FSU Jena
Fakultät für Mathematik und Informatik

Lineare Algebra für *-Informatik

FMI-MA0022
Wintersemester 2020/21

Simon King
12. Februar 2021

Inhaltsverzeichnis

Dies sind Vorlesungsnotizen, kein voll ausgearbeitetes Skript. Weitere Erklärungen und Details von Beispielen gibt es nur in der Vorlesung.

Literaturhinweise: Es gibt zahlreiche Lehrbücher zur linearen Algebra in verschiedenen Sprachen. Ich empfehle Ihnen daher, verschiedene Bücher zu lesen, um zu sehen, mit welchen Sie am besten zurecht kommen. Hier sind zwei in deutscher Sprache erschienene Standardwerke:

- Gerd Fischer, „Lineare Algebra“, vieweg+teubner
- Hans-Joachim Kowalski, „Lineare Algebra“, Walter de Gruyter

-1 Vorrede

Aufgabe von Schulen

Allgemeinbildende Schulen haben die Aufgabe, den Schülerinnen und Schülern eine *praktische* Einführung in die jeweiligen Fachmethoden zu vermitteln; mit diesen wird ein Einblick in die Entwicklung der Fachkenntnisse bis einschließlich der letzten 200 Jahre erarbeitet; und die Fachkenntnisse werden mit einigermaßen modernen Begriffsbildungen beschrieben.

Das gilt grundsätzlich für alle Schulformen und mit einer Ausnahme auch für alle Fächer, wie ein Blick in die Lehrpläne zeigt:

- An Grundschulen lernt man im Kunstunterricht, „Farben hinsichtlich ihrer Symbolwirkung und ihres Symbolwertes ... in der eigenen Gestaltungsarbeit anzuwenden“.

- An Regelschulen lernt man im Chemieunterricht das Periodensystem der Elemente kennen (Mitte des 19. Jahrhunderts entwickelt).

- An Realschulen schreibt man im Deutschunterricht Aufsätze und lernt verschiedene Textarten und ihre Funktion kennen.

- An Gymnasien lernt man im Chemieunterricht nicht etwa das aus der Alchemie stammende Konzept der „Ursäure“, sondern lernt die Säure-Base-Konzepte von Brønsted sowie von Lewis.

Die einzige Ausnahme ist Mathematik. Ich werde Ihnen zu den meisten in der Vorlesung behandelten Themen Jahreszahlen nennen, an denen Sie merken werden, dass der Schulunterricht nur höchst selten den Kenntnisstand des frühern 19. Jahrhunderts erreicht. Das ist noch vergleichsweise harmlos, denn die Fachmethoden werden noch nicht einmal auf dem Stand der griechischen Antike praktisch vermittelt. Grundbegriffe der Algebra fehlen an heutigen deutschen Schulen meist völlig. Grundbegriffe aus der Analysis werden meist so wie im 17. Jahrhundert behandelt. Zum Vergleich: Im 17. Jahrhundert trennte sich gerade erst die Chemie von der Alchemie.

Das ist einerseits ein kulturpolitischer Skandal, andererseits ein praktisches Problem für ein Studium von MINT-Fächern. Selbstverständlich haben mathematische Vorlesungen an Hochschulen ein vergleichbares Niveau wie Vorlesungen zu allen anderen Themen. Lediglich der Unterschied zum Schulniveau ist in Mathematik größer.

Inzwischen sind in manchen Bundesländern, etwa in Nordrhein-Westfalen, Deutsch- und Mathematikkurse für Lehramtsstudierende aller Schulformen und aller Fächer verpflichtend. Als gebildeter Mensch sollte man hoffen, dass davon auch etwas an den Schulen ankommt.

Arbeitsweise an Hochschulen

Das Unterrichtstempo an Hochschulen ist generell (nicht nur in Mathematik) höher als an Schulen. Ein Thema wird kurz in der Vorlesung behandelt und dann obliegt es den Studierenden, die in der Vorlesung behandelten Methoden **selbständig** einzuüben und mit Hilfe von Literatur das Detailverständnis zu vertiefen.

Zur Arbeitseinstellung an Hochschulen gehört auch die Bereitschaft, Dinge in Frage zu stellen. Fragem können an den Text gerichtet sein („Ist das Argument korrekt?“), an sich selbst („Verstehe ich die Methode wirklich?“), und natürlich auch an die Lehrpersonen („In Buch XY wird zur Lösung dieser Problemstellung eine andere Methode als in der Vorlesung verwendet; worin sehen Sie die Vor- und Nachteile beider Methoden?“).

Dieses In-Frage-Stellen sollte mit einer wohlwollenden Grundeinstellung erfolgen. Das heißt, man sollte nicht mit der Attitüde herangehen, beim kleinsten Fehler gleich alles zu verwerfen, sondern man sollte sich bewusst machen, welche Teile des Ganzen dem In-Frage-Stellen standhalten, und sich bemühen, die anderen Teile zu berichtigen.

Dazu ein extremes Beispiel: In seinem epochalen Werk „Elemente“ erhob Euklid von Alexandria [ca. 300 v.Chr] den Anspruch, zu Beginn des Buches alle verwendeten Definitionen, geometrischen Grundkonstruktionen (Postulate) und Grundtatsachen (Axiome) vollständig aufzulisten und nur diese in den Beweisen zu verwenden. Doch manche Definitionen waren gar keine (etwa: „Ein Punkt ist etwas, das keine Teile hat“ und „Eine gerade Linie ist eine solche, die zu den Punkten auf ihr gleichmäßig liegt.“) und bereits im Beweis seines ersten Satzes („Man kann über jeder gegebenen Strecke mit Zirkel und Lineal ein gleichseitiges Dreieck errichten.“) verwendete er implizite Zusatzannahmen, die er nicht als Axiom formulierte.

Erst mehr als 2000 Jahre später gelang es David Hilbert [1862–1943], die Lücken in Euklids Werk zu schließen — wodurch jedoch die herausragende Bedeutung Euklids nicht geschmälert wird! Die von ihm formulierten Ergebnisse waren korrekt, und auch wenn Euklid selbst seinen hohen Ansprüchen in den Beweisen nicht vollumfänglich gerecht wurde, war es für die Mathematik wegweisend, solche Ansprüche zu stellen. Hilbert vollendete Euklids ursprüngliches Vorhaben, doch „Vollendung“ ist nicht das Ende, denn Hilberts Werk hat weitere Entwicklungen der Geometrie angestoßen.

Mathematisches Grundstudium

Im Zentrum von Anfängervorlesungen der Mathematik stehen Lineare Algebra einerseits und Analysis andererseits. Wenn man beide nicht im gleichen Semester lernt, hat man meistens die Lineare Algebra *vor* der Analysis. Der Grund ist, dass die Lineare Algebra von den Begriffen her leichter als die Analysis ist. Die

Lineare Algebra verallgemeinert nämlich das Lösen linearer Gleichungssysteme, indem die dort auftretenden Phänomene begrifflich erfasst werden; dabei kommt man bereits mit Grundrechenarten sehr weit und kann viele Probleme mit leichten Verfahren lösen.
In der Analysis geht es hingegen um Folgen, Reihen, Grenzwerte, Ableitungen, Integrale etc., deren explizite Berechnung oft nur sehr schwer gelingt oder gar beweisbar unmöglich ist.

Dennoch haben nicht wenige Anfänger*innen das Gefühl, mit Analysis besser zurecht zu kommen, denn erstens wird das Lösen linearer Gleichungssysteme heute nicht mehr in nennenswertem Umfang an Schulen behandelt und zweitens kommen wie erwähnt zwar Grundbegriffe der Analysis, aber nicht der Algebra im Lehrplan vor. Doch tatsächlich halte ich dieses Gefühl für unangebracht. Es ist nämlich meines Erachtens leichter, das Begriffssystem der Algebra neu zu lernen, als das wie erwähnt *auf dem theoretisch unzulänglichen Stand des 17. Jahrhunderts* gelernte Begriffssystem der Analysis im Kopf zu reparieren.

Rechnerische Anforderungen

Wenn nicht ausdrücklich etwas anderes verlangt wird, sollen Sie exakt rechnen, also mit Brüchen und Wurzeln. **Gerundete Kommazahlen** sind viel zu kompliziert! Das mag Sie überraschen.

Tatsächlich gelten viele nützliche Rechengesetze, etwa das Assoziativgesetz $a \cdot (b \cdot c) = (a \cdot b) \cdot c)$ bei gerundetem Rechnen nicht mehr. Die Auswirkung von Rundungsfehlern wird in der ***Numerik*** untersucht. Für eine Rundungsfehleranalyse braucht man Vorkenntnisse aus der linearen Algebra und vor allem aus der Analysis.

0 Einige mathematische Grundlagen

In den meisten an Hochschulen unterrichteten Fächern gehört es zu den Fachmethoden, jemanden durch „schlüssiges Argumentieren“ von der Gültigkeit eines Sachverhalts zu überzeugen. Die einzelnen Fächer unterscheiden sich darin, was man konkret unter einem „schlüssigen Argument“ zu verstehen hat und was als gültiger Ausgangspunkt einer Argumentation gilt. In der Mathematik bedeutet „schlüssiges Argumentieren“ meist eine Abfolge ***deduktiver*** Argumente, für die man drei Zutaten benötigt:

- Feststellung eines bekannten Sachverhaltes (***Voraussetzungen***),
- Anwendung einer ***Schlussregel*** auf den festgestellten Sachverhalt,
- Formulierung der so entstehenden ***Schlussfolgerung***.

Das Alleinstellungsmerkmal der Mathematik ist, das Konzept des deduktiven Argumentierens konsequent zu Ende zu denken und zu formalisieren. Und dadurch werden aus Argumentationsketten *Beweise*. Die Formalisierung erfordert natürlich eine Formelsprache, die wir in diesem Kapitel entwickeln.

Zwar ist mathematisches Denken spätestens seit dem Neolithikum (ab 10 000 v.Chr. im Fruchtbaren Halbmond) archäologisch belegbar, doch ausformulierte Beweise kennt man erst später. In Babylon war zum Beispiel der Satz des Pythagoras empirisch im 2. Jtsd. v.Chr. bekannt, aber die ersten Beweise werden Thales von Milet [ca. 624–547 v.Chr.] zugeschrieben.

Auf Euklid von Alexandria [ca. 300 v.Chr] geht das Konzept des ***axiomatischen Beweises*** zurück. Seine Idee war, die benötigten Grundvoraussetzungen (***Axiome*** und ***Postulate***), Definitionen und Schlussregeln nicht erst während des Beweises aus der Luft zu greifen, sondern alles ganz zu Beginn vollständig aufzulisten. Daraus resultierte sein Hauptwerk, die „Elemente“, welches bis ins 20. Jhdt. hinein als Schulbuch gebräuchlich war.

Sie sollten sich klar machen, dass jeder Beweis ein Beispiel ist, nämlich ein Beispiel dafür, wie man mit den haptsächlichen Inhalten der Mathematik (den Begriffen!) umgeht. Darüber hinaus geht aus vielen Beweisen direkt hervor, wie ein bestimmter Aufgabentyp praktisch zu lösen ist.[1] Derartige Beweise sind also das allgemeinstmögliche Beispiel.

0.1 Logisch!

Aussagen

Die klassische zweiwertige ***Aussagenlogik*** geht bis auf Aristoteles [384–322 v.Chr] und Philon von Megara [ca. 4./3. Jhdt. v.Chr.] zurück. Formalisierungen der Aus-

[1] In Anlehnung an Euklid nenne ich derartige Ergebnisse „Problem“, während reine Sachverhaltsbeschreibungen als „Satz“ oder „Theorem“ bezeichnet werden.

sagenlogik schufen George Boole [1847], Gottlob Frege [1879] und Bertrand Russel [1908], allerdings werden diese hier keine Rolle spielen. Unter einer ***Aussage*** verstehen wir etwas, was wahr (W) oder falsch (F) sein kann. Was nicht wahr oder falsch sein kann (zum Beispiel eine Frage), ist keine Aussage. Wir benutzen Großbuchstaben $P, Q, ...$ als Platzhalter für Aussagen. Man kann aus gegebenen Aussagen P, Q neue Aussagen bilden:

Negation: $\neg P$, „nicht P“.

Konjunktion: $P \wedge Q$, „P und Q“.

Disjunktion: $P \vee Q$, „P oder Q“.

Implikation: $P \Rightarrow Q$, „Wenn P, dann Q“.

Äquivalenz: $P \Leftrightarrow Q$, „P genau dann wenn Q“.

Die Bedeutung der Zusammensetzung ist jeweils durch eine ***Wahrheitstafeln*** gegeben; dies geht bis auf Philon von Megara [ca. 4.-3. Jhdt. v.Chr.] zurück.

P	Q	$\neg P$	$P \wedge Q$	$P \vee Q$	$P \Rightarrow Q$	$P \Leftrightarrow Q$
F	F	W	F	F	W	W
F	W	W	F	W	W	F
W	F	F	F	W	F	F
W	W	F	W	W	W	W

Beim Rechnen gilt, wie Sie wissen, „Punkt- vor Strichrechnung“, es sei denn, es werden Klammern verwendet. Ähnliches gilt für die Logik: Es ist relativ üblich, dass $\neg$ vor $\wedge$, $\wedge$ vor $\vee$, $\vee$ vor $\Rightarrow$ und $\Rightarrow$ vor $\iff$ geht.

Beispiel Für Aussagen P, Q, R, S ist $P \vee Q \wedge \neg R \Rightarrow S$ als $(P \vee (Q \wedge (\neg R))) \Rightarrow S$ zu verstehen.

Jedoch ist diese Vorrangsregelung tendentiell verwirrend. Setzen Sie also im Zweifel lieber ein Klammerpaar zu viel.

Typische Missverständnisse treten auf, wenn man den Alltagsgebrauch obiger Konstruktionen mit ihrer logischen Definition (Wahrheitstafeln) verwechselt. Vermeiden Sie dies bitte! Die obigen Konstruktionen können übrigens *ausschließlich* auf Aussagen angewandt werden. In Klausuren las ich schon Dinge wie $\{1, 2, 3\} \Rightarrow 3$, doch das ist schlimmer als falsch: Es ist unsinnig.

Gleichbedeutende Aussagen

Haben zwei zusammengesetzte Aussagen P, Q die gleiche Wahrheitstafel, so bezeichnet man sie als ***gleichbedeutend***. In der Wahrheitstafel der Aussage $P \iff Q$ sind dann alle Einträge W, das heißt, die Aussage ist dann stets wahr. Eine immer wahre Aussage heißt auch ***Tautologie***[2]. Ich folge hier übrigens der Sprechweise, die auch in der Vorlesung über Diskrete Strukturen verwendet wird, obwohl von der griechischen Wortbedeutung her eine Tautologie eigentlich keine allwahre Aussage, sondern ein Paar gleichbedeutender Aussagen wäre.

Lemma 0.1
Seien P, Q Aussagen. Folgende Aussagenpaare sind gleichbedeutend:

a) P, $\neg(\neg P)$

b) $\neg(P \Rightarrow Q)$, $P \wedge (\neg Q)$

c) $\neg(P \wedge Q)$, $(\neg P) \vee (\neg Q)$

d) $\neg(P \vee Q)$, $(\neg P) \wedge (\neg Q)$

Die vier eben formulierten Regeln sind nützlich beim fehlerfreien Negieren von Aussagen. Die letzten beiden nennt man **de Morgansche Regeln**. Der Beweis erfolgt durch Hinschreiben der Wahrheitstafeln und wird Ihnen teilweise zur Übung überlassen.

Aussageformen und Quantoren

Die Aussagenlogik ist keine sehr ausdrucksstarke Sprache. Man erweitert sie daher zur ***Prädikatenlogik***. Ein n-stelliges ***Prädikat*** (auch: ***Aussageform***) ist ein Ausdruck n Variablen, aus dem durch Einsetzen konkreter Werte eine Aussage entsteht. Beispiel: Aus dem einstelligen Prädikat $x^2 > 1$ entsteht durch Einsetzen $2^2 > 1$ (wahre Aussage) und $\left(\frac{1}{2}\right)^2 > 1$ (falsche Aussage).

Auf die Variablen einer Aussageform kann man auch ***Quantoren***[3] anwenden, statt konkrete Werte einzusetzen. Intuitiv wurden Quantoren bereits im 4. Jhdt. v.Chr. verwendet, die Bezeichnung „Quantor“ und die heutigen Notationen stammen jedoch aus dem 19. und 20. Jahrhundert. Wir verwenden drei Quantoren[4]:

$\forall$ „für alle“
: Bsp: „$\forall x \in \mathbb{R}\colon x^2 \geq 0$“ ist die (wahre) Aussage, dass Quadrate reeller Zahlen größer oder gleich Null sind.

[2] Übersetzt aus dem Griechischen heißt das „dieselbe Aussage“

[3] Man sieht an den Beispielen, dass bisweilen der Wertebereich eines Quantors eingeschränkt wird. Die auftretenden Mengen sollten aus der Schule bekannt sein. Mehr über Mengen gibt es im nächsten Abschnitt.

[4] Weitere übliche Notation (siehe „Diskrete Strukturen“): $\bigvee x$ statt $\exists x$; $\bigwedge x$ statt $\forall x$.

$\exists$ „es gibt (mindestens) ein“
Bsp: „$\exists x \in \mathbb{Q}\colon x^2 = 2$“ ist die (falsche) Aussage, dass $\sqrt{2}$ rational ist.

$\exists!$ „es gibt genau ein“
Bsp: „$\exists! x \in \mathbb{Z}\colon x^2 = y$“ ist eine Aussageform mit Variable y, die für $y = 0$ wahr ist, aber für $y = 3$ bzw. $y = 4$ falsch ist, denn dort gibt es keine bzw. zwei Möglichkeiten für x.

Bemerkung 0.2
Die Variablen von Quantoren haben einen **Gültigkeitsbereich**, *den man im Zweifelsfall durch Klammern verdeutlichen sollte. So ist* $(\exists x \in \mathbb{R}\colon x > 0) \wedge (\exists x \in \mathbb{R}\colon x < 0)$ *wahr, denn in den beiden Teilen der Konjunktion bezeichnet* x *zwei* verschiedene *Variablen, die einfach „zufällig“ denselben Namen tragen. Hingegen ist* $\exists x \in \mathbb{R}\colon (x > 0) \wedge (x < 0)$ *falsch, denn diesmal ist der Gültigkeitsbereich des Existenzquantors die* ganze *Aussage.*

Vereinfacht kann man sich vorstellen, dass man einen logischen Ausdruck von links nach rechts liest und jeder Quantor wie in einem Computerprogramm einen Variablennamen neu deklariert oder, wenn der Name vorher schon gebraucht wurde, die alte Deklaration überschreibt.

Natürlich ist es schlechter Stil, verschiedene Variablen in verschiedenen Teilen einer Aussage gleich zu nennen — aber möglich *ist es!*

Hier sind einige Regeln für die Umformung von Ausdrücken mit Quantoren. Wohlgemerkt sind diese Axiome weder vollständig noch minimal (die Aussagen für $\exists$ folgen aus den Aussagen für $\forall$ und den obigen Wahrheitstafeln).

Axiom 0.3
Seien $P(x), Q(x, y), R(x)$ *Aussageformen.*

a) Vertauschung: $\forall x\colon \forall y\colon Q(x, y)$ *ist gleichbedeutend zu* $\forall y\colon \forall x\colon Q(x, y)$.
Ebenso ist $\exists x\colon \exists y\colon Q(x, y)$ *gleichbedeutend zu* $\exists y\colon \exists x\colon Q(x, y)$.
Kurzschreibweise: $\forall x, y\colon Q(x, y)$ *bzw.* $\exists x, y\colon Q(x, y)$.

b) Negation: $\neg(\forall x\colon P(x))$ *ist gleichbedeutend zu* $\exists x\colon \neg P(x)$.
Ebenso ist $\neg(\exists x\colon P(x))$ *gleichbedeutend zu* $\forall x\colon \neg P(x)$.

c) Distribution:
$(\forall x\colon P(x)) \wedge (\forall y\colon R(y))$ *ist gleichbedeutend zu* $\forall x\colon P(x) \wedge R(x)$.
Ebenso ist $(\exists x\colon P(x)) \vee (\exists y\colon R(y))$ *gleichbedeutend zu* $\exists x\colon P(x) \vee R(x)$.

Beispiel 0.4
a) $\forall x \in \mathbb{R}\colon \exists y \in \mathbb{R}\colon x + y = 0$ *ist wahr, aber* $\exists y \in \mathbb{R}\colon \forall x \in \mathbb{R}\colon x + y = 0$ *ist falsch. Nur gleichartige Quantoren können vertauscht werden!*

b) Vielleicht denken Sie, die Negation „Bei Nacht sind alle Katzen grau“ sei „Es gibt eine Katze, die nachts nicht grau ist“. Doch dabei hätten Sie stillschweigend angenommen, dass es manchmal Nacht ist: Obige Aussagen sind beide wahr, wenn es zwar Katzen, aber keine Nächte gibt. Bei der Analyse natürlicher Sprache sind Zusatzannahmen über das Weltwissen unverzichtbar, doch in der Mathematik sind sie unzulässig.

Die Aussage ist „$\forall t\colon\ N(t) \Rightarrow (\forall x\colon\ K(x,t) \Rightarrow g(x,t))$“, mit den Prädikaten $N(t)$ („zur Zeit t ist Nacht“), $K(x,t)$ („x ist zur Zeit t eine Katze“ — man muss prinzipiell auch mit Gestaltwandlern rechnen!) und $g(x,t)$ („x ist zur Zeit t grau“). Gleichbedeutend sind:

$$
\begin{aligned}
&\neg\,(\forall t\colon\ (N(t) \Rightarrow (\forall x\colon\ K(x,t) \Rightarrow g(x,t)))) && \\
&\quad \exists t\colon\ \neg\,(N(t) \Rightarrow (\forall x\colon\ K(x,t) \Rightarrow g(x,t))) && \text{nach Axiom 0.3.b)} \\
&\quad \exists t\colon\ (N(t) \wedge \neg\,(\forall x\colon\ K(x,t) \Rightarrow g(x,t))) && \text{nach Lemma 0.1.b)} \\
&\quad\ \exists t\colon\ (N(t) \wedge (\exists x\colon\ K(x,t) \wedge \neg g(x,t))) && \text{nach Ax. 0.3.b), Lem 0.1.b)}
\end{aligned}
$$

„Es gibt eine Nacht, in der es eine Katze gibt, die nicht grau ist.“

Beweisstrategieen Aus den Regeln der Aussagen- und Prädikatenlogik ergeben sich auch Strategieen für den Beweis von Aussagen. Beispielsweise kann man nachweisen, dass die Bedigungen, unter denen eine Aussage falsch würde, nicht eintreten kann. In der folgenden Zusammenstellung seien P, Q Aussagen bzw. Prädikate und M sei eine Menge (dazu mehr im nächsten Abschnitt).

$P \Rightarrow Q$: Zum Beweis einer Implikation bestehen folgende Möglichkeiten

a) ***Direkter Beweis***: Man beginnt mit der Voraussetzung, dass P gilt (würde nämlich P nicht gelten, wäre ja $P \Rightarrow Q$ wahr), und leitet daraus her, dass Q gilt.
Eine Herleitung ist übrigens oft konstruktiv, d.h. man kann direkte Beweise oft rechnerisch nachvollziehen.

b) ***Beweis durch Kontraposition***: Man ersetzt $P \Rightarrow Q$ durch die gleichbedeutende Aussage $\neg Q \Rightarrow \neg P$ und beweist diese direkt. Aus der Voraussetzung, dass Q nicht gilt, leitet man her, dass P nicht gilt.

c) ***Widerspruchsbeweis***: Man ersetzt $P \Rightarrow Q$ durch die gleichbedeutende Aussage $\neg(P \wedge \neg Q)$. Man weist also nach, dass nicht gleichzeitig P wahr und Q falsch sein kann. Anders gesagt: Man setzt zunächst P voraus; dann nimmt man an, dass Q falsch ist, und leitet aus der Voraussetzung und der Annahme einen Widerspruch (also F) her.
Widerspruchsbeweise sind meist nicht konstruktiv. Man sollte sie nach Möglichkeit vermeiden (aber es ist nicht immer möglich).

$\forall x \in M\colon\ P(x)$: Das ist eine Implikation mit einer Variable x, nämlich $\forall x\colon\ (x \in M \Rightarrow P(x))$. Es ist zu zeigen, dass die Implikation für alle x erfüllt ist. Dafür gibt es die oben genannten Möglichkeiten. Man könnte etwa unter der Voraussetzung, dass $P(x)$ falsch ist, herleiten, dass $x \notin M$ (Kontraposition).

$\exists x \in M\colon\ P(x)$: Logisch ist dies die Aussage $\exists x\colon\ (x \in M \wedge P(x))$.

a) Man könnte ein $x \in M$ mit $P(x)$ explizit (konstruktiv!) angeben.

b) Manchmal funktioniert wieder nur ein inkonstruktiver Ansatz. Man versucht zu beweisen, dass die Negation von $\exists x \in M\colon\ P(x)$ falsch ist (denn Q und $\neg\neg Q$ sind gleichbedeutend). Man würde also die Annahme $\forall x \in M\colon\ \neg P(x)$ zum Widerspruch führen.
Weist man die Existenz eines $x \in M$ mit $P(x)$ auf diese Weise nach, hat man unter Umständen nicht den geringsten Anhaltspunkt, wie man ein solches x tatsächlich finden kann!

Beispiel 0.5
Man beweise $\sqrt{2} \notin \mathbb{Q}$ (diese Aussage war schon den Pythagoräern bekannt, wurde allerdings geheim gehalten). Nun, zunächst einmal muss man die Definition von $\sqrt{2}$ einsetzen: Es handelt sich um $x \in \mathbb{R}$ mit $x \geq 0$ und $x^2 = 2$. Dann muss man die Definition von $\mathbb{Q}$ einsetzen: Es handelt sich um die Menge aller Brüche $\frac{m}{n}$ mit $m, n \in \mathbb{Z}$ und $n \neq 0$. Wie aus der Schule bekannt, kann man m, n hierbei teilerfremd wählen.

Die zu beweisende Aussage wurde also umgeformt zu $\forall m, n \in \mathbb{Z}$ mit $n \neq 0$ gilt: Wenn m, n teilerfremd, dann $\frac{m^2}{n^2} \neq 2$.

Wir führen einen Widerspruchsbeweis. Es seien also $m, n \in \mathbb{Z}$ teilerfremd mit $n \neq 0$ und nehmen an, dass $\frac{m^2}{n^2} = 2$. Dann $m^2 = n^2 \cdot 2$. Weil 2 eine Primzahl ist, folgt daraus m gerade, d.h. $\exists k \in \mathbb{Z}$ mit $m = 2\,k$. In der Primfaktorzerlegung von m^2 tritt 2 also mindestens zweimal auf, also auch in der Primfaktorzerlegung von $n^2 \cdot 2$. Also tritt die 2 mindestens einmal *in der Primfaktorzerlegung von n^2, also auch von n, auf. Widerspruch, denn m, n sollten teilerfremd sein. Die Annahme war also falsch, daher $\frac{m^2}{n^2} \neq 2$.* □

0.2 Mengen

Was ist eine „Menge“? In der Schule haben Sie vermutlich eine Vorstellung davon entwickelt (eine Menge ist etwas, was andere Dinge enthält), aber eine umfassende Behandlung der Mengenlehre wäre erst nach einigen Jahren Mathematikstudium möglich. Eine intuitive Vorstellung des Mengenbegriffs gibt folgendes Zitat von Georg Cantor [1845–1918], dem Begründer der Mengenlehre:

> *Unter einer ‚Menge‘ verstehen wir jede Zusammenfassung von bestimmten wohlunterschiedenen Objekten unserer Anschauung oder unseres Denkens zu einem Ganzen.* (G. Cantor, 1895)

Mit dieser Vorstellung kommt man in der Praxis recht weit, doch sobald man die Mengenlehre auf ein solideres (also axiomatisches) Fundament stellen möchte, zeigt sich, dass man gewisse Zusammenfassungen nicht als Menge auffassen kann.[5]

Statt Grundlagenfragen zu erörtern, stellen wir uns auf einen konstruktiven Standpunkt: Es gibt gewisse Mengen, die Sie bereits aus der Schule kennen (dazu unten mehr), und es gibt gewisse Konstruktionen, mit denen man aus bereits bekannten Mengen weitere Mengen gewinnt. Dass das alles funktioniert, stellen Axiome sicher, die wir hier nicht formulieren.

Mengen bestehen aus Elementen. Die Aussage $x \in M$ heißt „x ist ein Element der Menge M“. Statt $\neg(x \in M)$ schreibt man $x \notin M$. Beispielsweise ist $1 \in \mathbb{Z}$ und $\frac{1}{2} \notin \mathbb{Z}$. Folgendes Axiom ist inhaltlich nahe liegend. Weniger nahe liegend ist, dass man es als *Axiom* und nicht als *Definition* formuliert.

Axiom 0.6 (Extensionalitätsaxiom)
Zwei Mengen sind **gleich** *gdw. sie die gleichen Elemente haben.*

Definition 0.7
Eine Menge N heißt eine **Teilmenge** *einer Menge M :⇔ jedes Element von N ist auch ein Element aus M. Bezeichnung: $N \subseteq M$.*

Gleichbedeutend zu $N \subseteq M$ verwende ich meist einfach $N \subset M$. Kurzschreibweise: $N \subsetneq M$ (***echte Teilmenge***) heißt $(N \subset M) \wedge (N \neq M)$.

Beobachtung 0.8
Seien M, N Mengen. $M = N \iff (M \subseteq N \wedge N \subseteq M)$.

Definition 0.9
Sind M, N Mengen, so kann man folgende neue Mengen bilden:

a) Der **Schnitt** *(auch* **Durchschnitt** *oder* **Schnittmenge***) $M \cap N$ besteht aus allen Elementen von M, die zugleich Elemente von N sind.*

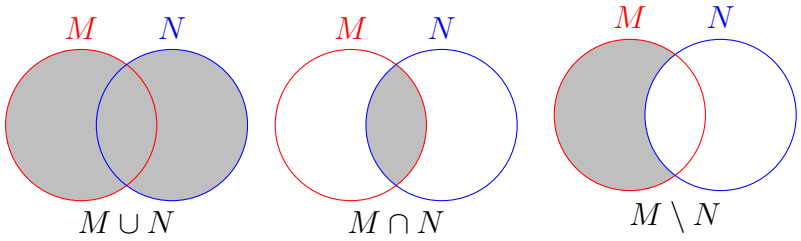
b) Die **Differenz** *$M \setminus N$ besteht aus allen Elementen von M, die zugleich* nicht *Elemente von N sind.*

c) Die **Vereinigung** *$M \cup N$ besteht aus den Elementen von M zusammen mit den Elementen von N (basiert auf* **Vereinigungsaxiom***)*

Die eben genannten Konstruktionen kann man gut mit so genannten ***Venn-***

[5] Die Gesamtheit aller Mengen bildet zum Beispiel *keine* Menge!

Diagrammen[6] veranschaulichen:



Venn-Diagramme sind allerdings nicht mehr als eine Veranschaulichung. Einen Beweis ersetzen sie nicht, aber die Veranschaulichung kann eine Idee für einen Beweis liefern. Etwas formaler kann man schreiben:

$$\begin{aligned} M \cup N &:= \{x \mid (x \in M) \vee (x \in N)\} \\ M \cap N &:= \{x \mid (x \in M) \wedge (x \in N)\} \\ M \setminus N &:= \{x \in M \mid x \notin N\} \end{aligned}$$

Übrigens: Die Notation „:=“ heißt, dass auf der Seite des Doppelpunkts etwas vormals unbekanntes steht, das durch die Gleichung definiert wird.

Eine ähnliche Notation wird auch in den folgenden drei nützlichen Wegen zur Angabe von Mengen verwendet:

a) Man listet alle Elemente auf, etwa $M := \{1, 3, 4, 6\}$. Das geht natürlich nur im Falle von *endlichen* Mengen. Bei unendlichen Mengen kann man manchmal mit Auslassungszeichen hinreichende Klarheit erreichen, etwa $\mathbb{Z} := \{\ldots, -2, -1, 0, 1, 2, \ldots\}$.

b) ***Aussonderungsaxiom***: Ist M eine Menge und P eine 1-stellige Aussageform, so ist die Gesamtheit $\{x \in M \mid P(x)\}$ der Elemente von M, durch die P wahr wird, eine Menge. Zum Beispiel $M := \{m \in \mathbb{Z} \mid \exists k \in \mathbb{N}\colon\ k^2 = m\}$.

c) ***Ersetzungsaxiom***: Ersetzt man jedes Element einer gegebenen Menge durch ein „Objekt unserer Anschauung oder unseres Denkens“, so entsteht wieder eine Menge. Zum Beispiel $\{n^2 \mid n \in \mathbb{Z}\}$.

Die gleiche Menge lässt sich normalerweise auf verschiedene Arten darstellen:

$$\begin{aligned} \emptyset &= \{x \in \mathbb{R} \mid x^2 < 0\} \\ \{n^2 \mid n \in \mathbb{Z}\} &= \{m \in \mathbb{Z} \mid \exists k \in \mathbb{N}\colon\ k^2 = m\} \end{aligned}$$

Schnitte und Vereinigungen kann man auch für beliebig viele Mengen (auch unendlich viele) verallgemeinern:

[6] John Venn [1834–1923]

Notation 0.10
Ist X eine Menge von Mengen, so sei

$$\bigcup_{A\in X} A := \{x \mid \exists A \in X\colon\ x \in A\} \qquad \textit{und} \qquad \bigcap_{A\in X} A := \{x \mid \forall A \in X\colon\ x \in A\}$$

Die obigen Konstruktionen setzten natürlich voraus, dass man bereits einige Mangen kennt. Die meisten der folgenden Mengen dürften aus der Schule bekannt sein, für die Existenz von manchen von ihnen benötigt man wieder Axiome (die wir nicht formulieren).

$\emptyset$, $\{\}$ Die *leere* Menge, die *keine* Elemente enthält. Dass es die leere Menge gibt, wird im ***Leermengenaxiom*** gefordert.

$\mathbb{N}$ Die Menge $\mathbb{N} := \{0, 1, 2, 3, 4, 5, \ldots\}$ aller natürlichen Zahlen — zumindest nach DIN–Norm 5473 gehört die Null zu den natürlichen Zahlen. Die Existenz von $\mathbb{N}$ folgt aus dem ***Unendlichkeitsaxiom***.

$\mathbb{N}^*$ Die Menge $\mathbb{N}^* := \{1, 2, 3, 4, 5, \ldots\}$ aller positiven natürlichen Zahlen — wieder nach DIN-Norm 5473.

$\mathbb{Z}$ Die Menge $\mathbb{Z} := \{\ldots, -2, -1, 0, 1, 2, \ldots\}$ aller ganzen Zahlen

$\mathbb{Q}$ Die Menge der rationalen Zahlen

$\mathbb{R}$ Die Menge der reellen Zahlen

$\mathbb{Q}_{>0}$ Menge der positiven rationalen Zahlen (analog $\mathbb{Q}_{\geq 0}$, $\mathbb{R}_{<0}$ etc.)

Intervalle: Für $a, b \in \mathbb{R}$, $a \leq b$, bezeichnet man $[a, b] := \{x \in \mathbb{R} \mid a \leq x \leq b\}$, $]a, b] := \{x \in \mathbb{R} \mid a < x \leq b\}$, analog $[a, b[$ und $]a, b[$, ferner $]-\infty, a] := \{x \in \mathbb{R} \mid x \leq a\}$ etc.

$\mathbb{C}$ Die Menge der komplexen Zahlen (kommt noch)

DIN–Norm 5473 wird häufig ignoriert. Sie sollten also immer auf den Kontext achten, ob $0 \in \mathbb{N}$. Wer $\mathbb{N} = \{1, 2, 3, 4, ...\}$ verwendet, der schreibt meistens $\mathbb{N}_0 = \{0, 1, 2, 3, ...\}$.

Notation 0.11
Die Anzahl der Elemente einer endlichen[7] *Menge M notiert man als $|M|$: Die* **Mächtigkeit** *oder* **Kardinalität** *von M.*

Definition 0.12
Das **kartesische Produkt** *(auch: direkte Produkt) zweier Mengen M und N besteht aus allen* **geordneten Paaren** *(m, n) mit $m \in M$ und $n \in N$:*

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

[7] Leider haben wir noch nicht definiert, was eine ***endliche Menge*** oder eine ***Anzahl*** ist.

„Geordnet“ heißt beispielsweise $(1,2) \neq (2,1)$; man beachte den Unterschied zu Mengen: $\{1,2\} = \{2,1\}$. Analog bezeichnet Ausdrücke der Form (a,b,c) als ***Tripel*** und $(x_1,...,x_n)$ mit $n \in \mathbb{N}$ als n-***Tupel***. Sind $M_1,...,M_n$ Mengen, so ist deren kartesisches Produkt $M_1 \times ... \times M_n := \{(m_1,...,m_n) \mid m_1 \in M_1,...,m_n \in M_n\}$. Statt $\underbrace{M \times ... \times M}_{n\text{-mal}}$ schreibt man auch M^n.

Beispiel In der Schule modellierten Sie den „Anschauungsraum“ als $\mathbb{R}^3$: Jeder Raumpunkt entspricht eindeutig einem Koordinatentripel (x,y,z) reeller Zahlen.

Definition 0.13 (und Potenzmengenaxiom)
Sei M eine Menge. Die Gesamtheit der Teilmengen von M bildet eine Menge, die **Potenzmenge** $\mathscr{P}(M) := \{T \mid T \subseteq M\}$.

Beispiel 0.14
$\mathscr{P}(\{1,2,3\}) = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1,2\}, \{1,3\}, \{2,3\}, \{1,2,3\}\}$. *Man kann sich relativ leicht überlegen: Ist M endlich, so gilt $|\mathscr{P}(M)| = 2^{|M|}$.*

Man beachte den Unterschied zwischen dem Aussonderungsaxiom und der Notation $\{T \mid T \subseteq M\}$: Es handelt sich um die Gesamtheit *aller* Objekte, die eine gewisses Prädikat erfüllen, statt nur Objekte aus einer vorgegebenen Menge. Der Zusammenhang von Mengen und Prädikaten hat sogar etwas Lokalkolorit. Gottlob Frege [1848–1925], der in Jena als Professor tätig war, entwickelte eine Mengenlehre auf Grundlage der Idee, dass es zu jedem Begriff eine Menge gebe, die genau die Objekte enthält, welche unter diesen Begriff fallen. Mit anderen Worten: Frege dachte, dass es für jede Aussageform $P(x)$ eine Menge gibt, welche genau diejenigen Objekte x enthält, für die $P(x)$ zu einer wahren Aussage wird. Dies nennt man das ***Abstraktionsprinzip***. Doch leider entdeckte Bertrand Russel [1872–1970], dass sich aus dem Abstraktionsprinzip ein Widerspruch ergibt. Russell wies Frege 1902 in einem Brief darauf hin und erschütterte damit die Grundlagen eines bereits im Druck befindlichen Buches von Frege.

Definition 0.15
Seien M, N Mengen. Eine **Abbildung** f *(auch* **Funktion** *genannt) von M (***Definitionsmenge***) nach N (***Zielmenge***) ordnet jedem Element $m \in M$ ein Element von N zu, welches man als $f(m)$ notiert. Notation: $f\colon M \to N$.*

Aus der Schule kennen Sie zum Beispiel die Sinusfunktion, $\sin\colon \mathbb{R} \to \mathbb{R}$, die Wurzelfunktion, $\sqrt{}\colon \mathbb{R}_{\geq 0} \to \mathbb{R}$, und hoffentlich auch die Logarithmusfunktion, $\ln\colon \mathbb{R}_{>0} \to \mathbb{R}$; es sei aber betont, dass nicht jede Abbildung durch eine „Formel“ definiert sein muss. Wie man sieht, ist nicht nötig, dass jedes Element der Zielmenge tatsächlich als Funktionswert auftritt. Man kann zwar den Sinus auch als Abbildung $\mathbb{R} \to [-1,1]$ auffassen — aber das wäre eine *andere* Funktion.

Nachträglich eingefügt wurde noch folgendes:

Alternativdefinition 0.15.b)
Seien M, N Mengen. Eine Teilmenge $f \subset M \times N$ heißt **Abbildung** *von M nach N (Notation: $f\colon M \to N$) $:\Leftrightarrow \forall m \in M\colon \exists! n \in N\colon (m,n) \in f$. Statt $(m,n) \in f$ schreibt man üblicherweise $f(m) = n$.*

Definition und Aufgabe
Seien M, N Mengen. Die Gesamtheit aller Abbildungen von M nach N wird mit N^M oder $\mathrm{Abb}(M,N)$ *bezeichnet und ist selbst eine Menge. Für $f, g \in N^M$ gilt $f = g \iff \forall m \in M\colon f(m) = g(m)$. Sind M, N endlich, so gilt $|N^M| = |N|^{|M|}$.*

0.3 Summen- und Produktzeichen

Sei $n \in \mathbb{N}$ und $a_1, ..., a_n \in \mathbb{R}$. Die ***Summe*** der $a_1, ..., a_n$ ist $\sum_{i=1}^{n} a_i = a_1 + ... + a_n$ und das ***Produkt*** der $a_1, ..., a_n$ ist $\prod_{j=1}^{n} a_j := a_1 \cdot ... \cdot a_n$. Analog für andere Indexbereiche.

Beispiele $\sum_{k=-2}^{3} k^3 = -8 - 1 + 0 + 1 + 8 + 27 = 27$; $\sum_{i=-100}^{100} 2 = 201 \cdot 2 = 402$.

Diese Definition ist aufgrund der auftretenden Auslassungszeichen nicht gut genug. Besser ist es, Summe bzw. Produkt ausgehend von einem Startwert für Null Elemente durch Hinzufügen weiterer Elemente zu erklären. So etwas nennt man eine ***rekursive Definition***.

Definition 0.16
Sei $n \in \mathbb{N}$ und $a_1, ..., a_n \in \mathbb{R}$. Wir definieren $\sum_{i=1}^{0} a_i := 0$ bzw. $\prod_{i=1}^{0} a_i := 1$. Für $n \in \mathbb{N}^*$ sei $\sum_{i=1}^{n} a_i := \left(\sum_{i=1}^{n-1} a_i\right) + a_n$ bzw. $\prod_{i=1}^{n} a_i := \left(\prod_{i=1}^{n-1} a_i\right) \cdot a_n$.

Beispiel 0.17

a) **Fakultät***: Für $n \in \mathbb{N}$ sei $n! := \prod_{i=1}^{n} i = 1 \cdot 2 \cdot ... \cdot (n-1) \cdot n$.*

b) **Potenzrechnung***: Für $n \in \mathbb{N}$ und $x \in \mathbb{R}$ sei $x^n := \prod_{i=1}^{n} x$ und $x^{-n} := \frac{1}{x^n}$.*
Daraus folgen die aus der Schule bekannten Potenzrechengesetze.

Bemerkung 0.18
Summen bzw. Produkte mit unendlich *vielen Summanden bzw. Faktoren sind nicht definiert! Allenfalls lassen sich Ausdrücke wie $\sum_{n \in \mathbb{N}^*} \frac{(-1)^n}{n}$ mit Hilfe von Grenzwerten, also mit Mitteln der Analysis, erfassen. Man nennt das dann nicht mehr Summe, sondern Reihe. Aber dann kann man im Allgemeinen nicht die Summanden beliebig vertauschen. Übrigens: Obige Reihe hat den Wert $-\ln(2)$, sofern man die Indizes aufsteigend $(1, 2, 3, 4, ...)$ sortiert.*

1 Ringe und Körper

Ich setze voraus, dass Sie die Grundrechenarten mit ganzen und reellen Zahlen (insbesondere auch Bruchrechnung) sowie mit Polynomen beherrschen. Tatsächlich beruhen wesentliche Teile der Linearen Algebra nur auf allgemeinen Eigenschaften der Grundrechenarten: Kommutativ-, Assoziativ- und Distributivgesetze. In der Schule sollten Sie allerdings auch Potenzrechnung, Wurzeln, Exponentialfunktion, Logarithmen sowie trigonometrische Funktionen kennen gelernt haben. Für die Arbeit mit komplexen Zahlen sind auch Sinus, Kosinus und Exponentialfunktion wichtig. Und in den späteren Kapiteln der Linearen Algebra müssen auch Nullstellen von Polynomen bestimmt werden.

In der Mathematik betrachtet man die wesentlichen Eigenschaften der Grundrechenarten als Werkzeugkasten: Man überlegt sich, was man mit jedem einzelnen Werkzeug anfangen kann, und entwickelt ein abgestuftes Begriffssystem. Gymnasien sollten im Hinblick auf „allgemeine Studierfähigkeit“ in begriffliches Denken einführen, doch sogar beim Rechnen „mit Buchstaben“ hat man meist *nur ein einzelnes Beispiel* im Hintergrund (nämlich die reellen Zahlen), das dazu auch noch auf einer hochkomplizierten Konstruktion basiert, die man in der Schule aber nicht erklärt; aus fachwissenschaftlicher Sicht ist das ein Unding.

Die aus der Schule bekannten Rechenbereiche sind zunächst einmal *Mengen*, nämlich Mengen von Zahlen. Sowohl durch Addition als auch durch Multiplikation wird jedem Paar von Zahlen eine neue Zahl zugeordnet. Dies gibt Anlass zu folgender Definition.

Definition 1.1
Eine **innere Verknüpfung** *auf einer Menge M ist eine Abbildung $M \times M \to M$.*

Anders als sonst für Abbildungen üblich wird nicht die Funktionenschreibweise $f(a, b)$ verwendet, sondern eines von vielen möglichen ***Verknüpfungssymbol***en, also zum Beispiel $a*b$, $a \otimes b$ etc. Ist die Menge endlich, so lässt sich die Verknüpfung durch eine Verknüpfungstafel angeben.

Definition 1.2

a) Ein **Ring** $(R, +, \cdot, 0, 1)$ *ist eine Menge R zusammen mit zwei inneren Verknüpfungen $+$ und $\cdot$, einem* **Nullelement** $0 \in R$ *und einem* **Einselement** $1 \in R$*, so dass die folgenden* **Ringaxiome** *erfüllt sind:*

Kommutativität von $+$: $\forall x, y \in R\colon\ x + y = y + x$

Assoziativität: $\forall x, y, z \in R\colon\ (x \cdot y) \cdot z = x \cdot (y \cdot z)$ *und* $(x+y)+z = x+(y+z)$

Distributivität: $\forall x, y, z \in R\colon\ (x+y) \cdot z = x \cdot z + y \cdot z$ *und* $x \cdot (y+z) = x \cdot y + x \cdot z$

Neutrale Elemente: $\forall x \in R\colon\ x + 0 = x$ *und* $x \cdot 1 = x = 1 \cdot x$

Negation: $\forall x \in R\colon\ \exists -x \in R\colon\ x + (-x) = 0$*; man nennt $-x$ das* **additive Inverse** *von x.*

Statt $a + (-b)$ schreiben wir wie üblich $a - b$ für $a, b \in R$. Den Multiplikationspunkt lässt man meist weg. Man setzt $R^* := R \setminus \{0\}$.

b) *Ein Ring R heißt* **kommutativ**, *gdw. $\forall x, y \in R\colon\ x \cdot y = y \cdot x$ (Kommutativität von $\cdot$).*

c) *Ein kommutativer Ring R heißt* **Körper**, *gdw. $1 \neq 0$ und $\forall x \in R^*\colon\ \exists x^{-1} \in R\colon\ xx^{-1} = x^{-1}x = 1$ (Existenz der Inversen). Man nennt x^{-1} auch das* **multiplikative Inverse** *von x.*
Die Ringaxiome zusammen mit der Kommutativität von $\cdot$ und der Existenz der multiplikativen Inversen nennt man auch **Körperaxiome**.

Meist sagt man einfach: „Sei R ein Ring“ — nur wenn nicht aus dem Kontext klar ist, welche Verknüpfungen für Addition und Multiplikation zu verwenden sind, sagt man explizit: „Sei $(R, +, \cdot, 0, 1)$ ein Ring“.

Beispiel 1.3

a) *Laut Ihrer Schulkenntnisse sind $\mathbb{Z}$ ebenso wie die Menge $\mathbb{R}[X]$ aller Polynome kommutative Ringe, jedoch keine Körper.*

b) *Hingegen sind $\mathbb{R}$ und $\mathbb{Q}$ Körper.*

c) *$\{0\}$ mit $0+0 = 0\cdot 0 = 0$ ist ein Ring, der* **triviale Ring** *(auch:* **Nullring***). Hier ist 0 gleichzeitig Null- und Einselement, $\{0\}$ ist also kein Körper.*

d) *$\mathbb{F}_2 := \{0, 1\}$ mit folgenden inneren Verknüpfungen ist ein Körper:*

$+$	0	1
0	0	1
1	1	0

$\cdot$	0	1
0	0	0
1	0	1

e) *$\mathbb{F}_3 := \{0, 1, 2\}$ mit folgenden inneren Verknüpfungen ist ein Körper:*

$+$	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

$\cdot$	0	1	2
0	0	0	0
1	0	1	2
2	0	2	1

f) *Da Computer zur Darstellung reeller Zahlen nur eine begrenzte Zahl von Stellen zur Verfügung stellen, ist es in vielen Anwendungen unvermeidbar, gerundet zu rechnen. Dabei gelten Distributiv- und Assoziativgesetze leider nicht, und nicht jede von Null verschiedene Zahl hat ein Inverses! Es wird in dieser Vorlesung an ausgewählten Stellen erste Einblicke in gerundetes Rechnen geben, aber eigentlich ist das erst Thema in der Numerik.*

g) *Die komplexen Zahlen bilden einen Körper — siehe nächstes Kapitel.*

h) Es ist möglich, dass es nur eine „Rechts-1“ gibt, aber keine „Links-1“, und dass auch die Distributivgesetze nur von einer Seite gelten. Ist etwa $\mathbb{Z}$ mit der gewohnten Addition, aber mit der Multiplikation $*$ definiert durch $\forall a, b \in \mathbb{Z}\colon\ a * b = a$, dann ist JEDES Element eine Rechts-1, aber es gibt keine Links-1.

Zudem gilt $\forall a, b, c \in \mathbb{Z}\colon\ (a + b) * c = a + b = a * c + b * c$, aber das andere Distributivgesetz ist verletzt, etwa $1 * (1 + 1) = 1 \neq 1 * 1 + 1 * 1 = 2$.

Aus den Körperaxiomen folgen die meisten aus der Schule bekannten Rechenregeln, beispielsweise die folgenden:

Satz 1.4
Sei R ein Ring.

a) Kürzungsregel: $\forall x, y, z \in R\colon\ (x + z = y + z \Rightarrow x = y)$.
Ist R ein Körper, so gilt zudem $\forall x, y, z \in R,\ z \neq 0\colon\ (xz = yz \Rightarrow x = y)$.

b) Eindeutigkeit des Einselements: $\forall x \in R\colon\ (\forall y \in R\colon\ y \cdot x = x \cdot y = y) \Rightarrow x = 1$

c) Eindeutigkeit des Nullelements: $\forall x \in R\colon\ (\exists y \in R\colon\ y + x = y) \Rightarrow x = 0$.
Ist R ein Körper, so gilt $\forall x \in R\colon\ (\exists y \in R^*\colon\ yx = y) \Rightarrow x = 1$.

d) $\forall x \in R\colon\ -(-x) = x$. Ist R ein Körper, dann auch $\forall x \in R^*\colon\ (x^{-1})^{-1} = x$.

e) $\forall x \in R\colon\ 0 \cdot x = 0$.

f) $\forall x \in R\colon\ (-1) \cdot x = -x$.

Beweis:

a) $x + z = y + z \Rightarrow (x + z) + (-z) = (y + z) + (-z)$, denn $+$ ist eine Abbildung, d.h. bei gleichen Eingabewerten ist auch die Ausgabe gleich. $\overset{\text{Assoz.}}{\Rightarrow} x + (z + (-z)) = y + (z + (-z)) \overset{\text{Negat.}}{\Rightarrow} x + 0 = y + 0 \overset{\text{Neutr.}}{\Rightarrow} x = y$. In einem Körper existiert das multiplikative Inverse von $z \in R^*$, also folgt analog die Kürzungsregel für die Multiplikation.

b) Sei $x \in R$ so, dass $\forall y \in R\colon\ y \cdot x = x \cdot y = y$. Dann gilt dies insbesondere für $y = 1$, also $x \overset{\text{Def.}}{=} x \cdot 1 \overset{\text{Vor.}}{=} 1$.

c) $y + x \overset{\text{Vor.}}{=} y \overset{\text{Neutr.}}{=} y + 0 \overset{a)}{\Rightarrow} x = 0$. Analog für Multiplikation im Fall von Körpern. Beachte: Das ist eine stärkere Aussage als die Eindeutigkeit des Einselements in Ringen!

d) $(-x) + (-(-x)) \overset{\text{Negat.}}{=} 0 \overset{\text{Negat.}}{=} x + (-x) \overset{\text{Komm.}}{=} (-x) + x \overset{a)}{\Rightarrow} x = -(-x)$. Analog für Multiplikation in Körpern.

e) $0 + 0 \cdot x \overset{\text{Neutr.}}{=} 0 \cdot x \overset{\text{Neutr.}}{=} (0+0) \cdot x \overset{\text{Distr.}}{=} 0 \cdot x + 0 \cdot x \overset{a)}{\Rightarrow} 0 = 0 \cdot x.$

f) $x + (-1) \cdot x \overset{\text{Neutr.}}{=} 1 \cdot x + (-1) \cdot x \overset{\text{Distr.}}{=} (1 + (-1)) \cdot x \overset{\text{Negat.}}{=} 0 \cdot x \overset{e)}{=} 0 = x + (-x) \overset{a)}{\Rightarrow} (-1) \cdot x = -x.$ □

Beispiel Es folgt die bekannte Regel $(-1) \cdot (-1) = -(-1) = 1$.

1.1 Restklassenringe

Das Thema dieses Abschnitts wurde in den Präsenzaufgaben gedanklich vorbereitet. Wir führen zunächst den Begriff der Relation ein.

Definition 1.5
Sei X eine Menge.[8] *Eine (binäre bzw. zweistellige)* Relation *auf X ist eine Abbildung $X \times X \to \{W, F\}$, man könnte auch sagen eine auf X beschränkte zweistellige Aussageform. Sind $x, y \in X$, so drückt man die meistens durch ein Relationssymbol aus (etwa $x \sim y$, $x \sim_R y$, $x \equiv y$, $x \cong y$). Wir verwenden hier $\sim$. Es ist auch üblich, eine Relation als eine Teilmenge von $X \times X$ aufzufassen, nämlich $\{(x, y) \in X \times X \mid x \sim y\}$.*

Eine Relation $\sim$ auf X heißt

- reflexiv, *falls* $\forall x \in X\colon\ x \sim x$.
- symmetrisch, *falls* $\forall x, y \in X\colon\ x \sim y \iff y \sim x$.
- transitiv, *falls* $\forall x, y, z \in X$ *gilt: Wenn* $x \sim y$ *und* $y \sim z$, *denn* $x \sim z$.

Eine reflexive symmetrische transitive Relation heißt **Äquivalenzrelation**.

Beispiel 1.6

a) $X =$ alle Städte Deutschlands: Relation „liegt im gleichen Bundesland wie“ ist eine Äquivalenzrelation.

b) Auf $\mathbb{R}$ ist die Relation $<$ weder reflexiv noch symmetrisch, aber sie ist transitiv; es handelt sich um eine **Ordnungsrelation**, *die ich hier zunächst nicht weiter thematisieren möchte.*

c) $X = \mathbb{R}^3$ mit der Relation „$(u_1, u_2, u_3) \bowtie (v_1, v_2, v_3) :\Leftrightarrow$ stimmen an mindestens zwei *Stellen überein“: Die Relation ist reflexiv und symmetrisch, aber nicht transitiv, denn $(1,1,1) \bowtie (1,1,2)$ und $(1,1,2) \bowtie (1,2,2)$, aber nicht $(1,1,1) \bowtie (1,2,2)$. Also ist es* keine *Äquivalenzrelation.*

Im Rest des Abschnitts sei R ein kommutativer Ring.

[8] Für echte Klassen hat man die gleiche Definition.

Definition 1.7
$\forall a, b \in R\colon\ a|b :\Leftrightarrow \exists c \in R\colon\ b = a \cdot c$ *(d.h. „a* **teilt** *b“).*

Sei $n \in R^*$. $\forall a, b \in R\colon\ a \equiv_n b :\Leftrightarrow n|(b-a)$ *(d.h. a ist* **kongruent** *zu b* **modulo** *n). Andere Schreibweise:* $a \equiv b \pmod{n}$.

Bemerkung In der Präsenzübung lernten Sie eine auf Division mit Rest basierende Definition von Kongruenz kennen. In $\mathbb{Z}$ sind beide Definitionen gleichbedeutend, aber in allgemeinen kommutativen Ringen steht eine Division mit Rest nicht zur Verfügung. Daher ist Definition 1.7 die Definition, mit der wir ab jetzt arbeiten werden.

Lemma 1.8
$\forall n \in R^*$: *„Kongruenz modulo n“ ist eine Äquivalenzrelation auf R.*

Beweis:
Reflexivität: $\forall a \in R\colon\ (a-a) = 0 = n \cdot 0$, also $n|(a-a)$, also $a \equiv_n a$.

Symmetrie: Sei $a, b \in R$. Für $q \in R$ ist $b - a = q \cdot n$ genau dann wenn $a - b = (-q) \cdot n$. Folglich $n|(b-a) \iff n|(a-b)$, und das heißt $a \equiv_n b \iff b \equiv_n a$.

Transitivität: Seien $a, b, c \in R$ mit $a \equiv_n b$ und $b \equiv_n c$. Dann $\exists q_1, q_2 \in R\colon\ (b-a) = q_1 \cdot n$ und $(c-b) = q_2 \cdot n$. Daher $(c-a) = (c-b)+(b-a) = q_2 \cdot n + q_1 \cdot n = (q_2 + q_1) \cdot n$, also $a \equiv_n c$. □

Definition 1.9
Ist $\sim$ *eine Äquivalenzrelation auf der Menge X und ist* $y \in X$*, so heißt* $[y] := [y]_\sim := \{x \in X \mid x \sim y\}$ *die* **Äquivalenzklasse** *von y. Ist K eine Äquivalenzklasse und* $x \in K$*, so heißt x ein* **Repräsentant** *von K.*

Lemma 1.10
Sei $\sim$ *eine Äquivalenzrelation auf der Menge X. Wegen Reflexivität gilt* $\forall x \in X\colon\ x \in [x]$*. Außerdem sind für* $x, y \in X$ *folgende drei Aussagen gleichbedeutend:*

a) $x \sim y$

b) $[x] \cap [y] \neq \emptyset$.

c) $[x] = [y]$

Beweis:
Wir zeigen a) ⇒ b) ⇒ c) ⇒ a).

a) ⇒ b) Aus $x \sim y$ folgt $x \in [y]$. Wegen Reflexivität ist aber auch $x \in [x]$. Daher ist $x \in [x] \cap [y]$, also $[x] \cap [y] \neq \emptyset$.

b) ⇒ c) Wenn $[x] \cap [y] \neq \emptyset$, dann gibt es ein $z \in [x] \cap [y]$. Wir wollen nun zeigen, dass $[x] \subseteq [y]$; wenn also $w \in [x]$, so wollen wir zeigen, dass auch $w \in [y]$. Aus $w \in [x]$ folgt $w \sim x$. Wegen $z \in [x]$ ist $z \sim x$, also $x \sim z$ wegen Symmetrie. Aus $w \sim x$ und $x \sim z$ folgt $w \sim z$ wegen Transitivität. Aufgrund von $z \in [y]$ gilt $z \sim y$, also aus $w \sim z$ folgt $w \sim y$ wegen Transitivität, also $w \in [y]$, was zu zeigen war.

Analog folgt $[y] \subseteq [x]$, also $[y] = [x]$, also c).

c) ⇒ a) Wenn $[x] = [y]$, dann $x \in [x] = [y]$, daher $x \sim y$. □

Satz 1.11 (und Definition)
Sei R ein kommutativer Ring und $n \in R^*$. Wir betrachten die durch $\forall x, y \in R$ durch $x \equiv_n y$ gegebene Äquivalenzrelation und definieren ${}^R/_{nR} := \{[x] \mid x \in R\}$.

a) Für $x, y \in R$ sind durch $[x]+[y] := [x+y]$ und $[x] \cdot [y] := [x \cdot y]$ zwei innere Verknüpfungen auf ${}^R/_{nR}$ definiert.

b) Durch diese beiden inneren Verknüpfungen wird ${}^R/_{nR}$ zu einem kommutativen Ring, dem so genannten **Restklassenring** *von R modulo n.*

Beweis:

a) Ist eine Definition von der Wahl der Repräsentanten unabhängig, so nennt man sie ***wohldefiniert***; wir müssen die Wohldefiniertheit nachweisen.

Seien $x', y' \in R$ mit $[x] = [x']$ und $[y] = [y']$. Nach dem vorigen Lemma[9] ist das gleichbedeutend zu $x \equiv_n x'$ und $y \equiv_n y'$. Nach Definition von Kongruenz modulo n und Teilbarkeit ist das gleichbedeutend zu $\exists q_x, q_y \in R$: $x' = x + q_x \cdot n$ und $y' = y + q_y \cdot n$.

Behauptung: $[x + y] = [x' + y']$ und $[x \cdot y] = [x' \cdot y']$.

- $x'+y' = x+q_x \cdot n+y+q_y \cdot n = (x+y)+(q_x+q_y) \cdot n$, also $x+y \equiv_n x'+y'$, was zu zeigen war.
- $x' \cdot y' = (x + q_x \cdot n) \cdot (y + q_y \cdot n) = x \cdot y + (x \cdot q_y + q_x \cdot y + q_x \cdot q_y \cdot n) \cdot n$, also $x \cdot y \equiv_n x' \cdot y'$, was zu zeigen war.

b) Die kommutativen Ringaxiome lassen sich für Äquivalenzklassen leicht durch Wahl von Repräsentanten nachweisen, da die entsprechenden Axiome in R gelten. Wir führen dies nur am Beispiel des Distributivgesetzes vor, den Rest kann man sich als Übung überlegen: Seien $X, Y, Z \in {}^R/_{nR}$ mit Repräsentanten x, y, z. Wir haben also $x, y, z \in R$ mit $[x] = X$, $[y] = Y$ und $[z] = Z$. Es gilt $X \cdot (Y+Z) = [x] \cdot ([y]+[z]) \overset{\text{Def}}{=} [x] \cdot [y+z] \overset{\text{Def}}{=} [x \cdot (y+z)] \overset{\text{Distr}}{=} [x \cdot y + x \cdot z] \overset{\text{Def}}{=} [x \cdot y] + [x \cdot z] \overset{\text{Def}}{=} [x] \cdot [y] + [x] \cdot [z]$. □

[9] Beachte: Die hier betrachtete Äquivalenzrelation wird nicht als $x \sim y$, sondern als $x \equiv_n y$ notiert.

Beispiel 1.12

a) Wenn Sie ausrechnen wollen, welcher Wochentag in 23 *Tagen ist, dann rechnen Sie in* $\mathbb{Z}/7\mathbb{Z}$*:* $[23] = [2]$*, wenn also heute Montag ist, dann ist in* 23 *Tagen zwei mehr als Montag, also Mittwoch.*

b) Sei $x \in \mathbb{N}$ *mit Dezimalziffern* $z_k, z_{k-1}, ..., z_0 \in \{0, ..., 9\}$*. Möglicherweise*[10] *kennen Sie aus der Schule folgende Teilbarkeitsregel:* x *ist durch* 11 *teilbar genau dann, wenn seine* **alternierende Quersumme** $z_0 - z_1 + z_2 - ... \pm z_k$ *durch* 11 *teilbar ist. Dies beweist man durch eine kleine Rechnung in* $\mathbb{Z}/11\mathbb{Z}$*: Wir haben* $x = \sum_{i=0}^{k} z_i \cdot 10^i$*. Also ist*

$$\begin{aligned}[x] &= \left[\sum_{i=0}^{k} z_i \cdot 10^i\right] = \sum_{i=0}^{k} [z_i] \cdot [10]^i \\ &= \sum_{i=0}^{k} [z_i] \cdot [-1]^i \qquad \textit{denn } 10 \equiv_{11} -1 \\ &= \left[\sum_{i=0}^{k} (-1)^i z_i\right] .\end{aligned}$$

Das bedeutet: Jede Zahl hat bei Division durch 11 *den gleichen Rest wie ihre alternierende Quersumme.*

c) In der Computeralgebra sind manchmal „modulare" Algorithmen zu finden: Wenn eine Rechnung in $\mathbb{Z}$ *zu schwer ist, so rechnet man in mehreren Restklassenringen. Aus den Ergebnissen in den Restklassenringen kann man dann das eigentlich gesuchte Ergebnis in* $\mathbb{Z}$ *rekonstruieren. Das Beispiel, das ich hier gebe, ist allerdings etwas künstlich, denn Quadratwurzeln würde man in der Computeralgebra anders ziehen.*

Gesucht sei $\sqrt{1369}$*; offenbar ist* $30 = \sqrt{900} < \sqrt{1369} < \sqrt{1600} = 40$*.* ***Wenn*** *das Ergebnis eine ganze Zahl ist (das müssen wir am Ende überprüfen), so könnte man wie folgt rechnen:*

- *In* $\mathbb{Z}/11\mathbb{Z}$*: Wegen b) ist* $1369 \equiv_{11} (9 - 6 + 3 - 1)$*, d.h.* $[1369] = [5]$*. Ferner ist* $[4]^2 = [7]^2 = [5]$*. Daher:* ***Wenn*** $\sqrt{1369}$ *ganzzahlig ist, hat es Rest* 4 *oder* 7 *bei Division durch* 11*.*
- *In* $\mathbb{Z}/10\mathbb{Z}$*: Ein Blick auf die letzte Ziffer zeigt* $[1389] = [9]$*, und* $[3]^2 = [7]^2 = [9]$*. Daher:* ***Wenn*** $\sqrt{1369}$ *ganzzahlig ist, hat es Rest* 3 *oder* 7 *bei Division durch* 10*.*

[10] Im Lehrplan steht sie glaube ich nicht mehr.

Wenn die Reste bezüglich 11 *und* 10 *vorgegeben sind, so gibt es nach dem* **Chinesischen Restsatz**[11]*, den ich hier allerdings nicht beweisen werde,* genau ein $w \in \{0, ..., 10 \cdot 11 - 1\}$ *mit den vorgegebenen Resten. Für die Reste* $(4, 3)$, $(7, 3)$, $(4, 7)$, $(7, 7)$ *ergibt sich* $w = 103$, $w = 73$, $w = 37$ *bzw.* $w = 7$. *Der einzige Kandidat zwischen* 20 *und* 40 *ist* 37, *und tatsächlich* $37^2 = 1369$.

Bemerkung 1.13
Sei $n \in \mathbb{N}^*$. $\mathbb{Z}/n\mathbb{Z}$ *ist genau dann ein Körper, wenn* n *eine Primzahl ist.*

Ein Teil dieser Aussage wird in den Übungen bewiesen. Übrigens: $n \in \mathbb{N}^*$ heißt ***Primzahl*** $:\Leftrightarrow n \nmid 1$ und $\forall a, b \in \mathbb{Z}\colon\ n|a \cdot b \ \Rightarrow\ (n|a) \vee (n|b)$. Der Primzahlbegriff, den Sie vermutlich in der Schule lernten („n heißt prim genau dann wenn n genau zwei Teiler hat“ oder „$n > 1$ heißt prim genau dann wenn n sich nicht als Produkt zweier von 1 verschiedener natürlicher Zahlen schreiben lässt“) heißt in der Algebra nicht „prim“ sondern ***irreduzibel***. Das ist in $\mathbb{Z}$ dasselbe, aber in manchen Ringen ist das ein Unterschied.

1.2 Komplexe Zahlen

Wir lernten im vorigen Abschnitt eine reichhaltige Quelle endlicher Körper kennen. In diesem Abschnitt geht es um einen Körper, der die reellen Zahlen umfasst. Diese Zahlbereichserweiterung wird aus mir unerfindlichen Gründen heute in der Schule meist nicht behandelt.

Sie lernten in der Schule, mit Polynomen zu rechnen und deren Nullstellen zu suchen. Nicht zu allen Polynomen war dies möglich. Schade! Glücklicherweise kann man die reellen Zahlen zu einem größeren Körper erweitern. Dazu führt man ein Symbol i ein (***imaginäre Einheit***, nach Norm DIN 1302. In der Elektrotechnik darf auch j als Symbol verwendet werden; fieserweise sind i und i verschiedene Symbole). Mit i rechnet man wie mit der Unbekannten von Polynomen, mit einem wichtigen Unterschied: Zusätzlich hat man $\mathrm{i}^2 = -1$.

Beispiel Es folgt $\mathrm{i}^3 = \mathrm{i} \cdot \mathrm{i}^2 = -\mathrm{i}$ und $\mathrm{i}^4 = (\mathrm{i}^2)^2 = (-1)^2 = 1$.

Auf diese Weise kann man alle höheren Potenzen von i beseitigen. Man erhält:

Definition 1.14
$\mathbb{C} := \{a + b\,\mathrm{i} \mid a, b \in \mathbb{R}\}$ *ist die Menge der* **komplexen Zahlen**. *Der* **Realteil** *von* $z = a + b\,\mathrm{i} \in \mathbb{C}$ *ist* $\mathrm{Re}(z) := a \in \mathbb{R}$, *der* **Imaginärteil** *ist* $\mathrm{Im}(z) := b \in \mathbb{R}$.

Beispiel 1.15 (und Definition)
In komplexen Zahlen kann man auch Wurzeln aus negativen Zahlen ziehen. Die Gleichung $x^2 = -1$ *hat nämlich in* $\mathbb{C}$ *die beiden Lösungen* $x_{1,2} = \pm\,\mathrm{i}$. *Man erweitert daher die Definition der Wurzelfunktion: Für* $a \in \mathbb{R}_{>0}$ *setzt man* $\sqrt{-a} := \mathrm{i}\sqrt{a}$. *Das ist sinnvoll, denn* $(\mathrm{i}\sqrt{a})^2 = \mathrm{i}^2(\sqrt{a})^2 = (-1) \cdot a = -a$.

[11] Die Aussage war im 3. Jhdt. in China und die algorithmische Lösung im 6. Jhdt. in Indien bekannt.

Rechenregeln für komplexe Zahlen
$\mathbb{C}$ *wird zu einem Körper, indem man für* $z_1 = a_1 + b_1\,\mathrm{i} \in \mathbb{C}$ *und* $z_2 = a_2 + b_2\,\mathrm{i} \in \mathbb{C}$ *mit* $a_1, a_2, b_1, b_2 \in \mathbb{R}$ *wie folgt rechnet:*

- $z_1 + z_2 = (a_1 + a_2) + (b_1 + b_2)\,\mathrm{i}$ *sowie* $z_1 - z_2 = (a_1 - a_2) + (b_1 - b_2)\,\mathrm{i}$
- $z_1 \cdot z_2 = (a_1a_2 - b_1b_2) + (a_1b_2 + a_2b_1)\,\mathrm{i}$
- *Falls* $a_2 + b_2\,\mathrm{i} \neq 0$*:* $\dfrac{z_1}{z_2} = \dfrac{a_1a_2 + b_1b_2}{a_2^2 + b_2^2} + \dfrac{a_2b_1 - a_1b_2}{a_2^2 + b_2^2}\,\mathrm{i}$.

Beweis: Übung. □

Definition 1.16
Für $z = a + b\,\mathrm{i} \in \mathbb{C}$ *mit* $a, b \in \mathbb{R}$ *ist* $\bar{z} := a - b\,\mathrm{i} \in \mathbb{C}$ *die* **konjugiert komplexe Zahl***; in der Physik schreibt man auch* z^* *statt* $\bar{z}$*. Ferner ist* $|z| := \sqrt{z \cdot \bar{z}} = \sqrt{a^2 + b^2} \in \mathbb{R}_{\geq 0}$ *der* **Betrag** *von* z*.*

Beobachtung 1.17
Weil stets $a^2 + b^2 \in \mathbb{R}_{\geq 0}$ *gilt, ist* $|z| \in \mathbb{R}_{\geq 0}$*. Ferner ist* $|z| = 0 \iff z = 0$*. Die Divisionsregel kann man sich wie folgt merken: Für* $z_1, z_2 \in \mathbb{C}$ *und* $z_2 \neq 0$ *gilt*

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

Bemerkung 1.18
Selbst wenn wir als bekannt voraussetzen, dass $\mathbb{R}[X]$ *einen kommutativen Ring bezüglich der üblichen Addition und Multiplikation von Polynomen bildet, ist nicht unmittelbar klar, dass die Zusatzregel* $\mathrm{i}^2 = -1$ *die Ringeigenschaften nicht zerstört. Ein solcher Bewei wurde in den Übungen geführt, indem die Ringaxiome für die hier verwendeten Rechenregel explizit nachgeprüft wurden. Es gibt noch mindestens zwei andere Wege,* $\mathbb{C}$ *konstruieren:*

a) Sei $R := \mathbb{R}[X]$ *und* $d := X^2 + 1$*. Dann kann man* $\mathbb{C} := {}^{R}\!/_{dR}$ *definieren. Die Restklassen bei Division durch* d *sind nämlich durch Elemente der Form* $a + bX$ *mit* $a, b \in \mathbb{R}$ *repräsentiert. Zudem ist* $X^2 \equiv_d -1$*.*

b) Man kann $\mathbb{C}$ *auch durch Multiplikation gewisser* (2×2)*-Matrizen beschreiben. Siehe nächstes Kapitel.*

Der folgende wichtige Satz lässt sich auf unterschiedlichen Wegen beweisen, doch dabei müssten Anleihen bei der Analysis in einem Umfang gemacht werden, der den Rahmen einer Anfängervorlesung sprengen würde.

Fundamentalsatz der Algebra
Jedes Polynom vom Grad ≥ 1 *mit Koeffizienten aus* $\mathbb{C}$ *hat mindestens eine Nullstelle in* $\mathbb{C}$*.* □

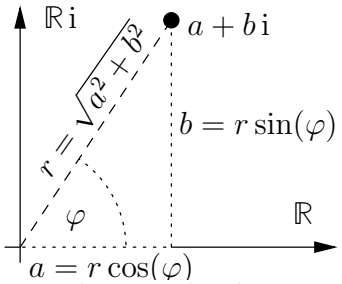
Ein zentrales Ergebnis der Galois-Theorie ist, dass es ab Grad 5 im Allgemeinen beweisbar unmöglich ist, die Nullstellen mit einer Lösungsformel (wie für quadratische Gleichungen) zu finden.

1.2.1 Die Gaußsche Zahlenebene

Bereits in der Mathematik des 16. oder 17. Jhdts. benutzte man komplexe Zahlen. Eine Veranschaulichung von $\mathbb{C}$ als Zahlenebene etablierte sich erst im 19. Jhdt.

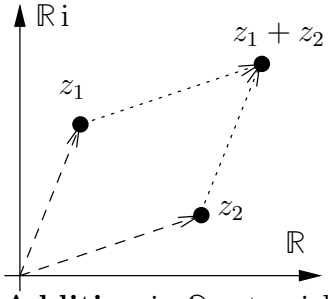
Definition 1.19
$\mathbb{C}$ mit der Ebene $\mathbb{R}^2$ identifiziert, indem man $a + b\,\mathrm{i} \in \mathbb{C}$ als $(a, b) \in \mathbb{R}^2$ darstellt. Dies heißt **Gaußsche Zahlenebene**[12] *oder* **Arganddiagramm**[13]*, wurde aber zuerst 1797 von Caspar Wessel [1745–1818] beschrieben.*

Darstellungsarten komplexer Zahlen
Sei $z = a + b\,\mathrm{i} \in \mathbb{C}$. Seien (r, φ) die Polarkoordinaten (Bogenmaß) des Punkts $(a, b) \in \mathbb{R}^2$, also $r \geq 0$ und $a = r\cos(\varphi)$, $b = r\sin(\varphi)$. Dann $z = a + b\,\mathrm{i} = r(\cos(\varphi) + \mathrm{i}\sin(\varphi))$. Nach dem Satz des Pythagoras ist $r^2 = a^2 + b^2$, also $r = |z|$.

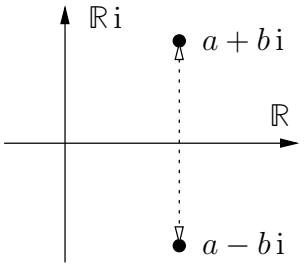


Man nennt φ das ***Argument*** $\arg(z)$ von z; es gibt verschiedene Konventionen, ob $\varphi \in]-\pi, \pi]$ oder $\varphi \in [0, 2\pi[$ gelten soll. Die Darstellung $z = r \cdot (\cos\varphi + \mathrm{i}\sin\varphi)$ heißt ***Polardarstellung*** oder ***trigonometrische Darstellung***, die Darstellung $z = a + b\,\mathrm{i}$ heißt ***Standarddarstellung*** oder ***kartesische Darstellung***. Übrigens: Oft lasse ich bei trigonometrischen Funktionen oder Logarithmus die Klammern weg, also $\sin\varphi$ statt $\sin(\varphi)$ oder $\ln\pi$ statt $\ln(\pi)$.

Veranschaulichung von Addition und Konjugation



Addition in $\mathbb{C}$ entspricht Vektoraddition in $\mathbb{R}^2$ (Kräfteparallelogramm).



Komplexe Konjugation ist Spiegelung an der reellen Achse.

Die ***Dreiecksungleichung*** besagt: Für $z_1, z_2 \in \mathbb{C}$ gilt $|z_1 + z_2| \leq |z_1| + |z_2|$. In der Mathematik müsste man solche Dinge eigentlich beweisen, doch hier berufen wir uns auf die Anschauung (siehe linkes Bild).

Rechnen in Polarkoordinaten
Addition und Konjugation sind in Standarddarstellung leicht, Multiplikation oder gar Wurzelziehen hingegen schwer. Zwar ist Addition in Polarkoordinaten schwer,

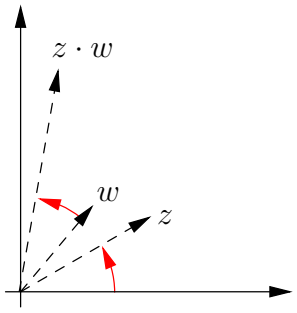
[12] Carl Friedrich Gauß [1777–1855] beschrieb sie in einem Brief von 1811.
[13] Jean–Robert Argand [1768–1822] beschrieb sie schon 1806

aber Konjugation ist leicht: Weil Konjugation die Spiegelung an der reellen Achse ist, gilt $|\bar{z}| = |z|$ und $\arg(\bar{z}) = -\arg(z)$ für alle $z \in \mathbb{C}$.

Auch Multiplikation, Division und sogar Wurzelziehen sind in Polardarstellung ziemlich leicht. Denn für $z = r(\cos\varphi + \mathrm{i}\sin\varphi)$ und $w = s(\cos\psi + \mathrm{i}\sin\psi)$, mit $r, s, \varphi, \psi \in \mathbb{R}$ ist nach den ***Additionstheoremen***[14] für Sinus und Kosinus, die wir hier nicht beweisen:

$$\begin{aligned} z \cdot w &= rs(\cos(\varphi) + \mathrm{i}\sin(\varphi))(\cos(\psi) + \mathrm{i}\sin(\psi)) \\ &= rs((\cos(\varphi)\cos(\psi) - \sin(\varphi)\sin(\psi)) + \mathrm{i}(\cos(\varphi)\sin(\psi) + \sin(\varphi)\cos(\psi))) \\ &= rs(\cos(\varphi + \psi) + \mathrm{i}\sin(\varphi + \psi)) \end{aligned}$$

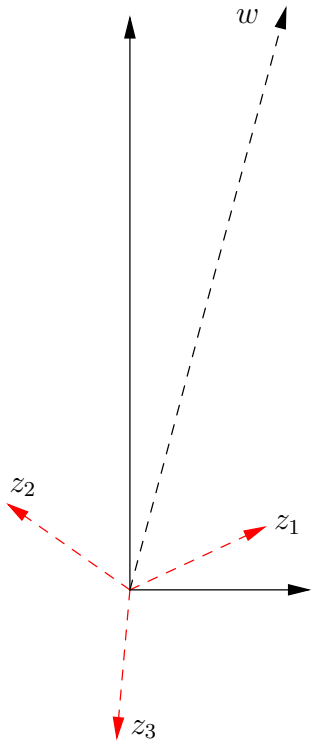
Also $|zw| = |z|\,|w|$ und $\arg(zw) = \arg(z) + \arg(w)$. Falls $w \neq 0$, erhält man ebenso $\left|\frac{z}{w}\right| = \frac{|z|}{|w|}$ und $\arg(\frac{z}{w}) = \arg(z) - \arg(w)$.
Im nebenstehenden Bild ist $|z| = 2$ und $|w| = 1.5$, daher $|zw| = 3$, und der Winkel zwischen positiver reeller Achse und z ist wie zwischen w und zw.



Beispiel 1.20 (Wurzelziehen)
Sei $\Phi := \{0, \frac{2\pi}{5}, \frac{4\pi}{5}, \frac{6\pi}{5}, \frac{8\pi}{5}\}$. *Dies sind alle Winkel, die mit* 5 *multipliziert ein Vielfaches von* 2π *ergeben. Daher gilt* $z^5 = 1$ *gdw.* $|z| = 1$ *und* $\arg(z) \in \Phi$, *oder anders formuliert:* $z = \cos(\varphi) + \mathrm{i}\sin(\varphi)$ *für* $\varphi \in \Phi$.
Ist nun $w \in \mathbb{C}$ *beliebig und* $n \in \mathbb{N}^*$, *so können wir alle Lösungen von* $z^n = w$ *(also die* n*-ten* **Wurzeln** *von* w*) finden: Es muss* $|z| = \sqrt[n]{|w|}$ *und* $n \cdot \arg(z) = \arg(w)$ *(bis auf Vielfache von* 2π*) gelten. Das bedeutet*

$$z = \sqrt[n]{|w|} \cdot \Big(\cos\big(\frac{\arg(w)}{n} + \varphi\big) + \mathrm{i}\sin\big(\frac{\arg(w)}{n} + \varphi\big)\Big)$$

mit $\varphi = \frac{2k\pi}{n}$ *und* $k \in \{0, \ldots, n-1\}$.
Im nebenstehenden Bild ist w *mit* $|w| = 8$ *und* $\arg(w) = 75°$. *Die Gleichung* $z^3 = w$ *hat die drei Lösungen* z_1, z_2, z_3 *mit* $|z_i| = 2$, $\arg(z_1) = 25°$, $\arg(z_2) = 25° + 120°$ *und* $\arg(z_3) = 25° + 240°$.



[14] In der Schule werden sie meist ausgelassen, obwohl sie elementargeometrisch behandelt werden können und für die Berechnung der Ableitung von Sinus und Kosinus (also $(\sin x)' = \cos x$ und $(\cos x)' = -\sin x$ — das ist noch Schulstoff!) gebraucht werden.

Das Beispiel zeigt, dass man aus jeder komplexen Zahl beliebige Wurzeln ziehen kann.

Umrechnungsformeln

Zwischen Standard- und Polardarstellung gelten folgende Umrechnungsformeln: Es ist $a + b\,\mathrm{i} = r(\cos\varphi + \mathrm{i}\sin\varphi)$ *genau dann wenn*

$$a = r \cdot \cos\varphi \qquad b = r \cdot \sin\varphi$$

$$r = \sqrt{a^2 + b^2} \qquad \varphi = \begin{cases} \arccos\left(\frac{a}{\sqrt{a^2+b^2}}\right) & \textit{falls } b \geq 0 \\ -\arccos\left(\frac{a}{\sqrt{a^2+b^2}}\right) & \textit{falls } b < 0 \end{cases}$$

Für das Argument gibt es auch Umrechnungsformeln, die auf dem Arkustangens basieren, aber eine Fallunterscheidung hat man immer.

Im Hinblick auf die Umrechnungsformeln ist folgende Tabelle praktisch, wobei man noch $\forall\varphi \in \mathbb{R}\colon\ \sin(-\varphi) = -\sin\varphi$, $\cos(-\varphi) = \cos\varphi$ sowie $\sin(\varphi + \frac{\pi}{2}) = \cos\varphi$ beachten muss.

Spezielle Werte von Sinus und Kosinus

φ	0	$\frac{\pi}{12}$	$\frac{\pi}{8}$	$\frac{\pi}{6}$	$\frac{\pi}{4}$	$\frac{\pi}{3}$	$\frac{5\pi}{12}$	$\frac{\pi}{2}$
$\sin\varphi$	0	$\frac{\sqrt{6}-\sqrt{2}}{4}$	$\frac{\sqrt{2-\sqrt{2}}}{2}$	$\frac{1}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{\sqrt{3}}{2}$	$\frac{\sqrt{6}+\sqrt{2}}{4}$	1
$\cos\varphi$	1	$\frac{\sqrt{6}+\sqrt{2}}{4}$	$\frac{\sqrt{2+\sqrt{2}}}{2}$	$\frac{\sqrt{3}}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{1}{2}$	$\frac{\sqrt{6}-\sqrt{2}}{4}$	0

2 Lineare Gleichungssysteme

In der Schule lernten Sie, lineare Gleichungssysteme zu lösen (vielleicht nicht unter diesem Namen), allerdings vermutlich nur mit höchstens drei Unbekannten. Als Lösungsmethoden sind Ihnen vermutlich Einsetzungs-, Gleichsetzungs- und Additionsverfahren ein Begriff. Mathematisch sind dies drei unterschiedlich geschickte Darstellungsweisen für ein und dasselbe Verfahren. Hinzu kommt eine ineffiziente Notationsweise für lineare Gleichungssysteme. Zunächst werden wir die Voraussetzungen für eine bessere Notation schaffen und dann allgemeine Lösungsmethoden vorstellen.

Zum Vergleich: Bereits 100 n.Chr. löste man in China (Kapitel 8 des Buches Jiŭ Zhāng Suànshù) Gleichungen mit bis zu drei Unbekannten, verwendete dafür aber die im folgenden dargestellte Matrix-Notation und sogar das zuerst von Gauß in voller Allgemeinheit formulierte Eliminationsverfahren.

Im gesamten Kapitel sei $\mathbb{K}$ ein Körper. Seine Elemente nennen wir ***Skalare***.

2.1 Zeilen, Spalten, Matrizen

Definition 2.1. *Seien $m, n \in \mathbb{N}$.*

*a) Eine $(m \times n)$-***Matrix** *A über $\mathbb{K}$ besteht aus $m \cdot n$ Skalaren, aufgestellt in m Zeilen und n Spalten. Wir schreiben $\mathbb{K}^{m\times n}$ für die Menge aller $(m \times n)$-Matrizen und diese Notation impliziert stets, dass m, n natürliche Zahlen sind. Für $1 \leq i \leq m$ und $1 \leq j \leq n$, so schreibt man $A_{i,j}$ für den Eintrag von A an der Stelle (i, j), d.h. in der i-ten Zeile und der j-ten Spalte.*

b) Eine Matrix heißt **quadratisch***, wenn Zeilen- und Spaltenzahl übereinstimmen. Wir schreiben $M_n(\mathbb{K}) := \mathbb{K}^{n\times n}$.*

c) Eine $(n \times 1)$-Matrix nennen wir **Spaltenvektor***, eine $(1 \times n)$-Matrix* **Zeilenvektor***.*

d) Elemente von $\mathbb{K}^n$ schreiben wir normalerweise als Spaltenvektor, wir identifizieren also $\mathbb{K}^n$ mit $\mathbb{K}^{n\times 1}$; insbesondere implizieren wir auch mit der Notation $\mathbb{K}^n$, dass $n \in \mathbb{N}$ gilt.

e) Den Eintrag in der k-ten Zeile von $\vec{v} \in \mathbb{K}^n$ bezeichne ich als v_k, wobei wichtig ist, dass der Pfeilakzent wegfällt.

f) Schreibt man Spaltenvektoren $\vec{v}_1, ..., \vec{v}_n \in \mathbb{K}^m$ nebeneinander, so entsteht die Matrix $(\vec{v}_1, ..., \vec{v}_n) \in \mathbb{K}^{m\times n}$.

g) Aus $A \in \mathbb{K}^{m\times n}$ entsteht die **transponierte Matrix** *$A^\top \in \mathbb{K}^{n\times m}$ durch Tausch von Zeilen und Spalten: $\forall i \in \{1, ..., m\},\ j \in \{1, ..., n\}\colon\ A^\top_{j,i} := A_{i,j}$.*

Beispiel 2.2

- $A := \begin{pmatrix} 1 & 1 & 1 \\ 2 & 3 & 1 \\ 1 & -1 & 3 \end{pmatrix} \in M_3(\mathbb{R})$ *hat etwa die Einträge* $A_{3,2} = -1$ *und* $A_{2,3} = 1$.
- $\vec{b} := \begin{pmatrix} 1 \\ 3 \\ -1 \end{pmatrix} \in \mathbb{R}^3$ *hat etwa den Eintrag* $b_3 = -1$.
- $\begin{pmatrix} 1 & 3 & 7 & 4 \\ 3 & 1 & 2 & 9 \\ 8 & 0 & 7 & 3 \end{pmatrix}^\top = \begin{pmatrix} 1 & 3 & 8 \\ 3 & 1 & 0 \\ 7 & 2 & 7 \\ 4 & 9 & 3 \end{pmatrix} \in \mathbb{R}^{4\times 3}$. *Man beachte:* $(A^\top)^\top = A$.

Bemerkung 2.3 (Notationsprobleme)
Es gibt leider viele verschiedene Notationskonventionen für Vektoren und Matrizen. Ich erläutere hier die Gründe für die von mir verwendete Notation.

- *Eigentlich ist immer aus dem Kontext klar, ob man einen Skalar ($a \in \mathbb{K}$) oder Vektor ($v \in \mathbb{K}^n$) betrachtet. Aber im Anfängerbereich halte ich es für sinnvoll, durch die Notation eine Verwechslung von Skalaren und Vektoren zu vermeiden. Ich verwende dafür einen Pfeilakzent ($\vec{v}$), was auch in der Physik recht üblich ist.*

  Ebenfalls üblich ist Fettdruck für Vektoren ($\mathbf{v}$). Seltener findet man einfachen bzw. doppelten Unterstrich für Vektoren bzw. Matrizen ($\underline{v}$, $\underline{\underline{A}}$). Kowalsky verwendet Frakturschrift ($\mathfrak{A}$, $\mathfrak{x}$) bzw. Sütterlin bei handschriftlichen Notizen, auch ich habe das als Student so gelernt. Fischer hat keinen Notationsunterschied zwischen Vektoren und Skalaren.

- *Häufig betrachtet man eine Liste von Vektoren ($\vec{v}_1, \vec{v}_2, \vec{v}_3, ...$) Bitte beachten Sie, dass dabei ein Pfeilakzent auftritt! $\vec{v}_2$ ist der zweite Vektor in einer Liste, hingegen ist v_2 der zweite Eintrag in einem einzelnen Vektor $\vec{v}$. Der Pfeilakzent hilft, eine Verwechslung zu vermeiden. Fischer verwendet keinen Pfeilakzent — die Verwechslungsgefahr ist dadurch groß.*

  Nur selten ist eine Notation für einzelne Einträge eines Vektors nötig. Wenn ich den i-ten Eintrag in Vektor $\vec{v}_j$ notieren müsste, würde ich $v_{i,j}$ schreiben, um konsistent mit der Notation für Matrixeinträge zu sein.

  Im Prinzip sinnvoll, aber sehr unüblich wäre, den Index für die Vektoreinträge links unten zu platzieren, also ${}_iv_j$ für den i-ten Eintrag von $\vec{v}_j$.

  Teile der theoretischen Physik basieren auf einer Verallgemeinerung von Vektoren und Matrizen, so genannten **Tensoren**. *Die Position von Indizes hat bei diesen eine inhaltliche Bedeutung, auf die ich hier nicht näher eingehen möchte. In Tensorschreibweise würde man ${A^i}_j$ statt $A_{i,j}$ schreiben und der i-te Eintrag eines Vektors $\vec{v}$ würde als v^i notiert.*

- *Manche Bücher verwenden griechische Kleinbuchstaben für Skalare. Das halte ich hier für unnötig, da durch den Pfeilakzent ohnehin ein Notationsunterschied zwischen Vektoren und Skalaren besteht.*

2.2 Matrixarithmetik

Addition von Matrizen sowie Multiplikation einer Matrix mit einem Körperelement ist genau so, wie Sie es aus der Schule für Vektoren kennen:

Definition 2.4
Für $c \in \mathbb{K}$, $A, B \in \mathbb{K}^{m\times n}$ *definiert man* $A + B := \begin{pmatrix} A_{1,1}+B_{1,1} & \dots & A_{1,n}+B_{1,n} \\ \vdots & & \vdots \\ A_{m,1}+B_{m,1} & \dots & A_{m,n}+B_{m,n} \end{pmatrix}$ *und* $c \cdot A := \begin{pmatrix} cA_{1,1} & \dots & cA_{1,n} \\ \vdots & & \vdots \\ cA_{m,1} & \dots & cA_{m,n} \end{pmatrix}$ *(***Skalarmultiplikation**[15]*).*

Da Spalten- und Zeilenvektoren ebenfalls Matrizen sind, ist dadurch die Addition und Skalarmultiplikation auf $\mathbb{K}^n$ ebenfalls definiert.

Definition 2.5 (Matrixmultiplikation)
Für $A \in \mathbb{K}^{m\times n}$ *und* $B \in \mathbb{K}^{n\times p}$ *definiert man* $AB \in \mathbb{K}^{m\times p}$ *durch*

$$(AB)_{i,k} := A_{i,1}B_{1,k} + A_{i,2}B_{2,k} + A_{i,3}B_{3,k} + \cdots + A_{i,n}B_{n,k} = \sum_{j=1}^{n} A_{i,j}B_{j,k}\,.$$

Damit das Produkt AB existiert, muss A genau so viel Spalten haben, wie B Zeilen hat — andernfalls ist das Matrixprodukt nicht definiert!

Beispiel 2.6
$\left(\begin{smallmatrix} 1 & 0 \\ 3 & 1 \end{smallmatrix}\right)\left(\begin{smallmatrix} 0 & 2 & 4 \\ 1 & 3 & 5 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 1\cdot0+0\cdot1 & 1\cdot2+0\cdot3 & 1\cdot4+0\cdot5 \\ 3\cdot0+1\cdot1 & 3\cdot2+1\cdot3 & 3\cdot4+1\cdot5 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 0 & 2 & 4 \\ 1 & 9 & 17 \end{smallmatrix}\right)$.

Definition 2.7

a) $\mathbb{0} \in \mathbb{K}^{m\times n}$ *bezeichnet die* **Nullmatrix** *mit* m *Zeilen und* n *Spalten, deren Einträge alle Null sind. Analog definiert man den* **Nullvektor** $\vec{0} \in \mathbb{K}^n$.

b) $D \in M_n(\mathbb{K})$ *heißt* **Diagonalmatrix** $:\Leftrightarrow \forall i \neq j \in \{1, ..., n\}\colon D_{i,j} = 0$. *Für* $\gamma_1, ..., \gamma_n \in \mathbb{K}$ *sei* $\operatorname{diag}(\gamma_1, ..., \gamma_n) := \begin{pmatrix} \gamma_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \gamma_n \end{pmatrix} \in M_n(\mathbb{K})$.

c) Die **Einsmatrix**[16] *ist* $\mathbb{1}_n = \operatorname{diag}(1, ..., 1) \in M_n(\mathbb{K})$.

Lemma 2.8
Seien $A \in \mathbb{K}^{\ell\times m}$, $B \in \mathbb{K}^{m\times n}$. *Dann* $(AB)^\top = B^\top A^\top$.

[15] Beachte: Es gibt auch *Skalarprodukte*, aber das ist etwas anderes!
[16] Auch ***Einheitsmatrix*** genannt.

Beweis: $\forall i \in \{1,...,\ell\},\ j \in \{1,...,n\}$:
$(AB)^\top_{i,j} = (AB)_{j,i} = \sum_{k=1}^{m} A_{j,k}B_{k,i} = \sum_{k=1}^{m} B^\top_{i,k}A^\top_{k,j} = (B^\top A^\top)_{i,j}$. □

Aufgabe 2.9
In den folgenden Aussagen sei $\gamma, \delta \in \mathbb{K}$ und A, B, C seien Matrizen, für die die angegebenen Rechenoperationen definiert sind (d.h. die Zeilen- und Spaltenzahlen „passen").

a) Assoziativität: $(\gamma\delta)A = \gamma(\delta A)$, $(\gamma A)B = \gamma(AB)$ und $(AB)C = A(BC)$; $(A+B)+C = A+(B+C)$.

b) Kommutativität der Addition: $A+B = B+A$. Multiplikation ist i.A. nicht kommutativ! Jedoch gilt $A(\gamma B) = \gamma(AB)$.

c) Distributivität: $\gamma(A+B) = \gamma A + \gamma B$, $A(B+C) = AB + AC$, $(A+B)C = AC + BC$ und $(c+d)A = cA + dA$. Beispielsweise

$$(A(B+C))_{i,j} = \sum_k A_{i,k}(B_{k,j}+C_{k,j}) = \sum_k A_{i,k}B_{k,j} + A_{i,k}C_{k,j} = (AB+AC)_{i,j}$$

d) Für $A \in \mathbb{K}^{m\times n}$ gelten $A \cdot \mathbb{1}_n = A$, $\mathbb{1}_m \cdot A = A$ und $\mathbb{0} + A = A$. Falls $\mathbb{0}A$ bzw. $A\mathbb{0}$ definiert ist, ist das Ergebnis $\mathbb{0}$.

e) $\forall A \in \mathbb{K}^{m\times n}\colon\ \exists -A \in \mathbb{K}^{m\times n}\colon\ A + (-A) = \mathbb{0}$.

Folglich ist $M_n(\mathbb{K})$ mit Matrixaddition und -multiplikation ein Ring, der für $n > 1$ allerdings niemals nullteilerfrei und niemals kommutativ ist. Im Allgemeinen gibt es NICHT für jedes $A \in M_n(\mathbb{K})$ mit $A \neq \mathbb{0}$ eine Inverse Matrix $A^{-1} \in M_n(\mathbb{K})$ mit $AA^{-1} = \mathbb{1}_n$. Das wird noch Thema werden.

Beobachtung 2.10
a) Für $A \in \mathbb{K}^{m\times n}$ und $\vec{v}_1, ..., \vec{v}_k \in \mathbb{K}^n$ gilt $A \cdot (\vec{v}_1, ..., \vec{v}_k) = (A\vec{v}_1, ..., A\vec{v}_k)$.

b) Wenn wir die i-te Zeile eine Matrix M mit $\underline{M}_i$ bezeichnen, gilt $\forall A \in \mathbb{K}^{\ell\times m}$, $B \in \mathbb{K}^{m\times n}$ und $\forall i \in \{1,...,\ell\}\colon\ (\underline{A}_i)B = \underline{(AB)}_i$.

2.3 Lösungsräume

In der Schule betrachteten Sie Gleichungssysteme der Art

$$x + y + z = 1 \quad \text{(I)}$$
$$2x + 3y + z = 3 \quad \text{(II)}$$
$$x - y + 3z = -1 \quad \text{(III)}$$

Man hat endlich viele Unbekannte (hier: x, y, z) und endlich viele Gleichungen. Die linke Seite jeder Gleichung ist eine Summe von Termen mit jeweils genau

einer Unbekannten, die zudem in erster Potenz auftritt; die rechte Seite jeder Gleichung ist eine Zahl. Derartige Gleichungen heißen linear, insgesamt hat man also ein lineares Gleichungssystem. Eine Lösung des Gleichungssystems weist den Unbekannten Werte so zu, dass alle Gleichungen simultan gelöst werden.

Sehr wahrscheinlich wurden an Ihrer Schule keine Gleichungssysteme mit mehr also drei Unbekannten behandelt — was auch an der ungeschickten Notation liegt.

a) Verwendet man mehr Unbekannte, ist bald das Alphabet ausgeschöpft. Besser: Durchnummerierung der Unbekannten, etwa $x_1, x_2, ...$ (es besteht aber kein Grund, Unbekannte immer als x zu bezeichnen!!).

b) Die zu einem Term gehörende Variable ergibt sich aus seiner Position. Also ist es unsinnig, in jeder Gleichung und bei jedem Umformungsschritt alle Variablennamen hin zu schreiben.

Matrixmultiplikation ergibt eine kurze Notation für große Gleichungssysteme.

Definition 2.11. *Sei $A \in \mathbb{K}^{m\times n}$ und $\vec{b} \in \mathbb{K}^m$.*

a) $A\vec{x} = \vec{b}$ *heißt* **lineares Gleichungssystem** *mit* **Koeffizientenmatrix** A, **Inhomogenität** $\vec{b}$ *und* **Unbekannter** $\vec{x} \in \mathbb{K}^n$.

b) Ist die Inhomogenität Null, also $A\vec{x} = \vec{0}$, so heißt das lineare Gleichungssystem **homogen**.

c) $\mathrm{LR}(A; \vec{b}) := \{\vec{x} \in \mathbb{K}^n \mid A\vec{x} = \vec{b}\}$ *heißt* **Lösungsraum** *des linearen Gleichungssystems $A\vec{x} = \vec{b}$.*

d) Die aus den Spalten von A und zusätzlich $\vec{b}$ gebildete Matrix $(A \mid \vec{b}) \in \mathbb{K}^{m\times(n+1)}$ heißt **erweiterte Matrix**. *Durch sie ist das lineare Gleichungssystem eindeutig bestimmt.*

Wenn wir im folgenden Kapitel lineare Gleichungssysteme allgemein lösen, werden wir stets nur mit A und $\vec{b}$ bzw. mit $(A|\vec{b})$ arbeiten. **Bitte halten Sie nicht an der Schulnotation fest und unterlassen Sie es, ständig irgendwelche überflüssigen Variablennamen zu schreiben.**

Beispiel $\left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 2 & 3 & 1 & 3 \\ 1 & -1 & 3 & -1 \end{array}\right) \in \mathbb{R}^{3\times 4}$ beschreibt das obige Gleichungssystem.

Der Lösungsraum linearer Gleichungssysteme weist folgende strukturelle Eigenschaften auf, die Anlass für die Begriffsbildungen der linearen Algebra bieten.

Satz 2.12 (Algebr. Eigenschaften von Lösungsräumen)
Sei $A \in \mathbb{K}^{m\times n}$ *und* $\vec{b} \in \mathbb{K}^m$.

a) $\forall \vec{x}_1, \vec{x}_2 \in \mathrm{LR}(A;\, \vec{0})\colon\ \vec{x}_1 + \vec{x}_2 \in \mathrm{LR}(A;\, \vec{0})$.

b) $\forall c \in \mathbb{K},\ \vec{x} \in \mathrm{LR}(A;\, \vec{0})\colon\ c\vec{x} \in \mathrm{LR}(A;\, \vec{0})$.

c) Sei $\vec{x}_{\mathrm{inh}} \in \mathrm{LR}(A;\, \vec{b})$. *Dann* $\mathrm{LR}(A;\, \vec{b}) = \left\{\vec{x}_{\mathrm{inh}} + \vec{x}_{\mathrm{h}} \mid \vec{x}_{\mathrm{h}} \in \mathrm{LR}(A;\, \vec{0})\right\}$.

Beweis:
a) $A \cdot (\vec{x}_1 + \vec{x}_2) = A \cdot \vec{x}_1 + A \cdot \vec{x}_2 = \vec{0} + \vec{0} = \vec{0}$.

b) $A \cdot (c\vec{x}) = c(A\vec{x}) = c\vec{0} = \vec{0}$.

c) „$\subset$“: Sei $\vec{x} \in \mathrm{LR}(A;\, \vec{b})$. Zu zeigen: $\vec{x}_{\mathrm{h}} := \vec{x} - \vec{x}_{\mathrm{inh}} \in \mathrm{LR}(A;\, \vec{0})$:

$$A \cdot (\vec{x} - \vec{x}_{\mathrm{inh}}) = A\vec{x} - A\vec{x}_{\mathrm{inh}} = \vec{b} - \vec{b} \overset{\checkmark}{=} \vec{0}$$

„$\supset$“: Sei $\vec{x}_{\mathrm{h}} \in \mathrm{LR}(A;\, \vec{0})$. Zu zeigen: $\vec{x} := \vec{x}_{\mathrm{inh}} + \vec{x}_{\mathrm{h}} \in \mathrm{LR}(A;\, \vec{b})$:

$$A \cdot (\vec{x}_{\mathrm{inh}} + \vec{x}_{\mathrm{h}}) = A\vec{x}_{\mathrm{inh}} + A\vec{x}_{\mathrm{h}} = \vec{b} + \vec{0} \overset{\checkmark}{=} \vec{b}$$ □

Definition 2.13 (und Beobachtung)
Sei $k \in \mathbb{N}^*$. *Die* **Linearkombination** *von* $\vec{v}_1, ..., \vec{v}_k \in \mathbb{K}^n$ *mit* **Koeffizienten** $c_1, ..., c_k \in \mathbb{K}$ *ist die Summe* $\sum_{j=1}^{k} c_j\vec{v}_j$.

Linearkombinationen lassen sich als Matrixprodukt darstellen. Für $\vec{c} \in \mathbb{K}^k$ *ist nämlich* $(\vec{v}_1, ..., \vec{v}_k) \cdot \vec{c} = \sum_{j=1}^{k} c_j\vec{v}_j$.

Unser Ziel ist, nicht nur Lösungsräume zu berechnen, sondern sie möglichst sparsam zu beschreiben. Wir werden einige so genannte Basislösungen berechnen, so dass jede Lösung von $A\vec{x} = \vec{0}$ eine *eindeutige* Beschreibung als Linearkombination von Basislösungen besitzt; dies entspricht einem Koordinatensystem für $\mathrm{LR}(A;\, \vec{0})$. Wenn wir danach noch eine einzige Lösung von $A\vec{x} = \vec{b}$ explizit berechnen (oder nachweisen, dass es keine Lösung gibt), ist dadurch wegen des vorigen Satzes $\mathrm{LR}(A; \vec{b})$ berechnet.

2.3.1 Zeilenstufenform

Wir behandeln zuerst einen Spezialfall, der eine relativ einfache Ablesung des Lösungsraums erlaubt. Danach wird der allgemeine Fall durch das *gaußsche Eliminationsverfahren* auf den Spezialfall zurückgeführt.

Definition 2.14. *Es sei $A \in \mathbb{K}^{m\times n}$.*
Für $i = 1, ..., m$ sei j_i die Nummer der ersten Spalte, in der ein von Null verschiedener Eintrag in Zeile i steht. A ist in **Zeilenstufenform** *(kurz:* **ZSF***), wenn es ein $1 \leq r \leq m$ gibt, so dass die Zeilen $1, ..., r$ nicht Null sind, die Zeilen $r+1, ..., m$ Null sind, und $j_1 < j_2 < ... < j_r$. Man nennt j_i die* **Pivotspalte** *der Zeile i.*

Beispiele
$$\begin{pmatrix} \boxed{0} & 2 & 1 & 0 & 3 & 4 \\ 0 & 0 & 0 & \boxed{3} & 1 & 1 \\ 0 & 0 & 0 & 0 & \boxed{3} & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \text{ und } \begin{pmatrix} \boxed{3} & 4 & 5 \\ 0 & \boxed{2} & 3 \\ 0 & 0 & \boxed{2} \end{pmatrix}$$

Ist $A \in \mathbb{K}^{m\times n}$ in ZSF, so berechnet man $\mathrm{LR}(A; \vec{0})$, indem man diejenigen Unbekannten frei wählt, die nicht den Pivotspalten der Matrix entsprechen, und dann nach den übrigen Unbekannten schrittweise „von unten nach oben“ auflöst.

Definition 2.15
Sei $A \in \mathbb{K}^{m\times n}$ in ZSF mit Pivotspalten $j_1 < j_2 < \cdots < j_r$. Sei $\vec{b} \in \mathbb{K}^m$. Wir nennen $x_{j_1}, x_{j_2}, ..., x_{j_r}$ **gebundene Variablen** *und die anderen Komponenten von $\vec{x}$* **freie Variablen** *des linearen Gleichungssystems $A\vec{x} = \vec{b}$.*

Problem 2.16 (Lösungsraum für Matrizen in Zeilenstufenform)
Sei $A \in \mathbb{K}^{m\times n}$ in ZSF mit Pivotspalten $j_1 < j_2 < \cdots < j_r$. Sei $\vec{b} \in \mathbb{K}^m$.

Wenn es $i \in \{r+1, ..., m\}$ mit $b_i \neq 0$ gibt, so ist $\mathrm{LR}(A; \vec{b}) = \emptyset$. Andernfalls ist durch jede Wahl von Werten aus $\mathbb{K}$ für die freien Variablen eine Lösung von $A\vec{x} = \vec{b}$ eindeutig bestimmt.

Lösung:
Wenn $\vec{b}$ in einer der Zeilen $r+1, r+2, ..., m$ nicht Null ist, so ist die betreffende Gleichung nicht erfüllbar (auf der linken Seite steht Null, auf der rechten nicht Null). Andernfalls sind die letzten $m - r$ Gleichungen einfach $0 = 0$, sind also sicher erfüllt. Wir streichen alle Nullzeilen und konzentrieren uns auf die obersten r Zeilen.

Wir betrachten Zeile i für $i = r, r-1, ..., 1$ (also von unten nach oben). In der i-ten Gleichung tritt genau eine noch nicht bestimmte Unbekannte auf, nämlich die gebundene Variable x_{j_i}. Alle $x_{j_i+1}, ..., x_n$ wurden als freie Variable gewählt oder bereits vorher als gebundene Variable berechnet.

Wir lösen die i-te Gleichung nach x_{j_i} auf (alle anderen Variablen auf die rechte Seite bringen, durch A_{i,j_i} dividieren) und können somit x_{j_i} einen eindeutigen Wert zuweisen, durch den die i-te Gleichung erfüllt ist. Durch diese so genannte ***Rückwärtssubstitution*** berechnet man die Werte aller gebundenen Variablen in Abhängigkeit von den freien Variablen.

Formell kann man die Auflösung von Gleichung i wie folgt schreiben: Gleichung i ist $A_{i,j_i} \cdot x_{j_i} + A_{i,j_i+1} \cdot x_{j_i+1} + \cdots + A_{i,n} \cdot x_n = b_i$, mit $A_{i,j_i} \neq 0$. Wenn also $x_{j_i+1}, x_{j_i+2}, \ldots, x_n$ gegeben sind, dann ist $x_{j_i} = \frac{1}{A_{i,j_i}} \left(b_i - \sum_{k=j_i+1}^{n} A_{i,k} x_k \right)$. □

Korollar 2.17 (und Definition)
Sei $A \in \mathbb{K}^{m \times n}$ in ZSF mit r Pivotspalten. Es gibt also $n-r$ freie und r gebundene Variablen.

a) *Es sei x_j eine freie Variable. Setze $x_j := 1$ und setze alle anderen freien Variablen auf Null. Dann liefert Problem 2.16 die Lösung $\vec{\beta}_j \in \mathrm{LR}(A; , \vec{0})$. Wir bezeichnen sie als* **Basislösung**. *Es sei $B_A \in \mathbb{K}^{n \times (n-r)}$ die Matrix, deren Spalten aus den Basislösungen bestehen.*

b) *Sei $\vec{b} \in \mathbb{K}^m$ mit $b_{r+1} = ... = b_m = 0$. Setzt man sämtliche freien Variablen auf Null, so liefert Problem 2.16 die Lösung $\vec{x}_{\text{spez}} \in \mathrm{LR}(A; \vec{b})$. Dann gilt*

$$\mathrm{LR}(A; \vec{b}) = \{\vec{x}_{\text{spez}} + B_A \cdot \vec{c} \mid \vec{c} \in \mathbb{K}^{n-r}\}$$

c) *Spezialfälle:*

- *Ist $\vec{b} = \vec{0}$, dann ist $\vec{x}_{\text{spez}} = \vec{0}$ und $\mathrm{LR}(A; \vec{0})$ besteht aus allen Linearkombinationen der Basislösungen.*
- *Ist $r = n$, dann ist $\mathrm{LR}(A; \vec{b}) = \{\vec{x}_{\text{spez}}\}$.*

Beweis:
Sei $\vec{c} \in \mathbb{K}^{n-r}$. Gemäß Satz 2.12 ist $B_A \cdot \vec{c} \in \mathrm{LR}(A; \vec{0})$, denn jede Basislösung liegt in $\mathrm{LR}(A; \vec{0})$, also auch alle ihre Linearkombinationen. Ferner ist $\vec{x}_{\text{spez}} + B_A \cdot \vec{c} \in \mathrm{LR}(A; \vec{b})$.

Die freien Variablen in $\vec{x}_{\text{spez}}$ sind 0. In einer Basislösung ist jeweils eine freie Variable 1, alle anderen 0. Daher sind die Werte der $n - r$ freien Variablen in $\vec{x}_{\text{spez}} + B_A \cdot \vec{c}$ gleich $c_1, ..., c_{n-r}$. Durch die Werte der freien Variablen ist eine Lösung aber eindeutig bestimmt. □

Es bleibt zu klären, wie man ein gegebenes Gleichungssystem in ZSF umformt, ohne den Lösungsraum zu ändern. Das ist Thema eines späteren Abschnitts.

Beispiel 2.18. $A = \left(\begin{array}{cccccc} \boxed{0} & 2 & 1 & 0 & 3 & 4 \\ 0 & \boxed{0} & 0 & 3 & 1 & 1 \\ 0 & 0 & \boxed{0} & \boxed{0} & 3 & 1 \\ 0 & 0 & 0 & 0 & \boxed{0} & 0 \\ 0 & 0 & 0 & 0 & 0 & \boxed{0} \end{array}\right) \in \mathbb{R}^{5 \times 6}$. *Falls* $\vec{b} = \begin{pmatrix} 2 \\ -1 \\ 3 \\ 0 \\ 1 \end{pmatrix}$, *so ist die letzte Gleichung nicht erfüllbar, also* $\mathrm{LR}(A; \vec{b}) = \emptyset$.

Sei nun $\vec{b} = \begin{pmatrix} 2 \\ -1 \\ 3 \\ 0 \\ 0 \end{pmatrix}$. *Diesmal sind alle Gleichungen erfüllbar. Wir haben* $r = 3$ *und streichen die letzten beiden Zeilen. Somit sind die Gleichungen*

$$\left(\begin{array}{cccccc|c} 0 & 2 & 1 & 0 & 3 & 4 & 2 \\ 0 & 0 & 0 & 3 & 1 & 1 & -1 \\ 0 & 0 & 0 & 0 & 3 & 1 & 3 \end{array}\right)$$

oder, wenn wir ausnahmsweise Schulschreibweise verwenden:

$$\begin{array}{rcrcrcrcr} 2x_2 & + & x_3 & & & + & 3x_5 & + & 4x_6 & = & 2 \\ & & & & 3x_4 & + & x_5 & + & x_6 & = & -1 \\ & & & & & & 3x_5 & + & x_6 & = & 3 \end{array}$$

Die gebundenen Variablen sind x_2, x_4, x_5, *und die freien sind* x_1, x_3, x_6.

Die dritte Gleichung ergibt $x_5 = \frac{3-x_6}{3} = 1 - \frac{1}{3}x_6$. *Die zweite ergibt* $x_4 = \frac{-1-x_6-x_5}{3} = -\frac{1}{3} - \frac{1}{3}x_6 - \frac{1}{3}\left(1 - \frac{1}{3}x_6\right) = -\frac{1}{3} - \frac{1}{3} + (\frac{1}{9} - \frac{1}{3})x_6 = -\frac{2}{3} - \frac{2}{9}x_6$. *Die erste Gleichung ergibt* $x_2 = \frac{2-4x_6-3x_5-x_3}{2} = 1 - 2x_6 - \frac{3}{2}\left(1 - \frac{1}{3}x_6\right) - \frac{1}{2}x_3 = -\frac{1}{2} + (\frac{1}{2} - 2)x_6 - \frac{1}{2}x_3 = -\frac{1}{2} - \frac{3}{2}x_6 - \frac{1}{2}x_3$.

Wir erhalten die Basislösungen $\vec{\beta}_1 := \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$, $\vec{\beta}_3 := \begin{pmatrix} 0 \\ -1/2 \\ 1 \\ 0 \\ 0 \end{pmatrix}$, $\vec{\beta}_6 := \begin{pmatrix} 0 \\ -3/2 \\ 0 \\ -2/9 \\ -1/3 \\ 1 \end{pmatrix}$

sowie die spezielle Lösung $\vec{x}_{spez} := \begin{pmatrix} 0 \\ -1/2 \\ 0 \\ -2/3 \\ 1 \\ 0 \end{pmatrix}$. *Die roten Einträge entsprechen den freien Variablen, deren Werte jeweils eingesetzt wurden; die anderen Einträge entsprechen den gebundenen Variablen, sie wurden berechnet.*

Den Lösungraum können wir nun auf verschiedene Arten angeben.

$$\begin{aligned} \mathrm{LR}(A; \vec{b}) &= \left\{ \begin{pmatrix} 0 \\ -1/2 \\ 0 \\ -2/3 \\ 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1/2 & -3/2 \\ 0 & 1 & 0 \\ 0 & 0 & -2/9 \\ 0 & 0 & -1/3 \\ 0 & 0 & 1 \end{pmatrix} \cdot \vec{c} \;\middle|\; \vec{c} \in \mathbb{R}^3 \right\} \\ &= \left\{ \begin{pmatrix} 0 \\ -1/2 \\ 0 \\ -2/3 \\ 1 \\ 0 \end{pmatrix} + x_1 \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} 0 \\ -1/2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + x_6 \begin{pmatrix} 0 \\ -3/2 \\ 0 \\ -2/9 \\ -1/3 \\ 1 \end{pmatrix} \;\middle|\; x_1, x_3, x_6 \in \mathbb{R} \right\} \\ &= \left\{ \begin{pmatrix} x_1 \\ -\frac{1}{2} - \frac{1}{2}x_3 - \frac{3}{2}x_6 \\ x_3 \\ -\frac{2}{3} - \frac{2}{9}x_6 \\ 1 - \frac{1}{3}x_6 \\ x_6 \end{pmatrix} \;\middle|\; x_1, x_3, x_6 \in \mathbb{R} \right\}. \end{aligned}$$

Die erste Darstellungsweise halte ich unter strukturellen Aspekten für die beste.

Bemerkung 2.19

Man beachte, dass wir in der Lösung von Problem 2.16 durch $A_{i,j_i} \neq 0$ *dividierten. Wir sind also darauf angewiesen, dass wir überhaupt dividieren* können. *Wir können, weil* $\mathbb{K}$ *ein Körper ist.*

2.4 Das Gaußsche Eliminationsverfahren

Das hier beschriebene Eliminationsverfahren ist von herausragender Bedeutung für die Mathematik. Das liegt daran, dass sich viele Aufgabenstellungen der Mathematik auf die Lösung von linearen Gleichungssystemen $A \cdot \vec{x} = \vec{b}$ mit $A \in \mathbb{K}^{m \times n}$ zurückführen lassen und das Gaußsche Eliminationsverfahren die Lösung erlaubt, indem es das Gleichungssystem in Zeilenstufenform umwandelt.

Algorithmus 2.20 (Das Gaußsche Eliminationsverfahren)
Durch **Zeilenoperationen** *wird ein vorgegebenes $A \in \mathbb{K}^{m \times n}$ auf Zeilenstufenform gebracht. Zu Beginn sei $i := 1$.*

1) Abbruch, falls die Zeilen $i, ..., m$ alle Null sind. Sonst: Sei $j \in \{1, ..., n\}$ minimal, so dass ein $A_{\ell,j} \neq 0$ mit $\ell \in \{i, ..., m\}$ existiert. Erzwinge $A_{i,j} \neq 0$, indem ggf. Zeilen i, ℓ für ein $\ell > i$ vertauscht werden. ***Anmerkung:*** *Hier besteht manchmal eine Wahlmöglichkeit.*

2) Optional: Ersetze Zeile i durch ihr $1/A_{i,j}$-Faches ($\rightsquigarrow A_{i,j} = 1$).

3) Für alle $\ell \in \{i+1, ..., m\}$: Ziehe das $\frac{A_{\ell,j}}{A_{i,j}}$-Fache der Zeile i von Zeile ℓ ab ($\rightsquigarrow A_{\ell,j} = 0$). Optional: Führe dies auch für alle $\ell \in \{1, ..., i-1\}$ durch.

4) Erhöhe i um 1 und gehe zurück zu Schritt 1).

Beispiel Gauß-Algorithmus für $A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 6 & 10 & 11 \\ -1 & -3 & -1 & 1 \\ 1 & 1 & 7 & 7 \end{pmatrix}$: Es ist $A_{1,1} \neq 0$. Wir ziehen die erste Zeile dreimal von der zweiten und einmal von der vierten Zeile ab und addieren die erste zur dritten: $\begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 0 & 1 & -1 \\ 0 & -1 & 2 & 5 \\ 0 & -1 & 4 & 3 \end{pmatrix}$. Jetzt $i := 2$: Wegen $A_{1,2} = 0$ vertauschen wir Zeilen 2 und 3 von A: $\begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & -1 & 2 & 5 \\ 0 & 0 & 1 & -1 \\ 0 & -1 & 4 & 3 \end{pmatrix}$. Jetzt ziehen wir die zweite von der vierten Zeile ab: $\begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & -1 & 2 & 5 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 2 & -2 \end{pmatrix}$ und wiederholen mit $i := 3$. Diesmal ist $j = 3$. Indem wir das Doppelte der dritten Zeile von der vierten abziehen, erhalten wir die Zeilenstufenform $\begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & -1 & 2 & 5 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$. Führt man zudem die optionalen Schritte aus, erhält man $\begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & -2 & -5 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$, $\begin{pmatrix} 1 & 0 & 7 & 14 \\ 0 & 1 & -2 & -5 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ und schließlich $\begin{pmatrix} 1 & 0 & 0 & 21 \\ 0 & 1 & 0 & -7 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$.

Der Gauß-Algorithmus verwendet nur die folgenden drei Typen von ***Zeilenoperationen***:

I) Zeilen i und ℓ miteinander vertauschen ($i \neq \ell$).

II) Zeile i mit $c \in \mathbb{K} \setminus \{0\}$ multiplizieren.

III) Das c-fache von Zeile i zu Zeile ℓ addieren ($c \in \mathbb{K}$, $i \neq \ell$).

Lemma 2.21

a) Seien $M \in \mathbb{K}^{m\times n}$, $N \in \mathbb{K}^{n\times \ell}$ Wenn M' aus M durch eine Zeilenoperation enteht, dann entsteht $M' \cdot N$ aus $M \cdot N$ durch dieselbe Zeilenoperation.

b) Aus $\mathbb{1}_m$ entstehe durch eine Abfolge von Zeilenoperationen die Matrix $X \in M_m(\mathbb{K})$. Ist $A \in \mathbb{K}^{m\times n}$, so entsteht $X \cdot A$ aus A durch dieselbe Abfolge von Zeilenoperationen.

Beweis:

a) Mit der Notation für die i-te Zeile aus Beobachtung 2.10.b) gilt $\underline{M \cdot N_i} = \underline{M_i} \cdot N$. Für die drei Typen der Zeilenoperationen folgt:

I) Sei $\underline{M'_i} = \underline{M_\ell}$ und $\underline{M'_\ell} = \underline{M_i}$. Dann $\underline{M' \cdot N_i} = \underline{M'_i} \cdot N = \underline{M_\ell} \cdot N = \underline{M \cdot N_\ell}$. Analog $\underline{M' \cdot N_\ell} = \underline{M \cdot N_i}$.

II) Sei $c \in \mathbb{K}^*$ und $\underline{M'_i} = c \cdot \underline{M_i}$. Dann $\underline{M' \cdot N_i} = (c \cdot \underline{M_i}) \cdot N = c \cdot \underline{M \cdot N_i}$.

III) Sei $c \in \mathbb{K}$ und $\underline{M'_\ell} = \underline{M_\ell} + c \cdot \underline{M_i}$. Dann $\underline{M' \cdot N_\ell} = \underline{M'_\ell} \cdot N = (\underline{M_\ell} + c \cdot \underline{M_i}) \cdot N = \underline{M_\ell} \cdot N + c \cdot \underline{M_i} \cdot N = \underline{M \cdot N_\ell} + c \cdot \underline{M \cdot N_i}$.

b) $\mathbb{1}_m \rightsquigarrow X$ durch eine Zeilenoperationen $\overset{\text{a)}}{\Rightarrow}$ $\mathbb{1}_m A = A \rightsquigarrow XA$ ebenso. □

Korollar 2.22

Sei $A \in \mathbb{K}^{m\times n}$, $\vec{b} \in \mathbb{K}^m$. Das Gauß-Verfahren bringt die erweiterte Matrix $B := (A|\vec{b})$ nach endlich vielen Schritten auf eine Zeilenstufenform $B' = (A'|\vec{b'})$ mit $A' \in \mathbb{K}^{m\times n}$ und $\vec{b'} \in \mathbb{K}^m$, und es gilt $\mathrm{LR}(A; \vec{b}) = \mathrm{LR}(A'; \vec{b'})$.

Beweis:

Schritte 1)–4) erfordern jeweils nur endlich viele Rechenoperationen und werden nacheinander auf Zeilen 1 bis $m-1$ angewandt. Spätestens dann erfolgt Abbruch.

Bei Anwendung der Schritte auf Zeile i sei $A_{i,j}$ das erste von Null verschiedene Element. Mit Induktion zeigt man: Spalten $1, ..., j-1$ sind auch unterhalb von Zeile i Null. Nach Anwendung von Zeilenoperationen sind sie weiterhin Null und zudem gilt $\forall \ell \in \{i+1, ..., m\}: A_{\ell,j} = 0$. Also entsteht eine Zeilenstufenform.

$X \in M_m(\mathbb{K})$ entstehe aus $\mathbb{1}_m$ durch die verwendeten Zeilenoperationen. Aus dem vorigen Lemma folgt $(A'|\vec{b'}) = X \cdot (A|\vec{b}) = (XA \mid X\vec{b})$, d.h. $A' = XA$ und $\vec{b'} = X\vec{b}$. Jede Zeilenoperation lässt sich durch eine Zeilenoperation rückgängig machen. Daher gibt es ein $Y \in M_n(\mathbb{K})$ mit $YA' = A$ und $Y\vec{b'} = \vec{b}$.
$\mathrm{LR}(A; \vec{b}) \subset \mathrm{LR}(A'; \vec{b'})$: $\vec{x} \in \mathrm{LR}(A; \vec{b}) \Rightarrow A\vec{x} = \vec{b} \Rightarrow XA\vec{x} = X\vec{b} \Rightarrow A'\vec{x} = \vec{b'}$.
$\mathrm{LR}(A'; \vec{b'}) \subset \mathrm{LR}(A; \vec{b})$: $\vec{x} \in \mathrm{LR}(A'; \vec{b}) \Rightarrow A'\vec{x} = \vec{b'} \Rightarrow YA'\vec{x} = Y\vec{b'} \Rightarrow A\vec{x} = \vec{b}$.□

Beispiel 2.23 (und Definition)

Führt man im Gauß-Algorithmus auch alle optionalen Schritte durch, nennt man dies den **Gauß-Jordan-Algorithmus**[17]*. Es entsteht eine Matrix in Zeilenstufenform, bei der zusätzlich jede Pivotspalte* genau ein *von Null verschiedenes Element enthält, und dieses hat den Wert* 1 *(***reduzierte Zeilenstufenform***).*

[17] Wilhelm Jordan [1842–1899] war ein deutscher Geodät.

Sei $A = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 6 & 10 \\ -1 & -3 & -1 \\ 1 & 1 & 7 \end{pmatrix}$ *und* $\vec{b} = \begin{pmatrix} 4 \\ 11 \\ 1 \\ 7 \end{pmatrix}$. *Dann* $B := (A|\vec{b}) = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 6 & 10 & 11 \\ -1 & -3 & -1 & 1 \\ 1 & 1 & 7 & 7 \end{pmatrix}$. *Wir sahen im obigen Beispiel, dass der Gauß-Algorithmus* $B' = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & -1 & 2 & 5 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ *bzw. der Gauß-Jordan-Algorithmus* $B' = \begin{pmatrix} 1 & 0 & 0 & 21 \\ 0 & 1 & 0 & -7 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ *ergibt, und man liest ab:* $\mathrm{LR}(A; \vec{b}) = \{\begin{pmatrix} 21 \\ -7 \\ -1 \end{pmatrix}\}$.

Es ist klar, dass ein Gleichungssystem in reduzierter Zeilenstufenform besonders einfach lösbar ist, allerdings wiegt dies nicht immer den zusätzlichen Aufwand für die optionalen Schritte auf. Rechnet man gerundet (z.B. in der Numerik) kann die Genauigkeit sehr wohl davon abhängen, ob man optionale Schritte durchführt und in welcher Reihenfolge man eliminiert.

3 Grundbegriffe der linearen Algebra

Auch in diesem Kapitel sei $\mathbb{K}$ ein Körper. Wir werden nun die Rechnungen des vorigen Kapitels begrifflich erfassen.

- (Unter-)Vektorräume: Abstrahiert von $\mathbb{K}^{m\times n}$ und Lösungsräumen homogener linearer Gleichungssysteme und ist dann auch auf die Menge der stetigen Funktionen und auf den Übergang von $\mathbb{Q}$ nach $\mathbb{R}$ nach $\mathbb{C}$ anwendbar.
- Lineare Abbildungen: Abstrahiert von der Matrixmultiplikation und ist dann auch auf Differential- und Integralrechnung sowie geometrische Abbildungen anwendbar.
- Basen: Abstrahiert von Koordinatensystemen sowie von der „sparsamen" Erzeugung von Lösungsräumen mittels Linearkombinationen; ist eine wichtige Grundlage für rechnerische Methoden.

3.1 Vektorräume

Vektorraumaxiome

Es sei $\mathbb{K}$ ein Körper mit Einselement $1 \in \mathbb{K}$. Ein $\mathbb{K}$–**Vektorraum** *V (oder deutlicher $(V, +, \cdot)$) besteht aus einer Menge V mit*

- *einer inneren Verknüpfung $+$ auf V (***Vektor-Addition***) zusammen mit einem Element $\vec{0} \in V$, so dass gilt*

  (V1) Kommutativ: $\vec{u} + \vec{v} = \vec{v} + \vec{u}$ für alle $\vec{u}, \vec{v} \in V$

  (V2) Assoziativ: $\vec{u} + (\vec{v} + \vec{w}) = (\vec{u} + \vec{v}) + \vec{w}$ für alle $\vec{u}, \vec{v}, \vec{w} \in V$

  (V3) Nullvektor: Für jedes $\vec{v} \in V$ gilt $\vec{v} + \vec{0} = \vec{v}$.

  (V4) Negation: Zu jedem $\vec{v} \in V$ gibt es $-\vec{v} \in V$ mit $\vec{v} + (-\vec{v}) = \vec{0}$.

- *einer* **Skalarmultiplikation**[18] *$\cdot\colon \mathbb{K} \times V \to V$, $(\lambda, \vec{v}) \mapsto \lambda\vec{v}$ (auch $\lambda \cdot \vec{v}$ geschrieben), so dass gilt:*

  (V5) Assoziativität: $(\lambda\mu)\vec{v} = \lambda(\mu\vec{v})$ für alle $\lambda, \mu \in \mathbb{K}$ und für alle $\vec{v} \in V$;

  (V6) Eins: $1\vec{v} = \vec{v}$ für alle $\vec{v} \in V$.

- *Addition in $\mathbb{K}$ ebenso wie Addition in V erfüllen mit der Skalarmultiplikation ein Distributivgesetz. Also für alle $\lambda, \mu \in \mathbb{K}$ und $\vec{u}, \vec{v} \in V$:*

$$\text{(V7)}\quad \lambda(\vec{u} + \vec{v}) = \lambda\vec{u} + \lambda\vec{v} \qquad \text{(V8)}\quad (\lambda + \mu)\vec{v} = \lambda\vec{v} + \mu\vec{v}\,.$$

Man verwendet für $+$ bzw. $\cdot$ in $\mathbb{K}$ und V meist die gleichen Symbole. Man sollte sich des Unterschieds bewusst sein!

[18] Das ist i.A. keine innere Verknüpfung!

Beispiel 3.1

a) Herkömmliche Vektoren in der Ebene, oder im Raum. Allgemeiner: $\mathbb{K}^{m\times n}$ *sowie die Lösungsmengen homogener linearer Gleichungssysteme.*

b) Kräfte in der Physik.

c) $\mathrm{Abb}(\mathbb{R},\mathbb{R})$ *mit* **punktweiser** *Addition und Skalarmultiplikation, d.h. für alle* $f,g\colon \mathbb{R}\to\mathbb{R}$ *und* $\lambda\in\mathbb{R}$ *sind* $f+g\colon \mathbb{R}\to\mathbb{R}$ *und* $\lambda f\colon \mathbb{R}\to\mathbb{R}$ *definiert durch*

$$(f+g)(x) := f(x)+g(x) \qquad\qquad (\lambda f)(x) := \lambda\cdot f(x)\,.$$

d) Der Vektorraum $\mathbb{R}[X]$ *aller (reellen) Polynome. Der Vektorraum* $\mathbb{R}[X]_n$ *aller Polynome von Grad* $\le n$*.*

e) Man sollte sich bei der Arbeit mit Begriffen stets auch fragen, ob die Axiome des Begriffs vielleicht redundant sind, also ob ein Axiom aus den anderen Axiomen folgt.

Zum Nachweis der Redundanzfreiheit sucht man nach „pathologischen Beispielen", in denen ein Axiom verletzt ist und alle anderen gelten. Hier zeige ich, dass (V6) nicht redundant ist:

Sei $V=\mathbb{R}^2$ *mit der üblichen Addition, aber einer neuen Skalarmultiplikation* ☺*:* $\forall\lambda,x,y\in\mathbb{R}\colon\ \lambda ☺ \binom{x}{y} := \binom{0}{0}$*. Dann gelten alle Axiome außer (V6).*

Lemma 3.2

Sei V *ein* $\mathbb{K}$*–Vektorraum.*

a) Für alle $\vec v\in V$ *und für alle* $\lambda\in\mathbb{K}$ *gilt* $0\cdot\vec v=\vec 0=\lambda\cdot\vec 0$*.*

b) Für alle $\vec v\in V$ *und für alle* $\lambda\in\mathbb{K}$ *gilt* $(-\lambda)\vec v=-(\lambda\vec v)=\lambda(-\vec v)$*.*

c) $\forall\lambda\in\mathbb{K}$ *und* $\forall\vec v\in V$*: Aus* $\lambda\vec v=\vec 0$ *folgt* $\lambda=0$ *oder* $\vec v=\vec 0$*.*

Beweis:

a) $\lambda\vec v+\vec 0=\lambda\vec v=(\lambda+0)\vec v=\lambda\vec v+0\vec v$. Also $0\vec v=\vec 0$ durch Addition von $-\lambda\vec v$ auf beiden Seiten. Außerdem ist $\lambda\vec 0+\vec 0=\lambda\vec 0=\lambda(\vec 0+\vec 0)=\lambda\vec 0+\lambda\vec 0$, daher $\lambda\vec 0=\vec 0$.

b) $\lambda\vec v+(-\lambda)\vec v=(\lambda-\lambda)\vec v=0\vec v=\vec 0$, und $\lambda\vec v+\lambda(-\vec v)=\lambda(\vec v-\vec v)=\lambda\vec 0=\vec 0$.

c) Wir zeigen: Ist $\lambda\vec v=\vec 0$ aber $\lambda\neq 0$, dann gilt $\vec v=\vec 0$. Wegen $\lambda\neq 0$ existiert $\frac{1}{\lambda}\in\mathbb{K}$. Dann $\vec 0=\frac{1}{\lambda}\vec 0=\frac{1}{\lambda}(\lambda\vec v)=\left(\frac{1}{\lambda}\lambda\right)\vec v=1\vec v=\vec v$. □

Notation 3.3

Sei V *ein* $\mathbb{K}$*-Vektorraum; dabei ist stets implizit vorausgesetzt, dass* $\mathbb{K}$ *ein Körper ist. Für* $\vec u,\vec v\in V$ *schreiben wir* $\vec u-\vec v:=\vec u+(-\vec v)$*.*

3.2 Untervektorräume

Definition 3.4
Sei V ein $\mathbb{K}$-Vektorraum. $U \subseteq V$ heißt **Untervektorraum** *von V (Notation: $U \leq V$), genau dann wenn folgendes gilt:*

(U1) $\vec{0} \in U$;

(U2) Für alle $\vec{u}, \vec{w} \in U$ *ist auch* $\vec{u} + \vec{w} \in U$;

(U3) Für alle $\vec{u} \in U$ *und alle* $\lambda \in \mathbb{K}$ *ist* $\lambda\vec{u} \in U$.

Beispiel 3.5

a) *Der Lösungsraum eines homogenen linearen Gleichungssystems über $\mathbb{K}$ mit n Variablen ist ein Untervektorraum von $\mathbb{K}^n$.*

b) *Für $d \in \mathbb{N}$ ist die Menge $\mathbb{R}[X]_d$ aller Polynome vom Grad $\leq d$ ein Untervektorraum von $\mathbb{R}[X]$.*

Lemma 3.6
Sei V ein $\mathbb{K}$-Vektorraum.

a) *$U \subseteq V$ ist ein Untervektorraum gdw. er ist ein $\mathbb{K}$-Vektorraum mit den von V „geerbten“ Verknüpfungen.*

b) *$U \subseteq V$ ist genau dann ein Untervektorraum, wenn $U \neq \emptyset$ und außerdem $\lambda\vec{u} + \mu\vec{v} \in U$ für alle $\lambda, \mu \in \mathbb{K}$, $\vec{u}, \vec{v} \in U$.*

Beweis:

a) „$\Leftarrow$“ ist klar. Für „$\Rightarrow$“: Diejenigen Axiome (V1)–(V8), in denen kein Existenzquantor vorkommt, gelten für *alle* Teilmengen von V. (V3) gilt für U wegen (U1). Nach Lemma 3.2.b) gilt $\forall \vec{u} \in U: \ -\vec{u} = (-1)\vec{u}$, und dies liegt nach (U3) in U; also gilt (V4) für U.

b) Untervektorraum $\Rightarrow$ Bedingungen: $U \neq \emptyset$ wegen (U1). Wegen (U2) liegen $\lambda\vec{u}$, $\mu\vec{v}$ in U. Also $\lambda\vec{u} + \mu\vec{v} \in U$ wegen (U3).

Bedingungen $\Rightarrow$ Untervektorraum: Wegen $U \neq \emptyset$ gibt es ein $\vec{u}_0 \in U$. Dann ist $\vec{0} = \vec{u}_0 - \vec{u}_0 = 1\vec{u}_0 + (-1)\vec{u}_0$ ein Element von U. Sind $\vec{u}, \vec{w} \in U$, dann liegen auch $\vec{u} + \vec{w} = 1\vec{u} + 1\vec{w}$ und $\lambda\vec{u} = \lambda\vec{u} + 1 \cdot \vec{0}$ in U. □

Wir formulieren den Begriff der Linearkombination neu.

Notation 3.7. *Seien M und I Mengen.*

a) *Eine* **Familie** $(a_i)_{i \in I} \subset M$ *mit* **Indexmenge** I *ist eine Abbildung $I \ni i \mapsto a_i \in M$. Im Fall $I = \{1, ..., n\}$ schreiben wir auch $[a_1, a_2, ..., a_n] \subset M$ statt $(a_i)_{i \in \{1,...,n\}} \subset M$.*

b) Ist $(c_i)_{i\in I} \subset \mathbb{K}$ so, dass $c_i \neq 0$ nur für endlich viele $i \in I$ gilt, dann sagt man $c_i = 0$ für **fast alle** *$i \in I$.*

Da in der Abbildung $I \to M$ Elemente von M mehrfach getroffen werden können, entspricht eine Familie in M einer „Teilmenge Wiederholungen".

Definition 3.8
Sei V ein $\mathbb{K}$-Vektorraum und $S = (\vec{v}_i)_{i\in I} \subset V$.

a) Sei $(c_i)_{i\in I} \subset \mathbb{K}$ und $c_i = 0$ für fast alle $i \in I$. Man nennt $\sum_{i\in I} c_i\vec{v}_i$ eine **Linearkombination** *mit den* **Koeffizienten** *$(c_i)_{i\in I}$. Statt „$\sum_{i\in I} c_i\vec{v}_i$ mit $c_i = 0$ für fast alle $i \in I$" schreiben wir einfach $\sum_{i\in I}{}' c_i\vec{v}_i$ (der Strich am Summenzeichen bedeutet, dass nur endlich viele Summanden auftreten).*

b) Wenn wir die Indexmenge nicht explizit benennen möchten, schreiben wir einfach $\sum_{\vec{v}\in S}{}' c_{\vec{v}}\vec{v}$ wobei zu beachten ist, dass der gleiche Vektor in der Familie S mehrfach auftreten kann.

c) Eine Linearkombination $\vec{0} = \sum_{\vec{v}\in S}{}' c_{\vec{v}}\vec{v}$ heißt **lineare Abhängigkeit** *für S, falls $\exists \vec{v} \in S\colon\ c_{\vec{v}} \neq 0$.*

d) S heißt **linear abhängig**, *wenn es eine lineare Abhängigkeit für S gibt, und heißt andernfalls* **linear unabhängig**.

e) Die Menge aller Linearkombinationen, die sich aus den Elementen von S bilden lassen, heißt **Erzeugnis** *von S. Notation:*

$$\mathrm{Span}(S) := \{\sum_{i\in I}{}' c_i\vec{v}_i \mid (c_i)_{i\in I} \subset \mathbb{K},\ c_i = 0 \text{ für fast alle } i \in I\}$$

Ist $V = \mathrm{Span}(S)$, so heißt S **Erzeugendensystem** *von V.*

Bemerkung: Es geht bei uns stets aus dem Kontext hervor, bezüglich welches Körpers die Linearkombinationen zu bilden sind. Will man es explizit festlegen, kann man $\mathrm{Span}_{\mathbb{K}}(S)$ schreiben.

Problem 3.9
Prüfe, ob $[\vec{v}_1, ..., \vec{v}_k] \subset \mathbb{K}^n$ linear unabhängig ist.

Lösung:
Sei $A := (\vec{v}_1, ..., \vec{v}_k) \in \mathbb{K}^{n\times k}$. Eine lineare Abhängigkeit für $[\vec{v}_1, ..., \vec{v}_k]$ ist ein $\vec{0} \neq \vec{c} \in \mathbb{K}^k$ mit $A\vec{c} = \vec{0}$. Und dies existiert genau dann, wenn $\mathrm{LR}(A; \vec{0}) \neq \{\vec{0}\}$.

Also: Bringe auf A auf ZSF A'. $[\vec{v}_1, ..., \vec{v}_k]$ ist genau dann linear unabhängig, wenn alle Spalten von A' Pivotspalten sind, denn dann hat das lineare Gliechungssystem $A\vec{c} = \vec{0}$ keine freien Variablen. □

Beispiel 3.10 (und Notation)

a) Sei V ein $\mathbb{K}$-Vektorraum und $S := (\vec{v}_i)_{i\in I} \subset V$.

- *Wenn es $i \neq j \in I$ mit $\vec{v}_i = \vec{v}_j$ gibt, dann ist $\vec{v}_i + (-1)\vec{v}_j = \vec{0}$ eine lineare Abhängigkeit.*
- *Sei $I = \{i\}$ ein-elementig. Es ist S linear unabhängig genau dann, wenn $\vec{v}_i \neq \vec{0}$. Eine Linearkombination ist nämlich einfach $c_i\vec{v}_i$ mit einem $c_i \in \mathbb{K}$, und $c_i\vec{v}_i = \vec{0} \iff (c_i = 0) \vee (\vec{v}_i = \vec{0})$.*

b) Die Basislösungen eines homogenen linearen Gleichungssystems bilden ein Erzeugendensystem seines Lösungsraums. Wir erhalten dadurch eine neue Möglichkeit, die Lösungen linearer Gleichungssysteme zu notieren, mit $A \in \mathbb{K}^{m\times n}$ und $\vec{b} \in \mathbb{K}^m$:

- *Ist $J \subset \{1, ..., n\}$ die Indexmenge der freien Variablen, dann ist*

$$\mathrm{LR}(A; \vec{0}) = \mathrm{Span}(\vec{\beta}_j)_{j\in J}\,.$$

- *Ist V ein $\mathbb{K}$-Vektorraum, $\vec{v} \in V$ und $U \subset V$, so schreiben wir $\vec{v}+U := \{\vec{v}+\vec{u} \mid \vec{u} \in U\}$. Damit können wir auch $\mathrm{LR}(A; \vec{b})$ kompakter schreiben: Ist $\vec{x}_{\mathrm{spez}} \in \mathrm{LR}(A; \vec{b})$ und $J \subset \{1, ..., n\}$ die Indexmenge der freien Variablen, dann ist $\mathrm{LR}(A; \vec{b}) = \vec{x}_{\mathrm{spez}} + \mathrm{Span}(\vec{\beta}_j)_{j\in J}$*

c) $\mathrm{Span}(\binom{1}{0}, \binom{1}{1}) = \mathbb{R}^2 = \mathrm{Span}(\binom{1}{0}, \binom{1}{1}, \binom{0}{-1})$

d) Es ist erlaubt, dass S unendlich viele Vektoren enthält, auch wenn in jeder Linearkombination nur endlich *viele Elemente von S verwendet werden. Auch $\mathbb{R}^2$ selbst ist ein Erzeugendensystem von $\mathbb{R}^2$.*

Lemma 3.11 (und Definition)

Sei V ein $\mathbb{K}$-Vektorraum und $S \subset V$ eine Familie. Dann ist $\mathrm{Span}(S) \leq V$. *Man nennt* $\mathrm{Span}(S)$ *daher auch den von S* **erzeugten Untervektorraum**.

Beweis:

Wir verwenden Lemma 3.6. Eine leere Summe hat den Wert Null, d.h. den Nullvektor erhält man als leere Linearkombination. Daher $\vec{0} \in \mathrm{Span}(S)$.

Seien $\vec{u}, \vec{v} \in \mathrm{Span}(S)$, also $\vec{u} = \sum_{\vec{w}\in S}{}' c_{\vec{w}}\vec{w}$ und $\vec{v} = \sum_{\vec{w}\in S}{}' d_{\vec{w}}\vec{w}$.

Weil $c_{\vec{w}} = d_{\vec{w}} = 0$ für fast alle $\vec{w} \in S$ gilt, gilt für $\lambda, \mu \in \mathbb{K}$ auch $\lambda c_{\vec{w}} + \mu d_{\vec{w}} = 0$ für fast alle $\vec{w} \in S$. Also $\lambda\vec{u} + \mu\vec{v} = \sum_{\vec{w}\in S}{}' (\lambda c_{\vec{w}} + \mu d_{\vec{w}})\vec{w} \in \mathrm{Span}(S)$ und wegen Lemma 3.6 folgt $\mathrm{Span}(S) \leq V$. □

3.3 Lineare Abbildungen

Wir kommen nun zu den strukturerhaltenden Abbildungen von Vektorräumen.

Definition 3.12
Seien V, W $\mathbb{K}$-Vektorräume. Eine Abbildung $f\colon V \to W$ heißt **linear** *(oder auch $\mathbb{K}$-linear oder* **Homomorphismus** *von $\mathbb{K}$-Vektorräumen) $:\Leftrightarrow \forall \vec{u}, \vec{v} \in V$ und $\forall \lambda \in \mathbb{K}$ gilt $f(\vec{u}+\vec{v}) = f(\vec{u}) + f(\vec{v})$ und $f(\lambda\vec{v}) = \lambda f(\vec{v})$.*

Beispiel 3.13 (und Notation)

a) *Für $A \in \mathbb{K}^{m\times n}$ sei $L_A\colon \mathbb{K}^n \to \mathbb{K}^m$ definiert durch $L_A(\vec{v}) := A \cdot \vec{v}$.*
L_A ist linear, denn $A \cdot (\vec{u}+\vec{v}) = A\vec{u} + A\vec{v}$ und $A \cdot (\lambda\vec{v}) = \lambda A\vec{v}$. Lineare Abbildungen verallgemeinern also Matrixmultiplikation.

b) *Parallelprojektionen von $V = \mathbb{R}^n$ auf $W \leq V$ sind lineare Abbildungen.*

c) *$\mathbb{R} \to \mathbb{R}$ mit $x \mapsto x^2$ ist* nicht *linear:* $(1+1) \mapsto 2^2 = 4 \neq 1^2 + 1^2 = 2$

d) *In der Schule haben Sie eine Funktion der Form wie $f(x) = 5\,x + 3$ vermutlich als „lineare Funktion“ bezeichnet. Das ist jedoch* keine *lineare Abbildung im Sinne der Definition, denn $f(1+1) = 13 \neq f(1) + f(1) = 8 + 8 = 16$.*

e) *Für differenzierbare Funktionen $f \in \mathscr{C}^1(\mathbb{R}, \mathbb{R})$ ist durch $D(f) := f'$ eine Abbildung $D\colon \mathscr{C}^1(\mathbb{R}, \mathbb{R}) \to \mathscr{C}^0(\mathbb{R}, \mathbb{R})$ definiert. Es gilt bekanntlich $(c_1 f_1 + c_2 f_2)' = c_1 f_1' + c_2 f_2'$ für $f_1, f_2 \in \mathscr{C}^1(\mathbb{R}, \mathbb{R})$ und $c_1, c_2 \in \mathbb{R}$. Also ist D linear.*

f) *Sei $\mathbb{K} := \mathbb{Z}/2\mathbb{Z}$ und sei $V := \mathbb{K}[X]$ der Vektorraum der Polynome über dem Körper mit zwei Elementen. Siehe Übungen: Die Abbildung $f\colon V \to V$ definiert durch $\forall p \in V\colon f(p) := p^2$ ist eine lineare Abbildung. Das ist erstaunlich, denn wenn man Polynome mit* reellen *Koeffizienten betrachtet, so ist die Abbildung $p \mapsto p^2$* nicht *linear!*

Lemma 3.14
Seien V, W $\mathbb{K}$-Vektorräume und $f\colon V \to W$ eine Abbildung.

a) *f ist linear $\iff \forall \vec{v}, \vec{w} \in V$ und $\forall \lambda, \mu \in \mathbb{K}$ gilt $f(\lambda\vec{v} + \mu\vec{w}) = \lambda f(\vec{v}) + \mu f(\vec{w})$.*

b) *Sei f linear, $\vec{v}_1, ..., \vec{v}_k \in V$, $\lambda_1, ..., \lambda_k \in \mathbb{K}$. Dann $f(\sum_{i=1}^{k} \lambda_i \vec{v}_i) = \sum_{i=1}^{k} \lambda_i f(\vec{v}_i)$.*

c) *Ist f linear, dann $f(\vec{0}) = \vec{0}$.*

Beweis:
a) „$\Rightarrow$“: $f(\lambda\vec{v} + \mu\vec{w}) = f(\lambda\vec{v}) + f(\mu\vec{w}) = \lambda f(\vec{v}) + \mu f(\vec{w})$.

„$\Leftarrow$“: $f(\vec{v}+\vec{w}) = f(1\vec{v}+1\vec{w}) = 1f(\vec{v}) + 1f(\vec{w}) = f(\vec{v}) + f(\vec{w})$, und $f(\lambda\vec{v}) = f(\lambda\vec{v} + 0\vec{v}) = \lambda f(\vec{v}) + 0f(\vec{v}) = \lambda f(\vec{v})$.

b) Klar (Induktion nach der Anzahl der Summanden).

c) $f(\vec{0}) = f(0 \cdot \vec{0}) = 0 f(\vec{0}) = \vec{0}$. □

3.4 Basen

Definition 3.15
Ein linear unabhängiges Erzeugendensystem eines $\mathbb{K}$-Vektorraums heißt **Basis**.

Beispiel 3.16

a) Die Basislösungen des Lösungsraumes eines homogenen linearen Gleichungssystems bilden eine Basis dieses Lösungsraums.

b) $[\binom{1}{0}, \binom{1}{1}]$ und $[\binom{1}{0}, \binom{0}{-1}]$ sind jeweils Basen von $\mathbb{R}^2$.

c) Für $i \in \{1, ..., n\}$ sei $\vec{e}_i \in \mathbb{K}^n$ der Spaltenvektor, der in der i-ten Zeile den Eintrag 1 und überall sonst die Einträge 0 hat. Dann ist $[\vec{e}_1, ..., \vec{e}_n]$ eine Basis von $\mathbb{K}^n$. Man bezeichnet sie als die **Standardbasis** *von $\mathbb{K}^n$.*

d) Der Vektorraum $\mathbb{R}[X]$ hat die Basis $(X^n)_{n \in \mathbb{N}}$: Polynome haben eine eindeutige Darstellung als Linearkombination $\sum_{n=0}^{m} a_n X^n$.

e) Für eine Menge I ist $\mathbb{K}^I_{fin} := \{(c_i)_{i \in I} \subset \mathbb{K} \mid c_i = 0 \text{ für fast alle } i \in I\}$ ein Untervektorräume von $\mathbb{K}^I$ (mit punktweiser Addition und Skalarmultiplikation). Wir nennen ihn den $\mathbb{K}$-Vektorraum der **Abbildungen mit endlichem Träger**. *Im Fall $I = \{1, ..., n\}$ ist $\mathbb{K}^I = \mathbb{K}^I_{fin} = \mathbb{K}^n$.*

$(i^*)_{i \in I}$ mit $i^* := \left(\begin{cases} 1 & (i = j) \\ 0 & (i \neq j) \end{cases} \right)_{j \in I} \subset \mathbb{K}$ ist eine Basis von $\mathbb{K}^I_{fin}$, aber für unendliches I nicht von $\mathbb{K}^I$.[19]

Problem 3.17 (Basisauswahl für Untervektorräume von $\mathbb{K}^m$)
Sei $[\vec{v}_1, ..., \vec{v}_n] \in \mathbb{K}^m$ und $V := \text{Span}(\vec{v}_1, ..., \vec{v}_n)$. Sei $[\vec{v}_1, ..., \vec{v}_k]$ linear unabhängig (ggf. $k = 0$). Berechne eine Auswahl von $[\vec{v}_1, ..., \vec{v}_n]$, die $[\vec{v}_1, ..., \vec{v}_k]$ umfasst und eine Basis von V bildet.

Lösung:
Bringe $A := (\vec{v}_1, ..., \vec{v}_n) \in \mathbb{K}^{m \times n}$ mittels Gauß-Elimination auf Zeilenstufenform mit Pivotspalten $j_1, ..., j_r$. Dann ist $[\vec{v}_{j_1}, ..., \vec{v}_{j_r}]$ die gesuchte Basis.

Begründung: Weil die ersten k Vektoren linear unabhängig sind, sind die ersten k Spalten der Zeilenstufenform Pivotspalten.

$V = \{A \cdot \vec{c} \mid \vec{c} \in \mathbb{K}^n\} = \{\vec{b} \in \mathbb{K}^m \mid \text{LR}(A; \vec{b}) \neq \emptyset\}$. Sei nun $\vec{b} \in V$. Nach unserem Lösungsalgorithmus für lineare Gleichungssysteme hat die Gleichung $A\vec{c} = \vec{b}$ eine

[19] Tatsächlich ist es in diesem Fall unmöglich, eine Basis von $\mathbb{K}^I$ explizit anzugeben.

Lösung $\vec{c}_{\text{spez}}$, bei der alle freien Variablen Null sind. Eine solche Lösung entspricht aber einer Linearkombination der $[\vec{v}_{j_1}, ..., \vec{v}_{j_r}]$.

Weil es zu jedem $\vec{b} \in V$ eine solche Linearkombination gibt, ist $[\vec{v}_{j_1}, ..., \vec{v}_{j_r}]$ ein Erzeugendensystem von V. Weil es *genau eine* gibt, gibt es keine lineare Abhängigkeit von $[\vec{v}_{j_1}, ..., \vec{v}_{j_r}]$

Beachte: Wähle die ursprünglichen Spalten, nicht die der ZSF! □

Beispiel 3.18
Die Vektoren $\vec{v}_1 = \begin{pmatrix}1\\2\\1\end{pmatrix}$, $\vec{v}_2 = \begin{pmatrix}-1\\1\\0\end{pmatrix}$, $\vec{v}_3 = \begin{pmatrix}2\\1\\1\end{pmatrix}$ *und* $\vec{v}_4 = \begin{pmatrix}2\\2\\2\end{pmatrix}$ *bilden ein Erzeugendensystem von* $\mathbb{R}^3$*, die ersten beiden Vektoren sind linear unabhängig.*
$\begin{pmatrix}1&-1&2&2\\2&1&1&2\\1&0&1&2\end{pmatrix} \rightsquigarrow \begin{pmatrix}1&-1&2&2\\0&3&-3&-2\\0&0&0&2/3\end{pmatrix}$. *Also bildet* $[\vec{v}_1, \vec{v}_2, \vec{v}_4]$ *eine Basis von* $\mathbb{R}^3$.

3.4.1 Charakterisierungen von Basen

Innerhalb einer Vorlesung oder eines Buches hat jeder Begriff genau eine Definition. Andere Definitionsmöglichkeiten heißen ***Charakterisierungen*** des Begriffs.

Lemma 3.19. *Sei* V *ein* $\mathbb{K}$*-Vektorraum,* $(\vec{v}_i)_{i\in I} \subset V$*,* $0 \in I$ *und* $I' := I \setminus \{0\}$*. Folgende drei Aussagen sind gleichbedeutend:*

a) Es gibt eine Linearkombination $\sum_{i\in I}{}' \lambda_i \vec{v}_i = \vec{0}$ *mit* $\lambda_0 \neq 0$.

b) $\vec{v}_0 \in \text{Span}(\vec{v}_i)_{i\in I'}$.

c) $\text{Span}(\vec{v}_i)_{i\in I} = \text{Span}(\vec{v}_i)_{i\in I'}$.

Zusatz: *Ist* $(\vec{v}_i)_{i\in I'}$ *linear unabhängig, so sind a), b), c) gleichbedeutend zu*

d) $(\vec{v}_i)_{i\in I}$ *ist linear abhängig.*

Beweis:
c) ⇒ b): Klar.

b) ⇒ a): Ist $\text{Span}(\vec{v}_i)_{i\in I'} \ni \vec{v}_0 = \sum_{i\in I'}{}' \lambda_i \vec{v}_i$, dann $\vec{0} = \sum_{i\in I}{}' \lambda_i \vec{v}_i$ mit $\lambda_0 = -1$.

a) ⇒ c): Aus jeder Linearkombination von Elementen von S kann man $\vec{v}_0$ löschen, indem man $\vec{v}_0 = \sum_{i\in I'}{}' \frac{-\lambda_i}{\lambda_0} \vec{v}_i$ einsetzt. Also $\subseteq$, und $\supseteq$ ist klar.

Zusatz: a) ⇒ d) gilt immer. Wir zeigen d) ⇒ a): Sei $(\vec{v}_i)_{i\in I'}$ linear unabhängig und $\vec{0} = \sum_{i\in I}{}' \lambda_i \vec{v}_i$ eine lineare Abhängigkeit. Angenommen, $\lambda_0 = 0$.
Dann $\vec{0} = \sum_{i\in I}{}' \lambda_i \vec{v}_i = \sum_{i\in I'}{}' \lambda_i \vec{v}_i$. Weil $(\vec{v}_i)_{i\in I'}$ linear unabhängig ist, folgt $\forall i \in I' \colon \lambda_i = 0$, also auch $\forall i \in I \colon \lambda_i = 0$ — Widerspruch. Also $\lambda_0 \neq 0$. □

Satz 3.20 (Charakterisierung von Basen). *Sei V ein $\mathbb{K}$-Vektorraum. Für eine Familie $\mathcal{S} \subseteq V$ ist gleichbedeutend:*

a) $\mathcal{S}$ ist eine Basis von V.

b) $\mathcal{S}$ ist ein minimales Erzeugendensystem von V: Entfernt man ein Element von $\mathcal{S}$, liegt danach kein Erzeugendensystem mehr vor.

c) $\mathcal{S}$ ist eine maximal linear unabhängige Menge in V: Fügt man ein Element zu $\mathcal{S}$ hinzu, liegt danach keine linear unabhängige Menge vor.

Beweis:

a) $\iff$ b) Sei $\mathcal{S}$ ein Erzeugendensystem. Zu zeigen ist: $\mathcal{S}$ ist linear unabhängig gdw. $\mathcal{S}$ ist ein *minimales* Erzeugendensystem.

$\mathcal{S}$ ist nicht minimal gdw. $\exists \vec{v}_0 \in \mathcal{S}$, so dass $V = \mathrm{Span}(\mathcal{S}) = \mathrm{Span}(\mathcal{S} \setminus \{\vec{v}_0\})$, gdw. $\mathcal{S}$ ist linear abhängig nach Lemma 3.19.

a) $\iff$ c) Sei $\mathcal{S}$ linear unabhängig. $\mathcal{S}$ ist *kein* Erzeugendensystem $\iff$ $\exists \vec{v}_0 \in V \setminus \mathrm{Span}(\mathcal{S})$ $\iff$ $\exists \vec{v}_0 \in V \setminus \mathcal{S}$: $\mathrm{Span}(\mathcal{S} \cup \{\vec{v}_0\}) \supsetneq \mathrm{Span}(\mathcal{S})$ $\iff$ $\exists \vec{v}_0 \in V \setminus \mathcal{S}$: $\mathcal{S} \cup \{\vec{v}_0\}$ ist linear unabhängig (Zusatz von Lemma 3.19). □

3.4.2 Invertierbare Matrizen

Soll man ein lineares Gleichungssystem $A\vec{x} = \vec{b}$ lösen, ist man vielleicht versucht, $\vec{x} = A^{-1}\vec{b}$ zu antworten. Problem ist die Invertierbarkeit: Nicht jede von Null verschiedene Matrix kann man invertieren.

Definition 3.21
Sei $A \in \mathbb{K}^{m \times n}$, welches mit Gauß-Elimination in eine Zeilenstufenform mit $0 \leq r \leq m$ von Null verschiedenen Zeilen (also auch mit $0 \leq r \leq n$ Pivotspalten) übergeht. Der **Rang** *von A ist* $\mathrm{Rang}(A) := r$.
$GL_n(\mathbb{K}) := \{A \in M_n(\mathbb{K}) \mid \mathrm{Rang}(A) = n\}$ *heißt* **allgemeine lineare Gruppe**.

Wir werden im nächsten Kapitel zeigen, dass diese Definition nicht von den Details des Gauß-Algorithmus abhängt.

Definition 3.22
Eine quadratische Matrix $A \in M_n(\mathbb{K})$ heißt **invertierbar** *oder* **regulär** *gdw.* $\mathrm{Rang}(A) = n$. *Andernfalls heißt A* **singulär**.

Problem 3.23 (Inverse Matrix)
Es sei $A \in M_n(\mathbb{K})$. Erkenne, ob A regulär ist und berechne ggf. eine reguläre Matrix $A^{-1} \in M_n(\mathbb{K})$ so dass $A^{-1} \cdot A = A \cdot A^{-1} = \mathbb{1}_n$.

Lösung:
Wir starten mit der erweiterten Matrix $(A, \mathbb{1}_n)$, wenden darauf den Gauß-Jordan-Algorithmus an und erhalten eine reduzierte Zeilenstufenform (A', B). Wenn $A' \neq \mathbb{1}_n$, dann hat A' mindestens eine nicht-Pivotspalte, also $\text{Rang}(A) < n$, d.h. A ist singular. Ist $A' = \mathbb{1}_n$, dann Rückgabe von $A^{-1} = B$.

Weil B durch die inversen Zeilenoperationen in die ZSF $\mathbb{1}_n$ über geht, ist $\text{Rang}(B) = n$. Nach Lemma 2.21.b) gibt es $X \in M_n(\mathbb{K})$ mit $\mathbb{1}_n = A' = X \cdot A$ und $B = X \cdot \mathbb{1}_n$. Also $X = B$ und $B \cdot A = \mathbb{1}_n$.

Sei $\vec{b}_i \in \mathbb{K}^n$ die i-te Spalte von B. Weil $\vec{e}_i \in \mathbb{K}^n$ die i-te Spalte von $\mathbb{1}_n$ ist, gilt $\text{LR}(A, \vec{e}_i) = \text{LR}(\mathbb{1}_n, \vec{b}_i) = \{\vec{b}_i\}$. Also $A \cdot \vec{b}_i = \vec{e}_i$, und das setzt sich spaltenweise zu $A \cdot B = \mathbb{1}_n$ zusammen. □

Beispiel Für $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$ ist $(A, \mathbb{1}_3) = \begin{pmatrix} 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \end{pmatrix}$ mit Zeilenstufenform $\begin{pmatrix} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -1 & 1 & 1 \end{pmatrix}$. Man erkennt, dass A singulär ist.

Für $A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$ ist $(A, \mathbb{1}_3) = \begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \end{pmatrix}$. Gauß-Jordan ergibt $\begin{pmatrix} 1 & 0 & 0 & 1 & 0 & -1 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & -1 & 1 & 1 \end{pmatrix}$. Also invertierbar, mit $A^{-1} = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 0 & 1 \\ -1 & 1 & 1 \end{pmatrix}$.

3.4.3 Rechnerische Zugänge

Mit Basen kann man Vektoren und lineare Abbildungen rechnerisch darstellen.

Lemma 3.24 (und Definition)
Sei V ein $\mathbb{K}$-Vektorraum. Eine Familie $B = (\vec{b}_i)_{i \in I} \subset V$ ist genau dann eine Basis von V, wenn jedes $\vec{v} \in V$ genau eine *Darstellung als Linearkombination $\vec{v} = \sum_{i \in I}{}' c_i \vec{b}_i$ hat. Ihre Koeffizienten nennt man die* **Koordinaten** *von $\vec{v}$ bzgl. B, die Familie ${}^B\vec{v} := (c_i)_{i \in I} \in \mathbb{K}^I_{fin}$ heißt* **Koordinatenvektor** *von $\vec{v}$ bezüglich B. Die Abbildung $\Phi_B : V \to \mathbb{K}^I_{fin}$ mit $\Phi_B(\vec{v}) := {}^B\vec{v}$ ist linear.*

Beweis:
B ist ein *Erzeugendensystem* genau dann, wenn jedes $\vec{v} \in V$ *mindestens* eine Darstellung als Linearkombination aus B hat.

B linear abhängig $\Rightarrow \vec{0}$ hat zwei Darstellungen als Linearkombination aus B. Umgekehrt: Sei $V \ni \vec{v} = \sum_{i \in I}{}' \lambda_i \vec{b}_i = \sum_{i \in I}{}' \mu_i \vec{b}_i$ und dabei $\lambda_i \neq \mu_i$ für mindestens ein $i \in I$. Dann ist $\sum_{i \in I}{}' (\lambda_i - \mu_i)\vec{b}_i = \vec{0}$ eine lineare Abhängigkeit, also B keine Basis.

Linearität von Φ_B: Sei $\lambda, \mu \in \mathbb{K}$ und $\vec{u}, \vec{v} \in V$ mit $\vec{u} = \sum_{i \in I}{}' c_i \vec{b}_i$ und $\vec{v} = \sum_{i \in I}{}' d_i \vec{b}_i$. Dann $\lambda\vec{u} + \mu\vec{v} = \sum_{i \in I}{}' (\lambda c_i + \mu d_i)\vec{b}_i$. □

Problem 3.25
Sei $B = [\vec{b}_1, ..., \vec{b}_n]$ eine Basis von $V \leq \mathbb{K}^m$ und sei $\vec{v} \in \mathbb{K}^m$. Prüfe, ob $\vec{v} \in V$ und berechne ggf. ${}^B\vec{v} \in \mathbb{K}^n$.

Lösung: $\mathrm{LR}\left((\vec{b}_1, ..., \vec{b}_n);\ \vec{v}\right) = \begin{cases} \{\,{}^B\vec{v}\} & \vec{v} \in V \\ \emptyset & \text{sonst} \end{cases}$. □

Satz von der Linearen Fortsetzung
Seien V, W $\mathbb{K}$-Vektorräume und sei $B = (\vec{b}_i)_{i\in I} \subset V$ eine Basis von V. Zu jedem $(\vec{w}_i)_{i\in I} \subset W$ gibt es genau eine lineare Abbildung $f\colon V \to W$ mit $\forall j \in I\colon f(\vec{b}_j) = \vec{w}_j$.

Beweis:
Eindeutigkeit: Seien $f\colon V \to W$ linear. Für alle $\vec{v} \in V$ gilt wegen Linearität von f: $f(\vec{v}) = f(\sum_{j\in I}{}' {}^B v_j\, \vec{b}_j) = \sum_{j\in I}{}' {}^B v_j\, f(\vec{b}_j)$, also ist f eindeutig dadurch bestimmt, wie es die Basisvektoren abbildet.

Existenz: Zeige, dass der einzig mögliche Kandidat $f(\vec{v}) := \sum_{j\in I}{}' {}^B v_j\, \vec{w}_j$ wirklich linear ist. Für $\lambda, \mu \in \mathbb{K}$ und $\vec{u}, \vec{v} \in V$ gilt $f(\lambda\vec{u}+\mu\vec{v}) = \sum_{j\in I}{}'(\lambda\, {}^B u_j + \mu\, {}^B v_j)\vec{w}_j =$

$\lambda\sum_{j\in I}{}' {}^B u_j\, \vec{w}_j + \mu\sum_{j\in I}{}' {}^B v_j\, \vec{w}_j$. □

3.5 Gruppen, Permutationen

Dieses Kapitel gehört logisch eigentlich sogar vor das Kapitel über Ringe und Körper. Es geht hier nämlich um Mengen mit nur *einer* (statt wie bei Ringen zwei) inneren Verknüpfung. Jede zusätzliche Struktur ist ein zusätzliches Werkzeug, daher sind Ringe oder gar Körper viel leichter zu untersuchen als Gruppen.

Allerdings handelt es sich nicht um einen eigentlichen Grundbegriff der *linearen Algebra*, sondern eher um einen Hilfsbegriff. **Im Hinblick auf die Verkürzung des Semesters werde ich diesen Abschnitt des Skripts daher in der Vorlesung nur stark verkürzt präsentieren und bitte Sie darum, die Details selbständig zu lesen.** Wieder geht es um eine Struktur und ihre strukturerhaltenden Abbildungen.

Definition 3.26
Eine **Gruppe** $(G, *)$ *(oder kurz G) ist eine Menge mit einer inneren Verknüpfung $*$, so dass die folgenden* **Gruppenaxiome** *gelten:*

(G1) $\forall a, b, c \in G\colon\ (a * b) * c = a * (b * c)$, *d.h. $*$ ist* **assoziativ**.

(G2) $\exists e \in G\colon\ \forall a \in G\colon\ a * e = a$*; man nennt* e **neutral**.

(G3) $\forall a \in G\colon\ \exists b \in G\colon\ a * b = e$*;* b *heißt zu* a **invers**.

Gilt $\forall a, b \in G\colon\ a * b = b * a$*, so heißt* G **abelsch**[20] *und* $*$ **kommutativ**.

Offenbar kennen wir schon viele abelsche Gruppen, nämlich die additive Gruppe jedes Rings. Wir werden noch zeigen, dass $GL_n(\mathbb{K})$ bezüglich Matrixmultiplikation eine Gruppe ist. Die folgenden Grundergebnisse kennen Sie im Prinzip bereits aus dem Kapitel über Ringe (inklusive Beweis):

Lemma 3.27 (und Notation)
Sei G *eine Gruppe mit Verknüpfung* $*$ *und neutralem Element* e*.*

a) $\forall a, b \in G$*: Wenn* $a * b = e$*, dann* $b * a = e$*.*

b) $\forall a \in G\colon\ e * a = a$*.*

c) Wenn $x \in G$*, so dass* $\exists a \in G\colon\ a * x = a$*, dann* $x = e$*.*

d) $\forall a, b, c \in G\colon\ a * b = e = a * c\ \Rightarrow\ b = c$*. Notation: Für* $a \in G$ *bezeichnet* $a^{-1} \in G$ *das durch* $a * a^{-1} = e$ *eindeutig bestimmte Inverse von* a*.*

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

f) $\forall a \in G\colon\ (a^{-1})^{-1} = a$*.*

Beweis:

a) Nach (G3) $\exists c \in G\colon\ b * c = e$. Dann $b * a = (b * a) * e = b * (a * e) = b * (a * (b * c)) = b * ((a * b) * c) = b * (e * c) = (b * e) * c = b * c = e$.

b) Nach (G3) und a) $\exists b \in G\colon\ a * b = b * a = e$. $e * a = (a * b) * a = a * (b * a) = a * e = a$.

c) Nach (G3) und a) $\exists b \in G\colon\ a * b = b * a = e$. $a * x = a\ \Rightarrow\ b * (a * x) = b * a\ \Rightarrow\ (b * a) * x = e\ \Rightarrow\ e * x = e\ \Rightarrow\ x = e$.

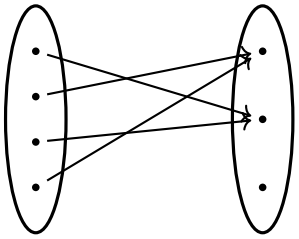
d) Nach (G3) und a) $\exists d \in G\colon\ a * d = d * a = e$. $a * b = a * c\ \Rightarrow\ d * (a * b) = d * (a * c)\ \Rightarrow\ e * b = e * c\ \Rightarrow\ b = c$.

e) Zu zeigen: $e \overset{!}{=} (a*b)*(b^{-1}*a^{-1}) = a*(b*(b^{-1}*a^{-1})) = a*((b*b^{-1})*a^{-1}) = a * (e * a^{-1}) = a * a^{-1} \overset{\checkmark}{=} e$.

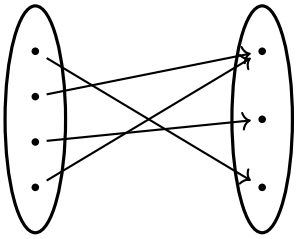
f) Wir wissen bereits $a^{-1} * a = e$, also ist wegen der Eindeutigkeit des inversen Elements auch $a = (a^{-1})^{-1}$. □

[20] Nach Niels Henrik Abel [1802–1829], einem der Begründer der Gruppentheorie.

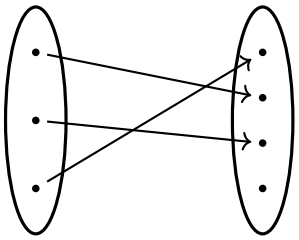
Ab jetzt lassen wir die Klammern in Gruppen meist weg, was wegen der Assoziativität erlaubt ist. Auch $*$ lassen wir meist weg (ab statt $a * b$), wenn die Verknüpfung aus dem Kontext klar ist.



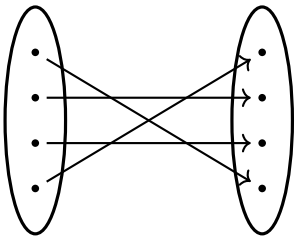
Weder surjektiv noch injektiv



Surjektiv aber nicht injektiv



Injektiv aber nicht surjektiv



Bijektiv

Definition 3.28
Sei $f\colon X \to Y$, $A \subseteq X$ und $B \subseteq Y$.

a) *$f(A)$ bezeichnet die* **Bildmenge** $f(A) := \{f(x) \mid x \in A\}$ *von A. Insbesondere nennt man* $\mathrm{Bild}(f) := f(X) = \{f(x) \mid x \in X\}$ *das* **Bild**[21] *von f.*

b) *Das* **Urbild**[22] *von B unter f ist $f^{-1}(B) := \{x \in X \mid f(x) \in B\} \subseteq X$.*

c) *f ist* **injektiv** $:\Leftrightarrow \forall y \in Y\colon\ |f^{-1}(\{y\})| \leq 1$.

d) *f ist* **surjektiv** $:\Leftrightarrow \forall y \in Y\colon\ |f^{-1}(\{y\})| \geq 1$.

e) *f ist* **bijektiv** $:\Leftrightarrow \forall y \in Y\colon\ |f^{-1}(\{y\})| = 1$.

Aufgabe 3.29 (und Definition). *Seien $f\colon X \to Y$ und $g\colon Y \to Z$ Abbildungen. Die* **Verknüpfung** *oder* **Komposition** *$g \circ f\colon X \to Z$ ist durch $\forall x \in X\colon\ (g \circ f)(x) := g(f(x))$ definiert.*

a) *Ist $g \circ f$ injektiv, dann ist f injektiv.*

b) *Ist $g \circ f$ surjektiv, dann ist g surjektiv.*

c) *Sind f und g injektiv (bzw. surjektiv), dann ist auch $g \circ f$ injektiv (bzw. surjektiv).*

[21] Auf Englisch "*image*"
[22] Auf Englisch "*preimage*". Andere Bezeichnung: „*Faser*“, englisch "*fibre*“

Die folgenden beiden Lemmas sind der Grund, warum man mit Abbildungen Gruppen bauen kann.

Lemma 3.30
Verknüpfung von Abbildungen ist assoziativ, d.h. sind $f\colon X \to Y$, $g\colon Y \to Z$ *und* $h\colon Z \to W$, *so gilt* $h \circ (g \circ f) = (h \circ g) \circ f$.

Beweis:
Für jedes $x \in X$ haben beide Abbildungen den Wert $h(g(f(x)))$. □

Definition 3.31 (und Axiom und Lemma)
Sei $f\colon X \to Y$.

a) Die **Identitätsabbildung** $\mathrm{Id}_X\colon X \to X$ *ist dadurch definiert, dass* $\forall x \in X\colon\ \mathrm{Id}_X(x) := x$.

b) **Auswahlaxiom***: Ist* f *surjektiv, so gibt es eine Abbildung* $g\colon Y \to X$ *mit* $f \circ g = \mathrm{Id}_Y$.

c) Ist f *bijektiv, so gibt*[23] *es eine durch* f *eindeutig bestimmte Abbildung* $f^{-1}\colon Y \to X$ *mit* $f \circ f^{-1} = \mathrm{Id}_Y$ *und* $f^{-1} \circ f = \mathrm{Id}_X$. *Man nennt sie die* **inverse Abbildung** *oder* **Umkehrabbildung** *von* f.

Trotz der ähnlichen Notation sind natürlich das Urbild einer Menge und die inverse Abbildung zwei verschiedene Dinge!

Beweis:
Zu b) sei angemerkt, dass wegen Surjektivität $\forall y \in Y\colon\ f^{-1}(\{y\}) \neq \emptyset$. Die Abbildung g hat die Eigenschaft $\forall y \in Y\colon\ g(y) \in f^{-1}(\{y\})$.

Wir beweisen c): Für alle $y \in Y$ besteht $f^{-1}(\{y\})$ wegen Bijektivität von f aus genau einem Element. Durch die Bedingung $f^{-1}(y) \in f^{-1}(\{y\})$ ist die Abbildung $f^{-1}\colon Y \to X$ also eindeutig und ohne Auswahlproblem bestimmt, woraus zunächst $f \circ f^{-1} = \mathrm{Id}_Y$ folgt. Ferner $\forall x \in X\colon\ f^{-1}(\{f(x)\}) = \{x\}$ und daher $f^{-1}(f(x)) = x$. Das bedeutet $f^{-1} \circ f = \mathrm{Id}_X$. □

3.5.1 Permutationen und die symmetrische Gruppe

Aufgabe 3.32 (und Definition)
Für eine Menge Ω *sei* $\mathrm{Sym}(\Omega) := \{f\colon \Omega \to \Omega \mid f \text{ ist bijektiv}\}$. $(\mathrm{Sym}(\Omega), \circ)$ *ist eine Gruppe, die* **symmetrische Gruppe** *von* Ω. *Das neutrale Element ist* Id_Ω, *das Inverse von* $\sigma \in \mathrm{Sym}(\Omega)$ *ist die Umkehrfunktion* σ^{-1}.

Die Elemente von $\mathrm{Sym}(\Omega)$ *nennt man auch* **Permutationen** *von* Ω. *Für* $n \in \mathbb{N}$ *sei* $S_n := \mathrm{Sym}(\{1, ..., n\})$.

[23] Auch ohne Verwendung des Auswahlaxioms!

Definition 3.33

a) *Sei $k \in \mathbb{N}_{\geq 2}$. Eine Permutation $\zeta \in \mathrm{Sym}(\Omega)$ heißt* **zyklisch** *der* **Länge** *$k \in \mathbb{N}^*$ oder kurz* **k-Zyklus** *gdw. $\exists a_0, ..., a_{k-1} \in \Omega$ paarweise verschieden, so dass $\forall i \in \{0, ..., k-2\}$: $\zeta(a_i) = a_{i+1}$, $\zeta(a_{k-1}) = a_0$ und $\forall c \in \Omega \setminus \{a_0, ..., a_{k-1}\}$: $\zeta(c) = c$. Schreibweise: $\zeta = (a_0\ a_1\ ...\ a_{k-1})$. Wählt man $a_0 = \min\{a_0, ..., a_{k-1}\}$, ist diese Schreibweise durch ζ eindeutig bestimmt. Wir bezeichnen $\mathrm{Tr}(\zeta) := \{a_0, ..., a_{k-1}\} \subseteq \Omega$ als* **Träger** *oder* **Bahn** *von ζ.*

b) *Zwei Zyklen $\zeta_1, \zeta_2 \in \mathrm{Sym}(\Omega)$ heißen* **disjunkt** *gdw. $\mathrm{Tr}(\zeta_1) \cap \mathrm{Tr}(\zeta_2) = \emptyset$.*

c) *2-Zyklen nennt man auch* **Transpositionen**.

Wir erlauben uns die Freiheit, für einen Zyklus nur den Träger zu notieren; auf welcher Gesamtmenge die Permutation operiert, geht aus dem Kontext hervor.

Beispiel 3.34

a) *Der 3-Zyklus $(1\ 2\ 4) = (2\ 4\ 1) = (4\ 1\ 2) \in S_4$ ist die Permutation $1 \mapsto 2$, $2 \mapsto 4$, $3 \mapsto 3$ und $4 \mapsto 1$.*

b) *$(3\ 4)$ und $(1\ 7)$ sind disjunkte Transpositionen.*

c) *$(1\ 2)(1\ 3) = (1\ 3\ 2) \neq (1\ 3)(1\ 2) = (1\ 2\ 3)$. Folglich ist S_n für $n \geq 3$ nicht-abelsch.*

Lemma 3.35 (und Notation)
Sind $\zeta_1, \zeta_2 \in \mathrm{Sym}(\Omega)$ disjunkte Zyklen, so gilt $\zeta_1\zeta_2 = \zeta_2\zeta_1$. Für alle $\sigma \in \mathrm{Sym}(\Omega)$ gibt es paarweise disjunkte Zyklen $\zeta_1, ..., \zeta_m \in \mathrm{Sym}(\Omega)$, so dass $\sigma = \zeta_1\zeta_2 \cdots \zeta_m$. Diese Darstellung ist bis auf Vertauschung der Produktreihenfolge eindeutig durch σ bestimmt. Man nennt dies die **Zyklendarstellung** *von σ. Beachte: Id_Ω hat die leere Zyklendarstellung.*

Beweisskizze:
Vertauschbarkeit disjunkter Zyklen: Ist $x \in \Omega \setminus \mathrm{Tr}(\zeta_1)$, so ist $\zeta_2(x) \in \Omega \setminus \mathrm{Tr}(\zeta_1)$. Daher $\zeta_1(\zeta_2(x)) = \zeta_2(x)$. Andererseits ist $\zeta_2(\zeta_1(x)) = \zeta_2(x)$. Und ist $x \in \mathrm{Tr}(\zeta_1)$, so ist insbesondere $x \in \Omega \setminus \mathrm{Tr}(\zeta_2)$, weshalb analog folgt $\zeta_1(\zeta_2(x)) = \zeta_2(\zeta_1(x))$.

Existenz der Zyklendarstellung: Wir verwenden vollst. Induktion nach der Mächtigkeit von Ω (Lücke: Wir müssten definieren, was Mächtigkeit wirklich ist, und zeigen, dass man die Mächtigkeiten endlicher Mengen mit den natürlichen Zahlen gleichsetzen kann; dazu braucht man das Auswahlaxiom).

- $\mathrm{Sym}(\emptyset) = \{\mathrm{Id}_\emptyset\}$, und $\mathrm{Id}_\emptyset$ hat die leere Zyklendarstellung.
- Sei $\Omega \neq \emptyset$. Wähle $a_0 \in \Omega$. Weil Ω endlich ist, ist die Abbildung $\mathbb{N} \to \Omega$ mit $m \mapsto a_m := \sigma^m(a_0)$ nicht injektiv. Sei $k \in \mathbb{N}^*$ minimal, so dass es ein $0 \leq \ell < k$ gibt mit $\sigma^\ell(a_0) = \sigma^k(a_0)$. Wäre $\ell > 0$, so würde gelten $a_0 = \sigma^0(a_0) = \sigma^{k-\ell}(a_0)$, was der Minimalität von k widerspräche. Also

sind die $a_0, ..., a_{k-1}$ paarweise disjunkt. Sei $A := \{a_0, ..., a_{k-1}\} \neq \emptyset$. Nach Induktionsvorraussetzung hat $\sigma|_{\Omega\setminus A} \in \mathrm{Sym}(\Omega \setminus A)$ eine Zyklendarstellung, $\sigma|_{\Omega\setminus A} = \zeta_1 ... \zeta_m$. Ist $k = 1$, dann ist $\sigma = \zeta_1 ... \zeta_m$. Ist $k > 1$, ist $\zeta_{m+1} := \sigma|_A = (a_0 \ \cdots \ a_{k-1})$ ein k-Zyklus, und weil dieser disjunkt zu $\zeta_1, ..., \zeta_m$ ist, gilt $\sigma = \zeta_1 \zeta_2 \cdots \zeta_{m+1}$.

Eindutigkeit der Zyklendarstellung: Ist $a \in \Omega$, so ist mit der obigen Konstruktion der Zyklus, in dessen Träger a liegt, eindeutig durch a bestimmt. □

Beispiel Die Permutation $\sigma \in S_{10}$ gegeben durch

$1 \mapsto 9$	$2 \mapsto 2$	$3 \mapsto 7$	$4 \mapsto 3$	$5 \mapsto 1$
$6 \mapsto 8$	$7 \mapsto 10$	$8 \mapsto 6$	$9 \mapsto 5$	$10 \mapsto 4$

hat die Zyklendarstellung (1 9 5)(3 7 10 4)(6 8).

3.5.2 Gruppenhomomorphismen

Definition 3.36
Seien $(G, *_G)$ bzw. $(H, *_H)$ Gruppen. Eine Abbildung $\varphi\colon G \to H$ heißt **Gruppenhomomorphismus** *gdw. $\forall g_1, g_2 \in G\colon \varphi(g_1 *_G g_2) = \varphi(g_1) *_H \varphi(g_2)$.*
Ist φ zudem bijektiv, heißt es **Gruppenisomorphismus**.

Man sagt auch kurz ***Homomorphismus*** bzw. ***Isomorphismus*** wenn aus dem Kontext klar ist, dass es um Gruppen geht.

Beispiel 3.37
$\exp\colon \mathbb{R} \to \mathbb{R}_{>0}$ *definiert durch* $\exp(x) := \mathrm{e}^x$ *ist ein Isomorphismus der Gruppen $(\mathbb{R}, +)$ und $(\mathbb{R}_{>0}, \cdot)$: Die Abbildung ist bijektiv (die Umkehrabbildung ist der natürliche Logarithmus), und $\forall s, t \in \mathbb{R}\colon \ \mathrm{e}^{s+t} = \mathrm{e}^s \cdot \mathrm{e}^t$.*

Lemma 3.38

a) Seien $(G, *_G)$ bzw. $(H, *_H)$ Gruppen mit neutralen Elementen $e_G \in G$ bzw. $e_H \in H$ und $\varphi\colon G \to H$ ein Homomorphismus. Dann gilt $\varphi(e_G) = e_H$ und $\forall g \in G\colon \varphi(g^{-1}) = (\varphi(g))^{-1}$.

b) Sind $\varphi\colon G_1 \to G_2$ und $\psi\colon G_2 \to G_3$ Gruppenhomomorphismen, dann auch $\psi \circ \varphi\colon G_1 \to G_3$.

c) Ist $\varphi\colon G \to H$ ein Gruppenisomomorphismus, so ist auch φ^{-1} ein Gruppenisomorphismus.

d) Id_G ein Isomorphismus.

Beweis:

a) $\varphi(e_G) = \varphi(e_G *_G e_G) = \varphi(e_G) *_H \varphi(e_G)$, also $\varphi(e_G) = e_H$. $\forall g \in G\colon e_H = \varphi(e_G) = \varphi(g *_G g^{-1}) = \varphi(g) *_H \varphi(g^{-1})$, also $\varphi(g^{-1}) = (\varphi(g))^{-1}$.

b) $\forall g_1, g_2 \in G_1\colon (\psi \circ \varphi)(g_1 * g_2) = \psi(\varphi(g_1 * g_2)) = \psi((\varphi(g_1) * \varphi(g_2)) = \psi(\varphi(g_1)) * \psi(\varphi(g_2))$.

c) Seien $h_1, h_2 \in H$. Weil φ bijektiv ist, gibt es $g_1 := \varphi^{-1}(h_1)$, $g_2 := \varphi^{-1}(h_2)$. Dann $\varphi(g_1 *_G g_2) = \varphi(g_1) *_H \varphi(g_2) = h_1 *_H h_2$. Wenden wir φ^{-1} auf beide Seiten der Gleichung an, folgt $g_1 *_G g_2 = \varphi^{-1}(h_1 *_H h_s)$. □

Wir konstruieren nun einen Homomorphismus $\operatorname{sgn}\colon S_n \to (\{\pm 1\}, \cdot)$. Die Idee: Ist $\sigma \in S_n$ ein Produkt von N Transpositionen, so ist $\operatorname{sgn}(\sigma) = (-1)^N$. Doch das ist in doppelter Hinsicht problematisch: Kann man jede Permutation als Produkt von Transpositionen schreiben? Und ist dann $\operatorname{sgn}(\sigma)$ wohldefiniert? Aufgrund dieser Probleme definiert man sgn stattdessen wie folgt:

Definition 3.39
Das **Vorzeichen** $\operatorname{sgn}(\sigma)$ *von* $\sigma \in S_n$ *ist* $\operatorname{sgn}(\sigma) := \prod_{1 \le i < j \le n} \frac{\sigma(j) - \sigma(i)}{j - i}$.

Satz 3.40
Sei $n \in \mathbb{N}$ *und* $\sigma \in S_n$.

a) *Es ist* $\operatorname{sgn}(\sigma) = (-1)^N$, *wobei* N *die Anzahl der Paare* i, j *mit* $i < j$ *und* $\sigma(i) > \sigma(j)$ *ist.*

b) $\operatorname{sgn}\colon S_n \to (\{\pm 1\}, \cdot)$ *ist ein Homomorphismus*

c) *Ist* σ *eine Transposition, dann* $\operatorname{sgn}(\sigma) = -1$.

d) *Ist* σ *ein* k*-Zyklus, dann* $\operatorname{sgn}(\sigma) = (-1)^{k-1}$.

e) σ *lässt sich als Produkt von (ggf. nicht-disjunkten) Transpositionen schreiben. Ist* σ *das Produkt von* N *Transpositionen, dann* $\operatorname{sgn}(\sigma) = (-1)^N$.

Beispiel $\operatorname{sgn}((1\ 4\ 7\ 2\ 9\ 11)(3\ 10\ 5)(6\ 8)) = (-1)^5 \cdot (-1)^2 \cdot (-1)^1 = +1$.

Beweis:

a) Jedes Paar $i < j$ gibt Anlass zum Faktor $j - i$ im Nenner. Ist $\sigma^{-1}(i) < \sigma^{-1}(j)$, so ergibt es auch den Faktor $j - i$ im Zähler; andernfalls ergibt es den Faktor $-(j - i)$ im Zähler.

b) Seien $\sigma, \tau \in S_n$. Die Faktoren, durch die das Vorzeichen definiert sind,

hängen wegen $\frac{\sigma(j)-\sigma(i)}{j-i} = \frac{\sigma(i)-\sigma(j)}{i-j}$ nur vom ungeordneten Paar i, j ab. Daher

$$\begin{aligned}\operatorname{sgn}(\sigma\tau) &= \prod_{\substack{\text{ungeordnete}\\ \text{Paare } i\neq j}} \frac{\sigma\tau(j)-\sigma\tau(i)}{j-i} \\ &= \prod_{\substack{\text{ungeordnete}\\ \text{Paare } i\neq j}} \frac{\sigma\tau(j)-\sigma\tau(i)}{\tau(j)-\tau(i)} \cdot \prod_{\substack{\text{ungeordnete}\\ \text{Paare } i\neq j}} \frac{\tau(j)-\tau(i)}{j-i} \\ &= \prod_{\substack{\text{ungeordnete}\\ \text{Paare } i\neq j}} \frac{\sigma(j)-\sigma(i)}{j-i} \cdot \prod_{\substack{\text{ungeordnete}\\ \text{Paare } i\neq j}} \frac{\tau(j)-\tau(i)}{j-i} = \operatorname{sgn}(\sigma)\cdot\operatorname{sgn}(\tau)\,.\end{aligned}$$

c) Sei $\sigma = (i\ j)$ mit $i < j$. Die Anzahl der in a) gezählten Paare ist ungerade, nämlich das Paar i, j sowie für jedes $i < \ell < j$ die Paare i, ℓ und ℓ, j.

d) $(a_1\ a_2\ \cdots\ a_k) = (a_1\ a_k)(a_1\ a_2\ \cdots\ a_{k-1})$. Die Behauptung folgt daraus durch die beiden vorigen Punkte des Satzes und Induktion nach k.

e) Jede Permutation ist ein Produkt von Zyklen und jeder Zyklus ist ein Produkt von Transpositionen, laut beweis des vorigen Punktes. Der Rest folgt, weil Transpositionen negatives Vorzeichen haben und das Vorzeichen ein Homomorphismus ist. □

4 Dimension

Wir kommen nun zu Vektorräumen, bei denen ein rechnerischer Zugang besteht. Wie gehabt sei $\mathbb{K}$ ein Körper.

Definition 4.1
Ein $\mathbb{K}$-Vektorraum heißt **endlich dimensional**, *wenn er ein endliches Erzeugendensystem hat.*

Praktisches Rechnen ist möglich, wenn sich zu jeder endlichen linear abhängigen Familie eine lineare Abhängigkeit effektiv berechnen lässt.

Beispiel 4.2
Für $n \in \mathbb{N}$ sind $\mathbb{K}^n$, $\mathbb{K}[X]_{\leq n}$, $M_n(\mathbb{K})$ endlich dimensional. Die linearen Abhängigkeiten von $[\vec{v}_1, ..., \vec{v}_k] \subset \mathbb{K}^n$ entsprechen $\mathrm{LR}\left((\vec{v}_1, ..., \vec{v}_k); \vec{0}\right) \setminus \{\vec{0}\}$, *lassen sich also berechnen.*

Beobachtung 4.3
Sei $B = [\vec{b}_1, ..., \vec{b}_n]$ eine endliche Basis eines $\mathbb{K}$-Vektorraums V. Wenn sich in V lineare Abhängigkeiten berechnen lassen, kann man für jedes $\vec{v} \in V$ den Koordinatenvektor ${}^B\vec{v} \in \mathbb{K}^n$ effektiv berechnen.

Beweis:
Weil B maximal linear unabhängig ist, gibt es eine effektiv berechenbare lineare Abhängigkeit $c\vec{v} + c_1\vec{b}_1 + ... + c_n\vec{b}_n = \vec{0}$. Weil B linear unabhängig ist, muss $c \neq 0$ gelten. Also ist ${}^Bv_i = -\frac{c_i}{c}$. □

4.1 Abbildungsmatrizen

Lemma 4.4 (und Definition)
Sei $f\colon V \to W$ $\mathbb{K}$-linear, $B = [\vec{b}_1, ..., \vec{b}_n]$ und $C = [\vec{c}_1, ..., \vec{c}_m]$ endliche Basen von V bzw. von W. Die **Abbildungsmatrix** *von f bzgl. B und C ist*

$$ {}^C_B f = ({}^C f(\vec{b}_1), ..., {}^C f(\vec{b}_n)) \in \mathbb{K}^{m \times n} $$

und $f(\vec{v})$ lässt sich durch Matrixmultiplikation beschreiben:

$$ {}^C f(\vec{v}) = {}^C_B f \cdot {}^B\vec{v} $$

Bemerkung Ich halte die hier verwendete Notation für geschickter als die in Wikipedia verwendete Notation $M_C^B(f)$.

Beweis:
Es sei $\Phi_C\colon W \to \mathbb{K}^m$ die lineare Abbildung gegeben durch $\Phi_C(\vec{w}) := {}^C\vec{w}$. Sei $\vec{v} \in V$. Definition des Koordinatenvektors: $\vec{v} = \sum_{i=1}^{n} {}^Bv_i\,\vec{b}_i$. Weil f und Φ_C linear sind, gilt ${}^C f(\vec{v}) = \Phi_C(f(\sum_{i=1}^{n} {}^Bv_i\,\vec{b}_i)) = \sum_{i=1}^{n} {}^Bv_i\,\Phi_C(f(\vec{b}_i)) = \left({}^C f(\vec{b}_1), ..., {}^C f(\vec{b}_n)\right) \cdot {}^B\vec{v}$.□

Beispiel 4.5
$f\colon \mathbb{R}^2 \to \mathbb{R}^2$ sei gegeben durch $\forall \binom{x}{y} \in \mathbb{R}^2\colon f\binom{x}{y} := \binom{x+2y}{2x+4y}$. Man sieht leicht, dass f linear ist. Sei $S = [\vec{e}_1, \vec{e}_2]$ die Standardbasis. Dann gilt also $f(\vec{e}_1) = \vec{e}_1 + 2\vec{e}_2$ und $f(\vec{e}_2) = 2\vec{e}_1 + 4\vec{e}_2$. Daher ${}^S_S f = \left(\begin{smallmatrix}1 & 2\\ 2 & 4\end{smallmatrix}\right)$.

Sei nun $\vec{b}_1 = \binom{1}{2}$ und $\vec{b}_2 = \binom{-2}{1}$. Dann ist $B = [\vec{b}_1, \vec{b}_2]$ eine Basis von $\mathbb{R}^2$, und wir können ${}^B_B f$ berechnen: $f(\vec{b}_1) = \binom{1+2\cdot 2}{2\cdot 1+4\cdot 2} = \binom{5}{10} = 5\vec{b}_1 + 0\vec{b}_2$ und $f(\vec{b}_2) = \binom{-2+2\cdot 1}{2\cdot(-2)+4\cdot 1} = \vec{0} = 0\vec{b}_1 + 0\vec{b}_2$. Daher ${}^B_B f = \left(\begin{smallmatrix}5 & 0\\ 0 & 0\end{smallmatrix}\right)$. Offenbar ist die Basis B besser zur Beschreibung von f geeignet als S.

Beobachtung 4.6 (und Definition)
Seien $B = [\vec{b}_1, ..., \vec{b}_n]$ und C endliche Basen eines $\mathbb{K}$-Vektorraums V. Die **Basiswechselmatrix** *für die Transformation von B nach C ist ${}^C_B\mathbb{T} := {}^C_B\mathrm{Id}_V \in \mathbb{K}^{|C|\times n}$. Wir werden noch zeigen, dass in diesem Fall $|C| = n$ gilt.*

a) Für alle $\vec{v} \in V$ gilt ${}^C\vec{v} = {}^C_B\mathbb{T} \cdot {}^B\vec{v}$.

b) Ist $f\colon V \to W$ linear, B_1, B_2 endliche Basen von V, C_1, C_2 endliche Basen von W, dann ${}^{C_2}_{B_2} f = {}^{C_2}_{C_1}\mathbb{T} \cdot {}^{C_1}_{B_1} f \cdot {}^{B_1}_{B_2}\mathbb{T}$.

Beispiel 4.7
In Beispiel 4.5 ist ${}^S_B\mathbb{T} = (\vec{b}_1, \vec{b}_2) = \left(\begin{smallmatrix}1 & -2\\ 2 & 1\end{smallmatrix}\right)$. Man berechnet $\mathrm{LR}((\vec{b}_1, \vec{b}_2), \vec{e}_1) = \{\frac{1}{5}\binom{1}{-2}\}$ und $\mathrm{LR}((\vec{b}_1, \vec{b}_2), \vec{e}_2) = \{\frac{1}{5}\binom{2}{1}\}$. Daher ${}^B_S\mathbb{T} = \frac{1}{5}\left(\begin{smallmatrix}1 & 2\\ -2 & 1\end{smallmatrix}\right)$. Die Abbildungsmatrix von f bezüglich der „günstigen" Basis B lässt sich also aus der gegebenen „ungünstigen" Basis berechnen durch $\frac{1}{5}\left(\begin{smallmatrix}1 & 2\\ -2 & 1\end{smallmatrix}\right) \cdot \left(\begin{smallmatrix}1 & 2\\ 2 & 4\end{smallmatrix}\right) \cdot \left(\begin{smallmatrix}1 & -2\\ 2 & 1\end{smallmatrix}\right) = \left(\begin{smallmatrix}5 & 0\\ 0 & 0\end{smallmatrix}\right)$.

4.1.1 Der Dimensionsbegriff

Dimensionssatz
Sind $B = [\vec{v}_1, \ldots, \vec{v}_m]$ und $C = [\vec{w}_1, \ldots, \vec{w}_n]$ Basen eines $\mathbb{K}$-Vektorraums V, dann ist $m = n$.

Dadurch wird die folgende Definition sinnvoll:

Definition 4.8. *Sei V ein endlich dimensionaler $\mathbb{K}$-Vektorraum.*
Hat V eine Basis $[\vec{v}_1, \ldots, \vec{v}_n]$, so hat V **Dimension** *n. Bezeichnung:* $\dim(V) = n$.

Beispiel $\dim(\mathbb{R}^n) = n$ bzw. $\dim(\mathbb{R}[X]_n) = n + 1$, da $[\vec{e}_1, \ldots, \vec{e}_n]$ bzw. $[X^0, ..., X^n]$ Basen sind.

Beweis des Dimensionssatzes:
Für $i \in \{1, ..., m\}$ bzw. $j \in \{1, ..., n\}$ gelten ${}^B\vec{v}_i = \vec{e}_i \in \mathbb{K}^m$ und ${}^C\vec{w}_j = \vec{e}_j \in \mathbb{K}^n$. Also ${}^B_B\mathbb{T}_{i,i} = 1 = {}^C_C\mathbb{T}_{j,j}$ und

$$m = \sum_{i=1}^{m} {}^B_B\mathbb{T}_{i,i} = \sum_{i=1}^{m}\sum_{j=1}^{n} {}^B_C\mathbb{T}_{i,j}\,{}^C_B\mathbb{T}_{j,i} = \sum_{j=1}^{n}\sum_{i=1}^{m} {}^C_B\mathbb{T}_{j,i}\,{}^B_C\mathbb{T}_{i,j} = \sum_{j=1}^{n} {}^C_C\mathbb{T}_{j,j} = n\,. \qquad \square$$

4.2 Rechnen mit Basen

Wir verallgemeinern nun Problem 3.17 für beliebige endlich dimensionale $\mathbb{K}$-Vektorräume.

Auswahlsatz
Sei V ein $\mathbb{K}$-Vektorraum mit einem endlichen Erzeugendensystem $[\vec{v}_1, \ldots, \vec{v}_n]$ und sei $[\vec{v}_1, \ldots, \vec{v}_k]$ linear unabhängig. Durch Weglassen einiger geeigneter Vektoren $\vec{v}_{k+1}, ..., \vec{v}_n$ erhält man eine Basis von V.

Der Beweis ist konstruktiv, sofern man lineare Abhängigkeiten berechnen kann. Ein Beispiel gab es bereits bei Problem 3.17.

Beweis:
Gibt es eine lineare Abhängigkeit $\vec{0} = \sum_{i=1}^{n} c_i \vec{v}_i$, so gibt es ein $j \in \{k+1, ..., n\}$ mit $c_j \neq 0$ (andernfalls wäre $[\vec{v}_1, ..., \vec{v}_k]$ linear abhängig). Nach Lemma 3.19 kann man $\vec{v}_j$ weglassen, ohne das Erzeugnis (also V) zu ändern. Man wiederholt, bis man ein linear unabhängiges Erzeugendensystem (also eine Basis) erhält. Das ist nach endlich vielen Schritten der Fall. □

Existenzsatz
Jeder endlich dimensionale $\mathbb{K}$-Vektorraum hat eine Basis.

Beweis:
Den Auswahlsatz auf ein endliches Erzeugendensystem anwenden. □

Bemerkung 4.9
Es ist äquivalent zum Auswahlaxiom, dass jeder Vektorraum eine Basis hat, also auch $\mathscr{C}^1(\mathbb{R}, \mathbb{R})$, $\mathbb{K}^I$ für $|I| = \infty$ und der $\mathbb{Q}$-Vektorraum $\mathbb{R}$.

Basisergänzungssatz
Jedes linear unabhängige System $\vec{v}_1, \ldots, \vec{v}_r$ in einem endlich dimensionalen $\mathbb{K}$-Vektorraum V lässt sich zu einer Basis von V fortsetzen.

Beweis:
Ist $\mathrm{Span}(\vec{w}_1, \ldots, \vec{w}_m) = V$, ist auch $\vec{v}_1, \ldots, \vec{v}_r, \vec{w}_1, \ldots, \vec{w}_m$ ein Erzeugendensystem. Wähle unter „Verschonung“ von $\vec{v}_1, ..., \vec{v}_r$ eine Basis aus. □

Beispiel 4.10
Ergänze $\vec{v}_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$, $\vec{v}_2 = \begin{pmatrix} 1 \\ 4 \\ 6 \end{pmatrix}$ zu einer Basis von $\mathbb{R}^3$:
$\left(\begin{smallmatrix} 1 & 1 & 1 & 0 & 0 \\ 2 & 4 & 0 & 1 & 0 \\ 3 & 6 & 0 & 0 & 1 \end{smallmatrix}\right) \rightsquigarrow \left(\begin{smallmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 2 & -2 & 1 & 0 \\ 0 & 0 & 0 & -3/2 & 1 \end{smallmatrix}\right)$, *also ist $[\vec{v}_1, \vec{v}_2, \vec{e}_2]$ eine Basis von $\mathbb{R}^3$.*

Korollar 4.11
Sei V ein n-dimensionaler Vektorraum.

a) Jedes Erzeugendensystem von V hat Länge $\geq n$. Jedes Erzeugendensystem der Länge n ist eine Basis.

b) Jedes linear unabhängige System in V hat Länge $\leq n$. Jedes linear unabhängige System der Länge n ist eine Basis.

Beweis:
a) Auswahlsatz: Länge n nach Auswahl einer Basis.

b) Basisergänzungssatz: Länge n nach Fortsetzung zu einer Basis. □

4.3 Untervektorräume und Dimensionsformel

Satz 4.12. *Sei V ein endlich dimensionaler $\mathbb{K}$-Vektorraum, $U \leq V$.*
U ist endl. dimensional, $\dim(U) \leq \dim(V)$. $\dim(U) = \dim(V) \iff U = V$.

Dazu eine Anmerkung: Auch für Gruppen betrachtet man Unterstrukturen, so genannte *Untergruppen*. Auch für (Unter-)Gruppen betrachtet man Erzeugendensysteme, wobei es manche Gruppen gibt, die kein endliches Erzeugendensystem besitzen. Es gibt Gruppen mit einem endlichen Erzeugendensystem, deren Untergruppen *nicht* alle ein endliches Erzeugendensystem besitzen. Vor diesem Hintergrund ist die Aussage des obigen Satzes alles andere als selbstverständlich.

Beweis von Satz 4.12:
Sei $n = \dim(V)$. Jede linear unabhängige Familie in V hat Länge $\leq n$. Sei $[\vec{u}_1, \ldots, \vec{u}_r] \subset U \subset V$ maximal linear unabhängig, also eine Basis von U (Charakterisierung von Basen). Nach Korollar 4.11 b) für V ist $\dim(U) = r \leq n$. Ist $r = n$, dann ist $[\vec{u}_1, \ldots, \vec{u}_r]$ nach Korollar 4.11 eine Basis von V, also $U = V$. □

Definition 4.13. *Sei V ein $\mathbb{K}$-Vektorraum und $U, W \leq V$.*

a) Die **Summe** *$U + W$ von U, W ist $U + W := \{\vec{u} + \vec{w} \mid \vec{u} \in U,\ \vec{w} \in W\}$.*

b) Die Summe von U, W heißt **direkt** *$:\Leftrightarrow U \cap W = \{\vec{0}\}$. In diesem Fall schreibt man $U \oplus W$ statt $U + W$. Bei mehr als zwei Summanden würde man von einer direkten Summe sprechen, wenn der Schnitt von je Summanden stets gleich $\{\vec{0}\}$ ist.*

c) Ist $U \oplus W = V$, so nennt man W ein **Komplement** *von U.*[24]

Aufgabe 4.14
Sei V ein $\mathbb{K}$-Vektorraum und $U, W \leq V$. Dann ist $U \cap W \leq V$, $U + W \leq V$, $U \leq U + W$ und $W \leq U + W$.

[24] Insbesondere gilt $U \cap W = \{\vec{0}\}$, sonst dürfte man ja nicht $\oplus$ schreiben. Natürlich ist dann auch U ein Komplement von W.

Beispiel 4.15

a) Sei $U \subseteq \mathbb{R}^5$ der Untervektorraum mit Basis $[\vec{e}_1, \vec{e}_2]$ und $W \subseteq \mathbb{R}^5$ der Untervektorraum mit Basis $[\vec{e}_4, \vec{e}_5]$. Dann ist $U \cap W = \{\vec{0}\}$, also ist die Summe $U + W$ direkt. Das heißt, $U \oplus W$ ist der Untervektorraum des $\mathbb{R}^5$ mit Basis $[\vec{e}_1, \vec{e}_2, \vec{e}_4, \vec{e}_5]$.

b) Sei $U \subseteq \mathbb{R}^3$ der Untervektorraum mit Basis $[\vec{e}_1, \vec{e}_2]$ und $W \subseteq \mathbb{R}^3$ der Untervektorraum mit Basis $[\vec{e}_2, \vec{e}_3]$. Dann ist $\vec{e}_2 \in U \cap W \neq \{\vec{0}\}$. Das heißt: die Summe $U + W$ existiert – es ist $U + W = \mathbb{R}^3$ –, aber diese Summe ist nicht direkt. In diesem Fall verwendet man die Notation $U \oplus W$ nicht.

Folgendes Ergebnis gilt zwar für allgemeine Vektorräume, unsere Mittel reichen aber nur für einen Beweis im endlich dimensionalen Fall.

Lemma 4.16
Sei V ein $\mathbb{K}$-Vektorraum. Jedes $U \leq V$ hat ein Komplement.

Beweis:
Sei B eine Basis von U; Basisergänzungssatz: V hat eine Basis C mit $B \subset C$ (diese Aussagen haben wir für endlich dimensionale Vektorräume bewiesen, unter Annahme des Auswahlaxioms gelten sie aber auch sonst). Behauptung: $W := \mathrm{Span}(C \setminus B)$ ist ein Komplement von U.

Weil C eine Basis von V ist, hat jedes $\vec{v} \in V$ eine eindeutige Darstellung als Linearkombination $\vec{v} = \sum_{\vec{b} \in B}{}' \beta_{\vec{b}} \vec{b} + \sum_{\vec{c} \in C \setminus B}{}' \gamma_{\vec{c}} \vec{c}$, insbesondere $U + W = V$.

Sei $\vec{v} \in U \cap W$. Wegen $\vec{v} \in U$ und Eindeutigkeit der Darstellung folgt $\forall \vec{c} \in C \setminus B\colon\ \gamma_{\vec{c}} = 0$. Wegen $\vec{v} \in W$ folgt ebenso $\forall \vec{b} \in B\colon\ \beta_{\vec{b}} = 0$. Also $\vec{v} = \vec{0}$, anders gesagt $U \cap W = \{\vec{0}\}$. □

Dimensionsformel
Sei V ein $\mathbb{K}$-Vektorraum und $U, W \leq V$ endlich dimensional. Dann

$$\dim(U + W) = \dim(U) + \dim(W) - \dim(U \cap W)$$

Korollar 4.17
Ist $U, W \leq V$ und $U \cap W = \{\vec{0}\}$, so gilt $\dim(U \oplus W) = \dim(U) + \dim(W)$.

Beweis:
Die leere Menge ist Basis von $U \cap W = \{\vec{0}\}$, daher $\dim(U \cap W) = 0$. □

Beispiel Seien U, W zwei 3-dimensionale Untervektorräume von $V = \mathbb{R}^5$, dann ist $\dim(U \cap W) \geq 1$: denn wegen $U + W \subseteq \mathbb{R}^5$ ist $\dim(U + W) \leq 5$, also

$$\dim(U \cap W) = \dim(U) + \dim(W) - \dim(U + W) \geq 3 + 3 - 5 = 1\,.$$

Beweis der Dimensionsformel:
Sei B eine Basis von $U \cap W$. Nach dem Basisergänzungssatz, den wir ja im endlich dimensionalen Fall vollständig bewiesen haben, kann man B zu Basen $C_U \cup B$ von U und $C_W \cup B$ von W ergänzen, wobei $C_U \cap B = C_W \cap B = \emptyset$. Offenbar wird $U + W$ von $B \cup C_U \cup C_W$ erzeugt.

Sei $\vec{0} = \sum_{\vec{b} \in B}{}' \alpha_{\vec{b}} \vec{b} + \sum_{\vec{u} \in C_U}{}' \beta_{\vec{u}} \vec{u} + \sum_{\vec{w} \in C_W}{}' \gamma_{\vec{w}} \vec{w}$. Dann folgt $\sum_{\vec{b} \in B}{}' \alpha_{\vec{b}} \vec{b} + \sum_{\vec{u} \in C_U}{}' \beta_{\vec{u}} \vec{u} = -\sum_{\vec{w} \in C_W}{}' \gamma_{\vec{w}} \vec{w} \in U \cap \mathrm{Span}(C_W)$. Wir sahen im Beweis des vorigen Lemmas, dass $\mathrm{Span}(C_W) \cap (U \cap W) = \{\vec{0}\}$, also wegen $\mathrm{Span}(C_W) \subset W$ bereits $U \cap \mathrm{Span}(C_W) = \{\vec{0}\}$. Weil C_W linear unabhängig ist, sind alle $\gamma_{\vec{w}} = 0$.

Also haben wir $\vec{0} = \sum_{\vec{b} \in B}{}' \alpha_{\vec{b}} \vec{b} + \sum_{\vec{u} \in C_U}{}' \beta_{\vec{u}} \vec{u}$ und es folgt $\sum_{\vec{b} \in B}{}' \alpha_{\vec{b}} \vec{b} = -\sum_{\vec{u} \in C_U}{}' \beta_{\vec{u}} \vec{u} \in (U \cap W) \cap \mathrm{Span}(C_U) = \{\vec{0}\}$, wieder wie im Beweis des vorigen Lemmas. Weil B und C_U linear unabhängig sind, sind auch alle $\alpha_{\vec{b}} = 0$ und $\beta_{\vec{u}} = 0$. Also hat $B \cup C_U \cup C_W$ keine lineare Abhängigkeit und ist somit eine Basis von $U + W$.

Insbesondere sind B, C_U, C_W paarweise disjunkt. Es folgt

$$\begin{aligned} \dim(U+W) &= |B \cup C_U \cup C_W| \overset{\text{disjunkt}}{=} |B| + |C_U| + |C_W| \\ &= |B \cup C_U| + |B \cup C_W| - |B| \\ &= \dim(U) + \dim(W) - \dim(U \cap W) \end{aligned}$$ □

4.4 Rang von Matrizen

Lemma 4.18 (und Definition)
Sei $A = (\vec{v}_1, ..., \vec{v}_n) \in \mathbb{K}^{m \times n}$ *und* $\mathrm{Spaltenraum}(A) := \mathrm{Span}(\vec{v}_1, ..., \vec{v}_n)$. *Es gilt* $\mathrm{Rang}(A) = \dim(\mathrm{Spaltenraum}(A))$. *Insbesondere ist* $\mathrm{Rang}(A)$ *wohldefiniert.*

Beweis:
Bringe A auf ZSF mit Pivotspalten $j_1, ..., j_r$. Nach Definition ist $\mathrm{Rang}(A) = r$. Wir wissen bereits (Basisauswahl), dass $[\vec{v}_{j_1}, ..., \vec{v}_{j_r}]$ eine Basis von $\mathrm{Span}(\vec{v}_1, ..., \vec{v}_n)$ ist. Also ist $\dim(\mathrm{Spaltenraum}(A)) = r = \mathrm{Rang}(A)$. □

Lemma 4.19
Für alle $A \in \mathbb{K}^{m \times n}$ *gilt* $\mathrm{Rang}(A) = \mathrm{Rang}(A^\top)$.

Beweis:
Sei $\mathrm{Zeilenraum}(A)$ das Erzeugnis der Zeilen von A. Bringe A mit Zeilenoperationen auf ZSF A' mit r Pivotspalten. Zeilenoperationen ändern den Zeilernraum nicht, denn die alten Zeilen sind Linearkombinationen der neuen Zeilen und umgekehrt. Also $\mathrm{Zeilenraum}(A) = \mathrm{Zeilenraum}(A')$. Offenbar bilden die r von Null verschiedenen Zeilen von A' eine Basis von $\mathrm{Zeilenraum}(A')$, daher $\mathrm{Rang}(A) = r = \dim(\mathrm{Zeilenraum}(A')) = \dim(\mathrm{Zeilenraum}(A)) = \dim(\mathrm{Spaltenraum}(A^\top)) = \mathrm{Rang}(A^\top)$. □

Lemma 4.20. *Seien $A \in \mathbb{K}^{m\times n}$ und $B \in \mathbb{K}^{n\times \ell}$.*
Es gelten $\mathrm{Rang}(AB) \leq \mathrm{Rang}(A)$ *und* $\mathrm{Rang}(AB) \leq \mathrm{Rang}(B)$.

Beweis:

- $\mathrm{Spaltenraum}(AB) = \{AB\vec{x} \mid \vec{x} \in \mathbb{K}^\ell\} \subset \{A\vec{y} \mid \vec{y} \in \mathbb{K}^n\} = \mathrm{Spaltenraum}(A)$.
- Sei $[\vec{b}_1, ..., \vec{b}_k]$ eine Basis von $\mathrm{Spaltenraum}(B)$. Dann ist $\mathrm{Spaltenraum}(AB) = \{AB\vec{x} \mid \vec{x} \in \mathbb{K}^\ell\} = \{A \cdot (\vec{b}_1, ..., \vec{b}_k) \cdot \vec{y} \mid \vec{y} \in \mathbb{K}^k\} = \mathrm{Span}(A\vec{b}_1, ..., A\vec{b}_k)$, also $\mathrm{Rang}(AB) \leq \mathrm{Rang}(B)$. □

Korollar 4.21

a) $GL_n(\mathbb{K})$ ist eine Gruppe bzgl. Matrixmultiplikation.

b) Seien $A \in \mathbb{K}^{m\times n}$ und $B \in \mathbb{K}^{n\times m}$. Wenn $AB = \mathbb{1}_m$ und $BA = \mathbb{1}_n$, dann gilt $n = m$, $A, B \in GL_n(\mathbb{K})$ und $B = A^{-1}$.

c) Seien C, D Basen eines n-dimensionalen $\mathbb{K}$-Vektorraums. Dann ${}^D_C\mathbb{T}$, ${}^C_D\mathbb{T} \in GL_n(\mathbb{K})$ und ${}^D_C\mathbb{T}^{-1} = {}^C_D\mathbb{T}$.

Beweis:

a) Seien $A, B \in GL_n(\mathbb{K})$. Im Hinblick auf frühere Ergebnisse bleibt nur noch zu zeigen, dass $AB \in GL_n(\mathbb{K})$. Nach Definition von $GL_n(\mathbb{K})$ ist also nur zu zeigen: $\mathrm{Rang}(AB) = n$.

Wir zeigten bereits, dass zu $A, B \in GL_n(\mathbb{K})$ inverse Matrizen existieren. Dann $n = \mathrm{Rang}(\mathbb{1}_n) = \mathrm{Rang}(ABB^{-1}A^{-1}) \leq \mathrm{Rang}(AB) \leq n$, die erste Ungleichung wegen des vorigen Lemmas, die zweite Ungleichung wegen $AB \in M_n(\mathbb{K})$.

b) $m = \mathrm{Rang}(\mathbb{1}_m) = \mathrm{Rang}(AB) \leq \mathrm{Rang}(B) \leq n$, $n = \mathrm{Rang}(\mathbb{1}_n) = \mathrm{Rang}(BA) \leq \mathrm{Rang}(A) \leq m$, also $m = n$ und $\mathrm{Rang}(A) = \mathrm{Rang}(B) = n$.

c) ${}^C_B\mathbb{T}$, ${}^B_C\mathbb{T} \in M_n(\mathbb{K})$ und ${}^C_B\mathbb{T} \cdot {}^B_C\mathbb{T} = {}^B_C\mathbb{T} \cdot {}^C_B\mathbb{T} = \mathbb{1}_n$ (nämlich: Wechsel von B-Koordinaten nach B-Koordinaten wird durch Multiplikation mit $\mathbb{1}_n$ beschrieben). □

Rangformel für Matrizen
Sei $A \in \mathbb{K}^{m\times n}$. Dann $n = \mathrm{Rang}(A) + \dim\Big(\mathrm{LR}(A;\, \vec{0})\Big)$.

Beweis:
Man bringe A auf Zeilenstufenform mit Pivotspalten $j_1 < ... < j_r$. Dann ist $\mathrm{Rang}(A) = r$ und es gibt $n - r$ Nichtpivotspalten, also auch $n - r$ Basislösungen. Diese bilden eine Basis von $\mathrm{LR}(A; \vec{0})$, also $\dim(\mathrm{LR}(A; \vec{0}) = n - r$. □

4.5 Rang linearer Abbildungen

Aufgabe 4.22 (und Definition)
Sei $f\colon V \to W$ $\mathbb{K}$*-linear,* $U_V \leq V$ *und* $U_W \leq W$.

a) $f(U_V) \leq W$. *Falls* $\mathrm{Bild}(f) \leq W$ *endlich dimensional ist, definieren wir* $\mathrm{Rang}(f) := \dim(\mathrm{Bild}(f))$.

b) $f^{-1}(U_W) \leq V$. *Wir definieren den* **Kern** *von* f *als* $\ker(f) := f^{-1}(\{\vec{0}\}) = \{\vec{v} \in V \mid f(\vec{v}) = \vec{0}\}$. *Es gilt also* $\ker(f) \leq V$.

Beispiel 4.23
a) Sei $D\colon \mathscr{C}^1(\mathbb{R}) \to \mathscr{C}^0(\mathbb{R})$ *definiert durch* $D(f) := f'$ *für jedes stetig differenzierbare* $f\colon \mathbb{R} \to \mathbb{R}$. *Es ist* $f \in \ker(D) \iff f' = 0$. *Mit anderen Worten:* $\ker(D)$ *ist die Menge der konstanten Funktionen.*

b) Sei $S\colon \mathbb{R}[t]_6 \to \mathbb{R}[x]_7$ *definiert durch*

$$S(p) := (x \mapsto \int_{-x}^{x} p(t)\,\mathrm{d}t)$$

Man überlegt sich leicht: Weil über ein symmetrisches Intervall integriert wird, fallen im Integral die geraden Potenzen von x *weg. Es ist* $\mathrm{Bild}(S) = \mathrm{Span}(x^1, x^3, x^5, x^7)$, *also* $\mathrm{Rang}(S) = 4$. *Ferner* $\ker(S) = \mathrm{Span}(t^1, t^3, t^5)$, *also* $\dim(\ker(S)) = 3$.

Problem 4.24
Sei $f\colon V \to W$ *linear und seien* B *bzw.* C *endliche Basen von* V *bzw. von* W. *Berechne* $\mathrm{Bild}(f)$ *und* $\ker(f)$.

Lösung:
- $\mathrm{Bild}(f) = \{\vec{w} \in W \mid {}^C\vec{w} \in \mathrm{Spaltenraum}({}^C_B f)\}$
- $\ker(f) = \{\vec{v} \in V \mid {}^B\vec{v} \in \mathrm{LR}({}^C_B f; \vec{0})\}$ □

Lemma 4.25
Eine lineare Abbildung $f\colon V \to W$ *ist injektiv gdw.* $\ker(f) = \{\vec{0}_V\}$.

Beweis:
Wir wissen bereits $f(\vec{0}_V) = \vec{0}_W$, d.h. $\vec{0}_V \in \ker(f)$.

- f injektiv $\Rightarrow |f^{-1}(\{\vec{0}_W\})| = |\ker(f)| = 1$.
- f nicht injektiv $\Rightarrow \exists \vec{v}_1 \neq \vec{v}_2 \in V\colon\ f(\vec{v}_1) = f(\vec{v}_2) \Rightarrow \vec{0}_W = f(\vec{v}_1) - f(\vec{v}_2) = f(\vec{v}_1 - \vec{v}_2) \Rightarrow \vec{0}_V \neq \vec{v}_1 - \vec{v}_2 \in \ker(f)$. □

Definition 4.26

a) *Eine bijektive $\mathbb{K}$-lineare Abbildung $f\colon V \to W$ nennt man* **Isomorphismus** *(oder genauer:* **$\mathbb{K}$-Vektorraumisomorphismus***). Notation: $f\colon V \xrightarrow{\cong} W$.*

b) *$\mathbb{K}$-Vektorräume heißen* **isomorph***, $V \cong W :\Leftrightarrow \exists f\colon V \xrightarrow{\cong} W$.*

Beispiel 4.27

Sei V ein $\mathbb{K}$-Vektorraum mit Basis $B = (\vec{v}_i)_{i\in I}$. Dann ist $\Phi_B\colon V \to \mathbb{K}^I_{fin}$ ein Isomorphismus. Wir wissen nämlich bereits, dass Φ_B linear ist. Es ist injektiv, denn jedes $\vec{v} \in V$ ist durch den Koordinatenvektor ${}^{\in}\vec{v}\,\mathbb{K}^I_{fin}$ eindeutig bestimmt. Es ist surjektiv, denn jedes Element von $\mathbb{K}^I_{fin}$ entspricht einer Linearkombination von B.

Lemma 4.28. *Sei $f\colon V \xrightarrow{\cong} W$.*

a) *Auch $f^{-1}\colon W \xrightarrow{\cong} V$.*

b) *$\cong$ ist eine Äquivalenzrelation.*

c) *Ist B eine Basis von V, dann ist $f(B)$ eine Basis von W. Insbesondere $\dim(V) = \dim(W)$, falls V endlich dimensional ist.*

Beweis:

a) Seien $\vec{w}_1, \vec{w}_2 \in W$ und $\alpha_1, \alpha_2 \in \mathbb{K}$. Seien $\vec{v}_1 := f^{-1}(\vec{w}_1)$ und $\vec{v}_2 := f^{-1}(\vec{w}_2)$. Wegen Linearität von f gilt $f\left(\alpha_1 f^{-1}(\vec{w}_1) + \alpha_2 f^{-1}(\vec{w}_2)\right) = f(\alpha_1\vec{v}_1 + \alpha_2\vec{v}_2) = \alpha_1 f(\vec{v}_1) + \alpha_2 f(\vec{v}_2) = \alpha_1\vec{w}_1 + \alpha_2\vec{w}_2$. Anwendung von f^{-1} auf beiden Seiten ergibt $\alpha_1 f^{-1}(\vec{w}_1) + \alpha_2 f^{-1}(\vec{w}_2) = f^{-1}(\alpha_1\vec{w}_1 + \alpha_2\vec{w}_2)$, das heißt f^{-1} ist linear. Außerdem ist f^{-1} bijektiv.

b) Übung.

c)
- f ist linear, daher $f(\sum'_{\vec{b}\in B} \alpha_{\vec{b}}\vec{b}) = \sum'_{\vec{b}\in B} \alpha_{\vec{b}} f(\vec{b})$. Das heißt, $f(B)$ ist ein Erzeugendensystem von $\mathrm{Bild}(f) = W$.
- Sei $\vec{0}_W = \sum'_{\vec{b}\in B} \alpha_{\vec{b}} f(\vec{b})$. Anwendung von f^{-1} ergibt $\vec{0}_V = f^{-1}(\vec{0}_W) = \sum'_{\vec{b}\in B} \alpha_{\vec{b}}\vec{b}$. Weil B linear unabhängig ist, folgt $\forall \vec{b} \in B\colon \alpha_{\vec{b}} = 0$, d.h. $f(B)$ besitzt keine lineare Abhängigkeit. □

Satz 4.29

Sei $f\colon V \to W$ linear und sei $U \leq V$ ein Komplement von $\ker(f) \leq V$. Dann gilt $U \cong \mathrm{Bild}(f)$.

Beweis:
Wir definieren $\varphi\colon U \to \text{Bild}(f)$ durch $\varphi(\vec{u}) := f(\vec{u})$. Offenbar ist φ linear. Wir müssen zeigen, dass φ bijektiv ist.

- Wegen $V = U \oplus \ker(f)$ gibt es für jedes $\vec{v} \in V$ ein $\vec{u} \in U$ und ein $\vec{v}_0 \in \ker(f)$ mit $\vec{v} = \vec{u} + \vec{v}_0$. Dann $f(\vec{v}) = \underbrace{f(\vec{u})}_{=\varphi(\vec{u})} + \underbrace{f(\vec{v}_0)}_{=\vec{0}} = \varphi(\vec{u})$. Also $\text{Bild}(f) = \text{Bild}(\varphi)$, d.h. φ ist surjektiv
- Sei $\vec{u} \in \ker(\varphi) \leq U$. Dann $\vec{0} = \varphi(\vec{u}) = f(\vec{u})$, also $\vec{u} \in \ker(f)$ und zudem $\vec{u} \in U$. Aber $U \cap \ker(f) = \{\vec{0}\}$, also $\vec{u} = \vec{0}$. Es folgt $\ker(\varphi) = \{\vec{0}\}$, also ist φ injektiv. □

Rangformel für lineare Abbildungen
Sei $f\colon V \to W$ $\mathbb{K}$-linear und V endlich dimensional. Dann gilt

$$\dim(V) = \text{Rang}(f) + \dim(\ker(f))\,.$$

Beweis:
$\ker(f)$ hat ein Komplement $U \leq V$ (Weil V endlich dimensional ist, haben wir dafür sogar einen Beweis geführt). Nach dem vorigen Satz und Korollar 4.17 gilt $\dim(V) = \dim(U \oplus \ker(f)) = \dim(U) + \dim(\ker(f)) = \dim(\text{Bild}(f)) + \dim(\ker(f))$. □

Beispiel 4.30
Wir betrachten wieder $S\colon \mathbb{R}[t]_6 \to \mathbb{R}[x]_7$ wie in Beispiel 4.23.b). Wir fanden $\text{Rang}(S) = 4$ *und* $\dim(\ker(S)) = 3$*, und in der Tat* $\dim(\mathbb{R}[t]_6) = 7 = 4 + 3 = \text{Rang}(S) + \dim(\ker(S))$.

5 Die Determinante

Sei $\mathbb{K}$ ein kommutativer Ring. Ziel: Ordne jedem $A \in M_n(\mathbb{K})$ die *Determinante* $\det(A) \in \mathbb{K}$ zu, so dass A invertierbar $\Leftrightarrow$ $\det(A)$ invertierbar in $\mathbb{K}$. Geometrische Deutung: Ist $\vec{v}_1, ..., \vec{v}_n \in \mathbb{R}^n$, so ist $|\det(\vec{v}_1, ..., \vec{v}_n)|$ das Volumen von $\{\sum_{i=1}^{n} \lambda_i \vec{v}_i \mid \lambda_1, ..., \lambda_n \in [0,1]\}$.

Nach Formulierung der Definition zeigen wir die *Eindeutigkeit* und eine Berechnungsformel. Weil die Berechnungsformel tatsächlich die definierenden Eigenschaften der Determinante erfüllt, folgt die *Existenz*. Die Formel ist hochgradig ineffizient. Daher kümmern wir uns um *effiziente* alternative Berechnungsmöglichkeiten und nicht zuletzt auch um *strukturelle Eigenschaften* der Determinante.

Definition 5.1
Sei $n \in \mathbb{N}$. Eine Abbildung $\det\colon (\mathbb{K}^n)^n \to \mathbb{K}$ *heißt* **Determinantenfunktion** *(kurz: „Determinante"), wenn sie folgende Eigenschaften hat:*

a) *Sie ist* **multilinear**: *Für jedes $j \in \{1, ..., n\}$ ist sie linear in Stelle j, also $\forall \lambda, \mu \in \mathbb{K}$, $\vec{a}_1, ..., \vec{a}_n, \vec{b} \in \mathbb{K}^n$:* $\det(\vec{a}_1, ..., \vec{a}_{j-1}, \lambda\vec{a}_j + \mu\vec{b}, \vec{a}_{j+1}, ...) = \lambda \det(\vec{a}_1, ..., \vec{a}_n) + \mu \det(\vec{a}_1, ..., \vec{a}_{j-1}, \vec{b}, \vec{a}_{j+1}, ...)$.

b) *Sie ist* **alternierend**: *Sie ist Null, wenn zwei Argumente gleich sind, d.h.* $\det(..., \vec{a}, ..., \vec{a}, ...) = 0$.

c) *Sie ist* **normiert**: $\det(\vec{e}_1, ..., \vec{e}_n) = 1$.

Ist $A \in M_n(\mathbb{K})$ mit Spalten $\vec{a}_1, ..., \vec{a}_n$, so setzen wir $\det(A) := \det(\vec{a}_1, ..., \vec{a}_n)$. *Für $n \geq 2$ ist auch die Notation $|A| := \det(A)$ üblich.*

Vorsicht: $\det(\lambda A) = \lambda^n \det(A)$ und im Allg. $\det(A + B) \neq \det(A) + \det(B)$. Für $A = (a) \in M_1(\mathbb{K})$ ist $|A| = a$ — bitte **NICHT** $|A| = |a|$!!

„Ausmultiplizieren" von multilinearen Abbildungen Wir bilden zu Vektoren $\vec{a}_1, ..., \vec{a}_n \in \mathbb{K}^n$ insgesamt n Linearkombinationen, wobei die j-te Linearkombination die Koeffizienten $b_{1,j}, ..., b_{n,j} \in \mathbb{K}$ hat. Wenn $f\colon (\mathbb{K}^n)^n \to \mathbb{K}$ multilinear ist, so gilt

$$\begin{aligned} f(\sum_{i_1=1}^{n} b_{i_1,1}\vec{a}_{i_1}, ..., \sum_{i_n=1}^{n} b_{i_n,n}\vec{a}_{i_n}) &= \sum_{i_1=1}^{n} b_{i_1,1} f(a_{i_1}, \sum_{i_2=1}^{n} b_{i_1,2}\vec{a}_{i_2}, ..., \sum_{i_n=1}^{n} b_{i_n,n}\vec{a}_{i_n}) \\ &= \sum_{i_1=1}^{n}\sum_{i_2=1}^{n} b_{i_1,1}b_{i_2,2} f(a_{i_1}, a_{i_2}, ..., \sum_{i_n=1}^{n} b_{i_n,n}\vec{a}_{i_n}) \end{aligned}$$

also insgesamt $f(\sum_{i_1=1}^{n} b_{i_1,1}\vec{a}_{i_1}, ..., \sum_{i_n=1}^{n} b_{i_n,n}\vec{a}_{i_n}) = \sum_{i_1=1}^{n} ... \sum_{i_n=1}^{n} b_{i_1,1} \cdots b_{i_n,n} f(\vec{a}_{i_1}, ..., \vec{a}_{i_n})$

Lemma 5.2
Determinanten sind **schiefsymmetrisch** *(auch: antisymmetrisch), d.h. Spaltentausch bewirkt Vorzeichenwechsel:* $\forall \vec{a}, \vec{b}$: $\det(..., \vec{a}, ..., \vec{b}, ...) = -\det(..., \vec{b}, ..., \vec{a}, ...)$.

Beweis:
Alternierend und multilinear $\Rightarrow$ $0 = \det(..., \vec{a} + \vec{b}, ..., \vec{a} + \vec{b}, ...)$
$= \underbrace{\det(..., \vec{a}, ..., \vec{a}, ...)}_{=0} + \det(..., \vec{a}, ..., \vec{b}, ...) + \det(..., \vec{b}, ..., \vec{a}, ...) + \underbrace{\det(..., \vec{b}, ..., \vec{b}, ...)}_{=0}$. □

Eindeutigkeit und Existenz der Determinante
Ist $\det\colon (\mathbb{K}^n)^n \to \mathbb{K}$ *eine Determinantenfunktion, so gilt für alle* $A \in M_n(\mathbb{K})$ *die* **Leibnizformel**[25] $\det(A) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{j=1}^{n} A_{\sigma(j),j}$. *Umgekehrt erfüllt die Leibnizformel die definierenden Eigenschaften einer Determinantenfunktion.*

Bemerkung Die Leibnizformel ist sehr ineffizient, da sie $n!$ Summanden hat.

Beweis:
Eigenschaften $\Rightarrow$ Formel: Für $j \in \{1, ..., n\}$ ist die j-te Spalte von A gleich $\sum_{i=1}^{n} A_{i,j} \vec{e}_i$. Multilinearität: $\det(A) = \sum_{i_1=1}^{n} ... \sum_{i_n=1}^{n} A_{i_1,1} \cdots A_{i_n,n} \det(\vec{e}_{i_1}, ..., \vec{e}_{i_n})$. Ist $j \mapsto i_j$ keine Permutation, so sind in $\det(\vec{e}_{i_1}, ..., \vec{e}_{i_n})$ zwei Spalten gleich, der betreffende Summand ist also Null (alternierend!).

Daher $\det(A) = \sum_{\sigma \in S_n} \det(\vec{e}_{\sigma(1)}, ..., \vec{e}_{\sigma(n)}) \prod_{j=1}^{n} A_{\sigma(j),j}$. Weil $\operatorname{sgn}(\sigma)$ Vertauschungen zählt, ist $\det(\vec{e}_{\sigma(1)}, ..., \vec{e}_{\sigma(n)}) = \operatorname{sgn}(\sigma) \det(\vec{e}_1, ..., \vec{e}_n) = \operatorname{sgn}(\sigma)$, wegen Normierung und Schiefsymmetrie.

Formel $\Rightarrow$ Eigenschaften:

Multilinear: $\prod_{j=1}^{n} A_{\sigma(j),j}$ enthält aus Spalte j genau einen Term.

Alternierend: Sind Spalten j, ℓ für $j \neq \ell$ gleich, so heben sich die Terme für $\sigma \in S_n$ und $(j\ \ell) \circ \sigma$ auf.

Normierung: Ist $A = \mathbb{1}_n$, bleibt von der Leibnizformel nur der Term für $\sigma = \mathrm{Id}$ und dieser liefert den Wert 1. □

Beispiel 5.3
Im Spezialfall $n \in \{2, 3\}$ *erhält man die* **Sarrus-Regel**[26]*:* $\left|\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right| = ad - bc$ *(siehe Hausaufgaben) und* $\left|\begin{smallmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{smallmatrix}\right| = a_1b_2c_3 + a_2b_3c_1 + a_3b_1c_2 - a_3b_2c_1 - a_1b_3c_2 - a_2b_1c_3$.

[25] Gottfried Wilhelm Leibniz [1646–1716]
[26] Pierre Frédéric Sarrus [1798–1861]

Bemerkung 5.4
Leider haben manche Studierende die Tendenz, die Sarrus-Regel als alleinseligmachend zu betrachten. Daher sei hier unmissverständlich betont: Die Sarrus-Regel ist sehr nützlich für 2×2-Matrizen. Sie ist korrekt, aber meist ineffizient, für 3×3-Matrizen. Und: Die Verallgemeinerung des Diagonalschemas der Sarrus-Regel auf $n \times n$-Matrizen mit $n > 3$ ist falsch.

Lemma 5.5
Sei $A \in M_n(\mathbb{K})$. Dann $\det(A^\top) = \det(A)$.

Beweis:
Induktion nach n, Verankerung $n = 0$ klar. Für $n > 0$ gilt

$$\begin{aligned}\det(A) &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{j=1}^{n} A_{\sigma(j),j} \overset{k=\sigma(j)}{=} \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{k=1}^{n} A_{k,\sigma^{-1}(k)} \\ &= \sum_{\sigma^{-1} \in S_n} \operatorname{sgn}(\sigma^{-1}) \prod_{k=1}^{n} A^\top_{\sigma^{-1}(k),k} = \det(A^\top)\,. \end{aligned}$$

□

Zwar ändert sich der Wert einer Determinante im Allgemeinen, wenn man Zeilenoperationen durchführt. Doch wegen des folgenden Lemmas ist der Gauß-Algorithmus der Schlüssel zu einer allgemeinen effizienten Berechnungsmethode für Determinanten, wenn $\mathbb{K}$ ein Körper ist. Diesmal lassen wir zusätzlich auch Spaltenoperationen zu.

Lemma 5.6 (Determinante und Zeilenoperationen)
Seien $A, A' \in M_n(\mathbb{K})$. Entsteht A' aus A durch ...

a) Multiplikation einer Zeile/Spalte mit $\lambda \in \mathbb{K}$, dann $\det(A') = \lambda \det(A)$.

b) Addition des μ-Fachen von Zeile/Spalte i zu Zeile/Spalte j ($i \neq j$, $\mu \in \mathbb{K}$), dann $\det(A') = \det(A)$.

c) Tausch zweier verschiedener Zeilen/Spalten, dann $\det(A') = -\det(A)$.

Beweis:
Wir betrachten Spalten. Für Zeilen folgt es aus Lemma 5.5.

a) Die Determinante ist multilinear.

b) $\det(..., \vec{a}_i, ..., \mu\vec{a}_i + \vec{a}_j, ...) = \mu \underbrace{\det(..., \vec{a}_i, ..., \vec{a}_i, ...)}_{=0} + \det(..., \vec{a}_i, ..., \vec{a}_j, ...)$.

c) Lemma 5.2. □

Ist also $\mathbb{K}$ ein Körper, lässt sich die Berechnung der Determinante von $A \in M_n(\mathbb{K})$ in $\mathcal{O}(n^3)$ $\mathbb{K}$-Operationen mit dem Gauß-Algorithmus auf die Berechnung der Determinante einer ZSF zurückführen. In der Praxis ist es oft sinnvoll, zusätzlich Spaltenoperationen anzuwenden und zudem die im folgenden betrachteten Spezialfälle zu nutzen.

Laplacescher Entwicklungssatz
Sei $A \in M_n(\mathbb{K})$ mit $n > 0$. Laplace[27]-Entwicklung nach Zeile i bzw. Spalte j:

$$\begin{aligned} \det(A) &= \sum_{j=1}^{n} (-1)^{i+j} A_{ij} \det(A(i,j)) && i \text{ fest} \\ &= \sum_{i=1}^{n} (-1)^{i+j} A_{ij} \det(A(i,j)) && j \text{ fest} \end{aligned}$$

Dabei entsteht $A(i,j) \in M_{n-1}(\mathbb{K})$ aus A durch Streichen von Zeile i und Spalte j. Die Determinante einer Untermatrix von A bezeichnet man übrigens als **Minore**.
Merkregel für die Vorzeichen: Schachbrett-Muster $\begin{pmatrix} + & - & + & - & \cdots \\ - & + & - & & \\ + & - & & & \\ - & & & & \\ \vdots & & & & \end{pmatrix}$

Rekursion mit Laplace-Entwicklung ist genau so ineffizient wie die Leibnizformel, denn sie führt auf $n!$ Summanden. Sie ist allerdings praktisch, wenn es eine Zeile gibt, in der nur ganz wenige Einträge von Null verschieden sind.

Beweis:
Wir zeigen zunächst die Aussage zur Entwicklung nach der i-ten Zeile. Wegen der Eindeutigkeit der Determinante genügt zu zeigen:
$A \mapsto \sum_{j=1}^{n} (-1)^{i+j} A_{i,j} \det(A(i,j))$ ist eine Determinantenfunktion.

Multilinear: Spalte ℓ von A sei $\lambda\vec{b} + \mu\vec{c}$ und es entstehe B, C aus A, indem Spalte ℓ durch $\vec{b}$ bzw. $\vec{c}$ ersetzt wird. Für $j \neq \ell$ ist $\det(A(i,j)) = \lambda \det(B(i,j)) + \mu \det(C(i,j))$. Und $A_{i,\ell} \det(A(i,\ell)) = \lambda B_{i,\ell} \det(B(i,\ell)) + \mu C_{i,\ell} \det(C(i,\ell))$ wegen $A(i,\ell) = B(i,\ell) = C(i,\ell)$.

Alternierend: Sei $\ell < m$, $\vec{a}_\ell = \vec{a}_m$. $A(i,m)$ entsteht aus $A(i,\ell)$ durch $m - \ell - 1$ Spaltenvertauschungen. Es gilt $\det(A(i,m)) = (-1)^{m-\ell-1} \det(A(i,\ell))$ und für $j \notin \{\ell, m\}$ ist $\det(A(i,j)) = 0$. Also $\sum_{j=1}^{n} (-1)^{i+j} A_{i,j} \det A(i,j) = (-1)^{i+\ell} A_{i,\ell} \det(A(i,\ell)) + (-1)^{i+m+m-\ell-1} A_{i,\ell} \det(A(i,\ell)) \overset{\checkmark}{=} 0$.

Normierung: $\mathbb{1}_n \mapsto (-1)^{i+i} \cdot 1 \cdot \det(\mathbb{1}_{n-1}) \overset{\checkmark}{=} 1$.

Die Spaltenentwicklung folgt daraus mit Lemma 5.5. □

[27] Pierre-Simon Marquis de Laplace [1749–1827]

Definition 5.7

a) Ist $A \in M_n(\mathbb{K})$ *mit* $A = \left(\begin{smallmatrix} B & C \\ \mathbb{0} & D \end{smallmatrix}\right) \in M_n(\mathbb{K})$ *oder* $A = \left(\begin{smallmatrix} B & \mathbb{0} \\ C & D \end{smallmatrix}\right)$ *mit quadratischen Matrizen* B, D*, so hat* A **Blockgestalt**.

b) $A \in M_n(\mathbb{K})$ *heißt* **obere Dreiecksmatrix** *(bzw.* **untere Dreiecksmatrix***), wenn* $\forall i, j \in \{1, ..., n\}$ *mit* $i > j$ *(bzw.* $i < j$*) gilt* $A_{i,j} = 0$.

Beispiel $\begin{pmatrix} 1&2&3&4&5 \\ 2&3&4&5&6 \\ 0&0&1&2&3 \\ 0&0&2&3&4 \\ 0&0&3&4&5 \end{pmatrix}$ hat Blockgestalt, $\begin{pmatrix} 1&2&3&4 \\ 0&5&6&7 \\ 0&0&0&8 \\ 0&0&0&9 \end{pmatrix}$ ist obere Dreiecksmatrix.

Lemma 5.8

Für alle $A \in M_n(\mathbb{K})$ *gilt:*

a) Hat $A = \left(\begin{smallmatrix} B & C \\ \mathbb{0} & D \end{smallmatrix}\right)$ *Blockgestalt, dann* $\det(A) = \det(B) \cdot \det(D)$.

b) Ist A *eine obere oder untere Dreiecksmatrix, dann* $\det(A) = \prod_{i=1}^{n} A_{i,i}$.

Beweis:

a) Sei $B \in M_m(\mathbb{K})$ mit $m < n$. Induktion nach $n-m$ mit Laplace-Entwicklung. Verankerung $m = n$: Die Determinante der 0×0-Matrix ist 1. Für $n-m > 0$ gilt nach Induktionsannahme $\det(A(m+1,\, j)) = \det(B) \cdot \det(D(1,\, j-m))$ für alle $j \in \{m+1, ..., n\}$. Laplace-Entwicklung nach Zeile $m+1$ liefert

$$\begin{aligned} \det(A) &= \sum_{j=1}^{n} (-1)^{m+1+j} A_{m+1,j} \det(A(m+1, j)) \\ &\overset{j=k+m}{=} \sum_{k=1}^{n-m} (-1)^{m+1+m+k} D_{1,k} \det(B) \det(D(1,k)) = \det(B) \det(D) \end{aligned}$$

b) Folgt per Induktion aus a). □

Beispiel 5.9

Es gibt meist mehrere Rechenwege; effizientes Rechnen erfordert Erfahrung.

$$\begin{aligned} \begin{vmatrix} 1&2&3&0&4 \\ 5&6&7&0&8 \\ \pi&e&13&3&-1 \\ 2&6&4&0&8 \\ 3&1&1&0&2 \end{vmatrix} &= -3 \cdot \begin{vmatrix} 1&2&3&4 \\ 5&6&7&8 \\ 2&6&4&8 \\ 3&1&1&2 \end{vmatrix} && \textit{Laplace 4. Spalte} \\ &= -3 \cdot 2 \cdot \begin{vmatrix} 1&2&3&4 \\ 5&6&7&8 \\ 1&3&2&4 \\ 3&1&1&2 \end{vmatrix} && \textit{Skalierung 3. Zeile} \\ &= -6 \cdot \begin{vmatrix} 1&2&3&4 \\ 0&-4&-8&-12 \\ 0&1&-1&0 \\ 0&-5&-8&-10 \end{vmatrix} && \textit{Gauß-Elimination} \\ &= 6 \cdot \begin{vmatrix} 1&2&3&4 \\ 0&1&-1&0 \\ 0&-4&-8&-12 \\ 0&-5&-8&-10 \end{vmatrix} && \textit{Zeilentausch} \end{aligned}$$

$$= 6 \cdot \begin{vmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & -12 & -12 \\ 0 & 0 & -13 & -10 \end{vmatrix} \qquad \textit{Gauß-Elimination}$$

$$= 6 \cdot 1 \cdot 1 \cdot \begin{vmatrix} -12 & -12 \\ -13 & -10 \end{vmatrix} \qquad \textit{Block-/Dreiecksgestalt}$$

$$= 6 \cdot (12 \cdot 10 - 12 \cdot 13) = 6 \cdot 12 \cdot (-3) = -216 \qquad \textit{Sarrus-Regel } (2 \times 2)$$

Produktregel
$\forall A, B \in M_n(\mathbb{K})\colon\ \det(A \cdot B) = \det(A) \cdot \det(B)$.

Beweis:
Für $A = (\vec{a}_1, ..., \vec{a}_n) \in M_n(\mathbb{K})$, $B =\in M_n(\mathbb{K})$ ist Spalte j von $A \cdot B$ gleich dem Produkt von A und der j-ten Spalte von B, und dies ist die Linearkombination der Spalten von A mit Koeffizienten $B_{1,j}, ..., B_{n,j}$, das heißt $\sum\limits_{i_j=1}^{n} \vec{a}_{i_j} B_{i_j,j}$.
Multilinear, alternierend $\Rightarrow \det(A \cdot B) = \sum\limits_{i_1=1}^{n} ... \sum\limits_{i_n=1}^{n} \det(\vec{a}_{i_1}, ..., \vec{a}_{i_n}) \cdot \prod\limits_{j=1}^{n} B_{i_j,j} =$
$\sum\limits_{\sigma \in S_n} \det(A) \operatorname{sgn}(\sigma) \cdot \prod\limits_{j=1}^{n} B_{\sigma(j),j} = \det(A) \cdot \det(B)$ □

Definition 5.10. *Die* **Adjunkte**[28] *von* $A \in M_n(\mathbb{K})$ *ist* $\operatorname{adj}(A) \in M_n(\mathbb{K})$ *mit* $\operatorname{adj}(A)_{j,i} = (-1)^{i+j} \det(A(i,j))$.

Lemma 5.11
$\forall A \in M_n(\mathbb{K})\colon\ A \cdot \operatorname{adj}(A) = \operatorname{adj}(A) \cdot A = \det(A) \cdot \mathbb{1}_n$.

Beweis:

- Sei $B := A \cdot \operatorname{adj}(A)$. Dann $B_{i,k} = \sum\limits_{j=1}^{n} A_{i,j} \cdot (-1)^{k+j} \det(A(k,j))$. Daraus folgt zunächst $B_{i,i} = \det(A)$ wegen der Laplace-Entwicklung nach Zeile i. Ist $i \neq k$, so entstehe $\tilde{A}$ aus A, indem man Zeile k durch eine Kopie von Zeile i ersetzt. Dann ist $B_{i,k} = \sum\limits_{j=1}^{n} \tilde{A}_{k,j} \cdot (-1)^{k+j} \det(\tilde{A}(k,j)) = \det(\tilde{A}) = 0$, wieder wegen Laplace-Entwicklung.

- Dass auch $\operatorname{adj}(A) \cdot A = \det(A) \cdot \mathbb{1}_n$ gilt, kann man mit Spaltenentwicklung nachrechnen. □

Korollar 5.12
$A \in M_n(\mathbb{K})$ ist invertierbar (d.h. es gibt $A^{-1} \in M_n(\mathbb{K})$ mit $A^{-1} \cdot A = A \cdot A^{-1} = \mathbb{1}_n$) genau dann, wenn $\det(A)$ *invertierbar in* $\mathbb{K}$ *ist. Dann gilt* $\det(A^{-1}) = (\det(A))^{-1}$.
$GL_n(\mathbb{K}) := \{A \in M_n(\mathbb{K}) \mid \det(A)$ invertierbar in $\mathbb{K}\}$ ist eine Gruppe bzgl. Matrixmultiplikation.

[28] Bitte nicht „Adjungierte“, das wäre etwas anderes.

Beweis:
Ist A invertierbar, dann wegen Produktregel $\det(A) \cdot \det(A^{-1}) = \det(A \cdot A^{-1}) = \det(\mathbb{1}_n) = 1 = \det(A^{-1} \cdot A) = \det(A^{-1}) \cdot \det(A)$, also $\det(A)$ invertierbar in $\mathbb{K}$ und zudem $(\det(A))^{-1} = \det(A^{-1})$.
Ist $\det(A)$ in $\mathbb{K}$ invertierbar, dann $A^{-1} = (\det(A))^{-1} \cdot \operatorname{adj}(A)$ nach Lemma 5.11.

Es bleibt zu zeigen: $A, B \in GL_n(\mathbb{K}) \Rightarrow AB \in GL_n(\mathbb{K})$. Wenn aber $\det(A)$ und $\det(B)$ in $\mathbb{K}$ invertierbar sind, dann ist auch $\det(AB) = \det(A)\det(B)$ invertierbar, mit Inversem $\det(B)^{-1}\det(A)^{-1}$. □

Beispiel In den Hausaufgaben sollten Sie eine Formel für das Inverse einer Matrix $A := \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in M_2(\mathbb{K})$ für den Fall finden, dass $\det(A) = ad - bc \neq 0$. Sie fanden $A^{-1} = \frac{1}{\det(A)} \left(\begin{smallmatrix} d & -b \\ -c & a \end{smallmatrix}\right)$. Das stimmt mit der obigen Formel für $\operatorname{adj}(A)$ überein, denn: $\operatorname{adj}(A)_{1,1} = (-1)^2 \det(A(1,1)) = d$, $\operatorname{adj}(A)_{1,2} = (-1)^3 \det(A(2,1)) = -b$, $\operatorname{adj}(A)_{2,1} = (-1)^3 \det(A(1,2)) = -c$, $\operatorname{adj}(A)_{2,2} = (-1)^4 \det(A(2,2)) = a$.

Zur Illustration leite ich eine explizite Formel für die eindeutige Lösung des Gleichungssystems $A \cdot \vec{x} = \vec{b}$ mit einer invertierbaren Matrix A her. Diese ist nach Gabriel Cramer [1704–1752] benannt, der sie 1750 beschrieb. Gottfried Wilhelm Leibniz [1646–1716] war sie sogar schon 1678 bekannt, aber einen Beweis der Formel lieferte erst 1815 Augustin Louis Cauchy [1789–1857].

Cramersche Regel
Sei $A \in GL_n(\mathbb{K})$ und $\vec{b} \in \mathbb{K}^n$. Für $i \in \{1, ..., n\}$ sei $A_i \in M_n(\mathbb{K})$ die Matrix, die durch Ersetzung der Spalte i von A durch $\vec{b}$ entsteht. Für die eindeutige Lösung $\vec{x} \in \mathbb{K}^n$ von $A \cdot \vec{x} = \vec{b}$ gilt $\forall i \in \{1, ..., n\}$: $x_i := \frac{\det(A_i)}{\det(A)}$.

Beispiel Für $A := \left(\begin{smallmatrix} 1 & 2 \\ 4 & 5 \end{smallmatrix}\right)$ und $\vec{b} = \left(\begin{smallmatrix} 3 \\ 6 \end{smallmatrix}\right)$ hat $A \cdot \left(\begin{smallmatrix} x_1 \\ x_2 \end{smallmatrix}\right) = \vec{b}$ die eindeutige Lösung

$$x_1 = \frac{\left|\begin{smallmatrix} 3 & 2 \\ 6 & 5 \end{smallmatrix}\right|}{\left|\begin{smallmatrix} 1 & 2 \\ 4 & 5 \end{smallmatrix}\right|} = \frac{3}{-3} = -1$$
$$x_2 = \frac{\left|\begin{smallmatrix} 1 & 3 \\ 4 & 6 \end{smallmatrix}\right|}{\left|\begin{smallmatrix} 1 & 2 \\ 4 & 5 \end{smallmatrix}\right|} = \frac{-6}{-3} = 2$$

Beweis der Cramerschen Regel:
Seien $\vec{a}_1, ..., \vec{a}_n \in \mathbb{K}^n$ die Spalten von A. Aus der eindeutigen Lösung des Gleichungssystems ergibt sich $x_1 \cdot \vec{a}_1 + ... + 1 \cdot (x_i \vec{a}_i - \vec{b}) + ... + x_n \cdot \vec{a}_n = \vec{0}$, d.h. die Vektoren auf der linken Seite sind linear abhängig. Also ist $C := (\vec{a}_1, ..., x_i\vec{a}_i - \vec{b}, ..., \vec{a}_n)$ nicht invertierbar, und wegen Korollar 5.12 und Multilinearität folgt

$$0 = \det(C) = x_i \det(A) - \det(\vec{a}_1, ..., \vec{a}_{i-1}, \vec{b}, \vec{a}_{i+1}, ..., \vec{a}_n). \qquad \square$$

Ich empfehle die Verwendung der Cramerschen Regel ausdrücklich *nicht*. Erstens nämlich nutzt sie nichts, wenn A nicht invertierbar ist. Und vor Allem ist der Rechenaufwand unsinnig groß.

6 Eigenwertprobleme

In diesem Abschnitt sei $\mathbb{K}$ wieder ein Körper. In diesem Kapitel wird untersucht, ob man zu linearen Selbstabbildungen eine Basis B finden kann, so dass die Darstellungsmatrix eine Diagonalmatrix ist.

6.1 Eigenwerte, -vektoren und -räume

Definition 6.1
Sei V ein $\mathbb{K}$-Vektorraum. Eine lineare Abbildung $\varphi\colon V \to V$ heißt **Endomorphismus** *von V. Sei im Folgenden φ ein Endomorphismus von V.*

a) Sei $\lambda \in \mathbb{K}$. $E_\lambda(\varphi) = \{\vec{v} \in V \mid \varphi(\vec{v}) = \lambda\vec{v}\}$ ist ein **Eigenraum** *von φ.*

b) $\vec{v} \in V$ heißt **Eigenvektor** *zum* **Eigenwert** *$\lambda \in \mathbb{K}$ von $\varphi :\Leftrightarrow \vec{v} \in E_\lambda(\varphi)\backslash\{\vec{0}\}$.*

c) Ist $A \in M_n(\mathbb{K})$, so ist $L_A\colon \mathbb{K}^n \to \mathbb{K}^n$ mit $L_A(\vec{v}) := A \cdot \vec{v}$ ein Endomorphismus. Für $\lambda \in \mathbb{K}$ sei $E_\lambda(A) := E_\lambda(L_A)$ Eigenraum von A. Eigenvektoren bzw. Eigenwerte von A sind gleich denen von L_A.

Der Nullvektor ist niemals ein Eigenvektor!!
Zwar ist stets $\vec{0} \in E_\lambda(A)$, aber λ ist nur Eigenwert, falls $E_\lambda(A) \neq \{\vec{0}\}$.

Beispiel 6.2

a) Zu den Eigenvektoren von $A = \left(\begin{smallmatrix} 2 & -1 \\ 3 & -2 \end{smallmatrix}\right)$ gehören unter anderen $\left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right)$ mit Eigenwert 1 und $\left(\begin{smallmatrix} 1 \\ 3 \end{smallmatrix}\right)$ mit Eigenwert -1, denn

$$\left(\begin{smallmatrix} 2 & -1 \\ 3 & -2 \end{smallmatrix}\right)\left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right) \qquad \left(\begin{smallmatrix} 2 & -1 \\ 3 & -2 \end{smallmatrix}\right)\left(\begin{smallmatrix} 1 \\ 3 \end{smallmatrix}\right) = \left(\begin{smallmatrix} -1 \\ -3 \end{smallmatrix}\right)$$

Dagegen ist $\left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)$ kein Eigenvektor von A, denn $\left(\begin{smallmatrix} 2 & -1 \\ 3 & -2 \end{smallmatrix}\right)\left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 2 \\ 3 \end{smallmatrix}\right)$ ist kein Skalarvielfaches von $\left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)$.

b) Wegen

$$\begin{pmatrix} 1 & 2 & 2 \\ 0 & -1 & 1 \\ -1 & 1 & -5 \end{pmatrix}\begin{pmatrix} 4 \\ -1 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

ist $\begin{pmatrix} 4 \\ -1 \\ -1 \end{pmatrix}$ ein Eigenvektor von $\begin{pmatrix} 1 & 2 & 2 \\ 0 & -1 & 1 \\ -1 & 1 & -5 \end{pmatrix}$ mit Eigenwert 0.

c) $A = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ hat keine reellen Eigenwerte, denn aus $A\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right) = \lambda\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right)$ folgen $-y = \lambda x$ und $x = \lambda y$, also $(\lambda^2 + 1)x = 0$, also $x = 0$ und $y = 0$. Erlaubt man aber komplexe Vektoren, so ist $\left(\begin{smallmatrix} \mathrm{i} \\ 1 \end{smallmatrix}\right)$ Eigenvektor zum Eigenwert i.

d) Hier ist ein Beispiel für Endomorphismen, das über die Begrifflichkeiten für Matrizen hinausgeht: Sei $V = \mathscr{C}^\infty(\mathbb{R}, \mathbb{R})$ und $\varphi\colon V \to V$ mit $\varphi(f) := f''$ (zweite Ableitung). Wegen $\sin''(x) = -\sin(x)$ und $\cos''(x) = -\cos(x)$ sind die Sinus- und Kosinusfunktionen dann Eigenvektoren von φ mit Eigenwert -1, wärend $\exp(x)$ ein Eigenvektor mit Eigenwert 1 ist. In diesem speziellen Beispiel gilt sogar, dass jede reelle Zahl Eigenwert von φ ist.

Die Suche nach Eigenwerten und Eigenvektoren hat zahlreiche Anwendungen; historisch gesehen wurden sogar *unendlichdimensionale* Eigenwertprobleme wie im obigen Beispiel d) zuerst betrachtet.

Eine kleine Auswahl von Anwendungen

- ***Hauptkomponentenanalyse*** zur Analyse experimenteller statistischer Daten: Aus den Daten wird zunächst die *Kovarianzmatrix* berechnet. An den Eigenvektoren und den zugehörigen Eigenwerten erkennt man, durch welche Einflussgrößen die Daten bestimmt werden (Hauptkomponenten). Kleine Eigenwerte der Kovarianzmatrix entsprechen statistischem Rauschen. Das lässt sich auch in der Datenkompression nutzen.

- Resonanzfrequenzen berechnet man durch Eigenwertprobleme.

- Die Eigenvektoren des Trägheitstensors sind die Rotationsachsen, bezüglich denen ein Körper keine Unwucht hat.

- Quantenmechanischen Systeme werden mit dem ***Hamilton-Operator*** beschrieben. Dessen Eigenwerte sind die möglichen Energiezustände und lassen sich im Spektrum beobachten. Manchmal erlauben Zusatzannahmen, die Quantenzustände angenähert als *endlich*dimensionale Eigenwertprobleme zu modellieren.

- Anfänglich verwendete Google bei der Bewertung von Suchergebnissen den ***PageRank-Algorithmus***. Webseiten und ihre gegenseitige Verlinkung werden in einer gigantischen Matrix kodiert. Man berechnet den Eigenvektor zum größten Eigenwert dieser Matrix. Die Webseite, die zum größten Koeffizienten dieses Eigenvektors gehört, ist das „beste“ Suchergebnis.

Mit den hier behandelten Methoden kann man kleinere Beispiele per Hand berechnen. Für große oder gar unendlichdimensionale Eigenwertprobleme benötigt man Methoden der Numerik oder der Analysis.

Definition 6.3
$\chi_A(X) := \det(X\mathbb{1}_n - A) \in \mathbb{K}[X]$ *heißt* **charakteristisches Polynom** *von* $A \in M_n(\mathbb{K})$.

Bemerkung 6.4
In der Literatur wird auch $\chi_A(X) = \det(A - X\mathbb{1}_n)$ *definiert. Die Definitionen unterscheiden sich um das Vorzeichen* $(-1)^n$. *Mit der hier gegebenen Definition ist* $\chi_A(X)$ **normiert**, *d.h. der Koeffizient des führenden Terms ist* 1.

Lemma 6.5
Sei $A \in M_n(\mathbb{K})$ *und* $\lambda \in \mathbb{K}$.

a) $E_\lambda(A) = \mathrm{LR}(\lambda\mathbb{1}_n - A;\ \vec{0}) \leq \mathbb{K}^n$.

b) λ *ist Eigenwert von* $A \iff \chi_A(\lambda) = 0$ *(***Säkulargleichung***).*

Beweis:
a): $A\vec{v} = \lambda\vec{v} = \lambda\mathbb{1}_n\vec{v} \iff \vec{0} = (\lambda\mathbb{1}_n - A)\vec{v} \iff \vec{v} \in \mathrm{LR}(\lambda\mathbb{1}_n - A;\ \vec{0})$.
b): λ E.-Wert $\overset{a)}{\iff} \dim \mathrm{LR}(\lambda\mathbb{1}_n - A; \vec{0}) > 0 \iff \mathrm{Rang}(\lambda\mathbb{1}_n - A) < n \iff \det(\lambda\mathbb{1}_n - A) = 0$. □

Definition 6.6
Die **Spur** *von* $A \in M_n(\mathbb{K})$ *ist* $\mathrm{Spur}(A) := \sum_{i=1}^n A_{i,i}$.

Beispiel $\mathrm{Spur}\left(\begin{smallmatrix} 2 & -1 \\ 3 & -2 \end{smallmatrix}\right) = 2 + (-2) = 0$.

Lemma 6.7
Für $A \in M_n(\mathbb{K})$ *hat* $\chi_A(X)$ *die Gestalt*

$$\chi_A(X) = X^n - \mathrm{Spur}(A)X^{n-1} + (\textit{Terme vom Grad } n-2 \geq r \geq 1) + (-1)^n \det(A)\,.$$

Beweis:
Sei $B = X\mathbb{1}_n - A$. Variable X kommt in jeder Spalte von B genau einmal vor. Nach der Leibnizformel ist $\det(B)$ eine Summe von Produkten von Matrixeinträgen (mit Vorzeichen), wobei jeder Summand genau einen Eintrag pro Zeile und genau einen pro Spalte enthält. Daher ist $\det(B)$ ein Polynom in X vom Grad $\leq n$. Setzt man $X = 0$, dann ist $\det(B) = \det(-A) = (-1)^n \det(A)$, woraus sich der konstante Term ergibt.

Enthält ein Produkt einen Eintrag außerhalb der Diagonale, dann auch noch einen weiteren daher enthält es X nur zur Potenz $\leq n-2$. Die Koeffizienten von X^n und X^{n-1} ergeben sich also einzig aus dem Produkt der Diagonalelemente:
$\prod_{i=1}^n B_{i,i} = \prod_{i=1}^n (X - A_{i,i}) = X^n - X^{n-1} \sum_{j=1}^n A_{j,j} + \text{Terme kleineren Grades.}$ □

Beispiel 6.8

a) Für $A \in M_2(\mathbb{K})$ *ergibt sich aus dem Lemma* $\chi_A(X) = X^2 - \mathrm{Spur}(A) \cdot X + \det(A)$ — ***eine nützliche Formel, die man sich merken sollte!***

Für $A = \left(\begin{smallmatrix} 2 & -1 \\ 3 & -2 \end{smallmatrix}\right)$ *ist* $\chi_A(X) = X^2 - 4 + 3 = X^2 - 1 = (X-1)(X+1)$. *Also sind* 1 *und* -1 *die einzigen Eigenwerte. Berechnung der Eigenvektoren:*

$\boxed{\lambda = 1}$ $B := \mathbb{1}_2 - A = \left(\begin{smallmatrix} -1 & 1 \\ -3 & 3 \end{smallmatrix}\right)$, $E_1(A) = \mathrm{LR}(B;\ \vec{0}) = \mathrm{Span}(\left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right))$.

$\boxed{\lambda = -1}$ $B := -\mathbb{1}_2 - A = \left(\begin{smallmatrix} -3 & 1 \\ -3 & 1 \end{smallmatrix}\right)$, $E_{-1}(A) = \mathrm{LR}(B;\ \vec{0}) = \mathrm{Span}(\left(\begin{smallmatrix} 1/3 \\ 1 \end{smallmatrix}\right))$.

b) Für $A = \left(\begin{smallmatrix} 0 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & -1 & 0 \end{smallmatrix}\right) \in M_3(\mathbb{R})$ *ist*

$$\chi_A(X) = \left|\begin{smallmatrix} X & -1 & 0 \\ 1 & X & -1 \\ 0 & 1 & X \end{smallmatrix}\right| = X \left|\begin{smallmatrix} X & -1 \\ 1 & X \end{smallmatrix}\right| + \left|\begin{smallmatrix} 1 & -1 \\ 0 & X \end{smallmatrix}\right| = X(X^2+1) + X = X(X^2+2)$$

Der einzige reelle Eigenwert ist 0*, mit* $E_0(A) = \mathrm{LR}(A;\ \vec{0}) = \mathrm{Span}(\begin{pmatrix}1\\0\\1\end{pmatrix})$. *In* $\mathbb{C}$ *gibt es noch die Eigenwerte* $\pm\sqrt{2}\,\mathrm{i}$.

c) Bei der Berechnung von $\det(X\mathbb{1} - A)$ *sollte man kreativ sein und Brüche von Polynomen möglichst vermeiden. Sei* $A = \begin{pmatrix}-3&-6&22&-4\\-6&-7&28&-6\\-2&-2&7&-2\\6&10&-38&7\end{pmatrix}$. *Dann ist*

$$\begin{aligned}
\chi_A(X) &= \begin{vmatrix} X+3 & 6 & -22 & 4 \\ 6 & X+7 & -28 & 6 \\ 2 & 2 & X-7 & 2 \\ -6 & -10 & 38 & X-7 \end{vmatrix} = \begin{vmatrix} X+3 & 3-X & -22 & 1-X \\ 6 & X+1 & -28 & 0 \\ 2 & 0 & X-7 & 0 \\ -6 & -4 & 38 & X-1 \end{vmatrix} \\
&= \begin{vmatrix} X+3 & -X+3 & -22 & -X+1 \\ 6 & X+1 & -28 & 0 \\ 2 & 0 & X-7 & 0 \\ X-3 & -X-1 & 16 & 0 \end{vmatrix} = (X-1)\cdot \begin{vmatrix} 6 & X+1 & -28 \\ 2 & 0 & X-7 \\ X+3 & 0 & -12 \end{vmatrix} \\
&= -(X-1)\cdot(X+1)\cdot \begin{vmatrix} 2 & X-7 \\ X+3 & -12 \end{vmatrix} \\
&= -(X-1)\cdot(X+1)\cdot(-X^2+4X-3) \\
&= (X-1)\cdot(X+1)\cdot(X-1)\cdot(X-3)
\end{aligned}$$

Die Eigenwerte von A *sind also* $1, -1, 3$. *Für jeden Eigenwert muss man nun noch eine Basis für den Eigenraum berechnen. Ergebnisse:*

$$\begin{aligned}
E_1(A) &= \mathrm{LR}(\mathbb{1}_4 - A;\ \vec{0}) = \mathrm{Span}(\begin{pmatrix}-2\\5\\1\\0\end{pmatrix}, \begin{pmatrix}-1\\0\\0\\1\end{pmatrix}) \\
E_{-1}(A) &= \mathrm{LR}(-\mathbb{1}_4 - A;\ \vec{0}) = \mathrm{Span}(\begin{pmatrix}-1\\-1\\0\\2\end{pmatrix}) \\
E_3 &= \mathrm{LR}(3\cdot\mathbb{1}_4 - A;\ \vec{0}) = \mathrm{Span}(\begin{pmatrix}-3\\-4\\-1\\5\end{pmatrix})
\end{aligned}$$

Rechnung für $E_1(A)$: $\mathbb{1}_4 - A = \begin{pmatrix}4&6&-22&4\\6&8&-28&6\\2&2&-6&2\\-6&-10&38&-6\end{pmatrix} \rightsquigarrow \begin{pmatrix}4&6&-22&4\\0&-1&5&0\\0&-1&5&0\\0&-1&5&0\end{pmatrix} \rightsquigarrow \begin{pmatrix}4&6&-22&4\\0&-1&5&0\\0&0&0&0\\0&0&0&0\end{pmatrix}$, *und* $E_1(A) = \mathrm{LR}(\mathbb{1}_4 - A;\ \vec{0})$ *wird aufgespannt von den beiden Basislösungen* $\vec{\beta}_3 = \begin{pmatrix}-2\\5\\1\\0\end{pmatrix}$ *und* $\vec{\beta}_4 = \begin{pmatrix}-1\\0\\0\\1\end{pmatrix}$.

Guter Rat: Behalten Sie die während der Berechnung von $\chi_A(X)$ gefundenen Faktoren bei. Also *NICHT* ausmultiplizieren, wenn es sich vermeiden lässt! Grund: Die Eigenwerte kann man an den Faktoren leicht ablesen.

Der Hintergrund dieses Ratschlags ist, dass zwar über den komplexen Zahlen jedes Polynom in Linearfaktoren zerfällt, dass es aber im Allgemeinen beweisbar unmöglich ist, diese Linearfaktoren durch eine Lösungsformel zu berechnen. Seien Sie also froh, wenn Sie eine Nullstelle gefunden haben!

Hauptsatz der Algebra
$\forall p \in \mathbb{C}[X]$ *mit* $\deg(p) = n$ *und führendem Term* X^n: $\exists \lambda_1, ..., \lambda_n \in \mathbb{C}$: $p(X) = (X-\lambda_1)\cdot ... \cdot (X-\lambda_n)$.

Definition 6.9
Sei $A \in M_n(\mathbb{K})$ und $\lambda \in \mathbb{K}$ ein Eigenwert von A.

a) λ hat die **algebraische Vielfachheit** *$k \in \mathbb{N} :\Leftrightarrow \chi_A(X)$ wird von $(X-\lambda)^k$ ohne Rest geteilt, aber nicht von $(X-\lambda)^{k+1}$,*

b) $\dim(E_\lambda(A))$ heißt die **geometrische Vielfachheit** *von λ.*

In der Schule lernten Sie Lösungsformeln für quadratische Gleichungen, jedoch nicht die aus dem 16. Jhdt. stammenden ***Formeln von Cardano***[29] zur Lösung algebraischer Gleichungen vom Grad 3 und 4. Jedoch gibt es darüber hinaus keine allgemeingültigen Lösungsformeln:

Satz von Abel-Ruffini
Die Nullstellen eines Polynoms vom Grad ≥ 5 lassen sich im Allgemeinen nicht *durch Grundrechenarten und Wurzelziehen berechnen. Dies gilt zum Beispiel für die Nullstelle $x = -1.1673...$ von $x^5 - x + 1$.*

Paolo Ruffini [1765–1822] hatte dafür 1799 einen lückenhaften Beweis. Niels Henrik Abel [1802–1829] gelang 1824 der erste vollständige Beweis. Heute wird der Satz (z.B. in Algebra–Vorlesungen) im Rahmen der *Galois-Theorie*[30] bewiesen.

Berechnen Sie Eigenwerte immer exakt! Wäre λ nur *näherungsweise* Eigenwert von $A \in M_n(\mathbb{K})$, so wäre $E_\lambda(A) = \{\vec{0}\}$: Man fände keinen Eigenvektor, auch nicht näherungsweise! In numerischen Lösungen des Eigenwertproblems werden daher die Eigenwerte und die Eigenvektoren *gleichzeitig* approximiert.

Folgende Tipps zur Nullstellenberechnung von Polynomen kennen Sie wahrscheinlich aus der Schule (ggf. mit anderen Begriffsbildungen).

Praktische Tipps zur Nullstellensuche
Sei $f \in \mathbb{R}[X]$.

a) Angenommen, die Koeffizienten von f sind ganze Zahlen. Wenn f eine Nullstelle $\lambda \in \mathbb{Z}$ besitzt, so teilt λ das Absolutglied von f. Auf diese Weise kann man ganzzahlige Nullstellen raten.

b) Kennt man eine Nullstelle λ von f, so sind die restlichen Nullstellen von f genau die Nullstellen von $\frac{f}{X-\lambda}$; die Polynomdivision geht ohne Rest auf (oder man hat einen Fehler gemacht).

[29] Nach Gerolamo Cardano [1501–1576]. Den komplizierten Prioritätenstreit mit Scipione del Ferro [1465–1526], Niccolò Fontana Tartaglia [1500–1557], und Lodovico Ferrari [1522–1565] übergehe ich hier.

[30] Das gesamte Werk von Évariste Galois [1811–1832] umfasst nur rund 60 Seiten, doch er begründet darin unabhängig von Abel die Gruppentheorie, gibt mit der „Galois-Theorie“ notwendige und hinreichende Bedingungen dafür, welche algebraischen Gleichungen eine Lösung durch Grundrechenarten und Wurzelziehen besitzen, und konstruiert alle endlichen Körper.

c) Kennt man bereits eine Faktorisierung eines Polynoms, sollte man sie sich merken! Sind $g_1, g_2 \in \mathbb{R}[X]$ vom Grad 2, so findet man die Nullstellen von $f = g_1 g_2$ durch Anwendung der p,q-Formel auf g_1 und g_2.

Bemerkung 6.10
Gleichheit von Polynomen ist <u>nicht</u> durch die Gleichheit aller Auswertungen definiert! *Zwei Polynome $p_1, p_2 \in \mathbb{K}[X]$ sind genau dann gleich, wenn alle ihre Koeffizienten gleich sind. Ist beispielsweise $\mathbb{K}$ endlich, dann ist $p := \prod_{a \in \mathbb{K}} (X - a) \in \mathbb{K}[X]$, $p \neq 0$, aber $\forall b \in \mathbb{K}\colon\ p(b) = 0$.*

Wenn aber $\mathbb{K}$ unendlich ist, so kann man zeigen, dass $\forall p_1, p_2 \in \mathbb{K}[X]\colon\ p_1 = p_2 \iff \forall x \in \mathbb{K}\colon\ p_1(x) = p_2(x)$.

6.2 Eigenräume sind komplementär

Lemma 6.11
Seien $\lambda_1, \ldots, \lambda_r$ paarweise verschiedene Eigenwerte von $A \in M_n(\mathbb{K})$.

a) Sind $\vec{u}_i \in E_{\lambda_i}(A) \setminus \{\vec{0}\}$ für $i = 1, ..., r$, dann ist $[\vec{u}_1, ..., \vec{u}_r]$ linear unabhängig.

b) Die Vereinigung von Basen von $E_{\lambda_1}(A), \ldots, E_{\lambda_r}(A)$ ist linear unabhängig. Insbesondere ist $\sum_{i=1}^{r} \dim E_{\lambda_i}(A) \leq n$.

Beweis:
b) folgt aus a).
a): Sei $\sum_{i=1}^{r} \alpha_i \vec{u}_i = \vec{0}$ mit $\alpha_1, ..., \alpha_r \in \mathbb{K}$. Zu zeigen: $\alpha_1 = ... = \alpha_r = 0$.
Induktion über r. Klar für $r = 1$. Ist $r \geq 2$, dann

$$\vec{0} = A\left(\sum_{i=1}^{r} \alpha_i \vec{u}_i\right) - \lambda_r \sum_{i=1}^{r} \alpha_i \vec{u}_i = \sum_{i=1}^{r} (\lambda_i - \lambda_r)\alpha_i \vec{u}_i = \sum_{i=1}^{r-1} (\lambda_i - \lambda_r)\alpha_i \vec{u}_i\,.$$

Also (Induktionsannahme) $(\lambda_i - \lambda_r)\alpha_i = 0$ für alle $i < r$. Wegen $\lambda_i \neq \lambda_r$ für $i < r$ folgt daraus $\alpha_i = 0$ für alle $i < r$. Dann folgt auch $\vec{0} = \sum_{i=1}^{r} \alpha_i \vec{u}_i = \alpha_r \vec{u}_r$ und damit $\alpha_r \neq 0$ wegen $\vec{u}_r \neq \vec{0}$. □

Beispiel $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$ hat in $\vec{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ einen Eigenvektor mit Eigenwert 3. Ferner ist $\mathrm{Rang}(A) = 1$, also ist $E_0(A) = \mathrm{LR}(A; \vec{0})$ 2-dimensional: Sei $[\vec{v}_2, \vec{v}_3]$ eine Basis von $E_0(A)$. Wende das Lemma auf den Eigenwerten $3, 0$ an: $[\vec{v}_1, \vec{v}_2, \vec{v}_3]$ ist linear unabhängig, daher eine Basis des $\mathbb{R}^3$. Also keine weitere Eigenwerte, sonst erhielte man nach dem Lemma ein linear unabhängiges System der Länge ≥ 4.

6.3 Basiswechsel

Sind B, C Basen eines endlichdimensionalen $\mathbb{K}$-Vektorraums V, so haben wir gemäß Beobachtung 4.6 Basiswechselmatrizen ${}^C_B\mathbb{T}$ und ${}^B_C\mathbb{T} = \left({}^C_B\mathbb{T}\right)^{-1}$.

Lemma 6.12
Sei $A \in M_n(\mathbb{K})$. Sei $F\colon \mathbb{K}^n \to \mathbb{K}^n$ der Endomorphismus $F = L_A$. Für eine Matrix $A' \in M_n(\mathbb{K})$ sind dann die folgenden beiden Aussagen äquivalent:

a) Es gibt $S \in GL_n(\mathbb{K})$ mit $A' = S^{-1}AS$.

b) Es gibt eine Basis B des $\mathbb{K}^n$ derart, dass $A' = {}^B_B F$ ist.

Konkret ist der Zusammenhang zwischen S und B wie folgt: $S = {}^E_B\mathbb{T}$, wobei E die Standardbasis ist; B besteht aus den Spalten von S.

Beweis:
$F = L_A$ heißt $A = {}^E_E F$, für E die Standardbasis. Sei $A' = S^{-1}AS$. $\operatorname{Rang}(S) = n$, also bilden die Spalten von S eine Basis B von $\mathbb{K}^n$. Dann ist $S = {}^E_B\mathrm{Id} = {}^E_B\mathbb{T}$ und $S^{-1} = {}^B_E\mathbb{T}$, also $A' = {}^B_E\mathbb{T}\ {}^E_E F\ {}^E_B\mathbb{T} = {}^B_B F$. Ist dagegen $A' = {}^B_B F$, dann ist mit $S = {}^E_B\mathbb{T}$ auch $A' = S^{-1}\,{}^E_E F\,S = S^{-1}AS$. □

Beispiel $A = \left(\begin{smallmatrix} 2 & -1 \\ 3 & -2 \end{smallmatrix}\right)$ hat Eigenvektoren $\vec{v}_1 = \left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right)$ und $\vec{v}_2 = \left(\begin{smallmatrix} 1 \\ 3 \end{smallmatrix}\right)$ mit Eigenwert 1 bzw. -1. Nun, $B := [\vec{v}_1, \vec{v}_2]$ ist eine Basis von $\mathbb{R}^2$. Wegen $F(\vec{v}_1) = A \cdot \vec{v}_1 = 1 \cdot \vec{v}_1$ und $F(\vec{v}_2) = (-1)\vec{v}_2$ ist ${}^B_B F = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$. Ferner ist ${}^E_B\mathbb{T} = \left(\begin{smallmatrix} 1 & 1 \\ 1 & 3 \end{smallmatrix}\right)$. Nach dem Lemma gilt $\left(\begin{smallmatrix} 1 & 1 \\ 1 & 3 \end{smallmatrix}\right)^{-1} \left(\begin{smallmatrix} 2 & -1 \\ 3 & -2 \end{smallmatrix}\right) \left(\begin{smallmatrix} 1 & 1 \\ 1 & 3 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$. Wir haben A *diagonalisiert*. Zur Kontrolle *kann* man die Gleichung direkt prüfen.

Aufgabe 6.13
Für $A' = S^{-1}AS$ gilt $\chi_{A'}(X) = \chi_A(X)$.

Korollar 6.14
Ist $\dim E_\lambda(A) = r$, dann ist $\chi_A(X)$ durch $(X - \lambda)^r$ teilbar. Die geometrische Vielfachheit eines Eigenwerts ist also höchstens so groß wie die algebraische Vielfachheit.

Beweis:
Sei $[\vec{v}_1, \ldots, \vec{v}_r]$ eine Basis von $E_\lambda(A)$. Basisergänzungssatz: Setze zu einer Basis $B = [\vec{v}_1, \ldots, \vec{v}_n]$ von $\mathbb{K}^n$ fort. Wie in Lemma 6.12 sei $F = L_A$, $S = {}^E_B\mathbb{T}$ und $A' = {}^B_B F = S^{-1}AS$. Nach Aufgabe 6.13 ist $\chi_{A'} = \chi_A$. Wegen $F(\vec{v}_i) = \lambda\vec{v}_i$ für $i \le r$ hat A' Blockgestalt $\left(\begin{smallmatrix} \lambda\mathbb{1}_r & C \\ 0 & D \end{smallmatrix}\right)$, also $X\mathbb{1}_n - A' = \left(\begin{smallmatrix} (X-\lambda)\mathbb{1}_r & -C \\ 0 & X\mathbb{1}_s - D \end{smallmatrix}\right)$ für $s = n - r$, und nach der Blockmatrix-Regel ist $\chi_{A'}(X) = (X - \lambda)^r \chi_D(X)$. □

Beispiel Sei $A = \left(\begin{smallmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{smallmatrix}\right)$. $X^2 \cdot (X - 3)$ teilt $\chi_A(X)$, denn $\dim E_0(A) = 2$ und $\dim E_3(A) \ge 1$. $\chi_A(X)$ normiert vom Grad 3 $\Rightarrow$ $\chi_A(X) = X^2(X - 3)$.

6.4 Diagonalisierung

Lemma 6.15 (Charakterisierungen von Diagonalisierbarkeit)
Für $A \in M_n(\mathbb{K})$ sind äquivalent:

a) Es gibt $S \in GL_n(\mathbb{K})$, so dass $S^{-1}AS$ diagonal ist. In diesem Fall nennt man A **diagonalisierbar** *und S eine* **diagonalisierende Matrix** *für A.*

b) $\mathbb{K}^n$ hat eine Basis B derart, dass ${}^B_B L_A$ eine Diagonalmatrix ist.

c) $\mathbb{K}^n$ hat eine Basis, die aus Eigenvektoren von A besteht.

d) Seien $\lambda_1, \ldots, \lambda_r \in \mathbb{K}$ alle paarweise verschiedenen Eigenwerte von A. Dann $\sum_{i=1}^{r} \dim E_{\lambda_i}(A) = n$.

Beweis:
c) $\Leftrightarrow$ d) wegen Lemma 6.11 b). a) $\Leftrightarrow$ b) wegen Lemma 6.12. b) $\Leftrightarrow$ c): ${}^B_B L_A$ ist genau dann diagonal, wenn B aus Eigenvektoren besteht. □

Problem 6.16 (Diagonalisierbarkeit)
Gegeben $A \in M_n(\mathbb{K})$, gesucht eine diagonalisierende Matrix, falls sie existiert.

Lösung:
Berechne alle paarweise verschiedenen Eigenwerte $\lambda_1, ..., \lambda_k \in \mathbb{K}$ von A und jeweils Basen von $E_{\lambda_1}(A), ..., E_{\lambda_k}(A)$. Wenn man nicht insgesamt n Basisvektoren findet, ist A nicht diagonalisierbar. Andernfalls bilden die Basen der Eigenräume zusammen eine Basis $[\vec{v}_1, ..., \vec{v}_n]$ von $\mathbb{K}^n$. Mit $S := (\vec{v}_1, ..., \vec{v}_n)$ ist $S^{-1}AS$ diagonal, das i-te Diagonalelement ist ist der Eigenwert von $\vec{v}_i$.
Spezialfälle, wenn nur nach der Existenz einer diagonalisierenden Matrix gefragt ist:

- Hat A n verschiedene Eigenwerte in $\mathbb{K}$, ist A diagonalisierbar, denn jeder Eigenraum ist mindestens eindimensional.
- Hat A Eigenwerte, die nicht in $\mathbb{K}$ liegen, ist A über $\mathbb{K}$ nicht diagonalisierbar.
- Ist die algebraische Vielfachheit eines Eigenwertes größer als seine geometrische Vielfachheit, ist A nicht diagonalisierbar. □

Beispiel Die Matrix aus Bsp. 6.8.c) ist diagonalisierbar. Mit $S := \begin{pmatrix} -2 & -1 & -1 & -3 \\ 5 & 0 & -1 & -4 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & 2 & 5 \end{pmatrix}$ ist $S^{-1} \cdot A \cdot S = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}$. Die entstehende Diagonalmatrix kann man anhand der Eigenwerte direkt angeben (ohne zusätzliche Rechnung!).

Beispiel $A = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 2 \end{smallmatrix}\right)$ hat Eigenvektoren $\left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)$ mit Eigenwert 1 und $\left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right)$ mit Eigenwert 2, also diagonalisierbar. Mit $S = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ ist $S^{-1}AS = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 2 \end{smallmatrix}\right)$.

Beispiel $A = \left(\begin{smallmatrix} 3 & -1 \\ 4 & -1 \end{smallmatrix}\right)$ ist wegen d) nicht diagonalisierbar, denn $\chi_A(X) = X^2 - 2X+1 = (X-1)^2$, d.h. 1 ist der einzige Eigenwert – und $E_1(A)$ ist eindimensional, mit Basis $[(\begin{smallmatrix} 1 \\ 2 \end{smallmatrix})]$.

Beispiel Für $A = (\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix})$ ist $\chi_A(X) = X^2 + 1$. Also nicht diagonalisierbar über $\mathbb{R}$ (aber schon über $\mathbb{C}$).

Beispiel Für $A = \left(\begin{smallmatrix} 1 & 5 & -2 \\ 0 & 4 & 7 \\ 0 & 0 & -1 \end{smallmatrix}\right)$ ist $\chi_A(X) = (X-1)(X-4)(X+1)$, denn $X\mathbb{1}_3 - A$ hat obere Dreiecksgestalt. Also drei verschiedene Eigenwerte $1, 4, -1$, daher diagonalisierbar.

Definition 6.17
$A \in M_n(\mathbb{K})$ heißt **symmetrisch**, *gdw. $A^\top = A$.*

Der folgende Satz wird hier nicht bewiesen. Mehr Hintergrund dazu wird es im Abschlusskapitel über euklidische Geometrie geben. Man beachte, dass der Satz nicht für $\mathbb{K} = \mathbb{Q}$ oder für endliche Körper gilt.

Satz 6.18
Jede symmetrische Matrix $A \in M_n(\mathbb{R})$ ist diagonalisierbar.

Es gilt sogar eine Verallgemeinerung für Matrizen über den komplexen Zahlen. Ist $A \in M_n(\mathbb{C})$ und $A = \overline{A^\top}$ (Erinnerung: $\overline{a + b\,\mathrm{i}} := a - b\,\mathrm{i}$ bezeichnet die komplexe Konjugation, und das soll hier auf jeden Eintrag von $A^\top$ angewandt werden), so heißt A ***hermitesch***[31]. Jede hermitesche Matrix ist diagonalisierbar und alle ihre Eigenwerte sind *reell* (obwohl ja die Matrixeinträge komplex sind!). Das hat wichtige Anwendungen in der Physik, denn beobachtbare physikalische Größen in der Quantenmechanik entsprechen Eigenwerten hermitescher Operatoren — sind also reell, und das ist gut, denn ein nicht-reeller Messwert wäre seltsam.

[31] Nach Charles Hermite [1822–1901]

7 Euklidische Räume

In diesem Kapitel müssen wir in der Lage sein, aus positiven Körperelementen Wurzeln zu ziehen. Weil es in $\mathbb{C}$ keinen Begriff von „Positivität“ gibt, scheidet $\mathbb{C}$ aus. Weil man in $\mathbb{Q}$ nicht immer Wurzeln ziehen kann, scheidet $\mathbb{Q}$ aus. In diesem Kapitel sei daher $\mathbb{K} = \mathbb{R}$.

Zur Einordnung: In diesem einleitenden Abschnitt gibt es keine Definition! Eine solche wird es erst im nächsten Abschnitt geben. Stattdessen wollen wir die geometrische Anschauung aufzeigen, durch die die Begriffsbildung motiviert ist.

Wir tun zunächst so, als seien für $\vec{v}, \vec{w} \in V := \mathbb{R}^n$ die Begriffe *Länge* $\|\vec{v}\|$ von $\vec{v}$ und *Winkel* $\sphericalangle(\vec{v}, \vec{w})$ zwischen $\vec{v}$ und $\vec{w}$ bereits bekannt. Aus der Schule ist bekannt, dass das Skalarprodukt $\langle \vec{v} | \vec{w} \rangle$ von $\vec{v}$ und $\vec{w}$ mit Längen und Winkel zu tun hat: $\langle \vec{v} | \vec{w} \rangle := \|\vec{v}\| \cdot \|\vec{w}\| \cdot \cos \sphericalangle(\vec{v}, \vec{w})$, das heißt, die senkrechte Projektion von $\vec{w}$ auf die durch $\vec{v}$ gegebene Gerade hat die Länge $|\langle \vec{v} | \vec{w} \rangle / \|\vec{v}\||$. Geometrische Überlegungen besagen:

a) $\langle \vec{v} | \vec{w} \rangle = \langle \vec{w} | \vec{v} \rangle$.

b) Sind $\vec{v}$, $\vec{w}$ zueinander senkrecht, so ist $\langle \vec{v} | \vec{w} \rangle = 0$.

c) $\|\vec{v}\|^2 = \langle \vec{v} | \vec{v} \rangle$.

d) $\forall c \in \mathbb{R}\colon\ \langle c\vec{v} | \vec{w} \rangle = \langle \vec{v} | c\vec{w} \rangle = c \langle \vec{v} | \vec{w} \rangle$

e) $\forall \vec{u}, \vec{v}, \vec{w} \in V\colon\ \langle \vec{u} | \vec{v} + \vec{w} \rangle = \langle \vec{u} | \vec{v} \rangle + \langle \vec{u} | \vec{w} \rangle$.

f) $\|\vec{v}\| \geq 0$. Außerdem $\vec{v} = \vec{0} \iff \|\vec{v}\| = 0$.

7.1 Skalarprodukte

Wir erinnern daran, dass wir Vektoren $\vec{v} \in \mathbb{R}^n$ als *Spalten*vektoren betrachten, d.h. $\vec{v} \in \mathbb{R}^{n \times 1}$, und dementsprendend ist $\vec{v}^\top \in \mathbb{R}^{1 \times n}$ der zugehörige Zeilenvektor.

Definition 7.1
Sei V ein $\mathbb{R}$-Vektorraum.

a) Eine **Bilinearform** *auf V ist eine multilineare Abbildung $b\colon\ V \times V \to \mathbb{R}$.*

b) Eine Bilinearform b auf V ist **symmetrisch**, *gdw. $\forall \vec{v}, \vec{w} \in V\colon\ b(\vec{v}, \vec{w}) = b(\vec{w}, \vec{v})$.*

c) Eine Bilinearform b auf V heißt **positiv definit**, *gdw.*

$$\forall \vec{v} \in V\colon \quad b(\vec{v}, \vec{v}) > 0 \iff \vec{v} \neq \vec{0}\,.$$

d) Ein **Skalarprodukt** *auf V ist eine positiv definite symmetrische Bilinearform. Ein Skalarprodukt notieren wir hier meist als $\langle \vec{v} \mid \vec{w} \rangle$ statt $b(\vec{v}, \vec{w})$.*

e) Ein reeller Vektorraum V mit einem Skalarprodukt bezeichnet man als **Skalarproduktraum**[32]*. Ein* **euklidischer Raum** *ist ein endlichdimensionaler Skalarproduktraum.*

Für Skalarprodukte gibt es verschiedene übliche Notationen; die hier verwendete ist auch in der Quantenmechanik sehr gebräuchlich und hat den Vorteil, dass eine Verwechslung mit dem Matrixprodukt vermieden wird.

Beispiel 7.2 (und Definition)

a) Das **Standardskalarprodukt** *auf $\mathbb{R}^n$ ist definiert durch $\langle\vec{v}|\vec{w}\rangle := \sum_{i=1}^{n} v_i w_i = \vec{w}^\top \cdot \vec{v}$. Es ist ein Skalarprodukt.*

b) Ist b eine Bilinearform auf einem endlich-dimensionalen $\mathbb{R}$-VR V und $B = [\vec{v}_1, ..., \vec{v}_n]$ eine Basis von V, so ist die **Darstellungsmatrix** *${}_Bb \in M_n(\mathbb{R})$ von b bzgl. B gegeben durch ${}_Bb_{i,j} = b(\vec{v}_j, \vec{v}_i)$. Offenbar ist b genau dann symmetrisch, wenn ${}_Bb$ symmetrisch ist. Das Standardskalarprodukt ist der Fall ${}_Eb = \mathbb{1}_n$ mit der Standardbasis E.*

Für $\vec{v}, \vec{w} \in V$ mit ${}^B\vec{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$ und ${}^B\vec{w} = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}$ gilt

$$b(\vec{v}, \vec{w}) = b\left(\sum_{j=1}^{n} v_j\vec{v}_j, \sum_{i=1}^{n} w_i\vec{v}_i\right) = \sum_{i,j=1}^{n} w_i\, {}_Bb_{i,j} v_j = {}^B\vec{w}^\top \cdot {}_Bb \cdot {}^B\vec{v}\,.$$

Ist umgekehrt $A \in M_n(\mathbb{R})$ vorgegeben, so gibt es eine eindeutig bestimmte Bilinearform b auf V mit ${}_Bb = A$, indem man nämlich obige Formel zur Definition von b nutzt.

c) $A = \left(\begin{smallmatrix} 1 & 1 \\ 1 & 2 \end{smallmatrix}\right)$ definiert die symmetrische Bilinearform

$$b\left(\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right), \left(\begin{smallmatrix} r \\ s \end{smallmatrix}\right)\right) = \left(\begin{smallmatrix} r & s \end{smallmatrix}\right)\left(\begin{smallmatrix} 1 & 1 \\ 1 & 2 \end{smallmatrix}\right)\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right) = xr + (xs + yr) + 2ys$$

auf $\mathbb{R}^2$ bzgl. der Standardbasis. Ferner ist b positiv definit, denn

$$b\left(\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right), \left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right)\right) = x^2 + 2xy + 2y^2 = (x + y)^2 + y^2\,.$$

Also ist b ein Skalarprodukt auf $\mathbb{R}^2$.

d) Die durch $A = \left(\begin{smallmatrix} 1 & -2 \\ -2 & 1 \end{smallmatrix}\right)$ definierte symmetrische Bilinearform ist kein Skalarprodukt: Zwar ist $b(\vec{e}_1, \vec{e}_1) = b(\vec{e}_2, \vec{e}_2) = 1 > 0$, aber für $\vec{v} = \left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right)$ ist $b(\vec{v}, \vec{v}) = 1 - 2 - 2 + 1 = -2 < 0$.

e) Wir werden (hoffentlich) noch zeigen: Eine durch $A \in M_n(\mathbb{R})$ gegebene symmetrische Bilinearform auf $\mathbb{R}^n$ ist genau dann ein Skalarprodukt, wenn alle Eigenwerte von A positive reelle Zahlen sind.

[32] Auch: ***Prä-Hilbertraum***; David Hilbert [1862–1943]

f) In der Analysis bestimmt man die Extremwerte einer Funktion $f\colon \mathbb{R} \to \mathbb{R}$ indem man die Gleichung $f'(x) = 0$ löst, um die stationären Punkten zu bestimmen, und dann in jedem stationären Punkt den Wert von $f''(x)$ untersucht. Insbesondere gilt: Wenn $f''(x) > 0$, dann ist der stationäre Punkt ein lokales Minimum.

Für Funktionen $f \in \mathscr{C}^2(\mathbb{R}^n, \mathbb{R})$ gibt es eine vergleichbare Vorgehensweise: Aufgrund des Satzes von Schwarz ist die **Hesse-Matrix** *H gegeben durch $H_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}$ symmetrisch; und die Bedingung $f'' > 0$ wird ersetzt durch die Bedingung, dass H positiv definit ist.*

g) Quantenmechanische Zustände werden durch eine so genannte Wellenfunktion *$\varphi\colon \mathbb{R}^3 \to \mathbb{C}$ beschrieben. Für Wellenfunktionen φ, ψ definiert man $\langle \varphi \mid \psi \rangle := \int_{\mathbb{R}^3} \bar{\varphi}(\vec{x})\psi(\vec{x})\,\mathrm{d}\vec{x}$. Das ist nicht symmetrisch, aber* **hermitesch***: $\langle \varphi \mid \psi \rangle = \overline{\langle \psi \mid \varphi \rangle}$ und linear in der zweiten Komponente. Wenn wir beweisen, dass symmetrische reelle Matrizen diagonalisierbar sind, werden wir auf „hermitesch“ näher eingehen.*

Definition 7.3. *Sei $(V, \langle|\rangle)$ ein Skalarproduktraum und $\vec{v}, \vec{w} \in V$.*

a) $\|\vec{v}\| := \sqrt{\langle \vec{v}|\vec{v}\rangle}$, die **Länge** *oder* **Norm** *von $\vec{v}$. Einen Vektor der Länge 1 nennt man auch* **normiert**.

b) Falls $\vec{v}, \vec{w} \neq \vec{0}$: $\sphericalangle(\vec{v}, \vec{w}) := \arccos \frac{\langle \vec{v}|\vec{w}\rangle}{\|\vec{v}\| \cdot \|\vec{w}\|}$, der **Winkel** *zwischen $\vec{v}$ und $\vec{w}$.*

c) $\vec{v} \perp \vec{w}$, d.h. $\vec{v}$ und $\vec{w}$ sind zueinander **orthogonal** *$:\Leftrightarrow \langle \vec{v}|\vec{w}\rangle = 0$.*

Beachte: $\vec{u} \perp \vec{v} \Leftrightarrow \vec{v} \perp \vec{u}$ wegen Symmetrie des Skalarprodukts.

Beispiel 7.4 (und Definition)
Betrachte $\mathbb{R}^3$ als euklidischen Raum mit dem Standardskalarprodukt. Für $\vec{v}, \vec{w} \in \mathbb{R}^3$ definiert man das **Kreuzprodukt** *(auch* **Vektorprodukt** *genannt) von $\vec{v}$ und $\vec{w}$ durch*

$$\vec{v} \times \vec{w} := \begin{pmatrix} v_2 w_3 - v_3 w_2 \\ v_3 w_1 - v_1 w_3 \\ v_1 w_2 - v_2 w_1 \end{pmatrix}$$

Es ist $\vec{v} \times \vec{w} = \vec{0}$ gdw. $\vec{v}, \vec{w}$ linear abhängig sind. Bezüglich des Standardskalarprodukts ist $\vec{v} \perp (\vec{v} \times \vec{w})$ und $\vec{w} \perp (\vec{v} \times \vec{w})$.
Ferner gilt $\vec{w} \times \vec{v} = -\vec{v} \times \vec{w}$ (anti-kommutativ). Das Assoziativgesetz gilt für das Kreuzprodukt im Allgemeinen nicht. Das Distributivgesetz hingegen gilt für das Kreuzprodukt.

Die folgenden geometrischen Eigenschaften des Kreuzprodukts sind nicht so einfach nachzurechnen. Da es sich nur um eine Illustration handelt, beweise ich sie nicht:

a) $\|\vec{v} \times \vec{w}\| = \|\vec{v}\|\|\vec{w}\| \,|\sin \alpha|$, wobei $\alpha := \sphericalangle(\vec{v}, \vec{w})$.

b) **Rechte-Hand-Regel***: Wenn der abgespreizte Daumen bzw. Zeigefinger der rechten Hand in Richtung $\vec{v}$ bzw. $\vec{w}$ zeigen, dann zeigt $\vec{v} \times \vec{w}$ senkrecht von der Handfläche weg.*

7.2 Orthonormalbasen

Vermutlich sagt Ihnen die aus der Schule mitgebrachte Anschauung, dass die Standardbasisvektoren paarweise aufeinander senkrecht stehen und zudem normiert sind. Dies ist allerding nur bezüglich des Standardskalarprodukts der Fall, so dass wir zur folgenden Begriffsbildung gelangen.

Definition 7.5. *Sei V ein Skalarproduktraum.*

a) $(\vec{v}_i)_{i\in I} \subset V$ *heißt* **Orthogonalsystem**, *gdw.* $\vec{v}_i \perp \vec{v}_j$ *für alle* $i \neq j \in I$.

b) Ein Orthogonalsystem $(\vec{v}_i)_{i\in I} \subset V$ *heißt* **Orthonormalsystem** $:\Leftrightarrow \forall i \in I\colon\ \|\vec{v}_i\| = 1$.

Beobachtung 7.6 (und Definition)

a) Ist $(V, \langle|\rangle)$ *ein Skalarproduktraum, so ist* $(\vec{v}_i)_{i\in I} \subset V$ *genau dann ein Orthonormalsystem wenn* $\forall i,j \in I\colon\ \langle \vec{v}_i \mid \vec{v}_j\rangle = \delta_{i,j} := \begin{cases} 1 & (i=j) \\ 0 & (i\neq j) \end{cases}$, *mit dem so genannten* **Kronecker-Delta**[33] $\delta_{i,j}$.

b) Eine **Orthonormalbasis** *eines euklidischen Raums V ist eine Basis von V, die zugleich ein Orthonormalsystem ist.*

c) Ist (V,b) ein euklidischer Raum mit einer ONB D, dann ${}_Db = \mathbb{1}_{\dim V}$. Verwendet man also bei vorgegebenen Skalarprodukt für alle Rechnungen eine ONB, so sind die Rechnungen genau wie beim Standardskalarprodukt. Später werden wir zeigen, wie man eine ONB berechnen kann.

Beispiel 7.7

a) Die Standardbasis $[\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n]$ ist eine ONB von $\mathbb{R}^n$ bzgl. Standardskalarprodukt.

b) In $\mathbb{R}^3$ mit dem Standardskalarprodukt gelten u.a.

i) $\begin{pmatrix}1\\2\\4\end{pmatrix} \perp \begin{pmatrix}2\\1\\-1\end{pmatrix}$ *und* $\begin{pmatrix}1\\3\\1\end{pmatrix} \not\perp \begin{pmatrix}2\\1\\-1\end{pmatrix}$.

ii) $\begin{pmatrix}2\\1\\1\end{pmatrix}, \begin{pmatrix}-1\\2\\0\end{pmatrix}, \begin{pmatrix}0\\0\\0\end{pmatrix}$ *ist ein orthogonales System.*

[33] Manchmal wird es auch Kronecker-Symbol genannt, aber leider wird „Kronecker-Symbol“ noch für einen ganz anderen Begriff verwendet!

iii) $\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\\frac{1}{\sqrt{2}}\\\frac{1}{\sqrt{2}}\end{pmatrix}$ *ist ein Orthonormalsystem.*

Lemma 7.8
Jedes Orthonormalsystem S in einem Skalarproduktraum V ist linear unabhängig.

Beweis:
Sei I die Indexmenge von S. Es ist $\forall j \in I\colon \left\langle \sum_{i\in I}{}' \lambda_i \vec{v}_i \mid \vec{v}_j \right\rangle = \sum_{i\in I}{}' \lambda_i \langle \vec{v}_i \mid \vec{v}_j\rangle = \lambda_j$ (diese Formel sollten Sie sich merken!). Daher: $\sum_{i\in I}{}' \lambda_i \vec{v}_i = \vec{0} \Rightarrow \lambda_i = 0$ für alle i.□

7.2.1 Orthonormalisierungsverfahren (Gram–Schmidt)

Beim Basisergänzungsproblem sollte ein linear unabhängiges System von Vektoren zu einer Basis fortgesetzt werden. Analog stellt sich nun das Problem, ein gegebenes Orthonormalsystem zu einer ONB fortzusetzen. Dazu gibt es verschiedene Verfahren. Eines der einfachsten ist das *Gram–Schmidt-Verfahren*. Es gibt andere Verfahren, die bei gerundetem Rechnen besser geeignet sind.

Problem 7.9
Gegeben sei eine Basis $[\vec{u}_1, \ldots, \vec{u}_d]$ eines euklidischen Vektorraums $(V, \langle|\rangle)$, so dass $[\vec{u}_1, ..., \vec{u}_k]$ für ein $0 \le k \le d$ ein Orthonormalsystem bilden.
Berechne eine ONB $[v_1, ..., v_d]$ von V, so dass $\vec{v}_1 = \vec{u}_1$, ..., $\vec{v}_k = \vec{u}_k$ und $\forall r \in \{k+1, ..., d\}\colon$ $\mathrm{Span}(\vec{v}_1, ..., \vec{v}_r) = \mathrm{Span}(\vec{u}_1, ..., \vec{u}_r)$.

Lösung (Gram-Schmidt-Verfahren):
Sei $\vec{v}_1 := \vec{w}_1 := \vec{u}_1$, ..., $\vec{v}_k := \vec{w}_k := \vec{u}_k$. Für $k < r \le d$ definiere nacheinander

$$\vec{w}_r := \vec{u}_r - \sum_{i=1}^{r-1} \frac{\langle \vec{u}_r | \vec{w}_i\rangle}{\langle \vec{w}_i | \vec{w}_i\rangle} \vec{w}_i \qquad \text{und} \qquad \vec{v}_r := \frac{\vec{w}_r}{\|\vec{w}_r\|}\ .$$

Hierbei kann man in Zwischenergebnissen die Vektoren $\vec{w}_r$ auch skalieren.
Die gesuchte ONB von V ist $\vec{v}_1, ..., \vec{v}_d$. Es gilt nämlich $\forall r \in \{1, ..., d\}\colon$

- $\mathrm{Span}(\vec{u}_1, ..., \vec{u}_r) = \mathrm{Span}(\vec{w}_1, ..., \vec{w}_r) = \mathrm{Span}(\vec{v}_1, ..., \vec{v}_r)$ (Induktion nach r),
- $\|\vec{v}_r\| = 1$
- $\forall j \in \{1, ..., r-1\}\colon \langle \vec{w}_r \mid \vec{w}_j\rangle = \langle \vec{v}_r \mid \vec{v}_j\rangle = 0$. Denn wenn $[\vec{v}_1, ..., \vec{v}_{r-1}]$ ein Orthonormalsystem ist, so gilt $\forall j \in \{1, ..., r-1\}$:

$$\begin{aligned}\left\langle \vec{u}_r - \sum_{i=1}^{r-1} \frac{\langle \vec{u}_r|\vec{w}_i\rangle}{\langle \vec{w}_i|\vec{w}_i\rangle}\vec{w}_i \,\middle|\, \vec{v}_j \right\rangle &= \langle \vec{u}_r|\vec{u}_j\rangle - \sum_{i=1}^{r-1} \frac{\langle \vec{u}_r|\vec{w}_i\rangle}{\langle \vec{w}_i|\vec{w}_i\rangle}\langle \vec{v}_i|\vec{v}_j\rangle \\ &= \langle \vec{u}_r|\vec{u}_j\rangle - \frac{\langle \vec{u}_r|\vec{w}_j\rangle}{\langle \vec{w}_j|\vec{w}_j\rangle}\langle \vec{v}_j|\vec{v}_j\rangle = 0 \qquad \square\end{aligned}$$

Beispiel Sei $V \leq \mathbb{R}^4$ der Lösungsraum der Gleichung $x_1 + x_2 + x_3 + x_4 = 0$. Konstruiere eine Orthonormalbasis von V bzgl. Standardskalarprodukt.

Sei $\vec{u}_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix}$, $\vec{u}_2 = \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix}$, $\vec{u}_3 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ -1 \end{pmatrix}$; dann ist $[\vec{u}_1, \vec{u}_2, \vec{u}_3]$ eine Basis von V. Zuerst setzen wir $\vec{w}_1 = \vec{u}_1$, also $\vec{w}_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix}$ und $\langle \vec{w}_1 | \vec{w}_1 \rangle = 2$. Jetzt setzen wir $\vec{w}_2 = \vec{u}_2 - \frac{\langle \vec{u}_2 | \vec{w}_1 \rangle}{\langle \vec{w}_1 | \vec{w}_1 \rangle} \vec{w}_1$, d.h. $\vec{w}_2 = \vec{u}_2 - \frac{1}{2}\vec{w}_1 = \begin{pmatrix} 1/2 \\ 1/2 \\ -1 \\ 0 \end{pmatrix}$. Um vorerst Brüche zu vermeiden, multiplizieren wir mit 2, also $\vec{w}_2 = \begin{pmatrix} 1 \\ 1 \\ -2 \\ 0 \end{pmatrix}$ und $\langle \vec{w}_2 | \vec{w}_2 \rangle = 6$.

Des Weiteren ist $\vec{w}_3 = \vec{u}_3 - \frac{\langle \vec{u}_3 | \vec{w}_1 \rangle}{\langle \vec{w}_1 | \vec{w}_1 \rangle} \vec{w}_1 - \frac{\langle \vec{u}_3 | \vec{w}_2 \rangle}{\langle \vec{w}_2 | \vec{w}_2 \rangle} \vec{w}_2 = \vec{u}_3 - \frac{1}{2}\vec{w}_1 - \frac{1}{6}\vec{w}_2 = \begin{pmatrix} 1/3 \\ 1/3 \\ 1/3 \\ -1 \end{pmatrix}$, mit $\langle \vec{w}_3 | \vec{w}_3 \rangle = \frac{4}{3}$. Dies ergibt die ONB $[\begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1/\sqrt{6} \\ 1/\sqrt{6} \\ -2/\sqrt{6} \\ 0 \end{pmatrix}, \begin{pmatrix} 1/(2\sqrt{3}) \\ 1/(2\sqrt{3}) \\ 1/(2\sqrt{3}) \\ -3/(2\sqrt{3}) \end{pmatrix}]$.

Als unmittelbare Folge erhält man:

Orthonormalisierungssatz
Jeder euklidische Vektorraum hat eine ONB und jedes Orthonormalsystem in einem euklidischen Vektorraum kann man zu einer ONB fortsetzen. □

7.2.2 Das orthogonale Komplement

In diesem Abschnitt sei $(V, \langle | \rangle)$ ein Skalarproduktraum.

Definition 7.10 (und Übung)
Für $M \subset V$ heißt $M^\perp := \{\vec{v} \in V \mid \forall \vec{u} \in M \colon \vec{v} \perp \vec{u}\} \leq V$ das **orthogonale Komplement** *von M.*

Beispiel Ist $U \subseteq \mathbb{R}^3$ ein zweidimensionaler Unterraum, d.h. eine Ebene durch den Ursprung, so ist $U^\perp$ eine zu U senkrechte Ursprungsgerade.

Lemma 7.11. *Sei $M \subset V$.*

a) $M_1 \subset M_2 \subset V \Rightarrow M_2^\perp \subset M_1^\perp$

b) $M^\perp = (\mathrm{Span}(M))^\perp$

c) $M \subset (M^\perp)^\perp$

Beweis:
a) Die Bedingungen in der Definition von $M_1^\perp$ gelten auch für alle Elemente von $M_2^\perp$.

b) $\supset$ folgt aus a). Sei $\vec{v} \in M^\perp$ und $\vec{u} = \sum_{\vec{b}\in M}{}' \lambda_{\vec{b}}\vec{b} \in \mathrm{Span}(M)$. Dann $\langle \vec{v} \mid \vec{u}\rangle = \sum_{\vec{b}\in M}{}' \lambda_{\vec{b}}\langle \vec{v}|\vec{b}\rangle = 0$.

c) Übung $\square$

Satz 7.12
Sei V euklidisch. Dann gilt für jedes $U \leq V$:

$$U \oplus U^\perp = V \qquad \textit{und} \qquad (U^\perp)^\perp = U$$

Bemerkung Für unendlichdimensionale Skalarprodukträume können die beiden Aussagen verletzt sein!

Beweis:
Sei $\dim(U) = r$ und $\dim(V) = n$. Wegen des Orthonormalisierungssatzes gibt es eine ONB $[\vec{v}_1, ..., \vec{v}_n]$ von V, so dass $[\vec{v}_1, ..., \vec{v}_r]$ eine ONB von U ist. Behauptung: $U^\perp = \mathrm{Span}(\vec{v}_{r+1}, ..., \vec{v}_n)$. Sei dazu $\vec{v} \in V$, $\vec{v} = \sum_{i=1}^{n} \lambda_i \vec{v}_i$.

- Wenn $\vec{v} \in U^\perp$, dann $\forall i \in \{1, ..., r\}\colon\ 0 = \langle \vec{v} \mid \vec{v}_i\rangle = \lambda_i$, also $\vec{v} \in \mathrm{Span}(\vec{v}_{r+1}, ..., \vec{v}_n)$.
- Wenn $\vec{v} \in \mathrm{Span}(\vec{v}_{r+1}, ..., \vec{v}_n)$, dann $\forall i \in \{1, ..., r\}\colon\ \langle \vec{v} \mid \vec{v}_i\rangle = \sum_{j=r+1}^{n} \lambda_j \langle \vec{v}_j|\vec{v}_i\rangle = 0$. Also $\vec{v} \in [\vec{v}_1, ..., \vec{v}_r]^\perp = U^\perp$.

Aus $U = \mathrm{Span}(\vec{v}_1, ..., \vec{v}_r)$ und $U^\perp = \mathrm{Span}(\vec{v}_{r+1}, ..., \vec{v}_n)$ folgt $U \oplus U^\perp = V$ und $(U^\perp)^\perp = U$. $\square$

Lemma 7.13 (und Definition)
Sei $U \leq V$ so, dass $U \oplus U^\perp = V$ (also zum Beispiel V euklidisch). Es gibt genau eine lineare Abbildung $\pi_U\colon V \to U$, so dass $\forall \vec{v} \in V\colon (\vec{v} - \pi_U(\vec{v})) \in U^\perp$. Man nennt π_U die **orthogonale Projektion** *auf U und nennt $\pi_U(\vec{v})$ den* **Lotfußpunkt** *von $\vec{v}$ in U.*
Ist B eine ONB von U, so gilt $\pi_U(\vec{v}) = \sum_{\vec{b}\in B}{}' \langle \vec{v}|\vec{b}\rangle \vec{b}$.

Beweis:
Eindeutigkeit: Sei $\vec{v} \in V$, $\vec{u}_1, \vec{u}_2 \in U$ mit $\vec{v} - \vec{u}_1 \in U^\perp$ und $\vec{v} - \vec{u}_2 \in U^\perp$. Für $\vec{u} := \vec{u}_2 - \vec{u}_1 \in U$ folgt

$$\begin{aligned} 0 =& \langle (\vec{v} - \vec{u}_1) \mid \vec{u}\rangle - \langle (\vec{v} - \vec{u}_2) \mid \vec{u}\rangle \\ & \langle (\vec{v} - \vec{u}_1 - \vec{v} + \vec{u}_2) \mid \vec{u}\rangle = \langle \vec{u}|\vec{u}\rangle = \|\vec{u}\|^2 \end{aligned}$$

und daher $\vec{u} = \vec{0}$.
Existenz: Wegen $U \oplus U^\perp = V$ gibt es für alle $\vec{v} \in V$ eine Darstellung $\vec{v} = \vec{u} + \vec{w}$ mit $\vec{u} \in U$ und $\vec{w} \in U^\perp$; dann sei $\pi_U(\vec{v}) := \vec{u}$. Wir zeigen, dass π_U linear ist: Seien $\vec{v}_1 = \pi_U(\vec{v}_1) + \vec{w}_1$ und $\vec{v}_2 = \pi_U(\vec{v}_2) + \vec{w}_2$ mit $\vec{w}_1, \vec{w}_2 \in U^\perp$. Für $\lambda, \mu \in \mathbb{R}$ folgt

$$\begin{aligned}\lambda\vec{v}_1 + \mu\vec{v}_2 =& \lambda(\pi_U(\vec{v}_1) + \vec{w}_1) + \mu(\pi_U(\vec{v}_2) + \vec{w}_2) \\ =& (\lambda\pi_U(\vec{v}_1) + \mu\pi_U(\vec{v}_2)) + (\lambda\vec{w}_1 + \mu\vec{w}_2)\ .\end{aligned}$$

Wegen $\lambda\vec{w}_1 + \mu\vec{w}_2 \in U^\perp$ (Vektorraum!) und $\lambda\pi_U(\vec{v}_1) + \mu\pi_U(\vec{v}_2) \in U$ folgt $\pi_U(\lambda\vec{v}_1 + \mu\vec{v}_2) = \lambda\pi_U(\vec{v}_1) + \mu\pi_U(\vec{v}_2)$.
Formel: Siehe Beweis der Korrektheit des Gram-Schmidt-Verfahrens. □

Lemma 7.14
Sei V euklidisch. Für alle $\vec{v}, \vec{w} \in V$ gelten:

- $|\langle\vec{v}|\vec{w}\rangle| \leq \|\vec{v}\| \cdot \|\vec{w}\|$ ***(Cauchy-Schwarz-Ungleichung)***
- $\|\vec{v} + \vec{w}\| \leq \|\vec{v}\| + \|\vec{w}\|$ ***(Dreiecksungleichung)***

In beiden Teilen gilt: Im Gleichheitsfall ist $[\vec{v}, \vec{w}]$ linear abhängig.

Beweis:
Wir nutzen $V = \mathrm{Span}(\vec{v}) \oplus (\mathrm{Span}(\vec{v}))^\perp$.
Cauchy-Schwarz: $\vec{w} = \lambda\vec{v} + \vec{w}'$ für ein $\vec{w}' \in (\mathrm{Span}(\vec{v}))^\perp$ und $\lambda \in \mathbb{R}$, also $\|\vec{w}\|^2 = |\lambda|^2 \|\vec{v}\|^2 + \|\vec{w}'\|^2$. Daher $|\langle\vec{v}|\vec{w}\rangle|^2 = |\lambda|^2 \|\vec{v}\|^4 \leq |\lambda|^2 \|\vec{v}\|^4 + \|\vec{v}\|^2 \|\vec{w}'\|^2 = \|\vec{v}\|^2 \|\vec{w}\|^2$. Für Gleichheit brauchen wir $\vec{v} = \vec{0}$ oder $\vec{w}' = \vec{0}$, in jedem Fall $[\vec{v}, \vec{w}]$ linear abhängig.
Dreieck: $\|\vec{v} + \vec{w}\|^2 = \|\vec{v}\|^2 + \|\vec{w}\|^2 + 2\langle\vec{v}|\vec{w}\rangle \leq \|\vec{v}\|^2 + \|\vec{w}\|^2 + 2\|\vec{v}\| \cdot \|\vec{w}\| = (\|\vec{v}\| + \|\vec{w}\|)^2$ wegen Cauchy-Schwarz. □

7.3 Besondere Endomorphismen

Sei $(V, \langle|\rangle)$ ein Skalarproduktraum.

Definition 7.15
Ein Endomorphismus $\varphi\colon V \to V$ heißt...

a) **orthogonal** $:\Leftrightarrow \forall\vec{v}, \vec{w} \in V\colon \langle\vec{v}|\vec{w}\rangle = \langle\varphi(\vec{v}) \mid \varphi(\vec{w})\rangle$. *Eine orthogonale Abbildung ist also längen- und winkelerhaltend.*

b) **selbstadjungiert** $:\Leftrightarrow \forall\vec{v}, \vec{w} \in V\colon \langle\vec{v} \mid \varphi(\vec{w})\rangle = \langle\varphi(\vec{v}) \mid \vec{w}\rangle$.

Lemma 7.16 (und Beispiel). *Sei V euklidisch und B eine ONB von V.*
Ein Endomorphismus φ von V ist genau dann orthogonal, wenn $\varphi(B)$ eine ONB ist. Dies ist etwa für Drehungen oder Spiegelungen im Anschauungsraum der Fall.

Beweis:
Sei $B := [\vec{v}_1, ..., \vec{v}_n]$. Wenn φ orthogonal ist, dann $\langle \varphi(\vec{v}_i) \mid \varphi(\vec{v}_j) \rangle = \langle \vec{v}_i | \vec{v}_j \rangle = \delta_{i,j}$, also ist $\varphi(B)$ Orthonormalsystem. Insbesondere ist es linear unabhängig, also nach dem Dimensionssatz eine Basis. Wenn $\varphi(B)$ eine ONB ist, dann ist $\forall \vec{v}, \vec{w} \in V$: $\langle \vec{v} | \vec{w} \rangle = {}^{B}\vec{w}^{\top} \, {}^{B}\vec{v} = {}^{\varphi(B)}\varphi(\vec{w})^{\top} \, {}^{\varphi(B)}\varphi(\vec{v}) = \langle \varphi(\vec{v}) \mid \varphi(\vec{w}) \rangle$. □

Beispiel 7.17
$V := \{f \in \mathscr{C}^{\infty}([0,1], \mathbb{R}) \mid f(0) = f(1) = 0\}$ ist mit $\langle f | g \rangle := \int_0^1 f(x)g(x)\,\mathrm{d}x$ ein Skalarproduktraum. Die Abbildung $\varphi\colon V \to V$ mit $\varphi(f) := -f''$ ist selbstadjungiert. Derartige Beispiele sind in der Quantenmechanik sehr wichtig.

Beweis:
Für alle $f, g \in V$ folgt mit partieller Integration:

$$\langle -f'' \mid g \rangle = \int_0^1 -f''(x)g(x)\,\mathrm{d}x = -f'(x)g(x)|_{x=0}^{x=1} - \int_0^1 (-f'(x))g'(x)\,\mathrm{d}x$$

$$= f(x)g'(x)|_{x=0}^{x=1} - \int_0^1 f(x)g''(x)\,\mathrm{d}x = \langle f \mid -g'' \rangle$$ □

Lemma 7.18
Sei V euklidisch, $\varphi\colon V \to V$ ein Endomorphismus, C eine Basis von V und A die Darstellungsmatrix des Skalarprodukts bezüglich C.

a) φ ist selbstadjungiert $\iff {}^{C}_{C}\varphi^{\top} \cdot A = A \cdot {}^{C}_{C}\varphi$.

b) φ ist orthogonal $\iff {}^{C}_{C}\varphi^{\top} \cdot A \cdot {}^{C}_{C}\varphi = A$.

Beweis:
Für alle $\vec{v}, \vec{w} \in V$ gilt:

a) $\langle \vec{v} | \varphi(\vec{w}) \rangle = \left({}^{C}_{C}\varphi \, {}^{C}\vec{w}\right)^{\top} \cdot A \cdot {}^{C}\vec{v}$ und $\langle \varphi(\vec{v}) | \vec{w} \rangle = {}^{C}\vec{w}^{\top} \cdot A \cdot {}^{C}_{C}\varphi \, {}^{C}\vec{v}$.

b) $\langle \varphi(\vec{v}) \mid \varphi(\vec{w}) \rangle = \left({}^{C}_{C}\varphi \, {}^{C}\vec{w}\right)^{\top} \cdot A \cdot {}^{C}_{C}\varphi \, {}^{C}\vec{v}$ und $\langle \vec{v} | \vec{w} \rangle = {}^{C}\vec{w}^{\top} \cdot A \cdot {}^{C}\vec{v}$. □

Lemma 7.19
Sei $(V, \langle | \rangle)$ ein euklidischer Raum, $\varphi\colon V \to V$ ein selbstadjungierter Endomorphismus und $\vec{u}, \vec{v} \in V$ Eigenvektoren von φ zu Eigenwerten $\lambda \neq \mu$. Dann $\vec{u} \perp \vec{v}$.

Beweis:
$\mu\langle \vec{u} | \vec{v} \rangle = \langle \vec{u} | \mu\vec{v} \rangle = \langle \vec{u} | \varphi(\vec{v}) \rangle = \langle \varphi(\vec{u}) | \vec{v} \rangle = \langle \lambda\vec{u} | \vec{v} \rangle = \lambda\langle \vec{u} | \vec{v} \rangle$.
Also $(\lambda - \mu) \cdot \langle \vec{u} | \vec{v} \rangle = 0$ und wegen $\lambda \neq \mu$ folgt $\langle \vec{u} | \vec{v} \rangle = 0$. □

Beispiel 7.20
Sei V und φ wie in Beispiel 7.17. Für $n \in \mathbb{N}$ sei $f_n := ((x \mapsto \sin(n\pi x))$. Dann ist $f_n \in V$ ein Eigenvektor von φ zum Eigenwert $n^2\pi^2$. Für $m \neq n \in \mathbb{N}$ gilt also nach dem vorigen Lemma: $0 = \langle f_m \mid f_n \rangle = \int_0^1 \sin(m\pi x)\sin(n\pi x)\,\mathrm{d}x$. Das könnte man natürlich auch direkt nachrechnen, aber das wäre erheblich aufwändiger als der abstrakter Zugang über Eigenvektoren einer selbstadjungierten linearen Abbildung.

7.4 Die Hauptachsentransformation

Wir werden noch zeigen, dass jede symmetrische Matrix diagonalisierbar ist. Aus den Ergebnissen des vorigen Abschnitts folgern wir nun, dass eine Diagonalisierung durch eine skalarproduktserhaltende Koordinatentransformation möglich ist.

Definition 7.21
a) $A \in GL_n(\mathbb{R})$ *heißt* **orthogonal** $:\Leftrightarrow A^{-1} = A^\top$.

b) $O_n := \{A \in GL_n(\mathbb{R}) \mid A^{-1} = A^\top\}$.

Lemma 7.22 (und Definition)
Für $A \in M_n(\mathbb{R})$ gilt:

a) *A orthogonal $\Rightarrow \det(A) = \pm 1$. Die Umkehrung gilt nicht.*

b) *A orthogonal $\Leftrightarrow$ die Spalten von A bilden eine ONB von $\mathbb{R}^n$ bezüglich des Standardskalarprodukts.*

*Man definiert noch $SO_n := \{A \in O_n \mid \det(A) = 1\}$ (***spezielle orthogonale Matrizen***). Es sind $SO_n \leq O_n \leq GL_n(\mathbb{R})$ Untergruppen.*

Beweis:
b) ist klar. Ist A orthogonal, dann $\det(A)^2 = \det(A) \cdot \det(A^\top) = \det(A \cdot A^\top) = \det(\mathbb{1}_n) = 1$. Nachweis der Untergruppeneigenschaft: Übung. □

Satz: Die Hauptachsentransformation (HAT)
Jede symmetrische Matrix $A \in M_n(\mathbb{R})$ hat eine diagonalisierende Matrix $S \in SO_n$. Dann ist also $S^\top AS$ diagonal, denn für $S \in O_n$ ist $S^{-1} = S^\top$.

Berechnung einer HAT:
Wir verweisen auf Satz 6.18, den wir später beweisen: Jede symmetrische Matrix ist diagonalisierbar. Es geht jetzt nur noch um die Frage, wie man eine *speziell-orthogonale* diagonalisierende Matrix konkret berechnen kann.

Berechne die Eigenwerte von A sowie jeweils eine ONB der Eigenräume (Gram-Schmidt). Weil A diagonalisierbar ist, ergeben die Orthonormalbasen der Eigenräume insgesamt eine Basis $C = [\vec{c}_1, ..., \vec{c}_n]$ von $\mathbb{R}^n$.

Weil A symmetrisch ist, ist nach Lemma 7.18 $L_A\colon \mathbb{R}^n \to \mathbb{R}^n$ selbstadjungiert bzgl. Standardskalarprodukt, also sind nach Lemma 7.19 Vektoren aus verschiedenen Eigenräumen zueinander orthogonal. Also ist C eine ONB von $\mathbb{R}^n$.

Nach Lemma 7.22 ist $S := (\vec{c}_1, ..., \vec{c}_n) \in O_n$. Ist $\det(S) = 1$, so ist $S \in SO_n$ die gesuchte diagonalisierende Matrix. Andernfalls sei $\tilde{S} := (-\vec{c}_1, \vec{c}_2, ..., \vec{c}_n)$. Dann ist $\tilde{S} \in SO_n$ die gesuchte diagonalisierende Matrix. □

Beispiel 7.23. *Sei* $A := \begin{pmatrix} -1 & 1 & 0 \\ 1 & -2 & 1 \\ 0 & 1 & -1 \end{pmatrix}$. *Es ist* $A^\top = A$.

- $\chi_A(X) = \begin{vmatrix} X+1 & -1 & 0 \\ -1 & X+2 & -1 \\ 0 & -1 & X+1 \end{vmatrix} = \begin{vmatrix} X+1 & -1 & -X-1 \\ -1 & X+2 & 0 \\ 0 & -1 & X+1 \end{vmatrix} = \begin{vmatrix} X+1 & -2 & 0 \\ -1 & X+2 & 0 \\ 0 & -1 & X+1 \end{vmatrix} =$
  $(X+1)\begin{vmatrix} X+1 & -2 \\ -1 & X+2 \end{vmatrix} = (X+1)\cdot(X^2+3\,X) = X(X+1)(X+3)$. *Also Eigenwerte* 0, -1, -3.

- $\lambda = 0$*:* $\begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{pmatrix}$, *also* $\mathrm{LR}(-A;\, \vec{0}) = \mathrm{Span}(\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix})$.

  $\lambda = -1$*:* $\begin{pmatrix} 0 & -1 & 0 \\ -1 & 1 & -1 \\ 0 & -1 & 0 \end{pmatrix} \rightsquigarrow \begin{pmatrix} -1 & 1 & -1 \\ 0 & -1 & 0 \\ 0 & -1 & 0 \end{pmatrix}$, *also* $\mathrm{LR}(-\mathbb{1}_3 - A;\, \vec{0}) = \mathrm{Span}(\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix})$

  $\lambda = -3$*:* $\begin{pmatrix} -2 & -1 & 0 \\ -1 & -1 & -1 \\ 0 & -1 & -2 \end{pmatrix} \rightsquigarrow \begin{pmatrix} -1 & -1 & -1 \\ -2 & -1 & 0 \\ 0 & -1 & -2 \end{pmatrix} \rightsquigarrow \begin{pmatrix} -1 & -1 & -1 \\ 0 & 1 & 2 \\ 0 & -1 & -2 \end{pmatrix}$, *und man liest ab:*
  $\mathrm{LR}(-3 \cdot \mathbb{1}_3 - A;\, \vec{0}) = \mathrm{Span}(\begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix})$.

- *Gram-Schmidt ist bei einem eindimensionalen Eigenraum unnötig, wir müssen lediglich normieren. Eine ONB ist also* $[\frac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{2}}\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{6}}\begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}]$.

- *Sei* $\tilde{S} := \begin{pmatrix} \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & -\frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \end{pmatrix}$. *Wir finden* $\det \tilde{S} = 1$. *Also setzen wir* $S := \tilde{S}$.
  Wer will, kann jetzt noch $S^\top AS = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -3 \end{pmatrix}$ *verifizieren.*

Bemerkung 7.24. *Folgendes sollten Sie bei HAT bzw. Diagonalisierung beachten:*

a) Ist nach einer HAT oder nach einer Diagonalisierung gefragt?

b) Findet man bei einer HAT einen nicht-reellen Eigenwert, so liegt ein Rechenfehler vor.

c) Findet man bei einer HAT einen Eigenwert, dessen geometrische Vielfachheit kleiner als seine algebraische Vielfachheit ist, so liegt ein Rechenfehler vor.

d) Generell kann eine Basis weder den Nullvektor noch einen Vektor $\vec{v}$ zweimal bzw. sein Negatives enthalten. Dasselbe gilt auch für die Spalten einer diagonalisierenden Matrix.

e) Wenn nicht nach einer HAT gefragt ist, so wäre die Anwendung des Gram-Schmidt-Verfahrens mindestens unnötig und fehleranfällig, eventuell sogar grundsätzlich falsch: Ist nämlich A nicht symmetrisch, dann kann eine diagonalisierende Matrix für A nicht orthogonal sein.

f) Das Gram-Schmidt-Verfahren sollte man bei einer HAT in jedem Eigenraum separat durchführen. Es wäre zwar nicht grundsätzlich falsch, es auf eine Basis aus Eigenvektoren des $\mathbb{R}^n$ anzuwenden, denn Eigenräume zu verschieden Eigenwerten sind ohnehin orthogonal, ein Rechenfehler bei Gram-Schmidt kann dies zunichte machen.

g) Findet man $\det(\tilde{S}) \notin \{\pm 1\}$ bei einer HAT, so liegt ein Rechenfehler vor.

7.5 Kriterien für positive Definitheit

Beobachtung 7.25 (und Definition)
$A \in M_n(\mathbb{R})$ heißt positiv definit:$\Leftrightarrow \forall \vec{v} \in \mathbb{R}^n \setminus \{\vec{0}\}: \ \vec{v}^\top \cdot A \cdot \vec{v} > 0$. Eine Bilinearform ist genau dann positiv definit, wenn ihre Darstellungsmatrizen positiv definit sind.

Wir zeigen nun, wie man die positive Definitheit einer **symmetrischen** reellen Matrix prüfen kann.

Lemma 7.26
Sei V ein endlichdimensionaler $\mathbb{R}$-Vektorraum mit Basen $C = [\vec{c}_1, \ldots, \vec{c}_n]$ und $D = [\vec{d}_1, \ldots, \vec{d}_n]$. Sei b eine Bilinearform auf V. Dann gilt ${}_Db = {}^C_D\mathbb{T}^\top \, {}_Cb \, {}^C_D\mathbb{T}$.

Beweis:
${}_Db_{i,j} = b(\vec{d}_j, \vec{d}_i) = {}^C\vec{d}_i^{\,\top} \cdot {}_Cb \cdot {}^C\vec{d}_j = \left({}^C_D\mathbb{T} \cdot \vec{e}_i\right)^\top \cdot {}_Cb \cdot {}^C_D\mathbb{T} \cdot \vec{e}_j = \vec{e}_i^\top \left({}^C_D\mathbb{T}^\top \, {}_Cb \, {}^C_D\mathbb{T}\right) \vec{e}_j = \left({}^C_D\mathbb{T}^\top \, {}_Cb \, {}^C_D\mathbb{T}\right)_{i,j}$. □

Lemma 7.27
Sei $A \in M_n(\mathbb{R})$ symmetrisch und $S \in GL_n(\mathbb{R})$, so dass $S^\top AS$ diagonal ist. A ist genau dann positiv definit, wenn alle Diagonaleinträge von $S^\top AS$ positiv sind.

Beweis:
Es seien $\mu_1, ..., \mu_n \in$ die Diagonaleinträge von $S^\top AS$. Für $\vec{v} \in \mathbb{R}^n \setminus \{\vec{0}\}$ sei $\vec{w} := S^{-1} \cdot \vec{v}$. Dann gilt $\vec{v}^\top A\vec{v} = \vec{w}^\top S^\top AS\vec{w} = \sum_{i=1}^{n} \mu_i w_i^2$. □

Korollar 7.28
$A \in M_n(\mathbb{R})$ mit $A^\top = A$ ist positiv definit $\Longleftrightarrow$ alle Eigenwerte von A sind positiv.

Beweis:
Hauptachsentransformation. □

Problem: Wir haben die Existenz der Hauptachsentransformation noch nicht bewiesen und Eigenwertberechnung ist schwer. Wir suchen einfachere Kriterien.

Notation 7.29
Sei $A \in M_n(\mathbb{K})$ und $k \in \{1, ..., n\}$. Mit $A_{\leq k} \in M_k(\mathbb{K})$ bezeichnen wir die aus den ersten k Zeilen der ersten k Spalten von A gebildeten Matrix und $\det(A_k)$ *nennt man eine* **führende Hauptminore** *von A.*

Problem 7.30
Sei $A \in M_n(\mathbb{R})$ symmetrisch. Berechne $S \in GL_n(\mathbb{R})$, so dass $S^\top AS$ diagonal ist. Wenn zudem A positiv definit ist, so soll S eine obere Dreiecksmatrix sein.

Bemerkung Bei der Diagonalisierung aus Kapitel 6 war $S^{-1}AS$ diagonal mit Eigenwerten auf der Diagonale. Hier ist $S^\top AS$ diagonal und von Eigenwerten ist keine Rede (allerdings kann man die Vorzeichen aller Eigenwerte ablesen, was wir aber nicht beweisen). Für die Hauptachsentransformation stimmen beide Arten der Transformation überein, denn dann ist $S \in SO_n(\mathbb{R})$, also $S^\top = S^{-1}$.

Lösung:
Sei $A^{(0)} := A$ und $B^{(0)} := \mathbb{1}_n$. Ziel: Für $k \in \{1, ..., n-1\}$ konstruiere $A^{(k)}, B^{(k)} \in M_n(\mathbb{R})$ mit $A^{(k)} = B^{(k)} A \left(B^{(k)}\right)^\top$ symmetrisch, so dass $A^{(n-1)}$ diagonal ist. Dann ist $S := \left(B^{(n-1)}\right)^\top$ die Lösung. Dies gelingt mit dem im Folgenden beschriebenen ***symmetrischen Gauß-Algorithmus***.

Sei $C^{(k)} := (A^{(k)}, B^{(k)}) \in \mathbb{R}^{n \times 2n}$. Per Induktion können wir voraussetzen, dass $A^{(k)}_{\leq k}$ diagonal ist und $\forall 1 \leq i \leq k,\ i < j \leq n$ gilt $A^{(k)}_{i,j} = A^{(k)}_{j,i} = 0$ (die Verankerung ist für $k = 0$). Wir werden jeweils Zeilenoperationen und zugleich die *entsprechenden* Spaltenoperationen auf $C^{(k)}$ anwenden, um zu $C^{(k+1)}$ zu gelangen.

a) Wenn $A^{(k)}_{k+1,k+1} = 0$: Wenn es $\ell > k+1$ gibt mit $A^{(k)}_{\ell,\ell} \neq 0$, dann vertausche die Zeilen $k+1$ und ℓ sowie die Spalten $k+1$ und ℓ von $C^{(k)}$. Andernfalls: Wenn es $\ell > k+1$ gibt mit $A^{(k)}_{\ell,k+1} \neq 0$, so addiere die ℓ-te Zeile/Spalte von $C^{(k)}$ zur $(k+1)$-ten Zeile/Spalte. Andernfalls: Setze $C^{(k+1)} := C^{(k)}$ und setze mit $k \mapsto k+1$ fort.

b) Nach a) ist $a := A^{(k)}_{k+1,k+1} \neq 0$. Für $i \in \{k+2, ..., n\}$: Subtrahiere das $\frac{A^{(k)}_{i,k+1}}{a}$-Fache der $(k+1)$-ten Zeile/Spalte von $C^{(k)}$ von der i-ten Zeile/Spalte. Dadurch entsteht $C^{(k+1)} = (A^{(k+1)}, B^{(k+1)})$. Setze mit $k \mapsto k+1$ fort.

Für Matrizen X, Y, M beobachten wir: Geht Y aus X durch Zeilenoperationen hervor, so geht YM aus XM durch die gleichen Zeilenoperationen hervor, und $MY^\top$ geht aus $MX^\top$ durch die entsprechenden Spaltenoperationen hervor. Wegen $A^{(0)} = B^{(0)} A \left(B^{(0)}\right)^\top$ folgt also aus dem obigen Algorithmus: $A^{(k)} = B^{(k)} A \left(B^{(k)}\right)^\top$ für $k \in \{0, ..., n-1\}$ und $A^{(n-1)}$ ist diagonal.

Wir beobachten: Ist $B^{(k)}$ eine untere Dreiecksmatrix und tritt Fall a) nicht auf, so ist $B^{(k+1)}$ nach Schritt b) ebenfalls eine untere Dreiecksmatrix. Sei nun A positiv definit. Für $\vec{v} := \left(B^{(k)}\right)^\top \vec{e}_{k+1}$ gilt dann $0 < \vec{v}^\top A\vec{v} = \vec{e}_{k+1}^\top A^{(k)}\vec{e}_{k+1} = A^{(k)}_{k+1,k+1}$. Also tritt Fall a) nicht auf. Per Induktion folgt, dass $B^{(n-1)}$ eine untere Dreiecksmatrix und damit S eine obere Dreiecksmatrix ist. □

Wenn man sich für die Transformationsmatrix S nicht interessiert, so genügt es, obigen Algorithmus auf $A^{(k)}$ statt auf $C^{(k)}$ anzuwenden.

Beispiel 7.31

a) Sei $A := \begin{pmatrix} -1 & -2 & -3 & 2 \\ -2 & -4 & -4 & 5 \\ -3 & -4 & -7 & 4 \\ 2 & 5 & 4 & -5 \end{pmatrix}$. $\begin{pmatrix} -1 & -2 & -3 & 2 & 1 & 0 & 0 & 0 \\ -2 & -4 & -4 & 5 & 0 & 1 & 0 & 0 \\ -3 & -4 & -7 & 4 & 0 & 0 & 1 & 0 \\ 2 & 5 & 4 & -5 & 0 & 0 & 0 & 1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} -1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 1 & -2 & 1 & 0 & 0 \\ 0 & 2 & 2 & -2 & -3 & 0 & 1 & 0 \\ 0 & 1 & -2 & -1 & 2 & 0 & 0 & 1 \end{pmatrix} \rightsquigarrow$
$\begin{pmatrix} -1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & -1 & -2 & 1 & 2 & 0 & 0 & 1 \\ 0 & -2 & 2 & 2 & -3 & 0 & 1 & 0 \\ 0 & 1 & 2 & 0 & -2 & 1 & 0 & 0 \end{pmatrix} \rightsquigarrow \begin{pmatrix} -1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 2 & 0 & 0 & 1 \\ 0 & 0 & 6 & 0 & -7 & 0 & 1 & -2 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}$. *Mit* $S := \begin{pmatrix} 1 & 2 & -7 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & -2 & 1 \end{pmatrix}$ *folgt* $S^\top AS = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 6 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$. *Es ist also A nicht positiv definit.*

b) Sei $A := \begin{pmatrix} 1 & -1 & -1 \\ -1 & 3 & 5 \\ -1 & 5 & 17 \end{pmatrix}$. *Wir arbeiten diesmal ohne die Transformationsmatrix:* $A \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 4 \\ 0 & 4 & 16 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 8 \end{pmatrix}$, *also positiv definit.*

Als Korollar formulieren wir das rechnerisch weniger effiziente aber gleichwohl wissenswerte:

Korollar 7.32 (und Definition)
Sei $A \in M_n(\mathbb{R})$ *mit* $A = A^\top$. **Kriterium von Sylvester**[34]*: A ist genau dann positiv definit, wenn alle führenden Hauptminoren von A positiv sind, also* $\forall k \in \{1, ..., n\}\colon \det(A_{\leq k}) > 0$.

Es sei noch einmal betont, dass das Kriterium ausschließlich für symmetrische Matrizen funktioniert.

Beweis:
Wir brechen den symmetrischen Gauß-Algorithmus an geeigneter Stelle ab und erhalten dadurch eine obere Dreiecksmatrix $S \in GL_n(\mathbb{R})$ und ein $\ell \in \{1, ..., n\}$, so dass mit $A' := S^\top AS$ gilt: $A'_{\leq \ell}$ ist diagonal und $\forall 1 \leq k < \ell\colon\ A'_{k,k} > 0$. Zudem: A ist genau dann *nicht* positiv definit, wenn $A'_{\ell,\ell} \leq 0$.

Wegen der oberen Dreiecksgestalt von S gilt $(S^\top AS)_{\leq k} = (S_{\leq k})^\top A_{\leq k} S_{\leq k}$ für alle $k \leq n$. Aus $S \in GL_n(\mathbb{R})$ folgt zudem $\det(S_{\leq k}) \in \mathbb{R}^*$.

Wegen Diagonalgestalt gilt: $\forall 1 \leq k < \ell\colon\ \det(A'_{\leq k}) > 0$ und $\det(A'_{\leq \ell})$ hat dasselbe Vorzeichen wie $A'_{\ell,\ell}$. Zudem haben $\det(A'_{\leq k})$ und $\det(A_{\leq k})$ dasselbe Vorzeichen, denn $\det(A'_{\leq k}) = \det(S^\top_{\leq k} A_{\leq k} S_{\leq k}) = \det(S_{\leq k})^2 \cdot \det(A_{\leq k})$. □

[34] James Joseph Sylvester [1814–1897]; auch ***Hurwitz-Kriterium*** oder ***Determinantenkriterium*** genannt.

Lemma 7.33 (und Definition)
$M \in M_n(\mathbb{R})$ heißt **schiefsymmetrisch** $:\Leftrightarrow M^\top = -M$.
In diesem Fall gilt $\forall \vec{v} \in \mathbb{R}^n\colon\ \vec{v}^\top M \vec{v} = 0$.

Beweis:
$\vec{v}^\top M \vec{v} = -\vec{v}^\top M^\top \vec{v} = -(M\vec{v})^\top \vec{v} = -\left((M\vec{v})^\top \vec{v}\right)^\top = -\vec{v}^\top M \vec{v}$. Also $2 \cdot \vec{v}^\top M \vec{v} = 0$, und darauf folgt wegen $2 \neq 0$ die Behauptung. □

Korollar 7.34 (Definitheitstest für beliebige Matrizen)
$A \in M_n(\mathbb{R})$ ist genau dann positiv definit, wenn $A + A^\top$ positiv definit ist. Da $A + A^\top$ symmetrisch ist, kann man mit dem symmetrischen Gauß-Algorithmus oder dem Sylvester-Kriterium die Definitheit von $A + A^\top$ und damit die von A effektiv prüfen.

Beweis:
Sei $A_+ := \frac{1}{2}(A + A^\top)$ und $A_- := \frac{1}{2}(A - A^\top)$. Dann ist $A = A_+ + A_-$. Weil A_- schiefsymmetrisch ist, gilt $\forall \vec{v} \in \mathbb{R}^n\colon\ \vec{v}^\top A \vec{v} = \vec{v}^\top A_+ \vec{v} + \vec{v}^\top A_- \vec{v} = \frac{1}{2}\vec{v}^\top (A + A^\top)\vec{v}$.□

7.6 Der Spektralsatz

Es sei $(V, \langle|\rangle)$ ein Euklidischer Raum. Wir beweisen in diesem Abschnitt, dass selbstadjungierte Endomorphismen von V diagonalisierbar sind. Unser Plan: Wir beweisen, dass alle Eigenwerte selbstadjungierter Endomorphismen reell sind. Nach dem Hauptsatz der Algebra (der erst in späteren Semestern bewiesen wird) gibt es einen reellen Eigenwert. Dann zeigen wir, dass man den selbstadjungierten Endomorphismus auch als Endomorphismus des orthogonalen Komplements des zugehörigen Eigenraums auffassen kann — dadurch ist Induktion nach $\dim(V)$ möglich. Dazu brauchen wir allerdings noch einen Begriff, der schon im Zusammenhang mit allgemeinen Abbildungen hätte definiert werden können.

Definition 7.35 (und Beobachtung)
Sei $f\colon\ A \to B$ eine Abbildung und $C \subset A$. Die **Einschränkung** *von f auf C ist $f\big|_C\colon\ C \to B$ mit $f\big|_C(x) := f(x)$.*

Lemma 7.36 (und Definition)
Für einen Endomorphismus $\varphi\colon\ V \to V$ sei $\sigma(\varphi) := \{\lambda \in \mathbb{C} \mid \chi_\varphi(\lambda) = 0\}$ das **Spektrum** *von φ; das ist also die Menge aller komplexen Eigenwerte von φ. Wenn φ selbstadjungiert ist, dann $\sigma(\varphi) \subset \mathbb{R}$.*

Beweis:
Erinnerung: χ_φ ist definiert als das charakteristische Polynom einer Darstellungsmatrix $A := {}_C^C\varphi \in M_n(\mathbb{R})$, was von der Wahl der Basis C unabhängig ist. Wir wählen C als ONB von V. Weil wir eine ONB wählten, können wir mit dem Standardskalarprodukt rechnen, d.h. $\forall \vec{u}, \vec{v} \in V\colon\ \langle \vec{u}|\vec{v}\rangle = {}^C\vec{v}^\top \cdot {}^C\vec{u}$. Weil φ selbstadjungiert ist, folgt $A = A^\top$ wegen Lemma 7.18.

Zwar sind alle Einträge von A reell, aber da Eigenwerte komplex sein können, betrachten wir den durch A gegebenen Endomorphismum von $\mathbb{C}$-Vektorräumen $f\colon \mathbb{C}^n \to \mathbb{C}^n$, wobei $\forall \vec{z} \in \mathbb{C}^n\colon f(\vec{z}) := A \cdot \vec{z}$. Wir erinnern an die komplexe Konjugation: Wenn wir komplexe Konjugation auf eine Matrix oder einen Vektor anwenden, ist damit die komplexe Konjugation jedes Eintrags gemeint. Insbesondere gilt $\bar{A} = A$, denn die Einträge sind reell, ändern sich also durch Konjugation nicht. Wir im Kapitel über komplexe Zahlen gesehen, ist die Konjugation einer Summe oder eines Produkts gleich der Summe bzw. dem Produkt der konjugierten Summanden bzw. Faktoren. Das überträgt sich auch auf Matrixprodukte. Ist $\vec{0} \neq \vec{z} \in \mathbb{C}^n$, dann ist $\vec{z}^\top \cdot \overline{\vec{z}} = \sum_{i=1}^{n} |z_i|^2 \in \mathbb{R}_{>0}$.

Ist $\lambda \in \sigma(\varphi)$, dann ist λ ein Eigenwert von A und es sei $\vec{z} \in \mathbb{C}^n$ ein Eigenvektor. Es ist also $A\vec{z} = \lambda\vec{z}$. Durch Konjugation folgt $\bar{A}\overline{\vec{z}} = \overline{A\vec{z}} = \overline{\lambda\vec{z}} = \bar{\lambda}\overline{\vec{z}}$. Daher ist $\overline{\vec{z}} \in \mathbb{C}^n$ ein Eigenvektor von A zum Eigenwert $\bar{\lambda}$.

Wegen $A = A^\top$ gilt $(A\vec{z})^\top\overline{\vec{z}} = \vec{z}^\top(A\overline{\vec{z}})$ und daher $\lambda\vec{z}^\top\overline{\vec{z}} = \bar{\lambda}\vec{z}^\top\overline{\vec{z}}$. Wegen $\vec{z}^\top\overline{\vec{z}} \neq 0$ können wir kürzen und es folgt $\lambda = \bar{\lambda}$. Das bedeutet aber $\lambda \in \mathbb{R}$. □

Definition 7.37. *Sei $\varphi\colon V \to V$ ein Endomorphismus.*
Ein Untervektorraum $U \leq V$ heißt **φ-invariant** $:\Leftrightarrow \varphi(U) \subset U$.

Lemma 7.38
Sei $\varphi\colon V \to V$ ein selbstadjungierter Endomorphismus. Wenn $U \leq V$ φ-invariant ist, dann ist auch $U^\perp$ φ-invariant.

Beweis:
$\vec{v} \in U^\perp \Rightarrow \forall \vec{u} \in U\colon \langle\vec{u}|\vec{v}\rangle = 0 \Rightarrow \forall \vec{u} \in U\colon \langle\vec{u} \mid \varphi(\vec{v})\rangle \overset{\text{selbstadj.}}{=} \langle\underbrace{\varphi(\vec{u})}_{\in U} \mid \vec{v}\rangle = 0$. □

Spektralsatz
Ist V ein euklidischer Raum und $\varphi\colon V \to V$ ein selbstadjungierter Endomorphismus, dann hat V eine ONB aus Eigenvektoren von φ.

Wie man diese ONB konkret berechnet, wissen Sie bereits.

Beweis:
Induktion nach $\dim(V)$. Verankerung für $\dim(V) = 1$: Sei $\vec{0} \neq \vec{u} \in V$ beliebig, dann ist $\varphi(\vec{u}) \in V = \mathrm{Span}(\vec{u})$, d.h. $\exists\lambda \in \mathbb{R}\colon \varphi(\vec{u}) = \lambda\vec{u}$, d.h. $\vec{u}$ ist ein Eigenvektor, und $[\frac{1}{\|\vec{u}\|}\vec{u}]$ ist eine ONB aus Eigenvektoren.

Sei nun $n := \dim V > 1$. Nach dem Hauptsatz der Algebra hat $\chi_A(X)$ mindestens eine Nullstelle in $\mathbb{C}$, d.h. φ hat einen Eigenwert $\lambda \in \mathbb{C}$. Nach Lemma 7.36 gilt sogar $\lambda \in \mathbb{R}$. Sei $U := E_\lambda(\varphi) \leq V$; weil λ ein Eigenwert ist, gilt $\dim(U) \geq 1$. Wegen $V = U \oplus U^\perp$ folgt $n = \dim(V) = \dim(U) + \dim(U^\perp)$, daher $r := \dim(U^\perp) < n$. Durch das auf V definierte Skalarprodukt wird auch $U^\perp$ zu einem euklidischen Raum.

Nach Lemma 7.38 ist $\varphi(U^\perp) \subseteq U^\perp$, d.h. $\varphi' := \varphi\big|_{U^\perp} : U^\perp \to U^\perp$ ist ein Endomorphismus von $U^\perp$. Nach Induktionsannahme hat $U^\perp$ eine aus Eigenvektoren von φ bestehende ONB $[\vec{v}_1, ..., \vec{v}_r]$. Weiter besitzt U eine ONB $[\vec{v}_{r+1}, ..., \vec{v}_n]$, und das sind alles Eigenvektoren von φ, weil U ein einzelner Eigenraum ist. Weil U und $U^\perp$ orthogonal sind, ist $[\vec{v}_1, ..., \vec{v}_n]$ die gesuchte ONB von V. □

7.7 SO_2 und SO_3

Vorbemerkung In diesem Abschnitt verwenden wir die Additionstheoreme für Sinus und Kosinus. Leider sind diese Ergebnisse nicht länger im Lehrplan — obwohl man sie braucht, um die Ableitungen von $\sin(x)$ und $\cos(x)$ bestimmen zu können, die weiterhin im Lehrplan sind. Ein Beweis steht in Abschnitt 7.7.1.

Additionstheoreme für Sinus und Kosinus

$$\forall \varphi, \vartheta \in \mathbb{R}\colon\ \sin(\vartheta+\varphi) = \sin(\vartheta)\cos(\varphi) + \cos(\vartheta)\sin(\varphi) \quad \textit{und}$$
$$\cos(\vartheta+\varphi) = \cos(\vartheta)\cos(\varphi) - \sin(\vartheta)\sin(\varphi)\,.$$

Lemma 7.39

a) $SO_2 = \left\{ \left(\begin{smallmatrix} \cos\vartheta & -\sin\vartheta \\ \sin\vartheta & \cos\vartheta \end{smallmatrix}\right) \mid \vartheta \in \mathbb{R} \right\}$. *Solche Matrizen nennt man* **Drehmatrizen**; *beschrieben wird dabei die Drehung um* $\vec{0}$ *im Winkel* ϑ *gegen den Uhrzeigersinn. Ist* $\sin\vartheta \neq 0$,*so gibt es keinen reellen Eigenwert.*

b) $O_2 \setminus SO_2 = \left\{ \left(\begin{smallmatrix} \cos\vartheta & \sin\vartheta \\ \sin\vartheta & -\cos\vartheta \end{smallmatrix}\right) \,\middle|\, \vartheta \in \mathbb{R} \right\}$. *Enspricht einer Spiegelung in der Gerade durch* $\vec{0}$ *mit Richtung* $(\cos\frac{\vartheta}{2}, \sin\frac{\vartheta}{2})$. *Eigenwerte* 1 *und* -1.

Beweis:

a) Die Matrix $\left(\begin{smallmatrix} \cos\vartheta & -\sin\vartheta \\ \sin\vartheta & \cos\vartheta \end{smallmatrix}\right)$ ist orthogonal mit Determinante 1. Für $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in SO_2$, ist $\left(\begin{smallmatrix} a & c \\ b & d \end{smallmatrix}\right) \cdot \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$ und $ad - bc = 1$, also

$$a^2 + c^2 = 1 \qquad b^2 + d^2 = 1 \qquad ab + cd = 0 \qquad ad - bc = 1$$

Aufgrund der ersten beiden Gleichungen gibt es Zahlen ϑ, φ mit $a = \cos\vartheta$, $c = \sin\vartheta$, $b = \cos\varphi$, $d = \sin\varphi$. Wegen $ad - bc = 1$ ist $\sin(\varphi - \vartheta) = 1$, also oBdA $\varphi = \vartheta + \frac{\pi}{2}$. Deshalb ist $b = -\sin\vartheta$, $d = \cos\vartheta$.

Das charakteristische Polynom ist $X^2 - 2\cos\vartheta X + 1 = (X - \cos\vartheta)^2 + \sin^2\vartheta$. Für $\sin\vartheta \neq 0$ hat das Polynom keine Nullstellen in $\mathbb{R}$.

b) Aus $A \in O_2 \setminus SO_2$ folgt $\det(A) = -1$. Sei $B = A \cdot \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$. Dann $A = B \cdot \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$, und $B \in SO_2$. Nach a) ist $A = \left(\begin{smallmatrix} \cos\vartheta & \sin\vartheta \\ \sin\vartheta & -\cos\vartheta \end{smallmatrix}\right)$. Dann ist $\begin{pmatrix} \cos\frac{\vartheta}{2} \\ \sin\frac{\vartheta}{2} \end{pmatrix}$ ein Eigenvektor mit Eigenwert 1:

$$\left(\begin{smallmatrix} \cos\vartheta & \sin\vartheta \\ \sin\vartheta & -\cos\vartheta \end{smallmatrix}\right) \cdot \begin{pmatrix} \cos\frac{\vartheta}{2} \\ \sin\frac{\vartheta}{2} \end{pmatrix} = \begin{pmatrix} \cos\left(\frac{\vartheta}{2}\right)\cos(\vartheta) + \sin\left(\frac{\vartheta}{2}\right)\sin(\vartheta) \\ -\cos(\vartheta)\sin\left(\frac{\vartheta}{2}\right) + \cos\left(\frac{\vartheta}{2}\right)\sin(\vartheta) \end{pmatrix}$$
$$= \begin{pmatrix} \cos\left(\frac{\vartheta}{2}\right)\cos(-\vartheta) - \sin\left(\frac{\vartheta}{2}\right)\sin(-\vartheta) \\ -\cos(-\vartheta)\sin\left(\frac{\vartheta}{2}\right) - \cos\left(\frac{\vartheta}{2}\right)\sin(-\vartheta) \end{pmatrix} = \begin{pmatrix} \cos\left(-\frac{\vartheta}{2}\right) \\ -\sin\left(-\frac{\vartheta}{2}\right) \end{pmatrix} = \begin{pmatrix} \cos\frac{\vartheta}{2} \\ \sin\frac{\vartheta}{2} \end{pmatrix}$$

Ebenso ist $\begin{pmatrix} -\sin\frac{\vartheta}{2} \\ \cos\frac{\vartheta}{2} \end{pmatrix}$ ein Eigenvektor mit Eigenwert -1. Man erkennt: A stellt eine Spiegelung an der Gerade $E_1(A)$ dar. □

Lemma 7.40. *Sei $A \in O_3$, $U_\pm := E_{\pm 1}(A) \leq \mathbb{R}^3$.*

a) Wenn $\det(A) = 1$ *und* $A \neq \mathbb{1}_3$, *dann* $\dim(U_+) = 1$.
A stellt eine **Drehung** *um die Achse U_+ dar.*

b) Wenn $\det(A) = -1$, *dann gilt eine der folgenden Aussagen:*

i) $A = -\mathbb{1}_3$ stellt die **Punktspiegelung** *in $\vec{0}$ dar.*

ii) $\dim(U_-) = 1$ *und A stellt die* **Drehspiegelung** *mit Achse U_- dar, also eine Drehung um U_- im Winkel ϑ gefolgt von der Spiegelung an der Ebene ${U_-}^\perp$.*
Im Spezialfall $\vartheta = 0$ stellt A die **Spiegelung** *an ${U_-}^\perp$ dar.*

Beweis:
$\deg(\chi_A(X)) = 3$ ist ungerade, also gibt es $\lambda \in \mathbb{R}$ mit $\chi_A(\lambda) = 0$. Wegen $A \in O_3$ gilt $\forall \vec{v} \in \mathbb{R}^3\colon\ \|A\vec{v}\| = \|\vec{v}\|$, dies gilt insbesondere, wenn $\vec{v}$ ein Eigenvektor zum Eigenwert λ ist. Daher $|\lambda| = 1$. Wenn möglich, so wähle $\lambda = 1$. Sei $[\vec{b}_1, \vec{b}_2, \vec{b}_3]$ eine ONB von $\mathbb{R}^3$ mit $\vec{b}_1$ ein Eigenvektor zum Eigenwert λ.

a) Setze $U = \vec{b}_1^{\perp}$. Für $\vec{u} \in U$ ist wegen A orthogonal

$$0 = \langle \vec{b}_1 \mid \vec{u} \rangle = \langle A\vec{b}_1 \mid A\vec{u} \rangle = \lambda \langle \vec{b}_1 \mid A\vec{u} \rangle .$$

Wegen $\lambda = \pm 1$ ist $\langle \vec{b}_1 | A\vec{u} \rangle = 0$, also $A\vec{u} \in U = \vec{b}_1^{\perp} = \mathrm{Span}(\vec{b}_2, \vec{b}_3)$, und $F := L_A\big|_U$ ist ein orthogonaler Endormorphismus von U. Bezüglich dieser Basis hat L_A die Matrix $\left(\begin{smallmatrix} \lambda & 0 \\ 0 & C \end{smallmatrix}\right)$, mit der Abbildungsmatrix C von F. Also $1 = \det A = \lambda \det(C)$, wegen $\lambda^2 = 1$ also $\det(F) = \lambda$.

Ist $\lambda = 1$, dann auch $\det(F) = 1$. Wegen $A \neq \mathbb{1}_3$ ist $F \neq \mathrm{Id}$ und nach Lemma 7.39.a) ist A eine echte Drehung der Ebene U um Achse $\mathrm{Span}(\vec{b}_1)$. Insbesondere $U_+ = \mathrm{Span}(\vec{b}_1)$, alle anderen Eigenräume haben Dimension 0.

$\lambda = -1$ ist unmöglich, denn sonst wäre $\det(F) = -1$ und nach Lemma 7.39.b) hätte F und damit auch A den Eigenwert $+1$ — wir hätten also oben bereits $\lambda = +1$ gewählt.

b) Hier ist $-A \in SO_3$. Wenn $A \neq -\mathbb{1}_3$, dann ist $U = E_1(-A) = E_{-1}(A)$ eindimensional und $-A$ ist eine Drehung um die Achse U. Somit ist A eine Drehspiegelung mit Achse U, bzw. eine Spiegelung, falls der Drehwinkel Null ist. □

Wir fassen zusammen:

Problem 7.41 (Klassifikation orthogonaler Endomorphismen von $\mathbb{R}^3$)
Sei $f\colon \mathbb{R}^3 \to \mathbb{R}^3$ ein Endomorphismus mit Abbildungsmatrix $F \in M_3(\mathbb{R})$ bezüglich der Standardbasis. Prüfe, ob f orthogonal bzgl. des Standardskalarprodukts ist, bestimme den Typ von f sowie ggf. Drehachse und den Betrag des Drehwinkels.

Lösung:
f ist orthogonal $\iff F^\top F = \mathbb{1}_3$. Wir setzen ab jetzt $F^\top F = \mathbb{1}_3$ voraus.

- $F = \mathbb{1}_3 \Rightarrow f = \mathrm{Id}_{\mathbb{R}^3}$.
- $F = -\mathbb{1}_3 \Rightarrow f$ ist Punktspiegelung.
- $\det(F) = 1 \Rightarrow f$ ist eine Drehung mit der eindimensionalen Achse $E_1(F)$. Der Drehwinkel φ hat den Betrag $\arccos\big(\frac{1}{2}(\mathrm{Spur}(F) - 1)\big)$: Ergänzt man nämlich einen normierten Eigenvektor mit Eigenwert 1 zu einer ONB D von $\mathbb{R}^3$, dann ${}^D_D f = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi & -\sin\varphi \\ 0 & \sin\varphi & \cos\varphi \end{pmatrix}$. Beim Basiswechsel ändert sich die Spur nicht, d.h. es gilt $\mathrm{Spur}({}^D_D f) = \mathrm{Spur}(F)$, und die Spurformel trifft für ${}^D_D f$ offenbar zu.
- $\det(F) = -1 \Rightarrow f$ ist eine Drehspiegelung mit der eindimensionalen Achse $E_{-1}(F)$. Der Drehwinkel φ hat den Betrag $\arccos\big(\frac{1}{2}(\mathrm{Spur}(F) + 1)\big)$: Ergänzt man nämlich einen normierten Eigenvektor mit Eigenwert -1 zu einer ONB D von $\mathbb{R}^3$, dann ${}^D_D f = \begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos\varphi & -\sin\varphi \\ 0 & \sin\varphi & \cos\varphi \end{pmatrix}$. Wieder $\mathrm{Spur}({}^D_D f) = \mathrm{Spur}(F)$ und die Spurformel trifft für ${}^D_D f$ zu. Spiegelebene: $(E_{-1}(F))^\perp$. □

Beispiel 7.42
Es sei $f\colon \mathbb{R}^3 \to \mathbb{R}^3$ eine lineare Abbildung mit der Abbildungsmatrix $F = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{1}{3} & -\frac{2}{3} \\ -\frac{2}{3} & \frac{2}{3} & -\frac{1}{3} \end{pmatrix}$ bezüglich der Standardbasis. Man prüft $F^\top F = \mathbb{1}_3$.

- $\det(F) = 1$. *Wegen $F \neq \mathbb{1}_3$ ist* $\dim(E_1(F)) = 1$*: $E_1(F) = \mathrm{LR}(\mathbb{1}_3 - F; \vec{0}) = \mathrm{Span}(\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix})$. Es handelt sich um eine Drehung um die Achse* $\mathrm{Span}(\begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{pmatrix})$.
- $\mathrm{Spur}(F) = 1/3$. *Der Drehwinkel ist* $\psi = \pm\arccos\big(\frac{1}{2}(\frac{1}{3} - 1)\big) = \arccos(-1/3) \approx 1.910633$ *oder circa* $109.47°$*; dies ist der so genannte* **Tetraederwinkel**.

Beispiel 7.43
Es sei $f\colon \mathbb{R}^3 \to \mathbb{R}^3$ eine lineare Abbildung mit der Abbildungsmatrix $F = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} & -\frac{2}{3} \\ \frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ -\frac{2}{3} & \frac{2}{3} & \frac{1}{3} \end{pmatrix}$ bezüglich der Standardbasis. Man prüft $F^\top F = \mathbb{1}_3$.

- $\det(F) = -1$. *Wegen $F \neq -\mathbb{1}_3$ ist* $\dim(E_{-1}(F)) = 1$*: $E_{-1}(F) = \mathrm{LR}(-\mathbb{1}_3 - F; \vec{0}) = \mathrm{Span}(\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix})$.*
- *Wegen* $\mathrm{Spur}(F) = 1$ *ist der Drehwinkel Null, es handelt sich also um die Spiegelung an der Ebene $E_1(F) = \mathrm{Span}(\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix})$.*

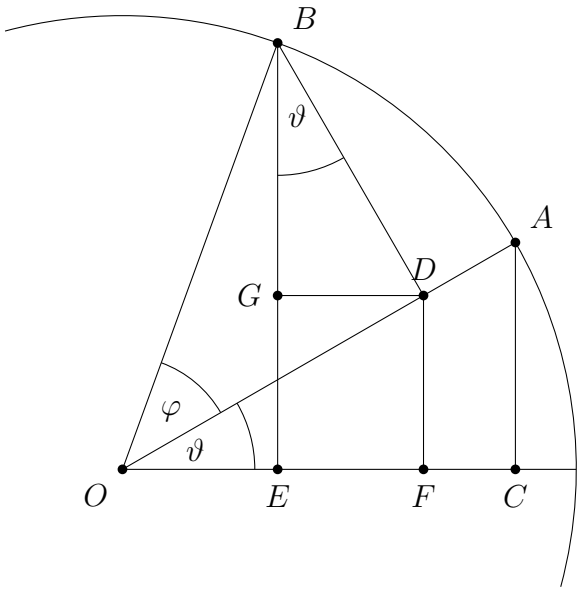


Abbildung 1: Beweis der Additionstheoreme

7.7.1 Additionstheoreme für Sinus und Kosinus

Beweis der Additionstheoreme:
Wir behandeln nur den Fall, wo ϑ, φ und $\vartheta + \varphi$ alle spitz sind. In Abbildung 1 sind C, D, E so gewählt, dass $\widehat{ACO}$, $\widehat{BDO}$ und $\widehat{BEO}$ rechte Winkel sind. Folglich gilt $\widehat{EBD} = \vartheta$. Ferner sind F, G so gewählt, dass $DGEF$ ein Rechteck ist. Für den Einheitskreis gelten

$$|OD| = \cos(\varphi) \quad |BD| = \sin(\varphi) \quad |OE| = \cos(\vartheta + \varphi) \quad |BE| = \sin(\vartheta + \varphi)\,.$$

Außerdem gelten

$$|BG| = |BD|\cos(\vartheta) = \cos(\vartheta)\sin(\varphi) \quad |GD| = |BD|\sin(\vartheta) = \sin(\vartheta)\sin(\varphi)$$
$$|OF| = |OD|\cos(\vartheta) = \cos(\vartheta)\cos(\varphi) \quad |DF| = |OD|\sin(\vartheta) = \sin(\vartheta)\cos(\varphi)\,.$$

Also

$$\sin(\vartheta+\varphi) = |BE| = |BG|+|GE| = |BG|+|DF| = \cos(\vartheta)\sin(\varphi)+\sin(\vartheta)\cos(\varphi)\,,$$

und

$$\cos(\vartheta + \varphi) = |OE| = |OF| - |EF| = |OF| - |GD| = \cos(\vartheta)\cos(\varphi) - \sin(\vartheta)\sin(\varphi)\,.$$

□

Der Rest dieses Abschnitts wurde in der Vorlesung nicht behandelt

Lemma 7.44
Für Bogenmaß gilt:

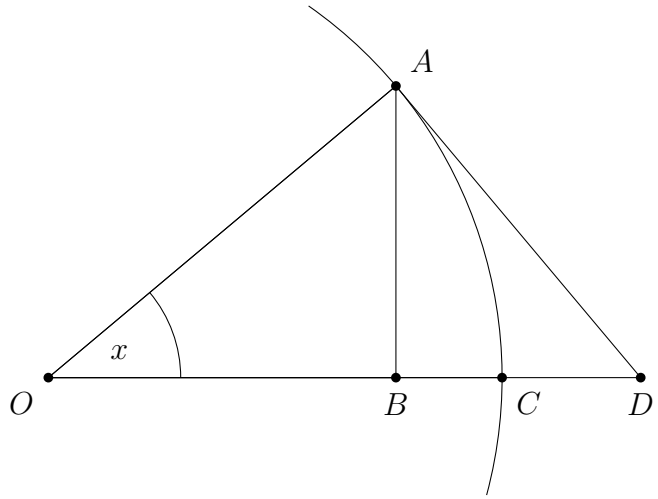


Abbildung 2: Beweis der Ableitungsregeln für Sinus und Kosinus

a) $\lim\limits_{x\to 0} \dfrac{\sin(x)}{x} = 1.$ *b)* $\lim\limits_{x\to 0} \dfrac{1-\cos(x)}{x} = 0.$

c) $\sin'(x) = \cos(x)$ *und* $\cos'(x) = -\sin(x).$

Beweis:

a): Zuerst betrachten wir den Fall $x > 0$. Da es um den Grenzwert $x \to 0^+$ geht, dürfen wir annehmen, dass x spitz ist. In Abbildung 2 ist B der Fußpunkt des Lots von A auf der x-Achse und D ist der Punkt, wo die Tangente in A die x-Achse schneidet. Somit sind $\widehat{ABO}$ und $\widehat{OAD}$ rechte Winkel. Für den Einheitskreis gelten $|OA| = 1$, $|AB| = \sin(x)$ und $|\text{Kreisbogen } AC| \underset{\text{Bogenmaß}}{=} x$. Ferner gilt

$$|AB| = |OA| \tan(x) = \tan(x)\,.$$

Nun, die Lotstrecke AB, der Kreisbogen AC und die Strecke AD sind drei verschiedene Wege von A zur x-Achse; und aus der Zeichnung erkennt man, dass

$$|AB| < |\text{Kreisbogen } AC| < |AD|\,,$$

das heißt $\sin(x) < x < \tan(x)$,also $\frac{1}{\tan(x)} < x < \frac{1}{\sin(x)}$; multipliziert mit $\sin(x) > 0$ ergibt sich $\cos(x) < \frac{\sin(x)}{x} < 1$. Aber $\lim\limits_{x\to 0} \cos(x) = 1$, also nach dem Einschließungssatz folgt $\lim\limits_{x\to 0^+} \dfrac{\sin(x)}{x} = 1$. Und folglich gilt $\lim\limits_{x\to 0} \dfrac{\sin(x)}{x} = 1$, denn $\dfrac{\sin(-x)}{-x} = \dfrac{\sin(x)}{x}$.

b): Sei $y = \frac{x}{2}$, dann

$$1 - \cos(x) = 1 - \cos(2y) = 1 - \cos^2(y) + \sin^2(y) \underset{\cos^2 + \sin^2 = 1}{=} 2\sin^2(y)\,.$$

Also

$$\frac{1 - \cos(x)}{x} = \frac{2\sin^2(y)}{2y} = \sin(y) \cdot \frac{\sin(y)}{y}\,.$$

Aber $\lim\limits_{y \to 0} \sin(y) = 0$, und nach a) folgt $\lim\limits_{y \to 0} \dfrac{\sin(y)}{y}$. Also

$$\lim_{x \to 0} \frac{1 - \cos(x)}{x} = \lim_{y \to 0} \sin(y) \cdot \frac{\sin(y)}{y} = 0 \cdot 1 = 0\,.$$

c):

$$\begin{aligned}
\sin'(x) &= \lim_{h \to 0} \frac{\sin(x+h) - \sin(x)}{h} \\
&= \lim_{h \to 0} \frac{\sin(x)(\cos(h) - 1) + \cos(x)\sin(h)}{h} \\
&= \cos(x) \lim_{h \to 0} \frac{\sin(h)}{h} - \sin(x) \lim_{h \to 0} \frac{1 - \cos(h)}{h} \\
&= 1 \cdot \cos(x) - 0 \cdot \sin(x) = \cos(x)
\end{aligned}$$

und

$$\begin{aligned}
\cos'(x) &= \lim_{h \to 0} \frac{\cos(x+h) - \cos(x)}{h} \\
&= \lim_{h \to 0} \frac{\cos(x)(\cos(h) - 1) - \sin(x)\sin(h)}{h} \\
&= -\cos(x) \lim_{h \to 0} \frac{1 - \cos(h)}{h} - \sin(x) \lim_{h \to 0} \frac{\sin(h)}{h} \\
&= -0 \cdot \cos(x) - 1 \cdot \sin(x) = -\sin(x)\,.
\end{aligned}$$

□

8 Eine Anwendung: Lineare Codes

Anmerkung: In der Vorlesung wurde hiervon nur eine Kurzfassung mit abweichender Nummerierung der Sätze vorgestellt.

Während es beim Chiffrieren darum geht, Daten so zu verschlüsseln, dass unbefugter Zugriff möglichst schwer ist, geht es beim Codieren um die effiziente Speicherung und Übertragung von Daten. Bekanntlich werden alle Daten im Computer als Folgen von 0 und 1 („Bit“) dargestellt. Meist sind diese gruppiert: Ein ***Wort*** entspricht genau n Bit. Offenbar entsprechen die Wörter den Vektoren aus $\mathbb{F}_2^n$; dieser Vektorraum hat 2^n Elemente, also kann man pro Wort 2^n verschiedene Daten darstellen.

Doch beim Übertragen der Wörter kann es zu Fehlern kommen: Statt einer 1 wird eine 0 oder umgekehrt übertragen.[35] Zur Vereinfachung nehmen wir an, dass pro Wort *höchstens* t Fehler auftreten. Daher wählt man eine Teilmenge $C \subset \mathbb{F}_2^n$ (den ***Code***) zu wählen, so dass für die Übertragung *nur* Wörter aus C (so genannte ***Codewörter***) verwendet werden. Erhält der Empfänger also ein Wort, welches kein Codewort ist, so muss es einen Übertragungsfehler gegeben haben. Das Ziel ist, C so zu wählen, dass t Übertagungsfehler pro Wort nicht nur erkannt, sondern auch automatisch korrigiert werden können. Ein solcher Code heißt t-***fehlerkorrigierend***.

Ein Beispiel: Mit vier Bit kann man bis zu 16 Wörter darstellen. Man könnte nun einen 8-Bit-Code konstruieren, so dass in jedem Codewort die ersten und die letzten 4 Bit übereinstimmen müssen. Es wäre also 11011101 ein Codewort, denn 1101 wird verdoppelt. Man überträgt also $n = 8$ Bit und wählt aus den 256 möglichen Wörtern 16 Codewörter aus. Wenn in den 8 Bit ein Fehler auftritt, etwa 11001101, so entsteht ein Wort, welches kein Codewort ist. Das Problem ist, dass man zwar sieht, dass ein Fehler vorliegt, man ihn aber nicht korrigieren kann: War 11011101 oder 11001100 gemeint?

Mit linearer Algebra kann man mit $n = 7$ (also einer kürzeren Wortlänge) 16 Codewörter so wählen, dass ein einzelner Übertragungsfehler in den 7 Bit nicht nur erkannt, sondern sogar korrigiert wird.

Im Folgenden sei $\mathbb{K} = \mathbb{F}_2$. Für die Analyse fehlerkorrigierender linearer Codes ist folgende Definition zentral:

Definition 8.1
Es seien $\vec{u}, \vec{v} \in \mathbb{K}^n$. Der **Hamming–Abstand**[36] *$d(\vec{u}, \vec{v}) \in \mathbb{N}$ ist die Anzahl der Komponenten, in denen $\vec{u}$ sich von $\vec{v}$ unterscheidet.*

Beispiel $d(11011100, 11011101) = 1$, $d(01000100, 11000101) = 2$.

[35] In der Realität könnten auch Bits ausgelassen oder eingefügt werden, aber derartige Fehler betrachten wir hier nicht.

[36] Richard Hamming [1915–1998]

Der Hamming–Abstand teilt wesentliche Eigenschaften mit dem euklidischen Abstand. Insbesondere gilt $\forall \vec{u}, \vec{v}, \vec{w} \in \mathbb{K}^n$

- $d(\vec{u}, \vec{v}) + d(\vec{v}, \vec{w}) \leq d(\vec{u}, \vec{w})$ (Dreiecksungleichung)
- $d(\vec{u}, \vec{v}) = d(\vec{u} - \vec{v}, \vec{0})$ (Verschiebungsinvarianz)
- $d(\vec{u}, \vec{v}) = d(\vec{v}, \vec{u})$ (Symmetrie).

Wir konzentrieren uns im Folgenden auf lineare Codes:

Definition 8.2
Ein Code $C \subset \mathbb{K}^n$ heißt **linear** *gdq. $C \leq \mathbb{K}^n$.*

Der Vorteil dieser Sichtweise ist: Man kann C durch eine Basis der Länge $d := \dim(C)$ darstellen (hat aber $|\mathbb{K}|^d$ Elemente), und jedes Codewort ist durch die d Koordinaten bezüglich dieser Basis gegeben (besteht aber aus $n \geq d$ Bit).

Beobachtung 8.3
$C \subset \mathbb{K}^n$ ist t-fehlerkorrigierend $\iff \forall \vec{v} \in C \setminus \{\vec{0}\}\colon\ d(\vec{v}, \vec{0}) \geq 2\,t + 1$.

8.0.1 Der Hamming–Code

Wir betrachten hier als Beispiel den so genannten „$(7, 4)$–Hamming–Code“. Die Zahlen geben die verwendete Wortlänge (hier $n = 7$) und $\dim(C)$ (hier 4) an. Für reale Anwendungen würde man größere Codes verwenden, etwa den $(63, 57)$–Hamming–Code. Eine Basis von C ist gegeben durch die Spalten der ***Generatormatrix*** $G = \begin{pmatrix} 1&0&0&0 \\ 0&1&0&0 \\ 0&0&1&0 \\ 0&0&0&1 \\ 0&1&1&1 \\ 1&0&1&1 \\ 1&1&1&0 \end{pmatrix}$. Durch Auflisten der $16 = 2^4$ Codewörter kann man nachprüfen, dass $d(\vec{c}, \vec{0}) \geq 3$ für alle $\vec{c} \in C \setminus \{\vec{0}\}$. Also ist C 1–fehlerkorrigierend.

Definition 8.4
Der **duale Code** *$C^\perp$ ist die Menge der Zeilenvektoren $\vec{z} \in \mathbb{K}^n$ mit $\vec{z} \cdot G = (0, 0)$.*

Offenbar ist $C^\perp$ durch ein System von $d = \dim(C)$ linearen Gleichungen mit n Unbekannten gegeben. Wegen $\mathrm{Rang}(G) = d$ folgt aus der Rangformel $\dim C^\perp = n - d$ (in unserem Beispiel also 3).

Definition 8.5
Eine Matrix $H \in \mathbb{K}^{(n-d)\times n}$, deren Zeilen eine Basis von $C^\perp$ bilden, heißt **Kontrollmatrix** *von C.*

In unserem Beispiel ist $H = \begin{pmatrix} 1&0&0&1&1&0&1 \\ 0&1&0&1&0&1&1 \\ 0&0&1&0&1&1&1 \end{pmatrix}$ eine Kontrollmatrix (das kann man als Übung nachweisen). Nach Definition gilt $H \cdot \vec{c} = \vec{0}$ für alle $\vec{c} \in C$. Darüber hinaus gilt:

Lemma 8.6
$\mathrm{LR}(H, \vec{0}) = C$, *d.h.* $\forall \vec{v} \in \mathbb{K}^n\colon\ \vec{v} \in C \iff H\vec{v} = \vec{0}$.

In diesem Sinne kann man mit der Kontrollmatrix kontrollieren, ob $\vec{v} \in C$.

Beweis:
Nach Definition von H ist $C \subseteq \mathrm{LR}(H, \vec{0})$. Da die Zeilen von H nach Definition linear unabhängig sind, gilt $\mathrm{Rang}(H) = n - d$. Also ist $\dim\left(\mathrm{LR}(H, \vec{0})\right) = n - (n - d) = d = \dim(C)$ und daher $\mathrm{LR}(H, \vec{0}) = C$. □

Definition 8.7
Das **Syndrom** *von* $\vec{v} \in \mathbb{K}^n$ *ist* $s(\vec{v}) = H \cdot \vec{v} \in \mathbb{K}^{n-d}$.

Lemma 8.8
Es sei $C \subset \mathbb{K}^n$ *ein* t*–fehlerkorrigierender linearer Code (in unserem Beispiel ist* $t = 1$*). Es seien* $\vec{v}_1, \vec{v}_2 \in \mathbb{K}^n$ *mit* $d(\vec{v}_i, \vec{0}) \leq t$ *für* $i = 1, 2$*. Wenn* $s(\vec{v}_1) = s(\vec{v}_2)$*, dann* $\vec{v}_1 = \vec{v}_2$.

Beweis:
Die Syndrome sind gleich, also $H \cdot \vec{v}_1 = H \cdot \vec{v}_2$. Es folgt $H \cdot (\vec{v}_1 - \vec{v}_2) = \vec{0}$, also $\vec{v}_1 - \vec{v}_2 \in C$ nach dem vorigen Lemma. Wäre $\vec{v}_1 \neq \vec{v}_2$, so ist nach Voraussetzung $d(\vec{v}_1, \vec{v}_2) = d(\vec{v}_1 - \vec{v}_2, \vec{0}) \geq 2 \cdot t + 1$. Nach der Dreiecksungleichung ist aber $d(\vec{v}_1, \vec{v}_2) \leq d(\vec{v}_1, \vec{0}) + d(0, \vec{v}_2) \leq 2 \cdot t$. Widerspruch, also $\vec{v}_1 = \vec{v}_2$. □

Wir können nun das fehlerkorrigierende Decodieren erklären. Es sei $\vec{c} \in C$ ein Codewort, das fehlerhaft übertragene Wort sei $\vec{w}$. Der Übertragungsfehler sei $\vec{f} := \vec{w} - \vec{c}$. Wir setzen voraus, dass es höchstens t Übertragungsfehler pro Codewort gibt. Es ist also $d(\vec{f}, \vec{0}) \leq t$. Es ist $s(\vec{w}) = H \cdot \vec{w} = H \cdot (\vec{c} + \vec{f}) = (H \cdot \vec{c}) + (H \cdot \vec{f}) = \vec{0} + s(\vec{f})$. Nach dem vorigen Lemma ist $\vec{f}$ der *einzige* mögliche Fehlervektor mit diesem Syndrom. Wir berechnen also $s(\vec{w})$, sehen dann in einer Tabelle den dazugehörigen Fehlervektor $\vec{f}$ nach, und erhalten das fehlerkorrigierte Codewort $\vec{c} = \vec{w} - \vec{f}$.

In unserem Beispiel ist dies die Tabelle mit den Syndromen der Fehlervektoren (als Bitfolgen geschrieben):

Fehlervektor	Syndrom
0000000	000
0000001	111
0000010	011
0000100	101
0001000	110
0010000	001
0100000	010
1000000	100

Wird das Wort $\vec{w} = 0010001$ empfangen, so ist $s(\vec{w}) = H \cdot \vec{w} = 110$. Der Fehlervektor ist also $\vec{f} = 0001000$. Das fehlerkorrigierte Codewort ist also $\vec{c} = \vec{w} - \vec{f} = 0011001$. Man kann überprüfen, dass dies die Summe der zweiten und dritten Spalte[37] der Generatormatrix G und damit eine Element von C ist. Man beachte, dass es viel aufwändiger wäre, für jedes der 16 Codewörter den Hamming–Abstand zu $\vec{w}$ zu berechnen und dann das Codewort mit dem geringsten Abstand zu wählen.

Die Hamming–Codes haben noch eine weitere schöne Eigenschaft: Zu jedem Wort $\vec{w}$ gibt es ein *eindeutig bestimmtes* Codewort, welches zu $\vec{w}$ minimale Hamming–Distanz hat. Das heißt, wenn es bei der Übertragung eines Codewortes $\vec{c}$ *zwei* Fehler gab und es als $\vec{w}$ übertragen wurde, dann gibt es ein anderes Codewort $\vec{\tilde{c}} \in C$ mit $d(\vec{\tilde{c}}, \vec{w}) = 1$. Allgemein definiert man:

Definition 8.9
Es sei $C \subset \mathbb{K}^n$ t ein t–fehlerkorrigierender Code. Wenn $\forall \vec{w} \in \mathbb{K}^n : \exists \vec{c} \in C : d(\vec{c}, \vec{w}) \leq t$, so heißt C **perfekt**.

Hamming–Codes sind perfekte 1–fehlerkorrigierende Codes.

8.0.2 Perfekte Codes und Sportwetten

Perfekte Codes dienen nicht nur der Datenübertragung: Man kann auch versuchen, mit ihnen Sportwetten zu „knacken“.

Beim Fußballtoto wird jede Woche eine Liste von n Fußballspielen ausgewählt, auf deren Ergebnis gewettet wird; für jedes Spiel gibt es drei mögliche Spielergebnisse, nämlich Unentschieden (0), Heimsieg (1) oder Auswärtssieg (2). Es gibt also 3^n mögliche Wetten, und der Wettgewinn hängt von der Anzahl der richtig vorhergesagten Spielergebnisse ab. Toto gab es auch in der DDR. In der BRD wurde Toto ab 1956 mit $n = 12$ (Zwölferwette), ab 1959 mit $n = 13$ (Dreizehnerwette), ab 1967 wieder als Zwölferwette, ab 1969 mit $n = 11$ (Elferwette) und ab 2004 wieder als Dreizehnerwette gespielt.

Der Einsatz beträgt 0,50€ pro Wette (wir vernachlässigen die Gebühr von 0,35€ pro Wettschein). Nach Daten der Westdeutschen Lotterie GmbH & Co OHG beträgt der theoretische Nettogewinn bei 13 richtig vorhergesagten Spielergebnissen rund 99.000€, bei 12 Richtigen rund 3.800€, bei 11 Richtigen rund 300€ und bei 10 Richtigen rund 40€.

Eine mögliche Gewinnstrategie für die Dreizehnerwette besteht nun darin, verschiedene Wetten so abzugeben, dass garantiert (also unabhängig von den tatsächlichen Spielergebnissen) mindestens eine Wette mit mindestens 11 Richtigen dabei ist. Dabei sollte natürlich die Anzahl der Wetten so klein sein, dass der Wetteinsatz kleiner als der erwartete Nettogewinn ist. Beim Platzieren dieser Wetten helfen perfekte Codes.

[37] Wir schreiben Codewörter als Bitfolge, obwohl sie „eigentlich“ Spaltenvektoren sind.

Da es drei mögliche Einzelergebnisse gibt, kann man diesmal $\mathbb{K} = \mathbb{F}_3$ wählen, also den Körper mit drei Elementen 0, 1, $2 = -1$. Die möglichen Wetten entsprechen den 3^n Elementen von $\mathbb{K}^n$. Auch in diesem Fall gibt es Hamming–Codes[38]. Sie sind perfekt, 1–fehlerkorrigierend, und es gilt $C \subset \mathbb{K}^n$ mit $n = \frac{3^k-1}{2}$ und $d = \dim(C) = n - k$ für ein $k \in \mathbb{N}$. Dabei ist weiterhin n die Anzahl der Spielergebnisse pro Wette, es werden 3^d Wetten abgegeben, und es ist garantiert, dass eine der Wetten $n - 1$ oder gar n Richtige hat.

Praktischerweise ist $13 = \frac{3^3-1}{2}$. Mit dem ternären Hamming–Code für $k = 3$ würde man also $3^{13-3} = 59049$ Wetten abgeben — der Wetteinsatz (ohne Gebühr) dafür beträgt 29.524,50€. Es gibt insgesamt $3^{13} = 1594323$ mögliche Wetten. Wenn man eine gleichmäßige Verteilung der Spielergebnisse annimmt, beträgt die Wahrscheinlichkeit für 13 Richtige $\frac{3^{10}}{3^{13}} = \frac{1}{27}$. Wenn man keine 13 Richtigen hat, so sind 12 Richtige *garantiert* — dafür ist die Wahrscheinlichkeit $1 - \frac{1}{27} = \frac{26}{27}$. Kleinere Gewinne vernachlässigen wir an dieser Stelle. Wir erwarten also einen Nettogewinn von $(\frac{1}{27} \cdot 99.000 + \frac{26}{27} \cdot 3.800)$€, also rund 7.325€. Schade — nach Abzug des Wetteinsatzes würde man einen herben Verlust machen!

Es gibt noch andere perfekte Codes, zum Beispiel den *ternären Golay*[39]*–Code*. Dies ist ein perfekter 2–fehlerkorrigierender Code mit $\mathbb{K} = \mathbb{F}_3$, $n = 11$ und $\dim(C) = 6$. Das würde eher zur Elfer– als zur Dreizehnerwette passen.[40]

Doch auch für die Dreizehnerwette kann man den Golay–Code nutzen. Mit „Expertenwissen“ könnte es nämlich möglich sein, das Ergebnis von zwei der dreizehn Spiele mit ziemlicher Sicherheit vorherzusagen. Auf diese zwei Spielergebnisse würde man also fest wetten und auf die restlichen 11 Spielergebnisse verschiedene Wetten gemäß des ternären Golay–Codes abgeben. Man erhält $3^6 = 729$ Wetten, also einen Wetteinsatz von 364,50€. Falls das „Expertenwissen“ zutrifft und die Spielergebnisse gleichverteilt sind, erhält man 13 Richtige mit der Wahrscheinlichkeit $\frac{3^6}{3^{11}} = \frac{1}{243}$. Für jedes Codewort des Golay–Codes gibt es 11 mögliche Stellen, an denen es einen Fehler geben kann, und wir haben $|\mathbb{K}| = 3$. Also gibt es für jedes Codewort $11 \cdot (|\mathbb{K}| - 1) = 22$ Wörter mit Hamming–Abstand 1. Weil der Golay–Code 2–fehlerkorrigierend ist, überlappen sich diese Wortmengen nicht. Das heißt: Mit Wahrscheinlichkeit $\frac{22}{243}$ hat man zwar keine 13, aber 12 Richtige. Weil der Golay–Code 2–fehlerkorrigierend und perfekt ist, hat man ansonsten 11 Richtige, also mit Wahrscheinlichkeit $1 - \frac{1}{243} - \frac{22}{243} = \frac{220}{243}$. Der erwartete Nettogewinn beträgt rund $(\frac{1}{243} \cdot 99.000 + \frac{22}{243} \cdot 3.800 + \frac{220}{243} \cdot 300)$€, das sind rund 1.023€ — das liegt deutlich über dem Wetteinsatz!

In dieser Analyse wurde einerseits Expertenwissen für zwei Spiele, andererseits eine Gleichverteilung der übrigen elf Spielergebnisse vorausgesetzt. Das sind nicht sehr realistische Annahmen, so dass Sie sich auf obige Gewinnstrategie nicht verlassen sollten. Wenn es zu viele Gewinner gibt, dann sinken außerdem die Quoten.

[38] Über $\mathbb{K} = \mathbb{F}_3$ nennt man sie *ternär*.

[39] Marcel Jules Edouard Golay [1902–1989], Schweizer Elektroingenieur

[40] Mir ist nicht bekannt, ob *deshalb* von der Elfer– zur Dreizehnerwette gewechselt wurde.

Suchen Sie sich also lieber einen richtigen Beruf — zum Beispiel InformatikerIn.

Index