Skript zu den Vorlesungen Lineare Algebra I am KIT im Wintersemester 2020/21

Rafael Dahmen
Wilderich Tuschmann

19. März 2021

Inhaltsverzeichnis

1. Zur Sprache der Mathematik **5**
- 1.1. Elementare Aussagenlogik . . . 5
- 1.2. (Naive) Mengenlehre . . . 7
- 1.3. Funktionen und Abbildungen . . . 13

2. Lineare Gleichungssysteme und Matrizen **21**
- 2.1. Lineare Gleichungssysteme mit reellen Koeffizienten . . . 21
- 2.2. Matrizenrechnung . . . 24
- 2.3. Homogene lineare Gleichungssysteme und Untervektorräume von $\mathbb{R}^n$ . . . 29
- 2.4. Affine Unterräume . . . 43
- 2.5. Der Gauß-Algorithmus . . . 49

3. Algebraische Strukturen **61**
- 3.1. Halbgruppen . . . 61
- 3.2. Gruppen . . . 65
- 3.3. Ringe und Körper . . . 75
- 3.4. Der Ring der ganzen Zahlen und seine Quotientenringe . . . 81
- 3.5. Der Körper der komplexen Zahlen . . . 91

4. Vektorräume und lineare Abbildungen **101**
- 4.1. Grundlegendes zu Vektorräumen und linearen Abbildungen . . . 101
- 4.2. Lineare Hülle, Basis, Dimension . . . 112
- 4.3. Lineare Abbildungen zwischen endlich dimensionalen Vektorräumen und Darstellungsmatrizen . . . 123
- 4.4. Direkte Summen und Komplemente . . . 146
- 4.5. Quotientenvektorräume . . . 152
- 4.6. Der Dualraum . . . 160

5. Endomorphismen **169**
- 5.1. Das Signum einer Permutation . . . 169
- 5.2. Alternierende Abbildungen . . . 174
- 5.3. Determinanten . . . 177
- 5.4. Polynome . . . 187
- 5.5. Endomorphismen und Ähnlichkeit . . . 202
- 5.6. Eigenwerte und Eigenvektoren . . . 208
- 5.7. Trigonalisierbarkeit und der Satz von Cayley-Hamilton . . . 222
- 5.8. Nilpotente Endomorphismen und die Jordansche Normalform . . . 228

A. Anhang **237**
- A.1. Das griechische Alphabet . . . 237
- A.2. Zusatzmaterial über endliche Körper . . . 238

Stichwortverzeichnis . . . 243

1. Zur Sprache der Mathematik

1.1. Elementare Aussagenlogik

Bemerkung 1.1.1 (Aussagen)**.**
Unter einer *(logischen) Aussage* verstehen wir einen Ausdruck, von dem eindeutig bestimmt ist, ob er *wahr* oder *falsch* ist. Beispiele für Aussagen sind z.B. „A ist der erste Buchstabe des lateinischen Alphabets“, „3 ist eine Primzahl“, „$3 > 5$“, „$1+1=4$“. Der Ausdruck „Nachts ist es kälter als draußen.“ ist keine Aussage in diesem Sinne, da unklar ist, was sie bedeuten soll; der Ausdruck „Alle Personen in diesem Raum sind sympathisch.“ ist ebenfalls keine Aussage in diesem Sinne, weil verschiedene Beobachter ihr verschiedene Wahrheitswerte zuweisen würden.

Wir können Aussagen miteinander verknüpfen, um aus ihnen neue Aussagen zu gewinnen:

Definition 1.1.2 (Verknüpfung von Aussagen)**.**
Gegeben seien Aussagen $\mathscr{A}$ und $\mathscr{B}$.

(a) Die Aussage $(\neg\mathscr{A})$ (gesprochen „nicht $\mathscr{A}$“) ist die Aussage „Die Aussage $\mathscr{A}$ ist falsch.“ und heißt die *Negation* von $\mathscr{A}$.

(b) Die Aussage $(\mathscr{A} \wedge \mathscr{B})$ (gesprochen „$\mathscr{A}$ und $\mathscr{B}$“) ist die Aussage „Die Aussage $\mathscr{A}$ ist wahr und die Aussage $\mathscr{B}$ ist wahr.“ und heißt die *Konjunktion* von $\mathscr{A}$ und $\mathscr{B}$.

(c) Die Aussage $(\mathscr{A} \vee \mathscr{B})$ (gesprochen „$\mathscr{A}$ oder $\mathscr{B}$“) ist die Aussage „Mindestens eine der beiden Aussagen $\mathscr{A}$ oder $\mathscr{B}$ ist wahr.“ Sie heißt die *(nicht-exklusive) Disjunktion* von $\mathscr{A}$ und $\mathscr{B}$.

(d) Die Aussage $(\mathscr{A} \implies \mathscr{B})$ (gesprochen „$\mathscr{A}$ impliziert $\mathscr{B}$“ oder „aus $\mathscr{A}$ folgt $\mathscr{B}$“) ist die Aussage „Wenn $\mathscr{A}$ wahr ist, dann ist auch $\mathscr{B}$ wahr.“ Sie heißt die *Implikation*.

Die Konjunktion ist also ein logisches „und“ und die Disjunktion ist ein logisches (nicht-exklusives) „oder“. Nach Definition sind die Wahrheitswerte dieser Aussagen (mit w = wahr und f = falsch) durch die folgende Tabelle gegeben.

$\mathscr{A}$	$\mathscr{B}$	$\neg\mathscr{A}$	$\mathscr{A} \wedge \mathscr{B}$	$\mathscr{A} \vee \mathscr{B}$	$\mathscr{A} \implies \mathscr{B}$
w	w	f	w	w	w
w	f	f	f	w	f
f	w	w	f	w	w
f	f	w	f	f	w

Gelegentlich werden auch andere Sprechweisen verwendet: Statt „$\mathscr{A}$ impliziert $\mathscr{B}$“ sagt man auch „wenn $\mathscr{A}$, dann $\mathscr{B}$“, „$\mathscr{B}$ folgt aus $\mathscr{A}$“, „$\mathscr{A}$ ist hinreichend für $\mathscr{B}$“ oder „$\mathscr{B}$ ist notwendig für $\mathscr{A}$“.

Bemerkung 1.1.3 (Wahrheitstafeln)**.**
Auch wenn es theoretisch möglich ist, mathematische Aussagen per Wahrheitstafel zu beweisen, so werden wir dies für gewöhnlich nicht tun.

(a) Eine mathematische Aussage der Form $\neg\mathscr{A}$ beweist man stets so, dass man annimmt, $\mathscr{A}$ sei wahr, und diese Annahme zu einem Widerspruch führt.

(b) Wenn wir eine Aussage der Form $\mathscr{A} \wedge \mathscr{B}$ beweisen wollen, werden wir für gewöhnlich so vorgehen, dass wir zuerst $\mathscr{A}$ beweisen und dann anschließend $\mathscr{B}$ (oder in umgekehrter Reihenfolge).

(c) Wenn wir eine Aussage der Form $\mathscr{A} \vee \mathscr{B}$ beweisen wollen, werden wir für gewöhnlich beweisen, dass $\mathscr{A}$ wahr ist oder dass $\mathscr{B}$ wahr ist.

(d) Eine mathematische Aussage der Form $\mathscr{A} \implies \mathscr{B}$ werden wir meist so beweisen: Wir nehmen an, $\mathscr{A}$ sei wahr, und folgern dann unter dieser zusätzlichen Annahme, dass $\mathscr{B}$ gilt.

Definition 1.1.4.
Gegeben seien Aussagen $\mathscr{A}, \mathscr{B}$. Die Aussage $\mathscr{A} \iff \mathscr{B}$ ist eine Kurzschreibweise für

$$((\mathscr{A} \implies \mathscr{B}) \wedge (\mathscr{B} \implies \mathscr{A}))$$

und heißt die *Äquivalenz* von $\mathscr{A}$ und $\mathscr{B}$. Ist diese Aussage wahr, so sagen wir „$\mathscr{A}$ ist äquivalent zu $\mathscr{B}$“. Statt „$\mathscr{A}$ ist äquivalent zu $\mathscr{B}$“ sagt man auch „$\mathscr{A}$ gilt, *genau dann, wenn* $\mathscr{B}$ gilt“.

Bemerkung 1.1.5.
Um eine Aussage der Form $\mathscr{A} \iff \mathscr{B}$ zu beweisen, ist es also notwendig, zwei Implikationen zu beweisen: Zuerst $\mathscr{A} \implies \mathscr{B}$ und dann $\mathscr{B} \implies \mathscr{A}$ (oder in umgekehrter Reihenfolge).

Wenn wir dann wissen, dass eine Äquivalenz $\mathscr{A} \iff \mathscr{B}$ wahr ist, so bedeutet das, dass beide Aussagen denselben Wahrheitswert besitzen, also entweder beide wahr oder beide falsch sind.
Notation: Wir verwenden zwischen Aussagen niemals ein Gleichheitszeichen.

Satz 1.1.6 (Elementare logische Umformungen)**.**
Für beliebige Aussagen $\mathscr{A}, \mathscr{B}, \mathscr{C}$ sind die folgenden Aussagen stets wahr:

(i) $(\neg(\neg\mathscr{A})) \iff \mathscr{A}$. (doppelte Negation)

(ii) $\mathscr{A} \vee (\neg\mathscr{A})$. (Tertium non datur.)

(iii) $\neg(\mathscr{A} \implies \mathscr{B}) \iff (\mathscr{A} \wedge \neg\mathscr{B})$ (Negation einer Implikation)

(iv) $(\mathscr{A} \implies \mathscr{B}) \iff ((\neg\mathscr{B}) \implies (\neg\mathscr{A}))$. (Kontrapositionsprinzip)

(v) $((\mathscr{A} \implies \mathscr{B}) \wedge (\mathscr{B} \implies \mathscr{C})) \implies (\mathscr{A} \implies \mathscr{C})$. (Transitivität der Implikation)

(vi) $((\mathscr{A} \wedge \mathscr{B}) \wedge \mathscr{C}) \iff (\mathscr{A} \wedge (\mathscr{B} \wedge \mathscr{C}))$
$((\mathscr{A} \vee \mathscr{B}) \vee \mathscr{C}) \iff (\mathscr{A} \vee (\mathscr{B} \vee \mathscr{C}))$. (Assoziativität von $\wedge$ und $\vee$)

(vii) $(\mathscr{A} \wedge \mathscr{B}) \iff (\mathscr{B} \wedge \mathscr{A})$
$(\mathscr{A} \vee \mathscr{B}) \iff (\mathscr{B} \vee \mathscr{A})$. (Kommutativität von $\wedge$ und $\vee$)

(viii) $\mathscr{A} \wedge (\mathscr{B} \vee \mathscr{C}) \iff (\mathscr{A} \wedge \mathscr{B}) \vee (\mathscr{A} \wedge \mathscr{C})$
$\mathscr{A} \vee (\mathscr{B} \wedge \mathscr{C}) \iff (\mathscr{A} \vee \mathscr{B}) \wedge (\mathscr{A} \vee \mathscr{C})$. (Distributivität)

(ix) $(\neg(\mathscr{A} \wedge \mathscr{B})) \iff ((\neg\mathscr{A}) \vee (\neg\mathscr{B}))$
$(\neg(\mathscr{A} \vee \mathscr{B})) \iff ((\neg\mathscr{A}) \wedge (\neg\mathscr{B}))$. (De Morgansche[1] Regeln)

[1] nach AUGUSTUS DE MORGAN, engl. Mathematiker, 1806–1871

Zusammenfassung von Abschnitt 1.1

(1) Aussagen sind Ausdrücke, denen ein eindeutiger Wahrheitswert („wahr“ oder „falsch“) zugeordnet werden kann.

(2) Aussagen können negiert und mittels Konjunktion, Disjunktion und Implikation zu komplizierteren verknüpft werden.

(3) Aussagen sind äquivalent, wenn sie sich gegenseitig implizieren, d.h. den gleichen Wahrheitswert besitzen.

1.2. (Naive) Mengenlehre

Fast die gesamte Mathematik basiert auf dem Begriff der *Menge*. Schauen wir uns also an, was eine Menge ist. Die folgende Definition geht auf CANTOR[2] zurück:

> Unter einer *Menge* verstehen wir jede Zusammenfassung M von bestimmten wohlunterschiedenen Objekten unserer Anschauung oder unseres Denkens (welche die *Elemente* von M genannt werden) zu einem Ganzen.

Falls x ein Element von M ist, schreiben wir „$x \in M$“, sonst „$x \notin M$“.

Aus Sicht der modernen Mathematik ist das keine brauchbare Definition. Zunächst ist unklar, was ein denkbares Objekt ist, aber in diese philosophischen Untiefen wollen wir uns hier lieber nicht begeben. Problematischer ist, dass man einer Zusammenfassung von Objekten nicht unbedingt ansehen kann, ob die Zugehörigkeit zu dieser Zusammenfassung zweifelsfrei entschieden werden kann. BERTRAND RUSSELL[3] hat hierzu das folgende Beispiel gegeben:

Beispiel 1.2.1 (Die Russellsche (Klasse oder) „Nichtmenge“)**.**
Es sei M die Zusammenfassung aller Mengen, die kein Element von sich selbst sind. Wäre M selbst eine Menge, so hätten wir folgendes Problem: Falls $M \in M$, so wäre M eine Element von sich selbst, also $M \notin M$. Falls $M \notin M$, so ist M nicht in sich selbst enthalten, also würde gelten $M \in M$. In beiden Fällen erhalten wir also einen Widerspruch. Daher kann M keine Menge sein.

Obgleich dieses Beispiel sehr konstruiert wirkt und im mathematischen „Alltag“ keine Bedeutung hat, so zeigt es doch, dass die „naive“ Mengendefinition zu Widersprüchen führt und somit mit Vorsicht genossen werden muss.[4]

Alle Mengen, die wir hier konstruieren werden, sind in diesem Sinne jedoch unproblematisch.

Bemerkung 1.2.2.
Wie können wir konkret eine Menge M angeben?

(i) Wir können bei endlichen Mengen (d.h. Mengen, die nur endlich viele Elemente enthalten) einfach alle Elemente der Menge angeben. Dazu benutzen wir die folgende Schreibweise:

$$A := \{1, a, \alpha\}$$

[2] nach GEORG FERDINAND LUDWIG PHILIPP CANTOR, dt. Mathematiker, 1845–1918.

[3] engl. Philosoph und Logiker, 1872–1970, Literaturnobelpreisträger 1950

[4] Die Lösung dieses Problems ist es, die Mengentheorie axiomatisch zu formalisieren. Dies werden wir hier aber nicht näher betrachten.

Hier ist „:=“ das *Zuweisungssymbol*; auf der linken Seite steht der Name der Menge („Definiendum“) und auf der rechten Seite („Definiens“) werden die Elemente der Menge von den Mengenklammern „{“ und „}“ eingerahmt. Beispiele für solcherart beschriebene endliche Mengen sind z.B.

$$A := \{0, 8, 15\}, \quad B := \{\text{Hund}, \text{Katze}, \text{Maus}, \text{Elefant}\}$$

Man beachte, dass Elemente einer Menge keine Zahlen sein müssen. Eine Menge kann auch Geraden, Punkte, Kreise, Funktionen, Vektoren, Tensoren, Matrizen oder sogar selbst wieder Mengen enthalten.

Spezialfall: Die Menge

$$\emptyset := \{\}$$

ohne Elemente heißt *leere Menge*.

(ii) Unendliche Mengen können wir auf diese Weise natürlich nicht beschreiben, aber wenn wir schreiben

$$\mathbb{N} := \{1,2,3,4,5,\ldots\} \quad \text{bzw.} \quad \mathbb{Z} := \{\ldots,-3,-2,-1,0,1,2,3,\ldots\}$$

dann werden Sie sicher verstehen, dass hier die Menge der *natürlichen Zahlen*[5] bzw. der *ganzen Zahlen* gemeint ist.

Grundsätzlich ist diese „Pünktchen“-Schreibweise allerdings nur dann zu verwenden, wenn man davon ausgehen kann, dass klar ist, wie es weitergeht. Beispielsweise ist (für die meisten)

$$A := \{1,2,4,8,16,\ldots\}$$

als Menge der Zweierpotenzen zu erkennen, aber die Definition

$$D := \{19,6,20,2,6,17,6,19,0,13,\ldots\}$$

ergibt keinen Sinn, weil nicht ersichtlich ist, was für eine Regelmäßigkeit hier gemeint ist. Im Zweifelsfall sollten Sie also versuchen, die gesuchte Menge ohne „Pünktchen“-Notation zu beschreiben.

(iii) Wir können Mengen auch in beschreibender oder *impliziter Form* angeben: Für die Menge aller Primzahlen schreiben wir z.B.

$$P := \{p \mid p \text{ ist eine Primzahl}\}.$$

Vor dem Trennstrich wird hier eine Variable eingeführt, und nach dem Trennstrich werden die Eigenschaften aufgezählt, die diese Variable erfüllen muss.

Der Name der Variablen ist hierbei völlig egal[6]. Das heißt, die folgende Definition definiert dieselbe Menge P:

$$P := \{\gamma \mid \gamma \text{ ist eine Primzahl}\}.$$

[5] Für uns gehört 0 nicht zu den natürlichen Zahlen. Dies ist vollkommen willkürlich und kann in anderen Vorlesungen oder Büchern anders sein.

[6] Man sagt auch, p ist hier eine *lokale Variable* – sie existiert nur innerhalb der Mengenklammern.

Manchmal steht auch links von dem Trennstrich ein komplizierter Ausdruck, z.B. bedeutet

$$Q := \{x^2 \mid x \text{ ist eine Primzahl}\},$$

dass wir jedes Objekt rechts von dem Trennstrich (also jede Primzahl) nehmen und dann mit diesem Objekt die Operation links durchführen (also quadrieren) und alle Ergebnisse in einer neuen, so definierten Menge zusammenfassen. Die Menge Q besteht in diesem Beispiel also aus allen Quadraten von Primzahlen, also $Q = \{2^2,3^2,5^2,7^2,\ldots\}$.

Definition 1.2.3 (Gleichheit von Mengen)**.**
Zwei Mengen sind gleich, wenn sie die gleichen Elemente besitzen: Für zwei Mengen A und B gilt also $A = B$ genau dann, wenn jedes Element in A auch ein Element in B ist und umgekehrt.

Bemerkung 1.2.4 (Gleichheit von Mengen)**.**
Beispielsweise sind die Mengen

$$A := \{0,8,15\} \quad \text{und} \quad B := \{8,15,8,0\}$$

gleich, denn jedes Element in der linken Menge ist auch Element der rechten Menge und jedes Element in der rechten Menge ist auch ein Element in der linken Menge. Insbesondere bedeutet dies, dass die Reihenfolge, in der wir die Elemente aufschreiben, keine Rolle spielt und dass es keinen Unterschied macht, ob wir die Zahl 8 einmal oder mehrfach in die Menge schreiben.

Bemerkung 1.2.5 (Aussageformen und Quantoren)**.**
Die Ausdrücke „p ist eine Primzahl." und „$x^2 = 4$" sind keine Aussagen, weil ihr Wahrheitswert davon abhängt, welche Zahl man für die Variablen p bzw. x einsetzt. Man nennt einen solchen Ausdruck eine *Aussageform* und die darin vorkommenden nicht erklärten Begriffe *freie Variablen*. Wenn man in einer Aussageform jede freie Variable durch einen passenden Begriff ersetzt, erhält man eine Aussage. Z.B. sind „$2^2 = 4$" und „$3^2 = 4$" Aussagen (die aber nicht notwendigerweise wahr sind).

Ist $\mathscr{A}(x)$ eine Aussageform über einer Menge M, so können wir zwei neue Aussagen definieren:

(i) Die *All-Aussage:* „Für jedes Element $x \in M$ ist $\mathscr{A}(x)$ wahr."

(ii) Die *Existenz-Aussage:* „Es gibt mindestens ein Element $x \in M$, für das $\mathscr{A}(x)$ wahr ist."

Wir schreiben diese Aussagen symbolisch als

$$\forall\, x \in M : \mathscr{A}(x) \quad \text{bzw.} \quad \exists\, x \in M : \mathscr{A}(x).$$

Das Symbol $\forall$ heißt der *Allquantor*, das Symbol $\exists$ heißt der *Existenzquantor*. Wir erhalten also Aussagen aus Aussageformen durch Anwenden solcher Quantoren. In der mathematischen Sprechweise bedeutet „Es gibt ein ..." immer „es gibt mindestens ein ...", d.h. die Anzahl der Elemente mit der gegebenen Eigenschaft ist größer oder gleich 1.

Beispiel 1.2.6.
Die Aussage $\forall\, x \in \mathbb{N} :\ x^2 = 4$ ist falsch, aber die Aussage $\exists\, x \in \mathbb{N} :\ x^2 = 4$ ist wahr.

Bei Aussageformen mit mehreren freien Variablen führt das Anwenden eines Quantors wieder zu einer Aussageform, und erst wenn alle Variablen quantifiziert sind, erhält man eine Aussage.

Bemerkung 1.2.7.
Beim Quantifizieren von Aussageformen mit mehreren freien Variablen ist die Reihenfolge der Quantoren entscheidend: Die Aussage $\forall x \in \mathbb{N}\, \exists y \in \mathbb{N} : x < y$ ist wahr, und die Aussage $\exists y \in \mathbb{N}\, \forall x \in \mathbb{N} : x < y$ ist falsch.

Bemerkung 1.2.8 (De Morgansche Regeln für Quantoren)**.**
Es sei $\mathscr{A}(x)$ eine Aussageform. Dann gelten die folgenden Regeln:

$$\neg(\forall x : \mathscr{A}(x)) \iff (\exists x : \neg\mathscr{A}(x)),$$
$$\neg(\exists x : \mathscr{A}(x)) \iff (\forall x : \neg\mathscr{A}(x)).$$

Beim Negieren einer Aussage werden also All-Quantoren zu Existenz-Quantoren und umgekehrt.

Definition 1.2.9 (Teilmenge)**.**
Eine Menge M heißt *Teilmenge* einer Menge N, falls gilt

$$\forall x : (x \in M \implies x \in N),$$

wenn also alle Elemente in M auch Elemente in N sind. Wir schreiben dann $M \subseteq N$.

Der Strich in „$\subseteq$“ deutet an, dass M und N auch gleich sein dürfen. Wenn wir sagen wollen, dass M eine Teilmenge von N, aber nicht gleich N ist, so schreiben wir

$$M \subsetneq N :\iff (M \subseteq N) \wedge (M \neq N).$$

Wir nennen dann M eine *echte Teilmenge* von N.

Bemerkung 1.2.10.
Achtung: Die Notation „$M \subset N$“ bedeutet – je nach Autor und Vorlesung/Buch – entweder $M \subseteq N$ oder $M \subsetneq N$.

Bemerkung 1.2.11 (Gleichheit vs. Inklusion von Mengen)**.**
Nach Definition 1.2.3 gilt für zwei Mengen M und N:

$$\begin{aligned}(M = N) &\iff (\forall x : (x \in M \iff x \in N))\\ &\iff (\forall x : (x \in M \implies x \in N) \wedge (x \in N \implies x \in M))\\ &\iff ((M \subseteq N) \wedge (N \subseteq M)).\end{aligned}$$

Um zu zeigen, dass zwei Mengen gleich sind, kann man also zeigen, dass sie jeweils ineinander enthalten sind (genauso, wie man die Gültigkeit einer Äquivalenz über den Nachweis zweier Implikationen beweisen kann).

Bemerkung 1.2.12 (Teilmengen mittels Aussageformen)**.**
Ist M eine Menge und $\mathscr{A}(x)$ eine Aussageform über M, so schreiben wir

$$\{x \in M \mid \mathscr{A}(x)\} := \{x \mid (x \in M) \wedge (\mathscr{A}(x) \text{ ist wahr})\}$$

für die Menge aller x aus M, die $\mathscr{A}(x)$ erfüllen.

Definition 1.2.13.
Gegeben seien Mengen M und N.

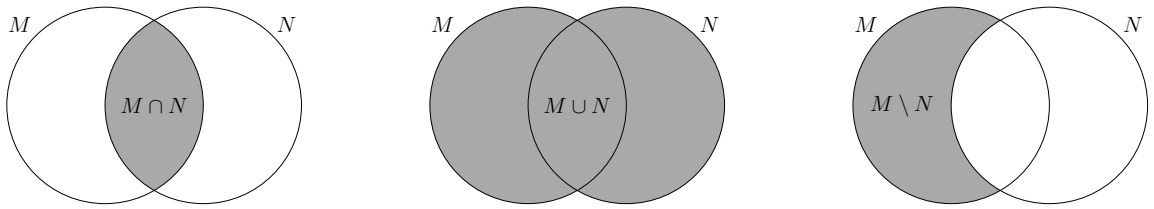


Abbildung 1.1.: Mengenoperationen

(i) $M \cap N := \{x \mid (x \in M) \wedge (x \in N)\}$ heißt der *Schnitt von M und N*.

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

(iii) $M \setminus N := \{x \mid (x \in M) \wedge (x \notin N)\}$ heißt das *Komplement von N in M*.

Wir können diese Definitionen für Teilmengen der Anschauungsebene anhand sogenannter *Venn-Diagramme* veranschaulichen, wie in Abbildung 1.1 dargestellt.

Bemerkung 1.2.14.
Nachdem wir also in Bemerkung 1.2.11 bereits gesehen haben, dass in gewisser Weise „Gleichheit von Mengen“ der „Äquivalenz von Aussagen“ entspricht und die „Teilmengenrelation“ der Implikation zwischen Aussagen zugeordnet werden kann, wissen wir nun, dass der „Schnitt“ in der Sprache der Mengenlehre der „Konjunktion“ und die „Vereinigung“ der „Disjunktion“ entspricht. Sind M und N Mengen, so sind $M \setminus N$, $M \cap N$ und $M \cup N$ wieder Mengen. Dagegen ist

$$\{x \mid x \notin N\}$$

keine Menge (siehe Beispiel 1.2.1). Es gibt also in der Sprache der Mengenlehre keine direkte Entsprechung für die Negation.

Dieses Problem kann man beheben, wenn man nur Mengen betrachtet, die sich innerhalb einer bereits gegebenen, umgebenden Menge befinden. In diesem Fall entspricht die logische Negation genau dem Komplement einer Menge A innerhalb der umgebenden Menge X:

$$\text{Für } A \subseteq X \text{ gilt:} \quad X \setminus A = \{x \in X \mid x \notin A\}.$$

Wir können nun alle elementaren logischen Umformungen aus Satz 1.1.6 in die Sprache der Mengenlehre übersetzen:

Satz 1.2.15 (Elementare Rechenregeln für Mengen)**.**
Gegeben seien Mengen X, A, B, C mit $A, B, C \subseteq X$. Dann gilt:

(i) $X \setminus (X \setminus A) = A$. (doppelte Negation)

(ii) $A \cup (X \setminus A) = X$. (Tertium non datur)

(iii) $\neg(A \subseteq B) \iff (\exists x : x \in A \wedge x \notin B)$. (Negation der Teilmengenrelation)

(iv) $(A \subseteq B) \iff (X \setminus B \subseteq X \setminus A)$. (Kontrapositionsprinzip)

(v) $((A \subseteq B) \wedge (B \subseteq C)) \Longrightarrow (A \subseteq C)$. (Transitivität der Teilmengenrelation)

(vi) $(A \cap B) \cap C = A \cap (B \cap C)$,
$(A \cup B) \cup C = A \cup (B \cup C)$. (Assoziativität von $\cap$ und $\cup$)

(vii) $A \cap B = B \cap A$,
$A \cup B = B \cup A$. (Kommutativität von $\cap$ und $\cup$)

(viii) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$,
$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$. (Distributivität)

(ix) $X \setminus (A \cap B) = (X \setminus A) \cup (X \setminus B)$,
$X \setminus (A \cup B) = (X \setminus A) \cap (X \setminus B)$. (De Morgansche Regeln)

Bemerkung 1.2.16 (Symmetrische Differenz)**.**
Gegeben seien zwei Mengen A, B. Dann ist die *symmetrische Differenz* von A und B definiert als

$$A \triangle B := (A \setminus B) \cup (B \setminus A).$$

Diese entspricht dem *exklusiven Oder* (XOR), das Sie eventuell aus der Informatik kennen.

Notation 1.2.17.
Sind M und N Mengen, und ist $m \in M$ und $n \in N$, so bezeichnet (m,n) das *geordnete Paar* bestehend aus $m \in M$ und $n \in N$. Zwei solche Paare (m_1,n_1) und (m_2,n_2) sind nach Definition genau dann gleich, wenn $m_1 = m_2$ und $n_1 = n_2$. Man schreibt

$$M \times N := \{(x,y) \mid x \in M, y \in N\}$$

und nennt $M \times N$ das *kartesische*[7] *Produkt* von M und N.

Die Kreuznotation ist durch die folgende Beobachtung motiviert:

Bemerkung 1.2.18.
Ist M eine endliche Menge, so bezeichnen wir mit $|M|$ oder $\#M$ die Anzahl der Elemente von M. Sind M und N endliche Mengen, so gilt

$$|M \times N| = |M| \cdot |N| \quad \text{und} \quad |M \cup N| = |M| + |N| - |M \cap N|.$$

Mittels Quantoren können wir nun auch den Schnitt und die Vereinigung von mehr als zwei Mengen definieren. (Auf Produkte von unendlich vielen Mengen werden wir hier zunächst nicht eingehen.)

Definition 1.2.19.
Es sei M eine Menge.

(i) Die Menge $\mathscr{P}(M) := \{N \mid N \subseteq M\}$ aller Teilmengen von M heißt die *Potenzmenge* von M.

(ii) Eine Teilmenge $\mathscr{A} \subseteq \mathscr{P}(M)$ heißt *Mengensystem* über M.

(iii) Ist $\mathscr{A} \subseteq \mathscr{P}(M)$ ein Mengensystem, so heißen

$$\bigcap \mathscr{A} := \bigcap_{A \in \mathscr{A}} A := \{x \in M \mid \forall\, A \in \mathscr{A} : x \in A\} \quad \text{und} \quad \bigcup \mathscr{A} := \bigcup_{A \in \mathscr{A}} A := \{x \in M \mid \exists A \in \mathscr{A} : x \in A\}$$

der *Schnitt* bzw. die *Vereinigung* über $\mathscr{A}$.

Ist M endlich, so hat $\mathscr{P}(M)$ genau $2^{|M|}$ Elemente.

[7] nach RENÉ DESCARTES, französischer Philosoph, Mathematiker und Naturgelehrter, 1596–1650

Bemerkung 1.2.20.
Beim Umgang mit Mengensystemen muss man sorgfältig zwischen den Symbolen „$\in$“ und „$\subseteq$“ unterscheiden. Z.B. gilt

$$(x \in M) \iff (\{x\} \subseteq M) \iff (\{x\} \in \mathscr{P}(M)),$$

aber weder $x \subseteq M$ noch $\{x\} \in M$ oder $\{x\} \subseteq \mathscr{P}(M)$ bedeuten dasselbe.

Zusammenfassung von Abschnitt 1.2

(1) Mengen sind gewisse gutartige Zusammenfassungen von Objekten – eine präzise Definition haben wir hier nicht gegeben.

(2) Durch die Operationen Schnitt, Vereinigung, Komplementbildung, kartesisches Produkt und Potenzmengenbildung kann man aus gegebenen Mengen neue Mengen erzeugen.

(3) Die Rechenregeln für diese Operationen ergeben sich aus den Regeln der Aussagenlogik.

(4) Um die Gleichheit zweier Mengen zu beweisen, genügt es zu zeigen, dass sie wechselseitig ineinander enthalten sind.

1.3. Funktionen und Abbildungen

Definition 1.3.1.
Eine *Funktion* oder[8] *Abbildung* ist ein Tripel (M,N,f) bestehend aus zwei Mengen M und N und einer Zuordnungsvorschrift f, die jedem $m \in M$ ein eindeutiges Element $f(m) \in N$ zuordnet.
M heißt dann *Definitionsbereich* der Funktion, N heißt ihr *Zielbereich* und (M,N,f) heißt eine Funktion *von M nach N*. Die Menge

$$\mathrm{gr}(f) := \{(m,n) \in M \times N \mid n = f(m)\} \subseteq M \times N$$

heißt der *Graph* der Funktion.

Bemerkung 1.3.2.

(1) Um zu sagen, dass (M,N,f) eine Funktion ist, die jedes $m \in M$ auf $f(m)$ in N abbildet, benutzt man auch die Schreibweise

$$f : M \to N, \quad m \mapsto f(m)$$

also schreibt man z.B.

$$p : \mathbb{N} \to \mathbb{N}, \quad n \mapsto 3n^2 + 7.$$

(2) Nach Definition sind zwei Funktionen (M,N,f) und (M',N',f') genau dann gleich , wenn $M = M'$, $N = N'$ und $\forall m \in M : f(m) = f'(m)$. Die Funktion

$$q : \mathbb{N} \to \mathbb{Z}, \quad n \mapsto 3n^2 + 7$$

ist also von der oben definierten Funktion p verschieden.

[8] Diese beiden Begriffe sind synonym. Oft wird der Begriff *Funktion* aber nur verwendet, wenn man betonen will, dass die Zielmenge $\mathbb{R}$ oder $\mathbb{C}$ ist und ansonsten wird *Abbildung* bevorzugt.

(3) Wir bezeichnen die Menge aller Funktionen von M nach N mit N^M. Sind M und N endliche Mengen, so gilt nämlich
$$|N^M| = |N|^{|M|}.$$

(4) Ist (M,N,f) eine Funktion, so ist die Funktionsvorschrift f durch den Graphen $\mathrm{gr}(f)$ eindeutig bestimmt. Es ist nämlich $f(m) = n$ genau dann, wenn $(m,n) \in \mathrm{gr}(f)$.

Notation 1.3.3.
Es sei $f : M \to N$ eine Funktion.

(1) Ist $A \subseteq M$, so heißt
$$f(A) := \{n \in N \mid \exists a \in A : f(a) = n\} = \{f(a) \mid a \in A\}$$
das *Bild* von A unter f.

(2) Speziell heißt $\mathrm{Bild}(M,N,f) := f(M)$ das *Bild* der Funktion (M,N,f).

(3) Ist $B \subseteq N$, so heißt
$$f^{-1}(B) := \{m \in M \mid f(m) \in B\}$$
das *Urbild* von B unter f.

(4) Ist $A \subseteq M$, so heißt die Funktion
$$f|_A : A \to N, \quad a \mapsto f(a)$$
die *Einschränkung* (oder *Restriktion*) von f auf A.

(5) Ist $\mathrm{Bild}(M,N,f) \subseteq B \subseteq N$, so heißt die Funktion
$$f|^B : M \to B, \quad m \mapsto f(m)$$
die *Koeinschränkung* (oder *Korestriktion*) von f auf B.

Bemerkung 1.3.4.
Wenn die Menge M aus dem Kontext klar ist, schreibt man häufig auch $\mathrm{Bild}(f)$ statt $\mathrm{Bild}(M,N,f)$. Dieses Bild hängt aber natürlich vom Definitionsbereich (nicht aber vom Zielbereich) ab.

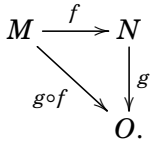
Im Folgenden seien M,N,O,P beliebige Mengen.

Definition 1.3.5 (Verkettung von Funktionen)**.**
Sind $f : M \to N$ und $g : N \to O$ Funktionen, so ist ihre *Verkettung* (oder *Komposition*) die Funktion „g nach f“ gegeben durch
$$g \circ f : M \to O, \quad m \mapsto g(f(m)).$$

Notation 1.3.6 (Pfeilnotation)**.**
Der französische Mathematiker JEAN LERAY[9] schlug in den 1940er Jahren vor, eine Funktion

[9] frz. Mathematiker, 1906–1998

$f : M \to N$ durch einen Pfeil $M \xrightarrow{f} N$ zu veranschaulichen. Sind $f : M \to N$ und $g : N \to O$ Funktionen, so erhalten wir ein Diagramm

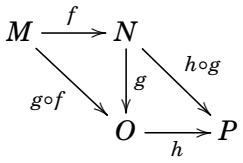


Starten wir von einem Element in M, so ist es egal, ob wir erst mittels f nach rechts, und dann mit g nach unten laufen, oder ob wir mittels $g \circ f$ diagonal abkürzen. Wir sagen, „das Diagramm kommutiert“. Viele wichtige Sätze der Mathematik lassen sich durch solche „kommutativen Diagramme“ ausdrücken.

Lemma 1.3.7 (Assoziativität der Verkettung[10])**.**
Es seien $f : M \to N$, $g : N \to O$ und $h : O \to P$ Abbildungen. Dann gilt

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

d.h. das Diagramm



kommutiert.

Beweis.
Für alle $m \in M$ gilt:

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

Folglich sind $(h \circ g) \circ f = h \circ (g \circ f)$ zwei Funktionen von M nach P mit derselben Zuordnungsvorschrift, also gleich. □

Notation 1.3.8.
Es sei M eine Menge. Dann heißt die Funktion

$$\mathrm{id}_M : M \to M, \quad m \mapsto m,$$

die jedes Element auf sich selbst abbildet die *Identitätsabbildung* (oder einfach *Identität*) auf der Menge M.

Definition 1.3.9.
Eine Funktion $f : M \to N$ heißt

(a) *injektiv*, falls für alle $m_1, m_2 \in M$ gilt :

$$(f(m_1) = f(m_2)) \Longrightarrow (m_1 = m_2);$$

(b) *surjektiv*, falls $\mathrm{Bild}(f) = f(M) = N$;

[10] Ein „Lemma“ ist ein Hilfssatz, der vor allem zum Beweis anderer Sätze verwendet wird.

(c) *bijektiv*, falls sie injektiv und surjektiv ist.

Bemerkung 1.3.10.
Für eine Funktion $f : M \to N$ gilt:

(a) Die Funktion f ist injektiv, wenn man m eindeutig daran erkennen kann, was $f(m)$ ist, d.h. für jedes $n \in N$ gibt es höchstens ein m mit $f(m) = n$. Formaler gesagt,

$$\forall\, n \in N : \left|f^{-1}(\{n\})\right| \leq 1.$$

(b) Die Funktion f ist surjektiv, wenn jedes Element von N ein Urbild hat, d.h. für jedes $n \in N$ gibt es mindestens ein $m \in M$ mit $f(m) = n$. Formal:

$$\forall\, n \in N : \left|f^{-1}(\{n\})\right| \geq 1.$$

(c) Aus (a) und (b) ergibt sich: Die Funktion f ist bijektiv genau dann, wenn es für jedes $n \in N$ genau ein $m \in M$ gibt mit $f(m) = n$. Die Elemente von M und N entsprechen sich unter f also Eins-zu-Eins, formal

$$\forall\, n \in N : \left|f^{-1}(\{n\})\right| = 1.$$

Ist f bijektiv, so gibt es also für jedes $n \in N$ ein eindeutig bestimmtes $m \in M$ mit $f(m) = n$. Wenn es eine bijektive Funktion zwischen zwei endlichen[11] Mengen M und N gibt, so folgt, dass M und N gleich viele Elemente haben, also $|M| = |N|$.

Definition 1.3.11.
Es sei $f : M \to N$ eine Funktion. Eine *Umkehrfunktion* von f ist eine Funktion

$$f^{-1} : N \to M,$$

mit der Eigenschaft $f^{-1} \circ f = \mathrm{id}_M$ und $f \circ f^{-1} = \mathrm{id}_N$

Satz 1.3.12 (Umkehrfunktion)**.**
Eine Funktion $f : M \to N$ hat genau dann eine Umkehrfunktion $f^{-1} : N \to M$, wenn f bijektiv ist. Wenn eine Umkehrfunktion existiert, ist sie stets eindeutig.

Beweis. Es sei $f : M \to N$ eine Funktion gegeben. Wir müssen zeigen:

$$f \text{ hat eine Umkehrfunktion } f^{-1} : N \to M \iff f \text{ is bijektiv.}$$

Um diese Äquivalenz zu zeige, müssen wir zwei Implikationen zeigen (siehe Definition 1.1.4):
„$\Longrightarrow$“:
Wir nehmen also an, es gibt eine Funktion $f^{-1} : N \to M$ mit $f^{-1} \circ f = \mathrm{id}_M$ und $f \circ f^{-1} = \mathrm{id}_N$. Wir müssen zeigen, dass f bijektiv, d.h. injektiv und surjektiv ist.

Zeigen wir zuerst, dass f injektiv ist. Es seinen dazu Elemente $m_1, m_2 \in M$ gegeben mit $f(m_1) = f(m_2)$. Wir müssen zeigen: $m_1 = m_2$.

[11] Für nicht notwendig endliche Mengen M und N sagt man, dass M und N *gleichmächtig* sind, falls es eine Bijektion von M nach N gibt. Man schreibt dann formal auch $|M| = |N|$.

Es gilt nun:

$$\begin{aligned} m_1 &= \mathrm{id}_M(m_1) \\ &= (f^{-1} \circ f)(m_1) \\ &= f^{-1}\big(f(m_1)\big) \\ &= f^{-1}\big(f(m_2)\big) \\ &= (f^{-1} \circ f)(m_2) \\ &= \mathrm{id}_M(m_2) \\ &= m_2. \end{aligned}$$

Also ist f injektiv.

Nun zeigen wir, dass f surjektiv ist. Es sei dazu ein Element $n \in N$ gegeben. Wir müssen zeigen: $\exists m \in M : f(m) = n$.

Wir definieren $m := f^{-1}(n) \in M$. Es bleibt zu zeigen, dass $f(m) = n$. Dies zeigen wir nun:

$$\begin{aligned} f\big(m\big) &= f\big(f^{-1}(n)\big) \\ &= (f \circ f^{-1})(n) \\ &= \mathrm{id}_N(n) \\ &= n. \end{aligned}$$

Dies zeigt, dass f surjektiv ist und mit der bereits bewiesenen Injektivität folgt die behauptete Bijektivität der Funktion f.

„$\Longleftarrow$“:

Für die andere Implikation nehmen wir jetzt an, dass f bijektiv, also injektiv und surjektiv ist. Wir wollen zeigen, dass es eine Funktion

$$f^{-1} : N \to M$$

gibt mit $f^{-1} \circ f = \mathrm{id}_M$ und $f \circ f^{-1} = \mathrm{id}_N$. Wir definieren die Funktion wie folgt:

$$f^{-1} : N \to M, \quad f(m) \mapsto h \text{ wenn } m \in M.$$

Die Frage, die sich nun stellt, ist: Ist f^{-1} *wohldefiniert*, d.h. ist f^{-1} wirklich eine Funktion, also wird jedem Element in N genau ein Element in M zugeordnet?

Wir müssen also zeigen:

- Jedes Element in N wird auf mindestens ein Element in M abgebildet.
- Jedes Element in N wird auf höchstens ein Element in M abgebildet.

Beginnen wir mit der ersten Aussage. Es sei $n \in N$ ein Element. Da die Funktion $f : M \to N$ surjektiv ist, gibt es (mindestens) ein $m \in M$ mit $f(m) = n$. Also taucht das Element $n = f(m)$ auf der linken Seite des Zuweisungspfeils „$\mapsto$“ auf und wird somit auf ein Element in M abgebildet.

Es bleibt zu zeigen, dass jedes $n \in N$ auf höchstens ein Element in M abgebildet wird. Nehmen wir also an, ein Element in N lässt sich als $f(m_1)$ und als $f(m_2)$ mit $m_1, m_2 \in M$ schreiben. Da f aber injektiv ist, folgt aus $f(m_1) = f(m_2)$ direkt, dass $m_1 = m_2$ ist und somit ist der Funktionswert von $f(m)$ eindeutig definiert.

Also ist $f^{-1} : N \to M$ wohldefiniert, d.h. wirklich eine Funktion.

Es bleibt zu zeigen, dass $f^{-1} \circ f = \mathrm{id}_M$ und $f \circ f^{-1} = \mathrm{id}_N$.
Wir beginnen mit $f^{-1} \circ f = \mathrm{id}_M$. Es sei dazu $m \in M$ ein Element. Dann gilt:

$$(f^{-1} \circ f)(m) = f^{-1}\big(f(m)\big) = m = \mathrm{id}_M(m).$$

Nun zeigen wir noch $f \circ f^{-1} = \mathrm{id}_N$. Es sei dazu $n \in N$ gegeben. Aus der Surjektivität von f folgt, dass es ein $n \in \mathbb{N}$ gibt mit $f(m) = n$. Dann gilt:

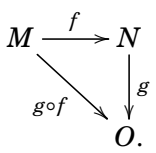
$$\begin{aligned}
(f \circ f^{-1})\big(n\big) &= (f \circ f^{-1})\big(f(m)\big) \\
&= f\big(f^{-1}\big(f(m)\big)\big) \\
&= f\big((f^{-1} \circ f)(m)\big) \\
&= f\big(\mathrm{id}_M(m)\big) \\
&= f\big(m\big) \\
&= n. \\
&= \mathrm{id}_N(n).
\end{aligned}$$

□

Bemerkung 1.3.13.
Jede Funktion kann – ohne den Definitionsbereich oder die Zuordnungsvorschrift abzuändern – surjektiv gemacht werden, indem man den Zielbereich auf das Bild einschränkt. Wenn also eine Funktion $f : M \to N$ gegeben ist, dann ist $f|^{f(M)} : M \to f(M), \quad m \mapsto f(m)$ eine surjektive Funktion, die sich von der Originalfunktion nur durch den veränderten Zielbereich unterscheidet.

Eine Funktion $f : M \to N$ hat nur dann eine Umkehrfunktion von N nach M, wenn f bijektiv ist. Es ist allerdings möglich, eine Funktion, die nur injektiv ist – wie eben beschrieben – künstlich surjektiv zu machen. Die so eingeschränkte injektive Funktion ist dann bijektiv und kann umgekehrt werden. In diesem erweiterten Sinne kann man also sogar jeder injektiven Funktion $f : M \to N$ eine Umkehrfunktion $(f|^{f(M)})^{-1} : f(M) \to M$ zuordnen.

Lemma 1.3.14.
Es seien M, N, O Mengen und $f : M \to N$ und $g : N \to O$ Funktionen.



Dann gilt:

(a) Wenn f und g beide injektiv sind, dann ist auch $g \circ f$ injektiv.

(b) Wenn f und g beide surjektiv sind, dann ist auch $g \circ f$ surjektiv.

(c) Wenn $g \circ f$ injektiv ist, dann ist f injektiv.

(d) Wenn $g \circ f$ surjektiv ist, dann ist g surjektiv.

Beweis. (a)
Gegeben seien $m_1, m_2 \in M$ mit $(g \circ f)(m_1) = (g \circ f)(m_2)$. Es ist zu zeigen, dass $m_1 = m_2$ gilt.

Nach der Definition der verketteten Abbildung $g \circ f$ gilt:

$$g\big(f(m_1)\big) = g\big(f(m_2)\big).$$

Da $g : N \to O$ injektiv ist, können wir daraus folgen:

$$f(m_1) = f(m_2).$$

Da aber auch $f : M \to N$ als injektiv vorausgesetzt wurde, gilt:

$$m_1 = m_2.$$

Das war zu zeigen.

(b)

Gegeben sei ein Element $o \in O$ im Zielbereich der Abbildung $g \circ f$. Es ist zu zeigen, dass es ein $m \in M$ gibt mit $(g \circ f)(m) = o$.

Da $g : N \to O$ surjektiv mit Zielbereich O ist und o ein Element in O ist, folgt: Es gibt ein $n \in N$ mit $g(n) = o$.

Da nun $f : M \to N$ surjektiv mit Zielbereich N ist und n ein Element in N, folgt: Es gibt ein $m \in M$ mit $f(m) = n$.

Insgesamt folgt also:

$$(g \circ f)(m) = g\big(f(m)\big) = g\big(n\big) = o.$$

Das war zu zeigen.

(c)

Gegeben seien $m_1, m_2 \in M$ mit $f(m_1) = f(m_2)$. Es ist zu zeigen, dass $m_1 = m_2$ gilt.

Wir wenden nun die Abbildung $g \circ f : M \to O$ auf das Element $m_1 \in M$ an und erhalten:

$$\begin{aligned}(g \circ f)(m_1) &= g\big(f(m_1)\big) \\ &= g\big(f(m_2)\big) \\ &= (g \circ f)(m_2).\end{aligned}$$

Also gilt $(g \circ f)(m_1) = (g \circ f)(m_2)$ und weil $g \circ f$ injektiv ist, folgt $m_1 = m_2$.

(d)

Gegeben sei ein Element $o \in O$ im Zielbereich der Abbildung g. Es ist zu zeigen, dass es ein $n \in N$ gibt mit $g(n) = o$.

Da O auch der Zielbereich der Abbildung $g \circ f$ ist und diese als surjektiv vorausgesetzt wurde, gibt es somit ein $m \in M$ mit $(g \circ f)(m) = o$. Wir setzen nun dieses m in f ein und erhalten

$$n := f(m).$$

Nun gilt:

$$g\big(n\big) = g\big(f(m)\big) = (g \circ f)(m) = o.$$

Das war zu zeigen. □

Zusammenfassung von Abschnitt 1.3

(1) Funktionen erlauben es Elemente verschiedener Mengen miteinander in Beziehung zu setzen.

(2) Funktionen mit passenden Definitions- und Zielbereichen können miteinander verkettet werden. Diese Verkettung (Komposition) ist assoziativ.

(3) Bijektive Funktionen besitzen eine Umkehrfunktion.

2. Lineare Gleichungssysteme und Matrizen

Im ersten Kapitel haben wir die Mengen $\mathbb{N} = \{1,2,\ldots\}$, $\mathbb{Z} = \{\ldots,-2,-1,0,1,2,\ldots\}$ kennengelernt. Im Folgenden wird auch die Bezeichnung $\mathbb{N}_0 := \mathbb{N} \cup \{0\} = \{0,1,2,\ldots\}$ öfters auftauchen.

Von nun an werden wir – ohne explizite Definition – auch die Menge $\mathbb{R}$ der *reellen Zahlen* als bekannt voraussetzen, deren Elemente wir uns geometrisch als Punkte auf der Zahlengeraden vorstellen. Für zwei reelle Zahlen $x, y \in \mathbb{R}$ sind die Summe $x+y \in \mathbb{R}$ und das Produkt $xy = x \cdot y \in \mathbb{R}$ definiert.

Die üblichen Rechenregeln für Summe und Produkt sowie deren Umkehrungen, Differenz und Quotient, setzen wir hier als bekannt voraus und verweisen auf die Veranstaltungen *Analysis I* bzw. *Höhere Mathematik I für Informatik*, in denen die Menge $\mathbb{R}$ sowie die Grundrechenarten auf dieser inklusive der gültigen Regeln auf eine solide mathematische Grundlage gestellt werden.

2.1. Lineare Gleichungssysteme mit reellen Koeffizienten

Beispiel 2.1.1 (Banales Beispiel)**.**
Eine lineare Gleichung mit einer Unbekannten sieht wie folgt aus:

$$ax = b, \tag{2.1.1}$$

wobei $a, b \in \mathbb{R}$ gegeben sind.

Falls $a \neq 0$ ist, können wir beide Seiten durch a teilen und erhalten $x = \frac{b}{a} \in \mathbb{R}$ als einzige Lösung. Die Lösungsmenge ist damit $\left\{\frac{b}{a}\right\}$. Es gibt also genau ein Element in der Lösungsmenge.

Falls $a = 0$ ist, ist dies nicht möglich, und wir erhalten stattdessen die Gleichung:

$$0x = b.$$

Falls nun $b = 0$ ist, ist diese Gleichung immer wahr (in Formeln: $\forall x \in \mathbb{R} : 0x = b$), also für alle $x \in \mathbb{R}$ erfüllt. Somit ist unsere Lösungsmenge gleich $\mathbb{R}$ und wir haben unendlich viele Lösungen. Falls dagegen $b \neq 0$ ist, ist die Gleichung niemals wahr (in Formeln: $\forall x \in \mathbb{R} : 0x \neq b$), die Lösungsmenge ist also die leere Menge $\emptyset$.

Wir sehen also: Je nachdem, wie a und b gewählt sind, hat die Lösungsmenge entweder 0, 1 oder unendlich viele Elemente. Dies wird uns auch bei linearen Gleichungssystemen in mehreren Variablen mit reellen Koeffizienten wieder begegnen.

Man kann diese Rechnung auch folgendermaßen verstehen: Wenn $a \in \mathbb{R}$ gegeben ist, betrachten wir die Funktion

$$\varphi : \mathbb{R} \to \mathbb{R}, \quad x \mapsto ax.$$

Wenn wir nun die Gleichung „$ax = b$“ lösen, dann suchen wir also alle x im Definitionsbereich von φ, die auf das gegebene b im Zielbereich abgebildet werden. Im Falle $a \neq 0$ ist $\varphi : \mathbb{R} \to \mathbb{R}$ bijektiv, es gibt also für jedes $b \in \mathbb{R}$ genau ein $x \in \mathbb{R}$ mit $\varphi(x) = b$. Deswegen brauchten wir im Fall $a \neq 0$ keine Fallunterscheidung für b.

Wenn aber $a = 0$ ist, dann ist φ nicht mehr surjektiv, also gibt es nicht mehr für alle $b \in \mathbb{R}$ eine Lösung. Außerdem ist φ dann auch nicht mehr injektiv, das heißt: Es kann mehrere Lösungen geben. Genau das haben wir oben gesehen.

Diese Sichtweise auf Gleichungen hat sich – sowohl in der Analysis wie auch in der (linearen) Algebra – als sehr sinnvoll erwiesen, da man Untersuchungen von Gleichungen zurückführen kann auf Untersuchungen von Funktionen.

Ausgehend von diesem eher langweiligen Beispiel gibt es nun mehrere Möglichkeiten, zu interessanteren Fragestellungen zu gelangen. Entweder man betrachtet *nicht-lineare* Gleichungen in einer Variablen, wie z.B. „$x^5 - 16x + 2 = 0$“ oder „$\cos(x) = x$“ oder „$e^x + x = 42$“. Diese werden sehr schnell sehr kompliziert und lassen sich oft nur numerisch (also näherungsweise) lösen[1].

Wir wollen in eine andere Richtung gehen: Wir bleiben bei linearen Gleichungen, aber wir erhöhen die Anzahl der Variablen:

Es sei $n \in \mathbb{N}$. Gegeben seien reelle Zahlen $a_1, \ldots, a_n \in \mathbb{R}$ sowie eine reelle Zahl $b \in \mathbb{R}$.

Eine *lineare Gleichung* ist eine Aussageform in den freien Variablen $x_1, \ldots, x_n$, die wie folgt aussieht:

$$a_1 x_1 + a_2 x_2 + \cdots + a_n x_n = b.$$

Dies lässt sich kompakter schreiben als

$$\sum_{j=1}^{n} a_j x_j = b.$$

Hierbei haben wir die *Summennotation* verwendet, die Sie schon in der Analysis, beziehungsweise HM1, gesehen haben.

Die Zahlen $a_1, \ldots, a_n$ nennen wir die *Koeffizienten* der linearen Gleichung.

Eine *Lösung* einer linearen Gleichung ist ein *Spaltenvektor*

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

mit $x_j \in \mathbb{R}$, sodass die Gleichung wahr wird.

Notation: Die Menge aller *Spaltenvektoren* mit n reellen Einträgen bezeichnen wir mit $\mathbb{R}^n$:

$$\mathbb{R}^n := \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \;\middle|\; \forall j \in \{1, \ldots n\} : x_j \in \mathbb{R} \right\}.$$

Die Lösungsmenge einer linearen Gleichung in n Unbekannten ist also immer eine Teilmenge der Menge $\mathbb{R}^n$.

Beispiel 2.1.2. Gegeben sei die lineare Gleichung

$$2x_1 - 3x_2 + 5x_3 = 7.$$

[1] Jede dieser drei Gleichungen hat mindestens eine reelle Lösung – doch wenn uns nicht alles täuscht, ist keine der Lösungen in einer geschlossenen Form darstellbar – aber Sie können es gerne trotzdem versuchen...

Wenn wir $x_1 = 1$, $x_2 = 0$ und $x_3 = 1$ setzen, sehen wir, dass $\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$ eine Lösung ist. Oft sind wir aber nicht nur an einer, sondern an allen Lösungen interessiert. In diesem Falle sind es aber unendlich viele und wir können sie nicht alle auflisten. Wir können allerdings die Gleichung nach x_1 auflösen:

$$x_1 = \frac{7}{2} + \frac{3}{2}x_2 - \frac{5}{2}x_3.$$

Diese Gleichung ist nun äquivalent zu der Ausgangsgleichung – hat also die gleiche Lösungsmenge. Hier können wir aber erkennen, dass zu jeder Wahl von $x_2 \in \mathbb{R}$ und $x_3 \in \mathbb{R}$ genau ein Wert von x_1 gehört.

Anders formuliert: Wenn wir $x_2 = s \in \mathbb{R}$ frei wählen und $x_3 = t \in \mathbb{R}$ frei wählen, können wir x_1 berechnen als $x_1 = \frac{7}{2} + \frac{3}{2}s - \frac{5}{2}t$. Die allgemeine Lösung der linearen Gleichung lautet also:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \frac{7}{2} + \frac{3}{2}s - \frac{5}{2}t \\ s \\ t \end{pmatrix},$$

wobei $s \in \mathbb{R}$ und $t \in \mathbb{R}$ frei wählbar sind. Somit können wir die Lösungsmenge angeben als:

$$\left\{ \begin{pmatrix} \frac{7}{2} + \frac{3}{2}s - \frac{5}{2}t \\ s \\ t \end{pmatrix} \middle| \, s, t \in \mathbb{R} \right\}.$$

Sobald wir die Konzepte *Basis* und *affiner Unterraum* eingeführt haben, werden wir die Lösungsmenge noch kompakter schreiben können.

Definition 2.1.3 (Lineares Gleichungssystem)**.** Ein *lineares Gleichungssystem (LGS) mit reellen Koeffizienten* ist gegeben durch eine Familie aus $m \in \mathbb{N}$ linearen Gleichungen, jede in $n \in \mathbb{N}$ Unbekannten. Wir notieren

$$\begin{aligned} a_{1,1}x_1 + \cdots + a_{1,n}x_n &= b_1 \\ a_{2,1}x_1 + \cdots + a_{2,n}x_n &= b_2 \\ &\vdots \\ a_{m,1}x_1 + \cdots + a_{m,n}x_n &= b_m. \end{aligned}$$

Eine Lösung x des linearen Gleichungssystems ist ein Spaltenvektor $x \in \mathbb{R}^n$, der gleichzeitig jede einzelne Gleichung löst[2].

Wenn alle rechten Seiten $b_j = 0$ sind, nennt man das LGS *homogen*, ansonsten nennt man es *inhomogen*.

[2] Die Lösungsmenge eines Gleichungssystems ist also die Schnittmenge der Lösungsmengen aller einzelnen Gleichungen.

Zusammenfassung von Abschnitt 2.1

(1) Ein Lineares Gleichungssystem (LGS) besteht aus endlich vielen linearen Gleichungen, die jeweils von mehreren Variablen abhängen können.

(2) Die Lösungsmenge eines LGS mit n Variablen ist eine Teilmenge von $\mathbb{R}^n$, der Menge aller Spaltenvektoren mit n reellen Einträgen.

2.2. Matrizenrechnung

Um Lineare Gleichungssysteme kompakter zu beschreiben, führen wir nun die Matrix-Schreibweise ein:

Definition 2.2.1 (Matrizen mit reellen Einträgen)**.** Gegeben seien $m, n \in \mathbb{N}$. Eine $(m \times n)$-*Matrix mit reellen Einträgen* ist ein rechteckiges Schema der Form

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} \end{pmatrix},$$

wobei die Einträge $a_{i,j}$ reelle Zahlen sind. Die Menge aller $(m \times n)$-Matrizen mit reellen Einträgen bezeichnen wir mit $\mathbb{R}^{m \times n}$.

Wir nennen eine Matrix *quadratisch*, wenn $m = n$ gilt.

Eine $(m \times 1)$-Matrix ist einfach ein Spaltenvektor, es gilt also $\mathbb{R}^{m \times 1} = \mathbb{R}^m$.

Eine $(1 \times n)$-Matrix nennen wir einen *Zeilenvektor*, die Menge aller Zeilenvektoren ist somit $\mathbb{R}^{1 \times n}$.

Eine (1×1)-Matrix ist einfach eine reelle Zahl, es gilt also:

$$\mathbb{R}^{1 \times 1} = \mathbb{R}^1 = \mathbb{R}.$$

Definition 2.2.2 (Transponierte)**.** Für eine $(m \times n)$-Matrix

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} \end{pmatrix} \in \mathbb{R}^{m \times n}$$

nennen wir die Matrix

$$A^\top = \begin{pmatrix} a_{1,1} & a_{2,1} & \dots & a_{m,1} \\ a_{1,2} & a_{2,2} & \dots & a_{m,2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1,n} & a_{2,n} & \dots & a_{m,n} \end{pmatrix} \in \mathbb{R}^{n \times m}$$

die *Transponierte* von A. Die Transponierte eines Spaltenvektors ist ein Zeilenvektor und umgekehrt.

Ziel ist es nun, ein LGS (Definition 2.1.3) in eine Form zu bringen, die zumindest rein optisch so einfach aussieht wie in Beispiel 2.1.1. Dazu ordnen wir einem LGS mit m Gleichungen in n Unbekannten

$$\begin{aligned} a_{1,1}x_1 + \cdots + a_{1,n}x_n &= b_1 \\ a_{2,1}x_1 + \cdots + a_{2,n}x_n &= b_2 \\ &\vdots \\ a_{m,1}x_1 + \cdots + a_{m,n}x_n &= b_m. \end{aligned}$$

die *Koeffizientenmatrix*

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} \end{pmatrix} \in \mathbb{R}^{m\times n}$$

und den Spaltenvektor

$$b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \in \mathbb{R}^m$$

zu. Dann würden wir das LGS gerne kurz schreiben als

$$Ax = b.$$

Um diese Schreibweise zu rechtfertigen, müssten wir definieren, wie man eine Matrix mit einem Spaltenvektor multipliziert. Weil wir diese Definition im Laufe dieser Veranstaltung sowieso noch oft benötigen, werden wir nun ganz allgemein Addition und Multiplikation von Matrizen definieren, sodass das gesuchte Matrix-Spaltenvektor-Produkt nur ein Spezialfall davon ist.

Definition 2.2.3 (Addition und skalares Vielfaches von Matrizen)**.**

(a) Für zwei Matrizen

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} \end{pmatrix} \in \mathbb{R}^{m\times n} \quad \text{und} \quad B = \begin{pmatrix} b_{1,1} & b_{1,2} & \dots & b_{1,n} \\ b_{2,1} & b_{2,2} & \dots & b_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m,1} & b_{m,2} & \dots & b_{m,n} \end{pmatrix} \in \mathbb{R}^{m\times n}$$

mit der gleichen Zeilen- und Spaltenanzahl definieren wir die *Summe* als

$$A + B := \begin{pmatrix} a_{1,1}+b_{1,1} & a_{1,2}+b_{1,2} & \dots & a_{1,n}+b_{1,n} \\ a_{2,1}+b_{2,1} & a_{2,2}+b_{2,2} & \dots & a_{2,n}+b_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1}+b_{m,1} & a_{m,2}+b_{m,2} & \dots & a_{m,n}+b_{m,n} \end{pmatrix} \in \mathbb{R}^{m\times n}.$$

(b) Für eine Matrix

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} \end{pmatrix} \in \mathbb{R}^{m\times n} \quad \text{und eine reelle Zahl} \quad \lambda \in \mathbb{R}$$

definieren wir das *skalare Vielfache* (oder die *Skalierung*) als

$$\lambda \cdot A := \lambda A := \begin{pmatrix} \lambda a_{1,1} & \lambda a_{1,2} & \dots & \lambda a_{1,n} \\ \lambda a_{2,1} & \lambda a_{2,2} & \dots & \lambda a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda a_{m,1} & \lambda a_{m,2} & \dots & \lambda a_{m,n} \end{pmatrix} \in \mathbb{R}^{m\times n}.$$

Matrizen kann man also nur addieren, wenn sie das gleiche Format (also die gleiche Anzahl sowohl an Zeilen als auch Spalten) haben. Die Summe der Matrizen hat dann dasselbe Format wie die beiden Summanden. Die Einträge der Summe sind einfach die Summen der Einträge der Summandenmatrizen.

Bemerkung 2.2.4 (Addition von Spaltenvektoren)**.** Da für uns ein Spaltenvektor $v \in \mathbb{R}^n = \mathbb{R}^{n\times 1}$ einfach nur eine Matrix mit n Zeilen und einer Spalte ist, haben wir mit Definition 2.2.3 auch gleich definiert, wie man Spaltenvektoren (mit der gleichen Zeilenanzahl) addiert und mit reellen Zahl skaliert. Im Falle von Spaltenvektoren im Raum $\mathbb{R}^2$ bzw. $\mathbb{R}^3$ entspricht diese Addition der Vektoraddition, die vielfache Anwendungen in Natur- und Ingenieurwissenschaft sowie in der analytischen Geometrie hat.

Bemerkung 2.2.5. Da reelle Zahlen verwendet werden können, um Spaltenvektoren und Matrizen *skalar zu multiplizieren*, werden sie in diesem Kontext auch *Skalare* genannt. Später[3] – wenn wir formal eingeführt haben, was ein *Vektorraum* ist –, werden wir auch den Begriff des Skalars nochmals genauer definieren.

Definition 2.2.6 (Nullmatrix)**.** Die $(m \times n)$-Matrix, deren Einträge alle Null sind, heißt die *Nullmatrix* und wird mit $\mathbf{0}_{m\times n}$ bezeichnet. Wenn das Format der Matrix aus dem Kontext klar ist, schreiben wir auch nur $\mathbf{0}$.

Im Spezialfall $(n \times 1)$ nennen wir den Spaltenvektor, der nur aus Nullen besteht, auch den *Nullvektor*: $0 := 0_n := \mathbf{0}_{n\times 1}$.

Lemma 2.2.7 (Eigenschaften der Matrixaddition und -skalierung)**.** *Gegeben seien natürliche Zahlen $m,n \in \mathbb{N}$ sowie Matrizen $A,B,C \in \mathbb{R}^{m\times n}$ und reelle Zahlen (Skalare) $\mu,\lambda \in \mathbb{R}$. Dann gilt*

(i) $(A+B)+C = A+(B+C)$

(ii) $A+\mathbf{0}_{m\times n} = \mathbf{0}_{m\times n}+A = A$

(iii) $A+(-1)A = (-1)A+A = \mathbf{0}_{m\times n}$

(iv) $A+B = B+A$

(v) $\lambda(A+B) = \lambda A+\lambda B$

[3] siehe Definition 4.1.1

(vi) $(\lambda+\mu)A = \lambda A + \mu A$

(vii) $(\lambda\cdot\mu)A = \lambda(\mu A)$

(viii) $1A = A$.

Bemerkung 2.2.8.

(a) Beachten Sie, dass in Lemma 2.2.7 das Symbol „+“ sowohl Addition von reellen Zahlen als auch Addition von Matrizen bedeutet – je nachdem, welche Objekte addiert werden.

(b) Es gilt „Punkt- vor Strichrechnung“, d.h. zum Beispiel bedeutet $\lambda A + \lambda B$ stets $(\lambda A)+(\lambda B)$.

Man beachte, dass wir bis jetzt noch nicht definiert haben, wie man zwei Matrizen miteinander multipliziert. Dies wollen wir jetzt nachholen:

Definition 2.2.9 (Matrixprodukt)**.** Zwei Matrizen $A \in \mathbb{R}^{m\times n}$ und $B \in \mathbb{R}^{q\times p}$ heißen *multiplizierbar* (oder *kompatibel*), wenn die Anzahl der Spalten von A gleich der Anzahl der Zeilen von B ist, wenn also $n = q$ gilt. Gegeben seien zwei multiplizierbare Matrizen

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} \end{pmatrix} \in \mathbb{R}^{m\times n} \quad \text{und} \quad B = \begin{pmatrix} b_{1,1} & b_{1,2} & \dots & b_{1,p} \\ b_{2,1} & b_{2,2} & \dots & b_{2,p} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n,1} & b_{n,2} & \dots & b_{n,p} \end{pmatrix} \in \mathbb{R}^{n\times p}.$$

Dann definieren wir das *Matrixprodukt* als

$$A\cdot B := AB := \begin{pmatrix} c_{1,1} & c_{1,2} & \dots & c_{1,p} \\ c_{2,1} & c_{2,2} & \dots & c_{2,p} \\ \vdots & \vdots & \ddots & \vdots \\ c_{m,1} & c_{m,2} & \dots & c_{m,p} \end{pmatrix} \in \mathbb{R}^{m\times p},$$

wobei jeder Eintrag $c_{i,k}$ definiert ist als

$$c_{i_k} := a_{i,1}b_{1,k} + \cdots a_{i,n}b_{n,k} = \sum_{j=1}^{n} a_{i,j}b_{j,k}.$$

Der Eintrag in Zeile i und Spalte k der Produktmatrix AB wird also berechnet mit Hilfe der i-ten Zeile der linken Matrix A und der k-ten Spalte der rechten Matrix B.

Beispiel 2.2.10. Die Matrizen

$$A = \begin{pmatrix} 1 & 2 \\ 0 & 0 \\ 3 & 4 \end{pmatrix} \quad \text{und} \quad B = \begin{pmatrix} 42 & 0 & 23 & 0 \\ -2 & 3 & 0 & 0 \end{pmatrix}$$

sind multiplizierbar, weil die Anzahl der Spalten von A gleich der Anzahl der Zeilen von B ist (nämlich 2). Also ist das Produkt AB definiert. Das Produkt hat 3 Zeilen und 4 Spalten, weil A genau 3 Zeilen und B genau 4 Spalten hat. Die Einträge lauten nun:

$$AB = \begin{pmatrix} 1 & 2 \\ 0 & 0 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 42 & 0 & 23 & 0 \\ -2 & 3 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 38 & 6 & 23 & 0 \\ 0 & 0 & 0 & 0 \\ 118 & 12 & 69 & 0 \end{pmatrix}.$$

Hierbei fällt auf: Wenn die *erste* Matrix eine *Nullzeile* hat (d.h., in dieser Zeile ist jeder einzelne Eintrag 0), so hat das Produkt ebenfalls eine Nullzeile. Ebenso gilt: Wenn die *zweite* Matrix eine *Nullspalte* hat, so hat die Produktmatrix ebenfalls eine Nullspalte. Eine Möglichkeit, die Rechnung schematisch darzustellen, ist hier durch die folgende Tabelle gegeben:

		42	0	23	0
		−2	3	0	0
1	2	38	6	23	0
0	0	0	0	0	0
3	4	118	12	69	0

Das Multiplizieren von Matrizen sollte man üben, bis man es im Schlaf beherrscht.

Definition 2.2.11 (Einheitsmatrix)**.** Für eine natürliche Zahl $n \in \mathbb{N}$ wird die Matrix

$$\mathbb{1}_n := \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} \in \mathbb{R}^{n \times n}$$

mit Einsen auf der Diagonale als *Einheitsmatrix* bezeichnet. Wenn die Zahl n aus dem Zusammenhang klar ist, schreiben wir auch nur $\mathbb{1}$.

Lemma 2.2.12 (Rechenregeln für Matrixprodukte)**.** *Gegeben seien natürliche Zahlen $m, n, p, r \in \mathbb{N}$. Dann gilt:*

(i)	$\forall A \in \mathbb{R}^{m \times n}, B \in \mathbb{R}^{n \times p}, C \in \mathbb{R}^{p \times r}:$	$(AB)C = A(BC).$	*(Assoziativität)*
(ii)	$\forall A \in \mathbb{R}^{m \times n}:$	$\mathbb{1}_m A = A \mathbb{1}_n = A.$	*(Neutralelement)*
(iii)	$\forall A_1, A_2 \in \mathbb{R}^{m \times n}, B \in \mathbb{R}^{n \times p}:$	$(A_1 + A_2)B = A_1 B + A_2 B.$	*(Distributivgesetz I)*
(iv)	$\forall A \in \mathbb{R}^{m \times n}, B_1, B_2 \in \mathbb{R}^{n \times p}:$	$A(B_1 + B_2) = AB_1 + AB_2.$	*(Distributivgesetz II)*
(v)	$\forall A \in \mathbb{R}^{m \times n}, B \in \mathbb{R}^{n \times p}, \lambda \in \mathbb{R}:$	$(\lambda A)B = A(\lambda B) = \lambda(AB).$	*(Bilinearität)*

Wir möchten nun noch auflisten, wie sich Summen und Produkte von Matrizen mit der in Definition 2.2.2 eingeführten Transponierten vertragen:

Lemma 2.2.13 (Rechenregeln für die transponierte Matrix)**.** *Gegeben seien natürliche Zahlen $m, n, p \in \mathbb{N}$. Dann gilt:*

(i)	$\forall A, B \in \mathbb{R}^{m \times n}:$	$(A + B)^\top = A^\top + B^\top.$	*(Transponieren und Summieren)*
(ii)	$\forall A \in \mathbb{R}^{m \times n}, \lambda \in \mathbb{R}:$	$(\lambda A)^\top = \lambda A^\top.$	*(Transponieren und Skalieren)*
(iii)	$\forall A \in \mathbb{R}^{m \times n}, B \in \mathbb{R}^{n \times p}:$	$(AB)^\top = B^\top A^\top.$	*(Transponieren und Produkte)*
(iv)		$(\mathbf{0}_{m \times n})^\top = \mathbf{0}_{n \times m}.$	*(Transponierte der Nullmatrix)*
(iv)		$(\mathbb{1}_n)^\top = \mathbb{1}_n.$	*(Transponierte der Einheitsmatrix)*
(v)	$\forall A \in \mathbb{R}^{m \times n}:$	$(A^\top)^\top = A.$	*(Transponierte der Transponierten)*

Mit Hilfe der Matrixmultiplikation können wir nun eine äquivalente Definition eines linearen Gleichungssystems mit reellen Koeffizienten (Definition 2.1.3) geben:

Definition 2.2.14 (Lineares Gleichungssystem mit reellen Koeffizienten)**.** Gegeben seien natürliche Zahlen $m, n \in \mathbb{N}$. Ein *lineares Gleichungssystem (LGS) mit reellen Koeffizienten* mit Koeffizientenmatrix $A \in \mathbb{R}^{m \times n}$ und rechter Seite $b \in \mathbb{R}^m$ ist eine Aussageform:

$$Ax = b$$

mit der freien Variable $x \in \mathbb{R}^n$. Die Lösungsmenge des Gleichungssystems

$$\{x \in \mathbb{R}^n \mid Ax = b\} \subseteq \mathbb{R}^n$$

ist eine Teilmenge der Menge $\mathbb{R}^n$. Wenn der Spaltenvektor $b = 0$ der Nullvektor ist, dann nennt man das LGS *homogen*, ansonsten nennt man es *inhomogen*.

Bemerkung 2.2.15 (Multiplikation mit der Nullmatrix)**.** Eine wichtige Eigenschaft der Matrixmultiplikation wurde bisher noch nicht angesprochen: Gegeben seien zwei Matrizen $A \in \mathbb{R}^{m\times n}$ und $B \in \mathbb{R}^{n\times p}$. Dann gilt:

$$\left(A = \mathbf{0}_{m\times n} \text{ oder } B = \mathbf{0}_{n\times p}\right) \Longrightarrow \left(AB = \mathbf{0}_{m\times p}\right).$$

Dies folgt unmittelbar aus der Definition des Matrizenproduktes. Achtung: Die Rückimplikation „$\Leftarrow$“ gilt nicht! Es gibt also Matrizen A, B die beide nicht die Nullmatrix sind, aber ihr Produkt ist die Nullmatrix (man sagt auch: Der Matrizenring ist nicht *nullteilerfrei*). Dies ist ein weiterer großer Unterschied zwischen dem Rechnen mit Matrizen und dem Rechnen mit reellen Zahlen.

Zusammenfassung von Abschnitt 2.2

(1) Eine Matrix $A \in \mathbb{R}^{m\times n}$ ist ein rechteckig angeordnetes Schema von reellen Zahlen.

(2) Matrizen können mit reellen Zahlen skaliert, addiert und ggf. auch multipliziert werden, wenn die Abmessungen verträglich sind.

(3) Spaltenvektoren $v \in \mathbb{R}^n = \mathbb{R}^{n\times 1}$ sind dasselbe wie $(n \times 1)$-Matrizen.

(4) Matrixmultiplikation ist assoziativ, aber im Allgemeinen nicht kommutativ.

2.3. Homogene lineare Gleichungssysteme und Untervektorräume von $\mathbb{R}^n$

Wir möchten nun speziell *homogene* lineare Gleichungssysteme untersuchen. Es sei dazu $A \in \mathbb{R}^{m\times n}$ eine Matrix. Dann bezeichnen wir die Lösungsmenge des homogenen linearen Gleichungssystems

$$Ax = 0$$

auch als den *Kern* der Matrix A und schreiben:

$$\ker A := \{x \in \mathbb{R}^n \mid Ax = 0\} \subseteq \mathbb{R}^n.$$

Der Kern entspricht also der *Nullstellenmenge* der Abbildung

$$\varphi : \mathbb{R}^n \to \mathbb{R}^m, \quad x \mapsto Ax.$$

Wir werden nun zeigen, dass die Lösungsmenge eines solchen homogenen LGS immer eine bestimmte Form hat:

Definition 2.3.1 (Linearer Unterraum / Untervektorraum)**.**
Es sei $n \in \mathbb{N}$. Eine Teilmenge $U \subseteq \mathbb{R}^n$ heißt *Untervektorraum* oder *linearer Unterraum* von $\mathbb{R}^n$, wenn folgende drei Bedingungen erfüllt sind:

(i)		$0 \in U$	(U enthält den Nullvektor)
(ii)	$\forall v, w \in U:$	$v + w \in U$	(U ist abgeschlossen unter Addition)
(iii)	$\forall v \in U, \lambda \in \mathbb{R}:$	$\lambda v \in U$	(U ist abgeschlossen unter Multiplikation mit Skalaren)

Lemma 2.3.2. *Gegeben sei eine Matrix $A \in \mathbb{R}^{m \times n}$, mit $m, n \in \mathbb{N}$. Dann ist der Kern*

$$\ker A := \left\{ x \in \mathbb{R}^n \;\middle|\; Ax = 0 \right\} \subseteq \mathbb{R}^n$$

immer ein Untervektorraum von $\mathbb{R}^n$.

Insbesondere haben wir damit gesehen, dass der Kern einer Matrix niemals leer ist. Ein homogenes LGS hat also stets mindestens eine Lösung, nämlich den Nullvektor.

Später[4] werden wir sehen, dass jeder Untervektorraum von $\mathbb{R}^n$ immer auch als Lösungsmenge eines homogenen LGS beschrieben werden kann.

Beweis von Lemma 2.3.2. Wir müssen zeigen, dass die Lösungsmenge des homogenen LGS

$$Ax = 0$$

die Eigenschaften (i),(ii) und (iii) aus Definition 2.3.1 erfüllt:

(i)

Es gilt $A0 = 0$ (nach Bemerkung 2.2.15). Also ist $0 \in \ker A$.

(ii)

Gegeben seien $v, w \in \ker A$, d.h. v und w sind Spaltenvektoren mit $Av = 0$ und $Aw = 0$. Nun können wir berechnen:

$$A(v + w) = Av + Aw = 0 + 0 = 0,$$

und zwar durch entsprechende Anwendung eines Matrixdistributivgesetzes (siehe Lemma 2.2.12 (iv)), der Tatsache, dass $v, w \in \ker A$ sind und der Rechenregel, dass $0 + 0 = 0$ (siehe Lemma 2.2.7(ii)) ist.

(iii)

Gegeben sei ein $v \in \ker A$ und eine reelle Zahl $\lambda \in \mathbb{R}$. Dann gilt:

$$A(\lambda v) = \lambda A v = \lambda 0 = 0.$$

Hier haben wir die Bilinearität der Matrixmultiplikation (Lemma 2.2.12 (v)) ausgenutzt. □

Definition 2.3.3 (Lineare Hülle)**.** Es sei $n \in \mathbb{N}$.

(a) Gegeben seien Vektoren $v_1, \ldots, v_r \in \mathbb{R}^n$ mit $r \in \mathbb{N}_0$. Eine *Linearkombination* der Vektoren $v_1, \ldots, v_r$ ist eine Summe der Form

$$\sum_{j=1}^{r} \lambda_j v_j = \lambda_1 v_1 + \cdots + \lambda_r v_r,$$

wobei die Skalare $\lambda_j \in \mathbb{R}$ beliebig gewählt sein dürfen.

[4] siehe Bemerkung 4.5.3

(b) Gegeben sei eine (endliche oder unendliche) Teilmenge $M \subseteq \mathbb{R}^n$. Wir definieren die *lineare Hülle* der Menge M als die Menge aller Linearkombinationen von Elementen aus M und schreiben:

$$\mathrm{LH}_{\mathbb{R}}(M) := \mathrm{LH}(M) := \left\{ \sum_{j=1}^{r} \lambda_j v_j \;\middle|\; r \in \mathbb{N}_0; \forall j \leq r : v_j \in M; \lambda_j \in \mathbb{R} \right\} \subseteq \mathbb{R}^n.$$

Eine leere Summe ist immer 0, deswegen ist $\mathrm{LH}(\emptyset) = \{0\}$.

Falls $M = \{v_1, v_2, \ldots, v_r\}$ eine endliche Menge ist, schreiben wir auch

$$\mathrm{LH}(v_1, v_2 \ldots, v_r) := \mathrm{LH}(\{v_1, v_2 \ldots, v_r\})$$

und es gilt:

$$\mathrm{LH}(v_1, v_2 \ldots, v_r) = \left\{ \lambda_1 v_1 + \cdots + \lambda_r v_r \;\middle|\; \lambda_j \in \mathbb{R} \right\}.$$

Neben dem Begriff *lineare Hülle* sind auch die Bezeichnungen *linearer Spann* von M, *linearer Aufspann* von M, oder *von M erzeugter Untervektorraum* gebräuchlich. Letztere Bezeichnung wird durch Lemma 2.3.4 gerechtfertigt.

Lemma 2.3.4. *Es sei $M \subseteq \mathbb{R}^n$ eine Teilmenge. Dann gilt*

(a) Die Menge $\mathrm{LH}(M)$ ist ein Untervektorraum von $\mathbb{R}^n$ und es gilt $M \subseteq \mathrm{LH}(M)$.

(b) Jeder andere Untervektorraum $U \subseteq \mathbb{R}^n$ mit $M \subseteq U$ enthält $\mathrm{LH}(M)$.

Man kann also sagen: $\mathrm{LH}(M)$ ist der kleinste Untervektorraum von $\mathbb{R}^n$, der M enthält.

Beweis.
(a)
Nach Definition 2.3.1 müssen wir nachrechnen, dass $0 \in \mathrm{LH}(M)$, dass M abgeschlossen unter Addition ist, und dass M abgeschlossen unter Multiplikation mit Skalaren aus $\mathbb{R}$ ist.

In Definition 2.3.3 haben wir die Linearkombination von $r = 0$ Vektoren erlaubt, somit ist $0 \in \mathrm{LH}(M)$.

Es seien $v, w \in \mathrm{LH}(M)$ gegeben. Wir müssen zeigen, dass $v + w \in \mathrm{LH}(M)$ ist. Da $v \in \mathrm{LH}(M)$ ist, gilt $v = \sum_{j=1}^{r} \lambda_j v_j$ mit $\lambda_j \in \mathbb{R}$ und $v_j \in M$. Da $w \in \mathrm{LH}(M)$ ebenfalls ist, gilt $w = \sum_{k=1}^{s} \mu_k w_k$ mit $\mu_k \in \mathbb{R}$ und $w_k \in M$.

Also gilt für die Summe:

$$v + w = \sum_{j=1}^{r} \lambda_j v_j + \sum_{k=1}^{s} \mu_k w_k = \lambda_1 v_1 + \cdots \lambda_j v_j + \mu_1 w_1 + \cdots + \mu_k w_k.$$

Dies ist eine Linearkombination aus Vektoren aus M mit Skalaren aus $\mathbb{R}$, also gilt $v + w \in \mathrm{LH}(M)$. Das war zu zeigen.

Es sei nun $v \in \mathrm{LH}(M)$ gegeben und $\mu \in \mathbb{R}$. Es ist zu zeigen, dass $\mu v \in \mathrm{LH}(M)$ ist. Da $v \in \mathrm{LH}(M)$ ist, gilt $v = \sum_{j=1}^{r} \lambda_j v_j$ mit $\lambda_j \in \mathbb{R}$ und $v_j \in M$. Also gilt:

$$\mu v = \mu \sum_{j=1}^{r} \lambda_j v_j = \sum_{j=1}^{r} (\mu \lambda_j) v_j$$

und dies ist auch in $\mathrm{LH}(M)$.

Also ist $\mathrm{LH}(M)$ ein Untervektorraum von $\mathbb{R}^n$.

Es bleibt zu zeigen, dass $M \subseteq \mathrm{LH}(M)$ gilt, aber dies folgt sofort, denn jedes $v \in M$ lässt sich schreiben als $v = 1v$ und ist somit eine Linearkombination von Elementen aus M und somit ist $v \in \mathrm{LH}(M)$.

(b)

Es sei nun $U \subseteq \mathbb{R}^n$ ein beliebiger Untervektorraum mit der Eigenschaft, dass $M \subseteq \mathrm{LH}(M)$. Wir müssen nun zeigen, dass $\mathrm{LH}(M) \subseteq U$.

Sei dazu $v \in \mathrm{LH}(M)$ gegeben. Dann ist $v = \lambda_1 v_1 + \cdots \lambda_r v_r$ für Skalare $\lambda_j \in \mathbb{R}$ und Vektoren $v_r \in M$. Da $M \subseteq U$ gilt, bedeutet dies, dass jedes v_j in U liegen. Nun ist U aber ein Untervektorraum und somit ist auch $\lambda_j v_r$ in U. Und da Untervektorräume auch unter Addition abgeschlossen sind, folgt somit

$$v = \sum_{j=1}^{r} \lambda_j v_j \in U.$$

Das war zu zeigen. □

Definition 2.3.5 (Erzeugendensystem)**.** Es sei $U \subseteq \mathbb{R}^n$ ein Untervektorraum. Dann heißt eine Teilmenge $M \subseteq U$ ein *Erzeugendensystem* für U, falls

$$\mathrm{LH}(M) = U,$$

wenn sich also jedes Element in U als Linearkombination von Elementen aus M schreiben lässt.

Bemerkung 2.3.6. Wenn zwei Teilmengen $M_1 \subseteq M_2 \subseteq U$ gegeben sind, und M_1 ein Erzeugendensystem ist, dann ist M_2 auch ein Erzeugendensystem. Insbesondere ist also der ganze Raum U immer ein Erzeugendensystem. Wir werden allerdings im Folgenden versuchen, zu gegebenem Untervektorraum $U \subseteq \mathbb{R}^n$ möglichst „kleine“ Erzeugendensysteme zu finden.

Wenn wir also ein homogenes LGS $Ax = b$ vorliegen haben, dann wissen wir nun, dass die Lösungsmenge $\ker A$ ein Untervektorraum ist und somit mit einem geeigneten Erzeugendensystem beschrieben werden kann. Bevor wir nun genauer darauf eingehen, wie man ein möglichst kleines Erzeugendensystem findet, wollen wir der Frage nachgehen, ob die Lösungsmenge überhaupt Vektoren enthält – außer dem Nullvektor, der immer im Kern enthalten ist. Wir werden nun ein hinreichendes Kriterium dafür finden, dass es mindestens eine weitere Lösung gibt.

Satz 2.3.7. *Es sei $A \in \mathbb{R}^{m \times n}$ eine Matrix mit weniger Zeilen als Spalten, d.h. $m < n$. Dann ist $\ker A \neq \{0\}$, d.h. es gibt mindestens eine nichttriviale Lösung des homogenen LGS $Ax = 0$.*

Beweis. Wir nutzen für diesen Beweis das Prinzip der *vollständigen Induktion*, das Sie aus der Vorlesung Analysis bzw. HM1 kennen.

Wir benutzen Induktion für die Anzahl der Zeilen, d.h. die Anzahl der Gleichungen.

Induktionsanfang $m = 1$**:**

In diesem Fall haben wir nur eine Gleichung:

$$a_{1,1}x_1 + \cdots + a_{1,n}x_n = 0.$$

Es gilt nun zwei Fälle zu unterscheiden:

Fall 1:$a_{1,1} = 0$

In diesem Fall ist der Vektor

$$x := \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in \mathbb{R}^n$$

eine Lösung des LGS und nicht der Nullvektor, also gilt $x \in \ker A$.

Fall 2:$a_{1,1} \neq 0$

Nach Voraussetzung ist $n > m = 1$, also ist $n \geq 2$, es gibt also mindestens zwei Unbekannte in unserem LGS. Setzen wir nun $x_1 := -a_{1,2}$, $x_2 := a_{1,1}$ und alle anderen Variablen auf 0, so erhalten wir eine Lösung

$$x := \begin{pmatrix} -a_{1,2} \\ a_{1,1} \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in \mathbb{R}^n.$$

Da die zweite Komponente des Vektors nicht 0 ist, ist dies nicht der Nullvektor.

Induktionsschritt:

Es sei $m \in \mathbb{N}$ so gewählt, dass die zu zeigende Aussage für alle Matrizen mit m Zeilen gilt. Es bleibt zu zeigen, dass die Aussage auch für eine Matrix A mit genau $(m+1)$ Zeilen gilt. Wieder dürfen wir annehmen, dass die Spaltenanzahl $n > m+1$ ist. Schauen wir nun die letzte Spalte der Matrix A an. Wenn diese Spalte der Nullvektor ist, dann ist

$$x := \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \in \mathbb{R}^n$$

eine Lösung des LGS, die nicht der Nullvektor ist.

Nehmen wir nun also an, dass es ein $i \in \{1, \ldots, m, m+1\}$ gibt mit $a_{i,n} \neq 0$. Dann können wir die i-te Gleichung unseres LGS

$$a_{i,1}x_1 + \cdots + a_{i,n}x_n = 0$$

nach x_n auflösen:

$$x_n = -\frac{1}{a_{i,n}} \left(a_{i,1}x_1 + \cdots + a_{i,n}x_{n-1} \right)$$

Wenn wir dies in die übrigen Gleichungen einsetzen, so erhalten wir ein Gleichungssystem mit m Gleichungen und $n-1$ Unbekannten. Nach Induktionsvoraussetzung hat dieses Gleichungssystem nun eine Lösung

$$\begin{pmatrix} x_1 \\ \vdots \\ x_{n-1} \end{pmatrix} \in \mathbb{R}^{n-1},$$

die nicht nur aus Nullen besteht. Wenn wir nun unseren Wert für x_n (der durchaus 0 sein darf) unten dranhängen, ergibt dies eine Lösung des Ausgangssystems

$$\begin{pmatrix} x_1 \\ \vdots \\ x_{n-1} \\ -\frac{1}{a_{i,n}} \left(a_{i,1}x_1 + \cdots + a_{i,n}x_{n-1} \right) \end{pmatrix} \in \mathbb{R}^n,$$

die nicht der Nullvektor ist. Das beendet den Beweis. □

Beispiel 2.3.8. Gegeben seien Vektoren $v_1, \dots v_r \in \mathbb{R}^n$. Dann ist die *triviale Linearkombination* die, bei der alle Skalare als 0 gewählt werden. Das Ergebnis ist dann immer der Nullvektor. Es kann aber auch vorkommen, dass eine nichttriviale Linearkombination den Nullvektor ergibt, wie z.B. bei den Vektoren

$$v_1 := \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}; \quad v_2 := \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}; \quad v_3 := \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}; \quad v_4 := \begin{pmatrix} 42 \\ -65 \\ 23 \end{pmatrix}.$$

Hier ist $1 \cdot v_1 + (-1) \cdot v_2 + (-1)v_3 + 0 \cdot v_4 = 0$, aber nicht alle Skalare sind 0. Es gibt also eine *Darstellung des Nullvektors als nichttriviale Linearkombination*.

Definition 2.3.9 (Lineare Unabhängigkeit)**.**
Gegeben sei eine endliche Menge $M = \{v_1, \dots, v_r\} \subseteq \mathbb{R}^n$ mit $r \in \mathbb{N}_0$ paarweise verschiedenen Vektoren $v_1, \dots, v_r$.

(a) Wir sagen $M = \{v_1, \dots, v_r\}$ ist *linear unabhängig*, wenn die einzige Darstellung des Nullvektors als Linearkombination aus M die triviale Linearkombination (alle Skalare null) ist. In Formeln:

$$M \text{ ist linear unabhängig } :\Longleftrightarrow \left(\forall \lambda_1, \dots, \lambda_r \in \mathbb{R} : \left(\sum_{j=1}^{r} \lambda_j v_j = 0 \implies (\forall j : \lambda_j = 0) \right) \right).$$

(b) Wir sagen $M = \{v_1, \dots, v_r\}$ ist *linear abhängig*, wenn M nicht linear unabhängig ist, d.h. wenn es (mindestens) eine Darstellung des Nullvektors als nichttriviale Linearkombination gibt. In Formeln:

$$M \text{ ist linear abhängig } :\Longleftrightarrow \left(\exists \lambda_1, \dots, \lambda_r \in \mathbb{R} : \left(\sum_{j=1}^{r} \lambda_j v_j = 0 \text{ und } (\exists j : \lambda_j \neq 0) \right) \right).$$

Unter Verwendung der De Morganschen Regeln für Quantoren (siehe Bemerkung 1.2.8) und entsprechenden Umformungsregeln (siehe Satz 1.1.6) können Sie sich leicht davon überzeugen, dass die Formel für lineare Abhängigkeit wirklich die Negation der linearen Unabhängigkeit ist.

Beachten Sie, dass Definition 2.3.9 nur für endliche Mengen funktioniert. Wir werden später[5] eine verallgemeinerte Definition der linearen Unabhängigkeit kennenlernen, die auch für unendliche Mengen einen Sinn ergibt.

Beispiel 2.3.10. (a) Die leere Menge $\emptyset$ ist linear unabhängig, weil es nicht möglich ist, aus Vektoren der leeren Menge eine nichttriviale Linearkombination zu finden.

(b) Es sei $M = \{v\}$ einelementig. Dann ist M linear unabhängig, falls $v \neq 0$ und linear abhängig, falls $v = 0$.

(c) Es sei $M = \{v, w\}$ zweielementig. Dann ist M linear abhängig, falls v ein Vielfaches von w ist oder umgekehrt – ansonsten ist M linear unabhängig:

$$\begin{aligned} \{v, w\} \text{ ist linear abhängig } &\Longleftrightarrow ((\exists \lambda \in \mathbb{R} : w = \lambda v) \text{ oder } (\exists \lambda \in \mathbb{R} : v = \lambda w)) \\ &\Longleftrightarrow ((v = 0) \text{ oder } (w = 0) \text{ oder } (\exists \lambda \in \mathbb{R} \setminus \{0\} : v = \lambda w)). \end{aligned}$$

[5] siehe Definition 4.2.3

(d) Die vier Vektoren in Beispiel 2.3.8 sind linear abhängig, obwohl je zwei von ihnen linear unabhängig sind.

Lemma 2.3.11. *Für eine endliche Teilmenge $M \subseteq \mathbb{R}^n$ sind die folgenden Aussagen äquivalent:*

(i) Die Menge M ist linear unabhängig.

(ii) Jeder Vektor $v \in \mathrm{LH}(M)$ hat eine eindeutige *Darstellung als Linearkombination aus Vektoren in M.*

(iii) Für jeden Vektor $v \in M$ ist $v \notin \mathrm{LH}(M \setminus \{v\})$.

Beweis. Es sei $M = \{v_1, \ldots, v_r\} \subseteq \mathbb{R}^n$ gegeben. Wir beweisen drei Implikationen, aus denen dann die Äquivalenz aller drei Aussagen folgt:
„(i) $\Longrightarrow$ (ii)“:
Wir nehmen an, M sei linear unabhängig. Es sei $v \in \mathrm{LH}(M)$. Dann ist v eine Linearkombination von Vektoren aus M:

$$v = \sum_{j=1}^{r} \lambda_j v_j.$$

Nehmen wir an, es gäbe noch eine weitere Darstellung:

$$v = \sum_{j=1}^{r} \mu_j v_j.$$

Dann können wir diese beiden Gleichungen voneinander subtrahieren und erhalten:

$$0 = \sum_{j=1}^{r} (\lambda_j - \mu_j) v_j.$$

Dies ergibt also eine Darstellung des Nullvektors als Linearkombination von Vektoren aus M. Da M aber nach Voraussetzung (i) linear unabhängig ist, ist die einzige solche Linearkombination die triviale, d.h. alle Skalare müssen 0 sein:

$$\forall j \in \{1, \ldots, r\} : \ \lambda_j - \mu_j = 0.$$

Also gilt $\lambda_j = \mu_j$ für alle j und damit ist gezeigt, dass v nur eine eindeutige Darstellung als Linearkombination aus M hat.
„(ii) $\Longrightarrow$ (iii)“:
Nehmen wir für diese Implikation an, dass jeder Vektor in $\mathrm{LH}(M)$ eine eindeutige Darstellung als Linearkombination hat. Es sei $v = v_k \in M = \{v_1, \ldots, v_r\}$ gegeben. Zu zeigen ist, dass sich v_k nicht als Linearkombination der übrigen Vektoren schreiben lässt. Nehmen wir deshalb per Widerspruch an, dass das möglich wäre, dass es also Skalare $\lambda_1, \ldots, \lambda_{k-1}, \lambda_{k+1}, \ldots, \lambda_r$ gäbe mit

$$v_k = \lambda_1 v_1 + \cdots + \lambda_{k-1} v_{k-1} + \lambda_{k+1} v_{k+1} + \cdots + \lambda_r v_r.$$

Dann ist dies ja insbesondere eine Darstellung von v_k als Linearkombination von Vektoren aus M. Natürlich kann man v_k aber auch folgendermaßen als Linearkombination schreiben:

$$v_k = 0v_1 + \cdots + 0v_{k-1} + 1v_k + 0v_{k+1} + \cdots 0v_r.$$

Diese beiden Linearkombinationen sind verschieden, weil der Faktor vor v_k in einem Fall 0 und im anderen Fall 1 ist. Somit haben wir einen Widerspruch zur Annahme (ii), dass Darstellungen als Linearkombination eindeutig sind.

„(iii) $\Longrightarrow$ (i)“:
Wir nehmen nun an, dass kein Vektor in M eine Linearkombination der anderen Vektoren ist und wollen zeigen, dass M linear unabhängig ist. Nach Definition der linearen Unabhängigkeit (Definition 2.3.9) nehmen wir dazu eine Darstellung des Nullvektors als Linearkombination von Vektoren aus M an

$$\sum_{j=1}^{r} \lambda_j v_j = 0$$

und müssen nun zeigen, dass alle Skalare 0 sind. Angenommen, für mindestens ein $k \in \{1,\dots,r\}$ gilt $\lambda_k \neq 0$. Dann können wir diese Gleichung nach v_k auflösen und erhalten

$$v_k = -\frac{1}{\lambda_k}(\lambda_1 v_1 + \cdots + \lambda_{k-1} v_{k-1} + \lambda_{k+1} v_{k+1} + \cdots + \lambda_r v_r),$$

was ein Widerspruch zu der Annahme ist, dass keiner der Vektoren aus M eine Linearkombination der übrigen ist.

Das zeigt die letzte der drei Implikationen und beendet damit den Beweis. □

Bemerkung 2.3.12. Wenn zwei Teilmengen $M_1 \subseteq M_2$ gegeben sind, und M_2 ist linear unabhängig, dann ist M_1 auch linear unabhängig. Wir werden allerdings im Folgenden versuchen, möglichst große linear unabhängige Teilmengen zu finden. Es gibt aber eine Obergrenze, wie groß eine linear unabhängige Teilmenge im $\mathbb{R}^n$ höchstens sein kann, wie der nächste Satz zeigt.

Satz 2.3.13. *Gegeben sei eine Menge linear unabhängiger Vektoren $M \subseteq \mathbb{R}^n$. Dann ist $|M| \leq n$.*

Beweis. Es sei $M = \{v_1,\dots,v_r\}$. Dann definieren wir die Matrix

$$A := \left(\begin{array}{c|c|c|c} & & & \\ v_1 & \cdots & v_r & \\ & & & \end{array} \right) \in \mathbb{R}^{n\times r},$$

die aus den Spalten $v_1,\dots,v_r$ besteht. Wir wollen zeigen, dass $r \leq n$. Per Widerspruch nehmen wir an, dass $r > n$. Nach Satz 2.3.7 existiert ein $x \in \ker A$ mit $x \neq 0$. Es gilt also:

$$\left(\begin{array}{c|c|c|c} & & & \\ v_1 & \cdots & v_r & \\ & & & \end{array} \right) \begin{pmatrix} x_1 \\ \vdots \\ x_r \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Das Matrixprodukt auf der linken Seite kann man umschreiben zu:

$$x_1 v_1 + \cdots + x_r v_r = 0$$

Da nicht alle $x_j = 0$ sind, bedeutet dies, dass $\{v_1,\dots,v_r\}$ linear abhängig ist. □

Definition 2.3.14 (Basis eines Untervektorraums von $\mathbb{R}^n$)**.** Gegeben sei ein Untervektorraum $U \subseteq \mathbb{R}^n$. Eine endliche Teilmenge $B \subseteq U$ heiße *Basis* von U, wenn sie die folgenden beiden Bedingungen erfüllt:

(i) Die Menge B ist ein Erzeugendensystem für U, d.h. $\mathrm{LH}(B) = U$.

(ii) Die Menge B ist linear unabhängig.

Lemma 2.3.15. *Es sei U ein Untervektorraum von $\mathbb{R}^n$ und $B = \{v_1, \ldots, v_r\} \subseteq U$ eine endliche Teilmenge. Dann sind die beiden folgenden Aussagen äquivalent:*

(i) B ist eine Basis von U.

(ii) Jeder Vektor in U besitzt eine eindeutige *Darstellung als Linearkombination aus Elementen aus B.*

Beweis. Dies folgt direkt aus Lemma 2.3.11. □

Beispiel 2.3.16. (a) Gegeben sei der folgende Untervektorraum

$$U := \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3 \;\middle|\; x_1 + x_2 + x_3 = 0 \right\} \subseteq \mathbb{R}^3.$$

Dann ist

$$B_1 := \left\{ \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \right\}$$

eine Basis für U.

Ebenso ist aber auch

$$B_2 := \left\{ \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}, \begin{pmatrix} 2 \\ -1 \\ -1 \end{pmatrix} \right\}$$

eine Basis für U. Ein Untervektorraum kann also mehr als eine Basis haben.

(b) Der ganze Raum $U = \mathbb{R}^2$ ist ein Untervektorraum von $\mathbb{R}^2$. Eine mögliche Basis ist

$$\left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \end{pmatrix} \right\}.$$

Eine andere ist

$$\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}.$$

Verallgemeinern wir das letzte Beispiel, so erhalten wir den folgenden Begriff:

Definition 2.3.17 (Standardbasis)**.** Die Menge

$$\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ldots, \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \right\}$$

bildet eine Basis für den Raum $U = \mathbb{R}^n$ und wird als *Standardbasis* des $\mathbb{R}^n$ bezeichnet. Die einzelnen Vektoren

$$e_1 := \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad e_2 := \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \cdots, \quad e_n := \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

werden dementsprechend *Standardbasisvektoren* genannt.

Wir wissen nun, dass der ganze Raum $\mathbb{R}^n$ eine Basis besitzt (und nicht nur eine). Es ist aber wichtig, sich klarzumachen, dass es überhaupt nicht offensichtlich ist, dass jeder Untervektorraum $U \subseteq \mathbb{R}^n$ eine Basis besitzt. Dies aber wollen wir nun beweisen:

Satz 2.3.18 (Basisauswahl- und ergänzungssatz)**.** *Gegeben seien*

- *ein Untervektorraum $U \subseteq \mathbb{R}^n$,*
- *ein Erzeugendensystem M von U,*
- *eine linear unabhängige Teilmenge $L \subseteq M$.*

Dann gibt es eine Basis B von U mit

$$L \subseteq B \subseteq M.$$

Beweis. Die Idee des Beweises besteht darin, die Menge L durch Hinzunahme von Elementen aus M immer größer zu machen, bis sie schließlich eine Basis von U wird.

Definieren wir dazu $L_r := L = \{v_1, \ldots, v_r\}$, wobei $r := |L| \in \mathbb{N}_0$.

Setzen wir $W_r := \mathrm{LH}(L)$, dann ist L ein linear unabhängiges Erzeugendensystem, also eine Basis, für W_r. Da $L_r \subseteq U$ ist, gilt nach Lemma 2.3.4, dass $W_r \subseteq U$. Es gibt nun zwei Fälle zu unterscheiden:

Erster Fall: $M \subseteq W_r$

Wenn jedes Element von M im Untervektorraum W_r liegt, dann gilt nach Lemma 2.3.4 auch $\mathrm{LH}(M) \subseteq W_r$. Also ist $U = W_r$. Das bedeutet, dass L_r eine Basis für U ist und mit der Wahl $B := L_r$ ist der Satz bewiesen.

Zweiter Fall: $M \nsubseteq W_r$

In diesem Fall gibt es mindestens ein $v_{r+1} \in M$, das nicht in $W_r = \mathrm{LH}(L)_r$ liegt, sich also nicht als Linearkombination aus Vektoren aus L_r schreiben lässt.

Wir behaupten nun $L_{r+1} := L_r \cup \{v_{r+1}\} = \{v_1, \ldots, v_r, v_{r+1}\}$ ist linear unabhängig. Es sei dazu

$$\lambda_1 v_1 + \cdots + \lambda_r v_r + \lambda_{r+1} v_{r+1} = 0$$

eine Darstellung des Nullvektors als Linearkombination. Falls $\lambda_{r+1} \neq 0$ ist, können wir diese Gleichung nach v_{r+1} auflösen und den Vektor v_{r+1} somit als Linearkombination aus Vektoren aus L_r darstellen. Das ist aber nicht möglich, weil wir v_{r+1} ja gerade so ausgewählt haben, dass $v_{r+1} \notin \mathrm{LH}(L)_r$. Also muss $\lambda_{r+1} = 0$ sein. Damit erhalten wir

$$\lambda_1 v_1 + \cdots + \lambda_r v_r = 0$$

Da aber die Menge $L = \{v_1, \ldots, v_r\}$ als linear unabhängig vorausgesetzt war, folgt damit, dass auch alle anderen $\lambda_j = 0$ sein müssen.

Also ist $L_{r+1} = L_r \cup \{v_{r+1}\}$ linear unabhängig. Die Menge L_{r+1} ist enthalten in M und enthält L_r als Teilmenge. Wenn wir nun also zum Beginn des Beweises zurückkehren und dasselbe Spiel mit L_{r+1} spielen, erhalten wir entweder eine Basis (1. Fall) oder wir können die Menge noch weiter vergrößern (zu L_{r+2}).

Auf diese Weise erhalten wir eine aufsteigende Folge $L_r \subsetneq L_{r+1} \subsetneq L_{r+2} \subseteq \cdots$, die aber nicht beliebig lange so fortgesetzt werden kann, weil nach Satz 2.3.13 eine linear unabhängige Menge in $\mathbb{R}^n$ höchstens n Elemente haben kann.

Also muss nach endlich vielen Schritten irgendwann Fall 1 auftreten und die Aussage ist bewiesen. □

Korollar 2.3.19. *Jeder Untervektorraum $U \subseteq \mathbb{R}^n$ besitzt eine Basis.*

Beweis. Wir wenden Satz 2.3.18 auf den Untervektorraum U, die linear unabhängige Teilmenge $L := \emptyset$ und das Erzeugendensystem $M := U$ an. □

Korollar 2.3.20. *Gegeben sind zwei Untervektorräume $W \subseteq U \subseteq \mathbb{R}^n$. Dann kann man jede Basis von W zu einer Basis von U vervollständigen.*

Beweis. Es sei L die gegebene Basis des Untervektorraums W. Wir wenden Satz 2.3.18 auf den Untervektorraum U, die linear unabhängige Teilmenge L und das Erzeugendensystem $M := U$ an. □

Insbesondere kann man also jede Basis eines Untervektorraums von $\mathbb{R}^n$ zu einer Basis des $\mathbb{R}^n$ vervollständigen.

Wir wissen also nun: Jeder Untervektorraum U hat eine Basis, aber es ist noch nicht klar, dass zwei unterschiedliche Basen des gleichen Unterraums dieselbe Anzahl an Elementen haben. Das wollen wir nun zeigen:

Satz 2.3.21 (Steinitzscher Austauschsatz[6])**.** *Es sei U ein Untervektorraum von $\mathbb{R}^n$ und $B = \{b_1, \ldots, b_r\}$ eine Basis von U. Gegeben ein Vektor $v \in U$ mit Darstellung*

$$v = \sum_{j=1}^{r} \lambda_j b_j.$$

Wenn $k \in \{1, \ldots, r\}$ so gewählt ist, dass $\lambda_k \neq 0$, dann ist auch $B' := (B \setminus \{b_k\}) \cup \{v\} = \{b_1, \ldots, b_{k-1}, v, b_{k+1}, \ldots, b_r\}$ eine Basis von U.

Beweis. Wir müssen zeigen, dass B' linear unabhängig ist und dass $\mathrm{LH}(B') = U$.

Warum ist B' linear unabhängig?

Es sei

$$\mu_1 b_1 + \cdots + \mu_{k-1} b_{k-1} + \mu_k v + \mu_{k+1} b_{k+1} + \cdots + \mu_r b_r = 0.$$

Wir wollen zeigen, dass alle Skalare $\mu_j = 0$ sind. Wenn wir nun für v den Ausdruck $\sum_{j=1}^{r} \lambda_j b_j$ einsetzen, ergibt dies:

$$\mu_1 b_1 + \cdots + \mu_{k-1} b_{k-1} + \mu_k \sum_{j=1}^{r} \lambda_j b_j + \mu_{k+1} b_{k+1} + \cdots + \mu_r b_r = 0.$$

Umgruppieren und Zusammenfassen ergibt:

$$(\mu_1 + \mu_k \lambda_1) b_1 + \cdots + (\mu_{k-1} + \mu_k \lambda_{k-1}) b_{k-1} + \mu_k \lambda_k b_k + (\mu_{k+1} + \mu_k \lambda_{k+1}) b_{k+1} + \cdots + (\mu_r + \mu_k \lambda_r) b_r = 0.$$

[6] nach Ernst Steinitz, deutscher Mathematiker, 1871–1928

Da die Menge $\{b_1,\dots,b_r\}$ linear unabhängig ist, sind alle Skalare in dieser Linearkombination Null, insbesondere der vor dem Vektor b_k. Es gilt also

$$\mu_k \lambda_k = 0.$$

In den reellen Zahlen ist ein Produkt nur dann Null, wenn einer der Faktoren Null ist. Nach Voraussetzung des Satzes ist aber $\lambda_k \neq 0$, also muss $\mu_k = 0$ sein. Wenn wir das in der obigen Linearkombination einsetzen, ergibt das:

$$\mu_1 b_1 + \cdots + \mu_{k-1} b_{k-1} + \mu_{k+1} b_{k+1} + \cdots + \mu_r b_r = 0.$$

Nochmaliges Verwenden der linearen Unabhängigkeit ergibt, dass alle anderen μ_j auch Null sein müssen. Das beendet den Beweis dafür, dass B' linear unabhängig ist.

Warum ist B' ein Erzeugendensystem von U?
Da $B \subseteq U$ und $v \in U$ sind, ist auf jeden Fall schon einmal geklärt, dass $B' \subseteq U$. Mit Lemma 2.3.4 folgt somit, dass $\mathrm{LH}(B') \subseteq U$.

Es bleibt zu zeigen $U \subseteq \mathrm{LH}(B')$, dass also jeder Vektor in U als Linearkombination von Vektoren in B' dargestellt werden kann.

Der Vektor v hat die folgende Darstellung:

$$v = \lambda_1 b_1 + \cdots + \lambda_{k-1} b_{k-1} + \lambda_k b_k + \lambda_{k+1} b_{k+1} + \cdots + \lambda_r b_r.$$

Da nach Voraussetzung $\lambda_k \neq 0$ gilt, können wir diese Gleichung nach b_k auflösen:

$$b_k = -\frac{\lambda_1}{\lambda_k} b_1 - \cdots - \frac{\lambda_{k-1}}{\lambda_k} b_{k-1} + \frac{1}{\lambda_k} v - \frac{\lambda_{k+1}}{\lambda_k} b_{k+1} - \cdots - \frac{\lambda_r}{\lambda_k} b_r \in \mathrm{LH}(B').$$

Dies zeigt, dass $B \subseteq \mathrm{LH}(B')$. Also ist (nach Lemma 2.3.4) auch $U = \mathrm{LH}(B) \subseteq \mathrm{LH}(B')$. □

Mit diesem Satz kann man also eine Basis mit r Elementen so abändern, dass man eine neue Basis mit r Elementen erhält. Es stellt sich die Frage: Haben alle Basen eines Untervektorraums die gleiche Anzahl von Elementen?

Satz 2.3.22 (Wohldefiniertheit der Dimension eines Untervektorraums)**.** *Gegeben sei ein Untervektorraum $U \subseteq \mathbb{R}^n$.*

Dann haben alle Basen von U gleich viele Elemente.

Beweis. Gegeben seien zwei Basen $B = \{b_1,\dots,b_r\}$ und $C = \{c_1,\dots,c_s\}$ mit $r,s \in \mathbb{N}_0$. Wir werden zeigen, dass $r = s$ gilt. Da für zwei Zahlen r,s immer gilt

$$r \leq s \text{ oder } s \leq r,$$

dürfen wir ohne Beschränkung der Allgemeinheit (o.B.d.A.) annehmen, dass $r \leq s$.

Es bleibt zu zeigen, dass $s \leq r$.

Wir werden so vorgehen, dass wir Schritt für Schritt einen Vektor aus B durch einen Vektor aus C ersetzen. Um dieses schrittweise Vorgehen formal korrekt abzubilden, führen wir eine Hilfsaussage ein:

Behauptung:

$$\forall k \in \mathbb{N}_0 : k \leq s \implies (\exists B_k \subseteq U : B_k \text{ ist eine Basis von } U \text{ mit } |B_k| = r \text{ und } \{c_1,\dots,c_k\} \subseteq B_k).$$

Sobald wir diese Behauptung gezeigt haben, folgt daraus direkt, was wir zeigen wollen, denn dann können wir für $k = s$ einsetzen und erhalten eine r-elementige Basis von U gibt, die die s-elementige Menge $\{c_1, \ldots, c_s\}$ als Teilmenge enthält. Dann muss also gelten: $s \leq r$.

Es bleibt die Frage, warum diese Behauptung wahr sein sollte. Wir beweisen sie nun mit vollständiger Induktion über k.

Induktionsanfang $k = 0$:
Wir müssen zeigen, dass es eine Basis von U gibt, die genauso viele Elemente wie B besitzt und $\emptyset$ als Teilmenge hat. Das ist einfach: Wir setzen einfach $B_0 := B$ und sind fertig.

Induktionsschritt:
Wir nehmen also an, $k \in \mathbb{N}_0$ ist so gewählt, dass die folgende Induktionsannahme

$$k \leq s \implies (\exists B_k \subseteq U : B_k \text{ ist eine Basis von } U \text{ mit } |B_k| = r \text{ und } \{c_1, \ldots, c_k\} \subseteq B_k)$$

gilt. Wir müssen nun zeigen, dass die folgende Aussage auch wahr ist:

$$k+1 \leq s \implies (\exists B_k \subseteq U : B_{k+1} \text{ ist eine Basis von } U \text{ mit } |B_{k+1}| = r \text{ und } \{c_1, \ldots, c_{k+1}\} \subseteq B_k).$$

Falls $k+1 > s$, ist die zu zeigende Aussage rein logisch korrekt (*Ex falsum quodlibet*) und wir können uns somit auf den Fall $k+1 \leq s$ beschränken. Insbesondere ist dann auch $k < s$ und unsere Induktionsannahme vereinfacht sich zu:

$$\exists B_k \subseteq U : B_k \text{ ist eine Basis von } U \text{ mit } |B_k| = r \text{ und } \{c_1, \ldots, c_k\} \subseteq B_k.$$

Wir wissen also, dass es eine Basis B_k von U gibt, die die Vektoren $c_1, \ldots, c_k$ enthält. Da $k+1 \leq s$ gibt es einen Vektor $c_{k+1} \in C \subseteq U$. Wir schreiben c_{k+1} als Linearkombination aus Vektoren aus der Basis B_k:

$$c_{k+1} = \lambda_1 c_1 + \lambda_2 c_2 + \cdots + \lambda_k c_k + \lambda_{k+1} b_{k+1} + \cdots + \lambda_r b_r.$$

Falls alle λ_j mit $j > k$ null wären, dann wäre $c_{k+1} \in \mathrm{LH}(\{c_1, \ldots, c_k\})$, was nicht sein kann, weil die Menge C linear unabhängig ist. Also gibt es mindestens ein $\lambda_j \neq 0$ mit $j > k$. Nun wenden wir den Steinitzschen Austauschsatz (Satz 2.3.21) auf den Vektor c_{k+1} an und erhalten eine Basis B_{k+1}, die nun $\{c_1, \ldots, c_{k+1}\}$ enthält und immer noch r Elemente hat.

Das beendet den Beweis. □

Der folgende Begriff der Dimension ist vielleicht der wichtigste in diesem Kapitel:

Definition 2.3.23 (Dimension eines Untervektorraums)**.** Es sei $U \subseteq \mathbb{R}^n$ ein Untervektorraum. Dann ist die *Dimension* von U definiert als die Anzahl der Elemente in einer Basis von U. Man schreibt

$$\dim U := |B| \in \mathbb{N}_0 \quad \text{für eine Basis } B \text{ von } U.$$

Die Dimension ist nur deshalb wohldefiniert, weil wir in Korollar 2.3.19 gesehen haben, dass jeder Untervektorraum eine Basis besitzt und weil wir in Satz 2.3.22 bewiesen haben, dass jede Basis gleich viele Elemente besitzt.

Bemerkung 2.3.24. Es ist nicht möglich, von *der* Basis eines Untervektorraums zu sprechen, da ein Untervektorraum viele unterschiedliche Basen hat. Es ist aber sehr wohl möglich, von *der* Dimension eines Untervektorraums zu sprechen.

Die Dimension eines Untervektorraums ist ein Maß, um anzugeben, wie *groß* er ist. Bei unendlichen Mengen ist die Angabe der Anzahl der Elemente oft nicht mehr nützlich (Beispielsweise haben die Mengen $\mathbb{R}^2$, $\mathbb{R}^3$ gleich viele Elemente, aber unterschiedliche Dimensionen.

Beispiel 2.3.25. (a) Die leere Menge $B := \emptyset$ ist eine Basis für den Untervektorraum $U := \{0\} \subseteq \mathbb{R}^n$. Somit gilt $\dim\{0\} = 0$. Man beachte, dass $\{0\}$ keine Basis für U ist, weil $\{0\}$ linear abhängig ist (siehe Beispiel 2.3.10 (b))

(b) Die Standardbasis des Raumes $\mathbb{R}^n$ hat n Elemente. Somit gilt

$$\dim \mathbb{R}^n = n.$$

(c) Für jeden Vektor $v \in \mathbb{R}^n$ mit $v \neq 0$ ist

$$U := \{\lambda v \mid \lambda \in \mathbb{R}\}$$

die Gerade durch die 0 mit Richtungsvektor v. Da $U = \mathrm{LH}(\{v\})$ und $v \neq 0$ gilt, ist $\{v\}$ eine Basis und somit ist $\dim U = 1$.

(d) Für zwei linear unabhängige Vektoren $v, w \in \mathbb{R}^n$ ist

$$U := \mathrm{LH}(v,w) = \left\{\lambda v + \mu w \;\middle|\; \lambda, \mu \in \mathbb{R}\right\}$$

die Ebene durch die 0 mit Richtungsvektoren v, w. Für eine solche Ebene gilt $\dim U = 2$.

Lemma 2.3.26 (Monotonie der Dimension)**.**
Gegeben seien zwei Untervektorräume $W \subseteq U \subseteq \mathbb{R}^n$. Dann gilt:

(a) $\dim W \leq \dim U$.

(b) $(W \subsetneq U) \implies (\dim W < \dim U)$.

Beweis. (a)
Nach Korollar 2.3.19 besitzt der Untervektorraum W eine Basis B_W. Diese Menge $B_W \subseteq W \subseteq U$ ist also insbesondere linear unabhängig. Nach Satz 2.3.18 lässt sich diese Menge also zu einer Basis B_U von U fortsetzen.

Da $B_W \subseteq B_U$ ist, gilt insbesondere

$$\dim W = |B_W| \leq |B_U| = \dim U.$$

(b)
Wir beginnen wieder mit Korollar 2.3.19 und beschaffen uns eine Basis B_W für den Untervektorraum W. Da $W \subsetneq U$ gilt, bedeutet dies, dass es ein $u \in U$ gibt, mit $u \notin W$. Somit ist $L := B_W \cup \{u\}$ linear unabhängig. Die Menge L ist somit eine Basis für $\mathrm{LH}(L) \subseteq U$.

Es gilt somit:

$$\dim(U) \geq \dim(\mathrm{LH}(L)) = |B_W \cup \{u\}| = |B_W| + 1 = \dim(W) + 1 > \dim(W). \qquad \square$$

Zusammenfassung von Abschnitt 2.3

(1) Die Lösungsmenge des homogenen linearen Gleichungssystems $Ax = 0$ (mit $A \in \mathbb{R}^{m \times n}$) wird mit $\ker A$ bezeichnet und ist immer ein Untervektorraum von $\mathbb{R}^n$.

(2) Eine Basis eines Untervektorraums ist ein Erzeugendensystem, das linear unabhängig ist.

(3) Jeder Untervektorraum hat eine Basis, diese ist aber nicht eindeutig.

(4) Die Dimension eines Untervektorraums ist die Anzahl der Elemente in einer Basis. Weil jede Basis gleich viele Elemente hat, ist dies wohldefiniert.

2.4. Affine Unterräume

Wir haben in Beispiel 2.3.25 gesehen, dass eine Gerade durch die 0 ein Untervektorraum der Dimension 1 ist.

Ebenso ist eine Ebene, die die 0 enthält, ein Untervektorraum der Dimension 2.

Ebenen oder Geraden, die nicht die 0 enthalten, können wir auf diese Weise nicht behandeln. Dies ist *algebraisch* gesehen sinnvoll, weil eine Gerade, die nicht die 0 enthält, nicht die Eigenschaften aus Definition 2.3.1 erfüllt (Man kann sogar leicht einsehen, dass keine der drei Bedingungen erfüllt ist). *Geometrisch* gesehen ist dies unbefriedigend, weil eine Gerade immer eine Gerade sein sollte, unabhängig davon, ob sie durch den Ursprung geht oder nicht. Deshalb wollen wir nun den Begriff des Untervektorraums so verallgemeinern, dass alle Geraden und Ebenen enthalten sind.

Notation 2.4.1. Es sei $M \subseteq \mathbb{R}^n$ eine Teilmenge und $p \in \mathbb{R}^n$. Dann ist

$$p + M := M + p := \{p + x \mid x \in M\} \subseteq \mathbb{R}^n$$

die um den Vektor p verschobene Menge M.

Definition 2.4.2 (Affiner Unterraum)**.** Es sei $n \in \mathbb{N}$. Eine Teilmenge $R \subseteq \mathbb{R}^n$ heiße *affiner Unterraum* von $\mathbb{R}^n$, wenn ein $p \in \mathbb{R}^n$ und ein Untervektorraum $U \subseteq \mathbb{R}^n$ existieren, sodass

$$R = p + U.$$

Den Vektor p in dieser Darstellung nennt man auch *Fußpunkt*.

Affine Unterräume sind also verschobene Untervektorräume.

Bemerkung 2.4.3. Da $p = 0$ ausdrücklich erlaubt ist, bedeutet dies, dass jeder Untervektorraum ein affiner Unterraum ist. Die Umkehrung gilt nicht, da es viele affine Unterräume gibt, die keine Untervektorräume sind.

Achtung: In der Darstellung $R = p + U$ ist der Fußpunkt p im Allgemeinen *nicht* eindeutig. So können Sie jedes Element von R als Fußpunkt verwenden und erhalten eine andere zulässige Darstellung. Bemerkenswerterweise ist aber der Untervektorraum U eindeutig durch die Menge R bestimmt, denn es gilt:

$$U = \{x - y \mid x, y \in R\}. \qquad (*)$$

Die Menge der Differenzvektoren zweier Punkt in R ergibt also den Untervektorraum U.

Beweis der Gleichheit $(*)$:

„$\subseteq$“:

Es sei $v \in U$ gegeben. Dann ist $x := p + v \in p + U = R$. Außerdem ist $y := p = p + 0 \in p + U = R$. Also gilt:

$$v = (p + v) - p = x - y.$$

„$\supseteq$“:

Es seien $x, y \in R$ gegeben. Da $R = p + U$ gilt, gibt es $u, w \in U$ mit $x = p + u$ und $y = p + w$. Also gilt:

$$x - y = (p + u) - (p + w) = u - w.$$

Da U ein Untervektorraum und $u, w \in U$, ist auch die Linearkombination $u - w$ ein Element in U. Das war zu zeigen.

Definition 2.4.4 (Dimension eines affinen Unterraums)**.** Gegeben sei ein affiner Unterraum $R \subseteq \mathbb{R}^n$. Wenn sich R schreiben lässt als $p+U$, wobei $p \in \mathbb{R}^n$ und U ein Untervektorraum von $\mathbb{R}^n$ ist, dann ist die *Dimension* von R definiert als die Dimension des dazugehörigen Untervektorraums U, d.h.

$$\dim R := \dim U.$$

Bemerkung 2.4.5. (a) Man beachte, dass dies wohldefiniert ist, weil in der Darstellung $R = p+U$ der Untervektorraum U eindeutig ist (siehe Bemerkung 2.4.3) – obwohl p nicht eindeutig ist.

(b) Jeder Untervektorraum ist auch ein affiner Unterraum. In diesem Fall ist die Dimension als affiner Unterraum (Definition 2.4.4) identisch mit der Dimension als Untervektorraum (Definition 2.3.23).

Beispiel 2.4.6. (a) Jede einelementige Menge $R := \{p\}$ lässt sich schreiben als $R = p+\{0\}$ und ist somit ein affiner Unterraum. Falls $p \neq 0$ gilt, ist dies aber kein Untervektorraum.

(b) Für zwei Punkte $a, b \in \mathbb{R}^n$ mit $a \neq b$ ist die *Gerade durch a und b* gegeben durch

$$R := \{\lambda b + (1-\lambda)a \mid \lambda \in \mathbb{R}\}.$$

Dies ist ein affiner Unterraum von $\mathbb{R}^n$. Um dies zu sehen, betrachten wir die Ursprungsgerade in Richtung $b-a$:

$$U := \mathrm{LH}(b-a) = \{\lambda(b-a) \mid \lambda \in \mathbb{R}\}$$

und schreiben R folgendermaßen um:

$$\begin{aligned} R &= \{\lambda b + (1-\lambda)a \mid \lambda \in \mathbb{R}\} \\ &= \{\lambda b + a - \lambda a \mid \lambda \in \mathbb{R}\} \\ &= \{a + \lambda(b-a) \mid \lambda \in \mathbb{R}\} \\ &= a + \{\lambda(b-a) \mid \lambda \in \mathbb{R}\} \\ &= a + U. \end{aligned}$$

Die Gerade R entspricht also der Ursprungsgerade U um den Vektor a verschoben.

(c) Für drei Punkte $a, b, c \in \mathbb{R}^n$, die nicht alle auf einer Geraden liegen, ist die *Ebene durch a, b, c* gegeben durch

$$R := \left\{a + \lambda(b-a) + \mu(c-a) \,\middle|\, \lambda, \mu \in \mathbb{R}\right\}.$$

Dies ist ein affiner Unterraum $p+U$ mit $p = a$ und $U = \mathrm{LH}(b-a, c-a)$. Die Dimension von U ist 2, da $\{b-a, c-a\}$ linear unabhängig sein müssen – sonst lägen a, b, c auf einer Geraden.

Definition 2.4.7 (Affinkombinationen)**.** Gegeben seien Vektoren $v_1, \ldots, v_r \in \mathbb{R}^n$ mit $r \in \mathbb{N}$. Eine *Affinkombination* der Vektoren $v_1, \ldots, v_r$ ist eine Linearkombination der Form

$$\sum_{j=1}^{r} \lambda_j v_j = \lambda_1 v_1 + \cdots + \lambda_r v_r,$$

wobei die Skalare sich zu 1 aufaddieren:

$$\sum_{j=1}^{r} \lambda_j = 1.$$

Jede Affinkombination ist eine Linearkombination, aber nicht jede Linearkombination ist eine Affinkombination.

Satz 2.4.8 (Charakterisierung affiner Unterräume)**.** *Für eine Teilmenge $R \subseteq \mathbb{R}^n$ sind die folgenden Aussagen äquivalent:*

(i) R *ist ein affiner Unterraum von* $\mathbb{R}^n$.

(ii) $R \neq \emptyset$ *und* $\forall x, y, z \in R, \lambda \in \mathbb{R} : x + \lambda(y - z) \in R$.

(iii) $R \neq \emptyset$ *und jede Affinkombination von endlich vielen Vektoren in* R *ist in* R.

Beweis. Wir zeigen wieder eine Reihe von Implikationen:
„(i) $\Longrightarrow$ (iii)“:
Wir nehmen an, R sei ein affiner Unterraum, d.h. es gibt $p \in \mathbb{R}^n$ und einen Untervektorraum $U \subseteq \mathbb{R}^n$ mit

$$R = p + U.$$

Da U ein Untervektorraum ist, gilt somit $0 \in U$ und daher $p \in R$. Also ist $R \neq \emptyset$.

Gegeben seien Vektoren $v_1, \ldots, v_r \in R$. Es bleibt zu zeigen, dass jede Affinkombination dieser Vektoren in R liegt. Seien dazu $\lambda_1, \ldots, \lambda_r \in \mathbb{R}$ gegeben mit $\lambda_1 + \cdots + \lambda_r = 1$. Wir möchten zeigen, dass $\sum_{j=1}^{r} \lambda_j v_j \in R$.

Da jedes $v_j \in R = p + U$ ist, gibt es ein $u_j \in U$ mit $v_j = p + u_j$. Damit ergibt sich:

$$\begin{aligned}
\sum_{j=1}^{r} \lambda_j v_j &= \sum_{j=1}^{r} \lambda_j (p + u_j) \\
&= \sum_{j=1}^{r} \lambda_j p + \sum_{j=1}^{r} \lambda_j u_j \\
&= \underbrace{\left(\sum_{j=1}^{r} \lambda_j \right)}_{=1} p + \underbrace{\sum_{j=1}^{r} \lambda_j u_j}_{\in U} \\
&\in p + U = R.
\end{aligned}$$

„(iii) $\Longrightarrow$ (ii)“:
Da $x + \lambda(y - z) = 1x + \lambda y + (-\lambda)z$ eine Affinkombination aus x, y, z ist, folgt die Aussage direkt.
„(ii) $\Longrightarrow$ (i)“:
Nach Voraussetzung ist $R \neq \emptyset$. Also gibt es ein $p \in R$. Wir definieren nun

$$U := \{x - p \mid x \in R\} = R - p \subseteq \mathbb{R}^n.$$

Dann ist $R = p + U$. Es bleibt zu zeigen, dass U ein Untervektorraum von $\mathbb{R}^n$ ist, d.h. dass U alle drei Eigenschaften aus Definition 2.3.1 erfüllt:

Es ist $0 = p - p \in U$, d.h. U enthält den Nullvektor.

Zeigen wir nun, dass U abgeschlossen unter Addition ist: Betrachten wir dazu $v, w \in U$ mit $v = x - p$ und $w = y - p$. Dann ist

$$v + w = (x - p) + (y - p) = (x + y - p) - p.$$

Der Ausdruck $x + y - p$ ist in R nach Voraussetzung (ii). Also ist $v + w \in U$.

Zeigen wir nun, dass U abgeschlossen unter skalarer Vielfachbildung ist: Betrachten wir dazu ein $v \in U$ mit $v = x - p$ und ein $\lambda \in \mathbb{R}$. Dann ist

$$\lambda v = p + \lambda v - p = \underbrace{p + \lambda(x-p)}_{\in R} - p.$$

Der Ausdruck $p + \lambda(x-p)$ ist in R nach Voraussetzung (ii). Also $\lambda v \in U$.

Also ist U ein Untervektorraum von $\mathbb{R}^n$ und somit $R = p + U$ ein affiner Unterraum. □

Bemerkung 2.4.9. Es gibt noch eine schöne geometrische Charakterisierung, die man auf eine sehr ähnliche Weise wie in Satz 2.4.8 zeigen kann:

Eine Teilmenge $R \subseteq \mathbb{R}^n$ ist ein affiner Unterraum genau dann, wenn R nicht leer ist und zu je zwei Punkten $x, y \in R$ mit $x \neq y$ die Gerade[7] durch x und y ganz in R liegt.

Wir werden diese Charakterisierung aber im Weiteren nicht verwenden.

In Lemma 2.3.2 haben wir gesehen, dass die Lösungsmenge eines *homogenen* linearen Gleichungssystems immer ein Untervektorraum ist. Wir werden nun sehen, dass die Lösungsmenge eines *inhomogenen* linearen Gleichungssystems immerhin ein affiner Unterraum ist – außer sie ist leer.

Satz 2.4.10 (Struktur der Lösungsmenge eines linearen Gleichungssystems)**.** *Gegeben seien eine Matrix $A \in \mathbb{R}^{m \times n}$ und ein Vektor $b \in \mathbb{R}^m$ sowie das dazugehörige lineare Gleichungssystem:*

$$Ax = b.$$

Wenn dieses Gleichungssystem mindestens eine Lösung $x_0 \in \mathbb{R}^n$ hat, dann ist die Lösungsmenge ein affiner Unterraum des $\mathbb{R}^n$ und es gilt:

$$\left\{x \in \mathbb{R}^n \;\middle|\; Ax = b\right\} = x_0 + \ker A.$$

Man sieht hier ein sehr allgemeines Prinzip, das einem nicht nur bei linearen Gleichungssystemen, sondern auch bei linearen Differentialgleichungen immer wieder begegnet: Die allgemeine Lösung x ist die Summe einer partikulären Lösung (also x_0) und der allgemeinen Lösung des dazugehörigen homogenen Problems (also $\ker A$).

Beweis. Angenommen $x_0 \in \mathbb{R}^n$ ist eine Lösung. Dann müssen wir die folgende Mengengleichheit zeigen:

$$\left\{x \in \mathbb{R}^n \;\middle|\; Ax = b\right\} = x_0 + \ker A.$$

„⊆“:
Es sei $x \in \mathbb{R}^n$ mit $Ax = b$. Wir setzen $u := x - x_0$. Somit gilt $x = x_0 + u$. Es bleibt zu zeigen, dass $u \in \ker A$.

$$Au = A(x - x_0) = Ax - Ax_0 = b - b = 0.$$

Also ist $u \in \ker A$ und somit $x \in x_0 + \ker A$.

„⊇“:
Es sei umgekehrt $x = x_0 + u$ mit $u \in \ker A$. Dann gilt

$$Ax = A(x_0 + u) = Ax_0 + Au = b + 0 = b.$$

Also ist $Ax = b$. □

[7] siehe Beispiel 2.4.6(b)

Satz 2.4.10 beantwortet nun also die Frage, wie die Lösungsmenge eines linearen Gleichungssystems aussieht – außer sie ist leer. Ob sie leer ist oder nicht, hängt von der rechten Seite b ab. Für eine Matrix $A \in \mathbb{R}^{m\times n}$ wollen wir nun untersuchen, für welche rechte Seite b dies der Fall ist:

Definition 2.4.11. Gegeben sei eine Matrix $A \in \mathbb{R}^{m\times n}$. Das *Bild* von A ist definiert als das Bild[8] der Abbildung

$$\mathbb{R}^n \to \mathbb{R}^m, \quad x \mapsto Ax.$$

Also ist $\text{Bild}(A) = \{Ax \mid x \in \mathbb{R}^n\} = \{b \in \mathbb{R}^m \mid \exists x \in \mathbb{R}^n : Ax = b\}$.

Ein lineares Gleichungssystem $Ax = b$ hat somit genau dann eine Lösung, wenn $b \in \text{Bild}(A)$.

Lemma 2.4.12. *Das Bild einer Matrix $A \in \mathbb{R}^{m\times n}$ ist die lineare Hülle der Spaltenvektoren von A. Insbesondere ist* $\text{Bild}(A)$ *immer ein Untervektorraum von* $\mathbb{R}^m$.

Beweis. Wir bezeichnen mit $v_1, v_2, \ldots, v_n$ die n Spalten der Matrix A und mit $U = \text{LH}(\{v_1, v_2, \ldots, v_n\})$ die lineare Hülle der Spaltenvektoren. Die Menge U ist ein Untervektorraum von $\mathbb{R}^m$ nach Lemma 2.3.4.

Es bleibt zu zeigen: $\text{Bild}(A) = U$.

„$\subseteq$“:

Es sei $b \in \text{Bild}(A)$. Dann gibt es ein $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$ mit $Ax = b$. Also gilt

$$b = Ax = \left(\begin{array}{c|c|c|c} & & & \\ v_1 & \cdots & v_n & \\ & & & \end{array} \right) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = x_1 v_1 + \cdots + x_n v_n \in U.$$

„$\supseteq$“:

Es sei $b \in U = \text{LH}(\{v_1, v_2, \ldots, v_n\})$. Dann gibt es Skalare $\lambda_1, \ldots \lambda_n \in \mathbb{R}$ mit:

$$b = \lambda_1 v_1 + \cdots \lambda_n v_n = \left(\begin{array}{c|c|c|c} & & & \\ v_1 & \cdots & v_n & \\ & & & \end{array} \right) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = Ax \in \text{Bild}(A),$$

wobei $x = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} \in \mathbb{R}^n$. □

Wir kommen nun zu einer sehr wichtigen Kennzahl einer Matrix:

Definition 2.4.13 (Der Rang einer Matrix)**.** Es sei $A \in \mathbb{R}^{m\times n}$. Dann ist der *Rang* von A definiert als die Dimension des Bildes von A, d.h.

$$\text{rg}(A) := \dim(\text{Bild}(A)).$$

Wegen Lemma 2.4.12 ist der Rang also die Dimension des Unterraums, der von den Spalten der Matrix aufgespannt wird, also die maximale Anzahl linear unabhängiger Spalten von A.

[8] siehe Notation 1.3.3(2)

Bemerkung 2.4.14. (a) Die Abbildung

$$\mathbb{R}^n \to \mathbb{R}^m, \quad x \mapsto Ax$$

ist genau dann surjektiv, wenn $\operatorname{rg}(A) = m$ ist. Falls dies der Fall ist, hat also für jedes $b \in \mathbb{R}^m$ das lineare Gleichungssystem $Ax = b$ eine nichtleere Lösungsmenge.

(b) Die Matrix $A \in \mathbb{R}^{m\times n}$ hat höchstens n verschiedene Spalten, somit kann die Dimension des Bildes, das von den Spalten aufgespannt wird, höchstens n sein. Andererseits ist $\operatorname{Bild}(A)$ ein Untervektorraum von $\mathbb{R}^m$. Also kann die Dimension nach Lemma 2.3.26 höchstens m sein. Insgesamt gilt also:

$$\operatorname{rg}(A) \le \min\{m, n\}.$$

(c) Wir werden in Satz 2.5.8 sehen, dass der Rang auch gleichzeitig die maximale Anzahl linear unabhängiger Zeilen ist. („Zeilenrang = Spaltenrang“).

Lemma 2.4.15. *Gegeben Matrizen $A \in \mathbb{R}^{m\times n}$ und $B \in \mathbb{R}^{n\times p}$. Dann ist*

$$\operatorname{rg}(AB) \le \operatorname{rg}(A) \quad \textit{und} \quad \operatorname{rg}(AB) \le \operatorname{rg}(B).$$

Beweis. Zuerst beweisen wir $\operatorname{rg}(AB) \le \operatorname{rg}(A)$.

Es sei $z \in \operatorname{Bild}(AB)$. Dann gibt es ein $x \in \mathbb{R}^p$ mit $ABx = z$. Setzen wir nun $y := Bx \in \mathbb{R}^n$, dann ist $z = Ay \in \operatorname{Bild}(A)$. Da z beliebig gewählt war, gilt somit die Teilmengenrelation:

$$\operatorname{Bild}(AB) \subseteq \operatorname{Bild}(A).$$

Aus der Monotonie der Dimension 2.3.26 folgt dann

$$\underbrace{\dim \operatorname{Bild}(AB)}_{=\operatorname{rg}(AB)} \le \underbrace{\dim \operatorname{Bild}(A)}_{=\operatorname{rg}(A)}.$$

Nun wollen wir $\operatorname{rg}(AB) \le \operatorname{rg}(B)$ beweisen. Die Menge

$$\operatorname{Bild}(B) = \left\{Bx \;\middle|\; x \in \mathbb{R}^p\right\}$$

ist ein Untervektorraum von $\mathbb{R}^n$ und hat nach Korollar 2.3.19 eine Basis $\{y_1, \dots, y_r\}$. Die Anzahl der Basisvektoren ist $r = \operatorname{rg}(B)$.

Es sei nun $z \in \operatorname{Bild}(AB)$. Dann gibt es – wie oben schon erwähnt – ein $x \in \mathbb{R}^p$ mit $ABx = z$. Der Vektor $Bx \in \operatorname{Bild}(B)$ ist als Linearkombination der Basisvektoren darstellbar:

$$Bx = \sum_{j=1}^{r} \lambda_j y_j.$$

Dann gilt:

$$z = ABx = A\sum_{j=1}^{r} \lambda_j y_j = \sum_{j=1}^{r} \lambda_j A y_j \in \operatorname{LH}(\{Ay_1, \dots, Ay_r\}).$$

Da z beliebig gewählt war, gilt somit die Teilmengenrelation:

$$\operatorname{Bild}(AB) \subseteq \operatorname{LH}(\{Ay_1, \dots, Ay_r\}).$$

Der Untervektorraum $\operatorname{LH}(\{Ay_1, \dots, Ay_r\})$ wird von r Elementen aufgespannt, aus denen sich eine Basis auswählen lässt. Somit ist seine Dimension höchstens r.

Es gilt also: $\operatorname{rg}(AB) = \dim \operatorname{Bild}(AB) \le \dim \operatorname{LH}(\{Ay_1, \dots, Ay_r\}) \le r$. □

Zusammenfassung von Abschnitt 2.4

(1) Ein *affiner Unterraum* ist ein verschobener Untervektorraum.

(2) Die Lösungsmenge eines linearen Gleichungssystems ist immer ein affiner Unterraum oder die leere Menge.

(3) Das Bild einer Matrix A besteht aus allen b, sodass $Ax = b$ mindestens eine Lösung x hat. Das Bild wird von den Spalten erzeugt.

(4) Der Rang einer Matrix ist die Dimension des Bildes.

2.5. Der Gauß-Algorithmus

Wir lernen nun den Gauß-Algorithmus[9] zum Lösen von linearen Gleichungssystemen kennen. Es gibt diesen Algorithmus in unzähligen verschiedenen Varianten und die hier gewählten Bezeichnungen werden keineswegs überall verwendet. Gemein ist allen Varianten, dass man zuerst ein gegebenes lineares Gleichungssystem in eine Art „Zeilenstufenform“ bringt, ohne dabei die Lösungsmenge zu verändern. Dann kann man dieser Stufenform ansehen, ob es überhaupt Lösungen gibt und falls es welche gibt, können diese dann in einem dritten Schritt ausgerechnet werden.

Ziel ist es, dieses Verfahren möglichst übersichtlich zu gestalten und in jedem Schritt sicher zu sein, dass keine Lösungen verloren gehen. Was genau eine *Zeilenstufenform* ist und wie genau man die Umformungsschritte schematisch aufschreibt, darüber gibt es unterschiedlichste Meinungen und entsprechend verschiedene Konventionen.

Gegeben sei ein lineares Gleichungssystem, bestehend aus m Gleichungen in n Unbekannten $x_1, \dots, x_n$:

$$\begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}.$$

Hierfür wird manchmal auch folgende Schreibweise verwendet:

x_1	x_2	$\cdots$	x_n	
$a_{1,1}$	$a_{1,2}$	$\dots$	$a_{1,n}$	b_1
$a_{2,1}$	$a_{2,2}$	$\dots$	$a_{2,n}$	b_2
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
$a_{m,1}$	$a_{m,2}$	$\dots$	$a_{m,n}$	b_m

Das grobe Vorgehen ist wie folgt:

[9] nach CARL FRIEDRICH GAUSS, deutscher Mathematiker, 1777–1855

Gaußsches Eliminationsverfahren (Gauß-Algorithmus), grober Plan

Der *Gauß-Algorithmus* oder das *Gaußsche Eliminationsverfahren* besteht aus drei bzw. vier Teilen:

(I) Gegeben ein lineares Gleichungssystem $Ax = b$, verwende man Zeilenumformungen, um ein äquivalentes lineares Gleichungssystem $Zx = c$ zu erhalten, sodass Z in Zeilenstufenform ist.

(II) Entscheide, ob es überhaupt eine Lösung gibt, wenn nicht: brich ab.

(IIb) Optional: Verwende Zeilenumformungen, um das Gleichungssystem in erweiterte Zeilenstufenform zu bringen.

(III) Führe freie Variable ein und löse nach den Pivot-Variablen auf.

Schritt (IIb) ist nicht unbedingt notwendig, vereinfacht die Rechnung aber häufig.

Bevor wir im Einzelnen beschreiben, wie die Schritte ausgeführt werden, eine Definition:

Definition 2.5.1 (Zeilenstufenform)**.** Gegeben sei eine Matrix $A \in \mathbb{R}^{m\times n}$ mit $m, n \in \mathbb{N}$.

(a) Eine *Nullzeile* einer Matrix ist eine Zeile, in der alle Einträge 0 sind.

(b) Wenn eine Zeile einer Matrix keine Nullzeile ist, dann nennen wir den ersten Eintrag dieser Zeile, der nicht 0 ist, den *Pivot-Eintrag* der Zeile. Eine *Pivot-Spalte* einer Matrix ist eine, in der sich ein Pivot-Eintrag befindet.

(c) Wir sagen, die Matrix A ist in *Zeilenstufenform*, wenn folgende beiden Eigenschaften erfüllt sind:

  (i) Wenn die i-te Zeile eine Nullzeile ist, dann sind alle Zeilen, die danach kommen, auch Nullzeilen.

  (ii) Wenn die i_1-te Zeile und die i_2-te Zeile keine Nullzeilen sind mit $i_1 < i_2$, dann befindet sich das Pivot-Element der i_2-ten Zeile rechts vom Pivot-Element der i_1-ten Zeile.

(d) Wir sagen die Matrix A hat *erweiterte Zeilenstufenform*, wenn zusätzlich gilt:

  (iii) Jedes Pivot-Element ist 1.

  (iv) Über jedem Pivot-Elementes stehen nur Nullen. (Äquivalent dazu ist: In den Pivot-Spalten stehen Standardbasisvektoren (siehe Definition 2.3.17))

(e) Ein lineares Gleichungssystem $Ax = b$ ist in (erweiterter) Zeilenstufenform, wenn die Koeffizientenmatrix A in (erweiterter) Zeilenstufenform ist.

Wie bereits oben geschrieben, gibt es in der Literatur auch andere Definitionen, was eine (erweiterte) Zeilenstufenform ist. Die Unterschiede sind aber oft eher gering.

Beispiel 2.5.2. Die folgenden Matrizen sind nicht in Zeilenstufenform:

$$\begin{pmatrix} 0 & 1 & 7 \\ -1 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 3 & 6 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 42 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 2 & 5 & 7 \\ 0 & 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}.$$

Die folgenden Matrizen sind in Zeilenstufenform, aber nicht in erweiterter Zeilenstufenform (die Pivot-Einträge sind eingekreist):

$$\begin{pmatrix} ① & 2 & 42 & -23 & 0 & \pi/2 \\ 0 & 0 & ⑦ & 128 & 256 & 0 \\ 0 & 0 & 0 & ① & 0 & \sqrt{13} \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} ① & 2 & 3 & 6 \\ 0 & 0 & ③ & 42 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & ② & 5 & 7 \\ 0 & 0 & ③ & 3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

$$\begin{pmatrix} ① & 1 & 1 \\ 0 & ① & 1 \\ 0 & 0 & ① \end{pmatrix}, \begin{pmatrix} ② \\ 0 \\ 0 \end{pmatrix}, ⑨.$$

Die $9 \in \mathbb{R} = \mathbb{R}^{1\times 1}$ ist hier als (1×1)-Matrix zu verstehen.

Hier nun noch ein paar Beispiele für Matrizen in erweiterter Zeilenstufenform (für beliebige $m, n \in \mathbb{N}$):

$$\begin{pmatrix} ① & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & ① & 0 & 256 & 0 \\ 0 & 0 & 0 & ① & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} ① & 2 & 0 & 6 \\ 0 & 0 & ① & 42 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & ① & 0 & 7 \\ 0 & 0 & ① & 3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \mathbb{1}_n, \mathbf{0}_{m\times n}, \begin{pmatrix} ① \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Auch hier sind die Pivot-Einträge eingekreist.

Gegeben ein lineares Gleichungssystem $Ax = b$, ist unser Ziel nun, dieses Gleichungssystem mit Äquivalenzumformungen in ein Gleichungssystem umzuformen, das in Zeilenstufenform ist. Hierzu werden wir nur die drei folgenden Operationen verwenden:

Definition 2.5.3 (Elementare Zeilenumformungen)**.** Die folgenden Operationen werden *elementare Zeilenumformungen* genannt:

(G1) Das Addieren des μ-fachen einer Zeile auf eine *andere* Zeile.

(G2) Das Vertauschen zweier Zeilen.

(G3) Das Multiplizieren einer Zeile mit einem Skalar $\lambda \neq 0$.

Die entsprechende Operation muss sowohl auf die Koeffizientenmatrix A als auch auf die rechte Seite b angewendet werden. Jede dieser drei Operationen ist umkehrbar, deshalb sind dies Äquivalenzumformungen und die Lösungsmenge ändert sich nicht.

Achtung: Spaltenumformungen, wie das Vertauschen zweier Spalten oder ähnliches, kann die Lösungsmenge ändern.

Lemma 2.5.4. *Jede der elementaren Zeilenumformungen lässt sich durch die Linksmultiplikation mit einer Matrix erreichen:*

(G1) Gegeben $i, j \in \{1, \dots, m\}$ mit $i \neq j$ und $\mu \in \mathbb{R}$. Das Addieren des μ-fachen der j-ten Zeile auf die i-te Zeile wird erreicht durch die Multiplikation von links mit der $(m \times m)$-Matrix

$$\begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & \ddots & & & \\ & & \mu & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix},$$

wobei der Eintrag μ in der i-ten Zeile und in der j-ten Spalte steht. Diese Matrix entsteht aus der Einheitsmatrix $\mathbb{1}_m$, indem man das μ-fache der j-ten Zeile auf die i-te Zeile addiert.

(G2) Gegeben $i, j \in \{1, \dots, m\}$ mit $i \neq j$. Das Vertauschen der i-ten Zeile mit der j-ten Zeile wird erreicht durch die Multiplikation von links mit der $(m \times m)$-Matrix

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

wobei auch diese Matrix aus der Einheitsmatrix $\mathbb{1}_m$ durch das Vertauschen der i-ten mit der j-ten Zeile entsteht.

(G3) Gegeben $i \in \{1, \dots, m\}$ und $\lambda \in \mathbb{R} \setminus \{0\}$. Das Multiplizieren der i-ten Zeile mit einem Skalar

$\lambda \neq 0$ wird erreicht durch die Multiplikation von links mit der Matrix

$$\begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & \lambda & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}.$$

Diese Matrix entsteht aus der Einheitsmatrix $\mathbb{1}_m$, wobei der i-te Eintrag auf der Diagonale durch λ ersetzt wurde.

Lemma 2.5.5. *Keine der elementaren Umformungen ändert den Rang einer Matrix.*

Beweis. Es sei $A \in \mathbb{R}^{m\times n}$ eine Matrix, die durch eine elementare Zeilenumformung zur Matrix $B \in \mathbb{R}^{m\times n}$ wird. Nach Lemma 2.5.4 gibt es eine Matrix $G \in \mathbb{R}^{m\times m}$ mit

$$B = GA.$$

Nach Lemma 2.4.15 gilt somit:

$$\mathrm{rg}B = \mathrm{rg}(GA) \leq \mathrm{rg}A.$$

Da Zeilenumformungen aber immer umkehrbar sind, gilt mit derselben Argumentation auch $\mathrm{rg}A \leq \mathrm{rg}B$.

Somit folgt: $\mathrm{rg}A = \mathrm{rg}B$. □

Gauß-Algorithmus Teil (I)

Es sei $A \in \mathbb{R}^{m\times n}$ eine Matrix und $b \in \mathbb{R}^m$.

- WENN die erste Spalte komplett nur Nullen enthält, ignoriere sie und wende diesen Algorithmus auf die Restmatrix an.
- WENN die erste Spalte einen Nichtnulleintrag erhält, wähle einen Nichtnulleintrag in der ersten Spalte und vertausche die entsprechende Zeile mit der ersten Zeile durch elementare Umformung (G2). (Falls der Eintrag ganz oben links nicht null ist, kann dieser Schritt entfallen).
- Benutze elementare Umformungen vom Typ (G1), um alle Einträge unter dem eben gewählten Pivot-Element auf Null zu setzen.
- Ignoriere von jetzt an die Spalte ganz links und die Zeile ganz oben und verfahre mit dem Rest wieder wie oben.

Die zu betrachtenden Restmatrizen werden immer kleiner und nach endlich vielen Schritten ist die Matrix in Zeilenstufenform. Jede der Umformungen muss auf die Matrix A und auf den Vektor b angewendet werden – nur so ist sichergestellt, dass die Lösungsmenge sich nicht ändert.

Nun wollen wir herausfinden, ob es eine Lösung gibt.

Wir haben gesehen, dass ein lineares Gleichungssystem $Ax = b$ genau dann eine Lösung hat, wenn $b \in \text{Bild}(A)$ (siehe Definition 2.4.11). Wenn die Matrix A in Zeilenstufenform ist, dann ist $\text{Bild}(A)$ besonders einfach zu beschreiben:

Lemma 2.5.6. *Es sei A eine Matrix in Zeilenstufenform mit $r \in \mathbb{N}_0$ Pivot-Einträgen. Dann sind die Pivot-Spalten der Matrix eine Basis für das Bild der Matrix*

$$\text{Bild}(A) = \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} \middle| \forall j > r : x_j = 0 \right\} = \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_r \\ 0 \\ \vdots \\ 0 \end{pmatrix} \middle| x_j \in \mathbb{R} \right\}.$$

und der Rang der Matrix ist die Anzahl der Pivot-Elemente, d.h.

$$\text{rg}(A) = r.$$

Beweis. Die Matrix $A \in \mathbb{R}^{m \times n}$ ist in der folgenden Form:

$$A = \begin{pmatrix} 0 & \dots & 0 & a_{1,j_1} & * & \dots & \dots & \dots & \dots & \dots & \dots & \dots & * \\ 0 & \dots & \dots & \dots & \dots & 0 & a_{2,j_2} & * & \dots & \dots & \dots & \dots & * \\ \vdots & & & & & & & & & & & & \vdots \\ 0 & \dots & \dots & \dots & \dots & \dots & \dots & \dots & 0 & a_{r,j_r} & * & \dots & * \\ 0 & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & 0 \\ \vdots & & & & & & & & & & & & \vdots \\ 0 & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & 0 \end{pmatrix}.$$

Aus der Definition der Zeilenstufenform folgt, dass die Pivot-Spalten linear unabhängig sind, d.h. sie bilden eine Basis für den Untervektorraum W, der von ihnen aufgespannt wird. Es gilt $W \subseteq \text{Bild}(A)$. Wir definieren

$$U := \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \middle| \forall j > r : x_j = 0 \right\}.$$

Aus der Definition der Zeilenstufenform folgt unmittelbar, dass sämtliche Spalten der Matrix in U liegen und somit die lineare Hülle der Spalten eine Teilmenge von U ist. Somit gilt $\text{Bild}(A) \subseteq U$. Wenn wir nun noch zeigen können, dass $U \subseteq W$, dann haben wir drei Inklusionen gezeigt und somit müssen alle drei Mengen gleich sein. Daraus folgen dann alle Behauptungen.

Es sei also $b \in U$, d.h. b ist ein Vektor in $\mathbb{R}^m$, bei dem nur die ersten r Komponenten ungleich Null sein dürfen, also

$$b = \begin{pmatrix} b_1 \\ \vdots \\ b_r \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Wenn wir nun das $\frac{b_r}{a_{r,j_r}}$-fache der r-ten Pivot-Spalte abziehen, erhalten wir einen Vektor, der in der r-ten Komponente Null ist. Nun subtrahieren wir ein entsprechendes Vielfaches der $(r-1)$-ten Pivot-Spalte und setzen dies solange fort, bis wir den Nullvektor erreichen. Das zeigt, dass man b durch Subtrahieren einer Linearkombination der Pivot-Spalten in den Nullvektor überführen kann. Also ist b eine Linearkombination der Pivot-Spalten. Das zeigt $b \in W$. □

Aus Lemma 2.5.6 folgt nun unmittelbar das folgende Verfahren, um festzustellen, ob ein lineares Gleichungssystem in Zeilenstufenform eine Lösung hat oder nicht:

Gauß-Algorithmus Teil (II)

Es sei $A \in \mathbb{R}^{m\times n}$ eine Matrix in Zeilenstufenform mit r Pivot-Zeilen und $b \in \mathbb{R}^m$.

- WENN $(\forall j > r : b_j = 0)$, DANN ist $b \in \text{Bild}(A)$ und es gibt es eine Lösung;
- WENN $(\exists j > r : b_j \neq 0)$, DANN ist $b \notin \text{Bild}(A)$ und es gibt keine Lösung. Algorithmus abbrechen.

Wenn $b \in \text{Bild}(A)$, dann sind wir in der folgenden Situation:

x_1	…	x_{j_1-1}	x_{j_1}	…	…	x_{j_2}	…	…	x_{j_r}	…	…	x_n	
0	…	0	a_{1,j_1}	$*$	…	…	…	…	…	…	…	$*$	b_1
0	…	…	…	…	0	a_{2,j_2}	$*$	…	…	…	…	$*$	b_2
⋮												⋮	⋮
0	…	…	…	…	…	…	…	0	a_{r,j_r}	$*$	…	$*$	b_r
0	…	…	…	…	…	…	…	…	…	…	…	0	0
⋮												⋮	0
0	…	…	…	…	…	…	…	…	…	…	…	0	0

Ebenfalls aus Lemma 2.5.6 folgt nun, wie man die Lösungen bekommt: Wenn $b \in \text{Bild}(A)$, dann ist b in der linearen Hülle der Pivot-Spalten und die Komponenten des Lösungsvektors sind genau die skalaren Vielfache, die man benötigt, um b mit den Pivot-Spalten darzustellen. Man kann also eine Lösung erhalten, indem man alle Nicht-Pivot-Variablen auf 0 setzt und dann von unten nach oben die einzelnen Gleichungen nach den Pivot-Variablen auflöst. Wenn wir aber gleich alle Lösungen finden wollen, gehen wir folgendermaßen vor:

Gauß-Algorithmus Teil (III)

Gegeben sei eine Matrix $A \in \mathbb{R}^{m\times n}$ in Zeilenstufenform mit r Pivot-Zeilen und $b \in \text{Bild}(A)$.

- Jede der $(n-r)$ Nicht-Pivot-Variablen – falls vorhanden – wird mit einem Parameter versehen.
- Löse jede der Gleichungen von unten nach oben nach den Pivot-Variablen auf.
- Stelle den allgemeinen Lösungsvektor $(x_1,\dots,x_n)^\top$ in Abhängigkeit der $(n-r)$ Parameter auf.
- Schreibe den Lösungsvektor als Summe eines konstanten Vektors x_0 (also ohne Parameter) und einer Linearkombination von Vektoren $v_1,\dots v_{n-r}$, wobei die Parameter die Skalare sind. Die Lösungsmenge ist nun der affine Unterraum

$$x_0 + \text{LH}(v_1,\dots,v_{n-r}).$$

Die Vektoren $v_1,\dots,v_{n-r}$ sind nach Konstruktion linear unabhängig, also eine Basis für $\text{LH}(v_1,\dots,v_{n-r})$.
Die Dimension der Lösungsmenge ist also $n-r$, die Differenz zwischen dem Rang der Matrix und der Anzahl der Spalten.

Bemerkung 2.5.7. (a) Falls das lineare Gleichungssystem homogen ist, d.h. falls $b = 0$, dann ist der Vektor $x_0 = 0$ und der affine Unterraum ist sogar ein Untervektorraum, nämlich der Kern von A.

(b) Falls $n = r$, so ist die Dimension der Lösungsmenge 0, d.h. sie besteht nur aus einem Punkt – falls überhaupt eine existiert. Anders formuliert bedeutet das, dass die Abbildung

$$\mathbb{R}^n \to \mathbb{R}^m, \quad x \mapsto Ax$$

für jeden Punkt $b \in \mathbb{R}^m$ in der Zielmenge höchstens ein Urbild hat. Die Abbildung ist also injektiv (siehe Bemerkung 1.3.10).

Oft ist es sinnvoll, zwischen Schritt (II) und (III) noch den folgenden Schritt einzuschieben, der eine Zeilenstufenform in eine erweiterte Zeilenstufenform überführt:

Gauß-Algorithmus Teil (IIb)

Es sei $A \in \mathbb{R}^{m\times n}$ eine Matrix in Zeilenstufenform mit r Pivot-Zeilen und $b \in \mathbb{R}^m$.

- Benutze elementare Umformungen vom Typ (G3), um alle Pivot-Einträge zu 1 zu machen.
- Benutze elementare Umformungen vom Typ (G1), um alle Einträge über den Pivot-Einträgen zu 0 zu machen. Beginne hierbei bei dem Letzten Pivot-Element.

Die Matrix ist nun in erweiterter Zeilenstufenform.

Beispiel. Gegeben seien die Matrix

$$A := \begin{pmatrix} 7 & 0 & -7 & 0 \\ 8 & 1 & -5 & -2 \\ 0 & 1 & -3 & 0 \\ 0 & 3 & -6 & -1 \end{pmatrix}$$

und der Vektor

$$b := \begin{pmatrix} 0 \\ 2 \\ 0 \\ 1 \end{pmatrix}.$$

Wir wollen nun alle $x \in \mathbb{R}^4$ finden mit $Ax = b$. Dazu verwenden wir den Gauß-Algorithmus:

x_1	x_2	x_3	x_4			
7	0	−7	0	0	$\vert \cdot 1/7$	$(G3)$
8	1	−5	−2	2		
0	1	−3	0	0		
0	3	−6	−1	1		
1	0	−1	0	0	$\vert \cdot(-8)$ auf Zeile 2	$(G1)$
8	1	−5	−2	2		
0	1	−3	0	0		
0	3	−6	−1	1		
①	0	−1	0	0		
0	1	3	−2	2	$\vert \cdot(-1)$ auf Zeile 3	$(G1)$
0	1	−3	0	0		
0	3	−6	−1	1		
①	0	−1	0	0		
0	1	3	−2	2	$\vert \cdot(-3)$ auf Zeile 4	$(G1)$
0	0	−6	2	−2		
0	3	−6	−1	1		
①	0	−1	0	0		
0	①	3	−2	2		
0	0	−6	2	−2	$\vert \cdot(-1/2)$	$(G3)$
0	0	−15	5	−5	$\vert \cdot(-1/5)$	$(G3)$
①	0	−1	0	0		
0	①	3	−2	2		
0	0	3	−1	1	$\vert \cdot(-1)$ auf Zeile 4	$(G1)$
0	0	3	−1	1		
①	0	−1	0	0		
0	①	3	−2	2		
0	0	③	−1	1		
0	0	0	0	0		

Nun ist unsere Koeffizientenmatrix in Zeilenstufenform (Schritt I ist abgeschlossen).

Da die Matrix eine Nullzeile hat, schauen wir, ob die rechte Seite in der entsprechenden Zeile auch 0 ist. Das ist der Fall, also können wir zu diesem Zeitpunkt schon sagen, dass das Gleichungssystem lösbar ist (Schritt II ist abgeschlossen).

Da wir $r = 3$ Nichtnullzeilen haben, hat die Matrix A Rang 3 und die Lösungsmenge ist 1-dimensional (als affiner Unterraum gesehen).

Nun könnten wir direkt zu Schritt III weitergehen. Stattdessen schieben wir jedoch noch den Schritt IIb dazwischen und transformieren das lineare Gleichungssystem in eine erweiterte Zeilenstufenform:

x_1	x_2	x_3	x_4			
①	0	-1	0	0		
0	①	3	-2	2		
0	0	③	-1	1	$\mid \cdot \frac{1}{3}$	$(G3)$
0	0	0	0	0		
①	0	-1	0	0		
0	①	3	-2	2		
0	0	①	$-\frac{1}{3}$	$\frac{1}{3}$	$\mid \cdot(-3)$ auf Zeile 2	$(G1)$
0	0	0	0	0		
①	0	-1	0	0		
0	①	0	-1	1		
0	0	①	$-\frac{1}{3}$	$\frac{1}{3}$	$\mid \cdot(1)$ auf Zeile 1	$(G1)$
0	0	0	0	0		
①	0	0	$-\frac{1}{3}$	$\frac{1}{3}$		
0	①	0	-1	1		
0	0	①	$-\frac{1}{3}$	$\frac{1}{3}$		
0	0	0	0	0		

Dies ist nun eine erweiterte Zeilenstufenform.

Nun zu Schritt III: Die Pivot-Variablen sind x_1, x_2, x_3. Es gibt genau eine Nicht-Pivot-Variable: x_4.

Wir setzen nun $x_4 = t \in \mathbb{R}$.

Dann lösen wir das lineare Gleichungssystem von unten nach oben auf: Die dritte Gleichung lautet:

$$1x_3 - \frac{1}{3}t = \frac{1}{3}.$$

Also ist $x_3 = \frac{1}{3} + t\frac{1}{3}$.

Die zweite Gleichung lautet:

$$1x_2 + (-1)t = 1.$$

Also ist $x_2 = 1 + t$.

Die erste Gleichung schließlich lautet:

$$1x_1 + \left(-\frac{1}{3}\right)t = \frac{1}{3}.$$

Also ist $x_1 = \frac{1}{3} + t\frac{1}{3}$.

Der Lösungsvektor lautet nun also:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} \frac{1}{3} + t\frac{1}{3} \\ 1+t \\ \frac{1}{3} + t\frac{1}{3} \\ t \end{pmatrix} = \begin{pmatrix} \frac{1}{3} \\ 1 \\ \frac{1}{3} \\ 0 \end{pmatrix} + \begin{pmatrix} t\frac{1}{3} \\ t \\ t\frac{1}{3} \\ t \end{pmatrix} = \begin{pmatrix} \frac{1}{3} \\ 1 \\ \frac{1}{3} \\ 0 \end{pmatrix} + t \begin{pmatrix} \frac{1}{3} \\ 1 \\ \frac{1}{3} \\ 1 \end{pmatrix}.$$

Also ist die Lösungsmenge des linearen Gleichungssystems der affine Untervektorraum

$$\left\{x \in \mathbb{R}^4 \;\middle|\; Ax = b\right\} = \begin{pmatrix} \frac{1}{3} \\ 1 \\ \frac{1}{3} \\ 0 \end{pmatrix} + \mathrm{LH}\left(\left\{\begin{pmatrix} \frac{1}{3} \\ 1 \\ \frac{1}{3} \\ 1 \end{pmatrix}\right\}\right).$$

Hierbei ist $\begin{pmatrix} \frac{1}{3} \\ 1 \\ \frac{1}{3} \\ 0 \end{pmatrix}$ eine spezielle Lösung und der Kern von A ist gegeben durch $\mathrm{LH}\left(\left\{\begin{pmatrix} \frac{1}{3} \\ 1 \\ \frac{1}{3} \\ 1 \end{pmatrix}\right\}\right)$. Da sich die lineare Hülle nicht ändert, wenn Vektoren mit einer Zahl ungleich 0 skaliert werden, können wir die gleiche Lösungsmenge auch anders darstellen:

$$\left\{x \in \mathbb{R}^4 \;\middle|\; Ax = b\right\} = \begin{pmatrix} \frac{1}{3} \\ 1 \\ \frac{1}{3} \\ 0 \end{pmatrix} + \mathrm{LH}\left(\left\{\begin{pmatrix} 1 \\ 3 \\ 1 \\ 3 \end{pmatrix}\right\}\right).$$

Zum Abschluss dieses Kapitels noch zwei praktische Sätze, die wir hier quasi als Nebenprodukt bekommen und die – je nach Quelle – beide als *Rangsatz* bezeichnet werden.

Satz 2.5.8 (Rangsatz)**.** *Für eine Matrix $A \in \mathbb{R}^{m\times n}$ gilt*

$$\mathrm{rg}(A) = \mathrm{rg}\left(A^\top\right).$$

*Anders formuliert: Die maximale Anzahl linear unabhängiger Spalten ist gleich der maximalen Anzahl linear unabhängiger Zeilen („*Spaltenrang = Zeilenrang*“).*

Beweis. Zuerst einmal überzeugen wir uns, dass die Aussage korrekt ist, falls A in Zeilenstufenform vorliegt. Es sei r die Anzahl der Nichtnullzeilen. Da jede dieser r Zeilen mehr führende Nullen als die Zeile darüber hat, sind diese Zeilen linear unabhängig. Somit hat der „Zeilenrang“ von A den Wert r. Da r außerdem auch der „Spalten-“Rang von A ist (Lemma 2.5.6), folgt daraus die Behauptung.

Die eigentliche Frage ist also: Wieso gilt diese Aussage auch für eine Matrix $A \in \mathbb{R}^{m\times n}$, die nicht in Zeilenstufenform ist?

Wir wissen aus Teil I des Gauß-Verfahrens, dass es möglich ist, endlich viele (sagen wir $N \in \mathbb{N}$ viele) elementare Zeilenumformungen auf A anzuwenden, um eine Matrix Z in Zeilenstufenform zu erhalten. Nach Lemma 2.5.4 lässt sich jede dieser elementaren Zeilenumformungen als Multiplikation mit einer Matrix $G_j \in \mathbb{R}^{m\times m}$ von links schreiben. Es gilt also:

$$Z = G_N \cdot G_{N-1} \cdots G_2 \cdot G_1 A.$$

Wir setzen nun $G := G_N \cdot G_{N-1} \cdots G_2 \cdot G_1 \in \mathbb{R}^{m \times m}$ und können somit schreiben:

$$Z = GA.$$

Nun gilt

$$\begin{aligned} \operatorname{rg}(A) &= \operatorname{rg}(Z) \quad \text{(Lemma 2.5.5)} \\ &= \operatorname{rg}(Z^\top) \quad \text{(weil } Z \text{ Zeilenstufenform hat)} \\ &= \operatorname{rg}((GA)^\top) \\ &= \operatorname{rg}(A^\top G^\top) \quad \text{(Lemma 2.2.13)} \\ &\leq \operatorname{rg}(A^\top). \quad \text{(Lemma 2.4.15)} \end{aligned}$$

Also gilt $\operatorname{rg}(A) \leq \operatorname{rg}\left(A^\top\right)$. Durch Anwenden derselben Argumentation auf $A^\top$ folgt dann

$$\operatorname{rg}\left(A^\top\right) \leq \operatorname{rg}(A). \qquad \square$$

Satz 2.5.9 (Dimensionsformel)**.**
Für eine Matrix $A \in \mathbb{R}^{m \times n}$ gilt

$$\operatorname{rg} A + \dim \ker A = n.$$

Anders formuliert: Die Dimension von Bild$A$ *und die Dimension von* $\ker A$ *addieren sich zur Dimension von* $\mathbb{R}^n$.

Beweis. Die Matrix $A \in \mathbb{R}^{m \times n}$ lässt sich durch endlich viele elementare Zeilenumformungen in Zeilenstufenform Z bringen.

Elementare Zeilenumformungen ändern die Lösungsmenge des homogenen linearen Gleichungssystems nicht. Also ist $\ker(A) = \ker(Z)$.

Elementare Zeilenumformungen ändern zwar eventuell das Bild, aber nicht den Rang (Lemma 2.5.5), also gilt: $\operatorname{rg}(A) = \operatorname{rg}(Z)$.

Die Matrix Z in Zeilenstufenform habe r Nichtnullzeilen. Dann gilt nach Lemma 2.5.6, dass $\operatorname{rg}(Z) = r$. Das Gleichungssystem hat nun $n - r$ Nicht-Pivot-Variablen, also ist $\dim \ker(Z) = n - r$. Insgesamt gilt also:

$$\operatorname{rg} A + \dim \ker A = \operatorname{rg} Z + \dim \ker Z = r + (n - r) = n. \qquad \square$$

Zusammenfassung von Abschnitt 2.5

(1) Jede Matrix lässt sich mit elementaren Zeilenumformungen in Zeilenstufenform bringen.

(2) Jede elementare Zeilenumformung entspricht einer Matrixmultiplikation von links.

(3) Der Rang einer Matrix ändert sich bei elementaren Zeilenumformungen nicht.

(4) Es gelten Rangsatz und Dimensionsformel.

3. Algebraische Strukturen

3.1. Halbgruppen

In diesem Kapitel möchten wir uns mit einigen algebraischen Strukturen befassen – bevor wir uns im nächsten Kapitel dann auf die für uns interessantesten Strukturen, nämlich Vektorräume, konzentrieren.

Definition 3.1.1 (Binäre Operation)**.** Es sei S eine Menge. Eine Abbildung

$$* : S \times S \to S$$

nennen wir eine *binäre Verknüpfung* (*zweistellige Operation*, *binäre Operation*, *zweistellige Verknüpfung*).

Eine binäre Verknüpfung ist also eine Operation, die zwei Elemente aus einer Menge S als Eingabe bekommt und dann ein Element aus S zurückgibt. Der Definitionsbereich von $*$ ist das kartesische Produkt[1] von S mit sich selbst:

$$S \times S = \{(x, y) \mid x, y \in S\}$$

Wenn $(x, y) \in S \times S$ im Definitionsbereich von $*$ gewählt ist, schreiben wir für das Bild unter der Abbildung $*$ für gewöhnlich

$$x * y := *(x, y).$$

Notation. Wenn X endlich ist, dann lässt sich eine binäre Verknüpfung eindeutig durch eine Tabelle, auch *Verknüpfungstafel* oder *Cayley-Tafel*[2] genannt, beschreiben. Nehmen wir als Beispiel an, die 6-elementige Menge $D := \{i, s_1, s_2, s_3, g, u\}$ sei gegeben, wobei i, s_1, s_2, s_3, g, u irgendwelche Objekte sind. Dann könnte eine binäre Operation

$$\star : D \times D \to D$$

zum Beispiel so aussehen:

$\star$	i	s_1	s_2	s_3	g	u
i	i	s_1	s_2	s_3	g	u
s_1	s_1	i	g	u	s_2	s_3
s_2	s_2	u	i	g	s_3	s_1
s_3	s_3	g	u	i	s_1	s_2
g	g	s_3	s_1	s_2	u	i
u	u	s_2	s_3	s_1	i	g

An dieser Tafel kann man nun zum Beispiel ablesen, dass $s_2 \star g = s_3$ und $u \star u = g$ gilt.

[1] siehe Notation 1.2.17

[2] nach ARTHUR CAYLEY, engl. Mathematiker, 1821–1895

Definition 3.1.2 (Halbgruppe)**.** (a) Es sei S eine Menge. Eine binäre Verknüpfung $* : S \times S \to S$ auf S heißt *assoziativ*, wenn

$$\forall x, y, z \in S : (x * y) * z = x * (y * z).$$

(b) Eine *Halbgruppe* ist ein Paar $(S, *)$ bestehend aus einer Menge S und einer assoziativen binären Verknüpfung auf S.

(c) Eine *kommutative Halbgruppe* ist eine Halbgruppe $(S, *)$, in der zusätzlich das Kommutativgesetz gilt, d.h.

$$\forall x, y \in S : x * y = y * x.$$

Beispiel 3.1.3. (a) Die reellen Zahlen mit der Addition $(\mathbb{R}, +)$ sind eine kommutative Halbgruppe, da Addition von reellen Zahlen assoziativ und kommutativ ist.

(b) Die reellen Zahlen mit der Multiplikation $(\mathbb{R}, \cdot)$ sind eine kommutative Halbgruppe, da Multiplikation von reellen Zahlen assoziativ und kommutativ ist.

(c) Für festes $m, n \in \mathbb{N}$ ist die Menge aller $(m \times n)$-Matrizen mit reellen Einträgen mit der Addition eine kommutative Halbgruppe: $(\mathbb{R}^{m \times n}, +)$, weil Matrixaddition assoziativ und kommutativ ist (Lemma 2.2.7).

(d) Als Spezialfall erhalten wir: Für $n \in \mathbb{N}$ ist $(\mathbb{R}^n, +)$ eine kommutative Halbgruppe.

(e) Für jedes $n \in \mathbb{N}$ ist Menge der $(n \times n)$-Matrizen mit der Multiplikation eine Halbgruppe: $(\mathbb{R}^{n \times n}, \cdot)$, weil Matrixmultiplikation assoziativ ist (Lemma 2.2.12). Für $n > 1$ ist diese Halbgruppe nicht kommutativ.

(f) Für $m \neq n$ ist $(\mathbb{R}^{m \times n}, \cdot)$ keine Halbgruppe, weil das Produkt von zwei Elementen nicht mehr in dieser Menge ist

(g) Die natürlichen Zahlen $\mathbb{N}$, die ganzen Zahlen $\mathbb{Z}$ und die rationalen Zahlen $\mathbb{Q}$ werden jeweils zu kommutativen Halbgruppen, wenn man sie entweder mit + oder $\cdot$ versieht.

(h) Die natürlichen Zahlen $(\mathbb{N}, \uparrow)$ mit der Operation $x \uparrow y = x^y$ sind keine Halbgruppe, weil Potenzieren nicht assoziativ ist.

(i) Die Menge $(\mathbb{R}, *)$ mit $x * y := \frac{x+y}{2}$ ist auch keine Halbgruppe, weil $* : \mathbb{R} \times \mathbb{R} \to \mathbb{R}, \quad (x, y) \mapsto \frac{x+y}{2}$ nicht assoziativ ist.

(j) Für eine Menge X sei $S := \mathscr{P}(X) := \{A \mid A \subseteq X\}$ die Menge aller Teilmengen von X. Dann ist $(\mathscr{P}(X), \cap)$ eine kommutative Halbgruppe, denn die Durchschnittsbildung $\cap : (\mathscr{P}(X)) \times (\mathscr{P}(X)) \to \mathscr{P}(X)$ ist assoziativ und kommutativ (Satz 1.2.15).

(k) Für eine Menge X sei $S := \mathscr{P}(X) := \{A \mid A \subseteq X\}$ die Menge aller Teilmengen von X. Dann ist $(\mathscr{P}(X), \cup)$ eine kommutative Halbgruppe, denn die Vereinigung $\cup : (\mathscr{P}(X)) \times (\mathscr{P}(X)) \to \mathscr{P}(X)$ ist assoziativ und kommutativ (Satz 1.2.15).

(l) Für eine Menge X sei $S := X^X = \{f \mid f : X \to X\}$ die Menge aller Selbstabbildungen von X. Dann ist $(X^X, \circ)$ eine Halbgruppe, denn Verkettung von Abbildungen $\circ : X^X \times X^X \to X^X$ ist assoziativ (Lemma 1.3.7). Diese Halbgruppe ist im Allgemeinen nicht kommutativ.

(m) Die binäre Verknüpfung $\star : D \times D \to D$ auf Seite 61 ist assoziativ. Dies lässt sich leider nicht direkt an der Cayley-Tafel ablesen, lässt sich aber mit etwas Arbeit überprüfen. Die Verknüpfung ist aber nicht kommutativ, weil sonst die Tafel symmetrisch sein müsste. Also ist $(\{i, s_1, s_2, s_3, g, u\}, \star)$ mit der dort definierten Verknüpfung $\star$ eine nichtkommutative Halbgruppe.

Definition 3.1.4. Gegeben sei eine Halbgruppe $(S, *)$. Dann heißt $e \in S$ *neutrales Element* (oder *Neutralelement*) für $*$, wenn

$$\forall x \in S : x * e = x = e * x.$$

Eine Halbgruppe $(S, *)$, die ein neutrales Element besitzt, heißt *Monoid*.

Lemma 3.1.5. *Wenn eine Halbgruppe ein neutrales Element besitzt, dann ist es eindeutig.*

Beweis. Es sei $(S, *)$ eine Halbgruppe und $e, f \in S$ seien beide neutrale Elemente für $*$. Dann gilt

$$e = e * f = f.$$ □

Bemerkung 3.1.6. Die Halbgruppe $(\mathbb{N}, +)$ besitzt kein neutrales Element. Alle anderen Halbgruppen aus Beispiel 3.1.3 besitzen jeweils ein neutrales Element, sind also sogar Monoide.

Definition 3.1.7 (Invertierbarkeit)**.** Gegeben sei ein Monoid $(S, *)$ mit neutralem Element e. Ein Element $x \in S$ heißt *invertierbar* bezüglich $*$, wenn es ein Element $y \in S$ gibt mit

$$x * y = e = y * x.$$

In diesem Fall heißt y *Inverses* von x bezüglich $*$. Die Menge aller invertierbaren Elemente bezeichnen wir mit $(S, *)^\times$. Für jedes $x \in S$ ist das Inverse von x eindeutig (siehe Lemma 3.1.8) und es wird für gewöhnlich mit x^{-1} bezeichnet.

Lemma 3.1.8. *Wenn ein Element x in einem Monoid invertierbar ist, dann ist das Inverse von x eindeutig.*

Beweis. Es sei $(S, *)$ ein Monoid mit neutralem Element e und $x \in S$ ein Element. Angenommen, es gibt zwei Inverse $y, z \in S$, d.h. es gilt:

$$x * y = e = y * x \quad \text{und} \quad x * z = e = z * x$$

Dann folgt daraus, dass $y = z$ gilt, weil

$$\begin{aligned} y &= y * e \\ &= y * (x * z) \\ &= (y * x) * z \\ &= e * z \\ &= z. \end{aligned}$$ □

Beispiel 3.1.9. Gegeben sei eine Menge X und das Monoid $(S = X^X, \circ)$ der Selbstabbildungen von X (Beispiel 3.1.3). Dann ist das neutrale Element bezüglich $\circ$ die identische Abbildung $\mathrm{id}_X \in S$ (siehe Notation 1.3.8). Eine Abbildung $f \in S$ ist invertierbar bezüglich $\circ$ genau dann, wenn f bijektiv ist (Satz 1.3.12). Das Inverse von f bezüglich $\circ$ ist dann genau die Umkehrabbildung (Definition 1.3.11).

Lemma 3.1.10. *Es sei $(S,*)$ ein Monoid. Für zwei invertierbare Elemente $x,y \in S$ gilt: Das Produkt $(x*y)$ ist auch invertierbar und*

$$(x*y)^{-1} = y^{-1} * x^{-1}.$$

Beweis. Gegeben seien zwei Elemente $x,y \in S^\times$. Dann existieren die dazugehörigen Inverse $x^{-1}, y^{-1} \in S$. Behauptung: $a := x*y$ ist invertierbar mit Inversem $b := y^{-1} * x^{-1}$. Dazu multiplizieren wir

$$\begin{aligned} a*b &= \big(x*y\big) * (y^{-1} * x^{-1}) \\ &= x * \big(y * (y^{-1} * x^{-1})\big) \\ &= x * \big((y * y^{-1}) * x^{-1}\big) \\ &= x * \big(e * x^{-1}\big) \\ &= x * x^{-1} \\ &= e. \end{aligned}$$

Ebenso sieht man:

$$\begin{aligned} b*a &= (y^{-1} * x^{-1}) * \big(x*y\big) \\ &= \big((y^{-1} * x^{-1}) * x\big) * y \\ &= \big(y^{-1} * (x^{-1} * x)\big) * y \\ &= \big(y^{-1} * e\big) * y \\ &= y^{-1} * y \\ &= e. \end{aligned}$$

Also haben wir gezeigt, dass $a*b = e = b*a$. Also ist a invertierbar mit $a^{-1} = b$. □

Achtung: Im Allgemeinen folgt aus der Existenz eines Linksinversen noch nicht die Existenz eines Inversen.

Beispiel 3.1.11. Gegeben seien die Funktionen

$$f : \mathbb{N} \to \mathbb{N}, \quad k \mapsto k+1$$

und

$$g : \mathbb{N} \to \mathbb{N}, \quad k \mapsto \begin{cases} k-1 & k \geq 2 \\ 1 & k = 1. \end{cases}$$

Die Funktionen f und g sind Elemente im Monoid $(\mathbb{N}^\mathbb{N}, \circ)$ aller Selbstabbildungen der natürlichen Zahlen. Beide Funktionen sind nicht invertierbar bezüglich $\circ$, weil sie beide nicht bijektiv sind. Nichtsdestotrotz gilt

$$g \circ f = \mathrm{id}_\mathbb{N}.$$

Dies zeigt, dass in einem Monoid im Allgemeinen $x*y = e$ nicht ausreicht, damit x und y invertierbar sind.

Ähnliche Beispiele gibt es bei Operatoren auf unendlich dimensionalen Räumen, wie sie zum Beispiel in der Quantenmechanik allgegenwärtig sind.

Glücklicherweise tritt dieses Phänomen im Monoid $(\mathbb{R}^{n\times n}, \cdot)$, der $(n \times n)$-Matrizen mit der Matrixmultiplikation nicht auf, d.h. bei einer Matrix folgt aus der Existenz eines „einseitigen Inversen“ bereits die Invertierbarkeit (dies werden wir in Satz 4.3.4 beweisen).

Notation 3.1.12 (Potenzen)**.** Es sei $(S, *)$ ein Monoid mit neutralem Element e und $x \in S$. Dann definieren wir

$$x^n := \underbrace{x * \cdots * x}_{n} \quad \text{für alle } n \in \mathbb{N}$$

Desweiteren setzen wir

$$x^0 := e$$

und – falls x invertierbar ist –

$$x^{-n} := (x^{-1})^n = \underbrace{x^{-1} * \cdots * x^{-1}}_{n} \quad \text{für alle } n \in \mathbb{N}.$$

Somit ist für invertierbare $x \in S$ und $k \in \mathbb{Z}$ die Potenz x^k definiert. Es gelten die folgenden Potenzgesetze:

$$x^{k+l} = x^k * x^l \quad \text{und} \quad (x^k)^l = x^{kl}.$$

Hierbei können k, l beliebige Zahlen aus $\mathbb{N}_0$ sein – im Falle eines invertierbaren Elements sogar aus $\mathbb{Z}$.

Achtung: $(x * y)^k$ ist im Allgemeinen nicht das Gleiche wie $x^k * y^k$.

Zusammenfassung von Abschnitt 3.1

(1) Eine Halbgruppe $(S, *)$ ist eine Menge S, zusammen mit einer assoziativen binären Verknüpfung $*$.

(2) Ein Monoid ist eine Halbgruppe mit neutralem Element, das immer eindeutig ist.

(3) Wenn ein Element invertierbar ist, dann ist das Inverse eindeutig.

3.2. Gruppen

Gruppen sind in der modernen Mathematik allgegenwärtig und werden zur Untersuchung von Symmetrien von geometrischen Objekten oder physikalischen Systemen eingesetzt.

Definition 3.2.1 (Gruppe)**.** Ein Monoid $(G, *)$, in dem jedes Element invertierbar ist, nennt man eine *Gruppe*. Das heißt: Eine *Gruppe* ist ein Paar $(G, *)$ bestehend aus einer Menge G und einer Abbildung $*$, sodass

- $* : G \times G \to G$
- $\forall x, y, z \in G : (x * y) * z = x * (y * z)$.
- $\exists e \in G : x * e = x = e * x$

- $\forall x \in G : \exists y \in G : x * y = e = y * x$.

Eine kommutative Gruppe nennt man auch eine *abelsche Gruppe*[3]

Lemma 3.2.2 (Einheitengruppe)**.** *Es sei $(S,*)$ ein Monoid. Dann ist $(S^\times,*)$ eine Gruppe, genannt die* Einheitengruppe *von $(S,*)$. Hierbei bezeichnet $S^\times := (S,*)^\times$ die Menge aller bezüglich $*$ invertierbaren Elemente und*

$$* : S^\times \times S^\times \to S^\times, \quad (x,y) \mapsto x * y$$

ist die Einschränkung[4] *der ursprünglichen Operation $* : S \times S \to S$.*

Beweis. Nach Lemma 3.1.10 ist für $(x,y) \in S^\times \times S^\times$ auch $x * y \in S^\times$. Also ist die Einschränkung (und Koeinschränkung)

$$* : S^\times \times S^\times \to S^\times, \quad (x,y) \mapsto x * y$$

eine wohldefinierte binäre Verknüpfung auf der Teilmenge $S^\times$. Da die ursprüngliche Verknüpfung assoziativ war, ist auch die eingeschränkte Verknüpfung assoziativ.

Das neutrale Element e ist invertierbar, weil $e * e = e = e * e$ gilt. Also gilt $e \in S^\times$ und $e^{-1} = e$.

Wenn $x \in S^x$, dann ist das Inverse x^{-1} wieder invertierbar mit Inversem $(x^{-1})^{-1} = x$. Also ist $S^\times$ abgeschlossen unter Inversenbildung.

Also ist $S^\times$ eine Gruppe. □

Beispiel 3.2.3. (a) Die Halbgruppe $(\mathbb{N},+)$ ist nicht einmal ein Monoid – also insbesondere keine Gruppe.

(b) Die Einheitengruppe von $(\mathbb{N}_0,+)$ ist die einelementige Gruppe $(\{0\},+)$.

(c) $(\mathbb{Z},+)$, $(\mathbb{Q},+)$, $(\mathbb{R},+)$ sind abelsche Gruppen, weil jedes Element additiv invertierbar ist.

(d) In der Halbgruppe $(\mathbb{Z},\cdot)$ sind genau zwei Elemente invertierbar: 1 und (-1). Also ist die Einheitengruppe die zweielementige Gruppe $(\mathbb{Z},\cdot)^\times = (\{-1,1\},\cdot)$.

(e) In den Halbgruppen $(\mathbb{Q},\cdot)$ und $(\mathbb{R},\cdot)$ sind alle Elemente außer 0 invertierbar. Es gilt also: $(\mathbb{Q},\cdot)^\times = (\mathbb{Q}\setminus\{0\},\cdot)$ und $(\mathbb{R},\cdot)^\times = (\mathbb{R}\setminus\{0\},\cdot)$.

(f) Die Halbgruppe $(D,\star)$ mit der Verknüpfung auf Seite 61 ist eine Gruppe. Die Assoziativität muss man von Hand überprüfen, aber man kann der Tabelle direkt ablesen, dass i das neutrale Element ist, da weder Linksmultiplikation noch Rechtsmultiplikation mit i irgendetwas ändert. Ferner sieht man, dass in jeder Zeile und in jeder Spalte das neutrale Element i auftaucht, was bedeutet, dass jedes Element invertierbar ist. Also ist $(\{i,s_1,s_2,s_3,u,g\},\star)$ eine Gruppe. Die Gruppe ist nicht kommutativ, weil die Cayley-Tafel nicht symmetrisch ist.

(g) Es sei X eine beliebige Menge. Wir haben bereits gesehen, dass für die Menge aller Selbstabbildungen $(X^X,\circ)$ ein Monoid ist, in dem eine Abbildung genau dann invertierbar bezüglich $\circ$ ist, wenn sie bijektiv ist. Also ist die Einheitengruppe von $(X^X,\circ)$ die Menge aller Bijektionen von X nach X:

$$\mathscr{S}(X) := (X^X,\circ)^\times := \{f : X \to X \mid f \text{ ist bijektiv}\}.$$

[3] nach NIELS HENRIK ABEL, norwegischer Mathematiker und Mitbegründer der modernen Gruppentheorie, der mit nur 26 Jahren an Tuberkulose starb, 1802–1829

[4] streng genommen: Einschränkung und Koeinschränkung

Diese Gruppe heißt die *symmetrische Gruppe* oder *Permutationsgruppe* der Menge X. Bijektive Selbstabbildungen einer Menge X nennt man auch *Permutationen* von X.

Man sieht leicht, dass $\mathscr{S}(X)$ unendlich viele Elemente hat, wenn X unendlich viele Elemente hat.

(h) Es sei $n \in \mathbb{N}$ eine natürliche Zahl[5] und $X := \{1,\dots,n\}$. Dann schreibt man auch $\mathscr{S}(n) := \mathscr{S}(X)$ für die symmetrische Gruppe auf n Elementen. Per vollständiger Induktion zeigt man leicht, dass die Gruppe $\mathscr{S}(n)$ genau $n! = n \cdot (n-1) \cdots 1$ viele Elemente besitzt.

Jedes Element $\sigma \in \mathscr{S}(n)$ ist eine Abbildung von der endlichen Menge $\{1,\dots,n\}$ in sich selbst und somit eindeutig beschrieben, wenn wir alle Funktionswerte explizit angeben. Beispielsweise ist

$$\sigma : \{1,2,3\} \to \{1,2,3\}, \quad k \mapsto \begin{cases} 2 & \text{falls } k = 1, \\ 3 & \text{falls } k = 2, \\ 1 & \text{falls } k = 3 \end{cases}$$

eine bijektive Selbstabbildung von $\{1,2,3\}$, also ein Element in der Gruppe $\mathscr{S}(3)$.

Für solche Permutationen gibt es viele unterschiedliche Notationen: Beispielsweise kann man die Permutation σ, die wir eben definiert haben, kürzer über ihre Wertetabelle beschreiben:

$$\sigma = \begin{array}{c|c|c} 1 & 2 & 3 \\ \hline 2 & 3 & 1 \end{array}$$

Hier wird explizit für jedes Element im Definitionsbereich (obere Zeile) der Funktionswert (untere Zeile) angegeben. Da der Definitionsbereich nur endlich viele Elemente hat, ist damit σ eindeutig beschrieben.

Eine andere übliche Notation ist die folgende:

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

Diese Schreibweise ähnelt einer Matrix, sollte aber nicht mit einer solchen verwechselt werden. Diese Ähnlichkeit ist rein zufällig. Nichts von dem, was wir über Matrizen gelernt haben, lässt sich auf diese *Matrix-Schreibweise* anwenden. Solche *Matrizen* lassen sich nicht addieren, multiplizieren oder transponieren.

Eine andere Schreibweise, die für Permutationen gebräuchlich ist, ist die *Zykelschreibweise*:

$$\sigma = (1,2,3).$$

Was hier aussieht, wie ein 3-Tupel oder ein Zeilenvektor sieht wieder nur so aus und bedeutet etwas ganz anderes: Man liest eine solche wie folgt:

Die Permutation $(x_1, x_2, \dots, x_m)$ bildet das erste Element in der Liste auf das zweite ab: $x_1 \mapsto x_2$, das zweite Element auf das dritte $x_2 \mapsto x_3$ und so weiter, bis zum letzten Element x_m, das wieder auf das erste Element abgebildet wird: $x_m \mapsto x_1$.

Wir könnten also ebenso schreiben $\sigma = (1,2,3) = (2,3,1) = (3,1,2)$.

[5] Wenn man sehr haarspalterisch ist, kann man diese Definition auch für $n = 0$ betrachten; dann ist $X = \emptyset$.

Falls ein Element in der Aufzählung gar nicht vorkommt, wird es auf sich selbst abgebildet. Nicht jede Permutation ist von dieser Form, aber jede Permutation (auf einer endlichen Menge) lässt sich als Verkettung von endlich vielen Zykel-Permutationen schreiben.

Solange n relativ klein ist, gibt es auch die Möglichkeit, eine Permutation durch ihren Graphen darzustellen. Bleiben wir bei dem obigen Beispiel $\sigma : \{1,2,3\} \to \{1,2,3\}$ und zeichnen den Graphen. Die drei Elemente des Definitionsbereichs 1,2,3 kommen auf die x-Achse, die wie gewöhnlich nach rechts zeigt. Die drei Elemente des Wertebereichs setzen wir auf die y-Achse und aus Gründen, die hoffentlich weiter unten klarer werden, machen wir das Ungewöhnliche und zeichnen die y-Achse nach unten. Der Graph von σ sieht dann so aus:

	1	2	3
1			•
2	•		
3		•	

Auch wenn diese Schreibweise für große Werte von n ziemlich unpraktisch wird, hat sie den Vorteil, dass das Verketten von Permutationen nun genau wie das Multiplizieren von Matrizen funktioniert. Wenn wir zum Beispiel die Permutation τ betrachten, die 1 und 2 vertauscht und 3 auf sich selbst abbildet:

$$\tau = \begin{array}{c|c|c} 1 & 2 & 3 \\ \hline 2 & 1 & 3 \end{array} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} = (1,2) = \begin{array}{c|ccc} & 1 & 2 & 3 \\ \hline 1 & & \bullet & \\ 2 & \bullet & & \\ 3 & & & \bullet \end{array}$$

Dann ist die Verkettung $\sigma \circ \tau$ gegeben durch:

$$\sigma \circ \tau = \left(\begin{array}{c|ccc} & 1 & 2 & 3 \\ \hline 1 & & & \bullet \\ 2 & \bullet & & \\ 3 & & \bullet & \end{array}\right) \circ \left(\begin{array}{c|ccc} & 1 & 2 & 3 \\ \hline 1 & & \bullet & \\ 2 & \bullet & & \\ 3 & & & \bullet \end{array}\right) = \begin{array}{c|ccc} & 1 & 2 & 3 \\ \hline 1 & & & \bullet \\ 2 & & \bullet & \\ 3 & \bullet & & \end{array},$$

was nicht nur zufällig genauso aussieht wie die Rechnung

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

Quadratische Matrizen, die aus einem Graphen einer Permutation auf diese Weise entstehen und somit in jeder Zeile und in jeder Spalte genau eine 1 haben und sonst nur Nullen, nennt man auch *Permutationsmatrizen*. Das neutrale Element $\mathrm{id}_{\{1,\dots,n\}}$ der Gruppe $\mathscr{S}(n)$ entspricht hierbei genau der Einheitsmatrix $\mathbb{1}_n$. Auf den Zusammenhang zwischen Permutationen und Matrizen werden später[6] beim Studium der Determinante noch genauer eingehen.

Beispiel 3.2.4. Es sei $\Delta \subseteq \mathbb{R}^2$ ein gleichseitiges Dreieck in der Ebene. Wir betrachten alle Symmetrieoperationen, d.h. alle Selbstabbildungen von $\mathbb{R}^2$, die entweder Drehungen oder Spiegelungen sind und das Dreieck auf sich selbst abbilden. Durch (systematisches) Herumprobieren

[6] siehe Proposition 5.3.13

stellen wir fest, dass es drei Spiegelungen gibt – die Spiegelachsen sind die drei Winkelhalbierenden – sowie drei Drehungen, eine um $120° = \frac{2\pi}{3}$, eine um $240° = 2 \cdot \frac{2\pi}{3}$ und eine um $360° = 2\pi$ jeweils in mathematisch positivem Drehsinn, d.h. *gegen* den Uhrzeigersinn.

Die Drehung um 120° *im* Uhrzeigersinn muss nicht extra aufgelistet werden, weil sie als Abbildung von $\mathbb{R}^2$ nach $\mathbb{R}^2$ identisch ist mit der Drehung um 240° gegen den Uhrzeigersinn.

Weiterhin fällt auf, dass die Drehung um 360° alle Punkte festhält und somit identisch ist mit der identischen Abbildung $\mathrm{id} : \mathbb{R}^2 \to \mathbb{R}^2$, die gar nichts tut.

Es sei nun G die Menge aller Symmetrieoperationen des gleichseitigen Dreiecks. Dann gilt also $|G| = 6$, es gibt drei Drehungen (inklusive der Identität) und drei Spiegelungen. Da jede dieser Symmetrieoperationen ja eine Abbildung der Ebene in sich ist, können wir Symmetrieoperationen auch verketten und erhalten wieder eine Symmetrieoperation. Es gibt also eine zweistellige Verknüpfung *zwischen den Symmetrien*:

$$\circ : G \times G \to G, \quad (\sigma, \tau) \mapsto \sigma \circ \tau.$$

Da die Verkettung von Abbildung assoziativ ist, G die identische Abbildung enthält und selbstverständlich jede Symmetrieoperation umkehrbar ist, wird $(G, \circ)$ zu einer Gruppe, genannt die *Symmetriegruppe des gleichseitigen Dreiecks*. Für diese sechselementige Gruppe ist auch die Notation D_3 gebräuchlich.

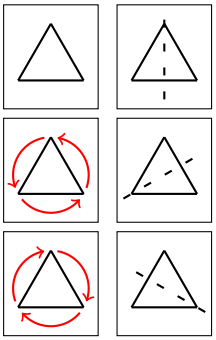


Abbildung 3.1.: Die sechs Elemente der Gruppe D_3

Wenn wir statt des gleichseitigen Dreiecks mit einem gleichseitigen n-Eck anfangen, erhalten wir die sogenannte *Diedergruppe* D_n, die immer aus n Drehungen und n Spiegelungen besteht, wobei wir die Identität immer als Drehung betrachten.

Ebenso kann man solche Symmetriegruppen für andere Objekte $A \subseteq \mathbb{R}^n$ betrachten, wobei man auch spezifizieren kann, welche Arten von Operationen erlaubt sind. So gibt es eine Symmetriegruppe des Würfels im $\mathbb{R}^3$, die aus 48 Elementen besteht, wobei es 24 Drehungen und 24 Spiegelungen gibt.

Es stellt sich heraus, dass die Verkettung von zwei Spiegelungen stets eine Drehung ergibt. Diese Aussagen werden wir hier jetzt nicht beweisen – aber auch nicht verwenden – sie dienen nur als Motivation, sich mit dem Gruppenbegriff zu beschäftigen.

Die Symmetriegruppe eines Kreises in der Ebene hat unendlich viele Elemente.

Bemerkung. In einer Gruppe $(G, *)$ gilt die folgende Kürzungsregel:

$$\forall x, y, a \in G : (a * x = a * y \implies x = y).$$

Dies sieht man zum Beispiel so:

$$\begin{aligned} x &= e_G * x \\ &= (a^{-1} * a) * x \\ &= a^{-1} * (a * x) \\ &= a^{-1} * (a * y) \\ &= (a^{-1} * a) * y \\ &= e_G * y \\ &= y. \end{aligned}$$

Analog zeigt man auch die folgende Kürzungsregel:

$$\forall x, y, z \in G : (x * a = y * a \implies x = y).$$

Definition 3.2.5 (Untergruppe)**.** Es sei $(G, *)$ eine Gruppe mit neutralem Element e_G und $H \subseteq G$ eine Teilmenge. Dann heißt H *Untergruppe* von G, wenn folgende drei Eigenschaften erfüllt sind:

(i) $e_G \in H$

(ii) $\forall a, b \in H : a * b \in H$

(iii) $\forall a \in H : a^{-1} \in H$.

Eine Untergruppe H einer Gruppe $(G, *)$ ist selbst wieder eine Gruppe, wenn sie mit der (Einschränkung und Koeinschränkung) der Gruppenmultiplikation von G versehen wird.

Beispiel 3.2.6. (a) $\mathbb{Z}$ ist eine Untergruppe von $(\mathbb{Q}, +)$ und $\mathbb{Q}$ ist eine Untergruppe von $(\mathbb{R}, +)$.

(b) Jeder Untervektorraum U von $\mathbb{R}^n$ ist eine Untergruppe von $(\mathbb{R}^n, +)$.

(c) Die Menge der positiven Zahlen $H := \{t \in \mathbb{R} \mid t > 0\}$ ist eine Untergruppe von $(\mathbb{R}, \cdot)^\times = (\mathbb{R} \setminus \{0\}, \cdot)$.

(d) Die Menge $\{i, g, u\}$ ist eine Untergruppe der Gruppe auf Seite 61.

Definition 3.2.7 (Gruppenhomomorphismus)**.** (a) Gegeben seien Gruppen $(G, *_G)$ und $(H, *_H)$. Dann heißt eine Abbildung

$$\varphi : G \to H$$

Homomorphismus von Gruppen (oder *Gruppenhomomorphismus*), wenn

$$\forall x, y \in G : \varphi(x *_G y) = \varphi(x) *_H \varphi(y)$$

gilt. Hieraus folgt dann bereits automatisch (siehe Lemma 3.2.8), dass

$$\varphi(e_G) = e_H$$

und

$$\forall x \in G : \varphi(x^{-1}) = (\varphi(x))^{-1}$$

gilt.

(b) Ein Gruppenhomomorphismus $\varphi : G \to H$ zwischen zwei Gruppen $(G, *)$ und $(H, *)$ heißt *Isomorphismus von Gruppen* (oder *Gruppenisomorphismus*), wenn φ bijektiv ist. In diesem Fall ist auch $\varphi^{-1} : H \to G$ ein Isomorphismus von Gruppen. Zwei Gruppen $(G, *)$ und $(H, *)$ heißen *isomorph*, wenn es einen Isomorphismus zwischen ihnen gibt. Wenn $(G, *)$ und $(H, *)$ isomorph sind, schreibt man auch $(G, *) \cong (H, *)$.

(c) Ein Gruppenhomomorphismus $\varphi : G \to G$ von einer Gruppe in sich selbst nennt man auch einen *Endomorphismus von Gruppen* (oder *Gruppenendomorphismus*). Ein Endomorphismus, der gleichzeitig ein Isomorphismus ist, ist ein *Automorphismus von Gruppen* (oder *Gruppenautomorphismus*).

Lemma 3.2.8. *Es seien $(G, *_G)$ und $(H, *_H)$ Gruppen mit neutralen Elementen e_G und e_H. Es sei $\varphi : G \to H$ ein Gruppenhomomorphismus.*

(a) Dann gilt

$$\varphi(e_G) = e_H \quad \text{und} \quad \forall x \in G : \varphi(x^{-1}) = (\varphi(x))^{-1}.$$

(b) Es sei $U \subseteq G$ eine Untergruppe von $(G, *)$. Dann ist $\varphi(U)$ eine Untergruppe von H. Insbesondere ist Bild$(\varphi) = \varphi(G)$ *eine Untergruppe von H.*

(c) Es sei $W \subseteq H$ eine Untergruppe von $(H, *)$. Dann ist $\varphi^{-1}(W)$ eine Untergruppe von G. Insbesondere ist der Kern $\ker(\varphi) := \varphi^{-1}(\{e_H\})$ *eine Untergruppe von G.*

Beweis. (a)
Es gilt:

$$\varphi(e_G) * \varphi(e_G) = \varphi(e_G * e_G) = \varphi(e_G) = \varphi(e_G) * e_H.$$

Nun wenden wir die Kürzungsregel (Seite 69) an und erhalten:

$$\varphi(e_G) = e_H,$$

was zu zeigen war.

Sei nun $x \in G$. Dann gilt:

$$\varphi(x) * \varphi(x^{-1}) = \varphi(x * x^{-1}) = \varphi(e_G) = e_H = \varphi(x)\left(\varphi(x)\right)^{-1}.$$

Nochmaliges Anwenden der Kürzungsregel (Seite 69) liefert dann:

$$\varphi(x^{-1}) = \left(\varphi(x)\right)^{-1}.$$

(b)
Es sei $U \subseteq G$ eine Untergruppe. Wir müssen zeigen:

- $e_H \in \varphi(U)$,
- $\forall a, b \in \varphi(U) : a * b \in \varphi(U)$,
- $\forall a \in \varphi(U) : a^{-1} \in \varphi(U)$.

Nach Teil (a) gilt: $e_H = \varphi(e_G) \in \varphi(U)$. Hier haben wir verwendet, dass die Untergruppe U das Neutralelement e_G enthält.

Gegeben seien nun $a, b \in \varphi(U)$. Dann gibt es $u, v \in U$ mit

$$a = \varphi(u) \quad \text{und} \quad b = \varphi(v).$$

Es gilt dann also:

$$a * b = \varphi(u) * \varphi(v) = \varphi(u * v) \in \varphi(U).$$

Hier haben wir verwendet, dass die Untergruppe U unter der Gruppenoperation $*$ abgeschlossen ist.

Schließlich sei $a \in \varphi(U)$ gegeben. Dann ist $a = \varphi(U)$ für ein Element $u \in U$. Unter Verwendung von Teil (a) gilt nun:

$$a^{-1} = (\varphi(u))^{-1} = \varphi(u^{-1}) \in \varphi(U).$$

Hier haben wir verwendet, dass die Untergruppe U unter Inversion abgeschlossen ist.

(c)

Es sei nun $W \subseteq H$ eine Untergruppe von H. Dann gilt:

$$\varphi(e_G) = e_H \in W,$$

also ist $e_G \in \varphi^{-1}(W)$.

Weiter gilt für $a, b \in \varphi^{-1}(W)$:

$$\varphi(a * b) = \underbrace{\varphi(a)}_{\in W} * \underbrace{\varphi(b)}_{\in W} \in W,$$

also ist $a * b \in \varphi^{-1}(W)$.

Schließlich sei $a \in \varphi^{-1}(W)$. Dann gilt:

$$\varphi(a^{-1}) = (\varphi(a))^{-1} \in W,$$

also ist auch $a^{-1} \in \varphi^{-1}(W)$. □

Beispiel 3.2.9. (a) Die Abbildung $\exp : (\mathbb{R}, +) \to \mathbb{R}^\times, \quad t \mapsto e^t$ ist ein Gruppenhomomorphismus.

(b) Die Abbildung $\exp : (\mathbb{R}, +) \to (\{s \in \mathbb{R} \mid s > 0\}, \cdot), \quad t \mapsto e^t$ ist ein Gruppenisomorphismus zwischen der additiven Gruppe der reellen Zahlen und der multiplikativen Gruppe der positiven reellen Zahlen.

(c) Die Abbildung $\varphi : (\mathbb{Z}, +) \to (\{-1, 1\}, \cdot), \quad k \mapsto (-1)^k$ ist ein Gruppenhomomorphismus.

(d) Allgemein: Es sei $(H, *)$ eine Gruppe und $a \in H$ ein Element. Dann ist

$$(\mathbb{Z}, +) \to (H, *), \quad k \mapsto a^k$$

ein Gruppenhomomorphismus (siehe Notation 3.1.12).

(e) Gegeben $m, n \in \mathbb{N}$. Dann ist die Abbildung

$$\varphi : \mathbb{R}^{m \times n} \to \mathbb{R}^{n \times m}, \quad A \mapsto A^\top$$

ein Isomorphismus zwischen den additiven Gruppen $(\mathbb{R}^{m \times n}, +)$ und $(\mathbb{R}^{n \times m}, +)$.

(f) Die Abbildung $\varphi : \mathbb{R}^\times \to \mathbb{R}^\times$, die x auf $\frac{1}{x}$ abbildet ist ein Automorphismus der Gruppe $\mathbb{R}^\times = (\mathbb{R} \setminus \{0\}, \cdot)$.

(g) Für eine Matrix $A \in \mathbb{R}^{m \times n}$ ist die Abbildung

$$\mathbb{R}^n \to \mathbb{R}^m, \quad x \mapsto Ax$$

ein Gruppenhomomorphismus von der Gruppe $(\mathbb{R}^n, +)$ in die Gruppe $(\mathbb{R}^m, +)$.

(h) Ist $H \subseteq G$ eine Untergruppe, so ist die Inklusionsabbildung

$$\iota : H \to G, \quad x \mapsto x$$

ein Gruppenhomomorphismus von $(H, *)$ nach $(G, *)$.

(i) Die konstante Abbildung $G \to H, \quad x \mapsto e_H$, die jedes Element aus G auf das neutrale Element von H abbildet, ist ein Gruppenhomomorphismus.

(j) Die Identität $\mathrm{id}_G : G \to G$ ist ein Automorphismus der Gruppe $(G, *)$.

(k) Es sei $(D_3, \circ)$ die Symmetriegruppe des gleichseitigen Dreiecks (siehe Beispiel 3.2.4). Jede dieser Symmetrieoperationen permutiert die drei Eckpunkte des Dreiecks. Wenn wir nun die Eckpunkte gegen den Uhrzeigersinn mit den Zahlen 1,2,3 versehen, dann gibt uns jede Symmetrie des Dreiecks eine Permutation der Eckpunkte. Wir erhalten also einen Gruppenhomomorphismus

$$\varphi : D_3 \to \mathscr{S}(3).$$

Dieser Homomorphismus ist injektiv, weil eine Symmetrieoperation des Dreiecks eindeutig beschrieben ist, wenn wir wissen, was mit den Eckpunkten passiert. Außerdem ist er surjektiv, weil jede Permutation der Eckenmenge durch eine Spiegelung oder eine Drehung des Dreiecks realisiert werden kann. Also ist φ ein Isomorphismus und die beiden Gruppen D_3 und $\mathscr{S}(3)$ sind isomorph, d.h. bis auf Umbenennung der Elemente kann man beide Gruppen als „gleich“ ansehen.

Man sollte nun aber nicht denken, dass die Symmetriegruppe eines regelmäßigen n-Ecks immer isomorph ist zur symmetrischen Gruppe auf n Elementen. Man kann zwar immer einen Gruppenhomomorphismus hinschreiben, der für alle $n \geq 3$ auch injektiv ist, aber die Surjektivität ist für $n > 3$ nicht gegeben, da es Permutationen der Eckenmenge gibt, die sich nicht durch eine Drehung oder Spiegelung des n-Ecks realisieren lassen.

Es sei nun $(D, \star)$ die Gruppe mit der Cayley-Tafel auf Seite 61. Auch diese Gruppe ist isomorph zu D_3 bzw. $\mathscr{S}(3)$. Dazu muss man sich nur davon überzeugen, dass die folgende

Bijektion ein Gruppenhomomorphismus ist:

$$\begin{aligned}
\psi : (D, \star) &\to (\mathscr{S}(3), \circ) \\
i \mapsto \mathrm{id}_{\{1,2,3\}} &= \begin{array}{c|c|c} 1 & 2 & 3 \\ \hline 1 & 2 & 3 \end{array} \\
s_1 \mapsto (2,3) &= \begin{array}{c|c|c} 1 & 2 & 3 \\ \hline 1 & 3 & 2 \end{array} \\
s_2 \mapsto (1,3) &= \begin{array}{c|c|c} 1 & 2 & 3 \\ \hline 3 & 2 & 1 \end{array} \\
s_3 \mapsto (1,2) &= \begin{array}{c|c|c} 1 & 2 & 3 \\ \hline 2 & 1 & 3 \end{array} \\
g \mapsto (1,2,3) &= \begin{array}{c|c|c} 1 & 2 & 3 \\ \hline 2 & 3 & 1 \end{array} \\
u \mapsto (3,2,1) &= \begin{array}{c|c|c} 1 & 2 & 3 \\ \hline 3 & 1 & 2 \end{array}
\end{aligned}$$

Die Gruppe auf Seite 61 ist somit isomorph zu $\mathscr{S}(3)$ und insbesondere auch isomorph zu D_3.

Später[7] werden wir noch viele weitere Beispiele für Gruppen und Gruppenhomomorphismen kennenlernen.

Lemma 3.2.10 (Injektive Gruppenhomomorphismen)**.** *Ein Gruppenhomomorphismus $\varphi : G \to H$ zwischen Gruppen $(G, *)$ und $(H, *)$ ist genau dann injektiv, wenn* $\ker \varphi = \{e_G\}$.

Beweis. Zuerst zeigen wir, dass aus der Injektivität von φ folgt, dass $\ker \varphi = \{e_G\}$.

Da φ ein Gruppenhomomorphismus ist, gilt nach Lemma 3.2.8, dass $\varphi(e_G) = e_H$. Somit ist $e_G \in \ker \varphi$. Sei umgekehrt $a \in \ker \varphi$. Dann gilt:

$$\varphi(a) = e_H = \varphi(e_G).$$

Also folgt aus der Injektivität von φ, dass $a = e_G$.

Nun zeigen wir die andere Implikation, dass also aus $\ker \varphi = \{e_G\}$ folgt, dass φ injektiv ist.

Gegeben seien also $a, b \in G$ mit $\varphi(a) = \varphi(b)$. Es ist zu zeigen, dass $a = b$.

$$\varphi(a * b^{-1}) = \varphi(a) * \varphi(b^{-1}) = \varphi(b) * \varphi(b^{-1}) = \varphi(b * b^{-1}) = \varphi(e_G) = e_H.$$

Also ist $a * b^{-1} \in \ker \varphi$. Da aber $\ker \varphi = \{e_G\}$ ist, gilt somit

$$a * b^{-1} = e_G.$$

Durch Multiplizieren mit b von rechts erhält man dann $a = b$. □

[7] Siehe z.B. Lemma 3.5.7, Notation 3.5.8, Satz 5.1.8 oder Korollar 5.3.11

Zusammenfassung von Abschnitt 3.2

(1) Eine Gruppe $(G,*)$ ist eine Menge mit einer binären Verknüpfung, die gewisse Eigenschaften erfüllen muss.

(2) Die Menge aller Permutationen $\mathscr{S}(()J)$ einer (endlichen oder unendlichen) Menge ist eine Gruppe mit der Verkettung als Operation.

(3) Ein Gruppenhomomorphismus ist eine strukturerhaltende Abbildung zwischen Gruppen.

(4) Zwei Gruppen sind isomorph, wenn es einen Gruppenisomorphismus zwischen ihnen gibt.

3.3. Ringe und Körper

Wir haben im Letzten Kapitel gesehen, dass Halbgruppen und Gruppen Mengen sind, zusammen mit einer binären Verknüpfung darauf. In vielen Fällen haben wir aber mehr als eine Verknüpfung auf einer Menge und gerade das Zusammenspiel dieser beiden Verknüpfungen ist interessant. Beispielsweise gibt auf der Menge $\mathbb{Z}$ der ganzen Zahlen die Addition $+:\mathbb{Z}\times\mathbb{Z}\to\mathbb{Z}$ und die Multiplikation $\cdot:\mathbb{Z}\times\mathbb{Z}\to\mathbb{Z}$. Additiv gesehen ist $(\mathbb{Z},+)$ eine Gruppe (mit neutralem Element 0). Multiplikativ ist $(\mathbb{Z},\cdot)$ nur eine Halbgruppe mit neutralem Element 1. Verbunden sind die beiden Strukturen durch das Distributivgesetz. Dies motiviert die folgende Definition:

Definition 3.3.1 (Ring)**.** (a) Ein *Ring* $(R,+,\cdot)$ besteht aus einer Menge R zusammen mit zwei zweistelligen Verknüpfungen

$$+:R\times R\to R \quad \text{und} \quad \cdot:R\times R\to R,$$

genannt *Addition* und *Multiplikation*, sodass die folgenden Eigenschaften erfüllt sind:

(i) $(R,+)$ ist eine kommutative Gruppe.

(ii) $(R,\cdot)$ ist ein Monoid.

(iii) Es gelten die folgenden Distributivgesetze:

$$\forall x,y,z\in R: x\cdot(y+z)=x\cdot y+x\cdot z \quad \text{und} \quad \forall x,y,z\in R:(y+z)\cdot x=y\cdot x+z\cdot x$$

Ein Ring $(R,+,\cdot)$ heißt *kommutativ*, wenn die Multiplikation $\cdot:R\times R\to R$ kommutativ ist.

(b) Es sei $(R,+,\cdot)$ ein Ring. Das neutrale Element der Addition wird mit 0_R oder einfach 0 bezeichnet und heißt *additives neutrales Element* oder *Nullelement*. Das Inverse von $x\in R$ in der additiven Gruppe $(R,+)$ wird mit $-x$ bezeichnet.

Das neutrale Element der Multiplikation wird normalerweise mit 1_R oder 1 bezeichnet und heißt *multiplikatives neutrales Element* oder *Einselement*. Ein Ringelement heißt *invertierbar*, wenn es bezüglich der Multiplikation invertierbar ist. Das Inverse eines invertierbaren Elements $x\in R^\times:=(R,\cdot)^\times$ wird gewöhnlich mit x^{-1} bezeichnet. Die Notation $\frac{1}{x}$ ist – vor allem in nichtkommutativen Ringen – nicht üblich.

Bemerkung 3.3.2. (a) In Definition 3.3.1 haben wir die bei Ringen übliche Punkt-Vor-Strich-Notation angewendet.

(b) Oft wird der „Multiplikationspunkt“ einfach weggelassen und man schreibt $xy = x \cdot y$.

(c) Die Bedingung, dass $(R, \cdot)$ ein neutrales Element besitzt, ist nicht einheitlich in der Literatur. In Texten, wo darauf verzichtet wird, wird ein Ring, wie wir ihn definiert haben als *Ring mit Eins* bezeichnet.

(d) Dass man die Kommutativität von $(R, +)$ in Definition 3.3.1 fordert, ist in der Literatur üblich, aber eigentlich überflüssig, da man die Kommutativität der Addition aus den restlichen Ringeigenschaften herleiten kann (zumindest, wenn man – wie wir hier – Ringe mit Eins betrachtet.

Beispiel 3.3.3. (1) $(\mathbb{Z}, +, \cdot), (\mathbb{Q}, +, \cdot), (\mathbb{R}, +, \cdot)$ sind Ringe, wobei + und $\cdot$ die gewöhnliche Addition bzw. Multiplikation ist.

(2) Für jedes $n \in \mathbb{N}$ ist $(\mathbb{R}^{n \times n}, +, \cdot)$ ein Ring (siehe Lemma 2.2.12). Hierbei ist + die Matrixaddition und $\cdot$ die Matrixmultiplikation.

(3) Für $m \neq n$ ist $(\mathbb{R}^{m \times n}, +, \cdot)$ kein Ring, weil die Multiplikation auf dieser Menge nicht definiert ist.

(4) $(\mathbb{N}, +, \cdot)$ und $(\mathbb{N}_0, +, \cdot)$ sind keine Ringe, weil $\mathbb{N}$ bzw. $\mathbb{N}_0$ bezüglich der Addition keine Gruppe bildet.

(5) $(2\mathbb{Z}, +, \cdot)$ mit $2\mathbb{Z} := \{2k \mid k \in \mathbb{Z}\}$ ist kein Ring nach unserer Definition, weil kein neutrales Element der Multiplikation existiert.

(6) $(\{0\}, +, \cdot)$ mit $0 + 0 = 0 \cdot 0 = 0$ ist ein Ring nach unserer Definition. Hier ist $1_R = 0_R$.

Lemma 3.3.4 (Rechenregeln in Ringen)**.** *Es sei $(R, +, \cdot)$ ein Ring mit additivem neutralen Element 0_R und multiplikativem neutrale Element 1_R. Dann gelten folgende Regeln:*

(a) $\forall x \in R : 0_R \cdot x = 0_R = x \cdot 0_R$.

(b) $\forall x \in R : -x = (-1_R) \cdot x = x \cdot (-1_R)$.

Beweis. (a) Es sei $x \in R$. Dann gilt nach dem Distributivgesetz:

$$(0_R \cdot x) + (0_R \cdot x) = (0_R + 0_R) \cdot x = 0_R \cdot x$$

Da $(R, +)$ eine Gruppe ist, können wir auf beide Seiten das additiven Inversen von $0_R \cdot x$ addieren und erhalten

$$0_R \cdot x = 0_R.$$

Die Formel $0_R = x \cdot 0_R$ zeigt man völlig analog.

(b) Es sei -1_R das additive Inverse des multiplikativen neutralen Elements 1_R und sei $x \in R$

beliebig. Dann gilt

$$\begin{aligned}(-1_R)\cdot x &= (-1_R)\cdot x + 0_R \\ &= (-1_R)\cdot x + (x + (-x)) \\ &= ((-1_R)\cdot x + x) + (-x) \\ &= ((-1_R)\cdot x + 1_R \cdot x) + (-x) \\ &= ((-1_R + 1_R)\cdot x) + (-x) \\ &= (0_R \cdot x) + (-x) \\ &= 0_R + (-x) \\ &= -x.\end{aligned}$$

Jeder einzelne Schritt ist durch je eine Rechenregel für Ringe gerechtfertigt – bis auf den Schritt $0_R \cdot x = 0_R$, der aus Teil (a) folgt. □

Notation 3.3.5. Es sei R ein Ring und $a \in R$. Dann ist für jedes $n \in \mathbb{N}$ definiert:

$$na := \underbrace{a + \cdots + a}_{n} \quad \text{und} \quad a^n := \underbrace{a \cdot \cdots \cdot a}_{n}$$

Außerdem definiert man $0a := 0_R$ und $a^0 := 1_R$. Es gelten dann die folgenden Potenzgesetze:

$$a^{m+n} = a^m \cdot a^n \quad \text{und} \quad (m+n)a = ma + na.$$

Für nichtkommutative Ringe gelten aber im Allgemeinen folgende Gesetze nicht:

$$(ab)^n = a^n b^n$$

Beachten Sie: Die Potenz a^b für Ringelemente $a, b \in R$ ist im Allgemeinen *nicht* definiert.[8]

Definition 3.3.6 (Unterring)**.** Es sei $(R, +, \cdot)$ ein Ring und $T \subseteq R$ eine Teilmenge. Dann heißt T *Unterring* von R, wenn folgende drei Eigenschaften erfüllt sind:

(i) T ist eine Untergruppe von $(R, +)$

(ii) $1_R \in T$

(iii) $\forall a, b \in T : a \cdot b \in T$.

Ein Unterring T ist selbst wieder ein Ring (mit demselben neutralen Element), wenn + und $\cdot$ entsprechend eingeschränkt werden.

Beispiel 3.3.7. Es sei $(R, +, \cdot) := (\mathbb{R}^{2\times 2}, +, \cdot)$ der Ring der (2×2)-Matrizen mit reellen Einträgen, wie üblich versehen mit der gewöhnlichen Addition und Multiplikation von Matrizen.

(a) Die Menge

$$C := \left\{ \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix} \,\middle|\, \alpha, \beta \in \mathbb{R} \right\}$$

ist ein Unterring von $(R, +, \cdot)$.

[8] Es gibt bereits in den reellen Zahlen, keine Möglichkeit, den Potenzen 0^{-42} oder $(-2)^{\sqrt{3}}$ sinnvoll eine reelle Zahl zuzuweisen (Letzteres ist im Bereich der komplexen Zahlen möglich, aber auch nicht auf eine wirklich zufriedenstellende, eindeutige Weise.

(b) Die Menge

$$\mathbb{Z}^{2\times 2} := \left\{ \begin{pmatrix} k & l \\ m & n \end{pmatrix} \middle| \, k,l,m,n \in \mathbb{Z} \right\}$$

ist ein Unterring von $(R,+,\cdot)$.

(c) Die Menge

$$N := \left\{ \begin{pmatrix} 0 & x \\ 0 & 0 \end{pmatrix} \middle| \, x \in \mathbb{R} \right\}$$

ist kein Unterring von $(R,+,\cdot)$, weil die Einheitsmatrix, das Einselement von $(R,+,\cdot)$, nicht in N liegt. Alle anderen Unterringeigenschaften sind erfüllt.

Wenn wir die Ringoperationen von R auf N einschränken, erfüllt $(N,+,\cdot)$ alle Ringeigenschaften bis auf die Existenz eines Einselementes.

(d) Wir setzen

$$T := \left\{ \begin{pmatrix} \alpha & 0 \\ 0 & 0 \end{pmatrix} \middle| \, \alpha \in \mathbb{R} \right\}.$$

Dann ist $(T,+,\cdot)$ ein Ring mit den eingeschränkten Ringoperationen von $R := (\mathbb{R}^{2\times 2},+,\cdot)$.

Trotzdem ist T kein Unterring von $(R,+,\cdot)$, weil das Einselement von $(R,+,\cdot)$ nicht in T liegt. Der Ring $(T,+,\cdot)$ besitzt zwar ein Einselement, aber dies ist nicht das Einselement von R.

Definition 3.3.8 (Ringhomomorphismus)**.** (a) Gegeben seien Ringe $(R,+,\cdot)$ und $(T,+,\cdot)$. Dann heißt eine Abbildung

$$\varphi : R \to T$$

Homomorphismus von Ringen (oder *Ringhomomorphismus*), wenn

$$\forall x,y \in R : \varphi(x+y) = \varphi(x) + \varphi(y)$$

$$\forall x,y \in R : \varphi(x\cdot y) = \varphi(x) \cdot \varphi(y)$$

$$\varphi(1_R) = 1_T.$$

gilt.

(b) Die Begriffsbildungen Isomorphismus, Endomorphismus, Automorphismus sind genauso wie bei Gruppen.

Völlig analog zu Lemma 3.2.8 zeigt man:

Lemma 3.3.9. *Es seien $(R,+,\cdot)$ und $(T,+,\cdot)$ Ringe und $\varphi : R \to T$ ein Ringhomomorphismus.*

(a) Es sei $U \subseteq R$ ein Unterring von $(R,+,\cdot)$. Dann ist $\varphi(U)$ ein Unterring von T. Insbesondere ist $\mathrm{Bild}(\varphi) = \varphi(R)$ *ein Unterring von T.*

(b) Es sei $W \subseteq T$ ein Unterring von $(T,+,\cdot)$. Dann ist $\varphi^{-1}(W)$ ein Unterring von R.

Bemerkung 3.3.10. Der *Kern* $\ker(\varphi) := \varphi^{-1}(\{0_T\})$ eines Ringhomomorphismus $\varphi : R \to T$ ist im Allgemeinen kein Unterring von $(R,+,\cdot)$. Dies ist kein Widerspruch zu Lemma 3.3.9, weil $\{0_T\}$ im Allgemeinen kein Unterring von $(T,+,\cdot)$ ist.

Definition 3.3.11 (Körper)**.** Ein *Körper* ist ein Ring $(K,+,\cdot)$ mit den beiden zusätzlichen Eigenschaften:

(i) Die Einheitengruppe ist gegeben durch: $K^\times = K \setminus \{0_K\}$.

(ii) Die Multiplikation ist kommutativ.

In einem Körper ist ein Element also invertierbar genau dann wenn es nicht 0_K ist.

Beispiel 3.3.12. Wichtige Beispiele für Körper sind der Körper der reellen Zahlen $(\mathbb{R},+,\cdot)$, der rationalen Zahlen $(\mathbb{Q},+,\cdot)$ sowie der komplexen Zahlen $(\mathbb{C},+,\cdot)$ (siehe Satz 3.5.4) sowie Körper der Form $(\mathbb{Z}/p\mathbb{Z},+,\cdot)$ (siehe Satz 3.4.19).

Die ganzen Zahlen $(\mathbb{Z},+,\cdot)$ sind ein wichtiges Beispiel für einen Ring, der kein Körper ist.

Ebenso sind die Matrizenringe $\mathbb{R}^{n\times n}$ für $n>1$ keine Körper.

Bemerkung 3.3.13. (a) In einem Körper gelten nun alle aus der Schule bekannten Rechenregeln, sofern sie nur die Grundrechenarten und keine Vergleiche verwenden. Beispielsweise gelten die Formeln:

$$(a+b)^2 = a^2+2ab+b^2 \quad \text{und} \quad (a+b)(a-b) = a^2-b^2.$$

(b) Ein Körper ist immer *nullteilerfrei*, d.h. aus $a\cdot b = 0_K$ folgt immer $a = 0_K$ oder $b = 0_K$. Dies ist z.B. im Ring $(\mathbb{R}^{2\times 2},+,\cdot)$ nicht der Fall.

(c) Wenn wir in Definition 3.3.11 auf die Bedingung der Kommutativität verzichten, erhalten wir einen *Divisionsring*. Ein Körper ist also ein kommutativer Divisionsring. Es gibt auch den Begriff *Schiefkörper*, der – je nach Autor – entweder synonym zu Divisionsring ist oder nur für solche Divisionsringe verwendet wird, die nicht kommutativ sind.

Wichtigstes Beispiel für einen nichtkommutativen Divisionsring ist der Ring der *Quaternionen*. Quaternionen sind in gewisser Weise „vierdimensionale Zahlen“ und eignen sich gut, um Drehungen im dreidimensionalen Raum zu beschreiben. Deshalb spielen sie auch eine wichtige Rolle in der Computergrafik und der Robotik. Wir können zum jetzigen Zeitpunkt nicht weiter auf dieses Thema eingehen, kommen aber vielleicht im zweiten Semester darauf zurück.

(d) Aus Teil (i) von Definition 3.3.11 folgt, dass 0_K nicht (multiplikativ) invertierbar ist. Da andererseits aber 1_K immer invertierbar ist, folgt hieraus, dass in einem Körper immer $0_K \neq 1_K$ sein muss.

Der Ring $R := (\{0\},+,\cdot)$ aus Beispiel 3.3.3 ist insbesondere kein Körper.

Beispiel 3.3.14 ($\mathbb{F}_2$ und $\mathbb{F}_4$)**.** Auf welchen Mengen kann man eine Körperstruktur definieren? Diese Frage ist in dieser Allgemeinheit zu schwierig, um sie in dieser Veranstaltung zu beantworten.

Da ein Körper auf jeden Fall 0_K und 1_K beinhalten muss und diese beiden Elemente nach Bemerkung 3.3.13 unterschiedlich sein müssen, folgt, dass ein Körper mindestens 2 Elemente haben muss.

Nehmen wir uns also eine Menge mit 2 Elementen her und nennen die beiden Elemente 0 und 1. Ist es möglich, auf $K = \{0,1\}$ so eine Addition und eine Multiplikation zu definieren, dass alle Körperaxiome erfüllt sind?

Wenn wir wollen, dass 0 wirklich das neutrale Element bezüglich der Addition wird, ist klar, dass das additive Inverse von 0 wieder 0 ist. Was ist also das additive Inverse von 1? Da unsere Menge nur zwei Elemente hat, und 0 bereits das additive Inverse von 0 ist, bleibt für 1 nur das Element 1. Also gilt

$$-1 = 1 \quad \text{und} \quad 1+1 = 0.$$

Da 0 das additive Inverse ist, gelten außerdem die Regeln $0+0=0$ und $0+1=1+0=1$. Somit ist die Addition eindeutig festgelegt.

Die Multiplikation ist noch einfacher: $1\cdot 1 = 1$, weil 1 das neutrale Element der Multiplikation ist und $0\cdot 1 = 1\cdot 0 = 0\cdot 0$ nach Lemma 3.3.4.

FALLS es also einen Körper mit zwei Elementen gibt, dann müssen die Rechenoperationen so aussehen:

$+$	0	1
0	0	1
1	1	0

und

$\cdot$	0	1
0	0	0
1	0	1

Nachrechnen jedes einzelnen Körperaxioms zeigt nun, dass diese Wahl auch tatsächlich einen Körper definiert.

Er wird auch mit $\mathbb{F}_2$ oder GF(2) bezeichnet. Später (Satz 3.4.19) werden wir sehen, dass auf einer Menge mit p Elementen immer eine Körperstruktur definiert werden kann, wenn p eine Primzahl (also 2,3,5,7,…) ist. Wie sieht es mit anderen Zahlen aus?

Auf einer vierelementigen Menge $\{0,1,A,B\}$ gibt es eine Körperstruktur, die durch folgende Tafeln definiert ist:

$+$	0	1	A	B
0	0	1	A	B
1	1	0	B	A
A	A	B	0	1
B	B	A	1	0

und

$\cdot$	0	1	A	B
0	0	0	0	0
1	0	1	A	B
A	0	A	B	1
B	0	B	1	A

Auch hier kann man von Hand nachrechnen, dass alle Körperaxiome erfüllt sind. Dieser Körper $(\{0,1,A,B\},+,\cdot)$ wird mit $\mathbb{F}_4$ = GF(4) bezeichnet. In diesen beiden Körpern gilt die ungewohnte Rechenregel $1+1=0$.

Es stellt sich heraus, dass auf einer endlichen Menge mit q Elementen genau dann eine Körperstruktur existiert, wenn q eine Primzahlpotenz (also 2,3,4,5,7,8,9,11,13,16,…) ist.

Wenn wir auch nicht in der Lage sein werden, in dieser Vorlesung zu zeigen, dass die Bedingung, dass q eine Primzahlpotenz hinreichend ist (die Konstruktion einer Körperstruktur erfordert mathematische Methoden, die wir dieses Semester nicht zur Verfügung haben werden), so werden wir später[9] mit Methoden der linearen Algebra zeigen, dass sie notwendig ist. Das bedeutet, wir werden sehen: wenn die Anzahl der Elemente einer endlichen Menge keine Primzahlpotenz ist, dann gibt es darauf keine Körperstruktur. Insbesondere werden wir damit zeigen, dass es keinen Körper mit 6 Elementen geben kann.

Notation 3.3.15 (Körperhomomorphismen und Unterkörper)**.** Ein Ringhomomorphismus zwischen zwei Körpern nennt man auch einen *Körperhomomorphismus* (oder einen *Homomorphismus von Körpern*). Ein Unterring K eines Körpers $(L,+\cdot)$ nennt man *Unterkörper* (oder *Teilkörper*), wenn K mit der Unterringstruktur selbst ein Körper wird.

[9] Siehe Satz A.2.5 im Anhang.

Beispielsweise ist $\mathbb{Q}$ ein Unterkörper von $\mathbb{R}$ und $\mathbb{F}_2$ ein Unterkörper von $\mathbb{F}_4$ (siehe Beispiel 3.3.14).

Wenn K ein Unterkörper von L ist, nennen wir L eine *Körpererweiterung* von K.

Zusammenfassung von Abschnitt 3.3

(1) Ein Ring $(R,+,\cdot)$ ist eine Menge mit zwei binären Verknüpfungen $+,\cdot$, die gewisse Eigenschaften erfüllen müssen.

(2) Ein Ringhomomorphismus ist eine strukturerhaltende Abbildung zwischen Ringen.

(3) Ein Körper ist ein spezieller Ring.

(4) Die additive Gruppe $(R,+)$ ist eine abelsche Gruppe mit der Addition.

(5) Die Einheitengruppe $(R,\cdot)^\times$ eines Ringes $(R,\cdot,\cdot)$ ist eine Gruppe mit der Multiplikation.

3.4. Der Ring der ganzen Zahlen und seine Quotientenringe

In diesem Kapitel wollen wir den Ring $(\mathbb{Z},+,\cdot)$ der ganzen Zahlen genauer untersuchen. Nur weil es sehr wenige multiplikativ invertierbare Elemente gibt (genau genommen nur zwei: -1 und 1), heißt das nicht, dass es nicht doch manchmal möglich ist, eine Zahl durch eine andere zu teilen:

Lemma 3.4.1 (Division mit Rest)**.** *Gegeben $a \in \mathbb{Z}$ und $m \in \mathbb{N}$, so gibt es eindeutige $k, r \in \mathbb{Z}$ mit*

$$a = mk + r \quad \text{und } 0 \le r < m.$$

Die Zahl $r \in \{0,\dots,m-1\}$ nennt man auch den Rest von a modulo m.

Beweis. Zuerst zeigen wir die Eindeutigkeit einer solchen Darstellung. Nehmen wir an, es gibt zwei Darstellungen

$$a = mk + r = ml + s$$

mit $k,r,l,s \in \mathbb{Z}$ und $0 \le r,s < m$. Dann bedeutet dies:

$$m(k-l) = r - s.$$

Nehmen wir auf beiden Seiten den Betrag, so ergibt dies:

$$m\,|k-l| = |r-s|.$$

Falls $|k-l| \neq 0$, dann ist die linke Seite dieser Gleichung mindestens m. Da aber $r,s \in \{0,\dots,m-1\}$, ist der Abstand von $|r-s|$ höchstens $m-1$. Die rechte Seite der Gleichung ist also höchstens $m-1$, also muss $|k-l| = 0$ sein, also $k = l$. Dann ist aber auch $|r-s| = 0$ und $r = s$. Dies zeigt die Eindeutigkeit der Darstellung.

Für $a \ge 0$ können wir die Aussage mit vollständiger Induktion beweisen: Induktionsanfang $a = 0$ ist klar: Wir setzen einfach $k = l = 0$.

Für den Induktionsschritt nehmen wir an, $a = mk + r$ mit $r \leq m-1$. Dann ist $a+1 = mk+(r+1)$ mit $r+1 \leq m$. Falls $r+1 < m$ ist, haben wir so eine Darstellung gefunden. Falls $r+1 = m$, dann können wir $a+1$ schreiben als

$$a+1 = mk + m = m(k+1) + 0.$$

Somit folgt die Aussage für alle $a \in \mathbb{N}_0$.

Für $a < 0$ nutzen wir aus, dass $|a| = -a > 0$ ist und somit eine Darstellung hat:

$$-a = mk + r$$

Es folgt somit

$$a = m(-k) - r$$

Falls $r = 0$, ist dies die gewünschte Darstellung. Falls $r \in \{1, \ldots, m-1\}$ gilt, dann ist auch $m - r \in \{1, \ldots, m-1\}$ und wir bekommen die gewünschte Darstellung durch

$$a = m(-k) - r = m(-k) - m + m - r = m(-1-k) + (m-r). \qquad \square$$

Notation 3.4.2 (Teilbarkeitsrelation)**.** Für $a, m \in \mathbb{Z}$ sagen wir *m teilt a* (oder *a ist durch m teilbar*) und schreiben

$$m|a,$$

wenn es ein $k \in \mathbb{Z}$ gibt mit $a = mk$.

Lemma 3.4.3. *Die Teilbarkeitsrelation $|$ hat folgende Eigenschaften:*

(a) $\forall a \in \mathbb{Z} : 1|a$.

(b) $\forall a \in \mathbb{Z} : (0|a \iff a = 0)$.

(c) $\forall a \in \mathbb{Z} : a|a$.

(d) $\forall a, b, c \in \mathbb{Z} : (a|b \text{ und } b|c) \implies a|c$.

(e) $\forall a, b \in \mathbb{Z} : (a|b \text{ und } b|a) \iff |a| = |b|$.

Beweis. (a)
Es sei $a \in \mathbb{Z}$ gegeben. Dann gilt $a = m \cdot 1$. Daraus folgt $1|a$.

(b)
Es sei $a \in \mathbb{Z}$ gegeben mit $0|a$. Dann heißt das, dass es ein $k \in \mathbb{Z}$ gibt mit $a = 0 \cdot k$. Also ist $a = 0$.

Für die Rückimplikation nehmen wir an, dass $a = 0$ gilt. Dann können wir schreiben:

$$a = 0 = 0 \cdot 42,$$

woraus sofort $0|a$ folgt.[10]

(c)
Es sei $a \in \mathbb{Z}$ gegeben. Da wir a schreiben können als

$$a = a \cdot 1,$$

gilt folglich $a|a$.

[10] Es ist streng genommen nicht notwendig, hier im Beweis die Zahl 42 zu verwenden.

(d)
Es seien $a,b,c \in \mathbb{Z}$ gegeben mit $a|b$ und $b|c$. Es ist zu zeigen, dass $a|c$.

Aus $a|b$ folgt, dass es ein $k \in \mathbb{Z}$ gibt mit $b = ak$. Aus $b|c$ folgt, dass es ein $\ell \in \mathbb{Z}$ gibt mit $c = b\ell$.

Fügen wir nun beides zusammen, erhalten wir:

$$c = b\ell = ak\ell = a(k\ell).$$

Also gilt $a|c$.

(e)
Es seien $a,b \in \mathbb{Z}$. Wir zeigen zwei Implikationen:
„$\Longleftarrow$“:
Angenommen $|a| = |b|$. Dann unterscheiden sich a und b höchstens durch ihr Vorzeichen. Es gibt also ein $k \in \{-1,1\}$ mit $b = ak$ und $a = bk$. Also gilt $a|b$ und $b|a$.

„$\Longrightarrow$“:
Angenommen wir haben: $a|b$ und $b|a$. Dann bedeutet dies ausgeschrieben, dass es Zahlen $k,\ell \in \mathbb{Z}$ gibt mit

$$b = ak \quad \text{und} \quad a = b\ell.$$

Wenn wir diese Aussagen ineinander einsetzen, erhalten wir:

$$a = b\ell = ak\ell.$$

Wenn wir in der Gleichung $a = ak\ell$ alles auf eine Seite bringen, erhalten wir:

$$a - a(k\ell) = 0$$
$$a(1 - k\ell) = 0.$$

In den ganzen Zahlen gilt: Ein Produkt ist 0, falls einer der Faktoren 0 ist (die ganzen Zahlen sind nullteilerfrei). Somit gilt:

$$a = 0 \quad \text{oder} \quad k\ell = 1.$$

Im ersten Fall $a = 0$ folgt sofort $b = ak = 0$ und somit ist $|a| = 0 = |b|$.

Im zweiten Fall $k\ell = 1$ folgt, dass k (multiplikativ) invertierbar im Ring $\mathbb{Z}$ ist und die einzigen invertierbaren Elemente sind 1 und -1. Also gilt: $|k| = 1$ und somit

$$|b| = |ak| = |a|\,|k| = |a| \cdot 1 = |a|.$$

□

Bis jetzt haben wir nur die multiplikative Struktur auf den ganzen Zahlen, also das Monoid $(\mathbb{Z},\cdot)$ untersucht. Nun bringen wir die additive Struktur hinzu:

Definition 3.4.4. Es seien $a,b \in \mathbb{Z}$ und $m \in \mathbb{Z}$ gegeben. Wir sagen a und b sind *kongruent modulo m*, wenn

$$m|(a-b).$$

Wir schreiben dafür auch

$$a \equiv_m b \quad \text{oder} \quad a \equiv b (\text{mod } m).$$

Bemerkung 3.4.5. Für $m = 0$ gilt: $(a \equiv_0 b) \iff a = b$.

Für $m = 1$ gilt: $\forall a,b \in \mathbb{Z} : a \equiv_m b$.

Für $m \in \mathbb{N}$ ist $a \equiv_m b$ gleichbedeutend damit, dass a und b den gleichen Rest modulo m haben.

Die in Definition 3.4.4 eingeführte Kongruenz-Relation ist ein Beispiel für ein viel allgemeineres Konzept:

Definition 3.4.6. Es sei X eine Menge.

(a) Eine Teilmenge $\sim \subseteq X \times X$ nennt man eine *Relation* auf X. Falls ein Paar (x,y) zur Teilmenge $\sim$ gehören, schreibt man auch $x \sim y$ anstelle $(x,y) \in \sim$.

(b) Eine Relation $\sim$ auf X heißt *reflexiv*, falls $\forall x \in X : x \sim x$.

(c) Eine Relation $\sim$ auf X heißt *symmetrisch*, falls $\forall x,y \in X : x \sim y \implies y \sim x$.

(d) Eine Relation $\sim$ auf X heißt *transitiv*, falls $\forall x,y,z \in X : (x \sim y \text{ und } y \sim z) \implies x \sim z$.

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

(f) Es sei $\sim$ eine Äquivalenzrelation auf X und es sei $x \in X$. Dann heißt

$$[x]_\sim := \{y \in X \mid x \sim y\}$$

die *Äquivalenzklasse* von x modulo $\sim$.

Bemerkung 3.4.7. Die Teilbarkeitsrelation $|$ auf $\mathbb{Z}$ (siehe Notation 3.4.2) ist reflexiv und transitiv (siehe Lemma 3.4.3), aber nicht symmetrisch und somit keine Äquivalenzrelation. Dasselbe gilt für die gewöhnliche Ordnungsrelation $\leq$ auf $\mathbb{R}$.

Lemma 3.4.8. *Es sei $m \in \mathbb{Z}$. Dann ist die Kongruenz modulo m eine Äquivalenzrelation auf der Menge $\mathbb{Z}$.*

Beweis. Zuerst zeigen wir, dass $\equiv_m$ reflexiv ist:
Es sei $a \in \mathbb{Z}$. Dann gilt

$$a - a = 0 = m \cdot 0.$$

Also gilt: $m|(a-a)$ und somit $a \equiv_m a$.

Nun zeigen wir Symmetrie:
Es seien dazu $a,b \in \mathbb{Z}$ gegeben mit $a \equiv_m b$, d.h. $m|(a-b)$. Ausgeschrieben bedeutet dies, dass es ein $k \in \mathbb{Z}$ gibt mit

$$a - b = mk,$$

woraus sofort

$$b - a = m(-k)$$

folgt. Da $-k \in \mathbb{Z}$ ist, gilt damit $m|(b-a)$ und somit $b \equiv_m a$, was zu zeigen war.

Als Letztes zeigen wir Transitivität: Es seien $a,b,c \in \mathbb{Z}$ gegeben mit $a \equiv_m b$ und $b \equiv_m c$. Wir müssen zeigen, dass $a \equiv_m c$ gilt.

Aus den Voraussetzungen folgt, dass es $k,\ell \in \mathbb{Z}$ gibt mit

$$a - b = mk \quad \text{und} \quad c - b = m\ell.$$

Dies können wir nun zusammensetzen zu:

$$a - c = (a-b) + (b-c) = mk + m\ell = m(k+\ell).$$

Also gilt $m|(a-c)$ und somit $a \equiv_m c$. □

Beispiel 3.4.9. Es sei $q : X \to Q$ eine Abbildung zwischen beliebigen Mengen X und Q. Dann ist die Relation $\sim_q$ auf X definiert durch

$$(x \sim_q y) :\iff q(x) = q(y)$$

eine Äquivalenzrelation auf X.

In gewisser Weise entsteht jede Äquivalenzrelation auf diese Weise: Wann immer wir eine Äquivalenzrelation auf einer Menge haben, können wir alle Elementen identifizieren, die bezüglich dieser Relation äquivalent sind und erhalten somit eine neue Menge, in der nicht mehr zwischen Elementen unterschieden wird, die vorher äquivalent waren. Diese Idee werden wir nun formalisieren:

Satz 3.4.10 (Quotientenmengen und Quotientenabbildungen)**.** *Es sei $\sim$ eine Äquivalenzrelation auf einer Menge X. Dann gibt es eine Menge $X/\sim$ (genannt* Quotientenmenge *oder* Faktormenge*) und eine surjektive Abbildung $q : X \to X/\sim$ (genannt* Quotientenabbildung*), sodass $\forall x, y \in X$: $x \sim y \iff q(x) = q(y)$.*

Beweis. Wir definieren: $X/\sim := \{[x]_\sim \mid x \in X\} \subseteq \mathscr{P}(X)$ die Menge aller Äquivalenzklassen von Elementen in x.

Die Abbildung

$$q : X \to X/\sim, \quad x \mapsto [x]_\sim,$$

die jedem Element aus X seine Äquivalenzklasse zuordnen, ist per Konstruktion surjektiv. Es bleibt zu zeigen: $\forall x, y \in X : x \sim y \iff [x]_\sim = [y]_\sim$.

„$\Longleftarrow$“:
Wir nehmen zwei Elemente $x, y \in X$ mit $[x]_\sim = [y]_\sim$ und wollen zeigen, dass $x \sim y$ gilt. Weil $\sim$ reflexiv ist, gilt $y \sim y$. Nach Definition von $[y]_\sim$ bedeutet dies, dass $y \in [y]_\sim$. Da wir nun aber $[x]_\sim = [y]_\sim$ vorausgesetzt haben, gilt demnach auch $y \in [x]_\sim$ und demnach $x \sim y$, was zu zeigen war.

„$\Longrightarrow$“:
Wir nehmen an, $x \sim y$ und müssen zeigen, dass $[x]_\sim = [y]_\sim$. Da dies eine Gleichheit von zwei Mengen ist, werden wir zwei Mengeninklusionen beweisen: Es sei $z \in [x]_\sim$ gegeben. Wir wollen zeigen, dass $z \in [y]_\sim$.

Aus $x \sim y$ und der Tatsache, dass $\sim$ symmetrisch ist, folgt $y \sim x$. Aus $z \in [x]_\sim$ folgt nun, dass $x \sim z$. Kombinieren dieser beiden Aussagen mit der Transitivität von $\sim$ ergibt dann $y \sim z$ und somit $z \in [y]_\sim$.

Die andere Mengeninklusion geht analog: $z \in [y]_\sim$ ergibt $y \sim z$, was zusammen mit $x \sim y$ und der Transitivität von $\sim$ schließlich $x \sim z$ und somit $z \in [x]_\sim$ ergibt.

Also sind die beiden Mengen $[x]_\sim = [y]_\sim$ gleich und dies endet den Beweis. □

Kommen wir nun zurück zum Ring der ganzen Zahlen. Für jedes $m \in \mathbb{Z}$ ist $\equiv_m$ eine Äquivalenzrelation auf $\mathbb{Z}$ und wir können die Quotientenmenge betrachten. Da für negative m keine neue Äquivalenzrelationen entstehen, beschränken wir uns im Folgenden auf $m \in \mathbb{N}_0$:

Definition 3.4.11. Es sei $m \in \mathbb{N}_0$. Dann definieren wir mit

$$\mathbb{Z}/m\mathbb{Z} := \mathbb{Z}/\equiv_m$$

die Menge aller Äquivalenzklassen modulo m.

Andere Notationen, die in der Literatur zu finden sind, sind: $\mathbb{Z}/_{m\mathbb{Z}} = \mathbb{Z}/_m = \mathbb{Z}_m$.

Die Äquivalenzklasse von x modulo m schreiben wir auch als $[x]_m := [x]_{\equiv_m}$.

Satz 3.4.12. *Es sei $m \in \mathbb{N}_0$. Dann gibt es auf der Quotientenmenge $\mathbb{Z}/m\mathbb{Z}$ genau eine Ringstruktur, sodass die Quotientenabbildung*

$$q : \mathbb{Z} \to \mathbb{Z}/m\mathbb{Z}, \quad x \mapsto [x]_m$$

zu einem Ringhomomorphismus wird.

Für $[x]_m, [y]_m, \in \mathbb{Z}/m\mathbb{Z}$ gilt dann $[x]_m + [y]_m = [x+y]_m$ und $[x]_m \cdot [y]_m = [x \cdot y]_m$. Das Einselement ist $[1]_m$, das Nullelement ist $[0]_m$.

Beweis. Die Quotientenabbildung $q : \mathbb{Z} \to \mathbb{Z}/m\mathbb{Z}$ ist surjektiv. Es ist also jedes Element in $\mathbb{Z}/m\mathbb{Z}$ von der Form $[x]_m$ für ein (eventuell nicht eindeutiges) $x \in \mathbb{Z}$.

Wenn also eine Ringstruktur $(\mathbb{Z}/m\mathbb{Z}, +, \cdot)$ existiert, sodass q ein Ringhomomorphismus wird, dann heißt das, dass für $[x]_m, [y]_m, \in \mathbb{Z}/m\mathbb{Z}$ die Formeln $[x]_m + [y]_m = [x+y]_m$ und $[x]_m \cdot [y]_m = [x \cdot y]_m$ gelten. Somit ist die Ringstruktur eindeutig bestimmt durch die Ringstruktur auf $\mathbb{Z}$.

Es bleibt also nur die Existenz zu zeigen. Die Idee ist hierbei, dass wir diese Formeln als Definition nehmen, dass wir also definieren:

$$+ : \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z} \to \mathbb{Z}/m, \quad ([x]_m, [y]_m) \mapsto [x+y]_m$$

und

$$\cdot : \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z} \to \mathbb{Z}/m, \quad ([x]_m, [y]_m) \mapsto [x \cdot y]_m.$$

Die Frage ist nun: Sind diese Abbildungen wohldefiniert?

Wohldefiniert bedeutet hierbei: Wird wirklich jedem Element im Definitionsbereich $\mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z}$ genau ein Element im Zielbereich $\mathbb{Z}/m\mathbb{Z}$ zugeordnet?

Da die Quotientenabbildung surjektiv ist, kann man jedes Element in $\mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z}$ als $([x]_m, [y]_m)$ für $x, y \in \mathbb{Z}$ schreiben. Da diese Darstellung aber nicht eindeutig ist (die Quotientenabbildung wird im Allgemeinen (fast nie) nicht injektiv sein), bleibt die Frage, ob das Ergebnis $[x+y]_m$, bzw. $[x \cdot y]_m$ von dieser Darstellung abhängig ist oder nicht. Wir werden nun zeigen, dass sowohl die Addition als auch die Multiplikation auf $\mathbb{Z}/m\mathbb{Z}$ wohldefiniert ist:

Wohldefiniertheit der Addition:
Nehmen wir an, ein Element im Definitionsbereich von $+ : \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z} \to \mathbb{Z}/m$ hat zwei Darstellungen, also

$$([x_1]_m, [y_1]_m) = ([x_2]_m, [y_2]_m) \quad \text{mit } x_1, y_1, x_2, y_2 \in \mathbb{Z}.$$

Dann wollen wir zeigen, dass $[x_1 + y_1]_m = [x_2 + y_2]_m$ ist.

Aus $([x_1]_m, [y_1]_m) = ([x_2]_m, [y_2]_m)$ folgt, dass $[x_1]_m = [x_2]_m$ und $[y_1]_m = [y_2]_m$. Mit Satz 3.4.10 folgt nun, dass $x_1 \equiv_m x_2$ und $y_1 \equiv_m y_2$ ist, was gleichbedeutend ist mit

$$m|(x_1 - x_2) \quad \text{und} \quad m|(y_1 - y_2).$$

Also gibt es Zahlen $k, l \in \mathbb{Z}$ mit

$$x_1 - x_2 = mk \quad \text{und} \quad y_1 - y_2 = ml.$$

Addieren ergibt dann

$$(x_1 + y_1) - (x_2 + y_2) = m(k + l).$$

Also ist $(x_1 + y_1) - (x_2 + y_2)$ durch m teilbar, es gilt also $x_1 + y_1 \equiv_m x_2 + y_2$ und somit schließlich $[x_1 + y_1]_m = [x_2 + y_2]_m$. Die Addition ist also wohldefiniert.

Wohldefiniertheit der Multiplikation:
Nehmen wir wieder an, ein Element im Definitionsbereich von $\cdot : \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z} \to \mathbb{Z}/m$ hat zwei Darstellungen, also

$$([x_1]_m, [y_1]_m) = ([x_2]_m, [y_2]_m) \quad \text{mit } x_1, y_1, x_2, y_2 \in \mathbb{Z}.$$

Wie oben folgt dann, dass Zahlen $k, l \in \mathbb{Z}$ gibt mit

$$x_1 - x_2 = mk \quad \text{und} \quad y_1 - y_2 = ml.$$

Wir wollen nun zeigen, dass $x_1 y_1 \equiv_m x_2 y_2$ gilt, dass also die Differenz durch m teilbar ist. Dies geschieht durch den beliebten Trick des „geschickten Einfügens einer Null“:

$$\begin{aligned} x_1 y_1 - x_2 y_2 &= x_1 y_1 - x_1 y_2 + x_1 y_2 - x_2 y_2 \\ &= x_1(y_1 - y_2) + (x_1 - x_2) y_2 \\ &= y_1 \cdot ml + mk \cdot y_2 \\ &= m(y_1 l + k y_2). \end{aligned}$$

Also teilt die Zahl m die Differenz $x_1 y_1 - x_2 y_2$, woraus $x_1 y_1 \equiv_m x_2 y_2$ und schließlich $[x_1 y_1]_m = [x_2 y_2]_m$ folgt.

Nachweis, dass $(\mathbb{Z}/m\mathbb{Z}, +, \cdot)$ ein Ring ist:
Wir wissen nun also, dass die binären Verknüpfungen + und $\cdot$ wohldefiniert sind und dass

$$q(x + y) = q(x) + q(y) \quad \text{und} \quad q(x \cdot y) = q(x) \cdot q(y)$$

gilt. Da $(\mathbb{Z}, +, \cdot)$ ein kommutativer Ring ist, und die Operationen auf $\mathbb{Z}/m\mathbb{Z}$ durch genau diese Operationen auf $\mathbb{Z}$ eingeführt wurden, übertragen sich sämtliche Eigenschaften auf $\mathbb{Z}/m\mathbb{Z}$. Wir erhalten, dass $\mathbb{Z}/m\mathbb{Z}$ ein Ring ist mit dem Einselement $q(1) = [1]_\sim$ und dem Nullelement $q(0) = [0]_\sim$. Dies beendet den Beweis. □

Lemma 3.4.13. *(a) Falls $m = 0$ ist die Abbildung $q : \mathbb{Z} \to \mathbb{Z}/0\mathbb{Z}$ ein Ringisomorphismus, das heißt: $\mathbb{Z}/0\mathbb{Z} \cong \mathbb{Z}$.*

(b) Falls $m \in \mathbb{N}$ ist die Abbildung

$$b : \{0, \dots, m-1\} \to \mathbb{Z}/m\mathbb{Z}, \quad r \mapsto q(x) = [r]_m$$

eine Bijektion. Also gilt insbesondere $|\mathbb{Z}/m\mathbb{Z}| = m$.

Beweis. (a)
Wir wissen bereits, dass $q : \mathbb{Z} \to \mathbb{Z}/m\mathbb{Z}$ ein surjektiver Ringhomomorphismus ist. Es bleibt also nur Injektivität zu zeigen. Wenn wir also $x, y \in \mathbb{Z}$ mit $q(x) = q(y)$ betrachten, dann bedeutet dies, dass $x \equiv_0 y$ und somit folgt $0|(x - y)$. Nach Lemma 3.4.3 folgt damit, dass $x - y = 0$ und somit $x = y$. Das zeigt: q ist injektiv.
(b)
Wir müssen zeigen, dass die Abbildung $b : \{0, \dots, m-1\} \to \mathbb{Z}/m\mathbb{Z}, \quad x \mapsto q(x) = [x]_m$ bijektiv ist. Für die Surjektivität nehmen wir uns ein Element in $\omega \in \mathbb{Z}/m\mathbb{Z}$. Aus der Surjektivität von q folgt, dass es ein $x \in \mathbb{Z}$ gibt mit $\omega = [x]_m$. Nach Lemma 3.4.1 können wir x schreiben als

$$x = mk + r \quad \text{und} \quad r \in \{0, \dots, m-1\}.$$

Also ist $[x]_m = [r]_m$ und somit $\omega = q(r)$.

Für die Injektivität nehmen wir zwei Elemente $r, s \in \{0, \ldots, m-1\}$ mit $[r]_m = [s]_m$. Dann bedeutet dies, dass $r \equiv_m s$ und somit $m|(r-s)$. Da aber $r, s \in \{0, \ldots, m-1\}$ kann $r-s$ nur durch m teilbar sein, wenn $r-s=0$ ist, wenn also $r=s$. □

Bemerkung 3.4.14. Wir sehen also: Für jedes $m \in \mathbb{N}$ ist $\mathbb{Z}/m\mathbb{Z}$ ein endlicher kommutativer Ring mit genau m Elementen. Insbesondere sind also

$$(\mathbb{Z}/m\mathbb{Z}, +)$$

abelsche Gruppen. Es gibt also für jede natürliche Zahl $m \in \mathbb{N}$ mindestens eine Gruppe mit genau m Elementen.

Endliche Ringe verhalten sich oft anders als unendliche. Beispielsweise gilt:

Lemma 3.4.15. *Es sei R ein endlicher kommutativer Ring mit $0_R \neq 1_R$. Dann sind äquivalent:*

(i) R ist ein Körper.

(ii) R ist nullteilerfrei, d.h.

$$\forall a, b \in R : (ab = 0_R) \implies (a = 0_R \text{ oder } b = 0_R).$$

Für unendliche Ringe sind diese beiden Eigenschaften nicht äquivalent, wie der Ring der ganzen Zahlen zeigt.

Beweis von Lemma 3.4.15. Die Implikation „(i) $\implies$ (ii)“ gilt auch ohne die Annahme der Endlichkeit (siehe Bemerkung 3.3.13). Zeigen wir nun „(ii) $\implies$ (i)“:

Nehmen wir an, dass $|R| = q \in \mathbb{N}$.

Wir müssen zeigen, dass $R^\times = R \setminus \{0_R\}$.

Da $0_R \neq 1_R$, bedeutet dies, dass 0_R nicht invertierbar ist. Das zeigt: „$\subseteq$“.

Es bleibt zu zeigen, dass jedes $a \in R \setminus \{0_R\}$ invertierbar ist. Es sei dazu $a \in R$. Wir betrachten die Teilmenge[11]

$$aR = \{ax \mid x \in R\} \subseteq R$$

und die Abbildung

$$\lambda_a : R \to aR, \quad x \mapsto ax.$$

Wir behaupten nun, dass $\lambda_a : R \to aR$ injektiv ist. Es sei dazu $\lambda_a(x) = \lambda_a(y)$. Das heißt:

$$\begin{aligned} ax &= ay \\ ax - ay &= 0 \\ a(x-y) &= 0 \end{aligned}$$

Also folgt mit (ii), dass $a = 0$ oder $x - y = 0$. Da aber nach Voraussetzung $a \neq 0$ ist, impliziert dies, dass $x = y$ ist.

Also ist $\lambda_a : R \to aR$ injektiv und weil die Abbildung offenbar auch surjektiv ist, ist λ_a eine Bijektion. Also gilt $|aR| = |R| = q$.

[11] Eine solche Teilmenge nennt man auch ein Hauptideal.

Die endliche Menge R hat q Elemente und die Teilmenge aR hat ebenfalls q Elemente. Dies ist bei endlichen Mengen nur möglich, wenn $R = aR$.

Insbesondere folgt somit, dass aR das Einselement enthält, also gibt es ein $x \in R$ mit $ax = 1_R$. Aus der Kommutativität von R folgt dann auch $xa = 1$ und somit ist a invertierbar. □

Wir möchten nun Lemma 3.4.15 auf unsere neu konstruierten Ringe $\mathbb{Z}/m\mathbb{Z}$ anwenden und entscheiden, wann $\mathbb{Z}/m\mathbb{Z}$ ein Körper ist.

Wir erinnern an den folgenden, wohlbekannten Begriff:

Definition 3.4.16. Eine Zahl $p \in \mathbb{Z}$ ist eine *Primzahl*, wenn

$$p > 1 \quad \text{und} \quad \forall m \in \mathbb{N} : (m|p \implies (m = 1 \text{ oder } m = p)).$$

Die Bedeutung der Primzahlen liegt in folgendem altbekannten Satz:

Satz 3.4.17 (Fundamentalsatz der Arithmetik)**.** *Jede natürliche Zahl $n \in \mathbb{N}$ lässt sich als (endliches) Produkt von Primzahlen schreiben. Diese Darstellung ist eindeutig bis auf Umordnung (Permutation) der Primfaktoren.*

Für uns wichtiger ist im Moment aber die folgende Aussage:

Lemma 3.4.18 (Lemma von Euklid[12])**.** *Es sei p eine Primzahl. Dann gilt für alle $a, b \in \mathbb{Z}$:*

$$p|(ab) \implies (p|a \text{ oder } p|b).$$

Lemma 3.4.18 folgt leicht aus Satz 3.4.17 (man schreibt einfach $ab = pk$ und zerlegt a, b und k in Primfaktoren. Dann folgt aus der Eindeutigkeit der Darstellung, dass p bereits einer der Faktoren von a oder b gewesen sein muss).

Fairerweise sollte man aber anmerken, dass man normalerweise Satz 3.4.17 (zumindest die Eindeutigkeit der Zerlegung) mit Hilfe von Lemma 3.4.18 beweist (man nimmt an, es gäbe zwei unterschiedliche Darstellungen einer Zahl. Dann gilt für jeden Primfaktor in der ersten Primfaktordarstellung, dass er auch einen der Faktoren der anderen Seite teilen muss). Um hier also nicht in einen Zirkelschluss zu geraten, müsste man eine der beiden Aussagen beweisen, ohne die andere zu verwenden. Dies werden wir aber in dieser Veranstaltung nicht tun, weil uns dies inhaltlich zu weit von der linearen Algebra fortführen würde. Wir verweisen auf andere Veranstaltungen wie z.B. Einführung in die Algebra und Zahlentheorie oder praktisch jedes Buch über elementare Zahlentheorie.

Satz 3.4.19. *Für eine natürliche Zahl $m \in \mathbb{N}$ sind äquivalent:*

(i) m ist eine Primzahl.

(ii) Der Ring $\mathbb{Z}/m\mathbb{Z}$ ist ein Körper.

Beweis. „(i) $\implies$ (ii)“:
Wenn m eine Primzahl ist, dann gilt, dass $m > 1$ ist. Das wiederum bedeutet, dass $[0]_m \neq [1]_m$ gilt. Nach Lemma 3.4.13 gilt, dass $|\mathbb{Z}/m\mathbb{Z}| = m$. Also ist $\mathbb{Z}/m\mathbb{Z}$ insbesondere ein endlicher Ring. Wir können also Lemma 3.4.15 anwenden und erhalten, dass der Ring $\mathbb{Z}/m\mathbb{Z}$ ein Körper ist, falls er nullteilerfrei ist. Es sei also

$$[a]_m \cdot [b]_m = [0]_m.$$

[12] nach EUKLID VON ALEXANDRIA, griechischer Mathematiker, hat wahrscheinlich im 3. Jahrhundert v. Chr. gelebt

Das Produkt auf der linken Seite ist $[ab]_m$. Wir erhalten also die Aussage, das m ein Teiler von ab ist. Nach (i) ist m eine Primzahl, also folgt mit dem Lemma von Euklid (Lemma 3.4.18), dass $m|a$ oder $m|b$, was wiederum gleichbedeutend ist mit $[a]_m = [0]_m$ oder $[b]_m = [0]_m$.

Also ist $\mathbb{Z}/m\mathbb{Z}$ nullteilerfrei und nach Lemma 3.4.15 ein Körper.

„(ii) $\Longrightarrow$ (i)“:
Da $\mathbb{Z}/m\mathbb{Z}$ ein Körper ist, gilt insbesondere $[0]_m \neq [1]_m$ (siehe Bemerkung 3.3.13). Also folgt, dass $m > 1$ sein muss. Um zu zeigen, dass m eine Primzahl ist, müssen wir annehmen, $n|m$ mit $n \in \mathbb{N}$ und folgern, dass $n = 1$ oder $n = m$ ist.

Da n ein Teiler von m ist, gibt es ein $k \in \mathbb{Z}$ (was aber auch in $\mathbb{N}$ liegen muss, da m, n beide positiv sind) mit

$$m = n \cdot k.$$

Wenn wir nun den Ringhomomorphismus $q : \mathbb{Z} \to \mathbb{Z}/m\mathbb{Z}, \quad x \mapsto [x]_m$ auf diese Gleichung anwenden, erhalten wir:

$$[m]_m = [n]_m \cdot [k]_m.$$

Da m durch m teilbar ist, ist die linke Seite der Gleichung gleich $[0]_m$. Da nach Voraussetzung (ii) der Ring $\mathbb{Z}/m\mathbb{Z}$ ein Körper ist und Körper nullteilerfrei sind, folgt somit

$$[n]_m = [0]_m \text{ oder } [k]_m = [0]_m,$$

was nichts anderes bedeutet, als dass $m|n$ oder $m|k$ gilt. Im ersten Falle heißt das, dass $m|n$ und $n|m$ und somit gilt $|m| = |n|$ nach Lemma 3.4.3. Da $m, n > 0$, folgt damit $n = m$.

Nehmen wir nun an, dass $m|k$. Aus $m = nk$ folgt, dass $k|m$ und somit $|m| = |k|$ nach Lemma 3.4.3. Wieder folgt: $m = k$ und somit $n = 1$. Das beendet den Beweis, dass m eine Primzahl ist. □

Bemerkung 3.4.20. Wir wissen nun also: Für jede Primzahl p gibt es einen Körper mit p Elementen, nämlich $\mathbb{Z}/p\mathbb{Z}$. Diese endlichen Körper spielen in der Informatik in der Kodierungstheorie und in der Kryptographie eine wesentliche Rolle.

Für die Primzahl $p = 2$ erhalten wir so einen Körper mit 2 Elementen. Wir haben bereits in Beispiel 3.3.14 einen Körper mit 2 Elementen kennengelernt: $\mathbb{F}_2$ und gesehen, dass es – bis auf Umbenennung der Elemente – keinen anderen Körper mit 2 Elementen geben kann. Es gilt also: $\mathbb{Z}/2\mathbb{Z}$ und $\mathbb{F}_2$ sind isomorph.

Interessanterweise ist $\mathbb{Z}/4\mathbb{Z}$ nicht isomorph zu $\mathbb{F}_4$. Dies sieht man direkt daran, dass $\mathbb{F}_4$ ein Körper ist, während der Ring $\mathbb{Z}/4\mathbb{Z}$ kein Körper ist (weil 4 keine Primzahl ist).

Wir sehen also: Die Körper $\mathbb{Z}/p\mathbb{Z}$ geben uns viele Beispiele für endliche Körper, aber es gibt auch endliche Körper, die nicht von dieser Form sind.

Zusammenfassung von Abschnitt 3.4

(1) Für jede Äquivalenzrelation $\sim$ auf einer Menge X gibt es eine Quotientenmenge $X/\sim$ und eine Quotientenabbildung $q : X \to X/\sim$.

(2) Für jedes $m \in \mathbb{N}$ ist $Z/m\mathbb{Z}$ ein endlicher kommutativer Ring mit genau m Elementen.

(3) Für jede Primzahl p ist $\mathbb{Z}/m\mathbb{Z}$ sogar ein Körper, aber nicht jeder endliche Körper ist von dieser Form (Beispiel $\mathbb{F}_4$).

(4) Für endliche kommutative Ringe mit $1 \neq 0$ ist die Eigenschaft, ein Körper zu sein äquivalent zu Nullteilerfreiheit.

3.5. Der Körper der komplexen Zahlen

Der Körper der rationalen Zahlen $(\mathbb{Q}, +, \cdot)$ ist ein Unterkörper von $\mathbb{R}$, d.h. mit den Grundrechenarten Addition, Multiplikation sowie deren Umkehrungen Subtraktion und Division kann man die rationalen Zahlen nicht verlassen. Alle Ausdrücke, die nur mit den Grundrechenarten aus rationalen Zahlen aufgebaut werden, sind selbst rationale Zahlen.

Dies ist für die lineare Algebra besonders wichtig, da dies bedeutet, dass alles, was wir für lineare Gleichungssysteme und Matrizen mit reellen Koeffizienten bzw. Einträgen gilt, analog auch für rationale Einträge gelten muss, da wir immer nur die Rechenregeln verwendet haben, die in jedem Körper, also auch in $\mathbb{Q}$ gelten.

Sobald man nicht-lineare Algebra betreibt, sieht die Sache anders aus:

Die Gleichung $x^2 = 2$ hat – obwohl alle Koeffizienten aus $\mathbb{Q}$ sind, keine Lösung in $\mathbb{Q}$, denn in den rationalen Zahlen gibt es keine Zahl, die quadriert 2 ergibt. In den reellen Zahlen $\mathbb{R}$ gibt es nun Lösungen für die nichtlineare Gleichung, nämlich $-\sqrt{2}$ und $\sqrt{2}$. Wir sehen also: Durch Übergang zu einer Körpererweiterung von $\mathbb{Q}$ sind nun plötzlich Gleichungen lösbar, die vorher nicht lösbar waren.

Nun kann man sich zwei Dinge fragen: Erstens: Gibt es auch noch Körper „zwischen“ $\mathbb{Q}$ und $\mathbb{R}$?

Die Antwort darauf ist Ja. Man kann beispielsweise die Menge $L := \{\alpha + \beta\sqrt{2} \mid \alpha, \beta \in \mathbb{Q}\}$ betrachten. Dann kann man zeigen: Dies ist echter Unterkörper von $\mathbb{R}$, der wiederum eine Körpererweiterung von $\mathbb{Q}$ ist: $\mathbb{Q} \subsetneq L \subsetneq \mathbb{R}$. In diesem so konstruieren Körper ist dann die Gleichung $x^2 = 2$ lösbar – allerdings sind ähnliche Gleichungen wie $x^2 = 3$ immer noch nicht lösbar.

Es gibt aber noch eine zweite Frage: Sind in $\mathbb{R}$ denn nun alle Gleichungen dieser Art lösbar?

Die Antwort hierauf ist leider: Nein.

Lemma 3.5.1. *Die Gleichung*

$$x^2 = -1$$

hat im Körper der reellen Zahlen keine Lösung.

Beweis. Es sei $x \in \mathbb{R}$. Dann ist $x > 0$ oder $x < 0$ oder $x = 0$.

Falls $x > 0$ ist, dann ist $x^2 = x \cdot x > 0$ und somit nicht gleich (-1).

Falls $x < 0$ ist, dann ist $-x > 0$ und somit $x^2 = (-x) \cdot (-x) > 0$ und somit nicht gleich (-1).

Falls $x = 0$ ist, dann ist $x^2 = 0$ und auch nicht gleich (-1). □

Die Frage ist nun: Gibt es eine Körpererweiterung K von $\mathbb{R}$, in der es ein Element $\mathrm{i} \in K$ gibt mit

$$\mathrm{i}^2 = -1?$$

Bevor wir die Frage beantworten, ob es eine solche Körpererweiterung gibt, nehmen wir einmal an, es gäbe sie.

Genau genommen nehmen wir also an, es gibt einen Körper $(K, +, \cdot)$, der $\mathbb{R} \subseteq K$ als Unterkörper enthält, aber in dem es ein Element $\mathrm{i} \in K$ gibt mit $\mathrm{i}^2 = -1$. Historisch war das der Ansatz, der lange Zeit verfolgt wurde. Man hat einfach mit diesem i (oft auch als $\sqrt{-1}$ notiert) gerechnet und sich nicht wirklich darum gekümmert, ob ein solcher Körper überhaupt „existiert“[13].

Wir können nun also beliebige Ausdrücke aus reellen Zahlen und diesem neuen Element i bilden, wie z.B.

$$13\mathrm{i}^4 - 5\mathrm{i}^3 + 12\mathrm{i}^2 - 6\mathrm{i} + 5$$

und die üblichen Körperregeln verwenden, wie Kommutativgesetze, Assoziativgesetze und das Distributivgesetz.

Da aber nach Voraussetzung $\mathrm{i}^2 = -1$ gilt, können wir den obigen Ausdruck folgendermaßen umformen:

$$13\mathrm{i}^4 - 5\mathrm{i}^3 + 12\mathrm{i}^2 + 6\mathrm{i} + 5 = 13\mathrm{i}^2 \cdot \mathrm{i}^2 - 5\mathrm{i}^2 \cdot \mathrm{i} + 12\mathrm{i}^2 + 6\mathrm{i} + 5 = 13(-1)(-1) - 5(-1)\mathrm{i} + 12(-1) + 6\mathrm{i} + 5 = 6 + 11\mathrm{i}.$$

Es sieht so aus, als ließe sich $6 + 11\mathrm{i}$ nicht weiter vereinfachen.

Lemma 3.5.2. *Es sei K eine Körpererweiterung von $\mathbb{R}$ und sei $\mathrm{i} \in K$ mit $\mathrm{i}^2 = -1$. Wenn sich ein Element $z \in K$ schreiben lässt als $z = a + \mathrm{i}b$ mit $a, b \in \mathbb{R}$, dann ist diese Darstellung eindeutig.*

Insbesondere gilt also

$$(a + \mathrm{i}b = 0) \iff (a = 0 \text{ und } b = 0).$$

Beweis. Angenommen es gibt zwei unterschiedliche Darstellungen

$$z = a + \mathrm{i}b = x + \mathrm{i}y \quad \text{mit } a, b, x, y \in \mathbb{R}.$$

Dann müssen wir zeigen, dass $a = x$ und $b = y$.

Wir beginnen mit der Gleichung $a + \mathrm{i}b = x + \mathrm{i}y$ und formen sie wie folgt um:

$$\begin{aligned} a + \mathrm{i}b &= x + \mathrm{i}y \\ a - x &= \mathrm{i}y - \mathrm{i}b \\ a - x &= \mathrm{i}(y - b) \\ (a - x)^2 &= \mathrm{i}^2(y - b)^2 \\ (a - x)^2 &= (-1)(y - b)^2. \end{aligned}$$

Die linke Seite der Gleichung ist nun das Quadrat einer reellen Zahl und somit größer oder gleich 0. Die rechte Seite ist (-1) mal das Quadrat einer reellen Zahl und somit kleiner oder gleich 0. Die einzige Möglichkeit, wie die beiden Seiten also wirklich gleich sein können, ist wenn beide Seiten gleich 0 sind.

Also ist $(a - x)^2 = 0$ und $(b - y)^2 = 0$, woraus (weil Körper nullteilerfrei sind) folgt, dass $a - x = 0$ und $b - y = 0$.

Das zeigt: $a = x$ und $b = y$. □

[13] Auch war das Konzept eines *Körpers* natürlich noch lange nicht vorhanden.

Lemma 3.5.3. *Es sei K eine Körpererweiterung von $\mathbb{R}$ und sei $\mathrm{i} \in K$ mit $\mathrm{i}^2 = -1$. Die Menge $\{a + \mathrm{i}b \mid a, b \in \mathbb{R}\}$ ist ein Unterkörper von K. Es gelten die folgenden Rechenregeln für $z_1 = a_1 + \mathrm{i}b_1$ und $z_2 = a_2 + \mathrm{i}b_2$:*

(a) $(a_1 + \mathrm{i}b_1) + (a_2 + \mathrm{i}b_2) = (a_1 + a_2) + \mathrm{i}(b_1 + b_2)$.

(b) $(a_1 + \mathrm{i}b_1) - (a_2 + \mathrm{i}b_2) = (a_1 - a_2) + \mathrm{i}(b_1 - b_2)$.

(c) $(a_1 + \mathrm{i}b_1) \cdot (a_2 + \mathrm{i}b_2) = (a_1a_2 - b_1b_2) + \mathrm{i}(a_1b_2 + a_2b_1)$.

(d) Für $a + \mathrm{i}b \neq 0$ gilt: $\frac{1}{a+\mathrm{i}b} = \frac{a-\mathrm{i}b}{(a+\mathrm{i}b)(a-\mathrm{i}b)} = \frac{a-\mathrm{i}b}{a^2+b^2} = \frac{a}{a^2+b^2} + \mathrm{i} \cdot \frac{-b}{a^2+b^2}$.

Beweis. Die Aussagen (a) und (b) folgen direkt aus den Körperaxiomen und den in allen Körpern gültigen Rechenregeln.

Aussage (c) ist eine einfache Anwendung des Distributivgesetzes (Ausmultiplizieren) und anschließender Verwendung von $\mathrm{i}^2 = -1$.

Schließlich folgt auch (d) direkt aus den allgemeingültigen Rechenregeln. Der einzige Schritt, der gerechtfertigt werden müsste, ist dass man mit $a - \mathrm{i}b$ erweitern darf, d.h. dass $a - \mathrm{i}b \neq 0$. Dies folgt allerdings aus $a + \mathrm{i}b \neq 0$ und Lemma 3.5.2. □

Wenn es also eine Körpererweiterung K von $\mathbb{R}$ gibt mit einem Element $\mathrm{i} \in K$, das quadriert (-1) ergibt, dann gibt es einen kleinsten mit dieser Eigenschaft.

Satz 3.5.4 (Existenz der komplexen Zahlen)**.** *Es existiert eine Körpererweiterung $\mathbb{C}$ von $\mathbb{R}$, die ein Element $\mathrm{i} \in \mathbb{C}$ enthält mit*

$$\mathrm{i}^2 = -1$$

und in dem sich jedes Element $z \in \mathbb{C}$ eindeutig in der Form

$$z = a + \mathrm{i}b \text{ mit } a, b, \in \mathbb{R}$$

schreiben lässt.

Diese Körpererweiterung nennen wir den Körper der komplexen Zahlen*.*

Beweis. Mal angenommen, wir wüssten schon, dass der Satz, den wir gerade beweisen wollen, wahr wäre, dann gäbe es eine Abbildung zwischen $\mathbb{R}^2$ und $\mathbb{C}$:

$$\psi : \mathbb{R}^2 \to \mathbb{C}, \quad \begin{pmatrix} a \\ b \end{pmatrix} \mapsto a + \mathrm{i}b.$$

Diese Abbildung wäre eine Bijektion mit $\psi^{-1}(1) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ und $\psi^{-1}(\mathrm{i}) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

Somit können wir die Menge $\mathbb{R}^2$ verwenden, um die gesuchte Menge $\mathbb{C}$ zu konstruieren.

Wir wissen bereits, dass $(\mathbb{R}^2, +)$ mit der Vektoraddition eine kommutative Gruppe ist. Wir definieren die Multiplikation

$$\bullet : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}^2, \quad \left(\begin{pmatrix} a_1 \\ b_1 \end{pmatrix}, \begin{pmatrix} a_2 \\ b_2 \end{pmatrix} \right) \mapsto \begin{pmatrix} a_1a_2 - b_1b_2 \\ a_1b_2 + a_2b_1 \end{pmatrix}.$$

Man sieht sofort, dass $\bullet$ kommutativ ist und durch direktes Nachrechnen sieht man, dass $\bullet$ assoziativ ist. Das Element $1_{\mathbb{C}} := \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ ist ein neutrales Element für $\bullet$, sodass $(\mathbb{R}^2, \bullet)$ ein kommutatives Monoid wird.

Das Distributivgesetz gilt auch (nachrechnen!), und somit ist $(\mathbb{R}^2, +, \bullet)$ ein kommutativer Ring. Wenn wir nun $\mathrm{i} := \begin{pmatrix}0\\1\end{pmatrix}$ setzen, sieht man, dass

$$a1_{\mathbb{C}} + b\mathrm{i} = a\begin{pmatrix}1\\0\end{pmatrix} + b\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}a\\0\end{pmatrix} + \begin{pmatrix}0\\b\end{pmatrix} = \begin{pmatrix}a\\b\end{pmatrix}$$

und dass

$$\mathrm{i}^2 = \mathrm{i} \bullet \mathrm{i} = \begin{pmatrix}0\\1\end{pmatrix} \bullet \begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}0\cdot 0 - 1\cdot 1\\0\cdot 1 + 1\cdot 0\end{pmatrix} = -1_{\mathbb{C}}.$$

Jedes Element $\begin{pmatrix}a\\b\end{pmatrix} = a1_{\mathbb{C}} + b\mathrm{i} \neq \begin{pmatrix}0\\0\end{pmatrix}$ ist invertierbar, weil

$$\begin{aligned}= \begin{pmatrix}a\\b\end{pmatrix} \bullet \left(\frac{1}{a^2+b^2}\begin{pmatrix}a\\-b\end{pmatrix}\right) &= \begin{pmatrix}a\\b\end{pmatrix} \bullet \begin{pmatrix}\frac{a}{a^2+b^2}\\\frac{-b}{a^2+b^2}\end{pmatrix}\\ &= \begin{pmatrix}a\cdot\frac{a}{a^2+b^2} - b\cdot\frac{-b}{a^2+b^2}\\a\cdot\frac{-b}{a^2+b^2} + b\cdot\frac{a}{a^2+b^2}\end{pmatrix}\\ &= \begin{pmatrix}\frac{a^2+b^2}{a^2+b^2}\\\frac{ab-ab}{a^2+b^2}\end{pmatrix} = \begin{pmatrix}1\\0\end{pmatrix} = 1_{\mathbb{C}}.\end{aligned}$$

Somit ist $\mathbb{C} := (\mathbb{R}^2, +, \bullet)$ ein Körper mit Nullelement $\begin{pmatrix}0\\0\end{pmatrix}$ und Einselement $1_{\mathbb{C}} = \begin{pmatrix}1\\0\end{pmatrix}$.

Die folgende Abbildung ist ein injektiver Ring-Homomorphismus, wie man leicht nachrechnet:

$$\Phi : \mathbb{R} \to \mathbb{C}, \quad a \mapsto a1_{\mathbb{C}} = \begin{pmatrix}a\\0\end{pmatrix}.$$

Also ist das $\Phi(\mathbb{R})$ ein Unterring von $\mathbb{C}$ isomorph zu $\mathbb{R}$. Es ist daher möglich, die Notation zu vereinfachen und $\mathbb{R}$ mit dem Bild unter dieser Einbettung zu identifizieren.

Somit können wir von nun an annehmen, dass $\mathbb{R} \subseteq \mathbb{C}$ ein Unterring (sogar ein Unterkörper) ist. Das beendet den Beweis. □

Bemerkung 3.5.5. Es gibt neben der Konstruktion $\mathbb{C} := \mathbb{R}^2$ auch noch andere Möglichkeiten, Satz 3.5.4 zu beweisen: Beispielsweise kann man sich folgenden Ring ansehen (siehe Beispiel 3.3.7):

$$C := \left\{ \begin{pmatrix}\alpha & -\beta\\\beta & \alpha\end{pmatrix} \,\middle|\, \alpha, \beta \in \mathbb{R} \right\}$$

Dies ist ein Unterring von $\mathbb{R}^{2\times 2}$ und somit ein Ring. Wir definieren

$$I := \begin{pmatrix}0 & -1\\1 & 0\end{pmatrix} \in \mathbb{C}$$

und stellen fest, dass jede Matrix $Z \in \mathbb{C}$ sich eindeutig schreiben lässt als

$$Z = \alpha \mathbb{1}_2 + \beta I.$$

Hieraus folgt direkt, dass der Ring C kommutativ ist (obwohl der größere Ring $\mathbb{R}^{2\times 2}$ nicht kommutativ ist). Außerdem ist $I \cdot I = -\mathbb{1}_2$ (direktes Nachrechnen).

Nun bleibt nur noch zu sehen, dass jedes $Z = \alpha \mathbb{1}_2 + \beta I \in C$ invertierbar ist, weil

$$Z \cdot \frac{1}{\alpha^2 + \beta^2} Z^\top = \mathbb{1}_2.$$

Also ist C ein Körper.

Wieder können wir $\mathbb{R}$ in C als Unterkörper einbetten über

$$\Phi : \mathbb{R} \to C, \quad a \mapsto a \mathbb{1}_2.$$

Auf diese Weise erhalten wir ein anderes, aber äquivalentes Modell der komplexen Zahlen.

Eine noch andere Möglichkeit, die komplexen Zahlen zu konstruieren, ist als Quotientenring eines Polynomrings, aber dazu vielleicht später mehr. . . .

Wichtig an dieser Stelle ist nur, dass es unwesentlich ist, wie genau man die komplexen Zahlen konstruiert, solange man weiß, wie man mit ihnen rechnet. Dasselbe Konzept haben wir auch schon bei den reellen Zahlen angewendet: Auch hier haben wir nie die Frage gestellt, was eine reelle Zahl *ist*. Wichtig ist nur, was man auf der Menge der reellen Zahlen für Strukturen hat.

Wie man konkret die reellen Zahlen konstruiert, ist nebensächlich.

Notation 3.5.6. Jede komplexe Zahl $z \in \mathbb{C}$ hat eine eindeutige Darstellung als $z = a + \mathrm{i}b$ mit $a, b \in \mathbb{R}$. Die Zahl a nennen wir den *Realteil* von z und schreiben $\operatorname{Re} z = a$. Die Zahl b nennen wir den *Imaginärteil* von z und schreiben $\operatorname{Im} z = b$.

Jede komplexe Zahl $z = a + \mathrm{i}b$ können wir geometrisch mit einem Punkt in einem 2-dimensionalen Koordinatensystem identifizieren. Die Menge $\mathbb{C}$ mit dieser geometrischen Interpretation nennt man auch die Gaußsche[14] Zahlenebene. Die x-Achse entspricht der Menge $\mathbb{R}$ und heißt dann auch die *reelle Achse*. Die y-Achse entspricht der Menge $\mathrm{i}\mathbb{R} = \{\mathrm{i}b \mid b \in \mathbb{R}\}$, genannt die *imaginäre Achse*. Komplexe Zahlen der Form $\mathrm{i}b$ nennt man auch *imaginäre Zahlen*.

Zu jeder komplexen Zahl $z = a + \mathrm{i}b$ definieren wir die *konjugiert komplexe Zahl*

$$\overline{z} := a - \mathrm{i}b$$

In der Gaußschen Zahlenebene entspricht dies der Spiegelung an der x-Achse (der reellen Achse).

Das Produkt $z\overline{z} = (a + \mathrm{i}b)(a - \mathrm{i}b) = a^2 - \mathrm{i}^2 b^2 = a^2 + b^2$ ist immer reell und nichtnegativ. Somit kann man daraus die (gewöhnliche reelle) Quadratwurzel ziehen und erhält dadurch den *Betrag*

$$|z| = \sqrt{z\overline{z}} = \sqrt{a^2 + b^2}.$$

Geometrisch gibt der Betrag $|z|$ den Abstand von z zum Ursprung (also zur komplexen Zahl 0) an. Dies ist eine Verallgemeinerung des reellen Betrags, weil für $x \in \mathbb{R}$ gilt:

$$|x| = \sqrt{x\overline{x}} = \sqrt{x^2} = \begin{cases} x & \text{für } x \geq 0 \\ -x & \text{für } x \leq 0. \end{cases}$$

Die Menge aller komplexen Zahlen mit Betrag 1 ist der *Einheitskreis*

$$S^1 := \{z \in \mathbb{C} \mid |z| = 1\}$$

in der Gaußschen Zahlenebene.

[14] benannt nach demselben GAUSS, den wir schon vom Gauß-Algorithmus kennen.

Lemma 3.5.7. *Die gerade eingeführten geometrischen Begriffe haben folgende algebraische Eigenschaften:*

(a) Die Abbildungen $\mathrm{Re}:\mathbb{C}\to\mathbb{R}$ *und* $\mathrm{Im}:\mathbb{C}\to\mathbb{R}$ *sind Gruppenhomomorphismen bezüglich der additiven Gruppen* $(\mathbb{C},+)$ *und* $(\mathbb{R},+)$*. Sie sind aber* keine *Ringhomomorphismen.*

(b) Die Menge der imaginären Zahlen $\mathrm{i}\mathbb{R}$ *ist eine Untergruppe von* $(\mathbb{C},+)$*, aber kein Unterring.*

(c) Die Konjugationsabbildung $\mathbb{C}\to\mathbb{C},\quad z\mapsto\overline{z}$ *ist ein Körperautomorphismus (also ein Ringautomorphismus auf einem Körper).*

(d) Die Betragsabbildung $\mathbb{C}\to[0,+\infty),\quad z\mapsto|z|$ *ist multiplikativ und wird eingeschränkt auf* $\mathbb{C}^\times$ *zu einem surjektiven Gruppenhomomorphismus*

$$\left(\mathbb{C}^\times,\cdot\right)\to((0,+\infty),\cdot),\quad z\mapsto|z|$$

Der Kern dieses Gruppenhomomorphismus ist der Einheitskreis $S^1\subseteq\mathbb{C}^\times$*. Also ist* S^1 *insbesondere eine Untergruppe der Gruppe* $(\mathbb{C}^\times)$*.*

Beweis. (a)
Es seien $z_1=a_1+\mathrm{i}b_1$ und $z_2=a_2+\mathrm{i}b_2$ gegeben. Dann gilt

$$\mathrm{Re}(z_1+z_2)=\mathrm{Re}(a_1+\mathrm{i}b_1+a_2+\mathrm{i}b_2)=\mathrm{Re}((a_1+a_2)+\mathrm{i}(b_1+b_2))=a_1+a_2=\mathrm{Re}(z_1)+\mathrm{Re}(z_2).$$

Dann gilt

$$\mathrm{Im}(z_1+z_2)=\mathrm{Im}(a_1+\mathrm{i}b_1+a_2+\mathrm{i}b_2)=\mathrm{Im}((a_1+a_2)+\mathrm{i}(b_1+b_2))=b_1+b_2=\mathrm{Im}(z_1)+\mathrm{Im}(z_2).$$

(b)
Die Menge $\mathrm{i}\mathbb{R}$ ist der Kern des Gruppenhomomorphismus $\mathrm{Re}:\mathbb{C}\to\mathbb{R}$ und somit eine Untergruppe. Da $\mathrm{i}\in\mathrm{i}\mathbb{R}$ liegt, aber $\mathrm{i}\cdot\mathrm{i}=\mathrm{i}^2=-1$ nicht, ist $\mathrm{i}\mathbb{R}$ kein Unterring.
(c)
Wenn wir die Konjugationsabbildung

$$\kappa:\mathbb{C}\to\mathbb{C},\quad z=a+\mathrm{i}b\mapsto\overline{z}=a-\mathrm{i}b$$

mit sich selbst verketten, erhalten wir für alle $z=a+\mathrm{i}b\in\mathbb{C}$:

$$(\kappa\circ\kappa)(a+\mathrm{i}b)=\overline{\overline{a+\mathrm{i}b}}=\overline{a-\mathrm{i}b}=a+\mathrm{i}b.$$

Somit gilt $\kappa\circ\kappa=\mathrm{id}_{\mathbb{C}}$ und somit ist κ bijektiv mit Umkehrabbildung $\kappa^{-1}=\kappa$.

Es bleibt zu zeigen, dass κ ein Ringhomomorphismus ist. Da κ alle reellen Zahlen festhält (also $\kappa(t)=t$ für alle $t\in\mathbb{R}$), wird insbesondere 1 auf 1 abgebildet. Es bleibt zu zeigen, dass κ die Addition und die Multiplikation erhält.

Es seien $z_1=a_1+\mathrm{i}b_1$ und $z_2=a_2+\mathrm{i}b_2$ gegeben. Dann gilt

$$\begin{aligned}\kappa(z_1+z_2)&=\overline{(a_1+\mathrm{i}b_1)+(a_2+\mathrm{i}b_2)}\\&=\overline{(a_1+a_2)+\mathrm{i}(b_1+b_2)}\\&=(a_1+a_2)-\mathrm{i}(b_1+b_2)\\&=(a_1-\mathrm{i}b_1)+(a_2-\mathrm{i}b_2)\\&=\overline{a_1+\mathrm{i}b_1}+\overline{a_2+\mathrm{i}b_2}\\&=\kappa(z_1)+\kappa(z_2).\end{aligned}$$

Es bleibt nun zu zeigen, dass $\kappa(z_1 \cdot z_2) = \kappa(z_1) \cdot \kappa(z_2)$. Wir vereinfachen die linke Seite der zu zeigenden Gleichung:

$$\begin{aligned}
\kappa(z_1 \cdot z_2) &= \overline{(a_1 + \mathrm{i}b_1) \cdot (a_2 + \mathrm{i}b_2)} \\
&= \overline{(a_1a_2 + \mathrm{i}a_1b_2 + \mathrm{i}b_1a_2 + \mathrm{i}^2 b_1b_2} \\
&= \overline{(a_1a_2 - b_1b_2) + \mathrm{i}(a_1b_2 + b_1a_2)} \\
&= (a_1a_2 - b_1b_2) - \mathrm{i}(a_1b_2 + b_1a_2) \\
&= (a_1a_2 - b_1b_2) + \mathrm{i}(-a_1b_2 - b_1a_2).
\end{aligned}$$

Und nun die rechte Seite:

$$\begin{aligned}
\kappa(z_1) \cdot \kappa(z_2) &= \overline{a_1 + \mathrm{i}b_1} \cdot \overline{a_2 + \mathrm{i}b_2} \\
&= (a_1 - \mathrm{i}b_1) \cdot (a_2 - \mathrm{i}b_2) \\
&= a_1a_2 - \mathrm{i}a_1b_2 - \mathrm{i}b_1a_2 - \mathrm{i}^2 b_1b_2 \\
&= (a_1a_2 - b_1b_2) + \mathrm{i}(-a_1b_2 - b_1a_2).
\end{aligned}$$

Dies beendet den Beweis, dass die Konjugationsabbildung ein Körperautomorphismus ist.

(d)

Nun zeigen wir, dass die komplexe Betragsfunktion multiplikativ ist: Gegeben seien $z, w \in \mathbb{C}$. Dann gilt

$$|zw|^2 = (zw)\overline{(zw)} = z\ w\ \overline{z}\ \overline{w} = (z\overline{z})(w\overline{w}) = |z|^2\,|w|^2 = (|z|\,|w|)^2.$$

Da sowohl $|zw|$ als auch $|z|\,|w|$ reell und nichtnegativ sind, wissen wir, dass Quadrieren bijektiv ist und somit gilt:

$$|zw| = |z|\,|w|\,. \qquad \square$$

Es gibt noch einen weiteren wichtigen Gruppenhomomorphismus, der immer auftaucht, wenn man mit den komplexen Zahlen arbeitet. Für die konkrete Konstruktion, sowie für den Nachweis der Gruppenhomomorphismuseigenschaft verweisen wir auf die Analysis- bzw. HM1-Vorlesung:

Notation 3.5.8. Die *komplexe Exponentialfunktion*

$$\exp : \mathbb{C} \to \mathbb{C}^\times, \quad z \mapsto e^z := \sum_{k=0}^{\infty} \frac{z^k}{k!}$$

ist ein surjektiver Gruppenhomomorphismus von der additiven Gruppe $(\mathbb{C}, +)$ in die multiplikative Gruppe $(\mathbb{C}^\times, \cdot)$. Konkret bedeutet dies, dass für $z, w \in \mathbb{C}$ gilt:

$$\exp(z + w) = \exp(z) \cdot \exp(w),$$

was sich auch suggestiv als $e^{z+w} = e^z \cdot e^w$ schreiben lässt.[15]

Es gilt $|\exp(z)| = \exp(\mathrm{Re}\, z)$ und insbesondere $|\exp(\mathrm{i}t)| = \exp(\mathrm{Re}(\mathrm{i}t)) = \exp(0) = 1$, also ist $\exp(\mathrm{i}t) \in S^1$.

[15] Man sollte sich aber klarmachen, dass diese Schreibweise e^z *keine* Potenz im Sinne von hintereinander durchgeführter Multiplikation (Notation 3.1.12) ist, da wir hier Exponenten erlauben, die keine ganzen Zahlen sein müssen. Aufpassen muss man auch, dass manche Potenzgesetze hier nicht mehr gelten, beispielsweise ist $(e^z)^w = e^{zw}$ eventuell nicht korrekt, bzw. eventuell nicht mal definiert.

Wenn wir also nur imaginäre Zahlen in exp einsetzen, erhalten wir nur Zahlen auf dem Einheitskreis. Dies ergibt einen neuen Gruppenhomomorphismus:

$$\varphi : (\mathbb{R}, +) \to (S^1, \cdot) \subseteq (\mathbb{C}^\times, \cdot), \quad t \mapsto e^{\mathrm{i}t.}$$

Für $t > 0$ kann man die Zahl $\varphi(t) = e^{\mathrm{i}t}$ geometrisch interpretieren als den Punkt auf dem Kreis, den man erhält, wenn man am Punkt 1 auf der reellen Achse anfängt und dann gegen den Uhrzeigersinn solange auf der Kreislinie läuft, bis man den Weg t zurückgelegt hat (oder äquivalent: den Winkel t im Bogenmaß). Für $t < 0$ geht man entsprechend *im* Uhrzeigersinn. Die *Eulersche Formel*[16] besagt[17]:

$$e^{\mathrm{i}t} = \cos(t) + \mathrm{i}\sin(t).$$

Der Homomorphismus φ ist surjektiv, was heißt, dass jede komplexe Zahl z mit Betrag 1 von der Form $e^{\mathrm{i}t}$ für ein $t \in \mathbb{R}$ ist.

Der Homomorphismus φ ist aber nicht injektiv, weil z.B. $\varphi(2\pi) = \varphi(4\pi) = \varphi(0) = 1$. Allgemeiner gilt:

$$\ker \varphi = 2\pi\mathbb{Z} = \{2\pi k \mid k \in \mathbb{Z}\}.$$

Bemerkung 3.5.9 (Einheitswurzeln in $\mathbb{C}$)**.** Es sei $m \in \mathbb{N}$.

- Die Menge $C_m := \{z \in \mathbb{C} \mid z^m = 1\}$ ist eine Untergruppe von $(S^1, \cdot) \subseteq (\mathbb{C}^\times, \cdot)$ mit m Elementen. Die Elemente heißen m-te *komplexe Einheitswurzeln*. Geometrisch liegen die Elemente von C_m auf den Ecken eines regelmäßigen m-Ecks (für $m \geq 3$).
- Die Gruppe $(C_m, \cdot)$ ist isomorph zur additiven Gruppe $(\mathbb{Z}/m\mathbb{Z}, +)$ mittels des folgenden Gruppenisomorphismus:

$$(\mathbb{Z}/m\mathbb{Z}, +) \to C_n, \quad [k]_m \mapsto \exp\left(\mathrm{i}k\frac{2\pi}{m}\right).$$

Wir wissen nun also: $\mathbb{C}$ ist eine Körpererweiterung von $\mathbb{R}$ und in $\mathbb{C}$ hat die Gleichung $x^2 = -1$ eine Lösung. Genau genommen gibt es sogar zwei Lösungen:

$$x^2 = -1 \iff x^2 = \mathrm{i}^2 \iff x^2 - \mathrm{i}^2 = 0 \iff (x - \mathrm{i})(x + \mathrm{i}) = 0.$$

Da $\mathbb{C}$ ein Körper ist und Körper nullteilerfrei sind, gilt: $x - \mathrm{i} = 0$ oder $x + \mathrm{i} = 0$. Es gibt also zwei Lösungen: $x = \mathrm{i}$ oder $x = -\mathrm{i}$.

Dies war relativ viel Aufwand, nur um eine einzige in $\mathbb{R}$ unlösbare Gleichung lösen zu können. Wir haben dann aber gesehen, dass auch viele andere Gleichungen in $\mathbb{C}$ lösbar sind (siehe Proposition 3.5.9). Es stellt sich nun die Frage: Gibt es Gleichungen, die man mit den Grundrechenarten in $\mathbb{C}$ formulieren kann, die man aber in $\mathbb{C}$ nicht lösen kann? Die verblüffende Antwort gibt der folgende berühmte Satz:

Satz 3.5.10 (Fundamentalsatz der Algebra)**.** *Es sei $p(X) = a_0 + a_1X + \cdots + a_kX^k$ ein Polynom mit komplexen Koeffizienten $a_j \in \mathbb{C}$ und $a_k \neq 0$ vom Grad $k \in \mathbb{N}$. Dann gibt es immer eine Lösung $z \in \mathbb{C}$ mit*

$$p(z) = 0.$$

[16] nach LEONHARD EULER, Schweizer Mathematiker, 1707–1783

[17] Je nachdem wie man die Analysis aufzieht, kann man die Eulersche Formel benutzen, um $e^{\mathrm{i}t}$ zu definieren, wenn man Sinus und Kosinus als bekannt voraussetzt oder man man kann Kosinus und Sinus als Real- bzw. Imaginärteil von $e^{\mathrm{i}t}$ definieren.

Ferner ist es sogar möglich, das Polynom in Linearfaktoren zu zerlegen, d.h. es gibt komplexe Zahlen $z_1, \ldots, z_k \in \mathbb{C}$ *mit*

$$p(X) = a_k(X - z_1)(X - z_2)\cdots(X - z_k).$$

Dieser Satz wird normalerweise mit Mitteln der Analysis oder der Topologie bewiesen und ist somit streng genommen kein Satz der Algebra. Auch wenn die Konstruktion von $\mathbb{C}$ aus $\mathbb{R}$ eine rein algebraische ist, so ist die Konstruktion des Körpers der reellen Zahlen selbst nicht algebraisch, d.h. um eine solche Aussage über $\mathbb{C}$ zu verwenden, muss man nichttriviale analytische oder topologische Aussagen von $\mathbb{R}$ verwenden, wie den Zwischenwertsatz oder die Aussage, dass stetige Funktionen auf kompakten Mengen Maxima und Minima anwenden.

Definition 3.5.11 (Algebraisch abgeschlossen)**.** Ein Körper K heiße *algebraisch abgeschlossen*, wenn jedes nichtkonstante Polynom[18] mit Koeffizienten aus K mindestens eine Nullstelle in K hat.

Somit kann man den Fundamentalsatz der Algebra auch kürzer formulieren:
Der Körper $\mathbb{C}$ der komplexen Zahlen ist algebraisch abgeschlossen.
Wir werden uns in Kapitel 5.4 noch ausführlicher mit Polynomen beschäftigen.

Zusammenfassung von Abschnitt 3.5

(1) Der Körper der komplexen Zahlen $\mathbb{C}$ ist eine Körpererweiterung von $\mathbb{R}$. Es gibt ein Element $\mathrm{i} \in \mathbb{C}$ mit $\mathrm{i}^2 = -1$.

(2) Jede komplexe Zahl $z \in \mathbb{C}$ lässt sich eindeutig als $z = a + \mathrm{i}b$ schreiben mit reellen Zahlen a, b, die Realteil und Imaginärteil genannt werden.

(3) Die Gaußsche Zahlenebene ist eine geometrische Darstellung der komplexen Zahlen.

(4) Der Betrag $|z|$ gibt den Abstand von z zur 0 an. Die komplexe Konjugation ist ein Körperautomorphismus von $\mathbb{C}$.

(5) Es gilt der Fundamentalsatz der Algebra: Jedes nicht konstante komplexe Polynom hat eine komplexe Nullstelle.

[18] Siehe Definition 5.4.6 für eine Definition, was ein Polynom eigentlich ist.

4. Vektorräume und lineare Abbildungen

4.1. Grundlegendes zu Vektorräumen und linearen Abbildungen

In Kapitel 2 haben wir untersucht, wie man lineare Gleichungssysteme mit reellen Koeffizienten lösen kann. Wir haben gelernt, dass die Lösungsmenge immer ein affiner Unterraum von $\mathbb{R}^n$ ist und im Fall eines homogenen linearen Gleichungssystems sogar ein Untervektorraum von $\mathbb{R}^n$.

Wie Sie sich leicht überzeugen können, können Sie jedes Vorkommen von $\mathbb{R}$ im Kapitel 2 durch einen beliebigen Körper $\mathbb{K}$ ersetzen und alles bleibt wahr. In allen Beweisen haben wir nur die Körpereigenschaften benutzt[1].

Auch der Gauß-Algorithmus lässt sich über jedem Körper (z.B. über $\mathbb{K} = \mathbb{F}_2, \mathbb{F}_4, \mathbb{Z}/p\mathbb{Z}, \mathbb{Q}, \mathbb{C}$) durchführen – sofern bekannt ist, wie in diesem konkreten Körper Addition, Subtraktion, Multiplikation und Division funktionieren.

Wir können also nun mit Spaltenvektoren im $\mathbb{K}^n$ rechnen und hier Untervektorräume, Dimensionen und Basen studieren. Wir gehen nun aber noch einen Schritt weiter und verallgemeinern das Konzept eines *Vektors* noch weiter.

Wir haben gesehen: Das entscheidende, was man mit Spaltenvektoren tun kann ist, sie zu addieren und mit Skalaren aus dem Grundkörper zu multiplizieren. Dies führt uns nun zu der folgenden Definition:

Definition 4.1.1 (Vektorraum über $\mathbb{K}$)**.** Es sei $\mathbb{K}$ ein Körper. Ein $\mathbb{K}$-*Vektorraum* $(V, +, \cdot)$ ist eine Menge V zusammen mit zwei Operationen

$$+ : V \times V \to V,$$

genannt Addition und

$$\cdot : \mathbb{K} \times V \to V,$$

genannt skalare Multiplikation, sodass die folgenden Eigenschaften erfüllt sind:

(i)		$(V, +)$ ist eine abelsche Gruppe	
(ii)	$\forall \lambda \in \mathbb{K}, v, w \in V:$	$\lambda \cdot (v + w) = \lambda \cdot v + \lambda \cdot w.$	(Skalares Distributivgesetz I)
(iii)	$\forall \lambda, \mu \in \mathbb{K}, v \in V:$	$(\lambda + \mu) \cdot v = \lambda \cdot v + \mu \cdot v.$	(Skalares Distributivgesetz II)
(iv)	$\forall \lambda, \mu \in \mathbb{K}, v \in V:$	$(\lambda \cdot \mu) \cdot v = \lambda \cdot (\mu \cdot v).$	(Skalares Assoziativgesetz)
(v)	$\forall v \in V:$	$1_{\mathbb{K}} \cdot v = v.$	(Wirkung des Einselementes)

Elemente der Menge V heißen dann auch *Vektoren* und Elemente aus $\mathbb{K}$ nennt man auch *Skalare*.

Anstelle $\mathbb{K}$-Vektorraum sagt man auch *Vektorraum über* $\mathbb{K}$.

Außerdem üblich sind die Bezeichnungen: *reeller Vektorraum* (falls $\mathbb{K} = \mathbb{R}$), *komplexer Vektorraum* (falls $\mathbb{K} = \mathbb{C}$) oder *rationaler Vektorraum* (falls $\mathbb{K} = \mathbb{Q}$).

[1] Und nicht z.B. die Ordnungsrelation auf den reellen Zahlen.

Bemerkung 4.1.2. (a) Die Vektorraumaxiome in Definition 4.1.1 sehen ähnlich aus wie die Ringaxiome (Definition 3.3.1). Ein entscheidender Unterschied ist: In einem Ring $(R,+,\cdot)$ ist die Multiplikation eine binäre Verknüpfung *auf* der Menge R.

In einem $\mathbb{K}$-Vektorraum haben wir im Allgemeinen keine Möglichkeit, zwei Vektoren miteinander zu multiplizieren. Alles, was möglich ist, ist einen Vektor mit einem Skalar von außen zu multiplizieren (skalieren). Die skalare Multiplikation ist also keine binäre Verknüpfung auf V. In diesem Sinne ist das „skalare Assoziativgesetz“ auch kein „echtes“ Assoziativgesetz, weil die darin vorkommenden $\cdot$ unterschiedliche Bedeutung haben: $\lambda \cdot \mu$ steht für die Multiplikation im Körper $(\mathbb{K},+,\cdot)$, aber $\mu \cdot v$ steht für die skalare Multiplikation, die Teil der Vektorraumstruktur ist.

(b) Es ist wichtig, dass der Begriff „Vektorraum“ nur sinnvoll ist, wenn man dazusagt (oder es aus dem Kontext klar ist), über welchem Körper man arbeitet. In diesem Sinne unterscheiden sich $\mathbb{K}$-Vektorräume von anderen algebraischen Strukturen, wie Gruppen oder Halbgruppen.

(c) Eine wichtige Sache, die oft falsch verstanden wird, ist die folgende: Die Angabe des Grundkörpers $\mathbb{K}$ sagt *nur* aus, aus welcher Menge die Skalare kommen, mit denen die Vektoren skalar multipliziert werden können. Sie sagt nichts darüber aus, wie die Elemente von V aussehen. Wie bei algebraischen Strukturen allgemein üblich, können die Elemente aus V beliebige Objekte sein. Beispielsweise ist es möglich, auf der Menge $V := \{\text{Hund}, \text{Katze}, \text{Maus}, \text{Telefon}\}$ die Struktur eines $\mathbb{F}_2$-Vektorraums zu definieren. (Wir nehmen hier an, dass Hund, Katze, Maus und Telefon vier unterschiedliche Objekte sind...)

(d) Die Vektorraumaxiome sind – wie auch schon die Ringaxiome in Definition 3.3.1 nicht minimal. Insbesondere folgt die Kommutativität der Addition aus den anderen Axiomen.

Beispiel 4.1.3. Es sei $\mathbb{K}$ ein beliebiger Körper.

(a) Für $m,n \in \mathbb{N}$ ist die Menge aller $(m \times n)$-Matrizen mit Einträgen aus $\mathbb{K}$

$$(\mathbb{K}^{m\times n},+,\cdot)$$

ein Vektorraum über $\mathbb{K}$. Die Definition von Matrixaddition und skalarem Vielfachen

$$+ : \mathbb{K}^{m\times n} \times \mathbb{K}^{m\times n} \to \mathbb{K}^{m\times n} \text{ und } \cdot : \mathbb{K} \times \mathbb{K}^{m\times n} \to \mathbb{K}^{m\times n}$$

aus Definition 2.2.3 überträgt sich direkt vom Fall $\mathbb{K} = \mathbb{R}$ auf den eines allgemeinen Körpers. Lemma 2.2.7 (bzw. die offensichtliche Verallgemeinerung auf allgemeine Körper $\mathbb{K}$) sagt uns, dass $(\mathbb{K}^{m\times n},+,\cdot)$ ein $\mathbb{K}$-Vektorraum ist.

(b) Als Spezialfall folgt, dass der Raum der Spaltenvektoren $\mathbb{K}^m$ für $m \in \mathbb{N}$ ein Vektorraum ist.

(c) Der Körper $\mathbb{K} = \mathbb{K}^1 = \mathbb{K}^{1\times 1}$ selbst ist auch ein Beispiel für einen $\mathbb{K}$-Vektorraum.

(d) Eine einelementige Menge $\{0\}$ ist ein $\mathbb{K}$-Vektorraum (es gibt nur genau eine Möglichkeit, $+$ und $\cdot$ zu definieren). Hierfür werden wir die Notation $\mathbb{K}^0$ benutzen:

$$\mathbb{K}^0 := \{0\}.$$

(e) Es sei J eine beliebige Menge und $f, g : J \to \mathbb{K}$ Funktionen. Dann lassen sich diese Funktionen punktweise addieren und wir erhalten eine neue Funktion :

$$f + g : J \to \mathbb{K}, \quad x \mapsto f(x) + g(x).$$

Ebenso kann man eine Funktion $f : J \to \mathbb{K}$ punktweise mit der Konstanten $\lambda \in \mathbb{K}$ multiplizieren:

$$\lambda f : J \to \mathbb{K}, \quad x \mapsto \lambda \cdot f(x).$$

Wir versehen nun

$$\mathbb{K}^J = \{f \mid f : J \to \mathbb{K}\},$$

die Menge aller $\mathbb{K}$-wertigen Abbildungen mit einer $\mathbb{K}$-Vektorraumstruktur: Die Addition ist die eben definierte punktweise Addition:

$$+ : \mathbb{K}^J \times \mathbb{K}^J \to \mathbb{K}^J, \quad (f, g) \mapsto f + g,$$

die skalare Multiplikation ist die eben definierte Multiplikation mit einer Konstanten:

$$\cdot : \mathbb{K} \times \mathbb{K}^J \to \mathbb{K}^J, \quad (\lambda, f) \mapsto \lambda f.$$

Dass die $\mathbb{K}$-Vektorraumaxiome erfüllt sind, folgt nun direkt aus der Eigenschaft, dass alle Operationen punktweise passieren und die Zielmenge $\mathbb{K}$ ein $\mathbb{K}$-Vektorraum ist.

(f) Im Spezialfall sei $\mathbb{K} = \mathbb{R}$ und $J = \mathbb{R}$. Dann ist der Raum der Funktionen von $\mathbb{R}$ nach $\mathbb{R}$ mit punktweiser Addition und skalarer Multiplikation ein reeller Vektorraum.

(g) Im Spezialfall sei $J = \mathbb{N}$. Dann ist $\mathbb{K}^{\mathbb{N}}$ die Menge aller Folgen in $\mathbb{K}$. Für $\mathbb{K} = \mathbb{R}$ sind das also alle reellen Zahlenfolgen, für $\mathbb{K} = \mathbb{C}$ alle komplexen Zahlenfolgen. Diese Vektorräume (und vor allem ihre Untervektorräume) sind für die Analysis sehr wichtig.

(h) Anstelle von skalarwertigen Funktionen oder Folgen kann man auch vektorwertige Funktionen oder Folgen untersuchen. Es sei dazu J wieder eine beliebige Menge (z.B. $J = \mathbb{N}$ oder $J = \mathbb{R}$) und V ein $\mathbb{K}$-Vektorraum (z.B. $V = \mathbb{K}^3$). Dann lässt sich die Menge aller V-wertigen Abbildungen

$$V^J = \{f \mid f : J \to V\}$$

mit der punktweisen Addition und skalaren Multiplikation zu einem Vektorraum machen. So kann man beispielsweise die Menge aller Folgen mit Werten in $\mathbb{R}^{3\times 4}$ zu einem Vektorraum machen.

(i) Es sei $(\mathbb{L}, +, \cdot)$ ein Körper (denken Sie z.B. an $\mathbb{L} = \mathbb{C}$). Dann ist die Multiplikation auf $\mathbb{L}$ bekanntermaßen eine binäre Verknüpfung

$$\cdot : \mathbb{L} \times \mathbb{L} \to \mathbb{L}, \quad (z, w) \mapsto zw.$$

Es sei nun $\mathbb{K}$ ein Unterkörper von $\mathbb{L}$ (oder was das gleiche ist: $\mathbb{L}$ eine Körpererweiterung von $\mathbb{K}$). Dann können wir die Körpermultiplikation von $\mathbb{L} \times \mathbb{L}$ auf die Teilmenge $\mathbb{K} \times \mathbb{L}$ einschränken und erhalten

$$\cdot|_{\mathbb{K}\times\mathbb{L}} : \mathbb{K} \times \mathbb{L} \to \mathbb{L}, \quad (r, w) \mapsto rw.$$

Die Menge $\mathbb{L}$ wird mit der gewöhnlichen Addition und dieser eingeschränkten Multiplikation zu einem Vektorraum über $\mathbb{K}$.

Insbesondere kann also $\mathbb{C}$ als Vektorraum über $\mathbb{R}$ angesehen werden (oder $\mathbb{R}$ als Vektorraum über $\mathbb{Q}$ oder $\mathbb{F}_4$ als Vektorraum über $\mathbb{F}_2$).

Der Begriff des Vektorraums über einem Körper $\mathbb{K}$ ist gerade so gewählt, dass man nur die Strukturen (Addition von Vektoren und skalare Vielfache von Vektoren) hat, die man braucht, um die entsprechenden Begriffsbildungen (wie Dimension, Kern, Basis, affiner Unterraum) definieren zu können. Begriffe, die man im $\mathbb{R}^n$ auch hätte einführen können, wie der Abstand zwischen zwei Punkten oder der Winkel zwischen Vektoren lassen sich in dieser Allgemeinheit nicht definieren und sind deshalb nicht Teil der Definition eines Vektorraums.

Notation 4.1.4. Ist $\mathbb{K}$ ein Körper und V ein $\mathbb{K}$-Vektorraum, so nennen wir das neutrale Element der additiven Gruppe $(V,+)$ den *Nullvektor* und schreiben dafür 0_V oder einfach 0. Das additive Inverse von v schreiben wir als $-v$ und die *Differenz* zweier Vektoren $v,w \in V$ durch

$$v - w := v + (-w).$$

Beim Rechnen mit Skalaren und Vektoren können wir wegen der Assoziativgesetze in vielen Ausdrücken auf Klammern verzichten. Für $u,v,w \in V$ schreiben wir z.B. einfach $u+v+w$ statt $(u+v)+w$ oder $u+(v+w)$. Darüber hinaus machen wir die übliche „Punkt- vor Strichrechnung"-Konvention, d.h. für $\lambda \in \mathbb{K}$ und $u,v \in V$ schreiben wir $\lambda u + v$ statt $(\lambda u)+v$.

Lemma 4.1.5. *Seien $\mathbb{K}$ ein Körper, $\lambda \in \mathbb{K}$, V ein $\mathbb{K}$-Vektorraum und $v \in V$.*

(i) $\lambda v = 0_V \iff (\lambda = 0_\mathbb{K})$ *oder* $(v = 0_V)$.

(ii) $(-1_\mathbb{K})v = -v$.

Beweis. (i) „$\Longleftarrow$": Für $v \in V$ gilt

$$0_\mathbb{K} v = (0_\mathbb{K} + 0_\mathbb{K})v = 0_\mathbb{K} v + 0_\mathbb{K} v \Longrightarrow 0_V = 0_\mathbb{K} v - 0_\mathbb{K} v = (0_\mathbb{K} v + 0_\mathbb{K} v) - 0_\mathbb{K} v = 0_\mathbb{K} v.$$

Analog gilt für $\lambda \in \mathbb{K}$, dass

$$\lambda 0_V = \lambda(0_V + 0_V) = \lambda 0_V + \lambda 0_V \Longrightarrow 0_V = \lambda 0_V - \lambda 0_V = (\lambda 0_V + \lambda 0_V) - \lambda 0_V = \lambda 0_V.$$

„$\Longrightarrow$": Es sei $\lambda v = 0_V$, aber $\lambda \neq 0_\mathbb{K}$. Dann hat λ ein multiplikatives Inverses λ^{-1} und wegen „$\Longleftarrow$" gilt

$$0_V = \lambda^{-1_\mathbb{K}} 0_V = \lambda^{-1_\mathbb{K}}(\lambda v) = (\lambda^{-1_\mathbb{K}} \lambda)v = 1_\mathbb{K} v = v.$$

(ii) Mit (i) erhalten wir

$$(v + (-1_\mathbb{K})v) = 1_\mathbb{K} v + (-1_\mathbb{K})v = (1_\mathbb{K} + (-1_\mathbb{K}))v = 0_\mathbb{K} v = 0_V. \qquad \square$$

Wie bei Gruppen und Ringen auch definieren wir nun Unterstrukturen:

Definition 4.1.6 (Untervektorraum)**.**
Es sei V ein Vektorraum über einem Körper $\mathbb{K}$. Eine Teilmenge $U \subseteq V$ heißt $\mathbb{K}$-*Untervektorraum* oder $\mathbb{K}$-*linearer Unterraum* von V, wenn folgende drei Bedingungen erfüllt sind:

(i)		$0_V \in U$	(U enthält den Nullvektor)
(ii)	$\forall v,w \in U:$	$v+w \in U$	(U ist abgeschlossen unter Addition)
(iii)	$\forall v \in U, \lambda \in \mathbb{K}:$	$\lambda v \in U$	(U ist abgeschlossen unter Multiplikation mit Skalaren aus $\mathbb{K}$)

Im Falle $\mathbb{K} = \mathbb{R}$ und $V = \mathbb{R}^n$ erhalten wir exakt Definition 2.3.1.

Wir haben gesehen, dass Untergruppe von Gruppen wieder Gruppen sind und dass Unterringe von Ringen wieder Ringe sind. Wenig überraschend ist also die folgende Aussage:

Lemma 4.1.7 (Untervektorräume als Vektorräume)**.** *Ist U ein $\mathbb{K}$-Untervektorraum eines $\mathbb{K}$-Vektorraums V, so wird $(U,+,\cdot)$ selbst ein $\mathbb{K}$Vektorraum, wenn wir $+$ und $\cdot$ entsprechend einschränken.*

Beweis. Es sei ein Untervektorraum U von einem $\mathbb{K}$-Vektorraum V gegeben. Zuerst behaupten wir, dass U eine Untergruppe von $(V,+)$ ist (siehe Definition 3.2.5). Das neutrale Element von $(V,+)$ ist der Nullvektor und dieser ist nach (i) in U enthalten. Als nächstes müssen wir überprüfen, dass U abgeschlossen unter der Gruppenoperation (in diesem Falle Addition) ist. Dies ist aber genau Bedingung (ii). Als drittes brauchen wir für eine Untergruppe, dass für jedes Element $v \in U$ auch das Inverse bezüglich $+$ in U enthalten ist. Nach Lemma 4.1.5 ist das Inverse $-x$ aber gleich $(-1)x$ und somit nach Bedingung (iii) in U enthalten.

Also ist U eine Untergruppe von $(V,+)$ und somit insbesondere eine Gruppe. Da V ein Vektorraum ist, ist die Addition kommutativ und somit ist $(V,+)$ eine abelsche Gruppe.

Wegen (iii) können wir

$$\mathbb{K} \times W \to W, \quad (\lambda, x) \mapsto \lambda x$$

definieren. Die Eigenschaften (ii) bis (v) aus Definition 4.1.1 folgen direkt aus denen von V. □

Beispiel 4.1.8. (a) Es seien $m,n \in \mathbb{N}$ natürliche Zahlen, und $A \in \mathbb{K}^{m\times n}$ eine $(m \times n)$-Matrix mit Einträgen aus einem Körper $\mathbb{K}$. Dann ist die Lösungsmenge des *homogenen* linearen Gleichungssystems:

$$\{x \in \mathbb{K}^n \mid Ax = 0\}$$

ein Untervektorraum von $\mathbb{K}^n$. Alle Argumente aus dem reellen Fall übertragen sich wortwörtlich.

Später[2] werden wir sehen, dass jeder Untervektorraum von $\mathbb{K}^n$ auf diese Weise entsteht.

(b) Es sei $n \in \mathbb{N}$ und $\mathbb{K}$ ein Körper. Eine Matrix $A \in \mathbb{K}^{n\times n}$ heiße *symmetrisch*, wenn $A^\top = A$. Die Menge aller symmetrischen Matrizen ist ein Untervektorraum des $\mathbb{K}$-Vektorraums $\mathbb{K}^{n\times n}$. Dies ist leicht zu überprüfen mit Hilfe von Lemma 2.2.13, das genauso für jeden beliebigen Körper $\mathbb{K}$ gilt.

(c) Es sei $\mathbb{L}$ eine Körpererweiterung von $\mathbb{K}$ (z.B. $\mathbb{K} = \mathbb{R}$ und $\mathbb{L} = \mathbb{C}$ oder $\mathbb{K} = \mathbb{Q}$ und $\mathbb{L} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{F}_2$ und $\mathbb{L} = \mathbb{F}_4$). Dann ist $\mathbb{L}$ ein Vektorraum über $\mathbb{K}$ (siehe Beispiel 4.1.3(i)). Die Menge $\mathbb{K}$ ist dann ein Untervektorraum von $\mathbb{L}$.

(d) Die Menge der imaginären Zahlen $i\mathbb{R} \subseteq \mathbb{C}$ ist ein reeller Untervektorraum von $\mathbb{C}$.

(e) Es sei $[a,b] \subseteq \mathbb{R}$ ein Intervall $(a,b \in \mathbb{R}; a < b)$. Der Raum $\mathbb{R}^{[a,b]}$ aller Funktionen von $[a,b]$ nach $\mathbb{R}$ ist ein $\mathbb{R}$-Vektorraum. Die Menge der *stetigen* Funktionen $C([a,b],\mathbb{R})$ ist ein Untervektorraum. Dasselbe gilt für die Menge der differenzierbaren Funktionen oder die Menge der Riemann-integrierbaren Funktionen von $[a,b]$ nach $\mathbb{R}$.

Diese Beispiele lassen sich entsprechend verallgemeinern auf offene Intervalle als Definitionsbereich oder auch mehrdimensionale (oder gekrümmte) Definitionsbereiche. Beispielsweise ist die Menge der stetigen reellwertigen Funktionen, die auf der Kugeloberfläche definiert sind, ein $\mathbb{R}$-Vektorraum.

[2] siehe Bemerkung 4.5.3

(f) Die Menge aller Folgen $\mathbb{R}^{\mathbb{N}}$ ist ein $\mathbb{R}$-Vektorraum. Die Menge $c(\mathbb{N},\mathbb{R}) \subseteq \mathbb{R}^{\mathbb{N}}$ aller *konvergenten* Folgen ist ein Untervektorraum. Die Menge $c_0(\mathbb{N},\mathbb{R})$ aller gegen 0 konvergierenden Folgen ist wiederum ein Untervektorraum von c.

Auch die Menge aller *beschränkten* Folgen (meistens mit $\ell^\infty(\mathbb{N},\mathbb{R})$ bezeichnet) ist ein Untervektorraum von $\mathbb{R}^{\mathbb{N}}$.

Das gleiche gilt für die Menge aller Folgen $(x_n)_{n\in\mathbb{N}}$ mit der Eigenschaft, dass die Reihe $\sum_{n=1}^{\infty} |x_n|$ konvergiert. Für diesen Raum ist die Bezeichnung $\ell^1(\mathbb{N},\mathbb{R})$ gebräuchlich. Es gelten die Untervektorrauminklusionen:

$$\ell^1(\mathbb{N},\mathbb{R}) \subseteq c_0(\mathbb{N},\mathbb{R}) \subseteq c(\mathbb{N},\mathbb{R}) \subseteq \ell^\infty(\mathbb{N},\mathbb{R}) \subseteq \mathbb{R}^{\mathbb{N}}$$

Solche Folgenräume spielen in der Funktionalanalysis eine große Rolle und lassen sich auch mit Werten in $\mathbb{C}$ statt mit Werten in $\mathbb{R}$ definieren. Die Vektorräume werden dann entsprechend $\mathbb{C}$-Vektorräume.

Diese Konzepte auf andere Körper als $\mathbb{R}$ oder $\mathbb{C}$ zu verallgemeinern wird allerdings schwierig, weil sich Begriffe wie Beschränktheit oder Konvergenz von Folgen nicht für beliebige Körper definieren lassen, sondern von der reellen (oder der komplexen) Betragsfunktion abhängen.

(g) Es sei J eine Menge und $\mathbb{K}$ ein Körper. Wir setzen

$$\mathbb{K}^{(J)} := \{f : J \to K \mid f \text{ ist nur an endlichen vielen Stellen ungleich } 0\} \subseteq \mathbb{K}^J.$$

Auch diese Menge ist ein Untervektorraum von $\mathbb{K}^J$. Falls J endlich ist, ist dies allerdings der ganze Raum $\mathbb{K}^J$.

Nachdem wir nun das Konzept eines Untervektorraums von $\mathbb{R}^n$ auf Untervektorräume von beliebigen $\mathbb{K}$-Vektorräumen verallgemeinert haben, tun wir das gleiche mit affinen Unterräumen:

Definition 4.1.9 (Affiner Unterraum)**.**
Es sei $\mathbb{K}$ ein Körper. Eine Teilmenge $R \subseteq V$ eines $\mathbb{K}$-Vektorraums V heiße *affiner Unterraum* (oder $\mathbb{K}$-affiner Unterraum) von V, wenn ein $p \in V$ und ein Untervektorraum $U \subseteq V$ existiert, sodass

$$R = p + U.$$

Den Vektor p in dieser Darstellung nennt man auch *Fußpunkt*.

Dies ist die natürliche, direkte Verallgemeinerung von Definition 2.4.2.

Bemerkung 4.1.10.
1. Wie schon bei affinen Unterräumen von $\mathbb{R}^n$ gilt auch hier, dass der Fußpunkt im Allgemeinen nicht eindeutig ist, der Untervektorraum U aber schon.
2. Jede einelementige Teilmenge $R := \{p\}$ ist ein affiner Unterraum.
3. Jeder Untervektorraum $U \subseteq V$ eines $\mathbb{K}$-Vektorraums V ist ein affiner Unterraum.
4. Wie im $\mathbb{R}^n$ lassen sich Affinkombinationen definieren.

Alle Argumentationen aus Kapitel 2.4 lassen sich wörtlich übertragen – mit Ausnahme von Bemerkung 2.4.9, die nicht mehr über jedem Grundkörper gilt. Sie sind herzlich eingeladen herauszufinden, an welcher Stelle genau die Argumentation zusammenbricht.[3]

[3] Sie müssen dafür natürlich zuerst das Konzept einer *Gerade* durch zwei Punkte sinngemäß in diesen allgemeinen Rahmen übertragen.

Der Beweis des folgenden Satzes überträgt sich auch wortwörtlich aus Satz 2.4.8:

Satz 4.1.11 (Charakterisierung affiner Unterräume)**.**
Es sei V ein Vektorraum über einem Körper $\mathbb{K}$. Für eine Teilmenge $R \subseteq V$ sind die folgenden Aussagen äquivalent:

(i) R *ist ein affiner Unterraum von* V.

(ii) $R \neq \emptyset$ *und* $\forall x, y, z \in R, \lambda \in \mathbb{K} : x + \lambda(y - z) \in R$.

(iii) $R \neq \emptyset$ *und jede Affinkombination von endlich vielen Vektoren in* R *ist in* R.

Beispiel 4.1.12. (a) Es seien $m, n \in \mathbb{N}$ natürliche Zahlen, und $A \in \mathbb{K}^{m \times n}$ eine $(m \times n)$-Matrix mit Einträgen aus einem Körper $\mathbb{K}$ und $b \in \mathbb{K}^m$. Dann ist die Lösungsmenge des *inhomogenen* linearen Gleichungssystems:

$$\{x \in \mathbb{K}^n \mid Ax = b\}$$

entweder leer oder ein affiner Unterraum von $\mathbb{K}^n$. Alle Argumente aus dem reellen Fall übertragen sich wortwörtlich.

(b) Es sei $[a, b] \subseteq \mathbb{R}$ ein Intervall $(a, b \in \mathbb{R}; a < b)$ und $f : [a, b] \to \mathbb{R}$ eine stetige Funktion. Nach Hauptsatz der Differential- und Integralrechnung gibt es eine Stammfunktion, also eine differenzierbare Funktion $F : [a, b] \to \mathbb{R}$ mit $F' = f$. Diese Funktion ist aber nicht eindeutig. Die Menge aller Stammfunktionen einer gegebenen Funktion f ist affiner Unterraum vom Vektorraum aller differenzierbarer Funktionen, der wiederum ein Untervektorraum vom Raum aller Funktionen $\mathbb{R}^{[a,b]}$ ist.

(c) Es sei $a \in \mathbb{R}$ eine reelle Zahl. Die Menge aller Folgen, die gegen a konvergieren ist ein affiner Unterraum vom Raum aller konvergenter Folgen, der wiederum ein Untervektorraum vom Raum aller Folgen $\mathbb{R}^{\mathbb{N}}$ ist.

Wie bei Gruppen und Ringen wollen wir nun strukturerhaltende Abbildungen zwischen $\mathbb{K}$-Vektorräumen betrachten, also Homomorphismen von $\mathbb{K}$-Vektorräumen.

Definition 4.1.13 (Lineare Abbildungen)**.**
Es sei $\mathbb{K}$ ein Körper.

(a) Gegeben seien Vektorräume V und W. Dann heißt eine Abbildung

$$\varphi : V \to W$$

Homomorphismus von $\mathbb{K}$*-Vektorräumen* (oder $\mathbb{K}$*-Vektorraumhomomorphismus*) oder einfach $\mathbb{K}$*-lineare Abbildung*, wenn

$$\forall x, y \in V : \varphi(x + y) = \varphi(x) + \varphi(y)$$

und

$$\forall x \in V, \lambda \in \mathbb{K} : \varphi(\lambda x) = \lambda \varphi(x).$$

Hieraus folgt dann bereits automatisch (siehe Lemma 4.1.14), dass

$$\varphi(0) = 0$$

gilt.

(b) Eine $\mathbb{K}$-lineare Abbildung $\varphi : V \to W$ zwischen zwei $\mathbb{K}$-Vektorräumen V und W heißt *Isomorphismus von Vektorräumen* (oder $\mathbb{K}$-*Vektorraumisomorphismus*), wenn φ bijektiv ist. In diesem Fall ist auch $\varphi^{-1} : W \to V$ ein Isomorphismus von Vektorräumen. Zwei $\mathbb{K}$-Vektorräume V und W heißen *isomorph*, wenn es einen Isomorphismus von $\mathbb{K}$-Vektorräumen zwischen ihnen gibt. Wenn V und W als $\mathbb{K}$-Vektorräume isomorph sind, schreibt man auch $V \cong_{\mathbb{K}} W$ oder einfach $V \cong W$, wenn der Grundkörper aus dem Zusammenhang klar ist.

(c) Eine $\mathbb{K}$-lineare Abbildung $\varphi : V \to V$ von einem $\mathbb{K}$-Vektorraum in sich selbst nennt man auch einen *Endomorphismus von Vektorräumen* (oder *Vektorraumendomorphismus*). Ein Endomorphismus, der gleichzeitig ein Isomorphismus ist, ist ein *Vektorraumautomorphismus*.

Wie bei Gruppen beweist man das folgende Lemma:

Lemma 4.1.14.
Es seien V und W Vektorräume über demselben Körper $\mathbb{K}$ und $\varphi : V \to W$ eine $\mathbb{K}$-lineare Abbildung.

(a) Dann gilt
$$\varphi(0_V) = 0_W.$$

(b) Es sei $U \subseteq V$ ein Untervektorraum von V. Dann ist $\varphi(U)$ ein Untervektorraum von W. Insbesondere ist Bild$(\varphi) = \varphi(V)$ *eine Untervektorraum von W.*

(c) Es sei $U \subseteq W$ ein Untervektorraum W. Dann ist $\varphi^{-1}(U)$ ein Untervektorraum von V. Insbesondere ist der Kern $\ker(\varphi) := \varphi^{-1}(\{0_W\})$ *ein Untervektorraum von V.*

Beweis. (a)
Jede lineare Abbildung ist insbesondere ein Gruppenhomomorphismus zwischen den additiven Gruppen. Somit folgt die Aussage direkt aus Lemma 3.2.8.

(b)
Da jeder Untervektorraum auch eine additive Untergruppe ist, folgt, dass $\varphi(U)$ eine additive Untergruppe von $(W, +)$ ist aus Lemma 3.2.8. Es bleibt zu zeigen, dass $\varphi(U)$ abgeschlossen ist unter Multiplikation mit Skalaren aus $\mathbb{K}$. Sei dazu $y \in \varphi(U)$ und $\lambda \in \mathbb{K}$. Es ist zu zeigen, dass $\lambda y \in \varphi(U)$ gilt.

Aus $y \in \varphi(U)$ folgt, dass es ein $x \in U$ gibt mit $\varphi(x) = y$. Da U als Untervektorraum von V vorausgesetzt wurde, gilt, dass $\lambda x \in U$. Nun gilt:
$$\lambda y = \lambda\varphi(x) = \varphi(\lambda x) \in \varphi(U).$$

(c)
Da jeder Untervektorraum auch eine additive Untergruppe ist, folgt, dass $\varphi^{-1}(U)$ eine additive Untergruppe von V ist aus Lemma 3.2.8. Es bleibt zu zeigen, dass $\varphi^{-1}(U)$ abgeschlossen ist unter Multiplikation mit Skalaren aus $\mathbb{K}$. Sei dazu $x \in \varphi^{-1}(U)$ und $\lambda \in \mathbb{K}$. Es ist zu zeigen, dass $\lambda x \in \varphi^{-1}(U)$ gilt.

Aus $x \in \varphi^{-1}(U)$ folgt, dass $\varphi(x) \in U$. Da U als Untervektorraum von W vorausgesetzt wurde, gilt, dass $\lambda\varphi(x) \in U$. Aus der $\mathbb{K}$-Linearität von φ folgt nun, dass $\varphi(\lambda x) = \lambda\varphi(x) \in U$. Also ist $\lambda x \in \varphi^{-1}(U)$. □

Durch Anwenden von Lemma 3.2.10 auf die additiven Gruppen der Vektorräume erhält man direkt:

Lemma 4.1.15.
Es sei $\mathbb{K}$ ein Körper. Eine $\mathbb{K}$-lineare Abbildung $\varphi : V \to W$ zwischen $\mathbb{K}$-Vektorräumen V und W ist genau dann injektiv, wenn $\ker \varphi = \{0_V\}$.

Beispiel 4.1.16. Es sei $\mathbb{K}$ ein Körper.

(a) Es seien $m, n \in \mathbb{N}$ und $A \in \mathbb{K}^{m \times n}$ gegeben. Dann ist die Abbildung

$$\Phi_A : \mathbb{K}^n \to \mathbb{K}^m, \quad x \mapsto Ax$$

$\mathbb{K}$-linear, also ein Vektorraumhomomorphismus zwischen von $\mathbb{K}$-Vektorräumen $\mathbb{K}^n$ und $\mathbb{K}^m$ (Dies folgt direkt aus Lemma 2.2.12[4]). Das Bild von Φ_A ist genau das Bild der Matrix A, wie definiert in Definition 2.4.11. Der Kern von Φ_A ist genau der Kern der Matrix A, wie definiert auf Seite 29.

Falls $m = n$, so ist Φ_A ein Vektorraumendomorphismus.

(b) Allgemeiner: Es seien $m, n, p \in \mathbb{N}$ und $A \in \mathbb{K}^{m \times n}$ und $B \in \mathbb{K}^{n \times p}$ Matrizen. Dann sind die Abbildungen

$$\mathbb{K}^{n \times p} \to \mathbb{K}^{m \times p}, \quad X \mapsto AX$$

und

$$\mathbb{K}^{m \times n} \to \mathbb{K}^{m \times p}, \quad X \mapsto XB$$

$\mathbb{K}$-linear.

(c) Die Abbildungen $\mathrm{Re} : \mathbb{C} \to \mathbb{R}$, $\mathrm{Im} : \mathbb{C} \to \mathbb{R}$ und $\mathbb{C} \to \mathbb{C}, \quad z \mapsto \overline{z}$ sind alle $\mathbb{R}$-linear, aber nicht $\mathbb{C}$-linear.

(d) Die Abbildung $\mathbb{K}^{m \times n} \to \mathbb{K}^{m \times n}, \quad A \mapsto A^\top$ ist ein Isomorphismus von $\mathbb{K}$-Vektorräumen. Insbesondere ist also der Raum der Zeilenvektoren $\mathbb{K}^{1 \times n}$ isomorph zum Raum der Spaltenvektoren $\mathbb{K}^n$. Es gilt also: $\mathbb{K}^{1 \times n} \cong_{\mathbb{K}} \mathbb{K}^n$.

(e) Die Abbildungen $\varphi : \mathbb{R} \to \mathbb{R}, \quad x \mapsto x^2$ und $\psi : \mathbb{R} \to \mathbb{R}, \quad x \mapsto x + 1$ sind nicht $\mathbb{R}$-linear.

(f) Eine Abbildung $\varphi : \mathbb{R} \to \mathbb{R}$ ist genau dann $\mathbb{R}$-linear, wenn $\varphi(x) = ax$ für eine Konstante $a \in \mathbb{R}$.

(g) Es sei $\varphi : \mathbb{R}^2 \to \mathbb{R}^2$ die Drehung mit Drehmittelpunkt 0 und Winkel $\alpha \in \mathbb{R}$ gegen den Uhrzeigersinn. Dann gilt für alle $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2$:

$$\varphi\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) = \begin{pmatrix} \cos(\alpha)x_1 - \sin(\alpha)x_2 \\ \sin(\alpha)x_1 + \cos(\alpha)x_2 \end{pmatrix} = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Nach Beispiel (a) ist diese Abbildung also $\mathbb{R}$-linear.

Da der Definitionsbereich von φ gleich dem Zielbereich ist, ist φ ein $\mathbb{R}$-Vektorraumendomorphismus und da eine Rotation um α immer durch eine Rotation um den Winkel $-\alpha$ rückgängig gemacht werden kann, ist φ sogar bijektiv und somit ein Automorphismus des reellen Vektorraums $\mathbb{R}^2$.

[4] verallgemeinert auf beliebige Körper $\mathbb{K}$ statt nur $\mathbb{R}$

(h) Es sei $G = \{(0)\,0t \mid t \in \mathbb{R}\} \subseteq \mathbb{R}^3$ die x_3-Achse im Raum $\mathbb{R}^3$ und $\alpha \in \mathbb{R}$. Eine Drehung $\varphi : \mathbb{R}^3 \to \mathbb{R}^3$ mit Drehachse G und Winkel $\alpha \in \mathbb{R}$ ist gegeben durch

$$\varphi\left(\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}\right) = \begin{pmatrix} \cos(\alpha)x_1 - \sin(\alpha)x_2 \\ \sin(\alpha)x_1 + \cos(\alpha)x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) & 0 \\ \sin(\alpha) & \cos(\alpha) & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

Nach Beispiel (a) ist dies also eine $\mathbb{R}$-lineare Abbildung. Wie in Beispiel (e) sieht man, dass es sich hierbei um einen Automorphismus des $\mathbb{R}$-Vektorraums $\mathbb{R}^3$ handelt.

(i) Da $\mathbb{C}$ eine Körpererweiterung von $\mathbb{R}$ ist, kann man $\mathbb{C}$ als $\mathbb{R}$-Vektorraum auffassen. Als solcher ist $\mathbb{C}$ zu $\mathbb{R}^2$ isomorph mittels des $\mathbb{K}$-Vektorraumisomorphismus

$$\mathbb{R}^2 \to \mathbb{C}, \quad \begin{pmatrix} a \\ b \end{pmatrix} \mapsto a + ib.$$

Ebenso ist $\mathbb{F}_4$ eine Körpererweiterung von $\mathbb{F}_2$ (siehe Beispiel 3.3.14). Als $\mathbb{F}_2$-Vektorraum ist $\mathbb{F}_4$ isomorph zum Raum $\mathbb{F}_2^2 = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \,\middle|\, x, y \in \mathbb{F}_2 \right\}$ über den $\mathbb{F}_2$-Vektorraumisomorphismus:

$$\begin{aligned} \mathbb{F}_4 &\to \mathbb{F}_2^2 \\ 0 &\mapsto \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\ 1 &\mapsto \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ A &\mapsto \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ B &\mapsto \begin{pmatrix} 1 \\ 1 \end{pmatrix}. \end{aligned}$$

(j) Es sei $c(\mathbb{N}, \mathbb{R})$ der $\mathbb{R}$-Vektorraum aller konvergenter Folgen in $\mathbb{R}$. Dann ist Abbildung

$$\lim : c(\mathbb{N}, \mathbb{R}) \to \mathbb{R}, \quad (x_n)_{n \in \mathbb{N}} \mapsto \lim_{n \to \infty} x_n,$$

die jeder konvergenter Folge ihren Grenzwert zuweist $\mathbb{R}$-linear. Der Kern dieser Abbildung ist genau der Raum $c_0(\mathbb{N}, \mathbb{R})$ der Nullfolgen.

(k) Es sei $C([a,b], \mathbb{R})$ der Raum aller stetigen Funktionen von $[a,b]$ nach $\mathbb{R}$ (hierbei sind $a, b \in \mathbb{R}$ mit $a < b$). Dann ist die Abbildung

$$\int : C([a,b], \mathbb{R}) \to \mathbb{R}, \quad f \mapsto \int_a^b f(t)dt,$$

die jeder stetigen Funktion ihr Integral zuweist $\mathbb{R}$-linear.

(l) Es sei $C^1([a,b], \mathbb{R})$ der Raum der stetig differenzierbaren[5] Funktionen. Dann ist der Ableitungsoperator[6]

$$\frac{d}{dt} : C^1([a,b], \mathbb{R}) \to C([a,b], \mathbb{R}), \quad f \mapsto f'$$

[5] Eine Funktion f ist stetig differenzierbar, wenn sie differenzierbar ist und ihre Ableitung stetig ist.

[6] das Wort *Operator* bedeutet nichts anderes als *Abbildung*, klingt aber komplizierter und wird (vermutlich aber nicht deshalb) in der Funktionalanalysis oft verwendet, für Abbildungen, die Funktionen auf Funktionen abbilden

eine $\mathbb{R}$-lineare Abbildung. Die Abbildung ist surjektiv (das besagt der Hauptsatz der Differential- und Integralrechnung), aber nicht injektiv. Der Kern besteht genau aus den konstanten Funktionen.

Proposition 4.1.17.
Es sei $\mathbb{K}$ ein Körper und V und W seien Vektorräume über $\mathbb{K}$. Dann ist die Menge aller $\mathbb{K}$-linearen Abbildungen von V nach W ein Untervektorraum des Raumes W^V aller Abbildungen von V nach W (siehe Beispiel 4.1.3 (h))

Den Raum aller $\mathbb{K}$-linearen Abbildungen von V nach W nennen wir $\mathrm{Hom}_{\mathbb{K}}(V,W)$.

Beweis. Wir müssen überprüfen, dass $\mathrm{Hom}_{\mathbb{K}}(V,W) \subseteq W^V$ ein Untervektorraum von W^V ist.

Das Nullelement von W^V ist die konstante 0-Abbildung, also die Abbildung $0 : V \to W, \quad x \mapsto 0_V$, die alle Elemente aus V auf den Nullvektor in W abbildet. Man sieht sofort, dass diese 0-Abbildung $\mathbb{K}$-linear ist. Also gilt $0 \in \mathrm{Hom}_{\mathbb{K}}(V,W)$.

Als nächstes wollen wir zeigen, dass die Summe von zwei linearen Abbildung wieder linear ist. Dazu seien $\varphi, \psi \in \mathrm{Hom}_{\mathbb{K}}(V,W)$ lineare Abbildungen. Dann ist zu zeigen, dass $\varphi + \psi$ auch wieder linear ist. Dazu nehmen wir $x, y \in V$ und erhalten:

$$(\varphi+\psi)(x+y) = \varphi(x+y)+\psi(x+y) = \varphi(x)+\varphi(y)+\psi(x)+\psi(y) = \varphi(x)+\psi(x)+\varphi(y)+\psi(y) = (\varphi+\psi)(x)+(\varphi+\psi)(y).$$

In dieser Gleichungskette haben wir die Definition der Addition von Abbildung verwendet (punktweise), die Tatsache, dass φ und ψ als linear vorausgesetzt waren und die Tatsache, dass Vektoraddition in W kommutativ und assoziativ ist. Wenn wir nun ein $x \in V$ und einen Skalar $\lambda \in \mathbb{K}$ hernehmen, erhalten wir:

$$(\varphi + \psi)(\lambda x) = \varphi(\lambda x) + \psi(\lambda x) = \lambda\varphi(x) + \lambda\psi(y) = \lambda(\varphi(x) + \psi(x)) = \lambda(\varphi + \psi)(x).$$

Dies zeigt, dass $\varphi+\psi : V \to W$ eine $\mathbb{K}$-lineare Abbildung ist. Also ist die Menge $\mathrm{Hom}_{\mathbb{K}}(V,W) \subseteq W^V$ abgeschlossen unter Addition von Abbildungen.

Es bleibt nur noch zu zeigen, dass die Menge $\mathrm{Hom}_{\mathbb{K}}(V,W) \subseteq W^V$ unter skalarer Multiplikation abgeschlossen ist. Sei dazu $\varphi \in \mathrm{Hom}_{\mathbb{K}}(V,W)$ eine lineare Abbildung und $\mu \in \mathbb{K}$. Dann ist zu zeigen, dass $\mu\varphi : V \to W$ eine lineare Abbildung ist. Es gilt aber für alle $x, y \in V$:

$$(\mu\varphi)(x + y) = \mu \cdot (\varphi(x + y)) = \mu \cdot (\varphi(x) + \varphi(y)) = \mu \cdot \varphi(x) + \mu \cdot \varphi(y) = (\mu\varphi)(x) + (\mu\varphi)(y).$$

Ebenso gilt für $x \in V$ und $\lambda \in \mathbb{K}$:

$$(\mu\varphi)(\lambda x) = \mu \cdot \varphi(\lambda x) = \mu(\lambda\varphi(x)) = (\mu\lambda)\varphi(x) = (\lambda\mu)\varphi(x) = \lambda(\mu\varphi(x)) = \lambda(\mu\varphi)(x).$$

Also ist die Abbildung $\mu\varphi : V \to W$ linear. Beachten Sie, wie hier die Kommutativität der Multiplikation im Körper $\mathbb{K}$ verwendet wird[7]. □

[7] Man kann einen Großteil der linearen Algebra auch ohne die Kommutativität von $\mathbb{K}$ aufziehen, indem man nicht fordert, dass $\mathbb{K}$ ein Körper, sondern nur ein Divisionsring (siehe Bemerkung 3.3.13) ist. Die entsprechenden „Vektorräume“ nennt man dann Linksvektorräume. Viele der Aussagen, die wir hier beweisen, gelten auch in diesem nichtkommutativen Setting – dieser Proposition ist ein Beispiel für eine Aussage, die *nicht* mehr gilt.

Zusammenfassung von Abschnitt 4.1

Es sei $\mathbb{K}$ ein Körper.

(1) Ein $\mathbb{K}$-Vektorraum $(V,+,\cdot)$ ist eine abelsche Gruppe $(V,+)$ mit einer skalaren Multiplikation $\cdot : \mathbb{K} \times V \to V$.

(2) Eine $\mathbb{K}$-lineare Abbildung ist ein Homomorphismus von $\mathbb{K}$-Vektorräumen.

(3) Bild und Kern einer $\mathbb{K}$-linearen Abbildung sind Untervektorräume von Ziel- bzw. Definitionsbereich.

(4) Die Menge $\mathrm{Hom}_{\mathbb{K}}(V,W)$ der $\mathbb{K}$-linearen Abbildungen zwischen zwei $\mathbb{K}$-Vektorräumen ist selbst wieder ein $\mathbb{K}$-Vektorraum.

4.2. Lineare Hülle, Basis, Dimension

Wir werden nun weitere Begriffe, die wir in Kapitel 2 im Spezialfall von $\mathbb{R}^n$ und seinen Untervektorräumen eingeführten Begriffe auf beliebige Vektorräume über beliebigen Körpern verallgemeinern:

Da wir in einem Vektorraum $(V,+,\cdot)$ Vektoren addieren und mit Skalaren aus $\mathbb{K}$ multiplizieren können, übertragen sich die Begriffe Linearkombination und lineare Hülle aus Definition 2.3.3 direkt:

Definition 4.2.1 (Lineare Hülle)**.**
Es sei V ein Vektorraum über einem Körper $\mathbb{K}$.

(a) Gegeben seien Vektoren $v_1,\dots,v_r \in V$ mit $r \in \mathbb{N}_0$. Eine *Linearkombination* der Vektoren $v_1,\dots,v_r$ ist eine Summe der Form

$$\sum_{j=1}^{r} \lambda_j v_j = \lambda_1 v_1 + \cdots + \lambda_r v_r,$$

wobei die Skalare $\lambda_j \in \mathbb{K}$ beliebig gewählt sein dürfen.

(b) Gegeben sei eine (endliche oder unendliche) Teilmenge $M \subseteq V$. Wir definieren die *lineare Hülle* der Menge M als die Menge aller Linearkombinationen von Elementen aus M und schreiben:

$$\mathrm{LH}_{\mathbb{K}}(M) := \mathrm{LH}(M) := \left\{ \sum_{j=1}^{r} \lambda_j v_j \;\middle|\; r \in \mathbb{N}_0; \forall j \le r : v_j \in M; \lambda_j \in \mathbb{K} \right\} \subseteq V.$$

Eine leere Summe ist immer 0, deswegen ist $\mathrm{LH}(\emptyset) = \{0\}$.

Wie in Lemma 2.3.4 sieht man, dass $\mathrm{LH}_{\mathbb{K}}(M)$ der kleinste Untervektorraum von V ist, der M enthält.

Falls $M = \{v_1, v_2, \dots, v_r\}$ eine endliche Menge ist, schreiben wir auch

$$\mathrm{LH}_{\mathbb{K}}(v_1, v_2 \dots, v_r) := \mathrm{LH}_{\mathbb{K}}(\{v_1, v_2 \dots, v_r\}).$$

und es gilt:

$$\mathrm{LH}_{\mathbb{K}}(v_1, v_2 \dots, v_r) = \left\{ \lambda_1 v_1 + \cdots + \lambda_r v_r \;\middle|\; \lambda_j \in \mathbb{K} \right\}.$$

(c) Eine Teilmenge $M \subseteq V$ heißt *Erzeugendensystem* für V, wenn

$$\mathrm{LH}_{\mathbb{K}}(M) = V.$$

Bemerkung 4.2.2. Sei $\mathbb{K}$ ein Körper. Für einen Vektorraum V über einem Körper $\mathbb{K}$ sind äquivalent:

- $\mathrm{LH}(\emptyset) = V$.
- $V \cong \mathbb{K}^0 := \{0\}$. (siehe Beispiel 4.1.3 (d)

Der Begriff der linearen Unabhängigkeit (Definition 2.3.9) für *endliche* Teilmengen überträgt sich analog. Den Fall unendlicher Teilmengen können wir auf den endlicher Teilmengen zurückführen:

Definition 4.2.3 (Lineare Unabhängigkeit)**.**

Gegeben sei ein Vektorraum V über einem Körper $\mathbb{K}$.

(a) Es sei $L = \{v_1, \ldots, v_r\} \subseteq V$ eine endliche Menge mit $r \in \mathbb{N}_0$ Elementen, d.h. $v_1, \ldots, v_r$ sind r paarweise verschiedene Vektoren aus V. Wir sagen L ist *linear unabhängig* (oder *linear unabhängig über* $\mathbb{K}$), wenn die einzige Darstellung des Nullvektors als Linearkombination aus L die triviale Linearkombination (alle Skalare null) ist. In Formeln:

$$L \text{ ist linear unabhängig } :\Longleftrightarrow \left(\forall \lambda_1, \ldots, \lambda_r \in \mathbb{K} : \left(\sum_{j=1}^{r} \lambda_j v_j = 0 \implies (\forall j : \lambda_j = 0)\right)\right).$$

(b) Eine unendliche Teilmenge $L \subseteq V$ heißt *linear unabhängig*, wenn jede endliche Teilmenge von L linear unabhängig ist.

Eine Teilmenge, die nicht linear unabhängig ist, heißt *linear abhängig*.

Wenn wir die Begriffe Erzeugendensysteme und lineare Unabhängigkeit eingeführt haben, können wir – wie in Definition 2.3.14 – definieren, was eine Basis ist:

Definition 4.2.4 (Basis)**.**
Es sei V ein Vektorraum über einem Körper $\mathbb{K}$. Eine Teilmenge $B \subseteq V$ heißt *Basis* von V, wenn folgende Bedingungen erfüllt sind:

(i) Die Menge B ist ein Erzeugendensystem für V, d.h. $\mathrm{LH}_{\mathbb{K}}(B) = V$.

(ii) Die Menge B ist linear unabhängig über $\mathbb{K}$.

Beispiel 4.2.5. (a) Die Menge

$$\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ldots, \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \right\}$$

bildet eine Basis für den Raum $V = \mathbb{K}^n$ und wird als *Standardbasis* des $\mathbb{K}^n$ bezeichnet. Die einzelnen Vektoren

$$e_1 := \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad e_2 := \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \cdots, \quad e_n := \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

werden dementsprechend *Standardbasisvektoren* genannt.

(b) Es sei $\mathbb{K}$ ein Körper. Dann ist die Menge

$$B := \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\} \subseteq \mathbb{K}^{2\times 2}$$

eine Basis des Vektorraums $\mathbb{K}^{2\times 2}$.

(c) Allgemeiner: Es seien $m, n \in \mathbb{N}$ und $\mathbb{K}$ ein Körper. Für jedes $k \leq m$ und $l \leq n$ sei $E_{k,l} \in \mathbb{K}^{m\times n}$ die Matrix, die nur aus Nullen besteht mit Ausnahme der Stelle (k,l), wo der Eintrag eine Eins ist. In Formeln:

$$E_{k,l} = \left(\delta_{i,k}\delta_{j,l}\right)_{i=1,\ldots,m;j=1,\ldots,n},$$

wobei $\delta_{\alpha,\beta}$ genau dann den Wert 1 hat, wenn $\alpha = \beta$ und sonst 0.

Die Matrizen der Form $E_{k,l}$ werden auch *Standardmatrizen* genannt.

Die Menge $\left\{E_{k,l} \,\middle|\, k \in \{1,\ldots,m\}, l \in \{1,\ldots,n\}\right\}$ ist eine Basis für den $\mathbb{K}$-Vektorraum $\mathbb{K}^{m\times n}$.

(d) Für jede reelle Zahl $\alpha \in \mathbb{R}$ sei $f_\alpha : \mathbb{R} \to \mathbb{R}$ gegeben durch:

$$f_\alpha : \mathbb{R} \to \mathbb{R}, \quad t \mapsto e^{\alpha t}.$$

Wir werden nun beweisen, dass $L := \{f_\alpha \mid \alpha \in \mathbb{R}\}$ eine linear unabhängige Teilmenge des reellen Vektorraums $\mathbb{R}^{\mathbb{R}}$ aller Funktionen von $\mathbb{R}$ nach $\mathbb{R}$ ist.

Es sei $(f_{\alpha_1}, \ldots, f_{\alpha_r})$ endlich viele Funktion und eine Linearkombination der Nullfunktion gegeben:

$$\sum_{k=1}^{r} \lambda_k f_{\alpha_k} = 0,$$

was bedeutet:

$$\forall t \in \mathbb{R} : \sum_{k=1}^{r} \lambda_k e^{\alpha_k t} = 0. \tag{4.2.1}$$

Es ist zu zeigen: $\forall k = 1, \ldots, r : \lambda_k = 0$.

Es sei $K = \{k \in \{1,\ldots,r\} \mid \lambda_k \neq 0\}$. Wir müssen zeigen, dass $K = \emptyset$ ist.

Wir zeigen dies per Widerspruch. Angenommen, J ist nicht leer. Dann können wir von allen $k \in K$ dasjenige k_0 auswählen, sodass α_{k_0} maximal ist, d.h.

$$\forall k \in K : \alpha_k \leq k_0.$$

Nun teilen wir beide Seiten von Gleichung 4.2.1 durch $e^{\alpha_0 t}$ und erhalten

$$\forall t \in \mathbb{R} : \sum_{k=1}^{r} \lambda_k e^{(\alpha_k - \alpha_{k_0})t} = 0.$$

Der Faktor im Exponenten $\alpha_k - \alpha_{k_0}$ ist immer negativ – außer für $k = k_0$. Da hat er den Wert 0.

$$\forall t \in \mathbb{R} : \sum_{k \neq k_0}^{r} \lambda_k e^{(\alpha_k - \alpha_{k_0})t} + \lambda_{k_0} e^0 = 0.$$

Wenn wir nun auf beiden Seiten den Grenzwert für $t \to \infty$ betrachten, steht auf der linken Seite der Gleichung nur noch λ_{k_0}, weil der Rest gegen 0 konvergiert und auf der rechten Seite steht 0. Also ist $\lambda_{k_0} = 0$, was ein Widerspruch zu der Annahme ist, dass $k_0 \in K$ ist.

(e) Es sei J eine Menge und $\mathbb{K}$ ein Körper. Es sei für jedes $j \in J$ die Funktion $e_j : J \to \mathbb{K}$ gegeben:

$$e_j : J \to \mathbb{K}, \quad x \mapsto \delta_{j,x},$$

wobei wieder $\delta_{j,x}$ genau dann den Wert 1 hat, wenn $j = x$ und sonst 0.

Die Funktion e_j hat also genau an einer Stelle den Wert 1, ist ansonsten gleich der konstanten Nullfunktion.

Die Menge

$$L := \{e_j \mid j \in J\} \subseteq \mathbb{K}^J$$

ist linear unabhängig.

Die lineare Hülle von L ist der Untervektorraum

$$\mathbb{K}^{(J)} = \{f : J \to K \mid f \text{ ist nur an endlichen vielen Stellen ungleich0}\} \subseteq \mathbb{K}^J,$$

bekannt aus Beispiel 4.1.8(g). Falls also J *endlich* ist, dann ist $\mathbb{K}^{(j)} = \mathbb{K}^J$ und L ist eine Basis für $\mathbb{K}^J$. Für unendliches J ist $\mathbb{K}^{(J)} \neq \mathbb{K}^J$ und L ist *keine* Basis für $\mathbb{K}^J$. In jedem Fall ist L aber eine Basis für $\mathbb{K}^{(J)}$.

Im Fall $J = \mathbb{N}$ besteht $\mathbb{K}^{(\mathbb{N})}$ also aus der Menge aller *abbrechenden Folgen* in $\mathbb{K}$, also solche Folgen, die nach endlich vielen Schritten konstant 0 werden.

(f) Da $\mathbb{C}$ eine Körpererweiterung von $\mathbb{R}$ ist, können wir $\mathbb{C}$ als Vektorraum über $\mathbb{R}$ auffassen. Die zweielementige Menge $\{1, \mathrm{i}\} \subseteq \mathbb{C}$ ist linear unabhängig (über $\mathbb{R}$) und ein Erzeugendensystem (über $\mathbb{R}$). Somit ist $\{1, \mathrm{i}\}$ eine $\mathbb{R}$-Basis für $\mathbb{C}$.

(g) Wir können $\mathbb{C}$ aber natürlich auch als Vektorraum über $\mathbb{C}$ auffassen. Dann ist die Menge $\{1, \mathrm{i}\}$ keine Basis, denn sie nicht linear unabhängig über $\mathbb{C}$. Es gilt nämlich:

$$\mathrm{i} \cdot 1 + (-1)\mathrm{i} = 0,$$

es gibt also eine Linearkombination von 1 und i mit *komplexen* Koeffizienten, die nicht alle 0 sind. Wir sehen also: Es kommt auf den Grundkörper an, ob eine Menge linear unabhängig ist oder nicht.

(h) Wir können auch $\mathbb{R}$ als Vektorraum über $\mathbb{Q}$ auffassen. Dann ist die Menge $\{1, \sqrt{2}\} \subseteq \mathbb{R}$ linear unabhängig. Beweis: Gegeben rationale Zahlen $\alpha, \beta \in \mathbb{Q}$ mit

$$\alpha \cdot 1 + \beta \cdot \sqrt{2} = 0.$$

Falls $\beta \neq 0$ ist, können wir nach $\sqrt{2}$ auflösen und erhalten $\sqrt{2} = -\frac{\alpha}{\beta}$. Da $\alpha, \beta \in \mathbb{Q}$ sind und $\mathbb{Q}$ ein Körper, wäre damit auch $\sqrt{2} \in \mathbb{Q}$, was aber bekanntermaßen nicht stimmt. Also muss $\beta = 0$ sein. Dann ergibt sich aber direkt, dass auch $\alpha = 0$ ist.

Es ist im Augenblick noch vollkommen unklar, ob ein gegebener Vektorraum V immer eine Basis hat. Was wir aber schnell sehen können, ist, dass sich Basen mit Hilfe von Vektorraumisomorphismen übertragen lassen:

Lemma 4.2.6 (Isomorphismen und Basen)**.**
Es seien V und W Vektorräume über demselben Körper $\mathbb{K}$ und $\varphi : V \to W$ sei eine lineare Abbildung.

(a) Wenn $M \subseteq V$ ein Erzeugendensystem von V ist und φ surjektiv, dann ist $\varphi(M)$ ein Erzeugendensystem für W.

(b) Wenn $L \subseteq V$ linear unabhängig ist und φ injektiv, dann ist $\varphi(L)$ linear unabhängig.

(c) Wenn $B \subseteq V$ eine Basis von V ist und φ bijektiv (also ein Isomorphismus von $\mathbb{K}$-Vektorräumen), dann ist $\varphi(B)$ eine Basis von W.

Beweis. (a)
Es sei $w \in W$ gegeben. Es ist zu zeigen, dass $w \in \mathrm{LH}_{\mathbb{K}}\big(\varphi(M)\big)$. Aus der Surjektivität von $\varphi : V \to W$ folgt nun, dass es ein $v \in V$ gibt mit $\varphi(v) = w$.

Der Vektor $v \in V = \mathrm{LH}_{\mathbb{K}}(M)$ ist eine Linearkombination von Vektoren aus M, also gibt es $\lambda_1, \ldots, \lambda_r \in \mathbb{K}$ und $v_1, \ldots, v_r \in M$ mit

$$v = \sum_{j=1}^{r} \lambda_j v_j.$$

Nun wenden wir φ auf beide Seiten an und nutzen aus, dass φ linear ist:

$$\varphi(v) = \varphi\left(\sum_{j=1}^{r} \lambda_j v_j\right) = \sum_{j=1}^{r} \lambda_j \varphi(v_j).$$

Dies zeigt, dass $w = \varphi(v) \in \mathrm{LH}_{\mathbb{K}}\big(\varphi(M)\big)$.

(b)
Um zu zeigen, dass $\varphi(L) \subseteq W$ linear unabhängig ist, müssen wir zeigen, dass jede endliche Teilmenge von $\varphi(L)$ linear unabhängig ist.

Seien also $w_1, \ldots, w_r \in \varphi(L)$ paarweise verschiedene Elemente gegeben. Wir nehmen an, es gibt Skalare $\lambda_1, \ldots, \lambda_r \in \mathbb{K}$ mit

$$\sum_{j=1}^{r} \lambda_j w_j = 0_W. \tag{4.2.2}$$

Wir müssen nun zeigen, dass alle λ_j Null sind.

Da jedes w_j in $\varphi(L)$ liegt, gibt es also $v_j \in L$ mit $\varphi(v_j) = w_j$. Somit können wir Formel (4.2.2) umschreiben zu:

$$\sum_{j=1}^{r} \lambda_j \varphi(v_j) = 0_W.$$

Aus der Linearität von φ folgt nun:

$$\varphi\left(\sum_{j=1}^{r} \lambda_j v_j\right) = 0_W,$$

was bedeutet, dass diese Summe im Kern von φ liegt.

Nun ist aber φ injektiv und nach Lemma 4.1.15 bedeutet dies, dass $\ker \varphi = \{0_V\}$.

Also gilt:

$$\sum_{j=1}^{r} \lambda_j v_j = 0_V.$$

Da die Vektoren $v_j \in L$ alle paarweise verschieden sind (dies folgt daraus, dass die w_j alle paarweise verschieden sind) und die Menge L linear unabhängig ist, folgt dass alle Skalare Null sein müssen:

$$\lambda_1 = \cdots = \lambda_r = 0_{\mathbb{K}}.$$

Das war zu zeigen.

(c)

Dies ist einfach nur eine Kombination aus den Teilen (a) und (b). □

Um eine Basis zu konstruieren, würden wir gerne den Basisauswahl- und -ergänzungssatz (Satz 2.3.18) verwenden. Allerdings haben wir in dem Beweis verwendet, dass es eine obere Schranke für die Größe linear unabhängiger Teilmengen gibt und das ist nicht in allen Vektorräumen der Fall.

Satz 4.2.7 (Basisauswahl- und -ergänzungssatz für endlich erzeugte Vektorräume)**.**
Es sei $\mathbb{K}$ ein Körper. Gegeben seien

- *ein $\mathbb{K}$-Vektorraum V,*
- *ein Erzeugendensystem M von V,*
- *eine linear unabhängige Teilmenge $L \subseteq M$.*

Wenn wir nun zusätzlich annehmen, dass jede linear unabhängige Teilmenge von M endlich ist, dann gibt es eine endliche Basis B von V mit

$$L \subseteq B \subseteq M.$$

Beweis. Wir können den Beweis von Satz 2.3.18 fast wörtlich übernehmen, die Idee ist also wieder: Wir fangen mit der Menge L an. Wenn sich mit dieser Menge alle Elemente aus M erzeugen lassen, sind wir fertig, weil L dann ein linear unabhängiges Erzeugendensystem, also eine Basis ist.

Wenn nicht, gibt es ein Element aus $M \setminus L$, das wir zu L hinzufügen können, sodass die so konstruierte größere Menge immer noch linear unabhängig ist. Details zu diesem Argument sind wörtlich aus dem Beweis von Satz 2.3.18 zu übernehmen.

Was wir noch brauchen, ist ein Argument, warum dieser Prozess nach endlich vielen Schritten abbricht. Nehmen wir per Widerspruch an, er würde nicht terminieren. Dann erhalten wir eine aufsteigende Folge $(L_k)_{k\in\mathbb{N}}$ von immer größer werdenden linear unabhängigen Teilmengen, die alle in M enthalten sind.

Wir nehmen nun die Vereinigung $L_\omega := \sum_{k\in\mathbb{N}} L_k$. Dies ist nun eine unendliche linear unabhängige Teilmenge von V, die in M enthalten ist, was nach Voraussetzung nicht sein kann. Dies zeigt, dass der Prozess nach endlich vielen Schritten abbricht und wir eine endliche Basis erhalten. □

Bemerkung. Es ist eine interessante Frage, ob dieser Satz auch ohne die zusätzliche Bedingung, dass jede linear unabhängige Teilmenge von M endlich sein soll, gilt und wir so (im Allgemeinen) unendliche Basen von Vektorräumen konstruieren können. Dies hängt davon

ab, ob wir bei der Axiomatisierung der Mengenlehre das sogenannte *Auswahlaxiom* hinzugenommen haben oder nicht. Wenn wir dieses Auswahlaxiom als wahr annehmen, dann gilt der Basisauswahl- und ergänzungssatz für jeden Vektorraum, insbesondere besitzt also jeder Vektorraum eine Basis. Ohne das Auswahlaxiom ist es allerdings nicht möglich, jedem Vektorraum eine Basis zuzuweisen. In jedem Falle ist es so, dass es Vektorräume gibt, wo es nicht möglich ist, eine Basis konkret anzugeben. Diese Subtilitäten sollen uns aber für diese Veranstaltung erst einmal egal sein.[8]

Die Begriffe Erzeugendensystem, linear unabhängige Teilmenge und Basis sind eng verwandt mit den Begriffen surjektiv, injektiv und bijektiv:

Satz 4.2.8.
Gegeben sei ein Vektorraum V über einem Körper $\mathbb{K}$, $n \in \mathbb{N}_0$ und paarweise verschiedenen Vektoren $v_1, \dots, v_n \in V$. Wir untersuchen die Abbildung:

$$\psi : \mathbb{K}^n \to V, \quad \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto \sum_{j=1}^{n} x_j v_j.$$

(a) Die Abbildung $\psi : \mathbb{K}^n \to V$ ist $\mathbb{K}$-linear, also ein Homomorphismus von $\mathbb{K}$-Vektorräumen.

(b) Die Abbildung $\psi : \mathbb{K}^n \to V$ genau dann injektiv, wenn $\{v_1, \dots, v_n\}$ linear unabhängig ist.

(c) Die Abbildung $\psi : \mathbb{K}^n \to V$ genau dann surjektiv, wenn $\{v_1, \dots, v_n\}$ ein Erzeugendensystem ist.

(d) Die Abbildung $\psi : \mathbb{K}^n \to V$ genau dann ein Isomorphismus von $\mathbb{K}$-Vektorräumen, wenn $\{v_1, \dots, v_n\}$ eine Basis von V ist.

Beachten Sie, dass der Fall $n = 0$ explizit erlaubt ist. In diesem Fall ist $\mathbb{K}^0 = \{0\}$ (siehe Beispiel 4.1.3(d)) und $\psi : \{0\} \to V, \quad 0 \mapsto 0_V$.

Beweis. (a)
Es gilt:

$$\psi\left(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}\right) = \sum_{j=1}^{n} (x_j + y_j) v_j = \sum_{j=1}^{n} x_j v_j + \sum_{j=1}^{n} y_j v_j = \psi\left(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}\right) + \psi\left(\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}\right)$$

und

$$\psi\left(\lambda \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}\right) = \sum_{j=1}^{n} (\lambda x_j) v_j = \lambda \sum_{j=1}^{n} x_j v_j = \lambda \psi\left(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}\right).$$

Also ist $\varphi : \mathbb{K}^n \to V$ linear. (b)
Die Vektoren $v_1, \dots, v_n$ sind linear unabhängig, genau dann, wenn die folgende Implikation gilt:

$$\forall x_1, \dots, x_n \in \mathbb{K} : \left(\left(\sum_{j=1}^{n} x_j v_j = 0\right) \implies \left(\forall j \in \{1, \dots, n\} : x_j = 0\right)\right).$$

[8] Es gibt allerdings ein Video über das Auswahlaxiom im ILIAS-Kurs.

Dies lässt sich äquivalent umschreiben als:

$$\forall \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{K}^n : \left(\left(\Psi \left(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \right) = 0 \right) \Longrightarrow \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = 0 \right),$$

was nichts anderes bedeutet, als $\ker(\psi) = \{0\}$. Und das ist nach Lemma 4.1.15 äquivalent dazu, dass $\psi : \mathbb{K}^n \to V$ injektiv ist.

(c)

Die lineare Hülle von $\{v_1, \ldots, v_n\}$ ist genau die Menge der Linearkombinationen der Vektoren und somit genau das Bild der Abbildung ψ.

(d)

Dies ist einfach die Kombination von (b) und (c). □

Satz 4.2.9.

Es sei V ein Vektorraum über einem Körper $\mathbb{K}$.

(a) Die folgenden Aussagen sind äquivalent:

(i) Es gibt eine endliche Menge $M \subseteq V$, die V erzeugt.

(ii) Es gibt eine Zahl $n \in \mathbb{N}_0$, sodass jede Menge mit mehr als n Elementen linear abhängig ist.

(iii) Es gibt eine endliche Basis $B \subseteq V$.

(iv) Es gibt eine Zahl $n \in \mathbb{N}_0$, sodass $V \cong_{\mathbb{K}} \mathbb{K}^n$.

Ein Vektorraum mit diesen Eigenschaften heißt endlich dimensional. *Ansonsten heißt er* unendlich dimensional.

(b) Sei V ein endlich dimensionaler $\mathbb{K}$-Vektorraum.

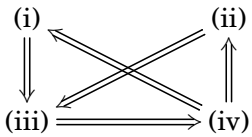
Dann hat jede Basis von V gleich viele Elemente. Die Anzahl der Elemente in einer Basis heißt Dimension *von V und wird mit* $\dim_{\mathbb{K}}(V) = \dim(V) \in \mathbb{N}_0$ *bezeichnet.*

Wenn V unendlich dimensional ist, schreibt man auch $\dim_{\mathbb{K}}(V) = \infty$.

(c) Die Dimension von V ist außerdem die minimale Anzahl von Elementen in einem Erzeugendensystem von V, die maximale Anzahl von Elementen in einer linear unabhängigen Teilmenge von V und die einzige Zahl $n \in \mathbb{N}_0$ mit $V \cong \mathbb{K}^n$.

Beweis. (a)

Wir müssen eine Reihe von Implikationen zeigen.:



Falls (i) gilt, können wir Satz 4.2.7 auf das Erzeugendensystem M und die linear unabhängige Teilmenge $\emptyset$ anwenden und erhalten eine endliche Basis $B \subseteq M$. Also gilt: (i) $\Longrightarrow$ (iii).

Auch falls (ii) gilt, können wir Satz 4.2.7 anwenden, diesmal auf das Erzeugendensystem V und die lineare unabhängige Teilmenge $\emptyset$. Da jede Menge mit mehr als n Elementen linear abhängig ist, sind insbesondere alle linear unabhängigen Mengen endlich. Dies zeigt (ii) $\Longrightarrow$ (iii).

Nehmen wir nun an, dass (iii) gilt und V eine endliche Basis B hat. Nach Satz 4.2.8 ist V isomorph zu $\mathbb{K}^n$. Im Falle, dass die Basis leer ist, ist V also isomorph zu $\mathbb{K}^0 = \{0\}$. Also gilt: (iii) $\Longrightarrow$ (iv).

Falls (iv) gilt, dann ist $V \cong \mathbb{K}^n$, d.h. entweder ist $V = \{0\}$ oder $V \cong \mathbb{K}^n$ für $n \in \mathbb{N}$. Im ersten Fall ist klar, dass (i) und (ii) gelten.

Nehmen wir also nun an, dass es einen Isomorphismus $\Phi : \mathbb{K}^n \to V$ gibt. Der Raum $\mathbb{K}^n$ hat eine Basis mit n Elementen, nämlich die Standardbasis. Also können wir den Isomorphismus Φ verwenden und die Basis auf den Raum V übertragen (siehe Lemma 4.2.6). Also hat V eine endliche Basis und somit insbesondere ein endliches Erzeugendensystem. Also gilt: (iv) $\Longrightarrow$ (i).

Für den Raum $\mathbb{K}^n$ haben wir gezeigt, dass jede Teilmenge mit mehr als n Elementen linear abhängig ist (Satz 2.3.13 für den Fall $\mathbb{K} = \mathbb{R}$, aber der Beweis funktioniert genauso auch für beliebige Körper). Da der Vektorraum V isomorph ist zu $\mathbb{K}^n$ und Isomorphismen lineare unabhängige Mengen auf linear unabhängige Mengen übertragen (siehe wieder Lemma 4.2.6) gilt die entsprechende Aussage auch für den Raum V. Somit gilt (iv) $\Longrightarrow$ (ii).

(b)

Wir wissen nach (a), dass es ein $n \in \mathbb{N}_0$ gibt, sodass $V \cong \mathbb{K}^n$. Es gibt also einen Isomorphismus von $\mathbb{K}$-Vektorräumen

$$\Phi : \mathbb{K}^n \to V.$$

Seien nun B und C Basen von V gegeben. Dann sind $\Phi^{-1}(B)$ und $\Phi^{-1}(C)$ Basen von $\mathbb{K}^n$. Nach Satz 2.3.22 haben zwei Basen von $\mathbb{K}^n$ immer gleich viele Elemente. Also sind auch B und C gleichmächtig und der Begriff der Dimension ist wohldefiniert.

(c)

Es sei M ein endliches Erzeugendensystem von V. Dann können wir nach Satz 4.2.7 eine Basis B von V finden mit $B \subseteq V$. Nach (b) haben alle Basen gleich viele Elemente und somit ist

$$|M| \geq |B| = \dim_{\mathbb{K}}(V).$$

Jede linear unabhängige Teilmenge L lässt sich nach Satz 4.2.7 zu einer Basis B vervollständigen und somit gilt:

$$|L| \leq |B| = \dim_{\mathbb{K}}(V).$$

Wenn V isomorph zu $\mathbb{K}^m$ ist, dann können wir die Standardbasis von $\mathbb{K}^m$ und den Isomorphismus zwischen V und $\mathbb{K}^n$ verwenden, um eine Basis von B mit genau m Elementen zu konstruieren. Also muss $m = \dim_{\mathbb{K}}(V)$ sein. □

Lemma 4.2.10 (Monotonie der Dimension)**.**
Es sei V ein Vektorraum über einem Körper $\mathbb{K}$ und $W \subseteq V$ ein Untervektorraum.

(a) Wenn V endlich dimensional ist, so ist auch W endlich dimensional und es gilt

$$\dim_{\mathbb{K}}(W) \leq \dim_{\mathbb{K}}(V).$$

(b) Wenn V endlich dimensional ist und $W \subsetneq V$, dann gilt sogar

$$\dim_{\mathbb{K}}(W) < \dim_{\mathbb{K}}(V).$$

(c) Wenn W unendlich dimensional ist, so ist auch V unendlich dimensional.

Beweis. Teil (c) folgt logisch direkt aus Teil (a) durch Kontraposition.

Teile (a) und (b) sind direkte Folgerungen aus dem Basisergänzungssatz und lassen sich genauso wie Lemma 2.3.26 beweisen. □

Beispiel 4.2.11. (a) Für jedes $n \in \mathbb{N}_0$ ist der Raum $\mathbb{K}^n$ endlich dimensional mit $\dim_{\mathbb{K}}(\mathbb{K}^n) = n$.

(b) Für jedes $m, n \in \mathbb{N}$ ist der Raum $\mathbb{K}^{m\times n}$ endlich dimensional mit $\dim_{\mathbb{K}}(\mathbb{K}^{m\times n}) = mn$.

(c) Es gilt $\dim_{\mathbb{R}}(\mathbb{C}) = 2$ und $\dim_{\mathbb{C}}(\mathbb{C}) = 1$.

(d) Für unendliches J ist der Raum $\mathbb{K}^{(J)}$ unendlich dimensional über $\mathbb{K}$ (nach Beispiel 4.2.5(e)). Der Raum $\mathbb{K}^J$ enthält $\mathbb{K}^{(J)}$ und ist somit erst recht unendlich dimensional über $\mathbb{K}$.

(e) Die Folgenräume $\ell^1(\mathbb{N},\mathbb{R}), c_0(\mathbb{N},\mathbb{R}), c(\mathbb{N},\mathbb{R}), \ell^\infty(\mathbb{N},\mathbb{R}), \mathbb{R}^{\mathbb{N}}$ enthalten alle den unendlich dimensionalen Raum $\mathbb{R}^{(\mathbb{N})}$ als Untervektorraum und sind somit alle unendlich dimensional über $\mathbb{R}$.

Jetzt, wo wir die Dimension eines beliebigen Vektorraums (als Element in $\mathbb{N}_0 \cup \{\infty\}$) definiert haben, bieten sich die folgenden Definitionen als naheliegende Verallgemeinerungen aus Kapitel 2 an:

Definition 4.2.12 (Dimension eines affinen Unterraums)**.**
Es sei $\mathbb{K}$ ein Körper, V ein $\mathbb{K}$-Vektorraum und R ein affiner Unterraum. Wenn wir $R = p + U$ schreiben mit einem Aufpunkt $p \in V$ und einem Untervektorraum $U \subseteq V$, dann sagen wir R ist endlich dimensional, wenn U endlich dimensional ist und definieren $\dim_{\mathbb{K}}(R) := \dim_{\mathbb{K}}(U) \in \mathbb{N}_0 \cup \{\infty\}$.

Definition 4.2.13 (Rang einer linearen Abbildung)**.**
Es sei $\mathbb{K}$ ein Körper, V und W Vektorräume über $\mathbb{K}$ und $\varphi : V \to W$ eine $\mathbb{K}$-lineare Abbildung. Der *Rang* von φ ist definiert als

$$\operatorname{rg}\varphi := \dim_{\mathbb{K}}(\operatorname{Bild}(\varphi)) = \dim_{\mathbb{K}}(\varphi(V)) \in \mathbb{N}_0 \cup \{\infty\}.$$

Satz 4.2.14 (Fortsetzungssatz)**.**
Es sei $\mathbb{K}$ ein Körper und V und W Vektorräume über $\mathbb{K}$.

(a) Es sei $M \subseteq V$ ein Erzeugendensystem für V und $f : M \to W$ irgendeine Abbildung. Dann gibt es höchstens eine lineare Abbildung $\varphi : V \to W$ mit $\varphi|_M = f$. Eine lineare Abbildung ist also eindeutig bestimmt, wenn man sie auf einem Erzeugendensystem kennt.

(b) Es sei $B \subseteq V$ eine Basis von V und $f : B \to W$ irgendeine Abbildung. Dann gibt es genau eine lineare Abbildung $\varphi : V \to W$ mit $\varphi|_B = f$. Es gilt sogar:

$$\operatorname{Hom}_{\mathbb{K}}(V, W) \to W^B, \quad \varphi \mapsto \varphi|_B$$

ist ein Vektorraumisomorphismus.

Beweis. (a)
Angenommen, es gäbe zwei lineare Abbildungen $\varphi, \psi : V \to W$ mit $\varphi|_M = \psi|_M$. Dann wollen wir zeigen, dass $\varphi = \psi$, d.h. dass für alle $v \in V$ gilt $\varphi(v) = \psi(v)$.

Es sei also $v \in V$ gegeben. Da $V = \mathrm{LH}_{\mathbb{K}}(M)$ gilt, gibt es endlich viele $v_1, \ldots, v_r \in M$ und Skalare $\lambda_1, \ldots, \lambda_r \in \mathbb{K}$ mit $v = \lambda_1 v_1 + \cdots + \lambda_r v_r$. Hieraus folgt dann:

$$\varphi(v) = \varphi\left(\sum_{j=1}^{r} \lambda_j v_j\right) = \sum_{j=1}^{r} \lambda_j \varphi\left(v_j\right) = \sum_{j=1}^{r} \lambda_j \psi\left(v_j\right) = \psi\left(\sum_{j=1}^{r} \lambda_j v_j\right) = \psi(v).$$

(b)
Da B ein Erzeugendensystem ist, gibt es für jedes $v \in V$ gibt es endlich viele $v_1, \ldots, v_r \in M$ und Skalare $\lambda_1, \ldots, \lambda_r \in \mathbb{K}$ mit $v = \lambda_1 v_1 + \cdots + \lambda_r v_r$. Diese Notation können wir wie folgt vereinfachen:

$$v = \sum_{b \in B} \lambda_b b,$$

wenn wir vereinbaren, dass nur endlich viele der Skalare λ_j ungleich 0 sind. Die Summe ist dann zwar formal eine potenziell unendliche, aber da nur endlich viele Summanden ungleich 0 sind, ist sie im Endeffekt eine endliche Summe.

Wir wollen nun definieren:

$$\varphi : V \to W, \quad \sum_{b \in B} \lambda_b b \mapsto \sum_{b \in B} \lambda_b f(b), \text{ wobei immer nur endlich viele der Skalare ungleich 0 sind.}$$

Es bleibt aber zu zeigen, dass diese Abbildung wirklich wohldefiniert ist. Wir haben gesehen, dass sich jedes $v \in V$ so schreiben lässt – aber ist diese Darstellung auch eindeutig?

Nehmen wir an,

$$v = \sum_{b \in B} \lambda_b b = \sum_{b \in B} \mu_b b,$$

wobei sowohl $\{b \in B \mid \lambda_b \neq 0\}$ als auch $\{b \in B \mid \lambda_b \neq 0\}$ endlich sind.

Dann gilt:

$$\sum_{b \in B} (\lambda_b - \mu) b = \sum_{b \in B} \lambda_b b - \sum_{b \in B} \mu_b b = v - v = 0.$$

Da die Menge B linear unabhängig ist, müssen alle $\lambda_b - \mu$ gleich 0 sein und es folgt:

$$\forall b \in B : \lambda_b = \mu_b.$$

Also ist die Darstellung eindeutig und die Abbildung $\varphi : V \to W$ somit wohldefiniert.

Es bleibt zu zeigen: φ ist linear und $\varphi | B = f$. Gegeben $v, w \in V$. Wir wollen zeigen $\varphi(v + w) = \varphi(v) + \varphi(w)$.

Aus $v, w \in V = \mathrm{LH}_{\mathbb{K}}(B)$ folgt: Es gilt:

$$v = \sum_{b \in B} \lambda_b b \quad \text{und} \quad w = \sum_{b \in B} \mu_b b$$

mit $\lambda_b, \mu_b \in \mathbb{K}$. Es folgt:

$$\begin{aligned}
\varphi(v + w) &= \varphi\left(\sum_{b \in B} \lambda_b b + \sum_{b \in B} \mu_b b\right) \\
&= \varphi\left(\sum_{b \in B} (\lambda_b + \mu_b) b\right) \\
&= \sum_{b \in B} (\lambda_b + \mu_b) f(b) \\
&= \sum_{b \in B} \lambda_b f(b) + \sum_{b \in B} \mu_b b f(b) \\
&= \varphi(v) + \varphi(w).
\end{aligned}$$

Es sei nun $v \in V$ und $\mu \in \mathbb{K}$. Wir wollen zeigen, dass $\varphi(\mu v) = \mu\varphi(v)$ gilt.

Aus $v \in V = \mathrm{LH}_{\mathbb{K}}(B)$ folgt: Es gilt:

$$v = \sum_{b \in B} \lambda_b b$$

mit $\lambda_b \in \mathbb{K}$. Es folgt:

$$\begin{aligned}
\varphi(\mu v) &= \varphi\left(\mu \sum_{b \in B} \lambda_b b\right) \\
&= \varphi\left(\sum_{b \in B} (\mu\lambda_b) b\right) \\
&= \sum_{b \in B} (\mu\lambda_b) f(b) \\
&= \mu \sum_{b \in B} \lambda_b f(b) \\
&= \mu\varphi(v).
\end{aligned}$$

Also ist $\varphi : V \to W$ eine $\mathbb{K}$-lineare Abbildung.

Es sei nun $b_0 \in B$. Dann können wir b_0 schreiben als:

$$b_0 = \sum_{b \in B} \lambda_b b,$$

wobei $\lambda_{b_0} = 1$ und alle anderen $\lambda_b = 0$ sind. Es gilt dann:

$$\varphi(b_0) = \varphi\left(\sum_{b \in B} \lambda_b b\right) = \sum_{b \in B} \lambda_b f(b) = f(b_0).$$

Also gilt $\varphi|_B = f$. Das war zu zeigen. □

Zusammenfassung von Abschnitt 4.2

(1) Die Konzepte Lineare Hülle, lineare Unabhängigkeit und Basen sind für beliebige Vektorräume definiert.

(2) Ein Vektorraum ist endlich dimensional, wenn er eine endliche Basis besitzt.

(3) Jeder endlich dimensionale $\mathbb{K}$-Vektorraum ist isomorph zu $\mathbb{K}^n$ für ein $n \in \mathbb{N}_0$.

(4) Eine lineare Abbildung ist eindeutig bestimmt, wenn man die Funktionswerte auf einer Basis des Definitionsbereichs kennt.

4.3. Lineare Abbildungen zwischen endlich dimensionalen Vektorräumen und Darstellungsmatrizen

In Beispiel 4.1.16(a) haben wir gesehen, dass die Linksmultiplikation mit einer Matrix $A \in \mathbb{K}^{m \times n}$ immer eine $\mathbb{K}$-lineare Abbildung von $\mathbb{K}^n$ nach $\mathbb{K}^m$ definiert. Abbildungen dieser Art haben wir in Kapitel 2 schon verwendet, um Lösungsmengen von (homogenen und inhomogenen)

linearen Gleichungssystemen zu studieren. Es stellt sich heraus, dass *jede* lineare Abbildung von $\mathbb{K}^n$ nach $\mathbb{K}^m$ von diesem Typ ist und dass die Matrix A, die zu dieser Abbildung gehört, eindeutig bestimmt ist:

Satz 4.3.1 (Darstellungsmatrix einer Abbildung von $\mathbb{K}^m$ nach $\mathbb{K}^n$)**.**
Es sei $\mathbb{K}$ ein Körper und $m,n,p \in \mathbb{N}$.

(a) Für jede Matrix $A \in \mathbb{K}^{m\times n}$ ist die Abbildung

$$\varphi_A : \mathbb{K}^n \to \mathbb{K}^m, \quad x \mapsto Ax$$

$\mathbb{K}$-linear, also $\varphi_A \in \mathrm{Hom}_{\mathbb{K}}(\mathbb{K}^n, \mathbb{K}^m)$.

In den Spalten der Matrix A stehen die Bilder der Standardbasisvektoren, d.h.

$$A = \left(\left| \varphi_A(e_1) \right| \cdots \left| \varphi_A(e_n) \right| \right),$$

mit den Standardbasisvektoren $e_1, \ldots, e_n$ aus Definition 2.3.17 und Beispiel 4.2.5(a).

(b) Umgekehrt gibt es zu jeder linearen Abbildung $\varphi : \mathbb{K}^n \to \mathbb{K}^m$ eine Matrix $A \in \mathbb{K}^{m\times n}$ mit $\varphi = \varphi_A$. Diese Matrix ist eindeutig und wird Darstellungsmatrix *von φ (bezüglich der Standardbasis) genannt und (zumindest in dieser Veranstaltung) mit $\mathrm{M}_{\mathsf{E},\mathsf{E}}(\varphi)$ bezeichnet.*

(c) Gegeben seien Matrizen $A \in \mathbb{K}^{m\times n}$ und $B \in \mathbb{K}^{n\times p}$. Dann ist

$$\varphi_A \circ \varphi_B = \varphi_{AB},$$

das heißt: Multiplizieren von Matrizen entspricht dem Verketten von Abbildungen.

Für lineare Abbildungen $\varphi : V \to W$ und $\psi : U \to V$ gilt also:

$$\mathrm{M}_{\mathsf{E},\mathsf{E}}(\varphi)\,\mathrm{M}_{\mathsf{E},\mathsf{E}}(\psi) = \mathrm{M}_{\mathsf{E},\mathsf{E}}(\varphi \circ \psi),$$

wobei auf das Produkt auf der linken Seite das gewöhnliche Matrixprodukt ist.

(d) Die Einheitsmatrix $\mathbb{1}_n \in \mathbb{K}^{n\times n}$ entspricht der Identitätsabbildung:

$$\varphi_{\mathbb{1}_n} = \mathrm{id}_{\mathbb{K}^n} \quad \textit{und} \quad \mathrm{M}_{\mathsf{E},\mathsf{E}}(\mathrm{id}_{\mathbb{K}^n}) = \mathbb{1}_n.$$

(e) Die Abbildung

$$\mathbb{K}^{m\times n} \to \mathrm{Hom}_{\mathbb{K}}(\mathbb{K}^n, \mathbb{K}^m), \quad A \mapsto \varphi_A,$$

die jede Matrix $A \in \mathbb{K}^{m\times n}$ auf die entsprechende Abbildung φ_A abbildet, ist ein Isomorphismus von $\mathbb{K}$-Vektorräumen mit Umkehrabbildung

$$\mathrm{Hom}_{\mathbb{K}}(\mathbb{K}^n, \mathbb{K}^m) \to \mathbb{K}^{m\times n}, \quad \varphi \mapsto \mathrm{M}_{\mathsf{E},\mathsf{E}}(\varphi).$$

(f) Eine lineare Abbildung $\varphi = \varphi_A : \mathbb{K}^n \to \mathbb{K}^m$ ist genau dann bijektiv (also ein Vektorraumisomorphismus), wenn $m = n$ und die quadratische Matrix $A = \mathrm{M}_{\mathsf{E},\mathsf{E}}(\varphi)$ im Ring $(\mathbb{K}^{n\times n}, +, \cdot)$ multiplikativ invertierbar ist . In diesem Fall gilt dann:

$$\varphi_{A^{-1}} = (\varphi_A)^{-1} \quad \textit{und} \quad \left(\mathrm{M}_{\mathsf{E},\mathsf{E}}(\varphi)\right)^{-1} = \mathrm{M}_{\mathsf{E},\mathsf{E}}(\varphi^{-1}).$$

Beweis. (a)
Dass $\varphi_A : \mathbb{K}^n \to \mathbb{K}^m$ linear ist, haben wir bereits in Beispiel 4.1.16(a) gesehen und folgt – wie dort schon gesagt – direkt aus Lemma 2.2.12. Es sei nun $r \in \{1, \dots, n\}$ und $e_r \in \mathbb{K}^n$ der r-te Standardbasisvektor, der in der r-ten Spalte eine 1 hat und sonst nur 0:

$$e_r = \left(\delta_{j,r}\right)_{j=1,\dots,n}.$$

mit der Notation

$$\delta_{i,r} = \begin{cases} 1 & \text{falls } i = r \\ 0 & \text{sonst.} \end{cases}$$

Nun multiplizieren wir diesen Vektor von links mit der Matrix $A = (a_{i,j})_{i=1,\dots,m;j=1,\dots,n}$ und erhalten

$$\begin{aligned} \varphi_A(e_r) &= Ae_r \\ &= (a_{i,j})_{i,j}(\delta_{j,r})_{j=1,\dots,n} \\ &= \left(\sum_{j=1}^{n} a_{i,j}\delta_{j,r}\right)_{i=1,\dots,m} \\ &= \left(a_{i,r}\right)_{i=1,\dots,m} \end{aligned}$$

und die ist genau die r-te Spalte der Matrix A.

Wir haben also gesehen:
In der r-ten Spalte der Matrix A steht das Bild der r-ten Standardbasisvektors e_r.

(b)
Es sei $\varphi : \mathbb{K}^n \to \mathbb{K}^m$ eine $\mathbb{K}$-lineare Abbildung gegeben. Wir wissen aus Teil (a), dass – wenn es eine Matrix A gibt mit $\varphi = \varphi_A$ – diese Matrix A eindeutig durch φ bestimmt ist, da in den Spalten von A ja die Bilder $\varphi(e_1), \dots, \varphi(e_n)$ stehen müssen. Die zeigt die Eindeutigkeit von A. Zu Existenz: Die einzige Wahl von A besteht darin, A wie folgt zu definieren:

$$A = \left(\left| \varphi(e_1) \right| \cdots \left| \varphi(e_n) \right| \right).$$

Es ist zu zeigen, dass $\varphi_A = \varphi$. Wir müssen also zeigen, dass für jedes $x = (x_j)_{j=1,\dots,n} \in \mathbb{K}^n$ die Werte $\varphi(x)$ und $\varphi_A(x)$ übereinstimmen:

$$\begin{aligned} \varphi_A(x) &= Ax \\ &= \left(\left| \varphi(e_1) \right| \cdots \left| \varphi(e_n) \right| \right) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \\ &= x_1\varphi(e_1) + \cdots + x_n\varphi(e_n) \\ &= \varphi(x_1e_1) + \cdots + \varphi(x_ne_n) \\ &= \varphi(x_1e_1 + \cdots + x_ne_n) \\ &= \varphi(x). \end{aligned}$$

(c)
Es gilt für alle $x \in \mathbb{K}^p$:

$$(\varphi_A \circ \varphi_B)(x) = \varphi_A\left(\varphi_B(x)\right) = A\left(Bx\right) = ABx = (AB)x = \varphi_{AB}(x).$$

Hier haben wir die Assoziativität des Matrixproduktes ausgenutzt.

(d)

Es gilt für alle $x \in \mathbb{K}^n$:

$$\varphi_{\mathbb{1}_n}(x) = \mathbb{1}_n x = x = \mathrm{id}_{\mathbb{K}^n}(x).$$

(e)

Wir wollen zeigen, dass

$$\Psi : \mathbb{K}^{m\times n} \to \mathrm{Hom}_{\mathbb{K}}(\mathbb{K}^n, \mathbb{K}^m), \quad A \mapsto \varphi_A$$

ein Isomorphismus von $\mathbb{K}$-Vektorräumen ist, d.h. Ψ ist bijektiv und $\mathbb{K}$-linear. Die Bijektivität haben wir in Teil (b) gezeigt. Es bleibt die Linearität: Gegeben seien dazu zwei Matrizen $A, B \in \mathbb{K}^{m\times n}$. Dann sind $\Psi(A+B), \Psi(A)+\Psi(B) \in \mathrm{Hom}_{\mathbb{K}}(\mathbb{K}^n, \mathbb{K}^m)$. Wir wollen zeigen, dass $\Psi(A+B) = \Psi(A)+\Psi(B)$. Es sei dazu ein $x \in \mathbb{K}^n$ gegeben. Es gilt:

$$\begin{aligned}(\Psi(A+B))(x) &= \varphi_{A+B}(x)\\ &= (A+B)x\\ &= Ax + Bx\\ &= \varphi_A(x) + \varphi_B(x)\\ &= (\varphi_A + \varphi_B)(x)\\ &= (\Psi(A) + \Psi(B))(x).\end{aligned}$$

Der einzige Schritt in dieser Rechnung, in der nicht nur Definitionen eingesetzt wurden, war die Verwendung des Distributivgesetzes $(A+B)x = Ax + Bx$ beim Matrixprodukt.

Wir müssen nun nur noch zeigen, dass für eine Matrix $A \in \mathbb{K}^{m\times n}$ und ein $\lambda \in \mathbb{K}$ gilt: $\Psi(\lambda A) = \lambda\Psi(A)$. Es sei dazu wieder ein $x \in \mathbb{K}^n$ gegeben. Dann gilt:

$$\begin{aligned}(\Psi(\lambda A))(x) &= \varphi_{\lambda A}(x)\\ &= (\lambda A)x\\ &= \lambda(Ax)\\ &= \lambda\varphi_A(x)\\ &= (\lambda\varphi_A)(x)\\ &= (\lambda\Psi(A))(x).\end{aligned}$$

Dies beendet den Beweis.

(f)

Wir müssen zwei Implikationen beweisen:

„$\Longleftarrow$“: Angenommen, $m = n$ und $A \in \mathbb{K}^{n\times n}$ ist eine invertierbare Matrix. Dann ergibt sich mit (c) und (d) direkt, dass

$$\varphi_A \circ \varphi_{A^{-1}} = \mathrm{id}_{\mathbb{K}^n} = \varphi_{A^{-1}} \circ \varphi_A.$$

Also ist $\varphi_{A^{-1}}$ eine Umkehrabbildung[9] von φ_A und somit bijektiv, also ein Vektorraumisomorphismus.

„$\Longrightarrow$“: Angenommen, $\varphi_A : \mathbb{K}^n \to \mathbb{K}^m$ ist ein Vektorraumisomorphismus. Dann heißt dies, dass $\mathbb{K}^m \cong \mathbb{K}^n$ und nach Satz 4.2.9(c) geht dies nur, wenn $m = n$ ist. Also ist A eine quadratische Matrix. Da $\varphi_A : \mathbb{K}^n \to \mathbb{K}^n$ bijektiv ist, gibt es eine Umkehrabbildung $(\varphi_A)^{-1} : \mathbb{K}^n \to \mathbb{K}^n$ mit

$$(\varphi_A)^{-1} \circ \varphi_A = \mathrm{id}_{\mathbb{K}^n} = \varphi_A \circ (\varphi_A)^{-1}.$$

[9] Siehe Definition 1.3.11

Da φ_A linear ist, ist $(\varphi_A)^{-1}$ ebenfalls linear und somit gibt es nach Teil (b) angewendet auf φ^{-1} eine Matrix $B \in \mathbb{K}^{n\times n}$ mit $(\varphi_A)^{-1} = \varphi_B$. Eingesetzt in obige Gleichung erhalten wir:

$$\varphi_B \circ \varphi_A = \mathrm{id}_{\mathbb{K}^n} = \varphi_A \circ \varphi_B,$$

was sich mit (c) und (d) vereinfacht zu:

$$\varphi_{BA} = \varphi_{\mathbb{1}_n} = \varphi_{AB}.$$

Aus der Eindeutigkeit der Darstellungsmatrix folgt schließlich:

$$BA = \mathbb{1}_n = AB.$$

Also ist A invertierbar im Ring $\mathbb{K}^{n\times n}$. □

Wir haben also gesehen, eine lineare Abbildung $\varphi : \mathbb{K}^n \to \mathbb{K}^n$ ist genau dann ein Isomorphismus von Vektorräumen, wenn $\varphi = \varphi_A$ mit einer *invertierbaren Matrix*, also einer Matrix, die als Element im Ring $(\mathbb{K}^{n\times n}, +, \cdot)$ multiplikativ invertierbar ist. Das ist Grund genug, um uns ein bisschen genauer mit invertierbare Matrizen beschäftigen:

Definition 4.3.2 (Invertierbare Matrizen)**.**
Eine quadratische Matrix $A \in \mathbb{K}^{n\times n}$ heißt *invertierbare Matrix*, wenn sie als Element im Ring $(\mathbb{K}^{n\times n}, +, \cdot)$ bezüglich der Matrixmultiplikation invertierbar ist, d.h. wenn eine Matrix $B \in \mathbb{K}^{n\times n}$ existiert mit

$$AB = \mathbb{1}_n = BA.$$

Die Gruppe aller invertierbaren $(n \times n)$-Matrizen mit reellen Einträgen wird *allgemeine lineare Gruppe* genannt und mit $\mathrm{GL}(n, \mathbb{R})$ bezeichnet.

Beispiel 4.3.3. (a) Die Einheitsmatrix $\mathbb{1}_n$ selbst ist invertierbar: $\mathbb{1}_n \in \mathrm{GL}(n, \mathbb{R})$. Allgemeiner ist das neutrale Element e in jedem Monoid $(S, *)$ invertierbar.

(b) Wenn A eine Nullzeile hat, ist A auf jeden Fall nicht invertierbar. Denn angenommen A hätte eine Inverse B, dann würde gelten

$$AB = \mathbb{1}_n = BA.$$

Da aber A eine Nullzeile hat, hat auch AB eine Nullzeile und das widerspricht $AB = \mathbb{1}_n$.

(c) Die Matrizen aus Lemma 2.5.4, die den elementaren Zeilenumformungen entsprechen, sind alle invertierbar:
Wenn G eine Matrix vom Typ (G1) ist, die das μ-fache der j-ten Zeile auf die i-te Zeile addiert, dann ist die Inverse von G gegeben durch die Matrix, die das $(-\mu)$-fache der j-ten Zeile auf die i-te Zeile addiert.
Wenn G eine Matrix vom Typ (G3) ist, die eine Zeile mit $\lambda \neq 0$ multipliziert, dann ist die Inverse gegeben durch die Matrix vom Typ (G3), die diese Zeile mit $\frac{1}{\lambda}$ multipliziert.
Matrizen vom Typ (G2) sind selbstinvers.

Wir haben in Beispiel 3.1.11 gesehen, dass es im Allgemeinen für Elemente in einem Monoid nicht ausreicht, ein Rechtsinverses zu haben, um (beidseitig) invertierbar zu sein. Es stellt sich heraus, dass dieses Problem, dass für allgemeine Monoide (und sogar für Operatoren auf unendlich dimensionalen Räumen) existiert, für Matrizen nicht auftritt: Für eine quadratische Matrix A gilt: Wenn es ein B gibt mit $AB = \mathbb{1}_n$, dann ist A invertierbar. Dies werden wir nun beweisen:

Satz 4.3.4 (Erste Charakterisierung der Invertierbarkeit von Matrizen)**.**
Es sei $n \in \mathbb{N}$ und $\mathbb{K}$ ein Körper. Für eine Matrix $A \in \mathbb{K}^{n\times n}$ sind äquivalent:

(i) A *ist invertierbar.*

(ii) $A^\top$ *ist invertierbar.*

(iii) $\exists B \in \mathbb{K}^{n\times n} : BA = \mathbb{1}_n$.

(iv) $\exists B \in \mathbb{K}^{n\times n} : AB = \mathbb{1}_n$.

(v) $\ker A = \{0\}$.

(vi) $\operatorname{rg} A = n$.

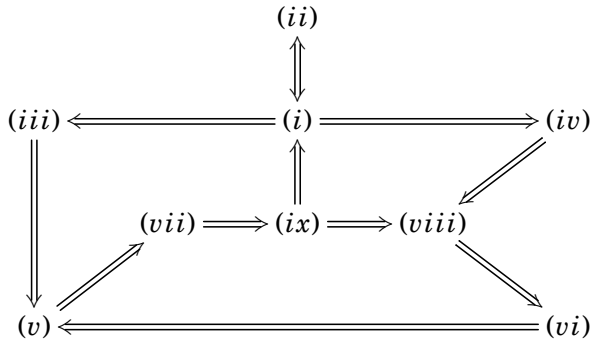
(vii) $\varphi_A : \mathbb{K}^n \to \mathbb{K}^n, \quad x \mapsto Ax$ *ist injektiv.*

(viii) $\varphi_A : \mathbb{K}^n \to \mathbb{K}^n, \quad x \mapsto Ax$ *ist surjektiv.*

(ix) $\varphi_A : \mathbb{K}^n \to \mathbb{K}^n, \quad x \mapsto Ax$ *ist bijektiv.*

Später, wenn wir Determinanten und Eigenwerte besprochen haben, werden wir noch weitere äquivalente Bedingungen zur Invertierbarkeit hinzufügen (siehe z.B. Satz 5.3.10).

Beweis. Wieder gibt es einige Implikationen zu beweisen. Der grundsätzliche Plan dieses Beweises sieht so aus:



„(i) $\Longleftrightarrow$ (ii)“:
Es sei A invertierbar, d.h. es gibt ein $B \in \mathbb{K}^{n\times n}$ mit

$$AB = \mathbb{1}_n = BA.$$

Dann können wir diese Gleichung transponieren und erhalten:

$$(AB)^\top = \mathbb{1}_n^\top = (BA)^\top,$$

was sich (siehe Lemma 2.2.13) umformen lässt zu:

$$B^\top A^\top = \mathbb{1}_n = A^\top B^\top.$$

Also ist $A^\top$ invertierbar mit Inversem $B^\top$.

Da sich zweimaliges Transponieren wieder aufhebt, folgt die Rückimplikation direkt.

„(i) $\Longrightarrow$ (iii)“:
Wenn A ein (beidseitiges) Inverses hat, dann insbesondere auch ein Linksinverses.
„(iii) $\Longrightarrow$ (v)“:
Nehmen wir an, $BA = \mathbb{1}_n$ und $x \in \ker A$. Dann gilt:

$$x = \mathbb{1}_n x = (BA)x = B(Ax) = B0 = 0.$$

„(v) $\Longrightarrow$ (vii)“:
Dies folgt aus Lemma 4.1.15.
„(vii) $\Longrightarrow$ (ix)“:
Es sei $V := \text{Bild}(A) = \text{Bild}(\varphi_A) = \varphi_A(\mathbb{K}^n)$. Dann ist V ein Untervektorraum von $\mathbb{K}^n$. Da die Koeinschränkung

$$\mathbb{K}^n \to V, \quad x \mapsto Ax$$

bijektiv ist, ist sie ein Vektorraumisomorphismus zwischen $\mathbb{K}^n$ und V. Somit gilt $\dim_{\mathbb{K}}(V) = \dim_{\mathbb{K}}(\mathbb{K}^n) = n$. Also hat $\mathbb{K}^n$ einen Untervektorraum mit Dimension n. Dies ist nur möglich, wenn $V = \mathbb{K}^n$, weil ein echter Untervektorraum eines endlich dimensionalen Vektorraums immer eine echt kleinere Dimension hat (siehe Lemma 4.2.10 oder Lemma 2.3.26). Also ist $\text{Bild}(\varphi_A) = \mathbb{K}^n$ und somit ist φ_A bijektiv.
„(ix) $\Longrightarrow$ (i)“:
Dies ist Satz 4.3.1(f)
„(i) $\Longrightarrow$ (iv)“:
Wenn A ein (beidseitiges) Inverses hat, dann insbesondere auch ein Rechtsinverses.
„(iv) $\Longrightarrow$ (viii)“:
Nehmen wir an, $AB = \mathbb{1}_n$ und $y \in \mathbb{K}^n$. Wir setzen $x := By \in \mathbb{K}^n$. Dann gilt:

$$y = \mathbb{1}_n y = (AB)y = A(By) = Ax \in \text{Bild}(A) = \text{Bild}(\varphi_A).$$

Also ist jedes Element $y \in \mathbb{K}^n$ im Bild von φ_A. Also ist φ_A surjektiv.
„(viii) $\Longrightarrow$ (vi)“:
Der Rang von A berechnet sich wie folgt:

$$\text{rg}A = \dim_{\mathbb{K}}(\text{Bild}(A)) = \dim_{\mathbb{K}}(\mathbb{K}^n) = n.$$

„(vi) $\Longrightarrow$ (v)“:
Die Dimensionsformel für Matrizen (Satz 2.5.9) besagt: $\dim_{\mathbb{K}}(\ker A) + \text{rg}A = n$. Da A aber Rang n hat, folgt somit $\dim_{\mathbb{K}}(\ker A) = 0$ und somit muss $\ker A = \{0\}$ gelten. □

Beispiel 4.3.5. Es sei $\mathbb{K}$ ein Körper und

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbb{K}^{2\times 2}.$$

Wir setzen

$$B = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \in \mathbb{K}^{2\times 2}.$$

Wenn wir nun AB berechnen, erhalten wir

$$AB = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \begin{pmatrix} ad-bc & 0 \\ 0 & ad-bc \end{pmatrix} = (ad-bc)\mathbb{1}_2.$$

Wenn also $ad - bc \neq 0$ ist, dann können wir dies umformen zu:

$$A \frac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \mathbb{1}_2$$

und wir sehen (mit Satz 4.3.4), dass A invertierbar ist und sich die Inverse schreiben lässt als

$$A^{-1} = \frac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Wenn allerdings $ad - bc = 0$ ist, dann bedeutet dies, dass die Spalten $\begin{pmatrix} a \\ c \end{pmatrix}$ und $\begin{pmatrix} b \\ d \end{pmatrix}$ linear abhängig sind und damit, dass A nicht Rang 2 hat und mit Satz 4.3.4 bedeutet dies, dass A nicht invertierbar ist.

Die Zahl $ad - bc$, die entscheidet, ob die Matrix invertierbar ist oder nicht, nennt man die *Determinante* von A. Determinanten gibt es nicht nur für (2×2)-Matrizen, sondern allgemein für quadratische Matrizen. Sie spielen auch nicht nur bei der Invertierbarkeit der Matrix eine Rolle, sondern haben auch eine wichtige geometrische Bedeutung. Wir werden uns in Kapitel 5.3 ausführlicher damit beschäftigen. Satz 5.3.17 bildet eine direkte Verallgemeinerung der Formel $A^{-1} = \frac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$, die allerdings für praktische Berechnungen ziemlich nutzlos ist.

Bemerkung 4.3.6 (Wie berechnet man Inverse?)**.** Die in Beispiel 4.3.5 vorgestellte Formel zur Berechnung Inversen ist sehr hilfreich, aber leider nur für (2×2)-Matrizen gültig. Es gibt für größere Matrizen auch Formeln, allerdings sind diese so unhandlich, dass man mit ihnen nicht sinnvoll arbeiten kann. Wir werden deshalb nun ein anderes Verfahren zum Invertieren von Matrizen kennenlernen, das uns – auch ohne Determinanten – auch sagt, ob eine Matrix invertierbar ist oder nicht und vielleicht wenig überraschend ist dies der Gauß-Algorithmus:

Bis jetzt haben wir den Gauß-Algorithmus verwendet, um lineare Gleichungssysteme der Form

$$Ax = b$$

zu lösen, wobei x und b Spaltenvektoren waren. Mit exakt demselben Verfahren kann man aber auch Gleichungen der Form

$$AX = B$$

lösen, wenn X und B Matrizen sind. Dies entspricht im Wesentlichen dem simultanen Lösen von mehreren linearen Gleichungssystemen mit mehreren rechten Seiten gleichzeitig.

Wir wissen nach Satz 4.3.4, dass A genau dann invertierbar ist, wenn es eine Matrix $X \in \mathbb{K}^{n \times n}$ gibt mit

$$AX = \mathbb{1}_n.$$

Diese Matrix X ist dann die Inverse A^{-1}. Wir wenden nun Zeilenumformungen auf A und die Matrix auf der rechten Seite an, bis wir nach endlich vielen Schritten entweder eine Nullzeile erhalten (Dann war A nicht invertierbar und wir können aufhören) oder die Matrix in erweiterte Zeilenstufenform ohne Nullzeile gebracht haben. Allerdings ist eine $(n \times n)$-Matrix mit vollem Rang in erweiterter Zeilenstufenform notwendigerweise die Einheitsmatrix:

$$\mathbb{1}_n X = C$$

und somit ist $X = C$ und wir sind fertig.

Beispiel 4.3.7. (a) Wir wollen die Inverse von

$$A = \begin{pmatrix} 3 & -6 & 5 \\ 0 & 3 & -4 \\ 3 & -6 & 6 \end{pmatrix} \in \mathbb{Q}^{3\times 3}$$

berechnen.

3	−6	5	1	0	0	$\vert\cdot(-1)$ auf Zeile 3	$(G1)$
0	3	−4	0	1	0		
3	−6	6	0	0	1		
3	−6	5	1	0	0		
0	3	−4	0	1	0		
0	0	1	−1	0	1	$\vert\cdot 4$ auf Zeile 2	$(G1)$
3	−6	5	1	0	0		
0	3	0	−4	1	4		
0	0	1	−1	0	1	$\vert\cdot(-5)$ auf Zeile 1	$(G1)$
3	−6	0	6	0	−5		
0	3	0	−4	1	4	$\vert\cdot 2$ auf Zeile 1	$(G1)$
0	0	1	−1	0	1		
3	0	0	−2	2	3	$\vert\cdot\frac{1}{3}$	$(G3)$
0	3	0	−4	1	4	$\vert\cdot\frac{1}{3}$	$(G3)$
0	0	1	−1	0	1		
1	0	0	$-\frac{2}{3}$	$\frac{2}{3}$	1		
0	1	0	$-\frac{4}{3}$	$\frac{1}{3}$	$\frac{4}{3}$		
0	0	1	−1	0	1		

Also ist A invertierbar und es gilt:

$$A^{-1} = \begin{pmatrix} -\frac{2}{3} & \frac{2}{3} & 1 \\ -\frac{4}{3} & \frac{1}{3} & \frac{4}{3} \\ -1 & 0 & 1 \end{pmatrix} = \frac{1}{3}\begin{pmatrix} -2 & 2 & 3 \\ -4 & 1 & 4 \\ -3 & 0 & 3 \end{pmatrix}.$$

(b) Es sei $\alpha \in \mathbb{Z}/5\mathbb{Z}$ ein beliebiges Element. Wir wollen über dem Körper $\mathbb{Z}/5\mathbb{Z}$ die Matrix

$$A = \begin{pmatrix} [0]_5 & [2]_5 & [2]_5 \\ [3]_5(\alpha + [1]_5) & [2]_5 & [0]_5 \\ [4]_5(\alpha + [1]_5) & [0]_5 & [3]_5 \end{pmatrix} \in \mathbb{Z}/5\mathbb{Z}^{3\times 3}$$

invertieren, falls sie invertierbar ist.

Hinweis zur Notation: Wir werden in der Rechnung die Restklassen $[\mathrm{cot}]_5$ weglassen, dies soll aber nicht darüber hinwegtäuschen, dass alle Zahlen für ihre jeweiligen Restklassen modulo 5 stehen.

0	2	2	1	0	0	| mit Zeile 3 vertauschen	$(G2)$
$3(\alpha+1)$	2	0	0	1	0		
$4(\alpha+1)$	0	3	0	0	1		
$4(\alpha+1)$	0	3	0	0	1	$\cdot 4$	$(G3)$
$3(\alpha+1)$	2	0	0	1	0		
0	2	2	1	0	0		
$\alpha+1$	0	2	0	0	4	$\cdot 2$ auf Zeile 2	$(G1)$
$3(\alpha+1)$	2	0	0	1	0		
0	2	2	1	0	0		
$\alpha+1$	0	2	0	0	4		
0	2	4	0	1	3	$\cdot 3$	$(G3)$
0	2	2	1	0	0		
$\alpha+1$	0	2	0	0	4		
0	1	2	0	3	4	$\cdot 3$ auf Zeile 3	$(G1)$
0	2	2	1	0	0		
$\alpha+1$	0	2	0	0	4		
0	1	2	0	3	4		
0	0	3	1	4	2	$\cdot 2$	$(G3)$
$\alpha+1$	0	2	0	0	4		
0	1	2	0	3	4		
0	0	1	2	3	4	$\cdot 3$ auf Zeile 2	$(G1)$
$\alpha+1$	0	2	0	0	4		
0	1	0	1	2	1		
0	0	1	2	3	4	$\cdot 3$ auf Zeile 1	$(G1)$
$\alpha+1$	0	0	1	4	1		
0	1	0	1	2	1		
0	0	1	2	3	4		

Wenn nun $\alpha+[1]_5 \neq [0]_5$ in $\mathbb{Z}/5\mathbb{Z}$, dann können wir die erste Zeile mit dem multiplikativen Inversen von $(\alpha+[1]_5)$ multiplizieren und erhalten

$$A^{-1} = \begin{pmatrix} (\alpha+[1]_5)^{-1} & [4]_5(\alpha+[1]_5)^{-1} & (\alpha+[1]_5)^{-1} \\ [1]_5 & [2]_5 & [1]_5 \\ [2]_5 & [3]_5 & [4]_5 \end{pmatrix}.$$

Falls $\alpha+1=0$ in $\mathbb{Z}/5\mathbb{Z}$, dann erhalten wir eine Nullspalte, also hat A nicht Rang 3 und ist somit nicht invertierbar.

Wir erhalten also: Für $\alpha \in \mathbb{Z}/5\mathbb{Z} \setminus \{[4]_5\} = \{[0]_5,[1]_5,[2]_5,[3]_5\}$ ist A invertierbar mit der oben angegebenen Inversen. Für $\alpha = [4]_5$ ist A nicht invertierbar.

Nach diesem kleinen Einschub über invertierbare Matrizen zurück zum eigentlichen Thema: Der Darstellung einer linearen Abbildung mit Hilfe einer Matrix.

Bemerkung 4.3.8. Wir wissen nun: Wenn wir lineare Abbildungen von $\mathbb{K}^n$ nach $\mathbb{K}^m$ untersuchen wollen, so reicht es, Matrizen zu untersuchen, weil jede Eigenschaft der linearen Abbildung in der dazugehörigen Matrix kodiert ist. Dies ist sehr wichtig und hilfreich, weil Matrizen häufig *konkretere* Objekte sind, mit denen man Rechnungen (z.B. den Gauß-Algorithmus) durchführen kann. Gewonnene Erkenntnisse kann man dann auf die lineare Abbildung übertragen.

Satz 4.3.1 ist somit ein wichtiges Werkzeug zum Studium linearer Abbildungen, aber aus zwei Gründen noch nicht ganz zufriedenstellend:

Zum Einen ist nicht jeder Vektorraum von der Form $\mathbb{K}^n$. Es gibt zum Beispiel Abbildungen, bei denen Definitions- oder Zielbereich Untervektorräume von $\mathbb{K}^n$ sind (etwa Geraden oder Ebenen im $\mathbb{R}^n$). Auch gibt es lineare Abbildungen, bei denen Definitions- oder Zielbereiche Räume von Matrizen (oder selbst Räume von Abbildungen) sind. Glücklicherweise haben wir in Satz 4.2.9 gesehen, dass jeder endlich dimensionale $\mathbb{K}$-Vektorraum isomorph ist zu $\mathbb{K}^n$ – allerdings ist so ein Isomorphismus nicht kanonisch gegeben, sondern hängt von der Wahl einer Basis ab.

Zum Anderen gibt es auch noch Verbesserungsbedarf bei linearen Abbildungen von $\mathbb{K}^n$ nach $\mathbb{K}^m$. Oft ist die Darstellungsmatrix, die zu einer linearen Abbildung gehört, viel zu kompliziert, um daran etwas ablesen oder damit rechnen zu können. Dies liegt daran, dass in der Konstruktion der Matrix die Standardbasis $(e_1,\dots,e_n)$ von $\mathbb{K}^n$ benutzt wurde und diese oft nicht die „beste Wahl“ ist. Oft ist eine andere Basis, die sich an der zu untersuchenden linearen Abbildung orientiert, besser geeignet.

Beide Probleme werden wir dadurch lösen, Abbildungsmatrizen von linearen Abbildungen zwischen endlich dimensionalen Vektorräumen bezüglich gegebener Basen in Definitions- und Zielbereich zu definieren. Allerdings ist dafür notwendig, auf einer Basis (die bisher einfach nur eine Menge ohne Ordnung war) eine Ordnung einzuführen.

Definition 4.3.9 (Geordnete Basis und Koordinatenvektoren)**.**
Es sei $\mathbb{K}$ ein Körper und V ein endlich dimensionaler Vektorraum über $\mathbb{K}$. Gegeben sei eine natürliche Zahl $n \in \mathbb{N}_0$.

(a) Eine *geordnete Basis* von V ist ein n-Tupel $\mathsf{B} = (b_1,\dots,b_n)$, bestehend aus n verschiedenen Vektoren aus V, sodass $\{b_1,\dots,b_n\}$ eine Basis für V ist.

(b) Es sei $\mathsf{B} = (b_1,\dots,b_n)$ eine geordnete Basis von V. Dann ist – nach Satz 4.2.8 – die lineare Abbildung

$$\mathbb{K}^n \to V, \quad \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto \sum_{k=1}^{n} x_k b_k$$

ein Isomorphismus von Vektorräumen. Die Umkehrabbildung bezeichnen wir mit

$$(\cdot)_{\mathsf{B}} : V \to \mathbb{K}^n, \quad v = \sum_{k=1}^{n} x_k b_k \mapsto (v)_{\mathsf{B}} := \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

Der Spaltenvektor $(v)_{\mathsf{B}} \in \mathbb{K}^n$ heißt der Koordinatenvektor von v bezüglich der geordneten Basis B.

Aus Gründen der Vollständigkeit ist hier $n = 0$ erlaubt – ein 0-Tupel besteht aus 0 Vektoren (die leere Liste sozusagen), und der einzige Vektorraum, der ein 0-Tupel als geordnete Basis hat ist der triviale Vektorraum $V = \{0_V\} \cong_{\mathbb{K}} \mathbb{K}^0$. Insofern ist $n = 0$ zugelassen, aber uninteressant.

Beispiel 4.3.10. (a) Es sei $\mathbb{K}$ ein Körper, $n \in \mathbb{N}$ und $V := \mathbb{K}^n$. Die *geordnete Standardbasis* von $\mathbb{K}^n$ ist

$$\mathsf{E} := (e_1,\dots,e_n),$$

wobei $e_1,\dots,e_n$ die Standardbasisvektoren sind.

Für jeden Vektor $x \in \mathbb{K}^n$ gilt: $(x)_{\mathsf{E}} = x$.

(b) Es sei $V := \mathbb{R}^{2\times 2}$ mit der geordneten Basis:

$$\mathsf{B} := \left(\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}, \begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}, \begin{pmatrix}0 & 0\\ 1 & 0\end{pmatrix}, \begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix}\right).$$

Dann hat die Matrix $A = \begin{pmatrix}a & b\\ c & d\end{pmatrix} \in \mathbb{R}^{2\times 2}$ bezüglich der geordneten Basis B den folgenden Koordinatenvektor:

$$\left(\begin{pmatrix}a & b\\ c & d\end{pmatrix}\right)_{\mathsf{B}} = \begin{pmatrix}a\\ b\\ c\\ d\end{pmatrix}.$$

Wenn wir stattdessen die folgende geordnete Basis betrachten:

$$\mathsf{C} := \left(\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}, \begin{pmatrix}0 & 0\\ 1 & 0\end{pmatrix}, \begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}, \begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix}\right),$$

so gilt:

$$\left(\begin{pmatrix}a & b\\ c & d\end{pmatrix}\right)_{\mathsf{C}} = \begin{pmatrix}a\\ c\\ b\\ d\end{pmatrix}.$$

Wenn wir wissen wollen, wir der Koordinatenvektor von $\begin{pmatrix}a & b\\ c & d\end{pmatrix}$ bezüglich der folgenden geordneten Basis

$$\mathsf{D} := \left(\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}, \begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}, \begin{pmatrix}1 & 0\\ 1 & 1\end{pmatrix}, \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\right)$$

aussieht, so müssen wir den Ansatz machen:

$$x_1\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix} + x_2\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix} + x_3\begin{pmatrix}1 & 0\\ 1 & 1\end{pmatrix} + x_4\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} = \begin{pmatrix}a & b\\ c & d\end{pmatrix}$$

und dann die Koordinaten x_1, x_2, x_3, x_4 bestimmen. Die linke Seite der Gleichung lässt sich zusammenfassen zu einer Matrix, was dann folgende Matrix-Gleichung ergibt:

$$\begin{pmatrix}x_1 + x_2 + x_3 + x_4 & x_2\\ x_3 & -x_1 + x_2 + x_3 + x_4\end{pmatrix} = \begin{pmatrix}a & b\\ c & d\end{pmatrix}$$

Dies entspricht dem inhomogenen linearen Gleichungssystem:

$$\begin{aligned} x_1 \;+x_2\; +x_3\; +x_4 &= a\\ x_2 &= b\\ x_3 &= c\\ -x_1 \;+x_2\; +x_3\; +x_4 &= d, \end{aligned}$$

was (Gauß-Algorithmus) die eindeutige Lösung

$$\begin{pmatrix}x_1\\ x_2\\ x_3\\ x_4\end{pmatrix} = \begin{pmatrix}a/2 - d/2\\ b\\ c\\ a/2 - b - c + d/2\end{pmatrix}$$

besitzt. Zusammengefasst kann man also sagen, dass die Matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ bezüglich der Basis D den Koordinatenvektor

$$\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right)_{\mathsf{D}} = \begin{pmatrix} a/2 - d/2 \\ b \\ c \\ a/2 - b - c + d/2 \end{pmatrix}$$

hat.

(c) Der Körper $\mathbb{C}$ ist eine Körpererweiterung von $\mathbb{R}$ und lässt sich somit als $\mathbb{R}$-Vektorraum auffassen. Wir betrachten die geordnete Basis:

$$\mathsf{B} := (1, \mathrm{i}).$$

Dann gilt für jede komplexe Zahl $z = a + \mathrm{i}b$, dass $(z)_{\mathsf{B}} = \begin{pmatrix} a \\ b \end{pmatrix}$.

(d) Es sei $\varphi : \mathbb{F}_2^4 \to \mathbb{F}_2$, $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \mapsto x_1 + x_2 + x_3 + x_4$ die $\mathbb{F}_2$-lineare Abbildung zur Matrix

$$A := \begin{pmatrix} 1 & 1 & 1 & 1 \end{pmatrix} \in \mathbb{F}_2^{1\times 4}.$$

Wir setzen $U := \ker \varphi = \ker A \subseteq \mathbb{F}_2^4$. Das folgende ist eine geordnete Basis von U:

$$\mathsf{B} := \left(\begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}\right).$$

Der Vektor $v := \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} \in U$ hat bezüglich B den Koordinatenvektor

$$(v)_{\mathsf{B}} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix},$$

weil

$$v = 0 \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} + 1 \cdot \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + 1 \cdot \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}.$$

(e) Der Raum $\mathbb{R}^{\mathbb{R}}$ aller Funktionen von $\mathbb{R}$ nach $\mathbb{R}$ ist ein unendlich dimensionaler $\mathbb{R}$-Vektorraum. Die Funktionen

$$f_1(x) = (x+1)^2; \quad f_2(x) = (x+2)^2; \quad f_3(x) = (x+3)^2$$

sind linear unabhängig (nachrechnen!). Das bedeutet also, dass

$$\mathsf{B} := (f_1, f_2, f_3)$$

eine geordnete Basis des Raums $V = \mathrm{LH}_{\mathbb{R}}(f_1, f_2, f_3)$ ist. Die Funktion $f_4 : \mathbb{R} \to \mathbb{R}$ mit

$$f_4(x) = (x+4)^2$$

ist ein Element in V, weil $f_4 = 1 \cdot f_1 - 3f_2 + 3f_3$ (nachrechnen!). Also können wir den Koordinatenvektor von f_4 bezüglich der Basis B schreiben als:

$$(f_4)_{\mathsf{B}} = \begin{pmatrix} 1 \\ -3 \\ 3 \end{pmatrix}.$$

Die Funktion $\mathrm{id}_{\mathbb{R}} : \mathbb{R} \to \mathbb{R}, \quad x \mapsto x$ gehört auch zum Raum V, weil $\mathrm{id}_{\mathbb{R}} = -\frac{5}{4}f_1 + 2f_2 - \frac{3}{4}f_3$. Es gilt somit:

$$(\mathrm{id}_{\mathbb{R}})_{\mathsf{B}} = \begin{pmatrix} -\frac{5}{4} \\ 2 \\ -\frac{3}{4} \end{pmatrix} = \frac{1}{4}\begin{pmatrix} -5 \\ 8 \\ -3 \end{pmatrix}.$$

(f) Es sei $\mathbb{K}$ ein Körper und $V := \mathbb{K}^3$. Das folgende ist eine geordnete Basis von $\mathbb{K}^3$:

$$\mathsf{B} := \left(\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right).$$

Der Vektor $\begin{pmatrix} a \\ b \\ c \end{pmatrix} \in \mathbb{K}^3$ hat bezüglich B die Koordinaten:

$$\left(\begin{pmatrix} a \\ b \\ c \end{pmatrix} \right)_{\mathsf{B}} = \begin{pmatrix} -b+c \\ -a+b \\ a \end{pmatrix},$$

weil

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} = (-b+c)\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} + (-a+b)\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} + a\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Diese Koordinaten findet man entweder durch Ausprobieren oder durch das Lösen des dazugehörigen lineares Gleichungssystems.

Wir wissen nun also, dass ein endlich dimensionaler Vektorraum mit einer geordneten Basis B immer isomorph zu einem Raum der Form $\mathbb{K}^n$ ist und nach Satz 4.3.1 lassen sich lineare Abbildungen zwischen solchen Räumen immer als Linksmultiplikation mit einer Matrix schreiben. Wenn wir diese beiden Ideen kombinieren, erhalten wir den folgenden für die lineare Algebra fundamentalen Satz:

Satz 4.3.11 (Darstellungsmatrix einer linearen Abbildung)**.**
Es sei $\mathbb{K}$ ein Körper und V, W endlich dimensionale Vektorräume über $\mathbb{K}$ mit $n = \dim_{\mathbb{K}}(V) \in \mathbb{N}$ und $m = \dim_{\mathbb{K}}(W) \in \mathbb{N}$. Wir nehmen an, es seien geordnete Basen B_V von V und B_W von W gegeben.

(a) *Es sei* $A \in \mathbb{K}^{m\times n}$ *eine Matrix und* $\varphi_A : \mathbb{K}^n \to \mathbb{K}^m, \quad x \mapsto Ax$ *die Linksmultiplikation mit* A*. Es bezeichne* $(\cdot)_{\mathsf{B}_V} : V \to \mathbb{K}^n$ *und* $(\cdot)_{\mathsf{B}_W} : W \to \mathbb{K}^m$ *die Vektorraumisomorphismen aus Definition 4.3.9.*

Dann ist das folgende eine lineare Abbildung zwischen V *und* W*:*

$$\varphi : V \to W, \quad v \mapsto ((\cdot)_{\mathsf{B}_W})^{-1}(A\,(v)_{\mathsf{B}_V}) = ((\cdot)_{\mathsf{B}_W})^{-1} \circ \varphi_A \circ ((\cdot)_{\mathsf{B}_V})(v).$$

Anders formuliert: Das folgende Diagramm kommutiert[10]*:*

$$\begin{array}{ccc} V & \xrightarrow{\varphi} & W \\ {\scriptstyle (\cdot)_{\mathsf{B}_V}}\downarrow & & \downarrow{\scriptstyle (\cdot)_{\mathsf{B}_W}} \\ \mathbb{K}^n & \xrightarrow[\varphi_A]{} & \mathbb{K}^m. \end{array}$$

In den Spalten von A *stehen die Bilder der Basisvektoren aus* B_V*, dargestellt bezüglich der Basis* B_W*; genauer: sei* b_j *der* j*-te Basisvektor der geordneten Basis* B*. Dann steht in der* j*-ten Spalte von* A *der Spaltenvektor* $\big(\varphi(b_j)\big)_{\mathsf{B}_W}$.

(b) *Für jede* $\mathbb{K}$*-lineare Abbildung* $\varphi : V \to W$ *gibt es eine eindeutige Matrix* $A \in \mathbb{K}^{m\times n}$ *mit*

$$\forall v \in V : \quad \big(\varphi(v)\big)_{\mathsf{B}_W} = A\,(v)_{\mathsf{B}_V}.$$

Die eindeutige Matrix $A \in \mathbb{K}^{m\times n}$ *wird* Darstellungsmatrix von φ bezüglich der geordneten Basen B_V und B_W *genannt und mit* $\mathrm{M}_{\mathsf{B}_W,\mathsf{B}_V}(\varphi)$ *bezeichnet*[11]*. Es gilt also:*

$$\forall v \in V : \quad \big(\varphi(v)\big)_{\mathsf{B}_W} = \mathrm{M}_{\mathsf{B}_W,\mathsf{B}_V}(\varphi)\,(v)_{\mathsf{B}_V}.$$

Im Falle, dass $V = \mathbb{K}^n$*,* $W = \mathbb{K}^m$ *und* B_V *und* B_W *die Standardbasen sind, ist dies genau die in Satz 4.3.1 eingeführte Matrix.*

(c) *Gegeben sei ein weiterer* $\mathbb{K}$*-Vektorraum* U *mit einer* p*-elementigen geordneten Basis* B_U*. Dann gilt für lineare Abbildungen* $\varphi : V \to W$ *und* $\psi : U \to V$*:*

$$\mathrm{M}_{\mathsf{B}_W,\mathsf{B}_V}(\varphi)\,\mathrm{M}_{\mathsf{B}_V,\mathsf{B}_U}(\psi) = \mathrm{M}_{\mathsf{B}_W,\mathsf{B}_U}(\varphi \circ \psi),$$

das heißt: Multiplizieren von Matrizen entspricht dem Verketten von Abbildungen.

(d) *Die Einheitsmatrix* $\mathbb{1}_n \in \mathbb{K}^{n\times n}$ *entspricht der Identitätsabbildung:*

$$\mathrm{M}_{\mathsf{B}_V,\mathsf{B}_V}(\mathrm{id}_V) = \mathbb{1}_n.$$

(e) *Die Abbildung*

$$\mathrm{Hom}_{\mathbb{K}}(V,W) \to \mathbb{K}^{m\times n}, \quad \varphi \mapsto \mathrm{M}_{\mathsf{B}_W,\mathsf{B}_V}(\varphi),$$

die jede lineare Abbildung von V *nach* W *auf ihre Darstellungsmatrix bezüglich der geordneten Basen* B_V *und* B_W *abbildet, ist ein Isomorphismus von* $\mathbb{K}$*-Vektorräumen.*

[10] Das Konzept von kommutativen Diagrammen wurde in 1.3.6 eingeführt.

[11] Diese Notation ist überhaupt nicht einheitlich in der Literatur. Weder der Buchstabe „M“ noch die Reihenfolge, in der die Basen B_V und B_W angegeben werden.

(f) *Eine lineare Abbildung $\varphi : V \to W$ ist genau dann bijektiv (also ein Vektorraumisomorphismus), wenn die Matrix $\mathrm{M}_{\mathsf{B}_W,\mathsf{B}_V}(\varphi)$ invertierbar ist (insbesondere ist dafür notwendig, dass $m = n$ ist). In diesem Fall gilt dann:*

$$\left(\mathrm{M}_{\mathsf{B}_W,\mathsf{B}_V}(\varphi)\right)^{-1} = \mathrm{M}_{\mathsf{B}_V,\mathsf{B}_W}\left(\varphi^{-1}\right).$$

(g) *Die Menge aller Vektorraumendomorphismen $\mathrm{End}_{\mathbb{K}}(V) := \mathrm{Hom}_{\mathbb{K}}(V,V)$ von V bildet mit $+$ und $\circ$ einen Ring. Die Abbildung*

$$\mathrm{End}_{\mathbb{K}}(V) \to \mathbb{K}^{n\times n}, \quad \varphi \mapsto \mathrm{M}_{\mathsf{B}_V,\mathsf{B}_V}(\varphi),$$

ist ein Isomorphismus von Ringen.

Beweis. (a)
Die Abbildung $\varphi = ((\cdot)_{\mathsf{B}_W})^{-1} \circ \varphi_A \circ ((\cdot)_{\mathsf{B}_V}) : V \to W$ ist als Verkettung von $\mathbb{K}$-linearen Abbildungen wieder $\mathbb{K}$-linear.

Es sei nun b_j der j-te Basisvektor aus der geordneten Basis B_V. Dann ist das Bild unter $\varphi : V \to W$ der Vektor $\varphi(b_j) \in W$. Diesen stellen wir nun zur Basis B_W dar und erhalten $\left(\varphi(b_j)\right)_{\mathsf{B}_W} \in \mathbb{K}^m$. Diesen Vektor kann man nun umschreiben zu:

$$\begin{aligned}
\left(\varphi(b_j)\right)_{\mathsf{B}_W} &= (\cdot)_{\mathsf{B}_W} \circ \varphi(b_j) \\
&= (\cdot)_{\mathsf{B}_W} \circ ((\cdot)_{\mathsf{B}_W})^{-1} \circ \varphi_A \circ ((\cdot)_{\mathsf{B}_V})(b_j) \\
&= \varphi_A \circ ((\cdot)_{\mathsf{B}_V})(b_j) \\
&= \varphi_A(\left(b_j\right)_{\mathsf{B}_V}) \\
&= \varphi_A(e_j) \\
&= Ae_j
\end{aligned}$$

und dies ist gerade die j-te Spalte der Matrix A. (b)
Nach Teil (a) wissen wir, dass die Matrix – falls sie existiert – eindeutig ist und dass in den Spalten der Matrix die Bilder der Basisvektoren stehen müssen. Setzen wir also:

$$A = \left(\left| \left(\varphi(b_1)\right)_{\mathsf{B}_W} \right| \cdots \left| \left(\varphi(b_n)\right)_{\mathsf{B}_W} \right| \right) \in \mathbb{K}^{m\times n}.$$

Es bleibt zu zeigen, dass

$$\forall v \in V : \quad \left(\varphi(v)\right)_{\mathsf{B}_W} = A\,(v)_{\mathsf{B}_V}.$$

Da beide Seiten dieser Gleichung linear in v sind, reicht es aus, diese Identität für Basisvektoren zu zeigen.

Es sei also $v = b_j$ einer der Basisvektoren. Dann ist

$$\left(\varphi(v)\right)_{\mathsf{B}_W} = \left(\varphi(b_j)\right)_{\mathsf{B}_W} = Ae_j = A\left(b_j\right)_{\mathsf{B}_V} = A\,(v)_{\mathsf{B}_V}.$$

(c)
Es sei $u \in U$. Dann gilt:

$$\begin{aligned}
\mathrm{M}_{\mathsf{B}_W,\mathsf{B}_U}(\varphi \circ \psi) \cdot (u)_{\mathsf{B}_U} &= \left((\varphi \circ \psi)(u)\right)_{\mathsf{B}_W} \\
&= \left(\varphi\left(\psi(u)\right)\right)_{\mathsf{B}_W} \\
&= \mathrm{M}_{\mathsf{B}_W,\mathsf{B}_V}(\varphi) \cdot \left(\psi(u)\right)_{\mathsf{B}_V} \\
&= \mathrm{M}_{\mathsf{B}_W,\mathsf{B}_V}(\varphi) \cdot \mathrm{M}_{\mathsf{B}_V,\mathsf{B}_U}(\psi) \cdot (u)_{\mathsf{B}_U}.
\end{aligned}$$

Also gilt:

$$\mathrm{M}_{\mathrm{B}_W,\mathrm{B}_U}\left(\varphi\circ\psi\right)=\mathrm{M}_{\mathrm{B}_W,\mathrm{B}_V}\left(\varphi\right)\cdot\mathrm{M}_{\mathrm{B}_V,\mathrm{B}_U}\left(\psi\right).$$

(d)
Es sei $v\in V$. Dann gilt:

$$\begin{aligned}\mathrm{M}_{\mathrm{B}_V,\mathrm{B}_V}\left(\mathrm{id}_V\right)\cdot(v)_{\mathrm{B}_V}&=\left(\mathrm{id}_V(v)\right)_{\mathrm{B}_V}\\&=(v)_{\mathrm{B}_V}\\&=\mathbb{1}_n\cdot(v)_{\mathrm{B}_V}\end{aligned}$$

Also gilt:

$$\mathrm{M}_{\mathrm{B}_V,\mathrm{B}_V}\left(\mathrm{id}_V\right)=\mathbb{1}_n.$$

(e)
Wir wollen zeigen, dass die folgende Abbildung ein Isomorphismus von $\mathbb{K}$-Vektorräumen ist:

$$\mathrm{Hom}_{\mathbb{K}}(V,W)\to\mathbb{K}^{m\times n},\quad\varphi\mapsto\mathrm{M}_{\mathrm{B}_W,\mathrm{B}_V}\left(\varphi\right).$$

Bijektivität wurde bereits in (a) und (b) gezeigt. Es bleibt die Linearität. Dazu müssen wir zeigen:

$$\forall\varphi,\psi\in\mathrm{Hom}_{\mathbb{K}}(V,W):\quad\mathrm{M}_{\mathrm{B}_W,\mathrm{B}_V}\left(\varphi+\psi\right)=\mathrm{M}_{\mathrm{B}_W,\mathrm{B}_V}\left(\varphi\right)+\mathrm{M}_{\mathrm{B}_W,\mathrm{B}_V}\left(\psi\right)$$

und

$$\forall\varphi\in\mathrm{Hom}_{\mathbb{K}}(V,W),\lambda\in\mathbb{K}:\quad\mathrm{M}_{\mathrm{B}_W,\mathrm{B}_V}\left(\lambda\varphi\right)=\lambda\mathrm{M}_{\mathrm{B}_W,\mathrm{B}_V}\left(\varphi\right).$$

Es seien $\varphi,\psi\in\mathrm{Hom}_{\mathbb{K}}(V,W)$ gegeben und $v\in V$. Dann gilt:

$$\begin{aligned}\mathrm{M}_{\mathrm{B}_W,\mathrm{B}_V}\left(\varphi+\psi\right)\cdot(v)_{\mathrm{B}_V}&=\left((\varphi+\psi)(v)\right)_{B_W}\\&=\left(\varphi(v)+\psi(v)\right)_{B_W}\\&=\left(\varphi(v)\right)_{B_W}+\left(\psi(v)\right)_{B_W}\\&=\mathrm{M}_{\mathrm{B}_W,\mathrm{B}_V}\left(\varphi\right)\cdot(v)_{B_V}+\mathrm{M}_{\mathrm{B}_W,\mathrm{B}_V}\left(\psi\right)\cdot(v)_{B_V}\\&=\left(\mathrm{M}_{\mathrm{B}_W,\mathrm{B}_V}\left(\varphi\right)+\mathrm{M}_{\mathrm{B}_W,\mathrm{B}_V}\left(\psi\right)\right)\cdot(v)_{B_V}.\end{aligned}$$

Also gilt:

$$\mathrm{M}_{\mathrm{B}_W,\mathrm{B}_V}\left(\varphi+\psi\right)=\mathrm{M}_{\mathrm{B}_W,\mathrm{B}_V}\left(\varphi\right)+\mathrm{M}_{\mathrm{B}_W,\mathrm{B}_V}\left(\psi\right).$$

Zu guter Letzt sei $\varphi\in\mathrm{Hom}_{\mathbb{K}}(V,W)$ und $\lambda\in\mathbb{K}$ gegeben und $v\in V$. Dann gilt:

$$\begin{aligned}\mathrm{M}_{\mathrm{B}_W,\mathrm{B}_V}\left(\lambda\varphi\right)\cdot(v)_{\mathrm{B}_V}&=\left((\lambda\varphi)(v)\right)_{B_W}\\&=\left(\lambda\varphi(v)\right)_{B_W}\\&=\lambda\left(\varphi(v)\right)_{B_W}\\&=\lambda\mathrm{M}_{\mathrm{B}_W,\mathrm{B}_V}\left(\varphi\right)\cdot(v)_{B_V}.\end{aligned}$$

Also gilt:

$$\mathrm{M}_{\mathrm{B}_W,\mathrm{B}_V}\left(\lambda\varphi\right)=\lambda\mathrm{M}_{\mathrm{B}_W,\mathrm{B}_V}\left(\varphi\right).$$

(f)
Es sei $A:=\mathrm{M}_{\mathrm{B}_W,\mathrm{B}_V}\left(\varphi\right)$. Dann kommutiert das folgende Diagramm

$$\begin{array}{ccc}V & \xrightarrow{\varphi} & W\\ {\scriptstyle(\cdot)_{\mathrm{B}_V}}\downarrow & & \downarrow{\scriptstyle(\cdot)_{\mathrm{B}_W}}\\ \mathbb{K}^n & \xrightarrow[\varphi_A]{} & \mathbb{K}^m.\end{array}$$

Somit folgt diese Aussage direkt mit Satz 4.3.1.

(g)

Die Menge $\mathrm{End}_{\mathbb{K}}(V) = \mathrm{Hom}_{\mathbb{K}}(V,V)$ bildet mit punktweiser Addition und skalarer Multiplikation einen Vektorraum (siehe Proposition 4.1.17). Wenn wir die skalare Multiplikation ignorieren, erhalten wir insbesondere, dass $(\mathrm{End}_{\mathbb{K}}(V),+)$ eine abelsche Gruppe bildet.

Da die Verkettung von linearen Abbildungen wieder linear ist, folgt, dass $\mathrm{End}_{\mathbb{K}}(V)$ abgeschlossen ist unter $\circ$. Verkettung ist assoziativ nach Lemma 1.3.7 mit dem Neutralelement id_V. Was noch fehlt, damit $(\mathrm{End}_{\mathbb{K}}(V),+,\circ)$ ein Ring wird, sind die beiden Distributivgesetze:

Es seien also $\varphi_1,\varphi_2,\psi \in \mathrm{End}_{\mathbb{K}}(V)$ gegeben und es sei $v \in V$ gegeben. Dann gilt:

$$\begin{aligned}\big((\varphi_1+\varphi_2)\circ\psi\big)(v) &= (\varphi_1+\varphi_2)\big(\psi(v)\big)\\ &= \varphi_1\big(\psi(v)\big)+\varphi_2\big(\psi(v)\big)\\ &= (\varphi_1\circ\psi)(v)+(\varphi_2\circ\psi)(v)\\ &= \big((\varphi_1\circ\psi)+(\varphi_2\circ\psi)\big)(v).\end{aligned}$$

Da $v \in V$ beliebig war, gilt somit:

$$(\varphi_1+\varphi_2)\circ\psi = (\varphi_1\circ\psi)+(\varphi_2\circ).$$

Für das andere Distributivgesetz seien wieder $\varphi_1,\varphi_2,\psi \in \mathrm{End}_{\mathbb{K}}(V)$ und $v \in V$ gegeben. Dann gilt:

$$\begin{aligned}\big(\psi\circ(\varphi_1+\varphi_2)\big)(v) &= \psi\big((\varphi_1+\varphi_2)(v)\big)\\ &= \psi\big(\varphi_1(v)+\varphi_2(v)\big)\\ &= \psi\big(\varphi_1(v)\big)+\psi\big(\varphi_2(v)\big)\\ &= (\psi\circ\varphi_1)(v)+(\psi\circ\varphi_2)(v)\\ &= \big((\psi\circ\varphi_1)+(\psi\circ\varphi_2)\big)(v).\end{aligned}$$

Da $v \in V$ beliebig war, gilt somit:

$$\psi\circ(\varphi_1+\varphi_2) = (\psi\circ\varphi_1)+(\psi\circ\varphi_2).$$

Damit ist $(\mathrm{End}_{\mathbb{K}}(V),+,\circ)$ ein Ring.[12]

Dass die Abbildung

$$\mathrm{End}_{\mathbb{K}}(V) \to \mathbb{K}^{n\times n},\quad \varphi \mapsto \mathrm{M}_{\mathsf{B}_V,\mathsf{B}_V}(\varphi),$$

bijektiv ist, haben wir in Teilen (a) und (b) gesehen. Dass sie die additive Struktur erhält, folgt aus Teil (e). Dass Verkettung auf Matrixmultiplikation abgebildet wird, folgt aus (c), dass das Einselement der linken Seite auf das Einselement der rechten Seite abgebildet wird haben wir in (d) gesehen. □

Definition 4.3.12.

Es sei $\mathbb{K}$ ein Körper und V ein $\mathbb{K}$-Vektorraum der Dimension $n \in \mathbb{N}$. Gegeben seien geordnete Basen B und C von V. Die Darstellungsmatrix

$$\mathrm{M}_{\mathsf{C},\mathsf{B}}(\mathrm{id}_V) \in \mathbb{K}^{n\times n}$$

von $\mathrm{id}_V : V \to V$ bezüglich der Basis B im Definitionsbereich und C im Zielbereich nennen wir auch *Basiswechselmatrix von* B *nach* C.

[12] Beachten Sie, dass in dieser ganzen Rechnung nur ein einziges Mal verwendet wurde, dass wir nur lineare Abbildungen betrachten. Wenn wir statt $\mathrm{End}_{\mathbb{K}}(V)$ die Menge V^V *aller* Abbildungen betrachtet hätten, so wäre alle Ringaxiome für $(V^V,+,\circ)$ mit Ausnahme eines der beiden Distributivgesetze erfüllt.

Bemerkung 4.3.13. (a) Angenommen, $v \in V$ ist ein Element in einem endlich dimensionalen Vektorraum V und der Koordinatenvektor von v bezüglich einer geordneten Basis B sei bekannt. Dann lässt sich der Koordinatenvektor von v bezüglich einer anderen geordneten Basis C mit Hilfe einer Basiswechselmatrix wie folgt schreiben:

$$(v)_{\mathsf{C}} = \mathrm{M}_{\mathsf{C},\mathsf{B}}(\mathrm{id})(v)_{\mathsf{B}}\,.$$

(b) Da die Abbildung $\mathrm{id}_V : V \to V$ bijektiv ist, folgt, dass eine Basiswechselmatrix immer invertierbar ist. Es gilt:

$$\left(\mathrm{M}_{\mathsf{C},\mathsf{B}}(\mathrm{id}_V)\right)^{-1} = \mathrm{M}_{\mathsf{B},\mathsf{C}}\left(\mathrm{id}_V^{-1}\right) = \mathrm{M}_{\mathsf{B},\mathsf{C}}(\mathrm{id}_V)\,,$$

das heißt: Wenn die Basiswechselmatrix in die eine Richtung bekannt ist, so lässt sich die Basiswechselmatrix in die andere Richtung durch Invertieren der Matrix berechnen.

(c) Gegeben eine lineare Abbildung $\varphi : V \to W$ zwischen endlich dimensionalen Vektorräumen V und W. Angenommen, die Darstellungsmatrix $\mathrm{M}_{\mathsf{B}_W,\mathsf{B}_V}(\varphi) \in \mathbb{K}^{n\times n}$ sei bekannt.

Wenn nun C_V und C_W andere geordnete Basen von V bzw. W sind, dann kann man die Darstellungsmatrix von φ bezüglich der neuen geordneten Basen aus der alten Darstellungsmatrix berechnen:

$$\mathrm{M}_{\mathsf{C}_W,\mathsf{C}_V}(\varphi) = \mathrm{M}_{\mathsf{C}_W,\mathsf{C}_V}\left(\mathrm{id}_W \circ \varphi \circ \mathrm{id}_V\right) = \mathrm{M}_{\mathsf{C}_W,\mathsf{B}_W}(\mathrm{id}_W)\,\mathrm{M}_{\mathsf{B}_W,\mathsf{B}_V}(\varphi)\,\mathrm{M}_{\mathsf{B}_V,\mathsf{C}_V}(\mathrm{id}_V)\,.$$

Man muss also nur die Darstellungsmatrix von links und rechts mit entsprechenden Basiswechselmatrizen multiplizieren.

Hierbei entspricht der Linksmultiplikation mit einer Basiswechselmatrix einem Basiswechsel im Zielbereich – und eine Rechtsmultiplikation einem Basiswechsel im Definitionsbereich.

Beispiel 4.3.14. Es sei

$$U := \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \in \mathbb{F}_2^4 \;\middle|\; x_1 + x_2 + x_3 + x_4 = 0 \right\}$$

der $\mathbb{F}_2$-Vektorraum aus Beispiel 4.3.10(d).

Es sei B die geordnete Basis:

$$\mathsf{B} := \left(\begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} \right).$$

Nun betrachten wir eine andere geordnete Basis für denselben Raum U:

$$\mathsf{C} := \left(\begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \right).$$

Wenn wir die Basiswechselmatrix $\mathrm{M}_{\mathsf{B},\mathsf{C}}(\mathrm{id}_U)$ berechnen wollen, nehmen wir uns also die Vektoren in der geordneten Basis C einer nach dem anderen vor:

Wir beginnen mit $c_1 := \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}$. Diesen Vektor müssen wir nun bezüglich der geordneten Basis B schreiben:

$$c_1 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} = 0 \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} + 1 \cdot \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + 1 \cdot \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}.$$

Also ist $(c_1)_\mathsf{B} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$.

Als nächstens nehmen wir den zweiten Vektor $c_2 := \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}$ aus der geordneten Basis C:

$$c_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} = 1 \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} + 1 \cdot \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + 0 \cdot \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}.$$

Damit gilt: $(c_2)_\mathsf{B} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$.

Und schließlich den dritten Vektor $c_3 := \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$:

$$c_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = 1 \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} + 1 \cdot \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + 1 \cdot \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix},$$

der somit den Koordinatenvektor $(c_3)_\mathsf{B} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$.

Für die Basiswechselmatrix gilt somit:

$$\mathrm{M}_{\mathsf{B},\mathsf{C}}(\mathrm{id}_U) = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}.$$

Die Basiswechselmatrix in die andere Richtung $\mathrm{M}_{\mathsf{C},\mathsf{B}}(\mathrm{id}_U)$ ist dann die Inverse davon:

$$\begin{array}{ccc|ccc}
0 & 1 & 1 & 1 & 0 & 0 \\
1 & 1 & 1 & 0 & 1 & 0 \\
1 & 0 & 1 & 0 & 0 & 1 \\
\hline
1 & 1 & 1 & 0 & 1 & 0 \\
0 & 1 & 1 & 1 & 0 & 0 \\
1 & 0 & 1 & 0 & 0 & 1 \\
\hline
1 & 1 & 1 & 0 & 1 & 0 \\
0 & 1 & 1 & 1 & 0 & 0 \\
0 & 1 & 0 & 0 & 1 & 1 \\
\hline
1 & 1 & 1 & 0 & 1 & 0 \\
0 & 1 & 1 & 1 & 0 & 0 \\
0 & 0 & 1 & 1 & 1 & 1 \\
\hline
1 & 1 & 1 & 0 & 1 & 0 \\
0 & 1 & 0 & 0 & 1 & 1 \\
0 & 0 & 1 & 1 & 1 & 1 \\
\hline
1 & 1 & 0 & 1 & 0 & 1 \\
0 & 1 & 0 & 0 & 1 & 1 \\
0 & 0 & 1 & 1 & 1 & 1 \\
\hline
1 & 0 & 0 & 1 & 1 & 0 \\
0 & 1 & 0 & 0 & 1 & 1 \\
0 & 0 & 1 & 1 & 1 & 1
\end{array}$$

Es gilt also:

$$\mathrm{M}_{\mathsf{C},\mathsf{B}}(\mathrm{id}_U) = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}.$$

Wir haben in Bemerkung 4.3.13 gesehen, dass jede Basiswechselmatrix invertierbar ist. Umgekehrt kann aber auch jede invertierbare Matrix als eine Basiswechselmatrix aufgefasst werden:

Lemma 4.3.15.
Es sei $\mathbb{K}$ ein Körper und V ein n-dimensionaler Vektorraum mit geordneter Basis B. Für jede invertierbare Matrix $S \in \mathrm{GL}(n,\mathbb{K})$ gibt es eine geordnete Basis C von V, sodass

$$S = \mathrm{M}_{\mathsf{B},\mathsf{C}}(\mathrm{id}_V).$$

Beweis. Da S eine invertierbare Matrix ist, ist die Abbildung

$$\varphi_S : \mathbb{K}^n \to \mathbb{K}^n, \quad x \mapsto Sx$$

ein Isomorphismus von $\mathbb{K}$-Vektorräumen. Ebenso ist die Abbildung

$$(\cdot)_{\mathsf{B}} : V \to \mathbb{K}^n, \quad v \mapsto (v)_{\mathsf{B}}$$

ein Isomorphismus. Somit können wir durch Verketten so einen Isomorphismus

$$\varphi := ((\cdot)_{\mathsf{B}})^{-1} \circ \varphi_S \circ (\cdot)_{\mathsf{B}} : V \to V$$

konstruieren, dass das folgende Diagramm kommutiert:

$$\begin{array}{ccc} V & \xrightarrow{\varphi} & V \\ {\scriptstyle (\cdot)_{\mathsf{B}}}\downarrow & & \downarrow{\scriptstyle (\cdot)_{\mathsf{B}}} \\ \mathbb{K}^n & \xrightarrow[\varphi_S]{} & \mathbb{K}^m. \end{array}$$

Die Matrix $S = \mathrm{M}_{\mathsf{B},\mathsf{B}}(\varphi)$ ist also die Darstellungsmatrix eines Vektorraumisomorphismus $\varphi : V \to V$ bezüglich der geordneten Basis

$$\mathsf{B} = (b_1, \ldots, b_n)$$

in Definitions- und Wertebereich. Ein Isomorphismus von Vektorräumen bildet Basen auf Basen ab (siehe Lemma 4.2.6). Somit ist

$$\mathsf{C} := \big(\varphi(b_1), \ldots, \varphi(b_n)\big)$$

ebenfalls eine geordnete Basis von V.

Da in den Spalten der Matrix $\mathrm{M}_{\mathsf{C},\mathsf{B}}(\varphi)$ immer die Bilder der Originalbasis B unter der Abbildung φ, dargestellt zur neuen Basis C stehen, folgt, dass die Matrix $\mathrm{M}_{\mathsf{C},\mathsf{B}}(\varphi)$ gerade die Einheitsmatrix ist:

$$\mathrm{M}_{\mathsf{C},\mathsf{B}}(\varphi) = \mathbb{1}_n.$$

Die Aussage folgt nun:

$$\begin{aligned} S &= \mathrm{M}_{\mathsf{B},\mathsf{B}}(\varphi) \\ &= \mathrm{M}_{\mathsf{B},\mathsf{B}}(\mathrm{id}_V \circ \varphi) \\ &= \mathrm{M}_{\mathsf{B},\mathsf{C}}(\mathrm{id}_V) \cdot \mathrm{M}_{\mathsf{C},\mathsf{B}}(\varphi) \\ &= \mathrm{M}_{\mathsf{B},\mathsf{C}}(\mathrm{id}_V) \cdot \mathbb{1}_n \\ &= \mathrm{M}_{\mathsf{B},\mathsf{C}}(\mathrm{id}_V). \end{aligned}$$

□

Bemerkung 4.3.16 (Zeilen- und Spaltenumformungen)**.** Wir haben in Lemma 2.5.4 gesehen, dass *elementare Zeilenumformungen* einer Matrix einer Multiplikation mit einer bestimmten invertierbaren Matrix (vom Typ $(G1)$,$(G2)$ oder $(G3)$) *von links* entsprechen. Ebenso entsprechen *Spaltenumformungen* einer Multiplikation mit einer invertierbaren Matrix (vom Typ $(G1)$,$(G2)$ oder $(G3)$) *von rechts*.

Zusammen mit Lemma 4.3.15 kann man also sagen: Eine Zeilenumformung entspricht einem Basiswechsel im Zielbereich – eine Spaltenumformung entspricht einem Basiswechsel im Definitionsbereich.

Satz 4.3.17 (Smithsche[13] Normalform)**.**
Es sei $\mathbb{K}$ ein Körper und V und W endlich dimensionale $\mathbb{K}$-Vektorräume. Zu jeder linearen Abbildung $\varphi : V \to W$ existieren geordnete Basen B_W von W und B_V von V, sodass

$$\mathrm{M}_{\mathsf{B}_W,\mathsf{B}_V}(\varphi) = \begin{pmatrix} 1 & & 0 & 0 & \cdots & 0 \\ & \ddots & & \vdots & & \vdots \\ 0 & & 1 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix}$$

[13] nach Henry John Stephen Smith, engl. Mathematiker, 1826–1883

gilt. Die Anzahl der Einsen entspricht genau dem Rang von φ. Eine Matrix von der obigen Form heißt eine Matrix in Smith-Normalform.

Beweis. Weil V und W endlich dimensionale Vektorräume sind, gibt es endliche geordnete Basen C_V von V und C_W von W. Bezüglich dieser Basen hat $\varphi : V \to W$ die Matrix $A := \mathrm{M}_{\mathsf{C}_W,\mathsf{C}_V}(\varphi) \in \mathbb{K}^{m\times n}$, wobei $n = \dim_{\mathbb{K}}(V)$ und $m = \dim_{\mathbb{K}}(W)$ ist.

Nun wenden wir den Gauß-Algorithmus auf A an, bis wir eine Matrix B erhalten, die in erweiterter Zeilenstufenform ist. Da die Zeilenumformungen des Gauß-Algorithmus Multiplikationen mit invertierbaren Matrizen von links entsprechen, entspricht dies einem Basiswechsel im Raum W. Anschließend wenden wir elementare Spaltenumformungen an, um die Matrix in Smith-Normalform zu bringen. Dies ändert zwar den Kern der Matrix, aber da wir ja kein lineares Gleichungssystem lösen wollen, ist das egal. Entscheidend ist, dass diese Spaltenumformungen nun einem Basiswechsel im Definitionsbereich V entsprechen. Insgesamt erhalten wir somit also eine Basis B_V von V und eine Basis B_W von W, sodass $\mathrm{M}_{\mathsf{B}_W,\mathsf{B}_V}(\varphi)$ in Smith-Normalform ist. Da Basiswechsel im Definitions- oder Zielbereich den Rang nicht ändern, ist der Rang von φ gleich dem Rang der Matrix in Smith-Normalform und damit gleich der Anzahl der Einsen auf der Diagonale. □

Bemerkung 4.3.18. Es seien $m,n \in \mathbb{N}$ und $\mathbb{K}$ ein Körper. Wir sagen: Zwei Matrizen $A,B \in \mathbb{K}^{m\times n}$ heißen *äquivalent*, wenn es invertierbare Matrizen $S \in \mathrm{GL}(m,\mathbb{K})$ und $T \in \mathrm{GL}(n,\mathbb{K})$ gibt mit

$$SAT = B.$$

In anderen Worten: Zwei Matrizen sind äquivalent, wenn sie die gleiche lineare Abbildung beschreiben, nur bezüglich unterschiedlicher Basen.

Satz 4.3.17 sagt also: Jede Matrix ist äquivalent zu einer Matrix in Smith-Normalform. Wenn wir also die Basen im Definitions- und Wertebereich beliebig wählen dürfen, so können wir jede Matrix in Smith-Normalform bringen.

Ein verwandter Begriff, der uns später[14] noch mehr beschäftigen wird, ist der folgende: Sind $A,B \in \mathbb{K}^{n\times n}$ zwei quadratische Matrizen derselben Größe, dann sagen wir A und B sind *ähnlich*, wenn eine invertierbare Matrix $S \in \mathrm{GL}(n,\mathbb{K})$ existiert, sodass

$$SAS^{-1} = B.$$

Man sieht dieser Formel an, dass ähnliche Matrizen auch äquivalent sind, allerdings ist dieser Begriff stärker, also nicht alle Matrizen, die äquivalent sind, sind auch ähnlich. Da die Matrix S^{-1} die Inverse der Matrix S ist, müssen A und B quadratisch sein, damit dieser Begriff sinnvoll ist.

Man kann sagen: Zwei quadratische Matrizen sind ähnlich, wenn sie den gleichen Endomorphismus beschreiben, nur bezüglich unterschiedlicher Basen. Der Unterschied zum schwächeren Begriff der Äquivalenz besteht darin, dass wir im Definitionsbereich und im Wertebereich des Endomorphismus immer dieselbe Basis wählen müssen.

Korollar 4.3.19.
Es seien $m,n \in \mathbb{N}$ und $\mathbb{K}$ ein Körper. Zwei Matrizen $A,B \in \mathbb{K}^{m\times n}$ sind genau dann äquivalent, wenn sie den gleichen Rang haben.

[14] Definition 5.5.4

Zusammenfassung von Abschnitt 4.3

(1) Jeder Vektor $v \in V$ in einem n-dimensionalen Vektorraum lässt sich zu einer geordneten Basis B als Spaltenvektor darstellen: $(v)_{\mathsf{B}} \in \mathbb{K}^n$

(2) Jede lineare Abbildung $\varphi : V \to W$ zwischen endlich dimensionalen Vektorräumen lässt sich bezüglich einer Basis B im Definitionsbereich und einer Basis C im Zielbereich als Matrix darstellen: $\mathrm{M}_{\mathsf{C},\mathsf{B}}(\varphi) \in \mathbb{K}^{m \times n}$.

(3) **In den Spalten der Matrix stehen die Bilder der Basisvektoren!**

(4) Verketten von linearen Abbildungen entspricht Multiplizieren von Matrizen.

(5) Eine Matrix ist genau dann invertierbar, wenn die dazugehörige lineare Abbildung bijektiv, also ein Isomorphismus ist.

4.4. Direkte Summen und Komplemente

Nachdem wir uns im Letzten Kapitel auf endlich dimensionale Vektorräume und geordnete Basen konzentriert haben, sind in diesem Kapitel alle Vektorräume erst einmal beliebig (also endlich dimensional oder unendlich dimensional). Viele interessante Beispiele sind aber natürlich trotzdem endlich dimensional.

Definition 4.4.1 (Direktes Produkt von Vektorräumen)**.**
Es sei $r \in \mathbb{N}$ und $V_1, \ldots, V_r$ Vektorräume über demselben Grundkörper $\mathbb{K}$. Dann definieren wir auf dem kartesischen Produkt[15] $V_1 \times \cdots V_r$ die Addition

$$(v_1, \ldots, v_r) + (w_1, \ldots, w_r) := (v_1 + w_1, \ldots, v_r + w_r)$$

und die skalare Multiplikation

$$\lambda \cdot (v_1, \ldots, v_r) := (\lambda v_1, \ldots, \lambda v_r).$$

Die Menge $(V_1 \times \cdots \times V_r, +, \cdot)$ versehen mit diesen Operationen wird selbst ein Vektorraum, den man auch das *direkte Produkt* von $V_1, \ldots, V_r$ nennt. Dass diese Operationen wirklich die Vektorraumaxiome (siehe Definition 4.1.1) erfüllen, sieht man leicht ein.

Analog kann man das direkte Produkt von Halbgruppen, Monoiden, Gruppen und Ringen definieren. Das direkte Produkt von zwei Körpern ist aber interessanterweise niemals ein Körper, weil es immer Nullteiler enthält.

Beispiel 4.4.2. Es sei $\mathbb{K}$ ein Körper, $r \in \mathbb{N}$ und jeder Vektorraum $V_j := \mathbb{K}$. Dann ist das kartesische Produkt

$$\underbrace{\mathbb{K} \times \cdots \times \mathbb{K}}_{r \text{ Faktoren}}$$

isomorph zu $\mathbb{K}^r$ über den natürlichen Isomorphismus

$$\mathbb{K} \times \cdots \times \mathbb{K} \to \mathbb{K}^r, \quad (x_1, \ldots, x_r) \mapsto \begin{pmatrix} x_1 \\ \vdots \\ x_r \end{pmatrix}.$$

[15] siehe Notation 1.2.17

Lemma 4.4.3.
Es seien V_1, V_2 Vektorräume über demselben Grundkörper $\mathbb{K}$. Weiterhin sei B eine Basis von V_1 und C eine Basis von V_2. Dann ist

$$\{(b,0) \mid b \in B\} \cup \{(0,c) \mid c \in C\}$$

eine Basis von $V_1 \times V_2$. Insbesondere ist das direkte Produkt von zwei endlich dimensionalen Vektorräumen V_1, V_2 wieder endlich dimensional mit

$$\dim_{\mathbb{K}}(V_1 \times V_2) = \dim_{\mathbb{K}}(V_1) + \dim_{\mathbb{K}}(V_2).$$

Beweis. Wir setzen

$$P := \{(b,0) \mid b \in B\} \cup \{(0,c) \mid c \in C\}.$$

Wir zeigen zuerst, dass P ein Erzeugendensystem für $V \times W$ ist. sei dazu $(v_1, v_2) \in V_1 \times V_2$. Dann folgt, dass $v_1 \in V_1 = \mathrm{LH}_{\mathbb{K}}(B)$ und somit lässt sich v_1 als Linearkombination von Elementen aus B schreiben:

$$v_1 = \sum_{j=1}^{r} \lambda_j b_j \quad \text{mit } b_j \in B, \lambda_j \in \mathbb{K}.$$

Andererseits gilt aber auch, dass $v_2 \in V_2 = \mathrm{LH}_{\mathbb{K}}(C)$ und somit lässt sich v_2 als Linearkombination von Elementen aus C schreiben:

$$v_2 = \sum_{k=1}^{s} \mu_k c_k \quad \text{mit } c_k \in C, \mu_k \in \mathbb{K}.$$

Dann gilt für (v_1, v_2) folgendes:

$$\begin{aligned}(v_1, v_2) &= \left(\sum_{j=1}^{r} \lambda_j b_j, \sum_{k=1}^{s} \mu_k c_k \right) \\ &= \left(\sum_{j=1}^{r} \lambda_j b_j, 0 \right) + \left(0, \sum_{k=1}^{s} \mu_k c_k \right) \\ &= \sum_{j=1}^{r} \lambda_j \left(b_j, 0\right) + \sum_{k=1}^{s} \mu_k (0, c_k) \in \mathrm{LH}_{\mathbb{K}}(P).\end{aligned}$$

Dies zeigt, dass P ein Erzeugendensystem ist. Nun zeigen wir noch, dass P linear unabhängig ist.

Eine Menge P ist linear unabhängig, wenn jede endliche Teilmenge davon linear unabhängig ist. Seien also endlich viele Elemente aus $P = \{(b,0) \mid b \in B\} \cup \{(0,c) \mid c \in C\}$ gegeben:

$$(b_1,0), \ldots, (b_r,0), (0,c_1), \ldots, (0,c_s) \in P.$$

Wir setzen nun eine endliche Linearkombination dieser Elemente gleich (0,0) und zeigen, dass

alle Skalare 0 sein müssen:

$$\begin{aligned}
\sum_{j=1}^{r} \lambda_j (b_j, 0) + \sum_{k=1}^{s} \mu_k (0, c_k) &= (0,0) \\
\left(\sum_{j=1}^{r} \lambda_j b_j, 0 \right) + \left(0, \sum_{k=1}^{s} \mu_k c_k \right) &= (0,0) \\
\left(\sum_{j=1}^{r} \lambda_j b_j, \sum_{k=1}^{s} \mu_k c_k \right) &= (0,0) \\
+ \left(0, \sum_{k=1}^{s} \mu_k c_k \right) &= (0,0) \\
\left(\sum_{j=1}^{r} \lambda_j b_j, \sum_{k=1}^{s} \mu_k c_k \right) &= (0,0) \\
\sum_{j=1}^{r} \lambda_j b_j = 0 \quad \text{und} \quad \sum_{k=1}^{s} \mu_k c_k &= 0 \\
\forall j \in \{1,\dots,r\} : \lambda_j = 0 \quad \text{und} \quad \forall k \in \{1,\dots,s\} : \mu_k &= 0
\end{aligned}$$

Also ist P linear unabhängig in $V_1 \times V_2$ und somit eine Basis. □

Lemma 4.4.4.
Es sei V ein Vektorraum über einem Körper $\mathbb{K}$. Dann ist der Durchschnitt von beliebig vielen Untervektorräumen wieder ein Untervektorraum.

Beweis. Es sei $\{U_J\} j \in J$ eine Menge von Untervektorräumen von V. Wir setzen

$$U := \bigcap_{j \in J} U_j = \left\{ v \in V \;\middle|\; \forall j \in J : v \in U_j \right\}.$$

Da 0_V in jedem Raum U_j liegt, liegt 0_V auch im Durchschnitt.

Es seien nun $v, w \in U$ gegeben und $j \in J$. Da $v \in U$ und $U \subseteq U_j$ gilt, folgt, dass $v \in U_j$. Ebenso gilt: Da $w \in U$ und $U \subseteq U_j$ gilt, folgt, dass $w \in U_j$. Nun ist aber U_j nach Voraussetzung ein Untervektorraum und somit abgeschlossen unter Addition. Es gilt also $v + w \in U_j$. Da $j \in J$ beliebig war, gilt damit also $v + w \in \cap_{j \in J} U_j = U$, also ist U abgeschlossen unter Addition.

Sei nun $v \in U, \lambda \in \mathbb{K}$ und $j \in J$. Da $v \in U$ und $U \subseteq U_j$ gilt, folgt, dass $v \in U_j$. Nun ist aber U_j ein Untervektorraum und somit gilt $\lambda v \in U_j$. Da $j \in J$ beliebig war, gilt damit also $\lambda v \in \cap_{j \in J} U_j = U$, also ist U abgeschlossen unter skalarer Multiplikation.

Dies zeigt, dass U ein Untervektorraum ist. □

Die Vereinigung von Untervektorräumen ist im Allgemeinen kein Untervektorraum. Es gilt aber:

Lemma 4.4.5.
Es sei V ein Vektorraum über einem Körper $\mathbb{K}$ und $U_1, \dots U_r \subseteq V$ seien Untervektorräume. Dann ist die Summe *der Untervektorräume*

$$U_1 + \cdots + U_r := \left\{ u_1 + \cdots + u_r \;\middle|\; \forall j \leq r : u_j \in U_j \right\} = \mathrm{LH}_{\mathbb{K}}(U_1 \cup \cdots \cup U_r)$$

wieder ein Untervektorraum von V.

Lemma 4.4.6.
Es sei V ein Vektorraum über einem Körper $\mathbb{K}$ und $r \in \mathbb{N}$.

(a) Die Additionsabbildung

$$+ : \underbrace{V \times \cdots \times V}_{r \text{ Faktoren}} \to V, \quad (v_1, \ldots, v_r) \mapsto v_1 + \cdots + v_r$$

ist $\mathbb{K}$-linear und surjektiv.

(b) Für Untervektorräume $U_1, \ldots, U_r$ von V gilt: Die $\mathbb{K}$-lineare Abbildung

$$U_1 \times \cdots \times U_r \to V, \quad (u_1, \ldots, u_r) \mapsto u_1 + \cdots + u_r$$

hat als Bild die Summe $U_1 + \cdots + U_r$.

(c) Für zwei Untervektorräume $U_1, U_2 \subseteq V$ gilt: Die $\mathbb{K}$-lineare Abbildung

$$\varphi : U_1 \times U_2 \to V, \quad (u_1, u_2) \mapsto u_1 + u_2$$

hat als Kern

$$\ker(\varphi) = \{(x, -x) \mid x \in U_1 \cap U_2\} \cong U_1 \cap U_2.$$

Definition 4.4.7.
Es sei V ein Vektorraum über einem Körper $\mathbb{K}$. Es seien $U_1, \ldots, U_r$ Untervektorräume von V. Wir sagen: V ist die *direkte Summe* der Untervektorräume $U_1, \ldots, U_r$ und schreiben

$$V = \bigoplus_{j=1}^{r} U_j = U_1 \oplus \cdots \oplus U_r,$$

wenn die Abbildung

$$U_1 \times \cdots \times U_r \to V, \quad (u_1, \ldots, u_r) \mapsto u_1 + \cdots + u_r$$

bijektiv, also ein Isomorphismus von $\mathbb{K}$-Vektorräumen ist, d.h. wenn jedes $v \in V$ auf genau eine Weise als Summe von Elementen aus den Untervektorräumen geschrieben werden kann.

Lemma 4.4.8 (Direkte Summe von zwei Untervektorräumen)**.**
Im Falle von genau zwei Untervektorräumen $U, W \subseteq V$ gilt:

$$V = U \oplus W \quad \Longleftrightarrow \quad (U + W = V \text{ und } U \cap W = \{0\})$$

Beweis. Die Abbildung

$$\varphi : U \times W \to V, \quad (u, w) \mapsto u + w.$$

ist genau dann surjektiv, wenn $U + W = V$.

Außerdem gilt, dass $\varphi : U \times W \to V$ genau dann injektiv ist, wenn $\ker(\varphi) = \{0\}$ ist. Dieser Kern ist aber isomorph zu $U \cap W$ nach Lemma 4.4.6(c).

Damit folgt die Behauptung. □

Definition 4.4.9 (Komplementäre Unterräume)**.**
Es sei V ein Vektorraum über einem Körper $\mathbb{K}$ und $U, W \subseteq V$ Untervektorräume. Dann heißt W ein *komplementärer Untervektorraum* oder *Vektorraumkomplement* von U, wenn

$$V = U \oplus W,$$

wenn sich also jeder Vektor $v \in V$ eindeutig schreiben lässt als $v = u + w$ mit $u \in U$ und $w \in W$.

Beispiel 4.4.10. Es sei $\mathbb{K} = \mathbb{R}$ und $V = \mathbb{R}^3$. Es sei $U = \mathrm{LH}_{\mathbb{R}}\left(\begin{pmatrix}1\\1\\1\end{pmatrix}\right) \subseteq \mathbb{R}^3$ ein eindimensionaler Untervektorraum.

Desweiteren sei $W = \mathrm{LH}_{\mathbb{R}}\left(\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}\right) \subseteq \mathbb{R}^3$ die x_1-x_2-Ebene.

Dann ist W ein Komplement von U in $V = \mathbb{R}^3$, weil sich jeder Vektor $v \in \mathbb{R}^3$ eindeutig schreiben lässt als ein Vektor in U und ein Vektor in W, es gilt also:

$$\mathbb{R}^3 = U \oplus W.$$

Anstelle von der Ebene W hätte man auch jede andere Ebene nehmen können, die nicht die Gerade U enthält. Die „natürlichste“ Wahl wäre die Ebene, die senkrecht auf U steht:

$$P := \left\{ \begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix} \in \mathbb{R}^3 \;\middle|\; x_1 + x_2 + x_3 = 0 \right\},$$

allerdings können wir in allgemeinen Vektorräumen über beliebigen Körpern geometrische Begriffe wie „senkrecht“ nicht verwenden, weil sie in allgemeinen Vektorräumen überhaupt nicht definiert sind.

Satz 4.4.11.
Es sei V ein endlich dimensionaler Vektorraum über einem Körper $\mathbb{K}$. Dann hat jeder Untervektorraum U von V ein Vektorraumkomplement W. Es gilt:

$$\dim_{\mathbb{K}}(W) = \dim_{\mathbb{K}}(V) - \dim_{\mathbb{K}}(U).$$

Beweis. Wir wenden Satz 4.2.7 auf den Vektorraum U mit Erzeugendensystem U und linear unabhängiger Teilmenge $\emptyset$ an und erhalten eine Basis B_U von U. Nun wenden wir Satz 4.2.7 nochmal an, diesmal aber auf V und ergänzen die linear unabhängige Teilmenge B_U zu einer Basis B_V von V. Wir setzen $W := \mathrm{LH}_{\mathbb{K}}(B_V \setminus B_U) \subseteq V$. Da B_V eine Basis von V ist, gilt

$$V = \mathrm{LH}_{\mathbb{K}}(B_V) = \mathrm{LH}_{\mathbb{K}}(B_U \cup (B_V \setminus B_U)) \subseteq \mathrm{LH}_{\mathbb{K}}(U \cup W) = U + W.$$

Da selbstverständlich auch die umgekehrte Mengeninklusion $U + W \subseteq V$ gilt, haben wir $V = U + W$.

Es sei nun $v \in U \cap W$ im Schnitt der beiden Untervektorräume. Dann können wir v sowohl als Linearkombination bestehend aus Elementen aus B_U, als auch als Linearkombination bestehend aus Elementen aus $B_V \setminus B_U$ schreiben. Da aber B_V linear unabhängig ist, muss dann $v = 0$ sein.

Also gilt:

$$V = U \oplus W \quad \Longleftrightarrow \quad (U + W = V \text{ und } U \cap W = \{0\})$$

und mit Lemma 4.4.8 folgt, dass W ein Vektorraumkomplement von V ist.

Es gilt nun, dass

$$V = U \oplus W \cong U \times W$$

und damit haben wir:

$$\dim_{\mathbb{K}}(V) = \dim_{\mathbb{K}}(U \times W) = \dim_{\mathbb{K}}(U) + \dim_{\mathbb{K}}(W)$$

und durch Subtrahieren von $\dim_{\mathbb{K}}(U)$ folgt die Behauptung. □

Beispiel 4.4.12. Es sei $\mathbb{K} = \mathbb{Q}$ und

$$V := \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbb{Q}^{2\times 2} \;\middle|\; a+b+c+d = 0 \right\}.$$

Wir betrachten den Untervektorraum

$$U := \left\{ A \in V \;\middle|\; A^\top = A \right\}.$$

Gesucht ist nun ein Vektorraumkomplement von U in V. Dazu suchen wir zuerst eine Basis des Raumes U: Eine Matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ ist in U, wenn die Summe aller Einträge 0 ist, d.h. $d = -a-b-c$ und wenn sie zusätzlich symmetrisch ist (wenn also $b = c$).

Eine allgemeine Matrix $A \in U$ ist also von der Form

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & b \\ c & -a-b-c \end{pmatrix} = \begin{pmatrix} a & b \\ b & -a-2b \end{pmatrix} = a \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + b \begin{pmatrix} 0 & 1 \\ 1 & -2 \end{pmatrix}.$$

Damit folgt:

$$U = \mathrm{LH}_{\mathbb{Q}} \left(\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & -2 \end{pmatrix} \right).$$

Da diese beiden Matrizen offensichtlich linear unabhängig sind, ist

$$B_U := \left\{ \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & -2 \end{pmatrix} \right\}$$

eine Basis von U und U ist somit 2 dimensional. Nun wollen wir die Basis B_U zu einer Basis von V fortsetzen.

Die Matrix $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ liegt in V, weil die Summe der Einträge 0 ist, aber nicht in U, weil sie nicht symmetrisch ist. Somit ist die Menge

$$\left\{ \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & -2 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \right\}$$

linear unabhängig. Da der Vektorraum V dreidimensional ist, ist dies nun eine Basis von V. Wenn wir nun ein Vektorraumkomplement von U in V wollen, können wir einfach

$$W := \mathrm{LH}_{\mathbb{Q}} \left(\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \right)$$

nehmen. Dann gilt

$$V = U \oplus W.$$

Alternativ hätten wir auch die Matrix $\begin{pmatrix} -1 & -1 \\ 1 & 1 \end{pmatrix}$ nehmen können und würden so eine andere Basis von V erhalten:

$$\left\{ \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & -2 \end{pmatrix}, \begin{pmatrix} -1 & -1 \\ 1 & 1 \end{pmatrix} \right\}$$

und somit auch ein anderes Vektorraumkomplement:

$$Z := \mathrm{LH}_{\mathbb{Q}} \left(\begin{pmatrix} -1 & -1 \\ 1 & 1 \end{pmatrix} \right).$$

Wir sehen also: W und Z sind beides Komplemente für U, ein Komplement ist also im Allgemeinen nicht eindeutig.

Zusammenfassung von Abschnitt 4.4

(1) Ein Vektorraum V ist die direkte Summe von Untervektorräumen $U_1,\dots,U_k$, wenn sich jeder Vektor $v \in V$ eindeutig als Summe von Vektoren aus $U_1,\dots,U_k$ schreiben lässt.

(2) Für zwei Untervektorräume gilt: Die Summe ist genau dann direkt, wenn der Schnitt $\{0\}$ ist. Für drei oder mehr Untervektorräume gilt diese Charakterisierung nicht.

(3) Wenn sich V als direkte Summe von U_1 und U_2 schreiben lässt, dann heißt U_2 Komplement von U_1 und U_1 ist ein Komplement von U_2

4.5. Quotientenvektorräume

Wir haben in Satz 3.4.12 gesehen, dass $\mathbb{Z}/m\mathbb{Z}$ ein Ring ist. Die Idee ist, dass wir alle Elemente in $m\mathbb{Z}$ identifiziert haben und gleichzeitig die Ringstruktur erhalten wollten. Dies führt zu der Äquivalenzrelation $x \equiv_m y :\iff x - y \in m\mathbb{Z}$.

Dieses Konzept einer Quotientenstruktur findet sich in fast allen algebraischen Disziplinen, es gibt beispielsweise Quotientengruppen und Quotientenringe. In diesem Kapitel möchten wir uns mit Quotienten von den algebraischen Strukturen beschäftigen, die für die lineare Algebra am Wichtigsten sind: Quotientenvektorräume.

Satz 4.5.1. *(Der Quotientenvektorraum)*
Es sei V ein Vektorraum über einem Körper $\mathbb{K}$ und $U \subseteq V$ ein Untervektorraum.

(a) Die Relation $\equiv_U$ auf V, definiert über

$$v \equiv_U w :\iff v - w \in U$$

ist eine Äquivalenzrelation auf der Menge V

(b) Auf der Quotientenmenge

$$V/U := V/\equiv_U$$

gibt es genau eine $\mathbb{K}$-Vektorraumstruktur, sodass die surjektive Quotientenabbildung

$$q : V \to V/U, \quad v \mapsto [v]_U := [v]_{\equiv_U}$$

zu einer $\mathbb{K}$-linearen Abbildung wird.

Der Vektorraum V/U heißt Quotientenvektorraum *(oder* Faktorraum*) von V modulo U.*

Für $[v]_U, [w]_U, \in V/U$ und $\lambda \in \mathbb{K}$ gilt dann $[v]_U + [w]_U = [v+w]_U$ und $\lambda \cdot [v]_U = [\lambda \cdot v]_U$. Der Nullvektor in V/U ist $[0]_U$.

(c) Es gilt:

$$\ker(q) = U.$$

Bemerkung 4.5.2 (Quotientenvektorräume und affine Unterräume)**.** Es sei U ein Untervektorraum eines Vektorraums V über einem Körper $\mathbb{K}$. Für einen Punkt $p \in V$ ist die Äquivalenzklasse $[p]_U$ gerade die Menge

$$[p]_U = \{x \in V \mid x - p \in U\} = p + U,$$

also der affine Unterraum parallel zu U mit Aufpunkt p. Man kann also sagen: Die Elemente in V/U sind genau alle affinen Unterräume parallel zu U, aber mit unterschiedlichen Aufpunkten. Die Nichteindeutigkeit des Aufpunkts p in der Darstellung $p+U$ entspricht somit exakt der Nichtinjektivität der Quotientenabbildung $q : V \to V/U$.

Bemerkung 4.5.3. Wir haben in Lemma 4.1.14 gesehen, dass der Kern einer linearen Abbildung immer ein Untervektorraum ist. Unklar war zu diesem Zeitpunkt noch, ob auch jeder Untervektorraum als Kern einer linearen Abbildung entsteht. Dies beantwortet nun Satz 4.5.1: Für jeden Untervektorraum U gilt $U = \ker(q)$, wobei $q : V \to V/U$ die Quotientenabbildung ist.

Sei nun konkret $V = \mathbb{K}^n$ mit $\mathbb{K}$ Körper und $n \in \mathbb{N}$. Dann kann jeder Untervektorraum U von V als Lösungsmenge eines homogenen linearen Gleichungssystems geschrieben werden. Wir wissen, dass U der Kern der linearen Abbildung $q : \mathbb{K}^n \to \mathbb{K}^n/U$ ist. Der Raum $\mathbb{K}^n/U$ ist als Bild eines endlich dimensionalen Vektorraums wieder endlich dimensional. Es gibt also eine geordnete Basis B von $\mathbb{K}^n/U$. Nun setzen wir

$$A := \mathrm{M}_{\mathsf{B},\mathsf{E}}(q) \in \mathbb{K}^{m\times n},$$

wobei $m = \dim_{\mathbb{K}}(\mathbb{K}^n/U)$ und E die Standardbasis auf $\mathbb{K}^n$ ist. Dann ist der Kern von A gleich dem Kern von q und somit gleich U. Anders formuliert: Das homogene lineare Gleichungssystem

$$Ax = 0$$

hat als Lösungsmenge genau den Vektorraum U.

Ebenso ist jeder affine Unterraum R von $\mathbb{K}^n$ die Lösungsmenge eines (im allgemeinen) inhomogenen linearen Gleichungssystems. Sei dazu $R = p+U$ mit Aufpunkt $p \in \mathbb{K}^n$ und $U \subseteq \mathbb{K}^n$ Untervektorraum. Dann gibt es nach obiger Überlegung eine Matrix $A \in \mathbb{K}^{m\times n}$ mit $U = \ker A$ und es gilt:

$$x \in R \iff x \in p+U \iff x-p \in U \iff x-p \in \ker A \iff A(x-p) = 0 \iff Ax = Ap.$$

Also ist R die Lösungsmenge des linearen Gleichungssystems

$$Ax = b$$

mit $b := Ap$.

Bemerkung 4.5.4. Quotientenvektorräume sind (nicht nur, aber) vor allem bei unendlich dimensionalen Vektorräumen in der Funktionalanalysis von Bedeutung. Beispielsweise sagt man eine Teilmenge $A \subseteq \mathbb{R}$ der reellen Zahlengerade ist eine Nullmenge, falls man ihr in einem gewissen Sinne das „Maß“ 0 zuordnen kann. Man sagt nun, eine Funktion $f : \mathbb{R} \to \mathbb{R}$ ist gleich 0 „fast überall“, falls die Menge der Punkte, wo die Funktionen nicht null ist, eine solche Nullmenge ist.

Funktionen, die fast überall 0 sind, spielen für die meisten Anwendungen keine Rolle und man würde gerne Funktionen $g, h : \mathbb{R} \to \mathbb{R}$ identifizieren, wenn sie sich nur durch eine solche Funktion unterscheiden, wenn sie also gleich sind „fast überall“.

Wir würden also als V den Vektorraum aller[16] Funktionen von $\mathbb{R}$ nach $\mathbb{R}$ nehmen und den Untervektorraum U aller Funktionen, die fast überall 0 sind, herausfaktorisieren. Der Quotientenraum V/U besteht dann nicht mehr aus Funktionen, sondern aus Äquivalenzklassen,

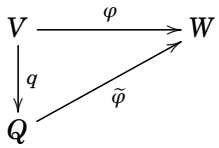
[16] Meistens nimmt man nicht alle Funktionen, sondern nur einen Untervektorraum wie z.B. den der Lebesguemessbaren Funktionen, was auch immer das bedeuten mag.

wobei wir Funktionen identifiziert haben, die fast überall gleich sind. Man faktorisiert sozusagen das weg, was einen für die Anwendung nicht interessiert und konzentriert sich auf das Wesentliche.

Ein analoger Begriff existiert in der Stochastik. Eine Zufallsvariable[17] heißt „fast sicher“ gleich 0, falls das Ereignis, dass sie nicht den Wert 0 annimmt, Wahrscheinlichkeit 0 hat. Man untersucht dann den reellen Vektorraum aller Zufallsvariablen modulo dem Untervektorraum derjenigen Zufallsvariablen, die „fast sicher“ 0 sind. Auf diese Weise identifiziert man zwei Zufallsvariablen, falls sie „fast sicher“ denselben Wert annehmen. Für die meisten Eigenschaften einer Zufallsvariable (wie Erwartungswert, Varianz, Verteilung) spielt diese Unterscheidung keine Rolle.

Beim Arbeiten mit Quotientenvektorräumen stellt sich oft die Frage, wie man Abbildungen definiert, die auf einem Quotientenvektorraum definiert sind. Die Antwort darauf gibt der folgende Satz:

Satz 4.5.5 (Faktorisieren von linearen Abbildung)**.**
Es seien V, W und Q Vektorräume über demselben Körper $\mathbb{K}$. Weiter seien lineare Abbildungen $q : V \to Q$ und $\varphi : V \to W$ gegeben. Wir nehmen an, $q : V \to Q$ sei surjektiv.



Dann sind die folgenden Aussagen äquivalent:

(i) $\ker(q) \subseteq \ker(\varphi)$

(ii) Es gibt eine Abbildung $\tilde{\varphi} : Q \to W$ mit $\varphi = \tilde{\varphi} \circ q$. Man sagt dann auch φ faktorisiert durch q.

Eine solche Abbildung $\tilde{\varphi} : Q \to W$ ist eindeutig, falls sie existiert, sie ist automatisch linear und es gilt $\mathrm{Bild}(\tilde{\varphi}) = \mathrm{Bild}(\varphi)$.

Beweis. Zuerst zeigen wir die Eindeutigkeit einer solchen Abbildung, wenn sie existiert. Es seien also $\tilde{\varphi}_1, \tilde{\varphi}_2 : Q \to W$ Abbildungen mit

$$\varphi = \tilde{\varphi}_1 \circ q \quad \text{und} \quad \varphi = \tilde{\varphi}_2 \circ q.$$

Wir wollen zeigen: $\tilde{\varphi}_1 = \tilde{\varphi}_2$. Dazu sei $u \in Q$ ein beliebiges Element. Da $q : V \to Q$ als surjektiv vorausgesetzt wurde, gibt es ein $v \in V$ mit $q(v) = u$. Dann gilt:

$$\tilde{\varphi}_1(u) = \tilde{\varphi}_1(q(v)) = (\tilde{\varphi}_1 \circ q)(v) = \varphi(v) = (\tilde{\varphi}_2 \circ q)(v) = \tilde{\varphi}_2(u).$$

Dies zeigt, dass – falls so eine Abbildung $\tilde{\varphi}$ existiert –, diese eindeutig ist.

Als nächstes zeigen wir, dass diese Abbildung – wenn sie existiert – linear ist: Es seien dazu $u_1, u_2 \in Q$. Aus der Surjektivität von $q : V \to Q$ folgt, dass es $v_1, v_2 \in V$ gibt mit $q(v_1) = u_1$ und

[17] Nein, wir werden Ihnen hier nicht formal definieren, was das genau ist.

$q(v_2) = u_2$. Wenn wir nun ausnutzen, dass $\varphi : V \to W$ und $q : V \to Q$ als linear angenommen wurden, dann gilt:

$$\begin{aligned}\widetilde{\varphi}\big(u_1 + u_2\big) &= \widetilde{\varphi}\big(q(v_1) + q(v_2)\big)\\ &= \widetilde{\varphi}\big(q(v_1 + v_2)\big)\\ &= (\widetilde{\varphi} \circ q)(v_1 + v_2)\\ &= \varphi(v_1 + v_2)\\ &= \varphi(v_1) + \varphi(v_2)\\ &= (\widetilde{\varphi} \circ q)(v_1) + (\widetilde{\varphi} \circ q)(v_2)\\ &= \widetilde{\varphi}\big(q(v_1)\big) + \widetilde{\varphi}\big(q(v_2)\big)\\ &= \widetilde{\varphi}\big(u_1\big) + \widetilde{\varphi}\big(u_2\big)\end{aligned}$$

Analog zeigt man $\widetilde{\varphi}(\lambda u) = \lambda\widetilde{\varphi}(u)$.

Wir zeigen nun: $\text{Bild}(\widetilde{\varphi}) = \text{Bild}(\varphi)$. Hier gibt es zwei Mengeninklusionen zu zeigen:

„$\subseteq$“:

Wenn $w \in \text{Bild}(\widetilde{\varphi})$ gilt, dann gibt es ein $a \in Q$ mit $\widetilde{\varphi}(a) = w$. Da $q : V \to Q$ surjektiv ist, gibt es ein $v \in V$ mit $q(v) = a$. Nun gilt:

$$\varphi(v) = (\widetilde{\varphi} \circ q)(v) = \widetilde{\varphi}\big(q(v)\big) = \widetilde{\varphi}\big(a\big) = w.$$

Also gilt: $w \in \text{Bild}(\varphi)$.

„$\supseteq$“:

Wenn $w \in \text{Bild}(\varphi)$ ist, dann gibt es ein $v \in V$ mit $\varphi(v) = w$. Nun gilt (auch ohne Surjektivität von q zu verwenden), dass:

$$\widetilde{\varphi}\big(q(v)\big) = (\widetilde{\varphi} \circ q)(v) = \varphi(v) = w.$$

Also gilt $w \in \text{Bild}(\widetilde{\varphi})$.

Es bleibt zu zeigen, dass (i) und (ii) äquivalent sind.

„(i) $\Longrightarrow$ (ii)“:

Wir definieren:

$$\widetilde{\varphi} : Q \to W, \quad q(v) \mapsto \varphi(v).$$

Wir müssen nun klären, dass diese Abbildung wohldefiniert ist. Aus der Surjektivität von $q : V \to Q$ folgt, dass jedes Element in Q von der Form $q(v)$ für ein $v \in V$ ist. Somit wird jedem Element in Q mindestens ein Element in W zugewiesen. Es bleibt zu zeigen, dass der Wert $\varphi(v)$ nicht von der Wahl von $v \in V$ abhängt. Es seien deshalb $v_1, v_2 \in V$ gegeben mit $q(v_1) = q(v_2)$. Wir müssen zeigen, dass $\varphi(v_1) = \varphi(v_2)$ ist.

Es gilt:

$$q(v_1 - v_2) = q(v_1) - q(v_2) = 0.$$

Also ist $v_1 - v_2 \in \ker(q)$. Nach Voraussetzung ist $\ker(q)$ enthalten in $\ker(\varphi)$, d.h.

$$\varphi(v_1) - \varphi(v_2) = \varphi(v_1 - v_2) = 0,$$

woraus sofort $\varphi(v_1) = \varphi(v_2)$ folgt.

Somit ist $\widetilde{\varphi} : Q \to W, \quad q(v) \mapsto \varphi(v)$ wohldefiniert. Nach Konstruktion gilt:

$$\forall v \in V : (\widetilde{\varphi} \circ q)(v) = \widetilde{\varphi}\big(q(v)\big) = \varphi(v).$$

„(ii) $\Longrightarrow$ (i)“:
Wir nehmen also an, es gebe eine solche Abbildung $\widetilde{\varphi}$. Nach der Vorüberlegung oben wissen wir, dass $\widetilde{\varphi}$ linear ist, also insbesondere 0 auf 0 abbildet.

Es sei $v \in \ker(q)$. Wir wollen zeigen: $v \in \ker(\varphi)$. Es gilt:

$$\varphi(v) = (\widetilde{\varphi} \circ q)(v) = \widetilde{\varphi}\big(q(v)\big) = \widetilde{\varphi}\big(0\big) = 0.$$

Also gilt: $v \in \ker(\varphi)$ und dies endet den Beweis. □

Beispiel 4.5.6. Es sei $V = C^\infty(\mathbb{R},\mathbb{R})$ der Vektorraum aller Funktionen von $\mathbb{R}$ nach $\mathbb{R}$, die beliebig oft differenzierbar sind und $U = \{f : \mathbb{R} \to \mathbb{R} \mid f \text{ ist konstant}\}$ der 1-dimensionale Untervektorraum aller konstanten Funktionen.

Für zwei Funktionen $f, g \in V$ gilt nun, dass sie kongruent modulo U sind, wenn $f - g \in U$ ist, wenn sich also f und g nur durch einen konstanten Summanden unterscheiden.

Es sei nun $\varphi : V \to \mathbb{R}, \quad f \mapsto f(42)$ die Abbildung, die eine Funktion an der Stelle 42 auswertet.

Man kann sich nun fragen: Ist es möglich diese Abbildung auf den Quotienten $Q := V/U$ zu übertragen, d.h. ist die Abbildung

$$\widetilde{\varphi} : V/U \to \mathbb{R}, \quad [f]_U \mapsto f(42)$$

wohldefiniert? Die Antwort gibt genau der Satz 4.5.5. Es ist möglich, φ in $\widetilde{\varphi} \circ q$ zu zerlegen, falls $\ker(q) \subseteq \ker(\varphi)$, d.h. falls für alle $f \in V$ die Implikation gilt:

$$q(f) = 0 \implies \varphi(f) = 0.$$

In Worten: Wenn f konstant ist, dann ist $f(42) = 0$. Aber das ist offenbar Unfug, denn es gibt viele konstante Funktionen, die an der Stelle 42 nicht den Wert 0 haben. Also gilt nach Satz 4.5.5, dass sich φ *nicht* über V/U faktorisieren lässt und dass die Abbildung $\widetilde{\varphi}$ nicht wohldefiniert, also gar keine Abbildung ist.

Betrachten wir nun stattdessen die folgende lineare Abbildung:

$$\psi : V \to \mathbb{R}, \quad f \mapsto f'(42),$$

die jede Funktion aus V auf ihre Ableitung an der Stelle 42 abbildet, dann gilt $\ker(q) \subseteq \ker(\psi)$, denn jede konstante Funktion hat an der Stelle 42 Ableitung 0. Also ist die folgende Abbildung auf dem Quotienten wohldefiniert

$$\widetilde{\psi} : V/U \to \mathbb{R}, \quad [f]_U \mapsto f'(42).$$

Man kann sich analog auch überlegen, dass die lineare Abbildung $D : V \to V, \quad f \mapsto f'$, die eine Funktion auf ihre Ableitungsfunktion abbildet durch V/U faktorisiert und dass die dazugehörige Abbildung

$$\widetilde{D} : V/U \to V, \quad [f]_U \mapsto f'$$

nicht nur wohldefiniert, sondern auch ein Isomorphismus von Vektorräumen ist.

Als Spezialfall von Satz 4.5.5 erhalten wir den folgenden Satz:

Korollar 4.5.7 (Homomorphiesatz)**.**
Es sei $\mathbb{K}$ ein Körper und $\varphi : V \to W$ eine $\mathbb{K}$-lineare Abbildung zwischen $\mathbb{K}$-Vektorräumen V und W. Dann ist die folgende Abbildung

$$\overline{\varphi} : V/\ker(\varphi) \to \text{Bild}(\varphi), \quad [v]_{\ker(\varphi)} \mapsto \varphi(v)$$

ein Isomorphismus von $\mathbb{K}$-Vektorräumen. Man kann also jede lineare Abbildung $\varphi : V \to W$ immer wie folgt zerlegen:

$$\varphi = \iota \circ \overline{\varphi} \circ q,$$

wobei $q : V \to V/\ker(\varphi)$ die (surjektive) Quotientenabbildung, $\overline{\varphi} : V/\ker(\varphi) \to \mathrm{Bild}(\varphi)$ ein (bijektiver) Vektorraumisomorphismus und $\iota : \mathrm{Bild}(\varphi) \to W$ die (injektive) Inklusionsabbildung ist.

$$\begin{array}{ccc} V & \xrightarrow{\quad\varphi\quad} & W \\ {\scriptstyle q}\downarrow & & \uparrow{\scriptstyle \iota} \\ V/\ker(\varphi) & \xrightarrow[\cong]{\overline{\varphi}} & \mathrm{Bild}(\varphi) \end{array}$$

Satz 4.5.8 (Quotienten und Komplemente)**.**
Es sei V ein Vektorraum über einem Körper $\mathbb{K}$ und $U \subseteq V$ ein Untervektorraum. Wir nehmen an, dass U ein Komplement W besitzt, dass also $V = U \oplus W$ gilt. Dann ist die folgende Abbildung ein Isomorphismus von $\mathbb{K}$-Vektorräumen:

$$q|_W : W \to V/U, \quad x \mapsto q(x) = [x]_U.$$

Wenn V endlich dimensional ist, dann gilt

$$\dim_{\mathbb{K}}(V/U) = \dim_{\mathbb{K}}(V) - \dim_{\mathbb{K}}(U).$$

Beweis. Die Abbildung $q : V \to V/U$ ist linear, folglich ist auch die Einschränkung auf den Unterraum W linear. Wir wollen nun zeigen, dass diese Abbildung

$$q|_W : W \to V/U, \quad x \mapsto q(x).$$

bijektiv ist.

Beginnen wir mit der Surjektivität. Sei dazu $a \in V/U$ gegeben. Da $q : V \to W$ surjektiv ist, gibt es ein $v \in V$ mit $q(v) = a$. Nun ist nach Voraussetzung V die Summe der Untervektorräume U und W. Also gibt es $u \in U$ und $w \in W$ mit $v = w + u$, also $w = v - u$. Dann gilt:

$$q|_W(w) = q(w) = q(v - u) = q(v) - q(u) = a - 0 = a.$$

Also ist $q|_W : W \to V/U$ surjektiv.

Nun zeigen wir die Injektivität. Sei dazu $w \in \ker(q|_W)$. Dann bedeutet dies ja, dass $w \in W$ und $q(w) = 0$. Somit ist $w \in W \cap \ker(q) = W \cap U$. Weil aber die Summe von W und U direkt ist, folgt mit Lemma 4.4.8, dass $W \cap U = \{0\}$. Also ist $w = 0$ und $q|_W : W \to V/U$ injektiv. □

Beispiel 4.5.9. Wir kehren zurück zu Beispiel 4.4.12: Es ist also $\mathbb{K} = \mathbb{Q}$,

$$V := \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbb{Q}^{2\times 2} \;\middle|\; a + b + c + d = 0 \right\}$$

und

$$U := \left\{ A \in V \;\middle|\; A^\top = A \right\}.$$

Wir wollen nun eine Basis des Quotientenraums V/U finden.

Nach Satz 4.5.8 ist der Quotientenraum V/U, dessen Elemente Äquivalenzklassen von Elementen aus V sind, als Vektorraum isomorph zu einem Komplement von U in V. Somit können

wir eine Basis von V/U finden, indem wir zuerst eine Basis eines Komplements von U in V finden und dann den Isomorphismus aus Satz 4.5.8 anwenden.

In Beispiel 4.4.12 haben wir gesehen, dass

$$W := \mathrm{LH}_{\mathbb{Q}}\left(\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\right) \subseteq V$$

ein Komplement von U in V ist. Somit ist

$$B_W := \left\{\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\right\}$$

eine Basis von W und schließlich

$$C := \left\{\left[\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\right]_U\right\}$$

eine Basis von V/U, jedes Element in V/U ist also ein Vielfaches von dieser einen Äquivalenzklasse $\left[\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\right]_U$.

Alternativ hätte man auch das Komplement

$$Z := \mathrm{LH}_{\mathbb{Q}}\left(\begin{pmatrix} -1 & -1 \\ 1 & 1 \end{pmatrix}\right) \subseteq V$$

nehmen können und würde dann

$$D := \left\{\left[\begin{pmatrix} -1 & -1 \\ 1 & 1 \end{pmatrix}\right]_U\right\}$$

als Basis für V/U bekommen.

Wir können also – wenn wir unterschiedliche Komplemente wählen – unterschiedliche Basen für den Quotientenraum bekommen.

Man kann sich nun fragen: Wie kann man die Äquivalenzklasse aus Basis D als Vielfaches des Äquivalenzklasse aus Basis C schreiben? Da C eine Basis ist, muss dies ja möglich sein. Wir müssten um diese Frage zu beantworten, eine Zahl $\mu \in \mathbb{Q}$ finden mit

$$\left[\begin{pmatrix} -1 & -1 \\ 1 & 1 \end{pmatrix}\right]_U = \mu\left[\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\right]_U.$$

Da die Quotientenabbildung $q : V \to V/U, \quad A \mapsto [A]_U$ linear ist, können wir den Skalar μ in die Äquivalenzklasse hineinziehen:

$$\left[\begin{pmatrix} -1 & -1 \\ 1 & 1 \end{pmatrix}\right]_U = \left[\mu\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\right]_U.$$

Nun nutzen wir aus, dass zwei Äquivalenzklassen modulo U genau dann gleich sind, wenn ihre

Differenz in U liegt:

$$\begin{aligned}
&\left[\begin{pmatrix}-1 & -1\\ 1 & 1\end{pmatrix}\right]_U = \left[\mu\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}\right]_U\\
\iff &\begin{pmatrix}-1 & -1\\ 1 & 1\end{pmatrix} - \mu\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix} \in U\\
\iff &\begin{pmatrix}-1 & -1-\mu\\ 1+\mu & 1\end{pmatrix} \in U\\
\iff &\begin{pmatrix}-1 & -1-\mu\\ 1+\mu & 1\end{pmatrix}^\top = \begin{pmatrix}-1 & -1-\mu\\ 1+\mu & 1\end{pmatrix}\\
\iff &\begin{pmatrix}-1 & 1+\mu\\ -1-\mu & 1\end{pmatrix} = \begin{pmatrix}-1 & -1-\mu\\ 1+\mu & 1\end{pmatrix}\\
\iff &1+\mu = -1-\mu\\
\iff &\mu = -1.
\end{aligned}$$

Hier haben wir außerdem ausgenutzt, dass $[A]_U = [0]_U \iff A \in U$.

Es gilt also:

$$\left[\begin{pmatrix}-1 & -1\\ 1 & 1\end{pmatrix}\right]_U = (-1)\left[\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}\right]_U.$$

Schließen möchten wir mit einen neuen Beweis für die Dimensionsformel (Satz 2.5.9) aus Kapitel 2.5:

Satz 4.5.10 (Dimensionsformel für lineare Abbildungen)**.**
Es sei $\mathbb{K}$ ein Körper und $\varphi : V \to W$ eine $\mathbb{K}$-lineare Abbildung zwischen $\mathbb{K}$-Vektorräumen V und W. Wir nehmen an, dass V endlich dimensional ist. Dann gilt:

$$\operatorname{rg}(\varphi) + \dim_{\mathbb{K}} \ker(\varphi) = \dim_{\mathbb{K}}(V),$$

wobei $\operatorname{rg}(\varphi) = \dim_{\mathbb{K}}(\operatorname{Bild}(\varphi))$.

Beweis. Nach dem Homomorphiesatz (Korollar 4.5.7) gilt:

$$V/\ker(\varphi) \cong \operatorname{Bild}(\varphi).$$

Folglich haben beide Seiten dieselbe Dimension:

$$\dim_{\mathbb{K}}\left(V/\ker(\varphi)\right) = \dim_{\mathbb{K}}(\operatorname{Bild}(\varphi)).$$

Die rechte Seite dieser Gleichung ist der Rang der Abbildung φ und die linke Seite ist $\dim_{\mathbb{K}}(V) - \dim_{\mathbb{K}}(\ker(\varphi))$ nach Satz 4.5.8. □

Korollar 4.5.11 (Dimensionsformel für Summe und Schnitt)**.**
Es sei V ein Vektorraum über einem Körper $\mathbb{K}$. Gegeben seien zwei endlich dimensionale Untervektorräume $U, W \subseteq V$. Dann gilt:

$$\dim_{\mathbb{K}}(U+W) = \dim_{\mathbb{K}}(U) + \dim_{\mathbb{K}}(W) - \dim_{\mathbb{K}}(U \cap W).$$

Beweis. Wir wenden die Dimensionsformel für lineare Abbildungen (Satz 4.5.10) auf die folgende lineare Abbildung an:

$$U \times W \to V, \quad (u,w) \mapsto u+w.$$

Der Kern ist isomorph zu $U \cap W$ nach Lemma 4.4.6(c) und das Bild ist genau $U+W$. □

Zusammenfassung von Abschnitt 4.5

(1) Wenn U ein Untervektorraum von V ist, dann existiert der Quotientenvektorraum V/U.

(2) Die Quotientenabbildung $q : V \to V/U, \quad v \mapsto [v]$ ist surjektiv und es gilt $\ker(q) = U$. Insbesondere ist also jeder Untervektorraum der Kern einer linearen Abbildung.

(3) Eine lineare Abbildung $\varphi : V \to W$ induziert genau dann eine lineare Abbildung $\widetilde{\varphi} : V/U \to W$, wenn $U \subseteq \ker \varphi$.

(4) Jede lineare Abbildung $\varphi : V \to W$ lässt sich zerlegen in $\varphi = \iota \circ \overline{\varphi} \circ q$, wobei $q : V \to V/U$ surjektiv, $\overline{\varphi} : V/U \to \mathrm{Bild}(\varphi)$ bijektiv und $\iota : \mathrm{Bild}(\varphi) \to W$ injektiv ist.

(5) Wenn W ein Komplement von U in V ist, dann gilt $V/U \cong W$.

(6) Es gilt die Dimensionsformel für lineare Abbildungen:
$\mathrm{rg}(\varphi) + \dim_{\mathbb{K}}(\ker(\varphi)) = \dim(V)$.

(7) Es gilt die Dimensionsformel für Untervektorräume:
$\dim_{\mathbb{K}}(U+W) = \dim_{\mathbb{K}}(U) + \dim_{\mathbb{K}}(W) - \dim_{\mathbb{K}}(U \cap W)$.

4.6. Der Dualraum

Wir haben in Proposition 4.1.17 gesehen, dass die Menge der $\mathbb{K}$-linearen Abbildungen von einem $\mathbb{K}$-Vektorraum V in einen $\mathbb{K}$-Vektorraum W mit punktweiser Addition und skalarem Vielfachen selbst ein Vektorraum wird. Interessant sind zwei Spezialfälle: Mit dem Fall $W = V$, also mit dem Raum aller Endomorphismen eines Vektorraums V werden wir uns in Kapitel 5 beschäftigen. Jetzt interessiert uns der Spezialfall $W = \mathbb{K}$, also der Raum aller linearen Abbildungen von V in den Grundkörper $\mathbb{K}$:

Definition 4.6.1.
Es sei $\mathbb{K}$ ein Körper und V ein Vektorraum. Eine *Linearform* σ auf V ist ein Element in $\mathrm{Hom}_{\mathbb{K}}(V,\mathbb{K})$, also eine lineare Abbildung $\sigma : V \to \mathbb{K}$.

Der Vektorraum $V^* := \mathrm{Hom}_{\mathbb{K}}(V,\mathbb{K})$ aller Linearformen auf V heißt der *Dualraum* von V.

Bemerkung 4.6.2. Dualräume spielen vor allem in der Funktionalanalysis und der Differentialgeometrie eine wichtige Rolle.

Es sei $V := C^\infty([-1,1],\mathbb{R})$ der $\mathbb{R}$-Vektorraum aller beliebig oft differenzierbaren Funktionen auf $[-1,1]$. Weiterhin sei $f : [-1,1] \to \mathbb{R}$ eine stetige Funktion. Dann ist die Abbildung

$$\sigma_f : V \to \mathbb{R}, \quad h \mapsto \int_{-1}^{1} f(t)h(t)dt$$

linear, also ein Element in $V^* = \mathrm{Hom}_{\mathbb{K}}(V, \mathbb{R})$.

Man kann zeigen, dass beim Übergang der Funktion f zu der Linearform σ_f keine Information verloren geht, dass man also f rekonstruieren kann, wenn man nur σ_f kennt. Man kann also eine nichtlineare Funktion f ersetzen durch eine lineare Abbildung, die dann allerdings auf einem unendlich dimensionalen Vektorraum definiert ist. Auch wenn es auf den ersten Blick nicht klar ist, warum man das tun sollte, so hat sich diese Sichtweise oft als hilfreich erwiesen, weil es viele Funktionale auf V gibt, die nicht von dieser Form sind, die man aber als „verallgemeinerte Funktionen“ auffassen kann.

Die bekannteste solche „verallgemeinerten Funktionen“, die in der Elektrotechnik und der Physik immer wieder auftaucht, ist die „Dirac-Delta-Funktion“ δ. Sie ist überall gleich 0, nur an der Stelle 0 ist sie unendlich groß und das Integral über diese „Funktion“ ist gleich 1. Obgleich sehr hilfreich in den Anwendungen, war es für die Mathematik lange Zeit nicht klar, wie man diese Idee formalisieren kann. Denn eine „echte“ Funktion, die überall 0 ist bis auf einen Punkt, kann niemals Integral 1 haben – das lässt die Integrationstheorie nicht zu. Allerdings wird diese Delta-„Funktion“ auch nie wie eine normale Funktion benutzt, sondern immer nur als Faktor in einem Integral der Form $\int_{-1}^{1} \delta(t)h(t)dt$, sodass man das Problem dann dadurch lösen konnte, indem man sagt, die Dirac-Delta-„Funktion“ und andere „verallgemeinerte Funktionen“ sind gar keine richtigen Funktionen, sondern nur Linearformen auf V (sogenannte *Distributionen*).

Dies ist z.B. in der Theorie der partiellen Differentialgleichungen von Vorteil, wo es wesentlich einfacher ist, eine Lösung im Raum der „verallgemeinerten Funktionen“ (Distributionen) zu finden als in dem kleineren Raum der echten Funktionen.

Proposition 4.6.3 (Dualisierung von linearen Abbildungen)**.**
Es seien V und W Vektorräume über demselben Grundkörper $\mathbb{K}$.

(a) Gegeben sei eine $\mathbb{K}$-lineare Abbildung $\varphi : V \to W$. Dann ist die Abbildung

$$\varphi^* : W^* \to V^*, \quad \sigma \mapsto \sigma \circ \varphi,$$

die eine Linearform auf W mit φ verkettet, selbst eine $\mathbb{K}$-lineare Abbildung.

(b) Ist $\varphi : V \to W$ surjektiv, so ist $\varphi^* : W^* \to V^*$ injektiv.

(c) Ist $\varphi : V \to W$ bijektiv, so ist $\varphi^* : W^* \to V^*$ bijektiv.

(d) Es gilt

$$(\mathrm{id}_V)^* = \mathrm{id}_{V^*}.$$

(e) Es sei U ein Vektorraum über $\mathbb{K}$ und $\varphi : V \to W$ und $\psi : U \to V$ lineare Abbildungen. Dann gilt:

$$\big(\varphi \circ \psi\big)^* = \psi^* \circ \varphi^*.$$

Proposition 4.6.4.
Es sei V ein Vektorraum über einem Körper $\mathbb{K}$ und $U \subseteq V$ ein Untervektorraum. Die Menge aller Linearformen auf V, die auf U verschwinden

$$\big\{\sigma \in V^* \;\big|\; \sigma|_U = 0\big\} \subseteq V^*$$

ist ein Untervektorraum von V^* und ist isomorph zum Dualraum des Faktorraums $(V/U)^*$ über den Isomorphismus

$$(V/U)^* \to \big\{\rho \in V^* \;\big|\; \rho|_U = 0\big\}, \quad \sigma \mapsto q^*(\sigma) = \sigma \circ q,$$

wobei $q : V \to V/U$ die Quotientenabbildung ist.

Beweis. Die Quotientenabbildung $q : V \to V/U$ ist linear und surjektiv (Satz 4.5.1). Nach Proposition 4.6.3 ist die dualisierte Abbildung

$$q^* : (V/U)^* \to V^*, \quad \sigma \mapsto \sigma \circ q$$

injektiv.

Wie jede andere injektive Abbildung auch kann man diese Abbildung bijektiv machen, indem wir sie auf ihr Bild koeinschränkt (Bemerkung 1.3.13), das folgende ist also ein Isomorphismus von $\mathbb{K}$-Vektorräumen:

$$(V/U)^* \to \mathrm{Bild}(q^*), \quad \sigma \mapsto \sigma \circ q.$$

Was ist aber nun das $\mathrm{Bild}(q^*)$? Eine Linearform $\rho \in V^*$ lässt sich genau dann als $\sigma \circ q$ schreiben (mit $\sigma \in (V/U)^*$), wenn $U = \ker(q) \subseteq \ker \rho$ (Satz 4.5.5) und das ist gleichbedeutend mit $\rho|_U = 0$. □

Lemma 4.6.5 (Auswertungsabbildung)**.**
Es sei V ein Vektorraum über einem Körper $\mathbb{K}$ und $v \in V$ ein Element. Dann ist die Auswertungsabbildung

$$\eta_V(v) : V^* \to \mathbb{K}, \quad \sigma \mapsto \sigma(v)$$

eine $\mathbb{K}$-Linearform definiert auf dem Dualraum V.

Proposition 4.6.6 (Bidual)**.**
Es sei V ein Vektorraum über einem Körper $\mathbb{K}$.

(a) Die Abbildung $\eta_V : V \to (V^*)^*, \quad v \mapsto \eta_V(v)$, die jeden Vektor auf die dazugehörige Auswertungsabbildung abbildet, ist $\mathbb{K}$-linear.

(b) Es sei W ein weiterer Vektorraum über $\mathbb{K}$ und $\varphi : V \to W$ eine $\mathbb{K}$-lineare Abbildung. Dann ist das folgende Diagramm kommutativ:

$$\begin{array}{ccc} V & \xrightarrow{\varphi} & W \\ {\scriptstyle \eta_V}\downarrow & & \downarrow{\scriptstyle \eta_W} \\ (V^*)^* & \xrightarrow[(\varphi^*)^*]{} & (W^*)^*, \end{array}$$

es gilt also:

$$(\varphi^*)^* \circ \eta_V = \eta_W \circ \varphi.$$

(c) Es gilt:

$$\big(\eta_V\big)^* \circ \eta_{V^*} = \mathrm{id}_{V^*}.$$

Der Raum $(V^*)^*$ wird auch als der Bidualraum *von V bezeichnet.*

Satz 4.6.7.
Es sei $\mathbb{K}$ ein Körper und V ein $\mathbb{K}$-Vektorraum mit Basis $B = \{b_i \mid i \in I\}$. Wir nehmen an, $b_i \neq b_j$ für $i \neq j$.

(a) Für jedes $j \in I$ ist die Koordinatenabbildung

$$b_j^* : V \to \mathbb{K}, \quad \sum_{i \in I} \lambda_i b_i \mapsto \lambda_j,$$

die einen Vektor auf seine b_j-Koordinate abbildet, linear. Für unendliches I ist diese Summe ist so zu verstehen, dass immer nur endlich viele Skalare ungleich 0 sind.

(b) *Die Menge* $\{b_i^* \mid i \in I\}$ *ist eine linear unabhängige Teilmenge von* V^*.

(c) *Der Dualraum* V^* *ist isomorph zu* $\mathbb{K}^B$, *dem Raum aller Funktionen von* B *nach* $\mathbb{K}$ *(und damit auch isomorph zu* $\mathbb{K}^I$*).*

(d) *Die Abbildung*

$$\eta_V : V \to (V^*)^*, \quad v \mapsto \sigma \mapsto \sigma(v),$$

ist injektiv.

Beweis. (a)
Für ein $j \in I$ betrachten wir die Funktion

$$f_j : B \to \mathbb{K}, \quad b_i \mapsto \begin{cases} 1 & \text{für } i = j \\ 0 & \text{sonst.} \end{cases}$$

Nach dem Fortsetzungssatz für lineare Abbildungen (Satz 4.2.14) gibt es eine lineare Fortsetzung $\varphi : V \to \mathbb{K}$ und man sieht leicht, dass dies genau die Koordinatenabbildung b_j^* ist, die somit wohldefiniert und linear ist.

(b)
Eine Menge ist linear unabhängig, wenn jede endliche Teilmenge linear unabhängig ist.

Es sei deshalb

$$\sum_{j \in I} \lambda_j b_j^* = 0, \tag{$*$}$$

wobei nur endlich viele λ_j nicht 0 sind. Wir müssen zeigen, dass alle $\lambda_j = 0$ sind.

Sei dazu $i \in I$. Es ist zu zeigen, dass $\lambda_i = 0$ gilt.

Wir wenden nun die Gleichung $(*)$ an auf den Vektor $b_j \in V$:

$$\sum_{j \in I} \lambda_j b_j^*(b_i) = 0.$$

Die linke Seite der Gleichung wird nun einfach zu λ_i und damit folgt die Behauptung.

(c)
Dies folgt durch direkte Anwendung von Satz 4.2.14.

(d)
Es sei $v \in \ker(\eta_V)$, d.h. $\eta_V(v) = 0$. Wir wollen zeigen, dass $v = 0$ ist.

Weil $V = \mathrm{LH}_{\mathbb{K}}(B)$, gilt:

$$v = \sum_{i \in I} \lambda_i b_i,$$

wobei wieder nur endlich viele λ_i ungleich 0 sind. Wir wollen nun zeigen, dass alle λ_i null sind.

Es sei dazu $j \in I$ gegeben. Dann ist $b_j^* \in V^*$ und $\eta_V(v) \in (V^*)^*$. Es ist somit möglich, $\eta_V(v)$ auf b_j^* anzuwenden:

$$(\eta_V(v))(b_j^*) = b_j^*(v) = b_j^*\left(\sum_{i \in I} \lambda_i b_i\right) = \lambda_i.$$

Andererseits ist aber $(\eta_V(v))(b_j^*) = 0$, weil $\eta_V(v)$ die konstante Nullabbildung ist. Daraus folgt die Behauptung. □

Kommen wir nun zu dem endlich dimensionalen Fall:

Satz 4.6.8.
Es sei $\mathbb{K}$ ein Körper und V ein endlich dimensionaler Vektorraum über $\mathbb{K}$ mit geordneter Basis

$$\mathsf{B} = (b_1, \ldots, b_n).$$

Dann bilden die Koordinatenabbildungen aus Satz 4.6.7

$$\mathsf{B}^* := \left(b_1^*, \ldots, b_n^*\right)$$

eine geordnete Basis des Dualraums V^*, genannt die duale Basis. *Insbesondere gilt also* $\dim_{\mathbb{K}}(V^*) = \dim_{\mathbb{K}}(V)$ *und* $V^* \cong_{\mathbb{K}} V$.

Für eine Linearform $\sigma : V \to \mathbb{K}$ gilt:

$$(\sigma)_{\mathsf{B}^*} = \left(\mathrm{M}_{\mathsf{E},\mathsf{B}}(\sigma)\right)^\top.$$

Beweis. Nach Satz 4.3.11 (e) ist die Abbildung

$$V^* \to \mathbb{K}^{1\times n}, \quad \sigma \mapsto \mathrm{M}_{\mathsf{E},\mathsf{B}}(\sigma)$$

ein $\mathbb{K}$-Vektorraumisomorphismus. Insbesondere gilt:

$$\dim_{\mathbb{K}}(V^*) = \dim_{\mathbb{K}}(\mathbb{K}^{1\times n}) = n.$$

Die Koordinatenabbildungen $b_1^*, \ldots, b_n^*$ sind linear unabhängig nach Satz 4.6.7. Eine linear unabhängige n-elementige Menge in einem n-elementigen Vektorraum ist eine Basis.

Es bleibt zu zeigen, dass für ein $\sigma \in V^*$ die folgende Formel gilt:

$$(\sigma)_{\mathsf{B}^*} = \left(\mathrm{M}_{\mathsf{E},\mathsf{B}}(\sigma)\right)^\top.$$

Wir setzen $\alpha_j := \sigma(b_j) \in \mathbb{K}$ für jedes $j \in \{1, \ldots, n\}$.

Da in den Spalten der Darstellungsmatrix die Bilder der Basisvektoren stehen, folgt somit:

$$\mathrm{M}_{\mathsf{E},\mathsf{B}}(\sigma) = \begin{pmatrix}\alpha_1 & \cdots & \alpha_n\end{pmatrix} \in \mathbb{K}^{1\times n}.$$

Definieren wir nun die Linearform $\rho := \alpha_1 b_1^* + \cdots + \alpha_n b_n^* \in V^*$, dann sehen wir durch direktes Einsetzen, dass

$$\rho(b_j) = b_j \quad \text{für alle } j \in \{1, \ldots, n\}.$$

Die linearen Abbildungen σ und ρ stimmen somit auf einer Basis von V überein, somit müssen sie gleich sein (siehe Satz 4.2.14). Es gilt also:

$$\sigma = \rho = \alpha_1 b_1^* + \cdots + \alpha_n b_n^*.$$

Also sind die Koordinaten von σ aufgefasst als Element im Vektorraum V^* bezüglich der Basis $B^* = \left(b_1^*, \ldots, b_n^*\right)$ genau die Skalare $\alpha_1, \ldots, \alpha_n$ und der Darstellungsvektor von σ bezüglich B^* ist also

$$(\sigma)_{\mathsf{B}^*} = \begin{pmatrix}\alpha_1 \\ \vdots \\ \alpha_n\end{pmatrix}$$

und dies endet den Beweis. □

Proposition 4.6.9 (Transponieren entspricht Dualisieren)**.**
Es seien V und W endlich dimensionale Vektorräume über einem Körper $\mathbb{K}$. Weiterhin sei eine geordnete Basis B von V, sowie eine geordnete Basis C von W gegeben. Die dazugehörigen dualen Basen von V^* und W^* werden mit B^* und C^* bezeichnet.

Für jede lineare Abbildung $\varphi : V \to W$ und ihre duale Abbildung $\varphi^* : W^* \to V^*$ gilt:

$$\mathrm{M}_{\mathsf{B}^*,\mathsf{C}^*}\left(\varphi^*\right) = \left(\mathrm{M}_{\mathsf{C},\mathsf{B}}\left(\varphi\right)\right)^\top.$$

Das Dualisieren einer Abbildung entspricht also dem Transponieren der Matrix.

Beweis. Es sei $\sigma \in W^*$ ein beliebiges Element im Dualraum von W^*. Dann gilt:

$$\begin{aligned}
\mathrm{M}_{\mathsf{B}^*,\mathsf{C}^*}\left(\varphi^*\right)(\sigma)_{C^*} &= \left(\varphi^*(\sigma)\right)_{\mathsf{B}^*} \\
&= \left(\sigma \circ \varphi\right)_{\mathsf{B}^*} \\
&= \left(\mathrm{M}_{\mathsf{E},\mathsf{B}}\left(\sigma \circ \varphi\right)\right)^\top \\
&= \left(\mathrm{M}_{\mathsf{E},\mathsf{C}}(\sigma) \cdot \mathrm{M}_{\mathsf{C},\mathsf{B}}\left(\varphi\right)\right)^\top \\
&= \left(\mathrm{M}_{\mathsf{C},\mathsf{B}}\left(\varphi\right)\right)^\top \cdot \left(\mathrm{M}_{\mathsf{E},\mathsf{C}}(\sigma)\right)^\top \\
&= \left(\mathrm{M}_{\mathsf{C},\mathsf{B}}\left(\varphi\right)\right)^\top \cdot (\sigma)_{C^*}.
\end{aligned}$$

Da $\sigma \in W^*$ beliebig war, gilt somit:

$$\mathrm{M}_{\mathsf{B}^*,\mathsf{C}^*}\left(\varphi^*\right) = \left(\mathrm{M}_{\mathsf{C},\mathsf{B}}\left(\varphi\right)\right)^\top. \qquad \square$$

Proposition 4.6.10 (Bidual endlich dimensional)**.**
Es sei V ein endlich dimensionaler Vektorraum über einem Körper $\mathbb{K}$. Dann ist die Abbildung

$$\eta_V : V \to (V^*)^*, \quad v \mapsto \eta(v) : \sigma \mapsto \sigma(v)$$

ein Isomorphismus von $\mathbb{K}$-Vektorräumen. Man sagt: Der Bidualraum eines endlich dimensionalen Vektorraum V ist natürlich isomorph zu V

Beweis. Nach Satz 4.6.7(d) ist $\eta_V : V \to (V^*)^*$ injektiv, d.h. Bild(η_V) ist isomorph zu V.

Nach Satz 4.6.8 gilt: $\dim_{\mathbb{K}}((V^*)^*) = \dim_{\mathbb{K}}(V^*) = \dim_{\mathbb{K}}(V)$.

Also ist Bild(η_V) ein Untervektorraum von $(V^*)^*$ mit der gleichen (endlichen) Dimension. Demnach muss Bild(η_V) bereits der ganze Raum $(V^*)^*$ sein und η_V ist auch surjektiv. $\square$

Bemerkung 4.6.11. Für einen endlich dimensionalen Vektorraum gilt auch $V \cong_{\mathbb{K}} V^*$. Trotzdem ist der Dualraum V^* nicht *natürlich isomorph* zu V. Dies liegt daran, dass ein solcher Isomorphismus immer von der Wahl einer Basis abhängt. Der Isomorphismus zwischen V und dem Bidualraum $(V^*)^*$ ist dagegen *natürlich*, weil man ihn explizit ohne willkürliche Wahl einer Basis schreiben kann.

Der Versuch, eine formal korrekte Definition dieser intuitive Idee einer *natürlichen* Abbildung zu geben führte zu einem Gebiet der Mathematik, das man als *Kategorientheorie* nennt. Dort werden Begriffe wie *Objekte*, *Morphismen* und *Funktoren* eingeführt, um schließlich eine korrekte Definition einer *natürlichen Transformation* geben zu können. All dies würde für diese Veranstaltung zu weit führen. Es ist aber – je nachdem, wie Ihre weiteres Studium verläuft – sehr wahrscheinlich, dass Sie sich früher oder später mit Kategorientheorie beschäftigen werden, da diese die Sprache ist, in der Großteile der Mathematik[18] formuliert werden.

[18] und auch Teile der theoretischen Informatik

Beispiel 4.6.12.
Es sei $\mathbb{K} = \mathbb{R}$ und $V := \mathbb{K}^{(\mathbb{N})}$ der Raum der abbrechenden Folgen (siehe Beispiel 4.2.5(e) oder Beispiel 4.2.11(d)). Dann ist $\{e_k \mid k \in \mathbb{N}\}$ eine unendliche Basis für diesen Vektorraum.

Die Koordinatenabbildungen

$$e_k^* : V \to \mathbb{R}, \quad (x_j)_{j\in\mathbb{N}} \mapsto x_k$$

bilden nun eine abbrechende Folge $(x_j)_{j\in\mathbb{N}}$ auf den k-ten Folgenterm ab. Nach Satz 4.6.7(b) ist die Menge $L := \{e_k^* \} k \in \mathbb{N}$ linear unabhängig in V^*. Allerdings ist L *keine* Basis für V^*, weil L kein Erzeugendensystem ist.

Beispielsweise ist die Linearform

$$\Sigma : V \to \mathbb{R}, \quad (x_j)_j \mapsto \sum_{j=1}^{\infty} x_j,$$

die eine abbrechende Folge auf ihre Summe abbildet, keine Linearkombination der Linearformen aus L.

Allgemein gibt es sehr viele Linearformen in V^*, die nicht in $\mathrm{LH}_{\mathbb{R}}(L)$ sind. Für jede beliebige (nicht notwendigerweise abbrechende) Folge $(y_k)_{k\in\mathbb{N}}$ ist das folgende eine Linearform auf V

$$\Sigma_{(y_j)_j} : V \to \mathbb{R}, \quad (x_j)_j \mapsto \sum_{j=1}^{\infty} y_j x_j.$$

Man beachte, dass alle auftauchenden Summen immer endliche Summen sind, weil immer nur endlich viele Terme nicht null sind. Man kann relativ leicht nachweisen, dass alle Linearformen auf V von diesem Typ sind, es gilt:

$$\mathbb{R}^{\mathbb{N}} \to V^*, \quad (y_j)_j \mapsto \Sigma_{(y_j)_j}$$

ist ein Isomorphismus.

Der Dualraum von $\mathbb{R}^{(\mathbb{N})}$ ist also isomorph zu $\mathbb{R}^{\mathbb{N}}$.

Allgemeiner gilt: Für jede Menge J und jeden Körper $\mathbb{K}$ ist der Dualraum von $\mathbb{K}^{(J)}$ der Raum $\mathbb{K}^J$, wenn wir also den Raum der Funktionen, die nur an endlich vielen Stellen ungleich 0 sind, dualisieren, erhalten wir den Raum *aller* Funktionen von J nach $\mathbb{K}$. Falls J unendlich viele Elemente hat, gibt es somit immer Elemente im Dualraum, die keine Linearkombination der Koordinatenabbildungen sind.

Kurz gesagt: Das Konzept einer *dualen Basis* existiert nur im Endlichdimensionalen.

Zusammenfassung von Abschnitt 4.6

(1) Zu jedem Vektorraum V ist der Dualraum V^* der Raum aller linearen Abbildungen von V in den Grundkörper $\mathbb{K}$.

(2) Zu jeder geordneten Basis B eines endlich dimensionalen Vektorraums V gibt es eine duale Basis B^* des Dualraums V^*. Es gilt: $V \cong V^*$ für endlich dimensionale Vektorräume.

(3) Es gilt die Formel: $(\sigma)_{\mathsf{B}^*} = \left(\mathrm{M}_{\mathsf{E},\mathsf{B}}(\sigma)\right)^\top$

(4) Im Unendlichdimensionalen gibt es keine duale Basis.

(5) Jede lineare Abbildung $\varphi : V \to W$ lässt sich dualisieren zu einer dualen linearen Abbildung $\varphi^* : W^* \to V^*$.

(6) Wenn wir für V und W geordnete Basen wählen und die Dualräume mit den entsprechenden dualen Basen versehen, so ist die Darstellungsmatrix der dualen Abbildung φ^* genau die Transponierte der Darstellungsmatrix von φ.

(7) Es gibt eine natürliche lineare Abbildung $\eta_V : V \to (V^*)^*$ von V in seinen Bidual $(V^*)^*$.

(8) $(\mathbb{K}^{(J)})^* \cong \mathbb{K}^J$.

5. Endomorphismen

In diesem Kapitel möchten wir uns nun mit *Endomorphismen* beschäftigen, also mit linearen Abbildungen $\varphi : V \to V$, die von einem $\mathbb{K}$-Vektorraum V in sich selbst abbilden. Ist V endlich dimensional, so können wir dem Endomorphismus φ eine Darstellungsmatrix

$$A := \mathrm{M}_{\mathrm{B},\mathrm{B}}(\varphi) \in \mathbb{K}^{n\times n}$$

zuordnen, die dann quadratisch ist. So ist das Studium von Vektorraumendomorphismen eng verknüpft mit der Theorie der $(n \times n)$-Matrizen.

Neben dem bereits eingeführten *Rang* einer Matrix ist die *Determinante* ein wichtiges Werkzeug zur Untersuchung von quadratischen Matrizen und Endomorphismen.

5.1. Das Signum einer Permutation

Bevor wir uns mit der Determinante eines Endomorphismus beschäftigen können, müssen wir ein wenig ausholen:

Es sei $n \in \mathbb{N}$. Eine *Permutation* auf der Menge $\{1,\dots,n\}$ ist eine bijektive Abbildung $\sigma : \{1,\dots,n\} \to \{1,\dots,n\}$. Die Menge aller Permutationen auf $\{1,\dots,n\}$ bilden (mit der Hintereinanderausführung $\circ$) eine Gruppe, die *symmetrische Gruppe* genannt und mit

$$\mathscr{S}(n) := \{\sigma : \{1,\dots,n\} \to \{1,\dots,n\} \mid \sigma \text{ ist bijektiv}\}$$

bezeichnet wird. Sie hat $|\mathscr{S}(n)| = n!$ viele Elemente. Wir haben uns in Beispiel 3.2.3(h) schon ausführlich mit solchen Permutationen beschäftigt.

Definition 5.1.1 (Transpositionen)**.**
Es sei $n \in \mathbb{N}$. Eine Permutation $\tau \in \mathscr{S}(n)$ heiße *Transposition*, wenn sie genau zwei Elemente $k,l \in \{1,\dots,n\}, k \neq l$ vertauscht und alle anderen Elemente festhält:

$$\tau : \{1,\dots,n\} \to \{1,\dots,n\}, \quad j \mapsto \begin{cases} l & \text{falls } j = k, \\ k & \text{falls } j = l, \\ j & \text{sonst.} \end{cases}$$

Lemma 5.1.2 (Jede Permutation ist ein Produkt von Transpositionen)**.**
Es sei $n \in \mathbb{N}$ und $\sigma \in \mathscr{S}(n)$. Dann gibt es ein $r \in \mathbb{N}_0$ und Transpositionen $\tau_1,\dots,\tau_r$, sodass

$$\sigma = \tau_r \circ \cdots \circ \tau_1.$$

Im Falle $r = 0$ ist die Verkettung von 0 vielen Transpositionen gerade die Identität $\mathrm{id}_{\{1,\dots,n\}}$*. Man sagt auch die Gruppe $\mathscr{S}(n)$ wird von den Transpositionen* erzeugt.

Beweis. Wir zeigen die Aussage per Induktion über $n \in \mathbb{N}$.

Induktionsanfang $n = 1$**:**
Eine Permutation $\sigma \in \mathscr{S}(1)$ ist eine Abbildung von $\{1\}$ nach $\{1\}$. Es gibt nur eine solche Abbildung. Also ist $\sigma = \mathrm{id}_{\{1\}}$.

Induktionsschritt:
Es sei nun $n \in \mathbb{N}$ so gewählt, dass sich jede Permutation $\sigma \in \mathscr{S}(n)$ als Verkettung von Transpositionen aus $\mathscr{S}(n)$ schreiben lässt.

Es sei nun $\sigma \in \mathscr{S}(n+1)$ gegeben. Dann sind zwei Fälle zu unterscheiden:

Fall 1:$\sigma(n+1) = n+1$
Betrachten wir $\sigma_0 := \sigma|_{\{1,\dots,n\}}$ die Einschränkung auf die Menge $\{1,\dots,n\}$. Da σ injektiv ist, ist σ_0 auch injektiv. Aus $\sigma(n+1) = n+1$ folgt, dass σ_0 ausschließlich Werte in $\{1,\dots,n\}$ annimmt. Wir können also (durch eine Koeinschränkung) σ_0 als Abbildung von $\{1,\dots,n\}$ nach $\{1,\dots,n\}$ auffassen.

Da das Bild der injektiven Abbildung σ_0 genau n Elemente hat, ist

$$\sigma_0 : \{1,\dots,n\} \to \{1,\dots,n\}, \quad j \mapsto \sigma(j)$$

sogar bijektiv, also gilt: $\sigma_0 \in \mathscr{S}(n)$. Nach Induktionsvoraussetzung ist σ_0 also eine Verkettung von endlich vielen Transpositionen der Menge $\{1,\dots,n\}$. Da σ den Punkt $n+1$ festhält, ist somit auch σ eine Verkettung von endlich vielen Transpositionen.

Fall 1:$\sigma(n+1) \neq n+1$
Wir setzen $k := \sigma(n+1) \in \{1,\dots,n\}$ und $\tau \in \mathscr{S}(n+1)$ sei die Transposition, die k und $n+1$ vertauscht.

Betrachten wir nun $\tau \circ \sigma \in \mathscr{S}(n+1)$. Diese Permutation hat nun die Eigenschaft, dass $n+1$ auf sich selbst abgebildet wird:

$$(\tau \circ \sigma)(n+1) = \tau(\sigma(n+1)) = \tau(k) = n+1.$$

Also gilt nach der Argumentation aus Fall 1, dass sich $\tau \circ \sigma$ als Produkt von endlich vielen Transpositionen schreiben lässt:

$$\tau \circ \sigma = \tau_r \circ \cdots \circ \tau_1$$

Wenn wir nun diese Gleichung mit der Transposition τ von links verketten, erhalten wir:

$$\tau \circ \tau \circ \sigma = \tau \circ \tau_r \circ \cdots \circ \tau_1$$

Da $\tau \circ \tau = \mathrm{id}_{\{1,\dots,n+1\}}$ gilt, folgt die Behauptung. □

Man beachte, dass der Beweis von Lemma 5.1.2 konstruktiv ist, in dem Sinne, dass er einen Algorithmus liefert, wie man eine gegebene Permutation in Transpositionen zerlegen kann. Man beachte, dass diese Zerlegung aber überhaupt nicht eindeutig ist.

Wir wollen nun jeder Permutation eine Zahl zuweisen, die eine wesentliche Eigenschaft der Permutation widerspiegelt:

Definition 5.1.3 (Signum einer Permutation)**.**
Es sei $n \in \mathbb{N}$ und $\sigma \in \mathscr{S}(n)$ eine Permutation auf $\{1,\dots,n\}$. Das *Signum* der Permutation σ ist definiert als:

$$\mathrm{sgn}(\sigma) = \prod_{\substack{\{i,j\} \subseteq \{1,\dots,n\} \\ i \neq j}} \frac{\sigma(i) - \sigma(j)}{i - j}.$$

Man beachte, dass dies wohldefiniert ist, denn wenn für $\{i,j\} = \{j,i\}$ folgt, dass

$$\frac{\sigma(i)-\sigma(j)}{i-j} = \frac{\sigma(j)-\sigma(i)}{j-i}.$$

Dies ist nicht die einzige Möglichkeit, das Signum zu definieren. Andere Möglichkeiten, die aber alle äquivalent zu dieser Definition sind, finden Sie in Bemerkung 5.1.9 und Proposition 5.3.13.<

Beispiel 5.1.4.
Es sei $\sigma : \{1,\dots,4\} \to \{1,\dots,4\}, \quad j \mapsto 5-j$. Die Abbildung σ ist bijektiv, also eine Permutation auf $\{1,\dots,4\}$, d.h. $\sigma \in \mathscr{S}(4)$.

$$\sigma = \left(\begin{array}{c|c|c|c} 1 & 2 & 3 & 4 \\ \hline 4 & 3 & 2 & 1 \end{array}\right) = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \end{pmatrix} = (1,4)\circ(2,3) = \left(\begin{array}{c|cccc} & 1 & 2 & 3 & 4 \\ \hline 1 & & & & \bullet \\ 2 & & & \bullet & \\ 3 & & \bullet & & \\ 4 & \bullet & & & \end{array}\right).$$

Das Signum von σ berechnet sich nun per Definition zu:

$$\begin{aligned} \operatorname{sgn}(\sigma) &= \frac{\sigma(1)-\sigma(2)}{1-2}\cdot\frac{\sigma(1)-\sigma(3)}{1-3}\cdot\frac{\sigma(1)-\sigma(4)}{1-4}\cdot\frac{\sigma(2)-\sigma(3)}{2-3}\cdot\frac{\sigma(2)-\sigma(4)}{2-4}\cdot\frac{\sigma(3)-\sigma(4)}{3-4} \\ &= \frac{4-3}{1-2}\cdot\frac{4-2}{1-3}\cdot\frac{4-1}{1-4}\cdot\frac{3-2}{2-3}\cdot\frac{3-1}{2-4}\cdot\frac{2-1}{3-4} \\ &= \frac{1}{-1}\cdot\frac{2}{-2}\cdot\frac{3}{-3}\cdot\frac{1}{-1}\cdot\frac{2}{-2}\cdot\frac{1}{-1} = 1. \end{aligned}$$

Lemma 5.1.5.
Für $n \in \mathbb{N}$ und $\sigma \in \mathscr{S}(n)$ gilt:

$$\operatorname{sgn}(\sigma) \in \{-1,1\}.$$

Beweis. Es sei $\sigma \in \mathscr{S}(n)$. Dann ist $\operatorname{sgn}(\sigma) \in \mathbb{R}$ und wir können darauf die Betragsfunktion anwenden:

$$\begin{aligned} |\operatorname{sgn}(\sigma)| &= \left| \prod_{\substack{\{i,j\}\subseteq\{1,\dots,n\} \\ i\neq j}} \frac{\sigma(i)-\sigma(j)}{i-j} \right| \\ &= \prod_{\substack{\{i,j\}\subseteq\{1,\dots,n\} \\ i\neq j}} \left| \frac{\sigma(i)-\sigma(j)}{i-j} \right| \\ &= \prod_{\substack{\{i,j\}\subseteq\{1,\dots,n\} \\ i\neq j}} \frac{|\sigma(i)-\sigma(j)|}{|i-j|} \\ &= \frac{\prod_{\substack{\{i,j\}\subseteq\{1,\dots,n\} \\ i\neq j}} |\sigma(i)-\sigma(j)|}{\prod_{\substack{\{i,j\}\subseteq\{1,\dots,n\} \\ i\neq j}} |i-j|} \end{aligned}$$

Im Nenner dieses Bruches steht nun das Produkt über alle Abstände $|i-j|$ zwischen je zwei verschiedenen Elementen $i,j \in \{1,\dots,n\}$.

Da die Abbildung $\sigma : \{1,\dots,n\} \to \{1,\dots,n\}$ bijektiv ist, steht exakt dieselbe Zahl auch im Zähler. Somit ist der Bruch gleich 1, woraus die Behauptung folgt. □

Lemma 5.1.6.
Für $n \in \mathbb{N}$ und $\tau \in \mathscr{S}(n)$ eine Transposition. Dann gilt: $\operatorname{sgn}(\tau) = -1$.

Lemma 5.1.7 (Signum ist ein Gruppenhomomorphismus)**.**
Die Abbildung

$$\operatorname{sgn} : (\mathscr{S}(n), \circ) \to \mathbb{Z}^\times = (\{-1,1\}, \cdot),$$

ist ein Gruppenhomomorphismus, d.h.

$$\forall \sigma, \tau \in \mathscr{S}(n) : \operatorname{sgn}(\sigma \circ \tau) = \operatorname{sgn}(\sigma) \cdot \operatorname{sgn}(\tau).$$

Beweis. Es seien $\sigma, \tau \in \mathscr{S}(n)$ gegeben. Dann gilt:

$$\begin{aligned}
\operatorname{sgn}\big(\sigma \circ \tau\big) &= \prod_{\substack{\{i,j\} \subseteq \{1,\dots,n\} \\ i \neq j}} \frac{(\sigma \circ \tau)(i) - (\sigma \circ \tau)(j)}{i - j} \\
&= \prod_{\substack{\{i,j\} \subseteq \{1,\dots,n\} \\ i \neq j}} \frac{\sigma\big(\tau(i)\big) - \sigma\big(\tau(j)\big)}{i - j} \\
&= \left(\prod_{\substack{\{i,j\} \subseteq \{1,\dots,n\} \\ i \neq j}} \frac{\sigma\big(\tau(i)\big) - \sigma\big(\tau(j)\big)}{\tau(i) - \tau(j)} \cdot \frac{\tau(i) - \tau(j)}{i - j} \right) \\
&= \prod_{\substack{\{i,j\} \subseteq \{1,\dots,n\} \\ i \neq j}} \left(\frac{\sigma\big(\tau(i)\big) - \sigma\big(\tau(j)\big)}{\tau(i) - \tau(j)} \right) \cdot \prod_{\substack{\{i,j\} \subseteq \{1,\dots,n\} \\ i \neq j}} \left(\frac{\tau(i) - \tau(j)}{i - j} \right) \\
&= \prod_{\substack{\{k,l\} \subseteq \{1,\dots,n\} \\ k \neq l}} \left(\frac{\sigma\big(k\big) - \sigma\big(l\big)}{k - l} \right) \cdot \prod_{\substack{\{i,j\} \subseteq \{1,\dots,n\} \\ i \neq j}} \left(\frac{\tau(i) - \tau(j)}{i - j} \right) \\
&= \operatorname{sgn}(\sigma) \cdot \operatorname{sgn}(\tau).
\end{aligned}$$

□

Satz 5.1.8 (Charakterisierung der Signum-Funktion)**.**
Es sei $n \in \mathbb{N}$ gegeben. Dann gibt es genau einen Gruppenhomomorphismus

$$\varphi : (\mathscr{S}(n), \circ) \to \mathbb{Z}^\times = (\{-1,1\}, \cdot),$$

der jede Transposition auf (-1) *abbildet. Dieser Gruppenhomomorphismus ist das Signum (Definition 5.1.3).*

Bemerkung 5.1.9. Sei σ eine Permutation, die sich als Verkettung von r Transpositionen schreiben lässt. Nach Satz 5.1.8 gilt nun: $\operatorname{sgn}(\sigma) = (-1)^r$.

Definition 5.1.10.
Eine Permutation $\sigma \in \mathscr{S}(n)$ heißt *gerade*, falls $\operatorname{sgn}(\sigma) = 1$ und *ungerade*, falls $\operatorname{sgn}(\sigma) = -1$.
Die Untergruppe

$$\mathscr{A}(n) := \ker(\operatorname{sgn}) = \{\sigma \in \mathscr{S}(n) \mid \operatorname{sgn}(\sigma) = 1\}$$

der geraden Permutationen heißt auch *alternierende Gruppe*.

Lemma 5.1.11 (Ungerade Permutationen)**.**
Es sei $n \in \mathbb{N}$ und $\tau \in \mathscr{S}(n)$ mit $\operatorname{sgn}(\tau) = -1$. *Dann gilt:*

$$\{\sigma \in \mathscr{S}(n) \mid \operatorname{sgn}(\sigma) = -1\} = \left\{\rho \circ \tau \;\middle|\; \rho \in \mathscr{A}(n)\right\}.$$

Insbesondere gibt es also genauso viele gerade wie ungerade Permutationen und es gilt:

$$|\mathscr{A}(n)| = \frac{n!}{2}.$$

Beweis. In jeder Gruppe $(G, *)$ ist die Rechtsmultiplikation mit einem Element $a \in G$ bijektiv:

$$G \to G, \quad u \mapsto u * a.$$

Also ist auch in $(\mathscr{S}(n), \circ)$ die Verkettung mit τ eine bijektive Abbildung:

$$\Phi : \mathscr{S}(n) \to \mathscr{S}(n), \quad \rho \mapsto \rho \circ \tau.$$

Da die Signumsabbildung $\operatorname{sgn} : (\mathscr{S}(n), \circ) \to (\{-1,1\}, \cdot)$ ein Gruppenhomomorphismus ist, folgt, dass Φ gerade Permutationen auf ungerade Permutationen und umgekehrt abbildet. Somit ist insbesondere die Abbildung

$$\mathscr{A}(n) \to \{\sigma \in \mathscr{S}(n) \mid \operatorname{sgn}(\sigma) = -1\}, \quad \rho \mapsto \rho \circ \tau$$

eine Bijektion, d.h. es gibt gleich viele gerade wie ungerade Permutationen.

Es sei $m := |\mathscr{A}(n)| = |\{\sigma \in \mathscr{S}(n) \mid \operatorname{sgn}(\sigma) = -1\}|$.

Da jede Permutation entweder gerade oder ungerade ist, folgt somit:

$$n! = |\mathscr{S}(n)| = |\mathscr{A}(n) \cup \{\sigma \in \mathscr{S}(n) \mid \operatorname{sgn}(\sigma) = -1\}| = |\mathscr{A}(n)| + |\{\sigma \in \mathscr{S}(n) \mid \operatorname{sgn}(\sigma) = -1\}| = m + m = 2m.$$

Hieraus folgt die Behauptung. □

Zusammenfassung von Abschnitt 5.1

(1) Jede Permutation auf $\{1, \ldots, n\}$ ist eine Verkettung von Transpositionen. Diese Darstellung ist nicht eindeutig.

(2) Es gibt genau einen Gruppenhomomorphismus $\operatorname{sgn} : (\mathscr{S}(n), \circ) \to (\{-1,1\}, \cdot)$, der alle Transpositionen auf -1 abbildet.

(3) Permutationen σ mit $\operatorname{sgn}(\sigma) = 1$ heißen *gerade*, solche mit $\operatorname{sgn}(\sigma) = -1$ heißen *ungerade*.

(4) Für $n \geq 2$ gibt es gleich viele gerade wie ungerade Permutationen.

(5) Die geraden Permutationen bilden die Untergruppe $\mathscr{A}(n) := \ker(\operatorname{sgn})$.

5.2. Alternierende Abbildungen

Definition 5.2.1.
Es seien V und W Vektorräume über einem Körper $\mathbb{K}$ und es sei $n \in \mathbb{N}$. Wir bezeichnen das n-fache kartesische Produkt von V mit

$$V^n := V \times \cdots \times V = \{(v_1, \ldots, v_n) \mid \forall j \in \{1, \ldots, n\} : v_j \in V\}$$

(a) Es sei $j \in \{1, \ldots, n\}$. Eine Abbildung $\omega : V^n \to W$ heiße linear in der j-ten Komponente, falls für fest gewählte $v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_n \in V$ die Abbildung

$$V \to W, \quad v \mapsto \omega(v_1, \ldots, v_{j-1}, v, v_{j+1}, \ldots, v_n)$$

linear ist, d.h. falls für $v, v' \in V$ und $\lambda \in \mathbb{K}$ gilt:

$$\omega(v_1, \ldots, v_{j-1}, v+v', v_{j+1}, \ldots, v_n) = \omega(v_1, \ldots, v_{j-1}, v, v_{j+1}, \ldots, v_n) + \omega(v_1, \ldots, v_{j-1}, v', v_{j+1}, \ldots, v_n)$$

und

$$\omega(v_1, \ldots, v_{j-1}, \lambda v, v_{j+1}, \ldots, v_n) = \lambda \omega(v_1, \ldots, v_{j-1}, v, v_{j+1}, \ldots, v_n).$$

(b) Eine Abbildung $\omega : V^n \to W$ heiße *multilinear*, falls sie in jeder einzelnen Komponente linear ist. Für $n = 2$ nennt man dies auch *bilinear*.

(c) Eine Abbildung $\omega : V^n \to W$ heiße *alternierend*, falls sie multilinear ist und jedes Tupel $(v_1, \ldots, v_n)$ auf 0 abgebildet wird, sobald mindestens zwei der Komponenten gleich sind, d.h. falls gilt:

$$\big(\exists i, j \in \{1, \ldots, n\} : \big(v_i = v_j \text{ und } i \neq j\big)\big) \implies \omega(v_1, \ldots, v_n) = 0.$$

Beispiel 5.2.2. (a) Für $n = 1$ ist jede lineare Abbildung $\varphi : V^1 = V \to W$ alternierend.

(b) Es sei $\mathbb{K} = \mathbb{R}$ und $V := \mathbb{R}^3$. Das *Kreuzprodukt* auf $\mathbb{R}^3$ ist definiert als

$$\times : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3, \quad \left(\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}\right) \mapsto \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \times \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} := \begin{pmatrix} x_2 y_3 - x_3 y_2 \\ x_3 y_1 - x_1 y_3 \\ x_1 y_2 - x_2 y_1 \end{pmatrix}$$

Diese Abbildung ist alternierend.

(c) Es sei $\mathbb{K} = \mathbb{R}$ und $V := \mathbb{R}^3$. Das *Standardskalarprodukt*[1] auf $\mathbb{R}^3$ ist definiert als

$$\langle \cdot, \cdot \rangle : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}, \quad \left(\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}\right) \mapsto \left\langle \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \right\rangle := x_1 y_1 + x_2 y_2 + x_3 y_3$$

Diese Abbildung ist bilinear, aber nicht alternierend.

(d) Die gewöhnliche Multiplikationsabbildung in einem Körper

$$\mathbb{K} \times \mathbb{K} \to \mathbb{K}, \quad (x, y) \mapsto xy$$

ist bilinear, aber nicht alternierend, weil z.B. $1 \cdot 1 \neq 0$ ist.

[1] Wir werden uns genauer mit Skalarprodukten in LA2 beschäftigen.

(e) Es sei $\mathbb{K} = \mathbb{R}$ und $V := \mathbb{R}^{\mathbb{R}}$ der Vektorraum aller Funktionen von $\mathbb{R}$ nach $\mathbb{R}$. Die Verkettungsabbildung

$$\circ : V \times V \to V, \quad (f,g) \mapsto f \circ g$$

ist linear in der ersten Komponente, aber nicht multilinear, weil sie nicht linear in der zweiten Komponente ist.

Lemma 5.2.3 (Rechenregeln für alternierende Abbildungen)**.**
Es seien V, W Vektorräume über einem Körper $\mathbb{K}$ und $n \in \mathbb{N}$. Weiterhin sei $\omega : V^n \to W$ eine alternierende Abbildung und $(v_1, \ldots, v_n) \in V^n$.

(G1) Wenn man das μ-fache der i-ten Komponente auf die j-te Komponente addiert ($i \neq j$), so ändert sich der Wert unter ω nicht:

$$\omega(v_1, \ldots, v_{j-1}, v_j + \mu v_i, v_{j+1}, \ldots, v_n) = \omega(v_1, \ldots, v_{j-1}, v_j, v_{j+1}, \ldots, v_n)$$

(G2) Wenn man den i-ten und den j-ten Eintrag vertauscht ($i \neq j$), so wird der Wert mit (-1) multipliziert:

$$\omega(v_1, \ldots, v_{i-1}, v_j, v_{i+1}, \ldots, v_{j-1}, v_i, v_{j+1}, \ldots, v_n) = (-1) \cdot \omega(v_1, \ldots, v_{i-1}, v_i, v_{i+1}, \ldots, v_{j-1}, v_j, v_{j+1}, \ldots, v_n)$$

(G3) Es ist möglich, aus jeder Komponente einen Koeffizienten $\lambda \in \mathbb{K}$ herauszuziehen

$$\omega(v_1, \ldots, v_{j-1}, \lambda v, v_{j+1}, \ldots, v_n) = \lambda \omega(v_1, \ldots, v_{j-1}, v, v_{j+1}, \ldots, v_n).$$

Beweis. (G1)
Nach Voraussetzung ist $\omega : V^n \to W$ linear in jeder Komponente, wenn man die anderen Komponenten fest hält. Insbesondere gilt also:

$$\omega(v_1, \ldots, v_{j-1}, v_j + \mu v_i, v_{j+1}, \ldots, v_n) = \omega(v_1, \ldots, v_{j-1}, v_j, v_{j+1}, \ldots, v_n) + \mu\omega(v_1, \ldots, v_{j-1}, v_i, v_{j+1}, \ldots, v_n)$$

Der zweite Summand ist aber 0, weil ω alternierend ist und zwei der Einträge identisch sind – nämlich v_i.

(G2)
Wir betrachten das Tupel

$$(v_1, \ldots, v_{i-1}, v_i + v_j, v_{i+1}, \ldots, v_{j-1}, v_i + v_j, v_{j+1}, \ldots, v_n) \in V^n,$$

das in der i-ten und in der j-ten Komponente den Eintrag $v_i + v_j$ hat. Da ω alternierend ist, wird dieses Tupel also auf 0 abgebildet:

$$\begin{aligned}
0 =& \omega(v_1, \ldots, v_{i-1}, v_i + v_j, v_{i+1}, \ldots, v_{j-1}, v_i + v_j, v_{j+1}, \ldots, v_n) \\
=& \omega(v_1, \ldots, v_{i-1}, v_i, v_{i+1}, \ldots, v_{j-1}, v_i + v_j, v_{j+1}, \ldots, v_n) \\
&+ \omega(v_1, \ldots, v_{i-1}, v_j, v_{i+1}, \ldots, v_{j-1}, v_i + v_j, v_{j+1}, \ldots, v_n) \\
=& \underbrace{\omega(v_1, \ldots, v_{i-1}, v_i, v_{i+1}, \ldots, v_{j-1}, v_i, v_{j+1}, \ldots, v_n)}_{=0} \\
&+ \omega(v_1, \ldots, v_{i-1}, v_i, v_{i+1}, \ldots, v_{j-1}, v_j, v_{j+1}, \ldots, v_n) \\
&+ \omega(v_1, \ldots, v_{i-1}, v_j, v_{i+1}, \ldots, v_{j-1}, v_i, v_{j+1}, \ldots, v_n) \\
&+ \underbrace{\omega(v_1, \ldots, v_{i-1}, v_j, v_{i+1}, \ldots, v_{j-1}, v_j, v_{j+1}, \ldots, v_n)}_{=0} \\
=& + \omega(v_1, \ldots, v_{i-1}, v_i, v_{i+1}, \ldots, v_{j-1}, v_j, v_{j+1}, \ldots, v_n) \\
&+ \omega(v_1, \ldots, v_{i-1}, v_j, v_{i+1}, \ldots, v_{j-1}, v_i, v_{j+1}, \ldots, v_n).
\end{aligned}$$

Hieraus folgt direkt die Aussage.

(G3)

Hierfür ist die Alterniertheit überhaupt nicht notwendig. Diese Eigenschaft folgt also direkt aus der Multilinearität. □

Wir können also elementare Umformungen – wie im Gauß-Algorithmus – auf die Argumente von ω anwenden, um den Wert zu berechnen.

Was passiert nun, wenn wir die Reihenfolge der Einträge verändern?

Lemma 5.2.4.
Es seien V, W Vektorräume über einem Körper $\mathbb{K}$ und $n \in \mathbb{N}$. Es sei $\omega : V^n \to W$ eine alternierende Abbildung. Weiter sei $\sigma \in \mathscr{S}(n)$ eine Permutation. Dann gilt:

$$\omega\left(v_{\sigma(1)}, \ldots, v_{\sigma(n)}\right) = \operatorname{sgn}(\sigma) \cdot \omega(v_1, \ldots, v_n).$$

Beweis. Nehmen wir zuerst einmal an, $\sigma = \tau$ wäre eine Transposition. Dann gilt $\operatorname{sgn}(\tau) = -1$ nach Lemma 5.1.6. Mit mit der Rechenregel (G2) aus Lemma 5.2.3 erhalten wir dann:

$$\omega\left(v_{\tau(1)}, \ldots, v_{\tau(n)}\right) = (-1) \cdot \omega(v_1, \ldots, v_n) = \operatorname{sgn}(\tau) \cdot \omega(v_1, \ldots, v_n).$$

Damit ist die Aussage bewiesen, wenn $\sigma = \tau$ eine Transposition ist.

Es sei nun $\sigma \in \mathscr{S}(n)$ eine beliebige Permutation. Nach Lemma 5.1.2 lässt sich σ schreiben als

$$\sigma = \tau_r \circ \cdots \circ \tau_1.$$

Es gilt (mit Lemma 5.1.7 und Lemma 5.1.6), dass:

$$\operatorname{sgn}(\sigma) = \operatorname{sgn}(\tau_r) \cdots \operatorname{sgn}(\tau_1) = (-1) \cdots (-1) = (-1)^r.$$

Wenn wir nun $\sigma = \tau_r \circ \cdots \circ \tau_1$ auf das Tupel $(v_1, \ldots, v_n)$ anwenden, so wenden wir einfach der Reihe nach alle Transpositionen an (von rechts nach links). Bei jedem Mal wird der Wert der alternierenden Abbildung mit (-1) multipliziert, sodass wir am Ende erhalten:

$$\omega\left(v_{\sigma(1)}, \ldots, v_{\sigma(n)}\right) = (-1)^r \omega(v_1, \ldots, v_n) = \operatorname{sgn}(\sigma) \omega(v_1, \ldots, v_n).$$ □

Wir haben bei linearen Abbildungen gesehen, dass es ausreicht, die Werte auf einem Erzeugendensystem zu kennen, um die Abbildung allgemein zu berechnen. Wir wollen nun eine analoge Aussage für alternierende Abbildungen formulieren:

Satz 5.2.5 (Transformationsformel für alternierende Abbildungen)**.**
Es seien V, W Vektorräume über einem Körper $\mathbb{K}$ und $n \in \mathbb{N}$. Weiterhin sei $\omega : V^n \to W$ eine alternierende Abbildung. Gegeben seien $u_1, \ldots, u_n \in V$ und $v_1, \ldots, v_n \in \mathrm{LH}_{\mathbb{K}}(u_1, \ldots, u_n)$. Dann gibt es ein $\Delta \in \mathbb{K}$ mit

$$\omega(v_1, \ldots, v_n) = \Delta \cdot \omega(u_1, \ldots, u_n).$$

Wenn wir jeden Vektor v_j als Linearkombination der $u_1, \ldots, u_n$ schreiben:

$$v_j = \sum_{i=1}^{n} \alpha_{i,j} u_i \quad \text{mit } \alpha_{i,j} \in \mathbb{K},$$

so lässt sich der Wert von Δ berechnen sich als

$$\Delta = \sum_{\sigma \in \mathscr{S}(n)} \operatorname{sgn}(\sigma) \cdot \alpha_{\sigma(1),1} \cdots \alpha_{\sigma(n),n}.$$

Insbesondere hängt der Wert nicht von der Abbildung ω ab.

Korollar 5.2.6.
Es seien V, W Vektorräume über einem Körper $\mathbb{K}$ und $n \in \mathbb{N}$. Weiterhin sei $\omega : V^n \to W$ eine alternierende Abbildung und $v_1, \ldots, v_n \in V$. Falls $\omega(v_1, \ldots, v_n) \neq 0$, so sind die Vektoren $v_1, \ldots, v_n$ paarweise verschieden und linear unabhängig.

Umgekehrt heißt dies, dass für linear abhängige Vektoren $v_1, \ldots, v_n$ gilt: $\omega(v_1, \ldots, v_n) = 0$.

Beweis. Wir zeigen die Implikation:

$(\omega(v_1, \ldots, v_n) \neq 0) \implies (v_1, \ldots, v_n$ sind paarweise verschieden und $\{v_1, \ldots, v_n\}$ ist linear unabhängig$)$.

durch Kontraposition (Siehe Satz 1.1.6(iv)), d.h. wir zeigen:

$(v_1, \ldots, v_n$ sind nicht paarweise verschieden oder $\{v_1, \ldots, v_n\}$ ist linear abhängig$) \implies (\omega(v_1, \ldots, v_n) = 0)$.

Falls die Vektoren $v_1, \ldots, v_n$ nicht paarweise verschieden sind, folgt $\omega(v_1, \ldots, v_n) = 0$ direkt aus der Definition einer alternierenden Abbildung (siehe Definition 5.2.1).

Nehmen wir also an, dass $\{v_1, \ldots, v_n\}$ linear abhängig sind, so gilt $m = \dim_{\mathbb{K}}(\mathrm{LH}_{\mathbb{K}}(v_1, \ldots, v_n)) < n$. Es gibt somit eine geordnete Basis $(u_1, \ldots, u_m)$ von $\mathrm{LH}_{\mathbb{K}}(v_1, \ldots, v_n)$. Für alle $j \in \{m+1, \ldots, n\}$ setze $u_j := 0$. Dann gilt:

$$v_1, \ldots, v_n \in \mathrm{LH}_{\mathbb{K}}(u_1, \ldots, u_n) = \mathrm{LH}_{\mathbb{K}}(u_1, \ldots, u_m, 0, \ldots, 0).$$

Also gilt nach Satz 5.2.5, dass es ein $\Delta \in \mathbb{K}$ gibt mit

$$\omega(v_1, \ldots, v_n) = \Delta \cdot \omega(u_1, \ldots, u_m, 0, \ldots, 0).$$

Da ω in der n-ten Komponente linear ist, ist dies gleich 0. □

Zusammenfassung von Abschnitt 5.2

(1) Eine alternierende Abbildung $\omega : V^n \to W$ ist eine multilineare Abbildung, die ein Tupel auf 0 abbildet, sobald mindestens zwei Einträge im Tupel gleich sind.

(2) Den Wert einer alternierenden Abbildung kann man mit Hilfe von „Gauß-Algorithmus“-Schritten auf den Einträgen berechnen.

(3) Beim Permutieren der Einträge wird der Funktionswert mit dem Signum der Permutation multipliziert.

(4) Es gilt die Transformationsformel, in dem die Determinante vorkommt.

5.3. Determinanten

Der Ausdruck Δ in Satz 5.2.5 bekommt nun einen eigenen Namen:

Definition 5.3.1 (Die Determinante einer Matrix)**.**
Es sei $n \in \mathbb{N}$ und $A = (a_{i,j})_{i,j \in \{1,\ldots,n\}} \in \mathbb{K}^{n \times n}$ eine quadratische Matrix mit Einträgen aus einem Körper $\mathbb{K}$. Dann ist die *Determinante* von A definiert als

$$\det(A) := \sum_{\sigma \in \mathcal{S}(n)} \mathrm{sgn}(\sigma) \cdot a_{\sigma(1),1} \cdots a_{\sigma(n),n}.$$

Diese Formel ist auch als *Leibniz-Formel*[2] bekannt.

Bemerkung 5.3.2. (a) Es gibt – neben der hier verwendeten Leibniz-Formel auch andere Möglichkeiten, die Determinante einer $(n \times n)$-Matrix zu definieren. Es stellt sich aber heraus, dass sie alle äquivalent sind (siehe Satz 5.3.12).

(b) Ausgehend von dieser Definition der Determinante ist erst einmal überhaupt nicht klar, warum dies – außer der Verwendung in der Transformationsformel (Satz 5.2.5) – ein sinnvolles oder interessantes Konzept sein sollte. Dies wird hoffentlich später[3] klar.

(c) Die hier vorgestellte Leibniz-Formel ist für praktische Berechnungen von Determinanten völlig ungeeignet, sobald $n > 3$ ist, da die Anzahl der Summanden $n!$ ist und dies schneller als exponentiell wächst. Glücklicherweise gibt es ein Verfahren, mit dem die Determinante von Matrizen beliebiger Größe in vertretbarer Zeit berechnen kann. Dies werden wir in Bemerkung 5.3.8 vorstellen.

Beispiel 5.3.3.
Es sei $\mathbb{K}$ ein Körper.

(a) Für $n = 1$ gilt $\det(a) = a$.

(b) Für $n = 2$ erhalten wir direkt die Formel

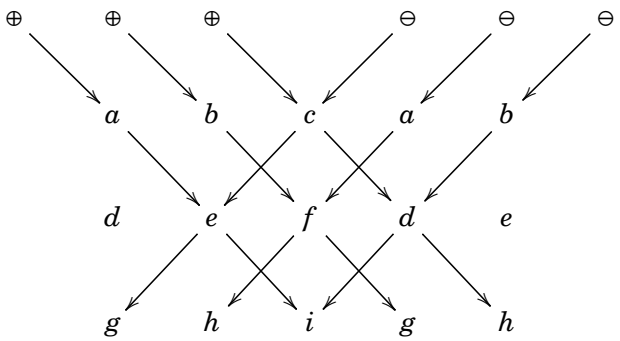
$$\det\begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc,$$

die bereits in Beispiel 4.3.5 gesehen haben.

(c) Für $n = 3$ erhalten wir die sogenannte *Regel von Sarrus*[4]:

$$\det\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = aei + bfg + cdh - gec - hfa - idb,$$

die man sich wie folgt merken kann:



WARNUNG! In Dimensionen ≥ 4 gilt keine Regel dieser Art!

[2] nach GOTTFRIED WILHELM LEIBNIZ, deutscher Philosoph und Mathematiker, 1646–1716
[3] z.B. in Satz 5.3.10
[4] PIERRE FRÉDÉRIC SARRUS, französischer Mathematiker, 1798–1861

Aus der Leibniz-Formel (Definition 5.3.1) folgt nun relativ direkt die folgende wichtige Rechenregel:

Proposition 5.3.4 (Determinanten ändern sich nicht beim Transponieren)**.**
Es sei $A \in \mathbb{K}^{n\times n}$ mit $n \in \mathbb{N}$ und $\mathbb{K}$ ein Körper. Dann gilt

$$\det\left(A^\top\right) = \det(A).$$

Beweis. Zu gegebener Matrix

$$A = (a_{i,j})_{i,j\in\{1,\dots,n\}} \in \mathbb{K}^{n\times n}$$

ist die Transponierte gegeben als

$$A^\top = (a_{j,i})_{i,j\in\{1,\dots,n\}}.$$

Die Determinante von $A^\top$ ist nach Definition 5.3.1 gegeben durch:

$$\det\left(A^\top\right) := \sum_{\sigma\in\mathscr{S}(n)} \operatorname{sgn}(\sigma)\cdot a_{1,\sigma(1)}\cdots a_{n,\sigma(n)}.$$

In jeder Gruppe $(G,*)$ gilt, dass die Inversionsabbildung

$$G \to G, \quad u \mapsto u^{-1}$$

bijektiv ist. Wenn wir also in der Summe über alle $\sigma \in \mathscr{S}(n)$ summieren, so können wir in den Summanden σ durch σ^{-1} ersetzen und ändern den Wert nicht:

$$\det\left(A^\top\right) := \sum_{\sigma\in\mathscr{S}(n)} \operatorname{sgn}\left(\sigma^{-1}\right)\cdot a_{1,\sigma^{-1}(1)}\cdots a_{n,\sigma^{-1}(n)}.$$

Da $\operatorname{sgn} : (\mathscr{S}(n),*) \to (\{-1,1\},\cdot)$ ein Gruppenhomomorphismus ist, gilt mit Lemma 3.2.8, dass

$$\operatorname{sgn}\left(\sigma^{-1}\right) = (\operatorname{sgn}(\sigma))^{-1} = \frac{1}{\operatorname{sgn}(\sigma)} = \operatorname{sgn}(\sigma),$$

weil die Zahlen 1 und -1 beide identisch mit ihren Kehrwerten (=multiplikativen Inversen) sind. Das heißt, wir haben nun:

$$\det\left(A^\top\right) := \sum_{\sigma\in\mathscr{S}(n)} \operatorname{sgn}(\sigma)\cdot a_{1,\sigma^{-1}(1)}\cdots a_{n,\sigma^{-1}(n)}.$$

Für gegebenes $\sigma \in \mathscr{S}(n)$ lautet der entsprechende Summand:

$$\operatorname{sgn}(\sigma)\cdot a_{1,\sigma^{-1}(1)}\cdots a_{n,\sigma^{-1}(n)}.$$

Da $\sigma^{-1} : \{1,\dots,n\} \to \{1,\dots,n\}$ bijektiv ist, wird jede Spalte $j \in \{1,\dots,n\}$ hier genau einmal als $\sigma^{-1}(i)$ vorkommen, nämlich für $i = \sigma(j)$. Wir können denselben Ausdruck – nach Umsortieren[5] – schreiben also als:

$$\operatorname{sgn}(\sigma)\cdot a_{\sigma(1),1}\cdots a_{\sigma(n),n}.$$

Wenn wir dies bei allen Summanden machen, erhalten wir schließlich:

$$\det\left(A^\top\right) := \sum_{\sigma\in\mathscr{S}(n)} \operatorname{sgn}(\sigma)\cdot a_{\sigma(1),1}\cdots a_{\sigma(n),n}.$$

Das war zu zeigen. □

[5] Glücklicherweise ist die Multiplikation in einem Körper ja kommutativ und assoziativ.

Für besonders schöne Matrizen lässt sich die Determinante ganz leicht berechnen:

Definition 5.3.5.
Es sei $n \in \mathbb{N}$ und $\mathbb{K}$ ein Körper. Eine quadratische Matrix $A = (a_{i,j})_{i,j\in\{1,\dots,n\}} \in \mathbb{K}^{n\times n}$ heiße

(a) *obere Dreiecksmatrix*, wenn alle Einträge unterhalb der Hauptdiagonale 0 sind, d.h. wenn gilt
$$\forall i > j : a_{i,j} = 0$$

(b) *untere Dreiecksmatrix*, wenn alle Einträge oberhalb der Hauptdiagonale 0 sind, d.h. wenn gilt
$$\forall i < j : a_{i,j} = 0$$

(c) *Diagonalmatrix*, wenn alle Einträge außer der auf der Hauptdiagonale 0 sind, d.h. wenn gilt
$$\forall i \neq j : a_{i,j} = 0$$

Proposition 5.3.6.
Es sei $n \in \mathbb{N}$ und $\mathbb{K}$ ein Körper. Angenommen eine Matrix $A = (a_{i,j})_{i,j\in\{1,\dots,n\}} \in \mathbb{K}^{n\times n}$ ist eine obere Dreiecksmatrix, eine untere Dreiecksmatrix oder eine Diagonalmatrix. Dann ist die Determinante das Produkt der Diagonaleinträge:

$$\det A = a_{1,1} \cdots a_{n,n}.$$

Insbesondere ist die Determinante der Einheitsmatrix gleich 1.

Beweis. Es reicht, die Aussage für obere Dreiecksmatrizen zu beweisen; jede Diagonalmatrix ist insbesondere eine obere (und untere) Dreiecksmatrix und der Fall der unteren Dreiecksmatrix folgt sofort aus Proposition 5.3.4.

Es sei nun also A eine obere Dreiecksmatrix und $\sigma \in \mathscr{S}(n)$ eine beliebige Permutation. Wir wollen den Ausdruck
$$a_{\sigma(1),1} \cdots a_{\sigma(n),n}$$
berechnen. Falls es ein $j \in \{1,\dots,n\}$ gibt mit $\sigma(j) > j$, so ist der entsprechende Term $a_{\sigma(j),j}$ gleich 0, weil A eine obere Dreiecksmatrix ist und somit wird das ganze Produkt gleich 0.

Es bleiben also nur solche Summanden übrig, bei denen $\sigma(j) \leq j$ für alle $j \in \{1,\dots,n\}$ gilt. Hieraus folgt aber mit der Injektivität von σ direkt, dass $\sigma = \mathrm{id}_{\{1,\dots,n\}}$ sein muss.

Das bedeutet, dass sich die Summe in der Leibniz-Formel (Definition 5.3.1) auf einen einzigen Summanden reduziert:

$$\det(A) := \sum_{\sigma\in\mathscr{S}(n)} \operatorname{sgn}(\sigma) \cdot a_{\sigma(1),1} \cdots a_{\sigma(n),n} = \underbrace{\operatorname{sgn}(\mathrm{id}_{\{1,\dots,n\}})}_{=1} \cdot a_{1,1} \cdots a_{n,n}.$$

Das war zu zeigen. □

Wir wollen nun alle Erkenntnisse über alternierende Abbildungen auf Determinanten anwenden. Dazu verwenden wir den folgenden Satz:

Satz 5.3.7.
Es sei $\mathbb{K}$ ein Körper und $n \in \mathbb{N}$. Dann gilt:

(a) Die Abbildung

$$\det : \mathbb{K}^{n\times n} \to \mathbb{K}$$

ist alternierend in den Spalten*, d.h. die Abbildung*

$$(\mathbb{K}^n)^n \to \mathbb{K}, \quad (v_1,\dots,v_n) \mapsto \det\left(\left|\begin{array}{c|c|c} v_1 & \cdots & v_m \end{array}\right|\right),$$

die n Spaltenvektoren aneinanderklebt und dann die Determinantenabbildung auf die so erzeugte quadratische Matrix anwendet, ist eine alternierende multilineare Abbildung (Definition 5.2.1).

(b) Die Abbildung

$$\det : \mathbb{K}^{n\times n} \to \mathbb{K}$$

ist außerdem alternierend in den Zeilen*, d.h. die Abbildung*

$$(\mathbb{K}^{1\times n})^n \to \mathbb{K}, \quad (z_1,\dots,z_n) \mapsto \det\begin{pmatrix} \overline{} \\ \hline z_1 \\ \hline \vdots \\ \hline z_n \\ \hline \end{pmatrix}$$

die n Zeilenvektoren aneinanderklebt und dann die Determinantenabbildung auf die so erzeugte quadratische Matrix anwendet, ist eine alternierende multilineare Abbildung (Definition 5.2.1).

Achtung: Die Abbildung $\det : \mathbb{K}^{n\times n} \to \mathbb{K}$ ist *nicht linear* (außer für $n = 1$)!

Beweis. Wir zeigen Teil (a); Teil (b) folgt dann sofort mit Proposition 5.3.4.

Nach Definition 5.3.1 gilt:

$$\det\begin{pmatrix} a_{1,1} & \dots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \dots & a_{n,n} \end{pmatrix} := \sum_{\sigma\in\mathscr{S}(n)} \operatorname{sgn}(\sigma)\cdot a_{\sigma(1),1}\cdots a_{\sigma(n),n}.$$

Für jedes $j \in \{1,\dots,n\}$ gilt: Jeder einzelne Summand enthält genau einen Faktor aus der j-ten Spalte. Somit ist jeder Summand linear in der j-ten Spalte.

Also ist die Determinante insgesamt linear in jeder Spalte, also multilinear in den Spalten.

Es bleibt zu zeigen, dass diese multilineare Abbildung auch alternierend in den Spalten ist. Nehmen wir dazu also an, dass es $k,l \in \{1,\dots,n\}$ gibt mit $k < l$, sodass die k-te und die l-te Spalte der Matrix A identisch sind. Wir werden dann zeigen, dass $\det(A) = 0$.

Es sei $\tau \in \mathscr{S}(n)$ die Transposition auf der Menge $\{1,\dots,n\}$, die die Zahlen k und l vertauscht.

Nach Lemma 5.1.6 gilt: $\det(\tau) = -1$.

Jede Permutation $\sigma \in \mathscr{S}(n)$ ist entweder gerade ($\operatorname{sgn}(\sigma) = 1$) oder ungerade ($\operatorname{sgn}(\sigma) = -1$). Die Menge der geraden Permutationen haben wir mit $\mathscr{A}(n)$ bezeichnet. Die Menge der ungeraden Permutationen lässt sich nach Lemma 5.1.11 schreiben als:

$$\{\sigma \in \mathscr{S}(n) \mid \operatorname{sgn}(\sigma) = -1\} = \{\sigma\circ\tau \mid \sigma \in \mathscr{A}(n)\}.$$

Nun benutzen wir diese Darstellung, um die Summe aus Definition 5.3.1 wie folgt zu zerlegen:

$$\begin{aligned}
\det A &= \sum_{\sigma\in\mathcal{S}(n)} \operatorname{sgn}(\sigma)\cdot a_{\sigma(1),1}\cdots a_{\sigma(n),n} \\
&= \sum_{\operatorname{sgn}(\sigma)=1} \operatorname{sgn}(\sigma)\cdot a_{\sigma(1),1}\cdots a_{\sigma(n),n} + \sum_{\operatorname{sgn}(\sigma)=-1} \operatorname{sgn}(\sigma)\cdot a_{\sigma(1),1}\cdots a_{\sigma(n),n} \\
&= \sum_{\sigma\in\mathcal{A}(n)} \operatorname{sgn}(\sigma)\cdot a_{\sigma(1),1}\cdots a_{\sigma(n),n} + \sum_{\sigma\in\mathcal{A}(n)} \operatorname{sgn}(\sigma\circ\tau)\cdot a_{(\sigma\circ\tau)(1),1}\cdots a_{(\sigma\circ\tau)(n),n} \\
&= \sum_{\sigma\in\mathcal{A}(n)} \operatorname{sgn}(\sigma)\cdot a_{\sigma(1),1}\cdots a_{\sigma(n),n} + \sum_{\sigma\in\mathcal{A}(n)} \operatorname{sgn}(\sigma)\cdot \underbrace{\operatorname{sgn}(\tau)}_{=-1}\cdot a_{\sigma(\tau((1)),1}\cdots a_{\sigma(\tau(n)),n} \\
&= \sum_{\sigma\in\mathcal{A}(n)} \operatorname{sgn}(\sigma)\cdot a_{\sigma(1),1}\cdots a_{\sigma(n),n} - \sum_{\sigma\in\mathcal{A}(n)} \operatorname{sgn}(\sigma)\cdot a_{\sigma(\tau(1)),1}\cdots a_{\sigma(\tau(k)),k}\cdots a_{\sigma(\tau(l)),l}\cdots a_{\sigma(\tau(n)),n}.
\end{aligned}$$

Weil die Transposition τ gerade k und l vertauscht und diese beiden Spalten gleich sind, gilt:

$$a_{\sigma(\tau(k)),k} = a_{\sigma(l),k} = a_{\sigma(l),l} \quad \text{und} \quad a_{\sigma(\tau(l)),l} = a_{\sigma(k),l} = a_{\sigma(k),k}$$

Wenn aber $j \in \{k,l\}$ ist, dann gilt $\tau(j) = j$. Also ist insgesamt:

$$a_{\sigma(\tau(1)),1}\cdots a_{\sigma(\tau(k)),k}\cdots a_{\sigma(\tau(l)),l}\cdots a_{\sigma(\tau(n)),n} = a_{\sigma(1),1}\cdots a_{\sigma(n),n}$$

und die Determinante wird 0. Das war zu zeigen. □

Bemerkung 5.3.8 (Berechnung einer Determinanten)**.**
Jetzt, wo wir wissen, dass die Determinante alternierend in den Zeilen (Satz 5.3.7(b)) ist, können wir – nach Lemma 5.2.3 – den Gauß-Algorithmus auf die Zeilen anwenden, um die Determinante zu berechnen. Bei Schritt (G1) (Addieren des μ-fachen einer Zeile) ändert sich die Determinante überhaupt nicht, bei Schritt (G2) (Vertauschen zweier Zeilen) wird die Determinante mit (-1) multipliziert und statt – wie beim Gauß-Algorithmus üblich – eine Zeile durch eine Zahl $\lambda \neq 0$ zu teilen, sollte man eher die Denkweise verwenden, dass man die Linearität in den Zeilen verwendet, und Skalare aus der Determinante nach vorne zieht.

Nach endlich vielen Schritten hat man die Matrix in Zeilenstufenform gebracht (erweiterte Zeilenstufenform ist hierfür nicht notwendig). Eine quadratische Matrix in Zeilenstufenform ist immer eine obere Dreiecksmatrix, sodass sich nun die Determinante direkt mit Proposition 5.3.6 berechnen lässt.

Dieses hier skizzierte Vorgehen funktioniert für jede Matrix und ist für große Matrizen (d.h. $n > 3$) deutlich schneller als direkt die Leibniz-Formel (Definition 5.3.1) zu verwenden und zwar sowohl beim Rechnen von Hand – als auch beim Berechnen mit einem Computer. Der Gauß-Algorithmus (und nichts anderes ist das hier ja) ist erstaunlich schnell und mit der richtigen Wahl der Pivot-Elemente auch numerisch einigermaßen stabil, d.h. beim gerundeten Rechnen mit Gleitkommazahlen lassen sich die Fehler einigermaßen im Griff halten. Die Leibniz-Formel auf der anderen Seite hat $n!$ Summanden und das ist schon für noch relativ kleine n für einen Computer nicht mehr in sinnvoller Zeit machbar.

Oft kann man die Rechnung auch noch weiter vereinfachen:
Da die Determinante ja auch alternierend in den Spalten (Satz 5.3.7(a)) ist, können wir während der Rechnung auch jederzeit Spaltenumformungen durchführen, solange wir dieselben Regeln wie bei den Zeilenumformungen berücksichtigen, d.h. Faktoren entsprechend vor die Determinante ziehen.

All diese möglichen Rechenschritte lassen sich auch noch mit den (noch nicht eingeführten) Laplace-Entwicklungen nach Zeilen oder Spalten (siehe Satz 5.3.15) kombinieren, sodass es eine Unzahl an unterschiedlichen Rechenwegen gibt, eine gegebene Determinante auszurechnen.

Außerdem kann man die Rechnung auch sofort beenden, wenn man eine Nullzeile oder -spalte sieht (Satz 5.3.7), sobald zwei Zeilen oder Spalten gleich sind oder sobald man sieht, dass die Zeilen oder Spalten linear abhängig sind (siehe Korollar 5.2.6).

Wir kommen nun zu den wirklich wichtigen Eigenschaften der Determinantenabbildung:

Satz 5.3.9 (Die Determinante ist multiplikativ)**.**
Es sei $\mathbb{K}$ ein Körper und $n \in \mathbb{N}$. Dann gilt:

$$\forall A, B \in \mathbb{K}^{n\times n} : \det(AB) = \det(A)\det(B).$$

Beweis. Wir bezeichnen die Spalten der Matrix A mit $u_1, \dots, u_n \in \mathbb{K}^n$:

$$A = \begin{pmatrix} | & & | \\ u_1 & \cdots & u_n \\ | & & | \end{pmatrix}.$$

Die Spalten der Produktmatrix AB bezeichnen wir mit $v_1, \dots, v_n \in \mathbb{K}^n$:

$$AB = \begin{pmatrix} | & & | \\ v_1 & \cdots & v_n \\ | & & | \end{pmatrix}.$$

Nach Definition des Matrixproduktes gilt nun:

$$\forall k \in \{1, \dots, n\} : v_k = \sum_{i=1}^{n} b_{i,j} u_i,$$

wobei $(b_{i,j})$ wie gewöhnlich die Einträge der Matrix B sind.

Nach Satz 5.3.7(a) ist die Abbildung

$$\omega : (\mathbb{K}^n)^n \to \mathbb{K}, \quad (w_1, \dots, w_n) \mapsto \det\begin{pmatrix} | & & | \\ w_1 & \cdots & w_n \\ | & & | \end{pmatrix}$$

alternierend. Wir können somit die Transformationsformel für alternierende Abbildungen (Satz 5.2.5) anwenden und erhalten:

$$\det(AB) = \omega(v_1, \dots, v_n) = \Delta \cdot \omega(u_1, \dots, u_n) = \Delta \det(A)$$

mit $\Delta = \det(B)$. □

Der folgende Satz ist die wohl wichtigste Anwendung der Determinanten für diese Veranstaltung:

Satz 5.3.10 (Kriterium für die Invertierbarkeit)**.**
Es sei $\mathbb{K}$ ein Körper und $n \in \mathbb{N}$. Dann gilt:

$$\forall A \in \mathbb{K}^{n\times n} : (A \text{ ist invertierbar}) \iff (\det(A) \neq 0).$$

Wenn A invertierbar ist, dann gilt: $\det(A^{-1}) = \frac{1}{\det(A)}$.

Die Gruppe der invertierbaren $(n \times n)$-Matrizen ist somit

$$\mathrm{GL}(n, \mathbb{K}) = \left\{ A \in \mathbb{K}^{n\times n} \;\middle|\; \det(A) \neq 0 \right\}.$$

Beweis. Wir müssen zwei Implikationen zeigen:
„$\Longrightarrow$“:
Wenn A invertierbar ist, dann gibt es ein $B \in \mathbb{K}^{n\times n}$ mit $AB = BA = \mathbb{1}_n$. Mit Satz 5.3.9 gilt nun:

$$\det(A)\det(B) = \det(AB) = \det(\mathbb{1}_n) = 1.$$

Also muss $\det(A), \det(B) \neq 0$ gelten und $\det(A) = \frac{1}{\det(B)}$.
„$\Longleftarrow$“:
Nach Satz 5.3.7(a) ist die Determinante alternierend in den Spalten. Da wir annehmen, dass die Determinante nicht 0 ist, folgt mit Korollar 5.2.6, dass die Spalten linear unabhängig sind, folglich hat die Matrix Rang n. Dies ist aber äquivalent zur Invertierbarkeit von A nach Satz 4.3.4. □

Korollar 5.3.11.
Es sei $\mathbb{K}$ ein Körper und $n \in \mathbb{N}$. Dann gilt:

$$(\mathrm{GL}(n,\mathbb{K}),\cdot) \to (\mathbb{K}^\times = \mathbb{K}\setminus\{0\},\cdot), \quad A \mapsto \det(A)$$

ist ein Gruppenhomomorphismus. Insbesondere ist der Kern

$$\mathrm{SL}(n,\mathbb{K}) = \left\{A \in \mathbb{K}^{n\times n} \;\middle|\; \det(A) = 1\right\}$$

eine Untergruppe von $\mathrm{GL}(n,\mathbb{K})$*, genannt die* spezielle lineare Gruppe.

Oft wird die Determinante nicht mit Hilfe der Leibniz-Formel (Definition 5.3.1), sondern axiomatisch eingeführt als die einzige Abbildung $\delta : \mathbb{K}^{n\times n} \to \mathbb{K}$, die alternierend in den Zeilen ist und die Einheitsmatrix auf 1 abbildet. Wir wollen nun zeigen, dass dies äquivalent zu dem von uns gewählten Ansatz ist:

Satz 5.3.12 (Charakterisierung der Determinantenabbildung)**.**
Es sei $\mathbb{K}$ ein Körper und $n \in \mathbb{N}$ und $\delta : \mathbb{K}^{n\times n} \to \mathbb{K}$ eine Abbildung. Dann sind äquivalent:

(i) *δ ist alternierend in den Spalten und $\delta(\mathbb{1}_n) = 1$.*

(ii) *δ ist alternierend in den Zeilen und $\delta(\mathbb{1}_n) = 1$.*

(iii) *$\delta = \det$.*

Beweis. Nach Satz 5.3.7 folgt (i) und (ii) aus (iii).
Zeigen wir nun „(i) $\Longrightarrow$ (iii)“:
Dann ist die Abbildung δ alternierend in den Spalten, d.h.

$$\omega : (\mathbb{K}^n)^n \to \mathbb{K}, \quad (w_1,\ldots,w_n) \mapsto \delta\left(\begin{vmatrix} & & \\ w_1 & \cdots & w_n \\ & & \end{vmatrix}\right)$$

ist alternierend. Dann folgt – genau wie im Beweis von Satz 5.3.9 – mit Hilfe der Transformationsformel für alternierende Abbildungen:

$$\delta(AB) = \det(B)\delta(A) \quad \text{für alle } A,B \in \mathbb{K}^{n\times n}.$$

Mit $A = \mathbb{1}_n$ folgt nun: $\delta(B) = \det(B)\underbrace{\delta(\mathbb{1}_n)}_{=1}$. Da B beliebig gewählt war, folgt $\delta = \det$.

Zeigen wir nun „(ii) $\Longrightarrow$ (iii)“:
Wir definieren $\eta : \mathbb{K}^{n\times n} \to \mathbb{K}, \quad A \mapsto \det\left(A^\top\right)$. Dann ist η alternierend in den Spalten und nach der soeben bewiesenen Rechnung gilt: $\eta = \det$.

Damit folgt für alle $A \in \mathbb{K}^{n\times n}$:

$$\delta(A) = \eta(A^\top) = \det(A^\top) = \det(A).$$

Das war zu zeigen. □

Wir hatten in Definition 5.3.1 die Determinante mit Hilfe des Signums definiert. Umgekehrt kann man aber auch das Signum mit Hilfe der Determinante berechnen:

Proposition 5.3.13 (Signum als Spezialfall der Determinante)**.**
Es sei $n \in \mathbb{N}$. Für $\sigma \in \mathcal{S}(n)$ definieren wir die Permutationsmatrix *von σ als*

$$A_\sigma := \left(\left| e_{\sigma(1)} \right| \cdots \left| e_{\sigma(n)} \right| \right) \in \mathbb{R}^{n\times n}.$$

In der j-ten Spalte von A steht also der $\sigma(j)$-te Standardbasisvektor.

Dann gilt:

$$\operatorname{sgn}(\sigma) = \det(A_\sigma).$$

Beweis. Wir setzen wieder

$$\omega : (\mathbb{K}^n)^n \to \mathbb{K}, \quad (v_1, \ldots, v_n) \mapsto \det\left(\left| v_1 \right| \cdots \left| v_n \right| \right).$$

Durch Anwenden von Satz 5.3.7 und Lemma 5.2.4 folgt dann sofort:

$$\det(A_\sigma) = \omega\left(a_{\sigma(1)}, \ldots, a_{\sigma(n)}\right) = \operatorname{sgn}(\sigma) \cdot \omega(e_1, \ldots, e_n) = \operatorname{sgn}(\sigma) \cdot \underbrace{\det(\mathbb{1}_n)}_{=1} = \operatorname{sgn}(\sigma).$$ □

Definition 5.3.14 (Streichungsmatrix)**.**
Es sei $n \in \mathbb{N}, n \geq 2$ und $A = (a_{i,j})_{i,j\in\{1,\ldots,n\}} \in \mathbb{K}^{n\times n}$ mit Einträgen aus einem Körper $\mathbb{K}$. Für jedes $(k,l) \in \{1,\ldots,n\} \times \{1,\ldots,n\}$ definieren wir die *Streichungsmatrix* von A bezüglich (k,l) als die $(n-1)\times(n-1)$-Matrix, die durch das Streichen der k-ten Zeile und der l-ten Spalte entsteht. In Formeln:

$$\mathrm{St}_{(k,l)}(A) := (b_{i,j})_{i,j\in\{1,\ldots,n-1\}} \in \mathbb{K}^{n-1\times n-1}$$

mit

$$b_{i,j} = \begin{cases} a_{i,j} & \text{für} i < k \text{ und } j < l \\ a_{i+1,j} & \text{für} i \geq k \text{ und } j < l \\ a_{i,j+1} & \text{für} i < k \text{ und } j \geq l \\ a_{i+1,j+1} & \text{für} i \geq k \text{ und } j \geq l \end{cases}$$

Satz 5.3.15 (Entwicklungssatz von Laplace[6])**.**
Es sei $\mathbb{K}$ ein Körper und $n \in \mathbb{N}$ mit $n \geq 2$. Gegeben sei eine Matrix $A = (a_{i,j})_{i,j\in\{1,\ldots,n\}} \in \mathbb{K}^{n\times n}$.

[6] nach Pierre-Simon Laplace, frz. Universalgelehrter, 1749–1827

(a) *Für jedes $l \in \{1,\dots,n\}$ gilt:*

$$\det(A) = \sum_{k=1}^{n} (-1)^{k+l} a_{k,l} \det\left(\mathrm{St}_{(k,l)}(A)\right)$$

Diese Art der Determinantenberechnung nennt man auch Entwicklung nach der l-ten Spalte.

(b) *Für jedes $k \in \{1,\dots,n\}$ gilt:*

$$\det(A) = \sum_{l=1}^{n} (-1)^{k+l} a_{k,l} \det\left(\mathrm{St}_{(k,l)}(A)\right)$$

Diese Art der Determinantenberechnung nennt man auch Entwicklung nach der k-ten Zeile.

Für praktische Berechnungen ist diese Laplace-Entwicklung nur sinnvoll, wenn eine Zeile (oder Spalte) gewählt wird, in der sehr viele Nullen stehen, da in diesem Falle viele der Unterdeterminanten $\det\left(\mathrm{St}_{(k,l)}(A)\right)$ gar nicht erst berechnet werden müssen.

Definition 5.3.16 (Adjunkte)**.**
Es sei $\mathbb{K}$ ein Körper und $n \in \mathbb{N}$ mit $n \geq 2$. Gegeben sei eine Matrix $A = (a_{i,j})_{i,j\in\{1,\dots,n\}} \in \mathbb{K}^{n\times n}$.
Die *Adjunkte* von A ist die folgende Matrix:

$$A^\sharp := (\alpha_{i,j})_{i,j\in\{1,\dots,n\}} \in \mathbb{K}^{n\times n}$$

mit

$$\alpha_{i,j} := (-1)^{i+j} \det(\mathrm{St}_{(j,i)}(A)).$$

Man beachte die Verdrehung der Indizes in der Definition von $\alpha_{i,j}$!

Beispiel.
Es sei $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbb{K}^{2\times 2}$. Dann gilt $\mathrm{St}_{(1,2)}(A) = c$, und daher $(-1)^{2+1}\det(\mathrm{St}_{(1,2)}(A)) = -c$. Durch vier solche Rechnungen erhält man

$$A^\sharp = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Ist A invertierbar, so haben wir andererseits in Beispiel 4.3.5 gesehen, dass

$$A^{-1} = \frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \frac{1}{\det(A)} \cdot A^\sharp.$$

Das gilt ganz allgemein:

Satz 5.3.17 (Cramersche[7] Regel)**.**

Es sei $\mathbb{K}$ ein Körper und $n \in \mathbb{N}$ mit $n \geq 2$. Ist $A = (a_{i,j})_{i,j\in\{1,\dots,n\}} \in \mathbb{K}^{n\times n}$, so gilt

$$A \cdot A^\sharp = \det(A) \cdot \mathbb{1}_n.$$

Ist speziell A invertierbar, so gilt

$$A^{-1} = \frac{1}{\det(A)} \cdot A^\sharp.$$

[7] nach GABRIEL CRAMER, Schweizer Mathematiker, 1704–1752

Zusammenfassung von Abschnitt 5.3

(1) Die Determinantenabbildung $\det : \mathbb{K}^{n\times n} \to \mathbb{K}$ ordnet jeder quadratischen Matrix ein Element im Grundkörper zu.

(2) Die Definition über die Leibniz-Formel ist für $n > 3$ nicht praktikabel.

(3) Die Determinante einer oberen oder unteren Dreiecksmatrix (oder sogar einer Diagonalmatrix) ist das Produkt der Diagonaleinträge.

(4) Die Determinantenabbildung ist alternierend in den Zeilen und in den Spalten. Insbesondere kann man Gauß-Schritte auf Zeilen und Spalten verwenden, um die Determinante zu berechnen.

(5) Man kann die Determinante auch durch Laplace-Entwicklung nach einer Zeile oder Spalte berechnen. Für praktische Berechnung ist dies nur sinnvoll, wenn die entsprechende Zeile oder Spalte viele Nullen enthält.

(6) Die Determinante ist multiplikativ: $\det(AB) = \det(A)\det(B)$ für quadratische Matrizen A, B.

(7) Die Determinante ändert sich nicht, wenn eine Matrix transponiert wird: $\det(A^\top) = \det(A)$.

(8) Eine Matrix ist genau dann invertierbar, wenn die Determinante ungleich 0 ist.

(9) Die Cramersche Regel zur Berechnung der Inversen ist für $n > 2$ nicht praktikabel.

5.4. Polynome

Bemerkung 5.4.1.
Wir betrachten die Menge $\mathbb{R}^\mathbb{R}$ aller Funktionen von $\mathbb{R}$ nach $\mathbb{R}$. Auf dieser Menge haben wir (mindestens) drei naheliegende Operationen:

- Die Addition $+ : \mathbb{R}^\mathbb{R} \times \mathbb{R}^\mathbb{R} \to \mathbb{R}^\mathbb{R}$, die zwei Funktionen $f, g : \mathbb{R} \to \mathbb{R}$ auf deren punktweise Summe $f + g : \mathbb{R} \to \mathbb{R}$ abbildet.
- Die Multiplikation $\bullet : \mathbb{R}^\mathbb{R} \times \mathbb{R}^\mathbb{R} \to \mathbb{R}^\mathbb{R}$, die zwei Funktionen $f, g : \mathbb{R} \to \mathbb{R}$ auf deren punktweises Produkt $fg : \mathbb{R} \to \mathbb{R}$ abbildet.
- Die skalare Multiplikation $\cdot : \mathbb{R} \times \mathbb{R}^\mathbb{R} \to \mathbb{R}^\mathbb{R}$, die eine Zahl $\lambda \in \mathbb{R}$ und eine Funktion $f : \mathbb{R} \to \mathbb{R}$ auf die skalierte Funktion $\lambda f : \mathbb{R} \to \mathbb{R}$ abbildet.

Mit der Addition und der skalaren Multiplikation wird $\mathbb{R}^\mathbb{R}$ ein $\mathbb{R}$-Vektorraum, mit der Addition und der punktweisen Multiplikation wird $\mathbb{R}^\mathbb{R}$ ein Ring.

Es stellt sich heraus, dass dies relativ häufig vorkommt, dass eine Menge sowohl eine Vektorraumstruktur trägt als auch eine Ringstruktur, wobei beide Strukturen dieselbe Addition verwenden und die beiden Multiplikationen in einer gewissen Weise verträglich sind. Deswegen lohnt es sich, eine neue algebraische Struktur einzuführen. Bei der Benennung dieser algebraischen Struktur waren die Algebraiker diesmal aber relativ einfallslos...

Definition 5.4.2 (Algebren)**.**
Es sei $\mathbb{K}$ ein Körper.

(a) Eine $\mathbb{K}$-*Algebra* ist ein Tupel $(\mathbf{A}, +, \bullet, \cdot)$ mit der Eigenschaft, dass

(i) $(\mathbf{A}, +, \bullet)$ ein Ring ist,

(ii) $(\mathbf{A}, +, \cdot)$ ein Vektorraum über $\mathbb{K}$ ist und

(iii) die beiden Multiplikationen auf folgende Weise verträglich sind:

$$\forall a, b \in \mathbf{A}, \forall \lambda \in \mathbb{K} : (\lambda \cdot a) \bullet b = a \bullet (\lambda \cdot b) = \lambda \cdot (a \bullet b).$$

Eine $\mathbb{K}$-Algebra $(\mathbf{A}, +, \bullet, \cdot)$ heiße *kommutativ*, wenn der Ring $(\mathbf{A}, +, \bullet)$ kommutativ ist. Eine $\mathbb{K}$-Algebra $(\mathbf{A}, +, \bullet, \cdot)$ heiße *endlich dimensional*, wenn der Vektorraum $(\mathbf{A}, +, \cdot)$ endlich dimensional ist. Die *Dimension* einer $\mathbb{K}$-Algebra ist dann die Dimension des zugrundeliegenden $\mathbb{K}$-Vektorraums.

(b) Eine $\mathbb{K}$-Unteralgebra ist genauso definiert, wie man es erwarten würde, wenn man das Konzept von Unterringen und Untervektorräumen schon gesehen hat: Es sei $(\mathbf{A}, +, \bullet, \cdot)$ eine $\mathbb{K}$-Algebra und $\mathbf{B} \subseteq \mathbf{A}$ eine Teilmenge. Dann ist $\mathbf{B}$ eine $\mathbb{K}$-Unteralgebra von $(\mathbf{A}, +, \bullet, \cdot)$, wenn $\mathbf{B}$ ein Unterring von $(\mathbf{A}, +, \bullet)$ und ein Untervektorraum von $(\mathbf{A}, +, \cdot)$ ist. Eine $\mathbb{K}$-Unteralgebra ist mit den eingeschränkten[8] Operationen wieder eine $\mathbb{K}$-Algebra.

(c) Eine Abbildung $\varphi : \mathbf{A} \to \mathbf{B}$ zwischen zwei $\mathbb{K}$-Algebren $(\mathbf{A}, +, \bullet, \cdot)$ und $(\mathbf{B}, +, \bullet, \cdot)$ ist ein $\mathbb{K}$-Algebra-Homomorphismus, wenn sie ein Ring-Homomorphismus zwischen $(\mathbf{A}, +, \bullet)$ und $(\mathbf{B}, +, \bullet)$ und ein $\mathbb{K}$-Vektorraum-Homomorphismus (also eine $\mathbb{K}$-lineare Abbildung) zwischen den Vektorräumen $(\mathbf{A}, +, \cdot)$ und $(\mathbf{B}, +, \cdot)$ ist.

Ein $\mathbb{K}$-Algebra-Isomorphismus ist ein bijektiver $\mathbb{K}$-Algebra-Homomorphismus und zwei $\mathbb{K}$-Algebren heißen isomorph, wenn es einen $\mathbb{K}$-Algebra-Isomorphismus zwischen ihnen gibt.

Bemerkung 5.4.3.
Diese Definition ist überhaupt nicht einheitlich. Manchmal wird auch deutlich weniger gefordert. Zum Beispiel ist es – wie schon bei Ringen – nicht einheitlich, ob ein multiplikatives Inverses gefordert wird.

Manchmal wird auch nicht gefordert, dass $(\mathbf{A}, +, \bullet)$ ein Ring ist, d.h. oft wird auf die Assoziativität von $\bullet$ verzichtet. Um sich von solchen Definitionen abzuheben, nennt man eine Algebra, wie wir sie hier definiert haben auch *assoziative Algebra mit Eins*.

Beispiel 5.4.4. (i) Der Grundkörper $\mathbb{K}$ ist immer eine 1-dimensionale kommutative $\mathbb{K}$-Algebra.

(ii) Jede Körpererweiterung $\mathbb{L}$ von $\mathbb{K}$ ist eine kommutative $\mathbb{K}$-Algebra. Insbesondere kann man also $\mathbb{F}_4$ als $\mathbb{F}_2$-Algebra auffassen, oder $\mathbb{R}$ als $\mathbb{Q}$-Algebra.

Der Raum der komplexen Zahlen kann entweder als (unendlich dimensionale) $\mathbb{Q}$-Algebra, als 2-dimensionale $\mathbb{R}$-Algebra oder als 1-dimensionale $\mathbb{C}$-Algebra aufgefasst werden.

(iii) Für jeden Körper $\mathbb{K}$ und jede natürliche Zahl $n \in \mathbb{N}$ ist $\mathbb{K}^{n \times n}$ eine n^2-dimensionale $\mathbb{K}$-Algebra, die für jedes $n > 1$ nichtkommutativ ist.

[8] und koeingeschränkten

(iv) Es sei $\mathbb{K}$ ein Körper und V ein Vektorraum über $\mathbb{K}$. Dann ist die Menge $(\mathrm{End}_{\mathbb{K}}(V), +, \circ)$ aller Vektorraumendomorphismen von V eine $\mathbb{K}$-Algebra. Falls V endlich dimensional ist $\mathsf{B} = (v_1, \ldots, v_n)$ eine geordnete Basis von V ist, dann ist $(\mathrm{End}_{\mathbb{K}}(V), +, \circ)$ als $\mathbb{K}$-Algebra isomorph zu $\mathbb{K}^{n\times n}$ über den $\mathbb{K}$-Algebra-Isomorphismus:

$$\mathrm{End}_{\mathbb{K}}(V) \to \mathbb{K}^{n\times n}, \quad \varphi \mapsto \mathrm{M}_{\mathsf{B},\mathsf{B}}(\varphi).$$

(v) Es sei J eine Menge (z.B. $J = \mathbb{R}$) und $\mathbb{K}$ ein Körper (z.B. $\mathbb{K} = \mathbb{R}$). Dann ist die Menge aller Funktionen von J nach $\mathbb{K}$ eine kommutative $\mathbb{K}$-Algebra, die genau dann endlich dimensional ist, wenn die Menge J endlich ist.

Bemerkung 5.4.5.
Wir kehren zurück zu dem Beispiel $(\mathbb{R}^{\mathbb{R}}, +, \bullet, \cdot)$ aus Bemerkung 5.4.1 der Funktionen von $\mathbb{R}$ nach $\mathbb{R}$. Diese Algebra ist für viele Anwendungen viel zu groß, weil sie viel zu viele Funktionen enthält. Wir wollen deshalb zu der folgenden interessanteren Algebra übergehen:

$$\mathbf{P} := \{f : \mathbb{R} \to \mathbb{R} \mid f \text{ ist eine Polynomfunktion}\}.$$

Eine *Polynomfunktion*[9] ist hierbei eine Funktion $f : \mathbb{R} \to \mathbb{R}$, die sich schreiben lässt als

$$f(t) = a_0 + a_1 t + a_2 t^2 + \cdots + a_n t^n \quad \text{mit } n \in \mathbb{N}_0 \text{ und } a_0, \ldots, a_n \in \mathbb{R}.$$

Man sieht leicht ein, dass Summen, skalare Vielfache und Produkte von Polynomfunktionen wieder Polynomfunktionen sind, also ist $\mathbf{P}$ eine $\mathbb{R}$-Unteralgebra von $\mathbb{R}^{\mathbb{R}}$.

Es ist außerdem möglich, eine Basis von $\mathbf{P}$ als $\mathbb{R}$-Vektorraum anzugeben:

Betrachten wir dazu das Element $X := (\mathrm{id}_{\mathbb{R}} : \mathbb{R} \to \mathbb{R}, \quad t \mapsto t) \in \mathbf{P}$ und die Potenzen davon:

$$X^0 : t \mapsto 1, X^1 : t \mapsto t, X^2 : t \mapsto t^2, X^3 : t \mapsto t^3, \ldots$$

Aus der Definition des Begriffs Polynomfunktion folgt sofort, dass die Menge $\{X^k \mid k \in \mathbb{N}_0\} = \{1, X, X^2, X^3, \ldots\}$ ein Erzeugendensystem für $\mathbf{P}$ als $\mathbb{R}$-Vektorraum ist, d.h. jede Polynomfunktion ist eine Linearkombination dieser Funktionen. Noch zu zeigen wäre, dass diese Darstellung eindeutig ist, d.h. dass $\{X^k \mid k \in \mathbb{N}_0\}$ linear unabhängig ist.

Es sei eine Polynomfunktion

$$f(t) = a_0 + a_1 t + \cdots a_n t^n = \sum_{k=0}^{n} a_k t^k$$

gegeben. Wir werden zeigen, dass diese Darstellung eindeutig ist, d.h. dass die Koeffizienten $a_0, \ldots, a_n$ bereits durch die Funktionswerte von f bestimmt sind.

Man sieht direkt durch Einsetzen, dass $f(0) = a_0$ ist. Folglich ist der konstante Koeffizient a_0 eindeutig bestimmt.

Nun leiten wir die Funktion einmal ab:

$$f'(t) = a_1 + \cdots n a_n t^{n-1} = \sum_{k=1}^{n} k a_k t^{k-1}$$

Wenn wir nun $t = 0$ einsetzen, erhalten wir $f'(0) = a_1$. Somit ist auch der Koeffizient a_1 eindeutig bestimmt.

[9] In deutschen Schulen ist hierfür manchmal auch der sehr irreführende Begriff *ganzrationale Funktion* gebräuchlich... Warum, ist mir bis heute nicht klar geworden...

Wenn wir diese Idee fortsetzen, erhalten wir eine Möglichkeit[10], den k-ten Koeffizienten von f mit Hilfe der k-ten Ableitung von f an der Stelle 0 zu berechnen:

$$a_k = \frac{f^{(k)}(0)}{k!}.$$

Hieraus folgt, dass die Koeffizienten eindeutig bestimmt sind.

Also gilt: Die $\{X^k \mid k \in \mathbb{N}_0\} = \{1, X, X^2, X^3, \ldots\}$ bilden eine Basis des reellen Vektorraums $\mathbf{P}$.

Falls Ihnen dieses Argument nicht gefallen hat, weil Sie keine Differentialrechnung mögen uns sich ein algebraisches Argument gewünscht hätten, verweisen wir auf Korollar 5.4.15(b), wo wir die lineare Unabhängigkeit der Funktionen X^k noch einmal ohne Ableitungen beweisen[11].

Insgesamt erhalten wir also die Erkenntnis:

Die Algebra der Polynomfunktionen $\mathbf{P}$ enthält ein Element $X \in P$, sodass die Potenzen $\{X^k \mid k \in \mathbb{N}_0\}$ eine Basis für den Vektorraum $\mathbf{P}$ bilden.

Dies nehmen wir als Motivation für die folgende Definition:

Definition 5.4.6 (Polynomalgebra)**.**
Es sei $\mathbb{K}$ ein Körper.

(a) Eine $\mathbb{K}$-Algebra $\mathbf{A}$ mit einem Element $X \in \mathbf{A}$ wird (formale) *Polynomalgebra* (oder *Algebra der formalen Polynome*) über $\mathbb{K}$ in der *Unbestimmten* (oder *formalen Variable*) X genannt, wenn die Potenzen $X^0, X^1, X^2, \ldots$ paarweise verschieden sind und $\{X^k \mid k \in \mathbb{N}_0\}$ eine Basis für $\mathbf{A}$ ist.

Wir werden die Bezeichnung $A = \mathbb{K}[X]$ verwenden[12].

(b) Ein Element in einer Polynomalgebra wird (formales) *Polynom* genannt. Es lässt sich also eindeutig als

$$p = \sum_{k=0}^{n} a_k X^k \quad \text{mit } a_0, a_1, \ldots, a_n \in \mathbb{K}$$

schreiben. Die Skalare $a_0, a_1, \ldots, a_n \in \mathbb{K}$ werden die Koeffizienten genannt.

(c) Es sei $p \in \mathbb{K}[X]$. Dann definieren wir den *Grad* von p als:

$$\deg(p) = \max\{k \in \mathbb{N}_0 \mid a_k \neq 0\} \in \mathbb{N}_0.$$

Für das *Nullpolynom* 0 definieren wir $\deg(0) := -\infty$.

Wenn man von der Polynomalgebra $\mathbb{K}[X]$ nur die Vektorraumstruktur betrachtet, nennt man $\mathbb{K}[X]$ auch *Polynomraum*. Ebenso wird das Wort *Polynomring* verwendet, wenn man nur an der Ringstruktur interessiert ist.

[10] Dieses Verfahren ist auch die Methode der Taylor-Koeffizienten bekannt und wird in der Analysis gern benutzt, um Funktionen, die keine Polynomfunktionen sind, durch solche anzunähern.

[11] Es stellt sich heraus: Die einzige Eigenschaft der reellen Zahlen, die wir wirklich brauchen ist die Tatsache, dass $\mathbb{R}$ unendlich viele Elemente hat.

[12] Wenn das spezielle Element nicht X, sondern Y oder irgendwie anders heißt, so wird die Algebra entsprechend mit $\mathbb{K}[Y]$ oder entsprechend bezeichnet.

Bemerkung 5.4.7. Nachdem wir in Definition 5.4.6 definiert haben, was eine Polynomalgebra ist, stellt sich nun die Frage, ob es über jedem Körper $\mathbb{K}$ eine solche gibt. In Bemerkung 5.4.5 haben wir gezeigt, dass die Algebra der Polynomfunktionen $\mathbf{P} \subseteq \mathbb{R}^{\mathbb{R}}$ eine Polynomalgebra über den reellen Zahlen bildet, für den Beweis der linearen Unabhängigkeit haben wir allerdings Methoden der reellen Analysis verwendet, die sich nicht ohne weiteres auf beliebige Körper übertragen lassen.

Es sei $\mathbb{K}$ nun ein endlicher Körper mit $q \in \mathbb{N}$ Elementen, wie z.B. $\mathbb{K} = \mathbb{F}_2$. Dann hat die Algebra aller Funktionen von $\mathbb{K}$ nach $\mathbb{K}$ nur endlich viele Elemente:

$$\left|\mathbb{K}^{\mathbb{K}}\right| = q^q.$$

Eine Polynomalgebra über $\mathbb{K}$ ist aber per Definition –sofern sie existiert – unendlich dimensional, weil sie eine unendliche Basis

$$\{X^0, X^1, X^2, \ldots\}$$

besitzt. Wenn es also über einem endlichen Körper $\mathbb{K}$ eine Polynomalgebra $\mathbb{K}[X]$ gibt, so lässt sie sich nicht aus Funktionen von $\mathbb{K}$ nach $\mathbb{K}$ konstruieren.

Satz 5.4.8 (Existenz und Eindeutigkeit der Polynomalgebra)**.**
Es sei $\mathbb{K}$ ein Körper.

(a) Es existiert eine Polynomalgebra $\mathbf{A} = \mathbb{K}[X]$ in einer Unbestimmten $X \in \mathbf{A}$.

(b) Die Polynomalgebra ist eindeutig bis auf eindeutigen Isomorphismus, d.h. Wenn $A = \mathbb{K}[X]$ und $B = \mathbb{K}[Y]$ Polynomalgebren über $\mathbb{K}$ sind, dann gibt es einen eindeutigen $\mathbb{K}$-Algebra-Isomorphismus

$$\varphi : \mathbb{K}[X] \to \mathbb{K}[Y],$$

mit $\varphi(X) = Y$.

In diesem Sinne ist es also erlaubt, von der *Polynomalgebra zu sprechen, da es im Wesentlichen nur eine solche gibt.*

Beweis. (a)

Wir betrachten den Vektorraum $\mathbb{K}^{(\mathbb{N}_0)}$ bestehend aus abbrechenden Folgen, also solchen Folgen $(x_n)_{n\in\mathbb{N}_0}$, die nach endlich vielen Termen konstant Null werden (siehe Beispiel 4.1.8(g)). Dieser Raum ist unendlich dimensional mit der Basis $\{e_m \mid m \in \mathbb{N}_0\}$ (siehe Beispiele 4.2.5(e) und 4.2.11(d)). Hier ist $e_m = \left(\delta_{m,k}\right)_{k\in\mathbb{N}_0}$ die Folge $(0,0,\ldots,0,1,0,\ldots)$, die an der Stelle m eine 1 stehen hat und sonst nur aus Nullen besteht.

Diesen $\mathbb{K}$-Vektorraum wollen wir nun zu einer $\mathbb{K}$-Algebra machen und müssen dafür noch eine Multiplikation definieren. Wir definieren dazu die *Faltung* von zwei solchen Folgen als

$$(a_k)_{k\in\mathbb{N}_0} * (b_l)_{l\in\mathbb{N}_0} := \left(\sum_{k+l=n} a_k b_l\right)_{n\in\mathbb{N}_0}$$

Die Summe geht über alle Paare (k,l) mit der Eigenschaft, dass $k+l=n$ ist. Da es davon – für festes $n \in \mathbb{N}$ – nur endlich viele gibt, ist dies eine endliche Summe. Dies zeigt, dass die Faltung von zwei Folgen wieder eine Folge ergibt. Nun müssen wir noch zeigen, dass die Faltung von zwei abbrechenden Folgen wieder eine abbrechende Folge ergibt.

Aus $(a_k)_{k\in\mathbb{N}_0}, (b_l)_{l\in\mathbb{N}_0} \in \mathbb{K}^{(\mathbb{N}_0)}$ folgt:

$$\exists k_0 \in \mathbb{N}_0 : \forall k > k_0 : a_k = 0 \quad \text{und} \quad \exists l_0 \in \mathbb{N}_0 : \forall l > l_0 : a_l = 0$$

Wir müssen nun zeigen, dass für $n > k_0 + l_0$, jeder Summand von

$$\sum_{k+l=n} a_k b_l$$

gleich 0 ist. Sei also $(k,l) \in \mathbb{N}_0 \times \mathbb{N}_0$ mit $k + l = n$. Dann gibt es zwei Fälle: Wenn $k > k_0$, dann ist $a_k = 0$ und somit auch das Produkt $a_k b_l$. Wenn aber $k \leq k_0$ ist, dann gilt:

$$l = (k+l) - k = \underbrace{n}_{>k_0+l_0} - \underbrace{k}_{\leq k_0} > (k_0 + l_0) - k_0 = l_0.$$

Also ist in diesem Falle $b_l = 0$ und somit auch das Produkt $a_k b_l$.

Dies zeigt, dass die Faltung von zwei Folgen in $\mathbb{K}^{(\mathbb{N}_0)}$ wieder in $\mathbb{K}^{(\mathbb{N}_0)}$. Wir erhalten also eine Abbildung

$$* : \mathbb{K}^{(\mathbb{N}_0)} \times \mathbb{K}^{(\mathbb{N}_0)} \to \mathbb{K}^{(\mathbb{N}_0)}.$$

Es bleibt zu zeigen, dass $(\mathbb{K}^{(\mathbb{N}_0)}, +, *)$ ein Ring ist und dass $*$ und $\cdot$ verträglich sind wie in Definition 5.4.2 gefordert.

Aus der Definition der Faltung sieht man sofort, dass die Faltung bilinear ist (siehe Definition 5.2.1). Somit gilt die Verträglichkeit mit der skalaren Multiplikation und das Distributivgesetz.

Aus

$$\begin{aligned}
\left((a_k)_{k\in\mathbb{N}_0} * (b_l)_{l\in\mathbb{N}_0}\right) * \left(c_j\right)_{j\in\mathbb{N}_0} &= \left(\sum_{k+l=n} a_k b_l\right)_{n\in\mathbb{N}_0} * \left(c_j\right)_{j\in\mathbb{N}_0} \\
&= \left(\sum_{n+j=p} \left(\sum_{k+l=n} a_k b_l\right) c_j\right)_{p\in\mathbb{N}_0} \\
&= \left(\sum_{k+l+j=p} a_k b_l c_j\right)_{p\in\mathbb{N}_0} \\
&= \left(\sum_{k+m=p} a_k \left(\sum_{l+j=m} b_l c_j\right)\right)_{p\in\mathbb{N}_0} \\
&= (a_k)_{k\in\mathbb{N}_0} * \left(\sum_{l+j=m} b_l c_j\right)_{m\in\mathbb{N}_0} \\
&= (a_k)_{k\in\mathbb{N}_0} * \left((b_l)_{l\in\mathbb{N}_0} * \left(c_j\right)_{j\in\mathbb{N}_0}\right)
\end{aligned}$$

folgt die Assoziativität der Faltung. Die Folge $e_0 = (=)_{n\in\mathbb{N}} \delta_{n,0} = (1,0,0,\ldots)$ ist ein Neutralelement

bezüglich der Faltung:

$$\begin{aligned}
e_0 * (b_l)_{l\in\mathbb{N}_0} &\left(\delta_{k,0}\right)_{k\in\mathbb{N}_0} * (b_l)_{l\in\mathbb{N}_0} \\
&= \left(\sum_{k+l=n} \delta_{k,0} b_l\right)_{n\in\mathbb{N}_0} \\
&= (b_n)_{n\in\mathbb{N}_0} \\
&= \left(\sum_{l+k=n} b_l \delta_{k,0} b_l\right)_{n\in\mathbb{N}_0} \\
&= (b_l)_{l\in\mathbb{N}_0} * \left(\delta_{k,0}\right)_{k\in\mathbb{N}_0} \\
&= (b_l)_{l\in\mathbb{N}_0} * e_0.
\end{aligned}$$

Also ist $(\mathbb{K}^{(\mathbb{N}_0)}, +, *)$ ein Ring und da $*$ und $\cdot$ verträglich sind, ist $(\mathbb{K}^{(\mathbb{N}_0)}, +, *, \cdot)$ eine $\mathbb{K}$-Algebra.

Es sei nun $X := e_1 = \left(\delta_{n,1}\right)_{n\in\mathbb{N}_0} = (0,1,0,0,\ldots)$. Man rechnet nach – zum Beispiel per vollständiger Induktion –, dass

$$\forall n \in \mathbb{N}_0 : X^n = X * \cdots * X = e_n = (0,\ldots,0,1,0,\ldots)$$

gilt. Somit gilt $\{X^n \mid n \in \mathbb{N}_0\} = \{e_n \mid n \in \mathbb{N}_0\}$ und dies ist eine Basis für $\mathbb{K}^{(\mathbb{N}_0)}$.

Damit haben wir gezeigt, dass $(\mathbb{K}^{(\mathbb{N}_0)}, +, *, \cdot)$ eine Polynomalgebra über dem Körper $\mathbb{K}$ ist:

$$(\mathbb{K}^{(\mathbb{N}_0)}, +, *, \cdot) = \mathbb{K}[X].$$

(b)
Wir zeigen nun die Eindeutigkeit bis auf eindeutigen Isomorphismus. Gegeben seien also zwei Polynomalgebren $(\mathbf{A}, +, \bullet, \cdot) = \mathbb{K}[X]$ und $(\mathbf{B}, +, \bullet, \cdot) = \mathbb{K}[Y]$. Die Menge $\{X^0, X^1, X^2, X^3, \ldots\} \subseteq \mathbf{A}$ ist eine Basis für $\mathbf{A}$ als $\mathbb{K}$-Vektorraum. Nach Satz 4.2.14 gibt es eine eindeutige lineare Abbildung $\varphi : \mathbf{A} \to \mathbf{B}$, die X^j auf Y^j abbildet. Da die Menge $\{Y^0, Y^1, Y^2, \ldots\}$ eine Basis von $\mathbf{B}$ ist, folgt, dass diese Abbildung auch bijektiv ist. Es bleibt zu zeigen, dass diese lineare Abbildung auch ein Ringhomomorphismus ist. Offenbar gilt $\varphi(1_{\mathbf{A}}) = \varphi(X^0) = Y^0 = 1_{\mathbf{B}}$, also wird das Einselement auf das Einselement geschickt. Warum erhält die Abbildung Produkte? Es seien $p, q \in \mathbf{A}$ beliebige Elemente. Aus $\mathbf{A} = \mathrm{LH}_{\mathbb{K}}\left(\left\{X^k \;\middle|\; k \in \mathbb{N}_0\right\}\right)$ folgt, dass sich p und q als Linearkombination von Potenzen von X schreiben lassen:

$$p = \sum_{j=0}^{m} a_k X^k \quad \text{und} \quad q = \sum_{l=0}^{n} b_l X^l.$$

Es gilt nun:

$$\begin{aligned}
\varphi(p \bullet q) &= \varphi\left(\left(\sum_{j=0}^{m} a_k X^k\right) \bullet \left(\sum_{l=0}^{n} b_l X^l\right)\right) \\
&= \varphi\left(\sum_{j=0}^{m}\sum_{l=0}^{n} a_k b_l X^k \bullet X^l\right) \\
&= \varphi\left(\sum_{j=0}^{m}\sum_{l=0}^{n} a_k b_l X^{k+l}\right) \\
&= \sum_{j=0}^{m}\sum_{l=0}^{n} a_k b_l \varphi\left(X^{k+l}\right) \\
&= \sum_{j=0}^{m}\sum_{l=0}^{n} a_k b_l Y^{k+l} \\
&= \sum_{j=0}^{m}\sum_{l=0}^{n} a_k b_l Y^k \bullet Y^l \\
&= \left(\left(\sum_{j=0}^{m} a_k Y^k\right) \bullet \left(\sum_{l=0}^{n} b_l Y^l\right)\right) \\
&= \left(\left(\sum_{j=0}^{m} a_k \varphi(X^k)\right) \bullet \left(\sum_{l=0}^{n} b_l \varphi(X^l)\right)\right) \\
&= \varphi\left(\sum_{j=0}^{m} a_k X^k\right) \bullet \varphi\left(\sum_{l=0}^{n} b_l X^l\right) \\
&= \varphi(p) \bullet \varphi(q).
\end{aligned}$$

Somit ist die Abbildung $\varphi : \mathbf{A} \to \mathbf{B}$ ein $\mathbb{K}$-Algebra-Homomorphismus und die Aussage ist bewiesen. □

Bemerkung 5.4.9. 1. Unsere Definition der Polynomalgebra (Definition 5.4.6) ähnelt der Definition des Körpers der komplexen Zahlen (Satz 3.5.4) insofern, als wir nicht explizit sagen, was ein Polynom/eine komplexe Zahl *ist*, sondern nur wie man mit Polynomen/komplexen Zahlen rechnen kann, d.h. welche Regeln die Menge aller Polynome erfüllen sollen. In einem Extraschritt muss man dann zeigen, dass ein solches Objekt (der Körper der Komplexen Zahlen bzw. die Polynomalgebra) wirklich (mathematisch) existiert. In so einem Existenzbeweis wird dann für gewöhnlich eine explizite Konstruktion angegeben (bei den komplexen Zahlen haben wir sogar zwei verschiedene angegeben, die aber isomorphe Körper ergeben).

   Auch bei den reellen Zahlen ist dies nicht anders. In der Analysis-/HM1-Vorlesung wurde auch nicht definiert, was eine reelle Zahl *ist*, sondern nur, dass die Menge aller reellen Zahlen ein Körper mit einer verträglichen vollständigen Ordnung ist.[13]

2. Der Algebrahomomorphismus von $\mathbb{K}$ nach $\mathbb{K}[X]$, der Körperelemente auf konstante Polynome abbildet, ist injektiv. Somit können und werden wir $\mathbb{K}$ als Unteralgebra von $\mathbb{K}[X]$ auffassen.

[13] Details lesen Sie bitte dort nach.

Für uns gilt ab jetzt also[14]: $\mathbb{K} \subseteq \mathbb{K}[X]$.

Somit ist ein Körperelement immer auch ein Polynom.

3. Neben der Konstruktion im Beweis von Satz 5.4.8, die die übliche ist, gibt es auch noch andere Möglichkeiten, eine Polynomalgebra $\mathbb{K}[X]$ über einem Körper $\mathbb{K}$ zu konstruieren. Beispielsweise kann man den Vektorraum $V := \mathbb{K}_0^{\mathbb{N}}$ aller Folgen betrachten und den *Rechts-Shift*-Operator
$$R : V \to V, \quad (x_0, x_1, x_2, \ldots) \mapsto (0, x_0, x_1, \ldots)$$
betrachten, der einfach alle Folgenglieder um eins nach rechts schiebt und vorne mit einer 0 auffüllt[15]. Da $R : V \to V$ linear ist, gilt $R \in \mathrm{End}_{\mathbb{K}}(V)$. Nun ist aber $(\mathrm{End}_{\mathbb{K}}(V), +, \circ, \cdot)$ eine $\mathbb{K}$-Algebra. Nun rechnet man nach, dass die Potenzen $R^0, R^1, R^2, \ldots$ in $(\mathrm{End}_{\mathbb{K}}(V), +, \circ, \cdot)$ paarweise verschieden und linear unabhängig sind. Daraus folgt, dass die Unteralgebra
$$\mathbb{K}[R] := \mathrm{LH}_{\mathbb{K}}\left(R^0, R^1, R^2, \ldots\right) \subseteq \mathrm{End}_K(V)$$
eine Polynomalgebra über $\mathbb{K}$ ist.

Lemma 5.4.10 (Der Grad eines Polynoms)**.**
Es sei $\mathbb{K}$ ein Körper und $p, q \in \mathbb{K}[X]$ Elemente in der Polynomalgebra.

(a) Es gilt: $\deg(p+q) \leq \max\{\deg(p), \deg(q)\}$.

(b) Falls $p, q \neq 0$, dann gilt $\deg(pq) = \deg(p) + \deg(q)$.

Falls man $-\infty + k := -\infty$ setzt, gilt (b) auch ohne die Bedingung $p, q \neq 0$.

Beweis. (a)
Es sei $k := \max\{\deg(p), \deg(q)\}$.
Falls $k = -\infty$, dann sind $p = 0$ und $q = 0$ und es gibt nichts zu zeigen.
Nehmen wir deshalb an, $k \in \mathbb{N}_0$. Dann können wir p und q schreiben als
$$p = \sum_{i=0}^{k} a_i X^i \quad \text{und} \quad q = \sum_{i=0}^{k} b_i X^i.$$
Die Summe $p + q$ ergibt dann:
$$p + q = \sum_{i=0}^{k} a_i X^i + \sum_{i=0}^{k} b_i X^i = \sum_{i=0}^{k} (a_i + b_i) X^i.$$
Hieraus folgt: $\deg(p+q) \leq k$.
(b)
Falls mindestens eins der Polynome p, q das Nullpolynom ist, steht auf beiden Seiten der Gleichung $-\infty$ und die Aussage ist bewiesen. Nehmen wir deshalb also an, dass beide Polynome ungleich 0 sind.
Wir setzen $m := \deg(p) \in \mathbb{N}_0$ und $n := \deg(q) \in \mathbb{N}_0$. Dann können wir schreiben:
$$p = \sum_{i=0}^{m} a_i X^i \quad \text{und} \quad q = \sum_{j=0}^{n} b_j X^j \quad \text{mit } a_m, b_n \neq 0.$$

[14] So, wie wir auch $\mathbb{R}$ als Teilmenge von $\mathbb{C}$ auffassen.
[15] Jetzt könnte ein guter Zeitpunkt sein, im Internet nach „Hilberts Hotel“ zu suchen…

Das Produkt pq ergibt dann:

$$pq = \left(\sum_{i=0}^{m} a_i X^i\right)\left(\sum_{j=0}^{n} b_j X^j\right) = \sum_{i=0}^{m}\sum_{j=0}^{n} a_i b_j X^{i+j}$$

Der höchste Exponent von X, der hier in dieser Summe auftaucht, ist $a_m b_n \neq 0$. Also gilt: $\deg(pq) = m + n$. □

Auch wenn ein Polynom im Sinne dieser Vorlesung keine Funktion von $\mathbb{K}$ nach $\mathbb{K}$ mehr ist, so ist es dennoch möglich in ein Polynom einen Wert „einzusetzen“. Es sei $p \in \mathbb{Q}[X]$ ein Polynom mit rationalen Koeffizienten. Dann ist es ohne weiteres möglich, ein Element a aus $\mathbb{R}$ oder $\mathbb{C}$ in p einzusetzen. Dies ist möglich, weil $\mathbb{R}$ und $\mathbb{C}$ Algebren über $\mathbb{Q}$ sind. Allgemein gilt:

Satz 5.4.11 (Auswertungshomomorphismus)**.** *Es sei $a \in \mathbf{A}$ ein Element in einer Algebra über einem Körper $\mathbb{K}$, wie üblich sei $\mathbb{K}[X]$ die Polynomalgebra über $\mathbb{K}$. Dann gibt es einen eindeutigen Homomorphismus von $\mathbb{K}$-Algebren*

$$\mathbb{K}[X] \to \mathbf{A}, \quad p \mapsto p(a),$$

der die Variable *X auf das Element a abbildet, genannt der* Auswertungshomomorphismus *an der Stelle a Wie die Notation $p(a)$ schon suggeriert, sagen wir auch, dass wir a in p einsetzen.*

Beweis. Der Beweis verläuft fast genauso wie der Beweis von Satz 5.4.8(b):

Weil $\mathbb{K}[X]$ eine Polynomalgebra über der *Variable* X ist, gilt: Die Menge $\{X^0, X^1, X^2, X^3, \ldots\} \subseteq \mathbb{K}[X]$ ist eine Basis für $\mathbb{K}[X]$ als $\mathbb{K}$-Vektorraum. Nach Satz 4.2.14 gibt es eine eindeutige lineare Abbildung $\varphi : \mathbb{K}[X] \to \mathbf{A}$, die X^j auf a^j abbildet. Es bleibt zu zeigen, dass diese lineare Abbildung auch ein Ringhomomorphismus ist. Offenbar gilt $\varphi(1_{\mathbb{K}[X]}) = \varphi(X^0) = a^0 = 1_A$, also wird das Einselement auf das Einselement geschickt. Warum erhält die Abbildung Produkte? Es seien $p, q \in \mathbf{A}$ beliebige Elemente. Aus $\mathbb{K}[X] = \mathrm{LH}_{\mathbb{K}}\left(\{X^k \mid k \in \mathbb{N}_0\}\right)$ folgt, dass sich p und q als Linearkombination von Potenzen von X schreiben lassen:

$$p = \sum_{j=0}^{m} a_k X^k \quad \text{und} \quad q = \sum_{l=0}^{n} b_l X^l.$$

Es gilt nun:

$$\begin{aligned}
\varphi(p \bullet q) &= \varphi\left(\left(\sum_{j=0}^{m} a_k X^k\right)\left(\sum_{l=0}^{n} b_l X^l\right)\right) \\
&= \varphi\left(\sum_{j=0}^{m}\sum_{l=0}^{n} a_k b_l X^k\ X^l\right) \\
&= \varphi\left(\sum_{j=0}^{m}\sum_{l=0}^{n} a_k b_l X^{k+l}\right) \\
&= \sum_{j=0}^{m}\sum_{l=0}^{n} a_k b_l \varphi\left(X^{k+l}\right) \\
&= \sum_{j=0}^{m}\sum_{l=0}^{n} a_k b_l a^{k+l} \\
&= \sum_{j=0}^{m}\sum_{l=0}^{n} a_k b_l a^k \bullet a^l \\
&= \left(\left(\sum_{j=0}^{m} a_k a^k\right) \bullet \left(\sum_{l=0}^{n} b_l a^l\right)\right) \\
&= \left(\left(\sum_{j=0}^{m} a_k \varphi(X^k)\right) \bullet \left(\sum_{l=0}^{n} b_l \varphi(X^l)\right)\right) \\
&= \varphi\left(\sum_{j=0}^{m} a_k X^k\right) \bullet \varphi\left(\sum_{l=0}^{n} b_l X^l\right) \\
&= \varphi(p) \bullet \varphi(q).
\end{aligned}$$

Somit ist die Abbildung $\varphi : \mathbb{K}[X] \to \mathbf{A}$ ein $\mathbb{K}$-Algebra-Homomorphismus und die Aussage ist bewiesen. □

Beispiel 5.4.12. Sei beispielsweise $\mathbb{K} = \mathbb{F}_2 = \{0,1\}$ und $p = X^3 + X^2 \in \mathbb{F}_2[X]$. Dann ist $p(0) = 0^3 + 0^2 = 0$ und $p(1) = 1^3 + 1^2 = 0$, d.h. p bildet jedes Element aus $\mathbb{F}_2$ auf 0 ab. Trotzdem ist p nicht das Nullpolynom, da nicht alle Koeffizienten 0 sind.

Betrachten wir nun die Algebra $\mathbf{A} := \mathbb{F}_2^{3\times 3}$ und das Element

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

Dann gilt

$$p\left(\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}\right) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}^3 + \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}^2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Wir sehen also: Obwohl das Polynom $p \in \mathbb{F}_2[X]$ jedes Element aus dem Grundkörper $\mathbb{F}_2$ auf 0 abbildet, bildet es nicht jede Matrix mit Einträgen aus $\mathbb{F}_2$ auf die Nullmatrix ab.

Lemma 5.4.13 (Nullstellen von Polynomen)**.**
Es sei $\mathbb{K}$ ein Körper, $p \in \mathbb{K}[X]$ und $\lambda \in \mathbb{K}$. Dann sind äquivalent:

(i) $p(\lambda) = 0$.

(ii) $\exists q \in \mathbb{K}[X] : p = (X - \lambda)q$.

Beweis. „(ii) $\Longrightarrow$ (i)“: Da das Auswerten an der Stelle λ insbesondere ein Ring-Homomorphismus ist, gilt

$$p(\lambda) = \big((X-\lambda)q\big)(\lambda) = (X(\lambda) - \lambda)q(\lambda) = (\lambda - \lambda)q(\lambda) = 0.$$

„(i) $\Longrightarrow$ (ii)“: Wir schreiben $p = \sum_{k=0}^{n} a_k X^k$.

Wir setzen $Y := X - \lambda \in \mathbb{K}[X]$. Dann gilt:$X = Y + \lambda$ und wir können p umschreiben als:

$$p = \sum_{k=0}^{n} a_k (Y+\lambda)^k = \sum_{j=0}^{n} b_j Y^j,$$

wobei wir einfach solange ausmultipliziert haben, bis nur noch Potenzen von Y übriggeblieben sind.

Nun wenden wir den Auswertungshomomorphismus an und nutzen aus, dass $Y(\lambda) = (X - \lambda)(\lambda) = 0$ ist:

$$0 = p(\lambda) = \left(\sum_{j=0}^{n} b_j Y^j\right)(\lambda) = \sum_{j=0}^{n} b_j 0^j = b_0.$$

Also ist $b_0 = 0$ und somit gilt:

$$p = b_1 Y + b_2 Y^2 + \cdots b_n Y^n = Y \cdot \big(b_1 + b_2 Y + \cdots b_n Y^{n-1}\big) = (X-\lambda)\underbrace{(b_1 + b_2(X-\lambda) + \cdots b_n (X-\lambda)^{n-1})}_{=q}.$$

□

Satz 5.4.14.
Es sei $\mathbb{K}$ *ein Körper und* $p \in \mathbb{K}[X]$ *ein Polynom mit* $\deg(p) = n \in \mathbb{N}_0$
Dann hat p *in* $\mathbb{K}$ *höchstens* n *verschiedene Nullstellen.*

Beweis. Wir beweisen dies per vollständiger Induktion.

Induktionsanfang $n = 0$**:**
Ein Polynom vom Grad 0 ist einfach nur eine Konstante: $p = \mu \in \mathbb{K}$. Diese Konstante ist nicht 0, denn sonst wäre der Grad $-\infty$. Also hat p gar keine Nullstelle in $\mathbb{K}$.

Induktionsschritt:
Es sei $n \in \mathbb{N}_0$ so gewählt, dass jedes Polynom vom Grad n höchstens n Nullstellen in $\mathbb{K}$ hat.

Wir nehmen nun an, $p \in \mathbb{K}[X]$ habe Grad $n+1$.

Falls p gar keine Nullstelle hat, dann ist die Aussage bewiesen, denn $0 \leq n+1$. Nehmen wir also an, p habe mindestens eine Nullstelle $\lambda \in \mathbb{K}$. Dann gilt nach Lemma 5.4.13, dass es ein $q \in \mathbb{K}[X]$ gibt mit

$$p = (X - \lambda)q.$$

Wenden wir nun Lemma 5.4.10 auf diese Gleichung an, dann erhalten wir:

$$\underbrace{\deg(p)}_{=n+1} = \underbrace{\deg(X-\lambda)}_{=1} + \deg(q)$$

und somit wissen wir, dass $\deg(q) = n$ ist.

Nach Induktionsvoraussetzung hat q somit höchstens n verschiedene Nullstelen in $\mathbb{K}$.

Es sei nun $\mu \in \mathbb{K}$ eine Nullstelle von p mit $\mu \neq \lambda$. Dann gilt:

$$0 = p(\mu) = \big((X-\lambda)q\big)(\mu) = \underbrace{(\mu-\lambda)}_{\neq 0} q(\mu).$$

Weil $\mathbb{K}$ ein Körper ist und somit nullteilerfrei, folgt: $q(\mu) = 0$.

Wir haben also gesehen, dass jede Nullstelle von p, die nicht λ ist auch eine Nullstelle von q ist und da q nach Voraussetzung höchstens n verschiedene Nullstellen haben kann, sehen wir dass p höchstens $n+1$ verschiedene Nullstellen haben kann. Damit ist die Aussage bewiesen. □

Korollar 5.4.15.
Es sei $\mathbb{K}$ ein Körper. Wir bezeichnen mit $\mathbf{P}_{\mathbb{K}}$ die Menge aller Polynomfunktionen von $\mathbb{K}$ nach $\mathbb{K}$. Dies ist eine Unteralgebra der Algebra $\mathbb{K}^{\mathbb{K}}$.

(a) Die Abbildung

$$\Phi : \mathbb{K}[X] \to \mathbb{K}^{\mathbb{K}}, \quad p \mapsto \big(\mathbb{K} \to \mathbb{K}, \quad t \mapsto p(t)\big),$$

die jedes (abstrakte) Polynom $p \in \mathbb{K}[X]$ auf die Polynomfunktion $\big(\mathbb{K} \to \mathbb{K}, \quad t \mapsto p(t)\big)$ abbildet, ist ein $\mathbb{K}$-Algebra-Homomorphismus. Es gilt: $\mathrm{Bild}(\Phi) = \mathbf{P}_{\mathbb{K}}$.

(b) Die Abbildung ist genau dann injektiv, wenn $\mathbb{K}$ unendlich viele Elemente hat. Dann ist also $\mathbb{K}[X]$ isomorph zu $\mathbf{P}_{\mathbb{K}}$ und insbesondere sind die Funktionen der Form

$$\big(\mathbb{K} \to \mathbb{K}, \quad t \mapsto t^n\big), n \in \mathbb{N}_0$$

paarweise verschieden und linear unabhängig.

(c) Die Abbildung Φ ist genau dann surjektiv, wenn $\mathbb{K}$ endlich viele Elemente hat. Dann gilt also $\mathbf{P}_{\mathbb{K}} = \mathbb{K}^{\mathbb{K}}$ und jede Funktion $f : \mathbb{K} \to \mathbb{K}$ ist eine Polynomfunktion.

Beweis. (a)
Die Menge $\mathbf{A} := \mathbb{K}^{\mathbb{K}}$ ist eine $\mathbb{K}$-Algebra nach Beispiel 5.4.4(v).

Wir betrachten die Funktion

$$f_0 : \mathbb{K} \to \mathbb{K}, \quad t \mapsto t.$$

Dann ist f ein Element in der Algebra $\mathbf{A}$.

Wir wenden nun Satz 5.4.11 an und erhalten einen eindeutigen $\mathbb{K}$-Algebra-Homomorphismus

$$\Phi : \mathbb{K}[X] \to \mathbf{A},$$

der die formale Variable X auf f_0 abbildet.

Für ein (formales) Polynom $p = \sum_{k=0}^{n} a_k X^k \in \mathbb{K}[X]$ ist also $\Phi(p) \in \mathbb{K}^{\mathbb{K}}$ eine Funktion von $\mathbb{K}$

nach $\mathbb{K}$:

$$\begin{aligned}\big(\Phi(p)\big)(t) &= \big(\Phi(\sum_{k=0}^{n} a_k X^k)\big)(t)\\ &= \big(\sum_{k=0}^{n} a_k \Phi(X)^k)\big)(t)\\ &= \big(\sum_{k=0}^{n} a_k f_0^k)\big)(t)\\ &= \sum_{k=0}^{n} a_k (f_0(t))^k)\\ &= \sum_{k=0}^{n} a_k t^k.\end{aligned}$$

Also ist $\Phi(p)$ die zu p gehörige Polynomfunktion.

(b)

Wir wollen zeigen:

$$\big(\Phi : \mathbb{K}[X] \to \mathbb{K}^{\mathbb{K}} \text{ ist injektiv}\big) \quad \Longleftrightarrow \quad (\mathbb{K} \text{ ist unendlich}).$$

„$\Longrightarrow$“: Angenommen, $\mathbb{K}$ sei nicht unendlich. Dann hieße dies $\mathbb{K}$ habe nur $|\mathbb{K}| = m \in \mathbb{N}$ viele Elemente:

$$\mathbb{K} = \{\lambda_1, \ldots, \lambda_m\}.$$

Betrachte das Polynom $p := (X-\lambda_1)\cdots(X-\lambda_m) \in \mathbb{K}[X]$. Dieses Polynom hat Grad m und ist somit insbesondere nicht das Nullpolynom, obwohl die Polynomfunktion $\Phi(p)$ jedes Körperelement auf die 0 abbildet.

Es gilt also $\Phi(p) = 0$, was der Injektivität von Φ widerspricht.

„$\Longleftarrow$“: Wir nehmen an $\mathbb{K}$ habe unendlich viele Elemente. Wir wollen zeigen: Φ ist injektiv. Es sei dazu $p \in \ker(\Phi)$ ein (formales) Polynom, das auf die konstante Nullfunktion abgebildet wird. Wir wollen zeigen, dass $p = 0$ ist. Angenommen, p sei nicht das Nullpolynom, dann würde gelten $n := \deg(p) \in \mathbb{N}_0$. Nach Satz 5.4.14 hätte p dann nur n viele Nullstellen. Nach Voraussetzung ist aber $\Phi(p)$ die Nullfunktion, d.h. jedes Körperelement ist eine Nullstelle. Da wir angenommen hatten, $\mathbb{K}$ habe unendlich viele Elemente, hätte also p unendlich viele Nullstellen. Das ist ein Widerspruch. Also muss $p = 0$ sein.

(c)

Wir wollen zeigen:

$$\big(\Phi : \mathbb{K}[X] \to \mathbb{K}^{\mathbb{K}} \text{ ist surjektiv}\big) \quad \Longleftrightarrow \quad (\mathbb{K} \text{ ist endlich}).$$

„$\Longrightarrow$“: Angenommen, Φ ist surjektiv. Wir wollen zeigen, dass $\mathbb{K}$ nur endlich viele Elemente hat. Nehmen wir per Widerspruch an, dass $\mathbb{K}$ unendlich viele Elemente hat. Die Funktion

$$h : \mathbb{K} \to \mathbb{K}, \quad t \mapsto \begin{cases} 1 & \text{für } t = 0 \\ 0 & \text{sonst.} \end{cases}$$

Diese Funktion hat unendlich viele Nullstellen, ist aber nicht die Nullfunktion. Nach Satz 5.4.14 ist so etwas für eine Polynomfunktion aber nicht möglich. Also gibt es Funktionen, die keine Polynomfunktionen sind und damit kann $\Phi : \mathbb{K}[X] \to \mathbb{K}^{\mathbb{K}}$ nicht surjektiv sein. Widerspruch.

„$\Longleftarrow$“:
Es sei $\mathbb{K}$ ein endlicher Körper:

$$\mathbb{K} = \{\lambda_1, \ldots, \lambda_m\}.$$

Für jedes $k \in \{1, \ldots, m\}$ betrachten wir das Polynom $p_k \in \mathbb{K}[X]$ vom Grad $m-1$, das wie folgt definiert ist:

$$p_k := \prod_{j \in \{1,\ldots,m\}; j \neq k} (X - \lambda_j).$$

Dann gilt $\Phi(p_k)$ ist eine Funktion, die alle Elemente aus $\mathbb{K}$ auf 0 abbildet, außer λ_k, was auf $p_k(\lambda_k) \neq 0$ abgebildet wird. Somit ist $\{\Phi(p_1), \ldots, \Phi(p_m)\}$ eine Basis für den Vektorraum $\mathbb{K}^{\mathbb{K}}$ und Φ ist surjektiv. □

Bemerkung 5.4.16. Alles, was wir in diesem Kapitel gemacht haben, lässt sich auch durchführen, wenn wir den Grundkörper $\mathbb{K}$ durch einen kommutativen Ring R ersetzen.

Zuerst stellt man fest, dass man in der Definition des Vektorraumbegriffs (siehe Definition 4.1.1) den Körper $\mathbb{K}$ durch einen kommutativen Ring R ersetzen kann. Das Konzept, was man dann erhält, nennt man einen *R-Modul*[16].

Begriffe wie lineare Unabhängigkeit, Erzeugendensystem und Basis übertragen sich direkt auf diesen allgemeineren Rahmen der R-Moduln über einem kommutativen Ring R. Einige Sätze (wie Satz 4.2.8) übertragen sich direkt auf R-Moduln, während andere (wie Satz 4.2.7) für R-Moduln einfach falsch sind.

Analog kann man Definition 5.4.2 verallgemeinern und dann R-Algebren definieren. Und schließlich ist es möglich, dann Polynomalgebren $R[X]$ über Ringen R zu definieren.

Insbesondere erhält man dann $\mathbb{Z}[X]$, den Ring der Polynome mit ganzzahligen Koeffizienten.

Ebenso kann man natürlich einen Polynomring $R[X]$ als Koeffizientenring für einen neuen Polynomring betrachten und erhält damit beispielsweise $R[X][Y] = R[X,Y]$ den Polynomring in zwei Veränderlichen. Dieser Vorgang lässt sich wiederholen und man kann damit Polynomringe in beliebig vielen Variablen $R[X_1, X_2, \ldots, X_n]$ definieren und untersuchen. Diese Polynome in mehreren Veränderlichen und ihre Nullstellenmengen sind Gegenstand der Untersuchung und Ausgangspunkt der *algebraischen Geometrie* eine große Rolle.

Für diese Veranstaltung reichen uns aber Polynome in einer Variablen mit Koeffizienten aus einem Körper $\mathbb{K}$.

[16] „Der Modul“ wird auf der ersten Silbe betont und der Plural ist „Moduln“. Es ist nicht zu verwechseln mit dem Wort „das Modul“ (wie in „Modulbeschreibung“ oder „Modulhandbuch“), das auf der zweiten Silbe betont wird und als Plural „Module“ hat.

Zusammenfassung von Abschnitt 5.4

(1) Eine $\mathbb{K}$-Algebra ist eine algebraische Struktur, die sowohl eine Ringstruktur als auch eine $\mathbb{K}$-Vektorraumstruktur beinhaltet.

(2) Ein Polynom ist eine formale Linearkombination der Potenzen von X, wobei X eine formale Variable ist. Die Menge aller Polynome $\mathbb{K}[X]$ mit Koeffizienten in $\mathbb{K}$ ist eine unendlich dimensionale $\mathbb{K}$-Algebra. Die Menge $\{X^k \mid k \in \mathbb{N}_0\}$ ist eine Basis für $\mathbb{K}[X]$ als $\mathbb{K}$-Vektorraum.

(3) Man kann jedes Elemente a einer $\mathbb{K}$-Algebra $\mathbf{A}$ in ein Polynom $p \in \mathbb{K}[X]$ einsetzen und erhält dann ein Element $p(a) \in \mathbb{K}[X]$. Insbesondere kann man Matrizen oder Endomorphismen in Polynome einsetzen.

(4) Wenn ein Polynom eine Nullstelle hat, so kann man diese Nullstelle als Linearfaktor ausklammern.

(5) Ein Polynom ist etwas anderes als eine Polynomfunktion. Die Unterscheidung ist insbesondere für endliche Körper sehr wichtig.

5.5. Endomorphismen und Ähnlichkeit

Es sei V ein Vektorraum über einem Körper $\mathbb{K}$ und $\varphi : V \to V$ ein Endomorphismus von V. Dann kann man für ein $v_0 \in V$ den Vektor $\varphi(v_0) \in V$ und dann $\varphi(\varphi(v_0)) \in V$ und so weiter berechnen.

Man erhält somit eine rekursiv definierte Folge: $v_{k+1} = \varphi(v_k)$. Eine naheliegende Frage ist nun: Was passiert für große Werte von k? Das kann sowohl von dem Endomorphismus φ als auch von dem Anfangswert v_0 abhängen.

Die erste einfache Erkenntnis ist: Wenn $v_0 = 0$ der Nullvektor ist, dann sind alle weiteren Folgenglieder auch 0. Es reicht also, sich auf Anfangsvektoren zu konzentrieren, die nicht der Nullvektor sind.

Beispiel 5.5.1. Es sei $V = \mathbb{R}^3$ und $\varphi : V \to V$ der Endomorphismus, der bezüglich der Standardbasis durch die Matrix $A \in \mathbb{R}^{3\times 3}$ gegeben ist.

(i) Wir nehmen an:

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

Dann gilt für jeden Anfangswert $v_0 = \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix}$, dass

$$\varphi(v_0) = \begin{pmatrix} \beta \\ \gamma \\ 0 \end{pmatrix} \quad \text{und} \quad \varphi(\varphi(v_0)) = \begin{pmatrix} \gamma \\ 0 \\ 0 \end{pmatrix}.$$

Wenn wir nun noch einmal φ anwenden, erhalten wir nur noch den Nullvektor.

Wir erhalten also – unabhängig vom Anfangswert – eine Folge, die nach endlich vielen Schritten konstant 0 wird und dann auch 0 bleibt.

(ii) Wir nehmen an:

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

Dann gilt für den Anfangswert $v_0 = e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$:

$$v_1 = \varphi(v_0) = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 3 \\ 0 \\ 0 \end{pmatrix} = 3 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

Wir sehen also: Der Vektor v_0 wird von φ auf das 3-fache von sich selbst abgebildet. Dies bedeutet, dass

$$\varphi(\varphi(v_0)) = \varphi(3v_0) = 3\varphi(v_0) = 3 \cdot 3v_0 = 9v_0.$$

Allgemein sieht man dann:

$$v_k = 3^k v_0 = \begin{pmatrix} 3^k \\ 0 \\ 0 \end{pmatrix}.$$

Wir sehen also: Der erste Eintrag in diesem Vektor (und damit auch seine Länge[17]) wächst also exponentiell an.

Wenn wir dieselbe Matrix A betrachten, aber einen Anfangsvektor $w_0 = \begin{pmatrix} 0 \\ \beta \\ \gamma \end{pmatrix}$, der in der ersten Komponente 0 ist, dann erhalten wir

$$w_1 = \varphi(w_0) = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 0 \\ \beta \\ \gamma \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \frac{1}{2}\gamma \end{pmatrix}.$$

Wiederholtes Anwenden von φ liefert die Folge

$$w_k = \begin{pmatrix} 0 \\ 0 \\ \left(\frac{1}{2}\right)^k \gamma \end{pmatrix},$$

die unabhängig von β und γ exponentiell gegen den Nullvektor konvergiert[18].

Für einen allgemeinen Vektor $v_0 = \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} \in \mathbb{R}^3$ und $k \in \mathbb{N}$ gilt:

$$\underbrace{\varphi \circ \cdots \circ \varphi}_{k \text{ mal}} \left(\begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} \right) = \begin{pmatrix} 3^k \alpha \\ 0 \\ \left(\frac{1}{2}\right)^k \gamma \end{pmatrix}.$$

[17] Wir haben zwar noch nicht formal definiert, was die Länge eines Vektors in $\mathbb{R}^3$ ist, aber es sollte klar sein, was gemeint ist

[18] Um dies zu formalisieren, müsste man erst einmal Abstände in $\mathbb{R}^3$ definieren, aber es sollte klar sein, dass der Letzte Eintrag exponentiell gegen 0 konvergiert.

(iii) Wir nehmen an:

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

Wenn wir einen Vektor $v_0 = \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} \in \mathbb{R}^3$ auswählen, so erhalten wir die Folgenglieder:

$$\begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} \longmapsto \begin{pmatrix} -\beta \\ \alpha \\ \gamma \end{pmatrix} \longmapsto \begin{pmatrix} -\alpha \\ -\beta \\ \gamma \end{pmatrix} \longmapsto \begin{pmatrix} \beta \\ -\alpha \\ \gamma \end{pmatrix} \longmapsto \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix}.$$

Wir haben also ein periodisches Verhalten. Nachdem wir φ viermal angewendet haben, sind wir wieder im Ursprungszustand des Systems.

(iv) Wir nehmen an:

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

Als Anfangswert nehmen wir den Vektor $e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \in \mathbb{R}^3$. Dann ergibt sich die Folge:

$$\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \longmapsto \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \longmapsto \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix} \longmapsto \begin{pmatrix} 0 \\ 3 \\ 1 \end{pmatrix} \longmapsto \cdots$$

Wir sehen also: Die Einträge wachsen an, aber nicht exponentiell, sondern nur linear.

Bemerkung 5.5.2. Wie man in Beispiel 5.5.1 gesehen hat, kann das Verhalten eines solchen Systems sehr unterschiedlich sein (nach endlich vielen Schritten konstant, exponentiell wachsend, exponentiell fallend, periodisch, langsamer als exponentiell wachsend). In diesem konkreten Fällen war es sehr leicht, durch Ausprobieren nach endlich vielen Schritten zu erkennen, was passiert. Das liegt daran, dass die Matrizen eine relativ schöne Gestalt hatten. Im Allgemeinen ist dies deutlich komplizierter. Deshalb wird es unser Ziel sein bei gegebenem Endomorphismus eine möglichst geeignete Basis des Raums zu finden, sodass die Matrix besonders schön ist.

Beispiel 5.5.3. Wir betrachten den Endomorphismus φ des reellen Vektorraums $\mathbb{R}^3$:

$$\varphi : \mathbb{R}^3 \to \mathbb{R}^3, \quad \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mapsto \begin{pmatrix} \frac{81x_1}{2} - 75x_2 + 180x_3 \\ 60x_1 - 112x_2 + 270x_3 \\ \frac{33x_1}{2} - 31x_2 + 75x_3 \end{pmatrix}.$$

Hier ist es überhaupt nicht ersichtlich, was passiert, wenn man diesen Endomorphismus mehrfach auf einen gegebenen Anfangsvektor anwendet. Es stellt sich die Frage, ob es möglich ist, das k-fache Anwenden in eine geschlossene Form zu bringen.

Da φ ein Endomorphismus eines endlich dimensionalen Vektorraums ist, können wir die Darstellungsmatrix von φ bestimmen. Dazu benötigen wir aber eine Basis von $\mathbb{R}^3$.

Wenn wir die Standardbasis $\mathsf{E} = (e_1, e_2, e_3)$ nehmen, erhalten wir die Matrix

$$A = \mathrm{M}_{\mathsf{E},\mathsf{E}}(\varphi) = \begin{pmatrix} \frac{81}{2} & -75 & 180 \\ 60 & -112 & 270 \\ \frac{33}{2} & -31 & 75 \end{pmatrix} \in \mathbb{R}^{3\times 3}.$$

Alle Informationen über den Endomorphismus φ sind nun in dieser Matrix kodiert. Leider kann man sagen, dass man hier immer noch nicht viel sehen kann. Man kann feststellen, dass φ kein Isomorphismus ist, indem man die Determinante von A ausrechnet, aber mehr wird erst einmal schwierig.

Hier hilft ein Basiswechsel. Wenn wir statt der Standardbasis diese Basis hier verwenden:

$$\mathsf{B} := \left(\begin{pmatrix} 12 \\ 18 \\ 5 \end{pmatrix}, \begin{pmatrix} 10 \\ 15 \\ 4 \end{pmatrix}, \begin{pmatrix} 3 \\ 4 \\ 1 \end{pmatrix} \right)$$

Dann erhalten wir die folgenden Basiswechselmatrix von der neuen Basis B in die alte Standardbasis E:

$$T := \mathrm{M}_{\mathsf{E},\mathsf{B}}(\mathrm{id}_{\mathbb{R}^3}) = \begin{pmatrix} 12 & 10 & 3 \\ 18 & 15 & 4 \\ 5 & 4 & 1 \end{pmatrix} \in \mathrm{GL}(3,\mathbb{R}).$$

Durch Invertieren dieser Matrix (Gauß!) erhalten wir die Basiswechselmatrix in die andere Richtung, also von der Standardbasis E in die neue Basis B:

$$T^{-1} := \mathrm{M}_{\mathsf{B},\mathsf{E}}(\mathrm{id}_{\mathbb{R}^3}) = \begin{pmatrix} 12 & 10 & 3 \\ 18 & 15 & 4 \\ 5 & 4 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & -2 & 5 \\ -2 & 3 & -6 \\ 3 & -2 & 0 \end{pmatrix} \in \mathrm{GL}(3,\mathbb{R}).$$

Dann können wir nun die Matrix von φ bezüglich der neuen Basis B darstellen:

$$\begin{aligned} \mathrm{M}_{\mathsf{B},\mathsf{B}}(\varphi) &= \underbrace{\mathrm{M}_{\mathsf{B},\mathsf{E}}(\mathrm{id}_{\mathbb{R}^3})}_{=T^{-1}} \underbrace{\mathrm{M}_{\mathsf{E},\mathsf{E}}(\varphi)}_{=A} \underbrace{\mathrm{M}_{\mathsf{E},\mathsf{B}}(\mathrm{id}_{\mathbb{R}^3})}_{=T} \\ &= T^{-1}AT \\ &= \begin{pmatrix} 1 & -2 & 5 \\ -2 & 3 & -6 \\ 3 & -2 & 0 \end{pmatrix} \begin{pmatrix} \frac{81}{2} & -75 & 180 \\ 60 & -112 & 270 \\ \frac{33}{2} & -31 & 75 \end{pmatrix} \begin{pmatrix} 12 & 10 & 3 \\ 18 & 15 & 4 \\ 5 & 4 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 3 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} \end{pmatrix}. \end{aligned}$$

Dies ist genau die Matrix aus Beispiel 5.5.1 (ii), die wir schon analysiert haben. Wir wissen also, wenn der Startvektor bezüglich der Basis B in der ersten Komponente 0 ist, dann konvergiert die iterierte Folge exponentiell gegen 0, ansonsten divergiert die iterierte Folge.

Wir sehen also: Um die Abbildung φ zu untersuchen, ist die Standardbasis E ungeeignet und die Basis B ist viel sinnvoller.

Es bleibt aber die Frage: Wie man eine solche Basis findet, unter der die Matrix besonders einfach wird, wenn sie nicht – wie hier – schon vorgegeben ist.

Motiviert durch Beispiel 5.5.3 definieren wir:

Definition 5.5.4 (Ähnlichkeit von Matrizen)**.**
Es sei $\mathbb{K}$ ein Körper und $n \in \mathbb{N}$. Dann heißen Matrizen A und B *ähnlich*, wenn eine invertierbare Matrix $S \in \mathrm{GL}(n,\mathbb{K})$ existiert, sodass

$$SAS^{-1} = B.$$

Wie wir bereits in Bemerkung 4.3.18 gesehen haben, bedeutet A ist ähnlich zu B genau, dass A und B durch einen Basiswechsel auseinander hervorgehen, **wenn wir im Definitions- und Zielbereich dieselbe geordnete Basis verwenden**. Ähnlichkeit ist stärker als die sogenannte Äquivalenz von Matrizen (siehe Bemerkung 4.3.18) für Details.

In Beispiel 5.5.3 haben wir gesehen:

$$\begin{pmatrix} \frac{81}{2} & -75 & 180 \\ 60 & -112 & 270 \\ \frac{33}{2} & -31 & 75 \end{pmatrix} \text{ ist ähnlich zu } \begin{pmatrix} 3 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} \end{pmatrix}.$$

Wie kann man überprüfen, ob zwei quadratische Matrizen derselben Größe ähnlich sind?

Wir führen nun eine notwendige Bedingung ein, damit zwei Matrizen ähnlich sind:

Definition 5.5.5 (Spur)**.** Es sei $\mathbb{K}$ ein Körper, $n \in \mathbb{N}$ und $A = (a_{i,j})_{i,j\in\{1,\dots,n\}} \in \mathbb{K}^{n\times n}$ eine quadratische Matrix. Dann definieren wir die Spur

$$\mathrm{tr}(A) := \sum_{j=1}^{n} a_{j,j} = a_{1,1} + \cdots + a_{n,n} \in \mathbb{K}$$

als die Summe der Diagonaleinträge.

Lemma 5.5.6. *Es sei $\mathbb{K}$ ein Körper.*

(a) Für jedes $n \in \mathbb{N}$ ist die Abbildung $\mathrm{tr} : \mathbb{K}^{n\times n} \to \mathbb{K}$ ist linear.

(b) Wenn A,B Matrizen sind, sodass $\mathrm{tr}(AB)$ definiert ist, dann ist auch $\mathrm{tr}(BA)$ definiert und es gilt: $\mathrm{tr}(AB) = \mathrm{tr}(BA)$.

Man beachte, dass im Allgemeinen *nicht* gilt: $\mathrm{tr}(AB) = \mathrm{tr}(A)\mathrm{tr}(B)$.

Proposition 5.5.7 (Ähnlichkeitsinvarianten)**.**
Es sei $\mathbb{K}$ ein Körper, $n \in \mathbb{N}$ und $A,B \in \mathbb{K}^{n\times n}$.
Wenn A ähnlich zu B ist, dann gilt

(a) $\mathrm{rg}(A) = \mathrm{rg}(B)$.

(b) $\det(A) = \det(B)$.

(c) $\mathrm{tr}(A) = \mathrm{tr}(B)$.

Wenn also eine dieser Bedingungen nicht erfüllt ist, sind die Matrizen nicht ähnlich.

Achtung: Falls zwei Matrizen den gleichen Rang, die gleiche Determinante und die gleiche Spur haben, folgt daraus noch nicht*, dass sie ähnlich sind!*

Beweis. Weil A ähnlich ist zu B, gibt es eine invertierbare Matrix $S \in \mathrm{GL}(n,\mathbb{K})$ mit

$$A = SBS^{-1}.$$

Der Rang einer Matrix A ist die Dimension der Bildes der zu A gehörigen linearen Abbildung. Ein Basiswechsel im Definitionsbereich ändert das Bild nicht und somit auch nicht den Rang. Ein Basiswechsel im Zielbereich ändert zwar das Bild, nicht aber dessen Dimension. Somit bleibt der Rang gleich bei beliebigen Basiswechseln im Definitions- und im Zielbereich (auch unabhängig voneinander, siehe auch den Korollar 4.3.19 zu Smith-Normalform).

Es gilt also insbesondere:

$$\mathrm{rg}(A) = \mathrm{rg}(SBS^{-1}) = \mathrm{rg}(SB) = \mathrm{rg}(B).$$

Nun zur Determinante: Es gilt:

$$\det(A) = \det(SBS^{-1}) = \det(S)\det(B)\det(S^{-1}) = \frac{\det(S)\det(B)}{\det(S)} = \det(B).$$

Hier haben wir die Multiplikativität der Determinantenfunktion (Satz 5.3.9) verwendet – sowie die Tatsache, dass $\det(S^{-1}) = \frac{1}{\det(S)}$ gilt (Satz 5.3.10).

Die Spur ist nicht multiplikativ. Deshalb müssen wir anders argumentieren:

$$\mathrm{tr}\big(A\big) = \mathrm{tr}\big(SBS^{-1}\big) = \mathrm{tr}\big((SB)\cdot S^{-1}\big) \overset{\text{Lemma } 5.5.6}{=} \mathrm{tr}\big(S^{-1}\cdot(SB)\big) = \mathrm{tr}\big(S^{-1}SB\big) = \mathrm{tr}\big(\mathbb{1}_n B\big) = \mathrm{tr}\big(B\big). \quad \square$$

Beispiel 5.5.8. Es seien die Matrizen $A_1, A_2, A_3, A_4 \in \mathbb{R}^{3\times 3}$ gegeben:

$$A_1 := \begin{pmatrix} 0 & 3 & 7 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}; A_2 := \begin{pmatrix} 1 & -1 & 0 \\ 1 & -1 & 0 \\ -2 & 2 & 0 \end{pmatrix}; A_3 := \begin{pmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \\ 2 & 4 & -3 \end{pmatrix}; A_4 := \begin{pmatrix} -6 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}; A_5 := \begin{pmatrix} 0 & 0 & -10 \\ 1 & 0 & -3 \\ 0 & 1 & 6 \end{pmatrix}$$

Keine zwei dieser fünf Matrizen sind zu einander ähnlich, weil bei je zwei von ihnen mindestens eine der Eigenschaften Rang, Determinante oder Spur unterschiedlich sind.

Definition 5.5.9 (Determinante und Spur eines Endomorphismus)**.**
Es sei V ein endlich dimensionaler Vektorraum über einem Körper $\mathbb{K}$. Es sei $\varphi \in \mathrm{End}_{\mathbb{K}}(V)$ ein Endomorphismus von V.

(i) Die *Determinante* von φ ist definiert als die Determinante von $\mathrm{M}_{\mathsf{B},\mathsf{B}}(\varphi)$ bezüglich irgendeiner geordneten Basis B von V:

$$\det(\varphi) := \det(\mathrm{M}_{\mathsf{B},\mathsf{B}}(\varphi)) \in \mathbb{K}.$$

(ii) Die *Spur* von φ ist definiert als die Spur von $\mathrm{M}_{\mathsf{B},\mathsf{B}}(\varphi)$ bezüglich irgendeiner geordneten Basis B von V:

$$\mathrm{tr}(\varphi) := \mathrm{tr}\left(\mathrm{M}_{\mathsf{B},\mathsf{B}}(\varphi)\right) \in \mathbb{K}.$$

Dies ist wohldefiniert, weil ähnliche Matrizen gleiche Determinante und gleiche Spur haben (Proposition 5.5.7).

Zusammenfassung von Abschnitt 5.5

(1) Die Darstellungsmatrix eines Endomorphismus eines endlich dimensionalen Vektorraums ist immer quadratisch.

(2) Zwei quadratische Matrizen sind ähnlich, wenn sie den gleichen Endomorphismus bezüglich unterschiedlicher Basen beschreiben – wobei wir im Definitionsbereich und im Zielbereich die gleiche Basis verwenden müssen.

(3) Ähnliche Matrizen haben die gleiche Spur, den gleichen Rang und die gleiche Determinante. So kann man auch Spur und Determinante von Endomorphismen von endlich dimensionalen Vektorräumen definieren.

5.6. Eigenwerte und Eigenvektoren

Definition 5.6.1 (Eigenwert und Eigenvektor)**.**
Es sei V ein Vektorraum über einem Körper $\mathbb{K}$ und $\varphi \in \mathrm{End}_{\mathbb{K}}(V)$ ein Endomorphismus.

(a) Ein Vektor $v \in V$ mit $v \neq 0$ heißt *Eigenvektor* von φ, wenn v auf ein Vielfaches von sich selbst abgebildet wird, d.h. wenn gilt:

$$\exists \lambda \in \mathbb{K} : \varphi(v) = \lambda v.$$

(b) Der Skalar λ, der zu einem Eigenvektor gehört heißt *Eigenwert*. Es gilt also: $\lambda \in \mathbb{K}$ ist ein Eigenwert zu φ, wenn gilt:

$$\exists v \in V \setminus \{0\} : \varphi(v) = \lambda v.$$

Im Falle $V = \mathbb{K}^n$ für ein $n \in \mathbb{N}$, so nennen wir die Eigenwerte von $\varphi_A : \mathbb{K}^n \to \mathbb{K}^n, \quad x \mapsto Ax$ auch die *Eigenwerte* der Matrix A. Ebenso sind die *Eigenvektoren* von A die Eigenvektoren der Abbildung $\varphi_A : \mathbb{K}^n \to \mathbb{K}^n, \quad x \mapsto Ax$.

Lemma 5.6.2 (Eigenwerte sind Ähnlichkeitsinvarianten)**.**
Es seien $\mathbb{K}$ ein Körper, $n \in \mathbb{N}$, $A, B \in \mathbb{K}^{n \times n}$.
Wenn A ähnlich ist zu B, dann haben A und B dieselben Eigenwerte:

$$\{\lambda \in \mathbb{K} \mid \lambda \text{ ist Eigenwert von } A\} = \{\lambda \in \mathbb{K} \mid \lambda \text{ ist Eigenwert von } B\}.$$

Proposition 5.6.3.
Es sei V ein Vektorraum über einem Körper $\mathbb{K}$ und $\mathsf{B} = (v_1, \ldots, v_n)$ eine geordnete Basis von V.
Für einen Endomorphismus $\varphi \in \mathrm{End}_{\mathbb{K}}(V)$ sind äquivalent:

(i) Die Darstellungsmatrix von φ bezüglich B ist in Diagonalform:

$$\mathrm{M}_{\mathsf{B},\mathsf{B}}(\varphi) = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}.$$

(ii) Die Basis $\mathsf{B} = (v_1, \ldots, v_n)$ *besteht aus Eigenvektoren von* φ*:*

$$\forall j \in \{1, \ldots, n\} : \varphi(v_j) = \lambda_j v_j.$$

Die Diagonaleinträge $\lambda_1, \ldots, \lambda_n$ *sind genau die Eigenwerte zu den Eigenvektoren* $v_1, \ldots, v_n$ *– in genau dieser Reihenfolge.*

Ein Endomorphismus $\varphi \in \mathrm{End}_{\mathbb{K}}(V)$ *eines endlich dimensionalen* $\mathbb{K}$*-Vektorraums* V *heißt* diagonalisierbar*, wenn es eine geordnete Basis* B *gibt, sodass die beiden äquivalenten Bedingungen erfüllt sind.*

Bemerkung 5.6.4. Gegeben ein Endomorphismus $\varphi : V \to V$ eines Vektorraums V über einem Körper $\mathbb{K}$. Es stellt sich die Frage: Wie finde ich Eigenwerte und Eigenvektoren von φ? Diese Frage kann man in drei Fragen unterteilen, die von oben nach unten schwieriger zu beantworten sind:

Frage I: Gegeben ein Eigenvektor $v \in V$**, wie viele Eigenwerte gehören dazu und wie finde ich diese?**

Frage II: Gegeben ein Eigenwert $\lambda \in \mathbb{K}$**, wie viele Eigenvektoren gehören dazu und wie finde ich diese?**

Frage III: Wie finde ich alle Eigenwerte – ohne die Eigenvektoren zu kennen. Wie viele Eigenwerte kann φ **haben?**

Diese Fragen wollen wir nun beantworten.

Lemma 5.6.5. *Gegeben ein Endomorphismus* $\varphi : V \to V$ *eines Vektorraums* V *über einem Körper* $\mathbb{K}$*. Zu jedem Eigenvektor* $v \in V$ *gibt es genau einen Eigenwert, d.h. der Skalar* $\lambda \in \mathbb{K}$ *in der Gleichung*

$$\varphi(v) = \lambda v$$

ist eindeutig bestimmt.

Beweis. Nach Definition 5.6.1(a) gibt es mindestens ein solches $\lambda \in \mathbb{K}$. Angenommen, es gäbe zwei unterschiedliche $\lambda_1, \lambda_2 \in \mathbb{K}$, dann gilt:

$$v = \frac{1}{\lambda_1 - \lambda_2}(\lambda_1 - \lambda_2)v = \frac{1}{\lambda_1 - \lambda_2}(\lambda_1 v - \lambda_2 v) = \frac{1}{\lambda_1 - \lambda_2}(\varphi(v) - \varphi(v)) = 0.$$

Da aber Eigenvektoren *per Definition* nicht 0 sind, folgt, dass $\lambda_1 = \lambda_2$ gelten muss. □

Damit ist geklärt, dass es – zu gegebenem Eigenvektor – einen eindeutigen Eigenwert gibt. Diesen berechnet direkt durch Einsetzen in die Gleichung

$$\varphi(v) = \lambda(v).$$

Damit ist Frage I aus Bemerkung 5.6.4 beantwortet.

Kommen wir nun zu Frage II: Wenn ein Eigenwert gegeben ist, wie viele Eigenvektoren gehören dazu und wie berechnen man sie?

Die Antwort ist gegeben durch folgende Proposition:

Proposition 5.6.6 (Eigenräume)**.**
Gegeben ein Endomorphismus $\varphi : V \to V$ *eines Vektorraums* V *über einem Körper* $\mathbb{K}$*.*

Gegeben sei ein Eigenwert $\lambda \in \mathbb{K}$*. Dann gilt:*

$$\{v \in V \mid v \text{ ist ein Eigenvektor zu } \lambda\} \cup \{0\} = \left\{v \in V \;\middle|\; \varphi(v) = \lambda v\right\} = \ker(\lambda \mathrm{id}_V - \varphi) = \ker(\varphi - \lambda \mathrm{id}) \subseteq V.$$

Diese Menge ist ein Untervektorraum von V und wir der Eigenraum *von φ zum Eigenwert λ genannt:*

$$E_\lambda(\varphi) := \ker(\varphi - \lambda\mathrm{id}) \subseteq V.$$

Die Dimension dieses Untervektorraums $\dim_{\mathbb{K}} E_\lambda(\varphi)$ *heißt* geometrische Vielfachheit *von λ.*

Beweis. Für jedes $\lambda \in \mathbb{K}$ und jeden Vektor $v \in V$ gelten die Äquivalenzen:

$$\begin{aligned}
&v \text{ ist ein Eigenvektor zu } \lambda \text{ oder } v = 0\\
\iff &\varphi(v) = \lambda v\\
\iff &\lambda v - \varphi(v) = 0\\
\iff &\lambda\mathrm{id}_V(v) - \varphi(v) = 0\\
\iff &\big(\lambda\mathrm{id}_V - \varphi\big)(v) = 0\\
\iff &v \in \ker\big(\lambda\mathrm{id}_V - \varphi\big)
\end{aligned}$$

Der Kern einer linearen Abbildung ist immer ein Untervektorraum. □

Damit haben wir die Antwort auf Frage II: Die Menge der Eigenvektoren zu gegebenem Eigenwert λ bilden – wenn wir den Nullvektor (der kein Eigenvektor ist!) hinzufügen – einen Untervektorraum, den Eigenraum $E_\lambda(\varphi)$ von V. Falls V endlich dimensional ist, so können wir die Menge – nach Wahl einer Basis – durch ein homogenes lineares Gleichungssystem lösen.

Bleibt nun noch Frage III: Wie finden wir die Eigenwerte, wenn wir noch keine Eigenvektoren kennen? Diese Frage werden wir nur beantworten, wenn V endlich dimensional ist. Zunächst holen wir etwas aus:

Definition 5.6.7 (Das charakteristische Polynom einer Matrix)**.**
Es sei $A \in \mathbb{K}^{n\times n}$ eine Matrix mit $n \in \mathbb{N}$ und $\mathbb{K}$ ein Körper.

Dann können wir die Matrix

$$\begin{aligned}
X\mathbb{1}_n - A &= \begin{pmatrix} X & & & \\ & X & & \\ & & \ddots & \\ & & & X \end{pmatrix} - \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & & a_{2,n} \\ \vdots & & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{pmatrix}\\
&= \begin{pmatrix} X - a_{1,1} & -a_{1,2} & \cdots & -a_{1,n} \\ -a_{2,1} & X - a_{2,2} & & -a_{2,n} \\ \vdots & & \ddots & \vdots \\ -a_{n,1} & -a_{n,2} & \cdots & X - a_{n,n} \end{pmatrix} \in (\mathbb{K}[X])^{n\times n}.
\end{aligned}$$

bilden. Das *charakteristische Polynom* der Matrix A ist definiert als die Determinante dieser Matrix:

$$p_A := p_A(X) := \det(X\mathbb{1}_n - A) \in \mathbb{K}[X].$$

Bemerkung 5.6.8. Wir haben bis jetzt immer nur Matrizen mit Einträgen aus einem Körper betrachtet und deshalb auch nur Determinanten von solchen Matrizen definiert und berechnet.

Es ist allerdings völlig unkritisch, die entsprechenden Definitionen auf kommutative Ringe auszudehnen. Die Definition über die Leibniz-Formel (Definition 5.3.1) lässt sich für beliebige kommutative Ringe übertragen. Ebenso die Laplace-Entwicklung nach Zeilen oder Spalten ist

möglich. Auch die „Gauß-Schritte“ (G1), (G2) und (G3) sind in diesem Rahmen gültig – man darf halt grundsätzlich nur durch Skalare teilen, die invertierbar sind.

In diesem allgemeineren Sinne ist die Determinante in Definition 5.6.7 zu verstehen.

Alternativ kann man auch den Polynomring $\mathbb{K}[X]$, der selbst kein Körper ist, in einen Körper einbetten – ebenso wie man den Ring $\mathbb{Z}$ in den Körper $\mathbb{Q}$ einbetten kann. Wenn man dies tut (was wir uns hier sparen), dann kann man die Determinante 5.6.7 auch wieder als eine Determinante mit Einträgen aus einem Körper verstehen – nur halt einem größeren als dem Körper $\mathbb{K}$, mit dem wir angefangen haben.

Proposition 5.6.9.
Es seien $\mathbb{K}$ ein Körper, $n \in \mathbb{N}$, $A, B \in \mathbb{K}^{n\times n}$.

Wenn A ähnlich ist zu B, dann haben A und B das gleiche charakteristische Polynom:

$$p_A = p_B \in \mathbb{K}[X].$$

Beweis. Es sei $A = SBS^{-1}$.

Dann gilt:

$$\begin{aligned}
p_A &= \det\big(X\mathbb{1}_n - A\big) \\
&= \det\big(X\mathbb{1}_n - SBS^{-1}\big) \\
&= \det\big(XSS^{-1} - SBS^{-1}\big) \\
&= \det\big(S(XS^{-1} - BS^{-1})\big) \\
&= \det\big(S(X\mathbb{1}_n - B)S^{-1}\big) \\
&= \det(S)\det\big(X\mathbb{1}_n - B\big)\underbrace{\det(S^{-1})}_{=1/\det(S)} \\
&= \det\big(X\mathbb{1}_n - B\big) \\
&= p_B.
\end{aligned}$$

□

Definition 5.6.10.
Es sei V ein Vektorraum mit *endlicher* Dimension $n \in \mathbb{N}$ über einem Körper $\mathbb{K}$.

Das *charakteristische Polynom* eines Endomorphismus $\varphi \in \mathrm{End}_{\mathbb{K}}(V)$ ist definiert als

$$p_\varphi := p_A \in \mathbb{K}[X], \quad \text{wobei } A = \mathrm{M}_{\mathsf{B},\mathsf{B}}(\varphi) \in \mathbb{K}^{n\times n}$$

die Darstellungsmatrix von φ bezüglich einer beliebigen geordneten Basis B von V ist. Dies ist wohldefiniert nach Proposition 5.6.9.

Was hat nun das charakteristische Polynom mit Frage III aus Bemerkung 5.6.4 zu tun?

Proposition 5.6.11 (Die Nullstellen des charakteristischen Polynoms)**.**
Gegeben ein Endomorphismus $\varphi : V \to V$ eines Vektorraums V der Dimension $n \in \mathbb{N}$ über einem Körper $\mathbb{K}$. Dann sind die Eigenwerte von φ genau die Nullstellen des charakteristischen Polynoms von φ:

$$\big\{\lambda \in \mathbb{K} \;\big|\; \lambda \textit{ ist ein Eigenwert von } \varphi\big\} = \big\{\lambda \in \mathbb{K} \;\big|\; p_\varphi(\lambda) = 0\big\}.$$

Beweis. Es sei $\mathcal{B}$ eine geordnete Basis von V. Für $\lambda \in \mathbb{K}$ sind äquivalent:

$$\begin{aligned}
\lambda \text{ ist ein Eigenwert} &\iff \exists v \neq 0 : \varphi(v) = \lambda v \\
&\iff \exists v \neq 0 : v \in \ker(\lambda \mathrm{id}_V - \varphi) \\
&\iff \ker(\lambda \mathrm{id}_V - \varphi) \neq \{0\} \\
&\iff \lambda \mathrm{id}_V - \varphi \text{ ist nicht injektiv} \\
&\iff \lambda \mathbb{1}_n - A \text{ ist nicht invertierbar} \\
&\iff \det(\lambda \mathbb{1}_n - A) = 0 \\
&\iff p_A(\lambda) = 0.
\end{aligned}$$

□

Nun wissen wir, wie wir Eigenwerte ausrechnen können: Wir müssen nur die Nullstellen des charakteristischen Polynoms bestimmen. Beachten Sie aber, dass das Finden von Nullstellen von Polynomen ein *nichtlineares* Problem ist. und dass es für $\deg(p_A) \geq 5$ keine geschlossene Lösungsformel zum Finden der Nullstellen gibt[19].

Lemma 5.6.12.
Es sei $A \in \mathbb{K}^{n\times n}$ eine Matrix mit $n \in \mathbb{N}$ und $\mathbb{K}$ ein Körper. Das charakteristische Polynom p_A hat Grad n und die höchste Potenz X^n hat Koeffizient 1.

Beweis. Wir beweisen die Aussage per Induktion nach $n \in \mathbb{N}$.

Induktionsanfang $n = 1$:
In diesem Fall gilt: $A = a_{1,1} \in \mathbb{K}^{1\times 1} = \mathbb{K}$ und

$$p_A = \det(X\mathbb{1}_1 - A) = \det(X - a_{1,1}) = X - a_{1,1} \in \mathbb{K}[X].$$

Induktionsschritt:
Es sei $n \in \mathbb{N}$ so gewählt, dass die zu zeigende Aussage für alle $(n \times n)$-Matrizen gilt. Sei nun $A \in \mathbb{K}^{n+1\times n+1}$.

Wir berechnen die Determinante $p_A = \det(X\mathbb{1}_{n+1} - A)$ durch Laplace-Entwicklung nach der ersten Zeile:

$$\begin{aligned}
p_A &= \det\begin{pmatrix} X - a_{1,1} & -a_{1,2} & \cdots & -a_{1,n+1} \\ -a_{2,1} & X - a_{2,2} & & -a_{2,n+1} \\ \vdots & & \ddots & \vdots \\ -a_{n+1,1} & -a_{n+1,2} & \cdots & X - a_{n+1,n+1} \end{pmatrix} \\
&= (X - a_{1,1}) \det\left(\mathrm{St}_{(1,1)}(X\mathbb{1}_{n+1} - A)\right) + \sum_{l=2}^{n+1} (-1)^{1+l}(-a_{1,l}) \det\left(\mathrm{St}_{(1,l)}(X\mathbb{1}_{n+1} - A)\right) \\
&= (X - a_{1,1}) p_B + \sum_{l=2}^{n+1} (-1)^{1+l}(-a_{1,l}) \underbrace{\det\left(\mathrm{St}_{(1,l)}(X\mathbb{1}_{n+1} - A)\right)}_{q_l \in \mathbb{K}[X]},
\end{aligned}$$

wobei $B := \mathrm{St}_{(1,1)}(A)$ die Matrix ist, die man durch Streichen der ersten Zeilen und Spalte erhält. Nach Induktionsvoraussetzung gilt somit $\deg p_B = n$ und der Leitkoeffizient von p_B ist 1. Somit hat das Polynom $(X - a_{1,1})p_B$ Grad $n + 1$ und ebenfalls Leitkoeffizient 1.

[19] Für Grad 3 und 4 existieren zwar Formeln, aber die sind praktisch nicht handhabbar und deshalb für uns nutzlos. Beachten Sie außerdem, dass die Lösungsformel für Grad 2 (abc-Formel, pq-Formel, bzw. *Mitternachtsformel* auch nicht über beliebigen Körpern funktioniert, da darin Quadratwurzeln vorkommen und diese nicht in jedem Körper definiert sind)

Jedes Polynom q_l hat Grad höchstens n, somit ändern diese Summanden den Grad und den Leitkoeffizienten nicht. □

Hieraus folgt sofort:

Proposition 5.6.13.
Es V ein Vektorraum der Dimension $n \in \mathbb{N}$ über einem Körper $\mathbb{K}$. Dann hat jeder Endomorphismus $\varphi \in \mathrm{End}_{\mathbb{K}}(V)$ höchstens *n verschiedene Eigenwerte in $\mathbb{K}$.*

Beweis. Die Eigenwerte von φ sind die Nullstellen von p_φ nach Proposition 5.6.11.

Das Polynom hat Grad n nach Lemma 5.6.12.

Ein Polynom von Grad n hat höchstens n verschiedene Nullstellen nach Satz 5.4.14. □

Wir werden nun ein paar hilfreiche Aussagen über Eigenvektoren und Eigenwerte beweisen, die uns helfen werden, zu entscheiden, ob eine gegebene Matrix diagonalisierbar ist oder nicht:

Satz 5.6.14. *Es sei V ein Vektorraum über einem Körper $\mathbb{K}$ und $\varphi \in \mathrm{End}_{\mathbb{K}}(V)$ ein Endomorphismus. Angenommen $\lambda_1, \dots, \lambda_k$ sind paarweise verschiedene Eigenwerte von φ. Dann ist die Summe $W := E_{\lambda_1}(\varphi) + \cdots + E_{\lambda_k}(\varphi)$ direkt, d.h. die Abbildung*

$$E_{\lambda_1}(\varphi) \times \cdots \times E_{\lambda_k}(\varphi) \to W, \quad (v_1, \dots, v_k) \mapsto v_1 + \cdots + v_k$$

ist bijektiv (siehe Definition 4.4.1).

Achtung: Dies heißt nicht, dass $W = V$ gilt.

Beweis. Wir zeigen die Aussage per vollständiger Induktion über $k \in \mathbb{N}$.

Induktionsanfang $k = 1$:
In diesem Fall ist $W = E_{\lambda_1}(\varphi)$ und die angegebene Abbildung ist einfach die Identität.

Induktionsschritt:
Es sei $k \in \mathbb{N}$ so gewählt, dass die Summe von k Eigenräumen zu verschiedenen Eigenwerten immer direkt ist. Wir werden zeigen, dass dies dann auch für $k+1$ Eigenräume gilt.

Die Abbildung

$$\Phi : E_{\lambda_1}(\varphi) \times \cdots \times E_{\lambda_{k+1}}(\varphi) \to W, \quad (v_1, \dots, v_{k+1}) \mapsto v_1 + \cdots + v_{k+1}$$

ist offensichtlich surjektiv.

Es bleibt, die Injektivität zu zeigen. Es sei also $(v_1, \dots, v_{k+1}) \in \ker(\Phi) \subseteq E_{\lambda_1}(\varphi) \times \cdots \times E_{\lambda_k}(\varphi)$, d.h. es gilt:

$$\sum_{j=1}^{k+1} v_j = 0 \tag{$*$}$$

und wir werden zeigen, dass alle Vektoren $v_j = 0$ sind.

Wir wenden nun den Endomorphismus $\varphi - \lambda_{k+1}\mathrm{id}_V$ auf den Nullvektor an:

$$\begin{aligned}
0 &= (\varphi - \lambda_{k+1}\mathrm{id}_V)(0) \\
&= \varphi(0) - \lambda_{k+1}0 \\
&= \varphi\left(\sum_{j=1}^{k+1} v_j\right) - \lambda_{k+1}\sum_{j=1}^{k+1} v_j \\
&= \sum_{j=1}^{k+1} \varphi\left(v_j\right) - \sum_{j=1}^{k+1} \lambda_{k+1} v_j \\
&= \sum_{j=1}^{k+1} \lambda_j v_j - \sum_{j=1}^{k+1} \lambda_{k+1} v_j \\
&= \sum_{j=1}^{k+1} (\lambda_j - \lambda_{k+1}) v_j \\
&= \sum_{j=1}^{k} (\lambda_j - \lambda_{k+1}) v_j.
\end{aligned}$$

Im Letzten Schritt haben wir verwendet, dass der $(k+1)$-te Summand 0 wird.

Wir haben also nun eine Darstellung des Nullvektors als Summe von k Vektoren, die aus paarweise verschiedenen Eigenräumen stammen. Nach Induktionsvoraussetzung ist aber die Summe von k verschiedenen Eigenräumen direkt, also müssen damit alle diese Vektoren 0 sein:

$$\forall j \in \{1,\dots,k\} : (\lambda_j - \lambda_{k+1}) v_j = 0.$$

Da aber $j < k+1$ und die Eigenwerte als paarweise verschieden angenommen wurden, gilt $\lambda_j - \lambda_{k+1} \neq 0$. Somit gilt: $v_1, \dots v_k = 0$. Da aber die Summe aller $k+1$ Vektoren Null ergibt:

$$\underbrace{v_1 + \dots + v_k}_{=0} + v_{k+1} = 0,$$

muss auch der Letzte Vektor v_{k+1} der Nullvektor sein. □

Korollar 5.6.15.
Es sei V ein Vektorraum über einem Körper $\mathbb{K}$ und $\varphi \in \mathrm{End}_{\mathbb{K}}(V)$ ein Endomorphismus. Wenn $v_1,\dots,v_k$ Eigenvektoren zu paarweise verschiedenen Eigenwerten $\lambda_1,\dots,\lambda_k$ sind, so sind die Vektoren $v_1,\dots,v_k$ paarweise verschieden und linear unabhängig.

Beweis. Dass die Vektoren paarweise verschieden sind, folgt direkt aus Lemma 5.6.5.

Da sie aus paarweise verschiedenen Untervektorräumen stammen, deren Summe direkt ist (Satz 5.6.14), folgt die lineare Unabhängigkeit. □

Hieraus folgt sofort die folgende Aussage:

Proposition 5.6.16 (Hinreichendes Kriterium zur Diagonalisierbarkeit)**.**
Es V ein Vektorraum der Dimension $n \in \mathbb{N}$ über einem Körper $\mathbb{K}$. Falls ein Endomorphismus $\varphi \in \mathrm{End}_{\mathbb{K}}(V)$ genau n verschiedene Eigenwerte in $\mathbb{K}$ besitzt, dann ist φ diagonalisierbar.

Achtung: Wenn φ weniger als n verschiedene Eigenwerte besitzt, so kann es trotzdem sein, dass φ diagonalisierbar ist.

Beweis. Nach Voraussetzung gibt es paarweise verschiedene Eigenwerte $\lambda_1,\dots,\lambda_n$.

Nach Definition 5.6.1 gilt: Zu jedem λ_j kann man einen Eigenvektor v_j auswählen.

Nach Korollar 5.6.15 ist die Menge $B = \{v_1,\dots,v_n\}$ linear unabhängig und hat genau n Elemente. Da der Vektorraum V Dimension n hat, muss B eine Basis für V sein.

Mit Proposition 5.6.3 folgt nun, dass φ diagonalisierbar ist. □

Beispiel 5.6.17. Es seien die folgenden Matrizen $A,B,C \in \mathbb{R}^{3\times 3}$ gegeben:

$$A := \begin{pmatrix} 3 & 0 & 0 \\ 1 & 2 & 0 \\ 42 & 23 & \pi \end{pmatrix}; \quad B := \begin{pmatrix} 3 & 0 & 0 \\ 5 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}; \quad C := \begin{pmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 5 & 2 \end{pmatrix}.$$

Wir möchten nun bestimmen, welche der Matrizen über $\mathbb{R}$ diagonalisierbar sind.

Beginnen wir mit der Matrix A: Das charakteristische Polynom von A lautet:

$$p_A = \det \begin{pmatrix} X-3 & 0 & 0 \\ -1 & X-2 & 0 \\ -42 & -23 & X-\pi \end{pmatrix} = (X-3)(X-2)(X-\pi).$$

Für die Berechnung der Determinanten haben wir hier Proposition 5.3.6 verwendet. Die Eigenwert von A lassen sich nun direkt ablesen: $3,2,\pi$. Dies sind 3 verschiedene Zahlen, unsere Matrix A hat 3 Zeilen und 3 Spalten und somit gilt nach Proposition 5.6.16, dass A diagonalisierbar ist. Es ist gibt also eine Matrix $\widetilde{A}$, die ähnlich zu A ist und Diagonalform hat:

$$\widetilde{A} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & \pi \end{pmatrix}.$$

Kommen wir zur Matrix B: Das charakteristische Polynom von B lautet:

$$p_B = \det \begin{pmatrix} X-3 & 0 & 0 \\ -5 & X-2 & 0 \\ 0 & 0 & X-2 \end{pmatrix} = (X-3)(X-2)^2.$$

Wir sehen also: Es gibt diesmal nur 2 verschiedene Eigenwerte: 2,3. Wir können somit *nicht* Proposition 5.6.16 verwenden, um Diagonalisierbarkeit zu zeigen, weil dies nur eine hinreichende Bedingung gibt.

Betrachten wir stattdessen den Eigenraum $\lambda = 2$:

$$E_2(B) = \ker \begin{pmatrix} 2-3 & 0 & 0 \\ -5 & 2-2 & 0 \\ 0 & 0 & 2-2 \end{pmatrix} = \ker \begin{pmatrix} 1 & 0 & 0 \\ -5 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \mathrm{LH}_{\mathbb{R}} \left(\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right),$$

wobei wir für die Berechnung des Kerns hier einen (sehr kurzen) Gauß-Algorithmus verwendet haben.

Wir sehen also: Der Eigenraum $E_2(B)$ ist 2-dimensional (die geometrische Vielfachheit von $\lambda = 2$ ist also 2.

Da die Eigenräume eine direkte Summe bilden (siehe Satz 5.6.14) gilt:

$$E_2(B) \oplus E_3(B) \subseteq \mathbb{R}^3 \text{ und somit } \dim E_2(B) + \dim E_3(B) \leq 3.$$

Da $\lambda = 3$ ein Eigenwert ist, gilt $\dim(E_3(B)) \geq 1$ und somit ist

$$\dim E_2(B) + \dim E_3(B) = 3$$

und somit ist $\mathbb{R}^3$ die direkte Summe der beiden Eigenräume und φ somit diagonalisierbar.

Kommen wir nun noch zur Matrix C. Das charakteristische Polynom von C lautet:

$$p_C = \det\begin{pmatrix} X-3 & 0 & 0 \\ 0 & X-2 & 0 \\ 0 & -5 & X-2 \end{pmatrix} = (X-3)(X-2)^2.$$

Es ist dasselbe charakteristische Polynom wie bei B und somit hat C die gleichen Eigenwerte wie die Matrix B, nämlich $\lambda = 2$ und $\lambda = 3$. Somit ist es wieder nicht möglich, Proposition 5.6.16 zu verwenden, um Diagonalisierbarkeit zu überprüfen.

Wenn wir die Eigenräume berechnen, erhalten wir

$$E_2(B) = \ker\begin{pmatrix} 2-3 & 0 & 0 \\ 0 & 2-2 & 0 \\ 0 & -5 & 2-2 \end{pmatrix} = \ker\begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & -5 & 0 \end{pmatrix} = \mathrm{LH}_{\mathbb{R}}\left(\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right)$$

und

$$E_3(B) = \ker\begin{pmatrix} 3-3 & 0 & 0 \\ 0 & 3-2 & 0 \\ 0 & -5 & 3-2 \end{pmatrix} = \ker\begin{pmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & -5 & -1 \end{pmatrix} = \mathrm{LH}_{\mathbb{R}}\left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}\right).$$

Wir sehen also: Jeder Eigenvektor von C ist entweder ein Vielfaches von $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ oder von $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$.

Also kann es keine Basis von $\mathbb{R}^3$ aus Eigenvektoren geben. Folglich ist C nicht diagonalisierbar. Wir haben also gezeigt: Es gibt keine Diagonalmatrix, die zu C ähnlich ist.

Da wir weiter oben gesehen hatten, dass B diagonalisierbar ist, haben wir somit zwei Matrizen, B und C gefunden, die *nicht* ähnlich sind, obwohl sie das gleiche charakteristische Polynom, und damit die gleichen Eigenwerte haben (und auch Rang, Determinante und Spur sind identisch).

Satz 5.6.18 (Diagonalisierbarkeit und geometrische Vielfachheit)**.**
Es sei V ein Vektorraum der Dimension $n \in \mathbb{N}$ über einem Körper $\mathbb{K}$ und $\varphi \in \mathrm{End}_{\mathbb{K}}(V)$ ein Endomorphismus. Dann sind äquivalent:

(i) *φ ist diagonalisierbar.*

(ii) *Der Vektorraum V ist die Summe der Eigenräume von φ, d.h. jedes $v \in V$ lässt sich als endliche Summe von Eigenvektoren schreiben.*

(iii) *Der Vektorraum V ist die direkte Summe der Eigenräume:*

$$V = \bigoplus_{\lambda} E_\lambda(\varphi)$$

(iv) *Die Summe der geometrischen Vielfachheiten ist gleich der Dimension von V:*

$$\sum_{\lambda} \dim(E_\lambda(\varphi)) = n.$$

Manchmal hat ein Endomorphismus allerdings überhaupt keine Eigenwerte, wie das folgende Beispiel zeigt:

Beispiel 5.6.19. Es sei $\varphi : \mathbb{R}^2 \to \mathbb{R}^2$ die Rotation um $90° = \pi/2$ gegen den Uhrzeigersinn um den Nullpunkt. Dann lässt sich die Matrix von φ bezüglich der Standardbasis auf einfache Weise[20] aufstellen:

$$A := \mathrm{M}_{\mathsf{E},\mathsf{E}}(\varphi) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Es stellt sich nun die Frage: Ist φ diagonalisierbar? Oder äquivalent gefragt: Ist A diagonalisierbar?

Suchen wir also die Eigenwerte und Eigenvektoren von φ. Per Definition ist ein Eigenvektor $v \in \mathbb{R}^2$ ein Vektor $v \neq 0$, der auf ein Vielfaches von sich selbst abgebildet wird, also entweder in die gleiche oder in die entgegengesetzte Richtung zeigt.

Dies ist bei dieser Abbildung aber unmöglich, weil jeder Vektor v auf einen gedrehten Vektor $\varphi(v)$ abgebildet wird, der natürlich immer linear unabhängig zu v ist. Somit hat φ keine Eigenvektoren, keine Eigenwerte und ist somit insbesondere nicht diagonalisierbar.

Versuchen wir nun das Problem rechnerisch zu lösen:

$$p_\varphi = \det \begin{pmatrix} X & 1 \\ -1 & X \end{pmatrix} = X^2 + 1 \in \mathbb{R}[X].$$

Dieses Polynom hat in den reellen Zahlen keine Lösung, weil für alle $t \in \mathbb{R} : t^2 + 1 > 0$ ist. Dies passt zu der geometrischen Argumentation, dass es keine Eigenwerte und keine Eigenvektoren gibt.

Es gibt aber komplexe Nullstellen von $p_\varphi = X^2 + 1 \in \mathbb{R}[X] \subseteq \mathbb{C}[X]$.

Im Ring $\mathbb{C}[X]$, der den Ring $\mathbb{R}[X]$ als Unterring enthält, gilt:

$$X^2 + 1 = X^2 - \mathrm{i}^2 = (X + \mathrm{i})(X - \mathrm{i}).$$

Es gibt also genau zwei komplexe Nullstellen: i und $-$i.

Wenn wir also die Matrix A nicht als Element in $\mathbb{R}^{2\times 2}$ betrachten, sondern als Element in $\mathbb{C}^{2\times 2}$, dann hat A zwei unterschiedliche Eigenwerte und ist nach Proposition 5.6.16 diagonalisierbar. Als *komplexe* Matrix ist A also diagonalisierbar und ähnlich zu der Matrix

$$\widetilde{A} = \begin{pmatrix} \mathrm{i} & 0 \\ 0 & -\mathrm{i} \end{pmatrix}.$$

Mit dieser Matrix lässt sich nun viel leichter arbeiten – eben, weil sie eine Diagonalmatrix ist. Leider ist der Nachteil, dass diese Matrix natürlich keine $\mathbb{R}$-lineare Abbildung von $\mathbb{R}^2$ nach $\mathbb{R}^2$ mehr beschreibt, sondern eine $\mathbb{C}$-lineare Abbildung von $\mathbb{C}^2$ nach $\mathbb{C}^2$. Diese kann man sich geometrisch natürlich nicht mehr so gut vorstellen, weil $\mathbb{C}^2$ als *reeller* Vektorraum vierdimensional ist und somit einer einfachen geometrischen Anschauung nicht zugänglich.

Dies ist ein Phänomen, das man häufig antrifft: Eine Matrix in $\mathbb{K}^{n\times n}$ ist *über* $\mathbb{K}$ nicht diagonalisierbar, aber über einer entsprechenden Körpererweiterung ist die Matrix diagonalisierbar.

Man kann sogar argumentieren, dass dies der tiefere Grund ist, warum man sich überhaupt mit komplexen Zahlen beschäftigt: Viele Phänomene, die rein reell sind (z.B. reelle Matrizen) lassen sich erst richtig verstehen, wenn man sie in einen größeren komplexen Rahmen einbettet (z.B. in den Raum der komplexen Matrizen).

[20] In den Spalten der Matrix stehen die Bilder der Basisvektoren

Hieraus lässt sich eine wichtige notwendige Bedingung für Diagonalisierbarkeit ableiten – nur brauchen wir dafür noch den folgenden Begriff aus der Theorie der Polynome:

Definition 5.6.20.
Es sei $\mathbb{K}$ ein Körper und $p \in \mathbb{K}[X]$ ein Polynom mit $p \neq 0$. Wir sagen: *p zerfällt in Linearfaktoren*, wenn sich p schreiben lässt als ein Produkt von Polynomen von Grad 1, also:

$$p = a(X-\lambda_1)^{r_1}\cdots(X-\lambda_k)^{r_k} \quad \text{mit } a \in \mathbb{K}^\times; k \in \mathbb{N}_0; \lambda_j \in \mathbb{K}; r_j \in \mathbb{N}.$$

Die Nullstellen $\lambda_1, \ldots, \lambda_k$ kann man als paarweise verschieden ansehen. Für Polynome p mit $\deg(p) = 0$ setzen wir $k = 0$.

Proposition 5.6.21 (Diagonalisierbarkeit und Linearfaktoren)**.**
Es sei V ein Vektorraum der Dimension $n \in \mathbb{N}$ über einem Körper $\mathbb{K}$ und $\varphi \in \mathrm{End}_{\mathbb{K}}(V)$ ein Endomorphismus. Wenn φ diagonalisierbar ist, dann zerfällt das charakteristische Polynom p_φ in $\mathbb{K}[X]$ in Linearfaktoren.

Wenn also ein charakteristisches Polynom über $\mathbb{K}$ nicht in Linearfaktoren zerfällt, so wissen wir sicher, dass der Endomorphismus nicht diagonalisierbar ist. Achtung: Wir haben in Beispiel 5.6.17 gesehen, dass es Matrizen gibt mit einem zerfallenden charakteristischen Polynom, die trotzdem nicht diagonalisierbar sind. In Beispiel 5.6.17 haben wir ja sogar zwei Matrizen (nämlich B und C) gesehen, die das gleiche charakteristische Polynom haben und trotzdem ist eine diagonalisierbar und die andere nicht. Das charakteristische Polynom allein ist also nicht ausreichend, um Diagonalisierbarkeit zu entscheiden.

Lemma 5.6.22.
Für einen Körper $\mathbb{K}$ sind äquivalent:

- $\mathbb{K}$ *ist* algebraisch abgeschlossen[21]*, d.h.*

$$\forall p \in \mathbb{K}[X] : \deg(p) \geq 1 \implies \exists \lambda \in \mathbb{K} : p(\lambda) = 0.$$

- *Jedes Polynom $p \in \mathbb{K}[X]$ mit $p \neq 0$ zerfällt in Linearfaktoren in $\mathbb{K}[X]$.*

Wenn wir also über einem algebraisch abgeschlossenen Körper arbeiten, so ist die notwendige Bedingung in Proposition 5.6.21 automatisch erfüllt. Leider ist der Körper, der in Anwendungen am Wichtigsten ist, nämlich der Körper $\mathbb{R}$, nicht algebraisch abgeschlossen. Auch sind alle endlichen Körper niemals algebraisch abgeschlossen.[22].

Es gibt aber auch eine gute Nachricht: Es gibt eine Körpererweiterung von $\mathbb{R}$, die algebraisch abgeschlossen ist: Der Körper der komplexen Zahlen $\mathbb{C}$. Dies ist genau die Aussage des Fundamentalsatzes der Algebra (Satz 3.5.10). Damit wissen wir also: Das charakteristische Polynom einer Matrix $A \in \mathbb{C}^{n\times n}$ zerfällt immer in Linearfaktoren.

Was fehlt nun noch für die Diagonalisierbarkeit?

Wir haben in Proposition 5.6.6 definiert, was die *geometrische Vielfachheit* eines Eigenwertes ist, nämlich die Dimension des dazugehörigen Eigenraums. Es gibt aber noch eine andere Möglichkeit, wie man die Vielfachheit eines Eigenwertes definieren kann und zwar mit Hilfe des charakteristischen Polynoms.

[21] siehe Definition 3.5.11

[22] Dies lässt sich am Einfachsten mit Korollar 5.4.15 zeigen, die Idee des Beweises ist: Jede Funktion – auch eine, die keine Nullstellen hat, ist eine Polynomfunktion. Also gibt es immer Polynome ohne Nullstellen.

Definition 5.6.23 (Algebraische Vielfachheit)**.**
Es sei V ein Vektorraum der Dimension $n \in \mathbb{N}$ über einem Körper $\mathbb{K}$ und $\varphi \in \mathrm{End}_{\mathbb{K}}(V)$ ein Endomorphismus. Wir bezeichnen wir üblich mit $p_\varphi \in \mathbb{K}[X]$ das charakteristische Polynom von φ.

Wir sagen: Ein Eigenwert $\lambda \in \mathbb{K}$ von φ hat die *algebraische Vielfachheit* $r \in \mathbb{N}$, wenn es ein $q \in \mathbb{K}[X]$ gibt mit

$$p = (X - \lambda)^r q \quad \text{mit } q(\lambda) \neq 0.$$

Bemerkung 5.6.24. Die Zahl r aus Definition 5.6.23 existiert immer und ist eindeutig bestimmt. Dies sieht man leicht durch wiederholtes Anwenden von Lemma 5.4.13.

Was ist nun die Beziehung zwischen den beiden Vielfachheiten?

Proposition 5.6.25.
Es sei V ein Vektorraum der Dimension $n \in \mathbb{N}$ über einem Körper $\mathbb{K}$ und $\varphi \in \mathrm{End}_{\mathbb{K}}(V)$ ein Endomorphismus.
Für jeden Eigenwert λ von φ gilt:
Die geometrische Vielfachheit ist kleiner oder gleich der algebraischen Vielfachheit.

In Beispiel 5.6.17 haben wir gesehen, dass bei den Matrizen B und C der Eigenwert $\lambda = 2$ einmal die geometrische Vielfachheit 2 und einmal die geometrische Vielfachheit 1 hatte. Die algebraische Vielfachheit ist in beiden Fällen 2. Und dies ist genau der Grund, warum die eine Matrix diagonalisierbar war und die andere nicht. Allgemein gilt:

Satz 5.6.26 (Diagonalisierbarkeit und algebraische Vielfachheit)**.**
Es sei V ein Vektorraum der Dimension $n \in \mathbb{N}$ über einem Körper $\mathbb{K}$ und $\varphi \in \mathrm{End}_{\mathbb{K}}(V)$ ein Endomorphismus. Dann sind äquivalent:

(i) *φ ist diagonalisierbar.*

(ii) *Das charakteristische Polynom zerfällt in Linearfaktoren und für jeden Eigenwert ist die algebraische Vielfachheit gleich der geometrischen Vielfachheit.*

Die Bedingung, dass das charakteristische Polynom in Linearfaktoren zerfällt ist automatisch erfüllt, wenn der Körper algebraisch abgeschlossen ist, z.B. wenn $\mathbb{K} = \mathbb{C}$.

Beispiel 5.6.27 (Fibonacci-Folge)**.** Die *Fibonacci-Folge* [23] 0,1,1,2,3,5,8,13,… ist die eindeutig bestimmte Folge (f_n) ganzer Zahlen mit $f_1 := 1; f_0 := 0$, die der Rekursionsvorschrift

$$f_{n+1} = f_n + f_{n-1} \quad \text{für } n \in \mathbb{N},$$

genügt. Definieren wir $v_n \in \mathbb{Q}^2$ durch

$$v_n := \begin{pmatrix} f_{n+1} \\ f_n \end{pmatrix} \quad \text{für } n \in \mathbb{N}_0,$$

so können wir die Rekursionsvorschrift schreiben als

$$v_{n+1} = \begin{pmatrix} f_{n+2} \\ f_{n+1} \end{pmatrix} = \begin{pmatrix} f_{n+1} + f_n \\ f_{n+1} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} f_{n+1} \\ f_n \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} v_n \quad \text{für } n \in \mathbb{N}_0,.$$

[23] Diese Folge geht auf ein klassisches Problem im *Liber abbaci* des mittelalterlichen Mathematikers Leonardo von Pisa (ca. 1170-1240), Sohn von Bonaccio („Fibonacci“) zurück.

Ist $A \in \mathbb{Q}^{2\times 2}$ gegeben durch

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

so erhalten wir also induktiv die Formel

$$v_{n+1} = Av_n = A^2 v_{n-1} = \cdots = A^n v_1 = A^{n+1} v_0.$$

Es gilt also:

$$v_n = A^n v_0 \quad \text{für alle } n \in \mathbb{N}_0.$$

Um eine geschlossene Formel für v_{n+1} zu erhalten, müssen wir also die Potenzen von A bestimmen. Dazu bietet es sich an, A zu diagonalisieren: Das charakteristische Polynom von A ist

$$p_A = \det \begin{pmatrix} X-1 & -1 \\ -1 & X \end{pmatrix} = (X-1)\cdot X - (-1)\cdot(-1) = X^2 - X - 1.$$

Da dieses Polynom keine rationalen Nullstellen hat, hat A keine rationalen Eigenwerte, also ist A über $\mathbb{Q}$ nicht diagonalisierbar. Wir können aber unseren Körper $\mathbb{Q}$ so erweitern, dass p_A über dem größeren Körper zerfällt. Im vorliegenden Fall können wir z.B. A als *reelle* Matrix auffassen. [24] In $\mathbb{R}$ (bzw. einer beliebigen hinreichend großen Körpererweiterung von $\mathbb{Q}$) hat p_A dann die Nullstellen

$$\lambda_{1,2} = \frac{1}{2} \pm \sqrt{\frac{1}{4} + 1} = \frac{1 \pm \sqrt{5}}{2}.$$

Die Zahl $\varphi := \frac{1+\sqrt{5}}{2}$ ist der berühmte *goldene Schnitt*, der in der Kunst seit der Proportionenlehre der Renaissance eine zentrale Rolle spielt. Die andere Nullstelle von p_A bezeichnen wir mit φ^*. Es gilt dann

$$\varphi^2 = \varphi + 1, \quad (\varphi^*)^2 = \varphi^* + 1 \quad \text{und} \quad \varphi\varphi^* = \varphi^*\varphi = \frac{(1+\sqrt{5})(1-\sqrt{5})}{4} = \frac{1-5}{4} = -1. \qquad (*)$$

Der Eigenraum zu $\lambda_1 = \varphi$ lässt sich berechnen als:

$$E_{\lambda_1}(A) = \ker \begin{pmatrix} \varphi - 1 & -1 \\ -1 & \varphi \end{pmatrix} = \ker \begin{pmatrix} \varphi^2 - \varphi & -\varphi \\ -1 & \varphi \end{pmatrix} = \ker \begin{pmatrix} 1 & -\varphi \\ -1 & \varphi \end{pmatrix} = \ker \begin{pmatrix} 1 & -\varphi \\ 0 & 0 \end{pmatrix} = \mathrm{LH}_{\mathbb{R}} \left(\begin{pmatrix} \varphi \\ 1 \end{pmatrix} \right).$$

Analog ist der Eigenraum zu $\lambda_2 = \varphi^*$ gegeben als:

$$E_{\lambda_1}(A) = \ker \begin{pmatrix} \varphi^* - 1 & -1 \\ -1 & \varphi^* \end{pmatrix} = \ker \begin{pmatrix} (\varphi^*)^2 - \varphi^* & -\varphi^* \\ -1 & \varphi^* \end{pmatrix} = \ker \begin{pmatrix} 1 & -\varphi^* \\ -1 & \varphi^* \end{pmatrix} = \ker \begin{pmatrix} 1 & -\varphi^* \\ 0 & 0 \end{pmatrix} = \mathrm{LH}_{\mathbb{R}} \left(\begin{pmatrix} \varphi^* \\ 1 \end{pmatrix} \right).$$

Wir erhalten also

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

Weiter gilt

$$\det \begin{pmatrix} \varphi & \varphi^* \\ 1 & 1 \end{pmatrix} = \varphi - \varphi^* = \sqrt{5} \quad \Longrightarrow \quad \begin{pmatrix} \varphi & \varphi^* \\ 1 & 1 \end{pmatrix}^{-1} = \frac{1}{\sqrt{5}} \begin{pmatrix} 1 & -\varphi^* \\ -1 & \varphi \end{pmatrix}.$$

Wir erhalten also

$$A^n = \frac{1}{\sqrt{5}} \begin{pmatrix} \varphi & \varphi^* \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \varphi^n & 0 \\ 0 & (\varphi^*)^n \end{pmatrix} \begin{pmatrix} 1 & -\varphi^* \\ -1 & \varphi \end{pmatrix}.$$

[24] Es würde auch genügen, A als Matrix über $\mathbb{Q}[\sqrt{5}] = \mathrm{LH}_{\mathbb{Q}}(1, \sqrt{5}) = \{\alpha + \beta\sqrt{5} \mid \alpha, \beta \in \mathbb{Q}\}$ zu betrachten.

Speziell gilt

$$\begin{aligned}\begin{pmatrix} f_{n+1} \\ f_n \end{pmatrix} &= v_n \\ &= A^n v_0 \\ &= \frac{1}{\sqrt{5}} \begin{pmatrix} \varphi & \varphi^* \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \varphi^n & 0 \\ 0 & (\varphi^*)^n \end{pmatrix} \begin{pmatrix} 1 & -\varphi^* \\ -1 & \varphi \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ &= \frac{1}{\sqrt{5}} \begin{pmatrix} \varphi^{n+1} & (\varphi^*)^{n+1} \\ \varphi^n & (\varphi^*)^n \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \\ &= \frac{1}{\sqrt{5}} \begin{pmatrix} \varphi^{n+1} - (\varphi^*)^{n+1} \\ \varphi^n - (\varphi^*)^n \end{pmatrix}\end{aligned}$$

Wenn wir nur die zweite Komponente dieses Vektors betrachten, erhalten wir die Formel:

$$f_n = \varphi^n - (\varphi^*)^n = \frac{1}{\sqrt{5}} \cdot \left(\left(\frac{1+\sqrt{5}}{2} \right)^n - \left(\frac{1-\sqrt{5}}{2} \right)^n \right). \tag{5.6.1}$$

Auf den ersten Blick ist nicht einmal ersichtlich, dass dies überhaupt eine ganze Zahl ist! Die Tatsache, dass in der geschlossenen Formel für die Fibonacci-Zahlen irrationale Zahlen auftauchen, ist ein weiteres Beispiel dafür, dass Körpererweiterungen unabdingbar sind, wenn man Matrizen diagonalisieren will.

Bemerkung 5.6.28 (Lineare Rekursionsgleichungen)**.** Die Fibonacci-Folge ist ein Beispiel für eine rekursiv definierte Folge. Solche Folgen spielen in der Informatik eine zentrale Rolle. Allgemeiner betrachtet man Folgen $(a_n)_{n\in\mathbb{N}_0} \in \mathbb{K}^{\mathbb{N}_0}$, die durch eine *lineare Rekursionsgleichung* der Form

$$a_{n+k} = \beta_1 a_{n+k-1} + \cdots + \beta_k a_n \quad (\beta_1, \ldots, \beta_k \in \mathbb{K})$$

und *Anfangswerte* $a_0, \ldots, a_{k-1} \in \mathbb{K}$ gegeben sind. Setzt man

$$v_n := \begin{pmatrix} a_{n+k-1} \\ \vdots \\ a_n \end{pmatrix} \quad \text{und} \quad A := \begin{pmatrix} \beta_1 & \cdots & \beta_{k-1} & \beta_k \\ 1 & & 0 & 0 \\ & \ddots & & \vdots \\ 0 & & 1 & 0 \end{pmatrix},$$

so gilt $v_{n+1} = A v_n$, und daher

$$\begin{pmatrix} a_{n+k-1} \\ \vdots \\ a_n \end{pmatrix} = A^n \begin{pmatrix} a_{k-1} \\ \vdots \\ a_0 \end{pmatrix}.$$

Um diese Gleichung zu lösen, empfiehlt es sich, die Matrix A falls möglich zu diagonalisieren.

Zusammenfassung von Abschnitt 5.6

(1) Ein Vektor v ist Eigenvektor zum Eigenwert λ, wenn gilt $v \neq 0$ und $\varphi(v) = \lambda v$.

(2) Zu gegebenem Eigenwert λ ist der Eigenraum $E_\lambda(\varphi)$ die Menge aller Eigenwerte, zusammen mit dem Nullvektor. Die Dimension des Eigenraums ist die geometrische Vielfachheit des Eigenwertes.

(3) Die Eigenwerte sind die Nullstellen des charakteristischen Polynoms. Die Vielfachheit der Nullstelle im charakteristischen Polynom ist die algebraische Vielfachheit des Eigenwertes.

(4) Eine Matrix (oder ein Endomorphismus) ist diagonalisierbar, wenn es eine Basis aus Eigenvektoren gibt. Dies ist äquivalent dazu, dass das charakteristische Polynom in Linearfaktoren zerfällt und für jeden Eigenwert die geometrische Vielfachheit gleich der geometrischen Vielfachheit ist.

(5) Es gibt Matrizen mit reellen Einträgen, die über $\mathbb{R}$ nicht diagonalisierbar sind, aber über $\mathbb{C}$ schon. Andere Matrizen sind nicht einmal über $\mathbb{C}$ diagonalisierbar.

(6) Eigenvektoren zu unterschiedlichen Eigenwerten sind linear unabhängig.

(7) Diagonalisierung einer Matrix kann hilfreich sein, um große Potenzen der Matrix zu berechnen, dies ist hilfreich für lineare Rekursionsgleichungen (wie z.B. für die Fibonacci-Folge).

5.7. Trigonalisierbarkeit und der Satz von Cayley-Hamilton

Wir haben im Letzten Kapitel gesehen, wann ein Endomorphismus diagonalisierbar ist und dass es Endomorphismen gibt, die nicht diagonalisierbar sind. Nun wollen wir untersuchen, wann ein Endomorphismus trigonalisierbar ist, d.h. wann es eine geordnete Basis gibt, sodass die Darstellungsmatrix eine Dreiecksform hat. Dies ist eine schwächere Bedingung und somit sollten es mehr Endomorphismen geben, für die dies möglich ist.

Satz 5.7.1 (Trigonalisierbarkeit)**.**
Es sei V ein Vektorraum der Dimension $n \in \mathbb{N}$ über einem Körper $\mathbb{K}$ und $\varphi \in \mathrm{End}_{\mathbb{K}}(V)$ ein Endomorphismus. Dann sind äquivalent:

(i) φ ist trigonalisierbar, d.h. es gibt eine geordnete Basis B von V, sodass $\mathrm{M}_{\mathsf{B},\mathsf{B}}(\varphi)$ eine obere Dreiecksmatrix ist.

(ii) Das charakteristische Polynom zerfällt in Linearfaktoren.

Die Bedingung, dass das charakteristische Polynom in Linearfaktoren zerfällt ist automatisch erfüllt, wenn der Körper algebraisch abgeschlossen ist, z.B. wenn $\mathbb{K} = \mathbb{C}$.

Beweis. Wir beginnen mit der einfacheren Implikation:
(i) $\Longrightarrow$ (ii):
Angenommen, es gibt eine geordnete Basis B von V, sodass φ bezüglich B Diagonalgestalt hat,

d.h.

$$A := \mathrm{M}_{\mathsf{B},\mathsf{B}}(\varphi) = \begin{pmatrix} \lambda_1 & * & \cdots & * \\ & \lambda_2 & \cdots & * \\ & & \ddots & \vdots \\ & & & \lambda_n \end{pmatrix}.$$

Da das charakteristische Polynom $p_\varphi \in \mathbb{K}[X]$ unabhängig von der Wahl der Basis ist, können wir jede beliebige geordnete Basis verwenden, um es zu berechnen. Insbesondere also die hier gewählte B:

$$\begin{aligned} p_\varphi &= p_A \\ &= \det(X\mathbb{1}_n - A) \\ &= \det\begin{pmatrix} X-\lambda_1 & * & \cdots & * \\ & X-\lambda_2 & \cdots & * \\ & & \ddots & \vdots \\ & & & X-\lambda_n \end{pmatrix}. \end{aligned}$$

Die Determinante einer oberen Dreiecksmatrix ist einfach das Produkt der Diagonalelemente (Proposition 5.3.6), also gilt:

$$p_\varphi = (X-\lambda_1)\cdots(X-\lambda_n)$$

und das charakteristische Polynom zerfällt in Linearfaktoren.

(ii) $\Longrightarrow$ (i):

Diese Implikation beweisen wir mit vollständiger Induktion nach $n = \dim(V)$:

Induktionsanfang $n = 1$:

Wenn V eindimensional ist, dann wählen wir ein beliebiges Element $v \in V$ mit $v \neq 0$ aus und erhalten $\mathsf{B} := (v)$ eine geordnete Basis von V. Die Matrix $\mathrm{M}_{\mathsf{B},\mathsf{B}}(\varphi) \in \mathbb{K}^{1\times 1} = \mathbb{K}$ ist einfach eine Zahl und somit eine Dreiecksmatrix.

Induktionsschritt:

Es sei $n \in \mathbb{N}$ so gewählt, dass jeder Endomorphismus eines n-dimensionalen $\mathbb{K}$-Vektorraums mit zerfallendem charakteristischen Polynom trigonalisierbar ist.

Es sei nun $\dim(V) = n+1$ Das charakteristische Polynom p_φ hat Grad $n+1 \geq 1$ (Lemma 5.6.12) und zerfällt in Linearfaktoren. Also gilt insbesondere, dass p_φ in $\mathbb{K}$ mindestens eine Nullstelle λ_1 hat. Da λ_1 eine Nullstelle des charakteristischen Polynoms von φ ist, bedeutet dies, dass λ_1 ein Eigenwert von φ ist, d.h. es gibt ein $v_1 \in V$ mit

$$\varphi(v_1) = \lambda_1 v_1 \quad \text{und} \quad v_1 \neq 0.$$

Aus $v_1 \neq 0$ folgt, dass der Untervektorraum $\mathrm{LH}_{\mathbb{K}}(v_1)$ eindimensional ist. Mit Satz 4.4.11 wählen wir nun ein Komplement W von $\mathrm{LH}_{\mathbb{K}}(v_1)$ in V aus:

$$V = \mathrm{LH}_{\mathbb{K}}(v_1) \oplus W.$$

Die Dimension von W ist dann genau $\dim(V) - 1 = n$. Wir wählen nun eine geordnete Basis $\mathsf{B}_W := (w_1, \ldots, w_n)$ von W. Dann ist

$$\mathsf{B}_1 := (v_1, w_1, \ldots, w_n)$$

eine Basis von V und wir haben

$$\mathrm{M}_{\mathsf{B}_1,\mathsf{B}_1}(\varphi) = \begin{pmatrix} \lambda_1 & * & \cdots & * \\ 0 & * & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & * & \cdots & * \end{pmatrix} = \left(\begin{array}{c|ccc} \lambda_1 & * & \cdots & * \\ \hline 0 & & & \\ \vdots & & C & \\ 0 & & & \end{array}\right)$$

mit einer Matrix $C \in \mathbb{K}^{n\times n}$. Nun können wir mit Hilfe dieser Matrix von φ das charakteristische Polynom bestimmen und erhalten:

$$p_\varphi = \det(X\mathbb{1}_{n+1} - A) = \left(\begin{array}{c|ccc} X-\lambda_1 & * & \cdots & * \\ \hline 0 & & & \\ \vdots & & X\mathbb{1}_n - C & \\ 0 & & & \end{array}\right) = (X-\lambda_1)p_C.$$

Da p_φ nach Voraussetzung in Linearfaktoren zerfällt, folgt daraus, dass auch p_C in Linearfaktoren zerfällt. Zu welcher Abbildung gehört nun diese Matrix C?

Betrachten wir zunächst die *Projektion* auf W:

$$\pi : V = \mathrm{LH}_{\mathbb{K}}(v_1) \oplus W \to W, \quad \mu v_1 + w \mapsto w$$

Diese Abbildung ist wohldefiniert und linear, weil W die direkte Summe von $\mathrm{LH}_{\mathbb{K}}(v_1)$ und W ist.

Wenn wir einen Vektor aus W nehmen und ihn in φ einsetzen, erhalten wir einen Vektor aus $V = \mathrm{LH}_{\mathbb{K}}(v_1) \oplus W$, der sich eindeutig schreiben lässt als Summe eines Vielfachen von v_1 und eines Vektors aus W. Umgekehrt gibt es eine Inklusionsabbildung

$$\iota : W \to V, \quad w \mapsto w.$$

Wir können diese Abbildungen nun in folgender Weise verketten:

$$\psi := \pi \circ \varphi \circ \iota : W \to W$$

und erhalten so einen Endomorphismus des Vektorraums W. Es gilt:

$$\begin{aligned} \mathrm{M}_{\mathsf{B}_W,\mathsf{B}_W}(\psi) &= \mathrm{M}_{\mathsf{B}_W,\mathsf{B}_1}(\pi) \cdot \mathrm{M}_{\mathsf{B}_1,\mathsf{B}_1}(\varphi) \cdot \mathrm{M}_{\mathsf{B}_1,\mathsf{B}_W}(\iota) \\ &= \left(\begin{array}{c|c} 0 & \\ \vdots & \mathbb{1}_n \\ 0 & \end{array}\right) \cdot \left(\begin{array}{c|ccc} \lambda_1 & * & \cdots & * \\ \hline 0 & & & \\ \vdots & & C & \\ 0 & & & \end{array}\right) \cdot \left(\begin{array}{c|ccc} 0 & \cdots & 0 \\ \hline & & \\ & \mathbb{1}_n & \\ & & \end{array}\right) \\ &= C. \end{aligned}$$

Zusammengefasst können wir also sagen: $\psi \in \mathrm{End}_{\mathbb{K}}(W)$ ist ein Endomorphismus eines n-dimensionalen $\mathbb{K}$-Vektorraums W und das charakteristische Polynom $p_\varphi = p_C$ zerfällt in Linearfaktoren. Also gibt es nach Induktionsvoraussetzung eine geordnete Basis $\widetilde{\mathsf{B}_W} = (u_1, \ldots, u_n)$ von W, sodass ψ bezüglich dieser Basis obere Dreiecksgestalt hat.

Wenn wir diese Basis nun von v_1 zu einer Basis von V zusammensetzen:

$$\mathsf{B}_\Delta := (v_1, u_1, \ldots, u_n),$$

so sehen wir, dass

$$\mathrm{M}_{\mathsf{B}_\Delta,\mathsf{B}_\Delta}(\varphi) = \left(\begin{array}{c|ccc} \lambda_1 & * & \cdots & * \\ \hline 0 & & & \\ \vdots & & \mathrm{M}_{\widetilde{\mathsf{B}_W},\widetilde{\mathsf{B}_W}}(\psi) & \\ 0 & & & \end{array}\right)$$

und dies ist eine obere Dreiecksmatrix. □

Alles, was nun kommt ist für die Klausur *Lineare Algebra I* im Wintersemester 2020/2021 nicht mehr klausurrelevant, die folgenden Inhalte sind aber wesentlich für die Veranstaltungen *Lineare Algebra II* und *Lineare Algebra II für Informatik* im kommenden Sommersemester.

Satz 5.7.2 (Cayley-Hamilton[25])**.** *Es sei V ein Vektorraum der Dimension $n \in \mathbb{N}$ über einem Körper $\mathbb{K}$ und $\varphi \in \mathrm{End}_{\mathbb{K}}(V)$ ein Endomorphismus. Dann gilt:*

$$p_\varphi(\varphi) = 0.$$

Ein Endomorphismus eingesetzt in sein eigenes charakteristisches Polynom ergibt also immer Null.

Beweis. Wir werden diesen Satz nur unter der Zusatzbedingung beweisen, dass sich der Körper $\mathbb{K}$ zu einem algebraisch abgeschlossenen Körper $\mathbb{L}$ erweitern lässt, d.h. es gibt einen algebraisch abgeschlossenen Körper $\mathbb{L}$, der $\mathbb{K}$ als Unterkörper enthält. Dies ist für die Körper $\mathbb{Q}, \mathbb{R}$ und $\mathbb{C}$ der Fall, da $\mathbb{Q} \subseteq \mathbb{R} \subseteq \mathbb{C}$ und $\mathbb{C}$ algebraisch abgeschlossen ist nach dem Fundamentalsatz der Algebra (Satz 3.5.10).

Man kann aber zeigen, dass sich *jeder* Körper in einen algebraisch abgeschlossenen Körper einbetten lässt.[26]

Wir wählen also zuerst eine Basis B_1 von V als $\mathbb{K}$-Vektorraum aus und erhalten eine Matrix

$$A := \mathrm{M}_{\mathsf{B}_1,\mathsf{B}_1}(\varphi) \in \mathbb{K}^{n\times n}.$$

Es bleibt zu zeigen, dass $p_A(A) = 0$.

Es gilt: $\mathbb{K}^{n\times n} \subseteq \mathbb{L}^{n\times n}$ und somit können wir A auch als Matrix mit Einträgen in $\mathbb{L}$ auffassen. Aus der Definition des charakteristischen Polynoms $p_\varphi = p_A = \det(X\mathbb{1}_n - A) \in \mathbb{K}[X] \subseteq \mathbb{L}[X]$ sieht man sofort, dass es egal ist, ob wir hier den Grundkörper $\mathbb{K}$ oder $\mathbb{L}$ betrachten.

Deshalb werden wir für den Rest des Beweises über dem algebraisch abgeschlossenen Körper $\mathbb{L}$ arbeiten.

Das charakteristische Polynom zerfällt hier in Linearfaktoren:

$$p_A = (X-\lambda_1)\cdots(X-\lambda_n) \quad \text{mit } \lambda_1,\ldots,\lambda_n \in \mathbb{L}.$$

Wir haben diesmal mehrfach vorkommende Linearfaktoren nicht zusammengefasst, sondern sie einzeln gelassen.

[25] nach ARTHUR CAYLEY (siehe Fußnote auf Seite 61) und WILLIAM ROWAN HAMILTON, irischer Mathematiker und Physiker, 1805–1865.

[26] In dieser Allgemeinheit benötigt man hierfür das sogenannte *Auswahlaxiom*, der Beweis ist nicht konstruktiv. Genau genommen reicht für das, was wir hier wirklich brauchen aber die schwächere Aussage aus, dass es eine Körpererweiterung von $\mathbb{K}$ gibt, in der p_φ in Linearfaktoren zerfällt. Dies lässt sich auch ohne Auswahlaxiom beweisen, würde uns hier aber zu sehr aufhalten.

Nach Satz 5.7.1 ist A ähnlich zu einer oberen Dreiecksmatrix

$$C = \begin{pmatrix} \lambda_1 & * & \cdots & * \\ & \lambda_2 & \cdots & * \\ & & \ddots & \vdots \\ & & & \lambda_n \end{pmatrix} \in \mathbb{L}^{n\times n}.$$

Da sich das charakteristische Polynom nicht ändert, wenn wir zu einer ähnlichen Basis übergehen, gilt $p_C = p_A = (X-\lambda_1)\cdots(X-\lambda_n)$. Es bleibt zu zeigen, dass $p_C(C) = 0$.

Wir setzen $C \in \mathbb{L}^{n\times n}$ in das Polynom $p_C = (X-\lambda_1 1)\cdots(X-\lambda_n 1) \in \mathbb{L}[X]$ ein und erhalten:

$$p(C) = (C-\lambda_1 \mathbb{1}_n)\cdots(C-\lambda_n \mathbb{1}_n) \in \mathbb{L}^{n\times n}.$$

Für jedes $k \in \{1,\dots,n\}$ setzen wir

$$W_k := \mathrm{LH}_{\mathbb{L}}(e_1,\dots,e_k) \subseteq \mathbb{L}^n.$$

Außerdem sei $W_0 := \{0\} \subseteq \mathbb{L}^n$. Wir haben dann die Kette[27] von Untervektorräumen:

$$\{0\} = W_0 \subsetneq W_1 \subsetneq \cdots \subsetneq W_{n-1} \subsetneq W_n = \mathbb{L}^n.$$

Es sei nun $k \in \{1,\dots,n\}$. Wir untersuchen die Matrix $C-\lambda_k \mathbb{1}_n$. Diese Matrix ist eine obere Dreiecksmatrix und hat in der Position (k,k) den Wert 0. Somit gilt für die ersten k Spalten, dass sie alle im Raum W_{k-1} liegen.

Da in den Spalten der Matrix C die Bilder der Standardbasisvektoren stehen, bedeutet dies, dass jedes $e_1,\dots,e_k$ in den Unterraum W_{k-1} abgebildet wird und folglich gilt:

$$\forall k \in \{1,\dots,n\},\ \forall v \in W_k: \quad (C-\lambda_k \mathbb{1}_n)v \in W_{k-1}.$$

Es sei nun $v \in \mathbb{L}^n = W_n$. Dann gilt:

$$(C-\lambda_n \mathbb{1}_n)v \in W_{n-1}.$$

Anwenden von $C-\lambda_{n-1}\mathbb{1}_n$ liefert dann:

$$(C-\lambda_{n-1}\mathbb{1}_n)(C-\lambda_n \mathbb{1}_n)v \in W_{n-2}.$$

Nun wenden wir der Reihe nach die weiteren Matrizen an, bis wir schließlich erhalten:

$$(C-\lambda_1)\cdots(C-\lambda_{n-1}\mathbb{1}_n)(C-\lambda_n \mathbb{1}_n)v \in W_0 = \{0\}.$$

Also gilt $(p(C))v = 0$, und weil $v \in \mathbb{L}^n$ beliebig war, muss die Matrix $p(C)$ die Nullmatrix sein.

Das war zu zeigen. □

Bemerkung 5.7.3. Es gibt auch einen kürzeren Beweis:

$$p_A(X) = \det(X\mathbb{1}_n - A), \quad \Longrightarrow \quad p_A(A) = \det(A\mathbb{1}_n - A) = \det(A-A) = \det(0) = 0.$$

Allerdings ist dieser Beweis so nicht zulässig! Der Schritt

$$p_A(X) = \det(X\mathbb{1}_n - A) \quad \Longrightarrow \quad p_A(A) = \det(A\mathbb{1}_n - A)$$

ist leider nicht erlaubt. Wieso?

[27] Eine solche Konstellation von Untervektorräumen nennt man auch eine *Fahne*.

Beispiel 5.7.4. (i) Der Satz von Cayley-Hamilton lässt sich manchmal verwenden, um Potenzen einer Matrix zu berechnen, wenn der Exponent nicht zu groß ist oder die Matrix nicht diagonalisiert werden kann.

Nehmen wir an, eine Matrix $A \in \mathbb{R}^{5\times 5}$ sei gegeben mit charakteristischem Polynom

$$p_A = X^5 - X - 1 \in \mathbb{R}[X].$$

Der Satz von Cayley-Hamilton (Satz 5.7.2) sagt uns nun:

$$A^5 - A - \mathbb{1}_5 = 0 \quad \text{oder} \quad A^5 = A + \mathbb{1}_5.$$

Wir wissen also, dass A^5 die gleiche Matrix ist wie $A + \mathbb{1}_5$. Angenommen, wir wollen A^{13} berechnen. Dann können wir zum Beispiel so vorgehen:

$$\begin{aligned} A^{13} &= A^5A^5A^3 = (A+\mathbb{1}_5)(A+\mathbb{1}_5)A^3 \\ &= (A^2+2A+\mathbb{1}_5)A^3 = A^5 + 2A^4 + A^3 = (A+\mathbb{1}_5) + 2A^4 + A^3 = 2A^4 + A^3 + A + \mathbb{1}_5. \end{aligned}$$

Man kann also die Rechenregel $A^5 = A + \mathbb{1}_5$ so oft anwenden, bis man alle Exponenten von A unter 5 gebracht hat[28], sodass man im Endeffekt nur die Potenzen $A^0 = \mathbb{1}_5, A, A^2, A^3, A^4$ benötigt, um beliebige Potenzen zu berechnen. Allerdings erhält man auf diese Weise keine geschlossene Form für die allgemeine Potenz A^n.

(ii) Eine weitere schöne Anwendung von Satz 5.7.2 ist die folgende: Gegeben sei eine invertierbare Matrix $A \in \mathbb{K}^{n\times n}$ mit charakteristischem Polynom

$$p_A = X^n + c_{n-1}X^{n-1} + \cdots + c_1X + c_0.$$

Wenn $c_0 = 0$ wäre, dann wäre 0 ein Eigenwert von A und dann wäre A nicht invertierbar. Also ist $c_0 \neq 0$. Nach Cayley-Hamilton gilt nun:

$$0 = A^n + c_{n-1}A^{n-1} + \cdots + c_1A + c_0\mathbb{1}_n.$$

Diese Gleichung multiplizieren wir nun mit der Inversen von A:

$$0 = A^{n-1} + c_{n-1}A^{n-2} + \cdots + c_1\mathbb{1}_n + c_0A^{-1}$$

und lösen nach A^{-1} auf – dies ist möglich, weil c_0 ja nicht 0 ist:

$$A^{-1} = -\frac{1}{c_0}\left(A^{n-1} + c_{n-1}A^{n-2} + \cdots + c_1\mathbb{1}_n\right).$$

Dies ist oft eine effektive Möglichkeit, die Inverse einer Matrix zu bestimmen – ohne Gauß-Algorithmus.[29]

Insbesondere, wenn das charakteristische Polynom bereits bekannt ist, kann dies eine sehr elegante Methode sein, die Inverse zu berechnen.

[28] Was wir eigentlich tun ist, das Polynom X^{13} modulo dem Polynom $X^5 - X - 1$ zu berechnen.

[29] Allerdings musste man für die Bestimmung des charakteristischen Polynoms ja bereits eine Determinante berechnen und hat dies ja vielleicht mit einer Variante des Gauß-Algorithmus gemacht.

Wenn also – wie oben – die Matrix $A \in \mathbb{R}^{5\times 5}$ das charakteristische Polynom $p_A = X^5 - X - 1$ besitzt, dann gilt:

$$\begin{aligned} A^5 - A - \mathbb{1}_5 &= 0 \\ A^5 - A &= \mathbb{1}_5 \\ A(A^4 - \mathbb{1}_5) &= \mathbb{1}_5 \\ A^4 - \mathbb{1}_5 &= A^{-1}. \end{aligned}$$

Interessant ist auch das theoretische Resultat: Wenn eine Matrix A invertierbar ist, dann ist die Inverse eine Linearkombination von Potenzen von A, also ein polynomieller Ausdruck in A. Das ist weder bei der Berechnung mit dem Gauß-Algorithmus – noch bei der Inversen-Formel mit der Adjunkten (Cramersche Regel, Satz 5.3.17) offensichtlich.

Zusammenfassung von Abschnitt 5.7

(1) Eine Matrix ist trigonalisierbar, wenn ihr charakteristisches Polynom in Linearfaktoren zerfällt.

(2) Jede (reelle oder komplexe) Matrix ist über $\mathbb{C}$ trigonalisierbar.

(3) Der Satz von Cayley-Hamilton besagt: Wenn man eine Matrix in ihr charakteristisches Polynom einsetzt, erhält man die Nullmatrix.

(4) Der Satz von Cayley-Hamilton kann benutzt werden, um die Inverse einer Matrix als Linearkombination von Potenzen der Matrix zu schreiben – wenn das charakteristische Polynom bekannt ist.

5.8. Nilpotente Endomorphismen und die Jordansche Normalform

Wir beginnen mit einer Definition:

Definition 5.8.1.
Es sei R ein Ring und $a \in R$ ein Element. Dann heißt a *nilpotent*, wenn es ein $k \in \mathbb{N}$ gibt mit

$$a^k = 0.$$

Insbesondere ist dieses Konzept interessant für den Ring $\mathbb{K}^{n\times n}$ der $(n \times n)$-Matrizen mit Einträgen aus einem Körper $\mathbb{K}$ oder dem Ring $\mathrm{End}_{\mathbb{K}}(V)$ der Endomorphismen eines $\mathbb{K}$-Vektorraums. Dies führt zum Konzept der *nilpotenten Matrix*, bzw. des *nilpotenter Endomorphismus*[nilpotenten Endomorphismus].

Beispiel 5.8.2. Die Matrix A aus Beispiel 5.5.1 (i) ist nilpotent. Es gilt nämlich

$$A^3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Proposition 5.8.3.
Es sei V ein Vektorraum der Dimension $n \in \mathbb{N}$ über einem Körper $\mathbb{K}$. Für einen Endomorphismus $\varphi \in \mathrm{End}_{\mathbb{K}}(V)$ sind äquivalent:

(i) φ *ist nilpotent, d.h.* $\exists k \in \mathbb{N} : \varphi^k = 0$.

(ii) $\varphi^n = 0$.

(iii) $p_\varphi = X^n$.

(iv) φ *hat Eigenwert* 0 *mit algebraischer Vielfachheit* n.

Da ein nilpotenter Endomorphismus nach endlich vielen Schritten der Nullendomorphismus wird, ist es von nilpotenten Endomorphismen immer besonders leicht, hohe Potenzen zu bestimmen. Wir haben somit zwei Klassen von Endomorphismen gefunden, von denen wir hohe Potenzen berechnen können: diagonalisierbare und nilpotente. Wie ist nun der Zusammenhang zwischen diesen beiden Klassen? Das beantworten wir nun:

Proposition 5.8.4.
Es sei V ein Vektorraum der Dimension $n \in \mathbb{N}$ über einem Körper $\mathbb{K}$. Für einen Endomorphismus $\varphi \in \mathrm{End}_{\mathbb{K}}(V)$ sind äquivalent:

(i) φ *ist nilpotent und diagonalisierbar.*

(ii) $\varphi = 0$.

Es gibt also – bis auf den Nullendomorphismus – keine Endomorphismen, die sowohl nilpotent als auch diagonalisierbar sind.

Wir wollen nun schauen, inwiefern wir diagonalisierbare und nilpotente Endomorphismen kombinieren können. Angenommen, $\varphi = \varphi_S + \varphi_N$ mit φ_S diagonalisierbar und φ_N nilpotent. Können wir dann das Wissen über die Potenzen von φ_S und φ_N verwenden, um Potenzen von φ zu bestimmen. Erinnern wir hier einmal an den *binomischen Lehrsatz*:

Satz 5.8.5 (Binomischer Lehrsatz)**.**
Es sei R ein Ring und $a, b \in R$ seinen Elemente, die kommutieren, d.h.

$$ab = ba.$$

Dann gilt für jedes $m \in \mathbb{N}_0$:

$$(a+b)^m = \sum_{k=0}^{m} \binom{m}{k} a^{m-k} b^k.$$

Hier sind die $\binom{m}{k} := \frac{m \cdot (m-1) \cdots (m-k+2) \cdot (m-k+1)}{k \cdot k-1 \cdots 2 \cdot 1}$ *die* Binomialkoeffizienten.

Beweis. Dieser Satz wird in der Analysis- bzw. HM1-Vorlesung normalerweise nur für $R = \mathbb{R}$ oder $R = \mathbb{C}$ bewiesen (mit vollständiger Induktion), der Beweis geht aber genauso für alle Ringe. Wichtig ist nur, dass a und b kommutieren. □

Bemerkung 5.8.6. Für $k > m$ ist $\binom{m}{k} = 0$, weil im Zähler dann eine Null steht. Mit dieser Erkenntnis können wir den binomischen Lehrsatz auch so umschreiben:

$$(a+b)^m = \sum_{k=0}^{\infty} \binom{m}{k} a^{m-k} b^k.$$

Dies sieht zwar formal wie eine unendliche Reihe aus, da aber für $k > n$ alle Terme Null werden, ergibt sich genau dieselbe endliche Summe wie oben.

Wir können also nun Potenzen von Endomorphismen berechnen, wenn sie sich als Summe von einem diagonalisierbaren und einem nilpotenten Summanden schreiben lassen, vorausgesetzt die beiden Summanden kommutieren:

Bemerkung 5.8.7.
Es sei $\mathbb{K}$ ein Körper und $n \in \mathbb{N}$. Gegeben seien Matrizen $S, N \in \mathbb{K}^{n\times n}$ mit S diagonalisierbar, N nilpotent und $SN = NS$. Dann gilt für jedes $m \in \mathbb{N}_0$:

$$(S+N)^m = \sum_{k=0}^{\infty} \binom{m}{k} S^{m-k} N^k.$$

Da N nilpotent ist, folgt mit Proposition 5.8.3, dass $N^n = 0$ ist. Somit gilt:

$$(S+N)^m = \sum_{k=0}^{n-1} \binom{m}{k} S^{m-k} N^k.$$

Die Anzahl der Summanden hängt nun also nicht mehr vom Exponenten m ab, sondern nur noch von der (konstanten) Dimension n. Die Potenzen S^{m-k} kann man nun berechnen, weil S als diagonalisierbar vorausgesetzt wurde.

Es bleibt die Frage, welche Matrizen/Endomorphismen sich so schreiben lassen. Diese Frage werden wir mit Satz 5.8.11 beantworten.

Zuvor wollen wir Beispiel 5.5.1 (i) verallgemeinern:

Definition 5.8.8 (Jordan-Block)**.**
Es sei $\mathbb{K}$ ein Körper und $n \in \mathbb{N}$.

(a) Der *nilpotente Jordan-Block* der Größe n ist definiert als

$$N_n := \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & & \ddots & & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix} \in \mathbb{K}^{n\times n}.$$

Er entspricht (bezüglich der Standardbasis) der linearen Abbildung:

$$v_n \mapsto v_{n-1} \mapsto v_{n-2} \mapsto \cdots \mapsto v_2 \mapsto v_1 \mapsto 0.$$

Der nilpotenter Jordan-Block der Größe 1 ist einfach nur das Nullelement des Körpers:

$$N_1 = 0 \in \mathbb{K}^{1\times 1} = \mathbb{K}.$$

(b) Der *Jordan-Block* der Größe n mit Eigenwert $\lambda \in \mathbb{K}$ ist definiert als

$$J_n(\lambda) := \lambda \mathbb{1}_n + N_n = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & & \ddots & & \vdots \\ 0 & 0 & 0 & \cdots & \lambda & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda \end{pmatrix} \in \mathbb{K}^{n\times n}.$$

Der Jordan-Block der Größe 1 mit Eigenwert λ ist einfach nur der Wert λ:

$$J_1(\lambda) = \lambda \in \mathbb{K}^{1\times 1} = \mathbb{K}.$$

Ein nilpotenter Jordan-Block ist gerade ein Jordan-Block mit Eigenwert 0.

Lemma 5.8.9. *Es sei V ein Vektorraum über einem Körper $\mathbb{K}$ und $\varphi \in \mathrm{End}_{\mathbb{K}}(V)$ ein Endomorphismus. Es sei $(v_1, \dots v_m)$ eine φ-Kette, d.h. $v_1, \dots, v_m$ sind Vektoren in $V \setminus \{0\}$ mit*

$$\varphi(v_1) = 0 \quad \textit{und} \quad \forall k \in \{1, \dots, m\} : \varphi(v_{k+1}) = v_k$$

Dann sind die Vektoren $v_1, \dots, v_m$ paarweise verschieden und $\{v_1, \dots, v_m\}$ ist linear unabhängig über $\mathbb{K}$.

Beweis. Wir beweisen dies per Induktion über $m \in \mathbb{N}$:

Induktionsanfang $m = 1$:
Nach Voraussetzung ist $v_1 \neq 0$. Also ist $\{v_1\}$ linear unabhängig.

Induktionsschritt:
Es sei $m \in \mathbb{N}$ so gewählt, dass alle φ-Ketten der Länge m aus m paarweise verschieden und linear unabhängig sind.

Gegeben sei nun eine φ-Kette $(v_1, \dots, v_{m+1})$.

Dass der Vektor v_{m+1} nicht identisch mit einem der Vektoren $v_1, \dots, v_m$ ist, folgt daraus, dass $\varphi^m(v_{m+1}) \neq 0$, während $\varphi^m(v_j) = 0$ für alle anderen v_j.

Es sei

$$\sum_{j=1}^{m+1} \alpha_j v_j = 0$$

eine Darstellung des Nullvektors als Linearkombination der Vektoren v_j. Wir wenden auf diese Gleichung φ an und erhalten:

$$\sum_{j=1}^{m} \alpha_{j+1} v_j = 0$$

Da die Vektoren $(v_1, \dots, v_m)$ linear unabhängig sind, folgt hieraus, dass alle Skalare $\alpha_2, \dots, \alpha_{m+1}$ Null sind. Damit vereinfacht sich die Formel

$$\sum_{j=1}^{m+1} \alpha_j v_j = 0$$

zu

$$\alpha_1 v_1 = 0$$

und weil $v_1 \neq 0$ ist, gilt $\alpha_1 = 0$. Das beendet den Beweis. □

Man beachte, dass es eigentlich ausgereicht hätte, zu fordern, dass $v_1 \neq 0$ ist. Daraus folgt, dass alle anderen Vektoren in der Kette auch ungleich 0 sind.

Satz 5.8.10 (Einfach nilpotente Endomorphismen)**.** *Es sei V ein Vektorraum der Dimension $n \in \mathbb{N}$ über einem Körper $\mathbb{K}$. Für einen Endomorphismus $\varphi \in \mathrm{End}_{\mathbb{K}}(V)$ sind äquivalent:*

(i) *φ ist* einfach nilpotent*, d.h. φ ist nilpotent und* $\dim(\ker(\varphi)) = 1$.

(ii) *Es gibt eine φ-Kette der Länge n, d.h. es gibt Vektoren $v_1, \dots, v_n$, die alle nicht* 0 *sind mit*

$$\varphi(v_1) = 0 \quad \textit{und} \quad \varphi(v_{k+1}) = v_k \textit{ für } k \in \{1, \dots, n\}.$$

(iii) Es gibt eine geordnete Basis B *von* V *mit*

$$\mathrm{M}_{\mathsf{B},\mathsf{B}}(\varphi) = N_n := \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & & \ddots & & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix} \in \mathbb{K}^{n\times n}.$$

Beweis. Die Implikation „(ii) $\Longrightarrow$ (iii)“ folgt aus Lemma 5.8.9 und der Tatsache, dass in den Spalten der Matrix die Bilder der Basisvektoren stehen. Die Implikation „(iii) $\Longrightarrow$ (i)“ ist direktes Nachrechnen.

Zeigen wir also „(i) $\Longrightarrow$ (ii)“:

Nach Voraussetzung ist $\ker(\varphi)$ ein 1-dimensionaler Untervektorraum von V. Es gibt somit eine Basis von $\ker(\varphi)$, die aus genau einem Element besteht: $\ker(\varphi) = \mathrm{LH}_{\mathbb{K}}(()v_1)$. Da ein Basisvektor niemals der Nullvektor ist, gilt $v_1 \neq 0$.

Falls es einen Vektor $v_2 \in V$ gibt mit $\varphi(v_2) = v_1$, wählen wir einen solchen aus und dann einen Vektor v_3 mit $\varphi(v_3) = v_2$ und so weiter, bis wir bei einem Vektor $v_m \in V$ angelangt sind mit $v_m \notin \mathrm{Bild}(\varphi)$.

Nach Lemma 5.8.9 ist eine solche Kette immer linear unabhängig. Da V die endliche Dimension $n \in \mathbb{N}$ hat, muss dieses Verfahren also nach $m \leq n$ Schritten enden und wir erhalten eine Kette $(v_1,\ldots,v_m)$, die sich nicht weiter verlängern lässt.

Wir wollen zeigen, dass die m-elementige linear unabhängige Teilmenge $\{v_1,\ldots,v_m\}$ ein Erzeugendensystem für V ist, d.h.

$$V \subseteq \mathrm{LH}_{\mathbb{K}}(v_1,\ldots,v_m)$$

Dann folgt, dass wir eine Basis mit m Elementen eines n-dimensionalen Vektorraums haben und somit $m = n$.

Bis jetzt haben wir noch nicht verwendet, dass φ nilpotent ist. Dies machen wir nun. Nach Proposition 5.8.3 gilt $\varphi^n = 0$ und somit $\ker(\varphi^n) = V$.

Wir werden nun die folgende Aussage per Induktion zeigen:

$$\forall k \in \mathbb{N}\colon\ \ker(\varphi^k) \subseteq \mathrm{LH}_{\mathbb{K}}(v_1,\ldots,v_m).$$

Wenn wir dies gezeigt haben, setzen wir einfach $k = n$ ein und die Aussage ist bewiesen.

Induktionsanfang $k = 1$**:**
Nach Wahl von v_1 gilt:

$$\ker(\varphi) = \mathrm{LH}_{\mathbb{K}}(v_1) \subseteq \mathrm{LH}_{\mathbb{K}}(v_1,\ldots,v_m).$$

Induktionsschritt:
Es sei nun $k \in \mathbb{N}$ so gewählt, dass $\ker(\varphi^k) \subseteq \mathrm{LH}_{\mathbb{K}}(v_1,\ldots,v_m)$ und $v \in \ker(\varphi^{k+1})$. Wir werden zeigen, dass $v \in \mathrm{LH}_{\mathbb{K}}(v_1,\ldots,v_m)$.

Es gilt nun $\varphi^k(\varphi(v)) = 0$, woraus folgt, dass $\varphi(v) \in \ker(\varphi^k)$. Nach Induktionsvoraussetzung gilt nun also:

$$\varphi(v) = \sum_{j=1}^{m} \alpha_j v_j \quad \text{mit } \alpha_j \in \mathbb{K}.$$

Diese Gleichung lösen wir nun nach $\alpha_m v_m$ auf:

$$\alpha_m v_m = \varphi(v) - \sum_{j=1}^{m-1} \alpha_j v_j.$$

Für jedes v_j auf der rechten Seite der Gleichung dürfen wir schreiben $v_j = \varphi(v_{j+1})$ und mit der Linearität von φ ergibt sich:

$$\alpha_m v_m = \varphi(v) - \sum_{j=1}^{m-1} \alpha_j \varphi(v_{j+1}) = \varphi\left(v - \sum_{j=1}^{m-1} \alpha_j v_{j+1}\right).$$

Falls $\alpha_m \neq 0$ ist, dann können wir beide Seiten durch α_m dividieren und erhalten $v_m \in \text{Bild}(\varphi)$, was nicht sein kann, weil wir angenommen haben, dass sie Kette $(v_1, \ldots, v_m)$ nicht weiter verlängert werden kann.

Also muss gelten: $\alpha_m = 0$ und folglich:

$$0 = \varphi\left(v - \sum_{j=1}^{m-1} \alpha_j v_{j+1}\right).$$

Demnach können wir schließen:

$$v - \sum_{j=1}^{m-1} \alpha_j v_{j+1} \in \ker(\varphi) = \text{LH}_{\mathbb{K}}(v_1).$$

Wenn wir dies nun wieder nach v auflösen, erhalten wir, dass v eine Linearkombination der Vektoren $v_1, \ldots, v_m$ ist. Das war zu zeigen. □

Satz 5.8.11 (Jordansche Normalform)**.** *Es sei V ein Vektorraum der Dimension $n \in \mathbb{N}$ über einem Körper $\mathbb{K}$. Gegeben sei ein Endomorphismus $\varphi \in \text{End}_{\mathbb{K}}(V)$, dessen charakteristisches Polynom $p_\varphi \in \mathbb{K}[X]$ über $\mathbb{K}$ in Linearfaktoren zerfällt.*[30]

Dann gibt es eine geordnete Basis B von V, sodass $\text{M}_{\mathsf{B},\mathsf{B}}(\varphi) \in \mathbb{K}^{n\times n}$ die folgende Blockstruktur hat:

$$A := \text{M}_{\mathsf{B},\mathsf{B}}(\varphi) = \begin{pmatrix} J_{n_1}(\lambda_1) & & & \\ & J_{n_2}(\lambda_2) & & \\ & & \ddots & \\ & & & J_{n_m}(\lambda_m) \end{pmatrix}$$

Diese Darstellung heißt Jordansche Normalform *und ist eindeutig bis auf Permutation der Blöcke. Wenn alle Blöcke Größe 1 haben, ist die Jordansche Normalform eine Diagonalmatrix.*

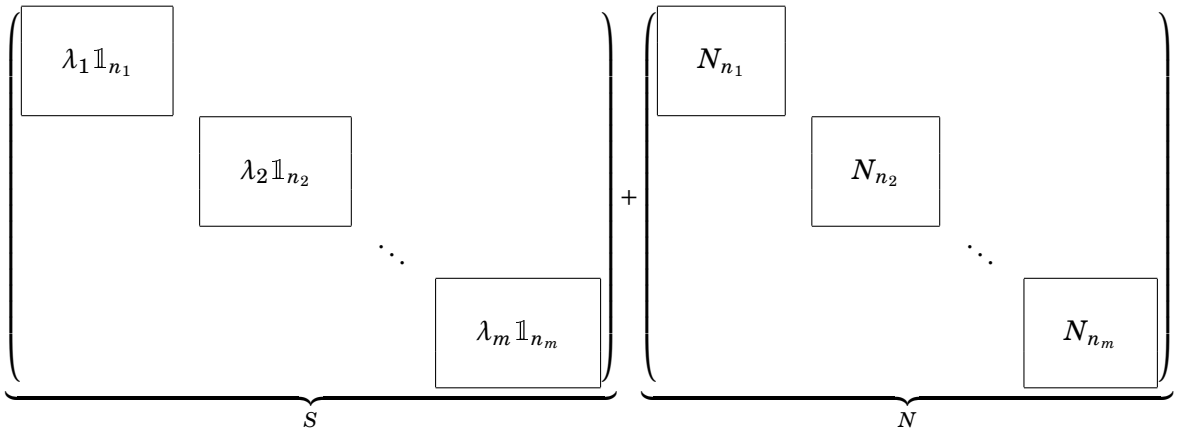
Eine Matrix in Jordanscher Normalform ist eine obere Dreiecksmatrix[31]*.*

Da sich jeder Jordan-Block zerlegen lässt in $J_{n_j}(\lambda_j) = \lambda_j \mathbb{1}_{n_j} + N_{n_j}$, erhalten wir auch eine

[30] Diese Voraussetzung ist immer erfüllt, wenn $\mathbb{K} = \mathbb{C}$.

[31] Insofern ist dies eine stärkere Aussage als Satz 5.7.1

Zerlegung der Matrix A als Summe einer Diagonalmatrix S und einer nilpotenten Matrix N:



Diese beiden Summanden kommutieren ($NS = SN$), weil dies in jedem Block gilt.

Insofern lässt sich also φ schreiben als Summe eines diagonalisierbaren und eines nilpotenten Endomorphismus. Dies ermöglicht eine effiziente Berechnung von hohen Potenzen der Matrix (siehe Bemerkung 5.8.7).

Die Jordansche Normalform ist sehr hilfreich, zum Einen um hohe Potenzen von Matrizen oder Endomorphismen zu berechnen; zum anderen, um zu entscheiden, ob zwei Matrizen ähnlich sind. Da die Jordansche Normalform eindeutig bis auf Permutation der Blöcke ist, folgt beispielsweise, dass die beiden folgenden Matrizen *nicht* ähnlich sind:

$$A_1 := \begin{pmatrix} 42 & 1 & 0 & 0 \\ 0 & 42 & 1 & 0 \\ 0 & 0 & 42 & 0 \\ 0 & 0 & 0 & 42 \end{pmatrix}; \quad A_2 := \begin{pmatrix} 42 & 1 & 0 & 0 \\ 0 & 42 & 0 & 0 \\ 0 & 0 & 42 & 1 \\ 0 & 0 & 0 & 42 \end{pmatrix}.$$

Beide Matrizen sind in Jordanscher Normalform, aber die Normalformen lassen sich nicht durch Vertauschen der Blöcke ineinander überführen, da A_1 einen Block der Größe 3 und einen der Größe 1 hat, während A_2 zwei Blöcke der Größe 2 besitzt. Dies ist beeindruckend, da alle Ähnlichkeitsinvarianten, die wir bis jetzt eingeführt hatten, für A_1 und A_2 gleich sind: Beide Matrizen haben die gleiche Determinante, die gleiche Spur, den gleichen Rang, das gleiche charakteristische Polynom. Beide haben nur den Eigenwert $\lambda = 42$ und bei beiden Matrizen hat $\lambda = 42$ die algebraische Vielfachheit 4 und die geometrische Vielfachheit 2. Trotzdem sind die beiden Matrizen nicht ähnlich.

Der Beweis von Satz 5.8.11 ist sehr aufwändig und wird hier in der Linearen Algebra I nicht geführt werden. Dazu müssten wir noch viel mehr Konzepte und Definitionen einführen. Wir verweisen hierfür auf die Veranstaltung „Lineare Algebra II (für Mathematik)“.

Zusammenfassung von Abschnitt 5.8

(1) Eine Matrix heißt nilpotent, wenn eine endliche Potenz der Matrix die Nullmatrix ist.

(2) Nilpotente Matrizen sind niemals diagonalierbar – außer der Nullmatrix selbst.

(3) Wenn eine Matrix trigonalisierbar ist, dann besitzt sie eine Jordansche Normalform, die eindeutig ist bis auf Permutation der Jordan-Blöcke.

(4) Für diagonalierbare Matrizen ist die Jordan-Normalform gleich der Diagonalform.

(5) Jede trigonalisierbare Matrix lässt sich mit Hilfe der Jordan-Normalform als die Summe einer diagonalierbaren und einer nilpotenten Matrix schreiben, sodass die beiden Summanden kommutieren. Dadurch ist es möglich, geschlossene Formeln für große Potenzen der Matrix zu erhalten, selbst wenn die Matrix selbst nicht diagonalierbar ist.

A. Anhang

A.1. Das griechische Alphabet

α			A	Alpha
β			B	Beta
γ			Γ	Gamma
δ			Δ	Delta
ε	oder	ϵ	E	Epsilon
ζ			Z	Zeta
η			H	Eta
ϑ	oder	θ	Θ	Theta
ι			I	Iota
κ			K	Kappa
λ			Λ	Lambda
μ			M	My
ν			N	Ny
ξ			Ξ	Xi
o			O	Omikron
π			Π	Pi
ρ			P	Rho
σ			Σ	Sigma
τ			T	Tau
υ			Y	Ypsilon
φ	oder	ϕ	Φ	Phi
χ			X	Chi
ψ			Ψ	Psi
ω			Ω	Omega

A.2. Zusatzmaterial über endliche Körper

Im Skript und in der Vorlesung wurde angekündigt, dass „später“ bewiesen wird, dass die Anzahl der Elemente in einem endlichen Körper $\mathbb{K}$ immer eine Primzahlpotenz ist. Dann sind wir leider doch nicht dazu gekommen. Deshalb holen wir den Beweis an dieser Stelle nach:

Der erste Schritt ist das folgende Lemma aus der elementaren Gruppentheorie, das eine Klassifikation aller Untergruppen der Gruppe $(\mathbb{Z},+)$ liefert:

Lemma A.2.1.
Für jede Untergruppe $U \subseteq (\mathbb{Z},+)$ gilt: Es gibt genau ein $m \in \mathbb{N}_0$ mit

$$U = m\mathbb{Z} = \{xk \mid k \in \mathbb{Z}\}.$$

Beweis. Die Eindeutigkeit von so einem m ist leicht zu sehen, wir beschränken uns in diesem Beweis auf die Existenz: Dazu betrachten wir die Menge $U \cap \mathbb{N}$.

1.Fall: $U \cap \mathbb{N} = \emptyset$:
Wir setzen $m := 0$ und zeigen, dass $U = 0\mathbb{Z} = \{0\}$ gilt. Die Mengeninklusion „$\supseteq$“ gilt einfach, weil U eine Untergruppe ist und somit die Zahl 0 enthält. Für die Richtung „$\subseteq$“ sei $u \in U$. Wir wollen zeigen, dass $u = 0$. Falls $u > 0$, dann gilt: $u \in U \cap \mathbb{N} = \emptyset$, das kann nicht sein. Und falls $u < 0$, dann ist $-u \in U$ (weil U eine Untergruppe ist) und somit ist $-u \in U \cap \mathbb{N} = \emptyset$, was auch ein Widerspruch ist. Also muss $u = 0$ sein.

2.Fall: $U \cap \mathbb{N} \neq \emptyset$:
In diesem Fall ist $U \cap \mathbb{N}$ eine nichtleere Teilmenge der natürlichen Zahlen und hat somit ein kleinstes Element, wir können also setzen:

$$m := \min(U \cap \mathbb{N}) \in \mathbb{N}.$$

Die Behauptung ist nun, dass die gegebene Untergruppe U genau die Menge $m\mathbb{Z}$ ist, d.h. wir behaupten:

$$m\mathbb{Z} = U.$$

Es gibt zwei Inklusionen zu zeigen:

„$\subseteq$“:
Es sei $x \in m\mathbb{Z}$, d.h. $x = mk$ mit $k \in \mathbb{Z}$. Falls $k = 0$, dann ist $x = 0$ und somit ist x in U, denn U enthält als Untergruppe von $(\mathbb{Z},+)$ natürlich auch das Neutralelement, also die Zahl 0. Falls $k \in \mathbb{N}$, dann ist $x = mk = \underbrace{m + \cdots + m}_{k \text{ Summanden}} \in U$, weil $m \in U$ und U als Untergruppe abgeschlossen ist unter der Gruppenoperation $+$. Falls $k \in -\mathbb{N}$, dann ist $x = -m(-k)$ und somit in U, weil U als Untergruppe abgeschlossen unter Inversenbildung ist.

„$\supseteq$“:
Es sei $u \in U$. Dann gibt es nach Lemma 3.4.1 eine Darstellung von u als

$$u = mk + r$$

mit $k \in \mathbb{Z}$ und $r \in \{0,\ldots,m-1\}$. Wie oben sieht man ein, dass $mk \in U$ gilt. Wenn wir diese Gleichung nach r auflösen, erhalten wir:

$$r = u - mk$$

und somit gilt auch $r \in U$, weil sowohl u als auch mk in U sind und U eine Untergruppe von $(\mathbb{Z},+)$ ist.

Falls nun $r \neq 0$, dann ist r eine natürliche Zahl und somit gilt $r \in U \cap \mathbb{N}$. Da m definiert ist als das Minimum der Menge $U \cap \mathbb{N}$, gilt somit $m \leq r$. Dies ist ein Widerspruch zu der Wahl von k und r im obigen Lemma.

Es muss also $r = 0$ sein und damit ist $u = mk \in m\mathbb{Z}$. Das beendet den Beweis. □

Als nächstes untersuchen wir den sog. *Primring*, d.h. den kleinsten Unterring eines Ringes. Wir werden sehen, dass so ein Primring immer existiert, dass man ihn schreiben kann als das Bild von einem Ring-Homomorphismus, und dass er immer isomorph ist zu $\mathbb{Z}/m\mathbb{Z}$ für ein $m \in \mathbb{N}_0$, was dann die *Charakteristik* genannt wird.

Lemma A.2.2.
Es sei R ein Ring. Dann gibt es genau einen Ring-Homomorphismus

$$\varphi_R : \mathbb{Z} \to R.$$

Der Kern von φ ist eine additive Untergruppe von $(\mathbb{Z}, +)$ und somit – nach Lemma A.2.1 – von der Form $m\mathbb{Z}$.

Das Bild Bild(φ_R) *ist der kleinste Unterring von R, genannt der* Primring *von R.*

Die Zahl $m \in \mathbb{N}_0$ heißt die Charakteristik *des Ringes.*

Weiter gilt: Die Abbildung

$$\widetilde{\varphi}_R : \mathbb{Z}/m\mathbb{Z} \to \text{Bild}(\varphi_R), \quad [k]_m \mapsto \varphi_R(k)$$

ist ein Ring-Isomorphismus. Jeder Primring ist somit isomorph zu einem Ring der Form $\mathbb{Z}/m\mathbb{Z}$.

Beweis. Wir definieren

$$\begin{aligned} \varphi_R(0) &= 0_R; \\ \varphi_R(k) &= \underbrace{1_R + \cdots + 1_R}_{k \text{ Summanden}} \text{ für } k > 0 \\ \text{und } \varphi_R(k) &= -\varphi_R(-k) \text{ für } k < 0. \end{aligned}$$

Man überprüft schnell, dass dies ein Ring-Homomorphismus ist und dass es keinen anderen Ring-Homomorphismus von $\mathbb{Z}$ nach R geben kann.

Das Bild eines Ring-Homomorphismus ist ein Unterring. Jedes Element im Bild ist eine Summe oder eine Differenz vom Eins-Element des Ringes. Somit enthält jeder Unterring von R den Unterring Bild(φ_R), und folglich ist Bild(φ_R) der kleinste Unterring von R. Der Kern $\ker(\varphi_R)$ ist bezüglich der Addition eine Untergruppe von $\mathbb{Z}$, also von der Form $m\mathbb{Z}$ mit einem eindeutigen $m \in \mathbb{N}_0$.

Die Abbildung

$$\widetilde{\varphi}_R : \mathbb{Z}/m\mathbb{Z} \to \text{Bild}(\varphi_R), \quad [k]_m \mapsto \varphi_R(k)$$

ist wohldefiniert, denn falls $[k]_m = [l]_m$ gilt, dann folgt daraus, dass $[k-l]_m = 0$, d.h. $k - l \in m\mathbb{Z} = \ker(\varphi_R)$, also ist

$$\varphi_R(k) = \varphi_R(l).$$

Die Eigenschaft, ein Ring-Homomorphismus zu sein, folgt nun direkt. Injektiv ist die Abbildung, weil wir gerade genau den Kern herausfaktorisiert haben und die Surjektivität folgt daraus, dass das Bild von φ_R gleich dem von $\widetilde{\varphi}_R$ ist. □

Beispiel A.2.3 (Charakteristiken von Ringen)**.** Hier einige interessante Informationen über die Charakteristik – diese sind aber für den eigentlichen Beweis nicht notwendig:

- Der Ring $\mathbb{Z}$ ist isomorph zu $\mathbb{Z}/0\mathbb{Z}$ und hat somit Charakteristik 0.
- Wenn R ein Unterring von S ist, dann haben R und S die gleiche Charakteristik. Damit folgt aus $\mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R} \subseteq \mathbb{C}$, dass alle diese Ringe auch Charakteristik 0 haben.
- Der Ring $\mathbb{Z}/m\mathbb{Z}$ hat immer Charakteristik m, insbesondere hat also $\mathbb{F}_2$ Charakteristik 2.
- Der Körper $\mathbb{F}_4$ enthält $\mathbb{F}_2$ als Unterring, also hat $\mathbb{F}_4$ auch Charakteristik 2 (und nicht etwa 4, wie man denken könnte!).
- Der Matrizenring $\mathbb{K}^{n\times n}$ über einem Körper $\mathbb{K}$ hat immer dieselbe Charakteristik wie der Körper $\mathbb{K}$.
- Der Polynomring $\mathbb{K}[X]$ über einem Körper $\mathbb{K}$ enthält $\mathbb{K}$ als Unterring und hat somit dieselbe Charakteristik wie $\mathbb{K}$. Beispielsweise hat $\mathbb{F}_2[X]$ Charakteristik 2. Es gibt also auch unendliche Ringe mit Charakteristik 2.
- Das kartesische Produkt $(\mathbb{Z}/4\mathbb{Z}) \times (\mathbb{Z}/6\mathbb{Z})$ mit komponentenweiser Addition und Multiplikation ist ein Beispiel für einen Ring mit Charakteristik 12 (einfache Übungsaufgabe).
- Ein Ring hat genau dann Charakteristik 1, wenn er isomorph zum Nullring $\{0\}$ ist.
- Ein Ring, der nicht isomorph zum Nullring ist, hat genau dann Charakteristik 2, wenn $1+1=0$ gilt.

Lemma A.2.4.
Es sei $\mathbb{L}$ ein Körper. Dann ist die Charakteristik entweder 0 oder eine Primzahl.

Beweis. Es sei $m \in \mathbb{N}_0$ die Charakteristik von $\mathbb{L}$ und es sei R_0 der Primring des Körpers $\mathbb{L}$. Da der Körper $\mathbb{L}$ nullteilerfrei ist und $1 \neq 0$ ist, gilt dies auch für jeden Unterring von $\mathbb{L}$. Also gilt: Im Ring R_0 ist das Einselement ungleich dem Nullelement und R_0 ist nullteilerfrei.

Wir haben aber gesehen, dass der Primring R_0 isomorph ist zu $\mathbb{Z}/m\mathbb{Z}$.

Falls $m = 1$, dann gilt $|R_0| = |\mathbb{Z}/m\mathbb{Z}| = 1$, also muss $1 = 0$ sein. Also kann m nicht 1 sein.

Falls $m \in \mathbb{N}$ mit einer Zerlegung $m = r \cdot s$ mit $r, s \in \mathbb{N}$ und $r, s > 1$, dann ist der Ring $R_0 \cong \mathbb{Z}/m\mathbb{Z}$ nicht nullteilerfrei, da in $\mathbb{Z}/m\mathbb{Z}$ gilt $[r]_m \cdot [s]_m = [c]_m = [0]_m$.

Folglich bleiben nur die Fälle, dass m entweder 0 oder eine Primzahl ist. □

Nun können wir dies zusammensetzen zu dem folgenden Satz:

Satz A.2.5.
Es sei $\mathbb{L}$ ein endlicher Körper. Dann gibt es eine Primzahl p und ein $k \in \mathbb{N}$ mit

$$|\mathbb{L}| = p^k.$$

Beweis. Es sei R_0 der Primring von $\mathbb{L}$ und es sei $m \in \mathbb{N}_0$ die Charakteristik von $\mathbb{L}$. Nach Lemma A.2.4 ist sie entweder 0 oder eine Primzahl. Falls $m = 0$, dann ist der Primring R isomorph zu $\mathbb{Z}/0\mathbb{Z}$. Dies ist aber isomorph zum Ring der ganzen Zahlen $\mathbb{Z}$. Also enthält der Körper $\mathbb{L}$ einen Unterring, der isomorph ist zu $\mathbb{Z}$. Dies ist aber nicht möglich, weil $\mathbb{Z}$ unendlich viele Elemente hat, aber $\mathbb{L}$ als endlich angenommen wurde. Also ist $p := m$ eine Primzahl.

Der Primring R_0 ist somit isomorph zu $\mathbb{Z}/p\mathbb{Z}$, also sogar ein Körper. Somit ist $\mathbb{L}$ eine Körpererweiterung von $\mathbb{Z}/p\mathbb{Z}$. Jede Körpererweiterung von $\mathbb{Z}/p\mathbb{Z}$ ist insbesondere ein Vektorraum über $\mathbb{Z}/p\mathbb{Z}$. Da $\mathbb{L}$ nur endlich viele Elemente besitzt, gibt es also ein endliches Erzeugendensystem und $\mathbb{L}$ ist somit endlich dimensional über $\mathbb{Z}/p\mathbb{Z}$. Also können wir eine (endliche!) Basis B auswählen. Wir setzen $k := \dim_{\mathbb{Z}/p\mathbb{Z}}(\mathbb{L}) = |B|$. Mit Hilfe dieser Basis erhalten wir einen Isomorphismus $\Psi : \mathbb{L} \to (\mathbb{Z}/p\mathbb{Z})^k$. Dies ist ein Isomorphismus als $\mathbb{Z}/p\mathbb{Z}$-Vektorraum. Insbesondere ist Ψ eine Bijektion, was heißt, dass $\mathbb{L}$ genau so viele Elemente hat wie der Raum $(\mathbb{Z}/p\mathbb{Z})^k$ der Spaltenvektoren mit Einträgen in $\mathbb{Z}/p\mathbb{Z}$. Also gilt:

$$|\mathbb{L}| = \left|(\mathbb{Z}/p\mathbb{Z})^k\right| = p^k. \qquad \square$$

Somit haben wir gezeigt, dass die Anzahl der Elemente in einem endlichen Körper immer eine Primzahlpotenz ist und das Kernargument war dabei, dass jeder endlich erzeugte Vektorraum eine Basis besitzt. Nun wissen wir insbesondere, dass es keinen Körper mit 6 oder mit 10 Elementen geben kann.

Es gibt über endliche Körper allerdings noch mehr zu sagen: Man kann zeigen, dass es zu jeder Primzahlpotenz $q = p^k$ immer auch einen Körper $\mathbb{F}_q$ gibt, der genau q Elemente hat[1]. Und mehr noch: Dieser Körper ist bis auf Isomorphie eindeutig. Es sind also zwei endliche Körper genau dann isomorph, wenn sie gleich viele Elemente haben. Eine solche Aussage ist für andere algebraische Strukturen falsch. Ringe oder Gruppen müssen noch lange nicht isomorph sein, nur weil sie gleich viele Elemente haben ($\mathbb{Z}/4\mathbb{Z}$ und $\mathbb{F}_4$ sind beides Ringe mit 4 Elementen).

Der Beweis dieses Klassifikationssatzes von endlichen Körpern verwendet meistens Galois-Theorie und geht weit über die lineare Algebra hinaus.

[1] Insbesondere gibt es also zu jeder Zweierpotenz einen endlichen Körper der Charakteristik 2. Diese Körper spielen in der Kryptographie eine wichtige Rolle.

Stichwortverzeichnis

Abbildung, 13
abbrechende Folge, 115
abelsche Gruppe, 66
additives neutrales Element, 75
Adjunkte, 186
affiner Unterraum, 43, 106
Affinkombination, 44
Ähnlichkeit von Matrizen, 145, 206
Algebra der formalen Polynome, 190
Algebra über einem Körper, 188
algebraisch abgeschlossen, 99, 218
algebraische Geometrie, 201
algebraische Vielfachheit, 219
allgemeine lineare Gruppe, 127
Allquantor, 9
alternierende Gruppe, 172
alternierende multilineare Abbildung, 174
Äquivalenz, 6
Äquivalenz von Matrizen, 145
Äquivalenzklasse, 84
Äquivalenzrelation, 84
assoziative Algebra mit Eins, 188
assoziative Verknüpfung, 62
Aussage, 5
Aussageform, 9
Auswahlaxiom, 118, 225
Auswertungshomomorphismus, 196
autologisches Wort, 243
Automorphismus von Gruppen, 71
Automorphismus von Ringen, 78
Automorphismus von Vektorräumen, 108

Basis, 113
Basis eines Untervektorraums, 36
Basisauswahlsatz, 38, 117
Basiswechselmatrix, 140
Betrag einer komplexen Zahl, 95
Bidualraum, 162
bijektiv, 16
Bild einer Funktion, 14
Bild einer Matrix, 47
bilineare Abbildung, 174
Binomialkoeffizienten, 229
binomischer Lehrsatz, 229
binäre Verknüpfung, 61

Cayley-Hamilton, 225
Cayley-Tafel, 61
Charakteristik, 239
charakteristisches Polynom einer Matrix, 210
charakteristisches Polynom eines Endomorphismus, 211

Darstellungsmatrix bezüglich Standardbasis, 124
Darstellungsmatrix einer linearen Abbildung bezüglich geordneter Basen, 137
De Morgansche Regeln, 6, 10, 12
Definitionsbereich, 13
Determinante einer (2x2)-Matrix, 130
Determinante einer Matrix, 177
Determinante eines Endomorphismus, 207
diagonalisierbar, 209
Diagonalmatrix, 180
Diedergruppe, 69
Dimension, 119
Dimension eines affinen Unterraums, 44
Dimension eines Untervektorraums, 41
Dimensionsformel, 60
direkte Summe, 149
direktes Produkt, 146
Disjunktion, 5
Distributionen, 161
Divisionsring, 79
duale Basis, 164
Dualraum, 160

Ebene durch drei Punkte, 44
echte Teilmenge, 10
Eigenraum, 210
Eigenvektor einer Matrix, 208
Eigenvektor eines Endomorphismus, 208
Eigenwert einer Matrix, 208
Eigenwert eines Endomorphismus, 208
einfach nilpotent, 231
Einheitengruppe, 66
Einheitskreis, 95
Einheitsmatrix, 28
Einschränkung, 14
Einselement, 75
elementare Zeilenumformungen, 51
endlich dimensional, 119
endlich dimensionale Algebra, 188
Endomorphismus von Gruppen, 71
Endomorphismus von Ringen, 78
Endomorphismus von Vektorräumen, 108
Entwicklung nach einer Spalte, 186
Entwicklung nach einer Zeile, 186
erweiterte Zeilenstufenform, 50
Erzeugendensystem, 32
erzeugter Untervektorraum, 31
Eulersche Formel, 98
Existenzquantor, 9
exklusives Oder, 12

Faktormenge, 85
Faktorraum, 152
falsch, 5
Faltung, 191
Fibonacci-Folge, 219
formale Variable, 190
Funktion, 13
Fußpunkt, 43, 106

ganze Zahlen, 8
Gauß-Algorithmus, 50
    Teil I, 53
    Teil II, 55
    Teil IIb, 56
    Teil III, 56
Gaußsches Eliminationsverfahren, 50
genau dann, wenn, 6
geometrische Vielfachheit, 210
geordnete Basis, 133
geordnete Standardbasis, 133
geordnetes Paar, 12
Gerade durch zwei Punkte, 44
gerade Permutation, 172
Gleichheit von Mengen, 9
goldener Schnitt, 220
Grad eines Polynoms, 190
Graph, 13
Gruppe, 65
Gruppenautomorphismus, 71
Gruppenendomorphismus, 71
Gruppenhomomorphismus, 70
Gruppenisomorphismus, 71

Halbgruppe, 62
homogenes LGS, 23, 29
Homomorphismus von Gruppen, 70
Homomorphismus von Körpern, 80
Homomorphismus von Ringen, 78
Homomorphismus von Vektorräumen, 107

Identität, 15
Identitätsabbildung, 15
imaginäre Achse, 95
imaginäre Zahlen, 95
Imaginärteil, 95
Implikation, 5
inhomogenes LGS, 23, 29
injektiv, 15
Inverses, 63
invertierbar, 63, 75
invertierbare Matrix, 127
isomorphe Gruppen, 71
isomorphe Vektorräume, 108
Isomorphismus von Gruppen, 71
Isomorphismus von Ringen, 78
Isomorphismus von Vektorräumen, 108

Jordan-Block, 230
Jordansche Normalform, 233

kartesisches Produkt, 12
Kategorientheorie, 165
Kern einer linearen Abbildung, 108
Kern einer Matrix, 29
Kern eines Gruppenhomomorphismus, 71
Kern eines Ringhomomorphismus, 78
Koeffizienten, 22
Koeffizientenmatrix, 25

Koeinschränkung, 14
kommutative Algebra, 188
kommutative Halbgruppe, 62
kommutativer Ring, 75
kommutatives Diagramm, 15
kompatibel, 27
Komplement, 11
komplementärer Untervektorraum, 149
komplexe Einheitswurzeln, 98
komplexe Exponentialfunktion, 97
komplexe Zahlen, 93
komplexer Vektorraum, 101
Komposition, 14
kongruent modulo, 83
konjugiert komplexe Zahl, 95
Konjunktion, 5
Kontrapositionsprinzip, 6
Korestriktion, 14
Kreuzprodukt, 174
Körper, 79
Körpererweiterung, 81
Körperhomomorphismus, 80

leere Menge, 8
Leibniz-Formel, 178
linear abhängig, 34, 113
linear unabhängig, 34, 113
lineare Abbildung, 107
Lineare Gleichung, 22
lineare Hülle, 31, 112
lineare Rekursionsgleichung, 221
linearer Aufspann, 31
linearer Spann, 31
linearer Unterraum, 30, 104
lineares Gleichungssystem (LGS) mit reellen Koeffizienten, 23, 28
Linearform, 160
Linearkombination, 30, 112

Matrix, 102
Matrix mit reellen Einträgen, 24
Matrixprodukt, 27
Menge, 7
Mengensystem, 12
Modul über einem Ring, 201
Monoid, 63
multilineare Abbildung, 174
multiplikatives neutrales Element, 75
multiplizierbar, 27

natürliche Zahlen, 8
Negation, 5
Neutralelement, 63
neutrales Element, 63
nilpotente Matrix, 228
nilpotenter Endomorphismus, 228
nilpotenter Jordan-Block, 230
nilpotentes Ringelement, 228
Nullelement, 75
Nullmatrix, 26
Nullpolynom, 190
nullteilerfrei, 29, 79
Nullvektor, 26, 104
Nullzeile, 50

obere Dreiecksmatrix, 180
Operator, 110

Paar, 12
Permutation, 67, 169
Permutationsgruppe, 67
Permutationsmatrix, 68, 185
Pivot-Eintrag, 50
Pivot-Spalte, 50
Polynom, 190
Polynomalgebra, 190
Polynomfunktion, 189
Polynomraum, 190
Polynomring, 190
Potenzen, 65
Potenzmenge, 12
Primring, 239
Primzahl, 89

quadratische Matrix, 24
Quantor, 9
Quaternionen, 79
Quotientenabbildung, 85
Quotientenmenge, 85
Quotientenvektorraum, 152

Rang einer linearen Abbildung, 121
Rang einer Matrix, 47
Rangsatz, 59
rationaler Vektorraum, 101
Realteil, 95

reelle Achse, 95
reelle Zahlen, 21
reeller Vektorraum, 101
reflexive Relation, 84
Regel von Sarrus, 178
Relation, 84
Rest modulo m, 81
Restriktion, 14
Ring, 75
Ring mit Eins, 76
Ringautomorphismus, 78
Ringendomorphismus, 78
Ringhomomorphismus, 78
Ringisomorphismus, 78

Sarrus-Regel, 178
Schiefkörper, 79
Schnitt, 11, 12
Signum einer Permutation, 170
Skalar, 26, 101
Skalares Vielfaches, 26
Skalierung einer Matrix, 26
Smith-Normalform, 145
Spaltenrang, 59
Spaltenvektor, 22
spezielle lineare Gruppe, 184
Spur eines Endomorphismus, 207
Standardbasis, 38, 114
Standardbasisvektoren, 38, 114
Standardmatrix, 114
Standardskalarprodukt, 174
Streichungsmatrix, 185
Summe von Matrizen, 25
Summe von Untervektorräumen, 148
Summennotation, 22
surjektiv, 15
Symmetriegruppe des gleichseitigen Dreiecks, 69
symmetrische Differenz, 12
symmetrische Gruppe, 67, 169
symmetrische Matrix, 105
symmetrische Relation, 84

teilbar, 82
Teilkörper, 80
Teilmenge, 10
Tertium non datur, 6, 11
transitive Relation, 84
Transponierte, 24
Transposition, 169
triviale Linearkombination, 34

Umkehrfunktion, 16
unendlich dimensional, 119
ungerade Permutation, 172
untere Dreiecksmatrix, 180
Untergruppe, 70
Unterkörper, 80
Unterring, 77
Untervektorraum, 30, 104
Urbild unter einer Funktion, 14

Vektor, 101
Vektorraum über einem Körper, 101
Vektorraumautomorphismus, 108
Vektorraumendomorphismus, 108
Vektorraumhomomorphismus, 107
Vektorraumisomorphismus, 108
Vektorraumkomplement, 149
Venn-Diagramm, 11
Vereinigung, 11, 12
Verkettung, 14
Verknüpfungstafel, 61

wahr, 5

Zeilenrang, 59
Zeilenstufenform, 50
Zeilenvektor, 24
zerfallen in Linearfaktoren, 218
Zielbereich, 13
Zykelschreibweise, 67