Universität Bremen

Mathematik 1 – Logik, Kombinatorik und Lineare Algebra

Notizen zur Vorlesung

Dr. Tim Haga

Wintersemester 2022/23 – the energy crisis war – Sláva Ukrayíni!

ALTA
Institute for Algebra, Geometry, Topology and their Applications

Disclaimer: Diese Notizen sind „work in progress“. Sie enthalten mit Sicherheit noch Fehler und Ungenauigkeiten.

Stand:**2022-10-28**
Version: 8fc909e

Verbesserungsvorschläge nehme ich gerne per E-Mail entgegen:

timhaga@uni-bremen.de

Ich danke herzlich Adil Ahmed, Lazhar Ben Amor, Jan Barfuß, Jan Blumenkamp, Steffen Brunßen, Felix Friedrich, Maurice Funk, Philipp Haker, Carl Hammann, Marieke Hennemann, Joana Holsten, Merlin Kannengießer, Hasan Köksal, Julia Kramer, Kim Marie Lankenau, Tim Lindemann, Lukas Monnerjahn, Laurens Müller-Groh, Lukas Nölke, Bino Nolting, Kai Renken, Marvin Schäcke, Ivo Florin Scheiber, Till Schlechtweg, Frederik Schlikker, Dennis Schürholz, Marcel Westenberg, und Maximilian Wick für ihre Hinweise zur Verbesserung dieser Notizen.

Jane-Loreen Krey danke ich für die Grafik zur Fallunterscheidungsregel.

Der Umfang und Inhalt dieser Notizen können von der Vorlesung abweichen. Im Zweifel gilt das in der Vorlesung Gesagte.

Änderungsverzeichnis:

v02	• Widerspruchsregel geändert und modifizierte Widerspruchsregel eingeführt. • Einige Bemerkungen beim logischen Schließen ergänzt. • kleinere Typos gefixt

Inhaltsverzeichnis

0 Vorbemerkung 1

0.1 Zu den Vorkenntnissen 1

0.2 Zur Notation 1

0.3 Zum Beweisen 3

1 Logik 5

1.1 Erste Begriffe 5

1.2 Sprachen 10

1.3 Junktoren 10

1.3.1 Negation von Aussagen 11

1.3.2 Konjunktion von Aussagen 11

1.3.3 Disjunktion von Aussagen 12

1.3.4 Subjunktion und Implikation 12

1.3.5 Bisubjunktion 13

1.3.6 Weitere Operatoren 13

1.3.7 Logische Terme 14

1.4 Normalformen 15

1.5 Logisches Schließen 17

1.6 Beweismethoden 22

2 Mengen 25

2.1 *Definitionen und Beispiele* 25
2.2 *Mengenoperationen* 28

3 *Relationen und Abbildungen* 33
3.1 *Grundbegriffe* 33
3.2 *Äquivalenzrelationen* 35
3.3 *Abbildungen* 37

4 *Natürliche Zahlen und Vollständige Induktion* 43
4.1 *Axiomatisierung der natürlichen Zahlen** 44
4.2 *Schwache und starke Induktion* 49

5 *Die ganzen Zahlen* 55
5.1 *Die ganzen Zahlen* 55
5.2 *Teilbarkeit* 58
5.3 *Division mit Rest* 59
5.4 *Primzahlen* 64

6 *Kombinatorik* 67
6.1 *Binomialkoeffizienten* 67
6.2 *Das Prinzip Inklusion-Exklusion* 68
6.3 *Das Schubfachprinzip und injektive Abbildungen* 69
6.4 *Bijektive Abbildungen und Permuationen* 72

7 *Gruppentheorie* 79
7.1 *Grundbegriffe* 79
7.2 *Homomorphismen* 84
7.3 *Untergruppen* 86
7.4 *Ringe und Körper* 88

8 Geometrie 93

8.1 Grundlagen 93

8.2 Standards 95

8.3 Winkel 97

8.4 Die Standardbasis im dreidimensionalen Raum 98

8.5 Kurven und Flächen 99

8.6 Höhere Dimensionen, Unterräume 101

8.7 Lineare Abbildungen 102

9 Vektorräume 105

9.1 Vektorraum, Basis, Lineare Unabhängigkeit 105

9.2 Winkel und Skalarprodukt 112

10 Lineare Gleichungssysteme 117

11 Matrizen 121

11.1 Grundbegriffe und Eigenschaften 121

11.2 Determinanten 123

12 Lineare Abbildungen 129

12.1 Definitionen und Grundbegriffe 129

12.2 Eigenwerte und Eigenvektoren 132

12.3 Eigenwerte symmetrischer Matrizen 137

12.4 Vektoriteration 137

12.5 Einschub: Normen von Matrizen 140

12.6 Anwendung: Googles PageRank 141

12.7 Anwendung: Orthogonale Matrizen 145

12.7.1 QR-Zerlegung 146

12.8 Das Gram-Schmidt-Verfahren 148

12.9 Basiswechsel 149

0
Vorbemerkung

In dieser Vorlesung werden wir uns mit einem Teil der mathematischen Grundlagen beschäftigen, welche Sie im Rahmen Ihres Studiums der Informatik, Wirtschaftsinformatik oder Digitalen Medien benötigen werden.

0.1 Zu den Vorkenntnissen

Wir werden *bei Null* anfangen, das heißt, es werden keine speziellen mathematischen Vorkenntnisse von Ihnen erwartet. Überspitzt könnte man sagen, die notwendigen Vorkenntnisse für diese Vorlesung sind lediglich Lesen und Schreiben. Auch wenn wir alle zentralen Konzepte sauber definieren werden, gehen wir tatsächlich davon aus, dass Sie mit dem Schulstoff halbwegs vertraut sind. Zumindest die Grundrechenarten, sowie die grundlegenden mathematischen Regeln sollten Sie kennen. Das heißt, die folgenden Aufgaben sollten Ihnen keine Schwierigkeiten machen:

Aufgabe 0.1. Berechnen Sie: $12 \cdot (13 - 24) \cdot (6 - 3)$.

Aufgabe 0.2. Schreiben Sie als einen Bruch: $\frac{5}{b-1} - \frac{6b}{b^2-1} - \frac{1-2b}{b+b^2}$.

Aufgabe 0.3. Lösen Sie nach x auf: $w = \frac{1}{2}v\left(1 - \frac{1+k}{1+\frac{a}{x}}\right)$.

Sollte Ihnen nicht klar sein, wie Sie die obigen Aufgaben lösen können, oder sollten Sie bei der Lösung Unsicherheiten haben, sollten Sie hier dringend nachbessern. Sie werden es im Laufe des Semesters nicht zu Beginn benötigen, aber erfahrungsgemäß ist die Stoffdichte im Studium zu groß, als dass Sie hierfür später noch Zeit haben werden.

0.2 Zur Notation

Mathematische Ausdrücke sind – ähnlich Programmiersprachen – Case-Sensitive, d. h. Klein- und Großbuchstaben haben verschiedene Bedeutungen. So sind n und N zwei verschiedene Symbole, die auch,

je nach Kontext, eine verschiedene Bedeutung haben. Tatsächlich ist Mathematik aber auch noch Font-Sensitive, d. h. auch die Schriftart kann eine Bedeutung haben. So haben etwa N und $\mathbb{N}$ eine verschiedene Bedeutung. Letzteres wird fast in jedem Kontext die Menge der natürlichen Zahlen symbolisieren. Daneben könnten Ihnen aber auch $\mathfrak{N}$, $\mathcal{N}$ und gegebenfalls auch noch andere Schriftarten begegnen. In der Literatur – aber nicht in diesem Skript – können Ihnen auch fettgedruckte Buchstaben begegnen, etwa $\mathbf{a}$. Auch dort wird $\mathbf{a}$ etwas anderes bedeuten als a oder $\mathfrak{a}$.

Erfahrungsgemäß macht zu Beginn des Studiums die Frakturschrift die größten Schwierigkeiten, daher sollten Sie sich mit dieser vertraut machen:

A	B	C	D	E	F	G	H	I	J	K	L	M
$\mathfrak{A}$	$\mathfrak{B}$	$\mathfrak{C}$	$\mathfrak{D}$	$\mathfrak{E}$	$\mathfrak{F}$	$\mathfrak{G}$	$\mathfrak{H}$	$\mathfrak{I}$	$\mathfrak{J}$	$\mathfrak{K}$	$\mathfrak{L}$	$\mathfrak{M}$
$\mathfrak{a}$	$\mathfrak{b}$	$\mathfrak{c}$	$\mathfrak{d}$	$\mathfrak{e}$	$\mathfrak{f}$	$\mathfrak{g}$	$\mathfrak{h}$	$\mathfrak{i}$	$\mathfrak{j}$	$\mathfrak{k}$	$\mathfrak{l}$	$\mathfrak{m}$
N	O	P	Q	R	S	T	U	V	W	X	Y	Z
$\mathfrak{N}$	$\mathfrak{O}$	$\mathfrak{P}$	$\mathfrak{Q}$	$\mathfrak{R}$	$\mathfrak{S}$	$\mathfrak{T}$	$\mathfrak{U}$	$\mathfrak{V}$	$\mathfrak{W}$	$\mathfrak{X}$	$\mathfrak{Y}$	$\mathfrak{Z}$
$\mathfrak{n}$	$\mathfrak{o}$	$\mathfrak{p}$	$\mathfrak{q}$	$\mathfrak{r}$	$\mathfrak{s}$	$\mathfrak{t}$	$\mathfrak{u}$	$\mathfrak{v}$	$\mathfrak{w}$	$\mathfrak{x}$	$\mathfrak{y}$	$\mathfrak{z}$

Tabelle 1: Das Alphabet in Fraktur.

Hinweis: An der Tafel lässt sich Druckfraktur nur schwer darstellen. Daher wird Fraktur an der Tafel üblicherweise in Sütterlin geschrieben.

Da Mathematikerinnen grundsätzlich unter chronischem Zeichenmangel leiden, werden wir neben dem lateinischen Alphabet auch noch das griechische Alphabet hinzuziehen.

Alpha	Beta	Gamma	Delta	Epsilon	Zeta	Eta	Theta	Iota	Kappa	Lambda	Mü
A	B	Γ	Δ	E	Z	E	Θ	I	K	Λ	M
α	β	γ	δ	ϵ, ε	ζ	η	θ	ι	κ	λ	μ
Nü	Xi	Omikron	Pi	Rho	Sigma	Tau	Ypsilon	Phi	Chi	Psi	Omega
N	Ξ	O	Π	R	Σ	T	Υ	Φ	X	Ψ	Ω
ν	ξ	o	π, ϖ	ρ	σ, ς	τ	υ	ϕ, φ	χ	ψ	ω

Tabelle 2: Das griechische Alphabet.

Offenbar sind einige Buchstaben des griechischen Alphabets auch in unserem lateinischen Alphabet enthalten. Um Verwirrungen zu vermeiden, werden wir diese daher nicht benutzen. (Hingegen war es früher in der Komplexitätstheorie üblich, klein und groß Omikron statt klein und groß O zu benutzen).

Einige kleinen griechischen Buchstaben kommen in zwei Varianten daher. Zumeist wird ε statt ϵ und π statt ϖ benutzt. Bei ϕ und φ scheint es keine Tendenz zu geben.

Da die griechischen Buchstaben nicht für alles ausreichen, haben Mathematikerinnen irgendwann angefangen, auch das hebräische Alphabet zu nutzen. Während der hebräische Buchstabe א (Aleph) in der Mengenlehre noch häufig vorkommt, sind die weiteren Buchstaben ב (Beth), ג (Gimel), ד (Daleth) usw. nur extrem selten und in spezieller Fachliteratur anzutreffen. In dieser Vorlesung werden wir

keine hebräischen Buchstaben benötigen, aber es könnte sein, dass sie Ihnen an anderer Stelle im Studium begegnen.

Darüber hinaus können Symbole oder Buchstaben mit zusätzlichen Zeichen versehen sein. Üblich sind etwa der Strich (a', lies: *a Strich*), der Überstrich ($\bar{a}$, lies: *a quer*), die Tilde ($\tilde{a}$, lies: *a Tilde / a Schlange* oder der Zirkumflex ($\hat{a}$, lies: *a Dach*). Der Strich kann auch mehrfach vorkommen (etwa a'', lies: *a Zweistrich*).

Zusätzlich könne an Symbolen noch Indizes angebracht sein. Neben Zahlen (a_{23}, lies: *a dreiundzwanzig* oder *a zwo drei*, je nach Kontext), können auch Buchstabenn (a_n, *lies: a n*) oder Symbole ($a_<$, lies: *a kleiner*) Indizes haben. Selten kommen auch hochgestellte Indizes vor (etwa: a^{23}), da aber die Verwechslung mit Potenzen recht groß ist, versuchen wir diese zu vermeiden. Auch Symbole können Indizes haben (etwa: $+_R$).

0.3 Zum Beweisen

Das Beweisen ist ein wesentliches Element der Mathematik und damit auch jeder Mathematikvorlesung. Beweise dienen dazu, sich von der Gültigkeit der getroffenen Aussagen zu überzeugen, ein mathematisches Theorem ohne einen Beweis ist eben kein Theorem, sondern nur eine Vermutung. Beweise nachzuvollziehen, aber auch selbst Beweise zu führen ist ein wesentlicher Teil jedes Studiums mathematischen Inhalts. Während das Nachvollziehen eines Beweises (*ich weiß, warum der Beweis richtig ist*) noch verhältnismäßig einfach ist, so ist das Finden eines Beweises sehr viel schwieriger. Insbesondere, da der Prozess der Beweisfindung meist nicht dokumentiert wird, lediglich wird der fertige Beweis präsentiert. Der nachfolgende Comic verdeutlicht dies.

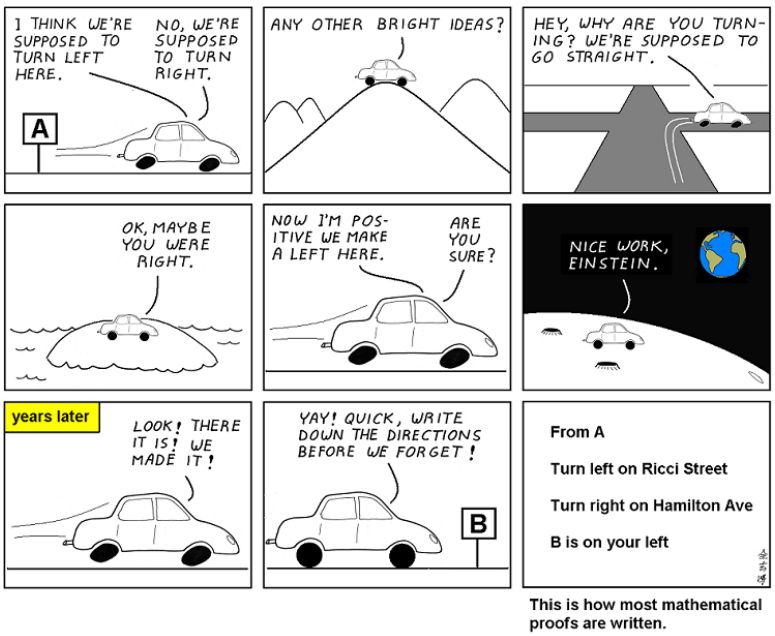


Abbildung 1:
http://abstrusegoose.com/230

1

Logik

Ein populärer Witz über Informatiker ist der Folgende:

> Ein Informatiker und seine Frau.
>
> Sie: „Schatz, wir haben kein Brot mehr. Könntest Du bitte zum Supermarkt gehen und eins holen? Und wenn sie Eier haben, bring 6 Stück mit."
>
> Nach kurzer Zeit kommt er wieder zurück und hat 6 Brote dabei.
>
> Sie: „Warum hast Du 6 Brote gekauft?"
>
> Er: „Sie hatten Eier."

In diesem Kapitel werden Sie lernen, warum dieser Witz nicht witzig ist, und warum eine präzise Sprache und rigide logische Regeln notwendig sind.

Ziel dieses Kapitels ist es, dass Sie

- wichtige Grundbegriffe der mathematischen Logik kennen und unterscheiden können;
- die wichtigsten logischen Operatoren anwenden können;
- in der Lage sind, für komplexe logische Ausdrücke die disjunktive und konjunktive Normalform aufzustellen.
- Beispiele für logische Schlussregeln und Beweismethoden kennen und diese in einfachen Kontexten anwenden können.

Logik ist die Sprache der Mathematik. Das Wort *Logik* stammt vom griechischen λόγος und bedeutet soviel wie Sinn oder Vernunft und beschreibt die Kunst des vernünftigen Schlussfolgerns.

1.1 Erste Begriffe

Wesentliche Fragestellungen der Logik sind

- Wann ist eine Aussage wahr?

- Wie schließt man richtig?
- Wie erhält man neue Erkenntnisse aus Alten?

Um diese Fragen fundiert beantworten können, müssen wir offensichtlich zunächst einige Begriffe klären, z. B. was eine *Aussage* ist, oder wann eine Aussage *wahr* oder *falsch* ist.

In der Mathematik nutzt man *Definitionen* um die Bedeutung von Begriffen festzulegen.

Definition 1.1. Eine *Aussage* ist ein grammatikalischer Ausdruck, von dem es sinnvoll ist zu fragen, ob er wahr oder falsch ist. Eine Aussage kann wahr oder falsch sein, aber nicht beides zugleich oder irgendwas dazwischen.

> Streng genommen interessieren sich Mathematikerinnen gar nicht dafür ob eine Aussage wahr oder falsch ist. So ist etwa „zwei mal drei ist sechs“ eine Aussage, und Sie würden vermutlich zustimmen, dass diese Aussage wahr bzw. richtig ist. Aber, wenn man diese Aussage beweisen wollen würde, müsste man zunächst definieren, was „zwei“ ist, was „drei“ ist, was „sechs“ ist, was „mal“ ist (und wie man damit umgeht) und daraus dann ableiten, dass zwei mal drei sechs ist. Aber ob das jetzt wahr ist, oder falsch ist, ist eher eine Frage für Philosophen. In der Mathematik interessiert uns lediglich, ob wir aus den Definitionen richtig abgeleitet haben.
> Natürlich wünschen wir uns, dass die Mathematik, die wir betreiben, die Wirklichkeit wie wir sie wahrnehmen, so gut wie es nur geht, beschreibt. Gleichwohl werden wir natürlich auch mit abstrakten Begriffen zu tun haben, die mit der Wirklichkeit nichts oder nur wenig zu tun haben. Der Sinn mag sich dann vielleicht nicht sofort erschließen, aber seien Sie versichert, dass es am Ende trotzdem nützlich sein wird.

Beispiel 1.2. „Der August hat 31 Tage“ ist eine Aussage (eine wahre). „Ein Schaltjahr hat 370 Tage“ ist ebenfalls eine Aussage (eine falsche). Hingegen ist „Das Schaf ist schwarz“ ohne nähere Spezifikation, welches Schaf genau gemeint ist, keine Aussage.

Es gibt Aussagen, von denen wir nicht wissen ob sie wahr oder falsch sind.

„Es gibt unendlich viele Paare von Primzahlen, deren Abstand gleich 2 ist.“ Dies ist eines der großen ungelösten Probleme der Mathematik. Gleichwohl wird diese Aussage entweder wahr sein, oder falsch. Welches von beiden wissen wir aber (noch) nicht. (Bemerkung: Es könnte auch sein, dass diese Aussage nicht entscheidbar ist, mehr dazu später.)

Ob etwas eine Aussage ist, hängt vom Kontext und den vorkommenden Objekten ab. Offenbar ist „$x > 3$“ keine Aussage, solange

wir nicht genau wissen, was die Variable x ist. Ist etwa $x = Schaf$, so ist „$x > 3$“ keine Aussage. Gleiches gilt, falls mit x tatsächlich der Buchstabe x des Alphabets gemeint ist. Ist hingegen $x = 3,1415926$, so ist „$x > 3$“ eine Aussage.

Hingegen ist „Für alle reellen Zahlen x gilt: $x > 3$“ eine (falsche) Aussage. Eine Aussage kann also durchaus Variablen enthalten.

Definition 1.3. Ein *Axiom* ist eine Aussage, die für Wahr erklärt wird.

Ein Axiom kann man also nicht beweisen, man kann aber z. B. beweisen, dass eine gewisse Struktur ein Axiom erfüllt. Im Aufbau der Mathematik nehmen Axiome das grundlegende Fundament ein, d. h. man gibt eine gewisse Menge von Axiomen vor und folgert daraus alle weiteren Aussagen. Dies ist nun nicht so einfach, wie es klingt, denn zunächst stellt sich etwa die Frage, ob die Auswahl der Axiome, welche man vorgegeben hat, sinnvoll sind. Es könnte ja sein, dass sich die Axiome widersprechen, oder sich eines der Axiome aus den anderen herleiten ließe (dann wäre es kein Axiom mehr). Dies ist ein erkenntnistheoretisches Problem, welches wir an dieser Stelle nicht vertiefen wollen.

Nachfolgend sehen Sie zwei Beispiele von Axiomen aus der klassischen Geometrie. Diese wurden schon von Euklid (*,† wsl. 3. Jhd. v. Chr.) postuliert.

Beispiel 1.4 (Axiom von der Geraden)**.** Zu zwei beliebigen, voneinander verschiedenen Punkten gibt es genau eine Gerade, welche diese beiden Punkte enthält.

Beispiel 1.5 (Parallelenaxiom)**.** Zu jeder Geraden, und jedem Punkt, welcher nicht auf dieser Geraden liegt, gibt es genau eine zu der Geraden parallele Gerade durch diesen Punkt.

Das Parallelenaxiom war lange Zeit in der Mathematik umstritten. Viele Mathematiker waren der Ansicht, dass es kein Axiom sei und das es möglich sein müsste es aus anderen Axiomen der Geometrie zu folgern. Erst gegen Ende des 19. Jahrhunderts konnte gezeigt werden, dass das Parallelenaxiom unabhängig von den anderen euklidischen Axiomen ist. Man kann also auch Geometrien studieren, für welche das Parallelaxiom nicht gilt (aber alle anderen euklidischen Axiome). Diese nennt man dann *nichteuklidische Geometrie*.

Beispiel 1.6 (Halbkreise)**.** Wir betrachten eine Gerade u in der Ebene (etwa die reelle Zahlengerade) und alle Punkte oberhalb dieser Geraden (man spricht auch von der oberen Halbebene). Als Geraden in unserer Geometrie der oberen Halbebene nehmen wir Strahlen die senkrecht auf u stehen, sowie die Halbkreise mit Mittelpunkt auf u. Nun kann man sich leicht davon überzeugen, dass es zu zwei verschiedenen Punkten in der oberen Halbebene genau einen Halbkreis mit Mittelpunkt auf u oder einen Strahl der senkrecht auf u steht gibt.

Hingegen ist das Parallelenaxiom nicht erfüllt.

Definition 1.7. Ein *Satz* (auch *Lemma, Korollar, Theorem, Proposition*) ist eine wahre Aussage.

Ein Satz wird üblicherweise aus anderen Sätzen oder aus Axiomen abgeleitet. Dies erfordert es, dass man eine Begründung angibt, warum dieser Satz wahr ist, d. h. man erläutert, wie sich der Satz aus den anderen wahren Aussagen ableiten lässt. Dies nennt man einen *Beweis*. Ein Beweis ist ein soziales Konstrukt, d. h. man kann nicht ohne weiteres sagen, was ein Beweis ist oder nicht. Etwas ist also dann ein Beweis, wenn es von anderen Mathematikern als Beweis anerkannt wird. Ein Lemma ist ein Hilfssatz, der nützlich zum Beweis eines Satzes ist. Ein Korollar ist eine Aussage, die sich aus einer Definition oder einem Satz ohne großen Beweisaufwand ergibt. Die Begriffe Theorem und Proposition werden oft Synonym mit dem Begriff Satz verwendet. Die Begriffe hängen ein wenig von den Autoren ab, welche sie benutzen, aber die Faustregel ist, dass eine Proposition üblicherweise keinen Namen hat, im Gegensatz zu Lemmata, Sätzen und Theoremen.

Im Wesentlichen geht es in der Mathematik um folgendes: Innerhalb eines vorgegebenen Systems von Axiomen möchte man so viele wahre Aussagen wie möglich aus diesen Axiomen ableiten. Wir betrachten dazu ein fiktives, umfangreicheres Beispiel:

Beispiel 1.8. An der Universität Bremen soll der neue Multi-Kombi-Bachelor (kurz: MuKoBa) eingeführt werden. Der Studiengang soll sich aus vielen verschiedenen Fächern zusammensetzen, jedes Fach muss aber von mindestens einem Studenten belegt werden. Außerdem hat man folgende Regeln festgelegt:

Axiom I: Jeder Student belegt mindestens ein Fach.

Axiom II: Zwei verschiedene Studenten belegen immer genau ein gemeinsames Fach.

Axiom III: Zu jedem Fach gibt es genau ein anderes Fach, so dass kein Student diese beiden Fächer belegt (sogenannte Komplementärfächer).

Diese drei Axiome (eigentlich sind es vier) bilden nun ein Axiomensystem, in welchem man nun Strukturuntersuchungen vornehmen kann. Tatsächlich versteckt sich hier eine interessante mathematische Struktur (auf welche wir in einem späteren Kapitel zurückkommen werden).

Man kann nun z. B. folgende Aussagen über den MuKoBa beweisen:

1. Jeder Student belegt mindestens zwei Fächer.
2. Jedes Fach wird von mindestens zwei Studenten belegt.
3. Es gibt mindestens sechs Fächer.

Man kann sogar ableiten, wieviele (!) Studenten den MuKoBa studieren.

Betrachten wir die erste Aussage. Wie kann man diese Beweisen, nur mittels der Axiome des MuKoBa?

Beweis. Sei A ein Student im neuen Studiengang. Dies können wir annehmen, da jedes Fach zumindest von einem Studenten belegt werden soll (0. Axiom). Dann sagt Axiom I, dass A ein Fach F studiert. Außerdem sagt uns Axiom III, dass es ein weiteres Fach G gibt, welches A nicht studiert. Insbesondere sagt uns Axiom III, dass jeder Student höchstens eines der Fächer F und G studiert. Aus dem 0. Axiom folgt, dass ein Student B dieses Fach G studiert. Da A nicht G studiert, sind A und B verschiedene Personen. Axiom II sagt, dass es ein Fach H gibt, welches von A und von B studiert wird. Da B das Fach H studiert, aber nicht das Fach F (denn B studiert schon G), sind F und H verschiedene Fächer. Also studiert A die beiden verschiedenen Fächer F und H und somit mindestens zwei Fächer. □

Die weiteren Aussagen bleiben der geneigten Leserin zur Übung empfohlen.

Wir haben in diesem Beispiel schon viel davon gesehen, was Mathematik ausmacht. Wir haben für einige Begriffe (Student, Fach) festgelegt, wie sie sich zueinander verhalten sollen (Axiome). Wir haben in dieser Begriffswelt eine neue Aussage formuliert und diese dann bewiesen. Dabei haben wir einen sogenannten *informellen Beweis* formuliert; Wir haben durch eine Kette von wahren Aussagen, die zu beweisende Aussage erhalten. Ob die Schlussfolgerungen im Beweis richtig sind, können wir durch logisches Denken erschließen. Für einen Computer hingegen ist das nicht so ohne weiteres möglich; Wir kommen darauf später in diesem Kapitel zurück.

Außerdem zeigt das Beispiel, das es zwar verhältnismäßig einfach ist, den Beweis nachzuvollziehen[1], wie der Beweis zu finden ist, d. h. wie man darauf kommt, dass verrät Ihnen der Beweis selbst aber nicht. Eine Aussage zu beweisen ist im Allgemeinen sehr viel schwieriger, als den Beweis auf seine Korrektheit zu prüfen (wobei dies in der Praxis auch sehr schwierig sein kann).

[1] Tun Sie dies bitte. Überlegen Sie, wie man an jeder Stelle des Beweises zur nächsten Schlussfolgerung gelangt. Das Nachvollziehen von Beweisen ist eine der wesentlichen eigenen Leistungen, die von Ihnen verlangt werden. Sollten Sie an einer Stelle nicht verstehen, wie der Beweis funktioniert, lesen Sie das Beispiel noch einmal in Ruhe von vorne durch. Ich empfehle Ihnen dringend, nicht weiter zu lesen, bevor alle Unklarheiten des Beweises für Sie beseitigt sind.

Wir wollen nun nocheinmal Ausdrücke mit Variablen näher betrachten.

Definition 1.9. Eine *Aussageform* ist ein Ausdruck in Variablen der zu einer Aussage wird wenn alle darin vorkommenden Variablen durch konkrete Objekte ersetzt werden. Diese Objekte müssen aus einer geeigneten Grundgesamtheit kommen.

Beispiel 1.10.
- x ist ohne Rest durch 7 teilbar. Eine geeignete Grundgesamtheit für x sind z. B. die natürliche Zahlen.
- $y = 4x + 5$. Geeignete Grundgesamtheiten für x und y sind z. B. die reellen Zahlen.

1.2 Sprachen

Um möglichst präzise formulieren zu können, müssen wir die Alltagssprache hinter uns lassen und uns der Sprache der Mathematik zuwenden. Mathematik lernen heißt auch, eine Sprache zu lernen. Tatsächlich gibt es viele formale Sprachen innerhalb der Mathematik, aber gewisse Grundprinzipien gelten in der gesamten Mathematik. In der Logik benutzt man z. B. sogenannte Sprachen der 1. Stufe. Andere Sprachen werden Sie in der theoretischen Informatik kennenlernen. Ähnlich wie in der Alltagssprache haben diese ein Alphabet, aus denen man Wörter und diese dann zu Ausdrücken formen kann. Dabei ist natürlich nicht jede beliebige Aneinanderreihung von Zeichen des Alphabets ein Wort und nicht jede Aneinanderreihung von Wörtern ein (sinnvoller) Ausdruck.

Wir werden daher folgende Zeichen in die Sprachen der Mathematik aufnehmen (und später weitere hinzufügen): $\neg$ (für „nicht“), $\wedge$ (für „und“), $\vee$ (für „oder“), $\rightarrow$ (für „wenn – dann“), $\leftrightarrow$ (für „genau dann wenn“), $\forall$ (für „für alle“), $\exists$ (für „es gibt“) und $=$ (als Gleichheitszeichen). Außerdem nehmen wir Variablen (für vorkommende Objekte, Strukturen, usw.) und Klammern als Hilfsymbole dazu.

1.3 Junktoren

Wir wollen nun untersuchen, wie sich Aussagen miteinander verknüpfen lassen und bestimmen, wie sich die Wahrheitswerte der zusammengesetzten Aussage verhalten.

In der Umgangssprache schwankt der Gebrauch insbesondere der Junktoren. Etwa das „oder“ wird manchmal nicht-exklusiv oder exklusiv, wie in „entweder-oder“, benutzt. Auch kann in der Umgangssprache die Bedeutung des Junktors vom Inhalt der verknüpften Aussagen abhängen. Vergleichen Sie etwa die beiden Aussagen „Ich sprang vor Freude in die Luft und stieß mir den Kopf“ und „Ich stieß mir den Kopf und sprang vor Freude in die Luft“. Der Wahrheitswert der Aussagen hängt hier also auch von der Reihenfolge der Aussagen und damit vom zeitlichen Bezug, also von ihrer Bedeutung.

Wir müssen also eine Normierung vornehmen. Diese ermöglicht uns, klar zu erkennen welchen Wahrheitswert eine zusammengesetzte Aussage hat, wenn wir nur den Wahrheitswert der Teilaussagen kennen (und nicht etwa auch den Inhalt der Aussagen kennen müssen).

Dabei werden wir Aussagen im Folgenden mit großen lateinischen Buchstaben bezeichnen, wobei uns, wie oben bemerkt, nicht die konkrete Aussage interessiert, sondern lediglich der Wahrheitswert der Aussage. Das bedeutet, dass wir für die im Folgenden auftretenden Variablen $A, B, C, \ldots$ jeweils konkrete Aussagen einsetzen können.

1.3.1 Negation von Aussagen

Definition 1.11. Für die *Verneinung* (auch: *Negation*) einer Aussage A schreiben wir $\neg A$. Die Negation hat die folgende Wahrheitstafel:

A	$\neg A$
W	F
F	W

Tabelle 1.1: Wahrheitstafel der Negation.

Die Tafel ist wie folgt zu lesen: A ist eine Aussage und diese kann die Wahrheitswerte W (kurz: für Wahr) oder F (kurz: für Falsch) annehmen. In der rechten Spalte sehen wir die Wahrheitswerte der Aussage $\neg A$, wobei wir zeilenweise ablesen: Ist A Wahr, so ist $\neg A$ Falsch bzw. ist A Falsch, so ist $\neg A$ Wahr.

Man kann die Wahrheitstafel also als eine Wertetabelle für die (einstellige) Funktion $\neg$ auffassen.

Beispiel 1.12. Aussage A: Die Zahl 7 ist gerade.
Negation von A: Die Zahl 7 ist ungerade.

Aussage B: Alle Schafe sind schwarz.
Negation von B: Nicht alle Schafe sind schwarz.

Äquivalente Formulierung: Es gibt mindestens ein Schaf, welches nicht schwarz ist.

Besondere Vorsicht ist bei sogenannten All- und Existenzaussagen geboten:

Beispiel 1.13. Aussage: Es gibt eine reelle Zahl x für die gilt $x > 7$.
Negation: Für alle reellen Zahlen x gilt: $x \leq 7$.

Aussage: Für alle reellen Zahlen x gilt: $x^2 \geq 0$.
Negation: Es gibt eine reelle Zahl x für die gilt: $x^2 < 0$.

Bei der Negation einer Allaussage ersetzt man also den Teil „Für alle“ mit „Es gibt (mindestens) ein“ und negiert dann den Teil der Aussage nach „(für die) gilt:“. Sinngemäß verfährt man bei der Negation einer Existenzaussage.

1.3.2 Konjunktion von Aussagen

Es seien A und B Aussagen. Wann ist die Aussage

Es gilt A und B (kurz: $A \wedge B$)

wahr?

Definition 1.14. Die *Verundung* oder *Konjunktion* zweier Aussagen A und B bezeichnen wir mit $A \wedge B$. Sie hat die folgende Wahrheitstafel:

Im Gegensatz zur Negation ist die Konjunktion eine zweistellige Funktion. Wir müssen also alle möglichen Kombinationen der möglichen Wahrheitswerte von A und B durchgehen um $\wedge$ definieren zu

A	B	$A \wedge B$
W	W	W
W	F	F
F	W	F
F	F	F

Tabelle 1.2: Wahrheitstafel der Konjunktion.

können. Da wir bei A und B jeweils zwei mögliche Wahrheitswerte haben, gibt es also insgesamt vier mögliche Kombinationen, wie der Tafel zu entnehmen ist.

1.3.3 Disjunktion von Aussagen

Wie oben seien A und B Aussagen. Wann ist die Aussage

Es gilt A oder B (kurz $A \vee B$)

wahr?

Definition 1.15. Die *Veroderung* oder *Disjunktion* zweier Aussagen A und B bezeichnen wir mit $A \vee B$. Sie hat die folgende Wahrheitstafel:

A	B	$A \vee B$
W	W	W
W	F	W
F	W	W
F	F	F

Tabelle 1.3: Wahrheitstafel der Disjunktion.

1.3.4 Subjunktion und Implikation

Seien A und B Aussagen. Wann ist die Aussage:

Wenn A gilt, dann gilt auch B (kurz: $A \to B$)

wahr?

Definition 1.16. Die *Wenn-Dann-Verknüpfung* oder *Subjunktion* zweier Aussagen A und B bezeichnen wir mit $A \to B$. Sie hat die folgende Wahrheitstafel:

A	B	$A \to B$
W	W	W
W	F	F
F	W	W
F	F	W

Tabelle 1.4: Wahrheitstafel der Implikation. **Achtung:** Die letzen beiden Zeilen sind nicht intuitiv. (Ex falso quodlibet!)

Es ist auf den ersten Blick vielleicht nicht intuitiv, warum die letzten beiden Zeilen der Wahrheitstafel so stimmen. Man sagt, dass aus Falschem Beliebiges folgt, lateinisch *ex falso (sequitur) quodlibet*.

Definition 1.17. Ist die verknüpfte Aussage $A \to B$ wahr, so nennt man dies eine *Implikation*, und schreibt

$$A \Rightarrow B.$$

Dies liest man als *aus A folgt B*, oder *A ist hinreichend für B*, oder *B ist notwendig für A*.

1.3.5 Bisubjunktion

Seien A und B Aussagen. Wann ist die Aussage:

Genau dann, wenn A gilt, gilt auch B (kurz: $A \leftrightarrow B$)

wahr?

Definition 1.18. Die *Bisubjunktion* oder *Genau-Dann-Wenn-Verknüpfung* zweier Aussagen A und B bezeichnen wir mit $A \leftrightarrow B$. Sie hat die folgende Wahrheitstafel:

A	B	$A \leftrightarrow B$
W	W	W
W	F	F
F	W	F
F	F	W

Tabelle 1.5: Wahrheitstafel der Bisubjunktion.

1.3.6 Weitere Operatoren

Neben den oben eingeführten Operatoren kann man nun noch weitere logische Operatoren einführen. Da eine Wahrheitstafel für zwei Aussagen insgesamt über vier Zeilen verfügt, und in jeder Zeile am Ende ein Wahr oder Falsch stehen kann, gibt es insgesamt $2^4 = 16$ binäre Operatoren, d. h. Verknüpfungen von zwei Aussagen. Viele dieser Verknüpfungen sind für uns nicht interessant, häufig kommt aber die folgende vor:

Definition 1.19. Das *Exklusive Oder* oder auch *xor* hat die folgende Wahrheitstafel:

A	B	$A \operatorname{xor} B$
W	W	F
W	F	W
F	W	W
F	F	F

Tabelle 1.6: Wahrheitstafel des Exklusiven Oders.

Die Verknüpfung $A \operatorname{xor} B$ wird also genau dann wahr, wenn A und B verschiedene Wahrheitswerte annehmen.

1.3.7 Logische Terme

Die oben eingeführten logischen Operatoren nennt man auch Junktoren.

Definition 1.20. Ein Ausdruck in Aussagen $A, B, C, \ldots$, welche durch Junktoren verknüpft sind, nennt man einen *(logischen) Term*. Dabei nutzt man Klammern um Teilterme, um deutlich zu machen, in welcher Reihenfolge die Junktoren ausgewertet werden sollen.

Man interessiert sich nun, dafür, ob ein Term durch einen anderen ersetzt werden kann, ohne den Wahrheitswert des Terms zu ändern.

Definition 1.21. Zwei Terme Z_1 und Z_2 heißen *logisch äquivalent*, wenn sie für alle Werte der darin vorkommenden Variablen denselben Wahrheitswert haben. Wir schreiben:

$$Z_1 \Leftrightarrow Z_2.$$

Beispiel 1.22. Sei A eine Aussage. Die beiden folgenden logischen Terme sind logisch äquivalent:

- A
- $\neg(\neg A)$.

Dies kann man mittels Wahrheitstafeln überprüfen: Da A und $\neg(\neg A)$

A	$\neg A$	$\neg(\neg A)$
W	F	W
F	W	F

die selbe Wahrheitstafel haben, sind die Ausdrücke logisch Äquivalent.

Von der folgenden Auflistung logisch aquivalenter Terme werden wir nur eine Äquivalenz beweisen. Der Rest bleibt der geneigten Leserin als Übungsaufgabe überlassen.

Satz 1.23. *Es gelten die folgenden Äquivalenzen*

i) $\neg(A \wedge B) \Leftrightarrow \neg A \vee \neg B$ *(Negation von UND)*

ii) $\neg(A \vee B) \Leftrightarrow \neg A \wedge \neg B$ *(Negation von ODER)*

iii) $(A \wedge B) \wedge C \Leftrightarrow A \wedge (B \wedge C)$ *(Assoziativgesetz)*

iv) $(A \vee B) \vee C \Leftrightarrow A \vee (B \vee C)$ *(Assoziativgesetz)*

v) $(A \wedge B) \vee C \Leftrightarrow (A \vee C) \wedge (B \vee C)$ *(Distributivgesetz)*

vi) $(A \vee B) \wedge C \Leftrightarrow (A \wedge C) \vee (B \wedge C)$ *(Distributivgesetz)*

vii) $A \wedge B \Leftrightarrow B \wedge A$ *(Kommutativgesetz)*

viii) $A \vee B \Leftrightarrow B \vee A$ *(Kommutativgesetz)*

ix) $(A \to B) \Leftrightarrow (\neg B \to \neg A)$ *(Kontrapositionsregel)*

x) $(A \leftrightarrow B) \Leftrightarrow (A \to B) \wedge (B \to A)$

Beweis. Wir beweisen exemplarisch Teilaussage i): Dazu betrachten wir die Wahrheitstafeln von $\neg(A \wedge B)$ und von $\neg A \vee \neg B$. Wir sehen,

A	B	$A \wedge B$	$\neg(A \wedge B)$	$\neg A$	$\neg B$	$\neg A \vee \neg B$
W	W	W	F	F	F	F
W	F	F	W	F	W	W
F	W	F	W	W	F	W
F	F	F	W	W	W	W

dass die beiden markierten Spalten identisch sind. Somit sind die beiden Terme logisch äquivalent. □

Äquivalenzen wie aus Satz 1.23 erlauben es, in komplizierten Termen bestimmte Teilterme durch andere zu ersetzen. Damit ist es unter Umständen möglich, logische Terme zu vereinfachen ohne deren Wahrheitswert zu verändern. Auch ist es so möglich, die logischen Operatoren $\vee, \to, \leftrightarrow$ nur durch die Nutzung von $\neg$ und $\wedge$ auszudrücken (Übung).

Abschließen definieren wir noch folgende Arten spezieller Terme:

Definition 1.24. Ein logischer Term heißt *Tautologie*, wenn er stets wahr ist, unabhängig davon welche Aussagen eingesetzt werden.

Ein logischer Term heißt *Kontradiktion*, wenn er stets falsch ist, egal welche Aussagen eingesetzt werden.

Ein Beispiel für eine Tautologie ist der *Satz vom ausgeschlossenen Dritten*:

$$A \vee \neg A.$$

Dieser wird auch *Tertium non Datur* genannt.

Eine bekannte Kontradiktion ist der *Satz vom ausgeschlossenen Widerspruch*:

$$A \wedge \neg A.$$

1.4 Normalformen

Es ist sinnvoll, sich bei logischen Termen auf eine standardisierte Form zu einigen. Dies ist die Idee von Normalformen, bei denen man sich auf spezielle Regeln für logische Terme einigt.

Definition 1.25. Sei Z ein logischer Term in Variablen $X_1, \ldots, X_n$. Z ist in *Disjunktiver Normalform* (kurz: *DNF*), falls

$$Z = Z_1 \vee \cdots \vee Z_k$$

gilt, wobei

$$Z_i = Y_1 \wedge \cdots \wedge Y_n, \quad \text{für alle } i = 1, \ldots, k$$

und jedes

$$Y_j = X_j \quad \text{oder } Y_j = \neg X_j$$

ist.

Z ist in *Konjunktiver Normalform* (kurz: *KNF*), falls

$$Z = Z_1 \wedge \cdots \wedge Z_k$$

gilt, wobei

$$Z_i = Y_1 \vee \cdots \vee Y_n, \quad \text{für alle } i = 1, \ldots, k$$

und jedes

$$Y_j = X_j \quad \text{oder } Y_j = \neg X_j$$

ist.

Die DNF ist also eine Veroderung von UND-Ausdrücken, während die KNF eine Verundung von ODER-Ausdrücken ist.

Beispiel 1.26. Wir betrachten den logischen Term $A \leftrightarrow B$. Dieser ist logisch äquivalent zur DNF $(A \wedge B) \vee (\neg A \wedge \neg B)$ und zur KNF $(\neg A \vee B) \wedge (A \vee \neg B)$.

Und ein etwas größeres Beispiel:

Beispiel 1.27. Der Ausdruck

$$Z = (A \wedge B \wedge C) \vee (\neg A \wedge B \wedge \neg C)$$

ist ein Ausdruck in DNF, und

$$Z' = (A \vee B \vee C) \wedge (\neg A \vee \neg B \vee \neg C) \wedge (\neg A \vee B \vee C)$$

ist ein Ausdruck in KNF.

Die Bedeutung von DNF und KNF wird durch den folgenden Satz klar.

Satz 1.28. *Jeder logische Term in den Variablen $X_1, \ldots, X_n$ ist zu einem logischen Term in disjunktiver (konjunktiver) Normalform logisch äquivalent. Letzterer ist bis auf die Reihenfolge der Teilterme, die durch ODER (UND) verknüpft sind, eindeutig bestimmt.*

Wir verzichten an dieser Stelle auf einen vollständigen Beweis und geben nur eine Beweisidee durch die explizite Konstruktion der Normalformen an.

Beispiel 1.29. Wir betrachten den logischen Term

$$Z = \neg\big((A \vee B \vee C) \wedge (\neg A \vee C)\big)$$

und möchten diesen nun in DNF bzw. KNF bringen. Dazu stellen wir zunächst die Wahrheitstafel von Z auf:

Dann betrachten wir die Zeilen in denen Z Wahr ist und stellen für jede dieser Zeile einen UND-Ausdruck in A, B, C auf, der mit der

A	B	C	$\neg A \vee C$	$A \vee B \vee C$	$(A \vee B \vee C) \wedge (\neg A \vee C)$	Z
W	W	W	W	W	W	F
W	W	F	F	W	F	W
W	F	W	W	W	W	F
W	F	F	F	W	F	W
F	W	W	W	W	W	F
F	W	F	W	W	W	F
F	F	W	W	W	W	F
F	F	F	W	F	F	W

Belegung dieser Zeile Wahr ist. Die erste Wahr-Zeile liefert $Z_1 = A \wedge B \wedge \neg C$. Entsprechend liefern die anderen beiden Wahr-Zeilen: $Z_2 = A \wedge \neg B \wedge \neg C$ und $Z_3 = \neg A \wedge \neg B \wedge \neg C$. Durch Veroderung von Z_1, Z_2, Z_3 erhält man einen zu Z äquivalenten Ausdruck in DNF:

$$Z \Leftrightarrow (A \wedge B \wedge \neg C) \vee (A \wedge \neg B \wedge \neg C) \vee (\neg A \wedge \neg B \wedge \neg C).$$

Analog erhält man einen zu Z äquivalenten Term in KNF, wenn man die Falsch-Zeilen betrachtet und jeweils einen ODER-Ausdruck in A, B, C aufstellt, welcher mit der Belegung der jeweiligen Zeile Falsch ist. Man erhält dann folgenden Ausdruck in KNF:

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

1.5 Logisches Schließen

In diesem Abschnitt wollen wir beleuchten, wie man formal aus alten Aussagen Neue erhält. Zunächst ist gar nicht klar, wie man aus gültigen Prämissen (so nennt man in der Logik, die Voraussetzungen oder Annahmen) gültige Schlussfolgerungen ziehen kann, denn dabei stößt man auf folgende Probleme:

(1) Sind die Prämissen wirklich wahr?

(2) Wie vermeidet man falsche Schlussfolgerungen?

Die Lösung des ersten Problems liegt darin, dass man *postuliert*, dass bestimmte Prämissen wahr sind. Wir nennen solche Prämissen bekanntlich *Axiome*. Das zweite Problem löst man, in dem man ein System von bestimmten Regeln aufstellt und dieses befolgt, so dass nur korrekte Schlussfolgerungen erlaubt sind.

Wenn man dies richtig handhabt, kann man aus einer kleinen Anzahl von Axiomen viele neue Aussagen herleiten.

Dann stellt sich die Frage, ob man so eigentlich alle möglichen gültigen Aussagen erhält. Der Mathematiker Kurt Gödel hat in den 1930er Jahren gezeigt, dass dies im Allgemeinen nicht möglich ist, dass es also sein kann, dass man in einem System von Axiomen Aussagen formulieren kann, welche innerhalb dieses Systems nicht beweisbar sind.

In diesem Abschnitt wollen wir nun lernen, wie man korrekte Schlussfolgerungen zieht, ohne uns allzu sehr mit Grundlagenfragen beschäftigen zu müssen.

Definition 1.30. Eine *Belegung* eines logischen Terms ist eine Zuweisung von Wahrheitswerten zu jeder Aussagenvariable des Terms.

Haben wir also einen logischen Term gegeben und betrachten die Wahrheitstafel des Terms, so entspricht jede Zeile der Tafel einer Belegung.

Definition 1.31. Eine *Interpretation* eines logischen Terms ist eine Zuordnung von konkreten Aussagen zu jeder Aussagenvariable des Terms. Eine Interpretation *erfüllt* einen logischen Term, wenn der Term bei der Interpretation wahr wird.

Jede Interpretation liefert also eine Belegung des logischen Terms.

Wir werden nun betrachten, wie man logische Schlussfolgerungen erhält und dabei ein wenig Notation verwenden, welche wir formal erst im Kapitel 2 einführen.

Im folgenden bezeichnet $\mathcal{K}$ eine Menge von logischen Termen und P einen logischen Termen. Mit $\{Q\}$ bezeichnen wir die Menge, welche nur den Ausdruck Q enthält, mit $\{P, Q\}$ entsprechend eine Menge, die nur die Terme P und Q enthält. Mit $\mathcal{K} \cup \{Q\}$ bezeichnen wir die Menge, welche neben allen Termen aus $\mathcal{K}$ auch den Term Q enthält.

Damit wir nur korrekte Schlussfolgerungen erhalten, muss das folgende *Korrektheitskriterium* gelten:

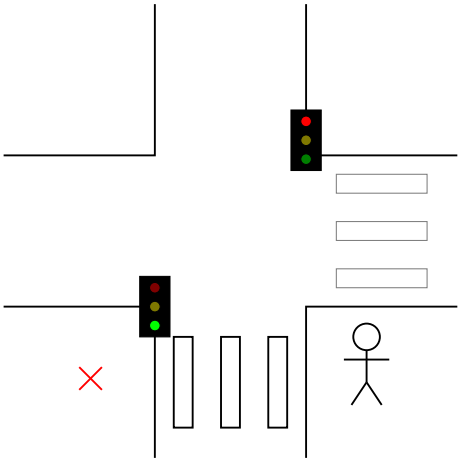
Der Term P kann nur dann aus $\mathcal{K}$ hergeleitet werden, wenn jede Interpretation, die alle Elemente von $\mathcal{K}$ erfüllt, auch P erfüllt.

Wir stellen nun einige Schlussregeln auf, welche wir nur intuitiv auf ihre Korrektheit hin überprüfen können (sogenannte Grundregeln).

Fallunterscheidungsregel (FU): Es seien $\mathcal{K}$ eine Menge von Termen, sowie P und Q Terme. Dann schreiben wir

$$\begin{array}{ll} \mathcal{K} \cup \{Q\} & P \\ \mathcal{K} \cup \{\neg Q\} & P \\ \hline \mathcal{K} & P \end{array}$$

und meinen damit: kann P sowohl aus $\mathcal{K} \cup \{Q\}$ und aus $\mathcal{K} \cup \{\neg Q\}$ hergeleitet werden, dann kann P auch schon aus $\mathcal{K}$ hergeleitet werden. Diese Regel ist korrekt, denn jede Interpretation die K erfüllt, erfüllt auch Q oder $\neg Q$., d. h. $\mathcal{K} \cup \{Q\}$ oder $\mathcal{K} \cup \{\neg Q\}$ In beiden Fällen erfüllt Sie nach Prämisse auch P.



Man kann sich die Situation vielleicht so vorstellen wie die Ampelschaltung an einer Kreuzung, siehe Abbildung. Möchte man als Fußgänger über die eine Straße gehen um das X zu erreichen, so kann man dies nur tun, wenn die zugehörige Ampel grün ist. Wenn also

die Ampel für die Straße, die ich überqueren möchte grün ist und die Ampel über die andere Straße grün ist, kann ich über die Straße gehen. Wenn aber die Ampel für die Straße, die ich überqueren möchte grün ist und die Ampel über die andere Straße rot ist, kann ich die Straße trotzdem überqueren. Die Ampel für die Überquerung der anderen Straße spielt also keine Rolle für die Möglichkeit das X zu erreichen.

Widerspruchsregel (Wid): Es seien $\mathcal{K}$ eine Menge von Termen, P und Q Terme. Dann lautet die Regel

$$\frac{\begin{array}{ll}\mathcal{K} \cup \{\neg P\} & Q \\ \mathcal{K} \cup \{\neg P\} & \neg Q\end{array}}{\begin{array}{ll}\mathcal{K} & P\end{array}}$$

Das bedeutet, dass wenn sowohl Q als auch $\neg Q$ aus $\mathcal{K} \cup \{\neg P\}$ hergeleitet werden kann, dann kann auch P aus $\mathcal{K}$ hergeleitet werden.

Diese Regel ist korrekt, denn jede Intepretation, die $\mathcal{K} \cup \{\neg P\}$ erfüllt, erfüllt sowohl Q als auch $\neg Q$. Ein Term ist aber entweder wahr oder nicht wahr. Darum gibt es gar keine Interpretation, welche $\mathcal{K} \cup \{\neg P\}$ erfüllt. Daher wird P von jeder Interpretation erfüllt, welche $\mathcal{K}$ erfüllt. Wenn man sich an die Schreibweise dieser Schlussregeln gewöhnt hat, dann sind die beiden folgenden beiden Regeln sofort klar.

$\vee$-Einführung im Sukzedenz ($\vee$Suk): Es seien $\mathcal{K}$ eine Menge von Termen, P und Q Terme. Dann gelten

$$\frac{\begin{array}{ll}\mathcal{K} & P\end{array}}{\begin{array}{ll}\mathcal{K} & P \vee Q\end{array}} \qquad \frac{\begin{array}{ll}\mathcal{K} & P\end{array}}{\begin{array}{ll}\mathcal{K} & Q \vee P\end{array}}$$

Diese Regel ist korrekt, denn jede Interpretation die $\mathcal{K}$ erfüllt, erfüllt auch P, und damit auch $P \vee Q$, bzw. $Q \vee P$.

Antezedensregel (Ant): Falls alle Terme von $\mathcal{K}$ auch in $\mathcal{K}'$ ($\mathcal{K} \subseteq \mathcal{K}'$) enthalten sind, dann

$$\frac{\begin{array}{ll}\mathcal{K} & P\end{array}}{\begin{array}{ll}\mathcal{K}' & P\end{array}}$$

Voraussetzungsregel (Vor): Falls P ein Term in $\mathcal{K}$ ist, dann

$$\frac{}{\begin{array}{ll}\mathcal{K} & P\end{array}}$$

$\vee$-Einführung im Antezedenz ($\vee$Ant): Es Sei $\mathcal{K}$ eine Menge von Termen, P, Q, R Terme.

$$\frac{\begin{array}{ll}\mathcal{K} \cup \{P\} & R \\ \mathcal{K} \cup \{Q\} & R\end{array}}{\begin{array}{ll}\mathcal{K} \cup \{P \vee Q\} & R\end{array}}$$

Diese Regel ist korrekt, denn R wird von jeder Interpretation erfüllt, die $\mathcal{K}$ und P erfüllt. Ebenso wird R von jeder Interpretation erfüllt, die $\mathcal{K}$ und Q erfüllt. Also wird R von jeder Interpretation erfüllt, die $\mathcal{K}$ und $P \vee Q$ erfüllt.

Man kann Schlussregeln auch aus anderen Regeln herleiten.

Tertium non datur (TND): Ohne jegliche Voraussetzungen können wir stets $P \vee \neg P$ folgern:

$$\frac{}{P \vee \neg P}$$

Rechtfertigung:

1.	$\{P\}$	P	(Vor)
2.	$\{P\}$	$P \vee \neg P$	($\vee$Suk) auf 1.
3.	$\{\neg P\}$	$\neg P$	(Vor)
4.	$\{\neg P\}$	$P \vee \neg P$	($\vee$Suk) auf 3.
5.		$P \vee \neg P$	(FU) auf 2. und 4.

Modifizierte Widerspruchsregel (Wid'): Es seien $\mathcal{K}$ eine Menge von Termen, P und Q Terme. Dann lautet die Regel

$$\frac{\begin{array}{cc} \mathcal{K} & Q \\ \mathcal{K} & \neg Q \end{array}}{\begin{array}{cc} \mathcal{K} & P \end{array}}.$$

Rechtfertigung: Übung.

Kettenschlussregel (KS):

$$\frac{\begin{array}{cc} \mathcal{K} & P \\ \mathcal{K} \cup \{P\} & Q \end{array}}{\begin{array}{cc} \mathcal{K} & Q \end{array}}$$

Rechtfertigung:

1.	$\mathcal{K}$	P	(Prämisse)
2.	$\mathcal{K} \cup \{\neg P\}$	P	(Ant) auf 1.
3.	$\mathcal{K} \cup \{\neg P\}$	$\neg P$	(Vor)
4.	$\mathcal{K} \cup \{\neg P\}$	Q	(Wid') auf 2. und 3.
5.	$\mathcal{K} \cup \{P\}$	Q	(Prämisse)
6.	$\mathcal{K}$	Q	(FU) auf 4. und 5.

Kontrapositionsregel (KP):

$$\frac{\begin{array}{cc} \mathcal{K} \cup \{\neg Q\} & \neg P \end{array}}{\begin{array}{cc} \mathcal{K} \cup \{P\} & Q \end{array}}$$

Rechtfertigung:

1.	$\mathcal{K} \cup \{\neg Q\}$	$\neg P$	(Prämisse)
2.	$\mathcal{K} \cup \{\neg Q\} \cup \{P\}$	$\neg P$	(Ant) auf 1.
3.	$\mathcal{K} \cup \{\neg Q\} \cup \{P\}$	P	(Vor)
4.	$\mathcal{K} \cup \{\neg Q\} \cup \{P\}$	Q	(Wid') auf 2. und 3.
5.	$\mathcal{K} \cup \{Q\} \cup \{P\}$	Q	(Vor)
6.	$\mathcal{K} \cup \{P\}$	Q	(FU) auf 4. und 5.

Bemerkung 1.32. Formal braucht man vier Kontrapositionsregeln, neben der obigen:

$$\frac{\mathcal{K} \cup \{Q\} \quad P}{\mathcal{K} \cup \{\neg P\} \quad \neg Q}$$

$$\frac{\mathcal{K} \cup \{\neg Q\} \quad P}{\mathcal{K} \cup \{\neg P\} \quad Q}$$

$$\frac{\mathcal{K} \cup \{Q\} \quad \neg P}{\mathcal{K} \cup \{P\} \quad \neg Q}$$

Sie lassen sich analog beweisen.

Es gibt noch viele weitere Regeln, etwa:

$$\frac{\begin{array}{cc}\mathcal{K} & P \vee Q \\ \mathcal{K} & \neg P\end{array}}{\begin{array}{cc}\mathcal{K} & Q\end{array}}$$

Rechtfertigung:

1.	$\mathcal{K}$	$P \vee Q$	(Prämisse)
2.	$\mathcal{K}$	$\neg P$	(Prämisse)
3.	$\mathcal{K} \cup \{P\}$	$\neg P$	(Ant) auf 2
4.	$\mathcal{K} \cup \{P\}$	P	(Vor)
5.	$\mathcal{K} \cup \{P\}$	Q	(Wid') auf 3 und 4
6.	$\mathcal{K} \cup \{Q\}$	Q	(Vor)
7.	$\mathcal{K} \cup \{P \vee Q\}$	Q	($\vee$Ant) auf 5 und 6
8.	$\mathcal{K}$	Q	(KS) auf 1 und 7

Manche Regeln, wie der Modus ponendo ponens sind sehr berühmt:

Modus ponendo ponens (kurz: Modus ponens):
$$\frac{\begin{array}{cc}\mathcal{K} & P \\ \mathcal{K} & P \to Q\end{array}}{\begin{array}{cc}\mathcal{K} & Q\end{array}}$$

Rechtfertigung: Übung.

Um ein vollständiges Kalkül von Schlussregeln zu haben, müsste man noch einige Grundregeln ergänzen, insbesondere für den Umgang mit Quantoren oder der Gleichheit von logischen Termen. [2]

[2] so, bräuchte man u. a. eine Regel

$$\frac{}{P=P},$$

und dann Regeln um aus $P = Q$ auch $Q = P$ zu schließen, etc.

Ein vollständiges Kalkül von Schlussregeln, erlaubt es, die Rechtfertigung von Schlussregeln durch einen Computer zu überprüfen. Dafür gibt es auch entsprechende Beweisassistenzsoftware, etwa `coq`.

1.6 Beweismethoden

In diesem Abschnitt beweisen wir beispielhafte, einfache Sätze, um verschiedene Beweismethoden zu illustrieren. Dabei setzen wir ein wenig Schulmathematik voraus.

Definition 1.33. Eine natürliche Zahl n heißt *gerade*, falls es eine natürliche Zahl k gibt, so dass $n = 2 \cdot k$ gilt. Andernfalls heißt die Zahl n ungerade.

Satz 1.34. *Ist n eine natürliche ungerade Zahl, so ist n^2 ebenfalls ungerade.*

Wir beweisen diesen Satz mittels eines *direkten Beweises*:

Beweis. Eine natürliche Zahl n ist genau dann ungerade, wenn es eine natürliche Zahl k gibt, so dass $n = 2k + 1$ gilt. Wir erhalten dann:

$$\begin{aligned} n^2 &= (2k+1)(2k+1) \\ &= 4k^2 + 4k + 1 \\ &= 2(2k^2 + 2k) + 1 \end{aligned}$$

Mit $k' = 2k^2 + 2k$, haben wir dann $n^2 = 2k' + 1$, also ist n^2 ungerade. □

Satz 1.35. *Ist n eine natürliche Zahl und n^2 gerade, so ist auch n gerade.*

Wir beweisen diesen Satz mittels *Kontraposition*

Beweis. Wir halten zunächst fest, dass eine Zahl entweder gerade oder ungerade ist. Durch die die Anwendung der Kontraposition genügt es zu zeigen, dass wenn n ungerade ist, auch n^2 ungerade ist. Dies haben wir oben gezeigt. □

Satz 1.36. *Die Zahl $\sqrt{2}$ lässt sich nicht als Bruch $\frac{a}{b}$ zweier natürlicher Zahlen a, b schreiben.*

Wir beweisen diesen Satz mit einem *Beweis durch Widerspruch*.

Beweis. Wir nehmen an, dass gilt $\sqrt{2} = \frac{a}{b}$, wobei der Bruch $\frac{a}{b}$ gekürzt ist. Dann würde auch gelten: $2b^2 = a^2$. Dann wäre a^2 eine gerade Zahl und nach obigen Satz ist dann auch a gerade, also $a = 2k$ für eine natürliche Zahl k. Dann ist aber

$$2b^2 = a^2 = (2k)^2 = 4k^2.$$

Das bedeutet, dass $b^2 = 2k^2$ gelten würde. Dann wäre aber b^2 und somit auch b eine gerade Zahl. Sind aber a und b gerade Zahlen,

könnte man im Bruch kürzen, da beide Zahlen durch 2 teilbar wären. Das ist ein Widerspruch zur Annahme, dass der Bruch gekürzt ist. (Also ist die Annahme falsch und das Gegenteil richtig.) □

2

Mengen

In diesem Kapitel wollen wir uns mit den grundlegendsten Objekten der Mathematik befassen, den Mengen. Aufbauend auf der Aussagenlogik ist die Mengenlehre quasi das Fundament der Mathematik. In praktisch allen Bereichen der Mathematik wird man es mit Mengen zu tun haben.

Wir werden in diesem Kapitel viele Analogien zur Aussagenlogik erkennen. Dies wird meist dadurch deutlich, dass die verwendeten Symbole in der Mengenlehre oftmals „abgerundete“ Versionen ihres logischen Pendants sind.

2.1 Definitionen und Beispiele

Wir wollen nun zunächst klären, was eine Menge eigentlich ist und uns dazu Beispiele anschauen. Historisch betrachtet gab es zwar ein intuitives Verständnis, was eine Menge sei, der Begriff wurde aber erst durch den deutschen Mathematiker Georg Cantor (*1845, †1918) Ende des 19. Jahrhunderts formal definiert.

Definition 2.1 (Cantor)**.** Eine *Menge* ist eine Zusammenfassung von bestimmten und wohlunterschiedenen Objekten unserer Anschauung oder unseres Denkens zu einem Ganzen. Die Objekte nennt man *Elemente* der Menge.

Objekte unserer Anschauung sind dabei z. B. Schafe, Murmeln, oder andere (materielle) Dinge, welche uns im Alltag begegnen. *Objekte unseres Denkens* sind hingegen Dinge, welche nicht materiell sein müssen, etwa Tugenden, Sinne, oder auch Zahlen. Der Ausdruck *bestimmt und wohlunterschieden* bedeutet, dass die Objekte einerseits konkret benannt werden können, andererseits auch von einander unterschieden werden können. So soll eine Menge ein Objekt nur einmal enthalten dürfen. *Zusammenfassung zu einem Ganzen* heißt, dass wir die Objekte danach zusammengefasst als ein ganzes oder auch eigenständiges Objekt betrachten können.

Diese Definition, welche Cantor im 19. Jahrhundert anführte, stellte sich später als ungenügend heraus. Durch die Arbeit des britischem

Mathematikers BERTRAND RUSSELL (*1872, †1970), der eine „unmögliche“ Menge konstruierte, wurde zu Beginn des 20. Jahrhundert die Grundlagenkrise der Mathematik ausgelöst, welche sich im Nachhinein als sehr fruchtbar erwies.

Beispiel 2.2 (Russellsche Antinomie)**.** Eine Menge, die nach Definition 2.1 erlaubt wäre, ist die Menge M aller Mengen, welche sich nicht selbst enthalten. Dies sieht zunächst harmlos aus, eine anschauliche Fassung davon ist: „Ein Barbier ist definiert als ein Mensch, der all diejenigen rasiert, welche sich nicht selbst rasieren“. Jetzt fragt man sich: Enthält M sich selbst? Oder anschaulich: Rasiert sich der Barbier selbst?

Man erhält ein Paradoxon. Rasiert sich der Barbier selbst, dann rasiert er sich nach Definition nicht selbst – ein Widerspruch. Angenommen, der Barbier rasiert sich nicht selbst, dann rasiert er sich nach Definition doch – wieder ein Widerspruch. Mengentheoretisch gesprochen: Enthält M sich selbst, dann enthält sich M nach Definition nicht selbst – und umgekehrt.

Um solche *unmöglichen Mengen* zu vermeiden, wandeln wir die Definition leicht ab:

Definition 2.3. Eine *Menge* ist eine wohldefinierte Zusammenfassung von bestimmten und wohlunterschiedenen Objekten unserer Anschauung oder unseres Denkens zu einem Ganzen. Die Objekte nennt man *Elemente* der Menge.

Der Terminus *wohldefinierte Zusammenfassung* bedeutet, dass von jedem Objekt klar sein soll, ob es zu einer gegebenen Menge gehört oder nicht.

Es gibt verschiedene Möglichkeiten Mengen zu notieren. Die Extensionale ist es, die Elemente der Menge aufzulisten und durch geschweifte Klammern zusammenzufassen:

Beispiel 2.4.

- $\{0,1\}$, die Menge, welche die Elemente 0 und 1 enthält.
- $\{a,b,c\}$, die Menge, welche die Elemente a, b und c enthält.
- $\{0, A, \text{Apfel}, \{\text{Pik}, 8\}\}$, die Menge, welche die Elemente 0, A, Apfel und $\{\text{Pik}, 8\}$ enthält. Offenbar können Mengen auch Elemente von anderen Mengen sein.
- $\{0,1,1\}$ – dies ist die selbe Menge wie $\{0,1\}$, da die Elemente einer Menge „wohlunterschieden“ sein sollen.

Streng genommen müssten wir überprüfen, ob die o. a. Beispiele tatsächlich unserer Definition von Mengen entsprechen. Denn zunächst ist überhaupt nicht klar, dass es soetwas wie Mengen überhaupt gibt. Wir könnten die Existenz von Mengen als Axiom fordern, verzichten an dieser Stelle aber darauf.

Häufig wiederkehrende Mengen sind die in der Mathematik viel benutzten Zahlmengen $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ und $\mathbb{C}$.

Notation 2.5. Ist a ein Element der Menge M, schreiben wir $a \in M$ oder $M \ni a$.

Eine weitere Möglichkeit Mengen zu notieren, ist die intensionale:

Beispiel 2.6. Wir betrachten die Menge aller derjenigen reellen Zahlen x, für die gilt, dass $x > 1$ und $x < 2$ ist. Wir schreiben dies verkürzt als:

$$\{x \in \mathbb{R} \mid x > 1 \text{ und } x < 2\}$$

Dabei liest man den senkrechten Strich $\mid$ als *für welche gilt*.

Haben wir Allgemein einen logischen Term $P(x)$ in einer freien Variablen x gegeben, so können wir die Menge

$$\{x \mid P(x)\}$$

konstruieren, welche alle Elemente x enthält, welche den Term $P(x)$ erfüllen.

Dabei kann es aber zu problematischen Konstruktionen kommen [1], daher sollte man immer angeben, aus welcher Menge die x stammen sollen. Ist hingegen A eine Menge, so ist die Menge

$$\{x \in A \mid P(x)\}$$

unproblematisch zu bilden.

[1] So ist die Konstruktion

$$M = \{x \mid \neg(x \in x)\}$$

problematisch. Das Konstrukt M enthält alle Mengen, welche sich nicht selbst enthalten. Wir haben also die Russellsche Antinomie konstruiert.

Um vernünftig mit Mengen „rechnen“ zu können, müssen wir zunächst erklären, wann zwei Mengen eigentlich identisch sind.

Axiom 2.7 (Extensionalitätsaxiom)**.** Zwei Mengen sind gleich, wenn sie dieselben Elemente enthalten.

Enthalten die Mengen sehr viele Objekte (ggf. sogar unendlich viele), oder sind sie in komplizierter Weise beschrieben, ist es nicht immer ganz leicht, zu erkennen, ob zwei Mengen gleich sind. Im nachfolgenden Beispiel ist die Gleichheit aber leicht einzusehen (wirklich!).

Beispiel 2.8. • $\{0, 1, 1\} = \{0, 1\} = \{1, 0\}$.

- $\{0\}$ und $\{\{0\}\}$ sind hingegen nicht gleich.
- $\{n \in \mathbb{N} \mid n < 3\} = \{0, 1, 2\}$.

Wir fordern im nächsten Axiom die Existenz einer etwas absonderlichen Menge, nämlich einer, welche keine Elemente enthält. Warum das notwendig ist, werden wir erst etwas später sehen.

Axiom 2.9 (Leermengenaxiom)**.** Es gibt eine Menge ohne Elemente, genannt die *leere Menge*:

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

Solche Definitionen oder Axiome, deren Sinn sich oftmals nicht sofort erschließt, und welche für Nicht-Mathematiker esoterisch anmuten, werden uns immer wieder begegnen. Meist führt man solche Objekte ein, um später einfachere Aussagen treffen zu können. Die Zahl Null wurde im Mittelalter auch als ein solches esoterisches Konstrukt betrachtet.

2.2 *Mengenoperationen*

Wir wollen nun anfangen mit Mengen zu „rechnen“. Dazu müssen wir Operationen definieren, welche wir auf die Mengen anwenden.

Definition 2.10. Eine Menge A heißt *Teilmenge* der Menge B (geschrieben $A \subseteq B$), wenn jedes Element von A auch Element von B ist:

$$x \in A \Longrightarrow x \in B.$$

Gilt $A \subseteq B$ und $A \neq B$, so schreibt man auch $A \subset B$ und nennt A *echte Teilmenge* von B.

Abbildung 2.1 zeigt eine Menge A (in rot), welche Teilmenge einer Menge B (in grün) ist.

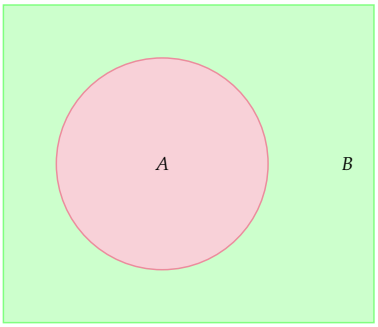


Abbildung 2.1: Die Menge A ist eine Teilmenge der Menge B: $A \subseteq B$.

Beispiel 2.11.

- Jede Menge ist Teilmenge von sich selbst: $A \subseteq A$.
- Ebenso enthält jede Menge die leere Menge als Teilmenge: $\emptyset \subseteq A$.
- $\{x, y\}$ ist keine Teilmenge der Menge $\big\{\{x, y\}, z\big\}$, sondern nur ein Element.

Man möchte nun alle möglichen Teilmengen einer Menge zu einer neuen Menge zusammenfassen. Die folgende Definition ist eigentlich ein Axiom und besagt, dass dies möglich ist.

Definition 2.12. Für jede Menge A gibt es eine Menge $\mathcal{P}(A)$, genannt die *Potenzmenge* von A, die alle Teilmengen von A als Elemente enthält.

Beispiel 2.13.

- $\mathcal{P}(\emptyset) = \{\emptyset\}$.
- $\mathcal{P}(\{1,2\}) = \big\{\emptyset, \{1\}, \{2\}, \{1,2\}\big\}$.

Notation 2.14. Ist A eine Menge, dann beschreibt $|A|$ die Anzahl der Elemente von A. Wir nennen $|A|$ auch die *Kardinalität* oder *Mächtigkeit* von A.

Beispiel 2.15.

1. $|\emptyset| = 0$,
2. $|\{1,2\}| = 2$,
3. $|\mathbb{N}| = \infty$ (unendlich).

Definition 2.16. Die *Vereinigung* $A \cup B$ zweier Mengen A und B, ist die Menge, die alle Elemente von A und aller Elemente von B enhält:

$$A \cup B = \{x \mid x \in A \vee x \in B\}.$$

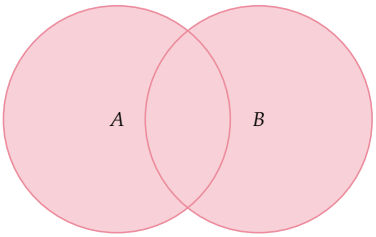


Abbildung 2.2: Vereinigung $A \cup B$ zweier Mengen A und B.

Abbildung 2.2 zeigt anschaulich, wie die Vereinigung zweier Mengen A und B aussieht. Die Vereinigungsmenge $A \cup B$ entspricht dem gefüllten Bereich.

Die Ähnlichkeit der Symbole $\vee$ und $\cup$ kommt nicht von ungefähr.

Beispiel 2.17.

- $A \cup \emptyset = A = A \cup A$
- $\{1,2\} \cup \{p,q\} = \{1,2,p,q\}$
- $\{1,2\} \cup \{2,3\} = \{1,2,3\}$

Definition 2.18. Der *Durchschnitt* $A \cap B$ zweier Mengen A und B ist die Menge aller Elemente, welche sowohl in A als auch in B enthalten sind:

$$A \cap B = \{x \mid x \in A \wedge x \in B\}.$$

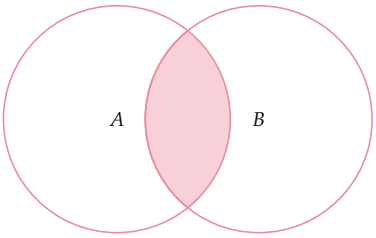


Abbildung 2.3: Durchschnitt $A \cap B$ zweier Mengen A und B.

Abbildung 2.3 zeigt anschaulich, wie der Durchschnitt zweier Mengen A und B aussieht. Die Vereinigungsmenge $A \cap B$ entspricht dem gefüllten Bereich.

Beispiel 2.19.

- $A \cap A = A$
- $A \cap \emptyset = \emptyset$
- $\{1,2,3\} \cap \{3,4,5\} = \{3\}$
- $\{1,2\} \cap \{p,q\} = \emptyset$

Definition 2.20. Die *Differenzmenge* $A \setminus B$ zweier Mengen A und B ist die Menge der Elemente von A, die nicht in B enthalten sind:

$$A \setminus B = \{x \in A \mid x \notin B\} = \{x | x \in A \wedge x \notin B\}.$$

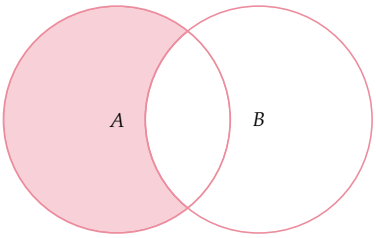


Abbildung 2.4: Differenz $A \setminus B$ zweier Mengen A und B.

Dabei bedeutet $a \notin A$, dass a nicht in A enthalten ist.

Abbildung 2.4 zeigt anschaulich, wie die Differenzmenge $A \setminus B$ zweier Mengen A und B aussieht. Die Differenzmenge $A \setminus B$ entspricht dem gefüllten Bereich.

Beispiel 2.21.

- $A \setminus A = \emptyset$
- $A \setminus \emptyset = A$
- $\{1,2,3\} \setminus \{3,4,5\} = \{1,2\}$

- $\{3,4,5\} \setminus \{1,2,3\} = \{4,5\}$

Im Gegensatz zur Vereinigung und dem Durchschnitt von Mengen ist es für die Bildung der Differenzmenge erheblich, welche Menge links und welche rechts des Operationssymbols steht. Im Allgemeinen gilt:

$$A \setminus B \neq B \setminus A.$$

Ist B eine Teilmenge von A, so nennt man $A \setminus B$ das *Komplement* von B in A und schreibt dafür $\complement B$, oder $\overline{B}$.

Abbildung 2.5 zeigt anschaulich, wie das Komplement $\complement A$ einer Menge A in einer Menge B aussieht. Das Komplement entspricht dem gefüllten Bereich.

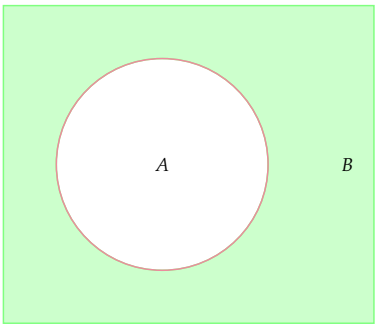


Abbildung 2.5: Das Komplement von A in B: $\complement A$.

Definition 2.22. Die *symmetrische Differenz* $A \triangle B$ zweier Mengen A und B ist die Menge aller Elemente von $A \cup B$, welche nicht in $A \cap B$ liegen:

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

Abbildung 2.6 zeigt anschaulich, wie die symmetrische Differenz zweier Mengen A und B aussieht. Die Menge $A \triangle B$ entspricht dem gefüllten Bereich.

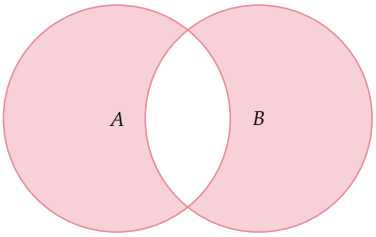


Abbildung 2.6: Symmetrische Differenz $A \triangle B$ zweier Mengen A und B.

Beispiel 2.23. $\{1,2,3\} \triangle \{3,4,5\} = \{1,2,4,5\}$.

Definition 2.24. Zwei Mengen A und B heißen *disjunkt*, falls gilt: $A \cap B = \emptyset$.

Proposition 2.25. *Seien A und B beliebige Mengen. Dann gelten:*

1. $A = (A \setminus B) \cup (A \cap B)$;
2. $(A \setminus B) \cap (A \cap B) = \emptyset$.

Beweis. 1. Zu zeigen ist die logische Äquivalenz

$$x \in A \iff x \in \big((A \setminus B) \cup (A \cap B)\big).$$

Durch eine Kette von logischen Äquivalenzumformungen zeigen wir dies nun.

$$\begin{aligned} & x \in (A \setminus B) \cup (A \cap B) \\ \Leftrightarrow\ & \big(x \in A \setminus B)\big) \vee \big(x \in (A \cap B)\big) \\ \Leftrightarrow\ & (x \in A \wedge x \notin B) \vee (x \in A \wedge x \in B) \\ \Leftrightarrow\ & (x \in A) \wedge (x \notin B \vee x \in B) \\ \Leftrightarrow\ & x \in A. \end{aligned}$$

2. Die zweite Aussage beweisen wir auf die gleiche Art:

$$\begin{aligned} & x \in \big((A \setminus B) \cap (A \cap B)\big) \\ \Leftrightarrow\ & x \in (A \setminus B) \wedge x \in (A \cap B) \\ \Leftrightarrow\ & (x \in A \wedge x \notin B) \wedge (x \in A \wedge x \in B) \\ \Leftrightarrow\ & x \in A \wedge x \notin B \wedge x \in A \wedge x \in B \\ \Leftrightarrow\ & (x \in A \wedge x \in A) \wedge \underbrace{(x \in B \wedge x \notin B)}_{\text{Kontradiktion}} \\ \Leftrightarrow\ & x \in \emptyset \end{aligned}$$

□

Während die Elemente in Mengen nicht geordnet sind, d. h. dass $\{a, b\}$ die selbe Menge wie $\{b, a\}$ ist, benötigt man oft eine Ordnung der Elemente um festzulegen, welches das erste, und welches das zweite Element (usw.) ist.

Definition 2.26. Für zwei Elemente x, y einer beliebigen Menge M ist das *geordnete Paar* (x, y) mit x als ersten Eintrag und y als zweiten Eintrag gegeben durch

$$(x, y) := \big\{\{x\}, \{x, y\}\big\}.$$

Durch diese Definition, ist die Reihenfolge der beiden Elemente klar. Die Menge $\big\{\{x\}, \{x, y\}\big\}$ heißt *Kuratowski-Paar*.

Satz 2.27. *Zwei geeordnete Paare (a, b) und (x, y) sind genau dann gleich, wenn $a = x$ und $b = y$ gilt:*

$$(a, b) = (x, y) \Longleftrightarrow a = x \wedge b = y$$

Beweis. Übungsaufgabe □

Beispiel 2.28. Aus der Schule kennt man die Koordinatendarstellung in der Ebene als Beispiel für geordnete Paare.

Definition 2.29. Das *kartesische Produkt* $A \times B$ zweier Mengen A und B ist die Menge aller geordneten Paare (a, b) mit $a \in A$ und $b \in B$:

$$A \times B := \big\{(a, b) \mid a \in A, b \in B\big\}.$$

Beispiel 2.30.

1. $\{1, 2\} \times \{a, b\} = \big\{(1, a), (1, b), (2, a), (2, b)\big\}$
2. $\mathbb{R} \times \mathbb{R} = \mathbb{R}^2 = \big\{(x, y) \mid x \in \mathbb{R}, y \in \mathbb{R}\big\}$
3. $A \times \emptyset = \emptyset \times A = \emptyset$.
4. Im Allgemeinen gilt: $A \times B \neq B \times A$.

Die bisherigen Mengenoperationen $\cup, \cap, \times$ haben wir stets für die Verknüpfung von zwei Mengen benutzt. Oft benötigt man aber mehrere Mengen, mit denen man arbeitet. Wir betrachten nun, wie man mit Systemen von Mengen rechnet.

Notation 2.31.

1. Haben wir nur endlich viele Mengen, so können wir diese „einfach“ aufzählen:
$$M_1, M_2, \ldots, M_n$$
2. Sind unendliche viele Mengen gegeben, so können wir diese nicht mehr aufzählen. Lassen sich die Mengen abzählen, so schreiben wir
$$M_1, M_2, \ldots$$
um anzudeuten, dass wir die Mengen mit natürlichen Zahlen indexieren.
3. Ist I allgemein irgendeine Menge, so können wir ein System von Mengen mit den Elementen von I indexieren und schreiben
$$(M_i \mid i \in I) \text{ oder } (M_i)_{i \in I}.$$

Definition 2.32. Ein Mengensystem $(M_i \mid i \in I)$ heißt *Familie von Mengen*.

Für Familien von Mengen können wir Durchschnitt und Vereinigungen bilden:

Definition 2.33.

1. Der Durchschnitt einer Familie von Mengen ist definiert als
$$\bigcap_{i \in I} M_i := \{x \mid x \in M_i, \text{ für alle } i \in I\}.$$
2. Analog defineren wir die Vereinigung einer Familie von Mengen:
$$\bigcup_{i \in I} M_i := \{x \mid \text{ es gibt ein } i \in I \text{ mit } x \in M_i\}.$$

Proposition 2.34. *Sei $(M_i \mid i \in I)$ eine Familie von Mengen. Dann gilt*
$$\bigcup_{i \in I} \mathcal{P}(M_i) \subseteq \mathcal{P}\left(\bigcup_{i \in I} M_i\right).$$

Beweis. Sei $A \in \bigcup_{i \in I} \mathcal{P}(M_i)$. Dann gibt es ein $i \in I$, für das gilt: $A \in \mathcal{P}(M_i)$. Also ist $A \subseteq M_i$ und damit gilt $A \subseteq \bigcup_{i \in I} M_i$ Daher ist $A \in \mathcal{P}\left(\bigcup_{i \in I} M_i\right)$. □

Die Gleichheit gilt im Allgemeinen nicht.

3
Relationen und Abbildungen

In diesem Kapitel geht es um *Relationen*, also Beziehungen zwischen zwei Objekten. Relationen tauchen in allen Bereichen der Mathematik auf, insbesondere die sogenannten *Äquivalenzrelationen* und *Ordnungsrelationen* sind besonders häufig anzutreffen. Auch die überall anzutreffenden *Abbildungen* sind spezielle Relationen. In der Informatik tauchen Relationen z.B. im Zusammenhang mit Datenbanken auf.

3.1 Grundbegriffe

Definition 3.1. Seien A und B Mengen. Dann heißt jede Teilmenge $R \subseteq A \times B$ *Relation* zwischen A und B. Ist $A = B$, also $R \subseteq A \times A$, dann heißt R *Relation auf* A. Gilt $(a, b) \in R$, so sagt man, „a steht in Relation zu b“.

Formal gesehen ist dies die Definition einer zweistelligen Relation. Wir können uns auch mehrstellige Relationen vorstellen, also Relationen zwischen mehr als zwei Objekten. In diesem Kapitel beschränken wir uns zunächst auf zweistellige Relationen.

Oft schreiben wir aRb statt $(a, b) \in R$. Dies machen wir insbesondere dann, wenn wir ein spezielles Relationssymbol benutzen. Zum Beispiel können wir auf einer Menge M die Gleichheitsrelation betrachten, also jene Relation, bei der zwei Elementen aus M in Relation zu einander stehen, wenn sie gleich sind. Dann ist es anschaulicher (und lesbarer) $a = b$ statt $(a, b) \in =$ und $= \subseteq M \times M$ zu schreiben.

Da jede Teilmenge des kartesischen Produkts von zwei Mengen eine Relation ist, kann man sich vorstellen, dass nicht jede Relation nützlich ist. Eine spezielle (und wenig nützliche) Relation ist die leere Relation $\emptyset = \emptyset \times \emptyset \subseteq A \times B$. Auch die Allrelation $A \times B \subseteq A \times B$, welche alle Elemente des kartesischen Produkts enthält ist meist wenig nützlich.

Einige besonders nützliche Relationen wollen wir im Folgenden beschreiben. Wir definieren zunächst einige Eigenschaften, welche eine Relation besitzen kann. Dabei beschränken wir uns für diesen Ab-

schnitt auf Relationen auf einer Menge, also Teilmengen $R \subseteq A \times A$.

Definition 3.2. Sei $R \subseteq A \times A$ eine Relation auf A. R heißt

a) *reflexiv*, genau dann wenn für alle $x \in A$ gilt: xRx;

b) *symmetrisch*, genau dann wenn für alle $x, y \in A$ gilt: $xRy \Rightarrow yRx$;

c) *antisymmetrisch*, genau dann wenn für alle $x, y \in A$ gilt:

$$xRy \wedge yRx \Rightarrow x = y;$$

d) *transitiv*, genau dann wenn für alle $x, y, z \in A$ gilt:

$$xRy \wedge yRz \Rightarrow xRz;$$

e) *total*, genau dann wenn für alle $x, y \in A$ gilt: $xRy \vee yRx$.

Bei der Verwendung der obigen Begriffe ist darauf zu achten, dass *antisymmetrisch* etwas anderes bedeutet als *nicht symmetrisch*. Insbesondere schließen sich die beiden Begriffe *symmetrisch* und *antisymmetrisch* nicht aus. Eine Relation kann symmetrisch und antisymmetrisch zugleich sein.

Beispiel 3.3. Die Gleichheitsrelation auf der Menge $M = \{1, 2, 3\}$ ist sowohl symmetrisch als auch antisymmetrisch. Die Relation hat folgende Gestalt:

$$\big\{(1,1), (2,2), (3,3)\big\}.$$

Daher gilt für alle $x, y \in M$:

$$x = y \Longrightarrow y = x \quad \text{, und}$$
$$x = y \wedge y = x \Longrightarrow x = y.$$

Im Übrigen ist die Relation auch reflexiv und transitiv. Reflexivität ist leicht einzusehen, für die Transitivität betrachtet man:

$$x = x \wedge x = x \Longrightarrow x = x,$$

für alle $x \in M$.

Wir kommen nun zu drei wichtigen Arten von Relationen, welche wir im Folgenden immer wieder benutzen werden.

Definition 3.4. Eine Relation R auf A heißt

1. *Äquivalenzrelation*, genau dann wenn R reflexiv, symmetrisch und transitiv ist;
2. *Halbordnung*, genau dann wenn R reflexiv, antisymmetrisch und transitiv ist;
3. *totale Ordnung*, genau dann wenn R eine Halbordnung und total ist.

Beispiel 3.5. 1. Die Relation „$=$“ auf einer beliebigen, nicht-leeren Menge A is eine Äquivalenzrelation.

2. Die Relation „$\leq$“ auf den natürlichen Zahlen ist eine totale Ordnung.

3. Die Relation „ist teilbar durch“ auf den natürlichen Zahlen ist eine Halbordnung.

4. Die Relation „ist Teilmenge von“ auf der Potenzmenge einer Menge ist eine Halbordnung.

3.2 *Äquivalenzrelationen*

In diesem Abschnitt werden wir wichtige Eigenschaften von Äquivalenzrelationen beleuchten, die oftmals nützlich sein werden.

Definition 3.6. Sei R eine Äquivalenzrelation auf A. Zu jedem $x \in A$ heißt die Menge

$$\overline{x} := \{y \in A \mid yRx\}$$

die *Äquivalenzklasse* von x.

Satz 3.7. *Sei A eine Menge und R eine Äquivalenzrelation auf A und xRy. Dann gilt: $\overline{x} = \overline{y}$.*

Beweis. Sei $z \in \overline{x}$. Dann ist zRx und wegen der Transitivität von R auch zRy. Nach Definition 3.6 ist dann $z \in \overline{y}$, also $\overline{x} \subseteq \overline{y}$.

Sei umgekehrt $z \in \overline{y}$, dann ist zRy. Nach Voraussetzung ist xRy, also wegen der Symmetrie von R auch yRx. Wegen der Transitivität von R ist dann zRx, und somit $z \in \overline{x}$. Damit ist $\overline{y} \subseteq \overline{x}$.

Aus $\overline{x} \subseteq \overline{y}$ und $\overline{y} \subseteq \overline{x}$ folgt $\overline{x} = \overline{y}$. □

Satz 3.8. *Sei A eine Menge und R eine Äquivalenzrelation auf A und $x, y \in A$. Dann gilt:*

$$\overline{x} = \overline{y} \quad \textit{oder} \quad \overline{x} \cap \overline{y} = \emptyset.$$

Zwei Äquivalenzklassen einer Äquivalenzrelation sind also entweder gleich oder disjunkt.

Beweis. Wir nehmen an, dass der Schnitt nicht leer ist. Dann gibt es ein $z \in \overline{x} \cap \overline{y}$. Dann gelten $z \in \overline{x}$ und $z \in \overline{y}$, also folgt zRx und zRy. Wegen der Symmetrie und Transitivität von R gilt dann xRy und nach Satz 3.7: $\overline{x} = \overline{y}$.

Ist der Schnitt also nicht leer, so sind die beiden Äquivalenzklassen gleich. □

Das folgende Korollar ist fast schon die nützlichere Aussage, wie wir gleich sehen werden.

Korollar 3.9. *Sei A eine Menge und R eine Äquivalenzrelation auf A. Dann ist A die disjunkte Vereinigung der Äquivalenzklassen von R:*

$$A = \bigcup_{x \in A} \overline{x}. \tag{3.1}$$

Beweis. Wegen der Reflexivität von R ist keine Äquivalenzklasse leer und jedes Element von A liegt in einer Äquivalenzklasse. Die Disjunktheit der Klassen folgt dann aus Satz 3.8 □

Definition 3.10. Sei A eine Menge. Die Menge $P \subseteq \mathcal{P}(A)$ heißt *Partition* von A, wenn gilt:

(P1) Für alle $M \in P$ gilt: $M \neq \emptyset$;

(P2) Für alle $M, N \in P$ gilt $M = N$ oder $M \cap N = \emptyset$;

(P3) Für alle $x \in A$ gilt: Es gibt ein $M \in P$ so dass $x \in M$.

Ein System von Mengen, welche die Eigenschaft (P3) erfüllt, nennt man auch *Überdeckung* von A.

Beispiel 3.11. Sei die Menge $A = \{1, 2, 3\}$ gegeben. Dann ist $\{\{1\}, \{2\}, \{3\}\}$ ist eine Partition von A. Ebenso ist $\{\{1, 2\}, \{3\}\}$ ist eine Partition von A. Hingegen ist $\{\{1, 2\}, \{2, 3\}\}$ keine Partition von A (oder sonst irgendeiner Menge).

Der folgende Satz stellt eine Beziehung zwischen Partitionen und Äquivalenzklassen her.

Satz 3.12. *Sei R eine Äquivalenzrelation auf der Menge A. Sei A/R die Menge aller Äquivalenzklassen von R. Dann gilt: A/R ist eine Partion von A.*

Beweis. Zu zeigen ist, dass $A/R = \{\overline{x} \mid x \in A\}$ die drei Eigenschaften aus Definition 3.10 besitzt.

Für alle $\overline{x} \in A/R$ gilt: $x \in \overline{x}$. Also sind alle Mengen in A/R nichtleer. Damit ist (P1) gezeigt.

Die Eigenschaft (P2) folgt aus Satz 3.8.

Die Menge aller Äquivalenzklassen A/R bildet eine Überdeckung von A, da für alle $x \in A$ gilt: $x \in \overline{x}$. Damit ist auch (P3) gezeigt. □

Jede Äquivalenzrelation liefert also eine Partition. Tatsächlich gilt auch die Umkehrung:

Satz 3.13. *Sei P eine Partition einer Menge A. Die Relation $R \subseteq A \times A$ sei gegeben durch:*

$$xRy \Leftrightarrow \exists M \in P : x \in M \wedge y \in M.$$

Dann ist R eine Äquivalenzrelation.

Der Beweis verbleibt zur Übung der geneigten Leserin überlassen.

3.3 *Abbildungen*

Abbildungen sind spezielle Relationen. Bisher haben wir nur Relationen auf einer Menge betrachtet. Nun widmen wir uns Relation zwischen zwei (nicht notwendigerweise) verschiedenen Mengen.

Definition 3.14. Eine Relation $R \subseteq M \times N$ zwischen zwei Mengen M und N heißt

- *funktional* (oder rechtseindeutig), wenn für alle $x \in M$ und $y, z \in N$ gilt:
$$xRy \wedge xRz \Rightarrow y = z.$$
- *linkstotal*, wenn gilt:
$$M = \{x \in M \mid \exists y \in N : xRy\}.$$

Funktional bedeutet also, dass jedem Element der Menge M zu höchstens einem Element der Menge N in Relation steht. Linkstotal bedeutet, dass jedes Element der Menge M zu mindestens einem Element der Menge N in Relation steht.

Definition 3.15. Eine linkstotale und rechtseindeutige Relation $F \subseteq M \times N$ zwischen zwei Mengen M und N heißt *Abbildung* von M nach N. Wir schreiben:
$$F \colon M \longrightarrow N.$$

Da F rechtseindeutig ist, schreiben wir statt xFy auch $y = F(x)$. In diesem Zusammenhang nennt man y das *Bild* von x (unter F) und x ein *Urbild* von y.

Offensichtlich gilt $F = \{(x, F(x)) \mid x \in M\}$.

Notation 3.16. Wir bezeichnen die zu Abbildungen gehörenden Relationen auch mit kleinen lateinischen Buchstaben (üblicherweise: $f, g, \ldots$). Wir schreiben
$$\begin{aligned} f \colon M &\longrightarrow N \\ x &\longmapsto f(x). \end{aligned}$$
und meinen damit, dass f die Abbildung von M nach N ist, welche x auf $f(x)$ abbildet. Dabei heißt $f(x)$ die *Abbildungsvorschrift* von f.

Beispiel 3.17. Wir betrachten einige Abbildungen. Auf den Nachweis, dass es sich tatsächlich um Abbildungen handelt, verzichten

wir an dieser Stelle.

$$\begin{aligned} a\colon \mathbb{N} &\longrightarrow \mathbb{N} \\ n &\longmapsto n\text{-te Nachkommastelle von } \pi = 3,1415926\ldots \end{aligned}$$

$$\begin{aligned} b\colon \mathbb{Z} &\longrightarrow \mathbb{N} \\ z &\longmapsto |z| = \begin{cases} z, & \text{falls } z \geq 0 \\ -z, & \text{sonst} \end{cases} \end{aligned}$$

$$\begin{aligned} c\colon \mathbb{Z} &\longrightarrow \mathbb{Z} \\ z &\longmapsto |z| \end{aligned}$$

$$\begin{aligned} d\colon \mathbb{Q} &\longrightarrow \mathbb{Q} \\ x &\longmapsto \begin{cases} \frac{|x|}{x}, & \text{falls } x \neq 0 \\ 0, & \text{sonst} \end{cases} \end{aligned}$$

$$\begin{aligned} e\colon \{1,2,3\} &\longrightarrow \mathbb{N} \\ k &\longmapsto k \end{aligned}$$

$$\begin{aligned} f\colon \mathbb{Z} &\longrightarrow \mathbb{N} \\ m &\longmapsto \begin{cases} 2m, & \text{falls } m \geq 0 \\ -(2m+1), & \text{sonst} \end{cases} \end{aligned}$$

$$\begin{aligned} g\colon \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\longmapsto x+1 \end{aligned}$$

$$\begin{aligned} h\colon \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\longmapsto x^2 \end{aligned}$$

$$\begin{aligned} i\colon \mathbb{Q} &\longrightarrow \mathbb{R} \\ x &\longmapsto 2^x \end{aligned}$$

$$\begin{aligned} l\colon \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\longmapsto x^3 - x \end{aligned}$$

Hingegen ist

$$\begin{aligned} m\colon \mathbb{N} &\longrightarrow \{1,2,3\} \\ n &\longmapsto n \end{aligned}$$

keine Abbildung. (Warum nicht?)

Abbildungen kann man verknüpfen. Sind

$$f\colon A \longrightarrow B$$

und

$$g\colon B \longrightarrow C$$

Abbildungen, dann ist

$$\begin{aligned} (g \circ f)\colon A &\longrightarrow C \\ x &\longmapsto g\big(f(x)\big) \end{aligned}$$

die *Komposition / Hintereinanderausführung* von f und g.

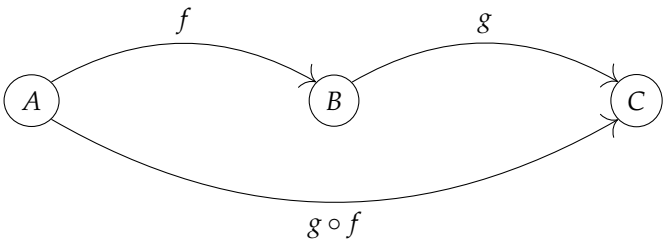


Abbildung 3.1: Komposition von Abbildungen.

Beispiel 3.18. Seien die Abbildungen

$$\begin{aligned} f\colon \mathbb{N} &\longrightarrow \mathbb{Z} \\ n &\longmapsto 2 \cdot n \end{aligned} \qquad \text{und}$$

$$\begin{aligned} g\colon \mathbb{Z} &\longrightarrow \mathbb{R} \\ m &\longmapsto m^2 \end{aligned}$$

gegeben. Dann ist die Abbildung $g \circ f\colon \mathbb{N} \longrightarrow \mathbb{R}$ gegeben durch:
$(g \circ f)(n) = g\big(f(n)\big) = (2 \cdot n)^2 = 4 \cdot n^2$

Ein anderes (lebensweltlicheres) Beispiel ist das Folgende.

Beispiel 3.19. Sei A die Menge von Studierenden, welche eine bestimmte Vorlesung besuchen, B die Menge der Tutorien zu dieser Vorlesung und C die Menge der Tutorinnen, welche die Tutorien leiten. Betrachten wir die Abbildung $f\colon A \longrightarrow B$, welche jeder Studierenden ein Tutorium zuweist, und die Abbildung $g\colon B \longrightarrow C$, welche jedem Tutorium eine Tutorin zuweist. Dann ist die Verknüpfung $g \circ f$ die Abbildung, welche jeder Studierenden die Tutorin ihres Tutoriums zuweist.

Definition 3.20. Sei $f\colon M \longrightarrow N$ eine Abbildung. Für eine Teilmenge $A \subseteq M$ ist das *Bild* von A (unter f) die Menge

$$f(A) := \{y \in N | \exists x \in A : f(x) = y\}.$$

Für eine Teilmenge $B \subseteq N$ ist die *Urbildmenge* von B die Menge

$$f^{-1}(B) := \{x \in M \mid f(x) \in B\}.$$

Definition 3.21. Die Abbildung $f\colon X \longrightarrow Y$ heißt

- *injektiv*, falls für alle $x_1, x_2 \in X$ gilt: $f(x_1) = f(x_2) \Rightarrow x_1 = x_2$.
- *surjektiv*, falls $f(X) = Y$ ist, d. h. falls es für jedes $y \in Y$ ein $x \in X$ gibt, so dass gilt: $f(x) = y$.
- *bijektiv*, falls f injektiv und surjektiv ist, d. h. falls es für jedes $y \in Y$ genau ein $x \in X$ gibt, so dass gilt: $f(x) = y$.

Beispiel 3.22. Wir kommen zurück zu den Abbildungen aus Beispiel 3.17. Die Abbildungen *a*, *c* und *d* sind weder injektiv noch surjektiv. Die Abbildungen *b* ist surjektiv, aber nicht injektiv. Die Abbildung *e* ist injektiv aber nicht surjektiv. Die Abbildungen f und g sind bijektiv. Über die anderen Abbildungen wollen wir an dieser Stelle nicht sprechen.

Oftmals möchte man Abbildungen graphisch darstellen.

Definition 3.23. Sei $f\colon X \longrightarrow Y$ eine Abbildung. Dann heißt

$$\{(x, f(x)) \mid x \in X\} \subseteq X \times Y$$

Graph der Abbildung f.

Der Graph einer Abbildung ist also zunächst einmal nur die Menge der geordneten Paare (man könnte auch Koordinaten sagen), welche zur Abbildung gehören. Es ist also nichts anderes als eine vollständige (ggf. unendlich lange) Wertetabelle der Abbildung. Wenn man nun in ein Koordinatensystem diese Koordinaten einträgt, erhält man ein *Schaubild des Graphen der Abbildung*.

Notation 3.24. Die Abbildung

$$\begin{aligned} id_X\colon X &\longrightarrow X \\ x &\longmapsto x \end{aligned}$$

heißt die *identische Abbildung* auf X.

Definition 3.25. Zwei Mengen X und Y heißen *gleichmächtig*, wenn es eine bijektive Abbildung $f\colon X \longrightarrow Y$ gibt. Wir schreiben dann $|X| = |Y|$.

Beispiel 3.26. Wie wir in Beispiel 3.22 gesehen haben, gibt es eine bijektive Abbildung von $\mathbb{Z}$ nach $\mathbb{N}$. Also gilt $|\mathbb{N}| = |\mathbb{Z}|$.

Satz 3.27. *Sei $f\colon X \longrightarrow Y$ eine Abbildung und A, B Teilmengen von X sowie C, D Teilmengen von Y. Dann gelten:*

1) $f(A \cap B) \subseteq f(A) \cap f(B)$,

2) $f(A \cup B) = f(A) \cup f(B)$,

3) $f(X \setminus A) \supseteq f(X) \setminus f(A)$,

4) $f^{-1}\big(f(A)\big) \supseteq A$,

5) $f^{-1}(C \cap D) = f^{-1}(C) \cap f^{-1}(D)$,

6) $f^{-1}(C \cup D) = f^{-1}(C) \cup f^{-1}(D)$,

7) $f^{-1}(Y \setminus C) = X \setminus f^{-1}(C)$,

8) $f\big(f^{-1}(C)\big) \subseteq C$.

Beweis. Wir beweisen an dieser Stelle nur einige der obigen Behauptungen. Den Rest wird die geneigte Leserin als leichte Übungsaufgabe lösen können.

1) Sei $y \in f(A \cap B)$. Dann gibt es ein $x \in A \cap B$, so dass gilt: $f(x) = y$. Also gilt $x \in A \wedge x \in B$ und somit:

$$f(x) \in f(A) \wedge f(x) \in f(B).$$

Damit ist $f(x) \in f(A) \cap f(B)$ und somit $y \in f(A) \cap f(B)$.

2) Sei $y \in f(A \cup B)$. Dann gibt es ein $x \in A \cup B$, so dass gilt: $f(x) = y$. Also gilt $x \in A \vee x \in B$ und somit:

$$f(x) \in f(A) \vee f(x) \in f(B).$$

Damit ist $f(x) \in f(A) \cup f(B)$ und somit $y \in f(A) \cup f(B)$.

Sei nun umgekehrt $y \in f(A) \cup f(B)$. Also gilt $y \in f(A) \vee y \in f(B)$. Somit gibt es ein $x \in A$ oder ein $x \in B$ mit $f(x) = y$. Damit gibt es ein $x \in A \cup B$ mit $f(x) = y$ und somit ist $y \in f(A \cup B)$.

4) Sei $a \in A$. Dann ist $f(a) \in f(A)$. Nach Definition des Urbildes ist $a \in f^{-1}\big(f(A)\big)$.

8) Sei $y \in f\big(f^{-1}(C)\big)$. Dann gibt es ein $x \in f^{-1}(C)$ mit $f(x) = y$. Nach Definition des Urbildes ist dann aber $f(x) \in C$ und somit gilt: $y \in C$.

□

Satz 3.28. *Seien $f\colon X \longrightarrow Y$ und $g\colon Y \longrightarrow Z$ Abbildungen. Dann gilt:*

1) *Sind f und g injektiv, dann auch $g \circ f$.*

2) *Ist $g \circ f$ injektiv, so auch f.*

3) *Ist $g \circ f$ surjektiv, so auch g.*

4) *Ist f surjektiv und $g \circ f$ injektiv, so ist g injektiv.*

5) *Ist g injektiv und $g \circ f$ surjektiv, so ist f surjektiv.*

Beweis. 1) Seien $x, y \in X$, wobei $x \neq y$ gelte. Da f injektiv ist, folgt $f(x) \neq f(y)$. Aus der Injektivität von g folgt dann $g\big(f(x)\big) \neq g\big(f(y)\big)$, d. h. $(g \circ f)(x) \neq (g \circ f)(y)$. Also ist $g \circ f$ injektiv.

2) Seien $x, y \in X$, wobei $x \neq y$ gelte. Da $g \circ f$ injektiv ist, ist $g\big(f(x)\big) \neq g\big(f(y)\big)$. Da es sich bei g um eine Abbildung handelt, folgt $f(x) \neq f(y)$. Also ist f injektiv.

3) Sei $z \in Z$. Zu zeigen ist, dass es ein $y \in Y$ für das $g(y) = z$ gilt. $g \circ f$ ist surjektiv. Also gibt es ein $x \in X$ für das gilt: $g\big(f(x)\big) = z$. Wir setzen $y := f(x)$, dann ist $g(y) = z$ und daher ist g surjektiv.

4) Diesen Teil überlassen wir der geneigten Leserin als Übungsaufgabe.

5) Sei $y \in Y$. Wir müssen zeigen, dass es ein $x \in X$ gibt, für das $f(x) = y$ gilt. Zu $g(y)$ gibt es ein $x \in X$ mit $(g \circ f)(x) = g(y)$, da $g \circ f$ surjektiv ist. Da g injektiv ist, folgt $y = f(x)$.

□

Wir werden in den nächsten Kapiteln immer wieder die unterschiedlichsten Abbildungen nutzen. In Mathematische Grundlagen 2 werden wir uns intensiv dem Studium der Abbildungen widmen, hier im Schwerpunkt der Abbildungen auf $\mathbb{R}$.

4
Natürliche Zahlen und Vollständige Induktion

In diesem Kapitel beschäftigen wir uns vorwiegend mit der Beweismethode der vollständigen Induktion. Viele Aussagen über natürliche Zahlen können damit bewiesen werden. Insbesondere werden wir die natürlichen Zahlen axiomatisieren und aus diesen Axiomen alle bekannten Rechengesetze der natürlichen Zahlen folgern.

Damit wir auch einige beispielhafte Sätze beweisen können, benötigen wir zunächst ein paar Grundbegriffe, welche wir nun einführen.

Notation 4.1. Sind $a_0, a_1, \ldots, a_n$ (natürliche, ganze, rationale, reelle, oder komplexe) Zahlen, dann schreiben wir statt $a_0 + a_1 + a_2 + \cdots + a_n$ oftmals

$$\sum_{i=0}^{n} a_i,$$

und lesen dies als die *Summe über die a_i von $i = 0$ bis n*. Statt $a_0 \cdot a_1 \cdots a_n$ schreiben wir oft

$$\prod_{i=0}^{n} a_i,$$

und lesen dies als das *Produkt über die a_i von $i = 0$ bis n*.

Statt bei $i = 0$ können wir auch bei jeder anderen natürlichen Zahl anfangen zu summieren (oder das Produkt bilden). Tatsächlich können wir den Index i auch aus einer beliebigen Indexmenge wählen, siehe Kapitel 2.

Wir sollten am Rande noch kurz darüber nachdenken, was wir erhalten, falls der Startindex größer ist, als der Endindex, z. B. bei

$$\sum_{i=1}^{0} a_i, \text{ oder } \prod_{i=1}^{0} a_i.$$

Diese Summe (bzw. dieses Produkt) enthält keine Summanden (bzw. Faktoren), wir sprechen in diesem Zusammenhang von der *leeren Summe* (bzw. dem *leeren Produkt*).

Sinnvollerweise soll die leere Summe gleich 0 und das leere Produkt gleich 1 sein.

Definition 4.2. Seien $a, b \in \mathbb{Z}$. Die Zahl a ist ein *Teiler* von b, wenn es ein $c \in \mathbb{Z}$ gibt, sodass $ac = b$ gilt. Man schreibt dann $a \mid b$. Ist a kein Teiler von b, so schreibt man $a \nmid b$.

Beispiel 4.3. Offensichtlich gelten die folgenden Beziehungen:

- $3 \mid 6$;
- $3 \nmid 5$;
- $3 \mid (-6)$;
- $(-3) \mid 6$;
- $(-3) \mid (-6)$;
- $m \mid 0$ für jedes $m \in \mathbb{Z}$.

*4.1 Axiomatisierung der natürlichen Zahlen**

Wir wollen für diesen Abschnitt einmal vergessen, dass wir wissen, was die natürlichen Zahlen sind und wie man mit Ihnen rechnet. Ausgehend von drei Axiomen wollen wir die natürlichen Zahlen konstruieren und ihre Eigenschaften untersuchen.

Peano-Axiome.

1) Es gibt eine Menge $\mathbb{N}$ mit einer injektiven Abbildung
$$S\colon \mathbb{N} \longrightarrow \mathbb{N}.$$
2) Es gibt ein Element „0“$\in \mathbb{N}$, welches nicht im Bild von S liegt.
3) Ist $M \subseteq \mathbb{N}$, sodass gilt
   a) $0 \in M$,
   b) Für alle $n \in \mathbb{N}$ gilt: $n \in M \Rightarrow S(n) \in M$,

   Dann ist $M = \mathbb{N}$.

Die Abbildung S heißt Nachfolgerabbildung und bildet jede natürliche Zahl auf ihren Nachfolger ab. Die Zahl 0 ist nicht Nachfolger irgendeiner Zahl. Die Natürlichen Zahlen sind also

$$0,\ S(0),\ S\big(S(0)\big),\ S\Big(S\big(S(0)\big)\Big), \ldots$$

Diese Bezeichnungen sind unhandlich, daher setzt man:

$$S(0) =: 1, \quad S(1) =: 2, \ldots.$$

Wir könnten jetzt zeigen, dass nur mit den Peanoaxiomen die Rechenregeln für natürliche Zahlen folgen:

1) Kommutativgesetz der Addition: Für alle $m, n \in \mathbb{N}$ gilt:
$$n + m = m + n.$$
2) Assoziativgesetz der Addition: Für alle $k, m, n \in \mathbb{N}$ gilt:
$$k + (m + n) = (k + m) + n.$$

3) Kommutativgesetz der Multiplikation: Für alle $m, m \in \mathbb{N}$ gilt:

$$m \cdot n = n \cdot m.$$

4) Distributivgesetz: Für alle $k, m, n \in \mathbb{N}$ gilt:

$$(k + m) \cdot n = k \cdot n + m \cdot n.$$

5) Assoziativgesetz der Multiplikation: Für alle $k, m, n \in \mathbb{N}$ gilt:

$$(k \cdot m) \cdot n = k \cdot (m \cdot n).$$

Bevor wir diese Gesetze beweisen, müssen wir zunächst eine Defintion der Addition und Multiplikation in der Sprache der Peanoaxiome angeben. Wir beginnen mit der Addition und beweisen die Rechengesetz für die Addition der natürlichen Zahlen.

Definition 4.4. Die Addition natürlicher Zahlen wird rekursiv definiert:

1) $n + 0 \coloneqq n$,

2) $n + S(k) = S(n + k)$.

Definition 4.5. Die Multiplikation natürlicher Zahlen wird rekursiv definiert:

1) $n \cdot 0 = 0 \cdot n = 0$,

2) $n \cdot S(m) \coloneqq n + n \cdot m$.

Während die rekursive Definition der Addition etwas seltsam anmutet, erinnert die Definition der Multiplikation an die Grundschuldefintion der Multiplikation als wiederholte Addition.

Zunächst halten wir folgende Eigenschaft der Nachfolgerfunktion fest:

Lemma 4.6 (Schaukellemma)**.** *Für alle natürlichen Zahlen m, n gilt: $m + S(n) = S(m) + n$.*

Beweis. Wir zeigen die Behauptung durch vollständige Induktion über n. Sei dazu m eine beliebige, aber fest gewählte natürliche Zahl.

Induktionsanfang $n = 0$: Es gilt

$$m + S(0) = S(m + 0) = S(m) = S(m) + 0.$$

Induktionsschritt: Es gelte also $m + S(n) = S(m) + n$ für ein beliebiges $n \in \mathbb{N}$. Wir müssen zeigen, dass nun auch $m + S\big(S(n)\big) = S(m) + S(n)$ gilt:

$$\begin{aligned} m + S\big(S(n)\big) &= S\big(m + S(n)\big) \\ &= S\big(S(m) + n\big) \quad \text{(nach Induktionsvoraussetzung)} \\ &= S(m) + S(n). \end{aligned}$$

Damit ist die Behauptung bewiesen. □

Lemma 4.7. *Für alle natürlichen Zahlen n gilt: $n + 0 = 0 + n$.*

Beweis. Induktion über n.

Induktionsanfang $n = 0$: $0 + 0 = 0 + 0$.

Induktionsschritt: Sei also für ein $n \in \mathbb{N}$ die Gleichung $n + 0 = 0 + n$ erfüllt. Wir müssen zeigen, dass $S(n) + 0 = 0 + S(n)$ gilt. Es gilt:

$$\begin{aligned} S(n) + 0 &= n + S(0) && \text{(nach Lemma 4.6)} \\ &= S(n + 0) \\ &= S(0 + n) && \text{(nach Induktionsvoraussetzung)} \\ &= 0 + S(n). \end{aligned}$$

Damit ist die Induktion abgeschlossen. □

Korollar 4.8 (0 ist neutrales Element der Addition)**.** *Für alle natürlichen Zahlen n gilt: $0 + n = n + 0 = n$.*

Satz 4.9 (Kommutativgesetz der Addition)**.** *Für alle natürlichen Zahlen m, n gilt: $m + n = n + m$.*

Beweis. Wir zeigen die Aussage durch Induktion über m. Sei dazu $n \in \mathbb{N}$ beliebig aber fest gewählt.

Induktionsanfang $m = 0$: Es gilt $n + 0 = 0 + n$ nach Lemma 4.7.

Induktionsschritt: Es gelte $m + n = n + m$ für ein $m \in \mathbb{N}$. Wir müssen zeigen, dass $n + S(m) = S(m) + n$ gilt:

$$\begin{aligned} n + S(m) &= S(n + m) \\ &= S(m + n) && \text{(nach Induktionsvoraussetzung)} \\ &= m + S(n) \\ &= S(m) + n && \text{(nach Lemma 4.6).} \end{aligned}$$

Damit ist der Satz bewiesen. □

Satz 4.10 (Assoziativgesetz der Addition)**.** *Für alle $k, m, n \in \mathbb{N}$ gilt: $k + (m + n) = (k + m) + n$.*

Beweis. Induktion über n. Seien k und m beliebig aber fest gewählt.

Induktionsanfang $n = 0$: Es gilt $k + (m + 0) = k + m = (k + m) + 0$ nach Definition 4.4.

Induktionsschritt: Es gelte $k + (m + n) = (k + m) + n$ für ein $n \in \mathbb{N}$. Wir müssen zeigen, dass gilt: $k + \big(m + S(n)\big) = (k + m) + S(n)$.

$$\begin{aligned} k + \big(m + S(n)\big) &= k + S(m + n) \\ &= S\big(k + (m + n)\big) \\ &= S\big((k + m) + n\big) && \text{(nach Induktionsvor.)} \\ &= (k + m) + S(n). \end{aligned}$$

□

Lemma 4.11 (1 ist neutrales Element der Multiplikation)**.** *Für alle $n \in \mathbb{N}$ gilt: $n \cdot 1 = n = 1 \cdot n$.*

Beweis. Den ersten Teil zeigen wir direkt:

Es gilt: $n \cdot 1 = n \cdot S(0) = n + n \cdot 0 = n + 0 = n$.

Den zweiten Teil zeigen wir per vollständiger Induktion über n:

Induktionsanfang $n = 0$: Es gilt nach Definition 4.5: $S(0) \cdot 0 = 1 \cdot 0 = 0$.

Induktionsschritt: Es gelte $S(0) \cdot n = n$ für ein $n \in \mathbb{N}$. Wir müssen zeigen, dass $S(0) \cdot S(n) = S(n)$ gilt.

$$\begin{aligned} 1 \cdot S(n) S(0) \cdot S(n) &= S(0) + S(0) \cdot n \\ &= S(0) + n \quad \text{(nach Induktionsvoraussetzung)} \\ &= S(0 + n) \\ &= S(n). \end{aligned}$$

□

Satz 4.12 (Distributivgesetz)**.** *Für alle natürlichen Zahlen k, m, n gilt:*

$$(k + m) \cdot n = k \cdot n + m \cdot n.$$

Beweis. Wir zeigen die Behauptung durch vollständige Induktion über n. Seien k, m fest aber beliebig gewählt.

Induktionsanfang $n = 0$: Es gilt: $(k + m) \cdot 0 = 0 = 0 + 0 = k \cdot 0 + m \cdot 0$.

Induktionsschritt: Es gelte $(k + m) \cdot n = k \cdot n + m \cdot n$ für ein $n \in \mathbb{N}$. Wir müssen zeigen: $(k + m) \cdot S(n) = k \cdot S(n) + m \cdot S(n)$.

$$\begin{aligned} (k + m) \cdot S(n) &= (k + m) + (k + m) \cdot n \\ &= (k + m) + k \cdot n + m \cdot n \quad \text{(nach Induktionsvor.)} \\ &= (k + k \cdot n) + (m + m \cdot n) \\ &= k \cdot S(n) + m \cdot S(n). \end{aligned}$$

□

Satz 4.13 (Kommutativgesetz der Multiplikation)**.** *Fur alle $m, n \in \mathbb{N}$ gilt: $m \cdot n = n \cdot m$.*

Beweis. Induktion über n. Sei m beliebig aber fest gewählt.

Induktionsanfang $n = 0$: Es gilt: $m \cdot 0 = 0 = 0 \cdot m$ nach Definition 4.5.

Induktionsschritt: Es gelte $m \cdot n = n \cdot m$ für ein $n \in \mathbb{N}$. Wir müssen

zeigen: $m \cdot S(n) = S(n) \cdot m$.

$$\begin{aligned} m \cdot S(n) &= m + m \cdot n \\ &= m + n \cdot m \quad \text{(nach Induktionsvoraussetzung)} \\ &= n \cdot m + m \\ &= n \cdot m + (m + 0) \\ &= n \cdot m + (m + 0 \cdot m) \\ &= n \cdot m + S(0) \cdot m \\ &= \big(n + S(0)\big) \cdot m \quad \text{(nach Satz 4.12)} \\ &= \big(S(n + 0)\big) \cdot m \\ &= S(n) \cdot m. \end{aligned}$$

□

Korollar 4.14 (Weiteres Distributivgesetz)**.** *Für alle $k, m, n \in \mathbb{N}$ gilt:*

$$k \cdot (m + n) = k \cdot m + k \cdot n.$$

Satz 4.15 (Assoziativgesetz der Multiplikation)**.** *Für alle $k, m, n \in \mathbb{N}$ gilt:*

$$(k \cdot m) \cdot n = k \cdot (m \cdot n).$$

Beweis. Vollständige Induktion über n. Seien k und m fest aber beliebig gewählt.

Induktionsanfang $n = 0$: Es gilt $(k \cdot m) \cdot 0 = 0 = k \cdot 0 = k \cdot (m \cdot 0)$.

Induktionsschritt: Es gelte also $(k \cdot m) \cdot n = k \cdot (m \cdot n)$ für ein $n \in \mathbb{N}$. Wir müssen zeigen: $(k \cdot m) \cdot S(n) = k \cdot \big(m \cdot S(n)\big)$.

$$\begin{aligned} (k \cdot m) \cdot S(n) &= k \cdot m + (k \cdot m) \cdot n \\ &= k \cdot m + k \cdot (m \cdot n) \quad \text{(nach Induktionsvor.)} \\ &= k \cdot (m + m \cdot n) \\ &= k \cdot \big(m \cdot S(n)\big). \end{aligned}$$

□

Abschließend wollen wir auch noch formal die kleiner-gleich-Relation auf den natürlichen Zahlen definieren.

Definition 4.16. Die *kleiner-Relation* wird auf den natürlichen Zahlen rekursiv definiert:

- Für alle $n \in \mathbb{N}$ gilt: $\neg(n < 0)$. Wir schreiben auch $n \not< 0$.
- Für alle $m, n \in \mathbb{N}$ gilt: $m < S(n)$ genau dann, wenn $m = n$ oder $m < n$.

Eine natürliche Zahl m heißt *kleiner gleich* einer natürlichen Zahl n, falls $m < n$ oder $m = n$ gilt. Wir schreiben dann: $m \leq n$.

Wir veranschaulichen die Definition an einem Beispiel:

Beispiel 4.17. Wir betrachen die natürlichen Zahlen 3 und 5. Gilt $3 \leq 5$? Wir überprüfen es durch Anwendung der Definition:

$$3 \leq 5 \Leftrightarrow 3 < 5 \vee 3 = 5.$$

Offensichtlich ist $3 \neq 5$, wir müssen also überprüfen, ob $3 < 5$ gilt. Es gilt: $5 = S(4) = S(S(3))$. Wir wenden die Definition an:

$$3 < 5 \Leftrightarrow 3 < S(4) \Leftrightarrow 3 = 4 \vee 3 < 4.$$

Offensichtlich ist $3 \neq 4$, wir müssen also überprüfen, ob $3 < 4$ ist. Wir wenden die Definition an:

$$3 < 4 \Leftrightarrow 3 < S(3) \Leftrightarrow 3 = 3 \vee 3 < 3.$$

Offensichtlich ist $3 = 3$, somit ist $3 < 5$ und somit auch $3 \leq 5$.

Wie sieht es mit $3 \leq 2$ aus? Gilt dies? Wir wenden die Definition an:

$$3 \leq 2 \Leftrightarrow 3 < 2 \vee 3 = 2.$$

Offensichtlich ist $3 \neq 2$, wir müssen also überprüfen, ob $3 < 2$ gilt. Es gilt: $3 = S(2) = S(S(1)) = S(S(S(0)))$. Wir wenden die Definition an:

$$3 < 2 \Leftrightarrow 3 < S(1) \Leftrightarrow 3 = 1 \vee 3 < 1.$$

Offensichtlich ist $3 \neq 1$, wir müssen also überprüfen, ob $3 < 1$ ist:

$$3 < 1 \Leftrightarrow 3 < S(0) \Leftrightarrow 3 = 0 \vee 3 < 0.$$

Offensichtlich ist $3 \neq 0$, wir müssen also überprüfen, ob $3 < 0$ ist. Nach Defintion von $<$ gilt: $\neg(3 < 0)$. Also ist $3 \not< 0$ und somit gilt dann auch $3 \not< 2$.

4.2 *Schwache und starke Induktion*

Oft hat man Aussagen der Form

$$\text{Für alle } x \in M \text{ gilt die Ausssage } A(x).$$

Ist M eine endliche Menge, kann man jeden Fall einzeln prüfen.

Beispiel 4.18. Für alle $n \in \{2, 3, 4\}$ gilt: $2^n \leq n^2$.

In diesem Beispiel ist $|M| = 3$ und man kann jeden Fall einzeln nachrechnen. Ist M eine unendliche Menge, geht dies nicht mehr.

Beispiel 4.19. Für alle $n \in \{m \in \mathbb{N} \mid m \geq 5\}$ gilt $2^n > n^2$.

Hat man eine Aussage der Form

$$\text{Für alle } n \in \mathbb{N} \text{ mit } n \geq n_0 \text{ gilt } A(n),$$

dann kann man diese mit der Methode der *vollständigen Induktion* beweisen (falls sie wahr ist):

1) Zeige dass $A(n_0)$ gilt, dass also die Aussage für die erste Zahl n_0 wahr ist. Dies nennt man den *Induktionsanfang*.

2) Zeige, dass $A(n) \Rightarrow A(n+1)$ für $n \geq n_0$ gilt, das also, falls die Aussage für ein $n \geq n_0$ gilt, sie auch für $n+1$ gilt. Dies nennt man den *Induktionsschritt*.

Die Aussage $A(n)$ heißt in diesem Zusammenhang auch *Induktionsvoraussetzung*.

Man kann sich die Methode der vollständigen Induktion durch eine unendlich lange Kette von Dominosteinen veranschaulichen. Fällt der erste Stein um, und mit jedem Stein auch sein jeweiliger Nachfolger, dann fallen alle Steine um.

Satz 4.20 (Kleiner Gauß)**.** *Für alle $n \in \mathbb{N}$ gilt:*

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

Beweis. Wir zeigen die Aussage durch vollständige Induktion über n. Induktionsanfang $n = 0$:

Offensichtlich gilt:

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

Induktionsschritt: Wir nehmen jetzt an, dass für $n \geq 0$ gilt:

$$\sum_{i=0}^{n} i = \frac{1}{2} n \cdot (n+1). \tag{4.1}$$

Daraus wollen wir nun folgern, dass

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

gilt. Dazu formen wir die Summe enstprechend so um, dass wir die Induktionsvoraussetzung (4.1) nutzen können.

$$\begin{aligned}
\sum_{i=0}^{n+1} i &= \left(\sum_{i=0}^{n} i\right) + (n+1) \\
&= \frac{1}{2} n \cdot (n+1) + (n+1) \quad \text{(nach Induktionsvoraussetzung)} \\
&= \frac{1}{2} n \cdot (n+1) + \frac{2 \cdot (n+1)}{2} \\
&= \frac{1}{2}\big(n \cdot (n+1) + 2 \cdot (n+1)\big) \\
&= \frac{1}{2}(n+1) \cdot (n+2).
\end{aligned}$$

□

Der Kern jedes Beweises mit Vollständiger Induktion ist, die Induktionvoraussetzung geschickt zu benutzen um den Induktionsschritt durchzuführen.

Satz 4.21. *Sei M eine endliche Menge mit n Elementen. Dann gilt:*

$$\left|\mathcal{P}(M)\right| = 2^n.$$

Beweis. Wir zeigen die Behauptung durch vollständige Induktion über $|M|$.

Induktionsanfang $n = 0$: Es gibt nur eine Menge mit Null Elementen, die leere Menge.

$$\left|\mathcal{P}(\emptyset)\right| = \left|\{\emptyset\}\right| = 1 = 2^0.$$

Induktionsschritt: Für $n \geq 0$ und $|M| = n$ gelte

$$\left|\mathcal{P}(M)\right| = 2^n.$$

Zu zeigen ist nun, dass für $|M| = n + 1$ gilt:

$$\left|\mathcal{P}(M)\right| = 2^{n+1}.$$

Sei also $M = \{m_1, m_2, \ldots, m_n, m_{n+1}\}$ und $A \subseteq M$. Wir unterscheiden zwei Fälle:

1) $m_{n+1} \notin A$, d. h. $A \subseteq \{m_1, \ldots, m_n\}$. Nach Induktionsvoraussetzung gibt es 2^n solcher Teilmengen.
2) $m_{n+1} \in A$, d. h. $A \setminus \{m_{n+1}\} \subseteq \{m_1 \ldots, m_n\}$. Nach Induktionsvoraussetzung gibt es ebenso 2^n solcher Teilmengen.

Insgesamt gibt es also $2^n + 2^n = 2^{n+1}$ Teilmengen von M. □

Satz 4.22. *Für alle natürlichen Zahlen n gilt:* 3 *ist ein Teiler von* $4n^3 - n$.

Beweis. Wir zeigen den Satz durch Induktion über n.

Induktionsanfang $n = 0$: Offensichtlich gilt: $4 \cdot 0^3 - 0 = 0$ und $3 \mid 0$.

Induktionsschritt: Für $n \geq 0$ gelte nun $3 \mid (4n^3 - n)$. Wir müssen zeigen, dass $3 \mid \big(4(n+1)^3 - (n+1)\big)$ gilt.

$$\begin{aligned} 4(n+1)^3 - (n+1) &= 4(n^3 + 3n^2 + 3n + 1) - (n+1) \\ &= 4n^3 + 12n^2 + 12n + 4 - n - 1 \\ &= (4n^3 - n) + (12n^2 + 12^n + 3) \\ &= (4n^3 - n) + 3 \cdot (4n^2 + 4n + 1). \end{aligned}$$

Der erste Summand in der letzten Zeile ist nach Induktionsvoraussetzung durch 3 teilbar, der zweite Summand ist offensichtlich ein Vielfaches von 3 und damit ebenfalls durch 3 teilbar. □

Manchmal ist es nicht möglich aus $A(n)$ direkt $A(n+1)$ zu folgern. Man benötigt dann:

$$A(k) \text{ ist wahr für alle } n_0 \leq k \leq n.$$

Man nennt dies starke Induktion (in Abgrenzung zur vorherigen, schwachen Induktion).

Vollständige Induktion (zweite Variante, starke Induktion*)*

1) Zeige, dass $A(n_0)$ wahr ist.
2) Folgere aus: $A(k)$ ist wahr für alle $n_0 \leq k \leq n$, dass auch $A(n+1)$ wahr ist.

Wir werden diese Methode im nächsten Lemma verdeutlichen.

Definition 4.23. Eine natürliche Zahl $p \in \mathbb{N}$ mit $p > 1$ heißt *Primzahl*, falls p nur die trivialen Teiler $1, -1, p, -p$ besitzt.

Lemma 4.24. *Jede natürliche Zahl $n \geq 2$ ist entweder eine Primzahl oder Produkt endlich vieler Primzahlen.*

Beweis. Induktionsanfang $n = 2$: 2 ist eine Primzahl.

Induktionsschritt: Seien also alle natürlichen Zahlen $2 \leq k \leq n$ Primzahlen oder Produkt endlich vieler Primzahlen. Betrachte $n+1$. Entweder ist $n+1$ eine Primzahl, dann sind wir fertig. Ansonsten besitzt $n+1$ einen Teiler $a \in \{2, \ldots, n\}$, d. h. es gibt ein $b \in \{2, \ldots, n\}$, so dass $ab = n+1$ ist. Nach Induktionsvoraussetzung sind a und b jeweils Produkt endlich vieler Primzahlen. Daher gilt dies auch für $a \cdot b = n+1$. □

Zum Abschluss wollen wir uns noch einen interessanten Anwendungsfall der vollständigen Induktion anschauen.

Beispiel 4.25 (Türme von Hanoi)**.** Das folgende Spiel wurde 1883 von Édouard Lucas erfunden. Gegeben sind drei Stäbe (A, B, C) zum stapeln und eine Anzahl von n gelochte Scheiben, die zu Beginn alle auf dem Stab A gelegt sind.

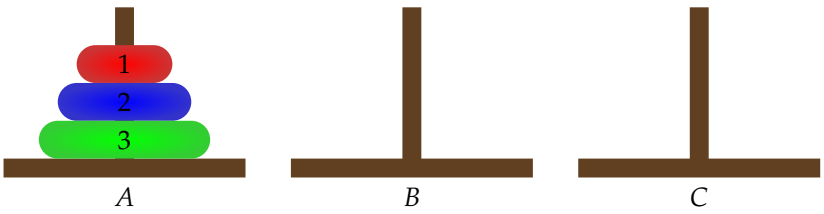


Abbildung 4.1: Ausgangssituation.

Dabei sind die Scheiben alle unterschiedlich groß und zu Beginn der Größe nach geordnet (die größte unten, die kleinste oben). Aufgabe ist es nun, die Scheiben von Stab A auf den Stab C zu legen.

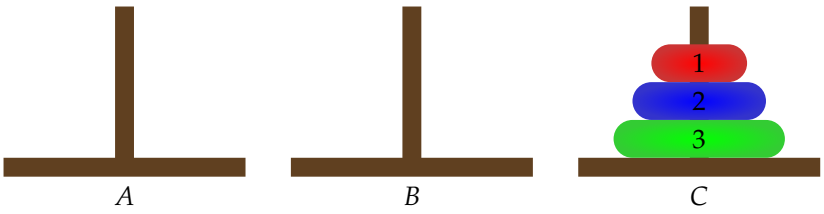


Abbildung 4.2: Ziel.

Dabei gelten die folgenden Regeln:

- Es darf nur eine Scheibe auf einmal bewegt werden.
- Eine Scheibe darf nur auf einen leeren Platz, oder auf eine größere Scheibe gelegt werden.

Behauptung: Es gibt einen Algorithmus, der die Aufgabe für jedes $n \in \mathbb{N}$ löst.

Beweis. Induktion über n.

Induktionsanfang $n = 0$. Es gibt einen Algorithmus für 0 Scheiben, da nichts zu tun ist.

Für die Anwendung etwas interessanterer Induktionsanfang $n = 1$: Bewege die Scheibe von A nach C. Fertig.

Induktionsschritt: Gegeben sei also ein Algorithmus, der die Aufgabe für n Scheiben löst. Wir müssen nun zeigen, dass es einen Algorithmus für $n + 1$ Scheiben gibt, der die Aufgabe löst.

Nach Induktionsvoraussetzung wissen wir, dass es einen Algorithmus für n Scheiben gibt. Nutze diesen, um die oberen n Scheiben von A nach B zu bewegen.

Bewege dann die $(n + 1)$-te Scheibe von A nach C.

Nutze nun wieder den Algorithmus für n Scheiben um die n Scheiben von B nach C zu bewegen. Fertig. □

Offensichtlich kann man mittels vollständiger Induktion nicht nur beweisen, dass bestimmte Algorithmen korrekt sind, sondern auch rekursive Algorithmen entwerfen.

5
Die ganzen Zahlen

In diesem Kapitel werden wir uns mit Zahlen, insbesondere den ganzen Zahlen beschäftigen. Hier werden wir uns zunächst die ganzen Zahlen konstruieren, bevor wir ihre Eigenschaften studieren. Dann werden wir speziellen Teilmengen der ganzen Zahlen und verwandten Zahlmengen untersuchen.

5.1 Die ganzen Zahlen

Wir wollen die ganzen Zahlen eingehender studieren. Dazu konstruieren wir die ganzen Zahlen aus den schon bekannten natürlichen Zahlen über eine Äquivanenzrelation. Die Motivation dahinter ist, dass wir gerne Gleichungen der Form

$$a + x = b$$

lösen können möchten. Dies ist aber in den natürlichen Zahlen nicht immer möglich, z. B. für $a = 5$ und $b = 3$ gibt es keine natürliche Zahl x, welche diese Gleichung erfüllt.

Definition 5.1. Die Relation $\sim$ auf $\mathbb{N} \times \mathbb{N}$ sei gegeben durch:

$$(m, n) \sim (m', n') :\Leftrightarrow m + n' = m' + n.$$

Proposition 5.2. *Die Relation* $\sim \subseteq (\mathbb{N} \times \mathbb{N}) \times (\mathbb{N} \times \mathbb{N})$ *ist eine Äquivalenzrelation.*

Beweis. Die Relation $\sim$ ist reflexiv, denn für alle $(a, b) \in \mathbb{N} \times \mathbb{N}$ gilt: $(a, b) \sim (a, b)$, da $a + b = b + a$. Die Symmetrie ergibt sich aus: $(a, b) \sim (c, d) \Leftrightarrow a + d = c + b \Leftrightarrow c + b = a + d \Leftrightarrow (c, d) \sim (a, b)$. Die Transitivität folgt aus:

$$\begin{aligned}(a, b) \sim (c, d) \wedge (c, d) \sim (e, f) &\Leftrightarrow a + d = c + b \wedge c + f = e + d \\ &\Rightarrow a + d + c + f = c + b + e + d \\ &\Rightarrow a + f = e + b \\ &\Leftrightarrow (a, b) \sim (e, f).\end{aligned}$$

□

Notation 5.3. Wir bezeichnen die Menge der Äquivalenzklassen von $\sim$

$$\{\overline{(m,n)} \mid m,n \in \mathbb{N}\}$$

mit $\mathbb{Z}$.

Wir wollen nun auf $\mathbb{Z}$ rechnen. Dazu führen wir die Addition und Multiplikation in $\mathbb{Z}$ auf die entsprechenden Operationen in den natürlichen Zahlen zurück.

Definition 5.4. Auf $\mathbb{Z}$ definieren wir die folgende Addition:

$$\overline{(a,b)} + \overline{(c,d)} = \overline{(a+c, b+d)}.$$

Außerdem definieren wir die folgende Multiplikation:

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

Das diese Definitionen sinnvoll sind, zeigt der nächste Satz.

Satz 5.5. *Die Addition und Multiplikation auf $\mathbb{Z}$ sind wohldefiniert, d. h. sie sind unabhängig von den gewählten Repräsentanten. Das heißt es gelten:*

$$(a,b) \sim (a',b') \wedge (c,d) \sim (c',d') \Leftrightarrow \overline{(a+c,b+d)} = \overline{(a'+c',b'+d')}.$$

und

$$(a,b) \sim (a',b') \wedge (c,d) \sim (c',d') \Leftrightarrow \overline{(ac+bd,ad+bc)} = \overline{(a'c'+b'd',a'd'+b'c')}.$$

Beweis. Wir zeigen die Behauptung für die Addition. Der Beweis für die Multiplikation verläuft analog und wird der geneigten Leserin zur Übung überlassen.

Es gilt: $(a,b) \sim (a',b')$, d. h. $a+b' = a'+b$. Außerdem gilt $c+d' = c'+d$ wegen $(c,d) \sim (c',d')$. Also ist

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

und deswegen $(a+c,b+d) \sim (a'+c',b'+d')$. Daraus folgt:

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

□

Diese Addition und Multiplikation haben die folgenden Eigenschaften.

Satz 5.6. *Für alle Elemente $\overline{(a,b)}$, $\overline{(c,d)}$ und $\overline{(e,f)}$ aus $\mathbb{Z}$ gilt:*

1) Assoziativität der Addition:

$$\overline{(a,b)} + \left(\overline{(c,d)} + \overline{(e,f)}\right) = \left(\overline{(a,b)} + \overline{(c,d)}\right) + \overline{(e,f)}.$$

2) *Kommutativität der Addition:*

$$\overline{(a,b)} + \overline{(c,d)} = \overline{(c,d)} + \overline{(a,b)}.$$

3) *Neutrales Element der Addition:* $\overline{(0,0)} + \overline{(a,b)} = \overline{(a,b)}$.

4) *Inverse Elemente der Addition:* $\overline{(a,b)} + \overline{(b,a)} = \overline{(0,0)}$.

5) *Assoziativität der Multiplikation:*

$$\overline{(a,b)} \cdot \big(\overline{(c,d)} \cdot \overline{(e,f)}\big) = \big(\overline{(a,b)} \cdot \overline{(c,d)}\big) \cdot \overline{(e,f)}.$$

6) *Kommutativität der Multiplikation:*

$$\overline{(a,b)} \cdot \overline{(c,d)} = \overline{(c,d)} \cdot \overline{(a,b)}.$$

7) *Neutrales Element der Multiplikation:* $\overline{(1,0)} \cdot \overline{(a,b)} = \overline{(a,b)}$.

8) *Distributivitätsgesetz:*

$$\overline{(a,b)} \cdot \big(\overline{(c,d)} + \overline{(e,f)}\big) = \overline{(a,b)} \cdot \overline{(c,d)} + \overline{(a,b)} \cdot \overline{(e,f)}.$$

Beweis. Seien also $a, b, c, d, e, f \in \mathbb{N}$. Die Eigenschaften für Elementen aus $\mathbb{Z}$ folgen direkt aus den Eigenschaften von Elementen aus $\mathbb{N}$.

1)

$$\begin{aligned}
\overline{(a,b)} + \big(\overline{(c,d)} + \overline{(e,f)}\big) &= \overline{(a,b)} + \overline{(c+e, d+f)} \\
&= \overline{\big(a + (c+e), b + (d+f)\big)} \\
&= \overline{\big((a+c)+e, (b+d)+f\big)}
\end{aligned}$$

(Wegen Assoz. der Addition auf $\mathbb{N}$)

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

2)

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

3)

$$\begin{aligned}
\overline{(0,0)} + \overline{(a,b)} &= \overline{(0+a, 0+b)} \\
&= \overline{(a,b)}.
\end{aligned}$$

4)

$$\begin{aligned}
\overline{(0,0)} &= \overline{(a+b, a+b)} \\
&= \overline{(a+b, b+a)} \\
&= \overline{(a,b)} + \overline{(b,a)}.
\end{aligned}$$

5) – 8) Übung.

□

Wir erweitern nun die Relation $\leq$ auf $\mathbb{Z}$.

Definition 5.7. Sind $\overline{(a,b)}$ und $\overline{(c,d)}$ ganze Zahlen, dann heißt $\overline{(a,b)}$ *kleiner gleich* $\overline{(c,d)}$ (symbolisch $\overline{(a,b)} \leq \overline{(c,d)}$), falls $a+d \leq b+c$ gilt. Die Zahl $\overline{(a,b)}$ heißt *kleiner* als $\overline{(c,d)}$ (symbolisch $\overline{(a,b)} < \overline{(c,d)}$), falls $\overline{(a,b)} \leq \overline{(c,d)}$ und $\overline{(a,b)} \neq \overline{(c,d)}$ gelten.

Man kann die natürlichen Zahlen in $\mathbb{Z}$ einbetten:

$$\begin{aligned} \mathbb{N} &\longrightarrow \mathbb{Z} \\ n &\longmapsto \overline{(n,0)}. \end{aligned}$$

Diese Abbildung ist offensichtlich injektiv und erlaubt es zusammen mit Satz 5.6 die Elemente von $\mathbb{Z}$ in einfacherer Weise zu schreiben. Statt $\overline{(n,0)}$ schreiben wir n und statt $\overline{(0,n)}$ schreiben wir $-n$. Offensichtlich finden wir für jede Äquivalenzklasse $\overline{(m,n)}$ einen Repräsentanten $(m-n,0)$ (falls $m \geq n$) oder $(0,n-m)$ (falls $m \leq n$), denn $0+m = m-n+n$, bzw. $m+n-m = 0+n$.

Definition 5.8. Der *Betrag* $|z|$ einer ganzen Zahl $z \in \mathbb{Z}$ ist gegeben durch:

$$|z| := \begin{cases} z, & \text{falls } z \geq 0 \\ -z, & \text{sonst} \end{cases}.$$

5.2 *Teilbarkeit*

Wir erinnern noch einmal an die Definition von Teilbarkeit für ganze Zahlen (siehe Kapitel 4):

Definition 5.9. Seien $a, b \in \mathbb{Z}$. Die Zahl a ist ein *Teiler* von b, wenn es ein $c \in \mathbb{Z}$ gibt, sodass $ac = b$ gilt. Man schreibt dann $a \mid b$. Ist a kein Teiler von b, so schreibt man $a \nmid b$.

Proposition 5.10. *Die Teilbarkeitsrelation auf* $\mathbb{N}$*, also die Relation* $R \subseteq \mathbb{N} \times \mathbb{N}$ *mit* $aRb :\Leftrightarrow a \mid b$ *ist eine Halbordnung auf* $\mathbb{N}$.

Beweis. Es gilt für alle $a \in \mathbb{N}$: $a \cdot 1 = a$ und damit $a \mid a$. Die Relation ist also reflexiv. Für die Antisymmetrie betrachten wir die Aussage $a \mid b \wedge b \mid a$. Dies bedeutet, dass es $c, d \in \mathbb{N}$ gibt, mit $ac = b$ und $bd = a$. Dann folgt $acd = a$ und daraus $cd = 1$, da a beliebig gewählt werden kann. Da c und d natürliche Zahlen sind, muss gelten $c = d = 1$, daraus folgt: $a = b$. Um die Transitivität zu zeigen, betrachten wir die Aussage $a \mid b \wedge b \mid c$. Dies bedeutet, dass es $d, e \in \mathbb{N}$ gibt, für die gilt: $ad = b$ und $be = c$. Daraus folgt, dass $ade = c$ gilt. Wir setzen $f := de$ und erhalten damit $a \mid c$. □

Die Teilbarkeitsrelation auf $\mathbb{Z}$ ist keine Halbordnung, denn z. B. gilt:

$$2 \mid (-2) \text{ und } (-2) \mid 2,$$

aber $2 \neq -2$.

Proposition 5.11. *Seien $k, l, m, n \in \mathbb{Z}$ und $k \neq 0$. Dann gelten:*

1) *Wenn $k \mid l$ und $k \mid m$ gelten, dann auch $k \mid (l + m)$.*
2) *Wenn $k \mid l$ und $k \mid m$ gelten, dann auch $k \mid (l - m)$.*
3) *Wenn $k \mid l$ gilt., dann gilt auch: $k \mid ln$.*
4) *Sei $m \neq 0$. Dann gilt: Aus $k \mid l$ und $m \mid n$ folgt $km \mid ln$.*

Beweis. 1) Die Voraussetzung ist äquivalent dazu, dass es $c, d \in \mathbb{Z}$ gibt, für die gilt $kc = l$ und $kd = m$. Daher gilt $l + m = kc + kd = k \cdot (c + d)$ und damit teilt k auch die Summe $l + m$.

2) Analog zum vorherigen Fall erhalten wir $l - m = kc - kd = k \cdot (c - d)$ und daher gilt $k \mid (l - m)$.

3) Aus $k \mid l$ folgt, dass es ein $c \in \mathbb{Z}$ gibt, mit $kc = l$. Dann gilt $(kc)n = k(cn) = ln$ und daher folgt $k \mid ln$.

4) Aus der Voraussetzung folgt, dass es $c, d \in \mathbb{Z}$ gibt, für die gilt: $kc = l$ und $md = n$. Damit erhalten wir $kc \cdot md = km \cdot (cd) = ln$ und daher $km \mid ln$.

□

5.3 *Division mit Rest*

In diesem Abschnitt untersuchen wir die Division mit Rest. Diese errinnert ein wenig an Grundschulmathematik, da wir hier Rechnungen der Form

$$7 \div 2 = 3 \text{ Rest } 1$$

betrachten. Oft interessiert man sich nur für den Rest, (z. B. bei der Frage ob eine Zahl gerade ist).

Lemma 5.12 (Division mit Rest in $\mathbb{Z}$)**.** *Seien $a, b \in \mathbb{Z}$ und $b \neq 0$. Dann gibt es eindeutig bestimmte $q, r \in \mathbb{Z}$ mit $0 \leq r < |b|$ für die gilt:*

$$a = qb + r.$$

Beweis. Sei $a \geq 0$. Wir zeigen die Existenz durch starke Induktion über a. Für den Induktionsanfang sei $0 \leq a < b$. Dann ist $a = 0 \cdot b + a$ in der gewünschten Form.

Im Induktionsschritt sei jetzt $a > b$ und für alle Zahlen, die kleiner als a sind gelte die Behauptung. Es ist $a - b < a$, also gibt es nach Induktionsvoraussetzung $q', r' \in \mathbb{Z}$ mit $0 \leq r' < |b|$ für die gilt:

$$a - b = q' \cdot b + r'.$$

Dann folgt

$$a = b + q' \cdot b + r' = (q' + 1) \cdot b + r',$$

und mit $q := q' + 1$ und $r := r'$ gilt die Behauptung für alle $a \geq 0$.

Sei jetzt $a < 0$. Dann gibt es nach der obigen Induktion $q', r' \in \mathbb{Z}$ mit $0 \leq r' < |b|$. für die gilt:

$$-a = q' \cdot b + r'.$$

Daher folgt:

$$a = \begin{cases} (-q' - 1) \cdot b + (b - r'), & \text{falls } r' \neq 0 \\ -q' \cdot b, & \text{falls } r' = 0. \end{cases}$$

Mit $q = -q' - 1$ und $r = b - r'$ bzw. $q = -q'$ und $r = 0$ folgt dann die Behauptung.

Die Darstellung ist eindeutig, denn angenommen es gäbe zwei Darstellungen:

$$a = q_1 b + r_1$$

und

$$a = q_2 b + r_2.$$

Dann gilt:

$$0 = a - a = (q_1 - q_2)b + (r_1 - r_2).$$

Daher ist b ein Teiler von $(r_1 - r_2)$. Da aber $|r_1 - r_2| < |b|$ ist, folgt $r_1 = r_2$ und daraus $q_1 = q_2$. □

Notation 5.13. Sind a, b ganze Zahlen und $a = q \cdot b + r$ mit $q, r \in \mathbb{Z}$ und $0 \leq r < |b|$, dann bezeichnen wir mit

$$a \bmod b := r \qquad (\text{a } \textit{modulo} \text{ b})$$

den Rest von a bei Division durch b, und mit

$$a \operatorname{div} b := q$$

die Ganzzahldivision von a durch b.

Die folgende Definition setzt Zahlen, die bei der Division den selben Rest lassen in Beziehung zueinander.

Definition 5.14. Seien a, b, m ganze Zahlen. Wir sagen, *a ist kongruent b modulo m*, falls a und b bei Division durch m den selben Rest lassen, d. h. falls gilt:

$$a \bmod m = b \bmod m.$$

Wir schreiben dann $a \equiv b \pmod{m}$.

Beispiel 5.15. Seien $a = -117$ und $b = 12$, dann ist $a = (-10)b + 3$ und somit $a \equiv 3 \pmod{12}$ und $a \operatorname{div} b = -10$.

Wir betrachten im Folgenden Mengen der Vielfachen einer ganzen Zahl. Diese Mengen werden im späteren Verlauf immer wieder auftauchen.

Definition 5.16. Für $n \in \mathbb{N}$ ist

$$n\mathbb{Z} := \{nx \mid x \in \mathbb{Z}\}$$

die *Menge der Vielfachen* von n.

Offenbar gilt für $m \mid n$:

$$n\mathbb{Z} \subseteq m\mathbb{Z}.$$

Wir wollen nun untersuchen, wie für gegebene $a, b \in \mathbb{Z}$ die Menge

$$M_{a,b} = a\mathbb{Z} + b\mathbb{Z} = \{ax + by \mid x, y \in \mathbb{Z}\}$$

der Summen der Vielfachen von a und b aussieht.

Bemerkung 5.17. Ist $m \in M_{a,b}$, so ist auch $-m \in M_{a,b}$.

Satz 5.18. *Sei $a, b \in \mathbb{Z}$. Dann gibt es ein $d \in \mathbb{N}$, sodass*

$$M_{a,b} = d\mathbb{Z}$$

gilt.

Beweis. Sei d das kleinste Element größer Null in $M_{a,b}$. Dann gibt es für alle $x, y \in \mathbb{Z}$ Zahlen $q, r \in \mathbb{Z}$ mit $0 \leq r < d$, sodass gilt:

$$ax + by = qd + r.$$

Nach Definition ist $d \in M_{a,b}$, also gibt es $x', y' \in \mathbb{Z}$ mit $d = ax' + by'$. Insbesondere ist dann

$$r = a(x - x'q) + b(y - y'q) \quad \in M_{a,b}.$$

Aber $M_{a,b}$ enthält keine positiven Zahlen kleiner als d und nach Bemerkung 5.17 gar keine Zahl $r \neq 0$ mit $0 < |r| < d$. Deshalb ist $r = 0$, also

$$ax + by = dq \quad \in d\mathbb{Z}.$$

Also ist $M_{a,b} \subseteq d\mathbb{Z}$.

Außerdem gilt für alle $q \in \mathbb{Z}$:

$$dq = a(x'q) + b(y'q) \in M_{a,b},$$

also $d\mathbb{Z} \subseteq M_{a,b}$ und somit gilt die Behauptung. □

Bemerkung 5.19. Die Zahl d ist die größte positive Zahl, die a und b teilt.

Definition 5.20. Die Zahl d aus Satz 5.18 wird $\mathrm{ggT}(a, b)$ genannt.

Es stellt sich nun die Frage, wie man den ggT berechnen kann.

Lemma 5.21. *Seien $a, b \in \mathbb{Z}$ und $a = qb + r$ für $q, r \in \mathbb{Z}$ mit $0 \leq r < |b|$. Dann gilt:*

$$\mathrm{ggT}(a, b) = \mathrm{ggT}(b, r).$$

Beweis. Es gilt: $\mathrm{ggT}(a,b)$ ist ein Teiler von a und von b. Mit Proposition 5.11 folgt dann: $\mathrm{ggT}(a,b)$ teilt auch r. Ungekehrt teilt $\mathrm{ggT}(b,r)$ sowohl b als auch r und nach Proposition 5.11 damit auch a. □

Satz 5.22. *Seien $a,b \in \mathbb{Z}$. Der folgende Algorithmus berechnet $\mathrm{ggT}(a,b)$.*

Listing 5.1: Euklidischer Algorithmus

```
euklid(a,b)
        WENN b = 0
                return a
        SONST
                return euklid(b, a mod b)
```

Wir schauen uns zunächst ein Beispiel an.

Beispiel 5.23. Seien $a = 128$ und $b = 84$. Dann gilt:

$$\begin{aligned}\mathrm{ggT}(a,b) &= \mathrm{ggT}(128,84)\\ &= \mathrm{ggT}(84,44)\\ &= \mathrm{ggT}(44,40)\\ &= \mathrm{ggT}(40,4)\\ &= \mathrm{ggT}(4,0)\\ &= 4\end{aligned}$$

Offenbar macht es keinen Unterschied, ob a oder b die größere Zahl ist:

$$\begin{aligned}\mathrm{ggT}(84,128) &= \mathrm{ggT}(128, 84 \mod 128)\\ &= \mathrm{ggT}(128,84)\end{aligned}$$

Man hat in diesem Fall also nur einen Schritt Mehraufwand.

Wir beweisen den Satz nicht, sondern zeigen eine stärkere Aussage.

Satz 5.24. *Seien $a,b \in \mathbb{Z}$. Dann bestimmt der folgende Algorithmus $d = \mathrm{ggT}(a,b)$ sowie $x,y \in \mathbb{Z}$ mit $\mathrm{ggT}(a,b) = ax + by$.*

Listing 5.2: Erweiterter euklidischer Algorithmus

```
erweiterter_euklid(a,b)
        WENN b = 0
                return (a,1,0)
        SONST
                (d',x',y') = erweiterter_euklid(b, a mod b)
                (d,x,y) = (d',y',x'- (a div b)y')
                return (d,x,y)
```

Beweis. Offenbar wird $d = \mathrm{ggT}(a,b)$ analog zum euklidischen Algorithmus berechnet. Bezeichne a_i, b_i die Eingabe im i-ten Schritt des Algorithmus. Da spätestens ab dem zweiten Schritt $a_i > b_i$ gilt und

damit $a_i > a_{i+1} = b_i$, gibt es einen k-ten Schritt, bei dem der Algorithmus mit $b_k = 0$ terminiert. Es bleibt somit zu zeigen, dass die berechneten x, y tatsächlich der Gleichung $ax + by = \text{ggT}(a, b)$ genügen.

Wir beweisen durch vollständige Induktion über i:

$$a_i = a \cdot x_i + b \cdot y_i$$

Induktionanfang $i = 0$: $a_0 = a \cdot x_0 + b \cdot y_0 = a \cdot 1 + b \cdot 0$.

$i = 1$: $a_1 = b = a \cdot x_1 + b \cdot y_1 = a \cdot 0 + b \cdot 1$.

Induktionsschritt: Sei für ein $1 \leq i < k$ die Gleichung $a_i = a \cdot x_i + b \cdot y_i$ erfüllt. Wir müssen zeigen, dass der Algorithmus im $(i+1)$-Schritt Werte x_{i+1}, y_{i+1} bestimmt, die $a_{i+1} = a \cdot x_{i+1} + b \cdot y_{i+1}$ erfüllen.

$$\begin{aligned}
a_{i+1} &= b_i \\
&= a_{i-1} \mod b_{i-1} \\
&= a_{i-1} - q_{i+1} b_{i-1} \quad \text{für ein } q_{i+1} \\
&= a_{i-1} - q_{i+1} a_i \\
&= (a \cdot x_{i-1} + b \cdot y_{i-1}) - q_{i+1}(a \cdot x_i + b \cdot y_i) \quad \text{(nach I.V.)} \\
&= a \cdot (x_{i-1} - q_{i+1} x_i) + b \cdot (y_{i-1} - q_{i+1} y_i)
\end{aligned}$$

Wenn der Algorithmus terminiert gilt also

$$a_{k-1} = \text{ggT}(a, b) = a \cdot x_k + b \cdot y_k.$$

□

Beispiel 5.25. Seien $a = 27$, $b = 15$. Es gilt:

$$\begin{aligned}
27 &= 1 \cdot 15 + 12 \\
15 &= 1 \cdot 12 + 3 \\
12 &= 4 \cdot 3 + 0.
\end{aligned}$$

Also $\text{ggT}(27, 15) = 3$. Wir setzen nun rückwärts ein:

$$\begin{aligned}
3 &= 1 \cdot 3 + 0 \cdot 12 \\
&= 1 \cdot 15 - 1 \cdot (27 - 15) \\
&= 1 \cdot 15 - 27 + 15 \\
&= 2 \cdot 15 - 27 \\
&= (-1) \cdot 27 + 2 \cdot 15.
\end{aligned}$$

Wir können den erweiterten euklidischen Algorithmus auch tabellarisch notieren (i ist der Index der Zeilen):

i	a	b	q	x	y
0	27	15	1	-1	2
1	15	12	1	1	-1
2	12	3	4	0	1
3	3	0		1	0

Dabei füllt man die ersten drei Spalten (a, b, q) von oben nach unten aus, die letzten beiden (x, y) hingegen von unten nach oben. Man nutzt dabei die folgenden Beziehungen aus dem euklidischen Algorithmus aus:

$a_0 = a$, $b_0 = b$, $q_0 = a \operatorname{div} b$.

Und für $i \geq 1$: $a_i = b_{i-1}$, $b_i = a_{i-1} \bmod b_{i-1}$, $q_i = a_i \operatorname{div} b_i$

Und dann rückwärts (Sei k der Index der letzten Zeile):

$x_k = 1$, $y_k = 0$.

$x_i = y_{i+1}$, $y_i = x_{i+1} - q_i \cdot y_{i+1}$

5.4 *Primzahlen*

In diesem Abschnitt betrachten wir einige Eigenschaften von Primzahlen, die später nützlich sein werden.

Lemma 5.26. *Sei p eine Primzahl und $a, b \in \mathbb{Z}$. Dann gilt:*

$$p \mid ab \quad \Rightarrow \quad p \mid a \text{ oder } p \mid b.$$

Beweis. Wir beweisen die Aussage durch Kontraposition. Gelte $p \nmid a$ und $p \nmid b$. Wir zeigen nun, dass $p \nmid ab$ gilt.

Aus $p \nmid a$ folgt $\mathrm{ggT}(p, a) = 1$. Ebenso folgt aus $p \nmid b$, dass $ggT(p, b) = 1$ gilt. Damit erhalten wir:

$$M_{p,a} = M_{p,b} = 1\mathbb{Z} = \mathbb{Z}.$$

Insbesondere gibt es $x, x', y, y' \in \mathbb{Z}$ mit $px + ay = px' + by' = 1$. Also ist

$$\begin{aligned} 1 &= (px + ay)(px' + by') \\ &= p(pxx' + bxy' + ayx') + abyy' \end{aligned}$$

Darum ist $1 \in M_{p,ab}$ und somit $ggT(p, ab) = 1$. Letztlich gilt damit $p \nmid ab$. □

Der folgende Satz ist in Teilen schon aus dem vorherigen Kapitel bekannt:

Satz 5.27 (Fundamentalsatz der Arithmetik)**.** *Jede natürliche Zahl $n \geq 2$ lässt sich (bis auf die Reihenfolge der Faktoren) auf eindeutige Weise als Produkt*

$$n = p_1^{m_1} \cdots p_k^{m_k}$$

schreiben, wobei die p_i positive Primzahlen und die m_i positive natürliche Zahlen sind.

Beweis. Das eine Zerlegung existiert haben wir in Kapitel 4 gesehen. Wir müssen also nur noch die Eindeutigkeit zeigen: Angenommen es gebe zwei Zerlegungen:

$$n = p_1^{m_1} \cdots p_k^{m_k} = q_1^{n_1} \cdots q_l^{n_l}.$$

Kommt eine Primzahl nur auf einer Seite vor, etwa p_1, dann gilt

$$p_1 \mid q_1^{n_1} \cdots q_k^{n_k},$$

aber $p_1 \nmid q_i$ für alle $i = 1, \ldots, l$ im Widerspruch zu Lemma 5.26. Deshalb gilt:

$$n = p_1^{m_1} \cdots p_k^{m_k} = p_1^{n_1} \cdots p_k^{n_k},$$

und wir müssen nur noch zeigen, dass $m_i = n_i$ für alle $i = 1, \ldots, k$ gilt.

Ist aber $m_1 \neq n_1$, etwa $m_1 < n_1$, dann wäre

$$p_2^{m_2} \cdots p_k^{m_k} = p_1^{n_1 - m_1} \cdots p_k^{n_k}$$

und p_1 würde $p_2^{m_2} \cdots p_k^{m_k}$ teilen, was ebenfalls im Widerspruch zu Lemma 5.26 stehen würde. □

6
Kombinatorik

In diesem Kapitel beschäftigen wir uns mit Kombinatorik, genauer gesagt, mit der abzählenden Kombinatorik. Wir wollen also „Dinge“ abzählen. Etwa Teilmengen einer bestimmten Mächtigkeit einer vorgegebenen Menge, oder Abbildungen bestimmten Typs zwischen bestimmten Mengen.

6.1 Binomialkoeffizienten

Ein wichtiges Werkzeug in der Kombinatorik sind die Binomialkoeffizienten. Wir führen sie im Folgenden ein und zeigen einige Anwendungen.

Definition 6.1. Für zwei Zahlen $n, k \in \mathbb{N}$ mit $k \leq n$ ist der *Binomialkoeffizient* $\binom{n}{k}$ definiert durch:

$$\begin{aligned}\binom{n}{0} &= \binom{n}{n} := 1\\ \binom{n}{k} &= \binom{n-1}{k-1} + \binom{n-1}{k}.\end{aligned}$$

Beispiel 6.2. $\binom{0}{0} = 1$, $\binom{1}{0} = 1$, $\binom{1}{1} = 1$, $\binom{2}{0} = 1$, $\binom{2}{1} = \binom{1}{0} + \binom{1}{1} = 2$, $\binom{2}{2} = 1$.

Man erhält das *Pascal'sche Dreieck*:

$$\begin{array}{ccccccc}
 & & & \binom{0}{0} & & & \\
 & & \binom{1}{0} & & \binom{1}{1} & & \\
 & \binom{2}{0} & & \binom{2}{1} & & \binom{2}{2} & \\
\binom{3}{0} & & \binom{3}{1} & & \binom{3}{2} & & \binom{3}{3}
\end{array}$$

Dabei ergibt sich ein Eintrag im Inneren des Dreiecks immer als Summe der beiden oberhalb stehenden Einträge.

Der folgende Satz ist an vielen Stellen nützlich:

Satz 6.3 (Binomischer Lehrsatz)**.** *Sei $n \in \mathbb{N}$. Dann gilt:*

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}.$$

Beweis. Durch vollständige Induktion – Übung. □

Beispiel 6.4. Es gilt:

$$\begin{aligned}(a+b)^2 &= \binom{2}{0}a^0b^2 + \binom{2}{1}a^1b^1 + \binom{2}{2}a^2b^0 \\ &= a^2 + 2ab + b^2.\end{aligned}$$

Die Binomialkoeffizienten haben auch eine Bedeutung in der Mengenlehre.

Satz 6.5. *Ist M eine endliche Menge mit n Elementen, dann gilt für jedes $k \in \mathbb{N}$ mit $k \leq n$: M besitzt genau $\binom{n}{k}$ k-elementige Teilmengen.*

Beweis. Vollständige Induktion zur Übung. □

Definition 6.6. Die *Fakultät* $n!$ einer natürlichen Zahl n ist definiert durch: $0! := 1$ und für $n > 0$: $n! = (n-1)! \cdot n = 1 \cdot 2 \cdot 3 \cdots n$

Satz 6.7. *Für $n, k \in \mathbb{N}$ mit $k \leq n$ gilt:*

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

Beweis. Zur Übung. □

Beispiel 6.8. Für das Lottospiel *6 aus 49* gibt es

$$\binom{49}{6} = \frac{49!}{6!43!} = 13983816$$

mögliche Ziehungen (wenn man die Reihenfolge der gezogenen Zahlen nicht betrachtet). Dies entspricht der Anzahl von 6-elementigen Teilmengen einer 49-elementigen Menge.

6.2 *Das Prinzip Inklusion-Exklusion*

Satz 6.9 (Inklusion-Exklusion)**.** *Seien $A_1, \ldots, A_k$ endliche Mengen. Dann gilt:*

$$\begin{aligned}|A_1 \cup \cdots \cup A_k| = \quad & |A_1| + \cdots + |A_k| \\ & - |A_1 \cap A_2| - |A_1 \cap A_3| - \cdots - |A_2 \cap A_3| - \cdots \\ & + |A_1 \cap A_2 \cap A_3| + |A_1 \cap A_2 \cap A_4| + \cdots \\ & \cdots + (-1)^{j+1}|A_{i_1} \cap \cdots \cap A_{i_j}| \\ & \cdots + (-1)^{k+1}|A_1 \cap \cdots A_k|.\end{aligned}$$

Mit der Abkürzung $[n] = \{1, 2, \cdots, n\}$ können wir schreiben:

$$|A_1 \cup \cdots A_k| = \sum_{\substack{I \subseteq [k] \\ I \neq \emptyset}} (-1)^{|I|+1} \left| \bigcap_{i \in I} A_i \right|.$$

Beweis. Zu jedem $x \in A_1 \cup \cdots \cup A_k$ gehört ein $I \subseteq [k]$, sodass $i \in I$ genau dann gilt, wenn $x \in A_i$. Dieses x wird

$$\binom{|I|}{1} - \binom{|I|}{2} + \binom{|I|}{3} - \cdots + (-1)^{|I|+1}\binom{|I|}{|I|}$$

mal in der Summe gezählt. Insgesamt wird x

$$\begin{aligned}
-\left(-\binom{|I|}{1} + \binom{|I|}{2} - \cdots\right) &= -\sum_{i=1}^{|I|}\binom{|I|}{i}(-1)^i \\
&= 1 - \sum_{i=0}^{|I|}\binom{|I|}{i}(-1)^i \\
&= 1 - \sum_{i=0}^{|I|}\binom{|I|}{i}(-1)^i 1^{|I|-i} \\
&= 1 - (1-1)^{|I|} \\
&= 1
\end{aligned}$$

mal gezählt. (Hier wurde der binonmische Lehrsatz benutzt) □

Beispiel 6.10. Seien die drei Mengen $A_1 = \{a, b\}$, $A_2 = \{b, c\}$ und $A_3 = \{c, d\}$ gegeben. Es gilt:

$$\begin{aligned}
|A_1 \cup A_2 \cup A_3| &= |A_1| + |A_2| + |A_3| - |A_1 \cap A_2| - |A_1 \cap A_3| - |A_2 \cap A_3| + |A_1 \cap A_2 \cap A_3| \\
&= 2 + 2 + 2 - 1 - 0 - 1 + 0 \\
&= 4.
\end{aligned}$$

6.3 *Das Schubfachprinzip und injektive Abbildungen*

Das Schubfachprinzip ist eines der Beispiele von mathematischen Aussagen, bei welcher sich Nichtmathematiker fragen, warum man sich damit beschäftigen soll, denn intuitiv ist klar, dass sie gilt. Allgemein besagt das Prinzip, dass wenn man $m > n$ Gegenstände auf n Schubfächer verteilt, man in mindestens ein Schubfach mehr als einen Gegenstand legen muss. Das klingt erst einmal simpel, gleichwohl führt es zu weitreichenden Konsequenzen, wie wir im Weiteren sehen werden.

Zunächst brauchen wir ein Vorbereitungslemma.

Lemma 6.11. *Seien X, Y Mengen und $f\colon X \longrightarrow Y$ eine Abbildung. Die Relation $\sim$ auf X gegeben durch*

$$x_1 \sim x_2 :\Leftrightarrow f(x_1) = f(x_2)$$

ist eine Äquivalenzrelation.

Beweis. Die Relation $\sim$ ist reflexiv, da $f(x) = f(x)$ für alle $x \in X$ gilt, da f eine Abbildung ist. Symmetrie und Transitivität folgen daraus, dass die Relation $=$ diese Eigenschaften hat. □

Satz 6.12 (Schubfachprinzip)**.** *Seien X, Y endliche Mengen, $f\colon X \longrightarrow Y$ eine Abbildung. Ist $|X| > |Y|$, dann gibt es ein $y \in Y$ mit*

$$\left|f^{-1}(y)\right| \geq 2.$$

Beweis. Die Äquivalenzklassen der Relation aus Lemma 6.11 sind von der Form $f^{-1}(y)$ für $y \in f(X) \subseteq Y$ und definieren eine Partition von X. Also gilt

$$X = \bigcup_{y \in f(x)} f^{-1}(y)$$

und $f^{-1}(y_1) \cap f^{-1}(y_2) = \emptyset$ für $y_1 \neq y_2$. Daher erhalten wir

$$|X| = \sum_{y \in f(X)} \left|f^{-1}(y)\right| \leq \sum_{y \in Y} \left|f^{-1}(y)\right|.$$

Wäre $\left|f^{-1}(y)\right| \leq 1$ für alle $y \in Y$, so würde gelten: $|X| \leq |Y| \cdot 1$, was im Widerspruch zur Voraussetzung stetht. □

Korollar 6.13. *Sind X, Y Mengen und gilt $|X| > |Y|$, so gibt es keine injektive Abbildung $f\colon X \longrightarrow Y$.*

Die beiden folgenden Sätze sind Anwendungen des Schubfachprinzips. Sie stammen aus dem Aufgabenkatalog der Internationalen Mathematikolympiade.

Satz 6.14. *Unter 101 Zahlen der Menge $M = \{1, 2, \ldots, 200\}$ gibt es stets zwei teilerfremde Zahlen.*

Beweis. Wir wählen folgende Partition von M:

$$\{1,2\} \cup \{3,4\} \cup \{5,6\} \cup \cdots \cup \{199,200\}.$$

Dies sind 100 Schubfächer. Also müssen zwei der 101 Zahlen im selben Fach liegen. Aufeinanderfolgende Zahlen sind aber teilerfremd. □

Satz 6.15. *Unter 101 Zahlen der Menge $M = \{1, 2, \ldots, 200\}$ gibt es stets Zahlen a, b, sodass $a \mid b$ gilt.*

Beweis. Jede natürliche Zahl lässt sich eindeutig schreiben als $2^k \cdot u$, wobei $k \in \mathbb{N}$ und u eine ungerade Zahl ist. Wähle nun 101 Zahlen $a_1 \ldots, a_{101}$ aus der Menge M aus und schreibe jeweils

$$a_i = 2^{k_i} \cdot u_i.$$

In M gibt es aber nur 100 ungerade Zahlen. Also gibt es unter den a_i zwei Zahlen a_x und a_y, mit $u_x = u_y$ und $a_x > a_y$. Also gilt $a_y \mid a_x$. □

Bemerkung 6.16. Sind X, Y Mengen und gilt $|X| \leq |Y|$, so gibt es wenigstens eine injektive Abbildung $f\colon X \longrightarrow Y$.

Es stellt sich nun die Frage, wie viele es wohl geben mag.

Notation 6.17. Sind X und Y endliche Mengen, so seien

$$\begin{aligned} A(X,Y) &:= \{\text{Abbildungen } X \longrightarrow Y\} \\ I(X,Y) &:= \{f \in A(X,Y) \mid f \text{ injektiv}\} \\ S(X,Y) &:= \{f \in A(X,Y) \mid f \text{ surjektiv}\} \\ B(X,Y) &:= \{f \in A(X,Y) \mid f \text{ bijektiv}\} \end{aligned}$$

Satz 6.18. *Sind X und Y Mengen und gelte*

$$1 \leq |X| = k \leq |Y| = n,$$

dann ist

$$\big|I(X,Y)\big| = \frac{n!}{(n-k)!}.$$

Beweis. Wir beweisen die Aussage durch Induktion über $n+k$.

Induktionsanfang $n+k=2$, also $n=k=1$: Es gibt in diesem Fall genau eine Abbildung, also gilt

$$\big|A(X,Y)\big| = \big|I(X,Y)\big| = 1 = \frac{1!}{0!}.$$

Induktionsschritt: Es ist $n+k \geq 2$. Wir wählen ein $x_0 \in X$ und schreiben $Y = \{y_1, \ldots, y_n\}$. Dann gilt:

$$\begin{aligned} I(X,Y) = \quad & \big\{f \in A(X,Y) \mid f(x_0) = y_1 \wedge \widetilde{f} \in I(\widetilde{X}, Y \setminus \{y_1\})\big\} \\ & \cup \big\{f \in A(X,Y) \mid f(x_0) = y_2 \wedge \widetilde{f} \in I(\widetilde{X}, Y \setminus \{y_2\})\big\} \\ & \cup \cdots \cup \big\{f \in A(X,Y) \mid f(x_0) = y_n \wedge \widetilde{f} \in I(\widetilde{X}, Y \setminus \{y_n\})\big\} \end{aligned}$$

Wobei $\widetilde{f}$ die Einschränkung von f auf $X \setminus \{x_0\} = \widetilde{X}$ ist. Diese Teilmengen sind disjunkt und bilden eine Partition von $I(X,Y)$. Daher gilt:

$$\big|I(X,Y)\big| = \sum_{y \in Y} \Big|\big\{f \in A(X,Y) \mid f(x_0) = y \wedge \widetilde{f} \in I(\widetilde{X}, Y \setminus \{y\})\big\}\Big|.$$

Nach Induktionsvoraussetzung gilt dann

$$\begin{aligned} \big|I(X,Y)\big| &= \sum_{y \in Y} \frac{(n-1)!}{\big((n-1)-(k-1)\big)!} \\ &= |Y| \cdot \frac{(n-1)!}{(n-k)!} \\ &= n \cdot \frac{(n-1)!}{(n-k)!} \\ &= \frac{n!}{(n-k)!}. \end{aligned}$$

□

Bemerkung 6.19. Für $|X| = 0$ gilt: Es gibt genau eine Abbildung $f\colon \emptyset \longrightarrow Y$ und diese ist injektiv, also $\big|I(X,Y)\big| = 1$.

Korollar 6.20. *Ist $|X| = |Y| = n$, so ist $I(X,Y) = B(X,Y)$ und damit $\big|B(X,Y)\big| = n!$. Also gilt:*

$$\big|B(X,Y)\big| = \begin{cases} n!, & \textit{wenn } |X| = |Y| = n \\ 0, & \textit{sonst}. \end{cases}$$

6.4 Bijektive Abbildungen und Permuationen

In diesem Abschnitt wollen wir bijektive Abbildungen auf endlichen Mengen untersuchen. Wie wir im vorherigen Abschnitt gesehen haben, kann es eine bijektive Abbildung nur zwischen gleichmächtigen Mengen geben.

Haben wir also eine Menge X und eine Menge Y gegeben mit $|X| = |Y| = n$, so können wir uns eine bijektive Abbildung

$$f\colon X \longrightarrow Y$$

vorstellen als eine Umbenennung der Elemente von Y durch die „Namen“ der Elemente von X.

Ist $X = \{x_1, \ldots, x_n\}$ eine endliche Menge mit n Elementen und hat man eine bijektive Abbildung

$$\pi\colon X \longrightarrow X$$

gegeben, so kann man sich diese Vorstellen als eine Vertauschung der Reihenfolge der Elemente von X. Das heißt, jedes x_i nimmt den Platz von $\pi(x_i)$ ein.

Definition 6.21. Eine bijektive Abbildung $\pi\colon X \longrightarrow X$ auf einer endlichen Menge X heißt *Permutation*.

Wenn X und Y endliche Mengen sind und $|X| = |Y|$ gilt, dann bedeuten injektiv, surjektiv und bijektiv das selbe.

Satz 6.22. *Es seien X und Y endliche Mengen mit $|X| = |Y| = n$ und $f\colon X \longrightarrow Y$ eine Abbildung. Dann ist f genau dann surjektiv, wenn f injektiv ist.*

Beweis. Sei f surjektiv, dann gilt $f(X) = Y$. Angenommen, f ist nicht injektiv, das heißt, es gibt $x_1, x_2 \in X$ mit $x_1 \neq x_2$ und $f(x_1) = f(x_2)$. Dann wäre $\big|f(X)\big| < n$ im Widerspruch zur Annahme, dass $f(X) = Y$ ist. Also muss f injektiv sein.

Sei f umgekehrt injektiv. Dann ist $\big|f(X)\big| = n$ und somit $f(X) = Y$. Also ist f surjektiv. □

Da es zu jeder Menge mit n Elementen eine bijektive Abbildung in die Menge $\{1, \ldots, n\}$ gibt, ist es für die weiteren Betrachtungen ausreichend, nur diese Menge zu betrachten.

Definition 6.23. Die Menge aller bijektiven Abbildungen

$$\sigma\colon \{1, \ldots, n\} \longrightarrow \{1, \ldots, n\}$$

bezeichnen wir mit S_n und nennen dies *symmetrische Gruppe*. Die Elemente von S_n, können wir wie folgt darstellen:

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{pmatrix}$$

Dies ist eine zweizeilige Matrix, wobei in der oberen Zeile die Elemente aus $\{1,\dots,n\}$ stehen und in der unteren Zeile steht jeweils das Bild $\sigma(i)$ von i unter i.

Beispiel 6.24. Für $n = 1,2,3$ sehen die S_n wie folgt aus:

$$S_1 = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}$$
$$S_2 = \left\{ \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \right\}$$
$$S_3 = \left\{ \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \right\}$$

Definition 6.25. Die Permutation, die jedes Element auf sich selbst abbildet, heißt *Identität* und wird mit id bezeichnet:

$$\mathrm{id} = \begin{pmatrix} 1 & 2 & \cdots & n \\ 1 & 2 & \cdots & n \end{pmatrix}$$

Bemerkung 6.26. Sind $\sigma, \tau \in S_n$, so ist auch das *Produkt* $\sigma\tau$ eine Permutation aus S_n. Es gilt: $\sigma\tau = \sigma \circ \tau$.

Im Allgemeinen gilt: $\sigma\tau \neq \tau\sigma$.

Offensichtlich gilt für alle $\sigma \in S_n$: $\mathrm{id} \circ \sigma = \sigma \circ \mathrm{id} = \sigma$.

Lemma 6.27. *Zu jedem $\sigma \in S_n$ gibt es eine Permutation $\sigma' \in S_n$, so dass $\sigma\sigma' = \sigma'\sigma = \mathrm{id}$ ist.*

Beweis. Da σ eine bijektive Abbildung der Form

$$\sigma\colon \{1,2,\dots,n\} \longrightarrow \{1,2,\dots,n\}$$
$$k \longmapsto \sigma(k)$$

ist, gibt es zu σ eine Umkehrabbildung σ^{-1} der Form

$$\sigma^{-1}\colon \{1,2,\dots,n\} \longrightarrow \{1,2,\dots,n\}$$
$$\sigma(k) \longmapsto k.$$

Für Umkehrabbildungen gilt: $\sigma \circ \sigma^{-1} = \sigma^{-1} \circ \sigma = \mathrm{id}$. Damit ist σ^{-1} das gesuchte σ'. □

Die Permutation σ^{-1} heißt *inverse Permutation* zu σ. Ist σ gegeben durch

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{pmatrix},$$

dann ist σ^{-1} diejenige Permutation, die wir erhalten, wenn wir die beiden Zeilen von σ tauschen:

$$\sigma^{-1} = \begin{pmatrix} \sigma(1) & \sigma(2) & \cdots & \sigma(n) \\ 1 & 2 & \cdots & n \end{pmatrix}.$$

Beispiel 6.28. Die inverse Permutation zu

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

ist

$$\sigma^{-1} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}.$$

Die Zweizeilige Darstellung von Permutationen ist etwas unhandlich. Wir führen daher eine einfachere Darstellung ein.

Definition 6.29. Eine Permutation $\sigma \in S_n$ heißt *Zyklus* der Länge l oder *l-Zyklus*, wenn es verschiedene $a_1, \ldots, a_l \in \{1, \ldots, n\}$ gibt, so dass

- $\sigma(a_i) = a_{i+1}$ für $i \in \{1, \ldots, l-1\}$,
- $\sigma(a_l) = a_1$,
- $\sigma(x) = x$ für $x \in \{1, \ldots, n\} \setminus \{a_1, \ldots, a_l\}$.

Einen solchen Zyklus schreiben wir als $\sigma = (a_1 \cdots a_l)$.

Zyklen der Länge 2 heißen *Transpositionen*

Die Darstellung eines Zyklus ist nicht eindeutig, wir können die Darstellung *zyklisch vertauschen*. Es gilt also:

$$(a_1 a_2 \cdots a_l) = (a_2 a_3 \cdots a_l a_1) = \cdots = (a_l a_1 \cdots a_{l-1}).$$

Üblich ist es die Zyklen so darzustellen, dass der kleinste Eintrag an erster Stelle steht.

Beispiel 6.30. Die Zyklusdarstellung von $\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$ ist

$$\sigma = (132) = (321) = (213).$$

Dies interpretieren wir als $1 \mapsto 3 \mapsto 2 \mapsto 1$. Die Zyklusdarstellung von $\tau = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$ ist $\tau = (13)$. Dies bedeutet $1 \mapsto 3 \mapsto 1$ und $2 \mapsto 2$. Ausführlich können wir schreiben $\tau = (13)(2)$. Üblich ist es aber, Zyklen der Länge 1 nicht in der Zyklusdarstellung aufzuführen.

Eine Ausnahme bildet der Zyklus (1), den wir in der Zyklusdarstellung für die Identität benutzen. Dies ist eine Konvention, denn offsichtlich gilt für die Identität in S_n:

$$\mathrm{id} = (1)(2)\cdots(n).$$

Diese Darstellung verkürzt man nun zu $\mathrm{id} = (1)$.

Notation 6.31. Transpositionen, welche die Zahlen i und j vertauschen, bezeichnen wir mit τ_{ij}.

Beispiel 6.32. In Zyklenschreibweise berechnet man Produkte wie folgt: In S_6 ist

$$(154623)(123)(45) = (13562)(4) = (13562).$$

Um diese Rechnung durchzuführen, geht man die Zyklen von rechts nach links durch. Üblich ist es, den ersten Zyklus mit einer 1 beginnen zu lassen, falls 1 nicht auf sich selbst abgebildet wird. Man sucht sich also den am weitest rechts stehenden Zyklus, der die 1 enthält. In diesem Beispiel wird die 1 auf die 2 abgebildet, im nächsten Zyklus wird die 2 auf die 3 abgebildet, daher wird insgesamt die 1 auf die 3 abgebildet. Die 3 wird im ersten Zyklus (welcher die 3 enthält, von rechts betrachtet) auf die 1 abgebildet, im nächsten Zyklus wird die 1 auf die 5 abgebildet. Daher wird insgesamt die 3 auf die 5 abgebildet, usw. Dies führt man durch, bis man wieder bei der 1 ankommt, dann nimmt man die kleinste Zahl aus $\{1, \dots, n\}$, welche noch nicht in diesem Zyklus steht und macht mit dieser weiter, usw.

Nicht jede Permutation lässt sich als Zyklus schreiben, z. B. ist $(12)(34) \in S_4$ Produkt zweier Zyklen, lässt sich aber nicht als ein Zyklus schreiben.

Lemma 6.33. *Es gilt für jeden Zyklus:* $(a_1 a_2 \dots a_l)^{-1} = (a_l a_{l-1} \dots a_1)$.

Beweis. Es gilt $(a_1 a_2 \dots a_l)^{-1}(a_1 a_2 \dots a_l) = (1)$. Daraus folgt die Behauptung. □

Definition 6.34. Zwei Zyklen $(a_1 \dots a_l)$ und $(b_1 \dots b_k)$ heißen *disjunkt*, falls die Mengen $\{a_1, \dots, a_l\}$ und $\{b_1, \dots, b_k\}$ disjunkt sind.

Der folgende Satz gibt einen Zusammenhang zwischen Permutationen und Zyklen an. Auf einen Beweis verzichten wir an dieser Stelle.

Satz 6.35. *Jedes Element* $(1) \neq \sigma \in S_n$ *ist Produkt paarweiser disjunkter Zyklen* $\neq (1)$. *Diese Darstellung ist (bis auf Reihenfolge der Faktoren) eindeutig.*

Da jede Permutation Produkt von Zyklen ist, wollen wir nun eine Darstellung aus Transpositionen abgeben:

Lemma 6.36. *Jedes Element von* S_n *lässt sich schreiben Produkt von Elementen aus den folgenden Mengen:*

1. *der Menge alle Transpositionen* $\{\tau_{ij} \mid 1 \leq i < j \leq n\}$,
2. *der Menge aller benachbarten Transpositionen:* $\{\tau_{i(i+1)} \mid 1 \leq i \leq n-1\}$,
3. *der Menge* $\{\tau_{12}, (123 \dots n), (123 \dots n)^{-1}\}$

Beweis. 1. Für jeden l-Zyklus $(a_1 a_2 a_3 \dots a_l)$ gilt:

$$(a_1 a_2 a_3 \dots a_l) = (a_1 a_l)(a_1 a_{l-1}) \cdots (a_1 a_3)(a_1 a_2),$$

2. Wir müssen zeigen, dass sich jede Transposition τ_{ij} schreiben lässt als Produkt von Elementen aus $\{\tau_{i(i+1)} \mid 1 \leq i \leq n-1\}$. Wir beweisen dies durch Induktion über $j-i$, wobei $1 \leq j-i \leq n-1$ ist. Im Induktionsanfang ist $j-i=1$, also $j=i+1$. Die Transposition $\tau_{i(i+1)}$ ist eine benachbarte Transposition, daher gilt die Behauptung. Sei jetzt also $j-i \geq 2$. Insbesondere ist $j-1 \neq i$. Dann gilt:
$$\tau_{(j-1)j}\tau_{i(j-1)}\tau_{(j-1)j} = \tau_{ij}$$
$\tau_{(j-1)i}$ ist eine benachbarte Transposition und da $(j-1)-i < j-i$ ist, lässt sich nach Induktionsvoraussetzung $\tau_{i(j-1)}$ als Produkt benachbarter Transpositionen schreiben, und damit auch τ_{ij}.

3. Sei $\sigma = (123\ldots n)$. Dann gilt: $\sigma^{i-1}\tau_{12}\sigma^{1-i} = \tau_{i(i+1)}$. Also können aus benachbarte Transpositionen als Produkte von Elementen aus $\{\tau_{12}, (123\ldots n), (123\ldots n)^{-1}\}$ geschrieben werden. Nach 2. können damit auch alle Permutationen so geschrieben werden.

□

Aus dem ersten Teil der Aussage von Lemma 6.36 erhalten wir das folgende Korollar:

Korollar 6.37. *Jede Permutation $\sigma \in S_n$ lässt sich als Produkt von $n-r$ Transpositionen schreiben, wobei r die Anzahl der Zyklen von σ sind, deren Länge größer 1 ist.*

Wir betrachten noch eine Beispielrechnung:

Beispiel 6.38. In S_8 gilt:

$$\begin{aligned}
(234)(4567)(1248) &= (1348)(2567)\\
&= (18)(14)(13)(27)(26)(25)\\
&= (12)(23)(34)(45)(56)(67)(78)(67)(56)(45)(34)\\
&\quad (23)(12)(12)(23)(34)(23)(12)(12)(23)(12)(23)\\
&\quad (34)(45)(56)(67)(56)(45)(34)(23)(23)(34)(45)\\
&\quad (56)(45)(34)(23)(23)(34)(45)(34)(23)\\
&= (12)(23)(34)(45)(56)(67)(78)(67)(56)(45)(12)\\
&\quad (23)(34)(45)(56)(67)(34)(23)
\end{aligned}$$

Wir haben also die Permutation zunächst in disjunkten Zyklen dargestellt, dann als Produkt von Transpositionen und zuletzt als Produkt von benachbarten Transpositionen. Wir sehen auch, dass die Anzahl an Transpositionen, die man benötigt um eine Permutation darzustellen, nicht eindeutig ist. Es gilt aber der folgende Satz:

Satz 6.39. *Ist $\sigma \in S_n$ eine Permutation auf verschiedene Weise dargestellt als Produkt von Transpositionen, so ist die Anzahl dieser Transpositionen entweder stets gerade oder ungerade.*

Beweis. Sei $\sigma \in S_n$ Produkt von s Transpositionen: $\sigma = \tau_1 \cdots \tau_s$. Außerdem sei eine vollständige Zerlegung von σ in disjkunkte Zyklen

(inkl. derer der Länge 1) gegeben: $\sigma = \zeta_1 \cdots \zeta_r$. Multipliziert man nun σ von links mit einer Transposition $\tau = (a\,b)$, so erhält man zwei mögliche Fälle:

1. Fall: a und b sind im selben Zyklus enthalten. Da die Zyklen disjunkt sind, können wir ohne Beschränkung der Allgemeinheit annehmen, dass dies der erste Zyklus $\zeta_1 = (a_1 \cdots a_k)$ ist und das gilt: $a = a_1$ und $b = a_i$. Wir betrachten die Wirkung von $\tau\sigma$:

$$\begin{aligned} a_1 &\overset{\sigma}{\longmapsto} a_2 \overset{\tau}{\longmapsto} a_2 \\ &\vdots \\ a_{i-1} &\overset{\sigma}{\longmapsto} a_i \overset{\tau}{\longmapsto} a_1 \\ a_i &\overset{\sigma}{\longmapsto} a_{i+1} \overset{\tau}{\longmapsto} a_{i+1} \\ &\vdots \\ a_k &\overset{\sigma}{\longmapsto} a_1 \overset{\tau}{\longmapsto} a_i \end{aligned}$$

Also ist $\tau\sigma = (a_1 \cdots a_{i-1})(a_i \cdots a_k)\zeta_2 \cdots \zeta_r$.

2. Fall: a und b sind in verschiedenen Zyklen enthalten. Ohne Beschränkung der Allgemeinheit können wir annehmen, dass dies die ersten beiden Zyklen $\zeta_1 = (a_1 \cdots a_k)$ und $\zeta_2 = (b_1 \cdots b_l)$ sind und das gilt: $a = a_1$ und $b = b_1$. Dann ist $\tau\sigma = (a_1 \cdots a_k b_1 \cdots b_l)\zeta_3 \cdots \zeta_r$.

Im ersten Fall hat sich die Anzahl der Zyklen um 1 erniedrigt, im zweiten Fall um 1 erhöht. Multipliziert man von links mit einer weiteren Transposition, so erhält man r, $r-2$ oder $r+2$ Zyklen. Wenn man q Transpositionen von links multipliziert, erhält man $r + t_q$ Zyklen, wobei $t_q \equiv q \pmod 2$ ist. Für das Produkt $\tau_s \cdots \tau_1\sigma$ erhält man also $r + t_s$ Zyklen, wobei $t_s \equiv s \pmod 2$ ist. Das Produkt ist aber die identische Permutations $\sigma^{-1}\sigma$, daher ist $r + t_s = n$ und somit gilt

$$s \equiv n - r \pmod 2$$

unabhängig davon, mit welcher Zerlegung in Transpositionen man gestartet ist. □

Definition 6.40. Für $\sigma \in S_n$ heißt

$$\operatorname{sign}\sigma = \prod_{1 \le i < j \le n} \frac{\sigma(j) - \sigma(i)}{j - i} \in \mathbb{Q}$$

Signum oder *Vorzeichen* von σ.

Lemma 6.41. *Das Signum besitzt die folgenden Eigenschaften:*

1. *Für $\sigma, \tau \in S_n$ gilt:* $\operatorname{sign}(\sigma\tau) = (\operatorname{sign}\sigma)(\operatorname{sign}\tau)$.
2. *Ist τ eine Transposition, so gilt* $\operatorname{sign}\tau = -1$.
3. *Ist σ Produkt von s Transpositionen, so gilt* $\operatorname{sign}\sigma = (-1)^s$.

4. *Ist σ ein l-Zyklus, so gilt* $\operatorname{sign}\sigma = (-1)^{l-1}$.

Beweis. 1.

$$\begin{aligned}\operatorname{sign}(\sigma\tau) &= \prod_{i<j} \frac{\sigma\tau(j) - \sigma\tau(i)}{j-i} \\ &= \prod_{i<j} \frac{\sigma\tau(j) - \sigma\tau(i)}{\tau(j) - \tau(i)} \prod_{i<j} \frac{\tau(j)\tau(i)}{j-i} \\ &= (\operatorname{sign}\sigma)(\operatorname{sign}\tau)\end{aligned}$$

2. Für $r \neq s$ gilt:

$$\operatorname{sign}\tau_{rs} = \frac{r-s}{s-r} \prod_{\substack{j=1 \\ j\neq r,s}}^{n} \frac{j-s}{j-r} \prod_{\substack{j=1 \\ j\neq r,s}}^{n} \frac{j-r}{j-s} = -1.$$

3. Folgt aus 1. und 2.

4. Folgt aus 3. und dem Fakt, dass $(a_1 a_2 \dots a_l) = (a_1 a_2)(a_2 a_3) \cdots (a_{l-1} a_l)$ gilt.

□

Definition 6.42. Permutationen $\sigma \in S_n$ mit $\operatorname{sign}\sigma = 1$ heißen *gerade*, solche mit $\operatorname{sign}\sigma = -1$ heißen *ungerade*.

Wie wir im nächsten Kapitel sehen werden, gibt es genau so viele gerade wie ungerade Permutationen, nämlich je $\frac{1}{2}n!$ viele in S_n.

7
Gruppentheorie

Gruppen verallgemeinern dass Konzept des *Rechnens in einer Menge*. Wenn wir zwei ganze Zahlen addieren, kommt dabei wieder eine ganze Zahl heraus; Wenn wir zwei Permutationen verknüpfen, erhalten wir wieder eine Permutation. Zwischen den beiden Operationen addieren von ganzen Zahlen und verknüpfen von Permutationen gibt es noch mehr Gemeinsamkeiten. Zum Beispiel gilt in beiden Fällen das Assoziativgesetz, es gibt ein neutrales Element und es gibt zu jedem Element ein Inverses. Es gibt aber auch einige Unterschiede. So ist z. B. die Addition auf $\mathbb{Z}$ kommutativ, es gilt also $a + b = b + a$, während die Verknüpfung von Permuationen dies nicht ist.

Wir wollen das *Rechnen in einer Menge* nun in eine einheitliche Theorie zusammenfassen.

7.1 Grundbegriffe

Definition 7.1. Sei M eine nichtleere Menge. Eine *Verknüpfung* $*$ auf M ist eine Abbildung

$$\begin{aligned} *\colon M \times M &\longrightarrow M \\ (a, b) &\longmapsto a * b. \end{aligned}$$

Beispiel 7.2. Wir kennen zum Beispiel schon: $+$, $\cdot$ auf $\mathbb{Z}$.

Auch auf der Menge der Abbildungen von einer Menge M in sich selbst ist die Hintereinanderausführung von Abbildungen eine Verknüpfung.

Auf der Potenzmenge $\mathcal{P}(M)$ einer Menge M kann man Verknüpfungen definieren:

$$\begin{aligned} \cup\colon \mathcal{P}(M) \times \mathcal{P}(M) &\longrightarrow \mathcal{P}(M) \\ (A, B) &\longmapsto A \cup B \end{aligned}$$

$$\begin{aligned} \cap\colon \mathcal{P}(M) \times \mathcal{P}(M) &\longrightarrow \mathcal{P}(M) \\ (A, B) &\longmapsto A \cap B \end{aligned}$$

Definition 7.3. Es sei G eine nichtleere Menge mit Verknüpfung $*$. $(G, *)$ heißt *Gruppe*, falls die folgenden drei Bedingungen erfüllt sind:

1. Für alle $a, b, c \in G$ gilt: $a * (b * c) = (a * b) * c$ (Assoziativgesetz)
2. Es gibt ein Element $e \in G$, sodass für alle $a \in G$ gilt: $e * a = a * e = a$ (Neutrales Element)
3. Für alle $a \in G$ gibt es ein $a' \in G$, so dass gilt: $a * a' = a' * a = e$ (Inverse Elemente)

Statt a' schreibt man oft a^{-1} für das inverse Element zu a. Ist die Verknüpfung auf G klar, so schreibt man oft auch ab statt $a * b$.

Die Anzahl der Elemente einer Gruppe G nennt man *Ordnung* von G und schreibt dafür $|G|$.

Definition 7.4. Eine Gruppe $(G, *)$ heißt *abelsch*, falls für alle $g, h \in G$ gilt: $g * h = h * g$.

Beispiel 7.5. Wichtige Beispiele für Gruppen sind:

1. $(\mathbb{Z}, +)$, dies ist eine abelsche Gruppe.
2. $(\mathbb{Q}, +)$, $(\mathbb{Q} \setminus \{0\}, \cdot)$ sind abelsche Gruppen.
3. $(\mathbb{Z}, \cdot)$ ist hingegen keine Gruppe, da z. B. 2 kein Inverses (in $\mathbb{Z}$) besitzt.
4. $(S_n, \circ)$ ist eine Gruppe, die *Symmetrische Gruppe*. Für $n \geq 3$ ist diese nicht abelsch.
5. Die Mengen $n\mathbb{Z}$ mit $n > 0$ sind mit der Addition als Verknüpfung Gruppen.

Für die Gruppen $n\mathbb{Z}$ gilt: $n\mathbb{Z} \subset \mathbb{Z}$. Sie sind also Teilmenge einer Gruppe und selbst wieder eine Gruppe. Wir verallgemeinern dies zu folgender Definition:

Definition 7.6. Ist $(G, *)$ eine Gruppe und $U \subseteq G$ eine Teilmenge von G. Dann heißt U *Untergruppe* von G, falls $(U, *)$ selbst wieder eine Gruppe ist. Wir schreiben: $U \leq G$.

Dabei ist zu beachten, dass die Gruppenverknüpfung für U dieselbe ist wie für G.

Beispiel 7.7. Wie schon gesehen gilt $n\mathbb{Z} \leq \mathbb{Z}$. Jede Gruppe G ist Untergruppe von sich selbst: $G \leq G$. Ist e das neutrale Element einer Gruppe G, so ist $\{e\}$ ebenfalls Untergruppe von G: $\{e\} \leq G$. Diese beiden Untergruppen heißen die *trivialen Untergruppen*

Die Gruppe $\mathbb{Z}_6$ besitzt folgende Untergruppen:

1. $\{0\}$,
2. $\{0, 3\}$,

3. $\{0, 2, 4\}$,

4. $\mathbb{Z}_6$.

Beispiel 7.8. Ein wichtiges Beispiel für Gruppen sind die *Restklassengruppen modulo n*:

$$(\mathbb{Z}_n, +) = \big(\{0, 1, \ldots, n-1\}, +\big),$$

wobei die Addition stets modulo n betrachtet wird. Diese Gruppe ist abelsch, 0 ist das neutrale Element und für jedes $0 \neq a \in \mathbb{Z}_n$ ist $n - a$ das inverse Element zu a. Für $0 \in \mathbb{Z}_n$ ist 0 selbst das inverse Element.

Hingegen ist $(\mathbb{Z}_n, \cdot)$ mit Multiplikation modulo n keine Gruppe, da 0 kein inverses besitzt. Wenn wir die 0 weglassen, $(\mathbb{Z}_n \setminus \{0\}, \cdot)$, erhalten wir für $n = 2, 3$ eine Gruppe, für $n = 4$ hingegen ist $2 \cdot 2 = 0 \notin \mathbb{Z}_n \setminus \{0\}$. Die Menge $\mathbb{Z}_n \setminus \{0\} = \{1, 2, 3\}$ ist also durch die Multiplikation nicht abgeschlossen, mithin ist dies also keine Verknüpfung auf der Menge.

Eine andere wichtige Klasse von Gruppen sind die *primen Restklassengruppen modulo n*:

$$(\mathbb{Z}_n^*, \cdot) = \big(a \in \mathbb{Z}_n \mid \mathrm{ggT}(a, n) = 1\big).$$

Die Multiplikation ist hierbei modulo n zu betrachten. Die Gruppen sind abelsch, 1 ist das neutrale Element und die inversen Element findet man mittels euklidischen Algorithmus.

Betrachte:

$$\mathbb{Z}_{14}^* = \{1, 3, 5, 9, 11, 13\}$$

Was ist das inverse Element zu 3?

Berechne dazu mit dem erweiterten euklidischen Algorithmus:

$$\begin{aligned} 1 &= \mathrm{ggT}(3, 14) \\ &= 3 \cdot 5 + (-1) \cdot 14. \end{aligned}$$

Also ist $3 \cdot 5 \equiv 1 \pmod{14}$. Somit ist 5 das Inverse zu 3. Analog findet man: $9^{-1} = 11$, $13^{-1} = 13$ in $\mathbb{Z}_{14}^*$.

Die Definition von Gruppen, die wir benutzen, ist recht stark formuliert. Es stellt sich heraus, dass man nicht alle Forderungen braucht um eine äquivalente Definition zu erhalten. Wir können die Definition von Gruppen auch schwächer formulieren:

Lemma 7.9. *Sei G eine Menge mit Verknüpfung $*$ auf G. Dann ist $(G, *)$ eine Gruppe, falls gilt:*

1. *Die Verknüpfung $*$ ist assoziativ, d. h. es gilt für alle $a, b, c \in G$: $a * (b * c) = (a * b) * c$.*

2. *Es gibt ein* linksneutrales *Element $e \in G$, so dass für alle $g \in G$ gilt: $e * g = g$.*

3. *Zu jedem Element $g \in G$ gibt es ein* linksinverses *Element $h \in G$ für das gilt: $h * g = e$.*

Beweis. Übung. □

In der uns vetrauten Gruppe der ganzen Zahlen $(\mathbb{Z}, +)$ haben wir nur ein neutrales Element, die 0. Es gibt also kein zweites Element $0' \in \mathbb{Z}$, so dass für alle $m \in \mathbb{Z}$ gilt $m + 0' = 0' + m = m$. Ebenso gibt es zu jedem Element $m \in \mathbb{Z}$ nur ein einziges Element m', so dass $m' + m = m' + m = 0$ ist ($-m$ ist dieses Element). Diese beiden Eigenschaften von $\mathbb{Z}$ sind kein Zufall, wir zeigen nun, dass das neutrale Element und die Inversen eindeutig sind.

Lemma 7.10. *Sei $(G, *)$ eine Gruppe. Dann gilt:*

1. *Sind $e_1, e_2 \in G$ beides neutrale Elemente, so gilt $e_1 = e_2$.*
2. *Sind $h_1, h_2 \in G$ beides Inverse zu $g \in G$, so gilt: $h_1 = h_2$.*

Beweis. Übung. □

Lemma 7.11. *Für alle $g, h \in G$ gilt:*

1. $\left(g^{-1}\right)^{-1} = g$ *und*
2. $(g * h)^{-1} = h^{-1} * g^{-1}$.

Beweis. 1. Für $h = g^{-1}$ gilt nach Definition $g * h = h * g = e$, also ist g das Inverse zu h, in Zeichen: $g = h^{-1}$. Durch einsetzen von h erhalten wir: $g = \left(g^{-1}\right)^{-1}$.

2. Für $k = g * h$ gilt:

$$\begin{aligned} k * (h^{-1} * g^{-1}) &= (g * h) * (h^{-1} * g^{-1}) \\ &= g * (h * h^{-1}) * g^{-1} \\ &= g * e * g^{-1} \\ &= g * g^{-1} \\ &= e. \end{aligned}$$

Also ist $k = g * h$ das Inverse zu $h^{-1} * g^{-1}$.

□

Lemma 7.12. *Es sei $G = (G, *)$ eine Gruppe. Dann gilt:*

1. *Für alle $a, b, c \in G$ gelten die Kürzungsregeln:*

$$\begin{aligned} ac = bc &\Rightarrow a = b \\ ca = cb &\Rightarrow a = b. \end{aligned}$$

2. *Für alle $a, b \in G$ gibt es genau ein $x \in G$, sodass gilt: $ax = b$, nämlich $x = a^{-1}b$ und es gibt genau ein $y \in G$, für das gilt: $ya = b$, nämlich $y = ba^{-1}$.*

Beweis. 1. Es gelte $ac = bc$. Dann gilt ebenfalls:

$$\begin{aligned} (ac)c^{-1} &= (bc)c^{-1} && \Rightarrow \\ a(cc^{-1}) &= b(cc^{-1}) && \Rightarrow \\ ae &= be && \Rightarrow \\ a &= b \end{aligned}$$

Die zweite Kürzungsregel zeigt man analog.

2. Mit $x = a^{-1}b \in G$ gilt:

$$\begin{aligned} ax &= a(a^{-1}b) \\ &= (aa^{-1})b \\ &= eb \\ &= b. \end{aligned}$$

Diese Lösung ist eindeutig, denn angenommen es gelte:

$$ax = b = ax',$$

Dann würde auch gelten:

$$ax = ax'$$

und nach der Kürzungsregel dann auch

$$x = x'.$$

Analog beweist man den zweiten Teil.

□

Lemma 7.13. *Sei G eine Menge und $*$ eine Verknüpfung auf G. Gelten zudem*

1. *$*$ ist assoziativ,*
2. *G ist endlich,*
3. *die Kürzungsregeln,*

dann ist $(G, *)$ eine Gruppe.

Beweis. Sei $g \in G$ beliebig aber fest gewählt. Die Abbildungen

$$\begin{aligned} \varphi_g \colon G &\longrightarrow G \\ h &\longmapsto g * h \end{aligned}$$

und

$$\begin{aligned} \psi_g \colon G &\longrightarrow G \\ h &\longmapsto h * g \end{aligned}$$

sind injektiv (wegen der Kürzungsregeln). Da G eine endliche Menge ist, sind φ_g und ψ_g sogar bijektiv. Somit gibt es ein Element $e_g \in G$

für das gilt: $\psi_g(e_g) = e_g * g = g$ und für jedes Element $h \in G$ gibt es ein $k \in G$ mit $\varphi_g(k) = h$. Damit:

$$\begin{aligned} e_g * h &= e_g * (g * k) \\ &= (e_g * g) * k \\ &= g * k \\ &= h. \end{aligned}$$

Also ist e_g linksneutrales Element in G. Da ψ_g bijektiv ist, gibt es ein $h \in G$ mit $\psi_g(h) = h * g = e$, d. h. h ist linksinverses Element zu g. Nach Lemma 7.9 ist $(G, *)$ also eine Gruppe. □

Für endliche Mengen nutzen wir oftmals eine *Verknüpfungstabelle*, um die Verknüpfung auf der Menge darzustellen:

$+$	0	1	2	3	4
0	0	1	2	3	4
1	1	2	3	4	0
2	2	3	4	0	1
3	3	4	0	1	2
4	4	0	1	2	3

$\cdot$	0	1	2	3	4
0	0	0	0	0	0
1	0	1	2	3	4
2	0	2	4	1	3
3	0	3	1	4	2
4	0	4	3	2	1

Tabelle 7.1: Additionstabelle für $(\mathbb{Z}_5, +)$ und Multiplikationstabelle für $(\mathbb{Z}_5, \cdot)$

Lemma 7.14. *Es gilt für endliche Mengen G mit assoziativer Verknüpfung $*$: $(G, *)$ ist eine Gruppe genau dann, wenn jedes $g \in G$ in der Verknüpfungstabelle genau einmal in jeder Zeile und jeder Spalte auftaucht.*

Beweis. Wir zeigen zunächst, dass die Kürzungsregeln äquivalent dazu sind, dass in jeder Zeile und jeder Spalte jedes Element nur einmal auftaucht.

Angenommen, die Kürzungsregeln gelten nicht. Dann gibt es Elemente $g, h, k \in G$, für die gilt: $g * h = g * k$ und $h \neq k$. Dann steht aber in der Zeile von g in den Spalten h und k der selbe Eintrag, im Widerspruch zur Annahme.

Gebe es nun umgekehrt eine Zeile g in der Verknüpfungstabelle, in der ein Eintrag zweimal auftaucht, z. B. in den Spalten h und k. Dann würde gelten: $g * h = g * k$ und $h \neq k$, und die Kürzungsregeln würden nicht gelten.

Also sind die Kürzungsregeln äquivalent dazu dass in jeder Zeile und jeder Spalte jedes Element nur einmal auftaucht.

Die Behauptung folgt nun aus Lemma 7.13 □

7.2 *Homomorphismen*

Wir wollen nun Abbildungen zwischen Gruppen untersuchen. Dabei sind wir besonders an Abbildungen interessiert, welche mit den Gruppenoperationen verträglich sind.

Definition 7.15. Seien $(G, *)$ und $(H, \circ)$ Gruppen. Eine Abbildung $\varphi\colon G \longrightarrow H$ heißt *((Gruppen-)Homomorphismus*, falls für alle $g, h \in G$ gilt:

$$\varphi(g * h) = \varphi(g) \circ \varphi(h).$$

Ist φ bijektiv, so nennt man φ *Isomorphismus*. Ist $\varphi\colon G \longrightarrow G$ ein Isomorphismus von G nach G, so heißt φ *Automorphismus*.

Beispiel 7.16. 1. $\varphi\colon (\mathbb{Z}, +) \longrightarrow (\mathbb{Z}, +)$ mit $\varphi(x) = 2x$ ist ein Homomorphismus.

2. $\varphi\colon (G, *) \longrightarrow (\{e\}, *)$ mit $\varphi(g) = e$ für alle $g \in G$ ist der triviale Homomorphismus.

3. Für abelsche Gruppen G ist $\varphi\colon (G, *) \longrightarrow (G, *)$ mit $\varphi(g) = g^{-1}$ ein Automorphismus, denn $\varphi(g * h) = (g * h)^{-1} = h^{-1} * g^{-1} = g^{-1} * h^{-1} = \varphi(g) * \varphi(h)$.

4. Für jede Gruppe $(G, *)$ mit festgewähltem $g \in G$ ist die *Konjugationsabbildung* $c_g\colon G \longrightarrow G$ mit $c_g(h) = g * h * g^{-1}$ ein Automorphismus. Das c_g ein Homomorphismus ist, folgt aus: $c_g(h * k) = g * h * k * g^{-1} = g * h * e * k * g^{-1} = g * h * g^{-1} * g * k * g^{-1} = c_g(h) * c_g(k)$

5. $\varphi\colon \mathbb{Z}_2 \longrightarrow \mathbb{Z}_4$ mit $\varphi(0) = 0$, $\varphi(1) = 2$ ist ein injektiver Homomorphismus.

6. $\varphi\colon \mathbb{Z}_4^* \longrightarrow \mathbb{Z}_2$ mit $\varphi(1) = 0$, $\varphi(3) = 1$ ist ein Isomorphismus, denn z. B. gilt $\varphi(1 \cdot 3) = \varphi(3) = 1 = \varphi(1) + \varphi(3)$

7. $(\{-1, 1\}, \cdot)$ ist eine abelsche Gruppe der Ordnung 2.

$$\begin{aligned} \operatorname{sign}\colon S_n &\longrightarrow \{-1, 1\} \\ \sigma &\longmapsto \operatorname{sign}\sigma \end{aligned}$$

ist ein Homomorphismus, der jeder Permutation ihr Signum zuweist, denn es gilt (siehe Kapitel 6): $\operatorname{sign}\sigma\tau = \operatorname{sign}\sigma \operatorname{sign}\tau$ für alle $\sigma, \tau \in S_n$.

Lemma 7.17. *Seien $(G, *)$ und $(H, \circ)$ Gruppen, e_G das neutrale Element von G, e_H das neutrale Element von H und $\varphi\colon G \longrightarrow H$ ein Gruppenhomomorphismus. Dann gelten:*

$$\varphi(e_G) = e_H,$$

und

$$\varphi\left(g^{-1}\right) = \left(\varphi(g)\right)^{-1}.$$

Beweis. Es gilt:

$$\begin{aligned} \varphi(e_G) &= \varphi(e_G) \circ e_H \\ &= \varphi(e_G * e_G) \circ e_H \\ &= \varphi(e_G) \circ \varphi(e_G) \circ e_H. \end{aligned}$$

Nach der Kürzungregel gilt dann:

$$e_H = \varphi(e_G) \circ e_H = \varphi(e_G).$$

Damit ist die erste Aussage gezeigt.

Es gilt:

$$\begin{aligned}\varphi\left(g^{-1}\right) \circ \varphi(g) &= \varphi\left(g^{-1} * g\right) \\ &= \varphi(e_G) \\ &= e_H.\end{aligned}$$

Da die Inversen eindeutig bestimmt sind, ist insbesondere das Inverse zu $\varphi(g)$ eindeutig bestimmt und ergibt sich nach obiger Rechnung zu $\varphi\left(g^{-1}\right)$. □

Definition 7.18. Sei $\varphi\colon G \longrightarrow H$ ein Gruppenhomomorphismus. Das *Bild* von φ ist definiert durch

$$\operatorname{Im}(\varphi) \coloneqq \{\varphi(g) \mid g \in G\}.$$

Der *Kern* von φ ist definiert durch

$$\ker(\varphi) \coloneqq \{g \in G \mid \varphi(g) = e_H\}.$$

Lemma 7.19. *Sind $\varphi\colon G \longrightarrow H$ und $\psi\colon H \longrightarrow K$ Gruppenhomomorphismen, so ist auch $\psi \circ \varphi\colon G \longrightarrow K$ ein Gruppenhomomorphismus. Sind φ und ψ Isomorphismen, so ist auch $\psi \circ \varphi$ ein Isomorphismus.*

Beweis. Übung. □

Das folgende Kriterium macht es einfach einen Gruppenhomomorphismus auf Injektivität zu überprüfen.

Satz 7.20. *Ist $\varphi\colon G \longrightarrow H$ ein Gruppenhomomorphismus, dann ist φ genau dann injektiv, wenn $\ker(\varphi) = \{e_G\}$ gilt.*

Beweis. Sei $g \in \ker(\varphi)$. Dann gilt: $\varphi(g) = e_H = \varphi(e_G)$. Wenn φ injektiv ist, dann ist $g = e_G$, also $\ker(\varphi) = \{e_G\}$.

Ist umgekehrt $\varphi(g) = \varphi(h)$ für $g, h \in G$, dann gilt:

$$e_H = \varphi(g)(\varphi(h))^{-1} = \varphi(g)\varphi\left(h^{-1}\right) = \varphi\left(gh^{-1}\right).$$

Also ist $gh^{-1} \in \ker(\varphi)$. Da $\ker(\varphi) = \{e_G\}$, muss $gh^{-1} = e_G$, und somit $g = h$ gelten. Also ist φ injektiv. □

7.3 Untergruppen

In diesem Abschnitt wollen wir Untergruppen genauer untersuchen. Dabei beschreiben wir Eigenschaften von Untergruppen und geben äquivalente Definitionen an.

Zunächst beschreiben wir die Untergruppen von $(\mathbb{Z}, +)$.

Satz 7.21. *Sei U eine Untergruppe von $\mathbb{Z}$. Dann ist U von der Form $n\mathbb{Z}$ mit $n \in \mathbb{N}$. Dabei ist n die kleinste natürliche Zahl in U.*

Beweis. Im Fall $U = \{0\}$ ist $n = 0$. Ist $U \neq \{0\}$, dann ist zu jedem Element $u \in U$ auch $(-u) \in U$. Daher gibt es natürliche Zahlen ungleich 0 in U. Sei n die kleinste positive Zahl in U und $m \in U$ beliebig. Der Satz über Division mit Rest liefert $q, r \in \mathbb{Z}$ mit $m = nq + r$ mit $0 \leq r < n$. Daher liegt auch $m - nq = r$ in U. Da $r < n$ gilt, und n die kleinste positive Zahl in U ist, folgt $r = 0$. Daher ist $m = nq$ und daher $U = n\mathbb{Z}$. □

Das nachfolgende Lemma gibt uns äquivalente Definitionen für Untergruppen, die gegebenenfalls einfacher zu überprüfen sind.

Lemma 7.22. *Sei G eine Gruppe, $U \subseteq G$, $U \neq \emptyset$. Dann sind äquivalent:*

1. $U \leq G$,
2. $a, b \in U \Rightarrow ab^{-1} \in U$
3. $a, b \in U \Rightarrow ab \in U \wedge a^{-1} \in U$.

Beweis. (1) $\Rightarrow$ (3) ist klar, ebenso (3) $\Rightarrow$ (2). Für die Richtung (2) $\Rightarrow$ (1) bleibt zu zeigen: U hat ein neutrales Element, zu jedem Element $a \in U$ gibt es ein Inverses $a^{-1} \in U$ und für zwei Elemente $a, b \in U$ ist auch $ab \in U$. Wir überlassen dies dem geneigten Leser zur Übung. □

Für endliche Untergruppen wird es noch einfacher.

Lemma 7.23. *Sei G eine Gruppe, $H \subseteq G$ eine endliche, nichtleere Teilmenge von G. Dann sind äquivalent:*

1. *H ist eine Untergruppe von G,*
2. *für alle $u, v \in H$ gilt: $uv \in H$.*

Beweis. Die Hinrichtung ist klar, da Untergruppen insbesondere Gruppen sind. Nach Lemma 7.13 bleibt nur noch zu zeigen, dass die Kürzungsregeln in H gelten. Dies ist aber klar, da die Kürzungsregeln schon in G gelten und sich damit auf H übertragen. □

Verschiedene Untergruppen einer Gruppe G haben immer wenigstens ein Element gemeinsam, das neutrale Element e_G. Es kann aber auch sein, dass sie mehrere Elemente gemeinsam haben. Im Folgenden beschreiben wir die Eigenschaften solcher Schnitte von Untergruppen.

Lemma 7.24. *Der Durchschnitt jeder Familie von Untergruppen einer Gruppe G ist eine Untergruppe von G.*

Beweis. Sei $(U_i)_{i \in I}$ eine Familie von Untergruppen und $U := \bigcap U_i$. Dann gilt: $e \in U$, da jede Untergruppe das neutrale Element enthält. Sind nun $a, b \in U$, dann ist auch $a, b \in U_i$ für alle i. Dann ist aber auch $ab^{-1} \in U_i$ für jedes i und daher gilt: $ab^{-1} \in U$. Nach Lemma 7.22 ist dann U eine Untegruppe. □

Wir betrachten nun die Wirkung von Homomorphismen auf Untergruppen.

Lemma 7.25. *Sei $\varphi\colon G \longrightarrow H$ ein Gruppenhomomorphismus. Dann gilt:*

1. *Ist $U \leq G$, so ist $\varphi(U) \leq H$.*
2. *Ist $V \leq H$, so ist $\varphi^{-1}(V) \leq G$.*

Beweis. 1. Seien $g, h \in U$. Da U eine Untergruppe ist, ist auch $gh \in U$ und $g^{-1} \in U$. Es gilt dann:

$$\varphi(g)\varphi(h) = \varphi(gh) \in \varphi(U) \subseteq H,$$

und

$$\big(\varphi(g)\big)^{-1} = \varphi\left(g^{-1}\right) \in \varphi(U).$$

Nach Lemma 7.22 ist $\varphi(U)$ also eine Untergruppe von H.

2. Seien $g, h \in \varphi^{-1}(V) \subseteq G$. Dann sind $\varphi(g)$, $\varphi(h)$ in V und somit gilt $\varphi(gh) = \varphi(g)\varphi(h) \in V$. Damit ist aber auch $gh \in \varphi^{-1}(V)$. Außerdem ist wegen $\varphi(g^{-1}) = (\varphi(g))^{-1} \in V$ auch $g^{-1} \in \varphi^{-1}(V)$.

□

Bemerkung 7.26. Insbesondere ist also $\mathrm{Im}(\varphi) \leq H$ und $\ker \varphi \leq G$.

In Abschnitt Lineare Algebra werden uns viele der obigen Konstruktionen wieder in ähnlicher Form begegnen.

7.4 Ringe und Körper

Bisher haben wir mit Gruppen Mengen mit einer Verknüpfung betrachtet. Nun nehmen wir noch eine zweite Verknüpfung hinzu.

Definition 7.27. Sei K eine Menge und

$$\begin{aligned} +\colon K \times K &\longrightarrow K, \quad \text{, sowie} \\ \cdot\colon K \times K &\longrightarrow K \end{aligned}$$

zwei Verknüpfungen auf K. Ist $(K, +)$ eine abelsche Gruppe mit neutralem Element 0 und $(K \setminus \{0\}, \cdot)$ ebenfalls eine abelsche Gruppe und gelten zudem die Distributivgesetze:

$$\begin{aligned} a \cdot (b + c) &= a \cdot b + a \cdot c \quad \text{und} \\ (a + b) \cdot c &= a \cdot c + b \cdot c, \end{aligned}$$

so heißt $(K, +, \cdot)$ *Körper*.

Beispiel 7.28. $(\mathbb{R}, +, \cdot)$ und $(\mathbb{Q}, +, \cdot)$ sind Körper. Hingegen ist $(\mathbb{Z}, +, \cdot)$ kein Körper, da $(\mathbb{Z} \setminus \{0\}, \cdot)$ keine Gruppe ist.

Es gibt auch endliche Körper:

Satz 7.29. *Sei p eine Primzahl, dann ist $(\mathbb{Z}_p, +, \cdot)$ ein Körper.*

Beweis. Das $(\mathbb{Z}_p, +)$ eine abelsche Gruppe ist, wissen wir schon. Ebenso wissen wir, dass $(\mathbb{Z}_p \setminus \{0\}, \cdot) = (\mathbb{Z}_p^*, \cdot)$ eine abelsche Gruppe ist. Die Distributivgesetze sind auch klar. □

Definition 7.30. Sei $(K, +, \cdot)$ ein Körper. Falls es eine natürliche Zahl n gibt, für die gilt

$$\underbrace{1 + 1 + \cdots + 1}_{n\text{-mal}} = 0,$$

so heißt die kleinste solche Zahl *Charakteristik* ($\mathrm{char}(K)$) von K. Gibt es keine solche Zahl, so setzt man $\mathrm{char}(K) = 0$.

Satz 7.31. *Sei K ein Körper. Dann gilt:* $\mathrm{char}(K) = 0$ *oder* $\mathrm{char}(K) = p$, *wobei p eine Primzahl ist.*

Beweis. Offenbar gibt es Körper mit Charakterstik 0, z. B. $\mathbb{R}$. Sei K also ein Körper mit $\mathrm{char}(K) \neq 0$. Angenommen es würde gelten: $\mathrm{char}(K) = pq$, mit $p, q \neq 1$, $p, q \in \mathbb{N}$. Dann gilt:

$$0 = \underbrace{1 + \cdots + 1}_{pq\text{-mal}} = (\underbrace{1 + \cdots + 1}_{p\text{-mal}}) \cdot (\underbrace{1 + \cdots + 1}_{q\text{-mal}})$$

Also gibt es Elemente $a, b \in K \setminus \{0\}$, für die $a \cdot b = 0$ gilt. Dann ist aber $(K \setminus \{0\}, \cdot)$ keine Gruppe mehr, im Widerspruch zur Voraussetzung, dass K ein Körper ist. Also muss $\mathrm{char}(K)$ eine Primzahl sein. □

Definition 7.32. Sei R eine Menge mit zwei assoziativen Verknüpfungen $+$ und $\cdot$ auf K. Dann heißt $(R, +, \cdot)$ *Ring*, falls $(R, +)$ eine abelsche Gruppe ist und für $+, \cdot$ die Distributivgesetze gelten. Ein Ring heißt *kommutativ*, falls $\cdot$ eine kommutative Verknüpfung ist. Gilt für alle $a, b \in R$: $ab = 0 \Rightarrow a = 0 \vee b = 0$, so heißt der Ring R *nullteilerfrei*. Gibt es ein neutrales Element bezüglich $\cdot$ (oft wird dieses 1 genannt), so heißt R *Ring mit Eins*.

Beispiel 7.33. 1. $(\mathbb{Z}, +, \cdot)$ ist ein kommutativer, nullteilerfreier Ring mit Eins.

2. Jeder Körper ist auch ein kommutativer, nullteilerfreier Ring mit Eins.

3. $(\mathbb{Z}_n, +\cdot)$ ist ein kommutativer Ring mit Eins, aber nicht nullteilerfrei, falls n keine Primzahl ist.

4. Ist $(G, +)$ eine abelsche Gruppe, so ist $(G, +, *)$ mit der trivialen Multiplikation $a * b = 0$ für alle $a, b \in G$ ein Ring.

Ein paar intuitiv klare Eigenschaften von Ringen wollen wir im Folgenden beweisen:

Lemma 7.34. *In jedem Ring $(R, +, \cdot)$ gilt für alle $a \in R$: $a \cdot 0 = 0 \cdot a = 0$.*

Beweis. Es gilt: $a \cdot 0 = a \cdot (0 + 0) = a \cdot 0 + a \cdot 0$. Daher ist

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

Die andere Gleichheit zeigt man analog. □

Lemma 7.35. *Sei $(R, +, \cdot)$ ein Ring und $a, b \in R$ Dann gilt: $a \cdot (-b) = (-a) \cdot b = -(a \cdot b)$*

Beweis. Es gilt:

$$\begin{aligned} a \cdot b + a \cdot (-b) &= a \cdot \big(b + (-b)\big) \\ &= a \cdot 0 \qquad\qquad\qquad\qquad = 0 \end{aligned}$$

Da das additiv Inverse eindeutig bestimmt ist, gilt $a \cdot (-b) = -(a \cdot b)$. Analog folgt $(-a) \cdot b = -(a \cdot b)$. □

Definition 7.36. Ist $(R, +, \cdot)$ ein Ring mit Eins. Ein Element $a \in R$ heißt *Einheit*, wenn es ein $b \in R$ gibt, so dass gilt: $ab = ba = 1$.

Einheiten sind also die multiplikativ invertierbaren Elemente von R. Wir bezeichnen die Menge der Einheiten von R mit R^*.

Lemma 7.37. *Sei $(R, +, \cdot)$ ein Ring. Dann ist $(R^*, \cdot)$ eine Gruppe, die* Einheitengruppe *von R.*

Beweis. Die Assoziativität von $\cdot$ wird direkt vom Ring vererbt, offensichtlich ist auch wegen $1 = 1 \cdot 1$ das neutrale Element in R^* enthalten. Zu jedem $a \in R^*$ gibt es nach Voraussetzung ein $b \in R$, so dass $a \cdot b = b \cdot a = 1$ ist. Also ist $a^{-1} = b$ das Inverse zu a.

Letztlich bleibt noch zu zeigen, dass R^* abgeschlossen ist bezüglich $\cdot$. Seien also $a_1, a_2 \in R^*$. Dann gibt es nach Voraussetzung $b_1, b_2 \in R^*$, so dass $a_1 b_1 = b_1 a_1 = 1$ und $a_2 b_2 = b_2 a_2 = 1$ gelten. Dann folgt:

$$\begin{aligned} (a_1 \cdot a_2) \cdot (b_2 \cdot b_1) &= a_1 \cdot 1 \cdot b_1 \\ &= a_1 \cdot b_1 \\ &= 1; \end{aligned}$$

bzw.

$$\begin{aligned} (b_2 \cdot b_1) \cdot (a_1 \cdot a_2) &= b_2 \cdot 1 \cdot a_2 \\ &= b_2 \cdot a_2 \\ &= 1. \end{aligned}$$

Also ist $a_1 \cdot a_2 \in R^*$. □

Definition 7.38. Sei R ein kommutativer Ring mit Eins und X eine Unbestimmte. Dann nennt man einen Ausdruck der Form

$$P = a_0 X^0 + a_1 X^1 + \cdots + a_n X^n$$

mit $n \in \mathbb{N}$ und $a_0, \ldots, a_n \in R$ ein *Polynom über R*. Die a_i heißen Koeffizienten des Polynoms. Der *Grad* des Polynoms P ist definiert als

$$\operatorname{grad} P = \begin{cases} -1 & \text{, falls } a_0 = a_1 = \cdots = a_n = 0 \\ \max\{i \mid a_i \neq 0\} & \text{, sonst.} \end{cases}$$

Zwei Polynome $P = a_0X^0 + a_1X^1 + \cdots + a_nX^n$ und $Q = b_0X^0 + b_1X^1 + \cdots + b_mX^m$, mit $m \leq n$, sind gleich, falls für alle $0 \leq i \leq m$ gilt: $a_i = b_i$ und für alle $m+1 \leq u \leq n$ gilt: $a_i = 0$. Die Menge aller Polynome über R bezeichnen wir mit $R[X]$.

Wir wollen nur solche Polynome betrachten, bei denen der höchste vorkommende Koeffizient ungleich 0 ist. Überhaupt lassen wir in der Darstellung der Polynome die Summanden mit Koeffizient 0 weg, d. h. wir identifizieren z. B. $0X^3 + 1X^2 + 0X + 2$ direkt mit $1X^2 + 2$.

Auf $R[X]$ definieren wir die folgenden Verknüpfungen:

$$\begin{aligned} +\colon R[X] \times R[X] &\longrightarrow R[X] \\ \left(\sum_{i=0}^{n} a_iX^i, \sum_{i=0}^{m} b_iX^i\right) &\longmapsto \sum_{i=0}^{\max\{n,m\}} (a_i + b_i)X^i \\ \cdot\colon R[X] \times R[X] &\longrightarrow R[X] \\ \left(\sum_{i=0}^{n} a_iX^i, \sum_{i=0}^{m} b_iX^i\right) &\longmapsto \sum_{i=0}^{n+m} \left(\sum_{j=0}^{i} a_jb_{i-j}\right) X^i \end{aligned}$$

Das neutrale Element der Addition ist $0X^0$, das der Multiplikation $1X^0$. Damit ist $R[X]$ ein kommutativer Ring mit Eins, genannt *Polynomring über R*.

Beispiel 7.39. Betrachte $\mathbb{Z}[X]$ und $P = 3 - 2X + 4X^2$, $Q = -1 + 3X + 2X^2$, dann ist $P + Q = 2 + X + 6X^2$, $P \cdot Q = -3 + 11X - 4X^2 + 8X^3 + 8X^4$.

Es gilt: $\operatorname{grad}(P \cdot Q) = \operatorname{grad}(P) + \operatorname{grad}(Q)$, falls R nullteilerfrei ist und $P, Q \in R[X] \setminus \{0\}$.

Betrachte $\mathbb{Z}_5[X]$, und $P = 3X^2$, $Q = 1 + 4X + 2X^2$, dann ist $P + Q = 1 + 4X$ und $P \cdot Q = 3X^2 + 2X^3 + X^4$.

Analog zu den ganzen Zahlen kann man unter gewissen Umständen eine Division mit Rest auf $R[X]$ erklären:

Satz 7.40. *Sei K ein Körper, und $f \in K[X] \setminus \{0\}$. Dann gibt es zu jedem $g \in K[X]$ eindeutig bestimmte Polynome q und r aus $K[X]$, mit $g = q \cdot f + r$ und* $\operatorname{grad}(r) < \operatorname{grad}(f)$.

Beispiel 7.41. Betrachte $\mathbb{Q}[X]$, $g = X^5 + 3X^4 + X^3 - 6X^2 - X + 1$ und $f = X^3 + 2X^2 + X - 1$.

$$
\begin{array}{rrrrl}
X^5 + 3X^4 & + X^3 - 6X^2 & - X + 1 = \left(X^3 + 2X^2 + X - 1\right)\left(X^2 + X - 2\right) & - 2X^2 + 2X - 1 \\
\underline{- X^5 - 2X^4} & \underline{- X^3 + X^2} & & \\
X^4 & - 5X^2 & - X & \\
\underline{- X^4 - 2X^3} & \underline{- X^2 + X} & & \\
- 2X^3 - 6X^2 & & + 1 & \\
\underline{2X^3 + 4X^2 + 2X - 2} & & & \\
- 2X^2 + 2X - 1 & & &
\end{array}
$$

Also gilt

$$g = (X^2 + X - 2) \cdot f + (-2X^2 + 2X - 1).$$

Definition 7.42. Zu jedem Polynom p aus $R[X]$ definiert man die *Polynomfunktion* $p(x)$

$$\begin{aligned} p \colon R &\longrightarrow R \\ a &\longmapsto p(a) \end{aligned}$$

in dem man $a \in R$ in das Polynomeinsetzt und den Ausdruck auswertet.

Offenbar können verschiedene Polynome die selbe Polynomfunktion besitzen:

Beispiel 7.43. Die Polynome X und X^5 aus $\mathbb{Z}_5[X]$ definieren die selbe Polynomfunktion.

Wir werden Polynomfunktion in der Veranstaltung *Mathematische Grundlagen 2* genauer untersuchen aber sie werden auch in der linearen Algebra wieder auftauchen.

Definition 7.44. Ist R ein Ring und $I \subseteq R$ eine nicht-leere Teilmenge, so heißt I *Ideal*, wenn die folgenden beiden Bedingungen erfüllt sind:

- Für alle $a, b \in I$ gilt: $a - b \in I$ (I ist eine Untergruppe bzgl. $+$),
- Für alle $a \in i$ und $r \in R$ gilt: $r \cdot a \in I$ und $a \cdot r \in I$.

Beispiel 7.45. Offensichtlich ist $\{0\}$ ein Ideal in jedem Ring und das Ideal, welches die 1 enthält ist gleich dem ganzen Ring. Dies sind die beiden trivialen Ideale, dass *Nullideal* und das *Einsideal* von R.

Die Ideale von $\mathbb{Z}$ sind genau die Mengen $n\mathbb{Z}$ mit $n \in \mathbb{N}$.

Alle Polynome P, für die $P(0) = 0$ gilt, bilden ein Ideal in $\mathbb{R}[X]$.

Ideale haben viele Anwendungen in der algebraischen Zahlentheorie, insbesonder in der Codierungstheorie und Kryptographie.

8

Geometrie

Wir halten zunächst einige Bezeichnungen fest.

Definition 8.1. Die Menge der *reellen Zahlen* bezeichnen wir mit $\mathbb{R}$. Für jede natürliche Zahl $n > 0$ ist $\mathbb{R}^n = \underbrace{\mathbb{R} \times \cdots \times \mathbb{R}}_{n\text{-mal}}$.

8.1 Grundlagen

In diesem Kapitel wollen wir die Ebene bzw. den Raum beschreiben, insbesondere um Orte/Positionen zu bestimmen, Längen und Winkel zu messen und Richtungen angeben zu können. Letztendlich wollen wir so Formen beschreiben können.

Zunächst müssen wir uns auf eine Referenz einigen. Dazu wählen wir

- einen Referenzpunkt O, *Ursprung* genannt,
- eine Referenzlänge,
- eine Referenzrichtung, welche zusammen mit der Referenzlänge einen gerichteten Maßstab bildet.
- einen Referenzwinkel, den Vollkreis.

Um den Vollkreis sinnvoll unterteilen zu können, gibt es verschiedene Winkelmaße. Das in Deutschland und der EU gesetzliche Winkelmaß ist:

$$1 \text{ Vollkreis} = 360^\circ = 360 \text{ Grad}.$$

Im Vermessungswesen nutzt man

$$1 \text{ Vollkreis} = 400^g = 400 \text{ gon}.$$

In der Mathematik ist es üblich im Bogenmaß zu rechnen:

$$1 \text{ Vollkreis} = 2\pi.$$

Wir werden in der Vorlesung in Grad und Bogenmaß rechnen.

Definition 8.2. Ein Winkel mit dem Winkelmaß 90° bzw. $\frac{\pi}{2}$ heißt auch *rechter Winkel*.

Definition 8.3. Ein (geometrischer) *Vektor* v ist eine Angabe von Länge und Richtung.

Sind a, b, c Vektoren, dann können wir diese

- aneinanderlegen, d. h. addieren,
- strecken oder skalieren, d. h. mit einem Skalar (einer reellen Zahl) multiplizieren.

Sind P, Q Punkte in der Ebene oder im Raum, so ist $\overrightarrow{PQ}$ die Angabe des Abstandes und der Richtung von P nach Q. Für jeden Punkt P ist $\overrightarrow{OP}$ der *Ortsvektor* von P. Umgekehrt ist zu jedem Vektor v der Punkt $P(v)$ genau der Punkt mit $\overrightarrow{OP(v)} = v$. Der Ort aller Punkte $P(xv)$ für $x \in \mathbb{R}$ ist eine *Gerade* (diese ist also eine Menge von Punkten).

Definition 8.4. Es seien Vektoren $v_1, \ldots, v_n$ der Ebene gegeben. Jede Wahl von $x_1, \ldots, x_n \in \mathbb{R}$ bestimmt einen Punkt $P(x_1v_1 + \cdots + x_nv_n)$ der Ebene.

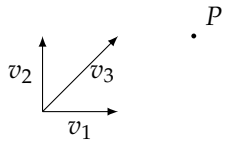
$v_1, \ldots, v_n$ heißen *Basis der Ebene*, wenn die Abbildung

$$\begin{aligned} \mathbb{R}^n &\longrightarrow \text{Ebene} \\ (x_1, \ldots, x_n) &\longmapsto P(x_1v_1 + \cdots + x_nv_n) \end{aligned}$$

bijektiv ist.

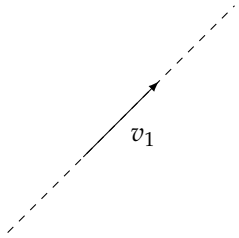
Bemerkung 8.5. Offenbar gilt für die obige Definition: $n = 2$.

Beispiel 8.6. Wir betrachten das folgende System von Vektoren und einem Punkt:



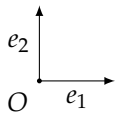
Offenbar werden unter der Abbildung aus Definition 8.4 die Tupel $(1, 2, 0)$ und $(0, 1, 1)$ auf P abgebildet. Die Abbildung ist also nicht injektiv.

Hingegen ist die Abbildung bei nur einem Vektor nicht surjektiv:



8.2 Standards

Definition 8.7. Die Standardbasis (der Ebene) bilden die beiden Vektoren e_1 und e_2, welche die Länge 1 haben und im rechten Winkel zueinander stehen:



Wenn nichts anderes gesagt wird, beziehen wir uns immer auf diese Basis, und schreiben kurz:

$$(x_1, x_2) \quad \text{für den Punkt} \quad P(x_1e_1 + x_2e_2),$$

und nennen dies die *Koordinaten* des Punktes. Wir schreiben auch

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad \text{für den Vektor} \quad x_1e_1 + x_2e_2,$$

und nennen dies die *Komponenten* des Vektors.

Wir können dann die Ebene mit dem $\mathbb{R}^2$ identifizieren, wobei der Ursprung den Koordinaten $(0,0)$ entspricht.

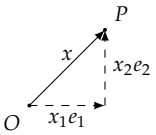
Offensichtlich hat der Ursprung O die Koordinaten $(0,0)$ und der Ortsvektor des Ursprungs $\overrightarrow{OO}$ die Komponenten $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$.

Lemma 8.8. *Mit diesen Definitionen gelten:*

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \end{pmatrix} \quad \textit{und}$$
$$\lambda \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \lambda x_1 \\ \lambda x_2 \end{pmatrix}.$$

Wir nehmen e_1 und e_2 als Grundmaßstäbe für die Bestimmung von Längen und Winkeln.

Definition 8.9. Der Abstand eines Punktes $P = (x_1, x_2)$ vom Ursprung ist die *Norm* des Vektors $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, nämlich:



$$\|x\| = \sqrt{x_1^2 + x_2^2}.$$

Satz 8.10 (Pythagoras)**.** *Seien a, b, c Vektoren in der Ebene, so dass zwischen a und b ein rechter Winkel ist und $\|c\| > \|a\|, \|b\|$. Dann gilt:*

$$\|a\|^2 + \|b\|^2 = \|c\|^2.$$

(Siehe Abbildung 8.1).

Beweis. Wir nehmen vier Kopien des Dreiecks und kleben Sie entsprechend aneinander:

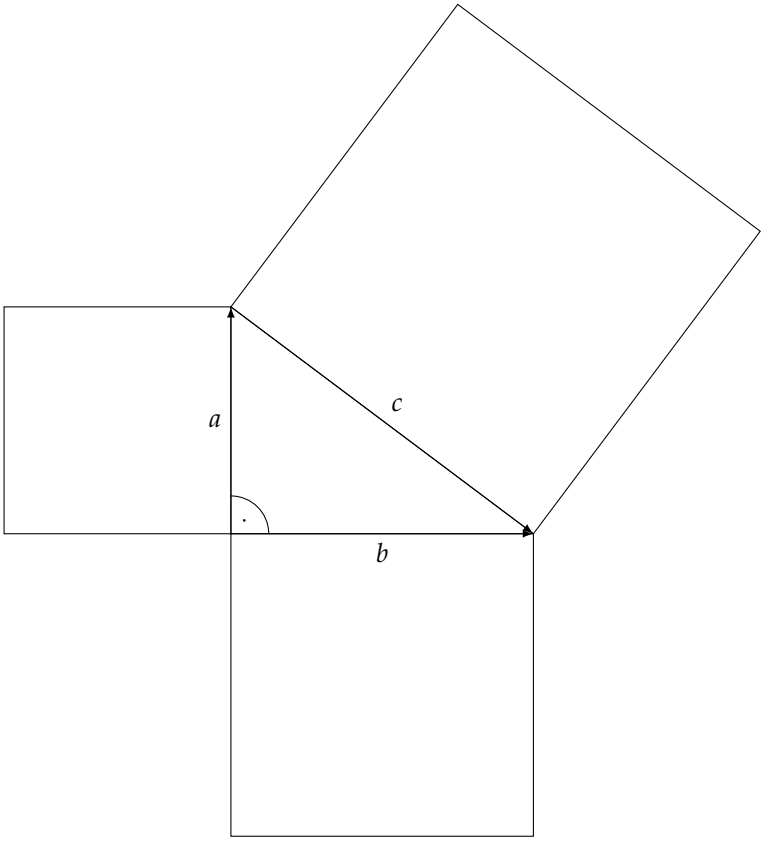
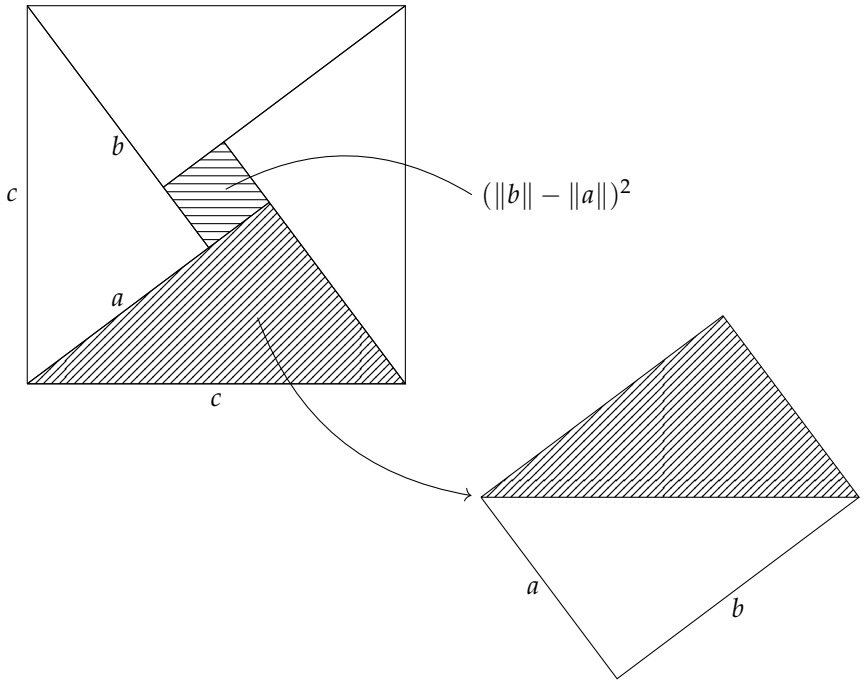


Abbildung 8.1: Zur Veranschaulichung des Satzes von Pythagoras.



Wir sehen, dass große Quadrat hat den Flächeninhalt $\|c\|^2$, das Rechteck aus zwei Dreiecken zusammengesetzt den Inhalt $\|a\|\|b\|$,

also hat ein Dreieck den Flächenihalt $\frac{\|a\|\|b\|}{2}$. Dann gilt:

$$\begin{aligned} c^2 &= 4 \cdot \frac{\|a\|\|b\|}{2} + (\|b\| - \|a\|)^2 \\ &= 2\|a\|\|b\| + \|b\|^2 - 2\|a\|\|b\| + \|a\|^2 \\ &= \|a\|^2 + \|b\|^2 \end{aligned}$$

□

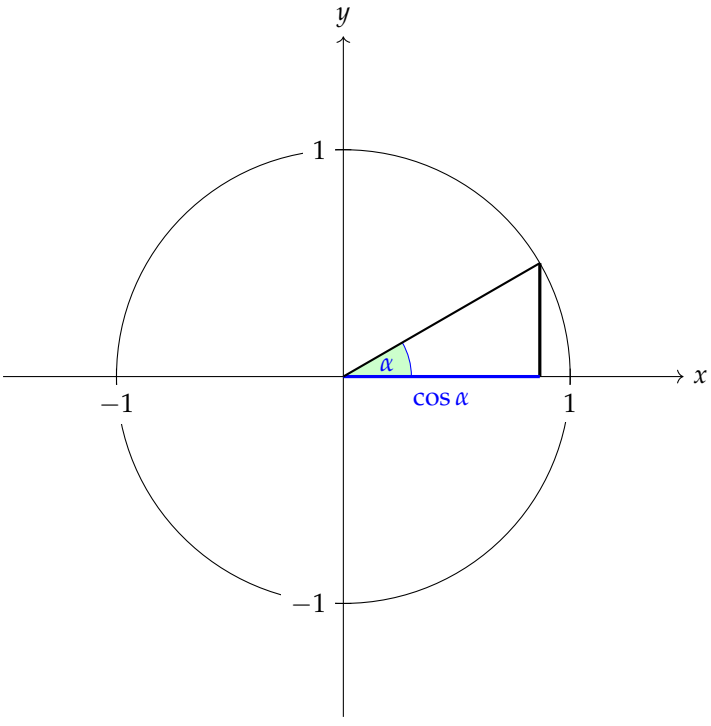
Der Abstand zwischen zwei Punkten P und Q ergibt sich wegen $\overrightarrow{PQ} = \overrightarrow{OQ} - \overrightarrow{OP}$ zu

$$\|\overrightarrow{PQ}\| = \left\| \begin{pmatrix} q_1 - p_1 \\ q_2 - p_2 \end{pmatrix} \right\| = \sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2}.$$

8.3 Winkel

Definition 8.11. Ein *Kreis mit Radius r um den Punkt P* ist die Menge aller Punkte Q in der Ebene, so dass $\|\overrightarrow{PQ}\| = r$ ist. Der *Einheitskreis* ist der Kreis um den Ursprung O mit Radius $r = 1$.

Definition 8.12. Zu einem Winkel α zwischen zwei Vektoren, definiert man den *Kosinus* von α, kurz $\cos(\alpha)$ am Einheitskreis:



Die Funktion $\cos(\alpha)$ nimmt offensichtlich Werte aus der Menge $\{x \in \mathbb{R} \mid -1 \leq x \leq 1\}$ an und ist 2π-periodisch. Außerdem gelten: $\cos(\alpha) = \cos(2\pi - \alpha)$ und $\cos(\alpha) = -\cos(\pi + \alpha)$ für alle $0 \leq \alpha \leq 2\pi$.

Wir notieren einige Standardwerte der Kosinusfunktion. Oft ist es hilfreich, diese auswendig zu können. Aufgrund der Periodizität und

α	$\cos(\alpha)$
0	1
$\frac{\pi}{6}$	$\frac{\sqrt{3}}{2}$
$\frac{\pi}{4}$	$\frac{\sqrt{2}}{2}$
$\frac{\pi}{3}$	$\frac{1}{2}$
$\frac{\pi}{2}$	0

Tabelle 8.1: Einige Standardwerte der Kosinusfunktion.

den Symmetrieeigenschaften der Kosinusfunktion kann man aus diesen wenigen Werten direkt viele weitere Werte ableiten, etwa $\cos\left(\frac{5\pi}{3}\right) = \frac{1}{2}$.

Definition 8.13. Zu Vektoren $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ und $y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ ist das *Skalarprodukt* $\langle x, y \rangle$ gegeben durch

$$\langle x, y \rangle = x_1 y_1 + x_2 y_2.$$

Lemma 8.14. *Für zwei Vektoren x und y gilt dann:*

$$\langle x, y \rangle = \|x\| \|y\| \cos(\alpha),$$

wobei α der Winkel zwischen x und y ist.

Insbesondere gilt für zwei Vektoren, die im rechten Winkel (Notation: $x \perp y$ zueinander stehen:

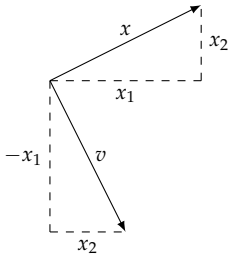
$$x \perp y \Leftrightarrow \langle x, y \rangle = 0.$$

Beispiel 8.15. Es sei $x = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ und $y = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Dann gilt:

$$\langle x, y \rangle = 1 \cdot 1 + 0 \cdot 1 = 1 = \|x\| \|y\| \cos(\alpha) = 1 \cdot \sqrt{2} \cos(\alpha).$$

Also ist hier $\cos(\alpha) = \frac{1}{\sqrt{2}} = \frac{\sqrt{2}}{2}$.

Lemma 8.16. *Sei* $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ *ein Vektor der Ebene. Dann steht der Vektor* $v = \begin{pmatrix} x_2 \\ -x_1 \end{pmatrix}$ *senkrecht auf x (also gilt* $\langle x, v \rangle = 0$*).*



Beweis. Es gilt:

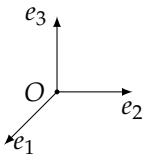
$$\left\langle \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} x_2 \\ -x_1 \end{pmatrix} \right\rangle = x_1 x_2 + (-x_1) x_2 = 0.$$

□

8.4 Die Standardbasis im dreidimensionalen Raum

Analog zum zweidimensionalen Fall:

Definition 8.17. Die Standardbasis bilden die Vektoren e_1, e_2 und e_3, welche die Länge 1 haben und paarweise im rechten Winkel zueinander stehen:



Es gilt die Rechte-Hand-Regel.

Wenn nichts anderes gesagt wird, beziehen wir uns immer auf diese Basis, und schreiben kurz:

$$(x_1, x_2, x_3) \quad \text{für den Punkt} \quad P(x_1e_1 + x_2e_2 + x_3e_3),$$

Wir schreiben auch

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \quad \text{für den Vektor} \quad x_1e_1 + x_2e_2 + x_3e_3,$$

Wir können dann den Raum mit dem $\mathbb{R}^3$ identifizieren, wobei der Ursprung den Koordinaten $(0, 0, 0)$ entspricht.

Lemma 8.18. *Mit diesen Definitionen gelten:*

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ x_3 + y_3 \end{pmatrix} \quad \textit{und}$$

$$\lambda \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \lambda x_1 \\ \lambda x_2 \\ \lambda x_3 \end{pmatrix}.$$

Definition 8.19. Zu Vektoren $x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ und $y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$ ist das *Skalarprodukt* $\langle x, y \rangle$ gegeben durch

$$\langle x, y \rangle = x_1y_1 + x_2y_2 + x_3y_3.$$

Analog zum zweidimensionalen Fall gilt auch hier:

$$\langle x, y \rangle = \|x\| \|y\| \cos(\alpha),$$

und

$$x \perp y \Leftrightarrow \langle x, y \rangle = 0.$$

8.5 Kurven und Flächen

Definition 8.20. Eine *Gerade* durch zwei gegebene Punkte P, Q besteht aus allen Punkten deren Ortsvektor sich schreiben lässt als

$$x = \overrightarrow{OP} + t\overrightarrow{PQ}, \quad \text{für ein } t \in \mathbb{R}.$$

In Koordinaten erhält man im $\mathbb{R}^3$ mit $P = (p_1, p_2, p_3)$ und $Q = (q_1, q_2, q_3)$:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix} + t \begin{pmatrix} q_1 - p_1 \\ q_2 - p_2 \\ q_3 - p_3 \end{pmatrix};$$

und im $\mathbb{R}^2$ mit $P = (p_1, p_2)$ und $Q = (q_1, q_2)$:

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} p_1 \\ p_2 \end{pmatrix} + t \begin{pmatrix} q_1 - p_1 \\ q_2 - p_2 \end{pmatrix}.$$

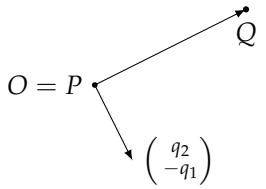
Das heißt, im $\mathbb{R}^2$ gilt dann für den Spezialfall $P = O$ und $Q \neq O$:

$$x_1 q_2 = t q_1 q_2 = x_2 q_1,$$

woraus folgt:

$$\left\langle \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} q_2 \\ -q_1 \end{pmatrix} \right\rangle = 0.$$

Die Gleichung hat dann folgende anschauliche Bedeutung:



Definition 8.21. Eine *Ebene* ist durch drei nicht kolineare Punkte A, B, C gegeben und besteht aus allen Punkten, deren Ortsvektor sich schreiben lässt als

$$x = \overrightarrow{OP} + t\overrightarrow{AB} + u\overrightarrow{AC}, \quad \text{für } t, u \in \mathbb{R}.$$

Wir schreiben auch

$$\mathcal{E} = \left\{ P(x) \,\middle|\, x = \overrightarrow{OA} + t\overrightarrow{AB} + u\overrightarrow{AC} \right\}$$

Lemma 8.22. *Für eine Ebene $\mathcal{E}$ im $\mathbb{R}^3$ gilt: Falls $A = O$ ist, so gibt es einen Vektor n, genannt der* Normalenvektor *zu der Ebene $\mathcal{E}$, so dass sich $\mathcal{E}$ schreiben lässt als*

$$\mathcal{E} = \{ P(x) \mid \langle x, n \rangle = 0 \}.$$

Beweis. Zur Übung! □

Definition 8.23. Durch einen Mittelpunkt M und einen Radius $r > 0$ wird im $\mathbb{R}^2$ ein *Kreis* (und analog im $\mathbb{R}^3$ eine Kugeloberfläche) wie folgt definiert:

$$K = \left\{ P(x) \,\middle|\, \left\| x - \overrightarrow{OM} \right\| = r \right\}.$$

8.6 Höhere Dimensionen, Unterräume

Wir haben gesehen, dass sich bei der Benutzung der Vektorschreibweise viele Dinge nicht ändern, wenn wir zwei- oder dreidimensionale Objekte betrachten. Wir können also im Allgemeinen auf die Intuition verzichten und den Sprung in höhere Dimensionen wagen.

Definition 8.24. Für jedes $n \in \mathbb{N}$ ist die Addition auf $\mathbb{R}^n$ definiert durch:

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix}.$$

Die *Skalarmultiplikation* ist definiert für $\lambda \in \mathbb{R}$ durch

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

Die Menge $\mathbb{R}^n$ mit diesen beiden Operationen wird *n-dimensionaler reeller Standardvektorraum* genannt.

Definition 8.25. Eine Basis des $\mathbb{R}^n$ ist ein Familie von Vektoren $b_1, \ldots, b_n$, sodass die Abbildung

$$\begin{aligned} \mathbb{R}^n &\longrightarrow \mathbb{R}^n \\ (\lambda_1, \ldots, \lambda_n) &\longmapsto \lambda_1 b_1 + \cdots + \lambda_n b_n \end{aligned}$$

bijektiv ist.

Bemerkung 8.26. Im Falle $n = 2, 3$ entspricht dies den vorher betrachteten Situationen.

Definition 8.27. Ein *(linearer) Unterraum* des $\mathbb{R}^n$ ist eine Teilmenge $U \subseteq \mathbb{R}^n$, für die gilt: Es gibt Vektoren $v_1, \ldots, v_m \in \mathbb{R}^n$, sodass

$$U = \{\lambda_1 v_1 + \cdots \lambda_m v_m \mid \lambda_i \in \mathbb{R}\} =: \operatorname{span}(v_1, \ldots, v_m)$$

gilt.

Lineare Unterräume sind also die Verallgemeinerung von Geraden und Ebenen durch den Ursprung.

Definition 8.28. Ein *affiner Unterraum* des $\mathbb{R}^n$ ist eine Teilmenge $A \subseteq \mathbb{R}^n$, für die gilt: Es gibt Vektoren $a, v_1, \ldots, v_m \in \mathbb{R}^n$, sodass

$$A = a + \operatorname{span}(v_1, \ldots, v_m)$$

gilt.

Affine Unterräume sind also die Verallgemeinerung von Geraden und Ebenen, die nicht durch den Ursprung gehen (müssen). Offenbar ist jeder lineare Unterraum auch ein affiner Unterraum (wobei a der Vektor der Länge Null ist).

8.7 Lineare Abbildungen

Definition 8.29. Eine *Bewegung* im $\mathbb{R}^2$ ist eine Abbildung $f\colon \mathbb{R}^2 \longrightarrow \mathbb{R}^2$, die Abstände und Winkel erhält:

Unter einer Bewegung wird also jeder Kreis um den Ursprung O auf einen Kreis um $f(O)$ abgebildet.

Falls $f(O) = O$ gilt, dann auch

- $f\big(P(\lambda x)\big) = \lambda f\big(P(x)\big)$, und
- $f\big(P(x+y)\big) = f\big(P(x)\big) + f\big(P(y)\big)$.

Deshalb können wir f als Abbildung von Ortsvektoren auffassen, mit

- $f(x+y) = f(x) + f(y)$
- $f(\lambda x) = \lambda f(x)$ für alle $\lambda \in \mathbb{R}$

Wir notieren dies allgemein als Definition:

Definition 8.30. Eine Abbildung zwischen Vektorräumen

$$f\colon \mathbb{R}^n \longrightarrow \mathbb{R}^m$$

heißt *linear*, wenn für alle $x, y \in \mathbb{R}^n$ und $\lambda \in \mathbb{R}$ gilt:

- $f(x+y) = f(x) + f(y)$ und
- $f(\lambda x) = \lambda f(x)$

Insbesondere gilt für $v_1, \ldots, v_n \in \mathbb{R}^n$ und $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$:

$$f\left(\sum_i \lambda_i v_i\right) = \sum_i \lambda_i f(v_i).$$

Sind $v_1, \ldots, v_n$ eine Basis, so bestimmen $f(v_1), \ldots, f(v_n)$ die Abbildung f vollständig:

Satz 8.31. *Eine lineare Abbildung $f\colon \mathbb{R}^n \longrightarrow \mathbb{R}^m$ ist durch die Angabe von $f(e_1), \ldots, f(e_n)$ eindeutig bestimmt.*

Beweis. Für alle $x \in \mathbb{R}^n$ gibt es eindeutige $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ mit

$$x = \lambda_1 e_1 + \cdots + \lambda_n e_n.$$

Also gilt

$$\begin{aligned} f(x) &= f(\lambda_1 e_1 + \cdots + \lambda_n e_n) \\ &= \lambda_1 f(e_1) + \cdots + \lambda_n f(e_n) \end{aligned}$$

Also genügt es $f(e_1), \ldots, f(e_n)$ zu kennen, um $f(x)$ bestimmen zu können. □

Tatsächlich gilt sogar:

$$f\left(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}\right) = x_1 f(e_1) + \cdots + x_n f(e_n).$$

Es ist üblich, die Werte $f(e_i)$ in einer *Matrix* darzustellen:

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} = \begin{pmatrix} f(e_1) & f(e_2) & \cdots & f(e_n) \end{pmatrix}.$$

Dies ist ein rechteckiges Schema von m Zeilen und n Spalten, eine $m \times n$-Matrix. Die i-te Spalten entspricht dabei dem Vektor $f(e_i)$. In Zusammenhang mit linearen Abbildungen spricht man von der *darstellenden Matrix* von f und bezeichnet sie mit M_f.

Beispiel 8.32. 1. $f\colon \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ sei die Drehung um den Ursprung um den Winkel $\pi/2$.

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

2. $f\colon \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ sei die Spiegelung an der Geraden λe_2.

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

3. $f\colon \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ sei die Drehung um die e_1-Achse um $\pi/2$.

$$M_f = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}$$

Für eine gegebene lineare Abbildung f mit darstellender Matrix M_f bestimmt man das Bild eines Vektors $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ unter f wie folgt:

$$\begin{aligned} M_f x = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} &:= x_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} + \cdots + x_n \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix} \\ &= \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots a + mnx_n \end{pmatrix}. \end{aligned}$$

Damit bestimmt jede $m \times n$-Matrix A eine lineare Abbildung

$$\begin{aligned} \mathbb{R}^n &\longrightarrow \mathbb{R}^m \\ x &\longmapsto Ax. \end{aligned}$$

9
Vektorräume

In diesem Kapitel beschäftigen wir uns mit Vektorräumen, einem wichtigen Konzept der linearen Algebra, mit welcher wir uns im zweiten Semester beschäftigen werden.

9.1 Vektorraum, Basis, Lineare Unabhängigkeit

Definition 9.1. Seien $(V, +)$ eine abelsche Gruppe und K ein Körper. V heißt *Vektorraum über K* (oder *K-Vektorraum*), falls es eine Abbildung

$$\begin{aligned} \cdot \colon K \times V &\longrightarrow V \\ (\lambda, v) &\longmapsto \lambda \cdot v \end{aligned}$$

(*Skalarmultiplikation* gennannt) gibt, so dass die folgenden Bedingungen erfüllt sind:

(V1) $\forall \lambda \in K \, \forall v, w \in V : \lambda \cdot (v + w) = \lambda \cdot v + \lambda \cdot w.$

(V2) $\forall \lambda, \mu \in K \, \forall v \in V : (\lambda + \mu) \cdot v = \lambda \cdot v + \mu \cdot v.$

(V3) $\forall \lambda, \mu \in K \, \forall v \in V : (\lambda \mu) \cdot v = \lambda \cdot (\mu \cdot v).$

(V4) $\forall v \in V : 1 \cdot v = v.$

Die Elemente von V heißen *Vektoren*.

Das folgende Beispiel ist unser Standardbeispiel für einen Vektorraum. Wir werden ihn immer wieder benutzen.

Beispiel 9.2. Ein häufig benutzter Vektorraum ist der Vektorraum $\mathbb{R}^n$. Die Elemente von $\mathbb{R}^n$ stellt man als Spaltenvektoren dar:

$$v = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$$

Aus Platzgründen schreibt man Vektoren manchmal als Zeilen, und macht dann durch ein hochgestelltes t deutlich, dass ein Spaltenvektor gemeint ist:

$$v = (v_1, v_2, \ldots, v_n)^t$$

Dabei bedeutet das t *transponiert*, ein Begriff, auf den wir später noch stoßen werden. Die Addition in $\mathbb{R}^n$ erfolgt komponentenweise:

$$\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} + \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} = \begin{pmatrix} v_1 + w_1 \\ v_2 + w_2 \\ \vdots \\ v_n + w_n \end{pmatrix}$$

Bezüglich der Addition ist der *Nullvektor*

$$\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

das neutrale Element, das Inverse zu einem Vektor

$$v = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} \text{ ist } \quad -v = \begin{pmatrix} -v_1 \\ -v_2 \\ \vdots \\ -v_n \end{pmatrix}.$$

Die Skalarmultiplikation für $\mathbb{R}^n$ ist gegeben durch

$$\begin{aligned} \cdot : \mathbb{R} \times \mathbb{R}^n &\longrightarrow \mathbb{R}^n \\ \left(\lambda, \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}\right) &\longmapsto \begin{pmatrix} \lambda v_1 \\ \vdots \\ \lambda v_n \end{pmatrix}. \end{aligned}$$

Üblicherweise identifiziert man $\mathbb{R}$ mit einer Geraden, $\mathbb{R}^2$ mit der Ebene, usw.:

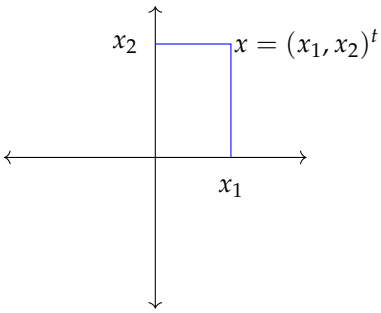


Abbildung 9.1: $\mathbb{R}^2$ kann als Koordinatenebene aufgefasst werden.

Lemma 9.3. *Sei V ein Vektorraum über K. Es gilt für alle $v \in V$:*

$$0 \cdot v = 0$$

$0 \in K$ $\qquad$ $0 \in V$

Beweis. Es gilt

$$0 \cdot v = \big(0+0\big) \cdot v = 0 \cdot v + 0 \cdot v$$

Also folgt:

$$0 = 0 \cdot v.$$

□

Lemma 9.4. *Sei V ein Vektorraum über K. Es ist $(-1) \cdot v = -v$ für alle $v \in V$*

Beweis. Es gilt:

$$0 = \big(1 + (-1)\big) \cdot v = v + (-1) \cdot v.$$

Da das Inverse eindeutig bestimmt ist, gilt $(-1) \cdot v = -v$. □

Definition 9.5. Sei V eine Vektorraum über K. $U \subseteq V$ heißt *Untervektorraum von V*, wenn U eine Untergruppe von V ist ($U \leq V$) und für alle $\lambda \in K$ und $v \in U$ gilt: $\lambda v \in U$.

Beispiel 9.6. • $\{0\} \subseteq V$ und V selbst sind Unterräume von V.

• Häufig: Für $m < n$ identifiziert man K^m als Unterraum von K^n vermöge

$$K^m \ni \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} \longmapsto \begin{pmatrix} x_1 \\ \vdots \\ x_m \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in K^n.$$

Satz 9.7. *Seien U, W Unterräume von V. Dann gilt: $U \cap W$ ist ein Unterraum von V.*

Beweis. Aus der Gruppentheorie wissen wir, dass gilt: $(U \cap W) \leq V$. Bleibt zu zeigen: Für $u \in U \cap W$ und $\lambda \in K$ ist $\lambda u \in U \cap W$.

Da U und W Unterräume sind, gilt für $u \in U$: $\lambda u \in U$. Ebenso für $u \in W$: $\lambda u \in W$. Somit ist $\lambda u \in U \cap W$. □

Satz 9.8. *Sei V ein Vektorraum über K und $S \subseteq V$ eine nicht-leere Teilmenge. Dann ist*

$$U_S \coloneqq \big\{\lambda_1 s_1 + \cdots + \lambda_k s_k \mid k \in \mathbb{N}, \lambda_i \in K, s_i \in S\big\}$$

ein Unterraum von V.

Terme der Form $\lambda_1 s_1 + \lambda_2 s_2 + \cdots$ heißen *Linearkombination* der s_i mit Koeffizienten λ_i. Die Menge U_S ist die Menge aller endlichen Linearkombinationen von Elementen von S.

Beweis. Seien $v, w \in U_S$. Dann lassen sich v und w schreiben als:

$$v = \lambda_1 s_1 + \cdots + \lambda_k s_k$$
$$w = \mu_1 t_1 + \cdots + \mu_l t_l$$

Da v und w endliche Linearkombinationen von Elementen von S sind, ist somit auch $v - w$ eine solche und somit Element von U_S.

Sei nun $\lambda \in K$. Dann ist

$$\lambda v = (\lambda\lambda_1)s_1 + \cdots + (\lambda\lambda_k)s_k,$$

und somit ebenfalls eine endliche Linearkombination von Elementen von S und daher ein Element von U_S. □

Definition 9.9. Der Raum U_S heißt der von S *erzeugte Unterraum* und wird mit $\operatorname{span}(S)$ bezeichnet. Die Menge S heißt in diesem Zusammenhang *Erzeugendensystem* von U_S.

Beispiel 9.10. Sei $V = \mathbb{R}^2$ und

$$S = \left\{ \begin{pmatrix} \lambda \\ 2\lambda \end{pmatrix} \middle| \, \lambda \in \mathbb{R} \right\}.$$

Dann ist S ein Unterraum von V und es gilt: $S = \operatorname{span}(S)$. (Der Leser möge sich davon selbst überzeugen.)

Beispiel 9.11. Sei $V = \mathbb{R}^2$ und

$$S = \left\{ \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 3 \\ 4 \end{pmatrix} \right\}.$$

Dann ist $\operatorname{span} S = \mathbb{R}^2$.

Offensichtlich ist $\operatorname{span} S \subseteq \mathbb{R}^2$; um $\mathbb{R}^2 \subseteq \operatorname{span} S$ zu zeigen, betrachten wir ein Element $x \in \mathbb{R}^2$, $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$.

Wir müssen nun zeigen, dass es $\lambda, \mu \in \mathbb{R}$ gibt, so dass gilt:

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \lambda \begin{pmatrix} 1 \\ 2 \end{pmatrix} + \mu \begin{pmatrix} 3 \\ 4 \end{pmatrix} = \begin{pmatrix} \lambda \\ 2\lambda \end{pmatrix} + \begin{pmatrix} 3\mu \\ 4\mu \end{pmatrix} = \begin{pmatrix} \lambda + 3\mu \\ 2\lambda + 4\mu \end{pmatrix}.$$

Wir müssen also das Gleichungssystem

$$x_1 = \lambda + 3\mu$$
$$x_2 = 2\lambda + 4\mu$$

lösen. Aus der ersten Gleichung erhalten wir $\lambda = x_1 - 3\mu$, welches wir in die zweite Gleichung einsetzen:

$$x_2 = 2(x_1 - 3\mu) + 4\mu = 2x_1 - 2\mu.$$

Also ist $\mu = \frac{2x_1 - x_2}{2}$ und damit

$$\lambda = x_1 - \frac{3}{2}(2x_1 - x_2) = -2x_1 + \frac{3}{2}x_2 = \frac{-4x_1 + 3x_2}{2}.$$

Mit den so bestimmten Skalaren $\lambda, \mu \in \mathbb{R}$ ist

$$x = \lambda \begin{pmatrix} 1 \\ 2 \end{pmatrix} + \mu \begin{pmatrix} 3 \\ 4 \end{pmatrix}.$$

Im Folgenden sei K stets ein Körper und V ein K-Vektorraum.

Definition 9.12. Eine Teilmenge $S \subseteq V$ heißt *linear unabhängig*, wenn gilt: Für beliebiege endliche Teilmengen $\{s_1, \dots, s_k\} \subseteq S$ besitzt die Gleichung

$$\lambda_1 s_1 + \cdots + \lambda_k s_k = 0 \qquad (\lambda_1, \dots, \lambda_k \in \mathbb{R})$$

nur die Lösung $\lambda_1 = \cdots = \lambda_k = 0$.

Beispiel 9.13. Betrachte $S = \{1, 2\} \subseteq \mathbb{R}^1$. Diese Menge ist nicht linear unabhängig, denn es gilt:

$$\underbrace{2}_{\neq 0} \cdot 1 + \underbrace{(-1)}_{\neq 0} \cdot 2 = 0.$$

Beispiel 9.14. Betrachte $S = \left\{ \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 3 \\ 6 \end{pmatrix} \right\}$. Diese Menge ist nicht linear unabhängig, denn es gilt:

$$3 \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix} + (-1) \cdot \begin{pmatrix} 3 \\ 6 \end{pmatrix} = 0.$$

Definition 9.15. Zu einem Vektorraum V sei die folgende Menge gegeben:

$$M := \left\{ n \in \mathbb{N} \,\middle|\, \begin{array}{l} \text{Es gibt eine } n\text{-elementige Teilmenge} \\ \text{linear unabhängiger Vektoren in } V \end{array} \right\}.$$

Besitzt M ein größtes Element, so heißt V *endlich-dimensionaler Vektorraum*, andernfalls *unendlich-dimensionaler Vektorraum*. Ist V endlich-dimensional, so gibt es also eine maximale Teilmenge linear unabhängiger Vektoren. Die Anzahl der Vektoren in dieser Teilmenge nennt man *Dimension* von V, symbolisch $\dim V$.

Beispiel 9.16. Im $\mathbb{R}^2$ ist die Menge

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

offenbar eine maximale Teilmenge linear unabhängiger Vektoren, denn für jeden Vektor $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2$ gilt:

$$1 \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + (-x_1) \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} + (-x_2) \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} = 0$$

Definition 9.17. Sei $B \subseteq V$ eine maximale Teilmenge linear unabhängiger Vektoren und gelte zudem $\operatorname{span} B = V$. Dann heißt B *Basis* von V.

Beispiel 9.18. Die Menge

$$B = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \right\} \subseteq \mathbb{R}^n$$

heißt *Standardbasis* des $\mathbb{R}^n$.

Oftmals spielt die Reihenfolge der Basisvektoren eine Rolle, z. B. wenn man etwas visualisieren möchte. Daher werden wir in Zukunft nur noch geordnete Basen betrachten, indem wir die Basisvektoren in einem Tupel zusammenfassen:

$$b_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, b_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ldots, b_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

und $B = (b_1, \ldots, b_n)$ ist dann die (geordnete) Standardbasis des $\mathbb{R}^n$.

Ein Vektorraum kann mehr als eine Basis besitzen, so ist z. B.

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

ebenfalls eine Basis für $\mathbb{R}^2$ (und eine andere als die Standardbasis)

Notation 9.19. Für die Standarbasis des $\mathbb{R}^n$ benutzen wir die Bezeichnungen

$$e_1 \coloneqq \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, e_2 \coloneqq \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \cdots, e_n \coloneqq \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}.$$

Satz 9.20. *Sei* $\dim V = n$ *und* $B \subseteq V$. *Dann sind äquivalent:*

1. *B ist eine Basis von V;*
2. *B ist eine maximale linear unabhängige Teilmenge von V, d. h. B ist linear unabhängig und für jedes* $v \in V \setminus B$ *ist* $B \cup \{v\}$ *linear abhängig;*
3. *B ist ein minimales Erzeugendensystem von V, d. h.* $\operatorname{span}(B) = V$ *und für jedes* $v \in B$ *gilt* $\operatorname{span}\big(B \setminus \{v\}\big) \neq V$.

Beweis. (1) $\Rightarrow$ (2): Sei B eine Basis von V, dann ist B eine linear unabhängige Teilmenge. Da $\operatorname{span}(B) = V$ gilt, ist $v \in V \setminus B$ auch Element von $\operatorname{span}(B)$. Insbesondere gibt es $\lambda_1, \ldots, \lambda_k \in K$ und $v_1, \ldots, v_k \in B$, so dass sich v schreiben lässt als

$$v = \sum_{i=1}^{k} \lambda_i v_i.$$

Also gilt:

$$\sum_{i=1}^{k} \lambda_i v_i + (-1)v = 0.$$

Also sind $v_1, \ldots, v_k, v$ linear abhängig, somit ist es auch $B \cup \{v\}$.

(2) $\Rightarrow$ (3): Sei B eine maximale lineare unabhängige Teilmenge von V. Sei ohne Beschränkung der Allgemeinheit $v \notin B$. (Warum dürfen

wir dies anehmen?) Nach Voraussetzung ist $B \cup \{v\}$ linear abhängig. Es gibt also $v_1, \dots, v_k \in B$ und $\lambda_1, \dots, \lambda_k, \lambda \in K$ so dass gilt:

$$\sum_{i=1}^{k} \lambda_i v_i + \lambda v = 0.$$

Insbesondere muss $\lambda \neq 0$ gelten. (Warum?) Also gilt

$$v = \sum_{i=1}^{k} \frac{-\lambda_i}{\lambda} v_i.$$

Also ist $v \in \operatorname{span} B$ und somit $V = \operatorname{span}(B)$. Es bleibt zu zeigen, dass B als Erzeugendensystem minimal ist, also dass für ein beliebiges $v \in B$ gilt: $\operatorname{span}(B \setminus \{v\} \neq V$. Angenommen es gilt $v \in \operatorname{span}(B \setminus \{v\})$. Dann gibt es $v_1, \dots, v_k \in B \setminus \{v\}$ und $\lambda_1, \dots, \lambda_k \in K$ mit

$$v = \sum_{i=1}^{k} \lambda_i v_i.$$

Daher erhalten wir

$$\sum_{i=1}^{k} \lambda_i v_i + (-1)v = 0.$$

Insbesondere sind die Vektoren $v_1, \dots, v_k, v$ linear abhängig. Dies ist aber im Widerspruch zur Voraussetzung, dass B eine linear unabhängige Teilmenge ist. Also ist B ein minimales Erzeugdensystem.

(3) $\Rightarrow$ (1): Sei B ein minimales Erzeugendensystem von V. Dann ist B insbesondere ein Erzeugendensystem. Es bleibt zu zeigen, dass B eine linear unabhängige Teilmenge ist. Angenommen es gäbe $\lambda_1 \dots, \lambda_k \in K$ und ein $j \in \{1, \dots, k\}$ mit $\lambda \not= 0$, so dass für $v_1, \dots, v_k \in B$ gilt:

$$\sum_{i=1}^{k} \lambda_i v_i = 0.$$

Dann erhalten wir:

$$v_j = \sum_{\substack{i=1 \\ i \neq j}}^{k} \frac{-\lambda_i}{\lambda_j} v_i,$$

d. h. v_j lässt sich als Linearkombination der anderen v_i darstellen. Dies steht im Widerspruch zur Minimalität von B. Also muss B eine linear unabhängige Teilmenge sein. □

Satz 9.20 sagt uns, dass jede Menge von n linear unabhängigen Vektoren eine Basis des $\mathbb{R}^n$ bildet. Gleichzeitig ist jede Menge von $m > n$ Vektoren aus $\mathbb{R}^n$ linear abhängig.

In den meisten Fällen werden wir mit der Standarbasis des $\mathbb{R}^n$ rechnen, manchmal ist es aber sinnvoll eine andere Basis zu benutzen.

Beispiel 9.21. Betrachte den Vektor

$$v = \begin{pmatrix} 3 \\ -2 \end{pmatrix} \in \mathbb{R}^2. \tag{9.1}$$

Bezüglich der Standardbasis lässt er sich schreiben als

$$v = 3 \cdot e_1 + (-2) \cdot e_2.$$

Allgemein lässt sich jeder Vektor

$$v = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} \in \mathbb{R}^n$$

bezüglich der Standardbasis schreiben als

$$v_1 e_1 + v_2 e_2 + \cdots v_n e_n.$$

Beispiel 9.22. Betrachete die Basis $B = (b_1, b_2) \subseteq \mathbb{R}^2$:

$$b_1 := \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \; b_2 := \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Gegeben seien die vier Vektoren $v_1, \ldots, v_4$ mit den Koordinaten

$$v_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, v_2 = \begin{pmatrix} 3 \\ -1 \end{pmatrix}, v_3 = \begin{pmatrix} 3 \\ -3 \end{pmatrix}, v_4 = \begin{pmatrix} 1 \\ -3 \end{pmatrix},$$

welche ein Quadrat in der Ebene aufspannen. Bezüglich der Standardbasis haben diese die Darstellungen

$$\begin{aligned} v_1 &= 1 \cdot e_1 + (-1) \cdot e_2 \\ v_2 &= 3 \cdot e_1 + (-1) \cdot e_2 \\ v_3 &= 3 \cdot e_1 + (-3) \cdot e_2 \\ v_4 &= 1 \cdot e_1 + (-3) \cdot e_2 \end{aligned}$$

Bezüglich der Basis B haben die v_i die Darstellung

$$\begin{aligned} v_1 &= 0 \cdot b_1 + 1 \cdot b_2 \\ v_2 &= 2 \cdot b_1 + 1 \cdot b_2 \\ v_3 &= 0 \cdot b_1 + 3 \cdot b_2 \\ v_4 &= (-2) \cdot b_1 + 3 \cdot b_2 \end{aligned}$$

9.2 *Winkel und Skalarprodukt*

Definition 9.23. Sei V ein Vektorraum über $\mathbb{R}$. Eine Abbildung

$$\langle \cdot, \cdot \rangle \colon V \times V \longrightarrow \mathbb{R}$$

heißt *Skalarprodukt*, falls für alle $u, v, w \in V$ und $\lambda, \mu \in \mathbb{R}$ gelten:

(SP1) $\langle v, v \rangle \geq 0$;

(SP2) $\langle v, v \rangle = 0$, genau dann wenn $v = 0$;

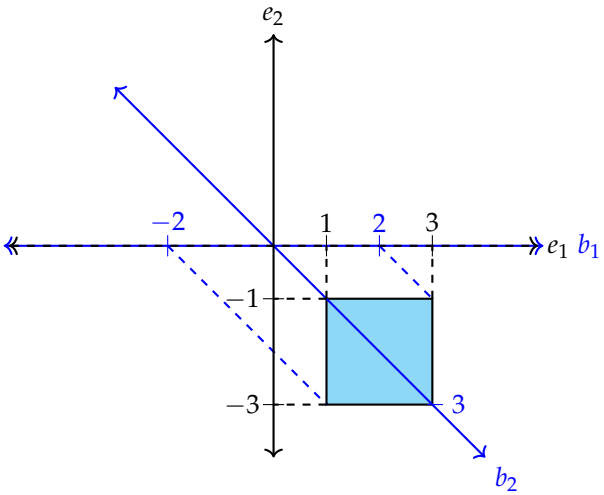


Abbildung 9.2: Koordinaten eines Vierecks bezügllich zweier verschiedener Basen.

(SP3) $\langle v, w\rangle = \langle w, v\rangle$;

(SP4) $\langle \lambda v, w\rangle = \langle v, \lambda w\rangle = \lambda\langle v, w\rangle$;

(SP5) $\langle u + v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$;

(SP6) $\langle u, v + w\rangle = \langle u, v\rangle + \langle u, w\rangle$;

Lemma 9.24. *Die Abbildung*

$$\begin{aligned}\langle\cdot,\cdot\rangle\colon \mathbb{R}^n \times \mathbb{R}^n &\longrightarrow \mathbb{R}\\ (v, w) &\longmapsto v_1w_1 + v_2w_2 + \cdots + v_nw_n\end{aligned}$$

ist ein Skalarprodukt Es heißt Standardskalarprodukt *auf* $\mathbb{R}^n$.

Beweis. Übung! □

Definition 9.25. Sei $V = \mathbb{R}^n$ und $\langle\cdot,\cdot\rangle$ das Standardskalarprodukt. Die *(Standard- oder* 2*-)Norm* eines Vektors $v \in \mathbb{R}^n$ ist gegeben durch

$$\|v\| := \sqrt{\langle v, v\rangle}.$$

Es gibt noch viele weitere Normen (auch auf $\mathbb{R}^n$). Die Norm $\|\cdot\|$ ohne Zusätze meint je nach Kontext entweder eine beliebige Norm, oder die Standardnorm. Die Standardnorm wird auch mit $\|\cdot\|_2$ bezeichnet.

Beispiel 9.26. Wir betrachten den Vektor $v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \in \mathbb{R}^2$:

Die Norm von v erhalten wir dann zu $\|v\| = \sqrt{v_1^2 + v_2^2}$. Nach dem Satz von Pythagoras entspricht dies der Länge des blauen Pfeils in Abbildung 9.3.

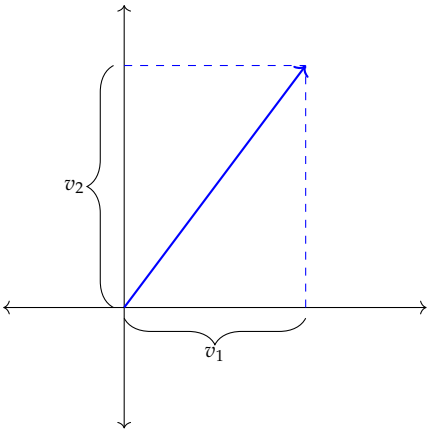


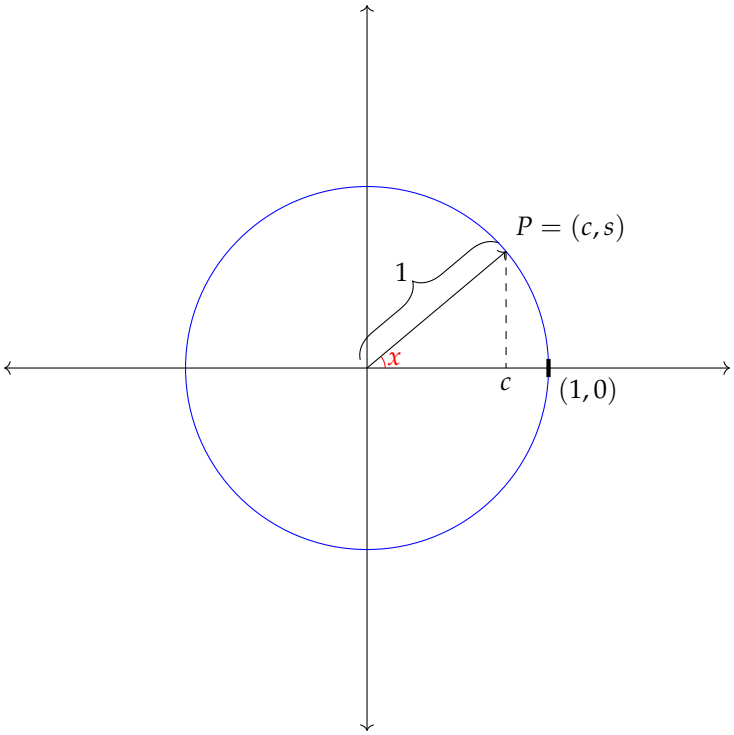
Abbildung 9.3: Zur Norm eines Vektors.

Dies war ein Beispiel für die Standardnorm. Es erschließt sich vielleicht warum diese auch mit $\|\cdot\|_2$ bezeichnet wird. Falls nicht, hier noch ein Hinweis: Die 3-Norm $\|\cdot\|_3$ ist im $\mathbb{R}^n$ wie folgt definiert:

$$\|v\|_3 = \sqrt[3]{|v_1|^3 + |v_2|^3 + \cdots + |v_n|^3}.$$

Erkennen Sie ein Muster? Was ist dann wohl die 4-Norm? Oder die π-Norm? Können Sie auch die 1-Norm oder die ∞-Norm definieren?

Definition 9.27. Sei x die Länge des Bogenstücks am Einheitskreis, gemessen vom Punkte $(1,0)$ beginnend, entgegen dem Uhrzeigersinn und $P = (c,s)$ der zugehörige Punkt. Dann heißt die erste Koordinate von P *Kosinus* von x, bzw. $\cos(x)$.

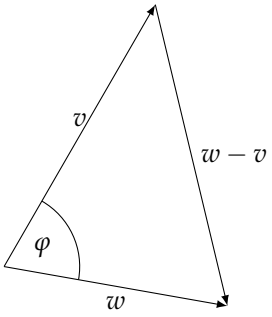


Satz 9.28. *Seien $v, w \in \mathbb{R}^n$, $v \neq 0$, $w \neq 0$ und $\varphi \in [0, 2\pi)$ der Winkel zwischen v und w. Dann gilt:*

$$\cos(\varphi) = \frac{\langle v, w \rangle}{\|v\|\|w\|}.$$

Beweis. Ohne Beschränkung der Allgemeinheit seien v und w linear unabhängig. Dann gilt $\operatorname{span}\{v, w\}$ ist ein zweidimensionaler Unterraum des $\mathbb{R}^n$. Insbesondere liegen v und w in einer Ebene und wir können Schulgeometrie verwenden. Nach dem Kosinussatz gilt:

Warum können wir dies Annehmen?



$$\|w - v\|^2 = \|v\|^2 + \|w\|^2 - 2\|v\|\|w\|\cos(\varphi)$$

Dies formen wir mittels Skalarprodukt um:

$$\langle w-v, w-v\rangle = \langle v, v\rangle + \langle w, w\rangle - 2\|v\|\|w\|\cos(\varphi)$$

Nach der Definition des Skalarprodukts gilt:

$$\begin{aligned}\langle w-v, w-v\rangle &= \langle w, w-v\rangle - \langle v, w-v\rangle \\ &= \langle w, w\rangle - \langle w, v\rangle - \langle v, w\rangle + \langle v, v\rangle \\ &= \langle w, w\rangle - \langle w, v\rangle - \langle w, v\rangle + \langle v, v\rangle \\ &= \langle w, w\rangle - 2\langle w, v\rangle + \langle v, v\rangle.\end{aligned}$$

Wir erhalten dann:

$$\langle w, w\rangle - 2\langle w, v\rangle + \langle v, v\rangle = \langle v, v\rangle + \langle w, w\rangle - 2\|v\|\|w\|\cos(\varphi)$$

Kürzen liefert:

$$\langle w, v\rangle = \|v\|\|w\|\cos(\varphi),$$

woraus sich die Behauptung ergibt. □

Definition 9.29. Sei V ein Vektorraum über $\mathbb{R}$. Eine Menge von Vektoren $v_1, \ldots, v_k \in V$ heißt *orthogonales System*, falls für alle $1 \leq i, j \leq k$, $i \neq j$ gilt:

$$\langle v_i, v_j\rangle = 0.$$

Ein orthogonales System $v_1, \ldots, v_k \in V$, dass zusätzlich eine Basis von V ist und für alle v_i die Bedingung $\|v_i\| = 1$ erfüllt, heißt *Orthonormalbasis* von V.

Satz 9.30. *Sei V ein Vektorraum über $\mathbb{R}$. Bilden $v_1, \ldots, v_k \in V$ ein orthogonales System, so sind $v_1, \ldots, v_k$ linear unabhängig.*

Beweis. Seien $\lambda_1, \ldots, \lambda_k \in \mathbb{R}$ so gewählt, dass

$$\lambda_1 v_1 + \cdots + \lambda_k v_k = 0$$

gilt. Sei $1 \leq i \leq k$ beliebig. Dann gilt:

$$\begin{aligned}\langle \lambda_1 v_1 + \cdots + \lambda_k v_k, v_i\rangle &= \langle 0, v_i\rangle \\ &= 0.\end{aligned}$$

Nach der Defintion des Skalarprodukts gilt:

$$0 = \lambda_1\langle v_1, v_i\rangle + \cdots + \lambda_i\langle v_i, v_i\rangle + \cdots + \lambda_k\langle v_k, v_i\rangle.$$

Da $v_1, \ldots, v_k$ ein orthogonales System bilden, ist für jedes $j \neq i$: $\langle v_j, v_i\rangle = 0$ und es bleibt in der Summe nur ein einziger Term übrig:

$$0 = \lambda_i\langle v_i, v_i\rangle = \lambda_i\|v_i\|^2.$$

Da $v_i \neq 0$, gilt auch $\|v_i\|^2 \neq 0$. Also muss $\lambda_i = 0$ gelten. Da wir i beliebig gewählt haben, erhalten wir:

$$\lambda_1 = \cdots = \lambda_k = 0.$$ □

10

Lineare Gleichungssysteme

Wenn wir eine Familie von Vektoren auf lineare Unabhängigkeit überprüfen wollen, müssen wir lineare Gleichungssysteme lösen können. Für wenige Gleichungen mit wenigen Unbekannten ist dies noch leicht durch Einsetzen machbar, bei größeren Systemen brauchen wir etwas besseres.

Definition 10.1. Ein *lineares Gleichungssystem (LGS)* aus m Gleichungen und n Unbekannten $x_1, \ldots, x_n \in K$ hat die Form

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

Die a_{ij} heißen Koeffizienten des LGS. Die b_i sind ebenfalls Elemente von K. Sind alle b_i gleich Null, so heißt das LGS *homogen*, ansonsten *inhomogen*.

Satz 10.2. *Ein inhomogenes lineares Gleichungssystem über $K = \mathbb{R}$ hat keine, genau eine, oder unendlich viele Lösungen.*

Beispiel 10.3. a)

$$\begin{aligned} x + y &= 2 \\ x - y &= 0 \end{aligned}$$

Aus der zweiten Gleichung folgt $x = y$ und durch einsetzen in die erste Gleichung erhalten wir:

$$2x = 2.$$

Somit ist $x = 1$ und daraus folgt $y = 1$. Es gibt also genau eine Lösung.

b)

$$\begin{aligned} x + y &= 2 \\ x + y &= 0 \end{aligned}$$

Aus der zweiten Gleichung folgt $x = -y$ und durch einsetzen in die erste Gleichung erhalten wir $0 = 2$, was eine falsche Aussage ist. Das Gleichungssystem hat also keine Lösung.

c)

$$\begin{aligned} x + y &= 2 \\ 2x + 2y &= 4 \end{aligned}$$

Aus der ersten Gleichung erhalten wir $y = 2 - x$ und durch einsetzen in die zweite Gleichung folgt

$$2x + 2(2 - x) = 2x + 4 - 2x = 4.$$

Wir erhalten also die Gleichung $4 = 4$, welche für alle Werte von x, y erfüllt ist. Das System hat also unendlich viele Lösungen, wir können x oder y beliebig wählen. Daher sind durch

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} t \\ 2 - t \end{pmatrix} \quad \text{mit } t \in \mathbb{R}$$

alle Lösungen des Systems beschrieben.

Wir können uns dies wie folgt in $\mathbb{R}^2$ veranschaulichen:

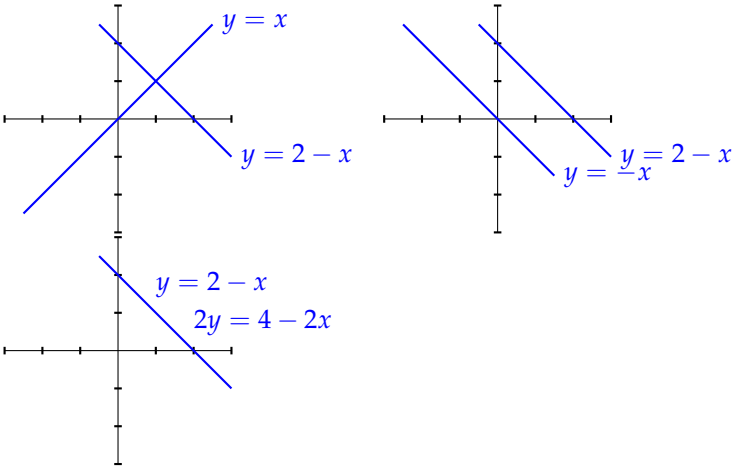


Abbildung 10.1: Visualisierung verschiedener Gleichungssysteme. Jede lineare Gleichung definiert einen Unterraum.

Bemerkung 10.4. Ein homogenes lineares Gleichungssystem hat immer mindestens eine Lösung, nämlich $x_1 = \cdots = x_n = 0$. Dies ist die *triviale Lösung*.

Satz 10.5. *Die folgenden Umformungen verändern nicht die Lösung eines linearen Gleichungssystems*

i) Vertauschen zweier Gleichungen;

ii) Multiplikation einer Gleichung mit $r \neq 0$.

iii) Addition des r-fachen einer Gleichung zu einer anderen Gleichung.

Bemerkung 10.6. Statt ein lineares Gleichungssystem in Gleichungen darzustellen, nutzt man oft die Darstellung als *erweiterte Koeffizientenmatrix*:

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & & & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right)$$

Definition 10.7. Eine erweiterte Koeffizientenmatrix ist in *Zeilenstufenform*, falls gelten:

i) In jeder Zeile ist die erste Zahl ungleich Null eine 1 und befindet sich rechts von der ersten nicht Null-Zahl der Zeile darüber.

ii) Alle Nullzeilen befinden sich am unteren Ende der Matrix

Die erweiterte Koeffizientenmatrix ist in *reduzierter Zeilenstufenform*, falls zusätzlich gilt

iii) über allen führenden Einsen stehen Nullen.

Ist eine erweiterte Koeffizientenmatrix (und somit das lineare Gleichungssystem) in Zeilenstufenform oder in reduzierter Zeilenstufenform, so kann man die Lösung des Systems leicht ablesen.

Satz 10.8. *Jedes lineare Gleichungssystem lässt sich durch endlich viele Zeilenumformungen in Zeilenstufenform bringen.*

Diese Umformungen kann man mittels des folgenden Algorithmus, welcher nach Carl Friedrich Gauss und dem deutschen Geodäten Wilhelm Jordan (*1842, †1899) benannt ist, vollziehen.

Algorithmus 10.9 (Gauß-Jordan-Algorithmus)**.** Der folgende Algorithmus liefert das Gewünschte.

```
i = j = 1
Gauss(i,j):
        WENN i=m oder j=n+1:
                return
        WENN a_ij = 0:
                Suche r>i mit a_rj ≠ 0
                WENN r existiert:
                        Tausche Zeilen r und i
                SONST:
                        Gauss(i,j+1)
        Teile i-te Zeile durch a_ij
        Fuer alle k>i:
                (Zeile k) - a_kj * (Zeile i)
        Gauss(i+1,j+1)
```

Bemerkung 10.10. Beim Lösen von Hand ist scharfes Hinsehen und geschicktes Umformen oft einfacher, als wortgetreues Umsetzen des Gauß-Jordan-Algorithus. Bei der Implementierung als Computerprogramm treten oft fehlerhafte Ergebnisse auf, die mit der Repräsentation reeller Zahlen auf dem Rechner zusammenhängen.

Beispiel 10.11. Sei $K = \mathbb{Z}_5$. Wir wollen das lineare Gleichungssystem mit der folgenden erweiterten Koeffizientenmatrix lösen.

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

$$\begin{array}{ll}
\begin{pmatrix} 2 & 3 & 4 & 0 \\ 3 & 2 & 2 & 3 \end{pmatrix} & \begin{array}{l} \leftarrow \\ \leftarrow \end{array} \\
\begin{pmatrix} 3 & 2 & 2 & 3 \\ 2 & 3 & 4 & 0 \end{pmatrix} & \begin{array}{l} \mid \cdot 2 \\ \mid \cdot 3 \end{array} \\
\begin{pmatrix} 1 & 4 & 4 & 1 \\ 1 & 4 & 2 & 0 \end{pmatrix} & \begin{array}{l} \mid \cdot 4 \\ \leftarrow \end{array} \; + \\
\begin{pmatrix} 1 & 4 & 4 & 1 \\ 0 & 0 & 3 & 4 \end{pmatrix} & \mid \cdot 2 \\
\begin{pmatrix} 1 & 4 & 4 & 1 \\ 0 & 0 & 1 & 3 \end{pmatrix} &
\end{array}$$

Wir können direkt aus der letzten Zeile ablesen: $x_3 = 3$. In die vorletzte Zeile setzen wir dies ein und lesen ab:

$$x_1 + 4x_2 + 2 = 1$$

Wir erhalten durch Umstellen:

$$4x_2 = 4 + 4x_1 \quad \Leftrightarrow \quad x_2 = 1 + x_1$$

Wir haben also die folgenden Lösungen:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \left\{ \begin{pmatrix} 0 \\ 1 \\ 3 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 2 \\ 3 \\ 3 \end{pmatrix}, \begin{pmatrix} 3 \\ 4 \\ 3 \end{pmatrix}, \begin{pmatrix} 4 \\ 0 \\ 3 \end{pmatrix} \right\}$$

11

Matrizen

11.1 Grundbegriffe und Eigenschaften

Definition 11.1. Ein rechteckiges Schema von Elementen a_{ij} aus einem Körper K in m Zeilen und n Spalten der Form

$$A = (a_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

heißt $m \times n$-*Matrix*. (m, n) heißt *Dimension* der Matrix, die a_{ij} *Koeffizienten*. Man schreibt $A \in K^{m \times n}$. Ist A eine quadratische Matrix (also eine $n \times n$-Matrix), so heißen die Elemente a_{ii} *Diagonalelemente* von A. Ist die Dimension klar, schreibt man kurz $A = (a_{ij})$ statt $A = (a_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}}$.

Mit Matrizen kann man rechnen. Seien $A = (a_{ij})$ und $B = (b_{ij})$ Matrizen gleicher Dimension. Dann kann man die beiden Matrizen addieren:

$$A + B = (a_{ij} + b_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}} = \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \cdots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \cdots & a_{2n} + b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & \cdots & a_{mn} + b_{mn} \end{pmatrix}$$

Lemma 11.2. *Bezüglich der Matrizenaddition und der Skalarmultiplikation*

$$kA = k(a_{ij}) = (ka_{ij}) \qquad \text{für } k \in K$$

ist $K^{m \times n}$ ein Vektorraum.

Beweis. Übung! □

Definition 11.3. Zu einer $m \times n$-Matrix $A = (a_{ij})$ heißt die $n \times m$ Matrix $A^t = (a_{ji})$ die *transponierte Matrix* zu A. Ihre Spalten sind gleich der Zeilen von A und ihre Zeilen sind die Spalten von A.

Beispiel 11.4.

$$A = \begin{pmatrix} 3 & 2 \\ 5 & -1 \\ 0 & 7 \end{pmatrix}, \qquad A^t = \begin{pmatrix} 3 & 5 & 0 \\ 2 & -1 & 7 \end{pmatrix}$$

$$B = \begin{pmatrix} 1 \\ 2 \\ 5 \end{pmatrix}, \qquad B^t = \begin{pmatrix} 1 & 2 & 5 \end{pmatrix}$$

Definition 11.5. Ist das transponierte einer $n \times n$-Matrix A wieder gleich A, so heißt A *symmetrische Matrix*.

Definition 11.6. Seien $A \in K^{m \times n}$ und $B \in K^{n \times r}$. Das *Produkt* $A \cdot B$ ist die Matrix $C \in K^{m \times r}$ mit den Koeffizienten

$$C = (c_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,r}} = \sum_{k=1}^{n} a_{ik} b_{kj}.$$

Das Produkt $A \cdot B$ ist nur definiert, falls die Anzahl der Spalten von A gleich der Anzahl der Zeilen von B ist.

Beispiel 11.7.

$$A = \begin{pmatrix} 4 & 2 & 0 \\ -1 & 3 & 5 \end{pmatrix}, \qquad B = \begin{pmatrix} 2 & 1 \\ 3 & 7 \\ 1 & 0 \end{pmatrix}, \qquad C = \begin{pmatrix} 7 & 1 \\ 2 & 5 \end{pmatrix}$$

$$AB = \begin{pmatrix} 14 & 18 \\ 12 & 20 \end{pmatrix}, \qquad BA = \begin{pmatrix} 7 & 7 & 5 \\ 5 & 27 & 35 \\ 4 & 2 & 0 \end{pmatrix}, \qquad BC = \begin{pmatrix} 16 & 7 \\ 35 & 38 \\ 7 & 1 \end{pmatrix}.$$

Die Produkte AC und CB sind nicht definiert.

Lemma 11.8. *Für die Multiplikation mit Matrizen gelten folgende Regeln: Seien A, B, C Matrizen passender Dimension und $k \in K$. Dann gelten:*

1. $(kA)B = k(AB)$
2. $A(BC) = (AB)C$
3. $(A + B)C = AC + BC$
4. $A(B + C) = AB + AC$
5. $(AB)^t = B^t A^t$

Im Allgemeinen gilt: $AB \neq BA$.

Definition 11.9. Die Matrix

$$\mathbb{1}_n = \begin{pmatrix} 1 & 0 & 0 & \cdots \\ 0 & 1 & 0 & \cdots \\ 0 & 0 & 1 & \\ \vdots & \vdots & & \ddots \end{pmatrix}$$

heißt $n \times n$-Einheitsmatrix.

Lemma 11.10. *Es gilt für $A \in K^{m\times n}$:*

$$\mathbb{1}_m \cdot A = A \cdot \mathbb{1}_n = A.$$

Lemma 11.11. *Bezüglich der Addition und Multiplikation von Matrizen bildet $K^{n\times n}$ einen Ring mit Eins.*

Lineare Gleichungssysteme kann man als Matrixgleichung schreiben. Statt

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

schreiben wir

$$Ax = b, \qquad \text{mit } A \in K^{m\times n},\ x \in K^n,\ b \in K^m.$$

Definition 11.12. Gibt es zu $A \in K^{n\times n}$ eine Matrix $A^{-1} \in K^{n\times n}$ mit

$$AA^{-1} = A^{-1}A = \mathbb{1}_n,$$

so heißt A *invertierbar* und A^{-1} die zu A *inverse Matrix*.

Lemma 11.13. *Ist ein lineares Gleichungssystem mit n Gleichungen und n Unbekannten der Form*

$$Ax = b$$

gegeben und A ist invertierbar, so ist

$$x = A^{-1}b$$

die einzige Lösung des linearen Gleichungssystems.

Lemma 11.14. *Sind $A, B \in K^{n\times n}$ invertierbar, so ist auch AB invertierbar und es gilt*

$$(AB)^{-1} = B^{-1}A^{-1}.$$

Beweis. Übung. □

11.2 Determinanten

Beispiel 11.15. Ist

$$A = \begin{pmatrix} 2 & 4 \\ -1 & 3 \end{pmatrix}$$

invertierbar? Wir suchen eine Matrix

$$B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$$

für die gilt

$$AB = \begin{pmatrix} 2 & 4 \\ -1 & 3 \end{pmatrix}\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} 2b_{11}+4b_{21} & 2b_{12}+4b_{22} \\ -b_{11}+3b_{21} & -b_{12}+3b_{22} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Wir müssen also ein lineares Gleichungssystem mit vier Gleichungen und vier Unbekannten lösen:

$$\begin{aligned}
2b_{11} \qquad\qquad &+4b_{21} &&= 1\\
-b_{11} \qquad\qquad &+3b_{21} &&= 0\\
2b_{12} \qquad\qquad &&4b_{22} &= 0\\
-b_{12} \qquad\qquad &&+3b_{22} &= 1
\end{aligned}$$

oder in Matrixschreibweise:

$$\begin{pmatrix} 2 & 0 & 4 & 0\\ -1 & 0 & 3 & 0\\ 0 & 2 & 0 & 4\\ 0 & -1 & 0 & 3 \end{pmatrix} \begin{pmatrix} b_{11}\\ b_{12}\\ b_{21}\\ b_{22} \end{pmatrix} = \begin{pmatrix} 1\\ 0\\ 0\\ 1 \end{pmatrix}$$

oder als zwei Gleichungssysteme mit der selben Koeffizientenmatrix:

$$\begin{pmatrix} 2 & 4\\ -1 & 3 \end{pmatrix} \begin{pmatrix} b_{11}\\ b_{21} \end{pmatrix} = \begin{pmatrix} 1\\ 0 \end{pmatrix} \qquad \text{und} \qquad \begin{pmatrix} 2 & 4\\ -1 & 3 \end{pmatrix} \begin{pmatrix} b_{12}\\ b_{22} \end{pmatrix} = \begin{pmatrix} 0\\ 1 \end{pmatrix}.$$

Wir können diese beiden Gleichungssysteme simultan mit dem Gauß-Jordan-Algorithmus lösen. Wir schreiben die Gleichungssysteme als erweiterte Koeffizientenmatrix

$$\left(\begin{array}{cc|cc} 2 & 4 & 1 & 0\\ -1 & 3 & 0 & 1 \end{array}\right) = A|\mathbb{1}_2$$

Mittels der erlaubten Äquivalenzumformungen versuchen wir nun die erweiterte Koeffizientenmatrix auf die Form

$$\mathbb{1}_2|B$$

zu bringen. Ist dies möglich, so ist B die Inverse zu A, ansonsten gibt es keine Inverse.

$$\begin{pmatrix} 2 & 4 & 1 & 0\\ -1 & 3 & 0 & 1 \end{pmatrix} \ | \div 2$$

$$\begin{pmatrix} 1 & 2 & \frac{1}{2} & 0\\ -1 & 3 & 0 & 1 \end{pmatrix} \ \begin{matrix}\urcorner\\ \leftarrow\!\lrcorner^{+}\end{matrix}$$

$$\begin{pmatrix} 1 & 2 & \frac{1}{2} & 0\\ 0 & 5 & \frac{1}{2} & 1 \end{pmatrix} \ | \div 5$$

$$\begin{pmatrix} 1 & 2 & \frac{1}{2} & 0\\ 0 & 1 & \frac{1}{10} & \frac{1}{5} \end{pmatrix} \ \begin{matrix}\leftarrow\!\!-\!\!-\!\!-\urcorner^{+}\\ |\cdot(-2)\ \lrcorner\end{matrix}$$

$$\begin{pmatrix} 1 & 0 & \frac{3}{10} & -\frac{2}{5}\\ 0 & 1 & \frac{1}{10} & \frac{1}{5} \end{pmatrix}$$

Also

$$A^{-1} = \begin{pmatrix} \frac{3}{10} & -\frac{2}{5}\\ \frac{1}{10} & \frac{1}{5} \end{pmatrix} = \frac{1}{10} \begin{pmatrix} 3 & -4\\ 1 & 2 \end{pmatrix}.$$

Es stellt sich nun die Frage, ob wir einer Matrix „ansehen“ können, ob sie invertierbar ist.

Wir betrachten dazu ein allgemeines 2×2 lineares Gleichungssystem:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= b_1 \\ a_{21}x_1 + a_{22}x_2 &= b_2 \end{aligned}$$

Wir multiplizieren die erste Gleichung mit a_{22} und die zweite mit $-a_{12}$ und erhalten

$$\begin{aligned} a_{11}a_{22}x_1 + a_{12}a_{22}x_2 &= a_{22}b_1 \\ -a_{12}a_{21}x_1 - a_{12}a_{22}x_2 &= -a_{12}b_2 \end{aligned}$$

Wir addieren beide Gleichungen und erhalten

$$a_{11}a_{22}x_1 - a_{12}a_{21}x_1 = (a_{11}a_{22} - a_{12}a_{21})x_1 = a_{22}b_1 - a_{12}b_2$$

Analog erhalten wir:

$$(a_{11}a_{22} - a_{12}a_{21})x_2 = -a_{21}b_1 + a_{11}b_2$$

Genau dann, wenn $(a_{11}a_{22} - a_{12}a_{21}) \neq 0$ ist, ist das System eindeutig lösbar und die Matrix A invertierbar.

Definition 11.16. Für eine 2×2-Matrix

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

heißt

$$\det(A) = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{21}a_{12}$$

die *Determinante* von A.

Die Verallgemeinerung der Determinante auf beliebige quadratische Matrizen ist schwieriger, die Herleitung würde den Rahmen dieser Notizen sprengen. Wir begnügen uns daher mit der folgenden Definition:

Definition 11.17 (Entwicklungssatz von Laplace)**.** Für $A \in K^{n \times n}$ gilt:

Entwicklung nach der i-ten Zeile:

$$det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij})$$

Entwicklung nach der j-ten Spalte:

$$det(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij})$$

Dabei ist A_{ij} jeweils die $(n-1) \times (n-1)$-Matrix, welche aus A hervorgeht indem man die i-te Zeile und die j-te Spalte streicht.

Der Entwicklungssatz von Laplace ist nach dem französischen Mathematiker PIERRE-SIMON LAPLACE (*1749, †1827) benannt.

Beispiel 11.18.

$$A = \begin{pmatrix} 4 & 2 & 4 & 4 \\ 5 & 6 & 1 & 4 \\ 0 & 0 & 2 & 0 \\ 5 & 5 & 3 & 6 \end{pmatrix} \in \mathbb{Z}_7^{4\times 4}$$

Wir wollen möglichst wenig rechnen und sehen, das in der dritten Zeile viele Nullen stehen. Wir entwickeln also nach der dritten Zeile:

$$\begin{aligned} \det(A) &= (-1)^{3+1} \cdot 0 \cdot A_{31} && +(-1)^{3+2} \cdot 0 \cdot A_{32} \\ &&& +(-1)^{3+3} \cdot 2 \cdot A_{33} \\ &&& +(-1)^{3+4} \cdot 0 \cdot A_{34} \\ &= (-1)^{3+3} \cdot 2 \cdot A_{33} \\ &= (-1)^6 \cdot 2 \cdot \det \begin{pmatrix} 4 & 2 & 4 \\ 5 & 6 & 4 \\ 5 & 5 & 6 \end{pmatrix} \\ &= 2 \cdot \left((-1)^{1+1} \cdot 4 \cdot \begin{vmatrix} 6 & 4 \\ 5 & 6 \end{vmatrix} \right. && +(-1)^{1+2} \cdot 2 \cdot \begin{vmatrix} 5 & 4 \\ 5 & 6 \end{vmatrix} \\ &&& \left. +(-1)^{1+3} \cdot 4 \cdot \begin{vmatrix} 5 & 6 \\ 5 & 5 \end{vmatrix} \right) \\ &= 2 \cdot \big(4 \cdot 2 - 2 \cdot 3 + 4 \cdot 2\big) \\ &= 2 \cdot 3 \\ &= 6 \end{aligned}$$

Satz 11.19. *Sind $A, B \in K^{n\times n}$, dann gelten:*

1. $\det(AB) = \det(A)\det(B)$
2. $\det(A) = \det(A^t)$
3. *A ist genau dann invertierbar, wenn $\det(A) \neq 0$ ist.*
4. $\big(\det(A)\big)^{-1} = \det(A^{-1})$, *falls A invertierbar ist.*
5. *Sind die Zeilen von A linear abhängig, so gilt $\det(A) = 0$.*

Korollar 11.20. *Sei $A \in K^{n\times n}$.*

1. $\det(\mathbb{1}) = 1$.
2. *Ist A' die Matrix, die entsteht, wenn man eine Zeile von A mit $\lambda \in K$ multipliziert, so gilt:* $\det(A') = \lambda \det(A)$.
3. *Ist A' die Matrix, die entsteht, wenn man zwei Zeilen von A tauscht, so gilt:* $\det(A') = -\det(A)$.
4. *Ist A' die Matrix, die entsteht, wenn man zu einer Zeile von A das λ-fache einer anderen Zeile addiert, so gilt:* $\det(A') = \det(A)$.

Beweis. 1. Wir können die erste Aussage beweisen, ohne die Laplace-Entwicklung zu nutzen: Es gilt: $\det(\mathbb{1}) = \det(\mathbb{1} \cdot \mathbb{1}) = \det(\mathbb{1}) \cdot \det(\mathbb{1})$. Also muss $\det(\mathbb{1}) = 1$ oder $\det(\mathbb{1}) = 0$ gelten. Die Einheitsmatrix ist aber invertierbar, daher muss $\det(\mathbb{1}) = 1$ gelten.

2. Multiplikation der i-ten Zeile von A mit $\lambda \in K$ entspricht der Multiplikation von links mit der Matrix

$$B = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & \lambda & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix},$$

wobei λ in der i-ten Zeile und Spalte steht. Es gilt also: $A' = BA$. Offensichtlich gilt $\det(B) = \lambda$. Nach dem vorherigen Satz folgt: $\det(A') = \det(BA) = \det(B)\det(A) = \lambda \det(A)$.

3. Übung!

4. Addition des λ-fachen der k-ten Zeile zur i-ten Zeile entspricht der Multiplikation von links mit der Matrix

$$B = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & \lambda & & \\ & & & \ddots & & & \\ & & & & \ddots & & \\ & & & & & 1 & \end{pmatrix},$$

wobei λ in der i-ten Zeile und k-ten Spalte steht. Es gilt also: $A' = BA$. Offensichtlich gilt $\det(B) = 1$. Nach dem vorherigen Satz folgt: $\det(A') = \det(BA) = \det(B)\det(A) = \det(A)$.

□

Dies erlaubt es, die Determinante einer Matrix auf einfachere Weise auszurechnen. Man formt mittels Äquivalenzumformungen die Matrix so um, dass Sie in Zeilenstufenform ist. Dann kann man die Determinante einfach als Produkt der Diagonalelemente ablesen. Dabei ist zu beachten, dass sich beim Tauschen zweier Zeilen das Vorzeichen der Determinante ändert. Hat man also insgesamt k-mal je zwei Zeilen vertauscht, muss man die erhaltene Determinante mit $(-1)^k$ multiplizieren.

Enthält eine Matrix eine Nullzeile oder Nullspalte, so ist die Determinante 0.

12

Lineare Abbildungen

12.1 *Definitionen und Grundbegriffe*

Definition 12.1. Sei K ein Körper und seien V, W Vektorräume über K. Eine Abbildung

$$\varphi\colon V \longrightarrow W$$

heißt *linear*, falls sie die folgenden beiden Bedingungen erfüllt:

(A1) Für alle $v, v' \in V$ gilt:

$$\varphi(v + v') = \varphi(v) + \varphi(v');$$

(A2) Für alle $v \in V$ und $\lambda \in K$ gilt:

$$\varphi(\lambda v) = \lambda\varphi(v).$$

Beispiel 12.2.

$$\begin{aligned} \varphi\colon \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\ \begin{pmatrix} x \\ y \end{pmatrix} &\longmapsto \begin{pmatrix} 2x + y \\ 3x \end{pmatrix} \end{aligned}$$

ist eine lineare Abbildung, denn

$$\begin{aligned} \varphi\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} + \varphi\begin{pmatrix} x_2 \\ y_2 \end{pmatrix} &= \begin{pmatrix} 2x_1 + y_1 \\ 3x_1 \end{pmatrix} + \begin{pmatrix} 2x_2 + y_2 \\ 3x_2 \end{pmatrix} \\ &= \begin{pmatrix} 2x_1 + y_1 + 2x_2 + y_2 \\ 3x_1 + 3x_2 \end{pmatrix} \\ &= \begin{pmatrix} 2(x_1 + x_2) + (y_1 + y_2) \\ 3(x_1 + x_2) \end{pmatrix} \\ &= \varphi\begin{pmatrix} x_1 + x_2 \\ y_1 + y_2 \end{pmatrix} \\ &= \varphi\left(\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} + \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}\right) \end{aligned}$$

und

$$\begin{aligned}\varphi\left(\lambda\begin{pmatrix}x\\y\end{pmatrix}\right) &= \varphi\begin{pmatrix}\lambda x\\ \lambda y\end{pmatrix}\\ &= \begin{pmatrix}2\lambda x+\lambda y\\ 3\lambda x\end{pmatrix}\\ &= \begin{pmatrix}\lambda(2x+y)\\ \lambda(3x)\end{pmatrix}\\ &= \lambda\begin{pmatrix}2x+y\\ 3x\end{pmatrix}\\ &= \lambda\varphi\begin{pmatrix}x\\y\end{pmatrix}\end{aligned}$$

$$\begin{aligned}\psi\colon \mathbb{R}^2 &\longrightarrow \mathbb{R}^2\\ \begin{pmatrix}x\\y\end{pmatrix} &\longmapsto \begin{pmatrix}x+3\\2y\end{pmatrix}\end{aligned}$$

ist keine lineare Abbildung, denn z. B.:

$$\psi\left(2\cdot\begin{pmatrix}1\\0\end{pmatrix}\right) = \psi\begin{pmatrix}2\\0\end{pmatrix} = \begin{pmatrix}5\\0\end{pmatrix}$$

aber

$$2\cdot\psi\begin{pmatrix}1\\0\end{pmatrix} = 2\cdot\begin{pmatrix}4\\2\end{pmatrix} = \begin{pmatrix}8\\4\end{pmatrix} \neq \begin{pmatrix}5\\0\end{pmatrix}$$

Satz 12.3. *Sei K ein Körper und seien $V = K^n, W = K^m$ jeweils mit der Standardbasis versehen. Eine Abbildung $\varphi\colon V \longrightarrow W$ ist genau dann linear, wenn sie in der Form $\varphi(x) = Ax$ mit $A \in K^{m\times n}$ geschrieben werden kann. Das heißt*

$$\varphi\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix} = \begin{pmatrix}a_{11}x_1+\cdots+a_{1n}x_n\\ \vdots\\ a_{m1}x_+\cdots+a_{mn}x_n\end{pmatrix}$$

Die Matrix A heißt darstellende Matrix *von φ und ist eindeutig bestimmt. Die Spalten von A sind die Bilder der Standardbasisvektoren $e_1,\ldots,e_n$:*

$$A = \big(\varphi(e_1),\ldots,\varphi(e_n)\big)$$

Beweis. Es gilt für $x, y \in V$:

$$\varphi(x+y) = A\cdot(x+y) = Ax+Ay = \varphi(x)+\varphi(y)$$

$$\varphi(\lambda x) = A(\lambda x) = \lambda(Ax) = \lambda\varphi(x)$$

Außerdem gilt $x = x_1e_1+\cdots+x_ne_n$ und daher

$$\begin{aligned}\varphi(x) &= \varphi(x_1e_1+\cdots+x_ne_n)\\ &= x_1\varphi(e_1)+\cdots+x_n\varphi(e_n)\\ &= \big(\varphi(e_1),\ldots,\varphi(e_n)\big)x\\ &= Ax\end{aligned}$$

□

Beispiel 12.4.

$$\varphi\colon \mathbb{R}^2 \longrightarrow \mathbb{R}^2$$
$$\begin{pmatrix} x \\ y \end{pmatrix} \longmapsto \begin{pmatrix} 2x+y \\ 3x \end{pmatrix}$$

Wir setzen die Standardbasisvektoren ein und erhalten:

$$\varphi\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) = \begin{pmatrix} 2 \\ 3 \end{pmatrix}, \qquad \varphi\left(\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

Also ist

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

Lemma 12.5. *Sei K ein Körper, $V = K^n$, $W = K^m$ und $U = K^r$. Seien $\varphi\colon V \longrightarrow W$, $\psi\colon U \longrightarrow V$ lineare Abbildungen mit $\varphi(y) = Ay$ und $\psi(x) = Bx$. Dann ist die Verkettung $\varphi \circ \psi\colon U \longrightarrow W$ wieder linear und gegeben durch*

$$(\varphi \circ \psi)(x) = ABx$$

Lemma 12.6. *Eine lineare Abbildung $\varphi\colon K^n \longrightarrow K^n$ ist genau dann umkehrbar, wenn die darstellende Matrix invertierbar ist. Die Umkehrabbildung ist ebenfalls linear und das Inverse der darstellenden Matrix von φ ist die darstellende Matrix der Umkehrabbildung.*

Definition 12.7. Seien V und W Vektorräume über K und $\varphi\colon V \longrightarrow W$ eine lineare Abbildung. Dann heißen

$$\ker \varphi := \{x \in V | \varphi(x) = 0\}$$

Kern von φ und

$$\operatorname{Im} \varphi := \varphi(V) = \{y \in W \mid \exists x \in V \text{ mit } \varphi(x) = y\}$$

Bild von φ

Dabei ist der Kern von φ ein Untervektorraum von V und das Bild von φ ist ein Untervektorraum von W. Der nächste Satz bringt die Dimensionen von $\ker \varphi$, $\operatorname{Im} \varphi$ und V in einen Zusammenhang.

Satz 12.8 (Rangsatz)**.** *Seien V und W endlichdimensionale Vektorräume über einem Körper K und $\varphi\colon V \longrightarrow W$ eine lineare Abbildung. Dann gilt:*

$$\dim \ker \varphi + \dim \operatorname{Im} \varphi = \dim V.$$

Definition 12.9. Der *Zeilenrang* einer Matrix A ist die Anzahl der linear unabhängigen Zeilen von A, er wird mit $\operatorname{rang}(A)$ bezeichnet. Der *Spaltenrang* einer Matrix A ist die Anzahl der linear unabhängigen Spalten von A.

Lemma 12.10. *Für jede Matrix $A \in K^{m \times n}$ gilt: Der Zeilenrang von A ist gleich dem Spaltenrang von A. Eine äquivalente Formulierung ist*

$$\operatorname{rang}(A) = \operatorname{rang}(A^t).$$

Beweis. Es sei

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}. \tag{12.1}$$

Sei r der Spaltenrang von A, also $r = \operatorname{rang} A^t$. Dann gibt es r Basisvektoren

$$b_1 = \begin{pmatrix} b_{11} \\ \vdots \\ b_{1m} \end{pmatrix}, \cdots, b_r = \begin{pmatrix} b_{r1} \\ \vdots \\ b_{rm} \end{pmatrix}$$

und die Spalten $a_{*1}, \ldots, a_{*n}$ von A lassen sich als Linearkombination der b_i schreiben:

$$\begin{aligned} a_{*1} &= \lambda_{11} b_1 + \cdots + \lambda_{1r} b_r \\ &\vdots \\ a_{*n} &= \lambda_{n1} b_1 + \cdots + \lambda_{nr} b_r \end{aligned}$$

Jeder Koeffizient a_{ij} von A lässt sich also schreiben als

$$a_{ij} = \lambda_{j1} b_{1i} + \cdots + \lambda_{jr} b_{ri}.$$

Die Koeffizienten einer Zeile a_{i*} von A können wir daher schreiben als

$$\begin{aligned} a_{i1} &= \lambda_{11} b_{1i} + \cdots + \lambda_{1r} b_{ri} \\ &\vdots \\ a_{in} &= \lambda_{n1} b_{1i} + \cdots + \lambda_{nr} b_{ri} \end{aligned}$$

Das heißt, wir können die Zeile a_{i*} als Linearkombination der r Zeilenvektoren

$$\lambda_j = (\lambda_{1j}, \ldots, \lambda_{nj}), \qquad j = 1, \ldots, r \tag{12.2}$$

darstellen:

$$a_{i*} = b_{1i}\lambda_1 + \cdots + b_{ri}\lambda_r.$$

Von den r Zeilenvektoren λ_j sind höchstens r linear unabhängig. Daher ist $\operatorname{rang}(A) \leq r$ und damit auch kleinergleich dem Spaltenrang von A.

Analog zeigt man $\operatorname{rang}(A^t) \leq \operatorname{rang}(A)$. □

Satz 12.11. *Es sei $\varphi\colon K^n \longrightarrow K^m$ eine lineare Abbildung mit darstellender Matrix A. Dann gilt:*

$$\dim \operatorname{Im} \varphi = \operatorname{rang}(A)$$

12.2 *Eigenwerte und Eigenvektoren*

Definition 12.12. Sei K ein Körper und $V = K^n$ ein Vektorraum über K und $\varphi\colon V \longrightarrow V$ eine lineare Abbildung. Ein Vektor $0 \neq v \in V$ heißt *Eigenvektor* zum *Eigenwert* λ, falls gilt

$$\varphi(v) = \lambda v.$$

Beispiel 12.13. Sei $\varphi\colon \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ gegeben durch die darstellende Matrix

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

Dann ist

$$v = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

ein Eigenvektor von φ zum Eigenwert $\lambda = 2$, denn:

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

Tatsächlich ist sogar $\begin{pmatrix} x \\ 0 \end{pmatrix}$ für alle $x \in \mathbb{R} \setminus \{0\}$ Eigenvektor zum Eigenwert 2. Analog ergibt sich, dass $\begin{pmatrix} 0 \\ y \end{pmatrix}$ für alle $y \in \mathbb{R} \setminus \{0\}$ Eigenvektor zum Eigenwert 3 ist.

Definition 12.14. Eine lineare Abbildung $\varphi\colon K^n \longrightarrow K^n$ heißt *diagonalisierbar*, falls es eine Basis des K^n aus Eigenvektoren von φ gibt.

Beispiel 12.15. Die Abbildung φ aus dem vorherigen Beispiel ist diagonalisierbar, da $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ und $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ Eigenvektoren von φ sind.

Satz 12.16. *Sei $\varphi\colon K^n \longrightarrow K^n$ eine lineare Abbildung mit darstellender Matrix A. Dann ist $\lambda \in K$ genau dann ein Eigenwert von φ, wenn* $\det(A - \lambda \mathbb{1}_n) = 0$ *ist.*

Beweis.

$$\begin{aligned}
\lambda \text{ Eigenwert von } \varphi : &\Leftrightarrow Av = \lambda v && \text{mit } v \neq 0 \\
&\Leftrightarrow Av - \lambda v = 0 && \text{mit } v \neq 0 \\
&\Leftrightarrow (A - \lambda \mathbb{1}_n)v = 0 && \text{mit } v \neq 0 \\
&\Leftrightarrow \det(A - \lambda \mathbb{1}_n) = 0 && \text{da } (A - \lambda \mathbb{1}_n) \text{ nicht invertierbar ist.}
\end{aligned}$$

□

Das Polynom $\chi_A(\lambda) = \det(A - \lambda \mathbb{1}_n)$ heißt *charakteristisches Polynom* von A (bzw. von φ).

Beispiel 12.17. Sei φ die lineare Abbildung mit der darstellenden Matrix

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

Dann ist $\chi_A(\lambda) = \lambda^2 - 5\lambda + 6$. Die Nullstellen von χ_A sind $\lambda_1 = 2$ und $\lambda_2 = 3$. Dies sind die Eigenwerte von φ. Um die Eigenvektoren zu finden müssen wir nun das lineare Gleichungssystem

$$(A - \lambda \mathbb{1}_n)x = 0$$

lösen für jeden Eigenwert λ von φ.

Für $\lambda_1 = 2$ erhalten wir das System:

$$\begin{pmatrix} 0 & 3 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Wir erhalten als Lösung $x_2 = 0$ und x_1 dürfen wir beliebig wählen. Also ist $\begin{pmatrix} x_1 \\ 0 \end{pmatrix}$ für alle $x_1 \in \mathbb{R}$ Eigenvektor zum Eigenwert $\lambda_1 = 2$.

Analog erhalten wir für $\lambda_2 = 3$:

$$\begin{pmatrix} 3x_2 \\ x_2 \end{pmatrix}$$

ist für alle $x_2 \in \mathbb{R}$ Eigenvektor zum Eigenwert $\lambda_2 = 3$.

Definition 12.18. Ist λ ein Eigenwert zu einer linearen Abbildung $\varphi\colon K^n \longrightarrow K^n$ mit darstellender Matrix A, so heißt die Lösungsmenge des linearen Gleichungssystems

$$(A - \lambda \mathbb{1}_n)v = 0 \tag{12.3}$$

Eigenraum zum Eigenwert λ. Dies ist ein Unterraum des K^n und jeder Vektor $v \neq 0$ aus diesem Unterraum ist ein Eigenvektor von φ zum Eigenwert λ.

Beispiel 12.19. Sei $\varphi\colon \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ die lineare Abbildung, welche jeden Punkt $x \in \mathbb{R}^2$ an der Geraden $\mu \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ ($\mu \in \mathbb{R}$) spiegelt.

Offenbar gilt $\varphi(e_1) = e_2$ und $\varphi(e_2) = e_1$, also ist die darstellende Matrix von φ gegeben durch

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

Das charakteristische Polynom ist $\lambda^2 - 1$, welches die Nullstellen $\lambda_1 = 1$ und $\lambda_2 = -1$ besitzt.

Der Eigenraum zum Eigenwert $\lambda_1 = 1$ ist

$$\left\{ \begin{pmatrix} x \\ x \end{pmatrix} \,\middle|\, x \in \mathbb{R} \right\}$$

Eine Basis dieses Eigenraums ist

$$\left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}$$

Der Eigenraum zum Eigenwert $\lambda_2 = -1$ ist

$$\left\{ \begin{pmatrix} x \\ -x \end{pmatrix} \,\middle|\, x \in \mathbb{R} \right\}$$

Eine Basis dieses Eigenraums ist

$$\left\{ \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\}$$

Also ist φ diagonalisierbar, da

$$\left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\}$$

eine Basis von $\mathbb{R}^2$ bilden.

Satz 12.20. *Ist eine Gerade durch den Ursprung in $\mathbb{R}^2$ gegeben, welche mit der x_1-Achse den Winkel $\alpha/2$ einschließt, so ist die lineare Abbildung, welche jeden Punkt $x \in \mathbb{R}^2$ an dieser Geraden spiegelt durch die darstellende Matrix*

$$A = \begin{pmatrix} \cos\alpha & \sin\alpha \\ sin\alpha & -\cos\alpha \end{pmatrix}$$

gegeben.

Wir bezeichnen Spiegelungen mit σ_α.

Satz 12.21. *Ist $\sigma_\alpha \colon \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ eine Spiegelung an der Ursprungsgeraden*

$$\left\{ \begin{pmatrix} \cos\frac{\alpha}{2} \\ \sin\frac{\alpha}{2} \end{pmatrix} \,\middle|\, \alpha \in \mathbb{R} \right\},$$

so hat σ_α die Eigenwerte $\lambda_1 = 1$ und $\lambda_2 = -1$.

Beweis. Die darstellende Matrix hat die Form

$$A = \begin{pmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix}$$

Das charakteristische Polynom lautet daher:

$$\begin{aligned} \chi_A(\lambda) &= \lambda^2 - \cos^2\alpha - \sin^2\alpha \\ &= \lambda^2 - (\cos^2\alpha + \sin^2\alpha) \\ &= \lambda^2 - 1 \\ &= (\lambda - 1)(\lambda + 1) \end{aligned}$$ □

Satz 12.22. *Sei σ_α wie im vorherigen Satz. Dann ist die darstellende Matrix diagonalisierbar und die Eigenräume sind orthogonal zueinander.*

Beweis. Wir berechnen die Eigenräume von σ_α für $\lambda_1 = 1$:

$$\left\{ \mu \begin{pmatrix} \cos\frac{\alpha}{2} \\ \sin\frac{\alpha}{2} \end{pmatrix} \,\middle|\, \mu \in \mathbb{R} \right\}$$

und für $\lambda_2 = -1$:

$$\left\{ \mu \begin{pmatrix} -\sin\frac{\alpha}{2} \\ \cos\frac{\alpha}{2} \end{pmatrix} \,\middle|\, \mu \in \mathbb{R} \right\}$$

Wählen wir jeweils $\mu = 1$, so erhalten wir eine Basis des jeweiligen Eigenraumes. Die beiden Vektoren

$$\left\{ \begin{pmatrix} \cos\frac{\alpha}{2} \\ \sin\frac{\alpha}{2} \end{pmatrix}, \begin{pmatrix} -\sin\frac{\alpha}{2} \\ \cos\frac{\alpha}{2} \end{pmatrix} \right\}$$

sind linear unabhängig. Sie bilden also eine Basis des $\mathbb{R}^2$. Somit ist σ_α diagonalisierbar. Wir berechnen jetzt:

$$\begin{aligned} &\left\langle \begin{pmatrix} \cos\frac{\alpha}{2} \\ \sin\frac{\alpha}{2} \end{pmatrix}, \begin{pmatrix} -\sin\frac{\alpha}{2} \\ \cos\frac{\alpha}{2} \end{pmatrix} \right\rangle \\ &= -\cos\frac{\alpha}{2}\sin\frac{\alpha}{2} + \sin\frac{\alpha}{2}\cos\frac{\alpha}{2} \\ &= 0. \end{aligned}$$

Also sind die Eigenräume orthogonal. □

Wir betrachten nun einige Beispiele für nicht-diagonalisierbare lineare Abbildungen:

Beispiel 12.23. Sei $\varphi\colon \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ die lineare Abbildung mit der darstellenden Matrix

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

Dann besitzt A das charakteristische Polynom $\chi_A(\lambda) = (1-\lambda)^2$. Dieses hat als einzige Nullstelle $\lambda_1 = 1$. Wir lösen das lineare Gleichungssystem

$$\left(\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right) \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

um die Eigenräume von A zu bestimmen:

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

hat $\left\{ \begin{pmatrix} x \\ 0 \end{pmatrix} \middle| x \in \mathbb{R} \right\}$ als Lösungsmenge. Dies sind alle Eigenvektoren von A und $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ ist eine Basis des Eigenraums zum Eigenwert 1. Dies ist allerdings keine Basis von $\mathbb{R}^2$. Somit gibt es keine Basis von $\mathbb{R}^2$ aus Eigenvektoren von φ und somit ist φ nicht diagonalisierbar.

Beispiel 12.24. Wir betrachten die lineare Abbildung ψ mit der darstellenden Matrix

$$B = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}.$$

Diese hat das charakteristische Polynom $\chi_B(\lambda) = (1-\lambda)^2 + 1 = \lambda^2 - 2\lambda + 2$. Dieses hat keine reellen Nullstellen. Damit hat ψ keine Eigenwerte und somit gibt es auch keine Eigenvektoren. Dies ist leicht einzusehen, wenn wir uns die Abbildung genauer anschauen.

Die Abbildung ψ dreht einen Vektor $v \in \mathbb{R}^2$ um einen Winkel von $\pi/4$ gegen den Uhrzeigersinn und streckt ihn dabei um den Faktor $\sqrt{2}$. So wird der Vektor $\begin{pmatrix}1\\0\end{pmatrix}$ auf den Vektor $\begin{pmatrix}1\\1\end{pmatrix}$ abgebildet.

12.3 *Eigenwerte symmetrischer Matrizen*

Symmetrische Matrizen, also Matrizen für die gilt $A = A^t$, kommen oft in praktischen Anwendungen vor. Wir wollen nun untersuchen, wie die Eigenwerte von solchen Matrizen aussehen.

Satz 12.25. *Sei $A \in \mathbb{R}^{n\times n}$ symmetrisch. Dann gilt: Die lineare Abbildung $\varphi(v) = Av$ besitzt n reelle Eigenwerte (Vielfachheiten mitgezählt) und Eigenvektoren zu verschiedenen Eigenwerten sind orthogonal.*

Beispiel 12.26. Sei $\varphi\colon \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ eine lineare Abbildung mit darstellender Matrix $A = \begin{pmatrix}1 & 2\\2 & 4\end{pmatrix}$. Dann gilt:

$$\begin{aligned}\chi_A(\lambda) &= \det(A - \lambda \mathbb{1}_2) = \lambda^2 - 5\lambda\\ &= \lambda(\lambda - 5).\end{aligned}$$

Die Abbildung φ besitzt also die beiden reellen Eigenwerte $\lambda_1 = 0$ und $\lambda_2 = 5$.

Zum Eigenwert λ_1 erhalten wir folgenden Eigenraum:

$$\left\{ \begin{pmatrix}2x\\-x\end{pmatrix} \middle| x \in \mathbb{R} \right\}$$

Eine Basis dieses Eigenraums und somit ein Eigenvektor von φ ist etwa $v_1 = \begin{pmatrix}2\\-1\end{pmatrix}$.

Zum Eigenwert λ_2 erhalten wir folgenden Eigenraum:

$$\left\{ \begin{pmatrix}x\\2x\end{pmatrix} \middle| x \in \mathbb{R} \right\}$$

Eine Basis dieses Eigenraums und somit ein Eigenvektor von φ ist etwa $v_2 = \begin{pmatrix}1\\2\end{pmatrix}$. Und tatsächlich gilt: $\langle v_1, v_2\rangle = 0$.

12.4 *Vektoriteration*

Das Vektoriterationsverfahren dient der Approximation der Eigenwerte und Eigenvektoren einer linearen Abbildung. Während für kleines n die Eigenwerte von $\varphi\colon \mathbb{R}^n \longrightarrow \mathbb{R}^n$ auch exakt relativ leicht zu bestimmen sind, ist dies für größere n nicht möglich. Wir benötigen daher ein Verfahren zur näherungsweisen Berechnung.

Dabei benötigen wir zunächst den Konvergenzbegriff im $\mathbb{R}^n$. Bisher haben wir nur Konvergenz von Zahlenfolgen untersucht, nun wollen wir dies für Folgen von Vektoren tun. Den Konvergenzbegriff können wir analog zu Zahlenfolgen definieren: Wir sagen, eine Folge von n-dimensionalen Vektoren $(x_k)_{k\in\mathbb{N}}$ konvergiert gegen einen Grenzwert $x \in \mathbb{R}^n$, falls gilt:

$$\forall \varepsilon > 0 \exists n_0 \in \mathbb{N} \forall n \geq n_0 : \|x_k - x\| < \varepsilon.$$

Außerdem definieren wir noch folgende Norm auf $\mathbb{R}^n$:

Definition 12.27. Für $x \in \mathbb{R}^n$, mit $x = (x_1, \ldots, x_n)^t$ heißt

$$\|x\|_\infty := \max\{|x_i|\}$$

die *Maximumsnorm* oder auch ∞*-Norm*.

Wir definieren nun zwei Folgen von Vektoren: Wir starten mit einem (fast) beliebigen Startvektor $x_0 \in \mathbb{R}^n$ und setzen:

$$y_0 = \frac{x_0}{\|x_0\|}.$$

Dann setzen wir:

$$\begin{aligned} y_k &= \frac{x_k}{\|x_k\|} \qquad \text{und} \\ x_{k+1} &= Ay_k \end{aligned}$$

Dabei spielt es keine Rolle, ob wir die euklidische oder die Maximumsnorm nehmen. In beiden Fällen sind die Einträge von y_k jeweils im Intervall $[-1, 1]$. Die Maximumsnorm erspart einem aber das (ggf. aufwendige) Wurzelziehen.

Satz 12.28. *Sei $A \in \mathbb{R}^{n\times n}$ die darstellende Matrix einer diagonalisierbaren linearen Abbildung φ. Sei $(v_1, \ldots, v_n)$ eine Basis des $\mathbb{R}^n$ aus Eigenvektoren von φ zu den Eigenwerten $\lambda_1, \ldots, \lambda_n$. Für die Eigenwerte gelte:*

$$|\lambda_1| > |\lambda_2| \geq \cdots \geq |\lambda_n|.$$

Sei weiterhin $0 \neq x_0 \in \mathbb{R}^n$ ein Vektor, für den in der Linearkombination

$$x_0 = \alpha_1 v_1 + \cdots + \alpha_n v_n$$

gilt: $\alpha_1 \neq 0$. Dann konvergiert die o. a. Folge $(y_k)_{k\in\mathbb{N}}$ gegen ein Vielfaches von v_1.

Beweis. Wir geben nur eine Beweisidee. Ohne Beschränkung der Allgemeinheit sei $\|v_1\| = 1$. Wir berechnen $x_1 = Ax_0$ und drücken die Koordinaten von x_1 bezüglich der Basis $v_1, \ldots, v_n$ aus:

$$\begin{aligned} x_1 = Ax_0 &= A(\alpha_1 v_1 + \cdots + \alpha_n v_n) \\ &= \alpha_1 Av_1 + \cdots + \alpha_n Av_n \\ &= \alpha_1 \lambda_1 v_1 + \cdots + \alpha_n \lambda_n v_n. \end{aligned}$$

Mittels vollständiger Induktion erhält man:

$$\begin{aligned} x_k &= \alpha_1 \lambda_1^k v_1 + \cdots + \alpha_n \lambda_n^k v_n \\ &= \alpha_1 \lambda_1^k \left(v_1 + \sum_{i=2}^{n} \left(\frac{\alpha_i}{\alpha_1} \right) \left(\frac{\lambda_i}{\lambda_1} \right)^k v_i \right) \end{aligned}$$

Da $|\lambda_1| > |\lambda_2| \geq \cdots |\lambda_n|$, gilt für große k:

$$\|x_k\| \approx \alpha_1 \lambda_1^k.$$

Dann erhält man:

$$\begin{aligned} y_k = \frac{x_k}{\|x_k\|} &\approx \frac{\sum_{i=1}^{n} \alpha_i \lambda_i^k v_i}{\alpha_1 \lambda_1^k} \\ &= v_1 + \sum_{i=2}^{n} \frac{\alpha_i}{\alpha_1} \left(\frac{\lambda_i}{\lambda_1} \right)^k v_i. \end{aligned}$$

Hier wird deutlich, warum $\alpha_1 \neq 0$ gefordert wurde. Weil λ_1 der betragsgrößte Eigenwert ist, wird y_k gegen v_1 konvergieren, und dies um so schneller, je größer $\left| \frac{\lambda_1}{\lambda_2} \right|$ ist. □

Beispiel 12.29. Sei

$$A = \begin{pmatrix} 245 & -254 & -252 & -46 & -224 \\ 161 & -168 & -174 & -32 & -148 \\ -39 & 40 & 45 & 7 & 38 \\ 27 & -28 & -32 & -6 & -26 \\ 110 & -113 & -110 & -21 & -101 \end{pmatrix}$$

Wir wollen das Vektoriterationsverfahren mit dem Startwert $x_0 := (1,1,1,1,1)^t$ durchführen.

Wir betrachten das Beispiel zunächst mit der eben eingeführten Maximumsnorm. Da der Vektor x_0 bereits 1 als maximalen Eintrag besitzt, ist $y_0 = \frac{x_0}{\|x_0\|_\infty} = x_0$ bereits normiert.

Im folgenden führen wir einige Schritte der Vektoriteration durch:

$$\begin{aligned} y_1 &= \frac{Ay_0}{\|Ay_0\|_\infty} \approx (-1.0000, -0.6798, 0.1714, -0.1224, -0.4426)^t \\ y_2 &= \frac{Ay_1}{\|Ay_1\|_\infty} \approx (-1.0000, -0.6693, 0.1717, -0.1124, -0.4431)^t \\ y_3 &= \frac{Ay_2}{\|Ay_2\|_\infty} \approx (-1.0000, -0.6688, 0.1683, -0.1126, -0.4438)^t \\ y_8 &= \frac{Ay_7}{\|Ay_7\|_\infty} \approx (-1.0000, -0.6667, 0.1667, -0.1111, -0.4444)^t \\ y_9 &= \frac{Ay_8}{\|Ay_8\|_\infty} \approx (-1.0000, -0.6667, 0.1667, -0.1111, -0.4444)^t \end{aligned}$$

Dabei stellen wir fest, dass sich der Vektor y_8 bereits nur noch sehr wenig von y_9 unterscheidet. Damit ist ein Eigenvektor angenähert

worden. Tatsächlich stellen wir fest, dass $Ay \approx 13y$ gilt für $y = (-1, -0.6667, 0.1667, -0.1111, -0.4444)^t$.

Schauen wir uns dasselbe Beispiel noch einmal mit der euklidischen Norm an. Die euklidische Norm eines Vektors $x = (x_1, x_2, x_3, x_4, x_5)^t$ ist definiert durch

$$\|x\|_2 = \sqrt{x_1^2 + x_2^2 + x_3^2 + x_4^2 + x_5^2}.$$

Nun müssen wir den Vektor $x_0 = (1, 1, 1, 1, 1)^t$ normieren. Es ergibt sich

$$y_0 = (0.4472, 0.4472, 0.4472, 0.4472, 0.4472)^t.$$

Auch hier führen wir die ersten Schritte der Vektoriteration durch:

$$\begin{aligned}
y_1 &= \frac{Ay_0}{\|Ay_0\|_2} \approx (-0.7664, -0.5211, 0.1313, -0.0938, -0.3392)^t \\
y_2 &= \frac{Ay_1}{\|Ay_1\|_2} \approx (-0.7701, -0.5154, 0.1322, -0.0866, -0.3412)^t \\
y_3 &= \frac{Ay_2}{\|Ay_2\|_2} \approx (-0.7703, -0.5152, 0.1297, -0.0868, -0.3419)^t \\
y_8 &= \frac{Ay_7}{\|Ay_7\|_2} \approx (-0.7710, -0.5140, 0.1285, -0.0857, -0.3427)^t
\end{aligned}$$

Mit dem Vektor $y = (-0.7710, -0.5140, 0.1285, -0.0857, -0.3427)^t$ rechnet man nach, dass wie oben gilt $Ay \approx 13y$, was bedeutet, dass auch mit der euklidischen Norm eine Approximation für den Eigenvektor gefunden wurde.

12.5 Einschub: Normen von Matrizen

Wir haben bereits die 2-Norm und die ∞-Norm auf K^n kennengelernt. Nun führen wir noch eine Vektornorm ein und betrachten dann Normen für Matrizen.

Definition 12.30. Sei K der Körper der reellen Zahlen. Für $x \in K^n$ setzt man

$$\|x\|_1 = \sum_{i=1}^{n} |x_i|,$$

und nennt dies die *Betragssummennorm* oder 1-Norm auf K^n.

Definition 12.31. Sei K der Körper der reellen Zahlen. Für eine Matrix $A \in K^{n \times n}$ setzt man

$$\|A\|_1 = \max_{\|x\|=1} \|Ax\|_1 = \max_j \sum_{i=1}^{n} |a_{ij}|$$

und nennt dies die *Spaltensummennorm* von A.

Lemma 12.32. *Es gilt:*

$$\|A\|_1 = \max_{x \neq 0} \frac{\|Ax\|_1}{\|x\|_1}$$

Beweis. Für jeden Vektor $x \neq 0$ aus K^n gibt es einen Vektor

$$x' := \frac{x}{\|x\|_1}$$

mit $\|x'\| = 1$. Daraus folgt die Behauptung. □

12.6 Anwendung: Googles PageRank

Eines der wesentlichen Kriterien, nach welchem Google Suchergebnisse sortiert ist der patentierte PageRank-Algorithmus, welcher von Larry Page (daher der Name) und Sergei Brin erfunden wurde. Eine Webseite hat einen hohen PageRank, wenn viele Seiten mit hohem PageRank auf diese verweisen. Dazu wird jeder Webseite ein PageRank zugewiesen, der nur auf der Linkstruktur des Webgraphen, nicht aber auf anderen Faktoren basiert.

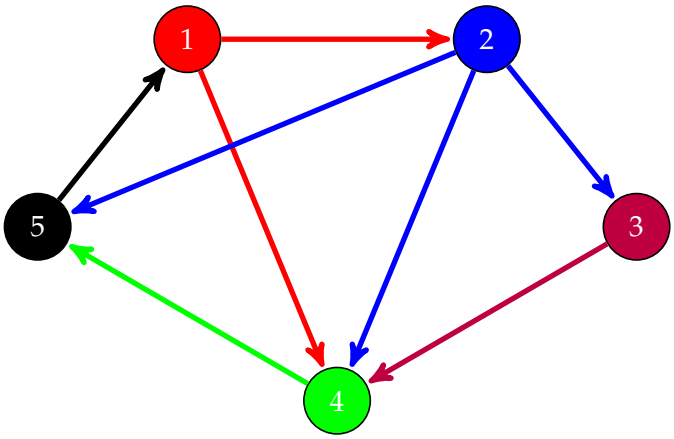
Wir nehmen dazu an, dass das World Wide Web ein Graph ist, die Knoten sind die Webseiten und die (gerichteten) Kanten sind Links von einer Webseite auf eine andere. Diesen Graphen können wir durch eine *Adjazenzmatrix* repräsentieren, genauer durch eine *spaltenstochastische Matrix*. Spaltenstochastisch bedeutet in diesem Zusammehang, dass sich alle Einträge einer Spalte zu 1 aufsummieren und alle Einträge der Matrix aus $[0,1]$ stammen. Wir tragen also in die Matrix $S \in \mathbb{R}^{n \times n}$ in den Eintrag s_{ij} ein:

$$s_{ij} = \begin{cases} 0 & \text{, falls es keinen Link von Seite } j \text{ auf Seite } i \text{ gibt;} \\ \frac{1}{k} & \text{, falls es einen Link von } j \text{ auf } i \text{ gibt und von } j \text{ insgesamt } k \text{ Links ausgehen.} \end{cases}$$

Beispiel 12.33. Für den folgenden Graphen ist die Matrix

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

die zugehörige spaltenstochastische Matrix.



Die nicht-Null-Einträge der Matrix entsprechen Links zwischen Webseiten. Ein nicht-Null-Eintrag in der Zeile i entspricht einem Link auf die i-te Webseite, ein nicht-Null-Eintrag in der Spalte j entspricht einem Link von der j-ten Webseite.

Satz 12.34. *Für jede spaltenstochastische Matrix $S \in \mathbb{R}^{n\times n}$ ist $\lambda = 1$ ein Eigenwert der zugehörigen linearen Abbildung.*

Beweis. Sei $v = (1, 1, \ldots, 1)^t \in \mathbb{R}^n$. Da sich die Einträge jeder Spalte von S zu 1 aufsummieren gilt:

$$S^t v = v = 1v.$$

Also ist 1 ein Eigenwert von S^t und damit auch von S. □

Lemma 12.35. *Für jede spaltenstochastische Matrix S mit nur positiven Einträgen gilt: Ist λ ein Eigenwert von S, so ist $|\lambda| \leq 1$.*

Beweis. Sei v ein Eigenvektor zum Eigenwert λ. Dann gilt:

$$\begin{aligned}
|\lambda| &= |\lambda| \frac{\|v\|_1}{\|v\|_1} \\
&= \frac{\|\lambda v\|_1}{\|v\|_1} \\
&= \frac{\|Sv\|_1}{\|v\|_1} \\
&\leq \max_{w\neq 0} \frac{\|Sw\|_1}{\|w\|_1} \\
&= \|S\|_1 = 1
\end{aligned}$$

□

Satz 12.36. *Sei $S \in \mathbb{R}^{n\times n}$ spaltenstochastisch und sei*

$$A = \frac{1}{n} \begin{pmatrix} 1 & \cdots & 1 \\ \vdots & & \vdots \\ 1 & \cdots & 1 \end{pmatrix} \in R^{n\times n}.$$

Dann ist $M = (1-\alpha)S + \alpha A$ für jedes $\alpha \in (0, 1]$ eine spaltenstochstische Matrix mit nur positiven Einträgen und es gelten:

1. *Der Eigenraum zum Eigenwert* 1 *hat Dimension* 1.
2. *Jeder Eigenvektor zum Eigenwert* 1 *hat nur positive oder nur negative Einträge.*

Aus 1. und 2. folgt, dass es genau einen Eigenvektor v mit nur positiven Einträgen gibt, für den gilt: $\sum_{i=1}^{n} v_i = 1$.

Beweis. Aus der Gleichung $Mv = v$ folgt: $v_i = \sum_{j=1}^{n} M_{ij} v_j$. Angenommen die v_i haben nicht alle das gleiche Vorzeichen, dann haben auch die $M_{ij} v_j$ nicht das gleiche Vorzeichen und es folgt:

$$|v_i| = \left| \sum_{j=1}^{n} M_{ij} v_j \right| < \sum_{j=1}^{n} M_{ij} |v_j|,$$

denn die M_{ij} sind alle positiv. Es folgt:

$$\begin{aligned} \sum_{i=1}^{n} |v_i| &< \sum_{i=1}^{n} \sum_{j=1}^{n} M_{ij} |v_j| \\ &= \sum_{j=1}^{n} \left(\sum_{i=1}^{n} M_{ij} \right) |v_j| \\ &= \sum_{i=1}^{n} |v_i|. \end{aligned}$$

Dies ist offensichtlich ein Widerspruch. Also müssen alle v_i das gleiche Vorzeichen haben.

Angenommen, der Eigenraum zum Eigenwert 1 hat Dimension größer als 1. Dann gibt es zwei linear unabhängige Vektoren v, w in diesem Eigenraum. Für alle $\lambda, \mu \in \mathbb{R}$ ist dann auch $x = \lambda v + \mu w$ in diesem Eigenraum. Nach der obigen Betrachtung müssen alle Einträge von x das selbe Vorzeichen haben. Aber für

$$\lambda = -\frac{\sum_{i=1}^{n} w_i}{\sum_{i=1}^{n} v_i}$$

und $\mu = 1$, ist $x \neq 0$ aber $\sum_{i=1}^{n} x_i = 0$. Also muss x Einträge mit verschiedenen Vorzeichen besitzen – ein Widerspruch. Also ist der Eigenraum 1-dimensional. □

Es gibt also einen eindeutigen Eigenvektor v mit nur positiven Einträgen, für welchen gilt:

$$\sum_{i=1}^{n} v_i = 1.$$

Wir wissen nun also, dass für jede spaltenstochastische Matrix mit nur positiven Einträgen 1 der größte Eigenwert ist. Damit können wir die Vektoriteration nutzen, um die Eigenwertaufgabe $Mv = v$ zu lösen.

Statt der Matrix S, welche die Linkverbindungen im Internet beschreibt, nutzt der PageRank-Algorithmus die Matrix $M = (1 - \alpha)S + \alpha A$ mit $\alpha \in (0, 1]$. Die Zahl α heißt Dämpfungsfaktor, ist ein

Betriebsgeheimnis von Google und liegt Studien zufolge etwa bei $\alpha = 0,15$.

Die Gleichung $v = Mv$ können wir auch schreiben als $v = (1 - \alpha)Sv + \alpha a$, wobei $a = (\frac{1}{n}, \ldots, \frac{1}{n})^t$.

Dies macht die Berechnungen einfacher.

Wir lösen also das Eigenwertproblem und erhalten einen Eigenvektor $x = (x_1, \ldots, x_n)^t$. Die Einträge von x entsprechen dem PageRank-Score der jeweiligen Webseite. Je höher der Score, desto höher steht die Webseite in den Suchergebnissen (vereinfacht ausgedrückt). In unserem Beispiel ergibt sich folgendes:

Beispiel 12.37. Wir berechnen zuerst die Matrix M und benutzen dafür den geschätzten Wert $\alpha = 0.15$. Damit ist

$$M = 0.85 \cdot \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ \frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{3} & 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{3} & 1 & 0 & 0 \\ 0 & \frac{1}{3} & 0 & 1 & 0 \end{pmatrix} + 0.15 \cdot \frac{1}{5} \cdot \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$
$$= \begin{pmatrix} 0.030 & 0.030 & 0.030 & 0.030 & 0.880 \\ 0.455 & 0.030 & 0.030 & 0.030 & 0.030 \\ 0.030 & 0.313 & 0.030 & 0.030 & 0.030 \\ 0.455 & 0.313 & 0.880 & 0.030 & 0.030 \\ 0.030 & 0.313 & 0.030 & 0.880 & 0.030 \end{pmatrix}.$$

Mit dieser Matrix und dem Startwert $x_0 = \left(\frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}\right)^t$ wird nun das Vektoriterationsverfahren durchgeführt. Wir nehmen dafür die 1-Norm (Betragssummennorm), da der Vektor x_0 in dieser Norm bereits normiert ist. Außerdem ist das Produkt aus einer spaltenstochastischen Matrix und einem Vektor x mit $||x||_1 = 1$ wieder ein Vektor, dessen Einträge sich zu 1 summieren.

Mit dem Vektoriterationsverfahren ergeben sich nun die folgenden Werte:

$$x_{10} = M^{10}x_0 \approx (0.2641, 0.1450, 0.0700, 0.2444, 0.2764)^t$$
$$x_{20} = M^{20}x_0 \approx (0.2658, 0.1429, 0.0705, 0.2434, 0.2774)^t$$
$$x_{30} = M^{30}x_0 \approx (0.2658, 0.1430, 0.0705, 0.2434, 0.2774)^t$$
$$x_{40} = M^{40}x_0 \approx (0.2658, 0.1430, 0.0705, 0.2434, 0.2774)^t$$

Wir haben damit unseren Zielvektor gefunden. Er lautet

$$y = \begin{pmatrix} 0.2658 \\ 0.1430 \\ 0.0705 \\ 0.2434 \\ 0.2774 \end{pmatrix},$$

und bedeutet, dass die Webseite 5 mit einem Ranking von 0.2774 am besten ist. Zwar hatte Webseite 4 drei eingehende Links und Webseite 5 nur zwei eingehende Links, man stellt jedoch fest, dass nach diesem Algorithmus die Webseite 3 ein so schlechtes Ranking besitzt, dass der Link von Seite 3 auf Seite 4 nur sehr wenig wert ist.

12.7 *Anwendung: Orthogonale Matrizen*

Wir betrachten nocheinmal eine Eigenschaft des Skalarprodukts:

Lemma 12.38. *Seien $x, y \in \mathbb{R}^n$ und $A \in \mathbb{R}^{n\times n}$. Dann gilt:*

$$\langle x, Ay\rangle = \langle A^t x, y\rangle$$

Beweis. Es gilt:

$$\begin{aligned}\langle x, Ay\rangle &= x^t Ay = (A^t x)^t y \\ &= \langle A^t x, y\rangle.\end{aligned}$$

□

Definition 12.39. Sei $U \in \mathbb{R}^{n\times n}$. Gilt $U^t = U^{-1}$, so heißt U *orthogonale Matrix*. Die zugehörige lineare Abbildung $\varphi(x) = Ux$ heißt *orthogonale Transformation*.

Lemma 12.40. *Seien $U \in \mathbb{R}^{n\times n}$ und $x, y \in \mathbb{R}^n$. Dann gilt:*

$$\langle Ux, Uy\rangle = \langle x, y\rangle.$$

Beweis. Da U orthogal ist, gilt:

$$\langle Ux, Uy\rangle = \langle x, U^t Uy\rangle = \langle x, y\rangle.$$

□

Lemma 12.41. *Ist U eine orthogonale Matrix, so ist auch U^t orthogonal.*

Beweis. Es gilt: $(U^t)^{-1} = (U^{-1})^t$. □

Satz 12.42. *Eine Matrix $U = (u_1, u_2, \ldots, u_n)$ ist genau dann orthogonal, wenn die Spalten $u_1, \ldots, u_n$ eine Orthonormalbasis bilden.*

Beweis. Seien $u_1, \ldots, u_n \in \mathbb{R}^n$ eine Orthonormalbasis, das heißt, für alle $i \neq j$ gilt:

$$\langle u_i, u_j\rangle = 0$$

und für alle i gilt

$$\langle u_i, u_i\rangle = 1.$$

Sei U die Matrix, deren Spalten diese u_i sind. Dann ist jeder Koeffizienten a_{ij} in der Matrix $U^t U$ gleich dem Skalarprodukt, der i-ten Zeile von U^t (d. h. die i-te Spalte von U mit der j-ten Spalte von U. Dann ist

$$a_{ij} = \begin{cases} 1, & \text{falls } i = j \\ 0, & \text{sonst.}\end{cases}$$

Also ist $U^t = U^{-1}$ und U orthogonal. □

Korollar 12.43. *Für jede orthogonale Matrix U gilt: $\left|det(U)\right| = 1$.*

Beweis. $1 = \det(\mathbb{1}_n) = \det(UU^t) = \det(U)\det(U^t) = \det(U)^2$. □

Korollar 12.44. *Sind U, V orthogonale Matrizen gleicher Dimension, dann ist auch UV ein orthogonale Matrix.*

Beweis. Es gilt:

$$(UV)^{-1} = V^{-1}U^{-1} = V^tU^t = (UV)^t.$$

□

Neben orthogonalen Matrizen (welche quadratisch sind), gibt es auch spaltenorthogonale Matrizen (welche nicht quadratisch sein müssen).

Definition 12.45. Eine Matrix $Q \in \mathbb{R}^{m\times n}$ heißt *spaltenorthogonal*, falls ihre Spalten ein Orthonormalsystem bilden, dass heißt, falls gilt:

$$Q^tQ = \mathbb{1}_n$$

Man beachte, dass im Falle $m \neq n$ die Matrix QQ^t nicht die Einheitsmatrix sein wird.

12.7.1 QR-Zerlegung

Orthogonale Matrizen haben eine Anwendung in der Lösung von linearen Gleichungssystemen. Bisher haben wir nicht-lösbare Gleichungssysteme nicht weiter betrachtet, dies wollen wir nun nachholen.

Angenommen wir haben ein lineares Gleichungssystem $Ax = b$, dann ist dieses genau dann lösbar, wenn b im Bild der linearen Abbildung $\varphi(x) = Ax$ ist. Dies ist natürlich nicht zwingend der Fall. Wir wollen nun für ein nicht-lösbares System eine „Quasi-Lösung“ finden, also einen Vektor b, der so nah wie möglich am Bild von φ liegt. Wir suchen also einen Vektor b, so dass $\|Ax - b\|$ möglichst klein ist. Dazu zerlegen wir den Vektor b in zwei Vektoren $b = b_{\|} + b_{\perp}$, wobei $b_{\|} \in \mathrm{Im}(\varphi)$ und $b_{\perp}$ orthogonal zu $\mathrm{Im}(\varphi)$ sein soll.

Dann ist das lineare Gleichungsystem $Ax = b_{\|}$ lösbar und wir erhalten:

$$\|Ax - b\|^2 = \|Ax - b_{\|}\|^2 + \|b_{\perp}\|^2.$$

Dabei hängt $\|b_{\perp}\|$ nicht von x ab und gibt den Fehler unserer Lösung an. Nun müssen wir nur noch klären, wie man die Zerlegung finden kann.

Satz 12.46. *Sei $e \in \mathbb{R}^n$ ein Einheitsvektor, dass heißt, $\|e\| = 1$. Dann lässt sich jeder Vektor $v \in \mathbb{R}^n$ bezüglich e in zwei zueinander orthogonale Komponenten zerlegen:*

$$v = v_{\|} + v_{\perp}.$$

Dabei ist

$$v_{\parallel} = \langle e, v\rangle e$$

die orthogonale Projektion von v in Richtung von e *und*

$$v_{\perp} = v - \langle e, v\rangle e.$$

das orthogonale Komplement von v in Richtung von e.

Beweis. Offensichtlich ist $v = v_{\parallel} + v_{\perp}$ und es gilt:

$$\begin{aligned}\langle v_{\parallel}, v_{\perp}\rangle &= \big\langle \langle e, v\rangle e, v - \langle e, v\rangle e\big\rangle \\ &= \big\langle \langle e, v\rangle e, v\big\rangle - \big\langle \langle e, v\rangle e, \langle e, v\rangle e\big\rangle \\ &= \langle e, v\rangle^2 - \langle e, v\rangle^2 \langle e, e\rangle \\ &= \langle e, v\rangle^2 - \langle e, v\rangle^2 = 0.\end{aligned}$$

□

Sei $\mathbb{R}^{m\times n} \ni Q = (u_1, \ldots, u_n)$ eine spaltenorthogonale Matrix. Dann gilt für $v \in \mathbb{R}^n$:

$$QQ^t v = Q \begin{pmatrix} \langle u_1, v\rangle \\ \vdots \\ \langle u_n, v\rangle \end{pmatrix} = \sum_{j=1}^{n} \langle u_j, v\rangle u_j = v_{\parallel}.$$

Damit wird v durch QQ^t orthogonal auf den von $u_1, \ldots, u_n$ aufgespannten Unterraum projiziert. Man nennt QQ^t auch *orthogonaler Projektor*:

Definition 12.47. Eine symmetrische Matrix P mit der Eigenschaft $P^2 = P$ heißt *orthogonaler Projektor*.

Offenbar erfüllt $P = QQ^t$ diese Definition, denn $P^t = (QQ^t)^t = (Q^t)^t Q^t = QQ^t = P$, also ist QQ^t symmetrisch und $P^2 = QQ^tQQ^t = Q\mathbb{1}_n Q^t = QQ^t$.

Dies hilft uns nun beim Finden der Zerlegung von $b = b_{\parallel} + b_{\perp}$.

Wir nehmen die Spalten der Koeffizientenmatrix A (aus dem LGS $Ax = b$) und bilden daraus (etwa mittels des Gram-Schmidt-Verfahrens) eine Orthonormalbasis und den zugehörigen orthogonalen Projektor QQ^t. Dann gilt: $QQ^tA = A$, denn die Spalten von A und die Spalten von Q erzeugen den selben Unterraum. Sodann berechnen wir:

$$\begin{aligned}Ax - b_{\parallel} &= QQ^tAx - QQ^tb \\ &= Q(Rx - Q^tb),\end{aligned}$$

wobei R die folgende Matrix ist:

$$R = Q^tA = \begin{pmatrix} \langle u_1, a_1\rangle & \cdots & \langle u_1, a_n\rangle \\ \vdots & & \vdots \\ \langle u_r, a_1\rangle & \cdots & \langle u_r, a_n\rangle \end{pmatrix}$$

Dabei sind die a_j die Spalten von A und r der Rang von A.

Werden $u_1, \ldots, u_r$ mittels des Gram-Schmidt-Verfahrens aus $a_1, \ldots, a_n$ erzeugt, können die a_j als Linearkombination der Vektoren u_i mit $i \leq j$ geschrieben werden. Daher gilt:

$$\langle u_i, a_j \rangle = 0$$

für $i > j$. Daher hat die Matrix R folgende Gestalt:

$$R = \begin{pmatrix} \langle u_1, a_1 \rangle & \langle u_1, a_2 \rangle & \cdots & \langle u_1, a_n \rangle \\ 0 & \langle u_2, a_2 \rangle & \cdots & \langle u_2, a_n \rangle \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \langle u_r, a_n \rangle \end{pmatrix}$$

Die Matrix R ist also eine obere rechte Dreiecksmatrix.

Satz 12.48 (QR-Zerlegung)**.** *Jede Matrix A kann als Produkt einer spaltenorthogonalen Matrix Q und einer oberen Dreiecksmatrix R geschrieben werden.*

Sind die Spalten von A linear unabhängig, so sind die Diagonalelemente von R alle ungleich Null und R ist invertierbar. Dann ist das Gleichungssystem $Rx = Q^t b$ besonders leicht zu lösen.

12.8 Das Gram-Schmidt-Verfahren

Im letzten Abschnitt mussten wir für die QR-Zerlegung zu einer Basis $v_1, \ldots, v_n$ eine Orthonormalbasis $u_1, \ldots, u_n$ finden, so dass beide Basen den selben Untervektorraum aufspannen. Wie wir so eine Basis finden wollen wir uns nun anschauen. Das Verfahren zur Orthonormalisierung einer Basis ist nach dem dänischen Mathematiker Jørgen Pedersen Gram (*1850, †1916) und dem deutschen Mathematiker Erhard Schmidt (*1876, †1959) benannt.

Satz 12.49 (Gram-Schmidt-Verfahren)**.** *Sei V ein reeller Vektorraum und $v_1, \ldots, v_n \in V$ linear unabhängige Vektoren. Die Vektoren*

$$u_k := \frac{v_k - \sum_{j=1}^{k-1} \langle u_j, v_k \rangle u_j}{\| v_k - \sum_{j=1}^{k-1} \langle u_j, v_k \rangle u_j \|}$$

bilden ein Orthonormalsystem mit $\operatorname{span}(u_1, \ldots, u_n) = \operatorname{span}(v_1, \ldots, v_n)$.

Wir verzichten an dieser Stelle auf einen Beweis.

Korollar 12.50. *Jeder reelle Vektorraum hat eine Orthonormalbasis.*

Beispiel 12.51. Wir betrachten die drei Vektoren

$$v_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, v_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, v_3 = \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix}$$

Den ersten Vektor der gesuchten Orthonormalbasis erhalten wir, in dem wir den Vektor v_1 normalisieren. Dazu teilen wir v_1 durch seine Norm:

$$u_1 = \frac{v_1}{\|v_1\|} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}.$$

Den Vektor u_2 erhalten wir dadurch, dass wir zunächst von v_2 die orthogonale Projektion in Richtung von u_1 abziehen

$$u_2' = v_2 - \langle u_1, v_2 \rangle u_1 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix}$$

und diesen Vektor dann normieren:

$$u_2 = \frac{u_2'}{\|u_2'\|} = \frac{1}{\sqrt{6}} \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix}.$$

Es verbleibt der dritte Vektor u_3 zu konstruieren. Wir gehen analog vor: Wir substrahieren von v_3 die orthogonale Projektion auf den von u_1, u_2 aufgespannten Vektorraum:

$$u_3' = v_3 - \langle u_1, v_3 \rangle u_1 - \langle u_2, v_3 \rangle u_2 = \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix} - \frac{5}{2} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} - \frac{7}{6} \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}$$

und normieren:

$$u_3 = \frac{u_3'}{\|u_3'\|} = \frac{1}{\sqrt{3}} \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}.$$

12.9 Basiswechsel

Das ganze Bestimmen von Orthonormalbasen oder überhaupt von anderen Basen außer der Standardbasis ist ja nur dann nützlich, wenn wir in der Lage sind, auch die Koordinaten eines Vektors bezüglich verschiedenen Basen zu berechnen, beziehungsweise die Koordinaten bezüglich der einen in die Koordinaten der anderen Basis umzurechnen.

Sei V ein reeller Vektorraum mit Basis $B = (v_1, \ldots, v_n)$. Dazu gehört ein Isomorphismus $\Phi_B \colon \mathbb{R}^n \longrightarrow V$ mit $\Phi_B(e_i) = v_i$ für alle $i = 1, \ldots, n$. Ist $x = (x_1, \ldots, x_n)^t \in \mathbb{R}^n$,so ist

$$\Phi_B(x) = \sum_{i=1}^{n} x_i v_i =: v.$$

Dabei heißt Φ_B das durch die Basis B bestimmte Koordinatensystem in V und $x = \Phi_B^{-1}(v) \in \mathbb{R}^n$ sind die Koordinaten von v.

Hat man nun zwei Basen $A = (v_1, \dots, v_n)$ und $B = (w_1, \dots, w_n)$ von V, so haben wir entsprechend zwei Isomorphismen: $\Phi_A, \Phi_B \colon \mathbb{R}^n \longrightarrow V$ und einen Isomorphismus

$$T_B^A := \Phi_B^{-1} \circ \Phi_A,$$

den man als Matrix aufgefasst die *Transformationsmatrix* des Basiswechsel genannt wird.

Satz 12.52. *Sei T_B^A die Transformationsmatrix des Basiswechsels, so gilt für*

$$v = \sum_{i=1}^{n} x_i v_i = \sum_{i=1}^{n} y_1 w_1 \in V :$$
$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = T_B^A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Mittels T_B^A kann man also die Koordinaten y bezüglich der Basis B aus den Koordinaten x bezüglich der Basis A bestimmen.

Beispiel 12.53 (Wichtigstes Beispiel)**.** Sei $V = K^n$ und A und B die Matrizen deren Spalten jeweils die Vektoren einer Basis sind. Dann erhalten wir die Transformationsmatrix $T = B^{-1}A$. Ist insbesondere A die Einheitsmatrix (also die Matrix für die Standardbasis), so ist $T = B^{-1}$.

Ein Wechsel der Basis beeinflusst auch lineare Abbildungen.

Satz 12.54. *Seien V ein reeller Vektorraum mit Basis A und W ein reeller Vektorraum mit Basis B. Sei $\varphi \colon V \longrightarrow W$ eine lineare Abbildung. Dann gilt ist die darstellende Matrix $M_B^A(\varphi)$ bezüglich der Basen A und B der Abbildung φ: $M_B^A(\varphi) = \Phi_B^{-1} \circ \varphi \circ \Phi_A$.*

Satz 12.55 (Transformationsformel)**.** *Sei $\varphi \colon V \longrightarrow W$ eine lineare Abbildung, und A, A' seinen Basen von V und B, B' seien Basen von W. Dann gilt für die darstellenden Matrizen von φ:*

$$M_{B'}^{A'}(\varphi) = T_{B'}^{B} \cdot M_B^A(\varphi) \cdot \left(T_{A'}^{A}\right)^{-1}.$$

Zeit für ein Beispiel:

Beispiel 12.56. Sei A die Basis bestehend aus den Vektoren

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, v_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, v_3 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

und B die Basis bestehend aus den Vektoren

$$w_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, w_2 = \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix}, w_3 = \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix}$$

Wir wollen den Vektor

$$v = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

bezüglich der beiden Basen darstellen.

Wir erhalten zunächst die Matrizen

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

und bestimmen ihre Inversen:

$$\Phi_A^{-1} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & -1 & 1 \\ 1 & 0 & -1 \end{pmatrix}, \Phi_B^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & \frac{1}{2} & -1 \end{pmatrix}.$$

Die Koordinaten von v bezüglich der Basis A ist dann

$$\Phi_A^{-1} v = \begin{pmatrix} 0 & 1 & 0 \\ 0 & -1 & 1 \\ 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

was sicherlich Sinn macht, da v der erste Basisvektor der Basis A ist. Für die Koordinaten bezüglich B erhalten wir analog:

$$\Phi_B^{-1} v = \begin{pmatrix} 1 \\ 1 \\ -\frac{1}{2} \end{pmatrix}.$$

und tatsächlich ist $1w_1 + 1w_2 - \frac{1}{2}w_3 = v$.

Für die Basiswechselmatrix erhalten wir:

$$T_B^A = \Phi_B^{-1} \Phi_A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ -\frac{1}{2} & -1 & 0 \end{pmatrix}.$$

Diese Matrix liefert uns den Basiswechsel von A nach B. Haben wir also einen Vektor in den Koordinaten von A gegeben, können wir diesen umrechnen in die Koordinaten von B.

Betrachte:

$$T_B^A \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ -\frac{1}{2} \end{pmatrix},$$

wie erwartet.

Nun zu linearen Abbildungen.

Beispiel 12.57. Sei φ die lineare Abbildung deren darstellende Matrix (bezüglich der Standardbasis) gegeben ist durch

$$M_E^E(\varphi) = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Dabei fassen wir φ auf als lineare Abbildung von $V = \mathbb{R}^3$ nach $W = \mathbb{R}^3$ (jeweils mit der Standardbasis). Nun fragen wir uns, wie ändert sich die darstellende Matrix von φ, wenn wir die Basis von V ändern auf A und die Basis von W ändern auf B?

Wir müssen zunächst die Basiswechselmatritzen bestimmen. Sei dazu E die Standardbasis. Dann sind

$$T_B^E = \Phi_B^{-1},\ T_A^E = \Phi_A^{-1}$$

und $\left(T_A^E\right)^-1 = \Phi_A$. Somit ist die darstellende Matrix von φ bezüglich der Basen A, B:

$$M_B^A(\varphi) = T_B^E M_E^E(\varphi)(T_A^E)^{-1} = \begin{pmatrix} 6 & 4 & 1 \\ 1 & 1 & 0 \\ -\frac{1}{2} & -1 & 0 \end{pmatrix}.$$

Exemplarisch überprüfen wir wie der Vektor v unter φ abgebildet wird. Bezüglich der Standardbasis erhalten wir:

$$\varphi(v) = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 6 \\ 1 \\ 1 \end{pmatrix}.$$

. Wir wissen schon, dass v bezüglich A die Koordinaten $(1, 0, 0)^t$ hat. Wir bilden also ab:

$$\begin{pmatrix} 6 & 4 & 1 \\ 1 & 1 & 0 \\ -\frac{1}{2} & -1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 6 \\ 1 \\ -\frac{1}{2} \end{pmatrix}.$$

Und tatsächlich ist $6w_1 + 1 + w_2 - \frac{1}{2}w_3 = \varphi(v)$.

THE END