universität
innsbruck
Institut für Mathematik

Lineare Algebra

Tim Netzer

∀¬6∃q⅄∀

A⌐9Eb⅄A

Inhaltsverzeichnis

1 Einleitung und Grundlagen **1**

1.1 Was ist lineare Algebra? 1
1.2 Notation 2
1.3 Logische Grundlagen 3
1.3.1 Aussagenlogik 3
1.3.2 Prädikatenlogik 3
1.3.3 Einige Beweismethoden 4
1.4 Mengen, Relationen und Abbildungen 7

2 Gauß-Algorithmus und Matrixrechnung **25**

2.1 Der Gauß-Algorithmus 25
2.2 Gruppen, Ringe und Körper 33
2.2.1 Grundbegriffe 33
2.2.2 Die komplexen Zahlen ★ 37
2.2.3 Endliche Körper 52
2.3 Matrixrechnung 56

3 Vektorräume und lineare Abbildungen **67**

3.1 Vektorräume 68
3.1.1 Grundbegriffe 68
3.1.2 Lineare Unabhängigkeit, Basen und Dimension 74
3.2 Lineare Abbildungen 84
3.2.1 Grundbegriffe 85
3.2.2 Die Dimensionsformel 92
3.2.3 Der Rang einer Matrix 96
3.2.4 Darstellungsmatrizen ★ 98
3.2.5 Der Dualraum ★ 104

4 Die Determinante **107**
4.1 Permutationen 107
4.2 Die Determinante und ihre Eigenschaften 113
4.3 Entwicklungsformeln und Cramersche Regel ★ 120

5 Eigenwerte und Eigenvektoren **125**
5.1 Eigenwerte und Eigenvektoren 125
5.2 Polynome und Nullstellen 130
5.3 Diagonalisierung 133
5.4 Trigonalisierung 139
5.5 Die Jordan'sche Normalform 142
5.6 Der Satz von Cayley-Hamilton und das Minimalpolynom 149

6 Skalarprodukte und Spektralsätze **155**
6.1 Skalarprodukte und Normen 155
6.2 Orthonormalbasen 160
6.3 Spektralsätze 166
6.4 Positiv semidefinite Matrizen 175

7 Bilinearformen **181**
7.1 Grundbegriffe 181
7.2 Symmetrische Bilinearformen 185
7.3 Quadratische Formen 192

8 Konvexität **197**
8.1 Grundbegriffe 197
8.2 Trennungssätze 203
8.3 Polyeder und Polytope 207

9 Verschiedenes **217**
9.1 Konstruktionen mit Vektorräumen 217
9.1.1 Summen und Produkte 217
9.1.2 Faktorräume 221
9.1.3 Tensorprodukte 224
9.2 Räume von unendlicher Dimension 228
9.2.1 Algebraische Sichtweise 228
9.2.2 Algebraisch-analytische Sichtweise 231
9.3 Kategorientheorie 233

Literaturverzeichnis **239**

Übungsaufgaben **241**

Kapitel 1

Einleitung und Grundlagen

1.1 Was ist lineare Algebra?

In vielen Bereichen der Mathematik beschäftigt man sich mit dem Lösen von Gleichungssystemen. Man will entweder Lösungen explizit finden (am besten sogar alle existierenden Lösungen), oder zumindest die Menge aller Lösungen geometrisch gut verstehen.

Die lineare Algebra befasst sich mit den denkbar einfachsten Gleichungssystemen, nämlich mit *linearen Gleichungssystemen* (über geeigneten Zahlbereichen). Hier hat man im Lauf der Zeit auf fast alle Fragen befriedigende Antworten gefunden. Da lineare Gleichungssysteme sowohl in der Mathematik als auch in anderen Wissenschaften sehr häufig auftreten, zählt die lineare Algebra zu den wichtigsten theoretischen Grundlagen in vielen Gebieten von Wissenschaft und Technik.

Die Lösungsmenge eines linearen Gleichungssystems hat eine einfache geometrische Struktur, die eines (affinen) Vektorraums. Darum können solche Lösungsmengen durch endlich viele Daten komplett angegeben werden, selbst im Fall unendlich vieler Lösungen. Zusätzlich existiert mit dem (vielleicht schon aus der Schule bekannten) Algorithmus von Gauß eine effiziente algorithmische Methode, um lineare Gleichungssysteme zu lösen. Alle diese Tatsachen lernen Sie im Verlauf dieser Vorlesung kennen. Damit sind dann viele praktische Fragen zu linearen Gleichungssystemen eigentlich bereits beantwortet. Die Mathematik fragt aber immer auch nach einem tieferen Verständnis und nach Verallgemeinerungen und Erweiterungen der bekannten Ergebnisse. Um lineare Gleichungssysteme in allen Aspekten gründlich zu verstehen, entwickeln wir die Theorie von

Vektorräumen und linearen Abbildungen. Das ist aber nicht nur Selbstzweck, denn aus dem tieferen Verständnis ergeben sich neue Sichtweisen und Erkenntnisse auch über die ursprünglichen konkreten Fragen. Auch aus Sicht der Geometrie sind Vektorraumtheorie und ihre Erweiterungen unerlässlich.

Die Vorlesung *Lineare Algebra* richtet sich an Studierende unterschiedlicher Fachrichtungen. Sie geht deshalb nicht an jeder Stelle so weit in die Tiefe, wie ein rein mathematischer Ansatz vielleicht nahelegen würde. Insbesondere äußert sich das dadurch, dass mit ★ gekennzeichnete Abschnitte erst im Vertiefungsteil der Vorlesung behandelt werden. Mit Vertiefung umfasst die lineare Algebra 1 den Stoff bis etwa zur Mitte von Kapitel 5. Der Rest des Skripts ist dann Stoff der Vorlesung *Lineare Algebra 2*.

Zur linearen Algebra gibt es unzählige Bücher und Skripten. Grundsätzlich stehen alle wichtigen Resultate in fast jedem dieser Texte. Wer zusätzliche Literatur konsultieren möchte, kann sich also ganz nach den eigenen Vorlieben richten. Als erste Anregungen empfehlen wir die Bücher von Bosch [1], Fischer [2] und Lang [3].

1.2 Notation

Im Skript und der Vorlesung werden durchgehend die folgende Symbole und Notationen verwendet:

$\wedge$	und
$\vee$	oder
$\Rightarrow$	wenn, dann
$\neg$	nicht
$\forall$	für alle
$\exists$	es existiert
$\exists!$	es existiert genau ein
$\mathbb{N}$	natürliche Zahlen (mit Null)
$\mathbb{Z}$	ganze Zahlen
$\mathbb{Q}$	rationale Zahlen
$\mathbb{R}$	reelle Zahlen
$\in$	ist Element von
$\notin$	ist kein Element von.

1.3 Logische Grundlagen

1.3.1 Aussagenlogik

In der Aussagenlogik weisen wir jeder Aussage einen sogenannten **Wahrheitswert** zu, der entweder $\mathbf{W}$ für *wahr* oder $\mathbf{F}$ für *falsch* lautet. Wenn wir dann Aussagen mit sogenannten **Junktoren** verknüpfen erhalten wir neue Aussage, deren Wahrheitswert sich aus den Wahrheitswerten der einzelnen Teilaussagen ergibt. Junktoren sind Verknüpfungen wie *und, oder, nicht....*
Sehr übersichtlich stellt man so etwas in einer **Wahrheitstafel** dar. Dabei stehen in der ersten Zeile zunächst die einzelnen Aussagen (oft $\mathbf{A}, \mathbf{B}$...) und danach die daraus konstruierten Aussagen. In den darunterliegenden Zeilen stehen alle möglichen Wahrheitswerte der ursprünglichen Aussagen und die sich ergebenden Werte für die konstruierten Aussagen:

A	B	$\neg A$ (nicht)	$A \wedge B$ (und)	$A \vee B$ (oder)	$A \Rightarrow B$ (Implikation)	$(\neg A) \vee B$
W	W	F	W	W	W	W
W	F	F	F	W	F	F
F	W	W	F	W	W	W
F	F	W	F	F	W	W

In einer solchen Tafel kann man auch ablesen, ob zwei Aussagen logisch äquivalent sind. Sie sind es genau dann, wenn für alle Wahrheitsbelegungen der ursprünglichen Aussagen jeweils beide dieselben Wahrheitswerte haben. So sieht man an der oberen Tabelle zum Beispiel, dass die Aussage “$A \Rightarrow B$” (*A impliziert B*) äquivalent ist zur Aussage “$(\neg A) \vee B$” (*(nicht A) oder B*).

1.3.2 Prädikatenlogik

Häufig hat man es mit logischen Aussagen zu tun, die nicht einfach wahr oder falsch sind, sondern die gewissen **Leerstellen/Variablen** enthalten. Eine solche Aussage kann nur dann als wahr oder falsch beurteilt werden, wenn etwas für die Leerstelle eingesetzt wird. Ein Beispiel ist die Aussage

x ist größer gleich Null (in Formeln: $x \geqslant 0$).

Die Wahrheit dieser Aussage kann nur entscheiden werden, wenn für x etwas eingesetzt wird. Die obere Aussage ist in den ganzen Zahlen etwa wahr wenn man für x die Zahl 1 einsetzt, und falsch wenn man die Zahl -1 einsetzt.

Durch Hinzufügen von **Quantoren** kann eine solche Aussage jedoch in eine Aussage umformuliert werden, die dann für sich selbst wahr oder falsch ist (natürlich abhängig davon, in welchem mathematischen Modell sie betrachtet wird). Quantoren sind für uns der **All-Quantor** $\forall$ (mit der Bedeutung *für alle*) und der **Existenz-Quantor** $\exists$ (mit der Bedeutung *es existiert*). So ergibt sich beispielsweise die Aussage

$$\forall x\colon x \geqslant 0$$

welche in den ganzen Zahlen falsch ist, und die Aussage

$$\exists x\colon x \geqslant 0$$

welche in den ganzen Zahlen wahr ist. Solche Aussagen können dann natürlich wie im ersten Abschnitt durch Junktoren weiter zu komplizierteren Aussagen zusammengesetzt werden.

1.3.3 Einige Beweismethoden

Um eine mathematische Aussage zu beweisen gibt es verschiedene Methoden (die jedoch für sich allein meistens nicht ausreichen, sondern zusätzliche mathematische Kreativität benötigen). Eine Auswahl stellen wir hier kurz vor.

(Gegen-)Beweis durch (Gegen-)Beispiel

Eine Existenzaussage kann man manchmal sehr leicht durch die Angabe eines Beispiels beweisen. So wird die Aussage

$$\exists x\colon x \geqslant 0$$

für die ganzen Zahlen dadurch bewiesen, dass man die Zahl 1 als Beispiel angibt. Genauso kann eine All-Aussage manchmal durch Angabe eines Gegenbeispiels widerlegt werden. Soll zum Beispiel die Aussage

$$\forall x\colon x \geqslant 0$$

für die ganzen Zahlen widerlegt werden, kann man einfach die Zahl -1 angeben. Aber Vorsicht: eine universelle Aussage wie $\forall x\colon x \geqslant 0$ ist jedoch durch die Angabe eines positiven Beispiels noch **nicht** bewiesen. Die Aussage besagt ja gerade etwas über *alle* Zahlen, nicht nur über eine. In einem endlichen Zahlbereich könnte man sie eventuell durch gezielte Betrachtung aller in Frage kommenden Belegungen

für x beweisen. Kommen unendlich viele Belegungen in Fragen, muss man sich etwas anderes einfallen lassen.
Ebenso lässt sich eine Existenzaussage im Allgemeinen nicht dadurch widerlegen, dass man ein Gegenbeispiel angibt. Man müsste dazu zeigen, dass *alle* in Frage kommenden Belegungen für x die Aussage nicht erfüllen.

Beweis durch Widerspruch

Wenn man eine Aussage beweisen möchte, kann man ihre Falschheit annehmen und daraus einen Widerspruch abzuleiten versuchen. Das nennt man *Beweis durch Widerspruch*.
Angenommen man möchte beweisen, dass es unendlich viele Primzahlen gibt. Man könnte das direkt dadurch tun, dass man eine Vorschrift angibt, anhand derer wirklich unendliche viele Primzahlen produziert werden. Für einen Widerspruchsbeweis würde man hingegen annehmen dass die Aussage falsch ist, dass also nur endlich viele Primzahlen existieren. Aus dieser Aussage leitet man dann mit logischen Schlüssen einen offensichtlichen Widerspruch her. Dadurch ist die ursprüngliche Aussage ebenso bewiesen.
Je nach Aussage ist ein Widerspruchsbeweis manchmal etwas leichter durchzuführen. Sehr häufig (nicht immer!) lässt sich ein Widerspruchsbeweis aber auch in einen direkten Beweis umformen und umgekehrt. Die meisten Mathematiker bevorzugen aus ästhetischen Gründen einen direkten Beweis gegenüber einem Widerspruchsbeweis.

Beweis durch Kontraposition

Möchte man die Aussage

$$A \Rightarrow B$$

beweisen (also "A impliziert B"), kann man stattdessen auch die Aussage

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

(also "(nicht B) impliziert (nicht A)") zeigen. Beide Aussagen sind nämlich logisch äquivalent, wie man durch die Angabe einer Wahrheitstafel leicht zeigen kann. Dieses Vorgehen nennt man *Beweis durch Kontraposition*.
Will man beispielsweise zeigen, dass jede durch 4 teilbare Zahl auch durch 2 teilbar ist, könnte man stattdessen zeigen, dass eine nicht durch 2 teilbare Zahl auch nicht durch 4 teilbar sein kann.

Vollständige Fallunterscheidung

Manchmal muss man eine allgemeine Aussage in mehrere Teile aufteilen, die jeweils verschiedene Beweise erfordern. Möchte man eine Aussage etwa für alle natürlichen Zahlen beweisen, kann man sie zunächst für die geraden und dann für die ungeraden Zahlen zeigen. Das nennt man **Fallunterscheidung**. Wichtig ist dabei, dass jede der möglichen Situationen in mindestens einem der Fälle abgedeckt ist.

Schubfachprinzip

Wenn man mindestens $n+1$ Objekte auf höchstens n Schubfächer aufteilt, müssen sich in mindestens einem der Fächer mindestens 2 Objekte befinden. Diese einfache Beobachtung nennt man **Schubfachprinzip**. Manchmal kann so ein sehr eleganter Beweis geführt werden, zum Beispiel dass unter 366 Personen mindestens zwei am gleichen Tag Geburtstag haben müssen.

Beweis durch vollständige Induktion

Wenn man eine Aussage für alle natürlichen Zahlen $0, 1, 2, 3, \ldots$ zeigen möchte, kann man das zum Beispiel durch **vollständige Induktion** tun.
Der erste Schritt dabei ist der **Induktionsanfang (IA)**. Dabei zeigt man die Gültigkeit der Aussage für die kleinstmögliche Zahl, etwa die 0 (manchmal soll die Aussage auch erst ab einer bestimmten Zahl gelten, dann startet man dort). Meistens ist das offensichtlich wahr, oder durch eine sehr einfache Rechnung einsichtig.
Dann nimmt man an dass die Aussage bereits für eine allgemeine natürliche Zahl n gilt (oder für alle Zahlen bis n), das nennt man die **Induktionsvoraussetzung (IV)**. Ausgehend davon beweist man die Aussage dann für die nächste Zahl $n+1$. Das ist der sogenannte **Induktionsschritt (IS)**.
Hat man das getan, ist die Aussage bewiesen. Das Prinzip kann durch den Dominoeffekt veranschaulicht werden. Wenn der erste Stein umgestoßen wird, und sichergestellt ist dass jeder Stein den nächsten umstößt, dann fällt irgendwann jeder Stein.
Als Beispiel wollen wir die Gleichung

$$0 + 1 + 2 + \cdots + n = \frac{n(n+1)}{2}$$

für jede natürliche Zahl n beweisen. Im Induktionsanfang rechnen wir die Gültigkeit für $n = 0$ nach. Links steht dann 0 und rechts offensichtlich auch. Die

Induktionsvoraussetzung ist nun, dass die Aussage für ein festes (aber allgemeines) n bereits stimmt. Im Induktionssschritt müssen wir die Aussage jetzt für $n+1$ zeigen, also die folgende Aussage beweisen:

$$0 + 1 + \cdots + n + (n+1) = \frac{(n+1)(n+2)}{2}.$$

Dazu gehen wir folgendermaßen vor:

$$\begin{aligned} 0 + 1 + \cdots + n + (n+1) &\overset{\textbf{IV}}{=} \frac{n(n+1)}{2} + (n+1) \\ &= \frac{n(n+1)}{2} + \frac{2(n+1)}{2} \\ &= \frac{n(n+1) + 2(n+1)}{2} \\ &= \frac{(n+1)(n+2)}{2}. \end{aligned}$$

Für die erste Gleichheit haben wir die Induktionsvoraussetzung benutzt, also dass die Summe aller Zahlen bis n gerade $\frac{n(n+1)}{2}$ ergibt. Damit ist die Aussage allgemein bewiesen.

1.4 Mengen, Relationen und Abbildungen

Der Begriff der **Menge** gehört zu den wichtigsten Begriffen der Mathematik überhaupt. Lange Zeit wurden Mengen einfach definiert als

> *Zusammenfassung bestimmter wohlunterschiedener Objekte der Anschauung oder des Denkens zu einem Ganzen. Die einzelnen Objekte heißen Elemente der Menge.*

Bemerkung 1.4.1 (★). Diese Definition ist eher philosophischer Natur und mathematisch nicht exakt. So könnte man etwa die Menge betrachten, deren Element alle Mengen sind:

$$A := \{M \mid M \text{ Menge}\}.$$

Es handelt sich hier offenbar um eine Zusammenfassung von wohlunterschiedenen Objekten zu einem Ganzen. Somit müsste A also laut Definition eine Menge sein, und damit ein Element von sich selbst:

$$A \in A.$$

Wem es nicht seltsam genug vorkommt, dass Mengen Elemente ihrer selbst sein können, der betrachte nun folgende Menge:

$$B := \{M \mid M \text{ Menge}, M \notin M\}.$$

Auch hier handelt es sich um eine Zusammenfassung wohlunterschiedener Objekte, also um eine Menge. Es gibt nun zwei Möglichkeiten. Entweder gilt $B \in B$. Aufgrund der Definition von B muss dann aber $B \notin B$ gelten, ein Widerspruch. Also muss die zweite Möglichkeiten eintreten, nämlich $B \notin B$. Damit erfüllt B aber die Bedingung in der Definition von B und es folgt $B \in B$, wieder ein Widerspruch.
Der einzige Weg aus diesem Dilemma besteht darin, B nicht als Menge zuzulassen. Damit entpuppt sich die angegebene Mengendefinition als nicht zufriedenstellend. Es gibt nun in der Tat mathematisch saubere Möglichkeiten, mit diesem Problem umzugehen und solche Widersprüche zu vermeiden. Allerdings ist die lineare Algebra nicht der richtige Ort, das zu besprechen. Wir bleiben deshalb im folgenden bei der nicht-exakten Mengendefinition und behalten nur im Hinterkopf, dass es theoretisch hier etwas zu bedenken gibt. △

Mengen können auf verschiedene Weise angegeben werden. Eine endliche Menge kann man beispielsweise einfach dadurch angeben, dass man alle ihre Elemente aufzählt (gewöhnlich verwendet man geschwungenen Klammern für Mengendefinitionen):

$$\{a, b, c, d\}, \quad \{\text{Klaus, Bello, Patscherkofel}\}, \quad \{1, \pi\}.$$

Bei unendlichen Mengen geht das nicht. Man kann diese Mengen beispielsweise angeben, indem man ihre Elemente andeutet und auf den Verstand der Leser hofft:

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

Oft werden Mengen auch dadurch angegeben, dass man in der Mengenklammer nach einem senkrechten Strich Bedingungen auflistet, die die Zugehörigkeit von Elementen definiert. Dabei kann auch auf bereits bekannte Mengen zurückgegriffen werden. Beispiele dafür sind

$$\mathbb{Q} := \left\{\frac{a}{b} \mid a, b \in \mathbb{Z}, b \neq 0\right\}, \quad \mathbb{P} := \{a \mid a \in \mathbb{N}, a \text{ Primzahl}\}.$$

Man beachte, dass in einer Menge keine Reihenfolge der Elemente vorgegeben oder relevant ist. So gilt etwa

$$\{a, b, c\} = \{c, b, a\}.$$

Außerdem kann jedes Element in einer Menge nur einmal enthalten sein, selbst wenn man es (überflüssigerweise) mehrfach aufzählt. Es gilt also

$$\{a, a, b, c\} = \{a, b, c\}.$$

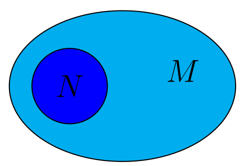
Vom Mengenbegriff ausgehend definiert man nun weitere wichtige Konzepte.

Definition 1.4.2. (i) Eine Menge N heißt **Teilmenge** einer Menge M, wenn jedes Element von N auch ein Element von M ist. Wir verwenden dafür die Notation

$$N \subseteq M.$$

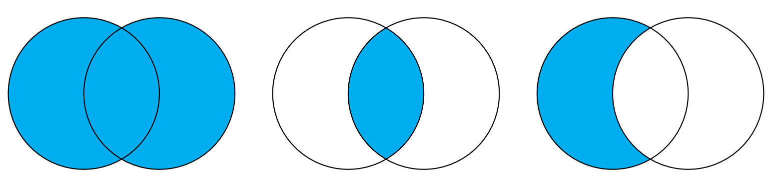
Als Formel geschrieben liest sich das so:

$$N \subseteq M \quad :\Leftrightarrow \quad \forall a\colon a \in N \Rightarrow a \in M.$$



(ii) Für zwei Mengen M, N bezeichnet $M \cup N$ die **Vereinigungs-**, $M \cap N$ die **Schnittmenge** sowie $M \setminus N$ die **Mengendifferenz**:

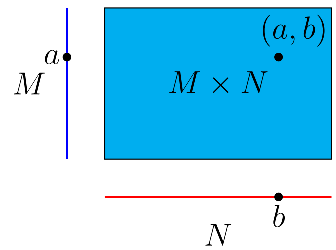
$$M \cup N := \{a \mid a \in M \vee a \in N\}$$
$$M \cap N := \{a \mid a \in M \wedge a \in N\}$$
$$M \setminus N := \{a \mid a \in M \wedge a \notin N\}.$$



$M \cup N$ $\qquad$ $M \cap N$ $\qquad$ $M \setminus N$

(iii) Für zwei Mengen M, N ist das **kartesische Produkt** von M und N die Menge aller geordneten Paare von Elementen aus M und N:

$$M \times N := \{(a, b) \mid a \in M, b \in N\}.$$



$\triangle$

Bemerkung/Beispiel 1.4.3. (i) Es gibt genau eine Menge ohne Elemente, die **leere Menge**. Wir verwenden dafür die Notation $\emptyset$:

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

(ii) Es gilt

$$\{a,b,c\} \cup \{a,7,\pi\} = \{a,b,c,7,\pi\}$$
$$\{a,b,c\} \cap \{a,7,\pi\} = \{a\}$$
$$\{a,b,c\} \setminus \{a,7,\pi\} = \{b,c\}.$$

(iii) Für $M = \{a,b,c\}$ und $N = \{1,7,\pi\}$ ergibt sich

$$M \times N = \{(a,1),(a,7),(a,\pi),(b,1),(b,7),(b,\pi),(c,1),(c,7),(c,\pi)\}.$$

(iv) Die Elemente von $M \times N$ sind *geordnete* Paare, es ist also beispielsweise (π, a) kein Element von $M \times N$ in (iii). Insbesondere sind $M \times N$ und $N \times M$ unterschiedliche Mengen.
(v) Hat M genau m und N genau n Elemente, so hat $M \times N$ genau $m \cdot n$ Elemente.
(vi) Das kartesische Produkt kann analog auch für mehr als zwei Mengen gebildet werden. Für Mengen $M_1, \ldots, M_n$ setzt man dabei

$$M_1 \times M_2 \times \cdots \times M_n := \{(a_1,\ldots,a_n) \mid a_1 \in M_1, \ldots, a_n \in M_n\}.$$

Für eine Menge M schreibt man auch

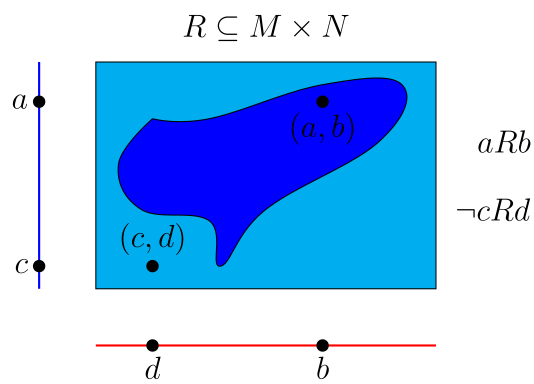
$$M^n := \underbrace{M \times \cdots \times M}_{n}.$$

Ein Beispiel dafür ist der aus der Schule bekannte $\mathbb{R}^3$. $\triangle$

Definition 1.4.4. Seien M, N Mengen. Eine **Relation** zwischen M und N ist eine Teilmenge

$$R \subseteq M \times N.$$

Statt $(a,b) \in R$ schreibt man manchmal auch aRb und sagt a *steht zu* b *in der Relation* R. Für $M = N$ sagt man auch, dass R eine Relation *auf* M ist.



△

Bemerkung/Beispiel 1.4.5. (i) Eine Relation gibt eine Beziehung zwischen Elementen an. Für jedes $a \in M$ und $b \in N$ kann diese Beziehung entweder gelten, oder auch nicht. Die Relation R als Menge enthält genau die Paare (a, b), für die a zu b in Relation steht.
(ii) Für $M = \{\text{Gabi}, \text{Monika}, \text{Erika}\}$ und $N = \{\text{Klaus}, \text{Martin}, \text{Horst}\}$ ist etwa eine Relation gegeben durch

$$R_1 := \{(\text{Gabi}, \text{Martin}), (\text{Monika}, \text{Horst}), (\text{Erika}, \text{Klaus})\}\,.$$

Eine andere Möglichkeit wäre aber zum Beispiel

$$R_2 = \{(\text{Gabi}, \text{Martin}), (\text{Monika}, \text{Martin}), (\text{Erika}, \text{Klaus}), (\text{Erika}, \text{Horst})\}\,.$$

Ohne das als gesellschaftspolitisches Statement zu verstehen, könnte in diesem Setup Monika beispielsweise nicht zu Erika in Relation stehen, denn Erika ist kein Element von N.
(iii) Mathematisch etwas relevanter ist die folgende Relation auf $\mathbb{R}$:

$$\leq := \left\{(a, b) \in \mathbb{R} \times \mathbb{R} \mid \exists c \in \mathbb{R} \colon a + c^2 = b\right\}.$$

Hier steht also a in Relation zu b genau dann wenn a kleiner gleich b ist. Hier kennen wir die alternative Notation $a \leq b$ statt $(a, b) \in \leq$ schon lange. △

Relationen können gewisse zusätzliche Eigenschaften haben, die von Bedeutung sind.

Definition 1.4.6. (i) Sei M eine Menge. Eine Relation $R \subseteq M \times M$ auf M heißt

reflexiv, falls	$\forall a \in M\colon$	$(a, a) \in R$,
symmetrisch, falls	$\forall a, b \in M\colon$	$(a, b) \in R \Rightarrow (b, a) \in R$,
antisymmetrisch, falls	$\forall a, b \in M\colon$	$(a, b) \in R \wedge (b, a) \in R \Rightarrow a = b$,
transitiv, falls	$\forall a, b, c \in M\colon$	$(a, b) \in R \wedge (b, c) \in R \Rightarrow (a, c) \in R$
vollständig, falls	$\forall a, b \in M\colon$	$(a, b) \in R \vee (b, a) \in R$.

(ii) Eine **Äquivalenzrelation** auf M ist eine reflexive, symmetrische und transitive Relation auf M.
(iii) Eine **partielle Ordnung** auf M ist eine reflexive, antisymmetrische und transitive Relation auf M. Die partielle Ordnung heißt **vollständige/totale Ordnung**, falls sie vollständig ist. △

Bemerkung/Beispiel 1.4.7. (i) Die in Beispiel 1.4.5 (iii) angegebene Relation $\leq$ ist eine vollständige Ordnung auf $\mathbb{R}$.
(ii) Auf $\mathbb{N}$ definieren wir folgende Relation (Teilbarkeit):

$$T = \{(a,b) \in \mathbb{N} \times \mathbb{N} \mid \exists c \in \mathbb{N} \colon a \cdot c = b\}.$$

Für $(a,b) \in T$ schreibt man oft auch $a \mid b$ und sagt a *teilt* b. Dann sieht man leicht, dass T eine partielle, aber keine total Ordnung auf $\mathbb{N}$ ist.
(iii) Auf jeder Menge M ist

$$\{(a,a) \mid a \in M\}$$

eine Äquivalenzrelation. Jedes Element steht nur zu sich selbst in Relation. Die Relation gibt also gerade die Gleichheit wieder.
(iv) Auf der Menge aller Personen im Raum definiert man folgende Relation:

$$H = \{(a,b) \mid a \text{ und } b \text{ haben dieselbe Haarfarbe}\}.$$

Dann ist H eine Äquivalenzrelation.
(v) Äquivalenzrelationen erlauben es also, einen verallgemeinerten Gleichheitsbegriff zu definieren. Es ist dabei nicht nur jedes Element zu sich selbst gleich, sondern kann im verallgemeinerten Sinne auch zu anderen Elementen gleich (äquivalent) sein. Die Eigenschaften reflexiv/symmetrisch/transitiv sind natürlich Eigenschaften, die man von einer verallgemeinerten Gleichheit erwarten würde.
(vi) Für Äquivalenzrelationen verwendet man häufig auch die Bezeichnung $\sim$. Statt $(a,b) \in R$ oder aRb schreibt man also $a \sim b$ (oder etwas genauer $a \sim_R b$). △

Definition 1.4.8. Sei R eine Äquivalenzrelation auf der Menge M. Für $a \in M$ definieren wir

$$[a] := \{b \in M \mid (a,b) \in R\},$$

und nennen es die **Äquivalenzklasse** von a bezüglich R. Es gilt also stets

$$a \in [a] \subseteq M.$$

△

Lemma 1.4.9 (★). *(i) Sei R eine Äquivalenzrelation auf M. Für $a, b \in M$ gilt dann stets*

$$[a] = [b] \text{ oder } [a] \cap [b] = \emptyset.$$

Die Äquivalenzklassen liefern also eine Zerlegung von M in paarweise disjunkte Teilmengen.
(ii) Sei I eine beliebige Indexmenge und für jedes $i \in I$ sei $M_i \subseteq M$ eine Teilmenge. Es gelte

$$M_i \cap M_j = \emptyset \text{ für } i \neq j \quad \text{sowie} \quad M = \bigcup_{i \in I} M_i.$$

Dann liefert folgende Setzung eine Äquivalenzrelation auf M, deren Äquivalenzklassen gerade die Mengen M_i sind:

$$R := \{(a, b) \in M \times M \mid \exists i \in I \colon a, b \in M_i\}.$$

Beweis. (i) Angenommen es gilt $[a] \cap [b] \neq \emptyset$. Wir zeigen, dass dann $[a] = [b]$ gelten muss.
Es gibt nun also $c \in M$ mit $c \in [a]$ und $c \in [b]$. Aus der Definition von Äquivalenzklassen folgt $(a, c) \in R$ und $(b, c) \in R$. Die Symmetrie von R impliziert $(c, b) \in R$ und aus der Transitivität folgt damit $(a, b) \in R$.
Sei nun $d \in [a]$ beliebig. Aus $(a, d) \in R$ folgt $(d, a) \in R$ und mit $(a, b) \in R$ dann wieder aus der Transitivität $(d, b) \in R$. Die Symmetrie impliziert $(b, d) \in R$, also $d \in [b]$. Wir haben gezeigt dass jedes Element aus $[a]$ auch zu $[b]$ gehört, also $[a] \subseteq [b]$. Die Inklusion $[b] \subseteq [a]$ zeigt man genauso (da die Voraussetzung $[a] \cap [b] = \emptyset$ aber völlig symmetrisch bezüglich a und b ist, muss eigentlich nichts mehr gezeigt werden). Damit gilt $[a] = [b]$.
Wegen $a \in [a]$ für jedes $a \in M$ liefern die Äquivalenzklassen also eine disjunkte Zerlegung von ganz M.
(ii) Die Relation R ist nach Konstruktion offensichtlich reflexiv und symmetrisch. Für die Transitivität gelte $(a, b) \in R$ und $(b, c) \in R$. Es gibt also $i, j \in I$ mit $a, b \in M_i$ und $b, c \in M_j$. Aus $i \neq j$ würde $M_i \cap M_j = \emptyset$ folgen, was wegen $b \in M_i \cap M_j$ nicht sein kann. Also gilt $i = j$, und somit $a, c \in M_i$, also $(a, c) \in R$. Damit ist gezeigt, dass R eine Äquivalenzrelation ist. Wegen

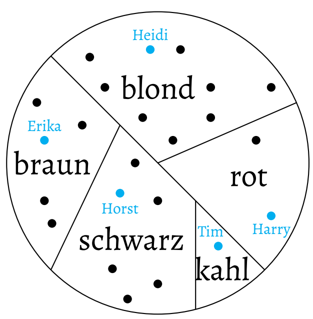
$$b \in [a] \;\Leftrightarrow\; (a, b) \in R \;\Leftrightarrow\; \exists i \in I \colon a, b \in M_i$$

folgt für $a \in M_i$ schon $[a] = M_i$. Die Mengen M_i sind also gerade die Äquivalenzklassen von R. □

Definition 1.4.10. Für eine Äquivalenzrelation R auf M nennen wir die Menge der Äquivalenzklassen auch **Faktormenge**:

$$M/R := \{[a] \mid a \in M\}. \qquad \triangle$$

Bemerkung/Beispiel 1.4.11. (i) Die Äquivalenzklasse $[a]$ enthält alle zu a äquivalenten Elemente. Es ist also die Zusammenfassung aller Elemente, die im verallgemeinerten Sinn als gleich wie a anzusehen sind. Für jedes $b \in [a]$ gilt $[a] = [b]$. Sowohl a als auch b sind also *Vertreter* der gleichen Äquivalenzklasse.
(ii) In Beispiel 1.4.7 (iv) enthält eine Äquivalenzklasse gerade alle Personen mit gleicher Haarfarbe.
(iii) Die Menge M/R hat Teilmengen von M als Elemente, nämlich die Äquivalenzklassen. Da verschiedene Elemente von M in der selben Äquivalenzklasse liegen können, enthält M/R potenziell weniger Elemente als M. Wenn in Beispiel 1.4.7 der Raum 100 Personen enthält, die insgesamt 5 verschiedene Haarfarben haben, ist M eine Menge mit 100 und M/R eine Menge mit 5 Elementen.



Menge M mit Äquivalenzrelation

$\{[\text{Heidi}], [\text{Erika}], [\text{Horst}], [\text{Harry}], [\text{Tim}]\}$ Faktormenge M/R $\triangle$

Konstruktion 1.4.12 (★). Wenn wir $\mathbb{Z}$ mit den Rechenregeln als bekannt voraussetzen, können wir nun eine exakte Definition von $\mathbb{Q}$ geben. Dazu definieren wir zunächst eine Äquivalenzrelation auf der Menge $\mathbb{Z} \times (\mathbb{Z} \setminus \{0\})$:

$$(a, b) \sim (c, d) \;:\Leftrightarrow\; ad = bc.$$

Man muss nun zunächst nachrechnen, dass es sich hier wirklich um eine Äquivalenzrelation handelt. Ist dies erledigt, definieren wir $\mathbb{Q}$ einfach als Faktormenge dieser Äquivalenzrelation:

$$\mathbb{Q} := (\mathbb{Z} \times (\mathbb{Z} \setminus \{0\})) \,/ \sim .$$

Für die Äquivalenzklasse von (a, b) schreibt man statt $[(a, b)]$ hier gewöhnlich $\frac{a}{b}$, was zur bekannte Notation

$$\mathbb{Q} = \left\{ \frac{a}{b} \mid a, b \in \mathbb{Z}, b \neq 0 \right\}$$

führt. Es gilt beispielsweise

$$(1, 2) \sim (2, 4), \text{ denn } 1 \cdot 4 = 2 \cdot 2.$$

Daraus folgt $[(1, 2)] = [(2, 4)]$, oder in anderer Schreibweise gerade

$$\frac{1}{2} = \frac{2}{4}.$$

Nun versieht man $\mathbb{Q}$ gewöhnlich mit *Addition* und *Multiplikation*, die folgendermaßen definiert wird:

$$\frac{a}{b} + \frac{c}{d} := \frac{ad + bc}{bd}$$
$$\frac{a}{b} \cdot \frac{c}{d} := \frac{ac}{bd}.$$

Hier taucht ein wichtiges Phänomen auf, nämlich das der *Wohldefiniertheit*. Wir definieren hier Addition und Multiplikation von Äquivalenzklassen, also streng genommen von Teilmengen von $\mathbb{Z} \times (\mathbb{Z} \setminus \{0\})$. Auf der rechten Seite der Definition wird aber Bezug auf einen *Vertreter* der Äquivalenzklasse Bezug genommen, also auf ein Element der jeweiligen Menge. Dieser Vertreter ist aber keineswegs eindeutig. Theoretisch könnte das zu Problem führen. Zum Beispiel erhalten wir mit der oberen Formel

$$\frac{1}{2} \cdot \frac{3}{5} = \frac{3}{10}.$$

Wir hätten aber statt $\frac{1}{2}$ auch $\frac{2}{4}$ schreiben können, und erhielten dann

$$\frac{2}{4} \cdot \frac{3}{5} = \frac{6}{20}.$$

Formal betrachtet haben wir im ersten Fall für die Äquivalenzklasse $\frac{1}{2}$ den Vertreter $(1, 2)$ in der Formel verwendet und im zweiten Fall den Vertreter $(2, 4)$. Erfreulicherweise gilt nun aber

$$\frac{3}{10} = \frac{6}{20},$$

wie man direkt sieht, d.h. die Elemente $(3, 10)$ und $(6, 20)$ liegen in der selben Äquivalenzklasse. Das kann (und muss) man nun noch ganz Allgemein für $+$ und $\cdot$ auf $\mathbb{Q}$ beweisen. $\triangle$

Definition 1.4.13. Seien M, N Mengen. Eine **Funktion/Abbildung** f von M nach N ist eine *linkstotale* und *rechtseindeutige* Relation zwischen M und N, d.h. eine Teilmenge $f \subseteq M \times N$ mit folgender zusätzlicher Eigenschaft:

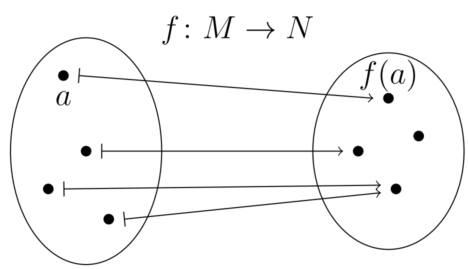
$$\forall a \in M \; \exists! b \in N \colon (a, b) \in f.$$

Dabei heißt M **Definitionsbereich** und N **Zielbereich** der Funktion. △

Bemerkung 1.4.14. (i) Die definierende Eigenschaft einer Funktion besagt gerade, dass für jedes Element a aus dem Definitionsbereich M *genau ein* Element b aus dem Zielbereich N existiert, welches zu a in Relation steht (also immer mindestens eins, aber auch nie mehr als eins). Aufgrund der Eindeutigkeit von b kann man eine Funktion also auch als *Zuordnungsvorschrift* verstehen, die jedem $a \in M$ ein $b \in N$ zuordnet.
(ii) Statt $f \subseteq M \times N$ schreibt man für Funktionen meistens $f \colon M \to N$, und statt $(a, b) \in f$ auch

$$b = f(a).$$

Man nennt dann b auch das **Bild von** a **unter** f.



(iii) Wenn man die Notation $f \colon M \to N$ wählt, nennt man die Darstellung $f \subseteq M \times N$ manchmal auch **Graph** von f. In strengen Sinn der Definition ist eine Funktion aber nichts anderes als ihr Graph.
(iv) Für $T \subseteq M$ nennt man die Menge

$$f(T) := \{f(a) \mid a \in T\}$$

das **Bild von** T **unter** f, und die Menge

$$\mathrm{Bild}(f) := f(M)$$

das **Bild von** f.

(v) Für $S \subseteq N$ nennt man

$$f^{-1}(S) := \{a \in M \mid f(a) \in S\}$$

das **Urbild von S unter f** und für $b \in N$

$$f^{-1}(b) := f^{-1}(\{b\}) = \{a \in M \mid f(a) = b\}$$

das **Urbild** von b. Man beachte aber, dass dabei f^{-1} im Allgemeinen nicht als Funktion von N nach M aufgefasst werden kann, denn das Urbild eines Punktes kann mehrere Punkte enthalten. Das widerspräche der Rechtseindeutigkeit. △

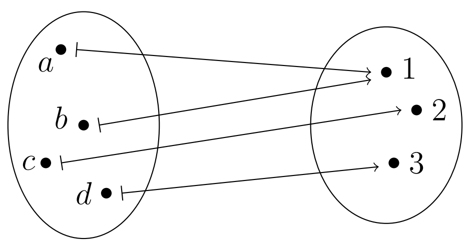
Beispiel 1.4.15. (i) Wir betrachten die Funktion $f\colon \{a,b,c,d\} \to \{1,2,3\}$ definiert durch

$$f(a) = 1,\ f(b) = 1,\ f(c) = 2,\ f(d) = 3.$$

In der formalen Darstellung als Graph erhalten wir

$$f = \{(a,1),(b,1),(c,2),(d,3)\}$$

und schematisch können wir f so darstellen:



Es gilt beispielsweise

$$f(\{a,c\}) = \{1,2\},\ f^{-1}(\{1,3\}) = \{a,b,d\},\ f^{-1}(3) = \{d\},\ f^{-1}(1) = \{a,b\}.$$

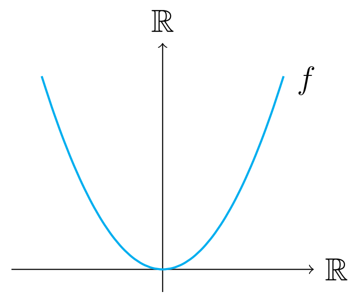
(ii) Auf jeder Menge M gibt es die **Identitätsfunktion**:

$$\begin{aligned} \mathrm{id}_M\colon M &\to M \\ a &\mapsto a. \end{aligned}$$

(iii) Die Funktion $f = \{(a,a^2) \in \mathbb{R} \times \mathbb{R} \mid a \in \mathbb{R}\}$ würde man gewöhnlich so angeben:

$$\begin{aligned} f\colon \mathbb{R} &\to \mathbb{R} \\ a &\mapsto a^2. \end{aligned}$$

Ihr Graph hat dann folgende Gestalt:



Es gilt $\mathrm{Bild}(f) = \mathbb{R}_{\geq 0} = f(\mathbb{R}_{\geq 0})$ und $f^{-1}(4) = \{-2, 2\}$. △

Definition 1.4.16. Sei $f\colon M \to N$ eine Abbildung. Dann heißt f

(i) **injektiv**, falls jedes Element im Zielbereich höchstens ein Urbild hat:

$$\forall b \in N\colon \#f^{-1}(b) \leq 1.$$

(ii) **surjektiv**, falls jedes Element im Zielbereich mindestens ein Urbild hat:

$$\forall b \in N\colon \#f^{-1}(b) \geq 1.$$

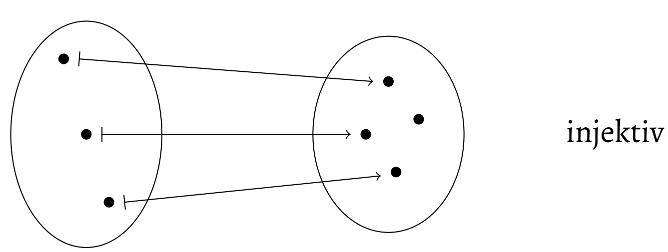
(iii) **bijektiv**, falls f injektiv und surjektiv ist:

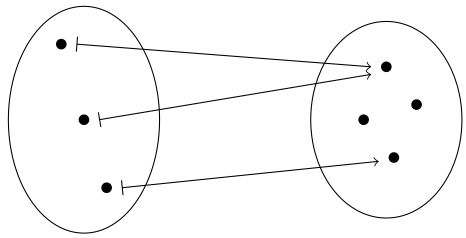
$$\forall b \in N\colon \#f^{-1}(b) = 1.$$ △

Bemerkung/Beispiel 1.4.17. (i) Die Injektivität von f kann man auch folgendermaßen ausdrücken:

$$\forall a, b \in M\colon a \neq b \Rightarrow f(a) \neq f(b).$$

Zwei verschiedene Elemente des Definitionsbereichs werden also stets auf verschiedene Elemente im Zielbereich abgebildet. In der schematischen Darstellung von oben bedeutet das, dass niemals zwei verschiedene Pfeile am selben Punkt enden. Man beachte, dass sich diese Eigenschaft von der Rechtseindeutigkeit (in der Funktionsdefinition) unterscheidet; diese besagt gerade, dass von jedem Element im Definitionsbereich genau ein Pfeil ausgeht.



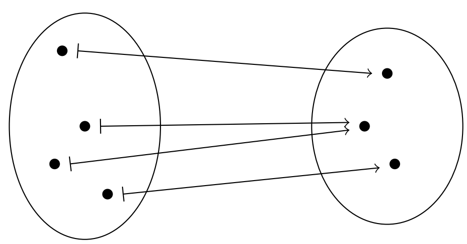


nicht injektiv

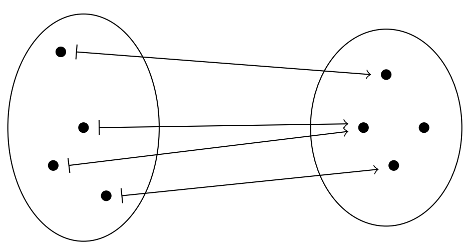
Wenn man die Injektivität einer Abbildung andeuten möchte, verwendet man manchmal folgende Notation:

$$f\colon M \hookrightarrow N.$$

(ii) Die Surjektivität besagt, dass jedes Element im Zielbereich von einem Pfeil getroffen wird, oder auch $\mathrm{Bild}(f) = N$:



surjektiv



nicht surjektiv

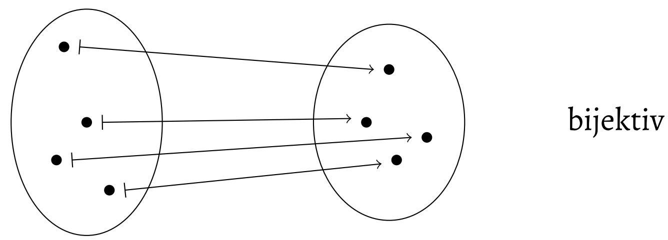
Um die Surjektivität anzudeuten, schreibt man manchmal

$$f\colon M \twoheadrightarrow N.$$

Die Surjektivität einer Abbildung kann man notfalls dadurch erzwingen, dass man den Zielbereich N durch das Bild $\mathrm{Bild}(f)$ der Abbildung ersetzt:

$$f\colon M \twoheadrightarrow \mathrm{Bild}(f).$$

(iii) Injektivität und Surjektivität sind voneinander unabhängig Bedingungen. Keine der beiden impliziert die andere oder schließt sie aus.
(iv) Die Bijektivität besagt, dass jedes Element im Zielbereich genau ein Urbild besitzt. Damit stellt f eine eins-zu-eins Zuordnung zwischen den Elementen von M und N her:



In dieser Situation kann man offensichtlich f^{-1} als Abbildung von N nach M auffassen:

$$f^{-1} := \{(b, a) \in N \times M \mid f(a) = b\}$$

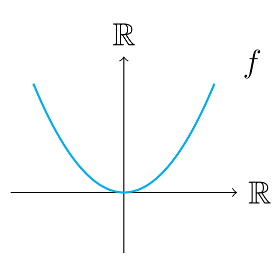
und man nennt f^{-1} dann **Umkehrabbildung** von f.
(v) Eine injektive Abbildung $f\colon M \to N$ wird bijektiv, wenn N durch $\mathrm{Bild}(f)$ ersetzt wird. In diesem Sinne existiert also für jede injektive Abbildung die Umkehrabbildung

$$f^{-1}\colon \mathrm{Bild}(f) \to M.$$

(vi) Die Identitätsfunktion $\mathrm{id}_M\colon M \to M$ ist stets bijektiv. Es gibt aber im Allgemeinen noch weitere bijektive Funktionen von M nach M. Beispiele sind Drehungen und Spiegelungen des $\mathbb{R}^2$. △

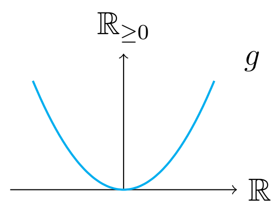
Beispiel 1.4.18. Wir betrachten die Funktion

$$\begin{aligned} f\colon \mathbb{R} &\to \mathbb{R} \\ a &\mapsto a^2. \end{aligned}$$



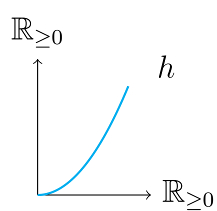
Es gilt $f(-1) = 1 = f(1)$, also ist f nicht injektiv. Es gilt $-1 \notin \mathbb{R}_{\geq 0} = \mathrm{Bild}(f)$, also ist f auch nicht surjektiv. Wir ersetzen nun den Zielbereich durch $\mathbb{R}_{\geq 0}$:

$$\begin{aligned} g\colon \mathbb{R} &\to \mathbb{R}_{\geq 0} \\ a &\mapsto a^2. \end{aligned}$$



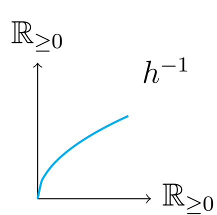
Die Abbildung g ist nun surjektiv, aber wegen $g(-1) = 1 = g(1)$ immer noch nicht injektiv. Wenn wir nun auch noch den Definitionsbereich zu $\mathbb{R}_{\geq 0}$ verkleinern, erhalten wir eine injektive und surjektive Abbildung, also eine Bijektion:

$$\begin{aligned} h\colon \mathbb{R}_{\geq 0} &\to \mathbb{R}_{\geq 0} \\ a &\mapsto a^2. \end{aligned}$$



Die Umkehrfunktion von h nennt man Wurzelfunktion:

$$\begin{aligned} h^{-1}\colon \mathbb{R}_{\geq 0} &\to \mathbb{R}_{\geq 0} \\ a &\mapsto \sqrt{a}. \end{aligned}$$



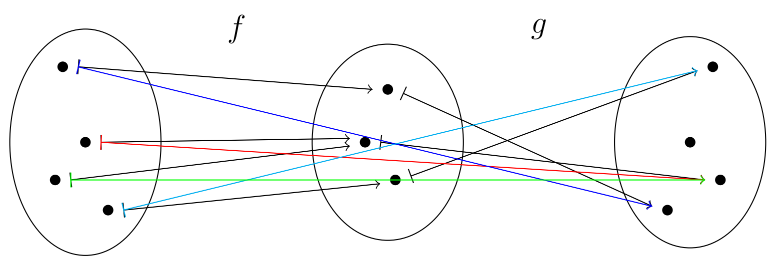
$\triangle$

Definition 1.4.19. Seien $f\colon M \to N$ und $g\colon N \to O$ Funktionen. Dann ist die **Hintereinanderausführung** von f und g definiert als

$$\begin{aligned} g \circ f\colon M &\to O \\ a &\mapsto g\left(f(a)\right). \end{aligned}$$

Dabei liest man das Symbol $\circ$ als *nach*, was auch die Reihenfolge $g \circ f$ erklärt.

$$\begin{tikzcd} M \arrow[r, "f"] \arrow[rr, bend right, "g\circ f"'] & N \arrow[r, "g"] & O \end{tikzcd}$$



$g \circ f$

$\triangle$

Bemerkung/Beispiel 1.4.20. (i) Sei $f\colon \{a, b, c\} \to \{1, 2\}$ definiert durch

$$f(a) = 1,\ f(b) = 2,\ f(c) = 1$$

und $g\colon \{1, 2\} \to \{\text{Klaus, Bello, Patscherkofel}\}$ definiert durch

$$g(1) = \text{Patscherkofel},\ g(2) = \text{Bello},$$

so erhalten wir für $g \circ f\colon \{a, b, c\} \to \{\text{Klaus, Bello, Patscherkofel}\}$

$$(g \circ f)(a) = \text{Patscherkofel}, (g \circ f)(b) = \text{ Bello},\ (g \circ f)(c) = \text{Patscherkofel}.$$

(ii) Für $f\colon M \to N$ gilt stets

$$f \circ \mathrm{id}_M = f, \quad \mathrm{id}_N \circ f = f.$$

(iii) Ist $f\colon M \to N$ bijektiv und $f^{-1}\colon N \to M$ die Umkehrabbildung, so gilt

$$f^{-1} \circ f = \mathrm{id}_M, \quad f \circ f^{-1} = \mathrm{id}_N\,.$$

(iv) Die Hintereinanderausführung von Abbildungen erfüllt das *Assoziativgesetz*, d.h. es gilt

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

für Abbildungen $f\colon M \to N, g\colon N \to O, h\colon O \to P$.
(v) Die Hintereinanderausführung von Abbildungen ist im Allgemeinen *nicht kommutativ*! Es gibt also Abbildungen $f\colon M \to M, g\colon M \to M$ mit

$$g \circ f \neq f \circ g.$$

Beispielsweise kann man das mit einer Drehung und einer Achsenspiegelung des $\mathbb{R}^2$ anschaulich leicht sehen. △

Kapitel 2

Gauß-Algorithmus und Matrixrechnung

In diesem Abschnitt lernen wir zunächst den Gauß-Algorithmus über $\mathbb{R}$ kennen. Mit ihm kann man lineare Gleichungssysteme algorithmisch vollständig lösen. Wir sehen dann, dass man statt $\mathbb{R}$ beispielsweise auch $\mathbb{Q}$ verwenden kann und das führt uns zur allgemeinen Definition eines *Körpers*. Das sind gerade die Zahlbereiche, über denen der Gauß-Algorithmus wie über $\mathbb{R}$ funktioniert. Wir lernen einige Beispiele dazu kennen, allen voran die *komplexen Zahlen*. Danach beschäftigen wir uns mit Matrixrechnung, die viel notationelle Vereinfachung und ein tieferes Verständnis des Gauß-Algorithmus liefert.

2.1 Der Gauß-Algorithmus

In diesem Abschnitt arbeiten wir ausschließlich mit $\mathbb{R}$, alle auftretenden Zahlen sollen also reell sein. In den meisten Anwendungen in Physik, Technik etc. wird auch mit den reellen Zahlen gearbeitet. Man sollte dabei nicht vergessen, dass die reellen Zahlen bei weitem kein einfaches oder offensichtliches Konstrukt sind. Mathematisch exakt wurden sie sogar erst gegen Ende des 19. Jahrhunderts gefasst. Für eine saubere Einführung in die reellen Zahlen verweisen wir auf die Analysis-Vorlesungen.

Beispiel 2.1.1. Aus der Schule bekannt sind meistens Systeme von linearen Gleichungen in zwei oder drei Variablen. Ein solches System ist beispielsweise

$$\begin{aligned} 2x + y &= 10 \\ x + y &= \ \ 6 \end{aligned}$$

Um das System zu lösen, möchte man alle (reellen) Werte für x und y bestimmen, die bei gemeinsamer Einsetzung in beide Gleichungen wahre Aussagen ergeben. Dies erreicht man gewöhnlich durch *Äquivalenzumformungen*, also Umformungen des Gleichungssystems, welche die Lösungsmenge nicht verändern. Am Ende sollte dabei die Lösungsmenge einfach abzulesen sein. Bei einfachen System kann man beispielsweise die Strategie des Auflösens und Einsetzens verwenden. Die zweite Gleichung ist offensichtlich äquivalent zu

$$y = 6 - x.$$

Setzt man das ins erste System ein, erhält man

$$2x + (6 - x) = 10,$$

also

$$x + 6 = 10.$$

Hier liest man die einzige Lösung $x = 4$ direkt ab, und erhält $y = 6 - 4 = 2$. Das System hat also die einelementige Lösungsmenge $L = \{(4, 2)\}$. △

Die Strategie des Auflösens und Einsetzens wird schnell unpraktikabel, wenn die Anzahl der Variablen und Gleichungen wächst. Bei 10 Gleichungen in 15 Variablen möchte eigentlich niemand mehr diese Rechnung durchführen. Wir gehen nun also etwas systematischer vor.

Definition 2.1.2. (i) Ein **lineares Gleichungssystem** ist ein Ausdruck der folgenden Form:

$$\begin{aligned} a_{11}x_1 + \ a_{12}x_2 + \cdots + \ a_{1n}x_n &= \ b_1 \\ a_{21}x_1 + \ a_{22}x_2 + \cdots + \ a_{2n}x_n &= \ b_2 \\ \vdots \qquad\qquad\qquad\qquad & \ \ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned} \tag{2.1}$$

Dabei sind a_{ij} und b_i vorgegebene *Koeffizienten* (hier aus $\mathbb{R}$) des Gleichungssystems. Die $x_1, \ldots, x_n$ sind *Variablen*.

(ii) Das lineare Gleichungssystem (2.1) heißt **homogen**, falls

$$b_1 = b_2 = \cdots = b_m = 0$$

gilt. Ansonsten heißt es **inhomogen**.
(iii) Die **Koeffizientenmatrix** des Gleichungssystems (2.1) ist folgendes rechteckiges Schema:

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix},$$

die **erweiterte Koeffizientenmatrix** ist

$$(A, b) = \begin{pmatrix} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{pmatrix}.$$

(iv) Die **Lösungsmenge** des Systems (2.1) ist

$$L(A,b) := \{(c_1, \ldots, c_n) \in \mathbb{R}^n \mid \text{rechts und links in (2.1) steht in jeder Gleichung dieselbe Zahl, wenn jedes } x_i \text{ durch } c_i \text{ ersetzt wird}\}. \qquad \triangle$$

Die Strategie des Algorithmus von Gauß ist nun, das Gleichungssystem (bzw. eigentlich nur seine erweiterte Koeffizientenmatrix) so umzuformen, dass sich einerseits die Lösungsmenge nicht ändert und andererseits die Lösungsmenge des neuen Systems leicht abzulesen ist. Leicht abzulesen ist die Lösungsmenge bei Koeffizientenmatrizen in *Zeilenstufenform*:

Definition 2.1.3. Eine $m \times n$-Matrix A ist in **Zeilenstufenform**, falls sie folgende Gestalt hat:

$$A = \begin{pmatrix} \circledast & & & & \\ & \circledast & & & \\ & & \circledast & & \\ & & & \ddots & \\ & 0 & & & \circledast \end{pmatrix}.$$

Dabei müssen alle Einträge $\circledast$ ungleich Null sein und unterhalb der Stufenlinien darf überall nur Null stehen. Die Einträge $\circledast$ nennt man **Pivots**. In exakterer (aber unverständlicherer) Ausdrucksweise besagt das:

(i) Es gibt ein $0 \leq r \leq m$, so dass alle Zeilen mit Index $r+1$ bis m nur Nullen enthalten.

(ii) In jeder Zeile mit Index $1 \leq i \leq r$ gibt es einen Eintrag $\neq 0$. Wenn j_i der kleinste Spaltenindex ist, für den der Eintrag in der i-ten Zeile $\neq 0$ ist, so gilt
$$j_1 < j_2 < \cdots < j_r. \qquad \triangle$$

Bemerkung 2.1.4. Hat die Koeffizientenmatrix A eines linearen Gleichungssystems Zeilenstufenform, kann man die Lösungsmenge des Systems direkt ablesen. Die erweiterte Koeffizientenmatrix sieht dann so aus:

$$\left(\begin{array}{ccccc|c} \circledast & & & & & b_1 \\ & \circledast & & & & b_2 \\ & & \circledast & & & b_3 \\ & & & \ddots & & \vdots \\ & & & & \circledast & b_r \\ \hline & 0 & & & & b_{r+1} \\ & & & & & \vdots \\ & & & & & b_m \end{array}\right).$$

Indem man das in Gedanken wieder in ein klassisches Gleichungssystem übersetzt, sieht man direkt:

(i) Das System hat genau dann eine Lösung, wenn $b_{r+1} = \cdots = b_m = 0$ gilt.

(ii) Ist das System lösbar, erhält man alle Lösungen auf folgende Weise:

Falls in der i-ten Spalte von A kein Pivot steht, kann für x_i eine beliebige Zahl eingesetzt werden. Solche Variablen nennt man **freie** Variablen.

Die Werte für alle anderen Variablen sind dann eindeutig bestimmt (diese Variablen nennt man **gebunden**). Man erhält sie, indem man von unten, beginnend mit der r-ten Zeile, die entsprechenden Gleichungen nach der Variable am Pivot auflöst.

Durch diese Prozedur erhält man eine *Parametrisierung* der gesamten Lösungsmenge durch $\mathbb{R}^{n-r}$, d.h. eine bijektive Abbildung
$$\varphi\colon \mathbb{R}^{n-r} \to L(A,b).$$

Dabei entspricht ein Tupel $d \in \mathbb{R}^{n-r}$ einer Wahl von Werten für die freien Variablen (davon gibt es $n-r$ viele). Das Tupel $\varphi(d)$ ist dann die eindeutig bestimmte Fortsetzung zu einer Lösung des Systems. $\triangle$

Beispiel 2.1.5. Wir betrachten folgende erweiterte Koeffizientenmatrix, in der A Zeilenstufenform hat:

$$(A,b) = \left(\begin{array}{cccc|c} \frac{1}{2} & 2 & -\frac{3}{4} & 0 & 9 \\ 0 & 0 & 2 & 1 & -4/3 \\ 0 & 0 & 0 & -1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

Hier ist $m = n = 4$ und $r = 3$. Wegen $b_4 = 0$ ist das System lösbar. Die einzige freie Variable ist x_2. Wenn wir für x_2 also den Wert $d \in \mathbb{R}$ wählen, ergibt sich aus der dritten Zeile der einzig mögliche Wert -2 für x_4, damit aus der zweiten Zeile $\frac{1}{3}$ für x_3 und schließlich aus der erste Zeile $\frac{37}{2} - 4d$ für x_1. Dadurch ergibt sich folgende Parametrisierung der Lösungsmenge:

$$\begin{aligned} \varphi\colon \mathbb{R} &\to L(A,b) \\ d &\mapsto \left(\tfrac{37}{2} - 4d, d, \tfrac{1}{3}, -2\right). \end{aligned}$$

Die Lösungsmenge kann auch einfach als

$$L(A,b) = \left\{ \left(\tfrac{37}{2} - 4d, d, \tfrac{1}{3}, -2\right) \mid d \in \mathbb{R} \right\}$$

angegeben werden. $\triangle$

Im Gauß-Algorithmus formen wir ein Gleichungssystem so lange um, bis es Zeilenstufenform hat. Wir geben die nötigen Definitionen nun in der Sprache der (erweiterte) Koeffizientenmatrix an.

Definition 2.1.6. Eine **elementare Zeilenumformung** einer Matrix ist eine der drei folgenden Operationen:

(I) Vertauschung von zwei Zeilen.

(II) Multiplikation (jedes Eintrags) einer Zeile mit einer Zahl $\neq 0$.

(III) Addition eines Vielfachen ($\neq$ 0) einer Zeile zu einer (echt) anderen (eintragsweise). $\triangle$

Der folgende Satz beinhaltet die wichtigsten Aussagen zum Gauß-Algorithmus. Der eigentliche Algorithmus wird im Beweis von Teil (ii) beschrieben.

Satz 2.1.7 (Algorithmus von Gauß). *Sei (A, b) die erweiterte Koeffizientenmatrix eines linearen Gleichungssystems.*

(i) *Entsteht (A', b') durch eine endliche Abfolge von elementaren Zeilenumformungen aus (A, b), so gilt*
$$L(A', b') = L(A, b).$$

(ii) *(A, b) lässt sich durch eine endliche Abfolge von elementaren Zeilenumformungen auf eine Gestalt (A', b') bringen, in der A' Zeilenstufenform hat.*

Beweis. (i): Das Vertauschen zweier Zeilen entspricht dem Vertauschen zweier Gleichungen, was die Lösungsmenge offensichtlich nicht ändert.
Das Multiplizieren der i-ten Zeile mit $\lambda \neq 0$ entspricht dem Multiplizieren aller Koeffizienten der i-ten Gleichung (einschließlich dem b_i) mit λ. Ist $c = (c_1, \ldots, c_n)$ eine Lösung der ursprünglichen Gleichung, so gilt

$$a_{i1}c_1 + \cdots + a_{in}c_n = b_i.$$

Daraus folgt

$$(\lambda a_{i1})c_1 + \cdots + (\lambda a_{in})c_n = \lambda(a_{i1}c_1 + \cdots + a_{in}c_n) = \lambda b_i,$$

als ist c auch eine Lösung des neuen Systems, die Lösungsmenge wird also höchstens größer. Da diese Umformung durch eine Umformung desselben Typs (Multiplikation mit λ^{-1}) wieder rückgängig gemacht werden kann und das System dabei abermals höchstens größer wird, ist es stets gleich geblieben.
Wir addieren nun das λ-fache der i-ten Zeile zur j-ten. Sei $c = (c_1, \ldots, c_n)$ eine Lösung des ursprünglichen Systems, insbesondere gelte also

$$\begin{aligned} a_{i1}c_1 + \cdots + a_{in}c_n &= b_i \\ a_{j1}c_1 + \cdots + a_{jn}c_n &= b_j. \end{aligned}$$

Daraus folgt

$$\begin{aligned} \lambda b_i + b_j &= \lambda\left(a_{i1}c_1 + \cdots + a_{in}c_n\right) + \left(a_{j1}c_1 + \cdots + a_{jn}c_n\right) \\ &= (\lambda a_{i1} + a_{j1})c_1 + \cdots + (\lambda a_{in} + a_{jn})c_n. \end{aligned}$$

Also löst c auch die neu entstandene Gleichung im neuen System. Also ist die Lösungsmenge auch in diesem Schritt höchstens größer geworden. Auch diese Umformung lässt sich durch eine vom selben Typ invertieren, nämlich durch Addition des $(-\lambda)$-fachen der i-ten Zeile zur j-ten. Also bleiben die Lösungsmengen stets gleich und wir haben (i) bewiesen.
(ii): Die Nullmatrix hat bereits Zeilenstufenform, also können wir annehmen, dass A mindestens einen Eintrag $\neq 0$ hat. Wir wählen nun eine Zeile, in der ein Eintrag $\neq 0$ in A mit kleinstem Spaltenindex j_1 auftritt. Wir vertauschen die erste Zeile von (A, b) dann mit dieser Zeile (Umformung vom Typ (I)). In allen Spalten von kleinerem Index als j_1 steht nun also nur Null und an der Stelle $(1, j_1)$ steht ein Eintrag $\neq 0$. Damit können wir mit Umformungen vom Typ (III) alle Einträge unterhalb des $(1, j_1)$-Eintrags ebenfalls eliminieren.
Nun ignorieren wir die erste Zeile und die ersten j_1 Spalten der Matrix und iterieren den Prozess mit der übrigen (kleineren) Matrix. Man beachte, dass die erste Zeile und die ersten j_1 Spalten der Matrix durch die folgenden Umformungen nicht mehr verändert werden. Nach endlich vielen Schritten erreichen wir also Zeilenstufenform. □

Bemerkung 2.1.8. (i) Der Gauß-Algorithmus und Bemerkung 2.1.4 ermöglichen es nun, für lineare Gleichungssysteme sämliche Lösungen zu bestimmen. Wir bringen die erweiterte Koeffizientenmatrix mit elementaren Zeilenumformungen auf Zeilenstufenform und lesen die Lösungen danach ab.
(ii) Am Beweis von (ii) in Satz 2.1.7 sehen wir, dass Umformungen von Typ (I) und (III) schon ausreichen, um Zeilenstufenform zu bekommen. Typ (II) benötigt man nur, wenn man alle Pivots auf 1 setzen möchte.
(iii) Man kann offensichtlich auch noch erreichen, dass alle Einträge oberhalb der Pivots 0 sind. Sie können einfach mit Umformungen vom Typ (III), von links mit den Pivots beginnend, eliminiert werden. Diese Form macht das Ablesen der Lösungen (wie im Bemerkung 2.1.4 beschrieben) nochmal einfacher. Man kann in jeder Gleichung direkt den Wert für (die einzige) gebundene Variable aus den Werten für die freien Variablen bestimmen, ohne iterativ von unten diese Werte auszurechnen. Sind zusätzlich die Pivots alle 1, muss dabei nicht einmal eine Division durchgeführt werden. Eine Zeilenstufenform mit Pivots $= 1$ und Nullen oberhalb der Pivots heißt **reduzierte Zeilenstufenform**.
(iv) Wichtig beim Gauß-Algorithmus ist es, immer die *erweiterte* Koeffizientenmatrix umzuformen. Ansonsten ändert sich die Lösungsmenge natürlich offensichtlich. Nur bei homogenen Systemen ist das nicht nötig, da sich die Nullspalte ganz hinten sowieso niemals ändert, wenn Zeilenumformungen vorgenommen

werden.
(v) Spaltenumformungen sind im Gauß-Algorithmus nicht ohne Probleme erlaubt. Vertauscht man etwa zwei Spalten, so ändert sich die Lösungsmenge, wenn auch nicht besonders dramatisch. Es müssen in der Lösung des neuen Systems zwei entsprechende Koordinaten vertauscht werden, um die Lösungsmenge des alten Systems zu erhalten. Um solche Komplikationen zu vermeiden, beschränkt man sich am besten ausschließlich auf Zeilentransformationen. △

Beispiel 2.1.9. Wir berechnen eine Zeilenstufenform:

$$
\begin{aligned}
(A,b) = \left(\begin{array}{cccc|c} 1 & 2 & 3 & 4 & -1 \\ 5 & 6 & 7 & 8 & -2 \\ 9 & 10 & 11 & 12 & -3 \end{array}\right)
&\overset{Z_2-5Z_1}{\rightsquigarrow} \left(\begin{array}{cccc|c} 1 & 2 & 3 & 4 & -1 \\ 0 & -4 & -8 & -12 & 3 \\ 9 & 10 & 11 & 12 & -3 \end{array}\right) \\
&\overset{Z_3-9Z_1}{\rightsquigarrow} \left(\begin{array}{cccc|c} 1 & 2 & 3 & 4 & -1 \\ 0 & -4 & -8 & -12 & 3 \\ 0 & -8 & -16 & -24 & 6 \end{array}\right) \\
&\overset{Z_3-2Z_2}{\rightsquigarrow} \left(\begin{array}{cccc|c} 1 & 2 & 3 & 4 & -1 \\ 0 & -4 & -8 & -12 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right).
\end{aligned}
$$

Hier sehen wir, dass das zugehörige Gleichungssystem eine Lösung besitzt. Die freien Variablen sind x_3 und x_4. Setzt man $a \in \mathbb{R}$ für x_3 und $b \in \mathbb{R}$ für x_4 ein, erhalten wir aus der zweiten Gleichung den einzig möglichen Wert $-\frac{3}{4} - 2a - 3b$ für x_2. Damit ergibt sich mit der ersten Gleichung für x_1 der Wert $\frac{1}{2} + a + 2b$. Insgesamt erhalten wir

$$
L(A,b) = \left\{ \left(\tfrac{1}{2} + a + 2b, -\tfrac{3}{4} - 2a - 3b, a, b\right) \mid a, b \in \mathbb{R} \right\}.
$$

Wir berechnen auch noch die reduzierte Zeilenstufenform:

$$
\begin{aligned}
\left(\begin{array}{cccc|c} 1 & 2 & 3 & 4 & -1 \\ 0 & -4 & -8 & -12 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)
&\overset{-\frac{1}{4}\cdot Z_2}{\rightsquigarrow} \left(\begin{array}{cccc|c} 1 & 2 & 3 & 4 & -1 \\ 0 & 1 & 2 & 3 & -\frac{3}{4} \\ 0 & 0 & 0 & 0 & 0 \end{array}\right) \\
&\overset{Z_1-2Z_2}{\rightsquigarrow} \left(\begin{array}{cccc|c} 1 & 0 & -1 & -2 & \frac{1}{2} \\ 0 & 1 & 2 & 3 & -\frac{3}{4} \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)
\end{aligned}
$$

Hier findet man die oben angegebene Beschreibung der Lösungsmenge noch einfacher. △

2.2 Gruppen, Ringe und Körper

Wir schauen uns nun genau an, welche Eigenschaften von $\mathbb{R}$ wir im letzten Abschnitt eigentlich verwendet haben.
Zunächst rechnen wir in $\mathbb{R}$ stets mit den beiden *Rechenoperationen* $+$ und $\cdot$. Dabei haben wir, vermutlich ohne es zu bemerken, immer wieder *Assoziativ-, Kommutativ- und Distributivgesetz* für Addition und Multiplikation verwendet. Im Beweis von Satz 2.1.7 (i) sieht man das beispielsweise sehr schön. Auch bereits in der Definition der elementaren Umformungsschritte war das Kommutativgesetz implizit vorausgesetzt. Ansonsten hätten wir ja z.B. die Reihenfolge bei der Addition angeben müssen.
Aber wir haben noch mehr verwendet. Um die Einträge unterhalb der Pivots zu eliminieren mussten wir verwenden, dass man zu jeder Zahl eine geeignete Zahl addieren kann, um Null zu erhalten. Das folgt aus der Existenz der *Subtraktion* von Zahlen. Um wirklich mit jedem Pivot jede andere Zahl elimieren zu können, muss man aber auch noch *dividieren* können. Das sieht man in Beispiel 2.1.9 gut. Weitere Eigenschaften wurden für die Existenz der Zeilenstufenform aber nicht verwendet. Auch die Extraktion der Lösungen (siehe Bemerkung 2.1.4) benötigt diese, aber keine weiteren Eigenschaften. Die rationalen Zahlen $\mathbb{Q}$ haben aber diese Eigenschaften ebenfalls. Wir können den Gauß-Algorithmus also genauso über $\mathbb{Q}$ durchführen, vorausgesetzt alle Koeffizienten des Gleichungssystems sind rationale Zahlen. Dann findet man die *rationale* Lösungsmenge des Systems.
Als Mathematiker nimmt man das zum Anlass, gleich alles *axiomatisch* zu formulieren. Das führt zum Begriff des *Körpers*.

2.2.1 Grundbegriffe

Definition 2.2.1. (i) Eine **Gruppe** ist eine Menge G zusammen mit einer zweistelligen Verknüpfung

$$*\colon G \times G \to G$$

die folgende Bedingungen erfüllt:

(Assoziativgesetz)	$\forall a, b, c$	$(a * b) * c = a * (b * c)$
(neutrales Element)	$\exists e \, \forall a$	$e * a = a * e = a$
(inverse Elemente)	$\forall a \, \exists b$	$a * b = b * a = e.$

(ii) Gilt zusätzlich

(Kommutativgesetz) $\forall a, b \quad a * b = b * a$

so nennt man G eine **kommutative** oder **abelsche** Gruppe. $\triangle$

Beispiel 2.2.2. (i) Beispiele für abelsche Gruppen sind $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$ mit der bekannten Addition $+$ und $e = 0$. $\mathbb{N}$ ist keine Gruppe bezüglich $+$, da nicht jedes Element ein inverses Element besitzt.
(ii) $\mathbb{Q} \setminus \{0\}$ und $\mathbb{R} \setminus \{0\}$ sind abelsche Gruppen bezüglich der bekannten Multiplikation $\cdot$ mit $e = 1$. Für $\mathbb{Z} \setminus \{0\}$ gilt das nicht, da nicht jedes Element ein inverses Element besitzt.
(iii) Für $n \in \mathbb{N}$ sei

$$\mathbb{Z}/n\mathbb{Z} := \{0, 1, \ldots, n-1\}$$

versehen mit der Addition wie auf einer Uhr. Wir addieren also wie gewöhnlich, identifizieren aber n mit 0, $n+1$ mit 1 usw... Auf diese Weise erhalten wir eine abelsche Gruppe mit n Elementen.
(iv) Für jede nichtleere Menge M bilden die bijektiven Abbildungen von M nach M eine Gruppe $S(M)$ bezüglich der Hintereinanderausführung $\circ$ von Funktionen. Das neutrale Element ist dabei die identische Abbildung id_M. Falls M mindestens 3 Elemente hat, ist $S(M)$ nicht abelsch. $S(M)$ wird als **Permutationsgruppe** oder **Symmetriegruppe** von M bezeichnet. $\triangle$

Bemerkung 2.2.3. (i) Das Assoziativgesetz erlaubt uns, Klammern bei komplizierteren Ausdrücken meistens ganz wegzulassen.
(ii) Die Definition der inversen Elemente nimmt Bezug auf das neutrale Element e. Das muss also zunächst bekannt sein, es ist eindeutig bestimmt. Falls nämlich $e, e' \in G$ beides neutrale Elemente sind, so gilt

$$e = e * e' = e'$$

wobei wir für die erste Gleichung die Neutralität von e' und für die zweite die Neutralität von e verwendet haben. Auch das inverse Element zu einem festen Element ist eindeutig bestimmt. Seien etwa b, b' beide invers zu a. Multiplizieren wir nun die Gleichung $a * b = e$ von links mit b' erhalten wir

$$\underbrace{b' * a}_{e} * b = b' * e = b',$$

also $b = b'$.
(iv) In abelschen Gruppen wird die Verknüpfung oft mit $+$ bezeichnet und das neutrale Element mit 0. Das zu a inverse Element wird dann mit $-a$ bezeichnet.

In nicht-abelschen Gruppen verwenden wir oft $\cdot$ und 1 oder $\circ$ und id für die Verknüpfung und das neutrale Element. Das zu a inverse Element wird dann mit a^{-1} bezeichnet. Oft lassen wir $\cdot$ auch einfach weg, schreiben also ab statt $a \cdot b$. △

Definition 2.2.4. (i) Ein **Ring** ist eine Menge R zusammen mit zwei Verknüpfungen

$$\begin{aligned} +\colon R \times R &\to R \\ \cdot\colon R \times R &\to R, \end{aligned}$$

genannt Addition und Multiplikation, sodass gilt:

- R ist bezüglich Addition eine abelsche Gruppe (das neutrale Element bezeichnen wir mit 0).
- Die Multiplikation ist assoziativ und besitzt ein neutrales Element (genannt 1). Es gelte stets $1 \neq 0$.
- Addition und Multiplikation erfüllen das Distributivgesetz, d.h. für alle $a, b, c \in R$ gilt

$$a \cdot (b + c) = (a \cdot b) + (a \cdot c) \text{ und } (b + c) \cdot a = (b \cdot a) + (c \cdot a).$$

(ii) Der Ring R heißt **kommutativ**, falls die Multiplikation kommutativ ist.
(iii) Die Menge

$$R^\times := \{a \in R \mid \exists b \in R\colon\ ab = ba = 1\}$$

heißt Menge der **Einheiten** von R. △

Beispiel 2.2.5. (i) Die einfachsten Beispiele von kommutativen Ringen sind $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$, jeweils mit der bekannten Addition und Multiplikation. Es gilt

$$\mathbb{Z}^\times = \{-1, 1\},\ \mathbb{Q}^\times = \mathbb{Q} \setminus \{0\},\ \mathbb{R}^\times = \mathbb{R} \setminus \{0\}.$$

(ii) Sei M eine nichtleere Menge. Dann bildet die Menge $\mathcal{F}(M, \mathbb{R})$ aller reellwertigen Abbildungen auf M einen kommutativen Ring bezüglich punktweise definierter Addition und Multiplikation:

$$\begin{aligned} (f + g)(m) &:= f(m) + g(m) \\ (f \cdot g)(m) &:= f(m) \cdot g(m). \end{aligned}$$

Hier kann statt $\mathbb{R}$ auch ein beliebiger anderer Ring R gewählt werden (wobei $\mathcal{F}(M, R)$ dann nicht mehr unbedingt kommutativ ist). △

Bemerkung 2.2.6. (i) Wie in Gruppen sind auch in Ringen die beiden neutralen Elemente $0, 1$ eindeutig bestimmt. Die inversen Elemente bezüglich $+$ sind ebenfalls eindeutig bestimmt. Für $a \in R^\times$ ist auch das multiplikativ inverse Element eindeutig bestimmt, wir bezeichnen es mit a^{-1}.
(ii) In Ringen gelten einige einfache Rechenregeln, wie zum Beispiel

$$0 \cdot a = 0$$

oder

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

für alle $a, b \in R$. Obwohl sie nicht zu den Axiomen gehören, kann man sie relativ leicht beweisen.
(iii) Um die Anzahl von Klammern zu verringern, vereinbaren wir, dass Multiplikation stärker bindet als Addition. So schreiben wir etwa

$$a \cdot b + c \cdot d$$

anstelle von

$$(a \cdot b) + (c \cdot d).$$

Wiederum lassen wir $\cdot$ oft einfach weg, schreiben also ab anstelle von $a \cdot b$. $\triangle$

Lemma 2.2.7. *Sei R ein Ring. Dann ist $R^\times$ eine Gruppe bezüglich $\cdot$.*

Beweis. Wir zeigen zunächst, dass $R^\times$ überhaupt abgeschlossen unter $\cdot$ ist. Seien dazu also $a, b \in R^\times$. Es existieren also $a^{-1}, b^{-1} \in R$ und es gilt

$$(ab)(b^{-1}a^{-1}) = a(bb^{-1})a^{-1} = a1a^{-1} = aa^{-1} = 1.$$

Genauso zeigt man $(b^{-1}a^{-1})(ab) = 1$. Also ist ab ebenfalls invertierbar in R, d.h. $ab \in R^\times$.
Die Multiplikation ist nun bereits aufgrund der Ringaxiome assoziativ. Offensichtlich gilt außerdem $1 \in R^\times$, also besitzt $R^\times$ ein neutrales Element bezüglich der Multiplikation. Wegen

$$aa^{-1} = a^{-1}a = 1$$

besitzt mit a auch a^{-1} ein inverses Element, nämlich a. Damit gilt $a^{-1} \in R^\times$ und a hat also in $R^\times$ ein inverses Element bezüglich $\cdot$. $\square$

Bemerkung 2.2.8. Der letzte Beweis zeigt, dass für $a, b \in R^\times$ stets gilt

$$(a^{-1})^{-1} = a \text{ sowie } (ab)^{-1} = b^{-1}a^{-1}. \qquad \triangle$$

Definition 2.2.9. Ein **Körper** ist ein kommutativer Ring K mit

$$K^\times = K \setminus \{0\}. \qquad \triangle$$

Beispiel 2.2.10. $\mathbb{Q}$ und $\mathbb{R}$ sind Körper. $\triangle$

Bemerkung 2.2.11. Wie bereits erwähnt, kann man den Algorithmus von Gauß für lineare Gleichungssysteme mit Koeffizienten aus einem beliebigen Körper durchführen, genau wie im letzten Abschnitt beschrieben, ohne die Lösungsmenge zu verändern. Man beachte, dass hier (per Definition)

$$L(A, b) \subseteq K^n$$

gilt. Es gelten Assoziativ-, Kommutativ- und Distributivgesetz für $+$ und $\cdot$. Die Bedingungung $K^\times = K \setminus \{0\}$ und die Existenz von additiv inversen Elementen erlaubt es uns, jeden Pivot so zu skalieren, dass man damit Einträge unterhalb eliminieren kann. Ist der Pivot etwa $a \neq 0$ und ein Eintrag unterhalb lautet b, so muss man die Pivotzeile mit $-ba^{-1}$ multiplizieren und nach unten addieren. An der Stelle von b steht dann

$$(-ba^{-1}) \cdot a + b = -b \cdot 1 + b = -b + b = 0.$$

Auch das Ablesen der Lösungsmenge funktioniert genau wie in Bemerkung 2.1.4. Man erhält eine Parametrisierung

$$\varphi\colon K^{n-r} \to L(A, b).$$

Über einem Ring lässt sich der Algorithmus jedoch im Allgemeinen nicht durchführen. Ist ein Pivot über $\mathbb{Z}$ beispielsweise 2, so kann man eine etwaige darunterstehende 1 damit nicht eliminieren (ohne den Zahlbereich $\mathbb{Z}$ zu verlassen).

2.2.2 Die komplexen Zahlen ★

Die komplexen Zahlen $\mathbb{C}$ bilden einen Körper, der eine Erweiterung von $\mathbb{R}$ ist. Wem das seltsam vorkommt, der rufe sich nochmal in Erinnerung, dass bereits die reellen Zahlen ein komplexes theoretisches Gebilde sind, das keineswegs natürlich gegeben ist. Es entstand aus $\mathbb{Q}$ in dem Versuch, gewisse Defizite beim Lösen von Gleichungen zu beheben. Es gibt nun keinen Grund, bei $\mathbb{R}$ mit solchen Konstruktionen aufzuhören. Es hat ja auch $\mathbb{R}$ noch gewisse Defizite, zum Beispiel besitzt die Gleichung $x^2 + 1 = 0$ dort keine Lösung.

Konstruktion 2.2.12. Als Menge sind die komplexen Zahlen einfach die reelle Ebene $\mathbb{R}^2$:

$$\mathbb{C} = \mathbb{R} \times \mathbb{R} = \{(a, b) \mid a, b \in \mathbb{R}\}.$$

Wir definieren nun Addition und Multiplikation folgendermaßen:

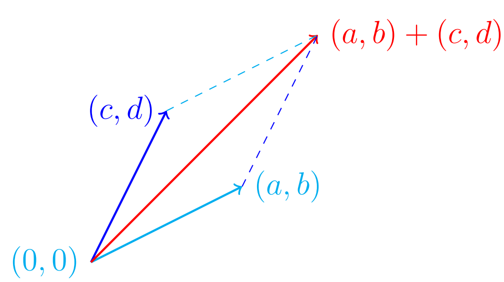
$$(a, b) + (c, d) = (a + c, b + d)$$

$$(a, b) \cdot (c, d) = (ac - bd, ad + bc).$$

Es ist also beispielsweise

$$(3, -2) + (2, 1) = (5, -1), \quad (3, -2) \cdot (2, 1) = (8, -1).$$

Die Addition entspricht einfach dem Aneinanderhängen von Pfeilen im $\mathbb{R}^2$:



Wir werden später sehen, dass auch die Multiplikation eine geometrische Interpretation besitzt. △

Satz 2.2.13. *$\mathbb{C}$ bildet mit $+$ und $\cdot$ einen Körper. Dabei ist $(0, 0)$ das additiv neutrale und $(1, 0)$ das multiplikativ neutrale Element. Das additiv inverse Element zu (a, b) ist $(-a, -b)$. Das multiplikativ inverse Element zu $z = (a, b) \neq (0, 0)$ ist*

$$z^{-1} = \left(\frac{a}{a^2 + b^2}, \frac{-b}{a^2 + b^2} \right).$$

Beweis. Wir rechnen nur einige Aussagen nach, der Rest ist Übungsaufgabe. Zum Beispiel ist

$$(a,b)+(0,0)=(a+0,b+0)=(a,b)$$

und

$$(a,b)\cdot(1,0)=(a\cdot 1-b\cdot 0,a\cdot 0+b\cdot 1)=(a,b).$$

Weiter gilt

$$\begin{aligned}(a,b)\cdot\left(\frac{a}{a^2+b^2},\frac{-b}{a^2+b^2}\right)&=\left(\frac{a^2}{a^2+b^2}-\frac{-b^2}{a^2+b^2},\frac{-ab}{a^2+b^2}+\frac{ab}{a^2+b^2}\right)\\&=\left(\frac{a^2+b^2}{a^2+b^2},0\right)=(1,0).\end{aligned}$$

□

Bemerkung 2.2.14. (i) Wir können $\mathbb{R}$ in $\mathbb{C}$ einbetten, indem wir $a\in\mathbb{R}$ mit $(a,0)\in\mathbb{C}$ identifizieren. Wir identifizieren $\mathbb{R}$ also mit der x-Achse in $\mathbb{R}^2$. Die neuen Verknüpfungen stimmen dann gerade mit den alten überein:

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

Das additiv neutrale Element ist also gerade die alte 0, das multiplikativ neutrale ist die alte 1. In diesem Sinne können wir ab sofort

$$\mathbb{R}\subseteq\mathbb{C}$$

schreiben. Besonders interessant ist nun das Element $(0,1)$, das nicht zu den reellen Zahlen gehört. Wir rechnen

$$(0,1)\cdot(0,1)=(0\cdot 0-1\cdot 1,0\cdot 1+1\cdot 0)=(-1,0)\widehat{=}-1\in\mathbb{R}.$$

Wir haben also eine Quadratwurzel aus -1 gefunden, die selbst allerdings keine reelle Zahl ist.

(ii) Gewöhnlich werden komplexe Zahlen $z=(a,b)$ nicht als Tupel geschrieben, sondern in der Form

$$z=a+bi.$$

Dabei ist

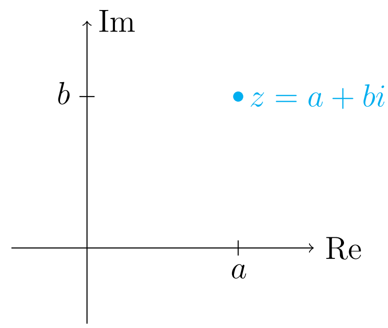
$$i=0+1\cdot i=(0,1)$$

und es gilt

$$i^2=-1.$$

Die Zahl a heißt **Realteil**, und b **Imaginärteil** von z. Wir schreiben auch

$$\mathrm{Re}(z)=a,\quad \mathrm{Im}(z)=b.$$



Die Schreibweise $a + bi$ hat den Vorteil, dass wir Produkte leichter ausrechnen können. Wir müssen uns nicht die Formel merken, sondern dürfen das Distributivgesetz verwenden, und müssen nur $i^2 = -1$ im Kopf behalten:

$$\begin{aligned}(a,b)\cdot(c,d) &= (a+bi)\cdot(c+di) = ac + adi + bci + bdi^2 \\ &= ac - bd + (ad+bc)i = (ac-bd, ad+bc).\end{aligned}$$

In dieser Schreibweise haben wir also

$$(a+bi)^{-1} = \frac{a}{a^2+b^2} - \frac{b}{a^2+b^2}i. \qquad \triangle$$

Beispiel 2.2.15. (i) Wir führen einige Beispielrechnungen mit komplexen Zahlen durch:

$$(3+2i) + (-5+i) = -2+3i$$

$$-(5-2i) = -5+2i$$

$$(5-2i) + (-5+2i) = 5-5+(-2+2)i = 0+0i = 0$$

$$(1+i)\cdot(1-i) = 1 - i + i - i^2 = 1+1 = 2$$

$$(1+i)^{-1} = \frac{1}{2} - \frac{1}{2}i$$

$$(1+i)\cdot\left(\frac{1}{2}-\frac{1}{2}i\right) = \frac{1}{2} - \frac{1}{2}i + \frac{1}{2}i - \frac{1}{2}i^2 = \frac{1}{2} + \frac{1}{2} = 1.$$

(ii) Wir lösen nun die Gleichung

$$(1+i)\cdot z = i$$

in $\mathbb{C}$. Dazu multiplizieren wir mit dem Inversen von $1 + i$ auf beiden Seiten. Das Inverse ist $\frac{1}{2} - \frac{1}{2}i$, wie wir im letzten Beispiel gesehen haben. Es ergibt sich

$$z = \left(\frac{1}{2} - \frac{1}{2}i\right) \cdot i = \frac{1}{2}i - \frac{1}{2}i^2 = \frac{1}{2} + \frac{1}{2}i$$

als Lösung.
(iii) Wir wollen die Gleichung

$$z^2 = -5$$

lösen. Wir wissen schon, dass $i^2 = -1$ gilt. Damit folgt mit $z = \sqrt{5}i$

$$z^2 = \sqrt{5}^2 i^2 = -5. \qquad \triangle$$

Auf den komplexen Zahlen gibt es weitere nützliche Operationen:

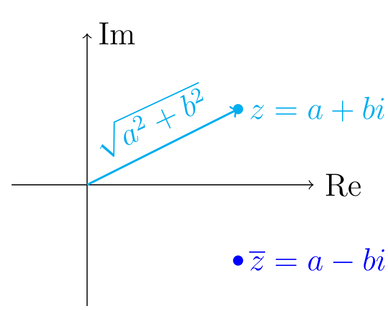
Definition 2.2.16. (i) Für eine komplexe Zahl $z = a + bi = (a, b)$ heißt

$$\overline{z} = a - bi = (a, -b)$$

die **komplex konjugierte Zahl**. Sie entsteht gerade durch Multiplikation des Imaginärteils mit -1. Geometrisch handelt es sich dabei um Spiegelung an der x-Achse.
(ii) Für $z = a + bi \in \mathbb{C}$ definieren wir den **Betrag** als

$$|z| = \sqrt{a^2 + b^2}.$$

Man beachte, dass $a^2 + b^2$ eine positive reelle Zahl ist, aus der wir eine positive reelle Quadratwurzel ziehen können. Der Betrag $|z|$ ist gerade die Länge von z als Vektor in $\mathbb{R}^2$, wie man mit dem Satz des Pythagoras sieht. Auch stimmt der Betrag $|a|$ für reelle Zahlen genau mit dem alten Betrag überein, d.h. $|a| = a$ falls $a \geq 0$ und $|a| = -a$ falls $a < 0$.



$\triangle$

Satz 2.2.17. *Für komplexe Zahlen z, z_1, z_2 gilt:*

(i) $\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}$ *und* $\overline{z_1 \cdot z_2} = \overline{z_1} \cdot \overline{z_2}$

(ii) $\overline{z^{-1}} = (\overline{z})^{-1}$

(iii) $\mathrm{Re}(z) = \frac{1}{2}(z + \overline{z})$ *und* $\mathrm{Im}(z) = -\frac{1}{2}i(z - \overline{z})$.

(iv) $|\mathrm{Re}(z)| \leq |z|$ *und* $|\mathrm{Im}(z)| \leq |z|$

(v) $|\overline{z}| = |z|$.

(vi) $z \cdot \overline{z} = |z|^2$*, also*
$$z^{-1} = \frac{1}{|z|^2}\overline{z}$$
für $z \neq 0$.

(vii) $|z_1 \cdot z_2| = |z_1| \cdot |z_2|$

(viii) $|z_1 + z_2| \leq |z_1| + |z_2|$ *(Dreiecksungleichung)*

Beweis. Die erste Aussage in (i) ist offensichtlich. Für die zweite setze $z_1 = a + bi$, $z_2 = c + di$. Dann ist $z_1 \cdot z_2 = ac - bd + (ad + bc)i$ und

$$\overline{z}_1 \cdot \overline{z}_2 = (a - bi) \cdot (c - di) = ac - bd + (-ad - bc)i = \overline{z_1 \cdot z_2}.$$

(ii) rechnet man entweder von Hand mit der Formel für das Inverse nach, oder man verwendet die folgende elegantere Methode: Es gilt $z \cdot z^{-1} = 1$. Wir konjugieren alles und erhalten mit (i)

$$\overline{z}\overline{z^{-1}} = \overline{z \cdot z^{-1}} = \overline{1} = 1.$$

Also ist $\overline{z^{-1}}$ das (eindeutige) multiplikativ Inverse zu $\overline{z}$. Das ist gerade die Aussage. Für (iii) setzen wir $z = a + bi$ und rechnen

$$z + \overline{z} = a + bi + a - bi = 2a = 2\mathrm{Re}(z),$$

und das ist gerade die erste Aussage. Weiter ist

$$z - \overline{z} = a + bi - (a - bi) = 2bi = 2\mathrm{Im}(z)i.$$

Wenn wir mit $-\frac{1}{2}i$ multiplizieren, erhalten wir rechts genau $\mathrm{Im}(z)$, da $i^2 = -1$.

Für (iv) berechnen wir

$$|\mathrm{Re}(z)| = |a| = \sqrt{a^2} \leq \sqrt{a^2 + b^2} = |z|,$$

und analog für $\mathrm{Im}(z) = b$. (v) ist offensichtlich.
Für (vi) rechnen wir

$$z \cdot \overline{z} = (a + bi) \cdot (a - bi) = a^2 - b^2 i^2 = a^2 + b^2 = |z|^2.$$

Für (vii) verwenden wir (vi) und (i):

$$|z_1 z_2|^2 = z_1 z_2 \overline{z_1 z_2} = z_1 \overline{z_1} z_2 \overline{z_2} = |z_1|^2 |z_2|^2.$$

Wenn wir auf beiden Seiten die Quadratwurzel ziehen, erhalten wir das Ergebnis.
Für $(viii)$ schließlich rechnen wir:

$$\begin{aligned} |z_1 + z_2|^2 &= (z_1 + z_2)\overline{(z_1 + z_2)} = z_1 \overline{z}_1 + z_1 \overline{z}_2 + z_2 \overline{z}_1 + z_2 \overline{z}_2 \\ &= |z_1|^2 + |z_2|^2 + z_1 \overline{z}_2 + \overline{z_1 \overline{z}_2} \\ &= |z_1|^2 + |z_2|^2 + 2\mathrm{Re}(z_1 \overline{z}_2). \end{aligned}$$

Dabei haben wir wieder (vi), (i) und (iii) verwendet. Weiter ist

$$\begin{aligned} (|z_1| + |z_2|)^2 &= |z_1|^2 + 2|z_1||z_2| + |z_2|^2 \\ &= |z_1|^2 + 2|z_1 z_2| + |z_2|^2, \end{aligned}$$

wobei wir (vii) verwenden. Als Differenz erhalten wir

$$(|z_1| + |z_2|)^2 - |z_1 + z_2|^2 = 2\left(|z_1 z_2| - \mathrm{Re}(z_1 \overline{z}_2)\right). \tag{2.2}$$

Da aber

$$|\mathrm{Re}(z_1 \overline{z}_2)| \leq |z_1 \overline{z}_2| = |z_1||\overline{z}_2| = |z_1||z_2| = |z_1 z_2|$$

nach (iv), (vii) und (v) gilt, ist (2.2) offensichtlich nichtnegativ. Damit gilt

$$|z_1 + z_2|^2 \leq (|z_1| + |z_2|)^2$$

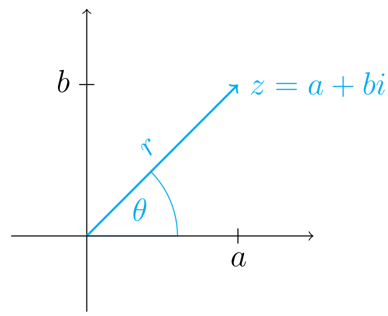
und nach Ziehen der Quadratwurzel ist das die Dreiecksungleichung. □

Nun wollen wir die Darstellung komplexer Zahlen in sogenannten *Polarkoordinaten* untersuchen. Wir erhalten damit eine geometrische Interpretation der Multiplikation. Bisher haben wir komplexe Zahlen z in kartesischen Koordinaten angegeben, d.h. in der Form

$$z = a + bi = (a, b).$$

Man kann die Punkte der Ebene $\mathbb{R}^2$ aber auch in *Polarkoordinaten* angeben.

Definition 2.2.18. Jeder Punkt $z \in \mathbb{C}$ ist eindeutig bestimmt durch seinen Winkel θ (gegen den Urzeigersinn von der positiven x-Achse aus gemessen), und seinen Abstand r zum Ursprung.



Dabei ist $\theta \in [0, 2\pi)$ und $r \geq 0$. Die Darstellung

$$z = (r, \theta)$$

nennt man **Polarkoordinatendarstellung** von z. △

Bemerkung 2.2.19. (i) Um Missverständisse zu vermeiden werden wir ab jetzt die kartesischen Koordinaten immer in der Form $z = a + bi$ angeben, und die Tupelschreibweise $z = (r, \theta)$ bleibt den Polarkoordinaten vorbehalten.
(ii) Für $z = a + bi = (r, \theta)$ gilt

$$r \cdot \sin(\theta) = b, \quad r \cdot \cos(\theta) = a, \quad r^2 = a^2 + b^2 = |z|^2.$$

Damit kann man die beiden Darstellungstypen ineinander umrechnen. Sind zum Beispiel die kartesischen Koordinaten a und b gegeben, erhält man

$$r = |z| = \sqrt{a^2 + b^2}, \quad \theta \in \sin^{-1}\left(\frac{b}{r}\right) \cap \cos^{-1}\left(\frac{a}{r}\right) \cap [0, 2\pi).$$

Sind hingegen die Polarkoordinaten r und θ gegeben, bekommt man direkt

$$a = r \cdot \cos(\theta), \quad b = r \cdot \sin(\theta).$$ △

Beispiel 2.2.20. (i) Sei $z = 1 + i$ in kartesischen Koordinaten gegeben. Wir berechnen $r = |z| = \sqrt{1 + 1} = \sqrt{2}$ sowie

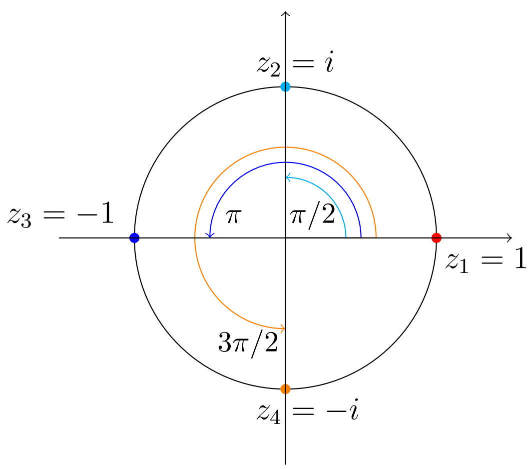
$$\theta = \arcsin\left(\frac{1}{\sqrt{2}}\right) = \pi/4.$$

Also ist

$$z = (\sqrt{2}, \pi/4)$$

die Darstellung in Polarkoordinaten.
(ii) Für $z_1 = 1, z_2 = i, z_3 = -1, z_4 = -i$ erhält man analog die Polarkoordinaten

$$z_1 = (1, 0),\ z_2 = (1, \pi/2),\ z_3 = (1, \pi),\ z_4 = (1, 3\pi/2).$$



(iii) Sei $z = (1, 3\pi/4)$ in Polarkoordinaten gegeben. Wir berechnen die kartesischen Koordinaten:

$$a = 1 \cdot \cos(3\pi/4) = -\frac{1}{\sqrt{2}}, \quad b = 1 \cdot \sin(3\pi/4) = \frac{1}{\sqrt{2}}. \qquad \triangle$$

Wie schon erwähnt liegt einer der Hauptvorteile der Polarkoordinaten in der besonders schönen Formel für die Multiplikation. Sie erlaubt ein unmittelbares geometrisches Verständnis:

Satz 2.2.21. *Die Multiplikation zweier komplexer Zahlen erfüllt in Polarkoordinaten die folgende Gleichung:*

$$(r, \theta) \cdot (s, \eta) = (r \cdot s, \theta + \eta).$$

Beweis. Sei $z_1 = (r, \theta)$ und $z_2 = (s, \eta)$. Wir rechnen zunächst in kartesische Koordinaten um:

$$z_1 = r\cos(\theta) + r\sin(\theta)i, \quad z_2 = s\cos(\eta) + s\sin(\eta)i.$$

Nun können wir die Multiplikation durchführen:

$$\begin{aligned} z_1 z_2 = {} & rs\cos(\theta)\cos(\eta) - rs\sin(\theta)\sin(\eta) \\ & + (rs\cos(\theta)\sin(\eta) + rs\sin(\theta)\cos(\eta))\, i. \end{aligned}$$

Es gilt also

$$\operatorname{Re}(z_1 z_2) = rs(\cos(\theta)\cos(\eta) - \sin(\theta)\sin(\eta)) = rs\cos(\theta + \eta)$$

und

$$\operatorname{Im}(z_1 z_2) = rs(\cos(\theta)\sin(\eta) + \sin(\theta)\cos(\eta)) = rs\sin(\theta + \eta).$$

Dabei verwenden wir die aus der Analysis als bekannt vorausgesetzten Additionstheoreme für Sinus und Cosinus. Diese Koordinaten rechnen wir nun wieder in Polarkoordinaten um. Wir erhalten als Betrag von $z_1 z_2$

$$|z_1 z_2| = \sqrt{(rs\cos(\theta + \eta))^2 + (rs\sin(\theta + \eta)^2} = \sqrt{(rs)^2} = rs,$$

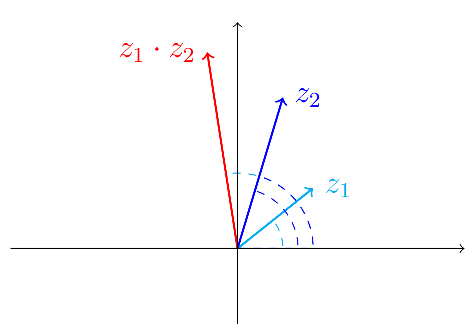
wobei wir $\sin(\varphi)^2 + \cos(\varphi)^2 = 1$ für alle φ und $r, s \geq 0$ verwenden. Für den Winkel von $z_1 z_1$ bekommen wir

$$\arcsin\left(\frac{rs\sin(\theta + \eta)}{rs}\right) = \arcsin(\sin(\theta + \eta)) = \theta + \eta.$$

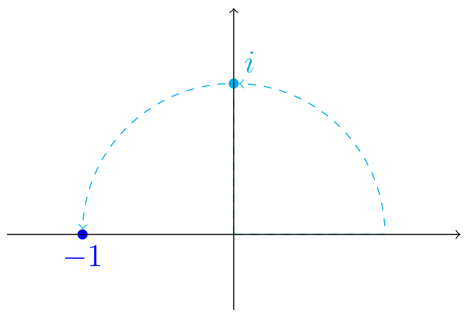
Genau das war zu zeigen. □

Bemerkung 2.2.22. Die geometrische Interpretation der Multiplikation ist also einfach:

Man addiert die Winkel der Elemente, und multipliziert ihre Längen.



Auf diese Weise sehen wir auch sofort, warum $i^2 = -1$ gilt:



△

In den reellen Zahlen kann man im Allgemeinen nicht beliebige Wurzeln aus Elementen ziehen. Wir wissen ja, dass -1 keine reelle Quadratwurzel besitzt und das war gerade die Motivation zur Konstruktion von $\mathbb{C}$. Aber auch wenn es Wurzeln gibt, gibt es eventuell nicht so viele, wie man vielleicht erwarten könnte. Wenn wir uns beispielsweise für die vierten Wurzeln von 1 interessieren, also für die Lösungen der Gleichung

$$x^4 - 1 = 0,$$

so finden wir in $\mathbb{R}$ gerade 1 und -1. Die Gleichung ist aber gerade vom Grad 4, und wir könnten deshalb sogar bis zu 4 verschiedene Lösungen erwarten. In der Tat sind ja i und $-i$ weitere Lösungen der Gleichung, die aber echt komplex sind.

Satz 2.2.23 (Existenz von Einheitswurzeln)**.** *Sei $n \geq 1$ eine positive ganze Zahl. Dann gibt es genau n verschiedene Zahlen $1 = \zeta_0, \zeta_1, \ldots, \zeta_{n-1} \in \mathbb{C}$ mit*

$$\zeta_j^n = 1$$

für alle $j = 0, \ldots, n-1$.

Beweis. Wir geben die ζ_j in Polarkoordinaten an:

$$\zeta_j := (1, 2\pi j/n).$$

Die ζ_j sind offensichtlich paarweise verschieden, da sie die paarweise verschiedenen Winkel

$$0,\ 2\pi/n,\ 4\pi/n,\ 6\pi/n,\ \ldots,\ 2\pi(n-1)/n$$

besitzen. Auch $\zeta_0 = 1$ ist klar. Mit Satz 2.2.21 gilt

$$\zeta_j^n = (1^n, n \cdot 2\pi j/n) = (1, 2\pi j) = 1,$$

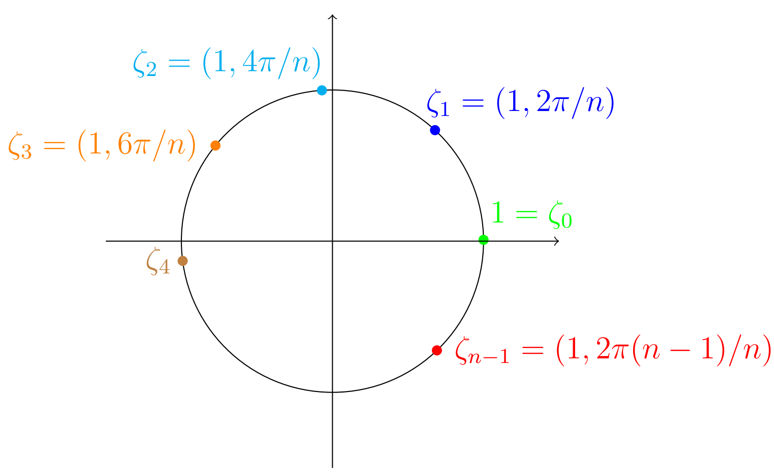
da Vielfache von 2π ja gerade dem Winkel 0 entsprechen.
Es bleibt noch zu zeigen, dass es keine weiteren n-ten Wurzeln aus 1 gibt. Falls $\zeta = (r, \varphi)$ mit $r \geq 0$ und $\varphi \in [0, 2\pi)$ aber

$$1 = \zeta^n = (r^n, n\varphi)$$

erfüllt, so muss $r^n = 1$ und $n\varphi = 2\pi j$ für ein $0 \leq j \leq n-1$ gelten. Daraus folgt $r = 1$ und $\varphi = 2\pi j/n$, also $\zeta = \zeta_j$. □

Definition 2.2.24. Die ζ_j aus dem letzten Satz werden **n-te Einheitswurzeln** genannt. △

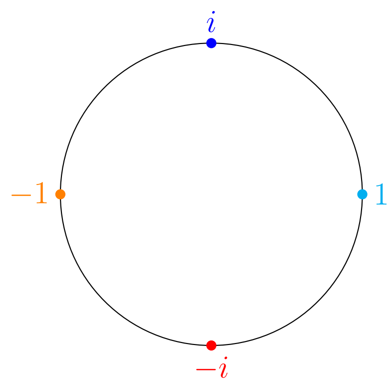
Bemerkung 2.2.25. Wir erhalten die n-ten Einheitswurzeln also gerade durch die Einteilung des Einheitskreises in $\mathbb{C}$ in n gleiche Stücke:



Wenn man ζ_1 mit sich selbst mutlipliziert, wandert es einen Schritt weiter zu ζ_2. Die dritte Potenz ist ζ_3, usw. In der n-ten Potenz ist man gerade einmal durch den Kreis gelaufen, und bei 1 gelandet.
Wenn man ζ_2 quadriert erhält man ζ_4. Aus der Formel $\zeta_2^n = (1, 4\pi)$ sieht man, dass man nach der n-ten Potenz gerade *zweimal* durch den Kreis gelaufen ist, und bei der 1 landet.
Allgemein wandert ζ_j beim potenzieren j-mal durch den Kreis und landet nach der n-ten Potenz bei 1. Allerdings kann ζ_j auch schon vorher einmal bei 1 gelandet sein. △

Definition 2.2.26. Eine n-te Einheitswurzel, die *das erste Mal* in der n-ten Potenz 1 ergibt, heißt **primitive n-te Einheitswurzel**. △

Beispiel 2.2.27. (i) Sei $n = 4$. Wir finden die 4-ten Einheitswurzeln $1, i, -1$ und $-i$:



Es gilt

$$1^1 = 1^2 = 1^3 = 1^4 = 1$$

$$i^1 = i,\ i^2 = -1,\ i^3 = -i,\ i^4 = 1$$

$$(-1)^1 = -1,\ (-1)^2 = 1,\ (-1)^3 = -1,\ (-1)^4 = 1$$

$$(-i)^1 = -i,\ (-i)^2 = -1,\ (-i)^3 = i,\ (-i)^4 = 1.$$

Die Zahlen 1 und -1 ergeben also schon in kleinerer Potenz als 4 einmal 1. Sie sind ja auch 2-te Einheitswurzeln. Die Zahlen i und $-i$ werden erst nach dem vierten Potenzieren das erste Mal 1.
(ii) Es sind i und $-i$ primitive 4-te Einheitswurzeln. Die -1 ist keine primitive 4-te Einheitswurzel, aber eine primitive 2-te Einheitswurzel. Die 1 ist eine (die!) primitive erste Einheitswurzel. $\triangle$

Man kann nun sehr einfach auch Wurzeln aus anderen Elementen ziehen:

Satz 2.2.28. *Sei $n \geq 1$ eine positive ganze Zahl und $0 \neq z \in \mathbb{C}$. Dann gibt es genau n verschiedene Zahlen $\omega_1, \ldots, \omega_n \in \mathbb{C}$ mit*

$$\omega_j^n = z$$

für $j = 1, \ldots, n$.

Beweis. Schreibe $z = (r, \varphi)$ in Polarkoordinaten, mit $r > 0$ und $\varphi \in [0, 2\pi)$. Die offensichtlichste n-te Wurzel aus z entsteht, indem man den Winkel durch n teilt und die n-te Wurzel aus dem (positiven) Radius zieht:

$$\omega_1 = (\sqrt[n]{r}, \varphi/n).$$

Seien nun $1 = \zeta_0, \zeta_1, \ldots, \zeta_{n-1}$ die n-ten Einheitswurzeln. Setze

$$\omega_j = \zeta_{j-1} \cdot \omega_1$$

für $j = 2, \ldots, n$. Wir multiplizieren ω_1 also gerade mit den n-ten Einheitswurzeln. Es entstehen dabei weitere n-te Wurzeln aus z:

$$\omega_j^n = \zeta_{j-1}^n \omega_1^n = 1 \cdot z = z.$$

Die ω_j sind außerdem paarweise verschieden, aufgrund der Kürzungsregel (die sich aus der Existenz von multiplikativ inversen Elementen ergibt).
Sei nun $\omega = (s, \psi)$ eine weitere n-te Wurzel aus z, mit $s \geq 0$ und $\psi \in [0, 2\pi)$. Es gilt also

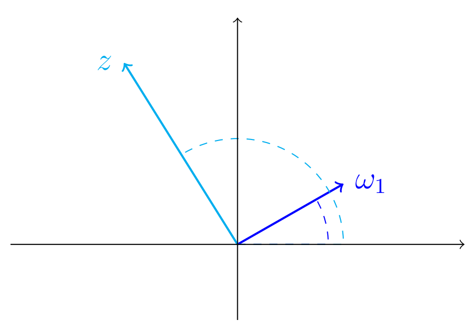
$$(r, \varphi) = z = \omega^n = (s^n, n \cdot \psi),$$

und damit $s = \sqrt[n]{r}$, $\psi = \frac{1}{n}(\varphi + 2\pi j)$ für ein $0 \leq j \leq n-1$. Damit ist

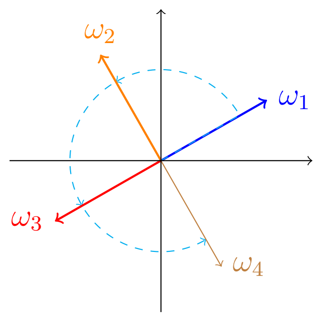
$$\omega = (\sqrt[n]{r}, \varphi/n + 2\pi j/n) = (1, 2\pi j/n) \cdot (\sqrt[n]{r}, \varphi/n) = \zeta_j \cdot \omega_1 = \omega_{j+1}.$$

Also gibt es keine weiteren n-ten Wurzeln außer den ω_j. □

Bemerkung 2.2.29. Wir sehen, wie man die Wurzeln geometrisch zu verstehen hat. Die offensichtlichste n-te Wurzel ω_1 aus z entsteht durch Teilung des Winkels durch n und Ziehen der n-ten Wurzel aus dem Radius (hier zum Beispiel $n = 4$):



Die weiteren n-ten Wurzeln aus z entstehen durch Multiplikation von ω_1 mit den ζ_j, also durch Addition der Vielfachen von $2\pi/n$ zum Winkel:



△

Wir haben gesehen, dass sich in $\mathbb{C}$ sehr viele polynomiale Gleichungen lösen lassen, wie etwa

$$x^n - a = 0,$$

mit $a \in \mathbb{C}$. Der *Fundamentalsatz der Algebra* besagt, dass sich *alle* polynomialen Gleichungen lösen lassen.

Satz 2.2.30 (Fundamentalsatz der Algebra)**.** *Seien $c_0, c_1, \dots, c_n \in \mathbb{C}$, wobei $n \geq 1$ und $c_n \neq 0$. Dann gibt es ein $z \in \mathbb{C}$ mit*

$$c_0 + c_1 z + c_2 z^2 + \dots + c_n z^n = 0.$$

Jede polynomiale Gleichung vom Grad $n \geq 1$ besitzt also eine Lösung.

Beweis. Sei $p = c_0 + c_1 t + c_2 t^2 + \dots + c_n t^n$ das gegebene Polynom, mit $n \geq 1$ und $c_n \neq 0$. Auf jeder abgeschlossenen Kreisschreibe

$$B_r = \{z \in \mathbb{C} \mid |z| \leq r\}$$

nimmt die stetige Funktion

$$|p| \colon z \mapsto |p(z)|$$

ein Minimum an. Wegen

$$p(z) = c_n z^n \left(\frac{c_0}{c_n z^n} + \dots + \frac{c_{n-1}}{c_n z} + 1 \right)$$

für alle $z \neq 0$ sieht man aber auch

$$|p(z)| \to \infty \text{ für } |z| \to \infty.$$

Deshalb gibt es ein globales Minimum $z_0 \in \mathbb{C}$ von $|p|$, es gilt also

$$|p(z_0)| \leq |p(z)| \quad \text{für alle } z \in \mathbb{C}.$$

Nach einer Translation können wir o.B.d.A. $z_0 = 0$ annehmen, also $p(z_0) = c_0$. Angenommen es gilt $0 \neq p(z_0) = c_0$. Sei m die kleinste Zahl > 0 mit $c_m \neq 0$, es gelte also

$$p = c_0 + c_m t^m + t^{m+1} q$$

für ein Polynom q. Wir wählen mit Satz 2.2.28 ein $z_1 \in \mathbb{C}$ mit

$$z_1^m = -\frac{c_0}{c_m}.$$

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

$$p(\lambda z_1) = c_0 \left(1 - \lambda^m + c_0^{-1} \lambda^{m+1} z_1^{m+1} q(\lambda z_1)\right)$$

und für $\lambda \in [0,1]$ also

$$|p(\lambda z_1)| \leq |c_0| \left(1 - \lambda^m + |c_0^{-1} z_1^{m+1} q(\lambda z_1)| \lambda^{m+1}\right) \leq |c_0| \left(1 - \lambda^m + C\lambda^{m+1}\right)$$

für ein genügend großes $C \in \mathbb{R}$. Es ist aber

$$0 < 1 - \lambda^m + C\lambda^{m+1} = 1 + \lambda^m (C\lambda - 1) < 1$$

für genügend kleines $\lambda > 0$. Also gilt

$$|p(\lambda z_1)| < |c_0| = |p(z_0)|$$

für solche λ, ein Widerspruch zur Wahl von z_0. Also gilt $p(z_0) = 0$, die gewünschte Aussage. □

2.2.3 Endliche Körper

Bisher kennen wir die Körper $\mathbb{Q}, \mathbb{R}$ und $\mathbb{C}$. Es gibt aber noch viele mehr. Besonders in Anwendungen in der Informatik spielen *endliche Körper* eine wichtige Rolle (die *Kodierungstheorie* ist beispielsweise größtenteils lineare Algebra über endlichen Körpern). Wir lernen einige wichtige endliche Körper jetzt kennen.

Konstruktion 2.2.31. Sei $n \geq 2$ eine natürliche Zahl. Wir definieren eine Äquivalenzrelation auf $\mathbb{Z}$ durch folgende Setzung:

$$a \sim b \;:\Leftrightarrow\; n \mid (b-a).$$

Zwei Zahlen sind also äquivalent, wenn sie sich um ein Vielfaches von n unterscheiden. Ordnen wir die Zahlen auf einer Uhr mit n Stunden an, so endet jeder volle Umlauf um die Uhr also in derselben Äquivalenzklasse. Wir setzen nun

$$\mathbb{Z}/n\mathbb{Z} := \mathbb{Z}/\sim = \{[a] \mid a \in \mathbb{Z}\} = \{[0],[1],\ldots,[n-1]\}.$$

Man beachte, dass wir $\mathbb{Z}/n\mathbb{Z}$ in Beispiel 2.2.2 bereits etwas salopp (als Gruppe mit $+$) eingeführt haben. Hier folgt also nun die exakte Definition. Wir definieren Addition und Multiplikation auf $\mathbb{Z}/n\mathbb{Z}$:

$$\begin{aligned}[a]+[b] &:= [a+b]\\ [a]\cdot[b] &:= [a\cdot b].\end{aligned}$$

Da die Definition auf die Vertreter der Äquivalenzklassen bezug nimmt, diese jedoch nicht eindeutig bestimmt sind, ist eine *Wohldefiniertheit* zu zeigen.
Für $+$ geht das so: Sei $[a]=[a']$ und $[b]=[b']$. Es gibt also $c,d \in \mathbb{Z}$ mit

$$a-a' = cn \text{ und } b-b' = dn.$$

Daraus folgt

$$(a+b)-(a'+b') = (a-a')+(b-b') = cn+dn = (c+d)n,$$

also $n \mid (a+b)-(a'+b')$, also $[a+b]=[a'+b']$. Die Wohldefiniertheit der Multiplikation zeigt man ähnlich. △

Satz 2.2.32. *Sei $n \geq 2$.*
(i) Es ist $\mathbb{Z}/n\mathbb{Z}$ ein kommutativer Ring mit genau n Elementen.
(ii) Es ist $\mathbb{Z}/n\mathbb{Z}$ ist genau dann ein Körper, wenn n eine Primzahl ist.

Beweis. (i) überträgt sich direkt von den entsprechenden Eigenschaften von $\mathbb{Z}$. Zum Beispiel gilt

$$[a]+[0] = [a+0] = [a],$$

also ist $[0]$ das neutrale Element bezüglich $+$. Weiter gilt

$$[-a]+[a] = [-a+a] = [0],$$

also besitzt $[a]$ das additiv inverse Element $[-a]$.
Für (ii) nehmen wir zunächst an, dass n eine Primzahl ist. Wir müssen die Existenz von multiplikativ inversen Elementen zeigen. Für $1 \leq a \leq n-1$ betrachten wir dann die (wohldefinierte) Abbildung

$$\begin{aligned} m_a \colon \mathbb{Z}/n\mathbb{Z} &\to \mathbb{Z}/n\mathbb{Z} \\ [b] &\mapsto [ab]. \end{aligned}$$

Wir zeigen dass m_a injektiv ist. Es bedeutet nämlich $m_a([b]) = m_a([b'])$ gerade

$$n \mid (ab - ab') = a(b - b').$$

Der *Primfaktor* n kommt also in der (eindeutigen) Primfaktorzerlegung von $a(b - b')$ in $\mathbb{Z}$ vor. Wegen $1 \leq a \leq n-1$ kann der Primfaktor n in der Primfaktorzerlegung von a nicht auftreten, also muss er in der Primfaktorzerlegung von $b - b'$ auftreten. Das bedeutet $n \mid (b - b')$, also $[b] = [b']$.
Eine injektive Abbildung von einer endlichen Menge in sich selbst muss aber offensichtlich surjektiv sein. Damit existiert ein $[b] \in \mathbb{Z}/n\mathbb{Z}$ mit

$$[1] = m_a([b]) = [ab] = [a] \cdot [b].$$

Da $[1]$ das multiplikativ neutrale Element ist, besitzt $[a]$ also ein multiplikativ Inverses. Es ist aber jedes Element $\neq 0$ aus $\mathbb{Z}/n\mathbb{Z}$ von der Gestalt $[a]$ mit $1 \leq a \leq n-1$. Also ist $\mathbb{Z}/n\mathbb{Z}$ ein Körper.
Sei nun schließlich n keine Primzahl, wir können also n schreiben als $n = ab$ mit $1 \leq a, b \leq n-1$. Wir nehmen an dass $\mathbb{Z}/n\mathbb{Z}$ trotzdem ein Körper ist. In $\mathbb{Z}/n\mathbb{Z}$ gilt nun

$$[a] \cdot [b] = [a \cdot b] = [n] = [0].$$

Es muss in $\mathbb{Z}/n\mathbb{Z}$ nun $[a]^{-1}$ existieren, denn $0 \neq [a]$. Daraus folgt

$$[b] = [1] \cdot [b] = \big([a]^{-1} \cdot [a]\big) \cdot [b] = [a]^{-1} \cdot ([a] \cdot [b]) = [a]^{-1} \cdot [0] = [0],$$

was ein Widerspruch zu $1 \leq b \leq n-1$ ist. Also ist $\mathbb{Z}/n\mathbb{Z}$ kein Körper. $\square$

Bemerkung/Beispiel 2.2.33. (i) In Zunkunft schreiben wir für Elemente aus $\mathbb{Z}/n\mathbb{Z}$ oft a statt $[a]$. Man muss sich aber im Klaren sein, dass die Elemente $a \in \mathbb{Z}$ und $a \in \mathbb{Z}/n\mathbb{Z}$ grundlegend unterschiedlich sind.

(ii) Addition und Multiplikation in $\mathbb{Z}/n\mathbb{Z}$ kann man für kleine n gut durch eine *Verknüpfungstabelle* angeben:

$$\mathbb{Z}/2\mathbb{Z} \qquad \begin{array}{c|cc} + & 0 & 1 \\ \hline 0 & 0 & 1 \\ 1 & 1 & 0 \end{array} \qquad \begin{array}{c|cc} \cdot & 0 & 1 \\ \hline 0 & 0 & 0 \\ 1 & 0 & 1 \end{array}$$

$$\mathbb{Z}/3\mathbb{Z} \qquad \begin{array}{c|ccc} + & 0 & 1 & 2 \\ \hline 0 & 0 & 1 & 2 \\ 1 & 1 & 2 & 0 \\ 2 & 2 & 0 & 1 \end{array} \qquad \begin{array}{c|ccc} \cdot & 0 & 1 & 2 \\ \hline 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 2 \\ 2 & 0 & 2 & 1 \end{array}$$

$$\mathbb{Z}/4\mathbb{Z} \qquad \begin{array}{c|cccc} + & 0 & 1 & 2 & 3 \\ \hline 0 & 0 & 1 & 2 & 3 \\ 1 & 1 & 2 & 3 & 0 \\ 2 & 2 & 3 & 0 & 1 \\ 3 & 3 & 0 & 1 & 2 \end{array} \qquad \begin{array}{c|cccc} \cdot & 0 & 1 & 2 & 3 \\ \hline 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 2 & 3 \\ 2 & 0 & 2 & 0 & 2 \\ 3 & 0 & 3 & 2 & 1 \end{array}$$

(iii) Zum Schluss lösen wir noch ein lineares Gleichungssystem über $\mathbb{Z}/2\mathbb{Z}$. Das klingt vielleicht erst kompliziert, ist aber in Wirklichkeit sogar deutlich leichter als über $\mathbb{Q}, \mathbb{R}$ oder $\mathbb{C}$. Es gibt nämlich nur die beiden möglichen Koeffizienten 0 und 1. Wegen $1 + 1 = 0$ ist auch die Elimination extrem einfach. Wir betrachten das folgende System linearer Gleichungen über $\mathbb{Z}/2\mathbb{Z}$:

$$\begin{aligned} x_1 \qquad\quad + x_3 + x_4 &= 1 \\ x_3 \qquad &= 0 \\ x_1 + x_2 \qquad\quad + x_4 &= 1 \end{aligned}$$

Wir bringen zunächst die Koeffizientenmatrix auf Zeilenstufenform:

$$(A, b) = \left(\begin{array}{cccc|c} 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 1 \end{array}\right) \overset{Z_1+Z_3}{\rightsquigarrow} \left(\begin{array}{cccc|c} 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \end{array}\right)$$
$$\overset{Z_2 \leftrightsquigarrow Z_3}{\rightsquigarrow} \left(\begin{array}{cccc|c} 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{array}\right).$$

Wir können also für x_4 einen belieben Wert aus $\mathbb{Z}/2\mathbb{Z}$ wählen und die Werte für die anderen Variablen damit eindeutig bestimmen. Es ergibt sich folgende Lösungsmenge

$$L(A,b) = \{(1,0,0,0),(0,0,0,1)\} \subseteq (\mathbb{Z}/2\mathbb{Z})^4 . \qquad \triangle$$

2.3 Matrixrechnung

Rechnen mit Matrizen erlaubt uns eine konzeptuell deutlich einfachere Definition und Behandlung von linearen Gleichungssystemen.

Definition 2.3.1. Sei R eine Menge und $m, n \in \mathbb{N}$.
(i) Eine $m \times n$-**Matrix über** R ist eine Abbildung

$$M\colon \{1,\ldots,m\} \times \{1,\ldots,n\} \to R.$$

Gewöhnlich schreibt man für den Wert $M(i,j)$ einfach m_{ij} und alle Werte gemeinsam in Form eines rechteckigen Schemas (wie früher bereits verwendet):

$$M = (m_{ij})_{i=1,\ldots,m;j=1,\ldots,n} = \begin{pmatrix} m_{11} & m_{12} & \cdots & m_{1n} \\ \vdots & & & \vdots \\ m_{m1} & m_{m2} & \cdots & m_{mn} \end{pmatrix}.$$

Laut unserer Konvention bezeichnet der erste Index also immer die Zeile, der zweite die Spalte der Matrix.
(ii) Eine $m \times 1$-Matrix wird auch **Spaltenvektor**, eine $1 \times n$-Matrix **Zeilenvektor** genannt. In Zeilen- oder Spaltenvektoren verwenden wir oft nur einen Index, schreiben einen Spaltenvektor also zum Beispiel als

$$v = \begin{pmatrix} v_1 \\ \vdots \\ v_m \end{pmatrix}.$$

(iii) Die Menge aller $m \times n$-Matrizen über R bezeichnen wir mit

$$\mathrm{Mat}_{m,n}(R).$$

Im Fall $m = n$ schreiben wir statt $\mathrm{Mat}_{m,m}(R)$ auch $\mathrm{Mat}_m(R)$. Statt $\mathrm{Mat}_{m,1}(R)$ schreiben wir auch R^m. Elemente des R^m sind bei uns also ab sofort stets als Spaltenvektoren zu verstehen.

(iv) Für $M = (m_{ij})_{i=1,\ldots,m;j=1,\ldots,n} \in \mathrm{Mat}_{m,n}(R)$ bezeichnen wir mit

$$M^t := (m_{ji})_{i=1,\ldots n, j=1\ldots m} \in \mathrm{Mat}_{n,m}(R)$$

die sogenannte **transponierte Matrix**. Sie entsteht aus M durch Vertauschung der Indizes, bzw. durch Spiegelung an der Diagonalen. Die Zeilen von M werden also zu den Spalten von M^t. △

Beispiel 2.3.2. Wir geben einige Matrizen und ihre Transponierten an:

$$M = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \in \mathrm{Mat}_3(\mathbb{N}), \quad M^t = \begin{pmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{pmatrix} \in \mathrm{Mat}_3(\mathbb{N})$$

$$N = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \in \mathrm{Mat}_{2,3}(\mathbb{N}), \quad N^t = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix} \in \mathrm{Mat}_{3,2}(\mathbb{N})$$

$$O = \begin{pmatrix} -1 \\ i \\ \pi \end{pmatrix} \in \mathbb{C}^3, \quad O^t = \begin{pmatrix} -1 & i & \pi \end{pmatrix} \in \mathrm{Mat}_{1,3}(\mathbb{C}).$$ △

Wir wollen nun Matrizen miteinander addieren und multiplizieren. Dazu werden die Einträge auf gewisse Weise miteinander verknüpft. Deshalb sei R ab sofort ein Ring.

Definition 2.3.3. Sei R ein Ring und $m, n, k \in \mathbb{N}$.
(i) Für $M = (m_{ij})_{i,j}, N = (n_{ij})_{i,j} \in \mathrm{Mat}_{m,n}(R)$ definieren wir

$$M + N := (m_{ij} + n_{ij})_{i,j} \in \mathrm{Mat}_{m,n}(R)$$

und erhalten so eine Abbildung

$$+\colon \mathrm{Mat}_{m,n}(R) \times \mathrm{Mat}_{m,n}(R) \to \mathrm{Mat}_{m,n}(R),$$

genannt die **Addition von Matrizen**.
(ii) Für $M = (m_{ij})_{i,j} \in \mathrm{Mat}_{m,k}(R)$ und $N = (n_{ij})_{i,j} \in \mathrm{Mat}_{k,n}(R)$ definieren wir

$$M \cdot N := \left(\sum_{r=1}^{k} m_{ir} n_{rj} \right)_{i,j} \in \mathrm{Mat}_{m,n}(R)$$

und erhalten so eine Abbildung

$$\cdot\colon \mathrm{Mat}_{m,k}(R) \times \mathrm{Mat}_{k,n}(R) \to \mathrm{Mat}_{m,n}(R),$$

gennant die **Matrixmultiplikation**. △

Bemerkung/Beispiel 2.3.4. (i) Addieren kann man Matrizen nur, wenn sie die gleichen Dimensionen haben. Dann ist die Addition einfach eintragsweise definiert, also zum Beispiel

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} + \begin{pmatrix} 6 & 3 \\ 5 & 2 \\ 4 & 1 \end{pmatrix} = \begin{pmatrix} 7 & 5 \\ 8 & 6 \\ 9 & 7 \end{pmatrix}.$$

(ii) Multiplizieren kann man zwei Matrizen nur, wenn die Spaltenanzahl der linken Matrix mit der Zeilenanzahl der rechten Matrix übereinstimmt. Anschaulich funktioniert die Multiplikation dann so: Man nimmt die erste Spalte der rechten Matrix N und legt sie auf die erste Zeile der linken Matrix M. Dann multipliziert man eintragsweise und summiert auf. Das Ergebnis steht an Stelle $(1,1)$ des Produkts. Man legt nun die erste Spalte von N auf die zweite Zeile von M und verfährt analog. Das ergibt die Stelle $(2,1)$ usw... Für die zweite Spalte des Produkts wählt man die zweite Spalte von N...

$$\begin{array}{c} i \rightarrow \begin{pmatrix} & & \\ & \square & \\ & & \end{pmatrix} \\ \uparrow \\ j \end{array} = \begin{array}{c} i \rightarrow \begin{pmatrix} \hline \quad\cdots\quad \\ \hline \end{pmatrix} \\ \\ \\ \end{array} \cdot \begin{array}{c} \begin{pmatrix} & \Bigg| & \vdots & \Bigg| & \end{pmatrix} \\ \uparrow \\ j \end{array}$$

Es gilt also zum Beispiel

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} = \begin{pmatrix} 9 & 12 & 15 \\ 19 & 26 & 33 \\ 29 & 40 & 51 \end{pmatrix}.$$

(iii) Die Matrix

$$I_m := \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \in \mathrm{Mat}_m(R)$$

heißt **Einheitsmatrix** (der Größe m). Für beliebige $M \in \mathrm{Mat}_{m,n}(R)$, $N \in \mathrm{Mat}_{n,m}(R)$ gilt offensichtlich

$$I_m \cdot M = M, \quad N \cdot I_m = N.$$

(iv) Seien K ein Körper, $m, n \in \mathbb{N}$ sowie $A \in \mathrm{Mat}_{m,n}(K)$ und $b \in K^m$. Mit Hilfe der Matrixmultplikation lässt sich die Lösungsmenge $L(A, b)$ des entsprechenden linearen Gleichungssystems ganz leicht definieren:

$$L(A, b) = \{c \in K^n \mid A \cdot c = b\}. \qquad \triangle$$

Satz 2.3.5. *Sei R ein Ring und $m, n \in \mathbb{N}$. Dann gilt:*

(i) $\mathrm{Mat}_{m,n}(R)$ ist eine abelsche Gruppe bezüglich $+$.

(ii) Die Multiplikation von Matrizen ist assoziativ, d.h. es gilt

$$(A \cdot B) \cdot C = A \cdot (B \cdot C)$$

wann immer die Größen der Matrizen diese Multiplikation erlauben.

(iii) Zwischen Addition und Multiplikation gilt das Distributivgesetz, d.h.

$$A \cdot (B + C) = (A \cdot B) + (A \cdot C), \quad (B + C) \cdot A = (B \cdot A) + (C \cdot A)$$

für alle Matrizen der geeigneten Größen.

(iv) $\mathrm{Mat}_m(R)$ ist ein Ring bezüglich $+$ und $\cdot$. Selbst wenn R kommutativ ist, ist es $\mathrm{Mat}_m(R)$ niemals, solange $m \geq 2$.

Beweis. (i) ergibt sich direkt aus den entsprechenden Eigenschaften der additiven Gruppe von R, da die Addition von Matrizen ja einfach eintragsweise definiert ist. Es ist beispielsweise die *Nullmatrix*

$$0_{m,n} := \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{pmatrix} \in \mathrm{Mat}_{m,n}(R)$$

das neutrale Element. Zu $M = (m_{ij})_{i,j} \in \mathrm{Mat}_{m,n}(R)$ ist

$$-M = (-m_{ij})_{i,j}$$

die additiv inverse Matrix. (ii) und (iii) ist Übungsaufgabe. Für (iv) bleibt nach (i),(ii), (iii) und Bemerkung 2.3.4 (iii) nur noch zu zeigen, dass $\mathrm{Mat}_m(R)$ für $m \geq$

2 niemals kommutativ ist. Für $m = 2$ folgt das mit $0 \neq 1$ aus der folgenden Rechnung:

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

Für größere m kann man alle hier auftretenden Matrizen nach unten rechts mit Nullen auffüllen und erhält dasselbe Ergebnis. □

Jetzt können wir sehr elegant beweisen, wie die Lösungsmenge eines inhomogenen linearen Gleichungssystems mit der Lösungsmenge des homogenen Systems zusammenhängt:

Satz 2.3.6. *Sei K ein Körper, $m, n \in \mathbb{N}$ sowie $A \in \mathrm{Mat}_{m,n}(K)$ und $b \in K^m$. Sei $\bar{c} \in L(A, b)$ ein beliebiges Element. Dann gilt*

$$L(A, b) = \{\bar{c} + c \mid c \in L(A, 0)\}.$$

Beweis. Für "$\subseteq$" sei $d \in L(A, b)$ beliebig gewählt. Wir setzen $c := d - \bar{c}$ und berechnen

$$Ac = A(d - \bar{c}) = Ad - A\bar{c} = b - b = 0.$$

Also gilt $c \in L(A, 0)$ und $d = \bar{c} + c$. Für "$\supseteq$" berechnen wir für $c \in L(A, 0)$ direkt

$$A(\bar{c} + c) = A\bar{c} + Ac = b + 0 = b,$$

also $\bar{c} + c \in L(A, b)$. □

Definition 2.3.7. Sei R ein Ring und $m \geq 1$. Dann heißen die Elemente von

$$\mathrm{GL}_m(R) := \mathrm{Mat}_m(R)^\times$$

(über R) invertierbare Matrizen. Eine Matrix $A \in \mathrm{Mat}_m(R)$ ist also invertierbar über R, falls eine Matrix $B \in \mathrm{Mat}_m(R)$ existiert mit

$$A \cdot B = B \cdot A = I_m.$$

Man nennt B dann die **inverse Matrix** zu A (sie ist eindeutig bestimmt) und verwenden die Bezeichnung $B = A^{-1}$. △

Bemerkung/Beispiel 2.3.8. (i) Die Matrix

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \in \mathrm{Mat}_2(\mathbb{Q})$$

ist invertierbar über $\mathbb{Q}$. Es gilt

$$A^{-1} = \begin{pmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{pmatrix} \in \mathrm{Mat}_2(\mathbb{Q}).$$

Über $\mathbb{Z}$ wäre A *nicht* invertierbar. Das sieht man daran, dass in $\mathrm{Mat}_2(\mathbb{Q})$ die inverse Matrix eindeutig ist, sie in unserem Fall aber echt rationale Einträge hat.
(ii) Hat eine Matrix A eine Nullzeile oder -spalte, so ist sie nicht invertierbar. Im Fall einer Nullzeile hat nämlich auch AB eine Nullzeile, für jede Matrix B (geeigneter Größe), im Fall einer Nullspalte stimmt dasselbe für BA. Die Einheitsmatrix hat jedoch keine Nullzeilen oder -spalten.
(iii) Ist die Koeffizientenmatrix A eines linearen Gleichungssystems quadratisch und invertierbar, so besitzt das System *genau eine Lösung*. Es ist $Ac = b$ nämlich äquivalent (durch Multiplikation von links mit A^{-1}) zu $c = A^{-1}b$.
(iv) Ist B eine invertierbare Matrix so gilt

$$L(A, b) = L(BA, Bb)$$

für alle linearen Gleichungssysteme (mit geeigneter Größe der Koeffizientenmatrix). Es bedeutet $c \in L(A, b)$ ja gerade

$$Ac = b$$

und daraus folgt nach Multiplikation von links mit B

$$BAc = Bb,$$

also $c \in L(BA, Bb)$. Die andere Inklusion folgt aus der Invertierbarkeit von B mit demselben Argument. Mit dieser Beobachtung können wir auch den Algorithmus von Gauß besser verstehen, wie wir gleich sehen werden. △

Definition 2.3.9. Sei K ein Körper und $m \in \mathbb{N}$. Eine **Elementarmatrix** (über K) ist eine Matrix aus $\mathrm{Mat}_m(K)$ von einem der folgenden drei Typen:

(I) Für $1 \leq i < j \leq m$

$$P_{ij} = \begin{array}{c} \begin{pmatrix} 1 &&&&&&&&&& \\ & \ddots &&&&&&&&& \\ && 1 &&&&&&&& \\ &&& 0 &&&& 1 &&& \\ &&&& 1 &&&&&& \\ &&&&& \ddots &&&&& \\ &&&&&& 1 &&&& \\ &&& 1 &&&& 0 &&& \\ &&&&&&&& 1 && \\ &&&&&&&&& \ddots & \\ &&&&&&&&&& 1 \end{pmatrix} \\ \begin{array}{ccccccccccc} & & & \uparrow & & & & \uparrow & & & \\ &&& i &&&& j &&& \end{array} \end{array}$$

(II) Für $1 \leq i \leq m$ und $0 \neq \lambda \in K$

$$S_i(\lambda) = \begin{array}{c} \begin{pmatrix} 1 &&&&&& \\ & \ddots &&&&& \\ && 1 &&&& \\ &&& \lambda &&& \\ &&&& 1 && \\ &&&&& \ddots & \\ &&&&&& 1 \end{pmatrix} \\ \begin{array}{ccccccc} & & & \uparrow & & & \\ &&& i &&& \end{array} \end{array}$$

(III) Für $1 \leq i, j \leq m, i \neq j, 0 \neq \lambda \in K$

$$M_j^i(\lambda) = \begin{array}{c} \left(\begin{array}{ccccccc} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & \ddots & & & \\ & & \lambda & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{array}\right) \\ \begin{array}{ccccccc} & & \uparrow & & \uparrow & & \\ & & i & & j & & \end{array} \end{array}$$

$\triangle$

Satz 2.3.10. *Seien K ein Körper, $m, n \in \mathbb{N}$ und $A \in \mathrm{Mat}_{m,n}(K)$. Dann gilt:*

(i)
- *$P_{ij}A$ entsteht aus A durch Vertauschen der i-ten und j-ten Zeile.*
- *$S_i(\lambda)A$ entsteht aus A durch Multiplikation der i-ten Zeile mit λ.*
- *$M_j^i(\lambda)A$ entsteht aus A durch Addition des λ-fachen der i-ten Zeile zur j-ten Zeile.*

(ii) Jede Elementarmatrix über K ist invertierbar über K. Es gilt

$$P_{ij}^{-1} = P_{ij}, \quad S_i(\lambda)^{-1} = S_i(\lambda^{-1}), \quad M_j^i(\lambda)^{-1} = M_j^i(-\lambda).$$

Beweis. (i) ist eine leichte Übung in Matrixmultiplikation. Aus (i) folgt direkt

$$I_m = P_{ij}P_{ij} = S_i(\lambda^{-1})S_i(\lambda) = M_j^i(-\lambda)M_j^i(\lambda).$$

Daraus folgt unmittelbar die Aussage (ii). $\square$

Bemerkung 2.3.11. Wir können nun den Algorithmus von Gauß auch folgendermaßen beschreiben. Jede elementare Zeilenumformung der erweiterten Koeffizientenmatrix entsteht durch Multiplikation mit einer Elementarmatrix von links. Jede dieser Elementarmatrizen ist invertierbar, also auch deren Produkt (vergleiche Lemma 2.2.7). Insgesamt erhalten wir also eine invertierbare Matrix B (das Produkt aller verwendeten Elementarmatrizen) und gehen von der erweiterten Koeffizientenmatrix (A, b) über zur Koeffizientenmatrix

$$B \cdot (A, b) = (BA, Bb).$$

Nach Bemerkung 2.3.8 (iv) ändert sich dabei aber die Lösungsmenge des Gleichungssystems nicht. $\triangle$

Satz 2.3.12. *Sei K ein Körper, $m \in \mathbb{N}$ und $A \in \mathrm{Mat}_m(K)$. Dann sind äquivalent:*

(*i*) $A \in \mathrm{GL}_m(K)$.

(*ii*) *A ist ein Produkt von endlich vielen Elementarmatrizen.*

(*iii*) *A lässt sich durch elementare Zeilenumformungen zur Einheitsmatrix transformieren.*

Beweis. Für (*i*)⇒(*ii*) bringen wir A mit dem Gauß-Algorithmus auf reduzierte Zeilenstufenform A'. Nach Bemerkung 2.3.11 existiert also ein Produkt B von endlich vielen Elementarmatrizen mit $BA = A'$. Da sowohl B als auch A invertierbar sind, ist es auch A'. Damit kann A' keine Nullzeile haben (vergleiche Bemerkung 2.3.8 (*ii*)). Eine quadratische Matrix in reduzierter Zeilenstufenform ohne Nullzeilen muss aber die Einheitsmatrix sein. Es gilt also $BA = I_m$ bzw. $A = B^{-1}$. Ist $B = E_1 \cdots E_n$ eine Darstellung als Produkt von Elementar-matrizen, so gilt

$$A = B^{-1} = E_n^{-1} \cdots E_1^{-1}.$$

Nach Satz 2.3.10 (*ii*) sind Inverse von Elementarmatrizen selbst wieder Elementarmatrizen und das beweist die Aussage.
Für (*ii*)⇒(*iii*) sei $A = E_1 \cdots E_n$ ein Produkt von Elementarmatrizen. Dann gilt

$$E_n^{-1} \cdots E_1^{-1} A = I_m$$

und alle diese Multiplikationen von links entsprechen elementaren Zeilenumformungen.
Für (*iii*)⇒(*i*) finden wir $B \in \mathrm{GL}_m(K)$ mit $BA = I_m$. Durch Multiplikation von links mit B^{-1} erhalten wir $A = B^{-1}$ und als Inverse einer invertierbaren Matrix ist A selbst invertierbar. □

Algorithmus 2.3.13 (Berechnung der inversen Matrix)**.** Sei K ein Körper, $m \in \mathbb{N}$ und $A \in \mathrm{GL}_m(K)$. Mit der folgenden Methode können wir die inverse Matrix A^{-1} berechnen:
Wir betrachten die Matrix $(A \mid I_m) \in \mathrm{Mat}_{m,2m}(K)$ und erzeugen mit elementaren Zeilenumformungen im linken Block die Einheitsmatrix:

$$(A \mid I_m) \rightsquigarrow (I_m \mid B).$$

Nach Satz 2.3.12 ist das möglich. Die Matrix B ist dann gerade die inverse Matrix von A. Es ist ja B gerade das Produkt der Elementarmatrizen, die A zu I_m transformieren: $BA = I_m$. Das bedeutet $B = A^{-1}$. △

Beispiel 2.3.14. Für
$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \in \mathrm{Mat}_2(\mathbb{Q})$$
berechnen wir:
$$\left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 3 & 4 & 0 & 1 \end{array}\right) \overset{Z_2 - 3Z_1}{\rightsquigarrow} \left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & -2 & -3 & 1 \end{array}\right)$$
$$\overset{-\frac{1}{2}Z_2}{\rightsquigarrow} \left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & 1 & \frac{3}{2} & -\frac{1}{2} \end{array}\right)$$
$$\overset{Z_1 - 2Z_2}{\rightsquigarrow} \left(\begin{array}{cc|cc} 1 & 0 & -2 & 1 \\ 0 & 1 & \frac{3}{2} & -\frac{1}{2} \end{array}\right).$$
Also gilt
$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}^{-1} = \begin{pmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{pmatrix}.$$
△

Kapitel 3

Vektorräume und lineare Abbildungen

Der Begriff eines Vektorraums kommt in fast allen Bereichen der Mathematik vor. Vektorräume bilden beispielsweise den Hintergrund für fast alle Erscheinungsformen der Geometrie, sie werden in der theoretischen Physik und Informatik benutzt und erlauben uns auch viele Begriffe zu entwickeln, die beim Verständnis von linearen Gleichungssystemen helfen. Man kann eine Beobachtung über die Lösungsmengen von homogenen linearen Gleichungssystemen als erste Rechtfertigung des Begriffs heranziehen: Hat man zwei Lösungen eines homogenen linearen Gleichungssystems, so bildet die Summe der beiden Lösungen wieder eine Lösung: aus $Ac = 0 = Ad$ folgt nämlich

$$A(c + d) = Ac + Ad = 0 + 0 = 0.$$

Ebenso kann man eine Lösung mit einer Zahl skalieren und das Ergebnis ist wieder eine Lösung:

$$A(\lambda c) = \lambda(Ac) = \lambda 0 = 0.$$

Diese beiden Eigenschaften bilden den Kern der Definition eines Vektorraums, die wir jetzt mathematisch exakt einführen. *Ab jetzt sei K stets ein beliebiger Körper.*

3.1 Vektorräume

3.1.1 Grundbegriffe

Definition 3.1.1. (*i*) Ein K-**Vektorraum** ist eine abelsche Gruppe $(V, +)$ zusammen mit einer Verknüpfung

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

genannt die **Skalarmultiplikation**, so dass für alle $\lambda, \gamma \in K$ und $v, w \in V$ gilt:

$$\begin{aligned}
\lambda \cdot (v + w) &= (\lambda \cdot v) + (\lambda \cdot w) \\
(\lambda + \gamma) \cdot v &= (\lambda \cdot v) + (\gamma \cdot v) \\
(\lambda\gamma) \cdot v &= \lambda \cdot (\gamma \cdot v) \\
1 \cdot v &= v.
\end{aligned}$$

(*ii*) Die Elemente von V heißen **Vektoren**, die Elemente von K **Skalare**. Das neutrale Element der abelschen Gruppe V heißt **Nullvektor**/**Ursprung**. △

Bemerkung 3.1.2. Die Skalarmultiplikation ist eine weitere Verknüpfung, zusätzlich zu den schon vorhandenen Verknüpfungen des Körpers K und der abelschen Gruppe V. Die mehrfache Bedeutung der Symbole $+$ und $\cdot$ ist aber unproblematisch, da sich deren Bedeutung eindeutig daraus ergibt, woher die Elemente stammen, die miteinander verknüpft werden. Wie üblich lassen wir das Symbol $\cdot$ oft weg und vereinbaren, dass $\cdot$ stärker bindet als $+$. △

Bemerkung/Beispiel 3.1.3. (*i*) Für $m \geq 1$ ist K^m ein K-Vektorraum. Die Addition von Vektoren ist komponentenweise definiert, ebenso die Skalarmultiplikation:

$$\lambda \cdot \begin{pmatrix} v_1 \\ \vdots \\ v_m \end{pmatrix} := \begin{pmatrix} \lambda v_1 \\ \vdots \\ \lambda v_m \end{pmatrix}.$$

Die Vektorraumaxiome ergeben sich dabei unmittelbar aus den Körperaxiomen.
Die aus der Schule bekannten Räume $\mathbb{R}^2$ und $\mathbb{R}^3$ sind also Beispiele für $\mathbb{R}$-Vektorräume. Man sieht hier, dass man sich die Addition von Vektoren als Hintereinanderhängung von Pfeilen und die Skalarmultiplikation als Streckung vorstellen kann.
(*ii*) Etwas allgemeiner ist für $m, n \geq 0$ stets $\mathrm{Mat}_{m,n}(K)$ ein K-Vektorraum. Die skalare Multiplikation ist ebenfalls eintragsweise definiert.

(iii) Sei M eine beliebige Menge. Die Menge

$$\mathcal{F}(M, K) = \{f : M \to K\}$$

aller K-wertigen Abbildungen auf M ist nicht nur ein Ring (wie in Beispiel 2.2.5 (ii)), sondern kann auch als K-Vektorraum aufgefasst werden. Dafür vergisst man die Multiplikation von Funktionen (die in einem Vektorraum ja nicht benötigt wird) und definiert stattdessen die Skalarmultiplikation punktweise durch die Vorschrift

$$(\lambda \cdot f)(m) := \lambda f(m)$$

für alle $\lambda \in K, f \in \mathcal{F}(M, K)$ und $m \in M$. Strenggenommen ist das eine erneute Verallgemeinerung von (ii), da Matrizen ja eigentlich als K-wertige Abbildungen auf $\{1, \ldots, m\} \times \{1, \ldots, n\}$ definiert sind.
(iv) Die Menge $\mathcal{C}([0, 1], \mathbb{R})$ aller stetigen reellwertigen Abbildungen auf dem reellen Intervall $[0, 1]$ ist ein $\mathbb{R}$-Vektorraum, mit den Verknüpfungen wie in (iii). Dafür muss man aber zunächst zeigen, dass Summen und skalare Vielfache von stetigen Funktionen wieder stetig sind.
(v) Die Menge

$$K[t] := \left\{c_0 + c_1 t + \cdots + c_d t^d \mid d \in \mathbb{N}, c_i \in K\right\}$$

aller Polynome mit Koeffizienten aus K ist ein K-Vektorraum. Dabei addiert man Polynome einfach koeffizientenweise. Auch die Skalarmultiplikation ist koeffizientenweise definiert.
(vi) Für jeden Körper K existiert der **Nullvektorraum**, i.e. $V = \{0\}$. Die Verknüpfungen sind dabei eindeutig bestimmt. Man nennt ihn auch den *trivialen* Vektorraum.
(vii) Sind V, W zwei K-Vektorräume, so kann man das kartesische Produkt $V \times W$ wieder als K-Vektorraum auffassen. Addition und Skalarmultiplikation sind komponentenweise definiert:

$$(v_1, w_1) + (v_2, w_2) := (v_1 + v_2, w_1 + w_2), \quad \lambda \cdot (v, w) := (\lambda v, \lambda w).$$

($viii$) In Vektorräumen gelten gewisse einleuchtende Rechenregeln, die man leicht aus den Axiomen ableiten kann, wie etwa

$$0 \cdot v = 0 \text{ oder } (-\lambda) \cdot v = -(\lambda \cdot v).$$

(ix) Für $0 \neq v \in V$ und $\lambda, \gamma \in K$ mit $\lambda \neq \gamma$ gilt stets

$$\lambda v \neq \gamma v.$$

Auch das kann man leicht anhand der Vektorraumaxiome zeigen. Insbesondere besitzt für jeden unendlichen Körper K jeder nichttriviale K-Vektorraum unendlich viele Elemente.
(x) Für $K = \mathbb{Z}/2\mathbb{Z}$ gilt

$$K^2 = \left\{(0,0)^t, (0,1)^t, (1,0)^t, (1,1)^t\right\}$$

und dieser Vektorraum ist offensichtlich endlich. △

Definition 3.1.4. Sei V ein K-Vektorraum und $U \subseteq V$ eine *nichtleere* Teilmenge.
(i) U heißt **K-Untervektorraum** von V, falls gilt:

$$\begin{aligned} v, w \in U &\Rightarrow v + w \in U \\ \lambda \in K, v \in U &\Rightarrow \lambda v \in U. \end{aligned}$$

(ii) U heißt **affiner K-Unterraum** von V, falls ein $v \in V$ existiert, so dass

$$v + U := \{v + u \mid u \in U\}$$

ein K-Untervektorraum von V ist. Dabei heißt dann $v + U$ **der zu U parallele Untervektorraum**. △

Bemerkung 3.1.5. (i) Jeder Untervektorraum U enthält den Nullvektor. Dazu wählt man zunächst ein beliebiges $u \in U$ (denn U ist nicht leer!). Aufgrund der zweiten Eigenschaft eines Untervektorraums gilt $0 = 0 \cdot u \in U$.
(ii) Jeder K-Untervektorraum von V ist mit den eingeschränkten Verknüpfungen selbst wieder ein K-Vektorraum. Sowohl Summen als auch Vielfache von Vektoren aus U liegen wieder in U und die Vektorraum-Axiome gelten dabei natürlich immer noch.
(iii) Der Durchschnitt einer beliebigen Familie von Untervektorräumen von V ist wieder ein Untervektorraum. Wegen (i) ist solch ein Durchschnitt stets nicht leer und die beiden Unterraum-Axiome gelten für den Durchschnitt offensichtlich.
(iv) Jeder Untervektorraum ist ein affiner Unterraum. Die Umkehrung gilt im Allgemeinen nicht. Genauer ist ein affiner Unterraum U genau dann auch ein Untervektorraum, wenn $0 \in U$ (Übungsaufgabe).
(v) Untervektorräume können wir uns als Geraden, Ebenen etc. durch den Ursprung vorstellen. Affine Unterräume entstehen aus Untervektorräumen durch Verschiebung. △

Beispiel 3.1.6. (i) Für $A \in \mathrm{Mat}_{m,n}(K)$ ist die Lösungsmenge

$$L(A,0) \subseteq K^n$$

des von A definierten *homogenen* Gleichungssystems ein K-Untervektorraum. Das ist genau die Bemerkung vom Anfang dieses Kapitels.
(ii) Für $A \in \mathrm{Mat}_{m,n}(K)$ und $b \in K^m$ ist die Lösungsmenge

$$L(A,b) \subseteq K^n$$

des inhomogenen Systems ein affiner K-Unterraum. Das folgt direkt aus (i) und Satz 2.3.6. Für jedes beliebige $\bar{c} \in L(A,b)$ gilt ja

$$(-\bar{c}) + L(A,b) = L(A,0).$$

(iii) Für $d \in \mathbb{N}$ ist

$$K[t]_{\leq d} := \left\{c_0 + c_1 t + \cdots + c_d t^d \mid c_i \in K\right\}$$

aller Polynome vom Grad höchstens d ein Untervektorraum des Vektorraums $K[t]$. △

Jede Teilmenge eines Vektorraums kann zu einem Untervektorraum erweitert werden. Dabei müssen einfach alle möglichen Vielfachen und Summen von Elementen hinzugenommen werden. Das führt zum Begriff der *Linearkombination*.

Definition 3.1.7. Seien V ein K-Vektorraum, $v_1, \ldots, v_r \in V$ und $\lambda_1, \ldots, \lambda_r \in K$. Dann heißt das Element

$$\lambda_1 v_1 + \cdots + \lambda_r v_r \in V$$

eine **Linearkombination** der Vektoren $v_1, \ldots, v_r$. Die leere Linearkombination (also den Fall $r = 0$) definieren wir als den Nullvektor. △

Lemma 3.1.8. *Sei V ein K-Vektorraum und $M \subseteq V$ eine beliebige Teilmenge. Dann gilt*

$$\bigcap_{\substack{M \subseteq U \subseteq V \\ U \text{ Untervektorraum}}} U = \left\{ \sum_{i=1}^{r} \lambda_i v_i \mid r \in \mathbb{N}, \lambda_i \in K, v_i \in M \right\}$$

und dies ist der kleinste Untervektorraum, der M enthält.

Beweis. Die Menge $U' := \{\sum_{i=1}^{r} \lambda_i v_i \mid r \in \mathbb{N}, \lambda_i \in K, v_i \in M\}$ aller Linearkombinationen von Elementen aus M ist offensichtlich ein Untervektorraum der M enthält. Also taucht U' im linken Schnitt auf, was "$\subseteq$" beweist. Die Inklusion "$\supseteq$" folgt aus der Tatsache, dass jeder über M liegende Untervektorraum aufgrund der Axiome auch alle Linearkombinationen von Elementen aus M enthalten muss. Der Schnitt links ist in jedem über M liegenden Untervektorraum enthalten, d.h. es ist der kleinste solche. □

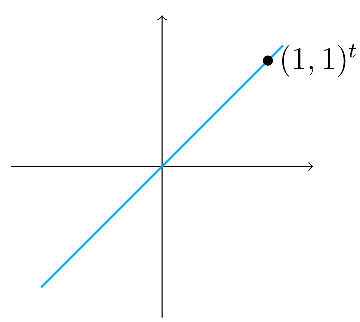
Definition 3.1.9. (i) Für $M \subseteq V$ bezeichnen wir die Menge aus Lemma 3.1.8 als K**-lineare Hülle** oder **Spann über** K von M und verwenden dafür auch die Bezeichnung

$$\mathrm{Span}_K(M).$$

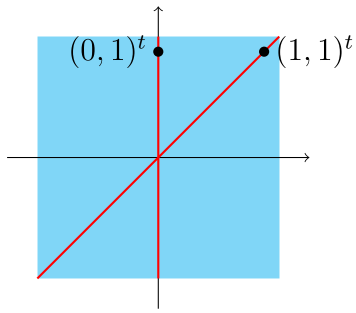
Es ist der kleinste K-Unterraum von V, der M enthält.
(ii) Ein K-Vektorraum heißt **endlich erzeugt** oder **endlich dimensional** wenn es eine *endliche* Teilmenge $M \subseteq V$ gibt mit

$$\mathrm{Span}_K(M) = V. \qquad \triangle$$

Beispiel 3.1.10. (i) Hat die Menge M nur ein Element, so ist jede Linearkombination von Elementen aus M nur ein Vielfaches dieses einen Vektors. Wir müssen uns $\mathrm{Span}_K(M)$ also als Ursprungsgerade vorstellen (außer im Fall, dass dieser eine Vektor der Nullvektor ist).
Wir betrachten beispielsweise den $\mathbb{R}$-Vektorraum $\mathbb{R}^2$ und darin die einelementige Menge $M = \{(1,1)^t\}$. Dann ist $\mathrm{Span}_{\mathbb{R}}(M)$ gerade die Ursprungsgerade durch den Punkt $(1,1)^t$:



(ii) Betrachtet man etwa $M = \{(1,1),(0,1)\} \subseteq \mathbb{R}^2$ so ergibt sich $\mathrm{Span}_{\mathbb{R}}(M) = \mathbb{R}^2$. Es reicht für die lineare Hülle ja nicht, nur Vielfache der beiden Vektoren zu betrachten, man muss auch Summen zulassen:



Will man das rechnerisch beweisen, so wählt man $(a,b)^t \in \mathbb{R}^2$ beliebig und setzt $\lambda = a, \gamma = b - a$. Dann rechnet man

$$\lambda \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \gamma \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \lambda \\ \lambda + \gamma \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix}$$

und hat so den Vektor $(a,b)^t$ als $\mathbb{R}$-Linearkombination der Elemente von M dargestellt.
(iii) Für jedes $m \in \mathbb{N}$ und jeden Körper K ist K^m ein endlich erzeugter K-Vektorraum. Dazu betrachten wir die Vektoren

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \; e_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \ldots, \; e_m = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

und sehen, dass offensichtlich

$$\mathrm{Span}_K\left(\{e_1, \ldots, e_m\}\right) = K^m$$

gilt.
(iv) Der K-Vektorraum $K[t]$ aller Polynome ist nicht endlich erzeugt. Endlich viele Polynome sind immer vom Grad beschränkt. Beim Übergang zum Spann entstehen niemals Polynome von höherem Grad. △

Beispiel 3.1.11. Sei $A \in \mathrm{Mat}_{m,n}(K)$ gegeben. Die n Spalten von A sind Elemente von K^m, wir bezeichnen sie mit $a_{\cdot 1}, \ldots, a_{\cdot n}$. Für $c = (c_1, \ldots, c_n)^t \in K^n$ gilt dann

$$Ac = c_1 a_{\cdot 1} + \cdots + c_n a_{\cdot n}$$

und somit gilt

$$\mathrm{Span}_K\left(\{a_{\cdot 1}, \ldots, a_{\cdot n}\}\right) = \{Ac \mid c \in K^n\}.$$

Diese Menge bezeichnet man als den **Spaltenraum** von A. Wir sehen hier, dass das durch $(A \mid b)$ definierte lineare Gleichungssystems genau dann eine Lösung besitzt, wenn b zum Spaltenraum von A gehört. △

Im Gegensatz zum Schnitt ist die Vereinigung zweier Untervektorräume normalerweise kein Untervektorraum mehr. Das haben wir in Beispiel 3.1.10 (ii) prinzipell schon gesehen. Man muss also den Spann der Vereinigung betrachten:

Lemma 3.1.12. *Sei V ein K-Vektorraum und U_1, U_2 Untervektorräume. Dann gilt*

$$\mathrm{Span}_K\left(U_1 \cup U_2\right) = \{u_1 + u_2 \mid u_1 \in U_1, u_2 \in U_2\}$$

und dies ist der kleinste K-Untervektorraum von V, der U_1 und U_2 enthält. Wir bezeichnen ihn auch mit $U_1 + U_2$.

Beweis. Die Menge $\{u_1 + u_2 \mid u_1 \in U_1, u_2 \in U_2\}$ auf der rechten Seite ist ein Untervektorraum von V, wie man direkt sieht (dabei verwendet man aber, dass U_1 und U_2 jeweils Untervektorräume sind). Gleichzeitig enthält sie aber sowohl U_1 als auch U_2 (da man immer entweder $u_1 = 0$ oder $u_2 = 0$ wählen kann). Als Untervektorraum der $U_1 \cup U_2$ enthält sie also auch $\mathrm{Span}_K\left(U_1 \cup U_2\right)$. Das beweist "$\subseteq$". Als Untervektorraum der U_1 und U_2 enthält, muss $\mathrm{Span}_K\left(U_1 \cup U_2\right)$ aber auch alle Summen vom Typ $u_1 + u_2$ mit $u_1 \in U_1, u_2 \in U_2$ enthalten. Das beweist "$\supseteq$". □

3.1.2 Lineare Unabhängigkeit, Basen und Dimension

Wir wollen in diesem Kapitel den Begriff der *Dimension* eines Vektorraums entwickeln. Der dafür grundlegende Begriff ist jener der *linearen Unabhängigkeit* von Vektoren:

Definition 3.1.13. Sei V ein K-Vektorraum und $v_1, \ldots, v_m \in V$.
(i) Das Tupel $(v_1, \ldots, v_m)$ heißt K**-linear unabhängig**, falls für alle $\lambda_1, \ldots, \lambda_m$ aus K gilt

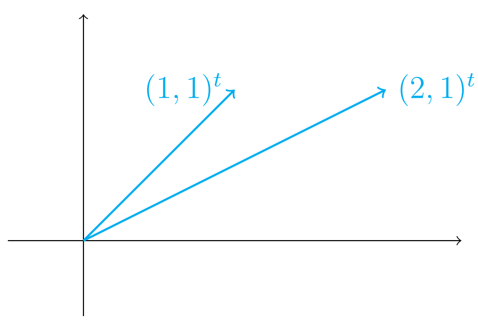
$$\lambda_1 v_1 + \cdots + \lambda_m v_m = 0 \quad \Rightarrow \quad \lambda_1 = \cdots = \lambda_m = 0.$$

Ansonsten heißt es K**-linear abhängig**.
(ii) $(v_1, \ldots, v_m)$ heißt K-**Erzeugendendsystem** von V, falls

$$\mathrm{Span}_K\left(\{v_1, \ldots, v_m\}\right) = V.$$

(iii) Das Tupel $(v_1, \ldots, v_m)$ heißt K**-Basis** von V, falls es linear unabhängig und ein Erzeugendensystem von V ist. △

Bemerkung/Beispiel 3.1.14. (i) Wir haben lineare Unabhängigkeit/Erzeugendensystem/Basis für *Tupel von Vektoren* definiert. Dabei kommt es allerdings nicht auf die Reihenfolge der Vektoren an, wie man leicht sieht. Später wird die Reihenfolge aber manchmal wichtig sein. Auch passiert es manchmal, dass ein Vektor in solch einem Tupel mehrfach auftritt (dann ist das Tupel aber nicht linear unabhängig). Deshalb sprechen wir von Tupel und nicht von einer *Menge von Vektoren*. In vielen Fällen ist es allerdings nicht nötig, diese Unterscheidung zu treffen und wir sind im folgenden nicht immer sehr streng.
(ii) Linear unabhängig sind Vektoren $v_1, \ldots, v_m$ also genau dann, wenn der Nullvektor *nur* als Linearkombination aus ihnen entsteht, *wenn alle Koeffizienten dabei Null sind* (und niemals auf andere Weise).
(iii) Im $\mathbb{R}^2$ sind die beiden Vektoren $v_1 = (1,1)^t, v_2 = (2,1)^t$ linear unabhängig. Man kann sie also nicht skalieren und addieren und damit den Nullpunkt erreichen (außer man skaliert beide auf Null).



Rechnerisch sieht man das folgendermaßen: Aus

$$0 = \lambda_1 v_1 + \lambda_2 v_2 = (\lambda_1 + 2\lambda_2, \lambda_1 + \lambda_2)$$

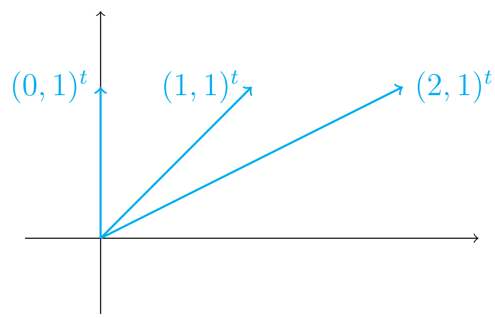
folgt $\lambda_1 + 2\lambda_2 = 0$ und $\lambda_1 + \lambda_2 = 0$. Dieses System linearer Gleichungen über $\mathbb{R}$ besitzt nur die triviale Lösung.
Zwei Vektoren sind allgemein genau dann linear abhängig, wenn einer ein Vielfaches des anderen ist. Aus $0 = \lambda_1 v_1 + \lambda_2 v_2$ mit etwa $\lambda_1 \neq 0$ folgt nämlich

$$v_1 = -\frac{\lambda_2}{\lambda_1} v_2$$

und umgekehrt folgt aus $v_1 = \gamma v_2$ direkt $0 = 1 \cdot v_1 - \gamma v_2$. Für zwei Vektoren bedeutet lineare Unabhängigkeit also anschaulich, dass sie in verschiedene Richtungen zeigen.

(iv) Ein einzelner Vektor $0 \neq v \in V$ ist stets linear unabhängig. Aus $\lambda v = 0$ und $\lambda \neq 0$ folgt nach Multiplikation mit λ^{-1} ja $v = 0$. Der Nullvektor ist der einzige Vektor, der für sich allein linear abhängig ist.
(v) Lineare Unabhängigkeit ist eine Bedingung, die nur für alle Vektoren $v_1, \ldots, v_m$ *gemeinsam* entschieden werden kann. So könnte durchaus jede $m-1$-elementige Teilmenge der Vektoren linear unabhängig sein, ohne dass es für alle gemeinsam zutrifft. Das sieht man beispielsweise, wenn man in (iii) noch einen dritten Vektor, etwa $v_3 = (0,1)^t$, hinzunimmt. Es gilt dann beispielsweise

$$2v_1 - v_2 - v_3 = 0.$$



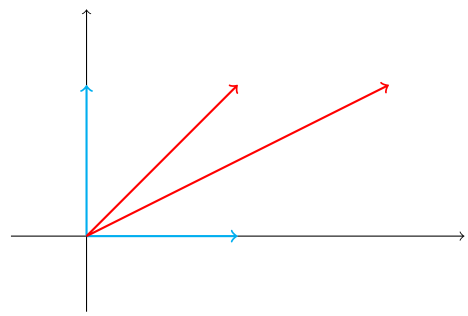
(vi) Sei $V = \mathcal{F}([0,1], \mathbb{R})$ der $\mathbb{R}$-Vektorraum der reellwertigen Funktionen. Die beiden Funktionen

$$f_1 \colon x \mapsto x, \quad f_2 \colon x \mapsto x^2$$

gehören zu V und sind $\mathbb{R}$-linear unabhängig. Dazu nehmen wir

$$\lambda_1 f_1 + \lambda_2 f_2 = 0$$

an und setzen in diese Gleichung auf beiden Seiten $x = 1$ ein. Das liefert $\lambda_1 + \lambda_2 = 0$. Setzen wir stattdessen $x = \frac{1}{2}$ ein, erhalten wir $\frac{1}{2}\lambda_1 + \frac{1}{4}\lambda_2 = 0$. Beide Gleichungen zusammen besitzen nur die triviale Lösung.
(vii) Der Begriff von linearer Unabhängigkeit kann auch für *unendlich viele* Vektoren definiert werden. Dafür betrachtet man aber keine unendlich langen Linearkombinationen, die ja in V überhaupt nicht definiert sind. Es müssen stattdessen je endlich viele der Vektoren linear unabhängig im Sinne unserer Definition sein. Wir beschränken uns in dieser Vorlesung aber auf endliche Mengen von Vektoren.
($viii$) Ein Vektorraum besitzt im Allgemeinen viele verschiedene Basen. Im $\mathbb{R}^2$ können wir beispielsweise $e_1 = (1,0)^t, e_2 = (0,1)^t$ oder auch $v_1 = (1,1)^t, v_2 = (2,1)^t$ wählen.



$\triangle$

Satz 3.1.15. *Sei V ein K-Vektorraum und $v_1, \ldots, v_m \in V$. Dann sind äquivalent:*

(i) $(v_1, \ldots, v_m)$ *bildet eine Basis von V.*

(ii) *Zu jedem $v \in V$ existieren eindeutig bestimmte Skalare $\lambda_1, \ldots, \lambda_m \in K$ mit $v = \lambda_1 v_1 + \cdots + \lambda_m v_m$.*

(iii) $(v_1, \ldots, v_m)$ *ist maximal linear unabhängig, d.h. das Hinzufügen eines beliebigen weiteren Vektors führt stets zur linearen Abhängigkeit.*

(iv) $(v_1, \ldots, v_m)$ *ist ein minimales Erzeugendensystem, d.h. die Vektoren spannen V auf, aber nach Wegnahme eines beliebigen der v_i stimmt das nicht mehr.*

Beweis. "$(i) \Rightarrow (ii)$": Da $(v_1, \ldots, v_m)$ als Basis insbesondere ein Erzeugendensystem von V bildet, existieren für jeden Vektor $v \in V$ Skalare $\lambda_1, \ldots, \lambda_m \in K$ mit $v = \lambda_1 v_1 + \cdots + \lambda_m v_m$. Für die Eindeutigkeit nehmen wir eine zweite solche Darstellung an:

$$v = \lambda_1 v_1 + \cdots + \lambda_m v_m = \gamma_1 v_1 + \cdots + \gamma_m v_m$$

mit allen $\gamma_i \in K$. Durch Subtraktion ergibt sich

$$0 = \lambda_1 v_1 + \cdots + \lambda_m v_m - (\gamma_1 v_1 + \cdots + \gamma_m v_m) = (\lambda_1 - \gamma_1) v_1 + \cdots + (\lambda_m - \gamma_m) v_m.$$

Da $(v_1, \ldots, v_m)$ als Basis linear unabhängig ist folgt hieraus

$$\lambda_1 - \gamma_1 = \cdots = \lambda_m - \gamma_m = 0,$$

also $\lambda_1 = \gamma_1, \ldots, \lambda_m = \gamma_m$.
"$(ii) \Rightarrow (iii)$": Die lineare Unabhängigkeit von $(v_1, \ldots, v_m)$ folgt aus der Tatsache, dass der Nullvektor natürlich immer die Darstellung $0 = 0v_1 + \cdots + 0v_m$ besitzt

und aufgrund der Eindeutigkeit also nur diese. Sei nun $v \in V$ ein weiterer Vektor. Dann gibt es eine Darstellung

$$v = \lambda_1 v_1 + \cdots + \lambda_m v_m$$

mit allen $\lambda_i \in K$. Damit erhalten wir

$$\lambda_1 v_1 + \cdots + \lambda_m v_m + (-1) \cdot v = 0$$

und das ist eine nichttriviale Darstellung des Nullvektors. Also ist $(v_1, \ldots, v_m, v)$ linear abhängig.
"$(iii) \Rightarrow (iv)$": Für $v \in V$ ist $(v_1, \ldots, v_m, v)$ linear abhängig, es gibt also $\lambda_1, \ldots, \lambda_m, \lambda \in K$, nicht alle Null, mit

$$\lambda_1 v_1 + \cdots + \lambda_m v_m + \lambda v = 0.$$

Da $(v_1, \ldots, v_m)$ jedoch linear unabhängig ist, muss $\lambda \neq 0$ gelten. Wir erhalten also

$$v = \tfrac{-\lambda_1}{\lambda} v_1 + \cdots + \tfrac{-\lambda_m}{\lambda} v_m.$$

Damit gezeigt, dass $(v_1, \ldots, v_m)$ ein Erzeugendensystem von V bildet. Es gilt nun

$$v_1 \notin \operatorname{Span}_K(\{v_2, \ldots, v_m\}),$$

denn eine Darstellung

$$v_1 = \lambda_2 v_2 + \cdots + \lambda_m v_m$$

würde

$$1 v_1 - \lambda_2 v_2 - \cdots - \lambda_m v_m = 0$$

implizieren, ein Widerspruch zur linearen Unabhängigkeit von $(v_1, \ldots, v_m)$. Ebenso kann kein anderes v_i durch die restlichen erzeugt werden kann. Damit ist die Minimalität bewiesen.
"$(iv) \Rightarrow (i)$": Nach Voraussetzung bildet $(v_1, \ldots, v_m)$ bereits ein Erzeugendensystem von V. Angenommen es gibt nun eine Darstellung

$$\lambda_1 v_1 + \cdots + \lambda_m v_m = 0$$

mit gewissen $\lambda_i \in K$, nicht alle Null. Wir nehmen ohne Beschränkung der Allgemeinheit $\lambda_1 \neq 0$ an. Dann gilt

$$v_1 = \tfrac{-\lambda_2}{\lambda_1} v_2 + \cdots + \tfrac{-\lambda_m}{\lambda_1} v_m.$$

Jede Linearkombination von $v_1, \ldots, v_m$ ist damit aber schon eine Linearkombination von $v_2, \ldots, v_m$:

$$\gamma_1 v_1 + \gamma_2 v_2 + \cdots + \gamma_m v_m = \left(\gamma_2 - \tfrac{\gamma_1 \lambda_2}{\lambda_1}\right) v_2 + \cdots + \left(\gamma_m - \tfrac{\gamma_1 \lambda_m}{\lambda_1}\right) v_m.$$

Mit $(v_1, \ldots, v_m)$ bildet damit aber auch $(v_2, \ldots, v_m)$ ein Erzeugendensystem von V, ein Widerspruch zur Minimalität. Also sind alle $\lambda_i = 0$ und $(v_1, \ldots, v_m)$ ist damit linear unabhängig. □

Beispiel 3.1.16. (i) Die Vektoren $e_1, \ldots, e_m$ aus Beispiel 3.1.10 (iii) sind eine K-Basis des K^m. Jede der vier Bedingungen des letzten Satzes kann man leicht überprüfen. Man nennt $(e_1, \ldots, e_m)$ die **Standardbasis** von K^m.
(ii) Für $d \in \mathbb{N}$ sei

$$K[t]_{\leq d} = \left\{c_0 + c_1 t + \cdots + c_d t^d \mid c_i \in K\right\}$$

der K-Vektorraum der Polynome vom Grad höchstens d. Er besitzt beispielsweise die Basis $(1, t, t^2, \ldots, t^d)$. △

Korollar 3.1.17. *Jeder endlich erzeugte K-Vektorraum besitzt eine Basis.*

Beweis. Da V endlich erzeugt ist, gibt es eine endliche Teilmenge $M \subseteq V$ mit $\mathrm{Span}_K(M) = V$. Wir wählen dabei ein solches M mit kleinstmöglicher Mächtigkeit. Dann kann man offensichtlich kein Element von M mehr weglassen, ohne die Bedingung *Erzeugendensystem* zu zerstören. Also erfüllt M Bedingung (iv) aus Satz 3.1.15 und ist also eine Basis von V. □

Satz 3.1.18 (Basisergänzungssatz)**.** *Sei V ein endlich erzeugter K-Vektorraum, sei $(v_1, \ldots, v_m)$ ein Erzeugendensystem von V und $(w_1, \ldots, w_n)$ linear unabhängig. Dann gilt:*

(i) $n \leq m$.

(ii) Je zwei Basen von V haben dieselbe Länge.

(iii) Es gibt unter $v_1, \ldots, v_m$ Vektoren $v_{i_1}, \ldots, v_{i_r}$ so, dass

$$(w_1, \ldots, w_n, v_{i_1}, \ldots, v_{i_r})$$

eine Basis von V bildet.

Beweis. Für (i) nehmen wir $m < n$ an. Da $v_1, \dots, v_m$ ganz V erzeugen, finden wir für jedes $j = 1, \dots, n$ Darstellungen

$$w_j = \sum_{i=1}^{m} \lambda_{ij} v_i$$

mit $\lambda_{ij} \in K$. Die Matrix

$$M = (\lambda_{ij})_{i,j} \in \mathrm{Mat}_{m,n}(K)$$

hat mehr Spalten als Zeilen, also besitzt das dazugehörige homogene lineare Gleichungssystem eine nichttriviale Lösung $\gamma = (\gamma_1, \dots, \gamma_n)^t \in K^n \backslash \{0\}$. Es gilt nun

$$\sum_{j=1}^{n} \gamma_j w_j = \sum_{j=1}^{n} \sum_{i=1}^{m} \gamma_j \lambda_{ij} v_i = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} \lambda_{ij} \gamma_j \right) v_i.$$

Die Koeffizienten in der Summe rechts sind genau die Einträge des Vektors $M\gamma$, also alle Null. Wir erhalten also $\sum_{j=1}^{m} \gamma_j w_j = 0$, ein Widerspruch zur linearen Unabhängigkeit von $(w_1, \dots, w_n)$.
(ii) Fasst man die eine gegebene Basis von V als Erzeugendensystem auf und die andere als Menge linear unabhängiger Elemente, so liefert (i) eine Ungleichung zwischen ihren Längen. Macht man das genau andersherum, bekommt man die andere Ungleichung.
(iii) Die Vektoren $w_1, \dots, w_n, v_1, \dots, v_m$ erzeugen offensichtlich ganz V. Durch eventuelles Entfernen gewisser v_i erreichen wir, dass die übrigen Vektoren

$$w_1, \dots, w_n, v_{i_1}, \dots, v_{i_r}$$

immer noch ganz V erzeugen, aber keines der v_{i_j} dazu noch weggelassen werden kann. Wir überlegen uns nun, dass dann auch keines der w_i mehr weggelassen werden kann, um V zu erzeugen. Dann handelt es sich um ein minimales Erzeugendensystem, also um eine Basis von V.
Wir nehmen im Gegenteil nun an, dass etwa $w_2, \dots, w_n, v_{i_1}, \dots, v_{i_r}$ auch noch ganz V erzeugen. Insbesondere gibt es eine Darstellung

$$w_1 = \sum_{i=2}^{n} \lambda_i w_i + \sum_{j=1}^{r} \gamma_j v_{i_j}.$$

Dabei können nicht alle $\gamma_j = 0$ sein, das widerspräche der linearen Unabhängigkeit der w_i. Wir nehmen also o.B.d.A. $\gamma_1 \neq 0$ an und lösen die Gleichung nach v_{i_1} auf:

$$v_{i_1} = \frac{1}{\gamma_1} w_1 - \sum_{i=2}^{n} \frac{\lambda_i}{\gamma_1} w_i - \sum_{j=2}^{r} \frac{\gamma_j}{\gamma_1} v_{i_j}.$$

Das zeigt

$$\operatorname{Span}_K(\{w_1, \ldots, w_n, v_{i_2}, \ldots, v_{i_r}\}) = \operatorname{Span}_K(\{w_1, \ldots, w_n, v_{i_1}, \ldots, v_{i_r}\}) = V.$$

Man hätte also statt w_1 auch v_{i_1} oben weglassen können, ein Widerspruch. □

Definition 3.1.19. Sei V ein endlich erzeugter K-Vektorraum. Die **Dimension** von V (über K) ist die Länge einer/jeder K-Basis von V. Wir verwenden dazu die Notation

$$\dim_K(V).$$

△

Beispiel 3.1.20. (*i*) Wir betrachten den K-Vektorraum K^m. Er besitzt die Standardbasis $(e_1, \ldots, e_m)$, die aus genau m Vektoren besteht. Somit besteht jede andere Basis ebenfalls aus m Vektoren und es gilt

$$\dim_K(K^m) = m.$$

Jedes Erzeugendensystem von K^m hat mindestens m Elemente und jede Menge linear unabhängiger Vektoren hat höchstens m Elemente.
(*ii*) Im Fall $m = 3$ betrachten wir $w_1 = (1, 1, 1)^t, w_2 = (2, -1, 1)^t \in \mathbb{Q}^3$. Diese beiden Vektoren sind linear unabhängig, wie man leicht nachrechnet. Zusätzlich verwenden wir die Standardbasis (e_1, e_2, e_3) als Erzeugendensystem von $\mathbb{Q}^3$. Wir können also (w_1, w_2) durch Hinzunahme eines der e_i zu einer Basis ergänzen. In diesem Fall funktioniert sogar jedes beliebige e_i.
(*iii*) Der Vektorraum $K[t]_{\leq d}$ besitzt die Basis $(1, t, \ldots, t^d)$, es gilt also

$$\dim_K(K[t]_{\leq d}) = d + 1.$$

Man sieht leicht, dass auch $p_1 = 1 + t, p_2 = 1 - t$ linear unabhängig sind. Man kann (p_1, p_2) also durch gewisse t^i zu einer Basis ergänzen. Hier gibt es nur eine Möglichkeit:

$$(p_1, p_2, t^2, t^3, \ldots, t^d).$$

△

Die folgende Aussage erlaubt es, mit Satz 2.3.12 (*iii*) zu entscheiden, ob gegebene Spaltenvektoren eine Basis des K^m bilden:

Proposition 3.1.21. *Die Vektoren $v_1, \ldots, v_m \in K^m$ bilden genau dann eine K-Basis von K^m, wenn die Matrix $M = (v_1, \ldots, v_m) \in \mathrm{Mat}_m(K)$ invertierbar über K ist.*

Beweis. Übungsaufgabe. □

Konstruktion 3.1.22. Sei (A, b) die erweiterte Koeffizientenmatrix eines linearen Gleichungssystems über K. Dann ist $L(A, b) \subseteq K^n$ ein affiner Unterraum und wir haben in Satz 2.3.6 gezeigt, dass für jedes beliebige $\bar{c} \in L(A, b)$ stets

$$L(A, b) = \bar{c} + L(A, 0)$$

gilt. Weiter ist $L(A, 0)$ ein Untervektorraum und wir können für ihn eine endliche Basis $(v_1, \ldots, v_s)$ finden. Dadurch hat man die (meist unendliche) Menge $L(A, b)$ durch die endlich vielen Daten $\bar{c}, v_1, \ldots, v_s$ komplett angegeben, denn es gilt dann

$$L(A, b) = \left\{ \bar{c} + \sum_{i=1}^{s} \lambda_i v_i \mid \lambda_i \in K \right\}.$$

Eine Basis von $L(A, 0)$ findet man folgendermaßen: zunächst bringt man A mit elementaren Zeilenumformungen auf Zeilenstufenform. Dabei ergeben sich r Pivotspalten und deswegen $n - r$ freie Variablen. Für jede freie Variable konstruiert man sich einen Lösungsvektor, indem man als Wert für diese Variable 1, für die anderen freien Variablen 0 wählt. Danach berechnet man die Werte für alle anderen Einträge. Insgesamt erhält man so $n - r$ Vektoren $v_1, \ldots, v_{n-r}$, die offensichtlich eine Basis von $L(A, 0)$ bilden. Damit ist $n - r$ gerade die Dimension von $L(A, 0)$. So sieht man auch, dass die Anzahl r der Pivots nicht von der genauen Durchführung des Gauß-Algorithmus abhängen. △

Proposition 3.1.23 (★). *Sei V ein endlich erzeugter K-Vektorraum und $U \subsetneq V$ ein echter Untervektorraum. Dann ist auch U endlich erzeugt und es gilt*

$$\dim_K(U) < \dim_K(V).$$

Beweis. Wenn $u_1, \ldots, u_r \in U$ linear unabhängig sind, so sind sie dies offensichtlich auch, wenn man sie als Elemente von V auffasst. Daraus folgt mit Satz 3.1.18 schon $r \leq \dim_K(V)$. Wir wählen uns in U nun ein linear unabhängiges Tupel von Vektoren von maximaler Länge, das nach Satz 3.1.15 dann eine Basis von U mit höchstens Länge $\dim_K(V)$ bildet. Insbesondere ist U endlich erzeugt mit $\dim_K(U) \leq \dim_K(V)$. Sei nun $(u_1, \ldots, u_{\dim_K(U)})$ eine Basis von U. Aus $U \subsetneq V$ folgt, dass es sich dabei nicht um ein Erzeugendensystem und deswegen erst recht nicht um eine Basis von V handelt. Man kann es aber mit Satz 3.1.18 zu einer Basis von V ergänzen. Dabei muss natürlich mindestens ein zusätzlicher Vektor hinzugenommen werden, woraus $\dim_K(U) < \dim_K(V)$ folgt. □

Bemerkung 3.1.24 (★). Eine aufsteigende Folge echter Untervektorräume in einem m-dimensionalen Vektorraum (z.B. im K^m) hat höchstens Länge $m+1$. Man beginnt mindestens mit dem 0-dimensionalen Unterraum $\{0\}$. In jedem weiteren Schritt wird die Dimension um mindestens eins größer und ist immer höchstens m.

Satz 3.1.25 (Dimensionsformel für Untervektorräume ★). *Sei V ein K-Vektorraum und U_1, U_2 zwei endlich-dimensionale Untervektorräume von V. Dann gilt*

$$\dim_K(U_1 + U_2) = \dim_K(U_1) + \dim_K(U_2) - \dim_K(U_1 \cap U_2).$$

Beweis. Zunächst wählen wir eine Basis $(u_1, \ldots, u_m)$ des Unterraums $U_1 \cap U_2$. Diese ergänzen wir dann einerseits zu einer Basis

$$(u_1, \ldots, u_m, v_1, \ldots, v_n)$$

von U_1, andererseits zu einer Basis

$$(u_1, \ldots, u_m, w_1, \ldots, w_r)$$

von U_2. Dann erzeugt

$$(u_1, \ldots, u_m, v_1, \ldots, v_n, w_1, \ldots, w_r)$$

offensichtlich ganz $U_1 + U_2$, ist aber auch linear unabhängig. Dazu nehmen wir

$$\sum_i \lambda_i u_i + \sum_i \gamma_i v_i + \sum_i \delta_i w_i = 0$$

an. Dann gilt

$$U_1 \ni \sum_i \lambda_i u_i + \sum_i \gamma_i v_i = -\sum_i \delta_i w_i \in U_2.$$

Da jedes Element aus $U_1 \cap U_2$ eine eindeutige Linearkombination der u_i und jedes Element aus U_1 eine eindeutige Linearkombination der u_i, v_i ist, folgt daraus $0 =$

$\gamma_1 = \cdots = \gamma_n$. Aus der linearen Unabhängigkeit des Tupels $(u_1, \ldots, u_m, w_1, \ldots, w_r)$ folgt dann direkt

$$0 = \lambda_1 = \cdots = \lambda_m = \delta_1 = \cdots = \delta_r.$$

Nun berechnen wir

$$\dim_K(U_1 + U_2) = m + n + r$$

und

$$\dim_K(U_1) + \dim_K(U_2) - \dim_K(U_1 \cap U_2) = (m+n) + (m+r) - m = m+n+r.$$

Das beweist die Aussage. $\square$

Beispiel 3.1.26 (★). Sind $U_1, U_2 \subseteq \mathbb{R}^3$ zwei unterschiedliche 2-dimensionale Unterräume, so gilt $U_1 \subsetneq U_1 + U_2$, also

$$2 = \dim_K(U_1) < \dim_K(U_1 + U_2) \leq 3.$$

Daraus folgt $U_1 + U_2 = \mathbb{R}^3$. Weiter gilt nun

$$\begin{aligned} 3 &= \dim_K(U_1 + U_2) \\ &= \dim_K(U_1) + \dim_K(U_2) - \dim_K(U_1 \cap U_2) \\ &= 4 - \dim_K(U_1 \cap U_2), \end{aligned}$$

woraus $\dim_K(U_1 \cap U_2) = 1$ folgt. Zwei unterschiedliche Ebenen in $\mathbb{R}^3$ schneiden sich also in einer Geraden. $\triangle$

3.2 Lineare Abbildungen

Vektorräume sind Mengen mit zusätzlicher Struktur, nämlich Addition und Skalarmultiplikation. Abbildungen zwischen Vektorräumen sollten sich sinnvollerweise mit dieser Struktur vertragen. Das führt zum Begriff einer *linearen Abbildung*. Das Konzept von linearen Abbildungen ermöglicht es uns außerdem, lineare Gleichungssysteme und Matrizen noch systematischer zu fassen und verstehen. Auch in diesem Abschnitt sei K stets wieder ein Körper.

3.2.1 Grundbegriffe

Definition 3.2.1. Seien V, W zwei K-Vektorräume.
(i) Eine Abbildung $\varphi\colon V \to W$ heißt **linear**, falls für alle $v_1, v_2 \in V, \lambda \in K$ stets

$$\varphi(v_1 + v_2) = \varphi(v_1) + \varphi(v_2) \quad \text{und} \quad \varphi(\lambda v_1) = \lambda\varphi(v_1)$$

gilt.
(ii) Eine lineare Abbildung $\varphi\colon V \to W$ heißt **Isomorphismus**, falls eine lineare Abbildung $\psi\colon W \to V$ existiert mit

$$\psi \circ \varphi = \mathrm{id}_V \quad \text{und} \quad \varphi \circ \psi = \mathrm{id}_W.$$

(iii) V und W heißen **isomorph**, falls ein Isomorphismus $\varphi\colon V \to W$ existiert. Wir schreiben dafür $V \cong W$. △

Bemerkung/Beispiel 3.2.2. (i) Jede Matrix $A \in \mathrm{Mat}_{m,n}(K)$ definiert durch Matrixmultiplikation von links eine lineare Abbildung

$$\begin{aligned} \mu_A\colon K^n &\to K^m \\ c &\mapsto Ac. \end{aligned}$$

In dieser Notation gilt

$$L(A, b) = \mu_A^{-1}(b).$$

Das lineare Gleichungssystem zu lösen bedeutet also, das Urbild des Vektors b unter dieser linearen Abbildung zu bestimmen.
(ii) Für $A \in \mathrm{Mat}_{m,n}(K)$ und $B \in \mathrm{Mat}_{r,m}(K)$ gilt

$$\mu_B \circ \mu_A = \mu_{BA}.$$

(iii) Die Differentiation

$$\begin{aligned} \partial\colon C^1(\mathbb{R}, \mathbb{R}) &\to C(\mathbb{R}, \mathbb{R}) \\ f &\mapsto f' \end{aligned}$$

ist eine lineare Abbildung. Die Linearität bedeutet ja gerade

$$(f + g)' = f' + g' \quad \text{und} \quad (\lambda f)' = \lambda f',$$

was aus der Analysis bekannt ist.

(iv) Zumindest Polynome kann man über jedem Körper (formal) ableiten. So erhält man die lineare Abbildung

$$\begin{aligned} \partial\colon K[t] &\to K[t] \\ \sum_{i=0}^{d} c_i t^i &\mapsto \sum_{i=1}^{d} i c_i t^{i-1}. \end{aligned}$$

Möchte man lieber endlich-dimensionale Vektorräume verwenden, kann man beispielsweise

$$\partial\colon K[t]_{\leq d} \to K[t]_{\leq d-1}$$

betrachten.
(v) Ein Isomorphismus ist offensichtlich immer bijektiv. Umgekehrt ist aber auch jede bijektive lineare Abbildung ein Isomorphismus. Aus der Bijektivität von φ folgt die Existenz der Umkehrabbildung $\varphi^{-1} =: \psi$. Man sieht nun relativ leicht, dass ψ automatisch auch linear ist. Das beweist die Aussage. Man könnte Isomorphismus als auch als *bijektive lineare Abbildung* definieren.
(vi) Isomorphie ist ein verallgemeinerter Gleichheitsbegriff für Vektorräume. Zwei isomorphe Vektorräume sind eigentlich völlig identisch, nur tragen ihre Elemente eventuell verschiedene Namen. Ein Isomorphismus $\varphi\colon V \to W$ gibt dabei einfach die Umbenennung der Elemente an. Grundsätzlich müssen wir zwischen isomorphen Vektorräumen nicht mehr unterscheiden. Jede sinnvolle mathematische Aussage, die für einen Vektorraum gilt, gilt auch für jeden dazu isomorphen Vektorraum. △

Lemma 3.2.3. *Seien V, W zwei K-Vektorräume und $\varphi\colon V \to W$ linear. Dann gilt*

$$\varphi(0) = 0 \quad \textit{und} \quad \varphi(-v) = -\varphi(v)$$

für alle $v \in V$.

Beweis. Übungsaufgabe. □

Bemerkung 3.2.4. (i) Zwei lineare Abbildungen $\varphi, \psi\colon V \to W$ kann man punktweise addieren:

$$\begin{aligned} \varphi + \psi\colon V &\to W \\ v &\mapsto \varphi(v) + \psi(v). \end{aligned}$$

Man rechnet dabei direkt nach, dass $\varphi + \psi$ selbst wieder eine lineare Abbildung ist. Ebenso kann man φ mit einem Skalar λ punktweise multiplizieren:

$$\begin{aligned} \lambda\varphi\colon V &\to W \\ v &\mapsto \lambda\varphi(v) \end{aligned}$$

und auch $\lambda\varphi$ ist wieder eine lineare Abbildung. Auf diese Weise wird die Menge

$$\mathrm{Lin}_K(V, W) := \{\varphi\colon V \to W \mid \varphi \text{ linear}\}$$

selbst zu einem K-Vektorraum.
(ii) Für $A, B \in \mathrm{Mat}_{m,n}(K)$ und $\lambda \in K$ gilt

$$\mu_A + \mu_B = \mu_{A+B} \text{ und } \lambda\mu_A = \mu_{\lambda A}.$$

Das rechnet man direkt punktweise nach:

$$\begin{aligned} (\mu_A + \mu_B)(v) = \mu_A(v) + \mu_B(v) = Av + Bv = (A+B)v = \mu_{A+B}(v) \\ (\lambda\mu_A)(v) = \lambda \cdot \mu_A(v) = \lambda(Av) = (\lambda A)v = \mu_{\lambda A}(v). \end{aligned} \qquad \triangle$$

Satz 3.2.5. *Sei V ein K-Vektorraum mit $\dim_K(V) = m$. Dann gilt*

$$V \cong K^m.$$

Beweis. Wir wählen eine Basis $(v_1, \ldots, v_m)$ von V und betrachten die lineare Abbildung

$$\begin{aligned} \varphi\colon K^m &\to V \\ (\lambda_1, \ldots, \lambda_m)^t &\mapsto \sum_{i=1}^m \lambda_i v_i. \end{aligned}$$

Da eine Basis insbesondere ein Erzeugendensystem von V ist, ist φ surjektiv. Da eine Basis linear unabhängig ist, ist φ injektiv. Damit ist φ ein Isomorphismus. $\square$

Bemerkung 3.2.6. (i) Der Isomorphismus zwischen K^m und V ist keineswegs eindeutig. Er hängt von der Wahl der Basis $(v_1, \ldots, v_m)$ ab und insbesondere von der gewählten Reihenfolge der v_i. Mit $\underline{v} := (v_1, \ldots, v_m)$ sollte man ihn also besser als

$$\varphi_{\underline{v}}\colon K^m \to V$$

bezeichnen. Hier sieht man, warum wir Basen als Tupel von Vektoren definiert haben.
(*ii*) Auch die Umkehrabbildung von $\varphi_{\underline{v}}$ kann man leicht angeben. Dazu entwickelt man $v \in V$ bezüglich der Basis $\underline{v}$ als

$$v = \sum_{i=1}^{m} \lambda_i v_i$$

und bildet v auf

$$\xi_{\underline{v}}(v) := (\lambda_1, \dots, \lambda_m)^t \in K^m$$

ab. Dadurch erhält man die Umkehrabbildung zu $\varphi_{\underline{v}}$:

$$\xi_{\underline{v}} \colon V \to K^m.$$

△

Definition 3.2.7. Sei V ein K-Vektorraum und $\underline{v} = (v_1, \dots, v_m)$ eine Basis von V. Für $v \in V$ mit $v = \sum_{i=1}^{m} \lambda_i v_i$ nennt man

$$\xi_{\underline{v}}(v) = (\lambda_1, \dots, \lambda_m)^t \in K^m$$

den **Koordinatenvektor** von v bezüglich $\underline{v}$. △

Beispiel 3.2.8. (*i*) Wir betrachten den Vektorraum

$$V = K[t]_{\leq 3} = \left\{ c_0 + c_1 t + c_2 t^2 + c_3 t^3 \mid c_0, c_1, c_2, c_3 \in K \right\}$$

aller Polynome vom Grad höchstens 3, mit der Basis

$$\underline{p} = (1, t, t^2, t^3).$$

Für ein Polynom $p = c_0 + c_1 t + c_2 t^2 + c_3 t^3 \in V$ gilt offensichtlich

$$\xi_{\underline{p}}(p) = (c_0, c_1, c_2, c_3)^t \in K^4.$$

Der Koordinatenvektor von $p = 1 - t^2 + 2t^3$ ist also

$$\xi_{\underline{p}}(p) = (1, 0, -1, 2)^t.$$

Insgesamt erhalten wir den Isomorphismus

$$\begin{aligned} K^4 &\leftrightarrow V \\ (c_0, c_1, c_2, c_3)^t &\leftrightarrow \sum_{i=0}^{3} c_i t^i \end{aligned}$$

(ii) Wir betrachten denselben Raum V wie in (i), wählen aber diesmal die Basis

$$\underline{q} = (1-t, 1+t, t^2, t^3).$$

Um für $p = 1 - t^2 + 2t^3$ nun den Koordinatenvektor $\xi_{\underline{q}}(p)$ zu bestimmen, entwickeln wir p in der Basis $\underline{q}$:

$$1 - t^2 + 2t^3 = \frac{1}{2}(1-t) + \frac{1}{2}(1+t) - t^2 + 2t^3.$$

Dadurch erhalten wir

$$\xi_{\underline{q}}(p) = \left(\tfrac{1}{2}, \tfrac{1}{2}, -1, 2\right).$$

Auch hier kann man den Isomorphismus explizit hinschreiben:

$$\begin{aligned} K^4 &\leftrightarrow V \\ (c_0, c_1, c_2, c_3)^t &\mapsto (c_0 + c_1) + (c_1 - c_0)t - c_2 t^2 + c_3 t^3 \\ \left(\frac{c_0 - c_1}{2}, \frac{c_0 + c_1}{2}, c_2, c_3\right)^t &\leftarrow\!\shortmid \sum_{i=0}^{3} c_i t^i. \end{aligned}$$

(iii) Wählt man in $V = K^m$ die Standardbasis $\underline{e} = (e_1, \ldots, e_m)$, so ist der Koordinatenvektor eines Tupels $v = (v_1, \ldots, v_m)^t \in K^m$ das Tupel v selbst:

$$\xi_{\underline{e}}(v) = v.$$

Aber natürlich kann man auch eine andere Basis $\underline{v}$ wählen und erhält damit einen Isomorphismus

$$\xi_{\underline{v}} \colon K^m \to K^m,$$

der nicht die Identität ist. Den Übergang von $v \in K^m$ zu $\xi_{\underline{v}}(v) \in K^m$ nennt man auch **Basiswechsel**. Wählen wir in $\mathbb{Q}^2$ beispielsweise die Basis

$$\underline{v} = \left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right)$$

so ergibt sich für $v = (a, b)^t \in \mathbb{Q}^2$

$$\xi_{\underline{v}}(v) = \begin{pmatrix} \frac{a+b}{2} \\ \frac{a-b}{2} \end{pmatrix}. \qquad \triangle$$

Satz 3.2.9. *Seien V und W zwei K-Vektorräume und $(v_1, \ldots, v_m)$ eine Basis von V. Dann gibt es für jede Wahl von $w_1, \ldots, w_m \in W$ genau eine lineare Abbildung $\varphi\colon V \to W$ mit*

$$\varphi(v_i) = w_i \quad \textit{für } i = 1, \ldots, m.$$

Beweis. Die Existenz beweisen wir, indem wir φ explizit angeben. Dazu entwickeln wir jedes $v \in V$ bezüglich der Basis $(v_1, \ldots, v_m)$ als

$$v = \sum_{i=1}^{m} \lambda_i v_i$$

und definieren

$$\varphi(v) := \sum_{i=1}^{m} \lambda_i w_i.$$

In kurzer Schreibweise mit Matrixmultiplikation könnte man dafür auch schreiben

$$\varphi(v) := \underline{w} \cdot \xi_{\underline{v}}(v).$$

Da $\xi_{\underline{v}}$ eine lineare Abbildung ist und die Matrixmultiplikation das Distributivgesetz erfüllt, definiert das eine lineare Abbildung mit der gewünschten Eigenschaft.

Für die Eindeutigkeit seien φ, φ' zwei solche linearen Abbildungen. Wir berechnen

$$\varphi(v) = \varphi\left(\sum_{i=1}^{m} \lambda_i v_i\right) = \sum_{i=1}^{m} \lambda_i \varphi(v_i) = \sum_{i=1}^{m} \lambda_i w_i = \sum_{i=1}^{m} \lambda_i \varphi'(v_i) = \varphi'(v),$$

wobei wir für die zweite und letzte Gleichheit die Linearität von φ bzw. φ' benutzt haben. Das zeigt $\varphi = \varphi'$. $\square$

Korollar 3.2.10. *Für jede lineare Abbildung $\varphi\colon K^m \to K^n$ gibt es genau eine Matrix $A \in \mathrm{Mat}_{n,m}(K)$ mit*

$$\varphi = \mu_A.$$

Die Zuordnung

$$\begin{aligned} \mu\colon \mathrm{Mat}_{n,m}(K) &\to \mathrm{Lin}_K(K^m, K^n) \\ A &\mapsto \mu_A \end{aligned}$$

ist ein Isormorphismus von K-Vektorräumen.

Beweis. Wir wählen die Standardbasis $\underline{e} = (e_1, \dots, e_m)$ von K^m, setzen $w_i := \varphi(e_i)$ und schreiben sie in die Spalten einer Matrix:

$$A = (w_1, \dots, w_m) \in \mathrm{Mat}_{n,m}(K).$$

Für jedes $i = 1, \dots, m$ gilt dann

$$\mu_A(e_i) = Ae_i = w_i = \varphi(e_i).$$

Die Eindeutigkeit aus Satz 3.2.9 zeigt $\mu_A = \varphi$. Die Eindeutigkeit von A ist klar. Damit ist gerade die Bijektivität der Abbildung μ gezeigt und die Linearität ist Bemerkung 3.2.4 (ii). □

Beispiel 3.2.11. (i) Wir betrachten die lineare Abbildung

$$\begin{aligned} \varphi\colon K^3 &\to K^2 \\ \begin{pmatrix} a \\ b \\ c \end{pmatrix} &\mapsto \begin{pmatrix} a-2c \\ a+b \end{pmatrix}. \end{aligned}$$

Es gilt

$$\varphi(e_1) = \begin{pmatrix} 1 \\ 1 \end{pmatrix},\ \varphi(e_2) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \varphi(e_3) = \begin{pmatrix} -2 \\ 0 \end{pmatrix}$$

und deshalb gilt für

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

gerade $\varphi = \mu_A$.

(ii) Wir betrachten eine Drehung des $\mathbb{R}^2$ um den Ursprung, gegen den Uhrzeigersinn um den Winkel $\theta \in [0, 2\pi)$. Es handelt sich dabei um eine lineare Abbildung

$$d_\theta\colon \mathbb{R}^2 \to \mathbb{R}^2,$$

wie man sich geometrisch leicht überzeugt. Dabei wird e_1 auf $(\cos(\theta), \sin(\theta))^t$ und e_2 auf $(-\sin(\theta), \cos(\theta))^t$ abgebildet. Die **Drehmatrix**

$$D_\theta := \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \in \mathrm{Mat}_2(\mathbb{R})$$

beschreibt also die Drehung, d.h. es gilt $d_\theta = \mu_{D_\theta}$. △

3.2.2 Die Dimensionsformel

Definition 3.2.12. Seien V, W zwei K-Vektorräume und $\varphi\colon V \to W$ eine lineare Abbildung. Dann heißt

$$\mathrm{Kern}(\varphi) := \varphi^{-1}(0) = \{v \in V \mid \varphi(v) = 0\}$$

der **Kern** von φ.

Bemerkung 3.2.13. Für eine Matrix $A \in \mathrm{Mat}_{m,n}(K)$ gilt offensichtlich

$$\mathrm{Kern}\,(\mu_A) = L(A, 0). \qquad \triangle$$

Lemma 3.2.14. *Seien V, W zwei K-Vektorräume und $\varphi\colon V \to W$ eine lineare Abbildung. Dann gilt*

(i) Es sind $\mathrm{Kern}(\varphi) \subseteq V$ und $\mathrm{Bild}(\varphi) \subseteq W$ jeweils Untervektorräume.

(ii) Es ist φ genau dann injektiv, wenn $\mathrm{Kern}(\varphi) = \{0\}$ gilt.

Beweis. (i) Die Unterraumaxiome rechnet man unter Zuhilfenahme der Linearität direkt nach, zum Beispiel zeigt Lemma 5.2.3 dass stets $0 \in \mathrm{Kern}(\varphi)$ und $0 \in \mathrm{Bild}(\varphi)$ gilt. Weiter gilt beispielsweise

$$\begin{aligned} v_1, v_2 \in \mathrm{Kern}(\varphi) &\Rightarrow \varphi(v_1) = \varphi(v_2) = 0 \\ &\Rightarrow \varphi(v_1 + v_2) = \varphi(v_1) + \varphi(v_2) = 0 + 0 = 0 \\ &\Rightarrow v_1 + v_2 \in \mathrm{Kern}(\varphi). \end{aligned}$$

$$\begin{aligned} w_1, w_2 \in \mathrm{Bild}(\varphi) &\Rightarrow \exists v_1, v_2 \in V\colon \varphi(v_1) = w_1, \varphi(v_2) = w_2 \\ &\Rightarrow \varphi(v_1 + v_2) = \varphi(v_1) + \varphi(v_2) = w_1 + w_2 \\ &\Rightarrow w_1 + w_2 \in \mathrm{Bild}(\varphi). \end{aligned}$$

In (ii) ist "$\Rightarrow$" klar, denn $0 \in \mathrm{Kern}(\varphi)$ mitsamt der Injektivität impliziert, das kein weiteres Element im Kern enthalten sein kann. Für "$\Leftarrow$" nehmen wir $\varphi(v_1) = \varphi(v_2)$ für gewisse $v_1, v_2 \in V$ an. Daraus folgt

$$\varphi(v_1 - v_2) = \varphi(v_1) - \varphi(v_2) = 0,$$

also $v_1 - v_2 \in \mathrm{Kern}(\varphi) = \{0\}$, also $v_1 = v_2$. $\square$

Beispiel 3.2.15. (i) Wir betrachten die lineare Abbildung

$$\begin{aligned}\varphi_1\colon \mathbb{R}^2 &\to \mathbb{R}^1\\ (a,b)^t &\mapsto a-b.\end{aligned}$$

Es ist

$$\mathrm{Kern}(\varphi_1) = \left\{(a,b)^t \in \mathbb{R}^2 \mid a=b\right\} = \{(a,a) \mid a \in \mathbb{R}\}$$

genau die Diagonale in $\mathbb{R}^2$, ein eindimensionaler Untervektorraum. Insbesondere ist φ_1 nicht injektiv. Es gilt $\mathrm{Bild}(\varphi_1) = \mathbb{R}^1$.
(ii) Wir betrachten die lineare Abbildung

$$\begin{aligned}\varphi_2\colon \mathbb{R}^2 &\to \mathbb{R}^2\\ (a,b)^t &\mapsto (a-b,a+b)^t.\end{aligned}$$

Es gilt

$$\mathrm{Kern}(\varphi_2) = \left\{(a,b)^t \in \mathbb{R}^2 \mid a-b=a+b=0\right\} = \{0\},$$

also ist φ_2 injektiv.

Satz 3.2.16 (Dimensionsformel für lineare Abbildungen)**.** *Sei V ein endlich-dimensionaler K-Vektorraum, W ein weiterer K-Vektorraum und $\varphi\colon V \to W$ eine lineare Abbildung. Dann gilt*

$$\dim_K(V) = \dim_K(\mathrm{Kern}(\varphi)) + \dim_K(\mathrm{Bild}(\varphi))\,.$$

Beweis. Seien zunächst $w_1,\dots,w_n \in \mathrm{Bild}(\varphi)$ linear unabhängig. Wir wählen $v_1,\dots,v_n \in V$ mit $\varphi(v_i) = w_i$ und zeigen, dass dann auch $v_1,\dots,v_n$ linear unabhängig sind. Sei dazu $0 = \sum_{i=1}^n \lambda_i v_i$ angenommen. Daraus folgt

$$0 = \varphi(0) = \varphi\left(\sum_{i=1}^n \lambda_i v_i\right) = \sum_{i=1}^n \lambda_i \varphi(v_i) = \sum_{i=1}^n \lambda_i w_i.$$

Aus der linearen Unabhängigkeit von $w_1,\dots,w_n$ folgt also $\lambda_i = 0$ für alle $i = 1,\dots,n$, die gewünschte Aussage. Insbesondere haben wir damit

$$\dim_K(\mathrm{Bild}(\varphi)) \le \dim_K(V) < \infty$$

gezeigt. Wir nehmen nun an, dass $(w_1,\dots,w_n)$ sogar eine Basis von $\mathrm{Bild}(\varphi)$ ist und wählen zusätzlich eine Basis $(u_1,\dots,u_m)$ von $\mathrm{Kern}(\varphi)$. Wir zeigen nun, dass dann

$$(v_1,\dots,v_n,u_1,\dots,u_m)$$

eine Basis von V bildet. Für die Eigenschaft *Erzeugendensystem* sei $v \in V$ beliebig gewählt. Wir finden dann $\lambda_1, \ldots, \lambda_n \in K$ mit

$$\varphi(v) = \lambda_1 w_1 + \cdots + \lambda_n w_n.$$

Für

$$v' := v - (\lambda_1 v_1 + \cdots + \lambda_n v_n) \in V$$

gilt dann

$$\varphi(v') = \varphi(v) - (\lambda_1 \varphi(v_1) + \cdots + \lambda_n \varphi(v_n)) = \varphi(v) - (\lambda_1 w_1 + \cdots + \lambda_n w_n) = 0,$$

also $v' \in \mathrm{Kern}(\varphi) = \mathrm{Span}_K(\{u_1, \ldots, u_m\})$ und damit

$$v = v' + \sum_{i=1}^{n} \lambda_i v_i \in \mathrm{Span}(\{v_1, \ldots, v_n, u_1, \ldots, u_m\}).$$

Für die lineare Unabhängigkeit nehmen wir

$$0 = \sum_{i=1}^{n} \lambda_i v_i + \sum_{i=1}^{m} \gamma_i u_i$$

an. Dadurch erhalten wir

$$\begin{aligned} 0 = \varphi(0) = \varphi\left(\sum_{i=1}^{n} \lambda_i v_i + \sum_{i=1}^{m} \gamma_i u_i\right) &= \sum_{i=1}^{n} \lambda_i \varphi(v_i) + \sum_{i=1}^{m} \gamma_i \underbrace{\varphi(u_i)}_{=0} \\ &= \sum_{i=1}^{n} \lambda_i w_i. \end{aligned}$$

Aus der linearen Unabhängigkeit der w_i erhalten wir also $\lambda_i = 0$ für alle $i = 1, \ldots, n$. Setzt man das in der oberen Linearkombination ein, erhält man anschließend aus der linearen Unabhängigkeit der u_i auch $\gamma_i = 0$ für alle $i = 1, \ldots, m$. Damit ist also gezeigt, dass $(v_1, \ldots, v_n, u_1, \ldots, u_m)$ eine Basis von V bildet. Nun ergibt sich

$$\dim_K(V) = n + m = \dim_K(\mathrm{Bild}(\varphi)) + \dim_K(\mathrm{Kern}(\varphi)). \qquad \square$$

Korollar 3.2.17. *Seien V, W zwei endlich-dimensionale K-Vektorräume mit*

$$\dim_K(V) = \dim_K(W)$$

sowie $\varphi\colon V \to W$ eine lineare Abbildung. Dann sind äquivalent:

(*i*) φ *ist injektiv.*

(*ii*) φ *ist surjektiv.*

(*iii*) φ *ist bijektiv.*

Beweis. "(*i*)⇒(*ii*)": Aus der Injektivität folgt $\mathrm{Kern}(\varphi) = \{0\}$ und mit der Dimensionsformel erhalten wir

$$\dim_K(W) = \dim_K(V) = \underbrace{\dim_K(\mathrm{Kern}(\varphi))}_{=0} + \dim_K(\mathrm{Bild}(\varphi)) = \dim_K(\mathrm{Bild}(\varphi)).$$

Aus Proposition 3.1.23 folgt damit $\mathrm{Bild}(\varphi) = W$, d.h φ ist surjektiv.
"(*ii*)⇒(*i*)" benutzt genau dieselbe Formel. Wenn man diesmal

$$\dim_K(\mathrm{Bild}(\varphi)) = \dim_K(W)$$

benutzt, folgt direkt $\dim_K(\mathrm{Kern}(\varphi)) = 0$. Das bedeutet $\mathrm{Kern}(\varphi) = \{0\}$ und das impliziert die Injektivität.
Damit implizieren aber sowohl (*i*) als auch (*ii*) immer schon (*iii*) und die Umkehrung ist sowieso klar. □

Hier sehen wir nun, warum die Sichtweise einer Matrix als lineare Abbildung sehr hilfreich sein kann:

Korollar 3.2.18. *Für eine quadratische Koeffizientenmatrix $A \in \mathrm{Mat}_m(K)$ sind äquivalent:*

(*i*) $L(A, 0) = \{0\}$.

(*ii*) *Für jedes $b \in K^m$ besitzt das lineare Gleichungssystem mit erweiterter Koeffizientenmatrix (A, b) eine Lösung.*

(*iii*) A *ist invertierbar.*

Beweis. Offensichtlich ist (*i*) äquivalent to $\mathrm{Kern}(\mu_A) = \{0\}$, also zur Injektivität von μ_A. Weiter ist (*ii*) offensichtlich äquivalent zur Surjektivität von μ_A. Damit folgt "(*i*)⇔(*ii*)" direkt aus Korollar 3.2.17. In diesen Fällen ist μ_A also ein Isomorphismus und die Umkehrabbildung ist nach Korollar 3.2.10 wieder durch eine Matrix $B \in \mathrm{Mat}_m(K)$ beschrieben:

$$\mu_A^{-1} = \mu_B.$$

Aus

$$\mu_{BA} = \mu_B \circ \mu_A = \mathrm{id}_{K^m} = \mu_A \circ \mu_B = \mu_{AB}$$

folgt $AB = BA = I_m$, also (*iii*). Die Aussage "(*iii*)⇒(*i*)" ist offensichtlich. □

3.2.3 Der Rang einer Matrix

Definition 3.2.19. Für $A \in \mathrm{Mat}_{m,n}(K)$ definieren wir den **Rang** als

$$\mathrm{rang}(A) := \dim_K(\mathrm{Bild}(\mu_A)). \qquad \triangle$$

Bemerkung 3.2.20. (i) Der Rang von A ist die Dimension des Bildes der von A definierten Abbildung $\mu_A\colon K^n \to K^m$. Laut Beispiel 3.1.11 ist das Bild von μ_A genau der von den Spalten von A aufgespannte Unterraum von K^m. Der Rang von A ist also genau die *maximale Anzahl linear unabhängiger Spalten* von A.
(ii) Der Rang ist ein Maß für die Surjektivität der Abbildung μ_A. Sie ist genau dann surjektiv, wenn $\mathrm{rang}(A) = m$ gilt. Genau dann ist auch das lineare Gleichungssystem (A, b) für jedes $b \in K^m$ lösbar.
(iii) Die Dimensionsformel für μ_A, zusammen mit der Beobachtung

$$\mathrm{Kern}(\mu_A) = L(A, 0)$$

besagt also

$$\dim(L(A,0)) + \mathrm{rang}(A) = n.$$

(iv) Bei elementaren Zeilenumformungen ändert sich der Rang von A nicht. Solche Umformungen entsprechen ja der Multiplikation von links mit invertierbaren Matrizen. In der Sprache der linearen Abbildungen führen wir also nach μ_A noch einen Isomorphismus aus. Das ändert laut Dimensionsformel die Dimension des Bildes nicht. Für eine Matrix in (reduzierter) Zeilenstufenform ist der Rang aber offensichtlich wegen (i) gerade die Anzahl der Pivots! Das liefert also einen Algorithmus zur Bestimmung des Rangs einer Matrix. $\triangle$

Beispiel 3.2.21. (i) Wir betrachten

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{pmatrix} \in \mathrm{Mat}_{3,4}(\mathbb{Q})$$

und wollen $\mathrm{rang}(A)$ bestimmen. Dazu bringen wir A auf Zeilenstufenform

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

und finden $\mathrm{rang}(A) = 2$. Das Bild der linearen Abbildung $\mu_A\colon \mathbb{Q}^4 \to \mathbb{Q}^3$ ist also ein zweidimensionaler Unterraum.

(ii) Um die endlichen Körper nicht völlig zu vergessen, betrachten wir dieselbe Matrix diesmal über $\mathbb{Z}/2\mathbb{Z}$. Dabei vereinfacht sich A natürlich zu

$$A = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \end{pmatrix} \in \mathrm{Mat}_{3,4}\left(\mathbb{Z}/2\mathbb{Z}\right).$$

Ganz ohne Rechnung sehen wir, dass die Spalten einen eindimensionalen Raum aufspannen. Es gilt also hier $\mathrm{rang}(A) = 1$ und die Abbildung

$$\mu_A \colon\ (\mathbb{Z}/2\mathbb{Z})^4 \to (\mathbb{Z}/2\mathbb{Z})^3$$

hat eindimensionales Bild. △

Satz 3.2.22. *Für $A \in \mathrm{Mat}_{m,n}(K)$ gilt*

$$\mathrm{rang}(A) = \mathrm{rang}(A^t) \leq \min\{m, n\}.$$

Beweis. Ist A in (reduzierter) Zeilenstufenform, so gilt die Aussage offensichtlich; sowohl die Pivotspalten als auch -zeilen sind jeweils maximal linear unabhängig unter den Spalten bzw. Zeilen. Wir haben außerdem schon gesehen, dass sich der Rang einer Matrix durch elementare Zeilenumformungen nicht ändert. Ist $T \in \mathrm{GL}_m(K)$ also eine Matrix, die A durch Multiplikation von links auf Zeilenstufenform Z bringt, so gilt

$$\mathrm{rang}(A) = \mathrm{rang}(TA) = \mathrm{rang}(Z) = \mathrm{rang}(Z^t) = \mathrm{rang}(A^tT^t) = \mathrm{rang}(A^t).$$

Die letzte Gleichung verwendet, dass auch Multiplikation von rechts mit einer invertierbaren Matrix (also Ausführung eines Isomorphismus vor μ_A) den Rang nicht ändert. Und natürlich ist mit T auch T^t invertierbar. □

Bemerkung 3.2.23. Die letzte Aussage besagt, dass der von den Spalten bzw. der von den Zeilen einer Matrix aufgespannte Raum stets dieselbe Dimension hat. Das ist umso bemerkenswerter, als es sich dabei um Unterräume in K^m bzw. K^n handelt, wobei $m \neq n$ gelten kann. △

Definition 3.2.24. Sei $A \in \mathrm{Mat}_{m,n}(K)$ eine Matrix. Eine **Untermatrix** von A ist eine Matrix B, die durch Streichung gewisser Zeilen und Spalten von A entsteht. △

Beispiel 3.2.25. Durch Streichung der ersten und dritten Spalte sowie der zweiten Zeile entsteht aus

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{pmatrix} \in \mathrm{Mat}_{3,4}(\mathbb{Q})$$

die Untermatrix

$$B = \begin{pmatrix} 2 & 4 \\ 10 & 12 \end{pmatrix} \in \mathrm{Mat}_2(\mathbb{Q}).$$

Proposition 3.2.26. *Für $A \in \mathrm{Mat}_{m,n}(K)$ ist $\mathrm{rang}(A)$ gleich der größten Größe einer invertierbaren Untermatrix von A.*

Beweis. Sei $r = \mathrm{rang}(A)$ und $B \in \mathrm{Mat}_{m,r}(K)$ eine Untermatrix von A mit linear unabhängigen Spalten. Dann gilt

$$r = \mathrm{rang}(B) = \mathrm{rang}(B^t)$$

und wir können eine Untermatrix $C \in \mathrm{Mat}_r(K)$ von B finden, die linear unabhängige Zeilen hat. Aus

$$r = \mathrm{rang}(C^t) = \mathrm{rang}(C)$$

folgt dann, dass C auch immer noch linear unabhängige Spalten hat. Mit Proposition 3.1.21 ist C also invertierbar.
Sei umgekehrt $C \in \mathrm{Mat}_s(K)$ eine invertierbare Untermatrix von A. Die Spalten von C sind also linear unabhängig. Also sind auch s Spalten von A linear unabhängig und das impliziert $s \leq \mathrm{rang}(A)$. $\square$

3.2.4 Darstellungsmatrizen ★

In vielen Fällen ist es sinnvoll, sich eine lineare Abbildung nicht unbedingt als Matrix vorzustellen, sondern einfach als Abbildung mit der Linearitätseigenschaft. Man denke dabei an Beispiel 3.2.2 (iii), in dem die Differentiation als lineare Abbildung aufgefasst wurde. Dabei waren Definitions- und Zielraum nicht explizit ein K^m und die Abbildung nicht von der Gestalt μ_A.
Andererseits möchte man manchmal explizite Berechnungen für eine lineare Abbildung durchführen, was mit einer Matrix deutlich leichter geht. Die Kombination von Satz 3.2.5 und Korollar 3.2.10 besagt, dass das auch immer möglich ist.

Konstruktion 3.2.27. Seien V, W zwei K-Vektorräume, jeweils mit einer fixierten Basis $\underline{v} = (v_1, \ldots, v_n)$ und $\underline{w} = (w_1, \ldots, w_m)$. Wir betrachten dazu die in Satz 3.2.5 konstruierten Isormophismen

$$\varphi_{\underline{v}} \colon K^n \to V \quad \text{und} \quad \varphi_{\underline{w}} \colon K^m \to W.$$

Sei nun $\varphi \colon V \to W$ eine lineare Abbildung. Dann ist die Abbildung

$$\varphi_{\underline{w}}^{-1} \circ \varphi \circ \varphi_{\underline{v}} \colon K^n \to K^m$$

ebenfalls linear und nach Korollar 3.2.10 gibt es (genau) eine Matrix $A \in \mathrm{Mat}_{m,n}(K)$ mit

$$\varphi_{\underline{w}}^{-1} \circ \varphi \circ \varphi_{\underline{v}} = \mu_A :$$

$$\begin{tikzcd} V \arrow[r, "\varphi"] & W \arrow[d, "\varphi_{\underline{w}}^{-1} = \xi_{\underline{w}}"] \\ K^n \arrow[u, "\varphi_{\underline{v}}"] \arrow[r, "\mu_A"] & K^m \end{tikzcd}$$

Diese Matrix A nennt man die **Darstellungsmatrix** von φ bezüglich der Basen $\underline{v}, \underline{w}$ und schreibt dafür auch

$$A = M_{\underline{v},\underline{w}}(\varphi).$$

Die i-te Spalte von A ist dabei offensichtlich genau

$$Ae_i = \mu_A(e_i) = (\varphi_{\underline{w}}^{-1} \circ \varphi \circ \varphi_{\underline{v}})(e_i) = \xi_{\underline{w}}\left(\varphi(v_i)\right).$$

Das liefert eine Methode zur Konstruktion der Darstellungsmatrix: man bildet v_i anhand von φ ab und entwickelt das Ergebnis bezüglich der Basis $\underline{w}$. Die Koeffizienten bilden die i-te Spalte von $M_{\underline{v},\underline{w}}(\varphi)$. $\triangle$

Beispiel 3.2.28. (i) Wir betrachten $V = \mathbb{Q}[t]_{\leq 3}, W = \mathbb{Q}[t]_{\leq 2}$ und die Differentiation

$$\partial \colon V \to W.$$

Wir versehen V mit der Basis $\underline{v} = (1, t, t^2, t^3)$ und W mit $\underline{w} = (1, t, t^2)$. Es gilt

$$\partial(t^i) = it^{i-1}$$

und das ist direkt die Darstellung bezüglich der Basis $\underline{w}$. Damit ergibt sich

$$M_{\underline{v},\underline{w}}(\partial) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

(ii) Wir wählen V und W wie oben, diesmal aber mit den Basen

$$\underline{v}' = (1+t, 1-t, t^2, t^3) \quad \text{und} \quad \underline{w}' = (1, t-1, t^2+1).$$

Wir berechnen

$$\begin{aligned} \partial(v_1') &= \partial(1+t) = 1 = w_1' \\ \partial(v_2') &= \partial(1-t) = -1 = -w_1' \\ \partial(v_2') &= \partial(t^2) = 2t = 2w_1' + 2w_2' \\ \partial(v_3') &= \partial(t^3) = 3t^2 = -3w_1' + 3w_3' \end{aligned}$$

und erhalten somit

$$M_{\underline{v}',\underline{w}'}(\partial) = \begin{pmatrix} 1 & -1 & 2 & -3 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

(iii) Wir wählen $V = W = \mathrm{Mat}_2(K)$ und in beiden Räumen dieselbe Basis

$$\underline{v} = \left(\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right).$$

Für zwei fest gewählte Matrizen

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \ B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \in \mathrm{Mat}_2(K)$$

betrachten wir dann die lineare Abbildung

$$\begin{aligned} \varphi\colon \mathrm{Mat}_2(K) &\to \mathrm{Mat}_2(K) \\ X &\mapsto AXB. \end{aligned}$$

Um die Darstellungmatrix $M_{\underline{v},\underline{v}}(\varphi)$ zu bestimmen berechnen wir

$$\begin{aligned} \varphi(v_1) = \varphi\left(\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \right) &= \begin{pmatrix} a_{11}b_{11} & a_{11}b_{12} \\ a_{21}b_{11} & a_{21}b_{12} \end{pmatrix} \\ &= a_{11}b_{11}v_1 + a_{11}b_{12}v_2 + a_{21}b_{11}v_3 + a_{21}b_{12}v_4. \end{aligned}$$

Die erste Spalte der Darstellungsmatrix ist also

$$(a_{11}b_{11}, a_{11}b_{12}, a_{21}b_{11}, a_{21}b_{12})^t \in K^4.$$

Führt man dieselbe Berechnung auch noch für v_2, v_3, v_4 durch erhält man insgesamt

$$M_{\underline{v},\underline{v}}(\varphi) = \begin{pmatrix} a_{11}b_{11} & a_{11}b_{21} & a_{12}b_{11} & a_{12}b_{21} \\ a_{11}b_{12} & a_{11}b_{22} & a_{12}b_{12} & a_{12}b_{22} \\ a_{21}b_{11} & a_{21}b_{21} & a_{22}b_{11} & a_{22}b_{21} \\ a_{21}b_{12} & a_{21}b_{22} & a_{22}b_{12} & a_{22}b_{22} \end{pmatrix}.$$

(iv) Natürlich kann man auch $V = K^n$ und $W = K^m$ benutzen, sowie eine Abbildung $\varphi = \mu_B$, die bereits durch eine Matrix gegeben ist. Je nach Wahl der Basen $\underline{v}$ und $\underline{w}$ erhält man trotzdem unterschiedliche Darstellungsmatrizen für μ_B :

$$\begin{tikzcd} V = K^n \arrow[r, "\mu_B"] & K^m = W \arrow[d, "\varphi_{\underline{w}}^{-1} = \xi_{\underline{w}}"] \\ K^n \arrow[u, "\varphi_{\underline{v}}"] \arrow[r, "\mu_A"] & K^m. \end{tikzcd}$$

Die Standardbasen $\underline{v} = \underline{e}$ und $\underline{w} = \underline{e}$ von K^n und K^m liefern als Isomorphismen $\varphi_{\underline{e}}$ beispielsweise die Identität und somit

$$M_{\underline{e},\underline{e}}(\mu_B) = B.$$

Wir können aber beispielsweise mit

$$B = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \in \mathrm{Mat}_{2,3}(\mathbb{Q})$$

die Abbildung

$$\begin{aligned} \mu_B \colon \mathbb{Q}^3 &\to \mathbb{Q}^2 \\ (a,b,c)^t &\mapsto (a+b+c, a+b+c) \end{aligned}$$

betrachten und in $V = \mathbb{Q}^3$ die Basis

$$\underline{v} = \big((1,-1,0)^t, (-1,0,1)^t, (0,0,-1)^t\big)$$

sowie in $W = \mathbb{Q}^2$ die Basis

$$\underline{w} = \big((-1,0)^t, (1,1)^t\big)$$

wählen. Wie vorhin berechnen wir die Darstellungsmatrix:

$$\begin{aligned}\mu_B(v_1) &= (0,0)^t = 0w_1 + 0w_2\\ \mu_B(v_2) &= (0,0)^t = 0w_1 + 0w_2\\ \mu_B(v_3) &= (-1,-1)^t = -w_2\end{aligned}$$

und damit

$$M_{\underline{v},\underline{w}}(\mu_B) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$ △

Proposition 3.2.29. *Seien V, W, X drei K-Vektorräume mit Basen $\underline{v}, \underline{w}, \underline{x}$. Seien $\varphi\colon V \to W$ und $\psi\colon W \to X$ lineare Abbildungen. Dann gilt*

$$M_{\underline{v},\underline{x}}(\psi \circ \varphi) = M_{\underline{w},\underline{x}}(\psi) \cdot M_{\underline{v},\underline{w}}(\varphi).$$

Beweis. Folgt direkt aus der Definition der Darstellungsmatrizen und Beispiel 3.2.2 (ii). □

Bemerkung 3.2.30. (i) Die Darstellungsmatrix $A = M_{\underline{v},\underline{w}}(\varphi)$, zusammen mit den Basen $\underline{v}, \underline{w}$, bestimmt die ursprüngliche Abbildung φ eindeutig; es gilt ja

$$\varphi = \varphi_{\underline{w}} \circ \mu_A \circ \varphi_{\underline{v}}^{-1}.$$

Beim Übergang zu einer Darstellungsmatrix geht also keine Information verloren. Gleichzeitig kann man manchmal eine Matrix leichter für Rechnungen verwenden als eine lineare Abbildung. Möchte man beispielsweise die Dimension von $\mathrm{Bild}(\varphi)$ bestimmen, so kann man stattdessen auch den Rang einer beliebigen Darstellungsmatrix von φ berechnen. Wir haben schon gesehen, dass man das leicht mit dem Gauß-Algorithmus tun kann.
(ii) Selbst wenn man eigentlich mit einer gegebenen Matrix B rechnen will, kann man theoretisch zunächst eine andere Darstellungsmatrix von μ_B bestimmen, mit der sich vielleicht leichter rechnen lässt. In Beispiel 3.2.28 (iv) ist $M_{\underline{v},\underline{w}}(\mu_B)$ eher einfacher als B.
(iii) Grundsätzlich lohnt es sich, bei der Wahl von Basen $\underline{v}, \underline{w}$ Mühe zu investieren, um eine Darstellungsmatrix von besonders einfacher Form zu erhalten. Damit werden wir uns im Rest der Vorlesung noch intensiver beschäftigen. Ein erstes Ergebnis dazu ist der folgende Satz. △

Satz 3.2.31. *Seien V, W zwei K-Vektorräume mit $\dim_K(V) = n, \dim_K(W) = m$, sowie $\varphi\colon V \to W$ eine lineare Abbildung. Dann gibt es Basen $\underline{v}, \underline{w}$ von V bzw. W mit*

$$M_{\underline{v},\underline{w}}(\varphi) = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} \in \mathrm{Mat}_{m,n}(K).$$

Dabei ist $r = \mathrm{rang}\left(M_{\underline{v},\underline{w}}(\varphi)\right) = \dim_K\left(\mathrm{Bild}(\varphi)\right)$.

Beweis. Zunächst wählen wir eine Basis $(u_1, \ldots, u_s)$ von $\mathrm{Kern}(\varphi)$ und ergänzen sie zu einer Basis

$$\underline{v} = (v_1, \ldots, v_t, u_1, \ldots, u_s)$$

von V. Wir zeigen zunächst, dass dann $\varphi(v_1), \ldots, \varphi(v_t)$ in W linear unabhängig sind. Dazu sei

$$0 = \lambda_1\varphi(v_1) + \cdots + \lambda_t\varphi(v_t) = \varphi\left(\lambda_1 v_1 + \cdots + \lambda_t v_t\right)$$

mit gewissen $\lambda_i \in K$ angenommen. Das bedeutet $\lambda_1 v_1 + \cdots + \lambda_t v_t \in \mathrm{Kern}(\varphi)$. Es gibt also eine Darstellung

$$\lambda_1 v_1 + \cdots + \lambda_t v_t = \gamma_1 u_1 + \cdots + \gamma_s u_s$$

und aus der linearen Unabhängigkeit von $(v_1, \ldots, v_t, u_1, \ldots, u_s)$ folgt direkt $\lambda_1 = \cdots = \lambda_t = 0$, die gewünschte Aussage.
Wir ergänzen nun $\varphi(v_1), \ldots, \varphi(v_t)$ zu einer Basis

$$\underline{w} = (\varphi(v_1), \cdots, \varphi(v_t), w_1, \ldots, w_k)$$

von W. Bezüglich der Basen $\underline{v}, \underline{w}$ hat die Darstellungsmatrix nun ganz offensichtlich die gewünschte Form. $\square$

Satz 3.2.32. *Für Matrizen $A, B \in \mathrm{Mat}_{m,n}(K)$ sind äquivalent:*

(i) *A und B sind beides Darstellungsmatrizen derselben linearen Abbildung, nur bezüglich (eventuell) verschiedener Basen.*

(ii) *Es gibt $P \in \mathrm{GL}_m(K)$ und $Q \in \mathrm{GL}_n(K)$ mit $PAQ = B$.*

(iii) $\mathrm{rang}(A) = \mathrm{rang}(B)$.

Beweis. Für "$(i)\Rightarrow(ii)$" liefert (i) die Existenz eines kommutativen Diagrams der folgenden Form:

$$\begin{tikzcd} K^n \arrow[r, "\mu_B"] \arrow[d, "\varphi_{\underline{v}'}"'] & K^m \\ V \arrow[r, "\varphi"] \arrow[d, "\xi_{\underline{v}}"'] & W \arrow[u, "\xi_{\underline{w}'}"'] \\ K^n \arrow[r, "\mu_A"] & K^m \arrow[u, "\varphi_{\underline{w}}"'] \end{tikzcd}$$

Es gilt also

$$\mu_B = \underbrace{(\xi_{\underline{w}'} \circ \varphi_{\underline{w}})}_{\mu_P} \circ \mu_A \circ \underbrace{(\xi_{\underline{v}} \circ \varphi_{\underline{v}'})}_{\mu_Q} = \mu_{PAQ}$$

wobei P und Q die entsprechenden Matrizen gemäß Korollar 3.2.10 sind, die für Isomorphismen natürlich invertierbar sind. Die Eindeutigkeit impliziert $B = PAQ$.
"$(ii)\Rightarrow(iii)$" erhält man offensichtlich aus der Tatsache, dass Isomorphismen vor oder nach einer linearen Abbildung deren Bilddimension nicht ändern.
Für "$(iii)\Rightarrow(i)$" wenden wir Satz 3.2.31 auf μ_A und μ_B an, mit $r = \mathrm{rang}(A) = \mathrm{rang}(B)$ und

$$J = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}:$$

$$\begin{tikzcd} K^n \arrow[r, "\mu_A"] & K^m \\ K^n \arrow[r, "\mu_J"] \arrow[u, "\varphi_{\underline{v}}"] \arrow[d, "\varphi_{\underline{v}'}"'] & K^m \arrow[u, "\varphi_{\underline{w}}"'] \arrow[d, "\varphi_{\underline{w}'}"] \\ K^n \arrow[r, "\mu_B"] & K^m. \end{tikzcd}$$

Es sind also A und B beides Darstellungsmatrizen von μ_J. □

Definition 3.2.33. Zwei Matrizen $A, B \in \mathrm{Mat}_{m,n}$ heißen **äquivalent**, wenn sie die Bedingungen aus Satz 3.2.32 erfüllen. △

3.2.5 Der Dualraum ★

Definition 3.2.34. Sei V ein K-Vektorraum. Dann heißt

$$V' := \mathrm{Lin}_K(V, K^1)$$

der **Dualraum** von V. Die Elemente von V' nennt man **Funktionale** oder **Linearformen** auf V. △

Bemerkung/Beispiel 3.2.35. (i) Nach Bemerkung 3.2.4 ist V' selbst wieder ein K-Vektorraum, mit punktweise definierten Verknüpfungen.
(ii) Nach Korollar 3.2.10 gilt

$$(K^m)' \cong \mathrm{Mat}_{1,m}(K) \cong K^m.$$

Zu jeder Linearform χ auf K^m existiert genau ein $w \in K^m$ mit

$$\chi(v) = w^t v \quad \text{für alle } v \in K^m.$$

(iii) Da jeder endlich-dimensionale K-Vektorraum isomorph zu einem K^m ist, ist in diesem Fall auch

$$\dim_K(V) = \dim_K(V')$$

gezeigt.
(iv) Auf $V = C([0,1], \mathbb{R})$ erhält man zum Beispiel folgende Linearform:

$$\begin{aligned} \int_0^1 \mathrm{d}t \colon C([0,1], \mathbb{R}) &\to \mathbb{R} \\ f &\mapsto \int_0^1 f(t)\mathrm{d}t. \end{aligned}$$

Für jeden Punkt $a \in [0,1]$ ist auch

$$\begin{aligned} \mathrm{e}_a \colon C([0,1], \mathbb{R}) &\to \mathbb{R} \\ f &\mapsto f(a) \end{aligned}$$

eine Linearform auf V. △

Lemma 3.2.36. *Für jede Basis $(v_1, \ldots, v_m)$ von V existiert eine Basis $(\chi_1, \ldots, \chi_m)$ von V' mit*

$$\chi_i(v_j) = \begin{cases} 1 & i = j \\ 0 & i \neq j. \end{cases}$$

Beweis. Nach Satz 3.2.9 gibt es für jedes $i = 1, \ldots, m$ genau ein $\chi_i \in V'$ mit der gewünschten Bedingung. Für die lineare Unabhängigkeit der χ_i nehmen wir

$$\lambda_1\chi_1 + \cdots + \lambda_m\chi_m = 0$$

and und setzen v_j ein. Dadurch ergibt sich direkt $\lambda_j = 0$. Wegen Bemerkung 3.2.35 (iii) handelt es sich bei $(\chi_1, \ldots, \chi_m)$ also schon um eine Basis von V'. □

Definition 3.2.37. Die Basis $(\chi_1, \ldots, \chi_m)$ von V' aus Lemma 3.2.36 nennt man die zu $(v_1, \ldots, v_m)$ **duale Basis**. △

Definition 3.2.38. Seien V, W zwei K-Vektorräume und $\varphi\colon V \to W$ eine lineare Abbildung. Dann ist die folgende Abbildung ebenfalls linear:

$$\begin{aligned}\varphi'\colon W' &\to V'\\ \chi &\mapsto \chi \circ \varphi.\end{aligned}$$

Dabei heißt φ' die zu φ **duale Abbildung**. △

Satz 3.2.39. *Seien V, W zwei K-Vektorräume mit Basen $\underline{v}$ und $\underline{w}$. Seien $\underline{v}'$ und $\underline{w}'$ die dazu dualen Basen von V' und W'. Dann gilt*

$$M_{\underline{w}',\underline{v}'}(\varphi') = M_{\underline{v},\underline{w}}(\varphi)^t.$$

Beweis. Schreiben wir

$$\varphi'(w_i') = \sum_j \gamma_j v_j'$$

so ergibt sich durch Einsetzung von v_k

$$\gamma_k = \sum_j \gamma_j v_j'(v_k) = \varphi'(w_i')(v_k) = (w_i' \circ \varphi)(v_k) = w_i'(\varphi(v_k)).$$

Schreiben wir nun

$$\varphi(v_k) = \sum_j \lambda_j w_j,$$

so ergibt sich insgesamt

$$\gamma_k = \sum_j \lambda_j w_i'(w_j) = \lambda_i.$$

Es ist aber γ_k gerade der (k, i)-Eintrag von $M_{\underline{w}',\underline{v}'}(\varphi')$ und λ_i der (i, k)-Eintrag von $M_{\underline{v},\underline{w}}(\varphi)$. Das beweist die Aussage. □

Kapitel 4

Die Determinante

In diesem Kapitel behandeln wir das wichtige Konzept der *Determinante* einer Matrix. Dabei handelt es sich um eine Zahl, die zwar offensichtlich nicht mehr die gesamte Information der gegebenen Matrix trägt, aber vor allem über die Invertierbarkeit noch Auskunft gibt. Dazu müssen wir uns zunächst ein wenig mit Permutationsgruppen beschäftigen.

4.1 Permutationen

In Beispiel 2.2.2 (iv) haben wir bereits die Permutations- oder Symmetriegruppe einer Menge M kennengelernt. Zur Erinnerung:

$$S(M) = \{f\colon M \to M \mid f \text{ bijektiv}\},$$

wobei wir die Hintereinanderausführung von Funktionen als Gruppenverknüpfung verwenden. Ab jetzt sei M stets eine endliche Menge. Wir können dabei o.B.d.A natürlich $M = \{1, \ldots, m\}$ annehmen und schreiben dann statt $S(M)$ auch

$$S_m.$$

Die Elemente $\sigma \in S_m$ sind also Bijektionen der Menge $\{1, \ldots, m\}$. Wir nennen sie auch **Permutationen** und können sie beispielsweise in **Matrixschreibweise** angeben, indem wir in die erste Zeile einer Matrix die Zahlen $1, \ldots, m$ schreiben und unter jede Zahl den ihr von σ zugeordneten Wert:

$$\sigma \mathrel{\widehat{=}} \begin{pmatrix} 1 & 2 & \cdots & m \\ \sigma(1) & \sigma(2) & \cdots & \sigma(m) \end{pmatrix}.$$

Die Hintereinanderausführung von zwei Permutationen liest man dann direkt von rechts nach links ab (hier handelt es sich *nicht* um Matrixmultiplikation!):

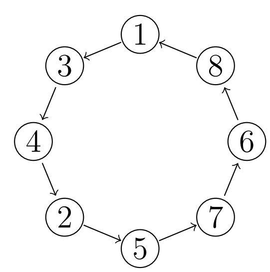
$$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}.$$

Die inverse Permutation zu σ erhält man in dieser Schreibweise einfach, indem man die beiden Zeilen vertauscht und dann die Spalten so ordnet, dass die Einträge in der ersten Zeile wieder in der richtigen Reihenfolge stehen (bzw. liest man die Permutation einfach von unten nach oben):

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

Es gibt aber auch noch eine andere Schreibweise von Permutationen, die **Zykelschreibweise**. Dabei startet man mit einer Zahl, gewöhnlich der 1, und schreibt der Reihe nach ihre iterierten Funktionswerte hintereinander, solange bis die ursprüngliche Zahl wieder auftritt. Damit ist ein Zykel abgeschlossen. Einen Zykel kann man sich also als kreisförmige Vertauschung der Zahlen vorstellen:

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 3 & 5 & 4 & 2 & 7 & 8 & 6 & 1 \end{pmatrix} =$$



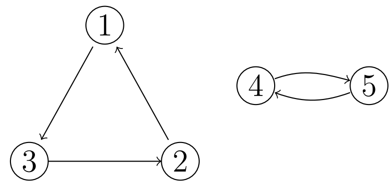
$$= (13425768)$$

Natürlich ist nicht jede Permutation ein *einzelner* Zykel. Aber offensichtlich kann man jede Permutation als Produkt von Zykeln schreiben. Dazu wählt man nach Erstellung des ersten Zykels eine gegebenenfalls noch nicht aufgetretene Zahl und startet von vorn. Alle so entstehenden Zykel schreibt man hintereinander. Die so auftretenden Zykel sind **elementfremd**, enthalten also keine gemeinsame Zahl. Man kann die Auflistung also auch als Hintereinanderausführung verstehen und die Reihenfolge spielt dabei keine Rolle. Für die Permutation

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 1 & 2 & 5 & 4 \end{pmatrix} \in S_5$$

erhält man beispielsweise die Zykelschreibweise $\sigma = (132)(45)$.

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



$$= (132)(45)$$

Sind Zykel nicht elementfremd, ist ihr Produkt natürlich im Allgemeinen von der Reihenfolge abhängig:

$$(132)(245) = (13245), \ (245)(132) = (13452).$$

Das Inverse eines Zykels entsteht einfach aus Umkehrung der der Reihenfolge der Zahlen:

$$(13245)^{-1} = (54231) = (15423).$$

Ein Zykel der Länge 2 heißt **Transposition**. Transpositionen sind selbstinvers:

$$(25)^{-1} = (52) = (25).$$

Zykel der Länge 1 lässt man gewöhnlich weg, zählt sie aber für die Gesamtanzahl der entstehenden Zykel mit.

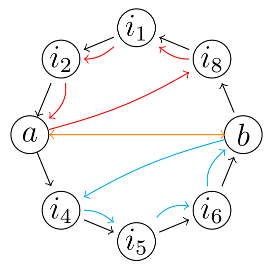
Satz 4.1.1. *(i) Jede Permutation $\sigma \in S_m$ lässt sich als Produkt von elementfremden Zykeln schreiben. Deren Anzahl ist durch σ eindeutig bestimmt.*
(ii) Ist eine Permutation $\sigma \in S_m$ ein Produkt von k elementfremden Zykeln und $\tau \in S_m$ eine Transposition, so besteht $\sigma\tau$ aus $k+1$ oder $k-1$ elementfremden Zykeln.
(iii) Jede Permutation lässt sich als Produkt von Transpositionen schreiben. Deren Anzahl ist für jede Permutation modulo 2 eindeutig bestimmt.

Beweis. Aussage (i) ist klar aufgrund der oben angegebenen Konstruktion. Für (ii) sei

$$\sigma = \rho_1 \rho_2 \cdots \rho_k$$

ein Produkt von elementfremden Zykeln und $\tau = (ab)$. Es gibt zwei Fälle: entweder kommen a und b im selben Zykel ρ_i vor, oder in zwei verschiedenen Zykeln ρ_i, ρ_j. Nachdem die Multiplikation von elementfremden Zykeln kommutativ ist, können wir im ersten Fall o.B.d.A. $i = k$ annehmen und betrachten nur $\rho_k\tau$:

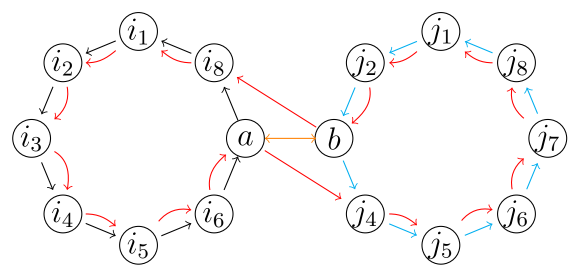
$\rho_k\tau =$



$= \rho'_k\rho''_k$

Man sieht unmittelbar, dass $\rho_k\tau$ ein zwei elementfremde Zykel zerfällt. Somit ist $\rho\tau$ ein Produkt von $k+1$ elementfremden Zykeln.
Im zweiten Fall können wir annehmen, dass a in ρ_{k-1} und b in ρ_k auftritt. Wir betrachten $\rho_{k-1}\rho_k\tau$:

$\rho_{k-1}\rho_k\tau =$



$= \rho$

Hier ist $\rho_{k-1}\rho_k\tau$ nun offensichtlich ein einziger Zykel. Somit besteht $\sigma\tau$ nun aus $k-1$ elementfremden Zykeln. Damit ist (ii) bewiesen.
Für die Existenz in (iii) reicht es nach (i) aus, nur einfache Zykel zu betrachten. Man sieht aber direkt, dass

$$(i_1 i_2 \cdots i_k) = (i_1 i_2)(i_2 i_3)\cdots(i_{k-1} i_k)$$

gilt. Die Eindeutigkeit der Länge modulo 2 folgt aus (ii): Man startet mit der Identität in S_m, die offensichtlich m Zyklen hat. Bei jeder Multiplikation mit einer Transposition ändert sich die Zyklenanzahl um genau 1. Also kann dieselbe Permutation nur höchstens dann auf zwei Weisen entstehen, wenn dabei die gleiche Anzahl modulo zwei von Transpositionen verwendet wurde. □

Beispiel 4.1.2. Wir betrachten die Permutation

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 2 & 4 & 5 & 1 & 3 & 6 & 8 & 7 \end{pmatrix} \in S_8.$$

Als Zerlegung in elementfremde Zykel erhält man

$$\sigma = (124)(35)(6)(78) = (124)(35)(78),$$

es tauchen in σ also 4 solche Zykel auf. Um eine Darstellung als Produkt von Transpositionen zu bekommen, muss man nur noch (124) zerlegen als $(12)(24)$ und bekommt

$$\sigma = (12)(24)(35)(78).$$

Es müssen also in jeder Darstellung von σ als Produkt von Transpositionen eine gerade Anzahl von Transpositionen auftreten. Natürlich kann man immer auch längere finden, zum Beispiel hier

$$\begin{aligned}\sigma &= (12)(24)(35)(78)\\ &= (12)(24)(35)(78)(18)(18)\\ &= (12)(24)(14)(35)(78)(14).\end{aligned}$$

$\triangle$

Definition 4.1.3. Für $\sigma \in S_m$ definieren wir die **Signatur** als

$$\operatorname{sgn}(\sigma) := (-1)^s$$

wenn σ ein Produkt von s Transpositionen ist. Nach Satz 4.1.1 (iii) ist das wohldefiniert. Auf diese Weise erhalten wir die **Signaturabbildung**

$$\operatorname{sgn}\colon S_m \to \{-1, 1\}.$$

$\triangle$

Bemerkung/Beispiel 4.1.4. (i) Die Permutation σ aus Beispiel 4.1.2 hat Signatur $(-1)^4 = 1$. Die Identität hat stets Signatur 1.
(ii) Ein Zykel hat genau dann Signatur 1, wenn er ungerade Länge hat. Das folgt aus der Beobachtung des Beweises von Satz 4.1.1 (iii), dass ein Zykel der Länge k ein Produkt von $k-1$ Transpositionen ist.
(iii) Schreibt man σ als ein Produkt von Zykeln und treten dabei r Zykel gerader Länge auf, so gilt $\operatorname{sgn}(\sigma) = (-1)^r$. Für die Berechnung der Signatur kann man Zykel von ungerader Länge also ignorieren.
(iv) Es gilt stets

$$\operatorname{sgn}(\sigma) = \operatorname{sgn}(\sigma^{-1}).$$

Aus einer Darstellung von σ als Produkt von s Transpositionen erhält man sofort eine Darstellung von σ^{-1} als Produkt von denselben s Transpositionen, nur in umgedrehter Reihenfolge. $\triangle$

Korollar 4.1.5. *Für $\sigma, \tau \in S_m$ gilt*

$$\operatorname{sgn}(\sigma\tau) = \operatorname{sgn}(\sigma)\operatorname{sgn}(\tau).$$

Beweis. Ist σ ein Produkt von r Transpositionen und τ ein Produkt von s Transpositionen, so ist $\sigma\tau$ offensichtlich ein Produkt von $r+s$ Transpositionen. Es gilt also
$$\operatorname{sgn}(\sigma\tau) = (-1)^{r+s} = (-1)^r \cdot (-1)^s = \operatorname{sgn}(\sigma)\operatorname{sgn}(\tau). \qquad \square$$

Beispiel 4.1.6. Wir betrachten das bekannte Schiebe-Puzzle, bei dem 15 kleine Quadrate in einem Rahmen mit 16 Positionen vorgegeben sind:

1	2	3	4
5	6	7	8
9	10	11	12
13	14	15	

Da eine Position leer ist, kann man manche Quadrate horizontal oder vertikal verschieben. Gewöhnlich ist das Spiel in einer ungeordneten Version vorgegeben und man möchte es mit solchen Verschiebungen wieder in die oben angegebene Ursprungsgestalt überführen.
Wir nummerieren die Positionen im Feld mit den Zahlen 1 bis 16. Auch die Quadrate werden mit 1 bis 16 bezeichnet, gemäß ihrer Beschriftung. Dabei stehe 16 für das leere Quadrat. Dann kann eine Position des Spiels durch eine Permutation $\sigma \in S_{16}$ angegeben werden, wobei $\sigma(i)$ die Zahl desjenigen Quadrats bezeichne, das an Position i steht. So beschreibt also

$$\sigma_1 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 & 16 \\ 3 & 11 & 16 & 14 & 2 & 1 & 15 & 9 & 7 & 4 & 8 & 12 & 6 & 5 & 10 & 13 \end{pmatrix}$$

etwa die folgende Konfiguration:

3	11		14
2	1	15	9
7	4	8	12
6	5	10	13

Die oben angegebene Ursprungskonfiguration entspricht gerade $\mathrm{id} \in S_{16}$. In dieser Formulierung entspricht nun jeder Zug der Multiplikation mit einer Transposition von links. Dabei treten sogar nur Transposition der Gestalt

$$(16\, i)$$

auf. Da man immer nur vertikal oder horizontal verschieben kann, weiß man je nach Ausgangs- und Endposition des leeren Quadrats auch die Anzahl der nötigen Transpositionen modulo 2. Möchte man beispielsweise die durch σ_1 angegebene Konfiguration in die Ursprungskonfiguration überführen, so geht das höchstens mit einer gerade Anzahl von Transpositionen. Es gilt nun $\operatorname{sgn}(\mathrm{id}) = 1$ und $\operatorname{sgn}(\sigma_1) = -1$, wie man leicht berechnet. Also kann id nicht aus σ_1 durch Multiplikation einer geraden Anzahl von Transpositionen erhalten werden. Somit lässt sich die Konfiguration σ_1 nicht in den Ursprungszustand überführen. △

4.2 Die Determinante und ihre Eigenschaften

Sei auch in diesem Abschnitt K stets ein Körper.

Definition 4.2.1 (Leibniz-Formel der Determinante)**.** Für $A = (a_{ij})_{i,j} \in \mathrm{Mat}_m(K)$ definieren wir die **Determinante** als

$$\det(A) := \sum_{\sigma \in S_m} \operatorname{sgn}(\sigma) \cdot a_{\sigma(1)1} \cdot a_{\sigma(2)2} \cdots a_{\sigma(m)m}.$$

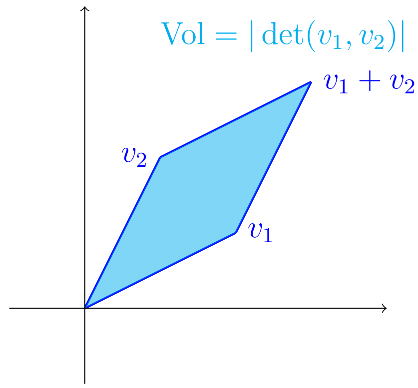
Wir erhalten dadurch die Determinantenabbildung

$$\det\colon \mathrm{Mat}_m(K) \to K.$$ △

Bemerkung/Beispiel 4.2.2. (i) In einem Summanden der Leibnizformel wählen wir aus jeder Spalte von A ein Element. Die Permutation σ gibt uns dabei vor, welches Element der jeweiligen Spalte wir wählen. Insgesamt haben wir dann auch aus jeder Zeile genau ein Element gewählt, denn σ ist bijektiv. Alle diese Element multiplizieren wir miteinander und mit der Signatur der Permutation. Alle Terme werden dann addiert.
(ii) Im Fall $m = 2$ gilt $S_2 = \{\mathrm{id}, (12)\}$ mit $\operatorname{sgn}(\mathrm{id}) = 1, \operatorname{sgn}((12)) = -1$. Wir erhalten also für $A \in \mathrm{Mat}_2(K)$

$$\det(A) = \det\left(\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}\right) = a_{11}a_{22} - a_{21}a_{12}.$$

Im Fall $K = \mathbb{R}$ gibt es hier eine geometrische Interpretation der Determinate. Wir fassen dazu die Spalten $v_1 = (a_{11}, a_{21})^t, v_2 = (a_{12}, a_{22})^t$ von A als Elemente von $\mathbb{R}^2$ auf. Die beiden Vektoren v_1, v_2 spannen ein Parallelogramm auf, dessen Volumen genau mit $|\det(A)|$ übereinstimmt (Übungsaufgabe):



(iii) Für allgemeines m besteht die Summe in der Leibnizformel aus $|S_m| = m!$ Summanden. Bereits für Matrizen der Größe $m = 7$ sind das mehr als 5000 Summanden. Für $m = 60$ stimmt $m!$ in etwa mit der geschätzten Zahl aller Atome im Universum überein. Wir werden also eine effizientere Methode zur Berechnung der Determinante finden müssen.
(iv) Eine Matrix $A = (a_{ij})_{i,j} \in \mathrm{Mat}_m(K)$ hat **obere Dreiecksgestalt**, wenn $a_{ij} = 0$ für alle $j < i$ gilt. Damit ein Summand

$$\mathrm{sgn}(\sigma) \cdot a_{\sigma(1)1} \cdot a_{\sigma(2)1} \cdots a_{\sigma(m)m}$$

nicht Null wird, muss also $\sigma(i) \leq i$ für alle $i = 1, \ldots, m$ gelten. Die einzige Permutation die das erfüllt ist $\sigma = \mathrm{id}$. Wegen $\mathrm{sgn}(\mathrm{id}) = 1$ gilt also

$$\det(A) = a_{11}a_{22} \cdots a_{mm}.$$

Die Determinate ist hier also genau das Produkt der Diagonaleinträge von A. Dasselbe Argument funktioniert auch mit Matrizen in *unterer Dreiecksgestalt*. $\triangle$

Proposition 4.2.3. *Für $A \in \mathrm{Mat}_m(K)$ und $v_1, \ldots, v_m \in K^m$ gilt:*

(i) $\det(I_m) = 1$.

(ii) Für jedes $i = 1, \ldots, m$ ist die Abbildung

$$\begin{aligned} K^m &\to K \\ w &\mapsto \det\left((v_1, \ldots, v_{i-1}, w, v_{i+1}, \ldots, v_m)\right) \end{aligned}$$

K-linear.

(iii) Besitzt A zwei identische Spalten, so gilt $\det(A) = 0$.

(*iv*) *Für $i \neq j$ gilt*

$$\det\left((v_1, \ldots, v_i, \ldots, v_j, \ldots, v_m)\right) = -\det\left((v_1, \ldots, v_j, \ldots, v_i, \ldots, v_m)\right).$$

(*v*) $\det(A) = \det(A^t)$. ***Insbesondere gelten Aussagen (i)-(iii) auch für Zeilen statt Spalten.***

(*vi*) *Die Determinante ist die einzige Funktion $\mathrm{Mat}_m(K) \to K$, welche (i)-(iii) erfüllt.*

Beweis. Aus Bemerkung 4.2.2 (*iv*) folgt (*i*) direkt. Für (*ii*) seien $w_1, w_2 \in K^m$ und $\lambda \in K$. Wir fixieren i und betrachten für die Matrix

$$(v_1, \ldots, v_{i-1}, w_1 + w_2, v_{i+1}, \ldots, v_m)$$

den Summanden zur Permutation σ in der Leibnizformel. Nur der Faktor mit Index $(\sigma(i), i)$ hängt überhaupt von w_1 und w_2 ab. Dieser Faktor ist aber gerade die Summe der $\sigma(i)$-ten Einträge von w_1 und w_2. Das stimmt für jedes $\sigma \in S_m$. Aufgrund des Distributivgesetzes in K gilt also

$$\begin{aligned}
&\det\left((v_1, \ldots, v_{i-1}, w_1 + w_2, v_{i+1}, \ldots, v_m)\right) \\
&= \det\left((v_1, \ldots, v_{i-1}, w_1, v_{i+1}, \ldots, v_m)\right) + \det\left((v_1, \ldots, v_{i-1}, w_2, v_{i+1}, \ldots, v_m)\right).
\end{aligned}$$

In der Leibnizformel für die Matrix $(v_1, \ldots, v_{i-1}, \lambda w_1, v_{i+1}, \ldots, v_m)$ steht in jedem Summanden der Skalar λ im i-ten Faktor und kann herausgehoben werden. Damit gilt

$$\det\left(v_1, \ldots, v_{i-1}, \lambda w_1, v_{i+1}, \ldots, v_m\right) = \lambda \cdot \det\left(v_1, \ldots, v_{i-1}, w_1, v_{i+1}, \ldots, v_m\right)$$

und (*i*) ist bewiesen.
Seien für (*iii*) in A die i-te und die j-te Spalte identisch, für $i \neq j$. Damit stimmen die Summanden der Leibnizformel zu den Permutationen σ und $\sigma \circ (ij)$ übereinein, haben nur jeweils anderes Vorzeichen (wir verwenden hier auch das Kommutativgesetz der Multiplikation in K). Sie heben sie also gerade gegenseitig auf.
Aussage (*iv*) sieht man beispielsweise mit (*ii*) und (*iii*) so:

$$\begin{aligned}
0 &= \det\left((v_1, \ldots, v_i + v_j, \ldots, v_i + v_j, \ldots, v_m)\right) \\
&= \det\left((v_1, \ldots, v_i, \ldots, v_i, \ldots, v_m)\right) \\
&\quad + \det\left((v_1, \ldots, v_i, \ldots, v_j, \ldots, v_m)\right) \\
&\quad + \det\left((v_1, \ldots, v_j, \ldots, v_i, \ldots, v_m)\right) \\
&\quad + \det\left((v_1, \ldots, v_j, \ldots, v_j, \ldots, v_m)\right) \\
&= \det\left((v_1, \ldots, v_i, \ldots, v_j, \ldots, v_m)\right) \\
&\quad + \det\left((v_1, \ldots, v_j, \ldots, v_i, \ldots, v_m)\right).
\end{aligned}$$

Das beweist die Aussage.
Für (v) sei wieder $A = (a_{ij})_{i,j}$. Die Leibnizformel für A^t lautet dann

$$\det(A^t) = \sum_{\sigma \in S_m} \operatorname{sgn}(\sigma) \cdot a_{1\sigma(1)} \cdots a_{m\sigma(m)}.$$

Wenn σ die Menge aller Permutationen durchläuft, durchläuft auch σ^{-1} die Menge aller Permutationen. Somit gilt

$$\begin{aligned}
\det(A^t) &= \sum_{\sigma \in S_m} \operatorname{sgn}(\sigma^{-1}) \cdot a_{1\sigma^{-1}(1)} \cdots a_{m\sigma^{-1}(m)} \\
&= \sum_{\sigma \in S_m} \operatorname{sgn}(\sigma) \cdot a_{\sigma(1)\sigma^{-1}(\sigma(1))} \cdots a_{\sigma(m)\sigma^{-1}(\sigma(m))} \\
&= \sum_{\sigma \in S_m} \operatorname{sgn}(\sigma) \cdot a_{\sigma(1)1} \cdots a_{\sigma(m)m} \\
&= \det(A).
\end{aligned}$$

Für die zweite Gleichheit haben wir die Faktoren im Produkt umgeordnet sowie $\operatorname{sgn}(\sigma^{-1}) = \operatorname{sgn}(\sigma)$ benutzt.
Sei für (vi) schließlich

$$\widetilde{\det}\colon \operatorname{Mat}_m(K) \to K$$

eine weitere Abbildung mit den Eigenschaften (i)-(iii). Für $A = (a_{ij})_{i,j} \in \operatorname{Mat}_m(K)$ gilt

$$A = \left(\sum_{i=1}^m a_{i1} e_i, \ldots, \sum_{i=1}^m a_{im} e_i \right)$$

und mit (i) also

$$\widetilde{\det}(A) = \sum_{i_1,\ldots,i_m=1}^m a_{i_1 1} \cdots a_{i_m m} \cdot \widetilde{\det}\left(e_{i_1}, \ldots e_{i_m}\right).$$

Laut Eigenschaft (iii) verschwinden alle Summanden, in denen zwei Indizes gleich sind. Damit kann man die Summe umschreiben zu

$$\widetilde{\det}(A) = \sum_{\sigma \in S_m} a_{\sigma(1)1} \cdots a_{\sigma(m)m} \cdot \widetilde{\det}\left(e_{\sigma(1)}, \ldots e_{\sigma(m)}\right).$$

Ist σ ein Produkt von s Transpositionen, kann die Matrix $(e_{\sigma(1)}, \dots, e_{\sigma(m)})$ durch s Spaltenvertauschungen in I_m überführt werden. Wegen Eigenschaften (iv) und (i) gilt also

$$\widetilde{\det}\left(e_{\sigma(1)}, \dots e_{\sigma(m)}\right) = (-1)^s = \operatorname{sgn}(\sigma).$$

Insgesamt ergibt sich

$$\widetilde{\det}(A) = \sum_{\sigma \in S_m} \operatorname{sgn}(\sigma) \cdot a_{\sigma(1)1} \cdots a_{\sigma(m)m} = \det(A). \qquad \square$$

Satz 4.2.4. *Für* $A, B \in \operatorname{Mat}_m(K)$ *gilt*

$$\det(AB) = \det(A) \cdot \det(B).$$

Beweis. Wir bezeichnen die Spalten von A wieder mit $a_{\cdot 1}, \dots a_{\cdot m}$. Es gilt dann

$$AB = \left(\sum_{i=1}^m b_{i1} a_{\cdot i}, \dots, \sum_{i=1}^m b_{im} a_{\cdot i} \right),$$

siehe dazu auch Beispiel 3.1.11. Wenn wir die Linearität der Determinante in jeder Spalte verwenden, erhalten wir

$$\det(AB) = \sum_{i_1, \dots, i_m = 1}^m b_{i_1 1} \cdots b_{i_m m} \cdot \det\left((a_{\cdot i_1}, \dots, a_{\cdot i_m})\right).$$

Wieder verschwinden alle Summanden, in denen zwei Indizes gleich sind. Damit kann man die Summe umschreiben zu

$$\det(AB) = \sum_{\sigma \in S_m} b_{\sigma(1)1} \cdots b_{\sigma(m)m} \cdot \det\left((a_{\cdot \sigma(1)}, \dots, a_{\cdot \sigma(m)})\right).$$

Genau wie im letzten Beweis folgt nun

$$\det\left((a_{\cdot \sigma(1)}, \dots, a_{\cdot \sigma(m)})\right) = (-1)^s \cdot \det(A) = \operatorname{sgn}(\sigma) \cdot \det(A).$$

Insgesamt erhalten wir

$$\begin{aligned} \det(AB) &= \sum_{\sigma \in S_m} b_{\sigma(1)1} \cdots b_{\sigma(m)m} \cdot \operatorname{sgn}(\sigma) \cdot \det(A) \\ &= \det(A) \cdot \sum_{\sigma \in S_m} \operatorname{sgn}(\sigma) \cdot b_{\sigma(1)1} \cdots b_{\sigma(m)m} \\ &= \det(A) \cdot \det(B). \end{aligned} \qquad \square$$

Korollar 4.2.5. *Für* $A, B \in \mathrm{Mat}_m(K)$ *und* $P \in \mathrm{GL}_m(K)$ *gilt*

$$\begin{aligned} \det(AB) &= \det(BA) \\ \det\left(P^{-1}\right) &= \det(P)^{-1} \\ \det\left(P^{-1}AP\right) &= \det(A). \end{aligned}$$

Beweis. Die erste Gleichung ist klar wegen Satz 4.2.4, die dritte Gleichung folgt direkt daraus. Die zweite Gleichung folgt aus

$$1 = \det(I_m) = \det\left(P^{-1}P\right) = \det\left(P^{-1}\right)\det(P). \qquad \square$$

Im folgenden Korollar verwenden wir wieder die Elementarmatrizen aus Definition 2.3.9.

Korollar 4.2.6. *Für* $A \in \mathrm{Mat}_m(K)$*,* $i \neq j$ *und* $0 \neq \lambda \in K$ *gilt:*

(i) $\det(P_{ij}A) = -\det(A)$.

(ii) $\det(S_i(\lambda)A) = \lambda\det(A)$.

(iii) $\det(M_j^i(\lambda)A) = \det(A)$.

Beweis. Wir verwenden Satz 4.2.4 und müssen somit nur die Determinanten der Elementarmatrizen ausrechnen. Dazu verwenden wir Proposition 4.2.3. Es gilt $\det(P_{ij}) = -1$, denn P_{ij} entsteht aus I_m durch die Vertauschung der i-ten und j-ten Spalte. Es gilt $\det(S_i(\lambda)) = \lambda$, denn $S_i(\lambda)$ entsteht aus I_m durch Multiplikation der i-ten Spalte mit λ. Es gilt $\det(M_j^i(\lambda)) = 1$, denn $M_j^i(\lambda)$ hat obere/untere Dreiecksgestalt mit Diagonaleinträgen gleich 1. □

Bemerkung 4.2.7. Korollar 4.2.6 kann man für einen effizienten Algorithmus zur Berechnung der Determinante verwenden. Man bringt A durch elementare Zeilenumformungen (am besten nur vom Typ (I) und (III)) auf obere Dreiecksgestalt und rechnet dann die Determinante als Produkt der Diagonaleinträge leicht aus. Hat man dabei eine ungerade Anzahl von Zeilenvertauschungen durchgeführt, muss man das Ergebnis noch mit -1 multiplizieren und erhält so die Determinante von A.
Man kann natürlich auch Umformungen vom Typ (II) verwenden. Für jede Multiplikation einer Zeile mit λ muss man das Ergebnis hinterher mit λ^{-1} multiplizieren, um die Determinante von A zu erhalten.
Wegen Proposition 4.2.3 (v) kann man genauso auch Spaltentransformationen verwenden. △

Beispiel 4.2.8. Wir wollen die Determinante von

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{pmatrix} \in \mathrm{Mat}_4(\mathbb{Q})$$

berechnen. Dazu bringen wir A auf Zeilenstufenform:

$$\begin{aligned} A &\rightsquigarrow \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & -4 & -8 & -12 \\ 0 & -8 & -16 & -24 \\ 0 & -12 & -24 & -36 \end{pmatrix} \\ &\rightsquigarrow \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & -4 & -8 & -12 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \end{aligned}$$

Dabei haben wir nur Umformungen vom Typ (III) verwendet. Also gilt

$$\det(A) = 1 \cdot (-4) \cdot 0 \cdot 0 = 0.$$

Die Leibnizformel hätte $4! = 24$ Summanden benötigt. △

Der folgende Satz ist einer der wichtigesten Gründe für die Einführung der Determinante.

Satz 4.2.9. *Für $A \in \mathrm{Mat}_m(K)$ gilt*

$$A \in \mathrm{GL}_m(K) \;\Leftrightarrow\; \det(A) \neq 0.$$

Beweis. "$\Rightarrow$" folgt aus

$$1 = \det(I_m) = \det\left(A^{-1}A\right) = \det\left(A^{-1}\right)\det\left(A\right).$$

Für "$\Leftarrow$" bringt man A durch elementare Zeilenumformungen auf reduzierte Zeilenstufenform. Dabei kann sich die Determinante ändern, aber laut Korollar 4.2.6 wird sie sicher nicht 0. Eine quadratische Matrix in reduzierter Zeilenstufenform mit Determinante $\neq 0$ muss aber die Einheitsmatrix sein. Also ist A nach Satz 2.3.12 invertierbar. □

Beispiel 4.2.10. Wir betrachten

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \in \mathrm{Mat}_2(\mathbb{Z}/7\mathbb{Z})$$

und berechnen

$$\det(A) = 4 - 6 = -2 = 5 \neq 0$$

A ist also über $\mathbb{Z}/7\mathbb{Z}$ invertierbar. Das lineare Gleichungssystem

$$\begin{aligned} x + 2y &= b_1 \\ 3x + 4y &= b_2 \end{aligned}$$

hat also für jede Wahl von $b_1, b_2 \in \mathbb{Z}/7\mathbb{Z}$ genau eine Lösung.
Über dem Körper $\mathbb{Z}/2\mathbb{Z}$ ist die Determinante Null, also ist A dort nicht invertierbar. Somit gibt es hier mehr als eine Lösung des homogenen Systems und mindestens ein inhomogenes System ist unlösbar. △

Korollar 4.2.11. *Für $A \in \mathrm{Mat}_{m,n}(K)$ ist $\mathrm{rang}(A)$ gerade die größte Größe einer quadratischen Untermatrix B von A mit $\det(B) \neq 0$.*

Beweis. Klar mit Satz 4.2.9 und Proposition 3.2.26. □

4.3 Entwicklungsformeln und Cramersche Regel ★

Wir lernen nun mit dem Laplaceschen Entwicklungssatz eine Methode kennen, um die Berechnung der Determinante einer $m \times m$-Matrix auf die Berechnung von Determinanten von Matrizen der Größe $m-1$ zu reduzieren. Theoretisch kann man das iterativ immer weiter anwenden und kommt dann zu Determinanten von Matrizen der Größe 1, die sehr leicht zu berechnen sind. Zum Ausrechnen ist dieses Vorgehen gewöhnlich nicht effizient genug, es ist aber manchmal für theoretische Überlegungen wichtig.
Für eine Matrix $A = (a_{ij})_{i,j} \in \mathrm{Mat}_m(K)$ sowie $i, j \in \{1, \ldots, m\}$ bezeichnen wir mit

$$A_{ij} \in \mathrm{Mat}_{m-1}(K)$$

die Untermatrix von A, durch Streichung der i-ten Zeile und j-ten Spalte entsteht.

Satz 4.3.1 (Laplacescher Entwicklungssatz). *Für $A \in \mathrm{Mat}_m(K)$ und jedes $i \in \{1, \dots, m\}$ gilt*

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

sowie

$$\det(A) = \sum_{j=1}^{m} (-1)^{i+j} a_{ji} \det(A_{ji}).$$

Beweis. Für festes i betrachten wir die Abbildung

$$\begin{aligned} \widetilde{\det} \colon \mathrm{Mat}_m(K) &\to K \\ A &\mapsto \sum_{j=1}^{m} (-1)^{i+j} a_{ij} \det(A_{ij}). \end{aligned}$$

Man sieht nun relativ leicht, dass $\widetilde{\det}$ die Eigenschaften (i)-(iii) der Determinante aus Proposition 4.2.3 hat. Mit Proposition 4.2.3 (vi) folgt damit $\widetilde{\det} = \det$. Die Formeln zur Spaltenentwicklung folgen zum Beispiel direkt aus Proposition 4.2.3 (v). □

Bemerkung/Beispiel 4.3.2. (i) Die Entwicklung der Determinante nach einer Zeile oder Spalte führt zu m Summanden, in denen jeweils die Determinante einer Matrix der Größe $m-1$ zu berechnen ist. Iteriert man das, erhält man eine Summe der Länge $m!$, genau wie in der Leibnizformel. Die Entwicklung ist also meistens nur dann sinnvoll, wenn die Matrix viele Nulleinträge hat, weil dann viele der Summanden wegfallen.
(ii) Wir wollen beispielsweise die Determinante der Matrix

$$A = \begin{pmatrix} 1 & 0 & 0 & 2 \\ 0 & -1 & 1 & 2 \\ 1 & 1 & 0 & 0 \\ 0 & -2 & 1 & 0 \end{pmatrix} \in \mathrm{Mat}_4(\mathbb{Q})$$

berechnen. Durch Entwicklung nach der ersten Zeile erhalten wir

$$\det(A) = 1 \cdot \det \begin{pmatrix} -1 & 1 & 2 \\ 1 & 0 & 0 \\ -2 & 1 & 0 \end{pmatrix} - 2 \cdot \det \begin{pmatrix} 0 & -1 & 1 \\ 1 & 1 & 0 \\ 0 & -2 & 1 \end{pmatrix}.$$

Die Determinante im ersten Summanden entwicklen wir nach der zweiten Zeile zu

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

und die Determinanten im zweiten Summanden nach der ersten Spalte zu

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

Insgesamt ergibt sich

$$\det(A) = 1 \cdot 2 - 2 \cdot (-1) = 4.$$

(iii) Für allgemeines

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \in \mathrm{Mat}_3(K)$$

erhalten wir durch Entwicklung nach der ersten Zeile

$$\begin{aligned}
\det(A) &= a_{11}\det\begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix} - a_{12}\det\begin{pmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{pmatrix} + a_{13}\begin{pmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix} \\
&= a_{11}\left(a_{22}a_{33} - a_{32}a_{23}\right) - a_{12}\left(a_{21}a_{33} - a_{31}a_{23}\right) + a_{13}\left(a_{21}a_{32} - a_{31}a_{22}\right) \\
&= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{31}a_{22}a_{13} - a_{32}a_{23}a_{11} - a_{33}a_{21}a_{12}.
\end{aligned}$$

Man kann sich das mit folgendem Schema veranschaulichen:

$$\begin{array}{ccc|cc}
+ & + & + & & \\
a_{11} & a_{12} & a_{13} & a_{11} & a_{12} \\
a_{21} & a_{22} & a_{23} & a_{21} & a_{22} \\
a_{31} & a_{32} & a_{33} & a_{31} & a_{32} \\
- & - & - & &
\end{array}$$

Dieses Schema zur Berechnung der Determinante nennt man die **Regel von Sarrus**. Sie ist allerdings *nur* für Matrizen der Größe 3×3 gültig! $\triangle$

Definition 4.3.3. Für $A \in \mathrm{Mat}_m(K)$ nennen wir

$$A^{\mathrm{adj}} := \big((-1)^{i+j}\det(A_{ji})\big)_{i,j} \in \mathrm{Mat}_m(K)$$

die zu A **adjunkte Matrix**. △

Satz 4.3.4. *Für jedes $A \in \mathrm{Mat}_m(K)$ gilt*

$$A \cdot A^{\mathrm{adj}} = A^{\mathrm{adj}} \cdot A = \det(A) \cdot I_m.$$

Aus $\det(A) \neq 0$ folgt daraus insbesondere

$$A^{-1} = \frac{1}{\det(A)} A^{\mathrm{adj}}.$$

Beweis. Der (r,j)-Eintrag von $A \cdot A^{\mathrm{adj}}$ lautet gerade

$$\sum_{i=1}^{m} (-1)^{i+j} a_{ri} \det(A_{ji}).$$

Dies ist genau die Laplace Entwicklung nach der j-ten Zeile der Matrix B, die aus A durch Ersetzung der j-ten Zeile durch die r-te Zeile entsteht. Für $r = j$ gilt damit $B = A$ und also $\det(B) = \det(A)$, für $i \neq j$ hat B zwei identische Zeilen, also erhalten wir als Ergebnis 0. Das beweist die Aussage. □

Beispiel 4.3.5. Für

$$A = \begin{pmatrix} i & 1 \\ 1+i & -i \end{pmatrix} \in \mathrm{Mat}_2(\mathbb{C})$$

erhalten wir $\det(A) = -i \neq 0$, also ist A invertierbar. Wir berechnen

$$A^{\mathrm{adj}} = \begin{pmatrix} -i & -1 \\ -1-i & i \end{pmatrix}$$

und erhalten

$$A^{-1} = iA^{\mathrm{adj}} = \begin{pmatrix} 1 & -i \\ 1-i & -1 \end{pmatrix}.$$ △

Die nun folgende **Cramersche Regel** gibt eine explizite Formel für die Lösung eines inhomogenen linearen Gleichungssystems an, bei invertierbarer Koeffizientenmatrix:

Satz 4.3.6 (Cramersche Regel)**.** *Sei* $A = (a_{\cdot 1}, \ldots, a_{\cdot m}) \in \mathrm{GL}_m(K)$ *und* $b \in K^m$. *Dann gilt für das (eindeutige)* $c = (c_1, \ldots, c_m)^t \in K^m$ *mit* $Ac = b$ *gerade*

$$c_i = \frac{\det(a_{\cdot 1}, \ldots, a_{\cdot i-1}, b, a_{\cdot i+1}, \ldots, a_{\cdot m})}{\det(A)}$$

für alle $i = 1, \ldots, m$.

Beweis. Sei $c = (c_1, \ldots, c_m)^t \in K^m$ der Spaltenvektor mit $Ac = b$, es gilt also

$$\sum_{j=1}^{m} c_j a_{\cdot j} = b.$$

Damit erhalten wir für jedes feste $i = 1, \ldots, m$

$$\begin{aligned}
&\det(a_{\cdot 1}, \ldots, a_{\cdot i-1}, b, a_{\cdot i+1}, \ldots, a_{\cdot m}) \\
=&\det\left(a_{\cdot 1}, \ldots, a_{\cdot i-1}, \sum_{j=1}^{m} c_j a_{\cdot j}, a_{\cdot i+1}, \ldots, a_{\cdot m}\right) \\
=&\sum_{j=1}^{m} c_j \det(a_{\cdot 1}, \ldots, a_{\cdot i-1}, a_{\cdot j}, a_{\cdot i+1}, \ldots, a_{\cdot m}) \\
=&c_i \det(a_{\cdot 1}, \ldots, a_{\cdot i-1}, a_{\cdot i}, a_{\cdot i+1}, \ldots, a_{\cdot m}) \\
=&c_i \det(A).
\end{aligned}$$

Hier haben wir wieder Eigenschaften (ii) und (iii) aus Proposition 4.2.3 benutzt. Das beweist die Aussage. $\square$

Beispiel 4.3.7. Wir wollen das lineare Gleichungssystem

$$\begin{aligned} x + 2y &= 3 \\ 3x + 4y &= 1 \end{aligned}$$

über $\mathbb{Z}/7\mathbb{Z}$ lösen. Wir haben in Beispiel 4.2.10 schon

$$\det\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = 5$$

berechnet. Es gilt nun

$$\det\begin{pmatrix} 3 & 2 \\ 1 & 4 \end{pmatrix} = 3 \text{ und } \det\begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix} = 6.$$

Wegen $\frac{1}{5} = 3$ erhalten wir also genau die Lösung $c = (2, 4)^t \in (\mathbb{Z}/7\mathbb{Z})^2$. $\triangle$

Kapitel 5

Eigenwerte und Eigenvektoren

Um lineare Abbildungen und deren Darstellungsmatrizen besser zu verstehen, führen wir nun den wichtigen Begriff eines Eigenwerts und -vektors einer linearen Abbildung ein. Diese Begriffe bilden die Grundlage für fast den gesamten Rest der Vorlesung. Sei auch hier wieder K ein beliebiger Körper. Wir müssen zuerst ein paar Fakten über Polynome in einer Variablen über K zusammenstellen.

5.1 Eigenwerte und Eigenvektoren

In diesem Abschnitt sei V stets ein endlich-dimensionaler K-Vektorraum und

$$\varphi\colon V \to V$$

eine lineare Abbildung. Eine solche lineare Abbildung von einem Raum in sich selbst nennt man auch **Endomorphismus**.

Definition 5.1.1. (i) Ein $\lambda \in K$ heißt **Eigenwert von** φ, falls es ein $0 \neq v \in V$ gibt mit

$$\varphi(v) = \lambda v.$$

Jedes solche v heißt dann **Eigenvektor** von φ zum Eigenwert λ.
(ii) Zu einem Eigenwert λ von φ heißt

$$\mathrm{Eig}(\varphi, \lambda) := \{v \in V \mid \varphi(v) = \lambda v\}$$

der zu λ gehörige **Eigenraum** von φ.
(iii) Die Menge aller Eigenwerte von φ heißt das **Spektrum** von φ. △

Bemerkung/Beispiel 5.1.2. (i) Ein Vektor $0 \neq v$ ist genau dann ein Eigenvektor von φ, wenn er durch φ nur skaliert, nicht aber in der Richtung verändert wird. Der Skalierungsfaktor ist dann genau der dazugehörende Eigenwert.
(ii) Die Bedingung $0 \neq v$ in Definition 5.1.1 (i) ist wichtig. Ohne sie wäre jedes $\lambda \in K$ ein Eigenwert von φ. Zum Eigenraum $\mathrm{Eig}(\varphi, \lambda)$ gehört der Nullvektor laut Definition aber immer.
(iii) Für jeden Eigenwert λ von φ ist $\mathrm{Eig}(\varphi, \lambda)$ ein Untervektorraum von V. Für $v, w \in \mathrm{Eig}(\varphi, \lambda)$ und $\gamma \in K$ gilt

$$\begin{aligned}\varphi(v+w) &= \varphi(v) + \varphi(w) = \lambda v + \lambda w = \lambda(v+w)\\ \varphi(\gamma v) &= \gamma\varphi(v) = \gamma\lambda v = \lambda(\gamma v).\end{aligned}$$

(iv) Eine Spiegelung des $\mathbb{R}^2$ an einer Ursprungsgeraden besitzt die Eigenwerte 1 und -1. Der Eigenraum zum Eigenwert 1 ist gerade die Spiegelungsgerade, die dazu orthogonale Ursprungsgerade ist der Eigenraum zum Eigenwert -1.
(v) Nicht jede lineare Abbildung besitzt Eigenwerte oder -vektoren. Eine Drehung des $\mathbb{R}^2$ um $\pi/2$ gegen den Uhrzeigersinn besitzt keinen Eigenwert, wie man geometrisch sofort erkennt. △

Bemerkung 5.1.3 (★). Um Eigenwerte und -vektoren wirklich auszurechnen, müssen wir meist zu einer Darstellungsmatrix übergehen, d.h. man wählt eine Basis $\underline{v}$ von V und berechnet zunächst

$$A := M_{\underline{v},\underline{v}}(\varphi).$$

Die Eigenwerte von φ stimmen dann genau mit den Eigenwerten von μ_A überein und die Eigenräume entsprechen sich via $\varphi_{\underline{v}}$ und $\xi_{\underline{v}}$. Dabei ist es aber wichtig, dass in V auf beiden Seiten dieselbe Basis $\underline{v}$ gewählt wurde.

$$\begin{tikzcd} V \arrow[r, "\varphi"] \arrow[d, "\xi_{\underline{v}}"] & V \arrow[d, "\xi_{\underline{v}}"] \\ K^m \arrow[r, "\mu_A"] & K^m \end{tikzcd}$$

Für einen Eigenwert λ von φ mit Eigenvektor $v \in V$ gilt nämlich

$$\mu_A(\xi_{\underline{v}}(v)) = \xi_{\underline{v}}(\varphi(v)) = \xi_{\underline{v}}(\lambda v) = \lambda\xi_{\underline{v}}(v).$$

Für einen Eigenwert γ von μ_A mit Eigenvektor $c \in K^m$ gilt

$$\varphi(\varphi_{\underline{v}}(c)) = \varphi_{\underline{v}}(\mu_A(c)) = \varphi_{\underline{v}}(\gamma c) = \gamma\varphi_{\underline{v}}(c).$$ △

Proposition 5.1.4. *Für $A \in \mathrm{Mat}_m(K)$ und $\lambda \in K$ sind äquivalent:*

(i) *λ ist Eigenwert von μ_A.*

(ii) $\mathrm{Kern}\,(\lambda \mathrm{id}_{K^m} - \mu_A) \neq \{0\}$.

(iii) *$\lambda I_m - A$ ist nicht invertierbar.*

(iv) $\det\,(\lambda I_m - A) = 0$.

Dabei gilt

$$\mathrm{Eig}(\mu_A, \lambda) = \mathrm{Kern}\,(\lambda \mathrm{id}_{K^m} - \mu_A) = \mathrm{Ker}(\mu_{(\lambda I_m - A)}) = L(\lambda I_m - A, 0).$$

Beweis. Wegen

$$\mu_A(v) = \lambda v \Leftrightarrow (\lambda \mathrm{id}_{K^m} - \mu_A)(v) = 0$$

ist die Äquivalenz von *(i)* und *(ii)* und der Zusatz klar. Die Äquivalenz von *(ii)* und *(iii)* folgt aus Lemma 3.2.14 und Korollar 3.2.17, denn

$$\lambda \mathrm{id}_{K^m} - \mu_A = \mu_{(\lambda I_m - A)}.$$

Die Äquivalenz von *(iii)* und *(iv)* folgt aus Satz 4.2.9. □

Definition 5.1.5. Für $A \in \mathrm{Mat}_m(K)$ heißt

$$p_A := \det\,(tI_m - A) \in K[t]$$

das **charakteristische Polynom** von A. (Streng genommen haben wir die Determinante nicht für Matrizen mit Einträgen aus dem kommutativen Ring $K[t]$ definiert. Man benutzt aber dieselbe Formel und alle wichtigen Eigenschaften bleiben erhalten.) △

Korollar 5.1.6. *Für $A \in \mathrm{Mat}_m(K)$ sind die Eigenwerte von μ_A gerade die Nullstellen von p_A in K. Somit hat μ_A höchstens m verschiedene Eigenwerte.*

Beweis. Die Aussage ist eine direkte Kombination aus Proposition 5.1.4 und Proposition 5.2.5. □

Beispiel 5.1.7. *(i)* Wir betrachten

$$A = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} \in \mathrm{Mat}_2(\mathbb{R})$$

und wollen die Eigenwerte von μ_A berechnen. Dazu berechnen wir zunächst das charakteristische Polynom

$$p_A = \det(tI_2 - A) = \det \begin{pmatrix} t+1 & -1 \\ -1 & t+1 \end{pmatrix} = (t+1)^2 - 1 = t^2 + 2t = t(t+2).$$

Die Nullstellen von p_A sind gerade 0 und -2. Es gilt $\operatorname{Eig}(\mu_A, 0) = L(-A, 0)$ und $\operatorname{Eig}(\mu_A, -2) = L(-2I_2 - A, 0)$ und wir können dafür Basen mit dem Gauß-Algorithmus finden:

$$\begin{aligned} \operatorname{Eig}(\mu_A, 0) &= \operatorname{Span}_{\mathbb{R}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ \operatorname{Eig}(\mu_A, -2) &= \operatorname{Span}_{\mathbb{R}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}. \end{aligned}$$

(ii) Wir betrachten nun

$$B = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in \operatorname{Mat}_2(\mathbb{R})$$

und berechnen

$$p_B = \det \begin{pmatrix} t & 1 \\ -1 & t \end{pmatrix} = t^2 + 1.$$

Dieses Polynom hat in $\mathbb{R}$ keine Nullstellen, also besitzt $\mu_B \colon \mathbb{R}^2 \to \mathbb{R}^2$ keine Eigenwerte und -vektoren. Es ist μ_B gerade die Drehung um den Winkel $\pi/2$ gegen den Uhrzeigersinn.
Betrachten wir B aber als Matrix über $\mathbb{C}$, so ergeben sich die Eigenwerte $-i, i$. Als Abbildung

$$\mu_B \colon \mathbb{C}^2 \to \mathbb{C}^2$$

gibt es also durchaus Eigenwerte. Man erhält

$$\begin{aligned} \operatorname{Eig}(\mu_B, -i) &= \operatorname{Span}_{\mathbb{C}} \begin{pmatrix} 1 \\ i \end{pmatrix} \\ \operatorname{Eig}(\mu_B, i) &= \operatorname{Span}_{\mathbb{C}} \begin{pmatrix} 1 \\ -i \end{pmatrix}. \end{aligned}$$

(iii) Wir betrachten

$$C = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \in \operatorname{Mat}_2(\mathbb{R})$$

und berechnen

$$p_C = \det\begin{pmatrix} t & -1 \\ 0 & t \end{pmatrix} = t^2.$$

Es besitzt μ_C also den Eigenwert 0. Der dazugehörende Eigenraum ist

$$\mathrm{Eig}(\mu_C, 0) = \mathrm{Span}_{\mathbb{R}}\begin{pmatrix} 1 \\ 0 \end{pmatrix}. \qquad \triangle$$

Proposition 5.1.8 (★). *Seien $\underline{v}$, $\underline{w}$ zwei Basen von V.*

(i) Es gibt ein $P \in \mathrm{GL}_m(K)$ mit $M_{\underline{w},\underline{w}}(\varphi) = P^{-1}M_{\underline{v},\underline{v}}(\varphi)P$.

(ii) $M_{\underline{v},\underline{v}}(\varphi)$ und $M_{\underline{w},\underline{w}}(\varphi)$ besitzen dasselbe charakteristische Polynom.

Beweis. (i) beweist man genau wie Satz 3.2.32; falls links und rechts jeweils dieselbe Basis gewählt wurde, sind die beiden dort auftretenden Matrizen P, Q gerade invers zu einander. Daraus ergibt sich (ii) mit folgender Rechnung:

$$\begin{aligned}
p_{M_{\underline{w},\underline{w}}(\varphi)} &= \det\left(tI_m - M_{\underline{w},\underline{w}}(\varphi)\right) \\
&= \det\left(tP^{-1}P - P^{-1}M_{\underline{v},\underline{v}}(\varphi)P\right) \\
&= \det\left(P^{-1}(tI_m - M_{\underline{v},\underline{v}}(\varphi))P\right) \\
&= \det\left(P^{-1}\right)\det\left(tI_m - M_{\underline{v},\underline{v}}(\varphi)\right)\det(P) \\
&= \det(P)^{-1}p_{M_{\underline{v},\underline{v}}(\varphi)}\det(P) \\
&= p_{M_{\underline{v},\underline{v}}(\varphi)}.
\end{aligned}$$

Dabei haben wir Satz 4.2.4 und Korollar 4.2.5 benutzt (genau genommen benutzen wir die Multiplikativität der Determinante für Matrizen mit Einträgen aus dem kommutativen Ring $K[t]$). □

Definition 5.1.9 (★). Das **charakteristische Polynom**

$$p_\varphi \in K[t]$$

von φ ist das charakteristische Polynom von $M_{\underline{v},\underline{v}}(\varphi)$, für eine beliebige Basis $\underline{v}$ von V. △

Korollar 5.1.10 (★). *Die Eigenwerte von φ sind gerade die Nullstellen von p_φ in K.*

Beweis. Klar mit Bemerkung 5.1.2 (vi) und Korollar 5.1.6. □

5.2 Polynome und Nullstellen

Wir haben Polynome bereits als formale Objekte der Gestalt

$$p = c_0 + c_1 t + c_2 t^2 + \cdots + c_d t^d$$

mit $c_0, \ldots, c_d \in K$ kennengelernt. Dabei heißen die c_i **Koeffizienten** des Polynoms. Man nennt c_0 den **konstanten Koeffizienten** von p. Ist $c_d \neq 0$ so nennt man

$$\deg(p) := d$$

den **Grad** des Polynoms und c_d den **Leitkoeffizienten**. Die Menge aller Polynome mit Koeffizienten aus K wird mit $K[t]$ bezeichnet, sie bildet einen kommutativen Ring.

Lemma 5.2.1. *Der Ring $K[t]$ ist* **nullteilerfrei**, *d.h. für $0 \neq p, q \in K[t]$ gilt $0 \neq pq$. Es gilt weiter*

$$\deg(pq) = \deg(p) + \deg(q).$$

In $K[t]$ gilt auch die Kürzungsregel, d.h. aus $qp_1 = qp_2$ und $q \neq 0$ folgt $p_1 = p_2$.

Beweis. Schreibe $p = b_0 + b_1 t \cdots + b_d t^d$ und $q = c_0 + c_1 t + \cdots + c_e t^e$ mit $d, e \geq 0, 0 \neq b_d, c_e$. Dann gilt

$$pq = b_0 c_0 + (b_0 c_1 + b_1 c_0)t + \cdots + \underbrace{b_d c_e}_{\neq 0} t^{d+e}.$$

Also gilt $pq \neq 0$ und $\deg(pq) = d + e = \deg(p) + \deg(q)$. Es gelte nun $qp_1 = qp_2$ mit $q \neq 0$. Das impliziert $0 = q(p_1 - p_2)$ und mit der Nullteilerfreiheit folgt $p_1 - p_2 = 0$, also $p_1 = p_2$. □

Bemerkung 5.2.2. Man kann ein Polynom $p \in K[t]$ auch als Vorschrift einer Abbildung

$$\begin{aligned} K &\to K \\ a &\mapsto p(a) = c_0 + c_1 a + c_2 a^2 + \cdots + c_d a^d \end{aligned}$$

verwenden. Dabei ist etwas Vorsicht geboten. Über *endlichen Körpern* können unterschiedliche Polynome dieselbe Abbildung definieren. In der Abbildung steckt also im Allgemeinen nicht mehr die gesamte Information über das Polynom, also über seine Koeffizienten. Für $K = \mathbb{Z}/2\mathbb{Z}$ gilt beispielsweise für

$$p = t^2 + t$$

zwar $0 \neq p \in K[t]$, p definiert aber die Nullabbildung auf K! Über unendlichen Körper passiert das jedoch nicht, wie wir gleich sehen werden. Man kann nämlich für jede Nullstelle einen *Linearfaktor* von p abspalten. △

Lemma 5.2.3. *Für $p \in K[t]$ und $a \in K$ sind äquivalent:*

(i) $p(a) = 0$.

(ii) Es gibt ein $q \in K[t]$ mit $p = (t-a)q$.

Beweis. "(ii)⇒(i)" ist klar:

$$p(a) = (a-a)q(a) = 0 \cdot q(a) = 0.$$

Für "(i)⇒(ii)" betrachten wir zunächst das Polynom

$$h(t) := p(t+a) \in K[t].$$

Es gilt

$$h(0) = p(0+a) = p(a) = 0$$

und deshalb hat h keinen konstanten Koeffizienten. Man kann nun offensichtlich t aus h herausheben, also h schreiben als $h = t \cdot g$ mit einem $g \in K[t]$. Damit gilt nun

$$p(t) = h(t-a) = (t-a) \cdot \underbrace{g(t-a)}_{=:q}. \qquad \square$$

Definition 5.2.4. Für $0 \neq p \in K[t]$ und $a \in K$ schreiben wir

$$p = (t-a)^m q$$

mit $m \geq 0, q \in K[t]$ und $q(a) \neq 0$. Dann heißt m die **Vielfachheit der Nullstelle** a von p. △

Proposition 5.2.5. *(i) Der Begriff von Vielfachheit ist wohldefiniert.*
(ii) Für $0 \neq p \in K[t]$ seien $a_1, \ldots, a_k$ paarweise verschiedene Nullstellen in K, mit Vielfachheiten $m_1, \ldots, m_k$. Dann gibt es ein $q \in K[t]$ mit

$$p = (t-a_1)^{m_1} \cdots (t-a_k)^{m_k} q$$

und $q(a_i) \neq 0$ für $i = 1, \ldots, k$.
(iii) Ein Polynom $0 \neq p \in K[t]$ hat höchstens $\deg(p)$ viele Nullstellen in K, gerechnet inklusive Vielfachheiten.
(iv) Ist K unendlich, so definieren zwei verschiedene Polynome auch verschiedene Abbildungen auf K.

Beweis. Für (i) schreiben wir

$$p = (t-a)^m q = (t-a)^n h$$

mit $q, h \in K[t]$, $q(a) \neq 0 \neq h(a)$ und nehmen dabei $m \leq n$ an. Dann gilt

$$0 = p - p = (t-a)^m q - (t-a)^n h = \underbrace{(t-a)^m}_{\neq 0} \left(q - (t-a)^{n-m} h\right).$$

Aus der Nullteilerfreiheit von $K[t]$ folgt $0 = q - (t-a)^{n-m}h$, also

$$q = (t-a)^{n-m} h.$$

Aus $0 \neq q(a) = (a-a)^{n-m} h(a)$ folgt nun $n = m$.
Für (ii) schreiben wir zunächst $p = (t-a_1)^{n_1} \tilde{q}$ mit einem $\tilde{q} \in K[t]$ und $\tilde{q}(a_1) \neq 0$. Aus

$$0 = p(a_2) = (\underbrace{a_2 - a_1}_{\neq 0})^{n_1} \tilde{q}(a_2)$$

folgt in K nun $\tilde{q}(a_2) = 0$. Wir iterieren den Prozess mit $\tilde{q}$ und allen weiteren a_i. Dadurch erhalten wir eine Darstellung

$$p = (t-a_1)^{n_1} \cdots (t-a_k)^{n_k} q$$

mit einem $q \in K[t]$ und $q(a_i) \neq 0$ für alle $i = 1, \ldots, k$. Für jedes $i = 1, \ldots, k$ gilt nun

$$p = (t-a_i)^{n_i} \cdot \underbrace{q \cdot \prod_{j \neq i} (t-a_j)^{n_j}}_{=:h}$$

mit $h(a_i) \neq 0$. Aus (i) folgt also $n_i = m_i$.
Für (iii) seien $a_1, \ldots, a_k$ Nullstellen von p in K mit Vielfachheiten $m_1, \ldots, m_k$. Wie in (ii) schreiben wir $p = (t-a_1)^{m_1} \cdots (t-a_k)^{m_k} q$ und erhalten mit Lemma 5.2.1

$$\deg(p) = m_1 + \cdots + m_k + \deg(q) \geq m_1 + \cdots + m_k.$$

Für (iv) seien $p, q \in K[t]$ mit $p(a) = q(a)$ für alle $a \in K$. Dann sind alle Elemente von K Nullstellen des Polynoms $p - q$. Da K unendlich ist folgt $p - q = 0$, also $p = q$. □

Definition 5.2.6. Ein Polynom $p \in K[t]$ **zerfällt über** K **in Linearfaktoren**, falls es $c, a_1, \ldots, a_k \in K$ sowie $m_1, \ldots, m_k \in \mathbb{N}$ gibt mit

$$p = c \cdot (t-a_1)^{m_1} \cdots (t-a_k)^{m_k}. \qquad \triangle$$

Beispiel 5.2.7. (i) Das Polynom $p = t^2 - 1 \in \mathbb{Q}[t]$ zerfällt über $\mathbb{Q}$ in Linearfaktoren. Es gilt ja
$$p = (t-1)(t+1).$$
Die beiden Nullstellen $-1, 1$ treten mit Vielfachheit 1 auf.
(ii) Das Polynom $q = 1 - t - t^4 + t^5$ zerfällt über $\mathbb{Q}$ nicht in Linearfaktoren. Man berechnet
$$q = (t-1)^2(t+1)(t^2+1)$$
und sieht, dass die Nullstelle 1 mit Vielfachheit 2 und die Nullstelle -1 mit Vielfachheit 1 auftritt. Weitere Nullstellen in $\mathbb{Q}$ hat q offensichtlich nicht, also kann es über $\mathbb{Q}$ auch nicht in Linearfaktoren zerfallen. △

Korollar 5.2.8. *Jedes Polynom $p \in \mathbb{C}[t]$ zerfällt über $\mathbb{C}$ in Linearfaktoren.*

Beweis. Ist p ein konstantes Polynom, so ist die Aussage klar. Ansonsten besitzt p nach Satz 2.2.30 eine Nullstelle $a \in \mathbb{C}$. Nach Lemma 5.2.3 erhalten wir also
$$p = (t-a)q$$
für ein $q \in \mathbb{C}[t]$. Wir iterieren den Prozess nun mit q und erhalten so das Ergebnis. □

Beispiel 5.2.9. Das Polynom $q = 1 - t - t^4 + t^5$ aus Beispiel 5.2.7 (ii) zerfällt über $\mathbb{C}$:
$$q = (t-1)^2(t+1)(t-i)(t+i).$$ △

5.3 Diagonalisierung

Auch in diesem Abschnitt sei V stets ein endlich-dimensionaler K-Vektorraum und $\varphi\colon V \to V$ eine lineare Abbildung.

Definition 5.3.1. Die lineare Abbildung φ heißt **diagonalisierbar**, falls eine Basis $\underline{v}$ von V existiert, so dass $M_{\underline{v},\underline{v}}(\varphi)$ eine Diagonalmatrix ist (d.h. höchstens auf der Diagonale stehen Einträge ungleich 0). △

Bemerkung 5.3.2. Ist v ein Eigenvektor zum Eigenwert λ von φ und v der i-te Vektor in einer Basis $\underline{v}$ von V, so besteht die i-te Spalte von $M_{\underline{v},\underline{v}}(\varphi)$ genau aus λe_i. Das ist unmittelbar klar aus Konstruktion 3.2.27 für Darstellungsmatrizen. Viele Eigenvektoren in einer Basis führen also zu einer einfachen Struktur der Darstellungsmatrix. △

Definition 5.3.3. Sei $\lambda \in K$ ein Eigenwert von φ.

(i) $\dim_K(\operatorname{Eig}(\varphi, \lambda))$ heißt **geometrische Vielfachheit** des Eigenwerts λ.

(ii) Die Vielfachheit von λ als Nullstelle von p_φ heißt **algebraische Vielfachheit** des Eigenwerts λ. △

Beispiel 5.3.4. (i) In Beispiel 5.1.7 (i) hat μ_A die beiden Eigenwerte 0 und -2. Beide haben sowohl geometrische als auch algebraische Vielfachheit 1.
(ii) In Beispiel 5.1.7 (ii) hat μ_B über $\mathbb{R}$ gar keine Eigenwerte, über $\mathbb{C}$ jedoch die Eigenwerte $-i, i$, mit jeweils sämtlichen Vielfachheiten gleich 1.
(iii) In Beispiel 5.1.7 (iii) hat μ_C nur den Eigenwert 0. Die algebraische Vielfachheit ist hier 2, die geometrische nur 1. △

Satz 5.3.5. *(i) Die geometrische Vielfachheit eines Eigenwerts ist stets kleiner gleich der algebraischen Vielfachheit.*
(ii) Sind $v_1, \ldots, v_k$ Eigenvektoren von φ zu paarweise unterschiedlichen Eigenwerten, so sind $v_1, \ldots v_k$ linear unabhängig.

Beweis. Für (i) wählen wir eine Basis $v_1, \ldots, v_r$ von $\operatorname{Eig}(\varphi, \lambda)$ und ergänzen sie zu einer Basis

$$\underline{v} = (v_1, \ldots, v_r, v_{r+1}, \ldots, v_m)$$

von V. Da $v_1, \ldots, v_r$ Eigenvektoren von φ zum Eigenwert λ sind, sieht man mit Bemerkung 5.3.2 sofort, dass die Darstellungsmatrix $M_{\underline{v},\underline{v}}(\varphi)$ folgende Gestalt hat:

$$M_{\underline{v},\underline{v}}(\varphi) = \left(\begin{array}{c|c} \lambda I_r & A \\ \hline 0 & B \end{array}\right).$$

Damit gilt

$$p_\varphi = \det\left(\begin{array}{c|c} (t-\lambda)I_r & -A \\ \hline 0 & C \end{array}\right) = \det\left((t-\lambda)I_r\right)\det(C) = (t-\lambda)^r \det(C).$$

Die Multiplikativität der Determinante in der zweiten Gleichheit sieht man analog zu Bemerkung 4.2.2 (iv) (Übungsaufgabe). Somit ist λ eine Nullstelle von p_φ mit Vielfachheit mindestens r und das ist genau die zu beweisende Aussage.
Aussage (ii) beweisen wir mit Induktion über k. Der Fall $k = 1$ ist klar, denn Eigenvektoren sind nach Definition stets ungleich 0. Sei nun jedes v_i jeweils Eigenvektor zum Eigenwert λ_i. Wir nehmen

$$\gamma_1 v_1 + \cdots + \gamma_k v_k = 0$$

für gewisse $\gamma_i \in K$ an. Einerseits multiplizieren wir diese Gleichung mit λ_k, andererseits wenden wir φ auf sie an. So erhalten wir die folgenden beiden Gleichungen:

$$0 = \gamma_1\lambda_k v_1 + \gamma_2\lambda_k v_2 + \cdots + \gamma_k\lambda_k v_k$$
$$0 = \gamma_1\lambda_1 v_1 + \gamma_2\lambda_2 v_2 + \cdots + \gamma_k\lambda_k v_k.$$

Ziehen wir sie voneinander ab, hebt sich der letzte Term der Summe weg:

$$0 = \gamma_1(\lambda_k - \lambda_1)v_1 + \cdots + \gamma_{k-1}(\lambda_k - \lambda_{k-1})v_{k-1}.$$

Nach Induktionsvoraussetzung erhalten wir

$$0 = \gamma_i(\lambda_k - \lambda_i)$$

für alle $i = 1, \dots, k-1$. Da die λ_i paarweise verschieden sind, folgt $\lambda_k - \lambda_i \neq 0$ und damit $\gamma_i = 0$ für alle $i = 1, \dots, k-1$. Daraus folgt schließlich auch $\gamma_k = 0$. □

Bemerkung/Beispiel 5.3.6. (i) Im Raum $C^1(\mathbb{R}, \mathbb{R})$ aller stetig differenzierbaren Funktionen sind

$$e^{\alpha_1 x}, \dots, e^{\alpha_k x}$$

jeweils Eigenvektoren der Differentiationsabbildung ∂ zu den Eigenwerten $\alpha_1, \dots, \alpha_k$. Sind die α_i paarweise verschieden, sind die Funktionen also linear unabhängig.
(ii) Sind $\lambda_1, \dots, \lambda_k$ die paarweise verschiedenen Eigenwerte von φ, wählen wir für jedes $i = 1, \dots, k$ eine Basis

$$\underline{v}^i = (v_1^i, \dots, v_{m_i}^i)$$

von $\mathrm{Eig}(\varphi, \lambda_i)$. Dann ist

$$\underline{\tilde{v}} = (\underline{v}^1, \dots, \underline{v}^k) = (v_1^1, \dots, v_{m_1}^1, \dots, v_1^k, \dots, v_{m_k}^k)$$

linear unabhängig. Aus

$$0 = \underbrace{\lambda_1^1 v_1^1 + \cdots + \lambda_{m_1}^1 v_{m_1}^1}_{\in \mathrm{Eig}(\varphi, \lambda_1)} + \cdots + \underbrace{\lambda_1^k v_1^k + \cdots + \lambda_{m_k}^k v_{m_k}^k}_{\in \mathrm{Eig}(\varphi, \lambda_k)}$$

folgt mit Satz 5.3.5 schon

$$\lambda_1^i v_1^i + \cdots + \lambda_{m_i}^i v_{m_i}^i = 0$$

für alle $i = 1, \ldots, k$ und daraus λ^i_j für $j = 1, \ldots, m_i$, da $\underline{v}^i$ linear unabhängig ist. Ergänzen wir $\tilde{\underline{v}}$ zu einer Basis $\underline{v}$ von V, so ergibt sich für die Darstellungsmatrix folgende Blockgestalt:

$$M_{\underline{v},\underline{v}}(\varphi) = \begin{pmatrix} \lambda_1 I_{m_1} & & & * \\ & \ddots & & * \\ & & \lambda_k I_{m_k} & * \\ & & & * \end{pmatrix}.$$

Sollte $\tilde{\underline{v}}$ bereits eine Basis von V sein, so ist φ diagonalisierbar.
(iii) Hat φ im Fall $\dim_K(V) = m$ genau m paarweise verschiedene Eigenwerte, so ist φ diagonalisierbar. Jeder Eigenraum enthält mindestens einen nichttrivialen Vektor, damit ergibt sich mit der Prozedur aus (ii) bereits eine Basis von V aus Eigenvektoren.
(iv) Die lineare Abbildung μ_A aus Beispiel 5.1.7 (i) besitzt zwei verschiedene Eigenwerte, ist also diagonalisierbar. Mit Proposition 5.1.8 bedeutet das also, dass ein $P \in \mathrm{GL}_2(\mathbb{R})$ existiert, so dass $P^{-1}AP$ Diagonalgestalt hat. $\triangle$

Satz 5.3.7. *Die folgenden Aussagen sind äquivalent:*

(i) *φ ist diagonalisierbar.*

(ii) *p_φ zerfällt über K in Linearfaktoren und für jeden Eigenwert stimmen geometrische und algebraische Vielfachheit überein.*

(iii) *V besitzt eine Basis aus Eigenvektoren von φ.*

In diesem Fall stehen in jeder diagonalen Darstellungsmatrix genau die Eigenwerte gemäß ihrer Vielfachheiten auf der Diagonalen.

Beweis. "$(i)\Rightarrow(ii)$": Ist $A = M_{\underline{v},\underline{v}}(\varphi) = \mathrm{diag}(\lambda_1, \ldots, \lambda_m)$ eine Diagonalmatrix, so gilt

$$p_\varphi = \det(tI_m - A) = (t - \lambda_1) \cdots (t - \lambda_m),$$

also zerfällt p_φ über K in Linearfaktoren. Dabei ist die algebraische Vielfachheit eines λ_i gerade die Häufigkeit, mit der es in der Diagonalen von A auftritt. Genauso viele verschiedene Standandardbasisvektoren sind aber dann auch Eigenvektoren von μ_A zum Eigenvektor λ_i. Standardbasisvektoren sind aber linear unabhängig, also stimmt die geometrische Vielfachheit von λ_i mit der algebraischen überein, für μ_A und somit auch für φ.

Für "$(ii)\Rightarrow(iii)$" beobachten wir, dass die Summe der geometrischen Vielfachheiten aller Eigenwerte gleich der Summe der algebraischen Vielfachheiten ist. Diese wiederum ergibt genau den Grad von p_φ und damit die Dimension von V. Wenn wir also für jeden Eigenraum eine Basis wählen, erhalten wir insgesamt $\dim_K(V)$ viele linear unabhängige Eigenvektoren (vergleiche Bemerkung 5.3.6 (ii)), also eine Basis von V.
"$(iii)\Rightarrow(i)$" ist klar, denn für eine Basis $\underline{v}$ aus Eigenvektoren von φ hat die Darstellungsmatrix $M_{\underline{v},\underline{v}}(\varphi)$ offensichtlich Diagonalgestalt (vergleiche auch Bemerkung 5.3.2). Die Einträge der Diagonalen sind dabei genau die Eigenwerte von φ. □

Beispiel 5.3.8. (i) Für die Matrix A aus Beispiel 5.1.7 (i) ist μ_A diagonalisierbar. Es existiert also ein $P \in \mathrm{GL}_2(\mathbb{R})$ mit

$$P^{-1}AP = \begin{pmatrix} 0 & 0 \\ 0 & -2 \end{pmatrix}.$$

(ii) Die lineare Abbildung μ_B aus Beispiel 5.1.7 (ii) besitzt in $\mathbb{R}$ gar keine Eigenwerte, ist also nicht diagonalisierbar. Über $\mathbb{C}$ besitzt sie aber die beiden Eigenwerte $-i, i$. Es existiert also ein $P \in \mathrm{GL}_2(\mathbb{C})$ mit

$$P^{-1}BP = \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix}.$$

(iii) Die lineare Abbildung μ_C aus Beispiel 5.1.7 (iii) besitzt den Eigenwert 0 mit algebraischer Vielfachheit 2 und geometrischer Vielfachheit 1. Also ist μ_C nicht diagonalisierbar. Daran ändert sich auch beim Übergang zu $\mathbb{C}$ (oder sonst einem Körper) nichts. △

Konstruktion 5.3.9. Oft ist eine Matrix $A \in \mathrm{Mat}_m(K)$ gegeben, die man diagonalisieren möchte. Gesucht ist also ein $P \in \mathrm{GL}_m(K)$ so dass $P^{-1}AP$ Diagonalgestalt hat. Dazu berechnet man zunächst p_A und alle seine Nullstellen in K. Gibt es (inklusive Vielfachheiten) weniger als m solche Nullstellen, ist μ_A (bzw. A) nicht diagonalisierbar, zumindest nicht über K.
Im anderen Fall berechnet man die Dimensionen aller Eigenräume (siehe Proposition 5.1.4). Stimmt eine dieser geometrischen Vielfachheiten nicht mit der algebraischen überein, ist A nicht diagonalisierbar (das hängt nicht vom Körper ab).
Im anderen Fall wählt man für jeden Eigenraum eine Basis und erhält so eine Basis $\underline{v} = (v_1, \ldots, v_m)$ von K^m aus Eigenvektoren von μ_A. Für die Matrix P wählen wir nun genau die v_i als Spalten:

$$P = (v_1, \ldots, v_m) \in \mathrm{GL}_m(K).$$

Dann ist P offenbar genau die Matrix des Basiswechsels der Standardbasis zu $\underline{v}$ und deshalb hat $P^{-1}AP$ Diagonalgestalt. Man kann diese Tatsache aber auch nochmals direkt nachrechnen. Für $i, j = \{1, \ldots, m\}$ gilt:

$$e_i^t \left(P^{-1}AP\right) e_j = e_i^t P^{-1} A v_j = e_i^t P^{-1} \lambda_j v_j = \lambda_j e_i^t e_j = \begin{cases} 0 & : \quad i \neq j \\ \lambda_j & : \quad i = j \end{cases}$$

und das ist einerseits genau der (i, j)-Eintrag von $P^{-1}AP$, andererseits der von $\operatorname{diag}(\lambda_1, \ldots, \lambda_m)$. △

Beispiel 5.3.10. (i) Wir betrachten

$$A = \begin{pmatrix} 1 & -1 & 1 \\ -1 & 1 & -1 \\ 1 & -1 & 1 \end{pmatrix} \in \mathrm{Mat}_3(\mathbb{Q})$$

und berechnen

$$p_A = \det \begin{pmatrix} t-1 & 1 & -1 \\ 1 & t-1 & 1 \\ -1 & 1 & t-1 \end{pmatrix} = t^2(t-3) \in \mathbb{Q}[t].$$

Die Nullstellen sind $\lambda_1 = 0$ und $\lambda_2 = 3$, wobei λ_1 algebraische Vielfachheit 2 und λ_2 algebraische Vielfachheit 1 hat. Wir berechnen nun

$$\mathrm{Eig}(\mu_A, \lambda_1) = L(-A, 0) = \mathrm{Span}_{\mathbb{Q}} \left\{ \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \right\},$$

wobei die beiden Vektoren rechts eine Basis des Eigenraums bilden. Die geometrischen Vielfachheit von λ_1 ist also ebenfalls 2. Weiter ist

$$\mathrm{Eig}(\mu_A, \lambda_2) = L(3I_3 - A, 0) = \mathrm{Span}_{\mathbb{Q}} \left\{ \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} \right\}.$$

Für

$$P = \begin{pmatrix} -1 & 1 & 1 \\ 0 & 1 & -1 \\ 1 & 0 & 1 \end{pmatrix}$$

gilt also $P \in \mathrm{GL}_3(\mathbb{Q})$ und

$$P^{-1}AP = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

(ii) Angenommen wir wollen eine allgemeine Formel für A^n finden, mit A wie in (i). Natürlich könnte man die ersten Potenzen ausrechnen, versuchen die richtige Formel zu erraten und sie dann allgemein mit Induktion über n zu beweisen. Wir machen das jetzt aber deutlich eleganter. Für $D = \mathrm{diag}(0,0,3)$ gilt $A = PDP^{-1}$ und damit für jedes $n \in \mathbb{N}$

$$A^n = \left(PDP^{-1}\right)^n = PD\underbrace{P^{-1}P}_{=I_3}D\underbrace{P^{-1}P}_{=I_3}\cdots\underbrace{P^{-1}P}_{=I_3}DP^{-1} = PD^nP^{-1}.$$

Offensichtlich ist aber $D^n = \mathrm{diag}(0,0,3^n)$ und wir erhalten

$$A^n = \frac{1}{3}\cdot\begin{pmatrix} -1 & 1 & 1 \\ 0 & 1 & -1 \\ 1 & 0 & 1 \end{pmatrix}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 3^n \end{pmatrix}\begin{pmatrix} -1 & 1 & 2 \\ 1 & 2 & 1 \\ 1 & -1 & 1 \end{pmatrix} = 3^{n-1}A. \qquad \triangle$$

5.4 Trigonalisierung

Wir wissen nun, dass man manche lineare Abbildungen/Matrizen nicht diagonalisieren kann, unabhängig vom Körper. Das einfachste Beispiel ist

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

Diese Matrix hat aber zumindest obere Dreiecksgestalt. Wir beschäftigen uns nun mit dieser Gestalt.

Definition 5.4.1. Die lineare Abbildung φ heißt **trigonalisierbar**, falls eine Basis $\underline{v}$ von V existiert, so dass $M_{\underline{v},\underline{v}}(\varphi)$ obere Dreiecksgestalt hat. $\triangle$

Satz 5.4.2. *Die folgenden Aussagen sind äquivalent:*

(i) φ *ist trigonalisierbar.*

(ii) p_φ *zerfällt über* K *in Linearfaktoren.*

Auf jeder Darstellungsmatrix von φ in oberer Dreiecksgestalt stehen dann die Eigenwerte von φ gemäß ihrer algebraischen Vielfachheiten auf der Diagonalen.

Beweis. "$(i) \Rightarrow (ii)$" ist leicht: Ist $\underline{v}$ eine Basis, bezüglich der

$$A = M_{\underline{v},\underline{v}}(\varphi) = \begin{pmatrix} \lambda_1 & * & * \\ & \ddots & * \\ 0 & & \lambda_m \end{pmatrix}$$

obere Dreiecksgestalt hat, so gilt

$$p_\varphi = \det(tI_m - A) = (t - \lambda_1) \cdots (t - \lambda_m).$$

Dieses Polynom zerfällt in Linearfaktoren und seine Nullstellen (also Eigenwerte von φ) sind genau die Diagonaleinträge von A.
Die Aussage "$(ii) \Rightarrow (i)$" beweisen wir per Induktion über m, der Fall $m = 1$ ist dabei klar. Im allgemeinen Fall schreiben wir

$$p_\varphi = (t - \lambda) \cdot q$$

für ein $q \in K[t]$, das nach Voraussetzung natürlich auch selbst über K wieder in Linearfaktoren zerfällt. Da λ als Nullstelle von p_φ ein Eigenwert von φ ist, wählen wir dazu einen Eigenvektor $0 \neq v_1 \in V$ und ergänzen ihn zu einer Basis $\underline{v} = (v_1, v_2, \ldots, v_m)$ von V. Dann hat die Darstellungsmatrix folgende Gestalt:

$$M_{\underline{v},\underline{v}}(\varphi) = \begin{pmatrix} \lambda & * & * & * \\ 0 & & & \\ \vdots & & A & \\ 0 & & & \end{pmatrix}$$

und es gilt

$$p_\varphi = (t - \lambda) \cdot p_A.$$

Aus der Kürzungseigenschaft in Lemma 5.2.1 folgt $p_A = q$. Nun betrachten wir

$$W := \operatorname{Span}_K\{v_2, \ldots, v_m\} \subseteq V$$

und definieren die Projektion von V auf W als folgende lineare Abbildung:

$$\pi \colon V \to W$$
$$\gamma_1 v_1 + \ldots + \gamma_m v_m \mapsto \gamma_2 v_2 + \cdots + \gamma_m v_m.$$

Damit erhalten wir die lineare Abbildung

$$\psi := \pi \circ \varphi_{|W} : W \to W.$$

Es gilt dabei mit $\underline{w} = (v_2, \ldots, v_m)$ offensichtlich

$$M_{\underline{w},\underline{w}}(\psi) = A$$

und somit zerfällt $p_\psi = p_A = q$ über K in Linearfaktoren. Wir können auf ψ also die Induktionsvoraussetzung anwenden und erhalten eine Basis $w_2, \ldots, w_m$ von W, bezüglich der die Darstellungsmatrix von ψ obere Dreiecksgestalt hat. Dann liefert aber $(v_1, w_2, \ldots, w_m)$ ebenfalls eine Darstellungsmatrix von ψ in oberer Dreiecksgestalt. □

Bemerkung 5.4.3. Theoretisch kann man eine geeignete Basis algorithmisch produzieren, für welche die Darstellungsmatrix von φ obere Dreiecksgestalt hat. Man muss dabei induktiv über die Dimension vorgehen, genau wie im letzten Beweis.

Korollar 5.4.4. *Im Fall $K = \mathbb{C}$ ist jede lineare Abbildung φ (auf einem endlichdimensionalem Raum V) trigonalisierbar.*

Beweis. Nach Satz 2.2.30 zerfällt jedes Polynom über $\mathbb{C}$ in Linearfaktoren. □

Beispiel 5.4.5. Wir betrachten $\mu_A : \mathbb{R}^2 \to \mathbb{R}^2$ für

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

Es gilt $p_A = t^2 + 1$ und dieses Polynom zerfällt über $\mathbb{R}$ nicht in Linearfaktoren. Also ist μ_A über $\mathbb{R}$ nicht trigonalisierbar.
Betrachten wir stattdessen $\mu_A : \mathbb{C}^2 \to \mathbb{C}^2$, so gilt

$$p_A = t^2 + 1 = (t - i)(t + i)$$

und μ_A ist trigonalisierbar. Nach Bemerkung 5.3.6 (iii) ist μ_A hier sogar diagonalisierbar mit Diagonalmatrix

$$\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$$ △

5.5 Die Jordan'sche Normalform

Wir haben bereits gesehen, dass man selbst über $\mathbb{C}$ nicht jede Abbildung/Matrix diagonalisieren kann. Für jedes $\lambda \in \mathbb{C}, m \in \mathbb{N}$ ist

$$J(\lambda, m) := \begin{pmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix} \in \mathrm{Mat}_m(\mathbb{C})$$

ein solches Beispiel. Es gilt

$$p_{J(\lambda,m)} = (t-\lambda)^m,$$

also besitzt $\mu_{J(\lambda,m)}$ nur den Eigenwert λ, in algebraischer Vielfachheit m. Für den Eigenraum gilt aber

$$\mathrm{Eig}(\mu_{J(\lambda,m)}, \lambda) = \mathrm{Span}\{e_1\},$$

die geometrische Vielfachheit ist also stets nur 1. Der Satz über die Jordan'schen Normalform besagt, dass über $\mathbb{C}$ zumindest nur diese Fälle auftreten können.

Definition 5.5.1. Die Matrix $J(\lambda, m) \in \mathrm{Mat}_m(K)$ heißt **Jordanblock der Größe m zum Eigenwert** λ. △

Bemerkung 5.5.2. (i) Berechnet man Potenzen von $J(0, m)$, so rücken die Einsen auf der Nebendiagonalen schrittweise immer weiter nach rechts oben. In der m-ten Potenz erhält man dann erstmals die Nullmatrix. Es ist $J(0, m)$ also eine sogenannte *nilpotente* Matrix. Der Dimension des Kerns wächst bei jeder Potenz genau um eins.
(ii) Wir betrachten die folgende Blockmatrix:

$$\left(\begin{array}{c|c} J(0,4) & 0 \\ \hline 0 & J(0,3) \end{array}\right).$$

Potenzen werden hier blockweise gebildet, also ist der untere rechte Block nach der dritten Potenz die Nullmatrix, der obere linke Block erst nach der vierten Potenz. Auch diese Matrix ist also nilpotent. Beim Potenzieren wächst die Kerndimension zunächst um jeweils 2, nach der dritten Potenz nur noch um 1. △

Proposition 5.5.3. *Sei V ein m-dimensionaler K-Vektorraum und $\psi\colon V \to V$ ein nilpotenter Endomorphismus, d.h. es gelte*

$$\psi^n := \underbrace{\psi \circ \psi \circ \cdots \circ \psi}_{n} = 0$$

für ein $n \in \mathbb{N}$. Dann gibt es eine Basis von V, bezüglich der die Darstellungsmatrix von ψ die Gestalt

$$\begin{pmatrix} 0 & \delta_1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & \delta_{m-1} \\ 0 & & & 0 \end{pmatrix}$$

mit $\delta_i \in \{0,1\}$ hat, also blockdiagonal mit Jordanblöcken zum Eigenwert 0 ist.

Beweis. Wir setzen $U_i := \mathrm{Kern}(\psi^i)$ und erhalten eine aufsteigende Kette von Untervektorräumen

$$\{0\} \subseteq U_1 \subseteq U_2 \subseteq \cdots$$

mit

$$\psi\left(U_{i+1}\right) \subseteq U_i$$

für alle i. Wir wählen s minimal mit $U_s = V$, ein solches $s \leq n$ existiert nach Voraussetzung. Wir konstruieren nun die gewünschte Basis von V durch folgende Methode:
Zunächst wählen wir irgendeine Basis von U_{s-1} und ergänzen sie durch Hinzunahme von Vektoren

$$\underline{v}^s := (v_1^s, \ldots, v_{r_s}^s)$$

zu einer Basis von $V = U_s$. Dann liegen alle $\psi(v_i^s)$ in U_{s-1}, aber nur die triviale Linearkombination von

$$\psi(\underline{v}^s) := (\psi(v_1^s), \ldots, \psi(v_{r_s}^s))$$

liegt in U_{s-2}. Aus

$$U_{s-2} \ni \sum_i \gamma_i \psi(v_i^s) = \psi\left(\sum_i \gamma_i v_i^s\right)$$

folgt nämlich $\sum_i \gamma_i v_i^s \in U_{s-1}$, was nach Wahl der v_i^s bereits $\gamma_i = 0$ für alle i impliziert. Insbesondere ist $\psi(\underline{v}^s)$ auch linear unabhängig.

Wir wählen dann irgendeine Basis von U_{s-2} und ergänzen sie zusammen mit $\psi(\underline{v}^s)$ zu einer Basis

$$\underline{v}^{s-1} := \big(\psi(\underline{v}^s), v_1^{s-1}, \ldots, v_{r_{s-1}}^{s-1}\big)$$

von U_{s-1}. Dieser Prozess wird nun iteriert:

$$\underline{v}^{i} := \big(\psi(\underline{v}^{i+1}), v_1^{i}, \ldots, v_{r_i}^{i}\big)$$

wird konstruiert als Ergänzung einer beliebigen Basis von U_{i-1} zu einer Basis von U_i. Wie wir oben schon gesehen habe ist das immer möglich. Nach Konstruktion erhalten damit dann insgesamt eine Basis $(\underline{v}^s, \underline{v}^{s-1}, \ldots, \underline{v}^1)$ von V. Diese Basis sortieren wir jetzt um, und zwar folgendermaßen:

$$\begin{array}{r}
\psi^{s-1}(v_1^s),\ \psi^{s-2}(v_1^s),\ \ldots,\ \psi(v_1^s), v_1^s \\
\psi^{s-1}(v_2^s),\ \psi^{s-2}(v_2^s),\ \ldots,\ \psi(v_2^s), v_2^s \\
\vdots \quad \\
\psi^{s-1}(v_{r_s}^s),\ \psi^{s-2}(v_{r_s}^s),\ \ldots,\ \psi(v_{r_s}^s), v_{r_s}^s \\
\psi^{s-2}(v_1^{s-1}),\ \ldots,\ \psi(v_1^{s-1}), v_1^{s-1} \\
\vdots \quad \\
v_1^1 \\
\vdots \quad \\
v_{r_1}^1.
\end{array}$$

Daraus ergibt sich die geforderte Darstellungsmatrix offensichtlich. □

Bemerkung 5.5.4. Die Konstruktion aus dem letzten Beweis liefert uns zusätzlich eine Aussage über die Anzahl der Nullen und Einsen auf der oberen Nebendiagonalen der Darstellungsmatrix von φ:
(i) Jeder der Vektoren aus $\underline{v}^s$ führt genau zu $s-1$ Einsen, also zu einem Jordanblock $J(0,s)$. Davon gibt es also genau r_s viele. Allgemein führt jeder Vektor aus $\underline{v}^i \setminus \psi(\underline{v}^{i+1})$ zu einem Jordanblock $J(0,i)$. Davon gibt es genau r_i viele. Daraus ergibt sich eindeutig die Struktur der Darstellungsmatrix.
(ii) Gewöhnlich rechnet man nicht die Zahlen r_i direkt aus, sondern die einzelnden Dimensionen $d_i := \dim(U_i)$. Damit kann man natürlich auch die r_i berechnen. Es gilt

$$r_s = d_s - d_{s-1}$$

und für $i < s$

$$\begin{aligned} r_i &= \dim(U_i) - \dim(U_{i-1}) - (\dim(U_{i+1} - \dim(U_i)) \\ &= 2d_i - d_{i-1} - d_{i+1}. \end{aligned}$$ △

Beispiel 5.5.5. Ist $\varphi\colon K^6 \to K^6$ nilpotent mit

$$d_1 = 2,\ d_2 = 4,\ d_3 = 5,\ d_4 = 6$$

so ergibt sich $r_4 = 1, r_3 = 0, r_2 = 1, r_1 = 0$. Wir erhalten also folgende Darstellungsmatrix von φ:

$$\left(\begin{array}{c|c} J(0,4) & 0 \\ \hline 0 & J(0,2) \end{array}\right)$$

Man kann hier die Dimensionen d_i auch nochmals direkt überprüfen. △

Proposition 5.5.6. *Es sei V ein endlich dimensionaler K-Vektorraum sowie $\varphi\colon V \to V$ ein Endomorphismus. Dann gibt es φ-invariante Untervektorräume $N, U \subseteq V$ mit*

$$V = N + U, \qquad N \cap U = \{0\}$$

und $\varphi_{|N}\colon N \to N$ nilpotent sowie $\varphi_{|U}\colon U \to U$ invertierbar. Dabei ist $\dim(N)$ die algebraische Vielfachheit von 0 als Eigenwert von φ.

Beweis. Wir setzen $N := \bigcup_{i\geq 1} \operatorname{Kern}(\varphi^i)$, wobei die Vereinigung nach Bemerkung 3.1.24 sogar endlich ist. So erhalten wir offensichtlich einen φ-invarianten Unterraum, auf dem φ nilpotent ist. Wenn $N = \operatorname{Kern}(\varphi^s)$ gilt, dann setzen wir $U := \operatorname{Bild}(\varphi^s)$. Dann gilt

$$\varphi(U) = \varphi^{s+1}(V) = \varphi^s\left(\varphi(V)\right) \subseteq \varphi^s(V) = U,$$

also ist U φ-invariant.
Es gilt nun $N \cap U = \{0\}$, denn für $v \in N \cap U$ gilt einerseits $v = \varphi^s(w)$ für ein $w \in V$, andererseits

$$0 = \varphi^s(v) = \varphi^{2s}(w),$$

also $w \in N$ und damit $0 = \varphi^s(w) = v$. Daraus folgt erst recht die Invertierbarkeit von φ auf U, denn

$$\operatorname{Kern}(\varphi) \cap U \subseteq N \cap U = \{0\}.$$

Nach Satz 3.2.16 und Satz 3.1.25 gilt

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

Daraus folgt $V = N + U$.
Wählen wir nun eine Basis von N wie in Proposition 5.5.3 und irgendeine Basis von U, so erhalten wir eine Darstellungsmatrix für φ der Form

$$\left(\begin{array}{c|c} A & 0 \\ \hline 0 & B \end{array}\right),$$

wobei A eine obere Dreiecksmatrix mit Nulldiagonale der Größe $\dim(N)$ und B invertierbar ist. Daraus ergibt sich

$$p_\varphi = t^{\dim(N)} \cdot p_B$$

und p_B hat 0 nicht als Nullstelle. Damit ist $\dim(N)$ gerade die algebraische Vielfachheit von 0 als Eigenwert von φ. □

Bemerkung 5.5.7. Wenn $\mathrm{Kern}(\varphi^j) = \mathrm{Kern}(\varphi^{j+1})$ gilt, dann gilt auch $\mathrm{Kern}(\varphi^j) = \mathrm{Kern}(\varphi^{j+i})$ für alle $i \geq 1$. Das sieht man induktiv sofort mit folgenden Argument:

$$\mathrm{Kern}(\varphi^{j+i}) = \varphi^{-1}(\mathrm{Kern}(\varphi^{j+i-1})) = \varphi^{-1}(\mathrm{Kern}(\varphi^j)) = \mathrm{Kern}(\varphi^{j+1}) = \mathrm{Kern}(\varphi^j).$$

Wenn die aufsteigende Kette der Unterräume also erstmals stabil bleibt, bleibt sie dauerhaft stabil. Das passiert aus Dimensionsgründen also spätestens nach n Schritten, wobei n die algebraische Vielfachheit von 0 als Eigenwert von φ, bzw. die Dimension von N ist. △

Bemerkung 5.5.8. Sei $\varphi\colon V \to V$ ein Endomorphismus mit Eigenwert $\lambda \in K$. Dann hat für jedes $\gamma \in K$ der Endomorphismus

$$\gamma \,\mathrm{id}_V - \varphi\colon V \to V$$

den Eigenwert $\gamma - \lambda$, mit gleicher algebraischer und geometrischer Vielfachheit wie λ für φ. △

Definition 5.5.9. Es sei V ein endlich-dimensionaler K-Vektorraum, $\varphi\colon V \to V$ ein Endomorphismus und $\lambda \in K$ ein Eigenwert von φ mit algebraische Vielfachheit n. Dann ist 0 ein Eigenwert von

$$\lambda \,\mathrm{id}_V - \varphi$$

mit algebraischer Vielfachheit n und wir nennen

$$\widetilde{\mathrm{Eig}}(\varphi, \lambda) := \bigcup_{i \geq 1} \mathrm{Kern}\left((\lambda \,\mathrm{id}_V - \varphi)^i\right) = \mathrm{Kern}\left((\lambda \,\mathrm{id}_V - \varphi)^n\right)$$

den **verallgemeinerten Eigenraum/Hauptraum** von φ zum Eigenwert λ. Dabei gilt

$$\dim\left(\widetilde{\operatorname{Eig}}(\varphi, \lambda)\right) = n. \qquad \triangle$$

Proposition 5.5.10. *Es sei V ein endlich dimensionaler K-Vektorraum und $\varphi\colon V \to V$ ein Endomorphismus, dessen charakteristisches Polynom über K in Linearfaktoren zerfällt. Wenn $\lambda_1, \dots, \lambda_s$ die paarweise verschiedenen Eigenwerte von φ sind, dann gilt*

$$V = \widetilde{\operatorname{Eig}}(\varphi, \lambda_1) + \dots + \widetilde{\operatorname{Eig}}(\varphi, \lambda_s).$$

Jeder verallgemeinerte Eigenraum $\widetilde{\operatorname{Eig}}(\varphi, \lambda_i)$ ist φ-invariant, und $\lambda_i \operatorname{id}_V - \varphi$ darauf eingeschränkt ist nilpotent.

Beweis. Zunächst zeigen wir die φ-Invarianz der verallgemeinerten Eigenräume. Ist n_i die algebraische Vielfachheit von λ_i, so gilt für $v \in \widetilde{\operatorname{Eig}}(\varphi, \lambda_i)$

$$\begin{aligned}(\lambda_i \operatorname{id} - \varphi)^{n_i}(\varphi(v)) &= (\lambda_i \operatorname{id} - \varphi)^{n_i}(\lambda_i v + \varphi(v) - \lambda_i v)\\ &= \lambda_i (\lambda_i \operatorname{id} - \varphi)^{n_i}(v) - (\lambda_i \operatorname{id} - \varphi)^{n_i+1}(v)\\ &= 0 + 0 = 0.\end{aligned}$$

Somit gilt wie gewünscht $\varphi(v) \in \widetilde{\operatorname{Eig}}(\varphi, \lambda_i)$. Per Definition ist $\lambda_i \operatorname{id}_V - \varphi$ auf $\widetilde{\operatorname{Eig}}(\varphi, \lambda_i)$ eingeschränkt nilpotent, hat deshalb natürlich dort nur den Eigenwert 0. Auf $\widetilde{\operatorname{Eig}}(\varphi, \lambda_j)$ eingeschränkt hat es deshalb nur den Eigenwert $\lambda_i - \lambda_j$, ist also für $j \neq i$ dort invertierbar. Daraus folgt unmittelbar

$$\widetilde{\operatorname{Eig}}(\varphi, \lambda_1) \cap \left(\widetilde{\operatorname{Eig}}(\varphi, \lambda_2) + \dots + \widetilde{\operatorname{Eig}}(\varphi, \lambda_s)\right) = \{0\}$$

Das kann man iterieren und erhält mit der Dimensionsformel

$$\begin{aligned}&\dim\left(\widetilde{\operatorname{Eig}}(\varphi, \lambda_1) + \dots + \widetilde{\operatorname{Eig}}(\varphi, \lambda_s)\right)\\ &= \dim\left(\widetilde{\operatorname{Eig}}(\varphi, \lambda_1)\right) + \dots + \dim\left(\widetilde{\operatorname{Eig}}(\varphi, \lambda_s)\right)\\ &= n_1 + \dots + n_s = \deg(p_\varphi) = \dim(V).\end{aligned}$$

Daraus folgt $V = \widetilde{\operatorname{Eig}}(\varphi, \lambda_1) + \dots + \widetilde{\operatorname{Eig}}(\varphi, \lambda_s)$. $\square$

Satz 5.5.11 (Jordan'sche Normalform)**.** *(i) Sei V ein endlich dimensionaler K-Vektorraum und $\varphi\colon V \to V$ ein Endomorphismus, dessen charakteristisches Polynom über K*

in Linearfaktoren zerfällt. Dann gibt es eine Basis $\underline{v}$ von V, bezüglich der die Darstellungsmatrix von φ Blockdiagonalgestalt hat. Die einzelnen Blöcke sind Jordan-Blöcke zu Eigenwerten von φ:

$$M_{\underline{v},\underline{v}}(\varphi) = \begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_r \end{pmatrix}$$

(ii) Das charakteristische Polynom von $A \in \mathrm{Mat}_m(K)$ zerfalle über K in Linearfaktoren. Dann existiert $P \in \mathrm{GL}_m(K)$ so, dass $P^{-1}AP$ eine Blocksumme aus Jordanblöcken zu den Eigenwerten von μ_A ist.
(iii) Die Anzahl und Größe der verschiedenen Jordanblöcke ist in (i) und (ii) jeweils eindeutig bestimmt.

Beweis. Wir betrachten die Zerlegung in die φ-invarianten Haupträume aus Proposition 5.5.10. Auf $\widetilde{\mathrm{Eig}}(\varphi, \lambda_i)$ ist natürlich auch $\varphi - \lambda_i \operatorname{id}$ nilpotent und wir wählen dafür eine Basis wie in Proposition 5.5.3. Wegen

$$\varphi = \lambda_i \operatorname{id} + (\varphi - \lambda_i \operatorname{id})$$

ergibt sich zusätzlich statt 0 der Eigenwert λ_i auf der Diagonalen. Insgesamt ergibt sich damit eine Basis wie gewünscht.
Aussage (ii) ist wie üblich klar aus (i). Die Eindeutigkeit in (iii) von Anzahl und Größe ist klar, denn daraus ergeben sich ja gerade die Dimensionen von

$$\mathrm{Kern}\left((\lambda_i \operatorname{id} - \varphi)^j\right),$$

wie wir in Bemerkung 5.5.2 gesehen haben. □

Bemerkung 5.5.12. Um die Jordan'sche Normalform wirklich zu berechnen, muss man wie in Bemerkung 5.5.4 vorgehen, aber für alle $\lambda_i \operatorname{id} - \varphi$ einzeln. Man berechnet die Dimensionen von allen Räumen

$$\mathrm{Kern}\left((\lambda_i \operatorname{id} - \varphi)^j\right)$$

und daraus die Anzahl und Größen der Jordankästchen zum Eigenwert λ_i. Alle zusammen ergeben dann die Jordan'sche Normalform.

Beispiel 5.5.13. Wir bestimmen die Jordan'sche Normalform von

$$A = \begin{pmatrix} 9 & 2 & 8 & 11 \\ 3 & -1 & -1 & 2 \\ 3 & -7 & 5 & 2 \\ -6 & 5 & -7 & -7 \end{pmatrix} \in \mathrm{Mat}_4(\mathbb{C}).$$

Dazu bestimmen wir das charakteristische Polynom

$$p_A = t^4 - 6t^3 - 27t^2 + 108t + 324 = (t+3)^2(t-6)^2 \in \mathbb{C}[t].$$

Es treten also die beiden Eigenwerte $6, -3$ jeweils mit algebraischer Vielfachheit 2 auf. Die geometrische Vielfachheit von 6 ist ebenfalls 2, also erhalten wir zwei Jordanblöcke der Größe 1 zum Eigenwert 6. Die geometrische Vielfachheit von -3 ist jedoch nur 1 und es gilt

$$\dim\left(\mathrm{Kern}((-3I_4 - A)^2)\right) = 2.$$

Damit gibt es einen Jordanblock der Größe 2 zum Eigenwert -3. Als Normalform erhalten wir

$$\begin{pmatrix} -3 & 1 & 0 & 0 \\ 0 & -3 & 0 & 0 \\ 0 & 0 & 6 & 0 \\ 0 & 0 & 0 & 6 \end{pmatrix}.$$

Definition 5.5.14. Zwei Matrizen $A, B \in \mathrm{Mat}_m(K)$ heißen **ähnlich**, falls $P \in \mathrm{GL}_m(K)$ existiert mit

$$B = P^{-1}AP.$$

Das bedeutet offensichtlich gerade, dass A, B Darstellungsmatrizen desselben Endomorphismus sind, nur bezüglich verschiedener Basen (aber vorne und hinten jeweils dieselbe). △

Korollar 5.5.15. *Zwei Matrizen $A, B \in \mathrm{Mat}_m(\mathbb{C})$ sind genau dann ähnlich, wenn sie dieselbe Jordan'sche Normalform besitzen.*

Beweis. Wenn A und B dieselbe Jordan'sche Normalform besitzen, sind sie zu ihr und damit zueinander ähnlich. Wenn A und B ähnlich sind, dann beschreiben sie dieselbe lineare Abbildung und nur davon hängt die Jordan'sche Normalform ab. □

5.6 Der Satz von Cayley-Hamilton und das Minimalpolynom

Die K-Vektorräume $\mathrm{Mat}_m(K)$ und $\mathrm{Lin}_K(V,V)$ sind sogar Ringe. Dabei verwendet man in $\mathrm{Mat}_m(K)$ die bekannte Matrixmultiplikation und in $\mathrm{Lin}_K(V,V)$ entsprechend die Hintereinanderausführung von Endormorphismen. Ist also

$\psi\colon V \to V$ ein Endomorphismus, $A \in \mathrm{Mat}_m(K)$ eine Matrix und $p \in K[t]$ ein Polynom, so ist $p(\varphi)$ und $p(A)$ jeweils ebenfalls ein Endomorphismus bzw. eine Matrix.

Satz 5.6.1 (Satz von Cayley-Hamilton)**.** *(i) Sei $\varphi\colon V \to V$ ein Endomorphismus des endlich dimensionalen K-Vektorraums V. Dann gilt*

$$p_\varphi(\varphi) = 0.$$

(ii) Für $A \in \mathrm{Mat}_m(K)$ gilt $p_A(A) = 0$.

Beweis. Wir nehmen zunächst an, dass p_φ über K in Linearfaktoren zerfällt:

$$p_\varphi = (t-\lambda_1)^{n_1} \cdots (t-\lambda_r)^{n_r}.$$

Statt mit φ können wir dann offensichtlich mit einer Darstellungsmatrix in Jordan'scher Normalform arbeiten:

$$J = \begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_r \end{pmatrix}.$$

Dabei vereine $J_i \in \mathrm{Mat}_{n_i}(K)$ alle Jordanblöcke zum Eigenwert λ_i. Wenn man nun J in p_φ einsetzt, eliminiert der Faktor $(t-\lambda_i)^{n_i}$ aber gerade den Block J_i. Als Produkt ergibt sich damit insgesamt die Nullmatrix.
Falls p_φ nicht in Linearfaktoren zerfällt, kann man den Körper K stets so erweitern, dass p_φ über dem größeren Körper in Linearfaktoren zerfällt (das wird in der Algebra bewiesen). Damit stimmt die Aussage über dem größeren Körper, und damit natürlich auch über K. $\square$

Beispiel 5.6.2. (i) Sei $\theta\colon \mathbb{R}^2 \to \mathbb{R}^2$ die Drehung um den Winkel $\pi/2$ gegen den Uhrzeigersinn. Es gilt $p_\varphi = t^2 + 1$ und somit

$$\varphi \circ \varphi + \mathrm{id} = 0 \quad \text{bzw.} \quad \varphi \circ \varphi = -\,\mathrm{id}\,.$$

(ii) Für

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \in \mathrm{Mat}_3(\mathbb{Q})$$

gilt $p_A = 18t + 15t^2 - t^3$ und somit

$$18 \cdot A + 15 \cdot A^2 - A^3 = 0. \qquad \triangle$$

Korollar 5.6.3. *Für $A \in \mathrm{GL}_m(K)$ ist A^{-1} stets eine K-Linearkombination von*

$$I_m, A, A^2, \ldots, A^{m-1}.$$

Dasselbe stimmt für invertierbare Endomorphismen eines m-dimensionalen Raums.

Beweis. Da A invertierbar ist, hat das charakteristische Polynom p_A einem nichttrivialen konstanten Koeffizienten. Nach geeigneter Normierung erhalten wir also mit dem Satz von Cayley-Hamilton eine Gleichung

$$I_m + a_1 A + a_2 A^2 + \cdots + a_m A^m = 0$$

mit allen $a_i \in K$. Daraus ergibt sich

$$A^{-1} = -a_1 I_m - a_2 A - \cdots - a_m A^{m-1},$$

die gewünschte Aussage. □

Beispiel 5.6.4. Wir betrachten

$$A = \begin{pmatrix} 1 & 2 \\ -2 & -2 \end{pmatrix}$$

und berechnen $p_A = t^2 + t + 2$. Es ergibt sich

$$I_2 + \tfrac{1}{2}A + \tfrac{1}{2}A^2 = 0$$

und daraus

$$A^{-1} = -\tfrac{1}{2}I_2 - \tfrac{1}{2}A = \begin{pmatrix} -1 & -1 \\ 1 & \frac{1}{2} \end{pmatrix}. \qquad \triangle$$

Lemma 5.6.5. *Sei $\varphi\colon V \to V$ ein Endomorphismus des endlich dimensionalen K-Vektorraums V. Dann gibt es ein eindeutig bestimmtes normiertes Polynom $q \in K[t]$ von minimalem Grad mit $q(\varphi) = 0$. Dasselbe stimmt für Matrizen in $\mathrm{Mat}_m(K)$.*

Beweis. Nach dem Satz von Cayley-Hamilton gibt es ein nichttrivales Polynom mit φ als Nullstelle, also gibt es auch ein normiertes solches Polynom von minimalem Grad. Für die Eindeutigkeit seien q_1, q_2 zwei solche Polynome. Da beide denselben (minimalen) Grad und Leitkoeffizient 1 haben, hat $q := q_1 - q_2$ echt kleineren Grad und es gibt $q(\varphi) = q_1(\varphi) - q_2(\varphi) = 0 - 0 = 0$. Wegen der Minimalität des Grades folgt $q = 0$, also $q_1 = q_2$. □

Definition 5.6.6. Sei $\varphi\colon V \to V$ ein Endomorphismus des endlich dimensionalen K-Vektorraums V und $A \in \mathrm{Mat}_m(K)$. Das normierte Polynom $q \in K[t]$ von minimalem Grad mit $q(\varphi) = 0$ bzw. $q(A) = 0$ heißt **Minimalpolynom** von φ bzw. A. Wir verwenden dafür die Bezeichnung

$$\tilde{p}_\varphi \quad \text{bzw.} \quad \tilde{p}_A. \qquad \triangle$$

Beispiel 5.6.7. Das Minimalpolynom kann sich vom charakteristischen Polynom unterscheiden. Für

$$A = \left(\begin{array}{c|c} J(0,2) & \\ \hline & J(0,1) \end{array}\right) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \in \mathrm{Mat}_3(K)$$

etwa gilt $p_A = t^3$ und $\tilde{p}_A = t^2$. Der kleine Jordanblock ist bereits in der ersten Potenz eliminiert, der größere nach der zweiten. Eine normierte Gleichung vom Grad 1 erfüllt A offensichtlich nicht. Für

$$A = J(0,3) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \in \mathrm{Mat}_3(K)$$

hingegen gilt $\tilde{p}_A = p_A = t^3$. Das Minimalpolynom hängt also von der Größe des jeweils größten Jordanblocks zu jedem Eigenwert ab. $\triangle$

Satz 5.6.8. *Sei $\varphi\colon V \to V$ ein Endomorphismus des endlich dimensionalen K-Vektorraums V.*
(i) Für jedes $q \in K[t]$ mit $q(\varphi) = 0$ gibt es ein $h \in K[t]$ mit $q = \tilde{p}_\varphi \cdot h$.
(ii) Das charakteristische Polynom von φ zerfalle über K in Linearfaktoren:

$$p_\varphi = (t - \lambda_1)^{n_1} \cdots (t - \lambda_r)^{n_r}.$$

Sei s_i die Größe des größten Jordanblocks zum Eigenwert λ_i von φ, also genau die minimale Zahl mit

$$\mathrm{Kern}\left((\lambda_i \operatorname{id}_V - \varphi)^{s_i}\right) = \mathrm{Kern}\left((\lambda_i \operatorname{id}_V - \varphi)^{s_i+1}\right).$$

Dann gilt

$$\tilde{p}_\varphi = (t - \lambda_1)^{s_1} \cdots (t - \lambda_r)^{s_r}.$$

(iii) Die Aussagen (i) und (ii) stimmen analog für Matrizen $A \in \mathrm{Mat}_m(K)$.

Beweis. Für (i) gelte $q(\varphi) = 0$. Wir machen Division mit Rest/Polynomdivision und erhalten eine Gleichung

$$q = h \cdot \tilde{p}_\varphi + r$$

mit $h, r \in K[t]$ und $\deg(r) < \deg(\tilde{p}_\varphi)$. Durch Einsetzung von φ erhalten wir

$$0 = q(\varphi) = h(\varphi)\underbrace{\tilde{p}_\varphi(\varphi)}_{=0} + r(\varphi) = r(\varphi)$$

und aus der Minimalität des Grades von $\tilde{p}_\varphi$ folgt $r = 0$.
(ii) Wenn wir mit der Jordan'schen Normalform J von φ rechnen, sehen wir dass mit $q = (t - \lambda_1)^{s_1} \cdots (t - \lambda_r)^{s_r}$ offensichtlich $q(J) = 0$ gilt. Wäre $\deg(\tilde{p}_\varphi) < \deg(q)$, so wäre $\tilde{p}_\varphi$ nach (i) ein echter Teiler von q, also von der Gestalt

$$\tilde{p}_\varphi = (t - \lambda_1)^{u_1} \cdots (t - \lambda_r)^{u_r}$$

mit mindestens einem $u_i < s_i$. Wenn o.B.d.A. $u_1 < s_1$ gilt und der obere linke Block in J gerade $J(\lambda_1, s_1)$ ist, so ergibt sich als oberer linker Block in $\tilde{p}_\varphi(J)$ gerade

$$J(0, s_1)^{u_1} J(\lambda_1 - \lambda_2, s_1)^{u_2} \cdots J(\lambda_1 - \lambda_r, s_1)^{u_r}.$$

Man überlegt sich nur relativ leicht, dass dieses Produkt nicht die Nullmatrix ist, ein Widerspruch. (iii) ist wie üblich klar. $\square$

Kapitel 6

Skalarprodukte und Spektralsätze

In diesem gesamten Kapitel steht $\mathbb{K}$ für die reellen Zahlen $\mathbb{R}$ oder die komplexen Zahlen $\mathbb{C}$. Über anderen Körpern betrachtet man Skalarprodukte gewöhnlich nicht. Auch sei V wie bisher immer ein $\mathbb{K}$-Vektorraum. Ein Skalarprodukt ist eine zusätzliche Struktur auf V, die beispielsweise die Definition von Winkeln, Längen und Abständen erlaubt. Verträgt sich eine lineare Abbildung auf geeignete Weise mit einem Skalarprodukt, kann man diese Tatsache für starke Sätze über ihre Eigenwerte und Darstellungsmatrizen verwenden, sogenannte *Spektralsätze*.

6.1 Skalarprodukte und Normen

Definition 6.1.1. (i) Ein **Skalarprodukt** auf V ist eine Abbildung

$$\begin{aligned} \langle \cdot, \cdot \rangle \colon V \times V &\to \mathbb{K} \\ (v, w) &\mapsto \langle v, w \rangle \end{aligned}$$

welche für alle $v_1, v_2, w \in V$ und $\lambda \in \mathbb{K}$ folgende Bedingungen erfüllt:

(Linearität im ersten Eintrag) $\quad \langle \lambda v_1 + v_2, w \rangle = \lambda \langle v_1, w \rangle + \langle v_2, w \rangle$

(Hermitizität/Symmetrie) $\quad \langle v, w \rangle = \overline{\langle w, v \rangle}$

(Positive Definitheit) $\quad \langle v, v \rangle \in \mathbb{R}_{>0} \quad$ für $v \neq 0$.

(ii) Ein endlich-dimensionaler $\mathbb{R}$-Vektorraum mit Skalarprodukt heißt **euklidischer Raum**.
(iii) Ein endlich-dimensionaler $\mathbb{C}$-Vektorraum mit Skalarprodukt heißt **unitärer Raum**. $\triangle$

Bemerkung 6.1.2. (i) Aus der Hermitizität und der Linearität im ersten Eintrag ergibt sich direkt die konjugierte Linearität im zweiten Eintrag:

$$\begin{aligned}\langle w, \lambda v_1 + v_2\rangle &= \overline{\langle \lambda v_1 + v_2, w\rangle}\\ &= \overline{\lambda\langle v_1, w\rangle + \langle v_2, w\rangle}\\ &= \overline{\lambda}\overline{\langle v_1, w\rangle} + \overline{\langle v_2, w\rangle}\\ &= \overline{\lambda}\langle w, v_1\rangle + \langle w, v_2\rangle.\end{aligned}$$

(ii) Im Fall $\mathbb{K} = \mathbb{R}$ gilt ja $\overline{\lambda} = \lambda$ für alle $\lambda \in \mathbb{K}$. Die zweite Bedingung wird also zur Symmetrie

$$\langle v, w\rangle = \langle w, v\rangle$$

und man erhält auch die klassische Linearität im zweiten Eintrag.
(iii) Auch im Fall $\mathbb{K} = \mathbb{C}$ gilt stets $\langle v, v\rangle \in \mathbb{R}$. Das folgt beispielsweise direkt aus der Hermitizität. Andererseits folgt es für $0 \neq v$ auch direkt aus der positiven Definitheit, aus der Linearität im ersten Eintrag und der Hermitizität erhält man zusätzlich

$$\langle 0, v\rangle = \langle v, 0\rangle = 0$$

für alle $v \in V$, also insbesondere $\langle 0, 0\rangle = 0$. △

Beispiel 6.1.3. (i) Für $\mathbb{K} = \mathbb{R}$ und $V = \mathbb{R}^m$ liefert die folgende Setzung das sogenannte *Standard-Skalarprodukt*:

$$\left\langle (v_1, \dots, v_m)^t, (w_1, \dots, w_m)^t\right\rangle := v_1 w_1 + \cdots + v_m w_m.$$

Alle Eigenschaften rechnet man direkt nach. Man kann diese Definition auch leicht abändern, zum Beispiel zu

$$\left\langle (v_1, \dots, v_m)^t, (w_1, \dots, w_m)^t\right\rangle := c_1 v_1 w_1 + \cdots + c_m v_m w_m.$$

Solange alle $c_i > 0$ sind, sind alle Axiome erfüllt.
(ii) Für $\mathbb{K} = \mathbb{C}$ und $V = \mathbb{C}^m$ muss man die Definition leicht variieren, um die Hermitizität und die positive Definitheit zu erhalten:

$$\left\langle (v_1, \dots, v_m)^t, (w_1, \dots, w_m)^t\right\rangle := v_1 \overline{w}_1 + \cdots + v_m \overline{w}_m.$$

(iii) Auf dem reellen Vektorraum $C\left([-1, 1], \mathbb{R}\right)$ erhält man durch folgende Setzung ein Skalarprodukt:

$$\langle f, g\rangle := \int_{-1}^{1} f(t) g(t) \mathrm{d}t.$$

(iv) Auf dem komplexen Vektorraum $C([-1,1],\mathbb{C})$ verwenden wir stattdessen:

$$\langle f, g\rangle := \int_{-1}^{1} f(t)\overline{g(t)}\mathrm{d}t. \qquad \triangle$$

Lemma 6.1.4 (Cauchy-Schwarz Ungleichung)**.** *Sei V ein Vektorraum mit Skalarprodukt. Dann gilt für alle $v, w \in V$*

$$|\langle v, w\rangle|^2 \leq \langle v, v\rangle \cdot \langle w, w\rangle.$$

Beweis. Für $w = 0$ gilt die Ungleichung offensichtlich, da dann auf beiden Seiten 0 steht. Wir nehmen also $w \neq 0$ und somit $\langle w, w\rangle > 0$ an. Für $\lambda \in \mathbb{K}$ berechnen wir

$$\begin{aligned} 0 &\leq \langle v - \lambda w, v - \lambda w\rangle \\ &= \langle v, v\rangle - \lambda\langle w, v\rangle - \overline{\lambda}\langle v, w\rangle + \lambda\overline{\lambda}\langle w, w\rangle. \end{aligned}$$

Wir setzen nun für λ die Zahl

$$\frac{\langle v, w\rangle}{\langle w, w\rangle}$$

ein und erhalten

$$\begin{aligned} 0 &\leq \langle v, v\rangle - \frac{\langle v, w\rangle}{\langle w, w\rangle}\langle w, v\rangle - \frac{\overline{\langle v, w\rangle}}{\langle w, w\rangle}\langle v, w\rangle + \frac{|\langle v, w\rangle|^2}{\langle w, w\rangle^2}\langle w, w\rangle \\ &= \langle v, v\rangle - \frac{|\langle v, w\rangle|^2}{\langle w, w\rangle} - \frac{|\langle v, w\rangle|^2}{\langle w, w\rangle} + \frac{|\langle v, w\rangle|^2}{\langle w, w\rangle} \\ &= \langle v, v\rangle - \frac{|\langle v, w\rangle|^2}{\langle w, w\rangle}. \end{aligned}$$

Daraus folgt die gewünschte Aussage unmittelbar. □

Definition 6.1.5. (i) Sei V ein $\mathbb{R}$-Vektorraum mit Skalarprodukt. Für $v, w \in V \setminus \{0\}$ definieren wir das **Winkelmaß** zwischen v und w als

$$\measuredangle(v, w) := \arccos\left(\frac{\langle v, w\rangle}{\sqrt{\langle v, v\rangle} \cdot \sqrt{\langle w, w\rangle}}\right).$$

(ii) In einem $\mathbb{K}$-Vektorraum V mit Skalarprodukt nennen wir v und w **orthogonal**, falls

$$\langle v, w\rangle = 0$$

gilt. Im reellen Fall ist das äquivalent zu $\measuredangle(v, w) = \frac{\pi}{2}$. △

Definition 6.1.6. Sei V ein Vektorraum mit Skalarprodukt und $S \subseteq V$. Dann ist

$$S^\perp := \{v \in V \mid \langle v, s\rangle = 0 \text{ für alle } s \in S\}$$

offensichtlich ein Untervektorraum von V. Man nennt $S^\perp$ das **orthogonale Komplement von S in V**. Es gilt stets

$$S \cap S^\perp \subseteq \{0\}. \qquad \triangle$$

Bemerkung 6.1.7. Sei $A \in \mathrm{Mat}_{m,n}(\mathbb{R})$ und $v \in \mathbb{R}^n$ mit $Av = 0$. Das bedeutet gerade, dass v auf allen Zeilen von A orthogonal steht, bezüglich des Standard-Skalarprodukts. Also ist $L(A, 0)$ gerade das orthogonale Komplement des Zeilenraums von A. $\triangle$

Definition 6.1.8. (i) Eine **Norm** auf V ist eine Abbildung

$$\begin{aligned} \|\cdot\| \colon V &\to \mathbb{R} \\ v &\mapsto \|v\| \end{aligned}$$

welche für alle $v, w \in V, \lambda \in \mathbb{K}$ folgende Bedingungen erfüllt:

(Positive Definitheit) $\quad \|v\| > 0 \quad$ für $v \neq 0$

(Homogenität) $\quad \|\lambda v\| = |\lambda| \cdot \|v\|$

(Dreiecksungleichung) $\quad \|v + w\| \leq \|v\| + \|w\|.$

(ii) Ein **normierter Raum** ist ein $\mathbb{K}$-Vektoraum mit Norm. $\triangle$

Bemerkung 6.1.9. (i) Beispielsweise aus der Homogenität folgt sofort

$$\|0\| = 0$$

für jede Norm. Somit nimmt eine Norm immer nur Werte ≥ 0 an.
(ii) Anschaulich muss man sich $\|v\|$ als die *Länge* des Vektors v, also seinen Abstand zum Nullvektor vorstellen. Die Axiome passen genau zu dieser Anschauung. $\triangle$

Proposition 6.1.10. *Sei $\langle\cdot,\cdot\rangle$ ein Skalarprodukt auf V. Dann definiert die Setzung*

$$\|v\| := \sqrt{\langle v, v\rangle}$$

eine Norm auf V.

Beweis. Zunächst ist $\|v\| = \sqrt{\langle v, v\rangle} \in \mathbb{R}_{\geq 0}$ wohldefiniert, denn $\langle v, v\rangle \geq 0$ gilt für alle $v \in V$. Für $0 \neq v$ gilt ausserdem $\langle v, v\rangle > 0$ und damit $\|v\| > 0$. Die Homogenität rechnen wir direkt nach:

$$\|\lambda v\|^2 = \langle \lambda v, \lambda v\rangle = \lambda\overline{\lambda}\langle v, v\rangle = |\lambda|^2\|v\|^2.$$

Die Aussage folgt nach Ziehen der Wurzel. Für die Dreiecksungleichung berechnen wir

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

wobei wir für die letzte Ungleichung die Cauchy-Schwarz Ungleichung benutzt haben. Nach Ziehen der Wurzel erhalten wir die Dreiecksungleichung. □

Bemerkung 6.1.11. Wie schon im Beweis von Proposition 6.1.10 gesehen, kann man die Cauchy-Schwarz Ungleichung auch als

$$|\langle v, w\rangle| \leq \|v\| \cdot \|w\|$$

formulieren. Dabei muss natürlich $\|\cdot\|$ die von $\langle\cdot,\cdot\rangle$ induzierte Norm sein. △

Beispiel 6.1.12. (i) Aus dem Standardskalarprodukt auf $\mathbb{R}^m$ erhalten wir die Norm

$$\|(v_1, \ldots, v_m)^t\| = \sqrt{v_1^2 + \cdots + v_m^2}.$$

Auf $\mathbb{C}^m$ erhalten wir analog

$$\|(v_1, \ldots, v_m)^t\| = \sqrt{|v_1|^2 + \cdots + |v_m|^2}.$$

(ii) Nicht jede Norm auf einem Vektorraum kommt von einem Skalarprodukt. Beispielsweise sind

$$\|(v_1, \ldots, v_m)^t\|_\infty := \max\{|v_1|, \ldots, |v_m|\}$$

und

$$\|(v_1, \ldots, v_m)^t\|_1 := |v_1| + \cdots + |v_m|$$

Normen auf $\mathbb{R}^m$, welche nicht von einem Skalarprodukt induziert werden (Übungsaufgabe). △

Bemerkung 6.1.13. Eine Norm $\|\cdot\|$ auf V kann auch zur Definition von Abständen zwischen Vektoren benutzt werden. Man definiert

$$d(v, w) := \|v - w\|$$

und kann sämtliche Axiome einer Metrik nachrechnen. Insgesamt ergibt sich also folgende Hierarchie der Strukturen:

$$\text{Skalarprodukt} \Rightarrow \text{Norm} \Rightarrow \text{Metrik.}$$

Dabei können die Pfeile im Allgemeinen nicht umgekehrt werden. △

6.2 Orthonormalbasen

Sei in diesem Abschnitt stets V ein endlich-dimensionaler $\mathbb{K}$-Vektorraum mit festgewählten Skalarprodukt $\langle\cdot,\cdot\rangle$ und induzierter Norm $\|\cdot\|$.

Definition 6.2.1. Eine Basis $\underline{v} = (v_1, \dots, v_m)$ von V heißt **Orthonormalbasis** (ONB), falls

$$\langle v_i, v_j\rangle = 0 \text{ für alle } i \neq j$$

und

$$\|v_i\| = 1 \text{ für alle } i$$

gilt. △

Konstruktion 6.2.2 (Gram-Schmidt'sches Orthonormalisierungsverfahren)**.** Seien $w_1, \dots, w_m$ linear unabhängige Vektoren in V. Der folgende Algorithmus produziert eine ONB $(v_1, \dots, v_m)$ von $\mathrm{Span}_{\mathbb{K}}\{w_1, \dots, w_m\}$:
(i) Setze $v_1 := \frac{1}{\|w_1\|} w_1$.
(ii) Angenommen $v_1, \dots, v_k$ sind bereits definiert. Dann setzen wir

$$\tilde{v}_{k+1} := w_{k+1} - \sum_{i=1}^{k} \langle w_{k+1}, v_i\rangle v_i$$

und

$$v_{k+1} := \frac{1}{\|\tilde{v}_{k+1}\|} \tilde{v}_{k+1}.$$

Beweis. Für jedes i gilt

$$\|v_i\| = \|\frac{1}{\|\tilde{v}_i\|}\tilde{v}_i\| = |\frac{1}{\|\tilde{v}_i\|}|\|\tilde{v}_i\| = 1.$$

Für $j \leq k$ gilt nun

$$\begin{aligned}\langle \tilde{v}_{k+1}, v_j\rangle &= \left\langle w_{k+1} - \sum_{i=1}^{k}\langle w_{k+1}, v_i\rangle v_i, v_j\right\rangle \\ &= \langle w_{k+1}, v_j\rangle - \sum_{i=1}^{k}\langle w_{k+1}, v_i\rangle\langle v_i, v_j\rangle.\end{aligned}$$

Wenn wir nun induktiv voraussetzen, dass $v_1, \dots, v_k$ bereits orthonormal sind, erhalten wir

$$\langle \tilde{v}_{k+1}, v_j\rangle = \langle w_{k+1}, v_j\rangle - \langle w_{k+1}, v_j\rangle = 0$$

und damit $\langle v_{k+1}, v_j\rangle = 0$ für alle $j \leq k$. Damit sind auch $v_1, \dots, v_{k+1}$ orthonormal.
An der Formel aus Schritt (ii) sieht man ebenfalls induktiv sofort, dass $v_1, \dots, v_k$ denselben Raum aufspannen wie $w_1, \dots, w_k$, für alle $k = 1, \dots, m$. □

Korollar 6.2.3. *V besitzt eine ONB.*

Beweis. Wir wählen mit Korollar 3.1.17 eine Basis von V und wenden darauf das Gram-Schmidt'sche Orthonormalisierungsverfahren an. Dadurch erhalten wir eine ONB von V. □

Beispiel 6.2.4. (i) Wir betrachten die linear unabhängigen Vektoren

$$w_1 = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}, w_2 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, w_3 = \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix}$$

im $\mathbb{R}^3$. Keiner der Vektoren hat Norm 1 und keine zwei der Vektoren sind orthogonal. Durch das Gram-Schmidt-Verfahren erhalten wir zunächst

$$v_1 = \frac{1}{\|w_1\|}w_1 = \frac{1}{\sqrt{5}}\begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}.$$

Im nächsten Schritt ergibt sich

$$\tilde{v}_2 = w_2 - \langle w_2, v_1 \rangle v_1 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} - \frac{1}{\sqrt{5}} \frac{1}{\sqrt{5}} \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} = \frac{1}{5} \begin{pmatrix} -6 \\ 0 \\ 3 \end{pmatrix}$$

und daraus

$$v_2 = \frac{\sqrt{5}}{3} \tilde{v}_2 = \frac{1}{\sqrt{5}} \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix}.$$

Im letzten Schritt erhalten wir

$$\tilde{v}_3 = w_3 - \langle w_3, v_1 \rangle v_1 - \langle w_3, v_2 \rangle v_2 = \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix}$$

und daraus

$$v_3 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}.$$

Wir haben dadurch die ONB

$$\left(\frac{1}{\sqrt{5}} \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}, \frac{1}{\sqrt{5}} \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right)$$

von $\mathbb{R}^3$ konstruiert.

(*ii*) In $C([-\pi, \pi], \mathbb{R})$ betrachten wir $f_1 \colon t \mapsto t^2$, $f_2 \colon t \mapsto \cos(t)$ und den Unterraum

$$V = \operatorname{Span}_{\mathbb{R}}\{f_1, f_2\}.$$

Wir versehen V mit dem Skalarprodukt aus Beispiel 6.1.3 (*iii*). Es gilt

$$\|f_1\|^2 = \int_{-\pi}^{\pi} t^4 \mathrm{d}t = \frac{2}{5} \pi^5$$

und wir erhalten

$$v_1 = \sqrt{\frac{5}{2\pi^5}} \cdot f_1$$

als erstes Element unserer ONB. Im zweiten Schritt berechnen wir

$$\begin{aligned}\tilde{v}_2 &= f_2 - \langle f_2, v_1\rangle v_1\\ &= f_2 - \frac{5}{2\pi^5}\int_{-\pi}^{\pi} t^2\cos(t)\mathrm{d}t \cdot f_1\\ &= f_2 + \frac{10}{\pi^4} f_1.\end{aligned}$$

Wenn man schließlich noch

$$\|\tilde{v}_2\|^2 = \pi - \frac{40}{\pi^3} = \frac{\pi^4 - 40}{\pi^3}$$

berechnet, erhält man

$$v_2 = \sqrt{\frac{\pi^3}{\pi^4 - 40}} f_2 + \frac{10}{\pi^2\sqrt{\pi(\pi^4 - 40)}} f_1. \qquad \triangle$$

Proposition 6.2.5. *Sei $U \subseteq V$ ein Untervektorraum. Dann gilt*

$$U + U^\perp = V$$

und jedes Element $v \in V$ hat eine eindeutige Darstellung

$$v = u_1 + u_2$$

mit $u_1 \in U, u_2 \in U^\perp$.

Beweis. Wir wählen eine ONB $v_1, \dots, v_n$ von U und definieren für $v \in V$

$$u := \sum_{i=1}^{n} \langle v, v_i\rangle v_i.$$

Nach Konstruktion liegt u in U und es gilt $v = u + (v - u)$. Wir zeigen nun dass $v - u \in U^\perp$ liegt. Dazu genügt es $\langle v - u, v_j\rangle = 0$ für alle $j = 1, \dots, n$ zu zeigen. Es gilt nun

$$\langle v - u, v_j\rangle = \langle v, v_j\rangle - \sum_i \langle v, v_i\rangle\langle v_i, v_j\rangle = \langle v, v_j\rangle - \langle v, v_j\rangle = 0,$$

wobei wir benutzt haben, dass $v_1, \dots, v_n$ eine ONB ist. Somit ist $U + U^\perp = V$ gezeigt.

Die Eindeutigkeit der Summendarstellung folgt direkt aus $U \cap U^\perp = \{0\}$: aus $v = u_1 + u_2 = u_1' + u_2'$ mit $u_1, u_1' \in U, u_2, u_2' \in U^\perp$ folgt

$$U \ni u_1 - u_1' = u_2' - u_2 \in U^\perp,$$

also $0 = u_1 - u_1' = u_2' - u_2$ und daraus $u_1 = u_1', u_2 = u_2'$. □

Definition 6.2.6. Sei $U \subseteq V$ ein Untervektorraum mit ONB $v_1, \dots, v_n$. Die lineare Abbildung

$$\begin{aligned} \pi_U \colon V &\to U \\ v &\mapsto \sum_{i=1}^n \langle v, v_i \rangle v_i \end{aligned}$$

heißt **orthogonale Projektion auf** U. Weil

$$v = \pi_U(v) + (v - \pi_U(v))$$

die eindeutige Darstellung von v in $U + U^\perp$ ist, hängt π_U nicht von der expliziten Wahl der ONB ab. △

Korollar 6.2.7. *Es ist $\pi_U(v)$ das (eindeutig bestimmte) Element von U mit kleinstem Abstand zu v:*

$$\|v - \pi_U(v)\| \le \|v - u\| \quad \textit{für alle } u \in U.$$

Beweis. Übungsaufgabe. □

Bemerkung 6.2.8. Mit dem Begriff der orthogonalen Projektion kann man auch das Verfahren von Gram-Schmidt noch anschaulicher verstehen. Die Gleichung

$$\tilde{v}_{k+1} := w_{k+1} - \sum_{i=1}^k \langle w_{k+1}, v_i \rangle v_i$$

besagt gerade, dass man w_{k+1} orthogonal auf den Span der bereits konstruierten Vektoren projiziert und das Ergebnis von w_{k+1} abzieht. Der Vektor $\tilde{v}_{k+1}$ steht also senkrecht auf den bisherigen Vektoren und muss nur noch normiert werden. △

Beispiel 6.2.9 (Regression)**.** Es seien $a = (a_1, \dots, a_m)^t, b = (b_1, \dots, b_m)^t \in \mathbb{R}^m$ gegeben (wobei etwa b_i das Ergebnis einer Messung zum Zeitpunkt a_i beschreibt). Wir haben nun einen Untervektorraum $U \subseteq \mathcal{F}(\mathbb{R}, \mathbb{R})$ gegeben und suchen eine Funktion $g \in U$, welche die Daten möglichst genau beschreibt. Es soll also $g(a_i)$ möglichst nahe an b_i liegen, für alle i. Wir betrachten dazu folgende lineare Abbildung

$$\begin{aligned} \varphi\colon U &\to \mathbb{R}^m \\ f &\mapsto (f(a_1), \ldots, f(a_m)) . \end{aligned}$$

Idealerweise finden wir ein Urbild von b unter φ, denn aus $\varphi(g) = b$ folgt $g(a_i) = b_i$ für alle i.
Es kann aber natürlich auch passieren, dass b nicht im Bild $\varphi(U)$ liegt. Korollar 6.2.7 legt dann nahe, b orthogonal auf $\mathrm{Bild}(\varphi)$ zu projizieren und dann ein Urbild von $\pi_{\varphi(U)}(b)$ zu wählen. Aus

$$\varphi(g) = \pi_{\varphi(U)}(b)$$

folgt dann, dass

$$\|\varphi(g) - b\|^2 = \sum_{i=1}^{m} (g(a_i) - b_i)^2$$

kleinstmöglich ist, unter allen Funktionen aus U (wir benutzen das Standard-Skalarprodukt auf $\mathbb{R}^m$). Die Summe der Fehlerquadrate wird hier also minimiert. Im Fall dass

$$U = \{t \mapsto r + st \mid r, s \in \mathbb{R}\}$$

der Unterraum aller affin linearen Abbildungen ist, nennt man das beschriebene Vorgehen auch **lineare Regression**. Da U von den beiden Funktionen $t \mapsto 1$ und $t \mapsto t$ aufgespannt wird, gilt

$$\varphi(U) = \mathrm{Span}_{\mathbb{R}} \{e, a\}$$

mit $e = (1, \ldots, 1)^t$. Als ONB von $\varphi(U)$ erhält man mit dem Gram-Schmidt Verfahren daraus

$$v_1 = \frac{1}{\sqrt{m}} e, \;\; v_2 = \frac{1}{\|c\|} c$$

mit

$$c = a - \frac{\langle a, e \rangle}{m} e.$$

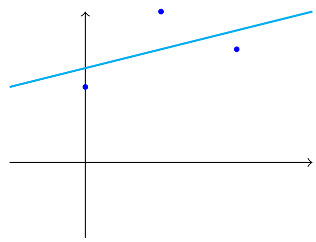
Damit ergibt sich

$$\begin{aligned} \pi_{\varphi(U)}(b) &= \frac{\langle b, e \rangle}{m} e + \frac{\langle b, c \rangle}{\|c\|^2} c \\ &= \frac{1}{m} \left(\langle b, e \rangle - \frac{\langle b, c \rangle \langle a, e \rangle}{\|c\|^2} \right) e + \frac{\langle b, c \rangle}{\|c\|^2} a. \end{aligned}$$

Als Urbild unter φ erhalten wir die Funktion

$$g\colon t \mapsto \frac{1}{m}\left(\langle b, e\rangle - \frac{\langle b, c\rangle\langle a, e\rangle}{\|c\|^2}\right) + \frac{\langle b, c\rangle}{\|c\|^2} t.$$

Im Fall $a = (0, 1, 2), b = (1, 2, \frac{3}{2})$ ergibt sich so beispielsweise die Funktion $g\colon t \mapsto \frac{5}{4} + \frac{1}{4}t$.



$\triangle$

6.3 Spektralsätze

In diesem Abschnitt seien V, W, X stets endlich-dimensionale $\mathbb{K}$-Vektorräume mit Skalarprodukt.

Proposition 6.3.1. *(i) Die folgende Abbildung ist bijektiv:*

$$\begin{aligned} \iota\colon V &\to V' \\ v &\mapsto \langle \cdot, v\rangle. \end{aligned}$$

Im Fall $\mathbb{K} = \mathbb{R}$ ist ι linear, also ein Isomorphismus. Im Fall $\mathbb{K} = \mathbb{C}$ ist ι additiv und konjugiert linear.
(ii) Ist $\underline{v} = (v_1, \ldots, v_m)$ eine ONB von V, so ist $\iota(\underline{v}) := (\iota(v_1), \ldots, \iota(v_m))$ genau die zu $\underline{v}$ duale Basis.

Beweis. (i): Für festes $v \in V$ ist die Abbildung

$$\begin{aligned} \langle \cdot, v\rangle\colon V &\to \mathbb{K} \\ x &\mapsto \langle x, v\rangle \end{aligned}$$

linear, also ein Element von V'. Die Aussagen über die Linearität von ι ist klar. Für die Injektivität nehmen wir $\iota(v_1) = \iota(v_2)$ an, also $\langle x, v_1\rangle = \langle x, v_2\rangle$ für alle $x \in V$. Daraus folgt

$$0 = \langle x, v_1\rangle - \langle x, v_2\rangle = \langle x, v_1 - v_2\rangle$$

und für $x := v_1 - v_2$ daraus $0 = \|v_1 - v_2\|^2$, also $v_1 = v_2$.
Für die Surjektivität wählen wir $0 \neq f \in V'$ (der Fall $f = 0$ ist klar). Dann gilt $\mathrm{Kern}(f) \subsetneq V$ und also existiert ein $0 \neq w \in \mathrm{Kern}(f)^\perp$. Wir setzen

$$v := \frac{\overline{f(w)}}{\|w\|^2} w \in \mathrm{Kern}(f)^\perp.$$

Dann gilt

$$\|v\|^2 = \frac{|f(w)|^2}{\|w\|^4}\|w\|^2 = \frac{|f(w)|^2}{\|w\|^2} = f(v).$$

Für jedes $x \in V$ gilt

$$x - \frac{f(x)}{f(v)} v \in \mathrm{Kern}(f),$$

wie man direkt sieht. Daraus folgt nun

$$0 = \langle x - \frac{f(x)}{f(v)} v, v\rangle = \langle x, v\rangle - \frac{f(x)}{f(v)}\|v\|^2 = \langle x, v\rangle - f(x).$$

Das zeigt $f = \iota(v)$.
(ii) ist offensichtlich wegen

$$\iota(v_i)(v_j) = \langle v_j, v_i\rangle = \begin{cases} 1: & i = j \\ 0: & i \neq j. \end{cases} \qquad \square$$

Definition 6.3.2. Für $A = (a_{ij})_{i,j} \in \mathrm{Mat}_{m,n}(\mathbb{C})$ bezeichnen wir mit

$$A^* := \overline{A}^t := (\overline{a_{ji}})_{i=1,\dots,n;j=1,\dots,m} \in \mathrm{Mat}_{n,m}(\mathbb{C})$$

die zu A **adjungierte Matrix**. Für reelle Matrizen stimmt A^* offensichtlich mit A^t überein. $\triangle$

Satz 6.3.3. *(i) Zu jeder linearen Abbildung $\varphi\colon V \to W$ existiert genau eine lineare Abbildung $\varphi^*\colon W \to V$ mit*

$$\langle \varphi(v), w\rangle = \langle v, \varphi^*(w)\rangle$$

für alle $v \in V, w \in W$ (man beachte, dass das Skalarprodukt links das in W, rechts dasjenige in V ist).
(ii) Sind $\underline{v}, \underline{w}$ ONBs von V bzw. W, so gilt

$$M_{\underline{w},\underline{v}}(\varphi^*) = M_{\underline{v},\underline{w}}(\varphi)^*.$$

Beweis. (i): Wir definieren
$$\varphi^* = \iota^{-1} \circ \varphi' \circ \iota,$$
wobei $\varphi' \colon W' \to V'$ die zu φ duale Abbildung ist (vergleiche Abschnitt 3.2.5):

$$\begin{tikzcd} V' \arrow[d, "\iota^{-1}"'] & W' \arrow[l, "\varphi'"'] \\ V & W \arrow[l, "\varphi^*"] \arrow[u, "\iota"'] \end{tikzcd}$$

Da in der Hintereinanderausführung im komplexen Fall zwei konjugiert lineare Abbildungen auftreten, ist φ^* linear. Offensichtlich erfüllt φ^* die Bedingung:

$$\langle \varphi(v), w\rangle = \iota(w)(\varphi(v)) = \varphi'(\iota(w))(v) = \langle v, \iota^{-1}(\varphi'(\iota(w)))\rangle = \langle v, \varphi^*(w)\rangle.$$

Umgekehrt besagt die geforderte Gleichung ja gerade $\varphi^* = \iota^{-1}\circ\varphi'\circ\iota$. Das beweist die Eindeutigkeit.
(ii) ist klar mit Satz 3.2.39, Korollar 6.3.1 (ii) und der Beobachtung dass ι^{-1} konjugiert linear ist. □

Definition 6.3.4. Die Abbildung φ^* aus Satz 6.3.3 heißt die zu φ **adjungierte Abbildung**. △

Beispiel 6.3.5. Wir betrachten $V = W = \mathbb{R}^2$ mit dem Standard-Skalarprodukt.
(i) Sei $\varphi \colon \mathbb{R}^2 \to \mathbb{R}^2$ die Spiegelung an der x-Achse. Sie hat bezüglich der Standardbasis die Darstellungsmatrix

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

Es gilt $S^* = S^t = S$, also ist $\varphi^* = \varphi$.
(ii) Die Drehung $d_\theta \colon \mathbb{R}^2 \to \mathbb{R}^2$ um den Winkel θ (gegen den Uhrzeigersinn) hat bezüglich der Standardbasis die Darstellungsmatrix

$$D_\theta := \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \in \mathrm{Mat}_2(\mathbb{R})$$

Es gilt
$$D_\theta^* = D_\theta^t = D_{-\theta} = D_\theta^{-1}.$$

Es ist d_θ^* also die Drehung um θ im Uhrzeigersinn, also gerade die Umkehrabbildung zu d_θ.

(*iii*) Sei

$$\begin{aligned} \varphi\colon \mathbb{R}^2 &\to \mathbb{R}^2 \\ (a,b) &\mapsto (b,0) \end{aligned}$$

die Verkettung einer Projektion und einer Spiegelung. Die Darstellungsmatrix von φ bezüglich der Standardbasis ist

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

und wir erhalten

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

Also gilt

$$\begin{aligned} \varphi^*\colon \mathbb{R}^2 &\to \mathbb{R}^2 \\ (a,b) &\mapsto (0,a). \end{aligned}$$

△

Lemma 6.3.6. *Für $\varphi\colon V \to W$ und $\psi\colon W \to X$ gilt:*

(*i*) $(\psi \circ \varphi)^* = \varphi^* \circ \psi^*$ *und* $(\varphi^*)^* = \varphi$.

(*ii*) $\mathrm{Kern}(\varphi^*) = \mathrm{Bild}(\varphi)^\perp$.

(*iii*) $\mathrm{Kern}(\varphi) = \mathrm{Bild}(\varphi^*)^\perp$.

Beweis. (*i*) rechnet man direkt anhand der Definition nach:

$$\langle(\psi\circ\varphi)(v), x\rangle = \langle\psi(\varphi(v)), x\rangle = \langle\varphi(v), \psi^*(x)\rangle = \langle v, \varphi^*(\psi^*(x))\rangle = \langle v, (\varphi^*\circ\psi^*)(x)\rangle$$

und

$$\langle\varphi^*(w), v\rangle = \overline{\langle v, \varphi^*(w)\rangle} = \overline{\langle\varphi(v), w\rangle} = \langle w, \varphi(v)\rangle.$$

Für (*ii*) sehen wir

$$\begin{aligned} w \in \mathrm{Kern}(\varphi^*) &\Leftrightarrow \varphi^*(w) = 0 \\ &\Leftrightarrow \langle v, \varphi^*(w)\rangle = 0 \quad \forall v \in V \\ &\Leftrightarrow \langle\varphi(v), w\rangle = 0 \quad \forall v \in V \\ &\Leftrightarrow w \in \mathrm{Bild}(\varphi)^\perp. \end{aligned}$$

(*iii*) folgt schließlich direkt aus (*i*) und (*ii*). □

Definition 6.3.7. (*i*) Sei $\varphi\colon V \to V$ ein Endomorphismus.

(1) φ heißt **selbstadjungiert** (bzw. **symmetrisch** im Fall $\mathbb{K} = \mathbb{R}$), falls $\varphi = \varphi^*$ gilt.

(2) φ heißt **unitär** (bzw. **orthogonal** im Fall $\mathbb{K} = \mathbb{R}$), falls φ bijektiv ist und $\varphi^{-1} = \varphi^*$ gilt.

(3) φ heißt **normal**, falls $\varphi \circ \varphi^* = \varphi^* \circ \varphi$ gilt.

(*ii*) Sei $A \in \mathrm{Mat}_m(\mathbb{K})$ eine Matrix.

(1) A heißt **Hermitesch** (bzw. **symmetrisch** im Fall $\mathbb{K} = \mathbb{R}$), falls $A = A^*$ gilt.

(2) A heißt **unitär** (bzw. **orthogonal** im Fall $\mathbb{K} = \mathbb{R}$), falls A invertierbar ist und $A^{-1} = A^*$ gilt.

(3) A heißt **normal**, falls $AA^* = A^*A$ gilt.

Offensichtlich ist jede Hermitesche/symmetrische und jede unitäre/orthogonale Abbildung/Matrix normal. △

Bemerkung 6.3.8. Sei $\varphi\colon V \to V$ ein Endomorphismus und $\underline{v}$ eine ONB von V. Dann ist φ nach Satz 6.3.3 (*ii*) genau dann selbstadjungiert/symmetrisch/ unitär/orthogonal/normal, wenn $M_{\underline{v},\underline{v}}(\varphi)$ jeweils die entsprechende Eigenschaft hat. △

Bemerkung/Beispiel 6.3.9. (*i*) In Beispiel 6.3.5 ist die Abbildung/Matrix aus (*i*) symmetrisch, die aus (*ii*) orthogonal und die aus (*iii*) nicht einmal normal. In (*iii*) gilt

$$AA^* = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \text{ und } A^*A = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

(*ii*) Eine Matrix $A \in \mathrm{Mat}_m(\mathbb{K})$ ist unitär/orthogonal, genau dann wenn die Spalten eine ONB von $\mathbb{K}^m$ bilden, bezüglich des Standard-Skalarprodukts. Das ist eine direkte Umformulierung der Bedingung $A^*A = I_m$.
(*iii*) Unitäre/orthogonale Abbildungen erhalten das Skalarprodukt:

$$\langle \varphi(v), \varphi(w) \rangle = \langle v, \varphi^*(\varphi(w)) \rangle = \langle v, w \rangle.$$

Erst recht erhalten sie auch die Norm:

$$\|\varphi(v)\|^2 = \langle \varphi(v), \varphi(v) \rangle = \langle v, v \rangle = \|v\|^2.$$ △

Proposition 6.3.10. *Sei $\varphi\colon V \to V$ ein normaler Endormorphismus mit Eigenwert $\lambda \in \mathbb{K}$. Dann gilt*

$$\mathrm{Eig}(\varphi, \lambda) = \mathrm{Eig}(\varphi^*, \overline{\lambda})$$

und insbesondere $\mathrm{Spec}(\varphi^*) = \overline{\mathrm{Spec}(\varphi)} := \left\{\overline{\lambda} \mid \lambda \in \mathrm{Spec}(\varphi)\right\}$.

Beweis. Für $v \in V$ gilt

$$\begin{aligned}\langle \varphi(v) - \lambda v, \varphi(v) - \lambda v\rangle &= \langle \varphi(v), \varphi(v)\rangle - \overline{\lambda}\langle \varphi(v), v\rangle - \lambda\langle v, \varphi(v)\rangle + |\lambda|^2\langle v, v\rangle \\ &= \langle v, \varphi^*(\varphi(v))\rangle - \overline{\lambda}\langle v, \varphi^*(v)\rangle - \lambda\langle \varphi^*(v), v\rangle + |\lambda|^2\langle v, v\rangle \\ &= \langle v, \varphi(\varphi^*(v))\rangle - \overline{\lambda}\langle v, \varphi^*(v)\rangle - \lambda\langle \varphi^*(v), v\rangle + |\lambda|^2\langle v, v\rangle \\ &= \langle \varphi^*(v), \varphi^*(v)\rangle - \overline{\lambda}\langle v, \varphi^*(v)\rangle - \lambda\langle \varphi^*(v), v\rangle + |\overline{\lambda}|^2\langle v, v\rangle \\ &= \langle \varphi^*(v) - \overline{\lambda} v, \varphi^*(v) - \overline{\lambda} v\rangle.\end{aligned}$$

Man beachte, dass wir für die dritte Gleichheit die Normalität von φ benutzt haben. Insbesondere gilt nun

$$v \in \mathrm{Eig}(\varphi, \lambda) \Leftrightarrow \|\varphi(v) - \lambda v\|^2 = 0 \Leftrightarrow \|\varphi^*(v) - \overline{\lambda} v\|^2 = 0 \Leftrightarrow v \in \mathrm{Eig}(\varphi^*, \overline{\lambda}),$$

die gewünschte Aussage. □

Satz 6.3.11 (Spektralsatz für normale Endomorphismen/Matrizen)**.**
(i) Sei $\varphi\colon V \to V$ normal und p_φ zerfalle über $\mathbb{K}$ in Linearfaktoren. Dann existiert eine ONB von V die aus Eigenvektoren von φ besteht. Insbesondere ist φ diagonalisierbar.
(ii) Sei $A \in \mathrm{Mat}_m(\mathbb{K})$ normal und p_A zerfalle über $\mathbb{K}$ in Linearfaktoren. Dann existiert eine unitäre/orthogonale Matrix $P \in \mathrm{GL}_m(\mathbb{K})$ für die $P^{-1}AP$ Diagonalgestalt hat.

Beweis. Wir beweisen (*i*) mit Induktion nach m. Der Induktionsanfang $m = 1$ ist klar. Im Fall von allgemeinem m wählen wir einen Eigenwert λ von φ und dazugehörigen Eigenvektor $v_1 \in V$. Diese existieren, da p_φ über $\mathbb{K}$ in Linearfaktoren zerfällt.
Wir betrachten nun $U := \{v_1\}^\perp$ und zeigen, dass $\varphi(U) \subseteq U$ gilt, also U ein sogenannten *invarianter Unterraum* von φ ist. Für $u \in U$ gilt

$$\langle \varphi(u), v_1\rangle = \langle u, \varphi^*(v_1)\rangle = \langle u, \overline{\lambda} v_1\rangle = \lambda\langle u, v_1\rangle = 0,$$

also $\varphi(u) \in U$. Für die zweite Gleichheit haben wir Proposition 6.3.10, also die Normalität von φ benutzt. Genauso sieht man auch $\varphi^*(U) \subseteq U$.
Wir betrachten nun $\psi := \varphi_{|U}\colon U \to U$. Es gilt $\psi^* = \varphi^*_{|U}$:

$$\langle \psi(u_1), u_2\rangle = \langle \varphi(u_1), u_2\rangle = \langle u_1, \varphi^*(u_2)\rangle.$$

Also ist auch ψ normal. Wie im Beweis von Satz 5.4.2 sieht man, dass p_ψ gerade der Kofaktor von $t-\lambda$ in p_φ ist, also selbst wieder über $\mathbb{K}$ zerfällt. Wegen $\dim(U) = m-1$ können wir nun die Induktionsvoraussetzung auf ψ anwenden, es existiert also eine ONB von U aus Eigenvektoren von ψ. Zusammen mit $\frac{1}{\|v_1\|}v_1$ erhalten wir damit eine ONB von V aus Eigenvektoren von φ.
(ii) ist nun mit Beispiel 6.3.9 (ii) klar. □

Satz 6.3.12 (Spektralsatz für selbstadjungierte/symmetrische Endormophismen/ Matrizen)**.**
(i) Sei $\mathbb{K} = \mathbb{C}$ und $\varphi\colon V \to V$ selbstadjungiert. Dann sind alle Eigenwerte von φ reell und es existiert eine ONB von V aus Eigenvektoren von φ. Insbesondere ist φ diagonalisierbar, mit einer reellen diagonalen Darstellungsmatrix.
(ii) Sei $A \in \mathrm{Mat}_m(\mathbb{C})$ Hermitesch. Dann existiert eine unitäre Matrix $P \in \mathrm{GL}_m(\mathbb{C})$ so dass $P^{-1}AP$ eine reelle Diagonalmatrix ist.
(iii) Sei $\mathbb{K} = \mathbb{R}$ und $\varphi\colon V \to V$ symmetrisch. Dann zerfällt p_φ über $\mathbb{R}$ in Linearfaktoren, es existiert eine ONB von V aus Eigenvektoren von φ und φ ist somit diagonalisierbar.
(iv) Sei $A \in \mathrm{Mat}_m(\mathbb{R})$ symmetrisch. Dann existiert eine orthogonale Matrix $P \in \mathrm{GL}_m(\mathbb{R})$ so dass $P^{-1}AP$ eine (reelle) Diagonalmatrix ist.

Beweis. Selbstadjungierte/symmetrische Abbildungen/Matrizen sind normal, wir können also Satz 6.3.11 anwenden.
Für (i) ist also nur noch zu zeigen, dass alle Eigenwerte von φ reell sind. Das folgt aber direkt aus Proposition 6.3.10. (ii) ist damit ebenfalls klar.
Für (iii) müssen wir nur zeigen, dass p_φ über $\mathbb{R}$ zerfällt. Für jede beliebige ONB $\underline{v}$ von V ist $M_{\underline{v},\underline{v}}(\varphi) \in \mathrm{Mat}_m(\mathbb{R})$ eine symmetrische Matrix. Als Element von $\mathrm{Mat}_m(\mathbb{C})$ ist sie deshalb selbstadjungiert und die Aussage folgt aus (ii). (iv) folgt direkt aus (iii). □

Für reelle orthogonale Abbildungen/Matrizen können wir offensichtlich nicht die Diagonalisierbarkeit beweisen, wie man an Drehungen sieht. Das ist aber die einzige Einschränkung, wie wir nun sehen.

Satz 6.3.13 (Spektralsatz für unitäre/orthogonale Endormophismen/Matrizen)**.**
(i) Sei $\mathbb{K} = \mathbb{C}$ und $\varphi\colon V \to V$ unitär. Dann haben alle Eigenwerte von φ Betrag 1 und es existiert eine ONB von V aus Eigenvektoren von φ. Insbesondere ist φ diagonalisierbar, mit einer diagonalen Darstellungsmatrix deren Diagonaleinträge Betrag 1 haben.
(ii) Sei $A \in \mathrm{Mat}_m(\mathbb{C})$ unitär. Dann existiert eine unitäre Matrix $P \in \mathrm{GL}_m(\mathbb{C})$ so dass $P^{-1}AP$ eine Diagonalmatrix ist, deren Diaognaleinträge Betrag 1 haben.

(iii) Sei $\mathbb{K} = \mathbb{R}$ und $\varphi\colon V \to V$ orthogonal. Dann existieren paarweise orthogonale und φ-invariante Unterräume $U_1, \ldots, U_d$ von V mit

$$V = U_1 + \cdots + U_d$$

und $\dim(U_i) = 1$ oder 2 für alle i. Im Fall $\dim(U_i) = 1$ ist $\varphi_{|U_i} = \pm \mathrm{id}_{U_i}$, im Fall $\dim(U_i) = 2$ ist $\varphi_{|U_i}$ eine Drehung.
(iv) Sei $A \in \mathrm{Mat}_m(\mathbb{R})$ orthogonal. Dann existiert eine orthogonale Matrix $P \in \mathrm{GL}_m(\mathbb{R})$ so dass $P^{-1}AP$ von der folgenden Gestalt ist:

$$\begin{pmatrix} 1 &&&&&&&& \\ & \ddots &&&&&&& \\ && 1 &&&&&& \\ &&& -1 &&&&& \\ &&&& \ddots &&&& \\ &&&&& -1 &&& \\ &&&&&& D_{\theta_1} && \\ &&&&&&& \ddots & \\ &&&&&&&& D_{\theta_r} \end{pmatrix}.$$

Dabei sind die $D_{\theta_i} \in \mathrm{Mat}_2(\mathbb{R})$ Drehmatrizen.

Beweis. Unitäre/orthogonale Abbildungen/Matrizen sind normal, also können wir Satz 6.3.11 anwenden.
Für (i) ist damit nur noch zu zeigen, dass die Eigenwerte von φ Betrag 1 haben. Das folgt aber aus Proposition 6.3.10, denn es gilt

$$\mathrm{Eig}\,(\varphi, \lambda) = \mathrm{Eig}\left(\varphi^*, \overline{\lambda}\right) = \mathrm{Eig}\left(\varphi^{-1}, \overline{\lambda}\right) = \mathrm{Eig}\left(\varphi, \overline{\lambda}^{-1}\right),$$

somit $\lambda = \overline{\lambda}^{-1}$ und deshalb $1 = \lambda\overline{\lambda} = |\lambda|^2$. Auch (ii) ist damit klar.
Für (iii) wählen wir eine ONB $\underline{v}$ von V und betrachten die orthogonale Darstellungsmatrix

$$A = M_{\underline{v},\underline{v}}(\varphi) \in \mathrm{Mat}_m(\mathbb{R}).$$

Als Element von $\mathrm{Mat}_m(\mathbb{C})$ ist A dann unitär und somit ist die lineare Abbildung

$$\mu_A\colon \mathbb{C}^m \to \mathbb{C}^m$$

ebenfalls unitär (bezüglich des Standard-Skalarprodukts auf $\mathbb{C}^m$). Mit (i) haben also alle Eigenwerte von μ_A Betrag 1 und wir erhalten eine ONB von $\mathbb{C}^m$ aus Eigenvektoren von μ_A. Weil p_A ein reelles Polynom ist, treten die Eigenwerte in konjugierten Paaren auf:

$$0 = p_A(\lambda) \;\Rightarrow\; 0 = \overline{0} = \overline{p_A(\lambda)} = p_A\left(\overline{\lambda}\right).$$

Dabei gilt

$$v \in \operatorname{Eig}(\mu_A, \lambda) \;\Rightarrow\; \overline{v} \in \operatorname{Eig}(\mu_A, \overline{\lambda}),$$

denn $\mu_A(\overline{v}) = A\overline{v} = \overline{Av} = \overline{\lambda v} = \overline{\lambda}\overline{v}$. Für $v \in \mathbb{C}^m$ gilt weiter

$$\operatorname{Span}_{\mathbb{C}}\{v, \overline{v}\} = \operatorname{Span}_{\mathbb{C}}\{\underbrace{v + \overline{v}, i(v - \overline{v})}_{\in \mathbb{R}^m}\}.$$

Deshalb kann man für die Eigenräume zu den reellen Eigenwerten (höchstens ± 1) eine Basis (und damit eine ONB) aus $\mathbb{R}^m$ wählen. Das übersetzt sich direkt in eindimensionale Unterräume U_i von V. Für einen echt komplexen Eigenwert λ mit Eigenvektor $v \in \mathbb{C}^m$ ist $\overline{v}$ ein Eigenvektor zu $\overline{\lambda} \neq \lambda$ und deshalb sind $v, \overline{v}$ orthogonal und insbesondere linear unabhängig. Es gilt also

$$2 = \dim_{\mathbb{C}} \operatorname{Span}_{\mathbb{C}}\{v, \overline{v}\} = \dim_{\mathbb{C}} \operatorname{Span}_{\mathbb{C}}\{v + \overline{v}, i(v - \overline{v})\}.$$

Also sind die beiden reellen Vektoren $v + \overline{v}, i(v - \overline{v})$ ebenfalls linear unabhängig und $\operatorname{Span}_{\mathbb{R}}\{v+\overline{v}, i(v-\overline{v})\}$ ist ein 2-dimensionaler invarianter $\mathbb{R}$-Unterraum von μ_A. Eine orthogonale Abbildung eines zweidimensionalen reellen Raums ohne reelle Eigenwerte muss aber eine Drehung sein, wie man sich leicht überlegt (zum Beispiel mit Bemerkung 6.3.9 (ii)). Damit ist (iii) bewiesen und (iv) folgt direkt daraus. $\square$

Bemerkung 6.3.14. Die Spektralsätze besagen anschaulich, dass es nur die offensichtlichen selbstadjungierten / symmetrischen... Endomorphismen gibt. Wenn man eine geeignete ONB wählt, skaliert eine solche Abbildung einfach in jede Richtung mit einem entsprechenden Faktor/Eigenwert. Nur im reellen orthogonalen Fall können noch zweidimensionale Drehungen auftreten. $\triangle$

Bemerkung 6.3.15 (Funktionalkalkül)**.** Manipuliert man die Diagonalelemente einer Diagonalmatrizen $D \in \operatorname{Mat}_m(\mathbb{C})$ auf gewisse Weise, so erhält man eine neue Diagonalmatrix, die sich oft wie erwartet verhält. Ist etwa $D = \operatorname{diag}(\lambda_1, \dots, \lambda_m)$ und setzen wir

$$\sqrt{D} := \operatorname{diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_m})$$

so gilt $\sqrt{D}\sqrt{D} = D$. Ist eine Matrix nicht diagonal, so geht das nicht mehr so leicht eintragsweise, da ja Matrixmultiplikation nicht eintragsweise funktioniert. Ist aber $P^{-1}AP = D$ eine Diagonalmatrix, so können wir

$$\sqrt{A} := P\sqrt{D}P^{-1}$$

setzen und erhalten

$$\sqrt{A}\sqrt{A} = P\sqrt{D}P^{-1}P\sqrt{D}P^{-1} = P\sqrt{D}\sqrt{D}P^{-1} = PDP^{-1} = A.$$

Man kann also aus jeder normalen komplexen Matrix eine Wurzel ziehen.
Allgemeiner erhält man mit jeder auf dem Spektrum von A definierten Funktion f die Matrix

$$f(A) := Pf(D)P^{-1}$$

mit den erwarteten Eigenschaften. $\triangle$

6.4 Positiv semidefinite Matrizen

Wir bezeichnen mit

$$\mathrm{Her}_m(\mathbb{C}) := \{A \in \mathrm{Mat}_m(\mathbb{C}) \mid A^* = A\}$$

die Menge aller Hermiteschen komplexen Matrizen und mit

$$\mathrm{Sym}_m(\mathbb{R}) := \left\{A \in \mathrm{Mat}_m(\mathbb{R}) \mid A^t = A\right\} = \mathrm{Her}_m(\mathbb{C}) \cap \mathrm{Mat}_m(\mathbb{R})$$

die Menge aller symmetrischen reellen Matrizen. Bei beiden Mengen handelt es sich (nur) um $\mathbb{R}$-Vektorräume.

Definition 6.4.1. Eine Matrix $A \in \mathrm{Her}_m(\mathbb{C})$ heißt **positiv semidefinit**, wenn alle ihre Eigenwerte ≥ 0 sind. Sie heißt **positiv definit**, wenn alle ihre Eigenwerte > 0 sind. Wir verwenden dafür die Notation

$$A \geqslant 0 \quad \text{bzw.} \quad A > 0. \qquad \triangle$$

Bemerkung 6.4.2. (i) Hermitesche Matrizen haben nur reelle Eigenwerte, nach Satz 6.3.12. Deren Nichtnegativität ist die zusätzliche Forderung in der letzten Definition.
(ii) Wegen $\mathrm{Sym}_m(\mathbb{R}) \subseteq \mathrm{Her}_m(\mathbb{C})$ gilt die letzte Definition natürlich auch für reelle symmetrische Matrizen. $\triangle$

Definition 6.4.3. Sei $A \in \mathrm{Mat}_m(K)$. Ein **Hauptminor** von A ist die Determinante einer quadratischen Untermatrix von A, die durch Streichung derselben Zeilen wie Spalten aus A entsteht.
Ein **führender Hauptminor** ist die Determinante der quadratischen Untermatrix, die durch Streichung der letzten k Zeilen und Spalten aus A entsteht, für ein $k = 0, \ldots, m-1$. △

Der folgende Satz fasst die wichtigsten alternativen Formulierungen für positive Semidefinitheit zusammen:

Satz 6.4.4. *Für $A \in \mathrm{Her}_m(\mathbb{C})$ sind äquivalent:*

(i) $A \geqslant 0$.

(ii) $A = B^*B$ *für ein* $B \in \mathrm{Mat}_m(\mathbb{C})$.

(iii) $v^*Av \geq 0$ *für alle* $v \in \mathbb{C}^m$.

(iv) Alle Hauptminoren von A sind nichtnegativ.

Ist $A \in \mathrm{Sym}_m(\mathbb{R})$, so kann in (ii) und (iii) jeweils $\mathbb{R}$ statt $\mathbb{C}$ gewählt werden.

Beweis. Für den Beweis wählen wir mit Satz 6.3.12 eine unitäre Matrix $P \in \mathrm{GL}_m(\mathbb{C})$ mit

$$P^{-1}AP = P^*AP = \mathrm{Diag}(\lambda_1, \ldots, \lambda_m) =: D,$$

wobei die $\lambda_i \in \mathbb{R}$ die Eigenwerte von A sind.
Für (i)⇒(ii) benutzen wir, dass man für $\lambda_i \geq 0$ in $\mathbb{R}$ eine Wurzel $\sqrt{\lambda_i} \geq 0$ finden kann. Wir setzen

$$\sqrt{D} := \mathrm{Diag}\left(\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_m}\right),$$

berechnen

$$A = PDP^{-1} = P\sqrt{D}\sqrt{D}P^* = \left(\sqrt{D}P^*\right)^* \sqrt{D}P^*$$

und haben (ii) bewiesen. Für (ii)⇒(iii) rechnen wir

$$v^*Av = v^*B^*Bv = (Bv)^* Bv = \|Bv\|^2 \geq 0,$$

wobei wir die Standardnorm auf $\mathbb{C}^m$ benutzt haben. Für (iii)⇒(iv) beobachten wir, dass (iii) natürlich auch für D gilt, denn

$$v^*Dv = (Pv)^*APv \geq 0.$$

Mit $v = e_i$ sieht man so, dass $\lambda_i \geq 0$ für alle i gelten muss. Damit gilt

$$\det(A) = \det(PDP^{-1}) = \det(D) = \lambda_1 \cdots \lambda_m \geq 0.$$

Mit A erfüllt auch jede Untermatrix, die durch Streichung derselben Zeilen wie Spalten entsteht, die Bedingung (iii). Damit sind alle Hauptminoren von A nichtnegativ.
Aussage $(iv)\Rightarrow(i)$ beweisen wir durch Induktion über m. Der Fall $m = 1$ ist klar. Im allgemeinen Fall zeigen wir die Aussage nun durch Kontraposition, nehmen also an, dass A nicht positiv semidefinit ist. Wenn genau eines der λ_i negativ und alle anderen positiv sind, gilt offensichtlich $\det(A) < 0$. Ansonsten wählen wir zwei orthonormale Eigenvektoren $v, w \in \mathbb{C}^m$ von μ_A zu Eigenwerten $\lambda_i < 0, \lambda_j \leq 0$. Dazu finden wir $r \in \mathbb{C}$, so dass

$$x := v + rw \in \mathbb{C}^m$$

mindestens einen Nulleintrag hat. Wenn wir in x nun diesen Eintrag und in A die entsprechende Zeile und Spalte streichen, erhalten wir eine Untermatrix $\tilde{A}$ und $\tilde{x} \in \mathbb{C}^{m-1}$ mit

$$\begin{aligned}
\tilde{x}^* \tilde{A} \tilde{x} &= x^* A x \\
&= (v + rw)^* A (v + rw) \\
&= v^* A v + \overline{r} w^* A v + r v^* A w + |r|^2 w^* A w \\
&= \lambda_i + |r|^2 \lambda_j < 0.
\end{aligned}$$

Also erfüllt $\tilde{A}$ die Bedingung (iii) nicht. Da $(i)\Rightarrow(iii)$ bereits bewiesen ist, erfüllt $\tilde{A}$ die Bedingung (i) nicht. Nach Induktionsvoraussetzung existiert also ein negativer Hauptminor von $\tilde{A}$ und somit natürlich auch von A. Die Aussage über symmetrische Matrizen ist klar aus dem Beweis. □

Beispiel 6.4.5. Die Matrix

$$A = \begin{pmatrix} 1 & 2 & -1 \\ 2 & 4 & -2 \\ -1 & -2 & 1 \end{pmatrix}$$

ist positiv semidefinit. Man kann zum Beispiel zuerst

$$p_A = t^3 - 6t^2 = t^2(t - 6)$$

und daraus die Eigenwerte $0, 0, 6$ berechnen. Eine ONB des $\mathbb{R}^3$ aus Eigenvektoren von μ_A ist zum Beispiel

$$v_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \; v_2 = \frac{1}{\sqrt{3}} \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}, v_3 = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}.$$

Mit $P = (v_1, v_2, v_3)$ erhält man dann

$$P^t AP = P^{-1} AP = \mathrm{Diag}(0, 0, 6) =: D$$

und für

$$B = \sqrt{D} P^t = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 2 & -1 \end{pmatrix}$$

gilt somit $B^t B = A$. Man kann von A aber auch alle 7 Hauptminoren ausrechnen, sie sind alle nichtnegativ. △

Auch die positive Definitheit kann man so ähnlich charakterisieren:

Satz 6.4.6. *Für $A \in \mathrm{Her}_m(\mathbb{C})$ sind äquivalent:*

(i) $A > 0$.

(ii) $A \geqslant 0$ *und A invertierbar.*

(iii) $A = B^* B$ *für ein* $B \in \mathrm{GL}_m(\mathbb{C})$.

(iv) $v^* Av > 0$ *für alle* $v \in \mathbb{C}^m \setminus \{0\}$.

(v) Alle führenden Hauptminoren von A sind strikt positiv.

(vi) Es ist $\langle v, w \rangle := w^* Av$ *ein Skalarprodukt auf* $\mathbb{C}^m$

Ist $A \in \mathrm{Sym}_m(\mathbb{R})$, so kann in (iii), (iv) und (vi) jeweils $\mathbb{R}$ statt $\mathbb{C}$ gewählt werden.

Beweis. Die Äquivalenz von (i)-(iv) ist klar, bzw. wird genau wie in Satz 6.4.4 bewiesen. Es sind weiter (iv) und (vi) offensichtlich äquivalent. Zu (v) kommt man direkt von (i) und der Beobachtung, dass sich (iv) und damit (i) auf Untermatrizen überträgt. Wir setzen schließlich (v) voraus und zeigen (iii) per Induktion über m. Der Induktionsanfang $m = 1$ ist dabei wieder klar. Im allgemeinen Fall schreiben wir

$$A = \left(\begin{array}{c|c} \tilde{A} & v \\ \hline v^* & a \end{array}\right)$$

mit $\tilde{A} \in \mathrm{Her}_{m-1}(\mathbb{C})$, $v \in \mathbb{C}^{m-1}$ und $a \in \mathbb{R}$. Dann sind natürlich auch die führenden Hauptminoren von $\tilde{A}$ strikt positiv und $\tilde{A}$ erfüllt nach Induktionsvoraussetzung (iii) und somit alle Bedingungen. Insbesondere ist $\tilde{A}$ invertierbar und die Spalten von $\tilde{A}$ spannen also $\mathbb{C}^{m-1}$ auf. Wir eliminieren nun mit Spaltentransformationen v und mit den dazu konjugierten Zeilentransformationen v^*. Es gibt also $P \in \mathrm{GL}_m(\mathbb{C})$ mit

$$P^*AP = \left(\begin{array}{c|c} \tilde{A} & 0 \\ \hline 0 & b \end{array}\right).$$

Es folgt

$$\det(\tilde{A}) \cdot b = \det(P^*AP) = |\det(P)|^2 \det(A) > 0$$

und daraus $b > 0$. Wir finden ein $\tilde{B} \in \mathrm{GL}_{m-1}(\mathbb{C})$ mit

$$\tilde{A} = \tilde{B}^*\tilde{B}$$

und erhalten mit

$$B = \left(\begin{array}{c|c} \tilde{B} & 0 \\ \hline 0 & \sqrt{b} \end{array}\right) \in \mathrm{GL}_m(\mathbb{C})$$

also $B^*B = P^*AP$, also $A = (BP^{-1})^* BP^{-1}$. Also erfüllt A die Eigeschaft (iii).

□

Bemerkung 6.4.7. Für positive Definitheit reicht es aus, die führenden Hauptminoren zu betrachten. Für positive Semidefinitheit stimmt das nicht, wie man an der Matrix

$$\begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix}$$

bereits sieht. △

Beispiel 6.4.8. Die Matrix

$$A = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}$$

ist positiv definit, denn die führenden Hauptminoren sind 1 und $\det(A) = 1$. Also erhalten wir durch folgende Setzung ein Skalarprodukt auf $\mathbb{R}^2$:

$$\langle (a,b)^t, (c,d)^t \rangle = (c,d)A(a,b)^t = ac - bc - ad + 2bd. \qquad \triangle$$

Kapitel 7

Bilinearformen

In diesem Kapitel seien V, W, X stets endlich dimensionale Vektorräume über dem Körper K. Statt linearer Abbildungen betrachten wir nun sogenannte Bilinearformen. Skalarprodukte im reellen Fall sind Beispiele hierfür, aber wir entwickeln die Theorie jetzt deutlich allgemeiner.

7.1 Grundbegriffe

Definition 7.1.1. (i) Eine Abbildung

$$\beta\colon V \times W \to X$$

heißt **(K)-bilinear**, falls für jedes $v \in V, w \in W$ die Abbildungen

$$\begin{aligned} \beta(v, \cdot)\colon W &\to X \\ \beta(\cdot, w)\colon V &\to X \end{aligned}$$

K-lineare Abbildungen sind.
(ii) Im Fall $X = K$ nennt man eine bilineare Abbildung auch eine **(K)-Bilinearform**.

$\triangle$

Beispiel 7.1.2. (i) Für jede Matrix $A \in \mathrm{Mat}_{m,n}(K)$ ist durch folgende Vorschrift eine Bilinearform definiert:

$$\begin{aligned} \beta_A\colon K^m \times K^n &\to K \\ (v, w) &\mapsto v^t A w. \end{aligned}$$

Für $A = (a_{ij})_{i,j}$ kann man das so ausschreiben:

$$\beta_A(v, w) = \sum_{i=1}^{m} \sum_{j=1}^{n} v_i w_j a_{ij}.$$

(ii) Ist V ein euklidischer Raum, so ist das Skalarprodukt

$$\langle \cdot, \cdot \rangle \colon V \times V \to \mathbb{R}$$

eine Bilinearform.
(iii) Sind $f \in V'$, $g \in W'$ Linearformen, so ist

$$\begin{aligned} f \cdot g \colon V \times W &\to K \\ (v, w) &\mapsto f(v)g(w) \end{aligned}$$

eine Bilinearform.
(iv) Die Matrixmultiplikation

$$\begin{aligned} \mathrm{Mat}_{m,r}(K) \times \mathrm{Mat}_{r,n}(K) &\to \mathrm{Mat}_{m,n}(K) \\ (A, B) &\mapsto AB \end{aligned}$$

ist eine bilineare Abbildung.
(v) Es ist

$$\begin{aligned} \det \colon K^2 \times K^2 &\to K \\ (v, w) &\mapsto \det(v, w) \end{aligned}$$

eine Bilinearform.
(vi) Die Menge

$$\mathrm{Bil}_K(V \times W, X)$$

aller bilinearen Abbildungen von $V \times W$ nach X bildet einen K-Vektorraum, bezüglich den punktweise definierten Verknüpfungen. $\triangle$

Für Bilinearformen gilt eine ähnliche Aussage wie Korollar 3.2.10 für lineare Abbildungen:

Satz 7.1.3. *Sei $\beta \colon K^m \times K^n \to K$ eine Bilinearform. Dann existiert genau ein $A \in \mathrm{Mat}_{m,n}(K)$ mit*

$$\beta = \beta_A.$$

Die Zuordnung

$$\begin{aligned} b\colon \operatorname{Mat}_{m,n}(K) &\to \operatorname{Bil}_K(K^m \times K^n, K) \\ A &\mapsto \beta_A \end{aligned}$$

ist ein Isomorphismus von K-Vektorräumen.

Beweis. Für $i = 1, \dots, m$ und $j = 1, \dots, n$ setzen wir

$$a_{ij} := \beta(e_i, e_j)$$

und damit $A = (a_{ij})_{i,j}$. Für $v \in K^m, w \in K^n$ gilt dann

$$\begin{aligned} \beta(v, w) &= \beta\left(\sum_{i=1}^m v_i e_i, \sum_{j=1}^n w_j e_j\right) \\ &= \sum_{i=1}^m \sum_{j=1}^n v_i w_j \beta(e_i, e_j) \\ &= \sum_{i=1}^m \sum_{j=1}^n v_i w_j a_{ij} = \beta_A(v, w). \end{aligned}$$

Für die zweite Gleichheit haben wir dabei die Bilinearität von β benutzt. Aus der Gleichung

$$a_{ij} = \beta_A(e_i, e_j) = \beta(e_i, e_j)$$

folgt direkt, dass A durch β eindeutig bestimmt ist. Somit ist b bijektiv und die Linearität ist klar. □

Konstruktion 7.1.4. Sei $\beta\colon V \times W \to K$ eine Bilinearform und $\underline{v}, \underline{w}$ Basen von V bzw. W. Dann ist

$$\begin{aligned} \varphi_{\underline{v}} \times \varphi_{\underline{w}}\colon K^m \times K^n &\to V \times W \\ (c, d) &\mapsto (\varphi_{\underline{v}}(c), \varphi_{\underline{w}}(d)) \end{aligned}$$

ein Isomorphismus und somit $\beta \circ (\varphi_{\underline{v}} \times \varphi_{\underline{w}})$ eine Bilinearform auf $K^m \times K^n$. Nach Satz 7.1.3 gibt es eine eindeutig bestimmte Matrix $A \in \operatorname{Mat}_{m,n}(K)$ mit

$$\beta \circ (\varphi_{\underline{v}} \times \varphi_{\underline{w}}) = \beta_A :$$

$$\begin{tikzcd} V \times W \arrow[r, "\beta"] & K \\ K^m \times K^n \arrow[u, "\varphi_{\underline{v}} \times \varphi_{\underline{w}}"] \arrow[ur, "\beta_A"'] & \end{tikzcd}$$

Diese Matrix heißt **Darstellungsmatrix** von β bezüglich $\underline{v}$ und $\underline{w}$. Wir verwenden dafür die Notation

$${}_{\underline{v}}M_{\underline{w}}(\beta).$$

Die Konstruktion von ${}_{\underline{v}}M_{\underline{w}}(\beta)$ ist leicht, der (i, j)-Eintrag ist gerade

$$\beta \circ (\varphi_{\underline{v}} \times \varphi_{\underline{w}})(e_i, e_j) = \beta(v_i, w_j).$$

Man erhält für $v \in V, w \in W$ gerade

$$\beta(v, w) = \xi_{\underline{v}}(v)^t \,{}_{\underline{v}}M_{\underline{w}}(\beta)\xi_{\underline{w}}(w). \qquad \triangle$$

Beispiel 7.1.5. (i) Sind $\underline{e}, \underline{e}'$ jeweils die Standardbasen von K^m und K^n, so gilt für jedes $M \in \mathrm{Mat}_{m,n}(K)$

$${}_{\underline{e}}M_{\underline{e}'}(\beta_A) = A.$$

(ii) Für eine Matrix $A = (a_{ij})_{i,j} \in \mathrm{Mat}_m(K)$ definieren wir die **Spur** als

$$\mathrm{tr}(A) := \sum_{i=1}^{m} a_{ii} \in K.$$

Für die Bilinearform

$$\begin{aligned} \beta\colon \mathrm{Mat}_2(K) \times \mathrm{Mat}_2(K) &\to K \\ (A, B) &\mapsto \mathrm{tr}(AB) \end{aligned}$$

und die Basis

$$\underline{v} = \left(\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right)$$

erhalten wir

$${}_{\underline{v}}M_{\underline{v}}(\beta) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \qquad \triangle$$

Lemma 7.1.6. *Für $A, B \in \mathrm{Mat}_{m,n}(K)$ sind äquivalent:*

(i) *A und B sind Darstellungsmatrizen derselben Bilinearform, nur bezüglich eventuell unterschiedlicher Basen.*

(ii) *Es gibt $P \in \mathrm{GL}_m(K)$ und $Q \in \mathrm{GL}_n(K)$ mit $P^t AQ = B$.*

(iii) *A und B sind äquivalent (siehe Definition 3.2.33).*

Beweis. (i) ist genau die Existenz eines folgenden kommutativen Diagramms:

$$\begin{tikzcd}
K^m \times K^n \arrow[dr, "\beta_A"] & \\
V \times W \arrow[u, "\xi_{\underline{v}} \times \xi_{\underline{w}}"] \arrow[r, "\beta"] & K \\
K^m \times K^n \arrow[u, "\varphi_{\underline{v}'} \times \varphi_{\underline{w}'}"] \arrow[ur, "\beta_B"'] \arrow[uu, bend left=70, "\mu_P \times \mu_Q"]
\end{tikzcd}$$

Für $c \in K^m, d \in K^n$ gilt also

$$c^t Bd = \beta_B(c,d) = \beta_A \circ (\mu_P \times \mu_Q)(c,d) = \beta_A(Pc, Qd) = (Pc)^t A(Qd) = c^t P^t AQd.$$

Daraus folgt direkt $B = P^t AQ$, also (ii).
Da mit P^t natürlich auch P invertierbar ist, folgt die Äquivalenz von (ii) und (iii).
Unter der Annahme von (ii) betrachten wir folgendes kommutatives Diagramm:

$$\begin{tikzcd}
K^m \times K^n \arrow[r, "\beta_A"] & K \\
K^m \times K^n \arrow[u, "\mu_P \times \mu_Q"] \arrow[ur, "\beta_B"']
\end{tikzcd}$$

Es ist A immer eine Darstellungsmatrix von β_A. Am Diagramm sieht man aber, dass auch B eine Darstellungsmatrix von β_A ist. Damit ist (i) gezeigt. $\square$

7.2 Symmetrische Bilinearformen

In diesem Abschnitt betrachten wir stets nur einen (endlich dimensionalen) K-Vektorraum V und Bilinearformen auf V:

$$\beta \colon V \times V \to K.$$

Nicht völlig überraschend wählen wir jetzt auch in beiden Kopien von V *dieselbe* Basis $\underline{v}$ und betrachten die Darstellungsmatrix

$$_{\underline{v}}M_{\underline{v}}(\beta) \in \mathrm{Mat}_m(K).$$

Zwei verschiedene Darstellungsmatrizen $A, B \in \mathrm{Mat}_m(K)$ von β erfüllen dabei gerade

$$P^t AP = B$$

für ein $P \in \mathrm{GL}_m(K)$, wie man am Beweis von Lemma 7.1.6 direkt sieht.

Definition 7.2.1. $A, B \in \mathrm{Mat}_m(K)$ heißen **kongruent**, wenn $P \in \mathrm{GL}_m(K)$ existiert mit

$$P^t AP = B.$$

Das bedeutet gerade, dass A und B Darstellungsmatrizen derselben Binlinearform auf einem Raum sind, nur bezüglich eventuell unterschiedlicher Basen (aber in beiden Kopien jeweils dieselbe). △

Bemerkung 7.2.2. Hier eine Übersicht über die bisher eingeführten Äquivalenzbegriffe für Matrizen und ihre jeweilige Bedeutung:

	$A, B \in \mathrm{Mat}_{m,n}(K)$	$A, B \in \mathrm{Mat}_m(K)$
lineare Abbildungen	$P^{-1}AQ = B$ *äquivalent*	$P^{-1}AP = B$ *ähnlich*
Bilinearformen	$P^t AQ = B$ *äquivalent*	$P^t AP = B$ *kongruent*

Äquivalenz von zwei Matrizen kann man nach Satz 3.2.32 einfach durch Vergleich des Rangs entscheiden. Ähnlichkeit kann man zumindest über $\mathbb{C}$ nach Korollar 5.5.15 durch Vergleich der Jordan'schen Normalform entscheiden. Mit Kongruenz beschäftigen wir uns im folgenden. △

Definition 7.2.3. Eine Bilinearform $\beta\colon V \times V \to K$ heißt **symmetrisch**, falls

$$\beta(v, w) = \beta(w, v)$$

für alle $v, w \in V$ gilt. △

Beispiel 7.2.4. (i) Die Bilinearform

$$\begin{aligned} \beta\colon \mathrm{Mat}_m(k) \times \mathrm{Mat}_m(K) &\to K \\ (A, B) &\mapsto \mathrm{tr}(AB) \end{aligned}$$

ist symmetrisch (Übungsaufgabe).
(ii) Für

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

ist die Bilinearform

$$\begin{aligned} \beta_A \colon \mathbb{Q}^2 \times \mathbb{Q}^2 &\to \mathbb{Q} \\ \big((a,b)^t, (c,d)^t\big) &\mapsto ad \end{aligned}$$

ist nicht symmetrisch. △

Lemma 7.2.5. *Die Bilinearform $\beta\colon V \times V \to K$ ist genau dann symmetrisch, wenn eine/jede Darstellungsmatrix ${}_{\underline{v}}M_{\underline{v}}(\beta)$ eine symmetrische Matrix ist.*

Beweis. Ist ${}_{\underline{v}}M_{\underline{v}}(\beta) = A = A^t$ eine symmetrische Darstellungsmatrix von β, so gilt für $v, w \in V$

$$\begin{aligned} \beta(v,w) &= \xi_{\underline{v}}(v)^t A \xi_{\underline{v}}(w) \\ &= \big(\xi_{\underline{v}}(v)^t A \xi_{\underline{v}}(w)\big)^t \\ &= \xi_{\underline{v}}(w)^t A^t \xi_{\underline{v}}(v) \\ &= \xi_{\underline{v}}(w)^t A \xi_{\underline{v}}(v) \\ &= \beta(w,v). \end{aligned}$$

Ist umgekehrt β symmetrisch, so gilt für jede Basis $\underline{v} = (v_1, \ldots, v_m)$ von V

$$\beta(v_i, v_j) = \beta(v_j, v_i),$$

also stimmen (i,j)- und (j,i)-Eintrag in der Darstellungsmatrix ${}_{\underline{v}}M_{\underline{v}}(\beta)$ überein. □

Satz 7.2.6 (Hauptachsentransformation)**.** *Sei K ein Körper mit $1 + 1 \neq 0$.*
(i) Zu jeder symmetrischen Bilinearform $\beta\colon V \times V \to K$ existiert eine Basis $\underline{v}$ von V, bezüglich der die Darstellungsmatrix ${}_{\underline{v}}M_{\underline{v}}(\beta)$ Diagonalgestalt hat.
(ii) Zu jeder symmetrischen Matrix $A = A^t \in \mathrm{Mat}_m(K)$ existiert ein $P \in \mathrm{GL}_m(K)$ so dass P^tAP Diagonalgestalt hat.

Beweis. Wir beweisen (i) mit Induktion über $m = \dim(V)$. Der Fall $m = 1$ ist klar. Im allgemeinen Fall müssen wir eine Basis $\underline{v}$ konstruieren, so dass für $i \neq j$ immer $\beta(v_i, v_j) = 0$ gilt.

Wir benutzen die sogenannte **Polarisationsformel**

$$2\beta(v,w) = \beta(v+w, v+w) - \beta(v,v) - \beta(w,w),$$

die man mit Bilinearität und Symmetrie sofort nachrechnet. Falls $\beta(v,v) = 0$ für alle $v \in V$ gilt, gilt damit auch $\beta(v,w) = 0$ für alle $v, w \in V$ und die Aussage ist klar (jede beliebige Basis führt zur Nullmatrix). Man beachte dass wir hier $1+1 \neq 0$ benutzt haben!
Es sei also $v_1 \in V$ mit $\beta(v_1, v_1) \neq 0$ gewählt. Dann ist

$$U := \{v \in V \mid \beta(v_1, v) = 0\}$$

ein echter Untervektorraum von V. Nach Induktionsvoraussetzung existiert also eine Basis $\underline{u}$ von U mit $\beta(u_i, u_j) = 0$ für alle $i \neq j$. Nach Definition von U gilt ausserdem $\beta(v_1, u_i) = 0$ für alle i. Wir müssen also nur noch zeigen, dass $\underline{v} = (v_1, \underline{u})$ eine Basis von V ist. Für $v \in V$ gilt aber

$$v = \frac{\beta(v_1, v)}{\beta(v_1, v_1)} v_1 + \left(v - \frac{\beta(v_1, v)}{\beta(v_1, v_1)} v_1 \right)$$

und der zweite Summand liegt offensichtlich in U. Somit ist $\underline{v}$ ein Erzeugendensystem von V. Aus $\lambda v_1 + u = 0$ für ein $u \in U$ folgt weiter

$$0 = \beta(v_1, \lambda v_1 + u) = \lambda \beta(v_1, v_1) + \beta(v_1, u) = \lambda \beta(v_1, v_1)$$

und daraus $\lambda = 0$. Das zeigt die lineare Unabhängigkeit von $\underline{v}$. Aussage (ii) ist wie üblich klar aus (i). $\square$

Bemerkung 7.2.7. (i) Der letzte Beweis ist konstruktiv, wenn man ihn iterativ anwendet. Für die symmetrische Bilinearform

$$\beta(A, B) = \mathrm{tr}(AB)$$

aus Beispiel 7.1.5 (ii) erhalten wir so zum Beispiel die Basis

$$v_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \; v_2 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \; v_3 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \; v_4 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

und damit als Darstellungsmatrix

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

(*ii*) Der Übergang von A zu P^tAP kann auch folgendermaßen aufgefasst werden. Wir wenden auf A elementare Zeilentransformationen an (das ist in der Multiplikation mit P^t von links kodiert), führen jede Operation aber simultan auch an den Spalten durch (das ist in der Multiplikation mit P von rechts kodiert). Auf diese Weise lässt sich A also auf Diagonalgestalt bringen. Das wollen wir auch **symmetrischen Gauß-Algorithmus** nennen (siehe Konstruktion 7.2.11 unten). Starten wir also mit der Matrix

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

aus Beispiel 7.1.5 (*ii*), können wir so die Matrix aus (*i*) erhalten.
(*iii*) Eine symmetrische *reelle* Matrix $A = A^t \in \mathrm{Mat}_m(\mathbb{R})$ ist *simultan* als Endomorphismus und als Bilinearform diagonalisierbar! Satz 6.3.12 (*iv*) liefert ja eine orthogonale Transformationsmatrix P, d.h. es gilt

$$P^{-1}AP = P^tAP = D.$$

(*iv*) Im Gegensatz zur Diagonalform eines Endomorphismus, wo in der Diagonalen stets die Eigenwerte stehen, ist eine diagonale Darstellungsmatrix einer Bilinearform nicht eindeutig bestimmt! Ersetzt man beispielsweise die Basisvektoren $v_1, \dots, v_m$ durch $\lambda_1 v_1, \dots, \lambda_m v_m$ für gewisse $\lambda_i \in K \setminus \{0\}$, so erhält man

$$\beta(\lambda_i v_i, \lambda_i v_i) = \lambda_i^2 \beta(v_i, v_i).$$

Die Diagonaleinträge einer diagonalen Darstellungsmatrix können also jederzeit mit Quadraten skaliert werden. Über $\mathbb{R}$ kann die Diagonalmatrix aus (*i*) also weiter vereinfacht werden zu

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

und über $\mathbb{C}$ sogar zu

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Dabei verwenden wir dass in $\mathbb{R}$ eine Wurzel aus $\frac{1}{2}$ und in $\mathbb{C}$ eine Wurzel aus jeder Zahl existiert. $\triangle$

Satz 7.2.8 (Sylvester'scher Trägheitssatz). *Sei $\mathbb{K} = \mathbb{R}$ oder $\mathbb{C}$ und V ein endlich dimensionaler $\mathbb{K}$-Vektorraum mit symmetrischer Bilinearform $\beta\colon V \times V \to \mathbb{K}$.*
(i) Ist $\mathbb{K} = \mathbb{C}$, so gibt es eine Basis $\underline{v}$ von V mit

$$_{\underline{v}}M_{\underline{v}}(\beta) = \mathrm{Diag}(1, \ldots, 1, 0, \ldots, 0).$$

Die Anzahl der auftretenden Einsen (und Nullen) ist durch β eindeutig bestimmt.
(ii) Ist $\mathbb{K} = \mathbb{R}$, so gibt es eine Basis $\underline{v}$ von V mit

$$_{\underline{v}}M_{\underline{v}}(\beta) = \mathrm{Diag}(1, \ldots, 1, -1, \ldots, -1, 0, \ldots, 0).$$

Die Anzahl der Einträge $1, -1$ und 0 ist dabei durch β jeweils eindeutig bestimmt.
(iii) Jedes $A = A^t \in \mathrm{Mat}_m(\mathbb{C})$ ist kongruent zu einer Matrix

$$\mathrm{Diag}(1, \ldots, 1, 0, \ldots, 0).$$

Die Anzahl der Einsen entspricht dem Rang von A. Insbesondere sind zwei symmetrische Matrizen $A, B \in \mathrm{Mat}_m(\mathbb{C})$ genau dann kongruent, wenn sie denselben Rang haben, also äquivalent sind.
(iv) Jedes $A = A^t \in \mathrm{Mat}_m(\mathbb{R})$ ist kongruent zu einer Matrix

$$\mathrm{Diag}(1, \ldots, 1, -1, \ldots, -1, 0, \ldots, 0).$$

Die Anzahl der Einsen entspricht dabei der Anzahl von strikt positiven Eigenwerten von A, die Anzahl von -1 der Anzahl von strikt negativen Eigenwerten (gerechnet inklusive Vielfachheiten). Insbesondere sind zwei symmetrische Matrizen $A, B \in \mathrm{Mat}_m(\mathbb{R})$ genau dann kongruent, wenn sie jeweils dieselbe Anzahl an positiven und negativen Eigenwerten haben.

Beweis. Die Existenz in (i) ist schon klar mit Satz 7.2.6 und Bemerkung 7.2.7 (iv). Offensichtlich ist die Anzahl der Einsen dabei der Rang von $_{\underline{v}}M_{\underline{v}}(\beta)$, welcher durch weitere Kongruenz-Transformationen offensichtlich nicht verändert werden kann. Damit ist auch (iii) bereits bewiesen.
Die Existenz in (ii) ist ebenfalls klar mit Satz 7.2.6 und Bemerkung 7.2.7 (iv). Für die Eindeutigkeit seien $(v_1, \ldots, v_m)$ und $(w_1, \ldots, w_m)$ zwei verschiedene Basen, die je zu einer Darstellungsmatrix in der erwünschten Form führen. Wir nehmen dabei o.B.d.A. folgende Sortierung an:

$$\begin{array}{llll} \beta(v_i, v_i) = 1 & i = 1, \ldots, r; & \beta(w_i, w_i) = 1 & i = 1, \ldots, r' \\ \beta(v_i, v_i) \leq 0 & i = r+1, \ldots, m; & \beta(w_i, w_i) \leq 0 & i = r'+1, \ldots, m. \end{array}$$

Dann sind
$$v_1, \ldots, v_r, w_{r'+1}, \ldots, w_m$$
linear unabhängig. Aus
$$0 = \underbrace{\sum_{i=1}^{r} \lambda_i v_i}_{=:x} + \underbrace{\sum_{j=r'+1}^{m} \gamma_j w_j}_{=:y}$$
folgt nämlich $x = -y$ und
$$0 \leq \sum_i \lambda_i^2 = \beta(x, x) = \beta(-y, -y) = \beta(y, y) = \sum_j \gamma_j^2 \underbrace{\beta(w_j, w_j)}_{\leq 0} \leq 0.$$
Daraus folgt $\lambda_i = 0$ für alle $i = 1, \ldots, r$ und daraus unmittelbar die Aussage. Wir erhalten damit $m = \dim(V) \geq r + m - r'$, also $r \leq r'$. Aus Symmetriegründen folgt $r = r'$, also stimmt die Anzahl der Einsen überein. Für -1 und damit 0 argumentiert man genau gleich. Aussage (iv) ist damit ebenfalls bewiesen, denn man kann zunächst orthogonal diagonalisieren (Satz 6.3.12 (iv)), wobei die Eigenwerte auf der Diagonalen auftreten. Danach skaliert man positive Eigenwerte auf 1 und negative Eigenwerte auf -1. □

Definition 7.2.9. Sei $A \in \mathrm{Sym}_m(\mathbb{R})$ eine reelle symmetrische Matrix. Sei s die Anzahl der positiven Eigenwerte und t die Anzahl der negativen Eigenwerte von A (gerechnet inklusive Vielfachheiten). Dann heißt das Paar (s, t) die **Signatur** von A. △

Bemerkung 7.2.10. (i) Zwei reelle symmetrische Matrizen sind genau dann kongruent, wenn sie dieselbe Signatur besitzen.
(ii) Die Signatur kann man berechnen, ohne die Eigenwerte zu berechnen. Dabei bringt man die Matrix mit dem symmetrischen Gauß-Algrithmus auf Diagonalform. Die Anzahl der positiven Diagonaleinträge ist die Anzahl der positiven Eigenwerte, analog für negative Eigenwerte und Null.
(iii) Ist (s, t) die Signatur von A, so ist $s + t$ gerade der Rang von A.
(iv) Eine reelle symmetrische Matrix ist genau dann positiv semidefinit, wenn sie Signatur $(r, 0)$ mit $r \leq m$ hat. Sie ist genau dann positiv definit, wenn sie Signatur $(m, 0)$ hat. △

Konstruktion 7.2.11 (Symmetrischer Gauß-Algorithmus)**.** Sei K ein Körper mit $1+1 \neq 0$ und $A = A^t \in \mathrm{Mat}_m(K)$. Der folgende Algorithmus verwendet simultane Zeilen- und Spaltentransformationen und produziert so eine zu A kongruente Diagonalmatrix:

(i) Gilt $A = 0$, sind wir fertig.

(ii) Gibt es einen Diagonaleintrag $a_{ii} \neq 0$, lässt dieser sich durch Vertauschung der i-ten und ersten Zeile und danach der i-ten mit der ersten Spalte an Position $(1,1)$ bringen. Gehe dann zu Schritt (iv).

(iii) Ist $A \neq 0$ und alle Diagonaleinträge sind gleich Null, gibt es aber einen Eintrag $a_{ij} \neq 0$ und $i < j$. Durch Addition der j-ten zur i-ten Zeile und der j-ten zur i-ten Spalte entsteht an der Stelle (i,i) der Eintrag $2a_{ij} \neq 0$. Gehe nun zu (ii).

(iv) Wir eliminieren alles unterhalb und rechts des $(1,1)$-Eintrags. Da die Matrix symmetrisch ist, werden jeweils dieselben Zeilen- wie Spaltentransformationen verwendet. Gehe dann zu (v).

(v) Gehe wieder zu (i), aber mit dem unteren rechten Block von A anstelle von A. △

Beispiel 7.2.12. Wir betrachten

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 0 \end{pmatrix} \in \mathrm{Mat}_3(\mathbb{R})$$

und führen den symmetrischen Gauß-Algorithmus durch. Die wesentlichen Schritte dabei sind:

$$A \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & -3 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 \\ 0 & -3 & -1 \\ 0 & -1 & 0 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & \frac{1}{3} \end{pmatrix}.$$

Die Signatur von A ist also $(2,1)$ und man kann A auch noch weiter kongruent transformieren zu $\mathrm{Diag}(1,1,-1)$. △

7.3 Quadratische Formen

Auch in diesem Abschnitt sei K stets ein Körper mit $1 + 1 \neq 0$ und V ein endlich dimensionaler K-Vektorraum.

Definition 7.3.1. Eine Abbildung

$$q\colon V \to K$$

heißt **quadratische Form** auf V, wenn es $r \in \mathbb{N}, f_1, \ldots, f_r, g_1, \ldots g_r \in V'$ gibt mit

$$q(v) = \sum_{i=1}^{r} f_i(v) g_i(v)$$

für alle $v \in V$. △

Proposition 7.3.2. *(i) Ist $q\colon V \to K$ eine quadratische Form, so ist*

$$\begin{aligned} \beta_q\colon V \times V &\to K \\ (v, w) &\mapsto \tfrac{1}{2}\left(q(v+w) - q(v) - q(w)\right) \end{aligned}$$

eine symmetrische Bilinearform auf V.
(ii) Ist $\beta\colon V \times V \to K$ eine symmetrische Bilinearform auf V, so ist

$$\begin{aligned} q_\beta\colon V &\to K \\ v &\mapsto \beta(v, v) \end{aligned}$$

eine quadratische Form auf V.
(iii) Die Konstruktionen aus (i) und (ii) sind invers zueinander.

Beweis. Übungsaufgabe. □

Satz 7.3.3. *Sei $q\colon V \to K$ eine quadratische Form.*
(i) Zu jeder Basis $\underline{v}$ von V existiert eine symmetrische Matrix $A = A^t \in \mathrm{Mat}_m(K)$ mit

$$q(v) = \xi_{\underline{v}}(v)^t A \xi_{\underline{v}}(v)$$

für alle $v \in V$.
(ii) Es existiert eine Basis $\underline{v}$ von V und eine Diagonalmatrix $D \in \mathrm{Mat}_m(K)$ mit

$$q(v) = \xi_{\underline{v}}(v)^t D \xi_{\underline{v}}(v)$$

für alle $v \in V$.

Beweis. Folgt direkt aus Proposition 7.3.2, Konstruktion 7.1.4 und Satz 7.2.6. □

Beispiel 7.3.4. Auf K^2 definieren wir folgende quadratische Form:

$$q((a,b)^t) := a^2 + ab - 2b^2 = (a,b)\begin{pmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & -2 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix}.$$

Wir bringen

$$A = \begin{pmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & -2 \end{pmatrix}$$

mit dem symmetrischen Gauß-Algorithmus auf folgende Diagonalgestalt:

$$D = \begin{pmatrix} 1 & 0 \\ 0 & -9 \end{pmatrix}.$$

Es gibt also eine invertierbare Matrix $P \in \mathrm{Mat}_2(K)$ mit

$$\tilde{q}((a,b)^t) := q(P(a,b)^t) = a^2 - 9b^2.$$

Gilt in K sogar $1+1+1 \neq 0$ so können wir dabei sogar die Form $a^2 - b^2$ erreichen. △

Beispiel 7.3.5 (Quadratische Gleichungen im $\mathbb{R}^2$)**.** Sei q eine quadratische Form, f eine Linearform auf $\mathbb{R}^2$ und $r \in \mathbb{R}$. Wir wollen in $\mathbb{R}^2$ alle Lösungen der quadratischen Gleichung

$$q(a,b) + f(a,b) + r = 0$$

bestimmen.
Bis auf Basiswechsel können wir q auf eine der folgenden Formen bringen (mit Hauptachsentransformation bzw. dem Sylvester'schem Trägheitssatz):

$$\begin{aligned} q(a,b) = \quad & 0 \\ & a^2, \quad -a^2, \\ & a^2 + b^2, \quad a^2 - b^2, \quad -a^2 - b^2. \end{aligned}$$

Dabei ändert sich f natürlich auch, bleibt aber eine Linearform. Für das Lösen der Gleichung können wir den letzten Fall in Zeile 2 und 3 ignorieren, indem wir alles mit -1 multiplizieren Wir schreiben (das neue) f dann als

$$f(a,b) = sa + tb$$

mit $s, t \in \mathbb{R}$ und stellen fest, dass man durch eine Translation

$$(a,b) \mapsto (a + \lambda, b + \gamma)$$

jeden linearen Term in f eliminieren kann, der quadratisch in q auftritt. Jeder nicht-eliminierbare lineare Term kann durch Skalierung der entsprechenden Koordinate dann noch normiert werden. Insgesamt ergeben sich (nach Basiswechsel und Translation) also folgende *Normalformen* für unsere Gleichungen:

$$a+b+r=0, \quad a^2+b+r=0, \quad a^2+b^2+r=0, \quad a^2-b^2+r=0.$$

Der erste Fall ist nicht besonders interessant, die Lösungsmenge ist eine (affine) Gerade in $\mathbb{R}^2$. Den zweiten Fall kann man durch $b \mapsto -b-r$ auf die Gestalt

$$b=a^2$$

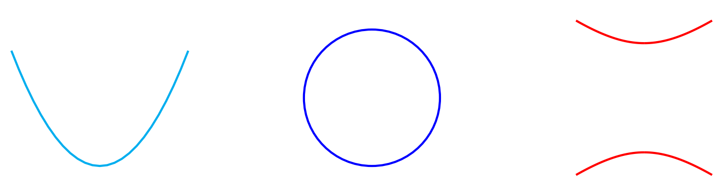
bringen, man erhält als Lösungsmenge also gerade die Standardparabel. Im dritten Fall erhält man für $r>0$ gar keine Lösung und für $r=0$ nur den Nullpunkt. Für $r<0$ kann man durch Skalierung der Koordinaten noch $r=-1$ erreichen und erhält die Standard-Kreisgleichung:

$$a^2+b^2=1.$$

Ist im letzten Fall $r=0$, erhält man wegen $a^2-b^2=(a+b)(a-b)$ eine Vereinigung von zwei Geraden als Lösungsmenge. Ansonsten kann man wieder $r=1$ oder $r=-1$ erreichen. Wenn man mit -1 multipliziert sowie a und b vertauschen, bleibt nur der Fall

$$a^2=1+b^2,$$

der gerade die Standard-Hyperbel als Lösungsmenge hat. Die Lösungsmengen der drei echt-quadratischen Fälle sind also genau die folgenden:



Parabel Kreis Hyperbel $\triangle$

Kapitel 8

Konvexität

In diesem Kapitel beschäftigen wir uns mit Systemen linearer *Ungleichungen*. Das führt zunächst zum Begriff des Halbraums und danach direkt zu dem einer *konvexen Menge*. Besonderen Schwerpunkt legen wir dann auf Polytope und Polyeder, die sowohl geometrisch schön sind, als auch in der Optimierung eine besonders wichtige Rolle spielen.
In diesem gesamten Kapitel arbeiten wir ausschließlich im reellen Vektorraum $V = \mathbb{R}^m$. Dabei verwenden wir wenn nötig das Standardskalarprodukt und die davon induzierte Norm.

8.1 Grundbegriffe

Definition 8.1.1. (i) Eine Teilmenge $S \subseteq \mathbb{R}^m$ heißt **konvex**, wenn für alle $a, b \in S$ und $\lambda \in [0,1]$ stets

$$\lambda a + (1-\lambda)b \in S$$

gilt.
(ii) Eine Teilmenge $K \subseteq \mathbb{R}^m$ heißt **konvexer Kegel**, wenn sie konvex ist und zusätzlich für $a \in K$ und $\lambda \geq 0$ stets $\lambda a \in K$ gilt. $\triangle$

Bemerkung 8.1.2. (i) Für $a \neq b \in \mathbb{R}^m$ ist die Menge

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

ein affiner Unterraum der Dimension 1, der sowohl a als auch b enthält. Die zweite Aussage erhält man für $\lambda = 0$ bzw. $\lambda = 1$, die erste folgt aus der Gleichung

$$\lambda a + (1-\lambda)b = a + (1-\lambda)(b-a).$$

Also ist $G_{a,b}$ gerade die Gerade durch a und b. Wenn man sich nun auf $\lambda \in [0,1]$ beschränkt, erhält man nur den Teil der Geraden, der *zwischen* a und b liegt. Es ist

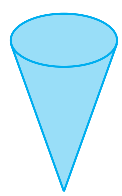
$$[a,b] := \{\lambda a + (1-\lambda)b \mid \lambda \in [0,1]\} \subseteq \mathbb{R}^m$$

also die Verbindungstrecke zwischen a und b.
Eine Menge ist also konvex, wenn sie mit je zwei Punkten auch jeden Punkt ihrer Verbindungsstrecke enthält:

nicht konvex konvex

(ii) Eine Menge K ist genau dann ein konvexer Kegel, wenn sie abgeschlossen unter positiver Skalierung und unter Summen ist:

$$\mathbb{R}_{\geq 0} \cdot K \subseteq K, \quad K + K \subseteq K.$$



△

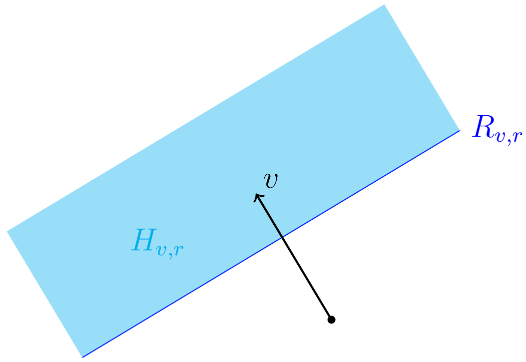
Definition 8.1.3. Für jedes $0 \neq v \in \mathbb{R}^m$ und $r \in \mathbb{R}$ ist

$$H_{v,r} := \left\{a \in \mathbb{R}^m \mid v^t a = \langle v, a\rangle \geq r\right\}$$

eine konvexe Menge, genannt **Halbraum**. Dabei heißt v **Normalenvektor** des Halbraums. Ein Halbraum ist genau die Lösungsmenge einer einzelnen (inhomogenen) linearen *Ungleichung*. Die Menge

$$R_{v,r} := \{a \in \mathbb{R}^m \mid \langle v, a\rangle = r\}$$

ist eine affine Hyperebene im Halbraum, genannt dessen **Rand**. Der Normalenvektor v zeigt in Richtung des Halbraums $H_{v,r}$, die Zahl r bestimmt die Weite der Verschiebung vom Ursprung in Richtung v.



△

Proposition 8.1.4. *(i) Jeder Halbraum in $\mathbb{R}^m$ ist konvex. Es ist $H_{v,r}$ genau dann ein konvexer Kegel, wenn $r = 0$ gilt.*
(ii) Der Durchschnitt von (beliebig vielen) konvexen Mengen in $\mathbb{R}^m$ ist wieder konvex. Dasselbe stimmt für konvexe Kegel.
(iii) Für jede Teilmenge $S \subseteq \mathbb{R}^m$ existiert die kleinste konvexe Obermenge $\operatorname{Konv}(S)$, wir bezeichnen sie als **konvexe Hülle** *von S. Es gilt*

$$\operatorname{Konv}(S) = \left\{ \sum_{i=1}^{n} \lambda_i a_i \mid n \in \mathbb{N}, a_i \in S, \lambda_i \geq 0, \sum_{i=1}^{n} \lambda_i = 1 \right\}.$$

(iv) Für $S \subseteq \mathbb{R}^m$ existiert der kleinste konvexe Oberkegel $\operatorname{KK}(S)$, wir nennen ihn den von S erzeugten **konvexen Kegel***. Es gilt*

$$\operatorname{KK}(S) := \left\{ \sum_{i=1}^{n} \lambda_i a_i \mid n \in \mathbb{N}, a_i \in S, \lambda_i \geq 0 \right\}.$$

(v) Das Innere und der Abschluss einer konvexen Menge in $\mathbb{R}^m$ ist wieder eine konvexe Menge. Der Abschluss eines konvexen Kegels ist wieder ein konvexer Kegel. Das Innere eines konvexen Kegels, zusammen mit dem Ursprung, ist wieder ein konvexer Kegel.

Beweis. Die Konvexität in (i) rechnet man direkt nach: für $a, b \in H_{v,r}$ und $\lambda \in [0, 1]$ gilt

$$\langle v, \lambda a + (1-\lambda) b \rangle = \lambda \langle v, a \rangle + (1-\lambda) \langle v, b \rangle \geq \lambda r + (1-\lambda) r = r,$$

also $\lambda a + (1-\lambda) b \in H_{v,r}$.
Da konvexe Kegel immer 0 enthalten, folgt $r \leq 0$ für jeden konvexen Kegel $H_{v,r}$. Wäre $r < 0$, so wäre $-\gamma v \in H_{v,r}$ für kleines positives γ. Dann wäre aber $-\lambda v \in H_{v,r}$ für alle $\lambda \geq 0$, ein offensichtlicher Widerspruch, denn

$$r \leq \langle v, -\lambda v \rangle = -\lambda \|v\|^2$$

kann nicht für alle $\lambda \geq 0$ gelten.
Aussage (ii) ist offensichtlich. Damit ist die Existenz von $\operatorname{Konv}(S)$ klar, es ist einfach der Durchschnitt aller konvexen Obermengen von S. Wir bezeichnen die Menge rechts in (iii) mit C. Die Konvexität von C sieht man direkt:

$$\lambda\left(\sum_i \lambda_i a_i\right) + (1-\lambda)\left(\sum_i \gamma_i b_i\right) = \sum_i \lambda\lambda_i a_i + \sum_i (1-\lambda)\gamma_i b_i$$

und das ist erneut eine Darstellung in C, denn es gilt

$$\sum_i \lambda\lambda_i + \sum_i (1-\lambda)\gamma_i = \lambda\sum_i \lambda_i + (1-\lambda)\sum_i \gamma_i = \lambda + (1-\lambda) = 1.$$

Da C offensichtlich S enthält, folgt daraus direkt $\operatorname{Konv}(S) \subseteq C$. Sei umgekehrt $a \in C$, also $a = \sum_{i=1}^n \lambda_i a_i$ mit den geforderten Eigenschaften. Wir zeigen $a \in \operatorname{Konv}(S)$ mit Induktion über n. Der Fall $n = 1$ ist klar. Im allgemeinen Fall betrachten wir die Darstellung

$$a = (1-\lambda_n)\sum_{i=1}^{n-1} \underbrace{\frac{\lambda_i}{1-\lambda_n}}_{=:\gamma_i} a_i + \lambda_n a_n$$

und berechnen

$$\sum_{i=1}^{n-1} \gamma_i = \sum_{i=1}^{n-1} \frac{\lambda_i}{1-\lambda_n} = \frac{\sum_{i=1}^{n-1} \lambda_i}{1-\lambda_n} = \frac{1-\lambda_n}{1-\lambda_n} = 1.$$

Somit können wir auf

$$b := \sum_{i=1}^{n-1} \gamma_i a_i \in C$$

die Induktionsvoraussetzung anwenden und erhalten $b \in \operatorname{Konv}(S)$. Wegen

$$a = \lambda_n a_n + (1-\lambda_n) b$$

enthält die konvexe Menge $\operatorname{Konv}(S)$ mit a_n und b also auch a.
Aussage (iv) folgt direkt aus (iii), da man zusätzlich nur noch mit positiven Zahlen skalieren muss. Aussage (v) ist Übungsaufgabe. $\square$

Satz 8.1.5 (Satz von Carathéodory)**.** *Für* $S \subseteq \mathbb{R}^m$ *gilt:*
(i) $\mathrm{KK}(S) = \{\sum_{i=1}^m \lambda_i a_i \mid a_i \in S, \lambda_i \geq 0\}$.
(ii) $\mathrm{Konv}(S) = \left\{\sum_{i=1}^{m+1} \lambda_i a_i \mid a_i \in S, \lambda_i \geq 0, \sum_{i=1}^{m+1} \lambda_i = 1\right\}$.

Beweis. (i): Wähle für $a \in \mathrm{KK}(S)$ eine Darstellung

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

von minimaler Länge (insbesondere sind alle $\lambda_i > 0$). Dann sind $a_1, \ldots, a_n$ linear unabhängig. Gäbe es nämlich eine Linearkombination

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

mit nicht allen $\gamma_i = 0$, so können wir o.B.d.A. folgendes annehmen:

$$\gamma_i \leq 0 \quad (i = 1, \ldots, r-1) \quad \text{und} \quad \gamma_i > 0 \quad (i = r, \ldots, n)$$

sowie

$$\frac{\lambda_i}{\gamma_i} \geq \frac{\lambda_n}{\gamma_n} \quad (i = r, \ldots, n).$$

Dann gilt

$$\begin{aligned} a &= \sum_{i=1}^{n-1} \lambda_i a_i + \lambda_n a_n \\ &= \sum_{i=1}^{n-1} \lambda_i a_i - \lambda_n \left(\sum_{i=1}^{n-1} \frac{\gamma_i}{\gamma_n} a_i \right) \\ &= \sum_{i=1}^{n-1} \left(\lambda_i - \frac{\lambda_n}{\gamma_n} \gamma_i \right) a_i \end{aligned}$$

und das ist eine Positivkombination mit weniger als n Summanden, ein Widerspruch. Aus der linearen Unabhängigkeit folgt natürlich $n \leq m$, die gewünschte Aussage.
(ii): Wir definieren

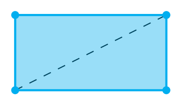
$$S^{\text{hom}} := \left\{ \begin{pmatrix} a \\ 1 \end{pmatrix} \mid a \in S \right\} \subseteq \mathbb{R}^{m+1}.$$

Dann gilt offensichtlich

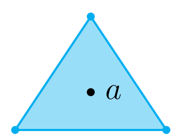
$$\mathrm{Konv}(S) = \left\{ a \in \mathbb{R}^m \mid \begin{pmatrix} a \\ 1 \end{pmatrix} \in \mathrm{KK}(S^{\mathrm{hom}}) \right\}$$

und da jedes Element von $\mathrm{KK}(S^{\mathrm{hom}})$ eine Positivkombination von $m+1$ Vektoren aus S^{hom} ist, gilt dasselbe für Elemente aus $\mathrm{Konv}(S)$. □

Bemerkung 8.1.6. (i) Jeder Punkt der konvexen Hülle von 4 Punkten im $\mathbb{R}^2$ ist bereits eine Konvexkombination von 3 der Punkte:



(ii) Die Schranke m bzw. $m+1$ kann im Allgemeinen nicht verbessert werden, wir man bereits an der konvexen Hülle von 3 Punkten im $\mathbb{R}^2$ sieht:



△

Korollar 8.1.7. *(i) Ist S endlich, so ist $\mathrm{KK}(S)$ abgeschlossen.*
(ii) Ist $S \subseteq \mathbb{R}^m$ kompakt, so auch $\mathrm{Konv}(S)$.

Beweis. (i): Nach dem Beweis von Satz 8.1.5 (i) gilt

$$\mathrm{KK}(S) = \bigcup_{S' \subseteq S \text{ linear unabhängig}} \mathrm{KK}(S').$$

Ist S endlich, so auch diese Vereinigung. Es genügt also, die Abgeschlossenheit für jedes solche S' zu zeigen. Jede linear unabhängige Menge von Vektoren lässt sich aber mit einem Isomorphismus des $\mathbb{R}^m$ zur einer der Mengen $E_r := \{e_1, \ldots, e_r\}$ mit $r \leq m$ transformieren. Da Isomorphismen (und ihre Umkehrabbildungen) stetig sind, genügt es, die Abgeschlossenheit von allen $\mathrm{KK}(E_r)$ zu zeigen. Die ist aber offensichtlich.
Für (ii) sei

$$\Delta := \left\{ (\lambda_1, \ldots, \lambda_{m+1}) \in \mathbb{R}^{m+1} \mid \lambda_i \geq 0, \sum_i \lambda_i = 1 \right\}.$$

Dann ist $\Delta \subseteq \mathbb{R}^{m+1}$ kompakt und somit auch

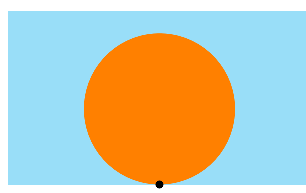
$$T := S \underbrace{\times \cdots \times}_{m+1} S \times \Delta \subseteq \mathbb{R}^{(m+1)^2}.$$

Nach Satz 8.1.5 (ii) ist $\operatorname{Konv}(S)$ aber das Bild von T unter der stetigen Abbildung

$$(a_1, \ldots, a_{m+1}, \lambda_1, \ldots, \lambda_{m+1}) \mapsto \sum_{i=1}^{m+1} \lambda_i a_i$$

und somit selbst kompakt. $\square$

Bemerkung 8.1.8. Für kompaktes S muss $\operatorname{KK}(S)$ nicht abgeschlossen sein. Ist beispielsweise $S \subseteq \mathbb{R}^2$ eine Kreisscheibe, deren Rand den Ursprung enthält, erhält man als erzeugten Kegel einen offenen Halbraum samt Ursprung im Rand:



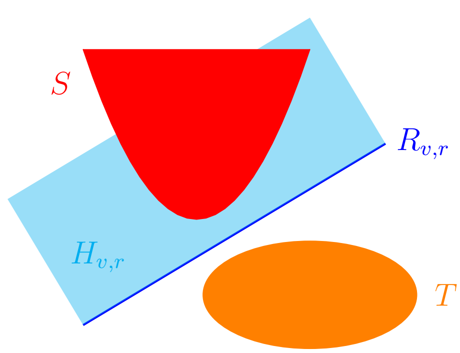
$\triangle$

8.2 Trennungssätze

Satz 8.2.1. *Seien $S, T \subseteq \mathbb{R}^m$ konvex, S abgeschlossen und T kompakt, mit*

$$S \cap T = \emptyset.$$

Dann existieren $0 \neq v \in \mathbb{R}^m, r \in \mathbb{R}$ mit $S \subseteq H_{v,r} \setminus R_{v,r}$ und $T \subseteq \mathbb{R}^m \setminus H_{v,r}$.



Beweis. Da S abgeschlossen und T kompakt ist, existieren Punkte $s_0 \in S, t_0 \in T$ mit minimalem Abstand zueinander:

$$d := \|s_0 - t_0\| \leq \|s - t\| \quad \forall s \in S, t \in T.$$

Aus $S \cap T = \emptyset$ folgt $s_0 \neq t_0$ und $d > 0$. Wir setzen nun $v := s_0 - t_0$. Wegen

$$0 < \|v\|^2 = \langle v, s_0 - t_0 \rangle = \langle v, s_0 \rangle - \langle v, t_0 \rangle$$

können wir ein $r \in \mathbb{R}$ mit

$$\langle v, t_0 \rangle < r < \langle v, s_0 \rangle$$

wählen und betrachten dazu $H_{v,r}$. Für jedes $s \in S$ und $\lambda \in [0,1]$ gilt

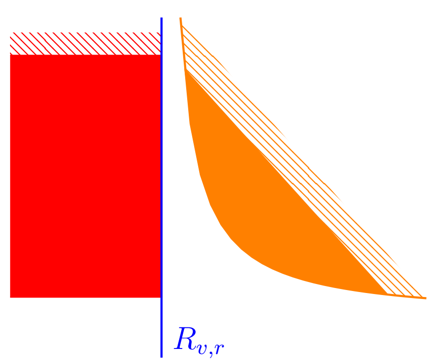
$$\begin{aligned} d^2 &\leq \| \underbrace{\lambda s + (1-\lambda) s_0}_{\in S} - t_0 \|^2 \\ &= \|\lambda(s - s_0) + v\|^2 \\ &= \lambda^2 \|s - s_0\|^2 + 2\lambda \langle s - s_0, v \rangle + \|v\|^2 \end{aligned}$$

mit Gleichheit für $\lambda = 0$. Deshalb muss die Ableitung des letzten Ausdrucks nach λ an der Stelle 0 nichtnegativ sein, d.h. $\langle s - s_0, v \rangle \geq 0$, also

$$\langle s, v \rangle \geq \langle s_0, v \rangle > r,$$

also $S \subseteq H_{v,r} \setminus R_{v,r}$. Ganz analog folgt $T \subseteq \mathbb{R}^m \setminus H_{v,r}$. □

Bemerkung 8.2.2. Die Kompaktheit einer der beiden Mengen ist notwendig für die strikte Trennungsaussage. Im folgenden Beispiel kommt für die Trennung nur ein Halbraum in Frage, die rote Menge berührt notwendigerweise dessen Rand:



△

Korollar 8.2.3. *(i) Sei $S \subseteq \mathbb{R}^m$ konvex und abgeschlossen. Für jedes $a \in \mathbb{R}^m \setminus S$ existieren dann $0 \neq v \in \mathbb{R}^m$ und $r \in \mathbb{R}$ mit*

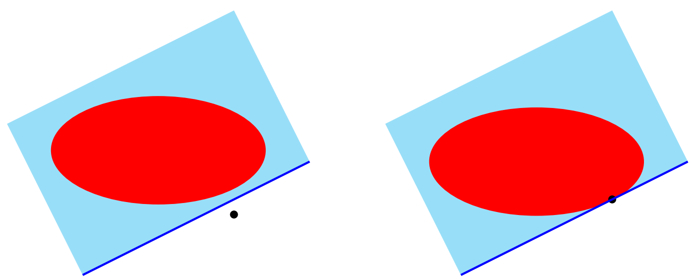
$$S \subseteq H_{v,r} \setminus R_{v,r} \quad \textit{und} \quad a \in \mathbb{R}^m \setminus H_{v,r}.$$

(ii) Sei $S \subseteq \mathbb{R}^m$ eine konvexe Menge und $a \in \partial S$. Dann existiert $0 \neq v \in \mathbb{R}^m$ und $r \in \mathbb{R}$ mit

$$S \subseteq H_{v,r} \quad \textit{und} \quad a \in R_{v,r}.$$

(iii) Für $S \subseteq \mathbb{R}^m$ gilt

$$\overline{\mathrm{Konv}(S)} = \bigcap_{S \subseteq H \,\textit{Halbraum}} H.$$



Beweis. (i) ist klar aus Satz 8.2.1, denn $T := \{a\}$ ist kompakt.
Für (ii) wählen wir eine Folge $(a_n)_{n \in \mathbb{N}}$ mit $a_n \to a$ und $a_n \notin \overline{S}$ für alle n. Nach (i) gibt es $0 \neq v_n \in \mathbb{R}^m, r_n \in \mathbb{R}$ mit

$$\langle v_n, a_n \rangle < r_n \quad \text{und} \quad S \subseteq H_{v_n, r_n}$$

für alle n. Beim Skalieren des Tupels (v_n, r_n) mit einer beliebigen positiven Zahl bleiben diese beiden Eigenschaften erhalten. Wir können also o.B.d.A. $\|v_n\| = 1$ für alle n annehmen. Aus $S \subseteq H_{v_n, r_n}$ folgt dann sofort auch die Beschränktheit der r_n (z.B. aus der Cauchy-Schwarz Ungleichung). Damit gilt o.B.d.A. $v_n \to v$ für ein $0 \neq v \in \mathbb{R}^m$ und $r_n \to r \in \mathbb{R}$. Aus Stetigkeitsgründen folgt $\langle v, a \rangle \leq r$ und $S \subseteq H_{v,r}$. Wegen $a \in \overline{S}$ folgt $a \in R_{v,r}$.
Für (iii) beobachten wir, dass mit jedem Halbraum auch der große Durchschnitt rechts konvex und abgeschlossen ist. Er enthält also $\overline{\mathrm{Konv}(S)}$. Ist umgekehrt $a \notin \overline{\mathrm{Konv}(S)}$, so gibt es nach (i) einen über S liegenden Halbraum, der a nicht enthält. Somit gehört a auch nicht zum Durchschnitt aller Halbräume über S. □

Bemerkung 8.2.4. Die Aussage von Korollar 8.2.3 (iii) kann man auch folgendermaßen formulieren: *Jede abgeschlossene konvexe Menge ist die Lösungsmenge eines (uendlichen) Systems (inhomogener) linearer Ungleichungen.* △

Definition 8.2.5. Ein $R_{v,r}$ wie in Korollar 8.2.3 (ii) heißt **Stützhyperebene** an S durch a. △

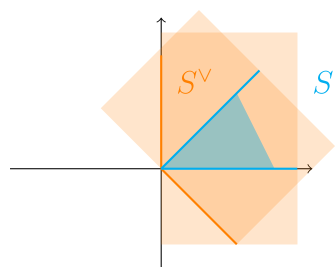
Definition 8.2.6. Für eine Menge $S \subseteq \mathbb{R}^m$ nennen wir

$$S^\vee := \{a \in \mathbb{R}^m \mid \forall s \in S\colon \langle s, a\rangle \geq 0\} = \bigcap_{s \in S} H_{s,0}$$

den zu S **dualen Kegel.** Offensichtlich handelt es sich dabei stets um einen abgeschlossenen konvexen Kegel. △

Beispiel 8.2.7. Für $S = \mathrm{KK}(\{(1,0)^t, (1,1)^t\}) \subseteq \mathbb{R}^2$ gilt

$$S^\vee = H_{(1,0)^t,0} \cap H_{(1,1)^t,0} = \mathrm{KK}((0,1)^t, (1,-1)^t).$$



△

Korollar 8.2.8 (Bidualsatz)**.** *Für $S \subseteq \mathbb{R}^m$ gilt*

$$\left(S^\vee\right)^\vee = \overline{\mathrm{KK}(S)}.$$

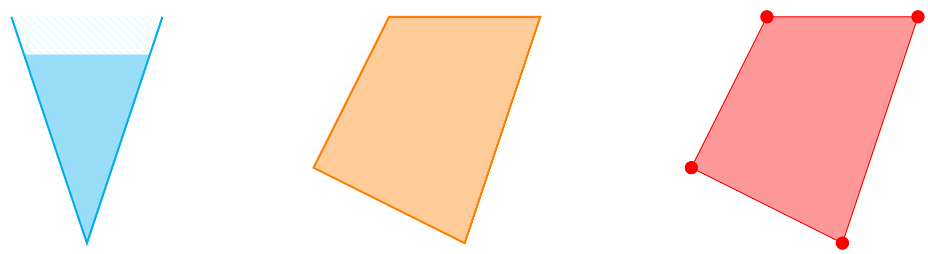
Beweis. Nach Definition gilt offensichtlich $S \subseteq \left(S^\vee\right)^\vee$ und hier ist die rechte Seite ein abgeschlossener konvexer Kegel. Das zeigt $\left(S^\vee\right)^\vee \supseteq \overline{\mathrm{KK}(S)}$. Sei umgekehrt $a \notin \overline{\mathrm{KK}(S)}$. Nach Korollar 8.2.3$(i)$ gibt es $0 \neq v \in \mathbb{R}^m$ und $r \in \mathbb{R}$ mit

$$\langle v, a\rangle < r \quad \text{ und } \langle v, b\rangle \geq r \text{ für alle } b \in \mathrm{KK}(S).$$

Es kann nun kein $b \in \mathrm{KK}(S)$ mit $\langle v, b\rangle < 0$ geben, denn $\langle v, \lambda b\rangle$ wäre dann nicht nach unten durch r beschränkt, für alle $\lambda > 0$. Aus $0 \in \mathrm{KK}(S)$ folgt außerdem $r \leq 0$. Wir können also r durch 0 ersetzen und erhalten die gewünschten Bedingungen. Das besagt aber gerade $v \in S^\vee$ und $a \notin \left(S^\vee\right)^\vee$, die gewünschte Aussage □

8.3 Polyeder und Polytope

Definition 8.3.1. (i) Ein **Polyeder** ist ein endlicher Durchschnitt von Halbräumen, bzw. die Lösungsmenge eines endlichen Systems von (inhomogenen) linearen Ungleichungen.
(ii) Ein **Polytop** ist die konvexe Hülle einer endlichen Menge.



$\triangle$

Der folgende Satz gibt ein Kriterium für die Lösbarkeit eines linearen Ungleichungssystems an (bzw. für die Frage, wann ein Polyeder nichtleer ist):

Satz 8.3.2 (Lemma von Farkas)**.** *Für* $v_1, \ldots, v_n \in \mathbb{R}^m \setminus \{0\}$ *und* $r_1, \ldots, r_n \in \mathbb{R}$ *sind äquivalent:*

(i) *Es gibt ein* $a \in \mathbb{R}^m$ *mit* $\langle v_1, a\rangle \geq r_1, \ldots, \langle v_n, a\rangle \geq r_n$.

(ii) *In* $\mathbb{R}^{m+1}$ *gilt*
$$e_{m+1} \notin \mathrm{KK}\left(\begin{pmatrix} v_1 \\ r_1 \end{pmatrix}, \ldots, \begin{pmatrix} v_n \\ r_n \end{pmatrix}\right).$$

Beweis. Nehmen wir zunächst an dass (i) gilt, (ii) aber nicht. Dann existieren $a \in \mathbb{R}^m$ mit
$$\langle v_1, a\rangle \geq r_1, \ldots, \langle v_n, a\rangle \geq r_n,$$
$\lambda_1, \ldots, \lambda_n \geq 0$ mit $\sum_i \lambda_i v_i = 0$ und $\sum_i \lambda_i r_i = 1$. Damit gilt
$$0 = \langle 0, a\rangle = \left\langle \sum_i \lambda_i v_i, a \right\rangle = \sum_i \lambda_i \langle v_i, a\rangle \geq \sum_i \lambda_i r_i = 1,$$
ein Widerspruch. Somit haben wir "$(i)\Rightarrow(ii)$" gezeigt.
Wenn umgekehrt (ii) gilt, dann gibt es nach Korollar 8.1.7 (i) und Korollar 8.2.3 (i) ein $(a, s)^t \in \mathbb{R}^{m+1} = \mathbb{R}^m \times \mathbb{R}$ und $r \in \mathbb{R}$ mit
$$\langle a, \lambda v_i\rangle + \lambda r_i s > r > s$$

für $i = 1, \dots, n$ und alle $\lambda \geq 0$. Für $\lambda = 0$ ergibt sich $s < 0$. Wenn wir die Ungleichung durch $\lambda > 0$ und $-s > 0$ dividieren, erhalten wir

$$\langle -\tfrac{1}{s}a, v_i\rangle - r_i > -\frac{1}{\lambda}$$

und daraus folgt

$$\langle -\tfrac{1}{s}a, v_i\rangle - r_i \geq 0,$$

die gewünschte Aussage (i). □

Beispiel 8.3.3. Wir betrachten das System

$$\begin{aligned} -x_1 + 2x_2 &\geq 1 \\ x_1 - x_2 &\geq 1 \\ x_1 + x_2 &\geq 0 \end{aligned}$$

Wir müssen also den von

$$x = \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix},\ y = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix},\ z = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

erzeugten Kegel in $\mathbb{R}^3$ betrachten. Es sind aber x, y, z linear unabhängig, also besitzt e_3 eine eindeutige Darstellung als

$$e_3 = \tfrac{2}{5}x + \tfrac{3}{5}y - \tfrac{1}{5}z$$

und der dritte Koeffizient ist negativ. Wegen $e_3 \notin \mathrm{KK}(x, y, z)$ besitzt das System von Ungleichungen also eine Lösung. △

Korollar 8.3.4. *Seien $v_1, \dots, v_n \in \mathbb{R}^m$ linear unabhängig. Dann besitzt das lineare Ungleichungssystem*

$$\langle v_1, x\rangle \geq r_1, \dots, \langle v_n, x\rangle \geq r_n$$

für jede Wahl der r_i eine Lösung.

Beweis. Es gilt offensichtlich $e_{m+1} \notin \mathrm{KK}\left(\begin{pmatrix} v_1 \\ r_1 \end{pmatrix}, \dots, \begin{pmatrix} v_n \\ r_n \end{pmatrix}\right)$, ansonsten wären $v_1, \dots, v_n$ linear abhängig. Die Lösbarkeit folgt also aus Satz 8.3.2. □

Nachdem wir nun ein Kriterium für die Lösbarkeit von linearen Ungleichungssystemen kennen, untersuchen wir im Folgenden, ob und wie man die Lösungsmenge auch mit endlich vielen Daten vollständig angeben kann. Dazu beweisen wir zunächst einige Hilfsaussagen.

Lemma 8.3.5. *Für $v_1, \dots, v_n \in \mathbb{R}^m \setminus \{0\}$ betrachten wir den konvexen Kegel*

$$S := \{(a,b) \in \mathbb{R}^m \times \mathbb{R}^n \mid \langle v_j, a\rangle \geq b_j, j = 1, \dots, n\}.$$

Mit

$$T = \{\pm(e_i, \langle v_1, e_i\rangle, \dots, \langle v_n, e_i\rangle) \mid i = 1, \dots, m\} \cup \{(0, -e_j) \mid j = 1, \dots, n\}$$

gilt dann

$$S = \mathrm{KK}(T).$$

Insbesondere ist S ein endlich erzeuger konvexer Kegel.

Beweis. Jedes Element von T gehört offensichtlich zu S. Da S offensichtlich ein konvexer Kegel ist, ist $S \supseteq \mathrm{KK}(T)$ gezeigt. Ein allgemeines Element von S schreiben wir als

$$(a,b) = (a, \langle v_1, a\rangle, \dots, \langle v_n, a\rangle) + (0, \underbrace{b_1 - \langle v_1, a\rangle}_{\leq 0}, \dots, \underbrace{b_n - \langle v_n, a\rangle}_{\leq 0}).$$

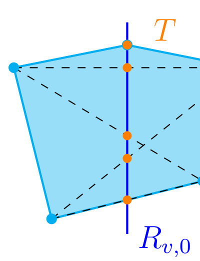
Der erste Summand ist dabei eine Positivkombination von Elementen des ersten Typs in T, der zweite eine solche von Elementen des zweiten Typs. Damit ist auch "$\subseteq$" gezeigt. $\square$

Lemma 8.3.6. *Sei $P = \mathrm{KK}(\{a_1, \dots, a_n\}) \subseteq \mathbb{R}^m$ ein endlich erzeugter Kegel und $0 \neq v \in \mathbb{R}^m$. Dann ist*

$$P \cap R_{v,0} = \mathrm{KK}(T)$$

wieder ein endlich erzeugter Kegel, wobei

$$T := \{a_i \mid a_i \in R_{v,0}\} \cup \{\langle v, a_i\rangle a_j - \langle v, a_j\rangle a_i \mid \langle v, a_i\rangle > 0, \langle v, a_j\rangle < 0\}.$$



Beweis. Man rechnet $T \subseteq P \cap R_{v,0}$ sofort nach. Daraus folgt natürlich direkt $\mathrm{KK}(T) \subseteq P \cap R_{v,0}$. Für die Umkehrung definieren wir

$$\begin{aligned}
\mathcal{P} &= \{i \mid \langle v, a_i \rangle > 0\} \\
\mathcal{Z} &= \{i \mid \langle v, a_i \rangle = 0\} \\
\mathcal{N} &= \{i \mid \langle v, a_i \rangle < 0\}\,.
\end{aligned}$$

Ein allgemeines Element $a \in P \cap R_{v,0}$ schreiben wir als

$$a = \sum_{i \in \mathcal{P}} \lambda_i a_i + \sum_{k \in \mathcal{Z}} \lambda_k a_k + \sum_{j \in \mathcal{N}} \lambda_j a_j$$

mit allen $\lambda_l \geq 0$. Aus $a \in R_{v,0}$ folgt

$$0 = \langle v, a \rangle = \sum_{i \in \mathcal{P}} \lambda_i \underbrace{\langle v, a_i \rangle}_{>0} + \sum_{k \in \mathcal{Z}} \lambda_k \underbrace{\langle v, a_k \rangle}_{=0} + \sum_{j \in \mathcal{N}} \lambda_j \underbrace{\langle v, a_j \rangle}_{<0}$$

und daraus

$$\sum_{i \in \mathcal{P}} \lambda_i \underbrace{\langle v, a_i \rangle}_{>0} = - \sum_{j \in \mathcal{N}} \lambda_j \underbrace{\langle v, a_j \rangle}_{<0} =: d \geq 0$$

Im Fall $d = 0$ ist die Aussage klar, da in der Darstellung von a nur Summanden des Typs $\mathcal{Z}$ auftreten, die entsprechenden a_k gehören alle zu T. Sei also $d > 0$. Dann gilt

$$\begin{aligned}
a &= \sum_{k \in \mathcal{Z}} \lambda_k a_k + \sum_{i \in \mathcal{P}} \tfrac{1}{d} d \lambda_i a_i + \sum_{j \in \mathcal{N}} \tfrac{1}{d} d \lambda_j a_j \\
&= \sum_{k \in \mathcal{Z}} \lambda_k a_k + \sum_{i \in \mathcal{P}, j \in \mathcal{N}} -\tfrac{1}{d} \lambda_j \lambda_i \langle v, a_j \rangle a_i + \sum_{j \in \mathcal{N}, i \in \mathcal{P}} \tfrac{1}{d} \lambda_i \lambda_j \langle v, a_i \rangle a_j \\
&= \sum_{k \in \mathcal{Z}} \lambda_k a_k + \sum_{i \in \mathcal{P}, j \in \mathcal{N}} \tfrac{1}{d} \lambda_j \lambda_i \left(\langle v, a_i \rangle a_j - \langle v, a_j \rangle a_i \right).
\end{aligned}$$

Die auftretenden Koeffizienten sind alle ≥ 0 und daraus folgt $a \in \mathrm{KK}(T)$. □

Satz 8.3.7 (Satz von Minkowski-Weyl, Teil 1). *(i) Für $v_1, \ldots, v_n \in \mathbb{R}^m$ ist*

$$C := \{a \in \mathbb{R}^m \mid \langle v_1, a \rangle \geq 0, \ldots, \langle v_n, a \rangle \geq 0\}$$

stets ein endlich erzeugter konvexer Kegel.
(ii) Jeder kompakte Polyeder ist ein Polytop.

Beweis. (i): Wir betrachten

$$S := \{(a,b) \in \mathbb{R}^m \times \mathbb{R}^n \mid \langle v_j, a\rangle \geq b_j, j = 1, \dots, n\}\,.$$

Nach Lemma 8.3.5 handelt es sich bei S um einen endlich erzeugten konvexen Kegel. Wir verwenden nun iterativ Lemma 8.3.6 und sehen, dass die folgende Menge ebenfalls ein endlich erzeugter konvexer Kegel ist:

$$\tilde{S} := \{(a,0) \in \mathbb{R}^m \times \mathbb{R}^n \mid \langle v_j, a\rangle \geq 0, j = 1, \dots, n\}\,.$$

Wenn er von Elementen

$$\begin{pmatrix} a_1 \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} a_r \\ 0 \end{pmatrix}$$

erzeugt wird, dann erzeugen die Elemente $a_1, \dots, a_r$ offensichtlich C.
Für (ii) sei

$$P = \{a \in \mathbb{R}^m \mid \langle v_i, a\rangle \geq r_i, i = 1, \dots, n\} \subseteq \mathbb{R}^m$$

ein kompakter Polyeder. Wir betrachten dann

$$P^{\text{hom}} := \left\{ \begin{pmatrix} \lambda a \\ \lambda \end{pmatrix} \mid a \in P, \lambda \geq 0 \right\} \subseteq \mathbb{R}^{m+1}.$$

Bei P^{hom} handelt es sich um einen konvexen Kegel, der von endlich vielen homogenen linearen Ungleichungen definiert wird. Für $\gamma > 0, b \in \mathbb{R}^m$ gilt nämlich

$$\begin{aligned}
\begin{pmatrix} b \\ \gamma \end{pmatrix} \in P^{\text{hom}} &\Leftrightarrow b/\gamma \in P \\
&\Leftrightarrow \langle v_i, b/\gamma\rangle \geq r_i,\ i = 1, \dots, n \\
&\Leftrightarrow \langle v_i, b\rangle - r_i\gamma \geq 0,\ i = 1, \dots, n \\
&\Leftrightarrow \left\langle \begin{pmatrix} v_i \\ -r_i \end{pmatrix}, \begin{pmatrix} b \\ \gamma \end{pmatrix} \right\rangle \geq 0,\ i = 1, \dots, n.
\end{aligned}$$

Weiter gilt für $b \in \mathbb{R}^m$

$$0 \leq \left\langle \begin{pmatrix} v_i \\ -r_i \end{pmatrix}, \begin{pmatrix} b \\ 0 \end{pmatrix} \right\rangle = \langle v_i, b\rangle \quad \Leftrightarrow \quad b = 0 \quad \Leftrightarrow \quad \begin{pmatrix} b \\ 0 \end{pmatrix} \in P^{\text{hom}}.$$

Für die erste der beiden Äquivalenzen benutzen wir die Kompaktheit von P, denn der Halbstrahl in Richtung b an jedem Punkt von P liegt unter der linken Annahme ganz in P.

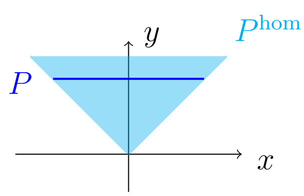
Nach (i) ist nun P^{hom} ein endlich erzeugter Kegel. Wenn wir Erzeuger der Form

$$\begin{pmatrix} a_1 \\ 1 \end{pmatrix}, \ldots, \begin{pmatrix} a_r \\ 1 \end{pmatrix}$$

wählen, gilt offensichtlich $P = \operatorname{Konv}(a_1, \ldots, a_r)$. □

Bemerkung 8.3.8. (i) Satz 8.3.7 (ii) besagt, dass man die Lösungsmenge eines Systems linearen Ungleichungen mit endlich vielen Daten vollständig angeben kann (zumindest im Fall einer kompakten Lösungsmenge). Man kann endlich viele Lösungen finden, aus denen jede andere Lösungen als Konvexkombination hervorgeht.
(ii) Der Beweis von Satz 8.3.7 ist konstruktiv. Wenn wir mit Ungleichungen für C oder P beginnen, erhalten wir konkrete Ungleichungen für S aus Lemma 8.3.5 und mit Lemma 8.3.6 explizite Erzeuger für C bzw. P. Wir haben also eine Methode gefunden, um ein System von linearen Ungleichungen vollständig lösen zu können (zumindest unter der Kompaktheitsannahme). △

Beispiel 8.3.9. (i) Um den Rechenaufwand möglichst gering zu halten, schauen wir uns ein sehr einfaches Beispiel an. Dazu sei $P \subseteq \mathbb{R}^1$ der von den Gleichungen $-1 \leq x$ und $x \leq 1$ definierte Polyeder. Der Polyeder $P^{\text{hom}} \subseteq \mathbb{R}^2$ wird dann von den beiden Gleichungen $x + y \geq 0$ und $-x + y \geq 0$ definiert.



Die dazugehörige Menge

$$S = \left\{(a, b) \in \mathbb{R}^2 \times \mathbb{R}^2 \mid a_1 + a_2 \geq b_1, a_2 - a_1 \geq b_2\right\}$$

wird nach Lemma 8.3.5 von der folgenden Menge von Spaltenvektoren erzeugt:

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

Wir eliminieren nun durch Konvexkombinationen von jeweils zwei der Vektoren alle Einträge in der dritten Zeile (wie in Lemma 8.3.6) und erhalten die folgenden Spalten (Nullspalten lassen wir gleich weg):

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

Anschließend eliminieren wir noch die letzte Zeile:

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

Die 3 Spaltenvektoren

$$\begin{pmatrix} -1 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

erzeugen also $P^{\hom}$ und die 3 Vektoren $-1, 0, 1$ erzeugen P. △

Satz 8.3.10 (Satz von Minkowski-Weyl, Teil 2)**.** *(i) Jeder endlich erzeugte konvexe Kegel ist durch endlich viele homogene lineare Ungleichungen definierbar.*
(ii) Jedes Polytop ist ein kompakter Polyeder.

Beweis. Für (i) sei $C = \mathrm{KK}(S)$ mit $S \subseteq \mathbb{R}^m$ endlich. Dann ist

$$C^\vee = \{a \in \mathbb{R}^m \mid \forall s \in S \colon \langle s, a\rangle \geq 0\}$$

ein durch endlich viele homogene lineare Ungleichungen definiert. Nach Satz 8.3.7 (i) gibt es also $a_1, \ldots, a_r \in \mathbb{R}^m$ mit

$$C^\vee = \mathrm{KK}(a_1, \ldots, a_r).$$

Es ist C als endlich erzeugter konvexer Kegel nach Korollar 8.1.7 abgeschlossen. Aus Satz 8.2.8 folgt also

$$C = (C^\vee)^\vee = \{a \in \mathbb{R}^m \mid \langle a_1, a\rangle \geq 0, \ldots, \langle a_r, a\rangle \geq 0\}.$$

Aussage (ii) ist Übungsaufgabe. □

Bemerkung 8.3.11. Ist $P \subseteq \mathbb{R}^m$ ein Polytop oder ein endlich erzeugter konvexer Kegel, sowie $\varphi\colon \mathbb{R}^m \to \mathbb{R}^n$ eine lineare Abbildung, so ist $\varphi(P) \subseteq \mathbb{R}^n$ wieder ein Polytop bzw. ein endlich erzeugter Kegel. Aus der Linearität folgt nämlich unmittelbar

$$\varphi\left(\mathrm{Konv}(S)\right) = \mathrm{Konv}\left(\varphi(S)\right) \quad \text{bzw.} \quad \varphi\left(\mathrm{KK}(S)\right) = \mathrm{KK}\left(\varphi(S)\right).$$

Für polyedrische Kegel oder kompakte Polyeder stimmt die Aussage wegen Satz 8.3.7 und Satz 8.3.10 also ebenfalls. Man kann sie aber auch für beliebige Polyeder zeigen und die Konstruktion im Beweis liefert sogar eine weitere Methode zur Entscheidung der Lösbarkeit eines Systems linearen Ungleichungen. △

Konstruktion 8.3.12 (Fourier-Motzkin Elimination)**.** Gegeben sei ein System von linearen Ungleichungen

$$\begin{aligned}
\langle v_1, x\rangle &= v_{11}x_1 + v_{12}x_2 + \cdots + v_{1m}x_m \geq r_1\\
\langle v_2, x\rangle &= v_{21}x_1 + v_{22}x_2 + \cdots + v_{2m}x_m \geq r_2\\
&\ \ \vdots \qquad\qquad\qquad\qquad\qquad\qquad \vdots\\
\langle v_n, x\rangle &= v_{n1}x_1 + v_{n2}x_2 + \cdots + v_{nm}x_m \geq r_n
\end{aligned}$$

mit Lösungsmenge $P \subseteq \mathbb{R}^m$, einem Polyeder. Wir betrachten die Projektion

$$\begin{aligned}
\pi\colon \mathbb{R}^m &\to \mathbb{R}^{m-1}\\
(a_1, \ldots, a_m) &\mapsto (a_1, \ldots, a_{m-1})
\end{aligned}$$

und wollen die Menge $Q := \pi(P) \subseteq \mathbb{R}^{m-1}$ beschreiben. Dazu sortieren wir die Ungleichungen so, dass

$$0 < v_{1m}, \ldots, v_{rm}, \quad 0 > v_{r+1,m}, \ldots, v_{sm}, \quad 0 = v_{s+1m} = \ldots = v_{nm}$$

gilt. Dann benutzen die letzten $n-s$ Ungleichungen die Variable x_m überhaupt nicht und können also direkt zur Beschreibung von Q verwendet werden. Für $1 \leq i \leq r$ schreiben wir die i-te Gleichung um zu

$$x_m \geq -\frac{v_{i1}}{v_{im}}x_1 - \cdots - \frac{v_{im-1}}{v_{im}}x_{m-1} + \frac{r_i}{v_{im}}.$$

Für $r+1 \leq j \leq s$ schreiben wir die j-te Gleichung als

$$-\frac{v_{j1}}{v_{jm}}x_1 - \cdots - \frac{v_{jm-1}}{v_{jm}}x_{m-1} + \frac{r_j}{v_{jm}} \geq x_m.$$

Für jeden Punkt $(a_1, \ldots, a_{m-1}) \in Q = \pi(P)$ existiert eine Zahl a_m, so dass $(a_1, \ldots, a_{m-1}, a_m)$ alle ursprünglichen Ungleichungen erfüllt. Für jedes $1 \leq i \leq r < j \leq s$ gilt also

$$-\frac{v_{j1}}{v_{jm}}a_1 - \cdots - \frac{v_{jm-1}}{v_{jm}}a_{m-1} + \frac{r_j}{v_{jm}} \geq a_m \geq -\frac{v_{i1}}{v_{im}}a_1 - \cdots - \frac{v_{im-1}}{v_{im}}a_{m-1} + \frac{r_i}{v_{im}}$$

und somit natürlich

$$-\frac{v_{j1}}{v_{jm}}a_1 - \cdots - \frac{v_{jm-1}}{v_{jm}}a_{m-1} + \frac{r_j}{v_{jm}} \geq -\frac{v_{i1}}{v_{im}}a_1 - \cdots - \frac{v_{im-1}}{v_{im}}a_{m-1} + \frac{r_i}{v_{im}}.$$

In der letzten Ungleichung tritt a_m nicht mehr auf, sie gilt also in $\mathbb{R}^{m-1}$ für alle Punkte aus Q.
Erfülle umgekehrt ein Punkt $(a_1, \ldots, a_{m-1})$ alle diese Gleichungen, für $1 \leq i \leq r < j \leq s$. Dann gilt natürlich

$$\begin{aligned} &\min_{r<j\leq s}\left\{-\frac{v_{j1}}{v_{jm}}a_1 - \cdots - \frac{v_{jm-1}}{v_{jm}}a_{m-1} + \frac{r_j}{v_{jm}}\right\} \\ \geq\ &\max_{1\leq i\leq r}\left\{-\frac{v_{i1}}{v_{im}}a_1 - \cdots - \frac{v_{im-1}}{v_{im}}a_{m-1} + \frac{r_i}{v_{im}}\right\} \end{aligned}$$

und es gibt eine Zahl $a_m \in \mathbb{R}$, die zwischen diesem Minimum und Maximum liegt. Offensichtlich erfüllt dann $(a_1, \ldots, a_m)$ alle ursprünglichen Gleichungen und liegt also in P. Somit gehört $(a_1, \ldots, a_{m-1})$ zu Q. Insgesamt haben wir endlich viele linearen Ungleichungen in den Variablen $x_1, \ldots, x_{m-1}$ konstruiert, die die Projektion $\pi(P)$ als Polyeder definieren. △

Korollar 8.3.13. *Sei $\varphi\colon \mathbb{R}^m \to \mathbb{R}^n$ eine lineare Abbildung und $P \subseteq \mathbb{R}^m$ ein Polyeder. Dann ist auch $\varphi(P) \subseteq \mathbb{R}^n$ ein Polyeder.*

Beweis. Falls φ eine Projektion ist, folgt die Aussage aus der Fourier-Motzkin Elimination 8.3.12. Wir überlegen uns nun noch, wie wir den Fall einer allgemeinen linearen Abbildung auf den einer Projektion zurückführen können. Für jede lineare Abbildung ist

$$\operatorname{Graph}(\varphi) = \{(a, \varphi(a)) \mid v \in \mathbb{R}^m\} \subseteq \mathbb{R}^m \times \mathbb{R}^n$$

ein Untervektorraum. Die Abbildung

$$\begin{aligned} \iota\colon \mathbb{R}^m &\to \operatorname{Graph}(\varphi) \\ a &\mapsto (a, \varphi(a)) \end{aligned}$$

ist ein Isomorphismus, also ist $\iota(P)$ offensichtlich ein Polyeder, denn es gilt ja

$$\langle \iota^{-*}(v), \iota(a) \rangle = \langle v, a \rangle$$

für beliebige Vektoren $v \in \mathbb{R}^m$. Die Abbildung φ ist aber die Komposition von ι mit der Projektion auf die zweite Komponente. Diese Projektion ist eine Iterationen von einfachen Projektionen wie in Konstruktion 8.3.12 und damit ist die Aussage bewiesen. $\square$

Bemerkung 8.3.14. Die Fourier-Motzkin Elimination kann zur Überprüfung der Lösbarkeit eines Systems linearer Ungleichungen verwendet werden. Man projiziert die Lösungsmenge iterativ bis auf $\mathbb{R}^1$ (oder sogar $\mathbb{R}^0$) und berechnet jeweils definierende Ungleichungen. Die Lösbarkeit ganz am Schluss der Prozedur kann man dann an den Gleichungen direkt ablesen. Und natürlich ist eine Menge genau dann nicht leer, wenn ihr Bild unter einer Abbildung nicht leer ist. $\triangle$

Beispiel 8.3.15. Wir betrachten nochmal das System

$$\begin{aligned} -x_1 + 2x_2 &\geq 1 \\ x_1 - x_2 &\geq 1 \\ x_1 + x_2 &\geq 0 \end{aligned}$$

aus Beispiel 8.3.3. Wir eliminieren die Variable x_2 und erhalten die beiden Ungleichungen

$$\begin{aligned} x_1 - 1 &\geq \tfrac{1}{2} + \tfrac{1}{2}x_1 \\ x_1 - 1 &\geq \qquad -x_1 \end{aligned}$$

beziehungsweise

$$x_1 \geq 3,\ x_1 \geq \tfrac{1}{2}.$$

Die Lösbarkeit dieses Systems sieht man sofort, also ist auch das ursprüngliche System lösbar. $\triangle$

Kapitel 9

Verschiedenes

Im letzten Kapitel sammeln wir noch einige Ergebnisse und Konstruktionen, die in verschiedenen Varianten in der Mathematik immer wieder auftreten und die wir bisher nicht systematisch eingeführt haben. Ebenso untersuchen wir auch noch den Fall von Vektorräumen mit unendlicher Dimension und sehen, wie man alle bisherigen Begriffe und Ergebnisse hier anpassen kann.

9.1 Konstruktionen mit Vektorräumen

In diesem Abschnitt sei stets K ein beliebiger aber fest gewählter Körper.

9.1.1 Summen und Produkte

Definition 9.1.1. Sei I eine beliebige Indexmenge und für jedes $i \in I$ sei V_i ein K-Vektorraum.
(i) Das **direkte/kartesische Produkt** der V_i ist der folgende Vektorraum:

$$\prod_{j \in I} V_j := \{(v_j)_{j \in I} \mid \forall j \in I \colon v_j \in V_j\}.$$

Dabei sind Addition und skalare Multiplikation komponentenweise definiert, die Vektorraumaxiome gelten dafür offensichtlich.
(ii) Für jedes $i \in I$ ist die kanonische **Projektion**

$$\pi_i \colon \prod_{j \in I} V_j \twoheadrightarrow V_i; \quad (v_j)_{j \in I} \mapsto v_i$$

eine surjektive lineare Abbildung. △

Beispiel 9.1.2. (i) Für $I = \{1, 2\}$ ist das direkte Produkt einfach das bereits bekannte kartesische Produkt $V_1 \times V_2$.
(ii) Für $I = \mathbb{N}$ und $V_i = \mathbb{R}$ für alle $i \in I$ erhalten wir als direktes Produkt den Raum aller reellen Folgen. △

Proposition 9.1.3. *(i) Das direkte Produkt hat folgende* **universelle Eigenschaft***: Für jeden Vektorraum W und jedes System linearer Abbildungen $\varphi_i \colon W \to V_i$ (für $i \in I$) gibt es genau eine lineare Abbildung $\varphi \colon W \to \prod_{j\in I} V_j$ mit*

$$\varphi_i = \pi_i \circ \varphi$$

für alle $i \in I$.

$$\begin{tikzcd}
 & & V_i \\
W \arrow[rr, bend left, "\varphi_i"] \arrow[r, "\exists!\,\varphi"'] & \prod_{j\in I} V_j \arrow[ur, "\pi_i"'] &
\end{tikzcd}$$

(ii) Die universelle Eigenschaft des direkten Produkts bestimmt das direkte Produkt eindeutig, bis auf eindeutige Isomorphie. Das heißt, ist V ein Vektorraum zusammen mit linearen Abbildungen $\tilde{\pi}_i \colon V \to V_i$ für alle $i \in I$, welcher die obere Eigenschaft ebenso erfüllt wie $\prod_{j\in I} V_j$, so gibt es genau einen Isomorphismus $\psi \colon V \to \prod_{j\in I} V_j$ mit $\pi_i \circ \psi = \tilde{\pi}_i$ für alle $i \in I$.

Beweis. (i) Wir definieren die Abbildung $\varphi \colon W \to \prod_{i\in I} V_i$ einfach durch folgende Vorschrift:

$$\varphi(w) := (\varphi_i(w))_{i\in I}\,.$$

Die Linearität von φ folgt direkt aus der Linearität der φ_i und der Wahl der Operationen im Produkt. Die Eigenschaft $\varphi_i = \pi_i \circ \varphi$ ist ebenfalls offensichtlich. Diese Eigenschaft sagt umgekehrt auch, dass man φ gar nicht anders als so definieren kann, also ist φ eindeutig.
(ii) Zunächst verwenden wir die universelle Eigenschaft für $\prod_{j\in I} V_j$, mit V anstelle von W und den $\tilde{\pi}_i$ anstelle der φ_i. Wir erhalten genau eine lineare Abbildung $\psi \colon V \to \prod_{j\in I} V_j$ mit $\pi_i \circ \psi = \tilde{\pi}_i$.

$$\begin{tikzcd}
 & V_i & \\
V \arrow[ur, "\tilde{\pi}_i"] \arrow[rr, bend left, "\psi"] & & \prod_{j\in I} V_j \arrow[ul, "\pi_i"'] \arrow[ll, bend left, "\tilde{\psi}"]
\end{tikzcd}$$

Umgekehrt erhalten wir genau eine Abbildung $\tilde{\psi}\colon \prod_{j\in I} V_j \to V$, welche das Diagramm ebenfalls kommutativ macht. Die Hintereinanderausführung

$$\gamma := \psi \circ \tilde{\psi}\colon \prod_{j\in I} V_j \to \prod_{j\in I} V_j$$

erfüllt dann

$$\pi_i \circ \gamma = \pi_i,$$

was auch für die Identität statt γ gilt. Aus der Eindeutigkeit für φ in der universtellen Eigenschaft von $\prod_{j\in I} V_j$ folgt $\gamma = \mathrm{id}_{\prod_{j\in I} V_j}$. Ganz analog folgt auch $\tilde{\psi} \circ \psi = \mathrm{id}_V$, also ist ψ ein Isomorphismus. □

Definition 9.1.4. Sei I eine beliebige Indexmenge und für jedes $i \in I$ sei V_i ein K-Vektorraum.
(i) Die **direkte Summe** der V_i ist der folgende Untervektorraum von $\prod_{j\in I} V_j$:

$$\bigoplus_{j\in I} V_j := \left\{ (v_j)_{j\in I} \in \prod_{j\in I} V_j \mid \text{ nur endlich viele } v_j \neq 0 \right\}.$$

(ii) Für jedes $i \in I$ ist die kanonische **Einbettung**

$$\iota_i\colon V_i \hookrightarrow \bigoplus_{j\in I} V_j$$
$$v \mapsto (0, \ldots, 0, \overset{\overset{i}{\downarrow}}{v}, 0, \ldots)$$

eine injektive lineare Abbildung. △

Beispiel 9.1.5. (i) Für jede endliche Indexmenge stimmt $\bigoplus_{j\in I} V_j$ mit $\prod_{j\in I} V_j$ überein.
(ii) Für $I = \mathbb{N}$ und $V_i = \mathbb{R}$ für alle $i \in I$ ist $\bigoplus_{j\in I} V_j$ ein echter Unterraum von $\prod_{j\in I} V_j$. Er besteht aus allen Folgen mit endlichem Träger. △

Proposition 9.1.6. *(i) Die direkte Summe hat folgende universelle Eigenschaft:
Für jeden Vektorraum W und jedes System linearer Abbildungen $\varphi_i\colon V_i \to W$ (für $i \in I$) gibt es genau eine lineare Abbildung $\varphi\colon \bigoplus_{j\in I} V_j \to W$ mit*

$$\varphi_i = \varphi \circ \iota_i$$

für alle $i \in I$.

$$\begin{tikzcd} V_i \arrow[dr, "\iota_i"'] \arrow[drr, bend left, "\varphi_i"] & & \\ & \bigoplus_{j\in I} V_j \arrow[r, "\exists!\varphi"] & W \end{tikzcd}$$

(ii) Die universelle Eigenschaft der direkten Summe bestimmt sie eindeutig, bis auf eindeutige Isomorphie.

Beweis. (i) Wir definieren die Abbildung $\varphi\colon\ \bigoplus_{j\in I} V_j \to W$ durch die Vorschrift

$$\varphi\left((v_j)_{j\in I}\right) := \sum_{j\in I} \varphi_j(v_j).$$

Dabei verwenden wir, dass nur endlich viele der v_j nicht Null und alle φ_j linear sind, die Summe also in Wirklichkeit nur eine endliche Summe ist. Dann ist φ offensichtlich linear und erfüllt $\varphi_i = \varphi \circ \iota_i$. Diese Eigenschaft, zusammen mit der Linearität von φ, zeigt aber dass man φ gar nicht anders definieren kann also so. Somit ist φ eindeutig bestimmt. (ii) ist Übungsaufgabe. □

Proposition 9.1.7. *Es gilt*

$$\left(\bigoplus_{j\in I} V_j\right)' \cong \prod_{j\in I} V_j' \text{ und } \bigoplus_{j\in I} V_j' \hookrightarrow \left(\prod_{j\in I} V_j\right)'.$$

Beweis. Wir betrachten die lineare Abbildung

$$\begin{aligned} \left(\bigoplus_{j\in I} V_j\right)' &\to \prod_{j\in I} V_j' \\ f &\mapsto (f \circ \iota_j)_{j\in I}. \end{aligned}$$

Sie ist offensichtlich injektiv, denn es gilt $f = \sum_{j\in I} (f \circ \iota_j)$, wie wir in Proposition 9.1.6 gezeigt haben. Die universelle Eigenschaft der direkten Summe liefert aber auch sofort die Surjektivität.

Umgekehrt betrachten wir die lineare Abbildung

$$\begin{aligned} \bigoplus_{j\in I} V_j' &\to \left(\prod_{j\in I} V_j\right)' \\ (f_j)_{j\in I} &\mapsto \left(\sum_{j\in I} f_j\colon (v_j)_{j\in I} \mapsto \sum_{j\in I} f_j(v_j)\right). \end{aligned}$$

Sie ist offensichtlich injektiv. Man beachte, dass die Abbildung im Allgemeinen nicht surjektiv, also kein Isomorphismus ist (Übungsaufgabe). □

9.1.2 Faktorräume

Konstruktion 9.1.8. Sei V ein K-Vektorraum und $U \subseteq V$ ein Untervektorraum. Wir definieren auf V eine Äquivalenzrelation:

$$v \sim_U w \quad :\Leftrightarrow \quad v - w \in U.$$

Aufgrund der Unterraumaxiome handelt es sich hier wirklich um eine Äquivalenzrelation. Die Äquivalenzklassen

$$[v]_U = v + U = \{v + u \mid u \in U\}$$

sind dann genau die zu U parallelen affinen Unterräume von V. Für die Faktormenge schreiben wir auch

$$V/U := \{v + U \mid v \in V\}$$

mit der Sprechweise V *modulo* U. Auf V/U definieren wir nun Addition und Skalarmultiplikation wie folgt:

$$\begin{aligned}(v + U) + (w + U) &:= (v + w) + U \\ \lambda \cdot (v + U) &:= (\lambda v) + U.\end{aligned}$$ △

Lemma 9.1.9. *(i) Die obigen Verknüpfungen sind wohldefiniert und machen V/U zu einem Vektorraum.*
(ii) Die kanonische Projektion

$$\begin{aligned}\pi_U\colon V &\to V/U \\ v &\mapsto v + U\end{aligned}$$

ist linear und surjektiv, mit $\mathrm{Kern}(\pi_U) = U$.

Beweis. Für (i) nehmen wir $v+U = v'+U, w+U = w'+U$, also $v-v', w-w' \in U$ an. Dann gilt

$$(v + w) - (v' + w') = (v - v') + (w - w') \in U,$$

wobei wir die Unterraumeigenschaft verwendet haben. Daraus folgt

$$(v+w)+U = (v'+w')+U,$$

also die Wohldefiniertheit der Addition. Die Wohldefiniertheit der Skalarmultiplikation geht analog. Die Vektorraumaxiome für V/U folgen nun unmittelbar aus denen von V. (ii) ist offensichtlich. □

Definition 9.1.10. Der Vektorraum V/U heißt **Faktorraum** von V modulo U. △

Korollar 9.1.11. *Ist V endlich dimensional, so gilt*

$$\dim(V/U) = \dim(V) - \dim(U).$$

Beweis. Folgt unmittelbar aus der Dimensionformel 3.2.16 für die Projektion $\pi_U \colon V \twoheadrightarrow V/U$. □

Proposition 9.1.12. *(i) Der Faktorraum V/U hat folgende universelle Eigenschaft: Für jede lineare Abbildung $\varphi\colon V \to W$ mit $U \subseteq \mathrm{Kern}(\varphi)$ existiert genau eine lineare Abbildung $\bar{\varphi}\colon V/U \to W$ mit $\varphi = \bar{\varphi} \circ \pi_U$.*

$$\begin{tikzcd} V \arrow[rr, "\varphi"] \arrow[dr, "\pi_U"'] & & W \\ & V/U \arrow[ur, "\exists!\bar{\varphi}"'] & \end{tikzcd}$$

(ii) Die universelle Eigenschaft definiert den Faktorraum eindeutig, bis auf eindeutige Isomorphie. Das heißt, ist X ein Raum und $\tilde{\pi}\colon V \to X$ surjektiv mit $\mathrm{Kern}(\tilde{\pi}) = U$, so dass für jede lineare Abbildung $\varphi\colon V \to W$ mit $U \subseteq \mathrm{Kern}(\varphi)$ genau eine lineare Abbildung $\bar{\varphi}\colon X \to W$ mit $\varphi = \bar{\varphi} \circ \tilde{\pi}$ existiert, so gibt es genau einen Isomorphismus $\psi\colon X \to V/U$ mit $\psi \circ \tilde{\pi} = \pi_U$.

Beweis. Wir definieren $\bar{\varphi}$ durch die Vorschrift

$$\bar{\varphi}(v+U) := \varphi(v).$$

Das ist wohldefiniert, denn aus $v+U = v'+U$ folgt $v - v' \in U$ und daraus

$$0 = \varphi(v-v') = \varphi(v) - \varphi(v'),$$

also $\varphi(v) = \varphi(v')$. Die Linearität und die gewünschten Eigenschaften von $\bar{\varphi}$ sind dann offensichtlich. Wegen

$$\varphi(v) = \bar{\varphi}(\pi_U(v)) = \bar{\varphi}(v+U)$$

können wir $\bar{\varphi}$ auch gar nicht anders definieren, wenn wir die gewünschte Eigenschaft erhalten wollen. Somit ist $\bar{\varphi}$ eindeutig bestimmt. (ii) ist wieder Übungsaufgabe. $\square$

Beispiel 9.1.13. Sei $V = C([-1,1], \mathbb{R})$ der Raum aller stetigen reellwertigen Funktionen auf $[-1,1]$. Sei $Y = [0,1]$ und

$$U := \{f \in V \mid f \equiv 0 \,\text{auf}\, Y\}.$$

Dann gilt

$$V/U \cong C(Y, \mathbb{R}).$$

Dazu betrachten wir zunächst die surjektive lineare Abbildung

$$\begin{aligned} r\colon V &\to C(Y, \mathbb{R}) \\ f &\mapsto f_{|_Y}. \end{aligned}$$

Offensichtlich gilt $\operatorname{Kern}(r) = U$ und wir erhalten die surjektive lineare Abbildung

$$\bar{r}\colon V/U \to C(Y, \mathbb{R}).$$

Aus

$$0 = \bar{r}(f + U) = r(f)$$

folgt aber $f \in U$, also $f + U = 0 + U = 0 \in V/U$. Also ist $\bar{r}$ auch injektiv. $\triangle$

Proposition 9.1.14. *Es gilt*

$$(V/U)' \cong \left\{ f \in V' \mid f_{|_U} = 0 \right\}.$$

Beweis. Wir bezeichnen den rechten Unterraum von V' mit X. Dann erhalten wir die lineare Abbildung

$$\begin{aligned} (V/U)' &\to X \\ f &\mapsto f \circ \pi_U. \end{aligned}$$

Aus der Surjektivität von π_U folgt die Injektivität dieser Abbildung, aus der universellen Eigenschaft des Faktorraums folgt die Surjektivität. $\square$

9.1.3 Tensorprodukte

Im Gegensatz zu den vorigen Konstruktionen führen wir das Tensorprodukt direkt über seine universelle Eigenschaft ein. Wir müssen dann zwar die Existenz hinterher zeigen, aber da die Konstruktion so technisch ist, arbeitet man in der Praxis grundsätzlich nur mit der universellen Eigenschaft.

Definition 9.1.15. Seien V, W zwei K-Vektorräume. Ein **Tensorprodukt** von V und W ist ein Vektorraum T, zusammen mit einer bilinearen Abbildung

$$\beta\colon V \times W \to T,$$

welcher die folgende universelle Eigenschaft erfüllt:
Für jeden Vektorraum X und jede bilineare Abbildung $\gamma\colon V \times W \to X$ existiert genau eine *lineare* Abbildung

$$\varphi\colon T \to X$$

mit $\gamma = \varphi \circ \beta$.

$$\begin{tikzcd} V \times W \arrow[r, "\beta"] \arrow[dr, "\gamma"'] & T \arrow[d, "\exists! \varphi"] \\ & X \end{tikzcd}$$

△

Proposition 9.1.16. *Für je zwei Vektorräume V, W existiert bis auf eindeutige Isomorphie genau ein Tensorprodukt.*

Beweis. Existenz: Wir betrachten die direkte Summe von $V \times W$-vielen Kopien von K:

$$F := \bigoplus_{V \times W} K \cong \left\{ \sum_{(v,w) \in V \times W \text{endlich}} \lambda_{(v,w)} \cdot (v, w) \mid \lambda_{(v,w)} \in K \right\}.$$

Man kann die Elemente von F als formale (endliche) Linearkombinationen von Tupeln (v, w) auffassen, was wir hier tun. Wir betrachten weiter den Untervektorraum U von F, der von allen Elementen der folgenden Gestalt aufgespannt wird, wobei $v, v' \in V, w, w' \in W, \lambda \in K$ gilt:

$$\begin{gathered} (v + v', w) - (v, w) - (v', w) \\ (v, w + w') - (v, w) - (v, w') \\ (\lambda v, w) - \lambda(v, w) \\ (v, \lambda w) - \lambda(v, w). \end{gathered}$$

Damit setzen wir

$$T := F/U$$

und definieren

$$\begin{aligned} \beta\colon V \times W &\to T \\ (v, w) &\mapsto (v, w) + U. \end{aligned}$$

Nach Definition von U ist β offensichtlich bilinear. Sei nun $\gamma\colon V \times W \to X$ eine weitere bilineare Abbildung. Wir erhalten damit zunächst eine wohldefinierte lineare Abbildung

$$\begin{aligned} \varphi\colon F &\to X \\ (v, w) &\mapsto \gamma(v, w). \end{aligned}$$

Aufgrund der Bilinearität von γ liegt U aber im Kern von φ, also gibt es die wohldefinierte lineare Abbildung

$$\begin{aligned} \bar{\varphi}\colon F/U &\to X \\ (v, w) + U &\mapsto \varphi(v, w) = \gamma(v, w), \end{aligned}$$

die offensichtlich $\gamma = \bar{\varphi} \circ \beta$ erfüllt. Durch diese Bedingung ist aber $\bar{\varphi}$ auch eindeutig bestimmt, da F/U von den Elementen

$$\beta(v, w) = (v, w) + U$$

aufgespannt wird. Also erfüllt T die universelle Eigenschaft des Tensorprodukts von V und W.

Eindeutigkeit: Seien

$$\beta_1\colon V \times W \to T_1 \text{ und } \beta_2\colon V \times W \to T_2$$

zwei Tensorprodukte von V und W. Da β_2 bilinear ist und T_1 die universelle Eigenschaft des Tensorprodukts erfüllt, existiert genau eine lineare Abbildung $\varphi\colon T_1 \to T_2$ mit $\beta_2 = \varphi \circ \beta_1$. Umgekehrt gibt es genau eine lineare Abbildung $\psi\colon T_2 \to T_1$ mit $\beta_1 = \psi \circ \beta_2$:

$$\begin{tikzcd} & T_1 \arrow[dd, bend left, "\varphi"] \\ V \times W \arrow[ur, "\beta_1"] \arrow[dr, "\beta_2"'] & \\ & T_2 \arrow[uu, bend left, "\psi"] \end{tikzcd}$$

Dann ist aber $\psi \circ \varphi \colon T_1 \to T_1$ eine lineare Abbildung mit

$$(\psi \circ \varphi) \circ \beta_1 = \psi \circ (\varphi \circ \beta_1) = \psi \circ \beta_2 = \beta_1,$$

und aus der Eindeutigkeit in der universellen Eigenschaft von T_1 folgt $\psi \circ \varphi = \mathrm{id}_{T_1}$. Genauso folgt $\varphi \circ \psi = \mathrm{id}_{T_2}$, und damit die Aussage. $\square$

Definition 9.1.17. Für K-Vektorräume V, W bezeichnen wir das Tensorprodukt mit

$$V \otimes_K W \quad \text{oder nur} \quad V \otimes W. \qquad \triangle$$

Bemerkung 9.1.18. (i) Die Restklasse $(v, w) + U$ bezeichnen wir auch mit $v \otimes w$, und nennen sie einen **Elementartensor**. Es ist jedoch nicht jedes Element in $V \otimes W$ ein Elementartensor, man muss auch Summen zulassen:

$$V \otimes W = \left\{ \sum_{i=1}^{d} v_i \otimes w_i \mid d \in \mathbb{N}, v_i \in V, w_i \in W \right\}.$$

Nach Konstruktion gelten die folgenden Rechenregeln in $V \otimes W$:

$$\begin{aligned}
(v + v') \otimes w &= v \otimes w + v' \otimes w \\
v \otimes (w + w') &= v \otimes w + v \otimes w' \\
(\lambda v) \otimes w &= \lambda \cdot (v \otimes w) = v \otimes (\lambda w).
\end{aligned}$$

(ii) Will man eine wohldefinierte lineare Abbildung $V \otimes W \to X$ angeben, gibt man gewöhnlich zunächst eine bilineare Abbildung $V \times W \to X$ an, und verwendet die universelle Eigenschaft des Tensorprodukts. Zum Beispiel gilt folgende Aussage: Sind

$$\varphi \colon V \to X \ \text{ und } \ \psi \colon W \to Y$$

linear, so gibt es eine lineare Abbildung

$$\varphi \otimes \psi \colon V \otimes W \to X \otimes Y$$

mit

$$(\varphi \otimes \psi)(v \otimes w) = \varphi(v) \otimes \psi(w).$$

Das sieht man am besten folgendermaßen. Die Abbildung

$$\begin{aligned}
\varphi \times \psi \colon V \times W &\to X \otimes Y \\
(v, w) &\mapsto \varphi(w) \otimes \psi(w)
\end{aligned}$$

ist bilinear. Mit der universellen Eigenschaft von $V \otimes W$ folgt nun die Existenz von $\varphi \otimes \psi$. $\triangle$

Beispiel 9.1.19. In manchen Fällen kann man das Tensorprodukt auch konkreter realisieren als im oberen Beweis.
(i) Es gilt stets
$$K^m \otimes_K K^n \cong \mathrm{Mat}_{m,n}(K) \cong K^{mn}.$$
Dazu betrachtet man die bilineare Abbildung
$$\begin{aligned} K^m \times K^n &\to \mathrm{Mat}_{m,n}(K) \\ (v,w) &\mapsto vw^t = (v_i w_j)_{i,j} \end{aligned}$$
und rechnet die universelle Eigenschaft nach (Übungsaufgabe). Die Elementartensoren sind dabei gerade die Matrizen vom Rang 1. Daraus erhält man sofort die Aussage
$$\dim(V \otimes W) = \dim(V) \cdot \dim(W)$$
für endlich dimensionale Räume.
(ii) Es gilt stets
$$\mathrm{Mat}_m(K) \otimes \mathrm{Mat}_n(K) \cong \mathrm{Mat}_{mn}(K).$$
Man kann dabei einen Elementartensor $M \otimes N$ von der linken Seite mit der Matrix
$$(m_{ij}N)_{i,j} \in \mathrm{Mat}_m(\mathrm{Mat}_n(K)) \cong \mathrm{Mat}_{mn}(K)$$
identifizieren. △

Bemerkung 9.1.20. Über das Tensorprodukt kann man auch sogenannte **Koeffizienten-Erweiterungen** durchführen. Sei dazu zunächst $K \subseteq F$ ein Erweiterung von Körpern. Dann kann man jeden F-Vektorraum natürlich auch als K-Vektorraum auffassen. Umgekehrt kann man einen K-Vektorraum V aber nicht notwendigerweise auch als F-Vektorraum auffassen. Deshalb bemerkt man zunächst, dass F selbst ein K-Vektorraum ist. Die skalare Multiplikation ist dabei einfach die eingeschränkte Körpermultiplikation von F :
$$\cdot\colon K \times F \to F.$$
Dann betrachtet man den K-Vektorraum
$$V_F := V \otimes_K F.$$
Auf V_F ist nun sogar eine Skalarmultiplikation aus F wohldefiniert:
$$\gamma \cdot \sum_i v_i \otimes \gamma_i := \sum_i v_i \otimes \gamma\gamma_i.$$
Auf diese Weise wird V_F zu einem F-Vektorraum. Zum Beispiel gilt
$$\mathbb{R}^n_{\mathbb{C}} = \mathbb{R}^n \otimes_{\mathbb{R}} \mathbb{C} \cong \mathbb{C}^n.$$ △

9.2 Räume von unendlicher Dimension

In diesem Abschnitt betrachten wir Vektorräume, die nicht notwendigerweise endlich dimensional sein müssen. Wir sehen wie man verschiedene Begriffe anpassen kann und welche Aussagen weiter gelten. Sei auch dazu stets K ein fest gewählter Körper.

9.2.1 Algebraische Sichtweise

Definition 9.2.1. Sei V ein K-Vektorraum, I eine beliebige Indexmenge und $\underline{v} = (v_i)_{i\in I}$ eine Familie von Elementen $v_i \in V$.
(i) Die Familie $\underline{v}$ heißt **linear unabhängig**, wenn für jede *endliche* Teilmenge $J \subseteq I$ die Familie $(v_j)_{j\in J}$ linear unabhängig ist (im Sinne von Definition 3.1.13). Ansonsten heißt $\underline{v}$ **linear abhängig**.
(ii) Die Familie $\underline{v}$ heißt **Erzeugendensystem** von V, falls

$$\mathrm{Span}_K\left(\{v_i \mid i \in I\}\right) = V$$

gilt.
(iii) Die Familie $\underline{v}$ heißt **Basis** von V, wenn sie ein linear unabhängiges Erzeugendensystem von V ist. △

Bemerkung 9.2.2. Auch hier werden wir statt der eigentlich exakten Notation von $\underline{v}$ als Familie von Elementen teilweise zur Betrachtung von $\underline{v}$ als Teilmenge von V übergehen. In einer Basis dürfen Elemente sowieso nicht mehrfach auftreten. Solange wir keine Koeffizientenvektoren aufstellen wollen, spielt auch die Reihenfolge der Basiselemente keine wirkliche Rolle. △

Beispiel 9.2.3. (i) Für den Raum $c_{00} := \bigoplus_{\mathbb{N}} \mathbb{R}$ der Folgen mit endlichem Träger ist die Familie

$$\underline{e} = (e_i)_{i\in\mathbb{N}}$$

eine Basis. Dabei sei e_i die Folge mit Eintrag 1 an Position i, und 0 überall sonst. Es handelt sich bei $\underline{e}$ aber *nicht* um eine Basis des deutlich größeren Raums

$$\prod_{i\in\mathbb{N}} \mathbb{R}$$

aller Folgen, noch nicht einmal des Teilraums $\ell^\infty(\mathbb{R})$ aller beschränkten Folgen oder des Teilraums $c_0(\mathbb{R})$ aller Nullfolgen. Jede Linearkombination muss ja stets endlich sein! Es ist übrigens

$$V \cong \mathbb{R}[t]$$

wobei e_i gerade dem Monom t^i entspricht.
(ii) Eine Basis für $\prod_{i\in\mathbb{N}} \mathbb{R}, \ell^\infty(\mathbb{R})$ oder $c_0(\mathbb{R})$ explizit anzugeben ist nicht möglich, man kann nur abstrakt deren Existenz beweisen. Es gibt hier auch keine abzählbare Basis. Das ist einer der Gründe für die algebraisch-analytische Herangehensweise im nächsten Abschnitt. $\triangle$

Bemerkung 9.2.4. (i) Ist $\underline{v}$ eine Basis von V, so lässt sich jedes $v \in V$ auf *eindeutige* Weise als eine Linearkombination

$$v = \sum_{j=1}^{m} \lambda_j v_{i_j}$$

mit $\lambda_j \in K$ schreiben. Es gibt dabei keine Beschränkung an die Länge der Linearkombination, nur endlich muss sie sein.
(ii) Man erhält jeweils eine eindeutige lineare Abbildung $\varphi\colon V \to W$, indem man die Werte auf einer Basis von V beliebig vorgibt. $\triangle$

Um die Existenz von Basen im allgemeinen Fall zu beweisen, benötigen wir das Zorn'sche Lemma. Es handelt sich dabei im wesentlichen um ein Axiom, das wir einfach benutzen und nicht weiter beweisen. Man kann es natürlich auf andere Axiome zurückführen, etwa auf das Auswahlaxiom. Zu diesem ist es sogar äquivalent.

Satz 9.2.5 (Zorn'sches Lemma). *Eine nichtleere partiell geordnete Menge, in der jede Kette eine obere Schranke besitzt, enthält mindestens ein maximales Element.*

Für den Begriff einer *partiellen Ordnung* $\leq$ auf einer Menge M siehe Definition 1.4.6. Eine *Kette* in M ist eine Familie $(m_i)_{i\in I}$ mit $m_i \in M$ für alle $i \in I$, so dass

$$m_i \leq m_j \text{ oder } m_j \leq m_i$$

für alle $i, j \in I$ gilt. Dabei darf I eine beliebige Indexmenge sein. Eine *obere Schranke* einer Kette ist ein $m \in M$ mit $m_i \leq m$ für alle $i \in I$. Ein *maximales Element* von M ist ein Element, welches sich bezüglich $\leq$ nicht überschreiten lässt.

Satz 9.2.6. *Jeder Vektorraum besitzt eine Basis.*

Beweis. Sei V ein beliebiger Vektorraum. Wir betrachten die Menge

$$\mathcal{L} := \{L \subseteq V \mid L \text{ linear unabhängig}\}$$

und wollen das Zorn'sche Lemma darauf anwenden. Dazu versehen wir $\mathcal{L}$ mit der Mengeninklusion als partielle Ordnung. Wegen $\emptyset \in \mathcal{L}$ ist $\mathcal{L}$ nicht leer. Sei nun $(L_i)_{i\in I}$ eine Kette in $\mathcal{L}$. Wir setzen

$$L := \bigcup_{i\in I} L_i$$

und erhalten, vorausgesetzt dass L in $\mathcal{L}$ liegt, damit offensichtlich eine obere Schranke der Kette. Um nun $L \in \mathcal{L}$ zu überprüfen, sei $L' \subseteq L$ endlich. Wir müssen zeigen, dass L' linear unabhängig ist. Aus der Endlichkeit von L' folgt

$$L' \subseteq L_{i_1} \cup \cdots \cup L_{i_n}$$

für gewisse $i_1, \ldots, i_n \in I$. Aus der Kettenbedingung folgt nun o.B.d.A.

$$L_{i_1} \subseteq \cdots \subseteq L_{i_n}$$

und damit $L' \subseteq L_{i_n}$. Wegen $L_{i_n} \in \mathcal{L}$ ist damit L' linear unabhängig, die gewünschte Aussage.
Wir haben nun die Voraussetzungen des Zorn'schen Lemmas überprüft und erhalten die Existenz eines maximalen Elements in $\mathcal{L}$, also einer maximalen linear unabhängigen Teilmenge $B \subseteq V$. Für $v \in V \setminus B$ ist dann $B \cup \{v\}$ linear abhängig, also gibt es eine endliche Teilmenge $B' \subseteq B$, so dass bereits $B' \cup \{v\}$ linear abhängig ist. Eine nichttriviale Linearkombination der Elemente zu Null muss aber v benutzen, denn B' ist linear unabhängig. Durch Auflösen solch einer Linearkombination nach v erhalten wir $v \in \mathrm{Span}(B') \subseteq \mathrm{Span}(B)$. Also ist B auch ein Erzeugendensystem von V, also eine Basis. □

Bemerkung 9.2.7. Auch im unendlich-dimensionalen Fall kann man einen Basisergänzungssatz beweisen. Es ist auch wahr, dass je zwei Basen von V dieselbe Mächtigkeit besitzen, es also eine Bijektion zwischen ihnen gibt. △

Korollar 9.2.8. *Jeder K-Vektorraum ist isomorph zu*

$$\bigoplus_I K$$

für eine Menge I.

Beweis. Wenn $\underline{v} = (v_i)_{i\in I}$ eine Basis von V ist, kann man jedes $v \in V$ eindeutig als

$$v = \sum_{j\in J} \lambda_j v_j$$

mit $J \subseteq I$ endlich schreiben. Wir definieren dann

$$\begin{aligned} \xi_{\underline{v}} \colon V &\to \bigoplus_I K \\ v &\mapsto \sum_{j \in J} \lambda_j e_j \end{aligned}$$

und erhalten so offensichtlich einen Isomorphismus. $\square$

Bemerkung 9.2.9. Viele der Aussagen von früher, die im Beweis essentiell den Dimensionsbegriff verwendet haben, stimmen im unendlich dimensionalen nicht mehr. Auf $\ell^\infty(\mathbb{C})$ gibt es beispielsweise den Rechts- und Linksshift:

$$\begin{aligned} \rho \colon (a_0, a_1, a_2, \dots) &\mapsto (0, a_0, a_1, \dots) \\ \lambda \colon (a_0, a_1, a_2 \dots) &\mapsto (a_1, a_2, a_3, \dots). \end{aligned}$$

Es ist ρ injektiv aber nicht surjektiv, λ surjektiv aber nicht injektiv. Es gilt

$$\lambda \circ \rho = \mathrm{id} \quad \text{und} \quad \rho \circ \lambda \neq \mathrm{id}\,.$$

Auch Begriffe wie Darstellungsmatrix, Determinante, Spur, charakteristisches Polynom etc... lassen sich im Allgemeinen nicht oder nur sehr abgewandelt definieren. $\triangle$

9.2.2 Algebraisch-analytische Sichtweise

Schon bei sehr einfach zu definierende Räumen, etwa Folgenräumen, kann man keine Basis explizit angeben, meistens gibt es noch nicht einmal eine abzählbare Basis. Deshalb wird die rein algebraische Sichtweise schnell zu kompliziert. Man bringt nun eine zusätzliche analytische Struktur ins Spiel, d.h. man fordert zumindest die Existenz einer Norm $\|\cdot\|$ auf dem Vektorraum V (siehe Definition 6.1.8). Man beachte, dass wir uns dazu auf $K = \mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ beschränken. Die Norm induziert eine Metrik

$$d(v, w) := \|v - w\|$$

und wir können also über Konvergenz etc... in V sprechen.

Definition 9.2.10. (i) Ein **Banachraum** ist ein normierter $\mathbb{K}$-Vektorraum, der bezüglich der induzierten Metrik vollständig ist (d.h. jede Cauchy-Folge besitzt in V einen Grenzwert).
(ii) Ein **Hilbertraum** ist $\mathbb{K}$-Vektorraum mit Skalarprodukt, der bezüglich der induzierten Norm ein Banachraum ist. $\triangle$

Beispiel 9.2.11. (i) Es ist $\mathbb{K}^m$ ein Hilbertraum bezüglich des Standardskalarprodukts. Das folgt sofort aus der Vollständigkeit von $\mathbb{K}$.
(ii) Der Raum $\ell^\infty(\mathbb{K})$ aller beschränkten $\mathbb{K}$-wertigen Folgen ist ein Banachraum, bezüglich der Supremumsnorm

$$\| (a_i)_{i\in\mathbb{N}} \|_\infty := \sup_{i\in\mathbb{N}} |a_i|.$$

Dasselbe stimmt für den Raum $c_0(\mathbb{K})$ aller Nullfolgen und den Raum $c(\mathbb{K})$ aller konvergenten Folgen.
(iii) Der Raum

$$\ell^2(\mathbb{K}) := \left\{ (a_i)_{i\in\mathbb{N}} \mid a_i \in \mathbb{K}, \sum_{i\in\mathbb{N}} |a_i|^2 < \infty \right\}$$

ist ein Hilbertraum, bezüglich des Skalarprodukts

$$\langle (a_i)_{i\in\mathbb{N}}, (b_i)_{i\in\mathbb{N}} \rangle := \sum_{i\in\mathbb{N}} a_i \overline{b_i}.$$

(iv) Für jeden kompakten Hausdorffraum X ist

$$C(X, \mathbb{K})$$

ein Banachraum bezüglich der Supremumsnorm

$$\|f\|_\infty := \sup_{x\in X} |f(x)|.$$

(v) Der Vektorraum $c_{00}(\mathbb{K})$ aller Folgen mit endlichem Träger ist *kein* Banachraum (bezüglich der Supremumsnorm), denn er ist nicht vollständig. △

Definition 9.2.12. (i) Sei V ein Banachraum. Eine Folge $(v_i)_{i\in\mathbb{N}}$ in V heißt **Schauderbasis von** V, falls für jedes $v \in V$ ein eindeutig bestimmte Folge $(\lambda_i)_{i\in\mathbb{N}}$ von Skalaren existiert, mit

$$v = \sum_{i=0}^{\infty} \lambda_i v_i.$$

Man beachte dass wir dabei eine Reihenfolge fixiert haben, der Wert der Reihe also nicht unabhängig von der Summationsreihenfolge sein muss.
(ii) Sei V ein Hilbertraum. Eine Folge $(v_i)_{i\in\mathbb{N}}$ in V heißt **Hilbertbasis von** V, wenn sie eine maximale orthonormale Menge von Vektoren bildet. △

Bemerkung 9.2.13. (i) Jede Schauderbasis von V ist offensichtlich linear unabhängig im algebraisches Sinn des letzten Abschnitts. Das folgt aus der geforderten Eindeutigkeit in der Basisdarstellung. Im algebraischen Sinn spannt sie aber nur einen *dichten* Unterraum von V auf.
(ii) Jede Hilbertbasis eines Hilbertraums V ist eine Schauderbasis von V, wenn man ihn als Banachraum betrachtet. Man kann nämlich zeigen, dass für $v \in V$ stets

$$v = \sum_{i \in \mathbb{N}} \langle v, v_i \rangle v_i$$

gilt, sogar unabhängig von der Summationsreihenfolge. △

Beispiel 9.2.14. (i) Die Familie $(e_i)_{i\in\mathbb{N}}$ ist eine (normale algebraische) Basis für den Raum $c_{00}(\mathbb{K})$ und eine Schauderbasis für den Raum $c_0(\mathbb{K})$ aller Nullfolgen, bezüglich Supremumsnorm. Der Raum $\ell^\infty(\mathbb{K})$ aller beschränkten Folgen besitzt keine Schauderbasis.
(ii) Die Familie $(e_i)_{i\in\mathbb{N}}$ ist eine Hilbertbasis für $\ell^2(\mathbb{K})$. △

Bemerkung 9.2.15. Die meisten unserer bisherigen Konzepte werden mit leichten Anpassungen ins algebraisch-analytische Setup übernommen. Von linearen Abbildungen zwischen Banachräumen fordert man zum Beispiel zusätzlich oft Stetigkeit (die nicht automatisch aus der Linearität folgen muss). Diese Theorie der linearen Algebra in unendlich-dimensionalen Räumen mit Skalarprodukten/Normen nennt man *Funktionalanalysis*. △

9.3 Kategorientheorie

Ganz zum Schluss lernen wir noch eine Sichtweise kennen, anhand derer man viele der bisherigen Konzepte, und viele Konzepte der Mathematik überhaupt, besser einordnen und verstehen kann.

Definition 9.3.1. (i) Eine **Kategorie** $\mathcal{C}$ besteht aus einer Klasse

$$\mathrm{Obj}(\mathcal{C})$$

(von sogenannten **Objekten**) und für alle $X, Y \in \mathrm{Obj}(\mathcal{C})$ jeweils einer Menge

$$\mathcal{C}(X, Y)$$

(von sogenannten **Morphismen**), sowie einer partiellen Verknüpfung von Morphismen

$$\begin{aligned} \mathcal{C}(X,Y) \times \mathcal{C}(Y,Z) &\to \mathcal{C}(X,Z) \\ (f,g) &\mapsto g \circ f \end{aligned}$$

welche folgenden zwei Bedingungen genügt:

(1) $\forall X \in \mathrm{Obj}(\mathcal{C})\ \exists \mathrm{id}_X \in \mathcal{C}(X,X)$ mit

$$\mathrm{id}_X \circ f = f,\ g \circ \mathrm{id}_X = g$$

für alle $f \in \mathcal{C}(Y,X), g \in \mathcal{C}(X,Y)$.

(2) Für alle $f \in \mathcal{C}(W,X), g \in \mathcal{C}(X,Y), h \in \mathcal{C}(Y,Z)$ gilt

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

(*ii*) Ein $f \in \mathcal{C}(X,Y)$ heißt **Isomorphismus**, falls $g \in \mathcal{C}(Y,X)$ existiert mit

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

(*iii*) Zwei Objekte $X, Y \in \mathrm{Obj}(\mathcal{C})$ heißen **isomorph** (in Zeichen $X \cong Y$), falls in $\mathcal{C}(X,Y)$ ein Isomorphismus existiert. △

Grafisch stellt man Ausschnitte aus Kategorien gewöhnlich durch (kommutative) Pfeildiagramme dar:

$$\begin{tikzcd} X \arrow[r, "f"] \arrow[dr, "g\circ f"'] & Y \arrow[d, "g"] \\ & Z \end{tikzcd}$$

Beispiel 9.3.2. (*i*) Die **Kategorie** $\mathcal{M}\mathrm{en}$ **der Mengen** hat als Objekte die Mengen und als Morphismen die Abbildungen, wobei die Verknüpfung gerade die Hintereinanderausführung ist. Ein Isomorphismus ist eine bijektive (also invertierbare) Abbildung, und zwei Mengen sind isomorph wenn sie die gleiche Mächtigkeit besitzen.
(*ii*) Für jeden festen Körper K hat die **Kategorie** K-$\mathcal{V}ec$ **der** K**-Vektorräume** als Objekte die K-Vektorräume und als Morphismen die K-linearen Abbildungen.

(iii) Die **Kategorie** $\mathcal{T}\mathrm{op}$ **der topologischen Räume** hat als Objekte die topologischen Räume und als Morphismen die stetigen Abbildungen. Ein Isomorphismus ist gerade ein Homöomorphismus, und isomorphe Objekte sind homöomorphe Räume.
(iv) Morphismen müssen nicht immer Abbildungen sein. Sei etwa $(G, \cdot)$ eine feste Gruppe. Wir definieren eine Kategorie $\mathcal{C}_G$ durch

$$\mathrm{Obj}(\mathcal{C}_G) := \{*\}$$

und

$$\mathcal{C}_G(*, *) := G.$$

Die Gruppenverknüpfung $\cdot$ dient uns dabei als Verknüpfung von Morphismen

$$\cdot\colon \mathcal{C}_G(*, *) \times \mathcal{C}_G(*, *) \to \mathcal{C}_G(*, *)$$

und man überprüft leicht die Bedingungen (1) und (2). Auf diese Weise lässt sich G als eigene Kategorie auffassen. Jeder Morphismus ist hier ein Isomorphismus. $\triangle$

Definition 9.3.3. Ein **kovarianter** (bzw. **kontravarianter**) **Funktor** von der Kategorie $\mathcal{C}$ in die Kategorie $\mathcal{D}$ besteht aus einer Abbildung

$$\mathcal{F}\colon \mathrm{Obj}(\mathcal{C}) \to \mathrm{Obj}(\mathcal{D})$$

sowie Abbildungen

$$\mathcal{F}\colon \mathcal{C}(X, Y) \to \mathcal{D}(\mathcal{F}(X), \mathcal{F}(Y))$$

$$(\text{bzw. } \mathcal{F}\colon \mathcal{C}(X, Y) \to \mathcal{D}(\mathcal{F}(Y), \mathcal{F}(X)))$$

mit

(1) $\mathcal{F}(\mathrm{id}_X) = \mathrm{id}_{\mathcal{F}(X)}$

(2) $\mathcal{F}(g \circ f) = \mathcal{F}(g) \circ \mathcal{F}(f)$ (bzw. $\mathcal{F}(g \circ f) = \mathcal{F}(f) \circ \mathcal{F}(g)$).

$$\begin{tikzcd} X \arrow[r, "f"] \arrow[dr, "g\circ f"'] & Y \arrow[d, "g"] \\ & Z \end{tikzcd} \qquad\qquad \begin{tikzcd} \mathcal{F}(X) \arrow[r, "\mathcal{F}(f)"] \arrow[dr, "\mathcal{F}(g\circ f)"'] & \mathcal{F}(Y) \arrow[d, "\mathcal{F}(g)"] \\ & \mathcal{F}(Z) \end{tikzcd}$$

$\triangle$

Beispiel 9.3.4. (i) Die Bildung des Dualraums ist ein kontravarianter Funktor von K-$\mathcal{V}ek$ in sich selbst.
(ii) Sei X ein fest gewählter K-Vektorraum. Dann ist die Zuordnung

$$V \mapsto X \otimes V$$

ein kovarianter Funktor von K-$\mathcal{V}ek$ in sich selbst. Eine lineare Abbildung $\varphi\colon V \to W$ wird dabei auf

$$\mathrm{id}_X \otimes \varphi\colon X \otimes V \to X \otimes W$$

abgebildet.
(iii) Es gibt den kovarianten Funktor $\mathcal{T}\mathrm{op} \to \mathcal{M}\mathrm{en}$, der jedem topologischen Raum die Menge seiner Zusammenhangskomponenten zuordnet. Stetige Abbildungen erhalten Zusammenhang, also induzieren sie Abbildungen zwischen den Mengen der Zusammenhangskomponenten.
(iv) Sei $\mathcal{C}$ eine Kategorie, in der die Objekte wirklich aus Mengen und die Morphismen wirklich aus Abbildungen zwischen diesen bestehen (z.B. $\mathcal{T}\mathrm{op}$, K-$\mathcal{V}ec, \ldots$). Der **Vergiss-Funktor** ist ein kovarianter Funktor $\mathcal{C} \to \mathcal{M}\mathrm{en}$, der einfach eine eventuelle zusätzliche Struktur auf den Objekten der Kategorie (z.B. eine Topologie, eine Vektorraumstruktur,...) vergisst. Ebenso vergisst er die Tatsache, dass Morphismen eventuell sehr spezielle Abbildungen sind, und betrachtet sie einfach nur noch als Abbildungen. △

Lemma 9.3.5. *Sei $\mathcal{F}\colon \mathcal{C} \to \mathcal{D}$ ein Funktor und in $\mathcal{C}$ gelte $X \cong Y$. Dann gilt in $\mathcal{D}$*

$$\mathcal{F}(X) \cong \mathcal{F}(Y).$$

Beweis. Sei o.B.d.A. $\mathcal{F}$ kovariant. Sei $f \in \mathcal{C}(X,Y)$ ein Isomorphismus mit inversem Morphismus $g \in \mathcal{C}(Y,X)$. Dann gilt $g \circ f = \mathrm{id}_X$ und nach Anwendung von $\mathcal{F}$ also

$$\mathrm{id}_{\mathcal{F}(X)} = \mathcal{F}(\mathrm{id}_X) = \mathcal{F}(g \circ f) = \mathcal{F}(g) \circ \mathcal{F}(f).$$

Analog bekommt man $\mathcal{F}(f) \circ \mathcal{F}(g) = \mathrm{id}_{\mathcal{F}(Y)}$ und damit $\mathcal{F}(X) \cong \mathcal{F}(Y)$. □

Bemerkung 9.3.6. Oft möchte man die Isomorphie von zwei Objekten einer Kategorie entscheiden (sind zwei Vektorräume isomorph, sind zwei topologische Räume homöomorph,....?) Wenn sie isomorph sind, kann man oft einen Isomorphismus angeben und hat die Frage damit entschieden. Wenn sie nicht isomorph sind, ist die Frage oft schwieriger. Man muss dann Eigenschaften der Objekte finden, die sie voneinander unterscheiden und die Isomorphie ausschließen. Etwas

konzeptioneller formuliert wendet man zunächst einen Funktor an. Sind die Objekte dann nicht isomorph, waren sie es vorher auch nicht.
Haben beispielsweise zwei topologische Räume eine unterschiedliche Anzahl von Zusammenhangskomponenten, können sie nicht homöoorph sein. Diese Argumentation entspricht der Anwendung des Funktors aus Beispiel 9.3.4 (iii). Sind die Dualräume von zwei Vektorräumen nicht isomorph, so sind es die Räume selbst auch nicht. Das entspricht dem Funktor aus Beispiel 9.3.4 (i). Noch banaler ist folgende Beobachtung: haben zwei Vektorräume nicht dieselbe Mächtigkeit (als Mengen), so sind sie nicht isomorph. Extrem abgehoben kann man sagen, dass man diese Beobachtung aus dem Vergissfunktor und Lemma 9.3.5 erhält. $\triangle$

Literaturverzeichnis

[1] Siegfried Bosch, *Lineare Algebra*, Springer, 3. Auflage, 2006.

[2] Gerd Fischer, *Lineare Algebra*, Springer, 18. Auflage, 2014.

[3] Serge Lang, *Linear Algebra*, Springer, Third Edition, 1987.

Übungsaufgaben

Aufgabe 1. Für eine Menge M nennen wir die Menge aller ihrer Teilmengen die *Potenzmenge* $\mathcal{P}(M)$ von M. In Formeln:

$$\mathcal{P}(M) := \{N \mid N \subseteq M\}.$$

Schreiben Sie für die folgenden Mengen M jeweils die Potenzmenge $\mathcal{P}(M)$ explizit durch Angabe aller Elemente auf:
(i) $M = \{1, a, \varphi\}$.
(ii) $M = \mathcal{P}(\mathcal{P}(\emptyset))$.
(iii) $M = \{1, 2\} \times \{a, b\}$.
Wieviele Elemente hat die Potenzmenge einer m-elementigen Menge?

Aufgabe 2. Für Mengen M, N definieren wir

$$M \Delta N := (M \cup N) \setminus (M \cap N).$$

Zeigen Sie, dass für beliebige Mengen M, N, O stets gilt:
(i) $(M \Delta N) \Delta O = M \Delta (N \Delta O)$.
(ii) $M \Delta N = N \Delta M$.
(iii) $M \Delta N = \emptyset \Leftrightarrow M = N$.

Aufgabe 3. (i) Sei M die Menge aller Menschen. Untersuchen Sie die folgenden Relationen auf M in Hinblick auf Symmetrie, Reflexivität und Transitivität:

1. $R_1 = \{(A, B) \mid A \text{ liebt } B\}$
2. $R_2 = \{(A, B) \mid A \text{ ist mit } B \text{ verheiratet}\}$
3. $R_3 = \{(A, B) \mid A \text{ ist Mutter von } B\}$
4. $R_4 = \{(A, B) \mid \exists C \in M \colon (C, A) \in R_3 \land (C, B) \in R_3\}$

(ii) Geben Sie für jede der drei Eigenschaften *reflexiv, symmetisch, transitiv* eine Relation auf einer Menge an, die diese Eigenschaft hat, die anderen beiden jedoch nicht.

Aufgabe 4. Sei $n \in \mathbb{N}$ gegeben. Auf $\mathbb{Z}$ definieren wir folgende Relation:

$$M_n := \{(a, b) \in \mathbb{Z} \times \mathbb{Z} \mid \exists c \in \mathbb{Z} \colon nc = b - a\}.$$

Zeigen Sie, dass es sich bei M_n um eine Äquivalenzrelation handelt und bestimmen Sie die Anzahl ihrer Äquivalenzklassen.

Aufgabe 5. Seien $f \colon M \to N, g \colon N \to O, h \colon O \to P$ Abbildungen. Zeigen Sie:
(i) Es gilt $h \circ (g \circ f) = (h \circ g) \circ f$.
(ii) Sind f und g injektiv, so auch $g \circ f$.
(iii) Sind f und g surjektiv, so auch $g \circ f$.
(iv) Sind f und g bijektiv, so auch $g \circ f$ und es gilt $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.

Aufgabe 6. Untersuchen Sie die folgenden Abbildungen auf Injektivität und Surjektivität:
(i) $f_1 \colon \mathbb{R}^2 \to \mathbb{R}; (a, b) \mapsto a + b$.
(ii) $f_2 \colon \mathbb{R}^2 \to \mathbb{R}; (a, b) \mapsto a^2 + b^2 - 1$.
(iii) $f_3 \colon \mathbb{R}^2 \to \mathbb{R}^2; (a, b) \mapsto (a + 2b, 2a - b)$.
(iv) $f_4 \colon \mathbb{R}^2 \to \mathbb{R}^3; (a, b) \mapsto (a^2, a - b, b)$.

Aufgabe 7. Sei $f \colon M \to N$ eine Abbildung. Zeigen Sie:
(i) Für $T \subseteq M$ und $S \subseteq N$ gilt $T \subseteq f^{-1}(f(T))$ sowie $f(f^{-1}(S)) \subseteq S$.
(ii) Für $S_1, S_2 \subseteq N$ gilt $f^{-1}(S_1 \cup S_2) = f^{-1}(S_1) \cup f^{-1}(S_2)$ sowie $f^{-1}(S_1 \cap S_2) = f^{-1}(S_1) \cap f^{-1}(S_2)$.
(iii) Für $T_1, T_2 \subseteq M$ gilt $f(T_1 \cup T_2) = f(T_1) \cup f(T_2)$.
(iv) Zeigen oder widerlegen Sie die Aussage $f(T_1 \cap T_2) = f(T_1) \cap f(T_2)$.

Aufgabe 8. Bestimmen Sie die (reelle) Lösungsmenge des linearen Gleichungssystems mit folgender erweiterter Koeffizientenmatrix:

$$(A, b) = \left(\begin{array}{ccccc|c} 1 & 3 & 2 & 5 & 4 & 0 \\ 0 & 0 & -2 & 0 & 4 & 1 \\ 0 & 0 & 0 & -10 & 2 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

Aufgabe 9. (i) Bestimmen Sie die reelle Lösungsmenge des folgenden linearen Gleichungssystems:

$$\begin{array}{rcrcrcrcl} x_1 & + & 2x_2 & + & 3x_3 & + & 4x_4 & = & 0 \\ -3x_1 & + & -6x_2 & + & -8x_3 & + & -9x_4 & = & 0 \\ 2x_1 & & & + & x_3 & + & 9x_4 & = & 0. \end{array}$$

(ii) Für welche $\lambda \in \mathbb{R}$ besitzt das folgende Gleichungssystem eine reelle Lösung? Wieviele Lösungen gibt es dann jeweils?

$$\begin{array}{rcrcrcl} 2x_1 & - & 2x_2 & + & \lambda x_3 & = & 3 \\ 4x_1 & + & 6x_2 & - & 3x_3 & = & -2 \\ 10x_1 & - & 10x_2 & + & 13x_3 & = & 0. \end{array}$$

Aufgabe 10. Finden Sie ein reelles Polynom vom Grad höchstens 4, also einen Ausdruck der Gestalt

$$p = c_0 + c_1 t + c_2 t^2 + c_3 t^3 + c_4 t^4$$

mit $c_0, \ldots, c_4 \in \mathbb{R}$, für das gilt:

$$p(-2) = 1,\ p(-1) = -1,\ p(0) = 2,\ p(1) = 1,\ p(2) = -1.$$

Aufgabe 11. Sei $(G, *)$ eine Gruppe. Das zu $a \in G$ inverse Element bezeichnen wir mit a^{-1}. Zeigen Sie, dass für $a, b, c \in G$ stets folgendes gilt:
(i) $(a * b)^{-1} = b^{-1} * a^{-1}$.
(ii) $(a^{-1})^{-1} = a$.
(ii) $a * c = b * c \ \Rightarrow\ a = b$.

Aufgabe 12. Sei $G = \{x, y\}$ eine zweielementige Menge. Bestimmen Sie alle Verknüpfungen

$$*\colon G \times G \to G,$$

welche G zu einer Gruppe machen (zum Beispiel durch Angabe der Verknüpfungstafeln).

Aufgabe 13. Sei M eine nichtleere Menge. Für $A, B \subseteq M$ definieren wir $A \Delta B$ wie in Aufgabe 2, sowie

$$A \circ B := M \setminus (A \Delta B)\,.$$

Entscheiden Sie:
(i) Sind $(\mathcal{P}(M), \Delta)$ und/oder $(\mathcal{P}(M), \circ)$ Gruppen?
(ii) Sind $(\mathcal{P}(M), \Delta, \cap)$ und/oder $(\mathcal{P}(M), \Delta, \cup)$ Ringe?

Aufgabe 14. Sei R ein Ring. Zeigen Sie, dass für alle $a, b, c \in R$ gilt:
(*i*) $0 \cdot a = a \cdot 0 = 0$.
(*ii*) $-(a \cdot b) = (-a) \cdot b = a \cdot (-b)$.
(*iii*) $(-a) \cdot (-b) = a \cdot b$.
Gilt auch immer $a \cdot b = a \cdot c \Rightarrow b = c$?

Aufgabe 15. Vervollständigen Sie den Beweis von Satz 2.2.13.

Aufgabe 16. Bestimmen Sie die Lösungsmenge (in $\mathbb{C}^2$) des folgenden linearen Gleichungssystems:

$$\begin{aligned} i \cdot x + y &= -i \\ x + i \cdot y &= 1 + i. \end{aligned}$$

Aufgabe 17. Lösen Sie das folgende lineare Gleichungssystem über dem Körper $\mathbb{Z}/3\mathbb{Z}$:

$$\begin{array}{ccccccccccc} x_1 & & & + & 2x_3 & + & x_4 & & & = & 2 \\ x_1 & + & x_2 & + & 2x_3 & & & + & x_5 & = & 0 \\ & & 2x_2 & + & x_3 & + & x_4 & + & 2x_5 & = & 1 \end{array}$$

Aufgabe 18. Zeigen Sie Assoziativ- und Distributivgesetze für die Matrixrechnung (d.h. die Aussagen von Satz 2.3.5 (*ii*) und (*iii*)).

Aufgabe 19. Bestimmen Sie alle Matrizen $X \in \mathrm{Mat}_{3,2}(\mathbb{C})$ mit

$$\begin{pmatrix} 1 & i & 0 \\ -i & -1 & 1 \\ 0 & 2 & 0 \end{pmatrix} \cdot X = \begin{pmatrix} 1 & -1 \\ i & -i \\ 0 & 1 \end{pmatrix}.$$

Aufgabe 20. Sei R ein Ring und $m \in \mathbb{N}$. Eine Matrix $M = (m_{ij})_{i,j=1,\ldots,m} \in \mathrm{Mat}_m(R)$ hat *obere Dreiecksgestalt*, wenn für $j < i$ stets $m_{ij} = 0$ gilt. Zeigen Sie: Haben $M, N \in \mathrm{Mat}_m(R)$ obere Dreiecksgestalt, so auch $M \cdot N$.

Aufgabe 21. Berechnen Sie die inverse Matrix der beiden folgenden Matrizen, jeweils im gegebenen Matrixring:

$$\begin{pmatrix} 1+i & -i & 1 \\ 2 & 1-2i & 0 \\ 0 & -1 & 1 \end{pmatrix} \in \mathrm{Mat}_3(\mathbb{C}) \text{ und } \begin{pmatrix} 1 & 2 & 4 \\ 2 & 3 & 6 \\ 5 & 0 & 1 \end{pmatrix} \in \mathrm{Mat}_3(\mathbb{Z}/7\mathbb{Z}).$$

Aufgabe 22. Sei K ein Körper und $A, B \in \mathrm{Mat}_m(K)$. Es gelte $AB = I_m$. Zeigen Sie, dass dann auch $BA = I_m$ gilt.
Hinweis: Wenn Sie die Invertierbarkeit von A oder B verwenden wollen, müssen Sie sie zuerst zeigen.

Aufgabe 23. Sei K ein Körper und V ein K-Vektorraum. Zeigen Sie, dass für alle $\lambda, \gamma \in K, v \in V$ gilt:
(i) $0 \cdot v = 0$.
(ii) $(-\lambda) \cdot v = -(\lambda \cdot v)$.
(iii) Aus $v \neq 0$ und $\lambda \neq \gamma$ folgt $\lambda \cdot v \neq \gamma \cdot v$.

Aufgabe 24. Sei $V = \mathcal{F}(\mathbb{R}, \mathbb{R})$ der $\mathbb{R}$-Vektorraum aller Abbildungen von $\mathbb{R}$ nach $\mathbb{R}$. Welche der folgenden Teilmengen sind Untervektorräume von V? (begründen Sie Ihre Aussage)
(i) $U_1 = \{f \in V \mid f(0) = 0\}$.
(ii) $U_2 = \{f \in V \mid f(0) = 1\}$.
(iii) $U_3 = \{f \in V \mid f \text{ stetig}\}$.
(iv) $U_4 = \{f \in V \mid \forall r \in \mathbb{R}\colon f(r) \geq 0\}$.
(v) $U_5 = \{f \in V \mid \forall r \in \mathbb{R}\colon |f(r)| \leq 1\}$.
(vi) $U_6 = \{f \in V \mid \exists C \in \mathbb{R}\, \forall r \in \mathbb{R}\colon |f(r)| \leq C\}$.

Aufgabe 25. Berechnen Sie zur gegebenen Matrix $A \in \mathrm{Mat}_3(\mathbb{Q})$ eine Zerlegung als Produkt von Elementarmatrizen:

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 0 \end{pmatrix}.$$

Was ändert sich, wenn unten rechts in A statt 0 eine 9 steht?

Aufgabe 26. Bestimmen Sie alle Untervektorräume von $\mathbb{C}$, wobei Sie $\mathbb{C}$ einmal als $\mathbb{C}$- und einmal als $\mathbb{R}$-Vektorraum auffassen (die Skalarmultiplikation ist einfach die bekannte Multiplikation von Zahlen).

Aufgabe 27. Sei V ein K-Vektorraum und $v_1, \ldots v_m \in V$ linear unabhängig. Zeigen Sie, dass für $v \in V \setminus \mathrm{Span}_K(\{v_1, \ldots, v_m\})$ auch

$$v_1, \ldots, v_m, v$$

linear unabhängig sind.

Aufgabe 28. Sei $V = \mathcal{F}(\mathbb{R}, \mathbb{R})$ und $n \geq 1$. Zeigen Sie, dass die Funktionen

$$\sin(x), \sin(2x), \ldots, \sin(nx)$$

als Elemente von V $\mathbb{R}$-linear unabhängig sind.

Aufgabe 29. Bestimmen Sie für die folgende Matrix $A \in \mathrm{Mat}_{3,5}(\mathbb{Q})$ eine Basis des Lösungsraums $L(A, 0) \subseteq \mathbb{Q}^5$:

$$A = \begin{pmatrix} 1 & -2 & 1 & 2 & 0 \\ 1 & 1 & -1 & 1 & 2 \\ 1 & 7 & -5 & -1 & 1 \end{pmatrix}.$$

Aufgabe 30. Bestimmen Sie für die folgenden Unterräume von $\mathbb{R}^4$ jeweils eine Basis und die Dimension:
(i) $U_1 = \{(a_1, a_2, a_3, a_4) \in \mathbb{R}^4 \mid a_1 = 0\}$.
(ii) $U_2 = \{(a_1, a_2, a_3, a_4) \in \mathbb{R}^4 \mid a_1 = a_2 + a_3 + a_4\}$.
(iii) $U_1 \cap U_2, U_1 + U_2$.

Aufgabe 31. Wir betrachten folgende Menge:

$$M := \{A \in \mathrm{Mat}_3(\mathbb{R}) \mid \text{jede Zeile und jede Spalte von } A \text{ summiert sich zu } 1\}.$$

(i) Zeigen Sie, dass M ein affiner Unterraum von $\mathrm{Mat}_3(\mathbb{R})$ ist und bestimmen Sie den zu M parallelen Untervektorraum $U \subseteq \mathrm{Mat}_3(\mathbb{R})$.
(ii) Bestimmen Sie eine Basis von U.

Aufgabe 32. Sei $\varphi\colon V \to W$ eine lineare Abbildung zwischen K-Vektorräumen. Entscheiden Sie, welche der folgenden Aussagen wahr sind:
(i) Sind $v_1, \ldots, v_n$ in V linear unabhängig, so sind $\varphi(v_1), \ldots, \varphi(v_n)$ in W linear unabhängig.
(ii) Sind $\varphi(v_1), \ldots, \varphi(v_n)$ in W linear unabhängig, so sind $v_1, \ldots, v_n$ in V linear unabhängig.

Aufgabe 33. Bestimmen Sie für die folgenden linearen Abbildungen jeweils eine Basis für Kern und Bild:
(i) $\varphi_1\colon \mathbb{R}^2 \to \mathbb{R}^2;\ (x_1, x_2) \mapsto (x_1 - x_2, x_2 + 2x_1)$.
(ii) $\varphi_2\colon \mathbb{R}^2 \to \mathbb{R}^3;\ (x_1, x_2) \mapsto (x_1 + x_2, x_1 - 2x_2, 2x_2 - x_1)$.
(iii) $\varphi_3\colon \mathbb{R}^3 \to \mathbb{R}^3;\ (x_1, x_2, x_3) \mapsto (7x_1 + x_2, x_1 + x_2 + x_3, x_3 - 6x_1)$.

Aufgabe 34. Sei $V = \mathbb{R}[t]_{\leq 4}$ der Vektorraum aller reellen Polynome vom Grad ≤ 4.
(i) Wählen Sie eine Basis $\underline{v}$ von V und bestimmen Sie den Koordinatenvektor $\xi_{\underline{v}}(p_i)$ für die folgenden Polynome:

$$p_1 = 1 - t^2 + 2t^3,\ p_2 = t^4,\ p_3 = (t-2)^2 - t^3.$$

(ii) Entscheiden Sie, welche der folgenden Abbildungen $\varphi\colon V \to V$ linear sind:

(a) $p \mapsto p(0)$ (Auswertung in 0)

(b) $p \mapsto p(1)$ (Auswertung in 1)

(c) $p \mapsto p'$ (Ableitung nach t)

(d) $p \mapsto p + 1$.

(iii) Finden Sie zu den linearen Abbildungen φ aus (ii) jeweils eine Matrix $A \in \mathrm{Mat}_5(\mathbb{R})$, so dass

$$\xi_{\underline{v}}(\varphi(p)) = A\xi_{\underline{v}}(p)$$

für alle $p \in V$ gilt.

Aufgabe 35. Seien V, W zwei K-Vektorräume und $\varphi\colon V \to W$ eine bijektive lineare Abbildung. Zeigen Sie:
(i) Die Umkehrabbildung $\varphi^{-1}\colon W \to V$ ist ebenfalls linear.
(ii) Es gilt $\dim_K(V) = \dim_K(W)$.

Aufgabe 36. Sei V ein K-Vektorraum und $\varphi\colon V \to V$ eine lineare Abbildung. Wir schreiben φ^n für die n-malige Hintereinanderausführung von φ, d.h.

$$\varphi^n = \underbrace{\varphi \circ \cdots \circ \varphi}_{n}.$$

Seien $v \in V$ und $n \in \mathbb{N}$ so, dass gilt:

$$\varphi^n(v) \neq 0 \text{ und } \varphi^{n+1}(v) = 0.$$

Zeigen Sie, dass dann $v, \varphi(v), \varphi^2(v), \ldots, \varphi^n(v)$ linear unabhängig sind.

Aufgabe 37. Seien $A \in \mathrm{Mat}_{m,n}(K)$ und $B \in \mathrm{Mat}_{n,r}(K)$. Zeigen Sie, dass

$$(AB)^t = B^t A^t$$

gilt.

Aufgabe 38. Seien $v_1, \ldots, v_m \in K^m$ und $A = (v_1, \ldots, v_m) \in \mathrm{Mat}_m(K)$. Zeigen Sie, dass $(v_1, \ldots, v_m)$ genau dann eine Basis von K^m ist, wenn A invertierbar ist.

Aufgabe 39. Bestimmen Sie den Rang der folgenden Matrix über dem Körper $\mathbb{Q}$:

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

Was ist der Rang, wenn wir A als Matrix über dem Körper $\mathbb{Z}/3\mathbb{Z}$ auffassen? Finden Sie jeweils eine invertierbare Untermatrix von maximaler Größe.

Aufgabe 40. (i) Schreiben Sie die Permutationen $\sigma_1, \sigma_2 \in S_8$ jeweils als Produkt von elementfremden Zyklen und als Produkt von Transpositionen. Bestimmen Sie ihre Signaturen.

$$\sigma_1 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 3 & 8 & 4 & 2 & 5 & 1 & 6 & 7 \end{pmatrix}, \sigma_2 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 6 & 8 & 5 & 1 & 7 & 4 & 3 & 2 \end{pmatrix}.$$

(ii) Zeigen Sie, dass für jedes $\sigma \in S_m$ (und jeden Körper K) genau eine Matrix $A_\sigma \in \mathrm{Mat}_m(K)$ existiert mit

$$A_\sigma \begin{pmatrix} c_1 \\ \vdots \\ c_m \end{pmatrix} = \begin{pmatrix} c_{\sigma(1)} \\ \vdots \\ c_{\sigma(m)} \end{pmatrix}$$

für alle $(c_1, \ldots, c_m)^t \in K^m$.

Aufgabe 41. (i) Berechnen Sie die Determinante der folgenden Matrix:

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 8 & 7 & 6 \\ 5 & 4 & 3 & 2 \end{pmatrix}.$$

Hängt das Ergebnis vom Körper ab, über dem die Matrix betrachtet wird?
(ii) Sei $A \in \mathrm{Mat}_m(\mathbb{Z})$. Zeigen Sie die Äquivalenz der folgenden beiden Aussagen:

(a) $A \in \mathrm{GL}_m(\mathbb{Q})$.

(b) $A \in \mathrm{GL}_m(\mathbb{Z}/p\mathbb{Z})$ für alle bis auf endlich viele Primzahlen p.

(iii) Sei $A \in \mathrm{Mat}_m(\mathbb{Q})$. Zeigen Sie die Äquivalenz der folgenden beiden Aussagen:

(a) $A \in \mathrm{GL}_m(\mathbb{C})$.

(b) $A \in \mathrm{GL}_m(\mathbb{Q})$.

Aufgabe 42. Sei K ein Körper. Zeigen Sie, dass für alle $a_0, \ldots, a_m, x \in K$ gilt:

$$\det \begin{pmatrix} x & 0 & \cdots & 0 & a_0 \\ -1 & x & \ddots & \vdots & a_1 \\ 0 & -1 & \ddots & 0 & \vdots \\ \vdots & \ddots & \ddots & x & a_{m-1} \\ 0 & \cdots & 0 & -1 & x + a_m \end{pmatrix} = x^{m+1} + a_m x^m + \cdots + a_1 x + a_0.$$

Aufgabe 43. Bestimmen Sie alle Werte $\lambda \in \mathbb{R}$, für welche die folgende Matrix über $\mathbb{R}$ invertierbar ist:

$$A = \begin{pmatrix} 3 & \lambda & -\lambda \\ 2 & -1 & 3 \\ \lambda + 10 & 1 & 1 \end{pmatrix}.$$

Aufgabe 44. Sei $A \in \mathrm{GL}_m(K)$, $B \in \mathrm{Mat}_m(K)$ und $\lambda \in K$. Zeigen Sie

$$\det(\lambda I_m - AB) = \det(\lambda I_m - BA).$$

Aufgabe 45. Berechnen Sie die Determinante der Matrix

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

indem Sie die Determinante von $A \cdot A^t$ berechnen und von dieser auf die Determinante der Matrix A schließen.

Aufgabe 46. (i) Sei $A \in \mathrm{Mat}_m(\mathbb{R})$, wobei m ungerade ist. Außerdem sei A schiefsymmetrisch, das heißt $A = -A^t$. Zeigen Sie $\det(A) = 0$.
(ii) Sei $A \in \mathrm{Mat}_m(\mathbb{R})$, wobei A mehr als $m^2 - m$ Elemente besitzt, die gleich 0 sind. Zeigen Sie $\det(A) = 0$.

Aufgabe 47. (i) Geben Sie zwei Matrizen $A,\ B \in \mathrm{Mat}_2(\mathbb{Z}/5\mathbb{Z})$ an, für die $A \cdot B \neq B \cdot A$ gilt.
(ii) Eine Matrix heißt symmetrisch, wenn $A = A^t$ und schiefsymmetrisch, wenn $A = -A^t$ gilt. Zeigen Sie, dass sich jede Matrix $A \in \mathrm{Mat}_m(K)$ als Summe einer symmetrischen Matrix A_s und einer schiefsymmetrischen Matrix A_t darstellen läßt.

Aufgabe 48. Es sei V ein $\mathbb{R}$-Vektorraum und $a, b, c \in V$. Außerdem seien

$$x := b + c,\ y := c + a \text{ und } z := a + b.$$

Zeigen Sie:
(i) Es ist $\mathrm{Span}_K(\{a, b, c\}) = \mathrm{Span}_K(\{x, y, z\})$.
(ii) Es sind a, b, c genau dann linear unabhängig, wenn x, y, z linear unabhängig sind.
(iii) Sind die obigen zwei Aussagen auch für Vektorräume über einem beliebigen Körper richtig?

Aufgabe 49. Sei K ein Körper und $x_1, \ldots, x_m \in K$. Wir betrachten

$$V(x_1, \ldots, x_m) := \begin{pmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{m-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{m-1} \\ & \vdots & & & \\ 1 & x_m & x_m^2 & \cdots & x_m^{m-1} \end{pmatrix} \in \mathrm{Mat}_m(K).$$

Zeigen Sie

$$\det\left(V(x_1, \ldots, x_m)\right) = \prod_{1 \leq i < j \leq m} (x_j - x_i).$$

Insbesondere ist $V(x_1, \ldots, x_m)$ genau dann invertierbar, wenn die x_i paarweise verschieden sind.

Aufgabe 50. Sei K ein Körper und $x_1, \ldots, x_m \in K$ paarweise verschiedene Elemente. Zeigen Sie, dass es dann für jede Wahl von $y_1, \ldots, y_m \in K$ genau ein Polynom $p \in K[t]$ vom Grad höchstens $m - 1$ gibt, mit $p(x_i) = y_i$ für alle $i = 1, \ldots, m$.

Aufgabe 51. Sei V ein K-Vektorraum und $\varphi\colon V \to V$ eine lineare Abbildung. Seien $\lambda_1, \ldots, \lambda_m \in K$ paarweise verschiedene Elemente und dazu $v_1, \ldots, v_m \in V \setminus \{0\}$ mit

$$\varphi(v_i) = \lambda_i v_i$$

für alle $i = 1, \ldots, m$. Zeigen Sie, dass $v_1, \ldots, v_m$ dann linear unabhängig sind.

Aufgabe 52. Seien V ein K-Vektorraum und $W_1, W_2 \subseteq V$ Untervektorräume mit $W_1 \cap W_2 = \{0\}$. Seien $x_1, \dots, x_m \in W_1$ und $y_1, \dots, y_n \in W_2$ jeweils linear unabhängig. Zeigen Sie, dass dann auch

$$x_1, \dots, x_m, y_1, \dots, y_n$$

linear unabhängig sind.

Aufgabe 53. Die Matrix $A \in \mathrm{Mat}_m(K)$ habe *obere Block-Dreiecksform*, d.h. sie sei von der Form

$$A = \left(\begin{array}{c|c} B & C \\ \hline 0 & D \end{array}\right)$$

mit $B \in \mathrm{Mat}_{m_1}(K), C \in \mathrm{Mat}_{m_1,m_2}(K), D \in \mathrm{Mat}_{m_2}(K)$ und $m_1 + m_2 = m$. Zeigen Sie, dass dann $\det(A) = \det(B) \cdot \det(D)$ gilt.

Aufgabe 54. Im $\mathbb{R}^5$ seien die folgenden Vektoren gegeben:

$$v_1 = \begin{pmatrix} 4 \\ 1 \\ 1 \\ 0 \\ -2 \end{pmatrix}, v_2 = \begin{pmatrix} 0 \\ 1 \\ 4 \\ -1 \\ 2 \end{pmatrix}, v_3 = \begin{pmatrix} 4 \\ 3 \\ 9 \\ -2 \\ 2 \end{pmatrix}, v_4 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, v_5 = \begin{pmatrix} 0 \\ -2 \\ -8 \\ 2 \\ -4 \end{pmatrix}.$$

Wir betrachten $V = \mathrm{Span}_{\mathbb{R}}\{v_1, \dots, v_5\} \subseteq \mathbb{R}^5$.
(i) Bestimmen Sie $\dim_{\mathbb{R}}(V)$.
(ii) Bestimmen Sie sämtliche Möglichkeiten, aus $v_1, \dots, v_5$ eine Basis für V auszuwählen.

Aufgabe 55. Für $m \geq 2$ betrachten wir

$$M = \begin{pmatrix} 0 & 1 & \cdots & 1 \\ & \ddots & \ddots & \vdots \\ & & \ddots & 1 \\ 0 & & & 0 \end{pmatrix} \in \mathrm{Mat}_m(\mathbb{Q}).$$

Zeigen Sie $M^{m-1} \neq 0$ und $M^m = 0$.

Aufgabe 56. Gibt es eine $\mathbb{C}$-Vektorraumstruktur auf $\mathbb{R}$, sodass die skalare Multiplikation

$$\mathbb{C} \times \mathbb{R} \to \mathbb{R}$$

eingeschränkt auf $\mathbb{R} \times \mathbb{R}$ die übliche Multiplikation auf $\mathbb{R}$ ist?

Aufgabe 57. Wir betrachten folgende lineare Abbildung:

$$\begin{aligned}\varphi\colon \mathbb{C}[t]_{\leq 4} &\to \mathbb{C}[t]_{\leq 3}\\ p &\mapsto p' + p'' + p(0).\end{aligned}$$

Bestimmen Sie die Darstellungmatrix $M_{\underline{v},\underline{w}}(\varphi)$ bezüglich Basen $\underline{v}$, $\underline{w}$ Ihrer Wahl.

Aufgabe 58. Bestimmen Sie für die folgende Matrix $A \in \mathrm{Mat}_3(\mathbb{R})$ das charakteristische Polynom, sowie die Eigenwerte und die Eigenräume von μ_A:

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

Aufgabe 59. Sei V ein K-Vektorraum und $\varphi\colon V \to V$ ein Endomorphismus. Seien $\lambda_1, \ldots, \lambda_n \in K$ paarweise verschiedene Eigenwerte von φ, und $\mathcal{B}_1, \ldots, \mathcal{B}_n$ dazu jeweils endliche Mengen von linear unabhängigen Eigenvektoren. Zeigen Sie, dass dann $\mathcal{B}_1 \cup \cdots \cup \mathcal{B}_n$ ebenfalls linear unabhängig ist.

Aufgabe 60. Für $A \in \mathrm{Mat}_n(\mathbb{R})$ suchen wir differenzierbare Funktionen $f_1, \ldots, f_n\colon \mathbb{R} \to \mathbb{R}$, welche die Differentialgleichung $f' = Af$ erfüllen. Dabei ist $f = (f_1, \ldots, f_n)^t$ und $f' = (f_1', \ldots, f_n')^t$. Zeigen Sie:
(i) Ist $0 \neq v \in \mathbb{R}^n$ und $\lambda \in \mathbb{R}$, so ist $f := e^{\lambda x} \cdot v$ eine Lösung der Differentialgleichung genau dann, wenn v ein Eigenvektor von A zum Eigenwert λ ist.
(ii) Sind $v_1, \ldots, v_r \in \mathbb{R}^n$ linear unabhängige Eigenvektoren zu den Eigenwerten $\lambda_1, \ldots, \lambda_r$ von A, so sind die Lösungen

$$e^{\lambda_1 x} v_1, \ldots, e^{\lambda_r x} v_r$$

linear unabhängig (im Raum der Funktionen von $\mathbb{R}$ nach $\mathbb{R}^n$).

Aufgabe 61. Sei V ein K-Vektorraum und $\varphi\colon V \to V$ ein Endomorphismus, für den jeder Vektor $0 \neq v \in V$ ein Eigenvektor ist. Zeigen Sie, dass dann

$$\varphi = \lambda \cdot \mathrm{id}_V$$

für ein $\lambda \in K$ gelten muss.

Aufgabe 62. Berechnen Sie (abhängig von $a \in \mathbb{R}$) die Eigenwerte und Eigenräume der folgenden Matrix $A \in \mathrm{Mat}_3(\mathbb{R})$:

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

Aufgabe 63. Sei $\varphi_1 \colon \mathbb{R}^3 \to \mathbb{R}^3$ die Drehung um $\pi/4$ um die x_3-Achse sowie $\varphi_2 \colon \mathbb{R}^3 \to \mathbb{R}^3$ die Spiegelung an der Ebene definiert durch $x_1 + 2x_2 = 0$.
(i) Bestimmen Sie die Darstellungsmatrizen $M_{\underline{v},\underline{v}}(\varphi_i)$ bezüglich Basen $\underline{v}$ Ihrer Wahl.
(ii) Berechnen Sie alle reellen Eigenwerte und Eigenräume der φ_i.

Aufgabe 64. Berechnen Sie die Potenzen A^k in allgemeiner Form für die folgende Matrix:

$$A = \begin{pmatrix} 3 & 6 \\ 6 & -2 \end{pmatrix}.$$

Aufgabe 65. Bestimmen Sie für die folgenden Polynome $p_i \in K[t]$ jeweils sämtliche Nullstellen samt Vielfachheiten in K:
(i) $p_1 = -4t^2 + 4t^3 - 2t^5 + t^6$ für $K = \mathbb{Q}, \mathbb{R}, \mathbb{C}$.
(ii) $p_2 = t^8 + 1$ für $K = \mathbb{Z}/2\mathbb{Z}$.

Aufgabe 66. Untersuchen Sie für die folgende Matrix $A \in \mathrm{Mat}_3(\mathbb{R})$ den Endomorphismus

$$\mu_A \colon \mathbb{R}^3 \to \mathbb{R}^3$$

auf Diagonalisierbarkeit und Trigonalisierbarkeit, in Abhängigkeit vom Parameter a:

$$A = \begin{pmatrix} -3 & 0 & 0 \\ 2 & a & 1 \\ 10 & 0 & 2 \end{pmatrix}.$$

Aufgabe 67. Diagonalisieren Sie die folgende Matrix $A \in \mathrm{Mat}_4(\mathbb{R})$, d.h. bestimmen Sie $P \in \mathrm{GL}_4(\mathbb{R})$, $D \in \mathrm{Mat}_4(\mathbb{R})$ diagonal, mit $P^{-1}AP = D$.

$$A = \begin{pmatrix} -5 & 1 & 6 & 6 \\ -12 & 2 & 12 & 12 \\ 1 & 1 & 0 & -2 \\ -4 & 0 & 4 & 6 \end{pmatrix}.$$

Aufgabe 68. Zeigen Sie für $A \in \mathrm{Mat}_m(\mathbb{C})$ die Äquivalenz der folgenden beiden Aussagen:

(i) Es gibt ein $n \in \mathbb{N}$ mit $A^n = 0$.

(ii) A hat nur den Eigenwert 0.

Zeigen Sie weiter, dass in diesem Fall stets $A^m = 0$ gilt.

Aufgabe 69. Finden Sie eine Matrix $A \in \mathrm{Mat}_9(\mathbb{Q})$ mit den folgenden Eigenschaften:

$$\dim\left(L(A,0)\right) = 4,\ \dim\left(L(A^2,0)\right) = 7$$
$$\dim\left(L(A^3,0)\right) = 8,\ \dim\left(L(A^4,0)\right) = 9.$$

Aufgabe 70. Bestimmen Sie für die Abbildung

$$\begin{aligned} \varphi\colon \mathbb{C}[t]_{\leq d} &\to \mathbb{C}[t]_{\leq d} \\ p &\mapsto p(t+1). \end{aligned}$$

eine Darstellungsmatrix in Jordan'scher Normalform.

Aufgabe 71. Sei

$$A = \begin{pmatrix} 2 & 2 & 0 & -1 \\ 0 & 0 & 0 & 1 \\ 1 & 5 & 2 & -1 \\ 0 & -4 & 0 & 4 \end{pmatrix} \in \mathrm{Mat}_4(\mathbb{C}).$$

Bestimmen Sie eine Basis von $\mathbb{C}^4$, bezüglich der die Darstellungsmatrix von μ_A Jordan'sche Normalform hat.

Aufgabe 72. Sei $A \in \mathrm{Mat}_5(\mathbb{C})$ eine Matrix mit charakteristischem Polynom $p_A = (t-1)^3(t+1)^2$. Weiter sei der Eigenraum zum Eigenwert 1 zweidimensional und der zum Eigenwert -1 eindimensional. Bestimmen Sie eine Jordan'sche Normalform für A.

Aufgabe 73. Bestimmen Sie das Minimalpolynom der folgenden Matrix in Abhängigkeit von $a \in \mathbb{C}$:

$$\begin{pmatrix} -1 & 0 & -1 & 0 \\ 1 & a & -a & 2 \\ 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & -1 \end{pmatrix}.$$

Aufgabe 74. Sei $A \in \mathrm{Mat}_m(\mathbb{C})$ beliebig, mit $m \geq 2$. Zeigen Sie, dass die Matrizen

$$I, A, A^2, \ldots, A^{m^2-1}$$

keine Basis von $\mathrm{Mat}_m(\mathbb{C})$ bilden.

Aufgabe 75. Sei V ein $\mathbb{K}$-Vektorraum mit Norm $\|\cdot\|$. Zeigen Sie:
(i) Ist $\|\cdot\|$ von einem Skalarprodukt auf V induziert, so gilt für alle $v, w \in V$ die sogenannte *Parallelogrammgleichung*:

$$\|v+w\|^2 + \|v-w\|^2 = 2\left(\|v\|^2 + \|w\|^2\right).$$

(ii) Die Normen $\|\cdot\|_\infty$ und $\|\cdot\|_1$ auf $\mathbb{R}^m$ sind nicht von einem Skalarprodukt induziert.

Aufgabe 76. Sei V ein $\mathbb{R}$-Vektorraum mit Norm $\|\cdot\|$, welche die Parallelogrammgleichung erfüllt. Zeigen Sie, dass die folgende Setzung dann ein Skalarprodukt auf V ist, welches $\|\cdot\|$ wiederum induziert:

$$\langle v, w\rangle = \frac{1}{2}\left(\|v+w\|^2 - \|v\|^2 - \|w\|^2\right).$$

Aufgabe 77. Auf dem $\mathbb{R}$-Vektorraum $V = \mathbb{R}[t]_{\leq 2}$ definieren wir

$$\langle p, q\rangle := p(0)q(0) + p(1)q(1) + p(2)q(2).$$

(i) Zeigen Sie, dass $\langle\cdot,\cdot\rangle$ ein Skalarprodukt auf V ist.
(ii) Bestimmen Sie eine ONB von V bezüglich $\langle\cdot,\cdot\rangle$.

Aufgabe 78. Beweisen Sie Korollar 6.2.7.

Aufgabe 79. Sei $A = A^t \in \mathrm{Mat}_m(\mathbb{R})$ eine reelle symmetrische Matrix. Wir definieren folgende Abbildung:

$$\begin{aligned}\beta_A \colon \mathbb{R}^m \times \mathbb{R}^m &\to \mathbb{R}\\ (v, w) &\mapsto v^t A w.\end{aligned}$$

(i) Zeigen Sie, dass β_A linear im ersten Eintrag und symmetrisch ist.
(ii) Geben Sie ein explizites Beispiel für A an, in dem β_A kein Skalarprodukt auf $\mathbb{R}^m$ ist.

Aufgabe 80. Sei V ein endlich dimensionaler $\mathbb{K}$-Vektorraum, $U \subseteq V$ ein Unterraum und sei $\pi_U \colon V \to V$ die orthogonale Projektion auf U. Zeigen Sie:
(i) Es gilt $\pi_U = \pi_U^* = \pi_U \circ \pi_U$.
(ii) Ist $\pi \colon V \to V$ ein Endomorphismus mit $\pi = \pi^* = \pi \circ \pi$, so ist π die orthogonale Projektion auf einen Unterraum von V.

Aufgabe 81. Berechnen Sie für die folgende Matrix $A \in \mathrm{Sym}_4(\mathbb{R})$ eine ONB von $\mathbb{R}^4$ aus Eigenvektoren von μ_A:

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

Aufgabe 82. Sei $A \in \mathrm{Mat}_m(\mathbb{R})$ orthogonal. Zeigen Sie:
(i) Es gilt $\det(A) \in \{-1, 1\}$.
(ii Aus $\det(A) = -1$ folgt, dass A den Eigenwert -1 hat.
(iii) Gelten (i) und (ii) auch für unitäre Matrizen?

Aufgabe 83. Seien $A, B \in \mathrm{Her}_m(\mathbb{C})$ positiv semidefinite Matrizen. Zeigen Sie:
(i) Für $\lambda, \gamma \geq 0$ gilt $\lambda A + \gamma B \geqslant 0$.
(ii) Für $W \in \mathrm{Mat}_{m,n}(\mathbb{C})$ gilt $W^*AW \geqslant 0$.
(iii) Es gibt $v_1 \ldots, v_m \in \mathbb{C}^m$ mit $A = \sum_{i=1}^m v_i v_i^*$.
(iv) Für $v \in \mathbb{C}^m$ gilt $v^*Av = 0 \Leftrightarrow Av = 0$.

Aufgabe 84. Seien V ein endlich dimensionaler $\mathbb{C}$-Vektorraum und $\varphi, \psi \colon V \to V$ zwei Endomorphismen mit $\psi \circ \varphi = \varphi \circ \psi$. Zeigen Sie:
(i) Für jeden Eigenwert $\lambda \in \mathbb{C}$ von φ ist der Unterraum $\mathrm{Eig}(\varphi, \lambda)$ ψ-invariant, d.h. es gilt

$$\psi_{|_{\mathrm{Eig}(\varphi,\lambda)}} \colon \mathrm{Eig}(\varphi, \lambda) \to \mathrm{Eig}(\varphi, \lambda).$$

(ii) Sind φ und ψ beide normal, so sind sie *simultan* diagonalisierbar (d.h. es gibt eine ONB Basis von V, bezüglich der beide Darstellungsmatrizen Diagonalgestalt haben).

Aufgabe 85. Sei $p \in \mathbb{R}[t]$ ein Polynom, das über $\mathbb{R}$ in Linearfaktoren zerfällt. Zeigen Sie die Äquivalenz der folgenden beiden Aussagen:

(i) Alle Nullstellen von p sind größer gleich Null.

(ii) Die Koeffizienten von p haben alternierendes Vorzeichen (d.h. sie sind abwechselnd ≥ 0 und ≤ 0).

Zeigen Sie damit, dass man die Bedingung "$A \geqslant 0$" für $A \in \mathrm{Her}_m(\mathbb{C})$ durch m polynomiale Bedingungen an die Koeffizienten von A ausdrücken kann.

Aufgabe 86. Zeigen Sie, dass für $A, B \in \mathrm{Her}_m(\mathbb{C})$ mit $A, B \geqslant 0$ stets gilt

$$L(A+B,0) = L(A,0) \cap L(B,0).$$

Stimmt die Aussage auch für beliebige Matrizen?

Aufgabe 87. Seien K ein Körper und $A \in \mathrm{Mat}_{m,r}(K), B \in \mathrm{Mat}_{r,m}(K)$. Zeigen Sie:
(i) Es gilt $\mathrm{tr}(AB) = \mathrm{tr}(BA)$.
(ii) Stimmt auch $\det(AB) = \det(BA)$?

Aufgabe 88. Sei

$$A = \begin{pmatrix} 4 & -2 & 1 \\ -2 & 4 & 2 \\ 1 & 2 & 0 \end{pmatrix} \in \mathrm{Mat}_3(\mathbb{R}).$$

Bestimmen Sie ein $P \in \mathrm{GL}_3(\mathbb{R})$ und eine Diagonalmatrix $D \in \mathrm{Mat}_3(\mathbb{R})$ mit $P^tAP = D$. Welche Signatur hat A?

Aufgabe 89. Beweisen Sie Proposition 7.3.2.

Aufgabe 90. Es sei K ein Körper mit $1+1 \neq 0$. Zeigen Sie, dass es genau eine symmetrische Bilinearform

$$\beta\colon \mathrm{Mat}_2(K) \times \mathrm{Mat}_2(K) \to K$$

gibt mit

$$\beta(A,A) = \det(A)$$

für alle $A \in \mathrm{Mat}_2(K)$. Stimmt das auch für größere Matrizen? Bestimmen Sie eine Darstellungsmatrix für β.

Aufgabe 91. Sei V ein m-dimensionaler K-Vektorraum. Zeigen Sie:
(i) Die Menge Q aller quadratischen Formen auf V bildet einen K-Untervektorraum des gesamten Funktionenraums $\mathcal{F}(V,K)$. Bestimmen Sie seine Dimension.
(ii) Für jedes $q \in Q$ ist $\mathcal{N}(q) := \{v \in V \mid q(v) = 0\}$ eine Vereinigung von Ursprungsgeraden.
(iii) Ist $N \subseteq V$ eine Vereinigung von höchstens $\binom{m+1}{2} - 1$ Ursprungsgeraden, so gibt es ein $0 \neq q \in Q$ mit $N \subseteq \mathcal{N}(q)$.

Aufgabe 92. Zeigen Sie, dass das Innere und der Abschluss einer konvexen Menge wieder konvex sind.

Aufgabe 93. Seien $\varphi\colon \mathbb{R}^m \to \mathbb{R}^n$ eine lineare Abbildung und $S \subseteq \mathbb{R}^m, T \subseteq \mathbb{R}^n$ zwei konvexe Mengen. Zeigen Sie:
(i) $\varphi(S)$ und $\varphi^{-1}(T)$ sind konvex.
(ii) $S \times T \subseteq \mathbb{R}^m \times \mathbb{R}^n = \mathbb{R}^{n+m}$ ist konvex.

Aufgabe 94. Zeigen Sie:
(i) Die Menge der positiv semidefiniten Matrizen ist ein abgeschlossener konvexer Kegel $P \subseteq \mathrm{Her}_m(\mathbb{C})$.
(ii) Das Innere von P besteht genau aus den positiv definiten Matrizen.

Aufgabe 95. Seien $p \in \mathbb{C}[t]$ ein Polynom und $z_1, \ldots, z_d \in \mathbb{C}$ seine Nullstellen. Zeigen Sie, dass die Nullstellen der Ableitung p' in der konvexen Hülle von $z_1, \ldots, z_d$ liegen, wenn man $\mathbb{C}$ als $\mathbb{R}^2$ auffasst.

Aufgabe 96. Sei $S \subseteq \mathbb{R}^d$ eine Menge mit mindestens $d + 2$ Elementen. Zeigen Sie, dass es disjunkte Teilmengen $R, B \subseteq S$ gibt mit

$$\mathrm{Konv}(R) \cap \mathrm{Konv}(B) \neq \emptyset.$$

Hinweis: Wählen Sie paarweise verschiedene Elemente $s_1, \ldots, s_{d+2}$ aus S und betrachten sie das lineare Gleichungssystem $\sum_i \lambda_i s_i = 0, \sum_i \lambda_i = 0$. Gruppieren Sie die s_i anhand der Vorzeichen der λ_i einer nichttrivialen Lösung in die zwei Mengen R und B.

Aufgabe 97. Seien $A, B \in \mathrm{Her}_m(\mathbb{C})$ zwei positiv semidefinite Matrizen. Zeigen Sie, dass $\mathrm{tr}(A) \geq 0$ und $\mathrm{tr}(AB) \geq 0$ gilt.

Aufgabe 98. Sei $K := \{a \in \mathbb{R}^m \mid a_1, \ldots, a_m \geq 0\} \subseteq \mathbb{R}^m$.
(i) Zeigen Sie, dass K ein endlich erzeugter konvexer Kegel ist.
(ii) Bestimmen Sie den dualen Kegel $K^\vee$.

Aufgabe 99. (i) Zeigen Sie, dass für jedes $A \in \mathrm{Her}_m(\mathbb{C})$ die folgenden beiden Bedingungen äquivalent sind:

(a) $\mathrm{tr}(AB) \geq 0$ für alle $B \in \mathrm{Her}_m(\mathbb{C})$ mit $B \geqslant 0$.

(b) $A \geqslant 0$.

(ii) Zeigen Sie, dass die Setzung $\langle A, B\rangle := \mathrm{tr}(AB)$ ein Skalarprodukt auf dem reellen Vektorraum $\mathrm{Her}_m(\mathbb{C})$ definiert und bestimmen Sie diesbezüglich den dualen Kegel des Kegels der positiv semidefinite Matrizen.

Aufgabe 100. Entscheiden Sie die Lösbarkeit des folgenden Systems linearer Ungleichungen über $\mathbb{R}$:

$$\begin{aligned} x_1 + 2x_2 - x_3 &\geq 1 \\ x_1 - x_2 + x_3 &\geq 1 \\ -x_1 + x_2 - 2x_3 &\geq 0 \\ x_3 &\geq 0 \end{aligned}$$

Aufgabe 101. Seien $A_1, \ldots, A_n \subseteq \mathbb{R}^m$ konvexe Mengen, so dass jeweils $m+1$ der Mengen einen nichtleeren Schnitt haben. Zeigen Sie, dass dann

$$A_1 \cap \cdots \cap A_n \neq \emptyset$$

gilt.
Hinweis: Gehen Sie per Induktion über n vor und benutzen Sie Aufgabe 96.

Aufgabe 102. Beweisen Sie Proposition 9.1.6 (ii).

Aufgabe 103. Zeigen Sie, dass die injektive Abbildung

$$\bigoplus_{j \in I} V_j' \hookrightarrow \left(\prod_{j \in I} V_j \right)'$$

aus Proposition 9.1.7 im Allgemeinen nicht surjektiv ist.

Aufgabe 104. Beweisen Sie Proposition 9.1.12 (ii).

Aufgabe 105. Sei $V = \mathbb{R}[t]$ der Vektorraum der reellen Polynome und

$$U = \{ p \in V \mid p(0) = 0 \} .$$

Zeigen Sie $V/U \cong \mathbb{R}$, indem Sie einen Isomorphismus explizit angeben.

Aufgabe 106. Zeigen Sie
(i) $K^m \otimes_K K^n \cong \mathrm{Mat}_{m,n}(K) \cong K^{mn}$.
(ii) $\mathrm{Mat}_m(K) \otimes \mathrm{Mat}_n(K) \cong \mathrm{Mat}_{mn}(K)$.