrix} a_1^1 & \dots & a_{j-1}^1 & a_{j+1}^1 & \dots & a_n^1 \\ \vdots & & \vdots & \vdots & & \vdots \\ a_1^{i-1} & \dots & a_{j-1}^{i-1} & a_{j+1}^{i-1} & \dots & a_n^{i-1} \\ a_1^{i+1} & \dots & a_{j-1}^{i+1} & a_{j+1}^{i+1} & \dots & a_n^{i+1} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_1^n & \dots & a_{j-1}^n & a_{j+1}^n & \dots & a_n^n \end{pmatrix}$$

Dann gilt für festes j, $1 \le j \le n$,

$$\det A = \sum_{i=1}^{n} (-1)^{i+j} a_j^i \det A_j^i$$

und für festes i, $1 \le i \le n$

$$\det A = \sum_{j=1}^{n} (-1)^{i+j} a_j^i \det A_j^i.$$

Beweis. Wegen $\det A = \det A^T$ genügt es, die erste Gleichheit nachzuweisen.

Beide Seiten sind in der j-ten Spalte linear. Daher genügt es, die Aussage für Matrizen der Form

$$A = \begin{pmatrix} a_1^1 & \dots & a_{j-1}^1 & 0 & a_{j+1}^1 & \dots & a_n^1 \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_1^{i-1} & \dots & a_{j-1}^{i-1} & 0 & a_{j+1}^{i-1} & \dots & a_n^{i-1} \\ a_1^i & \dots & a_{j-1}^i & a_j^i & a_{j+1}^i & \dots & a_n^i \\ a_1^{i+1} & \dots & a_{j-1}^{i+1} & 0 & a_{j+1}^{i+1} & \dots & a_n^{i+1} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_1^n & \dots & a_{j-1}^n & 0 & a_{j+1}^n & \dots & a_n^n \end{pmatrix}$$

nachzuweisen. Ohne Einschränkung können wir $a_j^i \neq 0$ annehmen und daher genügt es aufgrund der Eigenschaften der Determinante (die rechte Seite bleibt ohnehin unverändert), anzunehmen, dass A von der Form

$$A = \begin{pmatrix} a_1^1 & \dots & a_{j-1}^1 & 0 & a_{j+1}^1 & \dots & a_n^1 \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_1^{i-1} & \dots & a_{j-1}^{i-1} & 0 & a_{j+1}^{i-1} & \dots & a_n^{i-1} \\ 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ a_1^{i+1} & \dots & a_{j-1}^{i+1} & 0 & a_{j+1}^{i+1} & \dots & a_n^{i+1} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_1^n & \dots & a_{j-1}^n & 0 & a_{j+1}^n & \dots & a_n^n \end{pmatrix}$$

ist. Weiterhin dürfen wir die ausgezeichnete Spalte mit der rechts davon vertauschen. Auf diese Weise bringen wir die ausgezeichnete Spalte ganz nach rechts. Ebenso dürfen wir die ausgezeichnete Zeile sukzessive mit der darunterliegenden Zeile vertauschen. Bei den Vertauschungen ändert sich jeweils das Vorzeichen auf beiden Seiten. Wir dürfen also ohne Einschränkung annehmen, dass

$$A = \begin{pmatrix} a_1^1 & \dots & a_{n-1}^1 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ a_1^{n-1} & \dots & a_{n-1}^{n-1} & 0 \\ 0 & \dots & 0 & 1 \end{pmatrix}$$

gilt. Nun ist

$$\begin{aligned}
\det A &= \sum_{\sigma\in S_n} \operatorname{sign}\sigma \cdot a_1^{\sigma(1)} \cdots a_{n-1}^{\sigma(n-1)} \cdot a_n^{\sigma(n)} \\
&= \sum_{\substack{\sigma\in S_n \\ \sigma(n)=n}} \operatorname{sign}\sigma \cdot a_1^{\sigma(1)} \cdots a_{n-1}^{\sigma(n-1)} \cdot a_n^{\sigma(n)} \\
&= \sum_{\sigma\in S_{n-1}} (-1)^{n+n} \operatorname{sign}\sigma \cdot a_1^{\sigma(1)} \cdots a_{n-1}^{\sigma(n-1)} \cdot 1 \\
&= (-1)^{n+n} \det A_n^n \\
&= \sum_{i=1}^{n} (-1)^{i+n} a_n^i \det A_n^i.
\end{aligned}$$

Die Behauptung folgt. □

Die Determinante ist ein Gruppenhomomorphismus:

Theorem 6.3.12 (Determinantenmultiplikationssatz)**.** *Seien $A, B \in F^{n\times n}$. Dann gilt*

$$\det(AB) = \det A \cdot \det B.$$

Beweis. Seien $A = \left(a_j^i\right)$, $B = \left(b_j^i\right)$. Dann ist $AB = \left(\sum_{k=1}^{n} a_k^i b_j^k\right)$. Setzen wir $b^i := \left(b_j^i\right)$, den Vektor, der in der i-ten Zeile von B steht, so können wir

$$AB = \begin{pmatrix} a_1^1 & \dots & a_n^1 \\ \vdots & \ddots & \vdots \\ a_1^n & \dots & a_n^n \end{pmatrix} \begin{pmatrix} b^1 \\ \vdots \\ b^n \end{pmatrix} = \begin{pmatrix} \sum_{i_1=1}^{n} a_{i_1}^1 b^{i_1} \\ \vdots \\ \sum_{i_n=1}^{n} a_{i_n}^n b^{i_n} \end{pmatrix}$$

schreiben, wenn wir eine Kollektion aus Zeilenvektoren als Matrix interpretieren. Wir erhalten aufgrund der Multilinearität und der Symmetrieeigenschaften der Determinante

$$\begin{aligned}
\det(AB) &= \det \begin{pmatrix} \sum_{i_1=1}^{n} a_{i_1}^1 b^{i_1} \\ \vdots \\ \sum_{i_n=1}^{n} a_{i_n}^n b^{i_n} \end{pmatrix} \\
&= \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \dots \sum_{i_n=1}^{n} a_{i_1}^1 a_{i_2}^2 \cdots a_{i_n}^n \det \begin{pmatrix} b^{i_1} \\ \vdots \\ b^{i_n} \end{pmatrix} \\
&= \sum_{\sigma\in S_n} a_{\sigma(1)}^1 a_{\sigma(2)}^2 \cdots a_{\sigma(n)}^n \cdot \det \begin{pmatrix} b^{\sigma(1)} \\ \vdots \\ b^{\sigma(n)} \end{pmatrix} \\
&= \sum_{\sigma\in S_n} a_{\sigma(1)}^1 a_{\sigma(2)}^2 \cdots a_{\sigma(n)}^n \cdot \operatorname{sign}\sigma \cdot \det \begin{pmatrix} b^1 \\ \vdots \\ b^n \end{pmatrix} \\
&= \det A \cdot \det B.
\end{aligned}$$

□

Theorem 6.3.13. *Eine quadratische Matrix $A \in F^{n\times n}$ ist genau dann regulär, wenn* $\det A \neq 0$ *gilt.*

Beweis.

(i) Sei A regulär. Dann existiert nach Theorem 4.8.5 eine Inverse A^{-1}. Der Determinantenmultiplikationssatz liefert

$$1 = \det \mathbb{1} = \det\left(AA^{-1}\right) = \det A \cdot \det A^{-1}.$$

Somit ist $\det A \neq 0$.

(ii) Sei nun $\det A \neq 0$. Nach Theorem 6.3.4 sind daher die Spalten von A linear unabhängig. Somit ist A regulär. □

Der Beweis liefert auch

Korollar 6.3.14. *Sei A regulär. Dann gilt*

$$\det A^{-1} = \frac{1}{\det A}.$$

Korollar 6.3.15. *Ähnliche Matrizen haben die gleiche Determinante.*

Beweis. Sei $B = DAD^{-1}$. Dann folgt

$$\det B = \det D \cdot \det A \cdot \det D^{-1} = \det D \cdot (\det D)^{-1} \cdot \det A = \det A. \qquad \square$$

Da ähnliche Matrizen dieselbe Determinante haben, ist die Determinante eines Endomorphismusses zwischen endlichdimensionalen Vektorräumen wohldefiniert.

Definition 6.3.16. Sei $f\colon V \to V$ ein Endomorphismus. Sei A die Matrix zu f bezüglich einer beliebigen Basis. Setze

$$\det f := \det A.$$

Theorem 6.3.17. *Sei $A = \left(a^i_j\right)$ eine $(n\times n)$-Matrix. Sei A^i_j die $((n-1)\times(n-1))$-Matrix, die wir erhalten, wenn wir die i-te Zeile und die j-te Spalte wie in Theorem 6.3.11 streichen. Definiere die $(n\times n)$-Matrix B durch $b^i_j := (-1)^{i+j}\det A^j_i$. Dann gilt*

$$AB = \det A \cdot \mathbb{1}.$$

B heißt die Adjunkte von A, A^{adj}.

Ist A regulär, so gilt

$$A^{-1} = \frac{B}{\det A}.$$

Beweis. Es gilt

$$\sum_{j=1}^{n} a^i_j b^j_k = \sum_{j=1}^{n} a^i_j (-1)^{j+k} \det A^k_j = \delta^i_k \det A,$$

denn für $i = k$ steht in der Mitte gerade die Entwicklung (siehe Theorem 6.3.11, Entwicklungssatz von Laplace) von $\det A$ nach der j-ten Spalte und für $i \neq k$ die Entwicklung der Determinante einer Matrix mit gleicher i-ter und k-ter Zeile. □

6.4. **Charakteristisches Polynom.**

Theorem 6.4.1. *Sei $f\colon V \to V$ durch die Matrix A beschrieben. Dann ist $\lambda \in F$ genau dann ein Eigenwert von f, wenn*

$$\det(A - \lambda\mathbb{1}) = 0$$

gilt.

Beweis. Ist λ ein Eigenwert, so ist $A - \lambda\mathbb{1}$ singulär und somit $\det(A - \lambda\mathbb{1}) = 0$. Die Rückrichtung funktioniert genauso. □

Die Determinantenabbildung ist ein Polynom in den Einträgen der Matrix. Daher können wir Determinanten nicht nur für Matrizen mit Einträgen in F, sondern auch mit Einträgen $a^i_j \in F[X]$ oder $F[\lambda]$ nach der Leibnizformel definieren.

Lemma 6.4.2. *Sei $A = \left(a^i_j\right)$ eine $(n \times n)$-Matrix. Dann ist*

$$\chi_A(\lambda) := \det(A - \lambda \mathbb{1})$$

ein Polynom vom Grad n. Wir schreiben

$$\chi_A(\lambda) = (-1)^n \lambda^n + (-1)^{n-1} S_1 \lambda^{n-1} + \ldots + (-1)^0 S_n,$$

wobei hier zunächst $S_i \in F$ gilt.

Beweis. Da in jeder Zeile (oder Spalte) nur ein Term mit λ, nämlich $a^i_i - \lambda \delta^i_i$ vorkommt, folgt per Induktion (Leibnizformel), dass $\chi_A(\lambda)$ ein Polynom vom Grad $\leq n$ ist. Durch Entwicklung der Determinante sieht man, dass $\chi_A(\lambda)$ die Summe von

$$\left(a^1_1 - \lambda\right)\left(a^2_2 - \lambda\right) \cdots \left(a^n_n - \lambda\right)$$

und einem Polynom vom Grade $\leq n - 2$ ist. Somit ist $(-1)^n$ der Koeffizient vor λ^n. □

Definition 6.4.3 (Elementarsymmetrische Funktionen)**.** Seien $\mu_1, \ldots, \mu_n \in F$. Definiere $\sigma_k \equiv \sigma_k(\mu) \equiv \sigma_k(\mu_1, \ldots, \mu_n)$ durch

$$\begin{aligned}
\sigma_0 &:= 1, \\
\sigma_1 &:= \sum_{1 \leq i \leq n} \mu_i, \\
\sigma_2 &:= \sum_{1 \leq i_1 < i_2 \leq n} \mu_{i_1} \mu_{i_2}, \\
\sigma_k &:= \sum_{1 \leq i_1 < i_2 < \ldots < i_k \leq n} \mu_{i_1} \mu_{i_2} \cdots \mu_{i_k}, \quad 1 \leq k \leq n, \\
\sigma_n &:= \mu_1 \mu_2 \cdots \mu_n.
\end{aligned}$$

Sei $A = \left(a^i_j\right) \in F^{n \times n}$. Wir definieren

$$\begin{aligned}
S_0(A) &:= 1, \\
S_1(A) &:= \sum_{i=1}^{n} a^i_i =: \operatorname{tr} A, \qquad \text{(Spur von A)} \\
S_k(A) &:= \sum_{1 \leq i_1 < i_2 < \ldots < i_k \leq n} \det\left(a^i_j\right)_{i,j \in \{i_1, \ldots, i_k\}}, \\
S_n(A) &:= \det A.
\end{aligned}$$

Häufiger setzt man alle anderen Funktionen gleich Null: $0 = S_{-1} = S_{-2} = \ldots = S_{n+1} = S_{n+2} = \ldots = \sigma_{-1} = \sigma_{-2} = \ldots = \sigma_{n+1} = \sigma_{n+2} = \ldots$.

Bemerkung 6.4.4. Sei $\sigma \in S_n$, so gilt

$$\sigma_k(\mu_1, \ldots, \mu_n) = \sigma_k(\sigma(\mu_1), \ldots, \sigma(\mu_n)).$$

Ist $A \in F^{n \times n}$ diagonal, so ist

$$S_k(A) = \sigma_k\left(a^1_1, \ldots, a^n_n\right).$$

Lemma 6.4.5. *Sei $A \in F^{n \times n}$. Dann gilt*

$$\det(A - \lambda \mathbb{1}) = (-1)^n \lambda^n + (-1)^{n-1} S_1(A) \lambda^{n-1} + \ldots + (-1)^{n-k} S_k(A) \lambda^{n-k} + \ldots + S_n(A).$$

Ist A diagonal, so gilt

$$\begin{aligned}\det(A-\lambda\mathbb{1}) = &(-1)^n\lambda^n + (-1)^{n-1}\sigma_1\left(a_1^1,\ldots,a_n^n\right)\lambda^{n-1} + \ldots \\ &+ (-1)^{n-k}\sigma_k\left(a_1^1,\ldots,a_n^n\right)\lambda^{n-k} + \ldots + \sigma_n\left(a_1^1,\ldots,a_n^n\right).\end{aligned}$$

Beweis. Den allgemeinen Fall lassen wir als Übungsaufgabe.

Ist A diagonal und $n=1$, so ist $\det(A-\lambda\mathbb{1}) = a_1^1 - \lambda$ wie behauptet. Gelte die Behauptung bereits für n. Sei innerhalb dieses Beweises σ_k die n-dimensionale k-te elementarsymmetrische Funktion von $\left(a_1^1,\ldots,a_n^n\right)$ und $\tilde{\sigma}_k$ die $(n+1)$-dimensionale k-te elementarsymmetrische Funktion von $\left(a_1^1,\ldots,a_{n+1}^{n+1}\right)$. Es gilt für $1 \le k \le n$

$$\begin{aligned}\tilde{\sigma}_k &= \sum_{1\le i_1<\ldots<i_k\le n+1} a_{i_1}^{i_1}\cdots a_{i_k}^{i_k} \\ &= \sum_{1\le i_1<\ldots<i_{k-1}<n+1} a_{i_1}^{i_1}\cdots a_{i_{k-1}}^{i_{k-1}}\cdot a_{n+1}^{n+1} + \sum_{1\le i_1<\ldots<i_k<n+1} a_{i_1}^{i_1}\cdots a_{i_k}^{i_k} \\ &= \sigma_{k-1}\cdot a_{n+1}^{n+1} + \sigma_k.\end{aligned}$$

Dann folgt für die Diagonalmatrix $A \in F^{(n+1)\times(n+1)}$ mit den Diagonaleinträgen $a_1^1,\ldots,a_n^n,a_{n+1}^{n+1}$

$$\begin{aligned}\det(A-\lambda\mathbb{1}) &= \left(a_1^1-\lambda\right)\cdot\left(a_2^2-\lambda\right)\cdots\left(a_{n+1}^{n+1}-\lambda\right) \\ &= \left((-1)^n\lambda^n + (-1)^{n-1}\sigma_1\lambda^{n-1} + \ldots + (-1)^{n-k}\sigma_k\lambda^{n-k} + \ldots + \sigma_n\right)\cdot \\ &\quad\cdot\left(a_{n+1}^{n+1}-\lambda\right) \\ &= \left(\sum_{k=0}^{n}(-1)^{n-k}\sigma_k\lambda^{n-k}\right)\cdot\left(a_{n+1}^{n+1}-\lambda\right)\end{aligned}$$

und nach Indexverschiebung in der ersten Summe

$$\begin{aligned}&= \sum_{k=1}^{n+1}(-1)^{n+1-k}\sigma_{k-1}a_{n+1}^{n+1}\lambda^{n+1-k} + \sum_{k=0}^{n}(-1)^{n+1-k}\sigma_k\lambda^{n+1-k} \\ &= \underbrace{\sigma_n a_{n+1}^{n+1}}_{=\tilde{\sigma}_{n+1}} + \sum_{k=1}^{n}(-1)^{n+1-k}\underbrace{\left(\sigma_{k-1}a_{n+1}^{n+1}+\sigma_k\right)}_{=\tilde{\sigma}_k}\lambda^{n+1-k} + (-1)^{n+1}\sigma_0\lambda^{n+1} \\ &= \sum_{k=0}^{n+1}(-1)^{n+1-k}\tilde{\sigma}_k\lambda^{n+1-k}.\end{aligned}$$

Die Behauptung für Diagonalmatrizen folgt. □

Definition 6.4.6. Das Polynom

$$\chi_A(X) := \det(A - X\mathbb{1})$$

heißt **charakteristisches Polynom** der Matrix A. Sei f ein durch A dargestellter Endomorphismus. Setze

$$\chi_f(X) := \det(A - X\mathbb{1}).$$

Lemma 6.4.7. *Sei die Abbildung f ein Endomorphismus eines endlichdimensionalen Vektorraumes. Dann ist $\chi_f(X)$ wohldefiniert.*

Beweis. Sei S regulär und $B = SAS^{-1}$. Dann gilt

$$\begin{aligned}\chi_B(X) &= \det\left(SAS^{-1} - X\mathbb{1}\right) = \det\left(SAS^{-1} - SX\mathbb{1}S^{-1}\right) \\ &= \det\left(S(A-X\mathbb{1})S^{-1}\right) = \det S\cdot\det(A-X\mathbb{1})\cdot\det S^{-1} \\ &= \det(A-X\mathbb{1}) = \chi_A(X).\end{aligned}$$

□

Zusammenfassend erhalten wir

Theorem 6.4.8. *Die Eigenwerte eines linearen Endomorphismusses $f\colon V \to V$ zwischen endlichdimensionalen Vektorräumen sind genau die Nullstellen von $\chi_f(X)$ in F.*

Als Korollar zur Wohldefiniertheit von $\chi_f(X)$ erhalten wir

Korollar 6.4.9. *Seien A, A' ähnlich, so gilt $S_k(A) = S_k(A')$ für alle $1 \leq k \leq n$.*

Alternativ könnte man dieses Resultat für $k = 1$ mit dem folgenden Lemma beweisen.

Lemma 6.4.10. *Seien $A, B \in F^{n\times n}$. Dann gilt*

$$\operatorname{tr}(AB) = \operatorname{tr}(BA).$$

Beweis. Seien $A = \left(a^i_j\right)_{1\leq i,j\leq n}$ und $B = \left(b^i_j\right)_{1\leq i,j\leq n}$ Es gilt

$$\operatorname{tr}(AB) = \sum_{i,j=1}^{n} a^i_j b^j_i = \sum_{i,j=1}^{n} b^j_i a^i_j = \operatorname{tr}(BA). \qquad \square$$

Zur geometrischen Vielfachheit eines Eigenwertes gibt es auch ein algebraisches Analogon

Definition 6.4.11. Sei $\lambda \in F$ ein Eigenwert einer linearen Selbstabbildung $f\colon V \to V$ (zwischen endlichdimensionalen Räumen; in Zukunft wollen wir diesen Zusatz immer als automatisch gegeben ansehen, wenn Determinanten oder charakteristische Polynome auftreten). Dann heißt die Vielfachheit der Nullstelle λ von $\chi_f(X)$ die **algebraische Vielfachheit** von λ.

Theorem 6.4.12. *Sei λ ein Eigenwert von $f\colon V \to V$. Dann ist die geometrische Vielfachheit von λ kleiner oder gleich der algebraischen Vielfachheit von λ.*

Im Beispiel $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ ist die geometrische Vielfachheit des Eigenwertes 1 gleich eins, die algebraische ist 2.

Beweis. Sei $\dim E_\lambda = k$. Dann gibt es k linear unabhängige Eigenvektoren von f zum Eigenwert λ. Wir ergänzen diese zu einer Basis von V. Dann ist f bezüglich dieser Basis durch eine Matrix A der Form

$$\begin{pmatrix} \lambda & 0 & \dots & 0 & * & \dots & * \\ 0 & \lambda & \dots & 0 & * & \dots & * \\ \vdots & \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \lambda & * & \dots & * \\ 0 & 0 & \dots & 0 & b^1_1 & \dots & b^1_{n-k} \\ \vdots & \vdots & & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & b^{n-k}_1 & \dots & b^{n-k}_{n-k} \end{pmatrix}$$

mit geeigneten $b^i_j \in F$ dargestellt. Dabei bezeichnet $*$ einen beliebigen Eintrag aus F. Wir setzen $B := \left(b^i_j\right)_{1\leq i,j\leq n-k}$. Wir erhalten durch sukzessive Entwicklung nach der ersten, zweiten, …, k-ten Spalte $\chi_f(X) = \chi_A(X) = \det(A - X\mathbb{1}_n) = (\lambda - X)^k \det(B - X\mathbb{1}_{n-k})$. Somit ist die algebraische Vielfachheit von λ mindestens gleich k. Die Behauptung folgt. $\square$

6.5. **Diagonalisierbarkeit und Trigonalisierbarkeit.**

Definition 6.5.1. Eine lineare Selbstabbildung $f\colon V \to V$ eines endlichdimensionalen Vektorraumes V heißt **diagonalisierbar,** wenn es eine Basis gibt, so dass f bezüglich dieser Basis durch eine Diagonalmatrix dargestellt wird.

Lemma 6.5.2. *Sei $f\colon W \to W$ linear und sei $\dim W < \infty$. Sei U der Eigenraum zum Eigenwert λ und V der Eigenraum zum Eigenwert $\mu \neq \lambda$. Dann gilt*

$$\dim(U + V) = \dim U + \dim V.$$

(Die Verallgemeinerung auf endlich viele Eigenräume beweist man mit einem analogen Beweis.)

Beweis. Sei a_i, $1 \leq i \leq \dim U$, eine Basis von U und b_j, $1 \leq j \leq \dim V$ eine Basis von V. Wir behaupten, dass die Vektoren $\{(a_i)_{1\leq i\leq \dim U}, (b_j)_{1\leq j\leq \dim V}\}$ linear unabhängig sind. Sonst erhalten wir eine nichttriviale Linearkombination der Null

$$\underbrace{\sum_{i=1}^{\dim U} \lambda^i a_i}_{=:a} + \underbrace{\sum_{j=1}^{\dim V} \mu^j b_j}_{=:b} = 0,$$

wobei a ein Eigenvektor zum Eigenwert λ und b ein Eigenvektor zum Eigenwert μ ist. Nach Theorem 5.7.3 folgt nun $a = b = 0$. □

Theorem 6.5.3. *Eine lineare Abbildung $f\colon V \to V$ ist genau dann diagonalisierbar, wenn das charakteristische Polynom $\chi_f(X)$ in Linearfaktoren zerfällt, d. h. wenn es sich in der Form $a \cdot (X - a_1) \cdots (X - a_n)$ für $a, a_i \in F$ schreiben lässt, und die geometrische Vielfachheit jedes Eigenwertes von f mit seiner algebraischen Vielfachheit übereinstimmt.*

Beweis.

(i) Ist f diagonalisierbar, so ist die Behauptung klar.

(ii) Da $\chi_f(X)$ in Linearfaktoren zerfällt, summieren sich die geometrischen wie algebraischen Vielfachheiten der (einfach aufgeführten) Eigenwerte von f zu n auf. Wähle zu jedem Eigenraum eine Basis aus Eigenvektoren. Nach Lemma 6.5.2 bilden diese eine Basis eines n-dimensionalen Unterraumes, also von V selbst. Also besitzt V eine Basis aus Eigenvektoren von f. f ist somit diagonalisierbar. □

Definition 6.5.4. Eine lineare Selbstabbildung $f\colon V \to V$ eines endlichdimensionalen Raumes V heißt genau dann **trigonalisierbar,** wenn es eine Basis gibt, so dass f durch eine obere Dreiecksmatrix dargestellt wird, d. h. durch eine Matrix $A = \left(a^i_j\right)$ mit $a^i_j = 0$ für $j < i$.

Theorem 6.5.5. *Ein linearer Endomorphismus $f\colon V \to V$ ist genau dann trigonalisierbar, wenn das charakteristische Polynom $\chi_f(X)$ in Linearfaktoren zerfällt.*

Beweis. Ist f trigonalisierbar, so ist die Behauptung klar.

Zerfalle das charakteristische Polynom von f in Linearfaktoren. Wir wollen per Induktion zeigen, dass f trigonalisierbar ist. Sei λ_1 eine Nullstelle von $\chi(f)$ und sei f durch die Matrix A dargestellt. Dann ist $A - \lambda_1 \mathbb{1}$ singulär, besitzt also einen nichttrivialen Kern. Somit gibt es ein $a_1 \in \ker(A - \lambda_1 \mathbb{1})$, also einen Eigenwert von A zum Eigenwert λ_1. Bezüglich einer Basis $(a_1, \ldots, a_n)$, a_i für $i \geq 2$ noch nicht

fixiert, in der der erste Vektor a_1 ist, hat A die Gestalt

$$\begin{pmatrix} \lambda_1 & a_2^1 & \dots & a_n^1 \\ 0 & a_2^2 & \dots & a_n^2 \\ 0 & a_2^3 & \dots & a_n^3 \\ \vdots & \vdots & & \vdots \\ 0 & a_2^n & \dots & a_n^n \end{pmatrix}.$$

Betrachte nun $g\colon \langle a_2, \dots, a_n\rangle \to \langle a_2, \dots, a_n\rangle$, definiert durch

$$\begin{pmatrix} a_2^2 & \dots & a_n^2 \\ a_2^3 & \dots & a_n^3 \\ \vdots & & \vdots \\ a_2^n & \dots & a_n^n \end{pmatrix}.$$

Es gilt $\chi_f(X) = (\lambda_1 - X)\cdot\chi_g(X)$. Da χ_f in Linearfaktoren zerfällt, gilt dies auch für $\chi_g(X)$. Nach Induktionsvoraussetzung gibt es daher eine Basis $b_2, \dots, b_n$, bezüglich der g durch eine obere Dreiecksmatrix dargestellt ist. Somit ist f bezüglich der Basis $(a_1, b_2, \dots, b_n)$ durch eine obere Dreiecksmatrix dargestellt. □

7. Vektorräume mit Skalarprodukt

Sobald wir Skalarprodukte verwenden, wollen wir stets annehmen, dass wir $\mathbb{R}$- oder $\mathbb{C}$-Vektorräume betrachten.

7.1. Euklidische Vektorräume.

Ein euklidischer Vektorraum ist ein Vektorraum mit einem reellen Skalarprodukt.

Definition 7.1.1. Sei V ein $\mathbb{R}$-Vektorraum. Ein (reelles) **Skalarprodukt** auf V ist eine Funktion $\langle\cdot,\cdot\rangle\colon V \times V \to \mathbb{R}$, die folgendes erfüllt:

(i) $\langle a, b\rangle = \langle b, a\rangle$ für alle $a, b \in V$ (Symmetrie)
(ii) $\langle \lambda a + \mu b, c\rangle = \lambda\langle a, c\rangle + \mu\langle b, c\rangle$ für alle $a, b, c \in V$ und $\lambda, \mu \in \mathbb{R}$ (Linearität)
(iii) $\langle a, a\rangle \geq 0$ und $\langle a, a\rangle = 0 \iff a = 0$ für alle $a \in V$ (positive Definitheit)

Bemerkung 7.1.2.

(i) Aus des Symmetrie und der Linearität im ersten Argument folgt auch die Linearität im zweiten Argument

$$\langle c, \lambda a + \mu b\rangle = \langle \lambda a + \mu b, c\rangle = \lambda\langle a, c\rangle + \mu\langle b, c\rangle = \lambda\langle c, a\rangle + \mu\langle c, b\rangle.$$

Eine Funktion in zwei Argumenten, die in beiden Argumenten linear ist, heißt bilinear.

(ii) Per Induktion zeigt man für beliebige Linearkombinationen

$$\left\langle \sum_{i=1}^{m} \lambda^i a_i, \sum_{j=1}^{n} \mu^j b_j \right\rangle = \sum_{i=1}^{m}\sum_{j=1}^{n} \lambda^i\mu^j\langle a_i, b_j\rangle.$$

Somit ist ein Skalarprodukt durch seine Werte auf einer Basis eindeutig bestimmt.

(iii) Sei $V = \mathbb{R}^n$. Für Vektoren $\xi = (\xi^1, \dots, \xi^n)$ und $\eta = (\eta^1, \dots, \eta^n)$ definieren wir

$$\langle \xi, \eta\rangle := \sum_{i=1}^{n} \xi^i\eta^i.$$

Dies ist ein Skalarprodukt auf $\mathbb{R}^n$ (Übung), das Standard-Skalarprodukt.

(iv) Sei $V = \mathbb{R}^2$. Definiere für $\xi = (\xi^1, \xi^2)$ und $\eta = (\eta^1, \eta^2)$

$$\langle \xi, \eta \rangle := \xi^1\eta^1 + 5\xi^1\eta^2 + 5\xi^2\eta^2 + 26\xi^2\eta^2.$$

Dies ist ein weiteres Skalarprodukt auf $\mathbb{R}^2$. Es stimmt nicht mit dem Standardskalarprodukt auf $\mathbb{R}^2$ überein.

(v) Sei V der Vektorraum der auf $[a, b] \subset \mathbb{R}$ stetigen reellwertigen Funktionen, $V = C^0([a, b])$. Gelte bei diesen Beispielen stets $a < b$. Definiere für $f, g \in V$

$$\langle f, g \rangle = \int_a^b f(x)g(x)\,dx.$$

Dies ist ein Skalarprodukt auf V (vgl. Analysis-Vorlesung).

Definition 7.1.3. Sei $\langle \cdot, \cdot \rangle \colon V \times V \to \mathbb{R}$ ein Skalarprodukt und sei $a_1, \ldots, a_n$ eine Basis von V. Dann definieren wir

$$c_{ij} := \langle a_i, a_j \rangle \quad \text{für } 1 \leq i, j \leq n.$$

$C = (c_{ij})_{1 \leq i,j \leq n}$ heißt Matrix zum Skalarprodukt $\langle \cdot, \cdot \rangle$.

Beachte, dass wir hier beide Indices unten schreiben. In Matrizenform stellen wir C durch

$$\begin{pmatrix} c_{11} & c_{12} & \ldots & a_{1n} \\ c_{21} & c_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n1} & c_{n2} & \ldots & a_{nn} \end{pmatrix}$$

dar. Es ist nicht ideal, Matrizen der Form $\left(c^i_j\right)$ und (c_{ij}) graphisch gleich darzustellen, jedoch üblich.

Theorem 7.1.4. *Die Matrix $C = (c_{ij})_{1 \leq i,j \leq n}$ eines reellen Skalarproduktes ist symmetrisch, d. h. es gilt $c_{ij} = c_{ji}$ für alle $1 \leq i, j \leq n$ oder $C = C^T$ mit $C^T := (b_{ij})_{1 \leq i,j \leq n}$ mit $b_{ij} = c_{ji}$.*

Beweis. Sei $a_1, \ldots, a_n$ eine Basis. Dann gilt

$$c_{ij} = \langle a_i, a_j \rangle = \langle a_j, a_i \rangle = c_{ji}. \qquad \square$$

Beispiele 7.1.5.

(i) Bezüglich der Standardbasis ist die Matrix des Standardskalarproduktes auf $\mathbb{R}^n$ gleich $\mathbb{1} = I = (\delta_{ij})_{1 \leq i,j \leq n}$.

(ii) Das Skalarprodukt mit

$$\langle \xi, \eta \rangle := \xi^1\eta^1 + 5\xi^1\eta^2 + 5\xi^2\eta^1 + 26\xi^2\eta^2$$

ist bezüglich der Standardbasis des $\mathbb{R}^2$ durch die Matrix

$$\begin{pmatrix} 1 & 5 \\ 5 & 26 \end{pmatrix}$$

und bezüglich der Basis aus den Vektoren $(1, 0)$ und $(-5, 1)$ durch die Matrix $\mathbb{1}$ dargestellt.

(iii) Sei V der Vektorraum der Polynome vom Grad ≤ 3 mit Basis $(1, x, x^2, x^3)$. Dann ist das Skalarprodukt

$$\langle p, q \rangle = \int_0^1 p(x)q(x)\,dx$$

durch die Matrix

$$\begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \frac{1}{6} \\ \frac{1}{4} & \frac{1}{5} & \frac{1}{6} & \frac{1}{7} \end{pmatrix}$$

dargestellt.

Lemma 7.1.6. *Sei $\langle \cdot, \cdot \rangle$ ein Skalarprodukt, dem bezüglich der Basis $a_1, \ldots, a_n$ die Matrix $(c_{ij})_{1 \le i,j \le n}$ zugeordnet ist. Seien $\xi = \sum_{i=1}^{n} \xi^i a_i$ und $\eta = \sum_{j=1}^{n} \eta^i a_i$ beliebig. Dann gilt*

$$\langle \xi, \eta \rangle = \sum_{i,j=1}^{n} \xi^i c_{ij} \eta^j.$$

Beweis. Benutze die Linearität in beiden Argumenten wie in Bemerkung 7.1.2 (ii). □

Bemerkung 7.1.7. Wir wollen schließlich noch das Verhalten einer ein Skalarprodukt darstellenden Matrix unter Basistransformationen untersuchen.

Sei V ein reeller Vektorraum mit Basen S und T, $S = (a_1, \ldots, a_n)$ und $T = (b_1, \ldots, b_n)$. Gelte

$$b_k = \sum_{i=1}^{n} d_k^i a_i.$$

Das Skalarprodukt $\langle \cdot, \cdot \rangle \colon V \times V \to \mathbb{R}$ werde bezüglich S durch die Matrix $C = (c_{ij})_{1 \le i,j \le n}$ beschrieben. Dann gilt

$$\begin{aligned} m_{kl} := \langle b_k, b_l \rangle &= \left\langle \sum_{i=1}^{n} d_k^i a_i, \sum_{j=1}^{n} d_l^j a_j \right\rangle \\ &= \sum_{i,j=1}^{n} d_k^i d_l^j \langle a_i, a_j \rangle = \sum_{i,j=1}^{n} d_k^i d_l^j c_{ij}. \end{aligned}$$

Setze $M := (m_{ij})_{1 \le i,j \le n}$ und $D := (d_k^i)_{1 \le i,k \le n}$. Dann gilt

$$M = D^T C D.$$

7.2. **Unitäre Vektorräume.** Mit $\bar{z}$ bezeichnen wir die komplex konjugierte Zahl zu z. Ist also $z = x + iy$ mit $x, y \in \mathbb{R}$, so ist $\bar{z} = x - iy$.

Ein Vektorraum mit einem unitären Skalarprodukt heißt unitärer Vektorraum.

Definition 7.2.1. Sei V ein $\mathbb{C}$-Vektorraum. Ein **unitäres Skalarprodukt** auf V ist eine Funktion $\langle \cdot, \cdot \rangle \colon V \times V \to \mathbb{C}$, die die folgenden Eigenschaften erfüllt:

(i) $\langle a, b \rangle = \overline{\langle b, a \rangle}$ für alle $a, b \in V$ (hermitesch)

(ii) $\langle \lambda a + \mu b, c \rangle = \lambda \langle a, c \rangle + \mu \langle b, c \rangle$ für alle $a, b, c \in V$ und alle $\lambda, \mu \in \mathbb{C}$ (Linearität im ersten Argument)

(iii) $\langle a, a \rangle \ge 0$ und $\langle a, a \rangle = 0 \iff a = 0$ für alle $a \in V$ (positiv definit)

Bemerkung 7.2.2. Folgende Eigenschaften und Beispiele sind analog zum reellen Fall

(i) Seien $a, b, c \in V$ und $\lambda, \mu \in \mathbb{C}$. Dann gilt

$$\langle c, \lambda a + \mu b \rangle = \overline{\langle \lambda a + \mu b, c \rangle} = \bar{\lambda} \overline{\langle a, c \rangle} + \bar{\mu} \overline{\langle b, c \rangle} = \bar{\lambda} \langle c, a \rangle + \bar{\mu} \langle c, b \rangle.$$

Linearität im ersten Argument und dieses Verhalten im zweiten Argument bezeichnet man als Sesquilinearität.

Es gibt auch die umgekehrte Konvention, d. h. man definiert ein unitäres Skalarprodukt so, dass es im zweiten Argument statt im ersten Argument linear ist und dass die übrigen Eigenschaften unverändert gelten. Im ersten Argument werden dann Skalare komplex konjugiert nach außen gezogen.

(ii) Per Induktion folgt hieraus für beliebige Linearkombinationen

$$\left\langle \sum_{i=1}^{m} \lambda^i a_i, \sum_{j=1}^{n} \mu^j b_j \right\rangle = \sum_{i=1}^{m} \sum_{j=1}^{n} \lambda^i \overline{\mu^j} \langle a_i, b_j \rangle.$$

Daher ist ein unitäres Skalarprodukt durch seine Werte auf einer Basis bereits eindeutig bestimmt.

(iii) Ist $V = \mathbb{C}^n$ und seien $x = (x^1, \ldots, x^n)$ und $y = (y^1, \ldots, y^n)$ Vektoren in V, so ist durch

$$\langle x, y \rangle := \sum_{i=1}^{n} x^i \overline{y^i}$$

ein unitäres Skalarprodukt auf $\mathbb{C}^n$ definiert.

(iv) Sei V der komplexe Vektorraum der auf $[a, b]$ komplexwertigen stetigen Funktionen einer reellen Variablen. Dann definiert

$$\langle f, g \rangle := \int_a^b f(t) \overline{g(t)}\, dt$$

ein unitäres Skalarprodukt auf V. Vergleiche wieder eine Analysis-Vorlesung für die positive Definitheit.

(v) Achtung! Manchmal fordert man auch die Linearität im zweiten statt im ersten Argument.

Definition 7.2.3. Sei V ein komplexer Vektorraum und $\langle \cdot, \cdot \rangle \colon V \times V \to \mathbb{C}$ ein unitäres Skalarprodukt. Sei $(a_1, \ldots, a_n)$ eine Basis von V. Dann heißt die Matrix $C = (c_{ij})_{1 \leq i,j \leq n}$ mit

$$c_{ij} = \langle a_i, a_j \rangle, \quad 1 \leq i, j, \leq n$$

die Matrix des Skalarproduktes $\langle \cdot, \cdot \rangle$ bezüglich der Basis $a_1, \ldots, a_n$.

Theorem 7.2.4. *Die Matrix C eines unitären Skalarproduktes ist hermitesch, d. h. es gilt $C^T = \bar{C}$.*

Beweis. Es gilt

$$c_{ij} = \langle a_i, a_j \rangle = \overline{\langle a_j, a_i \rangle} = \overline{c_{ji}}.$$

Es folgt $C^T = \overline{C}$ wie behauptet. □

Wie im reellen Fall zeigt man:

Lemma 7.2.5. *Sei $\langle \cdot, \cdot \rangle$ ein unitäres Skalarprodukt, dem bezüglich einer Basis $a_1, \ldots, a_n$ die Matrix $(c_{ij})_{1 \leq i,j \leq n}$ zugeordnet ist. Seien $\xi = \sum_{i=1}^{n} \xi^i a_i$ und $\eta = \sum_{j=1}^{n} \eta^i a_i$ beliebig. Dann gilt*

$$\langle \xi, \eta \rangle = \sum_{i,j=1}^{n} \xi^i c_{ij} \bar{\eta^j}.$$

Bemerkung 7.2.6. Wir wollen wiederum das Verhalten einer ein Skalarprodukt darstellenden Matrix unter Basistransformationen untersuchen.

Sei V ein komplexer Vektorraum mit Basen S und T, $S = (a_1, \ldots, a_n)$ und $T = (b_1, \ldots, b_n)$. Gelte

$$b_k = \sum_{i=1}^{n} d_k^i a_i.$$

Das Skalarprodukt $\langle \cdot, \cdot \rangle \colon V \times V \to \mathbb{C}$ werde bezüglich S durch die Matrix $C = (c_{ij})_{1 \le i,j \le n}$ beschrieben. Dann gilt

$$m_{kl} := \langle b_k, b_l \rangle = \left\langle \sum_{i=1}^{n} d_k^i a_i, \sum_{j=1}^{n} d_l^j a_j \right\rangle$$
$$= \sum_{i,j=1}^{n} d_k^i \overline{d_l^j} \langle a_i, a_j \rangle = \sum_{i,j=1}^{n} d_k^i \overline{d_l^j} c_{ij}.$$

Setze $M := (m_{ij})_{1 \le i,j \le n}$ und $D := (d_k^i)_{1 \le i,k \le n}$. Dann gilt

$$M = D^T C \bar{D}.$$

Bemerkung 7.2.7. Reelles und komplexes Skalarprodukt verhalten sich sehr ähnlich. Daher werden wir häufiger nur den komplexen Sachverhalt untersuchen. Ein analoges reelles Resultat folgt dann analog.

7.3. **Norm.**

Theorem 7.3.1 (Cauchy-Schwarzsche Ungleichung)**.** *Sei V ein Vektorraum mit Skalarprodukt. Dann gilt*

$$|\langle a, b \rangle|^2 \le \langle a, a \rangle \cdot \langle b, b \rangle$$

für alle $a, b \in V$. Gleichheit gilt genau dann, wenn a und b linear abhängig sind.

Beweis. Der Fall $b = 0$ ist einfach (Übung).

Sei $\lambda \in \mathbb{C}$ beliebig. Dann gilt

$$0 \le \langle a - \lambda b, a - \lambda b \rangle = \langle a, a \rangle - \lambda \overline{\langle a, b \rangle} - \bar{\lambda} \langle a, b \rangle + \lambda \bar{\lambda} \langle b, b \rangle.$$

Setzen wir speziell $\lambda = \frac{\langle a,b \rangle}{\langle b,b \rangle}$, so folgt nach Multiplikation mit $\langle b, b \rangle$

$$0 \le \langle a, a \rangle \cdot \langle b, b \rangle - 2|\langle a, b \rangle|^2 + |\langle a, b \rangle|^2.$$

Die behauptete Ungleichung folgt.

Gilt Gleichheit, so gilt insbesondere auch in der ersten Ungleichung Gleichheit, also $0 = a - \lambda b$ aufgrund der positiven Definitheit. □

Korollar 7.3.2.

(i) Seien $a_1, \ldots, a_n$ und $b_1, \ldots, b_n$ reelle Zahlen. Dann gilt

$$\sum_{i=1}^{n} a_i b_i \le \left(\sum_{i=1}^{n} a_i^2 \right)^{1/2} \cdot \left(\sum_{j=1}^{n} b_j^2 \right)^{1/2}.$$

(ii) Seien $f, g \colon [a, b] \to \mathbb{C}$ stetig. Dann gilt

$$\left| \int_a^b f(t) g(t)\, dt \right|^2 \le \int_a^b |f(t)|^2\, dt \cdot \int_a^b |g(t)|^2\, dt.$$

Beweis. Dies folgt direkt aus der Cauchy-Schwarzschen Ungleichung für den Vektorraum

(i) $\mathbb{R}^n$ mit Standardskalarprodukt.
(ii) der stetigen Funktionen auf $[a, b]$ mit $\int fg$ als Skalarprodukt. □

Definition 7.3.3. Sei V ein reeller oder komplexer Vektorraum. Eine Funktion $\| \cdot \| \colon V \to \mathbb{R}$ heißt **Norm** auf V, wenn sie die folgenden Bedingungen erfüllt

(i) $\|\lambda a\| = |\lambda| \cdot \|a\|$ für alle $\lambda \in F$ und $a \in V$,
(ii) $\|a + b\| \le \|a\| + \|b\|$ für alle $a, b \in V$, (Dreiecksungleichung)
(iii) $\|a\| = 0 \iff a = 0$.

Theorem 7.3.4. *Sei V ein Vektorraum mit Skalarprodukt $\langle\cdot,\cdot\rangle$. Dann ist die Funktion $\|\cdot\|\colon V \to \mathbb{R}$, durch $\|a\| := \sqrt{\langle a,a\rangle}$ definiert, eine Norm.*

Beweis.

(i) $\|\lambda a\| = \sqrt{\langle \lambda a, \lambda a\rangle} = \sqrt{\lambda\bar{\lambda}\langle a,a\rangle} = |\lambda|\sqrt{\langle a,a\rangle} = |\lambda|\cdot\|a\|$.

(ii)
$$\begin{aligned}\|a+b\|^2 &= \langle a+b, a+b\rangle = \langle a,a\rangle + \langle a,b\rangle + \langle b,a\rangle + \langle b,b\rangle\\ &\le \langle a,a\rangle + 2|\langle a,b\rangle| + \langle b,b\rangle\\ &\le \|a\|^2 + 2\cdot\|a\|\cdot\|b\| + \|b\|^2 \qquad \text{(Cauchy-Schwarz)}\\ &= (\|a\|+\|b\|)^2.\end{aligned}$$

(iii) Die positive Definitheit der Norm folgt aus der positiven Definitheit des Skalarproduktes. □

In einem Skalarproduktraum (=Vektorraum mit Skalarprodukt) werden wir unter einer Norm immer $\|a\| := \sqrt{\langle a,a\rangle}$ verstehen.

7.4. **Orthonormalbasen.**

Definition 7.4.1. Sei V ein Skalarproduktraum.

(i) Dann heißt $a \in V$ **senkrecht** (oder orthogonal) zu $b \in V$, wenn $\langle a,b\rangle = 0$ gilt.

(ii) Eine Teilmenge $S \subset V$ heißt **orthogonal** (oder senkrecht) zu einer Teilmenge $T \subset V$, wenn $\langle s,t\rangle = 0$ für alle $(s,t) \in S \times T$ gilt.

Bemerkung 7.4.2.

(i) Ist a orthogonal zu b, so ist b orthogonal zu a.

(ii) Der Nullvektor ist zu allen anderen Vektoren orthogonal, da aus
$$\langle 0,a\rangle = \langle 0+0,a\rangle = \langle 0,a\rangle + \langle 0,a\rangle$$
$\langle 0,a\rangle = 0$ folgt. Dies ist auch der einzige Vektor mit dieser Eigenschaft, da aus $\langle a,a\rangle = 0$ bereits $a = 0$ folgt.

(iii) Zwei Teilmengen S und T sind genau dann orthogonal, wenn $\langle S\rangle$ und $\langle T\rangle$ orthogonal sind (Übung).

Definition 7.4.3.

(i) Eine (nichtleere) Familie/Teilmenge S von V heißt **orthogonal,** wenn je zwei verschiedene Elemente aus S orthogonal zueinander sind.

(ii) Eine orthogonale Familie S heißt **orthonormiert,** wenn $\langle a,a\rangle = 1$ für alle $a \in S$ gilt.

(iii) Eine orthonormierte Familie, die zugleich Basis von V ist, heißt **Orthonormalbasis** (ONB oder orthonormierte Basis) von V.

Bemerkung 7.4.4.

(i) Eine orthogonale Familie S von Vektoren mit $0 \notin S$ kann man zu einer orthonormalen Familie machen, indem man jeden Vektor $a \in S$ durch $\frac{a}{\|a\|}$ ersetzt. Die neue Familie ist dann durch Normieren aus der alten hervorgegangen (Übung).

(ii) Sei V der Vektorraum der auf $[-\pi,\pi]$ stetigen reellwertigen Funktionen mit Skalarprodukt
$$\langle f,g\rangle = \int_{-\pi}^{\pi} f(x)g(x)\,dx.$$
Dann ist
$$\{1, \sin x, \cos x, \sin 2x, \cos 2x, \sin 3x, \cos 3x, \ldots\}$$
eine orthogonale Familie (Übung). Dies spielt bei Fourierreihen eine Rolle.

Theorem 7.4.5. *Sei S eine orthogonale Familie mit $0 \notin S$. Dann ist S linear unabhängig.*

Beweis. Gelte

$$\sum_{i=1}^{n} \lambda^i a_i = 0$$

mit $\lambda^i \in F$ und $a_i \in S$ für $1 \leq i \leq n$ und ein $n \in \mathbb{N}^+$. Dann folgt für $1 \leq j \leq n$

$$0 = \langle 0, a_j \rangle = \left\langle \sum_{i=1}^{n} \lambda^i a_i, a_j \right\rangle = \sum_{i=1}^{n} \lambda^i \langle a_i, a_j \rangle = \lambda^j \langle a_j, a_j \rangle.$$

Wegen $a_j \neq 0$ folgt also $\lambda^j = 0$. Somit ist S linear unabhängig. □

Theorem 7.4.6. *Sei V ein endlichdimensionaler Skalarproduktraum. Sei $a_1, \ldots, a_n$ eine Orthonormalbasis. Dann gilt für jedes $b \in V$*

$$b = \sum_{i=1}^{n} \langle b, a_i \rangle a_i.$$

Beweis. Wir wissen bereits, dass sich b als Linearkombination der Form

$$b = \sum_{i=1}^{n} \lambda^i a_i$$

darstellen lässt. Hieraus folgt

$$\langle b, a_j \rangle = \left\langle \sum_{i=1}^{n} \lambda^i a_i, a_j \right\rangle = \sum_{i=1}^{n} \lambda^i \langle a_i, a_j \rangle = \lambda^j.$$

Wir erhalten die Behauptung. □

Der folgende Algorithmus erlaubt es, aus einer Basis eine Orthogonalbasis zu gewinnen.

Theorem 7.4.7 (Gram-Schmidtsches Orthogonalisierungsverfahren)**.** *Jeder endlichdimensionale Skalarproduktraum besitzt eine Orthonormalbasis.*

Beweis. Sei $a_1, \ldots, a_n$ eine beliebige Basis. Daraus konstruieren wir induktiv eine orthogonale Basis vermöge

$$b_{k+1} := a_{k+1} - \sum_{j=1}^{k} \frac{\overline{\langle b_j, a_{k+1} \rangle}}{\langle b_j, b_j \rangle} b_j.$$

Nach Definition handelt es sich weiterhin um eine Basis. Insbesondere gilt also $b_j \neq 0$ für alle j. Die Familie der b_j ist orthogonal, da wir für $i \in \{1, \ldots, k\}$ induktiv

$$\begin{aligned} \langle b_i, b_{k+1} \rangle &= \langle b_i, a_{k+1} \rangle - \sum_{j=1}^{k} \frac{\langle b_j, a_{k+1} \rangle}{\langle b_j, b_j \rangle} \langle b_i, b_j \rangle \\ &= \langle b_i, a_{k+1} \rangle - \langle b_i, a_{k+1} \rangle \\ &= 0 \end{aligned}$$

erhalten. Wir normieren nun die Vektoren b_i. Nach Theorem 7.4.5 sind sie linear unabhängig. Aufgrund ihrer Anzahl handelt es sich somit um eine Orthonormalbasis. □

Bemerkung 7.4.8. Um die Fälle $F = \mathbb{R}$ und $F = \mathbb{C}$ gleichzeitig behandeln zu können, setzen wir $A^* := \bar{A}^T$ für $A \in F^{n \times n}$. Im Reellen gilt $A^* = A^T$.

Theorem 7.4.9. *Sei V ein Skalarproduktraum, $\dim V < \infty$. Sei $\{b_1, \dots, b_n\}$ eine Orthonormalbasis. Definiere Vektoren d_k durch*

$$d_k = \sum_{i=1}^{n} a_k^i b_i.$$

Setze $A := \left(a_j^i\right)_{1 \le i,j \le n}$. Dann ist $\{d_1, \dots, d_n\}$ genau dann eine Orthonormalbasis von V, wenn

$$A^* A = I$$

gilt.

Beweis. Wie angekündigt zeigen wir nur den unitären Fall. Es gilt

$$\begin{aligned}\langle d_k, d_k \rangle &= \left\langle \sum_{i=1}^{n} a_k^i b_i, \sum_{j=1}^{n} a_l^j b_j \right\rangle = \sum_{i,j=1}^{n} a_k^i \overline{a_l^j} \langle b_i, b_j \rangle \\ &= \sum_{i,j=1}^{n} a_k^i \overline{a_l^j} \delta_{ij} = \sum_{i=1}^{n} a_k^i \overline{a_l^i}.\end{aligned}$$

Sei also $\{d_1, \dots, d_n\}$ orthonormal. Dann folgt

$$\delta_{kl} = \sum_{i=1}^{n} a_k^i \overline{a_l^i} \quad \text{oder} \quad A^T \bar{A} = I.$$

Durch komplexes Konjugieren erhalten wir $A^* A = I$.

Gelte umgekehrt $A^* A = I$. Mit der obigen Rechnung erhalten wir daher $\langle d_k, d_l \rangle = \delta_{kl}$. Somit ist $\{d_1, \dots, d_n\}$ orthonormiert und daher auch linear unabhängig; es handelt sich somit um eine Orthonormalbasis von V. □

Definition 7.4.10. Sei A eine $(n \times n)$-Matrix mit $A^* = A^{-1}$. Ist $A \in \mathbb{R}^{n \times n}$, so heißt A **orthogonal;** ist $A \in \mathbb{C}^{n \times n}$, so heißt A **unitär.**

Theorem 7.4.11. *Sei $A \in \mathbb{R}^{n \times n}$. Dann sind die folgenden Aussagen äquivalent:*

(i) A ist orthogonal
(ii) $A^T A = I$
(iii) $A A^T = I$
(iv) A vermittelt eine Basistransformation zwischen orthonormierten Basen eines n-dimensionalen euklidischen Vektorraumes.
(v) Die Spalten der Matrix A bilden eine orthonormierte Basis des Vektorraumes $\mathbb{R}^n$ mit Standardskalarprodukt.
(vi) Die Zeilen der Matrix A bilden eine orthonormierte Basis des Vektorraumes $\mathbb{R}^n$ mit Standardskalarprodukt.

Beweis. Nach Definition sind die Aussagen (i), (ii) und (iii) äquivalent. Theorem 7.4.9 impliziert, dass (ii) und (iv) äquivalent sind.

Sei $A = \left(a_j^i\right)_{1 \le i,j \le n}$. Dann folgt aus (ii)

$$\sum_{i=1}^{n} a_k^i a_l^i = \delta_{kl} \quad \text{für } 1 \le k, l \le n.$$

(v) folgt. Die Umkehrung „(v) $\Longrightarrow$ (ii)“ folgt ebenso aus dieser Gleichung.

Genauso wie man die Äquivalenz zwischen (ii) und (v) zeigt, erhält man auch die Äquivalenz von (iii) und (vi). □

Theorem 7.4.12. *Sei $A \in \mathbb{C}^{n \times n}$. Dann sind die folgenden Aussagen äquivalent:*

(i) A ist unitär
(ii) $A^* A = I$

(iii) $AA^* = I$
(iv) *A vermittelt eine Basistransformation zwischen orthonormierten Basen eines n-dimensionalen unitären Vektorraumes.*
(v) *Die Spalten der Matrix A bilden eine orthonormierte Basis des Vektorraumes $\mathbb{C}^n$ mit Standardskalarprodukt.*
(vi) *Die Zeilen der Matrix A bilden eine orthonormierte Basis des Vektorraumes $\mathbb{C}^n$ mit Standardskalarprodukt.*

Beweis. Vollständig analog zum reellen Fall. □

7.5. **Orthogonale und unitäre Endomorphismen.**

Definition 7.5.1. Sei V ein Skalarproduktraum. Dann heißt ein Endomorphismus $f\colon V \to V$ **orthogonal** (bzw. **unitär**), falls

$$\langle f(a), f(b)\rangle = \langle a, b\rangle \quad \text{für alle } a, b \in V$$

gilt.

Theorem 7.5.2. *Sei V ein endlichdimensionaler Skalarproduktraum. Sei $f\colon V \to V$ ein orthogonaler (bzw. unitärer) Endomorphismus. Dann gilt*

(i) $\|f(a)\| = \|a\|$
(ii) *Sind a und b orthogonal, so auch $f(a)$ und $f(b)$.*
(iii) *Ist λ ein Eigenwert von f, so gilt $|\lambda| = 1$.*
(iv) *f ist eine Isomorphismus.*

Beweis.
(i) Es ist
$$\|f(a)\| = \sqrt{\langle f(a), f(a)\rangle} = \sqrt{\langle a, a\rangle} = \|a\|.$$
(ii) Dies folgt aus
$$0 = \langle a, b\rangle = \langle f(a), f(b)\rangle.$$
(iii) Ist a ein Eigenvektor zum Eigenwert λ, so gilt nach nach (i)
$$\|a\| = \|f(a)\| = \|\lambda a\| = |\lambda| \cdot \|a\|.$$
Wegen $a \neq 0$ folgt die Behauptung.
(iv) Wegen (iii) ist 0 kein Eigenwert. Also folgt $\ker f = \{0\}$, f ist also injektiv. Da V endlichdimensional ist, ist f auch surjektiv und damit ein Isomorphismus.

□

Bemerkung 7.5.3. ⋆ Wir bemerken, dass (iv) für unendlichdimensionale Skalarprodukträume nicht aus unserer Definition folgt: Betrachte $f\colon l^2 \to l^2$ mit

$$\langle (a_i)_{i\in\mathbb{N}}, (b_j)_{j\in\mathbb{N}}\rangle := \sum_{i\in\mathbb{N}} a_i \bar{b}_i$$

und $l^2 \equiv l^2(\mathbb{N}, \mathbb{C}) \subset \mathbb{C}^{\mathbb{N}}$ der Teilmenge, auf der die zugehörige Norm endlich ist, sowie

$$f(a_0, a_1, a_2, a_3, \ldots) := (0, a_0, a_1, a_2, a_3, \ldots).$$

Dann ist f nicht surjektiv. Die anderen Teilaussagen gelten mit demselben Beweis auch für unendlichdimensionale Skalarprodukträume.

Theorem 7.5.4. *Sei $\{a_1, \ldots, a_n\}$ eine Orthonormalbasis von V. Dann ist ein Endomorphismus $f\colon V \to V$ genau dann orthogonal (bzw. unitär), falls*

$$\langle f(a_i), f(a_j)\rangle = \delta_{ij}$$

für alle $1 \leq i, j \leq n$ gilt.

Beweis. Es ist klar, dass aus der Orthogonalität (bzw. Unitarität) $\langle f(a_i), f(a_j)\rangle = \delta_{ij}$ folgt.

Gelte also $\langle f(a_i), f(a_j)\rangle = \delta_{ij}$. Wir wollen nachweisen, dass f „das Skalarprodukt erhält“: Seien $a = \sum_{i=1}^{n} \lambda^i a_i$ und $b = \sum_{j=1}^{n} \mu^j a_j$. Wir erhalten

$$\begin{aligned}\langle f(a), f(b)\rangle &= \left\langle f\left(\sum_{i=1}^{n} \lambda^i a_i\right), f\left(\sum_{j=1}^{n} \mu^j a_j\right)\right\rangle \\ &= \sum_{i,j=1}^{n} \lambda^i \overline{\mu^j} \underbrace{\langle f(a_i), f(a_j)\rangle}_{=\delta_{ij}=\langle a_i, a_j\rangle} \\ &= \left\langle \sum_{i=1}^{n} \lambda^i a_i, \sum_{j=1}^{n} \mu^j a_j \right\rangle = \langle a, b\rangle.\end{aligned}$$

Somit ist f orthogonal (bzw. unitär). □

Theorem 7.5.5. *Sei V ein endlichdimensionaler Skalarproduktraum. Sei S eine Orthonormalbasis. Dann ist ein Endomorphismus $f\colon V \to V$ genau dann orthogonal (bzw. unitär), wenn die Matrix A von f bezüglich S orthogonal (bzw. unitär) ist, d. h. wenn $A^*A = I$ gilt.*

Beweis. Sei $S = \{b_1, \dots, b_n\}$ und $A = \left(a^i_j\right)_{1\le i,j\le n}$, d. h. es gelte

$$f(b_k) = \sum_{i=1}^{n} a^i_k b_i.$$

Wir erhalten

$$\langle f(b_k), f(b_l)\rangle = \sum_{i,j=1}^{n} a^i_k \overline{a^j_l} \underbrace{\langle b_i, b_j\rangle}_{=\delta_{ij}} = \sum_{i=1}^{n} \overline{a^i_l} a^i_k.$$

Nach Theorem 7.5.4 ist f genau dann orthogonal, wenn $\langle f(b_k), f(b_l)\rangle = \delta_{kl}$ gilt. Aufgrund unserer Rechnung ist dies aber äquivalent zu $\delta_{kl} = \sum_{i=1}^{n} \overline{a^i_l} a^i_k$ und somit auch zu $A^*A = I$. Die Behauptung folgt. □

Beispiel 7.5.6. Bezüglich des Standardskalarproduktes des $\mathbb{R}^2$ sind die Matrizen

$$\begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}$$

für beliebige $\varphi \in \mathbb{R}$ orthogonal, insbesondere also auch I und $-I$.

Theorem 7.5.7. *Seien A, B orthogonale (bzw. unitäre) $(n \times n)$-Matrizen. Dann gelten*

(i) A^{-1} *ist orthogonal (bzw. unitär).*
(ii) AB *ist orthogonal (bzw. unitär).*
(iii) $|\det A| = 1$.

Die orthogonalen (bzw. unitären) $(n \times n)$-Matrizen bilden somit eine Untergruppe der invertierbaren Matrizen $GL_n(F)$,

- *die orthogonale Gruppe $O(n)$ bzw.*
- *die unitäre Gruppe $U(n)$.*

Beweis.

(i) Nach Definition ist $A^*A = I$. Weiterhin gilt $AA^* = I$. Also ist A^* die Inverse zu A, $A^{-1} = A^*$. Es folgt aus

$$\left(A^{-1}\right)^* A^{-1} = (A^*)^* A^* = AA^* = I,$$

dass auch A^{-1} orthogonal ist.

(ii) Dies folgt aus

$$(AB)^*(AB) = B^* (A^*A) B = B^*IB = B^*B = I.$$

(iii) Es gilt

$$\begin{aligned} 1 = \det I = \det (A^*A) = \det A^* \cdot \det A = \det \overline{A}^T \cdot \det A \\ = \det \overline{A} \cdot \det A = \overline{\det A} \cdot \det A = |\det A|^2. \end{aligned}$$

Die Behauptung folgt. □

Weiterhin definiert man:

Definition 7.5.8.

(i) Die **spezielle lineare Gruppe** $SL(n, F) \equiv SL_n(F)$ besteht aus den Elementen der "general linear group" $GL(n, F) \equiv GL_n(F)$ mit Determinante gleich 1.

(ii) Die **spezielle orthogonale Gruppe** $SO(n)$ besteht aus den orthogonalen Matrizen A mit $\det A = 1$.

Bemerkung 7.5.9. Der Determinantenmultiplikationssatz liefert, dass es sich damit tatsächlich um Untergruppen von $GL_n(F)$ bzw. $O(n)$ handelt.

7.6. **Orthogonale Komplemente.**

Definition 7.6.1. Sei V ein Skalarproduktraum und sei $S \subset V$ beliebig. Dann heißt

$$S^\perp := \{a \in V : \langle a, b\rangle = 0 \quad \text{für alle } b \in S\}$$

das **orthogonale Komplement** von S in V.

Theorem 7.6.2. *Sei V ein Skalarproduktraum und sei $S \subset V$. Dann gelten*

(i) $S^\perp$ *ist ein Unterraum von* V,
(ii) $\langle S\rangle \cap S^\perp = \{0\}$,
(iii) $\left(S^\perp\right)^\perp \supset \langle S\rangle$ *und*
(iv) $S^\perp = \langle S\rangle^\perp$.

Beweis.

(i) Seien $a, b \in S^\perp$ und $\lambda, \mu \in F$. Dann gilt für alle $c \in S$

$$\langle \lambda a + \mu b, c\rangle = \lambda\langle a, c\rangle + \mu\langle b, c\rangle = 0.$$

Somit ist auch $\lambda a + \mu b \in S^\perp$. Wegen $0 \in S^\perp$ ist $S^\perp$ nicht leer.

(ii) Sei $a \in S^\perp \cap \langle S\rangle$. Dann gibt es $\lambda^i \in F$ und $a_i \in S$ mit

$$\sum_{i=1}^{n} \lambda^i a_i = a.$$

Wegen $a \in S^\perp$ folgt hieraus

$$\langle a, a\rangle = \sum_{i=1}^{n} \overline{\lambda^i}\langle a, a_i\rangle = 0.$$

Somit ist $a = 0$ und wir erhalten die Behauptung.

(iii) Seien $a \in S$ und $b \in S^\perp$. Dann gilt $\langle a, b\rangle = \overline{\langle b, a\rangle} = 0$. Somit ist $S \subset \left(S^\perp\right)^\perp$. Da $\left(S^\perp\right)^\perp$ ein Unterraum ist, folgt sogar $\langle S\rangle \subset \left(S^\perp\right)^\perp$.

(iv) Wegen $S \subset \langle S\rangle$ ist klar, dass $\langle S\rangle^\perp \subset S^\perp$ gilt. Sei also $a \in S^\perp$ und $b = \sum_{i=1}^{n} \lambda^i b_i$ mit $b_i \in S$ ein beliebiges Element in $\langle S\rangle$. Dann gilt

$$\langle a, b\rangle = \sum_{i=1}^{n} \overline{\lambda^i}\langle a, b_i\rangle = 0.$$

Somit gilt auch $a \in \langle S\rangle^\perp$. □

Bemerkung 7.6.3. Ist $\dim V = \infty$, kann $\langle S\rangle \subsetneq \left(S^\perp\right)^\perp$ gelten: In $l^2 \equiv l^2(\mathbb{N}) \equiv l^2(\mathbb{N}, \mathbb{R}) \subset \mathbb{R}^{\mathbb{N}}$ ist $S = \{e_0, e_1, e_2, \ldots\}$ mit $S^\perp = \{0\}$ aber $\langle S\rangle \subsetneq l^2$ ein Gegenbeispiel, da $l^2(\mathbb{N}, \mathbb{R})$ auch Folgen mit unendlich vielen von Null verschiedenen Einträgen enthält.)

Theorem 7.6.4. *Sei V ein endlichdimensionaler Skalarproduktraum. Ist $U \subset V$ ein Unterraum, so gilt $V = U \oplus U^\perp$.*

Beweis. Nach Theorem 7.6.2 gilt $U \cap U^\perp = \{0\}$. Es genügt also nachzuweisen, dass sich jeder Vektor $c \in V$ in der Form $c = a + b$ mit $a \in U$ und $b \in U^\perp$ darstellen lässt. Sei $\{a_1, \ldots, a_r\}$ eine Orthonormalbasis von U. Definiere a durch

$$a := \langle c, a_1\rangle a_1 + \ldots + \langle c, a_r\rangle a_r$$

und setze $b := c - a$. Dann ist offensichtlicherweise $c = a + b$. Wir müssen also noch zeigen, dass $b \in U^\perp$ gilt: Es ist

$$\begin{aligned}\langle b, a_i\rangle &= \langle c - a, a_i\rangle \\ &= \langle c, a_i\rangle - \left\langle \sum_{j=1}^{r} \langle c, a_j\rangle a_j, a_i \right\rangle \\ &= \langle c, a_i\rangle - \sum_{j=1}^{r} \langle c, a_j\rangle \delta_{ji} = 0.\end{aligned}$$

Somit steht b orthogonal zu einer Basis von U. Wir erhalten $b \in U^\perp$ und die Behauptung folgt. □

Theorem 7.6.5. *Sei V ein endlichdimensionaler Skalarproduktraum. Ist $U \subset V$ ein Unterraum, so gelten*

(i) $\dim V = \dim U + \dim U^\perp$.

(ii) $\left(U^\perp\right)^\perp = U$,

Beweis.

(i) Dies folgt direkt aus Theorem 7.6.4 und der Dimensionsformel, Theorem 4.3.6.

(ii) Es gilt $V = U \oplus U^\perp$ und $V = \left(U^\perp\right)^\perp \oplus \left(\left(U^\perp\right)^\perp\right)^\perp$. Nach Theorem 7.6.2 gilt aber $U \subset \left(U^\perp\right)^\perp$ und ebenso $U^\perp \subset \left(\left(U^\perp\right)^\perp\right)^\perp$. Dies ist aber nur möglich, wenn bereits $U = \left(U^\perp\right)^\perp$ und $U^\perp = \left(\left(U^\perp\right)^\perp\right)^\perp$ gelten. □

Beispiel 7.6.6. Sei V der Vektorraum der auf $[-a, a]$, $a > 0$, stetigen reellwertigen Funktionen mit (L^2-)Skalarprodukt

$$\langle f, g\rangle := \int_{-a}^{a} f(x)g(x)\,dx.$$

Sei $U := \{f \in V\colon f(-x) = -f(x)\}$ der Raum der ungeraden Funktionen und $G := \{f \in V\colon f(-x) = f(x)\}$ der Raum der geraden Funktionen in V. Wir behaupten, dass $U^\perp = G$ gilt.

Beweis. Zunächst einmal ist klar, dass $G \subset U^\perp$ gilt. Sei nun $h \in U^\perp$ beliebig. Wir setzen $h_1(x) := \frac{1}{2}(h(x) + h(-x))$ und $h_2(x) := \frac{1}{2}(h(x) - h(-x))$. Dann gilt $h = h_1 + h_2$. Es ist $h_1 \in G$ und $h_2 \in U$. Wir sind fertig, wenn wir $h_2 \equiv 0$ zeigen können. Aus $h \in U^\perp$ und $h_2 \in U$ erhalten wir

$$0 = \langle h, h_2 \rangle = \int_{-a}^{a} h(x)h_2(x)\,dx = \int_{-a}^{a} h_1(x)h_2(x)\,dx + \int_{-a}^{a} h_2(x)h_2(x)\,dx.$$

Das erste Integral auf der rechten Seite mit einer geraden und einer ungeraden Funktion verschwindet. Da h und somit auch h_2 stetig ist, folgt also $h_2 \equiv 0$ wie behauptet. □

7.7. **Adjungierte Abbildungen.** Zu einem Endomorphismus $f\colon V \to V$ wollen wir eine lineare Abbildung $g\colon V \to V$ mit

$$\langle f(a), b \rangle = \langle a, g(b) \rangle$$

finden.

Definition 7.7.1. Sei V ein Skalarproduktraum. Zu $a \in V$ definieren wir $a^*\colon V \to \mathbb{R}$ (bzw. $a^*\colon V \to \mathbb{C}$) durch

$$a^*(b) = \langle b, a \rangle \quad \text{für } b \in V.$$

Bemerkung 7.7.2. Die Abbildung a^* ist linear, denn es gilt

$$a^*(\lambda b + \mu c) = \langle \lambda b + \mu c, a \rangle = \lambda \langle b, a \rangle + \mu \langle c, a \rangle = \lambda a^*(b) + \mu a^*(c).$$

Theorem 7.7.3. *Sei V ein endlichdimensionaler Skalarproduktraum. Dann gibt es zu jedem Funktional φ ein eindeutig bestimmtes $a \in V$ mit $\varphi(\xi) = \langle \xi, a \rangle$ für alle $\xi \in V$. Somit ist $\varphi = a^*$.*

Beweis. Sei $\{a_1, \ldots, a_n\}$ eine orthonormierte Basis von V. Definiere a durch

$$a := \sum_{i=1}^{n} \overline{\varphi(a_i)} a_i.$$

Wir erhalten

$$\langle a_j, a \rangle = \sum_{i=1}^{n} \overline{\overline{\varphi(a_i)}} \langle a_j a_i \rangle = \varphi(a_j).$$

Also stimmen a^* und φ auf einer Basis von V überein und sind daher gleich.

Es bleibt noch zu zeigen, dass a eindeutig bestimmt ist. Gelte also $\langle \xi, a \rangle = \langle \xi, a' \rangle$ für alle $\xi \in V$. Dies ist äquivalent zu $\langle \xi, a - a' \rangle = 0$ für alle $\xi \in V$ und gilt insbesondere für $\xi = a - a'$. Also folgt $a = a'$. □

Bemerkung 7.7.4. Aus diesem Beweis folgt insbesondere, dass $\langle \xi, a \rangle = 0$ für alle $\xi \in V$ nur für $a = 0$ erfüllt sein kann.

Bemerkung 7.7.5. Sei $f\colon V \to V$ ein Endomorphismus. Ist $a \in V$ fest, so definieren wir durch

$$\varphi(\xi) := \langle f(\xi), a \rangle$$

ein lineares Funktional $\varphi\colon V \to F$: Es gilt nämlich

$$\varphi(\lambda b + \mu c) = \langle f(\lambda b + \mu c), a \rangle = \lambda \langle f(b), a \rangle + \mu \langle f(c), a \rangle = \lambda \varphi(b) + \mu \varphi(c).$$

Nach Theorem 7.7.3 gibt es zu φ ein eindeutig bestimmtes $a_0 \in V$ mit $\varphi(\xi) = \langle \xi, a_0 \rangle$ für alle $\xi \in V$. Für fixiertes f ordnen wir auf diese Weise jedem $a \in V$ ein $a_0 \in V$ zu. Wir bezeichnen diese Zuordnung als $g\colon V \to V$, $g(a) := a_0$. g ist durch die Gleichung

$$\langle f(\xi), a \rangle = \langle \xi, g(a) \rangle \quad \text{für alle } \xi \in V$$

festgelegt. Wir behaupten, dass g eine lineare Abbildung ist. Aus der definierenden Gleichung erhalten wir

$$\begin{aligned}\langle \xi, g(\lambda a + \mu b)\rangle &= \langle f(\xi), \lambda a + \mu b\rangle = \overline{\lambda}\langle f(\xi), a\rangle + \overline{\mu}\langle f(\xi), b\rangle \\ &= \overline{\lambda}\langle \xi, g(a)\rangle + \overline{\mu}\langle \xi, g(b)\rangle = \langle \xi, \lambda g(a) + \mu g(b)\rangle.\end{aligned}$$

Da dies für alle $\xi \in V$ gilt, erhalten wir $g(\lambda a + \mu b) = \lambda g(a) + \mu g(b)$, also die Linearität von g.

Wir sagen, dass g die zu f adjungierte Abbildung ist und schreiben $g = f^*$.

Definition 7.7.6. Die zum Endomorphismus $f\colon V \to V$ **adjungierte Abbildung** $f^*\colon V \to V$ ist durch

$$\langle f(\xi), \eta\rangle = \langle \xi, f^*(\eta)\rangle \quad \text{für alle } \xi, \eta \in V$$

definiert.

Beispiele 7.7.7.

(i) Sei $f\colon \mathbb{R}^3 \to \mathbb{R}^3$ durch

$$f\left(x^1, x^2, x^3\right) := \left(x^1 - x^2, -x^1 + x^2 + 2x^3, x^2 + x^3\right)$$

gegeben. Nach Definition gilt $\langle f(x), y\rangle = \langle x, g(y)\rangle$. Somit erhalten wir aus

$$\begin{aligned}\langle f(x), y\rangle &= \left(x^1 - x^2\right) y^1 + \left(-x^1 + x^2 + 2x^3\right) y^2 + \left(x^2 + x^3\right) y^3 \\ &= x^1\left(y^1 - y^2\right) + x^2\left(-y^1 + y^2 + y^3\right) + x^3\left(2y^2 + y^3\right) \\ &= \langle x, f^*(y)\rangle\end{aligned}$$

die adjungierte Abbildung

$$f^*\left(y^1, y^2, y^3\right) := \left(y^1 - y^2, -y^1 + y^2 + y^3, 2y^2 + y^3\right).$$

Die darstellenden Matrizen bezüglich der Standardbasis sind

$$A_f = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 1 & 2 \\ 0 & 1 & 1 \end{pmatrix} \quad \text{sowie} \quad A_{f^*} = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 1 & 1 \\ 0 & 2 & 1 \end{pmatrix}.$$

(ii) Im Raum $V = L^2([a, b])$ der auf $[a, b]$ stetigen reellwertigen Funktionen mit L^2-Skalarprodukt ist

$$\begin{aligned}\Phi&\colon V \to V, \\ \Phi(f)(x) &:= x \cdot f(x)\end{aligned}$$

selbstadjungiert, d. h. es gilt $\Phi^* = \Phi$, denn es gilt

$$\langle \Phi(f), g\rangle = \langle f, \Phi(g)\rangle.$$

Den Zusammenhang zwischen der f und der f^* darstellenden Matrix haben wir bereits im Beispiel gesehen.

Theorem 7.7.8. *Sei V ein endlichdimensionaler Skalarproduktraum. Sei S eine Orthonormalbasis. Wird $f\colon V \to V$ bezüglich S durch die Matrix A dargestellt, so wird die adjungierte Abbildung $f^*\colon V \to V$ bezüglich S durch die Matrix A^* dargestellt.*

Beweis. Seien $S = (d_1, \ldots, d_n)$, $A = \left(a^i_j\right)_{1 \le i,j \le n}$ und $A_{f^*} = \left(b^i_j\right)_{1 \le i,j \le n}$. Wir erhalten

$$\begin{aligned}f(d_k) &= \sum_{i=1}^{n} a^i_k d_i, \\ f^*(d_l) &= \sum_{j=1}^{n} b^j_l d_j,\end{aligned}$$

$$\begin{aligned}
\langle f(d_k), d_l\rangle &= \langle d_k, f^*(d_l)\rangle,\\
\langle f(d_k), d_l\rangle &= \left\langle \sum_{i=1}^n a_k^i d_i, d_l \right\rangle = \sum_{i=1}^n a_k^i \langle d_i, d_l\rangle = \sum_{i=1}^n a_k^i \delta_{il} = a_k^l,\\
\langle d_k, f^*(d_l)\rangle &= \left\langle d_k, \sum_{j=1}^n b_l^j d_j \right\rangle = \sum_{j=1}^n \overline{b_l^j}\delta_{kj} = \overline{b_l^k}.
\end{aligned}$$

(In den letzten beiden Zeilen sind die jeweils letzten Gleichheiten aus Kovarianzgründen unschön geschrieben. Für die zugehörigen reellen bzw. komplexen Zahlen gilt die Gleichheit jedoch.) Also ist $a_k^l = \overline{b_l^k}$ für alle $1 \le k, l \le n$. Die Behauptung folgt. □

Theorem 7.7.9. *Sei V ein endlichdimensionaler Skalarproduktraum. Sei $f\colon V \to V$ linear. Dann gilt*

$$\ker f^* = (\operatorname{im} f)^\perp.$$

Beweis. Es gilt

$$\begin{aligned}
\ker f^* &= \{\xi \in V\colon f^*(\xi) = 0\}\\
&= \{\xi \in V\colon \langle \eta, f^*(\xi)\rangle = 0 \quad \text{für alle } \eta \in V\}\\
&= \{\xi \in V\colon \langle f(\eta), \xi\rangle = 0 \quad \text{für alle } \eta \in V\}\\
&= (\operatorname{im} f)^\perp.
\end{aligned}$$

□

Beispiel 7.7.10. Sei

$$\sum_{k=1}^n a_k^i x^k = b^i, \quad 1 \le i \le n$$

ein reelles lineares quadratisches Gleichungssystem. Sei $f\colon \mathbb{R}^n \to \mathbb{R}^n$ die (bezüglich der Standardbasis) durch $A = \left(a_j^i\right)$ dargestellte Abbildung. Dann ist das lineare Gleichungssystem genau dann lösbar, wenn $b = \left(b^1, \dots, b^n\right) \in \operatorname{im} f$ gilt. Auf $\mathbb{R}^n$ führen wir das Standardskalarprodukt ein und erhalten aus Theorem 7.7.9, dass das lineare Gleichungssystem genau dann lösbar ist, wenn $b \in (\ker f^*)^\perp$ gilt. Über $\mathbb{R}$ wird f^* durch die transponierte Matrix dargestellt. Somit folgt: Das lineare Gleichungssystem ist genau dann lösbar, wenn für jede Lösung $y = (y_1, \dots, y_n)$ von

$$\sum_{i=1}^n a_k^i y_i = 0 \quad \text{auch} \quad \sum_{i=1}^n y_i b^i = 0 \quad \text{gilt.}$$

Theorem 7.7.11. *Sei V ein endlichdimensionaler Skalarproduktraum. Sei $f\colon V \to V$ ein Endomorphismus. Dann haben f und f^* den gleichen Rang.*

Beweis. Es gilt

$$\begin{aligned}
\operatorname{rang} f &= \dim \operatorname{im} f = \dim V - \dim(\operatorname{im} f)^\perp\\
&= \dim V - \dim(\ker f^*) && \text{(Theorem 7.7.9)}\\
&= \dim(\operatorname{im} f^*) = \operatorname{rang} f^*.
\end{aligned}$$

(Alternativ kann man über $\mathbb{R}$ wegen $A^* = A^T$ benutzen, dass Zeilenrang und Spaltenrang übereinstimmen.) □

Korollar 7.7.12. *Sei $A \in \mathbb{C}^{n\times n}$. Dann gilt* $\operatorname{rang} A = \operatorname{rang} \overline{A}$.

Beweis. Setze $B := \overline{A}^T$. Da der Rang einer Matrix B mit dem von B^T übereinstimmt, folgt

$$\operatorname{rang} A = \operatorname{rang} B^* = \operatorname{rang} B = \operatorname{rang} B^T = \operatorname{rang} \overline{A}.$$

□

7.8. **Diagonalisierung von selbstadjungierten linearen Endomorphismen.** Wir wollen insbesondere zeigen, dass sich solche Endomorphismen sogar mit Hilfe einer Orthogonalbasis stets diagonalisieren lassen.

Definition 7.8.1. Sei V ein Skalarproduktraum. Der Endomorphismus $f\colon V \to V$ heißt **selbstadjungiert,** wenn

$$\langle f(\xi), \eta\rangle = \langle \xi, f(\eta)\rangle$$

für alle $\xi, \eta \in V$ gilt, also $f = f^*$ ist.

Theorem 7.8.2. *Sei V ein endlichdimensionaler Skalarproduktraum mit Orthonormalbasis S. Der Endomorphismus $f\colon V \to V$ werde bezüglich S durch die Matrix A beschrieben. Dann ist $f\colon V \to V$ genau dann selbstadjungiert, wenn $A = A^*$ gilt.*

Beweis. Dies folgt direkt aus Theorem 7.7.8. □

Theorem 7.8.3. *Sei V ein n-dimensionaler Skalarproduktraum. Sei $f\colon V \to V$ ein selbstadjungierter Endomorphismus. Dann sind alle Nullstellen des charakteristischen Polynoms $\chi_f(X)$ reell, es gilt also*

$$\chi_f(X) = \pm(X - \lambda_1)\cdots(X - \lambda_n)$$

für $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$.

Beweis. Wir betrachten zunächst den unitären Fall. Wie jedes Polynom zerfällt das charakteristische Polynom in Linearfaktoren und wir müssen lediglich nachweisen, dass die Eigenwerte von f reell sind. Sei also λ ein Eigenwert von f zum Eigenvektor ξ. Wir erhalten

$$\lambda\langle\xi,\xi\rangle = \langle\lambda\xi,\xi\rangle = \langle f(\xi),\xi\rangle = \langle\xi, f(\xi)\rangle = \langle\xi,\lambda\xi\rangle = \overline{\lambda}\langle\xi,\xi\rangle.$$

Wegen $\xi \neq 0$ bzw. $\langle\xi,\xi\rangle \neq 0$ folgt $\lambda = \overline{\lambda}$. Somit ist λ reell.

Im reellen Fall benutzen wir die sogenannte Komplexifizierung. Die zu $f\colon V \to V$ gehörige reelle symmetrische Matrix definiert eine Abbildung $\tilde{f}\colon \mathbb{C}^n \to \mathbb{C}^n$. Eine reelle symmetrische Matrix ist hermitesch, wenn wir sie als komplexe Matrix auffassen. Somit ist $\tilde{f}\colon \mathbb{C} \to \mathbb{C}$ bezüglich des Standardskalarproduktes auf $\mathbb{C}^n$ selbstadjungiert. Es gilt

$$\chi_f(X) = \chi_A(X) = \chi_{\tilde{f}}(X).$$

Somit stimmen die Eigenwerte von f und $\tilde{f}$ überein und sind aufgrund der obigen Überlegungen reell. □

Theorem 7.8.4. *Sei V ein endlichdimensionaler Skalarproduktraum. Sei $f\colon V \to V$ ein selbstadjungierter Endomorphismus. Dann besitzt V eine Orthonormalbasis aus Eigenvektoren von f.*

Beweis. Wir wissen, dass das charakteristische Polynom $\chi_f(X)$ in Linearfaktoren zerfällt. (Im Fall $F = \mathbb{C}$ folgt dies aus dem Fundamentalsatz der Algebra und für $F = \mathbb{R}$ (und $\mathbb{C}$) aus Theorem 7.8.3.)

Benutze Induktion nach $\dim V =: n$. Für $n = 1$ ist die Aussage wahr. Sei also $n \geq 2$. Sei a_1 ein Eigenvektor zum Eigenwert λ mit $\|a_1\| = 1$. Definiere $W := \{\xi \in V\colon \langle a_1, \xi\rangle = 0\}$. Nach Theorem 7.6.5 folgt $\dim W = n - 1$. W ist ein bezüglich f invarianter Unterraum, es gilt nämlich für $\xi \in W$

$$\langle a_1, f(\xi)\rangle = \langle f(a_1), \xi\rangle = \langle\lambda a_1, \xi\rangle = \lambda\langle a_1, \xi\rangle = 0.$$

Es folgt $f(\xi) \in W$. Nach Induktionsvoraussetzung existiert eine Orthonormalbasis $\{a_2, \ldots, a_n\}$ von W aus Eigenvektoren von $f|_W$ und somit auch von f. Also ist $\{a_1, a_2, \ldots, a_n\}$ wie behauptet eine Orthonormalbasis. □

Theorem 7.8.5. *Jeder selbstadjungierte Endomorphismus eines endlichdimensionalen Skalarproduktraumes ist diagonalisierbar.*

Beweis. Benutze eine Orthonormalbasis aus Eigenvektoren. Bezüglich dieser ist die zugehörige Matrix eine Diagonalmatrix mit den Eigenwerten entsprechend ihrer Vielfachheit auf der Diagonalen. □

Theorem 7.8.6. *Sei A eine reelle symmetrische (bzw. komplexe hermitesche) $(n \times n)$-Matrix. Dann gibt es eine orthogonale (bzw. unitäre) Matrix D mit*

$$DAD^* = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}$$

mit $\lambda_1, \dots, \lambda_n \in \mathbb{R}$.

Beweis. Dies folgt direkt aus den bisherigen Theoremen dieses Kapitels, da für orthogonale (bzw. unitäre) Matrizen $D^{-1} = D^*$ gilt. □

7.9. **Kästchenform orthogonaler Matrizen.** Wir wollen unitäre und orthogonale Endomorphismen bezüglich Orthonormalbasen durch „einfache“ Matrizen darstellen.

Theorem 7.9.1. *Sei $f\colon V \to V$ ein unitärer Endomorphismus eines endlichdimensionalen unitären Vektorraumes V. Dann gibt es eine Orthonormalbasis von V, die aus Eigenvektoren von f besteht.*

Beweis. Wir beweisen dies per Induktion nach $n = \dim V$. Der Induktionsanfang ist trivial. Sei $n \geq 2$. Das charakteristische Polynom zerfällt über $\mathbb{C}$ in Linearfaktoren

$$\chi_f(X) = \pm(X - \lambda_1)(X - \lambda_2) \cdots (X - \lambda_n).$$

Nach Theorem 7.5.2 gilt $|\lambda_i| = 1$. Sei a_1 ein Eigenvektor zum Eigenwert λ_1 mit $\|a_1\| = 1$. Sei W das orthogonale Komplement von a_1, $V = W \oplus \langle a_1 \rangle$. Für jedes $\xi \in W$ gilt

$$0 = \langle a_1, \xi \rangle = \langle f(a_1), f(\xi) \rangle = \lambda_1 \langle a_1, f(\xi) \rangle.$$

Da $\lambda_1 \neq 0$ gilt, folgt $f(\xi) \in W$. Somit induziert die unitäre Abbildung f eine unitäre Abbildung $f|_W\colon W \to W$. Nach Induktionsvoraussetzung gibt es für $f|_W$ eine Orthonormalbasis aus Eigenvektoren. Zusammen mit a_1 erhält man wie gewünscht eine Orthogonalbasis aus Eigenvektoren. □

Korollar 7.9.2. *Sei $A \in \mathbb{C}^{n \times n}$ unitär. Dann gibt es eine unitäre Matrix D mit*

$$DAD^* = \begin{pmatrix} e^{i\varphi_1} & 0 & \dots & 0 \\ 0 & e^{i\varphi_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & e^{i\varphi_n} \end{pmatrix} \quad \textit{und} \quad 0 \leq \varphi_i < 2\pi.$$

Lemma 7.9.3. *Sei $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ orthogonal. Fixiere eine Orthonormalbasis. Dann gibt es $0 \leq \varphi < 2\pi$, so dass f bezüglich dieser Basis durch*

$$\begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} \quad \textit{oder} \quad \begin{pmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{pmatrix}$$

dargestellt wird.

Beweis. Wir benutzen Theorem 7.4.11. Die Spalten und die Zeilen der Matrix bilden also eine Orthonormalbasis. Einen beliebigen Einheitsvektor in $\mathbb{R}^2$ können wir als

$$\left(\cos\varphi, \pm\sqrt{1-\cos^2\varphi}\right) = (\cos\varphi, \pm\sin\varphi)$$

schreiben. Somit ist f durch eine Matrix der Form

$$\begin{pmatrix} \cos\varphi & \pm\sin\varphi \\ \pm\sin\varphi & \pm\cos\varphi \end{pmatrix}$$

dargestellt. Man überzeugt sich leicht, dass die Vorzeichen für $\sin\varphi \neq 0$ bzw. $\cos\varphi \neq 0$ wie angegeben gewählt werden müssen um eine orthogonale Matrix zu erhalten. Ggf. ist φ durch $-\varphi + 2\pi k$, $k \in \mathbb{Z}$, zu ersetzen um Vorzeichen abzuändern. □

Durch eine spezielle Basiswahl können wir Kästchen der zweiten Form ausschließen.

Theorem 7.9.4. *Sei $f\colon V \to V$ eine orthogonale Selbstabbildung eines euklidischen Vektorraumes mit $\dim V < \infty$. Dann gibt es eine Orthonormalbasis von V, in der sich f in der Form*

$$\begin{pmatrix}
\mathbb{1}_k & 0 & 0 & 0 & \dots & 0 & 0 \\
0 & -\mathbb{1}_l & 0 & 0 & \dots & 0 & 0 \\
0 & 0 & \begin{matrix}\cos\varphi_1 & -\sin\varphi_1 \\ \sin\varphi_1 & \cos\varphi_1\end{matrix} & 0 & \dots & 0 & 0 \\
0 & 0 & 0 & & & & 0 \\
\vdots & \vdots & \vdots & & \ddots & & \vdots \\
0 & 0 & 0 & & & & 0 \\
0 & 0 & 0 & 0 & \dots & 0 & \begin{matrix}\cos\varphi_m & -\sin\varphi_m \\ \sin\varphi_m & \cos\varphi_m\end{matrix}
\end{pmatrix}$$

mit $k + l + 2m = \dim V$ darstellen lässt.

(In „Matrizensprache“ übersetzt erhalten wir: Zu jeder orthogonalen Matrix A gibt es also eine orthogonale Matrix D, so dass DAD^T von der angegebenen Gestalt ist.)

Beweis. Vermöge einer Matrixdarstellung $A \in \mathbb{R}^{\dim V \times \dim V}$ ordnen wir f wiederum eine Abbildung $\tilde{f}\colon \mathbb{C}^n \to \mathbb{C}^n$ zu. Wegen $A^*A = A^TA = \mathbb{1}$ ist A als Matrix in $\mathbb{C}^{n\times n}$ unitär. Nach Theorem 7.9.1 gibt es eine Orthonormalbasis aus Eigenvektoren für $\tilde{f}$ zu Eigenwerten λ_i mit $|\lambda_i| = 1$. Zu jedem λ_i mit $\lambda_i = \pm 1$ finden wir einen reellen Eigenvektor zu diesem Eigenwert, den Realteil des komplexen Eigenvektors: Aus $A\xi = \lambda_i\xi$ folgt $A\overline{\xi} = \lambda_i\overline{\xi}$ und $A\left(\xi + \overline{\xi}\right) = \lambda_i\left(\xi + \overline{\xi}\right)$. Daher ist $\operatorname{Re}\xi$ ein Eigenvektor oder der Nullvektor. Ist $\operatorname{Re}\xi$ der Nullvektor, so ist $i\xi$ ein reeller Eigenvektor.

Da das orthogonale Komplement eines Eigenvektors jeweils invariant unter f ist, erhalten wir induktiv die beiden Blöcke $\mathbb{1}_k$ und $\mathbb{1}_l$.

Sei ξ ein Eigenvektor zum Eigenwert $e^{i\varphi} \notin \{-1, 1\}$, $A\xi = e^{i\varphi}\xi$. Durch komplexe Konjugation erhalten wir einen zweiten Eigenvektor: $A\overline{\xi} = e^{-i\varphi}\overline{\xi}$. Setze $u := \xi + \overline{\xi}$ und $v := -i\left(\xi - \overline{\xi}\right)$. Da ξ und $\overline{\xi}$ wegen $e^{i\varphi} \notin \{-1, 1\}$ Eigenvektoren von A zu unterschiedlichen Eigenwerten sind, gilt $\xi \neq \pm\overline{\xi}$ und somit $u \neq 0 \neq v$. Wir erhalten

$$\begin{aligned}
Au &= A\xi + A\overline{\xi} = e^{i\varphi}\xi + e^{-i\varphi}\overline{\xi} = \frac{e^{i\varphi} + e^{-i\varphi}}{2}\left(\xi + \overline{\xi}\right) + \frac{e^{i\varphi} - e^{-i\varphi}}{2}\left(\xi - \overline{\xi}\right) \\
&= \cos\varphi \cdot u - \sin\varphi \cdot v, \\
-\frac{1}{i}Av &= A\xi - A\overline{\xi} = e^{i\varphi}\xi - e^{-i\varphi}\overline{\xi} = \frac{e^{i\varphi} - e^{-i\varphi}}{2}\left(\xi + \overline{\xi}\right) + \frac{e^{i\varphi} + e^{-i\varphi}}{2}\left(\xi - \overline{\xi}\right) \\
&= i\sin\varphi \cdot u + \cos\varphi \cdot \frac{v}{-i}, \\
Av &= \sin\varphi \cdot u + \cos\varphi \cdot v.
\end{aligned}$$

Wir wählen $\frac{1}{\sqrt{2}}u$ und $\frac{1}{\sqrt{2}}v$ als reellen Teil einer Orthonormalbasis. Beachte dazu, dass aus $\langle \xi, \overline{\xi} \rangle_{\mathbb{C}} = 0$ die Relationen $\left\| \xi + \overline{\xi} \right\|^2 = 2$, $\left\| -i \left(\xi - \overline{\xi} \right) \right\|^2 = 2$ und $\left\langle \xi + \overline{\xi}, -i \left(\xi - \overline{\xi} \right) \right\rangle = 0$ folgen, zunächst mit komplexen Skalarprodukten und Normen. Da die Einträge aber reell sind, gelten dieselben Beziehungen auch über $\mathbb{R}$ und zugehörigen Skalarprodukten und Normen. Dies liefert den ersten Block:

$$\begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix} = \begin{pmatrix} \cos\varphi_1 & -\sin\varphi_1 \\ \sin\varphi_1 & \cos\varphi_1 \end{pmatrix}, \quad \varphi_1 = -\varphi.$$

Da beide Zeilen und Spalten, die durch diesen Block verlaufen, bereits „in diesem Block Länge Eins haben“, enthalten sie sonst nur Nullen. Die restlichen Blöcke erhält man per Induktion. □

Literatur

1. Gerd Fischer, *Lineare Algebra*, fifth ed., Grundkurs Mathematik, vol. 17, Friedr. Vieweg & Sohn, Braunschweig, 1979.
2. Paul R. Halmos, *Naive Mengenlehre*, Vandenhoeck & Ruprecht, Göttingen, 1976.
3. Serge Lang, *Linear algebra*, third ed., Undergraduate Texts in Mathematics, Springer-Verlag, New York, 1989.
4. Falko Lorenz, *Lineare Algebra. I*, third ed., Bibliographisches Institut, Mannheim, 1992.
5. Oliver C. Schnürer, *Analysis I*, 2015, Skript zur Vorlesung.
6. Urs Stammbach, *Lineare Algebra*, Teubner Studienskripten, vol. 82, B. G. Teubner, Stuttgart, 1980.

Oliver C. Schnürer, Fachbereich Mathematik und Statistik, Universität Konstanz, 78457 Konstanz, Germany

E-mail address: Oliver.Schnuerer@uni-konstanz.de