Lineare Algebra 1 und 2

Wintersemester 2021/2022 und Sommersemester 2022

Catherine Meusburger
Department Mathematik
Friedrich-Alexander-Universität Erlangen-Nürnberg

(Stand: 22. August 2022)

Bei der Vorbereitung dieser Vorlesung habe ich die folgende Literatur benutzt, die ich auch den Studierenden als Lehrbücher empfehle:

- [F] G. Fischer, Lineare Algebra,
- [L] F. Lorenz, Lineare Algebra I und II,
- [G] W. Greub, Lineare Algebra,
- [B] E. Brieskorn, Lineare Algebra und analytische Geometrie,
- [S] H. Schichl, Roland Steinbauer, Einführung in das mathematische Arbeiten,

Zusätzlich habe ich auch die Vorlesungsskripten [K] von Andreas Knauf, [NVk, NLA1, NLA2] von Karl-Hermann Neeb und [Sw] von Christoph Schweigert verwendet.

Bitte schicken Sie Anmerkungen und Kommentare zu diesem Vorlesungsskript an:

`catherine.meusburger@math.uni-erlangen.de`.

Einleitung

Das Vorlesungsskript

Dieses Vorlesungsskript ersetzt nicht den Besuch der Vorlesung und Übungen, da es Ihnen im Gegensatz zu studentischen Tutor:innen, Assistent:innen und Professor:innen weder Fragen stellt noch Fragen beantwortet. Es erlaubt es Ihnen aber, in der Vorlesung nicht alles akribisch mitzuschreiben, sondern nur ausgewählte Informationen im Skript zu ergänzen. Sie können das Skript auch benutzen, um den Vorlesungsstoff zu wiederholen, zu üben und zu vertiefen. Dabei können ihnen die folgenden Elemente helfen:

- Am Ende des Skripts finden sich ein Notationsverzeichnis und ein Index, mit denen Sie Ihnen nicht (mehr) geläufige Symbole und Begriffe nachschlagen können.
- Am Ende jedes Kapitels finden Sie eine Übersicht über die wichtigsten Begriffe, Aussagen und Beispiele des Kapitels. Sie können die Übersicht auch benutzen, um sich selbst zu testen und zu sehen, ob Sie die wichtigsten Dinge im Kopf behalten haben.
- Das Skript ist relativ ausführlich mit viel erzählendem Text zwischen den mathematischen Bausteinen. Wenn Ihnen die nummerierten mathematischen Bausteine (Definition, Beispiel, Bemerkung, Lemma, Satz, Korollar und Beweis) ausreichen, um den Stoff zu verstehen, können Sie den Text zwischen diesen Inhalten einfach ignorieren. Die Mathematik ist in den nummerierten Inhalten konzentriert.
- Wenn Ihnen aus den nummerierten mathematischen Inhalten nicht direkt klar wird, worum es geht, können Sie den Text zwischen diesen Inhalten lesen, der Ihnen hoffentlich beim Verständnis hilft.
- Das Skript enthält relativ viele Beispiele, die aber zum Teil nur kurz skizziert werden. Dies ist absichtlich so, damit Sie sich mit den Beispielen beschäftigen und über sie nachdenken. Sie werden den Stoff besser verstehen, wenn Sie die fehlenden Schritte ergänzen und diese Beispiele im Detail verstehen.
- Am Ende des Vorlesungsskripts finden Sie eine umfangreiche Aufgabensammlung. Manche dieser Aufgaben werden in der Vorlesung behandelt. Andere können Sie benutzen, um sich auf die Klausur vorzubereiten.

Wie lernt man Mathematik?

Das ist am Anfang des Studiums oft relativ anstrengend und schwierig. Mathematik zu lernen ist eher einem Training für eine Sportart oder dem Erlernen eines Musikinstruments vergleichbar als dem Lernen in der Schule. Man muss *regelmäßig* trainieren und üben und sein Pensum vernünftig anpassen.

Trainieren und üben heißt in diesem Fall vor allem, die Vorlesung zu besuchen und nachzuarbeiten, die Präsenzübungen und die Hausaufgaben zu bearbeiten und gegen Ende des Semesters, rechtzeitig anzufangen, den Stoff zu wiederholen.

Das Pensum vernünftig anzupassen heißt, dass man an den Grenzen seiner Fähigkeiten arbeitet, denn nur so kann man die Grenzen verschieben. Es soll einem nicht leichtfallen, sondern ziemlich schwierig und anstrengend sein. Es soll aber auch Spaß bringen und nicht nur Verzweiflung. Es ist wichtig, dass man ehrlich zu sich selbst ist und sich klar macht, was man kann und was man noch üben muss.

Man sollte sich herausfordern, aber auch Geduld mit sich selbst haben, wenn man nicht alles sofort durchdringt. Etwas nicht zu verstehen oder nicht lösen zu können, ist im Mathematikstudium der Normalzustand. Es hilft, mit Kommiliton:innen und fortgeschritteneren Studierenden zu sprechen und sie nach ihren Erfahrungen und Strategien zu fragen.

Die Vorlesung nachzuarbeiten heißt, den Stoff nach der Vorlesung systematisch durchzugehen und ihn kritisch zu durchdenken. Dies betrifft vor allem die nummerierten mathematischen Inhalte, also Definitionen, Beispiele, Bemerkungen, Lemmata, Sätze, Korollare und Beweise.

Umgang mit mathematischen Strukturen

- **Definitionen und Beispiele:**
  Definitionen enthalten die Spielregeln der Mathematik. Stellt man sich mathematische Begriffe wie Spielfiguren in einem Brettspiel vor, so sind die Definitionen die Spielregeln, die beschrieben, welche Züge man damit machen kann. Um die Spielfiguren richtig einsetzen zu können, muss man präsent haben, welche Züge sie machen können und sollte es nicht jedes Mal neu nachlesen müssen. Es hilft, wenn man sich dazu die folgenden Fragen stellt: Was sind die Voraussetzungen der Definition? Warum braucht man sie? Was passiert, wenn man sie weglässt oder abändert? Fallen einem Beispiele und Gegenbeispiele dazu ein? Was wird dadurch alles erfasst?

  Oft braucht man aber nicht nur das abstrakte Wissen, welche Spielzüge möglich sind, sondern zusätzlich ein praktisches Wissen, um schnell zu entscheiden, welcher Zug sinnvoll ist oder sich in einer gegebenen Situation anbietet. Diese praktische Wissen kommt aus der Betrachtung von *Beispielen* und aus den bearbeiteten Übungsaufgaben. Dafür muss man das Spiel aber selbst spielen, also die Aufgaben *bearbeiten*, nicht nur erklären lassen oder die Lösung irgendwo nachlesen. Auch die Beispiele sollte man aktiv durchdenken und versuchen, eigene Beispiele zu finden. Kann man nachvollziehen, dass die Bedingungen aus der Definition erfüllt sind? Muss man hier noch etwas nachrechnen oder überprüfen?

- **Bemerkungen, Lemmata, Sätze und Korollare:**
  Nummerierte mathematische Inhalte, die Aussagen enthalten, also *Bemerkungen, Lemmata, Sätze und Korollare*, sind im wesentlichen Abkürzungen, die verschiedene Züge der Spielfiguren verknüpfen. Mit ihnen kann man das Brettspiel effizienter und schneller spielen. Sie geben einem ein Bild vom Zusammenspiel der mathematischen Strukturen. Um sie zu verstehen, hilft es, sich die folgenden Fragen zu stellen:

  Sind einem alle Begriffe und Symbole präsent, die in der Aussage auftreten? Wenn nicht, sollte man sie zunächst nachschlagen. Was sind die Voraussetzungen der Aussage? Wofür werden sie gebraucht? Was passiert, wenn man eine der Voraussetzungen weglässt? Kennt man ein Beispiel, das zeigt, dass die Aussage dann falsch wird? Wie kann man die Aussage mit Hilfe der Aussagenlogik noch anders formulieren? Kann man die Aussage visualisieren oder sich mit Hilfe eines Bildes besser merken? Wozu ist die Aussage gut, was hat man davon? Mathematische Aussagen sind im allgemeinen dafür da, etwas zu vereinfachen oder zu klären.

- **Beweise:**
  Jede mathematische Aussage muss *bewiesen* werden. Um *Beweise* zu verstehen und sie sich merken zu können, ist es hilfreich, sich vor dem Lesen des Beweises zu fragen, ob und wie man die Aussage selbst beweisen könnte. Ist einem etwa in einem einfachen Spezialfall klar, warum die Aussage gelten muss? Hat man eine Idee, woher die Aussage kommen

könnte oder eine Intuition, mit welcher Strategie man sie beweisen könnte? Wenn man die Aussage selbst beweisen kann, sollte man dies tun und anschließend den eigenen Beweis mit dem in der Vorlesung gegebenen Beweis vergleichen. Oft gibt es mehrere mögliche Vorgehensweisen.

Um einen Beweis aus der Vorlesung oder aus einem Lehrbuch zu verstehen, sollte man ihn zunächst Schritt für Schritt durchgehen: Kennt man alle darin benutzen Begriffe? Ist einem in jedem Schritt klar, was gefolgert wird? Woraus ergeben sich die Folgerungen? Welche vorher erzielten Ergebnisse werden dabei benutzt?

Anschließend sollte man den Beweis noch einmal mit etwas Abstand betrachten. Um welche Art von Beweis handelt es sich, z. B. vollständige Induktion, Widerspruchsbeweis, einfaches Nachrechnen von Axiomen? Wo genau im Beweis gehen die Voraussetzungen oder Annahmen des Lemmas, Satzes oder Korollars ein? Welcher Beweisschritt würde schiefgehen, wenn man eine der Voraussetzungen wegließe? Kann man die zentrale Idee des Beweises in wenigen Worten zusammenfassen? Wie könnte man sie anderen möglichst knapp erklären? Kann man sich einige Stichworte merken, mit denen man den Beweis selbst reproduzieren kann, ohne in ein Vorlesungsskript oder ein Lehrbuch zu schauen?

Anschauung und Bilder

Universitätsmathematik ist viel abstrakter als Schulmathematik und nimmt viel weniger auf Alltagsbeispiele und Anwendungen Bezug. Zu Beginn des Studiums muss man zunächst lernen, sich von der Anschauung etwas zu lösen. Die Spielregeln der mathematischen Strukturen stehen in den Definitionen. Erlaubt ist, was durch die Definition nicht verboten ist. Im Prinzip muss man die Spielregeln nur anwenden, ohne sich notwendigerweise etwas darunter vorzustellen. So kann man auch mit Dingen sicher umgehen, unter denen man sich nichts vorstellen kann und die man auch nicht visualisieren oder skizzieren kann.

Allerdings ist die Anschauung in der Mathematik trotzdem wichtig und nötig. Man braucht zwar keine Anschauung, um das Spiel Mathematik zu spielen, aber die Mathematik ist letztlich aus Anschauung und möglichen Anwendungen entstanden. Anschauung und mögliche Anwendungen entscheiden letztlich auch darüber, ob eine Definition akzeptiert wird, und wie nützlich sie ist. Sie helfen einem außerdem, sein Wissen über das Spiel Mathematik zu erinnern und organisieren und Lösungsansätze für mathematische Fragen zu finden.

Wichtig ist nur, dass man sich der Grenzen der Anschauung bewusst ist und eigene Vorstellungen oder Bilder von einer mathematischen Struktur kritisch hinterfragt. Worüber man spricht und was man darf, steht in der Definition. Die Anschauung kann die Definition ergänzen und illustrieren. Wenn sie aber im Widerspruch zur Definition steht, gewinnt die Definition.

Selbständig denken und kritisch hinterfragen

Wichtig ist, dass einem die Inhalte einleuchten und man damit umgehen kann. Das muss nicht unbedingt durch die Vorlesung und das Skript passieren. Prinzipiell kann man jedes Lehrbuch benutzen, das den Vorlesungsstoff abdeckt und mit dem man den Stoff gut versteht. Es bietet sich an, mehrere Bücher auszuprobieren und zu sehen, wie man damit zurechtkommt.

Generell sollte man in der Mathematik nichts unkritisch hinnehmen oder glauben, weder den Dozent:innen, noch den Vorlesungsskripten oder Lehrbüchern. Dozent:innen machen in der Regel Fehler, und auch Bücher und Skripten sind nie vollkommen fehlerfrei. Kritische Nachfragen sind daher in der Mathematik prinzipiell wichtig und willkommen. Man sollte nicht zu fragen aufhören, bevor es einem wirklich einleuchtet.

Inhaltsverzeichnis

1 Grundlagen: Mathematische Logik, Beweistechniken und Mengen

Im Gegensatz zur Schulmathematik werden in der Universitätsmathematik und der mathematischen Forschung alle Aussagen bewiesen. Dies liegt daran, dass die Mathematik nicht wie die Naturwissenschaften Experimente durchführen kann, mit denen sich Hypothesen überprüfen lassen. Man kann nicht messen, ob eine mathematische Aussage richtig ist. Man kann sie nur auf andere mathematische Aussagen oder auf gewisse Grundannahmen zurückführen, oder durch Angabe eines Gegenbeispiels zeigen, dass sie falsch ist.

Beweisen bedeutet zu zeigen, dass mathematische Aussagen aus anderen mathematischen Aussagen oder Axiomen (nicht beweisbare Grundannahmen) logisch zwingend folgen. Dabei ist ein hoher Grad an Genauigkeit und Sorgfalt notwendig.

Da mathematische Aussagen aufeinander aufbauen, kann ein kleiner Fehler in einem wichtigen Baustein zum Einsturz des ganzen Gedankengebäudes Mathematik führen. In größerem Maßstab ist dies in der Mathematik zum Glück erst ein einziges Mal passiert, nämlich in der sogenannten *Grundlagenkrise* zu Beginn des 20. Jahrhunderts, als die *Russelsche Antinomie*[1] fundamentale Probleme in der Mengenlehre offenlegte. Damals war eine gewisse Zeit lang unklar, welche Teile der Mathematik von diesem Problem betroffen waren, und welche Aussagen überhaupt noch Gültigkeit hatten. Das strenge und sorgfältige Beweisen von Aussagen hilft, solche Probleme zu vermeiden.

Die Grundlagen des Beweisens liegen in der Logik und der Mengenlehre. Beide Gebiete erscheinen anfangs manchmal abstrakt oder unnötig formal. Dies ist aber notwendig, da man in der Mathematik sehr schnell mit Konzepten in Kontakt kommt, bei denen die unmittelbare Anschauung oder der gesunde Menschenverstand versagt.

Andererseits hat die Abstraktion auch den Vorteil, dass sich abstrakte Konzepte und Methoden sehr allgemein anwenden lassen und nicht auf einzelne Beispiele oder Spezialfälle beschränkt sind. In diesem Vorkurs widmen wir uns zunächst der Logik, dem Schlussfolgern und dem präzisen Argumentieren. Anschließend behandeln wir dann einige grundlegende Beweistechniken und die Grundlagen der Mengenlehre.

1.1 Mathematische Logik

Definition 1.1.1: Eine **Aussage** ist ein sprachliches Gebilde, das entweder wahr oder falsch ist. Im ersten Fall sagt man, die Aussage habe den **Wahrheitswert** w und im zweiten sie habe den Wahrheitswert f.

Beispiel 1.1.2:

1. *“Die Zahl 10 ist eine Primzahl”* ist eine Aussage mit dem Wahrheitswert falsch (f).
2. *“Die Zahl 3 ist eine Primzahl”* ist eine wahre Aussage (w).
3. Die Ausdrücke “$5 + 7$” , “*Guten Tag*” und *“Wie heißen Sie?”* sind keine Aussagen.

[1] Siehe Abschnitt 1.3. Die Grundlagenkrise der Mathematik wird auf sehr ansprechende und unterhaltsame Weise in der Graphic Novel [DP] behandelt.

4. Der Satz *“Dieser Satz ist falsch.”* ist keine Aussage. Denn wäre er wahr, so wäre er falsch und umgekehrt. Man kann hier also keinen Wahrheitswert wahr oder falsch zuordnen[2].
5. Der Satz *“Dieser Satz ist wahr.”* ist keine Aussage. Dies ist weniger leicht einzusehen als im letzten Beispiel. Wir werden aber später sehen, dass man jede Aussage verneinen kann und so eine neue Aussage erhält. Diese wäre dann der Satz aus dem letzten Beispiel.
6. Die sogenannte wahre Aussage W ist eine Aussage, die immer den Wahrheitswert w annimmt.
7. Die sogenannte falsche Aussage F ist eine Aussage, die immer den Wahrheitswert f annimmt.

Sind eine oder mehrere Aussagen gegeben, so lassen sich daraus durch Verknüpfungen neue Aussagen bilden. Der zentrale Punkt dabei ist, dass der Wahrheitswert dieser neuen Aussagen durch den der alten Aussagen bereits festgelegt wird. Die Aussage *“Die Sonne scheint.”* ist beispielsweise keine Verknüpfung der Aussagen *“Der Vorkurs findet heute statt.”* und *“Gestern gab es zum Abendessen Pizza.”* , da die Tatsache, dass der Vorkurs stattfindet, und das Abendessen am Vortag keinerlei Rückschlüsse auf das heutige Wetter zulassen.

Definition 1.1.3: Für $n \in \{1, 2, 3...\}$ ist eine **n-stellige Verknüpfung** oder ein **n-stelliger Junktor** von gegebenen Aussagen $A_1, ..., A_n$ eine Aussage $V(A_1, ..., A_n)$, deren Wahrheitswert durch die Wahrheitswerte der gegebenen Aussagen $A_1, ..., A_n$ eindeutig bestimmt ist. Sie wird durch eine **Wahrheitstafel** beschrieben, die die Wahrheitswerte der Verknüpfung in Abhängigkeit der Wahrheitswerte der gegebenen Aussagen angibt.

A_1	A_2	$\dots$	A_n	$V(A_1, ..., A_n)$
w	w	$\dots$	w	?
f	w	$\dots$	w	?
w	f	$\dots$	w	?
$\vdots$	$\vdots$	$\vdots$	$\vdots$	?
f	f	$\dots$	f	?

Eine Wahrheitstafel für eine n-stelligen Verknüpfung ist eine Tabelle mit $n + 1$ Spalten und 2^n Zeilen. In die ersten n Spalten werden alle möglichen Kombinationen von Wahrheitswerten der gegebenen n Aussagen eingetragen, in die letzte Spalte kommt der Wahrheitswert für die Verknüpfung, der sich aus den jeweils in der gleichen Zeile angegebenen Wahrheitswerten der gegebenen Aussagen ergibt. Da hier für jede der 2^n Kombinationen von Wahrheitswerten der Aussagen $A_1, ..., A_n$ für $V(A_1, ..., A_n)$ entweder der Wahrheitswert w oder der Wahrheitswert f gewählt werden muss, gibt es insgesamt 2^{2^n} verschiedene Wahrheitstafeln, die n-stellige Verknüpfungen beschreiben.

Wichtige Beispiele von Verknüpfungen sind die sogenannten elementaren Verknüpfungen - besonders einfache ein- oder zweistellige Verknüpfungen, aus denen sich jede weitere Verknüpfung zusammensetzen lässt. Diese entsprechen jeweils bestimmten Konjunktionen oder Subjunktionen in der Alltagssprache, nämlich *NICHT*, *UND*, *ODER* sowie *WENN, DANN* und *GENAU DANN, WENN*. Sie können auch als die Grundbausteine oder Grundoperationen des logischen

[2]Der tiefere Grund, warum dieser Satz keine Aussage ist, ist seine *Selbstbezüglichkeit*, also die Tatsache, dass er etwas über sich selbst und nicht über einen anderen Sachverhalt aussagt. Selbstbezüglichkeit ist in der Logik oft problematisch, da sie zu Paradoxa führt. Auch die Russelsche Antinomie ist ein Beispiel eines aus Selbstbezüglichkeit entstehenden logischen Paradoxes. Als leicht zugängliche Hintergrundliteratur zum Thema *Selbstbezüglichkeit* empfehle ich den Artikel [P] und das Buch [Ho].

Denkens aufgefasst werden. Man beachte dabei, dass hier unter *ODER* immer ein *inklusives oder* (*“A oder B oder beides”*) verstanden wird, kein *exklusives oder* (*“entweder A oder B”*).

Definition 1.1.4: Seien A und B gegebene Aussagen. Dann definieren wir die folgenden elementaren Verknüpfungen durch ihre Wahrheitstafeln:

1. **Negation** - nicht A, in Zeichen: $\neg A$

A	$\neg A$
w	f
f	w

2. **Konjunktion** - A und B, in Zeichen: $A \wedge B$

A	B	$A \wedge B$
w	w	w
w	f	f
f	w	f
f	f	f

3. **Disjunktion** - A oder B (oder beide), in Zeichen: $A \vee B$

A	B	$A \vee B$
w	w	w
w	f	w
f	w	w
f	f	f

4. **Implikation** - aus A folgt B, in Zeichen: $A \Rightarrow B$

A	B	$A \Rightarrow B$
w	w	w
w	f	f
f	w	w
f	f	w

5. **Äquivalenz** - A äquivalent zu B, in Zeichen: $A \Leftrightarrow B$

A	B	$A \Leftrightarrow B$
w	w	w
w	f	f
f	w	f
f	f	w

Beispiel 1.1.5:

1. Die Aussage C *“Wenn es regnet, ist die Straße nass.”* ist ein Beispiel für eine Implikation. Hier ist die Aussage A: *“Es regnet.”* und Aussage B *“Die Straße ist nass.”* Aussage C ist wahr, wenn (1. Zeile) es regnet und die Straße nass ist, (3. Zeile) wenn es nicht regnet und die Straße nass ist oder (4. Zeile) wenn es weder regnet noch die Straße nass ist. Sie ist nur dann falsch, wenn (2. Zeile) es regnet, die Straße aber trocken ist. Während hier die erste und zweite Zeile intuitiv offensichtlich sind, ist dies für die dritte und vierte Zeile eventuell etwas schwerer einzusehen. Man beachte dazu, dass weder trockene Straßen in Abwesenheit von Regen noch andere Ursachen für nasse Straßen (z. B. Reinigungsfahrzeuge, Schnee) im Widerspruch dazu stehen, dass Regen die Straße nass macht.

2. Dies Aussage C *“Die Straße ist nass genau dann, wenn es regnet.”* ist ein Beispiel für eine Äquivalenz. Hier werden wieder die Aussagen A: *“Es regnet.”* und Aussage B *“Die Straße ist nass.”* kombiniert. Die Aussage C ist wahr, wenn (1. Zeile) es regnet und die Straße nass ist oder (4. Zeile) es nicht regnet und die Straße trocken ist. Sie ist falsch, wenn es (2. Zeile) regnet, die Straße aber trocken bleibt und (3. Zeile) die Straße nass ist, obwohl es nicht regnet.

3. Die Aussage C *“Wenn Kühe fliegen können, dann können Delfine klettern.”* ist die Implikation der Aussagen A *“Kühe können fliegen.”* und B *“Delfine können klettern.”* Aussage C ist wahr, weil sowohl A als auch B falsch sind (4. Zeile). Kühe können nicht fliegen und Delfine nicht klettern, aber es wird auch nicht behauptet, dass dies der Fall sei, sondern nur dass Delfine klettern könnten, *wenn* Kühe fliegen könnten.

4. Die Aussage *“Diese Person ist eine Studentin oder eine Schülerin.”* ist die Disjunktion der Aussagen A *“Diese Person ist eine Studentin.”* und B *“Diese Person ist eine Schülerin.”* Offensichtlich ist Aussage C wahr, wenn (1. Zeile) die Person sowohl eine Studentin ist als auch Schülerin, also eine Frühstudentin, wenn sie (2. Zeile) zwar eine Studentin ist, aber keine Schülerin oder (3. Zeile) zwar eine Schülerin ist, aber keine Studentin. Sie ist nur dann falsch, wenn die Person (4. Zeile) weder eine Studentin noch eine Schülerin ist.

5. Die Aussage *“Die Person ist entweder eine Studentin oder eine Schülerin.”* ist keine Disjunktion der Aussagen A *“Die Person ist eine Studentin.”* und B *“Die Person ist eine Schülerin.”* Denn sie ist wahr, wenn die Person zwar eine Studentin, aber keine Schülerin ist (2. Zeile) oder eine Schülerin aber keine Studentin (3. Zeile). Sie ist falsch, wenn die Person weder eine Schülerin noch eine Studentin ist (4. Zeile) oder sowohl eine Schülerin als auch eine Studentin (1. Zeile). Im letzten Fall müsste die Disjunktion aber wahr sein.

Merke: Mathematisches ODER ist NICHT-EXKLUSIVES ODER:
“A oder B” heißt immer “A oder B oder beides”, *nicht* “entweder A oder B”

Aufgabe 1: Spielkarten, deren Vorderseiten entweder rot oder blau sind und deren Rückseiten entweder mit einem Kreis oder mit einem Quadrat bedruckt sind, liegen mit einer beliebigen Seite nach oben auf einem Tisch. Es soll der Wahrheitswert der Implikation *“Wenn eine Seite der Karte blau ist, ist auf der anderen Seite ein Kreis.”* bestimmt werden. Welche Karten müssen dazu umgedreht werden?

Das 5. Beispiel ist ein Beispiel für eine weitere gebräuchliche logische Verknüpfung, nämlich das sogenannte **exklusive oder xor** das dem *“entweder ... oder”* in der Umgangssprache entspricht. Daneben gibt es noch die ebenfalls wichtige Verknüpfung **nand**, die fordert, dass maximal eine von zwei gegebenen Aussagen gelten darf.

Die praktische Bedeutung von **nand** rührt daher, dass sich diese Verknüpfung leicht technisch in Schaltkreisen realisieren lässt und man daraus unter Verwendung der wahren und der falschen Aussage *alle* elementaren Verknüpfungen erzeugen kann. Sie spielt daher in der Schaltkreistechnik eine wichtige Rolle und dient dort als Grundbaustein der logischen Verknüpfungen.

Definition 1.1.6: Seien A und B zwei Aussagen. Die Verknüpfungen A xor B und A nand B werden durch die folgenden Wahrheitstafeln definiert:

A	B	A xor B	A nand B
w	w	f	f
w	f	w	w
f	w	w	w
f	f	f	w

Aufgabe 2: Zeigen Sie, dass sich unter Verwendung der wahren und der falschen Aussage *alle* elementaren Verknüpfungen durch **nand** darstellen lassen.

Es ist offensichtlich, dass man durch Verschachteln der Verknüpfungen $\neg$, $\wedge$, $\vee$, $\Rightarrow$, $\Leftrightarrow$, xor und nand eine große Zahl von Verknüpfungen definieren kann. Allerdings sind nicht unbedingt alle dieser Verknüpfungen als verschieden zu betrachten. So erhält man z. B. durch Vertauschen der Aussagen A und B in einer Konjunktion eine Wahrheitstafel, bei der lediglich die ersten beiden Spalten vertauscht sind. Die letzte Spalte bleibt dagegen gleich. Die Verknüpfungen $A \wedge B$ und $B \wedge A$ nehmen also für *alle* Wahrheitswerte der Aussagen A und B *den gleichen Wahrheitswert an*. Man betrachtet sie daher als verschiedene Formulierungen ein- und derselben Aussage oder als logisch gleichwertig. Dieser Begriff lässt sich auch allgemeiner formulieren.

Definition 1.1.7: Seien für $n \in \{1, 2, 3....\}$ die Aussagen X und Y Verknüpfungen gegebener Aussagen A_1,..., A_n. Nehmen die Aussagen X und Y für alle Wahrheitswerte der Aussagen A_1,...,A_n jeweils den gleichen Wahrheitswert an, dann nennt man X und Y **logisch gleichwertig** und die Aussage $X \Leftrightarrow Y$ eine **Tautologie**.

Nehmen X und Y immer den gleichen Wahrheitswert an, so hat die Aussage $X \Leftrightarrow Y$ offensichtlich immer den Wahrheitswert w, unabhängig von den Wahrheitswerten der gegebenen Aussagen A_1, ..., A_n. Wir können also eine Tautologie als eine Aussage verstehen, die *immer*, *trivialerweise* und *a priori* wahr ist, d. h. unabhängig von den Wahrheitswerten irgendwelcher gegebener Aussagen.

Beispiel 1.1.8: Die Verknüpfungen $A \Leftrightarrow B$ und $(A\wedge B)\vee(\neg A\wedge\neg B)$ sind logisch gleichwertig, denn die Wahrheitstafel der Verknüpfung $(A \wedge B) \vee (\neg A \wedge \neg B)$ hat die Form

A	B	$\neg A$	$\neg B$	$A \wedge B$	$(\neg A \wedge \neg B)$	$(A \wedge B) \vee (\neg A \wedge \neg B)$
w	w	f	f	w	f	w
w	f	f	w	f	f	f
f	w	w	f	f	f	f
f	f	w	w	f	w	w.

Durch Vergleich der letzten Spalte mit der letzten Spalte der Wahrheitstafel für $A \Leftrightarrow B$ sieht man, dass diese übereinstimmen. Also ist $(A \Leftrightarrow B) \Leftrightarrow ((A \wedge B) \vee (\neg A \wedge \neg B))$ eine Tautologie.

Da es sehr viele Möglichkeiten gibt, die elementaren Verknüpfungen zu kombinieren, möchte man nicht nur einzelne Beispiele betrachten, sondern systematische Rechenregeln herleiten, die es einem erlauben, Ausdrücke mit ineinander verschachtelten elementare Verknüpfungen umzuformen und zu vereinfachen. Diese nehmen immer die Form von Tautologien an und können durch einen Vergleich der zugehörigen Wahrheitstafeln bewiesen werden.

Satz 1.1.9: (Rechenregeln für Verknüpfungen)

Für beliebige Aussagen A, B, C sind die folgenden Verknüpfungen Tautologien:

(i) **Doppelnegationsgesetz:** $\neg(\neg A) \Leftrightarrow A$,
(ii) **Kontrapositionsgesetz:** $(A \Rightarrow B) \Leftrightarrow ((\neg B) \Rightarrow (\neg A))$
(iii) **Kommutativgesetze:** $A \vee B \Leftrightarrow B \vee A$ und $A \wedge B \Leftrightarrow B \wedge A$,
(iv) **Assoziativgesetze:** $(A \vee B) \vee C \Leftrightarrow A \vee (B \vee C)$ und $(A \wedge B) \wedge C \Leftrightarrow A \wedge (B \wedge C)$
(v) **Distributivgesetze:** $A \vee (B \wedge C) \Leftrightarrow (A \vee B) \wedge (A \vee C)$ und $A \wedge (B \vee C) \Leftrightarrow (A \wedge B) \vee (A \wedge C)$
(vi) **de Morgansche Gesetze:** $\neg(A \wedge B) \Leftrightarrow (\neg A) \vee (\neg B)$ und $\neg(A \vee B) \Leftrightarrow (\neg A) \wedge (\neg B)$.

Beweis:
Wir stellen für beide Aussagen in den Tautologien Wahrheitstafeln auf und zeigen, dass die Spalten dieser Aussagen für alle Werte der Aussagen A und B übereinstimmen. Wir führen dies an den Beispielen des Doppelnegationsgesetzes und der de Morganschen Gesetze vor.
(i) Für das Doppelnegationsgesetz erhält man die Wahrheitstafel

A	$\neg A$	$\neg(\neg A)$
w	f	w
f	w	f

Da die erste und letzte Spalte der Tafel für alle möglichen Wahrheitswerte der Aussage A übereinstimmen, handelt es sich um eine Tautologie.

(vi) Die Wahrheitstafeln für das erste de Morgansche Gesetz nehmen die folgende Form an

A	B	$A \wedge B$	$\neg(A \wedge B)$
w	w	w	f
w	f	f	w
f	w	f	w
f	f	f	w

A	B	$\neg A$	$\neg B$	$(\neg A) \vee (\neg B)$
w	w	f	f	f
w	f	f	w	w
f	w	w	f	w
f	f	w	w	w

Da die letzten Spalten der zwei Tafeln für alle Werte der Aussagen A und B übereinstimmen, ist $\neg(A \wedge B) \Leftrightarrow (\neg A) \vee (\neg B)$ eine Tautologie. Für das zweite de Morgansche Gesetz erhält man die Wahrheitstafeln:

A	B	$A \vee B$	$\neg(A \vee B)$
w	w	w	f
w	f	w	f
f	w	w	f
f	f	f	w

A	B	$\neg A$	$\neg B$	$(\neg A) \wedge (\neg B)$
w	w	f	f	f
w	f	f	w	f
f	w	w	f	f
f	f	w	w	w

Da auch hier wieder die letzten Spalten der zwei Tafeln für alle Werte der Aussagen A und B übereinstimmen, ist $\neg(A \vee B) \Leftrightarrow (\neg A) \wedge (\neg B)$ eine Tautologie. □

Aufgabe 3: Beweisen Sie die Rechenregeln (ii) bis (v).

Wir zeigen nun, dass sich alle n-stelligen Verknüpfungen als verschachtelte elementare Verknüpfungen darstellen lassen. Genauer gesagt kommt man mit drei der fünf genannten elementaren Verknüpfungen aus, nämlich der Negation (NICHT), der Konjunktion (UND) und der Disjunktion (ODER).

Satz 1.1.10: Gegeben sei eine Wahrheitstafel mit 2^n Zeilen und $(n+1)$ Spalten, in deren ersten n Spalten die gegebenen Aussagen stehen und in deren letzter Spalte eine Verknüpfung dieser Aussagen steht. Die **disjunktive Normalform** dieser Verknüpfung ist die Verknüpfung, die aus der Wahrheitstafel durch die folgenden Schritte konstruiert wird.

- **Schritt 1:** Man streicht alle Zeilen, wo in der letzten Spalte der Wahrheitswert f steht.
- **Schritt 2:** Für jede verbleibende Zeile ersetzt man zuerst die Aussagen, die dort den Wahrheitswert f haben durch Ihre Negation und verknüpft dann mittels der Konjunktion alle gegebenen Aussagen. Die so entstehende Verknüpfung wird mit ihrem Wahrheitswerten als letzte Spalte an die Wahrheitstafel angefügt.
- **Schritt 3:** Hat man Schritt 2 für alle Zeilen durchgeführt, so klammert man die Ausdrücke in allen neu hinzugekommenen Spalten und verknüpft sie durch die Disjunktion.

Die disjunktive Normalform einer Verknüpfung ist zu der Verknüpfung logisch gleichwertig.
Beweis:
Ein Beweis dieses Satzes findet sich beispielsweise in [Si], Kapitel 1C, wo sich auch weitergehende Aussagen und Betrachtungen zur disjunktiven Normalform bereitgestellt werden. □

Beispiel 1.1.11: Wir betrachten die Äquivalenz $A \Leftrightarrow B$ und drücken sie mittels der disjunktiven Normalform durch die Verknüpfungen $\neg$ und $\wedge$ aus. Dazu stellen wir zuerst ihre Wahrheitstafel auf und wenden dann die einzelnen Schritte aus Satz 1.1.10 an.

$$\begin{array}{ccc} A & B & A\Leftrightarrow B \\ \hline w & w & w \\ w & f & f \\ f & w & f \\ f & f & w \end{array} \rightarrow \begin{array}{ccc} A & B & A\Leftrightarrow B \\ \hline w & w & w \\ f & f & w \end{array} \rightarrow \begin{array}{cccc} A & B & A\Leftrightarrow B & A\wedge B \\ \hline w & w & w & w \\ f & f & w & f \end{array}$$

$$\rightarrow \begin{array}{ccccc} A & B & A\Leftrightarrow B & A\wedge B & \neg A\wedge\neg B \\ \hline w & w & w & w & f \\ f & f & w & f & w \end{array} \rightarrow (A\wedge B)\vee(\neg A\wedge\neg B)$$

Also ist die Verknüpfung $A \Leftrightarrow B$ logisch gleichwertig zur Verknüpfung $(A \wedge B) \vee (\neg A \wedge \neg B)$.

Beispiel 1.1.12: Wir bestimmen die disjunktive Normalform der Verknüpfung $A \wedge (B \Rightarrow C)$.

$$\begin{array}{cccc} A & B & C & A\wedge(B\Rightarrow C) \\ \hline w & w & w & w \\ w & w & f & f \\ w & f & w & w \\ w & f & f & w \\ f & w & w & f \\ f & w & f & f \\ f & f & w & f \\ f & f & f & f \end{array} \rightarrow \begin{array}{cccc} A & B & C & A\wedge(B\Rightarrow C) \\ \hline w & w & w & w \\ w & f & w & w \\ w & f & f & w \end{array}$$

$$\rightarrow \begin{array}{ccccc} A & B & C & A\wedge(B\Rightarrow C) & A\wedge B\wedge C \\ \hline w & w & w & w & w \\ w & f & w & w & f \\ w & f & f & w & f \end{array}$$

$\rightarrow$

A	B	C	$A \wedge (B \Rightarrow C)$	$A \wedge B \wedge C$	$A \wedge \neg B \wedge C$
w	w	w	w	w	f
w	f	w	w	f	w
w	f	f	w	f	f

$\rightarrow$

A	B	C	$A \wedge (B \Rightarrow C)$	$A \wedge B \wedge C$	$A \wedge \neg B \wedge C$	$A \wedge \neg B \wedge \neg C$
w	w	w	w	w	f	f
w	f	w	w	f	w	f
w	f	f	w	f	f	w

$\rightarrow \quad (A \wedge B \wedge C) \vee (A \wedge \neg B \wedge C) \vee (A \wedge \neg B \wedge \neg C).$

Daraus erhält man, dass die Verknüpfung $A \wedge (B \Rightarrow C)$ logisch gleichwertig ist zur Verknüpfung $(A \wedge B \wedge C) \vee (A \wedge \neg B \wedge C) \vee (A \wedge \neg B \wedge \neg C)$.

Die disjunktive Normalform erlaubt es einem beliebige Verknüpfungen von Aussagen durch dazu gleichwertige Aussagen zu ersetzen, die nur aus $\neg$, $\wedge$, $\vee$ aufgebaut sind. In den so konstruierten Verknüpfungen stehen $\neg$ und $\wedge$ in den Klammern und $\vee$ zwischen den Klammern. Ebenso kann man zu einer gegebenen Verknüpfung aber auch die sogenannte *konjunktive Normalform* konstruieren, in denen aus $\neg$ und $\vee$ bestehende Ausdrücke in Klammern durch $\wedge$ zwischen den Klammern werden. Das geht ganz ähnlich.

Satz 1.1.13: Gegeben sei eine Wahrheitstafel mit 2^n Zeilen und $(n+1)$ Spalten, in deren ersten n Spalten die gegebenen Aussagen stehen und in deren letzter Spalte eine Verknüpfung dieser Aussagen steht. Die **konjunktive Normalform** dieser Verknüpfung ist die Verknüpfung, die aus der Wahrheitstafel wie folgt konstruiert wird.

- **Schritt 1:** Man streicht alle Zeilen, wo in der letzten Spalte der Wahrheitswert w steht.
- **Schritt 2:** Für jede verbleibende Zeile ersetzt man zuerst die Aussagen, die dort den Wahrheitswert w haben durch ihre Negation und verknüpft dann mittels der Disjunktion alle gegebenen Aussagen. Die so entstehende Verknüpfung wird mit ihrem Wahrheitswerten als letzte Spalte an die Wahrheitstafel angefügt.
- **Schritt 3:** Hat man Schritt 2 für alle Zeilen durchgeführt, so klammert man die Ausdrücke in allen neu hinzugekommenen Spalten und verknüpft sie durch die Konjunktion.

Die konjunktive Normalform einer Verknüpfung ist zu der Verknüpfung logisch gleichwertig.

Beweis:
Ein Beweis dieses Satzes findet sich beispielsweise in [Si], Kapitel 1C, wo sich auch weitergehende Aussagen und Betrachtungen zur konjunktiven Normalform bereitgestellt werden. □

Beispiel 1.1.14: Wir bestimmen die konjunktive Normalform der Äquivalenz $A \Leftrightarrow B$.

A	B	$A \Leftrightarrow B$
w	w	w
w	f	f
f	w	f
f	f	w

$\rightarrow$

A	B	$A \Leftrightarrow B$
w	f	f
f	w	f

$\rightarrow$

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

$$\rightarrow \quad \begin{array}{cccccc} A & B & A \Leftrightarrow B & \neg A \vee B & A \vee \neg B \\ \hline w & f & f & f & w \\ f & w & f & w & f \end{array} \quad \rightarrow \quad (\neg A \vee B) \wedge (A \vee \neg B)$$

Also ist die Verknüpfung $A \Leftrightarrow B$ logisch gleichwertig zur Verknüpfung $(\neg A \vee B) \wedge (A \vee \neg B)$.

Wir haben Verknüpfungen von Aussagen eingeführt, die sich mittels Wahrheitstafeln aus gewissen vorgegebenen Aussagen $A_1, ..., A_n$ konstruieren lassen. Stattdessen können wir eine n-stellige Verknüpfung von Aussagen A_1, ..., A_n aber auch als eine Zuordnung auffassen, die jeder geordneten Liste von n Wahrheitswerten - die ersten n Einträge einer Zeile der Wahrheitstafel - einen Wahrheitswert zuordnet - den Wahrheitswert der Verknüpfung im letzten Eintrag dieser Zeile. Bezeichnen wir mit $\{w, f\}^{\times n}$ die Menge der geordneten n-stelligen Listen oder n-Tupel von Wahrheitswerten, so können wir eine Verknüpfung q also auffassen als eine Abbildung $q : \{w, f\}^{\times n} \to \{w, f\}$.

Dieses Prinzip lässt sich verallgemeinern. Wir können statt der Menge von geordneten n-Tupeln von Wahrheitswerten beliebige Sammlungen mathematischer Gebilde betrachten zusammen mit Zuordnungen, die jedem Gebilde dieser Sammlung einen Wahrheitswert zuordnen. Die Gebilde in einer Sammlung nennen wir dann *Elemente* und die Zuordnung *Aussageformen* oder *Prädikate*. Dabei kann man auch aus einer Sammlung M auch immer eine neue Sammlung $M^{\times n}$ bilden, indem man geordnete Listen von n Elementen aus der Sammlung M betrachtet. Ein Prädikat auf der Sammlung $M^{\times n}$ solcher Listen nennt man auch *n-stelliges Prädikat* auf M.

Definition 1.1.15:

1. Eine **Aussageform** oder ein **Prädikat** auf M ist eine Zuordnung $P : M \to \{w, f\}$ eines Wahrheitswerts w oder f zu jedem Element in M.

2. Eine **n-stellige Aussageform** oder ein **n-stelliges Prädikat** auf M ist eine Zuordnung $P : M^{\times n} \to \{w, f\}$ von Wahrheitswerten zu geordneten Listen von n Elementen aus M.

Beispiel 1.1.16:

1. Wir bezeichnen mit $\mathbb{N} = \{1, 2, 3, ...\}$ die Menge der natürlichen Zahlen und betrachten das Prädikat $P : \{1, 2, 3...\} \to \{w, f\}$, das die Eigenschaft *“Die Zahl $n \in \mathbb{N}$ ist eine Primzahl.”* beschreibt. Er ordnet einer Zahl $n \in \mathbb{N}$ den Wahrheitswert w zu, wenn n eine Primzahl ist und ansonsten den Wahrheitswert f. Man schreibt auch
$$P(n) = \begin{cases} w & n \text{ Primzahl} \\ f & \textit{sonst.} \end{cases}$$
2. Eine n-stellige Verknüpfung ist ein Prädikat $P : \{w, f\}^{\times n} \to \{w, f\}$.
3. Die zweistellige Verknüpfung $A \wedge B$ ist ein Prädikat $P : \{w, f\}^{\times 2} \to \{w, f\}$, das dem Tupel (w, w) den Wahrheitswert w und den Tupeln (w, f), (f, w), (f, f) den Wahrheitswert f zuordnet.
4. Die zweistellige Verknüpfung $\neg A \Rightarrow B$ ist ein Prädikat $P : \{w, f\}^{\times 2} \to \{w, f\}$, das den Tupeln (f, w), (w, w) und (w, f) den Wahrheitswert w und dem Tupel (f, f) den Wahrheitswert f zuordnet.

Wie auch im Fall der Verknüpfungen kann man Prädikate mittels Negationen, Konjunktionen, Disjunktionen, Implikationen und Äquivalenzen kombinieren und erhält so neue Prädikate. Dabei lässt man die Verknüpfungen auf die zugeordneten Wahrheitswerte wirken.

Definition 1.1.17: Seien $P, Q : M \to \{w, f\}$ Prädikate. Dann definieren wir die folgenden Prädikate durch ihre Wahrheitswerte

1. Das Prädikat $\neg P : M \to \{w, f\}$ ordnet $x \in M$ den Wahrheitswert $\neg P(x)$ zu.
2. Das Prädikat $P \wedge Q : M \to \{w, f\}$ ordnet $x \in M$ den Wahrheitswert $P(x) \wedge Q(x)$ zu.
3. Das Prädikat $P \vee Q : M \to \{w, f\}$ ordnet $x \in M$ den Wahrheitswert $P(x) \vee Q(x)$ zu.
4. Das Prädikat $P \Rightarrow Q : M \to \{w, f\}$ ordnet $x \in M$ den Wahrheitswert $P(x) \Rightarrow Q(x)$ zu.
5. Das Prädikat $P \Leftrightarrow Q : M \to \{w, f\}$ ordnet $x \in M$ den Wahrheitswert $P(x) \Leftrightarrow Q(x)$ zu.

Wir haben bereits gezeigt, dass Verknüpfungen als Prädikate aufgefasst werden können. Umgekehrt können wir aus Prädikaten Verknüpfungen konstruieren, benötigen dazu aber zusätzliche Struktur, nämlich die sogenannten **Quantoren**. Die zwei wichtigsten sind der **Existenzquantor**, der mit $\exists$ bezeichnet wird und den Sätzen *" es gibt (mindestens) ein ..."* oder *"es existiert (mindestens) ein ..."* entspricht und der **Allquantor**, der mit $\forall$ bezeichnet wird und dem Satz *"für alle ...gilt"* entspricht. Es gibt auch noch den Quantor $\nexists$, der oft auch mit $\neg\exists$ bezeichnet wird und *"es gibt kein ..."* bedeutet und den Quantor $\exists!$, der für *"es gibt genau ein ..."* steht.

Definition 1.1.18: Sei $P : M \to \{w, f\}$ ein Prädikat. Dann definieren wir die folgenden Aussagen:

1. Die Aussage $(\forall x \in M : P(x))$ hat den Wahrheitswert w, wenn $P(x) = w$ für alle $x \in M$ gilt und ansonsten den Wahrheitswert f.
2. Die Aussage $(\neg\forall x \in M : P(x))$ hat den Wahrheitswert w, wenn es (mindestens) ein $x \in M$ mit $P(x) = f$ gibt und ansonsten den Wahrheitswert f.
3. Die Aussage $(\exists x \in M : P(x))$ hat den Wahrheitswert w, wenn es (mindestens) ein $x \in M$ mit $P(x) = w$ gibt und ansonsten den Wahrheitswert f.
4. Die Aussage $(\exists! x \in M : P(x))$ hat den Wahrheitswert w, wenn es genau ein $x \in M$ mit $P(x) = w$ gibt und ansonsten den Wahrheitswert f.

Beispiel 1.1.19: Für das Prädikat

$$P : \mathbb{N} \to \{w, f\}, \qquad P(n) = \begin{cases} w & n \text{ Primzahl} \\ f & \text{sonst} \end{cases}$$

gilt $(\forall n \in \mathbb{N} : P(n)) = f$, denn nicht jede Zahl $n \in \mathbb{N}$ ist eine Primzahl. Ausserdem gilt $(\neg\forall n \in \mathbb{N} : P(n)) = w$, denn es gibt eine Zahl, die keine Primzahl ist, $(\exists n \in \mathbb{N} : P(n)) = w$, denn es gibt (mindestens) eine Zahl in $\mathbb{N}$, die eine Primzahl ist, und $(\exists! x \in \mathbb{N} : P(n) = w) = f$, denn es gibt mehr als eine Primzahl.

Satz 1.1.20: Rechenregeln für Quantoren

Seien $P, Q : M \to \{w, f\}$ Prädikate und q eine Aussage. Dann sind die folgenden Paare von Aussagen logisch gleichwertig:

(i) $(\neg\forall x \in M : P(x))$ und $(\exists x \in M : \neg P(x))$
(ii) $(\neg\exists x \in M : P(x))$ und $(\forall x \in M : \neg P(x))$
(iii) $(\exists x \in M : (P \vee Q)(x))$ und $(\exists x \in M : P(x)) \vee (\exists x \in M : Q(x))$
(iv) $(\forall x \in M : (P \wedge Q)(x))$ und $(\forall x \in M : P(x)) \wedge (\forall x \in M : Q(x))$
(v) $(\forall x \in M : q \vee P(x))$ und $q \vee (\forall x \in M : P(x))$
(vi) $(\exists x \in M : q \wedge P(x))$ und $q \wedge (\exists x \in M : P(x))$

(vii) $(\forall x \in M : P(x)) \Rightarrow q$ und $(\exists x \in M : (P(x) \Rightarrow q))$
(viii) $(\exists x \in M : P(x)) \Rightarrow q$ und $(\forall x \in M : (P(x) \Rightarrow q))$
(ix) $q \Rightarrow (\forall x \in M : P(x))$ und $(\forall x \in M : (q \Rightarrow P(x)))$
(x) $q \Rightarrow (\exists x \in M : P(x))$ und $(\exists x \in M : (q \Rightarrow P(x)))$.

Beweis:
Wir beweisen Aussage (i) und (vii) als Beispiele für das allgemeine Vorgehen.

(i) Wenn $(\neg\forall x \in M : P(x))$ den Wahrheitswert w hat, dann gibt es nach Definition 1.1.18 ein $y \in M$ mit $P(y) = f$, was gleichbedeutend ist zu $\neg P(y) = w$. Also hat $(\exists x \in M : \neg P(x))$ ebenfalls den Wahrheitswert w.

Hat aber $(\neg\forall x \in M : P(x))$ den Wahrheitswert f, dann gilt $P(x) = w$ für alle $x \in M$ und somit $\neg P(x) = f$ für alle $x \in M$. Also gibt es kein $x \in M$ mit $\neg P(x) = w$ und somit hat $(\exists x \in M : \neg P(x))$ den Wahrheitswert f. Damit stimmen die Wahrheitstafeln der beiden Aussagen überein, und die Aussagen sind logisch gleichwertig.

(vii) Ist q wahr, dann sind sowohl die rechte als auch die linke Aussage immer wahr. Sei also jetzt q falsch.

1. Ist $(\forall x \in M : P(x)) \Rightarrow q$ wahr, dann muss es wegen $q = f$ ein $y \in M$ geben mit $P(y) = f$, da ja sonst q wahr sein müsste. Für dieses $y \in M$ ist dann auch $P(y) \Rightarrow q$ wahr und somit ist $(\exists x \in M : (P(x) \Rightarrow q))$ wahr.

2. Ist umgekehrt $(\exists x \in M : (P(x) \Rightarrow q))$ wahr, dann existiert wegen q falsch ein $y \in M$, so dass $P(y) \Rightarrow q$ wahr ist. Da q falsch ist, muss dann auch $P(y)$ falsch sein. Also ist $(\forall x \in M : P(x))$ falsch und somit ist $(\forall x \in M : P(x)) \Rightarrow q$ wahr.

3. In 1. und 2. wurde gezeigt, dass aus der Wahrheit der Aussage $(\forall x \in M : P(x)) \Rightarrow q)$ folgt, dass auch die Aussage $(\exists x \in M : (P(x) \Rightarrow q))$ wahr ist und aus der Wahrheit der Aussage $(\exists x \in M : (P(x) \Rightarrow q))$, folgt, dass auch $(\forall x \in M : P(x)) \Rightarrow q)$ wahr ist. Somit haben die Aussagen $(\forall x \in M : P(x)) \Rightarrow q)$ und $(\exists x \in M : (P(x) \Rightarrow q))$ immer den gleichen Wahrheitswert und sind somit logisch gleichwertig. $\square$

Bemerkung 1.1.21: Ist $M = \emptyset$ die leere Menge, so ist eine Aussage der Form $(\forall x \in M : ...)$ immer wahr und eine Aussage der Form $(\exists x \in M : ...)$ immer falsch. Dies ist jedoch *kein Widerspruch* zu den Rechenregeln in Satz 1.1.20 (vii) und (viii), denn ein Prädikat $P : \emptyset \to \{w, f\}$ ordnet *keinem Element* $x \in M$ einen Wahrheitswert zu. Das heißt die Ausdrücke $P(x)$ in (vii) und (viii) existieren nicht, und es wird keine Aussage über Elemente $x \in M$ gemacht.

Aufgabe 4: Beweisen Sie die Rechenregeln (ii)-(vi) und (viii)-(x).

Aufgabe 5: Der Computer meldet: *"Nicht alle Ihrer Dateien werden nicht gelöscht."*. Was passiert mit den Dateien?

Es ist nützlich, die gängigsten sprachlichen Formulierungen zu kennen, die die Erzeugung einer Aussage aus einem Prädikat mittels des Allquantors oder des Existenzquantors beschreiben. Da ein mathematischer Text, der vorwiegend aus geklammerten logischen Aussagen bestünde, sehr unübersichtlich wäre, werden die Aussagen der Form $\forall x \in M : P(x)$ oder $\exists x \in M : P(x)$ in mathematischen Definitionen, Sätzen und Beweisen häufig durch die folgenden sprachlichen Formulierungen ersetzt.

$\forall x \in M : P(x)$

- Für alle $x \in M$ gilt $P(x)$.
- Für jedes Element $x \in M$ gilt $P(x)$.
- Für ein beliebiges Element $x \in M$ gilt $P(x)$.
- Sei $x \in M$ (beliebig). Dann gilt $P(x)$.
- Ist $x \in M$, dann/so gilt $P(x)$.
- Wenn $x \in M$, dann folgt $P(x)$.
- Jedes Element von M erfüllt P.
- Alle Elemente von M erfüllen P.

$\exists x \in M : P(x)$

- Es gibt (mindestens) ein $x \in M$ mit $P(x)$.
- Es existiert (mindestens) ein $x \in M$ mit $P(x)$.
- Die Menge M hat ein Element x, das P erfüllt.
- Ein Element von M erfüllt P.
- Für ein geeignetes Element $x \in M$ gilt $P(x)$.
- Man kann ein $x \in M$ wählen, so dass $P(x)$ gilt.

Aufgabe 6: Sei M die Menge aller Menschen und N die Menge aller geordneten Paare von Menschen. Wir betrachten das logische Prädikat $P : N \to \{w, f\}$, das den Wert $P(a, b) = w$ annimmt, wenn der Mensch a den Menschen b liebt, und den Wert f, wenn das nicht der Fall ist. Übersetzen Sie die folgenden Aussagen in Alltagssprache:

(a) $\forall a \in M : \exists b \in M : P(a, b)$,
(b) $\forall a \in M : \exists b \in M : P(b, a)$,
(c) $\exists a \in M : \forall b \in M : P(a, b)$,
(d) $\exists a \in M : \forall b \in M : P(b, a)$,
(e) $\forall a \in M : (\exists b \in M : P(a, b)) \wedge (\exists c \in M : P(c, a))$,
(f) $\forall a \in M : \exists b \in M : P(a, b) \wedge P(b, a)$,
(g) $\forall a \in M : (\exists b \in M : P(a, b)) \vee (\exists c \in M : P(c, a))$,
(h) $\forall a \in M : \exists b \in M : P(a, b) \vee P(b, a)$,
(i) $\nexists a \in M : \forall b \in M : P(b, a)$,
(j) $\nexists a \in M : \nexists b \in M : P(b, a)$,
(k) $\nexists a \in M : \exists b \in M : P(a, b) \wedge P(b, a)$,
(l) $\forall a \in M : \exists! b \in M : P(a, b) \wedge P(b, a)$,
(m) $\exists a \in M : \nexists b \in M : P(b, a)$,
(n) $\exists! a \in M : \forall b \in M : P(a, b) \vee P(b, a)$,
(o) $\forall a \in M : \forall b \in M : P(a, b) \Rightarrow P(b, a)$.

1.2 Beweistechniken

Ein großer Teil der Mathematikvorlesungen ist dem Beweisen gewidmet. Anders als in der Schule werden im Mathematikstudium mathematische Aussagen immer bewiesen. Dazu reicht es nicht, einige Spezialfälle oder Beispiele zu untersuchen und festzustellen, dass diese Aussagen in diesen Fällen wahr sind. Dies macht es zwar plausibel, dass die Aussage gelten *könnte*, erlaubt es aber nicht, das zwingend zu folgern, da es ja immer möglich ist, dass man ein Gegenbeispiel übersehen oder vergessen hat.

Beweisen besteht also *nicht* aus dem Betrachten von Beispielen, auch wenn das oft eine nützliche Aufwärmübung ist, sondern daraus, die zu beweisenden Aussagen logisch zwingend auf andere Aussagen zurückzuführen, deren Wahrheit bekannt ist. Damit ein Beweis gelingt, ist es wichtig, klar, präzise und logisch stimmig zu argumentieren. Dabei ist es hilfreich, die folgenden Grundtechniken zu kennen.

Ein **konstruktiver Beweis** zeigt, dass ein Objekt mit bestimmten Eigenschaften existiert, indem entweder ein solches Objekt direkt angegeben wird oder eine Konstruktionsvorschrift, die es einem erlaubt, ein solches Objekt zu konstruieren. Ein **nicht konstruktiver Beweis** zeigt nur die Existenz eines solchen Objektes, ohne aber ein solches Objekt oder eine Konstruktionsvorschrift dafür anzugeben. Beispiele von konstruktiven Beweisen sind die folgenden.

Lemma 1.2.1: Gegeben ist ein Polynom $p = x^2 - 2px + q$ mit rationalen Koeffizienten p, q, so dass $p^2 - q > 0$ gilt. Dann hat dieses Polynom zwei verschiedene reelle Nullstellen.

Beweis:
Da $p^2 - q > 0$ ist $\sqrt{p^2 - q}$ definiert und $\sqrt{p^2 - q} > 0$. Also sind $y_+ = p + \sqrt{p^2 - q}$ und $y_- = p - \sqrt{p^2 - q}$ zwei verschiedene reelle Zahlen. Außerdem gilt

$$\begin{aligned} y_\pm^2 - 2py_\pm + q &= (p \pm \sqrt{p^2 - q})^2 - 2p(p \pm \sqrt{p^2 - q}) + q \\ &= p^2 + p^2 - q \pm 2p\sqrt{p^2 - q} \mp 2p\sqrt{p^2 - q} - 2p^2 + q = 0. \end{aligned}$$

Also sind y_+ und y_- zwei verschiedene reelle Nullstellen von p. □

Lemma 1.2.2: Das Quadrat einer geraden ganzen Zahl ist gerade.

Beweis:
Sei n eine gerade ganze Zahl. Da n gerade ist, gibt es eine ganze Zahl k mit $n = 2 \cdot k$. Daraus folgt $n^2 = (2 \cdot k)^2 = 4 \cdot k^2$ und somit ist $n^2 = 2 \cdot (2 \cdot k^2)$, wobei $2 \cdot k^2$ ganz ist, da k und 2 ganz sind. Also ist n^2 gerade. □

Die Beweise von Lemma 1.2.1 und Lemma 1.2.2 sind konstruktiv, da im Beweis von Lemma 1.2.1 die zwei verschiedenen Nullstellen explizit konstruiert werden und im Beweis von Lemma 1.2.2 eine ganze Zahl s mit $n = 2 \cdot s$ explizit angegeben wird, nämlich $s = 2 \cdot k^2$. Ein Beispiel für einen nicht-konstruktiven Beweis ist der folgende.

Lemma 1.2.3: Das Polynom $p = x^5 - x^4 + 2x^3 + x - 1$ hat eine Nullstelle im offenen Intervall $(-1, 1)$.

Beweis:
Es gilt $p(1) = 1 - 1 + 2 + 1 - 1 = 2$ und $p(-1) = -1 - 1 - 2 - 1 - 1 = -6$. Da also $p(1) > 0$ ist und $p(-1) < 0$ und Polynome stetig sind, muss es nach dem Mittelwertsatz eine Nullstelle im Intervall (-1,1) geben. □

Dieser Beweis ist nicht-konstruktiv, da er weder eine Nullstelle angibt noch eine Vorschrift zur Konstruktion einer Nullstelle. Man kann sogar beweisen, dass es für Polynome vom Grad ≥ 5 mit rationalen Koeffizienten eine solche Vorschrift nicht gibt, d. h. es ist *prinzipiell* unmöglich, eine Formel anzugeben, die die Nullstellen eines beliebigen Polynoms von Grad ≥ 5 mit rationalen Koeffizienten durch seine Koeffizienten ausdrückt.

Während ein konstruktiver Beweis also den Vorteil hat, dass er mehr Information enthält - er zeigt nicht nur, dass das gesuchte Objekt existiert sondern auch noch wie es konkret aussieht - sind nicht-konstruktive Beweise allgemeiner. Sie lassen sich auch in solchen Situationen führen, in denen es unmöglich oder sehr aufwändig wäre, das gesuchte Objekt explizit zu konstruieren.

Der **indirekte Beweis** oder **Widerspruchsbeweis** einer Aussage B wird geführt, indem man zunächst annimmt, die Aussage B sei falsch. Dies führt man dann zu einem Widerspruch. Man folgert mit Hilfe der Aussagenlogik, dass dann auch eine weitere Aussage A falsch sein müsste, von der aber bekannt ist, dass sie wahr ist. Somit muss die ursprüngliche Annahme falsch und die Aussage B wahr sein. Formal gesprochen zeigt man also, dass $\neg B \Rightarrow \neg A$, was logisch gleichwertig ist zu der Aussage $A \Rightarrow B$. Ist A wahr, so muss dann auch B wahr sein.

Beim indirekten Beweisen ist es daher essentiell, dass man mit Aussagenlogik, insbesondere der Verneinung von Aussagen, souverän umgehen kann. Ausserdem muss man stets überprüfen, ob die Aussage, zu der ein Widerspruch erzeugt wird, tatsächlich wahr ist. Wir betrachten nun einige Beispiele indirekter Beweise. Dabei setzen wir die Wahrheit der Aussagen voraus, dass sich jede rationale Zahl als gekürzter Bruch beschreiben lässt und dass sich jede natürliche Zahl außer der 1 in Primfaktoren zerlegen lässt. Sie werden in den Grundvorlesungen bewiesen.

Satz 1.2.4: $\sqrt{2}$ ist irrational: es gibt keine rationale Zahl q mit $q^2 = 2$.

Beweis:
Angenommen es gibt eine rationale Zahl q mit $q^2 = 2$. Dann können wir q als gekürzten Bruch ausdrücken $q = a/b$ mit teilerfremden ganzen Zahlen a, b. Aus $q^2 = 2$ folgt dann $a^2 = 2b^2$. Also teilt 2 das Produkt $a^2 = a \cdot a$, und da 2 eine Primzahl ist, muss sie einen der Faktoren im Produkt teilen. Also ist a gerade, und es gibt eine ganze Zahl c mit $a = 2c$. Einsetzen in die Gleichung $a^2 = 2b^2$ ergibt dann $b^2 = 2c^2$. Also teilt 2 das Produkt $b^2 = b \cdot b$, b ist gerade und es gibt eine ganze Zahl d mit $b = 2d$. Dies ist ein Widerspruch zur Annahme, dass der Bruch $q = a/b$ gekürzt ist, denn wenn $a = 2c$ und $b = 2d$ mit ganzen Zahlen c, d, so ist a/b nicht gekürzt. Also kann es keine rationale Zahl q mit $q^2 = 2$ geben. □

Satz 1.2.5 (Satz von Euklid)**:** Es gibt unendlich viele Primzahlen.

Beweis:
Angenommen es gibt nur n Primzahlen $p_1,..., p_n \in \mathbb{N}$. Dann betrachten wir die Zahl $m = p_1 \cdot ... \cdot p_n + 1$. Offensichtlich ist $m > 1$. Da sich jede natürliche Zahl außer der 1 als Produkt von Primfaktoren schreiben lässt, muss es also mindestens eine Primzahl geben, die m teilt. Dies kann aber nicht gelten, da Teilen von m durch *jede* Primzahl $p_1, ..., p_n$ den Rest 1 ergibt. Also ergibt sich ein Widerspruch zur Existenz der Primfaktorzerlegung, und es muss unendlich viele Primzahlen geben. □

Aufgabe 7: Beweisen Sie mit einem indirekten Beweis, dass das Quadrat einer ungeraden Zahl ungerade ist.

Ein weiteres wichtiges Beweisverfahren ist die **vollständige Induktion**. Diese kommt oft dann zur Anwendung, wenn eine Aussage der Form $(x \in \mathbb{N} : P(x))$ für ein Prädikat $P : \mathbb{N} \to \{w, f\}$ bewiesen werden soll. Beispiele solcher Aussagen sind oft Formeln oder Ungleichungen, deren Gültigkeit für alle natürlichen Zahlen (oder eine unendliche Teilmenge M davon) bewiesen werden soll. Dabei geht man wie folgt vor. Man zeigt zunächst, dass die Aussage für die kleinste natürliche Zahl in der Menge M wahr ist. Dies bezeichnet man als den **Induktionsanfang**. Danach zeigt man, dass aus $P(n) = w$ für eine Zahl $n \in M$ folgt, dass auch $P(m) = w$ für die nächstgrößere Zahl $m \in M$, also dass die **Induktionsbehauptung** wahr ist. Dies nennt man den **Induktionsschritt**. Er beweist die Wahrheit der Aussage für alle $m \in M$.

Lemma 1.2.6: Für alle natürlichen Zahlen n gilt

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

Beweis:
Beweis mit vollständiger Induktion.
1. Induktionsanfang: Die Aussage gilt für $n = 1$, denn $1 = (1 \cdot 2)/2$.
2. Induktionsschritt: Angenommen die Aussage gilt für die natürliche Zahl n. Dann ergibt sich für die Zahl $n + 1$

$$\begin{aligned} 1 + 2 + ... + n + n + 1 &= (1 + 2 + ... + n) + (n + 1) = \frac{n(n+1)}{2} + n + 1 \\ &= \frac{n(n+1) + 2(n+1)}{2} = \frac{(n+1)(n+2)}{2}. \end{aligned}$$

Also gilt dann die Aussage auch für die natürliche Zahl $n + 1$. Damit folgt, dass die Aussage für alle $n \in \mathbb{N}$ wahr ist. □

Lemma 1.2.7: Für alle natürlichen Zahlen n gilt

$$\sum_{k=0}^{n} k(k+1) = 0 \cdot 1 + 1 \cdot 2 + 2 \cdot 3 + ... + n(n+1) = \tfrac{1}{3} n(n+1)(n+2)$$

Beweis:
Beweis mit vollständiger Induktion.
1. Induktionsanfang: Die Aussage gilt für $n = 1$, denn $\sum_{k=0}^{1} k(k+1) = 0 \cdot 1 + 1 \cdot 2 = 2 = \frac{1}{3} \cdot 1 \cdot 2 \cdot 3$.
2. Induktionsschritt: Angenommen die Aussage gilt für die natürliche Zahl n. Dann ergibt sich für die Zahl $n + 1$

$$\begin{aligned} \sum_{k=0}^{n+1} k(k+1) &= 1 \cdot 2 + ... + n(n+1) + (n+1)(n+2) = \sum_{k=0}^{n} k(k+1) + (n+1)(n+2) \\ &= \tfrac{1}{3} n(n+1)(n+2) + (n+1)(n+2) = \tfrac{1}{3}(n+1)(n+2)(n+3) \end{aligned}$$

Also gilt die Aussage auch für $n + 1$ und somit für alle $n \in \mathbb{N}$. □

1.3 Mengenlehre - die Zermelo-Fraenkel Axiome

Neben der Logik ist die zweite Säule der Mathematik die Mengenlehre. Die Mengenlehre befasst sich mit *Mengen*, die man sich als eine Art Behälter vorstellen kann, die mit gewissen mathematischen Dingen gefüllt werden, den *Elementen*. Ein zentraler Punkt dabei ist, dass auch die Elemente selbst wieder Behälter sein können.

Der erste Versuch, dieses Konzept mathematisch zu formulieren war die naive Mengendefinition von Georg Cantor (1845-1918).

"Definition" 1.3.1 (Cantor)**:** "Eine Menge ist eine Zusammenfassung bestimmter, wohlunterschiedener Objekte unserer Anschauung oder unseres Denkens zu einem Ganzen."

Diese naive Definition einer Menge ist relativ nichtssagend, unpräzise und hat sich später als inadäquat herausgestellt. Sie enthält jedoch schon einige wichtige Eigenschaften von Mengen:

- Eine Menge besteht aus *Elementen* - den *Objekten unserer Anschauung oder unseres Denkens*. Diese Elemente können eine sehr allgemeine Gestalt annehmen - prinzipiell alles, was wir wahrnehmen oder denken können - und können insbesondere auch selbst wieder Mengen sein.
- Sie sind *bestimmt* - es lässt sich eindeutig entscheiden, ob etwas ein Element einer Menge ist oder nicht.
- Sie sind *wohlunterschieden* - es kommt nicht mehrmals das gleiche Element in einer Menge vor - und die Menge (das *Ganze*) ist durch ihre Elemente eindeutig charakterisiert.
- Obwohl dies in der Definition nicht explizit erwähnt wird, enthielt der damalige Mengenbegriff auch das Konzept der *leeren Menge* - eine Menge, die keine Elemente enthält.

Übertragen wir die Cantorsche Definition in eine moderne Sprache, so erhalten wir:

"Definition" 1.3.2:

1. Eine **Menge** ist etwas, das **Elemente** enthalten kann. Ist a ein Element einer Menge M, so schreibt man $a \in M$. Ist a kein Element der Menge M, so schreibt man $a \notin M$. Für eine endliche Menge M, die genau die Elemente a_1, a_2, ..., a_n enthält, schreibt man $M = \{a_1, a_2, ..., a_n\}$.
2. Es gibt eine ausgezeichnete Menge, die **leere Menge** $\emptyset$, die keine Elemente enthält.
3. Eine Menge ist durch ihre Elemente M eindeutig bestimmt. Zwei Mengen M, N sind **gleich** genau dann, wenn sie die gleichen Elemente enthalten:
$$(a \in M \Rightarrow a \in N) \wedge (a \in N \Rightarrow a \in M).$$
Man schreibt dann $M = N$ und ansonsten $M \neq N$.

Bemerkung 1.3.3:

1. Oft wird eine analoge Schreibweise $M = \{a_1, a_2, ...\}$ für unendliche Mengen benutzt, die aber eher als Abkürzung bzw. Konstruktionsvorschrift verstanden werden sollte.
2. Aus der letzten Aussage in der Definition ergibt sich:
   - Jedes Element ist nur einmal in einer Menge enthalten:
   $\{a_1, a_1, a_2, ..., a_n\} = \{a_1, a_2, ..., a_n\}$ und analog für unendliche Mengen. Denn sonst wäre eine Menge M durch die Angabe ihrer Elemente *nicht* eindeutig bestimmt; man müsste dann zusätzlich angeben, *wie oft* jedes Element in M enthalten ist.
   - Es kommt nicht auf die Reihenfolge der Elemente in einer Menge an:
   $\{a_1, a_2, a_3, ..., a_n\} = \{a_2, a_1, a_3, ..., a_n\} = \{a_3, a_1, a_2, ..., a_n\} = ...$ und analog für unendliche Mengen. Denn ansonsten müsste man nicht nur die Elemente einer Menge M angeben, um M festzulegen, sondern auch deren *Reihenfolge*.

Cantors Mengenbegriff ist auch in der Formulierung aus "Definition" 1.3.2 inadäquat. Sein Hauptproblem ist die **Russelsche Antinomie**. Diese beruht auf der (sogenannten) Menge R aller Mengen, die nicht in sich selbst als Element enthalten sind, $R = \{x | x \notin x\}$. Die Frage, ob die Menge R sich selbst als Element enthält, also ob $R \in R$ gilt, führt zu einem logischen Widerspruch: gilt $R \in R$, so folgt aus der Definition der Menge R, dass $R \notin R$. Umgekehrt folgt aus $R \notin R$ mit der Definition der Menge R, dass $R \in R$ gelten muss.

Dieser Widerspruch erschütterte die Mathematik in ihren Grundfesten und löste die sogenannte Grundlagenkrise der Mathematik aus. Im Laufe der Grundlagenkrise wurde der Cantorsche Mengenbegriff schließlich durch eine präziseren Mengenbegriff ersetzt - die heute gebräuchliche Definition der Menge durch die *Zermelo-Fraenkel Axiome*, inklusive des *Auswahlaxioms*.

Das Prinzip, mit dem diese Mengendefinition Unklarheiten und Paradoxa vermeidet, ist, dass eine Menge einfach als etwas definiert wird, was sich durch bestimmte Vorschriften aus anderen Mengen konstruieren lässt. Dadurch lassen sich alle Mengen letztendlich auf eine Menge zurückführen, nämlich die *leere Menge*. Ähnlich wie eine Spielfigur in einem Brettspiel, die ganz beliebig aussehen kann und letztendlich einfach etwas ist, was sich nach den Spielregeln auf bestimmte Weise auf dem Brett bewegen darf, *ist* eine Menge etwas, das man durch bestimmte Konstruktionen aus der leeren Menge herstellen kann.

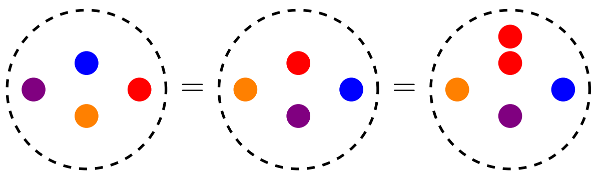
Zu diesem Zeitpunkt fehlt uns der Hintergrund, die Übung im Umgang mit mathematischen Konzepten und die Zeit, um diese Axiome vollständig und präzise zu erfassen, aber die folgende Definition kann zumindest einen Eindruck davon vermitteln und andeuten, wie diese die Probleme des Cantorschen Mengenbegriffs vermeiden. Um die Definition zu verstehen, ist es sinnvoll, sich eine Menge als einen Behälter vorzustellen, der entweder leer sein kann, oder gewisse Dinge enthält, die auch selbst wieder Behälter sein können.

Definition 1.3.4 (**Zermelo-Fraenkel Axiome**): Eine **Menge** ist etwas, das **Elemente** enthalten kann und durch die folgenden zehn Axiome charakterisiert ist.

1. **Bestimmtheitsaxiom:** Zwei Mengen M, N sind **gleich**, wenn $a \in M \Rightarrow a \in N$ und $a \in N \Rightarrow a \in M$ gilt. Man schreibt dann $M = N$ und ansonsten $M \neq N$. Eine Menge N heisst **Teilmenge** einer Menge M, in Zeichen: $N \subseteq M$, wenn $a \in N \Rightarrow a \in M$ gilt.

   Formel: $M = N \;\Leftrightarrow\; (a \in M \Rightarrow a \in N) \wedge (a \in N \Rightarrow a \in M)$

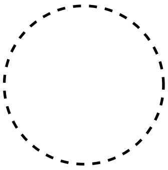
   Bild: *Ein Behälter ist durch seinen Inhalt eindeutig bestimmt. Zwei Behälter werden als gleich betrachtet, wenn sie den gleichen Inhalt haben. Insbesondere ist der Inhalt der Behälter nicht in irgendeiner Weise geordnet und jedes Ding ist maximal einmal in einem Behälter enthalten.*

2. **Axiom der leeren Menge:** Es gibt eine ausgezeichnete Menge, die **leere Menge**, die keine Elemente enthält und mit $\emptyset$ bezeichnet wird.

   Formel: $\exists \emptyset : \; \forall x : x \notin \emptyset.$

   Bild: *Es gibt einen leeren Behälter, der keine Dinge enthält.*

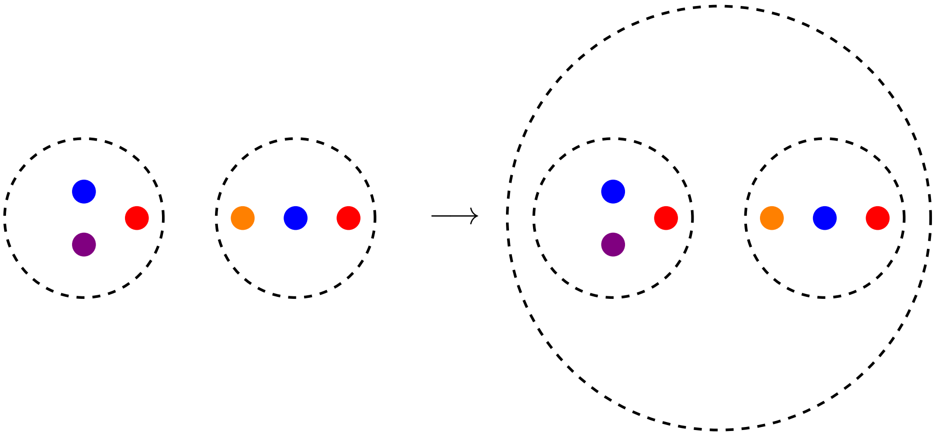


3. **Paarungssaxiom:** Zu zwei beliebigen Mengen M, N gibt es eine Menge X, die genau M und N als Elemente enthält. Man schreibt $X = \{M, N\}$ falls $M \neq N$ und $X = \{M\}$, wenn $M = N$.

   Formel: $\forall M \, \forall N \, \exists X : (x \in X) \Leftrightarrow (x = M \vee x = N)$

   Bild: *Man kann zwei Behälter mitsamt ihrem Inhalt in einen weiteren Behälter packen.*

   Beispiel: durch wiederholte Anwendung des Paarungsaxioms konstruiert man aus der leeren Menge die Mengen $\emptyset$, $\{\emptyset\}$, $\{\{\emptyset\}\}$, $\{\emptyset, \{\emptyset\}\}$, $\{\emptyset, \{\emptyset, \{\emptyset\}\}\}$,...

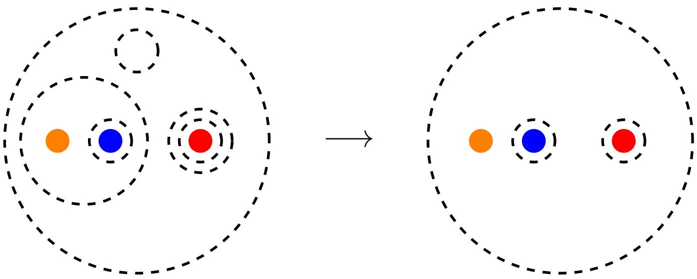


4. **Vereinigungssaxiom:** Zu jeder Menge M gibt es eine Menge X, die genau die Elemente der Elemente von M als Elemente enthält. Man schreibt $X = \cup M$ und statt $\cup\{A_1, A_2 ..., A_n\}$ auch $A_1 \cup A_2 \cup \ldots \cup A_n$.

   Formel: $\forall M \, \exists \cup M : (x \in \cup M) \;\Leftrightarrow\; (\exists W \in M : x \in W)$

   Bild: *Man kann alle in einem Behälter enthaltenen Behälter ausleeren und die dabei zum Vorschein kommenden Dinge in einen neuen Behälter packen.*

   Beispiel: Für die Menge $M = \{\emptyset, \{\emptyset, \{\emptyset\}\}\}$ gilt $\cup M = \{\emptyset, \{\emptyset\}\}$.



5. **Aussonderungsaxiom:** Für jede Menge M und jedes logische Prädikat $P : M \to \{w, f\}$ gibt es eine Menge X, die genau die Elemente von M mit $P(m) = w$ enthält. Man schreibt $X = \{m \in M | P(m)\}$ oder $X = \{m \in M : P(m)\}$.

   Formel: $\forall M \, \exists X : (m \in X) \Leftrightarrow (m \in M) \wedge (P(m) = w)$

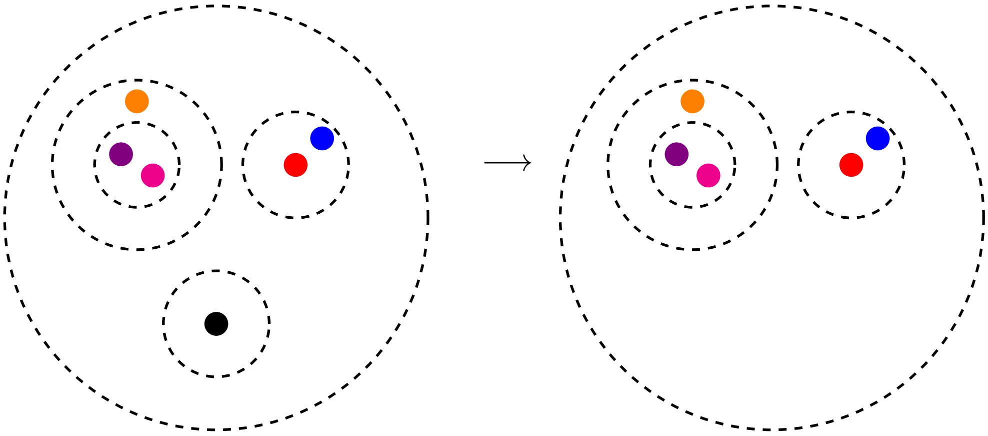
   Bild: *Man kann einen Behälter öffnen, die Dinge herausnehmen, die eine bestimmte Eigenschaft haben, und sie in einen neuen Behälter packen.*

Beispiel: Für $M = \{\emptyset, \{\emptyset, \{\emptyset\}\}\}$ und das Prädikat $P : M \to \{w, f\}$ mit $P(m) = w$ wenn m mindestens zwei Elemente hat und $P(m) = f$ sonst, gilt $\{m \in M | P(m)\} = \{\{\emptyset, \{\emptyset\}\}\}$.

Beispiel: Für $M = \{\emptyset, \{\emptyset, \{\emptyset\}\}\}$ und das Prädikat $P : M \to \{w, f\}$ mit $P(m) = w$ falls $m = \emptyset$ und $P(m) = f$ sonst, gilt $\{m \in M | P(m)\} = \{\emptyset\}$.

Beispiel: Für $P : M \to \{w, f\}$ mit M beliebig und $P(m) = f$ wenn $m \in M$ und $P(m) = w$ sonst ist $\{m \in M | P(m)\} = \emptyset$.

Beispiel: Für jedes Element $x \in M$ gibt es die Menge $\{x\}$, die x als einziges Element enthält, denn es gilt $\{x\} = \{m \in M | P(m)\}$ für $P : M \to \{w, f\}$ mit $P(m) = w$ wenn $m = x$ und $P(m) = f$ sonst.

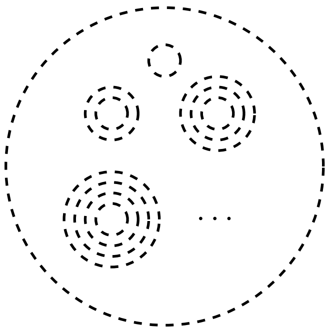


6. **Unendlichkeitsaxiom:** Es gibt eine Menge X, so dass $\emptyset \in X$ und für jedes Element $x \in X$ auch $\{x\} \in X$ gilt.

   Formel: $\exists X : (\emptyset \in X) \wedge (x \in X \Rightarrow \{x\} \in X)$

   Bild: *Es gibt einen Universalbehälter, der den leeren Behälter enthält und für jedes Ding, das er enthält, auch den Behälter enthält, der dieses Ding enthält.*

   Beispiel: Insbesondere enthält X alle Mengen der Form $\emptyset$, $\{\emptyset\}$, $\{\{\emptyset\}\}$, $\{\{\{\emptyset\}\}\}$,...

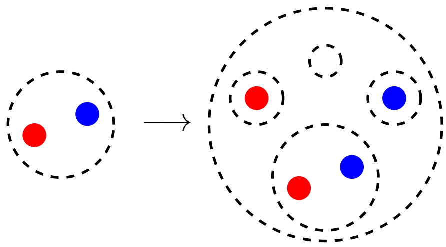


7. **Potenzmengenaxiom:** Für jede Menge M gibt es eine Menge $\mathrm{Pot}(M)$, die **Potenzmenge** von M, deren Elemente genau die Teilmengen von M sind.

   Formel: $\forall M\ \exists \mathrm{Pot}(M) : X \in \mathrm{Pot}(M) \Leftrightarrow (x \in X \Rightarrow x \in M)$

   Bild: *Man kann einen gegebenen Behälter ausleeren, einen Teil seines Inhalts auswählen und in einen neuen Behälter füllen - die Teilmenge. Dann kann man einen Behälter bauen, der alle auf diese Weise gefüllten Behälter enthält.*

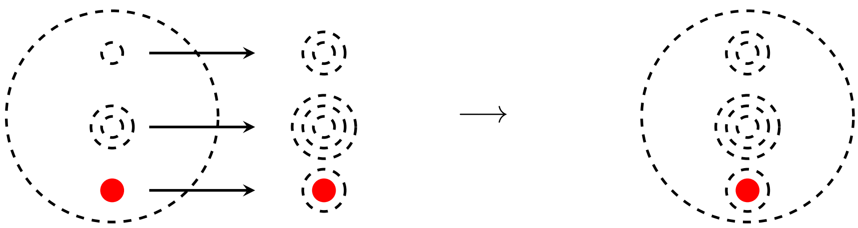
   Beispiel: Es gilt $\mathrm{Pot}(\emptyset) = \{\emptyset\}$, $\mathrm{Pot}(\{\emptyset, \{\emptyset\}\}) = \{\emptyset, \{\emptyset\}, \{\{\emptyset\}\}, \{\emptyset, \{\emptyset\}\}\}$.



8. **Ersetzungsaxiom:** Ist P ein zweistelliges logisches Prädikat, so dass es zu jeder Menge X genau eine Menge Y mit $P(X, Y) = w$ gibt, so gibt es zu jeder Menge M eine Menge X, die genau die Mengen x mit $P(m, x) = w$ für Elemente $m \in M$ enthält. Man schreibt $X = \{x \mid \exists m \in M : P(m, x) = w\}$ oder $X = \{x : \exists m \in M : P(m, x) = w\}$.

   Formel: $\forall M\ \exists X : (x \in X) \Leftrightarrow (\exists m \in M : P(m, x) = w)$

   Bild: *Hat man eine Bedingung an zwei Behälter, die für jeden gegebenen Behälter durch genau einen weiteren Behälter erfüllt wird, so kann man einen Behälter ausleeren, alle Behälter darin durch die entsprechenden Behälter ersetzen, die die Bedingung erfüllen, und diese Behälter mit anderen Behältern in einen neuen Behälter packen.*

   Beispiel: Für das Prädikat P mit $P(X, X) = w$ und $P(X, Y) = f$ für $X \neq Y$ und eine beliebige Menge M gilt $X = M$.

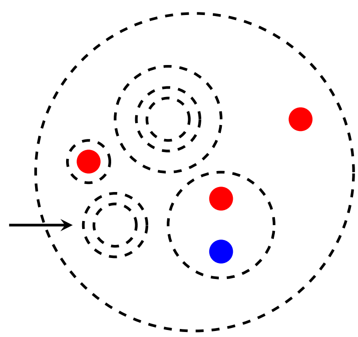
   Beispiel: Für das Prädikat P mit $P(X, \{X\}) = w$ und $P(X, Y) = f$ für $Y \neq \{X\}$ und die Menge $M = \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}$ ist $X = \{\{\emptyset\}, \{\{\emptyset\}\}, \{\{\emptyset, \{\emptyset\}\}\}\}$.

9. **Fundierungsaxiom:** In jeder nichtleeren Menge M gibt es ein Element $m \in M$, so dass m und M keine Elemente gemeinsam haben.

   Formel: $M \neq \emptyset \Rightarrow \exists m \in M : (x \in m \Rightarrow x \notin M)$

   Bild: *In jedem nicht-leeren Behälter gibt es einen Behälter, der keinen gemeinsamen Inhalt mit dem ersten Behälter hat.*

   Beispiel: Für $M = \{\emptyset, \{\emptyset\}, \{\{\emptyset\}\}\}$ ist $\emptyset$ ein solches Element, nicht aber $\{\emptyset\}$ oder $\{\{\emptyset\}\}$.

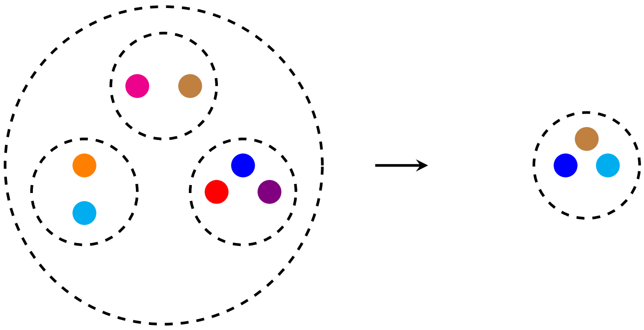
   

10. **Auswahlaxiom:** Ist M eine Menge, so dass alle Elemente von M nicht-leere Mengen sind und je zwei Elemente von M keine gemeinsamen Elemente haben, dann gibt es eine Menge X, die genau ein Element aus jedem Element $m \in M$ enthält.

Formel: $((m \in M \Rightarrow m \neq \emptyset) \wedge ((m \neq n \in M) \Rightarrow (x \in m \Rightarrow x \notin n)))$
$\Rightarrow \exists X : (\forall m \in M : \exists! x \in m : \ x \in X)$

Bild: *Hat man einen Behälter, der nur nicht-leere Behälter enthält, die untereinander keinerlei Inhalt gemeinsam haben, dann kann man aus jedem dieser nicht-leeren Behälter genau ein Ding auswählen und diese Dinge in einen neuen Behälter packen.*

Beispiel: Die Menge $M = \{\{\emptyset\}, \{\{\emptyset\}, \{\{\emptyset\}\}\}\}$ erfüllt die Bedingungen, und aus ihr lassen sich mit dem Auswahlaxiom die Mengen $\{\emptyset, \{\{\emptyset\}\}\}$ und $\{\emptyset, \{\emptyset\}\}$ konstruieren.



Bemerkung 1.3.5:

1. Die Zermelo-Fraenkel Axiome sind *redundant*, d. h. einige der Axiome sind unnötig, weil sie sich aus anderen ableiten lassen:
   - Das Aussonderungsaxiom ist ein Spezialfall des allgemeineren Ersetzungsaxioms.
   - Das Paarungsaxiom folgt aus dem Potenzmengenaxiom und dem Ersetzungsaxiom.
   - Das Leermengenaxiom folgt aus dem Aussonderungsaxiom und der Existenz irgendeiner Menge, die z. B. durch das Unendlichkeitsaxiom gesichert ist.

   Aus historischen Gründen werden diese Axiome aber trotzdem alle aufgeführt.
2. Das Unendlichkeitsaxiom wird oft auch in der folgenden Form formuliert: Es gibt eine Menge X, so dass $\emptyset \in X$ und für jedes $x \in X$ auch $\{x, \{x\}\} \in X$ gilt. Dieses Axiom ist zu dem oben angegebenen äquivalent, da man die beiden mit Hilfe der anderen Zermelo-Fraenkel Axiome ineinander umwandeln kann (Übung: wie?).

Bemerkung 1.3.6: Die Zermelo-Fraenkel-Axiome schließen die Russelsche Antinomie aus: Es gilt $M \notin M$ für alle Mengen M.

Beweis:
Für jede Menge M kann man mit dem Paarungsaxiom die Menge $\{M\}$ bilden, die als einziges Element die Menge M enthält. Nach dem Fundierungsaxiom, muss dann gelten, dass M mit $\{M\}$ kein Element gemeinsam hat: $x \in M \Rightarrow x \notin \{M\}$ oder, dazu äquivalent, $x \in \{M\} \Rightarrow x \notin M$. Da $M \in \{M\}$ gilt, muss somit $M \notin M$ gelten.

Damit kann es keine Menge geben, die alle Mengen als Elemente enthält, denn sie müsste sich selbst als Element enthalten. Ebenso kann es keine Menge geben, die alle Mengen als Elemente enthält, die sich nicht selbst als Element enthalten, denn diese wäre dann gleich der Menge, die alle Mengen als Elemente enthält. $\square$

Wir skizzieren nun, wie sich wichtige, zum Teil aus der Schule bekannte Konstruktionen mit Mengen aus den Zermelo-Fraenkel Axiomen ergeben. Diese Konstruktionen sind essentiell und durchziehen die gesamte Mathematik. Sie ermöglichen es uns außerdem, unseren Vorrat an Beispielen von Mengen zu erweitern.

- **Logik $\rightarrow$ Teilmengenbegriff:** Schon mit Hilfe der Logik können wir Teilmengen einer Menge M definieren. Nach dem Bestimmtheitsaxiom heißt eine Menge $N \subseteq M$ Teilmenge von M, wenn $x \in N \Rightarrow x \in M$.
- **Aussonderungsaxiom $\rightarrow$ Teilmengen aus logischen Prädikaten:** Um aus logischen Prädikaten $P : M \rightarrow \{w, f\}$ Teilmengen von M zu konstruieren, benötigen wir das Aussonderungsaxiom. Es stellt stellt sicher, dass für jedes solche Prädikat $\{m \in M|P(m)\}$ eine Teilmenge von M ist.
- **Spezialfall: Aussonderungsaxiom $\rightarrow$ Mengendifferenz:** Für zwei gegebene Mengen M, N können wir das Prädikat $P_N : M \rightarrow \{w, f\}$ mit $P_N(m) = w$ falls $m \in N$ und $P_N(m) = f$ sonst betrachten. Dies liefert die sogenannte **Mengendifferenz** $M \setminus N = \{m \in M|\neg P_N(m)\} = \{m \in M \mid P_N(m) = f\}$.
- **Potenzmengenaxiom $\rightarrow$ Potenzmenge:** Das Potenzmengenaxiom liefert uns die **Potenzmenge** einer Menge M, die Menge $\mathrm{Pot}(M)$ aller Teilmengen von M.
- **Ersetzungsaxiom $\rightarrow$ Indexmengen:** Das Ersetzungsaxiom ermöglicht es uns, Mengen zu indizieren, d. h. durch Elemente anderer Mengen aufzulisten. Ist I eine beliebige Menge und wird jedem Element $i \in I$ eine im Sinne des Ersetzungsaxioms eindeutig bestimmte Menge M_i zugeordnet, dann garantiert das Ersetzungsaxiom, dass die Menge $\{M_i|i \in I\}$ existiert. Man nennt dann I eine **Indexmenge** und sagt die Menge I **indiziere** die Familie von Mengen $(M_i)_{i\in I}$.
- **Ersetzungsaxiom+Aussonderungsaxiom $\rightarrow$ Schnittmengen:** Für eine durch eine Indexmenge $I \neq \emptyset$ indizierte Familie von Mengen $(M_i)_{i\in I}$ können wir ein beliebiges Element $j \in I$ auswählen und das Prädikat $P_\cap : M_j \rightarrow \{w, f\}$ mit $P_\cap(m) = w$ wenn $m \in M_i\ \forall i \in I$ und $P_\cap(m) = f$ sonst betrachten. Die **Schnittmenge** $\cap_{i\in I} M_i$ der Mengen M_i ist definiert als $\cap_{i\in I} M_i = \{m \in M_j|P_\cap(m)\}$. Man kann zeigen, dass sie von der Wahl des Elements $j \in I$ nicht abhängt.
- **Ersetzungsaxiom + Vereinigungsaxiom $\rightarrow$ Vereinigungen von Mengen:** Für eine durch eine Indexmenge $I \neq \emptyset$ indizierte Familie von Mengen $(M_i)_{i\in I}$ garantiert uns das Ersetzungsaxiom die Existenz der Menge $\{M_i|i \in I\}$. Das Vereinigungsaxiom erlaubt es uns dann die **Vereinigung** $\cup_{i\in I} M_i$ zu definieren als $\cup_{i\in I} M_i = \cup\{M_i|i \in I\}$.

1.4 Die Konstruktion der natürlichen Zahlen

Bis jetzt sieht es so aus, als könne man mit den Zermelo-Fraenkel Axiomen nur Ausdrücke produzieren, in der die Zeichen $\emptyset$ und $\{\}$ miteinander verschachtelt werden, nicht aber bekannte Strukturen wie Zahlen beschreiben. Wir zeigen nun, dass dies zur Beschreibung der natürlichen Zahlen völlig ausreichend ist. Dazu muss man zunächst klären, was die natürlichen Zahlen eigentlich sind.

Die naive Antwort, nämlich, dass es Ausdrücke sind, die sich aus den Ziffern 0,1,2,3,4,5,6,7,8,9 zusammensetzen lassen, hilft hier nicht weiter. Denn man könnte sie auch als römische Zahlen schreiben, ohne, dass sich an ihren Eigenschaften irgendetwas ändern würde. Auch die sprachlichen Bezeichnungen der natürlichen Zahlen variieren von Sprache zu Sprache, ohne dass es die

Zahlen selbst tun. Die Menge der natürlichen Zahlen kann daher nur bis auf *Umbenennungen* bestimmt sein. Verfügt man über ein Wörterbuch, das die Bezeichnungen oder Symbole der natürlichen Zahlen 1:1 in eine andere Bezeichnungen oder Symbole übersetzt, so kann man genauso mit den neuen Bezeichnungen oder Symbolen arbeiten, ohne, dass sich etwas ändert.

Das zeigt sich auch daran, dass es in der Mathematik zwei Konventionen für diese Zahlen gibt. Die eine bezeichnet mit den natürlichen Zahlen die Menge $\mathbb{N} = \{0, 1, 2, \ldots\}$, die andere die Menge $\mathbb{N} = \{1, 2, 3, \ldots\}$ und schreibt $\mathbb{N}_0 = \{0, 1, 2, \ldots\}$. In dieser Vorlesung wird die zweite Konvention benutzt.

Die entscheidende Eigenschaft der natürlichen Zahlen ist abstrakter, nämlich, dass sie einen Anfang haben, in unserer Konvention die Zahl 1, und sich alle anderen natürlichen Zahlen jeweils als Nachfolger von genau einer natürlichen Zahl ergeben. Daher charakterisiert man die natürlichen Zahlen abstrakt, als eine Menge, die die *Peano Axiome* erfüllt.

Definition 1.4.7: (**Peano Axiome**)
Die Menge $\mathbb{N}$ der natürlichen Zahlen ist eine Menge, die die folgenden Axiome erfüllt:

(P1) Es gibt eine ausgezeichnete natürliche Zahl $1 \in \mathbb{N}$.

(P2) Jede natürliche Zahl n hat einen eindeutigen Nachfolger $N(n)$: $\forall n \in \mathbb{N} \Rightarrow N(n) \in \mathbb{N}$.

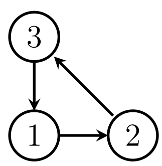
(P3) Die Zahl 1 ist nicht Nachfolger irgendeiner anderen Zahl: $\forall n \in \mathbb{N} : N(n) \neq 1$.

(P4) Verschiedene Zahlen haben verschiedene Nachfolger: $N(m) = N(n) \Rightarrow n = m$,

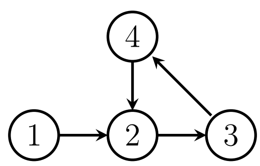
(P5) Ist M eine Menge, die die natürliche Zahl 1 und für jede darin enthaltene natürliche Zahl auch deren Nachfolger enthält, so enthält sie ganz $\mathbb{N}$:
$\forall M : (1 \in M) \wedge (\forall n \in \mathbb{N} : n \in M \Rightarrow N(n) \in M) \Rightarrow \mathbb{N} \subseteq M$.

Diese Charakterisierung erscheint zunächst unnötig kompliziert. Die ersten beiden Axiome (P1) und (P2) sind intuitiv einsichtig, aber man fragt sich, wofür der Rest gebraucht wird. Es stellt sich aber heraus, dass keines dieser Axiome unnötig ist. Lässt man eines davon weg, so gibt es viele weitere Mengen, die die Axiome erfüllen und nichts mit den natürlichen Zahlen zu tun haben. Dies sieht man an den folgenden Beispielen:

- Die Menge $\mathbb{N}' = \{1, 2, 3\}$ mit $N(1) = 2$, $N(2) = 3$ und $N(3) = 1$ erfüllt (P1), (P2), (P4) und (P5), aber nicht (P3).



- Die Menge $\mathbb{N}'' = \{1, 2, 3, 4\}$ mit $N(1) = 2$, $N(2) = 3$, $N(3) = 4$ und $N(4) = 2$ erfüllt die Axiome (P1), (P2) (P3) und (P5), aber nicht (P4).



- Die Menge $\mathbb{N}''' = \mathbb{N} \cup \{2^{-k} \mid k \in \mathbb{N}\}$ mit der Nachfolgern $N(n) = n + 1$ für $n \in \mathbb{N}$ und $N(2^{-k}) = 2^{-k-1}$ für $k \in \mathbb{N}$ erfüllt die Axiome (P1) bis (P4), aber nicht (P5). Denn für $M = \mathbb{N}$ gilt $1 \in \mathbb{N}$ und $N(n) \in M$ für alle $n \in \mathbb{N}''' \cap M$, aber nicht $\mathbb{N}''' \subseteq M$.

$$\tfrac{1}{2} \to \tfrac{1}{4} \to \tfrac{1}{4} \to \tfrac{1}{8} \to \cdots$$

$$1 \to 2 \to 3 \to 4 \to \cdots$$

Damit würde Weglassen eines Axioms sofort zu einer Uneindeutigkeit der natürlichen Zahlen bzw. zu einem Verlust ihrer Ordnung führen. Wir zeigen nun, dass die natürlichen Zahlen tatsächlich existieren, indem wir sie mit Hilfe der Zermelo-Fraenkel Axiome konstruieren.

Satz 1.4.8: (**Zermelosche Zahlenreihe**)
Die Menge $\mathbb{N} = \{\emptyset, \{\emptyset\}, \{\{\emptyset\}\}, \ldots\}$ mit $1 = \emptyset$ und $N(n) = \{n\}$ erfüllt die Peano Axiome.

Beweis:
Nach dem Leermengenaxiom existiert die leere Menge $\emptyset$ und nach dem Unendlichkeitsaxiom eine Menge X mit $\emptyset \in X$ und $\{x\} \in X$ für alle $x \in X$. Sie enthält $\mathbb{N}$, und mit dem Aussonderungsaxiom können wir daraus die Teilmenge $\mathbb{N} = \{\emptyset, \{\emptyset\}, \{\{\emptyset\}\}, \ldots\}$ bilden.

Wir setzen $1 = \emptyset$ und $N(x) = \{x\}$ für alle $x \in \mathbb{N}$. Dann sind (P1) und (P2) per Konstruktion der Menge $\mathbb{N}$ erfüllt. Da jede Menge der Form $\{x\}$ per Konstruktion x als Element enthält, gilt $1 = \emptyset \neq N(x)$ für alle $x \in \mathbb{N}$ (P3). Ist $N(n) = \{n\} = \{m\} = N(m)$ für $n, m \in \mathbb{N}$, so folgt aus dem Bestimmtheitsaxiom $n = m$ (P4).

Um (P5) nachzuweisen betrachten wir eine Menge M mit $1 = \emptyset \in M$ und $N(x) = \{x\} \in M$ für alle $x \in \mathbb{N} \cap M$. Wir schreiben $n = \{\{\ldots\{\emptyset\}\ldots\}\}$ für das Element von $\mathbb{N}$, das aus $\emptyset$ durch $(n-1)$-fache Klammerung konstruiert wird, also $2 = \{\emptyset\}$, $3 = \{\{\emptyset\}\}$ etc, und zeigen per vollständiger Induktion, dass $n \in M$ gilt. Per Annahme gilt $1 \in M$ (Induktionsanfang). Ist gezeigt, dass $n \in M$ gilt (Induktionsannahme), so folgt nach Voraussetzung an M auch $n + 1 = N(n) \in M$ (Induktionsschritt). Damit gilt $\mathbb{N} \subseteq M$ (P5). □

Das sich das Axiom (P5), oft auch *Induktionsaxiom* genannt, so leicht per vollständiger Induktion beweisen lässt, ist kein Zufall. Man kann zeigen, dass die Peano Axiome äquivalent zum Prinzip der vollständigen Induktion sind. Wir zeigen nun noch, dass es tatsächlich sinnvoll ist, von *den* natürlichen Zahlen zu sprechen, also dass die natürlichen Zahlen bis auf Umbenennung eindeutig bestimmt sind.

Satz 1.4.9: Die Menge $\mathbb{N}$ der natürlichen Zahlen ist durch die Peano Axiome bis auf Umbenennung eindeutig festgelegt:

1. Ist $\mathbb{N}'$ eine Menge, die die Peano-Axiome erfüllt, mit Nachfolger N' und ausgezeichnetem Element $1' \in \mathbb{N}'$, dann gibt es genau eine Zuordnung $f : \mathbb{N} \to \mathbb{N}'$ mit $f(1) = 1'$ und $N'(f(n)) = f(N(n))$ für alle $n \in \mathbb{N}$.
2. Die Zuordnung f ist eine Bijektion: Jedem Element von $\mathbb{N}$ wird genau ein Element von $\mathbb{N}'$ zugeordnet und jedes Element von $\mathbb{N}'$ wird genau einem Element von $\mathbb{N}$ zugeordnet.

Beweis:
1. Wir definieren $k' = N'^{k-1}(1')$ als den $(k-1)$ten Nachfolger von $1'$ in $\mathbb{N}'$, wobei $N'^0(1') = 1'$, und zeigen, dass $\mathbb{N}' = \{N'^k(1') \mid k \in \mathbb{N}_0\}$ gilt. Offensichtlich ist $M = \{N'^k(1') \mid k \in \mathbb{N}_0\} \subseteq \mathbb{N}'$ eine Teilmenge von $\mathbb{N}'$, die $1'$ per Konstruktion enthält. Aus $x \in M$ folgt $x = N'^k(1')$ für ein $k \in \mathbb{N}_0$, und damit ist auch $N'(x) = N'^{k+1}(1') \in M$. Mit (P5) folgt $\mathbb{N}' \subseteq M$ und damit $M = \mathbb{N}'$.

2. Ist $f : \mathbb{N} \to \mathbb{N}'$ eine Zuordnung mit $f(1) = 1'$ und $f(N(n)) = N'(f(n))$ für alle $n \in \mathbb{N}$, so folgt mit vollständiger Induktion $f(k) = k' = N'^{k-1}(1')$ für alle $k \in \mathbb{N}$:

Für $n = 1$ gilt es per Annahme an f (Induktionsanfang). Ist gezeigt, dass es für ein $k \in \mathbb{N}$ gilt (Induktionsannahme), so folgt $f(k+1) = f(N(k)) = N'(f(k)) = N'(N'^{k-1}(1')) = N'^k(1)$ (Induktionsschritt), und damit gilt es für alle $k \in \mathbb{N}$.

Also ist f durch diese Bedingungen eindeutig bestimmt und ordnet jedem Element von $\mathbb{N}$ genau ein Element von $\mathbb{N}'$ zu. Da $\mathbb{N}' = \{f(k) \mid k \in \mathbb{N}\}$, wird auch jedes Element von $\mathbb{N}'$ mindestens einem Element von $\mathbb{N}$ zugeordnet. Wir zeigen noch, dass es maximal einem Element von $\mathbb{N}$ zugeordnet wird. Seien dazu $m, n \in \mathbb{N}$ mit $m \leq n$ und $f(n) = N'^{n-1}(1') = N'^{m-1}(1') = f(m)$. Dann folgt mit (P4) zunächst $N'^{n-2}(1') = N'^{m-2}(1')$, dann $N'^{n-3}(1') = N'^{m-3}(1')$ und schließlich $N'^{n-m}(1') = N'^0(1') = 1'$. Daraus folgt mit (P3) dann $n = m$. □

Man kann die Flexibilität, die durch die Peano Axiome bestimmten natürlichen Zahlen umzubenennen, ausnutzen, um eine elegantere Konstruktion der natürlichen Zahlen aus den Zermelo-Fraenkel Axiomen anzugeben, die ihre Eigenschaften, nämlich die Ordnung, besser erfasst. Intuitiv erwartet man nämlich, dass die Mengen, die die natürlichen Zahlen beschreiben, mit den natürlichen Zahlen wachsen, also immer mehr Elemente enthalten sollten. Dies ist für die Zermelosche Zahlenreihe, in der alle Mengen einelementig sind, nicht der Fall. Mit Hilfe des Ersetzungsaxioms lässt sich aber leicht eine Variante der natürlichen Zahlen konstruieren, die auch diese Bedingung erfüllt.

Satz 1.4.10: (**von Neumannsche Zahlenreihe**)

1. Die Menge $\mathbb{N}' = \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}, \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}, \ldots\}$ mit $N'(x) = \cup\{x, \{x\}\}$ und $1' = \emptyset$ erfüllt die Peano Axiome.
2. Die $n$te natürliche Zahl in $n' \in \mathbb{N}'$ hat genau $n - 1$ Elemente.
3. Für $m', n' \in \mathbb{N}'$ gilt $m \leq n$ genau dann, wenn $m' \subseteq n'$.

Beweis:
1. Mit Hilfe des Paarungsaxioms können wir aus jeder Menge M die Menge $\{M\}$ konstruieren, und aus M und $\{M\}$ die Menge $\{M, \{M\}\}$. Wir setzen $1' = \emptyset$ und $(n+1)' = \cup\{n', \{n'\}\}$ für alle $n \in \mathbb{N}$. Wir bezeichnen mit $\mathbb{N} = \{\emptyset, \{\emptyset\}, \{\{\emptyset\}\}, \ldots\}$ die Zermelosche Zahlenreihe mit $1 = \emptyset$ und $N(n) = \{n\}$ für $n \in \mathbb{N}$.

Wir betrachten das zweistellige Prädikat P mit $P(n, n') = w$, $P(n, X) = f$ für $n \in \mathbb{N}$ und $X \neq n'$ und mit $P(U, U) = w$, $P(U, V) = f$ für $U \notin \mathbb{N}$ und $V \neq U$. Dieses erfüllt die Bedingung, dass es zu jeder Menge X genau eine Menge Y mit $P(X, Y) = w$ gibt. Mit dem Ersetzungsaxiom können wir also jedes Element $n \in \mathbb{N}$ durch das eindeutig bestimmte x mit $P(n, x) = w$ ersetzen, also durch $x = n'$. Damit erhalten wir die Menge $\mathbb{N}' = \{n' \mid n \in \mathbb{N}\}$.

Sie enthält offensichtlich das Element $\emptyset = 1'$ (P1). Für jedes $x \in \mathbb{N}'$ gibt es ein $n \in \mathbb{N}$ mit $x = n'$, und damit ist $N'(x) = N'(n') = \cup\{n', \{n'\}\} = (n+1)' \in \mathbb{N}'$ (P2). Wegen $x \in N(x)$ für alle $x \in \mathbb{N}'$ gilt $N(x) \neq \emptyset$ für alle $x \in \mathbb{N}'$ (P3).

Um (P4) nachzuweisen betrachten wir $x, y \in \mathbb{N}'$ mit $N'(x) = \cup\{x, \{x\}\} = \cup\{y, \{y\}\} = N'(y)$. Nun gilt entweder $x = y$ oder $x \neq y$. Im zweiten Fall müsste wegen $x \in N'(x) = N'(y)$ dann $x \in y$ gelten und wegen $y \in N'(x) = N'(y)$ auch $y \in x$. Damit wäre $x \in y \in x$, was in Übungsaufgabe 19 ausgeschlossen wurde. Also folgt $x = y$ (P4).

Ist M eine Menge mit $1' \in M$ und $N'(n') = (n+1)' \in M$ für alle $n \in \mathbb{N}$ mit $n' \in M$, so folgt wie in der Zermeloschen Zahlenreihe mit vollständiger Induktion $\mathbb{N}' = \{n' \mid n \in \mathbb{N}\} \subseteq M$ (P5).

2. Wir beweisen die Aussage per vollständiger Induktion: Offensichtlich enthält $1' = \emptyset$ kein Element (Induktionsanfang). Ist gezeigt, dass n' genau $n-1$ Elemente enthält (Induktionsannahme), so enthält $(n+1)' = \cup\{n', \{n'\}\}$ genau ein Element mehr, nämlich das Element n', und damit n Elemente (Induktionsschritt). Damit folgt die Behauptung.

3. Per Konstruktion der Mengen n' gilt $1' \subseteq 2' \subseteq 3' \subseteq \ldots$. Also folgt aus $m \leq n$, dass $m' \subseteq n'$. Sind umgekehrt $n, m \in \mathbb{N}$ mit $m' \subseteq n'$, so folgt $m \leq n$, denn ansonsten wäre $m > n$ und damit hätte m' nach 2. mehr Elemente als n', im Widerspruch zu $m' \subseteq n'$. $\square$

2 Algebraische Grundbegriffe

2.1 Mengen und Abbildungen

Im letzten Abschnitt haben wir Mengen mit Hilfe der Zermelo-Fraenkel Axiome definiert und gezeigt, dass sich damit bekannte Zahlenmengen wie die Menge der natürlichen Zahlen konstruieren lassen. Wir werden die Zermelo-Fraenkel Axiome und die Charakterisierung der natürlichen Zahlen über die Peano Axiome im Verlauf der Vorlesung kaum noch benutzen.

Trotzdem sind die Zermelo-Fraenkel Axiome wichtig. Sie garantieren nämlich, dass der Mengernbegriff im Gegensatz zum Cantorschen Mengenbegriff nicht sofort zu Paradoxa und Widersprüchen führen[3]. Außerdem ist es wichtig zu wissen, dass Mengen und Konstruktionen mit Mengen nicht einfach aus dem Nichts erscheinen, sondern aus Axiomen abgeleitet werden.

Im Folgenden reicht es, zu wissen, dass eine Menge *nie* Element von sich selbst sein kann, und dass es keine Menge gibt, die alle Mengen als Elemente enthält. Außerdem sollte man die aus der Schule bekannten genannten grundlegenden Konstruktionen mit Mengen beherrschen. Wir fassen diese Konstruktionen hier noch einmal zusammen.

Definition 2.1.1: Eine **Menge** ist etwas, das **Elemente** enthalten kann.

1. Ist a ein Element einer Menge M, so schreibt man $a \in M$. Ist a kein Element der Menge M, so schreibt man $a \notin M$.
2. Für eine endliche Menge M, die genau die Elemente a_1, a_2, ..., a_n enthält, schreibt man $M = \{a_1, a_2, ..., a_n\}$.
3. Es gibt eine ausgezeichnete Menge, die **leere Menge** $\emptyset$, die keine Elemente enthält.
4. Eine Menge ist durch ihre Elemente M eindeutig bestimmt. Zwei Mengen M, N sind **gleich** genau dann, wenn sie die gleichen Elemente enthalten:

$$M = N \quad \Leftrightarrow \quad (a \in M \Rightarrow a \in N) \wedge (a \in N \Rightarrow a \in M).$$

5. Eine Menge kann Mengen als Elemente enthalten, aber *nie* Element von sich selbst sein.
6. Die Anzahl der Elemente in einer Menge M wird mit $|M|$ bezeichnet und heißt auch **Kardinalität** von M.

Definition 2.1.2: Seien M, N Mengen und $(M_i)_{i \in I}$ eine durch eine Indexmenge I indizierte Familie von Mengen.

1. N heißt **Teilmenge** von M und M **Obermenge** von N genau dann, wenn alle Elemente von N auch Elemente von M sind. Man schreibt dann $N \subseteq M$:

$$N \subseteq M \Leftrightarrow ((a \in N) \Rightarrow (a \in M)).$$

2. Die **Potenzmenge** $\mathrm{Pot}(M)$ ist die Menge aller Teilmengen von M:

$$\mathrm{Pot}(M) = \{B \,|\, B \subseteq M\}.$$

[3]Allerdings kann die Widerspruchsfreiheit der Zermelo-Fraenkel Axiome nicht bewiesen werden. Das ist eine Konsequenz des 2. Gödelschen Unvollständigkeitssatzes. Man kann nur für einzelne Paradoxa des naiven Cantorschen Mengenbegriffs zeigen, dass sie durch die Zermelo-Fraenkel Axiome aufgelöst werden [Hf].

3. Die **Mengendifferenz** $M \setminus N$ ist die Menge der Elemente, die in M, aber nicht in N enthalten sind:
$$M \setminus N = \{x \,|\, (x \in M) \wedge (x \notin N)\}.$$
Ist $N \subseteq M$, dann nennt man $M \setminus N$ auch das **Komplement** von N in M.

4. Ist $I \neq \emptyset$, so ist die **Schnittmenge** $\cap_{i\in I} M_i$ die Menge der Elemente, die in allen Mengen M_i mit $i \in I$ enthalten sind:
$$\cap_{i\in I} M_i = \{x \mid x \in M_i \ \forall i \in I\}.$$
Für $I = \{1, 2, ..., n\}$ schreibt man auch $\cap_{i=1}^n M_i$ oder $M_1 \cap \ldots \cap M_n$ statt $\cap_{i\in I} M_i$. Sind M, N zwei Mengen mit $M \cap N = \emptyset$, so sagt man M und N seien **disjunkt**.

5. Die **Vereinigung** $\cup_{i\in I} M_i$ ist die Menge aller Elemente, die in mindestens einer der Mengen M_i mit $i \in I$ enthalten sind:
$$\cup_{i\in I} M_i = \{x \mid \exists i \in I : x \in M_i\}.$$
Für $I = \{1, 2, ..., n\}$ schreibt man auch $\cup_{i=1}^n M_i$ oder $M_1 \cup \ldots \cup M_n$ statt $\cup_{i\in I} M_i$.

Die Konstruktionen in 3. bis 5. heißen **elementare Mengenoperationen**.

Bemerkung 2.1.3:

1. Es gilt $M = N$ genau dann, wenn $M \subseteq N$ und $N \subseteq M$. Dies ist oft nützlich in Beweisen.

   Denn $M = N$ ist gleichbedeutend zu $(a \in M \Rightarrow a \in N) \wedge (a \in N \Rightarrow a \in M)$. Die Aussage $a \in M \Rightarrow a \in N$ ist gleichbedeutend zu $M \subseteq N$ und $a \in N \Rightarrow a \in M$ zu $N \subseteq M$.

2. Ist $I = \emptyset$, so ist $\cup_{i\in I} M_i = \{x|\exists i \in I : x \in M_i\} = \emptyset$, denn es existiert kein $i \in I$.

Wir machen uns nun mit einigen wichtigen Beispielen von Mengen vertraut und konstruieren daraus mit Definition 2.1.2 neue Beispiele.

Beispiel 2.1.4:

1. Die Menge $\mathbb{N} = \{1, 2, 3...\}$ der **natürlichen Zahlen** und die Menge $\mathbb{N}_0 = \{0, 1, 2, 3, ...\}$.
2. Die Menge $\mathbb{Z} = \{..., -2, -1, 0, 1, 2, ...\}$ der **ganzen Zahlen**. .
3. Die Menge $\mathbb{Q} = \{\frac{p}{q} \mid p \in \mathbb{Z}, q \in \mathbb{N}\}$ der **rationalen Zahlen**. .
4. Die in der Vorlesung *Analysis 1* eingeführte Menge $\mathbb{R}$ der **reellen Zahlen**.

Beispiel 2.1.5:

1. Es gilt $\emptyset \subseteq \mathbb{N} \subseteq \mathbb{N}_0 \subseteq \mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R}$.

2. Ist M eine Menge und $P : M \to \{w, f\}$ ein Prädikat, so sind $\{m \in M|\, P(m) = w\}$ und $\{m \in M|P(m) = f\}$ Teilmengen von M.

3. Für $M = \{\emptyset, \{\emptyset\}, \{1, 2\}\}$ gilt

$$\mathrm{Pot}(M) = \{\emptyset, \{\emptyset\}, \{\{\emptyset\}\}, \{\{1, 2\}\}, \{\emptyset, \{\emptyset\}\}, \{\emptyset, \{1, 2\}\}, \{\{\emptyset\}, \{1, 2\}\}, \{\emptyset, \{\emptyset\}, \{1, 2\}\}\}.$$

4. Für die Mengen $M = \{\emptyset, \{\emptyset\}, 1\}$ und $N = \{\emptyset, \{\{\emptyset\}\}, \{1\}\}$ gilt:
$M \cap N = \{\emptyset\}$, $M \setminus N = \{\{\emptyset\}, 1\}$, $N \setminus M = \{\{\{\emptyset\}\}, \{1\}\}$, $M \cup N = \{\emptyset, \{\emptyset\}, \{\{\emptyset\}\}, 1, \{1\}\}$.

5. Für die leere Menge $\emptyset$ gilt $\mathrm{Pot}(\emptyset) = \{\emptyset\}$ sowie $\emptyset \subseteq M$, $M \cap \emptyset = \emptyset \cap M = \emptyset$,
$M \cup \emptyset = \emptyset \cup M = M$, $M \setminus \emptyset = M$ und $\emptyset \setminus M = \emptyset$ für jede Menge M.

6. Die Potenzmenge einer endlichen Menge M mit n Elementen hat 2^n Elemente. Deswegen wird die Potenzmenge $\mathrm{Pot}(M)$ manchmal auch mit 2^M bezeichnet.

   Denn um eine Teilmenge der Menge M zu bilden, muss für jedes der n Elemente in M separat entschieden werden, ob es darin enthalten sein soll. Diese n Entscheidungen und bestimmen die Teilmenge eindeutig und liefern die 2^n Teilmengen von M.

Da man die elementaren Mengenoperationen sehr häufig benutzt, um aus gegebenen Mengen neue zu konstruieren, hat man es oft mit Ausdrücken zu tun, in denen gewisse Mengenoperationen verschachtelt werden. Diese möchte man systematisch auflösen und vereinfachen können. Dazu dienen die folgenden Rechenregeln.

Lemma 2.1.6: (Rechenregeln für Mengenoperationen) Seien A, B, C Mengen. Dann gilt:

1. **Kommutativgesetze:** $A \cup B = B \cup A$ und $A \cap B = B \cap A$.
2. **Distributivgesetze:** $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ und $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.
3. **de Morgansche Regeln:** $A \setminus (B \cup C) = (A \setminus B) \cap (A \setminus C)$ und $A \setminus (B \cap C) = (A \setminus B) \cup (A \setminus C)$.
4. **Komplementsgesetz:** Für jede Teilmenge $D \subseteq A$ gilt $A \setminus (A \setminus D) = D$.

Beweis:
1. Die Kommutativgesetze ergeben sich direkt aus der Definition von $A \cap B$ und $A \cup B$ und der Kommutativität der elementaren Verknüpfungen $\vee$ und $\wedge$:

$$A \cup B = \{a | (a \in A) \vee (a \in B)\} = \{a | (a \in B) \vee (a \in A)\} = B \cup A$$
$$A \cap B = \{a | (a \in A) \wedge (a \in B)\} = \{a | (a \in B) \wedge (a \in A)\} = B \cap A.$$

Nach Bemerkung 2.1.3, 1. reicht es zum Beweis der restlichen Aussagen, jeweils zu zeigen, dass die Menge auf der rechten Seite der Gleichung eine Teilmenge der Menge auf der linken Seite ist und umgekehrt. Wir führen dies exemplarisch für die erste Aussage in 2. und 3. durch.

2. Ist $a \in A \cap (B \cup C)$, so gilt per Definition $a \in A$ und $a \in B \cup C$. Ist $a \in B$, so folgt daraus $a \in A \cap B$ und somit auch $a \in (A \cap B) \cup (A \cap C)$. Ist $a \notin B$, so folgt aus $a \in B \cup C$, dass $a \in C$ und somit wegen $a \in A$ auch $a \in A \cap C$. Also gilt $a \in (A \cap B) \cup (A \cap C)$, und wir haben gezeigt, dass $A \cap (B \cup C) \subseteq (A \cap B) \cup (A \cap C)$.

Ist umgekehrt $a \in (A \cap B) \cup (A \cap C)$, dann gilt $a \in A \cap B$ oder $a \in A \cap C$. Ist $a \in A \cap B$, so folgt $a \in A$ und $a \in B$ und somit auch $a \in A \cap (B \cup C)$. Ist $a \notin A \cap B$, dann folgt aus $a \in (A \cap B) \cup (A \cap C)$, dass $a \in A \cap C$ und somit auch wieder $a \in A \cap (B \cup C)$. Also haben wir gezeigt, dass auch $(A \cap B) \cup (A \cap C) \subseteq A \cap (B \cup C)$ gilt, und es folgt die Behauptung.

3. Ist $a \in A \setminus (B \cup C)$, so gilt $a \in A$ and $a \notin B \cup C$. Aus $a \notin B \cup C$ folgt $a \notin B$ und $a \notin C$. Da $a \in A$, folgt aus diesen Aussagen $a \in A \setminus B$ und $a \in A \setminus C$ und somit $a \in (A \setminus B) \cap (A \setminus C)$. Also haben wir gezeigt, dass $A \setminus (B \cup C) \subseteq (A \setminus B) \cap (A \setminus C)$.

Ist umgekehrt $a \in (A \setminus B) \cap (A \setminus C)$ dann gilt $a \in A \setminus B$ und $a \in A \setminus C$. Aus der ersten Aussage folgt $a \in A$ und $a \notin B$ und aus der zweiten $a \in A$ und $a \notin C$. Aus $a \notin B$ und $a \notin C$ folgt $a \notin B \cup C$. Also folgt aus beiden Aussagen zusammen $a \in A$ und $a \notin B \cup C$ und somit $a \in A \setminus (B \cup C)$. Damit haben wir gezeigt, dass auch $(A \setminus B) \cap (A \setminus C) \subseteq A \setminus (B \cup C)$ gilt, und somit $A \setminus (B \cup C) = (A \setminus C) \cap (A \setminus B)$

Der Beweis der zweite Aussagen in 2. und 3. und der Aussagen in 4. ist eine hilfreiche Übung. □

Eine weitere wichtige Konstruktion, mit der man aus gegebenen Mengen neue Mengen konstruieren kann, sind die *kartesischen Produkte* von Mengen. Diese beruhen auf dem Begriff des **geordneten n-Tupels** von Elementen aus gegebenen Mengen $A_1, ..., A_n$. Mengentheoretisch definiert man ein geordnetes n-Tupel $(a_1, ..., a_n)$ mit $a_1 \in A_1$,..., $a_n \in A_n$ und $n \in \mathbb{N}$ wie folgt:

$$(a_1) = \{\emptyset, \{a_1\}\} \qquad (a_1, ..., a_n) = \{(a_1, ..., a_{n-1}), \{a_n\}\} \quad \forall n \in \mathbb{N}, n \geq 2.$$

Dies liefert $(a_1, a_2) = \{\{\emptyset, \{a_1\}\}, \{a_2\}\}$, $(a_1, a_2, a_3) = \{\{\{\emptyset, \{a_1\}\}, \{a_2\}\}, \{a_3\}\}$ etc. und es folgt die **charakteristische Eigenschaft** der geordneten n-Tupel (Übung):

$$\forall a_1, b_1 \in A_1, ..., \forall a_n, b_n \in A_n : \quad (a_1, ..., a_n) = (b_1, ..., b_n) \Leftrightarrow (a_1 = b_1) \wedge ... \wedge (a_n = b_n). \qquad (1)$$

Man kann damit ein n-Tupel als eine geordnete Liste von Elementen $a_1 \in A_1$,..., $a_n \in A_n$ auffassen. Zwei solche Listen werden als gleich betrachtet genau dann, wenn sie an allen Stellen die gleichen Einträge enthalten, also wenn (1) gilt. Ein 2-Tupel nennt man auch ein **Paar** und ein 3-Tupel ein **Tripel**. Das kartesische Produkt der Mengen A_1, ..., A_n ist dann definiert als die Menge aller geordneten n-Tupel $(a_1, ..., a_n)$ mit $a_1 \in A_1$,..., $a_n \in A_n$.

Definition 2.1.7: Das **kartesische Produkt** oder **direkte Produkt** von n gegebenen Mengen A_1,..., A_n ist die Menge der geordneten n-Tupel $(a_1, ..., a_n)$ mit $a_1 \in A_1$,..., $a_n \in A_n$. Es wird mit $A_1 \times \ldots \times A_n$ bezeichnet und für $A_1 = ... = A_n = A$ auch mit A^n oder $A^{\times n}$.

$$\begin{aligned} &A_1 \times \ldots \times A_n := \{(a_1, ..., a_n) | a_1 \in A_1, ..., a_n \in A_n\} \\ &A^n = A^{\times n} = A \times \ldots \times A := \{(a_1, ..., a_n) | a_1, ..., a_n \in A\}. \end{aligned}$$

Im Gegensatz zu einer n-elementigen Menge, bei der die Einträge vertauscht werden dürfen, kommt es bei einem n-Tupel auf die *Reihenfolge* der Einträge an. So gilt beispielsweise $(1, 0) \neq (0, 1)$ in $\mathbb{Z} \times \mathbb{Z}$. Ebenso dürfen in einem Tupel Einträge *wiederholt* werden, während in einer Menge jedes Element nur einmal auftritt.

Beispiel 2.1.8:

1. $\{0, 1, 2\} \times \{x, 1\} = \{(0, x), (0, 1), (1, x), (1, 1), (2, x), (2, 1)\}$.
2. Für jede Menge M ist $\emptyset \times M = M \times \emptyset = \emptyset$.
3. Es gilt:

$$\begin{aligned} \mathbb{N}^n &= \{(a_1, ..., a_n) | a_1, ..., a_n \in \mathbb{N}\} \qquad & \mathbb{Z}^n &= \{(a_1, ..., a_n) | a_1, ..., a_n \in \mathbb{Z}\} \\ \mathbb{Q}^n &= \{(a_1, ..., a_n) | a_1, ..., a_n \in \mathbb{Q}\} \qquad & \mathbb{R}^n &= \{(a_1, ..., a_n) | a_1, ..., a_n \in \mathbb{R}\} \end{aligned}$$

   und damit $\mathbb{N}^n \subseteq \mathbb{Z}^n \subseteq \mathbb{Q}^n \subseteq \mathbb{R}^n$.

4. Sind $A_1, ..., A_n$ endlich, so gilt $|A_1 \times \ldots \times A_n| = |A_1| \cdots |A_n|$.

Da man Punkte in der Ebene, im dreidimensionalen Raum oder, allgemeiner, im n-dimensionalen Raum durch geordnete Paare, Tripel und n-Tupel reeller Zahlen beschreiben kann, spielt das kartesische Produkt von Mengen eine wichtige Rolle in der Geometrie.

Außerdem kann man es benutzen, um *Beziehungen* zwischen Elementen verschiedener Mengen zu beschreiben, die sogenannten *Relationen*. Sind M und N Mengen, so können wir eine Beziehung zwischen bestimmten Elementen von M und von N beschreiben, indem wir einfach die Elemente in $M \times N$ angeben, zwischen denen die entsprechende Beziehung besteht. Dies entspricht der Auswahl einer Teilmenge $R \subseteq M \times N$, nämlich der Teilmenge R der Paare $(m, n) \in M \times N$, für die m mit n in der entsprechenden Beziehung steht.

Definition 2.1.9:

1. Eine **Relation** ist ein Tripel (M, N, R) aus zwei Mengen M, N und einer Teilmenge $R \subseteq M \times N$.
2. Gilt $(m, n) \in R$ für zwei Elemente $m \in M$, $n \in N$, sagt man, dass m zu n in Relation R steht und schreibt $m \sim_R n$.
3. Gilt $N = M$ so nennt man R auch eine **Relation auf** M.

Beispiel 2.1.10:

1. (i) *jemanden lieben*, (ii) *die gleichen Eltern*[4] *haben wie*, (iii) *mindestens einen Elternteil gemeinsam haben mit*, (iv) *eine Schwester sein von* sind Relationen auf der Menge M der Menschen. Die Menge $R \subseteq M \times M$ ist die Menge der geordneten Menschenpaare (m_1, m_2) für die gilt: (i) m_1 liebt m_2, (ii) m_1 hat die gleichen Eltern wie m_2, (iii) m_1 hat mindestens einen Elternteil mit m_2 gemeinsam, (iv) m_1 ist eine Schwester von m_2.

2. Der Ausdruck $(x_1, x_2, x_3) \sim_R (y_1, y_2, y_3) \Leftrightarrow x_1^2 + x_2^2 + x_3^2 = y_1^2 + y_2^2 + y_3^2$ ist eine Relation auf der Menge $\mathbb{R}^3$, nämlich die Relation *den gleichen Abstand vom Ursprung haben.*

3. Auf jeder Menge M gibt es die Relation $R = \{(m, m) |\, m \in M\} \subseteq M \times M$. Sie lässt sich äquivalent auch durch den Ausdruck $m \sim_R n \Leftrightarrow m = n$ beschreiben. Hier steht m also genau dann in Relation R mit n, wenn $m = n$.

4. Der Ausdruck $x \sim_R y \Leftrightarrow x < y$ ist eine Relation auf $\mathbb{R}$ mit $R = \{(a, b)|a, b \in \mathbb{R}, a < b\}$.

5. Für jedes $n \in \mathbb{N}$ ist $x \sim_R y \Leftrightarrow n$ teilt $x - y$ eine Relation auf $\mathbb{Z}$ mit $R = \{(x, y) \in \mathbb{Z} \times \mathbb{Z} \mid n \text{ teilt } x - y\}$. Statt $x \sim_R y$ schreibt man auch $x = y \bmod n$.

6. Ebenso ist für jedes $n \in \mathbb{N}$ die Menge $R = \{(x, y) \in \mathbb{R}^2 \mid (x - y)/n \in \mathbb{Z}\}$ eine Relation auf den reellen Zahlen.

7. $(a, b) \sim (c, d) \Leftrightarrow ad = bc$ definiert eine Relation auf der Menge $\{(a, b)|a, b \in \mathbb{Z} \text{ und } b \neq 0\}$.

Sind M und N zwei gegebene Mengen, so bilden die möglichen Relationen (M, N, R) über M und N eine Menge, nämlich die Potenzmenge $\text{Pot}(M \times N)$, die alle Teilmengen von $M \times N$ enthält. Eine Relation R über M und N ist ja gerade eine Teilmenge von $M \times N$.

[4]Hier und im Folgenden wird der Begriff "Eltern" für die genetischen Eltern benutzt. Soziale Elternschaft, Adoption, Leihmutterschaft und mitochondriale DNA werden hierbei nicht berücksichtigt. Das Gleiche gilt für die Begriffe "Schwester" und "Bruder".

Wie aus den Beispielen hervorgeht, können Relationen auf einer Menge ganz unterschiedlich aussehen. Manche sind gegenseitig oder symmetrisch, andere einseitig. Im allgemeinen muss also aus der Tatsache, dass ein Element in Relation zu einem anderen steht, nicht folgen, dass auch das zweite Element in Relation zum ersten steht. Auch muss ein Element einer Menge nicht unbedingt zu sich selbst in Relation stehen, und aus der Tatsache, dass zwei Elemente in Relation zu einem dritten stehen, kann man im allgemeinen nicht schließen, dass das erste zum zweiten in Relation steht oder umgekehrt. Es gibt jedoch ausgezeichnete Relationen, die diese drei Bedingungen erfüllen. Diese spielen eine besonders wichtige Rolle und durchziehen die gesamte Mathematik.

Definition 2.1.11: Sei M eine Menge und $R \subseteq M \times M$ eine Relation auf M.

1. Dann nennt man R eine **Äquivalenzrelation** auf M, wenn R für beliebige Elemente $m, n, p \in M$ die folgenden Bedingungen erfüllt:
   (i) **Reflexivität:** $m \sim_R m$.
   (ii) **Symmetrie:** aus $n \sim_R m$ folgt $m \sim_R n$.
   (iii) **Transitivität:** aus $p \sim_R n$ und $n \sim_R m$ folgt $p \sim_R m$.
2. Ist R eine Äquivalenzrelation auf M, dann nennt man für jedes $m \in M$ die Menge $[m] := \{n \in M |\, n \sim_R m\}$ die **Äquivalenzklasse** von m und ein Element $n \in [m]$ einen **Repräsentanten** von $[m]$.
3. Die Menge $\{[m] |\, m \in M\}$ der Äquivalenzklassen von Elementen $m \in M$ heißt **Quotientenmenge** von M bezüglich der Relation R und wird mit M/R oder $M/\sim_R$ bezeichnet.

Da sich Äquivalenzrelationen auf einer Menge M durch ein Prädikat $P : \mathrm{Pot}(M \times M) \to \{w, f\}$ aus der Menge $\mathrm{Pot}(M \times M)$ aller Relationen auf M auswählen lassen, bilden die Äquivalenzrelationen auf einer gegebenen Menge M eine Menge (Aussonderungsaxiom). Ebenso lassen sich die Äquivalenzklassen durch ein Prädikat $P : \mathrm{Pot}(M) \to \{w, f\}$ aus der Potenzmenge $\mathrm{Pot}(M)$ auswählen und bilden somit eine Menge.

Beispiel 2.1.12: Wir untersuchen die Relationen aus Beispiel 2.1.10, 1. darauf, ob sie Äquivalenzrelationen auf der Menge der Menschen sind.

(i) Die Relation *jemanden lieben* ist keine Äquivalenzrelation auf der Menge der Menschen, denn sie muss weder reflexiv sein (nicht jeder liebt sich selbst) noch symmetrisch (Liebe wird nicht immer erwidert) noch transitiv (aus a liebt b und b liebt c folgt im allgemeinen nicht, dass a auch c liebt).

(ii) Die Relation *die gleichen Eltern haben wie* ist eine Äquivalenzrelation auf der Menge der Menschen, denn man hat die gleichen Eltern wie man selbst (reflexiv). Wenn a die gleichen Eltern hat wie b, dann hat auch b die gleichen Eltern wie a (symmetrisch). Wenn a die gleichen Eltern hat wie b und b die gleichen Eltern wie c, dann hat auch a die gleichen Eltern wie c (transitiv). Die Äquivalenzklassen sind die Mengen der Kinder eines gegebenen Elternpaars.

(iii) Die Relation *mindestens einen Elternteil gemeinsam haben mit* ist keine Äquivalenzrelation auf der Menge der Menschen, denn sie ist zwar reflexiv (jeder hat mit sich selbst mindestens einen Elternteil gemeinsam) und symmetrisch (hat a mit b mindestens einen Elternteil gemeinsam, dann hat auch b mit a mindestens einen Elternteil gemeinsam), aber nicht transitiv. Denn es ist möglich, dass b die gleiche Mutter wie a und den gleichen Vater wie c, aber a und c weder die gleiche Mutter noch den gleichen Vater haben.

(iv) Die Relation *eine Schwester sein von* ist keine Äquivalenzrelation auf der Menge der Menschen, denn sie ist zwar transitiv (ist a eine Schwester von b und b eine Schwester von c, dann ist auch a eine Schwester von c), aber nicht reflexiv (nicht jeder ist seine eigene Schwester) und nicht symmetrisch (a kann die Schwester ihres Bruders b sein, aber der Bruder b ist dann keine Schwester von a).

Beispiel 2.1.13:

1. Die Relation $(x_1, x_2, x_3) \sim (y_1, y_2, y_3) \Leftrightarrow x_1^2 + x_2^2 + x_3^2 = y_1^2 + y_2^2 + y_3^2$ auf der Menge $\mathbb{R}^3$ ist eine Äquivalenzrelation.

   Denn sie ist reflexiv und symmetrisch, und aus $x_1^2 + x_2^2 + x_3^2 = y_1^2 + y_2^2 + y_3^2$ und $y_1^2+y_2^2+y_3^2 = z_1^2+z_2^2+z_3^2$ folgt $x_1^2+x_2^2+x_3^2 = z_1^2+z_2^2+z_3^2$ (transitiv). Die Äquivalenzklassen sind Sphären (Kugeloberflächen) im $\mathbb{R}^3$ mit dem Mittelpunkt im Ursprung sowie die Menge, die nur den Ursprung enthält. Jeder Punkt auf einer Sphäre ist ein Repräsentant der Sphäre. Der einzige Repräsentant der Äquivalenzklasse, die den Ursprung enthält, ist der Ursprung selbst.

2. Für jede Menge M ist $R = \{(m, m)|\, m \in M\} \subseteq M \times M$ eine Äquivalenzrelation auf M.

   Denn $(m, m) \in R$ für alle $m \in M$ (reflexiv), $(m, n) \in R \Rightarrow m = n \Rightarrow (n, m) \in R$ (symmetrisch) und $((p, n) \in R) \wedge ((n, m) \in R) \Rightarrow (p = n) \wedge (n = m) \Rightarrow p = m \Rightarrow (p, m) \in R$ (transitiv). Die Äquivalenzklassen sind die Mengen $\{m\}$ mit $m \in M$, und die Quotientenmenge ist $M/R = \{\{m\}|m \in M\}$.

3. Die Relation $R = \{(x, y) \in \mathbb{R} \times \mathbb{R}|\, x < y\}$ ist keine Äquivalenzrelation auf $\mathbb{R}$.

   Sie ist transitiv: $((x, y) \in R) \wedge ((y, z) \in R) \Rightarrow (x < y) \wedge (y < z) \Rightarrow x < z \Rightarrow (x, z) \in R$, aber nicht reflexiv: $(x, x) \notin R$, da $x \not< x$, und nicht symmetrisch: $(x, y) \in R \Rightarrow x < y \Rightarrow y \not< x \Rightarrow (y, x) \notin R$.

Beispiel 2.1.14: Für jedes $n \in \mathbb{N}$ ist $x \sim y \Leftrightarrow n$ teilt $x - y$ eine Äquivalenzrelation auf $\mathbb{Z}$.

Denn n teilt $x - x = 0$ (reflexiv). Teilt n die Zahl $x - y$ so teilt n auch $y - x = -(x - y)$ (symmetrisch). Teilt n sowohl $x - y$ als auch $y - z$, dann teilt n auch $x - z = (x - y) + (y - z)$ (transitiv). Die Äquivalenzklassen sind die Mengen

$$[z] = \{z + k \cdot n|\, k \in \mathbb{Z}\}$$

für $z \in \mathbb{Z}$, die auch als **Restklassen** bezeichnet werden. Da die möglichen Reste bei Division durch n gerade die Zahlen $0, 1, ..., n - 1$ sind gibt es genau n Restklassen, nämlich $[0]$, $[1]...$, $[n - 1]$. Die Quotientenmenge ist die Menge

$$\mathbb{Z}/n\mathbb{Z} = \{[0], [1], ..., [n - 1]\} = \{[z] \mid z \in \mathbb{Z}\}.$$

Für $n = 12$ ist die Äquivalenzrelation aus Beispiel 2.1.14 genau das Prinzip, das einer analogen Uhr zugrundeliegt. Wird der Zeiger zu einem willkürlich gewählten Zeitpunkt t_0 auf die Null (12) gesetzt, so zeigt er zu einer Zeit t_1 auf die Ziffer 2 genau dann, wenn zwischen t_1 und t_0 $..., -22, -10, 2, 14, 26, ...$ Stunden vergangen sind. Die Äquivalenzklassen $[0], [1], ..., [11]$ entsprechen also den Ziffern 12,1,...,11 auf der Uhr. Die Äquivalenzklasse $[z]$ mit $z \in \{0, ..., 11\}$ enthält genau die Zeitpunkte, zu denen der Zeiger auf z zeigt, also $[z] = \{z + 12k|k \in \mathbb{Z}\}$. Die Quotientenmenge ist die Menge der Uhrzeiten, also der möglichen Zeigerstände.

Ein weiteres wichtiges Beispiel von Äquivalenzrelationen, das schon aus der Schule bekannt ist, ist die Beschreibung von rationalen Zahlen durch Brüche. Dabei sind zwei Brüche a/b und c/d äquivalent genau dann, wenn sie durch Kürzen oder Erweitern auseinander hervorgehen, also wenn sie die selbe rationale Zahl beschreiben.

Beispiel 2.1.15: Die Relation $(a, b) \sim (c, d) \Leftrightarrow ad = bc$ ist eine Äquivalenzrelation auf der Menge $M = \{(a, b)|a, b \in \mathbb{Z}, b \neq 0\}$.

Denn wegen $ab = ba$ gilt $(a, b) \sim (a, b)$ (Reflexivität). Aus $(a, b) \sim (c, d)$ folgt $ad = bc$ und somit auch $cb = da$, also $(c, d) \sim (a, b)$ (Symmetrie). Aus $(a, b) \sim (c, d)$ und $(c, d) \sim (e, f)$ folgt $ad = bc$ und $cf = de$. Daraus ergibt sich

$$f(ad - bc) + b(cf - de) = fad - bcf + bcf - bde = fad - bde = d(af - be) = 0.$$

Da $d \neq 0$ folgt daraus $af = be$ und somit $(a, b) \sim (e, f)$ (Transitivität).

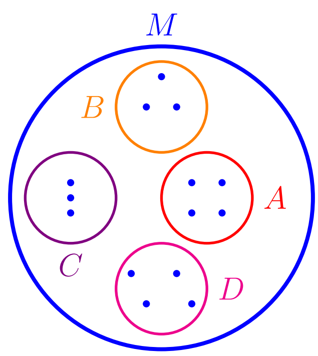
Schreibt man a/b statt (a, b), so erkennt man, dass die Äquivalenzklassen dieser Äquivalenzrelation gerade die rationalen Zahlen sind. Denn für zwei Brüche a/b und c/d mit $a, b, c, d \in \mathbb{Z}$ und $b, d \neq 0$ folgt aus $a/b = c/d$ durch Multiplizieren mit bd die Identität $ad = bc$. Umgekehrt folgt aus $ad = bc$ und $b, d \neq 0$ durch Division durch bd, dass $a/b = c/d$ gilt.

Also ist für alle $(a, b) \in M$ die Äquivalenzklasse $[(a, b)]$ die rationale Zahl, die durch den Bruch a/b und alle daraus durch Erweitern oder Kürzen hervorgehenden Brüche beschrieben wird. Die Quotientenmenge ist die Menge $M/\sim = \mathbb{Q}$ der rationalen Zahlen. Jeder Bruch a/b, der eine rationale Zahl $q \in \mathbb{Q}$ beschreibt, ist ein Repräsentant dieser Zahl. Man kann zeigen, dass jede Äquivalenzklasse $[(a, b)]$ genau einen *gekürzten Bruch* enthält, d. h. ein Element (c, d) mit $c, d \in \mathbb{Z}$, $d > 0$ und c, d teilerfremd (Übung). Die gekürzten Brüche bilden ein Repräsentandensystem.

Alternativ kann man eine Äquivalenzrelation auf einer Menge M auch als eine Zerlegung von M in nicht-leere disjunkte Teilmengen verstehen. Jedes Element von M ist in genau einer Menge in dieser Zerlegung enthalten, und jede Menge in der Zerlegung enthält mindestens ein Element von M. Eine solche Zerlegung heißt *Partition* von M.

Definition 2.1.16: Eine **Partition** einer Menge M ist eine Teilmenge $P \subseteq \mathrm{Pot}(M)$, so dass $\emptyset \notin P$ und jedes Element von M in genau einer Menge in P enthalten ist:

(P1) $\emptyset \notin P$,
(P2) $M = \cup P = \{x \in M \mid \exists A \in P : x \in A\}$,
(P3) Sind $A, B \in P$ mit $A \cap B \neq \emptyset$, so folgt $A = B$.



Eine Partition einer Menge M.

Veranschaulicht man sich Mengen als Behälter, dann kann man sich eine Partition einer Menge M wie folgt vorstellen: Man leert einen gegebenen Behälter (die Menge M) aus und verteilt seinen Inhalt auf neue Behälter, so dass keiner dieser neuen Behälter leer bleibt und nichts übrig bleibt, also alle ursprünglich in M enthaltenen Dinge in irgendeinem der neuen Behälter landen. Dann packt man diese neuen Behälter zusammen in einen Behälter (die Menge P). Da in dem ursprünglichen Behälter jedes Ding nur einmal vorhanden war und dann in genau einen Behälter einsortiert wurde, haben zwei verschiedene Behälter in der Menge P keinen gemeinsamen Inhalt und die Vereinigung ihrer Inhalte ergibt gerade den Inhalt von M.

Wir zeigen nun, dass jede Äquivalenzrelation $\sim$ auf einer Menge M eine Partition $P \subseteq \mathrm{Pot}(M)$ definiert, und umgekehrt jede Partition von M eine Äquivalenzrelation auf M. Im ersten Fall enthält die Partition genau die Äquivalenzklassen von $\sim$, ist also durch die Quotientenmenge $M/\sim$ gegeben. Im zweiten Fall definiert man eine Äquivalenzrelation, indem man zwei Elemente von M als äquivalent betrachtet, wenn sie in der selben Menge $A \in P$ enthalten sind.

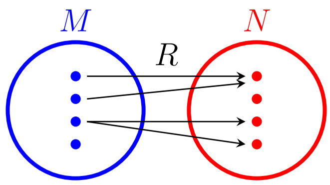
Satz 2.1.17: Sei M eine Menge.

1. Ist $\sim$ eine Äquivalenzrelation auf M, so ist die Quotientenmenge $M/\sim = \{[m] \mid m \in M\}$ eine Partition von M.
2. Ist $P \subseteq \mathrm{Pot}(M)$ eine Partition von M, so ist $m \sim n \Leftrightarrow \exists A \in P : m, n \in A$ eine Äquivalenzrelation auf M mit Quotientenmenge P.

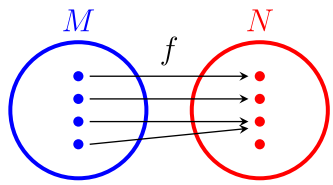
Beweis:
1⇒ 2: Wegen der Reflexivität von $\sim$ gilt $m \sim m$ und $m \in [m]$ für alle $m \in M$. Also ist $[m] \neq \emptyset$ für alle $m \in M$ (P1) und $\cup_{m \in M}[m] = M$ (P2). Sind $m, n \in M$ mit $[m] \cap [n] \neq \emptyset$, so gibt es ein $a \in [m] \cap [n]$, also ein $a \in M$ mit $a \sim m$ und $a \sim n$. Mit der Symmetrie der Äquivalenzrelation folgt dann $m \sim a$, und mit der Transitivität folgt aus $m \sim a$ und $a \sim n$ dann $m \sim n$. Ist nun $x \in [m]$, so ergibt sich $x \sim m$, und wegen $m \sim n$ folgt mit der Transitivität $x \sim n$, also $x \in [n]$. Damit ist gezeigt, dass $[m] \subseteq [n]$ gilt. Wegen der Symmetrie folgt aus $m \sim n$ aber auch $n \sim m$ und damit $[n] \subseteq [m]$ also $[n] = [m]$ (P3). Also ist $M/\sim = \{[m] \mid m \in M\}$ eine Partition von M.

2⇒1: Sei $P \subseteq \mathrm{Pot}(M)$ eine Partition von M und $m \sim n \Leftrightarrow \exists A \in P : m, n \in A$ die zugehörige Relation auf M. Diese ist per Definition symmetrisch. Sie ist auch reflexiv, denn zu jedem $m \in M$ existiert wegen $\cup P = M$ eine Menge $A \in P$ mit $m \in A$, und damit gilt $m \sim m$ für alle $m \in M$. Sind $m, n, r \in M$ mit $m \sim n$ und $n \sim r$, so gibt es Mengen $A, B \in P$ mit $m, n \in A$ und $n, r \in B$. Da $n \in A \cap B$ ist dann $A \cap B \neq \emptyset$, und mit (P3) folgt $A = B$. Also gibt es eine Menge $A = B \in P$ mit $m, r \in A$, und es folgt $m \sim r$. Also ist $\sim$ auch transitiv und damit eine Äquivalenzrelation auf M. Zu jedem Element $m \in M$ gibt es genau eine Menge $A \in P$ mit $m \in A$, und es folgt $[m] = \{n \in M \mid \exists A \in P : m, n \in A\} = A$. Also ist die Quotientenmenge gegeben durch $M/\sim = \{[m] \mid m \in M\} = P$. □

Während alle bisher betrachteten Beispiele von Relationen von der Form (M, M, R) waren, also Relationen *auf* einer gegebenen Menge M, lernen wir nun eine wichtige Klasse von Beispielen kennen, für die das nicht unbedingt der Fall sein muss, die *Abbildungen*. Diese sind Relationen der Form (M, N, R), die eine Zusatzbedingungen erfüllen, nämlich, dass zu jedem Element $m \in M$ *genau ein* Element $n \in N$ mit $(m, n) \in R$ existiert. Man ordnet also jedem Element $m \in M$ ein eindeutig bestimmtes Element $n \in N$ zu, nämlich das Element $n \in N$ mit $(m, n) \in R$. Die Menge R ist dann der Graph der durch diese Zuordnung bestimmten Abbildung.



eine Relation $R \subseteq M \times N$



eine Abbildung $f : M \to N$.

Definition 2.1.18: Eine **Abbildung** f von einer Menge M in eine Menge N ist eine Relation $f = (M, N, R)$, so dass zu jedem $m \in M$ genau ein $n \in N$ mit $(m, n) \in R$ existiert.

Man benutzt die folgenden Bezeichnungen:

- Statt $(m, n) \in R$ schreibt man auch $n = f(m)$ und nennt $f(m)$ den **Wert** von f in $m \in M$ oder das **Bild** von m unter f. Jedes Element $m \in M$ mit $f(m) = n$ heißt **Urbild** von $n \in N$.
- Die Abbildung $f = (M, N, R)$ wird oft auch mit $f : M \to N$, $m \mapsto f(m)$ bezeichnet.
- Die Menge M heißt **Definitionsbereich**, die Menge N **Bildbereich** oder **Wertebereich** und die Menge $R = \{(m, f(m)) | m \in M\}$ **Graph** der Abbildung f.
- Für eine Teilmenge $A \subseteq M$ heißt die Menge $f(A) = \{f(a) \mid a \in A\} \subseteq N$ das **Bild** von A unter der Abbildung f.
- Die Menge $f(M) = \{f(m) \,|\, m \in M\} \subseteq N$ wird auch als das **Bild der Abbildung** f und mit $\mathrm{im}(f)$ bezeichnet.
- Für eine Teilmenge $B \subseteq N$ heißt $f^{-1}(B) = \{m \in M | f(m) \in B\} \subseteq M$ das **Urbild** von B unter f.

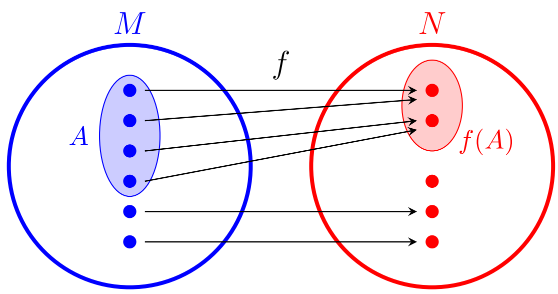
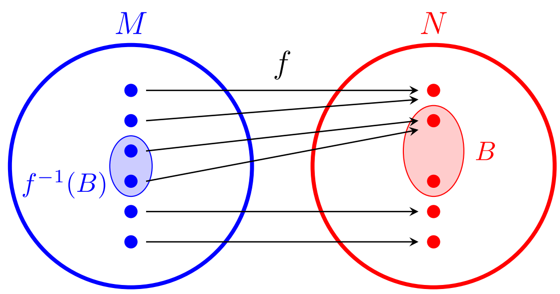


Bild $f(A)$ einer Menge $A \subseteq M$
unter einer Abbildung $f : M \to N$



Urbild $f^{-1}(B)$ einer Menge $B \subseteq M$
unter einer Abbildung $f : M \to N$.

Bemerkung 2.1.19:

1. Definitionsbereich und Wertebereich sind ein wesentlicher Teil der Abbildung und dürfen nicht weggelassen werden. Es ergibt keinen Sinn, von einer Abbildung zu sprechen, ohne ihren Definitions- und Wertebereich anzugeben!

   Insbesondere sind zwei Abbildungen $f : M \to N$ und $f' : M' \to N'$ gleich genau dann, wenn $M = M'$, $N = N'$ und $f(m) = f'(m)$ für alle $m \in M$ gilt.

2. Die Definition einer Abbildung sagt aus, dass jedem Element $m \in M$ nur *ein* Element von N zugeordnet werden darf. Zu einem Element $n \in N$ kann es dagegen beliebig viele Urbilder in M geben.

3. Der Graph einer Abbildung ist eine Verallgemeinerung des schon aus der Schule bekannten Konzepts des Graphens einer Funktion. Für eine Abbildung $f : \mathbb{R} \to \mathbb{R}$ ist der Graph der Abbildung f gerade die Menge $\Gamma_f = \{(x, f(x))|x \in \mathbb{R}\}$.

4. Für gegebene Mengen M und N bilden die Abbildungen $f : M \to N$ eine Menge. Dies folgt daraus, dass sie sich durch ein logisches Prädikat aus der Potenzmenge $\text{Pot}(M \times N)$ auswählen lassen. (Aussonderungsaxiom).

5. Aus der Definition einer Abbildung folgt, dass es zu jeder Menge N genau eine Abbildung $f : \emptyset \to N$ gibt, die sogenannte **leere Abbildung**.

   Denn eine Abbildung $f : \emptyset \to N$ ist eine Teilmenge $R \subseteq \emptyset \times N = \emptyset$, so dass es zu jedem Element $m \in \emptyset$ genau ein $n \in N$ mit $(m, n) \in R$ gibt. Da $\emptyset$ genau eine Teilmenge hat, nämlich $\emptyset$, und es keine Elemente $m \in \emptyset$ gibt, ist $R = \emptyset \subseteq \emptyset \times N$ eine Abbildung.

Beispiel 2.1.20:

1. Für jede Menge M ist $\text{id}_M : M \to M$, $m \mapsto m$ eine Abbildung, die **Identitätsabbildung** auf M.
2. Für jedes Polynom $p = a_0 + a_1x + \ldots + a_nx^n$ mit Koeffizienten $a_0, ..., a_n \in \mathbb{R}$ ist $f_p : \mathbb{R} \to \mathbb{R}$, $x \mapsto a_nx^n + ... + a_1x + a_0$ eine Abbildung.

3. Für jedes $n \in \mathbb{N}$ erhält man eine Abbildung $r_n : \mathbb{Z} \to \{0, 1, ..., n-1\}$, $z \mapsto r_n(z)$, die einer ganzen Zahl z den Rest $r_n(z)$ von z bei der Division durch n zuordnet.

4. Ein logisches Prädikat auf einer Menge M ist eine Abbildung $P : M \to \{w, f\}$ von der Menge M in die Menge $\{w, f\}$ der Wahrheitswerte.

5. Die **Dirichlet-Funktion** $\phi : \mathbb{R} \to \{0, 1\}$ ist die Abbildung definiert durch

$$\phi(x) = \begin{cases} 1 & x \in \mathbb{Q} \\ 0 & x \in \mathbb{R} \setminus \mathbb{Q}. \end{cases}$$

6. Eine Abbildung $f : M \to N$ heißt **konstant** auf einer Teilmenge $A \subseteq M$, wenn $f(a) = f(a')$ für alle $a, a' \in A$. Statt von einer auf M konstanten Abbildung spricht man auch oft einfach von einer **konstanten Abbildung**.
7. Eine Abbildung $f : \mathbb{N}_0 \to M$, $n \mapsto f(n)$ in eine Menge M nennt man auch eine **Folge** mit Werten in M und schreibt $x_n = f(n)$ und $(x_n)_{n\in\mathbb{N}_0}$ statt $f : \mathbb{N}_0 \to M$.

8. Die durch $M = \mathbb{Q}$, $N = \mathbb{Z}$ und $R = \{(q, z) \in \mathbb{Q} \times \mathbb{Z}|q \geq z\}$ gegebene Relation ist keine Abbildung. Denn zu jedem $q \in \mathbb{Q}$ gibt es unendlich viele ganze Zahlen $z \in \mathbb{Z}$ mit $q \geq z$, also mit $(q, z) \in R$.
9. Für jede Menge M gibt es genau eine Äquivalenzrelation auf M, die gleichzeitig eine Abbildung ist, nämlich die Äquivalenzrelation $m \sim m' \Leftrightarrow m = m'$ aus Beispiel 2.1.13, 2, die mit der Identitätsabbildung $\text{id}_M : M \to M$ übereinstimmt.

   Denn für diese Äquivalenzrelation existiert zu jedem $m \in M$ genau ein $m' \in M$ mit $(m, m') \in R$, nämlich $m' = m$. Ist umgekehrt $R \subseteq M \times M$ eine Abbildung, dann gibt es zu jedem $m \in M$ genau ein $m' \in M$ mit $(m, m') \in R$. Ist R außerdem eine Äquivalenzrelation, so gilt für jedes $m \in M$ wegen der Reflexivität $m \sim m$, d. h. $(m, m) \in R$ und somit $m = m'$. Also folgt $R = \{(m, m) | m \in M\}$.

Anhand dieser Beispiele sieht man, dass eine Abbildung nicht unbedingt durch eine Rechenvorschrift oder Formel definiert werden muss. Auch logische Prädikate, Beschreibungen oder Fallunterscheidungen wie in Beispiel 2.1.20, 5. können zur Definition einer Abbildung benutzt werden. Während diese *konkreten* Beispiele einen Eindruck von der Vielfalt der Abbildungen vermitteln, wollen wir nun noch einige *allgemeine* Konstruktionen betrachten, die es uns erlauben, aus gegebenen Abbildungen neue Abbildungen zu konstruieren.

Beispiel 2.1.21:

1. Ist M eine Menge und $A \subseteq M$ eine Teilmenge, so erhält man die **Inklusionsabbildung** $\iota_A : A \to M$, $a \mapsto a$. Diese unterscheidet sich von id_M nur durch ihren Definitionsbereich und von id_A nur durch den Wertebereich.

2. Ist $f : M \to N$, $m \mapsto f(m)$ eine Abbildung und $A \subseteq M$ eine Teilmenge, dann erhält man eine neue Abbildung $f|_A : A \to N$, $a \mapsto f(a)$. Diese heißt **Einschränkung** oder **Restriktion** von f auf A und unterscheidet sich von f nur durch ihren Definitionsbereich.

3. Ist $f : M \to N$, $m \mapsto f(m)$ eine Abbildung und $B \subseteq N$ eine Teilmenge mit $f(M) \subseteq B$, dann erhält man eine neue Abbildung $f^B : M \to B$, $m \mapsto f(m)$. Diese heißt **Koeinschränkung** von f auf B und unterscheidet sich von f nur durch ihren Wertebereich.

4. Ist $\sim$ eine Äquivalenzrelation auf einer Menge M, so erhält man eine Abbildung $\pi : M \to M/\sim$, $m \mapsto [m]$ von M in die Quotientenmenge $M/\sim$, die jedem Element $m \in M$ seine Äquivalenzklasse $[m] \in M/\sim$ zuordnet. Sie heißt **kanonische Surjektion**.

5. Ist $\sim$ eine Äquivalenzrelation auf M und $f : M \to N$ eine Abbildung, die auf den Äquivalenzklassen von $\sim$ konstant ist, also mit $m \sim m' \Rightarrow f(m) = f(m')$, so erhält man eine Abbildung $f_\sim : M/\sim \to N$, $[m] \mapsto f(m)$ von der Quotientenmenge nach N.

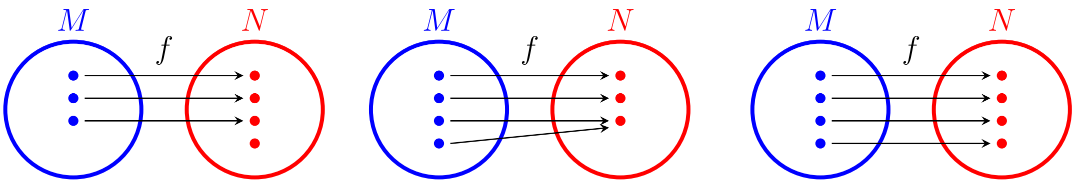
   Die Bedingung, dass f auf den Äquivalenzklassen konstant ist, ist dabei essentiell, denn sonst wäre die Abbildung nicht **wohldefiniert**. Ein und derselben Äquivalenzklasse $[m] = [m']$, die durch zwei verschiedene Repräsentanten $m, m' \in M$ beschrieben wird, könnten dann verschiedene Werte, nämlich $f(m) \in N$ und $f(m') \in N$ zugeordnet werden. Es muss also gesichert sein, dass $f_\sim$ nicht von der Wahl der Repräsentanten abhängt.

6. Ist $\sim$ die Äquivalenzrelation auf $\mathbb{R}^3$ aus Beispiel 2.1.13 mit $(x_1, x_2, x_3) \sim (y_1, y_2, y_3) \Leftrightarrow x_1^2 + x_2^2 + x_3^2 = y_1^2 + y_2^2 + y_3^2$ und $f : \mathbb{R}^3 \to \mathbb{R}$, $(x_1, x_2, x_3) \mapsto \sqrt{x_1^2 + x_2^2 + x_3^2}$ die Abbildung, die einem Punkt im $\mathbb{R}^3$ seinen Abstand vom Ursprung zuordnet, dann erfüllt f die Bedingung $(x_1, x_2, x_3) \sim (y_1, y_2, y_3) \Rightarrow f(x_1, x_2, x_3) = f(y_1, y_2, y_3)$. Die zugehörige Abbildung $f_\sim : M/\sim \to \mathbb{R}$ ordnet einer Sphäre mit dem Mittelpunkt im Ursprung ihren Radius zu und dem Ursprung den Wert 0.

Wir charakterisieren nun Abbildungen dadurch, wie viele Elemente $m \in M$ sie auf ein gegebenes Element $n \in N$ abbilden. Wie schon festgestellt, kann für eine Abbildung $f : M \to N$ ein gegebenes Element $n \in N$ entweder kein, genau ein oder mehrere *Urbilder* haben. Entsprechend hat die Gleichung $n = f(m)$ keine, genau eine oder mehrere Lösungen $(m, n) \in M \times N$.

Definition 2.1.22: Sei $f : M \to N$ eine Abbildung.

1. Die Abbildung f heißt **injektiv** oder **Injektion**, wenn jedes Element $n \in N$ höchstens ein Urbild hat:
$$\forall m, m' \in M : \; f(m) = f(m') \Rightarrow m = m'.$$
2. Die Abbildung f heißt **surjektiv** oder **Surjektion**, wenn jedes Element $n \in N$ mindestens ein Urbild hat:
$$\forall n \in N : \exists m \in M : n = f(m).$$
3. Die Abbildung f heißt **bijektiv** oder **Bijektion**, wenn jedes Element $n \in N$ genau ein Urbild hat, also genau dann, wenn sie injektiv und surjektiv ist:
$$\forall n \in N : \exists! m \in M : n = f(m),$$
Sind M und N zwei Mengen, so dass eine Bijektion $f : M \to N$ existiert, sagt man M und N sind **in Bijektion**.



eine injektive Abbildung $f : M \to N$ | eine surjektive Abbildung $f : M \to N$ | eine bijektive Abbildung $f : M \to N$

Beispiel 2.1.23:

1. Für jede Menge M ist die Identitätsabbildung $\mathrm{id}_M : M \to M$ bijektiv.
2. Für jede Teilmenge $A \subseteq M$ ist die Inklusionsabbildung $\iota_A : A \to M$ injektiv.
3. Für jede Abbildung $f : M \to N$ ist die Koeinschränkung $f : M \to f(M)$ surjektiv.
4. Ist $\sim$ eine Äquivalenzrelation auf M, so ist die kanonische Surjektion $\pi : M \to M/\sim$, $m \mapsto [m]$ surjektiv. Deswegen heißt sie kanonische Surjektion.
5. Für alle $m \in \mathbb{N}$ stehen die Mengen $\mathbb{Z}$ und $m\mathbb{Z} = \{mz | z \in \mathbb{Z}\}$ zueinander in Bijektion. Eine Bijektion ist gegeben durch $f : \mathbb{Z} \to m\mathbb{Z}$, $z \mapsto mz$.

Bemerkung 2.1.24:

1. Sind M, N *endliche* Mengen mit $|M| = |N|$, so gilt für alle Abbildungen $f : M \to N$: f injektiv $\Leftrightarrow$ f surjektiv $\Leftrightarrow$ f bijektiv.

   Denn f ist injektiv genau dann, wenn die Teilmenge $f(M) \subseteq N$ genau so viele Elemente wie M hat, was genau dann, der Fall ist, wenn $f(M) = N$ gilt, also f surjektiv ist.
2. Ist M eine *unendliche* Menge, so gibt es Abbildungen $f : M \to M$, die injektiv, aber nicht surjektiv und Abbildungen $g : M \to M$, die surjektiv, aber nicht injektiv sind.

   Ein Beispiel sind die Abbildungen $f : \mathbb{N} \to \mathbb{N}$, $n \mapsto n + 1$ und $g : \mathbb{N} \to \mathbb{N}$ mit $g(1) = 1$ und $g(n) = n - 1$ für $n \geq 2$.

3. Sind M, N *endliche* Mengen und ist $f : M \to N$ eine Bijektion, dann haben M und N gleich viele Elemente.

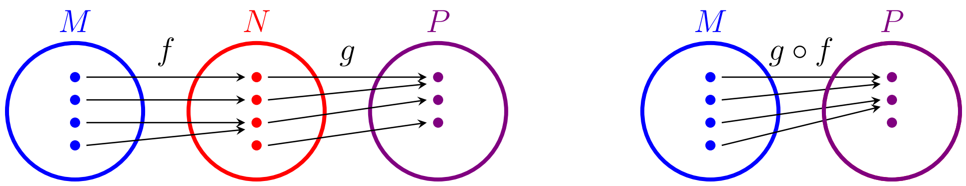
   Denn aus der Injektivität von f folgt, dass N mindestens so viele Elemente wie M haben muss, und aus der Surjektivität, dass N höchstens so viele Elemente wie M haben kann.

Mengen, die zueinander in Bijektion stehen, können alleine mit mengentheoretischen Mitteln nicht unterschieden werden. So muss man beispielsweise in Beispiel 2.1.23, 5. Teilbarkeitsaussagen oder andere Strukturen auf $\mathbb{Z}$ benutzen, um die beiden Mengen zu unterscheiden. Aus diesem Grund betrachtet man solche Mengen als mengentheoretisch gleich. Es ist dabei irrelevant, wie die einzelnen Elemente in den Mengen heißen oder welche Eigenschaften sie zusätzlich zu ihrer Rolle als Elemente gewisser Mengen besitzen. Dies bedeutet insbesondere, dass man endliche Mengen nur anhand der Anzahl ihrer Elemente unterscheidet.

Wir werden nun Abbildungen zu neuen Abbildungen zusammensetzen oder *verketten*. Die Verkettung reeller Funktionen ist schon aus der Schule bekannt. Hier erinnern wir aber noch einmal daran, dass Definitions- und Wertebereich ein essentieller Teil der Abbildungen sind. Zwei Abbildungen können nur dann verkettet werden, wenn der Definitionsbereich der einen mit dem Wertebereich der anderen übereinstimmt.

Definition 2.1.25: Seien $f : M \to N$ und $g : N \to P$ Abbildungen. Dann ist die **Verkettung** oder **Komposition** von f und g die Abbildung $g \circ f : M \to P$, $x \mapsto g(f(x))$.

Es ist eine hilfreiche Übung, die Verkettung zweier Abbildungen als Relation auszudrücken und zu beweisen, dass man so tatsächlich eine Abbildung im Sinn von Definition 2.1.18 erhält.



Abbildungen $f : M \to N$, $g : N \to P$ und Verkettung $g \circ f : M \to P$

Satz 2.1.26 (Satz über Verkettungen)**:**

1. Für jede Abbildung $f : M \to N$ gilt $\mathrm{id}_N \circ f = f \circ \mathrm{id}_M = f$.
2. Die Verkettung von Abbildungen ist **assoziativ**:
   für beliebige Abbildungen $f : M \to N$, $g : N \to P$ und $h : P \to Q$ gilt $(h \circ g) \circ f = h \circ (g \circ f)$.
3. Für beliebige Abbildungen $f : M \to N$ und $g : N \to P$ gilt:
   (i) sind f und g injektiv, dann ist auch $g \circ f$ injektiv,
   (ii) sind f und g surjektiv, dann ist auch $g \circ f$ surjektiv,
   (iii) sind f und g bijektiv, dann ist auch $g \circ f$ bijektiv.

Beweis:
1. Für alle $m \in M$ gilt: $(\mathrm{id}_N \circ f)(m) = \mathrm{id}_N(f(m)) = f(m) = f(\mathrm{id}_M(m)) = (f \circ \mathrm{id}_M)(m)$. Daraus folgt die Gleichheit der beiden Abbildungen.

2. Für beliebige Abbildungen $f : M \to N$, $g : N \to P$, $h : P \to Q$ und alle $m \in M$ gilt

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

Daraus folgt die Gleichheit der beiden Abbildungen.

3. (i) Seien $f : M \to N$ und $g : N \to P$ injektive Abbildungen, und seien $m, m' \in M$ mit $(g \circ f)(m) = g(f(m)) = g(f(m')) = (g \circ f)(m')$. Dann folgt aus der Injektivität von g, dass $f(m) = f(m')$ und aus der Injektivität von f, dass $m = m'$. Also ist $g \circ f : M \to P$ injektiv.

3. (ii) Seien $f : M \to N$ und $g : N \to P$ surjektive Abbildungen. Dann existiert wegen der Surjektivität von g zu jedem $p \in P$ ein $n \in N$ mit $g(n) = p$. Wegen der Surjektivität von f existiert dann zu $n \in N$ ein $m \in M$ mit $f(m) = n$. Daraus folgt $(g \circ f)(m) = g(f(m)) = g(n) = p$. Also existiert zu jedem $p \in P$ ein $m \in M$ mit $(g \circ f)(m) = p$, und $g \circ f : M \to P$ ist surjektiv.

3.(iii) Sind f und g bijektiv, so sind f und g injektiv und surjektiv. Aus 3.(i) folgt dann, dass die Abbildung $g \circ f : M \to P$ injektiv ist, und aus 3.(ii), dass sie surjektiv ist. Also ist die Abbildung $g \circ f : M \to P$ bijektiv. □

Die *Assoziativität* einer Verkettung erlaubt es uns, Klammern wegzulassen, wenn mehr als zwei Abbildungen verkettet werden. Alle möglichen Klammerungen einer Verkettung von Abbildungen liefern die gleiche Abbildung, und deswegen brauchen wir keine Klammern anzugeben. Die erste Aussage in Satz 2.1.26 erlaubt es uns, beim Verketten von Abbildungen beliebig viele Identitätsabbildungen "dazwischenzuschalten", ohne etwas am Ergebnis zu ändern. Der letzte Teil des Satzes sagt uns, dass die Eigenschaften Injektivität, Surjektivität und Bijektivität unter der Verkettung von Abbildungen erhalten bleiben.

Der Begriff der Verkettung von Abbildungen liefert außerdem eine alternative Beschreibung der Eigenschaften Bijektivität, Injektivität und Surjektivität. Wir können die Injektivität, Surjektivität und Bijektivität einer Abbildung $f : M \to N$ nämlich auch durch die Existenz gewisser Abbildungen $g : N \to M$ charakterisieren, die sich auf verschiedene Weise zu den Identitätsabbildungen id_M oder id_N verketten.

Satz 2.1.27: Sei $M \neq \emptyset$ und $f : M \to N$ eine Abbildung. Dann gilt:

1. Die Abbildung f ist surjektiv genau dann, wenn es eine Abbildung $g : N \to M$ mit $f \circ g = \mathrm{id}_N$ gibt. Die Abbildung g heißt dann ein **Rechtsinverses** von f.

2. Die Abbildung f ist injektiv genau dann, wenn es eine Abbildung $g : N \to M$ mit $g \circ f = \mathrm{id}_M$ gibt. Die Abbildung g heißt dann ein **Linksinverses** von f.

3. Die Abbildung f ist bijektiv genau dann, wenn es eine Abbildung $g : N \to M$ mit $f \circ g = \mathrm{id}_N$ und $g \circ f = \mathrm{id}_M$ gibt. Die Abbildung g ist dann eindeutig bestimmt und heißt die zu f **inverse Abbildung** oder die **Umkehrabbildung** von f. Man schreibt dann $g = f^{-1} : N \to M$.

Beweis:
1. $\Rightarrow$: Sei $f : M \to N$ surjektiv. Dann können wir zu jedem Element $n \in N$ ein Urbild $m_n \in M$ mit $f(m_n) = n$ wählen[5] und eine Abbildung $g : N \to M$, $n \mapsto m_n$ definieren. Es folgt $(f \circ g)(n) = f(m_n) = n$ für alle $n \in N$ und somit $f \circ g = \mathrm{id}_N$.

[5]Dies ist durch das Auswahlaxiom garantiert.

$\Leftarrow$: Seien $f : M \to N$ und $g : N \to M$ Abbildungen mit $f \circ g = \mathrm{id}_N$. Dann gilt $(f \circ g)(n) = f(g(n)) = n$ für alle $n \in N$. Also existiert zu jedem $n \in N$ ein Element $m \in M$ mit $f(m) = n$, nämlich $m = g(n)$, und somit ist f surjektiv.

2. $\Rightarrow$: Sei $f : M \to N$ injektiv. Für $n \in N$ gilt dann entweder $f^{-1}(\{n\}) = \emptyset$ oder es existiert genau ein $m_n \in M$ mit $f(m_n) = n$. Wir definieren eine Abbildung $g : N \to M$ durch

$$g(n) = \begin{cases} m_n \text{ mit } f(m_n) = n & \text{falls } f^{-1}(\{n\}) \neq \emptyset \\ m \in M & \text{falls } f^{-1}(\{n\}) = \emptyset, \end{cases}$$

wobei $m \in M$ ein beliebiges, fest gewähltes Element von M ist. Daraus folgt $(g \circ f)(x) = g(f(x)) = m_{f(x)} = x$ für alle $x \in M$, denn $f(x_{f(x)}) = f(x)$, und wegen der Injektivität von f folgt $x = x_{f(x)}$. Damit gilt $g \circ f = \mathrm{id}_M$.

$\Leftarrow$: Seien $f : M \to N$ und $g : N \to M$ Abbildungen mit $g \circ f = \mathrm{id}_M$. Dann folgt für alle $m, m' \in M$ aus $f(m) = f(m')$ auch $m = \mathrm{id}_M(m) = (g \circ f)(m) = g(f(m)) = g(f(m')) = (g \circ f)(m') = \mathrm{id}_M(m') = m'$. Somit ist f injektiv.

3.$\Rightarrow$: Ist f bijektiv, so ist f injektiv und surjektiv. Nach 1. existiert dann eine Abbildung $g : N \to M$ mit $f \circ g = \mathrm{id}_N$ und nach 2. eine Abbildung $g' : N \to M$ mit $g' \circ f = \mathrm{id}_M$. Um die Existenz von g zu beweisen, ist noch zu zeigen, dass $g = g'$ gilt. Dazu berechnen wir

$$g' = g' \circ \mathrm{id}_N = g' \circ (f \circ g) = (g' \circ f) \circ g = \mathrm{id}_M \circ g = g,$$

wobei im ersten und fünften Schritt Satz 2.1.26, 1. und im dritten Schritt die Assoziativität der Verkettung benutzt wurde. Die Eindeutigkeit der Abbildung g ergibt sich dann ganz ähnlich. Sind $g, g' : N \to M$ Abbildungen mit $g \circ f = g' \circ f = \mathrm{id}_M$ und $f \circ g = f \circ g' = \mathrm{id}_N$, so folgt aus dem letzten Schritt bereits $g = g'$.

$\Leftarrow$: Sind $f : M \to N$ und $g : N \to M$ Abbildungen mit $f \circ g = \mathrm{id}_N$ und $g \circ f = \mathrm{id}_M$, dann folgt aus 1., dass f surjektiv ist und aus 2., dass f injektiv ist. Also ist f bijektiv. $\square$

2.2 Gruppen

In diesem Abschnitt beschäftigen wir uns zunächst mit *Verknüpfungen* - Abbildungen, mit denen wir aus zwei gegebenen Elementen einer Menge ein neues Element bilden können. Anschließend lernen wir den wichtigen Begriff der *Gruppe* kennen - eine Menge mit einer Verknüpfung, die bestimmte zusätzliche Bedingungen erfüllt. Gruppen spielen eine sehr wichtige Rolle in der Mathematik und auch in der Physik, da sie Symmetrien beschreiben.

Definition 2.2.1: Sei M eine Menge. Eine **Verknüpfung** auf M ist eine Abbildung $\circ : M \times M \to M$, $(a, b) \mapsto a \circ b$, die jedem geordneten Paar (a, b) von Elementen $a, b \in M$ ein Element $a \circ b \in M$ zuordnet.

1. Eine Verknüpfung $\circ : M \times M \to M$ heißt **assoziativ**, wenn für alle $a, b, c \in M$ gilt: $a \circ (b \circ c) = (a \circ b) \circ c$.
2. Eine Verknüpfung $\circ : M \times M \to M$ heißt **kommutativ**, wenn für alle $a, b \in M$ gilt: $a \circ b = b \circ a$.

Beispiel 2.2.2:

1. Die Addition und die Multiplikation sind kommutative und assoziative Verknüpfungen auf den Mengen $\mathbb{N}$, $\mathbb{N}_0$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$.

2. Sei X eine Menge und $\mathrm{Abb}(X,X)$ die Menge der Abbildungen $f : X \to X$. Dann ist die Verkettung von Abbildungen $\circ : \mathrm{Abb}(X,X) \times \mathrm{Abb}(X,X) \to \mathrm{Abb}(X,X)$, $(g,f) \mapsto g \circ f$ mit $(g \circ f)(x) = g(f(x))$ eine Verknüpfung auf $\mathrm{Abb}(X,X)$, die nach Satz 2.1.26, 2. assoziativ ist. Hat X mehr als zwei Elemente, so ist sie nicht kommutativ (Beweis: Übung).

3. Sei X eine Menge. Dann sind $\circ : \mathrm{Pot}(X) \times \mathrm{Pot}(X) \to \mathrm{Pot}(X)$, $(U,V) \mapsto U \cap V$ und $\circ : \mathrm{Pot}(X) \times \mathrm{Pot}(X) \to \mathrm{Pot}(X)$, $(U,V) \mapsto U \cup V$ zwei verschiedene Verknüpfungen auf $\mathrm{Pot}(X)$. Diese sind nach Definition 2.1.2 assoziativ und nach Lemma 2.1.6 kommutativ.

4. Die Verknüpfung $\circ : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$, $(m,n) \mapsto 2^{n+m}$ ist kommutativ, aber nicht assoziativ. Denn es gilt $m \circ n = 2^{n+m} = n \circ m$ für alle $n, m \in \mathbb{N}$, aber für $m = n = 1$ und $p = 2$

$$(m \circ n) \circ p = 2^{n+m} \circ p = 2^{2^{n+m}+p} = 2^{2^2+2} = 2^6, \quad m \circ (n \circ p) = m \circ 2^{n+p} = 2^{m+2^{n+p}} = 2^{1+2^3} = 2^9.$$

Bemerkung 2.2.3:

1. Man muss eine Verknüpfung nicht unbedingt mit dem Symbol "$\circ$" bezeichnen. Oft werden stattdessen auch die Symbole "$*$", "$\bullet$", "$\star$", "$\cdot$" benutzt. Kommutative Verknüpfungen bezeichnet man oft auch mit "$+$".

2. Mit vollständiger Induktion kann man beweisen, dass für eine assoziative Verknüpfung $\circ$ auf einer Menge M alle möglichen Klammerungen eines Ausdrucks $a_1 \circ a_2 \circ \ldots \circ a_n$ das gleiche Element von M liefern. Aus diesen Gründen darf man die Klammern weglassen und stattdessen $a_1 \circ ... \circ a_n$ schreiben.

3. Eine Verknüpfung $\circ$ auf einer Menge M besteht immer aus zwei Dingen: einer *Verknüpfungsvorschrift*, die angibt, wie man aus zwei gegebenen Elementen $a, b \in M$ das Element $a \circ b$ berechnet, und der Menge M selbst. Dieselbe Vorschrift kann wie Beispiel 2.2.2 1. Verknüpfungen auf verschiedenen Mengen liefern. Umgekehrt kann es auf einer Menge M wie in Beispiel 2.2.2, 3. verschiedene Verknüpfungen geben.

Offensichtlich gibt es in manchen der Mengen in Beispiel 2.2.2 Elemente, die sich bezüglich der Verknüpfung *neutral* verhalten, also die Elemente, mit denen sie verknüpft werden, nicht ändern. So gilt beispielsweise für die Verknüpfung $+ : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ die Identität $r + 0 = 0 + r = r$ für alle $r \in \mathbb{R}$ und für die Verknüpfung $\circ : \mathrm{Pot}(X) \times \mathrm{Pot}(X) \to \mathrm{Pot}(X)$, $(U,V) \mapsto U \cap V$ die Identität $X \circ U = U \circ X = U$ für alle $U \in \mathrm{Pot}(X)$. Solche Elemente bezeichnet man als *neutrale Elemente*. Im ersten Fall gibt es außerdem zu jedem $r \in \mathbb{R}$ ein Element $s \in \mathbb{R}$ mit $r + s = s + r = 0$, nämlich $s = -r$. Solche Elemente bezeichnet man als *Inverse*.

Definition 2.2.4: Sei $\circ : M \times M \to M$ eine assoziative Verknüpfung auf M.

1. Ein Element $e \in M$ heißt **neutrales Element** für die Verknüpfung $\circ$, wenn für alle $a \in M$ gilt $e \circ a = a = a \circ e$.
2. Existiert ein neutrales Element $e \in M$, dann heißt ein Element $a' \in M$ **Inverses** eines Elements $a \in M$, wenn $a' \circ a = e = a \circ a'$. In diesem Fall nennt man a **invertierbar**.

Für manche Verknüpfungen in Beispiel 2.2.2 existieren neutrale Elemente oder Inverse und für andere nicht. Allerdings gibt es in keinem der betrachteten Beispiele mehr als ein neutrales Element oder mehr als ein Inverses zu einem gegebenen Element. Dass dies nicht an der Auswahl der Beispiele liegt, zeigt das folgende Lemma.

Lemma 2.2.5: Sei $\circ : M \times M \to M$ eine assoziative Verknüpfung auf M. Dann existiert höchstens ein neutrales Element $e \in M$ und zu jedem $a \in M$ höchstens ein Inverses. Dieses wird mit a^{-1} bezeichnet.

Beweis:
Sind $e \in M$ und $e' \in M$ zwei neutrale Elemente, so gilt nach Definition 2.2.4, 1.

$$e \overset{e' \text{ neutral}}{=} e \circ e' \overset{e \text{ neutral}}{=} e'.$$

Sei $e \in M$ das neutrale Element und $a', a'' \in M$ zwei Inverse zu $a \in M$. Dann gilt

$$a' \overset{e \text{ neutral}}{=} a' \circ e \overset{a \circ a''=e}{=} a' \circ (a \circ a'') \overset{\text{assoz.}}{=} (a' \circ a) \circ a'' \overset{a' \circ a=e}{=} e \circ a'' \overset{e \text{ neutral}}{=} a''. \qquad \square$$

Mit Hilfe der Konzepte eines neutralen Elements und eines Inversen können wir nun den Begriff der Gruppe, des Monoids und der Halbgruppe formulieren.

Definition 2.2.6: Sei M eine Menge und $\circ : M \times M \to M$ eine Verknüpfung. Dann nennt man $(M, \circ)$ eine **Gruppe**, wenn die folgenden drei Bedingungen erfüllt sind:

(G1) **Assoziativität:** für alle $a, b, c \in M$ gilt $a \circ (b \circ c) = (a \circ b) \circ c$.
(G2) **Neutrales Element:** es gibt ein Element $e \in M$ mit $a \circ e = a = e \circ a$ für alle $a \in M$.
(G3) **Inverse:** zu jedem Element $a \in M$ gibt es ein Element $a^{-1} \in M$ mit $a \circ a^{-1} = e = a^{-1} \circ a$.

Gilt nur (G1), so nennt man $(M, \circ)$ eine **Halbgruppe**. Gelten (G1) und (G2), so nennt man $(M, \circ)$ ein **Monoid**. Eine Gruppe mit einer kommutativen Verknüpfung nennt man auch eine **abelsche Gruppe**[6]. Die Anzahl der Elemente in einer Gruppe $(M, \circ)$ heißt die **Ordnung** von $(M, \circ)$ und wird mit $|M|$ bezeichnet.

Bemerkung 2.2.7: Die Verknüpfung einer Gruppe wird auch häufig mit $\cdot$ statt mit $\circ$ bezeichnet und *Gruppenmultiplikation* genannt. Manchmal lässt man das Verknüpfungszeichen auch ganz weg und schreibt ab statt $a \circ b$ oder $a \cdot b$. Das Einselement wird statt mit e auch oft mit 1 bezeichnet. In abelschen Gruppen schreibt man oft $+$ statt $\cdot$, 0 statt e und $-a$ statt a^{-1}. Diese Notation sollte *nur für abelsche Gruppen* benutzt werden.

Bemerkung 2.2.8: Man kann in der Definition einer Gruppe die Axiome (G2) und (G3) durch abgeschwächte Forderungen ersetzen, nämlich

(G2') Es gibt ein Element $e \in M$ mit $e \circ a = a$ für alle $a \in M$.
(G3') Zu jedem Element $a \in M$ gibt es ein Element $a^{-1} \in M$ mit $a^{-1} \circ a = e$.

[6] Nach dem norwegischen Mathematiker Niels Henrik Abel (1802-1829).

Denn zu $a^{-1} \in M$ mit $a^{-1} \circ a = e$ existiert nach (G3') ein $a' \in M$ mit $a' \circ a^{-1} = e$ und somit

$$a \circ a^{-1} \overset{(G2')}{=} e \circ a \circ a^{-1} \overset{(G3')}{=} a' \circ a^{-1} \circ a \circ a^{-1} \overset{(G3')}{=} a' \circ e \circ a^{-1} \overset{(G2')}{=} a' \circ a^{-1} = e. \qquad (2)$$

Daraus ergibt sich dann auch $a \cdot e \overset{(G3')}{=} a \circ a^{-1} \circ a \overset{(2)}{=} e \circ a \overset{(G2')}{=} a$. (*In einer Gruppe sind Linksinverse Rechtsinverse und Linksneutrale Rechtsneutrale*).

Die Forderungen $a \circ e = e$ und $a \circ a^{-1} = e$ in Definition 2.2.6 sind also unnötig, da sie aus (G2') und (G3') folgen. Allerdings gilt dies nur für Gruppen, nicht aber für Halbgruppen oder Monoide, weswegen man oft die kompaktere Formulierung in Definition 2.2.6 wählt.

Beispiel 2.2.9:

1. Bezeichnen wir mit + die Addition ganzer, rationaler oder reeller Zahlen, so sind sind die Mengen $(\mathbb{Z}, +)$, $(\mathbb{Q}, +)$, $(\mathbb{R}, +)$ abelsche Gruppen mit neutralem Element 0 und Inversen $a^{-1} = -a$.

2. Bezeichnen wir mit $\cdot$ die Multiplikation rationaler oder reeller Zahlen, so sind $(\mathbb{Q}, \cdot)$ und $(\mathbb{R}, \cdot)$ Monoide mit neutralem Element 1. Die Mengen $(\mathbb{R} \setminus \{0\}, \cdot)$, $(\mathbb{Q} \setminus \{0\}, \cdot)$ sind abelsche Gruppen mit neutralem Element 1 und Inversem $a^{-1} = 1/a$.

3. $(\mathbb{N}_0, +)$ ist ein Monoid, aber keine Gruppe, denn 0 ist ein neutrales Element, aber kein Element ausser 0 hat ein Inverses. Ebenso ist $(\mathbb{Z}, \cdot)$ ein Monoid, aber keine Gruppe, denn 1 ist ein neutrales Element, aber nur 1 und -1 haben Inverse.

4. Die Menge $(\mathbb{N}, +)$ ist eine Halbgruppe, aber kein Monoid, denn es existiert kein neutrales Element.

5. Jede ein-elementige Menge $M = \{m\}$ ist eine Gruppe mit der Verknüpfung $\circ : \{m\} \times \{m\} \to m$, $(m, m) \mapsto m \circ m = m$. Sie besteht nur aus dem neutralen Element $m \in M$. Eine solche Gruppe bezeichnet man auch als eine **triviale Gruppe**.

6. Sind $(G, \circ_G)$ und $(H, \circ_H)$ Gruppen so hat auch die Menge $G \times H$ die Struktur einer Gruppe mit der Verknüpfung $\circ : (G \times H) \times (G \times H) \to G \times H$, $(g, h) \circ (g', h') = (g \circ_G g', h \circ_H h')$ (Beweis: Übung). Diese Gruppe nennt man das **direkte Produkt** der Gruppen G und H.

7. Eine endliche Gruppe $(G, \circ)$ kann durch eine **Multiplikationstafel** beschrieben werden. Für eine n-elementige Gruppe $G = \{g_1, ..., g_n\}$ ist dies eine Tabelle mit n Zeilen und n Spalten, wo in die $i$te Zeile und $j$te Spalte das Produkt $g_i \circ g_j$ eingetragen wird. Dabei wählt man $g_1 = e$.

$\circ$	e	g_2	g_3	$\cdots$	g_n
e	e	g_2	g_3	$\cdots$	g_n
g_2	g_2	$g_2 \circ g_2$	$g_2 \circ g_3$	$\cdots$	$g_2 \circ g_n$
g_3	g_3	$g_3 \circ g_2$	$g_3 \circ g_3$	$\cdots$	$g_3 \circ g_n$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$
g_n	g_n	$g_n \circ g_2$	$g_n \circ g_3$	$\cdots$	$g_n \circ g_n$

Wie erkennt man an der Multiplikationstafel, ob eine Gruppe abelsch ist?

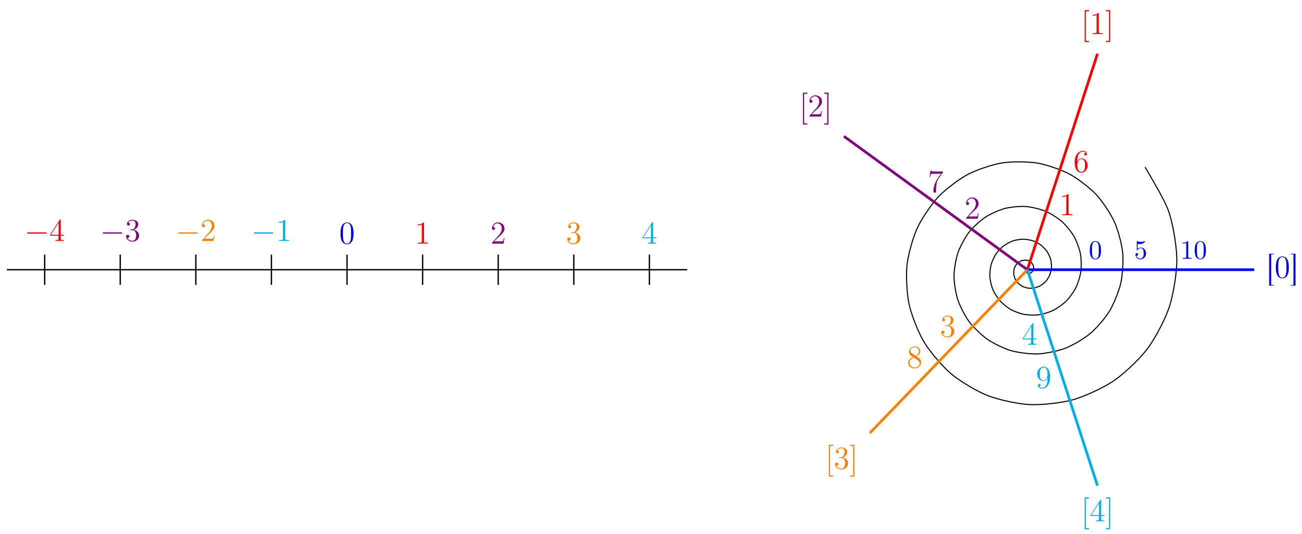


Abbildung 1: Die Quotientenmenge $\mathbb{Z}/5\mathbb{Z}$.

Besonders wichtige Beispiele von Gruppen sind die **Restklassengruppen** $(\mathbb{Z}/n\mathbb{Z}, +)$ für $n \in \mathbb{N}$. Wir fixieren dazu eine natürliche Zahl $n \in \mathbb{N}$ und betrachten auf $\mathbb{Z}$ die Äquivalenzrelation $x \sim y \Leftrightarrow n$ teilt $x - y$ aus Beispiel 2.1.14. Ihre Äquivalenzklassen sind die Mengen $[m] = \{m + kn | k \in \mathbb{Z}\}$ für $m \in \mathbb{Z}$, die auch als *Restklassen* bezeichnet werden. Es gibt genau n Restklassen, nämlich $[0]$, $[1]...$, $[n-1]$, und die Quotientenmenge ist gegeben durch $\mathbb{Z}/n\mathbb{Z} = \{[m] | m \in \mathbb{Z}\} = \{[0], [1], ..., [n-1]\}$.

Man kann sie sich als einen aufgerollten Zahlenstrahl vorstellen, wie in Abbildung 1 gezeigt. Markiert man auf einer Linie die Zahlen in $\mathbb{Z}$ und rollt diese Linie dann so um einen Kreis auf, dass genau die Zahlen übereinander liegen, deren Differenz durch n teilbar ist, oder, dazu äquivalent, die bei Division durch n den gleichen Rest ergeben, so erhält man die Menge $\mathbb{Z}/n\mathbb{Z}$.

Wir können der Quotientenmenge $\mathbb{Z}/n\mathbb{Z}$ die Struktur einer abelschen Gruppe geben, indem wir die Addition aus $\mathbb{Z}$ benutzen. Wir definieren also $[k] + [m] = [k + m]$ für alle $k, m \in \mathbb{Z}$. Stellt man sich $\mathbb{Z}/12\mathbb{Z}$ als eine Uhr mit den n Ziffern $[0]$,...,$[11]$ vor, so entspricht diese Addition dem Weiterstellen der Zeigers. Steht der Zeiger auf der Ziffer $[3]$, also 3 Uhr, und stellt man ihn 14 Stunden vor, so zeigt er auf $[3 + 14] = [17] = [5]$, also 5 Uhr. Vorstellen um 14 Stunden ergibt also die gleiche Zeigerstellung wie Vorstellen um 2 Stunden. Stellt man den Zeiger um 5 Stunden zurück, so zeigt er auf $[3 + (-5)] = [-2] = [10]$ Uhr.

Lemma 2.2.10: Für jedes $n \in \mathbb{N}$ ist $+ : \mathbb{Z}/n\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$, $([k], [m]) \mapsto [k + m]$ eine kommutative und assoziative Verknüpfung auf $\mathbb{Z}/n\mathbb{Z}$ mit neutralem Element $[0]$, und $(\mathbb{Z}/n\mathbb{Z}, +)$ ist eine abelsche Gruppe.

Beweis:
1. Zu zeigen ist zunächst, dass $+ : \mathbb{Z}/n\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$, $([k], [m]) \mapsto [k + m]$ überhaupt eine Abbildung ist, also einem Paar $([k], [m])$ nur ein Element aus $\mathbb{Z}/n\mathbb{Z}$ zuordnet. Das zugeordnete Element $[k + m]$ darf also nicht von der Wahl der Repräsentanten k, m abhängen, und es ist zu zeigen, dass aus $[k] = [k']$ und $[m] = [m']$ auch $[k + m] = [k' + m']$ folgt.

Ist $[k] = [k']$ und $[m] = [m']$, so gilt $k \sim k'$ und $m \sim m'$ und somit gibt es $a, b \in \mathbb{Z}$ mit $k - k' = na$ und $m - m' = nb$. Daraus ergibt sich $k + m = k' + m' + n(a + b)$ und $[k + m] = [k' + m]$.

2. Die Assoziativität und Kommutativität folgen sofort aus den entsprechenden Eigenschaften der Verknüpfung $+ : \mathbb{Z} \times \mathbb{Z} \to \mathbb{Z}$:

$$\begin{aligned} ([k] + [l]) + [m] &= [k + l] + [m] = [k + l + m] = [k] + [l + m] = [k] + ([l] + [m]) \\ [k] + [m] &= [k + m] = [m + k] = [m] + [k] \qquad\qquad\qquad\qquad \forall k, l, m \in \mathbb{Z}. \end{aligned}$$

Ebenso zeigt man, dass $[0]$ ein neutrales Element und $[-m]$ das Inverse von $[m]$ ist:

$$[0] + [m] = [0 + m] = [m] \qquad [m] + [-m] = [m + (-m)] = [m - m] = [0] \quad \forall m \in \mathbb{Z}.$$ □

Weitere besonders wichtige Beispiele von Gruppen erhält man, wenn man die bijektiven Abbildungen $f : X \to X$ von einer Menge X in sich selbst betrachtet. Die Verkettung solcher Abbildungen definiert eine Verknüpfung auf der Menge der Abbildungen $f : X \to X$ mit der Identitätsabbildung als neutrales Element. Die Bijektivität garantiert die Existenz von Umkehrabbildungen, die Inversen für diese Verknüpfung.

Beispiel 2.2.11: Sei X eine Menge, $\mathrm{Abb}(X, X)$ die Menge der Abbildungen $f : X \to X$ und $\mathrm{Bij}(X, X)$ die Menge der *bijektiven* Abbildungen $f : X \to X$.

1. Dann ist $\mathrm{Abb}(X, X)$ mit der Verkettung von Abbildungen $\circ : \mathrm{Abb}(X, X) \times \mathrm{Abb}(X, X) \to \mathrm{Abb}(X, X)$, $(g, f) \mapsto g \circ f$ als Verknüpfung ein Monoid.

   Denn nach Beispiel 2.2.2 ist die Verkettung eine assoziative Verknüpfung auf M, und nach Satz 2.1.26, 1. gilt $f \circ \mathrm{id}_X = \mathrm{id}_X \circ f = f$ für alle Abbildungen $f : X \to X$. Die Identitätsabbildung $\mathrm{id}_X : X \to X$ ist also ein neutrales Element.

2. Die Menge $\mathrm{Bij}(X, X)$ mit der Verkettung $\circ : \mathrm{Bij}(X, X) \times \mathrm{Bij}(X, X) \to \mathrm{Bij}(X, X)$, $(g, f) \mapsto g \circ f$ als Verknüpfung ist eine Gruppe.

   Denn nach Satz 2.1.26 3. (iii) ist die Verkettung zweier bijektiver Abbildungen bijektiv. Damit ist die Verkettung von Abbildungen eine assoziative Verknüpfung auf der Menge der *bijektiven* Abbildungen. Da die id_X bijektiv ist, ist sie ein neutrales Element. Nach Satz 2.1.27, 3. existiert zu jeder bijektiven Abbildung $f : X \to X$ eine bijektive Umkehrabbildung $f^{-1} : X \to X$ mit $f \circ f^{-1} = f^{-1} \circ f = \mathrm{id}_X$, also ein Inverses.

Ist X eine endliche Menge, so nennt man eine bijektive Abbildung $f : X \to X$ auch eine **Permutation** von X und die Gruppe der bijektiven Abbildungen $f : X \to X$ mit der Verkettung eine **Permutationsgruppe**. Durch Umbenennung der Elemente können wir jede endliche Menge X mit $n \in \mathbb{N}$ Elementen in die Menge $\{1, \ldots, n\}$ überführen. Statt der Permutationen der Menge X können wir daher auch die Permutationen der Menge $\{1, \ldots, n\}$ betrachten.

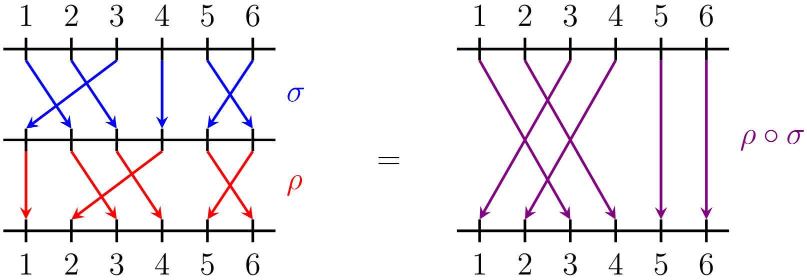
Definition 2.2.12: Sei $n \in \mathbb{N}$. Die **symmetrische Gruppe** S_n vom Grad n ist die Gruppe der bijektiven Abbildungen $\sigma : \{1, \ldots, n\} \to \{1, \ldots, n\}$ mit der Verkettung von Abbildungen als Verknüpfung.

Man kann zeigen, dass die symmetrische Gruppe S_n genau $n!$ Elemente enthält. Für $n \leq 2$ ist sie eine abelsche Gruppe, aber für $n \geq 3$ ist sie nicht abelsch (Übung).

Im Folgenden beschreiben wir Permutationen in S_n oft durch eine Wertetabelle

$$\sigma = \begin{pmatrix} 1 & 2 & \ldots & n \\ \sigma(1) & \sigma(2) & \ldots & \sigma(n) \end{pmatrix} \in S_n.$$

Wir können Permutationen $\sigma \in S_n$ auch durch ein Diagramm darstellen, das aus zwei parallelen horizontalen Linien besteht, die von links nach rechts mit den Zahlen $1, .., n$ markiert sind. Für jedes $x \in \{1, .., n\}$ wird ein gerader Pfeil von dem Punkt x auf der oberen Linie zum Punkt $\sigma(x) \in \{1, ..., n\}$ auf der unteren Linie gezeichnet. Die Verkettung von Permutationen entspricht dann dem vertikalen Aneinanderhängen von Diagrammen mit Begradigung der Pfeile.



$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 3 & 1 & 4 & 6 & 5 \end{pmatrix} \qquad \rho \circ \sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 4 & 1 & 2 & 5 & 6 \end{pmatrix}$$

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

In Beispiel 2.2.11 wurde aus einem Monoid eine Gruppe konstruiert, indem die Verknüpfung auf dem Monoid $\text{Abb}(X, X)$ auf die Elemente des Monoids eingeschränkt wurde, die ein Inverses besitzen - die bijektiven Abbildungen. Dies funktioniert analog für beliebige Monoide. Um dies zu beweisen, müssen wir uns genauer mit den Eigenschaften von invertierbaren Elementen in einem Monoid befassen, die auch als **Einheiten** des Monoids bezeichnet werden.

Lemma 2.2.13: (**Einheiten eines Monoids**)
Sei $(M, \circ)$ ein Monoid mit neutralem Element $e \in M$. Dann gilt:

1. Das neutrale Element e ist invertierbar mit Inversem $e^{-1} = e$.
2. Ist $a \in M$ invertierbar, dann ist auch a^{-1} invertierbar und $(a^{-1})^{-1} = a$.
3. Sind $a \in M$ und $b \in M$ invertierbar, dann ist auch $a \circ b$ invertierbar mit $(a \circ b)^{-1} = b^{-1} \circ a^{-1}$.
4. Ist $a \in M$ invertierbar, so gilt für alle $b, c \in M$:

$$a \circ b = a \circ c \Rightarrow b = c \qquad\qquad b \circ a = c \circ a \Rightarrow b = c.$$

Beweis:
1. Nach Definition 2.2.4, 1. gilt $e \circ e = e$. Damit ist e invertierbar mit Inversem $e^{-1} = e$.

2. Ist $a \in M$ invertierbar mit Inversem $a^{-1} \in M$, dann gilt $a \circ a^{-1} = e = a^{-1} \circ a$ nach Definition 2.2.4, 2. Mit Definition 2.2.4, 2. folgt, dass a^{-1} invertierbar ist mit Inversem $(a^{-1})^{-1} = a$.

3. Sind $a, b \in M$ invertierbar, dann gilt nach Definition 2.2.4, 2. $a \circ a^{-1} = e = a^{-1} \circ a$ und $b \circ b^{-1} = e = b^{-1} \circ b$. Daraus folgt unter Ausnutzung der Assoziativität

$$a \circ b \circ b^{-1} \circ a^{-1} \overset{(G3)}{=} a \circ e \circ a^{-1} \overset{(G2)}{=} a \circ a^{-1} \overset{(G3)}{=} e \overset{(G3)}{=} b^{-1} \circ b \overset{(G2)}{=} b^{-1} \circ e \circ b \overset{(G3)}{=} b^{-1} \circ a^{-1} \circ a \circ b.$$

Also ist nach Definition 2.2.4, 2. das Element $b^{-1} \circ a^{-1}$ invers zu $a \circ b$ und $a \circ b$ invertierbar.

4. Ist $a \in M$ invertierbar, dann gilt nach Definition 2.2.4, 2. $a^{-1} \circ a = a \circ a^{-1} = e$, und es folgt unter Ausnutzung der Assoziativität:

$$\begin{aligned} a \circ b = a \circ c \quad &\Rightarrow \quad b \overset{(G2)}{=} e \circ b \overset{(G3)}{=} a^{-1} \circ a \circ b = a^{-1} \circ a \circ c \overset{(G3)}{=} e \circ c \overset{(G2)}{=} c \\ b \circ a = c \circ a \quad &\Rightarrow \quad b \overset{(G2)}{=} b \circ e \overset{(G3)}{=} b \circ a \circ a^{-1} = c \circ a \circ a^{-1} \overset{(G3)}{=} c \circ e \overset{(G2)}{=} c. \end{aligned}$$

□

Korollar 2.2.14: In jedem Monoid $(M, \circ)$ ist die Menge $M^\times = \{a \in M \,|\, a \text{ invertierbar}\}$ mit der Verknüpfung $\circ : M^\times \times M^\times \to M^\times$ eine Gruppe. Sie heißt **Einheitengruppe** von $(M, \circ)$.

Beweis:
Sind $a, b \in M^\times$ so ist nach Lemma 2.2.13, 3. auch das Element $a \circ b \in M^\times$. Also ist $M^\times$ abgeschlossen unter $\circ$, und $\circ : M^\times \times M^\times \to M^\times$ ist eine Verknüpfung. Nach Lemma 2.2.13, 1. ist $e \in M^\times$, und somit besitzt $M^\times$ ein neutrales Element. Per Definition besitzt jedes Element $a \in M^\times$ ein Inverses $a^{-1} \in M$ und nach Lemma 2.2.13, 2. gilt $a^{-1} \in M^\times$. □

Beispiel 2.2.15:

1. Die Einheitengruppe des Monoids $(\mathbb{Z}, \cdot)$ ist $(\{1, -1\}, \cdot)$.
2. Die Einheitengruppen der Monoide $(\mathbb{Q}, \cdot)$ und $(\mathbb{R}, \cdot)$ sind $(\mathbb{Q} \setminus \{0\}, \cdot)$ und $(\mathbb{R} \setminus \{0\}, \cdot)$.
3. Ist $(M, \circ)$ bereits eine Gruppe, so gilt $M^\times = M$ und man erhält $(M^\times, \circ) = (M, \circ)$.

Bei der Untersuchung von Mengen wurde der Begriff einer Teilmenge eingeführt. Da Gruppen Mengen mit einer zusätzlichen Struktur sind - nämlich einer Verknüpfung, die die Gruppenaxiome (G1)-(G3) erfüllt - ist man bei der Betrachtung von Gruppen an Teilmengen interessiert, die mit der gegebenen Verknüpfung selbst wieder eine Gruppe bilden.

Definition 2.2.16: Sei $(G, \circ)$ eine Gruppe. Eine Teilmenge $H \subseteq G$ heißt **Untergruppe** von $(G, \circ)$, wenn

(i) H abgeschlossen unter der Verknüpfung $\circ$ ist: $a, b \in H \Rightarrow a \circ b \in H$.

(ii) H mit der Einschränkung der Verknüpfung $\circ$ eine Gruppe ist.

Bevor wir Beispiele von Untergruppen betrachten, verschaffen wir uns noch ein leicht handhabbares Kriterium, mit dem sich nachrechnen lässt, ob eine gegebene Teilmenge $H \subseteq G$ eine Untergruppe der Gruppe $(G, \circ)$ ist.

Lemma 2.2.17: Sei $(G, \circ)$ eine Gruppe. Dann ist eine Teilmenge $H \subseteq G$ eine Untergruppe genau dann, wenn sie die folgenden Bedingungen erfüllt:

(UG1) $H \neq \emptyset$,

(UG2) H ist abgeschlossen unter $\circ$: $a, b \in H \Rightarrow a \circ b \in H$,

(UG3) H ist abgeschlossen unter der Inversenbildung: $a \in H \Rightarrow a^{-1} \in H$.

Beweis:
Offensichtlich stimmt (UG2) mit Bedingung (i) in Definition 2.2.16 überein.

$\Leftarrow$: Sei $(G, \circ)$ eine Gruppe mit neutralem Element $e \in G$ und $H \subseteq G$ eine Teilmenge, die (UG1)-(UG3) erfüllt. Dann ist zu zeigen, dass H mit der Einschränkung von $\circ$ eine Gruppe ist. Nach (UG2) definiert $\circ$ einen assoziative Verknüpfung auf H (G1). Nach (UG1) gibt es ein Element $a \in H$. Nach (UG3) ist dann auch $a^{-1} \in H$, und mit (UG2) folgt $a \circ a^{-1} = e = a \circ a^{-1} \in H$. Also besitzt H das neutrale Element $e \in H$ (G2). Für jedes Element $a \in H$ ist nach (UG2) auch $a^{-1} \in H$ und damit ein Inverses zu a in H (G3). Damit ist $(H, \circ)$ eine Gruppe.

$\Rightarrow$: Sei $(G, \circ)$ eine Gruppe mit neutralem Element $e \in G$ und $H \subseteq G$ eine Untergruppe. Zu zeigen ist, dass (UG1) und (UG3) erfüllt sind. Nach Bedingung (ii) in Definition 2.2.16 ist $(H, \circ)$ eine Gruppe und enthält somit ein neutrales Element $e' \in H$. Daraus folgt $H \neq \emptyset$ (UG1).

Um zu zeigen, dass für jedes Element $a \in H$ auch das Inverse $a^{-1} \in G$ in H liegt, zeigt man zunächst, dass $e' = e$ gilt. Da $e' \in H$ das neutrale Element von H ist, gilt $e' \circ e' = e'$. Daraus folgt $e' \circ e' = e' \circ e$ in G, und da e' in G invertierbar ist, ergibt sich mit Lemma 2.2.13, 4. $e = e'$.

Da $(H, \circ)$ eine Gruppe mit neutralem Element $e' = e \in H$ ist, gibt es zu jedem Element $a \in H$ ein Inverses $a' \in H$ mit $a \circ a' = a' \circ a = e' = e$. Daraus folgt aber, dass a' auch das Inverse von a in G ist und wegen der Eindeutigkeit der Inversen in G damit $a' = a^{-1} \in H$, also (UG3). □

Beispiel 2.2.18:

1. Für jede Gruppe $(G, \circ)$ sind die Teilmengen $G \subseteq G$ und $\{e\} \subseteq G$ Untergruppen, denn sie sind nicht-leer, abgeschlossen unter $\circ$ und abgeschlossen unter Inversenbildung.

2. Ist $(G, \circ)$ eine Gruppe und sind $(H, \circ) \subseteq (G, \circ)$ und $(K, \circ) \subseteq (H, \circ)$ Untergruppen, so ist $(K, \circ)$ auch eine Untergruppe von $(G, \circ)$. Denn $K \subseteq H \subseteq G$ ist nicht-leer, abgeschlossen unter $\circ$ und abgeschlossen unter Inversenbildung.

3. Die Gruppen $(\mathbb{Z}, +) \subseteq (\mathbb{Q}, +) \subseteq (\mathbb{R}, +)$ und $(\mathbb{Q} \setminus \{0\}, \cdot) \subseteq (\mathbb{R} \setminus \{0\}, \cdot)$ sind jeweils Untergruppen.

   Es handelt sich offenbar um nicht-leere Teilmengen. Ausserdem sind für beliebige ganze (rationale) Zahlen x, y auch $x + y$ und $-x$ ganz (rational) und für beliebige $p, q \in \mathbb{Q} \setminus \{0\}$ sind auch $p \cdot q \in \mathbb{Q} \setminus \{0\}$ und $1/q \in \mathbb{Q} \setminus \{0\}$.

4. Für jede natürliche Zahl $n \in \mathbb{N}$ ist die Menge $n\mathbb{Z} = \{mn |\, m \in \mathbb{Z}\}$ der Vielfachen von n eine Untergruppe der Gruppe $(\mathbb{Z}, +)$.

   Denn es gilt $n\mathbb{Z} \neq \emptyset$. Sind $a, b \in n\mathbb{Z}$, so existieren $a', b' \in \mathbb{Z}$ mit $a = na'$, $b = nb'$ und somit gilt $a + b = na' + nb' = n(a' + b')$ mit $a' + b' \in \mathbb{Z}$ und $-a = -na' = n(-a')$ mit $-a' \in \mathbb{Z}$. Also ist $n\mathbb{Z}$ abgeschlossen unter $+$ und Inversenbildung.

Da Gruppen Mengen mit zusätzlicher Struktur sind, kann man natürlich auch Abbildungen zwischen Gruppen betrachten. Allerdings ist man weniger an allgemeinen Abbildungen interessiert als an *strukturerhaltenden Abbildungen*, also Abbildungen, die mit der Verknüpfung verträglich sind. Dies bedeutet, dass man das gleiche Ergebnis erhält, wenn man zuerst zwei Elemente verknüpft und dann das resultierende Element abbildet und wenn man die zwei Elemente zuerst abbildet und dann ihre Bilder verknüpft.

Definition 2.2.19: Seien $(G, \cdot_G)$ und $(H, \cdot_H)$ Gruppen. Eine Abbildung $\phi : G \to H$ heißt **Gruppenhomomorphismus** oder **Homomorphismus von Gruppen**, falls gilt

$$\phi(a \cdot_G b) = \phi(a) \cdot_H \phi(b) \qquad \forall a, b \in G$$

Ist die Abbildung ϕ außerdem bijektiv, so spricht man von einem **Gruppenisomorphismus**. Gibt es einen Gruppenisomorphismus $\phi : G \to H$, so nennt man die Gruppen $(G, \cdot_G)$ und $(H, \cdot_H)$ **isomorph** und schreibt $G \cong H$.

Im Fall $(G, \cdot_G) = (H, \cdot_H)$ spricht man statt einem Gruppenhomomorphismus auch von einem **Gruppenendomorphismus** und statt von einem Gruppenisomorphismus auch von einem **Gruppenautomorphismus**.

Beispiel 2.2.20:

1. Für jedes $g \in G$ ist die **Konjugationsabbildung** $C_g : G \to G$, $h \mapsto g \cdot h \cdot g^{-1}$ ein Gruppenautomorphismus. Denn es gilt für alle $a, b \in G$

$$C_g(a \cdot b) = g \cdot (a \cdot b) \cdot g^{-1} = g \cdot a \cdot e \cdot b \cdot g^{-1} = (g \cdot a \cdot g^{-1}) \cdot (g \cdot b \cdot g^{-1}) = C_g(a) \cdot C_g(b)$$
$$C_{g^{-1}} \circ C_g(a) = g^{-1} \cdot (g \cdot a \cdot g^{-1}) \cdot g = a = g \cdot (g^{-1} \cdot a \cdot g) \cdot g^{-1} = C_g \circ C_{g^{-1}}(a).$$

   Somit ist $C_g : G \to G$ ein Gruppenhomomorphismus mit Umkehrabbildung $C_{g^{-1}} : G \to G$ und damit ein Gruppenisomorphismus.

2. Für jedes $m \in \mathbb{Z}$ ist die Abbildung $\mu_m : \mathbb{Z} \to \mathbb{Z}$, $z \mapsto mz$ ein Gruppenendomorphismus von $(\mathbb{Z}, +)$. Denn es gilt $\mu_m(z) + \mu_m(z') = mz + mz' = m(z + z') = \mu_m(z + z')$ für alle $z, z' \in \mathbb{Z}$. Sie ist nur für $m \in \{1, -1\}$ ein Gruppenautomorphismus, denn ansonsten ist sie nicht surjektiv.

3. Für jedes $n \in \mathbb{N}$ ist die kanonische Surjektion $\pi : (\mathbb{Z}, +) \to (\mathbb{Z}/n\mathbb{Z}, +)$, $m \mapsto [m]$ ein Gruppenhomomorphismus, denn $\pi(k + m) = [k + m] = [k] + [m] = \pi(k) + \pi(m)$ für alle $k, m \in \mathbb{Z}$. Sie ist kein Gruppenisomorphismus, denn sie ist nicht injektiv: $\pi(n) = [n] = [0] = \pi(0)$.

4. Für jede Gruppe $(G, \cdot)$ und jedes $g \in G$ ist $\nu_g : (\mathbb{Z}, +) \to (G, \cdot)$, $m \mapsto g^m$ mit

$$g^m := \underbrace{g \cdot g \cdots g}_{m\times} \text{ für } m > 0, \qquad g^0 := e, \qquad g^m := \underbrace{g^{-1} \cdot g^{-1} \cdots g^{-1}}_{(-m)\times} \text{ für } m < 0$$

   ein Gruppenhomomorphismus. Denn es gilt $\nu_g(n + m) = g^{n+m} = g^n \cdot g^m = \nu_g(n) \cdot \nu_g(m)$ für alle $n, m \in \mathbb{Z}$.

5. Für beliebige Gruppen $(G, \cdot_G)$ und $(H, \cdot_H)$ gibt es immer mindestens einen Gruppenhomomorphismus $\phi : G \to H$, nämlich $\phi : G \to H$, $g \mapsto e_H$ für alle $g \in G$.

Da jede Gruppe ein ausgezeichnetes Element enthält - das neutrale Element - ist es naheliegend zu untersuchen, wie sich Gruppenhomomorphismen auf dieses ausgezeichnete Element auswirken und welche Eigenschaften das Urbild des neutralen Elements besitzt. Ebenso stellt sich die Frage, wie sich Gruppenhomomorphismen in Hinblick auf zueinander inverse Elemente verhalten und was sich über das Bild einer Gruppe unter einem Gruppenhomomorphismus aussagen lässt. Die Antwort auf diese Fragen liefert der folgende Satz.

Satz 2.2.21: Seien $(G, \cdot_G)$ und $(H, \cdot_H)$ Gruppen und $\phi : G \to H$ ein Gruppenhomomorphismus. Dann gilt:

1. ϕ bildet das neutrale Element auf das neutrale Element ab: $\phi(e_G) = e_H$.
2. ϕ bildet Inverse auf Inverse ab: $\phi(g^{-1}) = \phi(g)^{-1}$ für alle $g \in G$.
3. Der **Kern** $\ker(\phi) := \phi^{-1}(e_H) = \{g \in G | \phi(g) = e_H\}$ ist eine Untergruppe von $(G, \cdot_G)$.
4. Das **Bild** $\phi(G) := \{h \in H | \exists g \in G : \phi(g) = h\}$ ist eine Untergruppe von $(H, \cdot_H)$.

Beweis:
1. Für das neutrale Element $e_G \in G$ gilt $e_H \cdot_H \phi(e_G) = \phi(e_G) = \phi(e_G \cdot_G e_G) = \phi(e_G) \cdot_H \phi(e_G)$. Mit Lemma 2.2.13, 4. folgt daraus $\phi(e_G) = e_H$.

2. Für alle $g \in G$ ergibt sich mit 1.

$$\phi(g^{-1}) \cdot_H \phi(g) = \phi(g^{-1} \cdot_G g) = \phi(e_G) = e_H = \phi(e_G) = \phi(g \cdot_G g^{-1}) = \phi(g) \cdot_H \phi(g^{-1})$$

und somit ist $\phi(g^{-1}) = \phi(g)^{-1}$ für alle $g \in G$.

3. Nach 1. gilt $e_G \in \ker(\phi)$, also $\ker(\phi) \neq \emptyset$ (UG1). Sind $g, g' \in \ker(\phi)$, so gilt $\phi(g) = \phi(g') = e_H$. Daraus folgt $\phi(g \cdot_G g') = \phi(g) \cdot_H \phi(g') = e_H \cdot_H e_H = e_H$, also auch $g \cdot_G g' \in \ker(\phi)$. Also ist $\ker(\phi)$ abgeschlossen unter $\cdot_G$ (UG2). Ebenso gilt für jedes $g \in \ker(\phi)$ nach 2. $\phi(g^{-1}) = \phi(g)^{-1} = e_H^{-1} = e_H$, also auch $g^{-1} \in \ker(\phi)$. Damit ist $\ker(\phi)$ auch abgeschlossen unter der Inversenbildung (UG3) und somit eine Untergruppe von G.

4. Nach 1. ist $e_H = \phi(e_G) \in \phi(G)$ und somit $\phi(G) \neq \emptyset$ (UG1). Sind $h, h' \in \phi(G)$, so existieren $g, g' \in G$ mit $h = \phi(g)$ und $h' = \phi(g')$. Daraus folgt $h \cdot_H h' = \phi(g) \cdot_H \phi(g') = \phi(g \cdot_G g')$, also auch $h \cdot_H h' \in \phi(G)$ (UG2). Ebenso folgt aus $h = \phi(g)$ auch $h^{-1} = \phi(g)^{-1} = \phi(g^{-1})$ und somit $h^{-1} \in \phi(G)$ für alle $h \in \phi(G)$ (UG3). Also ist $\phi(G)$ eine Untergruppe von H. □

Beispiel 2.2.22:

1. Ein Gruppenhomomorphismus $\phi : G \to H$ ist injektiv genau dann, wenn $\ker(\phi) = \{e_G\}$. Denn für beliebige $g, g' \in G$ gilt:
   $\phi(g) = \phi(g') \Leftrightarrow e_H = \phi(g) \cdot_H \phi(g')^{-1} = \phi(g \cdot_G g'^{-1}) \Leftrightarrow g \cdot_G g'^{-1} \in \ker(\phi)$.

2. Für die Gruppenhomomorphismen $\mu_n : \mathbb{Z} \to \mathbb{Z}$, $k \mapsto nk$ und $\pi : \mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$, $k \mapsto [k]$ aus Beispiel 2.2.20, 2. und 3. ergibt sich $\ker(\pi) = n\mathbb{Z} = \mu_n(\mathbb{Z})$, denn es gilt:
   $k \in n\mathbb{Z} = \mu_n(\mathbb{Z}) \Leftrightarrow \exists a \in \mathbb{Z} : k = na \Leftrightarrow \pi(k) = [k] = [0]$.

3. Jeder Gruppenendomorphismus von $(\mathbb{Z}, +)$ ist von der Form $\phi = \mu_n : \mathbb{Z} \to \mathbb{Z}, z \mapsto n \cdot z$ für ein $n \in \mathbb{Z}$. Denn ist $\phi : \mathbb{Z} \to \mathbb{Z}$ ein Gruppenendomorphismus von $(\mathbb{Z}, +)$, so gilt $\phi(0) = 0$ nach Satz 2.2.21, 1. Für alle $k \in \mathbb{N}$ erhält man induktiv $\phi(k) = \phi(1 + ... + 1) = \phi(1) + ... + \phi(1) = k \cdot \phi(1)$ und mit Satz 2.2.21, 2. $\phi(-k) = -\phi(k) = -k \cdot \phi(1)$. Also gilt $\phi(k) = \phi(1) \cdot k$ für alle $k \in \mathbb{Z}$ und somit $\phi = \mu_n$ mit $n = \phi(1) \in \mathbb{Z}$.

Wir untersuchen nun, wie sich Gruppenhomomorphismen unter Verkettung und Inversenbildung verhalten. Diese ist unter anderem aus dem folgenden Grund relevant. Ähnlich wie man Mengen, die zueinander in Bijektion stehen, nicht mit mengentheoretischen Mitteln unterscheiden kann, so kann man isomorphe Gruppen nicht mit gruppentheoretischen Mitteln unterscheiden. Sie sind aus gruppentheoretischer Sicht gleich bzw. gehen durch eine bloße Umbenennung von Elementen auseinander hervor.

Dies suggeriert, dass es sich bei der Isomorphie von Gruppen um eine Äquivalenzrelation[7] handelt. Um zu zeigen, dass Isomorphie von Gruppen tatsächlich (i) reflexiv, (ii) symmetrisch und (iii) transitiv ist, benötigt man, dass (i) die Identitätsabbildungen Gruppenisomorphismen sind, dass (ii) für jeden Gruppenisomorphismus auch die inverse Abbildung ein Gruppenisomorphismus ist und dass (iii) die Verkettung zweier Gruppenisomorphismen ein Gruppenisomorphismus ist. Diese Aussagen ergeben sich aus aus dem folgenden Satz und der Tatsache, dass die Verkettung bijektiver Abbildungen bijektiv ist.

Satz 2.2.23:

1. Für jede Gruppe $(G, \cdot_G)$ ist $\mathrm{id}_G : G \to G$ ein Gruppenautomorphismus.
2. Sind $(G, \cdot_G)$ und $(H, \cdot_H)$ Gruppen und $\phi : G \to H$ ein Gruppenisomorphismus, dann ist auch die Umkehrabbildung $\phi^{-1} : H \to G$ ein Gruppenisomorphismus.

[7]Achtung: die Gruppen bilden keine Menge! Somit kann es sich also nicht um eine Äquivalenzrelation auf einer Menge handeln. Man kann aber den Begriff einer Äquivalenzrelation aber auch allgemeiner für andere mathematische Strukturen, die sogenannten Klassen, definieren.

3. Sind $(G, \cdot_G)$, $(H, \cdot_H)$ und $(K, \cdot_K)$ Gruppen und $\phi : G \to H$, $\psi : H \to K$ Gruppenhomomorphismen, so ist auch die Verkettung $\psi \circ \phi : G \to K$, $g \mapsto \psi(\phi(g))$ ein Gruppenhomomorphismus.
4. Die Gruppenautomorphismen $\phi : G \to G$ bilden mit der Verkettung eine Gruppe $\mathrm{Aut}(G)$.

Beweis:
1. Die Abbildung id_G ist bijektiv und für alle $g, g' \in G$ gilt $\mathrm{id}_G(g \cdot_G g') = g \cdot_G g' = \mathrm{id}_G(g) \cdot_G \mathrm{id}_G(g')$.

2. Da $\phi : G \to H$ ein Gruppenhomomorphimus ist, erhält man für alle $h, h' \in H$

$$\phi^{-1}(h) \cdot_G \phi^{-1}(h') = \phi^{-1}(\phi(\phi^{-1}(h) \cdot_G \phi^{-1}(h'))) = \phi^{-1}(\phi(\phi^{-1}(h)) \cdot_H \phi(\phi^{-1}(h'))) = \phi^{-1}(h \cdot_H h').$$

3. Da $\phi : G \to H$ und $\psi : H \to K$ Gruppenhomomorphismen sind, erhält man für alle $g, g' \in G$

$$(\psi \circ \phi)(g \cdot_G g') = \psi(\phi(g \cdot_G g')) = \psi(\phi(g) \cdot_H \phi(g')) = \psi(\phi(g)) \cdot_K \psi(\phi(g')).$$

4. Nach 3. ist die Verkettung von Gruppenautomorphismen ein Gruppenendomorphismus und nach Satz 2.1.26 bijektiv als Verkettung bijektiver Abbildungen, also ein Gruppenautomorphismus. Die Verkettung von Gruppenhomomorphismen ist nach Satz 2.1.26 assoziativ mit der Identitätsabbildung als neutrales Element und der Umkehrabbildung als Inverses. □

Mit Satz 2.2.23 wird es sinnvoll, isomorphe Gruppen als (gruppentheoretisch) äquivalent zu betrachten und sie nicht zu unterscheiden. Von diesem Standpunkt aus kennt man alle möglichen Gruppen, wenn man eine Liste aufstellen kann, so dass jede Gruppe zu genau einer Gruppe dieser Liste isomorph ist. Die Suche nach einer solchen Liste bezeichnet man als *Klassifikationsproblem* oder *Klassifikation* von Gruppen.

Während das Klassifikationsproblem für unendliche Gruppen ungelöst ist, ist es für *endliche* Gruppen seit 2002 im Wesentlichen gelöst. Man kann nämlich jede endliche Gruppe aus sogenannten *einfachen Gruppen* aufbauen, die bis auf Isomorphie vollständig klassifiziert sind. Die zugehörigen Beweise verteilen sich auf über 500 Publikationen mit insgesamt fast 15 000 Seiten, die von 1920 bis 2002 verfasst wurden.

2.3 Ringe und Körper

Offensichtlich verallgemeinert der Begriff der abelschen Gruppe das Konzept der *Addition* und liefert für die Gruppen $(\mathbb{Z}, +)$, $(\mathbb{Q}, +)$ und $(\mathbb{R}, +)$ genau die aus der Schule bekannte Addition ganzer, rationaler und reeller Zahlen. Neben der Addition gibt es dort noch eine Multiplikation.

Im Folgenden befassen wir uns daher mit Mengen, die mit zwei Verknüpfungen ausgestattet sind, die die Addition und Multiplikation verallgemeinern. Dabei fordern wir, dass sie auf ähnliche Art interagieren, wie die Addition und Multiplikation von ganzen, rationalen und reellen Zahlen, also ein Distributivgesetz erfüllen. Fordern wir für die Multiplikation zunächst nicht, dass sie kommutativ ist oder dass Inverse existieren, so erhalten wir das Konzept eines Rings.

Definition 2.3.1: Eine Menge R zusammen mit zwei Verknüpfungen $+ : R \times R \to R$, $(a, b) \mapsto a + b$ und $\cdot : R \times R \to R$, $(a, b) \mapsto a \cdot b$ heißt **Ring**, wenn gilt:

(R1) $(R, +)$ ist eine abelsche Gruppe,

(R2) $(R, \cdot)$ ist eine Halbgruppe,
(R3) **Distributivgesetz**:

$$a \cdot (b + c) = (a \cdot b) + (a \cdot c) \qquad (b + c) \cdot a = (b \cdot a) + (c \cdot a) \qquad \forall a, b, c \in R.$$

$(R, +, \cdot)$ heißt **Ring mit Eins** oder **unitaler Ring**, wenn zusätzlich gilt:

(UR) $(R, \cdot)$ ist ein Monoid.

Das neutrale Element der Gruppe $(R, +)$ wird mit 0 bezeichnet, das Inverse von a in $(R, +)$ mit $-a$, das neutrale Element von $(R, \cdot)$ mit 1 und Inverse von a bezüglich $\cdot$ mit $^{-1}$. Ist die Verknüpfung $\cdot$ kommutativ, so spricht man von einem **kommutativen Ring**.

Bemerkung 2.3.2:

1. In Ringen wählt man immer die Konvention *Punkt vor Strich*:
$$a \cdot b + c := (a \cdot b) + c \qquad a + b \cdot c := a + (b \cdot c) \qquad \forall a, b, c \in R.$$
2. Wie auch endlichen Gruppe kann man endliche Ringe durch Multiplikationstafeln beschreiben. Dabei benötigt man zwei Multiplikationstafeln, je eine für jede Verknüpfung.

Bevor wir Beispiele von Ringen betrachten, leiten wir zunächst noch einige nützliche Rechenregeln her, die in jedem Ring gelten und die aus der Schule bekannten Rechenregeln für ganze, rationale und reelle Zahlen verallgemeinern.

Lemma 2.3.3: (**Rechenregeln für Ringe**) In jedem Ring $(R, +, \cdot)$ gelten die Rechenregeln

$$0 \cdot a = a \cdot 0 = 0, \qquad -(a \cdot b) = (-a) \cdot b = a \cdot (-b), \qquad a \cdot b = (-a) \cdot (-b) \qquad \forall a, b \in R.$$

Beweis:
Es gilt $a \cdot 0 = a \cdot (0+0) = a \cdot 0 + a \cdot 0$, und mit Lemma 2.2.13, 4. folgt $a \cdot 0 = 0$. Daraus ergibt sich mit dem Distributivgesetz $a \cdot b + (-a) \cdot b = (a + (-a)) \cdot b = 0 \cdot b = 0 = a \cdot 0 = a \cdot ((-b) + b) = a \cdot (-b) + a \cdot b$. Also gilt $(-a) \cdot b = a \cdot (-b) = -a \cdot b$ und damit auch $a \cdot b = -(-(a \cdot b)) = -((-a) \cdot b) = (-a) \cdot (-b)$. □

Beispiel 2.3.4:

1. Ist $(R, +, \cdot)$ ein unitaler Ring mit $1 = 0$, so folgt aus den Rechenregeln $a = a \cdot 1 = a \cdot 0 = 0$ für alle $a \in R$, also $R = \{0\}$. Dieser Ring heißt **Nullring**.
2. $(\mathbb{Z}, +, \cdot)$, $(\mathbb{Q}, +, \cdot)$, $(\mathbb{R}, +, \cdot)$ sind kommutative unitale Ringe.
3. Für $m \in \mathbb{N}$ ist $(m\mathbb{Z}, +, \cdot)$ ein kommutativer Ring, besitzt aber kein Einselement für $m \neq 1$.
4. Für jede Menge $M \neq \emptyset$ und jeden Ring $(R, +, \cdot)$ ist die Menge $\mathrm{Abb}(M, R)$ der Abbildungen $f : M \to R$ mit den Verknüpfungen $+, \cdot : \mathrm{Abb}(M, R) \times \mathrm{Abb}(M, R) \to \mathrm{Abb}(M, R)$
$$(f + g)(m) := f(m) + g(m) \qquad (f \cdot g)(m) := f(m) \cdot g(m) \qquad \forall m \in M$$
ein Ring mit der Nullabbildung $0 : M \to R$, $m \mapsto 0$ als neutrales Element. Er ist kommutativ oder unital genau dann, wenn $(R, +, \cdot)$ kommutativ oder unital ist. Für $M = \emptyset$ ist $\mathrm{Abb}(M, R)$ der Nullring. (Beweis: Übung).

5. Aus 2. und 4. ergibt sich: Für jede Menge M sind $\mathrm{Abb}(M, \mathbb{Z})$, $\mathrm{Abb}(M, \mathbb{Q})$ und $\mathrm{Abb}(M, \mathbb{R})$ mit $+$ und $\cdot$ wie in 4. unitale kommutative Ringe.

6. Sind $(R, +, \cdot)$ und $(S, +', \cdot')$ (unitale) Ringe, so hat auch die Menge $R \times S$ die Struktur eines (unitalen) Rings mit den Verknüpfungen $+, \cdot : (R \times S) \times (R \times S) \to R \times S$
$$(r, s) + (r', s') := (r + r', s +' s') \qquad (r, s) \cdot (r', s') := (r \cdot r', s \cdot' s').$$
Sie heißt **direktes Produkt** der (unitalen) Ringe $(R, +, \cdot)$, $(S, +', \cdot')$ (Übung).

7. Ist G eine abelsche Gruppe, so bilden die Gruppenendomorphismen $\phi : G \to G$ mit der Verkettung und der punktweisen Addition einen Ring mit Einselement id_G (Übung).

Ein weiteres Beispiel eines kommutativen unitalen Rings ist der **Polynomring**. Reelle Polynome werden in der Schule üblicherweise als *Polynomabbildungen* eingeführt:
$$p : \mathbb{R} \to \mathbb{R}, \quad x \mapsto a_n x^n + ... + a_1 x + a_0 \qquad \text{mit} \quad a_0, ..., a_n \in \mathbb{R}.$$
In der Algebra fasst man ein Polynom allerdings als eine Folge auf, nämlich die Folge seiner Koeffizienten. Dies entspricht einer Abbildung $p : \mathbb{N}_0 \to \mathbb{R}$, $k \mapsto a_k$, die der Zahl $k \in \mathbb{N}_0$ den vor dem Monom x^k auftretenden Koeffizienten a_k zuordnet. Da ein Polynom $a_n x^n + ... + a_1 x + a_0$ nur *endlich viele* Monome x^k mit nichtverschwindenden Koeffizienten a_k enthält, muss man fordern, dass $p(k) = a_k = 0$ für *fast alle*, d. h. alle bis auf endlich viele $k \in \mathbb{N}_0$ gilt.

Da ein Polynom durch die Koeffizienten eindeutig festgelegt ist, können die Addition und Multiplikation von Polynomen anhand der Koeffizienten beschrieben werden. Für Polynome $p, q : \mathbb{N}_0 \to \mathbb{R}$ definiert man $p + q : \mathbb{N}_0 \to \mathbb{R}$ und $p \cdot q : \mathbb{N}_0 \to \mathbb{R}$ durch
$$(p + q)(k) = p(k) + q(k) \qquad (p \cdot q)(k) = \Sigma_{j=0}^{k}\, p(j) q(k - j) \qquad \forall k \in \mathbb{N}_0. \tag{3}$$
Schreibt man dann $p = \Sigma_{k \in \mathbb{N}_0} a_k x^k$ und $q = \Sigma_{j \in \mathbb{N}_0} b_j x^j$ mit $a_k := p(k)$, $b_j := q(j)$ so erhält man die bekannten Formeln für die Addition und Multiplikation von Polynomen
$$p + q = \Sigma_{k \in \mathbb{N}_0} (a_k + b_k) x^k \qquad p \cdot q = \Sigma_{k \in \mathbb{N}_0} \left(\Sigma_{j=0}^{k} a_j b_{k-j} \right) x^k,$$
wobei sich die zweite Formel mit dem Distributivgesetz durch Ausmultiplizieren von $p \cdot q$ ergibt:

$$
\begin{array}{cccccccccccc}
 & & a_0x^0 & + & a_1x^1 & + & a_2x^2 & + & a_3x^3 & + & \cdots & \\
b_0x^0 & & a_0b_0x^0 & + & a_1b_0x^1 & + & a_2b_0x^2 & + & a_3b_0x^3 & + & \cdots & \xrightarrow{p} \\
+ & & & & & & & & & & & \\
b_1x^1 & + & a_0b_1x^1 & + & a_1b_1x^2 & + & a_2b_1x^3 & + & a_3b_1x^4 & + & \cdots & \\
+ & & & & & & & & & & & \\
b_2x^2 & + & a_0b_2x^2 & + & a_1b_2x^3 & + & a_2b_2x^4 & + & a_3b_2x^5 & + & \cdots & \\
+ & & & & & & & & & & & \\
b_3x^3 & + & a_0b_3x^3 & + & a_1b_3x^4 & + & a_2b_3x^5 & + & a_3b_3x^6 & + & \cdots & \\
+ & & & & & & & & & & & \\
\vdots & q\big\downarrow & \vdots & & \vdots & & \vdots & & \vdots & & &
\end{array}
$$

Die Formeln in (3) können für Abbildungen $f : \mathbb{N}_0 \to R$ in einen beliebigen kommutativen unitalen Ring $(R, +, \cdot)$ betrachtet werden und geben der Menge solcher Abbildungen mit $f(n) = 0$ für fast alle $n \in \mathbb{N}_0$ die Struktur eines kommutativen unitalen Rings.

Beispiel 2.3.5: Ein **Polynom** mit Koeffizienten in einem kommutativen unitalen Ring $(R, +, \cdot)$ ist eine Abbildung $p : \mathbb{N}_0 \to R$ mit $p(n) = 0$ für **fast alle**, d. h. alle bis auf endliche viele $n \in \mathbb{N}_0$. Die Menge $R[x]$ der Polynome mit Koeffizienten in R bildet mit den Verknüpfungen

$$(p+q)(n) := p(n) + q(n)$$
$$(p \cdot q)(n) := \Sigma_{k=0}^{n}\, p(k) \cdot q(n-k) = p(0) \cdot q(n) + p(1) \cdot q(n-1) + ... + p(n-1) \cdot q(1) + p(n) \cdot q(0)$$

einen kommutativen unitalen Ring. Das Nullpolynom $0 : \mathbb{N}_0 \to R$, $n \mapsto 0$ für alle $n \in \mathbb{N}_0$ ist das neutrale Element für die Addition und das Polynom $1 : \mathbb{N}_0 \mapsto R$ mit $0 \mapsto 1$ und $n \mapsto 0$ für alle $n \in \mathbb{N}$ ist das neutrale Element für die Multiplikation. (Beweis: Übung)

- Der **Grad** eines Polynoms $p : \mathbb{N}_0 \to R$ ist $\deg(p) = \max\{n \in \mathbb{N}_0 \mid p(n) \neq 0\}$ für $p \neq 0$ und $\deg(0) = -\infty$. Der Koeffizient $l(p) := p(\deg(p)) \in R$ heißt **Leitkoeffizient** von p.
- Ein Polynom $p : \mathbb{N}_0 \to R$ heißt **normiert**, wenn es den Leitkoeffizienten $l(p) = 1$ hat.
- Das Polynom $p : \mathbb{N}_0 \to R$ mit $p(j) = 0$ für $j \neq k$ und $p(k) = 1$ heißt **Monom** vom Grad k und wird mit $p = x^k$ bezeichnet.

Weitere wichtige Beispiele kommutativer unitaler Ringe sind *Restklassenringe*. Dabei handelt es sich um die Menge $\mathbb{Z}/n\mathbb{Z}$ für $n \in \mathbb{N}$ aus Beispiel 2.1.14 und Lemma 2.2.10, die durch die Addition und Multiplikation im Ring $(\mathbb{Z}, +, \cdot)$ mit der Struktur eines kommutativen unitalen Rings versehen wird. Stellen wir uns $\mathbb{Z}/n\mathbb{Z}$ als einen aufgerollten Zahlenstrahl vor, so wird die Addition und Multiplikation auf dem aufgerollten Zahlenstrahl einfach über die Addition und Multiplikation auf dem Zahlenstrahl definiert, wobei durch das anschließende Aufrollen dann genau die Zahlen zusammenfallen, deren Differenz durch n teilbar ist.

Lemma 2.3.6: Für jedes $n \in \mathbb{N}$ ist $(\mathbb{Z}/n\mathbb{Z}, +, \cdot)$ mit $[k] + [m] := [k+m]$ und $[k] \cdot [m] := [km]$ ein kommutativer Ring mit Einselement $[1]$. Er wird als **Restklassenring** bezeichnet.

Beweis:
In Lemma 2.2.10 wurde bereits gezeigt, dass $(\mathbb{Z}/n\mathbb{Z}, +)$ eine abelsche Gruppe ist (R1). Die Multiplikation $\cdot : \mathbb{Z}/n\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$ ist wohldefiniert. Denn ist $[k] = [k']$ und $[m] = [m']$ für $k, k', m, m' \in \mathbb{Z}$, so gibt es $a, b \in \mathbb{Z}$ mit $k' = k + an$, $m' = m + bn$, und mit dem Distributibgesetz in $\mathbb{Z}$ folgt $k'm' = km + n(am + bk + nab)$, also $[km] = [k'm']$. Dass sie assoziativ (R2) und kommutativ ist, folgt direkt aus den entsprechenden Eigenschaften der Multiplikation in $\mathbb{Z}$:

$$([k] \cdot [l]) \cdot [m] = [kl] \cdot [m] = [klm] = [k] \cdot [lm] = [k] \cdot ([l] \cdot [m]), \qquad [k] \cdot [l] = [kl] = [lk] = [l] \cdot [k]$$

Für alle $k, l, m \in \mathbb{Z}$. Ebenso ergibt sich das Distributivgesetz (R3)

$$[k] \cdot ([l] + [m]) = [k] \cdot [l+m] = [k(l+m)] = [kl + km] = [kl] + [km] = [k] \cdot [l] + [k] \cdot [l]$$

und dass $[1]$ das Einselement für die Multiplikation ist: $[1] \cdot [m] = [1 \cdot m] = [m]\ \forall m \in \mathbb{Z}$. $\square$

Im Fall von Gruppen folgte auf die Definition der Gruppe die Definition einer Untergruppe als Teilmenge einer Gruppe, die mit der Einschränkung der Verknüpfung selbst wieder eine Gruppe bildet und die Definition eines Gruppenhomomorphismus als Abbildung zwischen Gruppen, die mit der Verknüpfung veträglich ist. Dieses Prinzip lässt sich auch auf Ringe übertragen. So ist ein Unterring eines Rings eine Teilmenge, die mit den Einschränkungen der zwei Verknüpfungen wieder einen Ring bildet, und ein Ringhomomorphismus eine Abbildung zwischen Ringen, die mit beiden Verknüpfungen verträglich ist.

Allerdings muss man sich dabei entscheiden, ob man nur unitale Ringe betrachten möchte oder auch Ringe ohne Einselement. Im ersten Fall ist es naheliegend, auch entsprechende Bedingungen für das Einselement aufzuerlegen, also zu fordern, dass das Einselement im Unterring erhalten ist und von einem Ringhomomorphismus auf ein Einselement abgebildet wird. Im zweiten Fall ergeben diese Forderungen keinen Sinn.

Definition 2.3.7: Sei $(R, +, \cdot)$ ein Ring.

1. Eine Teilmenge $U \subseteq R$ heißt **Teilring** oder **Unterring** von R, wenn U abgeschlossen unter den Verknüpfungen $+$ und $\cdot$ ist und mit den Einschränkungen dieser Verknüpfungen einen Ring bildet:

   (UR1) $(U, +) \subseteq (R, +)$ ist eine Untergruppe.

   (UR2) $a, b \in U \Rightarrow a \cdot b \in U$.

2. Ist $(R, +, \cdot)$ ein unitaler Ring, so heisst $U \subseteq R$ **unitaler Unterring** oder **Unterring mit Eins**, wenn U ein Unterring von R ist und $1_U = 1_R$ gilt. Das ist der Fall genau dann, wenn $1_R \in U$.

Definition 2.3.8: Seien $(R, +, \cdot)$ und $(S, +', \cdot')$ Ringe.

1. Eine Abbildung $\phi : R \to S$ heißt **Ringhomomorphismus**, wenn gilt

$$\phi(r + r') = \phi(r) +' \phi(r'), \quad \phi(r \cdot r') = \phi(r) \cdot' \phi(r') \qquad \forall r, r' \in R.$$

2. Sind $(R, +, \cdot)$ und $(S, +', \cdot')$ unitale Ringe, so heißt $\phi : R \to S$ **unitaler Ringhomomorphismus** oder **Homomorphismus von unitalen Ringen**, wenn ϕ ein Ringhomomorphismus ist und $\phi(1_R) = 1_S$ gilt.

3. Ein bijektiver (unitaler) Ringhomomorphismus heißt **(unitaler) Ringisomorphismus** oder **Isomorphismus von (unitalen) Ringen**. Existiert ein (unitaler) Ringisomorphismus $\phi : R \to S$ so nennt man die (unitalen) Ringe R, S isomorph und schreibt $R \cong S$.

4. Ist $(R, +, \cdot) = (S, +', \cdot')$ so spricht man statt von einem (unitalen) Ringhomomorphismus auch von einem (unitalen) **Ringendomorphismus** und statt einem (unitalen) Ringisomorphismus auch von einem (unitalen) **Ringautomorphismus**.

Beispiel 2.3.9:

1. $(\mathbb{Z}, +, \cdot) \subseteq (\mathbb{Q}, +, \cdot) \subseteq (\mathbb{R}, +, \cdot)$ sind unitale Unterringe.

2. Für $m \in \mathbb{N}$ ist $(m\mathbb{Z}, +, \cdot) \subseteq (\mathbb{Z}, +, \cdot)$ ein Unterring des unitalen Rings $(\mathbb{Z}, +, \cdot)$, aber für $m \neq 1$ kein unitaler Unterring, denn er besitzt kein Einselement.

3. Die Teilmenge $\{[0], [3]\} \subseteq \mathbb{Z}/6\mathbb{Z}$ ist ein Unterring. Sie ist ein unitaler Ring mit Nullelement $[0]$ und Einselement $[3]$, denn $[3] \cdot [3] = [9] = [3]$. Sie ist aber *kein unitaler Unterring*, denn das Einselement $[3]$ stimmt nicht mit dem Einselement von $\mathbb{Z}/6\mathbb{Z}$ überein.

4. Für jede Menge M und jeden (unitalen) Ring R ist die Menge der Abbildungen $f : M \to R$ mit $f(m) = 0$ für fast alle $m \in M$ ein (unitaler) Unterring des (unitalen) Rings $\mathrm{Abb}(M, R)$ aus Beispiel 2.3.4. 4. (Übung).

5. Ist S ein unitaler Unterring eines kommutativen unitalen Rings $(R,+,\cdot)$, so ist auch der Polynomring $S[x]$ ein unitaler Unterring des kommutativen unitalen Rings $R[x]$.

6. Für jeden kommutativen unitalen Ring $(R,+,\cdot)$ und jedes Element $b \in R$ ist die **Evaluationsabbildung** $\mathrm{ev}_b : R[x] \to R$, $p \mapsto \mathrm{ev}_b(p)$, die einem Polynom $p : \mathbb{N}_0 \to R$ mit $p(n) = 0$ für fast alle $n \in \mathbb{N}_0$ das Element $\mathrm{ev}_b(p) = \Sigma_{k\in\mathbb{N}_0} p(k)b^k$ zuordnet, ein unitaler Ringhomomorphismus. (Beweis: Übung).

7. Für jeden unitalen Ring $(R,+,\cdot)$ gibt es genau einen unitalen Ringhomomorphismus $\phi : \mathbb{Z} \to R$, der gegeben ist durch
$$\phi(m) = m1_R = \underbrace{1_R + \ldots + 1_R}_{m\times}, \quad \phi(-m) = -(m1_R) \text{ für } m \in \mathbb{N}, \quad \phi(0) = 0_R.$$
Denn ein unitaler Ringhomomorphismus $\phi : \mathbb{Z} \to R$ muss $\phi(1) = 1_R$ und $\phi(0) = 0_R$ erfüllen. Daraus folgt $\phi(m) = \phi(1 + \ldots + 1) = 1_R + \ldots + 1_R = m1_R$ und $\phi(-m) = -(m1_R)$ für alle $m \in \mathbb{N}$.

Wie auch im Fall von Gruppenhomomorphismen, kann man auch für (unitale) Ringhomomorphismen zeigen, das die Identitätsabbildung, die Umkehrabbildung eines (unitalen) Ringisomorphismus und die Verkettung zweier (unitaler) Ringhomomorphismen wieder (unitale) Ringhomomorphismen sind. Diese Aussagen garantieren wieder, dass die Eigenschaft "isomorph sein zu" für (unitale) Ringe reflexiv, symmetrisch und transitiv ist[8]. Dies macht es sinnvoll, isomorphe (unitale) Ringe als äquivalent zu betrachten und Ringe bis auf Isomorphie zu klassifizieren.

Satz 2.3.10:

1. Für alle (unitalen) Ringe $(R,+,\cdot)$ ist die Identitätsabbildung $\mathrm{id}_R : R \to R$ ein (unitaler) Ringisomorphismus.
2. Ist $\phi : R \to S$ ein (unitaler) Ringisomorphismus, so ist auch die Umkehrabbildung $\phi^{-1} : S \to R$ ein (unitaler) Ringisomorphismus.
3. Sind $\phi : R \to S$, $\psi : S \to T$ (unitale) Ringhomomorphismen, so ist auch die Verkettung $\psi \circ \phi : R \to T$, $r \mapsto \psi(\phi(r))$ ein (unitaler) Ringhomomorphismus.
4. Die (unitalen) Ringautomorphismen eines (unitalen) Rings $(R,,\cdot)$ bilden eine Gruppe $\mathrm{Aut}(R)$.

Beweis:
In Satz 2.2.23 wurde bereits gezeigt, dass $\mathrm{id}_R : R \to R$, $\phi^{-1} : S \to R$ Gruppenisomorphismen und $\psi \circ \phi : R \to T$ ein Gruppenhomomorphismus ist. Zu zeigen ist noch die entsprechende Bedingung für die Multiplikation und im Fall unitaler Ringe die Bedingung für die Einselemente. Seien also $(R,+,\cdot)$, $(S,+',\cdot')$ und $(T,+'',\cdot'')$ Ringe.

1. Dann gilt $\mathrm{id}_R(r \cdot r') = r \cdot r' = \mathrm{id}_R(r) \cdot \mathrm{id}_R(r')$ für alle $r \in R$ und gegebenenfalls $\mathrm{id}_R(1_R) = 1_R$.

2. Falls Einselemente 1_R, 1_S existieren folgt aus $\phi(1_R) = 1_S$ auch $\phi^{-1}(1_S) = 1_R$, und für die Multiplikation ergibt sich für alle $s, s' \in S$
$$\phi^{-1}(s) \cdot \phi^{-1}(s') = \phi^{-1}(\phi(\phi^{-1}(s) \cdot \phi^{-1}(s'))) = \phi^{-1}(\phi(\phi^{-1}(s)) \cdot' \phi(\phi^{-1}(s'))) = \phi^{-1}(s \cdot' s').$$

[8] Achtung! Auch bei der Isomorphie von (unitalen) Ringen handelt es sich nicht um eine Äquivalenzrelation auf einer Menge, sondern auf einer Klasse, denn die (unitalen) Ringe bilden keine Menge.

3. Falls $\phi : R \to S$, $\psi : S \to T$ unital sind, so ergibt sich $\phi(1_R) = 1_S$, $\psi(1_S) = 1_T$ und damit auch $(\psi \circ \phi)(1_R) = \psi(\phi(1_R)) = \psi(1_S) = 1_T$. Für die Multiplikation gilt für alle $r, r' \in R$

$$(\psi \circ \phi)(r \cdot r') = \psi(\phi(r \cdot r')) = \psi(\phi(r) \cdot' \phi(r')) = \psi(\phi(r)) \cdot'' \psi(\phi(r')) = (\psi \circ \phi)(r) \cdot'' (\psi \circ \phi)(r').$$

4. Die Verkettung von unitalen Ringautomorphismen definiert nach 3. eine assoziative Verknüpfung $\circ : \mathrm{Aut}(R) \times \mathrm{Aut}(R) \to \mathrm{Aut}(R)$. Nach 1. besitzt sie das neutrale Element id_R und nach 2. ist $\phi^{-1} : R \to R$ das Inverse eines (unitalen) Ringautomorphismus $\phi : R \to R$. $\square$

Das Konzept eines unitalen Rings ist allgemeiner als unser intuitiver Zahlenbegriff, da es nicht fordert, dass die Multiplikation kommutativ ist und auch keine Forderung bezüglich der Existenz von multiplikativen Inversen stellt. Fordert man, dass ein unitaler Ring kommutativ ist, $1 \neq 0$ gilt und jedes Element außer dem Nullelement 0 ein multiplikatives Inverses besitzt, so erhält man den Begriff eines Körpers.

Definition 2.3.11:
Ein Ring $(R, +, \cdot)$ heißt **Körper**, wenn $(R \setminus \{0\}, \cdot)$ eine abelsche Gruppe ist.

Aus der Forderung, dass $R \setminus \{0\}$ eine abelsche Gruppe ist, folgt direkt, dass der Ring $(R, +, \cdot)$ kommutativ und unital ist, und dass das neutrale Element $1 \in R$ der Multiplikation verschieden von dem neutralen Element $0 \in R$ der Addition ist. Damit hat ein Körper immer *mindestens zwei Elemente*, nämlich 0 und $1 \neq 0$. Insbesondere ist also der *Nullring kein Körper*.

Bevor wir Bespiele von Körpern betrachten, fassen wir alle Bedingungen aus Definition 2.3.11 noch einmal zu einer übersichtlichen Liste von Bedingungen zusammen, die ein Körper erfüllt.

Bemerkung 2.3.12: Eine Menge $\mathbb{K}$ mit zwei Verknüpfungen $+, \cdot : \mathbb{K} \times \mathbb{K} \to \mathbb{K}$ ist ein Körper genau dann, wenn die folgenden Bedingungen erfüllt sind:

(K1) **Assoziativität:** $(a + b) + c = a + (b + c)$ und $a \cdot (b \cdot c) = (a \cdot b) \cdot c$ für alle $a, b, c \in \mathbb{K}$.
(K2) **Kommutativität:** $a + b = b + a$ und $a \cdot b = b \cdot a$ für alle $a, b \in \mathbb{K}$.
(K3) **Distributivgesetz:** $a \cdot (b + c) = a \cdot b + a \cdot c$ für alle $a, b, c \in \mathbb{K}$.
(K4) **neutrale Elemente:** $\exists 0, 1 \in \mathbb{K}$ mit $1 \neq 0$, so dass $a + 0 = a$ und $a \cdot 1 = a$ für alle $a \in \mathbb{K}$.
(K5) **Inverse:** Für alle $a \in \mathbb{K}$ existiert ein Element $b \in \mathbb{K}$ mit $a + b = 0$, und für alle $a \in \mathbb{K} \setminus \{0\}$ existiert ein Element $c \in \mathbb{K}$ mit $a \cdot c = 1$.

Das Inverse eines Elements $a \in \mathbb{K}$ bezüglich $+$ bezeichnet man mit $-a$ und das Inverse eines Elements $a \in \mathbb{K} \setminus \{0\}$ bezüglich $\cdot$ mit a^{-1} oder $1/a$. Der Punkt für die Multiplikation wird oft weggelassen, und statt $a \cdot b^{-1}$ schreibt man oft a/b.

Körper sind also unitale Ringe, die bestimmte Bedingungen erfüllen, aber keine zusätzliche Struktur wie weitere Verknüpfungen enthalten. Ein Ring ist entweder ein Körper oder nicht. Daher ergibt sich die Definition eines Teilkörpers und eines Körperhomomorphismus direkt aus den entsprechenden Definitionen für unitale Ringe.

Definition 2.3.13: Seien $(\mathbb{K}, +, \cdot)$ und $(\mathbb{L}, +', \cdot')$ Körper.

1. Ein **Teilkörper** von $(\mathbb{K}, +, \cdot)$ ist ein unitaler Unterring von $(\mathbb{K}, +, \cdot)$, der selbst wieder ein Körper ist.

2. Ein **Körperhomomorphismus** $\phi : (\mathbb{K}, +, \cdot) \to (\mathbb{L}, +', \cdot')$ ist ein Homomorphismus von unitalen Ringen, ein **Körperisomorphismus** ein Isomorphismus von unitalen Ringen.

Beispiel 2.3.14:

1. $(\mathbb{Q}, +, \cdot)$ und $(\mathbb{R}, +, \cdot)$ sind Körper, und $(\mathbb{Q}, +, \cdot)$ ist ein Teilkörper von $(\mathbb{R}, +, \cdot)$
2. $(\mathbb{Z}, +, \cdot)$ ist ein kommutativer unitaler Ring, aber kein Köper, denn nur die Elemente 1 und -1 besitzen ein multiplikatives Inverses.
3. Für jedes $q \in \mathbb{Q}$ das kein Quadrat einer rationalen Zahl ist, ist $(\mathbb{Q}^2, +, \cdot)$ mit

$$(r, s) + (r', s') := (r + r', s + s') \qquad (r, s) \cdot (r', s') := (rr' + qss', rs' + r's)$$

ein Körper, der $\mathbb{Q} \times \{0\} \cong (\mathbb{Q}, +, \cdot)$ als Teilkörper enthält (Beweis: Übung).
Schreibt man $r + \sqrt{q}s := (r, s)$ so lauten das Additions- und Multiplikationsgesetz

$$(r+\sqrt{q}s)+(r'+\sqrt{q}s') = (r+r')+\sqrt{q}(s+s') \quad (r+\sqrt{q}s)\cdot(r'+\sqrt{q}s') = (rr'+qss')+\sqrt{q}(rs'+r's).$$

Man bezeichnet diesen Körper daher mit $\mathbb{Q}(\sqrt{q})$.

Während alle Körper in Beispiel 2.3.14 unendlich viele Elemente besitzen, gibt es auch Körper für die dies nicht der Fall ist. Die einfachsten Beispiele endlicher Körper sind die Restklassenringe $\mathbb{Z}/p\mathbb{Z}$ für *Primzahlen* $p \in \mathbb{N}$. Sie enthalten genau p Elemente.

Satz 2.3.15:
Der Restklassenring $(\mathbb{Z}/n\mathbb{Z}, +, \cdot)$ ist ein Körper genau dann, wenn $n \in \mathbb{N}$ eine Primzahl ist.

Beweis:
In Lemma 2.3.6 wurde bereits gezeigt, dass $(\mathbb{Z}/n\mathbb{Z}, +, \cdot)$ für $n \in \mathbb{N}$ ein kommutativer unitaler Ring ist, also dass die Bedingungen (K1)-(K4) aus Bemerkung 2.3.12 erfüllt sind und (K5) für die Addition. Ist $n = 1$, so ist $\mathbb{Z}/n\mathbb{Z}$ der Nullring und damit kein Körper. Zu zeigen ist also nur noch, dass für $n \geq 2$ jedes Element $[m] \in \mathbb{Z}/n\mathbb{Z} \setminus \{[0]\}$ ein multiplikatives Inverses besitzt genau dann, wenn $n \in \mathbb{Z}$ eine Primzahl ist.

$\Rightarrow$: Ist $n \geq 2$ keine Primzahl, so gibt es natürliche Zahlen $a, b \in \{2, ..., n-1\}$ mit $a \cdot b = n$. Daraus folgt $[a] \cdot [b] = [ab] = [n] = [0]$ und $[a], [b] \neq [0]$. Besäße $[a]$ ein multiplikatives Inverses $[a]^{-1}$ mit $[a]^{-1} \cdot [a] = [1]$, so würde mit der Assoziativität und Bemerkung 2.3.2, 1. folgen $[b] = [1] \cdot [b] = ([a]^{-1} \cdot [a]) \cdot [b] = [a]^{-1} \cdot ([a] \cdot [b]) = [a]^{-1} \cdot [0] = [0]$ - ein Widerspruch. Also kann $[a]$ kein Inverses haben und $(\mathbb{Z}/n\mathbb{Z}, +, \cdot)$ ist kein Körper.

$\Leftarrow$: Sei nun n eine Primzahl. Zu zeigen ist, dass für $m \in \{1, ..., n-1\}$ das Element $[m] \in \mathbb{Z}/n\mathbb{Z}$ ein multiplikatives Inverses besitzt, also dass ein $a \in \mathbb{Z}$ mit $[a] \cdot [m] = [1]$ existiert. Dazu betrachten wir die nicht-leere Menge $M = \{am + bn | a, b \in \mathbb{Z}, am + bn > 0\} \subseteq \mathbb{N}$ und definieren $d = \min M$ als die kleinste natürliche Zahl, die in M enthalten ist. Dann gilt $1 \leq d \leq m < n$, denn $m = 1 \cdot m + 0 \cdot n \in M$, und es gibt $a, b \in \mathbb{Z}$ mit $d = am + bn$.

Wäre $d = am + bn \geq 2$ so gäbe es ein $c \in \mathbb{Z}$, $r \in \{0, ..., d-1\}$ mit $n = cd + r$ (Division mit Rest durch d). Da $n > d$ eine Primzahl ist und $d \geq 2$ angenommen wurde, müsste $r \in \{1, ..., d-1\}$ gelten. Daraus ergäbe sich $r = n - cd = n - c(am + bn) = (-ac)m + (1-b)n$ mit $-ac, 1-b \in \mathbb{Z}$, also wegen $r > 0$ auch $r \in M$, im Widerspruch zur Minimalität von d.

Also muss $d = am + bn = 1$ gelten. Daraus folgt $[1] = [am + bn] = [am] + [bn] = [am] = [a] \cdot [m]$. Damit ist gezeigt, dass $[m]$ ein multiplikatives Inverses besitzt. Also ist für jede Primzahl n der Restklassenring $(\mathbb{Z}/n\mathbb{Z}, +, \cdot)$ ein Körper. □

Bemerkung 2.3.16:

1. Ist $(R, +, \cdot)$ ein kommutativer unitaler Ring, so bezeichnet man ein Element $a \in R \setminus \{0\}$ für das ein Element $b \in R \setminus \{0\}$ mit $a \cdot b = 0$ existiert als einen **Nullteiler** von R. Der erste Teil des Beweises von Satz 2.3.15 zeigt, dass Nullteiler keine multiplikativen Inversen haben. Damit kann ein Körper keine Nullteiler haben, ist also **nullteilerfrei**.

   Ein kommutativer nullteilerfreier unitaler Ring heißt auch **Integritätsring** oder **Integritätsbereich**. Wie das Beispiel des Integritätsrings $(\mathbb{Z}, +, \cdot)$ zeigt, ist zwar jeder Körper ein Integritätsring, aber nicht jeder Integritätsring mit $1 \neq 0$ ein Körper. *Endliche* Integritätsringe mit $1 \neq 0$ sind jedoch immer Körper (Beweis: Übung)

2. Aus Satz 2.3.15 folgt, dass für jede Primzahl $p \in \mathbb{N}$ ein Körper mit p Elementen existiert, nämlich der Körper $(\mathbb{Z}/p\mathbb{Z}, +, \cdot)$, der oft auch mit $\mathbb{F}_p$ bezeichnet wird. Allgemeiner kann man zeigen (siehe *Vorlesung Körpertheorie*), dass es zu jeder *Primpotenz* $n = p^d$ mit $d \in \mathbb{N}$ einen Körper $\mathbb{F}_{p^d}$ mit p^d Elementen gibt, und dass jeder *endliche* Körper isomorph zu genau einem dieser Körper $\mathbb{F}_{p^d}$ ist.

3. Für einen Integritätsbereich $(R, +, \cdot)$ definiert man die **Charakteristik** als $\mathrm{char}(R) = 0$, wenn $n \cdot 1 = 1 + \ldots + 1 \neq 0$ für alle $n \in \mathbb{N}$ und $\mathrm{char}(R) = \min\{n \in \mathbb{N} |\ n \cdot 1 = 1 + ... + 1 = 0\}$ sonst. Es gilt also $\mathrm{char}(\mathbb{R}) = \mathrm{char}(\mathbb{Q}) = 0$ und $\mathrm{char}(\mathbb{F}_p) = p$, denn $[0] = [p] = [1 + ... + 1] = [1] + ... + [1]$ und $[m] \neq [0]$ für $m \in \{1, ..., p-1\}$. Allgemeiner gilt, dass ein endlicher Körper $\mathbb{F}_{p^d}$ mit $d \in \mathbb{N}$ immer Charakteristik $\mathrm{char}(\mathbb{F}_{p^d}) = p$ hat.

Ein sehr wichtiger Körper ist der Körper $\mathbb{C}$ der komplexen Zahlen. Er ist besonders nützlich, weil polynomiale Gleichungen über dem Körper der komplexen Zahlen immer eine Lösung besitzen. Der Körper $\mathbb{C}$ spielt auch eine wichtige Rolle in der Geometrie. Die komplexen Zahlen werden beispielsweise in der *Vorlesung Körpertheorie* benutzt um zu beweisen, dass man einen Winkel im Allgemeinen nicht dritteln kann und die Quadratur des Kreises unmöglich ist. Komplexen Zahlen haben auch viele Anwendungen in der Physik, in der Elektrotechnik und den Ingenieurwissenschaften.

Satz 2.3.17: Die Menge $\mathbb{R}^2$ mit den Verknüpfungen $+, \cdot : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}^2$,

$$(a, b) + (c, d) := (a + c, b + d) \qquad (a, b) \cdot (c, d) := (ac - bd, ad + bc) \qquad \forall a, b, c, d \in \mathbb{R}$$

ist ein Körper. Er wird als der Körper der **komplexen Zahlen** und mit $\mathbb{C}$ bezeichnet. Der Körper $\mathbb{R}$ der reellen Zahlen ist isomorph zu dem Teilkörper $\mathbb{R} \times \{0\} \subseteq \mathbb{C}$.

Beweis:
$(\mathbb{R}^2, +)$ ist das direkte Produkt der abelschen Gruppe $(\mathbb{R}, +)$ mit sich selbst und somit eine abelsche Gruppe mit neutralem Element $(0, 0)$ und Inversen $-(a, b) = (-a, -b)$. Auch die Verknüpfung $\cdot$ ist offensichtlich kommutativ. Dass sie auch assoziativ ist folgt durch Nachrechnen

$$((a, b) \cdot (c, d)) \cdot (e, f) = (ac - bd, ad + bc) \cdot (e, f) = (ace - bde - adf - bcf, acf - bdf + ade + bce)$$
$$= (a, b) \cdot (ce - df, cf + de) = (a, b) \cdot ((c, d) \cdot (e, f)),$$

und ebenso beweist man das Distributivgesetz

$$\begin{aligned}&(a,b)\cdot((c,d)+(e,f))=(a,b)\cdot(c+e,d+f)=(ac+ae-bd-bf,ad+af+bc+be)\\&=(ac-bd,ad+bc)+(ae-bf,af+bd)=(a,b)\cdot(c,d)+(a,b)\cdot(e,f).\end{aligned}$$

Das Element $(1,0)$ erfüllt $(1,0)\cdot(a,b)=(1\cdot a-0\cdot b,1\cdot b+0\cdot a)=(a,b)$ für alle $a,b\in\mathbb{R}$ und ist somit das neutrale Element für die Multiplikation. Ist $(a,b)\neq(0,0)$, so gilt $a^2+b^2\neq 0$ und

$$\left(\frac{a}{a^2+b^2},\frac{-b}{a^2+b^2}\right)\cdot(a,b)=\left(\frac{a^2+b^2}{a^2+b^2},\frac{ab-ab}{a^2+b^2}\right)=(1,0). \tag{4}$$

Also besitzt jedes Element $(a,b)\in\mathbb{C}\setminus\{0\}$ ein multiplikatives Inverses, und $\mathbb{C}$ ist ein Körper. Um zu zeigen, dass $\mathbb{R}$ isomorph zu $\mathbb{R}\times\{0\}\subseteq\mathbb{C}$ ist, betrachtet man die bijektive Abbildung $\phi:\mathbb{R}\to\mathbb{R}\times\{0\}$, $a\mapsto(a,0)$. Da $\phi(1)=(1,0)$ und für alle $a,b\in\mathbb{R}$

$$\phi(a+b)=(a+b,0)=(a,0)+(b,0)=\phi(a)+\phi(b)\quad \phi(a\cdot b)=(a\cdot b,0)=(a,0)\cdot(b,0)=\phi(a)\cdot\phi(b),$$

ist $\phi(\mathbb{R})=\mathbb{R}\times\{0\}$ ein Teilkörper von $\mathbb{C}$, und $\phi:\mathbb{R}\to\mathbb{R}\times\{0\}$ ist ein Körperisomorphismus. □

Wie man am Beweis von Satz 2.3.17 sieht, ist die Notation für komplexe Zahlen relativ sperrig und unübersichtlich. Der Körperisomorphismus $\phi:\mathbb{R}\to\mathbb{R}\times\{0\}$, $a\mapsto(a,0)$ rechtfertigt die kürzere Schreibweise a für $(a,0)$. Definiert man zusätzlich $\mathrm{i}:=(0,1)$ und lässt das Symbol $\cdot$ in Produkten weg, so ergibt sich für alle $a,b\in\mathbb{R}$

$$a+\mathrm{i}b=(a,0)+(0,1)\cdot(b,0)=(a,0)+(0,b)=(a,b)$$

Jede komplexe Zahl lässt sich also schreiben als $a+\mathrm{i}b$ mit $a,b\in\mathbb{R}$, und es gilt $a+\mathrm{i}b=c+\mathrm{i}d$ genau dann, wenn $a=c$ und $b=d$. Die Formeln für die Verknüpfungen in Satz 2.3.17 lassen sich dann mit Hilfe der Identität $\mathrm{i}^2=(0,1)\cdot(0,1)=(-1,0)=-1$ schreiben als

$$\begin{aligned}&(a+\mathrm{i}b)+(c+\mathrm{i}d)=(a+c)+\mathrm{i}(b+d)\\&(a+\mathrm{i}b)\cdot(c+\mathrm{i}d)=ac+\mathrm{i}(bc+ad)+\mathrm{i}^2bd=(ac-bd)+\mathrm{i}(ad+bc)\end{aligned} \tag{5}$$

Wir fassen diese und noch einige weitere wichtige Bezeichnungen für komplexe Zahlen in einer Definition zusammen.

Definition 2.3.18: Im Körper $\mathbb{C}$ der komplexen Zahlen benutzt man die Bezeichnung $x+\mathrm{i}y$ für (x,y) und Formel (5) für die Verknüpfungen. Das Element $\mathrm{i}:=(0,1)$ mit $\mathrm{i}^2=-1$ heißt **imaginäre Einheit**, und eine Zahl der Form $\mathrm{i}y$, $y\in\mathbb{R}$ heißt **imaginär**.

Für $z=x+\mathrm{i}y\in\mathbb{C}$ mit $x,y\in\mathbb{R}$ definiert man

- den **Realteil** $\mathrm{Re}\,(z):=\mathrm{Re}\,(x+\mathrm{i}y)=x\in\mathbb{R}$,
- den **Imaginärteil** $\mathrm{Im}\,(z):=\mathrm{Im}\,(x+\mathrm{i}y)=y\in\mathbb{R}$,
- die zu z **komplex konjugierte** Zahl $\overline{z}:=\overline{x+\mathrm{i}y}=x-\mathrm{i}y\in\mathbb{C}$,
- den **Betrag** $|z|=|x+\mathrm{i}y|:=\sqrt{x^2+y^2}=\sqrt{z\overline{z}}\in\mathbb{R}$,
- das **Argument** $\arg(z)\in[0,2\pi)$ von $z=x+\mathrm{i}y\in\mathbb{C}\setminus\{0\}$ als den eindeutigen Winkel mit

$$\cos(\arg(z))=\frac{\mathrm{Re}\,(z)}{|z|}=\frac{x}{\sqrt{x^2+y^2}}\quad\wedge\quad\sin(\arg(z))=\frac{\mathrm{Im}\,(z)}{|z|}=\frac{y}{\sqrt{x^2+y^2}}.$$

Die Beziehungen zwischen den in Definition 2.3.18 eingeführten Größen und ihre wichtigsten Eigenschaften ergeben sich aus dem folgenden Lemma, das wir anschließend auch geometrisch interpretieren werden.

Lemma 2.3.19:

1. Die komplexe Konjugation ist ein Körperisomorphismus $\bar{}: \mathbb{C} \to \mathbb{C}$, der $\mathbb{R}$ fixiert. Es gilt:

$$\overline{z+w} = \overline{z} + \overline{w}, \qquad \overline{z \cdot w} = \overline{z} \cdot \overline{w}, \qquad \overline{x} = x, \qquad \overline{\overline{z}} = z \qquad \forall z, w \in \mathbb{C}, x \in \mathbb{R}$$

2. Für alle $z \in \mathbb{C}$ gelten die Identitäten

$$\mathrm{Re}\,(z) = \tfrac{1}{2}(z + \overline{z}), \qquad \mathrm{Im}\,(z) = \tfrac{\mathrm{i}}{2}(\overline{z} - z), \qquad z \cdot \overline{z} = |z|^2.$$

Beweis:
1. Sei $z = x + \mathrm{i}y$ und $w = u + \mathrm{i}v$ mit $x, y, u, v \in \mathbb{R}$. Dann ergibt sich:

$$\begin{aligned}
&\overline{z+w} = \overline{(x+u) + \mathrm{i}(y+v)} = x + u - \mathrm{i}(y+v) = x - \mathrm{i}y + u - \mathrm{i}v = \overline{z} + \overline{w} \\
&\overline{z \cdot w} = \overline{xu - yv + \mathrm{i}(xv + yu)} = xu - yv - \mathrm{i}(xv + yu) = (x - \mathrm{i}y)(u - \mathrm{i}v) = \overline{z} \cdot \overline{w} \\
&\overline{x} = x \qquad \overline{\overline{z}} = \overline{\overline{x + \mathrm{i}y}} = \overline{x - \mathrm{i}y} = x + \mathrm{i}y = z.
\end{aligned}$$

2. Es gilt für $z = x + \mathrm{i}y$ mit $x, y \in \mathbb{R}$:

$$\begin{aligned}
&\tfrac{1}{2}(z + \overline{z}) = \tfrac{1}{2}(x + \mathrm{i}y + x - \mathrm{i}y) = x \qquad \tfrac{\mathrm{i}}{2}(\overline{z} - z) = \tfrac{1}{2}\mathrm{i}(x - \mathrm{i}y - x - \mathrm{i}y) = -\mathrm{i}^2 y = y \\
&z \cdot \overline{z} = (x + \mathrm{i}y)(x - \mathrm{i}y) = x^2 + y^2 = |z|^2.
\end{aligned}$$

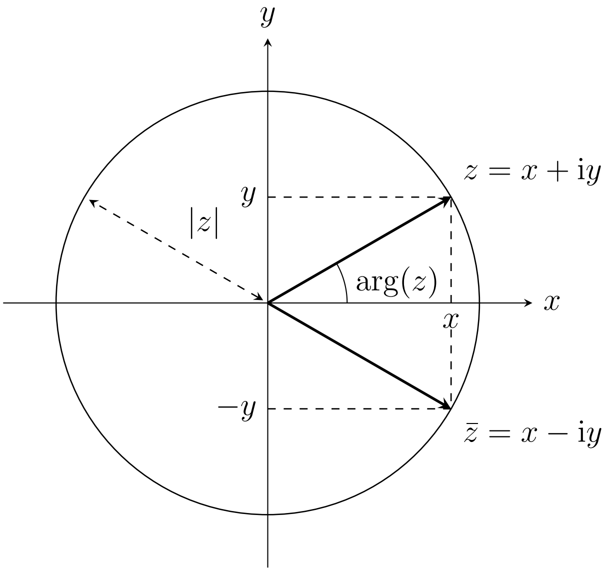
Aus $\overline{z+w} = \overline{z} + \overline{w}$, $\overline{z \cdot w} = \overline{z} \cdot \overline{w}$ und $\overline{1} = 1$ folgt, dass die komplexe Konjugation ein Körperhomomorphismus ist, und aus $\overline{\overline{z}} = z$ für alle $z \in \mathbb{C}$ folgt, dass es sich um einen Körperisomorphismus handelt. □

Beispiel 2.3.20: Wir bestimmen $\mathrm{Re}\,(1/z)$ und $\mathrm{Im}\,(1/z)$ für $z = 1 + 2\mathrm{i}$. Dazu könnten wir im Prinzip Formel (4) für die multiplikativen Inversen benutzen. Eine einfachere Methode, die allgemein für Brüche komplexer Zahlen funktioniert, ist das Erweitern des Bruchs mit dem komplex Konjugierten des Nenners:

$$\frac{1}{z} = \frac{1}{1+2\mathrm{i}} = \frac{1-2\mathrm{i}}{(1-2\mathrm{i})(1+2\mathrm{i})} = \frac{1-2\mathrm{i}}{1-4\mathrm{i}^2} = \frac{1-2\mathrm{i}}{5} \quad \Rightarrow \quad \mathrm{Re}\,(1/z) = \tfrac{1}{5}, \mathrm{Im}\,(1/z) = -\tfrac{2}{5}.$$

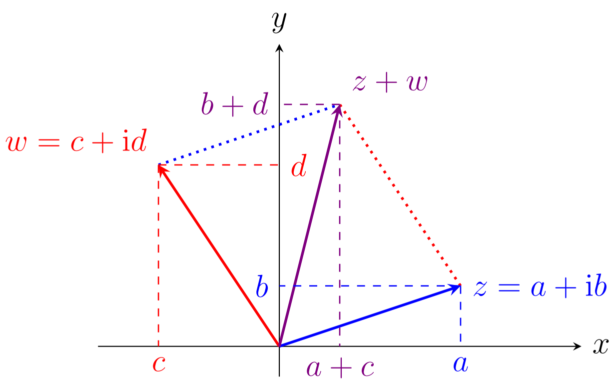
Da es sich bei komplexen Zahlen um Elemente der Menge $\mathbb{R}^2$ handelt, können wir sie als Punkte in der Ebene interpretieren. Dabei entspricht die komplexe Zahl $z = x + \mathrm{i}y$ dem Punkt (x, y).

- Der Punkt (x, y) hat nach dem Satz des Pythagoras vom Ursprung $(0, 0)$ den Abstand $|z| = \sqrt{x^2 + y^2}$.
- Die Gerade durch den Ursprung und (x, y) schließt mit der positiven x-Achse einen Winkel $\arg(z) \in [0, 2\pi)$ mit $\cos(\arg(z)) = x/\sqrt{x^2 + y^2}$ und $\sin(\arg(z)) = y/\sqrt{x^2 + y^2}$ ein.
- Die komplexe Zahl $\overline{z} = x - iy$ hat den selben Abstand vom Ursprung wie $z = x + \mathrm{i}y$ und schließt mit der positiven x-Achse den Winkel $\arg(\overline{z}) = 2\pi - \arg(z)$ ein. Die komplexe Konjugation entspricht also der Spiegelung an der x-Achse.



Betrag $|z|$ und Argument $\arg(z)$ einer komplexen Zahl $z = x + \mathrm{i}y$ und komplexe Konjugation.

Mit dieser geometrischen Beschreibung der komplexen Zahlen können wir auch ihrer Addition und Multiplikation geometrisch interpretieren. Die Formel $(a+\mathrm{i}b)+(c+\mathrm{i}d) = (a+c)+\mathrm{i}(b+d)$ bedeutet, dass die Addition komplexer Zahlen der Konstruktion eines Parallelogramms entspricht. Betrachtet man das von den Punkten $(0,0)$, (a,b) und (c,d) aufgespannte Parallelogramm, so ist dessen vierter Eckpunkt gerade der Punkt $(a+c, b+d)$, der der Summe der komplexen Zahlen $z = a + \mathrm{i}b$ und $w = c + \mathrm{i}d$ entspricht (Parallelogrammregel).

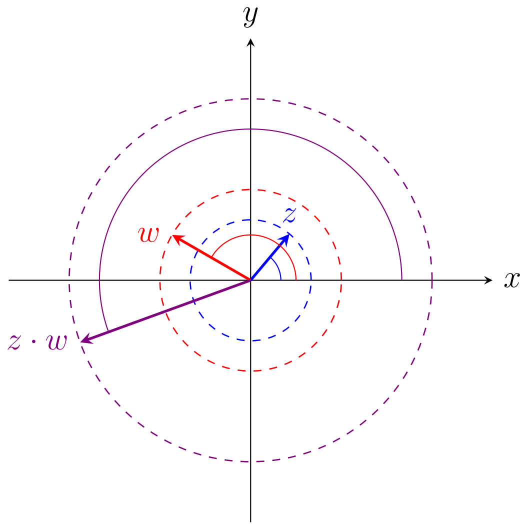


Addition zweier komplexer Zahlen mit der Parallelogrammregel.

Um die geometrische Bedeutung der Multiplikation komplexer Zahlen zu verstehen, betrachten wir zwei komplexe Zahlen $z = |z|(\cos\arg(z)+\mathrm{i}\sin\arg(z))$ und $w = |w|(\cos\arg(w)+\mathrm{i}\sin\arg(w))$. Dann ergibt sich mit Hilfe der Additionsformeln für Sinus und Kosinus für das Produkt $z \cdot w$

$$
\begin{aligned}
&z \cdot w = |z|(\cos\arg(z) + \mathrm{i}\sin\arg(z)) \cdot |w|(\cos\arg(w) + \mathrm{i}\sin\arg(w)) \\
&= |z||w|(\cos\arg(z)\cos\arg(w) - \sin\arg(z)\sin\arg(w) + \mathrm{i}(\sin\arg(z)\cos\arg(w) + \sin\arg(w)\cos\arg(z))) \\
&= |z||w|(\cos(\arg(z) + \arg(w)) + \mathrm{i}\sin(\arg(z) + \arg(w))).
\end{aligned}
$$

Also gilt $|z \cdot w| = |z| \cdot |w|$ und $\arg(z \cdot w) = \arg(z) + \arg(w)$. Das Produkt komplexer Zahlen multipliziert also deren Abstände vom Ursprung und addiert die zugehörigen Winkel mit der positiven x-Achse, letzteres bis auf Vielfache von 2π.



Definiert man
$$e^{a+ib} := e^a(\cos b + \mathrm{i}\sin b) \qquad \forall a, b \in \mathbb{R},$$
so erhält man die an die reelle Exponentialfunktion erinnernde Formel[9]
$$e^{a+\mathrm{i}b} \cdot e^{c+\mathrm{i}d} = e^a e^c(\cos b + \mathrm{i}\sin b)(\cos d + \mathrm{i}\sin d) = e^{a+c}e^{\mathrm{i}(b+d)} = e^{(a+\mathrm{i}b)+(c+\mathrm{i}d)} \qquad \forall a, b, c, d \in \mathbb{R}.$$

Korollar 2.3.21:

1. Jede komplexe Zahl $z \in \mathbb{C} \setminus \{0\}$ lässt sich eindeutig schreiben als
$$z = re^{\mathrm{i}\phi} = r(\cos\phi + \mathrm{i}\sin\phi) \qquad \text{mit} \quad r = |z| \in \mathbb{R}^+, \phi = \arg(z) \in [0, 2\pi).$$
Dies bezeichnet man als die **Polardarstellung** von z, während die Darstellung $z = x+\mathrm{i}y$ mit $x, y \in \mathbb{R}$ **kartesische Darstellung** oder **Normaldarstellung** von z heißt.

2. Für $z = re^{\mathrm{i}\phi}$ und $w = se^{\mathrm{i}\psi}$ mit $r, s \in \mathbb{R}$ und $\phi, \psi \in \mathbb{R}$ gilt $z \cdot w = rse^{\mathrm{i}(\phi+\psi)}$.

3. Es gilt die **de Moivresche Formel**
$$(\cos\phi + \mathrm{i}\sin\phi)^n = e^{\mathrm{i}n\phi} = \cos(n\phi) + \mathrm{i}\sin(n\phi) \qquad \forall n \in \mathbb{Z}, \phi \in [0, 2\pi)$$

Beispiel 2.3.22: Wir bestimmen die Polardarstellung der komplexen Zahl $z = 1 - \mathrm{i}$ und die kartesische Darstellung der komplexen Zahl $w = 3e^{\mathrm{i}\pi/3}$.

Für $z = 1 - \mathrm{i}$ gilt $|z| = \sqrt{1^2 + (-1)^2} = \sqrt{2}$ und damit auch
$$1/\sqrt{2} = \mathrm{Re}(z)/|z| = \cos(\arg(z)) = -\mathrm{Im(z)}/|z| = -\sin(\arg(z)) \quad \Rightarrow \quad \arg(z) = 7\pi/4.$$
Daraus folgt $z = \sqrt{2}e^{7\pi\mathrm{i}/4}$. Für w ergibt sich aus den Identitäten $\cos(\pi/3) = 1/2$ und $\sin(\pi/3) = \sqrt{3}/2$ die kartesische Darstellung $w = 3(\cos(\pi/3) + \mathrm{i}\sin(\pi/3)) = \frac{3}{2}(1 + \sqrt{3}\mathrm{i})$.

Eine wichtige Aussage über komplexe Zahlen ist der Fundamentalsatz der Algebra, dessen Beweis in der Vorlesung *Analysis* oder *Funktionentheorie* erfolgt. Er besagt, dass jedes nichtkonstante Polynom mit Koeffizienten in $\mathbb{C}$ über $\mathbb{C}$ in Linearfaktoren zerfällt.

[9]Dies ist kein Zufall. Wie auch in $\mathbb{R}$ kann man die Exponentialfunktion $\exp : \mathbb{C} \to \mathbb{C}$ als eine Potenzreihe $\exp(z) = \Sigma_{k=0}^{\infty} z^k/k!$ definieren. Diese konvergiert für alle $z \in \mathbb{C}$ und liefert genau diese Formel. Dies wird in der Vorlesung Analysis I behandelt.

Satz 2.3.23 (**Fundamentalsatz der Algebra**):

Jedes nicht-konstante Polynom mit Koeffizienten in $\mathbb{C}$ hat eine Nullstelle in $\mathbb{C}$ und zerfällt deshalb über $\mathbb{C}$ in **Linearfaktoren**: zu $p = a_n x^n + a_{n-1}x^{n-1} + ... + a_1 x + a_0$ mit $a_n, ..., a_0 \in \mathbb{C}$, $a_n \neq 0$ und $n \in \mathbb{N}$ gibt es $z_1, ..., z_n \in \mathbb{C}$ mit $p = a_n(x - z_1) \cdot (x - z_2) \cdots (x - z_n)$. Die Zahlen $z_1, ..., z_n \in \mathbb{C}$ sind genau die Nullstellen von p in $\mathbb{C}$.

Offensichtlich gilt diese Aussage nicht für reelle Zahlen. Das Polynom $x^2 + 1$ lässt sich nicht als Produkt $(x - r_1)(x - r_2)$ mit $r_1, r_2 \in \mathbb{R}$ schreiben. Denn sonst wären r_1, r_2 reelle Nullstellen, die das Polynom $x^2 + 1$ wegen $x^2 + 1 > 0$ für alle $x \in \mathbb{R}$ aber nicht besitzt. Über $\mathbb{C}$ zerfällt dieses Polynom jedoch in Linearfaktoren, denn es gilt $(x + \mathrm{i})(x - \mathrm{i}) = x^2 - \mathrm{i}^2 = x^2 - (-1) = x^2 + 1$.

Als Spezialfall nicht-konstanter Polynome mit *komplexen* Koeffizienten lässt sich also insbesondere jedes nicht-konstante Polynom mit *reellen* Koeffizienten als Produkt von Linearfaktoren $(x - z_1) \cdots (x - z_n)$ mit $z_j \in \mathbb{C}$ schreiben. In diesem Fall treten die nicht reellen Nullstellen in Paaren zueinander komplex konjugierter Nullstellen auf. Ein besonders wichtiger Fall ist dabei das relle Polynom $x^n - 1$, dessen Nullstellen sich leicht angeben lassen.

Lemma 2.3.24:

1. Ist $p = a_n x^n + ... + a_1 x + a_0$ ein Polynom mit reellen Koeffizienten $a_n, ..., a_0 \in \mathbb{R}$, so ist $z \in \mathbb{C}$ eine Nullstelle von p genau dann, wenn $\overline{z} \in \mathbb{C}$ eine Nullstelle von p ist.
2. Die Nullstellen des Polynoms $x^n - 1$ mit $n \in \mathbb{N}$ sind genau die komplexen Zahlen $z_k = e^{2\pi \mathrm{i} k/n}$ mit $k \in \{0, 1, ..., n-1\}$. Man nennt sie die $n$ten **Einheitswurzeln**.

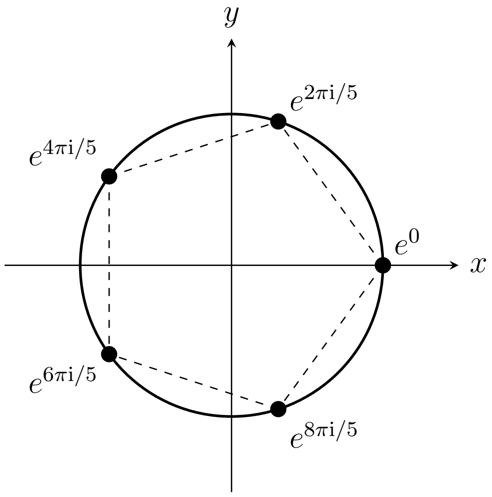
Beweis:
1. Ist $z \in \mathbb{C}$ ist eine Nullstelle von p so gilt $a_n z^n + ... + a_1 z + a_0 = 0$. Da $a_k \in \mathbb{R}$ gilt $\overline{a_k} = a_k$ für alle $k \in \{0, 1, ..., n\}$. Mit Lemma 2.3.19, 1. folgt, dass auch $\overline{z}$ eine Nullstelle ist:

$$0 = \overline{0} = \overline{a_n z^n + ... + a_1 z + a_0} = \overline{a}_n \overline{z}^n + ... + \overline{a}_1 \overline{z} + \overline{a}_0 = a_n \overline{z}^n + ... + a_1 \overline{z} + a_0.$$

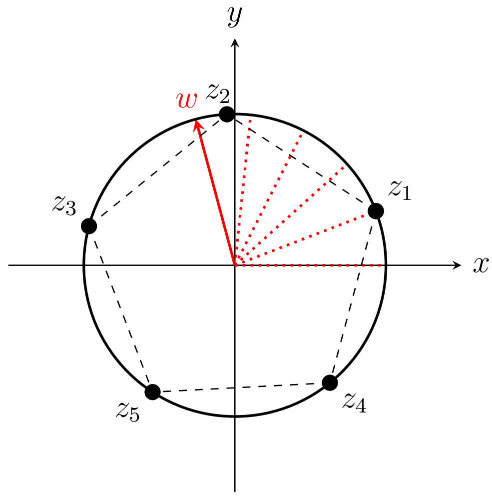
Ist umgekehrt $\overline{z} \in \mathbb{C}$ eine Nullstelle von p, so folgt daraus, dass $\overline{\overline{z}} = z$ eine Nullstelle von p ist.

2. Wir verwenden die Polardarstellung. Eine Zahl $z = re^{\mathrm{i}\phi}$ mit $r \in \mathbb{R}_+$ und $\phi \in [0, 2\pi)$ ist genau dann Lösung der Gleichung $z^n = 1$, wenn $z^n = r^n e^{\mathrm{i}n\phi} = r^n(\cos(n\phi) + \mathrm{i}\sin(n\phi)) = 1$ gilt. Dies ist genau dann der Fall, wenn $|z| = r = 1$, $\cos(n\phi) = 1$ und $\sin(n\phi) = 0$. Aus den letzten zwei Gleichungen folgt $n\phi = 2k\pi$ mit $k \in \mathbb{Z}$. Also ist jede Lösung der Gleichung $z^n = 1$ von der Form $z = e^{2k\pi \mathrm{i}/n}$ mit $k \in \mathbb{Z}$, und aus $\phi \in [0, 2\pi)$ folgt $k \in \{0, 1, ..., n-1\}$. □

Komplexe Nullstellen eine Polynoms p werden allgemein als **Wurzeln** von p bezeichnet, und auch die Lösungen der Gleichung $z^n = c$ mit $c \in \mathbb{C}$ nennt man oft **Wurzeln** von c. Ein Ausdruck der Form $\sqrt[n]{c}$ hat aber nur dann einen Sinn, wenn c eine *nicht-negative reelle Zahl* ist - in diesem Fall bezeichnet es die eindeutig bestimmte *reelle*, *nicht-negative* Lösung der Gleichung $x^n = c$. Ein Ausdruck der Form $\sqrt[n]{c}$ für $c \in \mathbb{R}$ und $c < 0$ oder $c \in \mathbb{C} \setminus \mathbb{R}$ ist ohne weitere Konventionen oder Angaben, welche Lösung gemeint ist, unsinnig!



Die 5ten Einheitswurzeln.



Die fünf Lösungen der Gleichung $z^5 = w$.

Die $n$ten Einheitswurzeln bilden die Ecken eines in den Einheitskreis $S^1 = \{z \in \mathbb{C} | \, |z| = 1\}$ einbeschriebenen regulären n-Ecks, dessen erste Ecke auf der Zahl 1 liegt. Allgemeiner erhält man für jede komplexe Zahl w mit $|w| = 1$ eine Lösung der Gleichung $z^n = w$, indem man den Winkel zwischen w und der positiven x-Achse in n gleiche Teile teilt und den Punkt auf dem Einheitskreis markiert, der mit der positiven x-Achse den resultierenden Winkel einschließt. Die n Lösungen sind dann genau die Ecken eines in den Einheitskreis einbeschriebenen regulären n-Ecks, dessen erste Ecke auf diesem Punkt liegt. (Beweis: Übung).

Die $n$ten Einheitswurzeln bilden mit der Multiplikation eine Gruppe - multipliziert man zwei $n$te Einheitswurzeln, so erhält man wieder eine $n$te Einheitswurzel. Wir zeigen nun, dass diese Gruppe isomorph zur Restklassengruppe $\mathbb{Z}/n\mathbb{Z}$ ist. Insbesondere können wir dann mit den $n$ten Einheitswurzeln die Gruppe $\mathbb{Z}/n\mathbb{Z}$ visualisieren.

Satz 2.3.25: Für alle $n \in \mathbb{N}$ ist die Menge $C_n = \{z \in \mathbb{C} | z^n = 1\}$ der $n$ten Einheitswurzeln eine Untergruppe der Gruppe $(\mathbb{C} \setminus \{0\}, \cdot)$. Die Gruppe $(C_n, \cdot)$ ist isomorph zu $(\mathbb{Z}/n\mathbb{Z}, +)$.

Beweis:
Offensichtlich ist $C_n \neq \emptyset$, da $1 \in C_n$ (UG1). Sind $z, w \in C_n$, so folgt $(z \cdot w)^n = z^n \cdot w^n = 1 \cdot 1 = 1$ und somit $z \cdot w \in C_n$ (UG2) und $(z^{-1})^n = (z^n)^{-1} = 1^{-1} = 1$, also $z^{-1} \in C_n$ (UG3). Damit ist $C_n \subseteq \mathbb{C} \setminus \{0\}$ eine Untergruppe von $(\mathbb{C} \setminus \{0\}, \cdot)$.

Nach Lemma 2.2.10 und 2.3.24 gilt

$$C_n = \left\{ e^{2\pi i k/n} | k \in \{0, 1, ..., n-1\} \right\} \qquad \mathbb{Z}/n\mathbb{Z} = \{[k] | k \in \{0, 1, ..., n-1\}\}.$$

Die Abbildung $f : C_n \to \mathbb{Z}/n\mathbb{Z}$, $e^{2\pi i k/n} \mapsto [k]$ ist wohldefiniert, denn aus $e^{2\pi i k/n} = e^{2\pi i l/n}$ folgt $k - l \in n\mathbb{Z}$ und damit auch $[k] = [l]$.

Sie ist ist injektiv, denn aus $f(e^{2\pi i k/n}) = [k] = [l] = f(e^{2\pi i l/n})$ folgt $k - l \in n\mathbb{Z}$ und damit $e^{2\pi i k/n} = e^{2\pi i l/n}$. Da die Mengen C_n und $\mathbb{Z}/n\mathbb{Z}$ gleich viele Elemente haben, ist f nach Bemerkung 2.1.24 bijektiv. Für alle $k, l \in \mathbb{Z}$ gilt

$$f(e^{2\pi i k/n} \cdot e^{2\pi i l/n}) = f(e^{2\pi i (k+l)/n}) = [k + l] = [k] + [l] = f(e^{2\pi i k/n}) + f(e^{2\pi i l/n}).$$

Also ist f ein Gruppenisomorphismus. $\square$

Bemerkung 2.3.26: Das sichere Rechnen mit trigonometrischen Funktionen und komplexen Zahlen wird im gesamten Studium vorausgesetzt. Dazu gehört es auch, dass man die wichtigsten Werte von Sinus und Kosinus ohne Hilfsmittel jederzeit reproduzieren kann.

Füllen Sie dazu die folgende Tabelle aus prägen Sie sich diese Werte des Sinus und Kosinus ein. Zeichnen Sie dann die zugehörigen Punkte $e^{\mathrm{i}\alpha} = \cos(\alpha) + \mathrm{i}\sin(\alpha)$ in der komplexen Ebene.

α	0	$\frac{\pi}{4}$	$\frac{\pi}{2}$	$\frac{3\pi}{4}$	π	$\frac{5\pi}{4}$	$\frac{3\pi}{2}$	$\frac{7\pi}{4}$	2π	$\frac{\pi}{6}$	$\frac{\pi}{3}$	$\frac{2\pi}{3}$	$\frac{5\pi}{6}$	$\frac{7\pi}{6}$	$\frac{4\pi}{3}$	$\frac{5\pi}{3}$	$\frac{11\pi}{6}$
$\sin(\alpha)$																	
$\cos(\alpha)$																	
$\tan(\alpha)$																	

Sie sollten auch wissen, dass $\sin(\alpha)^2 + \cos(\alpha)^2 = 1$ für alle $\alpha \in \mathbb{R}$ gilt, und die Additionsformeln für Sinus und Kosinus kennen:

$$\sin(\alpha \pm \beta) = \sin(\alpha)\cos(\beta) \pm \cos(\alpha)\sin(\beta) \qquad \cos(\alpha \pm \beta) = \cos(\alpha)\cos(\beta) \mp \sin(\alpha)\sin(\beta).$$

Es ist eine hilfreiche Übung, sich daraus entsprechende Additionsformeln für den Tangens und Kotangens herzuleiten.

An was man sich erinnern sollte:

Die wichtigsten Begriffe und Definitionen:

- Menge, (Äquivalenz)relation, Äquivalenzklasse, Abbildung,
- Definitionsbereich, Wertebereich, Bild, Urbild, injektiv, surjektiv, bijektiv
- Verknüpfung, kommutativ, assoziativ, neutrales Element, Inverse, Distributivgesetz
- Monoid, Gruppe, abelsche Gruppe, Untergruppe
- Ring, unitaler Ring, Körper, Teilring, Teilkörper
- Homo/iso/endo/automorphismus von Gruppen, (unitalen) Ringen und Körpern

Die wichtigsten Beispiele:

- **Gruppen:** triviale Gruppe, $(\mathbb{Z},+)$, $(\mathbb{Q},+)$, $(\mathbb{R},+)$, $(\mathbb{Q}\setminus\{0\},\cdot)$, $(\mathbb{R}\setminus\{0\},\cdot)$, $(\mathbb{Z}/n\mathbb{Z},+)$, $(\mathrm{Bij}(X,X),\circ)$, Permutationsgruppen, symmetrische Gruppe S_n, Einheitengruppe eines Monoids
- **Ringe:** Nullring, $(\mathbb{Z},+,\cdot)$, Restklassenring $(\mathbb{Z}/n\mathbb{Z},+,\cdot)$, Polynomring $R[x]$, $\mathrm{Abb}(M,R)$
- **Körper:** $(\mathbb{R},+,\cdot)$, $(\mathbb{Q},+,\cdot)$, $(\mathbb{C},+,\cdot)$, endliche Körper $(\mathbb{Z}/p\mathbb{Z},+,\cdot)$ mit $p\in\mathbb{N}$ prim
- **Homomorphismen von Gruppen, (unitalen) Ringen und Körpern :** Identitätsabbildung, $\mu_m:\mathbb{Z}\to\mathbb{Z}$, $z\mapsto mz$ für $m\in\mathbb{N}$

Die wichtigsten Aussagen:

- Beziehung zwischen Äquivalenzrelationen und Partitionen,
- Kriterien für Injektivität, Surjektivität und Bijektivität von Abbildungen, Verhalten unter Verkettung
- Kerne und Bilder von Gruppenhomomorphismen sind Untergruppen,
- Die Verkettung von Gruppen oder Ringhomomorphismen ist ein Gruppen bzw. Ringhomomorphismus,
- $\mathbb{Z}/n\mathbb{Z}$ ist ein Körper genau dann, wenn n eine Primzahl ist.
- Fundamentalsatz der Algebra.

3 Vektorräume

3.1 Der Begriff des Vektorraums

Die in Kapitel 2 eingeführten algebraischen Grundbegriffe (Mengen, Abbildungen, Äquivalenzrelationen, Verknüpfungen, Gruppen, Ringe, Körper,...) sind die Voraussetzung, um über das eigentliche Thema dieser Vorlesung - *Vektorräume* und *lineare Abbildungen* - überhaupt sinnvoll sprechen zu können. Im Folgenden werden wir sie voraussetzen und benutzen.

Wir konzentrieren uns aber ab jetzt auf eine einzige algebraische Struktur, den *Vektorraum* und die zugehörigen strukturerhaltenden Abbildungen, die *linearen Abbildungen*. Dabei beschäftigen wir uns zunächst allgemein mit dem Begriff eines Vektorraums und lernen Beispiele von Vektorräumen kennen. Anschliessend werden die Eigenschaften von Vektorräumen und die zugehörigen strukturerhaltenenden Abbildungen genauer betrachtet.

Vektorräume und lineare Abbildungen sind in der Mathematik, der Physik, in anderen Naturwissenschaften und in den Ingenieurwissenschaften allgegenwärtig und eines der mächtigsten und bestverstandenen mathematischen Werkzeuge. Stellt man bei einem konkreten Problem fest, dass es sich als eine Frage zu Vektorräumen und linearen Abbildungen formulieren lässt, hat man es schon fast gelöst.

Definition 3.1.1: Sei $\mathbb{K}$ ein Körper. Ein **Vektorraum** über $\mathbb{K}$ oder ein $\mathbb{K}$-**Vektorraum** ist eine abelsche Gruppe $(V,+)$ zusammen mit einer Abbildung $\cdot : \mathbb{K} \times V \to V$, die **Skalarmultiplikation**, so dass die folgenden Bedingungen erfüllt sind:

(V1) **Pseudoassoziativität:** $\lambda \cdot (\mu \cdot v) = (\lambda\mu) \cdot v$ für alle $\lambda, \mu \in \mathbb{K}$, $v \in V$,
(V2) **Skalares Distributivgesetz:** $(\lambda + \mu) \cdot v = (\lambda \cdot v) + (\mu \cdot v)$ für alle $\lambda, \mu \in \mathbb{K}$, $v \in V$,
(V3) **Vektorielles Distributivgesetz:** $\lambda \cdot (v + w) = (\lambda \cdot v) + (\lambda \cdot w)$ für alle $\lambda \in \mathbb{K}$, $v, w \in V$,
(V4) **Normierung:** $1_\mathbb{K} \cdot v = v$ für alle $v \in V$.

Die Elemente eines Vektorraums nennt man **Vektoren**. Die Verknüpfung $+$ wird oft **Vektoraddition** genannt. Das neutrale Element von $(V,+)$ wird mit 0_V bezeichnet und heißt **Nullvektor**. Ein Vektorraum über dem Körper $\mathbb{R}$ heißt auch **reeller Vektorraum**, ein Vektorraum über dem Körper $\mathbb{C}$ **komplexer Vektorraum**.

Bemerkung 3.1.2: Auch in Vektorräumen gilt die Konvention *Punkt vor Strich*, und die Punkte für die Skalarmultiplikation und die Multiplikation in $\mathbb{K}$ werden oft weggelassen. Wegen der Pseudoassoziativität schreibt man für $\lambda, \mu \in \mathbb{K}$ und $v \in V$ ausserdem auch oft $\lambda\mu\, v$, ohne Klammern zu setzen. Sowohl das neutrale Element 0_V von $(V,+)$ als auch das neutrale Element $0_\mathbb{K}$ von $(\mathbb{K},+)$ werden oft einfach mit 0 bezeichnet.

Lemma 3.1.3 (Rechenregeln für Vektorräume):
Sei V ein $\mathbb{K}$-Vektorraum und $u, v, w \in V$, $\lambda \in \mathbb{K}$ beliebig. Dann gilt:

(i) $u + v = u + w \Rightarrow v = w$.

(ii) $0_\mathbb{K} \cdot v = 0_V$ und $\lambda \cdot 0_V = 0_V$.

(iii) $(-1_\mathbb{K}) \cdot u = -u$.

(iv) $\lambda \cdot v = 0_V \Rightarrow (\lambda = 0_\mathbb{K}) \vee (v = 0_V)$.

Beweis:
Da $(V, +)$ eine abelsche Gruppe ist, folgt (i) aus Lemma 2.2.13, 4. Für (ii) berechnet man

$$0_\mathbb{K} \cdot v + 0_V \overset{(G2)}{=} 0_\mathbb{K} \cdot v \overset{(G2)}{=} (0_\mathbb{K} + 0_\mathbb{K}) \cdot v \overset{(V2)}{=} 0_\mathbb{K} \cdot v + 0_\mathbb{K} \cdot v \overset{(i)}{\Rightarrow} 0_\mathbb{K} \cdot v = 0_V \qquad \forall v \in V$$
$$\lambda \cdot 0_V + 0_V \overset{(G2)}{=} \lambda \cdot 0_V \overset{(G2)}{=} \lambda \cdot (0_V + 0_V) \overset{(V3)}{=} \lambda \cdot 0_V + \lambda \cdot 0_V \overset{(i)}{\Rightarrow} \lambda \cdot 0_V = 0_V \qquad \forall \lambda \in \mathbb{K}.$$

Ebenso ergibt sich (iii) direkt aus der folgenden Rechnung

$$(-1_\mathbb{K})\cdot u + u \overset{(V4)}{=} (-1_\mathbb{K})\cdot u + 1_\mathbb{K}\cdot u \overset{(V2)}{=} (-1_\mathbb{K}+1_\mathbb{K})\cdot u \overset{(K5)}{=} 0_\mathbb{K}\cdot u \overset{(ii)}{=} 0_V \overset{(G3)}{=} -u+u \overset{(i)}{\Rightarrow} -u = (-1_\mathbb{K})\cdot u.$$

Zum Beweis von (iv) betrachtet man $\lambda \in \mathbb{K}$ und $v \in V$ mit $\lambda \cdot v = 0_V$. Dann gilt entweder $\lambda = 0_\mathbb{K}$ oder $\lambda \neq 0_\mathbb{K}$. Im ersten Fall ist die Aussage wahr. Im zweiten gibt es ein multiplikatives Inverses $\lambda^{-1} \in \mathbb{K}$, und es folgt

$$v \overset{(V4)}{=} 1_\mathbb{K} \cdot v \overset{(G3)}{=} (\lambda^{-1}\lambda) \cdot v \overset{(V1)}{=} \lambda^{-1} \cdot (\lambda \cdot v) = \lambda^{-1} \cdot 0_V \overset{(ii)}{=} 0_V. \qquad \square$$

Beispiel 3.1.4:

1. Jede einelementige Menge $M = \{m\}$ ist ein Vektorraum über jedem beliebigen Körper $\mathbb{K}$ mit $+ : M \times M \to M$, $(m, m) \mapsto m$ und $\cdot : \mathbb{K} \times M \to M$, $(\lambda, m) \mapsto m$. Er besteht nur aus dem Nullvektor $0_M = m$ und heißt **Nullvektorraum** oder **trivialer Vektorraum**.

2. Sind $(V, +_V, \cdot_V)$ und $(W, +_W, \cdot_W)$ $\mathbb{K}$-Vektorräume, so hat auch die Menge $V \times W$ die Struktur eines $\mathbb{K}$-Vektorraums mit der Vektoraddition und Skalarmultiplikation

   $$(v, w)+(v', w') := (v+_V v', w+_W w'),\ \lambda\cdot(v, w) := (\lambda\cdot_V v, \lambda\cdot_W w)\ \ \forall v, v' \in V, w, w' \in W, \lambda \in \mathbb{K}.$$

   Er heißt **Produkt** der Vektorräume $(V, +_V, \cdot_V)$ und $(W, +_W, \cdot_W)$.

   Denn $(V \times W, +)$ ist das direkte Produkt der abelschen Gruppen $(V, +_V)$ und $(W, +_W)$ und somit eine abelsche Gruppe nach Beispiel 2.2.9, 6. Die Axiome (V1)-(V4) folgen aus den entsprechenden Axiomen für $(V, +_V, \cdot_V)$ und $(W, +_W, \cdot_W)$ (Übung).

3. Jeder Körper $\mathbb{K}$ hat die Struktur eines Vektorraums über sich selbst mit der Körperaddition und der Körpermultiplikation als Vektoraddition und Skalarmultiplikation.

   Denn $(\mathbb{K}, +)$ ist eine abelsche Gruppe, (V1) gilt wegen der Assoziativität der Körpermultiplikation, (V2), (V3) wegen des Distributivgesetzes in $\mathbb{K}$ und (V4), da 1 das neutrale Element von $(\mathbb{K} \setminus \{0\}, \cdot)$ ist.

4. Ist $\mathbb{K} \subseteq \mathbb{L}$ ein Teilkörper, so ist $\mathbb{L}$ mit der Körperaddition $+ : \mathbb{L} \times \mathbb{L} \to \mathbb{L}$ und der Einschränkung $\cdot : \mathbb{K} \times \mathbb{L} \to \mathbb{L}$ der Körpermultiplikation als Skalarmultiplikation ein Vektorraum über dem Körper $\mathbb{K}$.

   Denn da $(\mathbb{L}, +, \cdot)$ ein Körper ist, ist $(\mathbb{L}, +)$ eine abelsche Gruppe. Aus der Assoziativität der Multiplikation in $\mathbb{L}$ folgt die Pseudoassoziativität (V1) $k \cdot (k' \cdot l) = k \cdot k' \cdot l = (k \cdot k') \cdot l$ für alle $l \in \mathbb{L}$ und $k, k' \in \mathbb{K} \subseteq \mathbb{L}$. Aus dem Distributivgesetz (K3) in $\mathbb{L}$ folgen das skalare und vektorielle Distributivgesetz (V2) und (V3): für alle $l, l' \in \mathbb{L}$ und $k, k' \in \mathbb{K} \subseteq \mathbb{L}$ gilt $(k + k') \cdot l = k \cdot l + k' \cdot l$ und $k \cdot (l + l') = k \cdot l + k \cdot l'$. Aus der Tatsache, dass $1 = 1_\mathbb{L} = 1_\mathbb{K}$ das neutrale Element von $(\mathbb{L}, \cdot)$ ist, folgt die Normierung (V4): Für alle $l \in \mathbb{L}$ gilt $1 \cdot l = l$.

5. Aus 4. folgt: $\mathbb{C}$ hat die Struktur eines reellen Vektorraums. $\mathbb{R}$ hat die Struktur eines $\mathbb{Q}$-Vektorraums. Nach 3. sind $\mathbb{C}$ und $\mathbb{R}$ aber auch jeweils Vektorräume über sich selbst.

Besonders wichtige Beispiele von $\mathbb{K}$-Vektorräumen sind die $\mathbb{K}$-Vektorräume $\mathbb{K}^n = \mathbb{K} \times \ldots \times \mathbb{K}$ aus Beispiel 3.1.4, 2. für beliebige Körper $\mathbb{K}$. Diese sind einerseits deswegen wichtig, weil sie es uns erlauben, die Geometrie des zwei- und dreidimensionalen Raums durch die Vektorräume $\mathbb{R}^2$ und $\mathbb{R}^3$ zu beschreiben. Andererseits werden wir später sehen, dass sich sehr viele Vektorräume über einem Körper $\mathbb{K}$ durch Vektorräume $\mathbb{K}^n$ mit $n \in \mathbb{N}$ beschreiben lassen.

Beispiel 3.1.5: Für jeden Körper $\mathbb{K}$ und jedes $n \in \mathbb{N}$ hat die Menge $\mathbb{K}^n = \mathbb{K} \times ... \times \mathbb{K}$ die Struktur eines Vektorraums über $\mathbb{K}$ mit der Vektoraddition und Skalarmultiplikation

$$(a_1, ..., a_n) + (b_1, ..., b_n) := (a_1 + b_1, ..., a_n + b_n) \qquad \lambda \cdot (a_1, ..., a_n) := (\lambda a_1, ..., \lambda a_n)$$

für alle $\lambda, a_1, b_1, ..., a_n, b_n \in \mathbb{K}$. Der Nullvektor ist gegeben durch $(0, ..., 0)$. (Beweis: Übung)

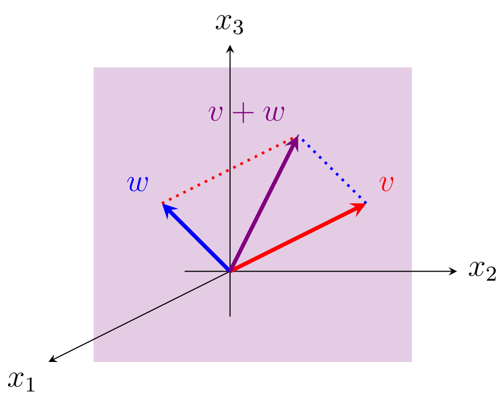
Es wird sich später als sinnvoll herausstellen, statt mit waagrechten Zahlentupeln (**Zeilenvektoren**) mit senkrechten Zahlentupeln (**Spaltenvektoren**) zu arbeiten. Dabei handelt es sich einfach um eine andere Schreibweise für n-Tupel im $\mathbb{K}^n$. Die Vektoraddition und Skalarmultiplikation sind dann gegeben durch

$$\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} + \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} a_1 + b_1 \\ \vdots \\ a_n + b_n \end{pmatrix} \qquad \lambda \cdot \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} \lambda a_1 \\ \vdots \\ \lambda a_n \end{pmatrix}.$$

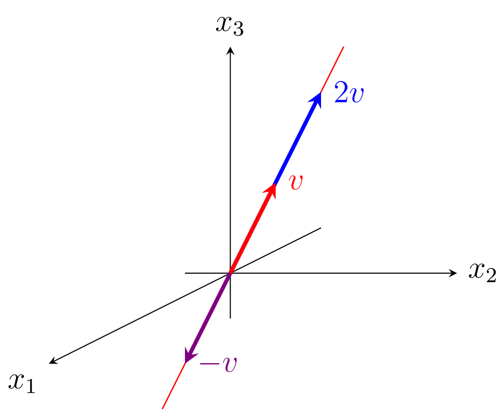
Im Text deuten wir durch die Bezeichnung $(a_1, ..., a_n)^T$ an, dass ein Spaltenvektor gemeint ist:

$$(a_1, ..., a_n)^T := \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$$

Ein Vektor $v = (a_1, ..., a_n)^T$ im $\mathbb{R}^n$ wird durch einen Pfeil vom Ursprung $(0, ..., 0) \in \mathbb{R}^n$ zu dem Punkt mit Koordinaten $(a_1, ..., a_n)$ beschrieben. Die Vektoraddition im $\mathbb{R}^n$ entspricht geometrisch dem *Aneinanderhängen von Pfeilen* (Parallelogrammregel) und die Skalarmultiplikation dem *Strecken* ($\lambda > 1$) bzw. *Stauchen* ($0 \leq \lambda < 1$) von Pfeilen oder einer Kombination von Strecken oder Stauchen mit einer *Punktspiegelung am Ursprung* ($\lambda < 0$).



Die Vektoraddition im $\mathbb{R}^3$



Die Skalarmultiplikation im $\mathbb{R}^3$

Wir können ein geordnetes n-Tupel $(a_1, ..., a_n)^T \in \mathbb{K}^n$ von Zahlen $a_1, ..., a_n \in \mathbb{K}$ auch als eine Abbildung $f : \{1, ..., n\} \to \mathbb{K}$, $k \mapsto a_k$ auffassen. Die Vektoraddition und Skalarmultiplikation im Vektorraum $\mathbb{K}^n$ entspricht dann gerade der punktweisen Addition solcher Abbildungen und ihrer punktweisen Multiplikation mit Elementen in $\mathbb{K}$. Für $f : \{1, ..., n\} \to \mathbb{K}$, $k \mapsto a_k$ und $g : \{1, ..., n\} \to \mathbb{K}$, $k \mapsto b_k$ erhält man so nämlich $(f + g)(k) = f(k) + g(k) = a_k + b_k$ und $(\lambda \cdot f)(k) = \lambda \cdot f(k) = \lambda a_k$ für alle $k \in \{1, ..., n\}$.

Dies lässt sich auf Abbildungen $f : M \to V$ von einer beliebigen Menge M in einen beliebigen $\mathbb{K}$-Vektorraum V verallgemeinern.

Lemma 3.1.6: Für jede Menge M und jeden $\mathbb{K}$-Vektorraum $(V, +_V, \cdot_V)$ hat die Menge $\mathrm{Abb}(M, V)$ der Abbildungen $f : M \to V$ die Struktur eines $\mathbb{K}$-Vektorraums mit der Vektoraddition und Skalarmultiplikation

$$(f + g)(m) := f(m) +_V g(m) \qquad (\lambda \cdot f)(m) := \lambda \cdot_V f(m) \qquad \forall m \in M.$$

Beweis:
Ist $M = \emptyset$, so gibt es genau eine Abbildung $f : \emptyset \to V$, die leere Abbildung, und es gilt $\mathrm{Abb}(M, V) = \{0\}$. Sei also jetzt $M \neq \emptyset$. In Beispiel 2.3.4, 4. wurde bereits gezeigt, dass für jede abelsche Gruppe $(V, +_V)$ und jede Menge M auch $(\mathrm{Abb}(M, V), +)$ mit der punktweisen Addition eine abelsche Gruppe ist. Um die Vektorraumaxiome (V1) bis (V4) nachzurechnen betrachten wir $f, g \in \mathrm{Abb}(M, V)$ und $\lambda, \mu \in \mathbb{K}$.

(V1) Aus der Pseudoassoziativität in V folgt

$$(\lambda \cdot (\mu \cdot f))(m) = \lambda \cdot_V ((\mu \cdot f)(m)) = \lambda \cdot_V (\mu \cdot_V f(m)) = (\lambda\mu) \cdot_V f(m) = ((\lambda\mu) \cdot f)(m)$$

für alle $m \in M$, also $\lambda \cdot (\mu \cdot f) = (\lambda\mu) \cdot f$.

(V2) Aus dem skalaren Distributivgesetz für V folgt

$$((\lambda + \mu) \cdot f)(m) = (\lambda + \mu) \cdot_V f(m) = \lambda \cdot_V f(m) +_V \mu \cdot_V f(m) = (\lambda \cdot f + \mu \cdot f)(m)$$

für alle $m \in M$ und damit $(\lambda + \mu) \cdot f = \lambda \cdot f + \mu \cdot f$.

(V3) Aus dem vektoriellen Distributivgesetz für V ergibt sich für V für alle $m \in M$

$$\begin{aligned}(\lambda \cdot (f + g))(m) &= \lambda \cdot_V (f + g)(m) = \lambda \cdot_V (f(m) +_V g(m)) = \lambda \cdot_V f(m) +_V \lambda \cdot_V g(m) \\ &= (\lambda \cdot f + \lambda \cdot g)(m),\end{aligned}$$

also $(\lambda + \mu) \cdot f = \lambda \cdot f + \mu \cdot f$.

(V4) Aus der Normierung von V erhält man für alle $m \in M$ die Identität $(1 \cdot f)(m) = 1 \cdot_V f(m) = f(m)$ und somit $1 \cdot f = f$. □

Als Spezialfälle dieses Beispiels ergeben sich viele weitere nützliche Beispiele von Vektorräumen.

Beispiel 3.1.7:

1. Für jede Menge M ist die Menge $\mathrm{Abb}(M, \mathbb{K})$ mit der Vektoraddition und Skalarmultiplikation aus Beispiel 3.1.6 ein $\mathbb{K}$-Vektorraum.

2. Die Abbildungen $f : \mathbb{R} \to \mathbb{R}$ bilden mit der punktweisen Addition und Skalarmultiplikation den reellen Vektorraum $\mathrm{Abb}(\mathbb{R}, \mathbb{R})$.

3. Für jeden $\mathbb{K}$-Vektorraum V ist die Menge der Folgen $(v_n)_{n\in\mathbb{N}_0}$ mit Werten in V ein $\mathbb{K}$-Vektorraum mit Vektoraddition und Skalarmultiplikation

$$(v_n)_{n\in\mathbb{N}_0} + (w_n)_{n\in\mathbb{N}_0} = (v_n + w_n)_{n\in\mathbb{N}_0} \qquad \lambda \cdot (x_n)_{n\in\mathbb{N}_0} = (\lambda x_n)_{n\in\mathbb{N}_0}.$$

   Denn eine Folge $(v_n)_{n\in\mathbb{N}_0}$ mit Werten in V ist nichts anderes als eine Abbildung $f : \mathbb{N}_0 \to V$, $n \mapsto v_n$, und die angegebene Vektoraddition und Skalarmultiplikation sind gerade die aus aus Beispiel 3.1.6. Also handelt es sich um den Vektorraum $\mathrm{Abb}(\mathbb{N}_0, V)$.

4. Für $V = \mathbb{K}$ erhält man aus 3. , dass die $\mathbb{K}$-wertigen Folgen mit der Vektoraddition und Skalarmultiplikation aus 3. den $\mathbb{K}$-Vektorraum $\mathrm{Abb}(\mathbb{N}_0, \mathbb{K})$ bilden. Für reelle Folgen erhält man den $\mathbb{R}$-Vektorraum $\mathrm{Abb}(\mathbb{N}_0, \mathbb{R})$, für rationale Folgen den $\mathbb{Q}$-Vektorraum $\mathrm{Abb}(\mathbb{N}_0, \mathbb{Q})$.

Bei Gruppen und Ringen folgten auf die Definition und ersten Beispiele der Begriff einer Untergruppe oder eines Unterrings. Dabei handelte es sich um Teilmengen der Gruppe oder des Rings, die mit der Einschränkungen der Verknüpfung(en) wieder eine Gruppe oder einen Ring bilden. Für Vektorräume definiert man analog den Begriff des Untervektorraums. Dabei werden Einschränkungen der Vektoraddition und Skalarmultiplikation betrachtet.

Definition 3.1.8: Eine Teilmenge $U \subseteq V$ eines $\mathbb{K}$-Vektorraums V heißt **Untervektorraum** von V oder **(linearer) Unterraum** von V wenn sie abgeschlossen unter Vektoraddition und Skalarmultiplikation ist und mit den Einschränkungen der Verknüpfung $+ : V \times V \to V$ und der Abbildung $\cdot : \mathbb{K} \times V \to V$ einen $\mathbb{K}$-Vektorraum bildet.

Aus der Definition geht hervor, dass ein Untervektorraum eines $\mathbb{K}$-Vektorraums V immer ein Vektorraum über dem Körper $\mathbb{K}$ ist. Mit Hilfe der folgenden Kriterien lässt sich leicht nachrechnen, ob es sich bei einer gegebenen Teilmenge $M \subseteq V$ eines $\mathbb{K}$-Vektorraums V um einen Untervektorraum handelt. Sie ähneln den Kriterien für Gruppen und Ringe, nur dass entsprechend die Vektoraddition und die Skalarmultiplikation des Vektorraums auftreten.

Lemma 3.1.9: Eine Teilmenge $U \subseteq V$ ist genau dann ein Untervektorraum eines $\mathbb{K}$-Vektorraums V, wenn:

(UV1) $0 \in U$,
(UV2) $u, v \in U \Rightarrow u + v \in U$,
(UV3) $u \in U$, $\lambda \in \mathbb{K} \Rightarrow \lambda \cdot u \in U$.

Beweis:
$\Rightarrow$: Ist $U \subseteq V$ ein Untervektorraum von V, so ist U abgeschlossen unter der Vektoraddition und Skalarmultiplikation, also $u + v \in U$ für alle $u, v \in U$ (UV2) und $\lambda u \in U$ für alle $u \in U$, $\lambda \in \mathbb{K}$ (UV3). Ausserdem ist dann $(U, +)$ eine Untergruppe der abelschen Gruppe $(V, +)$ und enthält somit nach Definition 2.2.16 das neutrale Element: $0_V \in U$ (UV1).

$\Leftarrow$: Ist $U \subseteq V$ eine Teilmenge, die (UV1) und (UV2) erfüllt, so ist $(U, +)$ ein kommutatives Monoid. Nach (UV3) ist aber für jedes $u \in U$ auch $(-1) \cdot u \in U$ und nach Lemma 3.1.3 gilt $(-1) \cdot u = -u$. Also ist $(U, +)$ eine Untergruppe von $(V, +)$ und damit eine abelsche Gruppe. Wegen (UV3) ist die Einschränkung der Skalarmultiplikation auf U eine Abbildung $\cdot : \mathbb{K} \times U \to U$, und die Axiome (V1)-(V4) für U folgen aus den entsprechenden Axiomen für V. $\square$

Beispiel 3.1.10:

1. Für jeden $\mathbb{K}$-Vektorraum V sind $\{0_V\} \subseteq V$ und $V \subseteq V$ Untervektorräume.

2. Für jeden $\mathbb{K}$-Vektorraum V und $v \in V$ ist $\mathbb{K}v = \{\alpha v | \, \alpha \in \mathbb{K}\} \subseteq V$ ein Untervektorraum.

   Denn es gilt $0_V = 0_{\mathbb{K}} \cdot v \in \mathbb{K}v$ (UV1). Aus $u, w \in \mathbb{K}v$ folgt, dass $\alpha, \beta \in \mathbb{K}$ existieren mit $u = \alpha v, w = \beta v$. Daraus ergibt sich $u + w = \alpha v + \beta v = (\alpha + \beta)v \in \mathbb{K}v$ (UV2). Ebenso erhält man $\lambda u = (\lambda\alpha)v \in \mathbb{K}v$ für alle $\lambda \in \mathbb{K}$ (UV3).

   Für $v = 0$ ist $\mathbb{K}v = \{0\}$. Für $v \neq 0$ bezeichnet man den Untervektorraum $\mathbb{K}v$ als **Gerade**. Im Vektorraum $\mathbb{R}^3$ ist dies nämlich nämlich eine Gerade durch den Ursprung.

3. Sei $\mathbb{K}$ ein Körper als Vektorraum über sich selbst. Dann sind die einzigen Untervektorräume von $\mathbb{K}$ der Nullvektorraum $\{0\}$ und $\mathbb{K}$ selbst.

   Denn enthält ein Untervektorraum $U \subseteq \mathbb{K}$ ein Element $v \in \mathbb{K} \setminus \{0\}$, so besitzt v ein multiplikatives Inverses v^{-1} und jedes Element $w \in \mathbb{K}$ lässt sich schreiben als $w = (wv^{-1}) \cdot v$ mit $wv^{-1} \in \mathbb{K}$, liegt also dann nach (UV3) in U.

4. Für jeden Körper $\mathbb{K}$ bilden die Polynomabbildungen $p : \mathbb{K} \to \mathbb{K}$, $x \mapsto \Sigma_{k=0}^n a_k x^k$ mit $n \in \mathbb{N}_0$, $a_k \in \mathbb{K}$ ein Untervektorraum $\mathrm{Pol}(\mathbb{K})$ des $\mathbb{K}$-Vektorraums $\mathrm{Abb}(\mathbb{K}, \mathbb{K})$.

   Denn die Nullabbildung $0 : \mathbb{K} \to \mathbb{K}$, $x \mapsto 0$ ist ein Polynomabbildung (UV1). Die Summe $p+q : \mathbb{K} \to \mathbb{K}$ zweier Polynomabbildungen $p, q : \mathbb{K} \to \mathbb{K}$ ist eine Polynomabbildung (UV2) und für alle Polynomabbildungen $p : \mathbb{K} \to \mathbb{K}$ und $\lambda \in \mathbb{K}$ ist die Abbildung $\lambda p : \mathbb{K} \to \mathbb{K}$ eine Polynomabbildung (UV3).

5. Für jeden Körper $\mathbb{K}$ sind die Polynome mit Koeffizienten in $\mathbb{K}$ ein Untervektorraum des $\mathbb{K}$-Vektorraums der $\mathbb{K}$-wertigen Folgen.

   Denn eine Folge ist eine Abbildung $f : \mathbb{N}_0 \to \mathbb{K}$, und ein Polynom ist eine Abbildung $p : \mathbb{N}_0 \to \mathbb{K}$ mit $p(k) = 0$ für fast alle $k \in \mathbb{N}_0$. Die Nullabbildung $0 : \mathbb{N}_0 \to \mathbb{K}$, $k \mapsto 0$ ist ein Polynom (UV1). Sind $p, q : \mathbb{N}_0 \to \mathbb{K}$ Polynome, so sind auch $p + q : \mathbb{N}_0 \to \mathbb{K}$ (UV2) und $\lambda p : \mathbb{N}_0 \to \mathbb{K}$ (UV3) Polynome, denn aus $p(k) = 0$ und $q(k) = 0$ für fast alle $k \in \mathbb{N}_0$ folgt $p(k) + q(k) = 0$ für fast alle $k \in \mathbb{N}_0$ und $\lambda p(k) = 0$ für fast alle $k \in \mathbb{N}_0$ und alle $\lambda \in \mathbb{K}$.

Im Fall von Gruppen wurde in einer Übungsaufgabe gezeigt, dass der Schnitt zweier Untergruppen einer gegebenen Gruppe G immer eine Untergruppe ist. Ein ähnliches Argument zeigt, dass der Schnitt von Untervektorräumen immer ein Untervektorraum ist.

Satz 3.1.11: Sei I eine Indexmenge und $(U_i)_{i \in I}$ eine Familie von Untervektorräumen eines $\mathbb{K}$-Vektorraums V. Dann ist auch die Menge $\cap_{i \in I} U_i$ ein Untervektorraum von V.

Beweis:
Nach (UV1) ist $0_V \in U_i$ für alle $i \in I$ und somit auch $0_V \in \cap_{i \in I} U_i$. Sind $u, v \in \cap_{i \in I} U_i$, so gilt $u, v \in U_i$ für alle $i \in I$. Mit (UV2) folgt $u + v \in U_i$ für alle $i \in I$ und somit $u + v \in \cap_{i \in I} U_i$. Mit (UV3) folgt $\lambda u \in U_i$ für alle $\lambda \in \mathbb{K}$ und $i \in I$ und somit $\lambda u \in \cap_{i \in I} U_i$. □

Die Vereinigung $U_1 \cup U_2$ von Untervektorräumen $U_1, U_2 \subseteq V$ eines $\mathbb{K}$-Vektorraums V ist im Allgemeinen kein Untervektorraum von V. Dies ist nur dann der Fall, wenn $U_1 \subseteq U_2$ oder $U_2 \subseteq U_1$ gilt (Übung). Man kann aber aus einer Vereinigung $U_1 \cup U_2$ einen Unterraum von V konstruieren, indem man zu $U_1 \cup U_2$ alle Vektoren der Form $u_1 + u_2$ mit $u_1 \in U_1$, $u_2 \in U_2$ hinzufügt. Dies ist ein Spezialfall einer allgemeineren Konstruktion, nämlich des sogenannten *Spanns* oder der *lineare Hülle* einer Teilmenge $M \subseteq V$.

Definition 3.1.12: Sei V ein $\mathbb{K}$-Vektorraum und $n \in \mathbb{N}$.

1. Eine **Linearkombination** von Vektoren $v_1, ..., v_n \in V$ ist ein Vektor der Form

$$\Sigma_{k=1}^n \lambda_k v_k := \lambda_1 v_1 + \lambda_2 v_2 + ... + \lambda_n v_n \qquad \text{mit} \qquad \lambda_1, ..., \lambda_n \in \mathbb{K}.$$

2. Für eine Teilmenge $M \subseteq V$ definiert man den **Spann** oder die **lineare Hülle** von M als

$$\begin{aligned}\mathrm{span}_{\mathbb{K}}(M) &= \{\Sigma_{k=1}^n \lambda_k v_k \mid n \in \mathbb{N}, v_1, ..., v_n \in M, \lambda_1, ..., \lambda_n \in \mathbb{K}\} \text{ für } M \neq \emptyset,\\ \mathrm{span}_{\mathbb{K}}(\emptyset) &= \{0\}.\end{aligned}$$

Man beachte, dass in Definition 3.1.12 nicht gefordert wird, dass die Vektoren $v_1, ..., v_n \in V$ paarweise verschieden sind und auch keine Einschränkungen an die natürliche Zahl $n \in \mathbb{N}$ auferlegt werden. In einer Linearkombination können also beliebig viele Vektoren auftreten, aber immer nur *endlich* viele.

Der Spann einer Teilmenge $M \subseteq V$ enthält genau die Elemente von V, die sich als Linearkombinationen von endlich vielen Vektoren aus M schreiben lassen. Insbesondere gilt für $v_1, ..., v_n \in V$ immer $\mathrm{span}_{\mathbb{K}}(\{v_1, ..., v_n\}) = \{\Sigma_{k=0}^n \lambda_k v_k | \lambda_1, ..., \lambda_n \in \mathbb{K}\}$. Denn jede Linearkombination endlich vieler Vektoren aus $\{v_1, ..., v_n\}$ lässt sich durch hinzufügen der darin nicht auftretenden Vektoren v_i mit Koeffizient 0 auch als eine Linearkombination $\Sigma_{k=0}^n \lambda_k v_k$ schreiben.

Wir zeigen nun, dass der Spann einer Teilmenge $M \subseteq V$ immer ein Untervektorraum von V ist. Dieser enthält offensichtlich die Menge $M \subseteq V$, denn für alle $m \in M$ ist $1 \cdot m$ eine Linearkombination von m und damit in $\mathrm{span}_{\mathbb{K}}(M)$ enthalten. Der Spann besitzt aber eine besondere Eigenschaft, die ihn von anderen Untervektorräumen $U \subseteq V$ mit $M \subseteq U$ unterscheidet. Er ist der *kleinste Untervektorraum* von V, der M enthält: Jeder Untervektorraum von V, der M als Teilmenge enthält, enthält auch $\mathrm{span}_{\mathbb{K}}(M)$ als Teilmenge. Damit ist $\mathrm{span}_{\mathbb{K}}(M)$ der Schnitt aller Untervektorräume von V, die M als Teilmenge enthalten.

Satz 3.1.13: Sei V ein $\mathbb{K}$-Vektorraum und $M \subseteq V$ eine beliebige Teilmenge. Dann gilt:

1. Die Teilmenge $\mathrm{span}_{\mathbb{K}}(M) \subseteq V$ ist ein Untervektorraum von V.
2. Ist $U \subseteq V$ ein Untervektorraum von V mit $M \subseteq U$, so folgt $\mathrm{span}_{\mathbb{K}}(M) \subseteq U$.
3. Der Untervektorraum $\mathrm{span}_{\mathbb{K}}(M)$ ist der Schnitt aller Untervektorräume, die M enthalten

$$\mathrm{span}_{\mathbb{K}}(M) = \bigcap_{\substack{U \subseteq V \text{ Untervektorraum} \\ M \subseteq U}} U.$$

Beweis:
1. Es gilt $\mathrm{span}_{\mathbb{K}}(\emptyset) = \{0\}$, und somit ist $\mathrm{span}_{\mathbb{K}}(\emptyset)$ ein Untervektorraum von V. Sei also jetzt $M \neq \emptyset$. Dann existiert ein $v \in M$ und somit gilt $0_V = 0 \cdot v \in \mathrm{span}_{\mathbb{K}}(M)$, denn $0 \cdot v$ ist eine Linearkombination von v (UV1). Sind $u, w \in \mathrm{span}_{\mathbb{K}}(M)$, so existieren Vektoren $u_1, ..., u_n, w_1, ..., w_m \in M$ und $\alpha_1, ..., \alpha_n, \beta_1, ..., \beta_m \in \mathbb{K}$ mit $u = \Sigma_{k=1}^n \alpha_k v_k$ und $w = \Sigma_{j=1}^m \beta_j w_j$. Dann ist auch $u + v = \Sigma_{k=1}^n \alpha_k u_k + \Sigma_{j=1}^m \beta_j w_j$ eine Linearkombination endlich vieler Vektoren aus M und somit $u + w \in \mathrm{span}_{\mathbb{K}}(M)$ (UV2). Ebenso ist $\lambda u = \Sigma_{k=1}^n (\lambda \alpha_k) u_k$ eine Linearkombination endlich vieler Vektoren aus M und damit $\lambda u \in \mathrm{span}_{\mathbb{K}}(M)$ für alle $\lambda \in \mathbb{K}$ (UV3).

2. Ist $U \subseteq V$ ein Untervektorraum von V mit $M \subseteq U$, so ist U mit der Einschränkung der Vektoraddition und Skalarmultiplikation ein Vektorraum mit $M \subseteq U$. Nach 1. ist $\mathrm{span}_{\mathbb{K}}(M)$ ein Untervektorraum von U.

3. Nach 1. ist $\mathrm{span}_{\mathbb{K}}(M)$ ein Untervektorraum von V, und es gilt $M \subseteq \mathrm{span}_{\mathbb{K}}(M)$, denn $1 \cdot m = m \in \mathrm{span}_{\mathbb{K}}(M)$ für alle $m \in M$. Also ist der Untervektorraum $\mathrm{span}_{\mathbb{K}}(M)$ am Schnitt der Unterräume $U \subseteq V$ mit $M \subseteq U$ beteiligt und daher

$$\mathrm{span}_{\mathbb{K}}(M) \supseteq \bigcap_{\substack{U \subseteq V \text{ Untervektorraum} \\ M \subseteq U}} U.$$

Für jeden Untervektorraum $U \subseteq V$ mit $M \subseteq U$ gilt nach 2. auch $\mathrm{span}_{\mathbb{K}}(M) \subseteq U$ und somit

$$\mathrm{span}_{\mathbb{K}}(M) \subseteq \bigcap_{\substack{U \subseteq V \text{ Untervektorraum} \\ M \subseteq U}} U.$$

□

Satz 3.1.13 erlaubt es uns, aus jeder Teilmenge $M \subseteq V$ eines $\mathbb{K}$-Vektorraums V einen minimalen Untervektorraum $\mathrm{span}_{\mathbb{K}}(M) \subseteq V$ zu *erzeugen*, der M als Teilmenge enthält. Ist $U \subseteq V$ ein Untervektorraum, so gilt $\mathrm{span}_{\mathbb{K}}(U) = U$, denn U ist am Schnitt der Untervektorräume, die U enthalten, beteiligt. Neben U selbst kann es aber noch andere Teilmengen $M \subseteq U$ mit $\mathrm{span}_{\mathbb{K}}(M) = U$ geben. Wir nennen solche Teilmengen M *Erzeugendensysteme* von U.

Definition 3.1.14: Sei V ein $\mathbb{K}$-Vektorraum. Gilt $U = \mathrm{span}_{\mathbb{K}}(M)$ für einen Untervektorraum $U \subseteq V$ und eine Teilmenge $M \subseteq V$, so sagt man, dass M den Untervektorraum U **erzeugt** und bezeichnet M als ein **Erzeugendensystem** von U.

Beispiel 3.1.15:

1. In jedem $\mathbb{K}$-Vektorraum V gilt $\mathbb{K}v = \mathrm{span}_{\mathbb{K}}(\{v\})$ für alle $v \in V$, $\mathrm{span}_{\mathbb{K}}(\{0\}) = \{0\}$ und $\mathrm{span}_{\mathbb{K}}(V) = V$.

2. Im reellen Vektorraum $\mathbb{R}^3$ ist $\mathrm{span}_{\mathbb{R}}(\{v\}) = \mathbb{K}v$ für $v \in \mathbb{R}^3 \setminus \{0\}$ die Gerade durch den Punkt v und den Ursprung. Für $v, w \in \mathbb{R}^3$ mit $v \notin \mathbb{R}w$ und $w \notin \mathbb{R}v$ ist $\mathrm{span}_{\mathbb{R}}(\{v, w\})$ die Ebene durch die Punkte v, w und den Ursprung.

3. Im Vektorraum der reellen Polynomabbildungen gilt für $p : \mathbb{R} \to \mathbb{R}$, $x \mapsto x^2 + 1$ und $q : \mathbb{R} \to \mathbb{R}$, $x \mapsto x^2 + x$
$$\mathrm{span}_{\mathbb{R}}(\{p, q\}) = \{f : \mathbb{R} \to \mathbb{R},\ x \mapsto (\alpha + \beta)x^2 + \beta x + \alpha \mid \alpha, \beta \in \mathbb{R}\}.$$

4. In jedem Körper $\mathbb{K}$ als Vektorraum über sich selbst, gilt $\mathrm{span}_{\mathbb{K}}(\{1\}) = \mathbb{K}$, denn jedes Element $k \in \mathbb{K}$ lässt sich als Linearkombination $k = k \cdot 1$ schreiben.

5. Im komplexen Vektorraum $\mathbb{C}$ gilt nach 4. $\mathrm{span}_{\mathbb{C}}(\{1\}) = \mathbb{C}$. Im reellen Vektorraum $\mathbb{C}$ gilt $\mathrm{span}_{\mathbb{R}}(\{1\}) = \mathbb{R} \subseteq \mathbb{C}$.

3.2 Basis und Dimension

Ein gegebener Untervektorraum $U \subseteq V$ eines $\mathbb{K}$-Vektorraums V kann sich als Spann vieler verschiedener Teilmengen $M \subseteq V$ ergeben. Denn fügt man zu einem Erzeugendensystem M von U weitere Elemente aus U hinzu, so bleibt M ein Erzeugendensystem von U. So gilt beispielsweise $\mathrm{span}_{\mathbb{R}}(M) = N = \mathrm{span}_{\mathbb{R}}(N)$ für die Mengen $M = \{e_1, e_2\} \subseteq \mathbb{R}^3$ und $N = \mathbb{R}^2 \times \{0\}$. Die Menge M erzeugt den Untervektorraum $N = \mathbb{R}^2 \times \{0\} \subseteq \mathbb{R}^3$ aber offensichtlich auf viel effizientere Weise als N selbst, da sie nur zwei Vektoren enthält, während N unendlich viele Vektoren enthält, die sich alle als Linearkombinationen der Vektoren e_1 und e_2 schreiben lassen.

Das Konzept, das erfasst, ob eine gegebene Teilmenge $M \subseteq V$ eines $\mathbb{K}$-Vektorraums V den Untervektorraum $\text{span}_{\mathbb{K}}(M)$ effizient erzeugt oder ob sie überflüssige Vektoren enthält, die sich als Linearkombinationen anderer Vektoren aus M schreiben lassen, ist die *lineare Unabhängigkeit.*

Definition 3.2.1: Sei V ein Vektorraum über $\mathbb{K}$.

1. Ein n-Tupel von Vektoren $(v_1, ..., v_n) \in V^{\times n}$ heißt **linear unabhängig**, falls gilt:

$$\Sigma_{j=1}^n \lambda_j v_j = \lambda_1 v_1 + ... + \lambda_n v_n = 0 \text{ mit } \lambda_1, ..., \lambda_n \in \mathbb{K} \quad \Rightarrow \quad \lambda_1 = \ldots = \lambda_n = 0.$$

   Ansonsten heißt es **linear abhängig**.

2. Eine Teilmenge $M \subseteq V$ heißt **linear unabhängig**, falls für jedes $n \in \mathbb{N}$ und paarweise verschiedene Vektoren $v_1, ..., v_n \in M$ das n-Tupel $(v_1, ..., v_n)$ linear unabhängig ist. Ansonsten heißt sie **linear abhängig**.

Offensichtlich spielt die Reihenfolge der Vektoren $(v_1, ..., v_n)$ in dem n-Tupel keine Rolle für die lineare Unabhängigkeit des n-Tupels, denn die Vektoraddition ist kommutativ. Der Grund warum man dennoch Tupel von Vektoren aus V und nicht nur Teilmengen von V auf lineare Unabhängigkeit untersucht, ist, dass in einem n-Tupel Vektoren mehrfach auftreten können, während dies in einer Teilmenge von V nicht der Fall ist. In diesem Fall ist das n-Tupel immer linear abhängig, die Menge der in dem n-Tupel enthaltenen Vektoren aber nicht unbedingt.

Beispiel 3.2.2:

1. In jedem $\mathbb{K}$-Vektorraum V ist $\emptyset \subseteq V$ linear unabhängig, denn die leere Menge enthält keine Vektoren.

2. In jedem $\mathbb{K}$-Vektorraum V ist jedes n-Tupel von Vektoren aus V, das den Nullvektor $0 \in V$ enthält, und jede Teilmenge $M \subseteq V$ mit $0 \in M$ linear abhängig, denn es gilt $0 = 0 \cdot 0 = 1 \cdot 0$ mit $1 \neq 0$. Eine einelementige Teilmenge $M = \{v\}$ ist linear abhängig genau dann, wenn $v = 0$. Denn für $v \neq 0$ folgt aus $\lambda v = 0$ mit Lemma 3.1.3, 4. $\lambda = 0$.

3. In jedem $\mathbb{K}$-Vektorraum V und für beliebige Vektoren $v \in V$ ist (v, v) linear abhängig, denn $1 \cdot v + (-1)v = v - v = 0$ nach Lemma 3.1.3. Die Menge $\{v, v\} = \{v\}$ ist jedoch nach 2. linear unabhängig für $v \neq 0$. Für *paarweise verschiedene* Vektoren $v_1, ..., v_n \in V$ ist $(v_1, ..., v_n)$ linear unabhängig genau dann, wenn $\{v_1, ..., v_n\}$ linear unabhängig ist.

4. Sei $V = \mathbb{K}^n$. Dann ist das n-Tupel $(e_1, ..., e_n) \in V^{\times n}$ mit

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

linear unabhängig, denn eine Linearkombination dieser Elemente ist von der Form

$$\Sigma_{j=1}^{n}\lambda_j e_j = \lambda_1 \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix} + \lambda_2 \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix} + \ldots + \lambda_n \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \vdots \\ \lambda_{n-1} \\ \lambda_n \end{pmatrix}.$$

Damit folgt aus $\Sigma_{j=1}^{n}\lambda_j e_j = 0 = (0, ..., 0)^T$ dann $\lambda_1 = \ldots = \lambda_n = 0$.

5. Im $\mathbb{R}^3$ ist das Tripel (u, v, w) mit

$$u = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \quad v = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}, \quad w = \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix}$$

linear abhängig, denn es gilt

$$2 \cdot u - 2 \cdot v - w = \begin{pmatrix} 2 \\ 4 \\ 0 \end{pmatrix} - \begin{pmatrix} 0 \\ 2 \\ -2 \end{pmatrix} - \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix} = \begin{pmatrix} 2-0-2 \\ 4-2-2 \\ 0+2-2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

6. Im $\mathbb{R}^3$ ist das Tripel (u, v, w) mit

$$u = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \quad v = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}, \quad w = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}$$

linear unabhängig, denn es gilt

$$\alpha u + \beta v + \gamma w = \begin{pmatrix} \alpha \\ 2\alpha \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ \beta \\ -\beta \end{pmatrix} + \begin{pmatrix} \gamma \\ 2\gamma \\ 2\gamma \end{pmatrix} = \begin{pmatrix} \alpha+\gamma \\ 2\alpha+\beta+2\gamma \\ 2\gamma-\beta \end{pmatrix}$$

Aus $\alpha u + \beta v + \gamma w = 0 = (0, 0, 0)^T$ folgt daher $\alpha + \gamma = 0$, $2(\alpha + \gamma) + \beta = 0$ und $2\gamma - \beta = 0$. Die ersten zwei Gleichungen ergeben $\alpha = -\gamma$ und $\beta = 0$, und Einsetzen in die letzte ergibt dann $\alpha = \beta = \gamma = 0$.

Wir geben nun noch einfach handhabbare Kriterien an, mit denen sich untersuchen lässt, ob eine Teilmenge $M \subseteq V$ eine $\mathbb{K}$-Vektorraums V linear unabhängig ist. Zum Beweis der *linearen Abhängigkeit* einer Menge $M \subseteq V$ reicht es aus, zu zeigen, dass ein Vektor in M existiert, der sich als Linearkombination der anderen schreiben lässt, dass M eine linear abhängige Teilmenge besitzt oder dass sich ein Vektor aus V auf verschiedene Weise als Linearkombination von Vektoren aus M schreiben lässt.

Lemma 3.2.3: Sei V ein $\mathbb{K}$-Vektorraum. Eine Teilmenge $M \subseteq V$ ist genau dann linear unabhängig, wenn eine der folgenden äquivalenten Bedingungen erfüllt ist:

(i) Für jeden Vektor $v \in M$ ist $v \notin \text{span}_{\mathbb{K}}(M \setminus \{v\})$.

(ii) Alle Teilmengen $N \subseteq M$ sind linear unabhängig.

(iii) Jeder Vektor $v \in \text{span}_{\mathbb{K}}(M)$ lässt sich *eindeutig* als Linearkombination von Vektoren in M schreiben.

Beweis:
(i) Ist $M \subseteq V$ linear abhängig, so existieren paarweise verschiedene Vektoren $v_1, ..., v_n \in M$ und $\lambda_1, ..., \lambda_n \in \mathbb{K}$ mit $\lambda_i \neq 0$ für ein $i \in \{1, ..., n\}$, so dass $\Sigma_{j=1}^n \lambda_j v_j = 0$. Daraus folgt

$$v_i = -\lambda_i^{-1}(\lambda_1 v_1 + ...\lambda_{i-1}v_{i-1} + \lambda_{i+1}v_{i+1} + ... + \lambda_n v_n) \in \mathrm{span}_{\mathbb{K}}(M \setminus \{v_i\}).$$

Also folgt aus $v \notin \mathrm{span}_{\mathbb{K}}(M \setminus \{v\})$ für alle $v \in V$, dass M linear unabhängig ist.

Gibt es einen Vektor $v \in M$ mit $v \in \mathrm{span}_{\mathbb{K}}(M \setminus \{v\})$, so existieren paarweise verschiedene Vektoren $v_1, ..., v_n \in M \setminus \{v\}$ und $\lambda_1, ..., \lambda_n \in \mathbb{K}$ mit $v = \Sigma_{j=1}^n \lambda_j v_j$. Also ist $0 = v - \Sigma_{j=1}^n \lambda_j v_j$ eine Linearkombination aus paarweise verschiedenen Vektoren aus M, in der nicht alle Koeffizienten verschwinden, und M ist linear abhängig.

(ii) Sind alle Teilmengen $N \subseteq M$ linear unabhängig, so ist auch $M \subseteq M$ linear unabhängig. Ist eine Teilmenge $N \subseteq M$ linear abhängig, so existieren paarweise verschiedene Vektoren $v_1, ..., v_n \in N$ und $\lambda_1, ..., \lambda_n \in \mathbb{K}$ mit $(\lambda_1, ..., \lambda_n) \neq (0, ..., 0)$ und $\Sigma_{j=1}^n \lambda_j v_j = 0$. Wegen $N \subseteq M$ gilt aber auch $v_1, ..., v_n \in M$ und somit ist auch M linear abhängig.

(iii) Jeder Vektor $v \in \mathrm{span}_{\mathbb{K}}(M)$ ist eine Linearkombination $v = \Sigma_{j=1}^n \lambda_j v_j$ mit $v_1, ..., v_n \in M$ und $\lambda_1, ..., \lambda_n \in \mathbb{K}$. Gibt es eine zweite Linearkombination $v = \Sigma_{j=1}^m \mu_j v'_j$ mit $v'_1, ..., v'_n \in M$ und $\mu_1, ..., \mu_n \in \mathbb{K}$, so können wir entsprechend Nullen in den Koeffizienten ergänzen und daher annehmen, dass in diesen Linearkombinationen die gleichen Vektoren auftreten. Sei also $m = n$ und $v_1 = v'_1, \ldots, v_n = v'_n$. Dann gilt $0 = v - v = \Sigma_{j=1}^n \lambda_j v_j - \Sigma_{j=1}^n \mu_j v_j = \Sigma_{j=1}^n (\lambda_j - \mu_j) v_j$. Ist M linear unabhängig, so folgt daraus $\lambda_1 = \mu_1, \ldots, \lambda_n = \mu_n$. Also ist die Darstellung von v als Linearkombination von Vektoren aus M eindeutig.

Lässt sich umgekehrt jeder Vektor aus $\mathrm{span}_{\mathbb{K}}(M)$ eindeutig als Linearkombination von Vektoren aus M schreiben, so gilt das auch für den Nullvektor $0 \in \mathrm{span}_{\mathbb{K}}(M)$. Sind also $v_1, ..., v_n \in M$ paarweise verschieden mit $\Sigma_{j=1}^n \lambda_j v_j = 0 = \Sigma_{j=1}^n 0\, v_j$, so folgt mit der Eindeutigkeit der Darstellung als Linearkombination $\lambda_1 = \ldots = \lambda_n = 0$. □

Wir kombinieren nun den Begriff des Erzeugendensystems aus Definition 3.1.14 mit dem Begriff der linearen Unabhängigkeit aus Definition 3.2.1. Dies führt auf das Konzept der *Basis*, einer Teilmenge eines $\mathbb{K}$-Vektorraums V, die V maximal effizient oder sparsam erzeugt.

Definition 3.2.4: Eine Teilmenge eines $\mathbb{K}$-Vektorraums V heißt **Basis** von V, wenn sie linear unabhängig ist und V erzeugt.

Beispiel 3.2.5:

1. Jeder Vektorraum ist ein Erzeugendensystem von sich selbst: $\mathrm{span}_{\mathbb{K}}(V) = V$, aber nie eine Basis von sich selbst, denn er enthält den Nullvektor und ist somit linear abhängig.

2. Die einzige Basis des Nullvektorraums ist die **leere Basis**, denn die Teilmenge $\emptyset \subseteq \{0\}$ ist linear unabhängig und per Definition gilt $\mathrm{span}_{\mathbb{K}}(\emptyset) = \{0\}$.

3. Für jeden Körper $\mathbb{K}$ ist $B = \{1\}$ eine Basis von $\mathbb{K}$ als Vektorraum über sich selbst. Denn sie ist wegen $1 \neq 0$ linear unabhängig und erzeugt $\mathbb{K}$, da $\lambda = \lambda \cdot 1$ für alle $\lambda \in \mathbb{K}$ gilt. Allgemeiner kann man zeigen, dass eine einelementige Teilmenge $B' = \{\lambda\}$ genau dann eine Basis von $\mathbb{K}$ ist, wenn $\lambda \neq 0$ gilt. (Übung)

4. Die Menge $B = \{e_1, ..., e_n\}$ mit

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

aus Beispiel 3.2.2, 4. ist eine Basis des $\mathbb{K}$-Vektorraums $\mathbb{K}^n$. Denn sie ist nach Beispiel 3.2.2, 4. linear unabhängig, und jedes Element aus $\mathbb{K}^n$ lässt sich als Linearkombination dieser Vektoren schreiben:

$$\begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \vdots \\ \lambda_n \end{pmatrix} = \lambda_1 \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \lambda_3 \begin{pmatrix} 0 \\ 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \ldots + \lambda_n \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} = \Sigma_{j=1}^n \lambda_j e_j.$$

Diese Basis wird als die **Standardbasis** des $\mathbb{K}^n$ bezeichnet.

5. Die Menge $M = \{1, \mathrm{i}\} \subseteq \mathbb{C}$ ist eine Basis des *reellen* Vektorraums $\mathbb{C}$, denn jedes $z \in \mathbb{C}$ lässt sich eindeutig als $z = x + \mathrm{i}y$ mit $x, y \in \mathbb{R}$ schreiben. Auch die Menge $\{1 + \mathrm{i}, 1 - \mathrm{i}\}$ ist eine Basis des *reellen* Vektorraums $\mathbb{C}$ (Beweis:Übung).

6. Die **Monome** $f_n : \mathbb{N}_0 \to \mathbb{K}$ mit $f_n(m) = 0$ für $m \neq n$ und $f_n(n) = 1$ für $n, m \in \mathbb{N}_0$ bilden eine Basis des $\mathbb{K}$-Vektorraums $\mathbb{K}[x]$ der Polynome mit Koeffizienten in $\mathbb{K}$.

   Die Menge $B = \{f_n | n \in \mathbb{N}_0\}$ erzeugt $\mathbb{K}[x]$, da sich jedes Polynom $p : \mathbb{N}_0 \to \mathbb{K}$ als Linearkombination $p = \Sigma_{j \in \mathbb{N}_0} p(j) f_j$ schreiben lässt. Die Bedingung $p(j) = 0$ für fast alle $j \in \mathbb{N}_0$ garantiert dabei, dass diese Linearkombination endlich ist. Ist $\Sigma_{j=0}^n \lambda_j f_j = 0$, so folgt $\Sigma_{j=0}^n \lambda_j f_j(m) = \lambda_m = 0$ für alle $m \in \mathbb{N}_0$, und somit ist B auch linear unabhängig.

7. Ist M eine *endliche* Menge und $\mathbb{K}$ ein Körper, so bilden die Abbildungen $\delta_m : M \to \mathbb{K}$ mit $\delta_m(m') := 0$ für $m' \neq m$ und $\delta_m(m) := 1$ eine Basis des $\mathbb{K}$-Vektorraums $\mathrm{Abb}(M, \mathbb{K})$.

   Denn jede Abbildung $f : M \to \mathbb{K}$ ist eine Linearkombination $f = \Sigma_{m \in M} f(m) \delta_m$. Gilt $\Sigma_{m \in M} a_m \delta_m = 0$, so folgt $\Sigma_{m \in M} a_m \delta_m(m') = a_{m'} = 0$ für alle $m' \in M$, und somit ist $B = \{\delta_m | \, m \in M\}$ linear unabhängig.

8. Achtung: Ist M *unendlich*, so bilden die Abbildungen $\delta_m : M \to \mathbb{K}$ mit $\delta_m(m') := 0$ für $m' \neq m$ und $\delta_m(m) := 1$ *keine* Basis des $\mathbb{K}$-Vektorraums $\mathrm{Abb}(M, \mathbb{K})$. (Beweis: Übung).

Wir geben jetzt noch zwei alternative Charakterisierungen einer Basis. Die erste zeigt, dass eine Basis, wie schon angedeutet, tatsächlich ein Erzeugendensystem ist, das so klein wie möglich gewählt ist. Das bedeutet, dass man keinen Vektor aus einer Basis entfernen kann, ohne die Eigenschaft, dass sie V erzeugt, zu verlieren. Die zweite zeigt, dass man eine Basis auch als eine linear unabhängige Teilmenge sehen kann, die so groß wie möglich gewählt ist. Man kann zu einer Basis keinen Vektor hinzufügen, ohne die lineare Unabhängigkeit zu verlieren.

Satz 3.2.6: Sei V ein $\mathbb{K}$-Vektorraum und $B \subseteq V$. Dann sind äquivalent:

(i) B ist eine Basis von V.

(ii) B ist ein minimales Erzeugendensystem von V:
für kein $b \in B$ ist $B \setminus \{b\}$ ein Erzeugendensystem von V.

(iii) B ist eine maximale linear unabhängige Teilmenge von V:
für jedes $v \in V \setminus B$ ist $B \cup \{v\}$ linear abhängig.

Beweis:
(i)$\Leftrightarrow$ (ii): Zu zeigen ist, dass ein Erzeugendensystem B von V linear unabhängig ist genau dann, wenn für kein $b \in B$ die Menge $B \setminus \{b\}$ den Vektorraum V erzeugt. Ist B linear unabhängig, so folgt mit Lemma 3.2.3 (i), dass für alle $b \in B$ gilt $b \notin \text{span}_{\mathbb{K}}(B \setminus \{b\})$, und somit ist für kein $b \in B$ die Menge $B \setminus \{b\}$ ein Erzeugendensystem. Ist B dagegen linear abhängig, so existiert nach Lemma 3.2.3 (i) ein $b \in B$ mit $b \in \text{span}_{\mathbb{K}}(B \setminus \{b\})$. Somit ist dann auch $B \setminus \{b\}$ ein Erzeugendensystem von V.

(i) $\Leftrightarrow$ (iii): Zu zeigen ist, dass eine linear unabhängige Teilmenge $B \subseteq V$ genau dann ein Erzeugendensystem von V ist, wenn für alle $v \in V \setminus B$ die Menge $B \cup \{v\}$ linear abhängig ist. Ist $B \subseteq V$ linear unabhängig mit $\text{span}_{\mathbb{K}}(B) = V$, so lässt sich jeder Vektor $v \in V \setminus B$ als Linearkombination von Vektoren aus B schreiben, aber auch als Linearkombination $v = 1 \cdot v$. Also ist die Darstellung von v als Linearkombination von Vektoren aus $B \cup \{v\}$ nicht eindeutig, und mit dem Lemma 3.2.3 (iii) folgt, dass $B \cup \{v\}$ linear abhängig ist.

Ist B linear unabhängig und $B \cup \{v\}$ linear abhängig für alle $v \in V \setminus B$, so gibt es für alle Vektoren $v \in V \setminus B$ Vektoren $b_1, ..., b_n \in B$ und Koeffizienten $\lambda, \lambda_1, ..., \lambda_k \in \mathbb{K}$ mit $\lambda v + \Sigma_{i=1}^n \lambda_i b_i = 0$ und $(\lambda, \lambda_1, ..., \lambda_n) \neq (0, ..., 0)$. Da B linear unabhängig ist, muss $\lambda \neq 0$ gelten, denn aus $\lambda = 0$ folgt $\Sigma_{j=1}^n \lambda_j b_j = 0$ und mit der linearen Unabhängigkeit von B auch $\lambda_1 = ... = \lambda_n = 0$. Also ist $v = \Sigma_{j=1}^n (-\lambda^{-1}\lambda_j) b_j \in \text{span}_{\mathbb{K}}(B)$ und B erzeugt V. $\square$

Offensichtlich besitzen manche der bisher betrachteten Vektorräume wie beispielsweise der $\mathbb{K}$-Vektorraum $\mathbb{K}^n$ *endliche* Erzeugendensysteme oder Basen, während dies bei anderen wie dem $\mathbb{K}$-Vektorraum $\mathbb{K}[x]$ nicht der Fall zu sein scheint. Dies ist ein wichtiger Unterschied. Wir werden später zeigen, dass sich jeder $\mathbb{K}$-Vektorraum mit einem endlichen Erzeugendensystem im Prinzip wie einer der Vektorräume $\mathbb{K}^n$ verhält, während die Vektorräume, die kein endliches Erzeugendensystem besitzen, viel komplizierter sein können.

Definition 3.2.7: Ein $\mathbb{K}$-Vektorraum V heißt **endlich erzeugt** wenn er ein endliches Erzeugendensystem besitzt, also eine endliche Teilmenge $M \subseteq V$ mit $V = \text{span}_{\mathbb{K}}(M)$.

Beispiel 3.2.8:

1. Für jeden Körper $\mathbb{K}$ sind der Nullvektorraum $\{0\}$ über $\mathbb{K}$ und die Vektorräume $\mathbb{K}^n$ für $n \in \mathbb{N}$ endlich erzeugt, denn die endliche Basis $\emptyset$ und die Standardbasis $\{e_1, .., e_n\}$ sind Erzeugendensysteme.

2. Ist M eine endliche Menge, so ist für jeden Körper $\mathbb{K}$ der Vektorraum $\text{Abb}(M, \mathbb{K})$ endlich erzeugt, denn $B = \{\delta_m | m \in M\}$ mit $\delta_m : M \to \mathbb{K}$, $\delta_m(m) := 1$ und $\delta_m(m') := 0$ für alle $m' \neq m$ aus Beispiel 3.2.5, 7. erzeugt $\text{Abb}(M, \mathbb{K})$.

3. Der Vektorraum $\mathbb{K}[x]$ der Polynome $p : \mathbb{N}_0 \to \mathbb{K}$ mit $p(n) = 0$ für fast alle $n \in \mathbb{N}_0$ ist nicht endlich erzeugt.

   Für jede endliche Teilmenge $\{p_1, ..., p_n\} \subseteq \mathbb{K}[x]$ gibt es ein $m \in \mathbb{N}$ mit $p_i(k) = 0$ für alle $i \in \{1, ..., n\}$ und alle $k > m$, denn jedes Polynom p_i erfüllt $p_i(k) = 0$ für alle bis auf endlich viele $k \in \mathbb{N}_0$. Daraus folgt für jedes Polynom $q = \Sigma_{i=1}^n \lambda_i p_i$ mit $\lambda_1, ..., \lambda_n \in \mathbb{K}$ auch $q(k) = 0$ für alle $k > m$ und somit $\text{span}_\mathbb{K}(\{p_1, ..., p_n\}) \subseteq \{q \in \mathbb{K}[x] | q(k) = 0 \forall k > m\}$. Also ist z. B. das Monom $f_{m+1} : \mathbb{N}_0 \to \mathbb{K}$ mit $f_{m+1}(m+1) = 1$ und $f_{m+1}(k) = 0$ für $k \neq m+1$ nicht in $\text{span}_\mathbb{K}(\{p_1, ..., p_n\}) \neq \mathbb{K}[x]$ enthalten und damit $\text{span}_\mathbb{K}(\{p_1, ..., p_n\}) \neq \mathbb{K}[x]$.

Besitzt ein $\mathbb{K}$-Vektorraum V eine endliche Basis, so ist er offensichtlich endlich erzeugt, denn eine endliche Basis ist insbesondere ein endliches Erzeugendensystem. Wir zeigen jetzt, dass auch die Umkehrung gilt, also jeder endlich erzeugte $\mathbb{K}$-Vektorraum auch eine endliche Basis besitzt. Dazu müssen wir uns genauer mit der Beziehung zwischen Erzeugendensystemen, linearer Unabhängigkeit und Basen beschäftigen. Zuerst stellen wir fest, dass in einem $\mathbb{K}$-Vektorraum, der nicht endlich erzeugt ist, zu jedem $n \in \mathbb{N}$ ein n-Tupel linear unabhängiger Vektoren existiert.

Lemma 3.2.9: Sei V ein $\mathbb{K}$-Vektorraum, der nicht endlich erzeugt ist. Dann existieren zu jedem $n \in \mathbb{N}$ Vektoren $v_1, ..., v_n \in V$ so dass $(v_1, ..., v_n)$ linear unabhängig ist.

Beweis:
Vollständige Induktion über n.

Induktionsanfang $n = 1$: Da V nicht endlich erzeugt ist, gilt $V \neq \{0\}$ und man kann einen Vektor $v \in V \setminus \{0\}$ wählen. Dann ist das Tupel (v) nach Beispiel 3.2.2, 2. linear unabhängig.

Induktionsschritt $n \to n + 1$: Sei gezeigt, dass Vektoren $v_1, ..., v_n \in V$ existieren, so dass $(v_1, ..., v_n)$ linear unabhängig ist. Da sonst $\{v_1, ..., v_n\}$ ein endliches Erzeugendensystem wäre, gibt es einen ein Vektor $v_{n+1} \in V \setminus \text{span}_\mathbb{K}(\{v_1, ..., v_n\})$. Sind $\lambda_1, ..., \lambda_{n+1} \in \mathbb{K}$ mit $\sum_{i=1}^{n+1} \lambda_i v_i = 0$, so muss $\lambda_{n+1} = 0$ gelten, denn sonst wäre $v_{n+1} = \Sigma_{i=1}^n (-\lambda_{n+1}^{-1} \lambda_i) v_i \in \text{span}_\mathbb{K}(\{v_1, ..., v_n\})$, im Widerspruch zur Annahme. Daraus folgt $\Sigma_{i=1}^n \lambda_i v_i = 0$ und wegen der linearen Unabhängigkeit von $(v_1, ..., v_n)$ dann auch $\lambda_1 = \ldots = \lambda_n = 0$. Also ist $(v_1, ..., v_{n+1})$ linear unabhängig. □

Wir zeigen nun, dass wir in einem beliebigen Vektorraum durch Weglassen von Elementen eines Erzeugendensystems immer eine Basis konstruieren können. Es gilt sogar noch eine stärkere Aussage, nämlich, dass zu einer vorgegebenen linear unabhängigen Teilmenge und einem Erzeugendensystem, das diese Teilmenge enthält, immer eine Basis existiert, die diese Teilmenge ebenfalls enthält und selbst Teilmenge des Erzeugendensystems ist.

Für endliche Erzeugendensysteme ist der Beweis dieser Aussage relativ simpel und erfolgt mit vollständiger Induktion. Im Fall unendlicher Erzeugendensysteme ist die Aussage eigentlich ein Postulat. Sie folgt nämlich aus dem Zornschen Lemma, von dem man zeigen kann, dass es äquivalent zum Auswahlaxiom ist.

Satz 3.2.10: (**Basisauswahlsatz**)
Sei V ein $\mathbb{K}$-Vektorraum, $E \subseteq V$ ein Erzeugendensystem von V und $M \subseteq E$ eine linear unabhängige Teilmenge von V. Dann gibt es eine Basis B von V mit $M \subseteq B \subseteq E$.

Beweis:
1. Wir beweisen die Aussage für endliche Erzeugendensysteme E durch vollständige Induktion

über die Anzahl n der Elemente in E.
Induktionsanfang $n = 0$: Erzeugt $E = \emptyset$ den Vektorraum V, dann ist $V = \mathrm{span}_{\mathbb{K}}(\emptyset) = \{0\}$ und $E = M = \emptyset$ ist eine Basis.

Induktionsschritt $n \to n+1$: Sei die Aussage bereits bewiesen für Erzeugendensysteme $E \subseteq V$ von V mit höchstens n Elementen. Sei $E = \{v_1, ..., v_{n+1}\}$ ein Erzeugendensystem von V und $M \subseteq E$ linear unabhängig.

Ist E linear unabhängig, so ist E eine Basis von V mit $M \subseteq E$. Ist E linear abhängig, so gibt es eine Linearkombination $\Sigma_{k=1}^{n+1} \lambda_k v_k = 0$ mit $(\lambda_1, ..., \lambda_{n+1}) \neq (0, ..., 0)$. Da M linear unabhängig ist, muss mindestens ein $j \in \{1, ..., n+1\}$ mit $v_j \in E \setminus M$ und $\lambda_j \neq 0$ existieren, da sonst $\Sigma_{k=1}^{n+1} \lambda_k v_k \in \mathrm{span}_{\mathbb{K}}(M)$ und damit $\lambda_1 = ... = \lambda_{n+1} = 0$ folgen würde.

Durch Umnummerieren der Vektoren v_i können wir erreichen, dass $v_1 \in E \setminus M$ mit $\lambda_1 \neq 0$ gilt. Dann folgt aus $\Sigma_{k=1}^{n+1} \lambda_k v_k = 0$, dass $v_1 = -\Sigma_{k=2}^{n+1} (\lambda_1^{-1} \lambda_k) v_k \in \mathrm{span}_{\mathbb{K}}(E \setminus \{v_1\})$ und $M \subseteq E \setminus \{v_1\}$. Damit ist auch $E' = E \setminus \{v_1\}$ ein Erzeugendensystem von V mit $M \subseteq E'$. Da E' nur n Elemente hat, existiert nach Induktionsvoraussetzung eine Basis $B \subseteq E' \subseteq E$ mit $M \subseteq B$.

2. Der Beweis für allgemeine Vektorräume und Erzeugendensysteme benutzt das *Zornsche Lemma*. Um das Zornsche Lemma zu formulieren, benötigen wir zunächst den Begriff einer *partiell geordneten Menge*.

Definition 3.2.11: Eine **partielle Ordnung** $\leq$ auf einer Menge X ist eine Relation $\leq$ auf X mit den folgenden Eigenschaften:

(PO1) reflexiv: $x \leq x$ für alle $x \in X$,
(PO2) transitiv: $x \leq y$, $y \leq z \Rightarrow x \leq z$,
(PO3) antisymmetrisch: $(x \leq y) \wedge (y \leq x) \Rightarrow x = y$.

Eine **Kette** in einer partiell geordneten Menge X ist eine Teilmenge $K \subseteq X$, so dass $(K, \leq)$ **total geordnet** ist: für alle $h, k \in K$ gilt $h \leq k$ oder $k \leq h$. Die Menge X heißt **induktiv geordnet**, wenn jede Kette K in X eine obere Schranke besitzt, also ein $m \in X$ existiert mit $k \leq m$ für alle $k \in K$.

Lemma 3.2.12: (**Zornsches Lemma**) Jede nichtleere induktiv geordnete Menge X besitzt ein maximales Element: es gibt ein $x \in X$, so dass kein $y \in X \setminus \{x\}$ mit $x \leq y$ existiert.

Wir betrachten die Menge $X(M, E) = \{A \subseteq E | A \text{ linear unabhängig}, M \subseteq A \subseteq E\} \subseteq \mathrm{Pot}(M)$, die wegen $M \in X(M, E)$ nicht leer ist. Dann ist $\leq := \subseteq$ eine partielle Ordnung auf $X(M, E)$, denn es gilt $A \subseteq A$, $A \subseteq B$, $B \subseteq C \Rightarrow A \subseteq C$ und $(A \subseteq B) \wedge (B \subseteq A) \Rightarrow A = B$.

1. Wir zeigen: $X(M, E)$ ist induktiv geordnet, denn für jede Kette $K \subseteq X(M, E)$ ist das Element $\cup K = \{x \in X \mid \exists A \in K : x \in A\} \in X(M, E)$ eine obere Schranke.

Da K eine Kette ist, gilt $A \subseteq B$ oder $B \subseteq A$ für alle $A, B \in K$. Offensichtlich folgt daraus auch $A \subseteq \cup K$ für alle $A \in K$, und aus $M \subseteq A \subseteq E$ für alle $A \in K$ ergibt sich $M \subseteq \cup K \subseteq B$.

Wir zeigen, dass $\cup K$ linear unabhängig ist. Seien dazu $\lambda_1, ..., \lambda_n \in \mathbb{K}$ und $v_1, ..., v_n \in \cup K$ mit $\Sigma_{j=1}^n \lambda_j v_j = 0$. Dann existieren $A_1, ..., A_n \in K$ mit $v_i \in A_i$. Durch Umnummerieren können wir $A_1 \subseteq A_2 \subseteq ... \subseteq A_n$ erreichen. Daraus folgt $v_1, ..., v_n \in A_n$, und aus der linearen Unabhängigkeit von A_n folgt $\lambda_1 = ... = \lambda_n = 0$.

Also ist $\cup K$ linear unabhängig und somit ist $\cup K \in X(M, E)$ eine obere Schranke der Kette K. Damit ist gezeigt, dass $X(M, E)$ induktiv geordnet ist.

2. Da $X(M, E)$ nicht leer und induktiv geordnet ist, existiert nach dem Zornschen Lemma ein maximales Element $B \in X(M, E)$. Dieses ist per Definition eine linear unabhängige Teilmenge von V mit $M \subseteq B \subseteq E$. Wir zeigen, dass B eine Basis ist, also dass $\text{span}_{\mathbb{K}}(B) = V$ gilt.

Die Maximalität von B bedeutet, dass keine linear unabhängige Teilmenge $C \subseteq V$ mit $M \subseteq B \subsetneq C \subseteq E$ existiert. Damit ist für jeden Vektor $e \in E \setminus B$ die Menge $B \cup \{e\}$ linear abhängig. Es gibt also $\lambda, \lambda_1, ..., \lambda_n \in \mathbb{K}$ mit $(\lambda, \lambda_1, ..., \lambda_n) \neq (0, ..., 0)$ und $b_1, .., b_n \in B$ so dass $\lambda e + \Sigma_{j=1}^n \lambda_j b_j = 0$. Wegen der linearen Unabhängigkeit von B ist $\lambda \neq 0$ und $e = -\Sigma_{j=1}^n \lambda^{-1} \lambda_i b_i \in \text{span}_{\mathbb{K}}(B)$. Daraus folgt $E = B \cup (E \setminus B) \subseteq \text{span}_{\mathbb{K}}(B)$ und $V = \text{span}_{\mathbb{K}}(E) \subseteq \text{span}_{\mathbb{K}}(\text{span}_{\mathbb{K}}(B)) = \text{span}_{\mathbb{K}}(B)$. Also ist B eine Basis von V mit $M \subseteq B \subseteq E$. □

Der Basisauswahlsatz garantiert insbesondere, dass jeder Vektorraum eine Basis besitzt. Denn man kann für das Erzeugendensystem E in Satz 3.2.10 den Vektorraum V selbst wählen und für die linear unabhängige Teilmenge $M \subseteq E$ in Satz 3.2.10 die leere Menge. Ebenso kann man jede linear unabhängige Teilmenge $M \subseteq V$ zu einer Basis $B \subseteq V$ ergänzen, indem man den Basisauswahlsatz auf $E = V$ anwendet.

Korollar 3.2.13: (**Existenz von Basen**) Jeder $\mathbb{K}$-Vektorraum V hat eine Basis.

Korollar 3.2.14: (**Basisergänzungssatz**) Jede linear unabhängige Teilmenge $M \subseteq V$ lässt sich zu einer Basis von V ergänzen: es gibt eine Basis $B \subseteq V$ mit $M \subseteq B$.

Nachdem die Existenz von Basen gesichert ist, stellt sich die Frage nach Ihrer Eindeutigkeit. Schon aus einfachen Beispielen ergibt sich, dass die Wahl einer Basis nur im Fall des Nullvektorraums eindeutig ist und in allen anderen Fällen hochgradig beliebig ist. Wir zeigen nun, dass man in einem endlich erzeugten $\mathbb{K}$-Vektorraum V aus einer beliebigen Basis B von V eine neue Basis konstruieren kann, indem man einen Teil der Elemente von B entfernt und durch gleich viele Elemente einer beliebigen linear unabhängigen Teilmenge $M \subseteq V$ ersetzt.

Lemma 3.2.15: (**Basisaustauschlemma**)
Sei V ein $\mathbb{K}$-Vektorraum mit Basis $B = \{v_1, ..., v_n\}$ und $w = \Sigma_{k=1}^n \mu_k v_k$ mit $\mu_1, ..., \mu_n \in \mathbb{K}$. Dann ist für alle $k \in \{1, ..., n\}$ mit $\mu_k \neq 0$ auch $B'_k = \{v_1, ..., v_{k-1}, w, v_{k+1}, ..., v_n\}$ eine Basis von V.

Beweis:
Sei $w = \Sigma_{k=1}^n \mu_k v_k$ mit $\mu_1, ..., \mu_n \in \mathbb{K}$ und $k \in \{1, ..., n\}$ mit $\mu_k \neq 0$. Dann können wir durch Umnummerieren der Vektoren $v_1, ..., v_n$ erreichen, dass $k = 1$ und $B'_1 = \{w, v_2, ..., v_n\}$ ist. Zu zeigen ist dann, dass (i) $\text{span}_{\mathbb{K}}(B'_1) = V$ und dass (ii) B'_1 linear unabhängig ist.

(i) Da B eine Basis ist, existieren zu jedem Vektor $v \in V$ eindeutig bestimmte $\lambda_1, .., \lambda_n \in \mathbb{K}$ mit $v = \Sigma_{k=1}^n \lambda_k v_k$. Da $\mu_1 \neq 0$, gilt $v_1 = \mu_1^{-1} w - \mu_1^{-1} \Sigma_{k=2}^n \mu_k v_k$ und somit auch

$$v = \lambda_1 v_1 + \Sigma_{k=2}^n \lambda_k v_k = \lambda_1 \mu_1^{-1} w - \lambda_1 \mu_1^{-1} \Sigma_{k=2}^n \mu_k v_k + \Sigma_{k=2}^n \lambda_k v_k = \lambda_1 \mu_1^{-1} w + \Sigma_{k=2}^n (\lambda_k - \lambda_1 \mu_1^{-1} \mu_k) v_k$$

Also ist $v \in \text{span}_{\mathbb{K}}(B'_1)$ und $\text{span}_{\mathbb{K}}(B'_1) = \text{span}_{\mathbb{K}}(B) = V$.

(ii) Seien $\lambda_1, ..., \lambda_n \in \mathbb{K}$ mit $\lambda_1 w + \Sigma_{k=2}^n \lambda_k v_k = 0$. Einsetzen von $w = \Sigma_{k=1}^n \mu_k v_k$ ergibt

$$0 = \lambda_1 \Sigma_{k=1}^n \mu_k v_k + \Sigma_{k=2}^n \lambda_k v_k = \lambda_1 \mu_1 v_1 + \Sigma_{k=2}^n (\lambda_1 \mu_k + \lambda_k) v_k.$$

Wegen der linearen Unabhängigkeit von B folgt daraus $\mu_1 \lambda_1 = 0$ und $\lambda_1 \mu_k + \lambda_k = 0$ für alle $k \in \{2, .., n\}$. Da nach Voraussetzung $\mu_1 \neq 0$ gilt, folgt $\lambda_1 = 0$ und damit auch $\lambda_2 = ... = \lambda_n = 0$. Also ist B'_1 linear unabhängig. □

Satz 3.2.16: (**Basisaustauschsatz**)
Sei V ein $\mathbb{K}$-Vektorraum mit Basis $B = \{v_1, ..., v_n\}$ und sei die Menge $\{w_1, ..., w_m\} \subseteq V$ linear unabhängig. Dann gilt $m \leq n$ und es gibt paarweise verschiedene $i_1, ..., i_m \in \{1, ..., n\}$, so dass $B' = (B \setminus \{v_{i_1}, ..., v_{i_m}\}) \cup \{w_1, ..., w_m\}$ eine Basis von V ist.

Beweis:
Vollständige Induktion über m.

Induktionsanfang $m = 1$**:** Ist $m = 1$, so ist $\{w_1\} \subseteq V$ linear unabhängig genau dann, wenn $w_1 \neq 0$. Da $B = \{v_1, ..., v_n\}$ eine Basis ist, existieren $\mu_1, .., \mu_n \in \mathbb{K}$ mit $w_1 = \Sigma_{k=1}^n \mu_k b_k$ und wegen $w_1 \neq 0$ existiert mindestens ein $k \in \{1, ..., n\}$ mit $\mu_k \neq 0$. Nach dem Basisaustauschlemma ist dann auch $(B \setminus \{v_k\}) \cup \{w_1\} = \{v_1, ..., v_{k-1}, w_1, v_{k+1}, ..., v_n\}$ eine Basis von V.

Induktionsschritt $m \to m+1$**:** Sei die Aussage bereits bewiesen für alle linear unabhängigen Teilmengen mit maximal m Elementen und sei $W := \{w_1, ..., w_{m+1}\} \subseteq V$ eine linear unabhängige Teilmenge. Dann ist nach Lemma 3.2.3 auch $\{w_1, ..., w_m\} \subseteq W$ linear unabhängig. Nach Induktionsvoraussetzung gilt also $m \leq n$, und es gibt paarweise verschiedene $i_1, ..., i_m \in \{1, ..., n\}$ so dass $B' = (B \setminus \{v_{i_1}, ..., v_{i_m}\}) \cup \{w_1, ..., w_m\}$ eine Basis von V ist. Durch Umnummerieren der Vektoren $v_1, ..., v_n$ können wir $i_1 = 1$,..., $i_m = m$ und $B' = \{w_1, .., w_m, v_{m+1}, ..., v_n\}$ erreichen.

Da $w_{m+1} \in \mathrm{span}_{\mathbb{K}}(B')$ wäre für $n = m$ auch $w_{m+1} \in \mathrm{span}_{\mathbb{K}}(\{w_1, ..., w_m\})$ - ein Widerspruch zur linearen Unabhängigkeit von W nach Lemma 3.2.3. Also muss $m + 1 \leq n$ gelten.

Da B' eine Basis von V ist, gibt es $\lambda_1, ..., \lambda_n$ mit $w_{m+1} = \Sigma_{k=1}^m \lambda_k w_k + \Sigma_{k=m+1}^n \lambda_k v_k$. Dabei muss $\lambda_k \neq 0$ für mindestens ein $k \in \{m+1, ..., n\}$ gelten, denn sonst wäre $w_{m+1} \in \mathrm{span}_{\mathbb{K}}(\{w_1, ..., w_m\})$, im Widerspruch zur linearen Unabhängigkeit von W. Durch Umnummerieren der Vektoren v_{m+1},...,v_n können wir erreichen, dass $\lambda_{m+1} \neq 0$. Nach dem Basisaustauschlemma ist dann auch $B'' = (B' \setminus \{v_{m+1}\}) \cup \{w_{m+1}\} = (B \setminus \{v_1, ..., v_{m+1}\}) \cup \{w_1, ..., w_{m+1}\}$ eine Basis von V. $\square$

Korollar 3.2.17: Ist ein $\mathbb{K}$-Vektorraum V endlich erzeugt, so ist jede Basis von V endlich, und je zwei Basen von V haben gleich viele Elemente. Ist V nicht endlich erzeugt, so sind alle Basen von V unendlich.

Beweis:
Ist V nicht endlich erzeugt, so kann V keine endliche Basis haben, und somit sind alle Basen von V unendlich. Ist V endlich erzeugt, so gibt es ein endliches Erzeugendensystem $E \subseteq V$ und nach dem Basisauswahlsatz für $M = \emptyset$ eine endliche Basis $B \subseteq E$.

Ist B eine endliche Basis von V mit $|B| = n \in \mathbb{N}_0$ Elementen und B' eine weitere Basis mit $|B'| = k \in \mathbb{N}_0 \cup \{\infty\}$ Elementen, so ist jede endliche Teilmenge $\{w_1, ..., w_m\} \subseteq B'$ linear unabhängig, und mit dem Basisaustauschsatz folgt $m \leq n$. Also ist B' endlich mit $k \leq n$ Elementen. Da aber auch B' eine Basis ist, folgt mit dem gleichen Argument auch $n \leq k$ und damit $k = n$. $\square$

Mit Korollar 3.2.17 wird es sinnvoll, einen endlich erzeugten $\mathbb{K}$-Vektorraum V durch die Zahl der Elemente in einer Basis von V zu charakterisieren, da diese Zahl nicht von der Wahl der Basis abhängt. Lassen wir für den Wert dieser Zahl auch ∞ zu, so können wir damit auch nicht endlich erzeugte Vektorräume erfassen.

Definition 3.2.18: Die **Dimension** $\dim_{\mathbb{K}}(V)$ eines $\mathbb{K}$-Vektorraums V ist definiert durch

$$\dim_{\mathbb{K}}(V) = \begin{cases} 0 & V = \{0\} \\ n & \text{falls } V \text{ eine Basis mit } n \in \mathbb{N} \text{ Elementen besitzt} \\ \infty & \text{falls } V \text{ nicht endlich erzeugt ist.} \end{cases}$$

Falls $\dim_{\mathbb{K}}(V) \in \mathbb{N}_0$, nennt man V **endlich-dimensional**, ansonsten **unendlich-dimensional**.

Wir werden später zeigen, dass die Dimension endlich-dimensionale Vektorräume über einem Körper $\mathbb{K}$ vollständig beschreibt. Schon hier wird aber deutlich, dass sich damit wichtige Eigenschaften von Vektorräumen erfassen lassen. Aus dem Basisauswahlsatz und Korollar 3.2.17 ergibt sich, dass ein Vektorraum genau dann endlich erzeugt ist, wenn er endliche Dimension $n \in \mathbb{N}_0$ hat. In diesem Fall folgt außerdem, dass jedes Erzeugendensystem mindestens n Elemente besitzen muss. Aus dem Basisaustauschsatz und Korollar 3.2.17 ergibt sich, dass ein linear unabhängiges Tupel von Vektoren dann höchstens n Vektoren enthalten kann.

Beispiel 3.2.19:

1. Der $\mathbb{K}$-Vektorraum $\mathbb{K}^n$ hat Dimension $\dim_{\mathbb{K}}(\mathbb{K}^n) = n$ für alle $n \in \mathbb{N}$, denn die Standardbasis $B = \{e_1, ..., e_n\}$ aus Beispiel 3.2.5, 4. enthält n Vektoren.
2. Insbesondere gilt $\dim_{\mathbb{K}}(\mathbb{K}) = 1$, denn nach Beispiel 3.2.5, 3. ist $\{1\}$ ist eine Basis von $\mathbb{K}$. Es folgt $\dim_{\mathbb{R}}(\mathbb{R}) = \dim_{\mathbb{Q}}(\mathbb{Q}) = \dim_{\mathbb{C}}(\mathbb{C}) = 1$.
3. Es gilt $\dim_{\mathbb{R}}(\mathbb{C}) = 2$, denn nach Beispiel 3.2.5, 5. ist $\{1, \mathrm{i}\}$ eine Basis von $\mathbb{C}$.
4. Für den Vektorraum $\mathbb{K}[x]$ der Polynome mit Koeffizienten in $\mathbb{K}$ gilt $\dim_{\mathbb{K}}(\mathbb{K}[x]) = \infty$, denn nach Beispiel 3.2.8, 3. ist $\mathbb{K}[x]$ nicht endlich erzeugt.

Beispiel 3.2.20:
Für jede unendliche Menge M und jeden Körper $\mathbb{K}$ gilt $\dim_{\mathbb{K}}(\mathrm{Abb}(M, \mathbb{K})) = \infty$.

Man betrachte für $m \in M$ die Abbildungen $\delta_m : M \to \mathbb{K}$ mit $\delta_m(m) = 1$ und $\delta_m(m') = 0$ für $m \neq m'$. Sind $m_1, ..., m_n \in M$ paarweise verschieden, so folgt aus $\Sigma_{j=1}^n a_j \delta_{m_j} = 0$ auch $0 = \Sigma_{j=1}^n a_j \delta_{m_j}(m_k) = a_k$ für alle $k \in \{1, ..., n\}$, und somit ist $\{\delta_{m_1}, ..., \delta_{m_n}\}$ linear unabhängig. Da M unendlich ist, existieren zu jedem $n \in \mathbb{N}$ paarweise verschiedene Elemente $m_1, .., m_n$ und daher eine n-elementige linear unabhängige Teilmenge $\{\delta_{m_1}, ..., \delta_{m_n}\} \subseteq \mathrm{Abb}(M, \mathbb{K})$. Wäre $\mathrm{Abb}(M, \mathbb{K})$ endlich-dimensional, so gäbe es nach dem Basisaustauschsatz aber ein $k \in \mathbb{N}$, so dass alle linear unabhängigen Teilmengen maximal k Elemente enthalten.

Beispiel 3.2.21: Es gilt $\dim_{\mathbb{Q}}(\mathbb{R}) = \infty$.

Man betrachte die Vektoren $v_i = \log(p_i)$, wobei $p_i \in \mathbb{N}$ die $i$te Primzahl ist. Für alle $n \in \mathbb{N}$ ist $\{v_1, ..., v_n\} \subseteq \mathbb{R}$ linear unabhängig. Denn sind $a_1, ..., a_n \in \mathbb{Q}$ mit $a_1 v_1 + ... + a_n v_n = 0$, so erhält man durch Multiplizieren mit den kleinsten gemeinsamen Vielfachen der Nenner von $a_1, ..., a_n$ ganze Zahlen $b_1, ..., b_n \in \mathbb{Z}$ mit $b_1 v_1 + ... + b_n v_n = 0$. Es folgt

$$1 = \exp(0) = \exp(b_1 v_1 + ... + b_n v_n) = \exp(b_1 \log(p_1) + ... + b_n \log(p_n)) = p_1^{b_1} \cdots p_n^{b_n},$$

und somit $b_1 = ... = b_n = 0$. Wäre der $\mathbb{Q}$-Vektorraum $\mathbb{R}$ endlich erzeugt, so gäbe es nach dem Basisaustauschsatz aber ein $k \in \mathbb{N}$, so dass alle linear unabhängigen Teilmengen von $\mathbb{R}$ maximal k Elemente enthalten.

Korollar 3.2.22: Sei V ein endlich erzeugter $\mathbb{K}$-Vektorraum. Dann ist auch jeder Untervektorraum $W \subseteq V$ endlich erzeugt. Es gilt $\dim_{\mathbb{K}}(W) \leq \dim_{\mathbb{K}}(V)$ sowie $\dim_{\mathbb{K}}(W) = \dim_{\mathbb{K}}(V)$ genau dann, wenn $W = V$.

Beweis:
Wäre W nicht endlich erzeugt, so gäbe es nach Lemma 3.2.9 $\dim_{\mathbb{K}}(V) + 1$ linear unabhängige Vektoren in W - ein Widerspruch zum Basisaustauschsatz. Also ist W endlich erzeugt und besitzt eine endliche Basis $B = \{w_1, ..., w_m\}$. Da B linear unabhängig ist, folgt mit dem Basisaustauschsatz $m = \dim_{\mathbb{K}}(W) \leq \dim_{\mathbb{K}}(V)$. Ist $\dim_{\mathbb{K}}(W) = \dim_{\mathbb{K}}(V)$ so muss B auch eine Basis von V sein, denn sonst gäbe es ein $v \in V \setminus W$, so dass $B \cup \{v\}$ linear unabhängig wäre - ein Widerspruch zum Basisaustauschsatz. □

Korollar 3.2.23: Sei V ein endlich erzeugter $\mathbb{K}$-Vektorraum mit $n = \dim_{\mathbb{K}}(V)$. Dann bildet jedes linear unabhängige n-Tupel $(v_1, ..., v_n) \in V^{\times n}$ eine Basis $\{v_1, ..., v_n\}$ von V.

Beweis:
Die linear unabhängige Menge $B = \{v_1, ..., v_n\}$ ist eine Basis des linearen Unterraums $\mathrm{span}_{\mathbb{K}}(B)$. Also gilt $\dim_{\mathbb{K}}(\mathrm{span}_{\mathbb{K}}(B)) = \dim_{\mathbb{K}}(V) = n$, und mit Korollar 3.2.22 folgt $\mathrm{span}_{\mathbb{K}}(B) = V$. □

3.3 Summen von Untervektorräumen

Wir beschäftigen uns nun noch mit der Frage, wie sich ein gegebener $\mathbb{K}$-Vektorraum V aus Untervektorräumen $W_1, W_2, ..., W_n \subseteq V$ aufbauen oder sich in solche Untervektorräume zerlegen lässt. Insbesondere ergibt sich die Frage, wie man aus Basen verschiedener Untervektorräume eine Basis von V gewinnt. Da jedes skalare Vielfache λw eines Vektors $w \in W_i$ mit $\lambda \in \mathbb{K}$ wieder in W_i enthalten ist, müssen wir dabei keine allgemeinen Linearkombinationen von Elementen aus $W_1, \ldots, W_n$ betrachten, sondern können uns auf Summen solcher Vektoren beschränken.

Definition 3.3.1: Sei V ein $\mathbb{K}$-Vektorraum und seien $W_1, W_2, ..., W_n \subseteq V$ Untervektorräume. Dann ist die **Summe** $W_1 + ... + W_n$ der Untervektorräume $W_1, ..., W_n \subseteq V$

$$W_1 + ... + W_n = \{w_1 + ... + w_n | w_1 \in W_1, ..., w_n \in W_n\}.$$

Lässt sich jedes Element $v \in W_1 + ... + W_n$ als Summe $v = w_1 + ... + w_n$ mit *eindeutig* bestimmten $w_1 \in W_1$, ... ,$w_n \in W_n$ schreiben, so spricht man von einer **(inneren) direkten Summe** und schreibt $W_1 \oplus \ldots \oplus W_n$ statt $W_1 + \ldots + W_n$.

Wir zeigen jetzt, dass Summen von Untervektorräumen $W_1, ..., W_n \subseteq V$ wieder Untervektorräume von V sind und dass die Dimension der Summe höchstens so groß ist wie die Summe ihrer Dimensionen. Dies folgt im Wesentlichen daraus, dass eine Summe $W_1 + ... + W_n$ nichts anderes ist als der Spann der Menge $W_1 \cup ... \cup W_n$.

Satz 3.3.2: Sei V ein $\mathbb{K}$-Vektorraum und $W_1, ..., W_n \subseteq V$ Untervektorräume. Dann gilt:

1. $W_1 + ... + W_n = \mathrm{span}_{\mathbb{K}}(W_1 \cup ... \cup W_n)$ ist ein Untervektorraum von V.
2. $\dim_{\mathbb{K}}(W_1 + ... + W_n) \leq \dim_{\mathbb{K}}(W_1) + \ldots + \dim_{\mathbb{K}}(W_n)$, wobei $\infty + \infty := \infty$, $\infty + n := \infty$ und $n < \infty$.

Beweis:
1. Es gilt per Definition $W_1+...+W_n \subseteq \mathrm{span}_{\mathbb{K}}(W_1 \cup ... \cup W_n)$. Sei nun $v \in \mathrm{span}_{\mathbb{K}}(W_1 \cup ... \cup W_n)$. Dann gibt es Vektoren $v_1^i, ..., v_{m_i}^i \in W_i$ und Skalare $\lambda_1^i, ..., \lambda_{m_i}^i \in \mathbb{K}$ mit $v = \Sigma_{i=1}^n(\Sigma_{j=1}^{m_i}\lambda_j^i v_j^i)$. Definiert man für $i = 1, ..., n$ die Vektoren $w_i = \Sigma_{j=1}^{m_i}\lambda_j^i v_j^i$, so gilt $w_i \in W_i$ und $v = w_1 + ... + w_n$, also $v \in W_1 + ... + W_n$. Damit ist gezeigt, dass auch $W_1 + ... + W_n \supseteq \mathrm{span}_{\mathbb{K}}(W_1 \cup ... \cup W_n)$ gilt.

2. Gilt $\dim_{\mathbb{K}}(W_i) = \infty$ für ein $i \in \{1, ..., n\}$, so ist die Aussage wahr. Sind alle Unterräume W_i endlich-dimensional, so existieren endliche Basen B_i von W_i mit jeweils $\dim_{\mathbb{K}}(W_i)$ Elementen. Dann enthält $B = B_1 \cup ... \cup B_n$ höchstens $\dim_{\mathbb{K}}(W_1) + ... + \dim_{\mathbb{K}}(W_n)$ Elemente und erzeugt $W_1 + ... + W_n$, denn $\mathrm{span}_{\mathbb{K}}(B) = \mathrm{span}_{\mathbb{K}}(B_1 \cup ... \cup B_n) = \mathrm{span}_{\mathbb{K}}(W_1 \cup ... \cup W_n) = W_1 + ... + W_n$. Daraus folgt $\dim_{\mathbb{K}}(W_1 + ... + W_n) \leq \dim_{\mathbb{K}}(W_1) + ... + \dim_{\mathbb{K}}(W_n)$. □

Betrachtet man nur zwei endlich-dimensionale Vektorräume, so kann man eine Formel herleiten, die die Dimension einer Summe $W_1 + W_2$, zu den Dimensionen der Untervektorräume $W_1 \cap W_2$, W_1 und W_2 in Beziehung setzt. Dazu ergänzt man eine Basis des Untervektorraums $W_1 \cap W_2$ zu Basen der Untervektorräume W_1 und W_2 und vereinigt diese zu einer Basis von $W_1 + W_2$.

Satz 3.3.3: Sei V ein $\mathbb{K}$-Vektorraum und $W_1, W_2 \subseteq V$ endlich-dimensionale Untervektorräume. Dann gilt: $\dim_{\mathbb{K}}(W_1 + W_2) = \dim_{\mathbb{K}}(W_1) + \dim_{\mathbb{K}}(W_2) - \dim_{\mathbb{K}}(W_1 \cap W_2)$.

Beweis:
Da W_1 und W_2 endlich-dimensional sind, ist auch $W_1 \cap W_2 \subseteq W_1$ endlich-dimensional nach Korollar 3.2.22. Sei $B_{12} = \{u_1, ..., u_m\}$ eine Basis von $W_1 \cap W_2$. Dann existieren nach dem Basisergänzungssatz $v_1, ..., v_r \in W_1 \setminus W_2$, so dass $B_1 = \{u_1, ..., u_m, v_1, ..., v_r\}$ eine Basis von W_1 ist und $w_1, ..., w_s \in W_2 \setminus W_1$, so dass $B_2 = \{u_1, ..., u_m, w_1, ..., w_s\}$ eine Basis von W_2 ist.

Wir zeigen, dass $B = B_1 \cup B_2 = \{u_1, ..., u_m, v_1, ..., v_r, w_1, ..., w_s\}$ eine Basis von $W_1 + W_2$ ist. Nach dem Beweis von Satz 3.3.2 ist B ein Erzeugendensystem von $W_1 + W_2$. Zum Nachweis der linearen Unabhängigkeit, betrachten wir $\lambda_1, ..., \lambda_m$, $\mu_1, ..., \mu_r$, $\nu_1, ..., \nu_s \in \mathbb{K}$ mit

$$\underbrace{\Sigma_{j=1}^m \lambda_j u_j}_{=:u \in W_1 \cap W_2} + \underbrace{\Sigma_{k=1}^r \mu_k v_k}_{=:v \in W_1} + \underbrace{\Sigma_{l=1}^s \nu_l w_l}_{=:w \in W_2} = u + v + w = 0.$$

Dann folgt aus $u + v \in W_1$ und $w \in W_2$, dass $u + v = -w \in W_1 \cap W_2$. Da B_1 eine Basis von W_1 ist und B_{12} eine Basis von $W_1 \cap W_2$, folgt $\mu_1 = ... = \mu_r = 0$ und $v = 0$. Daraus ergibt sich $w = -u \in W_1 \cap W_2$. Da B_2 eines Basis von W_2 ist und B_{12} eine Basis von $W_1 \cap W_2$, folgt $\nu_1 = ... = \nu_s = 0$ und $w = 0$. Also gilt auch $u = 0$. Da B_{12} eine Basis von $W_1 \cap W_2$ ist, folgt $\lambda_1 = ... = \lambda_m = 0$. Damit ist B auch linear unabhängig und eine Basis von $W_1 + W_2$. Es folgt

$$\dim_{\mathbb{K}}(W_1 + W_2) = m + r + s = (m + r) + (m + s) - m = \dim_{\mathbb{K}}(W_1) + \dim_{\mathbb{K}}(W_2) - \dim_{\mathbb{K}}(W_1 \cap W_2).$$

□

Wir verschaffen uns noch ein leicht überprüfbares Kriterium, mit dem wir feststellen können, wann eine gegebene Summe $W_1 + W_2$ von zwei Untervektorräumen eine direkte Summe ist. Es stellt sich heraus, dass das genau dann der Fall ist, wenn $W_1 \cap W_2 = \{0\}$ gilt.

Satz 3.3.4: Sei V ein $\mathbb{K}$-Vektorraum und $W_1, W_2 \subseteq V$ Untervektorräume. Dann gilt $V = W_1 \oplus W_2$ genau dann, wenn $V = W_1 + W_2$ und $W_1 \cap W_2 = \{0\}$.

Beweis:
$\Rightarrow$: Ist $V = W_1 \oplus W_2$, so lässt sich jedes $v \in V$ eindeutig schreiben als $v = w_1 + w_2$ mit $w_1 \in W_1$, $w_2 \in W_2$. Für jedes $v \in W_1 \cap W_2$, gibt es aber zwei solche Darstellungen, nämlich $v = v + 0 = 0 + v$, die gleich sind genau dann, wenn $v = 0$. Also muss $W_1 \cap W_2 = \{0\}$ gelten.

$\Leftarrow$: Ist umgekehrt $V = W_1 + W_2$ mit $W_1 \cap W_2 = \{0\}$, so lässt sich jedes $v \in V$ schreiben als $v = w_1 + w_2$ mit $w_i \in W_i$. Ist $v = w_1' + w_2'$ eine weitere solche Darstellung, so folgt $w_1 - w_1' = w_2' - w_2 \in W_1 \cap W_2 = \{0\}$, also $w_1 = w_1'$ und $w_2' = w_2$. $\square$

Beispiel 3.3.5: Wir betrachten $V = \mathbb{K}^3$ mit der Standardbasis und die Untervektorräume

$$W_1 = \mathrm{span}_{\mathbb{K}}(\{e_1, e_2\}), \qquad W_2 = \mathrm{span}_{\mathbb{K}}(\{e_3\}), \qquad W_3 = \mathrm{span}_{\mathbb{K}}(\{e_2, e_3\}).$$

Dann gilt $W_1 \cap W_2 = \{0\}$, $W_1 \cap W_3 = \mathbb{K}e_2$ und $W_2 \cap W_3 = \mathbb{K}e_3$. Daraus folgt $\mathbb{K}^3 = W_1 \oplus W_2$. Ebenso gilt $\mathbb{K}^3 = W_1 + W_3$, aber die Summe $W_1 + W_3$ ist nicht direkt.

Mit den Dimensionen $\dim_{\mathbb{K}}(W_2) = \dim_{\mathbb{K}}(W_1 \cap W_3) = \dim_{\mathbb{K}}(W_2 \cap W_3) = 1$, $\dim_{\mathbb{K}}(W_1 \cap W_2) = 0$ und $\dim_{\mathbb{K}}(W_1) = \dim_{\mathbb{K}}(W_3) = 2$ ergibt sich

$$\begin{aligned} 3 = \dim_{\mathbb{K}}(\mathbb{K}^3) &= \dim_{\mathbb{K}}(W_1 + W_2) = \dim_{\mathbb{K}}(W_1) + \dim_{\mathbb{K}}(W_2) - \dim_{\mathbb{K}}(W_1 \cap W_2) = 2 + 1 - 0 \\ &= \dim_{\mathbb{K}}(W_1 + W_3) = \dim_{\mathbb{K}}(W_1) + \dim_{\mathbb{K}}(W_3) - \dim_{\mathbb{K}}(W_1 \cap W_3) = 2 + 2 - 1. \end{aligned}$$

Man beachte, dass sich die Aussage von Satz 3.3.4 nicht auf Summen von mehr als zwei Untervektorräumen verallgemeinert. Ist $n > 2$ und sind $W_1, \ldots, W_n \subseteq V$ Untervektorräume mit $W_i \cap W_j = \{0\}$ für $i \neq j$, so muss die Summe $W_1 + \ldots + W_n$ trotzdem nicht direkt sein. Ein Besipiel sind die Untervektorräume $W_1 = \mathbb{K}e_1$, $W_2 = \mathbb{K}e_2$ und $W_3 = \mathbb{K}(e_1 + e_2)$ des Vektorraums $\mathbb{K}^2$. Hier gilt offensichtlich $W_1 \cap W_2 = W_1 \cap W_3 = W_2 \cap W_3 = \{0\}$, aber die Summe $W_1 + W_2 + W_3$ ist nicht direkt, da der Vektor $v = e_1 + e_2$ auf verschiedene Weisen als Linearkombination der Vektoren $e_1, e_2, e_1 + e_2$ geschrieben werden kann: $v = 1 \cdot e_1 + 1 \cdot e_2 + 0 \cdot (e_1 + e_2) = 0 \cdot e_1 + 0 \cdot e_2 + 1 \cdot (e_1 + e_2)$.

An was man sich erinnern sollte:

Die wichtigsten Begriffe und Definitionen:

- Vektorraum, Vektor, Untervektorraum, Spann, Erzeugendensystem
- linear (un)abhängig, Basis, Dimension
- Summe und direkte Summe von Untervektorräumen

Die wichtigsten Aussagen:

- Kriterien für lineare Unabhängigkeit, Kriterien für Basen,
- Basisauswahlsatz, Basisaustauschsatz und Basisergänzungssatz,
- Kriterien für die Direktheit von Summen und Dimensionsformeln für Summen,

Die wichtigsten Beispiele:

- **Vektorräume:** $\{0\}$, $\mathbb{K}^n$ für $\mathbb{K} = \mathbb{Q}, \mathbb{R}, \mathbb{C}, \mathbb{Z}/p\mathbb{Z}$, Abbildungen $\mathrm{Abb}(M, V)$ in einen Vektorraum V, Folgenraum $\mathrm{Abb}(\mathbb{N}_0, V)$, Polynomraum $\mathbb{K}[x]$, Körper $\mathbb{L}$ über Teilkörper $\mathbb{K} \subseteq \mathbb{L}$
- **Untervektorräume:** $\{0\} \subseteq V$, $V \subseteq V$, $\mathrm{span}_{\mathbb{K}}(M)$ für Teilmengen $M \subseteq V$, Geraden, Ebenen, $\mathbb{K}[x] \subseteq \mathrm{Abb}(\mathbb{N}_0, \mathbb{K})$, periodische Folgen, Teilkörper $\mathbb{K} \subseteq \mathbb{L}$
- **Basen:** $\emptyset \subseteq \{0\}$, Standardbasis des $\mathbb{K}^n$, Monome für $\mathbb{K}[x]$, $\{1\} \subseteq \mathbb{K}$.

4 Lineare Abbildungen und Matrizen

4.1 Lineare Abbildungen

Wir betrachten nun die zu dem Begriff des Vektorraums gehörenden strukturerhaltenden Abbildungen, die $\mathbb{K}$-*linearen Abbildungen* oder *Homomorphismen von* $\mathbb{K}$-*Vektorräumen*. Wie im Fall einer Gruppe und eines Rings, wo jeweils gefordert wurde, dass die zugehörigen strukturerhaltenden Abbildungen (Gruppenhomomorphismen und Ringhomomorphismen) mit den Verknüpfungen verträglich sind, werden auch im Fall des Vektorraums Verträglichkeitsbedingungen auferlegt, nämlich mit der Vektoraddition und der Skalarmultiplikation.

Die Verträglichkeitsbedingung mit der Vektoraddition ergibt sich dabei aus der Tatsache, dass $(V, +)$ für jeden Vektorraum $(V, +, \cdot)$ eine abelsche Gruppe ist. Man fordert also, dass eine $\mathbb{K}$-lineare Abbildung ein Gruppenhomomorphismus ist. Bezüglich der Skalarmultiplikation ist die einzig sinnvolle Forderung, dass es keinen Unterschied machen sollte, ob man einen Vektor zuerst mit einem Skalar multipliziert und dann abbildet oder zuerst abbildet und anschließend mit einem Skalar multipliziert. Dies liefert die folgende Definition.

Definition 4.1.1: Sei $\mathbb{K}$ ein Körper und V, W Vektorräume über $\mathbb{K}$. Eine Abbildung $\phi : V \to W$ heißt $\mathbb{K}$-**lineare Abbildung** oder $\mathbb{K}$-**Vektorraumhomomorphismus**, wenn gilt:

(L1) $\phi : (V, +_V) \to (W, +_W)$ ist ein Gruppenhomomorphismus:

$$\phi(v_1 +_V v_2) = \phi(v_1) +_W \phi(v_2) \qquad \forall v_1, v_2 \in V.$$

(L2) $\phi : V \to W$ ist mit der Skalarmultiplikation verträglich:

$$\phi(\lambda v) = \lambda\phi(v) \qquad \forall v \in V, \lambda \in \mathbb{K}.$$

Für $\mathbb{K}$-lineare Abbildungen $\phi : V \to W$ benutzt man auch die folgenden Bezeichnungen:

- **Monomorphismus von $\mathbb{K}$-Vektorräumen**, wenn $\phi : V \to W$ injektiv ist,
- **Epimorphismus von $\mathbb{K}$-Vektorräumen**, wenn $\phi : V \to W$ surjektiv ist,
- **Isomorphismus von $\mathbb{K}$-Vektorräumen**, wenn $\phi : V \to W$ bijektiv ist,
- **Endomorphismus von $\mathbb{K}$-Vektorräumen**, wenn $V = W$,
- **Automorphismus von $\mathbb{K}$-Vektorräumen**, wenn $V = W$ und $\phi : V \to V$ bijektiv ist.

Existiert ein $\mathbb{K}$-Vektorraumisomorphismus $\phi : V \to W$ dann nennt man die $\mathbb{K}$-Vektorräume V und W **isomorph** und schreibt $V \cong W$.

Bemerkung 4.1.2:

1. Da eine lineare Abbildung $\phi : V \to W$ ein Gruppenhomomorphismus bezüglich der Vektoraddition ist, bildet sie neutrale Elemente auf neutrale Elemente und Inverse auf Inverse ab: $\phi(0_V) = 0_W$ und $\phi(-v) = -\phi(v)$ für alle $v \in V$.

2. Eine Abbildung $\phi : V \to W$ zwischen $\mathbb{K}$-Vektorräumen V, W ist genau dann $\mathbb{K}$-linear, wenn sie Linearkombinationen auf Linearkombinationen abbildet:
$\phi(\Sigma_{j=1}^n \lambda_j v_j) = \Sigma_{j=1}^n \lambda_j \phi(v_j)$ für alle $n \in \mathbb{N}$, $v_1, ..., v_n \in V$ und $\lambda_1, ..., \lambda_n \in \mathbb{K}$.

Dies zeigt man mit vollständiger Induktion über n. Für $n = 1$ folgt $\phi(\lambda_1 v_1) = \lambda_1\phi(v_1)$ aus (L2). Ist die Aussage wahr für alle Linearkombinationen $\Sigma_{j=1}^{m}\lambda_j v_j$ mit $m \leq n$, so folgt

$$\begin{aligned}&\phi(\Sigma_{j=1}^{n+1}\lambda_j v_j) = \phi(\lambda_{n+1}v_{n+1} + \Sigma_{j=1}^{n}\lambda_j v_j) \overset{(L1)}{=} \phi(\lambda_{n+1}v_{n+1}) + \phi(\Sigma_{j=1}^{n}\lambda_j v_j)\\ &\overset{(L2)}{=} \lambda_{n+1}\phi(v_{n+1}) + \phi(\Sigma_{j=1}^{n}\lambda_j v_j) \overset{IV}{=} \lambda_{n+1}\phi(v_{n+1}) + \Sigma_{j=1}^{n}\lambda_j\phi(v_j) = \Sigma_{j=1}^{n+1}\lambda_j\phi(v_j).\end{aligned}$$

3. Nach 2. ist eine Abbildung $\phi : V \to W$ genau dann $\mathbb{K}$-linear, wenn sie lineare Hüllen auf lineare Hüllen abbildet: $\phi(\text{span}_{\mathbb{K}}(M)) = \text{span}_{\mathbb{K}}(\phi(M))$ für alle Teilmengen $M \subseteq V$.

Beispiel 4.1.3:

1. Für jeden $\mathbb{K}$-Vektorraum V und festes $\mu \in \mathbb{K}$ ist die **Streckung** $D_\mu : V \to V$, $v \mapsto \mu v$ eine $\mathbb{K}$-lineare Abbildung. Denn wegen (V1) und (V3) gilt:

$$\begin{aligned}&D_\mu(v + v') = \mu(v + v') = \mu v + \mu v' = D_\mu(v) + D_\mu(v')\\ &D_\mu(\lambda v) = \mu(\lambda v) = (\mu\lambda)v = \lambda(\mu v) = \lambda D_\mu(v) \qquad \forall v, v' \in V, \lambda \in \mathbb{K}.\end{aligned}$$

   Für $V = \mathbb{R}^3$, $V = \mathbb{R}^2$ oder $V = \mathbb{R}$ sieht man, dass es sich dabei um eine Streckung handelt.

2. Die komplexe Konjugation $\bar{\ } : \mathbb{C} \to \mathbb{C}$ ist eine $\mathbb{R}$-lineare Abbildung, denn es gilt $\overline{z + w} = \overline{z} + \overline{w}$, $\overline{\lambda z} = \lambda\overline{z}$ für alle $\lambda \in \mathbb{R}$ und $z, w \in \mathbb{C}$. Sie ist aber nicht $\mathbb{C}$-linear.

3. Im $\mathbb{R}^2$ ist die **Drehung** um den Ursprung um einen Winkel $\beta \in \mathbb{R}$ gegeben durch

$$R_\beta : \mathbb{R}^2 \mapsto \mathbb{R}^2, \quad \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x\cos\beta - y\sin\beta \\ y\cos\beta + x\sin\beta \end{pmatrix}$$

   Diese Abbildung ist $\mathbb{R}$-linear, denn es gilt

$$\begin{aligned}&R_\beta\left(\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} x' \\ y' \end{pmatrix}\right) = R_\beta\begin{pmatrix} x + x' \\ y + y' \end{pmatrix} = \begin{pmatrix} (x + x')\cos\beta - (y + y')\sin\beta \\ (y + y')\cos\beta + (x + x')\sin\beta \end{pmatrix}\\ &= \begin{pmatrix} x\cos\beta - y\sin\beta \\ y\cos\beta + x\sin\beta \end{pmatrix} + \begin{pmatrix} x'\cos\beta - y'\sin\beta \\ y'\cos\beta + x'\sin\beta \end{pmatrix} = R_\beta\begin{pmatrix} x \\ y \end{pmatrix} + R_\beta\begin{pmatrix} x' \\ y' \end{pmatrix}\\ &R_\beta\left(\lambda\begin{pmatrix} x \\ y \end{pmatrix}\right) = R_\beta\begin{pmatrix} \lambda x \\ \lambda y \end{pmatrix} = \begin{pmatrix} \lambda x\cos\beta - \lambda y\sin\beta \\ \lambda y\cos\beta + \lambda x\sin\beta \end{pmatrix} = \lambda\begin{pmatrix} x\cos\beta - y\sin\beta \\ y\cos\beta + x\sin\beta \end{pmatrix}\\ &= \lambda R_\beta\begin{pmatrix} x \\ y \end{pmatrix}.\end{aligned}$$

4. Für jedes $n \in \mathbb{N}$ sind die **Koordinatenprojektionen** $\pi_i : \mathbb{K}^n \to \mathbb{K}^n$, $\Sigma_{k=1}^{n}\lambda_k e_k \mapsto \lambda_i e_i$ $\mathbb{K}$-lineare Abbildungen, denn es gilt

$$\begin{aligned}&\pi_i(\Sigma_{k=1}^{n}\lambda_k e_k + \Sigma_{k=1}^{n}\mu_k e_k) = \pi_i(\Sigma_{k=1}^{n}(\lambda_k + \mu_k)e_k) = (\lambda_i + \mu_i)e_i = \pi_i(\Sigma_{k=1}^{n}\lambda_k e_k) + \pi_i(\Sigma_{k=1}^{n}\mu_k e_k)\\ &\pi_i(\lambda\Sigma_{k=1}^{n}\lambda_k e_k) = \pi_i(\Sigma_{k=1}^{n}(\lambda\lambda_k)e_k + \Sigma_{k=1}^{n}\mu_k e_k) = \lambda\lambda_i e_i = \lambda\pi_i(\Sigma_{k=1}^{n}\lambda_k e_k).\end{aligned}$$

5. Die Abbildung $d : \mathbb{K}[x] \to \mathbb{K}[x]$, $\Sigma_{k=0}^{\infty}a_k x^k \mapsto \Sigma_{k=0}^{\infty}k a_k x^{k-1}$ ist $\mathbb{K}$-linear. Sie heißt **formale Ableitung** von Polynomen. Man berechnet für $p = \Sigma_{k=0}^{n}a_k x^k, q = \Sigma_{k=0}^{n}b_k x^k \in \mathbb{K}[x]$

$$\begin{aligned}&d(p + q) = \Sigma_{k=0}^{\infty}k(a_k + b_k)x^{k-1} = \Sigma_{k=0}^{\infty}k a_k x^{k-1} + \Sigma_{k=0}^{\infty}k b_k x^{k-1} = d(p) + d(q)\\ &d(\lambda p) = \Sigma_{k=0}^{\infty}k\lambda a_k x^{k-1} = \lambda\Sigma_{k=0}^{\infty}k a_k x^{k-1} = \lambda d(p).\end{aligned}$$

Eine wichtige Eigenschaft von $\mathbb{K}$-linearen Abbildungen zwischen zwei gegebenen $\mathbb{K}$-Vektorräumen V, W ist, dass die punktweise Summe zweier $\mathbb{K}$-linearer Abbildungen und das punktweise Produkt einer $\mathbb{K}$-linearen Abbildung mit einem Skalar wieder $\mathbb{K}$-lineare Abbildungen sind. Die Menge der $\mathbb{K}$-linearen Abbildungen zwischen V und W ist also abgeschlossen unter der punktweisen Addition und punktweisen Multiplikation mit Skalaren und bildet somit einen Untervektorraum des $\mathbb{K}$-Vektorraums $\mathrm{Abb}(V, W)$ aus Lemma 3.1.6.

Satz 4.1.4: Seien V, W Vektorräume über einem Körper $\mathbb{K}$. Dann bilden die $\mathbb{K}$-linearen Abbildungen $\phi : V \to W$ mit der punktweisen Addition und Skalarmultiplikation

$$(\phi + \psi)(v) = \phi(v) + \psi(v) \qquad (\lambda\phi)(v) = \lambda\phi(v) \qquad \forall v \in V, \lambda \in \mathbb{K}$$

einen Untervektorraum des $\mathbb{K}$-Vektorraums $\mathrm{Abb}(V, W)$. Dieser wird mit $\mathrm{Hom}_{\mathbb{K}}(V, W)$ und für $V = W$ auch mit $\mathrm{End}_{\mathbb{K}}(V)$ bezeichnet.

Beweis:
Die Nullabbildung $0 : V \to W$, $v \mapsto 0_W$ ist wegen $0_W + 0_W = 0_W$ und $\lambda 0_W = 0_W$ für alle $\lambda \in \mathbb{K}$ eine $\mathbb{K}$-lineare Abbildung (UV1). Sind $\phi, \psi : V \to W$ $\mathbb{K}$-lineare Abbildungen, so gilt

$$\begin{aligned}
&(\phi + \psi)(v + v') = \phi(v + v') + \psi(v + v') = \phi(v) + \phi(v') + \psi(v) + \psi(v') = (\phi + \psi)(v) + (\phi + \psi)(v') \\
&(\phi + \psi)(\lambda v) = \phi(\lambda v) + \psi(\lambda v) = \lambda\phi(v) + \lambda\psi(v) = \lambda(\phi + \psi)(v) \qquad \forall v, v' \in V, \lambda \in \mathbb{K}
\end{aligned}$$

und somit ist auch $\phi + \psi : V \to W$ eine $\mathbb{K}$-lineare Abbildung (UV2). Für alle $\mu \in \mathbb{K}$ gilt

$$\begin{aligned}
&(\mu\phi)(v + v') = \mu\phi(v + v') = \mu(\phi(v) + \phi(v')) = \mu\phi(v) + \mu\phi(v') = (\mu\phi)(v) + (\mu\phi)(v') \\
&(\mu\phi)(\lambda v) = \mu\phi(\lambda v) = \mu\lambda\phi(v) = \lambda(\mu\phi)(v) \qquad \forall v, v' \in V, \lambda \in \mathbb{K}
\end{aligned}$$

und damit ist auch $\mu\phi : V \to W$ eine $\mathbb{K}$-lineare Abbildung (UV3). □

Für Gruppenhomomorphismen und Ringhomomorphismen wurde gezeigt, dass die Identitätsabbildung ein Gruppen- oder Ringautomorphismus ist, die Umkehrabbildung eines Gruppen- oder Ringisomorphismus wieder ein Gruppen- oder Ringisomorphismus ist, und die Verkettung zweier Gruppen- oder Ringhomomorphismen ein Gruppen- oder Ringhomomorphismus ist.

Wir zeigen nun die analogen Aussagen für $\mathbb{K}$-Vektorräume, nämlich dass die Identitätsabbildung ein Automorphismus von $\mathbb{K}$-Vektorräumen ist, für jeden Isomorphismus von $\mathbb{K}$-Vektorräumen auch die Umkehrabbildung wieder ein Isomorphismus von $\mathbb{K}$-Vektorräumen ist und die Verkettung zweier Homomorphismen von $\mathbb{K}$-Vektorräumen ein Homomorphismus von $\mathbb{K}$-Vektorräumen ist. Dies entspricht der Aussage, dass *isomorph sein zu* eine Äquivalenzrelation für Vektorräume[10] ist. Die drei Aussagen garantieren ihre Reflexivität, Symmetrie und Transitivität.

Satz 4.1.5: Sei $\mathbb{K}$ ein Körper und U, V, W Vektorräume über $\mathbb{K}$. Dann gilt:

1. Die Identitätsabbildung $\mathrm{id}_V : V \to V$ ist ein $\mathbb{K}$-Vektorraumautomorphismus.
2. Ist $\psi : V \to W$ ein $\mathbb{K}$-Vektorraumisomorphismus, so ist auch die Umkehrabbildung $\psi^{-1} : W \to V$ ein $\mathbb{K}$-Vektorraumisomorphismus.
3. Sind $\phi : U \to V$ und $\psi : V \to W$ $\mathbb{K}$-lineare Abbildungen, so ist auch die Verkettung $\psi \circ \phi : U \to W$ eine $\mathbb{K}$-lineare Abbildung.

[10] Achtung! Die Vektorräume bilden wie auch die Gruppen oder Ringe keine Menge. Es handelt sich also wieder um eine Äquivalenzrelation in einem etwas allgemeineren Sinn, nämlich auf einer Klasse.

Beweis:
In Satz 2.2.23 wurde schon gezeigt, dass id_V, ψ^{-1} und $\psi\circ\phi$ Gruppenhomomorphismen bezüglich der Vektoraddition sind. Zu zeigen ist die Verträglichkeit mit der Skalarmultiplikation.

1. Es gilt $\mathrm{id}_V(\lambda v) = \lambda v = \lambda \mathrm{id}_V(v)$ für alle $v \in V$, $\lambda \in \mathbb{K}$.

2. Ist $\psi : V \to W$ eine bijektive lineare Abbildung so folgt

$$\psi^{-1}(\lambda w) = \psi^{-1}(\lambda\psi(\psi^{-1}(w))) = \psi^{-1}(\psi(\lambda\psi^{-1}(w))) = \lambda\psi^{-1}(w) \qquad \forall w \in W, \lambda \in \mathbb{K}.$$

3. Sind $\phi : U \to V$ und $\psi : V \to W$ $\mathbb{K}$-lineare Abbildungen so folgt

$$(\psi \circ \phi)(\lambda u) = \psi(\phi(\lambda u)) = \psi(\lambda\phi(u)) = \lambda\psi(\phi(u)) = \lambda(\psi \circ \phi)(u) \qquad \forall u \in U, \lambda \in \mathbb{K}.$$ □

Im Gegensatz zu Satz 2.2.23 für Gruppen und Satz 2.3.10 für Ringe, wird in Satz 4.1.5 nicht erwähnt, dass die Automorphismen eines $\mathbb{K}$-Vektorraums V eine Gruppe bilden. Diese Aussage ist wahr, sie ergibt sich aber im Fall der Vektorräume aus einer allgemeineren Aussage.

Betrachten wir die Menge $\mathrm{End}_\mathbb{K}(V)$ der $\mathbb{K}$-linearen Abbildungen eines gegebenen $\mathbb{K}$-Vektorraum V in sich selbst, so erhalten wir zwei Verknüpfungen auf $\mathrm{End}_\mathbb{K}(V)$, nämlich die punktweise Addition und die Verkettung von $\mathbb{K}$-linearen Abbildungen. Die Menge $\mathrm{End}_\mathbb{K}(V)$ bildet mit diesen Verknüpfungen einen Rings einen unitalen Ring. Die Einheitengruppe dieses Rings aus Korollar 2.2.14 enthält genau die Automorphismen von V.

Satz 4.1.6: Sei V ein Vektorraum über $\mathbb{K}$.

1. Die $\mathbb{K}$-linearen Abbildungen $\phi : V \to V$ bilden mit der Verkettung und der punktweisen Addition einen unitalen Ring $(\mathrm{End}_\mathbb{K}(V), +, \circ)$ den **Endomorphismenring** von V.
2. Die Einheitengruppe des Monoids $(\mathrm{End}_\mathbb{K}(V), \circ)$ besteht genau aus den linearen Automorphismen: $\mathrm{End}_\mathbb{K}(V)^\times = \mathrm{Aut}_\mathbb{K}(V)$.

Beweis:
In Satz 4.1.4 wurde bereits gezeigt, dass $(\mathrm{End}_\mathbb{K}(V), +)$ eine abelsche Gruppe ist. In Satz 4.1.5 wurde gezeigt, dass die Verkettung von $\mathbb{K}$-linearen Abbildungen wieder $\mathbb{K}$-linear ist und die Identitätsabbildung eine $\mathbb{K}$-lineare Abbildung ist. Da die Verkettung von Abbildungen auch assoziativ ist, ist $(\mathrm{End}_\mathbb{K}(V), \circ)$ ein Monoid mit neutralem Element id_V.

Zu zeigen ist noch das Distributivgesetz. Für alle $v \in V$ und $\phi, \psi, \chi \in \mathrm{End}_\mathbb{K}(V)$ gilt:

$$((\psi + \phi) \circ \chi)(v) = (\psi + \phi)(\chi(v)) = \psi(\chi(v)) + \phi(\chi(v)) = (\psi \circ \chi + \phi \circ \chi)(v),$$
$$(\chi \circ (\psi + \phi))(v) = \chi((\phi + \psi)(v)) = \chi(\phi(v) + \psi(v)) = \chi(\phi(v)) + \chi(\psi(v)) = (\chi \circ \phi + \chi \circ \phi)(v),$$

also $(\psi + \phi) \circ \chi = \psi \circ \chi + \phi \circ \chi$ und $\chi \circ (\phi + \psi) = \chi \circ \phi + \chi \circ \psi$. Damit ist gezeigt, dass $(\mathrm{End}_\mathbb{K}(V), +, \circ)$ ein unitaler Ring ist.

Die Einheiten in $(\mathrm{End}_\mathbb{K}(V), \circ)$ sind per Definition die invertierbaren, also die bijektiven $\mathbb{K}$-linearen Abbildungen $\phi : V \to V$ und damit die Automorphismen. □

Kombiniert man Satz 4.1.6 mit Satz 4.1.4, so sieht man, dass die Menge $\mathrm{End}_{\mathbb{K}}(V)$ sowohl die Struktur eines $\mathbb{K}$-Vektorraums als auch die Struktur eines unitalen Rings hat, wobei in beiden Fällen die Addition durch die punktweise Addition von Endomorphismen gegeben ist, die Ringmultiplikation durch die Verkettung und die Skalarmultiplikation durch die punktweise Multiplikation mit Skalaren. Die Skalarmultiplikation ist verträglich mit der Verkettung, denn es gilt für alle $\lambda \in \mathbb{K}$, $\phi, \psi \in \mathrm{End}_{\mathbb{K}}(V)$ und $v \in V$

$$((\lambda\psi) \circ \phi)(v) = \lambda\psi(\phi(v)) = \lambda(\psi \circ \phi)(v) = \psi(\lambda\phi(v)) = (\psi \circ (\lambda\phi))(v),$$

also $(\lambda\psi) \circ \phi = \psi \circ (\lambda\phi) = \lambda(\psi \circ \phi)$. Eine solche Struktur bezeichnet man als eine $\mathbb{K}$-Algebra.

Definition 4.1.7: Sei $\mathbb{K}$ ein Körper.

1. Ein $\mathbb{K}$-Vektorraum $(V, +, \cdot)$ zusammen mit einer Verknüpfung $\circ : V \times V \to V$ heißt **Algebra** über $\mathbb{K}$ oder $\mathbb{K}$-**Algebra**, wenn $(V, +, \circ)$ ein unitaler Ring ist und die Verknüpfung $\circ$ mit der Skalarmultiplikation $\cdot$ verträglich ist:

$$(\lambda \cdot v) \circ w = v \circ (\lambda \cdot w) = \lambda \cdot (v \circ w) \qquad \forall v, w \in V,\ \lambda \in \mathbb{K}.$$

2. Eine Abbildung $\phi : A \to B$ zwischen $\mathbb{K}$-Algebren A und B heißt **Homomorphismus von $\mathbb{K}$-Algebren** oder $\mathbb{K}$-**Algebrahomomorphismus**, wenn sie gleichzeitig eine $\mathbb{K}$-lineare Abbildung und ein unitaler Ringhomomomorphismus ist.

3. Ist $A = B$ so spricht man auch von einem **Endomorphismus von $\mathbb{K}$-Algebren**. Ein bijektiver Homomorphismus von $\mathbb{K}$-Algebren heißt auch **Isomorphismus von $\mathbb{K}$-Algebren** und für $A = B$ auch **Automorphismus von $\mathbb{K}$-Algebren**.

Beispiel 4.1.8:

1. Für jeden $\mathbb{K}$-Vektorraum V ist $(\mathrm{End}_{\mathbb{K}}(V), +, \cdot, \circ)$ eine $\mathbb{K}$-Algebra.
2. Ist $\phi : V \to W$ eine bijektive $\mathbb{K}$-lineare Abbildung, so ist

$$F : \mathrm{End}_{\mathbb{K}}(V) \to \mathrm{End}_{\mathbb{K}}(W),\ \chi \mapsto \phi \circ \chi \circ \phi^{-1}$$

ein Homomorphismen von $\mathbb{K}$-Algebren (Übung).

3. Für jeden Körper $\mathbb{K}$ bilden die Polynome mit Koeffizienten in $\mathbb{K}$ mit der punktweisen Addition, Skalarmultiplikation und der Polynommultiplikation die $\mathbb{K}$-Algebra $\mathbb{K}[x]$.

Wir untersuchen nun das Verhalten von $\mathbb{K}$-linearen Abbildungen auf Untervektorräumen. Man fragt sich, ob sie Untervektorräume auf Untervektorräume abbilden und ob das Urbild jedes Untervektorraums ein Untervektorraum ist. Wichtige Spezialfälle sind dabei das Bild und der Kern einer $\mathbb{K}$-linearen Abbildung $\phi : V \to W$, also das Bild des Untervektorraums $V \subseteq V$ und das Urbild des Untervektorraums $\{0\} \subseteq W$.

Satz 4.1.9: Seien V, W $\mathbb{K}$-Vektorräume und sei $\phi : V \to W$ $\mathbb{K}$-linear. Dann gilt:

1. Für jeden Untervektorraum $U \subseteq V$ ist das Bild $\phi(U) \subseteq W$ ein Untervektorraum von W. Insbesondere ist das **Bild** $\mathrm{im}(\phi) = \phi(V)$ ein Untervektorraum von W.

2. Für jeden Untervektorraum $X \subseteq W$ ist das Urbild $\phi^{-1}(X) \subseteq V$ ein Untervektorraum von V. Insbesondere ist der **Kern** $\ker(\phi) = \phi^{-1}(\{0_W\})$ ein Untervektorraum von V.

Beweis:
1. Sei $U \subseteq V$ ein Untervektorraum. Dann gilt nach (UV1) $0_V \in U$ und damit ergibt sich $0_W = \phi(0_V) \in \phi(U)$ (UV1). Sind $w, w' \in \phi(U)$ so existieren $v, v' \in U$ mit $\phi(v) = w$ und $\phi(v') = w'$. Daraus folgt $w + w' = \phi(v) + \phi(v') = \phi(v + v')$. Da nach (UV2) auch $v + v' \in U$ gilt, folgt $w + w' \in \phi(U)$ (UV2). Für $\lambda \in \mathbb{K}$ folgt $\lambda w = \lambda\phi(v) = \phi(\lambda v)$. Da nach (UV3) auch $\lambda v \in U$ gilt, folgt $\lambda w \in \phi(U)$ (UV3). Also ist $\phi(U)$ ein Untervektorraum von W. Für $U = V$ folgt, dass $\mathrm{im}(\phi) = \phi(V) \subseteq W$ ein Untervektorraum ist.

2. Sei $X \subseteq W$ ein Untervektorraum. Dann gilt $0_W = \phi(0_V) \in X$ und damit $0_V \in \phi^{-1}(X)$ (UV1). Sind $v, v' \in \phi^{-1}(X)$ so gilt $\phi(v + v') = \phi(v) + \phi(v') \in X$, also auch $v + v' \in \phi^{-1}(X)$ (UV2). Für $\lambda \in \mathbb{K}$ folgt $\phi(\lambda v) = \lambda\phi(v) \in X$ und damit $\lambda v \in \phi^{-1}(X)$ (UV3). Also ist $\phi^{-1}(X) \subseteq V$ ein Untervektorraum von V. Für $X = \{0_W\}$ folgt, dass $\ker(\phi) = \phi^{-1}(\{0_W\}) \subseteq V$ ein Untervektorraum von V ist. □

Per Definition ist eine $\mathbb{K}$-lineare Abbildung $\phi : V \to W$ surjektiv genau dann, wenn $\mathrm{im}(\phi) = W$. Ebenso können wir zeigen, dass sie injektiv ist genau dann, wenn $\ker(\phi) = \{0\}$ gilt. Dies folgt bereits aus Beispiel 2.2.22, wo gezeigt wurde, dass ein Gruppenhomomorphismus injektiv ist genau dann, wenn sein Kern nur aus dem neutralen Element besteht. Da $\mathbb{K}$-lineare Abbildungen Gruppenhomomorphismen bezüglich der Vektoraddition sind, erhalten wir

Korollar 4.1.10: Eine $\mathbb{K}$-lineare Abbildung $\phi : V \to W$ ist injektiv genau dann, wenn $\ker(\phi) = \{0_V\}$ und surjektiv genau dann, wenn $\mathrm{im}(\phi) = W$ gilt.

Beispiel 4.1.11:

1. Für die Streckung $D_\mu : V \to V$, $v \mapsto \mu v$ aus Beispiel 4.1.3,1. gilt $\ker(D_\mu) = \{0\}$ und $\mathrm{im}(D_\mu) = V$ falls $\mu \in \mathbb{K} \setminus \{0\}$, denn dann ist D_μ invertierbar mit Inversem $D_{\mu^{-1}} : V \to V$. Für $\mu = 0$ erhält man $\ker(D_\mu) = V$ und $\mathrm{im}(D_\mu) = \{0\}$.

2. Für die komplexe Konjugation $\bar{} : \mathbb{C} \to \mathbb{C}$, $z \mapsto \bar{z}$ aus Beispiel 4.1.3,2. gilt $\ker(\bar{}) = \{0\}$ und $\mathrm{im}(\bar{}) = \mathbb{C}$, denn sie ist ihr eigenes Inverses und damit bijektiv.

3. Für die Drehung $R_\beta : \mathbb{R}^2 \mapsto \mathbb{R}^2$, $xe_1 + ye_2 \mapsto (x\cos\beta - y\sin\beta)e_1 + (y\cos\beta + x\sin\beta)e_2$ aus Beispiel 4.1.3,3. gilt $\ker(R_\beta) = \{0\}$ und $\mathrm{im}(R_\beta) = \mathbb{R}^2$, denn sie ist invertierbar mit Inversem $R_{-\beta} : \mathbb{R}^2 \to \mathbb{R}^2$. Aus dem Additionstheorem für Sinus und Kosinus folgt nämlich $R_\alpha \circ R_\beta = R_{\alpha+\beta}$ für beliebige $\alpha, \beta \in \mathbb{R}$ und damit $R_\beta \circ R_{-\beta} = R_{-\beta} \circ R_\beta = R_0 = \mathrm{id}_{\mathbb{R}^2}$.

4. Für die Koordinatenprojektionen $\pi_i : \mathbb{K}^n \to \mathbb{K}^n$, $\Sigma_{k=1}^n \lambda_k e_k \mapsto \lambda_i e_i$ aus Beispiel 4.1.3,4. gilt $\ker(\pi_i) = \mathrm{span}_\mathbb{K}(\{e_1, ..., e_n\} \setminus \{e_i\})$ und $\mathrm{im}(\pi_i) = \mathbb{K}e_i$.

5. Für die formale Ableitung $d : \mathbb{K}[x] \mapsto \mathbb{K}[x]$, $\Sigma_{k=0}^\infty a_k x^k \mapsto \Sigma_{k=0}^\infty k a_k x^{k-1}$ von Polynomen aus Beispiel 4.1.3,5. gilt $\ker(d) = \{\Sigma_{k=0}^\infty a_k x^k \in \mathbb{K}[x] | \, a_k = 0 \, \forall k \in \mathbb{N}\}$ und $\mathrm{im}(d) = \mathbb{K}[x]$.

   Denn für $p = \Sigma_{k=0}^\infty a_k x^k \in \mathbb{K}[x]$ gilt $d(p) = 0$ genau dann, wenn $a_k = 0$ für alle $k \geq 1$. Jedes Polynom $q = \Sigma_{k=0}^\infty b_k x^k$ mit $b_k = a_{k-1}/k$ für $k \geq 1$ und $b_0 \in \mathbb{K}$ beliebig erfülllt $d(q) = p$.

Allgemeiner können wir die Injektivität und Surjektivität von linearen Abbildungen durch ihr Verhalten auf Erzeugendensystemen und auf linear unabhängigen Teilmengen charakterisieren. Dabei ergibt sich, dass injektive Abbildungen die lineare Unabhängigkeit von Teilmengen erhalten und surjektive Abbildungen die Eigenschaft, ein Erzeugendensystem zu sein.

Satz 4.1.12: Seien V, W $\mathbb{K}$-Vektorräume und $\phi : V \to W$ eine lineare Abbildung. Dann gilt:

(i) ϕ injektiv $\Leftrightarrow$ $\phi(A)$ ist linear unabhängig für alle linear unabhängigen Teilmengen $A \subseteq V$.

(ii) ϕ surjektiv $\Leftrightarrow$ $\phi(A)$ erzeugt W für alle Erzeugendensysteme $A \subseteq V$ von V.

(iii) ϕ bijektiv $\Leftrightarrow$ $\phi(A)$ ist eine Basis von W für alle Basen $A \subseteq V$ von V.

Beweis:
(i) $\Leftarrow$ Ist $\phi(A)$ linear unabhängig für alle linear unabhängigen Teilmengen $A \subseteq V$, so ist insbesondere $\{\phi(v)\} = \phi(\{v\})$ linear unabhängig für alle $v \in V \setminus \{0\}$, denn $\{v\} \subseteq V$ ist linear unabhängig für $v \neq 0$. Daraus folgt $\phi(v) \neq 0$ für alle $v \in V \setminus \{0\}$, also $\ker(\phi) = \{0\}$, und nach Korollar 4.1.10 ist ϕ dann injektiv.

$\Rightarrow$ Sei nun ϕ injektiv und $A \subseteq V$ linear unabhängig. Um zu zeigen, dass auch $\phi(A)$ linear unabhängig ist, betrachten wir paarweise verschiedene $w_1, ..., w_n \in \phi(A)$ und $\lambda_1, .., \lambda_n \in \mathbb{K}$ mit $\Sigma_{j=1}^n \lambda_j w_j = 0$. Dann gibt es paarweise verschiedene $v_1, .., v_n \in A$ mit $w_i = \phi(v_i)$, und es folgt $0 = \Sigma_{j=1}^n \lambda_j \phi(v_j) = \phi(\Sigma_{j=1}^n \lambda_j v_j)$, also $\Sigma_{j=1}^n \lambda_j v_j \in \ker(\phi)$. Da ϕ injektiv ist, ist $\ker(\phi) = \{0\}$ nach Korollar 4.1.10 und damit $\Sigma_{j=1}^n \lambda_j v_j = 0$. Da A linear unabhängig ist, folgt $\lambda_1 = ... = \lambda_n = 0$. Also ist auch $\phi(A)$ linear unabhängig.

(ii) Nach Bemerkung 4.1.2 gilt $\phi(\mathrm{span}_{\mathbb{K}}(M)) = \mathrm{span}_{\mathbb{K}}(\phi(M))$ für alle Teilmengen $M \subseteq V$.

$\Leftarrow$ Ist $\phi(A)$ ein Erzeugendensystem von W für jedes Erzeugendensystem A von V, so gilt das auch für $A = V$, und es folgt $W = \mathrm{span}_{\mathbb{K}}(\phi(V)) = \phi(\mathrm{span}_{\mathbb{K}}(V)) = \phi(V)$, also ϕ surjektiv.

$\Rightarrow$ Ist ϕ surjektiv und $\mathrm{span}_{\mathbb{K}}(A) = V$, so folgt $\mathrm{span}_{\mathbb{K}}(\phi(A)) = \phi(\mathrm{span}_{\mathbb{K}}(A)) = \phi(V) = W$ und somit ist für jedes Erzeugendensystem A von V das Bild $\phi(A)$ ein Erzeugendensystem von W.

(iii) folgt aus (i) und (ii). □

Satz 4.1.15 charakterisiert die Injektivität und Surjektivität von linearen Abbildungen durch ihr Verhalten auf linear unabhängigen Teilmengen, Erzeugendensystemen und Basen. Bis jetzt haben wir aber in der Betrachtung von linearen Abbildungen den Dimensionsbegriff noch nicht berücksichtigt. Wir untersuchen daher allgemeiner den Zusammenhang zwischen der Dimension des Kerns, der Dimension des Bildes einer linearen Abbildung $\phi : V \to W$ und den Dimensionen der Vektorräume V und W.

Satz 4.1.13: (**Dimensionsformel für lineare Abbildungen**)
Seien V, W $\mathbb{K}$-Vektorräume und $\phi : V \to W$ eine $\mathbb{K}$-lineare Abbildung. Setzt man $\infty + n = \infty + \infty = \infty$ für alle $n \in \mathbb{N}_0$, so gilt

$$\dim_{\mathbb{K}}(V) = \dim_{\mathbb{K}}(\ker(\phi)) + \dim_{\mathbb{K}}(\mathrm{im}(\phi)).$$

Die Dimension $\dim_{\mathbb{K}}(\mathrm{im}(\phi))$ bezeichnet man als den **Rang** von ϕ und mit $\mathrm{rg}(\phi)$. Die Dimension des Kerns bezeichnet man auch als den **Defekt** von ϕ und mit $\mathrm{def}(\phi)$.

Beweis:
Ist $\mathrm{def}(\phi)$ oder $\mathrm{rg}(\phi)$ unendlich, so gilt dies auch für $\dim_{\mathbb{K}}(V)$, und es ist nichts mehr zu zeigen.

Seien nun also $d := \mathrm{def}(\phi) \in \mathbb{N}_0$ und $n := \dim_{\mathbb{K}}(V) \in \mathbb{N}_0$. Sei $D = \{v_1, ..., v_d\}$ eine Basis von $\ker(\phi)$. Mit Korollar 3.2.14 ergänzen wir sie zu einer Basis $B = \{v_1, ..., v_d, v_{d+1}, ..., v_n\}$ von V. Wir zeigen, dass $R = \{\phi(v_{d+1}), ..., \phi(v_n)\}$ eine Basis des Bildes $\mathrm{im}(\phi)$ ist.

• R erzeugt $\mathrm{im}(\phi)$:
Zu jedem Vektor $w \in \mathrm{im}(\phi)$ gibt es einen Vektor $v \in V$ mit $\phi(v) = w$. Da B eine Basis von V ist, gibt es $\lambda_1, ..., \lambda_n$ mit $v = \Sigma_{i=1}^n \lambda_i v_i$. Daraus folgt

$$w = \phi(v) = \phi(\Sigma_{i=1}^n \lambda_i v_i) = \Sigma_{i=1}^n \lambda_i \phi(v_i) = \Sigma_{i=d+1}^n \lambda_i \phi(v_i) \in \mathrm{span}_{\mathbb{K}}\{\phi(v_{d+1}), ..., \phi(v_n)\},$$

da $v_1, ..., v_d \in \ker(\phi)$. Damit ist R ein Erzeugendensystem von $\mathrm{im}(\phi)$.

• R ist linear unabhängig:
Seien $\lambda_{d+1}, ..., \lambda_n \in \mathbb{K}$ mit $\Sigma_{i=d+1}^n \lambda_i \phi(v_i) = 0$. Dann gilt $\phi(\Sigma_{i=d+1}^n \lambda_i v_i) = \Sigma_{i=d+1}^n \lambda_i \phi(v_i) = 0$ und damit $\Sigma_{i=d+1}^n \lambda_i v_i \in \ker(\phi)$. Da D eine Basis von $\ker(\phi)$ ist, gibt es $\lambda_1, ..., \lambda_d$ mit

$$\Sigma_{i=d+1}^n \lambda_i v_i = \Sigma_{i=1}^d \lambda_i v_i \quad \Rightarrow \quad \Sigma_{i=1}^d \lambda_i v_i + \Sigma_{i=d+1}^n (-\lambda_i) v_i = 0.$$

Wegen der linearen Unabhängigkeit von B folgt daraus $\lambda_1 = \ldots = \lambda_n = 0$. Also ist R linear unabhängig, und die Vektoren $\phi(v_{d+1}), ..., \phi(v_n)$ sind paarweise verschieden.

Damit ist R eine Basis von $\mathrm{im}(\phi)$. Es gilt $\dim_{\mathbb{K}}(\mathrm{im}(\phi)) = n - d$, $\dim_{\mathbb{K}}(\ker(\phi)) = d$ und $\dim_{\mathbb{K}}(V) = n = d + n - d = \dim_{\mathbb{K}}(\ker(\phi)) + \dim_{\mathbb{K}}(\mathrm{im}(\phi))$. □

Mit Satz 4.1.13 können wir eine bekannte und hilfreiche Aussage für endliche Mengen auf Vektorräume verallgemeinern. Ist $\phi : M \to N$ eine Abbildung zwischen *endlichen* Mengen M, N mit gleich vielen Elementen, so ist nach Bemerkung 2.1.24 die Abbildung ϕ injektiv genau dann, wenn sie surjektiv ist. Vektorräume über unendlichen Körpern wie $\mathbb{Q}$, $\mathbb{R}$ oder $\mathbb{C}$ haben zwar im Allgemeinen unendlich viele Elemente, aber eine analoge Aussage gilt für lineare Abbildungen $\phi : V \to W$, wenn V und W endlich-dimensional mit gleicher Dimension sind.

Korollar 4.1.14: Seien V, W Vektorräume über $\mathbb{K}$ mit $\dim_{\mathbb{K}}(V) = \dim_{\mathbb{K}}(W) < \infty$ und $\phi : V \to W$ eine $\mathbb{K}$-lineare Abbildung. Dann gilt ϕ injektiv $\Leftrightarrow$ ϕ surjektiv $\Leftrightarrow$ ϕ bijektiv.

Beweis:
ϕ injektiv $\overset{4.1.10}{\Leftrightarrow}$ $\ker(\phi) = \{0_V\}$ $\overset{4.1.13}{\Leftrightarrow}$ $\dim_{\mathbb{K}}(V) = \dim_{\mathbb{K}}(W) = \dim_{\mathbb{K}}(\mathrm{im}(\phi))$ $\overset{3.2.22}{\Leftrightarrow}$ $\mathrm{im}(\phi) = W$ $\overset{4.1.10}{\Leftrightarrow}$ ϕ surjektiv. □

Die Aussage von Korollar 4.1.14 gilt nicht für unendlich-dimensionale Vektorräume. Denn nach Beispiel 4.1.11 ist die formale Ableitung $d : \mathbb{K}[x] \to \mathbb{K}[x]$ von Polynomen mit Koeffizienten in $\mathbb{K}$ surjektiv, aber nicht injektiv.

Indem wir Korollar 4.1.10 mit der Dimensionsformel aus Satz 4.1.13 kombinieren, können wir im endlich-dimensionalen Fall auch noch allgemeinere Aussagen treffen, die aus der Existenz injektiver, surjektiver und bijektiver Abbildungen zwischen zwei Vektorräumen Einschränkungen an deren Dimensionen ableiten. Wir erhalten die folgende Verallgemeinerung der dritten Aussage in Bemerkung 2.1.24, deren Beweis eine Übungsaufgabe ist.

Korollar 4.1.15: Seien V, W endlich-dimensionale $\mathbb{K}$-Vektorräume und $\phi : V \to W$ eine $\mathbb{K}$-lineare Abbildung. Dann gilt:

(i) Ist ϕ injektiv, so folgt $\dim_{\mathbb{K}}(V) \leq \dim_{\mathbb{K}}(W)$.

(ii) Ist ϕ surjektiv, so folgt $\dim_{\mathbb{K}}(V) \geq \dim_{\mathbb{K}}(W)$.

(iii) Ist ϕ bijektiv, so folgt $\dim_{\mathbb{K}}(V) = \dim_{\mathbb{K}}(W)$.

4.2 Beschreibung durch Basen und Matrizen

Wir beschreiben nun $\mathbb{K}$-linearen Abbildungen $\phi : V \to W$, indem wir ihre Werte auf einer Basis des $\mathbb{K}$-Vektorraums V betrachten und diese als Linearkombinationen einer Basis des $\mathbb{K}$-Vektorraums W ausdrücken. Das zentrale Ergebnis dabei ist, dass eine $\mathbb{K}$-lineare Abbildung $\phi : V \to W$ durch ihre Werte auf einer Basis von V bereits eindeutig festgelegt ist, und dass man eine lineare Abbildung $\phi : V \to W$ konstruieren kann, indem man ihre Werte auf einer Basis von V beliebig vorgibt. Dies ergibt sich aus dem folgenden Satz.

Satz 4.2.1: Seien V, W Vektorräume über einem Körper $\mathbb{K}$, $M \subseteq V$ eine Teilmenge und $f : M \to W$ eine Abbildung. Dann gilt:

(i) Ist M ein Erzeugendensystem von V, so gibt es höchstens eine $\mathbb{K}$-lineare Abbildung $\phi : V \to W$ mit $\phi|_M = f$.

(ii) Ist M eine Basis von V, so gibt es genau eine $\mathbb{K}$-lineare Abbildung $\phi : V \to W$ mit $\phi|_M = f$.

(iii) Ist M linear unabhängig, so gibt es mindestens eine $\mathbb{K}$-lineare Abbildung $\phi : V \to W$ mit $\phi|_M = f$.

Beweis:
(i) Seien $\phi, \psi : V \to W$ lineare Abbildungen mit $\phi|_M = \psi|_M$, also $\phi(m) = \psi(m)$ für alle $m \in M$. Ist M ein Erzeugendensystem von V, so lässt sich jeder Vektor $v \in V$ als eine Linearkombination $v = \Sigma_{j=1}^n \lambda_j m_j$ mit $m_1, ..., m_n \in M$ und $\lambda_1, ..., \lambda_n \in \mathbb{K}$ schreiben. Es folgt

$$\phi(v) = \phi(\Sigma_{j=1}^n \lambda_j m_j) = \Sigma_{j=1}^n \lambda_j \phi(m_j) = \Sigma_{j=1}^n \lambda_j f(m_j) = \Sigma_{j=1}^n \lambda_j \psi(m_j) = \psi(\Sigma_{j=1}^n \lambda_j m_j) = \psi(v).$$

Also kann es höchstens eine lineare Abbildung $\phi : V \to W$ mit $\phi|_M = f$ geben.

(ii) Ist $M \subseteq V$ eine Basis von V, so ist M auch ein Erzeugendensystem von V und nach (i) existiert maximal eine lineare Abbildung $\phi : V \to W$ mit $\phi|_M = f$. Zu zeigen ist noch, dass mindestens eine lineare Abbildung $\phi : V \to W$ mit $\phi|_M = f$ existiert.

Dazu schreiben wir jeden Vektor $v \in V$ eindeutig als Linearkombination $v = \Sigma_{j=1}^n \lambda_j m_j$ mit $\lambda_1, ..., \lambda_n \in \mathbb{K}$ und $m_1, ..., m_n \in M$ und definieren eine lineare Abbildung $\phi : V \to W$ durch $\phi(\Sigma_{j=1}^n \lambda_j m_j) := \Sigma_{j=1}^n \lambda_j f(m_j)$ für alle $n \in \mathbb{N}$, $m_1, ..., m_n \in M$ und $\lambda_1, ..., \lambda_n \in \mathbb{K}$. Die Eindeutigkeit der Darstellung eines Vektors $v \in V$ als Linearkombination der Basisvektoren garantiert, dass $\phi : V \to W$ eine Abbildung ist, also jedem $v \in V$ genau einen Vektor in W zuordnet, und per Definition von ϕ gilt $\phi|_M = f$. Die Abbildung ϕ ist auch $\mathbb{K}$-linear, denn für $v = \Sigma_{j=1}^n \lambda_j m_j$ und $w = \Sigma_{j=1}^n \mu_j m_j$ mit $m_1, ..., m_n \in M$ und $\lambda, \lambda_1, ..., \lambda_n, \mu_1, ..., \mu_n \in \mathbb{K}$ erhält man

$$\begin{aligned}
\phi(v + w) &= \phi(\Sigma_{j=1}^n \lambda_j m_j + \Sigma_{j=1}^n \mu_j m_j) = \phi(\Sigma_{j=1}^n (\lambda_j + \mu_j) m_j) = \Sigma_{j=1}^n (\lambda_j + \mu_j) f(m_j) \\
&= \Sigma_{j=1}^n \lambda_j f(m_j) + \Sigma_{j=1}^n \mu_j f(m_j) = \phi(v) + \phi(w) \\
\phi(\lambda v) &= \phi(\lambda \Sigma_{j=1}^n \lambda_j m_j) = \phi(\Sigma_{j=1}^n (\lambda \lambda_j) m_j) = \Sigma_{j=1}^n \lambda \lambda_j f(m_j) = \lambda \Sigma_{j=1}^n \lambda_j f(m_j) = \lambda \phi(v).
\end{aligned}$$

(iii) Ist $M \subseteq V$ eine linear unabhängige Teilmenge, so existiert nach dem Basisergänzungssatz eine Basis C von V mit $M \subseteq C$. Wir definieren eine Abbildung $g : C \to W$, indem wir $g(m) = f(m)$ für alle $m \in M$ und $g(c) = 0$ für $c \in C \setminus M$ setzen. Nach (ii) existiert eine lineare Abbildung $\phi : V \to W$ mit $\phi|_C = g$, und wegen $M \subseteq C$ folgt daraus $\phi|_M = g|_M = f$. □

Korollar 4.2.2: Seien V, W endlich-dimensionale $\mathbb{K}$-Vektorräume, $v_1, ..., v_n \in V$ und $w_1, ..., w_n \in W$. Dann gilt:

(i) Ist $\{v_1, ..., v_n\}$ ein Erzeugendensystem von V, so gibt es höchstens eine $\mathbb{K}$-lineare Abbildung $\phi : V \to W$ mit $\phi(v_i) = w_i$ für alle $i \in \{1, ..., n\}$.

(ii) Ist $\{v_1, ..., v_n\}$ eine Basis von V, so gibt es genau eine $\mathbb{K}$-lineare Abbildung $\phi : V \to W$ mit $\phi(v_i) = w_i$ für alle $i \in \{1, ..., n\}$.

(iii) Ist $(v_1 ..., v_n)$ linear unabhängig, so gibt es mindestens eine $\mathbb{K}$-lineare Abbildung $\phi : V \to W$ mit $\phi(v_i) = w_i$ für alle $i \in \{1, ..., n\}$.

Beweis:
Wir definieren eine Abbildung $f : \{v_1, ..., v_n\} \to W$ durch $f(v_i) = w_i$ für alle $i \in \{1, ..., n\}$. Dann folgt die Aussage aus Satz 4.2.1. □

In Korollar 4.2.2 kommt es auf die Reihenfolge der Vektoren an, da nicht nur gefordert wird, dass die Menge $\{v_1, ..., v_n\}$ durch die lineare Abbildung $\phi : V \to W$ auf die Menge $\{w_1, ..., w_n\}$ abgebildet wird, sondern, dass jeweils der Vektor v_i auf w_i abgebildet wird. Insbesondere benötigt man für die zweite Aussage nicht einfach nur eine Basis von V, sondern eine Basis, deren Elemente nummeriert, also geordnet, sind. Deshalb bietet es sich an, statt mit einer Menge mit einem Tupel von Vektoren zu arbeiten.

Definition 4.2.3: Sei V ein n-dimensionaler $\mathbb{K}$-Vektorraum. Eine **geordnete Basis** von V ist ein n-Tupel $(v_1, ..., v_n)$ von Vektoren aus V, so dass $\{v_1, ..., v_n\}$ eine Basis von V ist.

Korollar 4.2.2 besagt, dass wir eine lineare Abbildung $\phi : V \to W$ zwischen endlich-dimensionalen Vektorräumen eindeutig definieren können, indem wir eine geordnete Basis von V wählen, und jedem der Basisvektoren einen beliebigen Vektor in W zuordnen. Lineare Abbildungen sind also durch ihre Werte auf einer geordneten Basis eindeutig bestimmt, und es gibt keine Einschränkungen an die Werte, die sie auf einer Basis annehmen können.

Mit Korollar 4.2.2 kann man auch auf einfache Weise Isomorphismen zwischen endlich-dimensionalen Vektorräumen der gleichen Dimension konstruieren. Dazu wählt man zwei geordnete Basen und bildet die eine auf die andere ab. So können wir insbesondere für jeden n-dimensionalen $\mathbb{K}$-Vektorraum V und jede geordnete Basis von V einen $\mathbb{K}$-Vektorraumisomorphismus in den $\mathbb{K}^n$ konstruieren.

Korollar 4.2.4: Sei V ein n-dimensionaler $\mathbb{K}$-Vektorraum, $(e_1, ..., e_n)$ die Standardbasis des $\mathbb{K}^n$ und $B = (v_1, ..., v_n)$ eine geordnete Basis von V.

1. Dann gibt es genau eine $\mathbb{K}$-lineare Abbildung ${}_BS : V \to \mathbb{K}^n$ mit ${}_BS(v_i) = e_i$ für $i = 1, ..., n$.
2. Die Abbildung ${}_BS$ ist bijektiv und ordnet einem Vektor $v = \Sigma_{i=1}^n \lambda_i v_i \in V$ den Spaltenvektor ${}_BS(v) = (\lambda_1, ..., \lambda_n)^T \in \mathbb{K}^n$ zu.

Die Einträge des Vektors ${}_BS(w)$ heißen **Koordinaten** von w bezüglich B.

Beweis:
Aus Korollar 4.2.2 folgen die Existenz und Eindeutigkeit von ${}_BS : V \to \mathbb{K}^n$. Die Abbildung ist offensichtlich surjektiv und wegen $\dim_{\mathbb{K}}(V) = n = \dim_{\mathbb{K}}(\mathbb{K}^n)$ nach Korollar 4.1.14 dann auch bijektiv. □

Korollar 4.2.5: Zwei endlich-dimensionale $\mathbb{K}$-Vektorräume sind genau dann isomorph, wenn sie die gleiche Dimension haben.

Beweis:
Sind V, W endlich-dimensionale $\mathbb{K}$-Vektorräume und $\phi : V \to W$ ein $\mathbb{K}$-Vektorraumisomorphismus, so ergibt sich aus der Dimensionsformel

$$\dim_{\mathbb{K}}(V) = \dim_{\mathbb{K}}(\mathrm{im}(\phi)) + \dim_{\mathbb{K}}(\ker(\phi)) = \dim_{\mathbb{K}}(\mathrm{im}(\phi)) = \dim_{\mathbb{K}}(W).$$

Ist umgekehrt $\dim_{\mathbb{K}}(V) = \dim_{\mathbb{K}}(W) = n$, so kann man geordnete Basen A von V und B von W wählen und erhält mit Korollar 4.2.4 $\mathbb{K}$-Vektorraumisomorphismen ${}_AS : V \to \mathbb{K}^n$ und ${}_BS : W \to \mathbb{K}^n$. Daraus ergibt sich mit Satz 4.1.5, dass die $\mathbb{K}$-lineare Abbildung ${}_BS^{-1} \circ {}_AS : V \to W$ ein $\mathbb{K}$-Vektorraumisomorphismus ist. □

Korollar 4.2.5 besagt, umgangssprachlich ausgedrückt, dass endlich-dimensionale Vektorräume im Gegensatz zu endlichen Gruppen oder Ringen keine Persönlichkeit haben. Während sich Gruppen oder Ringe mit der gleichen Anzahl von Elementen deutlich unterscheiden können - beispielsweise gibt es abelsche und nicht abelsche Gruppen mit der gleichen Anzahl von Elementen - sind alle n-dimensionalen $\mathbb{K}$-Vektorräume isomorph. Das bedeutet, dass sie im Wesentlichen durch bloße Umbenennung ihrer Elemente auseinander hervorgehen.

Damit sind auch alle endlich-dimensionalen Vektorräume vollständig klassifiziert. Denn man verfügt über eine Liste von endlich-dimensionalen $\mathbb{K}$-Vektorräumen, so dass jeder endlich-dimensionale $\mathbb{K}$-Vektorraum zu genau einem Vektorraum in dieser Liste isomorph ist. Diese Liste enthält genau die Vektorräume $\mathbb{K}^n$ mit $n \in \mathbb{N}$ und den Nullvektorraum.

Eine weitere Vereinfachung in der Beschreibung von $\mathbb{K}$-linearen Abbildungen $\phi : V \to W$ erhalten wir, wenn wir auch die Werte dieser Abbildung in W durch eine Basis des Vektorraums W beschreiben. Wir betrachten also endlich-dimensionale $\mathbb{K}$-Vektorräume V und W mit geordneten Basen $A = (v_1, ..., v_n)$ und $B = (w_1, ..., w_m)$. Für eine $\mathbb{K}$-lineare Abbildung $\phi : V \to W$ können wir dann die Bilder jedes Basisvektors v_i eindeutig als Linearkombination der Basisvektoren w_j schreiben. Es gibt also eindeutige Skalare $\phi_{ji} \in \mathbb{K}$ mit $\phi(v_i) = \Sigma_{j=1}^m \phi_{ji} w_j$ für alle $i \in \{1, ..., n\}$. Diese bestimmen die $\mathbb{K}$-lineare Abbildung $\phi : V \to W$ eindeutig. Umgekehrt gibt es zu jeder solchen Sammlung von Skalaren nach Korollar 4.2.2 genau eine $\mathbb{K}$-lineare Abbildung $\phi : V \to W$ mit $\phi(v_i) = \Sigma_{j=1}^m \phi_{ji} w_j$ für alle $i \in \{1, ..., n\}$.

Also können wir statt mit $\mathbb{K}$-linearen Abbildungen auch mit Tabellen arbeiten, die die Koeffizienten ϕ_{ji} enthalten, die *Matrizen*. Möchten wir nicht nur $\mathbb{K}$-lineare Abbildungen, sondern auch Bilder von allgemeinen Vektoren in V unter solchen Abbildungen betrachten, bietet es sich an, diese wie in Korollar 4.2.4 durch Spaltenvektoren im $\mathbb{K}^n$ zu beschreiben.

Definition 4.2.6: Sei $\mathbb{K}$ ein Körper.

1. Eine $(m \times n)$-**Matrix** A mit Einträgen in $\mathbb{K}$ ist eine Tabelle mit m Zeilen und n Spalten

$$A = (a_{ji}) = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix}$$

mit $a_{ji} \in \mathbb{K}$ für alle $j \in \{1, ..., m\}$ und $i \in \{1, ..., n\}$.

Die Menge der $(m \times n)$-Matrizen mit Einträgen in $\mathbb{K}$ wird mit $\mathrm{Mat}(m \times n, \mathbb{K})$ bezeichnet. Den Eintrag in der $j$ten Zeile und $i$ten Spalte von A bezeichnen wir mit a_{ji} oder mit A_{ji}.

2. Sind V, W $\mathbb{K}$-Vektorräume mit geordneten Basen $B = (v_1, ..., v_n)$ und $C = (w_1, ..., w_m)$ und $\phi : V \to W$ eine lineare Abbildung, so nennt man die eindeutig bestimmte Matrix

$$_C M_B(\phi) = (\phi_{ji}) = \begin{pmatrix} \phi_{11} & \phi_{12} & \dots & \phi_{1n} \\ \phi_{21} & \phi_{22} & \dots & \phi_{2n} \\ \vdots & \vdots & & \vdots \\ \phi_{m1} & \phi_{m2} & \dots & \phi_{mn} \end{pmatrix} \in \mathrm{Mat}(m \times n, \mathbb{K}) \quad \text{mit} \quad \phi(v_i) = \Sigma_{j=1}^m \phi_{ji} w_j$$

die **beschreibende** oder **darstellende Matrix** von ϕ bezüglich der Basen B und C.

3. Ist V ein $\mathbb{K}$-Vektorraum mit geordneter Basis $B = (v_1, ..., v_n)$ und $v \in V$, so nennt man den eindeutig bestimmten Spaltenvektor

$$_B S(v) = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} \in \mathbb{K}^n \quad \text{mit} \quad v = \Sigma_{j=1}^n \lambda_j v_j$$

den **beschreibenden** oder **darstellenden Spaltenvektor** von v bezüglich B.

Bemerkung 4.2.7:

1. Per Definition der beschreibenden Matrix und des beschreibenden Spaltenvektors enthält die $i$te Spalte der Matrix $_C M_B(\phi)$ gerade den beschreibenden Spaltenvektor $_C S(\phi(v_i))$ von $\phi(v_i)$ bezüglich der geordneten Basis C.

2. Der beschreibende Spaltenvektor $_B S(v)$ eines Vektors $v \in V$ bezüglich einer geordneten Basis $B = (v_1, ..., v_n)$ von V ist gleichzeitig die beschreibende Matrix der linearen Abbildung $L_v : \mathbb{K} \to V$, $\lambda \mapsto \lambda v$ bezüglich der geordneten Basen (1) von $\mathbb{K}$ und B von V. Denn aus $v = L_v(1) = \Sigma_{j=1}^n \lambda_j v_j$ folgt $_B S(v) = {_B M_{(1)}}(L_v) = (\lambda_1, ..., \lambda_n)^T$.

3. Ebenso kann man den Zeilenvektor $(\lambda_1, ..., \lambda_n) \in \mathrm{Mat}(1 \times n, \mathbb{K})$ als die beschreibende Matrix der eindeutigen linearen Abbildung $L^* : V \to \mathbb{K}$ mit $L^*(v_i) = \lambda_i$ bezüglich der geordneten Basen B von V und (1) von $\mathbb{K}$ auffassen. Denn per Definition der darstellenden Matrix gilt $_{(1)} M_B(L^*) = (\lambda_1, ..., \lambda_n)$.

Beispiel 4.2.8: Wir betrachten die Drehung um den Winkel $\beta \in \mathbb{R}$ aus Beispiel 4.1.3, 3.

$$R_\beta : \mathbb{R}^2 \mapsto \mathbb{R}^2, \qquad \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x\cos\beta - y\sin\beta \\ y\cos\beta + x\sin\beta \end{pmatrix}$$

1. Wir berechnen die beschreibende Matrix von R bezüglich der Basis $B = (e_1, e_2)$:

$$R_\beta(e_1) = \begin{pmatrix} 1\cdot\cos\beta - 0\cdot\sin\beta \\ 0\cdot\cos\beta + 1\cdot\sin\beta \end{pmatrix} = \begin{pmatrix} \cos\beta \\ \sin\beta \end{pmatrix} = \underbrace{\cos\beta}_{R_{11}}\, e_1 + \underbrace{\sin\beta}_{R_{21}}\, e_2$$

$$R_\beta(e_2) = \begin{pmatrix} 0\cdot\cos\beta - 1\cdot\sin\beta \\ 1\cdot\cos\beta + 0\cdot\sin\beta \end{pmatrix} = \begin{pmatrix} -\sin\beta \\ \cos\beta \end{pmatrix} = \underbrace{-\sin\beta}_{R_{12}}\, e_1 + \underbrace{\cos\beta}_{R_{22}}\, e_2$$

$$\Rightarrow \quad {_B M_B}(R_\beta) = \begin{pmatrix} \cos\beta & -\sin\beta \\ \sin\beta & \cos\beta \end{pmatrix}.$$

2. Wir berechnen die darstellende Matrix von R bezüglich der geordneten Basen $A = (e_1 + e_2, e_1 - e_2)$ und $C = (e_1 + e_2, e_2)$:

$$R_\beta(e_1 + e_2) = \begin{pmatrix} 1 \cdot \cos\beta - 1 \cdot \sin\beta \\ 1 \cdot \cos\beta + 1 \cdot \sin\beta \end{pmatrix} = \begin{pmatrix} \cos\beta - \sin\beta \\ \cos\beta + \sin\beta \end{pmatrix}$$
$$= (\cos\beta - \sin\beta)e_1 + (\cos\beta + \sin\beta)e_2 = \underbrace{(\cos\beta - \sin\beta)}_{R_{11}}(e_1 + e_2) + \underbrace{2\sin\beta}_{R_{21}}\, e_2$$

$$R_\beta(e_1 - e_2) = \begin{pmatrix} 1 \cdot \cos\beta - (-1) \cdot \sin\beta \\ (-1) \cdot \cos\beta + 1 \cdot \sin\beta \end{pmatrix} = \begin{pmatrix} \cos\beta + \sin\beta \\ \sin\beta - \cos\beta \end{pmatrix}$$
$$= (\cos\beta + \sin\beta)e_1 + (\sin\beta - \cos\beta)e_2 = \underbrace{(\cos\beta + \sin\beta)}_{R_{12}}(e_1 + e_2) + \underbrace{(-2\cos\beta)}_{R_{22}}\, e_2$$

$$\Rightarrow \quad {}_C M_A(R_\beta) = \begin{pmatrix} \cos\beta - \sin\beta & \cos\beta + \sin\beta \\ 2\sin\beta & -2\cos\beta \end{pmatrix}.$$

Beispiel 4.2.9: Wir betrachten den Vektorraum $\mathbb{K}[x]_{\leq n} = \{p = \Sigma_{k=0}^n a_k x^k \mid a_k \in \mathbb{K}\} \subseteq \mathbb{K}[x]$ der Polynome vom Grad $\leq n$ und die formale Ableitung aus Beispiel 4.1.3, 5.

$$d : \mathbb{K}[x]_{\leq n} \to \mathbb{K}[x]_{\leq n}, \quad \Sigma_{k=0}^n a_k x^k \mapsto \Sigma_{k=0}^n k a_k x^{k-1}.$$

Für die geordnete Basis $B = (1, x, ..., x^{n-1}, x^n)$ von $\mathbb{K}[x]_{\leq n}$ der Monome gilt dann $d(1) = 0$ und $d(x^k) = kx^{k-1}$ für alle $k \in \{1, ..., n\}$. Also ist die darstellende Matrix der formalen Ableitung bezüglich der geordneten Basis B von $\mathbb{K}[x]_{\leq n}$ gegeben durch

$$ {}_B M_B(d) = \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 2 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \ldots & 0 & n \\ 0 & 0 & \ldots & 0 & 0 \end{pmatrix}.$$

Wir untersuchen nun, wie sich die punktweise Addition von linearen Abbildungen, ihre punktweise Multiplikation mit Skalaren in $\mathbb{K}$ und ihre Verkettung durch die darstellenden Matrizen beschreiben lassen.

Dazu betrachten wir endlich-dimensionale $\mathbb{K}$-Vektorräume U, V, W mit fest gewählten geordneten Basen $A = (u_1, ..., u_p)$, $B = (v_1, ..., v_n)$ und $C = (w_1, ..., w_m)$ und $\mathbb{K}$-lineare Abbildungen $\phi, \psi : U \to V$ und $\chi : V \to W$. Wir berechnen jeweils die darstellende Matrix, indem wir die Bilder von Basisvektoren als Linearkombinationen von Basisvektoren schreiben und die Einträge der darstellenden Matrix daraus ablesen.

Betrachtung 4.2.10:

1. Sind $\phi, \psi : U \to V$ lineare Abbildungen mit $\phi(u_i) = \Sigma_{j=1}^n \phi_{ji} v_j$ und $\psi(u_i) = \Sigma_{j=1}^n \psi_{ji} v_j$, für alle $i \in \{1, ..., p\}$, so folgt

$$(\phi + \psi)(u_i) = \phi(u_i) + \psi(u_i) = \Sigma_{j=1}^n \phi_{ji} v_j + \Sigma_{j=1}^n \psi_{ji} v_j = \Sigma_{j=1}^n (\phi_{ji} + \psi_{ji}) v_j. \tag{6}$$

   Die Addition von linearen Abbildungen addiert also die Einträge ihrer beschreibenden Matrizen bezüglich der Basen A und B.

2. Ist $\phi : U \to V$ $\mathbb{K}$- linear mit $\phi(u_i) = \Sigma_{j=1}^n \phi_{ji} v_j$ für alle $i \in \{1, ..., p\}$ und $\lambda \in \mathbb{K}$, so folgt

$$(\lambda\phi)(u_i) = \lambda\phi(u_i) = \lambda\Sigma_{j=1}^n \phi_{ji} v_j = \Sigma_{j=1}^n (\lambda\phi_{ji}) v_j \tag{7}$$

   Die Multiplikation einer linearen Abbildung mit $\lambda \in \mathbb{K}$ multipliziert also jeden Eintrag ihrer beschreibenden Matrix bezüglich der Basen A und B mit λ.

3. Ist $\phi : U \to V$ eine lineare Abbildung mit $\phi(u_i) = \Sigma_{j=1}^n \phi_{ji} v_j$ für alle $i \in \{1, ..., p\}$ und $u = \Sigma_{i=1}^p \lambda_i u_i$ ein Vektor in U, so gilt

$$\phi(u) = \phi(\Sigma_{i=1}^p \lambda_i u_i) = \Sigma_{i=1}^p \lambda_i \phi(u_i) = \Sigma_{i=1}^p \Sigma_{j=1}^n \lambda_i \phi_{ji} v_j = \Sigma_{j=1}^n \left(\Sigma_{i=1}^p \phi_{ji} \lambda_i\right) v_j \tag{8}$$

   Der beschreibende Spaltenvektor ${}_BS(\phi(u))$ hat also als $j$ten Eintrag genau die Summe $\Sigma_{i=1}^p \phi_{ji} \lambda_i$. Diese erhält man, indem man jeweils den $i$ten Eintrag in der $j$ten Zeile der beschreibenden Matrix ${}_AM_B(\phi)$ mit dem $i$ten Eintrag des beschreibenden Spaltenvektors ${}_AS(u)$ multipliziert und die so entstehenden Terme addiert.

4. Sind $\phi : U \to V$ und $\chi : V \to W$ $\mathbb{K}$-lineare Abbildungen mit $\phi(u_i) = \Sigma_{j=1}^n \phi_{ji} v_j$ und $\chi(v_j) = \Sigma_{k=1}^m \chi_{kj} w_k$ für alle $i \in \{1, ..., p\}$ und $j \in \{1, .., n\}$, so folgt

$$\begin{aligned}(\chi \circ \phi)(u_i) &= \chi(\phi(u_i)) = \chi(\Sigma_{j=1}^n \phi_{ji} v_j) = \Sigma_{j=1}^n \phi_{ji} \chi(v_j) = \Sigma_{j=1}^n \phi_{ji} \left(\Sigma_{k=1}^m \chi_{kj} w_j\right) \\ &= \Sigma_{j=1}^n \Sigma_{k=1}^m \phi_{ji} \chi_{kj} w_k = \Sigma_{k=1}^m \left(\Sigma_{j=1}^n \chi_{kj} \phi_{ji}\right) w_k.\end{aligned} \tag{9}$$

   Den Eintrag in der $k$ten Zeile und $i$ten Spalte der beschreibenden Matrix ${}_CM_A(\chi \circ \psi)$ erhält man also indem man jeweils den $j$ten Eintrag in der $k$ten Zeile der beschreibenden Matrix ${}_CM_B(\chi)$ mit dem $j$ten Eintrag in der $i$ten Spalte der beschreibenden Matrix ${}_BM_A(\phi)$ multipliziert und die so entstehenden Terme addiert.

5. Die darstellende Matrix ${}_AM_A(\mathrm{id}_U)$ der Identitätsabbildung id_U bezüglich der geordneten Basis A ist gegeben durch $\phi(u_i) = u_i = \Sigma_{j=1}^n \delta_{ji} u_j$ mit $\delta_{ji} = 0$ für $i \neq j$ und $\delta_{ii} = 1$ und die darstellende Matrix der Nullabbildung durch $\phi(u_i) = 0$.

Wir definieren nun die Addition, Skalarmultiplikation und Multiplikation von Matrizen gerade so, dass dass diese der Addition, Skalarmultiplikation und Verkettung von linearen Abbildungen entsprechen. Zusätzlich führen wir ein Produkt von Matrizen mit Spaltenvektoren ein, das die Anwendung einer linearen Abbildung auf einen Vektor beschreibt, sowie die $n$ten Einheitsmatrizen, die die Identitätsabbildung bezüglich einer geordneten Basis beschreiben.

Definition 4.2.11: Sei $\mathbb{K}$ ein Körper.

1. Die **Summe** einer $(m \times n)$-Matrix A und einer $(m \times n)$-Matrix B ist die $(m \times n)$-Matrix

$$A + B = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} + \begin{pmatrix} b_{11} & \dots & b_{1n} \\ \vdots & & \vdots \\ b_{m1} & \dots & b_{mn} \end{pmatrix} := \begin{pmatrix} a_{11} + b_{11} & \dots & a_{1n} + b_{1n} \\ \vdots & & \vdots \\ a_{m1} + b_{m1} & \dots & a_{mn} + b_{mn} \end{pmatrix}.$$

2. Das **skalare Vielfache** einer $(m \times n)$-Matrix A mit $\lambda \in \mathbb{K}$ ist die $(m \times n)$-Matrix

$$\lambda A = \lambda \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} := \begin{pmatrix} \lambda a_{11} & \dots & \lambda a_{1n} \\ \vdots & & \vdots \\ \lambda a_{m1} & \dots & \lambda a_{mn} \end{pmatrix}.$$

3. Das **Produkt** einer $(m \times n)$-Matrix A mit einem Spaltenvektor $v \in \mathbb{K}^n$ ist

$$A \cdot v = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} := \begin{pmatrix} \Sigma_{j=1}^n a_{1j}\lambda_j \\ \vdots \\ \Sigma_{j=1}^n a_{mj}\lambda_j \end{pmatrix} \in \mathbb{K}^m$$

4. Das **Produkt** einer $(m \times n)$-Matrix A mit einer $(n \times p)$-Matrix B ist die $(m \times p)$-Matrix

$$A \cdot B = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} b_{11} & \dots & b_{1p} \\ \vdots & & \vdots \\ b_{n1} & \dots & b_{np} \end{pmatrix} := \begin{pmatrix} \Sigma_{j=1}^n a_{1j}b_{j1} & \dots & \Sigma_{j=1}^n a_{1j}b_{jp} \\ \vdots & & \vdots \\ \Sigma_{j=1}^n a_{mj}b_{j1} & \dots & \Sigma_{j=1}^n a_{mj}b_{jp} \end{pmatrix}.$$

5. Die $n$te **Einheitsmatrix** ist die Matrix $\mathbb{1}_n \in \mathrm{Mat}(n \times n, \mathbb{K})$ mit

$$\mathbb{1}_n = (\delta_{ji}) = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 1 \end{pmatrix}.$$

6. Die **Nullmatrix** ist die Matrix $0_{m\times n} \in \mathrm{Mat}(n \times n, \mathbb{K})$, für die jeder Eintrag Null ist.

Anhand der Formeln in Definition 4.2.11 sieht man, dass das Produkt einer Matrix mit einem Spaltenvektor im $\mathbb{K}^n$ das gleiche ist, wie das Produkt der Matrix mit der entsprechenden $(n\times 1)$-Matrix. Wir brauchen also nicht zwischen Spaltenvektoren im $\mathbb{K}^n$ und $(n \times 1)$-Matrizen zu unterscheiden. Dies ist eine Konsequenz von Bemerkung 4.2.7, 1.

Da Addition, Skalarmultiplikation und Multiplikation von Matrizen gerade so definiert wurden, dass sie mit der Addition, Skalarmultiplikation und Verkettung von linearen Abbildungen verträglich sind, wenn wir mit fest gewählten Basen arbeiten, erhalten wir

Satz 4.2.12: (**Satz über darstellende Matrizen**) Seien U, V, W $\mathbb{K}$-Vektorräume mit geordneten Basen $A = (u_1, ..., u_p)$, $B = (v_1, ..., v_n)$, $C = (w_1, ..., w_m)$.

Dann ist die Abbildung ${}_CM_B : \mathrm{Hom}_{\mathbb{K}}(V, W) \to \mathrm{Mat}(m \times n, \mathbb{K})$, $\chi \mapsto {}_CM_B(\chi)$ bijektiv.

Für die darstellenden Matrizen und darstellenden Spaltenvektoren gilt:

(i) ${}_BM_A(\phi + \psi) = {}_BM_A(\phi) + {}_BM_A(\psi)$ für alle $\mathbb{K}$-linearen Abbildungen $\phi, \psi : U \to V$,

(ii) ${}_BM_A(\lambda\phi) = \lambda\, {}_BM_A(\phi)$ für alle $\lambda \in \mathbb{K}$ und $\mathbb{K}$-linearen Abbildungen $\phi : U \to V$,

(iii) ${}_BS(\phi(u)) = {}_BM_A(\phi) \cdot {}_AS(u)$ für alle $u \in U$ und $\mathbb{K}$-linearen Abbildungen $\phi : U \to V$,

(iv) ${}_CM_A(\chi \circ \phi) = {}_CM_B(\chi) \cdot {}_BM_A(\phi)$ für alle $\mathbb{K}$-linearen Abbildungen $\phi : U \to V$, $\chi : V \to W$.

(v) ${}_AM_A(\mathrm{id}_U) = \mathbb{1}_p$ und ${}_BM_A(0) = 0_{n\times p}$.

Beweis:
Nach Korollar 4.2.2 existiert zu jeder Matrix $N = (n_{ji}) \in \mathrm{Mat}(m \times n, \mathbb{K})$ genau eine lineare Abbildung ${}^C\phi^B(N) : V \to W$ mit ${}^C\phi^B(N)(v_i) = \Sigma_{j=1}^m n_{ji}w_j$ für $i = 1, ..., n$. Diese erfüllt per Definition die Bedingung ${}_CM_B({}^C\phi^B(N)) = N$. Also hat ${}_CM_B$ die Umkehrabbildung

$${}_CM_B{}^{-1} : \mathrm{Mat}(m \times n, \mathbb{K}) \to \mathrm{Hom}_{\mathbb{K}}(V, W), \quad N \mapsto {}^C\phi^B(N) \text{ mit } {}^C\phi^N(N)(v_i) = \Sigma_{j=1}^n n_{ji}w_j$$

und ist somit bijektiv. Aus Gleichung (8) und Definition 4.2.6 folgt (i), aus (6) und Definition 4.2.6 folgt (ii), aus (7) folgt (iii) und aus (9) folgt (iv). Betrachtung 4.2.10, 5 impliziert (v). □

Wir untersuchen nun den Zusammenhang zwischen linearen Abbildungen und ihren darstellenden Matrizen sowie die Multiplikation von Matrizen aus Definition 4.2.11 an einem Beispiel.

Beispiel 4.2.13: Wir betrachten die Drehungen im $\mathbb{R}^2$ um den Winkel $\alpha \in \mathbb{R}$ aus Beispiel 4.2.8 mit darstellender Matrix bezüglich der Standardbasis $B = (e_1, e_2)$

$$_BM_B(R_\alpha) = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}.$$

Dann gilt für die Verkettung zweier Drehungen um die Winkel $\alpha, \beta \in \mathbb{R}$

$$_BM_B(R_\alpha \circ R_\beta) = {_BM_B}(R_\alpha) \cdot {_BM_B}(R_\beta) = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \cdot \begin{pmatrix} \cos\beta & -\sin\beta \\ \sin\beta & \cos\beta \end{pmatrix}$$
$$= \begin{pmatrix} \cos\alpha\cos\beta - \sin\alpha\sin\beta & -\cos\beta\sin\alpha - \cos\alpha\sin\beta \\ \sin\alpha\cos\beta + \sin\beta\cos\alpha & -\sin\alpha\sin\beta + \cos\alpha\cos\beta \end{pmatrix} = \begin{pmatrix} \cos(\alpha+\beta) & -\sin(\alpha+\beta) \\ \sin(\alpha+\beta) & \cos(\alpha+\beta) \end{pmatrix},$$

wobei im letzten Schritt die Additionsformeln für Sinus und Kosinus benutzt wurden. Die Verkettung zweier Drehungen um die Winkel $\alpha, \beta \in \mathbb{R}$ entspricht also gerade der Drehung um die Summe der beiden Winkel.

Für endlich-dimensionale $\mathbb{K}$-Vektorräume V und W mit geordneten Basen $B = (v_1, ..., v_n)$ und $C = (w_1, ..., w_n)$ können wir nach dem Beweis von Satz 4.2.12 jeder $(m \times n)$-Matrix N mit Einträgen in $\mathbb{K}$ eine eindeutig bestimmte lineare Abbildung ${^C\phi^B}(N) : V \to W$ zuordnen, deren darstellende Matrix bezüglich der Basen B und C gerade die Matrix N ist.

Insbesondere können wir dabei die $\mathbb{K}$-Vektorräume $V = \mathbb{K}^n$ und $W = \mathbb{K}^m$ mit den Standardbasen $B = (e_1, ..., e_n)$ und $C = (e_1, ..., e_m)$ betrachten. Dies liefert die lineare Abbildung $\phi_N : V \to W$ mit $\phi_N(e_i) = \Sigma_{j=1}^m n_{ji} e_j$. Sie erfüllt $\phi_N(\Sigma_{i=1}^n \lambda_i e_i) = \Sigma_{j=1}^m (\Sigma_{i=1}^n n_{ji}\lambda_i) e_j$ für alle $\lambda_1, ..., \lambda_n \in \mathbb{K}$, was gerade der Multiplikation von Spaltenvektoren im $\mathbb{K}^n$ mit der Matrix N entspricht. Man erhält also die lineare Abbildung $\phi_N : \mathbb{K}^n \to \mathbb{K}^m$, $v \mapsto N \cdot v$. Indem wir für jede Matrix $N \in \text{Mat}(m \times n, \mathbb{K})$ diese zugehörige lineare Abbildung betrachten, können wir Begriffe für lineare Abbildungen wie Bild, Kern und Rang auch für Matrizen definieren.

Definition 4.2.14: Die **Standardabbildung** für eine Matrix $N \in \text{Mat}(m \times n, \mathbb{K})$ ist die $\mathbb{K}$-lineare Abbildung $\phi_N : \mathbb{K}^n \to \mathbb{K}^m$, $v \mapsto N \cdot v$, die eindeutige $\mathbb{K}$-lineare Abbildung mit N als darstellende Matrix bezüglich der Standardbasen. Wir definieren

- den **Kern** von N als $\ker(N) := \ker(\phi_N)$,
- das **Bild** von N als $\text{im}(N) := \text{im}(\phi_N)$,
- den **Rang** von N als $\text{rg}(N) := \text{rg}(\phi_N) = \dim_{\mathbb{K}}(\text{im}(\phi_N))$.
- den **Defekt** von N als $\text{def}(N) := \text{def}(\phi_N) = \dim_{\mathbb{K}}(\ker(\phi_N))$.

Bemerkung 4.2.15:
Für eine Matrix $N = (v_1, ..., v_n) \in \text{Mat}(m \times n, \mathbb{K})$ mit Spaltenvektoren $v_1, ..., v_n \in \mathbb{K}^m$ gilt

$$\begin{aligned} \text{im}(N) &= \text{span}_{\mathbb{K}}(\{Ne_1, ..., Ne_n\}) = \text{span}_{\mathbb{K}}(\{v_1, ..., v_n\}) \\ \ker(N) &= \{x \in \mathbb{K}^n | N \cdot x = 0\} \\ \text{rg}(N) &= \dim_{\mathbb{K}}(\text{im}(N)) = \dim_{\mathbb{K}}(\text{span}_{\mathbb{K}}(\{v_1, ..., v_n\})). \end{aligned}$$

Da ihr Rang damit gleich der maximalen Anzahl linear unabhängiger Spaltenvektoren ist, spricht man oft auch vom **Spaltenrang** und schreibt $\mathrm{srg}(N) := \mathrm{rg}(N)$.

Mit Hilfe von Satz 4.2.12 können wir nun auch alle anderen Aussagen über lineare Abbildungen zwischen endlich-dimensionalen Vektorräumen in Matrizen übersetzen. Dass die Abbildung ${}_BM_A : \mathrm{Abb}(U, V) \to \mathrm{Mat}(n \times p, \mathbb{K})$, $\phi \mapsto {}_BM_A(\phi)$ bijektiv ist, bedeutet nämlich insbesondere, dass zwei lineare Abbildungen $\phi, \phi' : U \to V$ gleich sind genau dann, wenn ihre darstellenden Matrizen ${}_BM_A(\phi)$ und ${}_BM_A(\phi')$ bezüglich zweier fest gewählter Basen A, B gleich sind.

Außerdem können wir mit dieser bijektiven Abbildung aus den Eigenschaften der Verkettung, der Addition und der Skalarmultiplikation linearer Abbildungen entsprechende Rechenregeln für Matrizen ableiten. Diese kann man auch durch direktes Nachrechnen mit Hilfe der Formeln aus Definition 4.2.11 beweisen, aber der Beweis über die linearen Abbildungen ist aufschlussreicher.

Lemma 4.2.16: (**Rechenregeln für Matrizen**)
Seien $m, n, p, r \in \mathbb{N}$ und $\mathbb{K}$ ein Körper. Dann gilt für alle Skalare $\lambda, \mu \in \mathbb{K}$ und Matrizen $M, M', M'' \in \mathrm{Mat}(m \times n, \mathbb{K})$, $N, N' \in \mathrm{Mat}(n \times p, \mathbb{K})$ und $P \in \mathrm{Mat}(p \times r, \mathbb{K})$:

1. **Assoziativität:** $M \cdot (N \cdot P) = (M \cdot N) \cdot P$ und $M + (M' + M'') = (M + M') + M''$,
2. **Kommutativität:** $M + M' = M' + M$,
3. **Distributivgesetze:**

   (i) $M \cdot (N + N') = M \cdot N + M \cdot N'$, (ii) $(M + M') \cdot N = M \cdot N + M' \cdot N$
   (iii) $\lambda(N + N') = \lambda N + \lambda N'$, (iv) $(\lambda + \mu)N = \lambda N + \mu N$.

4. **Multiplikation mit Skalaren:** $\lambda(M \cdot N) = (\lambda M) \cdot N = M \cdot (\lambda N)$,
5. **Neutrale Elemente:** $\mathbb{1}_m \cdot M = M \cdot \mathbb{1}_n = M$ und $0_{m\times n} + M = M + 0_{m\times n} = M$.

Beweis:
Übung. □

Betrachten wir endlich-dimensionale $\mathbb{K}$-Vektorräume U, V mit geordneten Basen $A = (u_1, .., u_p)$ und $B = (v_1, ..., v_n)$, so können wir die Aussagen aus Satz 4.2.12 und Lemma 4.2.16 auch als Aussagen über die bijektive Abbildung ${}_BM_A : \mathrm{Hom}_{\mathbb{K}}(U, V) \to \mathrm{Mat}(n \times p, \mathbb{K})$ auffassen, die einer linearen Abbildung $\phi : U \to V$ ihre beschreibende Matrix ${}_BM_A(\phi)$ zuordnet. Damit erhalten wir die folgenden zwei Sätze.

Satz 4.2.17: Seien $m, n \in \mathbb{N}$ und $\mathbb{K}$ ein Körper.

1. Die Menge $\mathrm{Mat}(m \times n, \mathbb{K})$ mit der Skalarmultiplikation und Addition aus Definition 4.2.11, 1. und 2. ist ein $\mathbb{K}$-Vektorraum der Dimension mn.
2. Die **Elementarmatrizen** E_{ji}, die in der $j$ten Zeile und $i$ten Spalte den Eintrag 1 und ansonsten nur 0 enthalten, bilden eine Basis von $\mathrm{Mat}(m \times n, \mathbb{K})$.
3. Sind V, W $\mathbb{K}$-Vektorräume mit geordneten Basen $B = (v_1, ..., v_n)$ und $C = (w_1, ..., w_m)$, so ist die Abbildung ${}_CM_B : \mathrm{Hom}_{\mathbb{K}}(V, W) \to \mathrm{Mat}(m \times n, \mathbb{K})$, $\phi \mapsto {}_CM_B(\phi)$ ein $\mathbb{K}$-Vektorraumisomorphismus.
4. Die linearen Abbildungen $\phi^{(ji)} : V \to W$ mit $\phi^{(ji)}(v_i) = w_j$ und $\phi^{(ji)}(v_k) = 0$ für $k \neq i$ bilden eine Basis von $\mathrm{Hom}_{\mathbb{K}}(V, W)$. Insbesondere gilt $\dim_{\mathbb{K}}(\mathrm{Hom}_{\mathbb{K}}(V, W)) = mn$.

Beweis:
1. Dass $\mathrm{Mat}(m \times n, \mathbb{K})$ mit der Skalarmultiplikation und Addition aus Definition 4.2.11 einen Vektorraum über $\mathbb{K}$ ist, folgt aus Lemma 4.2.16. Konkret ergibt sich aus der zweiten Bedingung in Lemma 4.2.16, 1., aus 2. und aus der zweiten Bedingung in 5., dass $(\mathrm{Mat}(m \times n, \mathbb{K}), +)$ eine abelsche Gruppe mit neutralem Element $0_{m\times n}$ und Inversem $-M = (-1)\cdot M$ ist. Bedingungen 3. (iii) und (iv) sind das vektorielle und skalare Distributivgesetz (V2) und (V3). Die Pseudoassoziativität (V1) und die Normierungsbedingung $1\cdot M = M$ (V4) folgt direkt aus der Definition der skalaren Vielfachen in Definition 4.2.11.

2. Dass die Matrizen E_{ji} eine Basis von $\mathrm{Mat}(m \times n, \mathbb{K})$ bilden, ergibt sich aus der Definition der Matrixaddition und der skalaren Vielfachen in Definition 4.2.11. Da es genau mn solche Matrizen gibt, folgt $\dim_{\mathbb{K}}(\mathrm{Mat}(m \times n, \mathbb{K})) = mn$.

3. und 4. Dass ${}_CM_B : \mathrm{Hom}_{\mathbb{K}}(V, W) \to \mathrm{Mat}(m \times n, \mathbb{K})$ linear ist, ergibt sich aus Satz 4.2.12 (ii) und (iii). Da diese Abbildung nach Satz 4.2.12 außerdem bijektiv ist, ist sie ein $\mathbb{K}$-Vektorraumisomorphismus. Per Definition der darstellenden Matrix gilt ${}_CM_B(\phi^{(ji)}) = E_{ji}$. Mit Korollar 4.1.15 folgt dann, dass die Abbildungen $\phi^{(ji)}$ eine Basis von $\mathrm{Hom}_{\mathbb{K}}(V, W)$ bilden. □

Satz 4.2.18: Sei $n \in \mathbb{N}$ und $\mathbb{K}$ ein Körper.

1. Die Menge $\mathrm{Mat}(n \times n, \mathbb{K})$ bildet mit der Matrixmultiplikation, der Matrixaddition und der Skalarmultiplikation aus Definition 4.2.11 eine $\mathbb{K}$-Algebra mit Einheitselement $\mathbb{1}_n$.
2. Ist V ein $\mathbb{K}$-Vektorraum mit einer fest gewählten Basis $B = (v_1, ..., v_n)$, so ist die Abbildung ${}_BM_B : \mathrm{End}_{\mathbb{K}}(V) \to \mathrm{Mat}(n \times n, \mathbb{K})$, $\phi \mapsto {}_BM_B(\phi)$ ein Isomorphismus von $\mathbb{K}$-Algebren.

Beweis:
1. Nach Satz 4.2.17 ist $\mathrm{Mat}(n \times n, \mathbb{K})$ mit der Matrixaddition und Skalarmultiplikation ein Vektorraum über $\mathbb{K}$. Dass $\mathrm{Mat}(n \times n, \mathbb{K})$ mit der Matrixaddition und Matrixmultiplikation aus Definition 4.2.11 einen einen unitalen Ring mit Einheitselement $\mathbb{1}_n$ bildet, ergibt sich aus Lemma 4.2.16. Konkret besagt die zweite Bedingung in Lemma 4.2.16, 1. zusammen mit 2. und der zweiten Bedingung in 5. , dass $(\mathrm{Mat}(n \times n, \mathbb{K}), +)$ eine abelsche Gruppe mit neutralem Element $0_{m\times n}$ ist (R1). Bedingungen (i) und (ii) in Lemma 4.2.16, 3. sind die Distributivgesetze (R3). Die erste Bedingung in Lemma 4.2.16, 1. und die erste Bedingung in 5. besagen, dass $(\mathrm{Mat}(n \times n, \mathbb{K}), \cdot)$ ein Monoid mit neutralem Element $\mathbb{1}_n$ ist (UR). Die Kompatibilität von Skalarmultiplikation und Matrixmultiplikation entspricht Aussage 4. in Lemma 4.2.16.

2. Nach Satz 4.2.12 ist die Abbildung ${}_BM_B : \mathrm{End}_{\mathbb{K}}(V) \to \mathrm{Mat}(n \times n, \mathbb{K})$ bijektiv, und nach Satz 4.2.17 ist sie $\mathbb{K}$-linear. Aus Satz 4.2.12 (ii), (iv) und (v) folgt, dass sie ein unitaler Ringisomorphismus ist, und damit ein Isomorphismus von $\mathbb{K}$-Algebren. □

Bemerkung 4.2.19:

1. Die Multiplikation im Ring $\mathrm{Mat}(n \times n, \mathbb{K})$ ist für $n \geq 2$ *nicht kommutativ*. Beispielsweise gilt in $\mathrm{Mat}(2 \times 2, \mathbb{K})$
$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \qquad \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$
Die Ringe $\mathrm{Mat}(n \times n, \mathbb{K})$ für $n \geq 2$ und die dazu isomorphen Ringe $\mathrm{End}_{\mathbb{K}}(V)$ für n-dimensionale $\mathbb{K}$-Vektorräume V sind wichtige Beispiele nicht-kommutativer Ringe.

2. Die Einheitengruppe des Monoids $(\mathrm{Mat}(n \times n, \mathbb{K}), \cdot)$ ist[11]

$$\begin{aligned}\mathrm{GL}(n, \mathbb{K}) : &= \mathrm{Mat}(n \times n, \mathbb{K})^\times \\ &= \{A \in \mathrm{Mat}(n \times n, \mathbb{K}) \mid \exists A^{-1} \in \mathrm{Mat}(n \times n, \mathbb{K}) \text{ mit } A^{-1} \cdot A = A \cdot A^{-1} = \mathbb{1}_n\}.\end{aligned}$$

   Ist $A \in \mathrm{GL}(n, \mathbb{K})$, so nennt man A eine **invertierbare Matrix** und A^{-1} die zu A **inverse Matrix**. Es gelten die üblichem Rechenregeln und Identitäten für Einheiten in Monoiden (vgl. Lemma 2.2.13 und Korollar 2.2.14).

3. Für jeden n-dimensionalen $\mathbb{K}$-Vektorraum V ist nach Satz 4.2.18 die beschreibende Matrix ${}_BM_B(\phi)$ eines Endomorphismus $\phi : V \to V$ invertierbar genau dann, wenn der Endomorphismus ϕ invertierbar ist. Die Einheitengruppe $\mathrm{GL}(n, \mathbb{K})$ des Monoids $(\mathrm{Mat}(n \times n, \mathbb{K}), \cdot)$ ist also das Bild der Einheitengruppe $(\mathrm{Aut}_\mathbb{K}(V), \circ)$ des Monoids $(\mathrm{End}_\mathbb{K}(V), \circ)$ unter dem Algebraisomorphismus ${}_BM_B : \mathrm{End}_\mathbb{K}(V) \to \mathrm{Mat}(n \times n, \mathbb{K})$.

Mit Satz 4.2.18 können wir statt der Endomorphismenalgebra $\mathrm{End}_\mathbb{K}(V)$ eines n-dimensionalen $\mathbb{K}$-Vektorraums V die dazu isomorphe Algebra $\mathrm{Mat}(n\times n, \mathbb{K})$ betrachten. Man beachte aber, dass es nicht einen ausgezeichneten Algebraisomorphismus, sondern sehr viele verschiedene Isomorphismen zwischen den Algebren $\mathrm{Mat}(n \times n, \mathbb{K})$ und $\mathrm{End}_\mathbb{K}(V)$ gibt. Jede Wahl einer geordneten Basis B von V liefert einen Algebraisomorphismus ${}_BM_B : \mathrm{End}_\mathbb{K}(V) \to \mathrm{Mat}(n \times n, \mathbb{K})$.

4.3 Basiswechsel und Koordinatentransformationen

Die Beschreibung von endlich dimensionalen $\mathbb{K}$-Vektorräumen durch Spaltenvektoren im $\mathbb{K}^n$ und von $\mathbb{K}$-linearen Abbildungen durch Matrizen vereinfacht offensichtlich viele Probleme. Anderseits hat sie den Nachteil, dass die so erhaltenen darstellenden Spaltenvektoren und Matrizen von der Wahl der Basis abhängen.

Um die Ergebnisse für verschiedene Basen zu vergleichen, müssen wir untersuchen, wie sich ein Wechsel der Basis auf die beschreibenden Spaltenvektoren und Matrizen auswirkt. Dies ergibt sich als Spezialfall vom Satz 4.2.12. Denn aus Satz 4.2.12 (iv) folgt

$${}_{B'}M_{A'}(\phi) = {}_{B'}M_{A'}(\mathrm{id}_W \circ \phi \circ \mathrm{id}_V) = {}_{B'}M_B(\mathrm{id}_W) \cdot {}_BM_A(\phi) \cdot {}_AM_{A'}(\mathrm{id}_V).$$

für alle $\mathbb{K}$-lineare Abbildungen $\phi : V \to W$ mit $\dim_\mathbb{K} V, \dim_\mathbb{K} W < \infty$ und alle geordnete Basen A,A' von V und B, B' von W. Für $V = W$ und $\phi = \mathrm{id}_V$ folgt aus Satz 4.2.12 (iv) und (v)

$$\mathbb{1}_n = {}_{A'}M_{A'}(\mathrm{id}_V) = {}_{A'}M_{A'}(\mathrm{id}_V \circ \mathrm{id}_V) = {}_{A'}M_A(\mathrm{id}_V) \cdot {}_AM_{A'}(\mathrm{id}_V) \;\Rightarrow\; {}_AM_{A'}(\mathrm{id}_V) = {}_{A'}M_A(\mathrm{id}_V)^{-1}$$

Damit haben wir den folgenden Satz bewiesen.

Satz 4.3.1: (**Basistransformationsformel**)
Seien V, W endlich-dimensionale $\mathbb{K}$-Vektorräume mit geordneten Basen A, A' und B, B'.

1. Die **Basiswechselmatrizen** ${}_{A'}M_A(\mathrm{id}_V)$ und ${}_{B'}M_B(\mathrm{id}_W)$ sind invertierbar mit Inversen

$${}_{A'}M_A(\mathrm{id}_V)^{-1} = {}_AM_{A'}(\mathrm{id}_V) \qquad {}_{B'}M_B(\mathrm{id}_W)^{-1} = {}_BM_{B'}(\mathrm{id}_W).$$

[11]Die Abkürzung GL steht für deren englische Bezeichnung *General Linear Group*.

2. Für jede $\mathbb{K}$-lineare Abbildung $\phi : V \to W$ ist

$$_{B'}M_{A'}(\phi) = {}_{B'}M_B(\mathrm{id}_W) \cdot {}_BM_A(\phi) \cdot {}_AM_{A'}(\mathrm{id}_V).$$

3. Im Spezialfall $V = W$ erhält man für $A = B$ und $A' = B'$:

$$_{A'}M_{A'}(\phi) = {}_{A'}M_A(\mathrm{id}_V) \cdot {}_AM_A(\phi) \cdot {}_AM_{A'}(\mathrm{id}_V) = {}_{A'}M_A(\mathrm{id}_V) \cdot {}_AM_A(\phi) \cdot {}_{A'}M_A(\mathrm{id}_V)^{-1}.$$

Zur Übersicht fassen wir die Aussagen von Satz 4.2.12 (i), (iv) und Satz 4.3.1 in einem **kommutierenden Diagramm** von $\mathbb{K}$-Vektorräumen und linearen Abbildungen zusammen.

Dabei werden die linearen Abbildungen zwischen den beteiligten Vektorräumen V, W, $\mathbb{K}^n$ und $\mathbb{K}^m$ durch Pfeile beschrieben, die mit den Namen der entsprechenden Abbildung gekennzeichnet werden. Ein solcher Pfeil zeigt immer vom Definitionsbereich zum Wertebereich der Abbildung, und das Aneinanderhängen von Pfeilen entspricht ihrer Verkettung. Invertierbare Abbildungen werden oft durch das Symbol $\cong$ an ihrem Pfeil hervorgehoben.

Das *Kommutieren* des Diagramms bedeutet, dass alle durch Aneinanderhängen von Pfeilen entstehenden Wege im Diagramm von einem gegebenen Startpunkt zu einem gegebenen Zielpunkt die gleiche Abbildung liefern.

Die Aussagen von Satz 4.2.12 (i), (iv) und Satz 4.3.1 liefern das kommutierenden Diagramm

$$\begin{tikzcd}
\mathbb{K}^n \arrow[rrr, "x \mapsto {}_BM_A(\phi)\cdot x"] \arrow[ddd, "w \mapsto {}_{A'}M_A(\mathrm{id}_V)\cdot w"'] & & & \mathbb{K}^m \arrow[ddd, "y \mapsto {}_{B'}M_B(\mathrm{id}_W)\cdot y"] \\
& V \arrow[ul, "{}_AS", "\cong"'] \arrow[r, "\phi"] \arrow[ddl, "{}_{A'}S"', "\cong"] & W \arrow[ur, "{}_BS"', "\cong"] \arrow[ddr, "{}_{B'}S", "\cong"'] & \\
& & & \\
\mathbb{K}^n \arrow[rrr, "z \mapsto {}_{B'}M_{A'}(\phi)\cdot z"'] & & & \mathbb{K}^m
\end{tikzcd} \tag{10}$$

- Das Kommutieren des oberes Trapezes entspricht der Aussage

$$_BS(\phi(v)) = {}_BM_A(\phi) \cdot {}_AS(v) \qquad \forall v \in V,$$

- Das Kommutieren des unteren Trapezes entspricht der Aussage

$$_{B'}S(\phi(v)) = {}_{B'}M_{A'}(\phi) \cdot {}_{A'}S(v) \qquad \forall v \in V,$$

- Das Kommutieren des linken Dreiecks entspricht der Aussage

$$_{A'}S(v) = {}_{A'}M_A(\mathrm{id}_V) \cdot {}_AS(v) \qquad \forall v \in V,$$

- Das Kommutieren des rechten Dreiecks entspricht der Aussage

$$_{B'}S(w) = {}_{B'}M_B(\mathrm{id}_W) \cdot {}_BS(w) \qquad \forall w \in W,$$

- Das Kommutieren des äußeren Rechtecks entspricht der Aussage

$$_{B'}M_B(\mathrm{id}_W) \cdot {}_BM_A(\phi) = {}_{B'}M_{A'}(\phi) \cdot {}_{A'}M_A(\mathrm{id}_V).$$

Aus der Basistransformationsformel und dem zugehörigen kommutierenden Diagramm (10) wird deutlich, dass sehr viele verschiedene Matrizen dieselbe lineare Abbildung $\phi : V \to W$ beschreiben können. Beispielsweise kann man zeigen (Übung), dass jede invertierbare Matrix $N \in \mathrm{Mat}(n \times n, \mathbb{K})$ die Identitätsabbildung $\mathrm{id}_{\mathbb{K}^n} : \mathbb{K}^n \to \mathbb{K}^n$ bezüglich irgendwelcher Basen A, $A' \subseteq \mathbb{K}^n$ beschreibt: es gibt Basen A, A' von $\mathbb{K}^n$ mit $N = {}_{A'}M_A(\mathrm{id}_{\mathbb{K}^n})$. Dennoch beschreibt nicht jede Matrix jede beliebige lineare Abbildung. Beispielsweise ist die beschreibende Matrix der Nullabbildung $\phi : V \to W$, $v \mapsto 0$ bezüglich beliebiger Basen immer eine Nullmatrix.

Wir benötigen also ein Kriterium, das uns sagt, ob zwei gegebene Matrizen bezüglich irgendwelcher Basen dieselbe Abbildung beschreiben. Dabei müssen wir im Fall der $(n \times n)$-Matrizen entscheiden, ob wir darstellende Matrizen bezüglich zweier verschiedener Basen oder nur bezüglich einer Basis zulassen wollen. Die Kriterien, die sich daraus ergeben sind die Äquivalenz und Ähnlichkeit von Matrizen.

Definition 4.3.2: Sei $\mathbb{K}$ ein Körper.

1. Eine Matrix $M' \in \mathrm{Mat}(m \times n, \mathbb{K})$ heißt **äquivalent** zu einer Matrix $M \in \mathrm{Mat}(m \times n, \mathbb{K})$, wenn invertierbare Matrizen $A \in \mathrm{GL}(m, \mathbb{K})$, $B \in \mathrm{GL}(n, \mathbb{K})$ existieren mit $M' = A \cdot M \cdot B^{-1}$.

2. Eine Matrix $M' \in \mathrm{Mat}(n \times n, \mathbb{K})$ heißt **ähnlich** oder **konjugiert** zu einer Matrix $M \in \mathrm{Mat}(n \times n, \mathbb{K})$, wenn eine invertierbare Matrix $A \in \mathrm{GL}(n, \mathbb{K})$ existiert mit $M' = A \cdot M \cdot A^{-1}$.

Ist M' ähnlich zu M, so ist M' offensichtlich auch äquivalent zu M, denn dann kann man für die Matrix B in Definition 4.3.2, 1. $B = A$ wählen. Umgekehrt folgt aus M' äquivalent zu M aber *nicht*, dass M' ähnlich zu M ist - auch dann nicht wenn M und M' $(n \times n)$-Matrizen sind. Beispielsweise ist jede invertierbare Matrix $C \in \mathrm{GL}(n, \mathbb{K})$ *äquivalent* zur Einheitsmatrix, denn $\mathbb{1}_n = A \cdot C \cdot B^{-1}$ mit $A = \mathbb{1}_n$, $B = C$. Die einzige invertierbare Matrix $C \in \mathrm{GL}(n, \mathbb{K})$, die *ähnlich* zur Einheitsmatrix ist, ist aber die Einheitsmatrix selbst, denn aus $\mathbb{1}_n = A \cdot C \cdot A^{-1}$ folgt $C = A^{-1} \cdot \mathbb{1}_n \cdot A = \mathbb{1}_n$.

Wie die Namen Äquivalenz und Ähnlichkeit schon andeuten, handelt es sich dabei um Äquivalenzrelationen auf den Mengen $\mathrm{Mat}(m \times n, \mathbb{K})$ und $\mathrm{Mat}(n \times n, \mathbb{K})$.

Satz 4.3.3: Sei $\mathbb{K}$ ein Körper und $n, m \in \mathbb{N}$.

1. Äquivalenz von Matrizen ist eine Äquivalenzrelation auf der Menge $\mathrm{Mat}(m \times n, \mathbb{K})$.
2. Zwei Matrizen $N, N' \in \mathrm{Mat}(m \times n, \mathbb{K})$ sind äquivalent genau dann, wenn sie bezüglich *zweier Paare von Basen* die selbe lineare Abbildung beschreiben:
   es gibt einen n-dimensionalen $\mathbb{K}$-Vektorraum V mit geordneten Basen A, A', einen m-dimensionalen $\mathbb{K}$-Vektorraum W mit geordneten Basen B, B' und eine lineare Abbildung $\phi : V \to W$ mit $N = {}_BM_A(\phi)$ und $N' = {}_{B'}M_{A'}(\phi)$.
3. Ähnlichkeit von Matrizen ist eine Äquivalenzrelation auf der Menge $\mathrm{Mat}(n \times n, \mathbb{K})$.
4. Zwei Matrizen $N, N' \in \mathrm{Mat}(n \times n, \mathbb{K})$ sind ähnlich genau dann, wenn sie bezüglich *zweier Basen* den selben Vektorraumendomorphismus beschreiben:
   es gibt einen n-dimensionalen $\mathbb{K}$-Vektorraum V mit geordneten Basen A,A' und ein Endomorphismus $\phi : V \to V$ mit $N = {}_AM_A(\phi)$ und $N' = {}_{A'}M_{A'}(\phi)$.

Beweis:
1. Äquivalenz von Matrizen ist eine Äquivalenzrelation:

• **reflexiv:**
Jede Matrix $M \in \mathrm{Mat}(m \times n, \mathbb{K})$ ist äquivalent zu sich selbst, denn es gilt $M = \mathbb{1}_m \cdot M \cdot \mathbb{1}_n$.

• **symmetrisch:** Ist M' äquivalent zu M, so gibt es Matrizen $A \in \mathrm{GL}(m, \mathbb{K})$, $B \in \mathrm{GL}(n, \mathbb{K})$ mit $M' = A \cdot M \cdot B^{-1}$. Es folgt $A^{-1} \cdot M' \cdot B = A^{-1} \cdot A \cdot M \cdot B^{-1} \cdot B = M$ mit $A^{-1} \in \mathrm{GL}(m, \mathbb{K})$ und $B^{-1} \in \mathrm{GL}(n, \mathbb{K})$, und damit ist auch M äquivalent zu M'.

• **transitiv:** Ist M'' äquivalent zu M' und M' äquivalent zu M, dann existieren Matrizen $A, C \in \mathrm{GL}(m, \mathbb{K})$ und $B, D \in \mathrm{GL}(n, \mathbb{K})$ mit $M'' = A \cdot M' \cdot B^{-1}$ und $M' = C \cdot M \cdot D^{-1}$. Daraus folgt $M'' = A \cdot C \cdot M \cdot D^{-1} \cdot B^{-1} = (A \cdot C) \cdot M \cdot (B \cdot D)^{-1}$ mit $B \cdot D \in \mathrm{GL}(n, \mathbb{K})$ und $A \cdot C \in \mathrm{GL}(m, \mathbb{K})$. Also ist auch M'' äquivalent zu M.

2. $\Leftarrow$: Existieren $\mathbb{K}$-Vektorräume V, W mit geordneten Basen A, A' und B, B' und eine lineare Abbildung $\phi : V \to W$ mit ${}_BM_A(\phi) = N$ und ${}_{B'}M_{A'}(\phi) = N'$, so folgt aus Satz 4.3.1

$$N' = {}_{B'}M_{A'}(\phi) = {}_{B'}M_B(\mathrm{id}_W) \cdot {}_BM_A(\phi) \cdot {}_{A'}M_A(\mathrm{id}_V)^{-1} = {}_{B'}M_B(\mathrm{id}_W) \cdot N \cdot {}_{A'}M_A(\mathrm{id}_V)^{-1}$$

mit ${}_{B'}M_B(\mathrm{id}_W) \in \mathrm{GL}(m, \mathbb{K})$ und ${}_{A'}M_A(\mathrm{id}_V) \in \mathrm{GL}(n, \mathbb{K})$. Also sind N und N' äquivalent.

$\Rightarrow$: Sind umgekehrt $N, N' \in \mathrm{Mat}(m \times n, \mathbb{K})$ äquivalent, so existieren Matrizen $C \in \mathrm{GL}(m, \mathbb{K})$, $D \in \mathrm{GL}(n, \mathbb{K})$ mit $N' = C \cdot N \cdot D^{-1}$. Die beschreibende Matrix der $\mathbb{K}$-linearen Abbildung $\phi_N : \mathbb{K}^n \to \mathbb{K}^m$, $v \mapsto N \cdot v$ bezüglich der Standardbasen $A = (e_1, ..., e_n)$ und $B = (e_1, ..., e_m)$ ist dann nach Definition 4.2.14 und Bemerkung 4.2.15 gegeben durch ${}_BM_A(\phi_N) = N$.

Wir konstruieren Basesn A' des $\mathbb{K}^n$ und B' des $\mathbb{K}^m$ mit ${}_{B'}M_{A'}(\phi_N) = N'$. Wegen der Invertierbarkeit von D und C gilt $\mathrm{rg}(D) = \mathrm{rg}(D^{-1}) = n$ und $\mathrm{rg}(C) = \mathrm{rg}(C^{-1}) = m$. Denn aus der Invertierbarkeit von D folgt, dass $\phi_D : \mathbb{K}^n \to \mathbb{K}^n$ invertierbar ist mit Inversem $\phi_D^{-1} = \phi_{D^{-1}}$, und damit ist $\mathrm{rg}(D) = \mathrm{rg}(\phi_D) = \dim_{\mathbb{K}} \mathrm{im}(\phi_D) = n$ and analog für C. Damit bilden die Spaltenvektoren v_j von D^{-1} eine geordnete Basis $A' = (v_1, ..., v_n)$ des $\mathbb{K}^n$ und die Spaltenvektoren w_i von C^{-1} eine geordnete Basis $B' = (w_1, ..., w_m)$ des $\mathbb{K}^m$.

Die Spaltenvektoren sind gegeben durch $v_j = \Sigma_{k=1}^n D^{-1}_{kj} e_k$ und $w_i = \Sigma_{l=1}^m C^{-1}_{ki} e_k$ für $j \in \{1, ..., n\}$ und $i \in \{1, ..., m\}$. Damit sind die Basiswechselmatrizen für den Wechsel von A' zu A und den Wechsel von B' zu B die Matrizen ${}_AM_{A'}(\mathrm{id}_{\mathbb{K}^n}) = D^{-1}$ und ${}_BM_{B'}(\mathrm{id}_{\mathbb{K}^m}) = C^{-1}$. Mit Satz 4.3.1 ergibt sich dann

$${}_{B'}M_{A'}(\phi_N) = {}_BM_{B'}(\mathrm{id}_{\mathbb{K}^m})^{-1} \cdot {}_BM_A(\phi_N) \cdot {}_AM_{A'}(\mathrm{id}_{\mathbb{K}^n}) = (C^{-1})^{-1} \cdot n \cdot D^{-1} = C \cdot N \cdot D^{-1} = N'.$$

Die entsprechenden Beweise für Ähnlichkeit in 3. und 4. sind analog (Übung). $\square$

Satz 4.3.3 wandelt die Frage, ob zwei gegebene Matrizen N, N' die selbe lineare Abbildung bezüglich verschiedener Basen(paare) beschreiben, in die Begriffe der Äquivalenz und Ähnlichkeit von Matrizen um. Im Fall der Äquivalenz lässt sich schnell klären, welche Matrizen in der gleichen Äquivalenzklasse liegen: Zwei Matrizen $N, N' \in \mathrm{Mat}(m \times n, \mathbb{K})$ sind genau dann äquivalent, wenn sie den gleichen Rang haben.

Satz 4.3.4: (**Äquivalenz von Matrizen**)

1. Jede Matrix $N \in \mathrm{Mat}(m \times n, \mathbb{K})$ ist äquivalent zu genau einer Matrix der Form

$$E^r_{m \times n} = \begin{pmatrix} \mathbb{1}_r & 0_{r \times (n-r)} \\ 0_{(m-r) \times r} & 0_{(m-r) \times (n-r)} \end{pmatrix} \quad \text{mit } r = \mathrm{rg}(N) \in \{0, ..., \min(n, m)\},$$

2. Zwei Matrizen $N, N' \in \mathrm{Mat}(m\times n, \mathbb{K})$ sind genau dann äquivalent, wenn $\mathrm{rg}(N) = \mathrm{rg}(N')$.
3. Für jede lineare Abbildung $\phi : U \to V$, die bezüglich irgendwelcher Basen durch N beschrieben wird, gilt $\mathrm{rg}(\phi) = \mathrm{rg}(N)$.

Beweis:
1.1. Wir zeigen: Für äquivalente Matrizen N, N' gilt $\mathrm{rg}(N) = \mathrm{rg}(N')$:

Sind N und N' äquivalent, so gibt es invertierbare Matrizen $A \in \mathrm{Gl}(m, \mathbb{K})$ und $B \in \mathrm{GL}(n, \mathbb{K})$ mit $N' = A \cdot N \cdot B^{-1}$. Um den Rang von N und N' zu bestimmen, betrachten wir die Standardabbildungen $\phi_N : \mathbb{K}^n \to \mathbb{K}^m$, $v \mapsto N \cdot v$ und $\phi_{N'} : \mathbb{K}^n \to \mathbb{K}^m$, $v \mapsto N' \cdot v$ mit $\mathrm{rg}(N) = \mathrm{rg}(\phi_N)$ und $\mathrm{rg}(N') = \mathrm{rg}(\phi_{N'})$ aus Definition 4.2.14.

Da $N' = A \cdot N \cdot B^{-1}$ gilt $\phi_{N'} = \phi_A \circ \phi_N \circ \phi_{B^{-1}} : \mathbb{K}^n \to \mathbb{K}^m$, $v \mapsto A \cdot N \cdot B^{-1} \cdot v$. Wegen der Invertierbarkeit von A und B sind $\phi_A : \mathbb{K}^m \to \mathbb{K}^m$, $v \mapsto A \cdot v$ und $\phi_{B^{-1}} : \mathbb{K}^n \to \mathbb{K}^m$, $v \mapsto B^{-1} \cdot v$ bijektiv mit Inversen $\phi_{A^{-1}} : \mathbb{K}^m \to \mathbb{K}^m$, $v \mapsto A^{-1} \cdot v$ und $\phi_B : \mathbb{K}^n \to \mathbb{K}^m$, $v \mapsto B \cdot v$.

Nach Korollar 4.1.15 gilt für endlich-dimensionale $\mathbb{K}$-Vektorräume U, V, W, X, invertierbare $\mathbb{K}$-lineare Abbildungen $\xi : U \to V$ und $\psi : W \to X$ und jede $\mathbb{K}$-lineare Abbildung $\chi : V \to W$

$$\mathrm{rg}(\psi \circ \chi \circ \xi) = \dim_{\mathbb{K}}(\psi \circ \chi)(\xi(U)) = \dim_{\mathbb{K}}(\psi \circ \chi)(V) = \dim_{\mathbb{K}} \psi(\chi(V)) = \dim_{\mathbb{K}} \chi(V) = \mathrm{rg}(\chi).$$

Also folgt $\mathrm{rg}(N') = \mathrm{rg}(\phi_{N'}) = \mathrm{rg}(\phi_A \circ \phi_N \circ \phi_{B^{-1}}) = \mathrm{rg}(\phi_N) = \mathrm{rg}(N)$.

1.2. Wir zeigen: Jede Matrix $N \in \mathrm{Mat}(m \times n, \mathbb{K})$ ist äquivalent zu $E^r_{m\times n}$ mit $r = \mathrm{rg}(N)$:

Wir wählen eine geordnete Basis $(v_{r+1}, ..., v_n)$ von $\ker(\phi_N) \subseteq \mathbb{K}^n$ und ergänzen sie zu einer geordneten Basis $A = (v_1, ..., v_n)$ von $\mathbb{K}^n$. Dann ist $(\phi_N(v_1), ..., \phi_N(v_r))$ eine Basis von $\mathrm{im}(\phi_N) \subseteq \mathbb{K}^m$. Wir ergänzen sie zu einer Basis $B = (\phi_N(v_1), ..., \phi_N(v_r), w_{r+1}, ..., w_m)$ von $\mathbb{K}^m$. Die beschreibende Matrix von ϕ_N bezüglich der Basen A und B ist dann ${}_BM_A(\phi_N) = E^r_{m\times n}$, während N nach Definition 4.2.14 die beschreibende Matrix von ϕ_N bezüglich der Standardbasen ist. Nach Satz 4.3.3 sind N und $E^r_{m\times n}$ damit äquivalent.

2. Nach 1. gilt $\mathrm{rg}(N) = \mathrm{rg}(N')$ für äquivalente Matrizen $N, N' \in \mathrm{Mat}(m \times n, \mathbb{K})$. Ist umgekehrt $\mathrm{rg}(N) = \mathrm{rg}(N') = r$, so sind N und N' nach 2. äquivalent zu $E^r_{m\times n}$ und wegen der Transitivität der Äquivalenzrelation *äquivalent* dann auch zueinander äquivalent.

3. Ist $\phi : U \to V$ $\mathbb{K}$-linear mit $\dim_{\mathbb{K}}(U) = n$ und $\dim_{\mathbb{K}}(V) = m$ und sind A, B geordnete Basen von U, V mit ${}_BM_A(\phi) = N$, so gilt ${}_BS \circ \phi = \phi_N \circ {}_AS$ nach (10). Da die linearen Abbildungen ${}_BS : V \to \mathbb{K}^m$ und ${}_AS : U \to \mathbb{K}^n$ aus Korollar 4.2.4 $\mathbb{K}$-Vektorraumisomorphismen sind, folgt $\mathrm{rg}(\phi) = \mathrm{rg}({}_BS \circ \phi) = \mathrm{rg}(\phi_N \circ {}_AS) = \mathrm{rg}(\phi_N) = \mathrm{rg}(N)$. □

Satz 4.3.4 besagt insbesondere, dass die Äquivalenzrelation *äquivalent sein* auf $\mathrm{Mat}(m \times n, \mathbb{K})$ genau $1 + \min(n, m)$ Äquivalenzklassen besitzt, nämlich die Äquivalenzklassen der Matrizen $E^r_{m\times n}$ für alle möglichen Werte von r

$$\mathrm{Mat}(m \times n, \mathbb{K})/\sim_{\text{äqu}} = \{[E^r_{m\times n}] \,|\, r \in \{0, 1, ..., \min(n, m)\}.$$

Damit reduziert Satz 4.3.4 die Untersuchung der *Äquivalenz* zweier Matrizen auf die Frage, ob diese Matrizen den gleichen Rang besitzen. Satz 4.3.4 zeigt auch, dass alleine die beschreibende Matrix ohne Information über die zugrundeliegenden Basen sehr wenig über eine lineare Abbildung aussagt. Kennt man die Basen nicht, so kann man aus der beschreibenden Matrix nur die Dimension des Kerns und Bilds der linearen Abbildung rekonstruieren.

Festzustellen, ob zwei gegebene Matrizen $N, N' \in \mathrm{Mat}(n \times n, \mathbb{K})$ *ähnlich* sind, ist viel schwieriger als ihre Äquivalenz zu untersuchen. Wir werden dieses Problem erst in Kapitel 9 vollständig lösen. Allerdings können wir schon jetzt die linearen Abbildungen identifizieren, deren darstellende Matrix bezüglich *einer* Basis von der Form $E^r_{n\times n}$ ist und damit auch die Matrizen beschreiben, die ähnlich zu einer solchen Matrix sind.

Definition 4.3.5: Sei V ein $\mathbb{K}$-Vektorraum. Eine $\mathbb{K}$-lineare Abbildung $\phi : V \to V$ mit $\phi\circ\phi = \phi$ heißt **Projektionsabbildung** oder **Projektor** oder **Idempotent**.

Beispiel 4.3.6:

1. Im $\mathbb{K}^n$ sind die **Koordinatenprojektionen** $\pi_i : \mathbb{K}^n \to \mathbb{K}^n$, $\Sigma_{j=1}^n x_j e_j \mapsto x_i e_i$ Projektionsabbildungen, denn sie sind linear mit $\pi_i \circ \pi_i = \pi_i$ für $i \in \{1, ..., n\}$.

2. Im $\mathbb{R}^2$ ist für alle $\alpha \in \mathbb{R}$ die lineare Abbildung $\phi_\alpha : \mathbb{R}^2 \to \mathbb{R}^2$, $(x, y)^T \mapsto (x, \alpha x)^T$ eine Projektionsabbildung, denn sie ist linear und $\phi_\alpha \circ \phi_\alpha = \phi_\alpha$. Es handelt sich um die Projektion auf die Gerade $y = \alpha x$, wobei parallel zur y-Richtung projiziert wird.

3. Ist ein $\mathbb{K}$-Vektorraum V eine direkte Summe $V = U_1 \oplus U_2$ zweier Untervektorräume $U_1, U_2 \subseteq V$, so sind $\phi_1, \phi_2 : V \to V$ mit $\phi_1(u_1 + u_2) = u_1$ und $\phi_2(u_1 + u_2) = u_2$ für alle $u_1 \in U_1$ und $u_2 \in U_2$ Projektionsabbildungen.

Das letzte Beispiel zeigt, dass eine Zerlegung eines Vektorraums als direkte Summe zweier Untervektorräume Projektionsabbildungen liefert. Umgekehrt liefert eine Projektionsabbildung $\phi : V \to V$ immer eine Zerlegung von V als direkte Summe zweier Untervektorräume, nämlich ihres Kerns und Bilds. Damit können wir im endlich-dimensionalen Fall eine Basis konstruieren, bezüglich der eine Projektionsabbildung durch eine Matrix $E^r_{n\times n}$ beschrieben wird.

Lemma 4.3.7: Sei V ein $\mathbb{K}$-Vektorraum und $\phi : V \to V$ eine Projektionsabbildung.

1. Dann gilt $V = \ker(\phi) \oplus \mathrm{im}(\phi)$.
2. Ist $\dim_{\mathbb{K}}(V) = n$, gibt es eine geordnete Basis A von V mit ${}_AM_A(\phi) = E^r_{n\times n}$ für $r = \mathrm{rg}(\phi)$.

Beweis:
1. Wir können jeden Vektor $v \in V$ schreiben als $v = v - \phi(v) + \phi(v)$ mit $\phi(v) \in \mathrm{im}(\phi)$ und $v - \phi(v) \in \ker(\phi)$, denn es gilt $\phi(v - \phi(v)) = \phi(v) - \phi(\phi(v)) = \phi(v) - \phi(v) = 0$. Also ist $V = \ker(\phi) + \mathrm{im}(\phi)$. Ist $v \in \ker(\phi) \cap \mathrm{im}(\phi)$, so existiert ein $u \in V$ mit $v = \phi(u)$, und es folgt $0 = \phi(v) = \phi(\phi(u)) = \phi(u) = v$. Also ist $\ker(\phi) \cap \mathrm{im}(\phi) = \{0\}$ und damit $V = \ker(\phi) \oplus \mathrm{im}(\phi)$.

2. Ist $\dim_{\mathbb{K}} V = n$, so können wir eine geordnete Basis $(v_1, ..., v_r)$ mit $r = \mathrm{rg}(\phi)$ von $\mathrm{im}(\phi)$ und eine geordnete Basis $(v_{r+1}, ..., v_n)$ von $\ker(\phi)$ wählen. Dann ist $A = (v_1, ..., v_n)$ eine geordnete Basis von V, und es gilt $\phi(v_i) = v_i$ für $i \in \{1, .., r\}$ und $\phi(v_i) = 0$ für $i \in \{r + 1, ..., n\}$, also ${}_AM_A(\phi) = E^r_{n\times n}$. □

Mit Hilfe dieses Lemmas können wir nun zeigen, dass die $(n \times n)$-Matrizen, die *ähnlich* zu einer Matrix $E^r_{n\times n}$ sind, gerade die darstellenden Matrizen von Projektionsabbildungen sind. Daraus ergibt sich, dass genau die Matrizen $N \in \mathrm{Mat}(n \times n, \mathbb{K})$ ähnlich zu einer Matrix $E^r_{n\times n}$ sind, für die $N \cdot N = N$ gilt. Insbesondere ist damit gezeigt, dass es $(n \times n)$-Matrizen mit Einträgen in $\mathbb{K}$ gibt, die nicht ähnlich zu einer Matrix $E^r_{n\times n}$ sind, denn offensichtlich erfüllt nicht jede $(n \times n)$-Matrix diese Bedingung.

Satz 4.3.8: Für eine Matrix $N \in \text{Mat}(n \times n, \mathbb{K})$ sind äquivalent:

(i) N ist ähnlich zu einer Matrix $E^r_{n\times n}$.

(ii) $N \cdot N = N$.

(iii) N ist beschreibende Matrix einer Projektionsabbildung:
es gibt einen n-dimensionalen $\mathbb{K}$-Vektorraum V, eine geordnete Basis A von V und eine Projektionsabbildung $\phi : V \to V$ mit $N = {}_A M_A(\phi)$.

Beweis:
(i)⇒(ii): Ist $N \in \text{Mat}(n \times n, \mathbb{K})$ ähnlich zu einer Matrix $E^r_{n\times n}$, so existiert eine Matrix $A \in \text{GL}(n, \mathbb{K})$ mit $N = A \cdot E^r_{n\times n} \cdot A^{-1}$, und es folgt

$$N \cdot N = A \cdot E^r_{n\times n} \cdot A^{-1} \cdot A \cdot E^r_{n\times n} \cdot A^{-1} = A \cdot E^r_{n\times n} \cdot E^r_{n\times n} \cdot A^{-1} = A \cdot E^r_{n\times n} \cdot A^{-1} = N.$$

(ii)⇒(iii): Gilt $N \cdot N = N$, so ist die lineare Abbildung $\phi_N : \mathbb{K}^n \to \mathbb{K}^n$, $v \mapsto N \cdot v$ eine Projektionsabbildung, und N ist die beschreibende Matrix von ϕ_N bezüglich der Standardbasis.

(iii)⇒(i): Ist N beschreibende Matrix einer Projektionsabbildung $\phi : V \to V$ bezüglich einer Basis A von V, so ist N nach Satz 4.3.3 ähnlich zu der beschreibende Matrix von ϕ für jede andere geordnete Basis A' von V. Nach Lemma 4.3.7 existiert eine Basis von V, bezüglich der ϕ durch $E^r_{n\times n}$ beschrieben wird, und somit ist N ähnlich zu $E^r_{n\times n}$. □

4.4 Duale Abbildungen und Transposition

In der Beschreibung von linearen Abbildungen durch Matrizen spielen offensichtlich die Zeilen und Spalten der Matrix eine unterschiedliche Rolle. Sind $A = (v_1, ..., v_n)$ und $B = (w_1, ..., w_n)$ geordnete Basen von endlich-dimensionalen Vektorräumen V und W und $\phi : V \to W$ eine lineare Abbildung, so enthält die $i$te *Spalte* der beschreibenden Matrix ${}_B M_A(\phi)$ die Koeffizienten des Vektors $\phi(v_i)$ als Linearkombination von Vektoren in B.

Andererseits können wir aus einer $(m \times n)$-Matrix N eine $(n \times m)$-Matrix N^T konstruieren, indem wir die Einträge in der $i$ten Spalte von N als Einträge in der $i$te Zeile von N^T wählen. Die resultierende Matrix N^T, deren Spalten dann gerade die Zeilen von N sind, wird als die *Transponierte* von N bezeichnet.

Während die Matrix N eine lineare Abbildung $\phi_N : \mathbb{K}^n \to \mathbb{K}^m$, $v \mapsto N \cdot v$ definiert, erhält man für N^T eine lineare Abbildung $\phi_{N^T} : \mathbb{K}^m \to \mathbb{K}^n$, $w \mapsto N^T \cdot w$. Allerdings ist nicht offensichtlich, welche lineare Abbildung durch die Matrix N^T beschrieben wird, wenn N die beschreibende Matrix einer beliebigen linearen Abbildung $\phi : V \to W$ bezüglich beliebiger Basen ist. Um das zu klären, befassen wir uns zuerst mit den Eigenschaften der Transposition.

Definition 4.4.1: Sei $A = (a_{ji}) \in \text{Mat}(m \times n, \mathbb{K})$. Dann heißt die Matrix $A^T = (a_{ij}) \in \text{Mat}(n \times m, \mathbb{K})$ die zu A **transponierte Matrix**.

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \quad \Rightarrow \quad A^T = \begin{pmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{mn} \end{pmatrix}.$$

Eine Matrix $A = (a_{ji}) \in \text{Mat}(n \times n, \mathbb{K})$ heißt **symmetrisch**, wenn $A = A^T$, und **antisymmetrisch**, wenn $A^T = -A$ gilt.

Satz 4.4.2: Sei $\mathbb{K}$ ein Körper und $m, n, p \in \mathbb{N}$.

1. Die Transposition von Matrizen definiert einen $\mathbb{K}$-Vektorraumisomorphismus
$$T : \mathrm{Mat}(m \times n, \mathbb{K}) \to \mathrm{Mat}(n \times m, \mathbb{K}), \quad A \mapsto A^T.$$
2. Für beliebige Matrizen A in $\mathrm{Mat}(m \times n, \mathbb{K})$, $B \in \mathrm{Mat}(n \times p, \mathbb{K})$ gilt $(A \cdot B)^T = B^T \cdot A^T$.
3. Es gilt $\mathbb{1}_n^T = \mathbb{1}_n$ und $(A^{-1})^T = (A^T)^{-1}$ für alle $A \in \mathrm{GL}(n, \mathbb{K})$.

Beweis:
Die Transpositionsabbildung $T : \mathrm{Mat}(m \times n, \mathbb{K}) \to \mathrm{Mat}(n \times m, \mathbb{K})$ ist $\mathbb{K}$-linear, denn nach den Rechenregeln für Matrixmultiplikation und Skalarmultiplikation gilt für $A, B \in \mathrm{Mat}(m \times n, \mathbb{K})$, $\lambda \in \mathbb{K}$ und alle $i \in \{1, .., n\}$, $j \in \{1, ..., m\}$
$$(A+B)^T_{ji} = (A+B)_{ij} = A_{ij} + B_{ij} = A^T_{ji} + B^T_{ji} \qquad (\lambda A)^T_{ji} = (\lambda A)_{ij} = \lambda A_{ij} = \lambda A^T_{ji}.$$
Da per Definition der Transposition $(A^T)^T = A$ für alle $A \in \mathrm{Mat}(m \times n)$ gilt, ist die Transpositionsabbildung invertierbar und somit ein Vektorraumisomorphismus. Die Aussage 2. beweist man durch Nachrechnen: Für $A \in \mathrm{Mat}(m \times n, \mathbb{K})$ und $B \in \mathrm{Mat}(n \times p, \mathbb{K})$ erhält man
$$(A \cdot B)^T_{ji} = (A \cdot B)_{ij} = \Sigma_{k=1}^n A_{ik} B_{kj} = \Sigma_{k=1}^n A^T_{ki} B^T_{jk} = (B^T \cdot A^T)_{ji},$$
für alle $i \in \{1, ..., p\}$ und $j \in \{1, ..., m\}$. Die Aussage $\mathbb{1}_n^T = \mathbb{1}_n$ ist offensichtlich. Ist $A \in \mathrm{GL}(n, \mathbb{K})$, so erhält man damit $(A^{-1})^T \cdot A^T = (A \cdot A^{-1})^T = \mathbb{1}_n^T = \mathbb{1}_n = (A^{-1} \cdot A)^T = A^T \cdot (A^{-1})^T$ und mit der Eindeutigkeit der Inversen $(A^T)^{-1} = (A^{-1})^T$. □

Um ein Analogon der Transposition von Matrizen für lineare Abbildungen $\phi : V \to W$ zwischen endlich-dimensionalen Vektorräumen zu finden, erinnern wir uns an den Zusammenhang zwischen beschreibenden Matrizen linearer Abbildungen und Zeilen- und Spaltenvektoren aus Bemerkung 4.2.7. Ist $A = (v_1, ..., v_n)$ eine geordnete Basis eines $\mathbb{K}$-Vektorraums V und statten wir den Körper $\mathbb{K}$ mit der geordneten Basis (1) aus, so ist

- der Spaltenvektor $(\lambda_1, \ldots, \lambda_n)^T$ die beschreibende Matrix ${}_AM_{(1)}(L)$ der $\mathbb{K}$-linearen Abbildung $L : \mathbb{K} \to V$, $\lambda \mapsto \lambda \Sigma_{i=1}^n \lambda_i v_i$.
- der Zeilenvektor $(\lambda_1, ..., \lambda_n)$ die beschreibende Matrix ${}_{(1)}M_A(L^*)$ der $\mathbb{K}$-linearen Abbildung $L^* : V \to \mathbb{K}$, $\Sigma_{i=1}^n \lambda_i v_i \mapsto \Sigma_{i=1}^n \lambda_i$.

Um die Transposition zu verallgemeinern, müssen wir also nicht nur den Vektorraum V sondern auch den Vektorraum $\mathrm{Hom}_{\mathbb{K}}(V, \mathbb{K})$ der $\mathbb{K}$-linearen Abbildungen von V nach $\mathbb{K}$ betrachten.

Definition 4.4.3: Sei V ein Vektorraum über $\mathbb{K}$. Der **Dualraum** von V ist der Vektorraum $V^* := \mathrm{Hom}_{\mathbb{K}}(V, \mathbb{K})$. Elemente von V^* heißen **Linearformen**.

Bemerkung 4.4.4:

1. Ist V ein endlich-dimensionaler $\mathbb{K}$-Vektorraum, so gilt nach Satz 4.2.17 $\dim_{\mathbb{K}}(V^*) = \dim_{\mathbb{K}}(\mathbb{K}) \cdot \dim_{\mathbb{K}}(V) = \dim_{\mathbb{K}}(V)$ und damit sind V und V^* nach Korollar 4.2.5 isomorph.
2. Nach Satz 4.2.17 bilden dann für jede geordnete Basis $B = (v_1, ..., v_n)$ von V die linearen Abbildungen $\vartheta^i : V \to \mathbb{K}$ mit $\vartheta^i(v_i) = 1$ und $\vartheta^i(v_j) = 0$ für $i \neq j$ eine geordnete Basis $B^* = (\vartheta^1, ..., \vartheta^n)$ von V^*. Sie heißt die zu B **duale Basis**.

Beispiel 4.4.5: Ist $V = \mathbb{R}^3$ und $B = (e_1, e_2, e_3)$ die Standardbasis, so enthält die dazu duale Basis $B^* = (\vartheta^1, \vartheta^2, \vartheta^3)$ die Linearformen $\vartheta^i : \mathbb{R}^3 \to \mathbb{R}$, $(x_1, x_2, x_3)^T \mapsto x_i$, die einem Vektor im $\mathbb{R}^3$ seinen $i$ten Eintrag zuordnen, denn $\vartheta^i(e_i) = 1$ und $\vartheta^i(e_j) = 0$ für $i \neq j$.

Wählt man stattdessen die Basis $C = (c_1, c_2, c_3) = (e_1, e_1 + e_2, e_1 + e_2 + e_3)$, so erhält man die duale Basis $C^* = (\omega^1, \omega^2, \omega^3)$ mit

$$\omega^1 \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_1 - x_2, \quad \omega^2 \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_2 - x_3, \quad \omega^3 \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_3.$$

denn $(x_1, x_2, x_3)^T = (x_1 - x_2)c_1 + (x_2 - x_3)c_2 + x_3 c_3$ und $\omega^i(c_i) = 1$, $\omega^i(c_j) = 0$ für $i \neq j$.

Die Aussagen von Bemerkung 4.4.4 sind falsch für unendlich-dimensionale $\mathbb{K}$-Vektorräume V. Dort erhält man zwar ebenfalls für jedes Element $b \in B$ einer Basis B von V eine eindeutig bestimmte lineare Abbildung $\vartheta^b : V \to \mathbb{K}$ mit $\vartheta^b(b) = 1$ und $\vartheta^b(b') = 0$ für $b' \in B \setminus \{b\}$, aber die Menge $M = \{\vartheta^b | b \in B\} \subseteq V^*$ ist keine Basis von V^*. Denn aus der Endlichkeit der Linearkombinationen ergibt sich

$$\mathrm{span}_{\mathbb{K}}(M) = \{\alpha \in V^* | \alpha(b) = 0 \text{ für fast alle } b \in B\}.$$

Da man die Werte einer $\mathbb{K}$-linearen Abbildung auf B mit Satz 4.2.1 aber beliebig vorgeben kann, gibt es aber auch $\mathbb{K}$-lineare Abbildungen $\psi : V \to \mathbb{K}$, die die Bedingung $\psi(b) = 0$ für fast alle $b \in B$ nicht erfüllen, etwa die mit $\psi(b) = 1$ für alle $b \in B$. Dies illustriert das folgende Beispiel.

Beispiel 4.4.6: Für den Vektorraum $\mathbb{K}[x] = \{p \in \mathrm{Abb}(\mathbb{N}_0, \mathbb{K}) \mid p(n) = 0 \text{ für fast alle } n \in \mathbb{N}_0\}$ der Polynome über einem Körper $\mathbb{K}$ ist der Dualraum isomorph zum $\mathbb{K}$-Vektorraum $\mathrm{Abb}(\mathbb{N}_0, \mathbb{K})$ der Folgen in $\mathbb{K}$. Ein $\mathbb{K}$-Vektorraumisomorphismus ist die Abbildung

$$\phi : \mathrm{Abb}(\mathbb{N}_0, \mathbb{K}) \to \mathbb{K}[x]^*, f \mapsto \phi(f) \quad \text{mit} \quad \phi(f)(\Sigma_{n=0}^\infty a_n x^n) = \Sigma_{n=0}^\infty a_n f(n).$$

Beweis:
Die Abbildung ϕ ist wohldefiniert, denn für jedes Polynom $p = \Sigma_{n=0}^\infty a_n x^n$ gilt $a_n = 0$ für fast alle $n \in \mathbb{N}_0$. Sie ist $\mathbb{K}$-linear, denn

$$\begin{aligned} &\phi(f+g)(p) = \Sigma_{n=0}^\infty a_n (f(n) + g(n)) = \Sigma_{n=0}^\infty a_n f(n) + \Sigma_{n=0}^\infty a_n g(n) = \phi(f)(p) + \phi(g)(p) \\ &\phi(\lambda f)(p) = \Sigma_{n=0}^\infty a_n \lambda f(n) = \lambda \Sigma_{n=0}^\infty a_n f(n) = \lambda \phi(f)(p) \end{aligned}$$

und damit $\phi(f + g) = \phi(f) + \phi(g)$ und $\phi(\lambda f) = \lambda\phi(f)$ für jedes $\lambda \in \mathbb{K}$.

Sie ist injektiv, denn aus $\phi(f) = 0$ folgt $\phi(f)(x^n) = f(n) = 0$ für alle $n \in \mathbb{N}_0$ und damit $f = 0$. Sie ist surjektiv, denn jede $\mathbb{K}$-lineare Abbildung $\alpha : \mathbb{K}[x] \to \mathbb{K}$ ist durch ihre Werte den Monomen eindeutig bestimmt. Setzt man $f(n) = \alpha(x^n)$ für alle $n \in \mathbb{N}_0$, so ergibt sich

$$\phi(f)(p) = \Sigma_{n=0}^\infty a_n f(n) = \Sigma_{n=0}^\infty a_n \alpha(n) = \alpha(\Sigma_{n=0}^\infty a_n x^n) = \alpha(p) \;\forall p \in \mathbb{K}[x] \;\Rightarrow\; \phi(f) = \alpha.$$ □

Mit Hilfe des Dualraums können wir ein Gegenstück der Transposition für beliebige lineare Abbildungen $\phi : U \to V$ zwischen endlich-dimensionalen $\mathbb{K}$-Vektorräumen U und V definieren. Dies ist die *duale Abbildung* $\phi^* : V^* \to U^*$, $\alpha \mapsto \alpha \circ \phi$. Sie ist das Gegenstück der Transposition, da analoge Aussagen zu Satz 4.4.2 gelten und da die Transponierte der beschreibenden Matrix von ϕ bezüglich zweier beleibiger Basen gerade die beschreibende Matrix von ϕ^* bezüglich der dualen Basen ist.

Satz 4.4.7: Seien U, V, W Vektorräume über $\mathbb{K}$ und $\phi : U \to V$, $\psi : V \to W$ $\mathbb{K}$-lineare Abbildungen. Dann gilt:

1. Für jede $\mathbb{K}$-lineare Abbildung $\phi : U \to V$ ist auch die Abbildung $\phi^* : V^* \to U^*$, $\alpha \mapsto \alpha \circ \phi$ $\mathbb{K}$-linear. Sie heißt die zu ϕ **duale Abbildung**.

2. Dies definiert für alle $\mathbb{K}$-Vektorräume U, V eine injektive $\mathbb{K}$-lineare Abbildung
$$* : \mathrm{Hom}_{\mathbb{K}}(U, V) \to \mathrm{Hom}_{\mathbb{K}}(V^*, U^*), \quad \phi \mapsto \phi^*.$$

3. Es gilt $\mathrm{id}_V^* = \mathrm{id}_{V^*}$ sowie $(\psi \circ \phi)^* = \phi^* \circ \psi^*$ für alle lineare Abbildungen $\phi : U \to V$, $\psi : V \to W$ und $(\phi^{-1})^* = (\phi^*)^{-1}$ für alle invertierbaren linearen Abbildungen $\phi : U \to V$.

4. Sind U, V endlich-dimensional mit geordneten Basen A, B und A^* und B^* die dazu dualen Basen, so gilt für jede lineare Abbildung $\phi : U \to V$
$${}_{A^*}M_{B^*}(\phi^*) = {}_{B}M_A(\phi)^T.$$

Beweis:
1. Die Linearität von ϕ^* ergibt sich durch direktes Nachrechnen. Für alle $\alpha, \beta \in V^*$, $\lambda \in \mathbb{K}$ gilt
$$\phi^*(\alpha+\beta) = (\alpha+\beta)\circ\phi = \alpha\circ\phi + \beta\circ\phi = \phi^*(\alpha)+\phi^*(\beta), \; \phi^*(\lambda\alpha) = (\lambda\alpha)\circ\phi = \lambda(\alpha\circ\phi) = \lambda\phi^*(\alpha)$$
und somit ist $\phi^* : V^* \to U^*$ $\mathbb{K}$-linear.

2. Die $\mathbb{K}$-Linearität der Abbildung $*$ ergibt sich ebenfalls durch direktes Nachrechnen. Es gilt
$$\begin{aligned}
&(\phi+\psi)^*(\alpha) = \alpha\circ(\phi+\psi) \overset{\bullet}{=} \alpha\circ\phi + \alpha\circ\psi = \phi^*(\alpha)+\psi^*(\alpha) = (\phi^*+\psi^*)(\alpha)\\
&(\lambda\phi)^*(\alpha) = \alpha\circ(\lambda\phi) \overset{\bullet}{=} \lambda(\alpha\circ\phi) = \lambda\phi^*(\alpha) = (\lambda\phi^*)(\alpha)
\end{aligned}$$
für alle $\phi, \psi \in \mathrm{Hom}_{\mathbb{K}}(U, V)$, $\alpha \in V^*$ und $\lambda \in \mathbb{K}$, wobei in $\bullet$ die $\mathbb{K}$-Linearität von α benutzt wurde. Damit folgt $(\phi+\psi)^* = \phi^*+\psi^*$ und $(\lambda\phi)^* = \lambda\phi^*$.

Für die Injektivität von $*$ zeigen wir, dass für jede $\mathbb{K}$-lineare Abbildung $\phi : U \to V$ mit $\phi \neq 0$ ein $\alpha \in V^*$ mit $\phi^*(\alpha) = \alpha\circ\phi \neq 0$ existiert. Damit ist gezeigt, dass aus $\phi \neq 0$ auch $\phi^* \neq 0$ folgt und somit $\ker(*) = \{0\}$.

Ist $\phi \neq 0$, so gibt es ein $u \in U$ mit $\phi(u) \neq 0$. Wir ergänzen die linear unabhängige Menge $\{\phi(u)\} \subseteq V$ zu einer Basis B von V und definieren mit Satz 4.2.1 eine $\mathbb{K}$-lineare Abbildung $\alpha : V \to \mathbb{K}$, indem wir ihre Werte auf B vorgeben, nämlich als $\alpha(\phi(u)) = 1$ und $\alpha(b) = 0$ für $b \in B \setminus \{\phi(u)\}$. Dann ist $\phi^*(\alpha)(u) = (\alpha\circ\phi)(u) = \alpha(\phi(u)) = 1$ und damit $\phi^*(\alpha) = \alpha\circ\phi \neq 0$.

3. Die Identitäten in 3. folgen direktes Nachrechnen: für alle $\alpha \in V^*$ und $\beta \in W^*$ gilt
$$\begin{aligned}
&\mathrm{id}_V^*(\alpha) = \alpha\circ\mathrm{id}_V = \alpha \qquad \forall \alpha \in V^*\\
&(\phi\circ\psi)^*(\beta) = \beta\circ(\phi\circ\psi) = (\beta\circ\phi)\circ\psi = \psi^*(\beta\circ\phi) = \psi^*(\phi^*(\beta)) = (\psi^*\circ\phi^*)(\beta),
\end{aligned}$$
und damit ist $\mathrm{id}_V^* = \mathrm{id}_{V^*}$ und $(\psi\circ\phi)^* = \phi^*\circ\psi^*$. Für bijektive $\phi \in \mathrm{Hom}_{\mathbb{K}}(U, V)$ folgt
$$(\phi^{-1})^*\circ\phi^* = (\phi\circ\phi^{-1})^* = \mathrm{id}_V^* = \mathrm{id}_{V^*} \qquad \phi^*\circ(\phi^{-1})^* = (\phi^{-1}\circ\phi)^* = \mathrm{id}_U^* = \mathrm{id}_{U^*},$$
und damit $(\phi^{-1})^* = (\phi^*)^{-1}$.

4. Seien $A = (u_1, ..., u_m)$, $B = (v_1, ..., v_n)$ geordnete Basen von U, V und $A^* = (\omega^1, ..., \omega^m)$, $B^* = (\vartheta^1, ..., \vartheta^n)$ die dazu dualen Basen. Dann ist die beschreibende Matrix einer $\mathbb{K}$-linearen

Abbildung $\phi : U \to V$ mit $\phi(u_i) = \Sigma_{j=1}^n \phi_{ji} v_j$ gegeben durch ${}_B M_A(\phi)_{ji} = \phi_{ji}$. Daraus ergibt sich für alle $i \in \{1, ..., m\}$ und $j \in \{1, ..., n\}$

$$\phi^*(\vartheta^j)(u_i) = \vartheta^j(\phi(u_i)) = \vartheta^j(\Sigma_{k=1}^n \phi_{ki} v_k) = \Sigma_{k=1}^n \phi_{ki} \vartheta^j(v_k) = \phi_{ji} = \Sigma_{k=1}^m \phi_{jk} \omega^k(u_i)$$

und somit $\phi^*(\vartheta^j) = \Sigma_{k=1}^m \phi_{jk} \omega^k$. Also gilt ${}_{A^*} M_{B^*}(\phi^*)_{ij} = \phi_{ji} = {}_B M_A(\phi)_{ji}$ und ${}_{A^*} M_{B^*}(\phi^*) = {}_B M_A(\phi)^T$. □

Beispiel 4.4.8: Wir betrachten die Standardbasen $A = (e_1, e_2, e_3)$ des $\mathbb{R}^3$ und $B = (e_1, e_2)$ des $\mathbb{R}^2$, ihre dualen Basen $A^* = (\alpha^1, \alpha^2, \alpha^3)$ und $B^* = (\beta^1, \beta^2)$ sowie die $\mathbb{R}$-lineare Abbildung

$$\phi : \mathbb{R}^3 \to \mathbb{R}^2, \qquad \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mapsto \begin{pmatrix} x_1 - x_2 + x_3 \\ 2x_1 + x_2 + 2x_3 \end{pmatrix}$$

Dann ergibt sich die beschreibende Matrix

$${}_B M_A(\phi) = \begin{pmatrix} 1 & -1 & 1 \\ 2 & 1 & 2 \end{pmatrix}.$$

Zur Berechnung der beschreibenden Matrix von ϕ^* bezüglich B^* und A^* berechnet man die Bilder der Basisvektoren in B^* als Linearkombinationen von Vektoren aus A^*

$$\phi^*(\beta^1) = \beta^1 \circ \phi = \alpha^1 - \alpha^2 + \alpha^3 : \qquad \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \xrightarrow{\phi} \begin{pmatrix} x_1 - x_2 + x_3 \\ 2x_1 + x_2 + 2x_3 \end{pmatrix} \xrightarrow{\beta^1} x_1 - x_2 + x_3$$

$$\phi^*(\beta^2) = \beta^1 \circ \phi = 2\alpha^1 + \alpha^2 + 2\alpha^3 : \qquad \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \xrightarrow{\phi} \begin{pmatrix} x_1 - x_2 + x_3 \\ 2x_1 + x_2 + 2x_3 \end{pmatrix} \xrightarrow{\beta^2} 2x_1 + x_2 + 2x_3.$$

Daraus ergibt sich die beschreibende Matrix

$${}_{A^*} M_{B^*}(\phi^*) = \begin{pmatrix} 1 & 2 \\ -1 & 1 \\ 1 & 2 \end{pmatrix} = {}_B M_A(\phi)^T.$$

Die Aussagen 2. und 3. in Satz 4.4.7 sind das direkte Gegenstück von Aussagen 1., 2. und 3. in Satz 4.4.2 für lineare Abbildungen. Der einzige Unterschied ist, dass für unendlich-dimensionale $\mathbb{K}$-Vektorräume U, V die Abbildung $*$ im Gegensatz zur Transpositionsabbildung nicht notwendigerweise surjektiv ist. Aussage 4. in Satz 4.4.7 zeigt, dass sich durch Betrachtung der dualen Basen das Bilden der dualen Abbildung in die Transposition von Matrizen übersetzt. Damit verallgemeinert das Konzept der dualen Abbildung die Transposition von Matrizen.

Da für jede Matrix $A \in \mathrm{Mat}(m \times n, \mathbb{K})$ die Identität $(A^T)^T = A$ gilt, stellt sich nun die Frage, ob eine analoge Identität auch für duale Abbildungen existiert. Die doppelt duale Abbildung $\phi^{**} : U^{**} \to V^{**}$ einer $\mathbb{K}$-linearen Abbildung $\phi : U \to V$ ist aber zunächst keine Abbildung von U nach V ist. Wir müssen also zunächst eine Beziehung zwischen $V^{**} = (V^*)^*$ und V sowie $U^{**} = (U^*)^*$ und U herstellen und anschließend die linearen Abbildungen $\phi : U \to V$ und $\phi^{**} : U^{**} \to V^{**}$ vergleichen.

Dazu stellen wir fest, dass es für jeden $\mathbb{K}$-Vektorraum V eine *kanonische* $\mathbb{K}$-lineare Abbildung $\mathrm{kan}_V : V \to V^{**}$ gibt, also eine Abbildung, deren Definition keine willkürlichen Wahlen wie Basiswahlen erfordert. Wir können nämlich einem Vektor $v \in V$ die lineare Abbildung $f_v \in V^{**} = \mathrm{Hom}_{\mathbb{K}}(V^*, \mathbb{K})$ mit $f_v(\alpha) = \alpha(v)$ für alle $\alpha \in V^*$ zuordnen.

Satz 4.4.9: Zu jedem $\mathbb{K}$-Vektorraum V existiert eine kanonische $\mathbb{K}$-lineare Abbildung $\mathrm{kan}_V : V \to V^{**}$, $v \mapsto f_v$ mit $f_v(\alpha) = \alpha(v)$ für alle $\alpha \in V^*$, die **kanonische Abbildung**. Sie ist injektiv. Ist V endlich-dimensional, so ist sie ein Isomorphismus von $\mathbb{K}$-Vektorräumen.

Beweis:
1. Wir zeigen, dass $f_v : V^* \to \mathbb{K}$, $\alpha \mapsto \alpha(v)$ für festes $v \in V$ eine $\mathbb{K}$-lineare Abbildung ist, also ein Element von V^{**}:

$$f_v(\alpha+\beta) = (\alpha+\beta)(v) = \alpha(v) + \beta(v) = f_v(\alpha) + f_v(\beta) \qquad f_v(\lambda\alpha) = (\lambda\alpha)(v) = \lambda\alpha(v) = \lambda f_v(\alpha)$$

für alle $\alpha, \beta \in V^*$ und $\lambda \in \mathbb{K}$. Also ist $f_v \in \mathrm{Hom}_{\mathbb{K}}(V^*, \mathbb{K}) = V^{**}$ für alle $v \in V$. Wir erhalten also tatsächlich eine Abbildung $\mathrm{kan}_V : V \to V^{**}$, $v \mapsto f_v$.

2. Wir zeigen, dass die Abbildung $\mathrm{kan}_V : V \to V^{**}$, $v \mapsto f_v$ $\mathbb{K}$-linear ist: für alle $v, w \in V$, $\lambda \in \mathbb{K}$ und $\alpha \in V^*$ gilt:

$$f_{v+w}(\alpha) = \alpha(v+w) = \alpha(v) + \alpha(w) = f_v(\alpha) + f_w(\alpha) \qquad f_{\lambda v}(\alpha) = \alpha(\lambda v) = \lambda\alpha(v) = \lambda f_v(\alpha)$$

und somit für alle $v, w \in V$, $\lambda \in \mathbb{K}$

$$\mathrm{kan}_V(v+w) = f_{v+w} = f_v + f_w = \mathrm{kan}_V(v) + \mathrm{kan}_V(w), \qquad \mathrm{kan}_V(\lambda v) = f_{\lambda v} = \lambda f_v = \lambda \mathrm{kan}_V(v).$$

3. Wir zeigen, dass $\mathrm{kan}_V : V \to V^{**}$, $v \mapsto f_v$ injektiv ist, also dass $\ker(\mathrm{kan}_V) = \{0\}$ gilt. Dazu zeigen wir, dass für $v \in V \setminus \{0\}$ auch $\mathrm{kan}_V(v) = f_v \neq 0 : V^* \to \mathbb{K}$ ist, also dass es zu $v \in V \setminus \{0\}$ ein $\alpha \in V^*$ gibt mit $f_v(\alpha) = \alpha(v) \neq 0$.

Ist $v \in V \setminus \{0\}$, so können wir die linear unabhängige Menge $\{v\} \subseteq V$ zu einer Basis B von V ergänzen und mit Satz 4.2.1 eine lineare Abbildung $\alpha : V \to \mathbb{K}$ durch $\alpha(v) = 1$ und $\alpha(b) = 0$ für $b \in B \setminus \{v\}$ definieren. Also existiert für alle $v \in V \setminus \{0\}$ ein Element $\alpha \in V^*$ mit $f_v(\alpha) = \alpha(v) = 1$, und es folgt $f_v \neq 0$ für alle $v \in V \setminus \{0\}$. Damit ist gezeigt, dass $\ker(\mathrm{kan}_V) = \{0\}$ und $\mathrm{kan}_V : V \to V^{**}$ injektiv ist.

Ist V endlich-dimensional so folgt daraus mit $\dim_{\mathbb{K}}(V^{**}) = \dim_{\mathbb{K}}(V^*) = \dim_{\mathbb{K}}(V)$ und Korollar 4.1.14, dass $\mathrm{kan}_V : V \to V^{**}$ bijektiv ist. □

Mit Hilfe der kanonischen Abbildungen $\mathrm{kan}_U : U \to U^{**}$ und $\mathrm{kan}_V : V \to V^{**}$ können wir nun die Vektorräume U^{**} und V^{**} zu U und V in Beziehung setzen und die linearen Abbildungen $\phi : U \to V$ und $\phi^{**} : U^{**} \to V^{**}$ vergleichen.

Satz 4.4.10: Seien U, V Vektorräume über $\mathbb{K}$ mit Dualräumen U^*, V^* und U^{**}, V^{**} die Dualräume von U^*, V^*. Dann gilt für alle $\mathbb{K}$-linearen Abbildungen $\phi : U \to V$

$$\phi^{**} \circ \mathrm{kan}_U = \mathrm{kan}_V \circ \phi.$$

Sind U, V endlich-dimensional mit geordneten Basen $A = (u_1, ..., u_m)$, $B = (v_1, ..., v_n)$, so gilt

$${}_{B^{**}}M_{A^{**}}(\phi^{**}) = {}_BM_A(\phi) \text{ mit } A^{**} = (\mathrm{kan}_U(u_1), ..., \mathrm{kan}_U(u_m)),\ B^{**} = (\mathrm{kan}_V(v_1), ..., \mathrm{kan}_V(v_n)).$$

Beweis:
1. Für eine lineare Abbildung $\phi : U \to V$ ist die duale Abbildung $\phi^* : V^* \to U^*$, $\alpha \mapsto \alpha \circ \phi$ und die duale Abbildung der dualen Abbildung $\phi^{**} : U^{**} \to V^{**}$, $g \mapsto g \circ \phi^*$. Per Definition der dualen Abbildung un der kanonischen Abbildung gilt dann für alle $u \in U$ und $\alpha \in V^*$

$$(\phi^{**}(\mathrm{kan}_U(u)))(\alpha) = (\mathrm{kan}_U(u) \circ \phi^*)(\alpha) = \mathrm{kan}_U(u)(\phi^*(\alpha)) = \mathrm{kan}_U(u)(\alpha \circ \phi) = (\alpha \circ \phi)(u)$$
$$= \alpha(\phi(u)) = \mathrm{kan}_V(\phi(u))(\alpha).$$

Daraus folgt $\phi^{**}(\mathrm{kan}_U(u)) = \mathrm{kan}_V(\phi(u))$ für alle $u \in U$ und damit $\phi^{**} \circ \mathrm{kan}_U = \mathrm{kan}_V \circ \phi$. Ist U endlich-dimensional, so ist $\mathrm{kan}_U : U \to U^{**}$ invertierbar, und es folgt $\phi^{**} = \mathrm{kan}_V \circ \phi \circ \mathrm{kan}_U^{-1}$.

2. Ist $\dim_{\mathbb{K}} V < \infty$ und $B = (v_1, ..., v_n)$ eine geordnete Basis von V und $B^* = (\vartheta^1, ..., \vartheta^n)$ die dazu duale Basis, dann folgt $\mathrm{kan}_V(v_i)(\vartheta^i) = \vartheta^i(v_i) = 1$ und $\mathrm{kan}_V(v_j)(\vartheta^i) = \vartheta^i(v_j) = 0$ für $i \neq j$. Also ist $B^{**} = (\mathrm{kan}_V(v_1), ..., \mathrm{kan}_V(v_n))$ die duale Basis zu B^*. Ebenso ergibt sich für jede geordnete Basis $A = (u_1, ..., u_m)$ von U auch $A^{**} = (\mathrm{kan}_U(u_1), ..., \mathrm{kan}_U(u_m))$.

Mit Satz 4.4.7, 4. folgt dann ${}_{B^{**}}M_{A^{**}}(\phi^{**}) = {}_{A^*}M_{B^*}(\phi^*)^T = ({}_BM_A(\phi)^T)^T = {}_BM_A(\phi)$. □

Um unsere Betrachtungen zu Dualräumen und dualen Abbildungen abzuschließen betrachten wir den Rang von dualen Abbildungen zwischen endlich-dimensionalen Vektorräumen. Es zeigt sich, dass für jede $\mathbb{K}$-lineare Abbildung $\phi : U \to V$ die duale Abbildung $\phi^* : V^* \to U^*$ den gleichen Rang wie ϕ hat. Daraus lassen sich hilfreiche Aussagen über Matrizen ableiten, die den Rang einer Matrix und ihrer Transponierten in Beziehung setzen.

Korollar 4.4.11: Sind U, V endlich-dimensionale $\mathbb{K}$-Vektorräume und $\phi : U \to V$ eine $\mathbb{K}$-lineare Abbildung, so gilt $\mathrm{rg}(\phi) = \mathrm{rg}(\phi^*)$.

Beweis:
Sei $\phi : U \to V$ eine lineare Abbildung und $\phi^* : V^* \to U^*$, $\alpha \mapsto \alpha \circ \phi$ die dazu duale Abbildung. Wir wählen eine geordnete Basis $(b_1, ..., b_r)$ von $\mathrm{im}(\phi)$, ergänzen sie zu einer geordneten Basis $B = (b_1, ..., b_r, b_{r+1}, ..., b_n)$ von V und betrachten die dazu duale Basis $B^* = (\vartheta^1, ..., \vartheta^n)$ von V^*. Dann gilt $\ker(\phi^*) = \{\alpha \in V^* | \mathrm{im}(\phi) \subseteq \ker(\alpha)\} = \mathrm{span}_{\mathbb{K}}(\{\vartheta^{r+1}, ..., \vartheta^n\})$. Denn für $\alpha = \Sigma_{j=1}^n \lambda_j \vartheta^j$ gilt $\mathrm{im}(\phi) \subseteq \ker(\alpha)$ genau dann, wenn $\alpha(b_i) = \Sigma_{j=1}^n \lambda_j \vartheta^j(b_i) = \lambda_i = 0$ für alle $i \in \{1, ..., r\}$, also genau dann, wenn $\alpha \in \mathrm{span}_{\mathbb{K}}(\{\vartheta^{r+1}, ..., \vartheta^n\})$. Daraus folgt mit der Dimensionsformel $\mathrm{rg}(\phi^*) = n - \mathrm{def}(\phi^*) = n - (n - r) = r = \mathrm{rg}(\phi)$. □

Korollar 4.4.12: Für jede Matrix $A \in \mathrm{Mat}(m \times n, \mathbb{K})$ gilt: $\mathrm{rg}(A) = \mathrm{rg}(A^T)$.

Beweis:
Die Matrix A ist die beschreibende Matrix der linearen Abbildung $\phi_A : \mathbb{K}^n \to \mathbb{K}^m$, $v \mapsto A \cdot v$ bezüglich der Standardbasis, und nach Satz 4.4.7 ist A^T die beschreibende Matrix der dualen Abbildung $\phi_A^* : (\mathbb{K}^m)^* \to (\mathbb{K}^n)^*$ bezüglich der dazu dualen Basen. Per Definition des Rangs und nach Korollar 4.4.11 gilt dann $\mathrm{rg}(A) = \mathrm{rg}(\phi_A) = \mathrm{rg}(\phi_A^*) = \mathrm{rg}(A^T)$. □

Korollar 4.4.13: Für jede Matrix $A \in \mathrm{Mat}(m \times n, \mathbb{K})$ ist die maximale Anzahl $\mathrm{zrg}(A)$ linear unabhängiger Zeilen gleich der maximalen Anzahl $\mathrm{srg}(A)$ linear unabhängiger Spalten gleich dem Rang von A:

$$\mathrm{zrg}(A) = \mathrm{srg}(A) = \mathrm{rg}(A).$$

Die maximale Anzahl linear unabhängiger Zeilen in A heißt **Zeilenrang** von A.

Beweis:
Nach Bemerkung 4.2.15 gilt $\mathrm{rg}(A) = \mathrm{srg}(A)$. Da die Zeilen von A die Spalten von A^T sind und umgekehrt, folgt mit Korollar 4.4.12 $\mathrm{zrg}(A) = \mathrm{srg}(A^T) = \mathrm{rg}(A^T) = \mathrm{rg}(A) = \mathrm{srg}(A)$. □

4.5 Lineare Gleichungssysteme und Gaußscher Algorithmus

Wir betrachten nun noch eine Anwendung der Matrixrechnung, nämlich das Lösen von linearen Gleichungssystemen. Diese ist sehr wichtig, da viele Zusammenhänge in den Naturwissenschaften sowie Technik und Wirtschaft zumindest näherungsweise durch Systeme linearer Gleichungen beschrieben werden können. Im Gegensatz zu nicht-linearen Gleichungssystemen, deren Lösungen sehr schwierig zu finden sind, gibt es für lineare Gleichungssysteme Algorithmen, also systematische Lösungsverfahren, mit denen sich *alle* Lösungen konstruieren lassen.

Definition 4.5.1: Sei $\mathbb{K}$ ein Körper. Ein **lineares Gleichungssystem** über $\mathbb{K}$ in den Variablen $x_1, ..., x_n$ ist ein System von Gleichungen der Form

$$\begin{array}{ll} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 & a_{ji}, b_j \in \mathbb{K} \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 & \forall j \in \{1, ..., m\}, i \in \{1, ..., n\} \\ \vdots \quad \vdots \qquad \vdots \quad \vdots \quad \vdots & \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m. & \end{array} \tag{11}$$

Gilt $b_1 = \ldots = b_m = 0$, so nennt man das Gleichungssystem **homogen**, sonst **inhomogen**. Ein n-Tupel $(x_1, ..., x_n)^T \in \mathbb{K}^n$, das alle Gleichungen in (11) erfüllt, heißt **Lösung** von (11). Die Matrizen

$$A = \begin{pmatrix} a_{11} & a_{21} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix} \qquad (A, b) = \begin{pmatrix} a_{11} & a_{21} & \ldots & a_{1n} & b_1 \\ a_{21} & a_{22} & \ldots & a_{2n} & b_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} & b_m \end{pmatrix}$$

heißen **Koeffizientenmatrix** und **erweiterte Koeffizientenmatrix** des Gleichungssystems.

Offensichtlich können wir ein lineares Gleichungssystem auch als eine einzige vektorwertige Gleichung betrachten. Interpretiert man die Variablen $x_1, ..., x_n$ aus Gleichung (11) als die Einträge eines n-Tupels $x = (x_1, ..., x_n)^T \in \mathbb{K}^n$ und die Elemente $b_1, ..., b_m \in \mathbb{K}$ als Einträge eines m-Tupels $b = (b_1, ..., b_m)^T \in \mathbb{K}^m$, so lässt sich das Gleichungssystem (11) schreiben als

$$A \cdot x = \begin{pmatrix} a_{11} & a_{21} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} = b$$

Indem wir die Eigenschaften von Matrizen und Vektoren ausnutzen, können wir damit Aussagen über die Lösungen gewinnen.

Satz 4.5.2: Sei $\mathbb{K}$ ein Körper, $A \in \mathrm{Mat}(m \times n, \mathbb{K})$ und $b \in \mathbb{K}^m$.

1. Die Lösungsmenge des homogenen Gleichungssystems $Ax = 0$ ist der Untervektorraum $L(A, 0) = \ker(A) \subseteq \mathbb{K}^n$ der Dimension $\mathrm{def}(A) = n - \mathrm{rg}(A)$.

2. Ist $x \in \mathbb{K}^n$ eine Lösung des inhomogenen Gleichungssystems $Ax = b$, so ist die Lösungsmenge des inhomogenen Gleichungssystems $Ax = b$

$$L(A, b) = x + \ker(A) = \{x + z|\ z \in \ker(A)\} \subseteq \mathbb{K}^n.$$

Beweis:
Nach Definition 4.2.14 und Bemerkung 4.2.15 ist $L(A, 0) = \{x \in \mathbb{K}^n | A \cdot x = 0\} = \ker(A)$. Dies ist nach dem Dimensionssatz ein Untervektorraum des $\mathbb{K}^n$ der Dimension $\mathrm{def}(A) = n - \mathrm{rg}(A)$.

Ist $x \in \mathbb{K}^n$ eine Lösung des inhomogenen Gleichungssystems $A \cdot x = b$ und $z \in \mathbb{K}^n$ eine Lösung des homogenen Gleichungssystems $A \cdot z = 0$, so folgt $A \cdot (x + z) = A \cdot x + A \cdot z = b + 0 = b$ und damit $x + z \in L(A, b)$. Also gilt $x + \ker(A) \subseteq L(A, b)$.

Sind umgekehrt $x, x' \in \mathbb{K}^n$ Lösungen des inhomogenen Gleichungssystems, so ist $x - x'$ eine Lösung des zugehörigen homogenen Gleichungssystems, denn $A \cdot (x - x') = A \cdot x - A \cdot x' = b - b = 0$. Also gilt für jede Lösung x' des inhomogenen Gleichungssystems $x' = x + x' - x \in x + \ker(A)$ und somit $L(A, b) \subseteq x + \ker(A)$. $\square$

Die Lösungen eines *inhomogenen* Gleichungssystems $Ax = b$ bilden für $b \neq 0$ *keinen* Untervektorraum des $\mathbb{K}^n$, denn der Nullvektor ist dann nicht in $L(A, b)$ enthalten. Andererseits ist die Lösungsmenge eng verwandt mit einem Untervektorraum des $\mathbb{K}^n$, da die Differenz $y - y'$ zweier Lösungen $y, y' \in L(A, b)$ immer in dem Untervektorraum $L(A, 0) = \ker(A)$ liegt. Nichtleere Teilmengen eines Vektorraums mit dieser Eigenschaft bezeichnet man als *affine Unterräume*.

Definition 4.5.3: Sei V ein $\mathbb{K}$-Vektorraum.

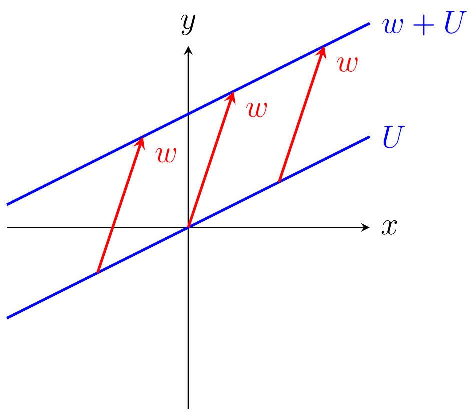
1. Ein **affiner Unterraum** von V ist eine Teilmenge der Form

$$w + U = \{w + u | u \in U\} \subseteq V$$

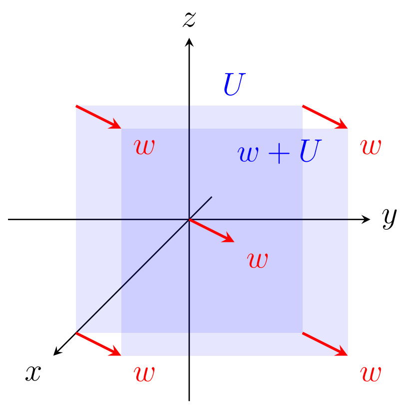
mit einem Untervektorraum $U \subseteq V$ und einem Vektor $w \in V$.

2. Die **Dimension** des affinen Unterraums A ist $\dim_{\mathbb{K}}(w + U) := \dim_{\mathbb{K}}(U)$.

Einen affinen Unterraum der Dimension 1 bezeichnet man auch als **affine Gerade**, einen affinen Unterraum der Dimension 2 als **affine Ebene** und einen affinen Unterraum der Dimension $\dim_{\mathbb{K}}(V) - 1$ für endlich-dimensionale Vektorräume V als **affine Hyperebene**.



affine Gerade im $\mathbb{R}^2$



affine Ebene im $\mathbb{R}^3$.

Nach dieser Definition können wir die Lösungsmenge jeder *Gleichung* $a_{k1}x_1 + \ldots + a_{kn}x_n = b_k$ im linearen Gleichungssystem (11), in der nicht alle Koeffizienten $a_{k1}, ..., a_{kn}$ verschwinden, als eine affine Hyperebene im $\mathbb{K}^n$ auffassen. Die Lösungsmenge des *Gleichungssystems* ist die Schnittmenge der Lösungsmengen der einzelnen Gleichungen und somit ein Schnitt von affinen Hyperebenen. Man kann zeigen (Übung), dass der Schnitt zweier affiner Unterräume eines $\mathbb{K}$-Vektorraums V entweder leer oder wieder ein affiner Unterraum von V ist. Dementsprechend ist die Lösungsmenge des inhomogenen Gleichungssystems entweder leer oder ein affiner Unterraum des $\mathbb{K}^n$. Welcher der beiden Fälle eintritt, ergibt sich aus dem folgenden Kriterium.

Satz 4.5.4: Sei $\mathbb{K}$ ein Körper, $A \in \mathrm{Mat}(m \times n, \mathbb{K})$ und $b \in \mathbb{K}^m$. Dann gilt $L(A, b) \neq \emptyset$ genau dann, wenn $\mathrm{rg}(A, b) = \mathrm{rg}(A)$.

Beweis:
Das Gleichungssystem $Ax = b$ hat genau dann eine Lösung, wenn $b \in \mathbb{K}^m$ im Bild der linearen Abbildung $\phi_A : \mathbb{K}^n \to \mathbb{K}^m$, $v \mapsto A \cdot v$ liegt. Dies ist genau dann der Fall, wenn $\mathrm{rg}(A, b) = \mathrm{srg}(A, b) = \dim_{\mathbb{K}}(\mathrm{im}(\phi_A) + \mathbb{K}b) = \dim_{\mathbb{K}}(\mathrm{im}(\phi_A)) = \mathrm{srg}(A) = \mathrm{rg}(A)$ gilt. □

Offensichtlich ist man nicht nur daran interessiert, Lösungen eines linearen Gleichungssystems abstrakt zu untersuchen, sondern möchte *konkrete Lösungen* finden. Gleichungssysteme, für die dies besonders einfach ist sind die Gleichungssysteme mit Koeffizientenmatrizen in der *speziellen Zeilenstufenform.*

Definition 4.5.5:

1. Eine Matrix $A \in \mathrm{Mat}(m \times n, \mathbb{K})$ ist in **Zeilenstufenform**, wenn für alle $j \in \{2, .., m\}$ gilt: sind für ein $k \in \{1, ..., n\}$ die ersten $(k-1)$ Einträge der $(j-1)$ten Zeile gleich Null, so sind auch die ersten k Einträge der $j$ten Zeile gleich Null.

2. Eine Matrix A ist in **spezieller Zeilenstufenform**, wenn sie in Zeilenstufenform ist und aus $a_{j1} = ... = a_{j(k-1)} = 0$ und $a_{jk} \neq 0$ folgt $a_{jk} = 1$.

Die Einträge a_{jk} mit $a_{j1} = ... = a_{j(k-1)} = 0$ und $a_{jk} \neq 0$ heißen **Pivots**.

Beispiel 4.5.6: Die folgenden Matrizen sind in spezieller Zeilenstufenform:

$$\begin{pmatrix} 0 & 1 & 7 \\ 0 & 0 & 1 \end{pmatrix} \quad \begin{pmatrix} 1 & 4 & 7 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 1 & 4 & 7 \\ 0 & 1 & 8 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 1 & 4 & 7 \\ 0 & 1 & 8 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

Die folgenden Matrizen sind in Zeilenstufenform, aber nicht in spezieller Zeilenstufenform:

$$\begin{pmatrix} 1 & 4 & 7 \\ 0 & 0 & 2 \end{pmatrix} \quad \begin{pmatrix} 1 & 4 & 7 \\ 0 & 0 & 8 \\ 0 & 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 1 & 4 & 7 \\ 0 & 2 & 8 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 1 & 4 & 7 \\ 0 & 1 & 8 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}$$

Die folgenden Matrizen sind nicht in Zeilenstufenform:

$$\begin{pmatrix} 0 & 4 & 7 \\ 0 & 1 & 1 \end{pmatrix} \quad \begin{pmatrix} 1 & 4 & 7 \\ 0 & 0 & 8 \\ 1 & 4 & 1 \end{pmatrix} \quad \begin{pmatrix} 1 & 4 & 7 \\ 0 & 1 & 8 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 1 & 4 & 7 \\ 0 & 1 & 8 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

Bemerkung 4.5.7:

1. Ist $A \in \mathrm{Mat}(m \times n, \mathbb{K})$ eine Matrix in Zeilenstufenform, so kann man A durch Vertauschen von Spalten in eine Matrix in Zeilenstufenform überführen, für die $a_{11} \neq 0$, $a_{22} \neq 0$,..., $a_{rr} \neq 0$ gilt und deren $(r+1)$te bis $m$te Zeile nur Nullen enthalten. Das entspricht einem Umnummerieren der Variablen $x_1, ..., x_n$ im zugehörigen linearen Gleichungssystem.

$$\begin{aligned} 2x_1 + 3x_2 + 4x_3 + 0x_4 &= 5 \\ -x_3 - 2x_4 &= 1 \end{aligned} \qquad \to \qquad \begin{aligned} 2x_1 + 4x_3 + 3x_2 + 0x_4 &= 5 \\ -x_3 + 0x_2 - 2x_4 &= 1 \end{aligned}$$

   Da sich dabei die Anzahl der Stufen nicht ändert, muss die Anzahl der Stufen gleich $\mathrm{rg}(A) = \mathrm{zrg}(A) = \mathrm{srg}(A) = r$ sein.

2. Ist A von dieser Form, so hat das zugehörige homogene Gleichungssystem die Form

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1r}x_r + \ldots + a_{1n}x_n &= 0 \\ a_{22}x_2 + \ldots + a_{1r}x_r + \ldots + a_{2n}x_n &= 0 \\ &\vdots \\ a_{rr}x_r + \ldots + a_{rn}x_n &= 0 \\ 0 &= 0 \\ &\vdots \\ 0 &= 0 \end{aligned}$$

   mit $a_{11}, ..., a_{rr} \in \mathbb{K} \setminus \{0\}$. Die Variablen $x_{r+1}, ..., x_n$ heißen dann **freie Variable** und die Variablen $x_1, ..., x_r$ **gebundene Variable**.

Ist die Koeffizientenmatrix eines linearen Gleichungssystems in Zeilenstufenform, so kann man direkt feststellen, ob eine Lösung existiert und gegebenenfalls eine konkrete Lösung angeben. Ob eine Lösung existiert, ergibt sich aus dem folgenden Satz, dessen Beweis einen Algorithmus zum Bestimmen der Lösung liefert.

Satz 4.5.8: Sei $A \in \mathrm{Mat}(m \times n, \mathbb{K})$ eine Matrix in Zeilenstufenform und $b \in \mathbb{K}^m$. Dann ist $L(A, b) = \emptyset$ genau dann, wenn es ein $j \in \{1, ..., m\}$ gibt mit $a_{j1} = ... = a_{jn} = 0$ und $b_j \neq 0$.

Beweis:
Gibt es ein $j \in \{1, ..., m\}$ mit $a_{j1} = ... = a_{jn} = 0$ und $b_j \neq 0$, so enthält das Gleichungssystem die widersprüchliche Gleichung $0 = b_j \neq 0$ und hat damit offensichtlich keine Lösung. Andererseits können wir für jede Matrix A in Zeilenstufenform durch Vertauschen der 1ten bis $n$ten Spalten erreichen, dass $a_{11} \neq 0$, $a_{22} \neq 0$,..., $a_{rr} \neq 0$ mit $r = \mathrm{rg}(A)$ gilt und die $(r+1)$te bis $m$te Zeile nur Nullen enthalten. Dies entspricht dem linearen Gleichungssystem

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1r}x_r + a_{1(r+1)}x_{r+1} + \ldots + a_{1n}x_n &= b_1 \\ a_{21}x_2 + \ldots + a_{2r}x_r + a_{2(r+1)}x_{r+1} + \ldots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{rr}x_r + a_{r(r+1)}x_{r+1} + \ldots + a_{rn}x_n &= b_r \\ 0 &= b_{r+1} \\ &\vdots \\ 0 &= b_m. \end{aligned}$$

Da dieses Vertauschen von Spalten einer Umnummerierung der Variablen entspricht, ändert sich die Anzahl der Lösungen dadurch nicht. Ist $b_{r+1} = ... = b_m = 0$, so erhalten wir eine Lösung, indem wir $x_{r+1} = ... = x_n = 0$ setzen und das resultierende Gleichungssystem mit r Gleichungen und r Variablen von unten nach oben lösen.

Dabei addieren wir jeweils Vielfache der $(k+1)$ten bis $r$ten Gleichung zur $k$ten Gleichung hinzu, so dass die Variablen $x_{k+1}, ..., x_r$ aus der $k$ten Gleichung verschwinden. Also existiert für $b_{r+1} = ... = b_m = 0$ mindestens eine Lösung. Da wir nur die ersten n Spalten vertauscht haben, ist die Bedingung $b_{r+1} = ... = b_m = 0$ gleichbedeutend dazu, dass in der ursprünglichen Matrix A aus $a_{j1} = ... = a_{jn} = 0$ für ein ein $j \in \{1, ..., m\}$ auch $b_j = 0$ folgt. □

Beispiel 4.5.9: Wir finden *eine* Lösung des Gleichungssystems in spezieller Zeilenstufenform

$$\begin{array}{rrrr}
x_1 - 3x_2 + 4x_3 + x_4 = 5 & x_1 - 3x_2 + 4x_3 = 5 & x_1 - 3x_2 = 13 & x_1 = 1 \\
x_2 - x_3 - 2x_4 = -2 & x_2 - x_3 = -2 & x_2 = -4 & x_2 = -4 \\
x_3 + 3x_4 = -2 & x_3 = -2 & x_3 = -2 & x_3 = -2 \\
 & x_4 = 0 & x_4 = 0 & x_4 = 0.
\end{array}$$

Hier ist $r = 3$, da drei Stufen vorhanden sind, und somit ist die einzige freie Variable die Variable x_4. Diese kann auf Null gesetzt werden, wenn man nur *eine* Lösung benötigt. Wir finden nun *alle* Lösungen, indem wir die freie Variable nicht entfernen, sondern als freien Parameter behandeln:

$$\begin{array}{rrrr}
x_1 - 3x_2 + 4x_3 + x_4 = 5 & x_1 - 3x_2 + 4x_3 = 5 - \lambda & x_1 - 3x_2 = 13 + 11\lambda & x_1 = 1 + 8\lambda \\
x_2 - x_3 - 2x_4 = -2 & x_2 - x_3 = -2 + 2\lambda & x_2 = -4 - \lambda & x_2 = -4 - \lambda \\
x_3 + 3x_4 = -2 & x_3 = -2 - 3\lambda & x_3 = -2 - 3\lambda & x_3 = -2 - 3\lambda \\
 & x_4 = \lambda & x_4 = \lambda & x_4 = \lambda.
\end{array}$$

Damit erhalten wir als Lösungsmenge einen affinen Unterraum des $\mathbb{R}^4$ der Dimension 1:

$$L = \left\{ \begin{pmatrix} 1 + 8\lambda \\ -4 - \lambda \\ -2 - 3\lambda \\ \lambda \end{pmatrix} | \lambda \in \mathbb{R} \right\} = \begin{pmatrix} 1 \\ -4 \\ -2 \\ 0 \end{pmatrix} + \mathbb{R} \begin{pmatrix} 8 \\ -1 \\ -3 \\ 1 \end{pmatrix}.$$

Der zugehörige eindimensionale Untervektorraum ist gerade die Lösungsmenge des entsprechenden homogenen linearen Gleichungssystems.

Da sich Gleichungssysteme, deren Koeffizientenmatrix in Zeilenstufenform ist, von unten nach oben leicht lösen lassen, können wir jedes Gleichungssystem lösen, wenn wir es in Zeilenstufenform überführen können, ohne dabei seinen Lösungsraum zu verändern.

Schon aus der Schule sind bestimmte Umformungen an Gleichungssystemen bekannt, die sich nicht auf den Lösungsraum auswirken, nämlich (i) das Multiplizieren von Gleichungen mit einer Konstante $\lambda \in \mathbb{K} \setminus \{0\}$, (ii) das Vertauschen zweier Gleichungen und (iii) das Hinzuaddieren eines Vielfachen einer Gleichung zu einer anderen.

Wir realisieren nun diese Umformungen an Gleichungssystemen durch die Linksmultiplikation seiner (erweiterten) Koeffizientenmatrix mit bestimmten invertierbaren Matrizen.

Definition 4.5.10: Wir definieren für $j, k \in \{1, ..., m\}$ mit $j \neq k$ und $\lambda \in \mathbb{K}\setminus\{0\}$ die folgenden $(m \times m)$-Matrizen in $\mathrm{GL}(m, \mathbb{K})$, die auch als **Elementarmatrizen** bezeichnet werden[12]

$$M_j(\lambda) = \mathbb{1}_m + (\lambda - 1)E_{jj} = \begin{pmatrix} 1 & 0 & \ldots & \ldots & \ldots & \ldots & 0 \\ 0 & \ddots & \ddots & & & & \vdots \\ \vdots & \ddots & 1 & \ddots & & & \vdots \\ \vdots & & \ddots & \lambda & \ddots & & \vdots \\ \vdots & & & \ddots & 1 & \ddots & \vdots \\ \vdots & & & & \ddots & \ddots & 0 \\ 0 & \ldots & \ldots & \ldots & \ldots & 0 & 1 \end{pmatrix}$$

$$V_{jk} = \mathbb{1}_m - E_{jj} - E_{kk} + E_{kj} + E_{jk} = \begin{pmatrix} 1 & 0 & \ldots & \ldots & \ldots & \ldots & 0 \\ 0 & \ddots & \ddots & & & & \vdots \\ \vdots & \ddots & 0 & \ldots & 1 & & \vdots \\ \vdots & & \vdots & 1 & \vdots & & \vdots \\ \vdots & & 1 & \ldots & 0 & \ddots & \vdots \\ \vdots & & & & \ddots & \ddots & 0 \\ 0 & \ldots & \ldots & \ldots & \ldots & 0 & 1 \end{pmatrix}$$

$$A_{jk}(\lambda) = \mathbb{1}_m + \lambda E_{jk} = \begin{pmatrix} 1 & 0 & \ldots & \ldots & \ldots & \ldots & 0 \\ 0 & \ddots & \ddots & & & & \vdots \\ \vdots & \ddots & 1 & \ddots & & \lambda & \vdots \\ \vdots & & \ddots & 1 & \ddots & & \vdots \\ \vdots & & & \ddots & 1 & \ddots & \vdots \\ \vdots & & & & \ddots & \ddots & 0 \\ 0 & \ldots & \ldots & \ldots & \ldots & 0 & 1 \end{pmatrix}$$

Nach der Definition sind die Einträge der Elementarmatrizen gegeben durch

$$M_j(\lambda)_{li} = \begin{cases} 1 & l = i \neq j \\ \lambda & l = i = j \\ 0 & \text{sonst} \end{cases} \qquad (V_{jk})_{li} = \begin{cases} 1 & l = i \notin \{j, k\} \\ 1 & \{l, i\} = \{j, k\} \\ 0 & \text{sonst} \end{cases} \qquad A_{jk}(\lambda)_{li} = \begin{cases} 1 & i = l \\ \lambda & l = j, i = k \\ 0 & \text{sonst.} \end{cases}$$

Eine einfache Rechnung zeigt, dass die Elementarmatrizen tatsächlich invertierbar sind mit Inversen $M_j(\lambda)^{-1} = M_j(\lambda^{-1})$, $V_{jk}^{-1} = V_{jk}$, $A_{jk}(\lambda)^{-1} = A_{jk}(-\lambda)$.

Mit Hilfe der Elementarmatrizen können wir nun die bekannten Umformungen (i)-(iii) an linearen Gleichungssystemen realisieren, ohne dabei den Lösungsraum zu verändern.

[12]Das ist ungünstig, da man diese Bezeichnung auch für die Matrizen E_{ij} benutzt, aber der allgemeine Sprachgebrauch.

Satz 4.5.11: Sei $N \in \mathrm{Mat}(m \times n, \mathbb{K})$. Dann entspricht

(i) die Linksmultiplikation $N \mapsto M_j(\lambda) \cdot N$ der Multiplikation der $j$ten Zeile von N mit dem Skalar $\lambda \in \mathbb{K} \setminus \{0\}$,

(ii) die Linksmultiplikation $N \mapsto V_{jk} \cdot N$ dem Vertauschen der $j$ten und $k$ten Zeile von N,

(iii) die Linksmultiplikation $N \mapsto A_{jk}(\lambda) \cdot N$ dem Hinzuaddieren der mit dem Skalar $\lambda \in \mathbb{K} \setminus \{0\}$ multiplizierten $k$ten Zeile von N zur $j$ten Zeile von N.

Diese linearen Abbildungen heißen **elementare Zeilenumformungen** von N.

Ist $N = (A, b)$ und entsteht $N' = (A', b')$ aus N durch elementare Zeilenumformungen, so haben die linearen Gleichungssysteme $Ax = b$ und $A'x = b'$ die gleiche Lösungsmenge.

Beweis:
Die Aussagen (i), (ii), (iii) ergeben sich durch eine direkte Rechnung (Übung). Daraus folgt insbesondere, dass die Inversen der Elementarmatrizen gegeben sind durch $M_j(\lambda)^{-1} = M_j(\lambda^{-1})$, $V_{jk}^{-1} = V_{jk}$ und $A_{jk}(\lambda) = A_{jk}(-\lambda)$, denn die inversen Zeilenumformungen sind, respektive, die Multiplikation der $j$ten Zeile mit $\lambda^{-1} \in \mathbb{K} \setminus \{0\}$, das Vertauschen der $j$ten und $k$ten Zeile und das Hinzuaddieren der mit $-\lambda$ multiplizierten $k$ten Zeile zur $j$ten Zeile.

Dass sich die Lösungsmenge eines linearen Gleichungssystems unter den Linksmultiplikation mit den Elementarmatrizen nicht ändert ergibt sich wie folgt. Wegen

$$(A,b) \cdot \begin{pmatrix} x \\ -1 \end{pmatrix} = \begin{pmatrix} a_{11} & \ldots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \ldots & a_{mn} & b_m \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \\ -1 \end{pmatrix} = \begin{pmatrix} \Sigma_{j=1}^n a_{1j} x_j - b_1 \\ \vdots \\ \Sigma_{j=1}^n a_{mj} x_j - b_m \end{pmatrix} = A \cdot x - b$$

gilt $x = (x_1, ..., x_n)^T \in L(A, b)$ genau dann, wenn $(x_1, .., x_n, -1)^T \in \ker(A, b)$.

Da die Elementarmatrizen invertierbar sind, ändert sich $\ker(A, b)$ unter Linksmultiplikation von (A, b) mit den Elementarmatrizen nicht, denn für $M \in \mathrm{GL}(m, \mathbb{K})$ und beliebige Matrizen $B \in \mathrm{Mat}(m \times p, \mathbb{K})$ gilt $\ker(B) = \ker(M \cdot B)$:

$$x \in \ker(B) \Rightarrow B \cdot x = 0 \Rightarrow M \cdot B \cdot x = M \cdot 0 = 0 \Rightarrow x \in \ker(M \cdot B)$$
$$x \in \ker(M \cdot B) \Rightarrow M \cdot B \cdot x = 0 \Rightarrow B \cdot x = M^{-1} \cdot M \cdot B \cdot x = M^{-1} \cdot 0 = 0 \Rightarrow x \in \ker(B).$$

□

Mit Hilfe der Elementarmatrizen und den zugehörigen elementaren Zeilenumformungen können wir nun einen Algorithmus angeben, der eine beliebige Matrix $A \in \mathrm{Mat}(m \times n, \mathbb{K})$ in spezielle Zeilenstufenform überführt, den *Gauß-Algorithmus*.

Satz 4.5.12: (**Gauß-Algorithmus für Matrizen**)

Sei $A \in \mathrm{Mat}(m \times n, \mathbb{K})$ eine beliebige Matrix. Dann kann A in eine Matrix in spezieller Zeilenstufenform überführt werden, indem man die folgenden Schritte durchführt:

1. Man vertauscht Zeilen, so dass in der ersten Zeile der erste Eintrag $\neq 0$ nicht weiter rechts steht, als in irgendeiner anderen Zeile.

2. Man multipliziert die erste Zeile und jede andere Zeile, deren erster Eintrag $\neq 0$ in der gleichen Spalte steht wie der der ersten Zeile mit einem Element $\lambda \in \mathbb{K} \setminus \{0\}$, so dass dieser Eintrag 1 wird.
3. Man subtrahiert die erste Zeile von diesen anderen Zeilen.
4. Ist die resultierende Matrix noch nicht in spezieller Zeilenstufenform, so wendet man 1.-3. auf die Matrix A' an, die durch Entfernen der ersten Zeile aus dieser Matrix entsteht.

So erhält man eine Matrix $A' \in \mathrm{Mat}(m \times n, \mathbb{K})$ in spezieller Zeilenstufenform mit $\mathrm{rg}(A') = \mathrm{rg}(A)$. Der Rang von A' ist die Anzahl der Pivots.

Beweis:
Es ist offensichtlich, dass dieser Algorithmus eine Matrix $A' \in \mathrm{Mat}(m \times n, \mathbb{K})$ in spezieller Zeilenstufenform liefert. Zu zeigen ist noch, dass $\mathrm{rg}(A') = \mathrm{rg}(A)$ gilt und gleich der Anzahl der Pivots ist.

Nach Satz 4.5.11 entspricht jeder Schritt im Gauß-Algorithmus der Linksmultiplikation mit einer invertierbaren Elementarmatrix Matrix $M \in \mathrm{GL}(m, \mathbb{K})$. Nach Definition 4.3.2 sind für alle die Matrizen $B \in \mathrm{Mat}(m \times n, \mathbb{K})$ die Matrizen B und $M \cdot B$ äquivalent, und nach Satz 4.3.4 haben äquivalente Matrizen den gleichen Rang.

Der Rang einer Matrix in spezieller Zeilenstufenform ist gleich der Anzahl der Pivots, da jede Pivot einen neuen Spaltenvektor liefert, der nicht im Spann der links davon stehenden Spaltenvektoren liegt, während Spaltenvektoren, die keine Pivot enthalten im Spann der links davon stehenden Spaltenvektoren enthalten sind. □

Beispiel 4.5.13:

$$\begin{pmatrix} 0 & 0 & 8 & 3 \\ 2 & 0 & -4 & -4 \\ 3 & 6 & 9 & 0 \end{pmatrix} \xrightarrow{V_{12}} \begin{pmatrix} 2 & 0 & -4 & -4 \\ 0 & 0 & 8 & 3 \\ 3 & 6 & 9 & 0 \end{pmatrix} \xrightarrow{M_1(\frac{1}{2})} \begin{pmatrix} 1 & 0 & -2 & -2 \\ 0 & 0 & 8 & 3 \\ 3 & 6 & 9 & 0 \end{pmatrix} \xrightarrow{M_3(\frac{1}{3})}$$
$$\begin{pmatrix} 1 & 0 & -2 & -2 \\ 0 & 0 & 8 & 3 \\ 1 & 2 & 3 & 0 \end{pmatrix} \xrightarrow{A_{31}(-1)} \begin{pmatrix} 1 & 0 & -2 & -2 \\ 0 & 0 & 8 & 3 \\ 0 & 2 & 5 & 2 \end{pmatrix} \xrightarrow{V_{23}} \begin{pmatrix} 1 & 0 & -2 & -2 \\ 0 & 2 & 5 & 2 \\ 0 & 0 & 8 & 3 \end{pmatrix} \xrightarrow{M_2(\frac{1}{2})}$$
$$\begin{pmatrix} 1 & 0 & -2 & -2 \\ 0 & 1 & \frac{5}{2} & 1 \\ 0 & 0 & 8 & 3 \end{pmatrix} \xrightarrow{M_3(\frac{1}{8})} \begin{pmatrix} 1 & 0 & -2 & -2 \\ 0 & 1 & \frac{5}{2} & 1 \\ 0 & 0 & 1 & \frac{3}{8} \end{pmatrix}$$

Indem man den Gauß- Algorithmus aus Satz 4.5.12 auf die erweiterte Koeffizientenmatrix eines linearen Gleichungssystems anwendet, erhält man dann ein Gleichungssystem in spezieller Zeilenstufenform, das die gleichen Lösungen besitzt wie das ursprüngliche Gleichungssystem.

Korollar 4.5.14: (**Gauß-Algorithmus für Gleichungssysteme**)
Sei $A \in \mathrm{Mat}(m \times n, \mathbb{K})$ und $b \in \mathbb{K}^m$. Dann kann man die Lösungen des Gleichungssystems $Ax = b$ finden, indem man die folgenden Schritte durchführt.

1. Man stellt die erweiterte Koeffizientenmatrix $(A, b) \in \mathrm{Mat}(m \times (n+1), \mathbb{K})$ auf.

2. Man überführt die erweiterte Koeffizientenmatrix (A, b) mit dem Algorithmus aus Satz 4.5.12 in spezielle Zeilenstufenform.
3. Man überprüft, ob Lösungen existieren. Dies ist nach Satz 4.5.8 genau dann der Fall, wenn die in 2. konstruierte Matrix keine Zeile hat, in der alle Einträge außer dem letzten $=0$ sind und der letzte $\neq 0$ ist.
4. Wenn Lösungen existieren, löst man das Gleichungssystem von unten nach oben, wie in Beispiel 4.5.9.

Beispiel 4.5.15: Wir lösen das lineare Gleichungssystem

$$\begin{aligned} x_1 - 3x_2 + 4x_3 + x_4 &= 5 \\ 2x_1 + 4x_2 - 2x_3 - 18x_4 &= -10 \\ x_1 + 7x_3 + 13x_4 &= -13. \end{aligned}$$

Die erweiterte Koeffizientenmatrix ist

$$(A, b) = \begin{pmatrix} 1 & -3 & 4 & 1 & 5 \\ 2 & 4 & -2 & -18 & -10 \\ 1 & 0 & 7 & 13 & -13 \end{pmatrix}$$

Wir überführen sie in spezielle Zeilenstufenform

$$\begin{aligned} &\begin{pmatrix} 1 & -3 & 4 & 1 & 5 \\ 2 & 4 & -2 & -18 & -10 \\ 1 & 0 & 7 & 13 & -13 \end{pmatrix} \xrightarrow{M_2(\frac{1}{2})} \begin{pmatrix} 1 & -3 & 4 & 1 & 5 \\ 1 & 2 & -1 & -9 & -5 \\ 1 & 0 & 7 & 13 & -13 \end{pmatrix} \\ &\xrightarrow{A_{21}(-1)} \begin{pmatrix} 1 & -3 & 4 & 1 & 5 \\ 0 & 5 & -5 & -10 & -10 \\ 1 & 0 & 7 & 13 & -13 \end{pmatrix} \xrightarrow{A_{31}(-1)} \begin{pmatrix} 1 & -3 & 4 & 1 & 5 \\ 0 & 5 & -5 & -10 & -10 \\ 0 & 3 & 3 & 12 & -18 \end{pmatrix} \\ &\xrightarrow{M_2(\frac{1}{5})} \begin{pmatrix} 1 & -3 & 4 & 1 & 5 \\ 0 & 1 & -1 & -2 & -2 \\ 0 & 3 & 3 & 12 & -18 \end{pmatrix} \xrightarrow{M_3(\frac{1}{3})} \begin{pmatrix} 1 & -3 & 4 & 1 & 5 \\ 0 & 1 & -1 & -2 & -2 \\ 0 & 1 & 1 & 4 & -6 \end{pmatrix} \\ &\xrightarrow{A_{32}(-1)} \begin{pmatrix} 1 & -3 & 4 & 1 & 5 \\ 0 & 1 & -1 & -2 & -2 \\ 0 & 0 & 2 & 6 & -4 \end{pmatrix} \xrightarrow{M_3(\frac{1}{2})} \begin{pmatrix} 1 & -3 & 4 & 1 & 5 \\ 0 & 1 & -1 & -2 & -2 \\ 0 & 0 & 1 & 3 & -2 \end{pmatrix} \end{aligned}$$

Das resultierende Gleichungssystem in spezieller Zeilenstufenform ist das aus Beispiel 4.5.9

$$\begin{aligned} x_1 - 3x_2 + 4x_3 + x_4 &= 5 \\ x_2 - x_3 - 2x_4 &= -2 \\ x_3 + 3x_4 &= -2. \end{aligned}$$

Also ist die Lösungsmenge des Gleichungssystems nach Beispiel 4.5.9

$$L(A, b) = \left\{ \begin{pmatrix} 1 + 8\lambda \\ -4 - \lambda \\ -2 - 3\lambda \\ \lambda \end{pmatrix} | \lambda \in \mathbb{R} \right\} = \begin{pmatrix} 1 \\ -4 \\ -2 \\ 0 \end{pmatrix} + \mathbb{R} \begin{pmatrix} 8 \\ -1 \\ -3 \\ 1 \end{pmatrix}.$$

Eine weitere sehr nützliche Anwendung des Gauß-Algorithmus und der elementaren Zeilenumformungen ist das Invertieren von Matrizen und die Untersuchung von Matrizen auf Invertierbarkeit. Dazu nutzt man aus, dass die Anzahl der Stufen für eine Matrix A in Zeilenstufenform gerade der Rang von A ist.

Also ist eine Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{K})$ in spezieller Zeilenstufenform genau dann invertierbar, wenn sie n Stufen aufweist oder, anders ausgedrückt, in der Diagonale keine Nullen sondern nur Einsen stehen. In diesem Fall kann man die Matrix durch weitere elementare Zeilenumformungen in die Einheitsmatrix überführen, indem man Vielfache der darunter liegenden Zeilen von jeder Zeile abzieht. Wendet man die gleichen Zeilenumformungen auf die Einheitsmatrix an, so erhält man die Inverse A^{-1}.

Satz 4.5.16: (**Invertieren von Matrizen mit dem Gauß-Algorithmus**)
Sei $A \in \mathrm{Mat}(n \times n, \mathbb{K})$. Dann kann man mit Hilfe des folgenden Algorithmus herausfinden, ob A invertierbar ist und in diesem Fall das Inverse von A bestimmen:

1. Man überführt A durch elementare Zeilenumformungen wie in Satz 4.5.12 in eine Matrix A' in spezieller Zeilenstufenform und führt die selben Umformungen an $\mathbb{1}_n$ durch.

2. Die Matrix A ist genau dann invertierbar, wenn A' in der Diagonale nur Einsen enthält.

3. Ist A invertierbar, so setzt man alle Einträge in der letzten Spalte von A' außer dem letzten auf 0, indem man entsprechende Vielfache der letzten Zeile von der 1. bis $(n-1)$ten Zeile abzieht. Die gleichen Zeilenoperationen werden jeweils auf die aus der Matrix $\mathbb{1}_n$ entstandene Matrix angewandt.

4. Man wiederholt den letzten Schritt von rechts nach links für jede Spalte. Dabei zieht man für die $k$te Spalte jeweils entsprechende Vielfache der $k$ten Zeile von der $(k-1)$ten bis zur 1. Zeile ab. Die gleichen Zeilenoperationen werden jeweils auf die aus der Matrix $\mathbb{1}_n$ entstandene Matrix angewandt.

Der Algorithmus überführt A in die Einheitsmatrix. Die dadurch aus der Einheitsmatrix entstandene Matrix ist die Inverse von A.

Beweis:
Die Schritte im Algorithmus entsprechen der Linksmultiplikation mit Elementarmatrizen. Also liefert der Algorithmus Elementarmatrizen $M_1, ..., M_k \in \mathrm{GL}(n, \mathbb{K})$ mit $\mathbb{1}_n = M_k \cdots M_1 \cdot A$. Daraus folgt $M_k \cdots M_1 \cdot \mathbb{1}_n = M_k \cdots M_1 = A^{-1}$. □

Um eine Matrix$A$ effizient zu invertieren, bietet es sich an, die Matrizen A und $\mathbb{1}_n$ entweder direkt nebeneinander zu schreiben, oder sie gleich in eine Matrix $(A, \mathbb{1}_n) \in \mathrm{Mat}(n \times (2n), \mathbb{K})$ zu kombinieren, an der dann die elementaren Zeilenumformungen durchgeführt werden.

Beispiel 4.5.17: Wir untersuchen die Matrix

$$A = \begin{pmatrix} 0 & 1 & -4 \\ 1 & 2 & -1 \\ 1 & 1 & 2 \end{pmatrix}.$$

auf Invertierbarkeit. In diesem Fall liefert der Algorithmus

$$
\begin{aligned}
(A, \mathbb{1}_3) = &\left(\begin{array}{ccc|ccc} 0 & 1 & -4 & 1 & 0 & 0 \\ 1 & 2 & -1 & 0 & 1 & 0 \\ 1 & 1 & 2 & 0 & 0 & 1 \end{array}\right) \xrightarrow{V_{13}} \left(\begin{array}{ccc|ccc} 1 & 1 & 2 & 0 & 0 & 1 \\ 1 & 2 & -1 & 0 & 1 & 0 \\ 0 & 1 & -4 & 1 & 0 & 0 \end{array}\right) \xrightarrow{A_{21}(-1)} \\
&\left(\begin{array}{ccc|ccc} 1 & 1 & 2 & 0 & 0 & 1 \\ 0 & 1 & -3 & 0 & 1 & -1 \\ 0 & 1 & -4 & 1 & 0 & 0 \end{array}\right) \xrightarrow{A_{32}(-1)} \left(\begin{array}{ccc|ccc} 1 & 1 & 2 & 0 & 0 & 1 \\ 0 & 1 & -3 & 0 & 1 & -1 \\ 0 & 0 & -1 & 1 & -1 & 1 \end{array}\right) \xrightarrow{M_3(-1)} \\
&\left(\begin{array}{ccc|ccc} 1 & 1 & 2 & 0 & 0 & 1 \\ 0 & 1 & -3 & 0 & 1 & -1 \\ 0 & 0 & 1 & -1 & 1 & -1 \end{array}\right) \xrightarrow{A_{23}(3)} \left(\begin{array}{ccc|ccc} 1 & 1 & 2 & 0 & 0 & 1 \\ 0 & 1 & 0 & -3 & 4 & -4 \\ 0 & 0 & 1 & -1 & 1 & -1 \end{array}\right) \xrightarrow{A_{13}(-2)} \\
&\left(\begin{array}{ccc|ccc} 1 & 1 & 0 & 2 & -2 & 3 \\ 0 & 1 & 0 & -3 & 4 & -4 \\ 0 & 0 & 1 & -1 & 1 & -1 \end{array}\right) \xrightarrow{A_{12}(-1)} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 5 & -6 & 7 \\ 0 & 1 & 0 & -3 & 4 & -4 \\ 0 & 0 & 1 & -1 & 1 & -1 \end{array}\right) = (\mathbb{1}_3 | A^{-1})
\end{aligned}
$$

und damit

$$
A^{-1} = \begin{pmatrix} 5 & -6 & 7 \\ -3 & 4 & -4 \\ -1 & 1 & -1 \end{pmatrix}.
$$

Der Gauß-Algorithmus für das Invertieren von Matrizen lässt sich auch algebraisch interpretieren. Der Gauß-Algorithmus und der Beweis von Satz 4.5.16 zeigen nämlich, dass sich jede invertierbare Matrix als Produkt von Elementarmatrizen schreiben lässt. Man sagt auch, dass die Elementarmatrizen die Gruppe $\mathrm{GL}(n, \mathbb{K})$ *erzeugen*.

Korollar 4.5.18:
Jede invertierbare Matrix $A \in \mathrm{GL}(n, \mathbb{K})$ ist ein Produkt der Elementarmatrizen $M_j(\lambda)$, V_{jk} und $A_{jk}(\lambda)$ aus Definition 4.5.10.

An was man sich erinnern sollte:

Die wichtigsten Begriffe und Definitionen:

- $\mathbb{K}$-lineare Abbildung, Mono-/Epi/-Iso/Endo-/Automorphismus von $\mathbb{K}$-Vektorräumen,
- Bild und Kern von linearer Abbildung,
- Rang und Defekt von linearer Abbildung,
- geordnete Basis,
- darstellender Spaltenvektor für Vektoren in endlich-dimensionalen $\mathbb{K}$-Vektorräumen,
- darstellende Matrix einer $\mathbb{K}$-linearen Abbildung zwischen endl.-dim. $\mathbb{K}$-Vektorräumen,
- Matrixaddition, Multiplikation, Multiplikation mit Skalaren,
- Invertierbarkeit von Matrizen, die Gruppe $\mathrm{GL}(n, \mathbb{K})$,
- Rang, Kern und Bild einer Matrix, Zeilenrang, Spaltenrang
- Basiswechselmatrizen, Ähnlichkeit und Äquivalenz von Matrizen,
- Projektionsabbildungen
- Transposition von Matrizen, symmetrische und antisymmetrische Matrix
- Dualraum, duale Basis, duale Abbildung, kanonische Abbildung
- (in)homogene lineare Gleichungssysteme, Lösungsraum, affiner Unterraum

Die wichtigsten Aussagen:

- (Ur)Bilder von Untervektorräumen unter $\mathbb{K}$-linearen Abbildungen sind Untervektorräume,
- die $\mathbb{K}$-linearen Abbildungen $\phi : V \to W$ bilden einen Vektorraum $\mathrm{Hom}_{\mathbb{K}}(V, W)$,
- die $\mathbb{K}$-linearen Abbildungen $\phi : V \to V$ bilden einen Ring $\mathrm{End}_{\mathbb{K}}(V)$ und eine $\mathbb{K}$-Algebra,
- Injektivitäts- und Surjektivitätskriterien für lineare Abbildungen,
- Dimensionsformel,
- jeder n-dimensionale $\mathbb{K}$-Vektorraum ist isomorph zum $\mathbb{K}^n$,
- Existenz und Eindeutigkeit von $\mathbb{K}$-lineare Abbildungen mit vorgegebenen Werten auf Basen, linear unabhängigen Mengen und Erzeugendensystemen.
- Charakterisierung von linearen Abbildungen durch beschreibende Matrizen, Basiswechsel.
- $\mathrm{Mat}(m \times n, \mathbb{K}) \cong \mathrm{Hom}_{\mathbb{K}}(\mathbb{K}^n, \mathbb{K}^m)$ ist ein $\mathbb{K}$-Vektorraum der Dimension mn,
- $\mathrm{Mat}(n \times n, \mathbb{K})$ ist ein Ring und eine $\mathbb{K}$-Algebra,
- Äquivalenz und Ähnlichkeit von Matrizen sind Äquivalenzrelationen,
- $M, N \in \mathrm{Mat}(m \times n, \mathbb{K})$ äquivalent $\Leftrightarrow \mathrm{rg}(M) = \mathrm{rg}(N) \Leftrightarrow M, N$ beschreiben die gleiche lineare Abbildung bzgl. zweier Basenpaare.
- $M, N \in \mathrm{Mat}(n \times n, \mathbb{K})$ ähnlich $\Leftrightarrow M, N$ beschreiben die gleiche gleiche lineare Abbildung bzgl. zweier Basen.
- $M \in \mathrm{Mat}(m \times n, \mathbb{K})$ äquivalent zu einer Matrix $E^r_{m\times n} \Leftrightarrow \mathrm{rg}(M) = r$.
- $M \in \mathrm{Mat}(n \times n, \mathbb{K})$ ähnlich zu einer Matrix $E^r_{n\times n} \Leftrightarrow (\mathrm{rg}(M) = r) \wedge (M \cdot M = M)$, $\Leftrightarrow$ $(\mathrm{rg}(M) = r) \wedge M$ beschreibende Matrix einer Projektionsabbildung bzgl. einer Basis,
- Für jede Projektionsabbildung $\phi : V \to V$ gilt $V = \ker(\phi) \oplus \mathrm{im}(\phi)$.
- Rechenregeln für die Transposition von Matrizen
- beschreibende Matrizen der dualen Abbildung
- $\mathrm{rg}(A) = \mathrm{zrg}(A) = \mathrm{srg}(A)$, $\mathrm{rg}(A) = \mathrm{rg}(A)^T$
- Gauß-Algorithmus zum Lösen von Gleichungssystemen und Invertieren von Matrizen.

Die wichtigsten Beispiele:

- **Lineare Abbildungen:** Nullabbildung, Identitätsabbildung, Drehung, Spiegelung, Streckung, Koordinatenprojektionen, (formale) Ableitung von Polynomen, $\mathbb{K}$-lineare Abbildungen mit vorgegebenen Werten auf einer Basis, Projektionsabbildungen.
- **Matrizen:** Einheitsmatrix, Elementarmatrizen, Matrizen $E^r_{m\times n}$.

5 Multilineare Abbildungen und Determinanten

In diesem Kapitel beschäftigen wir uns mit *n-linearen Abbildungen*. Dabei handelt es sich um Abbildungen $\phi : V \times \ldots \times V \to W$, $(v_1, ..., v_n) \mapsto \phi(v_1, ..., v_n)$, die n-Tupel von Vektoren aus einem $\mathbb{K}$-Vektorraum V auf Elemente eines $\mathbb{K}$-Vektorraums W abbilden und in jedem ihrer Argumente $\mathbb{K}$-linear sind. Besonders wichtige Beispiele von n-linearen Abbildungen sind die *n-Linearformen* für $W = \mathbb{K}$, also n-lineare Abbildungen $\phi : V \times \ldots \times V \to \mathbb{K}$.

Bei der Betrachtung von n-linearen Abbildungen ergibt sich automatisch die Frage, wie sie sich unter Vertauschungen oder, allgemeiner, Permutationen ihrer Argumente verhalten. Besonders einfache und wichtige Beispiele sind n-lineare Abbildungen, die *symmetrisch* oder *antisymmetrisch* unter Vertauschung ihrer Argumente sind.

Symmetrische n-Linearformen spielen bei der Beschreibung von Abständen und Winkeln im $\mathbb{R}^n$ eine besondere Rolle, bei Kegelschnitten wie Ellipsen, Parabeln, Hyperbeln und deren höherdimensionale Gegenstücke wie Ellipsoide, Paraboloide und Hyperboloide. Antisymmetrische n-Linearformen treten bei der Berechnung von Volumina im $\mathbb{R}^n$ auf.

Ist $\phi : \mathbb{K}^n \times \ldots \times \mathbb{K}^n \to \mathbb{K}$ eine antisymmetrische n-Linearform und fasst man n-Tupel von Vektoren im $\mathbb{K}^n$ als $(n \times n)$-Matrizen auf, so erhält man daraus eine Abbildung $\phi : \mathrm{Mat}(n \times n, \mathbb{K}) \to \mathbb{K}$, die antisymmetrisch unter dem Vertauschen von Spalten ist. Diese verschwindet auf Matrizen, deren Zeilen oder Spalten linear abhängig sind. Fordert man zusätzlich, dass eine solche Abbildung auf der Einheitsmatrix den Wert 1 annimmt, so ist sie eindeutig bestimmt, und man erhält die sogenannte *Determinante*. Sie beschreibt, wie sich das Volumen eines n-dimensionalen Parallelotops im $\mathbb{K}^n$ unter einer linearen Abbildung ändert.

5.1 Permutationen

Da das Verhalten von n-linearen Abbildungen unter Permutationen ihrer Argumente eine zentrale Rolle spielen wird, beschäftigen wir uns zunächst genauer mit den Eigenschaften von Permutationen. Sie wurden in Beispiel 2.2.11 und Definition 2.2.12 als bijektive Abbildungen $\sigma : \{1, ..., n\} \to \{1, ..., n\}$ für $n \in \mathbb{N}$ eingeführt. Wie in Abschnitt 2.2 beschreiben wir Permutationen durch Wertetabellen und Diagramme.

In in Abschnitt 2.2 wurde bereits gezeigt, dass die Permutationen für festes $n \in \mathbb{N}$ mit der Verkettung eine Gruppe bilden, die symmetrische Gruppe S_n. Diese enthält genau $n!$ Elemente. Es gilt also $S_1 = \{\mathrm{id}\}$. Die Permutationsgruppe S_2 enthält genau zwei Elemente, die Identitätsabbildung und eine Vertauschung, und ist isomorph zu $(\mathbb{Z}/2\mathbb{Z}, +)$. Für $n \geq 3$ ist die Permutationsgruppe S_n dagegen keine abelsche Gruppe.

Um die Struktur der symmetrischen Gruppe S_n besser zu verstehen, befassen wir uns mit ihren einfachsten Elementen. Neben der Identitätsabbildung sind die einfachsten Permutationen die *elementaren Vertauschungen*, die zwei benachbarte Zahlen $i, j \in \{1, ..., n\}$ vertauschen und alle anderen Zahlen $k \in \{1, ..., n\}$ auf sich selbst abbilden.

Ebenso interessieren wir uns dafür, wie Permutationen die Ordnung der Zahlen $1 < 2 < ... < n$ ändern. Ein Maß für die durch eine Permutation erzeugte *Unordnung* ist die Anzahl der Paare (i, j) mit $1 \leq i < j \leq n$ und $\sigma(j) > \sigma(i)$, der *Fehlstandspaare*.

Definition 5.1.1: Sei $n \in \mathbb{N}$.

1. Eine **Vertauschung** oder **Transposition** ist eine Permutation $\sigma_{ij} \in S_n$ mit $\sigma_{ij}(i) = j$, $\sigma_{ij}(j) = i$ und $\sigma_{ij}(k) = k$ für alle $k \in \{1, ..., n\} \setminus \{i, j\}$ für vorgegebene $1 \leq i < j \leq n$.

2. Eine Vertauschung σ_{ij} mit $j = i+1$ heißt **elementare Vertauschung** oder **elementare Transposition**.

3. Ein **Fehlstandspaar** einer Permutation $\sigma \in S_n$ ist ein Paar (i, j) mit $1 \leq i < j \leq n$ und $\sigma(i) > \sigma(j)$. Die Anzahl $F(\sigma)$ der Fehlstandspaare von σ heißt **Länge** von σ.

4. Das **Signum** einer Permutation $\sigma \in S_n$ ist $\mathrm{sgn}(\sigma) = (-1)^{F(\sigma)}$. Gilt $\mathrm{sgn}(\sigma) = 1$, so nennt man σ **gerade**, gilt $\mathrm{sgn}(\sigma) = -1$ so nennt man σ **ungerade**.

Beispiel 5.1.2:

1. Die Identitätsabbildung $\mathrm{id} \in S_n$ hat keine Fehlstandspaare und ist somit gerade.
2. Eine Vertauschung $\sigma_{ij} \in S_n$ mit $i < j$ wird beschrieben durch die Tabelle

$$\sigma_{ij} = \begin{pmatrix} 1 \ldots i-1 & i & i+1 \ldots j-1 & j & j+1 \ldots n \\ 1 \ldots i-1 & j & i+1 \ldots j-1 & i & j+1 \ldots n \end{pmatrix}$$

   Es gilt $\sigma_{ij} \circ \sigma_{ij} = \mathrm{id}$, und σ_{ij} hat genau $2(j-i)-1$ Fehlstandspaare, nämlich

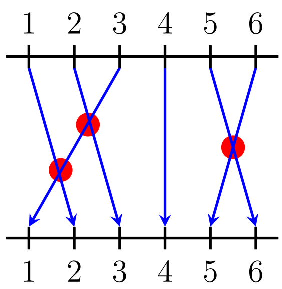
$$(i, i+1), (i, i+2), \ldots, (i, j-1), (i, j), (i+1, j), (i+2, j), \ldots, (j-1, j).$$

   Also gilt $\mathrm{sgn}(\sigma_{ij}) = -1$ für alle Vertauschungen $\sigma_{ij} \in S_n$.
3. Die Permutation

$$\sigma = \begin{pmatrix} 1 & 2 & \ldots & n-1 & n \\ 2 & 3 & \ldots & n & 1 \end{pmatrix} \in S_n$$

   hat genau $n-1$ Fehlstandspaare, nämlich $(1, n)$, $(2, n)$,...,$(n-1, n)$. Also ist σ gerade genau dann, wenn n ungerade ist. Es gilt $\sigma^n = \mathrm{id}$. Die Permutationen $\sigma, \sigma^2, ..., \sigma^{n-1}, \sigma^n$ heißen **zyklische Vertauschungen**.

Im Diagramm einer Permutation $\sigma \in S_n$ entspricht jedes Fehlstandspaar einem Kreuzungspunkt zweier Pfeile. Denn für $x, y \in \{1, ..., n\}$ mit $x < y$ kreuzen sich die Pfeile von x nach $\sigma(x)$ und von y nach $\sigma(y)$ genau dann, wenn $\sigma(x) > \sigma(y)$. Also ist die Anzahl der Kreuzungspunkte gerade die Länge einer Permutation σ. Dabei muss man allerdings in manchen Fällen durch ein leichtes Verschieben der Punkte auf der oberen oder unteren Linie sicherstellen, dass sich in jedem Kreuzungspunkt des *nur zwei Pfeile kreuzen*. Das zugehörige Fehlstandspaar findet man, indem man den sich kreuzenden Pfeilen bis zur oberen Linie folgt.



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

Die Fehlstandspaare der Permutation σ sind $(1, 3)$, $(2, 3)$, $(5, 6)$.

Die elementaren Vertauschungen spielen in der Permutationsgruppe eine besondere Rolle, weil sich jedes andere Element als Produkt elementarer Vertauschungen schreiben lässt. Man sagt, sie *erzeugen* die Permutationsgruppe S_n. Anders als die Darstellung eines Vektors als Linearkombination von Basisvektoren ist aber die Darstellung von Permutationen als Produkt elementarer Vertauschungen oft kompliziert und im allgemeinen auch nicht eindeutig.

Satz 5.1.3: Jede Permutation $\sigma \in S_n$ ist ein Produkt von $F(\sigma)$ elementaren Vertauschungen.

Beweis:
Induktion über die Länge $F(\sigma)$.

Induktionsanfang: Für $F(\sigma) = 0$ gilt $\sigma = \mathrm{id}$. Die Identitätsabbildung ist ein leeres Produkt von elementaren Vertauschungen.

Induktionsschritt: Sei bereits gezeigt, dass die Aussage für alle Permutationen maximal k Fehlstandspaaren gilt, und sei $\sigma \in S_n$ eine Permutation mit $F(\sigma) = k + 1$. Dann gibt es ein $i \in \{1, ..., n\}$ mit $\sigma(i) > \sigma(i+1)$, denn sonst wäre $\sigma(1) < \sigma(2) < \ldots < \sigma(n)$ und damit $\sigma = \mathrm{id}$. Dann ist $\sigma' = \sigma \circ \sigma_{i,i+1}$ eine Permutation mit $F(\sigma') = F(\sigma) - 1 = k$.

Denn das Fehlstandspaar $(i, i+1)$ verschwindet wegen $\sigma'(i) = \sigma(i+1) < \sigma(i) = \sigma'(i+1)$. Alle anderen Fehlstandspaare von σ stehen in Bijektion mit Fehlstandspaaren von σ':

- Ist (k, l) mit $k < l \notin \{i, i+1\}$ ein Fehlstandpaar von σ, so ist es auch ein Fehlstandspaar von σ' und umgekehrt.
- Fehlstandspaare der Form (i, j) oder (h, i) von σ mit $i + 1 < j$ und $h < i$ entsprechen Fehlstandspaaren der Form $(i + 1, j)$ und $(h, i + 1)$ von σ' und umgekehrt.

Nach Induktionsvoraussetzung ist σ' ein Produkt von $F(\sigma') = k$ elementaren Vertauschungen. Also ist $\sigma = \sigma' \circ \sigma_{i,i+1}$ ein Produkt von $F(\sigma) = k + 1$ elementaren Vertauschungen. □

Verkettet man eine Permutation $\tau \in S_n$ mit einer elementaren Vertauschung $\sigma \in S_n$, so ändert sich die Anzahl der Fehlstandspaare um eins, da σ entweder ein neues Fehlstandspaar erzeugt oder ein Fehlstandspaar vernichtet. Daraus folgt $\mathrm{sgn}(\sigma \circ \tau) = -\mathrm{sgn}(\tau)$. Also muss für eine Permutation τ, die sich als Verkettung von n elementaren Vertauschungen schreiben lässt, $\mathrm{sgn}(\tau) = (-1)^n$ gelten.

Da die Darstellung einer Permutation als Verkettung elementarer Vertauschungen aber im Allgemeinen nicht leicht zu finden ist, ist dies keine sehr nützliche Methode, um das Signum einer Permutation zu berechnen. Eine brauchbare Formel für das Signum liefert der folgende Satz.

Satz 5.1.4:

1. Für beliebige Permutationen $\sigma, \tau \in S_n$ gilt: $\mathrm{sgn}(\sigma \circ \tau) = \mathrm{sgn}(\sigma) \cdot \mathrm{sgn}(\tau)$.
2. Ist $\sigma \in S_n$ ein Produkt von $k \in \mathbb{N}_0$ Vertauschungen, so gilt $\mathrm{sgn}(\sigma) = (-1)^k$.
3. Für jede Permutation $\sigma \in S_n$ gilt:

$$\mathrm{sgn}(\sigma) = \prod_{1 \leq i < j \leq n} \frac{\sigma(j) - \sigma(i)}{j - i} = \frac{\Pi_{1 \leq i < j \leq n}\, \sigma(j) - \sigma(i)}{\Pi_{1 \leq i < j \leq n}\, j - i}. \tag{12}$$

Beweis:
1. Ist $\sigma \in S_n$ ein Produkt von k *elementaren* Vertauschungen, so gilt $\mathrm{sgn}(\sigma) = (-1)^k$, denn

Verketten mit einer elementaren Vertauschung ändert die Zahl der Fehlstandspaare einer Permutation um eins. Ist $\tau \in S_n$ ein Produkt von l elementaren Vertauschungen, so ist $\sigma \circ \tau$ ein Produkt von $k + l$ elementaren Vertauschungen und somit gilt

$$\mathrm{sgn}(\sigma \circ \tau) = (-1)^{k+l} = (-1)^k \cdot (-1)^l = \mathrm{sgn}(\sigma) \cdot \mathrm{sgn}(\tau).$$

2. Nach Beispiel 5.1.2 gilt $\mathrm{sgn}(\sigma_{ij}) = -1$ für jede Vertauschung $\sigma_{ij} \in S_n$. Ist $\sigma \in S_n$ ein Produkt von k Vertauschungen, so ergibt sich daraus mit 1. $\mathrm{sgn}(\sigma) = (-1)^k$.

3. Da $\sigma \in S_n$ eine bijektive Abbildung $\sigma : \{1, ..., n\} \to \{1, ..., n\}$ ist, tritt im Zähler des letzten Terms in (12) jeder Faktor $j - i$ mit $1 \leq i < j \leq n$ genau einmal auf, entweder mit negativem oder mit positivem Vorzeichen. Die Terme $j - i$ im Nenner sind natürliche Zahlen, der Term $\sigma(j) - \sigma(i)$ ist in $\mathbb{N}$, wenn (i, j) kein Fehlstandspaar von σ ist, und in $-\mathbb{N}$, wenn (i, j) ein Fehlstandspaar von σ ist. Daraus ergibt sich

$$\frac{\Pi_{1 \leq i < j \leq n}\, \sigma(j) - \sigma(i)}{\Pi_{1 \leq i < j \leq n}\, j - i} = (-1)^{F(\sigma)} \frac{\Pi_{1 \leq i < j \leq n}\, j - i}{\Pi_{1 \leq i < j \leq n}\, j - i} = (-1)^{F(\sigma)} = \mathrm{sgn}(\sigma).$$

□

Beispiel 5.1.5: Für die Permutationen

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 3 & 2 \end{pmatrix} \qquad \tau = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 1 & 4 \end{pmatrix}$$

erhält man

$$\mathrm{sgn}(\sigma) = \frac{(2-4)(2-1)(2-3)(3-4)(3-1)(1-4)}{(4-1)(4-2)(4-3)(3-1)(3-2)(2-1)} = 1$$

$$\mathrm{sgn}(\tau) = \frac{(4-3)(4-2)(4-1)(1-3)(1-2)(2-3)}{(4-1)(4-2)(4-3)(3-1)(3-2)(2-1)} = -1.$$

Neben einer Rechenmethode zur Bestimmung des Signums liefert Satz 5.1.4 auch eine algebraische Charakterisierung des Signums. Die erste Aussage in Satz 5.1.4 besagt nämlich, dass $\mathrm{sgn} : S_n \to (\{1, -1\}, \cdot)$, $\sigma \mapsto \mathrm{sgn}(\sigma)$ ein Gruppenhomomorphismus ist.

Die Gruppe $C_2 = (\{1, -1\}, \cdot)$ ist gerade die Gruppe der zweiten Einheitswurzeln. Nach Satz 2.3.25 ist die Abbildung $\phi : (C_2, \cdot) \to (\mathbb{Z}/2\mathbb{Z}, +)$, $-1 \mapsto [1]$, $1 \mapsto [0]$ ein Gruppenisomorphismus. Verkettet man ihn mit dem Gruppenhomomorphismus $\mathrm{sgn} : S_n \to C_2$, so erhält man einen Gruppenhomomorphismus $\mathrm{sgn}' = \phi \circ \mathrm{sgn} : S_n \to \mathbb{Z}/2\mathbb{Z}$, $\sigma \mapsto [F(\sigma)]$.

Der Kern $\ker(\mathrm{sgn}) = \ker(\mathrm{sgn}')$ ist die Menge aller geraden Permutationen und nach Satz 2.2.21 eine Untergruppe von S_n, die *alternierende Gruppe* A_n. Man kann zeigen (Übung), dass die alternierenden Gruppen A_1, A_2 und A_3 abelsch sind, während die alternierenden Gruppen A_n für $n \geq 4$ nicht-abelsch sind.

Korollar 5.1.6:

1. Für jedes $n \in \mathbb{N}$ induziert die Signumsabbildung einen Gruppenhomomorphismus $\mathrm{sgn}' : S_n \to \mathbb{Z}/2\mathbb{Z}$, $\sigma \mapsto [F(\sigma)]$.
2. Für $n \geq 2$ ist $A_n = \ker(\mathrm{sgn}) = \ker(\mathrm{sgn}') = \{\sigma \in S_n \mid \mathrm{sgn}(\sigma) = 1\} \subseteq S_n$ eine Untergruppe mit $\frac{1}{2}n!$ Elementen. Sie heißt die **alternierende Gruppe** vom Grad n.

Beweis:
1. Die Abbildung sgn' ist ein Gruppenhomomorphismus als Verkettung $\mathrm{sgn}' = \phi \circ \mathrm{sgn}$ der Gruppenhomomorphismen $\mathrm{sgn} : S_n \to C_2$ und $\phi : C_2 \to \mathbb{Z}/2\mathbb{Z}$.

2. Die Menge $A_n = \ker(\mathrm{sgn})$ ist als Kern des Gruppenhomomorphismus $\mathrm{sgn} : S_n \to C_2$ eine Untergruppe von S_n nach Satz 2.2.21. Um die Anzahl der Elemente von A_n zu bestimmen, betrachten wir für $n \geq 2$ eine Vertauschung $\sigma \in S_n$ und die bijektive Abbildung $f : S_n \to S_n$, $\tau \mapsto \tau \circ \sigma$ mit $f \circ f = \mathrm{id}$. Da nach Satz 5.1.4, 2. $\mathrm{sgn}(f(\tau)) = \mathrm{sgn}(\tau \circ \sigma) = \mathrm{sgn}(\tau) \cdot \mathrm{sgn}(\sigma) = -\mathrm{sgn}(\tau)$ für alle $\tau \in S_n$ gilt, ist $f(\ker(\mathrm{sgn})) = S_n \setminus \ker(\mathrm{sgn})$ und $f(S_n \setminus \ker(\mathrm{sgn})) = \ker(\mathrm{sgn})$. Also müssen $\ker(\mathrm{sgn})$ und $S_n \setminus \ker(\mathrm{sgn})$ gleich viele und damit jeweils $\frac{1}{2}n!$ Elemente enthalten. □

5.2 Multilineare Abbildungen

Nachdem die grundlegenden Eigenschaften der Permutationsgruppen untersucht wurden, beschäftigen wir uns nun mit den Eigenschaften *multilinearer Abbildungen.* Im Gegensatz zu einer linearen Abbildungen $\phi : V \to W$, die jedem Vektor $v \in V$ einen Vektor $\phi(v) \in W$ zuordnet, ordnet eine n-lineare Abbildung $\phi : V^{\times n} \to W$ einem n-Tupel $(v_1, ..., v_n)$ von Vektoren $v_i \in V$ einen Vektor in $\phi(v_1, ..., v_n) \in W$ zu, und zwar so, dass diese Zuordnung in jedem ihrer n Argumente linear ist. Solche Abbildungen werden später bei der Beschreibung von Flächeninhalten, Volumina, Winkeln und Abständen im $\mathbb{R}^n$ eine wichtige Rolle spielen.

Definition 5.2.1: Seien V, W Vektorräume über $\mathbb{K}$ und $n \in \mathbb{N}$. Eine Abbildung $\phi : V^{\times n} \to W$ heißt **n-linear** falls für alle $j \in \{1, ..., n\}$, $v_1, ..., v_n, v_1', ..., v_n' \in V$ und $\lambda \in \mathbb{K}$ gilt

$$\phi(..., v_j + v_j', ...) = \phi(..., v_j, ...) + \phi(..., v_j', ...) \qquad \phi(..., \lambda\, v_j, ...) = \lambda\, \phi(..., v_j, ...).$$

Für $n = 2$ spricht man auch von einer **bilinearen** Abbildung und für $n \geq 2$ von einer **multilinearen** Abbildung. Ist $W = \mathbb{K}$, so nennt man eine n-lineare Abbildung auch **n-Linearform**, eine bilineare Abbildung **Bilinearform** und eine multilineare Abbildung **Multilinearform**.

Stehen Summen oder skalare Vielfache in den Argumenten einer multilinearen Abbildung, so kann man diese also schrittweise auflösen. So erhält man beispielsweise für eine bilineare Abbildung $\phi : V^{\times 2} \to W$ für $u, u'v, v' \in V$ und $\mu, \mu', \lambda, \lambda' \in \mathbb{K}$

$$\begin{aligned}
\phi(\mu u + \mu' u', \lambda v + \lambda' v') =& \phi(\mu u, \lambda v + \lambda' v') + \phi(\mu' u', \lambda v + \lambda' v') \\
=& \phi(\mu u, \lambda v) + \phi(\mu u, \lambda' v') + \phi(\mu' u', \lambda v) + \phi(\mu' u', \lambda' v') \\
=& \mu \phi(u, \lambda v) + \mu \phi(u, \lambda' v') + \mu' \phi(u', \lambda v) + \mu' \phi(u', \lambda' v') \\
=& \mu \lambda \phi(u, v) + \mu \lambda' \phi(u, v') + \mu' \lambda \phi(u', v) + \mu' \lambda' \phi(u', v').
\end{aligned}$$

Bei der Betrachtung von multilinearen Abbildungen stellt sich sofort die Frage, wie sich ihr Wert ändert, wenn man Argumente vertauscht oder permutiert. Ein besonders einfaches Verhalten zeigen die *symmetrischen* und *alternierenden* multilinearen Abbildungen.

Definition 5.2.2: Seien V, W Vektorräume über $\mathbb{K}$ und $n \in \mathbb{N}$. Eine n-lineare Abbildung $\phi : V^{\times n} \to W$ heißt

- **symmetrisch**, wenn $\phi(..., v_i, ..., v_j, ...) = \phi(..., v_j, ..., v_i, ...)$ für alle $1 \leq i < j \leq n$,
- **alternierend**, wenn aus $v_i = v_j$ für zwei Indizes $1 \leq i < j \leq n$ folgt, dass $\phi(v_1, ..., v_n) = 0$.

Beispiel 5.2.3:

1. Ist $n = 1$, so ist jede n-lineare Abbildung $\phi : V^{\times n} \to W$ symmetrisch und alternierend.
2. Für beliebiges $n \in \mathbb{N}$ ist die Nullabbildung $0 : V^{\times n} \to W$, $(v_1, ..., v_n) \mapsto 0$ eine symmetrische und alternierende n-lineare Abbildung.
3. Die Abbildung $\phi : \mathbb{K}^{\times n} \to \mathbb{K}$, $(\lambda_1, ..., \lambda_n) \mapsto \lambda_1 \cdots \lambda_n$ ist eine symmetrische n-lineare Abbildung, denn es gilt das Distributivgesetz, und die Multiplikation im Körper $\mathbb{K}$ ist kommutativ.
4. Für beliebige $n, m \in \mathbb{N}$ ist die Abbildung $\phi : \mathrm{Mat}(m \times m, \mathbb{K})^{\times n} \to \mathrm{Mat}(m \times m, \mathbb{K})$, $(A_1, ..., A_n) \mapsto A_1 \cdots A_n$ eine n-lineare Abbildung, denn die Matrixmultiplikation erfüllt das Distributivgesetz und ist mit der Skalarmultiplikation kompatibel. Für $m, n \geq 2$ ist diese weder symmetrisch noch alternierend.
5. Der **Matrixkommutator** ist eine alternierende bilineare Abbildung
$$[\ ,\ ] : \mathrm{Mat}(m \times m, \mathbb{K}) \times \mathrm{Mat}(m \times m, \mathbb{K}) \to \mathrm{Mat}(m \times m, \mathbb{K}), \quad (A, B) \mapsto [A, B] := A \cdot B - B \cdot A$$
6. Für jedes $n \in \mathbb{N}$ und jeden Körper $\mathbb{K}$ ist
$$\langle\ ,\ \rangle : \mathbb{K}^n \times \mathbb{K}^n \to \mathbb{K}, \quad (u, v) \mapsto \langle u, v \rangle := \Sigma_{i=1}^n u_i v_i$$
eine symmetrische Bilinearform. Für $\mathbb{K} = \mathbb{R}$ heißt sie **euklidisches Skalarprodukt**.
7. Das **Vektorprodukt**
$$\wedge : \mathbb{K}^3 \times \mathbb{K}^3 \to \mathbb{K}^3, \ (v, w) \mapsto v \wedge w \qquad \text{mit} \quad \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \wedge \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} = \begin{pmatrix} v_2 w_3 - w_2 v_3 \\ v_3 w_1 - v_1 w_3 \\ v_1 w_2 - w_1 v_2 \end{pmatrix}$$
ist eine alternierende bilineare Abbildung.
8. Das **Spatprodukt**
$$\phi : \mathbb{K}^3 \times \mathbb{K}^3 \times \mathbb{K}^3 \to \mathbb{K}, \quad (u, v, w) \mapsto \langle u, v \wedge w \rangle$$
ist eine alternierende Multilinearform.

Betrachtet man Definition 5.2.2 so stellt sich die Frage, warum nicht einfach gefordert wurde, dass eine alternierende multilineare Abbildung analog zu einer symmetrischen die Bedingung $\phi(..., v_i, ..., v_j, ...) = -\phi(..., v_j, ..., v_i, ...)$ erfüllt und ob diese Bedingung zu der aus Definition 5.2.2 äquivalent ist. Letzteres ist der Fall, solange für den zugrundeliegenden Körper $\mathbb{K}$ gilt $\mathrm{char}(\mathbb{K}) \neq 2$, also $1 + 1 \neq 0$. Ist aber $\mathrm{char}(\mathbb{K}) = 2$, so gilt $\phi(..., v_i, ..., v_j, ...) = -\phi(..., v_j, ..., v_i, ...)$ oder, dazu äquivalent, $\phi(v_1, ..., v_n) + \phi(v_1, ..., v_n) = 0$ für jede *symmetrische* n-lineare Abbildung. Die Bedingung mit dem Minuszeichen ist also schwächer als die Bedingung in Definition 5.2.2.

Lemma 5.2.4: Sei $\mathbb{K}$ ein Körper der Charakteristik $\mathrm{char}(\mathbb{K}) \neq 2$ und $\phi : V^{\times n} \to W$ eine n-lineare Abbildung. Dann sind äquivalent:

(i) Die Abbildung ϕ ist alternierend.

(ii) $\phi(..., v_i, ..., v_j, ...) = -\phi(..., v_j, ..., v_i, ...)$ für alle $1 \leq i < j \leq n$.

(iii) $\phi(..., v_i, ..., v_j + \lambda v_i, ...) = \phi(..., v_i + \lambda v_j, ..., v_j, ...) = \phi(..., v_i, ..., v_j, ...)$ $\forall \lambda \in \mathbb{K}, 1 \leq i < j \leq n$.

Ist $\mathrm{char}(\mathbb{K}) = 2$, so gilt (i)$\Leftrightarrow$(iii) und (i)$\Rightarrow$(ii).

Beweis:
(i) $\Rightarrow$ (ii): Sei ϕ alternierend und $1 \leq i < j \leq n$. Dann folgt aus der Multilinearität von ϕ

$$\begin{aligned} 0 &\overset{(i)}{=} \phi(..., v_i + v_j, ..., v_i + v_j, ...) = \phi(..., v_i, .., v_i + v_j, ...) + \phi(..., v_j, ..., v_i + v_j, ...) \\ &= \phi(..., v_i, ..., v_j, ...) + \phi(..., v_i, ..., v_i, ...) + \phi(..., v_j, ..., v_j, ...) + \phi(..., v_i, ..., v_i, ...) \\ &\overset{(i)}{=} \phi(v_1, ..., v_i, ..., v_j, ..., v_n) + \phi(v_1, ..., v_j, ..., v_i, ..., v_n). \end{aligned}$$

(ii)$\Rightarrow$(i): Ist $\mathrm{char}(\mathbb{K}) \neq 2$, $\phi(..., v_k, ..., v_l, ...) = -\phi(..., v_l, ...v_k, ...)$ für alle $1 \leq k < l \leq n$, so folgt

$$0 = \phi(..., v_i, ..., v_i, ...) + \phi(..., v_i, ..., v_i, ...) = 2\phi(..., v_i, ..., v_i, ...) \quad \Rightarrow \quad \phi(..., v_i, ..., v_i, ...) = 0.$$

(i) $\Rightarrow$ (iii): Wegen der n-Linearität von ϕ gilt für alle mit $1 \leq i < j \leq n$ und $\lambda \in \mathbb{K}$

$$\begin{aligned} &\phi(..., v_i, ..., v_j + \lambda v_i, ...) = \phi(..., v_i, ..., v_j, ...) + \lambda\phi(..., v_i, ..., v_i, ...) \\ &\overset{(i)}{=} \phi(..., v_i, ..., v_j, ...) \overset{(i)}{=} \phi(..., v_i, ..., v_j, ...) + \lambda\phi(..., v_j, ..., v_j, ...) = \phi(..., v_i + \lambda v_j, ..., v_j, ...). \end{aligned}$$

(iii) $\Rightarrow$(i): Ist $v_i = v_j$ für zwei Indizes i, j mit mit $1 \leq i < j \leq n$, so folgt aus (iii)

$$\phi(..., v_i, ..., v_i, ...) = \phi(..., v_i, ..., 0 + 1v_i, ...) \overset{(iii)}{=} \phi(..., v_i, ..., 0, ...) = 0 \cdot \phi(..., v_i, ..., 0, ...) = 0.$$

□

Mit Hilfe von Lemma 5.2.4 lässt sich nun leicht klären, wie sich eine symmetrische oder alternierende n-lineare Abbildung unter beliebigen Permutationen ihrer Argumente verhält. Dazu nutzt man aus, dass sich jede Permutation $\sigma \in S_n$ als Produkt von Vertauschungen schreiben lässt und betrachtet die durch Anwendung von Vertauschungen erzeugten Vorzeichenwechsel.

Lemma 5.2.5: Sei $\mathbb{K}$ ein Körper, V, W Vektorräume über $\mathbb{K}$ und $\phi : V^{\times n} \to W$ eine n-lineare Abbildung. Dann gilt:

1. ϕ symmetrisch $\Rightarrow \phi(v_{\sigma(v_1)}, ..., v_{\sigma(n)}) = \phi(v_1, ..., v_n)$ für alle $\sigma \in S_n$ und $v_1, ..., v_n \in V$.
2. ϕ alternierend $\Rightarrow \phi(v_{\sigma(v_1)}, ..., v_{\sigma(n)}) = \mathrm{sgn}(\sigma)\, \phi(v_1, ..., v_n)$ für alle $\sigma \in S_n$ und $v_1, ..., v_n \in V$.

Beweis:
Ist $\sigma \in S_n$ eine Vertauschung, so gilt nach Definition 5.2.2 $\phi(v_{\sigma(1)}, ..., v_{\sigma(n)}) = \phi(v_1, ..., v_n)$ für symmetrische und $\phi(v_{\sigma(1)}, ..., v_{\sigma(n)}) = -\phi(v_1, ..., v_n) = \mathrm{sgn}(\sigma)\, \phi(v_1, ..., v_n)$ für alternierende ϕ nach Lemma 5.2.4 und Beispiel 5.1.2. Da sich jede Permutation $\tau \in S_n$ nach Satz 5.1.3 als Produkt von Vertauschungen schreiben lässt, folgt mit Satz 5.1.4 $\phi(v_{\tau(1)}, ..., v_{\tau(n)}) = \phi(v_1, ..., v_n)$ für symmetrische und $\phi(v_{\tau(1)}, ..., v_{\tau(n)}) = \mathrm{sgn}(\tau)\phi(v_1, ..., v_n)$ für alternierende ϕ. □

Wie untersuchen nun die algebraischen Strukturen auf der Menge der (alternierenden oder symmetrischen) n-linearen Abbildungen. Zunächst stellen wir fest, dass die Abbildungen $V \times \ldots \times V \to W$ für beliebige $\mathbb{K}$-Vektorräume V, W mit der punktweisen Addition und Skalarmultiplikation einen Vektorraum über $\mathbb{K}$ bilden. Dies gilt nach Beispiel 3.1.7 für die Abbildungen von einer beliebigen Menge M in einen beliebigen $\mathbb{K}$-Vektorraum.

Die n-linearen Abbildungen $V \times \ldots \times V \to W$ bilden einen Untervektorraum dieses Vektorraums. Da die symmetrischen und alternierenden n-linearen Abbildungen über Bedingungen definiert sind, die von der punktweisen Addition und Skalarmultiplikation Gebrauch machen, gilt dies ebenfalls für die Mengen der symmetrischen bzw. alternierenden n-linearen Abbildungen. Ebenso sieht man für $\mathrm{char}(\mathbb{K}) \neq 2$ und $n \geq 2$ leicht, dass die Nullabbildung die einzige n-lineare Abbildung ist, die gleichzeitig symmetrisch und alternierend ist.

Lemma 5.2.6: Sei $\mathbb{K}$ ein Körper und V, W Vektorräume über $\mathbb{K}$. Dann gilt für alle $n \in \mathbb{N}$:

1. Die n-linearen Abbildungen bilden einen Untervektorraum $\mathrm{L}(V^{\times n}, W) \subseteq \mathrm{Abb}(V^{\times n}, W)$.
2. Die symmetrischen n-linearen Abbildungen bilden einen Untervektorraum $S(V^{\times n}, W) \subseteq L(V^{\times n}, W)$.
3. Die alternierenden n-linearen Abbildungen bilden einen Untervektorraum $A(V^{\times n}, W) \subseteq L(V^{\times n}, W)$.
4. Ist $n \geq 2$ und $\mathrm{char}(\mathbb{K}) \neq 2$, so ist $A(V^{\times n}, W) \cap S(V^{\times n}, W) = \{0\}$.

Beweis:
1.-3. ergeben sich durch direktes Nachrechnen der Untervektorraumbedingungen (UV1)-(UV3) (Übung). Ist $n \geq 2$ und $\phi : V^{\times n} \to W$ eine symmetrische und alternierende n-lineare Abbildung, so ergibt sich für alle $1 \leq i < j \leq n$

$$\phi(v_1, ..., v_j, ..., v_i, ..., v_n) = \phi(v_1, ..., v_i, ..., v_j, ..., v_n) = -\phi(v_1, ..., v_i, ..., v_j, ..., v_n) \overset{\mathrm{char}(\mathbb{K})\neq 2}{\Rightarrow} 0. \quad \square$$

Ähnlich wie bei linearen Abbildungen kann man auch bei n-linearen Abbildungen die Linearität ausnutzen, um sie durch ihre Werte auf Basen zu charakterisieren. Der einzige Unterschied ist, dass bei einer n-linearen Abbildung $\phi : V^{\times n} \to W$ nun in jedes Argument von ϕ Basisvektoren eingesetzt werden müssen. Man muss also die Werte von ϕ auf n-Tupeln von Basisvektoren statt auf einzelnen Basisvektoren betrachten. So wie die Werte einer linearen Abbildung auf jedem Basisvektor beliebig vorgegeben werden können, können die Werte einer n-linearen Abbildung auf jedem n-Tupel von Basisvektoren beliebig vorgegeben werden.

Fordert man zusätzlich, dass die betrachteten n-linearen Abbildungen symmetrisch oder alternierend sind, so sind diese immer noch durch ihre Werte auf n-Tupeln von Basisvektoren eindeutig bestimmt. Diese Werte können aber nicht mehr beliebig vorgegeben werden, da sie das richtige Verhalten unter Permutationen der Argumente aufweisen müssen. Beliebig vorgegeben werden können dann nur noch ihre Werte auf n-Tupeln von Basisvektoren mit aufsteigenden oder streng aufsteigenden Indizes.

Satz 5.2.7: Sei $\mathbb{K}$ ein Körper, V, W Vektorräume über $\mathbb{K}$ und $B = (b_1, ..., b_m)$ eine geordnete Basis von V. Dann gilt:

1. Zu gegebenen Vektoren $w_{j_1 \ldots j_n} \in W$ für $1 \leq j_1, ..., j_n \leq m$, existiert genau eine n-lineare Abbildung $\phi : V^{\times n} \to W$ mit $\phi(b_{j_1}, ..., b_{j_n}) = w_{j_1 \ldots j_n}$ für alle $1 \leq j_1, ..., j_n \leq m$.

2. Zu gegebenen Vektoren $w_{j_1...j_n} \in W$ für $1 \leq j_1 \leq j_2 \leq ... \leq j_n \leq m$ existiert genau eine symmetrische n-lineare Abbildung $\phi : V^{\times n} \to W$ mit $\phi(b_{j_1}, ..., b_{j_n}) = w_{j_1...j_n}$ für alle $1 \leq j_1 \leq ... \leq j_n \leq m$.
3. Zu gegebenen Vektoren $w_{j_1...j_n} \in W$ für $1 \leq j_1 < j_2 < ... < j_n \leq m$ existiert genau eine alternierende n-lineare Abbildung $\phi : V^{\times n} \to W$ mit $\phi(b_{j_1}, ..., b_{j_n}) = w_{j_1...j_n}$ für alle $1 \leq j_1 < ... < j_n \leq m$.

Beweis:
1. Jede n-lineare Abbildung $\phi : V^{\times n} \to W$ ist durch die Werte $\phi(b_{i_1}, ..., b_{i_n})$ mit $1 \leq i_1, ..., i_n \leq m$ eindeutig bestimmt, denn zu jedem n-Tupel $(v_1, ..., v_n)$ von Vektoren aus V existieren Koeffizienten $\lambda_{ji} \in \mathbb{K}$, so dass $v_i = \Sigma_{j=1}^m \lambda_{ji} b_j$. Aus der Multilinearität von ϕ folgt dann

$$\phi(v_1, ..., v_n) = \phi(\Sigma_{j_1=1}^m \lambda_{j_1 1} b_{j_1}, ..., \Sigma_{j_n=1}^m \lambda_{j_n n} b_{j_n}) = \Sigma_{j_1=1}^n \cdots \Sigma_{j_n=1}^n \lambda_{j_1 1} \cdots \lambda_{j_n n}\, \phi(b_{j_1}, ..., b_{j_n}). \quad (13)$$

Umgekehrt kann man zu vorgegebenen Vektoren $w_{j_1...j_n} \in W$ für $1 \leq j_1, ..., j_n \leq m$ wie im Fall der linearen Abbildungen eine n-lineare Abbildung $\phi : V^{\times n} \to W$ definieren, indem man $\phi(b_{j_1}, ..., b_{j_n}) := w_{j_1...j_n}$ für alle $1 \leq j_1, ..., j_n \leq m$ setzt und sie für ein n-Tupel von Vektoren $v_i = \Sigma_{j=1}^m \lambda_{ji} b_j$ durch (13) definiert.

2. Eine symmetrische bilineare Abbildung ist nach 1. durch ihre Werte auf n-Tupeln $(b_{j_1}, ..., b_{j_n})$ mit $1 \leq j_1, ..., j_n \leq m$ bestimmt. Da nach Lemma 5.2.5 $\phi(b_{\sigma(j_1)}, ..., b_{\sigma(j_n)}) = \phi(b_{j_1}, ..., b_{j_n})$ für alle $\sigma \in S_n$ gilt und jedes n-Tupel $(j_1, ..., j_n) \in \{1, ..., m\}^{\times n}$ durch eine Permutation $\sigma \in S_n$ in genau ein n-Tupel $(i_1, ..., i_n)$ mit $1 \leq i_1 \leq \ldots \leq i_n \leq m$ überführt werden kann, ist sie bereits durch die Werte $\phi(b_{j_1}, ..., b_{j_n})$ mit $1 \leq j_1 \leq \ldots \leq j_n \leq m$ eindeutig bestimmt, und diese können beliebig vorgegeben werden.

3. Eine alternierende bilineare Abbildung ist nach 1. durch ihre Werte auf n-Tupeln $(b_{j_1}, ..., b_{j_n})$ mit $1 \leq j_1, ..., j_n \leq m$ bestimmt. Gilt $j_i = j_k$ für zwei Indizes i, k mit $1 \leq i \neq k \leq m$ so ist $\phi(b_{j_1}, ..., b_{j_n}) = 0$. Also ist ϕ durch die Werte $\phi(b_{j_1}, ..., b_{j_n})$ mit $1 \leq j_1, ..., j_n \leq m$ und $j_1, ..., j_m$ paarweise verschieden bereits eindeutig bestimmt. Nach Lemma 5.2.5 muss außerdem $\phi(b_{\sigma(j_1)}, ..., b_{\sigma(j_n)}) = \mathrm{sgn}(\sigma)\phi(b_{j_1}, ..., b_{j_n})$ für jede Permutation $\sigma \in S_n$ gelten. Da jedes n-Tupel $(j_1, ..., j_n) \in \{1, ..., m\}^{\times n}$ mit paarweise verschiedenen $j_1, ..., j_n$ durch Permutationen $\sigma \in S_n$ in genau ein n-Tupel $(i_1, ..., i_n)$ mit $1 \leq i_1 < \ldots < i_n \leq m$ überführt werden kann, ist ϕ durch die Werte $\phi(b_{j_1}, ..., b_{j_n})$ mit $1 \leq j_1 < \ldots < j_n \leq m$ eindeutig bestimmt, und diese können beliebig vorgegeben werden. □

Beispiel 5.2.8: Für $V = \mathbb{R}^3$ ist jede bilineare Abbildung $\phi : V \times V \to W$ durch die Werte auf den Paaren (e_1, e_1), (e_2, e_2), (e_3, e_3), (e_1, e_2), (e_2, e_1), (e_1, e_3), (e_3, e_1), (e_2, e_3), (e_3, e_2) vollständig bestimmt und diese können beliebig vorgegeben werden.

Ist ϕ symmetrisch, so können nur die Werte auf den Paaren (e_1, e_1), (e_2, e_2), (e_3, e_3), (e_1, e_2), (e_1, e_3), (e_2, e_3) beliebig vorgegeben werden, da aus der Symmetrie $\phi(e_2, e_1) = \phi(e_1, e_2)$, $\phi(e_3, e_1) = \phi(e_1, e_3)$ und $\phi(e_3, e_2) = \phi(e_2, e_3)$ folgt.

Ist ϕ alternierend, so können nur die Werte auf den Paaren (e_1, e_2), (e_1, e_3), (e_2, e_3) beliebig vorgegeben werden, da aus der der Tatsache, dass ϕ alternierend ist, folgt $\phi(e_2, e_1) = -\phi(e_1, e_2)$, $\phi(e_3, e_1) = -\phi(e_1, e_3)$ und $\phi(e_3, e_2) = -\phi(e_2, e_3)$ sowie $\phi(e_1, e_1) = \phi(e_2, e_2) = \phi(e_3, e_3) = 0$.

Insbesondere erhält man aus Satz 5.2.7 Aussagen über die Dimension des Vektorraums der n-Linearformen $V \times \ldots \times V \to \mathbb{K}$ für endlich-dimensionale $\mathbb{K}$-Vektorräume V und über die Dimension der Unterräume der alternierenden und symmetrischen n-Linearformen.

Korollar 5.2.9:
Sei $\mathbb{K}$ ein Körper und V ein n-dimensionaler Vektorraum über $\mathbb{K}$. Dann gilt für alle $p \in \mathbb{N}$:

$$\begin{aligned}
\dim_{\mathbb{K}} L(V^{\times p}, \mathbb{K}) &= n^p \\
\dim_{\mathbb{K}} A(V^{\times p}, \mathbb{K}) &= \begin{pmatrix} n \\ p \end{pmatrix} = \frac{n!}{p!(n-p)!} \\
\dim_{\mathbb{K}} S(V^{\times p}, \mathbb{K}) &= \begin{pmatrix} n+p-1 \\ p \end{pmatrix} = \frac{(n+p-1)!}{p!(n-1)!}
\end{aligned}$$

Insbesondere ergibt sich für $p = 2$:

$$\dim_{\mathbb{K}} L(V^{\times 2}, \mathbb{K}) = n^2 \qquad \dim_{\mathbb{K}} A(V^{\times 2}, \mathbb{K}) = \tfrac{1}{2} n(n-1) \qquad \dim_{\mathbb{K}} S(V^{\times 2}, \mathbb{K}) = \tfrac{1}{2} n(n+1).$$

Beweis:
Wir wählen eine geordnete Basis $B = (b_1, ..., b_n)$ von V. Dann bilden nach Satz 5.2.7 die p-Linearformen $\phi_{j_1 \cdots j_p} : V^{\times p} \to \mathbb{K}$ mit $\phi_{j_1 \cdots j_p}(b_{j_1}, ..., b_{j_p}) = 1$ und $\phi_{j_1 \cdots j_p}(b_{k_1}, ..., b_{k_p}) = 0$ für $(k_1, ..., k_p) \neq (j_1, ..., j_p)$ eine Basis von $L(V^{\times p}, \mathbb{K})$. Also ist die Dimension von $L(V^{\times p}, \mathbb{K})$ gleich der Anzahl der p-Tupel $(j_1, ..., j_p)$ mit $j_1, ..., j_p \in \{1, ..., n\}$, also n^p, da man jeden Eintrag des p-Tupels $(j_1, ..., j_p)$ unabhängig aus $\{1, ..., n\}$ wählen kann.

Ebenso bilden nach Satz 5.2.7 die alternierenden p-Linearformen $\psi_{j_1 \cdots j_p}$ mit $1 \leq j_1 < ... < j_p \leq n$ und $\psi_{j_1 \ldots j_p}(b_{k_1}, ..., b_{k_p}) = \mathrm{sgn}(\sigma)$ falls $k_i = \sigma(j_i)$ für ein $\sigma \in S_p$ und $\psi_{j_1 \ldots j_p}(b_{k_1}, ..., b_{k_p}) = 0$ sonst eine Basis von $A(V^{\times p}, \mathbb{K})$. Die Dimension von $A(V^{\times p}, \mathbb{K})$ ist also gleich der Anzahl der geordneten p-Tupel $(j_1, ..., j_p) \in \{1, ..., n\}^{\times p}$ mit $1 \leq j_1 < \ldots < j_p \leq n$. Diese ist wiederum gleich der Anzahl der Möglichkeiten p verschiedene Zahlen $j_1, ..., j_p$ aus der Menge $\{1, ..., n\}$ zu wählen, also p Elemente *ohne Zurücklegen* aus einer n-elementigen Menge zu wählen. Diese ist durch den angegebenen Binomialkoeffizienten gegeben.

Analog bilden nach Satz 5.2.7 für $1 \leq j_1 \leq j_2 \leq ... \leq j_p \leq n$ die symmetrischen p-Linearformen $\chi_{j_1 \cdots j_p}$ mit $\chi_{j_1 \ldots j_p}(b_{k_1}, ..., b_{k_p}) = 1$ falls $k_i = \sigma(j_i)$ für ein $\sigma \in S_p$ und $\chi_{j_1 \ldots j_p}(b_{k_1}, ..., b_{k_p}) = 0$ sonst eine Basis von $S(V^{\times p}, \mathbb{K})$. Die Dimension von $S(V^{\times p}, \mathbb{K})$ ist also gleich der Anzahl der geordneten Tupel $(j_1, ..., j_p) \in \{1, ..., n\}^{\times p}$ mit $1 \leq j_1 \leq \ldots \leq j_p \leq n$, also die Anzahl der Möglichkeiten p Elemente aus einer n-elementigen Menge *mit Zurücklegen* auszuwählen.

Da für jedes p-Tupel $(j_1, ..., j_p) \in \{1, ..., n\}^{\times p}$ mit $1 \leq j_1 \leq \ldots \leq j_p \leq n$ das p-Tupel $(k_1, ..., k_p) = (j_1, j_2 + 1, ..., j_p + p - 1)$ ein p-Tupel mit $1 \leq k_1 < \ldots < k_p \leq n + p - 1$ ist und für jedes p-Tupel $(k_1, ..., k_p)$ mit $1 \leq k_1 < \ldots < k_p \leq n + p - 1$ das p-Tupel $(j_1, ..., j_p) = (k_1, k_1 - 1, ..., k_p - p + 1)$ ein p-Tupel mit $1 \leq j_1 \leq ... \leq j_p \leq n$, ist dies die Anzahl der p-Tupel $(k_1, ..., k_p) \in \{1, ..., n + p - 1\}^{\times p}$ mit $1 < k_1 < \ldots < k_p \leq n + p - 1$. Diese ist nach 2. die Anzahl der Möglichkeiten aus $n + p - 1$ Zahlen *ohne Zurücklegen* p Zahlen zu ziehen, und gleich dem angegebenen Binomialkoeffizienten. Für $p = 2$ ergeben sich dann entsprechend die Ausdrücke im Korollar. $\square$

Eine Konsequenz aus Korollar 5.2.9 ist, dass für $p > n$ die Nullform die einzige alternierende p-Linearform auf einen n-dimensionalen Vektorraum V ist, und alle alternierenden n-Linearformen auf V skalare Vielfache voneinander sind. Denn es gilt $\dim_{\mathbb{K}} A(V^{\times n}, \mathbb{K}) = 0$ für $p > n$ und $\dim_{\mathbb{K}} A(V^{\times n}, \mathbb{K}) = 1$. Betrachtet man alternierende n-Linearformen auf dem Vektorraum $V = \mathbb{K}^n$ und wählt man den Skalar so, dass eine n-Linearform auf der Standardbasis den Wert 1 annimmt, so erhält man die sogenannte *Volumenform*.

Korollar 5.2.10: Für jeden Körper $\mathbb{K}$ und jedes $n \in \mathbb{N}$ gibt es genau eine alternierende n-lineare Abbildung $\mathrm{vol}_n : (\mathbb{K}^n)^{\times n} \to \mathbb{K}$ mit $\mathrm{vol}_n(e_1, ..., e_n) = 1$. Sie heißt **Volumenform**.

Beweis:
Nach Satz 5.2.7 ist eine alternierende n-lineare Abbildung $\phi : (\mathbb{K}^n)^{\times n} \to \mathbb{K}$ durch ihren Wert auf $(e_1, e_2, ..., e_n)$ eindeutig bestimmt, und $\phi(e_1, ..., e_n) \in \mathbb{K}$ kann beliebig vorgegeben werden. Also gibt es genau eine alternierende n-lineare Abbildung $\mathrm{vol}_n : (\mathbb{K}^n)^{\times n} \to \mathbb{K}$ mit $\phi(e_1, ..., e_n) = 1$. □

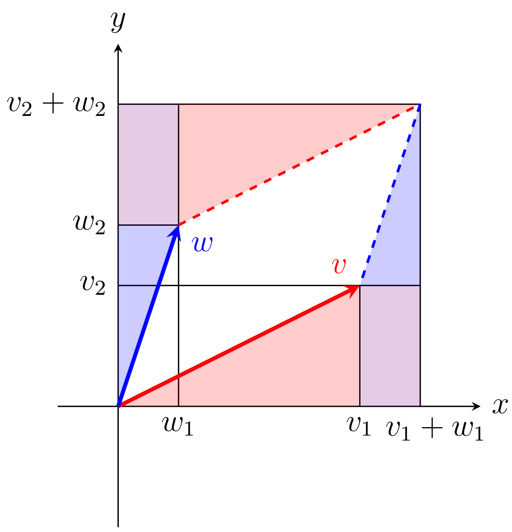
Die *Volumenform* vol_n definiert einen Volumenbegriff auf dem $\mathbb{R}^n$. Das von Vektoren $v_1, ..., v_n \in \mathbb{R}^n$ aufgespannten **Parallelotop** im $\mathbb{R}^n$ bezeichnen wir mit

$$P(v_1, ..., .v_n) = \left\{ \Sigma_{j=1}^n \lambda_j v_j \mid \lambda_1, ..., \lambda_n \in [0,1] \right\}.$$

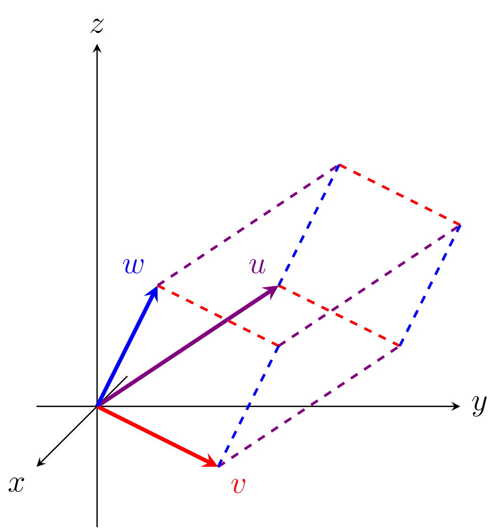
Für $n = 2$ handelt es sich dabei um ein Parallelogramm und für $n = 3$ um einen Spat.

Das Volumen eines solchen Parallelotops sollte sich linear in jedem Vektor reskalieren und sich unter Scherungen nicht ändern. Dies ergibt sich aus der Multilinearität der Volumenform. Ebenso sollte das Volumen des n-dimensionalen Einheitswürfels 1 sein und das n-dimensionale Volumen eines niedrig dimensionaleren Polytops null. Ersteres wird durch die Normierungsbedingung $\mathrm{vol}_n(e_1, ..., e_n) = 1$ und letzteres durch die Antisymmetrie der Volumenform garantiert.

Für $n = 2$ und $n = 3$ lässt sich nachrechnen, dass $|\mathrm{vol}_2(x, y)|$ und $|\mathrm{vol}_3(u, v, w)|$ den Flächeninhalt des von den Vektoren $x, y \in \mathbb{R}^2$ aufgespannten Parallelogramms und des von den Vektoren $u, v, w \in \mathbb{R}^3$ aufgespannten Spats angeben.



von $v, w \in \mathbb{R}^2$ aufgespanntes Parallelogramm



von $u, v, w \in \mathbb{R}^3$ aufgespannter Spat

Lemma 5.2.11:

1. Für beliebige Vektoren $v, w \in \mathbb{R}^2$ ist $|\mathrm{vol}_2(v, w)|$ der Flächeninhalt des von v und w aufgespannten Parallelogramms.
2. Für beliebige Vektoren $u, v, w \in \mathbb{R}^3$ ist $|\mathrm{vol}_3(u, v, w)|$ der Flächeninhalt des von u, v und w aufgespannten **Spats** oder **Parallelepipeds**.

Beweis:
1. Seien $v = v_1e_1 + v_2e_2, w = w_1e_1 + w_2e_2 \in \mathbb{R}^2$. Dann gilt wegen der Bilinearität und der Antisymmetrie der Volumenform

$$\begin{aligned}|\mathrm{vol}_2(v,w)| &= |v_1w_1\mathrm{vol}_2(e_1,e_1) + v_1w_2\mathrm{vol}_2(e_1,e_2) + v_2w_1\mathrm{vol}_2(e_2,e_1) + v_2w_2\mathrm{vol}_2(e_2,e_2)| \\ &= |v_1w_2 - v_2w_1|.\end{aligned}$$

Aus der Skizze ergibt sich der Flächeninhalt des Parallelogramms zu

$$F = |(v_1+v_2)\cdot(w_1+w_2) - 2\cdot\tfrac{1}{2}v_1v_2 - 2\cdot\tfrac{1}{2}w_1w_2 - 2\cdot w_1v_2| = |v_1w_2 - w_1v_2|.$$

2. Seien $u = u_1e_1 + u_2e_2 + u_3e_3$, $v = v_1e_1 + v_2e_2 + v_3e_3$ und $w = w_1e_1 + w_2e_2 + w_3e_3 \in \mathbb{R}^3$. Dann folgt aus der Multilinearität der Volumenform und der Tatsache, dass diese alternierend ist

$$\begin{aligned}|\mathrm{vol}_3(u,v,w)| &= |u_1v_2w_3\mathrm{vol}_3(e_1,e_2,e_3) + u_1v_3w_2\mathrm{vol}_3(e_1,e_3,e_2) + u_2v_1w_3\mathrm{vol}_3(e_2,e_1,e_3) \\ &\quad + u_2v_3w_2\mathrm{vol}_3(e_2,e_3,e_1) + u_3v_1w_2\mathrm{vol}_3(e_3,e_1,e_2) + u_3v_2w_1\mathrm{vol}_3(e_3,e_2,e_1)| \\ &= |u_1(v_2w_3 - v_3w_2) + u_2(v_3w_1 - w_3v_1) + u_3(v_1w_2 - v_2w_1)| = |\langle u, v\wedge w\rangle|.\end{aligned}$$

Dies entspricht gerade dem Spatprodukt aus Beispiel 5.2.3.

Um das Volumen des von $u, v, w \in \mathbb{R}^3$ aufgespannten Spats zu berechnen, betrachten wir zunächst den Fall $v, w \in \mathrm{span}_{\mathbb{R}}(\{e_1, e_2\})$. Dann gilt $|\mathrm{vol}_3(u,v,w)| = |u_3(v_1w_2 - v_2w_1)|$. Der Ausdruck $F = |v_1w_2 - v_2w_3|$ ist nach 1. gerade der Flächeninhalt des von $v = v_1e_1 + v_2e_2$ und $w = w_1e_1 + w_2e_2$ aufgespannten Parallelogramms, also dem Flächeninhalt der Grundfläche des Spats. Da $h = |u_3|$ gerade die Höhe dieses Spats ist, ist sein Volumen

$$V = h\cdot F = |\mathrm{vol}_3(u,v,w)| = |u_3(v_1w_2 - v_2w_1)|.$$

Zum Beweis des allgemeinen Falls überlegt man sich, dass man die Ebene $\mathrm{span}_{\mathbb{R}}(\{v, w\})$ durch eine Drehung in der x_2x_3-Ebene und eine Drehung in der x_1x_3-Ebene in die Ebene $\mathrm{span}_{\mathbb{R}}(\{e_1, e_2\})$ überführen kann. Wir werden später zeigen, dass sich die Volumenform unter einer solchen Drehung nicht ändert, und damit folgt die Aussage für den allgemeinen Fall. □

Da in Lemma 5.2.11 nur der Betrag der Volumenform auftritt, ergibt sich die Frage, ob auch ihr Vorzeichen eine geometrische Bedeutung besitzt. Dazu muss zunächst geklärt werden, dass die Volumenform auf einer geordneten Basis des $\mathbb{R}^n$ nie den Wert Null annimmt. Generell ist es naheliegend, dass das n-dimensionale Volumen genau dann verschwinden sollte, wenn die Vektoren in einer $(n-1)$-dimensionalen Hyperbene des $\mathbb{R}^n$ enthalten sind, also genau dann, wenn sie linear abhängig sind.

Lemma 5.2.12: Für jedes n-Tupel von Vektoren $(v_1, ..., v_n)$ mit $v_1, ..., v_n \in \mathbb{K}^n$ gilt $\mathrm{vol}_n(v_1, ..., v_n) = 0$ genau dann, wenn $(v_1, ..., v_n)$ linear abhängig ist.

Beweis:
Ist $(v_1, ..., v_n)$ linear abhängig, so lässt sich nach Lemma 3.2.3 einer der Vektoren als Linearkombination der übrigen ausdrücken. Wir nehmen ohne Beschränkung der Allgemeinheit an, dass es v_n ist. Also gibt es $\lambda_1, ..., \lambda_{n-1} \in \mathbb{K}$ mit $v_n = \Sigma_{i=1}^{n-1}\lambda_i v_i$. Aus der Multilinearität von vol_n und der Tatsache, dass vol_n alternierend ist, ergibt sich dann

$$\mathrm{vol}_n(v_1, ..., v_n) = \mathrm{vol}_n(v_1, ..., v_{n-1}, \Sigma_{i=1}^{n-1}\lambda_i v_i) = \Sigma_{i=1}^{n-1}\lambda_i \mathrm{vol}_n(v_1, ..., v_{n-1}, v_i) = 0.$$

Sei jetzt $(v_1, ..., v_n)$ linear unabhängig. Nach Lemma 5.2.4 ändert sich die Volumenform nicht, wenn man Vielfache eines Vektors v_i zu einem anderen Vektor v_j hinzuaddiert und ändert sich nur um ein Vorzeichen, wenn man Argumente vertauscht. Fasst man die Vektoren $v_1, ..., v_n$ als Spalten einer Matrix $M \in \mathrm{Mat}(n \times n, \mathbb{K})$ aus, so entspricht dieses Hinzuaddieren und Vertauschen von Vektoren gerade dem Hinzuaddieren eines Vielfachen einer Spalte zu einer anderen und dem Vertauschen von Spalten.

Nach Satz 4.5.16 (Gauß-Algorithmus zum Invertieren von Matrizen) lässt sich jede invertierbare Matrix durch Hinzuaddieren von Vielfachen von Spalten zu anderen Spalten und Vertauschen von Spalten auf Diagonalgestalt bringen[13], und in der Diagonale stehen dann nur Einträge $\neq 0$. Ist das n-Tupel $(v_1, ..., v_n)$ linear unabhängig, so ist die zugehörige Matrix invertierbar. Der Gauß-Algorithmus liefert dann ein n-Tupel $(v_1', ..., v_n')$ mit $v_i' = \lambda_i e_i$, $\lambda_i \in \mathbb{K} \setminus \{0\}$ und $\mathrm{vol}_n(v_1, ..., v_n) = \pm \mathrm{vol}_n(v_1', ..., v_n')$. Daraus ergibt sich

$$\mathrm{vol}_n(v_1, ..., v_n) = \pm \mathrm{vol}_n(\lambda_1 e_1, ..., \lambda_n e_n) = \pm \lambda_1 \cdots \lambda_n \, \mathrm{vol}_n(e_1, ..., e_n) = \pm \lambda_1 \cdots \lambda_n \neq 0.$$

□

Insbesondere gilt nach Lemma 5.2.12 für jede geordnete Basis $B = (b_1, ..., b_n)$ des $\mathbb{R}^n$ entweder $\mathrm{vol}_n(B) := \mathrm{vol}_n(b_1, ..., b_n) > 0$ oder $\mathrm{vol}_n(B) < 0$. Dies definiert eine Äquivalenzrelation auf der Menge $(\mathbb{R}^n)^n$ der geordneten Basen des $\mathbb{R}^n$, wobei zwei geordnete Basen B und C als äquivalent betrachtet werden, wenn $\mathrm{vol}_n(B)$ und $\mathrm{vol}_n(C)$ das gleiche Vorzeichen haben. Eine Äquivalenzklasse orientierter Basen des $\mathbb{R}^n$ bezeichnet man als eine *Orientierung* auf dem $\mathbb{R}^n$.

Definition 5.2.13:

1. Eine geordnete Basis $B = (b_1, ..., b_n)$ des $\mathbb{R}^n$ heißt **positiv orientiert**, wenn $\mathrm{vol}_n(b_1, ..., b_n) > 0$ und **negativ orientiert** wenn $\mathrm{vol}_n(b_1, ..., b_n) < 0$.
2. Eine **Orientierung** auf dem $\mathbb{R}^n$ ist eine Äquivalenzklasse orientierter Basen, wobei zwei Basen äquivalent sind, wenn sie beide positiv oder beide negativ orientiert sind.

Beispiel 5.2.14:

1. Die Standardbasis des $\mathbb{R}^n$ ist per Definition der Volumenform positiv orientiert.
2. Entsteht eine Basis $B' = (b_{\sigma(1)}, ..., b_{\sigma(n)})$ aus einer Basis $B = (b_1, ..., b_n)$ durch eine Permutation $\sigma \in S_n$ der Vektoren in B, so haben B und B' die gleiche Orientierung genau dann, wenn $\mathrm{sgn}(\sigma) = 1$. Denn nach Lemma 5.2.5 gilt

   $$\mathrm{vol}_n(B') = \mathrm{vol}_n(b_{\sigma(1)}, ..., b_{\sigma(n)}) = \mathrm{sgn}(\sigma)\mathrm{vol}_n(b_1, ..., b_n) = \mathrm{sgn}(\sigma)\mathrm{vol}_n(B).$$

3. Für Basen im $\mathbb{R}^2$ und $\mathbb{R}^3$ gilt die **Rechte-Hand-Regel**:

   Ist $B = (b_1, b_2)$ eine geordnete Basis des $\mathbb{R}^2$ und zeigt der Daumen der rechten Hand in Richtung b_1, der Zeigefinger in Richtung von b_2, so ist B positiv (negativ) orientiert genau dann, wenn der Mittelfinger zum (weg vom) Betrachter zeigt.

   Ist $B = (b_1, b_2, b_3)$ eine geordnete Basis des $\mathbb{R}^3$, zeigt der Daumen der rechten Hand in Richtung des Basisvektors b_1 und der Zeigefinger in Richtung des Basisvektors b_2 dann ist B positiv (negativ) orientiert genau dann, wenn der Mittelfinger in Richtung des Basisvektors b_3 $(-b_3)$ zeigt.

[13]Im Gauß-Algorithmus in Satz 4.5.16 wurden Vielfache von Zeilen zu anderen Zeilen hinzuaddiert und Zeilen vertauscht. Dies funktioniert aber analog mit Spalten, da man statt einer Matrix auch ihre Transponierte betrachten kann, wobei Zeilen durch Spalten ersetzt werden und umgekehrt.

5.3 Determinanten

Die Betrachtungen zur Volumenform und zur Orientierung legen es nahe, das Verhalten von Volumenformen unter der Anwendung einer linearen Abbildung auf jedes Argument systematisch zu untersuchen. So kann man beispielsweise ein Tripel von Vektoren $u, v, w \in \mathbb{R}^3$ als das Bild der Standardbasis unter einer linearen Abbildung $\phi : \mathbb{R}^3 \to \mathbb{R}^3$ auffassen . Der von u, v, w aufgespannte Spat ist dann das Bild des Einheitswürfels, also des von $e_1, e_2, e_3 \in \mathbb{R}^3$ aufgespannten Spats. Ebenso kann man aus dem Vorzeichen der Volumenform direkt ablesen, ob das Bild einer geordneten Basis unter einer invertierbaren linearen Abbildung $\phi : \mathbb{R}^n \to \mathbb{R}^n$ positiv oder negativ orientiert ist.

Während sich Multilinearformen allgemein unter linearen Abbildungen ziemlich kompliziert transformieren können, ist das Transformationsverhalten einer alternierenden n-Linearform auf einem n-dimensionalen Vektorraum V wesentlich einfacher. Wendet man eine lineare Abbildung $\phi : V \to V$ auf jedes Argument der n-Linearform an, so wird die n-Linearform nämlich einfach mit einem Faktor multipliziert. Da alle n-Linearformen auf V skalare Vielfache voneinander sind, ist dieser Faktor für jede n-Linearform gleich, hängt aber natürlich von der Abbildung ϕ ab. Dies führt auf das Konzept der *Determinante.*

Satz 5.3.1: Sei V ein $\mathbb{K}$-Vektorraum der Dimension $n \in \mathbb{N}$ und $\phi : V \to V$ eine lineare Abbildung. Dann gibt es genau eine Zahl $\lambda_\phi \in \mathbb{K}$, so dass für alle alternierenden n-Linearformen $\omega : V^{\times n} \to \mathbb{K}$ und alle Vektoren $v_1, ..., v_n \in V$ gilt

$$\omega(\phi(v_1), ..., \phi(v_n)) = \lambda_\phi\, \omega(v_1, ..., v_n).$$

Beweis:
Für jede alternierende n-Linearform $\omega : V^{\times n} \to \mathbb{K}$ und jede lineare Abbildung $\phi : V \to V$ ist auch die Abbildung $\phi^*\omega : V^{\times n} \to \mathbb{K}$, $(v_1, ..., v_n) \mapsto \omega(\phi(v_1), ..., \phi(v_n))$ wieder n-linear und alternierend, denn es gilt für alle $v, w \in V$, $\mu \in \mathbb{K}$

$$\begin{aligned}
\phi^*\omega(..., v + w, ...) &= \omega(..., \phi(v + w), ...) = \omega(..., \phi(v) + \phi(w), ...) = \omega(..., \phi(v), ...) + \omega(..., \phi(w), ...)\\
&= \phi^*\omega(..., v, ...) + \phi^*\omega(..., w, ...)\\
\phi^*\omega(..., \mu\, v, ...) &= \omega(..., \phi(\mu\, v), ...) = \omega(..., \mu\, \phi(v), ...) = \mu\, \omega(..., \phi(v), ...) = \mu\, \phi^*\omega(..., v, ...)\\
\phi^*\omega(..., v, ..., v, ...) &= \omega(..., \phi(v), ..., \phi(v), ...) = 0.
\end{aligned}$$

Nach Korollar 5.2.9 ist aber der $\mathbb{K}$-Vektorraum $A(V^{\times n}, \mathbb{K})$ der alternierenden n-Linearformen auf V eindimensional. Ist $\omega \in A(V^{\times n}, \mathbb{K}) \setminus \{0\}$, muss also ein Skalar $\lambda_\phi \in \mathbb{K}$ mit $\phi^*\omega = \lambda_\phi\, \omega$ existieren. Ebenso gibt es zu jeder anderen Linearform $\omega' \in A(V^{\times n}, \mathbb{K})$ einen Skalar $\mu \in \mathbb{K}$ mit $\omega' = \mu\omega$, und daraus folgt $\phi^*\omega' = \phi^*(\mu\omega) = \mu\phi^*\omega = \mu\lambda_\phi\omega = \lambda_\phi\omega'$ für alle $\omega' \in A(V^{\times n}, \mathbb{K})$. □

Definition 5.3.2: Sei $\mathbb{K}$ ein Körper und V ein $\mathbb{K}$-Vektorraum der Dimension $n \in \mathbb{N}$.

1. Die **Determinante** einer $\mathbb{K}$-linearen Abbildung $\phi : V \to V$ ist die eindeutig bestimmte Zahl $\det(\phi) = \lambda_\phi \in \mathbb{K}$ mit $\phi^*\omega = \lambda_\phi\, \omega$ aus Satz 5.3.1.
2. Die **Determinante** einer quadratischen Matrix $M \in \mathrm{Mat}(n \times n, \mathbb{K})$ ist die Determinante $\det(M) = \det(\phi_M)$ der zugehörigen linearen Abbildung $\phi_M : \mathbb{K}^n \to \mathbb{K}^n$, $v \mapsto M \cdot v$.

Bemerkung 5.3.3:

1. Die Determinante eines Endomorphismus $\phi : V \to V$ lässt sich berechnen, indem man die Bilder der Basisvektoren in einer geordneten Basis von V als Linearkombination der Basisvektoren ausdrückt. Ist beispielsweise (v_1, v_2) eine Basis von V und $\phi(v_1) = av_1 + cv_2$, $\phi(v_2) = bv_1 + dv_2$ so gilt für jede alternierende Bilinearform ω auf V

$$\begin{aligned}
&\phi^*\omega(v_1, v_2) = \omega(\phi(v_1), \phi(v_2)) = \omega(av_1 + cv_2, bv_1 + dv_2)\\
&= \omega(av_1, bv_1) + \omega(av_1, dv_2) + \omega(cv_2, bv_1) + \omega(cv_2, dv_2) = ad\omega(v_1, v_2) + bc\omega(v_2, v_1)\\
&= (ad - bc)\omega(v_1, v_2) \qquad \Rightarrow \qquad \det(\phi) = ad - bc.
\end{aligned}$$

2. Der Betrag der Determinante gibt an, wie sich Volumina unter linearen Abbildungen $\phi : \mathbb{R}^n \to \mathbb{R}^n$ ändern. Das Bild des Einheitswürfels unter ϕ ist das von den Vektoren $\phi(e_1), ..., \phi(e_n)$ aufgespannte n-dimensionale **Parallelotop**. Mit $\omega = \mathrm{vol}_n$ ergibt sich sein Volumen als $|\det(\phi)| = |\det(\phi)\,\mathrm{vol}_n(e_1, ..., e_n)| = |\mathrm{vol}_n(\phi(e_1), ..., \phi(e_n))|$. Das Vorzeichen der Determinante gibt an, ob ϕ die Orientierung einer Basis umkehrt.

   Analog ist das Volumen des von Vektoren $v_1, ..., v_n \in \mathbb{R}^n$ aufgespannten Parallelotops $|\det(M)|$, wobei $M = (v_1, ..., v_n) \in \mathrm{Mat}(n \times n, \mathbb{R})$. Ist $(v_1, ..., v_n)$ eine geordnete Basis des $\mathbb{R}^n$, so gibt das Vorzeichen von $\det(M)$ ihre Orientierung an.

3. Für $1 \leq i < j \leq n$ ist die Drehung in der $x_i x_j$-Ebene im $\mathbb{R}^n$ um einen Winkel $\theta \in \mathbb{R}$

$$R^\theta_{ij} : \mathbb{R}^n \to \mathbb{R}^n, \quad R^\theta_{ij}(e_k) = \begin{cases} e_i \cos\theta + e_j \sin\theta & k = i \\ e_j \cos\theta - e_i \sin\theta & k = j \\ e_k & k \notin \{i, j\} \end{cases}.$$

   Aus der Multilinearität der Volumenform ergibt sich

$$\begin{aligned}
&\det(R^\theta_{ij}) = \mathrm{vol}_n(R^\theta_{ij}(e_1), ..., R^\theta_{ij}(e_n)) = \mathrm{vol}_n(..., e_i \cos\theta + \sin\theta e_j, ..., e_j \cos\theta - e_i \sin\theta, ...)\\
&= \cos^2\theta\ \mathrm{vol}_n(..., e_i, ..., e_j, ...) + \cos\theta \sin\theta\ \mathrm{vol}_n(..., e_j, ..., e_j, ...)\\
&- \cos\theta \sin\theta\ \mathrm{vol}_n(..., e_i, ..., e_i, ...) - \sin^2\theta\ \mathrm{vol}_n(..., e_j, ...e_i, ...)\\
&= (\cos^2\theta + \sin^2\theta)\mathrm{vol}_n(e_1, ..., e_n) = 1.
\end{aligned}$$

   Drehungen erhalten also Volumina und die Orientierung.

4. Die Spiegelung an der Koordinaten-Hyperebene senkrecht zur x_j-Achse ist die lineare Abbildung $S_j : \mathbb{R}^n \to \mathbb{R}^n$ mit $S_j(e_j) = -e_j$ und $S_j(e_k) = e_k$ für $k \neq j$. Mit der Multilinearität der Volumenform folgt

$$\det(S_j) = \mathrm{vol}_n(S_j(e_1), ..., S_j(e_n)) = \mathrm{vol}_n(e_1, ..., e_{j-1}, -e_i, e_{j+1}, ..., e_n) = -\mathrm{vol}(e_1, ..., e_n) = -1.$$

   Spiegelungen erhalten also Volumina und kehren die Orientierung um.

Wir beschäftigen uns nun detaillierter mit den Eigenschaften der Determinante. Insbesondere ist zu klären, wie sich die Determinante unter der Addition, skalaren Multiplikation und Verkettung von linearen Abbildungen verhält. Anschaulich erwartet man dabei, dass ein n-dimensionales Volumen unter einer Streckung mit einem Skalar $\lambda \in \mathbb{R}$ mit einem Faktor $|\lambda|^n \in \mathbb{R}$ multipliziert werden sollte. Die Änderung eines Volumens unter der Verkettung zweier linearer Abbildungen sollte das Produkt der Volumenänderungen der einzelnen Abbildungen sein. Das Verhalten der Determinante unter der punktweisen Addition linearer Abbildungen ist dagegen kompliziert.

Satz 5.3.4: (**Eigenschaften der Determinante**)
Sei $\mathbb{K}$ ein Körper und V ein n-dimensionaler $\mathbb{K}$-Vektorraum. Dann gilt:

(i) $\det(\mu\,\phi) = \mu^n \det(\phi)$ für alle $\mathbb{K}$-linearen Abbildungen $\phi : V \to V$ und $\mu \in \mathbb{K}$.
(ii) $\det(\phi \circ \psi) = \det(\phi) \cdot \det(\psi)$ für alle $\mathbb{K}$-linearen Abbildungen $\phi, \psi : V \to V$.
(iii) $\det(\mathrm{id}_V) = 1$.
(iv) $\det(\phi^{-1}) = \det(\phi)^{-1}$ für alle invertierbaren linearen Abbildungen $\phi : V \to V$.

Aussage (ii) bezeichnet man als den **Determinantenmultiplikationssatz**.

Beweis:
Aus der Definition der Determinante in Satz 5.3.1 und der n-Linearität folgt

$$\begin{aligned}
\det(\mu\,\phi)\,\omega(v_1,...,v_n) &= \omega(\mu\phi(v_1),...,\mu\phi(v_n)) = \mu^n\omega(\phi(v_1),...,\phi(v_n)) = \mu^n \det(\phi)\,\omega(v_1,...,v_n)\\
\det(\phi\circ\psi)\,\omega(v_1,...,v_n) &= \omega(\phi(\psi(v_1)),...,\phi(\psi(v_n))) = \det(\phi)\,\omega(\psi(v_1),...,\psi(v_n))\\
&= \det(\phi)\cdot\det(\psi)\,\omega(v_1,...,v_n)\\
\det(\mathrm{id}_V)\,\omega(v_1,...,v_n) &= \omega(\mathrm{id}_V(v_1),...,\mathrm{id}_V(v_n)) = \omega(v_1,...,v_n)
\end{aligned}$$

für alle $\omega \in A(V^{\times n},\mathbb{K})$ und $v_1,...,v_n \in V$. Da $A(V^{\times n},\mathbb{K})$ eindimensional ist, gibt es ein $\omega \in A(V^{\times n},\mathbb{K})$ und Vektoren $v_1,...,v_n \in V$ mit $\omega(v_1,...,v_n) \neq 0$ und daraus folgen (i)-(iii). Für alle invertierbaren linearen Abbildungen $\phi : V \to V$ ergibt sich dann aus (ii) und (iii) $1 = \det(\mathrm{id}_V) = \det(\phi\circ\phi^{-1}) = \det(\phi)\cdot\det(\phi^{-1})$. □

Aus den Eigenschaften der Determinanten linearer Abbildungen und der Definition der Determinante einer Matrix $M \in \mathrm{Mat}(n\times n,\mathbb{K})$ als Determinante der linearen Abbildung $\phi_M : \mathbb{K}^n \to \mathbb{K}^n$, $v \mapsto M\cdot v$ ergeben sich nun direkt analoge Aussagen über die Determinanten von Matrizen.

Satz 5.3.5: (**Eigenschaften der Determinante**)
Sei $\mathbb{K}$ ein Körper und $n \in \mathbb{N}$. Dann gilt:

(i) $\det(\mu\,M) = \mu^n \det(M)$ für alle $M \in \mathrm{Mat}(n\times n,\mathbb{K})$ und $\mu \in \mathbb{K}$.
(ii) $\det(M\cdot N) = \det(M)\cdot\det(N)$ für alle $M,N \in \mathrm{Mat}(n\times n,\mathbb{K})$.
(iii) $\det(\mathbb{1}_n) = 1$.
(iv) $\det(M^{-1}) = \det(M)^{-1}$ für alle $M \in \mathrm{GL}(n,\mathbb{K})$.
(v) Ähnliche Matrizen haben die gleiche Determinante.

Aussage (ii) bezeichnet man als den **Determinantenmultiplikationssatz**.

Beweis:
Die Determinante einer Matrix $M \in \mathrm{Mat}(n\times n,\mathbb{K})$ ist definiert als die Determinante der linearen Abbildung $\phi_M : \mathbb{K}^n \to \mathbb{K}^n$, $v \mapsto M\cdot v$. Diese erfüllt die Bedingungen

$$\phi_{\mu M} = \mu\phi_M : v \mapsto \mu M\cdot v, \qquad \phi_{M\cdot N} = \phi_M\circ\phi_N : v\mapsto M\cdot N\cdot v, \qquad \phi_{\mathbb{1}_n} = \mathrm{id}_{\mathbb{K}^n} : v \mapsto v$$

für alle $\mu \in \mathbb{K}$ und $M,N \in \mathrm{Mat}(n\times n,\mathbb{K})$ sowie $\phi_{A^{-1}} = \phi_A^{-1} : v \mapsto A^{-1}\cdot v$ für alle $A \in \mathrm{GL}(n,\mathbb{K})$. Mit Satz 5.3.4 folgen dann Aussagen (i)-(iv):

(i) $\det(\mu M) = \det(\phi_{\mu M}) = \det(\mu\phi_M) = \mu^n\,\det(\phi_M) = \mu^n\det(M)$
(ii) $\det(M\cdot N) = \det(\phi_{M\cdot N}) = \det(\phi_M\circ\phi_N) = \det(\phi_M)\cdot\det(\phi_N) = \det(M)\cdot\det(N)$
(iii) $\det(\mathbb{1}_n) = \det(\phi_{\mathrm{id}_{\mathbb{K}^n}}) = 1$
(iv) $\det(A^{-1}) = \det(\phi_{A^{-1}}) = \det(\phi_A^{-1}) = \det(\phi_A)^{-1} = \det(A)^{-1}$.

Sind $M, N \in \mathrm{Mat}(n \times n, \mathbb{K})$ ähnlich, so existiert eine Matrix $S \in \mathrm{GL}(n, \mathbb{K})$ mit $M = S \cdot N \cdot S^{-1}$. Mit (ii) und (iv) folgt dann $\det(M) = \det(S \cdot N \cdot S^{-1}) = \det(S) \cdot \det(N) \cdot \det(S^{-1}) = \det(N)$. □

Mit der Determinante haben wir also ein weiteres Werkzeug, mit dem wir zeigen können, dass zwei Matrizen nicht ähnlich sind. Da ähnliche Matrizen die gleiche Determinante haben, sind zwei Matrizen mit verschiedenen Determinanten niemals ähnlich. Die Umkehrung gilt allerdings nicht: es gibt Matrizen die nicht ähnlich sind, aber die gleiche Determinante haben (Übung).

Eine weitere Konsequenz aus Satz 5.3.5 ist, dass die Determinante einer darstellenden Matrix einer linearen Abbildung $\phi : V \to V$ bezüglich jeder Basis B von V immer gleich der Determinante von ϕ ist.

Korollar 5.3.6: Sei $\mathbb{K}$ ein Körper, V ein endlich-dimensionaler $\mathbb{K}$-Vektorraum und $\phi : V \to V$ eine lineare Abbildung. Dann gilt für jede geordnete Basis B von V

$$\det({}_BM_B(\phi)) = \det(\phi).$$

Beweis:
Sei $\dim_{\mathbb{K}}(V) = n$ und $B = (b_1, ..., b_n)$ eine geordnete Basis von V. Wir betrachten die invertierbare $\mathbb{K}$-lineare Abbildung ${}_BS : V \to \mathbb{K}^n$, $\Sigma_{i=1}^n \lambda_i b_i \mapsto (\lambda_1, ..., \lambda_n)^T$ aus Korollar 4.2.4, die einem Vektor $v \in V$ seinen beschreibenden Spaltenvektor bezüglich B zuordnet. Damit definieren wir die alternierende n-Linearform

$$\omega = {}_BS^*\mathrm{vol}_n : V^{\times n} \to \mathbb{K}, \quad (v_1, ..., v_n) \mapsto \mathrm{vol}_n({}_BS(v_1), \ldots, {}_BS(v_n)).$$

Nach Satz 4.2.12, (iii) gilt ${}_BM_B(\phi) \cdot {}_BS(v) = {}_BS(\phi(v))$ für die beschreibende Matrix ${}_BM_B(\phi)$ eines Endomorphismus $\phi \in \mathrm{End}_{\mathbb{K}}(V)$ und jeden Vektor $v \in V$. Mit der Definition der Determinante ergibt sich dann

$$\begin{aligned}
&\det(\phi)\, \omega(v_1, ..., v_n) = \omega(\phi(v_1), ..., \phi(v_n)) = \mathrm{vol}_n({}_BS(\phi(v_1)), \ldots, {}_BS(\phi(v_n))) \\
&= \mathrm{vol}_n({}_BM_B(\phi) \cdot {}_BS(v_1), \ldots, {}_BM_B(\phi) \cdot {}_BS(v_n)) = \det({}_BM_B(\phi))\, \mathrm{vol}_n({}_BS(v_1), \ldots, {}_BS(v_n)) \\
&= \det({}_BM_B(\phi))\, \omega(v_1, \ldots, v_n).
\end{aligned}$$

Da ${}_BS$ invertierbar ist, gilt $\omega \neq 0$, und damit existieren $v_1, ..., v_n \in V$ mit $\omega(v_1, \ldots, v_n) \neq 0$. Nach Division durch $\omega(v_1, ..., v_n)$ folgt die Behauptung. □

Korollar 5.3.6 erlaubt es einem, die Determinante eines Vektorraumendomorphismus $\phi : V \to V$ zu berechnen, indem man eine beliebige Basis von V wählt und die Determinante der darstellenden Matrix ${}_BM_B(\phi)$ berechnet. Das wird sich als sehr nützlich erweisen, wenn wir effiziente Berechnungsverfahren für die Determinanten von Matrizen hergeleitet haben.

Eine weitere wichtige Konsequenz aus Satz 5.3.1 und Satz 5.3.4 ist, dass wir durch die Determinante das Verhalten von Automorphismen des $\mathbb{R}^n$ bezüglich der Orientierung charakterisieren können. Es ergibt sich aus dem Vorzeichen der Determinante.

Korollar 5.3.7: Sei $\phi : \mathbb{R}^n \to \mathbb{R}^n$ ein Vektorraumautomorphismus. Dann gilt:

1. Ist $\det(\phi) > 0\, (< 0)$, so bildet ϕ positiv orientierten Basen auf positiv (negativ) orientierte Basen ab und negativ orientierte auf negativ (positiv) orientierte. Im ersten Fall nennt man ϕ **orientierungserhaltend**, im zweiten **orientierungsumkehrend**.

2. Die orientierungserhaltenden Automorphismen $\phi : \mathbb{R}^n \to \mathbb{R}^n$ bilden mit der Verkettung eine Untergruppe $\mathrm{Aut}^+_{\mathbb{R}}(\mathbb{R}^n)$ der Gruppe $\mathrm{Aut}_{\mathbb{R}}(\mathbb{R}^n)$.
3. Die Matrizen $M \in \mathrm{GL}(n, \mathbb{R})$ mit $\det(M) > 0$ bilden eine Untergruppe $\mathrm{GL}(n, \mathbb{R})^+$ der Gruppe $\mathrm{GL}(n, \mathbb{R})$.

Beweis:
Per Definition der Determinante gilt für jede geordnete Basis $B = (v_1, ..., v_n)$ des $\mathbb{R}^n$ und jede lineare Abbildung $\phi : \mathbb{R}^n \to \mathbb{R}^n$ die Gleichung $\mathrm{vol}_n(\phi(v_1), ..., \phi(v_n)) = \det(\phi) \cdot \mathrm{vol}_n(v_1, ..., v_n)$.

Ist $\phi : \mathbb{R}^n \to \mathbb{R}^n$ invertierbar, so ist auch $\phi(B) = (\phi(v_1), ..., \phi(v_n))$ eine Basis des $\mathbb{R}^n$ und nach Lemma 5.2.12 gilt $\mathrm{vol}_n(v_1, ..., v_n) \neq 0$, $\mathrm{vol}_n(\phi(v_1), ..., \phi(v_n)) \neq 0$. Also haben sie das gleiche Vorzeichen, wenn $\det(\phi) > 0$ und verschiedene Vorzeichen, wenn $\det(\phi) < 0$ gilt.

Die zweite Aussage folgt aus dem Determinantenmultiplikationssatz. Für $\phi, \psi : \mathbb{R}^n \to \mathbb{R}^n$ gilt nämlich $\det(\phi \circ \psi) = \det(\phi) \cdot \det(\psi)$. Dieses Produkt ist immer positiv, wenn beide Faktoren positiv sind. Der Beweis der dritten Aussage ist analog. $\square$

Beispiel 5.3.8:

1. Die $\mathbb{R}$-lineare Abbildung $\phi : \mathbb{R}^n \to \mathbb{R}^n$, $v \mapsto -v$ ist orientierungserhaltend genau dann, wenn n gerade ist. Denn aus der Multilinearität der Volumenform folgt

$$\det(\phi) = \mathrm{vol}_n(\phi(e_1), ..., \phi(e_n)) = \mathrm{vol}_n(-e_1, ..., -e_n) = (-1)^n \mathrm{vol}_n(e_1, .., e_n) = (-1)^n.$$

2. Drehungen in der $x_i x_j$-Ebene sind sind orientierungserhaltend. Denn nach dem Beweis von Lemma 5.2.11 und Bemerkung 5.3.3 gilt $\det(R^\theta_{ij}) = 1$.
3. Jede Spiegelung an einer Koordinaten-Hyperbene des $\mathbb{R}^n$ ist orientierungsumkehrend. Denn nach Bemerkung 5.3.3 gilt für die Spiegelung $S_j : \mathbb{R}^n \to \mathbb{R}^n$ an der Koordinatenhyperebene senkrecht zu x_j die Formel $\det(S_j) = -1$.

Nachdem wir die wichtigsten Eigenschaften der Determinanten geklärt haben, leiten wir nun eine konkrete Formel her, mit der sich Determinanten berechnen lassen. Diese liefert insbesondere einfache Merkregeln für die Determinanten von (2×2)- und (3×3)-Matrizen.

Satz 5.3.9: (**Leibnizsche Formel**)
Sei $\mathbb{K}$ ein Körper und $n \in \mathbb{N}$. Dann gilt für jede Matrix $A = (a_1, ..., a_n) \in \mathrm{Mat}(n \times n, \mathbb{K})$

$$\det(A) = \mathrm{vol}_n(a_1, ..., a_n) = \Sigma_{\sigma \in S_n} \mathrm{sgn}(\sigma)\, A_{\sigma(1)1} \cdots A_{\sigma(n)n}.$$

Beweis:
Nach Satz 5.3.1 gilt für die lineare Abbildung $\phi_A : \mathbb{K}^n \to \mathbb{K}^n$, $v \mapsto A \cdot v$

$$\det(A) = \det(\phi_A) = \mathrm{vol}_n(Ae_1, ..., Ae_n) = \mathrm{vol}_n(a_1, ..., a_n),$$

denn wir können die alternierende n-Linearform $\omega = \mathrm{vol}_n$ in Satz 5.3.1 betrachten und $v_i = e_i$ setzen. Die Spaltenvektoren der Matrix A sind gegeben durch $a_i = \Sigma_{j=1}^n A_{ji} e_j$, und aus der n-Linearität der Volumenform folgt dann

$$\det(A) = \mathrm{vol}_n(\Sigma_{j_1=1}^n A_{j_1 1} e_{j_1}, ..., \Sigma_{j_n=1}^n A_{j_n n} e_{j_n}) = \Sigma_{j_1=1}^n \cdots \Sigma_{j_n=1}^n A_{j_1 1} \cdots A_{j_n n}\, \mathrm{vol}_n(e_{j_1}, ..., e_{j_n}).$$

Da die Volumenform alternierend ist, gilt $\mathrm{vol}_n(e_{j_1}, ..., e_{j_n}) = 0$ wenn zwei oder mehr der Indizes $j_1, ..., j_n$ gleich sind. Es tragen also nur Summanden mit paarweise verschiedenen Indizes $j_1, ..., j_n$ zur Summe bei. Die Indizes $j_1, ..., j_n$ sind genau dann paarweise verschieden, wenn es eine Permutation $\sigma \in S_n$ gibt mit $j_k = \sigma(k)$ für alle $1 \leq k \leq n$. Nach Lemma 5.2.5 gilt dann

$$\mathrm{vol}_n(e_{j_1}, ..., e_{j_n}) = \mathrm{vol}_n(e_{\sigma(1)}, ..., e_{\sigma(n)}) = \mathrm{sgn}(\sigma)\,\mathrm{vol}_n(e_1, ..., e_n) = \mathrm{sgn}(\sigma).$$

Wir können also die Summen über die Indizes $j_1, ..., j_n$ in der Formel für die Determinante durch eine Summe über alle Permutationen $\sigma \in S_n$ ersetzen, $j_k = \sigma(k)$ für alle $1 \leq k \leq n$ und $\mathrm{vol}_n(e_{j_1}, ..., e_{j_n}) = \mathrm{sgn}(\sigma)$ setzen. So erhalten wir

$$\det(A) = \Sigma_{j_1=1}^n \cdots \Sigma_{j_n=1}^n A_{j_1 1} \cdots A_{j_n n}\,\mathrm{vol}_n(e_{j_1}, ..., e_{j_n}) = \Sigma_{\sigma \in S_n} \mathrm{sgn}(\sigma)\, A_{\sigma(1)1} \cdots A_{\sigma(n)n}.$$

□

Beispiel 5.3.10:

1. Für (2×2)-Matrizen mit Einträgen in $\mathbb{K}$ berechnet sich die Determinante nach der Regel **Hauptdiagonale minus Gegendiagonale**:

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}a_{22} - a_{21}a_{12}.$$

   Denn $S_2 = \{\mathrm{id}, \sigma_{12}\}$, und mit der Leibnizschen Formel erhält man

$$\det(A) = \mathrm{sgn}(\mathrm{id})a_{11}a_{22} + \mathrm{sgn}(\sigma_{12})a_{21}a_{12} = a_{11}a_{22} - a_{21}a_{12}.$$

2. Für (3×3)-Matrizen mit Einträgen in $\mathbb{K}$ gilt die **Regel von Sarrus**

$$\det(A) = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}.$$

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

   Denn $S_3 = \{\mathrm{id}, \tau, \tau^2, \sigma_{12}, \sigma_{13}, \sigma_{23}\}$, wobei τ die zyklische Vertauschung mit $\tau(1) = 2$, $\tau(2) = 3$ und $\tau(3) = 1$ bezeichnet. Das Signum dieser Permutationen ist gegeben durch $\mathrm{sgn}(\tau) = \mathrm{sgn}(\tau^2) = \mathrm{sgn}(\mathrm{id}) = 1$ und $\mathrm{sgn}(\sigma_{12}) = \mathrm{sgn}(\sigma_{13}) = \mathrm{sgn}(\sigma_{23}) = -1$, und es folgt

$$\begin{aligned}\det(A) &= \mathrm{sgn}(\mathrm{id})a_{11}a_{22}a_{33} + \mathrm{sgn}(\sigma_{12})a_{21}a_{12}a_{33} + \mathrm{sgn}(\sigma_{13})a_{31}a_{22}a_{13} \\ &\quad + \mathrm{sgn}(\sigma_{23})a_{11}a_{32}a_{23} + \mathrm{sgn}(\tau)a_{21}a_{32}a_{13} + \mathrm{sgn}(\tau^2)a_{31}a_{12}a_{23} \\ &= a_{11}a_{22}a_{33} + a_{21}a_{32}a_{13} + a_{31}a_{12}a_{23} - a_{21}a_{12}a_{33} - a_{31}a_{22}a_{13} - a_{11}a_{32}a_{23} \\ &= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}.\end{aligned}$$

Die Merkregeln für Determinanten aus diesen Beispielen sollte man auswendig lernen, da einem dies langfristig viel Rechenarbeit spart. Für $(n \times n)$-Matrizen mit $n > 3$ ergeben sich jedoch keine brauchbaren Verallgemeinerungen der Regel von Sarrus. Die Berechnung der Determinante mit der Leibnizschen Formel wird dann ziemlich aufwändig. Da $|S_n| = n!$, wächst der Rechenaufwand exponentiell - für $n = 4$ erhält man $4! = 24$, für $n = 5$ bereits $5! = 120$ und für $n = 6$ schon $6! = 720$ Summanden.

Um effizientere Berechnungsverfahren für Determinanten größerer Matrizen zu finden, müssen wir uns genauer mit den Eigenschaften der Determinante befassen. Der folgende Satz liefert wichtige Rechenregeln.

Satz 5.3.11: (**Rechenregeln für Determinanten**)
Sei $\mathbb{K}$ ein Körper und $n \in \mathbb{N}$. Dann gilt für die Determinante $\det : \mathrm{Mat}(n \times n, \mathbb{K}) \to \mathbb{K}$:

1. **Multilinearität:** Die Determinante ist linear in jeder Zeile und Spalte der Matrix.
2. **Verschwindende Zeilen/Spalten:**
   Ist eine Spalte oder Zeile von $A \in \mathrm{Mat}(n \times n, \mathbb{K})$ gleich 0, so ist $\det(A) = 0$.
3. **Vertauschen von Zeilen/Spalten:** Entsteht $B \in \mathrm{Mat}(n \times n, \mathbb{K})$ aus $A \in \mathrm{Mat}(n \times n, \mathbb{K})$ durch Vertauschen zweier Zeilen oder Spalten, so ist $\det(B) = -\det(A)$.
4. **Addieren von Vielfachen von Zeilen/Spalten:** Entsteht $B \in \mathrm{Mat}(n \times n, \mathbb{K})$ aus $A \in \mathrm{Mat}(n \times n, \mathbb{K})$ durch Hinzuaddieren eines Vielfachen einer Spalte (Zeile) von A zu einer anderen Spalte (Zeile) von A, so ist $\det(B) = \det(A)$.
5. **Dreiecksmatrizen:** Ist $A \in \mathrm{Mat}(n \times n, \mathbb{K})$ eine **obere Dreiecksmatrix**, d. h. eine Matrix mit $A_{ji} = 0$ für alle $j > i$, oder eine **untere Dreiecksmatrix**, d. h. eine Matrix mit $A_{ji} = 0$ für alle $j < i$, so gilt $\det(A) = A_{11} \cdot A_{22} \cdots A_{nn}$.
6. **Blockmatrizen:** Ist $A \in \mathrm{Mat}(n \times n, \mathbb{K})$ eine Matrix in **Blockform**
   $$A = \begin{pmatrix} B & C \\ 0_{q\times p} & D \end{pmatrix} \quad \text{mit } B \in \mathrm{Mat}(p\times p, \mathbb{K}), C \in \mathrm{Mat}(p\times q, \mathbb{K}), D \in \mathrm{Mat}(q\times q, \mathbb{K}), n = p+q,$$
   so gilt $\det(A) = \det(B) \cdot \det(D)$.
7. **Transponierte:** Für alle $A \in \mathrm{Mat}(n \times n, \mathbb{K})$ gilt $\det(A) = \det(A^T)$.
8. **Rang:** Für alle $A \in \mathrm{Mat}(n \times n, \mathbb{K})$ gilt: $\det(A) \neq 0 \Leftrightarrow \mathrm{rg}(A) = n \Leftrightarrow A$ invertierbar.

Beweis:
Aussagen 1.-4. für die Spalten ergeben sich direkt aus der Formel $\det(A) = \mathrm{vol}_n(a_1, ..., a_n)$ für $A = (a_1, ..., a_n)$ und den Aussagen über alternierende n-Linearformen in Lemma 5.2.4.

5. Ist A eine obere Dreiecksmatrix, so ist $A_{ji} = 0$ für $j > i$, und somit $A_{\sigma(1)1} \cdots A_{\sigma(n)n} = 0$, wenn $\sigma \in S_n$ eine Permutation mit $\sigma(i) > i$ für ein $i \in \{1, ..., n\}$ ist. Also tragen nur Permutationen $\sigma \in S_n$ mit $\sigma(i) \leq i$ für alle $i \in \{1, ..., n\}$ zur Summe in der Determinante bei. Die einzige solche Permutation ist aber $\sigma = \mathrm{id}$, denn aus $\sigma(i) \leq i$ für alle $i \in \{1, ..., n\}$ folgt $\sigma(1) = 1$, $\sigma(2) = 2$ und dann induktiv $\sigma(i) = i$ für alle $i \in \{1, ..., n\}$. Also gilt

$$\det(A) = \Sigma_{\sigma\in S_n} \mathrm{sgn}(\sigma)\, A_{\sigma(1)1} \cdots A_{\sigma(n)n} = A_{11} \cdots A_{nn}.$$

6. Ist A eine Blockmatrix, so gilt $A_{ji} = 0$ für $i \in \{1, .., p\}$ und $j \in \{p + 1, ..., n\}$. Daraus folgt $A_{\sigma(1)1} \cdots A_{\sigma(n)n} = 0$ für alle $\sigma \in S_n$ mit $\sigma(i) > p$ für ein $i \in \{1, ..., p\}$. Also tragen nur die Permutationen $\sigma \in S_n$ mit $\sigma(\{1, ..., p\}) = \{1, ..., p\}$ zur Determinante bei. Für jede solche Permutation gilt aber auch $\sigma(\{p + 1, ..., n\}) = \{p + 1, ..., n\}$. Also existieren Permutationen $\pi \in S_p$ und $\tau \in S_q$ mit $\sigma(i) = \pi(i)$ für $i \in \{1, ..., p\}$ und $\sigma(i) = \tau(i-p)+p$ für $i \in \{p+1, ..., n\}$. Umgekehrt liefert jedes solche Paar von Permutationen $\pi \in S_p$, $\tau \in S_q$ eine Permutation $\sigma \in S_n$ mit $\sigma(\{1, .., p\}) = \{1, ..., p\}$ und $\sigma(\{p + 1, ..., n\}) = \{p + 1, ..., n\}$. Da für die Fehlstandspaare gilt $F(\sigma) = F(\pi) + F(\tau)$, folgt $\mathrm{sgn}(\sigma) = (-1)^{F(\pi)+F(\tau)} = \mathrm{sgn}(\pi) \cdot \mathrm{sgn}(\tau)$ und

$$\begin{aligned}
&\det(A) = \Sigma_{\sigma\in S_n} \mathrm{sgn}(\sigma)\, A_{\sigma(1)1} \cdots A_{\sigma(n)n} \\
&= \Sigma_{\pi\in S_p,\tau\in S_q} \mathrm{sgn}(\pi) \cdot \mathrm{sgn}(\tau) A_{\pi(1)1} \cdots A_{\pi(p)p} A_{(\tau(1)+p)(p+1)} \cdots A_{(\tau(q)+p)n} \\
&= \Sigma_{\pi\in S_p,\tau\in S_q} \mathrm{sgn}(\pi) \cdot \mathrm{sgn}(\tau) B_{\pi(1)1} \cdots B_{\pi(p)p} D_{\tau(1)1} \cdots D_{\tau(q)q} \\
&= \big(\Sigma_{\pi\in S_p,\tau\in S_q} \mathrm{sgn}(\pi) B_{\pi(1)1} \cdots B_{\pi(p)p}\big) \big(\Sigma_{\tau\in S_q} \mathrm{sgn}(\tau) D_{\tau(1)1} \cdots D_{\tau(q)q}\big) = \det(B) \cdot \det(D).
\end{aligned}$$

7. Für alle $\sigma \in S_n$ und $A \in \mathrm{Mat}(n \times n, \mathbb{K})$ gilt $A_{\sigma^{-1}(1)1} \cdots A_{\sigma^{-1}(n)n} = A_{1\sigma(1)} \cdots A_{n\sigma(n)}$, denn zu jedem $i \in \{1, ..., n\}$ existiert genau ein $j \in \{1, ..., n\}$ mit $\sigma(j) = i$ und damit $A_{i\sigma(i)} = A_{\sigma^{-1}(j)j}$. Also treten auf der rechten und linken Seite des Gleichheitszeichens die gleichen Faktoren auf. Da $1 = \mathrm{sgn}(\mathrm{id}) = \mathrm{sgn}(\sigma \circ \sigma^{-1}) = \mathrm{sgn}(\sigma) \cdot \mathrm{sgn}(\sigma^{-1})$ folgt $\mathrm{sgn}(\sigma) = \mathrm{sgn}(\sigma^{-1})$, und damit

$$\det(A^T) = \Sigma_{\sigma\in S_n} \mathrm{sgn}(\sigma)\, A_{1\sigma(1)} \cdots A_{n\sigma(n)} = \Sigma_{\sigma\in S_n} \mathrm{sgn}(\sigma^{-1})\, A_{\sigma^{-1}(1)1} \cdots A_{\sigma^{-1}(n)n} = \det(A).$$

Daraus folgen insbesondere Aussagen 1.-4. für die Zeilen von A, denn die Zeilen von A entsprechen den Spalten von A^T, und Aussage 5 für die unteren Dreiecksmatrizen, denn A ist eine obere Dreiecksmatrix genau dann, wenn A^T eine untere Dreiecksmatrix ist.

8. Die letzte Aussage folgt aus Lemma 5.2.12, der Formel $\det(A) = \mathrm{vol}_n(a_1, .., a_n)$ und der Tatsache, dass $\mathrm{rg}(A) = n$ genau dann, wenn die Spalten von A linear unabhängig sind. □

Kombinieren wir Aussagen 2.-4. und 8. aus diesem Satz, so sehen wir, dass der Gauß-Algorithmus ein effizientes Verfahren zur Berechnung von Determinanten liefert. Denn durch Hinzuaddieren von Vielfachen von Zeilen zu anderen Zeilen und Vertauschen von Zeilen lässt sich jede Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{K})$ in eine Matrix A' in Zeilenstufenform überführen, wobei die Anzahl der Stufen in A' gleich dem Rang von A ist.

Hat A' weniger als n Stufen, so gilt $\det(A') = 0$. Hat A' genau n Stufen, so ist A' eine obere Dreiecksmatrix und ihre Determinante ergibt sich als Produkt der Einträge in der Diagonalen. Die Determinante von A stimmt bis auf ein Vorzeichen mit der Determinante von A' überein, da das Hinzuaddieren von Vielfachen einer Zeile zu einer anderen die Determinante nicht ändert und das Vertauschen zweier Zeilen sie mit -1 multipliziert. Bringt man also die Matrix A mit dem Gauß-Algorithmus in Zeilenstufenform und multipliziert man bei jedem Vertauschen von Zeilen mit -1, so erhält man die Determinante von A.

Beispiel 5.3.12: (**Berechnung der Determinante mit dem Gauß-Verfahren**)

Wir berechnen die Determinante einer Matrix, indem wir sie durch Vertauschen von Zeilen und Addieren von Vielfachen von Zeilen zu anderen Zeilen in eine obere Dreiecksmatrix umwandeln:

$$\det\begin{pmatrix} 0 & 1 & -1 & 2 \\ 2 & 4 & 2 & 1 \\ 0 & 0 & 1 & 3 \\ 1 & -1 & 2 & 1 \end{pmatrix} \overset{Z1\leftrightarrow Z4}{=} -\det\begin{pmatrix} 1 & -1 & 2 & 1 \\ 2 & 4 & 2 & 1 \\ 0 & 0 & 1 & 3 \\ 0 & 1 & -1 & 2 \end{pmatrix} \overset{Z2\to Z2-2\cdot Z1}{=} -\det\begin{pmatrix} 1 & -1 & 2 & 1 \\ 0 & 6 & -2 & -1 \\ 0 & 0 & 1 & 3 \\ 0 & 1 & -1 & 2 \end{pmatrix}$$

$$\overset{Z2\leftrightarrow Z4}{=} \det\begin{pmatrix} 1 & -1 & 2 & 1 \\ 0 & 1 & -1 & 2 \\ 0 & 0 & 1 & 3 \\ 0 & 6 & -2 & -1 \end{pmatrix} \overset{Z4\mapsto Z4-6\cdot Z2}{=} \det\begin{pmatrix} 1 & -1 & 2 & 1 \\ 0 & 1 & -1 & 2 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 4 & -13 \end{pmatrix}$$

$$\overset{Z4\mapsto Z4-4\cdot Z3}{=} \det\begin{pmatrix} 1 & -1 & 2 & 1 \\ 0 & 1 & -1 & 2 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & -25 \end{pmatrix} = -25.$$

Oft ist es auch geschickt, eine Kombination mehrerer Aussagen aus Satz 5.3.11 zur Berechnung der Determinanten zu benutzen, also etwa durch Vertauschen von Zeilen oder Addieren von Vielfachen von Zeilen zu anderen Zeilen und Transponieren in Blockform zu bringen und Aussage 7. anzuwenden.

Beispiel 5.3.13: (**Berechnung einer Determinante mit der Blockform**)

Wir berechnen die Determinante einer Matrix, indem wir sie durch Vertauschen von Zeilen, Addieren von Vielfachen von Zeilen zu anderen Zeilen und Transponieren in eine Blockmatrix umwandeln:

$$\det\begin{pmatrix} 1 & -1 & 2 & 0 \\ -2 & 3 & -1 & 1 \\ -5 & 4 & 1 & 0 \\ 2 & 7 & 1 & -1 \end{pmatrix} \overset{Z1\mapsto Z1-2Z_3}{=} \det\begin{pmatrix} 11 & -9 & 0 & 0 \\ -2 & 3 & -1 & 1 \\ -5 & 4 & 1 & 0 \\ 2 & 7 & 1 & -1 \end{pmatrix}$$
$$\overset{Z2\mapsto Z2+Z4}{=} \det\begin{pmatrix} 11 & -9 & 0 & 0 \\ 0 & 10 & 0 & 0 \\ -5 & 4 & 1 & 0 \\ 2 & 7 & 1 & -1 \end{pmatrix} \overset{T}{=} \det\begin{pmatrix} 11 & 0 & -5 & 2 \\ -9 & 10 & 4 & 7 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$
$$\overset{6.}{=} \det\begin{pmatrix} 11 & 0 \\ -9 & 10 \end{pmatrix}\cdot\det\begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix} = (11\cdot 10+9\cdot 0)\cdot(-1\cdot 1-0\cdot 1) = 110\cdot(-1) = -110.$$

Ein weiteres nützliches Verfahren liefert der Entwicklungssatz von Laplace, der die Determinante einer $(n\times n)$-Matrix A auf die Determinanten der $(n-1)\times(n-1)$-Matrizen zurückführt, die durch Entfernen einer Zeile und Spalte aus A entstehen. Er ist besonders dann nützlich, wenn eine Zeile oder Spalte von A viele Nullen enthält.

Satz 5.3.14: (**Entwicklungssatz von Laplace**)
Sei $\mathbb{K}$ ein Körper und $n\in\mathbb{N}$. Dann gilt für $A=(a_{ji})\in \mathrm{Mat}(n\times n,\mathbb{K})$ und alle $i,j\in\{1,...,n\}$

$$\det(A) = \Sigma_{k=1}^n (-1)^{i+k} a_{ki}\det(A_{ki}^{str}) \qquad \text{(Entwicklung nach der } i\text{ten Spalte)}$$
$$\det(A) = \Sigma_{k=1}^n (-1)^{j+k} a_{jk}\det(A_{jk}^{str}) \qquad \text{(Entwicklung nach der } j\text{ten Zeile),}$$

wobei A_{ji}^{str} die (j,i)te-**Streichmatrix** bezeichnet, die aus A durch Entfernen der $i$ten Spalte und $j$ten Zeile entsteht:

$$A_{ji}^{str} = \begin{pmatrix} a_{11} & \cdots & a_{1(i-1)} & a_{1(i+1)} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{(j-1)1} & \cdots & a_{(j-1)(i-1)} & a_{(j-1)(i+1)} & \cdots & a_{(j-1)n} \\ a_{(j+1)1} & \cdots & a_{(j+1)(i-1)} & a_{(j+1)(i+1)} & \cdots & a_{(j+1)n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n(i-1)} & a_{n(i+1)} & \cdots & a_{nn} \end{pmatrix} \in \mathrm{Mat}((n-1)\times(n-1),\mathbb{K}).$$

Die Determinanten der Streichmatrizen bezeichnet man als **Minoren**.

Beweis:
Wir beweisen zunächst die Formel für die Entwicklung nach der $i$ten Spalte. Sei $A = (a_1,...,a_n)\in\mathrm{Mat}(n\times n,\mathbb{K})$ mit Spaltenvektoren $a_i=\Sigma_{k=1}^n a_{ki}e_k$. Dann gilt wegen der Multilinearität und Antisymmetrie der Determinante in den Spaltenvektoren

$$\begin{aligned}\det(A) &= \det(a_1,...,\Sigma_{j=1}^n a_{ji}e_j,...a_n) = \Sigma_{k=1}^n a_{ki}\det(a_1,...,e_k,...,a_n)\\ &= \Sigma_{k=1}^n(-1)^{i-1}a_{ki}\det(e_k,a_1,...,a_{i-1},a_{i+1},...,a_n)\\ &= \Sigma_{k=1}^n(-1)^{i+k-2}a_{ki}\det\begin{pmatrix} 1 & * \\ 0 & A_{ki}^{str}\end{pmatrix} = \Sigma_{k=1}^n(-1)^{i+k}a_{ki}\det(A_{ki}^{str}).\end{aligned}$$

Hier wurde beim Übergang zur zweiten Zeile die Permutation τ mit $\tau(1) = 2$, $\tau(2) = 3$,...,$\tau(i-1) = i$, $\tau(i) = 1$ und $\tau(l) = l$ für $l > i$ mit $\mathrm{sgn}(\tau) = (-1)^{i-1}$ auf die Spalten der Matrix angewendet und im Übergang zur 3. Zeile die Permutation σ mit $\sigma(1) = 2$, $\sigma(2) = 3$,..., $\sigma(k-1) = k$, $\sigma(k) = 1$ und $\sigma(l) = l$ für $l > k$ mit $\mathrm{sgn}(\sigma) = (-1)^{k-1}$ auf die Zeilen der Matrix. Die Formel für die Entwicklung nach der $j$ten Zeile ergibt sich durch Transposition

$$\begin{aligned}\det(A) =& \det(A^T) = \Sigma_{k=1}^n (-1)^{k+j} (A^T)_{kj} \det((A^T)^{str}_{kj}) = \Sigma_{k=1}^n (-1)^{k+j} a_{jk} \det((A^{str}_{jk})^T) \\ =& \Sigma_{k=1}^n (-1)^{k+j} a_{jk} \det(A^{str}_{jk}).\end{aligned}$$

□

Beispiel 5.3.15:

1. Enthält eine Spalte/Zeile einer $(n \times n)$-Matrix A nur einen Eintrag $a_{ki} \neq 0$, dann gilt nach dem Entwicklungssatz von Laplace $\det(A) = (-1)^{i+k} a_{ki} \det(A^{str}_{ki})$. Beispielsweise erhält man aus dem Entwicklungssatz von Laplace:

$$\det \begin{pmatrix} 0 & -1 & 1 \\ 1 & 0 & 0 \\ 0 & 2 & 1 \end{pmatrix} = -\det \begin{pmatrix} -1 & 1 \\ 2 & 1 \end{pmatrix} = -(-1-2) = 3.$$

2. Allgemein ist es sinnvoll, bei der Anwendung des Laplaceschen Entwicklungssatzes nach der Zeile oder Spalte mit der niedrigsten Anzahl von Einträgen $\neq 0$ zu entwickeln, da jeder solche Eintrag einen Summanden liefert. Im folgenden Beispiel ist das die 1. Spalte

$$\det \begin{pmatrix} 1 & 3 & 1 \\ 2 & 2 & 1 \\ 0 & -1 & 1 \end{pmatrix} = \det \begin{pmatrix} 2 & 1 \\ -1 & 1 \end{pmatrix} - 2 \cdot \det \begin{pmatrix} 3 & 1 \\ -1 & 1 \end{pmatrix} = 2 + 1 - 2 \cdot (3+1) = -5.$$

Eine weitere interessante Anwendung der Streichmatrizen ist die *adjungierte Matrix*, die für invertierbare Matrizen in engem Zusammenhang mit der inversen Matrix steht. Sie liefert eine Formel für die inverse Matrix, die diese durch die Determinanten der Streichmatrizen ausdrückt, die *Cramersche Regel*.

Definition 5.3.16: Die **adjungierte Matrix** einer Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{K})$ ist die Matrix $\mathrm{ad}(A)$ mit Einträgen $\mathrm{ad}(A)_{ji} = (-1)^{i+j} \det(A^{str}_{ji})$. Diese bezeichnet man auch als **Kofaktoren**.

Satz 5.3.17: (**Cramersche Regel**)
Für jede Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{K})$ gilt $A \cdot \mathrm{ad}(A)^T = \mathrm{ad}(A)^T \cdot A = \det(A)\, \mathbb{1}_n$. Für invertierbare Matrizen ergibt sich daraus $A^{-1} = \det(A)^{-1}\,\mathrm{ad}(A)^T$.

Beweis:
Wir berechnen zunächst die Einträge auf der Diagonalen der Matrix $A \cdot \mathrm{ad}(A)^T$

$$(A \cdot \mathrm{ad}(A)^T)_{ii} = \Sigma_{k=1}^n a_{ik} \mathrm{ad}(A)_{ik} = \Sigma_{k=1}^n (-1)^{k+i} a_{ik} \det(A^{str}_{ik}) \overset{5.3.14}{=} \det(A).$$

Um zu zeigen, dass die anderen Einträge von $A \cdot \mathrm{ad}(A)^T$ verschwinden, betrachten wir die Matrix $B = (b_{ji})$, die aus A entsteht, indem wir die $j$te Zeile von A durch die $i$te Zeile von A ersetzen, ohne die anderen Zeilen zu verändern. Dann gilt offensichtlich $\det(B) = 0$, da B zwei gleiche Zeilen hat, $b_{jk} = a_{ik}$ und $B^{str}_{jk} = A^{str}_{jk}$ für alle $k \in \{1, ..., n\}$, da hier die einzige Zeile, in der sich A und B unterscheiden, gestrichen wird. Damit erhalten wir

$$\begin{aligned}(A \cdot \mathrm{ad}(A)^T)_{ij} &= \Sigma_{k=1}^n a_{ik} \mathrm{ad}(A)_{jk} = \Sigma_{k=1}^n (-1)^{k+j} a_{ik} \det(A^{str}_{jk}) = \Sigma_{k=1}^n (-1)^{k+j} b_{jk} \det(B^{str}_{jk}) \\ &= (B \cdot \mathrm{ad}(B)^T)_{jj} = \det(B) = 0.\end{aligned}$$

Damit ist gezeigt, dass $A\cdot\text{ad}(A)^T = \det(A)\mathbb{1}_n$ gilt. Der Beweis der Formel $\text{ad}(A)^T\cdot A = \det(A)\,\mathbb{1}_n$ ergibt sich daraus durch Transposition. Ist $A \in \text{GL}(n,\mathbb{K})$, so hat $\det(A)$ ein multiplikatives Inverses, und es ergibt sich $A^{-1} = \det(A)^{-1}\text{ad}(A)^T$. □

Die Cramersche Regel liefert kein besonders schnelles oder effizientes Verfahren zur Berechnung der Inversen einer Matrix, da es sehr aufwändig ist, die Minoren zu berechnen. Im Allgemeinen führt das Gauß-Verfahren hier schneller zum Ziel. Die Cramersche Regel ist aber konzeptionell wichtig, weil sie eine explizite Formel für die inverse Matrix angibt und nicht nur einen Algorithmus zu ihrer Berechnung. Sie zeigt auch, dass sich das Berechnen der inversen Matrix auf Determinanten zurückführen lässt, und dass die Einträge der Inversen Matrix einer Matrix $A \in \text{Mat}(n\times n,\mathbb{K})$ mit Determinante $\det(A) = 1$ Polynome in den Einträgen von A sind.

Da es auch bei Lösen eines eindeutig bestimmten inhomogenen Gleichungssysteme $A\cdot x = b$ mit einer Matrix $A \in \text{GL}(n,\mathbb{K})$ und $b \in \mathbb{K}^n$ nur darum geht, das Inverse von A zu bestimmen, lässt sich die Cramersche Regel insbesondere auf solche inhomogenen Gleichungssysteme anwenden und liefert dann eine Lösungsformel für inhomogene Gleichungssysteme.

Satz 5.3.18: (**Cramersche Regel für lineare Gleichungssysteme**)

Sei $\mathbb{K}$ ein Körper, $n \in \mathbb{N}$ und $a_1, ..., a_n, b \in \mathbb{K}^n$. Ist $A = (a_1, ..., a_n) \in \text{GL}(n,\mathbb{K})$, so ist die eindeutige Lösung des inhomogenen linearen Gleichungssystems $Ax = b$ gegeben durch

$$x = \frac{1}{\det(A)} \begin{pmatrix} \det(b, a_2, ..., a_n) \\ \vdots \\ \det(a_1, ..., a_{n-1}, b) \end{pmatrix}$$

Beweis:
Ist $x = \Sigma_{j=1}^n x_j e_j \in \mathbb{K}^n$ die eindeutige Lösung des Gleichungssystems $Ax = b$, so gilt

$$b = \Sigma_{k=1}^n b_k e_k = \Sigma_{k=1}^n \Sigma_{j=1}^n a_{kj} x_j e_k = \Sigma_{j=1}^n x_j a_j.$$

Daraus ergibt sich für die Matrizen $A_i = (a_1, ..., a_{i-1}, b, a_{i+1}, ..., a_n) \in \text{Mat}(n\times n,\mathbb{K})$

$$\begin{aligned}\det(A_i) &= \det(a_1, ..., a_{i-1}, \Sigma_{j=1}^n x_j a_j, a_{i+1}, ..., a_n) = \Sigma_{j=1}^n x_j \det(a_1, ..., a_{i-1}, a_j, a_{i+1}, ..., a_n) \\ &= x_i \det(A),\end{aligned}$$

denn da für $j \neq i$ die Matrix $(a_1, ..., a_{i-1}, a_j, a_{i+1}, ..., a_n)$ zwei gleiche Zeilen hat, verschwindet für $j \neq i$ ihre Determinante. Da $A \in \text{GL}(n,\mathbb{K})$ ist $\det(A) \neq 0$, und es ergibt sich $x_i = \det(A)^{-1}\det(A_i)$ für alle $i \in \{1, ..., n\}$. □

An was man sich erinnern sollte:

Die wichtigsten Begriffe und Definitionen:

- Permutation, Transposition, elementare Transposition,
- Fehlstandspaar, Signum, gerade und ungerade Permutation, alternierende Gruppe
- multilineare Abbildung, Multilinearform, alternierend, symmetrisch,
- Volumenform, Volumen, Orientierung, orientierungserhaltend, orientierungsumkehrend,
- Determinante von Endomorphismus, Determinante von Matrix,
- adjungierte Matrix, Streichmatrix, Minoren

Die wichtigsten Aussagen:

- Jede Permutation ist ein Produkt von elementaren Vertauschungen.
- Das Signum ist ein Gruppenhomomorphismus.
- Die multilinearen Abbildungen $\phi : V^{\times n} \to W$ bilden einen Untervektorraum $L(V^{\times n}, W) \subseteq \mathrm{Abb}(V^{\times n}, W)$.
- Die alternierenden und die symmetrischen multilinearen Abbildungen bilden Untervektorräume $A(V^{\times n}, W) \subseteq L(V^{\times n}, W)$ und $S(V^{\times n}, W) \subseteq L(V^{\times n}, W)$
- Dimensionen von $L(V^{\times n}, \mathbb{K})$, $A(V^{\times n}, \mathbb{K})$, $S(V^{\times n}, \mathbb{K})$.
- Für einen n-dimensionalen $\mathbb{K}$-Vektorraum V ist $A(V^{\times n}, \mathbb{K})$ eindimensional.
- Existenz und Eindeutigkeit der Volumenform auf dem $\mathbb{K}^n$.
- $\mathrm{vol}_n(v_1, ..., v_n) = 0$ genau dann, wenn $(v_1, ..., v_n)$ linear abhängig.
- Berechnung der Determinante eines Endomorphismus mit der beschreibenden Matrix.
- Eigenschaften der Determinante, Rechenregeln für Determinanten.
- Berechnung der Determinante mit dem Gauß-Verfahren.
- Hauptdiagonale minus Gegendiagonale, Regel von Sarrus.
- Leibnizsche Formel, Laplacescher Entwicklungssatz, Cramersche Regel(n).

Die wichtigsten Beispiele:

- **multilineare Abbildungen:** euklidisches Skalarprodukt, Vektorprodukt, Volumenform.
- **Permutationen:** (elementare) Transpositionen, zyklische Vertauschungen.
- **Determinanten** Spiegelungen und Drehungen, $(n \times n)$-Matrizen vom Rang $< n$, obere Dreiecksmatrizen, Blockmatrizen.

6 Eigenwerte und Eigenvektoren

6.1 Eigenwerte, Eigenvektoren und Eigenräume

Mit der Determinante haben wir ein weiteres Werkzeug gewonnen, um die Frage der Ähnlichkeit von Matrizen zu untersuchen. Da ähnliche Matrizen die gleiche Determinante besitzen, können Matrizen mit verschiedenen Determinanten nie ähnlich sein.

Das Ziel ist es nun, eine *Normalform* zu finden, eine Liste von Matrizen aus $\mathrm{Mat}(n \times n, \mathbb{K})$, so dass jede Matrix aus $\mathrm{Mat}(n \times n, \mathbb{K})$ ähnlich zu genau einer Matrix aus dieser Liste ist. Das analoge Problem für die *Äquivalenz* von Matrizen wurde bereits in Satz 4.3.4 gelöst, wo gezeigt wurde, dass jede Matrix $A \in \mathrm{Mat}(m \times n, \mathbb{K})$ äquivalent zu einer Matrix der Form $E^r_{m\times n}$ mit $r = \mathrm{rg}(A)$ ist. Damit sind zwei Matrizen äquivalent genau dann, wenn sie den gleichen Rang haben. Das Problem der Ähnlichkeit ist deutlich schwerer, und wir werden es nur über bestimmten Körpern vollständig lösen können.

Ein wichtiger Schritt dahin sind die *Eigenvektoren* und *Eigenwerte*. Die Eigenvektoren eines Endomorphismus $\phi : V \to V$ sind die Vektoren $v \in V \setminus \{0\}$, die von ϕ auf skalare Vielfache $\phi(v) = \lambda v$ von sich selbst abgebildet werden. Die zugehörigen Skalare $\lambda \in \mathbb{K}$ nennt man Eigenwerte. Geometrisch bedeutet dies, dass der Endomorphismus die Gerade $\mathbb{K}v$ auf sich selbst abbildet, sie dabei aber strecken oder stauchen kann.

Wir versuchen nun, einen Endomorphismus $\phi : V \to V$ so weit wie möglich durch diese Fixgeraden $\mathbb{K}v$ und die zugehörigen Streckungsfaktoren $\lambda \in \mathbb{K}$ zu beschreiben. Dies hilft einerseits bei der Lösung des Ähnlichkeitsproblems. Andererseits können wir so Endomorphismen besser visualisieren und geometrisch interpretieren.

Definition 6.1.1: Sei V ein Vektorraum über $\mathbb{K}$ und $\phi : V \to V$ eine lineare Abbildung.

1. Eine Zahl $\lambda \in \mathbb{K}$ heißt **Eigenwert** von ϕ, wenn es ein $v \in V \setminus \{0\}$ gibt mit $\phi(v) = \lambda v$.
2. Den Vektor $v \in V \setminus \{0\}$ bezeichnet man dann als **Eigenvektor** von ϕ zum Eigenwert λ.
3. Der **Eigenraum** zum Eigenwert λ ist $E(\phi, \lambda) = \ker(\phi - \lambda\mathrm{id}_V) = \{v \in V | \phi(v) = \lambda v\} \subseteq V$.
4. Die **geometrische Vielfachheit** des Eigenwerts λ ist die Dimension des Eigenraums: $g(\phi, \lambda) = \dim_{\mathbb{K}} E(\phi, \lambda) = \mathrm{def}(\phi - \lambda\mathrm{id}_V)$.

Bemerkung 6.1.2:

1. Nach Definition 6.1.1 kann zwar $0 \in \mathbb{K}$ ein Eigenwert von ϕ sein, aber der Nullvektor $0 \in V$ ist *nie* ein Eigenvektor von ϕ.
2. Da $E(\phi, \lambda) = \ker(\phi - \lambda\mathrm{id}_V)$ gilt, ist der Eigenraum $E(\lambda, \phi)$ ein Untervektorraum von V. Der Nullvektor $0 \in V$ ist daher in jedem Eigenraum $E(\phi, \lambda)$ von ϕ enthalten.
3. Für einen Skalar $\lambda \in \mathbb{K}$ sind äquivalent:
   (i) λ ist ein Eigenwert von ϕ,
   (ii) Die Gleichung $(\phi - \lambda\mathrm{id}_V)(v) = 0$ hat eine nichttriviale Lösung $v \in V \setminus \{0\}$,
   (iii) $\ker(\phi - \lambda\mathrm{id}_V) \neq \{0\}$.

Beispiel 6.1.3:

1. Der Skalar $\alpha \in \mathbb{K}$ ist der einzige Eigenwert der **Streckung** $D_\alpha : V \to V$, $v \mapsto \alpha v$, und jeder Vektor $v \in V \setminus \{0\}$ ist ein Eigenvektor von D_α zum Eigenwert α.

Insbesondere besitzt die Identitätsabbildung $\mathrm{id}_V : V \to V$ als einzigen Eigenwert 1, und jeder Vektor $v \in V \setminus \{0\}$ ist ein Eigenvektor zum Eigenwert 1.

2. Ist $\phi : V \to V$ ein **Projektor**, so sind 0 und 1 die einzig möglichen Eigenwerte von ϕ.

   Denn es gilt $\phi \circ \phi = \phi$, und daraus folgt für jeden Eigenvektor $v \in V$ zum Eigenwert λ die Gleichung $\lambda^2 v = \lambda\phi(v) = \phi(\lambda v) = (\phi \circ \phi)(v) = \phi(\lambda v) = \lambda v$. Daraus ergibt sich $\lambda(\lambda - 1)v = 0$ und wegen $v \neq 0$ dann $\lambda \in \{0, 1\}$.

3. Ist $\phi : V \to V$ eine **Involution**, d. h. es gilt $\phi \circ \phi = \mathrm{id}_V$, so sind 1 und -1 die einzig möglichen Eigenwerte von ϕ.
   Insbesondere sind die einzig möglichen Eigenwerte einer **Spiegelung** an einer Hyperebene im $\mathbb{R}^n$ die Werte 1 und -1.

   Denn für jeden Eigenvektor $v \in V$ zum Eigenwert λ gilt $\lambda^2 v = \lambda\phi(v) = \phi(\lambda v) = (\phi \circ \phi)(v) = \mathrm{id}_V(v) = v$ und wegen $v \neq 0$ dann $0 = \lambda^2 - 1 = (\lambda + 1)(\lambda - 1)$.

4. Ist $\phi : V \to V$ eine lineare Abbildung **endlicher Ordnung**, d. h. es gibt ein $k \in \mathbb{N}$ mit $\phi^k = \phi \circ \ldots \circ \phi = \mathrm{id}_V$, so sind alle Eigenwerte von ϕ $k$te Einheitswurzeln.

   Denn für jeden Eigenvektor $v \in V$ zum Eigenwert λ gilt $\lambda^k v = \phi^k(v) = \mathrm{id}_V(v) = v$ und wegen $v \neq 0$ damit $\lambda^k = 1$. Insbesondere sind die komplexen Eigenwerte von Permutationsmatrizen immer Einheitswurzeln (Übung).

5. Ist $\phi : V \to V$ ein **nilpotenter Endomorphismus**, d. h. es existiert ein $n \in \mathbb{N}$ mit $\phi^n = \phi \circ \ldots \circ \phi = 0$, so ist 0 der einzig mögliche Eigenwert von ϕ.

   Denn für einen Eigenvektor $v \in V$ zum Eigenwert λ gilt $0 = \phi^n(v) = \phi^{n-1}(\lambda v) = \ldots = \lambda^{n-1}\phi(v) = \lambda^n v$. Da $v \neq 0$ folgt daraus $\lambda^n = 0$ und, da $\mathbb{K}$ ein Körper ist, $\lambda = 0$.

An den Beispielen in Beispiel 6.1.3 sieht man, dass die Eigenwerte viel über die geometrischen und algebraischen Eigenschaften von Endomorphismen aussagen. Leider lassen sie sich nicht für alle Endomorphismen so einfach berechnen. Deswegen ist es im endlich-dimensionalen Fall oft sinnvoll, eine Basis zu wählen und mit beschreibenden Matrizen zu arbeiten.

Das Prinzip zur Definition von Eigenwerten, Eigenvektoren und geometrischen Vielfachheiten für Matrizen ist das gleiche wie bei der Definition ihres Kerns, Bilds, Defekts und Rangs in Definition 4.2.14. Man betrachtet dabei für $A \in \mathrm{Mat}(n \times n, \mathbb{K})$ den Endomorphismus $\phi_A : \mathbb{K}^n \to \mathbb{K}^n$, $v \mapsto A \cdot v$ und definiert Eigenwerte, Eigenvektoren und Eigenräume von A als die Eigenwerte, Eigenvektoren und Eigenräume von ϕ_A.

Definition 6.1.4: Die **Eigenwerte**, **Eigenvektoren** und **Eigenräume** einer quadratischen Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{K})$ sind die Eigenwerte, Eigenvektoren und Eigenräume des zugehörigen Endomorphismus $\phi_A : \mathbb{K}^n \to \mathbb{K}^n$, $v \mapsto A \cdot v$.

Zur Berechnung von Eigenvektoren einer Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{K})$ zu einem gegebenen Eigenwert $\lambda \in \mathbb{K}$ muss man den Lösungsraums des homogenen Gleichungssystems $(A - \lambda \mathbb{1}_n) \cdot v = 0$ bestimmen, also den Kern $\ker(A - \lambda \mathbb{1}_n)$.

Möchte man nur wissen, ob $\lambda \in \mathbb{K}$ ein Eigenwert ist oder nicht, reicht es, festzustellen, ob der Rang $A - \lambda \mathbb{1}_n$ gleich n ist. Dazu reicht im endlich-dimensionalen Fall die Berechnung der Determinante. Da die Determinante jedes Endomorphismus $\phi : V \to V$ gleich der Determinante seiner beschreibenden Matrizen ist, reduziert sich dies dann auf Matrixrechnungen.

Lemma 6.1.5: Sei $\mathbb{K}$ ein Körper und V ein n-dimensionaler Vektorraum über $\mathbb{K}$.

1. Ein Skalar $\lambda \in \mathbb{K}$ ist genau dann ein Eigenwert eines Endomorphismus $\phi : V \to V$, wenn $\det(\phi - \lambda \mathrm{id}_V) = 0$ gilt.
2. Ein Skalar $\lambda \in \mathbb{K}$ ist genau dann ein Eigenwert einer Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{K})$, wenn $\det(A - \lambda \mathbb{1}_n) = 0$ gilt.
3. Die Eigenwerte von ϕ sind genau die Eigenwerte der darstellenden Matrix ${}_B M_B(\phi)$ bezüglich jeder geordneten Basis B von V.

Die Gleichungen $\det(\phi - \lambda \mathrm{id}_V) = 0$ und $\det(A - \lambda \mathbb{1}_n) = 0$ bezeichnet man als die **charakteristische Gleichung** von ϕ und von A.

Beweis:
1. Nach Bemerkung 6.1.2, 3. (iii) ist $\lambda \in \mathbb{K}$ genau dann ein Eigenwert von $\phi : V \to V$, wenn $\ker(\phi - \lambda \mathbb{1}_n) \neq \{0\}$ gilt. Dies ist genau dann der Fall, wenn $\mathrm{rg}(\phi - \lambda \mathbb{1}_n) = n - \mathrm{def}(\phi - \lambda \mathbb{1}_n) < n$, was nach Satz 5.3.11, 8. äquivalent ist zu $\det(\phi - \lambda \mathrm{id}_V) = 0$.

2. Daraus folgt für jede Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{K})$, dass $\lambda \in \mathbb{K}$ ein Eigenwert von A ist genau dann, wenn $\det(A - \lambda \mathbb{1}_n) = \det(\phi_{A - \lambda \mathbb{1}_n}) = \det(\phi_A - \lambda \mathrm{id}_{\mathbb{K}^n}) = 0$ gilt.

3. Nach Korollar 5.3.6 und Satz 4.2.12 gilt für jede geordnete Basis B von V

$$\det(\phi - \lambda \mathrm{id}_V) \overset{5.3.6}{=} \det({}_B M_B(\phi - \lambda \mathrm{id}_V)) \overset{(4.2.12)}{=} \det({}_B M_B(\phi) - \lambda \mathbb{1}_n).$$

Da nach 1. $\lambda \in \mathbb{K}$ genau dann ein Eigenwert von ϕ ist, wenn $\det(\phi - \lambda \mathrm{id}_V) = 0$ gilt und genau dann ein Eigenwert von ${}_B M_B(\phi)$, wenn $\det({}_B M_B(\phi) - \lambda \mathbb{1}_n) = 0$ ist, folgt die Behauptung. $\square$

Korollar 6.1.6: Ähnliche Matrizen haben die gleichen Eigenwerte mit den gleichen geometrischen Vielfachheiten.

Beweis:
Sind $A, B \in \mathrm{Mat}(n \times n, \mathbb{K})$ ähnlich, so existiert eine Matrix $S \in \mathrm{GL}(n, \mathbb{K})$ mit $A = S \cdot B \cdot S^{-1}$, und es folgt $\det(A - \lambda \mathbb{1}_n) = \det(S \cdot B \cdot S^{-1} - \lambda \mathbb{1}_n) = \det(S \cdot (B - \lambda \mathbb{1}_n) \cdot S^{-1}) = \det(B - \lambda \mathbb{1}_n)$. Also haben A und B nach Lemma 6.1.5 die gleichen Eigenwerte. Ebenso ergibt sich für die geometrischen Vielfachheiten $g(A, \lambda) = \mathrm{def}(A - \lambda \mathbb{1}_n) = \mathrm{def}(S \cdot (B - \lambda \mathbb{1}_n) \cdot S^{-1}) = \mathrm{def}(B - \lambda \mathbb{1}_n) = g(B, \lambda)$. $\square$

Mit diesen Aussagen können wir auch für kompliziertere Endomorphismen $\phi : V \to V$ als die in Beispiel 6.1.3 die Eigenwerte und Eigenvektoren bestimmen. Dazu wählt man zunächst eine geordnete Basis B von V, berechnet dann die beschreibende Matrix ${}_B M_B(\phi)$ und bestimmt dann die Eigenwerte als Lösungen der charakteristischen Gleichung $\det({}_B M_B(\phi) - \lambda \mathbb{1}) = 0$. Zur Berechnung der Eigenvektoren löst man dann für jeden Eigenwert λ das lineare Gleichungssystem ${}_B M_B(\phi) x = \lambda x$ oder, dazu äquivalent, bestimmt den Kern von ${}_B M_B(\phi) - \lambda \mathbb{1}_n$.

Beispiel 6.1.7: Die **Drehung** im $\mathbb{R}^2$ um einen Winkel $\alpha \in \mathbb{R}$

$$R_\alpha : \mathbb{R}^2 \to \mathbb{R}^2, \quad \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x \cos\alpha - y \sin\alpha \\ x \sin\alpha + y \cos\alpha \end{pmatrix} = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix}$$

hat nur für $\alpha \in \pi\mathbb{Z}$ reelle Eigenwerte und Eigenvektoren. Denn die quadratische Gleichung

$$\begin{aligned}0 = \det\begin{pmatrix} \cos\alpha - \lambda & -\sin\alpha \\ \sin\alpha & \cos\alpha - \lambda \end{pmatrix} &= (\cos(\alpha) - \lambda)^2 + \sin(\alpha)^2 \\ = \cos(\alpha)^2 + \sin(\alpha)^2 + \lambda^2 - 2\lambda\cos(\alpha) &= \lambda^2 - 2\lambda\cos(\alpha) + 1\end{aligned}$$

hat nur für $\cos(\alpha)^2 - 1 \geq 0$ eine reelle Lösung. Also hat R_α genau dann reelle Eigenwerte λ, wenn $|\cos(\alpha)| = 1$ gilt, also wenn $\alpha \in \pi\mathbb{Z}$. Für $\alpha = n\pi$ mit n gerade ist $R_\alpha = \mathrm{id}_{\mathbb{R}^2}$, und R_α besitzt den Eigenwert 1. Für $\alpha = n\pi$ mit n ungerade ist $R_\alpha = -\mathrm{id}_{\mathbb{R}^2}$ und R_α besitzt den reellen Eigenwert -1. In beiden Fällen gilt $E(R_\alpha, \lambda) = \mathbb{R}^2$.

Betrachten wir die zugehörige $\mathbb{C}$-lineare Abbildung

$$R'_\alpha : \mathbb{C}^2 \to \mathbb{C}^2, \quad \begin{pmatrix} z \\ w \end{pmatrix} \mapsto \begin{pmatrix} z\cos\alpha - w\sin\alpha \\ z\sin\alpha + w\cos\alpha \end{pmatrix} = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \cdot \begin{pmatrix} z \\ w \end{pmatrix}$$

so sind die Eigenwerte die *komplexen* Lösungen der Gleichung $\lambda^2 - 2\lambda\cos(\alpha) + 1 = 0$

$$\lambda_\pm = \cos(\alpha) \pm \sqrt{\cos(\alpha)^2 - 1} = \cos(\alpha) \pm \mathrm{i}\sqrt{1 - \cos(\alpha)^2} = \cos(\alpha) \pm \mathrm{i}\sin(\alpha) = e^{\pm\mathrm{i}\alpha}.$$

Man erhält also zwei komplexe Eigenwerte, die genau für $\alpha \in \pi\mathbb{Z}$ zu einem einzigen *reellen* Eigenwert zusammenfallen. Für $\alpha \in \pi\mathbb{Z}$ ist $E(R'_\alpha, e^{\pm i\alpha}) = \mathbb{C}^2$. Die zugehörigen Eigenvektoren für $\alpha \notin \pi\mathbb{Z}$ ergeben sich aus den Gleichungen

$$\begin{aligned} R'_\alpha\begin{pmatrix} z \\ w \end{pmatrix} = \begin{pmatrix} z\cos\alpha - w\sin\alpha \\ z\sin\alpha + w\cos\alpha \end{pmatrix} &= (\cos\alpha \pm \mathrm{i}\sin\alpha)\begin{pmatrix} z \\ w \end{pmatrix} \\ \Leftrightarrow \quad -w\sin\alpha = \pm\mathrm{i}z\sin\alpha \wedge z\sin\alpha = \pm\mathrm{i}w\sin\alpha \quad &\Leftrightarrow \quad w = \mp\mathrm{i}z. \end{aligned}$$

Also gilt für $\alpha \notin \pi\mathbb{Z}$:

$$E(R'_\alpha, e^{\mathrm{i}\alpha}) = \mathbb{C}(e_1 - \mathrm{i}e_2) \quad E(R'_\alpha, e^{-\mathrm{i}\alpha}) = \mathbb{C}(e_1 + \mathrm{i}e_2) \;\Rightarrow\; \mathbb{C}^2 = E(R'_\alpha, e^{\mathrm{i}\alpha}) \oplus E(R'_\alpha, e^{-\mathrm{i}\alpha}).$$

Beispiel 6.1.8: Wir bestimmen die Eigenwerte und Eigenvektoren der Matrix

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in \mathrm{Mat}(2 \times 2, \mathbb{K}).$$

Die charakteristische Gleichung lautet

$$\det(A - \lambda\mathbb{1}_2) = \det\begin{pmatrix} 1-\lambda & 1 \\ 0 & 1-\lambda \end{pmatrix} = (1-\lambda)^2 = 0 \quad \Leftrightarrow \quad \lambda = 1.$$

Also besitzt A nur den Eigenwert $\lambda = 1$. Für die Eigenvektoren ergibt sich

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x+y \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} \quad \Leftrightarrow \quad y = 0.$$

Also gilt $E(A, 1) = \mathbb{K}e_1$. Der Vektorraum $\mathbb{K}^2$ lässt sich also nicht als direkte Summe von Eigenräumen der Matrix A schreiben.

Die Beispiele zeigen, dass eine Matrix oder ein Endomorphismus keine Eigenwerte oder Eigenvektoren haben muss. Aus Beispiel 6.1.7 ergibt sich auch, dass die Existenz von Eigenwerten von dem betrachteten Körper abhängt. Ist nämlich $\mathbb{K} \subseteq \mathbb{L}$ ein Teilkörper (z. B. $\mathbb{K} = \mathbb{R}$, $\mathbb{L} = \mathbb{C}$ oder $\mathbb{K} = \mathbb{Q}$, $\mathbb{L} = \mathbb{R}$), so kann man eine Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{K})$ auch als eine Matrix in $\mathrm{Mat}(n \times n, \mathbb{L})$ auffassen und die Eigenwerte und Eigenvektoren in $\mathbb{K}$ und in $\mathbb{L}$ bestimmen. Jeder Eigenwert $\lambda \in \mathbb{K}$ ist auch ein Eigenwert in $\mathbb{L}$ und jeder Eigenvektor $v \in \mathbb{K}^n$ zum Eigenwert $\lambda \in \mathbb{K}$ ist auch ein Eigenvektor in $\mathbb{L}^n$. Beispiel 6.1.7 zeigt aber, dass die Umkehrung nicht gilt.

Zu einem gegebenen Eigenwert λ muss der zugehörige Eigenraum mindestens eindimensional sein, aber im Allgemeinen können wir keine weiteren Aussagen über dessen Dimension machen. Besonders einfach wird es, wenn der Vektorraum eine Basis aus Eigenvektoren besitzt. Beispiel 6.1.8 zeigt jedoch, dass es über jedem Körper Vektorräume V und Endomorphismen $\phi : V \to V$ gibt, die zwar Eigenwerte besitzen, jedoch keine Basis aus Eigenvektoren von ϕ.

6.2 Diagonalisierbarkeit und Trigonalisierbarkeit

Wir untersuchen nun, unter welchen Bedingungen zu einem Endomorphismus $\phi : V \to V$ eines n-dimensionalen $\mathbb{K}$-Vektorraums V eine Basis $B = (v_1, ..., v_n)$ aus Eigenvektoren von ϕ existiert. Dies ist äquivalent dazu, dass die beschreibende Matrix ${}_BM_B(\phi)$ eine Diagonalmatrix ist. Denn beide Bedingungen bedeuten, dass es $\lambda_1, ..., \lambda_n \in \mathbb{K}$ gibt mit $\phi(v_i) = \lambda_i v_i$ für alle $i \in \{1, ..., n\}$.

Da es nach Beispiel 6.1.8 allerdings über jedem Körper Vektorräume V und Endomorphismen gibt, für die Eigenvektoren von ϕ keine Basis bilden, untersuchen wir auch noch eine schwächere Bedingung, nämlich die Existenz einer geordneten Basis B, für die die beschreibende Matrix ${}_BM_B(\phi)$ zumindest eine obere Dreiecksmatrix ist wie in Beispiel 6.1.8.

Definition 6.2.1:

1. Ein Endomorphismus $\phi \in \mathrm{End}_{\mathbb{K}}(V)$ eines $\mathbb{K}$-Vektorraums V heißt **diagonalisierbar**, wenn V eine Basis aus Eigenvektoren von ϕ besitzt.
2. Eine Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{K})$ heißt **diagonalisierbar**, wenn sie ähnlich ist zu einer **Diagonalmatrix**, d. h. zu einer Matrix $B = (b_{ij}) \in \mathrm{Mat}(n \times n, \mathbb{K})$ mit $b_{ij} = 0$ für $i \neq j$.
3. Ein Endomorphismus $\phi \in \mathrm{End}_{\mathbb{K}}(V)$ eines $\mathbb{K}$-Vektorraums V heißt **trigonalisierbar**, wenn V eine geordnete Basis B besitzt, so dass ${}_BM_B(\phi)$ eine obere Dreiecksmatrix ist.
4. Eine Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{K})$ heißt **trigonalisierbar** wenn sie ähnlich ist zu einer **oberen Dreiecksmatrix**.

Definition 6.2.1 legt nahe, dass es einen Zusammenhang zwischen der Diagonalisierbarkeit oder Trigonalisierbarkeit eines Endomorphismus $\phi : V \to V$ und der Diagonalisierbarkeit oder Trigonalisierbarkeit seiner beschreibenden Matrizen gibt. Dies erlaubt es einem, bei der Untersuchung der Diagonalisierbarkeit oder Trigonalisierbarkeit von Endomorphismen endlich-dimensionaler Vektorräume sowohl mit Endomorphismen als auch mit beschreibenden Matrizen arbeiten.

Satz 6.2.2: Sei V ein n-dimensionaler $\mathbb{K}$-Vektorraum und $\phi \in \mathrm{End}_{\mathbb{K}}(V)$. Dann gilt:

1. Ist ϕ diagonalisierbar (trigonalisierbar), so ist auch die beschreibende Matrix ${}_BM_B(\phi)$ bezüglich jeder geordneten Basis B von V diagonalisierbar (trigonalisierbar).

2. Existiert eine geordnete Basis B von V, so dass die beschreibende Matrix ${}_BM_B(\phi)$ diagonalisierbar (trigonalisierbar) ist, so ist auch ϕ diagonalisierbar (trigonalisierbar).

Beweis:
1. Ist ϕ diagonalisierbar (trigonalisierbar), so existiert eine geordnete Basis C von V, so dass die beschreibende Matrix ${}_CM_C(\phi)$ eine Diagonalmatrix (obere Dreiecksmatrix) ist. Ist B eine beliebige geordnete Basis von V, so sind die beschreibenden Matrizen ${}_BM_B(\phi)$ und ${}_CM_C(\phi)$ nach Satz 4.3.3 ähnlich und damit ist ${}_BM_B(\phi)$ diagonalisierbar (trigonalisierbar).

2. Existiert umgekehrt eine geordnete Basis $B = (v_1, ..., v_n)$ von V, so dass die beschreibende Matrix ${}_BM_B(\phi)$ diagonalisierbar (trigonalisierbar) ist, so gibt es eine Diagonalmatrix (obere Dreiecksmatrix) $D \in \text{Mat}(n \times n, \mathbb{K})$ und eine Matrix $S \in \text{GL}(n, \mathbb{K})$ mit ${}_BM_B(\phi) = S \cdot D \cdot S^{-1}$. Da $S \in \text{GL}(n, \mathbb{K})$, ist auch $C = (w_1, ..., w_n)$ mit $w_i = \Sigma_{j=1}^n s_{ji} v_j$ eine Basis von V, und die Basiswechselmatrix ist ${}_BM_C(\text{id}_V) = S$. Daraus folgt

$$
{}_CM_C(\phi) = {}_BM_C(\text{id}_V)^{-1} \cdot {}_BM_B(\phi) \cdot {}_BM_C(\text{id}_V) = S^{-1} \cdot {}_BM_B(\phi) \cdot S = D.
$$

Die beschreibende Matrix von ϕ bezüglich C ist also die Diagonalmatrix (obere Dreiecksmatrix) D, und damit ist ϕ diagonalisierbar (trigonalisierbar). □

Bemerkung 6.2.3:

1. Diagonalisierbarkeit impliziert Trigonalisierbarkeit, aber die Umkehrung gilt nicht (siehe Beispiel 6.1.8).
2. Jeder trigonalisierbare Endomorphismus $\phi : V \to V$ besitzt mindestens einen Eigenwert. Denn ist $B = (v_1, ..., v_n)$ eine Basis, so dass die darstellende Matrix ${}_BM_B(\phi)$ eine obere Dreiecksmatrix ist, so folgt $\phi(v_1) = \lambda v_1$ für ein $\lambda \in \mathbb{K}$.
3. Ist $A = (a_{ji}) \in \text{Mat}(n \times n, \mathbb{K})$ eine obere Dreiecksmatrix, so sind die Eigenwerte von A genau die Diagonaleinträge a_{ii}. Denn die Determinante einer oberen Dreiecksmatrix ist nach Satz 5.3.11, 5. das Produkt der Diagonaleinträge. Damit verschwindet die Determinante der oberen Dreiecksmatrix $A - \lambda \mathbb{1}_n$ genau dann, wenn $\lambda = a_{ii}$ für ein $i \in \{1, ..., n\}$.

Beispiel 6.2.4:

1. Jede **Streckung** $\phi : V \to V$, $v \mapsto \alpha v$ mit $\alpha \in \mathbb{K}$ ist diagonalisierbar, denn jeder Vektor $v \in V \setminus \{0\}$ ist ein Eigenvektor von ϕ zum Eigenwert α. Insbesondere gilt das für die Identitätsabbildung $\text{id}_V : V \to V$.

2. Die **Drehung** $R_\alpha : \mathbb{R}^2 \to \mathbb{R}^2$ ist nach Beispiel 6.1.7 nur für $\alpha \in \pi\mathbb{Z}$ trigonalisierbar und in diesem Fall auch diagonalisierbar. Ansonsten hat sie keine Eigenwerte, aber jede trigonalisierbare Matrix muss nach Bemerkung 6.2.3 mindestens einen Eigenwert besitzen. Die zugehörige Abbildung $R'_\alpha : \mathbb{C}^2 \to \mathbb{C}^2$ ist für beliebige $\alpha \in \mathbb{R}$ diagonalisierbar.

3. Der Endomorphismus $\phi_A : \mathbb{K}^2 \to \mathbb{K}^2$, $v \mapsto A \cdot v$ mit A aus Beispiel 6.1.8 ist nicht diagonalisierbar, denn er hat nur den Eigenwert $\lambda = 1$ und $E(\phi, \lambda) = \mathbb{K}e_1 \neq \mathbb{K}^2$. Also gibt es keine Basis des $\mathbb{K}^2$ aus Eigenvektoren von ϕ.

4. Ist $\phi \in \text{End}_\mathbb{K}(V)$ **nilpotent** und $\phi \neq 0$, dann ist ϕ nicht diagonalisierbar.

   Denn nach Beispiel 6.1.3, 5. ist 0 der einzige Eigenwert von ϕ. Gäbe es eine Basis B von V aus Eigenvektoren von ϕ, so würde gelten $\ker(\phi) = \text{span}_\mathbb{K}(B) = V$, also $\phi = 0$.

5. Ist $\phi \in \mathrm{End}_{\mathbb{K}}(V)$ diagonalisierbar, so sind auch die linearen Abbildungen $\lambda\phi$, $\phi + \lambda\mathrm{id}_V$, ϕ^n für alle $\lambda \in \mathbb{K}$ und $n \in \mathbb{N}$ diagonalisierbar, und falls ϕ invertierbar ist, auch ϕ^{-1}.

   Denn ist $v \in V\setminus\{0\}$ ein Eigenvektor von ϕ zum Eigenwert μ, so ist v auch ein Eigenvektor von $\lambda\phi$ zum Eigenwert $\lambda\mu$, von $\phi+\lambda\mathrm{id}_V$ zum Eigenwert $\lambda+\mu$ und von ϕ^n zum Eigenwert μ^n. Ist ϕ invertierbar, so ist $\mu \neq 0$ und v ist ein Eigenvektor von ϕ^{-1} zum Eigenwert μ^{-1}.

6. Falls $\mathrm{char}(\mathbb{K}) \neq 2$ ist jede **Involution** $I : V \to V$ eines $\mathbb{K}$-Vektorraums V diagonalisierbar.

   Denn aus $I \circ I = \mathrm{id}_V$ folgt, dass sich jeder Vektor $v \in V$ schreiben lässt als $v = v_+ + v_-$ mit $v_\pm = \frac{1}{2}(v \pm I(v))$, und $I(v_+) = v_+$ und $I(v_-) = -v_-$. Damit ist jeder Vektor eine Summe von Eigenvektoren von I, und V besitzt ein Erzeugendensystem und damit auch eine Basis aus Eigenvektoren.

7. Jeder **Projektor** $P : V \to V$ eines n-dimensionalen $\mathbb{K}$-Vektorraums V ist diagonalisierbar.

   Denn nach Satz 4.3.8 ist P genau dann ein Projektor, wenn seine darstellende Matrix bezüglich jeder geordneten Basis B von V ähnlich zu einer Matrix der Form $E^r_{n\times n}$ mit $r = \mathrm{rg}(\phi)$ ist. Diese ist offensichtlich eine Diagonalmatrix.

Das letzte Beispiel wirft die Frage auf, ob sich diagonalisierbare Endomorphismen durch Projektoren beschrieben lassen. Offensichtlich ist die Bedingung, diagonalisierbar zu sein schwächer als die Bedingung, ein Projektor zu sein. So ist jedes skalare Vielfache $\lambda P : V \to V$ eines Projektors $P : V \to V$ nach Beispiel 6.2.4, 5. ebenfalls diagonalisierbar, für $\lambda \in \mathbb{K}\setminus\{0,1\}$ und $P \neq 0$ aber kein Projektor. Allgemein sind diagonalisierbare Endomorphismen Linearkombinationen kommutierender Projektoren, die sich zur Identitätsabbildung summieren.

Satz 6.2.5: Ein Endomorphismus $\phi : V \to V$ eines n-dimensionalen $\mathbb{K}$-Vektorraums V ist diagonalisierbar genau dann, wenn es Skalare $\lambda_1, ..., \lambda_m \in \mathbb{K}$ und Projektoren $P_1, ..., P_m : V \to V$ mit $P_iP_j = 0$ für $i \neq j$ und $P_1 + ... + P_m = \mathrm{id}_V$ gibt, so dass $\phi = \lambda_1 P_1 + ... + \lambda_m P_m$.

Beweis:
$\Rightarrow$: Ist $\phi : V \to V$ diagonalisierbar, so besitzt V eine geordnete Basis $B = (v_1, ..., v_n)$ aus Eigenvektoren von ϕ. Also gilt $V = \mathbb{K}v_1 \oplus ... \oplus \mathbb{K}v_n$, und es gibt $\lambda_1, .., \lambda_n \in \mathbb{K}$ mit $\phi(v_i) = \lambda_i v_i$ für $i \in \{1, ..., n\}$. Die linearen Abbildungen $P_i : V \to V$, $\Sigma_{j=1}^n \mu_j v_j \mapsto \mu_i v_i$ sind Projektoren mit $\mathrm{im}(P_i) = \mathbb{K}v_i$. Außerdem ergibt sich für alle $\mu_1, ..., \mu_n \in \mathbb{K}$

$$\begin{aligned}
&P_iP_j(\Sigma_{k=1}^n \mu_k v_k) = P_i(\mu_j v_j) = 0 && \Rightarrow \quad P_iP_j = 0 \text{ für } i \neq j,\\
&(P_1 + ... + P_n)(\Sigma_{k=1}^n \mu_k v_k) = \mu_1 v_1 + ... + \mu_n v_n = \Sigma_{k=1}^n \mu_k v_k && \Rightarrow \quad P_1 + ... + P_n = \mathrm{id}_V,\\
&\phi(\Sigma_{i=1}^n \mu_i v_i) = \Sigma_{i=1}^n \mu_i \lambda_i v_i = \Sigma_{j=1}^n \lambda_j P_j(\Sigma_{i=1}^n \mu_i v_i) && \Rightarrow \quad \phi = \Sigma_{j=1}^n \lambda_j P_j.
\end{aligned}$$

$\Leftarrow$: Seien nun $P_1, .., P_m : V \to V$ Projektoren mit $P_iP_j = 0$ für $i \neq j$ und $P_1 + ... + P_m = \mathrm{id}_V$. Dann folgt aus $P_iP_j = 0$ für $i \neq j$, dass $\mathrm{im}(P_j) \subseteq \ker(P_i)$ und $\mathrm{im}(P_i) \cap \mathrm{im}(P_j) = \{0\}$ für alle $i \neq j$. Da $P_1 + ... + P_m = \mathrm{id}_V$, gilt $V = \mathrm{im}(P_1) + ... + \mathrm{im}(P_m)$.

Wir zeigen, dass die Summe direkt ist. Ist $v = v_1 + \ldots + v_m = v_1' + \ldots + v_m'$ mit $v_i, v_i' \in \mathrm{im}(P_i)$, so folgt $0 = (v_1 - v_1') + \ldots + (v_m - v_m')$ mit $v_i - v_i' \in \mathrm{im}(P_i)$ für alle $i \in \{1, ..., m\}$. Damit ergibt sich $0 = P_i(0) = v_i - v_i'$ für alle $i \in \{1, ..., m\}$ und $V = \mathrm{im}(P_1) \oplus ... \oplus \mathrm{im}(P_m)$. Für beliebige Basen B_i vom $\mathrm{im}(P_i)$ ist $B = B_1 \cup ... \cup B_m$ dann eine Basis von V.

Ist nun $\phi = \Sigma_{i=1}^m \lambda_i P_i$ mit $\lambda_1, ..., \lambda_m \in \mathbb{K}$, so ergibt sich $\phi(v) = \Sigma_{i=1}^m \lambda_i P_i(v) = \lambda_j P_j(v) = \lambda_j v$ für jeden Vektor $v \in \mathrm{im}(P_j)$. Also ist jeder Vektor in $\mathrm{im}(P_j) \setminus \{0\}$ ein Eigenvektor von ϕ zum Eigenwert λ_j und B eine Basis von V aus Eigenvektoren von ϕ. □

Sind in Satz 6.2.5 in der Linearkombination $\phi = \lambda_1 P_1 + \ldots + \lambda_m P_m$ von Projektoren $P_i : V \to V$ die Koeffizienten $\lambda_i \in \mathbb{K}$ paarweise verschieden, so ist das Bild des Projektors P_j gerade der Eigenraum von ϕ zum Eigenwert λ_j. Dies kann man immer erreichen, indem man gegebenenfalls zwei oder mehr Projektoren in dieser Linearkombination zu einem neuen Projektor summiert. Die Summe der Eigenräume zu verschiedenen Eigenwerten ist damit direkt. Dies gilt allgemeiner für Endomorphismen beliebiger Vektorräume.

Satz 6.2.6: Sei V ein $\mathbb{K}$-Vektorraum, $\phi : V \to V$ ein Endomorphismus und $\lambda_1, ..., \lambda_m \in \mathbb{K}$ paarweise verschiedene Eigenwerte von ϕ. Dann gilt:

1. Sind $v_1, \ldots, v_m$ Eigenvektoren zu $\lambda_1, \ldots, \lambda_m$, so ist $(v_1, \ldots, v_m)$ linear unabhängig.
2. Die Summe $E(\phi, \lambda_1) + \ldots + E(\phi, \lambda_m) = E(\phi, \lambda_1) \oplus \ldots \oplus E(\phi, \lambda_m)$ ist direkt.

Beweis:
1. Wir beweisen die Aussage per Induktion über m:
$m = 1$: Die Aussage ist offensichtlich wahr, da $v_1 \neq 0$ für jeden Eigenvektor v_1 von ϕ.
$m - 1 \to m$: Sei die Aussage bewiesen für $m - 1$ verschiedene Eigenwerte $\lambda_1, ..., \lambda_{m-1}$. Seien $\lambda_1, ..., \lambda_m$ paarweise verschiedene Eigenwerte von ϕ und $v_1, ..., v_m$ Eigenvektoren zu $\lambda_1, .., \lambda_m$.

Wäre $(v_1, ..., v_m)$ linear abhängig, so gäbe es nach Lemma 3.2.3 einen Vektor v_i, der im Spann der anderen liegt, und durch Umbenennen der Eigenwerte und Eigenvektoren könnten wir $i = m$ erreichen. Aus $v_m = \Sigma_{j=1}^{m-1} \mu_j v_j$ mit $\mu_1, ..., \mu_{m-1} \in \mathbb{K}$ folgt aber

$$\lambda_m \Sigma_{j=1}^{m-1} \mu_j v_j = \lambda_m v_m = \phi(v_m) = \phi(\Sigma_{j=1}^{m-1} \mu_j v_j) = \Sigma_{j=1}^{m-1} \mu_j \lambda_j v_j \quad \Rightarrow \quad \Sigma_{j=1}^{m-1} (\lambda_j - \lambda_m) \mu_j v_j = 0.$$

Da nach Induktionsvoraussetzung $(v_1, ..., v_{m-1})$ linear unabhängig ist und $\lambda_j - \lambda_m \neq 0$ für $j \neq m$ ergäbe sich $\mu_1 = \ldots = \mu_{m-1} = 0$ und damit $v_m = 0$, ein Widerspruch.

2. Ansonsten gäbe es einen Vektor $v \in V$, der sich auf zwei verschiedene Weisen als Summe von Elementen in $E(\phi, \lambda_i)$ schreiben ließe: $v = v_1 + \ldots + v_m = v_1' + \ldots + v_m'$ für Vektoren $v_i, v_i' \in E(\phi, \lambda_i)$ und $(v_1, ..., v_m) \neq (v_1', ..., v_m')$. Daraus würde $0 = (v_1 - v_1') + \ldots + (v_m - v_m')$ folgen mit $v_i - v_i' \in E(\phi, \lambda_i)$ und $(v_1 - v_1', ..., v_m - v_m') \neq (0, ..., 0)$, im Widerspruch zu 1. □

Aus diesem Satz folgt, dass ein Endomorphismus $\phi : V \to V$ eines n-dimensionalen $\mathbb{K}$-Vektorraums V maximal n verschiedene Eigenwerte besitzen kann. Denn jeder Eigenraum $E(\phi, \lambda_i)$ ist mindestens eindimensional. Die Dimension einer direkten Summe von Untervektorräumen ist die Summe ihrer Dimensionen. Besitzt V genau n verschiedene Eigenwerte, so treten in dieser direkten Summe also genau n eindimensionale Summanden auf.

Korollar 6.2.7: Sei V ein $\mathbb{K}$-Vektorraum mit $\dim_{\mathbb{K}} V = n$ und $\phi \in \mathrm{End}_{\mathbb{K}}(V)$. Dann hat ϕ maximal n verschiedene Eigenwerte. Gibt es n verschiedene Eigenwerte, so ist ϕ diagonalisierbar.

Beweis:
Nach Satz 6.2.6 ist jedes m-Tupel von Eigenvektoren $(v_1, ..., v_m)$ zu den paarweise verschiedenen Eigenwerten $\lambda_1, ..., \lambda_m$ von ϕ linear unabhängig. Da $\dim_{\mathbb{K}}(V) = n$ kann ein linear unabhängiges Tupel von Vektoren aus V maximal n Vektoren enthalten und somit gilt $m \leq n$. Ist $m = n$, so ist $(v_1, ..., v_m)$ nach Korollar 3.2.23 eine geordnete Basis von V. □

Beispiel 6.2.8: Für alle $\alpha, \beta, \gamma \in \mathbb{Q}$ ist die Matrix

$$A = \begin{pmatrix} 1 & \alpha & \beta \\ 0 & -1 & \gamma \\ 0 & 0 & 2 \end{pmatrix} \in \mathrm{Mat}(3 \times 3, \mathbb{Q})$$

diagonalisierbar. Da A eine obere Dreiecksmatrix ist, sind nach Bemerkung 6.2.3 die drei verschiedenen Diagonaleinträge Eigenwerte von A. Also ist A nach Korollar 6.2.7 diagonalisierbar.

Natürlich können auch Endomorphismen $\phi : V \to V$ eines n-dimensionalen Vektorraums, die weniger als n Eigenwerte besitzen, diagonalisierbar sein. Dies gilt beispielsweise für die Identitätsabbildung $\mathrm{id}_V : V \to V$ mit $\lambda = 1$ als einzigem Eigenwert. Um zu überprüfen, ob ein gegebener Endomorphismus $\phi : V \to V$ mit weniger als n Eigenwerten diagonalisierbar ist, müssen wir untersuchen, ob die direkte Summe $E(\phi, \lambda_1) \oplus ... \oplus E(\phi, \lambda_m)$ der Eigenräume aus Satz 6.2.6 ganz V ist, also ob sich ihre Dimensionen zu $n = \dim_{\mathbb{K}}(V)$ summieren.

Satz 6.2.9: (**Diagonalisierbarkeitskriterium**)
Sei V ein n-dimensionaler $\mathbb{K}$-Vektorraum und $\phi \in \mathrm{End}_{\mathbb{K}}(V)$ ein Endomorphismus mit paarweise verschiedenen Eigenwerten $\lambda_1, ..., \lambda_m \in \mathbb{K}$. Dann ist ϕ genau dann diagonalisierbar, wenn für die geometrischen Vielfachheiten $g(\phi, \lambda_i) = \dim_{\mathbb{K}} E(\phi, \lambda_i)$ gilt:

$$g(\phi, \lambda_1) + ... + g(\phi, \lambda_m) = n.$$

Beweis:
Nach Satz 6.2.6 ist die Summe der Eigenräume $E(\phi, \lambda_i)$ zu verschiedenen Eigenwerten direkt. Damit gilt

$$n = \dim_K V \geq \dim_{\mathbb{K}}(E(\phi, \lambda_1) \oplus ... \oplus E(\phi, \lambda_m)) = \dim_{\mathbb{K}} E(\phi, \lambda_1) + \ldots + \dim_{\mathbb{K}} E(\phi, \lambda_m).$$

Also ist $g(\phi, \lambda_1) + \ldots + g(\phi, \lambda_m) = n$ genau dann, wenn $E(\phi, \lambda_1) \oplus ... \oplus E(\phi, \lambda_m) = V$, und dies ist gleichbedeutend zur Diagonalisierbarkeit von ϕ. □

Beispiel 6.2.10: Wir untersuchen, für welche Werte von $\alpha, \beta, \gamma \in \mathbb{Q}$ die Matrix

$$A = \begin{pmatrix} 1 & \alpha & \beta \\ 0 & 1 & \gamma \\ 0 & 0 & 2 \end{pmatrix} \in \mathrm{Mat}(3 \times 3, \mathbb{Q})$$

diagonalisierbar ist. Die Eigenwerte von A sind die Diagonaleinträge, also 1 und 2. Für jeden Eigenwert λ ist die geometrische Vielfachheit $g(A, \lambda) = \dim_{\mathbb{Q}} E(A, \lambda)$ der Defekt der Matrix $A_\lambda := A - \lambda \cdot \mathbb{1}_3$. Man berechnet

$$A_1 := A - \mathbb{1}_3 = \begin{pmatrix} 0 & \alpha & \beta \\ 0 & 0 & \gamma \\ 0 & 0 & 1 \end{pmatrix} \qquad A_2 := A - 2 \cdot \mathbb{1}_3 = \begin{pmatrix} -1 & \alpha & \beta \\ 0 & -1 & \gamma \\ 0 & 0 & 0 \end{pmatrix}.$$

Die Matrix A_2 ist bereits in Zeilenstufenform und weist zwei Stufen auf, unabhängig vom Wert von $\alpha, \beta, \gamma \in \mathbb{Q}$. Also gilt $g(A, 2) = \mathrm{def}(A_2) = 3 - \mathrm{rg}(A_2) = 3 - 2 = 1$.

Für $\alpha \neq 0$ kann die Matrix A_1 in Zeilenstufenform gebracht werden, indem man ein Vielfaches der dritten Zeile von der zweiten Zeile abzieht und die zweite und dritte Zeile vertauscht. Da dann zwei Stufen auftreten, ergibt sich $g(A, 1) = \mathrm{def}(A_1) = 3 - \mathrm{rg}(A_1) = 3 - 2 = 1$.

Ist $\alpha = 0$ so kann man A_1 durch Subtrahieren von Vielfachen der dritten Zeile von den ersten beiden Zeilen und Vertauschen der Zeilen in Zeilenstufenform bringen, und es gibt dann nur eine Stufe. In diesem Fall gilt also $g(A,1) = \mathrm{def}(A_1) = 3 - \mathrm{rg}(A_1) = 3 - 1 = 2$.

Die Matrix A ist nach Satz 6.2.9 genau dann diagonalisierbar, wenn $g(A,2) + g(A,1) = 3$ gilt, also genau dann, wenn $\alpha = 0$.

An was man sich erinnern sollte:

Die wichtigsten Begriffe und Definitionen:

- Eigenwert, Eigenvektor, Eigenraum von Matrix/Endomorphismus,
- geometrische Vielfachheit,
- diagonalisierbar, trigonalisierbar für Matrix/Endomorphismus,

Die wichtigsten Aussagen:

- Summe der Eigenräume ist direkt, Diagonalisierbarkeitskriterium,
- maximale Anzahl der Eigenwerte,
- ähnliche Matrizen haben die gleichen Eigenwerte und geometrische Vielfachheiten,
- Beschreibung der Diagonalisierbarkeit mit kommutierenden Projektoren,
- diagonalisierbar $\Rightarrow$ trigonalisierbar.

Die wichtigsten Beispiele:

- **Eigenvektoren und Eigenwerte:** Streckung, Projektor, Involution, Spiegelung, Drehung, nilpotenter Endomorphismus, Eigenwerte von Dreiecksmatrizen,
- **Diagonalisierbarkeit:** Endomorphismen mit $\dim_{\mathbb{K}}(V)$ verschiedenen Eigenwerten, Linearkombinationen $\Sigma_{i=1}^m \alpha_i P_i$ von Projektoren mit $P_i \circ P_j = 0$ für $i \neq j$, obere Dreiecksmatrizen mit paarweise verschiedenen Diagonaleinträgen,
- **Trigonalisierbarkeit:** diagonalisierbare Matrizen/Endomorphismen.

7 Trigonalisierbarkeit und die Hauptraumzerlegung

7.1 Die Hauptraumzerlegung

Nachdem wir geklärt haben, unter welchen Bedingungen ein Endomorphismus oder eine Matrix diagonalisierbar ist, wenden wir uns der Frage der Trigonalisierbarkeit zu. Nach Bemerkung 6.2.3, 2. brauchen wir dabei nur Endomorphismen und Matrizen zu betrachten, die mindestens einen Eigenwert besitzen. Wir betrachten zunächst Endomorphismen und Matrizen mit einem einzigen Eigenwert und setzen dann die Ergebnisse für verschiedene Eigenwerte zusammen.

Lemma 7.1.1: Sei V ein endlich-dimensionaler $\mathbb{K}$-Vektorraum und $\phi : V \to V$ ein Endomorphismus mit einem einzigen Eigenwert $\lambda \in \mathbb{K}$. Dann ist ϕ trigonalisierbar genau dann, wenn $\phi_\lambda := \phi - \lambda \mathrm{id}_V : V \to V$ **nilpotent** ist, also wenn ein $k \in \mathbb{N}$ mit $\phi_\lambda^k = 0$ existiert.

Beweis:
1. Zunächst überlegt man sich, dass die beschreibende Matrix ${}_BM_B(\phi)$ des Endomorphismus ϕ bezüglich einer geordneten Basis $B = (v_1, ..., v_n)$ genau dann eine obere Dreiecksmatrix ist, wenn $\phi(v_j) \in \mathrm{span}_{\mathbb{K}}\{v_1, ..., v_j\}$ für alle $j \in \{1, ..., n\}$ gilt. Denn der Eintrag in der $j$ten Spalte von ${}_BM_B(\phi)$ sind gerade die Koeffizienten in der Darstellung von $\phi(v_j)$ als Linearkombination der Vektoren v_i. Diese verschwinden für alle $i > j$ genau dann, wenn $\phi(v_j) \in \mathrm{span}_{\mathbb{K}}\{v_1, ..., v_j\}$.

2. $\Rightarrow$: Ist ϕ trigonalisierbar, gibt es eine geordnete Basis $B = (v_1, ..., v_n)$ von V so dass ${}_BM_B(\phi)$ eine obere Dreiecksmatrix ist. Nach Bemerkung 6.2.3, 3. ist jeder Diagonaleintrag von ${}_BM_B(\phi)$ ein Eigenwert von ϕ und damit gleich λ. Also gilt $\phi(v_k) = \lambda v_k + w_k$ mit $w_k \in \mathrm{span}_{\mathbb{K}}\{v_1, ..., v_{k-1}\}$ und $\phi_\lambda(v_k) = \phi(v_k) - \lambda v_k = w_k$. Daraus ergibt sich $\phi_\lambda(\mathrm{span}_{\mathbb{K}}\{v_1, ..., v_k\}) \subseteq \mathrm{span}_{\mathbb{K}}\{v_1, ..., v_{k-1}\}$ für alle $k \in \{1, ..., n\}$ und schließlich induktiv $\phi_\lambda^m(v_k) \in \mathrm{span}_{\mathbb{K}}\emptyset = \{0\}$ für alle $m \geq k$. Also gilt insbesondere $\phi_\lambda^n(v_k) = 0$ für alle $k \in \{1, .., n\}$ und damit $\phi_\lambda^n = 0$.

3. $\Leftarrow$: Ist ϕ_λ nilpotent, so gibt es ein $j \in \mathbb{N}$ mit $\phi_\lambda^j = 0$. Sei $m := \min\{j \in \mathbb{N} | \phi_\lambda^j = 0\}$, $V_0 := \{0\}$ und $V_k := \ker(\phi_\lambda^k)$ für $k \in \{1, ..., m\}$. Dann gilt $V_{k-1} \subseteq V_k$, denn aus $v \in V_{k-1} = \ker(\phi_\lambda^{k-1})$ folgt $\phi_\lambda^k(v) = \phi_\lambda(\phi_\lambda^{k-1}(v)) = \phi_\lambda(0) = 0$, also $v \in V_k$. Ebenso ergibt sich $\phi_\lambda(V_k) \subseteq V_{k-1}$ für alle $k \in \{1, ..., m\}$, denn aus $v \in V_k = \ker(\phi_\lambda^k)$ folgt $\phi_\lambda^k(v) = \phi_\lambda^{k-1}(\phi_\lambda(v)) = 0$, also $\phi_\lambda(v) \in V_{k-1}$. Wir erhalten also eine aufsteigende Kette von Kernen, die durch ϕ_λ ineinander abgebildet werden:

$$V_0 = \{0\} \quad \overset{\xleftarrow{\phi_\lambda}}{\subseteq} \quad V_1 = \ker(\phi_\lambda) \quad \overset{\xleftarrow{\phi_\lambda}}{\subseteq} \quad V_2 = \ker(\phi_\lambda^2) \quad \overset{\xleftarrow{\phi_\lambda}}{\subseteq} \quad \overset{\cdots}{\cdots} \quad \overset{\xleftarrow{\phi_\lambda}}{\subseteq} \quad V_m = \ker(\phi_\lambda^m) = V.$$

Wir wählen nun eine geordnete Basis $B_1 = (b_1^1, ..., b_{n_1}^1)$ von V_1, ergänzen sie zu einer geordneten Basis $B_2 = (b_1^1, ..., b_{n_1}^1, b_1^2, ..., b_{n_2}^2)$ von V_2 und induktiv zu Basen $B_k = (b_1^1, ..., b_{n_1}^1, ..., b_1^k, ..., b_{n_k}^k)$ von V_k. Dann ist $B := (b_1^1, ..., b_{n_1}^1, ..., b_1^m, ..., b_{n_m}^m)$ eine geordnete Basis von V. Aus $\phi_\lambda(V_k) \subseteq V_{k-1}$ ergibt sich $\phi_\lambda(b_j^k) \in \mathrm{span}_{\mathbb{K}}\{b_1^1, ..., b_{n_{k-1}}^{k-1}\}$ für alle $k \in \{1, ..., m\}$, $j \in \{1, ..., n_k\}$. Also ist ${}_BM_B(\phi_\lambda)$ eine obere Dreiecksmatrix und damit auch ${}_BM_B(\phi) = {}_BM_B(\phi_\lambda + \lambda \mathrm{id}_V) = {}_BM_B(\phi_\lambda) + \lambda \mathbb{1}_n$. $\square$

Bemerkung 7.1.2: Aus dem Beweis von Lemma 7.1.1 folgt, dass ein Endomorphismus $\phi : V \to V$ genau dann nilpotent ist, wenn es eine geordnete Basis B von V gibt, so dass ${}_BM_B(\phi)$ eine **echte obere Dreiecksmatrix** ist, also eine Matrix M mit $M_{ij} = 0$ für $i \geq j$. Ebenso ergibt sich dass eine Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{K})$ genau dann nilpotent ist, also ein $n \in \mathbb{N}$ mit $A^n = 0$ existiert, wenn sie ähnlich zu einer echten oberen Dreiecksmatrix ist.

Beispiel 7.1.3: Der Endomorphismus

$$\phi : \mathbb{R}^3 \to \mathbb{R}^3, \qquad \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x+y \\ y+z \\ z \end{pmatrix}$$

ist offensichtlich trigonalisierbar, denn seine darstellende Matrix bezüglich der Standardbasis ist bereits eine obere Dreiecksmatrix. Da 1 der einzige Eigenwert von ϕ ist, betrachten wir

$$\phi_1 = \phi - \mathrm{id}_{\mathbb{R}^3} : \mathbb{R}^3 \to \mathbb{R}^3, \qquad \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} y \\ z \\ 0 \end{pmatrix}.$$

Dieser ist nilpotent, denn es gilt

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

Vergleicht man die Aussage in Lemma 7.1.1 mit der entsprechenden Aussage für Diagonalisierbarkeit, erkennt man die Beziehung zwischen beiden Konzepten: Ein Endomorphismus $\phi : V \to V$ mit einem einzigen Eigenwert λ ist nämlich nach Satz 6.2.9 genau dann *diagonalisierbar*, wenn $\phi = \lambda \mathrm{id}_V$ gilt, also $\phi_\lambda = \phi - \lambda \mathrm{id}_V = 0$, während er nach Lemma 7.1.1 genau dann *trigonalisierbar* ist, wenn $\phi_\lambda = \phi - \lambda \mathrm{id}_V$ nilpotent ist.

Diagonalisierbarkeit entspricht also gerade dem Spezialfall, dass der nilpotente Endomorphismus ϕ_λ die Nullabbildung ist. Die Nilpotenz des Endomorphismus ϕ_λ ist äquivalent zu der Aussage, dass ein $k \in \mathbb{N}$ existiert mit $V = \ker(\phi_\lambda^k)$. Die Bedingung $\phi_\lambda = 0$ entspricht dem Spezialfall $k = 1$, also $V = \ker(\phi_\lambda)$.

Wir verallgemeinern nun das Trigonalisierbarkeitskriterium aus Lemma 7.1.1 auf Endomorphismen $\phi : V \to V$ mit mehreren Eigenwerten. Im Fall der Diagonalisierbarkeit ersetzt man beim Übergang von einem zu mehreren Eigenwerten die Bedingung $V = \ker(\phi - \lambda \mathrm{id}_V)$ durch die Bedingung $V = E(\phi, \lambda_1) \oplus ... \oplus E(\phi, \lambda_m) = \ker(\phi - \lambda_1 \mathrm{id}_V) \oplus ... \oplus \ker(\phi - \lambda_m \mathrm{id}_V)$ aus Satz 6.2.9.

Deswegen ist es im Fall der Trigonalisierbarkeit naheliegend, zu versuchen den Vektorraum V in die Kerne von Endomorphismen $(\phi - \lambda_i \mathrm{id}_V)^{k_i} : V \to V$ für die verschiedenen Eigenwerte λ_i von ϕ zu zerlegen. Da die Exponenten $k_i \in \mathbb{N}$ für verschiedene Eigenwerte nicht bekannt sind und auch verschieden sein können, können wir sie aber nicht fest vorgeben, sondern müssen die Kerne von $(\phi - \lambda_i \mathrm{id}_V)^k$ für alle $k \in \mathbb{N}$ betrachten.

Definition 7.1.4: Sei V ein $\mathbb{K}$-Vektorraum, $\phi : V \to V$ ein Endomorphismus und λ ein Eigenwert von ϕ. Der **Hauptraum** zum Eigenwert $\lambda \in \mathbb{K}$ ist die Menge

$$H(\phi, \lambda) = \{v \in V \mid \exists k \in \mathbb{N}_0 \,\mathrm{mit}\, (\phi - \lambda \mathrm{id}_V)^k(v) = 0\} = \cup_{k \in \mathbb{N}} \ker(\phi - \lambda \mathrm{id}_V)^k.$$

Anders als im Fall des Eigenraums $E(\phi, \lambda)$, der als Kern der linearen Abbildung $\phi - \lambda \mathrm{id}_V$ automatisch ein Untervektorraum von V ist, ist dies für den Hauptraum $H(\phi, \lambda)$ zunächst nicht gesichert, denn die Vereinigung von Untervektorräumen muss kein Untervektorraum sein.

Ebenso müssen wir nachprüfen, dass die Einschränkung von $\phi - \lambda \mathrm{id}_V$ auf den Hauptraum $H(\phi, \lambda)$ tatsächlich eine nilpotente Abbildung $\phi - \lambda \mathrm{id}_V|_{H(\phi,\lambda)} : H(\phi, \lambda) \to H(\phi, \lambda)$ definiert, während die entsprechende Aussage $\phi - \lambda \mathrm{id}_V|_{E(\phi,\lambda)} = 0$ für den Eigenraum offensichtlich ist. Diese zwei Aussagen ergeben sich aus dem folgenden Satz.

Satz 7.1.5: (**Fitting-Lemma**)
Sei V ein n-dimensionaler Vektorraum über $\mathbb{K}$ und $\phi : V \to V$ ein Endomorphismus. Dann gibt es eine eindeutig bestimmte Zahl $f \in \{0, 1, ..., n\}$, den **Fitting-Index** von ϕ, mit

$$\{0\} = \ker(\phi^0) \subsetneq \ker(\phi) \subsetneq \ker(\phi^2) \subsetneq \ldots \subsetneq \ker(\phi^{f-1}) \subsetneq \ker(\phi^f) = \ker(\phi^{f+1}) = \ker(\phi^{f+2}) = \ldots$$
$$V = \mathrm{im}(\phi^0) \supsetneq \mathrm{im}(\phi^1) \supsetneq \mathrm{im}(\phi^2) \supsetneq \ldots \supsetneq \mathrm{im}(\phi^{f-1}) \supsetneq \mathrm{im}(\phi^f) = \mathrm{im}(\phi^{f+1}) = \mathrm{im}(\phi^{f_2}) = \ldots$$

Es gilt:

(i) Für jedes $\psi \in \mathrm{End}_{\mathbb{K}}(V)$ mit $\psi \circ \phi = \phi \circ \psi$ ist $\psi(\ker \phi^f) \subseteq \ker \phi^f$ und $\psi(\mathrm{im}\, \phi^f) \subseteq \mathrm{im}\, \phi^f$.

(ii) Insbesondere: $\phi(\ker \phi^f) \subseteq \ker \phi^f$ und $\phi(\mathrm{im}\, \phi^f) \subseteq \mathrm{im}\, \phi^f$.

(iii) $\phi|_{\mathrm{im}\, \phi^f} : \mathrm{im}\, \phi^f \to \mathrm{im}\, \phi^f$ ist ein Isomorphismus.

(iv) $\phi|_{\ker \phi^f} : \ker \phi^f \to \ker \phi^f$ ist nilpotent mit $f = \min\{k \in \mathbb{N}_0 | \, (\phi|_{\ker \phi^f})^k = 0\}$.

(v) Die Summe von $\ker \phi^f$ und $\mathrm{im}\, \phi^f$ ist direkt und spannt ganz V auf: $\ker \phi^f \oplus \mathrm{im}\, \phi^f = V$.

Beweis:
1. Wir beweisen die Existenz und Eindeutigkeit des Fitting-Index. Es gilt

$$v \in \ker(\phi^k) \Rightarrow \phi^k(v) = 0 \Rightarrow \phi^{k+1}(v) = \phi(\phi^k(v)) = \phi(0) = 0 \Rightarrow v \in \ker(\phi^{k+1})$$
$$v \in \mathrm{im}(\phi^k) \Rightarrow \exists u \in V : v = \phi^k(u) \Rightarrow v = \phi^{k-1}(\phi(u)) \Rightarrow v \in \mathrm{im}(\phi^{k-1}).$$

Also bilden die Kerne der Potenzen von ϕ eine aufsteigende und die Bilder eine absteigende Kette von Untervektorräumen

$$\{0\} \subseteq \ker(\phi) \subseteq \ker(\phi^2) \subseteq .... \qquad V \supseteq \mathrm{im}(\phi) \supseteq \mathrm{im}(\phi^2) \supseteq ...$$

Da V endlich-dimensional ist, gibt es $k, l \in \mathbb{N}_0$ mit $0 \leq k, l \leq n$, so dass $\ker(\phi^k) = \ker(\phi^{k+1})$ und $\mathrm{im}(\phi^l) = \mathrm{im}(\phi^{l+1})$. Denn ansonsten wäre

$$\dim_{\mathbb{K}} \ker(\phi^{n+1}) \geq \dim_{\mathbb{K}} \ker(\phi^n) + 1 \geq ... \geq \dim_{\mathbb{K}} \ker(\phi) + n \geq n + 1 > \dim_{\mathbb{K}} V$$
$$\dim_{\mathbb{K}} \mathrm{im}(\phi) \geq \dim_{\mathbb{K}} \mathrm{im}(\phi^2) + 1 \geq ... \geq \dim_{\mathbb{K}} \mathrm{im}(\phi^{n+1}) + n \geq n + 1 > \dim_{\mathbb{K}} V.$$

Sei $f = \min\{k \in \mathbb{N}_0 | \ker(\phi^k) = \ker(\phi^{k+1})\} \leq n$ und $m = \min\{l \in \mathbb{N}_0 | \,\mathrm{im}(\phi^l) = \mathrm{im}(\phi^{l+1})\} \leq n$. Wir zeigen, dass $f = m$ gilt. Dazu berechnen wir mit der Dimensionsformel

$$\dim_{\mathbb{K}} \mathrm{im}(\phi^{f+1}) = \dim_{\mathbb{K}} V - \dim_{\mathbb{K}} \ker(\phi^{f+1}) = \dim_{\mathbb{K}} V - \dim_{\mathbb{K}} \ker(\phi^f) = \dim_{\mathbb{K}} \mathrm{im}(\phi^f)$$
$$\dim_{\mathbb{K}} \ker(\phi^{m+1}) = \dim_{\mathbb{K}} V - \dim_{\mathbb{K}} \mathrm{im}(\phi^{m+1}) = \dim_{\mathbb{K}} V - \dim_{\mathbb{K}} \mathrm{im}(\phi^m) = \dim_{\mathbb{K}} \ker(\phi^m).$$

Aus der ersten Identität folgt per Definition von f und m, dass $f \geq m$ und aus der zweiten $m \geq f$, also $m = f$. Das beweist die Existenz und Eindeutigkeit des Fitting-Index.

2. Wir beweisen Aussagen (i) bis (v):

(i) Aus $\phi \circ \psi = \psi \circ \phi$ folgt induktiv auch $\phi^k \circ \psi = \psi \circ \phi^k$ für alle $k \in \mathbb{N}$. Daraus ergibt sich

$$v \in \ker(\phi^f) \Rightarrow \phi^f(v) = 0 \Rightarrow \phi^f(\psi(v)) = \psi(\phi^f(v)) = \psi(0) = 0 \Rightarrow \psi(v) \in \ker(\phi^f)$$
$$v \in \mathrm{im}(\phi^f) \Rightarrow \exists u \in V : v = \phi^f(u) \Rightarrow \psi(v) = \psi(\phi^f(u)) = \phi^f(\psi(u)) \Rightarrow \psi(v) \in \mathrm{im}(\phi^f),$$

also $\psi(\ker \phi^f) \subseteq \ker \phi^f$ und $\psi(\mathrm{im}\, \phi^f) \subseteq \mathrm{im}\, \phi^f$.

(ii) Die Aussage ist ein Spezialfall von (i), nämlich der Fall $\psi = \phi$.

(iii) Die Aussage folgt aus (ii). Da $\phi(\mathrm{im}\,\phi^f) = \mathrm{im}(\phi^{f+1}) = \mathrm{im}(\phi^f)$, ist $\phi|_{\mathrm{im}\,\phi^f} : \mathrm{im}\,\phi^f \to \mathrm{im}\,\phi^f$ surjektiv und, da $\mathrm{im}\,\phi^f$ endlich-dimensional ist, damit auch ein Isomorphismus.

(iv) Die Aussage $\phi^f(u) = 0$ für alle $u \in \ker\phi^f$ gilt per Definition von $\ker\phi^f$. Gäbe es ein $l < f$ mit $\phi^l(u) = 0$ für alle $u \in \ker\phi^f$, so ergäbe sich $\ker(\phi^f) \subseteq \ker(\phi^l) \subseteq ... \subseteq \ker(\phi^{f-1})$, also $\ker(\phi^f) = \ker(\phi^{f-1})$, ein Widerspruch zur Definition des Fitting-Index, und damit folgt (iv).

(v) Wir zeigen zunächst, dass $\ker\phi^f \cap \mathrm{im}\,\phi^f = \{0\}$ gilt. Ist $v \in \ker\phi^f \cap \mathrm{im}\,\phi^f$, so folgt mit (iv) $\phi^f(v) = 0$. Da aber $\phi|_{\mathrm{im}\,\phi^f} : \mathrm{im}\,\phi^f \to \mathrm{im}\,\phi^f$ nach (iii) ein Isomorphismus ist, ist auch $(\phi|_{\mathrm{im}\,\phi^f})^f : \mathrm{im}\,\phi^f \to \mathrm{im}\,\phi^f$ ein Isomorphismus und aus $\phi^f(v) = 0$ und $v \in \mathrm{im}\,\phi^f$ folgt $v = 0$. Also ist $\ker\phi^f \cap \mathrm{im}\,\phi^f = \{0\}$, und die Summe $\ker\phi^f + \mathrm{im}\,\phi^f = \ker\phi^f \oplus \mathrm{im}\,\phi^f$ direkt. Mit der Dimensionsformel folgt dann

$$\dim_{\mathbb{K}}(V) = \dim_{\mathbb{K}}(\ker\phi^f) + \dim_{\mathbb{K}}(\mathrm{im}\,\phi^f) = \dim_{\mathbb{K}}(\ker\phi^f \oplus \mathrm{im}\,\phi^f) \Rightarrow \ker\phi^f \oplus \mathrm{im}\,\phi^f = V.$$

□

Beispiel 7.1.6: Für den Endomorphismus

$$\phi : \mathbb{R}^3 \to \mathbb{R}^3, \quad \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} y \\ z \\ 0 \end{pmatrix}$$

aus Beispiel 7.1.3 gilt

$$\phi^2 : \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} z \\ 0 \\ 0 \end{pmatrix}$$
$$\phi^3 : \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Daraus ergibt sich: $\{0\} \subseteq \ker(\phi) = \mathbb{K}e_1 \subseteq \ker(\phi^2) = \mathrm{span}_{\mathbb{K}}\{e_1, e_2\} \subseteq \ker(\phi^3) = \mathbb{R}^3$
$\mathbb{R}^3 \supseteq \mathrm{im}(\phi) = \mathrm{span}_{\mathbb{K}}\{e_1, e_2\} \supseteq \mathrm{im}(\phi^2) = \mathbb{K}e_1 \supseteq \mathrm{im}(\phi^3) = \{0\}$.

Also ist sein Fitting-Index $f = 3$.

Man sieht, dass für $k \in \{1, 2\}$ immer $\mathrm{im}(\phi^k) \cap \ker(\phi^k) \neq \{0\}$ und $\mathrm{im}(\phi^k) + \ker(\phi^k) \neq \mathbb{R}^3$ gilt. Erst für $k = f = 3$ erhält man $\mathrm{im}(\phi^3) \cap \ker(\phi^3) = \{0\}$ und $\mathrm{im}(\phi^3) + \ker(\phi^3) = \mathbb{R}^3$.

Indem wir das Fitting-Lemma nun auf die Endomorphismen $\phi_\lambda = \phi - \lambda \mathrm{id}_V : V \to V$ anwenden, die in Definition 7.1.4 der Haupträume auftreten, können wir wichtige Eigenschaften der Haupträume herleiten. Insbesondere können wir damit die Haupträume damit als Kerne von Endomorphismen charakterisieren und so zeigen, dass sie tatsächlich Untervektorräume von V bilden. Zusätzlich ergibt sich ähnlich eine Zerlegung von V als direkte Summe des Kerns und des Bilds eines Endomorphismus[14]. Da diese Unterräume invariant sind, können wir dann den Endomorphismus in seine Einschränkungen auf diese Unterräume zerlegen.

[14] Achtung: das bedeutet nicht, dass der Endomorphismus ein Projektor sein muss. Dass $\psi : V \to V$ ein Projektor ist, ist eine hinreichende, aber keine notwendige Bedingung für eine Zerlegung $V = \ker\psi \oplus \mathrm{im}\,\psi$. Auch für jeden Endomorphismus $\psi = \chi \circ P : V \to V$, der sich als Verkettung eines Projektors $P : V \to V$ und eines Endomorphismus $\chi : V \to V$ mit $\chi(\mathrm{im}\,P) = \mathrm{im}\,P$ schreiben lässt, ergibt sich eine solche Zerlegung.

Korollar 7.1.7: (**Eigenschaften des Hauptraums**)
Sei V ein n-dimensionaler Vektorraum über $\mathbb{K}$, $\phi \in \mathrm{End}_{\mathbb{K}}(V)$ und $\phi_\mu := \phi - \mu\,\mathrm{id}_V : V \to V$ für $\mu \in \mathbb{K}$. Dann gilt für jeden Eigenwert λ von ϕ

$$\{0\} \subsetneq E(\phi,\lambda) \subseteq H(\phi,\lambda) = \ker(\phi_\lambda^{f_\lambda}) \subseteq V,$$

wobei $f_\lambda \in \{0,1,...,n\}$ den Fitting-Index von ϕ_λ bezeichnet. Außerdem gilt:

(i) Ist $\psi \in \mathrm{End}_{\mathbb{K}}(V)$ mit $\phi\circ\psi = \psi\circ\phi$, so ist $\psi(H(\phi,\lambda)) \subseteq H(\phi,\lambda)$ und $\psi(\mathrm{im}(\phi_\lambda^{f_\lambda})) \subseteq \mathrm{im}(\phi_\lambda^{f_\lambda})$.

(ii) Insbesondere: $\phi_\mu(H(\phi,\lambda)) \subseteq H(\phi,\lambda)$ und $\phi_\mu(\mathrm{im}(\phi_\lambda^{f_\lambda})) \subseteq \mathrm{im}(\phi_\lambda^{f_\lambda})$ für alle $\mu \in \mathbb{K}$.

(iii) $\phi_\lambda|_{\mathrm{im}(\phi_\lambda^{f_\lambda})} : \mathrm{im}(\phi_\lambda^{f_\lambda}) \to \mathrm{im}(\phi_\lambda^{f_\lambda})$ ist ein Isomorphismus.

(iv) $\phi_\lambda|_{H(\phi,\lambda)} : H(\phi,\lambda) \to H(\phi,\lambda)$ ist nilpotent mit $f_\lambda = \min\{k \in \mathbb{N}_0 \mid \phi_\lambda^k = 0\}$.

(v) $V = H(\phi,\lambda) \oplus \mathrm{im}(\phi_\lambda^{f_\lambda})$.

Beweis:
Dass $H(\phi,\lambda) = \ker(\phi_\lambda^{f_\lambda})$ ergibt sich direkt aus der Definition des Fitting-Index in Satz 7.1.5 und der Definition des Hauptraums. Denn für alle $k \geq f_\lambda$ gilt $\ker(\phi_\lambda^k) = \ker(\phi_\lambda^{f_\lambda})$ und für $k < f_\lambda$ gilt $\ker(\phi_\lambda^k) \subsetneq \ker(\phi_\lambda^{f_\lambda})$. Also folgt $H(\phi,\lambda) = \cup_{k\in\mathbb{N}_0}\ker(\phi_\lambda^k) = \ker(\phi_\lambda^{f_\lambda})$. Als Kern des Endomorphismus $\phi_\lambda^{f_\lambda}$ ist $H(\phi,\lambda) \subseteq V$ ein Untervektorraum, und per Definition des Eigenraums gilt $\{0\} \neq E(\phi,\lambda) = \ker(\phi_\lambda) \subseteq \ker(\phi_\lambda^{f_\lambda}) = H(\phi,\lambda)$. Die übrigen Aussagen ergeben sich durch Anwendung von Satz 7.1.5 auf den Endomorphismus ϕ_λ und die Endomorphismen ϕ_μ. □

Mit Hilfe von Korollar 7.1.7 können wir eine Verallgemeinerung von Satz 6.2.6 für Haupträume beweisen, nämlich dass die Summe der Haupträume zu verschiedenen Eigenwerten direkt ist. Damit reduziert sich der allgemeine Fall auf die Situation in Lemma 7.1.1, und wir erhalten ein Gegenstück des Diagonalisierbarkeitskriteriums in Satz 6.2.9 für Trigonalisierbarkeit.

Zusätzlich erhalten wir hier aber noch ein stärkeres Resultat, nämlich ein ϕ-invariantes Komplement der direkten Summe der Haupträume. Damit können wir dann für nicht trigonalisierbare Endomorphismen $\phi : V \to V$ eine Basis angeben, bezüglich der die darstellende Matrix eine Blockdiagonalmatrix ist mit einem trigonalisierbaren Block und einem Block ohne Eigenwerte.

Satz 7.1.8: (**Hauptraumzerlegung**)
Sei V ein endlich-dimensionaler Vektorraum über $\mathbb{K}$ und $\phi : V \to V$ ein Endomorphismus mit paarweise verschiedenen Eigenwerten $\lambda_1, ..., \lambda_m \in \mathbb{K}$. Dann ist

$$V = H(\phi,\lambda_1) \oplus ... \oplus H(\phi,\lambda_m) \oplus W$$

mit einem ϕ-invarianten Unterraum $W \subseteq V$, also $\phi(W) \subseteq W$. Sind $f_1, .., f_m$ die Fitting-Indizes der Endomorphismen $\phi_{\lambda_i} = \phi - \lambda_i\,\mathrm{id}_V : V \to V$, so gilt

$$W = \mathrm{im}(\phi_{\lambda_1}^{f_1} \circ ... \circ \phi_{\lambda_n}^{f_m}) = \cap_{i=1}^m \mathrm{im}(\phi_{\lambda_i}^{f_i}) \qquad H(\phi,\lambda_1) \oplus ... \oplus H(\phi,\lambda_m) = \ker(\phi_{\lambda_1}^{f_1} \circ ... \circ \phi_{\lambda_n}^{f_m}).$$

Beweis:
1. Wir zeigen zunächst, dass $H(\phi,\lambda_i) \cap H(\phi,\lambda_j) = \{0\}$ für $i \neq j$ gilt:

Sei dazu $0 \neq v \in H(\phi,\lambda_i) \cap H(\phi,\lambda_j)$. Dann gilt auch $\phi_{\lambda_i}^l(v), \phi_{\lambda_j}^l(v) \in H(\phi,\lambda_i) \cap H(\phi,\lambda_j)$ für alle $l \in \mathbb{N}$ nach Korollar 7.1.7 (ii). Per Definition des Hauptraums $H(\phi,\lambda_i)$ gibt es ein $l \in \mathbb{N}_0$ mit $\phi_{\lambda_i}^l(v) = 0$, und wegen $v \neq 0$ ist $k := \min\{l \in \mathbb{N}_0 | \, \phi_{\lambda_i}^l(v) = 0\} \geq 1$.

Der Vektor $w := \phi_{\lambda_i}^{k-1}(v) \in H(\phi,\lambda_i) \cap H(\phi,\lambda_j) \setminus \{0\}$ ist dann ein Eigenvektor von ϕ zum Eigenwert λ_i, denn $\phi_{\lambda_i}(w) = \phi_{\lambda_i}^k(v) = 0$. Es folgt $\phi_{\lambda_j}^l(w) = (\lambda_i - \lambda_j)^l w$ für alle $l \in \mathbb{N}$. Da $w \in H(\phi,\lambda_j) \setminus \{0\}$ folgt $i = j$.

2. Nach Korollar 7.1.7 (v) ist $V = H(\phi, \lambda_i) \oplus W_i$ mit $W_i := \mathrm{im}(\phi_{\lambda_i}^{f_i})$ für jeden Eigenwert λ_i. Wir zeigen, dass $H(\phi, \lambda_j) \subseteq W_i$ und $W_i = H(\phi, \lambda_j) \oplus W_i \cap W_j$ für alle $i \neq j$ gilt.

Da $V = H(\phi, \lambda_i) \oplus W_i$, können wir für $i \neq j$ jeden Vektor $v \in H(\phi, \lambda_j)$ zerlegen als $v = v_i + w_i$ mit $v_i \in H(\phi, \lambda_i)$ und $w_i \in W_i$. Nach Korollar 7.1.7 (ii) gilt $\phi_{\lambda_j}^k(v_i) \in H(\phi, \lambda_i)$ und $\phi_{\lambda_j}^k(w_i) \in W_i$ für alle $k \in \mathbb{N}$. Da $V = H(\phi, \lambda_i) \oplus W_i$, folgt

$$0 = \phi_{\lambda_j}^{f_j}(v) = \phi_{\lambda_j}^{f_j}(v_i) + \phi_{\lambda_j}^{f_j}(w_i) \quad \Rightarrow \quad \phi_{\lambda_j}^{f_j}(v_i) = \phi_{\lambda_j}^{f_j}(w_i) = 0.$$

Damit ist $v_i \in H(\phi, \lambda_i) \cap H(\phi, \lambda_j) = \{0\}$. Es folgt $v = w_i \in W_i$ und $H(\phi, \lambda_j) \subseteq W_i$.

Wegen $V = H(\phi, \lambda_j) \oplus W_j$ lässt sich jedes $w \in W_i$ zerlegen als $w = v_j + w_j$ mit $v_j \in H(\phi, \lambda_j) \subseteq W_i$ und $w_j \in W_j$. Da $w, v_j \in W_i$ folgt auch $w_j = w - v_j \in W_i$ und damit $w \in W_i \cap W_j$. Also ist $W_i = H(\phi, \lambda_j) \oplus W_i \cap W_j$.

3. Wir beweisen die Hauptraumzerlegung:

Aus $W_i = H(\phi, \lambda_j) \oplus W_j$ für $i \neq j$ ergibt sich induktiv

$$V = H(\phi, \lambda_1) \oplus W_1 = H(\phi, \lambda_1) \oplus H(\phi, \lambda_2) \oplus W_1 \cap W_2 = ... = H(\phi, \lambda_1) \oplus ... \oplus H(\phi, \lambda_m) \oplus \cap_{i=1}^m W_i.$$

Ist $v = v_1 + ... + v_m + w$ mit $v_i \in H(\phi, \lambda_i)$ und $w \in W = \cap_{i=1}^m W_i$, so gilt wegen $\phi_{\lambda_i}^{f_i}|_{H(\phi,\lambda_i)} = 0$ und $\phi_{\lambda_i}^{f_i} \circ \phi_{\lambda_j}^{f_j} = \phi_{\lambda_j}^{f_j} \circ \phi_{\lambda_i}^{f_i}$ dann $\phi_{\lambda_1}^{f_1} \circ ... \circ \phi_{\lambda_n}^{f_m}(v) = \phi_{\lambda_1}^{f_1} \circ ... \circ \phi_{\lambda_n}^{f_m}(w)$. Da nach Korollar 7.1.7 (iii) $\phi_{\lambda_1}^{f_1} \circ ... \circ \phi_{\lambda_n}^{f_m}|_W : W \to W$ ein Isomorphismus ist, folgt

$$W = \cap_{i=1}^m W_i = \mathrm{im}(\phi_{\lambda_1}^{f_1} \circ ... \circ \phi_{\lambda_n}^{f_m}) \qquad H(\phi, \lambda_1) \oplus ... \oplus H(\phi, \lambda_m) = \ker(\phi_{\lambda_1}^{f_1} \circ ... \circ \phi_{\lambda_n}^{f_m}).$$

Die Invarianz von W ergibt sich aus Korollar 7.1.7 (ii), denn aus $\phi(W_i) \subseteq W_i$ folgt $\phi(W) = \phi(\cap_{i=1}^m W_i) \subseteq \cap_{i=1}^m \phi(W_i) \subseteq \cap_{i=1}^m W_i = W$. □

Ebenso wie Kerne, Bilder, Eigenwerte und Eigenvektoren können wir auch den Hauptraum und den Fitting-Index einer Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{K})$ definieren als den Hauptraum und den Fitting-Index des zugehörigen Endomorphismus $\phi_A : \mathbb{K}^n \to \mathbb{K}^n$, $v \mapsto A \cdot v$. Hierbei ergibt sich dann wieder ein direkter Zusammenhang zwischen dem Hauptraum eines Endomorphismus $\phi : V \to V$ und dem Hauptraum seiner beschreibenden Matrizen ${}_BM_B(\phi) - \lambda \mathbb{1}_n$ und zwischen den Dimensionen der Kernen von $(\phi - \lambda \mathrm{id}_V)^k$ und $({}_BM_B(\phi) - \lambda \mathbb{1}_n)^k$ für alle $k \in \mathbb{N}$.

Definition 7.1.9: Der **Hauptraum** einer Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{K})$ zu einem Eigenwert λ von A und der **Fitting-Index** von A sind der Hauptraum $H(A, \lambda) = H(\phi_A, \lambda)$ und der Fitting-Index des zugehörigen Endomorphismus $\phi_A : \mathbb{K}^n \to \mathbb{K}^n$, $v \mapsto A \cdot v$.

Korollar 7.1.10: Sei V ein n-dimensionaler $\mathbb{K}$-Vektorraum mit einer geordneten Basis B und $\phi : V \to V$ ein Endomorphismus.

1. Dann haben ϕ und ${}_BM_B(\phi)$ die gleichen Eigenwerte, und für alle Eigenwerte λ und $k \in \mathbb{N}$ gilt $\mathrm{def}(\phi - \lambda \mathrm{id}_V)^k = \mathrm{def}({}_BM_B(\phi) - \lambda \mathbb{1}_n)^k$. Insbesondere haben $\phi - \lambda \mathrm{id}_V$ und ${}_BM_B(\phi) - \lambda \mathbb{1}_n$ den gleichen Fitting-Index.

2. Sind $M, N \in \mathrm{Mat}(n \times n, \mathbb{K})$ ähnlich, so haben sie die gleichen Eigenwerte und für alle Eigenwerte λ und $k \in \mathbb{N}$ gilt $\mathrm{def}(M - \lambda \mathbb{1}_n)^k = \mathrm{def}(N - \lambda \mathbb{1}_n)^k$. Insbesondere haben $M - \lambda \mathbb{1}_n$ und $N - \lambda \mathbb{1}_n$ die gleichen Fitting-Indizes.

Beweis:
1. Die Aussagen über die Eigenwerte wurde bereits in Lemma 6.1.5 und Korollar 6.1.6 bewiesen. Für die Aussage über die Defekte betrachten wir den Isomorphismus ${}_BS : V \to \mathbb{K}^n$ aus Korollar 4.2.4. Dann gilt ${}_BM_B(\phi) \cdot {}_BS(v) = {}_BS \circ \phi(v)$ für alle $v \in V$ nach Satz 4.2.12, (iii). Daraus folgt mit Satz 4.2.12, (i)-(v) auch $({}_BM_B(\phi) - \lambda \mathbb{1}_n) \cdot {}_BS(v) = {}_BS \circ (\phi - \lambda \mathrm{id}_V)(v)$ und induktiv $({}_BM_B(\phi) - \lambda \mathbb{1}_n)^k \cdot {}_BS(v) = {}_BS \circ (\phi - \lambda \mathrm{id}_V)^k(v)$ für alle $k \in \mathbb{N}$, $\lambda \in \mathbb{K}$ und $v \in V$. Da ${}_BS : V \to \mathbb{K}^n$ ein Isomorphismus ist, folgt für alle $k \in \mathbb{N}$, $\lambda \in \mathbb{K}$

$$\mathrm{def}(\phi - \lambda \mathrm{id}_V)^k = \dim_{\mathbb{K}} \ker(\phi - \lambda \mathrm{id}_V)^k = \dim_{\mathbb{K}} \ker({}_BM_B(\phi) - \lambda \mathbb{1}_n)^k = \mathrm{def}({}_BM_B(\phi) - \lambda \mathbb{1}_n)^k.$$

2. Sind $M, N \in \mathrm{Mat}(n \times n, \mathbb{K})$ ähnlich, so gibt es einen Endomorphismus $\phi : V \to V$ und geordnete Basen B, C von V mit $M = {}_BM_B(\phi)$ und $N = {}_CM_C(\phi)$. Also haben nach 1. M und N die gleichen Eigenwerte wie ϕ und für alle Eigenwerte λ und $k \in \mathbb{N}$ gilt $\mathrm{def}(M - \lambda \mathbb{1}_n)^k = \mathrm{def}(\phi - \lambda \mathrm{id}_V)^k = \mathrm{def}(N - \lambda \mathbb{1}_n)^k$. □

Beispiel 7.1.11: Wir bestimmen die Hauptraumzerlegung der Matrix

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

Da A eine obere Dreiecksmatrix ist, sind ihre Eigenwerte gerade die Diagonaleinträge, also $\lambda_1 = 1$ und $\lambda_2 = 2$. Für die Matrix $A_1 := A - \lambda_1 \mathbb{1}_4 = A - \mathbb{1}_4$ gilt

$$A_1 = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix} \qquad \Rightarrow \quad \mathrm{def}(A_1) = 2$$

$$A_1^2 = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \qquad \Rightarrow \quad \mathrm{def}(A_1^2) = 3.$$

Da $\dim_{\mathbb{K}} H(A,1) + \dim_{\mathbb{K}} H(A,2) \leq 4$ gelten muss, und $\dim_{\mathbb{K}} H(A,2) \geq 1$ gilt, ergibt sich $\dim_{\mathbb{K}} H(A,1) \leq 3 = \mathrm{def}(A_1^2)$, und damit hat A_1 Fitting-Index 2. Es folgt

$$H(A,1) = \ker(A_1^2) = \mathrm{span}_{\mathbb{K}}\{e_1 - e_2, e_3, 2e_1 - e_4\} \qquad \mathrm{im}(A_1^2) = \mathbb{K}e_1.$$

Für die Matrix $A_2 = A - \lambda_2 \mathbb{1}_4 = A - 2\mathbb{1}_4$ ergibt sich

$$A_2 = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 0 & -1 & 0 & 1 \\ 0 & 0 & -1 & 2 \\ 0 & 0 & 0 & -1 \end{pmatrix} \qquad \Rightarrow \quad \mathrm{def}(A_2) = 1.$$

Da $\dim_{\mathbb{K}} H(A,1) + \dim_{\mathbb{K}} H(A,2) \leq 4$ und $\dim_{\mathbb{K}} H(A,1) = 3$, folgt $\dim_{\mathbb{K}} H(A,2) = 1 = \mathrm{def}(A_2)$ und A_2 hat Fitting-Index 1. Damit gilt

$$H(A,2) = E(A,2) = \ker(A_2) = \mathbb{K}e_1 = \mathrm{im}(A_1^2)$$
$$\mathrm{im}(A_2) = \mathrm{span}_{\mathbb{K}}\{e_1 - e_2, -e_3, e_1 + e_2 + 2e_3 - e_4\} = \mathrm{span}_{\mathbb{K}}\{e_1 - e_2, e_3, 2e_1 - e_4\} = H(A,1),$$

und es folgt $H(A,1) \oplus H(A,2) = \mathbb{R}^4$.

Korollar 7.1.10 liefert einen echten Mehrwert gegenüber der Aussage, dass ähnliche Matrizen die gleiche Spuren und Determinanten haben und ihre Eigenwerte sowie die Dimensionen der zugehörigen Eigenräume übereinstimmen. Es gibt nämlich Matrizen, die die gleichen Spuren, Determinanten, Eigenwerte, Eigenraumdimensionen und sogar die gleichen Hauptraumdimensionen haben, deren Fitting-Indizes sich aber trotzdem unterscheiden.

Beispiel 7.1.12: Sei $\mathbb{K}$ ein Körper und $\lambda \in \mathbb{K}$. Die Matrizen

$$A = \begin{pmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 0 \\ 0 & 0 & 0 & \lambda \end{pmatrix} \qquad B = \begin{pmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 0 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{pmatrix}$$

sind nicht ähnlich. Denn für die Matrix $A_\lambda = A - \lambda \mathbb{1}_4$ gilt

$$A_\lambda = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \qquad \mathrm{def} A_\lambda = g(A, \lambda) = 2$$

$$A_\lambda^2 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \qquad \mathrm{def} A_\lambda^2 = 3.$$

Für die Matrix $B_\lambda = B - \lambda \mathbb{1}_4$ ergibt sich aber

$$B_\lambda = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \qquad \mathrm{def} B_\lambda = g(B, \lambda) = 2$$

$$B_\lambda^2 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \qquad \mathrm{def} B_\lambda^2 = \dim_K H(B, \lambda) = 4.$$

Also ist $3 = \mathrm{def} A_\lambda^2 \neq \mathrm{def} B_\lambda^2 = 4$. Damit sind A und B nach Korollar 7.1.10 nicht ähnlich, obwohl sie die gleichen Eigenwerte, Eigenraumdimensionen und Hauptraumdimensionen haben.

Wir werden später in Korollar 9.1.7 zeigen, dass die notwendigen Bedingungen für die Ähnlichkeit von Matrizen aus Korollar 7.1.10 für trigonalisierbare Matrizen auch hinreichend sind: Zwei trigonalisierbare Matrizen $A, B \in \mathrm{Mat}(n \times n, \mathbb{K})$ sind genau dann ähnlich, wenn sie die gleichen Eigenwerte haben und für jeden Eigenwert λ und alle $k \in \mathbb{N}$ gilt $\mathrm{def}(A - \lambda \mathbb{1}_n)^k = \mathrm{def}(B - \lambda \mathbb{1}_n)^k$.

7.2 Trigonalisierbarkeit

Wir untersuchen nun mit Hilfe der Hauptraumzerlegung die Trigonalisierbarkeit eines Endomorphismus. Dazu betrachten wir die Dimensionen seiner Haupträume und erhalten das folgende Gegenstück von Satz 6.2.9 über Diagonalisierbarkeit.

Satz 7.2.1: (**Trigonalisierbarkeitskriterium**)
Sei V ein n-dimensionaler Vektorraum über $\mathbb{K}$ und $\phi : V \to V$ ein Endomorphismus mit paarweise verschiedenen Eigenwerten $\lambda_1, ..., \lambda_m \in \mathbb{K}$. Dann ist ϕ genau dann trigonalisierbar, wenn für die Dimensionen $h(\phi, \lambda) = \dim_{\mathbb{K}} H(\phi, \lambda)$ der Haupträume gilt

$$h(\phi, \lambda_1) + ... + h(\phi, \lambda_m) = n. \tag{14}$$

Beweis:
$\Leftarrow$: Wir benutzen die Notation $\phi_\mu := \phi - \mu \mathrm{id}_V : V \to V$ für alle $\mu \in \mathbb{K}$.
Erfüllen die Dimensionen der Haupträume die Bedingung (14), so folgt mit Satz 7.1.8

$$n = \dim_{\mathbb{K}} V = \dim_{\mathbb{K}} H(\phi, \lambda_1) + ... + \dim_{\mathbb{K}} H(\phi, \lambda_m) = \dim_{\mathbb{K}}(H(\phi, \lambda_1) \oplus ... \oplus H(\phi, \lambda_m)) \leq n$$

und somit $H(\phi, \lambda_1) \oplus ... \oplus H(\phi, \lambda_m) = V$.

Nach Korollar 7.1.7 (iv) ist $\phi_{\lambda_i}|_{H(\phi,\lambda_i)} : H(\phi, \lambda_i) \to H(\phi, \lambda_i)$ nilpotent und damit ist nach Lemma 7.1.1 $\phi|_{H(\phi,\lambda_i)} : H(\phi, \lambda_i) \to H(\phi, \lambda_i)$ trigonalisierbar. Also gibt es geordnete Basen $B_i = (b^i_1, ..., b^i_{h_i})$ von $H(\phi, \lambda_i)$, so dass die beschreibenden Matrizen $M_i := {}_{B_i}M_{B_i}(\phi|_{H(\phi,\lambda_i)})$ obere Dreiecksmatrizen sind. Da die Summe der Haupträume zu verschiedenen Eigenwerten direkt und gleich V ist, erhält man so eine geordnete Basis $B = (b^1_1, ..., b^1_{h_1}, ..., b^m_1, ..., b^m_{h_m})$ von V, so dass die beschreibende Matrix von ϕ bezüglich B eine blockdiagonale Matrix ist mit den oberen Dreicksmatrizen M_i in der Diagonalen

$${}_BM_B(\phi) = \begin{pmatrix} M_1 & 0 & \dots & 0 \\ 0 & M_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & M_m \end{pmatrix}.$$

$\Rightarrow$: Ist ϕ trigonalisierbar, so gibt es eine geordnete Basis B von V, so dass ${}_BM_B(\phi)$ eine obere Dreiecksmatrix ist. In der Diagonalen stehen nach Bemerkung 6.2.3, 3. die Eigenwerte von ϕ. Durch eine Permutation der Basisvektoren können wir erreichen, dass sie aufsteigend geordnet sind, also dass in der Diagonalen zunächst h_1 mal der Eigenwert λ_1 auftritt, dann h_2 mal der Eigenwert λ_2 und zuletzt h_m mal der Eigenwert λ_m.

Ist $B = (b^1_1, ..., b^1_{h_1}, ..., b^m_1, ..., b^m_{h_m})$, so folgt $\phi(b^i_j) = \lambda_i b^i_j + w^i_j$ mit $w^i_j \in \mathrm{span}_{\mathbb{K}}\{b^1_1, ..., b^i_{j-1}\}$ und $\phi_{\lambda_i}(b^i_j) = w^i_j \in \mathrm{span}_{\mathbb{K}}\{b^1_1, ..., b^i_{j-1}\}$. Damit ist $\phi^{h_i}_{\lambda_i}(b^i_j) \in \mathrm{span}_{\mathbb{K}}\{b^1_1, ..., b^{i-1}_{h_{i-1}}\}$. Indem man das gleiche Argument auf $b^{i-1}_{h_{i-1}}, \dots, b^1_{h_1}$ anwendet, erhält man $\phi^{h_1}_{\lambda_1} \circ \dots \circ \phi^{h_m}_{\lambda_m}(b^i_j) = 0$ für alle $i \in \{1, ..., m\}$ und $j \in \{1, ..., h_i\}$ und damit $V = \mathrm{span}_{\mathbb{K}} B \subseteq H(\phi, \lambda_1) \oplus \dots \oplus H(\phi, \lambda_m)$. Es folgt

$$\dim_{\mathbb{K}} V = h_1 + \dots + h_m \leq h(\phi, \lambda_1) + ... + h(\phi, \lambda_m) = \dim_{\mathbb{K}}(H(\phi, \lambda_1) \oplus \dots \oplus H(\phi, \lambda_m)) \leq \dim_{\mathbb{K}} V.$$

$\square$

Satz 7.1.8 über die Hauptraumzerlegung liefert noch allgemeinere Aussagen als das Trigonalisierbarkeitskriterium in Satz 7.2.1. Die Tatsache, dass die Haupträume $H(\phi, \lambda_i)$ und der Vektorraum W in Satz 7.1.8 invariant unter ϕ sind, erlaubt es uns nämlich, die beschreibende Matrix von ϕ in zwei Blöcke zu zerlegen, einen trigonalisierbaren Block für die Einschränkung von ϕ auf den Unterraum $H(\phi, \lambda_1) \oplus ... \oplus H(\phi, \lambda_n)$ und einen Block ohne Eigenwerte für die Einschränkung von ϕ auf den Unterraum W.

Korollar 7.2.2: Sei V ein endlich-dimensionaler Vektorraum über $\mathbb{K}$ und $\phi : V \to V$ ein nicht trigonalisierbarer Endomorphismus. Dann existiert eine geordnete Basis B von V mit

$$_BM_B(\phi) = \begin{pmatrix} M & 0_{p\times q} \\ 0_{q\times p} & A \end{pmatrix}$$

für eine obere Dreiecksmatrix $M \in \mathrm{Mat}(p \times p, \mathbb{K})$ und eine Matrix $A \in \mathrm{Mat}(q \times q, \mathbb{K})$ ohne Eigenwerte in $\mathbb{K}$.

Beweis:
Hat ϕ paarweise verschiedene Eigenwerte $\lambda_1, .., \lambda_m$ so gilt $V = H(\phi, \lambda_1) \oplus ... \oplus H(\phi, \lambda_m) \oplus W$ nach Satz 7.1.8, und die Untervektorräume $H := H(\phi, \lambda_1) \oplus ... \oplus H(\phi, \lambda_m)$ und W sind ϕ-invariant: $\phi(H) \subseteq H$ und $\phi(W) \subseteq W$. Daher enthält man durch Einschränken von ϕ Endomorphismen $\phi|_H : H \to H$ und $\phi|_W : W \to W$.

Der Endomorphismus $\phi|_H : H \to H$ ist nach Satz 7.2.1 trigonalisierbar. Also existiert eine geordnete Basis C von H, so dass $M := {}_CM_C(\phi|_H)$ eine obere Dreiecksmatrix ist.

Da $E(\phi, \lambda_i) \subseteq H(\phi, \lambda_i) \subseteq H$ für alle Eigenwerte λ_i von ϕ, hat $\phi|_W : W \to W$ keine Eigenwerte und somit gilt das auch für die darstellende Matrix $A := {}_DM_D(\phi|_W)$ bezüglich einer beliebigen geordneten Basis D von W. Da $V = H \oplus W$ ist $B = (C, D)$ eine geordnete Basis von V mit

$$_BM_B(\phi) = \begin{pmatrix} {}_CM_C(\phi|_H) & 0_{n\times p} \\ 0_{p\times n} & {}_DM_D(\phi|_W) \end{pmatrix} = \begin{pmatrix} M & 0_{p\times q} \\ 0_{q\times p} & A \end{pmatrix}. \qquad \square$$

Insbesondere folgt aus Satz 7.1.8 und Korollar 7.2.2, dass jeder Endomorphismus eines endlich-dimensionalen Vektorraums über $\mathbb{C}$ und jede Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{C})$ trigonalisierbar ist. Denn der Endomorphismus $\phi|_W : W \to W$ in Satz 7.1.8 und die Matrix A in Korollar 7.2.2 besitzen keine Eigenwerte in $\mathbb{K}$. Da die charakteristische Gleichung $\det(A - \lambda \mathrm{id}_V) = 0$ eine nicht-konstante polynomiale Gleichung in λ ist, folgt aber aus dem Fundamentalsatz der Algebra, dass sie für jede Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{C})$ eine Lösung $\lambda \in \mathbb{C}$ besitzt. Also kann in diesem Fall die Matrix A aus Korollar 7.2.2 nicht auftreten, und in Satz 7.1.8 ist $W = \{0\}$.

Analoge Aussagen gelten für alle Körper $\mathbb{K}$, in denen jedes nicht-konstante Polynom mindestens eine Nullstelle hat. Dies ist äquivalent zu der Bedingung, dass jedes nicht-konstante Polynom $p \in \mathbb{K}[x]$ in Linearfaktoren zerfällt, denn Nullstellen können induktiv abfaktorisiert werden. Solche Körper $\mathbb{K}$ bezeichnet man als *algebraisch abgeschlossen*.

Definition 7.2.3: Ein Körper $\mathbb{K}$ heißt **algebraisch abgeschlossen**, wenn jedes nicht-konstante Polynom $p \in \mathbb{K}[x]$ eine Nullstelle in $\mathbb{K}$ besitzt oder, dazu äquivalent, jedes nicht-konstante Polynom $p \in \mathbb{K}[x]$ in Linearfaktoren zerfällt.

Beispiel 7.2.4:

1. Der Körper $\mathbb{C}$ ist nach dem Fundamentalsatz der Algebra algebraisch abgeschlossen.
2. Die Körper $\mathbb{R}$ und $\mathbb{Q}$ sind nicht algebraisch abgeschlossen, denn das Polynom $x^2 + 1 \in \mathbb{Q}[x] \subseteq \mathbb{R}[x]$ zerfällt über $\mathbb{Q}$ und $\mathbb{R}$ nicht in Linearfaktoren.
3. Die endlichen Körper $\mathbb{Z}/p\mathbb{Z}$ sind nicht algebraisch abgeschlossen (siehe *Vorlesung Algebra*).
4. Man kann zeigen (siehe *Vorlesung Körpertheorie*), dass zu jedem Körper $\mathbb{K}$ ein algebraisch abgeschlossener Körper $\mathbb{L}$ existiert, der $\mathbb{K}$ als Teilkörper enthält. Da jedes Polynom $p \in \mathbb{K}[x]$ auch in $\mathbb{L}[x]$ enthalten ist, zerfällt es dann über $\mathbb{L}$ in Linearfaktoren.

Die Betrachtungen vor Definition 7.2.3 garantieren Trigonalisierbarkeit für Endomorphismen endlich-dimensionaler Vektorräume und Matrizen über algebraisch abgeschlossenen Körpern.

Korollar 7.2.5: Sei $\mathbb{K}$ ein algebraisch abgeschlossener Körper.
Dann ist jeder Endomorphismus $\phi : V \to V$ eines endlich-dimensionalen $\mathbb{K}$-Vektorraums V und jede Matrix $M \in \mathrm{Mat}(n \times n, \mathbb{K})$ trigonalisierbar.

Ist $\mathbb{K}$ nicht algebraisch abgeschlossen, so können wir mit Satz 7.1.8 und Satz 7.2.1 die Frage der Trigonalisierbarkeit einer Matrix $M \in \mathrm{Mat}(n \times n, \mathbb{K})$ ebenfalls vollständig klären:

1. Wir bestimmen ihre Eigenwerte $\lambda_1, ..., \lambda_m$ und die Matrizen $M_{\lambda_i} = M - \lambda_i \mathbb{1}_n$.
2. Mit dem Gauß-Verfahren bestimmen wir die Dimensionen der Kerne $\ker(M_{\lambda_i})$.
3. Summieren sie sich zu n, so ist M diagonalisierbar.
4. Wenn nicht, betrachten wir die Kerne höherer Potenzen $M^k_{\lambda_i}$.
5. Ab einem bestimmten Exponenten, dem Fitting-Index f_i von M_{λ_i}, ändern sich diese nicht mehr. Dann gilt $\dim_{\mathbb{K}} \ker(M^{f_i}_{\lambda_i}) = \dim_{\mathbb{K}} H(M, \lambda_i)$.
6. Addieren sich diese Dimensionen zu n, so ist die Matrix trigonalisierbar.
7. Wenn nicht, kann man sie in eine trigonalisierbare Matrix und eine Matrix ohne Eigenwerte zerlegen, wie in Korollar 7.2.2.

Alternativ kann man nach Beispiel 7.2.4, 3. auch zu einem algebraisch abgeschlossenen Körper $\mathbb{L}$ mit $\mathbb{K} \subseteq \mathbb{L}$ übergehen, und eine gegebene Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{K}) \subseteq \mathrm{Mat}(n \times n, \mathbb{L})$ dort trigonalisieren. Die resultierende obere Dreiecksmatrix kan dann aber Einträge aus $\mathbb{L} \setminus \mathbb{K}$ haben.

Satz 7.1.8 und 7.2.1 erlauben es außerdem, jeden trigonalisierbaren Endomorphismus $\phi : V \to V$ in einen diagonalisierbaren Anteil $\phi_D : V \to V$ und einen nilpotenten Anteil $\phi_N : V \to V$ zu zerlegen, so dass ϕ_D und ϕ_N kommutieren. Der diagonalisierbare Anteil erfüllt $\phi_D(v) = \lambda v$ und der nilpotente Anteil $\phi_N(v) = \phi_\lambda(v) = \phi(v) - \lambda v$ für alle $v \in H(\phi, \lambda)$.

Satz 7.2.6: (**Jordan-Zerlegung**)
Sei V ein endlich-dimensionaler $\mathbb{K}$-Vektorraum und $\phi : V \to V$ ein trigonalisierbarer Endomorphismus. Dann gibt es eindeutig bestimmte Endomorphismen $\phi_N, \phi_D : V \to V$ mit $\phi = \phi_N + \phi_D$, so dass ϕ_D diagonalisierbar ist, ϕ_N nilpotent und $\phi_D \circ \phi_N = \phi_N \circ \phi_D$.

Der Endomorphismus ϕ_D heißt **diagonalisierbare Jordan-Komponente** von ϕ und der Endomorphismus ϕ_N **nilpotente Jordan-Komponente** von ϕ.

Beweis:
1. **Existenz:** Ist ϕ trigonalisierbar mit paarweise verschiedenen Eigenwerten $\lambda_1, ..., \lambda_m$, so gilt $V = H(\phi, \lambda_1) \oplus ... \oplus H(\phi, \lambda_m)$ nach Satz 7.1.8 und 7.2.1. Also lässt sich jeder Vektor $v \in V$ eindeutig als Summe $v = v_1 + ... + v_m$ mit $v_i \in H(\phi, \lambda_i)$ schreiben. Wir definieren

$$\begin{aligned} &\phi_D : V \to V, && v_1 + \ldots + v_m \mapsto \lambda_1 v_1 + \ldots + \lambda_m v_m, \\ &\phi_N = \phi - \phi_D : V \to V, && v_1 + \ldots + v_m \mapsto (\phi(v_1) - \lambda_1 v_1) + \ldots + (\phi(v_m) - \lambda_m v_m). \end{aligned}$$

Dann gilt per Definition $\phi = \phi_D + \phi_N$, und ϕ_D ist diagonalisierbar mit $\phi_D|_{H(\phi,\lambda_i)} = \lambda_i \,\mathrm{id}_{H(\phi,\lambda_i)}$. Da $\phi(H(\phi, \lambda_i)) \subseteq H(\phi, \lambda_i)$, folgt auch $\phi_N(H(\phi, \lambda_i)) \subseteq H(\phi, \lambda_i)$. Nach Korollar 7.1.7 sind die Endomorphismen $\phi_N|_{H(\phi,\lambda_i)} = (\phi - \lambda_i \mathrm{id}_V)|_{H(\phi,\lambda_i)}$ nilpotent. Ist f_i der Fitting-Index von $(\phi - \lambda_i \mathrm{id}_V)|_{H(\phi,\lambda_i)}$, so gilt $\phi_N^k = 0$ für $k \geq \max\{f_1, ..., f_m\}$, und damit ist auch ϕ_N nilpotent. Außerdem gilt $\phi_D \circ \phi_N = \phi_N \circ \phi_D$, denn $V = H(\phi, \lambda_1) \oplus ... \oplus H(\phi, \lambda_m)$ und für $u \in H(\phi, \lambda_i)$ ist

$$\phi_D \circ \phi_N(u) = \phi_D \circ \phi(u) - \lambda_i \phi_D(u) = \lambda_i \phi(u) - \lambda_i^2 u = \lambda_i \phi_N(u) = \phi_N(\lambda_i u) = \phi_N \circ \phi_D(u).$$

2. **Eindeutigkeit:** Sei ϕ trigonalisierbar mit paarweise verschiedenen Eigenwerten $\lambda_1, ..., \lambda_m$ und $\phi = \psi_D + \psi_N$ eine weitere Zerlegung, so dass ψ_D diagonalisierbar ist, ψ_N nilpotent und $\psi_N \circ \psi_D = \psi_D \circ \psi_N$. Sind $\mu_1, ..., \mu_r$ die paarweise verschiedenen Eigenwerte von ψ_D, so gilt wegen der Diagonalisierbarkeit $H(\psi_D, \mu_i) = E(\psi_D, \mu_i)$ und $V = E(\psi_D, \mu_1) \oplus ... \oplus E(\psi_D, \mu_r)$.

Da $\psi_N \circ \psi_D = \psi_D \circ \psi_N$ ist $\psi_N(E(\psi_D, \mu_i)) \subseteq E(\psi_D, \mu_i)$ nach Korollar 7.1.7 (ii), und man erhält nilpotente Endomorphismen $\psi_N|_{E(\psi_D,\mu_i)} : E(\psi_D, \mu_i) \to E(\psi_D, \mu_i)$.

Also gibt es einen Vektor $v \in \ker(\psi_N) \cap E(\psi_D, \mu_i)$ mit $v \neq 0$. Dieser ist ein Eigenvektor von ϕ zum Eigenwert μ_i, denn $\phi(v) = \psi_D(v) + \psi_N(v) = \psi_D(v) = \mu_i v$. Also ist jeder Eigenwert von ψ_D auch ein Eigenwert von ϕ und somit $\{\mu_1, ..., \mu_r\} \subseteq \{\lambda_1, ..., \lambda_m\}$.

Da $\phi_N|_{E(\psi_D,\mu_i)} = \phi - \mu_i \,\mathrm{id}_V|_{E(\psi_D,\mu_i)}$ nilpotent ist, folgt $E(\psi_D, \mu_i) \subseteq H(\phi, \mu_i)$ und damit $V = E(\psi_D, \mu_1) \oplus ... \oplus E(\psi_D, \mu_r) \subseteq H(\phi, \mu_1) \oplus ... \oplus H(\phi, \mu_r) \subseteq V$. Also gilt $E(\psi_D, \mu_i) = H(\phi, \mu_i)$ und $\{\lambda_1, ..., \lambda_m\} = \{\mu_1, ..., \mu_r\}$. Daraus ergibt sich dann $\psi_D|_{H(\phi,\lambda_i)} = \psi_D|_{E(\psi_D,\lambda_i)} = \lambda_i \,\mathrm{id}_{H(\phi,\lambda_i)}$, und es folgt $\psi_D = \phi_D$ sowie $\psi_N = \phi - \psi_D = \phi - \phi_D = \phi_N$. □

Um die Jordan-Zerlegung eines Endomorphismus oder einer Matrix zu bestimmen, muss man nach dem Beweis von Satz 7.2.6 im Allgemeinen eine Basis des Hauptraums zu jedem Eigenwert bestimmen. Dies ist äquivalent dazu, eine Matrix durch Konjugation mit einer invertierbaren Matrix in eine Blockdiagonalmatrix zu überführen, die aus Blöcken von trigonalisierbaren Matrizen besteht, die jeweils nur einen einzigen Eigenwert besitzen.

Ist eine Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{K})$ bereits eine Blockdiagonalmatrix von dieser Form, so ist die Bestimmung der Jordan-Zerlegung einfach. Indem man von den Blöcken die mit den zugehörigen Eigenwerten multiplizierte Einheitsmatrix subtrahiert, erhält man den nilpotenten Anteil der Jordan-Zerlegung. Den Diagonalanteil erhält man, indem man die mit den Eigenwerten multiplizierten Einheitsmatrizen für die einzelnen Blöcke in eine Diagonalmatrix kombiniert.

Beispiel 7.2.7: Für den Endomorphismus

$$\phi : \mathbb{R}^4 \to \mathbb{R}^4, \qquad \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \mapsto \underbrace{\begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 2 & 0 & 0 \\ 0 & 0 & 5 & -1 \\ 0 & 0 & 1 & 3 \end{pmatrix}}_{=:M} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}.$$

erhält man wegen der Blockdiagonalität von M die charakteristische Gleichung

$$\det(M - \lambda \mathbb{1}_4) = \det \begin{pmatrix} -\lambda & 1 \\ -1 & 2-\lambda \end{pmatrix} \det \begin{pmatrix} 5-\lambda & -1 \\ 1 & 3-\lambda \end{pmatrix} = (\lambda-1)^2(\lambda-4)^2 = 0.$$

Damit sind die Eigenwerte 1 und 4 und die zugehörigen Haupträume $H(\phi, 1) = \mathrm{span}_{\mathbb{R}}\{e_1, e_2\}$ und $H(\phi, 4) = \mathrm{span}_{\mathbb{R}}\{e_3, e_4\}$. Die Endomorphismen $\phi_D, \phi_N : \mathbb{R}^3 \to \mathbb{R}^3$ sind gegeben durch

$$\phi_D : \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}, \quad \phi_N : \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \mapsto \begin{pmatrix} -1 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & -1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}.$$

Der diagonalisierbare Endomorphismus ergibt sich, indem man den Eigenwert 1 in die ersten zwei und den Eigenwert 4 in die letzten zwei Diagonaleinträge einer Diagonalmatrix einfügt. Der Endomorphismus ϕ_N ergibt sich, indem man ϕ_D von ϕ subtrahiert.

An was man sich erinnern sollte:

Die wichtigsten Begriffe und Definitionen:

- Hauptraum, Fitting-Index, nilpotent für Matrix/Endomorphismus,
- zerfallen in Linearfaktoren, algebraisch abgeschlossen
- Jordan-Zerlegung, Jordan-Komponenten.

Die wichtigsten Aussagen:

- Fitting-Lemma, Eigenschaften der Haupträume,
- Summe der Haupträume ist direkt, Hauptraumzerlegung,
- Trigonalisierbarkeitskriterium,
- $\mathbb{K}$ algebraisch abgeschlossen $\Rightarrow$ jd. Matrix/Endomorphismus trigonalisierbar,
- Jordan-Zerlegung.

Die wichtigsten Beispiele:

- **Trigonalisierbarkeit:** nilpotente Endomorphismen, diagonalisierbare Endomorphismen,
- **Trigonalisierbarkeit:** algebraisch abgeschlossene Körper, insbesondere $\mathbb{C}$,
- **Jordan-Zerlegung:** Matrix in Blockform, deren Blöcke den Haupträumen entsprechen.

8 Charakteristisches Polynom und Minimalpolynom

8.1 Polynome

In den letzten Kapiteln haben wir Verfahren kennengelernt, mit denen wir die Dimensionen der Haupträume und Eigenräume eines Endomorphismus $\phi : V \to V$ bestimmen und dessen Trigonalisierbarkeit und Diagonalisierbarkeit untersuchen können.

Man bestimmt dazu zunächst mit der charakteristischen Gleichung die Eigenwerte von ϕ und dann zu jedem Eigenwert λ den Eigenraum $E(\phi, \lambda) = \ker(\phi - \lambda \mathrm{id}_V)$ sowie die Kerne höherer Potzenzen $(\phi - \lambda \mathrm{id}_V)^k$ für $k = 2, 3, ...$, bis diese sich von k zu $k + 1$ nicht mehr ändern. Der letzte Kern ist dann der Hauptraum $H(\phi, \lambda)$. Der Endomorphismus ϕ ist trigonalisierbar (diagonalisierbar) genau dann, wenn die Haupträume (Eigenräume) zu verschiedenen Eigenwerten zusammen den ganzen Vektorraum V aufspannen. Dies ist der Fall genau dann, wenn sich Ihre Dimensionen zur Dimension von V addieren.

Das ist jedoch sehr aufwändig. Um ein weniger rechenintensives Verfahren zur Untersuchung der Trigonalisierbarkeit zu bekommen, befassen wir uns genauer mit der charakteristischen Gleichung $\det(A - \lambda \mathbb{1}_n) = 0$ für $A \in \mathrm{Mat}(n \times n, \mathbb{K})$, die eine polynomiale Gleichung in λ ist.

Dazu wiederholen wir zunächst die wichtigesten Ergebnisse zu Polynomen:

- Polynome wurden in Beispiel 2.3.5 abstrakt als Abbildungen $p : \mathbb{N}_0 \to R$ mit Werten in einem kommutativen unitalen Ring R definiert, die die Bedingung $p(n) = 0$ für alle bis auf endlich viele $n \in \mathbb{N}_0$ erfüllen.

- Mit der Multiplikation und Addition aus Beispiel 2.3.5 bilden diese einen kommutativen unitalen Ring $R[x]$.

- Ist $R = \mathbb{K}$ ein Körper, so bilden die Polynome mit Werten in $\mathbb{K}$ nach Beispiel 3.1.10 auch einen Vektorraum über $\mathbb{K}$, der ein Untervektorraum des Vektorraums der Folgen mit Werten in $\mathbb{K}$ ist. Die Monome bilden eine Basis $B = \{1, x, x^2, \ldots\}$ des Vektorraums $\mathbb{K}[x]$.

- Da Polynommultiplikation und die Skalarmultiplikation verträglich sind, ist $\mathbb{K}[x]$ nach Beispiel 4.1.8 eine kommutative $\mathbb{K}$-Algebra mit Einselement $1 = x^0$ und

$$\begin{aligned} \left(\Sigma_{j\in\mathbb{N}_0} a_j x^j\right) \cdot \left(\Sigma_{k\in\mathbb{N}_0} b_k x^k\right) &= \Sigma_{n\in\mathbb{N}_0}(\Sigma_{k=0}^n a_k b_{n-k}) x^n \\ \left(\Sigma_{j\in\mathbb{N}_0} a_j x^j\right) + \left(\Sigma_{k\in\mathbb{N}_0} b_k x^k\right) &= \Sigma_{n\in\mathbb{N}_0}(a_n + b_n) x^n \\ \lambda\left(\Sigma_{n\in\mathbb{N}_0} a_n x^n\right) &= \Sigma_{n\in\mathbb{N}_0}(\lambda a_n) x^n. \end{aligned}$$

- Für Polynome $p, q \in \mathbb{K}[x]$ erfüllt der Polynomgrad $\deg(p \cdot q) = \deg(p) + \deg(q)$ und $\deg(p + q) \leq \max\{\deg(p), \deg(q)\}$[15]. Insbesondere ist $\mathbb{K}[x]$ nullteilerfrei.

Die *Polynome* haben zunächst also nichts mit den aus der Schule bekannten *Polynomabbildungen* zu tun, sondern sind abstrakt, durch die Folge ihrer Koeffizienten definiert. Tatsächlich kann man aber aus Polynomen Polynomabbildungen konstruieren, indem man sie auf Elementen $\lambda \in \mathbb{K}$ auswertet. Man ersetzt dazu einfach in einem Polynom $p = \Sigma_{k\in\mathbb{N}_0} b_k x^k$ das Monom x^k durch das Element $\lambda^k \in \mathbb{K}$.

[15] Dabei setzt man $-\infty \cdot n = n \cdot -\infty = -\infty \cdot -\infty := -\infty$ für alle $n \in \mathbb{N}$.

Dieses Auswerten von Polynomen funktioniert in einem viel allgemeineren Kontext. Um den Term x in einem Polynom duch ein Element einer Menge A ersetzen zu können, braucht man auf A ein Konzept von Potenzen, also eine Multiplikation. Zusätzlich muss man Elemente von A mit den Koeffizienten $a_k \in \mathbb{K}$ multiplizieren können und Summen bilden können. Damit benötigt man eine Algebra A über $\mathbb{K}$. Das Ersetzen des Monoms x^k durch die entsprechende Potenz a^k eines Algebraelements $a \in A$ liefert dann einen Algebrahomomorphismus $\mathbb{K}[x] \to A$, der durch die Wahl von $a \in A$ eindeutig festgelegt ist.

Satz 8.1.1: (**Universelle Eigenschaft der Polynomalgebra**)
Zu jeder $\mathbb{K}$-Algebra A und jedem Element $a \in A$ existiert genau ein Homomorphismus von $\mathbb{K}$-Algebren $\mathrm{ev}_a : \mathbb{K}[x] \to A$ mit $\mathrm{ev}_a(x) = a$, nämlich

$$\mathrm{ev}_a : \mathbb{K}[x] \to A, \quad \Sigma_{k\in\mathbb{N}_0} b_k x^k \mapsto \Sigma_{k\in\mathbb{N}_0} b_k a^k.$$

Für $p \in \mathbb{K}[x]$ schreiben wir $p(a) := \mathrm{ev}_a(p)$.

Beweis:
Sei $\phi : \mathbb{K}[x] \to A$ ein Homomorphismus von $\mathbb{K}$-Algebren. Da ϕ $\mathbb{K}$-linear ist, ist ϕ durch seine Werte auf den Monomen x^k bereits eindeutig bestimmt, denn diese bilden eine Basis des $\mathbb{K}$-Vektorraums $\mathbb{K}[x]$. Da ϕ außerdem ein Ringhomomorphismus ist, gilt $\phi(x \cdot x) = \phi(x) \cdot \phi(x)$, und induktiv folgt $\phi(x^k) = \phi(x)^k$ für alle $k \in \mathbb{N}$. Da ϕ ein Homomorphismus von *unitalen* Ringen ist, gilt $\phi(1) = 1_A$. Also ist jeder unitale Ringhomomorphismus $\phi : \mathbb{K}[x] \to A$ durch $\phi(x) = a$ bereits eindeutig bestimmt.

Umgekehrt lässt sich zu jedem $a \in A$ ein Homomorphismus $\mathrm{ev}_a : \mathbb{K}[x] \to A$ von $\mathbb{K}$-Algebren konstruieren, indem man einem Polynom $p = \Sigma_{k\in\mathbb{N}_0} b_k x^k$ das Element $\mathrm{ev}_a(p) = \Sigma_{k\in\mathbb{N}_0} b_k a^k$ mit $a^0 = 1_A$ zuordnet. Dieser erfüllt die Bedingungen $\mathrm{ev}_a(1) = 1_A$, $\mathrm{ev}_a(p \cdot q) = \mathrm{ev}_a(p) \cdot \mathrm{ev}_a(q)$, $\mathrm{ev}_a(\lambda p) = \lambda \mathrm{ev}_a(p)$ und $\mathrm{ev}_a(p+q) = \mathrm{ev}_a(p) + \mathrm{ev}_a(q)$ für alle $p, q \in \mathbb{K}[x]$ und $\lambda \in \mathbb{K}$ und ist somit ein Homomorphismus von $\mathbb{K}$-Algebren. □

Beispiel 8.1.2:

1. Ist $A = \mathrm{Mat}(n \times n, \mathbb{K})$, so ist für jede Matrix $M \in \mathrm{Mat}(n \times n, \mathbb{K})$ die Evaluationsabbildung gegeben durch $\mathrm{ev}_M : \mathbb{K}[x] \to \mathrm{Mat}(n \times n, \mathbb{K})$, $\Sigma_{j\in\mathbb{N}_0} a_n x^n \mapsto \Sigma_{n\in\mathbb{N}_0} a_n M^n$.

2. Ist V ein $\mathbb{K}$-Vektorraum und $A = \mathrm{End}_{\mathbb{K}}(V)$, so ordnet $\mathrm{ev}_\phi : \mathbb{K}[x] \to \mathrm{End}_{\mathbb{K}}(V)$ für einen Endomorphismus $\phi \in \mathrm{End}_{\mathbb{K}}(V)$ einem Polynom $p = \Sigma_{j\in\mathbb{N}_0} a_j x^j$ den Endomorphismus $p(\phi) = \mathrm{ev}_\phi(p) = \Sigma_{j\in\mathbb{N}_0} a_j \phi^j$ mit $\phi^j = \phi \circ \ldots \circ \phi$ zu.

3. Wählt man $A = \mathbb{K}[x]$ und $q = \Sigma_{l\in\mathbb{N}_0} b_l x^l \in \mathbb{K}[x]$, so ordnet die Evaluationsabbildung $\mathrm{ev}_q : \mathbb{K}[x] \to \mathbb{K}[x]$, $p \mapsto p(q)$ einem Polynom $p = \Sigma_{k\in\mathbb{N}_0} a_k x^k \in \mathbb{K}[x]$ das Polynom $p(q) = \Sigma_{k\in\mathbb{N}_0} a_k (\Sigma_{l\in\mathbb{N}_0} b_l x^l)^k$ zu, das durch Einsetzen von q in p entsteht.

4. Für $A = \mathbb{K}$ ordnet die Evaluationsabbildung $\mathrm{ev}_\lambda : \mathbb{K}[x] \to \mathbb{K}$, $p \mapsto p(\lambda)$ einem Polynom $p \in \mathbb{K}[x]$ seine Auswertung an der Stelle $\lambda \in \mathbb{K}$ zu. Die Abbildung $f_p : \mathbb{K} \to \mathbb{K}$, $\lambda \mapsto p(\lambda)$ ist dann die zugehörige Polynomabbildung zu p.

   Die Abbildung $f : \mathbb{K}[x] \to \mathrm{Pol}(\mathbb{K})$, $p \mapsto f_p$, die einem Polynom $p \in \mathbb{K}[x]$ die zugehörige Polynomabbildung $f_p : \mathbb{K} \to \mathbb{K}$ in $\mathrm{Pol}(\mathbb{K}) \subseteq \mathrm{Abb}(\mathbb{K}, \mathbb{K})$ zuordnet, ist ein Homomorphismus von $\mathbb{K}$-Algebren. Dieser ist injektiv genau dann, wenn $\mathbb{K}$ unendlich ist. (Beweis: Übung).

Die Tatsache, dass der Homomorphismus von $\mathbb{K}$-Algebren $f : \mathbb{K}[x] \to \mathrm{Pol}(\mathbb{K})$ aus Beispiel 8.1.2, 4. für endliche Körper nicht injektiv ist, ist der Grund, warum man zwischen Polynomen und Polynomabbildungen unterscheiden muss. So sind beispielsweise die Polynome x und x^2 in $\mathbb{Z}/2\mathbb{Z}[x]$ verschieden. Die Polynomabbildungen $f_x, f_{x^2} : \mathbb{Z}/2\mathbb{Z} \to \mathbb{Z}/2\mathbb{Z}$ sind aber gleich, denn

$$f_x([0]) = [0] = [0] \cdot [0] = f_{x^2}([0]) \qquad f_x([1]) = [1] = [1] \cdot [1] = f_{x^2}([1]).$$

Dass in unendlichen Körpern solche Probleme nicht auftreten liegt daran, dass ein Polynom vom Grad $n \in \mathbb{N}_0$ dort maximal n Nullstellen hat, was $\ker(f) = \{0\}$ impliziert. Um dies zu präzisieren und beweisen, definieren wir zunächst, was eine Nullstelle ist und befassen uns dann mit dem Abspalten von Nullstellen. Dies lässt sich mit Hilfe der Auswertungsabbildung abstrakt für allgemeine Polynome formulieren.

Definition 8.1.3: Eine **Nullstelle** oder **Wurzel** eines Polynoms $p \in \mathbb{K}[x]$ ist ein Element $\lambda \in \mathbb{K}$ mit $p(\lambda) = \mathrm{ev}_\lambda(p) = 0$.

Satz 8.1.4: Ist $\lambda \in \mathbb{K}$ eine Nullstelle von $p \in \mathbb{K}[x]$, so existiert ein eindeutig bestimmtes Polynom $q \in \mathbb{K}[x]$ vom Grad $\deg(q) = \deg(p) - 1$ mit $p = (x - \lambda) \cdot q$. Insbesondere kann ein Polynom vom Grad $n \in \mathbb{N}_0$ maximal n Nullstellen besitzen.

Beweis:
Mit der binomischen Formel ergibt sich für alle $n \in \mathbb{N}_0$

$$x^n = (x - \lambda + \lambda)^n = \sum_{k=0}^{n} \binom{n}{k} \lambda^{n-k}(x-\lambda)^k = \sum_{k=0}^{n} \frac{n!\lambda^{n-k}(x-\lambda)^k}{k!(n-k)!}.$$

Für $p = \Sigma_{n=0}^m a_n x^n$ ergibt sich daraus durch Einsetzen

$$p = \sum_{n=0}^{m} a_n x^n = \sum_{n=0}^{m} a_n \left(\sum_{k=0}^{n} \frac{n!\lambda^{n-k}(x-\lambda)^k}{k!(n-k)!} \right) \overset{(*)}{=} \sum_{k=0}^{m} \left(\sum_{n=k}^{m} \frac{a_n \lambda^{n-k} n!}{k!(n-k)!} \right) (x-\lambda)^k = \sum_{k=0}^{m} b_k (x-\lambda)^k,$$

wobei in $(*)$ eine Umordnung der Summe vorgenommen wurde. Ist $p(\lambda) = 0$, so ergibt sich daraus $0 = \Sigma_{k=0}^m b_k(\lambda - \lambda)^k = b_0 \cdot 1 = \Sigma_{n=0}^m a_n \lambda^n$, und man erhält

$$p = \sum_{k=1}^{m} b_k (x-\lambda)^k = (x-\lambda) \cdot \sum_{k=1}^{m} b_k (x-\lambda)^{k-1} = (x-\lambda) \cdot q \quad \text{mit} \quad \deg(q) = \deg(p) - 1.$$

Dass ein Polynom p vom Grad n höchstens n verschiedene Nullstellen besitzt, ergibt sich dann durch vollständige Induktion über den Polynomgrad:

$n = 0$: Ist $\deg(p) = 0$, so ist p konstant, $p \neq 0$ und besitzt damit keine Nullstellen.

$n - 1 \to n :$ Sei die Aussage bewiesen für alle Polynome $q \in \mathbb{K}[x]$ vom Grad $\deg(q) \leq n - 1$ und sei $p \in \mathbb{K}[x]$ ein Polynom vom Grad $\deg(p) = n$. Ist $\lambda \in \mathbb{K}$ eine Nullstelle von p, so existiert ein eindeutig bestimmtes Polynom $q \in \mathbb{K}[x]$ mit $\deg(q) = n - 1$ und $p = (x - \lambda) \cdot q$. Nach Induktionsvoraussetzung hat q höchstens $n - 1$ Nullstellen. Da $x - \lambda \neq 0$ für alle $x \in \mathbb{K} \setminus \{\lambda\}$ hat p damit höchstens n Nullstellen. □

Insbesondere kann man von einem Polynom $p \in \mathbb{K}[x]$ eine gegebene Nullstelle λ von p so oft abspalten, dass ein Polynom q mit $q(\lambda) \neq 0$ zurückbleibt. Dies führt auf das Konzept der *Vielfachheit* einer Nullstelle.

Korollar 8.1.5: Sei $p \in \mathbb{K}[x]$ ein Polynom und $\lambda \in \mathbb{K}$ eine Nullstelle von p. Dann existieren eindeutig bestimmte $k \in \mathbb{N}$ und $q \in \mathbb{K}[x]$ mit $q(\lambda) \neq 0$, $\deg(q) = \deg(p) - k$ und $p = (x - \lambda)^k \cdot q$. Die Zahl k heißt **Vielfachheit** der Nullstelle λ.

Beweis:
Ist $\lambda \in \mathbb{K}$ eine Nullstelle von $q_0 := p$, so gibt es nach Satz 8.1.4 ein eindeutig bestimmtes Polynom $q_1 \in \mathbb{K}[x]$ mit $\deg(q_1) = \deg(q_0) - 1$ und $q_0 = (x-\lambda)q_1$. Ist $q_1(\lambda) \neq 0$, so ist die Aussage bewiesen. Ansonsten gibt es nach Satz 8.1.4 ein eindeutig bestimmtes Polynom $q_2 \in \mathbb{K}[x]$ mit $\deg(q_2) = \deg(q_1) - 1$ und $q_1 = (x - \lambda)q_2$. Iterativ konstruiert man Polynome $q_j \in \mathbb{K}[x]$ mit $q_j = (x - \lambda)q_{j-1}$ und $\deg(q_j) = \deg(q_{j-1}) - 1 = \deg(p) - j$. Nach maximal $k = \deg(p)$ Schritten bricht das Verfahren ab, und man erhält ein Polynom q_k mit $p = (x - \lambda)^k \cdot q_k$, $q_k(\lambda) \neq 0$ und $\deg(q_k) = \deg(p) - k$. □

Indem man Korollar 8.1.5 nacheinander auf die verschiedenen Nullstellen $\lambda_i \in \mathbb{K}$ eines Polynoms $p \in \mathbb{K}[x]$ anwendet, erhält man eine Faktorisierung von p in ein Polynom q ohne Nullstellen in $\mathbb{K}$ und Faktoren $(x - \lambda_i)^{k_i}$ für die verschiedenen Nullstellen λ_i von p.

Korollar 8.1.6: Sei $p \in \mathbb{K}[x]$ ein Polynom mit paarweise verschiedenen Nullstellen $\lambda_1, ..., \lambda_m \in \mathbb{K}$ und $k_1, ..., k_m \in \mathbb{N}$ ihre Vielfachheiten.

Dann gibt es ein Polynom $q \in \mathbb{K}[x]$ vom Grad $\deg(q) = \deg(p) - k_1 - ... - k_m$ ohne Nullstellen in $\mathbb{K}$ mit $p = (x-\lambda_1)^{k_1} \cdots (x-\lambda_m)^{k_m} \cdot q$. Ist $\deg(q) = 0$ so sagt man, p **zerfalle in Linearfaktoren**.

Ob ein gegebenes Polynom $p \in \mathbb{K}[x]$ über $\mathbb{K}$ in Linearfaktoren zerfällt, hängt offensichtlich davon ab, ob es Nullstellen in $\mathbb{K}$ besitzt und ob sich die Vielfachheiten dieser Nullstellen zu seinem Grad addieren. In algebraisch abgeschlossenen Körpern wie $\mathbb{K} = \mathbb{C}$ ist dies immer der Fall.

Aber auch in dem nicht algebraisch abgeschlossenen Körper $\mathbb{R}$ zerfallen Polynome weitgehend, und zwar in lineare oder quadratische Faktoren. Dies liegt daran, dass für jede Nullstelle $\mu \in \mathbb{C} \setminus \mathbb{R}$ von p auch die komplex konjugierte Zahl $\bar{\mu} \in \mathbb{C} \setminus \mathbb{R}$ eine Nullstelle ist. Die entsprechenden Linearfaktoren kombinieren sich dann zu einem quadratischen Polynom in $\mathbb{R}[x]$.

Lemma 8.1.7: Jedes Polynom $p \in \mathbb{R}[x]$ zerfällt in lineare und quadratische Faktoren:

$$p = a(x - \lambda_1)^{k_1} \cdots (x - \lambda_m)^{k_m}(x^2 - 2\mathrm{Re}\,(\mu_1)x + |\mu_1|^2)^{l_1} \cdots (x^2 - 2\mathrm{Re}\,(\mu_r)x + |\mu_r|^2)^{l_r}$$

wobei $a \in \mathbb{R}$, $\lambda_1, ..., \lambda_m \in \mathbb{R}$ die reellen Nullstellen von p mit Vielfachheiten $k_1, ..., k_m$ sind und $\mu_1, ..., \mu_r, \bar{\mu}_1, ..., \bar{\mu}_r$ die Nullstellen in $\mathbb{C} \setminus \mathbb{R}$ mit Vielfachheiten $l_1, ..., l_r$.

Beweis:
Ist $p \in \mathbb{R}[x]$ mit paarweise verschiedenen Nullstellen $\lambda_1, ..., \lambda_m$ und Vielfachheiten $k_1, ..., k_m$, so gibt es nach Korollar 8.1.6 ein Polynom $q = \Sigma_{k\in\mathbb{N}_0} a_k\, x^k \in \mathbb{R}[x]$ ohne Nullstellen in $\mathbb{R}$ mit $p = (x - \lambda_1)^{k_1} \cdots (x - \lambda_m)^{k_m} \cdot q$. Da $q \in \mathbb{R}[x] \subseteq \mathbb{C}[x]$, zerfällt q über $\mathbb{C}$ in Linearfaktoren. Ist $\mu \in \mathbb{C} \setminus \mathbb{R}$ eine Nullstelle von q, so ist auch $\overline{\mu} \in \mathbb{C} \setminus \mathbb{R}$ eine Nullstelle, denn

$$\Sigma_{k\in\mathbb{N}_0} a_k\, \mu^k = q(\mu) = 0 = \overline{q(\mu)} = \overline{\Sigma_{k\in\mathbb{N}_0} a_k\, \mu^k} = \Sigma_{k\in\mathbb{N}_0} \overline{a_k}\, \overline{\mu}^k \stackrel{a_k \in \mathbb{R}}{=} \Sigma_{k\in\mathbb{N}_0} a_k\, \overline{\mu}^k = q(\overline{\mu}).$$

Also ist q von der Form $q = (x - \mu)(x - \bar{\mu}) \cdot r = (x^2 - (\mu + \bar{\mu})x + \mu\bar{\mu}) \cdot r = (x^2 - 2\mathrm{Re}(\mu)x + |\mu|^2) \cdot r$ mit $r \in \mathbb{R}[x]$. Da q über $\mathbb{C}$ in Linearfaktoren zerfällt, folgt mit Korollar 8.1.6 die Behauptung.
□

Neben dem Abspalten von Nullstellen hat man im Polynomring auch ein Konzept von Division mit Rest, die *Polynomdivision*. Dies funktioniert ähnlich wie die Division mit Rest in $\mathbb{Z}$. Die zugrundeliegende Größe, die dies ermöglicht, ist der Polynomgrad. Er spielt bei der Polynomdivision eine ähnliche Rolle wie der Betrag einer ganzen Zahl bei der Division mit Rest in $\mathbb{Z}$.

Satz 8.1.8: (**Polynomdivision**)
Zu zwei Polynomen $p, q \in \mathbb{K}[x]$ mit $q \neq 0$ existieren eindeutig bestimmte Polynome $r, s \in \mathbb{K}[x]$ mit $p = s \cdot q + r$ und $\deg(r) < \deg(q)$. Das Polynom r heißt **Rest** des Polynoms p bei der Division durch q. Ist $r = 0$, so nennt man q einen **Teiler** von p und p ein **Vielfaches** von q.

Beweis:
Eindeutigkeit: Gilt $p = s \cdot q + r = s' \cdot q + r'$ mit $s, s', r, r' \in \mathbb{K}[x]$ und $\deg(r), \deg(r') < \deg(q)$, so folgt $(s - s') \cdot q = r - r'$. Ist $s - s' \neq 0$ ergibt sich daraus

$$\max\{\deg(r), \deg(r')\} \geq \deg(r - r') = \deg((s - s') \cdot q) = \deg(q) + \deg(s - s') \geq \deg(q)$$

ein Widerspruch zur Annahme. Also muss $s - s' = 0$ gelten, und das impliziert $r - r' = 0 \cdot q = 0$.

Existenz: Induktion über den Polynomgrad $\deg(p)$.
$n = -\infty$: Ist $\deg(p) = -\infty$, so ist $p = 0$ und es gilt $0 = p = 0 \cdot q + 0$ für beliebige $q \in \mathbb{K}[x] \setminus \{0\}$.

$n = 0$: Ist $\deg(p) = \deg(q) = 0$, so folgt $p = \lambda \cdot q + 0$ für ein $\lambda \in \mathbb{K}$. Für $0 = \deg(p) < \deg(q)$ ergibt sich $p = 0 \cdot q + p$.

$n - 1 \to n$: Sei die Aussage nun bewiesen für alle Polynome $p \in \mathbb{K}[x]$ mit $\deg(p) \leq n - 1$ und beliebige Polynome $q \in \mathbb{K}[x] \setminus \{0\}$. Sei $p = \Sigma_{k=0}^{n} a_k x^k$ mit $a_n \neq 0$ und $q = \Sigma_{j=0}^{m} b_j x^j$ mit $b_m \neq 0$. Ist $n = \deg(p) < m = \deg(q)$, so ist $p = 0 \cdot q + p$ mit $\deg(p) < \deg(q)$ und wir können $s = 0$ und $r = p$ wählen. Ist $n \geq m$ so betrachten wir das Polynom $p' = p - a_n b_m^{-1} x^{n-m} \cdot q$. Dann gilt $\deg(p') \leq n - 1$ und nach Induktionsvoraussetzung existieren eindeutig bestimmte $s, r \in \mathbb{K}[x]$ mit $r < \deg(q)$ und $p' = s \cdot q + r$. Daraus ergibt sich

$$p = p' + a_n b_m^{-1} x^{n-m} \cdot q = s \cdot q + r + a_n b_m^{-1} x^{n-m} \cdot q = (s + a_n b_m^{-1} x^{n-m}) \cdot q + r = s' \cdot q + r$$

mit $s' = s + a_n b_m^{-1} x^{n-m}, r \in \mathbb{K}[x]$ und $\deg(r) < q$. □

Der Beweis von Satz 8.1.8 liefert nicht nur eine Existenz- und Eindeutigkeitsaussage für die Polynomdivision, sondern auch einen Algorithmus zur Bestimmung der Polynome $r, s \in \mathbb{K}[x]$ mit $p = s \cdot q + r$ und $\deg(r) < \deg(q)$.

1. Start: Man setzt $p_0 = p$.
2. Schritt 1: Man prüft ob $\deg(p_0) < \deg(q)$. Ist das der Fall, so bricht der Algorithmus ab, und man erhält $p_0 = 0 \cdot q + p_0$. Ist $\deg(p_0) \geq q$ setzt man

$$p_1 := p_0 - l(p_0) l(q)^{-1} x^{\deg(p_0) - \deg(q)} \cdot q$$

   wobei $l(p_0)$ und $l(q)$ die Leitkoeffizienten von p_0 und q bezeichnen.
3. Schritt k: Man wendet Schritt 1. auf das Polynom p_{k-1} an. Ist $\deg(p_{k-1}) < \deg(q)$, so bricht der Algorithmus ab und $p_{k-1} = 0 \cdot q + p_{k-1}$. Ist $\deg(p_{k-1}) \geq q$ setzt man

$$p_k := p_{k-1} - l(p_{k-1}) l(q)^{-1} x^{\deg(p_{k-1}) - \deg(q)} \cdot q.$$

Dies liefert eine Folge $p = p_0, p_1, ..., p_n$ von Polynomen mit $\deg(p_0) > \deg(p_1) > ... > \deg(p_n)$ und $\deg(p_n) < \deg(q)$. Es folgt

$$p = s \cdot q + r \quad \text{mit} \quad r = p_n, \quad q \cdot s = p_0 - p_1 + p_1 - p_2 + ... + p_{n-1} - p_n = p_0 - p_n.$$

Beispiel 8.1.9: Wir betrachten die reellen Polynome

$$p = 24x^5 - 24x^4 + 16x^3 + 8x - 32 \qquad q = 2x^3 - x^2 + 1.$$

Dann liefert der Algorithmus:

$$\begin{aligned}
p_0 &= p = 24x^5 - 24x^4 + 16x^3 + 8x - 32 \\
p_1 &= p_0 - 12x^2 \cdot q = 24x^5 - 24x^4 + 16x^3 + 8x - 32 - 24x^5 + 12x^4 - 12x^2 \\
&= -12x^4 + 16x^3 - 12x^2 + 8x - 32 \\
p_2 &= p_1 - (-6x) \cdot q = -12x^4 + 16x^3 - 12x^2 + 8x - 32 + 12x^4 - 6x^3 + 6x \\
&= 10x^3 - 12x^2 + 14x - 32 \\
p_3 &= p_2 - 5 \cdot q = 10x^3 - 12x^2 + 14x - 32 - 10x^3 + 5x^2 - 5 \\
&= -7x^2 + 14x - 37
\end{aligned}$$

$$\begin{aligned}
\Rightarrow \quad & r = p_3 = -7x^2 + 14x - 37 \\
& s = (p_0 - p_3)/q = 12x^2 - 6x + 5.
\end{aligned}$$

Also erhält man $p = (12x^2 - 6x + 5) \cdot q + (-7x^2 + 14x - 37)$.

Bemerkung 8.1.10:

1. Die Konzepte von Teilern und Vielfachen lassen sich in jedem kommutativen nullteilerfreien Ring R mit Eins definieren: Ein Element $b \in R \setminus \{0\}$ heißt **Teiler** eines Elements $a \in R$ und a heißt **Vielfaches** von b, wenn es ein Element $s \in R$ mit $a = s \cdot b$ gibt.

2. Das Konzept der **Division mit Rest**, lässt sich allgemeiner für **euklidische Ringe** definieren. Dies sind Integritätsbereiche, also kommutative nullteilerfreie unitale Ringe R, zusammen mit einer Abbildung $\deg : R \setminus \{0\} \to \mathbb{N}_0$, so dass es für alle $a \in R$, $b \in R \setminus \{0\}$ Elemente $r, s \in R$ gibt mit $a = s \cdot b + r$ und $r = 0$ oder $\deg(r) < \deg(b)$.

   Wichtige Beispiele sind der Ring $\mathbb{Z}$ mit $\deg : \mathbb{Z} \setminus \{0\} \to \mathbb{N}$, $z \mapsto |z|$, der Polynomring $\mathbb{K}[x]$ mit dem Polynomgrad sowie die **Gaußschen Zahlen** $\mathbb{Z}(\mathrm{i}) = \{a + \mathrm{i}b | a, b \in \mathbb{Z}\}$ mit $\deg : \mathbb{Z}(i) \setminus \{0\} \to \mathbb{N}$, $\deg(a + \mathrm{i}b) = a^2 + b^2$.

Eine abstraktere Sicht auf Vielfache und Teiler in einem kommutativen Ring R mit Eins, die sich mit kleinen Änderungen auch auf nicht-kommutative Ringe verallgemeinern lässt, bietet das Konzept des *Ideals*. Durch Ideale erfasst man nicht nur die Menge der Vielfachen *eines* Elements $r \in R$, sondern auch Vielfache von Produkten und Summen mehrerer Elemente.

Ein Ideal ist eine Teilmenge $I \subseteq R$, die ähnliche Eigenschaften hat wie die Mengen $m\mathbb{Z}$ der Vielfachen einer Zahl m im Ring $\mathbb{Z}$. Man fordert, dass das Ideal I abgeschlossen unter der Summenbildung ist, und dass das Produkt eines Elements im Ideal mit einem beliebigen Ringelement wieder in I liegt, ähnlich wie die Summe von zwei Vielfachen einer ganzen Zahl m ein Vielfaches von m ist, und das Produkt eines Vielfachen von m mit einer beliebigen ganzen Zahl ein Vielfaches vom m ist.

Definition 8.1.11: Sei R ein kommutativer Ring. Eine Teilmenge $I \subseteq R$ heißt **Ideal**, wenn $(I,+)$ eine Untergruppe von $(R,+)$ ist und $r \cdot i \in I$ für alle $r \in R$, $i \in I$ gilt.

Beispiel 8.1.12: Sei R ein kommutativer Ring.

1. Die Teilmengen $\{0\} \subseteq R$ und $R \subseteq R$ sind Ideale in R.

2. Für jede Familie $(I_j)_{j\in J}$ von Idealen $I_j \subseteq R$ ist auch $\cap_{j\in J} I_j$ ein Ideal in R.

   Denn der Schnitt $\cap_{j\in J} I_j$ von Untergruppen $(I_j,+) \subseteq (R,+)$ ist eine Untergruppe von $(R,+)$. Ist $i \in \cap_{j\in J} I_j$ und $r \in R$, so folgt $i \in I_j$ und $r \cdot i \in I_j$ für alle für alle $j \in J$, da I_j ein Ideal ist, also auch $r \cdot i \in \cap_{j\in J} I_j$.

3. Für jede Teilmenge $M \subseteq R$ ist die Menge $(M) = \cap_{M\subseteq I, I\subseteq R \text{ Ideal}} I$ ein Ideal in R. Es heißt das **von M erzeugte Ideal** und ist das kleinste Ideal, das M enthält: für jedes Ideal $I \subseteq R$ mit $M \subseteq I$ folgt $(M) \subseteq I$. (Übung).

4. Ist $M = \{r\}$ für ein Element $r \in R$, so schreibt man $(r) := (\{r\})$. In diesem Fall ist $(r) = \{s \cdot r | s \in R\}$ die Menge der Vielfachen von r.

   Denn da $(r) \subseteq R$ ein Ideal mit $r \in (r)$ ist, folgt $s \cdot r \in (r)$ für alle $s \in R$ und somit $I = \{s \cdot r | s \in R\} \subseteq (r)$. Andererseits ist $I \subseteq R$ ein Ideal mit $r = 1 \cdot r \in I$. Aus $a, b \in I$ folgt nämlich, dass es $c, d \in R$ mit $a = c \cdot r$ und $b = d \cdot r$ gibt, und somit ist auch $a + b = c \cdot r + d \cdot r = (c + d) \cdot r \in I$ und $s \cdot a = (sc) \cdot r \in I$ für alle $s \in R$. Also nimmt I am Schnitt der Ideale in der Definition von (r) teil und somit gilt $(r) \subseteq I$.

5. Ist $\phi : R \to S$ ein Ringhomomorphismus, so ist $\ker(\phi) = \phi^{-1}(\{0\}) \subseteq R$ ein Ideal in R.

   Denn $\phi : (R,+) \to (S,+)$ ist ein Gruppenhomomorphismus, und damit ist $\ker(\phi) \subseteq R$ eine Untergruppe von $(R,+)$ nach Satz 2.2.21. Für alle $r \in R$ und $p \in \ker(\phi)$ folgt $\phi(r \cdot p) = \phi(r) \cdot \phi(p) = \phi(r) \cdot 0 = 0$ und damit $r \cdot p \in \ker(\phi)$. Also ist $\ker(\phi)$ ein Ideal.

Aus den betrachteten Beispielen ergibt sich, dass die Ideale der Form $(r) = \{s \cdot r | s \in R\}$, die von einem einzigen Element $r \in R$ erzeugt werden, besonders einfach zu handhaben sind. Für ein Ideal, das von einer beliebigen Teilmenge $M \subseteq R$ erzeugt wird, ist es oft schwieriger zu sehen, welche Elemente genau darin enthalten sind. Kommutative Ringe, in denen jedes Ideal von einem einzigen Element erzeugt wird, haben daher besonders gute Eigenschaften.

Definition 8.1.13: Sei $R \neq \{0\}$ ein kommutativer nullteilerfreier Ring. Ein Ideal $I \subseteq R$ heißt **Hauptideal**, wenn es ein Element $r \in R$ mit $I = (r)$ gibt. Einen Ring, in dem alle Ideale Hauptideale sind, bezeichnet man als **Hauptidealring**.

Beispiel 8.1.14:

1. Der Ring $(\mathbb{Z},+,\cdot)$ ist ein Hauptidealring.

   Denn ist $I \subseteq \mathbb{Z}$ ein Ideal, so gilt entweder $I = \{0\} = (0)$ oder $\{z \in I | z > 0\} \neq \emptyset$. Im zweiten Fall setzt man $d = \min\{z \in I | z > 0\}$. Dann gilt $(d) \subseteq I$ nach Beispiel 8.1.12, 3. Zu jedem Element $z \in I$ gibt es ganze Zahlen $r, s \in \mathbb{Z}$ mit $z = s \cdot d + r$ und $0 \leq r < d$. Da $z, d \in I$ und I ein Ideal ist, ist auch $s \cdot d \in I$ und $r = z - s \cdot d \in I$. Nun ist aber d das kleinste positive Element aus I und $0 \leq r < d$. Also muss $r = 0$ gelten, und es folgt $z = s \cdot d \in (d)$. Also gilt auch $I \subseteq (d)$ und damit $I = (d)$.

2. Auf ähnliche Weise kann man zeigen, dass jeder euklidische Ring ein Hauptidealring ist.
3. Der Ring $\mathbb{Z}[x]$ der Polynome mit Koeffizienten in $\mathbb{Z}$ ist kein Hauptidealring. Denn das von der Menge $M = \{2, x\}$ erzeugte Ideal $I = (M)$ ist kein Hauptideal (Übung).

Bemerkung 8.1.15: Ist $R \neq 0$ ein kommutativer nullteilerfreier Ring mit Eins, so ist das Element $r \in R$ mit $I = (r)$ in Definition 8.1.13 eindeutig bis auf Multiplikation mit Einheiten.

Denn ist $r' = r \cdot s$ mit $r, s \in R$ so folgt $(r') = \{r' \cdot t | t \in R\} = \{r \cdot s \cdot t | t \in R\} \subseteq (r)$. Ist $s \in R^\times$ eine Einheit, so folgt daraus aber auch $r = s^{-1} \cdot r'$ und damit $(r) \subseteq (r')$, was $(r) = (r')$ impliziert. Umgekehrt folgt aus $(r) = (r')$ für $r, r' \in R \setminus \{0\}$, dass es $s, t \in R$ mit $r' = s \cdot r$ und $r = t \cdot r'$ gibt. Daraus folgt $(1 - t \cdot s) \cdot r = r - r = 0$. Da $r \neq 0$ und R nullteilerfrei ist, folgt daraus $t \cdot s = 1$ und damit sind $s, t \in R$ Einheiten.

Wir zeigen nun, dass für jeden Körper $\mathbb{K}$ der Polynomring $\mathbb{K}[x]$ ein Hauptidealring ist. Der Beweis ist ähnlich wie der aus Beispiel 8.1.14, 1. für den Ring $\mathbb{Z}$, nur dass der Polynomgrad die Rolle des Betrags einer Zahl $z \in \mathbb{Z}$ in Beispiel 8.1.14, 1. übernimmt. Das Element p mit $I = (p)$ ist also nicht mehr die kleinste positive ganze Zahl, die in I enthalten ist, sondern ein Polynom minimalen Grades, das in I enthalten ist.

Man kann nicht nur zeigen, dass es zu jedem Ideal $I \subseteq \mathbb{K}[x]$ ein Polynom $p \in \mathbb{K}[x]$ mit $I = (p)$ gibt, sondern dass für $I \neq \{0\}$ sogar ein *normiertes* Polynom p mit dieser Eigenschaft existiert. Das ergibt sich aus Bemerkung 8.1.15, denn die Einheiten im Polynomring $\mathbb{K}[x]$ sind die Polynome von Grad 0, also die konstanten Polynome $\neq 0$. Indem man ein Polynom $p \in \mathbb{K}[x] \setminus \{0\}$ mit einer solchen Konstante multipliziert, kann man es normieren.

Satz 8.1.16: Für jeden Körper $\mathbb{K}$ ist der Polynomring $\mathbb{K}[x]$ ein Hauptidealring. Jedes Ideal $\{0\} \neq I \subseteq \mathbb{K}[x]$ enthält genau ein normiertes Polynom $q \in \mathbb{K}[x]$ von minimalem Grad, und dieses erzeugt $I = (q)$.

Beweis:
1. Existenz von q:

Ist $I = \{0\}$, so gilt $I = (0)$, und das Nullpolynom $0 \in \mathbb{K}[x]$ ist das gesuchte Polynom. Ist $I \neq \{0\}$, so ist die Menge $M = \{n \in \mathbb{N}_0 | \exists p \in I \text{ mit } \deg(p) = n\}$ nicht leer. Also gibt es ein $k \in \mathbb{N}_0$ mit $k = \min M$ und ein normiertes Polynom $q \in I$ mit $\deg(q) = k$. Sei $(q) \subseteq \mathbb{K}[x]$ das davon erzeugte Ideal. Dann gilt offensichtlich $(q) = \{r \cdot q | r \in \mathbb{K}[x]\} \subseteq I$, denn aus $q \in I$ folgt $r \cdot q \in I$ für alle $r \in \mathbb{K}[x]$.

Wir zeigen, dass auch $I \subseteq (q)$ und damit $I = (q)$ gilt. Zu jedem Polynom $p \in I$ gibt es nach Satz 8.1.8 eindeutig bestimmte Polynome $s, r \in \mathbb{K}[x]$ mit $p = s \cdot q + r$ und $\deg(r) < \deg(q)$. Da $I \subseteq \mathbb{K}[x]$ ein Ideal ist und $p, q \in I$ folgt, dass auch $r = p - s \cdot q \in I$ gilt. Da $\deg(r) < \deg(q)$ und $\deg(q) = \min\{\deg(t) | t \in I, t \neq 0\}$, muss $\deg(r) = -\infty$ und damit $r = 0$ gelten. Also ist $p = s \cdot q \in (q)$ und $I \subseteq (q)$.

2. Eindeutigkeit von q:

Ist $q' \in \mathbb{K}[x]$ ein weiteres normiertes Polynom mit $\deg(q') = \deg(q) = k$, so gibt es wegen $I = (q)$ ein $s \in \mathbb{K}[x]$ mit $q' = s \cdot q$. Da $\deg(q') = \deg(q) + \deg(s) = \deg(q)$, folgt $\deg(s) = 0$, also s konstant, und da q und q' normiert sind, folgt $s = 1$ und $q = q'$. $\square$

8.2 Das charakteristische Polynom

Wir benutzen nun Polynome, um Fragen der Diagonalisierbarkeit und der Trigonalisierbarkeit zu untersuchen. Ausgangspunkt dabei ist die charakteristische Gleichung $\det(A-\lambda\mathbb{1}_n)=0$, die eine polynomiale Gleichung in λ ist. Sie ordnet jeder Matrix und jedem Endomorphismus eines endlich-dimensionalen Vektorraums ein Polynom zu, das *charakteristische Polynom*.

Die Eigenschaften dieses Polynoms, wie die Anzahl seiner Nullstellen, deren Vielfachheiten und das Zerfallen in Polynome kleineren Grades, liefern Informationen über die zugrundeliegende Matrix oder den zugrundeliegenden Endomorphismus. Es liefert die Dimensionen der Haupträume und zeigt damit, ob eine Matrix oder ein Endomorphismus trigonalisierbar ist.

Um das charakteristische Polynom zu definieren, ordnen wir jeder Matrix $A \in \mathrm{Mat}(n\times n,\mathbb{K})$ die Matrix $A - x\mathbb{1}_n \in \mathrm{Mat}(n\times n,\mathbb{K}[x])$ mit Einträgen im Polynomring zu und berechnen ihre Determinante $\det(A-x\mathbb{1}_n)\in \mathrm{Mat}(n\times n,\mathbb{K}[x])$. Dabei nutzt man aus, dass man alle Matrixverknüpfungen wie Addition, Matrixmultiplikation, Multiplikation mit Skalaren völlig analog auch für Matrizen mit Einträgen in einem kommutativen Ring mit Eins wie $\mathbb{K}[x]$ definieren kann. Das gleiche gilt für die Transposition, Determinante und Cramersche Regel, da fort nach der Leibnizschen Formel nur Summen von Produkten von Einträgen der Matrix auftreten.

Da die Determinante eine Linearkombination von Produkten der Einträge der Matrix ist, ist dann auch $p_A=\det(A-x\mathbb{1}_n)$ ein Polynom in $\mathbb{K}[x]$, das sogenannte *charakteristische Polynom* von A. Die Nullstellen von p_A sind genau die Elemente $\lambda\in\mathbb{K}$ mit $p_A(\lambda)=\det(A-\lambda\mathbb{1}_n)=0$, also die Eigenwerte der Matrix A. Sind zwei Matrizen $A,B\in\mathrm{Mat}(n\times n,\mathbb{K})$ ähnlich, so gibt es eine Matrix $S\in\mathrm{GL}(n,\mathbb{K})$ mit $A=S\cdot B\cdot S^{-1}$, und es folgt

$$p_A=\det(A-x\mathbb{1}_n)=\det(S\cdot B\cdot S^{-1}-x\mathbb{1}_n)=\det(S\cdot(B-x\mathbb{1}_n)\cdot S^{-1})=\det(B-x\mathbb{1}_n)=p_B.$$

Ähnliche Matrizen haben also das gleiche charakteristische Polynom. Das bedeutet, dass wir auch jedem Endomorphismus $\phi: V\to V$ eines endlich-dimensionalen $\mathbb{K}$-Vektorraums V ein charakteristisches Polynom p_ϕ zuordnen können, indem wir eine geordnete Basis B von V wählen und $p_\phi=\det({}_BM_B(\phi)-x\mathbb{1}_n)$ setzen. Da die darstellenden Matrizen von ϕ bezüglich verschiedener Basen ähnlich sind, hängt dieses Polynom nicht von der Wahl der Basis ab.

Definition 8.2.1: Sei $\mathbb{K}$ ein Körper und V ein n-dimensionaler Vektorraum über $\mathbb{K}$.

1. Das **charakteristische Polynom** einer Matrix $A\in\mathrm{Mat}(n\times n,\mathbb{K})$ ist das Polynom $p_A=\det(A-x\mathbb{1}_n)\in\mathbb{K}[x]$.
2. Das **charakteristische Polynom** eines Endomorphismus $\phi: V\to V$ ist das Polynom $p_\phi=\det({}_BM_B(\phi)-x\mathbb{1}_n)$, wobei B eine beliebige geordnete Basis von V ist.
3. Die Vielfachheit einer Nullstelle $\lambda\in\mathbb{K}$ von p_A oder p_ϕ heißt **algebraische Vielfachheit** des Eigenwerts λ und wird mit $a(A,\lambda)$ oder $a(\phi,\lambda)$ bezeichnet.

Rein rechnerisch besteht kein Unterschied zwischen dem Bestimmen der Nullstellen des charakteristischen Polynoms und dem Lösen der charakteristischen Gleichung. Wir erhalten die schon vorher bekannten Ausdrücke, die jetzt als ein Polynom aufgefasst werden. Der Vorteil ist, dass dieses Polynom Informationen über alle Eigenwerte und die zugehörigen Haupträume enthält.

Beispiel 8.2.2:

1. Für eine (2×2)-Matrix $A = (a_{ji}) \in \mathrm{Mat}(2 \times 2, \mathbb{K})$ ergibt sich

$$\begin{aligned} p_A &= \det \begin{pmatrix} a_{11} - x & a_{12} \\ a_{21} & a_{22} - x \end{pmatrix} = (a_{11} - x)(a_{22} - x) - a_{12}a_{21} \\ &= x^2 - (a_{11} + a_{22})x + a_{11}a_{22} - a_{12}a_{21} = x^2 - \mathrm{Tr}(A)x + \det(A). \end{aligned}$$

2. Ist $A \in \mathrm{Mat}(n \times n, \mathbb{K})$ eine obere Dreiecksmatrix, so gilt

$$p_A = \det(A - x\mathbb{1}_n) = \det \begin{pmatrix} a_{11} - x & * & * & * \\ 0 & a_{22} - x & * & * \\ \vdots & \ddots & \ddots & * \\ 0 & \ldots & 0 & a_{nn} - x \end{pmatrix} = (a_{11} - x) \cdot (a_{22} - x) \cdots (a_{nn} - x).$$

3. Ist V ein n-dimensionaler $\mathbb{K}$-Vektorraum und $\phi : V \to V$ ein nilpotenter Endomorphismus, so gibt es nach Lemma 7.1.1 eine geordnete Basis B von V, so dass ${}_BM_B(\phi)$ eine obere Dreiecksmatrix ist, die in der Diagonale nur Nullen enthält. Es folgt

$$p_\phi = \det({}_BM_B(\phi) - x\mathbb{1}_n) = \det \begin{pmatrix} -x & * & * & * \\ 0 & -x & \ddots & * \\ \vdots & \ddots & \ddots & * \\ 0 & \ldots & 0 & -x \end{pmatrix} = (-1)^n x^n$$

   Damit ist $\lambda = 0$ die einzige Nullstelle von p_ϕ mit algebraischer Vielfachheit $a(\phi, 0) = n$.

Das charakteristische Polynom p_A einer Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{K})$ oder eines Endomorphismus $\phi : V \to V$ eines n-dimensionalen Vektorraums V ist immer ein Polynom vom Grad n. Dass es wichtige Informationen über die zugrundeliegende Matrix oder den zugrundeliegenden Endomorphismus liefert, sieht man schon daran, dass es die Determinante und die Spur als Koeffizienten enthält. Das charakteristische Polynom ist also eine stärkere Invariante als die Spur oder Determinante einer Matrix alleine und kann mehr nicht-ähnliche Matrizen unterscheiden.

Satz 8.2.3: (**Eigenschaften des charakteristischen Polynoms**)
Das charakteristische Polynom einer Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{K})$ ist gegeben durch

$$p_A = (-1)^n x^n + (-1)^{n-1} \mathrm{tr}(A) x^{n-1} + \ldots + \det(A).$$

Beweis:
Mit der Leibnizschen Formel erhält man

$$p_A = \det(A - x\mathbb{1}_n) = \Sigma_{\sigma \in S_n} \mathrm{sgn}(\sigma)(A - x\mathbb{1}_n)_{\sigma(1)1} \cdots (A - x\mathbb{1}_n)_{\sigma(n)n}.$$

Der konstante Term von p_A ergibt sich, indem man $x = 0$ setzt, zu $\det(A - 0 \cdot \mathbb{1}_n) = \det(A)$. Da $(A - x\mathbb{1}_n)_{ji} = A_{ji} - \delta_{ji}x$, trägt nur $\sigma = \mathrm{id}$ zu den Koeffizienten von x^n und x^{n-1} bei und

$$\begin{aligned} &(A - x\mathbb{1}_n)_{11} \cdots (A - x\mathbb{1}_n)_{nn} = (A_{11} - x) \cdots (A_{nn} - x) \\ &= (-1)^n x^n + (-1)^{n-1} (\Sigma_{j=1}^n A_{jj}) \cdot x^{n-1} + ... = (-1)^n x^n + (-1)^{n-1} \mathrm{tr}(A) x^{n-1} + ... \end{aligned}$$

Einsetzen in die Leibnizsche Formel ergibt dann die Behauptung. $\square$

Mit dem charakteristischen Polynom vereinfachen wir nun die Kriterien für Diagonalisierbarkeit und Trigonalisierbarkeit aus Satz 6.2.9 und Satz 7.2.1. Neben der Information über die Eigenwerte, die als Nullstellen des charakteristischen Polynoms auftreten, enthält das charakteristische Polynom noch weitere Informationen, nämlich deren Vielfachheiten. Diese sind gerade die Dimensionen der zugehörigen Haupträume. Damit lässt sich Trigonalisierbarkeit direkt am charakteristischen Polynom ablesen.

Satz 8.2.4: (**Trigonalisierbarkeitskriterium und Diagonalisierbarkeitskriterium**)
Sei $\mathbb{K}$ ein Körper, V ein n-dimensionaler $\mathbb{K}$-Vektorraum und $\phi : V \to V$ ein Endomorphismus mit paarweise verschiedenen Eigenwerten $\lambda_1, ..., \lambda_m \in \mathbb{K}$. Dann ist das charakteristische Polynom

$$p_\phi = (-1)^n (x - \lambda_1)^{h_1} \cdots (x - \lambda_m)^{h_m} \cdot r_\phi \qquad h_i = \dim_{\mathbb{K}} H(\phi, \lambda_i) = a(\phi, \lambda_i),$$

wobei $r_\phi \in \mathbb{K}[x]$ ein Polynom ohne Nullstellen in $\mathbb{K}$ ist. Insbesondere gilt:

1. Der Endomorphismus ϕ ist trigonalisierbar genau dann, wenn p_ϕ in Linearfaktoren zerfällt, und dies ist äquivalent zu $r_\phi = 1$.

2. Es gilt $1 \leq g(\phi, \lambda_i) \leq a(\phi, \lambda_i)$, und ϕ ist diagonalisierbar genau dann, wenn $r_\phi = 1$ und $g(\phi, \lambda_i) = a(\phi, \lambda_i)$ für alle $i \in \{1, ..., m\}$.

Beweis:
Nach Korollar 7.2.2 gibt es im nicht-trigonalisierbaren Fall eine Basis B von V, so dass

$${}_B M_B(\phi) = \begin{pmatrix} M & 0_{p\times q} \\ 0_{q\times p} & A \end{pmatrix},$$

wobei $M \in \mathrm{Mat}(p \times p, \mathbb{K})$ eine obere Dreiecksmatrix ist und $A \in \mathrm{Mat}(q \times q, \mathbb{K})$ eine Matrix ohne Eigenwerte in $\mathbb{K}$. Der trigonalisierbare Fall entspricht dabei $p = n$ und $q = 0$, also der Abwesenheit der Matrix A. Daraus folgt für das charakteristische Polynom

$$p_\phi = \det({}_B M_B(\phi) - x\mathbb{1}_n) = \det\begin{pmatrix} M - x\mathbb{1}_p & 0_{p\times q} \\ 0_{q\times p} & A - x\mathbb{1}_q \end{pmatrix} = \det(M - x\mathbb{1}_p) \cdot \det(A - x\mathbb{1}_q) = p_M \cdot p_A.$$

Da A keine Eigenwerte in $\mathbb{K}$ hat, hat $p_A \in \mathbb{K}[x]$ keine Nullstellen in $\mathbb{K}$. Da M und damit auch $M - x\mathbb{1}_p$ eine obere Dreiecksmatrix ist, gilt $p_M = (M_{11} - x) \cdots (M_{pp} - x)$. Nach dem Beweis von Satz 7.2.1 enthält die Diagonale von M genau die Eigenwerte $\lambda_1, ..., \lambda_m$, wobei der Eigenwert λ_i genau $h_i = \dim_{\mathbb{K}} H(\phi, \lambda_i)$ mal auftritt. Also folgt

$$p_\phi = (-1)^n (x - \lambda_1)^{h_1} \cdots (x - \lambda_m)^{h_m} \cdot r_\phi$$

mit $r_\phi = (-1)^q p_A$ und $\deg(r_\phi) = q$. Offensichtlich gilt $r_\phi = 1$ genau dann, wenn $h_1 + \ldots + h_m = n$, also nach Satz 7.2.1 genau dann, wenn ϕ trigonalisierbar ist.

Nach Satz 7.1.7 gilt $\{0\} \neq E(\phi, \lambda_i) \subseteq H(\phi, \lambda_i)$ und damit $1 \leq g(\phi, \lambda_i) \leq h_i = a(\phi, \lambda_i)$. Nach Satz 6.2.9 ist ϕ diagonalisierbar genau dann, wenn $g(\phi, \lambda_1) + \ldots g(\phi, \lambda_m) = n$, also genau dann, wenn $r_\phi = 1$ und $g(\phi, \lambda_i) = a(\phi, \lambda_i)$ für alle $i \in \{1, ..., m\}$ gilt. □

Anhand des Trigonalisierbarkeitskriteriums aus Satz 8.2.4 sieht man, dass Trigonalisierbarkeit vorwiegend eine Frage des zugrundeliegenden Körpers ist. Insbesondere ergibt sich für einen algebraisch abgeschlossenen Körper $\mathbb{K}$ aus Satz 8.2.4, dass jedes Polynom aus $\mathbb{K}[x]$ über $\mathbb{K}$

in Linearfaktoren zerfällt und damit jeder Endomorphismus trigonalisierbar ist. Dies wurde bereits in Korollar 7.2.5 bewiesen. Für die Diagonalisierbarkeit benötigt man zusätzlich, dass geometrische und algebraische Vielfachheit jedes Eigenwertes übereinstimmen.

Eine weitere nützliche Aussage über trigonalisierbare Matrizen, die sich direkt aus dem charakteristischen Polynom ergibt, ist eine Formel, die die Spur und Determinante einer trigonalisierbaren Matrix durch die Eigenwerte und deren algebraische Vielfachheiten ausdrückt.

Korollar 8.2.5: Ist $A \in \mathrm{Mat}(n \times n, \mathbb{K})$ trigonalisierbar mit paarweise verschiedenen Eigenwerten $\lambda_1, ..., \lambda_m$ und algebraischen Vielfachheiten $h_1, ..., h_m$, so ist

$$\mathrm{tr}(A) = h_1\lambda_1 + ... + h_m\lambda_m \qquad \det(A) = \lambda_1^{h_1} \cdots \lambda_m^{h_m}.$$

Beweis:
Nach Satz 8.2.3 gilt für das charakteristische Polynom p_A einer Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{K})$

$$p_A = (-1)^n x^n + (-1)^{n-1}\mathrm{tr}\,(A)x^{n-1} + a_{n-2}x^{n-2} + ... + a_1 x + \det(A).$$

Ist A trigonalisierbar, so gilt nach Satz 8.2.4

$$p_A = (-1)^n(x-\lambda_1)^{h_1} \cdots (x-\lambda_m)^{h_m} = (-1)^n x^n + (-1)^{n-1}(h_1\lambda_1 + ... + h_m\lambda_m)x^{n-1} + ... + \lambda_1^{h_1} \cdots \lambda_m^{h_m},$$

wobei $\lambda_1, ..., \lambda_m$ die verschiedenen Eigenwerte von A und $h_1, ..., h_m$ deren algebraische Vielfachheiten bezeichnen. Aus einem Koeffizientenvergleich folgt die Behauptung. □

Die Aussagen in Satz 8.2.4 liefern ein effizienteres Verfahren zur Untersuchung der Diagonalisierbarkeit und Trigonalisierbarkeit einer Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{K})$, das in Abbildung 2 in einem Flussdiagramm dargestellt ist. Man kann nun darauf verzichten, die Dimensionen der Haupträume explizit zu berechnen. Sie sind ergeben sich aus der Vielfachheit der entsprechenden Nullstelle des charakteristischen Polynoms.

1. Ist $\mathbb{K}$ algebraisch abgeschlossen (z. B. $\mathbb{K} = \mathbb{C}$), so ist A immer trigonalisierbar. Man geht zu Schritt 3.
2. Ist $\mathbb{K}$ nicht algebraisch abgeschlossen, so untersucht man, ob das charakteristische Polynom $p_A = \det(A - x\mathbb{1}_n)$ über $\mathbb{K}$ in Linearfaktoren zerfällt. Wenn nein, so ist A nicht trigonalisierbar und auch nicht diagonalisierbar. Das Verfahren bricht ab.
3. Zerfällt p_A über $\mathbb{K}$ in Linearfaktoren, so ist A trigonalisierbar. Zur Untersuchung der Diagonalisierbarkeit bestimmt man die Nullstellen des charakteristischen Polynoms p_A und deren Vielfachheiten $a(A, \lambda)$. Ist $a(A, \lambda) = 1$ für alle Nullstellen λ, so ist A diagonalisierbar.
4. Ansonsten bestimmt man mit dem Gauß-Verfahren die Defekte der Matrizen $A - \lambda\mathbb{1}_n$ für jede Nullstelle λ von p_A mit $a(\lambda, 1) > 1$. Gilt $\mathrm{def}(A - \lambda\mathbb{1}_n) = a(A, \lambda)$ für all diese Nullstellen, so ist A diagonalisierbar. Ansonsten nicht.

Beispiel 8.2.6: Wir untersuchen die Matrix

$$A = \begin{pmatrix} 1 & -1 & 7 & 2 \\ 1 & 1 & 2 & 5 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & -1 & 4 \end{pmatrix} \in \mathrm{Mat}(4 \times 4, \mathbb{R})$$

auf Trigonalisierbarkeit über $\mathbb{R}$ und bestimmen die Dimensionen ihrer Haupträume. Da A eine Blockmatrix ist, erhält man

$$p_A = \det\begin{pmatrix} 1-x & -1 & 7 & 2 \\ 1 & 1-x & 2 & 5 \\ 0 & 0 & 2-x & 1 \\ 0 & 0 & -1 & 4-x \end{pmatrix} = \det\begin{pmatrix} 1-x & -1 \\ 1 & 1-x \end{pmatrix} \cdot \det\begin{pmatrix} 2-x & 1 \\ -1 & 4-x \end{pmatrix}$$
$$= ((1-x)^2+1)\cdot((2-x)(4-x)+1) = (x^2-2x+2)(x-3)^2.$$

Da das Polynom $x^2-2x+2 = (x-1)^2+1$ keine Nullstellen in $\mathbb{R}$ hat, ist A nicht trigonalisierbar über $\mathbb{R}$. Der einzige Eigenwert ist $\lambda = 3$ und $\dim_{\mathbb{R}} H(A,3) = a(A,3) = 2 < 4$.

Mit Hilfe des charakteristischen Polynoms sieht man, welche Eigenschaften einer Matrix körperabhängig sind und welche nicht. Ob eine Matrix $A \in \mathrm{Mat}(\mathbb{K}\times\mathbb{K}, n)$ trigonalisierbar ist, hängt nur davon ab, ob das charakteristische Polynom p_A über $\mathbb{K}$ in Linearfaktoren zerfällt. Dies ist im Wesentlichen eine Frage des Körpers.

Insbesondere existiert zu jedem Körper $\mathbb{K}$ ein algebraisch abgeschlossener Körper $\mathbb{L}$ mit $\mathbb{K} \subseteq \mathbb{L}$, über dem sich jede Matrix mit Einträgen aus $\mathbb{K}$ trigonalisieren lässt. Jeder Hauptraum entspricht dabei einer Nullstelle, und seine Dimension ist ihre algebraische Vielfachheit. Letztere hängt nicht vom betrachteten Körper ab.

Die geometrischen Vielfachheiten der Nullstellen, also die Dimensionen der Eigenräume werden durch das charakteristische Polynom dagegen *nicht* erfasst. Ebensowenig finden sich darin Informationen über die Fitting-Indizes der Eigenwerte. So hat beispielsweise in jedem Körper die Nullmatrix $0_{n\times n}$ das gleiche charakteristische Polynom wie jede nilpotente Matrix $A \in \mathrm{Mat}(n\times n, \mathbb{K})$, nämlich $(-1)^n x^n$. Die Dimension des Eigenraums von A zum Eigenwert 0 und der zugehörige Fitting-Index können aber jeden Wert von 1 bis n annehmen.

8.3 Der Satz von Cayley-Hamilton und das Minimalpolynom

In diesem Abschnitt verbessern wir die Beschreibung von Endomorphismen $\phi : V \to V$ endlichdimensionaler Vektorräume V durch Polynome, so dass wir auch den Fitting-Index f_λ des Endomorphismus $\phi_\lambda = \phi - \lambda \mathrm{id}_V$ für jeden Eigenwert λ aus einem Polynom ablesen können.

Daraus ergibt sich dann insbesondere ein Diagonalisierbarkeitskriterium. Nach Satz 8.2.4 ist ein trigonalisierbarer Endomorphismus $\phi : V \to V$ nämlich genau dann diagonalisierbar, wenn für jeden Eigenwert λ die geometrische Vielfachheit $g(\phi,\lambda) = \dim_{\mathbb{K}} \ker(\phi - \lambda\mathrm{id}_V)$ mit der algebraischen Vielfachheit $a(\phi,\lambda) = \dim_{\mathbb{K}} H(\phi,\lambda) = \dim_{\mathbb{K}} \ker(\phi-\lambda\mathrm{id}_V)^{f_\lambda}$ übereinstimmt. Dies ist genau dann der Fall, wenn $f_\lambda = 1$ gilt.

Wir betrachten für einen Endomorphismus $\phi \in \mathrm{End}_{\mathbb{K}}(V)$ die Menge I_ϕ der Polynome $p \in \mathbb{K}[x]$ mit $p(\phi) = 0$, also den Kern der Evaluationsabbildung $\mathrm{ev}_\phi : \mathbb{K}[x] \to \mathrm{End}_{\mathbb{K}}(V)$, $p \mapsto p(\phi)$. Schon aus Dimensionsgründen muss I_ϕ für einen n-dimensionalen $\mathbb{K}$-Vektorraum V mehr als nur das Nullpolynom enthalten. Denn es gilt $\dim_{\mathbb{K}} \mathrm{End}_{\mathbb{K}}(V) = n^2$ und somit ist die Menge $\{\phi^0, \phi, ..., \phi^{n^2}\} \subseteq \mathrm{End}_{\mathbb{K}}(V)$ linear abhängig. Jede Linearkombination $\Sigma_{k=0}^m a_k\phi^k = 0$, in der nicht alle Koeffizienten verschwinden, liefert ein Polynom $0 \neq p = \Sigma_{k=0}^m a_k x^k \in I_\phi$.

Da $I_\phi = \ker \mathrm{ev}_\phi$ ein Ideal in $\mathbb{K}[x]$ ist, gibt es nach Satz 8.1.16 ein Polynom q_ϕ, das I_ϕ erzeugt.

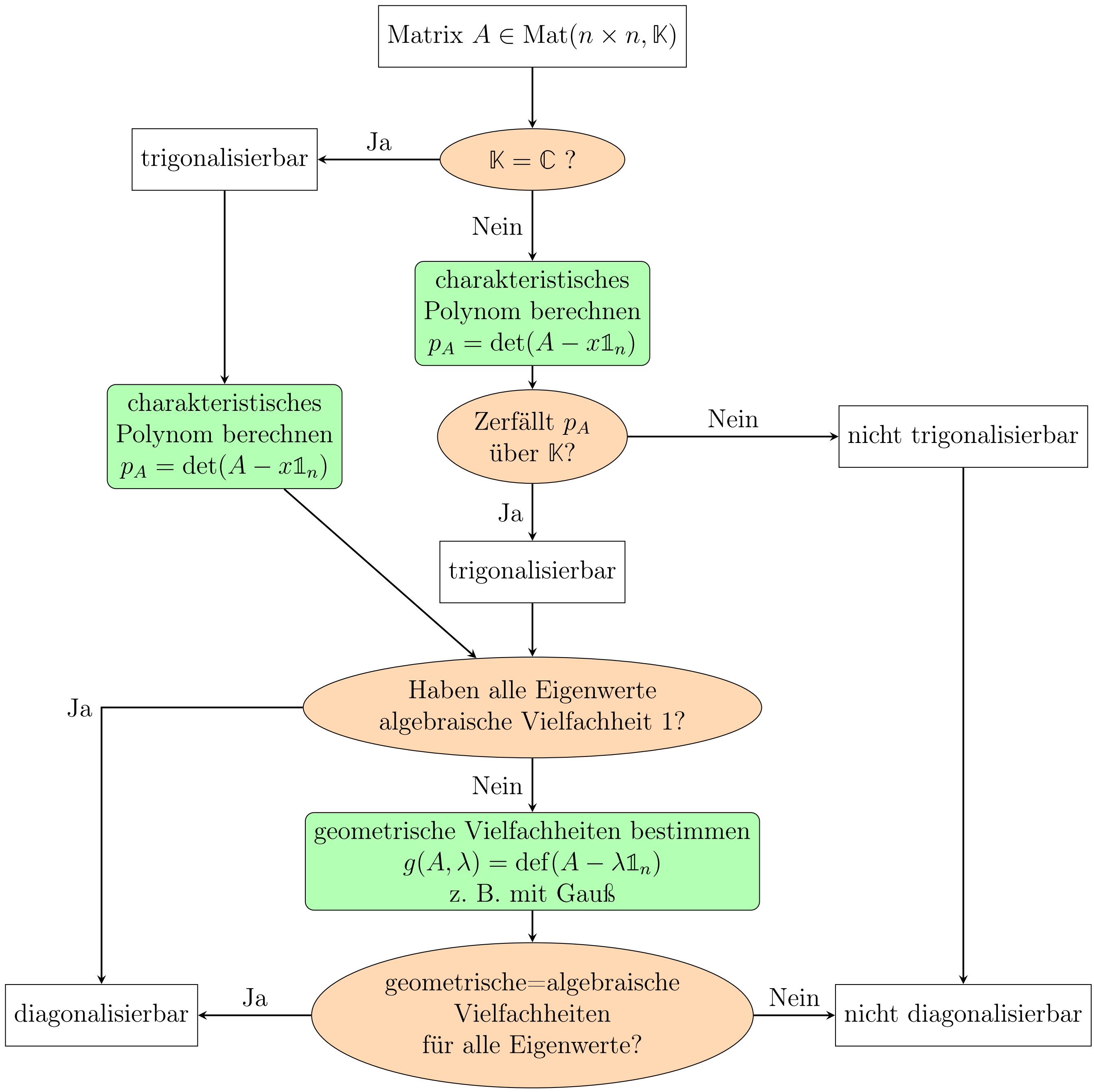


Abbildung 2: Untersuchung der Diagonalisierbarkeit und Trigonalisierbarkeit.

Satz 8.3.1: Sei $\mathbb{K}$ ein Körper, $n \in \mathbb{N}$ und V ein n-dimensionaler Vektorraum über $\mathbb{K}$. Dann gilt für jeden Endomorphismus $\phi : V \to V$ und jede Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{K})$

1. Die Teilmengen $I_\phi = \ker(\mathrm{ev}_\phi)$ und $I_A = \ker(\mathrm{ev}_A)$ sind Ideale in $\mathbb{K}[x]$.
2. Es gibt eindeutige normierte Polynome $q_\phi \in I_\phi$ und $q_A \in I_A$ minimalen Grades. Es gilt
$$I_\phi = (q_\phi) = \{s \cdot q_\phi \mid s \in \mathbb{K}[x]\} \qquad I_A = (q_A) = \{s \cdot q_A \mid s \in \mathbb{K}[x]\}.$$
3. Für jede geordnete Basis B von V gilt $I_\phi = I_M$ und $q_\phi = q_M$ mit $M := {}_B M_B(\phi)$.
4. Sind $M, N \in \mathrm{Mat}(n \times n, \mathbb{K})$ ähnlich, so gilt $I_M = I_N$ und $q_M = q_N$.

Die Polynome q_ϕ und q_A bezeichnet man als die **Minimalpolynome** von ϕ und von A.

Beweis:
Wir beweisen 1. und 2. für Endomorphismen $\phi : V \to V$. Die entsprechenden Aussagen für Matrizen folgen aus Betrachtung der Endomorphismen $\phi_A : \mathbb{K}^n \to \mathbb{K}^n$, $v \mapsto A \cdot v$.

Nach Satz 8.1.1 ist die Evaluationsabbildung $\mathrm{ev}_\phi : \mathbb{K}[x] \to \mathrm{End}_\mathbb{K}(V)$, $p \mapsto p(\phi)$ ein Homomorphismus von $\mathbb{K}$-Algebren und damit insbesondere ein unitaler Ringhomomorphismus. Also ist nach Beispiel 8.1.12 ihr Kern $I_\phi = \ker(\mathrm{ev}_\phi) = \{p \in \mathbb{K}[x] | \, p(\phi) = 0\}$ ein Ideal in $\mathbb{K}[x]$. Nach Satz 8.1.16 enthält I_ϕ ein eindeutiges normiertes Polynom q_ϕ minimalen Grades, und $I_\phi = (q_\phi)$.

3. Nach Satz 4.2.18 ist ${}_B M_B : \mathrm{End}_\mathbb{K}(V) \to \mathrm{Mat}(n \times n, \mathbb{K})$, $\phi \mapsto {}_B M_B(\phi)$ für jede geordnete Basis B von V ein Isomorphismus von $\mathbb{K}$-Algebren. Daraus folgt für $p = \Sigma_{k=0}^n a_k x^k \in \mathbb{K}[x]$
$$p({}_B M_B(\phi)) = \Sigma_{k=0}^n a_k \, {}_B M_B(\phi)^k = \Sigma_{k=0}^n a_k \, {}_B M_B(\phi^k) = {}_B M_B(\Sigma_{k=0}^n a_k \phi^k) = {}_B M_B(p(\phi)).$$
Ist $p(\phi) = 0$, so folgt daraus $p({}_B M_B(\phi)) = {}_B M_B(p(\phi)) = 0$. Wegen der Injektivität von ${}_B M_B$ folgt umgekehrt aus $p({}_B M_B(\phi)) = {}_B M_B(p(\phi)) = 0$ auch $p(\phi) = 0$. Damit ist $p \in \ker(\mathrm{ev}_\phi)$ genau dann, wenn $p \in \ker(\mathrm{ev}_M)$ und somit $I_\phi = I_M$ für $M = {}_B M_B(\phi)$. Da q_ϕ und q_M durch die Ideale I_ϕ und I_M eindeutig bestimmt sind, folgt $q_\phi = q_M$.

4. Sind $M, N \in \mathrm{Mat}(n \times n, \mathbb{K})$ ähnlich, so gibt es einen Endomorphismus $\phi \in \mathrm{End}_\mathbb{K}(V)$ und geordnete Basen B, C von V mit $M = {}_B M_B(\phi)$, $N = {}_C M_C(\phi)$. Mit 3. folgt $q_M = q_\phi = q_N$. $\square$

Beispiel 8.3.2: Sei V ein n-dimensionaler Vektorraum über $\mathbb{K}$ mit $n \in \mathbb{N}$.

1. Es gilt $\deg(q_\phi) \geq 1$ für alle Endomorphismen $\phi : V \to V$.

   Denn das Nullpolynom ist nicht normiert, und das einzige normierte Polynom von Grad 0 ist das konstante Polynom $x^0 = 1$. Für $V \neq 0$ gilt aber $x^0(\phi) = \mathrm{id}_V \neq 0$.

2. Für eine Streckung $\phi = \alpha \, \mathrm{id}_V$ ist $q_\phi = x - \alpha$. Denn $x - \alpha$ ist normiert, $q_\phi(\phi) = \phi - \alpha \, \mathrm{id}_V = 0$ und $\deg(x - \alpha) = 1$.

3. Ist $\phi : V \to V$ eine Involution, so gilt $\phi^2 = \mathrm{id}_V$. Daraus folgt $t(\phi) = 0$ für das Polynom $t = x^2 - 1 = (x-1)(x+1)$. Ist $\phi = \pm \mathrm{id}_V$, so gilt nach 2. $q_\phi = x \mp 1$. Ist $\phi \neq \pm \mathrm{id}_V$, so folgt $q_\phi = x^2 - 1$.

4. Ist $\phi : V \to V$ ein Projektor, so ist $\phi^2 = \phi$ und damit $t(\phi) = 0$ für $t = x^2 - x = x(x-1)$. Ist $\phi = 0$, so folgt $q_\phi = x$, und ist $\phi = \mathrm{id}_V$, so folgt $q_\phi = x - 1$. Ansonsten ist $q_\phi = x(x-1)$.

Die Beispiele zeigen, dass man das Minimalpolynom von $\phi : V \to V$ bestimmen kann, indem man ein Polynom $p \in \mathbb{K}[x] \setminus \{0\}$ mit $p(\phi) = 0$ findet, und dann für alle Teiler q von p untersucht, ob auch $q(\phi) = 0$ gilt. Für einen Endomorphismus $\phi : V \to V$ müssen wir dazu aber zunächst ein Polynom p mit $p(\phi) = 0$ finden. Das liefert der Satz von Cayley-Hamilton.

Satz 8.3.3: (**Satz von Cayley-Hamilton**)
Sei $\mathbb{K}$ ein Körper, $n \in \mathbb{N}$ und V ein n-dimensionaler Vektorraum über $\mathbb{K}$. Dann gilt für jeden Endomorphismus $\phi : V \to V$ und jede Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{K})$

$$\mathrm{ev}_\phi(p_\phi) = p_\phi(\phi) = 0 \quad \text{und} \quad \mathrm{ev}_A(p_A) = p_A(A) = 0.$$

Also ist das Minimalpolynom q_ϕ oder q_A ein Teiler des charakteristischen Polynoms p_ϕ oder p_A.

Beweis:
1. Sei zunächst $\phi : V \to V$ trigonalisierbar mit paarweise verschiedenen Eigenwerten $\lambda_1, ..., \lambda_m$. Das charakteristische Polynom p_ϕ ist dann nach Satz 8.2.4 gegeben durch

$$p_\phi = (-1)^n (x - \lambda_1)^{h_1} \cdots (x - \lambda_m)^{h_m} \quad \text{mit} \quad h_i = \dim_{\mathbb{K}} H(\phi, \lambda_i).$$

Nach Satz 7.1.8 gilt gilt $V = \ker(\phi_{\lambda_1}^{f_1} \circ ... \circ \phi_{\lambda_m}^{f_m})$, wobei f_i der Fitting-Index des Endomorphismus $\phi_{\lambda_i} = \phi - \lambda_i \,\mathrm{id}_V : V \to V$ ist. Also ergibt sich

$$\phi_{\lambda_1}^{f_1} \circ ... \circ \phi_{\lambda_m}^{f_m} = (\phi - \lambda_1)^{f_1} \circ ... \circ (\phi - \lambda_m)^{f_m} = 0 \quad \Rightarrow \quad t_\phi = (x - \lambda_1)^{f_1} \cdots (x - \lambda_m)^{f_m} \in I_\phi.$$

Da nach dem Fitting-Lemma und Korollar 7.1.7 für die Fitting-Indizes gilt $1 \leq f_i \leq h_i$, ist das charakteristische Polynom $p_\phi = (-1)^n (x - \lambda_1)^{h_1 - f_1} \cdots (x - \lambda_m)^{h_m - f_m} \cdot t_\phi$ ein Vielfaches von t_ϕ und somit ebenfalls im Ideal I_ϕ enthalten. Also gilt $p_\phi(\phi) = 0$. Die Aussage für trigonalisierbare Matrizen $A \in \mathrm{Mat}(n \times n, \mathbb{K})$ ergibt sich dann aus der Betrachtung von $\phi_A : \mathbb{K}^n \to \mathbb{K}^n$, $v \mapsto A \cdot v$.

2. Sei nun $\phi : V \to V$ ein beliebiger Endomorphismus und $M := {}_C M_C(\phi)$ seine beschreibende Matrix bezüglich einer beliebigen geordneten Basis C von V. Dann gibt es einen Körper $\mathbb{L}$, so dass $\mathbb{K} \subseteq \mathbb{L}$ ein Teilkörper ist und $p_M = p_\phi \in \mathbb{K}[x] \subseteq \mathbb{L}[x]$ über $\mathbb{L}$ in Linearfaktoren zerfällt.

Damit ist M nach Satz 8.2.4 trigonalisierbar über $\mathbb{L}$. Es gibt eine Matrix $S \in \mathrm{GL}(n, \mathbb{L})$, so dass $N = S \cdot M \cdot S^{-1} \in \mathrm{Mat}(n \times n, \mathbb{L})$ eine obere Dreiecksmatrix ist. Nach 1. ist $p_M(M) = 0$, also $p_M \in I_M$, und da nach Satz 8.3.1, 3. $I_\phi = I_M$ gilt, ist $p_\phi = p_M \in I_\phi = I_M$ und somit $p_\phi(\phi) = 0$. □

Der Satz von Cayley-Hamilton liefert als Kandidaten für das Minimalpolynom eines Endomorphismus $\phi : V \to V$ alle Teiler seines charakteristischen Polynoms p_ϕ. Wie Beispiel 8.3.2 zeigt, ist dabei aber zunächst nicht offensichtlich, welche Teiler des Polynoms p_ϕ in Frage kommen. Außerdem wünscht man sich eine Charakterisierung des Minimalpolynoms durch bekannte Konzepte wie Haupträume, Eigenräume und Fitting-Indizes.

Satz 8.3.4: Sei V ein n-dimensionaler $\mathbb{K}$-Vektorraum und $\phi : V \to V$ ein Endomorphismus mit paarweise verschiedenen Eigenwerten $\lambda_1, ..., \lambda_m \in \mathbb{K}$. Dann ist

$$q_\phi = (x - \lambda_1)^{f_1} \cdots (x - \lambda_m)^{f_m} \cdot s_\phi$$

wobei f_i der Fitting-Index von $\phi_{\lambda_i} = \phi - \lambda_i \,\mathrm{id}_V$ ist und $s_\phi \in \mathbb{K}[x]$ keine Nullstellen in $\mathbb{K}$ hat.

Beweis:
Nach Satz 7.1.8 gilt $V = U \oplus W$, wobei $U = H(\phi, \lambda_1) \oplus ... \oplus H(\phi, \lambda_M) = \ker(\phi_{\lambda_1}^{f_1} \circ ... \circ \phi_{\lambda_m}^{f_m})$, $\phi|_U : U \to U$ ein trigonalisierbarer Endomorphismus und $\phi|_W : W \to W$ ein Isomorphismus ohne Eigenwerte in $\mathbb{K}$ ist. Nach Satz 8.2.4 gilt mit $h_i = \dim_{\mathbb{K}} H(\phi, \lambda_i)$

$$p_\phi = p_{\phi|_U} \cdot p_{\phi|_W} = (-1)^{h_1+...+h_m}(x - \lambda_1)^{h_1} \cdots (x - \lambda_m)^{h_m} \cdot p_{\phi|_W}$$

wobei $p_{\phi|_W}$ ein Polynom ohne Nullstellen in $\mathbb{K}$ ist.

Da das Minimalpolynom q_ϕ per Definition jedes Polynom $t \in \mathbb{K}[x]$ mit $t(\phi) = 0$ teilt, aber kein Polynom $t \in \mathbb{K}[x]$ mit $t(\phi) \neq 0$, reicht es zu zeigen, dass für die Polynome

$$t = q \cdot p_{\phi|_W} = (x - \lambda_1)^{f_1} \cdots (x - \lambda_m)^{f_m} \cdot p_{\phi|_W} \qquad t_i = t/(x - \lambda_i)$$

$t(\phi) = 0$ und $t_i(\phi) \neq 0$ für alle $i \in \{1, ..., m\}$ gilt. Denn daraus folgt für das Minimalpolynom $q_\phi = (x - \lambda_1)^{f_1} \cdots (x - \lambda_m)^{f_m} \cdot s_\phi$, wobei s_ϕ ein Teiler von $p_{\phi|_W}$ ist, also ohne Nullstellen in $\mathbb{K}$.

Nach dem Satz von Cayley-Hamilton gilt $p_{\phi|_W}(\phi|_W) = 0$, also $p_{\phi|_W}(\phi)(w) = 0$ für $w \in W$. Aus $U = \ker(\phi_{\lambda_1}^{f_1} \circ ... \circ \phi_{\lambda_m}^{f_m})$ folgt $q(\phi)(u) = \phi_{\lambda_1}^{f_1} \circ ... \circ \phi_{\lambda_m}^{f_m}(u) = 0$ für alle $u \in U$. Da $V = U \oplus W$ gibt es zu $v \in V$ Elemente $u \in U$ und $w \in W$ mit $v = u + w$, und es folgt

$$t(\phi)(v) = q(\phi) \circ p_{\phi|_W}(\phi)(u + w) = p_{\phi|_W}(\phi)(q(\phi)(u)) + q(\phi)(p_{\phi|_W}(\phi)(w)) = 0 + 0 = 0.$$

Per Definition der Fitting-Indizes gibt es zu jedem $i \in \{1, ..., m\}$ ein $v_i \in H(\phi, \lambda_i)$ mit $v_i' = \phi_{\lambda_i}^{f_i - 1}(v_i) \neq 0$. Dann ist v_i' ein Eigenvektor von ϕ zum Eigenwert λ_i und

$$\begin{aligned} t_i(\phi)(v_i) &= \phi_{\lambda_1}^{f_1} \circ ... \circ \phi_{\lambda_{i-1}}^{f_{i-1}} \circ \phi_{\lambda_i}^{f_i - 1} \circ \phi_{\lambda_{i+1}}^{f_{i+1}} \circ ... \circ \phi_{\lambda_m}^{f_m} \circ p_{\phi|_W}(\phi)(v_i) \\ &= \phi_{\lambda_1}^{f_1} \circ ... \circ \phi_{\lambda_{i-1}}^{f_{i-1}} \circ \phi_{\lambda_{i+1}}^{f_{i+1}} \circ ... \circ \phi_{\lambda_m}^{f_m} \circ p_{\phi|_W}(\phi) \circ \phi_{\lambda_i}^{f_i - 1}(v_i) \\ &= \phi_{\lambda_1}^{f_1} \circ ... \circ \phi_{\lambda_{i-1}}^{f_{i-1}} \circ \phi_{\lambda_{i+1}}^{f_{i+1}} \circ ... \circ \phi_{\lambda_m}^{f_m} \circ p_{\phi|_W}(\phi)(v_i') \\ &= (\lambda_i - \lambda_1)^{f_1} \cdots (\lambda_i - \lambda_{i-1})^{f_{i-1}} (\lambda_i - \lambda_{i+1})^{f_{i+1}} \cdots (\lambda_i - \lambda_m)^{f_m} p_{\phi|_W}(\lambda_i)\, v_i' \neq 0, \end{aligned}$$

also $t_i(\phi) \neq 0$ für alle $i \in \{1, ..., m\}$. □

Mit Hilfe von Satz 8.3.4 können wir nun auch die Diagonalisierbarkeit eines Endomorphismus ϕ durch die Eigenschaften eines Polynoms charakterisieren, nämlich des Minimalpolynoms q_ϕ. Es stellt sich heraus, dass ϕ genau dann diagonalisierbar ist, wenn das Minimalpolynom q_ϕ in Linearfaktoren zerfällt und jede Nullstelle von q_ϕ Vielfachheit 1 hat.

Korollar 8.3.5: Sei $\mathbb{K}$ ein Körper, V ein endlich-dimensionaler $\mathbb{K}$-Vektorraum und $\phi : V \to V$ ein Endomorphismus mit paarweise verschiedenen Eigenwerten $\lambda_1, ..., \lambda_m$. Sei q_ϕ das Minimalpolynom von ϕ und f_i der Fitting-Index von $\phi_{\lambda_i} = \phi - \lambda_i \operatorname{id}_V : V \to V$. Dann gilt:

1. ϕ ist trigonalisierbar genau dann, wenn $q_\phi = (x - \lambda_1)^{f_1} \cdots (x - \lambda_m)^{f_m}$.
2. ϕ ist diagonalisierbar genau dann, wenn $q_\phi = (x - \lambda_1) \cdots (x - \lambda_m)$.

Beweis:
1. Ist ϕ trigonalisierbar, so zerfällt nach Satz 8.2.4 das charakteristische Polynom p_ϕ in Linearfaktoren und damit auch sein Teiler q_ϕ. Mit Satz 8.3.4 folgt $q_\phi = (x - \lambda_1)^{f_1} \cdots (x - \lambda_m)^{f_m}$.

Ist umgekehrt $q_\phi = (x-\lambda_1)^{f_1}\cdots(x-\lambda_m)^{f_m}$, so ist $q_\phi(\phi) = \phi_{\lambda_1}^{f_1}\circ ... \circ \phi_{\lambda_m}^{f_m} = 0$. Mit Satz 7.1.8 folgt $V = \ker(\phi_{\lambda_1}^{f_1}\circ ... \circ \phi_{\lambda_m}^{f_m}) = H(\phi,\lambda_1)\oplus ... \oplus H(\phi,\lambda_m)$, und somit ist ϕ trigonalisierbar.

2. Ein Endomorphismus ϕ ist nach Satz 8.2.4 genau dann diagonalisierbar, wenn er trigonalisierbar ist und für alle Eigenwerte λ_i gilt

$$\mathrm{def}(\phi-\lambda_i\mathrm{id}_V) = \dim_{\mathbb{K}} E(\phi,\lambda_i) = g(\phi,\lambda_i) = a(\phi,\lambda_i) = \dim_{\mathbb{K}} H(\phi,\lambda_i) = \mathrm{def}(\phi-\lambda_i\mathrm{id}_V)^{f_i}.$$

Ersteres ist nach 1. genau dann der Fall, wenn $q_\phi = (x-\lambda_1)^{f_1}\cdots(x-\lambda_m)^{f_m}$ gilt, und die zweite Bedingung ist äquivalent zu $f_i = 1$ für alle $i\in\{1,...,m\}$. □

Beispiel 8.3.6: Für die Matrix

$$A = \begin{pmatrix} 1 & 2 & 0 & 1\\ 0 & 1 & 1 & -1\\ 0 & 0 & -1 & 4\\ 0 & 0 & 1 & -1\end{pmatrix} \in \mathrm{Mat}(4\times 4,\mathbb{R})$$

ist das charakteristische Polynom gegeben durch

$$p_A = \det\begin{pmatrix} 1-x & 2 & 0 & 1\\ 0 & 1-x & 1 & -1\\ 0 & 0 & -1-x & 4\\ 0 & 0 & 1 & -1-x\end{pmatrix} = \det\begin{pmatrix} 1-x & 2\\ 0 & 1-x\end{pmatrix}\cdot\det\begin{pmatrix} -1-x & 4\\ 1 & -1-x\end{pmatrix}$$
$$= (1-x)^2((1+x)^2-4) = (x-1)^2(x^2+2x-3) = (x-1)^3(x+3).$$

Also sind die Eigenwerte $1, -3$, und $\dim_{\mathbb{K}} H(\phi,1) = 3$, $\dim_{\mathbb{K}} H(A,-3) = \dim_{\mathbb{K}} E(A,-3) = 1$. Damit kommen nur die Polynome $(x-1)(x-3)$, $(x-1)^2(x-3)$ und $(x-1)^3(x-3)$ als Minimalpolynom q_A in Frage.

Zur Bestimmung von q_A berechnet man den Fitting-Index von $A_1 = A - \mathbb{1}_4$. Dabei ergibt sich

$$A_1 = \begin{pmatrix} 0 & 2 & 0 & 1\\ 0 & 0 & 1 & -1\\ 0 & 0 & -2 & 4\\ 0 & 0 & 1 & -2\end{pmatrix} \qquad\Rightarrow\qquad \mathrm{def}(A_1) = 1$$

$$A_1^2 = \begin{pmatrix} 0 & 2 & 0 & 1\\ 0 & 0 & 1 & -1\\ 0 & 0 & -2 & 4\\ 0 & 0 & 1 & -2\end{pmatrix}\cdot\begin{pmatrix} 0 & 2 & 0 & 1\\ 0 & 0 & 1 & -1\\ 0 & 0 & -2 & 4\\ 0 & 0 & 1 & -2\end{pmatrix} = \begin{pmatrix} 0 & 0 & 3 & -4\\ 0 & 0 & -3 & 6\\ 0 & 0 & 8 & -16\\ 0 & 0 & -4 & 8\end{pmatrix} \qquad\Rightarrow\qquad \mathrm{def}(A_1^2) = 2$$

Da $\dim_{\mathbb{K}} H(A,1) = 3$ gilt $f_1 \leq 3$ für den Fitting-Index f_1 von A_1. Da $H(A,1) = \ker(A_1^{f_1})$ und $\mathrm{def} A_1^2 = 2 < \dim_{\mathbb{K}} H(A,1)$ gilt $f_1 \geq 3$, und damit $f_1 = 3$.

Alternativ oder zur Kontrolle kann man A_1^3 berechnen und erhält

$$A_1^3 = \begin{pmatrix} 0 & 0 & 3 & -4\\ 0 & 0 & -3 & 6\\ 0 & 0 & 8 & -16\\ 0 & 0 & -4 & 8\end{pmatrix}\cdot\begin{pmatrix} 0 & 2 & 0 & 1\\ 0 & 0 & 1 & -1\\ 0 & 0 & -2 & 4\\ 0 & 0 & 1 & -2\end{pmatrix} = \begin{pmatrix} 0 & 0 & -10 & 20\\ 0 & 0 & 12 & -24\\ 0 & 0 & -32 & 64\\ 0 & 0 & 16 & -32\end{pmatrix} \qquad\Rightarrow\qquad \mathrm{def}(A_1^3) = 3.$$

Also ist der Fitting-Index von A_1 gleich 3 und $q_A = (x-1)^3(x+3)$. Damit ist A zwar trigonalisierbar, aber nicht diagonalisierbar über $\mathbb{R}$.

An was man sich erinnern sollte:

Die wichtigsten Begriffe und Definitionen:

- Polynom, Monom, Polynomgrad, Leitkoeffizient, normiert
- Evaluationsabbildung, Nullstelle, Vielfachheit
- Polynomdivision, Rest, Teiler, Vielfaches,
- Ideal, von $M \subseteq \mathbb{K}[x]$ erzeugtes Ideal, Hauptideal, Hauptidealring
- charakteristisches Polynom, algebraische Vielfachheit
- Minimalpolynom

Die wichtigsten Aussagen:

- Abspalten von Nullstellen, Polynomdivision,
- $\mathbb{K}[x]$ ist ein Hauptidealring,
- Diagonalisierbarkeitskriterium, Trigonalisierbarkeitskriterium,
- Formeln für Spur und Determinante trigonalisierbarer Matrix,
- Satz von Cayley-Hamilton,
- Berechnung von charakteristischen und Minimalpolynom,
- Formeln für charakteristisches und Minimalpolynom im trigonalisierbaren Fall

Die wichtigsten Beispiele:

- **Zerfallen in Linearfaktoren:** $x^2 + 1$, $x^2 - 2$
- **Trigonalisierbarkeit:** Trigonalisierbarkeit über $\mathbb{C}$
- **Minimalpolynom:** Streckung, Projektor, Involution, nilpotente Endomorphismen

9 Die Jordan-Normalform

9.1 Jordanketten und Jordanbasen

Nachdem wir die Frage der Trigonalisierbarkeit und Diagonalisierbarkeit von Endomorphismen und Matrizem weitgehend geklärt haben, wenden wir uns nun wieder der Frage der Klassifikation von Matrizen und dem Problem der Ähnlichkeit zu. Gesucht ist eine leicht handhabbare Bedingung, so dass zwei Matrizen genau dann ähnlich sind, wenn diese Bedingung erfüllt ist, oder, dazu äquivalent, eine Liste von Matrizen in $\mathrm{Mat}(n\times n,\mathbb{K})$, so dass jede Matrix $A \in \mathrm{Mat}(n\times n,\mathbb{K})$ ähnlich zu genau einer Matrix in der Liste ist.

Korollar 7.1.10 liefert uns notwendige Bedingungen für die Ähnlichkeit zweier Matrizen $M, N \in \mathrm{Mat}(n \times n,\mathbb{K})$, nämlich, dass sie

(i) die gleichen Eigenwerte haben,
(ii) $\mathrm{def}(M - \lambda\mathbb{1}_n)^k = \mathrm{def}(N - \lambda\mathbb{1}_n)^k$ für jeden Eigenwert λ und jedes $k \in \mathbb{N}$.

In diesem Abschnitt zeigen wir, dass diese Bedingungen für trigonalisierbare Matrizen auch hinreichend sind. Damit können wir für jeden algebraisch abgeschlossenen Körper $\mathbb{K}$ dann eine Liste von Matrizen angeben, so dass jede Matrix mit Einträgen in $\mathbb{K}$ ähnlich zu genau einer Matrix in der Liste ist.

Dazu betrachten wir zunächst Endomorphismen und wählen eine an die Struktur der Haupträume angepasste Basis. In Lemma 7.1.1 wurde gezeigt, dass der Hauptraum $H(\phi,\lambda)$ für jeden Eigenwert λ eines Endomorphismus $\phi : V \to V$ eine aufsteigende Kette von Kernen der Form

$$\{0\} = \ker(\phi_\lambda^0) \overset{\xleftarrow{\phi_\lambda}}{\subseteq} \ker(\phi_\lambda) \overset{\xleftarrow{\phi_\lambda}}{\subseteq} \ker(\phi_\lambda^2) \overset{\xleftarrow{\phi_\lambda}}{\subseteq} \overset{...}{...} \overset{\xleftarrow{\phi_\lambda}}{\subseteq} \ker(\phi_\lambda^{f_\lambda-1}) \overset{\xleftarrow{\phi_\lambda}}{\subseteq} \ker(\phi_\lambda^{f_\lambda}) = H(\phi,\lambda)$$

enthält, wobei f_λ den Fitting-Index von $\phi_\lambda = \phi - \lambda\mathrm{id}_V$ bezeichnet. Da $\phi_\lambda(\ker(\phi_\lambda^k)) \subseteq \ker(\phi_\lambda^{k-1})$, ist es naheliegend, für jeden Vektor $v \in \ker(\phi_\lambda^k)$ in der Basis auch all seine nicht verschwindenden Bilder $\phi_\lambda^l(v) \in \ker(\phi_\lambda^{k-l})$ in die Basis aufzunehmen. Die beschreibende Matrix bekommt so eine besonders einfache Form.

Definition 9.1.1: Sei $\mathbb{K}$ ein Körper, V ein endlich-dimensionaler Vektorraum über $\mathbb{K}$ und $\phi : V \to V$ ein Endomorphismus.

1. Eine **Jordankette** der Länge $m \in \mathbb{N}$ zum Eigenwert λ von ϕ ist ein m-Tupel $(v_1, ..., v_m)$ von Vektoren $v_i \in V \setminus \{0\}$ mit $\phi(v_i) = \lambda v_i + v_{i-1}$ für $i \in \{2, ..., m\}$ und $\phi(v_1) = \lambda v_1$.
2. Eine **Jordanbasis** für ϕ ist eine Basis $B = (v_1^1, ..., v_{m_1}^1, v_1^2, ..., v_{m_2}^2, ..., v_1^k, ..., v_{m_k}^k)$ von V, so dass $(v_1^j, ..., v_{m_j}^j)$ eine Jordankette ist für alle $j \in \{1, ..., k\}$.

Wir konstruieren nun für einen trigonalisierbaren Endomorphismus $\phi : V \to V$ eine Jordanbasis. Dazu muss zunächst gezeigt werden, dass jede Jordankette von ϕ linear unabhängig ist. Da man außerdem verschiedene Jordanketten in eine Basis kombinieren möchte, muss auch untersucht werden, unter welchen Bedingungen zwei verschiedene Jordanketten von ϕ linear unabhängig sind. Für Jordanketten zu verschiedenen Eigenwerten folgt die lineare Unabhängigkeit direkt aus der Tatsache, dass eine Jordankette zum Eigenwert λ im Hauptraum $H(\phi,\lambda)$ enthalten ist und $H(\phi,\lambda) \cap H(\phi,\mu) = \{0\}$ für $\lambda \neq \mu$ gilt. Für verschiedene Jordanketten zum selben Eigenwert λ ist dies jedoch zunächst nicht offensichtlich.

Lemma 9.1.2: Sei V ein endlich-dimensionaler $\mathbb{K}$-Vektorraum und $\phi : V \to V$ ein Endomorphismus. Dann gilt für jeden Eigenwert λ von ϕ:

1. Jede Jordan-Kette $(v_1, ..., v_m)$ zum Eigenwert λ ist linear unabhängig.
2. Sind $(v_1^1, ..., v_{m_1}^1)$, ..., $(v_1^k, ..., v_{m_k}^k)$ Jordanketten zum Eigenwert λ so dass $(v_1^1, ..., v_1^k)$ linear unabhängig ist, so ist auch $(v_1^1, ..., v_{m_1}^1, v_1^2, ..., v_{m_2}^2, ..., v_1^k, ..., v_{m_k}^k)$ linear unabhängig.

Beweis:
1. Die Aussage folgt per Induktion über m:

$m = 1$: Eine Jordan-Kette (v_1) ist linear unabhängig, da per Definition $v_1 \neq 0$ gilt.

$m-1 \to m$: Sei die Aussage bewiesen für alle Jordanketten der Länge $\leq m-1$ und sei $(v_1, ..., v_m)$ eine Jordankette zum Eigenwert λ der Länge m. Dann gilt $\phi_\lambda(v_1) = 0$ und $\phi_\lambda(v_k) = v_{k-1}$ für $k \in \{2, ..., m\}$ und $\phi_\lambda = \phi - \lambda \mathrm{id}_V$. Sind $\mu_1, ..., \mu_m \in \mathbb{K}$ mit $\Sigma_{k=1}^m \mu_k v_k = 0$, so folgt

$$0 = \phi_\lambda(0) = \phi_\lambda(\Sigma_{k=1}^m \mu_k v_k) = \Sigma_{k=1}^m \mu_k \phi_\lambda(v_k) = \Sigma_{k=2}^m \mu_k v_{k-1} = \Sigma_{j=1}^{m-1} \mu_{k+1} v_k.$$

Da auch $(v_1, ..., v_{m-1})$ eine Jordankette zum Eigenwert λ ist, ist diese nach Induktionsvoraussetzung linear unabhängig, und es folgt $\mu_2 = ... = \mu_m = 0$. Also gilt $\Sigma_{k=1}^m \mu_k v_k = \mu_1 v_1 = 0$, und mit $v_1 \neq 0$ folgt $\mu_1 = ... = \mu_m = 0$. Damit ist auch $(v_1, ..., v_m)$ linear unabhängig.

2. Ist $C := (v_1^1, ..., v_{m_1}^1, v_1^2, ..., v_{m_2}^2, ..., v_1^k, ..., v_{m_k}^k)$ linear abhängig, so gibt es $\mu_{ij} \in \mathbb{K}$, die nicht alle verschwinden, mit $\Sigma_{j=1}^k \Sigma_{i=1}^{m_j} \mu_{ij} v_i^j = 0$. Also existiert $n = \max\{i \in \mathbb{N} | \exists j \in \{1, ..., k\} : \mu_{ij} \neq 0\}$. Da $\phi_\lambda^s(v_r^j) = v_{r-s}^j$ für $s < r$ und $\phi_\lambda^s(v_r^j) = 0$ für $s \geq r$, folgt dann

$$0 = \phi_\lambda^{n-1}(\Sigma_{j=1}^k \Sigma_{i=1}^{m_j} \mu_{ij} v_i^j) = \Sigma_{j=1}^k \Sigma_{i=1}^{m_j} \mu_{ij} \phi_\lambda^{n-1}(v_i^j) = \Sigma_{j=1}^k \Sigma_{i=n}^{m_j} \mu_{ij} v_{i-n+1}^j = \Sigma_{j=1}^k \mu_{nj} v_1^j,$$

wobei $\phi^0 = \mathrm{id}_V$ und im letzten Schritt benutzt wurde, dass $\mu_{ij} = 0$ für $i > n$. Da nach Voraussetzung $(v_1^1, ..., v_1^k)$ linear unabhängig ist, folgt $\mu_{nj} = 0$ für alle $j \in \{1, ..., k\}$, im Widerspruch zur Definition von n. Also ist C linear unabhängig. $\square$

Da Jordanketten und Jordanbasen an die Struktur der Haupträume $H(\phi, \lambda)$, also der aufsteigenden Kette $\{0\} \subsetneq \ker(\phi_\lambda) \subsetneq ... \subsetneq \ker(\phi_\lambda^f) = H(\phi, \lambda)$, angepasst sind, nimmt die beschreibende Matrix des Endomorphismus $\phi : V \to V$ bezüglich einer Jordanbasis eine besonders einfache Form an. Dies führt auf den Begriff des *Jordanblocks* und auf die *Jordan-Normalform.*

Jordanblöcke sind die beschreibende Matrizen der Einschränkungen von ϕ auf Jordanketten. Sie enthalten in der Diagonalen den zugehörigen Eigenwert und in der Diagonalreihe oberhalb der Diagonalen nur Einsen. Kombiniert man mehrere Jordanblöcke zu einer blockdiagonalen Matrix, so erhält man eine Matrix in Jordan-Normalform.

Definition 9.1.3: Sei $\mathbb{K}$ ein Körper.

1. Ein **Jordanblock** oder **Jordan-Kästchen** der Länge $m \in \mathbb{N}$ ist eine Matrix der Form

$$J_m(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \lambda & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & \dots & 0 & \lambda & 1 \\ 0 & \dots & \dots & 0 & \lambda \end{pmatrix} \in \mathrm{Mat}(m \times m, \mathbb{K}).$$

2. Eine Matrix $M \in \mathrm{Mat}(n \times n, \mathbb{K})$ ist in **Jordan-Normalform**, wenn

$$M = \begin{pmatrix} J_{m_1}(\lambda_1) & 0 & \dots & 0 \\ 0 & J_{m_2}(\lambda_2) & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & J_{m_k}(\lambda_k) \end{pmatrix}$$

mit Jordanblöcken $J_{m_i}(\lambda_i) \in \mathrm{Mat}(m_i \times m_i, \mathbb{K})$ und $m_1 + ... + m_k = n$.

Lemma 9.1.4: Sei V ein endlich-dimensionaler $\mathbb{K}$-Vektorraum und $\phi : V \to V$ ein Endomorphismus. Dann gilt:

1. Ist $A = (v_1, ..., v_m)$ eine Jordankette für ϕ zum Eigenwert λ, so ist $U := \mathrm{span}_{\mathbb{K}}(A)$ ein ϕ-invarianter Unterraum, und die beschreibende Matrix von $\phi|_U : U \to U$ bezüglich A ist ein Jordanblock:

$$_A M_A(\phi|_U) = J_m(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \lambda & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & \dots & 0 & \lambda & 1 \\ 0 & \dots & \dots & 0 & \lambda \end{pmatrix}.$$

2. Ist $B = (v_1^1, ..., v_{m_1}^1, v_1^2, ..., v_{m_2}^2, ..., v_1^k, ..., v_{m_k}^k)$ eine Jordanbasis für ϕ, so ist die beschreibende Matrix von ϕ bezüglich B in Jordan-Normalform

$$_B M_B(\phi) = \begin{pmatrix} J_{m_1}(\lambda_1) & 0 & \dots & 0 \\ 0 & J_{m_2}(\lambda_2) & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & J_{m_k}(\lambda_k) \end{pmatrix}$$

wobei λ_j der Eigenwert für die Jordankette $(v_1^j, ..., v_{m_j}^j)$ ist.

Beweis:
1. Per Definition gilt für eine Jordankette $A = (v_1, ..., v_m)$ zum Eigenwert λ

$$\phi(v_i) = \lambda v_i + v_{i-1} \text{ für } i \in \{2, .., m\} \qquad \phi(v_1) = \lambda v_1$$

Daraus folgt insbesondere $\phi(U) \subseteq U$ für $U := \mathrm{span}_{\mathbb{K}}(A)$. Also ist die Einschränkung von ϕ auf U ein Endomorphismus $\phi|_U : U \to U$. Da nach Lemma 9.1.2 jede Jordankette linear unabhängig ist, ist A eine Basis von U, und die beschreibende Matrix von $\phi|_U$ ist $_A M_A(\phi|_U) = J_m(\lambda)$.

2. Ist $B = (v_1^1, ..., v_{m_1}^1, ..., v_1^2, ..., v_{m_2}^2, ..., v_1^k, ..., v_{m_k}^k)$ eine Jordanbasis für ϕ mit Jordanketten $A_i = (v_1^i, ..., v_{m_i}^i)$ zum Eigenwert λ_i, so ist nach 1. jeder Unterraum $U_i := \mathrm{span}_{\mathbb{K}}(A_i)$ invariant unter ϕ und die darstellende Matrix von $\phi|_{U_i} : U_i \to U_i$ bezüglich A_i ist der Jordanblock $M_i := {_{A_i}M_{A_i}}(\phi|_{U_i}) = J_{m_i}(\lambda_i)$. □

Wir zeigen nun, dass zu jedem trigonalisierbaren Endomorphismus $\phi : V \to V$ eines endlich-dimensionalen $\mathbb{K}$-Vektorraums V eine Jordanbasis B von V existiert. Die beschreibende Matrix von ϕ bezüglich B ist dann eine Matrix in Jordan-Normalform.

Satz 9.1.5: (**Existenz der Jordan-Normalform**)
Sei V ein endlich-dimensionaler $\mathbb{K}$-Vektorraum und $\phi : V \to V$ ein trigonalisierbarer Endomorphismus. Dann gibt es eine Jordanbasis B für ϕ und es gilt

$$ {}_B M_B(\phi) = \begin{pmatrix} J_{m_1}(\lambda_1) & 0 & \ldots & 0 \\ 0 & J_{m_2}(\lambda_2) & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & J_{m_k}(\lambda_k) \end{pmatrix}, $$

wobei $\lambda_1, ..., \lambda_k \in \mathbb{K}$ nicht notwendigerweise verschiedene Eigenwerte von ϕ sind, $J_{m_i}(\lambda_i)$ Jordanblöcke und $m_1 + ... + m_k = \dim_{\mathbb{K}}(V)$ gilt.

Diese Matrix heißt **Jordan-Normalform** von ϕ. Eine **Jordan-Normalform** einer Matrix $A \in \mathrm{Mat}(n\times n, \mathbb{K})$ ist eine **Jordan-Normalform** des Endomorphismus $\phi_A : \mathbb{K}^n \to \mathbb{K}^n$, $v \mapsto A\cdot v$.

Beweis:
Es reicht zu beweisen, dass jeder trigonalisierbare Endomorphismus $\phi : V \to V$ eine Jordanbasis besitzt. Die Aussage über die beschreibende Matrix von ϕ folgt dann aus Lemma 9.1.4.

1. Wir beweisen die Existenz einer Jordanbasis für den Fall, dass der trigonalisierbare Endomorphismus $\phi : V \to V$ nur einen einzigen Eigenwert λ besitzt.

Nach Lemma 7.1.1 ist ϕ trigonalisierbar genau dann, wenn $\phi_\lambda := \phi - \lambda \mathrm{id}_V : V \to V$ nilpotent ist. Jordanketten von ϕ_λ zum Eigenwert 0 entsprechen Jordanketten von ϕ zum Eigenwert λ und umgekehrt. Also ist eine geordnete Basis B genau dann eine Jordanbasis von ϕ, wenn sie eine Jordanbasis von ϕ_λ ist. Damit reicht es, den Fall $\lambda = 0$ zu betrachten und zu zeigen, dass jeder nilpotente Endomorphismus $\phi : V \to V$ eine Jordanbasis besitzt. Der Beweis erfolgt durch Induktion über den Fitting-Index f von ϕ.

$f = 1$: Ist $f = 1$, so gilt wegen der Nilpotenz von ϕ und per Definition des Fitting-Index $\ker(\phi) = V$, und jeder Vektor $v \in V$ ist ein Eigenvektor von ϕ zum Eigenwert 0. Also ist jede Basis von V eine Jordanbasis für ϕ, die nur Jordanketten der Länge 1 enthält.

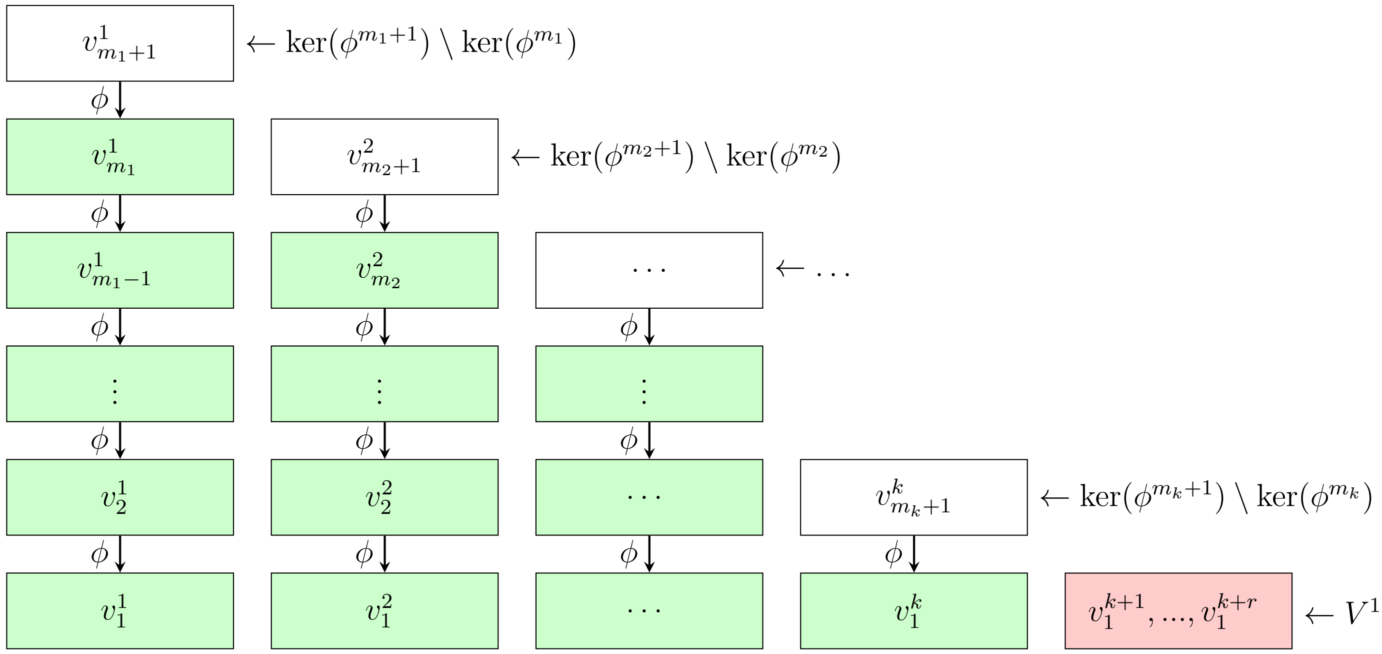
$f - 1 \to f$: Sei die Aussage bewiesen für alle $\mathbb{K}$-Vektorräume V und nilpotenten Endomorphismen $\phi : V \to V$ mit Fitting-Index $\leq f - 1$. Sei $\phi : V \to V$ nilpotent mit Fitting-Index f. Dann gilt wegen der Nilpotenz von ϕ und per Definition des Fitting-Index $V = \ker(\phi^f)$, und wir erhalten eine aufsteigende Kette von Kernen wie im Beweis von Lemma 7.1.1

$$ \{0\} = \ker(\phi^0) \underset{\neq}{\overset{\xleftarrow{\phi}}{\subseteq}} \ker(\phi) \underset{\neq}{\overset{\xleftarrow{\phi}}{\subseteq}} \overset{...}{...} \underset{\neq}{\overset{\xleftarrow{\phi}}{\subseteq}} \ker(\phi^{f-1}) \underset{\neq}{\overset{\xleftarrow{\phi}}{\subseteq}} \ker(\phi^f) = V $$

Wir konstruieren nun mit dem in Abbildung 3 illustrierten Verfahren eine Jordanbasis von ϕ:

(i) Wir betrachten den Untervektorraum $\mathrm{im}(\phi) \subseteq V$. Offensichtlich gilt $\phi(\mathrm{im}(\phi)) \subseteq \mathrm{im}(\phi)$, und $\phi|_{\mathrm{im}(\phi)} : \mathrm{im}(\phi) \to \mathrm{im}(\phi)$ ist nilpotent mit Fitting-Index $\leq f - 1$. Denn aus $v = \phi(u) \in \mathrm{im}(\phi)$ folgt $\phi^{f-1}(v) = \phi^f(u) = 0$. Also ist $\mathrm{im}(\phi) \subseteq \ker(\phi^{f-1})$. Nach Induktionsvoraussetzung hat $\phi|_{\mathrm{im}(\phi)}$ eine Jordanbasis $B = (v_1^1, .., v_{m_1}^1, ..., v_1^k, ..., v_{m_k}^k)$.

(ii) Wir verlängern nun alle Jordanketten dieser Basis, indem wir hinten einen weiteren Vektor hinzufügen. Zu jedem Vektor $0 \neq v_{m_j}^j \in \mathrm{im}(\phi)$ gibt es einen Vektor $0 \neq v_{m_j+1}^j \in V$ mit $\phi(v_{m_j+1}^j) = v_{m_j}^j$. Also ist $(v_1^j, ..., v_{m_j+1}^j)$ eine Jordankette für ϕ für alle $j \in \{1, ..., k\}$. Da B eine Basis von $\mathrm{im}(\phi)$ ist, ist das k-Tupel $(v_1^1, ..., v_1^k)$ linear unabhängig, und per Definition der Jordankette gilt außerdem $v_1^j \in \ker(\phi)$ für alle $j \in \{1, ..., k\}$.



- grüne Kästchen: die Basisvektoren aus der Jordanbasis B von $\mathrm{im}(\phi)$ aus Schritt (i).
- weiße Kästchen: die hinzugefügten Vektoren $v^j_{m_j+1}$ mit $\phi(v^j_{m_j+1}) = v^j_{m_j}$ aus Schritt (ii).
- rotes Kästchen: die hinzugefügten Vektoren in $\ker(\phi)$ aus Schritt (iii).

Der Übersichtlichkeit halber sind hier die Jordanketten nach absteigender Länge geordnet.

Abbildung 3: Konstruktion einer Jordanbasis

(iii) Wir ergänzen $(v^1_1, ..., v^k_1)$ zu einer Basis $(v^1_1, ..., v^k_1, v^{k+1}_1, ..., v^{k+r}_1)$ von $\ker(\phi)$. Dann sind (v^{k+1}_1),...(v^{k+r}_1) Jordanketten der Länge 1 für ϕ. Da $(v^1_1, ..., v^k_1, v^{k+1}_1, ..., v^{k+r}_1)$ linear unabhängig ist, ist nach Lemma 9.1.2 auch $C = (v^1_1, .., v^1_{m_1+1}, ..., v^k_1, ..., v^k_{m_k+1}, v^{k+1}_1, ..., v^{k+r}_1)$ linear unabhängig.

Da $B = (v^1_1, .., v^1_{m_1}, ..., v^k_1, ..., v^k_{m_k})$ eine Basis von $\mathrm{im}(\phi)$ ist, gilt $\mathrm{rg}(\phi) = m_1 + ... + m_k$, und da $(v^1_1, ..., v^k_1, v^{k+1}_1, ..., v^{k+r}_1)$ eine Basis von $\ker(\phi)$ ist, gilt $\mathrm{def}(\phi) = k + r$. Daraus ergibt sich $n = \dim_{\mathbb{K}} V = \mathrm{rg}(\phi) + \mathrm{def}(\phi) = m_1 + ... + m_k + k + r$. Da C aus Jordanketten besteht, linear unabhängig ist und $n = \dim_{\mathbb{K}} V$ Vektoren enthält, ist C eine Jordanbasis von V.

Das zeigt die Existenz für Endomorphismen $\phi : V \to V$ mit einem einzigen Eigenwert λ.

2. Ist ϕ trigonalisierbar mit paarweise verschiedenen Eigenwerten $\lambda_1, ..., \lambda_m$, so ist $V = H(\phi, \lambda_1) \oplus ... \oplus H(\phi, \lambda_m)$ nach Satz 7.1.8 und 7.2.1. Nach Korollar 7.1.7 gilt für alle Eigenwerte $\phi(H(\phi, \lambda_i)) \subseteq H(\phi, \lambda_i)$, und die Endomorphismen $\phi|_{H(\phi,\lambda_i)} : H(\phi, \lambda_i) \to H(\phi, \lambda_i)$ besitzen als einzigen Eigenwert λ_i. Nach 1. existieren damit Jordanbasen B_i von $H(\phi, \lambda_i)$ und wegen $V = H(\phi, \lambda_1) \oplus ... \oplus H(\phi, \lambda_m)$ ist $B = (B_1, ..., B_m)$ dann eine Jordanbasis von V. $\square$

Da eine Permutation der Jordanketten in einer Jordanbasis wieder eine Jordanbasis liefert und einer Permutation der Jordanblöcke in der darstellenden Matrix entspricht, kann die Jordan-Normalform höchstens eindeutig bis auf Permutation der Jordanblöcke sein. Um zu zeigen, dass sie tatsächlich in diesem Sinn eindeutig ist, reicht es, zu zeigen, dass die Anzahl der Jordanblöcke $J_k(\lambda)$ durch den Endomorphismus $\phi : V \to V$ eindeutig bestimmt ist.

Satz 9.1.6: (**Eindeutigkeit der Jordan-Normalform**)
Sei V ein endlich-dimensionaler $\mathbb{K}$-Vektorraum, $\phi : V \to V$ ein trigonalisierbarer Endomorphismus und B eine Jordanbasis für ϕ.

Dann ist die Anzahl $n(k, \lambda)$ der Jordanketten der Länge k in B zum Eigenwert λ

$$n(k,\lambda) = 2\mathrm{def}(\phi_\lambda^k) - \mathrm{def}(\phi_\lambda^{k+1}) - \mathrm{def}(\phi_\lambda^{k-1}) \quad \text{mit} \quad \phi_\lambda := \phi - \lambda \mathrm{id}_V$$

und die Gesamtzahl der Jordanketten zum Eigenwert λ ist $g(\phi, \lambda) = \dim_{\mathbb{K}} E(\phi, \lambda)$.

Insbesondere ist die Jordan-Normalform von ϕ eindeutig bis auf Permutation der Jordanblöcke.

Beweis:
Sei B eine Jordanbasis für ϕ. Da jede Jordankette der Länge k in B zum Eigenwert λ von ϕ einem Jordanblock $J_k(\lambda)$ in der Jordan-Normalform von ϕ entspricht, ist die Jordan-Normalform durch die Anzahl $n(k, \lambda)$ der Jordanketten der Länge $k \in \mathbb{N}$ zum Eigenwert λ eindeutig bestimmt.

Sind $(v_1^1, ..., v_{m_1}^1)$,... ,$(v_1^s, ..., v_{m_s}^s)$ die Jordanketten von ϕ zum Eigenwert λ, so bilden ihre Vektoren eine Basis des Hauptraums $H(\phi, \lambda)$, und es gilt

$$\ker(\phi_\lambda^k) = \mathrm{span}_{\mathbb{K}}\{v_1^1, ..., v_{r_1}^1, ..., v_1^s, ..., v_{r_s}^s\} \quad \text{mit} \quad r_i = \min\{k, m_i\}.$$

Denn aus $v = \Sigma_{i=1}^s \Sigma_{j=1}^{m_i} \mu_{ij} v_j^i \in H(\phi, \lambda)$ mit $\mu_{ij} \in \mathbb{K}$ folgt $\phi_\lambda^k(v) = \Sigma_{i=1}^s \Sigma_{j=k+1}^{m_i} \mu_{ij} v_j^i$, also $\phi_\lambda^k(v) = 0$ genau dann, wenn $v \in \mathrm{span}_{\mathbb{K}}\{v_1^1, ..., v_{r_1}^1, ..., v_1^s, ..., v_{r_s}^s\}$. Jede Jordankette der Länge $m < k$ zu λ trägt also m Vektoren in $\ker(\phi_\lambda^k)$ bei und jede Jordankette der Länge $m \geq k$ genau k Vektoren. Daraus ergibt sich

$$\mathrm{def}(\phi_\lambda^k) = k \cdot \Sigma_{m=k+1}^f n(m,\lambda) + \Sigma_{m=1}^k m \cdot n(m,\lambda) \quad \Rightarrow \quad g(\phi,\lambda) = \mathrm{def}(\phi_\lambda) = \Sigma_{m=1}^f n(m,\lambda) = s,$$

und es folgt

$$\begin{aligned}
&2\,\mathrm{def}(\phi_\lambda^k) - \mathrm{def}(\phi_\lambda^{k+1}) - \mathrm{def}(\phi_\lambda^{k-1}) \\
&= 2k\Sigma_{m=k+1}^f n(m,\lambda) - (k+1)\Sigma_{m=k+2}^f n(m,\lambda) - (k-1)\Sigma_{m=k}^f n(m,\lambda) \\
&+ 2\Sigma_{m=1}^k m \cdot n(m,\lambda) - \Sigma_{m=1}^{k+1} m \cdot n(m,\lambda) - \Sigma_{m=1}^{k-1} m \cdot n(m,\lambda) \\
&= k(2\Sigma_{m=k+1}^f n(m,\lambda) - \Sigma_{m=k+2}^f n(m,\lambda) - \Sigma_{m=k}^f n(m,\lambda)) - \Sigma_{m=k+2}^f n(m,\lambda) + \Sigma_{m=k}^f n(m,\lambda) \\
&+ 2kn(k,\lambda) - (k+1)n(k+1,\lambda) - kn(k,\lambda) \\
&= k(2n(k+1,\lambda) - n(k,\lambda) - n(k+1,\lambda)) + n(k,\lambda) + n(k+1,\lambda) + kn(k,\lambda) - (k+1)n(k+1,\lambda) \\
&= k(n(k+1,\lambda) - n(k,\lambda)) + (k+1)n(k,\lambda) - kn(k+1,\lambda) = n(k,\lambda).
\end{aligned}$$
□

Insbesondere folgt daraus, dass in den Jordan-Normalformen zweier ähnlicher trigonalisierbarer Matrizen $M, N \in \mathrm{Mat}(n \times n, \mathbb{K})$ jeweils die gleiche Anzahl von Jordanblöcke $J_k(\lambda)$ für alle $k \in \mathbb{N}$ und $\lambda \in \mathbb{K}$ auftreten müssen. Denn nach Korollar 7.1.10 haben zwei ähnliche trigonalisierbare Matrizen $M, N \in \mathrm{Mat}(n \times n, \mathbb{K})$ die selben Eigenwerte und für alle Eigenwerte λ und $k \in \mathbb{N}$ gilt $\mathrm{def}(M - \lambda \mathbb{1}_n)^k = \mathrm{def}(n - \lambda \mathbb{1}_n)^k$. Dies bedeutet, dass die Jordan-Normalformen von M und N bis auf Permutationen von Jordan-Blöcken übereinstimmen.

Umgekehrt sind zwei Matrizen M, N, deren Jordan-Normalformen bis auf Permutationen übereinstimmen ähnlich. Denn ihre Jordan-Normalformen sind ähnlich, da sie durch Permutationen von Jordanketten in der Jordanbasis ineinander überführt werden können, und Ähnlichkeit ist eine Äquivalenzrelation.

Korollar 9.1.7: Zwei trigonalisierbare Matrizen $M, N \in \mathrm{Mat}(n \times n, \mathbb{K})$ sind ähnlich genau dann, wenn eine der folgenden äquivalenten Bedingungen gilt:

(i) In ihren Jordan-Normalformen treten jeweils die gleiche Anzahl an Jordanblöcken $J_k(\lambda)$ für $k \in \mathbb{N}$ und $\lambda \in \mathbb{K}$ auf.

(ii) Sie besitzen die gleichen Eigenwerte, und für jeden Eigenwert λ und alle $k \in \mathbb{N}$ gilt $\mathrm{def}(M - \lambda \mathbb{1}_n)^k = \mathrm{def}(N - \lambda \mathbb{1}_n)^k$.

Ist $\mathbb{K}$ algebraisch abgeschlossen (z. B. $\mathbb{K} = \mathbb{C}$), so gilt dies für alle Matrizen $M, N \in \mathrm{Mat}(n \times n, \mathbb{K})$.

Im Fall $\mathbb{K} = \mathbb{R}$ kann man auch für nicht trigonalisierbare Matrizen eine Normalform konstruieren (siehe Aufgabe 188). Dies liegt daran, dass sich das charakteristische Polynom in lineare und quadratische Faktoren faktorisieren lässt, wobei letztere jeweils eine Nullstelle $\lambda \in \mathbb{C} \setminus \mathbb{R}$ mit der komplex konjugierten Nullstelle $\overline{\lambda} \in \mathbb{C} \setminus \mathbb{R}$ kombinieren (vgl. Lemma 8.1.7). Durch geschickte Basiswahl kann man dann erreichen, dass die beschreibende Matrix eines Endomorphismus eine Blockform annimmt, die der Jordan-Normalform ähnelt.

Für die Körper $\mathbb{Q}$ oder $\mathbb{Z}/p\mathbb{Z}$ ist dies jedoch nicht möglich, da es über diesen Körpern Polynome beliebig hohen Grades gibt, die sich nicht in Polynome kleineren Grades mit Koeffizienten in $\mathbb{Q}$ oder $\mathbb{Z}/p\mathbb{Z}$ faktorisieren lassen.

Die Ergebnisse zur Jordan-Normalform liefern auch ein Verfahren, um die Jordan-Normalform einer Matrix oder eines Endomorphismus effizient zu berechnen. Wichtig ist, dass man nur die Dinge berechnet, die wirklich erforderlich sind. So ist etwa die Konstruktion der Jordanbasen für $n \geq 5$ relativ aufwändig, aber für die Bestimmung der Jordan-Normalform reicht die Berechnung der Defekte von $(A - \lambda \mathbb{1}_n)^k$ für $1 \leq k \leq a(A, \lambda)$, wobei $a(A, \lambda)$ die algebraische Vielfachheit des Eigenwerts λ von A ist. Auch das kann oft entfallen, etwa, wenn $a(g, \lambda) = 1$ gilt.

Das allgemeine Verfahren ist das Folgende:

1. **Bestimmung des charakteristischen Polynoms p_A:**

   Man berechnet das charakteristische Polynom $p_A = \det(A - x\mathbb{1}_n)$. Zerfällt p_A nicht in Linearfaktoren, so ist A nicht trigonalisierbar und hat damit auch keine Jordan-Normalform. Das Verfahren bricht ab.

   Ansonsten hat A eine Jordan-Normalform. Die Nullstellen $\lambda \in \mathbb{K}$ von p_A sind die Eigenwerte von A, und ihre Vielfachheiten die Hauptraumdimensionen $h(A, \lambda) = a(A, \lambda)$. Hat jede Nullstelle Vielfachheit $a(A, \lambda) = 1$, so ist die Jordan-Normalform von A eine Diagonalmatrix, in deren Diagonale jeder Eigenwert λ jeweils einmal auftritt.

2. **Bestimmung der Defekte $\mathrm{def}(A_\lambda^k)$:**

   Für jeden Eigenwert $\lambda \in \mathbb{K}$ von A mit $a(A, \lambda) > 1$ berechnet man $A_\lambda := A - \lambda \mathbb{1}_n$ und dann $g(A, \lambda) = \mathrm{def}(A_\lambda)$, etwa mit dem Gauß-Verfahren. Für die Eigenwerte λ mit $a(A, \lambda) = 1$ gilt $g(A, \lambda) = a(A, \lambda) = 1$. Ist die Summe der geometrischen Vielfachheiten gleich n, so ist A diagonalisierbar, und die Jordan-Normalform von A ist eine Diagonalmatrix, in deren Diagonale der Eigenwert λ jeweils $g(A, \lambda)$ mal auftritt.

   Ist die Summe der geometrischen Vielfachheiten $< n$, so berechnet man für die Eigenwerte λ mit $a(A, \lambda) > g(A, \lambda)$ beginnend bei $k = 2$ die $k$ten Potenzen A_λ^k und bestimmt die Defekte $\mathrm{def}(A_\lambda^k)$, etwa mit dem Gauß-Verfahren. Man stoppt, wenn zwei aufeinanderfolgende Potenzen den gleichen Defekt haben, was spätestens bei $k = a(A, \lambda)$ passiert.

Der Fitting-Index von A_λ ist dann $f_\lambda = \min\{k \in \mathbb{N} | \operatorname{def}(A_\lambda^k) = \operatorname{def}(A_\lambda^{k+1})\}$. Für die Eigenwerte λ mit $a(A, \lambda) = g(A, \lambda)$ ist der Fitting-Index immer 1.

3. **Bestimmung der Anzahl der Jordanketten:**

   Für jeden Eigenwert λ von A ist die Anzahl $n(k, \lambda)$ der Jordanblöcke $J_k(\lambda)$ der Länge k gegeben durch $n(k, \lambda) = 2\operatorname{def}(A_\lambda^k) - \operatorname{def}(A_\lambda^{k+1}) - \operatorname{def}(A_\lambda^{k-1})$. Die längste Jordankette zu λ hat Länge f_λ, und die Gesamtzahl der Jordanketten zu λ ist $g(A, \lambda)$.

   Damit ist die Jordan-Normalform J bestimmt.

4. **Konstruktion der Jordanketten:**

   Ist auch eine Jordanbasis oder eine Matrix $S \in \mathrm{GL}(n, \mathbb{K})$ gesucht, so dass $A = S \cdot J \cdot S^{-1}$, muss man die Jordanketten explizit bestimmen.

   (a) Dazu bestimmt man für jeden Eigenwert λ von ϕ zunächst $k_\lambda = n(f_\lambda, \lambda)$ Vektoren $v^1_{f_\lambda}, ..., v^{k_\lambda}_{f_\lambda}$ in $\ker(A_\lambda^{f_\lambda})$ so dass $\ker(A_\lambda^{f_\lambda - 1}) \cap \operatorname{span}_{\mathbb{K}}\{v^1_{f_\lambda}, ..., v^{k_\lambda}_{f_\lambda}\} = \{0\}$ gilt. Man setzt $v^j_i := A_\lambda^{f_\lambda - i}(v^j_{f_\lambda})$ und erhält Jordanketten $(v^1_1, ..., v^1_{f_\lambda})$,..., $(v^{k_\lambda}_1, ..., v^{k_\lambda}_{f_\lambda})$ der Länge f_λ.

   (b) Treten auch Jordanketten kleinerer Länge auf, sucht man beginnend beim maximalen $k \in \{1, ..., f_\lambda - 1\}$ mit $n(k, \lambda) \neq 0$ jeweils $n(k, \lambda)$ Vektoren $w^1_k, ..., w^{n(k,\lambda)}_k$ in $\ker(A^k_\lambda)$, so dass der Schnitt von $\operatorname{span}_{\mathbb{K}}\{w^1_k, ..., w^{n(k,\lambda)}_k\}$ mit dem Spann aller vorher konstruierten Vektoren und aller Vektoren in $\ker(A_\lambda^{k-1})$ Null ist. Dann konstruiert man durch Anwendung von A_λ jeweils Vektoren $w^j_i = A_\lambda^{k-i}(w^j_k)$ und erhält so die Jordanketten $(w^1_1, ..., w^1_k)$,..., $(w^{n(k,\lambda)}_1, ..., w^{n(k,\lambda)}_k)$ der Länge k.

   (c) Wurde dies für alle k mit $n(k, \lambda) \neq 0$ und Eigenwerte λ von A durchgeführt, so bilden die zugehörigen Jordanketten eine Jordanbasis von A. Die Matrix $S \in \mathrm{GL}(n, \mathbb{K})$, die als Spaltenvektoren die Vektoren der Jordanbasis enthält , ist dann eine invertierbare Matrix mit $A = S \cdot J \cdot S^{-1}$, wobei J die Jordan-Normalform ist.

Beispiel 9.1.8: Wir zeigen, dass die Matrix

$$A = \begin{pmatrix} 4 & 1 & 1 & 1 & 1 \\ -1 & 2 & 1 & 1 & -1 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 3 \end{pmatrix} \in \mathrm{Mat}(5 \times 5, \mathbb{R})$$

trigonalisierbar ist, bestimmen ihre Jordan-Normalform J und eine Matrix $S \in \mathrm{GL}(n, \mathbb{R})$, so dass $A = S \cdot J \cdot S^{-1}$.

1. **Bestimmung charakteristischen Polynoms p_A:**
   Mit dem Laplaceschen Entwicklungssatz (Entwicklung nach der 1. Spalte) erhält man die charakteristische Gleichung

$$\det(A - \lambda \mathbb{1}_5) = \det \begin{pmatrix} 4-\lambda & 1 & 1 & 1 & 1 \\ -1 & 2-\lambda & 1 & 1 & -1 \\ 0 & 0 & 3-\lambda & 0 & 0 \\ 0 & 0 & 0 & 3-\lambda & 1 \\ 0 & 0 & 0 & 0 & 3-\lambda \end{pmatrix}$$

$$= (4-\lambda)\det\begin{pmatrix} 2-\lambda & 1 & 1 & -1 \\ 0 & 3-\lambda & 0 & 0 \\ 0 & 0 & 3-\lambda & 1 \\ 0 & 0 & 0 & 3-\lambda \end{pmatrix} + \det\begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 3-\lambda & 0 & 0 \\ 0 & 0 & 3-\lambda & 1 \\ 0 & 0 & 0 & 3-\lambda \end{pmatrix}$$
$$= (4-\lambda)(2-\lambda)(3-\lambda)^3 + (3-\lambda)^3 = (3-\lambda)^3(\lambda^2-6\lambda+9) = (3-\lambda)^5$$

Also besitzt die Matrix A nur einen einzigen Eigenwert in $\mathbb{R}$, nämlich $\lambda = 3$. Da das charakteristische Polynom p_A über $\mathbb{R}$ zerfällt, ist A trigonalisierbar.

2. **Bestimmung der Defekte** $\operatorname{def}(A_\lambda^k)$**:**
Für $A_3 := A - 3\mathbb{1}_5$ ergibt sich

$$A_3 = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ -1 & -1 & 1 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$
$$\Rightarrow \quad \operatorname{def}(A_3) = 2.$$

$$A_3^2 = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ -1 & -1 & 1 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ -1 & -1 & 1 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 2 & 2 & 1 \\ 0 & 0 & -2 & -2 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$
$$\Rightarrow \quad \operatorname{def}(A_3^2) = 3.$$

$$A_3^3 = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ -1 & -1 & 1 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 & 2 & 2 & 1 \\ 0 & 0 & -2 & -2 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 & -2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$
$$\Rightarrow \quad \operatorname{def}(A_3^3) = 4.$$

Da $\mathbb{R}^5 = H(A,3) > \operatorname{def}(A_3^3)$, muss für den Fitting-Index $f_3 \geq 4$ gelten. Daraus folgt $\operatorname{def}(A_3^4) \geq \operatorname{def}(A_3^3) + 1 = 4 + 1 = 5$ und damit $f_3 = 4$. Alternativ oder zur Kontrolle kann man dafür natürlich auch A_3^4 explizit berechnen:

$$A_3^4 = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ -1 & -1 & 1 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 & -2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$
$$\Rightarrow \quad \operatorname{def}(A_3^4) = 5.$$

3. **Bestimmung der Anzahl der Jordanketten:**
Per Definition des Fitting-Index, muss es mindestens eine Jordankette der Länge $f_3 = 4$ geben. Da die Vektoren in einer Jordanbasis linear unabhängig sind und $\dim_{\mathbb{R}} \mathbb{R}^5 = 5$

gilt, folgt sofort, dass es genau eine Jordankette der Länge 4 und noch eine weitere Jordankette der Länge 1 gibt.

Alternativ oder zur Kontrolle kann man die Anzahlen $n(k,3)$ der Jordanketten der Länge k explizit berechnen. Sie sind

$$\begin{aligned}
n(4,3) &= 2\mathrm{def} A_3^4 - \mathrm{def} A_3^5 - \mathrm{def} A_3^3 = 10 - 5 - 4 = 1\\
n(3,3) &= 2\mathrm{def} A_3^3 - \mathrm{def} A_3^4 - \mathrm{def} A_3^2 = 8 - 5 - 3 = 0\\
n(2,3) &= 2\mathrm{def} A_3^2 - \mathrm{def} A_3^3 - \mathrm{def} A_3 = 6 - 4 - 2 = 0\\
n(1,3) &= 2\mathrm{def} A_3 - \mathrm{def} A_3^2 = 4 - 3 = 1.
\end{aligned}$$

Die Jordan-Normalform von A enthält also einen Jordanblock der Länge 4 und einen Jordanblock der Länge 1 zum Eigenwert 3 und ist (bis auf Permutation der Jordanblöcke)

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

4. **Konstruktion der Jordanketten:**
Man berechnet zunächst die Kerne $\ker(A_3)$, $\ker(A_3^2)$, $\ker(A_3^3)$, $\ker(A_3^4)$:

$$\begin{aligned}
\ker(A_3) &= \mathrm{span}_{\mathbb{K}}\{e_1 - e_2, e_3 - e_4\} & \ker(A_3^2) &= \mathrm{span}_{\mathbb{K}}\{e_1, e_2, e_3 - e_4\}\\
\ker(A_3^3) &= \mathrm{span}_{\mathbb{K}}\{e_1, e_2, e_3, e_4\} & \ker(A_3^4) &= \mathbb{R}^5.
\end{aligned}$$

Zum Auffinden der längsten Jordankette wählt man einen Vektor v in $\ker(A_3^4) = \mathbb{R}^5$, der nicht in $\ker(A_3^3)$ enthalten ist, also etwa $v = e_5$, und berechnet

$$A_3 \cdot e_5 = e_1 - e_2 + e_4 \qquad A_3^2 \cdot e_5 = e_1 + e_2 \qquad A_3^3 \cdot e_5 = 2(e_1 - e_2)$$

Also ist $K_4 = (2(e_1 - e_2), e_1 + e_2, e_1 - e_2 + e_4, e_5)$ eine Jordankette der Länge 4. Zur Konstruktion der Jordankette der Länge 1 wählt man einen Vektor $w \in \ker(A_3)\setminus\mathrm{span}_{\mathbb{K}} K_4$, beispielsweise $w = e_3 - e_4$. Damit ist $(e_3 - e_4)$ eine Jordankette der Länge 1. Also ist

$$B = (2(e_1 - e_2), e_1 + e_2, e_1 - e_2 + e_4, e_5, e_3 - e_4)$$

eine Jordanbasis von A, so dass J die beschreibende Matrix bezüglich B ist. Eine Matrix $S \in \mathrm{GL}(5,\mathbb{R})$ mit $A = S \cdot J \cdot S^{-1}$ erhält man, indem man die Vektoren der Jordanbasis in dieser Reihenfolge als Spaltenvektoren einsetzt

$$S = (2(e_1 - e_2), e_1 + e_2, e_1 - e_2 + e_4, e_5, e_3 - e_4) = \begin{pmatrix} 2 & 1 & 1 & 0 & 0\\ -2 & 1 & -1 & 0 & 0\\ 0 & 0 & 0 & 0 & 1\\ 0 & 0 & 1 & 0 & -1\\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

Für Matrizen mit $n \leq 4$ sind alle möglichen Jordan-Normalformen in der Tabelle in Abbildung 4 dargestellt. In vielen Fällen ergeben sich aus Dimensionsbetrachtungen wesentliche Vereinfachungen in der Berechnung der Jordan-Normalform. Einige Regeln, die bei diesen Dimensionsbetrachtungen helfen, sind die folgenden.

	$n=2$	$n=3$	$n=4$
λ_1 $f_1=1$	$\begin{pmatrix}\lambda_1 & 0\\ 0 & \lambda_1\end{pmatrix}$	$\begin{pmatrix}\lambda_1 & 0 & 0\\ 0 & \lambda_1 & 0\\ 0 & 0 & \lambda_1\end{pmatrix}$	$\begin{pmatrix}\lambda_1 & 0 & 0 & 0\\ 0 & \lambda_1 & 0 & 0\\ 0 & 0 & \lambda_1 & 0\\ 0 & 0 & 0 & \lambda_1\end{pmatrix}$
λ_1 $f_1=2$	$\begin{pmatrix}\lambda_1 & 1\\ 0 & \lambda_1\end{pmatrix}$	$\begin{pmatrix}\lambda_1 & 1 & 0\\ 0 & \lambda_1 & 0\\ 0 & 0 & \lambda_1\end{pmatrix}$	$\begin{pmatrix}\lambda_1 & 1 & 0 & 0\\ 0 & \lambda_1 & 0 & 0\\ 0 & 0 & \lambda_1 & 0\\ 0 & 0 & 0 & \lambda_1\end{pmatrix}, \begin{pmatrix}\lambda_1 & 1 & 0 & 0\\ 0 & \lambda_1 & 0 & 0\\ 0 & 0 & \lambda_1 & 1\\ 0 & 0 & 0 & \lambda_1\end{pmatrix}$
λ_1 $f_1=3$	-	$\begin{pmatrix}\lambda_1 & 1 & 0\\ 0 & \lambda_1 & 1\\ 0 & 0 & \lambda_1\end{pmatrix}$	$\begin{pmatrix}\lambda_1 & 1 & 0 & 0\\ 0 & \lambda_1 & 1 & 0\\ 0 & 0 & \lambda_1 & 0\\ 0 & 0 & 0 & \lambda_1\end{pmatrix}$
λ_1 $f_1=4$	-	-	$\begin{pmatrix}\lambda_1 & 1 & 0 & 0\\ 0 & \lambda_1 & 1 & 0\\ 0 & 0 & \lambda_1 & 1\\ 0 & 0 & 0 & \lambda_1\end{pmatrix}$
λ_1, λ_2 $f_1=f_2=1$ $h_1=1$	$\begin{pmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{pmatrix}$	$\begin{pmatrix}\lambda_1 & 0 & 0\\ 0 & \lambda_2 & 0\\ 0 & 0 & \lambda_2\end{pmatrix}$	$\begin{pmatrix}\lambda_1 & 0 & 0 & 0\\ 0 & \lambda_2 & 0 & 0\\ 0 & 0 & \lambda_2 & 0\\ 0 & 0 & 0 & \lambda_2\end{pmatrix}$
λ_1, λ_2 $f_1=1,\ f_2=2$ $h_1=1$	-	$\begin{pmatrix}\lambda_1 & 0 & 0\\ 0 & \lambda_2 & 1\\ 0 & 0 & \lambda_2\end{pmatrix}$	$\begin{pmatrix}\lambda_1 & 0 & 0 & 0\\ 0 & \lambda_2 & 1 & 0\\ 0 & 0 & \lambda_2 & 0\\ 0 & 0 & 0 & \lambda_2\end{pmatrix}$
λ_1, λ_2 $f_1=1,\ f_2=3$ $h_1=1$	-	-	$\begin{pmatrix}\lambda_1 & 0 & 0 & 0\\ 0 & \lambda_2 & 1 & 0\\ 0 & 0 & \lambda_2 & 1\\ 0 & 0 & 0 & \lambda_2\end{pmatrix}$
λ_1, λ_2 $f_1=f_2=1$ $h_1=2$	-	$\begin{pmatrix}\lambda_1 & 0 & 0\\ 0 & \lambda_1 & 0\\ 0 & 0 & \lambda_2\end{pmatrix}$	$\begin{pmatrix}\lambda_1 & 0 & 0 & 0\\ 0 & \lambda_1 & 0 & 0\\ 0 & 0 & \lambda_2 & 0\\ 0 & 0 & 0 & \lambda_2\end{pmatrix}$
λ_1, λ_2 $f_1=2,\ f_2=1$ $h_1=2$	-	$\begin{pmatrix}\lambda_1 & 1 & 0\\ 0 & \lambda_1 & 0\\ 0 & 0 & \lambda_2\end{pmatrix}$	$\begin{pmatrix}\lambda_1 & 1 & 0 & 0\\ 0 & \lambda_1 & 0 & 0\\ 0 & 0 & \lambda_2 & 0\\ 0 & 0 & 0 & \lambda_2\end{pmatrix}$
λ_1, λ_2 $f_1=f_2=2$ $h_1=2$	-	-	$\begin{pmatrix}\lambda_1 & 1 & 0 & 0\\ 0 & \lambda_1 & 0 & 0\\ 0 & 0 & \lambda_2 & 1\\ 0 & 0 & 0 & \lambda_2\end{pmatrix}$
$\lambda_1, \lambda_2, \lambda_3$ $f_1=f_2=f_3=1$ $h_1=h_2=1$	-	$\begin{pmatrix}\lambda_1 & 0 & 0\\ 0 & \lambda_2 & 0\\ 0 & 0 & \lambda_3\end{pmatrix}$	$\begin{pmatrix}\lambda_1 & 0 & 0 & 0\\ 0 & \lambda_2 & 0 & 0\\ 0 & 0 & \lambda_3 & 0\\ 0 & 0 & 0 & \lambda_3\end{pmatrix}$
$\lambda_1, \lambda_2, \lambda_3$ $f_1=f_2=1,\ f_3=2$ $h_1=h_2=1$	-	-	$\begin{pmatrix}\lambda_1 & 0 & 0 & 0\\ 0 & \lambda_2 & 0 & 0\\ 0 & 0 & \lambda_3 & 1\\ 0 & 0 & 0 & \lambda_3\end{pmatrix}$
$\lambda_1, \lambda_2, \lambda_3, \lambda_4$ $f_i=1$ $h_i=1$	-	-	$\begin{pmatrix}\lambda_1 & 0 & 0 & 0\\ 0 & \lambda_2 & 0 & 0\\ 0 & 0 & \lambda_3 & 0\\ 0 & 0 & 0 & \lambda_4\end{pmatrix}$

Abbildung 4: Jordan-Normalformen

Für jeden Eigenwert λ...

- ... gibt es mindestens eine Jordankette oder einen Jordanblock der Länge f_λ.
- ... gilt $\dim_{\mathbb{K}} H(A,\lambda) = \Sigma_{k=1}^{f_\lambda} k \cdot n(k,\lambda)$, wobei $n(k,\lambda)$ die Anzahl der Jordanketten oder Jordanblöcke der Länge k zu λ bezeichnet.
- ... ist $g(A,\lambda) = \dim_{\mathbb{K}} E(A,\lambda) = \Sigma_{k=1}^{f_\lambda} n(k,\lambda)$ die Zahl der Jordanketten oder Jordanblöcke zu λ, denn jede Jordankette liefert einen Vektor in $E(A,\lambda)$.

Zusätzlich ist es hilfreich, im Kopf zu behalten, dass die Reihenfolge der Jordanblöcke in der Jordanbasis der Reihenfolge der Jordanblöcke in der Jordan-Normalform entspricht, und dass $\mathrm{def}(A_\lambda^k) = \dim_{\mathbb{K}} H(A,\lambda) - 1$ direkt $f_\lambda = k+1$ impliziert.

9.2 Anwendungen

Im letzten Abschnitt haben wir mit Hilfe der Jordan-Normalform das Ähnlichkeitsproblem für Matrizen über algebraisch abgeschlossenen Körpern vollständig gelöst. Die Jordan-Normalform und verwandte Konstruktionen wie die Jordan-Zerlegung haben aber noch weitere interessante Anwendungen. Insbesondere sind sie nützlich, wenn man Polynome von Matrizen berechnen muss, also Ausdrücke der Form $p(A) = a_n A^n + a_{n-1}A^{n-1} + ... + a_1 A + a_0 \mathbb{1}_n$ für $A \in \mathrm{Mat}(n \times n, \mathbb{K})$ und ein Polynom $p = \Sigma_{k\in\mathbb{N}_0} a_k x^k \in \mathbb{K}[x]$.

Solche Ausdrücke treten beispielsweise bei induktiv definierter Zahlenfolgen auf. Dort liefern sie eine geschlossene, also nicht induktive, Formel für die Folgenglieder und oft auch den Grenzwert. Die Berechnungen kann man für diagonalisierbare oder trigonalisierbare Matrizen mit Hilfe des folgenden Lemmas stark vereinfachen.

Lemma 9.2.1: Sei $\mathbb{K}$ ein Körper und $p = \Sigma_{k=0}^n a_k x^k \in \mathbb{K}[x]$ ein Polynom.

1. Ist $A \in \mathrm{Mat}(n \times n, \mathbb{K})$ diagonalisierbar mit $A = S \cdot D \cdot S^{-1}$ für eine Matrix $S \in \mathrm{GL}(n, \mathbb{K})$ und eine Diagonalmatrix $D = (d_{ij}) \in \mathrm{Mat}(n \times n, \mathbb{K})$, so ist
$$p(A) = \Sigma_{k=0}^n a_k A^k = S \cdot \begin{pmatrix} p(d_{11}) & 0 & \dots & 0 \\ 0 & p(d_{22}) & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & p(d_{nn}) \end{pmatrix} \cdot S^{-1}.$$

2. Ist $A \in \mathrm{Mat}(n \times n, \mathbb{K})$ trigonalisierbar mit $A = S \cdot D \cdot S^{-1}$ für eine Matrix $S \in \mathrm{GL}(n, \mathbb{K})$ und eine obere Dreiecksmatrix $D \in \mathrm{Mat}(n \times n, \mathbb{K})$, so ist
$$p(A) = \Sigma_{k=0}^n a_k A^k = S \cdot \begin{pmatrix} p(d_{11}) & * & * & * \\ 0 & p(d_{22}) & * & * \\ \vdots & \ddots & \ddots & * \\ 0 & \dots & 0 & p(d_{nn}) \end{pmatrix} \cdot S^{-1}.$$

Beweis:
Ist $A = S \cdot D \cdot S^{-1}$ mit $S \in \mathrm{GL}(n,\mathbb{K})$ und einer beliebigen Matrix $D \in \mathrm{Mat}(n \times n, \mathbb{K})$, so ergibt sich induktiv $A^k = (S \cdot D \cdot S^{-1}) \cdot (S \cdot D \cdot S^{-1}) \cdots (S \cdot D \cdot S^{-1}) = S \cdot D^k \cdot S^{-1}$ für alle $k \in \mathbb{N}$. Also folgt $p(A) = S \cdot p(D) \cdot S^{-1}$ für alle Polynome $p \in \mathbb{K}[x]$. Die Aussage ergibt sich dann aus

den Identitäten

$$\begin{pmatrix} d_{11} & 0 & \dots & 0 \\ 0 & d_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & d_{nn} \end{pmatrix}^k = \begin{pmatrix} d_{11}^k & 0 & \dots & 0 \\ 0 & d_{22}^k & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & d_{nn}^k \end{pmatrix}$$

$$\begin{pmatrix} d_{11} & * & * & * \\ 0 & d_{22} & * & * \\ \vdots & \ddots & \ddots & * \\ 0 & \dots & 0 & d_{nn} \end{pmatrix}^k = \begin{pmatrix} d_{11}^k & * & * & * \\ 0 & d_{22}^k & * & * \\ \vdots & \ddots & \ddots & * \\ 0 & \dots & 0 & d_{nn}^k \end{pmatrix}$$

für Diagonalmatrizen und obere Dreiecksmatrizen, die man per Induktion über $k \in \mathbb{N}$ beweist. □

Beispiel 9.2.2: Die **Pell-Folge** ist die reelle Folge $(p_n)_{n\in\mathbb{N}_0}$, die durch $p_0 = 0$, $p_1 = 1$ und die Rekursionsrelation

$$p_n = 2p_{n-1} + p_{n-2} \qquad \text{für } n \geq 2$$

definiert ist. Die Rekursionsrelation kann auch als Matrixgleichung formuliert werden

$$\begin{pmatrix} p_n \\ p_{n-1} \end{pmatrix} = M \cdot \begin{pmatrix} p_{n-1} \\ p_{n-2} \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} p_{n-1} \\ p_{n-2} \end{pmatrix}.$$

Daraus ergibt sich für das $n$te und $(n-1)$te Folgenglied

$$\begin{pmatrix} p_n \\ p_{n-1} \end{pmatrix} = M^{n-1} \cdot \begin{pmatrix} p_1 \\ p_0 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix}^{n-1} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Zur Berechnung der Folgenglieder zeigt man zunächst, dass M diagonalisierbar ist. Die charakteristische Gleichung ergibt

$$0 = \det \begin{pmatrix} 2-\lambda & 1 \\ 1 & -\lambda \end{pmatrix} = \lambda(\lambda - 2) - 1 = \lambda^2 - 2\lambda - 1.$$

Man erhält zwei verschiedene Eigenwerte $\lambda_\pm = 1 \pm \sqrt{2}$, und somit ist M diagonalisierbar. Die Eigenvektoren ergeben sich aus der Gleichung

$$\begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2x + y \\ x \end{pmatrix} = \lambda_\pm \begin{pmatrix} x \\ y \end{pmatrix},$$

die erfüllt ist genau dann, wenn $x = \lambda_\pm y$ und $2x + y = \lambda_\pm x$ oder, dazu äquivalent, $x = \lambda_\pm y$ und $(\lambda_\pm^2 - 2\lambda_\pm - 1)y = 0$. Daraus folgt

$$E(M, \lambda_\pm) = \mathbb{R} \begin{pmatrix} \lambda_\pm \\ 1 \end{pmatrix} = \mathbb{R} \begin{pmatrix} 1 \pm \sqrt{2} \\ 1 \end{pmatrix}.$$

Ist B die Standardbasis und $C = (v_+, v_-)$ eine Basis aus Eigenvektoren zu den Eigenwerten $\lambda_\pm$ von M, so ist die darstellende Matrix des Endomorphismus $\phi_M : \mathbb{R}^2 \to \mathbb{R}^2$, $v \mapsto M \cdot v$ bezüglich B gerade die Matrix $M = {}_BM_B(\phi_M)$. Die Matrix ${}_CM_C(\phi_M)$ ist diagonal, und es gilt

$$M = {}_BM_B(\phi_M) = {}_BM_C(\mathrm{id}_{\mathbb{K}^{k+1}}) \cdot {}_CM_C(\phi_M) \cdot {}_BM_C(\mathrm{id}_{\mathbb{K}^{k+1}})^{-1}.$$

Die Spaltenvektoren der Basiswechselmatrix $S := {}_B M_C(\mathrm{id}_{\mathbb{K}^{k+1}})$ sind gerade die Eigenvektoren. Daraus ergibt sich

$$S = \begin{pmatrix} 1+\sqrt{2} & 1-\sqrt{2} \\ 1 & 1 \end{pmatrix} \qquad\qquad S^{-1} = \frac{1}{2\sqrt{2}} \begin{pmatrix} 1 & \sqrt{2}-1 \\ -1 & 1+\sqrt{2} \end{pmatrix}.$$

Mit Lemma 9.2.1 erhält man für $n \geq 2$

$$\begin{aligned}
\begin{pmatrix} p_n \\ p_{n-1} \end{pmatrix} &= M^{n-1} \begin{pmatrix} p_1 \\ p_0 \end{pmatrix} = S \cdot \begin{pmatrix} 1+\sqrt{2} & 0 \\ 0 & 1-\sqrt{2} \end{pmatrix}^{n-1} \cdot S^{-1} \begin{pmatrix} p_1 \\ p_0 \end{pmatrix} \\
&= \frac{1}{2\sqrt{2}} \begin{pmatrix} 1+\sqrt{2} & 1-\sqrt{2} \\ 1 & 1 \end{pmatrix} \begin{pmatrix} (1+\sqrt{2})^{n-1} & 0 \\ 0 & (1-\sqrt{2})^{n-1} \end{pmatrix} \begin{pmatrix} 1 & \sqrt{2}-1 \\ -1 & 1+\sqrt{2} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\
&= \frac{1}{2\sqrt{2}} \begin{pmatrix} (1+\sqrt{2})^n - (1-\sqrt{2})^n \\ (1+\sqrt{2})^{n-1} - (1-\sqrt{2})^{n-1} \end{pmatrix}
\end{aligned}$$

und somit ist die $n$te Pell-Zahl für $n \geq 2$ gegeben durch die Formel:

$$p_n = \frac{(1+\sqrt{2})^n - (1-\sqrt{2})^n}{2\sqrt{2}} \qquad n \geq 2.$$

Mit Hilfe dieser Formel kann man zeigen, dass die Folge $(q_n)_{n\in\mathbb{N}_0}$ der Quotienten $q_n = p_{n+1}/p_n$ aufeinanderfolgender Pell-Zahlen gegen einen Grenzwert $q = \lim_{n\to\infty} q_n$ konvergiert. Eine Methode, den Grenzwert zu bestimmen, ergibt sich aus der Rekursionsbedingung

$$\begin{aligned}
1 &= \frac{p_{n+1}}{p_{n+1}} = 2\frac{p_n}{p_{n+1}} + \frac{p_{n-1}}{p_{n+1}} = 2\frac{p_n}{p_{n+1}} + \frac{p_{n-1}}{p_n} \cdot \frac{p_n}{p_{n+1}} = \frac{2}{q_n} + \frac{1}{q_n \cdot q_{n-1}} \\
\Rightarrow \quad 1 &= 2 \lim_{n\to\infty} \frac{2}{q_n} + \lim_{n\to\infty} \frac{1}{q_n \cdot q_{n-1}} = \frac{2}{q} + \frac{1}{q^2} \quad \Rightarrow \quad q^2 - 2q - 1 = 0.
\end{aligned}$$

Also ist der Grenzwert $q = \lim_{n\to\infty} q_n$ der Quotientenfolge eine Lösung der charakteristischen Gleichung von M und damit ein Eigenwert. Da $p_n > 0$ für alle $n \in \mathbb{N}$ folgt auch $q_n > 0$ für alle $n \in \mathbb{N}$ und somit $q = \lim_{n\to\infty} q_n \geq 0$. Also gilt $q = 1 + \sqrt{2}$.

Unter Ausnutzung der Jordan-Zerlegung und von Lemma 9.2.1 können wir nicht nur Polynome von Matrizen sondern auch Potenzreihen von trigonalisierbaren Matrizen über $\mathbb{R}$ oder $\mathbb{C}$ effizient berechnen, sofern diese konvergieren. Dies ist beispielsweise dann der Fall, wenn die relevante Potenzreihe auf ganz $\mathbb{R}$ oder $\mathbb{C}$ konvergiert.

Interessant ist hier vor allem die Exponentialreihe für Matrizen mit Einträgen aus $\mathbb{C}$. Denn nach Korollar 7.2.5 ist jede Matrix mit Einträgen aus $\mathbb{C}$ und damit auch jede reelle Matrix über $\mathbb{C}$ trigonalisierbar, und die Berechnungen vereinfachen sich dadurch deutlich. Um die Exponentialreihe zu definieren, benötigen wir zunächst Konvergenzbegriffe für Folgen oder Reihen komplexer Matrizen. Wir wiederholen dazu noch einmal die üblichen Konvergenzbegriffe in $\mathbb{C}$.

Definition 9.2.3:

1. Eine Folge $(z_n)_{n\in N_0}$ mit $z_n \in \mathbb{C}$ heißt **konvergent** mit Grenzwert $z \in \mathbb{C}$, wenn die reellen Folgen $(\mathrm{Re}(z_n))_{n\in\mathbb{N}_0}$ und $(\mathrm{Im}(z_n))_{n\in\mathbb{N}_0}$ gegen $\mathrm{Re}(z)$ und $\mathrm{Im}(z)$ konvergieren.

2. Eine Reihe $\Sigma_{k=0}^{\infty} z_k$ mit $z_k \in \mathbb{C}$ heißt **konvergent**, wenn die Folge der Partialsummen $(\Sigma_{k=0}^{n} z_k)_{n\in\mathbb{N}}$ konvergiert, und **absolut konvergent**, wenn die Reihe $\Sigma_{k=0}^{\infty}|z_n|$ konvergiert.

Wie bei reellen Reihen folgt aus absoluter Konvergenz einer komplexen Reihe ihre Konvergenz, und absolut konvergente Reihen in $\mathbb{C}$ dürfen beliebig umgeordnet werden. Auch die **Cauchysche Produktformel** behält ihre Gültigkeit. Damit ergibt sich beispielsweise

$$\begin{aligned}\exp(a+\mathrm{i}b) &= \sum_{n=0}^{\infty}\frac{(a+\mathrm{i}b)^n}{n!} = \sum_{n=0}^{\infty}\sum_{k=0}^{n}\frac{a^k(\mathrm{i}b)^{n-k}}{k!(n-k)!} = \left(\sum_{k=0}^{\infty}\frac{a^k}{k!}\right)\cdot\left(\sum_{j=0}^{\infty}\frac{(\mathrm{i}b)^j}{j!}\right) = e^a\left(\sum_{j=0}^{\infty}\frac{\mathrm{i}^j b^j}{j!}\right)\\ &= e^a\cdot e^{\mathrm{i}b} = e^a\left(\sum_{k=0}^{\infty}\frac{(-1)^k b^{2k}}{(2k)!} + \mathrm{i}\sum_{k=0}^{\infty}\frac{(-1)^k b^{2k+1}}{(2k+1)!}\right) = e^a(\cos b + \mathrm{i}\sin b)\end{aligned}$$

für alle $a, b \in \mathbb{R}$. Dies rechtfertigt nachträglich die Schreibweise aus Korollar 2.3.21. Für Folgen und Reihen von Matrizen definieren wir nun die Konvergenz komponentenweise, also indem wir die einzelnen Einträge der Matrizen betrachten. Die absolute Konvergenz einer Reihe definieren wir über das Maximum des Betrags der einzelnen Einträge. Damit lassen sich dann analog Folgen und Reihen von Matrizen definieren und auf Konvergenz untersuchen.

Definition 9.2.4: Sei $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$ und $n, m \in \mathbb{N}$.

1. Eine Folge $(M_k)_{k\in\mathbb{N}_0}$ von Matrizen $M_k \in \mathrm{Mat}(n \times m, \mathbb{K})$ heißt **konvergent** mit **Grenzwert** $M \in \mathrm{Mat}(n \times m, \mathbb{K})$, wenn für alle $i \in \{1, ..., n\}$, $j \in \{1, ..., m\}$ gilt $\lim_{k\to\infty}(M_k)_{ij} = M_{ij}$.
2. Eine Reihe $\Sigma_{r=0}^{\infty} M_r$ von Matrizen $M_r \in \mathrm{Mat}(n\times m, \mathbb{K})$ heißt **konvergent**, wenn die Folge der Partialsummen $(\Sigma_{r=0}^{k} M_r)_{k\in\mathbb{N}_0}$ konvergiert und **absolut konvergent**, wenn die reelle Reihe $\Sigma_{r=0}^{\infty} m_r$ mit $m_r = \max\{|(M_r)_{ij}|\,|\, 1 \le i \le n, 1 \le j \le m\}$ konvergiert.

Auch hier folgt aus der absoluten Konvergenz einer Reihe ihre Konvergenz, und absolut konvergente Reihen dürfen beliebig umgeordnet werden. Mit Hilfe dieser Definition können wir nun Potenzreihen für quadratische Matrizen definieren, sofern die zugehörige Potenzreihe in $\mathbb{R}$ oder $\mathbb{C}$ absolut konvergiert. Denn die absolute Konvergenz einer Potenzreihe $\Sigma_{k=0}^{\infty} a_k x^k$ auf ganz $\mathbb{C}$ impliziert die absolute Konvergenz der zugehörigen Potenzreihe in $\mathrm{Mat}(n \times n, \mathbb{C})$. Insbesondere gilt dies für die Exponentialreihe $\exp(x) = \Sigma_{k=0}^{\infty} x^k/k!$.

Um das Exponential einer Matrix zu berechnen, bietet es sich an, sie mit Hilfe der Jordan-Zerlegung in kommutierende diagonalisierbare und einen nilpotente Anteile zu zerlegen. Ist sie bereits in Jordan-Normalform, so ist ersterer einfach ihre Diagonale und letzterer die Matrix, die durch Subtrahieren der Diagonale entsteht. Das Exponential des diagonalisierbaren Anteils ergibt sich direkt aus den Exponentialen der Eigenwerte. Für den nilpotenten Anteil bricht die Exponentialreihe ab und reduziert sich auf eine endliche Summe.

Satz 9.2.5: Die **Exponentialreihe** $\exp(A) = \Sigma_{k=0}^{\infty} A^k/k!$ konvergiert absolut für alle $A \in \mathrm{Mat}(n \times n, \mathbb{C})$, und es gilt:

1. Ist $B = S \cdot A \cdot S^{-1}$ mit $S \in \mathrm{GL}(n, \mathbb{C})$, so folgt $\exp(B) = S \cdot \exp(A) \cdot S^{-1}$.
2. Ist $A \cdot B = B \cdot A$, so folgt $\exp(A + B) = \exp(A) \cdot \exp(B)$.
3. Für jede Diagonalmatrix $D = (d_{ij}) \in \mathrm{Mat}(n \times n, \mathbb{C})$ gilt $\exp(D)_{ij} = \exp(d_{ii})\delta_{ij}$.
4. Ist $N \in \mathrm{Mat}(n, \mathbb{C})$ nilpotent mit Fitting-Index f, so folgt $\exp(N) = \Sigma_{k=0}^{f-1} N^k/k!$.

Beweis:
1. Ist $B = S \cdot A \cdot S^{-1}$, so folgt mit Lemma 9.2.1

$$\exp(B) = \lim_{n\to\infty} \sum_{k=0}^{n} \frac{(S \cdot A \cdot S)^k}{k!} = S \cdot \left(\lim_{n\to\infty} \sum_{k=0}^{n} \frac{A^k}{k!} \right) \cdot S^{-1} = S \cdot \exp(A) \cdot S^{-1}.$$

2. Ist $A \cdot B = B \cdot A$, so folgt induktiv $B^l \cdot A^k = A^k \cdot B^l$ für alle $k, l \in \mathbb{N}$. Mit vollständiger Induktion zeigt man dann, dass die binomische Formel auch für Matrizen $A, B \in \mathrm{Mat}(n \times n, \mathbb{C})$ gilt, sofern diese die Bedingung $A \cdot B = B \cdot A$ erfüllen (Übung)

$$(A + B)^n = \Sigma_{k=0}^{n} \begin{pmatrix} n \\ k \end{pmatrix} A^k B^{n-k} \qquad \forall n \in \mathbb{N}_0.$$

Mit dem Cauchy-Produkt für Reihen erhält man dann

$$\begin{aligned}\exp(A+B) &= \lim_{n\to\infty} \sum_{k=0}^{n} \frac{(A+B)^k}{k!} = \lim_{n\to\infty} \sum_{k=0}^{n} \sum_{j=0}^{k} \begin{pmatrix} k \\ j \end{pmatrix} \frac{A^j B^{k-j}}{k!} = \lim_{n\to\infty} \sum_{j=0}^{n} \sum_{l=0}^{n-j} \frac{A^j B^l}{j! l!} \\ &= \left(\lim_{n\to\infty} \sum_{j=0}^{n} \frac{A^j}{j!} \right) \cdot \left(\lim_{m\to\infty} \sum_{l=0}^{m} \frac{B^l}{l!} \right) = \exp(A) \cdot \exp(B).\end{aligned}$$

3. Ist D diagonal, so folgt mit Lemma 9.2.1

$$\exp(D)_{ij} = \lim_{n\to\infty} \left(\sum_{k=0}^{n} \frac{D^k}{k!} \right)_{ij} = \delta_{ij} \lim_{n\to\infty} \sum_{k=0}^{n} \frac{d_{ii}^k}{k!} = \delta_{ij} \exp(d_{ii}).$$

4. Ist N nilpotent mit Fitting-Index f, so ist $N^k = 0$ für alle $k \geq f$ und damit folgt

$$\exp(N) = \lim_{n\to\infty} \sum_{k=0}^{n} \frac{N^k}{k!} = \lim_{n\to\infty} \sum_{k=0}^{f-1} \frac{N^k}{k!} = \sum_{k=0}^{f-1} \frac{N^k}{k!}.$$

□

Beispiel 9.2.6: Wir betrachten eine Matrix in Form eines Jordanblocks:

$$A = \begin{pmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & \ddots & 0 \\ \vdots & \ddots & \ddots & 1 \\ 0 & \ldots & 0 & \lambda \end{pmatrix} = \underbrace{\begin{pmatrix} \lambda & 0 & \ldots & 0 \\ 0 & \lambda & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & \lambda \end{pmatrix}}_{=:D} + \underbrace{\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ \vdots & \ddots & \ddots & 1 \\ 0 & \ldots & 0 & 0 \end{pmatrix}}_{=:N} \in \mathrm{Mat}(n \times n, \mathbb{C})$$

Da A in Jordan-Normalform ist, ist der Diagonalanteil D der Jordan-Zerlegung die Diagonale und der nilpotente Anteil N der Anteil über der Diagonalen. Man zeigt auch leicht durch Nachrechnen, dass $D \cdot N = N \cdot D$ und N nilpotent ist mit Fitting-Index n. Daraus folgt mit Satz 9.2.5 für alle $t \in \mathbb{R}$

$$\exp(tA) = \begin{pmatrix} e^{t\lambda} & 0 & 0 & \ldots & 0 \\ 0 & e^{t\lambda} & 0 & \ddots & \vdots \\ 0 & 0 & e^{t\lambda} & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & 0 & e^{t\lambda} \end{pmatrix} \cdot \begin{pmatrix} 1 & t & \frac{t^2}{2} & \ldots & \frac{t^{n-1}}{(n-1)!} \\ 0 & 1 & \ddots & \ddots & \vdots \\ 0 & 0 & \ddots & t & \frac{t^2}{2} \\ \vdots & \ddots & \ddots & 1 & t \\ 0 & \ldots & 0 & 0 & 1 \end{pmatrix} = e^{\lambda t} \begin{pmatrix} 1 & t & \frac{t^2}{2} & \ldots & \frac{t^{n-1}}{(n-1)!} \\ 0 & 1 & \ddots & \ddots & \vdots \\ 0 & 0 & \ddots & t & \frac{t^2}{2} \\ \vdots & \ddots & \ddots & 1 & t \\ 0 & \ldots & 0 & 0 & 1 \end{pmatrix}$$

Eine wichtige Anwendung des Matrixexponentials ergibt sich im Kontext von Differentialgleichungssystemen mit konstanten Koeffizienten. Hier lassen sich mit Mitteln der linearen Algebra Lösungen konstruieren, indem man das Differentialgleichungssystem durch eine Matrix beschreibt und diese dann exponenziert.

Beispiel 9.2.7: Sei $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$.

1. Ein System gewöhnlicher **Differentialgleichungen erster Ordnung mit konstanten Koeffizienten** und **Anfangswerten** $v_1, ..., v_n \in \mathbb{K}$ ist ein Gleichungssystem der Form

$$\begin{aligned} f_1' &= a_{11}f_1 + a_{12}f_2 + \ldots + a_{1n}f_n \quad (15) \\ f_2' &= a_{21}f_1 + a_{22}f_2 + \ldots + a_{2n}f_n \\ &\vdots \\ f_n' &= a_{n1}f_1 + a_{n2}f_2 + \ldots + a_{nn}f_n \end{aligned} \qquad \begin{aligned} f_1(0) &= v_1 \quad (16) \\ f_2(0) &= v_2 \\ &\vdots \\ f_n(0) &= v_n, \end{aligned}$$

   wobei f_i' die Ableitung einer differenzierbaren Funktion $f_i : \mathbb{R} \to \mathbb{K}$ bezeichnet und $a_{ij} \in \mathbb{K}$.

2. Eine **Lösung** des Differentialgleichungssystems ist ein n-Tupel differenzierbarer Funktionen $f_i : \mathbb{R} \to \mathbb{K}$, das die Gleichungen (15) erfüllt.

3. Eine **Lösung zum Anfangswert** $v \in \mathbb{K}^n$ ist eine Lösung von (15), die auch (16) erfüllt.

Das Differentialgleichungssystem und die Anfangsbedingungen lassen sich auch als vektorwertige Gleichungen schreiben:

$$\begin{pmatrix} f_1' \\ \vdots \\ f_n' \end{pmatrix} = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \ldots & a_{nn} \end{pmatrix} \cdot \begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix} \qquad (17) \qquad \begin{pmatrix} f_1(0) \\ \vdots \\ f_n(0) \end{pmatrix} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \qquad (18)$$

mit einer zu bestimmenden differenzierbaren Funktion $f = (f_1, .., f_n)^T : \mathbb{R} \to \mathbb{K}^n$. Man kann zeigen, dass das Differentialgleichungssystem

$$f' = A \cdot f \qquad (19) \qquad\qquad f(0) = v \qquad (20)$$

zu jedem $A \in \mathrm{Mat}(n \times n, \mathbb{K})$ und jedem $v \in \mathbb{K}^n$ genau eine Lösung $f = (f_1, .., f_n)^T : \mathbb{R} \to \mathbb{K}^n$ besitzt, nämlich

$$\begin{pmatrix} f_1(t) \\ \vdots \\ f_n(t) \end{pmatrix} := \exp(tA) \cdot \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}.$$

Denn für jede Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{K})$ und jedes $t \in \mathbb{R}$ gilt

$$\frac{d}{dt} \exp(tA) = \frac{d}{dt} \sum_{k=0}^{\infty} \frac{t^k A^k}{k!} = \sum_{k=1}^{\infty} \frac{t^{k-1} A^k}{(k-1)!} = A \cdot \sum_{k=0}^{\infty} \frac{t^k A^k}{k!} = A \cdot \exp(tA) = \exp(tA) \cdot A,$$

und daraus folgt für alle $t \in \mathbb{R}$

$$\begin{pmatrix} f_1'(t) \\ \vdots \\ f_n'(t) \end{pmatrix} = \frac{d}{dt} \exp(tA) \cdot \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = A \cdot \exp(tA) \cdot \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = A \cdot \begin{pmatrix} f_1(t) \\ \vdots \\ f_n(t) \end{pmatrix}$$
$$\begin{pmatrix} f_1(0) \\ \vdots \\ f_n(0) \end{pmatrix} = \exp(0) \cdot \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}.$$

Besonders einfach wird die Lösung des Differentialgleichungssystems, wenn die Matrix A eine Diagonalmatrix ist. In diesem Fall erhält man n separate Differentialgleichungen, die jeweils die Lösung $f_i(t) = e^{a_{ii}t} v_i$ besitzen. Man spricht von einem **entkoppelten System**.

Ist A diagonalisierbar, so kann man das Differentialgleichungssystem durch Konjugation mit einer invertierbaren Matrix S in ein solches entkoppeltes System überführen. Die Lösungen sind dann wieder bestimmte Linearkombinationen von Exponentialfunktionen. Ist A trigonalisierbar, aber nicht diagonalisierbar, so haben die Lösungen eine kompliziertere Gestalt.

Beispiel 9.2.8: Das Differentialgleichungssystem mit Anfangswerten

$$\begin{aligned} f_1' &= \lambda f_1 + f_2 \qquad & f_1(0) &= v_1 \\ f_2' &= \lambda f_2 + f_3 & f_2(0) &= v_2 \\ &\vdots & &\vdots \\ f_{n-1}' &= \lambda f_{n-1} + f_{n-2} & f_{n-1}(0) &= v_{n-1} \\ f_n' &= \lambda f_n & f_n(0) &= v_n \end{aligned}$$

mit $\lambda, v_1, ..., v_n \in \mathbb{C}$ wird durch die Matrixgleichung

$$\begin{pmatrix} f_1' \\ f_2' \\ \vdots \\ f_{n-1}' \\ f_n' \end{pmatrix} = \begin{pmatrix} \lambda & 1 & 0 & \ldots & 0 \\ 0 & \lambda & 1 & \ddots & \vdots \\ 0 & 0 & \ddots & \ddots & 0 \\ \vdots & \vdots & \ddots & \lambda & 1 \\ 0 & 0 & \ldots & 0 & \lambda \end{pmatrix} \cdot \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_{n-1} \\ f_n \end{pmatrix} \qquad \begin{pmatrix} f_1(0) \\ f_2(0) \\ \vdots \\ f_{n-1}(0) \\ f_n(0) \end{pmatrix} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_{n-1} \\ v_n \end{pmatrix}$$

beschrieben und hat damit nach Beispiel 9.2.7 und Beispiel 9.2.6 die eindeutige Lösung

$$\begin{pmatrix} f_1(t) \\ f_2(t) \\ \vdots \\ f_{n-1}(t) \\ f_n(t) \end{pmatrix} = e^{\lambda t} \begin{pmatrix} 1 & t & \frac{t^2}{2} & \cdots & \frac{t^{n-1}}{(n-1)!} \\ 0 & 1 & \ddots & \ddots & \vdots \\ 0 & 0 & \ddots & t & \frac{t^2}{2} \\ \vdots & \ddots & \ddots & 1 & t \\ 0 & \ldots & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_{n-1} \\ v_n \end{pmatrix} = \begin{pmatrix} e^{\lambda t}(v_1 + tv_2 + \frac{t^2}{2} v_3 + ... + \frac{t^{n-1}}{(n-1)!} v_n) \\ e^{\lambda t}(v_2 + tv_3 + \frac{t^2}{2} v_4 + ... + \frac{t^{n-2}}{(n-2)!} v_n) \\ \vdots \\ e^{\lambda t}(v_{n-1} + tv_n) \\ e^{\lambda t} v_n \end{pmatrix}.$$

Beispiel 9.2.8 liefert die Lösung eines Anfangswertproblems für ein Differentialgleichungssystem, das durch einen Jordanblock beschrieben wird. Über dem Körper $\mathbb{C}$ liefert dies bereits die vollständige Lösung eines allgemeinen Differentialgleichungssystems mit Anfangswerten.

Denn zu der beschreibenden Matrix $A \in \text{Mat}(n \times n, \mathbb{C})$ eines Differentialgleichungssystems $f' = A \cdot f$ mit Anfangswert $f(0) = v$ gibt es eine Matrix $J \in \text{Mat}(n \times n, \mathbb{C})$ in Jordan-Normalform und eine Matrix $S \in \text{GL}(n, \mathbb{C})$ mit $A = S \cdot J \cdot S^{-1}$.

Indem man statt der Abbildung $f : \mathbb{R} \to \mathbb{C}^n$ die Abbildung $g = S^{-1} \cdot f$ betrachtet, erhält man dann das Differentialgleichungssystem $g' = J \cdot g$ mit Anfangswert $g(0) = S^{-1} \cdot v$. Dieses kann man für jeden Jordanblock mit Beispiel 9.2.8 separat lösen, da die Differentialgleichungen für die verschiedenen Jordanblöcke entkoppeln. Anschließend erhält man durch Multiplikation mit S^{-1} eine Lösung $f = S^{-1} \cdot g$ des ursprünglichen Anfangswertproblems.

An was man sich erinnern sollte:

Die wichtigsten Begriffe und Definitionen:

- Jordankette, Jordanbasis,
- Jordanblock, Jordan-Normalform
- rekursive Folge, Differentialgleichungssystem mit konstanten Koeffizienten.

Die wichtigsten Aussagen:

- Jordanketten sind linear unabhängig,
- Existenz und Eindeutigkeit der Jordan-Normalform,
- Jordan-Normalform und Ähnlichkeit von Matrizen,
- Anzahl der Jordanketten der Länge n,
- Berechnung der Jordan-Normalform.
- Berechnung von Polynomen und Exponentialen von Matrizen mit der Jordan-Zerlegung,
- Anwendung auf rekursive Folgen,
- Lösen von Differentialgleichungssystemen mit konstanten Koeffizienten.

Die wichtigsten Beispiele:

- **Jordan-Normalform:** Beispiele für $\leq 4 \times 4$-Matrizen
- rekursive Folgen, Differenzialgleichungssysteme.

10 Bilinearformen und Sesquilinearformen

10.1 Bilinearformen und Sesquilinearformen

Während wir bisher die Eigenschaften von Vektorräumen und linearen Abbildungen untersucht haben, betrachten wir jetzt Vektorräume, die mit zusätzlicher Struktur ausgestattet sind, nämlich mit *Bilinearformen* oder *Sesquilinearformen*, und die zugehörigen linearen Abbildungen, die mit dieser Struktur verträglich sind. Bilinearformen wurden bereits in Kapitel 5.2 eingeführt. Sesquilinearformen unterscheiden sich von Bilinearformen nur dadurch, dass sie in einem der beiden Argumente nicht ganz linear sind, sondern nur bis auf komplexe Konjugation.

Bi- und Sesquilinearformen werden einerseits dazu benötigt, reelle und komplexe Vektorräume mit einem Begriff von Längen und Winkeln auszustatten, also um geometrische Größen auf Vektorräumen zu beschreiben. Andererseits können wir damit später auch geometrische Kurven wie Kreise, Ellipsen und Hyperbeln und bestimmte Flächen im dreidimensionalen Raum wie Ellipsoide und Hyperboloide erfassen.

Neben diesen geometrischen Anwendungen spielen Bilinearformen und Sesquilinearformen auch in der Analysis und Funktionalanalysis eine wichtige Rolle. Darüber hinaus besitzen sie viele Anwendungen in der Physik, etwa in der klassischen Mechanik (symplektische Formen), der speziellen Relativitätstheorie (Minkowski-Metrik) und in der Quantenmechanik (Hilbertraum).

Definition 10.1.1:

1. Sei $\mathbb{K}$ ein Körper und V ein Vektorraum über $\mathbb{K}$. Eine **Bilinearform** β auf V ist eine Abbildung $\beta : V \times V \to \mathbb{K}$, so dass für alle $v, v', w, w' \in V$ und $\alpha \in \mathbb{K}$ gilt:
$$\begin{aligned}\beta(v+v', w) &= \beta(v,w) + \beta(v', w)\\ \beta(v, w+w') &= \beta(v,w) + \beta(v,w')\\ \beta(\alpha v, w) &= \beta(v, \alpha w) = \alpha\beta(v,w)\end{aligned}$$

2. Sei $\mathbb{K} = \mathbb{C}$ oder $\mathbb{K} = \mathbb{R}$ und V ein Vektorraum über $\mathbb{K}$. Eine **Sesquilinearform** β auf V ist eine Abbildung $\beta : V \times V \to \mathbb{K}$, so dass für alle $v, v', w, w' \in V$ und $\alpha \in \mathbb{K}$ gilt
$$\begin{aligned}\beta(v+v', w) &= \beta(v,w) + \beta(v', w)\\ \beta(v, w+w') &= \beta(v,w) + \beta(v,w')\\ \beta(\overline{\alpha} v, w) &= \beta(v, \alpha w) = \alpha\beta(v,w),\end{aligned}$$
wobei $\bar{}: \mathbb{C} \to \mathbb{C}$, $x + \mathrm{i}y \mapsto x - \mathrm{i}y$ die komplexe Konjugation bezeichnet.

Bilinearformen und Sesquilinearformen unterscheiden sich also einerseits durch ihre Allgemeinheit - erstere sind über allen Körpern definiert, letztere nur über $\mathbb{C}$ und $\mathbb{R}$ - andererseits dadurch, dass beim Herausziehen eines Skalaren aus dem ersten Argument einer Sesquilinearform der Skalar komplex konjugiert wird[16]. Eine Bilinearform auf einen *reellen* Vektorraum V werden wir im Folgenden immer auch als eine Sesquilinearform auf V interpretieren, wobei die komplexe Konjugation auf $\mathbb{R}$ die Identitätsabbildung ist. Dies liefert uns eine kompaktere Notation

[16] Achtung: hier gibt es in der Literatur unterschiedliche Konventionen. Viele Lehrbücher bzw. Skripten vertauschen die Rolle der beiden Argumente gegenüber der hier gewählten Konvention und fordern Linearität im ersten Argument statt im zweiten.

mit weniger Fallunterscheidungen. Eine *Bilinearform* auf einem *komplexen* Vektorraum W ist dagegen etwas völlig anderes als eine *Sesquilinearform* auf W.

Besonders wichtige Beispiele von Bi-oder Sesquilinearformen sind Bi-oder Sesquilinearformen, die eine oder mehrere der folgenden Eigenschaften besitzen. Diese sind viel einfacher zu handhaben als der allgemeine Fall.

Definition 10.1.2: Sei V ein Vektorraum über $\mathbb{K}$.

1. Eine Sesquilinearform β auf V heißt:
   - **hermitesch**, wenn $\beta(w,v) = \overline{\beta(v,w)}$ für alle $v, w \in V$,
   - **antihermitesch**, wenn $\beta(w,v) = -\overline{\beta(v,w)}$ für alle $v, w \in V$.
2. Eine Bilinearform β auf V heißt:
   - **symmetrisch**, wenn $\beta(w,v) = \beta(v,w)$ für alle $v, w \in V$,
   - **antisymmetrisch**, wenn $\beta(w,v) = -\beta(v,w)$ für alle $v, w \in V$,
   - **alternierend**, wenn $\beta(v,v) = 0$ für alle $v \in V$.
3. Eine Bilinearform oder Sesquilinearform β auf V heißt:
   - **nicht ausgeartet**, wenn $\beta(v,w) = 0 \,\forall\, v \in V \Rightarrow w = 0$.

Wie bereits in Kapitel 5 diskutiert, stimmen der Begriff der antisymmetrischen und alternierenden Bilinearform für Körper mit $\mathrm{char}(\mathbb{K}) \neq 2$ überein, unterscheiden sich aber für $\mathrm{char}(\mathbb{K}) = 2$.

Ebenso geht aus der Definition hervor, dass eine Sesquilinearform auf einem reellen Vektorraum genau dann (anti)hermitesch ist, wenn sie (anti)symmetrisch ist. Für komplexe Vektorräume unterscheiden sich (anti)hermitesche Sesquilinearformen und (anti)symmetrische Bilinearformen dagegen erheblich. *Antisymmetrische Sesquilinearformen* auf komplexen Vektorräumen sind irrelevant, weil das einzige Beispiel die Nullabbildung ist (Beweis: Übung).

Beispiel 10.1.3:

1. Für jede Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{K})$ ist $\beta_A : \mathbb{K}^n \times \mathbb{K}^n \to \mathbb{K}$ mit

   $$\beta_A(v,w) = v^T \cdot A \cdot w = \Sigma_{i,j=1}^n v_i w_j A_{ij}$$

   eine Bilinearform auf dem $\mathbb{K}^n$. Sie ist symmetrisch genau dann, wenn A symmetrisch ist ($A^T = A$), antisymmetrisch genau dann, wenn A antisymmetrisch ist ($A^T = -A$) und nicht ausgeartet genau dann, wenn $\mathrm{rg}(A) = n$. (Übung)

2. Für jede Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{C})$ ist $\beta'_A : \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$ mit

   $$\beta'_A(v,w) = v^\dagger \cdot A \cdot w := \overline{v}^T \cdot A \cdot w = \Sigma_{i,j=1}^n \overline{v_i} w_j A_{ij}$$

   eine Sesquilinearform auf dem $\mathbb{C}^n$. Sie ist hermitesch genau dann, wenn $A^\dagger := \overline{A}^T = A$, antihermitesch genau dann, wenn $A^\dagger = -A$ und nicht ausgeartet genau dann, wenn $\mathrm{rg}(A) = n$. (Übung)

3. Für jede Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{K})$ definiert

   $$\beta_A(M,N) := \mathrm{tr}\,(M^T \cdot A \cdot N) = \Sigma_{i,j,k=1}^n A_{kj} M_{ki} N_{ji}$$

eine Bilinearform auf $\mathrm{Mat}(n \times n, \mathbb{K})$, und für jede Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{C})$ definiert

$$\beta'_A(M, N) := \mathrm{tr}\,(M^\dagger \cdot A \cdot N) = \Sigma_{i,j,k=1}^n A_{kj} \overline{M_{ki}} N_{ji}$$

eine Sesquilinearform auf $\mathrm{Mat}(n \times n, \mathbb{C})$. β_A ist (anti)symmetrisch genau dann, wenn $A^T = A$ ($A^T = -A$) und β'_A ist (anti)hermitesch genau dann, wenn $A^\dagger = A$ ($A^\dagger = -A$).

4. Für jede stetige Abbildung $f : [0,1] \to \mathbb{R}$ definiert

$$\beta_f(g, h) := \int_0^1 f(x)g(x)h(x)\,dx$$

eine symmetrische Bilinearform auf dem reellen Vektorraum $C([0,1], \mathbb{R})$ der stetigen Abbildungen von $[0,1]$ nach $\mathbb{R}$ und

$$\beta'_f(g, h) := \int_0^1 f(x)\overline{g(x)}h(x)\,dx$$

eine hermitesche Sesquilinearform auf dem komplexen Vektorraum $C([0,1], \mathbb{C})$ der stetigen Abbildungen von $\mathbb{R}$ nach $\mathbb{C}$.

5. Für jede Bi- oder Sesquilinearform β auf einem $\mathbb{K}$-Vektorraum V und jeden Untervektorraum $U \subseteq V$ ist die Einschränkung $\beta|_U : U \times U \to \mathbb{K}$, $(u, v) \mapsto \beta(u, v)$ eine Bi- oder Sesquilinearform auf U. Diese kann ausgeartet sein, auch wenn β nicht ausgeartet ist.

6. Die symmetrische Bilinearform $\beta : \mathbb{K}^2 \times \mathbb{K}^2 \to \mathbb{K}$ mit $\beta(v, w) = v_1 w_1 - v_2 w_2$ ist nicht ausgeartet, denn bereits aus $\beta(e_1, v) = \beta(e_2, v) = 0$ folgt $v = 0$. Ihre Einschränkungen auf die Untervektorräume $U_\pm = \mathbb{K}(e_1 \pm e_2)$ sind aber ausgeartet, denn es gilt

$$\beta(\lambda(e_1 \pm e_2), \mu(e_1 \pm e_2)) = \lambda \cdot \mu - (\pm\lambda) \cdot (\pm\mu) = 0 \;\; \forall \lambda, \mu \in \mathbb{K} \quad \Rightarrow \beta|_{U_\pm} = 0.$$

7. Die Einschränkung einer *alternierenden* Bilinearform β auf einen eindimensionalen Untervektorraum $U \subseteq V$ ist immer ausgeartet. Denn es gilt $U = \mathbb{K}u$ für ein $u \in U$ und $\beta(u, u) = 0$.

Besonders wichtige Beispiele von Sesquilinearformen sind *Skalarprodukte* auf reellen oder komplexen Vektorräumen. Für reelle Vektorräume liefern diese Längen- und Winkelbegriffe. Für komplexe Vektorräume haben sie keine unmittelbare geometrische Interpretation, besitzen aber viele Anwendungen, beispielsweise in der Quantenmechanik. Die Eigenschaften des Skalarprodukts sind dort essentiell für eine sinnvolle Beschreibung von Messwahrscheinlichkeiten.

Definition 10.1.4: Sei $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$, V ein $\mathbb{K}$-Vektorraum und $\beta : V \times V \to \mathbb{K}$ eine Sesquilinearform auf V. Dann heißt β:
- **positiv semidefinit**, wenn $\beta(v, v) \geq 0$ für alle $v \in V$,
- **negativ semidefinit**, wenn $\beta(v, v) \leq 0$ für alle $v \in V$,
- **positiv definit**, wenn $\beta(v, v) > 0$ für alle $v \in V \setminus \{0\}$,
- **negativ definit**, wenn $\beta(v, v) < 0$ für alle $v \in V \setminus \{0\}$,
- **Skalarprodukt**, wenn β positiv definit und hermitesch ist.

Einen $\mathbb{K}$-Vektorraum V zusammen mit einem Skalarprodukt bezeichnet man als einen **euklidischen Vektorraum** für $\mathbb{K} = \mathbb{R}$ oder als einen **unitären Vektorraum** für $\mathbb{K} = \mathbb{C}$. Das Skalarprodukt wird oft mit $\langle\,,\,\rangle$ statt mit β bezeichnet.

Bemerkung 10.1.5:

1. Der Begriff des Skalarprodukts benötigt die *Anordenbarkeit* des Körpers $\mathbb{R}$. Man kann zeigen, dass ein Körper der Charakteristik $\mathrm{char}(\mathbb{K}) \neq 0$ nie anordenbar sein kann. Damit gibt es keine Skalarprodukte auf Vektorräumen über endlichen Körpern.
2. Skalarprodukte sind nicht ausgeartet, denn für alle Vektoren $w \in V \setminus \{0\}$ gilt $\langle w, w\rangle > 0$. Also folgt aus $\langle v, w\rangle = 0$ für alle $v \in V$ immer $w = 0$.
3. Ist $\langle\, ,\, \rangle$ ein Skalarprodukt auf V, so ist für jeden Untervektorraum $U \subseteq V$ die Einschränkung $\langle\, ,\, \rangle|_U$ ein Skalarprodukt auf U. Insbesondere ist $\langle\, ,\rangle|_U$ also nicht ausgeartet für alle Untervektorräume $U \subseteq V$.

Beispiel 10.1.6:

1. Das **Standardskalarprodukt** auf dem reellen Vektorraum $\mathbb{R}^n$ ist die Bilinearform aus Beispiel 10.1.3, 1. für $A = \mathbb{1}_n$
$$\langle x, y\rangle = x^T \cdot y = \Sigma_{i=1}^n x_i\, y_i.$$
2. Das **Standardskalarprodukt** auf dem komplexen Vektorraum $\mathbb{C}^n$ ist die Sesquilinearform aus Beispiel 10.1.3, 2. für $A = \mathbb{1}_n$
$$\langle x, y\rangle = \overline{x}^T \cdot y = \Sigma_{i=1}^n \overline{x_i}\, y_i.$$
3. Ein Skalarprodukt auf dem reellen Vektorraum $\mathrm{Mat}(n \times n, \mathbb{R})$ ist gegeben durch
$$\langle M, N\rangle_{\mathbb{R}} = \mathrm{tr}\,(M^T \cdot N) = \Sigma_{i,j=1}^n M_{ij} N_{ij}$$
und ein Skalarprodukt auf dem komplexen Vektorraum $\mathrm{Mat}(n \times n, \mathbb{C})$ durch
$$\langle M, N\rangle_{\mathbb{C}} = \mathrm{tr}\,(\overline{M}^T \cdot N) = \Sigma_{i,j=1}^n \overline{M_{ij}} N_{ij}.$$
Dies sind die Bilinearform und der Sesquilinearform aus Beispiel 10.1.3, 3. für $A = \mathbb{1}_n$.
4. Die Bilinearform aus Beispiel 10.1.3, 4. für $f = 1$
$$\beta(g, h) = \int_0^1 g(x)h(x)\, dx$$
ist ein Skalarprodukt auf $C([0, 1], \mathbb{R})$. Denn für jede stetige Abbildung $g : [0, 1] \to \mathbb{R}$ mit $g \neq 0$ gilt $\beta(g, g) = \int_0^1 g(x)^2\, dx > 0$. Analog zeigt man, dass die Sesquilinearform
$$\beta(g, h) = \int_0^1 \overline{g(x)}h(x)\, dx$$
aus Beispiel 10.1.3, 4. ein Skalarprodukt auf $C([0, 1], \mathbb{C})$ ist.
5. Die hermitesche Sesquilinearform $\beta : C(\mathbb{R}, \mathbb{C}) \times C(\mathbb{R}, \mathbb{C}) \to \mathbb{C}$ mit
$$\beta(g, h) = \int_0^1 \overline{g(x)}h(x)\, dx$$
ist kein Skalarprodukt auf $C(\mathbb{R}, \mathbb{C})$. Sie ist zwar positiv semidefinit, aber nicht positiv definit. Denn für die stetige Abbildung $g : \mathbb{R} \to \mathbb{C}$ mit $g(x) = 0$ für $x \leq 2$ und $g(x) = x-2$ für $x \geq 2$ gilt $\beta(g, g) = \int_0^1 |g(x)|^2\, dx = \int_0^1 0\, dx = 0$.

6. Die **Minkowski-Metrik** auf dem $\mathbb{R}^n$ für $n \geq 2$

$$\eta(x, y) = -x_1 y_1 + \Sigma_{i=2}^n x_i\, y_i$$

ist eine symmetrische Bilinearform auf dem $\mathbb{R}^n$. Sie ist nicht ausgeartet, aber nicht positiv semidefinit, denn es gilt $\eta(e_1, e_1) = -1 < 0$.

7. Eine Bilinearform $\beta \neq 0$ auf einem komplexen Vektorraum V ist nie positiv semidefinit. Denn ist $\beta(v, v) > 0$ für einen Vektor $v \in V$, so folgt $\beta(\mathrm{i}v, \mathrm{i}v) = \mathrm{i}^2\beta(v, v) = -\beta(v, v) < 0$.

Wie schon bei allen vorher eingeführten mathematischen Strukturen betrachten wir nun die zugehörigen strukturerhaltenden Abbildungen. Da es hier um Vektorräume geht, die mit zusätzlicher Struktur ausgestattet sind, müssen sie verträglich mit der Vektorraumstruktur sein, also lineare Abbildungen. Zusätzlich benötigen wir eine Verträglichkeitsbedingung mit den Bi- oder Sesquilinearformen. Sind V und W Vektorräume über $\mathbb{K}$, $\phi : V \to W$ eine $\mathbb{K}$-lineare Abbildung und β und γ Bi-oder Sesquilinearformen auf V und W, so sollte es keinen Unterschied machen, ob wir zwei Vektoren $v, v' \in V$ in β einsetzen oder sie zunächst mit ϕ in den Vektorraum W transportieren und ihre Bilder dann in γ einsetzen.

Definition 10.1.7: Seien V, W Vektorräume über $\mathbb{K}$ und β und γ Bi- oder Sesquilinearformen auf V und W. Eine $\mathbb{K}$-lineare Abbildung $\phi : V \to W$ heißt **verträglich** mit β und γ, wenn

$$\gamma(\phi(v), \phi(v')) = \beta(v, v') \qquad \forall v, v' \in V.$$

Ist $V = W$ und $\beta = \gamma$, so heißt ein mit β verträglicher Endomorphismus von V auch β-**unitär**.

Sind β und γ Skalarprodukte, so nennt man eine mit β und γ verträgliche $\mathbb{K}$-lineare Abbildung auch einen **Homomorphismus von unitären oder euklidischen Vektorräumen**.

Bemerkung 10.1.8:

1. Ist β nicht ausgeartet, so sind mit β verträgliche lineare Abbildungen $\phi : V \to W$ injektiv.

   Denn ist ϕ verträglich mit β, γ und ist $v \in \ker(\phi)$, so folgt $\beta(v', v) = \gamma(\phi(v'), \phi(v)) = \gamma(\phi(v'), 0) = 0$ für alle $v' \in V$, also $v = 0$.

2. Ist V endlich-dimensional und β nicht ausgeartet, so ist jeder β-unitäre Endomorphismus $\phi : V \to V$ bijektiv. Denn nach 1. ist ϕ injektiv und wegen $\dim_{\mathbb{K}} V < \infty$ auch bijektiv.

3. Aus 2. folgt: Jeder Homomorphismus von unitären oder euklidischen Vektorräumen ist injektiv. Jeder Endomorphismus eines endlich-dimensionalen unitären oder euklidischen Vektorraums, der mit dem Skalarprodukt verträglich ist, ist bijektiv.

Insbesondere können wir für einen gegebenen Vektorraum V, der mit einer Bi- oder Sesquilinearform β ausgestattet ist, die Automorphismen von V betrachten, die verträglich mit β sind. Es zeigt sich, dass diese eine Untergruppe der Automorphismengruppe bilden.

Satz 10.1.9: Sei V ein $\mathbb{K}$-Vektorraum und β eine Bilinearform oder Sesquilinearform auf V.

1. Die β-unitären Automorphismen von V bilden eine Untergruppe $U(V, \beta) \subseteq (\mathrm{Aut}_{\mathbb{K}}(V), \circ)$, die **unitäre Gruppe** von (V, β).

   Für eine symmetrische Bilinearform β heißt sie auch die **orthogonale Gruppe** $\mathrm{O}(V, \beta)$ und für eine antisymmetrische Bilinearform β die **symplektische Gruppe** $\mathrm{Sp}(V, \beta)$.

2. Ist V endlich-dimensional, so bilden die β-unitären Endomorphismen $\phi : V \to V$ mit $\det(\phi) = 1$ eine Untergruppe der Gruppe $\mathrm{U}(V, \beta)$, die mit $\mathrm{SU}(V, \beta)$ und für symmetrische Bilinearformen mit $\mathrm{SO}(V, \beta)$ bezeichnet wird.

Beweis:
1. Wir rechnen die Untergruppenkriterien nach:

(UG1) Es gilt $\beta(\mathrm{id}_V(v), \mathrm{id}_V(w)) = \beta(v, w)$ für $v, w \in V$ und $\mathrm{id}_V \in \mathrm{Aut}_{\mathbb{K}}(V)$, also $\mathrm{id}_V \in U(V, \beta)$.
(UG2) Aus $\phi, \psi \in U(V, \beta)$ folgt $\beta(\phi(\psi(v)), \phi(\psi(w))) = \beta(\psi(v), \psi(w)) = \beta(v, w)$ für $v, w \in V$ und $\phi \circ \psi \in \mathrm{Aut}_{\mathbb{K}}(V)$, also $\phi \circ \psi \in U(V, \beta)$.
(UG3) Es gilt $\beta(v, w) = \beta(\phi(\phi^{-1}(v)), \phi(\phi^{-1}(w))) = \beta(\phi^{-1}(v), \phi^{-1}(w))$ für alle $v, w \in V$ und $\phi^{-1} \in \mathrm{Aut}_{\mathbb{K}}(V)$, also $\phi^{-1} \in U(V, \beta)$. Also ist $\mathrm{U}(V, \beta) \subseteq \mathrm{Aut}_{\mathbb{K}}(V)$ eine Untergruppe.

2. Dass im endlich-dimensionalen Fall die β-unitären Endomorphismen mit $\det(\phi) = 1$ eine Untergruppe von $\mathrm{U}(V, \beta)$ bilden, folgt direkt aus den Eigenschaften der Determinante. Denn es gilt $\det(\mathrm{id}_V) = 1$, also $\mathrm{id}_V \in \mathrm{SU}(V, \beta)$ (UG1). Ist $\det(\phi) = \det(\psi) = 1$, so folgt $\det(\psi \circ \phi) = \det(\psi) \cdot \det(\phi) = 1$, also $\psi \circ \phi \in \mathrm{SU}(V, \beta)$ für alle $\phi, \psi \in \mathrm{SU}(V, \beta)$ (UG2) und $\det(\phi^{-1}) = \det(\phi)^{-1} = 1$, also $\phi^{-1} \in \mathrm{SU}(V, \beta)$ für alle $\phi \in \mathrm{SU}(V, \beta)$ (UG3). $\square$

10.2 Beschreibung durch Matrizen

Wie sich bereits in den betrachteten Beispielen und in Kapitel 5.2 gezeigt hat, ist eine Bi- oder Sesquilinearform β auf einem $\mathbb{K}$-Vektorraum V eindeutig durch ihre Werte auf einer Basis von V bestimmt. Für endlich-dimensionale Vektorräume V können wie sie also durch eine Matrix beschreiben, die als Einträge die Werte $\beta(v_i, v_j)$ für eine geordnete Basis $B = (v_1, ..., v_n)$ von V enthält. Dies reduziert die Betrachtung von Bi- und Sesquilinearformen auf *endlich-dimensionalen* Vektorräumen dann auf die Bilinearformen in Beispiel 10.1.3, 1. und 2.

Definition 10.2.1: Sei V ein n-dimensionaler $\mathbb{K}$-Vektorraum, β eine Bilinearform oder Sesquilinearform auf V. Die **beschreibende Matrix** von β bezüglich einer geordneten Basis $B = (b_1, .., b_n)$ von V ist die Matrix ${}^B\beta^B \in \mathrm{Mat}(n \times n, \mathbb{K})$ mit

$${}^B\beta^B{}_{ij} := \beta(b_i, b_j) \qquad \forall i, j \in \{1, ..., n\}.$$

Beispiel 10.2.2:

1. Für die Bilinearform β auf dem $\mathbb{R}^n$ aus Beispiel 10.1.3, 1. mit $\beta(x, y) = x^T \cdot A \cdot y$ und die Sesquilinearform β auf dem $\mathbb{C}^n$ aus Beispiel 10.1.3, 2. mit $\beta(x, y) = x^\dagger \cdot A \cdot y$ ist die beschreibende Matrix bezüglich der Standardbasis $B = (e_1, ..., e_n)$ gerade die Matrix ${}^B\beta^B = A$, denn es gilt $\beta(e_i, e_j) = e_i^T \cdot A \cdot e_j = e_i^\dagger \cdot A \cdot e_j = A_{ij}$.

2. Insbesondere erhält man für die beschreibende Matrix der Standardskalarprodukte auf $\mathbb{R}^n$ und $\mathbb{C}^n$ bezüglich der Standardbasis ${}^B\beta^B = \mathbb{1}_n$ (vgl. Beispiel 10.1.6).

3. Wir betrachten den Vektorraum $\mathrm{Pol}_{\leq n}(\mathbb{R})$ der Polynomabbildungen $p : \mathbb{R} \to \mathbb{R}$ vom Grad $\leq n$ und die symmetrische Bilinearform

$$\beta(p, q) = \int_0^1 p(x)q(x)\, dx.$$

Dann ist die beschreibende Matrix von β bezüglich der Basis $(1, x, ..., x^n)$ gegeben durch

$$\beta(x^k, x^l) = \int_0^1 x^k \cdot x^l \, dx = \int_0^1 x^{k+l} \, dx = \tfrac{1}{k+l+1} \left(1^{k+l+1} - 0^{k+l+1}\right) = \tfrac{1}{k+l+1}$$
$$\Rightarrow \quad {}^B\beta^B = \begin{pmatrix} 1 & \cdots & \frac{1}{n+1} \\ \vdots & & \vdots \\ \frac{1}{n+1} & \cdots & \frac{1}{2n+1} \end{pmatrix}.$$

Man beachte, dass die beschreibende Matrix einer Bi- oder Sesquilinearform etwas anderes ist, als die beschreibende Matrix einer linearen Abbildung. Die einzige Gemeinsamkeit ist, dass in beiden Fällen die mathematische Struktur durch eine Tabelle mit n Spalten und Zeilen, also eine quadratische Matrix, beschrieben wird. Die Bedeutung und die Eigenschaften dieser Tabelle sind jedoch völlig verschieden.

Wie im Fall der beschreibenden Matrix einer linearen Abbildung liegt die Nützlichkeit der beschreibenden Matrix einer Bi- oder Sesquilinearform β darin, dass sie alle Information über β enthält und sich die Werte $\beta(v, w)$ für beliebige Vektoren $v, w \in V$ leicht daraus berechnen lassen. Ebenso kann man wichtige Eigenschaften von β, wie etwa symmetrisch, alternierend oder (nicht) ausgeartet zu sein, leicht aus ihrer Matrix ablesen. Dies reduziert Bi- oder Sesquilinearformen auf endlich-dimensionalen Vektorräumen auf Beispiel 10.1.3, 1. und 2.

Satz 10.2.3: Sei V ein n-dimensionaler $\mathbb{K}$-Vektorraum und $B = (b_1, ..., b_n)$ eine Basis von V.

1. Jede Bi- oder Sesquilinearform β auf V ist durch ihre beschreibende Matrix bezüglich B eindeutig bestimmt. Es gilt für alle $v, w \in V$:
$$\beta(v, w) = \begin{cases} {}_BS(v)^T \cdot {}^B\beta^B \cdot {}_BS(w) & \beta \text{ bilinear} \\ {}_BS(v)^\dagger \cdot {}^B\beta^B \cdot {}_BS(w) & \beta \text{ sesquilinear.} \end{cases} \tag{21}$$
Dabei bezeichnet ${}_BS : V \to \mathbb{K}^n$, $\Sigma_{j=1}^n \lambda_j b_j \mapsto (\lambda_1, ..., \lambda_n)^T$ den Isomorphismus aus Korollar 4.2.4 und $M^\dagger$ die zu einer Matrix $M \in \mathrm{Mat}(n \times m, \mathbb{C})$ **hermitesch konjugierte Matrix**
$$M^\dagger := \overline{M}^T \in \mathrm{Mat}(m \times n, \mathbb{C}).$$

2. Für jede Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{K})$ und jeden n-dimensionalen $\mathbb{K}$-Vektorraum V gibt es eine Bi- oder Sesquilinearform β auf V mit ${}^B\beta^B = A$.

3. Für jede Bi- oder Sesquilinearform β auf V gilt:
   - β hermitesch $\Leftrightarrow ({}^B\beta^B)^\dagger = {}^B\beta^B$,
   - β antihermitesch $\Leftrightarrow ({}^B\beta^B)^\dagger = -{}^B\beta^B$
   - β symmetrisch $\Leftrightarrow ({}^B\beta^B)^T = {}^B\beta^B$
   - β antisymmetrisch $\Leftrightarrow ({}^B\beta^B)^T = -{}^B\beta^B$
   - β nicht ausgeartet $\Leftrightarrow \mathrm{rg}({}^B\beta^B) = n$.

Beweis:
1. und 2: Die Eindeutigkeit folgt aus den angegebenen Formeln.
Sind $v = \Sigma_{i=1}^n \lambda_i b_i$ und $w = \Sigma_{j=1}^n \mu_j b_j$, so ergibt sich für eine Sesquilinearform β auf V

$$\begin{aligned}\beta(v, w) &= \beta(\Sigma_{i=1}^n \lambda_i b_i, \Sigma_{j=1}^n \mu_j b_j) = \Sigma_{i,j=1}^n \overline{\lambda_i} \mu_j \, \beta(b_i, b_j) = \Sigma_{i,j=1}^n \overline{\lambda_i} \mu_j \, {}^B\beta^B{}_{ij} \\ &= (\overline{\lambda_1}, ..., \overline{\lambda_n}) \cdot {}^B\beta^B \cdot (\mu_1, ... \mu_n)^T = {}_BS(v)^\dagger \cdot {}^B\beta^B \cdot {}_BS(w).\end{aligned}$$

Der Beweis für Bilinearformen ist analog und folgt durch Weglassen der komplexen Konjugation. Umgekehrt definiert man für $A \in \text{Mat}(n \times n, \mathbb{K})$ eine Abbildung $\beta : V \times V \to \mathbb{K}$ durch

$$\beta(v,w) = \begin{cases} {}_BS(v)^T \cdot A \cdot {}_BS(w) & \beta \text{ bilinear} \\ {}_BS(v)^\dagger \cdot A \cdot {}_BS(w) & \beta \text{ sesquilinear,} \end{cases}$$

und eine direkte Rechnung zeigt, dass β eine Bi- bzw. Sesquilinearform ist.

3. Aus 1. ergibt sich, dass eine Sesquilinearform β (anti)hermitesch ist genau dann, wenn $\overline{\beta(b_j, b_i)} = \beta(b_i, b_j)$ ($\overline{\beta(b_j, b_i)} = -\beta(b_i, b_j)$) für alle $i, j \in \{1, ..., n\}$. Dies ist äquivalent zu

$$({}^B\beta^B)^\dagger_{ij} = (\overline{{}^B\beta^B})_{ji} = \overline{\beta(b_j, b_i)} = (-)\beta(b_i, b_j) = (-)({}^B\beta^B)_{ij}.$$

Der Beweis für (anti)symmetrische Bilinearformen ist analog und ergibt sich durch Entfernen der komplexen Konjugation. In beiden Fällen folgt aus 1. , dass β nicht ausgeartet ist, genau dann wenn $\ker({}^B\beta^B) = \{0\}$, also genau dann, wenn $\text{rg}({}^B\beta^B) = n$ gilt. □

Möchte man Bi- oder Sesquilinearformen mit Hilfe von Basen beschreiben, so ergibt sich natürlich die Frage, wie sich die beschreibende Matrix einer Bi- oder Sesquilinearform unter einem Basiswechsel verhält. Anders als im Fall der beschreibenden Matrix einer linearen Abbildung ergibt sich hier jedoch keine Konjugation mit der Basiswechselmatrix. Statt der Inversen in der Transformationsformel für beschreibende Matrizen von Endomorphismen treten hier Transponierte oder hermitesch Konjugierte auf.

Satz 10.2.4: (**Transformationsformel**)
Sei V ein n-dimensionaler $\mathbb{K}$-Vektorraum, $B = (b_1, ..., b_n)$, $C = (c_1, ..., c_n)$ Basen von V und β eine Bilinearform oder Sesquilinearform auf V. Dann gilt:

$${}^B\beta^B = \begin{cases} {}_CM_B(\text{id}_V)^T \cdot {}^C\beta^C \cdot {}_CM_B(\text{id}_V) & \beta \text{ bilinear} \\ {}_CM_B(\text{id}_V)^\dagger \cdot {}^C\beta^C \cdot {}_CM_B(\text{id}_V) & \beta \text{ sesquilinear.} \end{cases}$$

Beweis:
Seien $B = (b_1, ..., b_n)$ und $C = (c_1, ..., c_n)$ Basen von V mit $b_i = \Sigma_{j=1}^n M_{ji} c_j$. Dann ist ${}_CM_B(\text{id}_V) = M = (M_{ji})$, und für jede Sesquilinearform β auf V erhält man

$$\begin{aligned} {}^B\beta^B{}_{ij} &= \beta(b_i, b_j) = \beta(\Sigma_{k=1}^n M_{ki} c_k, \Sigma_{l=1}^n M_{lj} c_l) = \Sigma_{k,l=1}^n \overline{M_{ki}}\, M_{lj}\, \beta(c_k, c_l) \\ &= ({}_CM_B(\text{id}_\mathbb{K})^\dagger \cdot {}^C\beta^C \cdot {}_CM_B(\text{id}_\mathbb{K}))_{ij} \qquad \forall i, j \in \{1, ..., n\}. \end{aligned}$$

Der Beweis für Bilinearformen ergibt sich durch Entfernen der komplexen Konjugation. □

Mit Hilfe der beschreibenden Matrix einer Bi-oder Sesquilinearform β auf V lässt sich insbesondere deren unitäre Gruppe $U(V, \beta)$ beschreiben. Die Bedingung, dass ein Automorphismus $\phi : V \to V$ unitär bezüglich β ist, übersetzt sich dabei in eine Verträglichkeitsbedingung zwischen seiner beschreibenden Matrix und der beschreibenden Matrix von β.

Satz 10.2.5: (**Beschreibung der unitären Gruppe mit Matrizen**)
Sei V ein n-dimensionaler $\mathbb{K}$-Vektorraum, $\phi : V \to V$ ein Endomorphismus und β eine Bi- oder Sesquilinearform auf V. Dann gilt für jede geordnete Basis B von V:

$$\phi \in \text{U}(V, \beta) \quad \Leftrightarrow \quad \begin{cases} {}_BM_B(\phi)^\dagger \cdot {}^B\beta^B \cdot {}_BM_B(\phi) = {}^B\beta^B & \beta \text{ sesquilinear,} \\ {}_BM_B(\phi)^T \cdot {}^B\beta^B \cdot {}_BM_B(\phi) = {}^B\beta^B & \beta \text{ bilinear.} \end{cases}$$

Beweis:
Wir beweisen die Aussage für Sesquilinearformen. Die Beweise für Bilinearformen sind analog und ergeben sich durch Entfernen der komplexen Konjugation.

Sei dazu $B = (b_1, ..., b_n)$ eine geordnete Basis von V und $\phi : V \to V$ ein Endomorphismus mit $\phi(b_i) = \Sigma_{j=1}^n \phi_{ji} b_j$. Dann ist ${}_BM_B(\phi) = (\phi_{ji})$, und es folgt

$$\beta(\phi(b_i), \phi(b_j)) = \beta(\Sigma_{k=1}^n \phi_{ki} b_k, \Sigma_{l=1}^n \phi_{lj} b_l) = \Sigma_{k,l=1}^n \overline{\phi_{ki}}\, \phi_{lj}\, \beta(b_k, b_l) = ({}_BM_B(\phi)^\dagger \cdot {}^B\beta^B \cdot {}_BM_B(\phi))_{ij}.$$

Ist $\phi \in U(\beta, V)$, so ist $\beta(\phi(b_i), \phi(b_j)) = \beta(b_i, b_j)$ und damit ${}_BM_B(\phi)^\dagger \cdot {}^B\beta^B \cdot {}_BM_B(\phi) = {}^B\beta^B$.

Ist umgekehrt $\phi : V \to V$ ein Endomorphismus mit ${}_BM_B(\phi)^\dagger \cdot {}^B\beta^B \cdot {}_BM_B(\phi) = {}^B\beta^B$, so folgt mit Satz 10.2.3 und der Linearität von ϕ, dass $\beta(\phi(v), \phi(w)) = \beta(v, w)$ für alle $v, w \in V$ und somit $\phi \in U(V, \beta)$. □

Satz 10.2.5 reduziert die Bestimmung der unitäre Gruppe $U(V, \beta)$ für jede Bi- oder Sesquilinearform auf einem endlich-dimensionalen $\mathbb{K}$-Vektorraum V auf eine Rechnung mit Matrizen. Ein besonders einfacher Fall liegt sicher dann vor, wenn die beschreibende Matrix ${}^B\beta^B$ der Bilinearform β die Einheitsmatrix ist. Dies gilt nach Beispiel 10.2.2 insbesondere für den $\mathbb{R}^n$ und C^n mit dem Standardskalarprodukt und der Standardbasis. Da diese eine besonders wichtige Rolle spielen, erhalten die zugehörigen Matrixgruppen eigene Namen und eine spezielle Notation.

Korollar 10.2.6:

1. Eine lineare Abbildung $\phi : \mathbb{C}^n \to \mathbb{C}^n$, $v \mapsto U \cdot v$ ist unitär bezüglich des Standardskalarprodukts genau dann, wenn $U^\dagger \cdot U = \mathbb{1}_n$ gilt.

2. Eine lineare Abbildung $\phi : \mathbb{R}^n \to \mathbb{R}^n$, $v \mapsto U \cdot v$ ist unitär bezüglich des Standardskalarprodukts genau dann, wenn $U^T \cdot U = \mathbb{1}_n$ gilt.

Definition 10.2.7:

1. Eine Matrix $U \in \mathrm{Mat}(n \times n, \mathbb{C})$ heißt **unitär**, wenn $U^\dagger \cdot U = \mathbb{1}_n$ gilt. Die Gruppe der unitären $(n \times n)$-Matrizen wird mit $\mathrm{U}(n, \mathbb{C})$ oder $\mathrm{U}(n)$ bezeichnet.

2. Eine Matrix $O \in \mathrm{Mat}(n \times n, \mathbb{R})$ heißt **orthogonal**, wenn $O^T \cdot O = \mathbb{1}_n$ gilt. Die Gruppe der orthogonalen reellen $(n \times n)$-Matrizen wird mit $\mathrm{O}(n, \mathbb{R})$ oder $\mathrm{O}(n)$ bezeichnet.

3. Die Untergruppen der unitären bzw. der orthogonalen $(n \times n)$-Matrizen mit Determinante eins werden mit $\mathrm{SU}(n, \mathbb{C})$ oder $\mathrm{SU}(n)$ bzw. mit $\mathrm{SO}(n, \mathbb{R})$ oder $\mathrm{SO}(n)$ bezeichnet[17].

Betrachtet man als einfachen Spezialfall den $\mathbb{R}^2$, so lässt sich durch eine explizite Rechnung leicht zeigen, dass jeder orthogonale Automorphismus eine Verkettung einer Drehung und einer Spiegelung ist. Dies gilt auch für orthogonale Automorphismen des $\mathbb{R}^n$ mit $n > 2$, wenn man Drehungen um beliebige Achsen und Spiegelungen an einer Hyperebene zulässt, aber der Beweis wird dann etwas aufwändiger.

[17] Die Buchstaben U und O stehen dabei für die englischen Bezeichnungen *unitary* und *orthogonal*, der Buchstabe S charakterisiert die Matrizen mit Determinante Eins und steht für die englische Bezeichnung *special*.

Beispiel 10.2.8: Wir bestimmen die Gruppe $\mathrm{O}(2,\mathbb{R})$. Für eine reelle 2×2-Matrix A gilt

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \Rightarrow \quad A^T \cdot A = \begin{pmatrix} a & c \\ b & d \end{pmatrix} \cdot \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a^2+c^2 & ab+cd \\ ab+cd & b^2+d^2 \end{pmatrix}.$$

Also ist A orthogonal genau dann, wenn $a^2+c^2 = b^2+d^2 = 1$, $ab+cd = 0$. Dies ist der Fall genau dann, wenn $a = \cos(\phi)$, $c = \sin(\phi)$, $b = \cos(\psi)$, $c = \sin(\psi)$ für zwei Winkel $\phi, \psi \in [0, 2\pi)$ gilt und $0 = ab+cd = \cos(\phi)\cos(\psi) + \sin(\phi)\cos(\psi) = \cos(\phi-\psi)$. Aus der letzten Bedingung folgt $\phi - \psi = \frac{(2n+1)\pi}{2}$ mit $n \in \mathbb{N}_0$. Daraus ergibt sich $\cos(\psi) = -\sin(\phi)$, $\sin(\psi) = \cos(\phi)$ oder $\cos(\psi) = \sin(\phi)$, $\sin(\psi) = -\cos(\phi)$ und damit

$$A = \begin{pmatrix} \cos(\phi) & \mp\sin(\phi) \\ \sin(\phi) & \pm\cos(\phi) \end{pmatrix} = \begin{pmatrix} \cos(\phi) & -\sin(\phi) \\ \sin(\phi) & \cos(\phi) \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & \pm 1 \end{pmatrix}$$

Jede orthogonale 2×2-Matrix ist also entweder eine Drehung um einen Winkel $\phi \in [0, 2\pi)$ oder eine Verkettung einer Drehung mit einer Spiegelung an der x_1-Achse. Die Drehungen sind genau die Matrizen in $M \in \mathrm{O}(2,\mathbb{R})$ mit $\det(M) = 1$. Sie bilden die Untergruppe $\mathrm{SO}(2,\mathbb{R}) \subseteq \mathrm{O}(2,\mathbb{R})$.

Auch für andere Bi- oder Sesquilinearformen β auf endlich-dimensionalen Vektorräumen V lässt sich mit Satz 10.2.5 die unitäre Gruppe $U(V,\beta)$ explizit bestimmen.

Beispiel 10.2.9: Wir betrachten die Minkowski-Metrik $\eta(x,y) = -x_1y_1 + \Sigma_{i=2}^n x_iy_i$ auf dem $\mathbb{R}^n$. Dann ist die darstellende Matrix bezüglich der Standardbasis $B = (e_1, ..., e_n)$ gegeben durch

$${}^B\eta^B = \begin{pmatrix} -1 & 0 & \ldots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & 1 \end{pmatrix}.$$

Jeder Endomorphismus $\phi: \mathbb{R}^n \to \mathbb{R}^n$ ist von der Form $\phi: v \mapsto M \cdot v$, wobei $M = {}_BM_B(\phi)$ die beschreibende Matrix von ϕ bezüglich der Standardbasis bezeichnet. Man erhält

$$\begin{aligned} M^T \cdot {}^B\eta^B \cdot M &= \begin{pmatrix} m_{11} & m_{21} & \ldots & m_{n1} \\ m_{12} & m_{22} & \ldots & m_{n2} \\ \vdots & \vdots & & \vdots \\ m_{1n} & m_{2n} & \ldots & m_{nn} \end{pmatrix} \cdot \begin{pmatrix} -1 & 0 & \ldots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} m_{11} & m_{12} & \ldots & m_{1n} \\ m_{21} & m_{22} & \ldots & m_{2n} \\ \vdots & \vdots & & \vdots \\ m_{n1} & m_{n2} & \ldots & m_{nn} \end{pmatrix} \\ &= \begin{pmatrix} \eta(m_1,m_1) & \eta(m_1,m_2) & \ldots & \eta(m_1,m_n) \\ \eta(m_2,m_1) & \eta(m_2,m_2) & \ldots & \eta(m_2,m_n) \\ \vdots & \vdots & & \vdots \\ \eta(m_n,m_1) & \eta(m_n,m_2) & \ldots & \eta(m_n,m_n) \end{pmatrix}. \end{aligned}$$

wobei $m_1, ..., m_n \in \mathbb{R}^n$ die Spaltenvektoren von M bezeichnen. Also ist nach Satz 10.2.5 ϕ genau dann unitär bezüglich η, wenn für seine Spaltenvektoren gilt:

$$\eta(m_i, m_j) = ({}^B\eta^B)_{ij} = \begin{cases} -1 & i = j = 1 \\ 1 & i = j \in \{2, ..., n\} \\ 0 & i \neq j. \end{cases}$$

Wie schwierig sich die Berechnung der unitären Gruppe $U(V,\beta)$ gestaltet, hängt offensichtlich stark davon ab, wie kompliziert die beschreibende Matrix von β bezüglich der gewählten Basis ist. Gesucht ist also jeweils eine Basis, so dass die beschreibende Matrix von β eine möglichst einfache Form annimmt. Während es nur für Skalarprodukte möglich ist, eine Einheitsmatrix als beschreibende Matrix zu erhalten, ist es aber in vielen Fällen möglich, durch geschickte Basiswahlen eine Diagonalmatrix zu erhalten. Dies führt auf die Begriffe der *Orthonormalbasis* und der *Orthogonalbasis*.

Definition 10.2.10: Sei V ein $\mathbb{K}$-Vektorraum und β eine Bilinearform oder Sesquilinearform auf V. Ein m-Tupel von Vektoren $(v_1, ..., v_m)$ mit $v_i \in V$ heißt

- **Orthogonalsystem**, wenn $v_i \neq 0$ und $\beta(v_i, v_j) = 0$ für alle $i, j \in \{1, ..., m\}$ mit $i \neq j$,
- **Orthonormalsystem**, wenn $\beta(v_i, v_j) = 0$ und $\beta(v_i, v_i) = 1$ für $i, j \in \{1, ..., m\}$, $i \neq j$,
- **Orthogonalbasis**, wenn es ein Orthogonalsystem und eine Basis von V ist,
- **Orthonormalbasis**, wenn ein Orthonormalsystem und eine Basis von V ist.

Beispiel 10.2.11:

1. Die Standardbasis $B = (e_1, ..., e_n)$ ist eine Orthonormalbasis für das Standardskalarprodukt auf dem $\mathbb{R}^n$ und dem $\mathbb{C}^n$.
2. Die Standardbasis $B = (e_1, ..., e_n)$ des $\mathbb{R}^n$ ist eine Orthogonalbasis für die Minkowski-Metrik $\eta(x, y) = -x_1 y_1 + \Sigma_{i=2}^n x_i\, y_i$, aber keine Orthonormalbasis, denn $\eta(e_1, e_1) = -1$.
3. Ist $\beta \neq 0$ alternierend, so kann es keine Orthogonalbasis zu β geben. Denn gäbe es eine Orthogonalbasis $B = (v_1, ..., v_n)$, so wäre $\beta(v_i, v_j) = 0$ für $i \neq j$ und $\beta(v_i, v_i) = 0$ für alle $i \in \{1, ..., n\}$, da β alternierend ist. Daraus würde $\beta = 0$ folgen.

Eine *Orthonormalbasis* einer Bi- oder Sesquilinearform β auf einem endlich-dimensionalen Vektorraum V ist also eine Basis B, so dass die beschreibende Matrix ${}^B\beta^B$ die Einheitsmatrix ist, und eine *Orthogonalbasis* ist eine Basis B, so dass ${}^B\beta^B$ eine Diagonalmatrix ist. Die Beispiele zeigen, dass es zu Bi- oder Sesquilinearformen auf endlich-dimensionalen Vektorräumen nicht unbedingt eine Orthogonal- oder Orthonormalbasis geben muss. Wir werden in den folgenden Abschnitten aber zeigen, dass für symmetrische Bilinearformen und hermitesche Sesquilinearformen immer eine Orthogonalbasis existiert.

Um das Konzept der Orthogonalität besser zu verstehen und mit der geometrischen Anschauung in Verbindung zu bringen, untersuchen wir zunächst den Spezialfall, dass β ein Skalarprodukt ist, und behandeln anschließend den allgemeinen Fall. Der Fall euklidischer und unitärer Vektorräume ist auch deswegen wichtig, weil hier eine direkte Verbindung zwischen geometrischen Begriffen und den Vektorraumstrukturen besteht.

10.3 Euklidische und unitäre Vektorräume

Da Skalarprodukte positiv definit sind, statten sie den zugrundeliegenden Vektorraum mit einem Längenbegriff aus. Die Länge eines Vektors ergibt sich dabei als Wurzel aus dem Skalarprodukt des Vektors mit sich selbst. Offensichtlich ist die positive Semidefinitheit notwendig, damit diese Wurzel überhaupt definiert ist. Die positive Definitheit sorgt dann dafür, dass der Nullvektor der einzige Vektor der Länge Null ist.

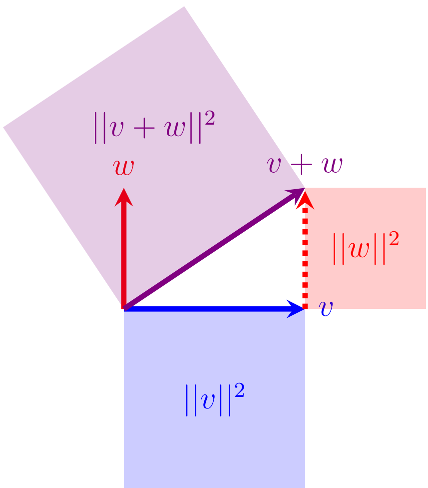
Definition 10.3.1: Sei V mit $\langle\, ,\, \rangle$ ein euklidischer oder unitärer Vektorraum.

1. Die **Länge** eines Vektors $v \in V$ ist $||v|| := \sqrt{\langle v, v\rangle} \in \mathbb{R}_{\geq 0}$.
2. Ein Vektor $v \in V$ heißt **normiert**, wenn $||v|| = 1$.
3. Zwei Vektoren $v, w \in V$ heißen **orthogonal**, wenn $\langle v, w\rangle = 0$.

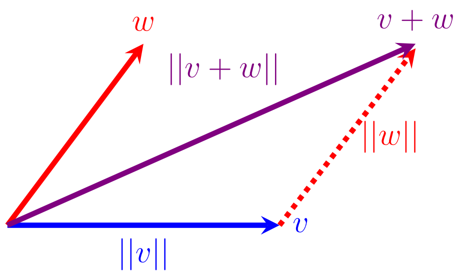
Ein Längenbegriff ist mehr als eine Zuordnung einer nicht-negativen Zahl zu jedem Vektor. Dass die Länge $||v||$ eines Vektors $v \in V$ tatsächlich als eine geometrische Länge interpretiert werden kann und Orthogonalität von Vektoren tatsächlich etwas mit dem geometrischen Orthogonalitätsbegriff zu tun hat, ergibt sich aus dem folgenden Satz.

Satz 10.3.2: Sei V mit $\langle\, ,\, \rangle$ ein euklidischer oder unitärer Vektorraum. Dann gilt:

1. **Satz des Pythagoras**:
$$\langle v, w\rangle = 0 \quad \Rightarrow \quad ||v+w||^2 = ||v||^2 + ||w||^2,$$
2. **Cauchy-Schwartz-Ungleichung**:
$$|\langle v, w\rangle| \leq ||v|| \cdot ||w|| \quad \text{und} \quad |\langle v, w\rangle| = ||v|| \cdot ||w|| \Leftrightarrow v, w \text{ linear abhängig},$$
3. **Dreiecksungleichung**:
$$||v+w|| \leq ||v|| + ||w|| \quad \text{und} \quad ||v+w|| = ||v|| + ||w|| \Leftrightarrow v, w \text{ linear abhängig}.$$



Der Satz des Pythagoras



Die Dreiecksungleichung.

Beweis:
1. Ist $\langle v, w\rangle = \overline{\langle w, v\rangle} = 0$, so ergibt sich aus der Definition der Länge der Satz des Pythagoras
$$||v+w||^2 = \langle v+w, v+w\rangle = \langle v, v\rangle + \langle w, w\rangle + \langle v, w\rangle + \langle w, v\rangle = ||v||^2 + ||w||^2.$$
2. Aus der positiven Definitheit folgt für alle $\lambda \in \mathbb{K}$ und $v, w \in V$
$$0 \leq \langle v - \lambda w, v - \lambda w\rangle = \langle v, v\rangle - \lambda\langle v, w\rangle - \overline{\lambda}\langle w, v\rangle + |\lambda|^2\langle w, w\rangle = ||v||^2 - \lambda\langle v, w\rangle - \overline{\lambda\langle v, w\rangle} + |\lambda|^2||w||^2.$$
Ist $\langle v, w\rangle = 0$, so ist die Cauchy-Schwartzsche Ungleichung offensichtlich erfüllt. Ansonsten betrachtet man für $\mu \in \mathbb{R}$
$$\lambda = \frac{\overline{\langle v, w\rangle}\mu}{|\langle v, w\rangle|} \quad \Rightarrow \quad 0 \leq \langle v - \lambda w, v - \lambda w\rangle = ||v||^2 - 2\mu|\langle v, w\rangle| + \mu^2||w||^2. \tag{22}$$

Da es sich dabei um ein quadratisches Polynom in μ handelt, kann diese Bedingung nur dann erfüllt sein, wenn dieses Polynom maximal eine Nullstelle in $\mathbb{R}$ hat. Also muss seine Diskriminante ≤ 0 sein, und es folgt die Cauchy-Schwartzsche Ungleichung

$$\frac{|\langle v,w\rangle|^2}{||w||^4}-\frac{||v||^2}{||w||^2}\leq 0 \quad\Rightarrow\quad |\langle v,w\rangle|\leq ||v||\cdot||w||.$$

Gilt $|\langle v,w\rangle| = ||v||\cdot||w||$, so ist entweder $||w||=0$ oder das Polynom in (22) hat die Nullstelle $\mu = ||v||/||w||$, was $v = \langle v,w\rangle w/||w||^2$ impliziert. In beiden Fällen sind v, w linear abhängig.

3. Die Dreiecksungleichung ergibt sich direkt aus der Cauchy-Schwartzschen Ungleichung:

$$\begin{aligned}||v+w||^2 =&\langle v+w,v+w\rangle = \langle v,v\rangle+\langle w,w\rangle+\langle v,w\rangle+\langle w,v\rangle = ||v||^2+||w||^2+\langle v,w\rangle+\overline{\langle v,w\rangle}\\ =&||v||^2+||w||^2+2\mathrm{Re}\,\langle v,w\rangle \overset{2.}{\leq} ||v||^2+||w||^2+2||v||\cdot||w|| = (||v||+||w||)^2.\end{aligned}$$

□

Die Cauchy-Schwartzsche Ungleichung ermöglicht es uns, Winkel zwischen Vektoren eines euklidischen oder unitären Vektorraums zu definieren. Insbesondere folgt aus der Cauchy-Schwartzschen Ungleichung in Satz 10.3.2

$$-1\leq \frac{\mathrm{Re}\,\langle v,w\rangle}{||v||\cdot||w||}\leq 1$$

für alle Vektoren $v,w\in V\setminus\{0\}$, und dass jeder Wert in $[-1,1]$ für gewisse Vektoren $v,w\in V\setminus\{0\}$ angenommen wird. Daher können wir diesen Ausdruck als den Kosinus eines Winkels interpretieren, und dieser Winkel ist eindeutig bestimmt, wenn wir ihn in $[0,\pi]$ wählen.

Definition 10.3.3: Sei V ein euklidischer oder unitärer Vektorraum. Der **Winkel** $\theta\in[0,\pi]$ zwischen zwei Vektoren $v,w\in V\setminus\{0\}$ ist definiert durch

$$\cos\theta = \frac{\mathrm{Re}\,\langle v,w\rangle}{||v||\cdot||w||}.$$

Beispiel 10.3.4: Für das Standardskalarprodukt auf dem $\mathbb{R}^2$

$$\langle v,w\rangle = v_1w_1+v_2w_2$$

ergibt sich für die Länge eines Vektors $v=(v_1,v_2)^T$ aus Definition 10.3.1 der bekannte Ausdruck

$$||v|| = \sqrt{v_1^2+v_2^2}.$$

Sind $v=(v_1,v_2)^T$ und $w=(w_1,w_2)^T$ zwei Vektoren in $\mathbb{R}^2\setminus\{0\}$, so sind die Winkel α und β, die v und w mit der positiven x-Achse einschließen, gegeben durch

$$\cos\alpha=\frac{v_1}{||v||}\qquad \sin\alpha=\frac{v_2}{||v||}\qquad \cos\beta=\frac{w_1}{||w||}\qquad \sin\beta=\frac{w_2}{||w||}.$$

Der Winkel zwischen v und w ist dann $\beta-\alpha$ oder $\alpha-\beta$, und mit der Additionsformel folgt

$$\cos(\beta-\alpha)=\cos(\alpha-\beta)=\cos\beta\cos\alpha+\sin\beta\sin\alpha=\frac{v_1w_1+v_2w_2}{||v||\cdot||w||}=\frac{\langle v,w\rangle}{||v||\cdot||w||}.$$

Das Beispiel zeigt, dass der Längenbegriff aus Definition 10.3.1 mit dem geometrischen Längenbegriff im $\mathbb{R}^2$ übereinstimmt und der Winkel aus Definition 10.3.3 tatsächlich den Winkel zwischen zwei Vektoren im $\mathbb{R}^2$ angibt. Insbesondere sind zwei Vektoren $v, w \in \mathbb{R}^2 \setminus \{0\}$ orthogonal im Sinn von Definition 10.3.1 genau dann, wenn sie geometrisch orthogonal sind.

In diesem Fall hat das von den Vektoren v, w und $v+w$ gebildete Dreieck einen rechten Winkel zwischen v und w. Da die Quadrate seiner Seitenlängen $||v||^2$, $||w||^2$ und $||v+w||^2$ sind, ist die Aussage in Satz 10.3.2, 1. dann gerade der Satz des Pythagoras.

Für komplexe oder unendlich-dimensionale reelle Vektorräume V erhält man keine direkte geometrische Interpretation von Längen und Winkeln. Aber auch in diesen Fällen liefert die Länge aus Definition 10.3.1 eine Norm auf V.

Satz 10.3.5: Sei V mit $\langle\, ,\, \rangle$ ein euklidischer oder unitärer Vektorraum. Dann ist $||\ || : V \to \mathbb{R}$, $v \mapsto \sqrt{\langle v, v\rangle}$ eine **Norm** auf V und V somit ein **normierter Vektorraum**, d. h. es gilt:

(N1) **positive Definitheit:** $||v|| \geq 0$ für alle $v \in V$ und $||v|| = 0$ genau dann, wenn $v = 0$,
(N2) **Homogenität:** $||\lambda v|| = |\lambda|\, ||v||$ für alle $v \in V$, $\lambda \in \mathbb{K}$,
(N3) **Dreiecksungleichung:** $||v + w|| \leq ||v|| + ||w||$ für alle $v, w \in V$.

Außerdem gilt die **Polarisierungsidentität**

$$\begin{aligned} \mathbb{K} = \mathbb{R}: &\qquad \langle v, w\rangle = \tfrac{1}{2}||v+w||^2 - \tfrac{1}{2}||v||^2 - \tfrac{1}{2}||w||^2 \\ \mathbb{K} = \mathbb{C}: &\qquad \langle v, w\rangle = \tfrac{1}{4}||v+w||^2 - \tfrac{1}{4}||v-w||^2 - \tfrac{\mathrm{i}}{4}||v+\mathrm{i}w||^2 + \tfrac{\mathrm{i}}{4}||v-\mathrm{i}w||^2. \end{aligned}$$

Beweis:
Dass für jedes Skalarprodukt $\langle\, ,\rangle$ auf V die Abbildung $||\ || : V \to \mathbb{R}$, $v \mapsto \sqrt{\langle v, v\rangle}$ eine Norm auf v ist, ergibt sich direkt aus den Eigenschaften des Skalarprodukts. Die positive Definitheit der Norm folgt aus der positiven Definitheit des Skalarprodukts, die Homogenität aus der Sesquilinearität des Skalarprodukts und die Dreiecksungleichung gilt nach Satz 10.3.2. Die Polarisierungsidentität ergibt sich durch Nachrechnen. Für $\mathbb{K} = \mathbb{R}$ erhält man

$$\begin{aligned} &||v+w||^2 - ||v||^2 - ||w||^2 = \langle v+w, v+w\rangle - \langle v, v\rangle - \langle w, w\rangle \\ =&||v||^2 + \langle w, v\rangle + \langle v, w\rangle + ||w||^2 - ||v||^2 - ||w||^2 = \langle v, w\rangle + \langle w, v\rangle = 2\langle v, w\rangle, \end{aligned}$$

und für $\mathbb{K} = \mathbb{C}$ ergibt sich

$$\begin{aligned} &||v+w||^2 - ||v-w||^2 - \mathrm{i}||v+\mathrm{i}w||^2 + \mathrm{i}||v-\mathrm{i}w||^2 \\ &= \langle v+w, v+w\rangle - \langle v-w, v-w\rangle - \mathrm{i}\langle v+\mathrm{i}w, v+\mathrm{i}w\rangle + \mathrm{i}\langle v-\mathrm{i}w, v-\mathrm{i}w\rangle \\ &= ||v||^2 + \langle w, v\rangle + \langle v, w\rangle + ||w||^2 - (||v||^2 - \langle w, v\rangle - \langle v, w\rangle + ||w||^2) \\ &- \mathrm{i}(||v||^2 - \mathrm{i}\langle w, v\rangle + \mathrm{i}\langle v, w\rangle - ||w||^2) + \mathrm{i}(||v||^2 + \mathrm{i}\langle w, v\rangle - \mathrm{i}\langle v, w\rangle - ||w||^2) \\ &= 2\langle v, w\rangle + 2\overline{\langle v, w\rangle} - 2\overline{\langle v, w\rangle} + 2\langle v, w\rangle = 4\mathrm{Re}\,\langle v, w\rangle + 4\mathrm{i}\mathrm{Im}\,\langle v, w\rangle = 4\langle v, w\rangle. \end{aligned}$$

$\square$

Ein Skalarprodukt auf einem reellen oder komplexen Vektorraum V liefert also eine Norm auf V, aber nicht alle Normen auf V entstehen aus Skalarprodukten. Ob eine gegebene Norm $||\ ||$ auf V durch ein Skalarprodukt induziert wird, lässt sich mit Hilfe der Polarisierungsidentität testen. Die rechte Seite der Polarisierungsidentität ist nämlich für jede beliebige Norm definiert. So erhält man für jede Norm auf V eine Abbildung $\langle\ ,\ \rangle : V \times V \to \mathbb{R}$ und kann überprüfen, ob es sich dabei um ein Skalarprodukt handelt.

Die Axiome (N1)-(N3) kann man als die Mindestanforderungen an einen *Längenbegriff* sehen, der mit der Vektorraumstruktur, also insbesondere der Skalarmultiplikation kompatibel ist. Eine Länge eines Vektors sollte sicherlich eine nicht-negative reelle Zahl sein und Null genau dann, wenn es sich um den Nullvektor handelt (N1). Reskalieren eines Vektors mit einem Skalar $\lambda \in \mathbb{K}$ sollte einem Reskalieren der Länge mit $|\lambda| \in \mathbb{R}_{\geq 0}$ entsprechen (N2), und die Länge der Summe zweier Vektoren sollte maximal so groß sein, wie die Summe ihrer Längen (N3).

Insbesondere liefert ein solcher Längenbegriff einen *Abstandsbegriff* auf dem Vektorraum V, wobei der Abstand zweier Punkte $x, y \in V$ als $||x-y||$ definiert ist. Zu den Mindestanforderungen an einen Abstandsbegriff auf einer Menge M gehört sicherlich, dass der Abstand zweier Elemente $x, y \in M$ eine nicht-negative reelle Zahl ist, und Null genau dann, wenn die Elemente gleich sind. Der Abstand eines Elements $x \in M$ von einem Element $y \in M$ sollte außerdem gleich dem Abstand des Elements y von x sein. Der Abstand eines Elements $x \in M$ von $z \in M$ sollte maximal so groß werden wie die Summe der Abstände von $x \in M$ zu $y \in M$ und von $y \in M$ zu $z \in M$. Dies sind genau die Axiome für einen *metrischen Raum*.

Bemerkung 10.3.6: Ist V mit $||\ ||$ ein normierter Vektorraum, so ist V mit der **Abstandsfunktion** $d : V \times V \to \mathbb{R}$, $(x, y) \mapsto ||x-y||$ ein **metrischer Raum**, d. h. die Abstandsfunktion erfüllt die Bedingungen:

(M1) **Positive Definitheit**: $d(x, y) \geq 0$ für alle $x, y \in M$ und $d(x, y) = 0 \Leftrightarrow x = y$.
(M2) **Symmetrie**: $d(x, y) = d(y, x)$ für alle $x, y \in M$,
(M3) **Dreiecksungleichung**: $d(x, z) \leq d(x, y) + d(y, z)$ für alle $x, y, z \in M$.

Ein metrischer Raum ist also eine Menge mit einem *Abstandsbegriff*. Ein normierter Vektorraum ist ein Vektorraum mit einem *Abstandsbegriff* bzw. *Längenbegriff*, der mit der Vektorraumstruktur kompatibel ist. Ein Skalarprodukt liefert einen solchen Längenbegriff und zusätzlich noch einen Begriff von *Winkel*.

Ist β eine beliebige symmetrische Bilinearform auf einem $\mathbb{K}$-Vektorraum V, so kann man ähnlich wie in der Definition der Norm eine Abbildung $Q_\beta : V \times V \to \mathbb{K}$, $v \mapsto \beta(v, v)$ definieren. Die Wurzel aus $\beta(v, v)$ zu ziehen, ergibt allerdings nur über $\mathbb{R}$ oder $\mathbb{C}$ Sinn und auch dort nur dann, wenn β positiv semidefinit ist. Stattdessen erhält man für $\mathrm{char}(\mathbb{K}) \neq 2$ eine sogenannte *quadratische Form*.

Bemerkung 10.3.7: Sei $\mathbb{K}$ ein Körper der Charakteristik $\mathrm{char}(\mathbb{K}) \neq 2$ und V ein Vektorraum über $\mathbb{K}$. Eine Abbildung $Q : V \to \mathbb{K}$ heißt **quadratische Form**, wenn gilt:

(Q1) $Q(\lambda v) = \lambda^2 Q(v)$ für alle $v \in V$, $\lambda \in \mathbb{K}$,
(Q2) Die Abbildung $\beta_Q : V \times V \to \mathbb{K}$ mit $\beta_Q(v, w) = \frac{1}{2}Q(v+w) - \frac{1}{2}Q(v) - \frac{1}{2}Q(w)$ ist bilinear.

Ist β eine symmetrische Bilinearform auf V, so ist $Q_\beta : V \to \mathbb{K}$, $v \mapsto \beta(v, v)$ eine quadratische Form. Umgekehrt ist für jede quadratische Form $Q : V \to \mathbb{K}$ die Abbildung $\beta_Q : V \times V \to \mathbb{K}$ eine symmetrische Bilinearform auf V mit $Q_\beta = Q$ als zugehörige quadratische Form.

Mit Hilfe dieser geometrischen Anschauung können wir nun den Begriff der Orthogonalität genauer untersuchen und zeigen, dass im Fall euklidischer oder unitärer Vektorräume immer Orthonormalbasen existieren. Da auch jede Einschränkung des Skalarprodukts auf einen Untervektorraum nach Bemerkung 10.1.5 wieder ein Skalarprodukt liefert, gilt dies insbesondere auch für Untervektorräume euklidischer und unitärer Vektorräume. Als ersten Schritt stellen wir dazu fest, dass Orthogonalsysteme für Skalarprodukte immer linear unabhängig sind.

Lemma 10.3.8: Sei V mit $\langle\, ,\, \rangle$ ein euklidischer oder unitärer Vektorraum. Dann gilt:

1. Jedes Orthogonalsystem $(v_1, ..., v_m)$ ist linear unabhängig.
2. Ist $\dim_{\mathbb{K}}(V) = n \in \mathbb{N}$, so ist jedes n-elementige Orthogonalsystem (Orthonormalsystem) eine Orthogonalbasis (Orthonormalbasis) von V.

Beweis:
Die zweite Aussage folgt aus der ersten, da in einem Vektorraum V der Dimension $\dim_{\mathbb{K}}(V) = n$ jede n-elementige linear unabhängige Teilmenge eine Basis ist. Sei $(v_1, ..., v_m)$ ein Orthogonalsystem und $\Sigma_{i=1}^m \lambda_i v_i = 0$ mit $\lambda_i \in \mathbb{K}$. Dann gilt

$$0 = \langle 0, 0\rangle = \langle \Sigma_{i=1}^m \lambda_i v_i, \Sigma_{j=1}^m \lambda_j v_j\rangle = \Sigma_{i,j=1}^m \overline{\lambda_i}\, \lambda_j\, \langle v_i, v_j\rangle = \Sigma_{i=1}^n \overline{\lambda_i}\, \lambda_i \langle v_i, v_i\rangle = \Sigma_{i=1}^m |\lambda_i|^2 \langle v_i, v_i\rangle.$$

Da $|\lambda_i|^2 > 0$ für alle $\lambda_i \in \mathbb{K} \setminus \{0\}$ und wegen $v_i \neq 0$ auch $\langle v_i, v_i\rangle > 0$ für alle $i \in \{1, ..., m\}$ gilt, folgt $\lambda_1 = ... = \lambda_m = 0$. Also ist $(v_1, ..., v_m)$ linear unabhängig. □

Die Aussagen aus diesem Lemma sind im Allgemeinen falsch für Bi- oder Sesquilinearformen, die keine Skalarprodukte sind. Es ist nämlich die *positive Definitheit* des Skalarprodukts, die im Beweis von Lemma 10.3.8 die lineare Unabhängigkeit des Orthogonalsystems sicherstellt. So ist beispielsweise $(e_1 + e_2, e_1 + e_2)$ ein Orthogonalsystem bezüglich der Minkowski-Metrik auf dem $\mathbb{R}^2$, aber offensichtlich nicht linear unabhängig.

Da ein Orthogonalsystem in einem unitären oder euklidischen Vektorraum linear unabhängige Teilmengen dieses Vektorraums liefert, hat man sicherlich ein Interesse daran, Orthogonalsysteme zu konstruieren. Dies lässt sich effizient mit einem Algorithmus erledigen, dem sogenannten *Gram-Schmidt-Verfahren*. Dieses Verfahren konstruiert aus beliebigen linear unabhängigen m-Tupeln von Vektoren eines unitären oder euklidischen Vektorraums ein Orthonormalsystem, das den selben Untervektorraum aufspannt, wie das betrachtete m-Tupel.

Satz 10.3.9: (**Gram-Schmidt-Verfahren**)
Sei V mit $\langle\, ,\, \rangle$ ein euklidischer oder unitärer Vektorraum und $(v_1, ..., v_m)$ ein linear unabhängiges m-Tupel von Vektoren aus V. Dann liefert der folgende Algorithmus eine Orthonormalbasis $B = (b_1, ..., b_m)$ des Untervektorraums $U = \mathrm{span}_{\mathbb{K}}\{v_1, ..., v_m\}$:

$$\begin{aligned}
&u_1 := v_1 && b_1 = \frac{u_1}{||u_1||}\\
&u_2 := v_2 - \langle b_1, v_2\rangle\, b_1 = v_2 - \frac{\langle u_1, v_2\rangle}{\langle u_1, u_1\rangle}\, u_1 && b_2 = \frac{u_2}{||u_2||}\\
&\vdots && \vdots\\
&u_k := v_k - \sum_{j=1}^{k-1} \langle b_j, v_k\rangle\, b_j = v_k - \sum_{j=1}^{k-1} \frac{\langle u_j, v_k\rangle}{\langle u_j, u_j\rangle} u_j && b_k = \frac{u_k}{||u_k||}\\
&\vdots && \vdots\\
&u_m := v_m - \sum_{j=1}^{m-1} \langle b_j, v_m\rangle\, b_j = v_m - \sum_{j=1}^{m-1} \frac{\langle u_j, v_m\rangle}{\langle u_j, u_j\rangle} u_j && b_m = \frac{u_m}{||u_m||}.
\end{aligned}$$

Beweis:
Induktion über m.
$m = 1$: Dann gilt $v_1 \neq 0$ wegen (v_1) linear unabhängig, und man setzt $b_1 := v_1/||v_1||$. Offensichtlich ist (b_1) dann eine Orthonormalbasis von $\mathbb{K}v_1 = \mathbb{K}b_1$.

$m-1 \to m$: Sei gezeigt, dass das Verfahren für alle linear unabhängigen k-Tupel $(v_1, ..., v_k)$ mit $k \leq m-1$ ein Orthonormalsystem $(b_1, ..., b_k)$ mit $\mathrm{span}_{\mathbb{K}}\{b_1, ..., b_k\} = \mathrm{span}_{\mathbb{K}}\{v_1, ..., v_k\}$ liefert.

Sei $(v_1, ..., v_m)$ linear unabhängig. Dann ist auch $(v_1, ..., v_{m-1})$ linear unabhängig, und nach Induktionsvoraussetzung ist $(b_1, ..., b_{m-1})$ ein Orthonormalsystem mit $\mathrm{span}_{\mathbb{K}}\{b_1, ..., b_{m-1}\} = \mathrm{span}_{\mathbb{K}}\{v_1, ..., v_{m-1}\}$. Sei nun $u_m := v_m - \Sigma_{i=1}^{m-1}\langle b_j, v_m\rangle b_j$.

Dann ist $u_m \neq 0$, denn ansonsten wäre $v_m \in \mathrm{span}_{\mathbb{K}}\{b_1, ..., b_{m-1}\} = \mathrm{span}_{\mathbb{K}}\{v_1, ..., v_{m-1}\}$, ein Widerspruch zur linearen Unabhängigkeit von $(v_1, ..., v_m)$. Also können wir $b_m = u_m/||u_m||$ setzen und erhalten $||b_m|| = ||u_m||/||u_m|| = 1$.

Da alle Vektoren b_j Linearkombinationen der Vektoren v_j sind, gilt $\mathrm{span}_{\mathbb{K}}\{b_1, .., b_m\} \subseteq \mathrm{span}_{\mathbb{K}}\{v_1, ..., v_m\}$. Andererseits ist $v_m = u_m + \Sigma_{j=1}^{m-1}\langle b_j, v_m\rangle b_j \in \mathrm{span}_{\mathbb{K}}\{b_1, ..., b_m\}$ und damit $\mathrm{span}_{\mathbb{K}}\{v_1, ..., v_m\} = \mathrm{span}_{\mathbb{K}}\{b_1, ..., b_m\}$. Außerdem gilt

$$\begin{aligned}
||u_m||\langle b_j, b_m\rangle &= \langle b_j, u_m\rangle = \langle b_j, v_m - \Sigma_{k=1}^{m-1}\langle b_k, v_m\rangle b_k\rangle = \langle b_j, v_m\rangle - \Sigma_{k=1}^{m-1}\langle b_k, v_m\rangle\langle b_j, b_k\rangle \\
&= \langle b_j, v_m\rangle - \Sigma_{k=1}^{m-1}\langle b_k, v_m\rangle\delta_{jk} = \langle b_j, v_m\rangle - \langle b_j, v_m\rangle = 0 \qquad \forall j \in \{1, ..., m-1\}.
\end{aligned}$$

Da $(b_1, ..., b_{m-1})$ bereits ein Orthonormalsystem ist und $||b_m|| = 1$, ist damit auch $(b_1, ..., b_m)$ ein Orthonormalsystem. □

Korollar 10.3.10: Jeder endlich-dimensionale Untervektorraum eines euklidischen oder unitären Vektorraums besitzt eine Orthonormalbasis. Insbesondere besitzt jeder endlich-dimensionale euklidische oder unitäre Vektorraum eine Orthonormalbasis.

Beispiel 10.3.11: Wir betrachten $\mathbb{C}^4$ mit dem Standardskalarprodukt und das linear unabhängige Tripel (v_1, v_2, v_3) mit

$$v_1 = \begin{pmatrix} \mathrm{i} \\ \mathrm{i} \\ 0 \\ 0 \end{pmatrix} \qquad v_2 = \begin{pmatrix} 1 \\ 0 \\ \mathrm{i} \\ 0 \end{pmatrix} \qquad v_3 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ \mathrm{i} \end{pmatrix}.$$

Mit dem Gram-Schmidtschen Orthogonalisierungsverfahren erhält man:

$$\begin{aligned}
&u_1 := v_1 \quad \Rightarrow \quad ||u_1||^2 = 2 \quad \Rightarrow \quad b_1 = \tfrac{1}{\sqrt{2}}\, u_1 \\
&u_2 := v_2 - \frac{\langle u_1, v_2\rangle}{||u_1||^2} u_1 = \begin{pmatrix} 1 \\ 0 \\ \mathrm{i} \\ 0 \end{pmatrix} + \tfrac{\mathrm{i}}{2} \begin{pmatrix} \mathrm{i} \\ \mathrm{i} \\ 0 \\ 0 \end{pmatrix} = \tfrac{1}{2} \begin{pmatrix} 1 \\ -1 \\ 2\mathrm{i} \\ 0 \end{pmatrix} \\
&\Rightarrow ||u_2||^2 = \tfrac{3}{2} \quad \Rightarrow \quad b_2 = \sqrt{\tfrac{2}{3}}\, u_2 \\
&u_3 := v_3 - \frac{\langle u_1, v_3\rangle}{||u_1||^2} u_1 - \frac{\langle u_2, v_3\rangle}{||u_2||^2} u_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ \mathrm{i} \end{pmatrix} + \tfrac{\mathrm{i}}{2} \begin{pmatrix} \mathrm{i} \\ \mathrm{i} \\ 0 \\ 0 \end{pmatrix} + \tfrac{1+2\mathrm{i}}{6} \begin{pmatrix} 1 \\ -1 \\ 2\mathrm{i} \\ 0 \end{pmatrix} = \tfrac{1}{3} \begin{pmatrix} \mathrm{i}-1 \\ 1-\mathrm{i} \\ 1+\mathrm{i} \\ 3\mathrm{i} \end{pmatrix} \\
&\Rightarrow ||u_3||^2 = \tfrac{5}{3} \quad \Rightarrow \quad b_3 = \sqrt{\tfrac{3}{5}}\, u_3
\end{aligned}$$

$$\Rightarrow \quad b_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} \mathrm{i} \\ \mathrm{i} \\ 0 \\ 0 \end{pmatrix} \qquad b_2 = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ -1 \\ 2\mathrm{i} \\ 0 \end{pmatrix} \qquad b_3 = \frac{1}{\sqrt{15}} \begin{pmatrix} \mathrm{i}-1 \\ 1-\mathrm{i} \\ 1+\mathrm{i} \\ 3\mathrm{i} \end{pmatrix}.$$

Da jede geordnete Basis $B = (v_1, ..., v_n)$ eines n-dimensionalen $\mathbb{K}$-Vektorraums V einen Vektorraumisomorphismus ${}_BS : V \to \mathbb{K}^n$, $\Sigma_{i=1}^n \lambda_i v_i \mapsto (\lambda_1, ..., \lambda_n)^T$ definiert, stellt sich die Frage, wie sich die Tatsache, dass B nicht nur eine geordnete Basis sondern eine Orthonormalbasis ist, auf diesen Isomorphismus auswirkt. Wir zeigen, dass dies zwei konkrete Vereinfachungen mit sich bringt: Die erste ist, dass die Koeffizienten λ_i durch Skalarprodukte $\lambda_i = \langle v_i, v\rangle$ gegeben sind. Die zweite ist, dass der Isomorphismus ${}_BS : V \to \mathbb{K}^n$ in diesem Fall nicht nur ein Isomorphismus von Vektorräumen ist, sondern ein Isomorphismus von unitären oder euklidischen Räumen wird, wenn man $\mathbb{C}^n$ oder $\mathbb{R}^n$ mit dem Standardskalarprodukt ausstattet.

Satz 10.3.12: Sei $\mathbb{K} = \mathbb{C}$ oder $\mathbb{K} = \mathbb{R}$, V mit $\langle\,,\,\rangle$ ein n-dimensionaler unitärer oder euklidischer Vektorraum und $B = (v_1, ..., v_n)$ eine Orthonormalbasis von V. Dann gilt für alle $v \in V$

$$v = \Sigma_{j=1}^n \langle v_j, v\rangle\, v_j.$$

Versieht man $\mathbb{K}^n$ mit dem Standardskalarprodukt, so ist ${}_BS : V \to \mathbb{K}^n$, $v \mapsto (\langle v_1, v\rangle, ..., \langle v_n, v\rangle)^T$ ein Isomorphismus von unitären oder euklidischen Räumen.

Beweis:
Jeder Vektor $v \in V$ lässt sich eindeutig als Linearkombination $v = \Sigma_{i=1}^n \lambda_i v_i$ der Vektoren v_i darstellen, und es folgt

$$\Sigma_{j=1}^n \langle v_j, v\rangle\, v_j = \Sigma_{j=1}^n \langle v_j, \Sigma_{i=1}^n \lambda_i v_i\rangle\, v_j = \Sigma_{i,j=1}^n \lambda_i \langle v_j, v_i\rangle\, v_j = \Sigma_{i,j=1}^n \lambda_i \delta_{ji}\, v_j = \Sigma_{i=1}^n \lambda_i\, v_i = v.$$

Sind $v = \Sigma_{i=1}^n \lambda_i v_i$ und $w = \Sigma_{j=1}^n \mu_j v_j$ Vektoren in V, so gilt ${}_BS(v) = (\lambda_1, ..., \lambda_n)^T$ und ${}_BS(w) = (\mu_1, ..., \mu_n)^T$, und man erhält

$$\langle v, w\rangle = \langle \Sigma_{i=1}^n \lambda_i v_i, \Sigma_{j=1}^n \mu_j w_j\rangle = \Sigma_{i,j=1}^n \overline{\lambda_i} \mu_j \langle v_i, v_j\rangle = \Sigma_{i,j=1}^n \overline{\lambda_i} \mu_j \delta_{ij} = (\overline{\lambda}_1, ..., \overline{\lambda}_n) \cdot (\mu_1, ..., \mu_n)^T.$$

□

Nach Satz 10.3.12 ist also insbesondere für jede Orthonormalbasis $B = (v_1, ..., v_n)$ von V die Identitätsabbildung auf V gegeben durch $\mathrm{id}_V : V \to V$, $v \mapsto \Sigma_{i=1}^n \langle v_i, v\rangle\, v_i$. Wählt man nun statt einer Orthonormalbasis ein *Orthonormalsystem* $B_U = (u_1, ..., u_m)$ in V, also eine Orthogonalbasis des Untervektorraums $U = \mathrm{span}_{\mathbb{K}}\{u_1, ..., u_m\}$, so stimmt die entsprechende Abbildung für dieses Orthonormalsystem B_U offensichtlich nicht mehr mit der Identitätsabbildung überein. Stattdessen erhält man eine Projektionsabbildung mit Bild U, deren Kern senkrecht auf U steht, die sogenannte *Orthogonalprojektion* auf U.

Satz 10.3.13: Sei V mit $\langle\,,\,\rangle$ ein euklidischer oder unitärer Vektorraum, $U \subseteq V$ ein m-dimensionaler Untervektorraum und $(u_1, ..., u_m)$ eine Orthonormalbasis von U. Dann gilt:

1. Die lineare Abbildung $P_U : V \to V$, $v \mapsto \Sigma_{j=1}^m \langle u_j, v\rangle\, u_j$ ist ein Projektor mit $\mathrm{im}(P_U) = U$ und $\ker(P_U) = U^\perp = \{v \in V \mid \langle u, v\rangle = 0\, \forall u \in U\}$.

2. Für alle $v \in V$ ist $P_U(v)$ der Vektor in U mit dem kleinsten Abstand von v:

$$||v - P_U(v)|| \leq ||v - u|| \;\; \forall u \in U \quad \text{und} \quad ||v - P_U(v)|| = ||v - u|| \Leftrightarrow u = P_U(v).$$

Die Abbildung P_U heißt **Orthogonalprojektion** auf U und der Untervektorraum $U^\perp$ **orthogonales Komplement** von U.

Beweis:
$P_U : V \to V$ ist $\mathbb{K}$-linear mit $P_U(v) \in U$ für ale $v \in V$ und $P_U(u_i) = \Sigma_{j=1}^m \langle u_j, u_i \rangle u_j = u_i$ für alle $i \in \{1, ..., m\}$ wegen der Orthonormalität von $(u_1, ..., u_m)$. Daraus folgt $\mathrm{im}(P_U) = U$

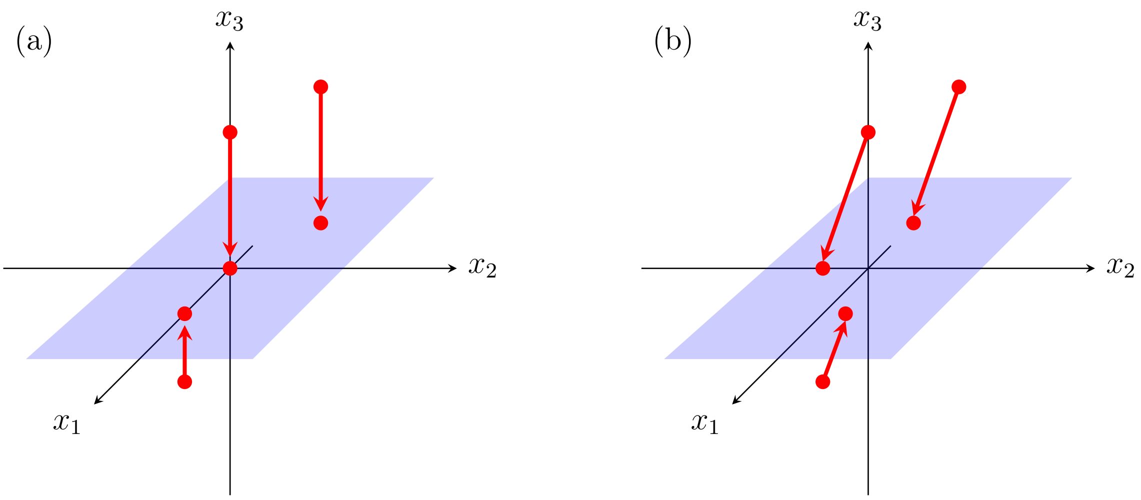
$$P_U(P_U(v)) = P_U(\Sigma_{j=1}^m \langle u_j, v \rangle\, u_j) = \Sigma_{j=1}^m \langle u_j, v \rangle P_U(u_j) = \Sigma_{j=1}^m \langle u_j, v \rangle u_j = P_U(v) \qquad \forall v \in V.$$

Also ist V ein Projektor. Wegen der linearen Unabhängigkeit von $(u_1, ..., u_m)$ ist $P_U(v) = 0$ genau dann, wenn $\langle u_j, v \rangle = 0$ für alle $j \in \{1, ..., m\}$, was wegen der Sesquilinearität von $\langle\, ,\, \rangle$ gleichbedeutend ist zu $\langle u, v \rangle = 0$ für alle $u \in U$. Damit ist $\ker(P_U) = U^\perp = \{u_1, ..., u_m\}^\perp$.

Da P_U ein Projektor ist, gilt $v - P_U(v) \in \ker(P_U) = U^\perp$ für alle $v \in V$. Damit sind für alle $v \in V$ und $u \in U$ die Vektoren $P_U(v) - u \in U$ und $v - P_U(v) \in U^\perp$ orthogonal. Mit dem Satz des Pythagoras folgt

$$||v - u||^2 = ||v - P_U(v) + P_U(v) - u||^2 = ||v - P_U(v)||^2 + ||P_U(v) - u||^2 \geq ||v - P_U(v)||^2$$

und $||v - u||^2 = ||v - P_U(v)||^2$ genau dann, wenn $||P_U(v) - u||^2 = 0$, also $u = P_U(v)$. $\square$



(a) orthogonale Projektion auf $\mathrm{span}_\mathbb{R}\{e_1, e_2\}$
$P_U : x_1e_1 + x_2e_2 + x_3 \mapsto x_1e_1 + x_2e_2$

(b) nicht-orthogonale Projektion auf $\mathrm{span}_\mathbb{R}\{e_1, e_2\}$
$P' : x_1e_1 + x_2e_2 + x_3 \mapsto x_1e_1 + (x_2 - \frac{1}{3}x_3)e_2$

Beispiel 10.3.14: Wir betrachten den $\mathbb{R}^3$ mit dem Standardskalarprodukt.

1. Die Orthogonalprojektion auf die x_1x_2-Ebene $U = \mathrm{span}_\mathbb{R}\{e_1, e_2\}$ ist nach Satz 10.3.13

$$P_U : \mathbb{R}^3 \to \mathbb{R}^3, \quad x = x_1e_1 + x_2e_2 + x_3e_3 \mapsto \langle e_1, x \rangle e_1 + \langle e_2, x \rangle e_2 = x_1e_1 + x_2e_2$$

Dabei handelt es sich um den Projektor, der einen Punkt $x \in \mathbb{R}^3$ parallel zur x_3-Achse in die x_1x_2-Ebene projiziert. Es gilt $\ker(P_U) = \mathbb{R}e_3 = \{e_1, e_2\}^\perp = \mathrm{im}(P_U)^\perp$. Der Punkt in U mit dem minimalen Abstand von x ist $P_U(x) = x_1e_1 + x_2e_2$, denn es gilt für $u = u_1e_1 + u_2e_2$

$$||x - u||^2 = (x_1 - u_1)^2 + (x_2 - u_2)^2 + x_3^2 \geq x_3^2 = ||x - P_U(x)||^2$$

2. Die lineare Abbildung
$$P' : \mathbb{R}^3 \to \mathbb{R}^3, \quad x = x_1e_1 + x_2e_2 + x_3e_3 \mapsto x_1e_1 + (x_2 - x_3)e_2$$
erfüllt $P' \circ P' = P'$ und ist damit auch ein Projektor mit $\mathrm{im}(P') = \mathrm{span}_{\mathbb{K}}\{e_1, e_2\} = U$. Er projiziert einen Punkt $x \in \mathbb{R}^3$ entlang der Geraden $\mathbb{K}(e_2 + e_3)$ in die x_1x_2-Ebene. Allerdings ist er nicht orthogonal, denn $\ker(P') = \mathbb{R}(e_2 + e_3) \neq U^\perp$. Der Bildpunkt $P'(x) = x_1e_1 + (x_2 - x_3)e_2$ des Punktes $x \in \mathbb{R}^3$ ist nicht der Punkt in U mit dem kleinsten Abstand von x, denn es gilt
$$||x - P'(x)|| = ||x_3(e_2 + e_3)|| = \sqrt{2}|x_3| > |x_3| = ||x_3e_3|| = ||x - P_U(x)||.$$

Allgemein erlaubt es einem eine Basiswahl in einem endlich-dimensionalen $\mathbb{K}$-Vektorraum V den Vektorraum V durch den $\mathbb{K}^n$ und seine Endomorphismen durch Matrizen zu beschreiben. Im Fall eines euklidischen oder unitären Vektorraums V kann man analog durch Wahl einer *Orthonormalbasis* den Vektorraum V durch den euklidischen Vektorraum $\mathbb{R}^n$ oder den unitären Vektorraum $\mathbb{C}^n$ mit dem *Standardskalarprodukt* beschreiben. Insbesondere erhält man so eine einfache und anschauliche Charakterisierung der unitären Endomorphismen, also der Endomorphismen, die mit dem Skalarprodukt auf V verträglich sind, und von unitären Matrizen.

Satz 10.3.15: Sei V mit $\langle\, , \rangle$ ein n-dimensionaler unitärer oder euklidischer Vektorraum. Dann ist ein Endomorphismus $\phi : V \to V$ genau dann unitär bezüglich $\langle\, , \rangle$, wenn er Orthonormalbasen auf Orthonormalbasen abbildet.

Beweis:
Nach Satz 10.2.5 ist ein Automorphismus $\phi \in \mathrm{Aut}_{\mathbb{K}}(V)$ unitär bezüglich $\langle\, , \rangle$ genau dann, wenn für eine beliebige Basis $B = (v_1, \ldots, v_n)$ gilt $\langle \phi(v_i), \phi(v_j)\rangle = \langle v_i, v_j\rangle$ für alle $i, j \in \{1, ..., n\}$. Da ein Skalarprodukt nach Bemerkung 10.1.5, 2. nicht ausgeartet ist, gilt dies nach nach Bemerkung 10.1.8, 2. auch für jeden Endomorphismus. Wählt man für B eine Orthonormalbasis, so ist diese Bedingung gleichbedeutend zu $\langle \phi(v_i), \phi(v_j)\rangle = \delta_{ij}$ für alle $i, j \in \{1, ..., n\}$, also zur Aussage, dass auch $B' = (\phi(v_1), \ldots, \phi(v_n))$ eine Orthonormalbasis ist. □

Satz 10.3.16: (**Charakterisierung unitärer Matrizen**)
Sei $\mathbb{K} = \mathbb{C}$ oder $\mathbb{K} = \mathbb{R}$. Eine Matrix $U \in \mathrm{Mat}(n \times n, \mathbb{K})$ ist genau dann unitär (orthogonal), wenn eine der folgenden äquivalenten Bedingungen gilt:

(i) Der Endomorphismus $\phi_U : \mathbb{K}^n \to \mathbb{K}^n$, $v \mapsto U \cdot v$ ist unitär (orthogonal) bezüglich des Standardskalarprodukts $\langle\, , \rangle$ auf dem $\mathbb{K}^n$.

(ii) Die Spaltenvektoren von U bilden eine Orthonormalbasis.

(iii) Die Zeilenvektoren von U bilden eine Orthonormalbasis.

Beweis:
Wir beweisen die Aussagen für $\mathbb{K} = \mathbb{C}$. Der Beweis für $\mathbb{K} = \mathbb{R}$ ist analog und ergibt sich durch Weglassen der komplexen Konjugation. Dass Aussage (i) äquivalent zur Unitarität von U ist, wurde bereits in Korollar 10.2.6 bewiesen. Zum Beweis von (ii) berechnet man für eine Matrix $U = (u_1, ..., u_n) \in \mathrm{Mat}(n \times n, \mathbb{C})$ mit Spaltenvektoren $u_1, ..., u_n \in \mathbb{C}^n$
$$U^\dagger \cdot U = \begin{pmatrix} \overline{u}_1^T \\ \vdots \\ \overline{u}_n^T \end{pmatrix} \cdot (u_1, ..., u_n) = \begin{pmatrix} \overline{u}_1^T \cdot u_1 & \ldots & \overline{u}_1^T \cdot u_n \\ \vdots & & \vdots \\ \overline{u}_1^T \cdot u_1 & \ldots & \overline{u}_1^T \cdot u_n \end{pmatrix} = \begin{pmatrix} \langle u_1, u_1\rangle & \ldots & \langle u_1, u_n\rangle \\ \vdots & & \vdots \\ \langle u_n, u_1\rangle & \ldots & \langle u_n, u_n\rangle \end{pmatrix},$$

wobei $\langle\,,\rangle$ das Standardskalarprodukt auf dem $\mathbb{C}^n$ bezeichnet. Also gilt $U^\dagger \cdot U = \mathbb{1}_n$ genau dann, wenn $\langle u_i, u_j\rangle = \delta_{ij}$ für alle $i, j \in \{1, ..., n\}$ ist, also genau dann, wenn $(u_1, ..., u_n)$ eine Orthonormalbasis des $\mathbb{C}^n$ mit dem Standardskalarprodukt $\langle\,,\rangle$ ist. Zum Beweis von (iii) benutzt man, dass die Zeilenvektoren von U die Spaltenvektoren von U^T sind und

$$U \text{ unitär} \Leftrightarrow U^\dagger \cdot U = \mathbb{1}_n \Leftrightarrow U \cdot U^\dagger = \mathbb{1}_n \Leftrightarrow \overline{U} \cdot U^T = \mathbb{1}_n \Leftrightarrow (U^T)^\dagger \cdot U^T = \mathbb{1}_n \Leftrightarrow U^T \text{ unitär.}$$

□

Beispiel 10.3.17: Wir betrachten den $\mathbb{R}^n$ mit dem Standardskalarprodukt.

1. Die Spiegelung $S_j : \mathbb{R}^n \to \mathbb{R}^n$ an der Hyperebene senkrecht zur x_j-Achse ist unitär bezüglich $\langle\ ,\ \rangle$. Denn aus $S_j(e_j) = -e_j$ und $S_j(e_k) = e_k$ für alle $k \neq j$ folgt $\langle S(e_i), S(e_k)\rangle = \delta_{ik}$ für alle $i, k \in \{1, ..., n\}$.

2. Die Drehung $D_\alpha : \mathbb{R}^n \to \mathbb{R}^n$ um den Winkel α in der $x_i x_j$-Ebene für $i < j$ ist unitär bezüglich $\langle\ ,\ \rangle$, denn

   $$D_\alpha(e_i) = \cos(\alpha)e_i + \sin(\alpha)e_j, \quad D_\alpha(e_j) = \cos(\alpha)e_j - \sin(\alpha)e_i, \quad D_\alpha(e_k) = e_k \;\; \forall k \notin \{i, j\}.$$

   Daraus ergibt sich $\langle D_\alpha(e_k), D_\alpha(e_l)\rangle = \langle e_k, e_l\rangle = \delta_{kl}$ für alle $k, l \notin \{i, j\}$ und

   $$\begin{aligned}
   \langle D_\alpha(e_i), D_\alpha(e_i)\rangle &= \langle \cos(\alpha)e_i + \sin(\alpha)e_j, \cos(\alpha)e_i + \sin(\alpha)e_j\rangle \\
   &= \cos(\alpha)^2\langle e_i, e_i\rangle + 2\sin(\alpha)\cos(\alpha)\langle e_i, e_j\rangle + \sin(\alpha)^2\langle e_j, e_j\rangle \\
   &= \cos(\alpha)^2 + \sin(\alpha)^2 = 1 \\
   \langle D_\alpha(e_j), D_\alpha(e_j)\rangle &= \langle \cos(\alpha)e_j - \sin(\alpha)e_j, \cos(\alpha)e_j - \sin(\alpha)e_i\rangle \\
   &= \cos(\alpha)^2\langle e_j, e_j\rangle - 2\sin(\alpha)\cos(\alpha)\langle e_i, e_j\rangle + \sin(\alpha)^2\langle e_i, e_i\rangle \\
   &= \cos(\alpha)^2 + \sin(\alpha)^2 = 1 \\
   \langle D_\alpha(e_i), D_\alpha(e_j)\rangle &= \langle \cos(\alpha)e_i + \sin(\alpha)e_j, \cos(\alpha)e_j - \sin(\alpha)e_i\rangle \\
   &= \cos(\alpha)^2\langle e_i, e_j\rangle + \sin(\alpha)\cos(\alpha)(\langle e_j, e_j\rangle - \langle e_i, e_i\rangle) - \sin(\alpha)^2\langle e_i, e_j\rangle = 0 \\
   \langle D_\alpha(e_j), D_\alpha(e_k)\rangle &= \langle D_\alpha(e_i), D_\alpha(e_k)\rangle = 0
   \end{aligned}$$

Allgemein ist eine lineare Abbildung $\phi : \mathbb{R}^n \to \mathbb{R}^n$ genau dann unitär bezüglich des Standardskalarprodukts, wenn sie alle Längen und Winkel erhält. Denn Längen und Winkel sind durch die Skalarprodukte $\langle v, w\rangle$ von Vektoren $v, w \in V$ gegeben. Umgekehrt lässt sich jedes Skalarprodukt durch Längen und Winkel ausdrücken und wird damit erhalten, wenn Längen und Winkel erhalten werden. Für unitäre Vektorräume erhält man keine so offensichtliche geometrische Interpretation von unitären Abbildungen. Sie besitzen jedoch wichtige Anwendungen. In der Quantenmechanik sind Meßwahrscheinlichkeiten durch Skalarprodukte gegeben und unitäre Abbildungen beschreiben die Zeitentwicklung von quantenmechanischen Systemen.

10.4 Orthogonalität von Vektoren und Untervektorräumen

Wir beschäftigen uns nun mit der Existenz von Orthogonalbasen und Orthogonalsystemen für allgemeinere symmetrische Bilinearformen und hermitesche Sesquilinearformen. In diesem Fall erhält man auch für $\mathbb{K} = \mathbb{C}$ oder $\mathbb{K} = \mathbb{R}$ weder einen Begriff von Länge noch einen Begriff von Winkel. Denn es kann Vektoren $v \in V \setminus \{0\}$ mit $\langle v, v\rangle \leq 0$ geben und Vektoren $v, w \in V \setminus \{0\}$ mit $|\langle v, w\rangle|^2 > \langle v, v\rangle\langle w, w\rangle$. Man erhält aber weiterhin ein Konzept von *Orthogonalität* und *orthogonalem Komplement*, da die Bedingung $\beta(v, w) = 0$ für Bilinearformen β über beliebigen Körpern $\mathbb{K}$ und jede Sesquilinearform β formuliert werden kann.

Definition 10.4.1: Sei V ein $\mathbb{K}$-Vektorraum und β eine Bi- oder Sesquilinearform auf V. Ein Vektor $v \in V$ heißt **orthogonal** zu einem Vektor $w \in V$ bezüglich β, wenn $\beta(v, w) = 0$ gilt. Der **Orthogonalraum** oder das **orthogonale Komplement** einer Teilmenge $M \subseteq V$ ist

$$M^\perp = \{v \in V | \; \beta(m, v) = 0 \; \forall m \in M\}.$$

Der Orthogonalitätsbegriff aus Definition 10.4.1 stimmt für allgemeine Bi- oder Sesquilinearformen nicht mehr unbedingt mit der geometrischen Anschauung überein, und er kann sich deutlich von Orthogonalität in euklidischen oder unitären Vektorräumen unterscheiden. So können etwa Vektoren existieren, die orthogonal zu sich selbst sind, und Untervektorräume können ihre orthogonalen Komplemente nicht-trivial schneiden. Daher müssen wir uns zunächst systematisch mit den Eigenschaften des orthogonalen Komplements befassen.

Satz 10.4.2: (**Eigenschaften des orthogonalen Komplements**)
Sei V ein $\mathbb{K}$-Vektorraum und β eine Bilinearform oder Sesquilinearform auf V. Dann gilt:

1. Für alle Teilmengen $M \subseteq V$ ist $M^\perp = \text{span}_\mathbb{K}(M)^\perp$ ein Untervektorraum von V.
2. Ist β **orthosymmetrisch**, d. h. $\beta(v, w) = 0 \Leftrightarrow \beta(w, v) = 0$, so gilt $M \subseteq (M^\perp)^\perp$.
3. Für jeden Untervektorraum $U \subseteq V$ gilt: $\beta|_U$ nicht ausgeartet $\Leftrightarrow$ $U \cap U^\perp = \{0\}$. Insbesondere: β nicht ausgeartet $\Leftrightarrow V^\perp = \{0\}$.
4. Ist $U \subseteq V$ endlich-dimensional und $\beta|_U$ nicht ausgeartet, so gilt $V = U \oplus U^\perp$.
5. Ist β ein Skalarprodukt, so gilt $V = U \oplus U^\perp$ und $(U^\perp)^\perp = U$ für alle endlich-dimensionalen Untervektorräume $U \subseteq V$.

Beweis:
1. Da $M \subseteq \text{span}_\mathbb{K}(M)$ folgt aus $v \in \text{span}_\mathbb{K}(M)^\perp$ auch $v \in M^\perp$. Ist umgekehrt $v \in M^\perp$, so folgt $\beta(\Sigma_{j=1}^n \lambda_j m_j, v) = \Sigma_{j=1}^n \overline{\lambda_j} \beta(m_j, v) = 0$ für alle $n \in \mathbb{N}$, $m_1, ..., m_n \in M$ und $\lambda_1, ..., \lambda_n \in \mathbb{K}$, also $v \in \text{span}_\mathbb{K}(M)^\perp$. Damit ist $M^\perp = \text{span}_k(M)^\perp$. Sind $v, w \in M^\perp$, so ergibt sich für alle $m \in M$ $\beta(m, v + w) = \beta(m, v) + \beta(m, w) = 0$, also $v + w \in M^\perp$ und $\beta(m, \lambda v) = \lambda\beta(m, v) = 0$ für alle $\lambda \in \mathbb{K}$, also $\lambda v \in M^\perp$. Damit ist $M^\perp \subseteq V$ ein Untervektorraum.

2. und 3. folgen direkt aus der Definition. Ist nämlich $m \in M$ und β orthosymmetrisch, so folgt $\beta(m, v) = \beta(v, m) = 0$ für alle $v \in M^\perp$ und damit $m \in (M^\perp)^\perp$. Außerdem gilt für jeden Untervektorraum $U \subseteq V$ per Definition $U \cap U^\perp = \{u \in U \mid \beta(v, u) = 0 \; \forall v \in U\}$.

4. Da $\beta|_U$ nicht ausgeartet ist, folgt $U \cap U^\perp = \{0\}$ mit 3. Zu zeigen ist, dass $U + U^\perp = V$. Wir betrachten die lineare Abbildung $\phi : V \to V$, $v \mapsto \Sigma_{i=1}^n \beta(u_i, v) \, u_i$ für eine Basis $(u_1, ..., u_n)$ von U. Dann gilt $\ker(\phi) = \{u_1, ..., u_n\}^\perp = U^\perp$ wegen der linearen Unabhängigkeit von $(u_1, ..., u_n)$. Da $\phi(U) \subseteq U$ und $U \cap U^\perp = \{0\}$, ist $\phi|_U : U \to U$ injektiv und, da U endlich-dimensional ist, bijektiv. Also ist $\text{im}(\phi) = \phi(V) = \phi(U) = U$. Zu jedem $v \in V$ gibt es damit ein $u \in U$ mit $\phi(v) = \phi(u)$, und damit ist $v - u \in \ker(\phi) = U^\perp$ und $v = u + v - u \in U + U^\perp$, also $V = U + U^\perp$.

5. Ist β ein Skalarprodukt, so ist β orthosymmetrisch und $\beta|_W$ nicht ausgeartet für alle Unterräume $W \subseteq V$ nach Bemerkung 10.1.5. Ist $U \subseteq V$ ein endlich-dimensionaler Unterraum, so folgt $V = U \oplus U^\perp$ mit 4. und $U \subseteq (U^\perp)^\perp$ mit 2. Also ist $(U^\perp)^\perp = U \oplus (U^\perp \cap (U^\perp)^\perp)$, aber da auch $\beta|_{U^\perp}$ nicht ausgeartet ist, folgt mit 3. $U^\perp \cap (U^\perp)^\perp = \{0\}$ und damit $(U^\perp)^\perp = U$. $\square$

Beispiel 10.4.3:

1. Im $\mathbb{R}^n$ mit dem Standardskalarprodukt $\langle x,y\rangle = \Sigma_{i=1}^n x_i\, y_i$ gilt für alle $k \in \{1,...,n\}$

$$\{e_k\}^\perp = (\mathbb{R}e_k)^\perp = \{x \in \mathbb{R}^n | \, \langle e_k, x\rangle = x_1 = 0\} = \mathrm{span}_\mathbb{R}\{e_1,...,e_{k-1},e_{k+1},...e_n\}.$$

   Das orthogonale Komplement der x_k-Achse ist also die von den x_j-Achsen mit $j \neq k$ aufgespannte Hyperebene. Umgekehrt ist nach Satz 10.4.2, 5. das orthogonale Komplement dieser Hyperbene die x_k-Achse.

2. Ist β eine alternierende Bilinearform, so gilt $\mathbb{K}v \subseteq (\mathbb{K}v)^\perp$ für alle $v \in V$. Jeder Vektor $v \in V$ ist also orthogonal zu sich selbst.

3. Der Entartungsraum $V_{deg} = \{v \in V | \beta(w,v) = 0 \; \forall w \in V\}$ einer Bi-oder Sesquilinearform β auf V ist für alle Teilmengen $M \subseteq V$ im Orthogonalraum $M^\perp$ enthalten.

4. Wir betrachten für $n \geq 2$ den $\mathbb{R}^n$ mit der Minkowski-Metrik $\eta(x,y) = -x_1y_1 + \Sigma_{i=2}^n x_i\, y_i$. Dann sind die orthogonalen Komplemente der Geraden $g = \mathbb{K}e_1$, $h = \mathbb{K}(e_1+e_2)$ und $k = \mathbb{K}e_2$ gegeben durch

$$g^\perp = \mathrm{span}_\mathbb{R}\{e_2,...,e_n\}, \quad h^\perp = \mathrm{span}_\mathbb{R}\{e_1+e_2, e_3,...,e_n\}, \quad k^\perp = \mathrm{span}_\mathbb{R}\{e_1,e_3,...,e_n\}.$$

Wir erweitern nun unsere Suche nach Orthogonalbasen auf allgemeine symmetrische Bi- und hermitesche Sesquilinearformen β, die nicht unbedingt Skalarprodukte sein müssen. Ein wichtiger Grund, Orthogonalbasen zu suchen und anschließend die beschreibenden Diagonalmatrizen möglichst weit zu vereinfachen, ist die *Klassifikation* von Bi- oder Sesquilinearformen.

Ähnlich wie im Fall der beschreibenden Matrizen von Endomorphismen möchte man eine Liste von beschreibenden Matrizen von symmetrischen Bi- oder hermiteschen Sesquilinearformen finden, so dass nach einer geschickten Basiswahl jede solche Bi- oder Sesquilinearform durch genau eine Matrix aus dieser Liste gegeben ist. Damit hätte man dann eine vollständige Liste aller symmetrischen Bi- oder hermiteschen Sesquilinearformen auf endlich-dimensionalen Vektorräumen, bis auf Isomorphie. Die Matrizen in dieser Liste nennt man *Normalformen*.

Mit den Vorarbeiten zu orthogonalen Komplementen aus Satz 10.4.2 können wir jetzt für beliebige symmetrische Bilinearformen oder hermitesche Sesquilinearformen auf endlich-dimensionalen Vektorräumen direkt eine Normalform angeben.

Satz 10.4.4: **(Normalformen)**
Sei $\mathbb{K}$ ein Körper und V ein n-dimensionaler $\mathbb{K}$-Vektorraum.

1. Ist $\mathrm{char}(\mathbb{K}) \neq 2$, so existiert für jede symmetrische Bilinearform β auf V eine Orthogonalbasis, d. h. eine geordnete Basis B von V mit

$$_B\beta_B = \begin{pmatrix} \beta_{11} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \beta_{nn} \end{pmatrix}.$$

   Ist $\mathbb{K}$ algebraisch abgeschlossen (z. B. $\mathbb{K} = \mathbb{C}$), so kann man $\beta_{ii} \in \{0,1\}$ fordern.

2. Ist $\mathbb{K} = \mathbb{C}$ oder $\mathbb{K} = \mathbb{R}$ und β eine eine hermitesche Sesquilinearform auf V, so gibt es eine Orthogonalbasis B von V und $p, q \in \mathbb{N}_0$ mit $p + q \in \{0, 1, ..., n\}$, so dass

$$^B\beta^B = \begin{pmatrix} \mathbb{1}_p & 0 & 0 \\ 0 & -\mathbb{1}_q & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Beweis:
Ist $\beta = 0$ ist nichts zu zeigen. Für $\beta \neq 0$ ist der Beweis per Induktion über $n = \dim_{\mathbb{K}}(V)$.

$n = 1$: Dann gilt $V = \mathbb{K}u$ mit $u \in V$ und $\beta(u, u) \neq 0$.

• Ist $\mathbb{K}$ algebraisch abgeschlossen und β eine symmetrische Bilinearform, so gibt es ein $\lambda \in \mathbb{K} \setminus \{0\}$ mit $\lambda^2 = \beta(u, u) \neq 0$, denn die Gleichung $\lambda^2 - \beta(u, u) = 0$ hat eine Lösung in $\mathbb{K}$. Für $v := \lambda^{-1}u$ folgt dann $\beta(v, v) = \lambda^{-2}\beta(u, u) = 1$.

• Ist $\mathbb{K} = \mathbb{C}$ oder $\mathbb{K} = \mathbb{R}$ und β eine hermitesche Sesquilinearform auf V, so gilt für jedes $u \in V$ $\beta(u, u) = \overline{\beta(u, u)} \in \mathbb{R}$, und es gibt ein $\lambda \in \mathbb{R}$ mit $\lambda^2 = |\beta(u, u)| > 0$. Für $v = \lambda^{-1}u$ folgt $\beta(v, v) = |\lambda|^{-2}\beta(u, u) \in \{\pm 1\}$.

$(n - 1) \to n$: Seien 1. und 2. bewiesen für alle $\mathbb{K}$-Vektorräume der Dimension $\leq n - 1$ und sei V ein $\mathbb{K}$-Vektorraum der Dimension n.

Sei $\beta \neq 0$ eine symmetrische Bilinearform oder eine hermitesche Sesquilinearform auf V. Dann gibt es Vektoren $x, y \in V$ mit $\beta(x, y) \neq 0$. Da $\mathrm{char}(\mathbb{K}) \neq 2$ gelten die Polarisierungsidentitäten

$$\beta(x, y) = \tfrac{1}{2}\beta(x + y, x + y) - \tfrac{1}{2}\beta(x, x) - \tfrac{1}{2}\beta(y, y),$$
$$\beta(x, y) = \tfrac{1}{4}\beta(x + y, x + y) - \tfrac{1}{4}\beta(x - y, x - y) - \tfrac{\mathrm{i}}{4}\beta(x + \mathrm{i}y, x + \mathrm{i}y) + \tfrac{\mathrm{i}}{4}\beta(x - \mathrm{i}y, x - \mathrm{i}y)$$

Wäre $\beta(v, v) = 0$ für alle $v \in V$, so wäre die rechte und damit auch die linke Seite dieser Gleichungen für alle $x, y \in V$ gleich Null, ein Widerspruch zu $\beta(x, y) \neq 0$. Also gibt es einen Vektor $v \in V$ mit $\beta(v, v) \neq 0$.

Wie im Fall $n = 1$ können wir $v \in V$ so wählen, dass $\beta(v, v) = 1$, falls β eine symmetrische Bilinearform und $\mathbb{K}$ algebraisch abgeschlossen ist, und so, dass $\beta(v, v) \in \{\pm 1\}$ falls $\mathbb{K} = \mathbb{C}$ oder $\mathbb{K} = \mathbb{R}$ und β eine hermitesche Sesquilinearform ist.

Da wegen $\beta(v, v) \neq 0$ die Einschränkung $\beta|_{\mathbb{K}v}$ nicht ausgeartet ist, gilt $V = \mathbb{K}v \oplus (\mathbb{K}v)^\perp$ nach Satz 10.4.2, 4. Da $(\mathbb{K}v)^\perp$ ein Untervektorraum der Dimension $n - 1$ ist, existiert eine Orthogonalbasis $C = (b_2, ..., b_n)$ von $(\mathbb{K}v)^\perp$, die die Bedingungen in 1. und 2. erfüllt. Damit ist $B = (v, b_2, ..., b_n)$ eine Orthogonalbasis von V, so dass die darstellende Matrix ${}_BM_B(\phi)$ bis auf Permutationen der Basisvektoren die angegebene Form hat. □

Am Beweis von Satz 10.4.4 sieht man, dass die Normierbarkeit von Vektoren in einer Orthogonalbasis vom zugrundeliegenden Körper abhängt. Sie entspricht nämlich dem Lösen der quadratischen Gleichung $\lambda^2 = c$ für eine Konstante $c \in \mathbb{K}$. Eine solche Gleichung besitzt in $\mathbb{C}$ immer mindestens eine Lösung. In $\mathbb{R}$ besitzt sie eine Lösung, wenn $c > 0$ und in den Körpern $\mathbb{Z}/p\mathbb{Z}$ muss keine Lösung existieren. So hat in $\mathbb{Z}/3\mathbb{Z}$ die Gleichung $\lambda^2 = [2]$ keine Lösung.

Außerdem wird im Beweis von Satz 10.4.4 deutlich, warum das Gram-Schmidt-Verfahren für allgemeine hermitesche Sesquilinearformen oder symmetrische Bilinearformen β nicht analog funktioniert. In diesem Fall kann nämlich eine linear unabhängige Teilmenge von V Vektoren $v \in V$ mit $\beta(v, v) = 0$ enthalten. Der Beweis von Satz 10.4.4 besteht im Wesentlichen daraus, mit der Polarisierungsidentität zu zeigen, dass man immer einen Vektor $v \in V$ mit $\beta(v, v) \neq 0$ finden kann. Das erübrigt sich für Skalarprodukte, die $\beta(v, v) \neq 0$ für alle $v \in V \setminus \{0\}$ erfüllen.

Satz 10.4.4 zeigt, dass zu jeder symmetrischen Bilinearform auf einem reellen oder hermiteschen Sesquilinearform auf einem komplexen n-dimensionalen Vektorraum V eine Orthogonalbasis existiert. Die beschreibenden Matrizen bezüglich dieser Orthogonalbasis können sich dann nur noch durch die Anzahl der Einträge 1, -1 und 0 in der Diagonalen unterscheiden. Nun stellt sich die Frage, ob die Anzahl der Einträge 0,1,-1 durch die Bi- oder Sesquilinearform festgelegt ist, oder ob sich durch geschickte Basiswahlen noch weitere Vereinfachungen erzielen lassen.

Satz 10.4.5: (**Trägheitssatz von Sylvester**)

1. Die Zahl der Diagonaleinträge $\beta_{ii} \neq 0$ in Satz 10.4.4, 1. hängt nicht von der Basis ab.
2. Ist $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$, so hängen die Zahlen $p, q \in \mathbb{N}_0$ in Satz 10.4.4, 2. nicht von der Wahl der Basis ab. Das Paar (p, q) heißt **Signatur** von β und $p{-}q$ **Trägheitsindex** von β.

Beweis:
1. Die Anzahl der Diagonaleinträge mit $\beta_{ii} \neq 0$ in Satz 10.4.4, 1. ist gleich dem Rang $\mathrm{rg}({}^B\beta^B)$. Nach Satz 10.2.4 gilt ${}^B\beta^B = {}_CM_B(\mathrm{id}_V)^{T,\dagger} \cdot {}^C\beta^C \cdot {}_CM_B(\mathrm{id}_V)$ für jede geordnete Basis C von V. Da Multiplikation mit einer invertierbaren Matrix den Rang einer Matrix nicht ändert und ${}_CM_B(\mathrm{id}_V)$, ${}_CM_B(\mathrm{id}_V)^T$ invertierbar sind, folgt $\mathrm{rg}({}^C\beta^C) = \mathrm{rg}({}^B\beta^B)$.

2. Die Zahl $p \in \mathbb{N}_0$ in Satz 10.4.4, 2. ist die maximale Dimension eines Untervektorraums $U_+ \subseteq V$, so dass $\beta|_{U_+}$ positiv definit ist. Analog ergibt sich, dass die Zahl $q \in \mathbb{N}_0$ in Satz 10.4.4, 2. die maximale Dimension eines Untervektorraums $U_- \subseteq V$ ist, so dass $\beta|_{U_-}$ negativ definit ist, und $n = p - q$ die maximale Dimension eines Untervektorraums $U_0 \subseteq V$, so dass $\beta|_{U_0} = 0$ ist. Damit hängen diese Zahlen nicht von der Wahl der Basis ab. □

Korollar 10.4.6: Sei V ein endlich-dimensionaler Vektorraum über $\mathbb{K}$.

1. Ist $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$, so ist jede hermitesche Sesquilinearform auf V durch die maximalen Dimensionen von Untervektorräumen $U_0, U_+, U_- \subseteq V$ mit $\beta|_{U_+}$ positiv definit $(-\beta)|_{U_-}$ positiv definit und $\beta|_{U_0} = 0$ bis auf Basiswahlen eindeutig bestimmt.
2. Ist $\mathbb{K} = \mathbb{C}$, so ist jede symmetrische Bilinearform auf V durch die maximale Dimension eines Untervektorraums $U_0 \subseteq V$ mit $\beta|_{U_0} = 0$ bis auf Basiswahlen eindeutig bestimmt.

An was man sich erinnern sollte:

Die wichtigsten Begriffe und Definitionen:

- Bilinearform, Sesquilinearform,
- (anti)hermitesch, (anti)symmetrisch, alternierend, ausgeartet,
- positiv (semi)definit, Skalarprodukt,
- Verträglichkeit mit Bi- oder Sesquilinearformen,
- unitäre Gruppe bezüglich einer Bi- oder Sesquilinearform,
- orthogonale Gruppe, symplektische Gruppe,
- beschreibende Matrix von Bi- oder Sesquilinearform.
- hermitesch konjugierte Matrix, unitäre Matrix, orthogonale Matrix,
- Orthogonalsystem, Orthonormalsystem, Orthogonalbasis, Orthonormalbasis
- Norm, normiert, Metrik, quadratische Form,
- Länge von Vektor, Winkel zwischen Vektoren,
- Gram-Schmidt-Verfahren, Orthogonalprojektion,
- orthogonal, Orthogonalraum, orthogonales Komplement,

Die wichtigsten Aussagen:

- Transformationsformel für Bi - oder Sesquilinearformen,
- Beschreibung der unitären Gruppe von Bi - oder Sesquilinearformen mit Matrizen,
- Pythagoras, Cauchy-Schwartz-Ungleichung, Dreiecksungleichung,
- Orthogonalsysteme in euklidischen oder unitären Vektorräumen sind linear unabhängig,
- Orthonormalbasen und Orthogonalprojektionen aus dem Gram-Schmidt-Verfahren,
- Charakterisierung unitärer Matrizen durch Orthonormalbasen,
- Polarisierungsidentität,
- Eigenschaften des Orthogonalraums,
- Normalformen von symmetrischen Bi- oder hermiteschen Sesquilinearformen,
- Trägheitssatz von Sylvester.

Die wichtigsten Beispiele:

- Standardskalarprodukt auf $\mathbb{R}^n$ und $\mathbb{C}^n$,
- Minkowski-Metrik,
- Skalarprodukt auf $\mathrm{Mat}(n \times n, \mathbb{K})$,
- Bi- und Sesquilinearformen aus Matrizen,
- Die Gruppen $\mathrm{U}(n)$, $\mathrm{SU}(n)$, $\mathrm{O}(n)$, $\mathrm{SO}(n)$

11 Normale Endomorphismen

11.1 Die adjungierte Abbildung

In diesem Abschnitt betrachten wir Bi- bzw. Sesquilinearformen unter einem anderen Gesichtspunkt, nämlich als Strukturen, die eine Beziehung zwischen einem Vektorraum und seinem Dualraum herstellen.

Der Dualraum V^* eines $\mathbb{K}$-Vektorraums V wurde in Kapitel 4.4 definiert als der Vektorraum der Linearformen auf V, also der Vektorraum der linearen Abbildungen $\alpha : V \to \mathbb{K}$ mit der punktweisen Addition und Skalarmultiplikation. Da der Dualraum V^* eines endlich-dimensionalen $\mathbb{K}$-Vektorraums V nach Bemerkung 4.4.4 die gleiche Dimension hat wie V, sind V und V^* in diesem Fall isomorph. Allerdings sind sie nicht *kanonisch isomorph*. Das bedeutet, dass man *Strukturen wählen* muss, um einen Isomorphismus $\phi : V \to V^*$ zu erhalten, und dass auf diese Weise viele unterschiedliche Isomorphismen $\phi : V \to V^*$ entstehen.

In Kapitel 4.4 wurden ein endlich-dimensionaler Vektorraum V und sein Dualraum V^* in Beziehung gesetzt, indem Basen von V gewählt und die dazu dualen Basen von V^* betrachtet wurden. Wir werden nun sehen, dass auch eine nicht ausgeartete orthosymmetrische Bilinearform $\beta : V \times V \to \mathbb{K}$ auf V einen Isomorphismus zwischen V und V^* definiert. Man erhält damit nämlich zu jedem Vektor $v \in V$ eine Linearform $v^\flat : V \to \mathbb{K}$, $w \mapsto \beta(v, w)$, und jedes Element von V^* ist von dieser Form. Für Sesquilinearformen funktioniert das analog, aber die Zuordnung $v \to v^\flat$ ist nur noch linear bis auf komplexe Konjugation.

Satz 11.1.1: (**Darstellungssatz von Fréchet-Riesz**)
Sei V ein n-dimensionaler $\mathbb{K}$-Vektorraum und β eine nicht ausgeartete orthosymmetrische Bi- oder Sesquilinearform auf V. Dann gilt:

1. Jeder Vektor $v \in V$ definiert eine Linearform $v^\flat : V \to \mathbb{K}$, $w \mapsto \beta(v, w)$, und zu jeder Linearform $\alpha : V \to \mathbb{K}$ gibt es genau einen Vektor $\alpha^\sharp \in V$ mit $(\alpha^\sharp)^\flat = \alpha$.
2. Für eine Bilinearform β definiert dies zueinander inverse lineare Abbildungen $\flat : V \to V^*$, $v \mapsto v^\flat$ und $\sharp : V^* \to V$, $\alpha \mapsto \alpha^\sharp$, die **musikalischen Isomorphismen**.
3. Für eine Sesquilinearform β definiert dies zueinander inverse **semilineare Abbildungen** $\flat : V \to V^*$, $v \mapsto v^\flat$ und $\sharp : V^* \to V$, $\alpha \mapsto \alpha^\sharp$. Für alle $v, w \in V$, $\alpha, \beta \in V^*$, $\lambda \in \mathbb{C}$ gilt
$$(v + w)^\flat = v^\flat + w^\flat, \quad (\alpha + \beta)^\sharp = \alpha^\sharp + \beta^\sharp, \quad (\lambda v)^\flat = \overline{\lambda} v^\flat, \quad (\lambda\alpha)^\sharp = \overline{\lambda}\alpha^\sharp.$$

Beweis:
Wegen der Bilinearität von β ist für jeden Vektor $v \in V$ die Abbildung $v^\flat : V \to \mathbb{K}$, $w \mapsto \beta(v, w)$ eine Linearform auf V, denn es gilt für alle $x, y \in V$ und $\lambda \in \mathbb{K}$

$$v^\flat(x+y) = \beta(v, x+y) = \beta(v, x)+\beta(v, y) = v^\flat(x)+v^\flat(y) \quad v^\flat(\lambda x) = \beta(v, \lambda x) = \lambda\beta(v, x) = \lambda v^\flat(x).$$

Wir erhalten also eine Abbildung $\flat : V \to V^*$, $v \mapsto v^\flat$. Da β nicht ausgeartet und orthosymmetrisch ist, erfüllt diese

$$\flat^{-1}(\{0\}) = \{v \in V | v^\flat = 0\} = \{v \in V | \beta(v, w) = 0 \,\forall w \in V\} = \{v \in V | \beta(w, v) = 0 \,\forall w \in V\} = \{0\}.$$

Ist β bilinear, so ist die Abbildung $\flat : V \to V^*$ linear, denn für alle $v, w, x \in V$ und $\lambda \in \mathbb{K}$ gilt

$$(v + w)^\flat(x) = \beta(v + w, x) = \beta(v, x) + \beta(w, x) = (v^\flat + w^\flat)(x)$$
$$(\lambda v)^\flat(x) = \beta(\lambda v, x) = \lambda\beta(v, x) = \lambda v^\flat(x).$$

Ist β sesquilinear, so ist $\flat$ semilinear, denn für alle $v, w, x \in V$ und $\lambda \in \mathbb{K}$ gilt

$$(v+w)^\flat(x) = \beta(v+w, x) = \beta(v,x) + \beta(w,x) = (v^\flat + w^\flat)(x)$$
$$(\lambda v)^\flat(x) = \beta(\lambda v, x) = \overline{\lambda}\beta(v,x) = \overline{\lambda} v^\flat(x).$$

Auch für semilineare Abbildungen $\phi : V \to W$ zwischen zwei komplexen Vektorräumen V, W kann man zeigen (Übung), dass $\ker(\phi) = \phi^{-1}(\{0\}) \subseteq V$ und $\mathrm{im}(\phi) = \phi(V) \subseteq W$ Untervektorräume sind und $\dim_{\mathbb{K}} \ker(\phi) + \dim_{\mathbb{K}} \mathrm{im}(\phi) = \dim_{\mathbb{K}}(V)$ gilt.

Damit ist in beiden Fällen die Abbildung $\flat$ bijektiv, denn wegen $\dim_{\mathbb{K}}(V) = \dim_{\mathbb{K}}(V^*) = n$ folgt mit der Dimensionsformel $\dim_{\mathbb{K}} \mathrm{im}(\flat) = n - \dim_{\mathbb{K}} \ker(\flat) = n = \dim_{\mathbb{K}} V = \dim_{\mathbb{K}} V^*$. Bezeichnen wir mit $\sharp : V^* \to V$, $\alpha \mapsto \alpha^\sharp$ ihre Umkehrabbildung, so ist diese linear für jede Bilinearform β und semilinear für jede Sesquilinearform β. □

Mit der durch eine Bi- oder Sesquilinearform β auf V definierten Abbildung $\sharp : V^* \to V$ kann man nun insbesondere Endomorphismen $\phi : V \to V$ mit den zugehörigen *dualen Endomorphismen* $\phi^* : V^* \to V^*$, $\alpha \mapsto \alpha \circ \phi$ vergleichen. Dazu konjugieren wir letztere mit $\sharp : V^* \to V$ und erhalten so einen Endomorphismus $\phi^\dagger = \sharp \circ \phi^* \circ \flat : V \to V$, die *adjungierte Abbildung*.

Dies funktioniert für jede nicht ausgeartete orthosymmetrische Bi- oder Sesquilinearform auf V. Allerdings werden wir diese Konstruktion im Folgenden nur für den Fall einer nicht ausgearteten symmetrischen Bilinearform oder hermiteschen Sesquilinearform benötigen und beschränken uns daher auf diese Fälle.

Satz 11.1.2: Sei V ein endlich-dimensionaler $\mathbb{K}$-Vektorraum und β eine nicht ausgeartete hermitesche Sesquilinearform oder eine nicht ausgeartete symmetrische Bilinearform auf V. Dann existiert zu jedem Endomorphismus $\phi : V \to V$ genau eine Abbildung $\phi^\dagger : V \to V$ mit

$$\beta(v, \phi(w)) = \beta(\phi^\dagger(v), w) \qquad \forall v, w \in V. \tag{23}$$

Diese ist ein Vektorraumendomorphismus und heißt der zu ϕ **adjungierte Endomorphismus** oder die zu ϕ **adjungierte Abbildung**. Es gilt

$$\phi^\dagger = \sharp \circ \phi^* \circ \flat$$

wobei $\phi^* : V^* \to V^*$, $\alpha \mapsto \alpha \circ \phi$ die duale Abbildung aus Satz 4.4.7 und $\flat : V \to V^*$, $v \mapsto v^\flat$ und $\sharp : V^* \to V$, $\alpha \mapsto \alpha^\sharp$ die bijektiven (semi)linearen Abbildungen aus Satz 11.1.1 sind.

Beweis:
Die Eindeutigkeit der Adjungierten ergibt sich direkt aus der Tatsache, dass β nicht ausgeartet ist. Denn sind $\chi, \chi' : V \to V$ zwei Adjungierte von ϕ, so folgt für alle $v, w \in V$

$$\beta(\chi(v) - \chi'(v), w) = \beta(\chi(v), w) - \beta(\chi'(v), w) = \beta(v, \phi(w)) - \beta(v, \phi(w)) = 0.$$

Da β nicht ausgeartet ist, folgt $\chi(v) - \chi'(v) = 0$ für alle $v \in V$ und damit $\chi = \chi'$.

Für die Existenz betrachten wir die bijektiven (semi)linearen Abbildungen $\sharp : V^* \to V$, $\alpha \mapsto \alpha^\sharp$ und $\flat = \sharp^{-1} : V \to V^*$, $v \mapsto v^\flat$ aus Satz 11.1.1 mit (i) $v^\flat(w) = \beta(v, w)$ und (ii) $\beta(\alpha^\sharp, w) = \alpha(w)$ für alle $v, w \in V$, $\alpha \in V^*$. Auch im semilinearen Fall ist $\sharp \circ \phi^* \circ \flat : V \to V$ linear für jeden Endomorphismus $\phi : V \to V$, denn dann gilt

$$\phi^\dagger(\lambda v) = \sharp(\phi^*(\flat(\lambda v))) = \sharp(\phi^*(\bar{\lambda}\flat(v))) = \sharp(\bar{\lambda}\phi^*(\flat(v))) = \lambda \sharp(\phi^*(\flat(v))) = \lambda\, \phi^\dagger(v) \quad \forall v \in V, \lambda \in \mathbb{K}.$$

Für alle $v, w \in V$ ergibt sich aus der Definition der dualen Abbildung und aus (i) und (ii)

$$\beta(\sharp \circ \phi^* \circ \flat(v), w) = \beta((\phi^*(v^\flat))^\sharp, w) \overset{\text{(ii)}}{=} \phi^*(v^\flat)(w) \overset{\text{Def. } \phi^*}{=} v^\flat \circ \phi(w) = v^\flat(\phi(w)) \overset{\text{(i)}}{=} \beta(v, \phi(w)).$$

Damit ist $\phi^\dagger = \sharp \circ \phi^* \circ \flat : V \to V$ der adjungierte Endomorphismus von $\phi : V \to V$. □

Auch für unendlich-dimensionale $\mathbb{K}$-Vektorräume V definiert man die Adjungierte eines Endomorphismus $\phi \in \mathrm{End}_{\mathbb{K}}(V)$ bezüglich einer hermiteschen Sesquilinearform oder symmetrischen Bilinearform β durch die Gleichung (23). Die Eindeutigkeit folgt dann wie im endlichdimensionalen Fall, aber ϕ muss keine Adjungierte besitzen. Denn in diesem Fall ist die lineare Abbildung $\beta : V \to V^*$, $v \mapsto v^\beta$ zwar definiert und injektiv, aber nicht notwendigerweise surjektiv, und man kann $\phi^\dagger$ nicht durch $\phi^\dagger = \sharp \circ \phi^* \circ \flat$ definieren.

Bevor wir konkrete Beispiele betrachten, untersuchen wir zunächst allgemein die Eigenschaften der Adjungierten. Offensichtliche Fragen sind, wie sich die Adjungierte unter Addition, skalarer Multiplikation und Verkettung von Abbildungen verhält. Ebenso interessiert man sich dafür, was passiert, wenn man Adjungierte von Adjungierten bildet, und was der Zusammenhang zwischen der beschreibenden Matrix einer linearen Abbildung $\phi : V \to V$ und der beschreibenden Matrix der Adjungierten $\phi^\dagger : V \to V$ bezüglich einer Basis von V ist.

Satz 11.1.3: Sei V ein n-dimensionaler $\mathbb{K}$-Vektorraum, β eine nicht ausgeartete hermitesche Sesquilinearform oder eine nicht ausgeartete symmetrische Bilinearform auf V und $\phi, \psi : V \to V$ Endomorphismen. Dann gilt für die adjungierten Endomorphismen

$$(\phi+\psi)^\dagger = \phi^\dagger + \psi^\dagger \quad (\phi \circ \psi)^\dagger = \psi^\dagger \circ \phi^\dagger \quad (\phi^\dagger)^\dagger = \phi \quad \mathrm{id}_V^\dagger = \mathrm{id}_V \quad (\lambda\phi)^\dagger = \begin{cases} \lambda\,\phi^\dagger & \beta \text{ bilinear} \\ \overline{\lambda}\,\phi^\dagger & \beta \text{ sesquilinear,} \end{cases}$$

sowie $(\phi^\dagger)^{-1} = (\phi^{-1})^\dagger$ falls ϕ invertierbar ist. Für jede geordnete Basis B von V gilt

$$_BM_B(\phi) = \begin{cases} ({}^B\beta^B)^{-1} \cdot {}_BM_B(\phi^\dagger)^T \cdot {}^B\beta^B & \beta \text{ bilinear} \\ ({}^B\beta^B)^{-1} \cdot {}_BM_B(\phi^\dagger)^\dagger \cdot {}^B\beta^B & \beta \text{ sesquilinear.} \end{cases}$$

Beweis:
Wir beweisen die Aussagen für hermitesche Sesquilinearformen. Der Beweis für symmetrische Bilinearformen ist analog und ergibt sich durch Weglassen der komplexen Konjugation.

1. Die erste Aussage folgt direkt aus der definierenden Gleichung (23) und der Eindeutigkeit der Adjungierten. Für alle Endomorphismen $\phi, \psi : V \to V$ und $v, w \in V$ ergibt sich nämlich

$$\beta(\phi^\dagger(v) + \psi^\dagger(v), w) = \beta(\phi^\dagger(v), w) + \beta(\psi^\dagger(v), w) = \beta(v, \phi(w)) + \beta(v, \psi(w)) = \beta(v, \phi(w) + \psi(w))$$
$$\beta(\psi^\dagger(\phi^\dagger(v)), w) = \beta(\phi^\dagger(v), \psi(w)) = \beta(v, \phi(\psi(w)))$$
$$\beta(\mathrm{id}_V(v), w) = \beta(v, w) = \beta(v, \mathrm{id}_V(w))$$
$$\beta(\phi(v), w) = \overline{\beta(w, \phi(v))} = \overline{\beta(\phi^\dagger(w), v)} = \beta(v, \phi^\dagger(w))$$
$$\beta(\overline{\lambda}\,\phi^\dagger(v), w) = \lambda\,\beta(\phi^\dagger(v), w) = \lambda\,\beta(v, \phi(w)) = \beta(v, \lambda\,\phi(w)) = \beta(v, \lambda\,\phi(w)).$$

Da die Adjungierte durch (23) eindeutig bestimmt ist, beweist dies die erste Aussage. Ist ϕ invertierbar, so ergibt sich $(\phi^{-1})^\dagger \circ \phi^\dagger = (\phi \circ \phi^{-1})^\dagger = \mathrm{id}_V^\dagger = (\phi^{-1} \circ \phi)^\dagger = \phi^\dagger \circ (\phi^{-1})^\dagger$ und damit $(\phi^{-1})^\dagger = (\phi^\dagger)^{-1}$.

2. Ist $\phi : V \to V$ ein Endomorphismus mit Adjungierter $\phi^\dagger : V \to V$ und $B = (b_1, ..., b_n)$ eine geordnete Basis von V, so gibt es Koeffizienten $\phi_{ji}, \phi^\dagger_{ji} \in \mathbb{K}$ mit $\phi(b_i) = \Sigma_{j=1}^n \phi_{ji} b_j$ und $\phi^\dagger(b_i) = \Sigma_{j=1}^n \phi^\dagger_{ji} b_j$ für alle $i, j \in \{1, ..., n\}$. Die beschreibenden Matrizen von ϕ und $\phi^\dagger$ sind dann gegeben durch ${}_BM_B(\phi)_{ji} = \phi_{ji}$ und ${}_BM_B(\phi^\dagger)_{ji} = \phi^\dagger_{ji}$. Einsetzen in Gleichung (23) ergibt

$$\begin{aligned}
\beta(\phi^\dagger(b_i), b_j) &= \beta(\Sigma_{k=1}^n \phi^\dagger_{ki} b_k, b_j) = \Sigma_{k=1}^n \overline{\phi^\dagger_{ki}}\, \beta(b_k, b_j) = \Sigma_{k=1}^n \overline{\phi^\dagger_{ki}}\, {}^B\beta^B{}_{kj} = ({}_BM_B(\phi^\dagger)^\dagger \cdot {}^B\beta^B)_{ij} \\
\beta(b_i, \phi(b_j)) &= \beta(b_i, \Sigma_{k=1}^n \phi_{kj} b_k) = \Sigma_{k=1}^n \phi_{kj}\, \beta(b_i, b_k) = \Sigma_{k=1}^n \phi_{kj}\, {}^B\beta^B{}_{ik} = ({}^B\beta^B \cdot {}_BM_B(\phi))_{ij}
\end{aligned}$$

Da diese beiden Ausdrücke per Definition der Adjungierten für alle $i, j \in \{1, ..., n\}$ übereinstimmen müssen, ergibt sich ${}_BM_B(\phi^\dagger)^\dagger \cdot {}^B\beta^B = {}^B\beta^B \cdot {}_BM_B(\phi)$. Da β nicht ausgeartet ist, gilt nach Satz 10.2.3, 3. $\mathrm{rg}({}^B\beta^B) = n$ und somit ist die Matrix ${}^B\beta^B$ invertierbar. Durch Linksmultiplikation der Gleichung mit $({}^B\beta^B)^{-1}$ folgt die Behauptung. □

Am Beweis erkennt man, dass die Voraussetzung, dass β hermitesch bzw. symmetrisch ist, nur für die Aussage $(\phi^\dagger)^\dagger = \phi$ benötigt wird. Diese führt aber zu wesentlichen Vereinfachungen. Besonders offensichtlich wird dies im Fall von euklidischen oder unitären Vektorräumen, wo sie es einem erlaubt, Kerne und Bilder des adjungierten Endomorphismus als die orthogonalen Komplemente von Bildern und Kernen des Endomorphismus zu beschreiben. Insbesondere klärt sich so auch, warum die Bezeichnung † sowohl für den adjungierten Endomorphismus als auch für die hermitesch konjugierte Matrix aus Satz 10.2.3 verwendet wird, und welcher Zusammenhang zwischen den beiden Konzepten besteht.

Satz 11.1.4: Sei V mit $\langle\, ,\, \rangle$ ein endlich-dimensionaler unitärer oder euklidischer Vektorraum und $\phi : V \to V$ ein Endomorphismus. Dann gilt

$$\ker(\phi^\dagger) = \mathrm{im}(\phi)^\perp \quad \mathrm{im}(\phi^\dagger) = \ker(\phi)^\perp \quad \mathrm{im}(\phi) = \ker(\phi^\dagger)^\perp \quad \ker(\phi) = \mathrm{im}(\phi^\dagger)^\perp.$$

Für jede Orthonormalbasis B von V gilt für die beschreibenden Matrizen ${}_BM_B(\phi^\dagger) = {}_BM_B(\phi)^\dagger$ (unitärer Fall) oder ${}_BM_B(\phi^\dagger) = {}_BM_B(\phi)^T$ (euklidischer Fall).

Beweis:
Aus der Definition des orthogonalen Komplements und der Adjungierten folgt

$$w \in \mathrm{im}(\phi^\dagger)^\perp \Leftrightarrow 0 = \langle \phi^\dagger(v), w\rangle = \langle v, \phi(w)\rangle \;\; \forall v \in V \Leftrightarrow \phi(w) = 0 \Leftrightarrow w \in \ker(\phi),$$

also $\mathrm{im}(\phi^\dagger)^\perp = \ker(\phi)$. Da V endlich-dimensional und $\langle\, ,\, \rangle$ ein Skalarprodukt ist, folgt mit Satz 10.4.2, 5. dann auch $\mathrm{im}(\phi^\dagger) = (\mathrm{im}(\phi^\dagger)^\perp)^\perp = \ker(\phi)^\perp$. Die anderen zwei Gleichungen ergeben sich aus Betrachtung von $\phi^\dagger$ statt ϕ und der Identität $(\phi^\dagger)^\dagger = \phi$.

Ist B eine Orthonormalbasis, so gilt ${}^B\langle\, ,\, \rangle^B = \mathbb{1}_n$. Also ist die beschreibende Matrix des adjungierten Endomorphismus nach Satz 11.1.3 gegeben durch ${}_BM_B(\phi^\dagger) = {}_BM_B(\phi)^\dagger$ für unitäre und durch ${}_BM_B(\phi^\dagger) = {}_BM_B(\phi)^T$ für euklidische Vektorräume. □

Da sowohl die Adjungierte einer linearen Abbildung $\phi : V \to V$ bezüglich einer Bi- oder Sesquilinearform β als auch die β-unitären Endomorphismen $\phi : V \to V$ aus Definition 10.1.7 durch eine Bedingung charakterisiert sind, in der die Bi- oder Sesquilinearform β und Endomorphismen $\phi : V \to V$ auftreten, ist es naheliegend, dass sich das Konzept der Unitarität ebenfalls mit Hilfe der Adjungierten beschreiben lassen sollte.

In der Tat findet man, dass unitäre Endomorphismen genau die Endomorphismen sind, die invers zu ihrer Adjungierten sind. Dies verallgemeinert den Zusammenhang zwischen unitären Endomorphismen des $\mathbb{R}^n$ bzw. $\mathbb{C}^n$ und unitären Matrizen aus Korollar 10.2.6.

Satz 11.1.5: Sei V ein endlich-dimensionaler $\mathbb{K}$-Vektorraum, β eine nicht ausgeartete hermitesche Sesquilinearform oder eine nicht ausgeartete symmetrische Bilinearform auf V. Dann ist ein Endomorphismus $\phi : V \to V$ β-unitär genau dann, wenn ϕ invertierbar ist mit $\phi^\dagger = \phi^{-1}$.

Beweis:
Ein Endomorphismus $\phi : V \to V$ ist nach Bemerkung 10.1.8 und Satz 10.1.9 genau dann in $U(V, \beta)$ enthalten, wenn

$$\beta(v, w) = \beta(\phi(v), \phi(w)) = \beta((\phi^\dagger \circ \phi)(v), w) \;\; \forall v, w \in V \quad \Leftrightarrow \quad \beta((\phi^\dagger \circ \phi)(v) - v, w) = 0 \;\; \forall v, w \in V.$$

Da β nicht ausgeartet ist, ist dies der Fall genau dann, wenn $(\phi^\dagger \circ \phi)(v) = v$ für alle $v \in V$ gilt, also wenn $\phi^\dagger \circ \phi = \mathrm{id}_V$. Da in jeder Gruppe Linksinverse Rechtsinverse sind und umgekehrt, ist das gleichbedeutend mit $\phi^\dagger = \phi^{-1}$. □

11.2 Normale Endomorphismen

In diesem Abschnitt betrachten wir Endomorphismen $\phi : V \to V$, die auf bestimmte Weise mit ihrem adjungierten Endomorphismus $\phi^\dagger : V \to V$ verträglich sind. Da sich der adjungierte Endomorphismus durch Konjugation der dualen Abbildung $\phi^* : V^* \to V^*$ mit dem Isomorphismus $\sharp : V^* \to V$ aus Satz 11.1.1 ergibt, können wir dies auch als Verträglichkeit zwischen dem Endomorphismus $\phi : V \to V$ und dem dualen Endomorphismus $\phi^* : V^* \to V^*$ auffassen.

Zunächst stellt sich die Frage, was überhaupt eine sinnvolle Verträglichkeitsbedingung zwischen einem Endomorphismus $\phi : V \to V$ und dem dazu adjungierten Endomorphismus $\phi^\dagger : V \to V$ sein kann. In Kapitel 7.2 wurde gezeigt, dass es in Bezug auf die Eigen- und Haupträume zweier Endomorphismen $\phi, \psi : V \to V$ günstig ist, wenn diese kommutieren, also wenn $\phi \circ \psi = \psi \circ \phi$ gilt. In diesem Fall erhält ψ nach Korollar 7.1.7 alle Eigen- und Haupträume von ϕ und umgekehrt.

Daher ist es naheliegend, Endomorphismen zu betrachten, die mit ihrer Adjungierten kommutieren. Offensichtlich ist diese Bedingung für unitäre Endomorphismen erfüllt, also Endomorphismen mit $\phi^\dagger = \phi^{-1}$, und ebenso für Endomorphismen mit $\phi^\dagger = \phi$ oder mit $\phi^\dagger = -\phi$.

Definition 11.2.1: Sei V ein n-dimensionaler $\mathbb{K}$-Vektorraum, β eine nicht ausgeartete hermitesche Sesquilinearform oder eine nicht ausgeartete symmetrische Bilinearform auf V.

Ein Endomorphismus $\phi : V \to V$ heißt

- β-**normal**, wenn $\phi^\dagger \circ \phi = \phi \circ \phi^\dagger$ gilt,
- β-**selbstadjungiert**, wenn $\phi^\dagger = \phi$ gilt,
- β-**antiselbstadjungiert**, wenn $\phi^\dagger = -\phi$ gilt.

Wichtige Beispiele normaler Endomorphismen liefert das folgende Lemma, das auch die Beziehung zwischen normalen, unitären und (anti)selbstadjungierten Endomorphismen klärt.

Lemma 11.2.2: Sei und V ein n-dimensionaler $\mathbb{K}$-Vektorraum, β eine nicht ausgeartete hermitesche Sesquilinearform oder eine nicht ausgeartete symmetrische Bilinearform auf V.

Dann gilt für alle Endomorphismen $\phi, \psi : V \to V$:

1. $\phi \circ \phi^\dagger$ und $\phi^\dagger \circ \phi$ sind β-selbstadjungiert.
2. Ist ϕ β-unitär oder β-(anti)selbstadjungiert, so ist ϕ auch β-normal.
3. Ist ϕ β-normal, so sind auch $\lambda\phi$ und $\phi - \lambda \mathrm{id}_V$ β-normal für alle $\lambda \in \mathbb{K}$.
4. Ist ϕ ist β-normal und ψ β-unitär, so ist $\psi \circ \phi \circ \psi^{-1}$ β-normal.

Beweis:
Wir beweisen die Aussagen für Sesquilinearformen β. Der entsprechende Beweis für Bilinearformen ergibt sich durch Weglassen der komplexen Konjugation.
1. Mit den Rechenregeln aus Satz 11.1.3 erhält man $(\phi \circ \phi^\dagger)^\dagger = (\phi^\dagger)^\dagger \circ \phi^\dagger = \phi \circ \phi^\dagger$ und $(\phi^\dagger \circ \phi)^\dagger = \phi^\dagger \circ (\phi^\dagger)^\dagger = \phi^\dagger \circ \phi$.

2. Ist ϕ β-unitär, so gilt $\phi^\dagger = \phi^{-1}$ nach Satz 11.1.5 und damit $\phi^\dagger \circ \phi = \phi \circ \phi^\dagger = \mathrm{id}_V$. Ist ϕ β-selbstadjungiert oder β-antiselbstadjungiert, so gilt $\phi^\dagger = \pm\phi$ und damit $\phi \circ \phi^\dagger = \pm\phi \circ \phi = \phi^\dagger \circ \phi$.

3. Ist ϕ β-normal, so ist $\phi^\dagger \circ \phi = \phi \circ \phi^\dagger$, und mit Satz 11.1.3 folgt für alle $\lambda \in \mathbb{K}$

$$\begin{aligned}
&(\lambda\phi)^\dagger \circ (\lambda\phi) = (\overline{\lambda}\,\phi^\dagger) \circ (\lambda\phi) = |\lambda|^2 \phi^\dagger \circ \phi = |\lambda|^2 \phi \circ \phi^\dagger = (\lambda\phi) \circ (\overline{\lambda}\,\phi^\dagger) = (\lambda\phi) \circ (\lambda\phi)^\dagger \\
&(\phi - \lambda \mathrm{id}_V)^\dagger \circ (\phi - \lambda \mathrm{id}_V) = (\phi^\dagger - \overline{\lambda}\mathrm{id}_V) \circ (\phi - \lambda \mathrm{id}_V) = \phi^\dagger \circ \phi - \lambda\phi^\dagger - \overline{\lambda}\phi - |\lambda|^2 \mathrm{id}_V \\
&\ = \phi \circ \phi^\dagger - \lambda\phi^\dagger - \overline{\lambda}\phi - |\lambda|^2 \mathrm{id}_V = (\phi - \lambda \mathrm{id}_V) \circ (\phi^\dagger - \overline{\lambda}\mathrm{id}_V) = (\phi - \lambda \mathrm{id}_V) \circ (\phi - \lambda \mathrm{id}_V)^\dagger.
\end{aligned}$$

4. Ist ϕ β-normal und ψ β-unitär, so folgt mit den Rechenregeln aus Satz 11.1.3

$$\begin{aligned}
&(\psi \circ \phi \circ \psi^{-1})^\dagger \circ (\psi \circ \phi \circ \psi^{-1}) = (\psi^{-1})^\dagger \circ \phi^\dagger \circ \psi^\dagger \circ \psi \circ \phi \circ \psi^{-1} = \psi \circ \phi^\dagger \circ \phi \circ \psi^{-1} \\
&= \psi \circ \phi \circ \phi^\dagger \circ \psi^{-1} = \psi \circ \phi \circ \psi^{-1} \circ (\psi^{-1})^\dagger \circ \phi^\dagger \circ \psi^{-1} = (\psi \circ \phi \circ \psi^{-1}) \circ (\psi \circ \phi \circ \psi^{-1})^\dagger
\end{aligned}$$ □

Wir untersuchen nun, wie sich normale und (anti)selbstadjungierte Endomorphismen $\phi : V \to V$ in einem euklidischen oder unitären Vektorraum durch die beschreibende Matrix ${}_BM_B(\phi)$ bezüglich einer Basis B charakterisieren lassen. Dabei ist es günstig, Orthonormalbasen zu betrachten, da dort die beschreibenden Matrizen ${}^B\beta^B$ eine besonders einfache Form annehmen.

Satz 11.2.3: Sei V ein n-dimensionaler euklidischer oder unitärer Vektorraum und B eine Orthonormalbasis von V. Dann gilt für jeden Endomorphismus $\phi : V \to V$

- ϕ normal $\Leftrightarrow {}_BM_B(\phi)^\dagger \cdot {}_BM_B(\phi) = {}_BM_B(\phi) \cdot {}_BM_B(\phi)^\dagger$ (unitärer Fall)
- ϕ normal $\Leftrightarrow {}_BM_B(\phi)^T \cdot {}_BM_B(\phi) = {}_BM_B(\phi) \cdot {}_BM_B(\phi)^T$ (euklidischer Fall)
- ϕ selbstadjungiert $\Leftrightarrow {}_BM_B(\phi)^\dagger = {}_BM_B(\phi)$ (unitärer Fall)
- ϕ selbstadjungiert $\Leftrightarrow {}_BM_B(\phi)^T = {}_BM_B(\phi)$ (euklidischer Fall)
- ϕ antiselbstadjungiert $\Leftrightarrow {}_BM_B(\phi)^\dagger = -{}_BM_B(\phi)$ (unitärer Fall)
- ϕ antiselbstadjungiert $\Leftrightarrow {}_BM_B(\phi)^T = -{}_BM_B(\phi)$ (euklidischer Fall).

Beweis:
Wir beweisen die Aussage für unitäre Vektorräume. Da B eine Orthonormalbasis von V ist, gilt ${}^B\beta^B = \mathbb{1}_n$, und aus Satz 11.1.3 folgt ${}_BM_B(\phi^\dagger) = {}_BM_B(\phi)^\dagger$.

Da ${}_BM_B : \mathrm{End}_\mathbb{K}(V) \to \mathrm{Mat}(n\times n, \mathbb{C})$ bijektiv ist, ist ϕ somit (anti)selbstadjungiert genau dann, wenn ${}_BM_B(\phi) = (-)_BM_B(\phi^\dagger) = (-)_BM_B(\phi)^\dagger$ gilt. Da ${}_BM_B : \mathrm{End}_\mathbb{K}(V) \to \mathrm{Mat}(n \times n, \mathbb{C})$ ein Algebraisomorphismus ist, ergibt sich

$$
\begin{aligned}
{}_BM_B(\phi^\dagger \circ \phi) &= {}_BM_B(\phi^\dagger) \cdot {}_BM_B(\phi) = {}_BM_B(\phi)^\dagger \cdot {}_BM_B(\phi) \\
{}_BM_B(\phi \circ \phi^\dagger) &= {}_BM_B(\phi) \cdot {}_BM_B(\phi^\dagger) = {}_BM_B(\phi) \cdot {}_BM_B(\phi)^\dagger,
\end{aligned}
$$

und durch Gleichsetzen der zwei Gleichungen folgt die Behauptung für ϕ normal. □

Dieser Satz erlaubt es uns, die Selbstadjungiertheit und Normalität von Endomorphismen endlich-dimensionaler Vektorräume durch deren beschreibende Matrizen bezüglich Orthonormalbasen zu erfassen. Dies reduziert den allgemeinen Fall auf die Betrachtung von Endomorphismen des $\mathbb{C}^n$ oder $\mathbb{R}^n$ mit dem Standardskalarprodukt und motiviert die folgende Definition.

Definition 11.2.4:
Eine Matrix $M \in \mathrm{Mat}(n \times n, \mathbb{C})$ heißt

- **normal**, wenn $M^\dagger \cdot M = M \cdot M^\dagger$.
- **selbstadjungiert** oder **hermitesch** wenn $M^\dagger = M$,
- **antiselbstadjungiert** oder **antihermitesch** wenn $M^\dagger = -M$.

Eine Matrix $M \in \mathrm{Mat}(n \times n, \mathbb{R})$ heißt

- **normal**, wenn $M^T \cdot M = M \cdot M^T$,
- **selbstadjungiert** oder **symmetrisch**, wenn $M^T = M$,
- **antiselbstadjungiert** oder **antisymmetrisch** wenn $M^T = -M$.

11.3 Diagonalisierung normaler Endomorphismen

Eine wesentliche Motivation für die Einführung der normalen Endomorphismen im letzten Abschnitt war es, dass das Kommutieren eines Endomorphismus $\phi : V \to V$ mit seiner Adjungierten $\phi^\dagger : V \to V$ in Bezug auf die Eigenräume und Haupträume günstig sein sollte, da jeder mit ϕ kommutierende Endomorphismus die Eigenräume und Haupträume von ϕ erhält.

In diesem Abschnitt untersuchen wir die Eigenwerte, Eigenräume und Eigenvektoren normaler Abbildungen und leiten Diagonalisierbarkeitsaussagen her. Dabei beschränken wir uns auf Endomorphismen euklidischer und unitärer Vektorräume, da dies die wichtigsten Fälle sind und dort stärkere Aussagen möglich sind als für den allgemeinen Fall. Dafür stellen wir zunächst eine Beziehung zwischen den Eigenwerten eines normalen Endomorphismus ϕ und untersuchen die Beziehung zwischen den zugehörigen Eigenvektoren.

Satz 11.3.1: Sei V mit $\langle\, ,\, \rangle$ ein endlich-dimensionaler euklidischer oder unitärer Vektorraum und $\phi : V \to V$ ein normaler Endomorphismus. Dann gilt:

1. Ein Vektor $v \in V$ ist ein Eigenvektor von ϕ zum Eigenwert λ genau dann, wenn v ein Eigenvektor von $\phi^\dagger$ zum Eigenwert $\overline{\lambda}$ ist.
2. Ist v ein Eigenvektor von ϕ zum Eigenwert λ, so ist das orthogonale Komplement $\{v\}^\perp \subseteq V$ ϕ-invariant: $\phi(\{v\}^\dagger) \subseteq \{v\}^\dagger$.
3. Sind $v_1, v_2 \in V$ Eigenvektoren von ϕ zu verschiedenen Eigenwerten λ_1, λ_2, ist $\langle v_1, v_2\rangle = 0$.

Beweis:
1. Ist $\phi : V \to V$ normal, so ist nach Lemma 11.2.2 auch $\phi - \lambda \mathrm{id}_V$ normal für alle $\lambda \in \mathbb{K}$. Für alle normalen Endomorphismen $\psi \in \mathrm{End}_{\mathbb{K}}(V)$ und $v \in V$ folgt mit (23) und Satz 11.1.3

$$||\psi^\dagger(v)||^2 = \langle \psi^\dagger(v), \psi^\dagger(v)\rangle = \langle v, \psi \circ \psi^\dagger(v)\rangle = \langle v, \psi^\dagger \circ \psi(v)\rangle = \langle (\psi^\dagger)^\dagger(v), \psi(v)\rangle = \langle \psi(v), \psi(v)\rangle = ||\psi(v)||^2.$$

Daraus ergibt sich $E(\phi, \lambda) = E(\phi^\dagger, \bar{\lambda})$:

$$\begin{aligned} &v \in E(\phi, \lambda) \Leftrightarrow (\phi - \lambda \mathrm{id}_V)(v) = 0 \Leftrightarrow 0 = ||(\phi - \lambda \mathrm{id}_V)(v)||^2 = ||(\phi - \lambda \mathrm{id}_V)^\dagger(v)||^2 \\ &\Leftrightarrow 0 = ||(\phi^\dagger - \bar{\lambda}\mathrm{id}_V)(v)||^2 \Leftrightarrow (\phi^\dagger - \bar{\lambda}\mathrm{id}_V)(v) = 0 \Leftrightarrow v \in E(\phi^\dagger, \bar{\lambda}). \end{aligned}$$

2. Ist $v \in V$ ein Eigenvektor von ϕ zum Eigenwert λ, so ist v nach 1. auch ein Eigenvektor von $\phi^\dagger$ zum Eigenwert $\overline{\lambda}$. Für $w \in \{v\}^\perp$ folgt daraus $\langle v, \phi(w)\rangle = \langle \phi^\dagger(v), w\rangle = \langle \overline{\lambda} v, w\rangle = \lambda \langle v, w\rangle = 0$, also $\phi(w) \in \{v\}^\perp$.

3. Sind $v_1, v_2 \in V$ Eigenvektoren von ϕ zu verschiedenen Eigenwerten $\lambda_1, \lambda_2 \in \mathbb{K}$, so folgt mit 1. $0 = \langle v_2, \phi(v_1)\rangle - \langle \phi^\dagger(v_2), v_1\rangle = \langle v_2, \lambda_1 v_1\rangle - \langle \overline{\lambda_2} v_2, v_1\rangle = (\lambda_1 - \lambda_2)\langle v_1, v_2\rangle$. Da $\lambda_1 - \lambda_2 \neq 0$ gilt, ergibt sich daraus $\langle v_1, v_2\rangle = 0$. □

Im Fall von normalen Endomorphismen $\phi : V \to V$ *unitärer* Vektorräume erlaubt es einem die Invarianz des orthogonalen Komplements jedes Eigenvektors, den Endomorphismus ϕ zu diagonalisieren. Dabei erhält man nicht nur eine Basis von V aus Eigenvektoren, sondern sogar eine Orthonormalbasis.

Denn jeder Endomorphismus eines endlich-dimensionalen komplexen Vektorraums besitzt mindestens einen Eigenvektor. Wir können also einen Eigenvektor v von ϕ wählen und den Vektorraum V als $V = \mathbb{K}v \oplus \{v\}^\perp$ zerlegen. Da das orthogonale Komplement $\{v\}^\perp$ invariant unter ϕ ist, erhält man durch Einschränken von ϕ auf $\{v\}^\perp$ einen weiteren normalen Endomorphismus, auf den man das selbe Verfahren anwenden kann, bis man schließlich den gesamten Vektorraum V als direkte Summe von Eigenräumen zerlegt hat.

Satz 11.3.2: (**Spektralsatz, Diagonalisierung normaler Endomorphismen**)
Sei V mit $\langle\, ,\, \rangle$ ein n-dimensionaler unitärer Vektorraum und $\phi : V \to V$ ein Endomorphismus. Dann ist ϕ genau dann normal, wenn eine Orthonormalbasis aus Eigenvektoren von ϕ existiert.

Beweis:
$\Rightarrow$: Induktion über n.
$n = 1$: Dann ist $\phi = \lambda \mathrm{id}_V$, und jeder Vektor $v \in V$ mit $||v|| = 1$ bildet eine Orthonormalbasis aus Eigenvektoren von ϕ.

$n - 1 \to n$: Sei die Aussage bewiesen für unitäre Vektorräume der Dimension $\leq n - 1$. Sei V ein unitärer Vektorraum der Dimension n und $\phi : V \to V$ ein normaler Endomorphismus.

Da V ein Vektorraum über $\mathbb{C}$ ist, hat das charakteristische Polynom von ϕ eine Nullstelle $\lambda_1 \in \mathbb{C}$, und es existiert ein Eigenvektor $u_1 \in V$ zum Eigenwert λ_1. Dann ist der Vektor $v_1 := u_1/||u_1||$ ein normierter Vektor mit $\phi(v_1) = \lambda_1 v_1$. Nach Satz 10.4.2 gilt $V = \mathbb{K}v_1 \oplus \{v_1\}^\perp$ und nach Satz 11.3.1 2. gilt $\phi(\{v_1\}^\perp) \subseteq \{v_1\}^\perp$. Die Einschränkung $\phi|_{\{v_1\}^\perp} : \{v_1\}^\perp \to \{v_1\}^\perp$ ist ebenfalls normal. Da $\dim_{\mathbb{K}}\{v_1\}^\perp = n - 1$, existiert nach Induktionsvoraussetzung eine Orthonormalbasis $(v_2, ..., v_n)$ von $\{v_1\}^\perp$ aus Eigenvektoren von ϕ. Damit ist $(v_1, ..., v_n)$ eine Orthonormalbasis von V aus Eigenvektoren von ϕ.

$\Leftarrow$: Ist C eine Orthonormalbasis von V aus Eigenvektoren von ϕ, so ist die beschreibende Matrix ${}^C\langle\,,\rangle^C$ des Skalarprodukts bezüglich C die Einheitsmatrix und die beschreibende Matrix ${}_CM_C(\phi)$ ist diagonal. Aus Satz 11.1.3 folgt ${}_CM_C(\phi^\dagger) = {}_CM_C(\phi)^\dagger$, und die beschreibende Matrix von $\phi^\dagger$ ist ebenfalls diagonal. Da diagonale Matrizen kommutieren, folgt ${}_CM_C(\phi^\dagger\circ\phi) = {}_CM_C(\phi^\dagger)\cdot{}_CM_C(\phi) = {}_CM_C(\phi)\cdot{}_CM_C(\phi^\dagger) = {}_CM_C(\phi\circ\phi^\dagger)$, also $\phi^\dagger\circ\phi = \phi\circ\phi^\dagger$. $\square$

Indem man als Spezialfall den unitären Vektorraum $V = \mathbb{C}^n$ mit dem Standardskalarprodukt und einen Endomorphismus $\phi_A : \mathbb{C}^n \to \mathbb{C}^n$, $v \mapsto A\cdot v$ mit $A \in \mathrm{Mat}(n\times n,\mathbb{C})$ betrachtet, kann man die Aussagen aus Satz 11.3.2 in Aussagen über die Diagonalisierbarkeit normaler Matrizen umwandeln. Die Normalität von ϕ_A entspricht dann der Normalität der Matrix A und die Diagonalisierbarkeit von ϕ_A der Diagonalisierbarkeit von A. Dass ein normaler Endomorphismus ϕ_A sogar diagonalisierbar bezüglich einer *Orthonormalbasis* ist, wird zu der Aussage, dass die zugehörige Matrix A diagonalisierbar mittels einer *unitären* Matrix ist.

Satz 11.3.3: Für eine Matrix $A \in \mathrm{Mat}(n\times n,\mathbb{C})$ sind äquivalent:

(i) A normal: $A^\dagger\cdot A = A\cdot A^\dagger$.

(ii) Es gibt eine Orthonormalbasis des $\mathbb{C}^n$ aus Eigenvektoren von A.

(iii) A ist diagonalisierbar mittels einer unitären Matrix:
Es gibt eine Matrix $U \in \mathrm{U}(n,\mathbb{C})$ und eine Diagonalmatrix D mit $A = U\cdot D\cdot U^{-1}$.

Beweis:
(i)$\Leftrightarrow$(ii): Da die Standardbasis eine Orthonormalbasis des $\mathbb{C}^n$ ist, ist die Matrix A nach Satz 11.2.3 normal genau dann, wenn der Endomorphismus $\phi_A : \mathbb{C}^n \to \mathbb{C}^n$, $v\mapsto A\cdot v$ normal ist. Dies ist nach Satz 11.3.2 äquivalent zur Existenz einer Orthonormalbasis des $\mathbb{C}^n$ aus Eigenvektoren von ϕ_A, also zur Existenz einer Orthonormalbasis aus Eigenvektoren von A.

(ii)$\Leftrightarrow$(iii) Eine Matrix $U = (u_1, ..., u_n)$ mit Spaltenvektoren $u_1, ..., u_n \in \mathbb{C}^n$ ist nach Satz 10.3.16 unitär genau dann, wenn $(u_1, ..., u_n)$ eine Orthonormalbasis des $\mathbb{C}^n$ ist. Ist $A = U\cdot D\cdot U^{-1}$ mit $U \in \mathrm{U}(n,\mathbb{C})$ und D diagonal, so folgt $A\cdot u_i = A\cdot U\cdot e_i = U\cdot D\cdot e_i = d_{ii}\,U\cdot e_i = d_{ii}u_i$ für alle $i \in \{1, ..., n\}$ und somit ist $(u_1, ..., u_n)$ eine Orthonormalbasis aus Eigenvektoren von A. Ist umgekehrt $(u_1, ..., u_n)$ eine Orthonormalbasis aus Eigenvektoren von A mit $Au_i = \lambda_i u_i$ so ist die Matrix $U = (u_1, ..., u_n)$ unitär und für alle $i \in \{1, .., n\}$ gilt $A\cdot U\cdot e_i = A\cdot u_i = \lambda_i u_i = \lambda_i U\cdot e_i$. Also folgt $A = U\cdot D\cdot U^{-1}$ mit $D = (\lambda_i\,\delta_{ij})$ diagonal. $\square$

Insbesondere können wir Satz 11.3.2 und Satz 11.3.3 auf die Spezialfälle unitärer Endomorphismen und unitärer Matrizen anwenden, die nach Lemma 11.2.2 normal sind. Indem wir dann bezüglich einer Orthonormalbasis diagonalisieren, können wir unitäre Endomorphismen bzw. unitäre Matrizen dann als normale Endomorphismen bzw. Matrizen charakterisieren, deren Eigenwerte auf dem komplexen Einheitskreis $S^1 = \{z \in \mathbb{C} \mid |z| = 1\} \subseteq \mathbb{C}$ liegen.

Satz 11.3.4: Sei V mit $\langle\,,\rangle$ ein endlich-dimensionaler unitärer Vektorraum und $\phi : V \to V$ ein Endomorphismus. Dann sind äquivalent:

(i) ϕ ist unitär.

(ii) ϕ ist normal, und alle Eigenwerte von ϕ liegen auf dem Einheitskreis S^1.

(iii) V hat eine Orthonormalbasis aus Eigenvektoren von ϕ und mit Eigenwerten in S^1.

Beweis:
(ii) $\Leftrightarrow$ (iii) folgt direkt aus Satz 11.3.2.

(i)$\Rightarrow$(ii) Ist ϕ unitär, so ist ϕ nach Lemma 11.2.2 auch normal. Ist $v \in V$ ein Eigenvektor von ϕ zum Eigenwert λ, so folgt $0 = ||v||^2 - ||\phi(v)||^2 = ||v||^2 - ||\lambda\, v||^2 = ||v||^2 - |\lambda|^2||v||^2 = (|\lambda|^2 - 1)||v||^2$, und wegen $v \neq 0$ ergibt sich $|\lambda| = 1$.

(iii)$\Rightarrow$(i): Ist $B = (b_1, ..., b_n)$ eine Orthonormalbasis von V aus Eigenvektoren mit $\phi(b_i) = \lambda_i b_i$ und $|\lambda_i| = 1$ für $i \in \{1, ..., n\}$, so folgt $\langle \phi(b_i), \phi(b_j)\rangle = \langle \lambda_i b_i, \lambda_j b_j\rangle = \overline{\lambda_i}\, \lambda_j\, \langle b_i, b_j\rangle = |\lambda_i|^2\, \delta_{ij} = \delta_{ij}$. Also ist auch $(\phi(b_1), ..., \phi(b_n))$ eine Orthonormalbasis von V und ϕ nach Satz 10.3.15 unitär. $\square$

Korollar 11.3.5: Für eine Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{C})$ sind äquivalent:

(i) A ist unitär.

(ii) A ist normal, und alle Eigenwerte von A liegen auf dem Einheitskreis S^1.

(iii) A ist diagonalisierbar mittels einer unitären Matrix und mit Eigenwerten in S^1.

Beweis:
Da die Standardbasis eine Orthonormalbasis des $\mathbb{C}^n$ ist, ist die Matrix A ist nach Satz 11.2.3 normal genau dann, wenn der Endomorphismus $\phi_A : \mathbb{C}^n \to \mathbb{C}^n$, $v \mapsto A \cdot v$ normal ist und nach Korollar 10.2.6 unitär genau dann, wenn der Endomorphismus ϕ_A unitär ist. Da A und ϕ_A die gleichen Eigenwerte haben, folgt die Behauptung aus Satz 11.3.4 und Satz 11.3.3. $\square$

Bemerkung 11.3.6: Normale *reelle* Matrizen und damit auch orthogonale *reelle* Matrizen sind nach Satz 11.3.3 und Korollar 11.3.5 immer diagonalisierbar *über* $\mathbb{C}$, aber nicht notwendigerweise *über* $\mathbb{R}$. Ein Gegenbeispiel ist die Drehmatrix

$$D(\phi) = \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix},$$

die nach Beispiel 10.3.17 orthogonal, und damit insbesondere normal ist. In Beispiel 6.1.7 wurde aber gezeigt, dass sie nur dann diagonalisierbar über $\mathbb{R}$ ist, wenn $\phi \in \pi\mathbb{Z}$, während sie über $\mathbb{C}$ für alle Winkel $\phi \in \mathbb{R}$ diagonalisierbar ist.

Dass die Resultate von Satz 11.3.2 bis Korollar 11.3.5 für euklidische Vektorräume nicht gelten, wird auch am Beweis von Satz 11.3.2 deutlich. Dort wird im Induktionsschritt die Existenz eines Eigenvektors von ϕ benötigt, die zwar in einem komplexen Vektorraum immer gegeben ist, nicht aber in einem reellen Vektorraum. Um zu klären, unter welchen Zusatzbedingungen eine reelle orthogonale Matrix nicht nur diagonalisierbar über $\mathbb{C}$ sondern auch diagonalisierbar über $\mathbb{R}$ ist, benutzen wir das folgende Lemma.

Lemma 11.3.7: Eine über $\mathbb{C}$ diagonalisierbare reelle Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{R})$ ist diagonalisierbar über $\mathbb{R}$ genau dann, wenn all ihre komplexen Eigenwerte in $\mathbb{R}$ enthalten sind.

Beweis:
Hat A komplexe Eigenwerte in $\mathbb{C} \setminus \mathbb{R}$, so zerfällt das charakteristische Polynom p_A über $\mathbb{R}$ nicht in Linearfaktoren und A kann damit nicht diagonalisierbar über $\mathbb{R}$ sein. Ist A diagonalisierbar über $\mathbb{C}$ und alle komplexen Eigenwerte von A sind reell, so gibt es eine Matrix $S \in \mathrm{GL}(n, \mathbb{C})$ und

eine Diagonalmatrix $D \in \mathrm{Mat}(n \times n, \mathbb{R})$ mit $A = S \cdot D \cdot S^{-1}$. Da das charakteristische Polynom p_A dann auch über $\mathbb{R}$ in Linearfaktoren zerfällt, ist A ist trigonalisierbar über $\mathbb{R}$. Also gibt es eine reelle Matrix $R \in \mathrm{GL}(n, \mathbb{R})$ und eine Matrix $J \in \mathrm{Mat}(n \times n, \mathbb{R})$ in Jordan-Normalform mit $A = R \cdot J \cdot R^{-1}$. Wegen der Eindeutigkeit der Jordan-Normalform über $\mathbb{C}$ muss dann aber auch J eine Diagonalmatrix sein und bis auf Permutation der Diagonalelemente mit D übereinstimmen. Also ist A diagonalisierbar über $\mathbb{R}$. □

Mit Lemma 11.3.7 können wir nun die orthogonalen Matrizen $A \in \mathrm{Mat}(n \times n, \mathbb{R})$ charakterisieren, die auch über $\mathbb{R}$ diagonalisierbar sind. Dies sind nach Lemma 11.3.7 genau die orthogonalen Matrizen, deren Eigenwerte nicht nur auf dem komplexen Einheitskreis, sondern auch auf der reellen Achse liegen. Damit kommen nur noch ± 1 als Eigenwerte in Frage.

Korollar 11.3.8: Für eine Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{R})$ sind äquivalent:

(i) A ist orthogonal, und alle komplexen Eigenwerte von A sind reell.

(ii) A ist normal, und die einzigen komplexen Eigenwerte von A sind ± 1.

(iii) A ist diagonalisierbar über $\mathbb{R}$ mittels einer orthogonalen Matrix und mit Eigenwerten ± 1.

Beweis:
Da $S^1 \cap \mathbb{R} = \{z \in \mathbb{R} | z^2 = 1\} = \{1, -1\}$ sind die Bedingungen (i)-(iii) in Korollar 11.3.8 äquivalent zu den Bedingungen in Korollar 11.3.5 zusammen mit der Bedingung, dass alle komplexen Eigenwerte von A reell sind. Mit Korollar 11.3.5 und Lemma 11.3.7 folgt dann die Behauptung. □

Neben den unitären Endomorphismen hatten wir als zweiten wichtigen Spezialfall normaler Endomorphismen die selbstadjungierten Endomorphismen betrachtet. Die zugehörigen Matrizen sind nach Satz 11.2.3 die selbstadjungierten oder hermiteschen Matrizen. Indem wir Satz 11.3.2 und Satz 11.3.3 auf diese Spezialfälle anwenden und bezüglich einer Orthonormalbasis diagonalisieren, können wir selbstadjungierte Endomorphismen oder Matrizen dann als normale Endomorphismen oder Matrizen charakterisieren, deren Eigenwerte reell sind.

Der antiselbstadjungierte Fall ergibt sich direkt aus dem selbstadjungierten. Denn ein Endomorphismus ϕ eines unitären Vektorraums oder eine komplexe Matrix A ist antiselbstadjungiert, genau dann, wenn $\mathrm{i}\phi$ oder $\mathrm{i}A$ selbstadjungiert sind. In diesem Fall erhält man normale Endomorphismen oder Matrizen mit rein imaginären Eigenwerten.

Satz 11.3.9: Sei V mit $\langle\, ,\, \rangle$ ein endlich-dimensionaler unitärer Vektorraum und $\phi : V \to V$ ein Endomorphismus. Dann sind äquivalent:

(i) ϕ ist selbstadjungiert (antiselbstadjungiert).

(ii) ϕ ist normal, und alle Eigenwerte von ϕ sind reell (imaginär).

(iii) V besitzt eine Orthonormalbasis aus Eigenvektoren von ϕ mit nur reellen (imaginären) Eigenwerten.

Beweis:
Da ϕ selbstadjungiert ist, genau dann, wenn $\mathrm{i}\phi$ antiselbstadjungiert ist, reicht es, den selbstadjungierten Fall zu betrachten. (ii) $\Leftrightarrow$ (iii) folgt direkt aus Satz 11.3.2.

(i) $\Rightarrow$ (ii): Ist ϕ selbstadjungiert, so ist ϕ normal nach Lemma 11.2.2, und für jeden Eigenvektor $v \in V$ zum Eigenwert $\lambda \in \mathbb{C}$ gilt

$$0 = \langle v, \phi(v)\rangle - \langle \phi^\dagger(v), v\rangle = \langle v, \phi(v)\rangle - \langle \phi(v), v\rangle = \langle v, \lambda v\rangle - \langle \lambda v, v\rangle = (\lambda - \overline{\lambda})||v||^2.$$

Da $v \neq 0$ folgt daraus $\lambda - \bar{\lambda} = 0$ und somit ist der Eigenwert λ reell.

(iii) $\Rightarrow$ (i): Ist ϕ normal und $B = (b_1, ..., b_n)$ eine Orthonormalbasis von V aus Eigenvektoren von ϕ mit $\phi(b_i) = \lambda_i b_i$ und $\lambda_i \in \mathbb{R}$, so ergibt sich

$$\langle \phi^\dagger(b_i) - \phi(b_i), b_j\rangle = \langle b_i, \phi(b_j)\rangle - \langle \phi(b_i), b_j\rangle = \langle b_i, \lambda_j b_j\rangle - \langle \lambda_i b_i, b_j\rangle = (\lambda_j - \overline{\lambda_i})\,\delta_{ij} = 0$$

für alle $i, j \in \{1, ..., n\}$. Da B eine Basis von V ist, folgt $\phi = \phi^\dagger$. □

Korollar 11.3.10: Für eine Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{C})$ sind äquivalent:

(i) A ist hermitesch: $A = A^\dagger$.

(ii) A ist normal, und alle Eigenwerte von A sind reell.

(iii) A ist diagonalisierbar mittels einer unitären Matrix und alle Eigenwerte von A sind reell.

Beweis:
Da die Standardbasis eine Orthonormalbasis des $\mathbb{C}^n$ ist, ist die Matrix A ist nach Satz 11.2.3 normal genau dann, wenn der Endomorphismus $\phi_A : \mathbb{C}^n \to \mathbb{C}^n$, $v \mapsto A \cdot v$ normal ist und selbstadjungiert genau dann, wenn der Endomorphismus ϕ_A selbstadjungiert ist. Da A und ϕ_A die gleichen Eigenwerte haben, folgt die Behauptung aus Satz 11.3.9 und Satz 11.3.3. □

Anders als im Fall unitärer Matrizen überträgt sich im selbstadjungierten Fall das Diagonalisierbarkeitsresultat für hermitesche komplexe Matrizen direkt auf reelle symmetrische Matrizen. Die Bedingung, dass alle Eigenwerte reell sind, impliziert nämlich, dass letztere nicht nur diagonalisierbar über $\mathbb{C}$ sondern auch diagonalisierbar über $\mathbb{R}$ sind.

Korollar 11.3.11: Für eine Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{R})$ sind äquivalent:

(i) A ist symmetrisch: $A^T = A$.

(ii) A ist normal, und alle komplexen Eigenwerte von A sind reell.

(iii) A ist diagonalisierbar über $\mathbb{R}$ mittels einer orthogonalen Matrix.

Beweis:
(i) $\Leftrightarrow$ (ii) und (iii) $\Rightarrow$ (i) folgt direkt aus Korollar 11.3.10.
(i) $\Rightarrow$ (iii): Da jede symmetrische Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{R})$ auch hermitesch ist, ist sie nach Korollar 11.3.10 diagonalisierbar über $\mathbb{C}$ mittels einer unitären Matrix. Da aus der Hermitizität folgt, dass alle Eigenwerte von A reell sind, folgt mit Lemma 11.3.7, dass sie auch diagonalisierbar über $\mathbb{R}$ mittels einer orthogonalen Matrix ist. □

Bemerkung 11.3.12: *Komplexe* symmetrische Matrizen sind nicht notwendigerweise normal und damit auch nicht unbedingt diagonalisierbar mittels einer unitären Matrix. Ein Gegenbeispiel ist die Matrix

$$A = \begin{pmatrix} \mathrm{i} & 1 \\ 1 & 0 \end{pmatrix} \quad \text{mit} \quad A^T = A \quad \text{und}$$

$$A^\dagger \cdot A = \begin{pmatrix} -\mathrm{i} & 1 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} \mathrm{i} & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 2 & -\mathrm{i} \\ \mathrm{i} & 1 \end{pmatrix} \neq \begin{pmatrix} 2 & \mathrm{i} \\ -\mathrm{i} & 1 \end{pmatrix} = \begin{pmatrix} \mathrm{i} & 1 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} -\mathrm{i} & 1 \\ 1 & 0 \end{pmatrix} = A \cdot A^\dagger.$$

Antisymmetrische reelle Matrizen sind zwar normal, aber im allgemeinen *nicht* diagonalisierbar über $\mathbb{R}$. Nach Satz 11.3.9 liegen nämlich alle komplexen Eigenwerte einer antisymmetrischen reellen Matrix auf der imaginären Achse. Daraus folgt mit der komplexen Diagonalisierbarkeit der Matrix und mit Lemma 11.3.7, dass die einzige über $\mathbb{R}$ diagonalisierbare antisymmetrische Matrix die Nullmatrix ist.

Korollar 11.3.11 können wir insbesondere auf die beschreibende Matrix einer symmetrischen Bilinearform β auf einem endlich-dimensionalen euklidischen Vektorraum V anwenden. Denn nach Satz 10.2.3 ist diese für jede Basis von V symmetrisch. Korollar 11.3.11 garantiert dann die Existenz einer Orthonormalbasis von V, die gleichzeitig eine Orthogonalbasis für β ist.

Damit können wir die symmetrischen Bilinearformen und daher auch die quadratischen Formen auf V klassifizieren. Denn nach Bemerkung 10.3.7 definiert jede symmetrische Bilinearform β auf V eine quadratische Form $Q_\beta : V \to \mathbb{R}$, $v \mapsto \beta(v, v)$. Umgekehrt ist nach Bemerkung 10.3.7 die Abbildung $\beta_Q : V \times V \to \mathbb{R}$, $(v, w) \mapsto \frac{1}{2}Q(v + w) - \frac{1}{2}Q(v) - \frac{1}{2}Q(w)$ eine symmetrische Bilinearform auf V für jede quadratische Form Q auf V.

Korollar 11.3.13: (**Hauptachsentransformation**)
Sei V mit $\langle\,,\,\rangle$ ein endlich-dimensionaler euklidischer Vektorraum und β eine symmetrische Bilinearform auf V. Dann gibt es eine Orthonormalbasis B von V bezüglich des Skalarprodukts $\langle\,,\,\rangle$, so dass die beschreibende Matrix ${}^B\beta^B$ von β bezüglich B diagonal ist.

Beweis:
Nach Korollar 10.3.10 existiert eine Orthonormalbasis $C = (c_1, ..., c_n)$ von V bezüglich $\langle\,,\,\rangle$, also eine Basis C, für die die beschreibende Matrix des Skalarprodukts $\langle\,,\,\rangle$ die Einheitsmatrix ist. Nach Satz 10.2.3 ist die beschreibende Matrix ${}^C\beta^C$ symmetrisch. Also existiert nach Korollar 11.3.11 eine orthogonale Matrix $O \in \mathrm{O}(n, \mathbb{R})$ mit $n = \dim_{\mathbb{R}}(V)$, so dass $D = O^T \cdot {}^C\beta^C \cdot O$ diagonal ist. Wir setzen $b_i = \Sigma_{j=1}^n O_{ji} c_j$. Dann gilt ${}_CM_B(\mathrm{id}_V) = O$ und

$$
\begin{aligned}
{}^B\beta^B &= {}_CM_B(\mathrm{id}_V)^T \cdot {}^C\beta^C \cdot {}_CM_B(\mathrm{id}_V) = O^T \cdot {}^C\beta^C \cdot O = D \\
{}^B\langle\,,\,\rangle^B &= {}_CM_B(\mathrm{id}_V)^T \cdot {}^C\langle\,,\,\rangle^C \cdot {}_CM_B(\mathrm{id}_V) = O^T \cdot \mathbb{1}_n \cdot O = O^T \cdot O = \mathbb{1}_n.
\end{aligned}
$$

Also ist auch $B = (b_1, ..., b_n)$ eine Orthonormalbasis von V bezüglich des Skalarprodukts $\langle\,,\,\rangle$, und die beschreibende Matrix ${}^B\beta^B$ ist diagonal. □

Es ist instruktiv Korollar 11.3.13 mit den entsprechenden Aussagen aus Satz 10.4.4 und Satz 10.4.5 für symmetrische Bilinearformen über reellen Vektorräumen zu vergleichen. In Satz 10.4.4 wird bewiesen, dass zu jeder symmmetrischen Bilinearform auf einem rellen Vektorraum eine Basis existiert, so dass die beschreibende Matrix diagonal mit Diagonaleinträgen $0, 1, -1$ ist. In Satz 10.4.5 wird gezeigt, dass die Anzahl der Einträge $0, 1, -1$ nicht von der Basiswahl abhängt.

Die Voraussetzungen von Sätzen 10.4.4 und 10.4.5 sind also schwächer als die von Korollar 11.3.13, wo zusätzlich ein Skalarprodukt gefordert wird. Andererseits sind auch die Aussagen von Sätzen 10.4.4 und 10.4.5 schwächer, da dort nur die Existenz einer Basis bewiesen wird, nicht aber einer Orthonormalbasis für ein euklidisches Skalarprodukt. Deswegen können die Diagonaleinträge dort auch durch Reskaliserung der Basisvektoren auf die Werte $0, 1, -1$ beschränkt werden, während dies in Korollar 11.3.13 nicht möglich ist, ohne die Orthonormalität der Basis zu zerstören.

Der Name *Hauptachsentransformation* in Korollar 11.3.13 erklärt sich dadurch, dass die Vektoren der darin konstruierten Orthonormalbasis für $V = \mathbb{R}^n$ mit dem Standardskalarprodukt $\langle\,,\,\rangle$ die Symmetrieachsen von bestimmten Flächen im $\mathbb{R}^n$ beschreiben. Dies sind die sogenannten *quadratischen Hyperflächen* oder *Quadriken*, die durch eine symmetrische Bilinearform oder, dazu äquivalent, eine quadratische Form auf dem $\mathbb{R}^n$ definiert werden (vgl. Bemerkung 10.3.7).

Definition 11.3.14: Sei $\beta : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ eine symmetrische Bilinearform auf dem $\mathbb{R}^n$. Dann ist die durch β definierte **Quadrik** oder **quadratische Hyperfläche** die Menge

$$M_\beta = \{x \in \mathbb{R}^n \,|\, \beta(x, x) = 1\} \subseteq \mathbb{R}^n.$$

Ist ${}^B\beta^B = (q_{ij})$ die beschreibende Matrix einer symmetrischen Bilinearform β auf dem $\mathbb{R}^n$ für die Standardbasis $B = (e_1, ..., e_n)$, so lautet die Gleichung für die Punkte in M_β

$$\beta(x, x) = x^T \cdot {}^B\beta^B \cdot x = \Sigma_{i,j=1}^n q_{ij} x_i x_j = 1 \quad \text{mit} \quad q_{ij} = \beta(e_i, e_j).$$

Da β symmetrisch ist, können wir diese Gleichung mit Korollar 11.3.13 vereinfachen und die quadratischen Hyperflächen im $\mathbb{R}^n$ weitgehend klassifizieren.

Korollar 11.3.15: Sei $\beta : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ eine symmetrische Bilinearform. Dann gibt es eine orthogonale Matrix $O \in \mathrm{O}(n, \mathbb{R})$ und $d_{11}, d_{22}, \ldots, d_{nn} \in \mathbb{R}$, so dass

$$M_\beta = \{O \cdot x \mid x \in \mathbb{R}^n \text{ mit } \Sigma_{i=1}^n d_{ii} x_i^2 = 1\}.$$

Beweis:
Nach Korollar 11.3.13 gibt es eine Orthonormalbasis C des $\mathbb{R}^n$, so dass ${}^C\beta^C = D$ diagonal ist. Die Abbildung, die die Standardbasis B auf C abbildet, ist orthogonal, da beide Orthonormalbasen sind, und wird damit durch eine orthogonale Matrix $O \in \mathrm{O}(n, \mathbb{R})$ mit ${}^B\beta^B = O \cdot D \cdot O^T$ beschrieben. Daraus folgt

$$\begin{aligned} M_\beta &= \{x \in \mathbb{R}^n |\, x^T \cdot {}^B\beta^B \cdot x = 1\} = \{x \in \mathbb{R}^n |\, x^T \cdot O \cdot D \cdot O^T \cdot x = 1\} \\ &= \{x \in \mathbb{R}^n |\, (O^T x)^T \cdot D \cdot (O^T x) = 1\} = \{Oy \in \mathbb{R}^n | y^T \cdot D \cdot y = 1\} = \{Oy \in \mathbb{R}^n | \Sigma_{i=1}^n d_{ii} y_i^2 = 1\}, \end{aligned}$$

wobei $y = O^T \cdot x$ substituiert wurde und ausgenutzt wurde, dass $O \cdot O^T = \mathbb{1}_n$ gilt. □

Damit ist jede Quadrik M_Q das Bild einer durch eine Diagonalmatrix definierten Quadrik

$$M_D = \{x \in \mathbb{R}^n | x^T \cdot D \cdot x = 1\} = \{x \in \mathbb{R}^n |\, \Sigma_{i=1}^n d_{ii} x_i^2 = 1\} \tag{24}$$

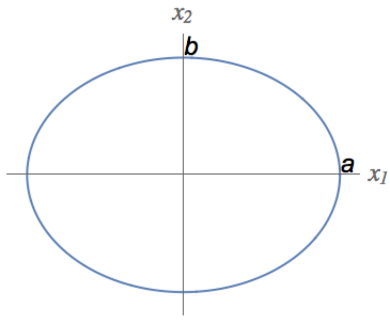
unter einer orthogonalen Abbildung $\phi_O : \mathbb{R}^n \to \mathbb{R}^n$, $v \mapsto O \cdot v$. Bis auf lineare Abbildungen, die Längen und Winkel erhalten, sind also alle Quadriken im $\mathbb{R}^n$ durch Diagonalmatrizen D und die Gleichung in (24) gegeben. Damit können wir die Quadriken auf dem $\mathbb{R}^n$ bis auf orthogonale Automorphismen klassifizieren, indem wir die Quadriken aus Gleichung (24) betrachten.

Da sich jede Diagonalmatrix D durch Streckungen in Richtung der x_i-Achsen überführen lässt in eine Diagonalmatrix D mit $d_{ii} \in \{0, \pm 1\}$ kommt es bei dieser Klassifikation im Wesentlichen nur auf das Vorzeichen der Diagonaleinträge von D an, also auf die Vorzeichen der Eigenwerte der beschreibenden Matrix ${}^B\beta^B$. Wir führen dies für die Fälle $n = 2$ und $n = 3$ durch.

$n = 2$: quadratische Kurven

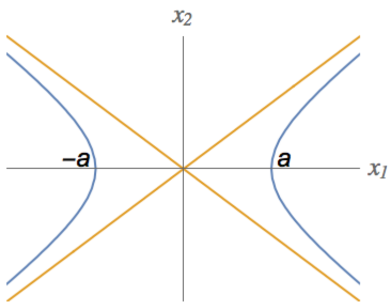
(a) $d_{11} > 0$, $d_{22} > 0$: Wir setzen $a := 1/\sqrt{d_{11}}$, $b = 1/\sqrt{d_{22}}$ und erhalten die Gleichung einer **Ellipse** mit den **Halbachsen** a, b

$$\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} = 1.$$



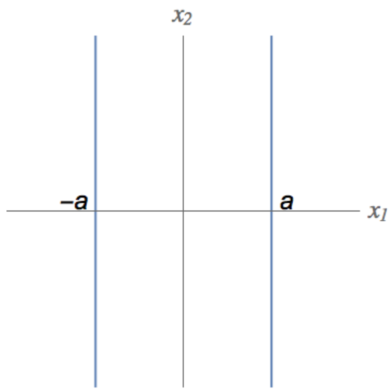
(b) $d_{11} > 0$, $d_{22} < 0$: Wir setzen $a := 1/\sqrt{d_{11}}$, $b = 1/\sqrt{-d_{22}}$ und erhalten die Gleichung einer **Hyperbel** mit den **Asymptoten** $x_2 = \pm\frac{b}{a} \cdot x_1$

$$\frac{x_1^2}{a^2} - \frac{x_2^2}{b^2} = 1.$$



(c) $d_{11} > 0$, $d_{22} = 0$: Wir setzen $a := 1/\sqrt{d_{11}}$ und erhalten die Gleichungen zweier zur x_2-Achse paralleler **Geraden**

$$x_1 = \pm a, \quad x_2 \in \mathbb{R}.$$

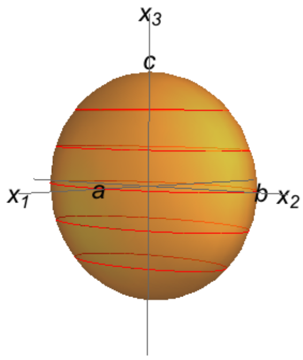


(d) $d_{11} \leq 0$, $d_{22} \leq 0$: Die Gleichung (24) der Quadrik hat keine reelle Lösung.

$n = 3$: quadratische Flächen

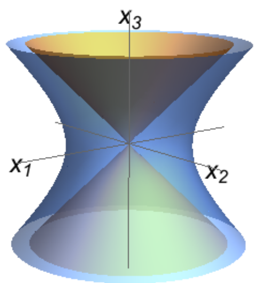
(a) $d_{11} > 0$, $d_{22} > 0$, $d_{33} > 0$: Wir setzen $a := 1/\sqrt{d_{11}}$, $b = 1/\sqrt{d_{22}}$ und $c = 1/\sqrt{d_{33}}$ und erhalten die Gleichung eines **Ellipsoids** mit den **Halbachsen** a, b, c

$$\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} + \frac{x_3^2}{c^2} = 1.$$



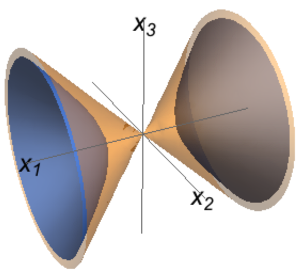
(b) $d_{11} > 0$, $d_{22} > 0$, $d_{33} < 0$: Wir setzen $a := 1/\sqrt{d_{11}}$, $b = 1/\sqrt{d_{22}}$ und $c = 1/\sqrt{-d_{33}}$ und erhalten die Gleichung eines **einschaligen Hyperboloids**

$$\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} - \frac{x_3^2}{c^2} = 1.$$



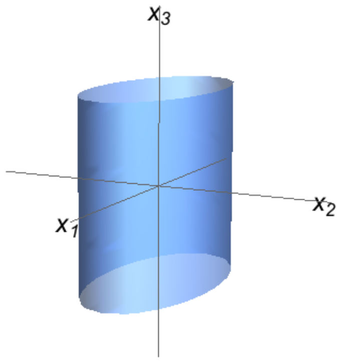
(c) $d_{11} > 0$, $d_{22} < 0$, $d_{33} < 0$: Wir setzen $a := 1/\sqrt{d_{11}}$, $b = 1/\sqrt{-d_{22}}$ und $c = 1/\sqrt{-d_{33}}$ und erhalten die Gleichung eines **zweischaligen Hyperboloids**

$$\frac{x_1^2}{a^2} - \frac{x_2^2}{b^2} - \frac{x_3^2}{c^2} = 1.$$



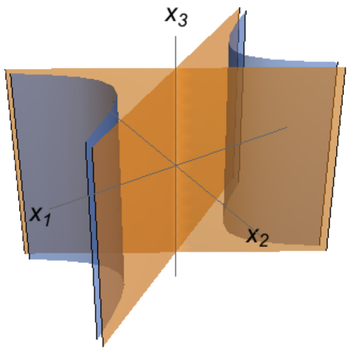
(d) $d_{11} > 0$, $d_{22} > 0$, $d_{33} = 0$: Wir setzen $a := 1/\sqrt{d_{11}}$, $b = 1/\sqrt{d_{22}}$ und erhalten die Gleichung eines **elliptischen Zylinders**

$$\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} = 1, \quad x_3 \in \mathbb{R}$$



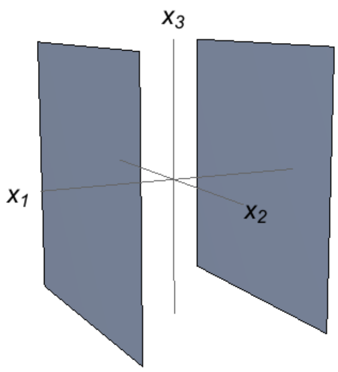
(e) $d_{11} > 0$, $d_{22} < 0$, $d_{33} = 0$: Wir setzen $a := 1/\sqrt{d_{11}}$, $b = 1/\sqrt{-\lambda_2}$ und erhalten die Gleichung eines **hyperbolischen Zylinders**

$$\frac{x_1^2}{a^2} - \frac{x_2^2}{b^2} = 1, \quad x_3 \in \mathbb{R}.$$



(f) $d_{11} > 0$, $d_{22} = d_{33} = 0$: Wir setzen $a := 1/\sqrt{d_{11}}$ und erhalten die Gleichungen zweier **Ebenen** parallel zur x_2x_3-Ebene

$$x_1 = \pm a, \quad x_2, x_3 \in \mathbb{R}.$$



(g) $d_{11} \leq 0$, $d_{22} \leq 0$, $d_{33} \leq 0$: Die Gleichung (24) der Quadrik hat keine reelle Lösung.

Mit dieser Betrachtung haben wir die durch symmetrische 2×2 und 3×3-Matrizen definierten Quadriken vollständig klassifiziert. Um für eine symmetrische Matrix $Q \in \mathrm{Mat}(n \times n, \mathbb{R})$ mit $n \in \{2, 3\}$ den Typ der zugehörigen Quadrik zu bestimmen, müssen lediglich die *Vorzeichen der Eigenwerte* von Q bestimmt werden. Der Typ der Quadrik lässt sich dann aus der folgenden Tabelle ablesen.

	Vorzeichen der Eigenwerte von Q	Typ der Quadrik
$n=2$	$(+,+)$	Ellipse
	$(+,-)$	Hyperbel
	$(+,0)$	2 parallele Geraden
	$(-,0)$, $(0,0)$	$\emptyset$
$n=3$	$(+,+,+)$	Ellipsoid
	$(+,+,-)$	einschaliges Hyperboloid
	$(+,-,-)$	zweischaliges Hyperboloid
	$(+,+,0)$	elliptischer Zylinder
	$(+,-,0)$	hyperbolischer Zylinder
	$(+,0,0)$	2 parallele Ebenen
	$(0,0,0)$, $(-,0,0)$, $(-,-,0)$, $(-,-,-)$	$\emptyset$

An was man sich erinnern sollte:

Die wichtigsten Begriffe und Definitionen:

- adjungierter Endomorphismus, adjungierte Abbildung
- normaler Endomorphismus, selbstadjungierter Endomorphismus
- normale Matrix
- (anti)selbstadjungierte bzw. (anti)hermitesche, (anti)symmetrische Matrix
- Quadrik, quadratische Hyperfläche, quadratische Fläche, quadratische Kurve
- Ellipse, Hyperbel,
- Ellipsoid, ein- und zweischaliges Hyperboloid, elliptischer und hyperbolischer Zylinder

Die wichtigsten Aussagen:

- Darstellungssatz von Fréchet-Riesz
- Eigenschaften der Adjungierten, beschreibende Matrix der Adjungierten,
- Bilder und Kerne von Adjungierten für euklidische und unitäre Vektorräume,
- Charakterisierung von Unitarität, Normalität, Selbstadjungiertheit mit Adjungierter,
- Charakterisierung von normalen und selbstadjungierten Endomorphismen mit Matrizen
- Eigenwerte und Eigenvektoren von normalen Endomorphismen
- Diagonalisierbarkeitsaussagen für normale, unitäre und selbstadjungierte Endomorphismen und Matrizen über $\mathbb{C}$ und über $\mathbb{R}$
- Hauptachsentransformation
- Klassifikation von Quadriken

Die wichtigsten Beispiele:

- $\phi^\dagger \circ \phi$ und $\phi \circ \phi^\dagger$ sind selbstadjungiert
- ϕ normal $\Rightarrow \lambda\phi$ und $\phi - \lambda\mathrm{id}$ normal
- ϕ normal, ψ unitär $\Rightarrow \psi \circ \phi \circ \psi^{-1}$ normal
- beschreibende Matrizen von symmetrischen Bilinearformen
- beschreibende Matrizen von hermiteschen Sesquilinearformen

12 Konstruktionen mit Vektorräumen

In diesem Kapitel befassen wir uns systematisch damit, aus gegebenen $\mathbb{K}$-Vektorräumen neue $\mathbb{K}$-Vektorräume zu konstruieren. Wir lernen diese Konstruktionen zwar im Kontext von Vektorräumen kennen, aber es gibt analoge Konstruktionen auch für Mengen, topologische Räume, (abelsche) Gruppen, (unitale) Ringe, Algebren etc.

Diese erhält man, indem man in den Konstruktionen für Vektorräume den Begriff des Vektorraums durch Menge, topologischen Raum, (abelsche) Gruppe, (unitalen) Ring, Algebra etc. ersetzt und den Begriff der linearen Abbildung durch Abbildung, stetige Abbildung, Gruppenhomomorphismus, (unitalen) Ringhomomorphismus, Algebrahomomorphismus etc. In manchen dieser Fälle muss man auch noch bestimmte Zusatzvoraussetzungen fordern.

Man kann sich das wie eine Art Grammatik für mathematische Strukturen vorstellen, die hier für die Sprache der Vektorräume untersucht wird. Auch die Sprachen der Mengen, topologischen Räume, (abelschen) Gruppen, (unitalen) Ringe, Algebren besitzen eine Grammatik, die sowohl Gemeinsamkeiten mit der der Vektorräume aufweist als auch Besonderheiten.

Der zentrale Punkt bei diesen Konstruktionen ist es, sie systematisch und übersichtlich zu formulieren, so dass man sich nicht in einer Fülle von Details verliert, sondern die zugrundeliegenden Prinzipien klar erkennt. Dabei ist es nützlich, die Konstruktionen für Vektorräume durch die Eindeutigkeit und Existenz gewisser linearer Abbildungen zu charakterisieren, die sogenannten *universellen Eigenschaften* dieser Konstruktionen.

Dies ist einerseits deswegen sinnvoll, weil man letztendlich nie isolierte Vektorräume sondern immer auch lineare Abbildungen zwischen ihnen betrachten will. Andererseits führt dies zu einer übersichtlicheren Beschreibung, bei der die wesentlichen Eigenschaften der Konstruktionen klar hervortreten und sich leicht auf andere mathematische Strukturen verallgemeinern lassen. Zudem ersparen es einem die universellen Eigenschaften, immer wieder die gleichen Rechnungen durchzuführen, um die Wohldefiniertheit gewisser Abbildungen zu beweisen.

12.1 Quotientenräume

Die erste Konstruktion, die wir kennenlernen ist der *Quotientenraum* - das Vektorraumanalogon der Restklassenringe $\mathbb{Z}/n\mathbb{Z}$. Das Prinzip dieser Konstruktion ist analog zur Konstruktion von $\mathbb{Z}/n\mathbb{Z}$, nur dass wir statt des Ideals $n\mathbb{Z} \subseteq \mathbb{Z}$ einen Untervektorraum $U \subseteq V$ eines $\mathbb{K}$-Vektorraums V betrachten. Analog zur Konstruktion des Restklassenrings $\mathbb{Z}/n\mathbb{Z}$, wo zwei Zahlen $z, w \in \mathbb{Z}$ als äquivalent betrachtet wurden, wenn ihre Differenz in $n\mathbb{Z}$ liegt, führen wir eine Äquivalenzrelation ein, bei der zwei Vektoren als äquivalent betrachtet werden, wenn ihre Differenz im Untervektorraum U liegt. Die Bedingung, dass $U \subseteq V$ ein Untervektorraum ist, garantiert nicht nur, dass wir so eine Äquivalenzrelation auf V erhalten, sondern auch dass die zugehörige Quotientenmenge die Struktur eines $\mathbb{K}$-Vektorraums erhält.

Satz 12.1.1: Sei V ein Vektorraum über einem Körper $\mathbb{K}$ und $U \subseteq V$ ein Untervektorraum.

1. Die Relation $v \sim w \Leftrightarrow v - w \in U$ ist eine Äquivalenzrelation auf V.
2. Die Quotientenmenge $V/U = \{[v] \,|\, v \in V\} = \{v+U | v \in V\}$ bildet mit den Verknüpfungen
$$+ : V/U \times V/U \to V/U,\ ([v],[w]) \mapsto [v+w] \qquad \cdot : \mathbb{K} \times V/U \to V/U,\ (\lambda,[v]) \mapsto [\lambda v]$$
einen Vektorraum über $\mathbb{K}$, den **Quotientenraum** von V bezüglich U.

Beweis:
1. **Reflexivität:** da U ein Untervektorraum ist, gilt $v - v = 0 \in U$, also $v \sim v$ für alle $v \in V$.
Symmetrie: Sind $v, w \in V$ mit $v \sim w$, so folgt $v - w \in U$. Da U ein Untervektorraum ist, auch $w - v = (-1)(v - w) \in U$, also $w \sim v$.
Transitivität: Sind $u, v, w \in V$ mit $u \sim v$ und $v \sim w$, so gilt $u - v \in U$, $v - w \in U$. Da U ein Untervektorraum ist, folgt $u - w = (u - v) + (v - w) \in U$ und damit $u \sim w$.
Also ist $\sim$ eine Äquivalenzrelation auf V.

2. Zu zeigen ist zunächst, dass $+$ und $\cdot$ wohldefiniert sind:

Sind $v, v', w, w' \in V$ mit $[v] = [v']$ und $[w] = [w']$, so gilt $v \sim v'$ und $w \sim w'$, also $v - v' \in U$ und $w - w' \in U$. Da U ein Untervektorraum ist, folgt daraus $(v+w)-(v'+w') = v-v'+w-w' \in U$, also $v + w \sim v' + w'$ und damit $[v] + [w] = [v + w] = [v' + w'] = [v'] + [w']$. Analog folgt $\lambda v - \lambda v' = \lambda(v - v') \in U$ für alle $\lambda \in \mathbb{K}$, also $\lambda v \sim \lambda v'$ und $\lambda[v] = [\lambda v] = [\lambda v'] = \lambda[v']$.

Die Vektorraumaxiome für V/U ergeben sich direkt aus den Vektorraumaxiomen für V:

Zunächst ist $(V/U, +)$ eine abelsche Gruppe, denn für alle $u, v, w \in V$ gilt

$$[u] + ([v] + [w]) = [u] + [v + w] = [u + (v + w)] = [(u + v) + w] = [u + v] + [w] = ([u] + [v]) + [w]$$
$$[v] + [0] = [v + 0] = [v] = [0 + v] = [0] + [v]$$
$$[v] + [-v] = [v + (-v)] = [0] = [(-v) + v] = [-v] + [v]$$
$$[v] + [w] = [v + w] = [w + v] = [w] + [v].$$

Analog folgen die Vektorraumaxiome (VR1) - (VR4): für alle $u, v, w \in V$, $\lambda, \mu \in \mathbb{K}$

(VR1) $\lambda \cdot (\mu \cdot [v]) = \lambda \cdot [\mu v] = [\lambda(\mu v)] = [(\lambda\mu)v] = (\lambda\mu) \cdot [v]$
(VR2) $(\lambda + \mu) \cdot [v] = [(\lambda + \mu) \cdot v] = [\lambda \cdot v + \mu \cdot v] = [\lambda v] + [\mu v] = \lambda \cdot [v] + \mu \cdot [v]$
(VR3) $\lambda \cdot ([v] + [w]) = \lambda \cdot [v + w] = [\lambda(v + w)] = [\lambda v + \lambda w] = [\lambda v] + [\lambda w] = \lambda \cdot [v] + \lambda \cdot [w]$
(VR4) $1 \cdot [v] = [1 \cdot v] = [v]$.

$\square$

Per Definition der Äquivalenzrelation in Satz 12.1.1 sind ihre Äquivalenzklassen affine Unterräume von V. Der Vektorraum V/U besteht also aus affinen Unterräumen $v + U$, die addiert und mit Skalaren multipliziert werden, indem man die zugehörigen Vektoren v addiert und mit Skalaren multipliziert. Obwohl diese Vektoren nicht eindeutig sind, hängt der resultierende Untervektorraum nicht von ihrer Wahl ab, wie in Abbildung 5 dargestellt.

Im Allgemeinen möchte man natürlich nicht nur isolierte Quotientenräume V/U betrachten, sondern auch lineare Abbildungen in Quotientenräume und aus Quotientenräumen. Erstere lassen sich durch Verkettung von linearen Abbildungen $f : W \to V$ mit der *kanonischen Surjektion* $\pi : V \to V/U$, $v \mapsto [v]$ konstruieren. Man erhält die lineare Abbildung $\pi \circ f : W \to V/U$, $w \mapsto [f(w)]$.

Um aus einer linearen Abbildung $\phi : V \to W$ eine lineare Abbildung $\tilde{\phi} : V/U \to W$ zu konstruieren, möchte man $\tilde{\phi}([v]) := \phi(v)$ für alle $v \in V$ setzen, also $\tilde{\phi} \circ \pi = \phi$ fordern. Da die linke Seite dieser Gleichung auf U verschwindet, kann das nur dann funktionieren, wenn $U \subseteq \ker(\phi)$ gilt. Da man mit Äquivalenzklassen arbeitet, ist hier jedes Mal noch Wohldefiniertheit von $\tilde{\phi}$ zu zeigen. Um hier nicht immer wieder die selben Beweise durchzuführen, bietet es sich an, die Existenz und Eindeutigkeit solcher linearer Abbildungen systematisch zu untersuchen. Das führt auf die *universelle Eigenschaft* des Quotientenraums.

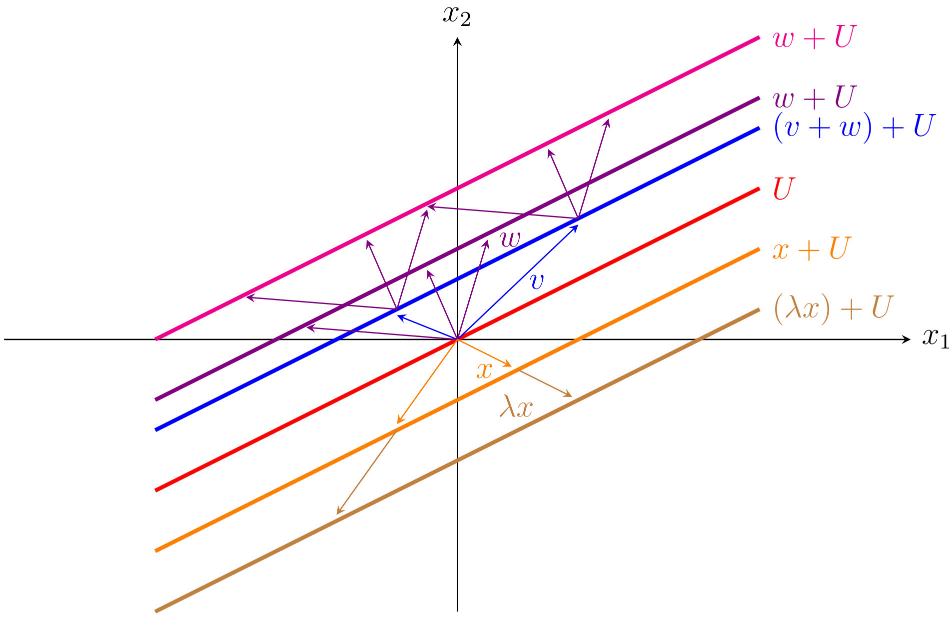


Abbildung 5: Die Unabhängigkeit der Vektorraumstruktur auf V/U von den Repräsentanden.

Satz 12.1.2: (**universelle Eigenschaft des Quotientenraums**)
Sei V ein Vektorraum über $\mathbb{K}$ und $U \subseteq V$ ein Untervektorraum. Dann gilt:

1. Die kanonische Surjektion $\pi : V \to V/U$, $v \mapsto [v]$ ist eine surjektive $\mathbb{K}$-lineare Abbildung.
2. Zu jeder $\mathbb{K}$-linearen Abbildung $\phi : V \to W$ mit $U \subseteq \ker(\phi)$ existiert genau eine $\mathbb{K}$-lineare Abbildung $\tilde{\phi} : V/U \to W$ mit $\tilde{\phi} \circ \pi = \phi$.

$$\begin{tikzcd} V \arrow[r, "\phi", "U\subseteq\ker(\phi)"'] \arrow[d, "\pi"'] & W \\ V/U. \arrow[ur, dashed, "\exists!\tilde{\phi}"'] & \end{tikzcd}$$

3. $\tilde{\phi}$ ist injektiv genau dann, wenn $\ker(\phi) = U$.

Beweis:
1. Die Abbildung $\pi : V \to V/U$, $v \mapsto [v]$ ist surjektiv per Definition der Quotientenmenge und linear per Definition der Vektorraumstruktur auf V/U, denn es gilt für alle $v, w \in V$, $\lambda \in \mathbb{K}$

$$\pi(v+w) = [v+w] = [v] + [w] = \pi(v) + \pi(w) \qquad \pi(\lambda v) = [\lambda v] = \lambda \cdot [v] = \lambda\pi(v).$$

2. **Eindeutigkeit:** Ist $\phi : V \to W$ eine $\mathbb{K}$-lineare Abbildung und sind $\tilde{\phi}_1, \tilde{\phi}_2 : V/U \to W$ zwei $\mathbb{K}$-lineare Abbildungen mit $\tilde{\phi}_1 \circ \pi = \tilde{\phi}_2 \circ \pi = \phi$, so folgt $(\tilde{\phi}_1 - \tilde{\phi}_2) \circ \pi = \tilde{\phi}_1 \circ \pi - \tilde{\phi}_2 \circ \pi = \phi - \phi = 0$. Da $\pi : V \to V/U$ surjektiv ist, folgt $\tilde{\phi}_1 - \tilde{\phi}_2 = 0$.

Existenz: Sei $\phi : V \to W$ eine $\mathbb{K}$-lineare Abbildung mit $U \subseteq \ker(\phi)$. Sind $v, v' \in V$ mit $v \sim v'$, so ist $v - v' \in U$ und damit $\phi(v) - \phi(v') = \phi(v - v') = 0$, da $U \subseteq \ker(\phi)$. Also folgt aus $[v] = [v']$ auch $\phi(v) = \phi(v')$ und die Abbildung $\tilde{\phi} : V/U \to W$, $[v] \mapsto \phi(v)$ ist wohldefiniert. Sie ist außerdem linear, denn es gilt für alle $v, w \in V$ und $\lambda \in \mathbb{K}$

$$\tilde{\phi}([v] + [w]) = \tilde{\phi}([v+w]) = \phi(v+w) = \phi(v) + \phi(w) = \tilde{\phi}([v]) + \tilde{\phi}([w])$$
$$\tilde{\phi}(\lambda \cdot [v]) = \tilde{\phi}([\lambda v]) = \phi(\lambda v) = \lambda\phi(v) = \lambda\tilde{\phi}([v]).$$

3. Ist $x \in \ker(\tilde{\phi})$, so gibt es wegen der Surjektivität von π ein $v \in V$ mit $x = [v]$. Es folgt $0 = \tilde{\phi}([v]) = (\tilde{\phi} \circ \pi)(v) = \phi(v)$, also $x = [v]$ mit $v \in \ker(\phi)$. Ist $\ker(\phi) = U$, so impliziert das $x = [v] = [0]$, und $\tilde{\phi}$ ist injektiv. Ist umgekehrt $\tilde{\phi}$ injektiv, so folgt aus $\phi(v) = \tilde{\phi}([v]) = 0$, dass $x = [v] = 0$ gilt, also $v \in U$, und damit $\ker(\phi) = U$. □

Bemerkung 12.1.3: Die universelle Eigenschaft des Quotientenraums V/U ist äquivalent zu der Aussage, dass für jeden $\mathbb{K}$-Vektorraum W die $\mathbb{K}$-Vektorräume $\mathrm{Hom}_{\mathbb{K}}(V/U, W)$ und $\mathrm{Hom}_U(V, W) = \{f \in \mathrm{Hom}_{\mathbb{K}}(V, W) \,|\, U \subseteq \ker(f)\}$ isomorph sind. (Beweis: Übung).

Die universelle Eigenschaft des Quotientenraums bzw. die dazu äquivalente Aussage in Bemerkung 12.1.3 erlaubt es einem, direkt lineare Abbildungen aus Quotientenräumen in andere Vektorräume zu konstruieren, ohne umständliche Berechnungen mit Äquivalenzklassen durchzuführen. Um eine lineare Abbildung $\phi : V/U \to W$ zu konstruieren, reicht es nach Satz 12.1.2 aus, eine lineare Abbildung $\phi : V \to W$ mit $U \subseteq \ker(\phi)$ anzugeben. Ein Beispiel dafür ist der *Homomorphiesatz*, der für jede $\mathbb{K}$-lineare Abbildung $\phi : V \to W$ eine Beziehung zwischen dem Quotientenraum $V/\ker(\phi)$ und dem Vektorraum $\mathrm{im}(\phi)$ herstellt.

Korollar 12.1.4: (**Homomorphiesatz**)
Sei $\phi : V \to W$ eine $\mathbb{K}$-lineare Abbildung und $\pi : V \to V/\ker(\phi)$ die kanonische Surjektion. Dann ist die $\mathbb{K}$-lineare Abbildung $\tilde{\phi} : V/\ker(\phi) \to W$ mit $\tilde{\phi} \circ \pi = \phi$ injektiv, und ihre Koeinschränkung $\tilde{\phi} : V/\ker(\phi) \to \mathrm{im}(\phi)$ ist ein Vektorraumisomorphismus.

Die Darstellung $\phi = \tilde{\phi} \circ \pi$ nennt man **kanonische Faktorisierung** von ϕ.

Beweis:
Nach der universellen Eigenschaft des Quotienten gibt es genau eine lineare Abbildung $\tilde{\phi} : V/\ker(\phi) \to W$ mit $\tilde{\phi} \circ \pi = \phi$ und diese ist nach Satz 12.1.2, 3. injektiv. Damit ist die Koeinschränkung $\tilde{\phi} : V/\ker(\phi) \to \mathrm{im}(\phi)$ ein Isomorphismus. □

Der Bezeichnung *kanonische Faktorisierung* erfasst, dass man für jeden Vektorraumhomomorphismus $\phi : V \to W$ eine Faktorisierung $\phi = \tilde{\phi} \circ \pi$ in einen surjektiven Vektorraumhomomorphismus $\pi : V \to V/U$ und einen injektiven Vektorraumhomomorphismus $\tilde{\phi} : V/U \to W$ erhält, die keine zusätzliche Wahl von Strukturen erfordert und von keinen willkürlichen Wahlen abhängt.

Allgemein ist die universelle Eigenschaft des Quotientenraums nützlich, um Isomorphieaussagen zu beweisen. Man kann mit ihr direkt Isomorphismen zwischen Quotientenräumen konstruieren, ohne Äquivalenzklassen oder Wohldefiniertheit zu betrachten. Beispiele sind die folgenden Aussagen zu Quotienten von direkten Summen und von Quotienten.

Beispiel 12.1.5: Ist ein $\mathbb{K}$-Vektorraum V die direkte Summe $V = U \oplus W$ von Untervektorräumen $U, W \subseteq V$, so induziert die lineare Abbildung $P_W : V \to W$ mit $P_W(u + w) = w$ für $u \in U$, $w \in W$ einen Vektorraumisomorphismus $\tilde{P}_W : V/U \to W$, $[w + u] \mapsto w$.

Beweis:
P_W ist surjektiv und $\ker(P_W) = U$. Satz 12.1.2 liefert den Isomorphismus $\tilde{P}_W : V/U \to W$. □

Beispiel 12.1.6: Sei V ein $\mathbb{K}$-Vektorraum und $T \subseteq U \subseteq V$ Untervektorräume von V. Dann ist $U/T \subseteq V/T$ ein Untervektorraum und $V/U \cong (V/T)/(U/T)$.

Beweis:
Wir betrachten die kanonischen Surjektionen $\pi_{V/T} : V \to V/T$ und $\pi_{V/U} : V \to V/U$. Dann gilt $U/T = \pi_{V/T}(U)$ und damit ist $U/T \subseteq V/T$ ein Untervektorraum als Bild des Untervektorraums $U \subseteq V$ unter der linearen Abbildung $\pi_{V/T}$. Also können wir den Quotienten $(V/T)/(U/T)$ betrachten und seine kanonische Surjektion $\pi_{(V/T)/(U/T)} : V/T \to (V/T)/(U/T)$. Verkettet man die kanonischen Surjektionen $\pi_{V/T}$ und $\pi_{(V/T)/(U/T)}$, so erhält man eine lineare Abbildung $\pi_{(V/T)/(U/T)} \circ \pi_{V/T} : V \to (V/T)/(U/T)$, die als Verkettung zweier surjektiver Abbildungen wieder surjektiv ist. Es gilt

$$\ker(\pi_{(V/T)/(U/T)} \circ \pi_{V/T}) = \{v \in V | \pi_{(V/T)/(U/T)} \circ \pi_{V/T}(v) = 0\} = \{v \in V | \pi_{V/T}(v) \in U/T\} = U.$$

Da $\ker(\pi_{V/U}) = U$ gibt es dann nach Satz 12.1.2 genau einen Vektorraumhomomorphismus $f : V/U \to (V/T)/(U/T)$ mit $f \circ \pi_{V/U} = \pi_{(V/T)/(U/T)} \circ \pi_{V/T}$. Dieser ist wegen der Surjektivität von $\pi_{(V/T)/(U/T)} \circ \pi_{V/T}$ surjektiv und wegen $U = \ker(\pi_{V/U}) = \ker(\pi_{(V/T)/(U/T)} \circ \pi_{V/T})$ nach Satz 12.1.2 auch injektiv, also ein Isomorphismus. □

Ein weiterer wichtiger Vorteil von universellen Eigenschaften ist, dass durch universelle Eigenschaften charakterisierte Strukturen immer *eindeutig bis auf eindeutige Isomorphie* sind. Eindeutigkeit *bis auf eindeutige Isomorphie* ist eine viel stärkere Form der Eindeutigkeit als Eindeutigkeit *bis auf Isomorphie*.

Einelementige Vektorräume bzw. Nullvektorräume sind eindeutig bis auf eindeutige Isomorphie. Denn sind $V = \{v\}$ und $W = \{w\}$ zwei einelementige Vektorräume über $\mathbb{K}$, so existiert genau ein Vektorraumisomorphismus $\phi : V \to W$, $v \mapsto w$. Daher gibt es bis auf eindeutige Isomorphie nur *einen* Vektorraum der Dimension 0, den Nullvektorraum. Der eindeutig bestimmte Isomorphismus entspricht einfach einem Umbenennen des Elements. Die Eindeutigkeit bis auf eindeutige Isomorphie ist der Grund, warum man überhaupt von *dem Nullvektorraum* sprechen kann und nicht von *einem Nullvektorraum*.

Dagegen sind für $n \in \mathbb{N}$ die n-dimensionalen Vektorräume über $\mathbb{K}$ nur *eindeutig bis auf Isomorphie*. Jeder n-dimensionale $\mathbb{K}$-Vektorraum ist isomorph zu $\mathbb{K}^n$. Es handelt sich aber nicht um *eindeutige Isomorphie*, denn es gibt viele verschiedene Isomorphismen zwischen zwei gegebenen n-dimensionalen $\mathbb{K}$-Vektorräumen. Deswegen sprechen wir nicht von *dem*, sondern von *einem* n-dimensionalen $\mathbb{K}$-Vektorraum.

Der Beweis der Eindeutigkeit bis auf eindeutige Isomorphie funktioniert für alle durch universelle Eigenschaften charakterisierte Strukturen analog. Sind zwei Strukturen gegeben, die beide die universelle Eigenschaft besitzen, so benutzt man die universelle Eigenschaft, um eine lineare Abbildung von der ersten in die zweite und eine lineare Abbildung von der zweiten in die erste zu konstruieren. Dann verkettet man diese Abbildungen und zeigt mit der universellen Eigenschaft, dass die resultierenden Abbildungen Identitätsabbildungen sind.

Satz 12.1.7: (**Eindeutigkeit des Quotientenraums**)
Sei V ein Vektorraum über $\mathbb{K}$ und $U \subseteq V$ ein Untervektorraum. Dann ist der Quotient V/U **eindeutig bis auf eindeutige Isomorphie**:

Ist (Q, π') ein Paar aus einem $\mathbb{K}$-Vektorraum Q und einer $\mathbb{K}$-linearen Abbildung $\pi' : V \to Q$ mit $U \subseteq \ker(\pi')$, so dass zu jeder $\mathbb{K}$-linearen Abbildung $\phi : V \to W$ mit $U \subseteq \ker(\phi)$ genau eine $\mathbb{K}$-lineare Abbildung $\tilde{\phi} : Q \to W$ mit $\tilde{\phi} \circ \pi' = \phi$ existiert, so gibt es genau eine $\mathbb{K}$-lineare Abbildung $\tilde{\pi}' : V/U \to Q$ mit $\tilde{\pi}' \circ \pi = \pi'$, und diese ist ein Isomorphismus.

$$\begin{tikzcd} & V \arrow[ld, "\pi"'] \arrow[rd, "\pi'"] & \\ V/U \arrow[rr, dashed, "\cong", "\exists! \tilde{\pi}'"'] & & Q. \end{tikzcd}$$

Beweis:
Nach der universellen Eigenschaft von $(V/U, \pi)$, angewendet auf $\phi = \pi' : V \to Q$ mit $U \subseteq \ker(\pi')$, gibt es genau eine $\mathbb{K}$-lineare Abbildung $\tilde{\pi}' : V/U \to Q$ mit $\tilde{\pi}' \circ \pi = \pi'$. Zu zeigen ist noch, dass sie ein Isomorphismus ist.

Dazu wenden wir die universelle Eigenschaft von (Q, π') auf die lineare Abbildung $\pi : V \to V/U$ mit $\ker(\pi) = U$ an und erhalten eine lineare Abbildung $\tilde{\pi} : Q \to V/U$ mit $\tilde{\pi} \circ \pi' = \pi$

$$\begin{tikzcd} & V \arrow[ld, "\pi'"'] \arrow[rd, "\pi"] & \\ Q \arrow[rr, dashed, "\exists! \tilde{\pi}"'] & & V/U. \end{tikzcd}$$

Verketten wir $\tilde{\pi}'$ und $\tilde{\pi}$, so erhalten wir zwei lineare Abbildungen $\tilde{\pi} \circ \tilde{\pi}' : V/U \to V/U$ und $\tilde{\pi}' \circ \tilde{\pi} : Q \to Q$ mit $\tilde{\pi} \circ \tilde{\pi}' \circ \pi = \tilde{\pi} \circ \pi' = \pi$ und $\tilde{\pi}' \circ \tilde{\pi} \circ \pi' = \tilde{\pi}' \circ \pi = \pi'$

$$\begin{tikzcd} & V \arrow[ld, "\pi"'] \arrow[d, "\pi'"] \arrow[rd, "\pi"] & \\ V/U \arrow[r, "\tilde{\pi}'"] \arrow[rr, bend right, "\mathrm{id}_{V/U}"'] & Q \arrow[r, "\tilde{\pi}"] & V/U \end{tikzcd} \qquad \begin{tikzcd} & V \arrow[ld, "\pi'"'] \arrow[d, "\pi"] \arrow[rd, "\pi'"] & \\ Q \arrow[r, "\tilde{\pi}"] \arrow[rr, bend right, "\mathrm{id}_Q"'] & V/U \arrow[r, "\tilde{\pi}'"] & Q. \end{tikzcd}$$

Nun sind aber auch die Identitätsabbildungen $\mathrm{id}_{V/U} : V/U \to V/U$ und $\mathrm{id}_Q : Q \to Q$ lineare Abbildungen mit $\mathrm{id}_Q \circ \pi' = \pi'$ und $\mathrm{id}_{V/U} \circ \pi = \pi$. Da es nach der universellen Eigenschaft von $(V/U, \pi)$ nur eine $\mathbb{K}$-lineare Abbildung $\tilde{\phi} : V/U \to V/U$ mit $\tilde{\phi} \circ \pi = \pi$ gibt, folgt $\tilde{\pi} \circ \tilde{\pi}' = \mathrm{id}_{V/U}$. Da es nach der universellen Eigenschaft von (Q, π') nur eine $\mathbb{K}$-lineare Abbildung $\tilde{\phi} : Q \to Q$ mit $\tilde{\phi} \circ \pi' = \pi'$ gibt, folgt $\tilde{\pi}' \circ \tilde{\pi} = \mathrm{id}_Q$.

Also gilt $\tilde{\pi} = \tilde{\pi}'^{-1}$ und $\tilde{\pi}' : V/U \to Q$ ist ein $\mathbb{K}$-Vektorraumisomorphismus. □

12.2 Direkte Summen und Produkte

Wir lernen nun zwei weitere wichtige Konstruktionen für Vektorräume kennen, die es einem erlauben aus einer durch eine beliebige Indexmenge I indizierten Familie $(V_i)_{i \in I}$ von $\mathbb{K}$-Vektorräumen V_i zwei neue Vektorräume zu konstruieren, die *(äußere) direkte Summe* und das *Produkt*. Dazu erinnern wir zunächst an zwei schon betrachtete Beispiele, die sich später als Spezialfälle der Produkte und direkten Summen erweisen werden.

In Lemma 3.1.6 wurde gezeigt, dass für jede Menge I und jeden $\mathbb{K}$-Vektorraum V die Menge $\mathrm{Abb}(I, V)$ der Abbildungen $f : I \to V$ mit der punktweisen Addition und Skalarmultiplikation die Struktur eines $\mathbb{K}$-Vektorraums hat. Als Spezialfall dieser Konstruktion kann man insbesondere den Vektorraum $\mathrm{Abb}(I, \mathbb{K})$ betrachten. In einer Übung und in Beispiel 3.1.10, 5. wurde gezeigt, dass die Teilmenge

$$\langle I\rangle_{\mathbb{K}} := \{f : I \to \mathbb{K} | f(i) = 0 \text{ für fast alle } i \in I\} \subseteq \mathrm{Abb}(M, \mathbb{K})$$

ein Untervektorraum von $\mathrm{Abb}(I, \mathbb{K})$ ist. Wie in Beispiel 3.2.5 kann man zeigen, dass die Abbildungen $\delta_i : I \to \mathbb{K}$ mit $\delta_i(i) = 1$ und $\delta_i(j) = 0$ für alle $i \neq j$ eine Basis von $\langle I\rangle_{\mathbb{K}}$ bilden. Die Basiselemente δ_i stehen also in Bijektion mit den Elementen der Menge I. Da dieser Vektorraum im Folgenden eine wichtige Rolle spielen wird, erhält er einen eigenen Namen.

Definition 12.2.1: Sei I eine Menge und $\mathbb{K}$ ein Körper. Der Untervektorraum

$$\langle I\rangle_{\mathbb{K}} = \{f : I \to \mathbb{K} | f(i) = 0 \text{ für fast alle } i \in I\} \subseteq \mathrm{Abb}(I, \mathbb{K})$$

heißt der von I **frei erzeugte** $\mathbb{K}$-Vektorraum. Die Abbildung $\iota : I \to \langle I\rangle_{\mathbb{K}}$, $i \mapsto \delta_i$ wird als die **Inklusionsabbildung** bezeichnet.

Wie der Quotientenraum lässt sich auch der von einer Menge I frei erzeugte $\mathbb{K}$-Vektorraum $\langle I\rangle_{\mathbb{K}}$ durch eine universelle Eigenschaft charakterisieren. In diesem Fall garantiert die universelle Eigenschaft die eindeutige Fortsetzbarkeit von Abbildungen $\phi : I \to V$ in einen $\mathbb{K}$-Vektorraum V zu *linearen* Abbildungen $\tilde{\phi} : \langle I\rangle_{\mathbb{K}} \to V$. Die entsprechenden Aussagen wurden im Prinzip bereits in Satz 4.2.1 bewiesen und werden hier nur noch einmal leicht umformuliert.

Satz 12.2.2: (**universelle Eigenschaft des frei erzeugten Vektorraums**)
Sei I eine Menge und $\mathbb{K}$ ein Körper. Dann gibt es zu jeder Abbildung $\phi : I \to V$ in einen $\mathbb{K}$-Vektorraum V genau eine $\mathbb{K}$-lineare Abbildung $\tilde{\phi} : \langle I\rangle_{\mathbb{K}} \to V$ mit $\tilde{\phi} \circ \iota = \phi$.

$$\begin{tikzcd} I \arrow[r, "\phi"] \arrow[d, "\iota"'] & V \\ \langle I\rangle_{\mathbb{K}} \arrow[ur, dashed, "\exists!\tilde{\phi}"'] & \end{tikzcd}$$

Beweis:
Die Menge $B = \iota(I) = \{\delta_i : I \to \mathbb{K} | i \in I\}$ ist eine Basis von $\langle I\rangle_{\mathbb{K}}$. Nach Satz 4.2.1 gibt es daher zu jeder Abbildung $\phi : B \to V$ genau eine $\mathbb{K}$-lineare Abbildung $\tilde{\phi} : \langle I\rangle_{\mathbb{K}} \to V$ mit $\tilde{\phi}|_B = \phi$. Die Bedingung $\tilde{\phi}|_B = \phi$ ist gleichbedeutend zu $\tilde{\phi} \circ \iota = \phi$. □

Bemerkung 12.2.3: Die universelle Eigenschaft von frei erzeugten Vektorräumen lässt sich auch alleine durch Vektorräume und lineare Abbildungen ausdrücken:

Für jeden $\mathbb{K}$-Vektorraum V sind die $\mathbb{K}$-Vektorräume $\mathrm{Abb}(I, V)$ der Abbildungen $\phi : I \to V$ und $\mathrm{Hom}_{\mathbb{K}}(\langle I\rangle_{\mathbb{K}}, V)$ der $\mathbb{K}$-linearen Abbildungen $\tilde{\phi} : \langle I\rangle_{\mathbb{K}} \to V$ isomorph.

Wie bei allen Konstruktionen, die durch universelle Eigenschaften charakterisiert sind, bestimmt die universelle Eigenschaft den frei erzeugten Vektorraum eindeutig bis auf eindeutige Isomorphie. Der Beweis ist völlig analog zum Beweis für die Eindeutigkeit des Quotientenraums.

Satz 12.2.4: (**Eindeutigkeit des frei erzeugten Vektorraums**)
Sei I eine Menge und $\mathbb{K}$ ein Körper. Dann ist der durch I frei erzeugte $\mathbb{K}$-Vektorraum eindeutig bis auf eindeutige Isomorphie:

Ist W ein $\mathbb{K}$-Vektorraum und $\iota' : I \to W$ eine Abbildung, so dass zu jeder Abbildung $\phi : I \to V$ genau eine $\mathbb{K}$-lineare Abbildung $\tilde{\phi} : W \to V$ mit $\tilde{\phi} \circ \iota' = \phi$ existiert, so gibt es genau eine $\mathbb{K}$-lineare Abbildung $\tilde{\iota}' : \langle I \rangle_{\mathbb{K}} \to W$ mit $\tilde{\iota}' \circ \iota = \iota'$, und diese ist ein Isomorphismus.

$$\begin{tikzcd} & I \arrow[dl, "\iota"'] \arrow[dr, "\iota'"] & \\ \langle I \rangle_{\mathbb{K}} \arrow[rr, dashed, "\cong", "\exists! \tilde{\iota}'"'] & & W. \end{tikzcd}$$

Beweis:
Nach der universellen Eigenschaft des frei erzeugten Vektorraums gibt es zu der Abbildung $\phi = \iota' : I \to W$ genau eine $\mathbb{K}$-lineare Abbildung $\tilde{\iota}' : \langle I \rangle_{\mathbb{K}} \to W$ mit $\tilde{\iota}' \circ \iota = \iota'$

$$\begin{tikzcd} I \arrow[r, "\iota'"] \arrow[d, "\iota"'] & W \\ \langle I \rangle_{\mathbb{K}}. \arrow[ur, dashed, "\exists! \tilde{\iota}'"'] & \end{tikzcd}$$

Zu zeigen ist noch, dass sie ein Isomorphismus ist. Dazu nutzt man aus, dass nach der universellen Eigenschaft von (W, ι') zu der Abbildung $\phi = \iota : I \to \langle I \rangle_{\mathbb{K}}$ genau eine $\mathbb{K}$-lineare Abbildung $\tilde{\iota} : W \to \langle I \rangle_{\mathbb{K}}$ mit $\tilde{\iota} \circ \iota' = \iota$ existiert

$$\begin{tikzcd} I \arrow[r, "\iota'"] \arrow[d, "\iota"'] & W \arrow[dl, dashed, "\exists! \tilde{\iota}"] \\ \langle I \rangle_{\mathbb{K}}. & \end{tikzcd}$$

Durch Verketten erhält man $\mathbb{K}$-lineare Abbildungen $\tilde{\iota} \circ \tilde{\iota}' : \langle I \rangle_{\mathbb{K}} \to \langle I \rangle_{\mathbb{K}}$ und $\tilde{\iota}' \circ \tilde{\iota} : W \to W$ mit $\tilde{\iota} \circ \tilde{\iota}' \circ \iota = \tilde{\iota} \circ \iota' = \iota$ und $\tilde{\iota}' \circ \tilde{\iota} \circ \iota' = \tilde{\iota}' \circ \iota = \iota'$. Nun sind aber auch $\mathrm{id}_{\langle I \rangle_{\mathbb{K}}} : \langle I \rangle_{\mathbb{K}} \to \langle I \rangle_{\mathbb{K}}$ und $\mathrm{id}_W : W \to W$ $\mathbb{K}$-lineare Abbildungen mit $\mathrm{id}_{\langle I \rangle_{\mathbb{K}}} \circ \iota = \iota$ und $\mathrm{id}_W \circ \iota' = \iota'$.

$$\begin{tikzcd} & I \arrow[dl, "\iota"'] \arrow[d, "\iota'"] \arrow[dr, "\iota"] & \\ \langle I \rangle_{\mathbb{K}} \arrow[r, "\tilde{\iota}'"] \arrow[rr, bend right, "\mathrm{id}_{\langle I \rangle_{\mathbb{K}}}"'] & W \arrow[r, "\tilde{\iota}"] & \langle I \rangle_{\mathbb{K}} \end{tikzcd} \qquad \begin{tikzcd} & I \arrow[dl, "\iota'"'] \arrow[d, "\iota"] \arrow[dr, "\iota'"] & \\ W \arrow[r, "\tilde{\iota}"] \arrow[rr, bend right, "\mathrm{id}_W"'] & \langle I \rangle_{\mathbb{K}} \arrow[r, "\tilde{\iota}'"] & W. \end{tikzcd}$$

Nach den universellen Eigenschaften von $(\langle I \rangle_{\mathbb{K}}, \iota)$ und (W, ι') kann es aber jeweils nur eine solche $\mathbb{K}$-lineare Abbildung geben. Also folgt $\tilde{\iota} \circ \tilde{\iota}' = \mathrm{id}_{\langle I \rangle_{\mathbb{K}}}$ und $\tilde{\iota}' \circ \tilde{\iota} = \mathrm{id}_W$ und damit $\tilde{\iota} = \tilde{\iota}'^{-1}$. Also ist $\tilde{\iota}' : \langle I \rangle_{\mathbb{K}} \to W$ ein $\mathbb{K}$-Vektorraumisomorphismus mit $\tilde{\iota}' \circ \iota = \iota'$. □

Auch der Vektorraum $\mathrm{Abb}(I, \mathbb{K})$ besitzt eine universelle Eigenschaft, die ähnlich zu der des frei erzeugten Vektorraum $\langle I \rangle_{\mathbb{K}}$ ist, und ist damit eindeutig bis auf eindeutige Isomorphie. Dies werden wir später als Spezialfall einer allgemeineren Konstruktion untersuchen.

Wir verallgemeinern nun den von einer Menge I frei erzeugten Vektorraum $\langle I\rangle_{\mathbb{K}}$ und den Vektorraum $\mathrm{Abb}(I, \mathbb{K})$, indem wir statt $\mathbb{K}$ eine Familie $(V_i)_{i\in I}$ von $\mathbb{K}$-Vektorräumen V_i betrachten und statt Abbildungen $f : I \to \mathbb{K}$ Abbildungen mit $f(i) \in V_i$ für alle $i \in I$. Um ihren Wertebereich zu definieren, benötigen wir die disjunkte Vereinigung von Mengen.

Ist I eine Indexmenge und $(M_i)_{i\in I}$ eine durch I indizierte Familie von nicht notwendigerweise verschiedenen Mengen M_i, können wir durch eine einfache Konstruktion erreichen, dass wir die Mengen M_i als disjunkt behandeln können. Dazu betrachten wir für $i \in I$ die Produktmengen $M_i \times \{i\}$. Für diese gilt $(M_i \times \{i\}) \cap (M_j \times \{j\}) = \emptyset$ für alle $i \neq j$, auch wenn $M_i \cap M_j \neq \emptyset$. Durch Hinzufügen des Index i haben wir die Mengen M_i also durch paarweise disjunkte Mengen $M_i \times \{i\}$ ersetzt, die in Bijektion mit den Mengen M_i stehen.

Definition 12.2.5: Sei $(M_i)_{i\in I}$ eine durch I indizierte Familie von Mengen. Dann heißt die Menge $\dot\cup_{i\in I} M_i = \cup_{i\in I} M_i \times \{i\}$ die **disjunkte Vereinigung** der Mengen M_i.

Betrachten wir nun eine Familie $(V_i)_{i\in I}$ von $\mathbb{K}$-Vektorräumen V_i, so haben auch die Mengen $V_i \times \{i\}$ die Struktur von $\mathbb{K}$-Vektorräumen mit Vektoraddition und Skalarmultiplikation

$$(v, i) + (w, i) = (v + w, i) \qquad \lambda(v, i) = (\lambda v, i) \qquad \forall v, w \in V_i, \lambda \in \mathbb{K}, i \in I.$$

Abbildungen $f : I \to \dot\cup_{i\in I} V_i$, $i \mapsto (f_i, i)$ und $g : I \to \dot\cup_{i\in I} V_i$, $i \mapsto (g_i, i)$ mit $f_i, g_i \in V_i$ für alle $i \in I$ können dann punktweise addiert und mit Skalaren $\lambda \in \mathbb{K}$ multipliziert werden:

$$f + g : I \to \dot\cup_{i\in I} V_i,\;\; i \mapsto (f_i + g_i, i) \qquad \lambda f : I \to \dot\cup_{i\in I} V_i,\;\; i \mapsto (\lambda f_i, i).$$

Man zeigt durch Nachrechnen der Vektorraumaxiome, dass dies der Menge der Abbildungen $f : I \to \dot\cup_{i\in I} V_i$, $i \mapsto (f_i, i)$ die Struktur eines Vektorraums über $\mathbb{K}$ verleiht. Nachrechnen zeigt auch, dass die Abbildungen mit $f_i = 0$ für fast alle $i \in I$ einen Untervektorraum bilden.

Um die Notation zu vereinfachen, schreiben wir statt $f : I \to \dot\cup_{i\in I} V_i$, $i \mapsto (f_i, i)$ auch $f = (f_i)_{i\in I}$. Damit erhalten wir die Vektoraddition und Skalarmultiplikation

$$(f_i)_{i\in I} + (g_i)_{i\in I} = (f_i + g_i)_{i\in I} \qquad \lambda\,(f_i)_{i\in I} = (\lambda f_i)_{i\in I}. \tag{25}$$

Definition 12.2.6:
Sei $(V_i)_{i\in I}$ eine durch eine Indexmenge I indizierte Familie von $\mathbb{K}$-Vektorräumen.

1. Das **Produkt** der Vektorräume V_i ist die Menge

$$\Pi_{i\in I} V_i = \{(f_i)_{i\in I} \,|\, f_i \in V_i \,\forall i \in I\}$$

mit der Vektoraddition und Skalarmultiplikation (25). Die Abbildungen $\pi_j : \Pi_{i\in I} V_i \to V_j$, $(f_i)_{i\in I} \mapsto f_j$ für $j \in I$ heißen **kanonische Projektionen**.

2. Die **(äußere) direkte Summe** der Vektorräume V_i ist die Menge

$$\oplus_{i\in I} V_i = \{(f_i)_{i\in I} \,|\, f_i \in V_i \,\forall i \in I, f_i = 0 \text{ für fast alle } i \in I\}$$

mit der Vektoraddition und Skalarmultiplikation (25). Die Abbildungen $\iota_j : V_j \to \oplus_{i\in I} V_i$, $v \mapsto (v\delta_{ij})_{i\in I}$ heißen **kanonische Inklusionen**.

Bemerkung 12.2.7:

1. Ist I endlich, so gilt $\oplus_{i\in I} V_i = \Pi_{i\in I} V_i$. Denn jedes Element $(f_i)_{i\in I} \in \Pi_{i\in I} V_i$ erfüllt die Bedingung $f_i = 0$ für fast alle $i \in I$.

2. Ist I unendlich und $V_i \neq \{0\}$ für unendlich viele $i \in I$, so gilt $\oplus_{i\in I} V_i \subsetneq \Pi_{i\in I} V_i$.

   Denn ist B_i eine Basis von V_i, so bilden die Elemente $(b_i\delta_{ij})_{i\in I}$ mit $j \in I$ und $b_i \in B_i$ eine Basis von $\oplus_{i\in I} V_i$ (Übung). Da Linearkombinationen endlich sind, kann aber beispielsweise ein Element $(b_i)_{i\in I} \in \Pi_{i\in V} V_i$ mit $b_i \in B_i$ für alle $i \in I$ nicht als Linearkombination dieser Elemente dargestellt werden.

Die Definition von direkten Summen und Produkten von Vektorräumen mit Hilfe von Abbildungen in die disjunkte Vereinigung der zugrundeliegenden Mengen ist relativ kompliziert und unhandlich, wie sich schon an der Notation zeigt. Insbesondere muss man mit Abbildungen von Abbildungen arbeiten, wenn man Abbildungen in oder aus direkten Summen oder Produkten von Vektorräumen betrachten will. Deswegen ist es effizienter, direkte Summen und Produkte durch eine universelle Eigenschaft zu beschreiben und in Beweisen dann nur noch diese universelle Eigenschaft zu benutzen.

Satz 12.2.8: (**universelle Eigenschaften von Summen und Produkten**)
Sei $(V_i)_{i\in I}$ eine durch eine Indexmenge I indizierte Familie von $\mathbb{K}$-Vektorräumen. Dann gilt:

1. Die Projektionsabbildungen $\pi_j : \Pi_{i\in I} V_i \to V_j$ sind $\mathbb{K}$-linear. Zu jeder Familie $(\phi_i)_{i\in I}$ von $\mathbb{K}$-linearen Abbildungen $\phi_i : U \to V_i$ gibt es genau eine $\mathbb{K}$-lineare Abbildung $\phi : U \to \Pi_{i\in I} V_i$ mit $\pi_j \circ \phi = \phi_j$ für alle $j \in I$.

$$\begin{tikzcd} V_j & U \arrow[l, "\phi_j"'] \arrow[dl, dashed, "\exists!\phi"] \\ \Pi_{i\in I} V_i. \arrow[u, "\pi_j"] \end{tikzcd}$$

   Dies bezeichnet man als die **universelle Eigenschaft des Produkts**.

2. Die Inklusionsabbildungen $\iota_j : V_j \to \oplus_{i\in I} V_i$ sind $\mathbb{K}$-linear. Zu jeder Familie $(\psi_i)_{i\in I}$ von $\mathbb{K}$-linearen Abbildungen $\psi_i : V_i \to W$ gibt es genau eine $\mathbb{K}$-lineare Abbildung $\psi : \oplus_{i\in I} V_i \to W$ mit $\psi \circ \iota_j = \phi_j$ für alle $j \in I$

$$\begin{tikzcd} V_j \arrow[r, "\psi_j"] \arrow[d, "\iota_j"'] & W \\ \oplus_{i\in I} V_i. \arrow[ur, dashed, "\exists!\psi"'] \end{tikzcd}$$

   Dies bezeichnet man als die **universelle Eigenschaft der direkten Summe**.

Beweis:
1. Die $\mathbb{K}$-Linearität der Projektionsabbildungen ergibt sich direkt aus der Definition:

$$\begin{aligned} &\pi_j((f_i)_{i\in I} + (g_i)_{i\in I}) = \pi_j((f_i + g_i)_{i\in I}) = f_j + g_j = \pi_j((f_i)_{i\in I}) + \pi_j((g_i)_{i\in I}) \\ &\pi_j(\lambda(f_i)_{i\in I}) = \pi_j((\lambda f_i)_{i\in I}) = \lambda f_j = \lambda \pi_j((f_i)_{i\in I}) \end{aligned}$$

für alle $(f_i)_{i\in I}, (g_i)_{i\in I} \in \Pi_{i\in I} V_i$, $\lambda \in \mathbb{K}$ und $j \in J$. Ist $(\phi_i)_{i\in I}$ eine Familie von $\mathbb{K}$-linearen Abbildungen $\phi_i : U \to V_i$, so definieren wir $\phi : U \to \Pi_{i\in I} V_i$ durch $\phi(u) = (\phi_i(u))_{i\in I}$. Dann ist

ϕ $\mathbb{K}$-linear, denn es gilt für alle $u, u' \in U$, $\lambda \in \mathbb{K}$

$$\phi(u+u') = (\phi_i(u+u'))_{i\in I} = (\phi_i(u)+\phi_i(u'))_{i\in I} = (\phi_i(u))_{i\in I} + (\phi_i(u'))_{i\in I} = \phi(u)+\phi(u')$$
$$\phi(\lambda u) = (\phi_i(\lambda u))_{i\in I} = (\lambda\phi_i(u))_{i\in I} = \lambda\,(\phi_i(u))_{i\in I} = \lambda\phi(u).$$

Direkt aus der Definition ergibt sich $\pi_j \circ \phi(u) = \pi_j((\phi_i(u))_{i\in I}) = \phi_j(u)$ für alle $u \in U$, $j \in I$.

Ist $g: U \to \Pi_{i\in I} V_i$, $u \mapsto (g_i(u))_{i\in I}$ eine weitere $\mathbb{K}$-lineare Abbildung mit $\pi_j \circ g = \phi_j$ für alle $j \in I$, so folgt $\pi_j \circ g(u) = \pi_j((g_i(u))_{i\in I}) = g_j(u) = \phi_j(u)$ für alle $u \in U$ und damit $g = \phi$.

2. Die $\mathbb{K}$-Linearität der Inklusionsabbildungen ergibt sich direkt aus der Definition, denn es gilt

$$\iota_j(v+v') = ((v+v')\delta_{ij})_{i\in I} = (v\delta_{ij}+v'\delta_{ij})_{i\in I} = (v\delta_{ij})_{i\in I} + (v'\delta_{ij})_{i\in I} = \iota_j(v)+\iota_j(v')$$
$$\iota_j(\lambda v) = ((\lambda v)\delta_{ij})_{i\in I} = \lambda(v\delta_{ij})_{i\in I} = \lambda\iota_j(v)$$

für alle $v, v' \in V_j$, $\lambda \in \mathbb{K}$. Ist $(\psi_i)_{i\in I}$ eine Familie von $\mathbb{K}$-linearen Abbildungen $\psi_i : V_i \to W$, so definieren wir $\psi : \oplus_{i\in I} V_i \to W$ durch $\psi((f_i)_{i\in I}) = \Sigma_{i\in I}\psi_i(f_i)$ für alle $i \in I$. Dann ist ψ wohldefiniert, denn wegen $f_i = 0$ für fast alle $i \in I$ ist die Summe endlich. Außerdem ist ψ linear, denn es gilt für alle $(f_i)_{i\in I}, (g_i)_{i\in I} \in \oplus_{i\in I} V_i$ und $\lambda \in \mathbb{K}$

$$\psi((f_i)_{i\in I} + (g_i)_{i\in I}) = \psi_i((f_i+g_i)_{i\in I}) = \Sigma_{i\in I}\psi_i(f_i+g_i) = \psi((f_i)_{i\in I}) + \psi((g_i)_{i\in I})$$
$$\psi(\lambda(f_i)_{i\in I}) = \psi((\lambda f_i)_{i\in I}) = \Sigma_{i\in I}\lambda\psi_i(f_i) = \lambda\psi((f_i)_{i\in I}).$$

Per Definition ist $\psi \circ \iota_j(v) = \psi((v\delta_{ij})_{i\in I}) = \psi_j(v)$ für alle $v \in V$, also $\psi \circ \iota_j = \psi_j$ für alle $j \in I$.

Ist $g : \oplus_{i\in I} V_i \to W$ eine weitere $\mathbb{K}$-lineare Abbildung mit $g \circ \iota_j = \psi_j$ für alle $j \in I$, so folgt $g \circ \iota_j(v) = g((v\delta_{ij})_{i\in I}) = \psi_j(v)$ für alle $j \in I$ und $v \in V_j$. Da jedes Element in $\oplus_{i\in I} V_i$ eine Linearkombination von Elementen $(v\delta_{ij})_{i\in I}$ mit $v \in V_j$ ist, stimmen g und ψ dmait auf einer basis von $\oplus_{i\in I} V_i$ überein, und es folgt $g = \psi$. □

Bemerkung 12.2.9: Die universellen Eigenschaften von Produkten und direkten Summen von $\mathbb{K}$-Vektorräumen lassen sich äquivalent auch wie folgt formulieren:

1. Für jeden $\mathbb{K}$-Vektorraum U ist $F_U : \mathrm{Hom}_{\mathbb{K}}(U, \Pi_{i\in I} V_i) \to \Pi_{i\in I}\mathrm{Hom}_{\mathbb{K}}(U, V_i)$, $\phi \mapsto (\pi_i \circ \phi)_{i\in I}$ ein Vektorraumisomorphismus.

2. Für jeden $\mathbb{K}$-Vektorraum W ist $G_W : \mathrm{Hom}_{\mathbb{K}}(\oplus_{i\in I} V_i, W) \to \Pi_{i\in I}\mathrm{Hom}_{\mathbb{K}}(V_i, W)$, $\psi \mapsto (\psi\circ\iota_i)_{i\in I}$ ein Vektorraumisomorphismus.

Die universellen Eigenschaften von Produkten und direkten Summen sind offensichtlich völlig analog und gehen auseinander hervor, indem man in den entsprechenden kommutierenden Diagrammen die Pfeile umdreht. Man sagt, die zwei Konstruktionen seien *zueinander dual*.

Wie auch schon im Fall von Quotientenräumen und frei erzeugten Vektorräumen garantiert die universelle Eigenschaft von Produkten und direkten Summen deren Eindeutigkeit bis auf eindeutige Isomorphie. Der Beweis wird wieder völlig analog zu den entsprechenden Beweisen für Quotientenräume und frei erzeugte Vektorräume mit der universellen Eigenschaft geführt.

Satz 12.2.10: (**Eindeutigkeit von Summen und Produkten**)
Sei $\mathbb{K}$ ein Körper und $(V_i)_{i\in I}$ eine durch eine Indexmenge I indizierte Familie von $\mathbb{K}$-Vektorräumen. Dann sind $\Pi_{i\in I} V_i$ und $\oplus_{i\in I} V_i$ eindeutig bis auf eindeutige Isomorphie:

1. Ist X ein $\mathbb{K}$-Vektorraum und $(\pi'_i)_{i\in I}$ ein Familie von $\mathbb{K}$-linearen Abbildungen $\pi'_i : X \to V_i$, so dass zu jeder Familie $(\phi_i)_{i\in I}$ von linearen Abbildungen $\phi_i : U \to V_i$ genau eine lineare Abbildung $\tilde{\phi} : U \to X$ mit $\pi'_j \circ \tilde{\phi} = \phi_j$ für alle $j \in I$ existiert, so gibt es genau eine $\mathbb{K}$-lineare Abbildung $f : X \to \Pi_{i\in I}V_i$ mit $\pi_j \circ f = \pi'_j$ für alle $j \in I$, und diese ist ein Isomorphismus
$$\begin{tikzcd}
& V_j & \\
\Pi_{i\in I}V_i \arrow[ur, "\pi_j"] & & X. \arrow[ul, "\pi'_j"'] \arrow[ll, dashed, "\cong", "\exists! f"']
\end{tikzcd}$$
2. Ist Y ein $\mathbb{K}$-Vektorraum und $(\iota'_i)_{i\in I}$ ein Familie von $\mathbb{K}$-linearen Abbildungen $\iota'_i : V_i \to Y$, so dass zu jeder Familie $(\phi_i)_{i\in I}$ von linearen Abbildungen $\phi_i : V_i \to W$ genau eine lineare Abbildung $\tilde{\phi} : Y \to W$ mit $\tilde{\phi} \circ \iota'_j = \phi_j$ für alle $j \in I$ existiert, so gibt es genau eine $\mathbb{K}$-lineare Abbildung $g : Y \to \oplus_{i\in I}V_i$ mit $g \circ \iota'_j = \iota_j$ für alle $j \in I$, und diese ist ein Isomorphismus
$$\begin{tikzcd}
& V_j \arrow[dl, "\iota_j"'] \arrow[dr, "\iota'_j"] & \\
\oplus_{i\in I}V_i & & Y. \arrow[ll, dashed, "\cong", "\exists! g"']
\end{tikzcd}$$

Beweis:
Wir beweisen die Aussage für Produkte. Der Beweis für Summen ist analog (Übung). Zu der Familie $(\pi'_i)_{i\in I}$ existiert nach der universellen Eigenschaft des Produkts genau eine $\mathbb{K}$-lineare Abbildung $\pi' : X \to \Pi_{i\in I}V_i$ mit $\pi_j \circ \pi' = \pi'_j$ für alle $j \in I$

$$\begin{tikzcd}
V_j & X \arrow[l, "\pi'_j"'] \arrow[dl, dashed, "\exists!\pi'"] \\
\Pi_{i\in I}V_i. \arrow[u, "\pi_j"] &
\end{tikzcd}$$

Umgekehrt ergibt sich aus der universellen Eigenschaft von $(X, (\pi'_i)_{i\in I})$, dass es zu der Familie $(\pi_i)_{i\in I}$ genau eine $\mathbb{K}$-lineare Abbildung $\pi : \Pi_{i\in I}V_i \to X$ mit $\pi'_j \circ \pi = \pi_j$ gibt

$$\begin{tikzcd}
V_j & \Pi_{i\in I}V_i \arrow[l, "\pi_j"'] \arrow[dl, dashed, "\exists!\pi"] \\
X. \arrow[u, "\pi'_j"] &
\end{tikzcd}$$

Verketten liefert $\mathbb{K}$-lineare Abbildungen $\pi \circ \pi' : X \to X$ und $\pi' \circ \pi : \Pi_{i\in I}V_i \to \Pi_{i\in I}V_i$ mit $\pi'_j \circ \pi \circ \pi' = \pi_j \circ \pi' = \pi'_j$ und $\pi_j \circ \pi' \circ \pi = \pi'_j \circ \pi = \pi_j$. Andererseits sind aber auch $\mathrm{id}_X : X \to X$ und $\mathrm{id}_{\Pi_{i\in I}V_i} : \Pi_{i\in I}V_i \to \Pi_{i\in I}V_i$ $\mathbb{K}$-lineare Abbildungen mit $\pi'_j \circ \mathrm{id}_X = \pi'_j$ und $\pi_j \circ \mathrm{id}_{\Pi_{i\in I}V_i} = \pi_j$.

$$\begin{tikzcd}
& V_j & \\
\Pi_{i\in I}V_i \arrow[ur, "\pi_j"] & X \arrow[l, "\pi'"'] \arrow[u, "\pi'_j"] & \Pi_{i\in I}V_i \arrow[l, "\pi"'] \arrow[ul, "\pi_j"'] \arrow[ll, bend left, "\mathrm{id}_{\Pi_{i\in I}V_i}"]
\end{tikzcd}
\qquad
\begin{tikzcd}
& V_j & \\
X \arrow[ur, "\pi'_j"] & \Pi_{i\in I}V_i \arrow[l, "\pi"'] \arrow[u, "\pi_j"] & X \arrow[l, "\pi'"'] \arrow[ul, "\pi'_j"'] \arrow[ll, bend left, "\mathrm{id}_X"]
\end{tikzcd}$$

Aus den universellen Eigenschaften von $\Pi_{i\in I}V_i$ und X folgt $\pi \circ \pi' = \mathrm{id}_X$ und $\pi' \circ \pi = \mathrm{id}_{\Pi_{i\in I}V_i}$. Also gilt $\pi' = \pi^{-1}$ und $f = \pi' : X \to \Pi_{i\in I}V_i$ ist ein Isomorphismus mit $\pi_i \circ f = \pi'_i$. $\square$

Mit den universellen Eigenschaften von Produkten und direkten Summen können wir nun leicht Vektorraumisomorphismen konstruieren. Insbesondere können wir so zeigen, dass Produkte und direkte Summen für die Familie $(\mathbb{K})_{i\in I}$ isomorph zu den $\mathbb{K}$-Vektorräumen $\mathrm{Abb}(I,\mathbb{K})$ und $\langle I\rangle_{\mathbb{K}}$ sind, also tatsächlich diese Vektorräume verallgemeinern.

Satz 12.2.11: Ist $V_i = \mathbb{K}$ für alle $i \in I$, so ist $\Pi_{i\in I}\mathbb{K} \cong \mathrm{Abb}(I,\mathbb{K})$ und $\oplus_{i\in I}\mathbb{K} = \langle I\rangle_{\mathbb{K}}$.

Beweis:
1. Wir betrachten $X = \mathrm{Abb}(I,\mathbb{K})$ mit den Projektionen $\pi'_j : \mathrm{Abb}(I,\mathbb{K}) \to \mathbb{K}$, $f \mapsto f(j)$, die offensichtlich $\mathbb{K}$-linear sind. Zu zeigen ist, dass $(X, (\pi'_i))_{i\in I}$ die universelle Eigenschaft des Produkts aus Satz 12.2.10 hat.

Ist $(\phi_i)_{i\in I}$ eine Familie von $\mathbb{K}$-linearen Abbildungen $\phi_i : U \to \mathbb{K}$, so erhalten wir eine Abbildung $\phi : U \to \mathrm{Abb}(I,\mathbb{K})$, $u \mapsto f_u$, indem wir $f_u(i) = \phi_i(u)$ für alle $i \in I$ setzen. Diese ist $\mathbb{K}$-linear, denn für alle $u, u' \in U$, $\lambda \in \mathbb{K}$ und $i \in I$ gilt

$$\begin{aligned}
&\phi(u+u')(i) = f_{u+u'}(i) = \phi_i(u+u') = \phi_i(u) + \phi_i(u') = f_u(i) + f_{u'}(i) = \phi(u)(i) + \phi(u')(i)\\
&\phi(\lambda u)(i) = f_{\lambda u}(i) = \phi_i(\lambda u) = \lambda\phi_i(u) = \lambda f_u(i) = \lambda\phi(u)(i).
\end{aligned}$$

Außerdem gilt $\pi'_i \circ \phi(u) = \pi_i(f_u) = f_u(i) = \phi_j(u)$ für alle $u \in U$ und $i \in I$ und damit $\pi'_i \circ \phi = \phi_i$ für alle $i \in I$. Ist $\psi : U \to \mathrm{Abb}(I,\mathbb{K})$, $u \mapsto g_u$ eine weitere $\mathbb{K}$-lineare Abbildung mit $\pi'_i \circ \psi = \phi_i$ für alle $i \in I$, so folgt $\pi'_i \circ \psi(u) = g_u(i) = \phi_i(u)$ für alle $u \in U$, $i \in I$ und damit $\psi = \phi$.

Also hat der Vektorraum $\mathrm{Abb}(I,\mathbb{K})$ mit den Projektionsabbildungen π'_j die universelle Eigenschaft des Produkts, und mit Satz 12.2.10 folgt $\mathrm{Abb}(I,\mathbb{K}) \cong \Pi_{i\in\mathbb{N}_0}\mathbb{K}$.

2. Wir betrachten $Y = \langle I\rangle_{\mathbb{K}}$ mit den Inklusionsabbildungen $\iota'_i : \mathbb{K} \to \langle I\rangle_{\mathbb{K}}$, $\lambda \mapsto \lambda\delta_i$, die offensichtlich $\mathbb{K}$-linear sind. Zu zeigen ist, dass $(Y, (\iota'_i))_{i\in I}$ die universelle Eigenschaft der direkten Summe aus Satz 12.2.10 hat.

Ist $(\psi_i)_{i\in I}$ eine Familie von $\mathbb{K}$-linearen Abbildungen $\psi_i : \mathbb{K} \to W$, so können wir eine Abbildung $\psi : \langle I\rangle_{\mathbb{K}} \to W$ definieren, indem wir ihre Werte auf der Basis $B = \{\delta_i : I \to \mathbb{K}, j \mapsto \delta_{ij} \,|\, i \in I\}$ von $\langle I\rangle_{\mathbb{K}}$ vorgeben. Wir setzen $\psi(\delta_i) = \psi_i(1)$ für alle $i \in I$. Dies liefert eine eindeutig bestimmte $\mathbb{K}$-lineare Abbildung $\psi : \langle I\rangle_{\mathbb{K}} \to W$ mit $\psi \circ \iota'_i(\lambda) = \psi(\lambda\delta_i) = \lambda\psi(\delta_i) = \lambda\psi_i(1) = \psi_i(\lambda)$ für alle $\lambda \in \mathbb{K}$ und $i \in I$, also $\psi \circ \iota'_i = \psi_i$ für alle $i \in I$.

Also hat der $\mathbb{K}$-Vektorraum $\langle I\rangle_{\mathbb{K}}$ mit den Inklusionsabbildungen $\iota'_i : \mathbb{K} \to \langle I\rangle_{\mathbb{K}}$ die universelle Eigenschaft der direkten Summe. Mit Satz 12.2.10 folgt $\langle I\rangle_{\mathbb{K}} \cong \oplus_{i\in I}\mathbb{K}$. □

Beispiel 12.2.12:

1. Ist $I = \{1, ..., n\}$ und $V_i = \mathbb{K}$ für alle $i \in I$, so ist $\Pi_{i=1}^n\mathbb{K} = \oplus_{i=1}^n\mathbb{K} \cong \mathbb{K}^n$.

   Die Projektionsabbildungen sind gegeben durch $\pi_j : \mathbb{K}^n \to \mathbb{K}$, $(\lambda_1, ..., \lambda_n)^T \mapsto \lambda_j$ und die $\mathbb{K}$-lineare Abbildung $\phi : U \to \mathbb{K}^n$ mit $\pi_j \circ \phi = \phi_j$ zu einer Familie von $\mathbb{K}$-linearen Abbildungen $\phi_j : U \to \mathbb{K}$ durch $\phi : U \to \mathbb{K}^n$, $u \mapsto (\phi_1(u), ..., \phi_n(u))^T$.

2. Ist $I = \mathbb{N}_0$ und $V_i = \mathbb{K}$ für alle $i \in \mathbb{N}_0$, so ist $\Pi_{i\in I}V_i = \Pi_{n=0}^\infty\mathbb{K}$ isomorph zum Vektorraum $\mathrm{Abb}(\mathbb{N}_0,\mathbb{K})$ der Folgen in $\mathbb{K}$ und $\oplus_{i\in I}V_i = \oplus_{i=0}^\infty\mathbb{K}$ ist isomorph zum $\mathbb{K}$-Vektorraum $\mathbb{K}[x]$ der Polynome mit Koeffizienten in $\mathbb{K}$.

Mit Hilfe der universellen Eigenschaften können wir die (äußeren) direkten Summen aus Definition 12.2.6 mit den (inneren) direkten Summen aus Definition 3.3.1 in Beziehung bringen, also direkten Summen von Untervektorräumen eines $\mathbb{K}$-Vektorraums V. Ist $(V_i)_{i\in I}$ eine Familie von Untervektorräumen eines $\mathbb{K}$-Vektorraums V mit $V_i \cap \mathrm{span}_{\mathbb{K}}(\cup_{j\in I\setminus\{i\}} V_j) = \{0\}$ für alle $i \in I$, so ist die innere direkte Summe nach Satz 3.3.2 gegeben durch $\oplus'_{i\in I} V_i = \mathrm{span}_{\mathbb{K}}(\cup_{i\in I} V_i)$. Mit der universellen Eigenschaften zeigen wir, dass sie isomorph zur (äußeren) direkten Summe aus Definition 12.2.6 ist, was die gemeinsame Bezeichnung *direkte Summe* rechtfertigt.

Satz 12.2.13: Innere direkte Summen und äußere direkte Summen sind isomorph:
Ist V ein $\mathbb{K}$-Vektorraum und $(V_i)_{i\in I}$ eine Familie von Unterräumen $V_i \subseteq V$, so dass für alle $i \in I$ gilt $V_i \cap \mathrm{span}_{\mathbb{K}}(\cup_{j\in I\setminus\{i\}} V_j) = \{0\}$, so ist $\mathrm{span}_{\mathbb{K}}(\cup_{i\in I} V_i) \cong \oplus_{i\in I} V_i$.

Beweis:
Wir betrachten $\oplus'_{i\in I} V_i = \mathrm{span}_{\mathbb{K}}(\cup_{i\in I} V_i)$ und die $\mathbb{K}$-linearen Abbildungen $\iota_i : V_i \to \oplus'_{i\in I} V_i$, $v \mapsto v$. Zu zeigen ist, dass sie die universelle Eigenschaft der direkten Summe haben.

Ist $(\psi_i)_{i\in I}$ eine Familie von $\mathbb{K}$-linearen Abbildungen $\psi_i : V_i \to W$, so erhalten wir eine $\mathbb{K}$-lineare Abbildung $\psi : \oplus_{i\in I} V_i \to W$ mit $\psi \circ \iota_i = \psi_i$, indem wir $\psi(\Sigma_{i=1}^n v_i) = \Sigma_{i=1}^n \psi_i(v_i)$ für alle $n \in \mathbb{N}$ und $v_i \in V_i$ setzen. Diese Abbildung ist wohldefiniert wegen $V_i \cap \mathrm{span}_{\mathbb{K}}(\cup_{j\in I\setminus\{i\}} V_j) = \{0\}$ für alle $i \in I$. Ist umgekehrt $\psi' : \oplus'_{i\in I} V_i \to W$ eine $\mathbb{K}$-lineare Abbildung mit $\psi' \circ \iota_i = \psi_i$, so folgt $\psi'(v_i) = \psi_i(v)$ für alle $v \in V_i$ und $i \in I$ und damit $\psi' = \psi$. Damit besitzt $\oplus'_{i\in I} V_i$ mit den Inklusionsabbildungen $\iota_i : V_i \to \oplus'_{i\in I} V_i$ die universelle Eigenschaft der direkten Summe und nach Satz 12.2.10 gilt $\oplus'_{i\in I} V_i \cong \oplus_{i\in I} V_i$. □

Eine weitere Konsequenz der universellen Eigenschaften von Produkten und direkten Summen ist die Tatsache, dass jede Familie $(\phi_i)_{i\in I}$ von $\mathbb{K}$-linearen Abbildungen $\phi_i : V_i \to W_i$ $\mathbb{K}$-lineare Abbildungen zwischen den Summen und Produkten der Vektorräume V_i und W_i induziert. Dies ergibt sich durch Verketten von $\phi_i : V_i \to W_i$ mit den Projektions- und Inklusionsabbildungen.

Satz 12.2.14: Seien $(V_i)_{i\in I}$ und $(W_i)_{i\in I}$ Familien von $\mathbb{K}$-Vektorräumen und $(\phi_i)_{i\in I}$ eine Familie von $\mathbb{K}$-linearen Abbildungen $\phi_i : V_i \to W_i$.

1. Zu den Produkten $(\Pi_{i\in I} V_i, (\pi_i)_{i\in I})$ und $(\Pi_{i\in I} W_i, (\pi'_i)_{i\in I})$ gibt es genau eine $\mathbb{K}$-lineare Abbildung $\Pi_{i\in I}\phi_i : \Pi_{i\in I} V_i \to \Pi_{i\in I} W_i$ mit $\pi'_j \circ (\Pi_{i\in I}\phi_i) = \phi_j \circ \pi_j$ für alle $j \in I$. Sie heißt **Produkt** der $\mathbb{K}$-linearen Abbildungen ϕ_i.

2. Zu den Summen $(\oplus_{i\in I} V_i, (\iota_i)_{i\in I})$ und $(\oplus_{i\in I} W_i, (\iota'_i)_{i\in I})$ gibt es genau eine $\mathbb{K}$-lineare Abbildung $\oplus_{i\in I}\phi_i : \oplus_{i\in I} V_i \to \Pi_{i\in I} W_i$ mit $(\oplus_{i\in I}\phi_i) \circ \iota_j = \iota'_j \circ \phi_j$ für alle $j \in I$. Sie heißt **direkte Summe** der $\mathbb{K}$-linearen Abbildungen ϕ_i.

Für $I = \{1, ..., n\}$ schreibt man auch $\phi_1 \times ... \times \phi_n := \Pi_{i=1}^n \phi_i$ und $(\phi_1, ..., \phi_n) := \oplus_{i=0}^n \phi_i$.

Beweis:
Die universellen Eigenschaften der Produkte und Summen liefern kommutierende Diagramme

$$\begin{tikzcd} V_j \arrow[r, "\phi_j"] & W_j \\ \Pi_{i\in I} V_i \arrow[u, "\pi_j"] \arrow[ur, "\phi_j\circ\pi_j"] \arrow[r, dashed, "\exists!\Pi_{i\in I}\phi_i"'] & \Pi_{i\in I} W_i \arrow[u, "\pi'_j"'] \end{tikzcd} \qquad \begin{tikzcd} V_j \arrow[r, "\phi_j"] \arrow[d, "\iota_j"'] \arrow[dr, "\iota'_j\circ\phi_j"] & W_j \arrow[d, "\iota'_j"] \\ \oplus_{i\in I} V_i \arrow[r, dashed, "\exists!\oplus_{i\in I}\phi_i"'] & \oplus_{i\in I} W_i. \end{tikzcd}$$

Diese definieren das Produkt und die direkte Summe der $\mathbb{K}$-linearen Abbildungen ϕ_i. □

Ebenso lassen sich mit Hilfe der universellen Eigenschaft Isomorphieaussagen für verschiedene direkte Summen beweisen. Die folgenden Isomorphieaussagen können als Analoga der Eigenschaften eines kommutativen Monoids aus Definition 2.2.6 interpretiert werden, wobei die direkte Summe die Rolle der Verknüpfung einnimmt und die Vektorräume die Rolle der Elemente.

Satz 12.2.15: Seien U, V, W Vektorräume über $\mathbb{K}$. Dann gilt:

1. **Assoziativität der direkten Summe:** $(U \oplus V) \oplus W \cong U \oplus (V \oplus W)$,
2. **Kommutativität der direkten Summe:** $V \oplus W \cong W \oplus V$,
3. **neutrales Element**: $V \oplus \{0\} = \{0\} \oplus V \cong V$

Beweis:
Der Beweis dieser Aussagen ist eine hilfreiche Übung. Wir geben hier nur die Diagramme an, die die Isomorphismen definieren.

(i)
$$\begin{tikzcd}
& U \oplus V \arrow[r, "\iota_{U\oplus V}"] \arrow[dl, dashed, "\exists! g"'] & (U \oplus V) \oplus W \\
U \arrow[ur, "\iota_U"] \arrow[d, "\iota'_U"'] & V \arrow[u, "\iota_V"] \arrow[d, "\iota'_V"'] & W \arrow[u, "\iota_W"'] \arrow[dl, "\iota'_W"] \\
U \oplus (V \oplus W) & V \oplus W \arrow[l, "\iota'_{V\oplus W}"'] \arrow[uur, dashed, "\exists! f"] &
\end{tikzcd}$$

$$\begin{tikzcd}
& U \oplus V \arrow[r, "\iota_{U\oplus V}"] \arrow[ddl, "g"] & (U \oplus V) \oplus W \\
U \arrow[ur, "\iota_U"] \arrow[d, "\iota'_U"'] & & W \arrow[u, "\iota_W"'] \arrow[dl, "\iota'_W"] \\
U \oplus (V \oplus W) \arrow[uurr, dashed, "\exists! \tilde{f}"] & V \oplus W \arrow[l, "\iota'_{V\oplus W}"'] \arrow[uur, "f"] &
\end{tikzcd}
\qquad
\begin{tikzcd}
& U \oplus V \arrow[r, "\iota_{U\oplus V}"] \arrow[ddl, "g"] & (U \oplus V) \oplus W \arrow[ddll, dashed, "\exists! \tilde{g}"] \\
U \arrow[ur, "\iota_U"] \arrow[d, "\iota'_U"'] & & W \arrow[u, "\iota_W"'] \arrow[dl, "\iota'_W"] \\
U \oplus (V \oplus W) & V \oplus W \arrow[l, "\iota'_{V\oplus W}"'] \arrow[uur, "f"] &
\end{tikzcd}$$

(ii)
$$\begin{tikzcd}
& V \oplus W \arrow[dd, dashed, "\exists! f"] & \\
V \arrow[ur, "\iota_V"] \arrow[dr, "\iota'_V"'] & & W \arrow[ul, "\iota_W"'] \arrow[dl, "\iota'_W"] \\
& W \oplus V &
\end{tikzcd}
\qquad
\begin{tikzcd}
& V \oplus W & \\
V \arrow[ur, "\iota_V"] \arrow[dr, "\iota'_V"'] & & W \arrow[ul, "\iota_W"'] \arrow[dl, "\iota'_W"] \\
& W \oplus V \arrow[uu, dashed, "\exists! g"] &
\end{tikzcd}$$

(iii)
$$\begin{tikzcd}
& V \oplus \{0\} \arrow[dd, dashed, "\exists! f"] & \\
V \arrow[ur, "\iota_V"] \arrow[dr, "\mathrm{id}_V"'] & & \{0\} \arrow[ul, "\iota_0"'] \arrow[dl, "0"] \\
& V &
\end{tikzcd}
\qquad
\begin{tikzcd}
& \{0\} \oplus V \arrow[dd, dashed, "\exists! g"] & \\
\{0\} \arrow[ur, "\iota_0"] \arrow[dr, "0"'] & & V \arrow[ul, "\iota_V"'] \arrow[dl, "\mathrm{id}_V"] \\
& V. &
\end{tikzcd}$$

□

12.3 Tensorprodukte

Wir lernen nun eine weitere wichtige Konstruktion mit Vektorräumen kennen, die durch eine universelle Eigenschaft charakterisiert ist, nämlich *Tensorprodukte* von Vektorräumen. Statt Tensorprodukte konkret zu konstruieren und dann die universelle Eigenschaft nachzuweisen, gehen wir diesmal umgekehrt vor. Wir definieren Tensorprodukte über eine universelle Eigenschaft, die dann auch direkt deren Eindeutigkeit bis auf eindeutige Isomorphie beweist. Anschließend weisen wir durch eine konkrete Konstruktion nach, dass tatsächlich ein Tensorprodukt existiert. Letzteres ist essentiell, da die universelle Eigenschaft zwar immer die Eindeutigkeit der dadurch definierten Struktur liefert, aber nie deren Existenz.

Die universelle Eigenschaft des Tensorprodukts setzt *bilineare* Abbildungen $\phi : V \times W \to X$ in einen $\mathbb{K}$-Vektorraum X in Beziehung mit *linearen* Abbildungen $\tilde{\phi} : V \otimes_{\mathbb{K}} W \to X$, wobei $V \otimes_{\mathbb{K}} W$ ein Vektorraum ist, der aus V und W konstruiert wird, das *Tensorprodukt* von V und W. Damit können wir bilineare Abbildungen durch lineare Abbildungen ersetzen.

Definition 12.3.1: Seien V, W Vektorräume über $\mathbb{K}$. Ein Paar (U, τ) aus einem $\mathbb{K}$-Vektorraum U und einer $\mathbb{K}$-bilinearen Abbildung $\tau : V \times W \to U$ heißt **Tensorprodukt** von V und W über $\mathbb{K}$, wenn zu jeder $\mathbb{K}$-bilinearen Abbildung $\phi : V \times W \to X$ in einen $\mathbb{K}$-Vektorraum X genau eine $\mathbb{K}$-lineare Abbildung $\tilde{\phi} : U \to X$ mit $\tilde{\phi} \circ \tau = \phi$ existiert.

$$\begin{tikzcd} V \times W \arrow[r, "\phi"] \arrow[d, "\tau"'] & X \\ U \arrow[ur, dashed, "\exists! \tilde{\phi}"'] & \end{tikzcd}$$

Dies bezeichnet man als die **universelle Eigenschaft des Tensorprodukts**.

Da Tensorprodukte über eine universelle Eigenschaft definiert sind, folgt ihre Eindeutigkeit bis auf eindeutige Isomorphie wieder völlig analog zu der von Quotientenräumen, frei erzeugten Vektorräumen, Produkten und direkten Summen.

Satz 12.3.2: (**Eindeutigkeit von Tensorprodukten**)
Tensorprodukte sind eindeutig bis auf eindeutige Isomorphie: Sind V, W Vektorräume über $\mathbb{K}$ und (U_1, τ_1) und (U_2, τ_2) Tensorprodukte von V und W über $\mathbb{K}$, so gibt es genau eine $\mathbb{K}$-lineare Abbildung $f : U_1 \to U_2$ mit $f \circ \tau_1 = \tau_2$, und diese ist ein Vektorraumisomorphismus.

Beweis:
Da (U_1, τ_1) ein Tensorprodukt von V und W ist, gibt es zu der $\mathbb{K}$-bilinearen Abbildung $\tau_2 : V \times W \to U_2$ genau eine $\mathbb{K}$-lineare Abbildung $\tilde{\tau}_2 : U_1 \to U_2$ mit $\tilde{\tau}_2 \circ \tau_1 = \tau_2$.

$$\begin{tikzcd} V \times W \arrow[r, "\tau_2"] \arrow[d, "\tau_1"'] & U_2 \\ U_1 \arrow[ur, dashed, "\exists! \tilde{\tau}_2"'] & \end{tikzcd}$$

Da auch (U_2, τ_2) ein Tensorprodukt von V und W ist, gibt es zu der $\mathbb{K}$-bilinearen Abbildung $\tau_1 : V \times W \to U_1$ genau eine $\mathbb{K}$-lineare Abbildung $\tilde{\tau}_1 : U_2 \to U_1$ mit $\tilde{\tau}_1 \circ \tau_2 = \tau_1$.

$$\begin{tikzcd} V \times W \arrow[r, "\tau_1"] \arrow[d, "\tau_2"'] & U_1 \\ U_2 \arrow[ur, dashed, "\exists! \tilde{\tau}_1"'] & \end{tikzcd}$$

Durch Verketten von $\tilde{\tau}_1$ und $\tilde{\tau}_2$ erhält man $\mathbb{K}$-lineare Abbildungen $\tilde{\tau}_1 \circ \tilde{\tau}_2 : U_1 \to U_1$ und $\tilde{\tau}_2 \circ \tilde{\tau}_1 : U_2 \to U_2$ mit $\tilde{\tau}_1 \circ \tilde{\tau}_2 \circ \tau_1 = \tilde{\tau}_1 \circ \tau_2 = \tau_1 = \mathrm{id}_{U_1} \circ \tau_1$ und $\tilde{\tau}_2 \circ \tilde{\tau}_1 \circ \tau_2 = \tilde{\tau}_2 \circ \tau_1 = \tau_2 = \mathrm{id}_{U_2} \circ \tau_2$.

$$\begin{tikzcd}
& V \times W \arrow[dl, "\tau_1"'] \arrow[d, "\tau_2"] \arrow[dr, "\tau_1"] & \\
U_1 \arrow[r, "\tilde{\tau}_2"] \arrow[rr, bend right, "\mathrm{id}_{U_1}"'] & U_2 \arrow[r, "\tilde{\tau}_1"] & U_1
\end{tikzcd}
\qquad
\begin{tikzcd}
& V \times W \arrow[dl, "\tau_2"'] \arrow[d, "\tau_1"] \arrow[dr, "\tau_2"] & \\
U_2 \arrow[r, "\tilde{\tau}_1"] \arrow[rr, bend right, "\mathrm{id}_{U_2}"'] & U_1 \arrow[r, "\tilde{\tau}_2"] & U_2.
\end{tikzcd}$$

Da (U_i, τ_i) Tensorprodukte von V und W sind, gibt es aber zu der $\mathbb{K}$-bilinearen Abbildung $\phi_i = \tau_i : V \times W \to U_i$ nur eine $\mathbb{K}$-lineare Abbildung $\tilde{\phi}_i : U_i \to U_i$ mit $\tilde{\phi}_i \circ \tau_i = \tau_i$, nämlich $\tilde{\phi}_i = \mathrm{id}_{U_i}$. Also gilt $\tilde{\tau}_1 \circ \tilde{\tau}_2 = \mathrm{id}_{U_1}$ und $\tilde{\tau}_2 \circ \tilde{\tau}_1 = \mathrm{id}_{U_2}$, also $\tilde{\tau}_1 = \tilde{\tau}_2^{-1}$. Damit ist $f = \tilde{\tau}_2 : U_1 \to U_2$ ein $\mathbb{K}$-Vektorraumisomorphismus. □

Um nun nachzuweisen, dass Tensorprodukte tatsächlich existieren, müssen wir mit Hilfe einer konkreten Konstruktion zu zwei gegebenen Vektorräumen V und W einen Vektorraum U und eine $\mathbb{K}$-bilineare Abbildung $\tau : V \times W \to U$ konstruieren, so dass das Paar (U, τ) die universelle Eigenschaft des Tensorprodukts besitzt. Dazu betrachten wir zunächst den von $V \times W$ frei erzeugten $\mathbb{K}$-Vektorraum und bilden anschließend einen Quotienten, der es uns erlaubt eine $\mathbb{K}$-bilineare Abbildung $\phi : V \times W \to X$ in eine $\mathbb{K}$-lineare Abbildung umzuwandeln.

Satz 12.3.3: (**Existenz von Tensorprodukten**)
Seien V, W Vektorräume über $\mathbb{K}$ und $U \subseteq \langle V \times W \rangle_{\mathbb{K}}$ der von den Vektoren

$$\begin{aligned} &\delta_{(v+v',w)} - \delta_{(v,w)} - \delta_{(v',w)}, &\qquad &\delta_{(v,w+w')} - \delta_{(v,w)} - \delta_{(v,w')} \\ &\delta_{(\lambda v,w)} - \lambda\delta_{(v,w)} & &\delta_{(v,\lambda w)} - \lambda\delta_{(v,w)} \end{aligned} \tag{26}$$

für $v, v' \in V$, $w, w' \in W$, $\lambda \in \mathbb{K}$ aufgespannte Untervektorraum.

Dann ist der Quotientenraum $V \otimes_{\mathbb{K}} W := \langle V \times W \rangle_{\mathbb{K}} / U$ mit $\tau : V \times W \to V \otimes_{\mathbb{K}} W$, $(v, w) \mapsto [\delta_{(v,w)}]$ ein Tensorprodukt von V und W. Man schreibt $v \otimes w := [\delta_{(v,w)}]$ für alle $v \in V$, $w \in W$.

Beweis:
Wir bezeichnen mit $\iota : V \times W \to \langle V \times W \rangle_{\mathbb{K}}$, $(v, w) \mapsto \delta_{(v,w)}$ die Inklusionsabbildung aus Definition 12.2.1 und mit $\pi : \langle V \times W \rangle_{\mathbb{K}} \to \langle V \times W \rangle_{\mathbb{K}} / U$, $x \mapsto [x]$ die kanonische Surjektion aus Satz 12.1.2. Die Abbildung $\tau = \pi \circ \iota : V \times W \to V \otimes_{\mathbb{K}} W$, $(v, w) \mapsto [\delta_{(v,w)}]$ ist $\mathbb{K}$-bilinear, denn per Definition von U gilt für alle $v, v' \in V$, $w \in W$ und $\lambda \in \mathbb{K}$

$$\begin{aligned} \tau(v + v', w) &= [\delta_{(v+v',w)}] = [\delta_{(v+v',w)} - \delta_{(v,w)} - \delta_{(v',w)} + \delta_{(v,w)} + \delta_{(v',w)}] = [\delta_{(v,w)} + \delta_{(v',w)}] \\ &= [\delta_{(v,w)}] + [\delta_{(v',w)}] = \tau(v, w) + \tau(v', w) \\ \tau(\lambda v, w) &= [\delta_{(\lambda v,w)}] = [\delta_{(\lambda v,w)} - \lambda\delta_{(v,w)} + \lambda\delta_{(v,w)}] = [\lambda\delta_{(v,w)}] = \lambda[\delta_{(v,w)}] = \lambda\tau(v, w). \end{aligned}$$

Eine analoge Rechnung für das zweite Argument zeigt, dass auch $\tau(v, w+w') = \tau(v, w) + \tau(v, w')$ und $\tau(v, \lambda w) = \lambda\tau(w)$ für alle $v \in V$, $w, w' \in W$, $\lambda \in \mathbb{K}$ gilt.

Zu jeder Abbildung $\phi : V \times W \to X$ existiert nach der universellen Eigenschaft des frei erzeugten Vektorraums in Satz 12.2.2 genau eine $\mathbb{K}$-lineare Abbildung $\phi' : \langle V \times W \rangle_{\mathbb{K}} \to X$ mit $\phi' \circ \iota = \phi$, was gleichbedeutend ist zu $\phi'(\delta_{(v,w)}) = \phi(v, w)$ für alle $v \in V$, $w \in W$. Ist $\phi : V \times W \to X$ $\mathbb{K}$-bilinear, so ergibt sich für alle $v, v' \in V$, $\lambda \in \mathbb{K}$

$$\begin{aligned} \phi'(\delta_{(v+v',w)} - \delta_{(v,w)} - \delta_{(v',w)}) &= \phi'(\delta_{(v+v',w)}) - \phi'(\delta_{(v,w)}) - \phi'(\delta_{(v',w)}) \\ &= \phi(v + v', w) - \phi(v, w) - \phi(v', w) = 0 \\ \phi'(\delta_{(\lambda v,w)} - \lambda\delta_{(v,w)}) &= \phi'(\delta_{(\lambda v,w)}) - \lambda\phi'(\delta_{(v,w)}) = \phi(\lambda v, w) - \lambda\phi(v, w) = 0. \end{aligned}$$

Eine analoge Rechnung zeigt, dass $\phi'(\delta_{(v,w+w')} - \delta_{(v,w)} - \delta_{(v,w')}) = 0$ und $\phi'(\delta_{(v,\lambda w)} - \lambda\delta_{(v,w)}) = 0$ für alle $v \in V$, $w, w' \in W$, $\lambda \in \mathbb{K}$ gilt. Also ist $U \subseteq \ker(\phi')$, und nach der universellen Eigenschaft des Quotientenraums aus Satz 12.1.7 existiert damit eine eindeutig bestimmte $\mathbb{K}$-lineare Abbildung $\tilde{\phi} : V\otimes_{\mathbb{K}}W \to X$ mit $\tilde{\phi} \circ \pi = \phi'$. Für die $\mathbb{K}$-lineare Abbildung $\tau = \pi \circ \iota$ folgt dann $\tilde{\phi} \circ \tau = \tilde{\phi} \circ \pi \circ \iota = \phi' \circ \iota = \phi$. Wir erhalten das folgende kommutierende Diagramm:

$$\begin{tikzcd}
V \times W \arrow[d, "\iota"'] \arrow[dr, "\phi"] \arrow[dd, bend right=60, "\tau"'] & \\
\langle V \times W\rangle_{\mathbb{K}} \arrow[r, dashed, "\exists!\phi'"'] \arrow[d, "\pi"'] & X \\
V\otimes_{\mathbb{K}}W \arrow[ur, dashed, "\exists!\tilde{\phi}"'] &
\end{tikzcd}$$

Ist $\psi : V\otimes_{\mathbb{K}}W \to X$ eine weitere $\mathbb{K}$-lineare Abbildung mit $\psi \circ \tau = \psi \circ \pi \circ \iota = \phi$, so ist $\psi' = \psi \circ \pi : \langle V \times W\rangle_{\mathbb{K}} \to X$ eine $\mathbb{K}$-lineare Abbildung mit $\psi' \circ \iota = \phi = \phi' \circ \iota$, und aus der universellen Eigenschaft des frei erzeugten Vektorraums folgt $\psi \circ \pi = \psi' = \phi' = \tilde{\phi} \circ \pi$. Aus der universellen Eigenschaft des Quotienten folgt dann $\psi = \tilde{\phi}$. Damit ist gezeigt, dass das Paar $(V\otimes_{\mathbb{K}}W, \tau)$ die universelle Eigenschaft aus Definition 12.3.1 besitzt. □

Bemerkung 12.3.4: In der Konstruktion des Tensorprodukts $V\otimes_{\mathbb{K}}W = \langle V \times W\rangle_{\mathbb{K}}/U$ wandeln sich die Elemente in (26), die den Untervektorraum U aufspannen, in Rechenregeln für das Tensorprodukt um. Für alle $v, v' \in V$, $w, w' \in W$ und $\lambda \in \mathbb{K}$ gilt

$$(v + v')\otimes w = v\otimes w + v'\otimes w \qquad v\otimes(w + w') = v\otimes w + v\otimes w' \qquad (\lambda v)\otimes w = v\otimes(\lambda w) = \lambda\, v\otimes w.$$

Bemerkung 12.3.5: Alternativ kann man die universelle Eigenschaft des Tensorprodukts $V\otimes_{\mathbb{K}}W$ folgendermaßen formulieren: Für alle Vektorräume X über $\mathbb{K}$ ist die Abbildung

$$F_X : L(V \times W, X) \to \mathrm{Hom}_{\mathbb{K}}(V\otimes_{\mathbb{K}}W, X), \quad \phi \mapsto \tilde{\phi}$$

ein Isomorphismus von $\mathbb{K}$-Vektorräumen. Insbesondere ist $L(V \times W, \mathbb{K})$ isomorph zum Dualraum $(V\otimes_{\mathbb{K}}W)^* = \mathrm{Hom}_{\mathbb{K}}(V\otimes_{\mathbb{K}}W, \mathbb{K})$ und der $\mathbb{K}$-Vektorraum $L(V \times V, \mathbb{K})$ der Bilinearformen auf V ist isomorph zum Dualraum $(V\otimes V)^*$. (Beweis: Übung).

Da nach Satz 12.3.3 die Elemente $v\otimes w$ mit $v \in V$ und $w \in W$ den Vektorraum $V\otimes_{\mathbb{K}}W$ erzeugen und nach Satz 4.2.1 jede $\mathbb{K}$-lineare Abbildung $\phi : V\otimes_{\mathbb{K}}W \to X$ durch ihre Werte auf einem Erzeugendensystem von $V\otimes_{\mathbb{K}}W$ eindeutig bestimmt ist, spezifiziert man eine lineare Abbildung ϕ häufig auch durch die Abbildungsvorschrift $\phi : v\otimes w \mapsto \phi(v\otimes w)$, auch wenn ein allgemeines Element von $V\otimes_{\mathbb{K}}W$ nicht von dieser Form, sondern eine endliche Summe solcher Elemente ist.

Dies garantiert jedoch nicht die Widerspruchsfreiheit der so angegebenen Abbildungsvorschrift, da die Werte einer linearen Abbildung nur auf einer Basis, nicht aber auf einem beliebigen Erzeugendensystem frei vorgegeben werden können. Aus diesem Grund möchte man auch für Tensorprodukte $V\otimes_{\mathbb{K}}W$ eine Basis angeben können, und es ist naheliegend, diese aus einer Basis von V und einer Basis von W zu konstruieren.

Satz 12.3.6: Seien V, W Vektorräume über $\mathbb{K}$. Ist B_V eine Basis von V und B_W eine Basis von W, so ist $B_{V\otimes W} = \{b\otimes c | b \in B_V, c \in B_W\}$ eine Basis von $V\otimes_{\mathbb{K}}W$.

Sind V und W endlich-dimensional, so folgt

$$\dim_{\mathbb{K}}(V\otimes_{\mathbb{K}}W) = \dim_{\mathbb{K}}(V) \cdot \dim_{\mathbb{K}}(W).$$

Beweis:
Per Definition ist $D = \{\delta_{(v,w)} | v \in V, w \in W\}$ eine Basis von $\langle V \times W\rangle_{\mathbb{K}}$. Da die Abbildung $\pi : \langle V\times W\rangle_{\mathbb{K}} \to V\otimes_{\mathbb{K}}W$, $\delta_{(v,w)} \mapsto v\otimes w$ surjektiv und $\mathbb{K}$-linear ist, ist $D' = \{v\otimes w | v \in V, w \in W\}$ ein Erzeugendensystem von $V\otimes_{\mathbb{K}}W$. Da B_V und B_W Basen von V und W sind, existieren zu $v \in V$ und $w \in W$ Skalare $\lambda_1, ..., \lambda_n, \mu_1, ..., \mu_m \in \mathbb{K}$ und Vektoren $b_1, ..., b_n \in B_V$, $c_1, ..., c_m \in B_W$ mit $v = \Sigma_{i=1}^n \lambda_i b_i$ und $w = \Sigma_{j=1}^m \mu_i c_i$. Daraus folgt

$$v\otimes w = (\Sigma_{i=1}^n \lambda_i b_i)\otimes(\Sigma_{j=1}^m \mu_j c_j) = \Sigma_{i=1}^n \Sigma_{j=1}^m \lambda_i \mu_j \, b_i\otimes c_j.$$

Also lässt sich jeder Vektor in D' als Linearkombination von Vektoren in $B_{V\otimes W}$ darstellen, und damit ist auch $B_{V\otimes W}$ ein Erzeugendensystem von $V\otimes_{\mathbb{K}}W$.

Seien nun $\mu_{ij} \in \mathbb{K}$, $b_1, ..., b_n \in B_V$ und $c_1, ..., c_m \in B_W$ mit $\Sigma_{i=1}^n \Sigma_{j=1}^m \mu_{ij} \, b_i\otimes c_j = 0$. Dann setzen wir $d_j = \Sigma_{i=1}^n \mu_{ij} b_i \in V$ und erhalten $\Sigma_{j=1}^m d_j\otimes c_j = 0$. Daraus folgt $\Sigma_{j=1}^m \delta_{(d_j,c_j)} \in U$, wobei U den Untervektorraum aus Satz 12.3.3 bezeichnet. Da das m-Tupel $(c_1, ..., c_m)$ linear unabhängig ist, kann das aber nur dann gelten, wenn $d_j = 0$ für alle $j \in \{1, ..., m\}$ gilt. Da auch das n-Tupel $(b_1, ..., b_n)$ linear unabhängig ist, folgt daraus $\mu_{ij} = 0$ für alle $i \in \{1, ..., n\}$, $j \in \{1, ..., m\}$ und somit ist $B_{V\otimes W}$ linear unabhängig. Also ist $B_{V\otimes W}$ eine Basis von $V\otimes_{\mathbb{K}}W$.

Die Dimensionsformel ergibt sich durch Abzählen der Basisvektoren. □

Während wir für die Konstruktion einer Basis die konkrete Realisierung des Tensorprodukts als Quotient des von der Menge $V \times W$ frei erzeugten Vektorraums benutzt haben, lassen sich Isomorphieaussagen für Tensorprodukte von Vektorräumen wieder abstrakt und allgemein mit der universellen Eigenschaft beweisen. Insbesondere erhält man daraus eine Charakterisierung von Vektorräumen $\mathrm{Hom}_{\mathbb{K}}(V, W)$ und Endomorphismenräumen $\mathrm{End}_{\mathbb{K}}(V)$ durch Tensorprodukte.

Beispiel 12.3.7: Seien V, W endlich-dimensionale Vektorräume über $\mathbb{K}$. Dann gilt:

$$V^*\otimes_{\mathbb{K}}W \cong \mathrm{Hom}_{\mathbb{K}}(V, W) \qquad V^*\otimes V \cong \mathrm{End}_{\mathbb{K}}(V).$$

Beweis: Die Abbildung $\phi : V^* \times W \to \mathrm{Hom}_{\mathbb{K}}(V, W)$, $(\alpha, w) \mapsto \phi_{\alpha,w}$ mit $\phi_{\alpha,w}(v) = \alpha(v)\, w$ für alle $v \in V$ ist $\mathbb{K}$-bilinear, denn für alle $\alpha, \beta \in V^*$, $w, w' \in W$ und $v \in V$ gilt

$$\begin{aligned}
&\phi_{\alpha+\beta,w}(v) = (\alpha + \beta)(v)\, w = \alpha(v)\, w + \beta(v)\, w = \phi_{\alpha,w}(v) + \phi_{\beta,w}(v) = (\phi_{\alpha,w} + \phi_{\beta,w})(v)\\
&\phi_{\alpha,w+w'}(v) = \alpha(v)\,(w + w') = \alpha(v)\, w + \alpha(v)\, w' = \phi_{\alpha,w}(v) + \phi_{\alpha,w'}(v) = (\phi_{\alpha,w} + \phi_{\alpha,w'})(v)\\
&(\phi_{\lambda\alpha,w})(v) = (\lambda\,\alpha)(v)\, w = \lambda\alpha(v)\, w = (\lambda\phi_{\alpha,w})(v) = \alpha(v)\,(\lambda w) = \phi_{\alpha,\lambda w}(v).
\end{aligned}$$

Also existiert nach der universellen Eigenschaft des Tensorprodukts genau eine $\mathbb{K}$-lineare Abbildung $\tilde{\phi} : V^*\otimes_{\mathbb{K}}W \to \mathrm{Hom}_{\mathbb{K}}(V, W)$ mit $\tilde{\phi} \circ \tau = \phi$. Diese ist injektiv, denn

$$\tilde{\phi}(\alpha\otimes w)(v) = \phi_{\alpha,w}(v) = \alpha(v)w = 0\, \forall v \in V \Rightarrow (\alpha(v) = 0\, \forall v \in V) \vee (w = 0) \Rightarrow \alpha\otimes w = 0.$$

Da $\dim_{\mathbb{K}}(V^*\otimes_{\mathbb{K}}W) = \dim_{\mathbb{K}}(V^*) \cdot \dim_{\mathbb{K}}(W) = \dim_{\mathbb{K}}(V) \cdot \dim_{\mathbb{K}}(W) = \dim_{\mathbb{K}} \mathrm{Hom}_{\mathbb{K}}(V, W)$ für alle endlich-dimensionalen $\mathbb{K}$-Vektorräume V, W, ist $\tilde{\phi}$ in diesem Fall ein Isomorphismus.

Eine weitere Möglichkeit, lineare Abbildungen zwischen Tensorprodukten zu konstruieren und Aussagen über die Isomorphie verschiedener Tensorprodukte beweisen, ist die Benutzung der Basen aus Satz 12.3.6. Diese erlaubt es uns, lineare Abbildungen durch die Vorgabe ihrer Werte auf der Basis eindeutig zu definieren. Ähnlich wie in Satz 12.2.15 erhalten wir so Isomorphieregeln für Tensorprodukte, die zusammen mit denen für direkte Summen den Axiomen eines unitalen Rings ähneln. Man kann sie nämlich als Analoga von Assoziativität, Kommutativität, Existenz von neutralen Elementen und Distributivgesetz auffassen. Dabei ersetzt man die Ringaddition durch die direkte Summe und die Ringmultiplikation durch das Tensorprodukt.

Satz 12.3.8: (**Isomorphieregeln für Tensorprodukte**) Seien U, V, W Vektorräume über einem Körper $\mathbb{K}$. Dann erhält man die folgenden Isomorphismen von $\mathbb{K}$-Vektorräumen:

(i) **Assoziativität:** $a_{U,V,W} : U\otimes_{\mathbb{K}}(V\otimes_{\mathbb{K}}W) \to (U\otimes_{\mathbb{K}}V)\otimes_{\mathbb{K}}W$, $u\otimes(v\otimes w) \mapsto (u\otimes v)\otimes w$
(ii) **Kommutativität:** $c_{V,W} : V\otimes_{\mathbb{K}}W \cong W\otimes_{\mathbb{K}}V$, $v\otimes w \mapsto w\otimes v$
(iii) **neutrales Element:** $l_V : \mathbb{K}\otimes_{\mathbb{K}}V \to V$, $\lambda\otimes v \mapsto \lambda v$ und $r_V : V\otimes_{\mathbb{K}}\mathbb{K} \to V$, $v\otimes\lambda \mapsto \lambda v$
(iv) **Distributivgesetze:** $l_{U,V,W} : (U\oplus V)\otimes_{\mathbb{K}}W \to U\otimes_{\mathbb{K}}W\oplus V\otimes_{\mathbb{K}}W$, $(u+v)\otimes w \mapsto u\otimes w+v\otimes w$ und $r_{U,V,W} : U\otimes_{\mathbb{K}}(V \oplus W) \to U\otimes_{\mathbb{K}}V \oplus U\otimes_{\mathbb{K}}W$, $u\otimes(v + w) \mapsto u\otimes v + u\otimes w$.

Beweis:
(i) Für beliebige Basen B_U, B_V, B_W von U, V, W sind nach Satz 12.3.6

$$B_{U\otimes_{\mathbb{K}}(V\otimes_{\mathbb{K}}W)} = \{u\otimes(v\otimes w) \,|\, u \in B_U, v \in B_V, w \in B_W\}$$
$$B_{(U\otimes_{\mathbb{K}}V)\otimes_{\mathbb{K}}W} = \{(u\otimes v)\otimes w) \,|\, u \in B_U, v \in B_V, w \in B_W\}$$

Basen von $U\otimes_{\mathbb{K}}(V\otimes_{\mathbb{K}}W)$ und $(U\otimes_{\mathbb{K}}V)\otimes_{\mathbb{K}}W$. Nach Satz 4.2.1 ist eine lineare Abbildung durch ihre Werte auf einer Basis eindeutig bestimmt, und diese können beliebig vorgegeben werden. So definieren wir $\mathbb{K}$-lineare Abbildungen $a_{U,V,W} : U\otimes_{\mathbb{K}}(V\otimes_{\mathbb{K}}W) \to (U\otimes_{\mathbb{K}}V)\otimes_{\mathbb{K}}W$ und $a'_{U,V,W} : (U\otimes_{\mathbb{K}}V)\otimes_{\mathbb{K}}W \to U\otimes_{\mathbb{K}}(V\otimes_{\mathbb{K}}W)$ durch die Bedingungen $a_{U,V,W}(u\otimes(v\otimes w)) = (u\otimes v)\otimes w$ und $a'_{U,V,W}((u\otimes v)\otimes w) = u\otimes(v\otimes w)$ für alle $u \in B_U, v \in B_V, w \in B_W$.

Mit den Rechenregeln in Bemerkung 12.3.4 lässt sich nachrechnen, dass diese Bedingungen dann für alle $u \in U$, $v \in V$ und $w \in W$ erfüllt sind, insbesondere also nicht von der Wahl der Basen abhängen. Ebenso folgt daraus $a'_{U,V,W} = a^{-1}_{U,V,W}$, und somit ist $a_{U,V,W}$ ist ein Isomorphismus.

(ii) Analog zu (i) definiert man die lineare Abbildung $c_{V\otimes W}$ und ihre Inverse durch ihre Werte auf den Basen $B_{V\otimes W} = \{v\otimes w|v \in B_V, w \in B_W\}$ und $B_{W\otimes V} = \{w\otimes v|v \in B_V, w \in B_W\}$. Setzt man $c_{V,W}(v\otimes w) = w\otimes v$ und $c'_{V,W}(w\otimes v) = v\otimes w$ für $v \in B_V, w \in B_W$, so erhält man mit Rechenregeln in Bemerkung 12.3.4 die angegebene Abbildung $c_{V,W}$ und $c_{V,W} \circ c'_{W,V} = \mathrm{id}_{W\otimes_{\mathbb{K}}V}$ sowie $c'_{V,W} \circ c_{W,V} = \mathrm{id}_{V\otimes_{\mathbb{K}}W}$. Dies zeigt, dass $c_{V,W}$ ein Isomorphismus von $\mathbb{K}$-Vektorräumen ist.

(iii) Analog zu (i) und (ii) können wir die linearen Abbildungen l_V, r_V und ihre Inversen durch ihre Werte auf Basen definieren. Setzt man $l_V(1\otimes v) = v = r_V(v\otimes 1)$, $l'_V(v) = 1\otimes v$ und $r'_V(v) = v\otimes 1$ für alle $v \in B_V$, so erhält man mit den Rechenregeln für Tensorprodukte die angegeben Isomorphismen l_V und r_V sowie ihre Inversen.

(iv) Der Beweis von (iv) ist analog zu (i)-(iii), nur dass man für $l_{U,V,W}$ die Werte auf den Basen

$$B_{(U\oplus V)\otimes_{\mathbb{K}}W} = \{(u + v)\otimes w \,|\, u \in B_U, v \in B_V, w \in B_W\}$$
$$B_{U\otimes_{\mathbb{K}}W\oplus V\otimes_{\mathbb{K}}W} = \{u\otimes w + v\otimes w \,|\, u \in B_U, v \in B_V, w \in B_W\}$$

vorgibt. Aus Bemerkung 12.3.4 ergibt sich dann der angegebene Isomorphismus und sein Inverses. Der Beweis für $r_{U,V,W}$ ist analog. □

12.4 Die Tensoralgebra

In diesem Abschnitt befassen wir uns mit multiplen Tensorprodukten von Vektorräumen und zeigen, dass sich unter Verwendung von multiplen Tensorprodukten und direkten Summen von Vektorräumen aus jedem $\mathbb{K}$-Vektorraum V eine Algebra konstruieren lässt, die *Tensoralgebra* von V. Diese besitzt eine wichtige universelle Eigenschaft, die es uns erlaubt, sie als Verallgemeinerung der Polynomalgebra $\mathbb{K}[x]$ zu interpretieren.

Wir beginnen mit der Definition multipler Tensorprodukte. Diese definieren wir induktiv, wobei wir uns für eine Klammerung entscheiden müssen. Auf die genaue Wahl kommt es nach Lemma 12.3.8 nicht an, da alle Wahlen isomorphe Ergebnisse liefern. Die Definition von mehrfachen Tensorprodukten erfordert aber eine Klammerung.

Definition 12.4.9: Sei $n \in \mathbb{N}$, $V_1, .., V_n$ Vektorräume über $\mathbb{K}$ und $v_i \in V$ für $i \in \{1, ..., n\}$. Dann definiert man

$$V_1 \otimes_{\mathbb{K}} ... \otimes_{\mathbb{K}} V_n := (V_1 \otimes_{\mathbb{K}} ... \otimes_{\mathbb{K}} V_{n-1}) \otimes_{\mathbb{K}} V_n \qquad v_1 \otimes ... \otimes v_n := (v_1 \otimes ... \otimes v_{n-1}) \otimes v_n.$$

Das n-fache Tensorprodukt eines $\mathbb{K}$-Vektorraums V mit sich selbst bezeichnet man auch mit $V^{\otimes n} := V \otimes_{\mathbb{K}} ... \otimes_{\mathbb{K}} V$, und man setzt $V^{\otimes 0} := \mathbb{K}$ und $V^{\otimes 1} := V$.

Aus der universellen Eigenschaft des Tensorprodukts in Definition 12.3.1 lassen sich nun induktiv entsprechende universelle Eigenschaften für n-fache Tensorprodukte herleiten. Diese sind völlig analog zu der universellen Eigenschaft in Definition 12.3.1. Es werden lediglich die bilinearen Abbildungen in Definition 12.3.1 durch n-lineare Abbildungen ersetzt.

Satz 12.4.10: **(universelle Eigenschaft von multiplen Tensorprodukten)**
Seien $V_1, ..., V_n$ Vektorräume über $\mathbb{K}$. Dann ist die Abbildung

$$\tau : V_1 \times .... \times V_n \to V_1 \otimes_{\mathbb{K}} ... \otimes_{\mathbb{K}} V_n, \quad (v_1, ..., v_n) \mapsto v_1 \otimes ... \otimes v_n$$

n-linear, und zu jeder n-linearen Abbildung $\phi : V_1 \times ... \times V_n \to W$ in einen $\mathbb{K}$-Vektorraum W existiert genau eine lineare Abbildung $\tilde{\phi} : V_1 \otimes_{\mathbb{K}} ... \otimes_{\mathbb{K}} V_n \to W$ mit $\tilde{\phi} \circ \tau = \phi$.

$$\begin{tikzcd} V_1 \times ... \times V_n \arrow[r, "\phi"] \arrow[d, "\tau"'] & W \\ V_1 \otimes_{\mathbb{K}} ... \otimes_{\mathbb{K}} V_n \arrow[ur, dashed, "\exists! \tilde{\phi}"'] & \end{tikzcd}$$

Beweis:
Induktion über n. Für $n = 2$ ist die Aussage die universelle Eigenschaft aus Definition 12.3.1.

$n - 1 \to n$: Seien die Aussagen für Tensorprodukte von $\leq n - 1$ Vektorräumen bewiesen. Nach Induktionsvoraussetzung ist dann die Abbildung

$$\tau_{n-1} : V_1 \times ... \times V_{n-1} \to V_1 \otimes_{\mathbb{K}} ... \otimes_{\mathbb{K}} V_{n-1}, \quad (v_1, ..., v_n) \mapsto v_1 \otimes ... \otimes v_n$$

$(n - 1)$-linear. Per Definition des Tensorprodukts ist die Abbildung

$$\tau' : V_1 \otimes_{\mathbb{K}} ... \otimes_{\mathbb{K}} V_{n-1} \times V_n \to V_1 \otimes_{\mathbb{K}} ... \otimes_{\mathbb{K}} V_n, \quad (x, v_n) \to x \otimes v_n$$

bilinear. Also ist die Abbildung $\tau = \tau' \circ (\tau_{n-1} \times \mathrm{id}_{V_n}) : V_1 \times ... \times V_n \to V_1 \otimes_{\mathbb{K}} ... \otimes_{\mathbb{K}} V_n$, $(v_1, ..., v_n) \to v_1 \otimes ... \otimes v_n$ linear in jedem der ersten $(n - 1)$ Argumente als Verkettung zweier linearer Abbildungen und linear im letzten Argument, da τ' linear im zweiten Argument ist. Damit ist τ eine n-lineare Abbildung.

Ist $\phi : V_1 \times ... \times V_n \to W$ n-linear, so ist $\phi_{v_n} : V_1 \times ... \times V_{n-1} \to W$, $(v_1, ..., v_{n-1}) \mapsto \phi(v_1, ..., v_n)$ eine $(n-1)$-lineare Abbildung für alle $v_n \in V$. Nach Induktionsvoraussetzung existiert genau eine lineare Abbildung $\tilde{\phi}_{v_n} : V_1 \otimes_{\mathbb{K}} ... \otimes_{\mathbb{K}} V_{n-1} \to W$ mit $\tilde{\phi}_{v_n} \circ \tau_{n-1} = \phi_{v_n}$. Dann ist die Abbildung $\phi' : V_1 \otimes_{\mathbb{K}} ... \otimes_{\mathbb{K}} V_{n-1} \times V_n \to W$, $(x, v_n) \mapsto \tilde{\phi}_{v_n}(x)$ bilinear. Also existiert nach der universellen Eigenschaft des Tensorprodukts genau eine lineare Abbildung $\tilde{\phi} : V_1 \otimes_{\mathbb{K}} ... \otimes_{\mathbb{K}} V_n \to W$ mit $\tilde{\phi} \circ \tau' = \phi'$. Daraus folgt $\tilde{\phi} \circ \tau = \tilde{\phi} \circ \tau' \circ (\tau_{n-1} \times \mathrm{id}_{V_n}) = \phi' \circ (\tau_{n-1} \times \mathrm{id}_{V_n}) = \phi$. Also ist $\tilde{\phi} : V_1 \times ... \times V_n \to W$ die eindeutig bestimmte lineare Abbildung mit $\tilde{\phi} \circ \tau = \phi$. □

Wir betrachten nun die direkte Summe $\oplus_{n=0}^{\infty} V^{\otimes n}$ der $\mathbb{K}$-Vektorräume $V^{\otimes n}$. Da $V^{\otimes n} \cap V^{\otimes m} = \emptyset$ für $m \neq n$ gilt, können wir die Elemente $(x_n)_{n \in \mathbb{N}_0} \in \oplus_{n=0}^{\infty} V^n$ mit $x_n \in V^{\otimes n}$ und $x_n = 0$ für fast alle $n \in \mathbb{N}_0$ auch ohne Verwechslungsgefahr mit $\Sigma_{n=0}^{\infty} x_n$ bezeichnen. Mit dieser Notation erhält man insbesondere $(x\delta_{in})_{n \in \mathbb{N}_0} = x$ für alle $x \in V^{\otimes n}$. Dadurch vereinfacht sich die Notation erheblich, und wir können zeigen, dass das "Aneinanderhängen" von Elementen $v_1 \otimes ... \otimes v_n$ eine Algebrastruktur auf der direkten Summe $T(V) = \oplus_{n=0}^{\infty} V^{\otimes n}$ definiert.

Satz 12.4.11: Sei V ein $\mathbb{K}$-Vektorraum und $T(V) = \oplus_{n=0}^{\infty} V^{\otimes n}$. Dann gibt es genau eine $\mathbb{K}$-bilineare Abbildung $\cdot : T(V) \times T(V) \to T(V)$ mit

$$\begin{aligned}
&(v_1 \otimes ... \otimes v_n) \cdot (v_{n+1} \otimes ... \otimes v_{n+m}) = v_1 \otimes ... \otimes v_{n+m} \\
&1_{\mathbb{K}} \cdot (v_1 \otimes ... \otimes v_n) = (v_1 \otimes ... \otimes v_n) \cdot 1_{\mathbb{K}} = v_1 \otimes ... \otimes v_n \\
&1_{\mathbb{K}} \cdot 1_{\mathbb{K}} = 1_{\mathbb{K}}
\end{aligned}$$

für alle $n, m \in \mathbb{N}$ und $v_1, ..., v_{n+m} \in V$.

Der $\mathbb{K}$-Vektorraum $T(V)$ mit dieser Abbildung $\cdot$ ist eine eine assoziative Algebra mit Eins über $\mathbb{K}$ und heißt die **Tensoralgebra** von V. Die Abbildung $\iota_V : V \to T(V)$, $v \mapsto v$ ist injektiv und $\mathbb{K}$-linear und heißt **Inklusionsabbildung**.

Beweis:
Ist B eine Basis von V, so ist nach Satz 12.3.6 die Menge $B_n = \{b_1 \otimes ... \otimes b_n \,|\, b_1, ..., b_n \in B\}$ eine Basis von $V^{\otimes n}$ für alle $n \in \mathbb{N}$ und $B_0 = \{1_{\mathbb{K}}\}$ ist eine Basis von $V^{\otimes 0} = \mathbb{K}$. Per Definition der direkten Summe ist dann $B_{\otimes} = \{b_1 \otimes ... \otimes b_n \,|\, b_i \in B, n \in \mathbb{N}_0\}$ eine Basis von $T(V)$, wenn wir $b_1 \otimes ... \otimes b_n = 1_{\mathbb{K}}$ für $n = 0$ setzen.

Eine bilineare Abbildung $\cdot : T(V) \times T(V) \to T(V)$ ist nach Satz 5.2.7 durch ihre Werte auf Paaren von Basisvektoren eindeutig bestimmt, und diese können beliebig vorgegeben werden. Also existiert genau eine bilineare Abbildung $\cdot : T(V) \times T(V) \to T(V)$ mit $(v_1 \otimes ... \otimes v_n) \cdot (v_{n+1} \otimes ... \otimes v_{n+m}) = v_1 \otimes ... \otimes v_{n+m}$, $1_{\mathbb{K}} \cdot (v_1 \otimes ... \otimes v_n) = (v_1 \otimes ... \otimes v_n) \cdot 1_{\mathbb{K}} = v_1 \otimes ... \otimes v_n$ und $1_{\mathbb{K}} \cdot 1_{\mathbb{K}} = 1_{\mathbb{K}}$ für alle $n, m \in \mathbb{N}$ und $v_1, ..., v_{n+m} \in B$.

Aus der n-Linearität des n-fachen Tensorprodukts und der Bilinearität von $\cdot$ folgt, dass diese Identitäten für alle $n, m \in \mathbb{N}_0$ und $v_1, ..., v_{n+m} \in V$ gelten. Die Assoziativität von $\cdot$ und die Tatsache, dass $1_{\mathbb{K}}$ das neutrale Element für $\cdot$ ist, ergibt sich direkt aus der Definition. Die Distributivgesetze und die Verträglichkeit der Multiplikation mit der Skalarmultiplikation folgen aus der Bilinearität der Multiplikation. Die Injektivität und $\mathbb{K}$-Linearität der Inklusionsabbildung ι_V folgt direkt aus der Definition. □

Auch die Tensoralgebra lässt sich durch eine universelle Eigenschaft beschreiben, die an die universelle Eigenschaft der Polynomalgebra erinnert. In Satz 8.1.1 wurde gezeigt, dass zu jedem Element $a \in A$ einer $\mathbb{K}$-Algebra A genau ein Algebrahomomorphismus $\phi : \mathbb{K}[x] \to A$ mit $\phi(x) = a$ existiert. Im Fall der Tensoralgebra gibt man stattdessen eine $\mathbb{K}$-lineare Abbildung $\phi : V \to A$ vor und die Bedingung $\phi(x) = a$ wird durch die Bedingung $\tilde{\phi}|_V = \phi$ ersetzt.

Satz 12.4.12: (**universelle Eigenschaft der Tensoralgebra**)
Sei V ein $\mathbb{K}$-Vektorraum und A eine assoziative $\mathbb{K}$-Algebra mit Eins. Dann existiert zu jeder $\mathbb{K}$-linearen Abbildung $\phi : V \to A$ genau ein Algebrahomomorphismus $\tilde{\phi} : T(V) \to A$ mit $\tilde{\phi} \circ \iota = \phi(v)$

$$\begin{tikzcd} V \arrow[r, "\phi"] \arrow[d, "\iota"'] & A \\ T(V) \arrow[ur, dashed, "\tilde{\phi}"'] & \end{tikzcd}$$

Beweis:
Da für jede Basis B von V die Menge $B_\otimes = \{b_1 \otimes ... \otimes b_n \,|\, n \in \mathbb{N}_0, b_i \in B\}$ eine Basis von $T(V)$ ist, ist jeder Algebrahomomorphismus $\tilde{\phi} : T(V) \to A$ durch seine Werte auf $B_\otimes$ eindeutig bestimmt. Außerdem erfüllt jeder Algebrahomomorphismus $\tilde{\phi} : T(V) \to A$ die Bedingungen $\phi(1) = 1_A$ und $\tilde{\phi}(b_1 \otimes ... \otimes b_n) = \tilde{\phi}(b_1 \cdot ... \cdot b_n) = \tilde{\phi}(b_1) \cdot ... \cdot \tilde{\phi}(b_n)$ für alle $n \in \mathbb{N}$ und $b_i \in B$. Also ist jeder Algebrahomomorphismus $\tilde{\phi} : T(V) \to A$ durch seine Werte auf B eindeutig bestimmt, und diese können beliebig vorgegeben werden. Damit gibt es genau einen Algebrahomomorphismus $\tilde{\phi} : T(V) \to A$ mit $\tilde{\phi}(b) = \phi(b)$ für alle $b \in B$. Also stimmen die $\mathbb{K}$-linearen Abbildungen $\phi : V \to A$ und $\tilde{\phi}|_V : V \to A$ auf einer Basis von V überein, und es folgt $\tilde{\phi}(v) = \tilde{\phi} \circ \iota_V(v) = \phi(v)$ für alle $v \in V$. □

Auch der Beweis der universellen Eigenschaft der Tensoralgebra erinnert stark an den entsprechenden Beweis der universellen Eigenschaft der Polynomalgebra in Satz 8.1.1. Dies ist kein Zufall, sondern ergibt sich aus der Tatsache, dass die Polynomalgebra ein Spezialfall der Tensoralgebra $T(V)$ für $V = \mathbb{K}$ ist. In diesem Fall ist eine $\mathbb{K}$-lineare Abbildung $\phi : V \to A$ durch ihren Wert auf $1_\mathbb{K}$ eindeutig bestimmt, und dieser Wert kann beliebig vorgegeben werden. Damit entspricht die Wahl einer $\mathbb{K}$-linearen Abbildung $\phi : \mathbb{K} \to A$ der Wahl eines Elements $a = \phi(1_\mathbb{K})$ in A. Dass die Polynomalgebra $\mathbb{K}[x]$ tatsächlich isomorph zur Tensoralgebra $T(\mathbb{K})$ ist, ergibt sich dann direkt aus den universellen Eigenschaften.

Satz 12.4.13: Für jeden Körper $\mathbb{K}$ sind die Algebren $T(\mathbb{K})$ und $\mathbb{K}[x]$ isomorph.

Beweis:
Zu der $\mathbb{K}$-linearen Abbildung $\phi : \mathbb{K} \to \mathbb{K}[x]$, $\lambda \mapsto \lambda x$ gibt es genau einen Algebrahomomorphismus $\tilde{\phi} : T(\mathbb{K}) \to \mathbb{K}[x]$ mit $\tilde{\phi} \circ \iota_\mathbb{K} = \phi$. Umgekehrt gibt es nach der universellen Eigenschaft der Polynomalgebra aber auch genau einen Algebrahomomorphismus $\psi : \mathbb{K}[x] \to T(\mathbb{K})$ mit $\psi(x) = \iota_\mathbb{K}(1)$. Verkettet man diese, erhält man Algebrahomomorphismen $\psi \circ \phi : T(\mathbb{K}) \to T(\mathbb{K})$ mit $(\psi \circ \phi) \circ \iota_\mathbb{K} = \iota_\mathbb{K}$ und $\phi \circ \psi : \mathbb{K}[x] \to \mathbb{K}[x]$ mit $(\phi \circ \psi)(x) = x$.

Da auch die zwei Identitätsabbildungen $\mathrm{id}_{T(\mathbb{K})} : T(\mathbb{K}) \to T(\mathbb{K})$ und $\mathrm{id}_{\mathbb{K}[x]} : \mathbb{K}[x] \to \mathbb{K}[x]$ die Bedingungen $\mathrm{id}_{T(\mathbb{K})} \circ \iota_\mathbb{K} = \iota_\mathbb{K}$ und $\mathrm{id}_{\mathbb{K}[x]}(x) = x$ erfüllen, muss nach der universellen Eigenschaft der Tensoralgebra $\psi \circ \phi = \mathrm{id}_{T(\mathbb{K})}$ und nach der universellen Eigenschaft der Polynomalgebra $\phi \circ \psi = \mathrm{id}_{\mathbb{K}[x]}$ gelten. Daraus folgt $\psi = \phi^{-1}$ und damit ist $\phi : T(\mathbb{K}) \to \mathbb{K}[x]$ ein Algebraisomorphismus. □

An was man sich erinnern sollte:

Die wichtigsten Begriffe und Definitionen:

- universelle Eigenschaft,
- Eindeutigkeit bis auf eindeutige Isomorphie,
- Quotientenraum,
- frei erzeugter Vektorraum,
- direkte Summe und Produkt von Vektorräumen,
- Tensorprodukt von Vektorräumen,
- Tensoralgebra.

Die wichtigsten Aussagen:

- universelle Eigenschaften liefern Eindeutigkeit bis auf eindeutige Isomorphie
- universelle Eigenschaft von Quotientenräumen,
- universelle Eigenschaft von frei erzeugten Vektorräumen,
- universelle Eigenschaft von direkten Summen und Produkten,
- universelle Eigenschaft von (multiplen) Tensorprodukten,
- Homomorphiesatz, kanonische Faktorisierung,
- innere und äußere direkte Summen sind isomorph,
- Isomorphieregeln für direkte Summen,
- Rechenregeln für Tensorprodukte,
- Basen und Dimensionen von Tensorprodukten,
- Isomorphieregeln für Tensorprodukte,
- universelle Eigenschaft der Tensoralgebra.

Die wichtigsten Beispiele:

- $\oplus_{i\in I}\mathbb{K} \cong \langle I\rangle_{\mathbb{K}}$,
- $\Pi_{i\in I}\mathbb{K} \cong \mathrm{Abb}(I,\mathbb{K})$,
- $L(V\times V,\mathbb{K}) \cong (V\otimes_{\mathbb{K}}V)^*$,
- $\mathrm{Hom}_{\mathbb{K}}(V,W) \cong V^*\otimes_{\mathbb{K}}W$,
- $\mathrm{End}_{\mathbb{K}}(V) \cong V^*\otimes_{\mathbb{K}}V$,
- $T(\mathbb{K}) \cong \mathbb{K}[x]$.

13 Aufgaben

13.1 Aufgaben zu Kapitel 1

Aufgabe 1: Welche der folgenden Schlussfolgerungen ist korrekt?

(a) Falls es anfängt zu regnen, wird die Straße nass. Da die Straße nicht nass werden wird, wird es auch nicht regnen.
(b) Einige Politiker:innen sind ehrlich. Einige Frauen sind Politiker:innen. Also sind einige weibliche Politiker:innen ehrlich.
(c) Eine Politiker:in, die sich bestechen lässt, ist eine Verbrecher:in. Der Regierungschef ist unbestechlich. Daher ist er kein Verbrecher.

Aufgabe 2:

1. Negieren Sie die folgenden Aussagen:
   (a) Jedes Auto, das am Samstag um 9:00 auf dem Parkplatz parkte, war rot.
   (b) Mindestens ein Auto, das am Samstag um 9:00 auf dem Parkplatz parkte, war rot.
   (c) Es gibt keine größte ganze Zahl.
2. Drücken Sie die folgenden Aussagen in Worten aus und, falls eine Aussage falsch sein sollte, ersetzen Sie sie dabei durch ihre Negation.
   (a) $\forall m \in \mathbb{N}\ \exists n \in \mathbb{N} : m = n + n$
   (b) $\exists m \in \mathbb{N}\ \exists n \in \mathbb{N} : (m \neq n) \wedge (m^n = n^m)$

Aufgabe 3:

1. Schreiben Sie die folgenden Aussagen als Formel. Welche der Aussagen sind wahr?
   (a) Für jede natürliche Zahl gibt es eine natürliche Zahl, die doppelt so groß ist.
   (b) Es gibt keine größte natürliche Zahl.
   (c) Ist die Summe zweier ganzer Zahlen gerade, so ist es auch ihre Differenz.
2. Schreiben Sie die folgenden Aussagen als Sätze. Welche der Aussagen sind wahr?
   (a) $\forall x \in \mathbb{Z} : x > 7 \Rightarrow x > 5$
   (b) $\exists x \in \mathbb{Z} : \neg (x \in \mathbb{N})$
   (c) $\forall m \in \mathbb{N}\ \forall n \in \mathbb{N}\ \exists k \in \mathbb{N} : n^k > m$

Aufgabe 4:

1. Schreiben Sie die Negation der folgenden Aussagen mittels Quantoren:
   (a) $\forall n \in \mathbb{N}\ \exists m \in \mathbb{N}\colon m > n$.
   (b) $\exists n \in \mathbb{N}\ \forall m \in \mathbb{N}\colon m > n$.
   (c) $\forall \varepsilon > 0\ \exists N \in \mathbb{N}\ \forall n \geq N\colon \frac{1}{n} < \varepsilon$.
2. Formulieren Sie die folgenden Aussagen mittels Quantoren in mathematischer Formelschreibweise. Welche der Aussagen sind wahr?
   (a) Zu jeder natürlichen Zahl n gibt es eine reelle Zahl, deren Quadrat kleiner ist als n.
   (b) Es gibt genau eine reelle Zahl, die mit ihrem Quadrat übereinstimmt.
   (c) Das Quadrat einer reellen Zahl r ist genau dann gleich 1, wenn die Zahl r gleich -1 oder 1 ist.

Aufgabe 5:

1. Zeigen Sie mittels Wahrheitstafeln, dass die folgenden Äquivalenzen Tautologien sind:
   (a) Assoziativgesetz: $(A \vee (B \vee C)) \Leftrightarrow ((A \vee B) \vee C)$.
   (b) Distributivgesetz: $(A \wedge (B \vee C)) \Leftrightarrow ((A \wedge B) \vee (A \wedge C))$.
   (c) Kommutativgesetz: $(A \vee B) \Leftrightarrow (B \vee A)$.
   (d) Kontrapositionsgesetz: $(A \Rightarrow B) \Leftrightarrow ((\neg B) \Rightarrow (\neg A))$.

2. Formulieren Sie die folgenden Aussagen mittels Quantoren in mathematischer Formelschreibweise. Welche der Aussagen sind wahr?
   (a) Das Quadrat keiner natürlichen Zahl n ist kleiner als n selbst.
   (b) Es gibt eine reelle Zahl r, die größer ist als jede natürliche Zahl n.
   (c) Es gibt genau eine natürliche Zahl n, die mit ihrem Kehrwert übereinstimmt.

Aufgabe 6: Zeigen Sie mit Wahrheitstafeln, dass die Äquivalenzen Tautologien sind:

(a) Assoziativgesetz: $(A \wedge (B \wedge C)) \Leftrightarrow ((A \wedge B) \wedge C)$.
(b) Distributivgesetz: $(A \vee (B \wedge C)) \Leftrightarrow ((A \vee B) \wedge (A \vee C))$.
(c) Kommutativgesetz: $(A \wedge B) \Leftrightarrow (B \wedge A)$.
(d) De Morgansche Regel: $(\neg (A \vee B)) \Leftrightarrow ((\neg A) \wedge (\neg B))$.

Aufgabe 7: Sie sind Gastgeber:in einer Party und verfügen über folgende Informationen:

- Andreas erscheint vielleicht, aber nur wenn Bernd kommt.
- Bernd erscheint nicht, wenn nicht auch Clara da ist.
- Wenn Clara erscheint, dann auch Daniel.
- Wenn Bernd und Daniel beide da sind, dann erscheint Clara sicher nicht.
- Daniel erscheint nur, wenn Andreas oder Clara da sind.
- Mindestens eine Person erscheint.

Wer besucht die Party?

Aufgabe 8: Die Einwohner von Radersi sind strikt in zwei Gruppen getrennt: Solche, die sich stets selbst rasieren und solche, die sich nur von Roberto, dem Barbier, rasieren lassen. Roberto rasiert also genau diejenigen Dorfbewohner, die sich nicht selbst rasieren. Begründen Sie, dass Roberto nicht in Radersi wohnt.

Aufgabe 9: "Schneiders werden uns heute Abend besuchen", kündigt Frau Müller an. "Die ganze Familie, also Herr und Frau Schneider mit ihren drei Söhnen Tim, Kay und Uwe?", fragt Herr Müller. Darauf antwortet Frau Müller: "Nein, ich will es dir so erklären: Wenn Frau Schneider kommt, dann bringt sie auch ihren Mann mit. Mindestens einer der beiden Söhne Uwe und Kay kommt. Entweder kommt Herr Schneider oder Tim. Entweder kommen Tim und Kay oder beide nicht. Und wenn Uwe kommt, dann auch Kay und Frau Schneider. So, jetzt weißt du, wer uns heute abend besuchen wird." Geben Sie die Gästeliste an.

Aufgabe 10: Ein Kommissar hat drei Tatverdächtige: Paula, Quentin und Ralf. Er weiß:

- Wenn sich Quentin oder Ralf als Täter herausstellen, ist Paula unschuldig.

- Ist aber Paula oder Ralf unschuldig, dann muss Quentin ein Täter sein.
- Ist Ralf schuldig, so ist Paula Mittäterin.

Wer ist schuldig?

Aufgabe 11: In Ihrem mathematischen Leben werden Sie häufig mit folgender Situation konfrontiert sein: Gegeben sei eine (wahre) Aussage A, aus der die zu beweisende Aussage B gefolgert werden soll. Formal zeigen Sie also $A \Rightarrow B$. Wir diskutieren exemplarisch die folgenden Beweistechniken:

1. **Direkter Beweis:** Ausgehend von einer gegebene Aussage A, wird auf direktem Wege mit gültigen Schlüssen die zu beweisende Aussage B gefolgert. Beweisen Sie auf diese Weise den Satz: *Die Summe zweier gerader ganzer Zahlen ist gerade.*

2. **Kontrapositionsprinzip:** Es beruht auf der folgenden Tautologie:

$$(A \Rightarrow B) \Leftrightarrow (\neg B \Rightarrow \neg A)\,. \tag{27}$$

   (a) Zeigen Sie, dass es sich bei (27) tatsächlich um eine Tautologie handelt.
   (b) Beweisen Sie mit Hilfe des Kontrapositionsprinzips den folgenden Satz: *Sei $a \in \mathbb{Z}$ und a^2 ungerade, dann ist a ungerade.*

3. **Widerspruchsbeweis:** Der Widerspruchsbeweis beruht auf der folgenden Tautologie:

$$(A \Rightarrow B) \Leftrightarrow (\neg B \wedge A \Rightarrow \neg A)\,. \tag{28}$$

   (a) Zeigen Sie, dass es sich bei (28) tatsächlich um eine Tautologie handelt.
   (b) Führen Sie einen Widerspruchsbeweis, um folgenden Satz zu beweisen: *Die Summe einer geraden und einer ungeraden ganzen Zahl ist ungerade.*

Aufgabe 12: Beweisen Sie mittels vollständiger Induktion:

(a) Für alle $n \in \mathbb{N}$ gilt

$$\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}.$$

(b) Für alle $n \in \mathbb{N}_0$ gilt

$$\sum_{k=0}^{n} 2^k = 2^{n+1} - 1.$$

(c) Für alle $n \in \mathbb{N}, n \geq 2$, gilt

$$\prod_{k=2}^{n} \left(1 - \frac{1}{k}\right) = \frac{1}{n}.$$

(d) Für alle $n \in \mathbb{N}$ ist der folgende Term durch 47 teilbar

$$7^{2n} - 2^n.$$

Aufgabe 13: Beweisen Sie mittels vollständiger Induktion:

(a) Für alle $n \in \mathbb{N}_{\geq 2}$ gilt

$$\frac{4^n}{n+1} < \frac{(2n)!}{(n!)^2}.$$

(b) Seien $a, b \in \mathbb{R}$ mit $a + b > 0$ und $a \neq b$. Für alle $n \in \mathbb{N}_{\geq 2}$ gilt

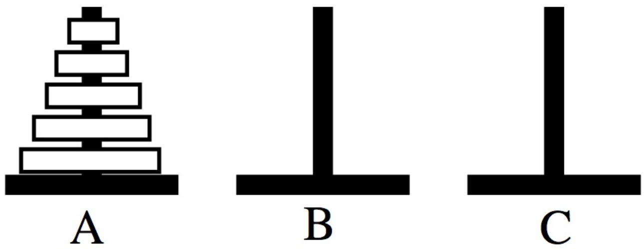
$$(a+b)^n < 2^{n-1}(a^n + b^n).$$

Aufgabe 14: Wir beweisen mit vollständiger Induktion: *Wenn sich unter n Tieren ein Elefant befindet, dann sind alle diese Tiere Elefanten.*

Induktionsanfang: $n = 1$: Wenn eines der Tiere ein Elefant ist, dann sind sie alle Elefanten.
Induktionsannahme: Die Behauptung sei wahr für alle natürlichen Zahlen $\leq n$.
Induktionsschluss: Sei unter $n + 1$ Tieren eines ein Elefant. Wir stellen die Tiere so in eine Reihe, dass sich dieser Elefant unter den ersten n Tieren befindet. Nach Induktionsannahme sind dann alle diese ersten n Tiere Elefanten. Damit befindet sich aber auch unter den letzten n Tieren ein Elefant, womit diese auch alle Elefanten sein müssen. Also sind alle $n + 1$ Tiere Elefanten.

Was ist falsch an diesem Argument, oder gilt mathematische Induktion nicht für Elefanten?

Aufgabe 15: Ein aus $n \in \mathbb{N}$ nach oben kleiner werdenden Scheiben bestehender Turm soll vom Ort A zum Ort B transportiert werden. Es gibt einen Ort C, wo die Scheiben zwischengelagert werden können. Die Bedingung ist, dass in keinem Zwischenschritt eine größere Scheibe auf einer kleineren abgelegt werden darf. Ist es so immer möglich den Turm unter diesen Bedingungen von A nach B zu transportieren, oder hängt dies möglicherweise von der Anzahl Scheiben ab, aus denen der Turm besteht? Begründen Sie Ihre Antwort durch Beweis (z.B. vollständige Induktion).



Aufgabe 16: Bestimmen Sie die Potenzmenge $\mathrm{Pot}(M)$ der Menge $M = \{1, 2, 3\}$.

Aufgabe 17: Wahr oder falsch? Begründen oder widerlegen Sie die Aussagen.

(a) Die leere Menge $\emptyset$ ist Teilmenge jeder Menge.
(b) Die Potenzmenge der leeren Menge ist $\mathrm{Pot}(\emptyset) = \emptyset$.
(c) Es gilt $\mathrm{Pot}(X) \neq \emptyset$ für jede Menge X.
(d) Die Potenzmenge einer endlichen Menge X hat immer mehr Elemente als X.
(e) Für jede Menge X gilt $\cup\{M \in \mathrm{Pot}(X) \mid M \neq X\} = X$.

Aufgabe 18: Wie viele Elemente hat die Menge M?

(a) $M = \{\emptyset,\ \{\emptyset\},\ \{\emptyset, \{\emptyset\}\},\ \{\emptyset, \{\{\emptyset\}\}\}\}$
(b) $M = \{\emptyset,\ \{\emptyset\}, \{\emptyset\},\ \{\{\emptyset\}, \{\{\emptyset\}\}\}\}$
(c) $M = \cup\{\emptyset,\ \{\emptyset\},\ \{\emptyset, \{\emptyset\}\},\ \{\emptyset, \{\{\emptyset\}\}\}\}$
(d) $M = \{X,\ \mathrm{Pot}(X),\ \cup\mathrm{Pot}(X),\ \{X\} \cup \mathrm{Pot}(X)\}$ für eine beliebige Menge X

Aufgabe 19: Zeigen Sie mit den Zermelo-Fraenkel Axiomen: Für $n \in \mathbb{N} = \{1, 2, ...\}$ gibt es keine Mengen $M_1, ..., M_n$ mit $M_1 \in M_2 \in M_3 \in \ldots \in M_n \in M_1$.

Hinweis: Betrachten Sie die Menge $\{M_1, ..., M_n\}$ (siehe Aufgabe 21).

Bild: *Es gibt keine Folge von Behältern, so dass der erste im zweiten enthalten ist, der zweite im dritten, der dritte im vierten etc. und der letzte wieder im ersten.*

Aufgabe 20: Zeigen Sie mit den Zermelo-Fraenkel Axiomen: Es gibt keine Mengen $M_1, M_2, M_3, ...$ mit $M_1 \ni M_2 \ni M_3 \ni \ldots$.

Hinweis: Betrachten Sie die Menge $\{M_1, M_2, \ldots\} = \{M_i \mid i \in \mathbb{N}\}$ (siehe Aufgabe 21).

Bild: *Es gibt keine unendliche Folge von Behältern, so dass jeder Behälter einen noch kleineren Behälter enthält.*

Aufgabe 21: Sei $n \in \mathbb{N}$. Zeigen Sie mit den Zermelo-Fraenkel Axiomen:

(a) Zu n vorgegebenen Mengen $M_1, ..., M_n$ gibt es eine Menge $M = \{M_1, ..., M_n\}$, die genau $M_1, ..., M_n$ als Elemente enthält.
(b) Zu vorgegebenen Mengen M_i für $i \in \mathbb{N}$ gibt es eine Menge M, die genau die Mengen M_i als Elemente enthält: $M = \{M_1, M_2, \ldots\} = \{M_i \mid i \in \mathbb{N}\}$.

Aufgabe 22: Wir betrachten die Menge $\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$ der ganzen Zahlen. Geben Sie ein ausgezeichnetes Element und eine Nachfolgerfunktion an, so dass die Peano-Axiome erfüllt sind.

Aufgabe 23: Zeigen Sie, dass eine endliche Menge nie die Peano Axiome erfüllen kann.

13.2 Aufgaben zu Kapitel 2

Aufgabe 1: Sei $A_1 = \{4, 8, 12\}$, $A_2 = \{3, 6, 9\}$, $A_3 = \{0, 2, 4, 6\}$ und $A_4 = \{6, 12\}$.
Bestimmen Sie die Mengen:

(a) $((A_1 \cup A_2) \cap A_3) \setminus A_4$,
(b) $A_1 \cap (A_2 \cup A_3 \cup A_4)$,
(c) $((A_1 \setminus A_2) \cap A_3) \cap A_4$,
(d) $(A_3 \setminus A_2) \setminus (A_1 \setminus A_4)$.

Aufgabe 2:
Seien $n, k \in \mathbb{N}$ natürliche Zahlen und M eine Menge mit n Elementen. Bestimmen Sie die Anzahl der Teilmengen von M, die genau k Elemente enthalten.

Aufgabe 3: Wir betrachten die Mengen

$$M = \{\{1,2\},1,3\} \qquad N = \{\{\{1\}\},\{1,2\},\{\emptyset\},\{1\}\}$$

(a) Geben Sie an, wie viele Elemente die Mengen M und N haben.
(b) Bestimmen Sie alle Teilmengen der Mengen M und N.
(c) Bestimmen Sie die Potenzmengen $\mathrm{Pot}(M)$ und $\mathrm{Pot}(N)$.
(d) Bestimmen Sie die Mengen $\cup M$ und $\cup N$.
(e) Bestimmen Sie die Menge $M \cap N$.
(f) Bestimmen Sie die Menge $M \cup N$.
(g) Bestimmen Sie die Mengen $M \setminus N$ und $N \setminus M$.
(h) Bestimmen Sie die Mengen $M \setminus (N \setminus M)$ und $N \setminus \mathrm{Pot}(M)$.

Aufgabe 4: Seien A, B, C Mengen.

(a) Beweisen Sie das Komplementsgesetz: Für jede Teilmenge $D \subseteq A$ gilt $A \setminus (A \setminus D) = D$.
(b) Beweisen Sie das zweite de Morgansche Gesetz: $A \setminus (B \cap C) = (A \setminus B) \cup (A \setminus C)$

Aufgabe 5: Seien M, N, P, Q Mengen. Beweisen Sie die folgenden Aussagen oder widerlegen Sie sie durch ein Gegenbeispiel.
(a) Aus $M \setminus N = M$ folgt $N = \emptyset$.
(b) Aus $N \subseteq M$ und $P \subseteq N$ folgt $P \subseteq M$.
(c) Aus $M \setminus N = N \setminus M$ folgt $M = N$.
(d) Aus $M \cup N = M$ folgt $N = \emptyset$.
(e) Aus $M \times N = P \times Q$ folgt $M = P$ und $N = Q$.

Aufgabe 6: Wir betrachten die Mengen

$$S^1 = \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 = 1\} \qquad Q = \{(x,y) \in \mathbb{R}^2 : |x| \leq 1, |y| \leq 1\},$$
$$I = \{x \in \mathbb{R} : |x| < 1\} \qquad W = \{(x,y,z) \in \mathbb{R}^3 : |x| \leq 1, |y| \leq 1, |z| \leq 1\}$$

(a) Skizzieren Sie diese Mengen.
(b) Bestimmen Sie die Mengen $\{0\} \times I$, $I \times I$, $I \times Q$, $Q \times I$, $I \times S^1$ und skizzieren Sie diese.
(c) Bestimmen Sie die Mengen $Q \setminus (I \times I)$, $Q \cap (I \times I)$, $W \setminus (I \times Q)$, $W \setminus (Q \times I)$ und skizzieren Sie diese.

Aufgabe 7: Bestimmen Sie:
(a) die Anzahl der Relationen auf einer Menge mit $n \in \mathbb{N}$ Elementen,
(b) die Anzahl der Abbildungen $f : M \to M$ für eine Menge M mit $n \in \mathbb{N}$ Elementen,
(c) die Anzahl der Äquivalenzrelationen auf einer Menge mit 3 Elementen.

Aufgabe 8: Aussage (falsch!): "Jede Relation auf einer Menge M, die symmetrisch und transitiv ist, ist auch reflexiv und somit eine Äquivalenzrelation."

"Beweis": Sei $\sim$ eine symmetrische und transitive Relation auf einer Menge M. Seien $x, y \in M$ beliebig mit $x \sim y$. Dann gilt wegen der Symmetrie auch $y \sim x$ und mit der Transitivität folgt aus $x \sim y$ und $y \sim x$ auch $x \sim x$. Also ist die Relation $\sim$ auch reflexiv und damit eine Äquivalenzrelation.

Finden Sie den Fehler im "Beweis" und geben Sie eine Relation auf einer Menge an, die zwar symmetrisch und transitiv, aber nicht reflexiv ist.

Aufgabe 9: Bestimmen Sie:

(a) die Anzahl der Relationen auf einer Menge mit $n \in \mathbb{N}$ Elementen,
(b) die Anzahl der Abbildungen $f : M \to M$ für eine Menge M mit $n \in \mathbb{N}$ Elementen,
(c) Die Anzahl der bijektiven Abbildungen $f : M \to M$ für eine Menge M mit $n \in \mathbb{N}$ Elementen,
(d) die Anzahl der Äquivalenzrelationen auf einer Menge mit 3 Elementen.

Aufgabe 10: Untersuchen Sie, ob es sich bei den folgenden Relationen auf der Menge M um Äquivalenzrelationen handelt. Falls ja, geben Sie die Äquivalenzklassen an.

(a) $M = \mathbb{R}$, $a \sim b \Leftrightarrow a^2 - b^2 = 0$.
(b) $M = \mathbb{Z}$, $a \sim b \Leftrightarrow a + b = 1$.
(c) $M = \mathbb{Z}$, $R = \{(a, b) \in \mathbb{Z} \times \mathbb{Z} \mid a \cdot b \geq 0\}$.
(d) $M = \mathbb{N}_0$, $R = \{(a, b) \in \mathbb{N}_0 \times \mathbb{N}_0 | \, \exists c \in \mathbb{N}_0 : a + b > c\}$.

Aufgabe 11: Untersuchen Sie, ob es sich bei den folgenden Relationen um Äquivalenzrelationen auf der Menge $\mathbb{Z}$ handelt, und bestimmen Sie gegebenenfalls die Äquivalenzklassen und die Quotientenmenge:

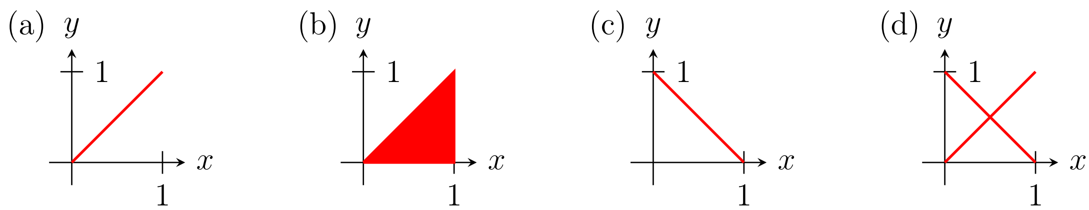
(a) $a \sim b \Leftrightarrow a + b$ gerade.
(b) $R = \{(a, b) \in \mathbb{Z} \times \mathbb{Z}|$ a+b ist nicht durch 7 teilbar$\}$

Aufgabe 12: Sei M eine Menge und $R, S \subseteq M \times M$ Äquivalenzrelationen auf M. Untersuchen Sie, ob dann auch

(a) $S \cap R \subseteq M \times M$,
(b) $S \cup R \subseteq M \times M$

Äquivalenzrelationen auf M sind.

Aufgabe 13: Welche der dargestellten Teilmengen $R \subseteq [0, 1] \times [0, 1]$ beschreiben Äquivalenzrelationen auf $[0, 1]$? Geben Sie jeweils die Äquivalenzklassen von $x \in [0, 1]$ an.



Aufgabe 14: Wir betrachten die Relation

$$(x, y) \sim (x', y') \quad \Leftrightarrow \quad \exists c \in \mathbb{R}, c > 0 : \ (x' = cx) \wedge (y' = cy)$$

auf der Menge $M = \mathbb{R}^2 \setminus \{(0, 0)\}$.

(a) Beweisen Sie, dass diese Relation eine Äquivalenzrelation auf der Menge M ist.
(b) Bestimmen Sie die Äquivalenzklassen und die Quotientenmenge $M/\sim$ und skizzieren Sie die Äquivalenzklassen.
(c) Zeigen Sie, dass es eine bijektive Abbildung $g : S^1 \to M/\sim$ gibt, wobei S^1 den Einheitskreis: $S^1 = \{(x, y) \mid x^2 + y^2 = 1\}$ bezeichnet. Bestimmen Sie ihre Umkehrfunktion.

Aufgabe 15:
Seien A, B, C Mengen und $f : A \to B$, $g : B \to C$ Abbildungen. Beweisen Sie:

(a) Ist die Abbildung $g \circ f : A \to C$ injektiv, dann ist auch f injektiv.
(b) Ist die Abbildung $g \circ f : A \to C$ surjektiv, dann ist auch g surjektiv.
(c) Ist die Abbildung $g \circ f : A \to C$ bijektiv, dann ist f injektiv und g surjektiv.
(d) Geben Sie ein Beispiel an, in dem $g \circ f$ bijektiv, g nicht injektiv und f nicht surjektiv ist.

Aufgabe 16: Sei $f : M \to N$ eine Abbildung und $M_1, M_2 \subseteq M$ and $N_1, N_2 \subseteq N$ Teilmengen. Beweisen Sie die folgenden Aussagen oder widerlegen Sie sie durch ein Gegenbeispiel.

(a) $f(M_1 \cup M_2) = f(M_1) \cup f(M_2)$,
(b) $f(M_1 \cap M_2) = f(M_1) \cap f(M_2)$,
(c) $f^{-1}(N_1 \cup N_2) = f^{-1}(N_1) \cup f^{-1}(N_2)$,
(d) $f^{-1}(N_1 \cap N_2) = f^{-1}(N_1) \cap f^{-1}(N_2)$.

Aufgabe 17: Seien M, N Mengen und $\mathrm{Pot}(M)$ die Potenzmenge von M. Beweisen Sie:

(a) Eine Abbildung $f : M \to N$ ist surjektiv genau dann, wenn für beliebige Abbildungen $g, g' : N \to P$ gilt: $g \circ f = g' \circ f \Rightarrow g = g'$.
(b) Eine Abbildung $f : M \to N$ ist injektiv genau dann, wenn für beliebige Abbildungen $h, h' : L \to M$ gilt: $f \circ h = f \circ h' \Rightarrow h = h'$.
(c) Es gibt keine surjektive Abbildung $f : M \to \mathrm{Pot}(M)$.

Hinweis: Betrachten Sie in (c) für $f : M \to \mathrm{Pot}(M)$ die Menge $\{x \in M | \, x \notin f(M)\}$.

Aufgabe 18: Sind die folgenden Aussagen wahr oder falsch? Begründen Sie oder widerlegen Sie durch ein Gegenbeispiel.

(a) Sind M, N endliche Mengen und $f : M \to N$ eine injektive Abbildung, so hat M mindestens so viele Elemente wie N.
(b) Ist $g : \mathbb{N} \to \mathbb{N}$ surjektiv, so existiert eine Abbildung $f : \mathbb{N} \to \mathbb{N}$ mit $f \circ g = \mathrm{id}_{\mathbb{N}}$.
(c) Ist $g : \mathbb{N} \to \mathbb{N}$ surjektiv, so existiert eine Abbildung $f : \mathbb{N} \to \mathbb{N}$ mit $g \circ f = \mathrm{id}_{\mathbb{N}}$.
(d) Ist $g : \mathbb{R} \to \mathbb{R}$ eine Abbildung, so dass für jede Abbildung $f : \mathbb{R} \to \mathbb{R}$ mit $g \circ f = 0$ folgt, dass auch $f = 0$ gilt, so ist g injektiv.
(e) Sind M, N, P beliebige Mengen und $f : M \to N$, $g : N \to P$ Abbildungen, so dass $g \circ f : M \to P$ bijektiv ist, so ist f surjektiv und g injektiv.
(f) Sind M, N, P beliebige Mengen und $f : M \to N$, $g : N \to P$ Abbildungen, so dass $g \circ f : M \to P$ bijektiv ist, so ist f injektiv und g surjektiv.

(g) Sind M, N, P beliebige Mengen, $f : M \to N$ eine injektive Abbildung und $g : N \to P$ eine surjektive lineare Abbildung, so ist $g \circ f : M \to P$ bijektiv.
(h) Jede surjektive Abbildung $f : \{1, ..., n\} \to \{1, ..., n\}$ ist bijektiv.
(i) Eine Abbildung $f : \mathbb{N} \to \mathbb{N}$ ist bijektiv genau dann, wenn sie injektiv ist.
(j) Es gibt unendlich viele injektive Abbildungen $f : \mathbb{R} \to \mathbb{R}$.
(k) Zu jeder Menge M gibt es genau eine Abbildung $f : \emptyset \to M$.
(l) Zu jeder Menge M gibt es genau eine Abbildung $f : M \to \emptyset$.

Aufgabe 19: Wir betrachten die Verknüpfung $\circ : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}^2$ mit $(a, b) \circ (c, d) = (a + c, b)$ für alle $a, b, c, d \in \mathbb{R}$. Untersuchen Sie diese Verknüpfung auf

(a) Assoziativität,
(b) Kommutativität,
(c) Existenz eines neutralen Elements.

Aufgabe 20: Wir betrachten die Menge $\mathbb{R}^2$ mit den Verknüpfungen $+, \cdot : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}^2$

$$(a, b) + (c, d) = (a + c, b + d) \qquad (a, b) \cdot (c, d) = (ac + \lambda bd, ad + bc) \qquad \text{für } \lambda \in \{0, 1\}.$$

(a) Zeigen Sie, dass $(\mathbb{R}^2, +, \cdot)$ ein kommutativer unitaler Ring ist.
(b) Bestimmen Sie für $\lambda = 0$ und $\lambda = 1$ die Nullteiler von $(\mathbb{R}^2, +, \cdot)$.
(c) Bestimmen Sie für $\lambda = 0$ und $\lambda = 1$ die Einheitengruppe von $(\mathbb{R}^2, \cdot)$.

Aufgabe 21: Seien $(G, \circ)$ und $(H, \circ')$ Gruppen. Beweisen Sie, dass dann auch die Menge $G \times H$ mit der Verknüpfung $\circ : (G \times H) \times (G \times H) \to G \times H$, $(g_1, h_1) \circ (g_2, h_2) = (g_1 \circ g_2, h_1 \circ' h_2)$ für alle $g_1, g_2 \in G$, $h_1, h_2 \in H$ eine Gruppe ist.

Aufgabe 22: Stellen Sie Multiplikationstafeln für die Gruppen $(\mathbb{Z}/2\mathbb{Z}, +)$ und $(\mathbb{Z}/3\mathbb{Z}, +)$ auf. Verwenden Sie dabei in der Multiplikationstafel für $(\mathbb{Z}/n\mathbb{Z}, +)$ nur die Ausdrücke $[0], ..., [n-1]$.

Aufgabe 23: Wir betrachten die Menge $\mathbb{Z}/8\mathbb{Z}$ mit der Verknüpfung $\cdot : \mathbb{Z}/8\mathbb{Z} \times \mathbb{Z}/8\mathbb{Z} \to \mathbb{Z}/8\mathbb{Z}$, $([k], [m]) \mapsto [km]$.

(a) Zeigen Sie, dass $(\mathbb{Z}/8\mathbb{Z}, \cdot)$ ein Monoid ist.
(b) Bestimmen Sie die Einheitengruppe $((\mathbb{Z}/8\mathbb{Z})^\times, \cdot)$ des Monoids $(\mathbb{Z}/8\mathbb{Z}, \cdot)$.
(c) Geben Sie eine Multiplikationstafel für die Einheitengruppe $\mathbb{Z}/8\mathbb{Z}$ an.

Aufgabe 24: Sei $(G, \cdot)$ eine Gruppe.

(a) Zeigen Sie: $h \sim k \Leftrightarrow \exists g \in G : g \cdot h \cdot g^{-1} = k$ ist eine Äquivalenzrelation auf G.
(b) Bestimmen Sie die Äquivalenzklasse des neutralen Elements $e \in G$.
(c) Bestimmen Sie alle Äquivalenzklassen für den Fall, dass G eine abelsche Gruppe ist.

Aufgabe 25: Zeigen Sie, dass für $n \in \mathbb{N}$ die Untergruppen der Gruppe $(\mathbb{Z}/n\mathbb{Z}, +)$ genau die Teilmengen der Form $G_m = \{[km] |\ k \in \mathbb{Z}\}$ mit $m \in \{0, 1, ..., n-1\}$ und m teilt n sind.

Aufgabe 26:
Untersuchen Sie, ob es sich bei den angegebenen Teilmengen $H \subseteq G$ um Untergruppen der Gruppe $(G, \circ)$ handelt:

(a) $(G,\circ) = (\mathbb{Q},+)$, $H = \{\frac{k}{3} | k \in \mathbb{Z}\}$,
(b) $(G,\circ) = (\mathbb{Q}\setminus\{0\},\cdot)$, $H = \{\frac{k}{3} |\, k \in \mathbb{Z}\setminus\{0\}\}$,
(c) $(G,\circ) = (S_n,\circ)$, $H = \{\pi \in S_n | \pi(k)$ gerade für alle geraden $k \in \{1,...,n\}\}$,
(d) $(G,\circ)$ beliebig, $H = \{g \in G | g \circ g = g\}$.

Aufgabe 27: Wir betrachten für $n \in \mathbb{N}$ die symmetrische Gruppe S_n, also die Gruppe der bijektiven Abbildungen $\pi : \{1,2,...,n\} \to \{1,2,...,n\}$.

(a) Zeigen Sie, dass S_n abelsch ist genau dann, wenn $n \leq 2$.
(b) Zeigen Sie, dass die Menge $\{\pi \in S_n |\, \pi(n) = n\}$ eine Untergruppe von S_n ist.
(c) Geben Sie einen injektiven Gruppenhomomorphismus $\phi : S_{n-1} \to S_n$ an.

Aufgabe 28: Sei G eine Gruppe und $G_1 \subseteq G$, $G_2 \subseteq G$ Untergruppen. Beweisen Sie die folgenden Aussagen:

(a) $G_1 \cap G_2$ ist eine Untergruppe von G.
(b) $G_1 \cup G_2$ ist genau dann eine Untergruppe von G, wenn gilt $G_1 \subseteq G_2$ oder $G_2 \subseteq G_1$.

Aufgabe 29:

(a) Sei $(G,\cdot)$ eine Gruppe mit neutralem Element e, so dass für alle $g \in G$ gilt $g \cdot g = e$. Beweisen Sie, dass G dann abelsch ist.
(b) Sei $G = \{g_1,...,g_n\}$ mit der Verknüpfung $\cdot : G \times G \to G$ eine endliche abelsche Gruppe mit neutralem Element e. Beweisen Sie, dass dann gilt $g_1^2 \cdot g_2^2 \cdots g_n^2 = e$.

Aufgabe 30: Wahr oder falsch? Beweisen Sie die Aussage oder widerlegen Sie sie durch ein Gegenbeispiel.

(a) Zu jedem $n \in \mathbb{N}$ gibt es genau einen Gruppenhomomorphismus $\phi : (\mathbb{Z}/n\mathbb{Z},+) \to (\mathbb{Z},+)$.
(b) Ist $(G,\cdot)$ eine endliche Gruppe und $(H,\cdot)$ eine Gruppe mit unendlich vielen Elementen, so gibt es genau einen Gruppenhomomorphismus $\phi : G \to H$.
(c) Ist $(A,+)$ eine abelsche Gruppe und enthält eine Gruppe $(G,\cdot)$ Elemente $g,h \in G$ mit $g \cdot h \cdot g^{-1} \cdot h^{-1} \neq e$, so gibt es keinen surjektiven Gruppenhomomorphismus $\phi : A \to G$.

Aufgabe 31:

(a) Zeigen Sie, dass $\phi : (\mathbb{Z}/4\mathbb{Z},+) \to (\mathbb{Z}/4\mathbb{Z},+)$, $[m] \mapsto [2m]$ ein Gruppenhomomorphismus ist, und bestimmen Sie den Kern $\ker(\phi)$ und das Bild $\phi(\mathbb{Z}/4\mathbb{Z})$.
(b) Zeigen Sie, dass das Bild $\phi(\mathbb{Z}/4\mathbb{Z})$ isomorph zur Gruppe $\mathbb{Z}/2\mathbb{Z}$ ist.

Aufgabe 32:
Untersuchen Sie, ob die folgenden Gruppen isomorph sind:

(a) S_4 und $\mathbb{Z}/4\mathbb{Z}$
(b) $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ und $\mathbb{Z}/4\mathbb{Z}$

Aufgabe 33: Sei $(G,\cdot_G)$ eine Gruppe. Zeigen Sie, dass die Menge $\mathrm{Aut}(G)$ der Gruppenautomorphismen $\phi : G \to G$ eine Untergruppe der Gruppe $(\mathrm{Bij}(G,G),\circ)$ ist.

Aufgabe 34: Wahr oder falsch? Beweisen Sie oder widerlegen Sie sie durch ein Gegenbeispiel.

(a) Das Bild einer abelschen Gruppe unter einem beliebigen Gruppenhomomorphismus ist abelsch.
(b) Haben zwei endliche Gruppen die gleiche Anzahl von Elementen, so sind sie isomorph.
(c) Sind G, H, K Gruppen mit $G \cong H$ und $H \cong K$, so folgt $G \cong K$.
(d) In jeder Gruppe $(\mathbb{Z}/n\mathbb{Z}, +)$ mit $n \geq 3$ gibt es ein Element $a \in \mathbb{Z}/n\mathbb{Z} \setminus \{[1]\}$ mit $a \cdot a = [1]$.

Aufgabe 35: Sei $n \in \mathbb{N}$ ungerade.

(a) Zeigen Sie, dass zu jedem Element $a \in \mathbb{Z}/n\mathbb{Z}$ ein Element $b \in \mathbb{Z}/n\mathbb{Z}$ mit $a = b + b$ existiert.
(b) Sei nun $(A, +)$ eine abelsche Gruppe mit n Elementen. Zeigen Sie, dass zu jedem $a \in A$ ein $b \in A$ mit $a = b + b$ existiert.

Aufgabe 36: Der Mengenring
Sei X eine Menge mit Potenzmenge $\mathrm{Pot}(X)$. Wir betrachten die Verknüpfungen

$$+ : \mathrm{Pot}(X) \times \mathrm{Pot}(X) \to \mathrm{Pot}(X), \qquad (M, N) \mapsto (M \cup N) \setminus (M \cap N)$$
$$\circ : \mathrm{Pot}(X) \times \mathrm{Pot}(X) \to \mathrm{Pot}(X), \qquad (M, N) \mapsto M \cap N.$$

Zeigen Sie, dass $(\mathrm{Pot}(X), +, \circ)$ ein kommutativer Ring mit Einselement ist.

Aufgabe 37: Der Polynomring
Sei $(R, +_R, \cdot_R)$ ein kommutativer Ring mit Einselement. Wir betrachten die Menge $R[x]$ der Abbildungen $p : \mathbb{N}_0 \to R$ mit $p(n) = 0$ für alle bis auf endliche viele $n \in \mathbb{N}_0$ und die Verknüpfungen $+ : R[x] \times R[x] \to R[x]$, $(p, q) \mapsto p + q$ und $\cdot : R[x] \times R[x] \to R[x]$, $(p, q) \mapsto p \cdot q$

$$\begin{aligned}(p + q)(n) &:= p(n) +_R q(n)\\ (p \cdot q)(n) &:= \Sigma_{k=0}^{n}\, p(k) \cdot_R q(n - k)\\ &= p(0) \cdot_R q(n) +_R p(1) \cdot_R q(n - 1) +_R ... +_R p(n - 1) \cdot_R q(1) +_R p(n) \cdot_R q(0)\end{aligned}$$

für alle $n \in \mathbb{N}_0$. Beweisen Sie, dass $(R[x], +, \cdot)$ ein kommutativer Ring mit Einselement ist.

Aufgabe 38: Sei $M \neq \emptyset$ und $(R, +_R, \cdot_R)$ ein Ring. Wir betrachten die Menge $\mathrm{Abb}(M, R)$ der Abbildungen $f : M \to R$ mit den Verknüpfungen $+, \cdot : \mathrm{Abb}(M, R) \times \mathrm{Abb}(M, R) \to \mathrm{Abb}(M, R)$

$$(f + g)(m) := f(m) +_R g(m) \qquad (f \cdot g)(m) := f(m) \cdot_R g(m) \qquad \forall m \in M.$$

(a) Zeigen Sie, dass $(\mathrm{Abb}(M, R), +, \cdot)$ ein Ring ist.
(b) Zeigen Sie, dass $(\mathrm{Abb}(M, R), +, \cdot)$ kommutativ ist genau dann, wenn $(R, +, \cdot)$ kommutativ ist und unital genau dann, wenn $(R, +, \cdot)$ unital ist.
(c) Hat auch die Menge $\mathrm{Abb}(\emptyset, R)$ die Struktur eines Rings? Wenn ja, ist er kommutativ und unital?

Aufgabe 39: Untersuchen Sie, welche Elemente im Restklassenring $\mathbb{Z}/10\mathbb{Z}$ multiplikative Inverse haben, und bestimmen Sie diese multiplikativen Inversen.

Aufgabe 40: Beweisen Sie, dass in jedem endlichen kommutativen nullteilerfreien unitalen Ring $(R, +, \cdot)$ ein $m \in \mathbb{N}$ existiert mit $m1_R = 1_R + ... + 1_R = 0_R$. Zeigen Sie, dass für $R \neq \{0\}$ die kleinste solche Zahl $m \in \mathbb{N}$ eine Primzahl ist.

Aufgabe 41: Seien $(R, \cdot, +)$ und $(S, +', \cdot')$ Ringe. Zeigen Sie, dass dann auch die Menge $R \times S$ mit den Verknüpfungen $\dagger, \circ : (R \times S) \times (R \times S) \to R \times S$

$$(r, s) \dagger (r', s') := (r + r', s +' s') \qquad (r, s) \circ (r', s') := (r \cdot r', s \cdot' s') \qquad \forall r, r' \in R, s, s' \in S$$

ein Ring ist und dass dieser unital ist, wenn R und S unital sind. Folgt, dass auch $R \times S$ ein Körper ist, wenn R und S Körper sind?

Aufgabe 42: Sei $(A, +)$ eine abelsche Gruppe. Zeigen Sie, dass die Menge $\mathrm{End}(A)$ der Gruppenendomorphismen $\phi : A \to A$ mit den Verknüpfungen $+, \circ : \mathrm{End}(A) \times \mathrm{End}(A) \to \mathrm{End}(A)$

$$(\phi + \psi)(a) = \phi(a) + \psi(a) \qquad (\phi \circ \psi)(a) = \phi(\psi(a)) \qquad \forall a \in A$$

einen unitalen Ring bildet.

Aufgabe 43:

(a) Zeigen Sie, dass jeder Körperhomomorphismus injektiv ist.
(b) Zeigen Sie, dass jeder *endliche* kommutative nullteilerfreie unitale Ring R mit $1 \neq 0$ ein Körper ist.

Hinweis: Betrachten Sie in (b) für $r \in R \setminus \{0\}$ die Abbildung $\phi_r : R \to R$, $r' \mapsto r \cdot r'$.

Aufgabe 44: Wir betrachten den Körper $(\mathbb{Z}/7\mathbb{Z}, +, \cdot)$

(a) Bestimmen Sie für jedes Element von $\mathbb{Z}/7\mathbb{Z} \setminus \{[0]\}$ das multiplikative Inverse.
(b) Untersuchen Sie, welche Elemente des Körpers $\mathbb{Z}/7\mathbb{Z}$ Quadrate sind, d. h. sich als $[m] \cdot [m]$ mit $m \in \{0, 1, ..., 6\}$ schreiben lassen.

Aufgabe 45: Sei $q \in \mathbb{Q}$ eine Zahl, die nicht Quadrat einer rationalen Zahl ist: $q \neq r^2$ für alle $r \in \mathbb{Q}$. Zeigen Sie, dass $\mathbb{Q}^2$ mit den Verknüpfungen
$+, \cdot : \mathbb{Q}^2 \times \mathbb{Q}^2 \to \mathbb{Q}^2$

$$(a, b) + (c, d) = (a + b, c + d) \qquad (a, b) \cdot (c, d) = (ac + qbd, ad + bc) \qquad \forall a, b, c, d \in \mathbb{Q}$$

ein Körper und $\phi : \mathbb{Q}^2 \to \mathbb{R}$, $(a, b) \to a + b\sqrt{q}$ ein Körperhomomorphismus ist.

Aufgabe 46: Sei $R \neq \{0\}$ ein Integritätsbereich, d. h. kommutativer, nullteilerfreier Ring mit Einselement. Wir betrachten die Menge $Q = \{(r, s) \,|\, r \in R, s \in R \setminus \{0\}\}$ mit der Relation $(r, s) \sim (r', s') \Leftrightarrow rs' = r's$.

(a) Zeigen Sie, dass es sich dabei um eine Äquivalenzrelation auf Q handelt.
(b) Auf der Quotientenmenge $Q/_\sim$ definieren wir Verknüpfungen $+, \cdot : Q/_\sim \times Q/_\sim \to Q/_\sim$

$$[(r, s)] + [(r', s')] := [(rs' + r's, ss')] \qquad [(r, s)] \cdot [(r', s')] := [(rr', ss')]$$

für alle $r, r' \in R$, $s, s' \in R \setminus \{0\}$. Zeigen Sie, dass $Q/_\sim$ mit diesen Verknüpfungen ein kommutativer unitaler Ring ist.
(c) Zeigen Sie, dass jedes Element in $Q/_\sim$ außer der Null ein multiplikatives Inverses besitzt und $Q/_\sim$ somit ein Körper ist. Diesen bezeichnet man als den **Quotientenkörper** des Integritätsbereichs R.

(d) Bestimmen Sie den Quotientenkörper $Q/_\sim$ für die Integritätsbereiche $R = \mathbb{Z}$ und $\mathbb{R} = \mathbb{R}[x]$.

Aufgabe 47: Wahr oder falsch? Begründen Sie die Aussage, oder widerlegen Sie sie durch ein Gegenbeispiel.

(a) In jedem Körper $(\mathbb{K}, +, \cdot)$ gilt: ist $a \in \mathbb{K}$ mit $a \cdot a = 1$, so folgt $a = 1$ oder $a = -1$.
(b) In jedem Körper $(\mathbb{K}, +, \cdot)$ und für jedes $a \in \mathbb{K} \setminus \{0\}$ gilt: die Abbildung $\phi_a : \mathbb{K} \to \mathbb{K}$, $b \mapsto a \cdot b$ ist bijektiv.
(c) In jedem kommutativen unitalen Ring $(R, +, \cdot)$ und für jedes $a \in R \setminus \{0\}$ gilt: ist die Abbildung $\phi_a : R \to R$, $b \mapsto a \cdot b$ surjektiv, so hat a ein multiplikatives Inverses.
(d) In jedem Körper $(\mathbb{K}, +, \cdot)$ gilt: die Abbildung $\phi : \mathbb{K} \to \mathbb{K}$, $k \mapsto k + k + k$ ist bijektiv.

Aufgabe 48: Skizzieren Sie die folgenden Teilmengen $M \subseteq \mathbb{C}$:

(a) $M = \{z \in \mathbb{C} | \mathrm{Re}\,(z) < \mathrm{Im}\,(z)\}$
(b) $M = \{z \in \mathbb{C} | \arg(z) \in [\frac{\pi}{4}, \frac{3\pi}{8})\}$
(c) $M = \{z \in \mathbb{C} || z - 1 + i| \leq 1\}$
(d) $M = \{z \in \mathbb{C} || 3z - 1| < 2\}$
(e) $M = \{z \in \mathbb{C} | \mathrm{Re}\,(z \cdot (1 + \mathrm{i})) = 0\}$
(f) $M = \{z \in \mathbb{C} | \mathrm{Im}\,(z \cdot (1 + \mathrm{i})) = 3\}$
(g) $M = \{z \in \mathbb{C} | \mathrm{Im}\,(z \cdot \mathrm{i}) \leq 0\}$
(h) $M = \{z \in \mathbb{C} | \arg(z^n) \in [0, \pi)\}$
(i) $M = \{z \in \mathbb{C} || \bar{z} - \mathrm{i}| = 3\}$
(j) $M = \{z \in \mathbb{C} | z^3 = 8\}$
(k) $M = \{z \in \mathbb{C} || z - \mathrm{i} - 1| = |z + \mathrm{i} + 1|\}$
(l) $M = \{z \in \mathbb{C} | z^2 = \mathrm{i}\}$

Aufgabe 49:

(a) Bestimmen Sie den Real- und Imaginärteil der folgenden komplexen Zahlen:

$$z_1 = \frac{1}{4 - 3\mathrm{i}}, \qquad z_2 = 2e^{\mathrm{i}\pi/6}, \quad z_3 = \frac{2 - 2\mathrm{i}}{1 + \mathrm{i}}, \quad z_4 = \frac{1}{(1 + \mathrm{i})^4}$$

(b) Geben Sie die Polardarstellung der folgenden komplexen Zahlen an:

$$w_1 = -\mathrm{i}, \qquad w_2 = \sqrt{3} - \mathrm{i}, \qquad w_3 = \frac{1}{-5 + 5\mathrm{i}}, \qquad w_4 = \frac{1}{\mathrm{i}} + \frac{\mathrm{i}}{1 + \mathrm{i}}$$

Aufgabe 50: Geben Sie die folgenden komplexen Zahlen in der Form $x + \mathrm{i}y$ mit $x, y \in \mathbb{R}$ an und bestimmen Sie Ihren Betrag.

$$z_1 = \frac{3 - 5\mathrm{i}}{2 + 2\mathrm{i}},\ z_2 = \frac{2 - \mathrm{i}}{3 + 4\mathrm{i}} + \frac{\mathrm{i}}{1 + \mathrm{i}},\ z_3 = (2 - \mathrm{i})^3,\ z_4 = \left(\frac{1 - \mathrm{i}}{1 + \mathrm{i}}\right)^2,\ z_5 = \frac{1 + 3\mathrm{i}}{1 - \mathrm{i}} \cdot \frac{4 + \mathrm{i}}{2 - \mathrm{i}},\ z_6 = (\sqrt{3} - \mathrm{i})^{10}$$

Aufgabe 51: Sei $w = re^{\mathrm{i}\psi}$ mit $r > 0$ und $\psi \in [0, 2\pi)$. Bestimmen Sie für $n \in \mathbb{N}$ alle komplexen Lösungen der Gleichung $z^n = w$ und skizzieren Sie diese.

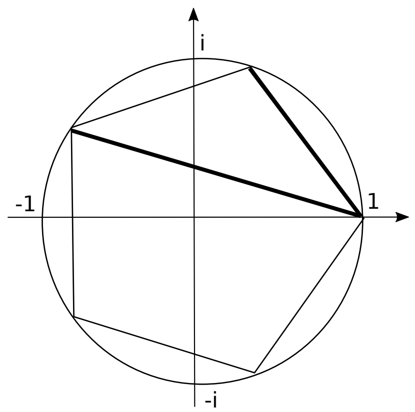
Aufgabe 52: Stellen Sie die folgenden komplexen Zahlen in der Form $x + \mathrm{i}y$ mit $x, y \in \mathbb{R}$ dar, geben Sie ihre Beträge und Argumente an und skizziere Sie sie in der komplexen Zahlenebene:

1. $(1-\mathrm{i})^4$,
2. $(1-\mathrm{i}\sqrt{3})^3/8$,
3. $(7-3\mathrm{i})/(1+\mathrm{i})$,
4. i^n für $n \in \mathbb{N}$.

Aufgabe 53: Beweisen Sie, dass für jedes $n \in \mathbb{N}$ und $k \in \{1, ..., n-1\}$ die Zahl $z = e^{2\pi \mathrm{i}k/n}$ die Gleichung $1 + z + z^2 + ... + z^{n-1} = 0$ erfüllt.

Aufgabe 54: Sei $\mathbb{H} = \{z \in \mathbb{C} | \mathrm{Im}\,(z) > 0\}$ die komplexe obere Halbebene. Wir betrachten die Cayley-Abbildung $f : \mathbb{H} \to \mathbb{C}$, $z \mapsto (z-\mathrm{i})/(z+\mathrm{i})$. Zeigen Sie, dass f injektiv ist und bestimmen Sie die Bildmenge von f.

Aufgabe 55: Zwei Streckenlängen $b, c \in \mathbb{R}$ mit $b > c$ stehen im Verhältnis des Goldenen Schnitts, wenn das Verhältnis $a = b/c$ der längeren Strecke b zur kürzeren Strecke c gleich dem Verhältnis $a' = (b+c)/b$ der Summe der beiden Strecken zur längeren Strecke ist. Wir beweisen mit Hilfe komplexer Zahlen, dass die Diagonalen und die Seiten eines regelmäßigen Fünfecks im Verhältnis des Goldenen Schnitts stehen.



(a) Zeigen Sie zunächst, dass zwei Strecken $b, c \in \mathbb{R}$ mit $b > c$ genau dann im Verhältnis des goldenen Schnitts stehen, wenn $a = b/c$ die Gleichung $a^2 - a - 1 = 0$ erfüllt.
(b) Überprüfen Sie anhand der Skizze, dass das gesuchte Verhältnis gegeben ist durch $a = |1-\zeta^2|/|1-\zeta|$ mit $\zeta = e^{2\pi\mathrm{i}/5}$.
(c) Zeigen Sie, dass folgende Gleichungen gelten: $a = |1+\zeta|$, $\bar{\zeta} = \zeta^4$, $1 + \zeta + \zeta^2 + \zeta^3 + \zeta^4 = 0$.
(d) Nutzen Sie (a)-(c) um zu beweisen, dass die die Diagonalen und die Seiten eines regelmäßigen Fünfecks im Verhältnis des Goldenen Schnitts stehen.

Aufgabe 56: Sei $z = \frac{1}{2} + \frac{\mathrm{i}}{2}\sqrt{3}$. Berechnen Sie z^{2015} und $z^{2015} + z^{2017} + z^{2019}$.

Aufgabe 57: Beweisen Sie, dass für alle $z, w \in \mathbb{C}$ gilt

$$|z+w|^2 = |z|^2 + |w|^2 + 2\mathrm{Re}\,(\bar{z}w) \qquad |z+w|^2 + |z-w|^2 = 2(|z|^2 + |w|^2).$$

Aufgabe 58: Sei $(R, +, \cdot)$ ein Ring und $f : R \to S$ eine bijektive Abbildung. Beweisen Sie, dass es genau eine Ringstruktur $(S, +', \cdot')$ auf S gibt, die $f : R \to S$ zu einem Ringisomorphismus macht. Zeigen Sie, dass $(S, +', \cdot')$ unital ist, genau dann, wenn $(R, +, \cdot)$ unital ist.

Aufgabe 59: Wir betrachten die Menge $\mathbb{R}^2$ mit den Verknüpfungen $+, \cdot : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}^2$

$$(a,b) + (c,d) = (a+c, b+d) \qquad (a,b)\cdot(c,d) = (ac + \lambda bd, ad + bc) \qquad \text{für } \lambda \in \{0,1\}.$$

Überzeugen Sie sich (ohne dies aufzuschreiben), dass $(\mathbb{R}^2, +, \cdot)$ ein kommutativer Ring mit Einselement $(1,0)$ ist.

(a) Bestimmen Sie für $\lambda = 0$ und $\lambda = 1$ die Nullteiler von $(\mathbb{R}^2, +, \cdot)$.
(b) Bestimmen Sie für $\lambda = 0$ und $\lambda = 1$ die Einheitengruppe von $(\mathbb{R}^2, \cdot)$.

Aufgabe 60: Wahr oder falsch? Begründen Sie oder widerlegen Sie die Aussage.

(a) Der Ring $(\mathbb{Z}/18\mathbb{Z}, +, \cdot)$ ist ein Körper.
(b) Die Gruppen S_3 und $(\mathbb{Z}/6\mathbb{Z}, +)$ sind isomorph.
(c) Alle zweielementigen Gruppen sind isomorph.
(d) Die Gruppen $(\mathbb{Z}/6\mathbb{Z}, +)$ und $(\mathbb{Z}/8\mathbb{Z}, +)$ sind isomorph.
(e) Es gibt unendlich viele Gruppenhomomorphismen $\phi : \mathbb{Z} \to \mathbb{Z}$.
(f) Für jedes $n \in \mathbb{N}$ gibt es genau einen Gruppenhomomorphismus $\phi : \mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$.
(g) Zu zwei Gruppen G, H gibt es immer einen Gruppenhomomorphismus $\phi : G \to H$.
(h) Zu jedem Ring $(R, +, \cdot)$ mit Einselement gibt es genau einen unitalen Ringhomomorphismus $\phi : (\mathbb{Z}, +, \cdot) \to (R, +, \cdot)$.

Aufgabe 61: Wir betrachten für $n \in \mathbb{N}$, $n \geq 2$ die $n$ten Einheitswurzeln, also die Lösungen der Gleichung $z^n = 1$ in $\mathbb{C}$ und bezeichnen die Menge der $n$ten Einheitswurzeln mit G_n.

(a) Zeigen Sie, dass die $n$ten Einheitswurzeln mit der Multiplikation in $\mathbb{C}$ eine abelsche Gruppe bilden.
(b) Zeigen sie, dass für jeden Teiler m von n die Menge $H_m = \{e^{2\pi i mk/n} | k \in \mathbb{Z}\}$ eine Untergruppe von $(G_n, \cdot)$ ist.
(c) Zeigen Sie, dass die Abbildung $\phi_n : \mathbb{Z}/n\mathbb{Z} \to G_n$, $[k] \mapsto e^{2\pi i k/n}$ wohldefiniert ist und ein Gruppenisomorphismus von $(\mathbb{Z}/n\mathbb{Z}, +)$ nach $(G_n, \cdot)$ ist.
(d) Bestimmen Sie die Urbilder der Gruppen H_m unter ϕ_n.

13.3 Aufgaben zu Kapitel 3

Aufgabe 62: Wie viele Elemente hat der Vektorraum $\mathbb{K}^n$ über einem endlichen Körper $\mathbb{K}$ mit q Elementen?

Aufgabe 63: Wir betrachten den Vektorraum $\mathrm{Abb}(\mathbb{R}, \mathbb{R})$. Untersuchen Sie, ob die folgenden Teilmengen Untervektorräume von $\mathrm{Abb}(\mathbb{R}, \mathbb{R})$ sind:

(a) $M = \{f : \mathbb{R} \to \mathbb{R} | f(0) = 0\}$
(b) $M = \{f : \mathbb{R} \to \mathbb{R} | f(0) = 1\}$
(c) $M = \{f : \mathbb{R} \to \mathbb{R} | f(\mathbb{R}) \subseteq \mathbb{Q}\}$
(d) $M = \{f : \mathbb{R} \to \mathbb{R} | f(x) = 0 \ \forall x \in [2,4]\}$
(e) $M = \{f : \mathbb{R} \to \mathbb{R} | f(1) = f(2)\}$
(f) $M = \{f : \mathbb{R} \to \mathbb{R} | \exists c \in \mathbb{R} : |f(x)| \leq c \ \forall x \in \mathbb{R}\}$
(g) $M = \{f : \mathbb{R} \to \mathbb{R} | f(x) \leq f(y) \ \forall x, y \in \mathbb{R}, x < y\}$

(h) $M = \{f : \mathbb{R} \to \mathbb{R} | f(x) \geq 0 \; \forall x \in \mathbb{R}\}$
(i) $M = \{f : \mathbb{R} \to \mathbb{R} | f(2x) = 3f(x) \; \forall x \in \mathbb{R}\}$
(j) $M = \{f : \mathbb{R} \to \mathbb{R} | f(ax) = a^2 f(x) \; \forall a, x \in \mathbb{R}\}$

Aufgabe 64: Wir betrachten den reellen Vektorraum $\mathrm{Pol}(\mathbb{R}) \subseteq \mathrm{Abb}(\mathbb{R}, \mathbb{R})$ der Polynomabbildungen $p : \mathbb{R} \to \mathbb{R}$. Der **Grad** einer Polynomabbildung $p : \mathbb{R} \to \mathbb{R}$, $x \mapsto \Sigma_{k \in \mathbb{N}_0} a_k x^k$ mit $a_k = 0$ für fast alle $k \in \mathbb{N}_0$ ist definiert als $\deg(p) = -\infty$ wenn $a_k = 0 \; \forall k \in \mathbb{N}_0$ und $\deg(p) = \max\{k \in \mathbb{N}_0 | a_k \neq 0\}$ sonst. Untersuchen Sie

(a) für welche $n \in \mathbb{N} \cup \{-\infty\}$ die Polynomabbildungen vom Grad n einen Untervektorraum von $\mathrm{Pol}(\mathbb{R})$ bilden,
(b) für welche $n \in \mathbb{N} \cup \{-\infty\}$ die Polynomabbildungen vom Grad $\leq n$ einen Untervektorraum von $\mathrm{Pol}(\mathbb{R})$ bilden.

Aufgabe 65: Wir betrachten den Vektorraum $\mathrm{Abb}(\mathbb{N}_0, \mathbb{Q})$ der Folgen mit Koeffizienten in $\mathbb{Q}$. Untersuchen Sie, ob die folgenden Teilmengen von $\mathrm{Abb}(\mathbb{N}_0, \mathbb{Q})$ Untervektorräume sind:

(a) $M = \{f : \mathbb{N}_0 \to \mathbb{Q} | f(k) = 0$ für $k \in \mathbb{N}_0$ gerade$\}$
(b) $M = \{f : \mathbb{N}_0 \to \mathbb{Q} | f(k) = 1$ für $k \in \mathbb{N}_0$ gerade$\}$
(c) $M = \mathbb{Q}[x]$
(d) $M = \{f : \mathbb{N}_0 \to \mathbb{Q} | f(k) = f(k+n) \; \forall k \in \mathbb{N}_0\}$ für ein festes $n \in \mathbb{N}$.

Aufgabe 66: Sei V ein $\mathbb{K}$-Vektorraum und $U_1, U_2, U_3 \subseteq V$ Untervektorräume.

(a) Zeigen Sie, dass $U_1 \cup U_2$ genau dann ein Untervektorraum von V ist, wenn $U_1 \subseteq U_2$ oder $U_2 \subseteq U_1$ gilt.
(b) Sei $U_1 \cup U_2 \cup U_3$ ein Untervektorraum von V. Folgt daraus, dass $U_1 \subseteq U_2 \cup U_3$ oder $U_2 \subseteq U_1 \cup U_3$ oder $U_3 \subseteq U_1 \cup U_2$ gilt? (Beweis oder Gegenbeispiel).

Hinweis: Betrachten Sie in (b) den Vektorraum $(\mathbb{Z}/2\mathbb{Z})^2$ über dem Körper $\mathbb{Z}/2\mathbb{Z}$.

Aufgabe 67: Wir betrachten den reellen Vektorraum $\mathbb{R}^3$ und für $s \in \mathbb{R}$ die Teilmenge $E_s = \{(x, y, z)^T \in \mathbb{R}^3 | y + sz = 0\}$.

(a) Zeigen Sie, dass für alle $s \in \mathbb{R}$ die Teilmenge $E_s \subseteq \mathbb{R}^3$ ein Untervektorraum des $\mathbb{R}^3$ ist.
(b) Bestimmen Sie den Schnitt $E_s \cap G$ von E_s mit der Geraden $G = \{(0, \alpha, \alpha)^T | \alpha \in \mathbb{R}\}$.
(c) Bestimmen Sie den Untervektorraum $\cap_{s \in \mathbb{R}} E_s$.
(d) Skizzieren Sie E_s, G, $E_s \cap G$ und $\cap_{s \in \mathbb{R}} E_s$.

Aufgabe 68: Welche der folgenden Teilmengen M der angegebenen $\mathbb{K}$-Vektorräume V sind Untervektorräume?

(a) $\mathbb{K} = \mathbb{R}$, $V = \mathbb{C}$, $M = \{z \in \mathbb{C} | \mathrm{Re}(z) = 0\}$,
(b) $\mathbb{K} = \mathbb{R}$, $V = \mathbb{R}[x]$, $M = \{p : \mathbb{N}_0 \to \mathbb{R} | \, p(0) \cdot p(1) = 0\}$,
(c) $\mathbb{K} = \mathbb{Z}/11\mathbb{Z}$, $V = \mathbb{Z}/11\mathbb{Z}$, $M \subseteq \mathbb{Z}/11\mathbb{Z}$ ist eine Teilmenge mit 5 Elementen,
(d) $\mathbb{K}$ ist ein beliebiger Körper, $V = \mathbb{K}^2$, $M = \mathbb{K}e_1 \cup \mathbb{K}e_2$.

Aufgabe 69: Sei M eine Menge. Wir betrachten den $\mathbb{K}$-Vektorraum $\mathrm{Abb}(M, \mathbb{K})$ für einen Körper $\mathbb{K}$. Für $m \in M$ definieren wir $\delta_m : M \to \mathbb{K}$ durch $\delta_m(m) = 1$ und $\delta_m(m') = 0$ für alle $m' \in M \setminus \{m\}$. Bestimmen Sie $\mathrm{span}_{\mathbb{K}}(\{\delta_m | m \in M\})$.

Aufgabe 70: Sei $\mathbb{K}$ ein endlicher Körper mit q Elementen. Wir betrachten für $n \in \mathbb{N}$ den $\mathbb{K}$-Vektorraum $\mathbb{K}^n$. Wie in der Vorlesung ist eine Gerade in $\mathbb{K}^n$ ein Untervektorraum der Form $\mathbb{K}v = \{\alpha v | \alpha \in \mathbb{K}\}$ mit $v \in \mathbb{K}^n \setminus \{0\}$. Wie viele verschiedene Geraden gibt es im $\mathbb{K}^n$?

Aufgabe 71: Sei V ein $\mathbb{K}$-Vektorraum und $v_1, v_2, v_3 \in V$, so dass (v_1, v_2), (v_1, v_3) und (v_2, v_3) linear unabhängig sind. Folgt dann auch, dass (v_1, v_2, v_3) linear unabhängig ist?

Aufgabe 72: Wir betrachten den $\mathbb{Q}$-Vektorraum $\mathbb{Q}^4$. Untersuchen Sie, ob die Vektoren $v_1, ..., v_4$ in $\operatorname{span}_{\mathbb{Q}}(M)$ enthalten sind und schreiben Sie sie in diesem Fall als Linearkombinationen von Vektoren in M.

$$M = \left\{ \begin{pmatrix} 5 \\ 0 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 3 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 1 \\ 1 \end{pmatrix} \right\}$$

$$v_1 = \begin{pmatrix} 3 \\ 2 \\ 5 \\ 0 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 5 \\ 2 \\ 4 \\ 0 \end{pmatrix}, \quad v_3 = \begin{pmatrix} -4 \\ -3 \\ -9 \\ -5 \end{pmatrix}, \quad v_4 = \begin{pmatrix} 0 \\ -2 \\ -4 \\ 0 \end{pmatrix}.$$

Aufgabe 73: Sei p eine Primzahl und $n \geq 2$. Zeigen Sie, dass der Vektorraum $(\mathbb{Z}/p\mathbb{Z})^n$ Vereinigung von $p+1$ Untervektorräumen $U_0, U_1, ..., U_p \subseteq V$ mit $U_i \neq V$ ist.

Aufgabe 74: Untersuchen Sie, ob es sich bei den Teilmengen $\{v_1, v_2, v_3\} \subseteq \mathbb{R}^3$ um Basen des $\mathbb{R}^3$ handelt. Wenn ja, stellen Sie die Vektoren e_1, e_2 und e_3 als Linearkombination der Vektoren v_1, v_2, v_3 dar.

(a)

$$v_1 = \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 4 \\ 1 \\ 1 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 3 \\ -1 \\ 2 \end{pmatrix}$$

(b)

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}$$

(c)

$$v_1 = \begin{pmatrix} 1 \\ 3 \\ -1 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 2 \\ 2 \\ 3 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 1 \\ 4 \\ 1 \end{pmatrix}$$

Aufgabe 75: Untersuchen Sie, ob die Mengen $M \subseteq V$ in den angegebenen $\mathbb{K}$-Vektorräumen V linear unabhängig sind:

(a) $\mathbb{K} = \mathbb{Q}$, $V = \mathbb{R}$, $M = \{\sqrt{2}, 1\}$,
(b) $\mathbb{K} = \mathbb{R}$, $V = \mathbb{C}$, $M = \{1 + 3\mathrm{i}, e^{3\pi\mathrm{i}/2}\}$,
(c) $\mathbb{K} = \mathbb{Q}$, $V = \mathbb{C}$, $M = \{\mathrm{i}, \sqrt{2}, 1 + 2\mathrm{i}\}$,

(d) $\mathbb{K} = \mathbb{R}$, $V = \mathrm{Abb}(\mathbb{R}, \mathbb{R})$, $M = \{f, g, h : \mathbb{R} \to \mathbb{R}\}$ mit $f : x \mapsto \cos(2x)$, $g : x \mapsto \cos(x)^2$, $h : x \mapsto 1$,

(e) $\mathbb{K} = \mathbb{Z}/5\mathbb{Z}$, $V = (\mathbb{Z}/5\mathbb{Z})^3$, $M = \{v_1, v_2, v_3\}$ mit

$$v_1 = \begin{pmatrix} [1] \\ [3] \\ [2] \end{pmatrix}, \qquad v_2 = \begin{pmatrix} [4] \\ [3] \\ [1] \end{pmatrix}, \qquad v_3 = \begin{pmatrix} [0] \\ [3] \\ [4] \end{pmatrix},$$

(f) $\mathbb{K} = \mathbb{C}$, $V = \mathbb{C}^3$, $M = \{v_1, v_2, v_3\}$

$$v_1 = \begin{pmatrix} 1 + \mathrm{i} \\ 2 \\ -\mathrm{i} \end{pmatrix}, \qquad v_2 = \begin{pmatrix} e^{5\pi\mathrm{i}/4} \\ -1 \\ \sqrt{2}\mathrm{i} \end{pmatrix}, \qquad v_3 = \begin{pmatrix} 0 \\ \mathrm{i}(1 - \sqrt{2}) \\ 1/\sqrt{2} \end{pmatrix}.$$

Aufgabe 76: Untersuchen Sie, ob die folgenden Teilmengen der angegebenen $\mathbb{K}$-Vektorräume linear unabhängig sind:

(a) $\mathbb{K} = \mathbb{R}$, $V = \mathbb{R}^3$

$$M = \left\{ \begin{pmatrix} 1 \\ -2 \\ \sqrt{2} \end{pmatrix}, \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\},$$

(b) $\mathbb{K} = V = \mathbb{Z}/23\mathbb{Z}$, $M = \{[21], [17], [-5]\}$,

(c) $\mathbb{K} = \mathbb{C}$, $V = \mathrm{Abb}(\mathbb{C}, \mathbb{C})$, $M = \{f, g, h : \mathbb{C} \to \mathbb{C}\}$ mit $f : z \mapsto z^2$, $g : z \mapsto 1$, $h : z \mapsto z + 1$,

(d) $\mathbb{K} = \mathbb{R}$, $V = \mathbb{R}^5$

$$M = \left\{ \begin{pmatrix} 1 \\ \sqrt{3} \\ 26\sqrt{7} \\ -1 \\ 4 \end{pmatrix}, \begin{pmatrix} -2 \\ 0 \\ \sqrt[7]{19} \\ -\sqrt{10} \\ \pi \end{pmatrix}, \begin{pmatrix} e^3 \\ -4 \\ 0 \\ \sqrt{11} \\ 2 \end{pmatrix}, \begin{pmatrix} \sqrt{\pi} \\ 67 \\ \sqrt[3]{4} \\ 1 \\ 3 \end{pmatrix}, \begin{pmatrix} -6 \\ \sqrt[6]{2} \\ 6 \\ 3 \\ 1 \end{pmatrix}, \begin{pmatrix} \sin(2) \\ -1 \\ \sqrt{7} \\ \sqrt{5} \\ \cos(5) \end{pmatrix} \right\}.$$

Aufgabe 77: Sei V ein $\mathbb{K}$-Vektorraum, $v_1, ..., v_n \in V$ und $w_k = \Sigma_{j=1}^k v_j$ für $k \in \{1, ..., n\}$. Zeigen Sie, das das Tupel $(v_1, ..., v_n)$ linear unabhängig ist genau dann, wenn das Tupel $(w_1, ..., w_n)$ linear unabhängig ist.

Aufgabe 78: Sei $\mathbb{K}$ ein Körper. Wir betrachten für festes $n \in \mathbb{N}$ den Untervektorraum

$$U = \{f : \mathbb{N}_0 \to \mathbb{K} |\, f(k) = f(k + n)\ \forall k \in \mathbb{N}_0\} \subseteq \mathrm{Abb}(\mathbb{N}_0, \mathbb{K})$$

mit der punktweisen Addition und Skalarmultiplikation. Geben Sie eine Basis von U an und bestimmen Sie $\dim_{\mathbb{K}}(U)$.

Aufgabe 79: Untersuchen Sie, ob die folgenden Teilmengen $M \subseteq \mathbb{K}^3$ eine Basis des $\mathbb{K}$-Vektorraums $\mathbb{K}^3$ bilden. Wenn ja, drücken Sie die Vektoren in der Standardbasis als Linearkombinationen von Vektoren in M aus.

(a) $\mathbb{K} = \mathbb{Z}/7\mathbb{Z}$ und $M = \{u_1, u_2, u_3\}$ mit

$$u_1 = \begin{pmatrix} [3] \\ [2] \\ [6] \end{pmatrix} \qquad u_2 = \begin{pmatrix} [4] \\ [-2] \\ [1] \end{pmatrix}, \quad u_3 = \begin{pmatrix} [0] \\ [0] \\ [1] \end{pmatrix}.$$

(b) $\mathbb{K} = \mathbb{C}$ und $M = \{v_1, v_2, v_3\}$ mit

$$v_1 = \begin{pmatrix} 1 \\ \mathrm{i} \\ 0 \end{pmatrix} \qquad v_2 = \begin{pmatrix} 1 \\ \mathrm{i} \\ -\sqrt{2} \end{pmatrix}, \quad v_3 = \begin{pmatrix} e^{\mathrm{i}\pi/2} \\ e^{-\mathrm{i}\pi/2} \\ 0 \end{pmatrix}.$$

Aufgabe 80: Wir betrachten den Körper $\mathbb{F}_2 = \mathbb{Z}/2\mathbb{Z}$.

(a) Geben Sie alle Basen von $\mathbb{F}_2^2$ an.
(b) Bestimmen Sie die Anzahl der Basen B von $\mathbb{F}_2^3$ mit $\{e_1, e_2\} \subseteq B$.

Aufgabe 81: Wir betrachten die Untervektorräume

$$U_1 = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \middle| x + y + z = 0 \right\} \qquad U_2 = \mathrm{span}_{\mathbb{R}} \left\{ \begin{pmatrix} -2 \\ 3 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix}, \begin{pmatrix} 2 \\ 7 \\ 11 \end{pmatrix} \right\}$$

Geben Sie jeweils eine Basis von U_1, U_2 und $U_1 \cap U_2$ an.

Aufgabe 82: Seien V, W Vektorräume über $\mathbb{K}$ und $V \times W$ das Produkt der Vektorräume V, W. Sei $B \subseteq V$ eine Basis von V und $C \subseteq W$ eine Basis von W. Zeigen Sie, dass dann

$$M = \{(b, 0) | b \in B\} \cup \{(0, c) | c \in C\}$$

eine Basis von $V \times W$ ist. Geben Sie für endlich-dimensionale Vektorräume V, W die Dimension von $V \times W$ in Abhängigkeit von den Dimensionen von V und W an.

Aufgabe 83: Wir betrachten den reellen Vektorraum $\mathrm{Abb}(\mathbb{R}, \mathbb{R})$ und die Teilmengen

$$\begin{aligned} \mathrm{Abb}(\mathbb{R}, \mathbb{R})_+ &= \{f : \mathbb{R} \to \mathbb{R} | f(-x) = f(x) \,\forall x \in \mathbb{R}\} \\ \mathrm{Abb}(\mathbb{R}, \mathbb{R})_- &= \{f : \mathbb{R} \to \mathbb{R} | f(-x) = -f(x) \,\forall x \in \mathbb{R}\} \end{aligned}$$

Zeigen Sie, dass $\mathrm{Abb}(\mathbb{R}, \mathbb{R})_+$ und $\mathrm{Abb}(\mathbb{R}, \mathbb{R})_-$ Untervektorräume von $\mathrm{Abb}(\mathbb{R}, \mathbb{R})$ sind, und dass $\mathrm{Abb}(\mathbb{R}, \mathbb{R}) = \mathrm{Abb}(\mathbb{R}, \mathbb{R})_+ \oplus \mathrm{Abb}(\mathbb{R}, \mathbb{R})_-$ gilt.

Aufgabe 84: Wir betrachten im $\mathbb{R}^n$ die Untervektorräume $U_1 = \{(x, ..., x)^T \in \mathbb{R}^n | x \in \mathbb{R}\}$ und $U_2 = \{(x_1, ..., x_n)^T \in \mathbb{R}^n | \Sigma_{k=1}^n x_k = 0\}$. Berechnen Sie $\dim_{\mathbb{R}}(U_1 + U_2)$.

Aufgabe 85: Wir betrachten die Abbildungen

$$f : \mathbb{Z}/3\mathbb{Z} \to \mathbb{Z}/3\mathbb{Z}, x \mapsto x + [1], \quad g : \mathbb{Z}/3\mathbb{Z} \to \mathbb{Z}/3\mathbb{Z}, x \mapsto x^2, \quad h : \mathbb{Z}/3\mathbb{Z} \to \mathbb{Z}/3\mathbb{Z}, x \mapsto x^3 + [1].$$

Ist die Teilmenge $\{f, g, h\} \subseteq \mathrm{Abb}(\mathbb{Z}/3\mathbb{Z}, \mathbb{Z}/3\mathbb{Z})$ linear unabhängig?

Aufgabe 86: Sei V ein Vektorraum *über einem Körper* $\mathbb{L}$ und $\mathbb{K} \subseteq \mathbb{L}$ ein Teilkörper von $\mathbb{L}$. Sei $B \subseteq V$ eine Basis von V als Vektorraum *über* $\mathbb{L}$ und $C \subseteq \mathbb{L}$ eine Basis von $\mathbb{L}$ als Vektorraum *über* $\mathbb{K}$.

(a) Zeigen Sie, dass V dann auch die Struktur eines Vektorraums *über dem Körper* $\mathbb{K}$ hat.
(b) Konstruieren Sie aus den Basen B und C eine Basis des Vektorraums V als *Vektorraum über* $\mathbb{K}$.
(c) Bestimmen Sie die Dimension $\dim_{\mathbb{K}}(V)$ von V als Vektorraum *über* $\mathbb{K}$ in Abhängigkeit von $\dim_{\mathbb{L}}(V)$ und $\dim_{\mathbb{K}}(\mathbb{L})$.
(d) Folgern Sie aus (c) dass für Teilkörper $\mathbb{F} \subseteq \mathbb{K}$ und $\mathbb{K} \subseteq \mathbb{L}$ der Körper $\mathbb{L}$ ein Vektorraum über $\mathbb{F}$ der Dimension $\dim_{\mathbb{F}}(\mathbb{L}) = \dim_{\mathbb{F}}(\mathbb{K}) \cdot \dim_{\mathbb{K}}(\mathbb{L})$ ist, falls $\dim_{\mathbb{F}}(\mathbb{K})$ und $\dim_{\mathbb{K}}(\mathbb{L})$ endlich sind.

Hinweis: Die Aussage (d) ist der sogenannte **Gradsatz** der Körpertheorie. Ist $\mathbb{K} \subseteq \mathbb{L}$ ein Teilkörper, so nennt man $\mathbb{L}$ auch eine **Körpererweiterung** von $\mathbb{K}$, und $\dim_{\mathbb{K}}(\mathbb{L})$ den **Grad der Körpererweiterung**.

Aufgabe 87: Sei V ein 8-dimensionaler Vektorraum über einem Körper $\mathbb{K}$ und $U_1, U_2, U_3 \subseteq V$ Untervektorräume der Dimension $\dim_{\mathbb{K}}(U_1) = \dim_{\mathbb{K}}(U_2) = \dim_{\mathbb{K}}(U_3) = 3$ mit $U_1 + U_2 + U_3 = V$. Zeigen Sie, dass dann $U_1 \cap U_2 \cap U_3 = \{0\}$ gilt.

Aufgabe 88: Sei V ein $\mathbb{K}$-Vektorraum und $U, W \subseteq V$ Untervektorräume. Begründen Sie die folgenden Aussagen oder widerlegen Sie sie durch ein Gegenbeispiel.

(a) Ist $A \subseteq U$ ein Erzeugendensystem von U und $B \subseteq W$ ein Erzeugendensystem von W, so ist $A \cap B$ ein Erzeugendensystem von $U \cap W$.
(b) Ist $A \subseteq U$ ein Erzeugendensystem von U und $B \subseteq W$ ein Erzeugendensystem von W, so ist $A \cup B$ ein Erzeugendensystem von $U + W$.
(c) Ist $A \subseteq U$ eine Basis von U und $B \subseteq W$ eine Basis von W, so ist $A \cup B$ eine Basis von $U + W$.
(d) Sind $A \subseteq U$ und $B \subseteq W$ linear unabhängig und $U \cap W = \{0\}$, so ist auch $A \cup B$ linear unabhängig.

Aufgabe 89: Zeigen Sie, dass für $\mathbb{K} = \mathbb{Z}/5\mathbb{Z}$ der $\mathbb{K}$-Vektorraum $\mathbb{K}^{11}$ keinen Untervektorraum mit 50 Elementen besitzt.

13.4 Aufgaben zu Kapitel 4

Aufgabe 90: Untersuchen Sie, ob die folgenden Abbildungen zwischen den angegebenen $\mathbb{K}$-Vektorräumen $\mathbb{K}$-linear sind.

(a) $\mathbb{K} = \mathbb{R}$ $\quad \phi : \mathbb{R}^2 \to \mathbb{R}^2, \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} xy \\ 3x + 3y \end{pmatrix}$

(b) $\mathbb{K} = \mathbb{R}$ $\quad \phi : \mathbb{R}^2 \to \mathbb{R}^3, \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} y \\ 2x + 3y \\ -3x + 2 \end{pmatrix}$

(c) $\mathbb{K} = \mathbb{R}$ $\quad \phi : \mathbb{R}^2 \to \mathbb{R}^2$ mit $\begin{pmatrix} 1 \\ 2 \end{pmatrix} \mapsto \begin{pmatrix} 3 \\ 4 \end{pmatrix}, \begin{pmatrix} 7 \\ 8 \end{pmatrix} \mapsto \begin{pmatrix} 3 \\ 4 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \mapsto \begin{pmatrix} -1 \\ 0 \end{pmatrix}$

(d) $\mathbb{K} = \mathbb{C}$ $\quad \phi : \mathbb{C} \to \mathbb{C},\ x + \mathrm{i}y \mapsto (x + y) + \mathrm{i}x$

(e) $\mathbb{K} = \mathbb{C}$ $\quad \phi : \mathbb{C} \to \mathbb{C},\ x + \mathrm{i}y \mapsto (3x \cos\phi + 3y \sin\phi) + \mathrm{i}(3y \cos\phi - 3x \sin\phi)$ mit $\phi \in \mathbb{R}$

(f) $\mathbb{K} = \mathbb{Z}/3\mathbb{Z}$ $\quad \phi : \mathbb{Z}/3\mathbb{Z} \to \mathbb{Z}/3\mathbb{Z},\ x \mapsto x^3$.

Aufgabe 91: Wir betrachten für $\alpha \in \mathbb{R}$ die Abbildung

$$\phi_\alpha : \mathbb{R}^2 \to \mathbb{R}^2, \quad \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x \cos\alpha + y \sin\alpha \\ x \sin\alpha - y \cos\alpha \end{pmatrix}.$$

(a) Zeigen Sie, dass die Abbildung ϕ_α $\mathbb{R}$-linear ist.
(b) Bestimmen Sie die Menge $\{(x,y)^T \in \mathbb{R}^2 | \phi_\alpha((x,y)^T) = (x,y)^T\}$ und zeigen Sie, dass es sich dabei um einen Untervektorraum des $\mathbb{R}^2$ handelt.
(c) Interpretieren Sie die Abbildung ϕ_α geometrisch.
(d) Berechnen Sie für $\alpha, \beta \in \mathbb{R}$ die Abbildung $\phi_\alpha \circ \phi_\beta$ und interpretieren Sie sie geometrisch.

Aufgabe 92: Sei V ein Vektorraum über einem Körper $\mathbb{K}$. Eine lineare Abbildung $\phi : V \to V$ heißt *nilpotent*, wenn ein $n \in \mathbb{N}$ existiert mit $\phi^n = 0$, wobei $\phi^n = \phi \circ ... \circ \phi$ die n-fache Verkettung von ϕ mit sich selbst bezeichnet. Beweisen Sie, dass für jede nilpotente Abbildung $\phi : V \to V$ die Abbildungen $\mathrm{id}_V - \phi : V \to V$ und $\mathrm{id}_V + \phi : V \to V$ invertierbar sind.

Aufgabe 93: Sei $\mathbb{K}$ ein Körper, U, V, W, X Vektorräume über $\mathbb{K}$ und $\phi : U \to V$, $\chi : W \to X$ $\mathbb{K}$-lineare Abbildungen. Wir betrachten $F_{\chi,\phi} : \mathrm{Hom}_\mathbb{K}(V, W) \to \mathrm{Hom}_\mathbb{K}(U, X)$, $f \mapsto \chi \circ f \circ \phi$.

(a) Zeigen Sie, dass $F_{\chi,\phi}$ eine $\mathbb{K}$-lineare Abbildung ist.
(b) Zeigen Sie, dass $F_{\chi,\phi}$ injektiv ist, wenn $\phi : U \to V$ surjektiv und $\chi : W \to X$ injektiv ist.
(c) Bestimmen Sie $\ker(F_{\chi,\phi})$.

Aufgabe 94: Sei $(V, +, \cdot)$ ein Vektorraum über einem Körper $\mathbb{K}$, W eine Menge und $\phi : V \to W$ eine bijektive Abbildung. Beweisen Sie, dass es genau eine $\mathbb{K}$-Vektorraumstruktur auf W gibt, die $\phi : V \to W$ zu einem Isomorphismus von $\mathbb{K}$-Vektorräumen macht.

Aufgabe 95: Geben Sie eine $\mathbb{R}$-lineare Abbildung ϕ mit den folgenden Eigenschaften an, oder begründen Sie, warum eine solche Abbildung nicht existiert:

(a) $\phi : \mathbb{R}^3 \to \mathbb{R}^3$ mit $\phi(e_1) = e_2$ und $\phi(e_2) = e_3$,

(b) $\phi : \mathbb{R}^3 \to \mathbb{R}^3$ mit $e_1, e_2 \in \ker(\phi)$, $e_1 \in \mathrm{im}(\phi)$,
(c) $\phi : \mathbb{R}^4 \to \mathbb{R}^3$ surjektiv mit $\ker(\phi) = \mathrm{span}_{\mathbb{R}}\{e_1, e_3\}$,
(d) $\phi : \mathbb{R}^3 \to \mathbb{R}^3$ mit $\phi(e_1) = e_2$, $\phi(e_1 + e_2) = e_3$ und $\phi(e_2) = e_1$,
(e) $\phi : \mathbb{R}^3 \to \mathbb{R}^3$ mit $0 \neq \ker(\phi) \subseteq \mathrm{im}(\phi) \neq \mathbb{R}^3$,
(f) $\phi : \mathbb{R}^3 \to \mathbb{R}^3$ mit $\ker(\phi) = \mathbb{R}e_1$ und $\mathrm{im}(\phi) = \mathbb{R}e_2$.

Aufgabe 96: Seien V, W endlich-dimensionale Vektorräume über $\mathbb{K}$. Sind die folgenden Aussagen wahr oder falsch? Begründen Sie oder widerlegen Sie durch ein Gegenbeispiel.

(a) Existiert eine surjektive lineare Abbildung $\phi : V \to W$, so gilt $\dim_{\mathbb{K}}(V) \geq \dim_{\mathbb{K}}(W)$.
(b) Ist $\dim_{\mathbb{K}}(V) \geq \dim_{\mathbb{K}}(W)$, so gibt es eine surjektive lineare Abbildung $\phi : V \to W$.
(c) Ist $\dim_{\mathbb{K}}(V) > \dim_{\mathbb{K}}(W)$, so gibt es eine injektive lineare Abbildung $\phi : V \to W$.
(d) Existiert eine injektive lineare Abbildung $\phi : V \to W$, so gilt $\dim_{\mathbb{K}}(V) \leq \dim_{\mathbb{K}}(W)$.
(e) Existiert eine surjektive lineare Abbildung $\phi : V \to W$ und eine surjektive lineare Abbildung $\psi : W \to V$, so gilt $V \cong W$.
(f) Existiert eine injektive lineare Abbildung $\phi : V \to W$ und eine surjektive lineare Abbildung $\psi : W \to V$, so gilt $V \cong W$.
(g) Existiert eine surjektive lineare Abbildung $\phi : V \to W$ und eine injektive lineare Abbildung $\phi' : V \to W$, so gilt $V \cong W$.
(h) Gilt $\dim_{\mathbb{K}}(V) = \dim_{\mathbb{K}}(W)$ so gibt es eine injektive lineare Abbildung $\phi : V \to W$ und eine injektive lineare Abbildung $\psi : W \to V$.
(i) Ist $\phi : V \to W$ eine surjektive lineare Abbildung und $\psi : W \to V$ eine injektive lineare Abbildung, so ist $\psi \circ \phi$ ein $\mathbb{K}$-Vektorraumisomorphismus.
(j) Ist $\phi : V \to W$ eine surjektive lineare Abbildung und $\psi : W \to V$ eine injektive lineare Abbildung, so ist $\phi \circ \psi$ ein $\mathbb{K}$-Vektorraumisomorphismus.
(k) Ist $\dim_{\mathbb{K}}(V) = \dim_{\mathbb{K}}(W)$, so existiert genau eine bijektive lineare Abbildung $\phi : V \to W$.
(l) Ist $\dim_{\mathbb{K}}(V) = \dim_{\mathbb{K}}(W) > 0$, so existieren unendlich viele bijektive lineare Abbildungen $\phi : V \to W$.
(m) Sind $\phi : V \to W$ und $\psi : W \to V$ lineare Abbildungen, so dass $\psi \circ \phi : V \to V$ ein $\mathbb{K}$-Vektorraumisomorphismus ist, so ist ψ surjektiv und ϕ injektiv.
(n) Sind $\phi : V \to W$ und $\psi : W \to V$ lineare Abbildungen, so dass $\psi \circ \phi : V \to V$ ein $\mathbb{K}$-Vektorraumisomorphismus ist, so ist ϕ surjektiv und ψ injektiv.
(o) Ist $\phi : V \to W$ eine injektive lineare Abbildung und $\psi : W \to V$ eine surjektive lineare Abbildung, so ist $\psi \circ \phi : V \to V$ eine bijektive lineare Abbildung.
(p) Ist $\phi : V \to W$ eine injektive lineare Abbildung und $\psi : W \to V$ eine surjektive lineare Abbildung, so ist $\phi \circ \psi : W \to W$ eine bijektive lineare Abbildung.
(q) Ist $\phi : V \to W$ eine injektive lineare Abbildung, so gilt $\mathrm{rg}(\phi) = \dim_{\mathbb{K}}(V)$.
(r) Sind $\phi : V \to W$ und $\psi : W \to V$ lineare Abbildungen, so gilt $\mathrm{rg}(\psi \circ \phi) \leq \mathrm{rg}(\phi)$ und $\mathrm{rg}(\psi \circ \phi) \leq \mathrm{rg}(\psi)$.
(s) Ist $\phi : V \to W$ eine surjektive lineare Abbildung, so ist $\mathrm{im}(\phi) = \phi(V)$ ein Erzeugendensystem von W.
(t) Ist $\phi : V \to W$ eine injektive lineare Abbildung und B eine Basis von V, so ist $\phi(B)$ ein Erzeugendensystem von W.
(u) Ist $\phi : V \to W$ eine injektive lineare Abbildung und B eine Basis von V, so ist $\phi(B)$ linear unabhängig.
(v) Ist $\phi : V \to W$ eine lineare Abbildung, B eine Basis von V und $\phi(B)$ linear unabhängig, so ist ϕ injektiv.
(w) Für jedes $n \in \mathbb{N}$ gibt es unendlich viele bijektive $\mathbb{R}$-lineare Abbildungen $\phi : \mathbb{K}^n \to \mathbb{K}^n$.
(x) Für $m, n \in \mathbb{N}$ mit $n > m$ gibt es genau eine injektive $\mathbb{R}$-lineare Abbildung $\phi : \mathbb{K}^n \to \mathbb{K}^m$.

(y) Es gibt unendlich viele bijektive $\mathbb{Q}$-lineare Abbildungen $\phi : \mathbb{R} \to \mathbb{R}$.
(z) Es gibt genau eine bijektive $\mathbb{C}$-lineare Abbildung $\phi : \mathbb{C} \to \mathbb{C}$.

Aufgabe 97: Seien U, V, W endlich-dimensionale $\mathbb{K}$-Vektorräume und $\phi : U \to V$, $\psi : V \to W$ lineare Abbildungen. Beweisen Sie:

(a) $\mathrm{rg}(\psi \circ \phi) \leq \min(\mathrm{rg}(\phi), \mathrm{rg}(\psi))$.
(b) $\mathrm{rg}(\phi) + \mathrm{rg}(\psi) - \dim_{\mathbb{K}}(V) \leq \mathrm{rg}(\psi \circ \phi)$
(c) Finden Sie für beide Fälle lineare Abbildungen, für die in diesen Ungleichungen Gleichheit gilt, und lineare Abbildungen, für die strikte Ungleichheit gilt.

Aufgabe 98: Berechnen Sie die folgenden Produkte von Matrizen mit Einträgen in $\mathbb{K}$:

1. $\mathbb{K} = \mathbb{Q}$

$$\begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 & -1 \end{pmatrix}$$

2. $\mathbb{K} = \mathbb{Z}/5\mathbb{Z}$

$$\begin{pmatrix} [1] & [3] & [1] \\ [2] & [3] & [4] \end{pmatrix} \cdot \begin{pmatrix} [3] & [4] \\ [0] & [0] \\ [0] & [2] \end{pmatrix}$$

3. $\mathbb{K} = \mathbb{C}$

$$\begin{pmatrix} \mathrm{i} & 1 - 2\mathrm{i} \\ 1 + \mathrm{i} & 2e^{\mathrm{i}\pi/3} \end{pmatrix} \cdot \begin{pmatrix} \sqrt{2}e^{\mathrm{i}\pi/4} & 1 \\ e^{5\pi\mathrm{i}/6} & -1 \end{pmatrix}$$

Aufgabe 99: Das **Kroneckersche δ-Symbol** ist die Abbildung

$$\delta : \mathbb{N} \times \mathbb{N} \to \{0, 1\},\ (i, j) \mapsto \delta_{ij} \quad \text{mit } \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j. \end{cases}$$

Wir betrachten die **Elementarmatrizen** E_{lk}, die in der $l$ten Zeile und $k$ten Spalte den Eintrag 1 und ansonsten nur Nullen enthalten.

(a) Machen Sie sich klar, dass die Einträge der Elementarmatrix $E_{lk} \in \mathrm{Mat}(m \times n, \mathbb{K})$ gegeben sind durch

$$(E_{lk})_{ji} = \delta_{jl}\delta_{ik} \qquad \forall i, k \in \{1, ..., n\}, j, l \in \{1, ..., m\}$$

und zeigen Sie, dass für $E_{ji} \in \mathrm{Mat}(m \times n, \mathbb{K})$, $E_{lk} \in \mathrm{Mat}(n \times p, \mathbb{K})$ gilt:

$$E_{ji} \cdot E_{lk} = \delta_{il}\, E_{jk} \qquad \forall j \in \{1, ..., m\}, i, l \in \{1, ..., n\}, k \in \{1, ..., p\}.$$

(b) Berechnen Sie $A \cdot E_{sr}$ und $E_{cb} \cdot A$ für $A = (a_{ji}) \in \mathrm{Mat}(m \times n, \mathbb{K})$, $E_{cb} \in \mathrm{Mat}(l \times m, \mathbb{K})$ und $E_{sr} \in \mathrm{Mat}(n \times p, \mathbb{K})$, $r \in \{1, ..., p\}$, $s \in \{1, ..., n\}$, $b \in \{1, ..., m\}$, $c \in \{1, ..., l\}$. Interpretieren Sie Ihr Ergebnis.

(c) Berechnen Sie $(\mathbb{1}_m + \lambda E_{jk}) \cdot A$ und $(\mathbb{1}_m - E_{jj} - E_{kk} + E_{jk} + E_{kj}) \cdot A$ für $A = (a_{ji}) \in \mathrm{Mat}(m \times n, \mathbb{K})$, $E_{jk}, E_{kj}, E_{jj}, E_{kk} \in \mathrm{Mat}(m \times m, \mathbb{K})$ und $j, k \in \{1, ..., m\}$. Interpretieren Sie Ihr Ergebnis.

Aufgabe 100: Bestimmen Sie die darstellende Matrix der folgenden $\mathbb{R}$-linearen Abbildungen $\phi : V \to V$ bezüglich der angegebenen geordneten Basis B:

(a) $V = \mathbb{R}^2$, $B = (e_1 + e_2, e_1 - e_2)$, $\phi : \mathbb{R}^2 \to \mathbb{R}^2$ ist die Spiegelung an der x-Achse.

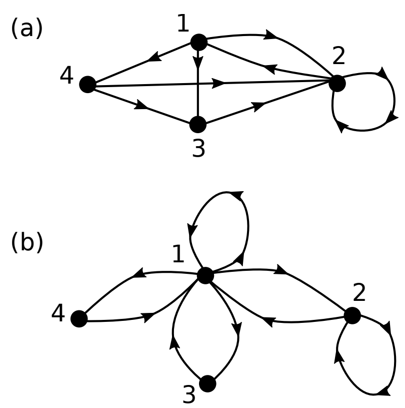
(b) $V = \mathrm{Pol}_{\leq 2}(\mathbb{R}) = \{\text{Polynomabbildungen } p : \mathbb{R} \to \mathbb{R}, |\deg(p) \leq 2\}$, $B = (1, x, x^2)$

$$\phi : \mathrm{Pol}_{\leq 2}(\mathbb{R}) \to \mathrm{Pol}_{\leq 2}(\mathbb{R}), p \mapsto \phi(p) \quad \text{mit } \phi(p)(x) := p(x-1),$$

Aufgabe 101: Sei V ein endlich-dimensionaler Vektorraum über $\mathbb{K}$ mit einer geordneten Basis $B = (v_1, ..., v_n)$. Wir betrachten die $\mathbb{K}$-linearen Abbildungen $\phi : V \to V$, deren beschreibende Matrix ${}_BM_B(\phi)$ bezüglich der Basis B in jeder Zeile und Spalte genau eine 1 und ansonsten nur 0 enthält.

(a) Charakterisieren Sie diese Abbildungen, indem Sie ihr Verhalten auf B betrachten.
(b) Zeigen Sie, dass diese linearen Abbildungen mit der Verkettung eine Gruppe bilden.
(c) Bestimmen Sie die Anzahl solcher linearen Abbildungen.

Aufgabe 102: Ein **gerichteter Graph** (V, E) mit k Vertizes besteht aus der endlichen Menge $V = \{1, 2, ..., k\}$ und einer Teilmenge $E \subseteq V \times V$. Elemente von V heißen **Vertizes**. Ein Element $(j, i) \in E$ heißt **Kante** vom Vertex $j \in V$ zum Vertex $i \in V$. Ein gerichteter Graph wird visualisiert, indem man für jeden Vertex $j \in V$ einen Punkt und für jede Kante $(j, i) \in E$ einen Pfeil vom Punkt j zum Punkt i zeichnet.



Die **Adjazenzmatrix** eines gerichteten Graphen (V, E) mit k Vertizes ist die Matrix

$$A = (a_{ji}) \in \mathrm{Mat}(k \times k, \mathbb{Q}) \qquad a_{ji} = \begin{cases} 1 & (j,i) \in E \\ 0 & (j,i) \notin E \end{cases}$$

Ein **Weg** der Länge $n \in \mathbb{N}$ vom Vertex $i_1 \in V$ zum Vertex $i_{n+1} \in V$ ist ein n-Tupel von Kanten

$$((i_1, i_2), (i_2, i_3), ..., (i_{n-1}, i_n), (i_n, i_{n+1})) \quad \text{mit} \quad i_m \in V, (i_m, i_{m+1}) \in E \; \forall m \in \{1, ..., n\}.$$

(a) Geben Sie die Adjazenzmatrix und alle Wege der Länge 3 von Vertex 1 zum Vertex 2 für die zwei abgebildeten Graphen an.

(b) Sei $L_n(j, i)$ die Anzahl der Wege der Länge n vom Vertex $j \in V$ zum Vertex $i \in V$. Zeigen Sie, dass für einen gerichteten Graphen (V, E) mit k Vertizes gilt

$$L_n(j, i) = \Sigma_{l=1}^k L_m(j, l) \cdot L_{n-m}(l, i) \qquad \forall m \in \{1, ..., n-1\}.$$

(c) Beweisen Sie unter Ausnutzung von (b), dass für jeden gerichteten Graphen (V, E) gilt

$$L_n(j,i) = (A^n)_{ji},$$

wobei $A^n := A \cdots A$ das n-fache Produkt der Adjazenzmatrix A mit sich selbst bezeichnet.

(d) Leiten Sie eine Formel her, die die Anzahl der Wege der Länge $\leq n$ vom Vertex j zum Vertex i in einem gerichteten Graphen (V, E) durch seine Adjazenzmatrix A ausdrückt.

(e) Ein **ungerichteter Graph** (V, E) ist ein gerichteter Graph (V, E) mit $(j, i) \in E \Leftrightarrow (i, j) \in E$ für alle $i, j \in V$. Sei jetzt (V, E) ein ungerichteter Graph, so dass zu jedem Vertex $j \in V$ mindestens eine Kante $(j, i) \in E$ existiert. Zeigen Sie, dass dann

$$j \sim i \quad \Leftrightarrow \quad \text{es existiert ein Weg (beliebiger Länge } n \in \mathbb{N}\text{) von } j \text{ nach } i$$

eine Äquivalenzrelation auf V ist.

Aufgabe 103: Wir betrachten den Vektorraum $V = \mathrm{Mat}(2 \times 2, \mathbb{R})$ und die Abbildung $\phi : V \to V$, $X \mapsto AXA$ mit

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

(a) Zeigen Sie, dass ϕ linear ist.
(b) Geben Sie eine Basis von $\ker(\phi)$ und $\mathrm{im}(\phi)$ an.
(c) Ergänzen Sie eine Basis von $\ker(\phi)$ zu einer Basis von V

Aufgabe 104: Für $x, y, z \in \mathbb{R}$ betrachten wir die Matrix

$$A(x,y,z) = \begin{pmatrix} 1 & x & y \\ 0 & 1 & z \\ 0 & 0 & 1 \end{pmatrix} \in \mathrm{Mat}(3 \times 3, \mathbb{R})$$

Zeigen Sie, dass $M = \{A(x,y,z) | x, y, z \in \mathbb{R}\}$ mit der Matrixmultiplikation eine Gruppe bildet.

Aufgabe 105: Berechnen Sie die Basiswechselmatrizen ${}_BM_A(\mathrm{id}_V)$ und ${}_BM_A(\mathrm{id}_V)$ für die angegebenen Basen der $\mathbb{K}$-Vektorräume V. Kontrollieren Sie Ihr Ergebnis, indem Sie ${}_BM_A(\mathrm{id}_V) \cdot {}_AM_B(\mathrm{id}_V)$ und ${}_AM_B(\mathrm{id}_V) \cdot {}_BM_A(\mathrm{id}_V)$ berechnen.

(a) $\mathbb{K} = \mathbb{C}$, $V = \mathbb{C}^3$, $A = (e_1, e_2, e_3)$, $B = (e_1 - \mathrm{i}e_3, e_1 - \mathrm{i}e_2, e_2 + e_3)$.

(b) $\mathbb{K} = \mathbb{R}$, $V = \mathrm{Pol}_{\leq 2}(\mathbb{R})$, $A = (f_1, f_2, f_3)$, $B = (g_1, g_2, g_3)$ mit

$$f_1 = g_1 : x \mapsto 1, \quad f_2 : x \mapsto x, \quad f_3 : x \mapsto x^2, \quad g_2 : x \mapsto x + 1, \quad g_3 : x \mapsto x^2 + x + 1.$$

Aufgabe 106: Wir betrachten den reellen Vektorraum $\mathbb{R}^3$ mit den geordneten Basen $A = (e_1, e_2, e_3)$, $B = (e_1 + e_2, e_1 - e_3, e_3)$ und die linearen Abbildungen

$$\phi : \mathbb{R}^3 \to \mathbb{R}^3, \quad \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} x + z \\ y - z \\ z - x \end{pmatrix}$$

$$\psi : \mathbb{R}^3 \to \mathbb{R}^3, \quad \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} 2x \\ 2x + 2y - z \\ z \end{pmatrix}.$$

(a) Berechnen Sie die darstellenden Matrizen ${}_B M_A(\phi)$, ${}_A M_B(\psi)$ und ${}_A M_A(\psi \circ \phi)$. Rechnen Sie nach, dass diese Matrizen die Bedingung ${}_A M_A(\psi \circ \phi) = {}_A M_B(\psi) \cdot {}_B M_A(\phi)$ erfüllen.

(b) Berechnen Sie die Basiswechselmatrizen ${}_B M_A(\mathrm{id}_{\mathbb{R}^3})$, ${}_A M_B(\mathrm{id}_{\mathbb{R}^3})$. Zeige Sie, dass diese Matrizen die Bedingung ${}_B M_A(\mathrm{id}_{\mathbb{R}^3}) \cdot {}_A M_B(\mathrm{id}_{\mathbb{R}^3}) = {}_A M_B(\mathrm{id}_{\mathbb{R}^3}) \cdot {}_B M_A(\mathrm{id}_{\mathbb{R}^3}) = \mathbb{1}_3$ erfüllen.

(c) Bestimmen Sie die darstellenden Matrizen ${}_B M_B(\phi)$, ${}_A M_A(\phi)$ und rechnen Sie nach, dass für Ihre Matrizen gilt

$$\begin{aligned}
{}_B M_B(\phi) &= {}_B M_A(\phi) \cdot {}_A M_B(\mathrm{id}_{\mathbb{R}^3}) \\
{}_A M_A(\phi) &= {}_A M_B(\mathrm{id}_{\mathbb{R}^3}) \cdot {}_B M_A(\phi) \\
{}_B M_B(\phi) &= {}_B M_A(\mathrm{id}_{\mathbb{R}^3}) \cdot {}_A M_A(\phi) \cdot {}_A M_B(\mathrm{id}_{\mathbb{R}^3})
\end{aligned}$$

Aufgabe 107: Untersuchen Sie, ob die reellen Matrizen A, B äquivalent oder ähnlich sind:

(a)

$$A = \begin{pmatrix} 1 & 2 & -1 \\ 2 & 3 & -1 \\ 2 & 4 & -2 \end{pmatrix} \qquad B = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

(b)

$$A = \begin{pmatrix} 1 & 2 & -1 \\ 0 & 3 & -1 \\ 1 & 0 & 1 \end{pmatrix} \qquad B = \begin{pmatrix} 1 & 0 & -1 \\ 2 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$$

(c)

$$A = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix} \qquad B = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

Aufgabe 108: Sei V ein endlich-dimensionaler Vektorraum über $\mathbb{K}$ und $\phi : V \to V$ eine $\mathbb{K}$-lineare Abbildung. Wir bezeichnen mit $\phi^k := \phi \circ \ldots \circ \phi$ für $k \in \mathbb{N}$ die k-fache Verkettung von ϕ mit sich selbst und setzen $\phi^0 := \mathrm{id}_V$. Zeigen Sie, dass ein $n \in \mathbb{N}$ und $\lambda_0, \lambda_1, ..., \lambda_{n-1} \in \mathbb{K}$ existieren mit

$$\phi^n + \lambda_{n-1}\phi^{n-1} + \ldots + \lambda_1 \phi + \lambda_0 \mathrm{id}_V = 0.$$

Zeigen Sie, dass diese Aussage für unendlich-dimensionale Vektorräume falsch ist, indem Sie ein Gegenbeispiel angeben.

Aufgabe 109: Bestimmen Sie den Rang der folgenden komplexen Matrizen:

$$A = \begin{pmatrix} 1 & 0 & -2 \\ 3 & \mathrm{i} & 1 \\ \mathrm{i}+1 & 2 & 0 \\ -1 & 2 & 2 \end{pmatrix} \qquad B = \begin{pmatrix} 0 & 2\mathrm{i} & -1 \\ \mathrm{i} & 2 & 3+\mathrm{i} \\ 0 & 0 & 1 \end{pmatrix} \qquad C = \begin{pmatrix} 1 & \mathrm{i} & 2 \\ -2\mathrm{i} & 2 & -4\mathrm{i} \\ 0 & 1 & \mathrm{i} \end{pmatrix}$$

Aufgabe 110: Bestimmen Sie jeweils eine Basis des Kerns und des Bildes der folgenden Matrizen $A \in \mathrm{Mat}(n \times n, \mathbb{K})$:

(a) $\mathbb{K} = \mathbb{Q}$

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

(b) $\mathbb{K} = \mathbb{C}$

$$A = \begin{pmatrix} \mathrm{i} & 2 & 0 \\ -1 & 1+\mathrm{i} & 0 \\ 1 & 1 & 0 \end{pmatrix},$$

(c) $\mathbb{K} = \mathbb{C}$

$$A = \begin{pmatrix} 1 & -1 & 1 & 2 \\ 2\mathrm{i} & -2\mathrm{i} & 2\mathrm{i} & 4\mathrm{i} \\ \mathrm{i} & 0 & -2\mathrm{i} & 0 \\ 1+3\mathrm{i} & -1 & 1-6\mathrm{i} & 2 \end{pmatrix},$$

(d) $\mathbb{K} = \mathbb{Z}/5\mathbb{Z}$

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

Aufgabe 111: Wir betrachten den $\mathbb{R}$-Vektorraum $\mathrm{Mat}(n \times n, \mathbb{R})$ und die linearen Abbildungen

$$\begin{aligned} \phi_+ &: \mathrm{Mat}(n \times n, \mathbb{R}) \to \mathrm{Mat}(n \times n, \mathbb{R}), \quad A \mapsto \tfrac{1}{2}(A + A^T) \\ \phi_- &: \mathrm{Mat}(n \times n, \mathbb{R}) \to \mathrm{Mat}(n \times n, \mathbb{R}), \quad A \mapsto \tfrac{1}{2}(A - A^T) \end{aligned}$$

(a) Zeigen Sie, dass ϕ_+ und ϕ_- Projektoren sind.
(b) Bestimmen Sie $\ker(\phi_\pm)$ und $\mathrm{im}(\phi_\pm)$.
(c) Geben Sie Basen von $\mathrm{im}(\phi_+)$ und $\ker(\phi_+)$ an, und bestimmen Sie $\mathrm{rg}(\phi_+)$ und $\mathrm{def}(\phi_+)$.

Aufgabe 112:
Untersuchen Sie die Matrizen $A, B \in \mathrm{Mat}(3 \times 3, \mathbb{R})$ auf Äquivalenz und Ähnlichkeit:

(a)

$$A = \begin{pmatrix} 1 & 2 & -1 \\ 2 & 3 & -1 \\ 2 & 4 & -2 \end{pmatrix} \qquad B = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

(b)

$$A = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix} \qquad B = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

Aufgabe 113: Untersuchen Sie die folgenden Matrizen $A, B \in \mathrm{Mat}(3 \times 3, \mathbb{R})$ auf Äquivalenz und Ähnlichkeit:

(a)

$$A = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \qquad B = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

(b)

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

(c)

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \qquad B = \begin{pmatrix} 2 & 1 & -1 \\ 1 & -1 & 1 \\ 1 & 2 & -1 \end{pmatrix}$$

(d)

$$A = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \qquad B = \begin{pmatrix} 2 & 1 & -1 \\ 1 & -1 & 1 \\ 1 & 2 & -1 \end{pmatrix}$$

Aufgabe 114: Die **Spur** einer Matrix $A = (a_{ji}) \in \mathrm{Mat}(n \times n, \mathbb{K})$ ist definiert als

$$\mathrm{tr}(A) = \Sigma_{i=1}^n a_{ii} \,.$$

Zeigen Sie:

(a) Die Spur definiert eine $\mathbb{K}$-lineare Abbildung $\mathrm{tr} : \mathrm{Mat}(n \times n, \mathbb{K}) \to \mathbb{K}$, $A \mapsto \mathrm{tr}(A)$.
(b) Für alle Matrizen $A, B \in \mathrm{Mat}(n \times n, \mathbb{K})$ gilt: $\mathrm{tr}(A \cdot B) = \mathrm{tr}(B \cdot A)$.
(c) Sind $A, B \in \mathrm{Mat}(n \times n, \mathbb{K})$ ähnlich, so gilt $\mathrm{tr}(A) = \mathrm{tr}(B)$.
(d) Es gibt Matrizen, deren Spuren gleich sind, die aber nicht ähnlich sind.

Aufgabe 115: Sei $n \in \mathbb{N}$. Für jede Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{K})$ definieren wir die Abbildung $\phi_A : \mathrm{Mat}(n\times n, \mathbb{K}) \to \mathbb{K}$, $M \mapsto \mathrm{tr}(A\cdot M)$. Zeigen Sie, dass die Abbildungen ϕ_A $\mathbb{K}$-linear sind und den ganzen Dualraum $\mathrm{Mat}(n \times n, \mathbb{K})^*$ aufspannen: $\mathrm{Mat}(n \times n, \mathbb{K})^* = \{\phi_A \,|\, A \in \mathrm{Mat}(n \times n, \mathbb{K})\}$.

Aufgabe 116: Sei V ein endlich-dimensionaler $\mathbb{K}$-Vektorraum.

(a) Bestimmen Sie alle injektiven Projektoren $P : V \to V$.
(b) Zeigen Sie, dass eine $\mathbb{K}$-lineare Abbildung $P : V \to V$ genau dann ein Projektor ist, wenn $\mathrm{id}_V - P : V \to V$ ein Projektor ist.
(c) Zeigen Sie: Sind $P, Q : V \to V$ Projektoren mit $\mathrm{im}(P) \subseteq \ker(Q)$ und $\mathrm{im}(Q) \subseteq \ker(P)$, so ist auch $P + Q : V \to V$ ein Projektor.
(d) Zeigen Sie durch ein Gegenbeispiel, dass im Allgemeinen die Summe $P+Q : V \to V$ zweier Projektoren $P, Q : V \to V$ kein Projektor sein muss.

Aufgabe 117: Sei V ein beliebiger $\mathbb{K}$-Vektorraum. Eine $\mathbb{K}$-lineare Abbildung $I : V \to V$ heißt **Involution**, wenn $I \circ I = \mathrm{id}_V$ gilt.

(a) Zeigen Sie, dass für jede Involution $I : V \to V$ die linearen Abbildungen $I_+ = \frac{1}{2}(\mathrm{id}_V + I)$ und $I_- = \frac{1}{2}(\mathrm{id}_V - I)$ Projektoren sind und $\ker(I_+) = \mathrm{im}(I_-)$, $\ker(I_-) = \mathrm{im}(I_+)$ gilt.
(b) Zeigen Sie, dass für jeden Projektor $P : V \to V$ die linearen Abbildungen $I_P = \mathrm{id}_V - 2P$ und $-I_P = 2P - \mathrm{id}_V$ Involutionen sind.
(c) Bestimmen Sie die Involutionen $I_{\pi_i} : \mathbb{K}^n \to \mathbb{K}^n$ für die Koordinatenprojektionen $\pi_i : \mathbb{K}^n \to \mathbb{K}^n$, $\Sigma_{j=1}^n x_j e_j \mapsto x_i e_i$ und interpretieren Sie sie für $\mathbb{K} = \mathbb{R}$ und $n = 2, 3$ geometrisch.
(d) Bestimmen Sie die Projektoren $I_\pm$ für die Involution $I : \mathbb{R}^2 \to \mathbb{R}^2$, $(x, y)^T \mapsto (y, x)^T$ und interpretieren Sie sie geometrisch.

Aufgabe 118: Wir betrachten für $n \in \mathbb{N}$ die Matrix

$$A_n = \begin{pmatrix} 1 & 2 & \dots & n \\ n+1 & n+2 & \dots & 2n \\ \vdots & \vdots & & \vdots \\ n^2-n+1 & n^2-n+2 & \dots & n^2 \end{pmatrix}$$

Bestimmen Sie den Rang von A_n in Abhängigkeit von $n \in \mathbb{N}$.

Aufgabe 119: Invertieren Sie die folgenden Matrizen mit dem Gauß-Verfahren:

$$A = \begin{pmatrix} 0 & 4 & -2 \\ 2 & -1 & 1 \\ 3 & 1 & -1 \end{pmatrix} \qquad B = \begin{pmatrix} \frac{1}{2} & -\frac{1}{3} & 0 \\ 1 & -1 & 2 \\ \frac{1}{3} & \frac{1}{9} & -1 \end{pmatrix} \qquad C = \begin{pmatrix} 0 & 1 & -1 \\ 3 & 0 & 1 \\ 0 & -3 & 1 \end{pmatrix}$$

Aufgabe 120: Bringen Sie die folgenden Matrizen in spezielle Zeilenstufenform und bestimmen Sie ihren Rang:

$$D = \begin{pmatrix} 0 & 2 & 1 \\ 1 & -1 & 2 \\ 1 & 3 & 4 \end{pmatrix} \qquad E = \begin{pmatrix} 0 & 1 & -2 & 3 \\ 1 & 0 & -1 & 1 \\ 2 & 1 & 0 & 0 \\ 2 & -1 & 1 & 1 \end{pmatrix} \qquad F = \begin{pmatrix} 0 & -1 & 1 & 2 \\ 1 & 2 & -1 & 1 \\ -1 & 3 & 5 & -1 \\ 0 & 2 & 7 & 6 \end{pmatrix}$$

Aufgabe 121: Bestimmen Sie die alle Lösungen der folgenden linearen Gleichungssysteme:

(a) $\mathbb{K} = \mathbb{R}$

$$\begin{aligned} x_1 - 2x_2 + x_3 - x_4 &= 1 \\ x_1 + x_2 + x_3 + x_4 &= 2 \\ x_1 - x_2 + x_3 + x_4 &= 5 \\ 2x_1 + x_2 + x_3 + x_4 &= 3 \end{aligned}$$

(b) $\mathbb{K} = \mathbb{Q}$

$$\begin{aligned} x_1 - 2x_2 + x_3 - x_4 &= 1 \\ x_1 + x_2 + x_3 + x_4 &= 2 \\ 3x_1 - 3x_2 + 3x_3 - x_4 &= 4 \\ x_1 + x_2 + x_3 + 2x_4 &= 3 \end{aligned}$$

(c) $\mathbb{K} = \mathbb{Z}/2\mathbb{Z}$

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

(d) $\mathbb{K} = \mathbb{Z}/5\mathbb{Z}$

$$\begin{aligned} x_1 + x_3 &= [0] \\ x_1 + x_2 + x_3 + x_4 &= [2] \\ x_2 + x_4 &= [0] \\ [3]x_1 - [4]x_2 + [2]x_3 &= [0] \end{aligned}$$

13.5 Aufgaben zu Kapitel 5

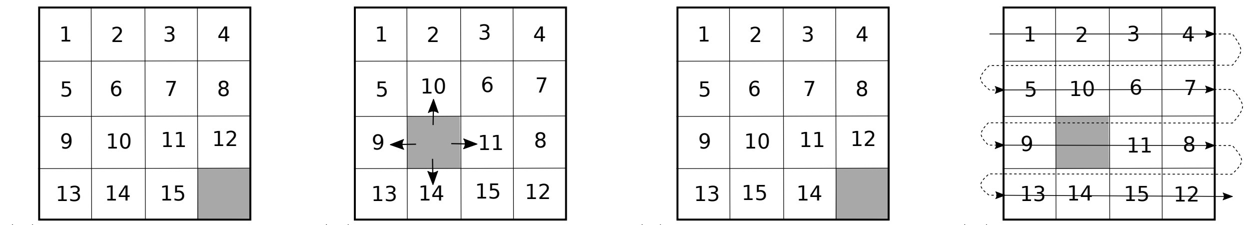
Aufgabe 122: Eine Permutation $\sigma \in S_n$ heißt **Zykel der Länge** $k \in \mathbb{N}$ oder k**-Zykel**, wenn es $i_1, ..., i_k \in \{1, ..., n\}$ gibt mit $\sigma(i_1) = i_2,..., \sigma(i_{k-1}) = i_k$, $\sigma(i_k) = i_1$ und $\sigma(j) = j$ für alle $j \in \{1, ..., n\} \setminus \{i_1, .., i_k\}$. Zeigen Sie, dass sich jeder k-Zykel als Produkt von $k-1$ Transpositionen schreiben lässt.

Aufgabe 123: Zeigen Sie, dass für $n \geq 2$ jede Transposition $\sigma \in S_n$ zu der elementaren Transposition $\sigma_{12} \in S_n$ konjugiert ist, d. h. dass eine Permutation $\pi \in S_n$ mit $\sigma = \pi \circ \sigma_{12} \circ \pi^{-1}$ existiert. Geben Sie eine solche Permutation $\pi \in S_n$ an.

Aufgabe 124: Zeigen Sie, dass die alternierende Gruppe $A_n = \ker(\mathrm{sgn}) \subseteq S_n$ abelsch ist genau dann, wenn $n \leq 3$ gilt.

Aufgabe 125: (Das 15er Spiel)

Auf einem quadratischen Tablett, das 16 quadratische Plättchen fasst, sind 15 mit den Zahlen von 1 bis 15 nummerierte quadratische Plättchen in der Startposition (a) angeordnet. Die Leerstelle (grau) darf horizontal und vertikal verschoben werden wie in (b), aber die einzelnen Plättchen dürfen nicht aus dem Tablett entfernt und wieder eingesetzt werden. Beweisen Sie, dass es *nicht* möglich ist, das Spiel durch Verschieben von Plättchen aus der Startposition (a) in die Zielposition (c) zu bringen.



(a) (b) (c) (d)

Hinweis: Bringen Sie das Problem in Zusammenhang mit Permutationen, indem sie die Nummern auf den Plättchen nach dem Schema in (d) anordnen.

Aufgabe 126: Wir betrachten den Vektorraum $V = (\mathbb{Z}/2\mathbb{Z})^2$ über dem Körper $\mathbb{Z}/2\mathbb{Z}$.

(a) Geben Sie alle symmetrischen und alle alternierenden Bilinearformen auf V an.

(b) Geben Sie alle Bilinearformen ω auf V an, die die Bedingung $\omega(v, w) = -\omega(w, v)$ für alle $v, w \in V$ erfüllen.

Aufgabe 127: Sei V ein endlich-dimensionaler reeller Vektorraum.

(a) Zeigen Sie, dass der Vektorraum $L(V^{\times 2}, \mathbb{R})$ der Bilinearformen auf V die direkte Summe der Unterräume der symmetrischen und der alternierenden Bilinearformen ist

$$L(V^{\times 2}, \mathbb{R}) = A(V^{\times 2}, \mathbb{R}) \oplus S(V^{\times 2}, \mathbb{R})$$

(b) Zeigen Sie, dass die analoge Aussage für n-Linearformen mit $n > 3$ falsch ist, indem Sie einen geeigneten reellen Vektorraum V wählen und eine n-Linearform ω auf V angeben, die nicht Summe einer alternierenden und einer symmetrischen n-Linearform ist.

Aufgabe 128: Sei $\mathbb{K}$ ein Körper der Charakteristik 0 und V, W endlich-dimensionale Vektorräume über $\mathbb{K}$. Für eine n-lineare Abbildung $\mu : V^{\times n} \to W$ definieren wir die n-linearen Abbildungen $\mu_s, \mu_a : V^{\times n} \to W$ durch

$$\mu_s(v_1, ..., v_n) = \tfrac{1}{n!}\Sigma_{\sigma\in S_n}\mu(v_{\sigma(1)}, ..., v_{\sigma(n)}) \qquad \mu_a(v_1, ..., v_n) = \tfrac{1}{n!}\Sigma_{\sigma\in S_n}\mathrm{sgn}(\sigma)\mu(v_{\sigma(1)}, ..., v_{\sigma(n)})$$

für alle $v_1, ..., v_n \in V$. Zeigen Sie, dass die Abbildungen

$$\mathrm{Sym} : L(V^{\times n}, W) \to L(V^{\times n}, W), \quad \mu \mapsto \mu_s \qquad \mathrm{Asym} : L(V^{\times n}, W) \to L(V^{\times n}, W), \quad \mu \mapsto \mu_a$$

Projektoren auf den Untervektorraum $S(V^{\times n}, W) \subseteq L(V^{\times n}, W)$ der symmetrischen und den Untervektorraum $A(V^{\times n}, W) \subseteq L(V^{\times n}, W)$ der alternierenden n-linearen Abbildungen von V nach W sind. Warum ist die Voraussetzung $\mathrm{char}(\mathbb{K}) \neq 0$ notwendig?

Aufgabe 129: Wir betrachten das sogenannte Vektorprodukt $\wedge : \mathbb{K}^3 \times \mathbb{K}^3 \to \mathbb{K}^3$

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \wedge \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} x_3y_3 - x_3y_2 \\ x_3y_1 - y_3x_1 \\ x_1y_2 - x_2y_1 \end{pmatrix}$$

und das euklidische Skalarprodukt $\langle\, ,\, \rangle : \mathbb{R}^2 \to \mathbb{R}^2$, $\langle x, y\rangle = \Sigma_{j=1}^3 x_j y_j$.

(a) Beweisen Sie die Graßmann Identität

$$x \wedge (y \wedge z) = \langle x, z\rangle y - \langle x, y\rangle z.$$

(b) Beweisen Sie die Lagrange Identität

$$\langle w \wedge x, y \wedge z\rangle = \langle w, y\rangle\langle x, z\rangle - \langle w, z\rangle\langle x, y\rangle.$$

(c) Folgern Sie aus (a), dass das Vektorprodukt nicht assoziativ ist.

Aufgabe 130: Berechnen Sie das Volumen des von den Vektoren $x, y, z \in \mathbb{R}^3$ aufgespannten Parallelepipeds:

(a)

$$x = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad y = \begin{pmatrix} 3 \\ -1 \\ 2 \end{pmatrix}, \quad z = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}$$

(b)

$$x = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad y = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \quad z = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

Aufgabe 131: Seien V, W endlich-dimensionale Vektorräume über einem Körper $\mathbb{K}$. Für eine n-Linearform $\omega : W^{\times n} \to \mathbb{K}$ auf W definiert man den **Pullback** oder die **Zurückziehung** von ω auf V mit einer linearen Abbildung $\phi : V \to W$ als

$$\phi^*\omega : V^{\times n} \to \mathbb{K}, \quad (v_1, ..., v_n) \mapsto \omega(\phi(v_1), ..., \phi(v_n)).$$

(a) Zeigen Sie, dass $\phi^*\omega$ eine n-Linearform auf V ist, und dass $\phi^*\omega$ alternierend (symmetrisch) ist, wenn ω alternierend (symmetrisch) ist.

(b) Beweisen Sie: $(\mu\phi)^*\omega = \mu^n(\phi^*\omega)$ für alle $\mu \in \mathbb{K}$ und $(\phi \circ \psi)^*\omega = \psi^*(\phi^*\omega)$ für alle linearen Abbildungen $\phi : V \to W$, $\psi : U \to V$.

Aufgabe 132: Sei $\mathbb{K}$ ein Körper und $n \in \mathbb{N}$. Wir betrachten die Abbildung

$$\omega : \mathrm{Mat}(n \times n, \mathbb{K})^{\times 3} \to \mathbb{K}, \quad (A, B, C) \mapsto \mathrm{tr}\,(A \cdot [B, C]).$$

wobei $\mathrm{Tr}(D) = \Sigma_{j=1}^n D_{jj}$ die Spur und $[B, C] = B \cdot C - C \cdot B$ den Matrixkommutator bezeichnet.

(a) Zeigen Sie, dass ω eine alternierende 3-Linearform auf $\mathrm{Mat}(n \times n, \mathbb{K})$ ist.

(b) Sei jetzt $n = 2$. Bestimmen Sie $\omega(v_i, v_j, v_k)$ für alle $i, j, k \in \{1, ..., 4\}$ und

$$v_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad v_4 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

(c) Sei $V = \mathrm{span}_{\mathbb{K}}(\{v_1, v_2, v_3\}) \subseteq \mathrm{Mat}(2 \times 2, \mathbb{K})$ und $T : V \to V$ die Einschränkung der Transpositionsabbildung auf V. Berechnen Sie die Determinante $\det(T)$ mit Hilfe der alternierenden 3-Linearform $\omega|_V : V^{\times 3} \to \mathbb{K}$.

Hinweis: Erinnern Sie sich an die Identität $\mathrm{tr}\,(A \cdot B) = \mathrm{tr}\,(B \cdot A)$, die in den Präsenzübungen in der Linearen Algebra 1 bewiesen wurde. Überlegen Sie sich in (b), für welche Tripel von Matrizen Sie die Werte berechnen müssen, damit die Rechnung nicht zu aufwändig wird.

Aufgabe 133: Sei V ein endlich-dimensionaler Vektorraum über einem Körper $\mathbb{K}$ und $\phi : V \to V$ eine lineare Abbildung. Beweisen Sie:

(a) Ist ϕ eine Involution, d. h. es gilt $\phi \circ \phi = \mathrm{id}_V$, so folgt $\det(\phi) \in \{1, -1\}$.

(b) Ist $\mathbb{K} = \mathbb{C}$ und $n \in \mathbb{N}$ mit $\phi^n = \phi \circ ... \circ \phi = \mathrm{id}_V$, so ist $\det(\phi) = e^{2\pi i k/n}$ für ein $k \in \{0, 1, ..., n-1\}$.

(c) Ist ϕ ein Projektor, so gilt $\det(\phi) \in \{0, 1\}$ und $\det(\phi) = 1$ genau dann, wenn $\phi = \mathrm{id}_V$.

(d) Ist ϕ nilpotent, d. h. es gibt ein $k \in \mathbb{N}$ mit $\phi^k = \phi \circ ... \circ \phi = 0$, so gilt $\det(\phi) = 0$.

Aufgabe 134: Sei $\mathbb{K}$ ein Körper, $n \in \mathbb{N}$ und V ein endlich-dimensionaler Vektorraum über $\mathbb{K}$. Begründen Sie die folgenden Aussagen oder widerlegen Sie sie durch ein Gegenbeispiel:

(a) Enthält $M \in \mathrm{Mat}(n \times n, \mathbb{R})$ nur Einträge aus $\{0, 1, -1\}$, so ist auch $\det(M) \in \{0, 1, -1\}$.

(b) Hat $M \in \mathrm{Mat}(n \times n, \mathbb{R})$ nur Einträge aus $\mathbb{Z}$, so ist auch $\det(M) \in \mathbb{Z}$.

(c) Enthält eine Matrix $M \in \mathrm{Mat}(n \times n, \mathbb{K})$ in jeder Zeile und Spalte genau einen Eintrag $\neq 0$, so gilt $\det(M) \neq 0$.

(d) Die Determinante definiert einen Gruppenhomomorphismus $\det : \mathrm{GL}(n, \mathbb{K}) \to (\mathbb{K} \setminus \{0\}, \cdot)$.

(e) Ist $M \in \mathrm{Mat}(n \times n, \mathbb{C})$ hermitesch, d. h. $\overline{M}^T = M$, so ist $\det(M) \in \mathbb{R}$.

Aufgabe 135: Sei $n \in \mathbb{N}$ und $\mathbb{K}$ ein Körper. Berechnen Sie die Determinanten der folgenden Endomorphismen $\phi : V \to V$.

(a) $V = \mathrm{Pol}_{\leq n}(\mathbb{R}) \subseteq \mathrm{Abb}(\mathbb{R}, \mathbb{R})$ ist der Vektorraum der reellen Polynomabbildungen vom Grad $\leq n$ und $\phi : V \to V$, $p \mapsto p'$ ist die Ableitung.

(b) $V = \operatorname{span}_{\mathbb{R}}(\{f_1, ..., f_n\}) \subseteq \operatorname{Abb}(\mathbb{R}, \mathbb{R})$ mit $f_k(x) = e^{kx}$ für alle $x \in \mathbb{R}$, $k \in \{1, ..., n\}$ und $\phi : V \to V$, $f \mapsto f'$ ist die Ableitung.

(c) $V = \operatorname{Mat}(n \times n, \mathbb{K})$ für ein $n \geq 2$, $\phi : V \to V$ ist die Abbildung, die die erste und zweite Zeile einer Matrix vertauscht.

(d) $V = \operatorname{Mat}(n \times n, \mathbb{K})$, $\phi : V \to V$, $M \mapsto M^T$ ist die Transpositionsabbildung.

(e) $V = \operatorname{Mat}(n \times n, \mathbb{K})$, $\phi : V \to V$, $(a_1, .., a_n) \mapsto (\lambda a_1, a_2, ..., a_n)$ ist die Abbildung, die die erste Spalte einer Matrix mit dem Skalar $\lambda \in \mathbb{K}$ multipliziert.

(f) $V = \operatorname{Mat}(n \times n, \mathbb{K})$, $\phi : V \to V$, $M \mapsto M - A \cdot M \cdot A^{-1}$ wobei $A \in \operatorname{GL}(n, \mathbb{K})$ eine fest gewählte Matrix ist.

Aufgabe 136: Berechnen Sie die Determinanten der folgenden Matrizen in $\operatorname{Mat}(4 \times 4, \mathbb{K})$

(a) $\mathbb{K} = \mathbb{Q}$

$$M = \begin{pmatrix} 2 & 5 & -1 & 0 \\ 2 & 2 & 1 & 1 \\ -1 & 1 & 1 & 0 \\ 3 & 1 & -1 & 0 \end{pmatrix}$$

(b) $\mathbb{K} = \mathbb{Q}$

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

(c) $\mathbb{K} = \mathbb{Z}/5\mathbb{Z}$

$$M = \begin{pmatrix} [2] & [3] & [1] & [3] \\ [4] & [0] & [2] & [4] \\ [1] & [2] & [4] & [1] \\ [3] & [4] & [2] & [2] \end{pmatrix}$$

(d) $\mathbb{K} = \mathbb{C}$

$$M = \begin{pmatrix} 1 & e^{\mathrm{i}\pi/2} & -\mathrm{i} & 1+\mathrm{i} \\ e^{2\pi\mathrm{i}/3} & 1+\mathrm{i} & -1 & e^{-\mathrm{i}\pi/2} \\ e^{2\pi\mathrm{i}/3} & -1 & 2\mathrm{i} & 1+2\mathrm{i} \\ 1 & 0 & -\mathrm{i} & 1+\mathrm{i} \end{pmatrix}$$

Aufgabe 137: Sei $\mathbb{K}$ ein Körper, $A \in \operatorname{Mat}(m \times m, \mathbb{K})$, $D \in \operatorname{GL}(n, \mathbb{K})$, $B \in \operatorname{Mat}(m \times n, \mathbb{K})$ und $C \in \operatorname{Mat}(n \times n, \mathbb{K})$. Zeigen Sie:

(a) $\det \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det(D)(\det A - BD^{-1}C)$

(b) $\det \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det(AD - BC)$ falls $n = m$ und $CD = DC$ gilt.

Aufgabe 138: Sei $\mathbb{K}$ ein Körper, $n \in \mathbb{N}$ und $X, Y \in \operatorname{Mat}(n \times n, \mathbb{K})$ beliebige Matrizen. Berechnen Sie die Determinante des Endomorphismus

$$\phi : \operatorname{Mat}(n \times n, \mathbb{K}) \to \operatorname{Mat}(n \times n, \mathbb{K}), \quad M \mapsto X \cdot M \cdot Y.$$

Hinweis: Drücken Sie die Bilder der Fundamentalmatrizen $E_{ij} \in \operatorname{Mat}(n \times n, \mathbb{K})$ mit $(E_{ij})_{kl} = \delta_{ik}\delta_{jl}$ für alle $i, j, k, l \in \{1, ..., n\}$ als Linearkombination der Fundamentalmatrizen aus.

Aufgabe 139: Sei $\mathbb{K}$ ein Körper und $x_1, ..., x_n \in \mathbb{K}$.

(a) Beweisen Sie:

$$\det \begin{pmatrix} 1 & x_1 & x_1^2 & \dots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \dots & x_2^{n-2} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & x_n^2 & \dots & x_n^{n-2} \end{pmatrix} = \prod_{1 \le i < j \le n} (x_j - x_i)$$

(b) Zeigen Sie mit (a), dass für jeden unendlichen Körper $\mathbb{K}$ eine unendliche Teilmenge $M \subseteq \mathbb{K}^n$ existiert, so dass alle n-Tupel von Vektoren aus M linear unabhängig sind.

Aufgabe 140: Sei $\mathbb{K}$ ein Körper, $n \ge 4$ und $x_1, ..., x_n \in \mathbb{K}$. Berechnen Sie

$$\det \begin{pmatrix} x_1^2 & (x_1+1)^2 & \dots & (x_1+n-1)^2 \\ \vdots & \vdots & & \vdots \\ x_n^2 & (x_n+1)^2 & \dots & (x_n+n-1)^2 \end{pmatrix}$$

Aufgabe 141: Sei $\mathbb{K}$ ein Körper, $p \in \mathbb{K}[x]$ ein Polynom vom Grad $\le n-2$ und $x_1, ..., x_n$, $y_1, .., y_n \in \mathbb{K}$. Berechnen Sie

$$\det \begin{pmatrix} p(x_1+y_1) & \dots & p(x_1+y_n) \\ \vdots & & \vdots \\ p(x_n+y_1) & \dots & p(x_n+y_n) \end{pmatrix}$$

Aufgabe 142: Die **Fibonacci-Zahlen** sind die Glieder der Fibonacci-Folge $(f_n)_{n\in\mathbb{N}_0}$ in $\mathbb{N}_0$, die rekursiv durch $f_0 = f_1 = 1$ und $f_{n+1} = f_n + f_{n-1}$ definiert ist. Zeigen Sie, dass für $n \ge 2$ das $n$te Folgenglied der Fibonacci-Folge gegeben ist durch $f_n = \det(B_n)$, wobei $B_n \in \mathrm{Mat}(n \times n, \mathbb{Q})$ die Matrix mit Einträgen

$$B_n = \begin{pmatrix} 1 & 1 & 0 & \dots & 0 \\ -1 & 1 & 1 & \ddots & \vdots \\ 0 & -1 & 1 & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & 1 \\ 0 & \dots & 0 & -1 & 1 \end{pmatrix} \qquad (B_n)_{ji} = \delta_{ij} + \delta_{i(j+1)} - \delta_{j(i+1)} = \begin{cases} 1 & i = j \lor i = j+1 \\ -1 & j = i+1 \\ 0 & \text{sonst} \end{cases}$$

Aufgabe 143: Seien $a, b \in \mathbb{R}$ und

$$A = \begin{pmatrix} 3a & 1 & 1 \\ -2 & 0 & b-1 \\ 1 & 1 & 1 \end{pmatrix}.$$

(a) Untersuchen Sie, für welche $a, b \in \mathbb{R}$ die Matrix A invertierbar ist.

(b) Bestimmen Sie die adjungierte Matrix $\mathrm{ad}(A)$ und zeigen Sie durch Nachrechnen, dass

$$\mathrm{ad}(A)^T \cdot A = \det(A)\mathbb{1}_3.$$

(c) Seien nun $a, b \in \mathbb{R}$ so, dass A invertierbar ist. Bestimmen Sie mit der Cramerschen Regel die eindeutige Lösung des inhomogenen linearen Gleichungssystems

$$Ax = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

13.6 Aufgaben zu Kapitel 6

Aufgabe 144: Wir betrachten die Matrizen

$$A_1 = \begin{pmatrix} 2 & -1 \\ 3 & -2 \end{pmatrix} \qquad A_2 = \begin{pmatrix} -1 & 1 \\ -2 & 1 \end{pmatrix}$$
$$A_3 = \begin{pmatrix} 3 & 2 & -2 \\ 2 & 3 & -2 \\ 6 & 6 & -5 \end{pmatrix} \qquad A_4 = \begin{pmatrix} 4 & -1 & 3 \\ -2 & 1 & -1 \\ -2 & 3 & -3 \end{pmatrix}.$$

(a) Bestimmen Sie die Eigenwerte der Matrizen A_i über $\mathbb{R}$ und über $\mathbb{C}$.
(b) Geben Sie für jeden Eigenwert λ von A_i jeweils eine Basis des Eigenraums $E(A_i, \lambda)$ an.
(c) Entscheiden Sie, ob die Matrizen A_i über $\mathbb{R}$ und über $\mathbb{C}$ diagonalisierbar sind.

Aufgabe 145: (Staatsexamen 2015)
Wir betrachten den Körper $\mathbb{K} = \mathbb{Z}/2\mathbb{Z}$ und den Endomorphismus

$$\phi : \mathbb{K}^3 \to \mathbb{K}^3, \qquad \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} z \\ y \\ x \end{pmatrix}$$

Bestimmen Sie alle Eigenwerte von ϕ und geben Sie jeweils eine Basis der zugehörigen Eigenräume $E(\phi, \lambda)$ an. Untersuchen Sie, ob ϕ trigonalisierbar ist.

Aufgabe 146: Bestimmen Sie für die folgenden Matrizen $A \in \mathrm{Mat}(n \times n, \mathbb{K})$ eine Diagonalmatrix $D \in \mathrm{Mat}(n \times n, \mathbb{K})$ und eine Matrix $S \in \mathrm{GL}(n, \mathbb{K})$ mit $A = S \cdot D \cdot S^{-1}$:

(a) $\mathbb{K} = \mathbb{R}$

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

(b) $\mathbb{K} = \mathbb{Z}/5\mathbb{Z}$

$$A = \begin{pmatrix} [1] & [3] & [0] \\ [0] & [0] & [0] \\ [1] & [2] & [2] \end{pmatrix},$$

(c) $\mathbb{K} = \mathbb{C}$

$$A = \begin{pmatrix} 1-\mathrm{i} & \mathrm{i} \\ -5\mathrm{i} & 1+\mathrm{i} \end{pmatrix},$$

(d) $\mathbb{K} = \mathbb{R}$

$$A = \begin{pmatrix} -5 & -7 & 4 \\ -7 & -5 & -4 \\ 4 & -4 & -2 \end{pmatrix}.$$

Aufgabe 147: Bestimmen Sie zu jedem Eigenwert von A die Haupträume der folgenden Matrizen $A \in \mathrm{Mat}(n \times n, \mathbb{K})$

(a) $\mathbb{K} = \mathbb{R}$

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

(b) $\mathbb{K} = \mathbb{Z}/3\mathbb{Z}$

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

(c) $\mathbb{K} = \mathbb{R}$

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

(d) $\mathbb{K} = \mathbb{C}$

$$A = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & -9 & 4 \end{pmatrix}.$$

Aufgabe 148: In dieser Aufgabe verallgemeinern wir das Verfahren, das in Beispiel 6.1.3 benutzt wurde, um die Eigenwerte von Projektoren, Involutionen und nilpotenten Abbildungen zu bestimmen.

Sei dazu V ein endlich-dimensionaler $\mathbb{K}$-Vektorraum und $\phi : V \to V$ ein Endomorphismus. Für ein Polynom $p = a_n x^n + a_{n-1}x^{n-1} + ... + a_1 x + a_0 \in \mathbb{K}[x]$ betrachten wir den Endomorphismus

$$p(\phi) := a_n \phi^n + a_{n-1}\phi^{n-1} + ... + a_1\phi + a_0\phi^0 : V \to V \quad \text{mit} \quad \phi^k := \underbrace{\phi \circ ... \circ \phi}_{k\times},\ \phi^0 := \mathrm{id}_V$$

(a) Zeigen Sie: Ist $p \in \mathbb{K}[x]$ ein Polynom mit $p(\phi) = 0$, so ist jeder Eigenwert λ von ϕ eine Nullstelle von p, d. h. $p(\lambda) = a_n\lambda^n + a_{n-1}\lambda^{n-1} + ... + a_1\lambda + a_0 = 0$.

(b) Geben Sie ein Polynom $p \in \mathbb{K}[x]$ mit $p \neq 0$ und $p(\phi) = 0$ an, für die Fälle, dass ϕ (i) eine Involution, (ii) ein Projektor, (iii) eine Streckung mit $\alpha \in \mathbb{K}$ oder (iv) nilpotent ist.

(c) Geben Sie für $\mathbb{K} = \mathbb{R}$ ein Polynom $p \in \mathbb{R}[x]$, $p \neq 0$ und einen Endomorphismus $\phi : V \to V$ mit $p(\phi) = 0$ an, so dass ϕ keine reellen Eigenwerte besitzt.

(d) Zeigen Sie: Ist $\phi : V \to V$ diagonalisierbar und $p \in \mathbb{K}[x]$ ein Polynom, so dass alle Eigenwerte von ϕ Nullstellen von p sind, so gilt $p(\phi) = 0$.

(e) Zeigen Sie: Zu jedem Endomorphismus $\phi : V \to V$ gibt es ein Polynom $p \in \mathbb{K}[x]$ mit $p \neq 0$ und $p(\phi) = 0$.

Aufgabe 149: (**Permutationsmatrizen**)
Wir betrachten für $n \in \mathbb{N}$ die Abbildung

$$\phi : S_n \to \mathrm{GL}(n, \mathbb{C}),\ \ \sigma \mapsto M_\sigma \quad \text{mit} \quad (M_\sigma)_{ji} = \begin{cases} 1 & j = \sigma(i) \\ 0 & j \neq \sigma(i) \end{cases}$$

.

(a) Zeigen Sie, dass $\phi : S_n \to \mathrm{GL}(n, \mathbb{C})$ ein injektiver Gruppenhomomorphismus ist. Folgern Sie, dass die Matrizen in $\mathrm{GL}(n, \mathbb{C})$, die in jeder Zeile und Spalte genau eine Eins und ansonsten nur Nullen enthalten eine zu S_n isomorphe Untergruppe $G \subseteq \mathrm{GL}(n, \mathbb{C})$ bilden.

(b) Zeigen Sie, dass $\det(M_\sigma) = \mathrm{sgn}(\sigma)$ für alle $\sigma \in S_n$ gilt.

(c) Zeigen Sie, dass für jede Permutation $\sigma \in S_n$ jeder Eigenwert von M_σ eine $k$te Einheitswurzel mit $k \leq n$ ist.

(d) Zeigen Sie, dass für jede Permutation $\sigma \in S_n$ die Matrix M_σ den Eigenwert 1 besitzt und geben Sie einen Eigenvektor zum Eigenwert 1 an.

Aufgabe 150: Wir betrachten die Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{K})$ die über der Diagonalen eine (diagonale) Reihe von Einsen und ansonsten nur Nullen enthält, also die Matrix

$$A = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 0 & 1 \\ 0 & \dots & 0 & 0 & 0 \end{pmatrix} \in \mathrm{Mat}(n \times n, \mathbb{K}) \qquad A_{ji} = \delta_{j+1,i} = \begin{cases} 1 & i = j+1 \\ 0 & i \neq j+1. \end{cases}$$

(a) Zeigen Sie, dass A nilpotent ist, indem Sie die Potenzen A^k für $k \in \mathbb{N}$ explizit berechnen, und bestimmen Sie das kleinste $k \in \mathbb{N}$ mit $A^k = 0$.

(b) Bestimmen Sie alle Eigenwerte und Eigenvektoren von A^k für $k \in \mathbb{N}$.

Aufgabe 151: Sei $\mathbb{K}$ ein Körper, $n \in \mathbb{N}$ und V ein endlich-dimensionaler Vektorraum über $\mathbb{K}$. Zeigen Sie:

(a) Sind $A, B \in \mathrm{Mat}(n \times n, \mathbb{K})$ ähnlich und ist A nilpotent, so ist auch B nilpotent.

(b) Ist ein Endomorphismus $\phi : V \to V$ nilpotent, so ist auch die beschreibende Matrix ${}_BM_B(\phi)$ bezüglich jeder geordneten Basis B von V nilpotent.

(c) Existiert eine geordnete Basis B von V, so dass ${}_BM_B(\phi)$ nilpotent ist, so ist auch der Endomorphismus $\phi : V \to V$ nilpotent.

(d) Ist ein Endomorphismus $\phi : V \to V$ bzw. eine Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{K})$ nilpotent, so ist ϕ bzw. A diagonalisierbar genau dann, wenn $\phi = 0$ bzw. $A = 0$ gilt.

Aufgabe 152: Wir betrachten den Untervektorraum $V = \mathrm{span}_\mathbb{R}\{f_1, ..., f_5\} \subseteq \mathrm{Abb}(\mathbb{R}, \mathbb{R})$

$$f_1 : x \mapsto e^x, \qquad f_2 : x \mapsto xe^x, \qquad f_3 : x \mapsto x^2 e^x, \qquad f_4 : x \mapsto e^{2x}, \qquad f_5 : x \mapsto xe^{2x}$$

Überlegen Sie sich, ohne das aufzuschreiben, dass die Ableitung einen Endomorphismus $d : V \to V$, $f \mapsto f'$ definiert.

(a) Berechnen Sie die beschreibende Matrix ${}_BM_B(d)$ für die geordnete Basis $B = (f_1, ..., f_5)$.

(b) Bestimmen Sie die Eigenwerte von d.

(c) Untersuchen Sie d auf Diagonalisierbarkeit.

(d) Bestimmen Sie für jeden Eigenwert λ von d den Eigenraum $E(d, \lambda)$ und den Hauptraum $H(d, \lambda)$.

Hinweis: Lösen Sie (b), (c), (d) mit Hilfe der beschreibenden Matrix ${}_BM_B(d)$.

Aufgabe 153: Wir betrachten die Matrix

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & -2 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 2 & 2 & 2 & -1 \end{pmatrix} \in \mathrm{Mat}(4 \times 4, \mathbb{Q}).$$

(a) Bestimmen Sie die Eigenwerte von A.
(b) Untersuchen Sie A auf Diagonalisierbarkeit über $\mathbb{Q}$.
(c) Zeigen Sie, dass A trigonalisierbar ist, und geben Sie eine obere Dreiecksmatrix B an, die ähnlich zu A ist.

Aufgabe 154: Sei V mit $\langle\, ,\, \rangle$ ein endlich-dimensionaler euklidischer oder unitärer Vektorraum und $P : V \to V$ ein Projektor. Zeigen sie, dass P selbstadjungiert ist genau dann, wenn P eine *Orthogonalprojektion* ist, d. h. $\mathrm{im}(P) = \ker(P)^{\perp}$ gilt.

Aufgabe 155:
• Zwei Endomorphismen $\phi, \psi \in \mathrm{End}_{\mathbb{K}}(V)$ eines endlich-dimensionalen $\mathbb{K}$-Vektorraums V heißen **simultan diagonalisierbar**, wenn es eine geordnete Basis B von V gibt, so dass ${}_BM_B(\phi)$ *und* ${}_BM_B(\psi)$ diagonal sind.

• Zwei Matrizen $M, N \in \mathrm{Mat}(n \times n, \mathbb{K})$ heißen **simultan diagonalisierbar**, wenn es eine Matrix $S \in \mathrm{GL}(n, \mathbb{K})$ gibt, so dass $S^{-1}MS$ *und* $S^{-1}NS$ diagonal sind.

Zeigen Sie, dass zwei diagonalisierbare Endomorphismen $\phi, \psi \in \mathrm{End}_{\mathbb{K}}(V)$ (zwei diagonalisierbare Matrizen $M, N \in \mathrm{Mat}(n \times n, \mathbb{K})$) genau dann simultan diagonalisierbar sind, wenn sie kommutieren, d. h. wenn $\phi \circ \psi = \psi \circ \phi$ $(M \cdot N = N \cdot M)$ gilt.

Aufgabe 156: Sei V ein n-dimensionaler Vektorraum über $\mathbb{K}$ und $\phi : V \to V$ ein Endomorphismus mit Fitting-Index f. Zeigen Sie:

(a) Ist ϕ nilpotent, so ist $f = \min\{k \in \mathbb{N}_0 |\, \phi^k = 0\}$.
(b) Es gilt $f = 0$ genau dann, wenn ϕ ein Isomorphismus ist.
(c) Es gilt $f = 1$ genau dann, wenn 0 ein Eigenwert von ϕ ist und $E(\phi, 0) = H(\phi, 0)$.
(d) Es gilt $f = n$ genau dann, wenn ϕ nilpotent mit $\phi^k \neq 0$ für $k < n$ ist.

Aufgabe 157: Sei V ein endlich-dimensionaler $\mathbb{K}$-Vektorraum und $\phi \in \mathrm{End}_{\mathbb{K}}(V)$ ein Endomorphismus. Wir betrachten den Endomorphismus

$$F_\phi : \mathrm{End}_{\mathbb{K}}(V) \to \mathrm{End}_{\mathbb{K}}(V), \quad \psi \mapsto \phi \circ \psi.$$

(a) Zeigen Sie, dass jeder Eigenwert von F_ϕ auch ein Eigenwert von ϕ ist.
(b) Zeigen Sie, dass jeder Eigenwert von ϕ auch ein Eigenwert von F_ϕ ist, und bestimmen Sie den zugehörigen Eigenraum $E(F_\phi, \lambda)$.
(c) Gilt für die geometrischen Vielfachheiten $g(F_\phi, \lambda) = g(\phi, \lambda)$ für jeden Eigenwert λ von ϕ?
(d) Zeigen Sie, dass F_ϕ diagonalisierbar ist genau dann, wenn ϕ diagonalisierbar ist.

Aufgabe 158: Wir betrachten den Endomorphismus

$$\phi : \mathbb{R}^3 \to \mathbb{R}^3, \qquad \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} 3 & -1 & 2 \\ 0 & 2 & 0 \\ -1 & -1 & 0 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 3x - y + 2z \\ 2y \\ -x - y \end{pmatrix}$$

(a) Bestimmen Sie die Haupträume von ϕ.
(b) Bestimmen sie die Jordan-Zerlegung von ϕ.

Aufgabe 159: Sei V ein endlich-dimensionaler $\mathbb{K}$-Vektorraum und $\phi, \psi : V \to V$ Endomorphismen. Zeigen Sie:

(a) Ist $\phi \circ \psi = \psi \circ \phi$, so ist für alle Eigenwerte λ von ϕ der Eigenraum $E(\phi, \lambda)$ ψ-invariant: $\psi(E(\phi, \lambda)) \subseteq E(\phi, \lambda)$.

(b) Sind ϕ, ψ diagonalisierbar und $\psi \circ \phi = \phi \circ \psi$, so gibt es einen Basis von V aus gemeinsamen Eigenvektoren von ϕ und ψ.

(c) Zeigen Sie: sind ϕ, ψ diagonalisierbar und es gibt es eine Basis aus gemeinsamen Eigenvektoren von V zu ϕ und ψ, so gilt $\phi \circ \psi = \psi \circ \phi$.

Aufgabe 160: Sei $\mathbb{K}$ ein Körper und $n \in \mathbb{N}$. **Wahr oder falsch?** Begründen Sie die Aussage oder widerlegen Sie sie mit einem Gegenbeispiel.

(a) $A, B \in \mathrm{Mat}(n \times n, \mathbb{K})$ nilpotent $\Rightarrow A + B$ nilpotent.
(b) $A, B \in \mathrm{Mat}(n \times n, \mathbb{K})$ nilpotent und $A \cdot B = B \cdot A \Rightarrow A + B$ nilpotent.
(c) $A, B \in \mathrm{Mat}(n \times n, \mathbb{K})$ nilpotent $\Rightarrow A \cdot B$ nilpotent.
(d) $A, B \in \mathrm{Mat}(n \times n, \mathbb{K})$ nilpotent und $A \cdot B = B \cdot A \Rightarrow A \cdot B$ nilpotent.
(e) $A \in \mathrm{Mat}(n \times n, \mathbb{K})$ nilpotent $\Rightarrow A^n = 0$.
(f) $A \in \mathrm{Mat}(n \times n, \mathbb{K})$ nilpotent $\Rightarrow \mathrm{Tr}(A) = 0$.
(g) $A \in \mathrm{Mat}(n \times n, \mathbb{K})$ nilpotent $\Rightarrow \det(\mathbb{1}_n + A) = \det(\mathbb{1}_n - A) = 1$.
(h) $A, B \in \mathrm{Mat}(n \times n, \mathbb{K})$ mit A nilpotent und $A \cdot B = B \cdot A \Rightarrow \det(A + B) = \det(B)$.

Aufgabe 161: Sei $n \in \mathbb{N}$ und $A \in \mathrm{Mat}(n \times n, \mathbb{C})$. Beweisen Sie, dass

$$\det(\exp(A)) = \exp(\mathrm{Tr}(A))$$

Aufgabe 162: Die **Fibonacci-Zahlen** sind die Glieder der **Fibonacci-Folge** $(f_n)_{n \in \mathbb{N}_0}$, die rekursiv durch $f_0 = f_1 = 1$ und $f_{n+1} = f_n + f_{n-1}$ definiert ist.

(a) Beschreiben Sie die Rekursionsrelation mit einer Matrix $A \in \mathrm{Mat}(2 \times 2, \mathbb{R})$, zeigen Sie, dass A über $\mathbb{R}$ diagonalisierbar ist und bestimmen Sie eine Diagonalmatrix $D \in \mathrm{Mat}(2 \times 2, \mathbb{R})$ und eine Matrix $S \in \mathrm{GL}(2, \mathbb{R})$, so dass $A = S \cdot D \cdot S^{-1}$.

(b) Leiten Sie eine nicht-rekursive Formel für die $k$te Fibonacci-Zahl her, indem Sie die Potenzen A^k für $k \in \mathbb{N}$ berechnen.

(c) Zeigen Sie, dass die Folge $(q_n)_{n \in \mathbb{N}_0}$ mit $q_n = f_{n+1}/f_n$ einen Grenzwert besitzt, und dass dieser ein Eigenwert von A ist. Geben Sie eine Folge $(a_n)_{n \in \mathbb{N}_0}$ an, die sich durch die Fibonacci-Zahlen ausdrücken lässt und deren Grenzwert der andere Eigenwert von A ist.

Aufgabe 163: Wir betrachten die reelle Folge $(f_n)_{n\in\mathbb{N}_0}$, die durch vorgegebene Startwerte $f_0, ..., f_k \in \mathbb{R}$ und die Rekursionsbedingung $f_{n+1} = a_0 f_n + ... + a_k f_{n-k}$ für $n \geq k$ gegeben ist. Diese kann als ein inhomogenes lineares Gleichungssystem formuliert werden

$$\begin{pmatrix} f_{n+1} \\ f_n \\ \vdots \\ f_{n-k+1} \end{pmatrix} = \begin{pmatrix} a_0 f_n + a_1 f_{n-1} + ... + a_k f_{n-k} \\ f_n \\ \vdots \\ f_{n-k+1} \end{pmatrix} = \underbrace{\begin{pmatrix} a_0 & a_1 & \dots & a_k \\ 1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ 0 & 0 & 1 & 0 \end{pmatrix}}_{:=M} \cdot \begin{pmatrix} f_n \\ f_{n-1} \\ \vdots \\ f_{n-k} \end{pmatrix}.$$

(a) Zeigen Sie, dass die charakteristische Gleichung der Matrix M gegeben ist durch

$$\lambda^{k+1} - a_0\lambda^k - a_1\lambda^{k-1} - ... - a_{k-1}\lambda - a_k = 0.$$

(b) Zeigen Sie, dass die Eigenräume $E(M, \lambda)$ zu jedem Eigenwert λ von M eindimensional sind, und geben Sie eine Basis von $E(M, \lambda)$ an.

(c) Zeigen Sie, dass M diagonalisierbar ist genau dann, wenn M genau $k + 1$ verschiedene Eigenwerte $\lambda_1, ..., \lambda_{k+1} \in \mathbb{R}$ besitzt. Geben Sie für diesen Fall eine Matrix $S \in \mathrm{GL}(k + 1, \mathbb{R})$ als Funktion der Eigenwerte $\lambda_1, ..., \lambda_{k+1}$ an, so dass $S \cdot M \cdot S^{-1}$ diagonal ist.

(d) Sei nun M diagonalisierbar. Geben Sie Matrizen A, B als Funktion der Eigenwerte $\lambda_1, ..., \lambda_{k+1}$ an, so dass

$$\begin{pmatrix} f_{k+n} \\ \vdots \\ f_n \end{pmatrix} = A \cdot B^{-1} \begin{pmatrix} f_k \\ \vdots \\ f_0 \end{pmatrix}$$

(e) Zeigen Sie: ist $f_n \neq 0$ für alle $n \in \mathbb{N}_0$ und konvergiert die Quotientenfolge $(q_n)_{n\in\mathbb{N}_0}$ mit $q_n = f_{n+1}/f_n$ gegen den Grenzwert $q = \lim_{n\to\infty} q_n$, so ist q ein Eigenwert von M.

Aufgabe 164: Berechnen Sie die Exponentiale der folgenden Matrizen

(a) $M = \begin{pmatrix} 0 & \mathrm{i} \\ -\mathrm{i} & 0 \end{pmatrix}$ (b) $N = \begin{pmatrix} 2 & 0 \\ 2 & -\mathrm{i}\pi \end{pmatrix}$ (c) $P = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ (d) $Q = \begin{pmatrix} 0 & \mathrm{i} \\ 1 & 0 \end{pmatrix}$.

Aufgabe 165: Sei $n \in \mathbb{N}$ und $\mathbb{K}$ ein Körper.

(a) Zeigen Sie, dass eine Matrix $M \in \mathrm{Mat}(n \times n, \mathbb{K})$ mit allen oberen Dreiecksmatrizen $D \in \mathrm{Mat}(n \times n, \mathbb{K})$ kommutiert genau dann, wenn $M = \lambda \mathbb{1}_n$ mit $\lambda \in \mathbb{K}$.

(b) Sei $M \in \mathrm{Mat}(n \times n, \mathbb{K})$ eine beliebige *untere* Dreiecksmatrix. Geben Sie eine Matrix $S \in \mathrm{GL}(n, \mathbb{K})$ an, so dass $S \cdot M \cdot S^{-1}$ eine *obere* Dreiecksmatrix ist.

Aufgabe 166: (Staatsexamen 2010) Bestimmen Sie alle Lösungen des Differentialgleichungssystems

$$\begin{aligned} x' &= 8x + 10y \quad & x(0) &= -1 \\ y' &= -5x - 6y \quad & y(0) &= 2 \end{aligned}$$

Aufgabe 167: (Staatsexamen 2010) Bestimmen Sie alle Lösungen des Differentialgleichungssystems

$$\begin{pmatrix} f_1' \\ f_2' \\ f_3' \end{pmatrix} = \begin{pmatrix} -1 & 1 & -2 \\ 0 & -1 & 4 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} f_1 \\ f_2 \\ f_3 \end{pmatrix}.$$

Aufgabe 168: (Staatsexamen 2012) Bestimmen Sie alle Lösungen des Differentialgleichungssystems

$$\begin{pmatrix} f_1' \\ f_2' \\ f_3' \end{pmatrix} = \begin{pmatrix} 1 & -1 & 1 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{pmatrix} \cdot \begin{pmatrix} f_1 \\ f_2 \\ f_3 \end{pmatrix}.$$

Aufgabe 169: Wir betrachten für $a, b \in \mathbb{R}$, $v = (v_1, v_2)^T \in \mathbb{R}^2$ das System von Differentialgleichungen erster Ordnung mit konstanten Koeffizienten

$$\begin{aligned} f_1' &= f_2 & \qquad f_1(0) &= v_1 \\ f_2' &= -af_1 + 2bf_2 & \qquad f_2(0) &= v_2. \end{aligned}$$

(a) Stellen Sie die Matrixgleichung

$$\begin{pmatrix} f_1' \\ f_2' \end{pmatrix} = A \cdot \begin{pmatrix} f_1 \\ f_2 \end{pmatrix} \qquad \text{mit} \quad A \in \mathrm{Mat}(2 \times 2, \mathbb{C})$$

auf und bestimmen Sie die Matrix $A \in \mathrm{Mat}(2 \times 2, \mathbb{C})$.

(b) Berechnen Sie die Eigenwerte der Matrix $A \in \mathrm{Mat}(2 \times 2, \mathbb{C})$ in Abhängigkeit von $a, b \in \mathbb{R}$ und untersuchen Sie A auf Diagonalisierbarkeit über $\mathbb{R}$ und $\mathbb{C}$.

(c) Bestimmen Sie für die Werte von $a, b \in \mathbb{R}$, für die A über $\mathbb{C}$ diagonalisierbar ist, eine Matrix $S \in \mathrm{GL}(2, \mathbb{C})$ und eine Diagonalmatrix $D \in \mathrm{Mat}(2 \times 2, \mathbb{C})$ mit $D = S \cdot A \cdot S^{-1}$.

(d) Bestimmen Sie für die Werte von $a, b \in \mathbb{R}$, für die A über $\mathbb{C}$ trigonalisierbar, aber nicht diagonalisierbar ist, eine Matrix $S \in \mathrm{GL}(2, \mathbb{C})$ und eine obere Dreiecksmatrix D mit $D = S \cdot A \cdot S^{-1}$.

(e) Berechnen Sie mit (c) und (d) das Exponential $\exp(tA)$ für $t \in \mathbb{R}$ und bestimmen Sie die Lösung des Differentialgleichungssystems mit $f_1(0) = v_1$ und $f_2(0) = v_2$.

Aufgabe 170: Sei $\mathbb{K} = \mathbb{C}$ oder $\mathbb{K} = \mathbb{R}$. Eine Differentialgleichung $n$ter Ordnung mit konstanten Koeffizienten $a_0, ..., a_{n-1} \in \mathbb{K}$ und Anfangswerten $v_1, .., v_n \in \mathbb{K}$ ist eine Gleichung der Form

$$f^{(n)} + a_{n-1}f^{(n-1)} + ... + a_1 f^{(1)} + a_0 f = 0 \qquad f(0) = v_1, f'(0) = v_2, ..., f^{(n-1)}(0) = v_n,$$

wobei $f^{(n)}$ die $n$te Ableitung einer n-mal differenzierbaren Funktion $f : \mathbb{R} \to \mathbb{K}$ bezeichnet.

Eine Lösung ist eine n-mal differenzierbare Funktion $f : \mathbb{R} \to \mathbb{K}$, die die Gleichungen erfüllt. Man kann zeigen, dass zu jedem Vektor $v = (v_1, ..., v_n)^T \in \mathbb{K}^n$ genau eine Lösung existiert.

(a) Wandeln Sie die angegebene Differentialgleichung $n$ter Ordnung in ein Differentialgleichungssystem 1. Ordnung mit konstanten Koeffizienten um, so dass jede Lösung des Differentialgleichungssystems zum Anfangswert $v = (v_1, ..., v_n)^T \in \mathbb{K}^n$ eine Lösung der Differentialgleichung $n$ter Ordnung mit Anfangswerten $v_1, ..., v_n$ liefert und umgekehrt.

(b) Stellen Sie die Matrixgleichung für dieses Differentialgleichungssystem 1. Ordnung auf und die charakteristische Gleichung für die darin auftretende Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{C})$. Zeigen Sie, dass jede Zahl $\lambda \in \mathbb{C}$ als Eigenwert von $A \in \mathrm{Mat}(n \times n, \mathbb{C})$ auftreten kann.

Aufgabe 171: Wir betrachten für $M \in \mathrm{Mat}(n \times n, \mathbb{C})$, $b \in \mathbb{C}^n$ das inhomogene lineare Differentialgleichungssystem mit konstanten Koeffizienten

$$f' = M \cdot f + b$$

und das zugehörige homogene Differentialgleichungssystem

$$f' = M \cdot f.$$

(a) Zeigen Sie, dass die Lösungen $f : \mathbb{R} \to \mathbb{C}^n$ des homogenen Differentialgleichungssystems einen n-dimensionalen Vektorraum L_M über $\mathbb{C}$ bilden, und dass die Abbildung $\phi : L_M \to \mathbb{C}^n$, $f \mapsto f(0)$ ein Vektorraumisomorphismus ist.
(b) Zeigen Sie, dass die Lösungen $f : \mathbb{R} \to \mathbb{C}^n$ des inhomogenen Differentialgleichungssystem einen n-dimensionalen affinen Raum A_M über $\mathbb{C}$ bilden.
(c) Zeigen Sie, dass die Ableitung einer Lösung $f : \mathbb{R} \to \mathbb{C}^n$ des inhomogenen Differentialgleichungssystems eine Lösung des homogenen Differentialgleichungssystems ist. Konstruieren Sie damit aus den Lösungen des homogenen Differentialgleichungssystems Lösungen des inhomogenen Differentialgleichungssystems für invertierbare Matrizen M.

Hinweis: Sie dürfen voraussetzen, dass zu jedem Anfangswert $f(0) = v \in \mathbb{C}^n$ genau eine Lösung existiert.

13.7 Aufgaben zu Kapitel 8

Aufgabe 172: Der **größte gemeinsame Teiler** zweier Polynome $p, q \in \mathbb{K}[x]$ ist das eindeutig bestimmte normierte Polynom $t \in \mathbb{K}[x]$ maximalen Grades, das sowohl p als auch q teilt. Man schreibt $t = \mathrm{ggT}(p, q)$. Die Polynome p, q heißen **teilerfremd**, wenn $\mathrm{ggT}(p, q) = 1$.

Seien $p, q \in \mathbb{K}[x]$ und $I := (\{p, q\})$ das von p und q erzeugte Ideal. Zeigen Sie:

(a) Es gilt $I = \{r \cdot p + s \cdot q | r, s \in \mathbb{K}[x]\}$.
(b) $\mathrm{ggT}(p, q)$ teilt alle Elemente in I.
(c) Es gilt $I = (\mathrm{ggT}(p, q))$.
(d) Sind p, q teilerfremd, so gibt es Polynome $r, s \in \mathbb{K}[x]$ mit $r \cdot p + s \cdot q = 1$.

Aufgabe 173:

(a) Bestimmen Sie alle normierten Teiler der Polynome p_1, p_2 in $\mathbb{Q}[x]$, $\mathbb{R}[x]$ und $\mathbb{C}[x]$:

$$p_1 = (x^2 - x + 1)(x^2 - x - 1) \qquad p_2 = (x^2 - 4)(x + 2)$$

(b) Bestimmen Sie Polynome $s, r \in \mathbb{R}[x]$ mit $q = s \cdot p + r$ und $\deg(r) < \deg(p)$ für

$$q = x^4 + 5x - 6,\ p = x - 1 \qquad\qquad q = x^5 + 3x^3 - x^2 + 4,\ p = x^2 - x + 1.$$

Aufgabe 174: Untersuchen Sie, ob die folgenden Polynome $p \in \mathbb{R}[x]$ in dem von dem Polynom $q = x^2 + x - 2 \in \mathbb{R}[x]$ erzeugten Ideal $(q) \subseteq \mathbb{R}[x]$ enthalten sind:

(a) $p = (x^2 + x - 2)(x^7 - 19x^2 + 51) + 17x + 19$,
(b) $p = (x - 1)(x + 2)(x - 3)(x + 4)$,
(c) $p = x^5 - x^3 + 2x^2 + 4$,
(d) $p = 2x^6 - x^5 - 7x^4 - 5x^3 - 3x + 2$.

Aufgabe 175: In dieser Aufgabe zeigen wir, dass nicht nur der Kern eines Ringhomomorphismus ein Ideal ist, sondern auch jedes Ideal Kern eines Ringhomomorphismus.

Sei $\mathbb{K}$ ein Körper, R ein kommutativer Ring mit Eins und $I \subseteq R$ ein Ideal.

(a) Zeigen Sie, dass $r \sim s \Leftrightarrow r - s \in I$ eine Äquivalenzrelation auf R ist.

(b) Zeigen Sie, dass die Verknüpfungen $+, \cdot : R/I \times R/I \to R/I$ mit $[r] + [s] := [r + s]$ und $[r] \cdot [s] := [r \cdot s]$ wohldefiniert sind und der Quotientenmenge R/I die Struktur eines kommutativen Rings geben. Folgern Sie, dass die Abbildung $\phi : R \to R/I$, $r \mapsto [r]$ ein Ringhomomorphismus mit $\ker(\phi) = I$ ist.

Aufgabe 176: Sei $\mathbb{K}$ ein Körper und $p \in \mathbb{K}[x]$ ein normiertes Polynom, das ausser sich selbst und 1 keine normierten Teiler in $\mathbb{K}[x]$ hat.

(a) Wir betrachten das von p erzeugte Ideal $I = (p) \subseteq \mathbb{K}[x]$. Zeigen Sie, dass $\mathbb{L} := \mathbb{K}[x]/(p)$ ein Körper ist.

(b) Zeigen Sie, dass $\mathbb{L}$ den Körper $\mathbb{K}$ als Teilkörper enthält und dass das Polynom $p \in \mathbb{K}[x]$ eine Nullstelle in $\mathbb{L}$ hat.

Hinweis: Benutzen Sie für (a) Aufgabe 175. Überlegen Sie sich, dass es für die Existenz des multiplikativen Inversen von $[q] \neq 0$ mit $q \in \mathbb{K}[x]$ ausreicht, Polynome $a, b \in \mathbb{K}[x]$ mit $a \cdot p + b \cdot q = 1$ zu finden. Benutzen Sie dazu Aufgabe 172.

Aufgabe 177: Beweisen Sie, dass die Abbildung $f : \mathbb{K}[x] \to \mathrm{Pol}(\mathbb{K})$, die einem Polynom $p \in \mathbb{K}[x]$ die Polynomabbildung $f_p : \mathbb{K} \to \mathbb{K}$, $\lambda \mapsto p(\lambda) = \mathrm{ev}_\lambda(p)$ zuordnet, injektiv ist genau dann, wenn der Körper $\mathbb{K}$ unendlich ist.

Aufgabe 178: Geben Sie ein Beispiel eines unendlich-dimensionalen reellen Vektorraums V und eines Endomorphismus $\phi : V \to V$ an mit $p(\phi) = \mathrm{ev}_\phi(p) \neq 0$ für alle Polynome $p \in \mathbb{R}[x]$.

Aufgabe 179: Sei $\mathbb{K}$ ein Körper und $n \in \mathbb{N}$. Sei V ein n-dimensionaler Vektorraum über $\mathbb{K}$, A eine $\mathbb{K}$-Algebra mit Eins und $M \subseteq A$ eine Teilmenge.

(a) Zeigen Sie, dass $\mathrm{Ann}(M) = \{p \in \mathbb{K}[x] \,|\, \mathrm{ev}_m(p) = 0 \; \forall m \in M\}$ ein Ideal in $\mathbb{K}[x]$ ist.
(b) Bestimmen Sie ein Polynom $p \in \mathbb{K}[x]$ mit $\mathrm{Ann}(M) = (p)$ für die folgenden Fälle:
  (i) $A = \mathbb{K}$ und $M = \{0, 1, 2\}$,
  (ii) $A = \mathrm{End}_\mathbb{K}(V)$, $M = \{0 : V \to V\}$,
  (iii) $A = \mathrm{End}_\mathbb{K}(V)$, $M = \{P \in \mathrm{End}_\mathbb{K}(V) \,|\, P \circ P = P\}$,
  (iv) $A = \mathbb{K}[x]$ und $M = \{x^3 - x^2 + 2x - 1, x + 2\}$,

Aufgabe 180: Bestimmen Sie das charakteristische Polynom der folgenden Matrizen $A \in \mathrm{Mat}(n \times n, \mathbb{K})$ mit dem *Laplaceschen Entwicklungssatz* und alle Eigenwerte in $\mathbb{K}$:

(a) $\mathbb{K} = \mathbb{R}$

$$A = \begin{pmatrix} -1 & 0 & 1 \\ -1 & -1 & 7 \\ 2 & 0 & 2 \end{pmatrix},$$

(b) $\mathbb{K} = \mathbb{Q}$

$$A = \begin{pmatrix} -1 & 1 & 1 & 2 \\ 0 & 3 & 0 & 0 \\ 0 & 5 & -1 & 0 \\ 5 & 4 & 2 & 3 \end{pmatrix},$$

(c) $\mathbb{K} = \mathbb{Z}/7\mathbb{Z}$

$$A = \begin{pmatrix} [6] & [3] & [0] \\ [4] & [5] & [0] \\ [1] & [2] & [6] \end{pmatrix},$$

(d) $\mathbb{K} = \mathbb{C}$

$$A = \begin{pmatrix} -1 & 5 & 1 & 0 \\ -1 & 3 & 2 & 0 \\ 0 & 0 & 1 & 0 \\ 5 & 4 & 2 & 0 \end{pmatrix}.$$

Aufgabe 181: Sei $\mathbb{K}$ ein Körper. Bestimmen Sie das Minimalpolynom des Endomorphismus

$$\phi : \mathrm{Mat}(4 \times 4, \mathbb{K}) \to \mathrm{Mat}(4 \times 4, \mathbb{K}), \qquad A \mapsto A \cdot \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Aufgabe 182: Sei $\mathbb{K}$ ein Körper, $\lambda \in \mathbb{K}\backslash\{0\}$ und $n \in \mathbb{N}$. Zeigen Sie: Ist $M \in \mathrm{Mat}(n \times n, \mathbb{K})$ eine Matrix mit Minimalpolynom $q_M \in \mathbb{K}[x]$, so ist das Minimalpolynom der Matrix $A = M - \lambda \mathbb{1}_n$ gegeben durch $q_A(x + \lambda)$ und das Minimalpolynom von $B = \lambda M$ durch $q_B = q_M(x/\lambda)$.

Aufgabe 183: Bestimmen Sie die charakteristischen Polynome und Minimalpolynome der folgenden Matrizen und untersuchen Sie sie auf Trigonalisierbarkeit und Diagonalisierbarkeit über $\mathbb{R}$:

$$A = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 1 & 1 \\ 2 & 1 & 0 \end{pmatrix} \qquad B = \begin{pmatrix} 4 & 0 & 0 \\ 4 & 2 & -2 \\ 0 & 3 & 1 \end{pmatrix} \qquad C = \begin{pmatrix} -3 & 1 & 0 \\ 1 & -1 & -2 \\ -1 & 1 & -2 \end{pmatrix}$$

Aufgabe 184: Wir betrachten den reellen Vektorraum $\mathrm{Pol}_{\leq n}(\mathbb{R})$ der Polynomabbildungen $p : \mathbb{R} \to \mathbb{R}$ vom Grad $\leq n$ und den Endomorphismus $\phi : \mathrm{Pol}_{\leq n} \to \mathrm{Pol}_{\leq n}$, $p \mapsto -p + p'$. Bestimmen Sie das Minimalpolynom von ϕ.

Aufgabe 185: Bestimmen Sie die charakteristischen Polynome und Minimalpolynome der folgenden Matrizen in $\mathrm{Mat}(4 \times 4, \mathbb{Q})$ und untersuchen Sie die Matrizen auf Trigonalisierbarkeit und Diagonalisierbarkeit über $\mathbb{Q}$:

$$A = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 2 & 1 \end{pmatrix} \qquad B = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 2 & -1 & 1 & 0 \\ 1 & -1 & 0 & -1 \end{pmatrix} \qquad C = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}$$

Aufgabe 186:

(a) Sei $M \in \mathrm{Mat}(n \times n, \mathbb{R})$ eine Matrix mit Minimalpolynom $q_M = x^2 + bx - c$ mit $b, c \in \mathbb{R}$. Bestimmen Sie die Minimalpolynome der Matrizen $A = M - \mathbb{1}_n$ und $B = \frac{1}{2}M$.

(b) Bestimmen sie das Minimalpolynom der linearen Abbildung

$$\phi : \mathrm{Mat}(n \times n, \mathbb{K}) \to \mathrm{Mat}(n \times n, \mathbb{K}), \qquad \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \mapsto \begin{pmatrix} a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \\ a_{11} & a_{12} & \dots & a_{1n} \end{pmatrix}$$

Aufgabe 187: Beweisen Sie die folgenden Aussagen:

(a) Ist $m \in \mathbb{N}$ ungerade, so hat jede Matrix $M \in \mathrm{Mat}(m \times m, \mathbb{R})$ mindestens einen Eigenwert.

(b) Ist das Minimalpolynom von $M \in \mathrm{Mat}(n \times n, \mathbb{K})$ gleich dem charakteristischen Polynom von M, so sind alle Eigenräume von M eindimensional.

(c) Für jede Matrix $M \in \mathrm{Mat}(n \times n, \mathbb{K})$ hat $\phi : \mathrm{Mat}(n \times n, \mathbb{K}) \to \mathrm{Mat}(n \times n, \mathbb{K})$, $A \mapsto M \cdot A$ das gleiche Minimalpolynom wie M.

(d) Für jede Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{K})$ gilt $\deg(q_A) = \dim_{\mathbb{K}}(\mathrm{span}_{\mathbb{K}}\{A^k | k \in \mathbb{N}_0\})$.

13.8 Aufgaben zu Kapitel 9

Aufgabe 188: (Die reelle Jordan-Normalform)

In dieser Aufgabe leiten Sie eine Normalform für reelle Matrizen her, die die Jordan-Normalform für komplexe Matrizen verallgemeinert.

(a) Sei V ein komplexer Vektorraum und $\phi : V \to V$ ein Endomorphismus. Seien $(v_1, ..., v_s)$ und $(w_1, ..., w_s)$ Jordanketten von ϕ zu den Eigenwerten $\mu \in \mathbb{C}$ und $\overline{\mu} \in \mathbb{C}$ und $U = \mathrm{span}_{\mathbb{C}}\{v_1, ..., v_s, w_1, ..., w_s\}$. Berechnen Sie die darstellende Matrix von $\phi|_U : U \to U$ bezüglich der Basis $B = (v_1 + w_1, \mathrm{i}(v_1 - w_1), ..., v_s + w_s, \mathrm{i}(v_s - w_s))$.

(b) Zeigen Sie, dass für jede Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{R})$ gilt: $\mathrm{def}(A - \mu\mathbb{1}_n)^k = \mathrm{def}(A - \overline{\mu}\mathbb{1}_n)^k$ für alle $\mu \in \mathbb{C}$ und $k \in \mathbb{N}$.

(c) Da $A \in \mathrm{Mat}(n \times n, \mathbb{R})$ trigonalisierbar über $\mathbb{C}$ ist, existiert eine Matrix $J \in \mathrm{Mat}(n \times n, \mathbb{C})$ in Jordan-Normalform und eine Matrix $S \in \mathrm{GL}(n, \mathbb{C})$ mit $A = S \cdot J \cdot S^{-1}$. Folgern Sie aus (b), dass die Anzahl der Jordanblöcke $J_r(\mu)$ der Länge $r \in \mathbb{N}$ zum Eigenwert $\mu \in \mathbb{C}$ von A gleich der Anzahl der Jordanblöcke $J_r(\overline{\mu})$ der Länge r zum Eigenwert $\overline{\mu} \in \mathbb{C}$ ist.

(d) Kombinieren Sie (a) und (c) um zu folgern, dass jede Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{R})$ ähnlich ist über $\mathbb{R}$ zu einer Matrix der Form

$$J^{\mathbb{R}} = \begin{pmatrix} J_{n_1}(\lambda_1) & 0 & 0 & \dots & 0 & 0 \\ 0 & \ddots & \ddots & & \vdots & \vdots \\ 0 & \ddots & J_{n_k}(\lambda_k) & 0 & 0 & 0 \\ \vdots & \ddots & 0 & J^{\mathbb{R}}_{m_1}(\mu_1) & \ddots & 0 \\ 0 & \dots & 0 & 0 & \ddots & 0 \\ 0 & 0 & \dots & 0 & 0 & J^{\mathbb{R}}_{m_l}(\mu_l) \end{pmatrix} \in \mathrm{Mat}(n \times n, \mathbb{R}),$$

wobei $\lambda_1, ..., \lambda_k \in \mathbb{R}$ reelle Eigenwerte von A sind, $\mu_1, ..., \mu_l \in \mathbb{C}$ komplexe Eigenwerte von A mit $\mathrm{Im}\,(\mu_j) > 0$, und die Matrizen $J^{\mathbb{R}}_n(\mu)$ gegeben sind durch

$$J^{\mathbb{R}}_n(\mu) = \begin{pmatrix} R_\mu & \mathbb{1}_2 & 0 & \dots & 0 \\ 0 & R_\mu & \mathbb{1}_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & R_\mu & \mathbb{1}_2 \\ 0 & \dots & 0 & 0 & R_\mu \end{pmatrix} \in \mathrm{Mat}(2n \times 2n, \mathbb{R}), \quad R_\mu = \begin{pmatrix} \mathrm{Re}\,(\mu) & -\mathrm{Im}\,(\mu) \\ \mathrm{Im}\,(\mu) & \mathrm{Re}\,(\mu) \end{pmatrix}.$$

Zeigen Sie, dass $J^{\mathbb{R}}$ eindeutig ist, bis auf Permutationen der Blöcke $J_{n_i}(\lambda_i)$ und $J^{\mathbb{R}}_{m_j}(\mu_j)$.

Aufgabe 189: Bestimmen Sie für die folgenden Matrizen $M \in \mathrm{Mat}(n \times n, \mathbb{K})$ für alle Eigenwerte $\lambda \in \mathbb{K}$ die geometrischen und die algebraische Vielfachheiten, den Fitting-Index von $M - \lambda \mathbb{1}_n$, die Defekte $\mathrm{def}(M - \lambda \mathbb{1}_n)^k$ für alle $k \in \mathbb{N}$ sowie das charakteristische Polynom und das Minimalpolynom. Geben Sie jeweils eine Diagonalmatrix D und eine nilpotente Matrix N an, so dass $M = D + N$ und $D \cdot N = N \cdot D$.

(a) $M = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$ (b) $M = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$

(c) $M = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$ (d) $M = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$

Aufgabe 190: Sind die folgenden Matrizen $A, B \in \mathrm{Mat}(4 \times 4, \mathbb{R})$ ähnlich? Geben Sie, falls A

und B ähnlich sind, eine Matrix S mit $B = S \cdot A \cdot S^{-1}$ an.

(a) $A = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}$ $\qquad B = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix}$

(b) $A = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}$ $\qquad B = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix}$

(c) $A = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}$ $\qquad B = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 1 & 2 \end{pmatrix}$

(d) $A = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix}$ $\qquad B = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 1 & 2 \end{pmatrix}.$

Aufgabe 191: Wir betrachten die Matrix

$$A = \begin{pmatrix} -2 & -1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ -3 & -1 & 0 & -1 & 0 \\ -1 & -3 & 1 & -2 & 0 \\ 0 & 0 & 0 & 0 & -1 \end{pmatrix} \in \mathrm{Mat}(5 \times 5, \mathbb{Q}).$$

Bestimmen Sie die Jordan-Normalform von A und eine Matrix $S \in \mathrm{GL}(5, \mathbb{Q})$, so dass $S^{-1} \cdot A \cdot S$ Jordan-Normalform hat.

Aufgabe 192: Wir betrachten den von den folgenden Abbildungen $f_1, ..., f_5 : \mathbb{R} \to \mathbb{R}$ aufgespannten Untervektorraum U des reellen Vektorraums $\mathrm{Abb}(\mathbb{R}, \mathbb{R})$

$$f_1 : x \mapsto 1 + x^2, \quad f_2 : x \mapsto x, \quad f_3 : x \mapsto x^2 - x, \quad f_4 : x \mapsto e^{2x}, \quad f_5 : x \mapsto (x-1)e^{2x}$$

und die Einschränkung $d|_U : U \to U$ der Ableitung auf U. Bestimmen Sie eine Jordan-Basis für $d|_U$ und die zugehörige Jordan-Normalform von $d|_U$.

Aufgabe 193: Beweisen Sie, dass eine Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{K})$ trigonalisierbar ist genau dann, wenn ihre Transponierte A^T trigonalisierbar ist, und dass in diesem Fall die Jordan-Normalformen von A^T und A übereinstimmen.

Aufgabe 194: (Staatsexamen 2014) Wir betrachten das Differentialgleichungssystem

$$\dot{f} = A \cdot f = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & -1 \\ 0 & 0 & -1 \end{pmatrix} \cdot f \qquad f(0) = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$$

(a) Bestimmen Sie die Jordan-Normalform der Matrix $A \in \mathrm{Mat}(3 \times 3, \mathbb{C})$ und eine Matrix $S \in \mathrm{GL}(3, \mathbb{C})$, so dass $S^{-1} \cdot A \cdot S$ Jordan-Normalform hat.

(b) Bestimmen Sie die Lösung des Differentialgleichungssystems zum angegebenen Anfangswert mit Hilfe von (a).

13.9 Aufgaben zu Kapitel 10

Aufgabe 195: Untersuchen Sie, ob die folgenden Bilinearformen bzw. Sesquilinearformen β auf den angegebenen $\mathbb{K}$-Vektorräumen V (i) symmetrisch bzw. hermitesch, (ii) alternierend, (iii) positiv (semi)definit, (iv) nicht ausgeartet oder (v) Skalarprodukte sind.

(a) $\mathbb{K}$ beliebig, $V = \mathbb{K}^4$, $\beta(x,y) = x_1y_2 - x_2y_1 + x_3y_4 - x_4y_3$,
(b) $\mathbb{K}$ beliebig, $V = \mathbb{K}^3$, $\beta(x,y) = x_2y_3 - x_2y_3$,
(c) $\mathbb{K} = \mathbb{C}$, $V = \mathrm{Abb}(\mathbb{R},\mathbb{C})$, $\beta(f,g) = \overline{f(0)}g(0)$,
(d) $\mathbb{K}$ beliebig, $V = \mathbb{K}^3$, $\beta(x,y) = x_1y_1 + x_2y_3 - x_3y_2$,
(e) $\mathbb{K} = \mathbb{R}$, $V = C^1([0,1],\mathbb{R})$, $\beta(f,g) = \int_0^1 f(x)g(x) + f'(x)g'(x)\,dx$,
(f) $\mathbb{K} = \mathbb{C}$, $V = C^1([0,1],\mathbb{C})$, $\beta(f,g) = \int_0^1 \overline{f'(x)}g'(x)\,dx$,
(g) $\mathbb{K} = \mathbb{C}$, $V = C^1([0,1],\mathbb{C})$, $\beta(f,g) = \int_0^1 f(x)g(x)\,dx$,
(h) $\mathbb{K} = \mathbb{C}$, $V = C^1([0,1],\mathbb{C})$, $\beta(f,g) = \int_0^1 \overline{f(x)}g(x) + \overline{f'(x)}g'(x)\,dx$,
(i) $\mathbb{K} = \mathbb{R}$, $V = C^1([0,1],\mathbb{R})$, $\beta(f,g) = \int_0^1 f(x)g'(x) + f'(x)g(x)\,dx$.

Hier bezeichnet $C^1([0,1],\mathbb{K})$ für $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$ den $\mathbb{K}$-Vektorraum der einmal stetig differenzierbaren Abbildungen $f : [0,1] \to \mathbb{K}$ und f' die Ableitung von f.

Aufgabe 196: Sei V ein komplexer Vektorraum. Zeigen Sie, dass die Nullabbildung die einzige alternierende Sesquilinearform auf V ist.

Aufgabe 197: Wir bezeichnen mit $|| \; ||$ die euklidische Norm auf dem $\mathbb{R}^n$. Für eine Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{R})$ definieren wir

$$||A||_{sup} = \sup\{||A \cdot v|| \,|v \in \mathbb{R}^n, ||v|| = 1\}$$

Zeigen Sie, dass $|| \; ||_{sup} : \mathrm{Mat}(n \times n, \mathbb{R}) \to \mathbb{R}$ eine Norm auf $\mathrm{Mat}(n \times n, \mathbb{R})$ ist. Untersuchen Sie, ob es ein Skalarprodukt $\langle \, , \, \rangle$ auf $\mathrm{Mat}(n \times n, \mathbb{R})$ gibt mit $||A||_{sup}^2 = \langle A, A\rangle$ für alle $A \in \mathrm{Mat}(n \times n, \mathbb{R})$.

Aufgabe 198: Geben Sie ein Beispiel einer symmetrischen Bilinearform β auf einem n-dimensionalen reellen Vektorraum V an, für die eine Basis $B = (b_1, ..., b_n)$ von V existiert mit $\beta(b_i, b_i) > 0$ für alle $i \in \{1, ..., n\}$, die aber nicht positiv definit ist.

Aufgabe 199: Sei V mit $\langle \, , \, \rangle$ ein euklidischer Vektorraum und $d : V \times V \to \mathbb{R}$, $(v,w) \mapsto d(v,w) = ||v - w||$ die zugehörige Abstandsfunktion. Zeigen Sie, dass zu $v, w \in V$ immer ein Vektor $m \in V$ mit $d(m,v) = d(m,w) = \frac{1}{2}d(v,w)$ existiert.

Aufgabe 200: Sei V ein $\mathbb{K}$-Vektorraum und $\beta : V \times V \to \mathbb{K}$ eine Bi- oder Sesquilinearform auf V.

(a) Zeigen Sie, dass für jeden Vektor $v \in V \setminus \{0\}$ und jeden Skalar $\lambda \in \mathbb{K}$ die Teilmenge $E_{v,\lambda} = \{w \in V | \beta(v,w) = \lambda\}$ ein affiner Unterraum von V ist.
(b) Bestimmen Sie die Dimension von E_λ falls β ein Skalarprodukt auf V ist.
(c) Bestimmen Sie für $V = \mathbb{R}^3$ mit dem Standardskalarprodukt den Schnitt $E_{v,1} \cap E_{w,1}$ für $v = e_1 + e_2$ und $w = e_1 - e_3$.

Aufgabe 201: Sei V mit $\langle\, , \rangle$ ein euklidischer Vektorraum, $|| \; ||$ die zugehörige Norm und $\phi : V \to V$ eine Abbildung mit $\phi(0) = 0$ und $||\phi(x) - \phi(y)|| = ||x - y||$ für alle $x, y \in V$. Zeigen Sie, dass die Abbildung ϕ eine $\mathbb{R}$-lineare Abbildung ist.

Aufgabe 202: Bestimmen Sie für die folgenden $\mathbb{K}$-Vektorräume V mit den angegebenen Bi- oder Sesquilinearformen β und Untervektorräumen $U \subseteq V$ jeweils die orthogonalen Komplemente $U^\perp$, $(U^\perp)^\perp$ sowie den Schnitt $U \cap U^\perp$.

(a) $V = \mathbb{R}^4$, $\beta(x, y) = -x_1y_1 - x_2y_2 + x_3y_3$, $U = \mathrm{span}_\mathbb{R}\{e_1 + e_3, e_2 + e_4\}$,
(b) $V = \mathbb{C}^3$, $\beta(x, y) = \overline{x_1}y_1 + \overline{x_2}y_2$, $U = \mathrm{span}_\mathbb{C}\{e_1, e_3\}$,
(c) $V = \mathbb{R}^4$, $\beta(x, y) = x_1y_2 - x_2y_1 + x_3y_4 - x_4y_3$, $U = \mathrm{span}_\mathbb{R}\{e_1, e_3\}$,
(d) $V = \mathbb{R}^4$, $\beta(x, y) = x_1y_2 - x_2y_1 + x_3y_3 + x_4y_4$, $U = \mathrm{span}_\mathbb{R}\{e_1, e_3\}$.

Aufgabe 203: Sei V mit $\langle\, , \,\rangle$ ein euklidischer oder unitärer Vektorraum über $\mathbb{K}$. Beweisen Sie die folgenden Eigenschaften des orthogonalen Komplements

(a) $A \subseteq B \Rightarrow B^\perp \subseteq A^\perp$
(b) $A \subseteq B^\perp \Leftrightarrow B \subseteq A^\perp$
(c) $((A^\perp)^\perp)^\perp = A$

Aufgabe 204: Sei V ein Vektorraum über $\mathbb{K}$ und V^* sein Dualraum. Zeigen Sie, dass zu jedem linear unabhängigen n-Tupel $(\alpha^1, ..., \alpha^n)$ von Vektoren $\alpha^i \in V^*$ und zu beliebigen $\lambda_1, ..., \lambda_n \in \mathbb{K}$ ein Vektor $v \in V$ existiert mit $\alpha^i(v) = \lambda_i$ für alle $i \in \{1, ..., n\}$.

Hinweis: Betrachten Sie die Abbildung $\phi : V \to \mathbb{K}^n$, $v \mapsto (\alpha^1(v), ..., \alpha^n(v))^T$ und die nicht ausgeartete symmetrische Bilinearform $\beta : \mathbb{K}^n \times \mathbb{K}^n \to \mathbb{K}$, $\beta(v, w) = \Sigma_{i=1}^n v_i w_i$.

Aufgabe 205: (**Der Hilbertsche Folgenraum**)
Wir betrachten den komplexen Vektorraum $\mathrm{Abb}(\mathbb{N}_0, \mathbb{C})$ der komplexen Folgen $(a_n)_{n\in\mathbb{N}_0}$ mit der üblichen Vektoraddition und Skalarmultiplikation:

$$(a_n)_{n\in\mathbb{N}_0} + (b_n)_{n\in\mathbb{N}_0} = (a_n + b_n)_{n\in\mathbb{N}_0} \qquad \lambda(a_n)_{n\in\mathbb{N}_0} = (\lambda a_n)_{n\in\mathbb{N}_0} \quad \forall \lambda \in \mathbb{C}.$$

(a) Zeigen Sie, dass die Folgen $(a_n)_{n\in\mathbb{N}_0}$, für die die Reihe $\Sigma_{n=0}^\infty |a_n|^2$ konvergiert, einen Untervektorraum $\ell^2 \subseteq \mathrm{Abb}(\mathbb{N}_0, \mathbb{C})$ bilden.
(b) Zeigen Sie, dass für alle Folgen $(a_n)_{n\in\mathbb{N}_0}, (b_n)_{n\in\mathbb{N}_0} \in \ell^2$ die Reihe $\Sigma_{n=0}^\infty \overline{a_n} \cdot b_n$ konvergiert.
(c) Zeigen Sie, dass die folgende Abbildung ein Skalarprodukt auf ℓ^2 ist

$$\langle\, , \,\rangle : \ell^2 \times \ell^2 \to \mathbb{C}, \quad \langle (a_n)_{n\in\mathbb{N}_0}, (b_n)_{n\in\mathbb{N}_0} \rangle = \Sigma_{n=0}^\infty \overline{a_n} \cdot b_n.$$

(d) Wir betrachten für $k \in \mathbb{N}_0$ die Folgen $e_k = (\delta_{nk})_{n\in\mathbb{N}_0} \in \ell^2$. Zeigen Sie, dass jede endliche Teilmenge der Menge $E = \{e_k | k \in \mathbb{N}_0\}$ ein Orthonormalsystem ist.
(e) Zeigen Sie, dass $E^\perp = \{0\}$ gilt, und folgern Sie, dass $\mathrm{span}_\mathbb{C}(E) \oplus (\mathrm{span}_\mathbb{C}(E))^\perp \neq \ell^2$.

Aufgabe 206: Seien $(V, \langle\, , \rangle_V)$ und $(W, \langle\, , \rangle_W)$ euklidische oder unitäre Vektorräume, $|| \; ||_V$ und $|| \; ||_W$ die zugehörigen Normen und $\phi : V \to W$ eine lineare Abbildung. Beweisen Sie die Äquivalenz der Aussagen (i) und (ii)

(i) $\langle \phi(v), \phi(v') \rangle_W = \langle v, v' \rangle_V$ für alle $v, v' \in V$,
(ii) $||\phi(v)||_W = ||v||_V$ für alle $v \in V$.

Aufgabe 207: Wir betrachten den reellen Vektorraum $\mathrm{Pol}(\mathbb{R})$ der Polynomabbildungen $p : \mathbb{R} \to \mathbb{R}$ und $\langle \, , \, \rangle : \mathrm{Pol}(\mathbb{R}) \times \mathrm{Pol}(\mathbb{R}) \to \mathbb{R}$

$$\langle p, q \rangle = \tfrac{1}{2} \int_{-1}^{1} p(x) q(x) \, dx.$$

(a) Zeigen Sie, dass $\langle \, , \, \rangle$ ein Skalarprodukt auf $\mathrm{Pol}(\mathbb{R})$ ist.
(b) Berechnen Sie die Längen aller Monome $x^k : \mathbb{R} \to \mathbb{R}$, $r \mapsto r^k$ für $k \in \mathbb{N}_0$.
(c) Sei β_{kl} der Winkel zwischen den Monomen x^{2k} und x^{2l} für $l, k \in \mathbb{N}_0$. Zeige Sie, dass $\lim_{k\to\infty} \beta_{kl} = \frac{\pi}{2}$ für alle $l \in \mathbb{N}_0$ und $\lim_{k\to\infty} \beta_{k(n+k)} = 0$ für alle $n \in \mathbb{N}_0$ gilt.
(d) Bestimmen Sie das orthogonale Komplement der Menge

$$M = \{ p \in \mathrm{Pol}(\mathbb{R}) | p = x^k \text{ mit } k \in \mathbb{N}_0 \text{ gerade} \}.$$

Aufgabe 208: Wir betrachten den reellen Vektorraum $\mathbb{R}^2$ mit der symplektischen Form $\beta(x, y) = x_1 y_2 - x_2 y_1$. Bestimmen Sie die symplektische Gruppe $\mathrm{Sp}(2, \mathbb{R}) = \mathrm{Sp}(\mathbb{R}^2, \beta)$.

Aufgabe 209: Untersuchen Sie, ob die folgenden komplexen Matrizen unitär sind:

(a) $A = \begin{pmatrix} \mathrm{i}/\sqrt{3} & 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{3} & -\mathrm{i}/\sqrt{2} & \mathrm{i}/\sqrt{2} \\ 1/\sqrt{3} & 0 & 0 \end{pmatrix}$,
(b) $B = \begin{pmatrix} 1/\sqrt{2} & -1\sqrt{2} & 0 \\ 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \\ 1/\sqrt{6} & 1/\sqrt{6} & -2/\sqrt{6} \end{pmatrix}$,
(c) $C = \begin{pmatrix} \mathrm{i}\cos\phi & -\sin\phi & 0 \\ \mathrm{i}\sin(\phi) & \cos\phi & 0 \\ 0 & 0 & -\mathrm{i} \end{pmatrix}$,
(d) eine beliebige Permutationsmatrix $D \in \mathrm{Mat}(n \times n, \mathbb{C})$.

Aufgabe 210: Wir betrachten den komplexen Vektorraum $\mathrm{Mat}(n \times n, \mathbb{C})$ mit dem Skalarprodukt $\langle M, N \rangle = \mathrm{Tr}(M^\dagger \cdot N)$.

(a) Zeigen Sie, dass die Elementarmatrizen $E_{ij} \in \mathrm{Mat}(n \times n, \mathbb{C})$ mit $(E_{ij})_{kl} = \delta_{ik}\delta_{jl}$ eine Orthonormalbasis von $\mathrm{Mat}(n \times n, \mathbb{C})$ bilden.
(b) Bestimmen Sie eine Orthonormalbasis des Untervektorraums

$$U = \{ M \in \mathrm{Mat}(n \times n, \mathbb{C}) | M^T = M \} \subseteq \mathrm{Mat}(n \times n, \mathbb{C}).$$

(c) Bestimmen Sie das orthogonale Komplement $U^\perp$.
(d) Bestimmen Sie die Matrix $B \in U$ mit dem kleinsten Abstand von $A = E_{11} + E_{12} + ... + E_{1n}$ und berechnen Sie $||A - B||$.
(e) Sei $n \geq 2$. Berechnen Sie den Winkel zwischen den Matrizen $E_{11} + E_{12}$ und $2E_{11} + 2E_{21}$.

Aufgabe 211: Wir betrachten den euklidischen Vektorraum $V = \mathrm{Mat}(2 \times 2, \mathbb{R})$ mit dem Skalarprodukt $\langle M, N \rangle = \mathrm{Tr}(M^T \cdot N)$ und die Matrizen

$$M_1 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad M_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad M_3 = \begin{pmatrix} 1 & 0 \\ 1 & -1 \end{pmatrix} \quad M_4 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

(a) Bestimmen Sie die Matrix im Untervektorraum $U = \mathrm{span}_{\mathbb{R}}\{M_2, M_4\}$, die den kleinsten Abstand von der Matrix $A = \Sigma_{i=1}^4 a_i M_i$ hat.
(b) Bestimmen Sie die Orthogonalprojektion auf $W = \mathrm{span}_{\mathbb{R}}\{M_1, M_3\}$.

Aufgabe 212: Sei V ein endlich-dimensionaler euklidischer oder unitärer Vektorraum und $P_U : V \to V$ die Orthogonalprojektion auf den Untervektorraum $U \subseteq V$. Zeigen Sie, dass dann $\mathrm{id}_V - P_U : V \to V$ die Orthogonalprojektion auf das orthogonale Komplement $U^\perp$ ist.

Aufgabe 213: Bestimmen Sie die QR-Zerlegung der Matrix

$$A = \begin{pmatrix} 1 & -1 & 2 & \mathrm{i} \\ 1 & 1 & 1 & 0 \\ \mathrm{i} & 0 & -2\mathrm{i} & -1 \\ 0 & \mathrm{i} & \mathrm{i} & -1 \end{pmatrix}.$$

Aufgabe 214: Wir betrachten den $\mathbb{R}^4$ mit der Minkowski-Metrik

$$\eta(x, y) = -x_1 y_1 + x_2 y_2 + x_3 y_3 + x_4 y_4.$$

(a) Berechnen Sie die beschreibenden Matrizen ${}^C\eta^C$ und ${}^B\eta^B$ bezüglich der geordneten Basen $C = (e_1, e_2, e_3, e_4)$ und $B = (e_1 + e_2, e_1 - e_2, e_1 + e_3, e_1 + e_4)$.
(b) Berechnen Sie die Basiswechselmatrix ${}_CM_B(\mathrm{id}_{\mathbb{R}^4})$ und verifizieren Sie durch Nachrechnen die Transformationsformel aus Satz 9.3.4:

$${}^B\eta^B = {}_CM_B(\mathrm{id}_{\mathbb{R}^4})^T \cdot {}^C\eta^C \cdot {}_CM_B(\mathrm{id}_{\mathbb{R}^4}).$$

Aufgabe 215: Eine **Gruppenwirkung** einer Gruppe G auf einer Menge M ist eine Abbildung $\rho : G \times M \to M$ mit $\rho(g, \rho(h, m)) = \rho(g \cdot h, m)$ für alle $g, h \in G$, $m \in M$. Zeigen Sie:

(a) Ist $\rho : G \times M \to M$ eine Gruppenwirkung, so ist $m \sim m' \Leftrightarrow \exists g \in G : m' = \rho(g, m)$ eine Äquivalenzrelation auf M.
(b) Für jedes Element $m \in M$ ist der **Stabilisator** von m $\mathrm{Stab}(m) = \{g \in G \mid \rho(g, m) = m\}$ eine Untergruppe von G.
(c) Ist $m \sim m'$, so sind die Gruppen $\mathrm{Stab}(m)$ und $\mathrm{Stab}(m')$ isomorph.
(d) Die folgenden Abbildungen $\rho : \mathrm{GL}(n, \mathbb{K}) \times \mathrm{Mat}(n \times n, \mathbb{K}) \to \mathrm{Mat}(n \times n, \mathbb{K})$ sind Gruppenwirkungen von $\mathrm{GL}(n, \mathbb{K})$ auf $\mathrm{Mat}(n \times n, \mathbb{K})$:
(i) $(S, M) \mapsto S \cdot M \cdot S^{-1}$, (ii) $(S, M) \mapsto S \cdot M \cdot S^T$, (iii) $(S, M) \mapsto S \cdot M \cdot S^\dagger$
(e) Sei V ein n-dimensionaler $\mathbb{K}$-Vektorraum, B eine Basis von V und β eine Bilinearform (Sesquilinearform) auf V. Dann ist die Abbildung ${}_BM_B : U(V, \beta) \to \mathrm{Stab}({}^B\beta^B)$ ein Gruppenisomorphismus , wenn $\mathrm{Mat}(n \times n, \mathbb{K})$ mit der Gruppenwirkung (ii) aus (d) (mit der Gruppenwirkung (iii) aus (d)) ausgestattet wird.

13.10 Aufgaben zu Kapitel 11

Aufgabe 216: Wir betrachten $\mathbb{R}^2$ mit der Minkowski-Metrik $\eta(x,y) = -x_1y_1 + x_2y_2$.

(a) Bestimmen Sie die Adjungierte von $\phi_M : \mathbb{R}^2 \to \mathbb{R}^2$, $v \mapsto M \cdot v$ mit $M \in \mathrm{Mat}(2 \times 2, \mathbb{R})$.
(b) Bestimmen Sie alle η-selbstadjungierten und alle η-normalen Endomorphismen $\phi : \mathbb{R}^2 \to \mathbb{R}^2$.
(c) Bestimmen Sie die unitäre Gruppe $\mathrm{O}(1,1) = U(\mathbb{R}^2, \eta)$.
(d) Zeigen Sie, dass jeder η-unitäre Endomorphismus von der Form $\phi_M : \mathbb{R}^2 \to \mathbb{R}^2$, $v \mapsto M \cdot v$ ist mit

$$M = \begin{pmatrix} \pm\cosh\theta & \sinh\theta \\ \sinh\theta & \pm\cosh\theta \end{pmatrix} \quad \text{oder} \quad M = \begin{pmatrix} \pm\cosh\theta & -\sinh\theta \\ \sinh\theta & \mp\cosh\theta \end{pmatrix} \qquad \text{und } \theta \in \mathbb{R}.$$

Hinweis: Nutzen Sie in (b)-(d) aus, dass jede $\mathbb{R}$-lineare Abbildung $\phi : \mathbb{R}^2 \to \mathbb{R}^2$ von der in (a) angegebenen Form ist und leiten Sie Bedingungen an die Matrix M her.

Aufgabe 217: Wir betrachten den komplexen Vektorraum $\mathrm{Mat}(n \times n, \mathbb{C})$ mit dem Skalarprodukt $\langle M, N\rangle = \mathrm{Tr}(M^\dagger \cdot N)$ und für eine Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{C})$ den Endomorphismus

$$\phi_A : \mathrm{Mat}(n \times n, \mathbb{C}) \to \mathrm{Mat}(n \times n, \mathbb{C}), \quad M \mapsto A \cdot M.$$

Zeigen Sie, dass der adjungierte Endomorphismus gegeben ist durch

$$(\phi_A)^\dagger = \phi_{A^\dagger} : \mathrm{Mat}(n \times n, \mathbb{C}) \to \mathrm{Mat}(n \times n, \mathbb{C}), \quad M \mapsto A^\dagger \cdot M.$$

Folgern Sie: ist A (i) normal, (ii) unitär, (iii) (anti)selbstadjungiert, so ist auch ϕ_A (i) normal, (ii) unitär, (iii) (anti)selbstadjungiert.

Aufgabe 218: Wir betrachten den $\mathbb{K}$-Vektorraum $\mathrm{Mat}(n \times n, \mathbb{K})$, seinen Dualraum und für $A \in \mathrm{Mat}(n \times n, \mathbb{K})$ die Abbildung $\mu_A : \mathrm{Mat}(n \times n, \mathbb{K}) \to \mathbb{K}$, $M \mapsto \mathrm{tr}\,(A \cdot M)$.

(a) Zeigen Sie, dass für alle $A \in \mathrm{Mat}(n \times n, \mathbb{K})$ gilt $\mu_A \in \mathrm{Mat}(n \times n, \mathbb{K})^*$.
(b) Zeige Sie, dass $\mathrm{Mat}(n \times n, \mathbb{K})^* = \{\mu_A | A \in \mathrm{Mat}(n \times n, \mathbb{K})\}$.

Aufgabe 219: Wir betrachten einen n-dimensionalen $\mathbb{K}$-Vektorraum V mit einer nicht ausgearteten symmetrischen Bilinearform $\beta : V \times V \to \mathbb{K}$ und seinen Dualraum $V^* = \mathrm{Hom}_{\mathbb{K}}(V, \mathbb{K})$.

- Die zu einer Basis $B = (v_1, ..., v_n)$ von V **duale Basis** ist die Basis $B^* = (\vartheta^1, ..., \vartheta^n)$ von V^*, wobei $\vartheta^i : V \to \mathbb{K}$ die lineare Abbildung mit $\vartheta^i(v_j) = \delta_{ij}$ bezeichnet.
- Der zu einem Endomorphismus $\phi : V \to V$ **duale Endomorphismus** $\phi^* : V^* \to V^*$ ist gegeben durch $\phi^*(\alpha) = \alpha \circ \phi$ für alle $\alpha \in V^*$.

Seien $\flat : V \to V^*$, $v \mapsto v^\flat$ und $\sharp = \flat^{-1} : V^* \to V$, $\alpha \mapsto \alpha^\sharp$ die in der Vorlesung eingeführten linearen Abbildungen mit $v^\flat(w) = \beta(v, w)$ und $\beta(\alpha^\sharp, w) = \alpha(w)$ für alle $v, w \in V$ und $\alpha \in V^*$. Sei $\phi^\dagger = \sharp \circ \phi^* \circ \flat : V \to V$ der zu ϕ adjungierte Endomorphismus.

(a) Zeigen Sie, dass die Abbildungen $\flat : V \to V^*$ und $\sharp = \flat^{-1} : V^* \to V$ gegeben sind durch

$$\flat(v_i) = v_i^\flat = \Sigma_{k=1}^n M_{ik} \vartheta^k \qquad \sharp(\vartheta^i) = (\vartheta^i)^\sharp = \Sigma_{k=1}^n M_{ik}^{-1} v_k \qquad \forall i \in \{1, ..., n\}.$$

wobei $M = {}^B\beta^B$ die beschreibende Matrix von β bezüglich B bezeichnet. Folgern Sie:

$${}_{B^*}M_B(\flat) = ({}^B\beta^B)^T = {}^B\beta^B \qquad {}_B M_{B^*}(\sharp) = ({}^B\beta^B)^{-1}.$$

(b) Sei $\phi : V \to V$ ein Endomorphismus mit beschreibender Matrix $A = {}_BM_B(\phi)$. Zeigen Sie, dass für den dualen Endomorphismus $\phi^* : V^* \to V^*$ gilt

$$\phi^*(\vartheta^i) = \Sigma_{k=1}^n A_{ik}\vartheta^k \qquad \forall i \in \{1, ..., n\}.$$

Folgern Sie, dass ${}_{B^*}M_{B^*}(\phi^*) = {}_BM_B(\phi)^T$.

(c) Zeigen Sie mit (a) und (b), dass für den adjungierten Endomorphismus $\phi^\dagger : V \to V$ gilt

$$\phi^\dagger(v_i) = \Sigma_{j,k,l=1}^n M_{ij}A_{jk}M_{kl}^{-1}v^l \qquad \forall i \in \{1, ..., n\}.$$

Folgern Sie, dass die beschreibende Matrix von $\phi^\dagger$ die Bedingung aus Satz 10.1.3 erfüllt:

$$({}^B\beta^B)^{-1} \cdot {}_BM_B(\phi^\dagger)^T \cdot {}^B\beta^B = {}_BM_B(\phi).$$

Aufgabe 220: Sei V mit $\langle\, , \rangle$ ein unendlich-dimensionaler euklidischer der unitärer Vektorraum. Dann heißt eine Abbildung $\phi^\dagger : V \to V$ adjungiert zu einem Endomorphismus $\phi : V \to V$ wenn $\langle \phi^\dagger(v), w\rangle = \langle v, \phi(w)\rangle$ für alle $v, w \in V$ gilt. Die Linearität und Eindeutigkeit der Adjungierten folgt wie im endlich-dimensionalen Fall, falls eine Adjungierte existiert.

Wir betrachten den euklidischen Vektorraum $C_c^\infty(\mathbb{R})$ der unendlich oft stetig differenzierbaren Abbildungen $f : \mathbb{R} \to \mathbb{R}$ mit **kompaktem Träger** und das Skalarprodukt $\langle\, , \rangle$

$$C_c^\infty(\mathbb{R}) = \{f \in C^\infty(\mathbb{R}) \,|\, \exists c_f \in \mathbb{R} : f(x) = 0\, \forall x \in \mathbb{R} \text{ mit } |x| \geq c_f\} \qquad \langle f, g\rangle = \int_{-\infty}^{\infty} f(x)g(x)\, dx.$$

(a) Bestimmen Sie die Adjungierte der Ableitung $d : C_c^\infty(\mathbb{R}) \to C_c^\infty(\mathbb{R})$.

(b) Bestimmen Sie die Adjungierte von $\phi_\alpha : C_c^\infty(\mathbb{R}) \to C_c^\infty(\mathbb{R})$, $f \mapsto f_\alpha$ mit $f_\alpha(x) = f(\alpha x)$ für ein festes $\alpha \in \mathbb{R} \setminus \{0\}$.

(c) Zeigen Sie, dass $F_\psi : C_c^\infty(\mathbb{R}) \to C_c^\infty(\mathbb{R})$, $f \mapsto f(0)\psi$ für ein fest gewähltes $\psi \in C_c^\infty(\mathbb{R}) \setminus \{0\}$ keine Adjungierte besitzt.

Aufgabe 221: Sei V ein Vektorraum über $\mathbb{K}$. Eine Bilinearform oder Sesquilinearform β auf V heißt **orthosymmetrisch** wenn gilt: $\beta(x, y) = 0 \Leftrightarrow \beta(y, x) = 0$. Zeigen Sie, dass jede orthosymmetrische Bilinearform (Sesquilinearform) entweder symmetrisch (hermitesch) oder antisymmetrisch (antihermitesch) ist.

Aufgabe 222: Wir betrachten für $n, m \in \mathbb{N}$ den komplexen Vektorraum $V = \mathrm{Mat}(m \times n, \mathbb{C})$ mit der Sesquilinearform $\beta_A(M, N) = \mathrm{Tr}(M^\dagger \cdot A \cdot N)$ für und $A \in \mathrm{Mat}(m \times m, \mathbb{C})$.

(a) Zeigen sie, dass β_A hermitesch ist genau dann, wenn A hermitesch ist, und nicht ausgeartet genau dann, wenn $\mathrm{rg}(A) = m$.

(b) Sei A hermitesch vom Rang m. Bestimmen Sie die Adjungierten der Endomorphismen $L_B, R_C : \mathrm{Mat}(m \times n, \mathbb{C}) \to \mathrm{Mat}(m \times n, \mathbb{C})$ mit $L_B(M) = B \cdot M$ und $R_C(M) = M \cdot C$ für $B \in \mathrm{Mat}(m \times m, \mathbb{C})$ und $C \in \mathrm{Mat}(n \times n, \mathbb{C})$.

(c) Sei A hermitesch vom Rang m. Bestimmen Sie die Matrizen $B \in \mathrm{Mat}(m \times m, \mathbb{C})$ und $C \in \mathrm{Mat}(n \times n, \mathbb{C})$, für die L_B und R_C (i) unitär, (ii) selbstadjungiert und (iii) normal sind.

Aufgabe 223: Sei V mit $\langle\, , \rangle$ ein endlich-dimensionaler euklidischer oder unitärer Vektorraum und $P : V \to V$ eine Projektionsabbildung. Zeigen Sie, dass P selbstadjungiert ist genau dann, wenn P eine *Orthogonalprojektion* ist, d. h. $\mathrm{im}(P) = \ker(P)^\perp$ gilt.

Aufgabe 224: Wir betrachten die symmetrische Bilinearform $\beta_A(x, y) = x^T \cdot A \cdot y$ auf dem $\mathbb{R}^3$ für die symmetrische Matrix

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

(a) Bestimmen Sie die Signatur von β_A.
(b) Bestimmen Sie den Typ der Quadrik $Q = \{x \in \mathbb{R}^3 : \beta_A(x, x) = 1\}$ und ihre Achsen. Skizzieren Sie die Quadrik.
(c) Bestimmen Sie eine orthogonale Matrix $S \in O(3, \mathbb{R})$, so dass $S^{-1} \cdot A \cdot S$ diagonal ist.

Aufgabe 225: Sei $n \in \mathbb{N}$ und $A \in \mathrm{Mat}(n \times n, \mathbb{C})$. Zeigen Sie:

(a) Ist A normal, so ist auch $\exp(A)$ normal.
(b) Ist A selbstadjungiert, so ist auch $\exp(A)$ selbstadjungiert.
(c) Ist A selbstadjungiert, so ist $\mathrm{i}A$ anti-selbstadjungiert und $\exp(\mathrm{i}A)$ unitär.
(d) Es gibt eine selbstadjungierte Matrix A_+ und eine anti-selbstadjungierte Matrix A_- mit $A = A_+ + A_-$.
(e) A ist normal genau dann, wenn $A_+ \cdot A_- = A_- \cdot A_+$.

Aufgabe 226: Beweisen Sie die folgenden Aussagen:

(a) Eine Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{C})$ ist genau dann antihermitesch, wenn $\mathrm{i}A$ hermitesch ist.
(b) Eine Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{C})$ ist genau dann antihermitesch, wenn sie normal ist und alle ihre komplexen Eigenwerte in $\mathrm{i}\mathbb{R}$ liegen.
(c) Eine Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{C})$ ist genau dann antihermitesch, wenn es eine unitäre Matrix $U \in \mathrm{U}(n, \mathbb{C})$ und eine Diagonalmatrix $D \in \mathrm{Mat}(n \times n, \mathbb{C})$ mit Diagonaleinträgen in $\mathrm{i}\mathbb{R}$ gibt, so dass $A = U \cdot D \cdot U^{-1}$.
(d) Eine antisymmetrische Matrix $B \in \mathrm{Mat}(n \times n, \mathbb{R})$ ist genau dann diagonalisierbar über $\mathbb{R}$, wenn sie die Nullmatrix ist.
(e) Eine antisymmetrische Matrix $B \in \mathrm{Mat}(n \times n, \mathbb{R})$ ist genau dann trigonalisierbar über $\mathbb{R}$, wenn sie die Nullmatrix ist.

Aufgabe 227: Zeigen Sie, dass eine antisymmetrische Matrix $M \in \mathrm{Mat}(n \times n, \mathbb{C}) \setminus \{0\}$ nie durch eine orthogonale Matrix diagonalisierbar ist.

Aufgabe 228: Untersuchen Sie die folgenden Matrizen auf Diagonalisierbarkeit über den angegebenen Körpern, ohne deren charakteristische Polynome zu berechnen.

(a) $\mathbb{K} = \mathbb{C}$

$$A = \begin{pmatrix} \mathrm{i} & -2 & \mathrm{i} & 1+\mathrm{i} \\ 2 & -2\mathrm{i} & -1 & 1-\mathrm{i} \\ \mathrm{i} & 1 & -\mathrm{i} & 2\mathrm{i} \\ \mathrm{i}-1 & -1-\mathrm{i} & 2\mathrm{i} & -\mathrm{i} \end{pmatrix}$$

(b) $\mathbb{K} = \mathbb{R}$ und $\mathbb{K} = \mathbb{C}$,

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

(c) $\mathbb{K} = \mathbb{R}$ und $\mathbb{K} = \mathbb{C}$,

$$A = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & \sqrt{3} & -1 \\ 0 & 0 & 1 & \sqrt{3} \end{pmatrix}$$

(d) $\mathbb{K} = \mathbb{R}$ und $\mathbb{K} = \mathbb{C}$,

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

Aufgabe 229: Sei $A \in \mathrm{Mat}(n \times n, \mathbb{R})$ eine symmetrische Matrix.

(a) Zeigen Sie, dass die durch A definierte symmetrische Bilinearform β_A auf dem $\mathbb{R}^n$ mit $\beta_A(x, y) = x^T \cdot A \cdot y$ genau dann positiv definit ist, wenn alle Eigenwerte von A positiv sind.
(b) Zeigen Sie, dass die folgende Matrix $A \in \mathrm{Mat}(6 \times 6, \mathbb{R})$ keine positiv definite Bilinearform β_A auf dem $\mathbb{R}^6$ definiert

$$A = \begin{pmatrix} \sqrt{2} & e^{\sqrt{7}} & -\cos 2 & \sqrt{23} & \pi & 1 \\ e^{\sqrt{7}} & -1 & -\ln 2 & 2 & 4 & 1 + \tan(6) \\ -\cos 2 & -\ln 2 & \pi - \sqrt{2} & \cos 3 & \sqrt{19} & 0 \\ \sqrt{23} & 2 & \cos 3 & 2 - \pi & \sqrt{13} & -1 \\ \pi & 4 & \sqrt{19} & \sqrt{13} & 3 & -2 \\ 1 & 1 + \tan(6) & 0 & -1 & -2 & -4 \end{pmatrix}$$

Aufgabe 230: Für eine symmetrische Matrix $Q \in \mathrm{Mat}(n \times n, \mathbb{R})$ betrachten wir

$$N_Q = \{x \in \mathbb{R}^n \mid x^T \cdot Q \cdot x = 0\} \subseteq \mathbb{R}^n$$

Klassifizieren Sie für $n \in \{2, 3\}$ die Mengen N_Q analog zur Klassifikation der Quadriken in der Vorlesung. Bestimmen Sie dabei jeweils eine Gleichung, die N_Q beschreibt, geben Sie an, um welche geometrische Figur es sich handelt, und skizzieren Sie diese.

Aufgabe 231: Wir betrachten den $\mathbb{R}^2$ mit dem Standardskalarprodukt.

(a) Zeigen Sie, dass für $d > c > 0$ die Menge

$$E = \{x \in \mathbb{R}^2 \mid ||x - ce_1|| + ||x + ce_1|| = 2d\}$$

der Punkte, für die die Summe der Abstände von den Punkten ce_1 und $-ce_1$ den Wert $2d$ annimmt, eine Ellipse ist. Die Punkte $\pm ce_1$ heißen **Brennpunkte** der Ellipse. Bestimmen Sie die Halbachsenlängen $a, b > 0$ in der Ellipsengleichung $(x_1/a)^2 + (x_2/b)^2 = 1$ als Funktion von c und d. Skizzieren Sie die Ellipse.
(b) Zeigen Sie, dass für $c > d > 0$ die Menge

$$H = \{x \in \mathbb{R}^2 \mid ||x - ce_1|| - ||x + ce_1|| = 2d\}$$

der Punkte, für die die Differenz der Abstände von den Punkten ce_1 und $-ce_1$ den Wert $2d$ annimmt, eine Hyperbel ist. Die Punkte $\pm ce_1$ heißen **Brennpunkte** der Hyperbel. Bestimmen Sie die Parameter $a, b > 0$ in der Hyperbelgleichung $(x_1/a)^2 - (x_2/b)^2 = 1$ als Funktion von c und d. Skizzieren Sie die Hyperbel.

(c) Zeigen Sie, dass für $c > 0$ die Menge

$$P = \{x \in \mathbb{R}^2 \mid ||x - ce_1|| = |x_1 + c|\}$$

der Punkte, die den gleichen Abstand vom Punkt ce_1 und der Geraden $x_1 = -c$ haben, eine Parabel ist. Der Punkt ce_1 heißt **Brennpunkt** und die Gerade $x_1 = -c$ **Leitlinie** der Parabel. Bestimmen Sie den Parameter $a \in \mathbb{R}$ in der Parabelgleichung $x_1 = ax_2^2$ in Abhängigkeit von c. Skizzieren Sie die Parabel.

(d) Wir betrachten die Gleichung $x_2^2 = 2px_1 + (\epsilon^2 - 1)x_1^2$ mit $\epsilon \geq 0$, $p > 0$. Der Parameter ϵ heißt **Exzentrizität**. Untersuchen Sie, welche der in (a)-(c) betrachteten Kurven durch diese Gleichung für verschiedene Werte von ϵ definiert werden. Skizzieren Sie diese Kurven für $p = 1$ und $\epsilon \in \{0, 1/2, 1, 2\}$.

Aufgabe 232: Wir betrachten die reelle symmetrische Matrix

$$A = \begin{pmatrix} 1 & 3 & 0 \\ 3 & 1 & 4 \\ 0 & 4 & 1 \end{pmatrix}$$

(a) Bestimmen Sie die Eigenwerte und Eigenvektoren von A.
(b) Bestimmen Sie eine orthogonale Matrix $O \in \mathrm{O}(3, \mathbb{R})$, so dass $O^T \cdot A \cdot O$ diagonal ist.
(c) Bestimmen Sie die Signatur der symmetrischen Bilinearform β_A mit $\beta_A(x, y) = x^T \cdot A \cdot y$.
(d) Bestimmen Sie den Typ der zugehörigen Quadrik $M_A = \{x \in \mathbb{R}^3 | x^T \cdot A \cdot x = 1\}$.

Aufgabe 233: Wir betrachten die reelle symmetrische Matrix

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

(a) Bestimmen Sie die Eigenwerte und Eigenvektoren von A.
(b) Bestimmen Sie eine orthogonale Matrix $O \in \mathrm{O}(3, \mathbb{R})$, so dass $O^T \cdot A \cdot O$ diagonal ist.
(c) Bestimmen Sie die Signatur der symmetrischen Bilinearform β_A mit $\beta_A(x, y) = x^T \cdot A \cdot y$.
(d) Bestimmen Sie den Typ der zugehörigen Quadrik $M_A = \{x \in \mathbb{R}^3 | x^T \cdot A \cdot x = 1\}$.

13.11 Aufgaben zu Kapitel 12

Aufgabe 234: Seien U, V Vektorräume über einem Körper $\mathbb{K}$, $X \subseteq U$, $Y \subseteq V$ Untervektorräume und $\pi_U : U \to U/X$ und $\pi_V : V \to V/Y$ die kanonischen Surjektionen. Zeigen Sie, dass zu jeder $\mathbb{K}$-linearen Abbildung $\phi : U \to V$ mit $\phi(X) \subseteq Y$ genau eine $\mathbb{K}$-lineare Abbildung $\tilde{\phi} : U/X \to V/Y$ mit $\pi_V \circ \phi = \tilde{\phi} \circ \pi_U$ existiert.

Aufgabe 235: Sei V ein $\mathbb{K}$-Vektorraum und $\phi : V \to V$ ein Endomorphismus. Für $k \in \mathbb{N}_0$ betrachten wir die Untervektorräume $V_k = \ker(\phi^k)$, die Quotienten $U_k = V_{k+1}/V_k$ und die kanonischen Surjektionen $\pi_k : V_{k+1} \to U_k$.

(a) Zeigen Sie, dass zu jedem $k \in \mathbb{N}$ genau eine $\mathbb{K}$-lineare Abbildung $\tilde{\phi}_k : U_k \to U_{k-1}$ mit $\tilde{\phi}_k \circ \pi_k = \pi_{k-1} \circ \phi|_{V_{k+1}}$ existiert.
(b) Stellen Sie das zugehörige kommutierende Diagramm auf.
(c) Zeigen Sie, dass $\mathrm{im}(\tilde{\phi}_k) \subseteq \ker(\tilde{\phi}_{k-1})$ für alle $k \in \mathbb{N}$ gilt.
(d) Sei nun V endlich-dimensional und ϕ nilpotent mit einer Jordanbasis, die aus einer einzigen Jordankette $K = (b_1, ..., b_f)$ besteht. Bestimmen Sie Basen der Vektorräume U_k für $k \in \mathbb{N}$, indem Sie Äquivalenzklassen von Basisvektoren der Jordankette betrachten. Bestimmen Sie die beschreibende Matrizen der Abbildungen $\tilde{\phi}_k$ bezüglich dieser Basen.

Aufgabe 236: Sei V mit $\langle\, , \rangle$ ein n-dimensionaler euklidischer Vektorraum und $U \subseteq V$ ein Untervektorraum der Dimension $n-1$. Zeigen Sie, dass dann jede Äquivalenzklasse in V/U genau einen Vektor $v \in V$ mit $v \in U^\perp$ enthält.

Aufgabe 237: Sei $V = U \oplus W$ ein Vektorraum über $\mathbb{K}$ und $X \subseteq V$ ein Untervektorraum. Zeigen Sie, dass der Vektorraum V/X isomorph ist zu $U/(U \cap X) \oplus V/(V \cap X)$.

Aufgabe 238: Sei V ein Vektorraum über einem Körper $\mathbb{K}$ und β eine symmetrische Bilinearform auf V. Zeigen Sie, dass durch $\tilde{\beta}([v],[w]) := \beta(v,w)$ für alle $v, w \in V$ eine nicht ausgeartete symmetrische Bilinearform $\tilde{\beta}$ auf dem Quotientenraum $V/V^\perp$ definiert wird.

Aufgabe 239: Sei $\mathbb{K}$ ein Körper, I, J beliebige Mengen und $\iota_I : I \to \langle I \rangle_{\mathbb{K}}$, $i \mapsto i$, $\iota_J : J \to \langle J \rangle_{\mathbb{K}}$, $j \mapsto j$ die Inklusionsabbildungen. Zeigen Sie, dass zu jeder Abbildung $f : I \to J$ genau eine $\mathbb{K}$-lineare Abbildung $\tilde{f} : \langle I \rangle_{\mathbb{K}} \to \langle J \rangle_{\mathbb{K}}$ mit $\tilde{f} \circ \iota_i = \iota_j \circ f$ existiert.

Aufgabe 240: Sei I eine Menge, $(U_i)_{i\in I}$, $(V_i)_{i\in I}$ und $(W_i)_{i\in I}$ Familien von Vektorräumen über $\mathbb{K}$ und $(f_i : U_i \to V_i)_{i\in I}$ und $(g_i : V_i \to W_i)_{i\in I}$ Familien $\mathbb{K}$-linearer Abbildungen. Beweisen Sie mit Hilfe der universellen Eigenschaften, dass für die dadurch induzierten Abbildungen zwischen den direkten Summen und Produkten dieser Vektorräume gilt

$$\Pi_{i\in I}(f_i \circ g_i) = (\Pi_{i\in I} f_i) \circ (\Pi_{i\in I} g_i) : \Pi_{i\in I} U_i \to \Pi_{i\in I} W_i$$
$$\oplus_{i\in I}(f_i \circ g_i) = (\oplus_{i\in I} f_i) \circ (\oplus_{i\in I} g_i) : \oplus_{i\in I} U_i \to \oplus_{i\in I} W_i$$
$$\Pi_{i\in I}\mathrm{id}_{V_i} = \mathrm{id}_{\Pi_{i\in I} V_i} \qquad \oplus_{i\in I}\,\mathrm{id}_{V_i} = \mathrm{id}_{\oplus_{i\in I} V_i}.$$

Aufgabe 241: Seien V, W, X, Y Vektorräume über $\mathbb{K}$ und $V \otimes_{\mathbb{K}} W$ mit $\tau : V \times W \to V \otimes_{\mathbb{K}} W$ und $X \otimes_{\mathbb{K}} Y$ mit $\tau' : X \times Y \to X \otimes_{\mathbb{K}} Y$ Tensorprodukte.

(a) Zeigen Sie, dass zu jedem Paar $\mathbb{K}$-linearer Abbildungen $f : V \to X$ und $g : W \to Y$ genau eine $\mathbb{K}$-lineare Abbildung $f \otimes g : V \otimes_{\mathbb{K}} W \to X \otimes_{\mathbb{K}} Y$ mit $(f \otimes g) \circ \tau = \tau' \circ (f \times g)$ existiert. Diese heißt das **Tensorprodukt** der Abbildungen f und g.
(b) Zeigen Sie, dass für alle Endomorphismen $f : V \to V$ und $g : W \to W$ das Tensorprodukt die Bedingung $f \otimes g = (f \otimes \mathrm{id}_W) \circ (\mathrm{id}_V \otimes g) = (\mathrm{id}_V \otimes g) \circ (f \otimes \mathrm{id}_W)$ erfüllt.

Aufgabe 242: Sei V, W reelle Vektorräume. Wir interpretieren $\mathbb{C}$ als einen zweidimensionalen reellen Vektorraum und betrachten die Tensorprodukte $V_{\mathbb{C}} := \mathbb{C} \otimes_{\mathbb{R}} V$ und $W_{\mathbb{C}} := \mathbb{C} \otimes_{\mathbb{R}} W$.

(a) Zeigen Sie, dass $V_{\mathbb{C}}$ die Struktur eines komplexen Vektorraums besitzt, und dass $\dim_{\mathbb{C}}(V_{\mathbb{C}}) = \dim_{\mathbb{R}}(V)$ gilt.

(b) Zeigen Sie, dass zu jeder $\mathbb{R}$-linearen Abbildung $f : V \to W$ genau eine $\mathbb{C}$-lineare Abbildung $f_{\mathbb{C}} : V_{\mathbb{C}} \to W_{\mathbb{C}}$ mit $f_{\mathbb{C}}(1 \otimes v) = f(v)$ für alle $v \in V$ existiert.
(c) Zeigen Sie, dass analoge Aussagen zu (a) und (b) für beliebige Körper $\mathbb{K}$ statt $\mathbb{C}$ und Teilkörper $\mathbb{F} \subseteq \mathbb{K}$ statt $\mathbb{R}$ durchgeführt werden können.

Aufgabe 243: Seien V, W Vektorräume über einem Körper $\mathbb{K}$ mit Tensoralgebren $T(V)$ und $T(W)$ und $\iota_V : V \to T(V)$, $\iota_W : W \to T(W)$ die zugehörigen Inklusionsabbildungen. Zeigen Sie, dass zu jeder $\mathbb{K}$-linearen Abbildung $\phi : V \to W$ genau ein Algebrahomomorphismus $\tilde{\phi} : T(V) \to T(W)$ mit $\tilde{\phi} \circ \iota_V = \iota_W \circ \phi$ existiert.

13.12 Aufgaben zur Wiederholung

Sei in diesem Abschnitt $\mathbb{K}$ ein Körper, $n \in \mathbb{N}$, V ein endlich-dimensionaler Vektorraum über $\mathbb{K}$ und W ein beliebiger Vektorraum über $\mathbb{K}$.

Aufgabe 244: (Permutationen und Permutationsmatrizen)
Wahr oder Falsch? Begründen oder widerlegen Sie die Aussagen.

(a) Die Permutationen $\sigma \in S_n$ mit $\mathrm{sgn}(\sigma) = 1$ bilden eine Untergruppe der Gruppe S_n.
(b) Sei $n \geq 3$ und $\sigma, \tau \in S_n$ zwei elementare Transpositionen. Dann gilt $\sigma \circ \tau = \tau \circ \sigma$.
(c) Jede Permutationsmatrix $A \in \mathrm{Mat}(n \times n, \mathbb{C})$ ist orthogonal.
(d) Für jede Permutationsmatrix $A \in \mathrm{Mat}(n \times n, \mathbb{K})$ gilt $\det(A) \in \{1, -1\}$.
(e) Die Permutationsmatrizen $A \in \mathrm{Mat}(n \times n, \mathbb{K})$ bilden eine Untergruppe der Gruppe $\mathrm{GL}(n, \mathbb{K})$.
(f) Jede Permutationsmatrix $A \in \mathrm{Mat}(n \times n, \mathbb{R})$ ist diagonalisierbar über $\mathbb{R}$.

Aufgabe 245: (Eigenwerte, Spuren und Determinanten)
Wahr oder Falsch? Begründen oder widerlegen Sie die Aussagen.

(a) Ist 0 kein Eigenwert von $A \in \mathrm{Mat}(n \times n, \mathbb{K})$, so ist A invertierbar.
(b) Ist $A \in \mathrm{Mat}(3 \times 3, \mathbb{Z}/2\mathbb{Z})$ so hat A mindestens einen Eigenwert mit algebraischer Vielfachheit ≥ 2.
(c) Ist $A \in \mathrm{Mat}(3 \times 3, \mathbb{Z}/2\mathbb{Z})$ trigonalisierbar, so hat A mindestens einen Eigenwert mit algebraischer Vielfachheit ≥ 2.
(d) Für alle $A, B \in \mathrm{Mat}(n \times n, \mathbb{K})$ haben $A \cdot B$ und $B \cdot A$ die gleichen Eigenwerte.
(e) Sei $\phi : V \to V$ ein Projektor und $f \in \mathbb{K}[x]$. Dann hat $f(\phi)$ die Eigenwerte $f(0)$ und $f(1)$.
(f) Sei $\phi : V \to V$ mit $\phi \neq 0$ ein Projektor mit $\det(\phi) = 0$ und $f \in \mathbb{K}[x]$. Dann hat $f(\phi)$ die Eigenwerte $f(0)$ und $f(1)$.
(g) Ist $\mathrm{tr}\,(A) = 0$ für eine Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{K})$, so ist A nicht invertierbar.
(h) Seien $A, B \in \mathrm{Mat}(n \times n, \mathbb{K})$ obere Dreiecksmatrizen. Dann gilt $\mathrm{tr}\,(AB) = \mathrm{tr}\,(A)\mathrm{tr}\,(B)$.
(i) Seien $A, B \in \mathrm{Mat}(n \times n, \mathbb{K})$ obere Dreiecksmatrizen mit positiven Diagonaleinträgen. Dann gilt $\mathrm{tr}\,(AB) > 0$.
(j) Seien $A, B \in \mathrm{Mat}(n \times n, \mathbb{K})$. Falls A, B und $B^{-1} + (AB)^{-1}B$ invertierbar sind so gilt $\det(A + B) \neq 0$.
(k) Für jede Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{R})$ haben A und A^T die gleichen Eigenwerte.
(l) Für jede Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{C})$ haben A und $A^\dagger$ die gleichen Eigenwerte.
(m) Sei ϕ: $\mathrm{Mat}(n \times n, \mathbb{C}) \to \mathrm{Mat}(n \times n, \mathbb{C})$, $A \mapsto A^T$. Dann ist $\det(\phi) = 1$.
(n) Sei $A \in \mathrm{Mat}(n \times n, \mathbb{C})$. Dann ist $\overline{\det(A)} = \det(A^\dagger)$.
(o) Sei $A \in \mathrm{Mat}(n \times n, \mathbb{R})$ antisymmetrisch. Dann ist $\det(A) = 0$.

(p) Jede reelle Matrix $A \in \mathrm{Mat}(7 \times 7, \mathbb{R})$ hat mindestens einen Eigenwert.
(q) Jede reelle Matrix $A \in \mathrm{Mat}(4 \times 4, \mathbb{R})$ hat mindestens einen Eigenwert.

Aufgabe 246: (Ähnlichkeit und Äquivalenz von Matrizen)
Wahr oder Falsch? Begründen oder widerlegen Sie die Aussagen.

(a) Jede Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{K})$ ist äquivalent zu einer Diagonalmatrix.
(b) Jede Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{K})$ ist ähnlich zu einer Diagonalmatrix.
(c) Jede Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{C})$ ist ähnlich zu einer unteren Dreiecksmatrix.
(d) Jede Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{C})$ ist ähnlich zu einer Diagonalmatrix.
(e) Sind $A, B \in \mathrm{Mat}(n \times n, \mathbb{K})$ zwei zueinander ähnliche Diagonalmatrizen, so gilt $A = B$.
(f) Haben $A, B \in \mathrm{Mat}(n \times n, \mathbb{K})$ die gleichen Fitting-Indizes, so sind sie ähnlich.
(g) Haben $A, B \in \mathrm{Mat}(n \times n, \mathbb{K})$ die gleichen Eigenwerte, so sind sie ähnlich.
(h) Haben $A, B \in \mathrm{Mat}(n \times n, \mathbb{K})$ das gleiche Minimalpolynom, so sind sie ähnlich.
(i) Haben $A, B \in \mathrm{Mat}(n \times n, \mathbb{K})$ das gleiche charakteristische Polynom, so sind sie ähnlich.
(j) Stimmen die Minimalpolynome von $A, B \in \mathrm{Mat}(n \times n, \mathbb{K})$ überein, so sind A und B ähnlich.

Aufgabe 247: (Diagonalisierbarkeit und Trigonalisierbarkeit)
Wahr oder Falsch? Begründen oder widerlegen Sie die Aussagen.

(a) Sei $A \in \mathrm{Mat}(n \times n, \mathbb{K})$ eine Matrix, die mit einer diagonalisierbaren Matrix $B \in \mathrm{Mat}(n \times n, \mathbb{K})$ kommutiert. Dann ist A diagonalisierbar.
(b) Sei $A \in \mathrm{Mat}(n \times n, \mathbb{R})$. Ist A^2 diagonalisierbar über $\mathbb{R}$, so gilt das auch für A.
(c) Sei V ein endlich-dimensionaler komplexer Vektorraum. Dann ist jede Involution $I : V \to V$ diagonalisierbar.
(d) Sei V ein endlich-dimensionaler Vektorraum über $\mathbb{K} = \mathbb{Z}/2\mathbb{Z}$. Dann ist jede Involution $I : V \to V$ diagonalisierbar.
(e) Sei V ein endlich-dimensionaler $\mathbb{K}$-Vektorraum. Dann ist jeder Projektor $P : V \to V$ diagonalisierbar.
(f) Jede Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{R})$ ist trigonalisierbar über $\mathbb{R}$.
(g) Jede Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{R})$ ist trigonalisierbar über $\mathbb{C}$.
(h) Jede komplexe Matrix ist trigonalisierbar über $\mathbb{C}$.
(i) Sind $A, B \in \mathrm{Mat}(n \times n, \mathbb{K})$ trigonalisierbar, so ist auch $A + B$ trigonalisierbar.
(j) Sind $A, B \in \mathrm{Mat}(n \times n, \mathbb{K})$ diagonalisierbar, so ist auch $A + B$ diagonalisierbar.
(k) Jede untere Dreiecksmatrix ist trigonalisierbar.
(l) Sei $A \in \mathrm{Mat}(n \times n, \mathbb{C})$ und $m \in \mathbb{N}$. Falls $A^m = \mathbb{1}_n$ so ist A diagonalisierbar.
(m) Sei A eine Matrix, deren charakteristisches Polynom in paarweise verschiedene Linearfaktoren zerfällt und die mit einer Matrix B kommutiert. Dann ist B diagonalisierbar.
(n) Jeder Projektor $\phi : V \to V$ ist diagonalisierbar.

Aufgabe 248: (Fitting-Index, Haupträume und Jordan-Normalform)
Wahr oder Falsch? Begründen oder widerlegen Sie die Aussagen.

(a) Es gibt eine nilpotente Matrix $A \in \mathrm{Mat}(3 \times 3, \mathbb{C})$ mit Fitting-Index 4.
(b) Für jede Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{C})$ ist die Länge des längsten Jordanblocks in der Jordan-Normalform von A gleich der größten Hauptraumdimension von A.
(c) Stimmen alle Eigenwerte von $A, B \in \mathrm{Mat}(n \times n, \mathbb{K})$ und alle Dimensionen der zugehörigen Haupträume überein, so sind A und B ähnlich.

(d) Es gibt eine Matrix $A \in \mathrm{Mat}(4 \times 4, \mathbb{C})$, so dass die Jordan-Normalform von $A^\dagger \cdot A$ nur einen Jordanblock enthält.
(e) Für eine unitäre Matrix $U \in \mathrm{Mat}(n \times n, \mathbb{C})$ hat jeder Jordanblock Länge 1.
(f) Stimmen das Minimalpolynom und das charakteristische Polynom eines Endomorphismus überein, so gibt es zu jedem Eigenwert λ genau einen Jordanblock.
(g) Sei λ ein Eigenwert des Endomorphismus $\phi\colon V \to V$ über $\mathbb{C}$. Dann ist der Fitting-Index von $\phi_\lambda := \phi - \lambda\,\mathrm{id}_V$ stets kleiner als die geometrische Vielfachheit des Eigenwerts λ.

Aufgabe 249: (Matrixexponentiale und Differentialgleichungen)
Wahr oder Falsch? Begründen oder widerlegen Sie die Aussagen.

(a) Sei $A \in \mathrm{Mat}(n \times n, \mathbb{C})$ und $m \in \mathbb{N}$. Dann gilt $\exp(A)^m = \exp(mA)$.
(b) Ist $A = \exp(B)$ mit $B \in \mathrm{Mat}(n \times n, \mathbb{C})$, so ist A invertierbar.
(c) Ist $A = \exp(B)$ und $B \in \mathrm{Mat}(n \times n, \mathbb{C})$ diagonalisierbar, so ist auch A diagonalisierbar.
(d) Sei $A \in \mathrm{Mat}(n \times n, \mathbb{R})$ eine Matrix mit $A^2 = 0$. Dann ist $x(t) := (tA + \mathbb{1}_n)x_0$ die Lösung des Anfangswertproblems $\dot{x} = Ax$, $x(0) = x_0$.
(e) Sei $A \in \mathrm{Mat}(n \times n, \mathbb{R})$. Dann hat $\exp(A)$ nur positive Einträge.
(f) Sei $A \in \mathrm{Mat}(n \times n, \mathbb{C})$ mit $A^\dagger = \pi i \mathbb{1}_n + A$. Dann ist $\exp(\exp(A))$ unitär.
(g) Für jede trigonalisierbare Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{C})$ gilt $\det(\exp(A)) \neq 0$.

Aufgabe 250: (nilpotente Matrizen und nilpotente Abbildungen)
Wahr oder Falsch? Begründen oder widerlegen Sie die Aussagen.

(a) Jede nilpotente Matrix ist eine obere Dreiecksmatrix.
(b) Jede nilpotente Matrix ist eine *echte* obere Dreiecksmatrix.
(c) Jede nilpotente Matrix ist ähnlich zu einer *echten* unteren Dreiecksmatrix.
(d) Die einzige diagonalisierbare nilpotente Matrix $N \in \mathrm{Mat}(n \times n, \mathbb{K})$ ist die Nullmatrix.
(e) Eine *echte* obere Dreiecksmatrix A ist diagonalisierbar genau dann, wenn $A = 0$.
(f) Ist $N \in \mathrm{Mat}(n \times n, \mathbb{K})$ eine nilpotente Matrix, so ist 1 der einzige Eigenwert von $\exp(N)$.
(g) Sei $A \in \mathrm{Mat}(n \times n, \mathbb{K})$ nilpotent. Dann ist $\mathbb{1}_n - A$ invertierbar.
(h) Sind $A, B \in \mathrm{Mat}(n \times n, \mathbb{K})$ nilpotent, so ist auch $A + B$ nilpotent.
(i) Sind $A, B \in \mathrm{Mat}(n \times n, \mathbb{K})$ nilpotent, so ist auch $A \cdot B$ nilpotent.
(j) Ist $A \in \mathrm{Mat}(n \times n, \mathbb{K})$ mit $\det(A) \neq 0$, so ist A nicht nilpotent.
(k) Sind $A, B \in \mathrm{Mat}(n \times n, \mathbb{K})$ nilpotent mit $A \cdot B = B \cdot A$, so ist auch $A + B$ nilpotent.
(l) Sind $A, B \in \mathrm{Mat}(n \times n, \mathbb{K})$ nilpotent mit $A \cdot B = B \cdot A$, so ist auch $A \cdot B$ nilpotent.
(m) Ist V ein endlich-dimensionaler $\mathbb{K}$-Vektorraum und $\phi \in \mathrm{End}_{\mathbb{K}}(V) \setminus \{0\}$ nilpotent, so ist ϕ nicht diagonalisierbar.
(n) Alle nilpotenten Matrizen $A \in \mathrm{Mat}(2 \times 2, \mathbb{R})$ sind zueinander ähnlich.
(o) Für jede nilpotente Matrix zerfällt das charakteristische Polynom in Linearfaktoren.

Aufgabe 251: (Ideale)
Wahr oder Falsch? Begründen oder widerlegen Sie die Aussagen.

(a) Sei R ein kommutativer unitaler Ring und $I \subseteq R$ ein Ideal. Dann ist $I = R \Leftrightarrow 1 \in I$.
(b) Sei R ein kommutativer Ring und $I, J \subseteq R$ Ideale. Dann ist auch $I \cap J$ ein Ideal.
(c) Sei R ein kommutativer Ring und $I, J \subseteq R$ Ideale. Dann ist auch $I \cup J$ ein Ideal.
(d) Sei $M = \{-2, 3, 6\} \subseteq \mathbb{Z}$. Dann ist das von M erzeugte Ideal $(M) = (12)$.

Aufgabe 252: (charakteristisches Polynom)
Wahr oder Falsch? Begründen oder widerlegen Sie die Aussagen.

(a) Eine Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{K})$ ist genau dann die Nullmatrix, wenn ihr charakteristisches Polynom p_A das Nullpolynom ist.
(b) Es gibt eine Matrix $A \in \mathrm{Mat}(2 \times 2, \mathbb{C})$ mit Eigenwerten in $\mathbb{N}$, Determinante $\det(A) = 7$ und Spur $\mathrm{Tr}(A) = 9$.
(c) Es gibt eine Matrix $A \in \mathrm{Mat}(4 \times 4, \mathbb{R})$ mit charakteristischen Polynom $p_A = -x^3 - 2x + 1$.
(d) Es gibt eine Matrix $A \in \mathrm{Mat}(4 \times 4, \mathbb{Q})$ mit charakteristischen Polynom $p_A = x^4 - \sqrt{2}x^3 - 2$.
(e) Es gibt eine Matrix $A \in \mathrm{Mat}(3 \times 3, \mathbb{R})$ mit charakteristischen Polynom $p_A = -x^3 - ix$.
(f) Es gibt eine Matrix $M \in \mathrm{GL}(3, \mathbb{R})$ mit charakteristischem Polynom $p_M = -x^3 + 7x^2 - 12x$.
(g) Es gibt eine Matrix $M \in \mathrm{Mat}(3 \times 3, \mathbb{C})$ mit Spur $\mathrm{Tr}(M) = 0$ und charakteristischem Polynom $p_M = -x^3 + x^2 - 1$.
(h) Ist das charakteristische Polynom von $A \in \mathrm{Mat}(2 \times 2, \mathbb{R})$ gleich $x^2 + 3x + 2$, so ist 3 kein Eigenwert von A.
(i) Hat $A \in \mathrm{Mat}(n \times n, \mathbb{R})$ das charakteristische Polynom $p_A = (-1)^n x^n + 2x^{n-1} + x + 1$, so ist A invertierbar.
(j) Hat $A \in \mathrm{Mat}(3 \times 3, \mathbb{C})$ das charakteristische Polynom $-x^3 + 5x - 4$, so hat das lineare Gleichungssystem $A \cdot x = b$ für alle $b \in \mathbb{C}^3$ eine Lösung.

Aufgabe 253: (Minimalpolynom)
Wahr oder Falsch? Begründen oder widerlegen Sie die Aussagen.

(a) Das Minimalpolynom eines Projektors $P : V \to V$ ist immer $q_P = x(x-1)$.
(b) Es gibt eine Involution $I : V \to V$ mit Minimalpolynom $q_I = x + 1$.
(c) Sei $M \in \mathrm{Mat}(n \times n, \mathbb{C})$ mit Minimalpolynom q_M und $A := aM + b\mathbb{1}_n$ für $a, b \in \mathbb{R}$. Dann ist das Minimalpolynom von A gegeben durch $q_A = a^{\deg(q_M)} q_M(\frac{x-b}{a})$.
(d) Sei $M \in \mathrm{Mat}(n \times n, \mathbb{C})$ mit Minimalpolynom q_M und $A := f(M)$ für $f \in \mathbb{C}[x]$. Dann ist das Minimalpolynom q_A von A gegeben durch $q_A = q_M(f(x))$.
(e) Sei $M \in \mathrm{Mat}(n \times n, \mathbb{C})$ mit Minimalpolynom q_M und $A := f(M)$ für $f \in \mathbb{C}[x]$. Dann ist das Minimalpolynom q_A von A gegeben durch $q_A = q_M(f(x))$.
(f) Sei $\phi \in \mathrm{End}_{\mathbb{K}}(V)$ mit Minimalpolynom $q_\phi = \sum_{k=0}^{l} a_k x^k$. Dann ist das Minimalpolynom der adjungierten Abbildung gegeben durch $q_{\phi^\dagger} = \sum_{k=0}^{l} \overline{a_k} x^k$.
(g) Zu jedem beliebigen $\mathbb{K}$-Vektorraum W und jedem Endomorphismus $\phi : W \to W$ gilt es ein Polynom $p \in \mathbb{K}[x]$ mit $p(\phi) = \mathrm{ev}_\phi(p) = 0$.
(h) Sei $A \in \mathrm{Mat}(n \times n, \mathbb{C})$ mit $A^n - \mathbb{1}_n = 0$. Dann hat das Minimalpolynom q_A den Grad n.
(i) Seien $p, q \in \mathbb{C}[X]$, so dass q und $(-1)^{\deg(p)} p$ normiert sind. Falls q ein Teiler von p ist, so gibt es eine Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{C})$ mit Minimalpolynom $q_A = q$ und charakteristischem Polynom $p_A = p$.
(j) Sei $q \in \mathbb{C}[X]$ normiert und $n \in \mathbb{N}$. Dann gibt es eine Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{C})$ mit Minimalpolynom $q_A = q$ und charakteristischem Polynom $p_A = (-1)^{n \deg(q)} q^n$ ist.
(k) Das Minimalpolynom einer Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{K})$ ist immer ein Polynom vom Grad n.
(l) Ist $A \in \mathrm{Mat}(n \times n, \mathbb{K})$ eine untere Dreiecksmatrix, so zerfällt das Minimalpolynom von A in Linearfaktoren.
(m) Ein Endomorphismus $\phi \in \mathrm{End}_{\mathbb{K}}(V)$ ist diagonalisierbar genau dann, wenn das Minimalpolynom und das charakteristische Polynom von ϕ übereinstimmen.

Aufgabe 254: (Bilinearformen und Sesquilinearformen)
Wahr oder Falsch? Begründen oder widerlegen Sie die Aussagen.

(a) Sei V ein endlich-dimensionaler reeller Vektorraum und β eine Bilinearform auf V. Dann gilt $V = U \oplus U^\perp$ für jeden Untervektorraum $U \subseteq V$.
(b) Sei V ein endlich-dimensionaler reeller Vektorraum und β ein Skalarprodukt auf V. Dann gilt $V = U \oplus U^\perp$ für jeden Untervektorraum $U \subseteq V$.
(c) Sei V ein reeller Vektorraum und β ein Skalarprodukt auf V. Dann gilt $V = U \oplus U^\perp$ für jeden Untervektorraum $U \subseteq V$.
(d) Sei V ein endlich-dimensionaler reeller Vektorraum und β eine Bilinearform auf V. Dann gilt $(U^\perp)^\perp = U$ für jeden Untervektorraum $U \subseteq V$.
(e) Sei V ein $\mathbb{K}$-Vektorraum und β eine Bilinearform auf V. Gibt es eine Basis $B = (b_1, ..., b_n)$ von V mit $\beta(v_i, v_i) > 0$ für alle $i \in \{1, ..., n\}$, so ist β positiv definit.
(f) Sei V ein endlich-dimensionaler reeller Vektorraum. Dann gibt es zu jeder alternierenden Bilinearform β auf V eine Orthogonalbasis.
(g) Eine Bilinearform auf einem komplexen Vektorraum $V \neq \{0\}$ ist nie positiv definit.
(h) Eine hermitesche Sesquilinearform β auf einem komplexen Vektorraum V ist positiv definit genau dann, wenn $\beta(v, v) \neq 0$ für alle $v \in V \setminus \{0\}$.
(i) Eine positiv semidefinite hermitesche Sesquilinearform β auf einem komplexen Vektorraum V ist positiv definit genau dann, wenn $\beta(v, v) \neq 0$ für alle $v \in V \setminus \{0\}$.
(j) Sei β eine Bi- oder Sesquilinearform auf V. Dann gilt $U^\perp \cap U = \{0\}$ für jeden Untervektorraum $U \subseteq V$.
(k) Zu der symmetrischen Bilinearform $\eta(x, y) = -x_1y_1 + \Sigma_{i=2}^n x_iy_i$ auf dem $\mathbb{R}^n$ gibt es eine Orthonormalbasis.
(l) Ist β eine symmetrische Bilinearform auf V und $v \in V \setminus \{0\}$ ein Vektor mit $\beta(v, v) = 0$, so ist β ausgeartet.
(m) Zu jeder alternierenden Bilinearform auf dem $\mathbb{R}^n$ gibt es eine Orthogonalbasis.
(n) Sei β eine symmetrische Bilinearform auf V der Signatur (p, q) mit $p, q \geq 1$. Dann gibt es einen Vektor $v \in V \setminus \{0\}$ mit $\beta(v, v) = 0$.
(o) Ist β eine symmetrische Bilinearform auf dem $\mathbb{C}^n$ und gibt es einen Vektor $v \in \mathbb{C}^n$ mit $\beta(v, v) > 0$, so gibt es auch einen Vektor $w \in \mathbb{C}^n$ mit $\beta(w, w) < 0$.
(p) Ist $A \in \mathrm{Mat}(n \times n, \mathbb{R})$ die beschreibende Matrix einer symmetrischen Bilinearform auf dem $\mathbb{R}^n$, so ist A diagonalisierbar.
(q) Ist $A \in \mathrm{Mat}(n \times n, \mathbb{R})$ symmetrisch, so ist $\eta(x, y) = x^T \cdot A \cdot y$ eine symmetrische Bilinearform auf dem $\mathbb{R}^n$.
(r) Ist $A \in \mathrm{Mat}(n \times n, \mathbb{C})$ hermitesch, so ist $\eta(x, y) = x^\dagger \cdot A \cdot y$ eine hermitesche Bilinearform auf dem $\mathbb{C}^n$.

Aufgabe 255: (unitäre und orthogonale Matrizen)
Wahr oder Falsch? Begründen oder widerlegen Sie die Aussagen.

(a) Sei $A \in \mathrm{Mat}(n \times n, \mathbb{C})$ unitär mit $\det(A) = -1$. Dann hat A den Eigenwert -1.
(b) Sei $A \in \mathrm{Mat}(n \times n, \mathbb{R})$ orthogonal mit $\det(A) = -1$. Dann hat A den Eigenwert -1.
(c) Ist $U \in \mathrm{Mat}(n \times n, \mathbb{C})$ unitär, so gilt $\det(U) = 1$.
(d) Ist $U \in \mathrm{Mat}(n \times n, \mathbb{C})$ unitär, so gilt $|\det(U)| = 1$.
(e) Sind $A, B \in \mathrm{Mat}(n \times n, \mathbb{R})$ orthogonal, so ist auch $A \cdot B$ orthogonal.
(f) Ist $U \in \mathrm{Mat}(3 \times 3, \mathbb{C})$ unitär, so hat U keine reellen Eigenwerte.
(g) Jede orthogonale Matrix $O \in \mathrm{Mat}(n \times n, \mathbb{R})$ ist diagonalisierbar.
(h) Jede unitäre Matrix ist ähnlich zu einer oberen Dreiecksmatrix.
(i) Ist $A \in \mathrm{Mat}(n \times n, \mathbb{R})$ orthogonal, so ist 0 kein Eigenwert von A.
(j) Ist $A \in \mathrm{Mat}(n \times n, \mathbb{C})$ unitär, so hat das Gleichungssystem $A \cdot x = b$ für alle $b \in \mathbb{C}^n$ genau eine Lösung.

Aufgabe 256: ((anti)selbstadjungierte Matrizen und Abbildungen)
Wahr oder Falsch? Begründen oder widerlegen Sie die Aussagen.

(a) Seien ϕ_1, ϕ_2 zwei selbstadjungierte Endomorphismen eines euklidischen oder unitären Vektorraums mit $\phi_1 \circ \phi_2 = \phi_2 \circ \phi_1$. Dann ist $\phi_1 \circ \phi_2$ selbstadjungiert.
(b) Ist $A \in \mathrm{Mat}(n \times n, \mathbb{C})$ antisymmetrisch, so ist 0 ein Eigenwert von A.
(c) Sei $A \in \mathrm{Mat}(n \times n, \mathbb{C})$ antiselbstadjungiert. Dann ist $\det(\exp(A)) = 1$.
(d) Ist $A \in \mathrm{Mat}(n \times n, \mathbb{R})$ symmetrisch, so bilden die Matrizen $M \in \mathrm{GL}(n, \mathbb{R})$ mit $M^T \cdot A \cdot M = A$ eine Gruppe.
(e) Ist $A \in \mathrm{Mat}(n \times n, \mathbb{R})$ symmetrisch mit $\mathrm{rg}(A) = n$, so bilden die Matrizen $M \in \mathrm{Mat}(n \times n, \mathbb{R})$ mit $M^T \cdot A \cdot M = A$ eine Gruppe.
(f) Die Matrizen $A \in \mathrm{GL}(n, \mathbb{C})$ mit $A^\dagger = A$ bilden eine Untergruppe von $\mathrm{GL}(n, \mathbb{C})$.
(g) Die einzigen Matrizen $A \in \mathrm{Mat}(3 \times 3, \mathbb{C})$, die gleichzeitig selbstadjungiert und unitär sind, sind die Matrizen $\mathbb{1}_3$ und $-\mathbb{1}_3$.
(h) Ist $A \in \mathrm{Mat}(n \times n, \mathbb{R})$ symmetrisch und ähnlich zu einer echten oberen Dreiecksmatrix, so gilt $A = 0$.
(i) Für jede Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{C})$ sind $A^\dagger \cdot A$ und $A \cdot A^\dagger$ hermitesch.
(j) Für jede Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{C})$ sind $\mathrm{i}A^\dagger \cdot A$ und $\mathrm{i}A \cdot A^\dagger$ antihermitesch.
(k) Zu jeder antisymmetrischen Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{R})$ gibt es eine orthogonale Matrix $O \in \mathrm{Mat}(n \times n, \mathbb{R})$, so dass $O^T \cdot A \cdot O$ diagonal ist.
(l) Jede antihermitesche Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{C})$ ist ähnlich zu einer Diagonalmatrix mit rein imaginären Diagonaleinträgen.
(m) Jede antihermitesche Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{C})$ von Rang n ist ähnlich zu einer Diagonalmatrix mit rein imaginären Diagonaleinträgen.
(n) Sind $A, B \in \mathrm{Mat}(n \times n, \mathbb{K})$ symmetrisch, so ist auch $A \cdot B$ symmetrisch.
(o) Ist $A \in \mathrm{Mat}(n \times n, \mathbb{C})$ selbstadjungiert, so ist $\det(A) \in \mathbb{R}$.
(p) Ist $A \in \mathrm{Mat}(n \times n, \mathbb{C})$ anti-selbstadjungiert, so ist $\det(A) \in \mathrm{i}\mathbb{R}$.
(q) Ist $A \in \mathrm{Mat}(n \times n, \mathbb{C})$ antihermitesch, so ist $\exp(A)$ unitär.
(r) Ist $A \in \mathrm{Mat}(n \times n, \mathbb{C})$ hermitesch, so ist auch $\exp(A)$ hermitesch.

Aufgabe 257: (normale Matrizen)
Wahr oder Falsch? Begründen oder widerlegen Sie die Aussagen.

(a) Alle diagonalisierbaren komplexen Matrizen sind normal.
(b) Alle über $\mathbb{R}$ diagonalisierbaren reellen Matrizen sind normal.
(c) Alle unitären Matrizen $A \in \mathrm{Mat}(n \times n, \mathbb{C})$ sind normal.
(d) Alle antihermitesche Matrizen $A \in \mathrm{Mat}(n \times n, \mathbb{C})$ sind normal.
(e) Alle hermiteschen Matrizen $A \in \mathrm{Mat}(n \times n, \mathbb{C})$ sind normal.
(f) Sind $A, B \in \mathrm{Mat}(n \times n, \mathbb{C})$ normal und $A \cdot B = B \cdot A$, so ist auch $A \cdot B$ normal.
(g) Ist $A \in \mathrm{Mat}(n \times n, \mathbb{R})$ normal, so ist A diagonalisierbar über $\mathbb{R}$.
(h) Eine Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{C})$ ist normal genau dann, wenn die Matrizen $A + A^\dagger$ und $A - A^\dagger$ kommutieren.
(i) Die einzige Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{C})$, die gleichzeitig normal und nilpotent ist, ist die Nullmatrix.
(j) Ist $A \in \mathrm{Mat}(n \times n, \mathbb{C})$ normal, so enthält jede Jordan-Normalform von A nur Jordanblöcke der Länge 1.

Literatur

[B] Albrecht Beutelspacher: Das ist o.B.d.A. trivial, Vieweg.

[DP] Apostolos Doxiadis, Christos H. Papadimitriou: Logicomix: An epic search for truth, Bloomsbury USA.

[B] Egbert Brieskorn: Lineare Algebra und analytische Geometrie, Vieweg.

[F] Gerd Fischer: Lineare Algebra, 16. Auflage, Vieweg+Teubner.

[G] Werner Greub: Lineare Algebra, Springer.

[L] Falko Lorenz: Lineare Algebra I und II, BI Verlag.

[Hf] Dirk W. Hoffmann, Grenzen der Mathematik, Springer Spektrum.

[Ho] Douglas R. Hofstadter: Gödel, Escher, Bach: An Eternal Golden Braid, Basic Books.

[K] Andreas Knauf, Vorlesungsskript zum Vorkurs Mathematik, Wintersemester 2013/14.

[NVk] Karl-Hermann Neeb, Vorlesungsskript Vorkurs Mathematik, WS 2010/11.

[NLA1] Karl-Hermann Neeb, Vorlesungsskript Lineare Algebra 1, WS 2010/11.

[NLA2] Karl-Hermann Neeb, Vorlesungsskript Lineare Algebra 2, SS 2011.

[P] Christoph Pöppe, Hinrichtung mit Überraschung, Spektrum der Wissenschaft, Juni 2015, S. 54–57, Spektrum Verlag.

[Si] Dirk Siefkes, Formalisieren und Beweisen: Logik für Informatiker, Vieweg.

[S] Hermann Schichl, Roland Steinbauer: Einführung in das mathematische Arbeiten, Springer.

[Sw] Christoph Schweigert, Vorlesungsskript Lineare Algebra 2005/2006 und 2007/2008

Notationsverzeichnis

$\mathbb{N}$	natürliche Zahlen (ohne Null) $\mathbb{N} = \{1, 2, 3, \ldots\}$
$\mathbb{N}_0$	natürliche Zahlen mit Null $\mathbb{N}_0 = \{0, 1, 2, \ldots\}$
$\mathbb{Z}$	ganze Zahlen $\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$
$\mathbb{R}$	reelle Zahlen
$\mathbb{C}$	komplexe Zahlen
$\mathbb{K}$	allgemeiner Körper
$\mathbb{F}_q$	endlicher Körper mit q Elementen
$\mathbb{Z}/n\mathbb{Z}$	(für $n \in \mathbb{N}$) Menge der Restklassen $\mathbb{Z}/n\mathbb{Z} = \{[0], [1], .., [n-1]\}$
$\mathbb{K}[x]$	Menge der Polynome mit Koeffizienten in einem Körper $\mathbb{K}$
$M \cup N$	Vereinigung der Mengen M und N
$M \cap N$	Schnitt der Mengen M und N
$M \setminus N$	Menge der Elemente der Menge M, die nicht in N enthalten sind
$M \times N$	kartesisches Produkt der Mengen M und N
M^n, $M^{\times n}$	(für $n \in \mathbb{N}$) n-faches kartesisches Produkt der Menge M mit sich selbst
$\sim$	äquivalent zu (Äquivalenzrelation)
$[x]$	Äquivalenzklasse von x bezüglich einer Äquivalenzrelation
$X/\sim$	Quotientenmenge bezüglich einer Äquivalenzrelation $\sim$ auf X
$\mathrm{Abb}(M, N)$	Menge der Abbildungen $f : M \to N$
$\mathrm{Bij}(M, N)$	Menge der Bijektionen $f : M \to N$
$\mathrm{im}(f)$	Bild einer Abbildung $f : M \to N$
$f(A)$	Bild einer Menge $A \subseteq M$ unter einer Abildung $f : M \to N$
$f^{-1}(B)$	Urbild einer Menge $B \subseteq N$ unter einer Abildung $f : M \to N$
S_n	(für $n \in \mathbb{N}$) symmetrische Gruppe: Gruppe der Permutationen $\pi : \{1, ..., n\} \to \{1, .., n\}$.
S^n	(für $n \in \mathbb{N}$) n-Sphäre $S^n = \{(x_1, ..., x_{n+1})^T \in \mathbb{R}^{n+1} \vert x_1^2 + \ldots + x_{n+1}^2 = 1\}$
$\mathrm{Re}\,(z)$	Realteil einer komplexen Zahl $z \in \mathbb{C}$
$\mathrm{Im}\,(z)$	Imaginärteil einer komplexen Zahl $z \in \mathbb{C}$
$\vert z\vert$	Betrag einer komplexen Zahl $z \in \mathbb{C}$
$\overline{z}$	komplex konjugiertes einer komplexen Zahl $z \in \mathbb{C}$
$\arg(z)$	Argument einer komplexen Zahl $z \in \mathbb{C}$
$\mathbb{K}^n$	(für $n \in \mathbb{N}$) der $\mathbb{K}$-Vektorraum $\mathbb{K}^n$
$\dim_{\mathbb{K}}(V)$	Dimension eines $\mathbb{K}$-Vektorraums V
$\mathrm{span}_{\mathbb{K}}(M)$	Spann einer Teilmenge $M \subseteq V$ eines $\mathbb{K}$-Vektorraums V
$\mathbb{K}v$	$\mathbb{K}v = \mathrm{span}_{\mathbb{K}}(\{v\})$ für $v \in V$ in einem $\mathbb{K}$-Vektorraum V
$U + V$	Summe der Untervektorräume $U, V \subseteq W$ eines $\mathbb{K}$-Vektorraums W
$U \oplus V$	direkte Summe der Untervektorräume $U, V \subseteq W$ eines $\mathbb{K}$-Vektorraums W
$\mathrm{Hom}_{\mathbb{K}}(U, V)$	Vektorraum der $\mathbb{K}$-linearen Abbildungen $\phi : U \to V$ für $\mathbb{K}$-Vektorräume U, V
$\mathrm{End}_{\mathbb{K}}(V)$	Algebra der $\mathbb{K}$-linearen Abbildungen $\phi : V \to V$ für einen $\mathbb{K}$-Vektorraum V
$\ker(\phi)$	Kern einer linearen Abbildung $\phi : U \to V$
$\mathrm{im}(\phi)$	Bild einer linearen Abbildung $\phi : U \to V$
$\mathrm{rg}(\phi)$, $\mathrm{rg}(A)$	Rang einer linearen Abbildung $\phi : U \to V$, einer Matrix A
$\mathrm{def}(\phi)$, $\mathrm{def}(A)$	Defekt einer linearen Abbildung $\phi : U \to V$, einer Matrix A
$\mathrm{Mat}(m \times n, \mathbb{K})$	Menge der $(m \times n)$-Matrizen mit Einträgen in $\mathbb{K}$
$\mathrm{GL}(n, \mathbb{K})$	Menge der invertierbaren $(n \times n)$-Matrizen mit Einträgen in $\mathbb{K}$
E_{ji}	Elementarmatrizen mit Eintrag 1 in der $j$ten Zeile und $i$ten Spalte und 0 sonst
A^T	Transponierte einer Matrix A
V^*	Dualraum eines $\mathbb{K}$-Vektorraums V

ϕ^*	duale Abbildung zu linearer Abbildung ϕ
sgn	Signum einer Permutation σ
σ_{ij}	elementare Vertauschungen $\sigma_{ij} \in S_n$
A_n	alternierende Gruppe
$L(V^{\times n}, W)$	Vektorraum der n-linearen Abbildungen $\phi : V \to W$
$S(V^{\times n}, W)$	Vektorraum der symmetrischen n-linearen Abbildungen $\phi : V \to W$
$A(V^{\times n}, W)$	Vektorraum der alternierenden n-linearen Abbildungen $\phi : V \to W$
vol_n	Volumenform auf dem $\mathbb{R}^n$
det	Determinante
$\mathrm{Aut}^+_{\mathbb{R}}(V)$	Gruppe der orientierungserhaltenden Automorphismen eines reellen Vektorraums V
$\mathrm{GL}(n, \mathbb{R})^+$	Gruppe der reellen $n \times n$-Matrizen mit positiver Determinante
$\mathrm{SL}(n, \mathbb{R})$	Gruppe der reellen $n \times n$-Matrizen mit Determinante 1
A^{str}_{ij}	ijte Streichmatrix
$\mathrm{ad}(A)$	adjungierte Matrix
$E(\phi, \lambda)$	Eigenraum von ϕ zum Eigenwert λ
$g(\phi, \lambda)$	geometrische Vielfachheit des Eigenwert λ von ϕ
$H(\phi, \lambda)$	Hauptraum von ϕ zum Eigenwert λ
$a(\phi, \lambda)$	algebraische Vielfachheit des Eigenwert λ von ϕ
$\mathbb{K}[x]$	Polynomalgebra über dem Körper $\mathbb{K}$
p_ϕ	charakteristisches Polynom von ϕ
q_ϕ	Minimalpolynom von ϕ
$J_k(\lambda)$	Jordanblock der Länge k zum Eigenwert λ
$n(k, \lambda)$	Anzahl der Jordanblöcke der Länge k zum Eigenwert λ
$\langle\, ,\, \rangle$	Skalarprodukt auf reellem oder komplexem Vektorraum
$\|\; \|$	Norm auf reellem oder komplexem Vektorraum, Länge eines Vektors
$\mathrm{U}(V, \beta)$	unitäre Gruppe eines $\mathbb{K}$-Vektorraums V mit Bi- oder Sesquilinearform β
$\mathrm{U}(n)$, $\mathrm{U}(n, \mathbb{C})$	unitäre Gruppe $\{U \in \mathrm{Mat}(n \times n, \mathbb{C}) \,\vert\, U^\dagger \cdot U = \mathbb{1}_n\}$
$\mathrm{O}(V, \beta)$	orthogonale Gruppe eines reellen Vektorraums V mit symmetrischer Bilinearform β
$\mathrm{O}(n)$, $\mathrm{O}(n, \mathbb{R})$	orthogonale Gruppe $\{O \in \mathrm{Mat}(n \times n, \mathbb{R}) \,\vert\, O^T \cdot O = \mathbb{1}_n\}$
$\mathrm{Sp}(V, \beta)$	symplektische Gruppe eines $\mathbb{K}$-Vektorraums V mit alternierender Bilinearform β
$\mathrm{SU}(V, \beta)$	Untergruppe $\{\phi \in \mathrm{U}(V, \beta) \vert\ \det(\phi) = 1\}$
$\mathrm{SO}(V, \beta)$	Untergruppe $\{\phi \in \mathrm{U}(V, \beta) \vert\ \det(\phi) = 1\}$
$\mathrm{SU}(n)$	Untergruppe $\{M \in \mathrm{U}(n) \vert\ \det(M) = 1\}$
$\mathrm{SO}(n)$	Untergruppe $\{M \in \mathrm{O}(n) \vert\ \det(M) = 1\}$
$A^\dagger$	hermitesch konjugierte Matrix einer Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{C})$
$\phi^\dagger$	Adjungierte eines Endomorphismus ϕ
$M^\perp$	orthogonales Komplement von M
${}^B\beta^B$	beschreibende Matrix einer Bi- oder Sesquilinearform β bezüglich einer Basis B
V/U	Quotientenraum bezüglich eines Untervektorraums $U \subseteq V$
$\langle I \rangle_{\mathbb{K}}$	von Menge I frei erzeugter $\mathbb{K}$-Vektorraum
$\oplus_{i \in I} V_i$	direkte Summe einer Familie $(V_i)_{i \in I}$ von Vektorräumen
$\Pi_{i \in I} V_i$	Produkt einer Familie $(V_i)_{i \in I}$ von Vektorräumen
$V \otimes_{\mathbb{K}} W$	Tensorprodukt von $\mathbb{K}$-Vektorräumen V, W
$T(V)$	Tensoralgebra eines Vektorraus V

Index