

# Skript zur Vorlesung Mathematik 1

Martin Redlof

Stand: 18. März 2025

# Inhaltsverzeichnis

<b>0</b>	<b>Einleitung</b>	<b>7</b>
0.1	Zum Skript . . . . .	7
0.1.1	Inhalt . . . . .	7
0.1.2	Begriffe . . . . .	7
0.1.3	Organisation . . . . .	8
0.2	Zum Studium . . . . .	8
<b>1</b>	<b>Grundlagen</b>	<b>10</b>
1.1	Aussagenlogik: ((To Be) or (not(To Be))) . . . . .	10
1.1.1	Elementare Begriffe . . . . .	10
1.1.2	Prädikate . . . . .	15
1.1.3	Prädikatenlogik . . . . .	16
1.1.4	Implikation, Äquivalenz . . . . .	19
1.1.5	Indirekte Beweise . . . . .	22
1.2	Mengenlehre . . . . .	23
1.2.1	Grundbegriffe . . . . .	23
1.2.2	Teilmengen und Potenzmenge . . . . .	25
1.2.3	Bezug zur Aussagenlogik: Mengenoperationen und Rechenregeln . . . . .	27
1.2.4	Konstruktion von Mengen . . . . .	30
1.2.5	Kartesisches Produkt . . . . .	32
1.3	Wichtige Zahlenmengen . . . . .	33
1.3.1	Natürliche Zahlen $\mathbb{N}$ . . . . .	34
1.3.2	Ganze Zahlen $\mathbb{Z}$ . . . . .	34
1.3.3	Rationale Zahlen $\mathbb{Q}$ . . . . .	35
1.3.4	Reelle Zahlen $\mathbb{R}$ . . . . .	35
1.4	(Un-)Gleichungen und Äquivalenzumformungen . . . . .	36
1.4.1	(Un-)Gleichungen sind logische Aussagen . . . . .	36
1.4.2	(Un-)Gleichungen durch Äquivalenzumformung lösen . . . . .	38
1.4.3	Lineare Gleichungen in einer reellen Veränderlichen . . . . .	41
1.4.4	Lineare Ungleichungen in einer reellen Veränderlichen . . . . .	42
1.4.5	Binomische Formeln und quadratische Gleichungen in einer reellen Variablen . . . . .	43
1.5	Betrag, Summen- und Produktzeichen . . . . .	46
1.6	Fakultätsfunktion und Binomialkoeffizienten . . . . .	49
<b>2</b>	<b>Teilbarkeit und Primzahlen</b>	<b>52</b>
2.1	Konzepte . . . . .	52
2.1.1	Teiler . . . . .	52
2.1.2	Primzahlen . . . . .	54
2.1.3	Division mit Rest . . . . .	55
2.2	Größter gemeinsamer Teiler . . . . .	57
2.2.1	Gemeinsame Teiler . . . . .	57
2.2.2	Euklidischer Algorithmus . . . . .	61
2.3	Anwendungen . . . . .	65
2.3.1	Lemma von Bezout . . . . .	65
2.3.2	Lemma von Euklid . . . . .	69
2.3.3	Teilerfremdheit bei Produkten ganzer Zahlen . . . . .	69

<b>3</b>	<b>Restklassen</b>	<b>72</b>
3.1	Äquivalenz modulo $m$	72
3.2	Rechenoperationen mit Restklassen	74
3.2.1	Addition und Subtraktion modulo $m$	74
3.2.2	Multiplikation modulo $m$	75
3.2.3	Zur Division modulo $m$	76
3.2.4	Potenzieren modulo $m$	77
3.3	Multiplikation und Division mit Restklassen	78
3.3.1	Multiplikationstabellen modulo $m$ und Nullteiler	78
3.3.2	Multiplikatives Inverses modulo $m$	79
3.3.3	Prime Restklassensysteme	82
3.3.4	Kleiner Satz von Fermat	83
3.3.5	Auffinden der multiplikativen Inversen	84
<b>4</b>	<b>Funktionen</b>	<b>85</b>
4.1	Allgemeine Eigenschaften	85
4.1.1	Funktionsbegriff	85
4.1.2	Abbildungseigenschaften	88
4.1.3	Verkettung (Hintereinanderausführung) von Funktionen	92
4.1.4	Addition, Multiplikation, Division	94
4.2	Basisfunktionen	94
4.2.1	Potenzfunktionen	95
4.2.2	Rationale Funktionen	95
4.2.3	Trigonometrische Funktionen (mehr dazu in Mathematik 2/Analysis)	98
4.2.4	Exponentialfunktion (mehr dazu in Mathematik 2/Analysis)	99
4.3	Permutationen (Einführung)	100
4.3.1	Permutationsfunktionen	100
4.3.2	Zyklen und Transpositionen	101
4.3.3	Verkettung von Permutationen	106
4.3.4	Inverse Permutation	108
<b>5</b>	<b>Algebra</b>	<b>111</b>
5.1	Motivation/Grundlagen	111
5.2	Gruppenartige Strukturen	112
5.2.1	Abgeschlossenheit $\rightsquigarrow$ Magma	112
5.2.2	Assoziativität $\rightsquigarrow$ Halbgruppe	113
5.2.3	Neutrales Element $\rightsquigarrow$ Monoid	113
5.2.4	Inverse Elemente $\rightsquigarrow$ Gruppe	115
5.2.5	Untergruppen	117
5.2.6	Symmetrische Gruppe $S_n$	118
5.3	Ringartige Strukturen	120
5.3.1	Halbringe	121
5.3.2	Ringe	121
5.3.3	Körper	123
5.4	Polynome in einer Veränderlichen	125
5.4.1	Definitionen und Beispiele	126
5.4.2	Polynomfunktion vs. Polynom	127
5.4.3	Rechenregeln, Polynomring	128
5.4.4	Polynomdivision	135
5.4.5	Nullstellen	143
<b>6</b>	<b>Lineare Algebra: Vektoren</b>	<b>147</b>
6.1	Vektorräume	147
6.1.1	Koordinatensysteme und Dimensionen	147
6.1.2	Motivation: Ortsvektoren	148
6.1.3	Vektorraumbegriff	149
6.1.4	Kartesische Produkträume	150
6.1.5	Skalarprodukt bei $\mathbb{R}$ -Vektorräumen	153
6.1.6	Norm	154
6.1.7	Metrik	155
6.2	Die reellen Vektorräume $\mathbb{R}^n$	156

6.2.1	Kanonisches Skalarprodukt in $\mathbb{R}^n$ . . . . .	156
6.2.2	Geometrische Bedeutung des kanonischen Skalarprodukts . . . . .	160
6.2.3	Kreuzprodukt in $\mathbb{R}^3$ . . . . .	162
6.2.4	Geometrische Bedeutung des Kreuzprodukts . . . . .	163
6.3	Lineare Abhängigkeit . . . . .	166
6.3.1	Definitionen . . . . .	166
6.3.2	Basis und Dimension . . . . .	172
6.3.3	Kronecker-Delta und Standardbasis von $\mathbb{R}^n$ . . . . .	173
6.3.4	Orthogonale und Orthonormale Basen bei Innenprodukträumen . . . . .	176
6.3.5	Koordinaten eines Vektors in einer Orthonormalbasis . . . . .	177
<b>7</b>	<b>Lineare Algebra: Lineare Abbildungen und Matrizen</b>	<b>178</b>
7.1	Lineare Abbildungen . . . . .	178
7.2	Matrizen . . . . .	180
7.2.1	Definition . . . . .	180
7.2.2	Zusammenhang mit linearen Abbildungen . . . . .	182
7.2.3	Bestimmung der Abbildungsmatrix in der Praxis . . . . .	183
7.2.4	Produkt aus Matrix und Vektor . . . . .	185
7.2.5	Zusammenfassung zu linearen Abbildungen . . . . .	186
7.2.6	Matrizenräume . . . . .	188
7.3	Matrixprodukt . . . . .	189
7.3.1	Komposition linearer Abbildungen . . . . .	189
7.3.2	Herleitung des Matrixprodukts . . . . .	190
7.3.3	Eigenschaften des Matrixprodukts . . . . .	192
7.3.4	Quadratische Matrizen . . . . .	193
7.4	Matrixtransposition . . . . .	194
7.4.1	Definition . . . . .	194
7.4.2	Eigenschaften . . . . .	195
7.4.3	Kanonisches Skalarprodukt . . . . .	196
7.4.4	Symmetrische Matrizen . . . . .	197
7.5	Inverse Matrizen (Einführung) . . . . .	198
7.6	Orthogonale Matrizen (Einführung) . . . . .	201
7.7	Permutationsmatrizen . . . . .	205
<b>8</b>	<b>Lineare Algebra: Gleichungssysteme und Determinanten</b>	<b>209</b>
8.1	Lineare Gleichungssysteme . . . . .	209
8.1.1	Definition und Einordnung . . . . .	209
8.1.2	Lösung von LGS: Überlegungen . . . . .	211
8.1.3	Lösung von LGS: Vorgehen . . . . .	213
8.1.4	Schematisches Beispiel . . . . .	214
8.1.5	Beispiel: Eindeutig lösbares LGS . . . . .	216
8.1.6	Beispiel: Mehrdeutig lösbares LGS . . . . .	217
8.1.7	Beispiel: Unlösbares LGS . . . . .	219
8.1.8	Bestimmtheit und Lösbarkeit von LGS . . . . .	219
8.1.9	Kern einer Matrix/Linearen Abbildung . . . . .	220
8.1.10	Rang einer Matrix/Linearen Abbildung . . . . .	221
8.2	Permutationen Revisited . . . . .	222
8.2.1	Recap . . . . .	222
8.2.2	Zyklen und Transpositionen . . . . .	222
8.2.3	Vorzeichen einer Permutation . . . . .	225
8.3	Determinanten . . . . .	228
8.3.1	Allgemeine Eigenschaften . . . . .	228
8.3.2	Permutationsmatrizen . . . . .	230
8.3.3	Leibniz-Formel . . . . .	231
8.3.4	Transponierte Matrizen . . . . .	233
8.3.5	Dreiecksmatrizen (und Diagonalmatrizen) . . . . .	234
8.3.6	Laplace-Entwicklung . . . . .	235
8.3.7	Vereinfachung von Determinantenberechnungen mit Zeilen- und Spaltenoperationen . . . . .	237
8.3.8	Determinanten und LGS . . . . .	241
8.3.9	Determinanten-Produktsatz . . . . .	241

<b>9</b>	<b>Lineare Algebra: Matrizen Revisited</b>	<b>243</b>
9.1	Invertierung reeller quadratischer Matrizen . . . . .	243
9.1.1	Eigenschaften inverser Matrizen (Vervollständigung) . . . . .	243
9.1.2	Berechnung der Inversen mit Gauß . . . . .	245
9.1.3	Die lineare Gruppe . . . . .	248
9.2	Orthogonale Matrizen (Fortsetzung) . . . . .	248
9.2.1	Eigenschaften . . . . .	248
9.2.2	Drehmatrizen . . . . .	249
9.2.3	Aktive Drehungen in $\mathbb{R}^2$ . . . . .	249
9.2.4	Passive Drehungen . . . . .	251
9.2.5	Drehungen in $\mathbb{R}^3$ . . . . .	251
9.2.6	Orthogonale Gruppe . . . . .	251
9.3	Eigenwertproblem . . . . .	252
9.3.1	Motivation und Formulierung . . . . .	252
9.3.2	Eigenräume und geometrische Vielfachheit . . . . .	253
9.3.3	Lösung des Eigenwertproblems . . . . .	254
9.3.4	Berechnung von Eigenräumen . . . . .	257
9.3.5	Weitere Eigenschaften von Eigenvektoren und Eigenwerten . . . . .	261
9.3.6	Dreiecks- und Diagonalmatrizen . . . . .	263
9.3.7	Drehmatrizen für $\mathbb{R}^3$ . . . . .	265
9.4	Diagonalisierung . . . . .	265
9.4.1	Ähnliche Matrizen . . . . .	265
9.4.2	Diagonalisierbarkeit . . . . .	267
9.4.3	Zusammenfassung Diagonalisierbarkeit . . . . .	269
9.4.4	Symmetrische Matrizen und Spektralsatz . . . . .	269
<b>A</b>	<b>Exkurse</b>	<b>271</b>
A.1	Dieder-Gruppe $D_5$ . . . . .	271
A.1.1	Intro . . . . .	271
A.1.2	Permutationen für die Symmetrieoperationen . . . . .	272
A.1.3	Nachweis der Gruppeneigenschaft . . . . .	272
A.1.4	Abgeschlossenheit . . . . .	272
A.1.5	Gruppentafel von $D_5$ . . . . .	273
A.1.6	Nebenbemerkung: Die restlichen Permutationen in $S_5$ . . . . .	274
A.2	Horner-Schema . . . . .	274
A.2.1	Polynom-Auswertung . . . . .	274
A.2.2	Stellenwertsysteme . . . . .	276
A.3	Geraden und Ebenen . . . . .	276
A.3.1	Geraden in $\mathbb{R}^2$ . . . . .	277
A.3.2	Geraden in $\mathbb{R}^n$ , $n > 2$ . . . . .	282
A.3.3	Ebenen in $\mathbb{R}^n$ , $n > 2$ . . . . .	283
A.3.4	Abstandsprobleme . . . . .	286
A.3.5	Schnittprobleme . . . . .	287
A.3.6	Hyperebenen (kein Vorlesungsstoff) . . . . .	295
A.4	Gram-Schmidt-Orthogonalisierung . . . . .	296
A.5	Vektorraum der Polynome . . . . .	300
A.6	Basiswechsel mit Matrizen . . . . .	301
A.6.1	Invertierbare Matrizen als Basis von $\mathbb{R}^n$ . . . . .	301
A.6.2	Basiswechsel zwischen Standardbasis und $\mathcal{A}$ und $\mathcal{B}$ . . . . .	302
A.6.3	Basiswechsel allgemein . . . . .	302
A.6.4	Zahlenbeispiel in $\mathbb{R}^3$ . . . . .	302
A.7	Matrizenoperationen für das Gaußverfahren . . . . .	304
A.7.1	Vertauschen von Zeilen . . . . .	305
A.7.2	Skalieren von Zeilen mit $c \neq 0$ . . . . .	306
A.7.3	Addition der mit $c$ skalierten Zeile $j$ zu Zeile $k \neq j$ . . . . .	307
A.7.4	Zusammenfassung . . . . .	308
A.8	Regel von Cramer für LGS . . . . .	308

<b>B</b>	<b>Ausgewählte Beweise</b>	<b>311</b>
B.1	Für Kapitel 1 . . . . .	311
B.2	Für Kapitel 2 . . . . .	313
B.3	Für Kapitel 3 . . . . .	317
B.4	Für Kapitel 5 . . . . .	320
B.5	Für Kapitel 6 . . . . .	321
B.6	Für Kapitel 7 . . . . .	327
B.7	Für Kapitel 8 . . . . .	332
	<b>Literaturverzeichnis</b>	<b>351</b>

# Kapitel 0

## Einleitung

### 0.1 Zum Skript

#### 0.1.1 Inhalt

Dieses Skript behandelt die Themen der Vorlesung Mathematik 1:

- Aussagen und Mengen

Ohne diese lässt sich kaum ein mathematischer Zusammenhang formulieren.

- Relationen und Funktionen

Weiteres “Handwerkszeug”. Für die Informatik ist die *modulo-Arithmetik* (also das Rechnen mit Restklassen) von besonderer Bedeutung: Computer haben nur eine bestimmte Registerbreite, und auch Datentypen für Zahlen sind in der Regel in der Größe beschränkt<sup>1</sup>. Sind Zahlencodierungen zu groß, muss das gesondert behandelt werden. Bei Ganzzahlen (integers) wird hierbei oft “abgeschnitten”, was effektiv eine modulo-Rechnung bedeutet.

Dem Rechnen mit Restklassen ist ein eigenes Kapitel gewidmet.

- Algebra (Gruppen, Ringe, Körper, Polynome)

Gerade zu Anfang des Kapitels: Einige mathematische “Buchhaltung”. Die Konzepte der Algebra sind überall dort wichtig, wo symbolisch (mit Variablen, nicht wie in der Arithmetik mit Zahlenwerten) gerechnet werden muss. Wichtig z.B. für Kryptografie oder Codierungstheorie.

- Lineare Algebra (Vektorräume und Matrizen)

Hier lernen Sie mit Vektoren und Matrizen zu hantieren. Wichtig z.B. für Robotik, Computergrafik, oder im Bereich Machine Learning.

Einige zusätzliche, nicht prüfungsrelevante Abschnitte sind im Anhang A zu finden; diese sind für Neugierige gedacht, bzw. können im Zeitrahmen der Vorlesung nicht behandelt werden.

#### 0.1.2 Begriffe

In der Mathematik ist es recht gut möglich, genau zu formulieren – das bedeutet aber auch, dass einige feste Begriffe vereinbart werden, die dann zu benutzen sind. Beispiele hierfür sind:

- *Axiom (das)*: Eine Grundtatsache, die als richtig gefordert und angesehen wird. Axiome bilden das Fundament der Mathematik; aus ihnen werden alle anderen Zusammenhänge abgeleitet. Wenn wir vermeiden wollen, uns logisch im Kreis zu drehen (“das gilt, weil es gilt”), müssen wir akzeptieren, dass bestimmte Ausgangs-Zusammenhänge sich nicht beweisen lassen. Ziel ist es natürlich, mit möglichst wenigen solchen Ansatzpunkten auszukommen.
- *Satz*: Ein Text, der einen mathematischen Zusammenhang beschreibt. Dazu gehört in der Regel ein *Beweis*, der logisch “wasserdicht” den behaupteten Zusammenhang begründet (etwa durch Zurückführen auf andere bereits bewiesene Sätze, oder auf Axiome).

---

<sup>1</sup>Ein Typ wie `BigInteger` in Java lässt auch größere Zahlen zu; allerdings sind die Rechnungen dann sicher nicht mehr durch einzelne Prozessorbefehle ausführbar. Und selbst “beliebig große” Zahlen müssen irgendwo in einem – endlich großen – Speicher abgelegt werden!

- *Definition:* Hier wird mathematisches Vokabular vereinbart, um Zusammenhänge zu benennen oder abkürzend zu beschreiben. Oft werden Definitionen benutzt, um Objekte mit gewissen gemeinsamen Eigenschaften semantisch (“ideen-mäßig”) zu gruppieren, um in der Folge genau solche Objekte präzise ansprechen zu können.

Vorsicht: Mathematische und umgangssprachliche Begriffe heißen manchmal gleich, stehen aber für verschiedene Konzepte! Zum Beispiel ist “eine Menge Steine” mathematisch interpretiert eine Art mathematischer Behälter, der Steine enthält – aber das wäre auch dann richtig, wenn es sich dabei bloß um zwei oder drei Steine handeln würde. Umgangssprachlich gemeint war hier vermutlich eher “eine große Anzahl von Steinen”. In der Mathematik sind “Anzahl” und “Menge” zwei verschiedene Konzepte.

### 0.1.3 Organisation

Die hier vorgestellten Definitionen und Sätze sind (pro Kapitel) durchnummeriert. Sie sollten sich jedoch für Ihr eigenes Verständnis eher die Titel der Sätze oder die Inhalte der Definitionen merken, selbst wenn die Querverweise im Skript die Nummern verwenden. Der Sinn solcher Nummern hängt entscheidend davon ab, ob die Ressource auch allgemein verfügbar ist. Querverweise im Skript sind insofern unproblematisch – aber schon dort kann sich die Nummerierung ändern, wenn eine aktualisierte Fassung entsteht und z.B. Dinge hinzugefügt oder anders angeordnet werden. Wenn Sie vom “Satz des Pythagoras” sprechen, wird man Sie überall verstehen – bei “Satz 2.1.7” käme hingegen bestenfalls die Frage “aus welchem Buch denn?”.

Zu einigen Sätzen sind Beweise ausgeführt; diese finden sich im Anhang B. In der Prüfung wird von Ihnen nicht verlangt, solche Beweise aus dem Stegreif auszuführen. Trotzdem hier die Empfehlung, sich die dargestellten Beweise mindestens durchlesen. Ein Zusammenhang, der begründet wurde und also nachweisbar mathematisch sinnvoll ist, lässt sich deutlich leichter verwenden (oder überhaupt erst: merken), als wenn die Aussage des Satzes ohne Beweis akzeptiert und (für die Prüfung) auswendig gelernt wird.

Darüber hinaus sollten Sie bedenken, dass im späteren Berufsleben die interessanten Probleme eben nicht in Gestalt einer Prüfungsaufgabe daher kommen, die Sie direkt mit der passenden (vielleicht sogar vorgegebenen) Methode lösen können. Meist ist zunächst das mathematische Problem aus einem größeren Kontext zu extrahieren. Um solch ein mathematisches Modell zu erstellen, ist es sehr nützlich, über die wörtlichen Aussagen von Sätzen hinaus auch noch nützlich zu wissen, wann sich diese sinnvoll anwenden lassen. Das Nachvollziehen der Beweise kann hierbei enorm hilfreich sein – denn genau dort wird schlüssig argumentiert, warum unter gewissen Voraussetzungen bestimmte Sachverhalte (nämlich die in der Aussage eines Satzes ausgedrückten) gelten.

Ihr Ziel sollte sein, die in den Sätzen ausgedrückten Konzepte zumindest im Ansatz zu verstehen, damit Sie sie auch nach der Prüfung noch abrufen können. Natürlich erwartet niemand, dass Sie den kompletten Stoff dieser Vorlesung noch im sechsten Semester oder während einer Bachelor-/Masterthesis ohne Zögern abrufen und wiedergeben können. Aber Sie werden, je nach Laufbahn/Fachgebiet, durchaus in Situationen kommen, in denen Sie sich Zusammenhänge aus diesen Bereichen nochmal “hervor holen” müssen. Das gelingt umso besser und schneller, je solider Sie hier arbeiten.

## 0.2 Zum Studium

Der Umstieg von der weiterführenden Schule auf die Hochschule ist oft eine ziemliche Umstellung. Neben der privaten Neuorganisation betrifft das auch Ihr Lern- und Arbeitsverhalten.

Zuallererst: In den Vorlesungen gibt es keine Note für die mündliche Mitarbeit. Frontalunterricht ist, gerade an einer HAW, nicht angestrebt; und man wird also versuchen, Sie einzubinden – mindestens über die Frage, ob das gerade Behandelte verstanden wurde; aber durchaus auch, was Ihre Ideen für die nächsten Schritte zum Lösen eines Problems sind. Der Unterschied ist aber, dass Sie schweigen dürfen, ohne dass Sie das etwas kostet.

Zur Vorlesung gibt es auch in der Regel keine Hausaufgaben (anders sieht es in den Laboren und Übungen aus).

Hier sind nun Sie gefragt, sich beim Studieren selbst einzuschätzen. Kommen Sie momentan gut mit dem Stoff zurecht? Gibt es grundsätzliche Verständnisprobleme? Sind Sie vielleicht nur noch nicht dazu gekommen, die Inhalte der letzten Woche nachzuholen?

Es ergeben sich drei wichtige Aufgaben für Sie:



- Bleiben Sie zeitlich am Ball. Die Inhalte in Mathematik-Vorlesungen bauen in aller Regel aufeinander auf und lassen sich daher nicht beliebig aufschieben. Auch ist der Stoff zu umfangreich und komplex, als dass er sich ein oder zwei Tage vor der Prüfung schnell lernen lässt. Natürlich sind Sie durch den Studienplan mit vielen Veranstaltungen versorgt und haben nur endlich viel Zeit. Aber gerade die Mathematik (und verwandte Fächer) lassen sich kaum in mehrfacher Geschwindigkeit oder “auf Lücke” lernen.
- Beobachten Sie kontinuierlich, wie Sie den Stoff bewältigen. Je nach schulischem Hintergrund oder Talent im Bereich Mathematik werden einige von Ihnen anfangs kaum Schwierigkeiten verspüren. Dieser Zeitpunkt kommt aber für so gut wie alle Leute irgendwann; das liegt einfach an Stoffmenge und Geschwindigkeit.

Sobald Sie Schwierigkeiten beim Bewältigen des Stoffs haben, sollten Sie aktiv gegensteuern. Fragen Sie Ihre Studienkolleg(inn)en, in Ihrer Lerngruppe, und natürlich auch während oder nach der Vorlesung. Falls die Zeit nicht reicht, um das Problem live zu entschärfen, können Sie außerdem die Sprechstunde besuchen oder um einen gesonderten Termin bitten – leider kann während der Vorlesung aus Zeitgründen nicht immer jedes individuelle Problem gelöst werden.

Nutzen Sie bitte außerdem die Angebote des Lernzentrums Mathematik sowie die Übung zur Vorlesung.

Haken Sie aber am Besten sofort nach, wenn Sie den Eindruck haben, dass in den Folien oder dem Tafelanschrieb ein Fehler ist, oder wenn Sie eine bestimmte Formulierung zu unklar fanden. Meist sind Sie dann nicht die einzige Person mit diesem Problem – und solche Situationen lassen sich, Ihr Feedback voraus gesetzt, meist schnell einfangen.

- Rechnen Sie bitte damit, den Vorlesungsstoff nachzubereiten – mit dem Skript oder gerne auch mit Lehrbüchern<sup>2</sup> zum Thema. Die Geschwindigkeit des Stoff-Durchsatzes ist verglichen mit der Schulmathematik deutlich höher. Wegen des begrenzten Zeitkontingents für Vorlesungen lassen sich auch nicht mehr alle Konzepte im Dialog zusammen entwickeln.

Die Vorlesung soll definitiv mehr als ein bloßer Stichwortgeber sein – aber die Substanz müssen Sie (meistens im Nachhinein, je nach Motiviertheit natürlich auch gerne schon im Voraus) aktiv mit erzeugen.

Versuchen Sie dabei, nicht (nur) auf Beispielaufgaben (ggf. sogar mit verfügbarer Musterlösung) hin zu lernen. Gerade weil die mathematischen Konzepte der Vorlesung auch später noch benötigt werden (mindestens in diversen späteren Lehrveranstaltungen), lohnt es sich auf lange Sicht nicht, den Fokus ausschließlich auf die anstehende Prüfungsklausur zu legen – auch wenn dies verständlicherweise zunächst Ihr Hauptanliegen ist.

Idealerweise diskutieren Sie Lernstoff auch in einer Lerngruppe. Davon haben nicht nur die Leute etwas, die etwas erklärt bekommen, sondern auch die, die anderen etwas erklären (denn dieser Perspektivwechsel vertieft das eigene Verständnis der Sache nochmal).

Noch ein Wort zu Computeralgebrasystemen: Ja – die gibt es! Und sie sind teilweise extrem praktisch, um (gerade längliche) Rechnungen, wie sie z.B. bei Determinanten, inversen Matrizen oder Eigenvektoren auftreten, zu überprüfen. Dabei ist allerdings Ihre Disziplin gefragt – denn in der Prüfung müssen Sie ohne solche Programme auskommen. Außerdem gibt es die Gefahr, allzu “zielorientiert” zu arbeiten, wenn Sie die Lösung einer Aufgabe schon im Voraus haben ausrechnen lassen.

Außerdem gehört im mathematischen Bereich zum Lösen eines Problems meistens auch der Lösungsweg mit dazu, d.h. es sollte nachvollziehbar sein, dass eine Lösung erarbeitet (nicht erraten) wurde. Ausnahmen (z.B. Auffinden einzelner Nullstellen von Polynomen) bestätigen die Regel!

---

<sup>2</sup>Für unsere Zwecke eignet sich z.B. der Blick in die angelsächsische Literatur, oder deutschsprachige Werke “für Ingenieure” oder “für die Informatik” – dort sind die Erklärungen oft etwas ausführlicher, dafür aber weniger detailverliebt als in der Fachliteratur aus dem Universitätsumfeld. Auf Beweise wird dafür oft verzichtet.

# Kapitel 1

## Grundlagen

### 1.1 Aussagenlogik: ((To Be) or (not(To Be)))

#### 1.1.1 Elementare Begriffe

Wir beginnen mit einigen grundlegenden Definitionen, um zu beschreiben, in welcher Art wir in Zukunft über logische Aussagen sprechen wollen:

**Definition 1.1** (Zuweisung). *Falls ein Symbol  $x$  ein bestimmtes mathematisches Objekt (z.B. eine Zahl) repräsentieren soll, führen wir diese Zuweisung mit folgender Schreibweise ein:*

$$x := \dots$$

Links steht dabei der neue Name, der ab sofort bekannt ist und synonym mit dem Objekt auf der rechten Seite verwendet werden darf. Die Auslassungspunkte rechts stehen für einen *Ausdruck* – das kann ein mathematisches Objekt, oder auch eine Verknüpfung mehrerer Objekte sein. Der Operator “:=”, der das neue Symbol mit dem Ausdruck verknüpft, bedeutet soviel wie “sei (ab sofort) definiert als”.

**Beispiele:**

- Die Zuweisung

$$x := 7$$

macht die Zahl 7 ab sofort unter dem Namen  $x$  bekannt.

- Mit

$$y := 3 \cdot x + 12$$

wird der Name  $y$  mit dem (zusammen gesetzten) Ausdruck auf der rechten Seite belegt. Falls  $x$  schon bekannt ist (etwa nach der Zuweisung s.o.), lässt sich der Zahlenwert von  $y$  berechnen; hier würde er 33 betragen.

---

**Definition 1.2** (Gleichheit). *Falls zwei mathematische Objekte (oder allgemeiner: Ausdrücke)  $x$  und  $y$  gleich sind, so schreiben wir*

$$x = y$$

*Sind  $x$  und  $y$  hingegen nicht gleich, so nennen wir sie ungleich und würden schreiben:  $x \neq y$*

Die Kriterien für Gleichheit sind hierbei abhängig von den Objekttypen; letztere müssen zumindest in einer Art verträglich miteinander sein, dass es Sinn hat, von der Gleichheit von Objekten zu sprechen<sup>1</sup>.

---

<sup>1</sup>In Java wäre dies entweder das, was sich mit dem Operator `==` prüfen lässt, oder das Resultat einer `equals()`-Methode.

### Beispiele:

•

$$4 = 7 - \frac{12}{4}$$

Beide Ausdrücke links und rechts vom Gleichheitszeichen haben denselben Zahlenwert.

•

$$(\text{Hallo}) \neq 42$$

Eine Zeichenkette hat (jedenfalls ohne weiteren Kontext sicher) keinen Zahlenwert. Hier stimmt auch der mathematische Typ der Objekte nicht überein.

•

$$4 \cdot x = 7 \cdot y + 3$$

Wenn wir annehmen, dass  $x$  und  $y$  Zahlen sind, dann kann die Gleichheit der beiden Ausdrücke durchaus gegeben sein. Hier liegt eine *Gleichung* vor (mehr dazu in Abschnitt 1.4 s.u.). Zum Beispiel würden die Zahlen  $x = 6$  und  $y = 3$  der Gleichung genügen; der Zahlenwert wäre dann auf beiden Seiten 24. Auch richtig wäre die Kombination aus den Zahlen  $x = \frac{5}{2}$  und  $y = 1$ ; dann wäre der Zahlenwert auf beiden Seiten 10.

Dagegen wären für die Zahlen  $x = 3$  und  $y = 9$  die Zahlenwerte nicht übereinstimmen – in dem Fall liegt dann auch keine Gleichheit vor, da  $12 \neq 66$ .

---

Nach diesen sehr allgemeinen Definitionen wenden wir uns nun der Logik zu:

**Definition 1.3** (Logische Aussage). *Eine logische Aussage ist eine (ggf. mit mathematischer Notation) formulierte textuelle Aussage, die einen definierten Wahrheitswert besitzt, also immer entweder wahr oder falsch ist (dazwischen gibt es allerdings nichts!).*

*Textuell formulierte Aussagen schreiben wir in runden Klammern.*

*Die Wahrheitswerte seien mit den Symbolen  $\mathcal{W}$  und  $\mathcal{F}$  bezeichnet<sup>2</sup>.*

*Dabei heißt die spezielle Aussage  $\mathcal{W}$  die wahre Aussage;  $\mathcal{F}$  hingegen heißt die falsche Aussage.*

### Beispiele:

•

$$A := (\text{Franz ist ein Vorname})$$

$A$  ist eine wahre logische Aussage.

•

$$B := (\text{Franz ist ein Nachname})$$

Auch  $B$  ist eine wahre logische Aussage.

•

$$C := 4$$

$C$  ist *keine* logische Aussage. Das Symbol  $C$  hat einen Zahlenwert, aber keinen Wahrheitswert.

•

$$D := (4 < 7)$$

Die Aussage  $D$  ist logisch wahr.

•

$$E := (4 > 7)$$

Die Aussage  $E$  hingegen ist logisch falsch.

•

$$F := <$$

$F$  ist *keine* logische Aussage, sondern steht für ein anderes mathematisches Objekt – den “kleiner”-Operator.

---

<sup>2</sup>andere häufige Bezeichnungen sind  $t, f$  oder **true, false** oder 1,0. Letztere Notation ist besonders riskant, da sie Zahlenwerte suggeriert – um so genauer muss der Kontext bekannt sein!

•

$$G := (\text{Hallo!})$$

Es handelt sich bei  $G$  nicht um eine logische Aussage, da  $G$  kein Wahrheitswert zugeordnet werden kann.

**Definition 1.4** (Äquivalenz logischer Ausdrücke). *Zwei Ausdrücke  $A$  und  $B$  heißen (logisch) äquivalent, falls sie den gleichen Wahrheitswert besitzen. Man schreibt dann:*

$$A \equiv B \quad \text{oder} \quad A \Leftrightarrow B$$

**Bemerkungen:**

- Warum zwei verschiedene Symbole, um Äquivalenz auszudrücken? Die Behauptung, dass  $A$  und  $B$  den gleichen Wahrheitswert besitzen, ist ebenfalls eine logische Aussage – sie kann zutreffen oder nicht (und nichts dazwischen); wir notieren sie üblicherweise als  $(A \Leftrightarrow B)$ . Diese Aussage selbst hat nun wiederum einen Wahrheitswert, nämlich  $\mathcal{W}$ , falls  $A$  und  $B$  tatsächlich den gleichen Wahrheitswert hätten, oder  $\mathcal{F}$ , falls  $A$  und  $B$  verschiedene Wahrheitswerte hätten.

Daher kann es sinnvoll sein, zwei verschiedene Zeichen zum Ausdrücken der logischen Äquivalenz zur Verfügung zu haben. Wir wollen das Symbol  $\equiv$  für Aussagen hauptsächlich im aussagenlogischen Kontext verwenden, d.h. wenn wir uns primär mit den logischen Aspekten befassen. Später im Anwendungsfall (vor allem beim Lösen von Gleichungen oder Gleichungssystemen) wird dagegen meist  $\Leftrightarrow$  benutzt.

- Alle logisch wahren Aussagen sind logisch äquivalent zur wahren Aussage  $\mathcal{W}$ ; alle logisch falschen sind äquivalent zu  $\mathcal{F}$ .
- Man findet auch oft Formulierungen, in denen die logische Äquivalenz mit einem Gleichheitszeichen ausgedrückt wird. Man kann das durchaus so vereinbaren.

Für diese Vorlesung wollen wir jedoch darauf verzichten, und “=” für Vergleiche von Zahlenwerten (oder damit verwandten Objekten) reservieren, um Missverständnisse zu vermeiden. Eine Formulierung wie “ $x \Leftrightarrow 7$ ” ist nämlich definitiv inkorrekt, da 7 keinen Wahrheitswert besitzt. Hier müsste es “ $x = 7$ ” heißen.

**Beispiele:**

- Mit den oben eingeführten Aussagen gilt:

$$A \equiv B \equiv D \equiv \mathcal{W} \quad \text{und} \quad E \equiv \mathcal{F}$$

- Hier ein Beispiel für die Verwendung beider Äquivalenzzeichen:

$$(A \Leftrightarrow E) \equiv \mathcal{F}$$

Der geklammerte Ausdruck ist eine Aussage, nämlich, dass  $A$  und  $E$  (mit den Zuweisungen wie oben vereinbart) den gleichen Wahrheitswert besitzen. Das gilt jedoch nicht. Betrachtet man diese Aussage also insgesamt, so ist ihr Wahrheitswert  $\mathcal{F}$ .

Nach dieser Einführung wollen wir nun drei Operatoren definieren, mit denen sich aus gegebenen logischen Aussagen neue Aussagen erzeugen lassen:

**Definition 1.5** (Verknüpfung logischer Aussagen). *Seien  $A, B$  logische Aussagen<sup>3</sup>. Dann definieren wir die Konjunktion (“Und-Verknüpfung”) von  $A$  und  $B$  mit dem Operator  $\wedge$  per*

$$A \wedge B$$

*Diese Aussage hat genau dann den Wahrheitswert  $\mathcal{W}$ , wenn sowohl  $A$  als auch  $B$  den Wahrheitswert  $\mathcal{W}$  besitzen – andernfalls ist  $A \wedge B \equiv \mathcal{F}$ .*

<sup>3</sup>Wenn Symbole derart mit Kommata getrennt aufgelistet werden, soll sich das folgende immer auf jedes Element der Auflistung beziehen.

Die Disjunktion (“Oder-Verknüpfung”) von  $A$  und  $B$  wird mit dem Operator  $\vee$  erklärt. Die Aussage

$$A \vee B$$

ist genau dann logisch wahr, wenn mindestens einer der Operanden  $A, B$  logisch wahr ist.

Die Negation (“Verneinung” oder “Komplement”) von  $A$  ist die Aussage, die den gegenteiligen Wahrheitswert von  $A$  besitzt; wir schreiben sie als

$$\neg A$$

#### Bemerkungen:

- Man beachte, dass die Disjunktion nicht ausschließend (“entweder-oder”) gemeint ist – sie ist also auch dann logisch wahr, wenn beide Operanden logisch wahr sind. Beim “ausschließenden Oder” (**xor**) müssen dagegen die Operanden unterschiedliche Wahrheitswerte besitzen, damit die Verknüpfung als ganzes logisch wahr ist. Dazu folgt später noch eine Bemerkung im Unterabschnitt zur Äquivalenzverknüpfung (s.u.).
- Der Komplement-Operator bindet stärker als Konjunktion oder Disjunktion. Das ist eine Konvention analog zur bekannten Regel “Punktrechnung vor Strichrechnung”. Statt z.B.  $(A \wedge (\neg B))$  zu schreiben, wäre  $(A \wedge \neg B)$  ebenfalls zulässig.

#### Beispiele:

•

$$(4 < 7) \wedge (12 = 14 - 2) \equiv \mathcal{W} \wedge \mathcal{W} \equiv \mathcal{W}$$

•

$$(4 < 7) \wedge (4 > 7) \equiv \mathcal{W} \wedge \mathcal{F} \equiv \mathcal{F}$$

•

$$(4 < 7) \vee (4 > 7) \equiv \mathcal{W} \vee \mathcal{F} \equiv \mathcal{W}$$

•

$$(4 < 7) \wedge \neg(4 > 7) \equiv (4 < 7) \wedge (4 \leq 7) \equiv \mathcal{W} \wedge \mathcal{W} \equiv \mathcal{W}$$

•

$$(x < y) \vee (x \geq y) \equiv \mathcal{W}$$

Dies gilt für alle angeordneten Zahlen unabhängig von deren Wert, denn:

$$(x < y) \vee (x \geq y) \equiv (x < y) \vee \neg(x < y)$$

Wenn aber eine Aussage in Disjunktion zu ihrem eigenen Komplement genommen wird, ist diese Verknüpfung immer wahr. Für jede beliebige logische Aussage  $A$  gilt nämlich (s.u.)  $A \vee \neg A \equiv \mathcal{W}$ : Wenn  $A$  eine logische Aussage ist, muss sie wahr oder falsch sein – ist sie falsch, so ist aber genau ihr Komplement wahr. In beiden Fällen ist die Disjunktion also wahr.

---

Mit den Operatoren aus Definition 1.5 lassen sich logische Aussagen auf mannigfache Weise verknüpfen. Um solche zusammen gesetzten Aussagen miteinander vergleichen zu können, hilft die folgende

**Definition 1.6** (Wahrheitstafel). *Sei  $B$  eine Verknüpfung von  $n$  logischen Aussagen  $A_1, \dots, A_n$ . Dann ist die Wahrheitstafel von  $B$  eine Tabelle, die für sämtliche möglichen Belegungen der Symbole  $A_1, \dots, A_n$  mit entsprechenden Wahrheitswerten die jeweiligen Wahrheitswerte von  $B$  auflistet.*

*Falls  $C$  eine zweite Verknüpfung der  $n$  Aussagen ist, so sind  $B$  und  $C$  genau dann logisch äquivalent, falls ihre Wahrheitstafeln identisch sind<sup>4</sup>.*

---

<sup>4</sup>Bei gleicher Reihenfolge der Tabellenzeilen

### Bemerkungen:

- Oft ist es hilfreich, die Ausdrücke noch in Teilausdrücke zu zerlegen, die dann leichter umrechenbar sind.
- Bei  $n$  unabhängigen Aussagen  $A_1, \dots, A_n$  hat die Wahrheitstafel genau  $2^n$  Zeilen, die sämtliche Kombinationen von möglichen Wahrheitswerten der  $A_i$  enthalten.

### Beispiele:

- Wir betrachten zwei Teilaussagen  $A_1, A_2$  und geben verschiedene Wahrheitstafeln an. Im linken abgesetzten Bereich der Tabelle befinden sich die Teilaussagen. Dann folgen vier Wahrheitstafeln für die Verknüpfungen, die in Definition 1.5 eingeführt wurden, und, wiederum abgesetzt und unterhalb noch markiert, vier Wahrheitstafeln, die sich aus den Basisverknüpfungen ergeben und für die folgenden beiden Beispiele markiert sind.

$A_1$	$A_2$	$\neg A_1$	$\neg A_2$	$A_1 \wedge A_2$	$A_1 \vee A_2$	$\neg A_1 \wedge \neg A_2$	$\neg A_1 \vee \neg A_2$	$\neg(A_1 \wedge A_2)$	$\neg(A_1 \vee A_2)$
$\mathcal{W}$	$\mathcal{W}$	$\mathcal{F}$	$\mathcal{F}$	$\mathcal{W}$	$\mathcal{W}$	$\mathcal{F}$	$\mathcal{F}$	$\mathcal{F}$	$\mathcal{F}$
$\mathcal{W}$	$\mathcal{F}$	$\mathcal{F}$	$\mathcal{W}$	$\mathcal{F}$	$\mathcal{W}$	$\mathcal{F}$	$\mathcal{W}$	$\mathcal{W}$	$\mathcal{F}$
$\mathcal{F}$	$\mathcal{W}$	$\mathcal{W}$	$\mathcal{F}$	$\mathcal{F}$	$\mathcal{W}$	$\mathcal{F}$	$\mathcal{W}$	$\mathcal{W}$	$\mathcal{F}$
$\mathcal{F}$	$\mathcal{F}$	$\mathcal{W}$	$\mathcal{W}$	$\mathcal{F}$	$\mathcal{F}$	$\mathcal{W}$	$\mathcal{W}$	$\mathcal{W}$	$\mathcal{W}$
						$\underbrace{\hspace{1.5cm}}$	$\underbrace{\hspace{1.5cm}}$	$\underbrace{\hspace{1.5cm}}$	$\underbrace{\hspace{1.5cm}}$
						$D$	$B$	$C$	$E$

- Für  $B := \neg A_1 \vee \neg A_2$  und  $C := \neg(A_1 \wedge A_2)$  gilt  $B \equiv C$ . Beweis durch Wahrheitstafel siehe voriges Beispiel. Die Wahrheitstafeln für  $B$  und  $C$  sind identisch.
- Für  $D := \neg A_2 \wedge \neg A_1$  und  $E := \neg(A_1 \vee A_2)$  gilt  $D \equiv E$ . Analog siehe erstes Beispiel.

Mit Wahrheitstafeln zeigt man leicht (aber mit einiger Schreiarbeit) den folgenden

**Satz 1.7** (Rechenregeln für logische Aussagen). *Für  $A, B, C$  logische Aussagen gelten folgende Rechenregeln:*

Name	Variante 1	Variante 2
Kommutativgesetze	$A \wedge B \equiv B \wedge A$	$A \vee B \equiv B \vee A$
Distributivgesetze	$A \wedge (B \vee C) \equiv (A \wedge B) \vee (A \wedge C)$	$A \vee (B \wedge C) \equiv (A \vee B) \wedge (A \vee C)$
Neutralitätsgesetze	$A \wedge \mathcal{W} \equiv A$	$A \vee \mathcal{F} \equiv A$
Komplementärgesetze	$A \wedge (\neg A) \equiv \mathcal{F}$	$A \vee (\neg A) \equiv \mathcal{W}$
Assoziativgesetze	$A \wedge (B \wedge C) \equiv (A \wedge B) \wedge C$	$A \vee (B \vee C) \equiv (A \vee B) \vee C$
Idempotenzgesetze	$A \wedge A \equiv A$	$A \vee A \equiv A$
Eliminationsgesetze	$A \wedge \mathcal{F} \equiv \mathcal{F}$	$A \vee \mathcal{W} \equiv \mathcal{W}$
Absorptionsgesetze	$A \wedge (A \vee B) \equiv A$	$A \vee (A \wedge B) \equiv A$
Regeln von deMorgan	$\neg(A \wedge B) \equiv (\neg A) \vee (\neg B)$	$\neg(A \vee B) \equiv (\neg A) \wedge (\neg B)$
Doppelte Negation	$\neg(\neg A) \equiv A$	

### Bemerkungen:

- Die ersten vier Zeilen der Tabelle sind auch als Huntington-Axiome bekannt (benannt nach E. V. Huntington); die anderen Regeln lassen sich daraus mit diversen Umformungsschritten herleiten (siehe dazu z.B. [2]).
- Die Formeln in den zwei Varianten sind jeweils zueinander dual; man erhält sie durch Tauschen der beiden Verknüpfungsoperatoren und der konstanten Wahrheitswerte

### Beispiele:

- Das Neutralitätsgesetz (Variante 1) ergibt sich direkt aus Definition 1.5 (Verknüpfung logischer Aussagen): Die konjunktive Verknüpfung von  $A$  mit der wahren Aussage ist genau dann wahr, wenn  $A$  selbst wahr ist. Anderenfalls (also wenn  $A$  falsch ist), ist auch die Verknüpfung logisch falsch. In beiden Fällen reproduziert die Konjunktion also genau den Wahrheitswert von  $A$ .
- Die beiden Regeln von deMorgan wurden schon in den Beispielen zu Definition 1.6 mit Wahrheitstafeln gezeigt.

### 1.1.2 Prädikate

**Definition 1.8** (Prädikat). Eine logische Aussage  $A$ , deren Wahrheitswert von  $n$  äußeren Parametern  $x_1, x_2, \dots, x_n$  abhängt, heißt  $n$ -stelliges Prädikat. Wir schreiben:

$$A(x_1, x_2, \dots, x_n)$$

Ein Prädikat, das für sämtliche möglichen Kombinationen von Parameter-Werten stets logisch wahr ist, heißt Tautologie. Eines dass stets logisch falsch ist, heißt Kontradiktion.

**Bemerkung:** Die Parameterliste von  $A$  ist hierbei *geordnet*, d.h. die Reihenfolge der einzelnen Parameter ist stets dieselbe.

**Beispiele:**

- Die beiden Verknüpfungen  $B$  und  $C$  aus dem obigen Beispiel zur Definition 1.6 (Wahrheitstafel) sind jeweils zweistellige ("binäre") Prädikate mit den Parametern  $A_1$  und  $A_2$ . Hier exemplarisch als Prädikate notiert:

$$\begin{aligned} B(A_1, A_2) &\equiv ((\neg A_1) \vee (\neg A_2)) \\ C(A_1, A_2) &\equiv \neg(A_1 \wedge A_2) \end{aligned}$$

- Mit  $B$  und  $C$  wie oben eingesetzt, ist das zweistellige Prädikat  $d$  mit

$$d(A_1, A_2) := (((\neg A_1) \vee (\neg A_2)) \Leftrightarrow \neg(A_1 \wedge A_2))$$

eine Tautologie. Im erwähnten Beispiel wurde dies schon nachgerechnet.

- Die Rechenregeln in Satz 1.7 (Rechenregeln für logische Aussagen) lassen sich jeweils als ein- bis dreistellige Prädikate verstehen und sind allesamt Tautologien<sup>5</sup>.
- Die Parameter eines Prädikats müssen natürlich keine logischen Aussagen sein. Ein binäres Prädikat wäre z.B. auch:

$$a(x, y) := (x < y)$$

Hier hängt der Wahrheitswert von  $a$  nun von konkreten Werten für  $x, y$  ab – wobei wir voraussetzen, dass diese Parameter Zahlen oder andere mathematische Objekte sind, die sich in der Größe vergleichen lassen. Es gilt z.B.:

- $a(3, 7) \equiv \mathcal{W}$
  - $a(4, 4) \equiv \mathcal{F}$
  - $a(\frac{3}{4}, \frac{12}{16}) \equiv \mathcal{F}$
  - $a(\frac{3}{4}, \frac{121}{160}) \equiv \mathcal{W}$
- Mit  $a$  wie eben können wir ein einstelliges ("unäres") Prädikat definieren per

$$b(x) := a(x, 13) \equiv (x < 13)$$

Im Vergleich zu  $a$  ist nun einer der beiden Parameter von  $a$  fest gehalten. Dann wäre z.B.:

- $b(7) \equiv \mathcal{W}$
  - $b(13.42) \equiv \mathcal{F}$

(Wir schreiben Kommazahlen hier mit Dezimalpunkt)
- Statt der symbolischen Namen sind manchmal ausführlichere Bezeichner für Prädikate sinnvoll, z.B. für natürliche Zahlen  $n$  die folgende Fallunterscheidung:

$$\text{istGerade}(n) := \begin{cases} \mathcal{W}, & \text{falls es eine natürliche Zahl } k \text{ gibt, sodass } n = 2 \cdot k \\ \mathcal{F}, & \text{sonst} \end{cases}$$

Und damit auch:

$$\text{istUngerade}(n) := \neg \text{istGerade}(n)$$

<sup>5</sup>Das sollte für allgemeine Rechenregeln auch so sein – diese sollen schließlich immer gelten.

### 1.1.3 Prädikatenlogik

**Definition 1.9** (Prädikatenlogische Aussagen). *Falls für sämtliche möglichen Objekte  $x$ , für die ein Prädikat  $a(x)$  definiert ist, gilt, dass  $a(x) \equiv \mathcal{W}$  erfüllt ist, drücken wir dies als (in Prädikatenlogik formulierte) allgemeine Aussage aus und schreiben:*

$$\forall x : a(x)$$

*Das Symbol  $\forall$  heißt Allquantor. Gelesen: “Für alle  $x$  gilt  $a(x)$ ”.*

*Ist hingegen das Prädikat  $a$  für alle  $x$  logisch falsch, so notieren wir die Aussage als*

$$\forall x : \neg a(x)$$

*Gilt  $a(x) \equiv \mathcal{W}$  für mindestens ein  $x$  (möglicherweise für mehrere), so schreiben wir eine (in Prädikatenlogik formulierte) Existenzaussage per*

$$\exists x : a(x)$$

*Das Symbol  $\exists$  heißt Existenzquantor. Gelesen: “Es existiert (mindestens) ein  $x$ , für das  $a(x)$  gilt”.*

*Oder, falls es ein  $x$  gibt, für das  $a(x)$  gerade nicht erfüllt ist:*

$$\exists x : \neg a(x)$$

**Bemerkung:** Wir lernen erst in nächsten Abschnitt den Gebrauch von Mengen (d.h. Ansammlungen von Objekten) kennen<sup>6</sup>. Meist werden prädikatenlogische Aussagen nicht für sämtliche (mathematischen?) Objekte  $x$ , sondern eher für Objekte aus bestimmten Mengen formuliert. Dann gehört in obiger Definition zwischen  $x$  und den Doppelpunkt noch die Spezifikation, von welcher Menge die Objekte  $x$  gezogen werden. Die folgenden Beispiele lassen sich mit Mengen eleganter formulieren.

**Beispiele:**

- Die allgemeine Aussage

$$A := \left( \forall n : ((n \text{ ist keine natürliche Zahl}) \vee (5 \cdot n > n)) \right)$$

ist in Prädikatenlogik formuliert, und logisch wahr, also  $A \equiv \mathcal{W}$ . Für sämtliche  $n$  gilt nämlich:

- Falls  $n$  gar keine natürliche Zahl ist, dann trifft der linke Teil der Disjunktion zu, und nach dem Eliminationsgesetz (siehe Satz 1.7) hat die ganze Disjunktion den Wahrheitswert  $\mathcal{W}$ .
- Falls  $n$  jedoch eine natürliche Zahl ist, trifft der linke Teil der Disjunktion nicht zu. Nach dem Neutralitätsgesetz (ebenda) entspricht der Wahrheitswert der Disjunktion dann dem Wahrheitswert ihres rechten Operanden. Aber für alle natürlichen Zahlen (die Null gehört bei uns *nicht* dazu!) ist die Aussage  $(5 \cdot n > n)$  richtig.

Wir beobachten hier außerdem noch folgendes in Bezug auf die Parameter:

- $n$  ist *kein* äußerer Parameter der Aussage  $A$ , d.h.  $A$  ist kein Prädikat mit Parameter  $n$ , sondern ein null-stelliges Prädikat – eine Aussage ohne Parameter.
- $n$  ist jedoch ein *innerer* Parameter von  $A$ , d.h. dieser Name gilt nur innerhalb der Aussage  $A$  und wird dort neu eingeführt – nämlich dadurch, dass über  $n$  eine prädikatenlogische Aussage formuliert wird.
- Ab dem Doppelpunkt in obiger prädikatenlogischer Aussage ist  $n$  jedoch bekannt (und jeweils als fest anzusehen).
- Der rechte Teil der prädikatenlogischen Aussage von oben kann als Prädikat  $p(n)$  ausgedrückt werden:

$$p(n) := ((n \text{ ist keine natürliche Zahl}) \vee (5 \cdot n > n))$$

Dann lautet  $A$  wie folgt:

$$A \equiv (\forall n : p(n))$$

---

<sup>6</sup>Was die Reihenfolge der Einführung von Aussagenlogik und Mengenlehre angeht, besteht ein gewisses Henne-Ei-Problem. Dazu siehe die Bemerkungen im Abschnitt zur Mengenlehre.



- Bezogen auf  $p(n)$  ist  $n$  hier also ein äußerer Parameter. Innerhalb von  $p$  ist  $n$  nämlich bekannt und darf auch nicht überschrieben werden.  
Allgemein gilt für alle prädikatenlogisch formulierten Aussagen: mit Symbolen benannte Objekte sind entweder äußere Parameter (wurden also irgendwo weiter links eingeführt) oder innere Parameter, welche links vom Doppelpunkt eingeführt und rechts davon mit einem Prädikat beschrieben werden. Ein Symbol darf, bezogen auf die gleiche Hierarchieebene einer Aussage nicht gleichzeitig äußerer und innerer Parameter sein.
- Auch Prädikate dürfen in Prädikatenlogik formulierte Aussagen enthalten. Zum Beispiel:

$$b(n) := (n \text{ ist keine natürliche Zahl}) \vee \left( \exists k : ((k \text{ ist keine natürliche Zahl}) \vee (k > n)) \right)$$

Hier wollen wir einige neue Namen einführen, um die Parameter genauer unterscheiden zu können. Es sei also:

$$\begin{aligned} q(n) &:= (n \text{ ist keine natürliche Zahl}) \\ r(n) &:= (\exists k : s(n, k)) \\ s(n, k) &:= (t(n, k) \vee u(n, k)) \\ t(n, k) &:= (k \text{ ist keine natürliche Zahl}) \\ u(n, k) &:= (k > n) \end{aligned}$$

Wir stellen also fest:

- Das etwas komplexe Prädikat  $b(n)$  lässt sich in einfachere Teilprädikate zerlegen. Genauer:

$$b(n) \equiv (q(n) \vee r(n))$$

Die Variable  $n$  ist hier für alle drei auftretenden Prädikate ein äußerer Parameter.

- Das Prädikat  $r(n)$  ist in Prädikatenlogik formuliert. Innerhalb wird ein innerer Parameter  $k$  eingeführt – das  $k$ , auf das sich die Existenzaussage bezieht.
- Innerhalb von  $r(n)$  und nach dem Doppelpunkt, der dort den Parameter  $k$  eingeführt hat, sind nun  $n$  und  $k$  bekannt. Das Prädikat hinter dem Doppelpunkt,  $s(n, k)$  darf daher auch auf beide Objekte  $n$  und  $k$  Bezug nehmen.
- $s(n, k)$  lässt sich als zusammen gesetztes Prädikat noch in zwei einfachere Teilprädikate zerlegen.
- $t(n, k)$  nimmt keinen Bezug auf  $n$ , dürfte dies aber in einem anders gelagerten Beispiel durchaus tun. (Schlecht wäre hier jedoch, wenn  $t(n, k)$  Bezug nähme auf eine Variable, die dort gar nicht bekannt ist, z.B. ein  $x$ , das nirgendwo links davon eingeführt wurde.)
- $u(n, k)$  ist ein binäres Prädikat, das Zahlenwerte (oder anderweitig auf Zahlenwerte abbildbare Größenverhältnisse) vergleicht. Es verwendet beide hier bekannten Parameter, und ordnet jeder Parameterkombination eindeutig einen Wahrheitswert zu.

Was ist nun die Aussage des Prädikats  $b(n)$ ? Das Prädikat ist schon dann wahr, wenn  $n$  keine natürliche Zahl ist (dann wäre nämlich  $q(n) \equiv \mathcal{W}$ ). Falls jedoch  $n$  eine natürliche Zahl ist, müsste, damit  $b(n)$  gilt, der rechte Teil der Disjunktion, also  $r(n)$  erfüllt sein. In  $r(n)$  wird behauptet, dass es ein  $k$  gibt (mindestens eins, wenn nicht mehr), das  $(t(n, k))$  keine natürliche Zahl sein könnte – das aber, wenn es eine natürliche Zahl ist, größer ist als  $n$ .

Und das gilt! Etwas kompakter ausgedrückt (und mit leichtem Vorgriff auf die Mengenlehre) lässt sich nämlich das obige umschreiben als: “Falls  $n$  eine natürliche Zahl ist, so gibt es eine natürliche Zahl  $k$ , welche größer ist als  $n$ .” Solch ein  $k$  findet man immer; z.B. wäre schon  $(n + 1)$  ein geeigneter Wert für  $k$ .

Wir kommen auf diese beiden Beispiele nach der “offiziellen” Einführung der natürlichen Zahlen nochmal zurück und formulieren sie dann entsprechend um.

- Das Prädikat **istGerade**( $n$ ) für natürliche Zahlen aus dem Beispiel zu Definition 1.8 lässt sich nun schreiben als

$$\text{istGerade}(n) \equiv (\exists k : (k \text{ ist natürliche Zahl}) \wedge (n = 2 \cdot k))$$

Wir befassen uns nun noch mit der Verneinung von prädikatenlogischen Aussagen:

**Satz 1.10** (Verneinung prädikatenlogischer Aussagen). *Sei  $a(x)$  ein unäres Prädikat, und die Aussagen  $A, B$  gegeben per*

$$A := (\forall x : a(x)) \quad \text{und} \quad B := (\exists x : a(x))$$

Dann gilt:

$$\neg A \equiv (\exists x : \neg a(x)) \quad \text{und} \quad \neg B \equiv (\forall x : \neg a(x))$$

(Beweis auf Seite 311)

**Bemerkung:** Wie oben gesehen, lassen sich komplexere Prädikataussagen in einfachere Teilaussagen zerlegen. Mit dem obigen Satz können damit auch die Verneinungen von geschachtelten prädikatenlogischen Aussagen berechnet werden. Das Schema bleibt immer das gleiche: Quantoren tauschen ( $\forall$  in  $\exists$  und umgekehrt) und die Prädikate rechts vom Doppelpunkt durch ihre Komplemente (Verneinungen) ersetzen.

**Beispiele:**

- Wir betrachten zunächst die Aussage  $A$  aus dem Beispiel zur Definition 1.9:

$$A := (\forall n : ((n \text{ ist keine natürliche Zahl}) \vee (5n > n)))$$

Dann bilden wir das Komplement von  $A$ ; dazu benutzen wir außerdem die Regel von deMorgan:

$$\begin{aligned} \neg A &\equiv \neg(\forall n : ((n \text{ ist keine natürliche Zahl}) \vee (5n > n))) \\ &\equiv \exists n : \neg((n \text{ ist keine natürliche Zahl}) \vee (5n > n)) \\ &\equiv \exists n : \neg(n \text{ ist keine natürliche Zahl}) \wedge \neg(5n > n) \\ &\equiv \exists n : (n \text{ ist eine natürliche Zahl}) \wedge (5n \leq n) \end{aligned}$$

Im letzten Schritt haben wir noch die Negationen der beiden Teilaussagen eingesetzt.

Da  $A$  besagt, dass alle natürlichen Zahlen die Eigenschaft haben, dass ihre Fünffachen größer sind als sie selbst (was stimmt!), lautet die negierte Aussage: “Es gibt eine natürliche Zahl, deren Fünffaches höchstens so groß ist wie sie selbst” (was natürlich nicht korrekt ist).

- Auch das Prädikat  $b(n)$  aus dem anderen Beispiel zu Definition 1.9 wollen wir verneinen:

$$b(n) := (n \text{ ist keine natürliche Zahl}) \vee (\exists k : ((k \text{ ist keine natürliche Zahl}) \vee (k > n)))$$

Hierzu benutzen wir die Aufspaltung in Teilprädikate von oben, beginnend mit

$$b(n) \equiv (q(n) \vee r(n))$$

Nun verneinen wir schrittweise (“top-down”, also von außen nach innen) die Kaskade einfacher Teilprädikate (wieder bei Bedarf unter Benutzung von deMorgan):

$$\begin{aligned} \neg b(n) &\equiv \neg q(n) \wedge \neg r(n) \\ \neg q(n) &\equiv \neg(n \text{ ist keine natürliche Zahl}) \equiv (n \text{ ist eine natürliche Zahl}) \\ \neg r(n) &\equiv \neg(\exists k : s(n, k)) \equiv \forall k : \neg s(n, k) \\ \neg s(n, k) &\equiv \neg(t(n, k) \vee u(n, k)) \equiv \neg t(n, k) \wedge \neg u(n, k) \\ \neg t(n, k) &\equiv \neg(k \text{ ist keine natürliche Zahl}) \equiv (k \text{ ist eine natürliche Zahl}) \\ \neg u(n, k) &\equiv \neg(k > n) \equiv k \leq n \end{aligned}$$

Nun setzen wir die Zwischenergebnisse wieder zusammen und erhalten:

$$\neg b(n) \equiv (n \text{ ist eine natürliche Zahl}) \wedge (\forall k : ((k \text{ ist eine natürliche Zahl}) \wedge (k \leq n)))$$

Dieses Resultat ließe sich mit etwas Übung auch direkt aus  $b(n)$  erhalten, aber zum Trainieren empfiehlt sich dringend die Aufspaltung in Teilprädikate (die sich dafür dann jeweils leicht negieren lassen). Insgesamt besagt die Verneinung von  $b(n)$  also, dass für eine beliebige natürliche Zahl  $n$  (denn  $n$  ist ja ein Parameter) sämtliche natürlichen Zahlen  $k$  höchstens so groß sind wie  $n$ . (Das stimmt natürlich nicht, denn schon die Zahl  $(n + 1)$  wäre dann eine natürliche Zahl, die größer ist als  $n$ .)

- Für die geschachtelte prädikatenlogische Aussage

$$A := (\forall x : \exists y : q(x, y))$$

(kein Bezug zum  $A$  aus dem ersten Beispiel; hier: “Für sämtliche  $x$  gibt es jeweils (mindestens) ein  $y$ , sodass  $q(x, y)$  erfüllt ist”) führen wir zunächst ein Prädikat  $p(x)$  ein, das die innere Aussage repräsentiert:

$$p(x) := (\exists y : q(x, y))$$

Dann verneinen wir kaskadiert von außen nach innen. Wegen

$$A \equiv (\forall x : p(x))$$

gilt dann nach dem obigen Satz:

$$\neg A \equiv (\exists x : \neg p(x))$$

Und für die Verneinung des Zwischenprädikats, wiederum nach dem Satz:

$$\neg p(x) \equiv (\forall y : \neg q(x, y))$$

Also insgesamt, die Zwischenergebnisse eingesetzt und zusammen gefasst:

$$\neg A \equiv (\exists x : \forall y : \neg q(x, y))$$

Die Verneinung lautet also: “Es gibt (mindestens) ein  $x$ , für welches mit sämtlichen  $y$   $q(x, y)$  nicht erfüllt ist”, oder gleichwertig “Es gibt (mindestens) ein  $x$ , für welches mit keinem  $y$  die Eigenschaft  $q(x, y)$  erfüllt ist”.

**Bemerkung:** Das eingeklammerte Wort “mindestens” in Existenzaussagen wird später meist weggelassen (aber immer mit gedacht!). Hier soll damit betont werden, dass eben nicht “Es gibt *genau ein* ...” gemeint ist. Falls mehr als ein Objekt der Existenzaussage genügt, ist das kein Widerspruch zur Aussage.

### 1.1.4 Implikation, Äquivalenz

Die Implikation ist eine der wichtigsten Beziehungen zwischen logischen Aussagen, da sie das logische Schließen (*Deduktion*) erlaubt – ist eine bestimmter Sachverhalt gegeben, so folgt daraus, dass ein weiterer Sachverhalt gegeben ist.

**Definition 1.11** (Implikation). Für zwei Aussagen  $A, B$  wird die Verknüpfung

$$\neg A \vee B$$

als Implikation bezeichnet, auch geschrieben als

$$A \Rightarrow B$$

Dabei heißt  $A$  Prämisse und  $B$  Konklusion. Gelesen: “ $A$  impliziert  $B$ ” oder “Aus  $A$  folgt  $B$ ” oder “Wenn  $A$  gilt, dann gilt auch  $B$ ”.

**Bemerkung:** Die Wahrheitstafel der Implikation ist damit (Verknüpfung von oben eingesetzt):

$A$	$B$	$A \Rightarrow B$
$\mathcal{W}$	$\mathcal{W}$	$\mathcal{W}$
$\mathcal{W}$	$\mathcal{F}$	$\mathcal{F}$
$\mathcal{F}$	$\mathcal{W}$	$\mathcal{W}$
$\mathcal{F}$	$\mathcal{F}$	$\mathcal{W}$

Insbesondere ist die Implikation schon dann erfüllt, wenn die Prämisse logisch falsch ist!

## Beispiele:

- Mit dem Prädikat  $\text{istGerade}(n)$  für natürliche Zahlen (und seinem Komplement) können wir formulieren:

$$\text{istGerade}(n) \Rightarrow \text{istUngerade}(n+1)$$

Man macht sich leicht klar, dass die gleichwertige Aussage “Jeder Nachfolger einer geraden natürlichen Zahl ist ungerade” stimmt; z.B. ist 4 gerade, und der Nachfolger 5 ist ungerade.

Nicht möglich ist dagegen, dass der Nachfolger einer geraden Zahl *nicht* ungerade ist (also: erfüllte Prämisse, aber nicht erfüllte Konklusion). Denn dann wäre die “wenn-dann”-Beziehung verletzt, die in der Implikation ausgedrückt ist. Daher hat die Implikation in diesem Fall den Wahrheitswert  $\mathcal{F}$ .

Und falls  $n$  gar keine gerade Zahl ist? Dann könnte  $n$  z.B. eine ungerade Zahl sein, und der Nachfolger  $(n+1)$  wäre gerade – aber das ist kein Widerspruch zur obigen Implikation; daher hat die (hier relevante) unterste Zeile der obigen Wahrheitstafel den Eintrag  $\mathcal{W}$ .

Oder  $n$  könnte außerhalb der natürlichen Zahlen liegen, z.B.  $n = 0$ . Der Nachfolger von  $n$  wäre damit 1 und ungerade. Auch hier liegt kein Widerspruch zur behaupteten Implikationsbeziehung vor, daher hat auch die vorletzte Zeile der obigen Wahrheitstafel den Eintrag  $\mathcal{W}$ .

- Wir betrachten folgende Aussage: “Jedes von Donalds Häusern hat mindestens zwei Golfplätze”. Wir drücken dies mit (zugegeben sehr spezifischen) Prädikaten aus:

$$\text{istDonaldsHaus}(x) \Rightarrow \text{hatMindestensZweiGolfplätze}(x)$$

Wir wollen annehmen, dass die behauptete Aussage stimmt, und betrachten nochmal die vier möglichen Fälle, die als Einträge in der Wahrheitstafel auftauchen:

- $x$  ist ein Haus von Donald, d.h. die Prämisse ist  $\mathcal{W}$ . Da die Behauptung ja stimmt, hat das Haus  $x$  auch mindestens zwei Golfplätze, z.B. sieben Stück.
- $x$  ist ein Haus von Donald, aber hat bloß einen Golfplatz (oder am Ende – wie peinlich ist das! – gar keinen). In dem Fall ist, trotz erfüllter Prämisse, die Konklusion nicht erfüllt, *im Widerspruch* zur behaupteten Implikation – daher hier der Wahrheitswert  $\mathcal{F}$ .
- $x$  ist keins von Donalds Häusern, sondern vielleicht das Haus von Joe, d.h. die Prämisse ist  $\mathcal{F}$ . Trotzdem wäre es möglich, dass die Konklusion (für sich genommen) erfüllt ist, falls Joes Haus zwei (oder mehr) Golfplätze hat. Da die Behauptung sich aber nur auf Häuser von Donald bezieht, ist dies kein Widerspruch – daher Wahrheitswert  $\mathcal{W}$  für die Implikation.
- $x$  ist keins von Donalds Häusern, und hat auch keine zwei Golfplätze – wieder kein Widerspruch zur Behauptung, also Wahrheitswert  $\mathcal{W}$  für die Implikation.

Dieser Fall gilt für alle Häuser, die nicht Donald gehören und höchstens einen Golfplatz haben. Aber außerdem auch noch für alle  $x$ , die gar keine Häuser sind, z.B. für alle natürlichen Zahlen  $x$ , oder alle Kürbisse  $x$ . Diese Eventualität müssen wir hier berücksichtigen, weil wir den Mengenbegriff noch nicht benutzen können, der  $x$  z.B. nur auf Häuser einschränken könnte.

*Die Konklusion einer Implikationsbeziehung kann also unabhängig von der Prämisse zutreffen oder nicht – das einzige, was ausgeschlossen ist, wäre, dass sie unerfüllt ist, obwohl die Prämisse gilt.*

---

Mit Wahrheitstafeln zeigt man schnell die folgende Tatsache:

**Satz 1.12** (Kontrapositorische Formulierung der Implikation). *Seien  $A, B$  logische Aussagen. Dann gilt:*

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

Mit anderen Worten: Wenn die Implikation  $A \Rightarrow B$  erfüllt ist, und  $B$  ist für sich genommen nicht erfüllt, dann kann auch  $A$  nicht erfüllt sein. (Denn wenn  $A$  erfüllt wäre, *müsste* nach der Implikation ja  $B$  gelten!)

### Bemerkungen:

- Daher spricht man bei  $A \Rightarrow B$  auch davon, dass  $A$  “hinreichend für  $B$ ” ist, und  $B$  “notwendig für  $A$ ”. Diese Notwendigkeit äußert sich eben darin, dass  $A$  ohne  $B$  nicht erfüllt sein kann.
- Man beachte die Asymmetrie:  $B$  darf nämlich durchaus ohne  $A$  erfüllt sein (dritte Zeile der obigen Wahrheitstafel).

**Beispiel:** Ein Haus, zu dem nicht wenigstens zwei Golfplätze gehören, kann (nach der Annahme im Beispiel zu Definition 1.11) unmöglich eins von Donalds Häusern sein.

---

Falls jedoch die Implikationsbeziehung in beiden Richtungen gilt, gelangen wir zu einer neuen Beziehung:

**Definition 1.13** (Äquivalenz von Aussagen). *Zwei Aussagen  $A, B$ , die sich gegenseitig implizieren, heißen äquivalent; geschrieben als*

$$(A \Leftrightarrow B) := ((A \Rightarrow B) \wedge (B \Rightarrow A))$$

*Gelesen: “Genau dann, wenn  $A$  gilt, gilt auch  $B$ ” oder “ $A$  gilt dann und nur dann, wenn  $B$  gilt” oder “ $A$  ist notwendig und hinreichend für  $B$ ”.*

### Bemerkungen:

- Statt  $B \Rightarrow A$  dürfen wir synonym schreiben:  $A \Leftarrow B$ . So sieht man leichter, dass die Implikationsbeziehung hier in beiden Richtungen gilt.
- Wir vervollständigen die Wahrheitstafel der Implikation von oben:

$A$	$B$	$A \Rightarrow B$	$A \Leftarrow B$	$A \Leftrightarrow B$
$\mathcal{W}$	$\mathcal{W}$	$\mathcal{W}$	$\mathcal{W}$	$\mathcal{W}$
$\mathcal{W}$	$\mathcal{F}$	$\mathcal{F}$	$\mathcal{W}$	$\mathcal{F}$
$\mathcal{F}$	$\mathcal{W}$	$\mathcal{W}$	$\mathcal{F}$	$\mathcal{F}$
$\mathcal{F}$	$\mathcal{F}$	$\mathcal{W}$	$\mathcal{W}$	$\mathcal{W}$

- Äquivalenz liegt also (analog zur Definition 1.4 (Äquivalenz logischer Ausdrücke) gerade dann vor, wenn  $A$  und  $B$  jeweils den gleichen Wahrheitswert besitzen. Hier kann es also nicht sein, dass  $B$  unabhängig von  $A$  erfüllt ist (oder umgekehrt).
- Wir können mit den Formeln aus Satz 1.7 (Rechenregeln für logische Aussagen) die folgende Kette von logischen Äquivalenzen aufstellen (die aber hauptsächlich mit Blick auf die Vorlesung zur Technischen Informatik interessant ist):

$$\begin{aligned}(A \Leftrightarrow B) &\equiv (A \Rightarrow B) \wedge (B \Rightarrow A) \\ &\equiv (\neg A \vee B) \wedge (\neg B \vee A) \\ &\equiv ((\neg A \vee B) \wedge \neg B) \vee ((\neg A \vee B) \wedge A) \\ &\equiv ((\neg A \wedge \neg B) \vee (B \wedge \neg B)) \vee ((\neg A \wedge A) \vee (B \wedge A)) \\ &\equiv (\neg A \wedge \neg B) \vee (B \wedge \neg B) \vee (\neg A \wedge A) \vee (B \wedge A) \\ &\equiv (\neg A \wedge \neg B) \vee \mathcal{F} \vee \mathcal{F} \vee (B \wedge A) \\ &\equiv (\neg A \wedge \neg B) \vee (B \wedge A) \\ &\equiv (A \wedge B) \vee (\neg A \wedge \neg B)\end{aligned}$$

Hierbei haben wir zunächst die Definition der beiden Äquivalenzbeziehungen eingesetzt. In den nächsten Schritten wurde zweimal das Distributivgesetz benutzt. Dann das Assoziativgesetz – denn dieses erlaubt es, auf die Klammern zu verzichten, wenn auf gleicher Hierarchieebene nur Verknüpfungen des gleichen Typs aneinander gereiht sind (hier: Disjunktionen)<sup>7</sup>. Mit dem Komplementärgesetz werden nun zwei der einzelnen Unterausdrücke zu  $\mathcal{F}$  ausgewertet, und im nächsten Schritt, da sie disjunktiv verknüpft sind, mit dem Neutralitätsgesetz

---

<sup>7</sup>Genau genommen erlaubt das Assoziativgesetz dann beliebige Klammerungen, aber das läuft darauf hinaus, dass man die Klammern dieser Hierarchieebene auch weglassen kann.

entfernt. Mit dem Kommutativgesetz kann man dann noch (kosmetisch) umstellen und erhält die letzte Zeile.

Insgesamt also:

$$(A \Leftrightarrow B) \equiv (A \wedge B) \vee (\neg A \wedge \neg B)$$

Die rechte Schreibweise<sup>8</sup> gibt alle Kombinationen wieder, für die die Wahrheitstafel der Äquivalenz den Wert  $\mathcal{W}$  annimmt.

---

Das Gegenteil der Äquivalenz ist auch als **xor**-Verknüpfung bekannt:

**Definition 1.14** (Antivalenz). *Zwei Aussagen  $A, B$  heißen antivalent genau dann, wenn sie nicht äquivalent zueinander sind; geschrieben als*

$$(A \nleftrightarrow B) := \neg(A \Leftrightarrow B)$$

**Bemerkungen:**

- In diesem Fall gilt, analog zu oben:

$$(A \nleftrightarrow B) \equiv (A \wedge \neg B) \vee (\neg A \wedge B)$$

- Dies ist genau das *ausschließende Oder*; damit dieses erfüllt ist, muss genau einer der beiden verknüpften Ausdrücke wahr sein, der andere falsch.

### 1.1.5 Indirekte Beweise

Der *indirekte Beweis* einer Aussage  $A$  besteht darin, das Gegenteil von  $A$  anzunehmen und daraus einen Widerspruch zu folgern (über eine Implikationskette). Man zeigt dadurch:

$$(\neg A \Rightarrow \dots \Rightarrow \mathcal{F})$$

Nun kann man diese Kette von Implikationen auch kontrapositorisch lesen (siehe Satz 1.12); dann erhält man:

$$\begin{aligned} (\neg A \Rightarrow \dots \Rightarrow \mathcal{F}) &\equiv (\neg \mathcal{F} \Rightarrow \dots \Rightarrow \neg \neg A) \\ &\equiv (\mathcal{W} \Rightarrow \dots \Rightarrow A) \end{aligned}$$

Folgt also aus dem Gegenteil von  $A$  die falsche Aussage  $\mathcal{F}$ , so bedeutet das, dass  $A$  aus der wahren Aussage  $\mathcal{W}$  folgt. Aber alles, was sich aus der Wahrheit folgern lässt, muss wahr sein – somit ist  $A$  bewiesen!

**Beispiele:**

- Wir betrachten nochmal die Aussage “Jede gerade natürliche Zahl hat einen ungeraden Nachfolger” (die wir uns hier nicht als Implikation denken wollen, sondern als allgemeine prädikatenlogische Aussage interpretieren). Hier ein beispielhafter Beweistext:

Angenommen, die Aussage träfe nicht zu – dann gibt es eine gerade natürliche Zahl, die einen geraden Nachfolger hat.

Sei  $x$  solche eine Zahl. Da  $x$  gerade ist, gibt es eine natürliche Zahl  $k$ , für die gilt:

$$x = 2 \cdot k$$

Der Nachfolger dieser Zahl ist dann:

$$(x + 1) = 2 \cdot k + 1$$

Dieser Nachfolger ist aber ungerade, d.h. nicht durch 2 teilbar, im Widerspruch zur Annahme  $\nmid$

Also muss die Annahme falsch sein – die obige Aussage ist demnach wahr. ■

---

<sup>8</sup>die sog. *disjunktive Normalform* der Äquivalenzbeziehung

Man erkennt an den zwischendurch verwendeten Wörtern “dann”, dass eine Implikationskette aufgebaut wird. Am Schluss führt diese auf einen Widerspruch, d.h. auf  $\mathcal{F}$ , was beim indirekten Beweis meist durch einen Blitz gekennzeichnet wird.

(Das schwarze Kästchen am Ende ist eine Abkürzung für den Ausdruck “q.e.d.” (quod erat demonstrandum), den man notiert, um den Abschluss des Beweises zu kennzeichnen. Ein nicht ausgefülltes Kästchen bedeutet das gleiche).

- Noch ein Beispiel, das allerdings noch etwas mehr vorgreift und die Begriffe der Teilbarkeit und der Primzahlen verwendet. Behauptung:  $A :=$  (Jede Primzahl größer als 2 ist ungerade). Wir beweisen dies, indem wir das Gegenteil annehmen, und von dort aus einen Widerspruch ableiten. Aber dann kann die *Annahme* (nämlich, dass die *Behauptung* falsch sei) nicht stimmen. Und weil es in der Aussagenlogik nur zwei Möglichkeiten für Wahrheitswerte gibt, muss das bedeuten, dass die ursprüngliche Behauptung (A) wahr ist.

Nehmen wir also an, es gebe eine Primzahl  $p > 2$ , die gerade ist. Da jede gerade Zahl durch 2 teilbar ist, ist 2 also ein Teiler von  $p$ . Es sind aber auch  $p$  und 1 Teiler von  $p$ . Aber dann hätte  $p$  einen nicht-trivialen Teiler, nämlich 2, und wäre nicht prim – im Widerspruch zur Annahme  $\mathcal{F}$ . Also ist die Annahme falsch, und damit muss die Behauptung  $A$  wahr sein. ■

## 1.2 Mengenlehre

Wir hatten schon in der Einleitung gesehen, dass “Menge” und “Anzahl” für die Mathematik zwei verschiedene Begriffe sind; daher müssen wir hier zuerst klar stellen, was wir in der Mathematik unter Mengen verstehen wollen. Tatsächlich ist das gar nicht so leicht, wenn dies absolut stichfest sein soll. Mit Blick auf den Praxisbezug können wir uns hier aber eine gewisse Lockerheit leisten.

### 1.2.1 Grundbegriffe

**Definition 1.15** (Mengen). *Eine Ansammlung von (mathematischen) Objekten, die als Ganzes unter einem neuen Namen verstanden werden soll, heißt Menge. Die Objekte innerhalb einer Menge werden Elemente genannt. Falls  $M$  eine Menge ist und  $x$  ein Element aus  $M$ , so schreiben wir mit dem Element-Operator ‘ $\in$ ’:*

$$x \in M$$

*Falls  $x$  dagegen kein Element aus  $M$  ist, schreiben wir:  $x \notin M$ .*

*Elemente können nicht mehrfach in Mengen vorkommen.*

*Wir notieren die Elemente von Mengen innerhalb geschweiften Klammern ‘ $\{, \}$ ’.*

*Zwei Mengen heißen gleich, falls sie genau die gleichen Elemente enthalten.*

#### Bemerkungen:

- Man beachte, dass Mengen *nicht* geordnet sind: Für die Frage, ob  $x$  zu einer Menge  $M$  gehört oder nicht, ist keine Anordnung der Elemente innerhalb der Menge nötig – und wird auch nicht vorausgesetzt!
- Oft werden wir es mit Mengen von Zahlen zu tun haben, aber nicht ausschließlich. Auch andere Objekte dürfen Elemente von Mengen sein.
- Mengen selbst dürfen wiederum ebenfalls Elemente von (anderen) Mengen sein.
- Anders als bei den Klassen, die aus objektorientierten Programmiersprachen bekannt sind, sind mathematische Mengen nicht zwingend typisiert (in der Art, dass alle ihre Elemente gleichen Typ haben).
- Dass Elemente nicht mehrfach in Mengen vorkommen dürfen, setzt voraus, dass wir die Elemente unterscheiden können. Innerhalb einer Menge sind keine der Elemente gleich.

### Beispiele:

- Eine Menge von Zahlen (mit fünf Elementen) wäre z.B.:

$$M := \{3, 1, 4, 2, 9\}$$

Es gelten z.B. folgende Element-Beziehungen (die logische Aussagen sind!):  $4 \in M$ ,  $2 \in M$ , aber  $7 \notin M$ .

- Die Menge  $M$  aus dem vorigen Beispiel kann anders notiert werden, da (s.o.) die Anordnung der Elemente innerhalb der Menge nicht fest gelegt ist:

$$M = \{3, 1, 4, 2, 9\} = \{1, 2, 3, 4, 9\} = \{2, 1, 3, 4, 9\}$$

- Auch das hier ist eine Menge (mit  $M$  wie im ersten Beispiel):

$$N := \{3, a, M\}$$

Wir können  $M$  noch einsetzen und erhalten:

$$N = \{3, a, \{1, 2, 3, 4, 9\}\}$$

Die Menge  $N$  enthält *drei* Elemente: die einzelne 3, das Symbol  $a$ , und die Menge  $M$ . Hier haben wir ein Beispiel für eine Menge ohne einheitlichen Elementtyp.

- Auch das folgende ist eine Menge:

$$O := \left\{ \left\{ \{42\} \right\} \right\}$$

$O$  enthält nur genau ein Element, nämlich die Menge  $\{\{42\}\}$ . Letztere enthält wiederum genau ein Element, nämlich die Menge, die 42 enthält.

Also insbesondere:  $O \neq \{42\}$

- *Keine* Menge wäre allerdings

$$\{4, 1, 2, 4\}$$

Denn das Element 4 wäre hier doppelt vertreten.

- Ebenso ist dies keine Menge (mit  $M$  wie oben):

$$\{3, a, M, \{2, 9, 4, 3, 1\}\}$$

Denn hier wäre das Element  $M$  doppelt vertreten. Die Anordnung der Elemente von  $M$  spielt wie gesagt keine Rolle.

---

Zum Abschluss definieren wir noch zwei Begriffe, die mit der Anzahl der Elemente einer Menge zusammen hängen:

**Definition 1.16** (Mächtigkeit einer Menge). *Unter der Mächtigkeit (auch: Kardinalität) einer Menge  $M$  versteht man die Anzahl ihrer Elemente. Wir schreiben diese Anzahl symbolisch mit Betragsstrichen:*

$$|M|$$

*Ist  $|M|$  eine natürliche Zahl, so ist  $M$  eine endliche Menge.*

*$M$  heißt abzählbar unendlich, falls sie keine endliche Menge ist, aber ihre Elemente durch eine Vorschrift einzeln nacheinander aufgezählt werden könnten. Ist auch dies nicht möglich, so heißt  $M$  überabzählbar.*

*Ist  $M$  eine unendliche Menge, so schreiben wir  $|M| = \infty$ .*

### Bemerkungen:

- Offenbar kann  $|M|$  nicht negativ sein (eine Eigenschaft, die auch für den Betrag einer Zahl gilt).
- Außerdem kann  $|M|$  keine gebrochene (oder sogar irrationale) Zahl sein. Ein Objekt kann nur ganz oder gar nicht zu einer Menge gehören, nicht z.B. zu drei Vierteln.



**Beispiel:** Mit den Mengen  $M, N, O$  wie in den Beispielen zu Definition 1.15 gilt:

$$|M| = 5; \quad |N| = 3; \quad |O| = 1$$

---

**Definition 1.17** (Leere Menge). *Die leere Menge ist die Menge, welche kein Element enthält. Wir notieren sie mit*

$$\emptyset$$

**Bemerkungen:**

- Man findet auch manchmal die Notation  $\{\}$  für die leere Menge – auf diese wollen wir hier jedoch verzichten.
- Es gilt  $|\emptyset| = 0$ . Nach Definition 1.15 sind sämtliche Mengen mit null Elementen gleich – daher ist es sinnvoll, von *der* leeren Menge zu sprechen anstatt von *einer* leeren Menge.
- Da  $\emptyset$  keine Elemente enthält, ist für beliebige  $x$  die Aussage

$$x \in \emptyset$$

stets falsch (eine Kontradiktion).

- Man beachte: Die Menge

$$\{\emptyset\}$$

ist *nicht* die leere Menge! Es handelt sich vielmehr um eine Menge mit genau einem Element. Dieses Element wiederum ist die leere Menge.

Die leere Menge enthält nur keine Elemente – aber sie selbst ist durchaus ein mathematisches Objekt, mit dem man rechnen kann. Insbesondere ist  $\emptyset$  nicht *nichts*.

### 1.2.2 Teilmengen und Potenzmenge

**Definition 1.18** (Teilmenge). *Eine Menge  $A$  heißt Teilmenge einer Menge  $B$ , falls alle Elemente von  $A$  auch in  $B$  enthalten sind. Wir notieren dies mit dem Operator ' $\subseteq$ ':*

$$(A \subseteq B) := (\forall x : (x \in A) \Rightarrow (x \in B))$$

*Die Menge  $B$  heißt dann Obermenge von  $A$ .*

*Falls  $A \subseteq B$ , aber  $A \neq B$ , heißt  $A$  echte Teilmenge von  $B$ .*

**Bemerkungen:**

- Man findet für die Teilmengenbeziehung auch das Symbol ' $\subset$ '. Ähnlich wie beim Symbol ' $\leq$ ' soll jedoch bei uns darauf hingewiesen sein, dass die Mengen  $A, B$  auch gleich sein dürfen.
- Mit dem Konzept der Mengen können wir prädikatenlogische Aussagen (siehe Definition 1.9) nun kompakter formulieren, indem wir die Objekte, über die ausgesagt wird, durch Zugehörigkeit zu einer Menge einschränken. Konkret könnten wir z.B. die Teilmengenbeziehung von oben auch so formulieren:

$$(A \subseteq B) \equiv (\forall x \in A : x \in B)$$

Hier befassen wir uns also nicht mehr mit allen möglichen mathematischen Objekten  $x$ , sondern nur mit allen Elementen von  $A$ . Rechts vom Doppelpunkt steht ein Prädikat in  $x$ , denn  $x$  kann entweder zu  $B$  gehören oder nicht.

- *Insbesondere ist jede Menge auch Teilmenge von sich selbst.*
- Da die Aussage  $x \in \emptyset$  stets logisch falsch ist (siehe die Bemerkung bei Definition 1.17), und da die Implikation schon dann wahr ist, wenn ihre Prämisse verletzt wird (siehe Definition 1.11), können wir für jede beliebige Menge  $A$  notieren:

$$((x \in \emptyset) \Rightarrow (x \in A)) \equiv (\emptyset \subseteq A)$$

*Die leere Menge ist also Teilmenge jeder beliebigen Menge.*

- Für echte Teilmengen schreibt man manchmal  $A \subsetneq B$ , um die Nicht-Gleichheit auszudrücken.

### Beispiele:

- Die Menge  $M := \{2, 3, 4\}$  ist Teilmenge von  $\{1, 2, 3, 4, 5\}$ .
- $M$  ist auch Teilmenge von  $\{2, 3, 4\} = M$ , da jede Menge auch Teilmenge von sich selbst ist.
- $M$  ist aber nicht Teilmenge von  $N := \{1, 2, 3\}$ , da zwar  $4 \in M$ , aber  $4 \notin N$ . Hier wird die Implikationsbeziehung aus der Definition verletzt. Man notiert hier konkret:

$$M \not\subseteq N$$

---

Mit dem Teilmengenbegriff erhalten wir eine weitere Formulierung für die Gleichheit zweier Mengen:

**Satz 1.19** (Gleichheit von Mengen). *Zwei Mengen  $A, B$  sind gleich, falls sie jeweils Teilmengen voneinander sind:*

$$(A = B) \equiv ((A \subseteq B) \wedge (B \subseteq A))$$

(Der Beweis erfolgt direkt durch Einsetzen der Definitionen 1.18 sowie 1.13.)

---

Angenommen, eine Menge  $M$  hätte  $n$  Elemente, also  $|M| = n$ . Wie viele mögliche Teilmengen könnten wir aus  $M$  konstruieren – also Mengen, die nur Elemente aus  $M$  enthalten, aber keine aus anderen Mengen?

Zunächst ist klar, dass keine dieser Teilmengen mehr als  $n$  Elemente haben wird. Null Elemente wären allerdings möglich, da die leere Menge “automatisch” Teilmenge von  $M$  ist. Außerdem gibt es genau eine Teilmenge mit  $n$  Elementen, nämlich  $M$  selbst. Siehe zu beiden Behauptungen die Bemerkungen bei Definition 1.18.

Falls  $n > 1$ , hat es Sinn, die Teilmengen mit je einem Element zu betrachten. Davon gibt es genau  $n$  verschiedene, da die Elemente aus  $M$  alle paarweise verschieden zueinander sind. Ähnlich dazu gibt es genau  $n$  Teilmengen mit  $(n - 1)$  Elementen, nämlich durch Entfernen von je genau einem der  $n$  Elemente aus  $M$ .

Es stellt sich heraus, dass es insgesamt genau  $2^n = 2^{|M|}$  verschiedene Teilmengen von  $M$  gibt. Das gilt auch dann, wenn  $M = \emptyset$ , also  $n = 0$ .

Für unendliche Mengen gibt es schon trivialerweise<sup>9</sup> unendlich viele Teilmengen mit je einem Element. Die Anzahl sämtlicher Teilmengen ist also auf jeden Fall unendlich groß (und damit keine natürliche Zahl mehr). (Neugierige mögen dazu den Begriff *Kardinalzahl* nachschlagen.)

Wir begnügen uns hier mit folgender

**Definition 1.20** (Potenzmenge). *Unter der Potenzmenge einer Menge  $A$  versteht man die Menge, die sämtliche Teilmengen von  $A$  als Elemente enthält. Sie wird notiert als*

$$\mathcal{P}(A)$$

### Bemerkungen:

- Wegen der obigen Überlegungen (die allerdings hier nicht bewiesen wurden!) über die Mächtigkeit der Potenzmenge findet man teilweise für  $\mathcal{P}(A)$  auch die Bezeichnung  $2^A$  – auf diese wollen wir aber hier verzichten.
- Die Elemente einer Potenzmenge sind grundsätzlich stets Mengen.
- Falls eine Menge  $A$  Teilmenge einer Menge  $B$  ist, dann ist sie immer auch ein Element der Potenzmenge von  $B$ :

$$(A \subseteq B) \equiv (A \in \mathcal{P}(B))$$

---

<sup>9</sup>“trivial” ist ein vornehmer Ausdruck für “billig” oder “offensichtlich”, wird aber in der Mathematik offiziell benutzt, um zu beschreiben, dass ein Sachverhalt ohne substantielle weitere Denkarbeit sofort einzusehen ist.

### Beispiele:

- Für die Menge  $M := \{1, 2, 7\}$  finden wir die folgende Potenzmenge (systematisch nach Anzahl der Elemente in den Teilmengen notiert):

$$\mathcal{P}(M) = \{\emptyset, \{1\}, \{2\}, \{7\}, \{1, 2\}, \{1, 7\}, \{2, 7\}, \{1, 2, 7\}\}$$

- Für die Menge  $N := \{3, \{a\}\}$  mit zwei Elementen erhalten wir vier Teilmengen:

$$\mathcal{P}(N) = \{\emptyset, \{3\}, \{\{a\}\}, \{3, \{a\}\}\}$$

- Wir bilden noch die Potenzmenge von  $\mathcal{P}(N)$ , also die Menge der Teilmengen der Potenzmenge von  $N$ . Das werden insgesamt sechzehn Elemente, die wir sorgfältig notieren müssen, da schon die ursprüngliche Menge  $N$  ein Element hatte, das selbst eine Menge ist. Wir gehen wieder systematisch vor und notieren die Potenzmenge in fünf Zeilen (nach Anzahl der jeweils in den Teilmengen enthaltenen Elemente von 0 bis 4:

$$\begin{aligned} \mathcal{P}(\mathcal{P}(N)) &= \left\{ \emptyset, \right. \\ &\quad \{\emptyset\}, \{\{3\}\}, \{\{\{a\}\}\}, \{\{3, \{a\}\}\}, \\ &\quad \{\emptyset, \{3\}\}, \{\emptyset, \{\{a\}\}\}, \{\emptyset, \{3, \{a\}\}\}, \{\{3\}, \{\{a\}\}\}, \{\{3\}, \{3, \{a\}\}\}, \{\{\{a\}\}, \{3, \{a\}\}\}, \\ &\quad \{\emptyset, \{3\}, \{\{a\}\}\}, \{\emptyset, \{3\}, \{3, \{a\}\}\}, \{\emptyset, \{\{a\}\}, \{3, \{a\}\}\}, \{\{3\}, \{\{a\}\}, \{3, \{a\}\}\}, \\ &\quad \left. \{\emptyset, \{3\}, \{\{a\}\}, \{3, \{a\}\}\} \right\} \end{aligned}$$

- Die Potenzmenge der leeren Menge, also der Menge mit null Elementen, kann nur ein Element enthalten, nämlich die leere Menge, die trivial in jeder Potenzmenge enthalten ist:

$$\mathcal{P}(\emptyset) = \{\emptyset\}$$

- Die Potenzmenge der leeren Menge hat also Mächtigkeit 1. Dann erhalten wir zwei Elemente für deren Potenzmenge – nämlich die leere Menge und die Menge selbst (das ist so für alle einelementigen Mengen der Fall):

$$\mathcal{P}(\mathcal{P}(\emptyset)) = \{\emptyset, \{\emptyset\}\}$$

### 1.2.3 Bezug zur Aussagenlogik: Mengenoperationen und Rechenregeln

Über die Elementbeziehung können wir Mengenoperationen und Rechenregeln für Mengen auf die schon bekannten Zusammenhänge für logische Aussagen zurück führen. Zunächst also:

**Definition 1.21** (Verknüpfung von Mengen). *Seien  $A, B$  Mengen. Dann definieren wir den Durchschnitt (oder: Schnittmenge, Schnitt) von  $A$  und  $B$  mit dem Operator  $\cap$  per*

$$(x \in (A \cap B)) :\Leftrightarrow ((x \in A) \wedge (x \in B))$$

*Der Schnitt enthält genau die Elemente, die sowohl in  $A$  als auch in  $B$  enthalten sind.*

*Die Vereinigung von  $A$  und  $B$  wird mit dem Operator  $\cup$  erklärt:*

$$(x \in (A \cup B)) :\Leftrightarrow ((x \in A) \vee (x \in B))$$

*Die Vereinigung enthält alle Elemente, die in (mindestens) einer der beiden Mengen enthalten sind.*

*Die Differenz  $A$  und  $B$  enthält alle Elemente, die in  $A$  enthalten sind, aber nicht in  $B$ , und wird mit  $\setminus$  ausgedrückt:*

$$(x \in (A \setminus B)) :\Leftrightarrow ((x \in A) \wedge (x \notin B))$$

*Falls  $B$  eine Teilmenge von  $A$  ist, nennt man  $A \setminus B$  auch das Komplement von  $B$  bezüglich der Obermenge  $A$ ; in diesem Fall schreibt man alternativ noch:*

$$(A \setminus B) =: \mathbb{C}_A(B) =: \overline{B}$$

### Bemerkungen:

- Die logische Zuweisung “ $:\Leftrightarrow$ ” ist eine Variante der Zuweisung aus Definition 1.1, die nur für logische Aussagen zulässig ist. Hier heißt es also nicht “sei definiert als”, sondern spezieller “sei äquivalent zu” oder “soll genau dann gelten, wenn”.

Im folgenden Unterabschnitt zur Konstruktion von Mengen lernen wir noch eine Schreibweise kennen, mit der wir obige Definitionen wie üblich per “ $:=$ ” hätten einführen können.

- Beim Komplement wird zwar gern die Schreibweise mit dem Oberstrich verwendet – jedoch muss dann aus dem Kontext klar sein, auf welche Obermenge sich das Komplement bezieht. Nur dann ist das Komplement in einer Weise eindeutig definiert, die mit der logischen Negation vergleichbar ist.
- Grafisch veranschaulicht mit *Venn-Diagrammen*:

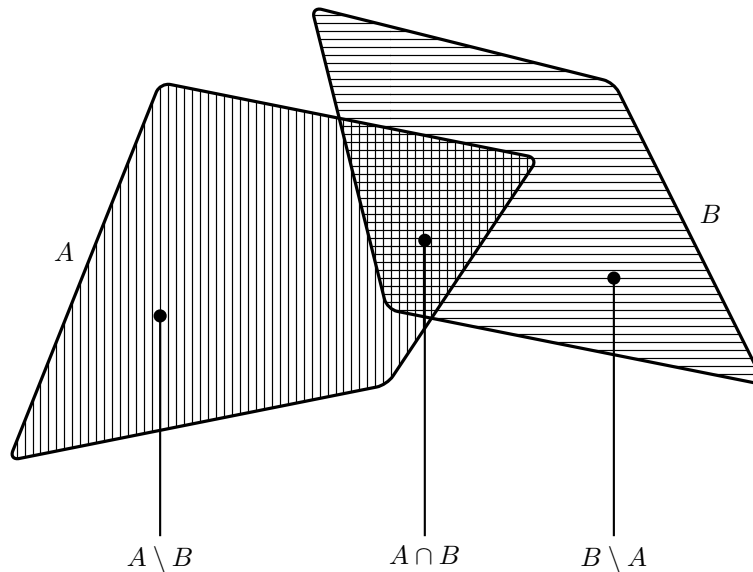


Abbildung 1.1: Mengenoperationen grafisch dargestellt

- Nach der Skizze ist auch ersichtlich, dass:

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

**Beispiel:** Für die Obermenge  $\Omega := \{1, 2, \dots, 10\}$  und die beiden Mengen  $A := \{2, 3, 7, 9\}$  und  $B := \{3, 4, 7, 8\}$  (siehe Abbildung 1.2, S. 29), gilt:

- $A \cap B = \{3, 7\}$
- $A \cup B = \{2, 3, 4, 7, 8, 9\}$
- $A \setminus B = \{2, 9\}$
- $B \setminus A = \{4, 8\}$
- $\overline{A} = \{1, 4, 5, 6, 8, 10\}$
- $\overline{B} = \{1, 2, 5, 6, 9, 10\}$
- $\overline{A} \cap \overline{B} = \{1, 5, 6, 10\} = \overline{A \cup B}$
- $\overline{A} \cup \overline{B} = \{1, 2, 4, 5, 6, 8, 9, 10\} = \overline{A \cap B}$

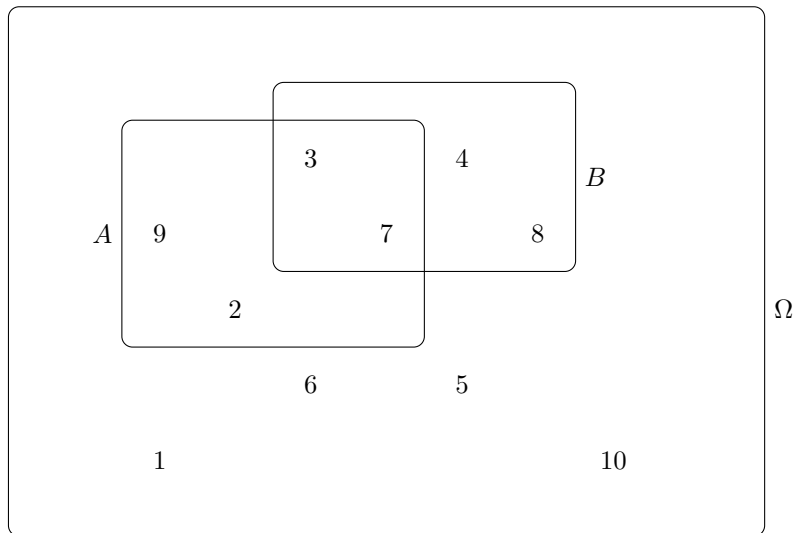


Abbildung 1.2: Mengen zur Beispielaufgabe

Die folgenden Rechenregeln gelten für Mengen (Beweis durch Zurückführen auf die aussagenlogischen Äquivalente und Benutzung von Satz 1.7 (Rechenregeln für logische Aussagen)):

**Satz 1.22** (Rechenregeln für Mengen). *Für Mengen  $A, B, C$ , die alle Teilmengen einer gemeinsamen Obermenge  $\Omega$  sind, gelten folgende Rechenregeln (alle Komplemente beziehen sich auf  $\Omega$ ):*

Name	Variante 1	Variante 2
Kommutativgesetze	$A \cap B = B \cap A$	$A \cup B = B \cup A$
Distributivgesetze	$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$	$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$
Neutralitätsgesetze	$A \cap \Omega = A$	$A \cup \emptyset = A$
Komplementärgesetze	$A \cap \bar{A} = \emptyset$	$A \cup \bar{A} = \Omega$
Assoziativgesetze	$A \cap (B \cap C) = (A \cap B) \cap C$	$A \cup (B \cup C) = (A \cup B) \cup C$
Idempotenzgesetze	$A \cap A = A$	$A \cup A = A$
Eliminationsgesetze	$A \cap \emptyset = \emptyset$	$A \cup \Omega = \Omega$
Absorptionsgesetze	$A \cap (A \cup B) = A$	$A \cup (A \cap B) = A$
Regeln von deMorgan	$\overline{A \cap B} = \bar{A} \cup \bar{B}$	$\overline{A \cup B} = \bar{A} \cap \bar{B}$
Doppeltes Komplement	$\overline{\bar{A}} = A$	

#### Bemerkungen:

- Teilmengen einer Obermenge und logische Aussagen sind zwei Ausprägungen einer gemeinsamen Struktur, die *Boolesche Algebra* genannt wird – deswegen sind die Operationen und Regeln so leicht übertragbar. Für diese Vorlesung nicht weiter von Belang.
- Man beachte die Verknüpfungen mit der leeren Menge: Vereinigung bringt keine Änderung (eine “neutrale Operation”); Schnitt führt (immer) auf die leere Menge.

Zum Abschluss dieses Unterabschnitts betrachten wir nochmal die Teilmengenbeziehung (die auf die logische Implikation zurück zu führen ist; siehe Definition 1.18). Es liege also eine Situation wie in Abbildung 1.3 (Seite 30):

$$A \subseteq B \subseteq \Omega$$

Dann gilt auch die Implikation

$$(x \in A) \Rightarrow (x \in B)$$

Die Abbildung zeigt genau die drei Bereiche, die mit diesem Sachverhalt verträglich sind:

- Für alle Elemente aus  $A$  gilt, dass sie auch in  $B$  enthalten sind.

- Für alle Elemente aus  $B \setminus A$  gilt, dass sie nicht in  $A$  enthalten sind, aber noch in  $B$ .
- Für alle Elemente außerhalb  $B$  gilt, dass sie weder in  $A$  noch in  $B$  enthalten sind. Auch das ist verträglich mit obiger Teilmengenbeziehung.
- Nur der Fall, dass ein Element in  $A$  liegt, aber nicht in  $B$ , ist nicht einzurichten. Das ist gerade der Fall, bei dem die Prämisse der Implikation erfüllt ist, aber die Konklusion nicht zutrifft. In der Abbildung mit Venn-Diagrammen findet sich dieser Fall nirgendwo – und in der Wahrheitstabelle der Implikation würde hier der Wahrheitswert  $\mathcal{F}$  stehen.

Wir sehen auch die kontrapositorische Formulierung der Implikation:

$$\begin{aligned}
 (A \subseteq B) &\equiv ((x \in A) \Rightarrow (x \in B)) \\
 &\equiv (\neg(x \in B) \Rightarrow \neg(x \in A)) \\
 &\equiv ((x \notin B) \Rightarrow (x \notin A)) \\
 &\equiv ((x \in \overline{B}) \Rightarrow (x \in \overline{A})) \\
 &\equiv (\overline{B} \subseteq \overline{A})
 \end{aligned}$$

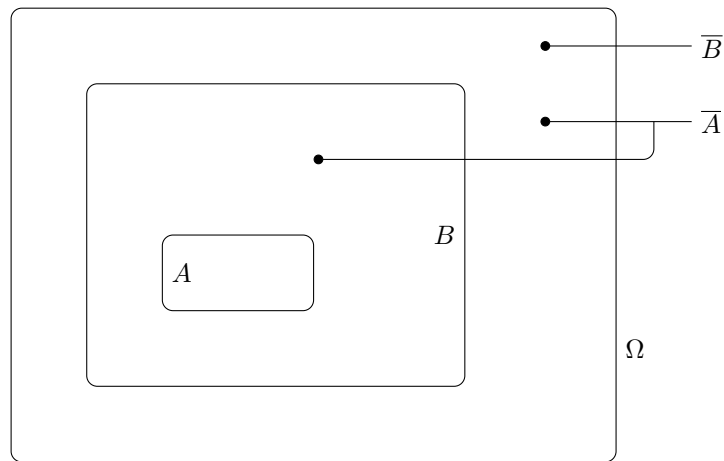


Abbildung 1.3: Teilmengenbeziehung

### 1.2.4 Konstruktion von Mengen

Mengen lassen sich auf verschiedene Weise konstruieren – entweder über die einzelnen Elemente oder aus anderen (bereits definierten) Mengen. Hier eine Übersicht gängiger Verfahren:

#### Angabe aller Elemente innerhalb geschweiften Klammern

Das ist möglich für endliche Mengen.

**Beispiel:**

$$\left\{4, 3, -1, \frac{\pi}{2}\right\}$$

#### Angabe einiger Elemente und einer Auslassung für abzählbare Mengen

**Beispiele:**

- $\{1, 3, 5, 7, \dots\}$
- $\{1, 2, 3, 4, \dots\}$
- $\{1, 2, 3, \dots, 10\}$
- $\{2, 3, 5, 7, 11, 13, \dots\}$

Das ist allerdings mathematisch *nicht exakt* und vor allem auch nicht eindeutig! Hier hängt es vom guten Willen der Lesenden ab, ob die Auslassungspunkte in Gedanken wie beabsichtigt ergänzt werden. Offenbar soll die erste Menge die ungeraden Zahlen darstellen, die zweite die natürlichen Zahlen und die dritte die natürlichen Zahlen von 1 bis 10.

Bei der vierten Menge muss man schon etwas schärfer hinsehen, um zu erkennen, dass hier die Primzahlen gemeint sind. Aber sind sie das wirklich? In der Mathematik dürfen wir uns darauf nicht verlassen! Deswegen wollen wir nach Möglichkeit auf Auslassungspunkte verzichten.

## Verknüpfung bereits definierter Mengen

**Beispiele:**

- $A := \{1, 4\} \cup \{15, 41\}$
- $B := A \setminus \{4\}$
- $C := \mathcal{P}(B)$

## Filtern einer Grundmenge

Aus einer bereits definierten Menge werden mit einem Prädikat die Elemente übernommen, für die das Prädikat zutrifft. Wir notieren ebenfalls mit geschweiften Klammern und mit einem Trennstrich. Links vom Trennstrich geben wir an, welche Grundmenge wir betrachten. Rechts steht ein Prädikat auf der Grundmenge.

**Beispiele:**

- $\Omega := \{3, 7, 2, 14, 9\}$ ,  $A := \{x \in \Omega \mid x > 5\}$ . Dann enthält  $A$  alle Elemente von  $\Omega$ , die größer sind als 5; das führt hier auf

$$A = \{7, 14, 9\}$$

- Falls wir mit dem selben  $\Omega$  nun alle Quadratzahlen filtern wollen, könnten wir schreiben:

$$B := \{x \in \Omega \mid \exists k : (k \text{ ist natürliche Zahl}) \wedge (x = k^2)\}$$

Das führt, da 9 die einzige Quadratzahl in  $\Omega$  ist, auf

$$B = \{9\}$$

- Hat das Filterprädikat für kein Element der Grundmenge den Wahrheitswert  $\mathcal{W}$ , so ergibt sich die leere Menge, z.B. (mit selbem  $\Omega$ ) bei

$$C := \{x \in \Omega \mid x^2 < 0\}$$

**Bemerkung:** Manchmal wird statt des Trennstrichs auch ein Doppelpunkt oder ein Komma notiert – wir wollen hier aber die Notation mit Trennstrich beibehalten.

## Filtern einer Grundmenge, mit Funktionsvorschrift

Eine Variante der vorigen Technik: Wenn die Elemente einer Grundmenge durch ein Prädikat selektiert wurden, kann man diese Elemente noch mit einer gemeinsamen Funktionsvorschrift abbilden – nämlich auf die Objekte, die man eigentlich in die neue Menge aufnehmen möchte.

Dabei kann dann allerdings links vom Trennstrich nur noch die Funktionsvorschrift stehen – aus welcher Menge selektiert wird, muss in das Filterprädikat einfließen. Das ist aber immer möglich durch eine einfache Und-Verknüpfung (da die Zugehörigkeit zu einer Menge mit dem Element-Operator als Prädikat formulierbar ist).

**Beispiele:**

- Mit selbem  $\Omega$  wie oben könnten wir z.B. sämtliche Elemente aus  $\Omega$  quadrieren:

$$D := \{x^2 \mid x \in \Omega\} = \{4, 9, 49, 81, 196\}$$

- Es spricht auch nichts dagegen, ein kombiniertes Filterprädikat zu benutzen:

$$E := \{x^2 \mid (x \in \Omega) \wedge (x > 5)\} = \{49, 81, 196\}$$

- Falls die Elemente einer Grundmenge (hier selbes  $\Omega$ ) mit einer konstanten Zahl (hier z.B. 3) multipliziert werden sollen, dann dürfen wir das auch so schreiben wie am rechten Ende folgender Zeile:

$$F := \{3 \cdot x \mid x \in \Omega\} = \{6, 9, 21, 27, 42\} =: 3\Omega$$

- Man beachte aber, dass die Notation aus dem vorigen Beispiel nur ausnahmsweise für das Herausziehen eines gemeinsamen Faktors zulässig ist. Bei anderen Operationen, z.B. dem Quadrieren, wäre es *falsch*, zu notieren:

$$\Omega^2 = \dots = \{4, 9, 49, 81, 196\} \quad \text{!}$$

Denn der Ausdruck  $\Omega^2$  auf der linken Seite entspricht eben *nicht* der Menge aller quadrierten Elemente von  $\Omega$ , sondern wird für das kartesische Produkt  $\Omega \times \Omega$  verwendet, das schon strukturell etwas völlig anderes ist (siehe nächster Unterabschnitt!).

### 1.2.5 Kartesisches Produkt

**Definition 1.23** (Tupel). *Für eine natürliche Zahl  $n$  wird eine geordnete Liste aus  $n$  Objekten (hier Elemente oder Komponenten genannt) als  $n$ -Tupel bezeichnet.*

*Wir notieren solche Tupel im allgemeinen als kommaseparierte Liste in runden Klammern.*

*Die Position innerhalb eines Tupels wird Index genannt und von 1 aufsteigend von links aus gezählt.*

*Zwei  $n$ -Tupel sind gleich, falls für jeden Index  $j$  von 1 bis  $n$  die beiden jeweiligen Komponenten mit Index  $j$  gleich sind.*

#### Bemerkungen:

- Bei  $n = 2$  spricht man auch von *Paaren*, bei  $n = 3$  von *Tripeln*.
- Anders als bei Mengen dürfen Objekte mehrmals in einem Tupel vorkommen.
- Die Formulierung “geordnete Liste” bedeutet, dass die Position eines Tuppelements fest ist. Rückt eine Komponente an eine andere Indexposition, so erhält man meist ein Tupel, das dem ursprünglichen nicht gleicht.

#### Beispiele:

- $(4, 1, 2)$  und  $(1, 2, 4)$  sind beides Tripel von natürlichen Zahlen. Sie sind *nicht* gleich, da es Komponenten gibt, die ungleich sind (in diesem Fall sogar alle drei:  $4 \neq 1$ ;  $1 \neq 2$ ;  $2 \neq 4$ ).
- Für  $A := (4, 1, 2)$  lauten die drei Komponenten von  $A$ :  $A_1 = 4$ ;  $A_2 = 1$ ;  $A_3 = 2$ .
- $G := (V, \Sigma, P, S)$  ist ein 4-Tupel. In der Vorlesung Theoretische Informatik 1 ist dies die allgemeine Notation für formale Grammatiken (siehe dort).
- $(7, \text{“Hallo”})$  ist ein Paar. Man könnte es z.B. verwenden, um der Zahl 7, etwa in einem Computerprogramm, den Text “Hallo” zuzuordnen.
- Allgemein sind alle Schlüssel-Wert-Paare (die oft ähnlich aussehen wie im vorigen Beispiel) Paare, also 2-Tupel.

**Definition 1.24** (Kartesisches Produkt von Mengen). *Für  $n$  gegebene Mengen  $M_1, M_2, \dots, M_n$  heißt die Menge aller möglichen  $n$ -Tupel das kartesische Produkt<sup>10</sup> dieser Mengen:*

$$M_1 \times M_2 \times \dots \times M_n := \{(x_1, x_2, \dots, x_n) \mid (x_1 \in M_1) \wedge (x_2 \in M_2) \wedge \dots \wedge (x_n \in M_n)\}$$

*Falls  $k > 1$  der aufeinander folgenden Mengen  $M_j, M_{j+1}, \dots, M_{j+(k-1)}$  gleich sind (z.B. gleich  $N$ ), kann man sie beim Notieren des kartesischen Produkts mit dem Ausdruck  $N^k$  zusammen fassen.*

<sup>10</sup>Nach R. Descartes, frz. Mathematiker und Naturwissenschaftler



**Bemerkung:** Falls die einzelnen Mengen alle endlich sind, gibt es auch nur endlich viele  $n$ -Tupel über ihnen; d.h. dann ist auch ihr kartesisches Produkt endlich; seine Mächtigkeit entspricht dem Produkt der Mächtigkeiten der einzelnen Mengen.

**Beispiele:**

- Für  $\Omega = \{2, 3, 7, 9, 14\}$  wie oben ist

$$\Omega \times \{4\} = \{(2, 4), (3, 4), (7, 4), (9, 4), (14, 4)\}$$

das kartesische Produkt von  $\Omega$  mit der einelementigen Menge  $\{4\}$ .

- Für die Mengen  $A := \{a, r, t\}$  und  $B := \{2, 7\}$  ist

$$A \times B = \{(a, 2), (a, 7), (r, 2), (r, 7), (t, 2), (t, 7)\}$$

das kartesische Produkt von  $A$  und  $B$ .

Das Produkt von  $B$  mit  $A$  hingegen wäre:

$$B \times A = \{(2, a), (2, r), (2, t), (7, a), (7, r), (7, t)\}$$

Auch hier finden wir sechs Elemente, da  $3 \cdot 2 = 2 \cdot 3$ . Aber die Elemente sind offenbar nicht die gleichen wie in  $A \times B$ .

- Für  $\Omega$  wie oben ergibt sich:

$$\begin{aligned} \Omega^2 &= \Omega \times \Omega \\ &= \{(2, 2), (2, 3), (2, 7), (2, 9), (2, 14), \\ &\quad (3, 2), (3, 3), (3, 7), (3, 9), (3, 14), \\ &\quad (7, 2), (7, 3), (7, 7), (7, 9), (7, 14), \\ &\quad (9, 2), (9, 3), (9, 7), (9, 9), (9, 14), \\ &\quad (14, 2), (14, 3), (14, 7), (14, 9), (14, 14)\} \end{aligned}$$

Dies ist schon strukturell nicht die Menge der Quadratzahlen der Elemente aus  $\Omega$  (s.o.).

- Der Ausdruck

$$A \times B^3 \times C \times A^2$$

ist eine abkürzende Notation für das kartesische Produkt

$$A \times B \times B \times B \times C \times A \times A$$

## 1.3 Wichtige Zahlenmengen

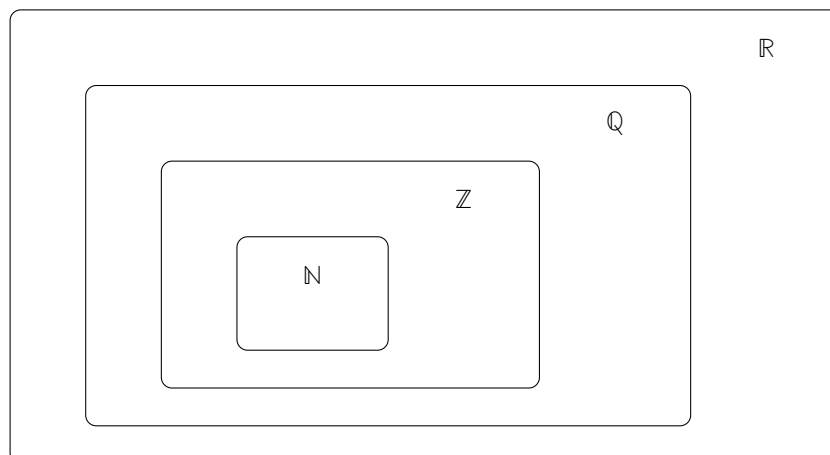


Abbildung 1.4: Hierarchie der für die Vorlesung wichtigen Zahlenmengen

Abbildung 1.4 zeigt eine Teilmengenbeziehung von vier bedeutenden Zahlenmengen, die in dieser Vorlesung häufig verwendet werden. Diese werden nun stichpunktartig vorgestellt.

### 1.3.1 Natürliche Zahlen $\mathbb{N}$

- $\mathbb{N} = \{1, 2, 3, \dots\}$
- Umgangssprachlich: “Alles, was sich mit (genügend Händen und) Fingern abzählen lässt”. Also insbesondere *Anzahlen* von unteilbaren, oder zumindest ungeteilten, Objekten.
- Formal: Definition durch die Axiome von Peano<sup>11</sup> (in der vollen Form nicht prüfungsrelevant). Für uns aber besonders wichtig:
  - 1 ist eine natürliche Zahl
  - Falls  $n$  eine natürliche Zahl ist, dann ist der Nachfolger  $(n+1)$  stets auch eine natürliche Zahl.
- $\mathbb{N}$  ist *abgeschlossen*<sup>12</sup> bzgl. *Addition*: Jede Summe von zwei natürlichen Zahlen ist wieder eine natürliche Zahl.
- abgeschlossen bzgl. *Multiplikation*: Jedes Produkt von zwei natürlichen Zahlen ist ebenfalls wieder eine natürliche Zahl.
- *nicht abgeschlossen* bzgl. Subtraktion oder Division! Beispiele:
  - 4 und 9 sind natürliche Zahlen, aber  $(4 - 9)$  nicht.
  - 2 und 3 sind natürliche Zahlen, aber  $(2/3)$  nicht.
- Falls die 0 (die keine natürliche Zahl ist) auch einbezogen wird, schreiben wir:

$$\mathbb{N}_0 := \{0\} \cup \mathbb{N} = \{0, 1, 2, 3, \dots\}$$

- Eine kompakte Beschreibung der natürlichen Zahlen als algebraische Struktur<sup>13</sup> werden wir im Algebra-Kapitel einführen.

**Beispiele:** Unsere drei prädikatenlogischen Beispiele zu Definition 1.9 (S. 16) lauten nun deutlich vereinfacht:

- $A \equiv (\forall n \in \mathbb{N} : (5 \cdot n > n))$   
Für sämtliche natürlichen Zahlen gilt, dass ihr Fünffaches größer ist als die jeweilige Zahl selbst.
- $b(n)$  sei ein Prädikat auf  $\mathbb{N}$ , d.h. nur für die natürlichen Zahlen wird folgendes betrachtet:

$$b(n) \equiv (\exists k \in \mathbb{N} : (k > n))$$

Da  $b(n)$  für alle natürlichen Zahlen richtig war, können wir auch formulieren:

$$\forall n \in \mathbb{N} : (\exists k \in \mathbb{N} : (k > n))$$

Für jede natürliche Zahl  $n$  existiert eine größere natürliche Zahl  $k$ .

- Das Prädikat *istGerade* lässt sich nun als Prädikat auf den natürlichen Zahlen so formulieren:

$$\text{istGerade}(n) \equiv (\exists k \in \mathbb{N} : n = 2 \cdot k)$$

### 1.3.2 Ganze Zahlen $\mathbb{Z}$

- $\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}$
- Erweiterung der natürlichen Zahlen  $\mathbb{N}$ , sodass die neue Zahlenmenge bzgl. Subtraktion ebenfalls abgeschlossen ist.
- Formal:
  - 0 ist eine ganze Zahl. Für jede ganze Zahl  $m$  gilt: Addiert man 0 hinzu, so bleibt die Zahl gleich, d.h.  $(m + 0) = m$ .
  - Für jede natürliche Zahl  $n \in \mathbb{N}$  definieren wir eine *negative Zahl*  $(-n) \in \mathbb{Z}$ , sodass

$$(n + (-n)) = (n - n) = 0$$

---

<sup>11</sup>G. Peano, ital. Mathematiker

<sup>12</sup>Wir kommen auf das Konzept der Abgeschlossenheit einer Operation nochmal im Algebra-Kapitel zurück.

<sup>13</sup>Auch dieser Begriff wird im Algebra-Kapitel eingeführt

### 1.3.3 Rationale Zahlen $\mathbb{Q}$

- Beschreibung aller möglichen Aufteilungen von ganzen Zahlen in gleich große Teile.
- Erweiterung der ganzen Zahlen  $\mathbb{Z}$ , sodass die neue Zahlenmenge bzgl. Division ebenfalls abgeschlossen ist (durch 0 darf allerdings nicht geteilt werden).
- Formal:
$$\mathbb{Q} := \left\{ \frac{z}{n} \mid z \in \mathbb{Z} \wedge n \in \mathbb{N} \wedge (z, n \text{ nicht kürzbar}) \right\}$$
- Darstellung alternativ auch als Kommazahl. Für rationale Zahlen sind diese Kommazahlen immer endlich lang oder periodisch (dazu mehr in Technischer Informatik 1).
- Für die Zahlenwerte gelten die Regeln der Bruchrechnung, die aus der Schule bekannt sind.

#### Beispiele:

- $\frac{3}{4} \in \mathbb{Q}$
- $\frac{7}{1} = 7 \in \mathbb{Q}$
- $\frac{-2}{1} = -\frac{2}{1} = -2$  ist auch rational. Alle ganzen Zahlen sind auch rationale Zahlen (Umkehrung gilt aber nicht!)
- $\frac{5}{0}$  ist *keine* rationale Zahl, da der Nenner nicht aus  $\mathbb{N}$  ist. Und durch 0 dürfte ohnehin nicht geteilt werden!
- $0 = \frac{0}{1}$  ist hingegen eine rationale Zahl.
- Nach obiger Definition wäre der Bruch  $\frac{36}{42}$  keine rationale Zahl. Der gekürzte Bruch  $\frac{6}{7}$  mit gleichem Zahlenwert ist allerdings aus  $\mathbb{Q}$ .

### 1.3.4 Reelle Zahlen $\mathbb{R}$

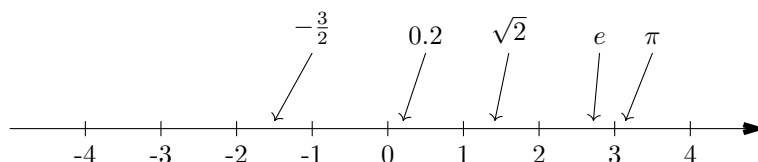


Abbildung 1.5: Reelle Zahlen auf dem Zahlenstrahl

- Umgangssprachlich: “Alles, was sich als Abstand zum Nullpunkt eines Zahlenstrahls vorstellen lässt” (siehe Bild 1.5; Vorzeichen durch Richtung des Zahlenstrahls gegeben). Im übertragenen Sinn also auch die geeignete Zahlenmenge für alle physikalischen Größen, die sich mit Messgeräten mit kontinuierlich beweglichem Zeiger darstellen lassen.
- Das sind sicher alle rationalen Zahlen – aber auch noch andere, die *irrationalen* Zahlen heißen und alle *nicht* als Brüche (bzw. endliche oder periodische Kommazahlen) darstellbar sind.
  - Zahlen, die sich als Nullstellen von Polynomen (siehe dazu den Abschnitt über Polynome im Algebra-Kapitel) mit ganzzahligen Koeffizienten schreiben lassen, z.B.  $\sqrt{2}$  als Lösung von  $x^2 - 2 = 0$ . (diese Zahlen heißen *algebraisch*)
    - Beweis der Irrationalität für  $\sqrt{2}$ : möglich als indirekter Beweis, unter Verwendung der Teilbarkeit und
  - *transzendente* Zahlen wie  $\pi$  (Kreiszahl) und  $e$  (Eulersche Zahl). Diese sind nur als Grenzwerte definierbar (Mathematik 2/Analysis), aber nicht als Nullstellen von Polynomen mit ganzzahligen Koeffizienten.
- Anders als die natürlichen, ganzen und rationalen Zahlen sind die reellen Zahlen auch nicht mehr *abzählbar*. Sie liegen “dicht” bzw. kontinuierlich auf dem Zahlenstrahl.

- Das Symbol  $\infty$  bezeichnet die (positive) Unendlichkeit (als Konzept in Mathematik 2 behandelt). Jede reelle Zahl ist echt kleiner als  $\infty$  und echt größer als  $-\infty$ . Daher sind  $\pm\infty$  *keine* reellen Zahlen.

---

Nach der Beschreibung von  $\mathbb{R}$  noch eine Definition zum Abschluss:

**Definition 1.25** (Intervall). *Eine Teilmenge der reellen Zahlen heißt (abgeschlossenes) Intervall, falls sie mit  $a \leq b$  auf die folgende Weise beschreibbar ist:*

$$[a, b] := \{x \in \mathbb{R} \mid (a \leq x) \wedge (x \leq b)\}$$

*Ist die Gleichheit an einer der Intervallgrenzen  $a, b$  nicht erfüllt, so notiert man statt der eckigen Klammer an der betreffenden Grenze eine runde Klammer; dann spricht man von einem (ggf. halb-) offenen Intervall.*

**Bemerkungen:**

- Die Symbole  $\pm\infty$  können (s.o.) nur an offenen Intervallgrenzen stehen.
- (nicht prüfungsrelevant) Jede reelle Zahl lässt sich als Grenzwert einer *Intervallschachtelung* von rationalen Zahlen beschreiben – also als diejenige Zahl, die in einer Abfolge von immer schmäler werdenden Intervallen, die jeweils in Teilmengenbeziehung stehen, auch dann noch enthalten ist, wenn die Intervallbreite gegen null strebt.

**Beispiele:**

- Das Intervall  $[4, 7]$  enthält sämtliche reellen Zahlen zwischen 4 und 7, einschließlich der beiden Intervallgrenzen.
- Die 4 gehört jedoch nicht zum Intervall  $(4, 7]$  (dieses ist halboffen), und auch nicht zum (offenen) Intervall  $(4, 7)$ ; zu letzterem gehört auch die 7 nicht dazu.
- Das Intervall  $[4, \infty)$  bezeichnet die Menge

$$\{x \in \mathbb{R} \mid x \geq 4\}$$

- Das Intervall  $(-\infty, \infty)$  steht synonym für sämtliche reellen Zahlen, also für  $\mathbb{R}$ .

## 1.4 (Un-)Gleichungen und Äquivalenzumformungen

### 1.4.1 (Un-)Gleichungen sind logische Aussagen

Wir hatten oben bereits über den Begriff der Gleichheit gesprochen. Hier wollen wir die Gleichheit (oder Ungleichheit) von Ausdrücken betrachten, die noch keinen konkreten Zahlenwert besitzen, sondern von zunächst unbekannten Zahlenwerten abhängen, die wir während der Rechnung symbolisch mit Variablen bezeichnen:

**Definition 1.26** ((Un-)Gleichung). *Eine Prädikat  $p(x_1, x_2, \dots, x_n)$  mit  $n$  Parametern heißt Gleichung mit  $n$  Unbekannten, falls es eine Gleichheit ausdrückt, also genau ein Gleichheitszeichen verwendet. Links und rechts des Gleichheitszeichens werden die Unbekannten mit festen Zahlenwerten und (verträglichen) Rechenoperationen verknüpft.*

*Die Unbekannten heißen auch Variablen oder Veränderliche.*

*Eine Ungleichung liegt dann vor, wenn statt des Gleichheitszeichens ein Symbol aus*

$$\{<, \leq, >, \geq\}$$

*verwendet wird.*

### Bemerkungen:

- Wenn nicht anders vereinbart, verstehen sich Gleichungen in dieser Vorlesung mit reellen Unbekannten.
- Die Symbole für Ungleichungen sind nur dann zulässig, wenn die verglichenen Ausdrücke aus einer *angeordneten* Zahlenmenge stammen. Reelle Zahlen (und Teilmengen davon) sind angeordnet; damit auch  $\mathbb{Q}, \mathbb{Z}, \mathbb{N}$ .
- Alle fünf vorgestellten Vergleichsrelationen sind *transitiv*, lassen sich also zu Kaskaden schachteln (solange immer dasselbe Vergleichssymbol verwendet wird). Wir sprechen dann von einer *(Un-)Gleichungskette*.
  - Es dürfen auch verschiedene Vergleichssymbole verwendet werden, solange diese nicht die generelle “Bewegungsrichtung” auf dem Zahlenstrahl ändern, z.B.:

$$a \leq \dots \leq \dots = \dots < \dots = \dots \leq b$$

Hier wird im Verlauf der Kaskade der Ausdruck zweimal mit Gleichheitszeichen umgeformt. Da die Kaskade ein “<” enthält, gilt insgesamt  $a < b$ .

- Wäre oben statt “<” ebenfalls “≤” notiert gewesen, würde gelten:  $a \leq b$
- Keine Aussage über den Vergleich zwischen  $a$  und  $b$  könnten wir jedoch treffen, falls die Kaskade auch “>” oder “≥” enthalten würde.
- Würde die Kaskade nur Gleichheitszeichen zum Vergleichen enthalten, so wäre  $a = b$ .

### Beispiele:

- Eine Gleichung mit drei reellen Unbekannten  $x, y, z$  wäre z.B.:

$$4 \cdot x + 3 = 10 \cdot y^5 - 12 \cdot x + z^2$$

- Hier eine Gleichung mit einer Unbekannten  $m \in \mathbb{N}$ :

$$3 \cdot m = 7 \cdot m - 9$$

Die Gleichung würde zwar auch für reelle Zahlen funktionieren, wird hier aber auf  $\mathbb{N}$  eingeschränkt.

- Eine Ungleichung mit zwei (reellen) Unbekannten:

$$4 \cdot x - 3 \cdot y \leq 12$$

---

Nun haben Prädikate zwei mögliche Wahrheitswerte. Setzt man z.B. in die erste obige Beispielgleichung die Zahlenwerte 2, 1 und 5 für die Variablen  $x, y, z$  ein, so erhält man:  $8 + 3 = 10 - 24 + 25$ , was nichts anderes bedeutet als  $11 = 11$ . Also ist, falls  $p$  das Prädikat ist, das diese Gleichung ausdrückt, auch  $p(2, 1, 5) \equiv \mathcal{W}$ . Das führt uns auf die folgende

**Definition 1.27** (Lösung einer (Un-)Gleichung). *Falls ein  $n$ stelliges Prädikat  $p(x_1, x_2, \dots, x_n)$  eine (Un-)Gleichung ausdrückt, so heißt jede konkrete Kombination von Variablen, für die  $p$  den Wahrheitswert  $\mathcal{W}$  annimmt, eine Lösung der (Un-)Gleichung.*

*Die Lösungsmenge der (Un-)Gleichung ist dann formal durch folgenden Ausdruck gegeben:*

$$\{(x_1, x_2, \dots, x_n) \mid p(x_1, x_2, \dots, x_n) \equiv \mathcal{W}\}$$

*Falls die Lösungsmenge der leeren Menge entspricht, heißt die (Un-)Gleichung unlösbar oder unerfüllbar.*

### Bemerkungen:

- Die Frage der Lösbarkeit kann mit den Zahlenmengen zusammen hängen, die für die Variablen vereinbart wurden.
- Die Lösungen müssen nicht als Tupel angegeben werden, sondern dürfen auch als Liste von Belegungen notiert sein, z.B. in der Form “ $x = 4; y = 19; \dots$ ”. Da Gleichungen selbst meist nicht formal als Prädikat notiert, sondern direkt ausgeschrieben werden, ist das sogar geschickter, da dann die Reihenfolge der Prädikats-Parameter nicht beachtet werden muss.

### Beispiele:

- Die erste oben aufgeführte Gleichung hat viele verschiedene Lösungen. Neben der Kombination  $(2, 1, 5)$  wäre z.B. auch die Kombination  $(20\frac{3}{8}, 2, 3)$  eine Lösung. Die Kombination  $(20, 2, 3)$  hingegen gehört nicht zur Lösungsmenge.
- Die zweite obige Gleichung ist für eine Variable  $m$  aus den natürlichen Zahlen definiert worden; unter der Bedingung ist ihre Lösungsmenge leer; die Gleichung ist also “unerfüllbar in  $\mathbb{N}$ ”.
- Wäre die Gleichung in den reellen (oder rationalen) Zahlen gegeben gewesen, so wäre der Bruch  $m = \frac{9}{4} = 2\frac{1}{4}$  eine Lösung.
- Die Ungleichung aus dem dritten obigen Beispiel hat (unendlich) viele Lösungen, z.B. die Paare  $(0, 0)$  und  $(1, 1)$ . Das Paar  $(30, 2)$  gehört jedoch nicht zur Lösungsmenge, denn der linke Ausdruck würde dann den Wert 114 besitzen, was größer als 12 ist.

---

Zum Abschluss definieren wir noch, was wir unter linearen und quadratischen Gleichungen verstehen:

**Definition 1.28** (Lineare und quadratische (Un-)Gleichungen). *Eine (Un-)Gleichung heißt linear, falls alle ihre Unbekannten höchstens mit der Potenz 1 auftreten. Sie heißt quadratisch, falls die Unbekannten höchstens mit der Potenz 2 auftreten.*

### Beispiele:

- Die erste Beispielgleichung bei Definition 1.26 ist weder linear noch quadratisch, da die Unbekannte  $y$  in fünfter Potenz auftritt.
- Die zweite Beispielgleichung ist linear, da sie keine höheren Potenzen als  $m^1 = m$  enthält.
- Auch die Ungleichung im dritten obigen Beispiel ist linear.
- Eine quadratische Gleichung mit zwei Variablen wäre z.B.:

$$4 \cdot y^2 - 3 \cdot y + x = 12 \cdot x^2 - 19$$

- Die Gleichung

$$\sin(x + 3) = \frac{1}{\sqrt{3}}$$

ist *nicht* linear. Zwar sehen wir  $x$  nur in erster Potenz, aber die Sinusfunktion ist de facto eine Art unendliches Polynom<sup>14</sup> und enthält sämtliche ungeraden Potenzen des Arguments (hier von  $(x + 3)$ ).

### 1.4.2 (Un-)Gleichungen durch Äquivalenzumformung lösen

Da (Un-)Gleichungen als Prädikate geschrieben werden können, können wir versuchen, diese in einer Art *umzuformen*, dass sich der Wahrheitswert des jeweiligen Prädikats dabei nicht ändert. Ziel ist hierbei, die Ausdrücke so zu vereinfachen, dass sich die Lösung direkt ablesen lässt.

Wenn die zugehörigen Prädikate für jeden Umformungsschritt alle den gleichen Wahrheitswert behalten, sind sie zueinander logisch äquivalent (siehe Definition 1.4). Daher:

**Definition 1.29** (Äquivalenz von (Un-)Gleichungen). *Zwei (Un-)Gleichungen heißen äquivalent genau dann, wenn sie die gleiche Lösungsmenge besitzen.*

*Die Äquivalenz zwischen (Un-)Gleichungen wird mit dem Symbol “ $\Leftrightarrow$ ” ausgedrückt.*

---

<sup>14</sup>Genauer: eine Potenzreihe. Nicht prüfungsrelevant in Mathematik 1; siehe Mathematik 2 / Analysis. Polynome behandeln wir im Algebra-Kapitel (s.u.)

### Bemerkungen:

- Führt eine Kette von Äquivalenzumformungen von der ursprünglichen (Un-)Gleichung auf einen Widerspruch (Kontradiktion), dann entspricht der Wahrheitswert der ursprünglichen (Un-)Gleichung aufgrund der Äquivalenzzeichen ebenfalls  $\mathcal{F}$ . Dann gibt es für die (Un-)Gleichung gar keine Kombination von Parametern, die eine Lösung wäre, d.h. die Lösungsmenge ist leer; die (Un-)Gleichung ist unerfüllbar.
- Führt eine Kette von Äquivalenzumformungen dagegen auf eine Tautologie (" $0 = 0$ "), dann besteht die Lösungsmenge der ursprünglichen (Un-)Gleichung aus allen zugelassenen Werten für die Unbekannten.
- Wegen Definition 1.13 ist eine Kette von Äquivalenzumformungen stets in beiden Richtungen als Kette von Implikationen lesbar – also folgt nicht nur das Resultat (in üblicher Leserichtung) aus der ursprünglichen (Un-)Gleichung, sondern aus dem Resultat folgt auch wieder (in umgekehrter Leserichtung) die ursprüngliche (Un-)Gleichung.

---

Wir betrachten nun Beispiele von Äquivalenzumformungen anhand einiger wichtiger Techniken:

**Skalieren von Gleichungen:** Falls für die Zahlen und Ausdrücke auf beiden Seiten eines Gleichheitszeichens eine Multiplikation mit reellen Zahlen definiert ist, dann ist für jedes  $c \in \mathbb{R}$ ,  $c \neq 0$ , auch das  $c$ -fache der linken Seite gleich mit dem  $c$ -fachen der rechten Seite.

### Beispiel:

$$9 = 4 \cdot m \quad \Leftrightarrow \quad \frac{9}{4} = m$$

Hier wurde die erste Gleichung mit dem Faktor  $\frac{1}{4}$  skaliert. Also Ausführlicher:

$$9 = 4 \cdot m \quad \Leftrightarrow \quad \frac{1}{4} \cdot 9 = \frac{1}{4} \cdot 4 \cdot m \quad \Leftrightarrow \quad \frac{9}{4} = m$$

Warum die Forderung  $c \neq 0$ ? Nach obiger Bemerkung zur beidseitigen Implikationsbeziehung müsste aus der resultierenden Gleichung auch wieder die ursprüngliche folgen, wenn man die Skalierung rückgängig macht – und das geht durch Multiplikation mit dem Kehrwert  $\frac{1}{c}$ , der dafür aber existieren muss!

Daher:

$$9 = 4 \cdot m \quad \Rightarrow \quad 0 \cdot 9 = 0 \cdot 4 \cdot m \quad \Leftrightarrow \quad 0 = 0$$

Die Implikation von links nach rechts ist hier richtig: Falls die linke Gleichung gilt, dürfen wir selbstverständliche ihre beiden Seiten mit null multiplizieren. Aus der resultierenden (tautologischen) Tatsache, dass null gleich null ist, lässt sich hingegen *nicht* in umgekehrter Richtung wieder auf die ursprüngliche Gleichung folgern.

Skalieren wir aber im obigen positiven Beispiel die Ergebnisgleichung wieder mit dem Kehrwert von  $\frac{1}{4}$ , also mit 4, dann erhalten wir, von rechts nach links gelesen, wieder die ursprüngliche Gleichung von der linken Seite.

**Skalieren von Ungleichungen:** Falls keine Gleichung, sondern eine Ungleichung vorliegt, können wir ebenfalls mit Faktoren ungleich null skalieren – allerdings ändert sich bei negativen Faktoren die Richtung der Ungleichheit.

### Beispiel:

$$\begin{aligned} -4x &\geq 2 \\ \Leftrightarrow -x &\geq \frac{1}{2} \\ \Leftrightarrow x &\leq -\frac{1}{2} \end{aligned}$$

Hier wurde die erste Ungleichung zunächst mit dem Faktor  $\frac{1}{4}$  skaliert; dabei blieb die Richtung der Ungleichheit erhalten. Danach dann die Skalierung mit  $(-1) < 0$ , wodurch sich das Vergleichssymbol umdreht.

Auch in umgekehrter Richtung sind die Skalierungen ausführbar, sodass die Äquivalenzpfeile richtig sind – hier würde man zunächst mit  $(-1)$  skalieren, danach mit 4, und käme wieder bei der oberen Ungleichung heraus.

**Addition von Gleichungen** Die linken bzw. rechten Seiten von zwei Gleichungen lassen sich jeweils addieren (solange für die Objekte in den Gleichungen eine Addition definiert ist): Angenommen, Gleichung (1) drücke in Zahlenwerten den Zusammenhang  $x = 3$  aus, und Gleichung (2) den Zusammenhang  $y = 7$ .

Achtung: Wenn diese beiden Gleichungen beide erfüllt sein sollen, müssen wir das auch kenntlich machen – und zwar mit einer Konjunktion. Dann liegen die Gleichungen *gekoppelt* vor, oder auch als *Gleichungssystem*<sup>15</sup>.

Für unser theoretisches Beispiel gilt dann folgende Implikationskette (an den geeigneten Stellen wird auf die verwendeten Gleichungen verwiesen, indem über dem Umformungspfeil deren Nummer angegeben ist):

$$\underbrace{(x = 3)}_{(1)} \wedge \underbrace{(y = 7)}_{(2)} \xrightarrow{(1)} x + y = 3 + y \xrightarrow{(2)} x + y = 3 + 7 = 10$$

Gilt denn auch die umgekehrte Implikationsrichtung – sodass wir eine Äquivalenzumformung konstruiert hätten? Leider nein – denn das Resultat der obigen Implikationskette ist eine Gleichung mit zwei Variablen, die nur die Summe aus  $x$  und  $y$  beschreibt, aber nicht länger deren konkrete Zahlenwerte. Es gibt unendlich viele Möglichkeiten, die Gleichung  $x + y = 10$  zu erfüllen, z.B. auch  $x = 2$  und  $y = 8$ , im Widerspruch zum ursprünglichen Gleichungssystem!

Um von  $x + y = 10$  wieder auf die mittlere Gleichung zurück zu gelangen, benötigen wir Gleichung (2), damit klar ist, dass sich die Zahl 10 als  $(3 + y)$  schreiben lässt. Dort angekommen, könnten wir durch Addition von  $(-y)$  auf beide Seiten immerhin Gleichung (1) rekonstruieren.

Oder aber, wir könnten aus  $x + y = 10$  mit Gleichung (1) wieder auf  $(3 + y) = 10$  gelangen, woraus wir Gleichung (2) durch Addition von  $(-3)$  auf beide Seiten erhalten.

Eine der beiden ursprünglichen Gleichungen ist also nötig, um von rechts nach links zurück zu gelangen. Folgendes wäre also eine korrekte Äquivalenzumformung:

$$(x = 3) \wedge (y = 7) \Leftrightarrow (x = 3) \wedge (x + y = 10)$$

Wir merken uns (und das ist auch später für die linearen Gleichungssysteme wichtig!): *Durch Äquivalenzumformungen lässt sich ein Gleichungssystem nicht auf weniger Gleichungen als ursprünglich reduzieren, es sei denn, es werden triviale Gleichungen entfernt.* Mit anderen Worten: Wenn zwei Gleichungen addiert werden, erhalten wir nur eine Äquivalenzumformung, wenn wir eine der beiden ursprünglichen Gleichungen “mitnehmen” (eine genügt).

Eine triviale Gleichung ist eine mit tautologischem Gehalt. Wäre Gleichung (2) von oben z.B. nicht  $y = 7$ , sondern  $7 = 7$ , dann wäre Gleichung (2) eine Tautologie. Addition mit der Gleichung (1) würde ergeben:  $(x + 7) = 3 + 7 = 10$ . Daraus lässt sich, durch Addition mit der Tautologie  $(-7) = (-7)$  sofort die ursprüngliche Gleichung 1 wieder herstellen.

Auch als Gleichungssystem wäre alles richtig, denn die Gleichung  $7 = 7$  ist immer logisch wahr – und die Konjunktion der Gleichung (1) mit  $\mathcal{W}$  ergibt nach dem Neutralitätsgesetz (siehe Satz 1.7) äquivalent die Gleichung (1). Also gilt:

$$x = 3 \Leftrightarrow (x = 3) \wedge \mathcal{W} \Leftrightarrow (x = 3) \wedge (7 = 7) \Leftrightarrow x + 7 = 3 + 7 \Leftrightarrow x + 7 = 10$$

Das bleibt auch richtig, falls die Tautologie Unbekannte enthält: Für das zweite Beispiel zur Definition 1.26 betrachten wir zusätzlich die tautologische Gleichung

$$9 - 3 \cdot m = 9 - 3 \cdot m$$

Also können wir schreiben:

$$\begin{aligned} 3 \cdot m &= 7 \cdot m - 9 \\ \Leftrightarrow (9 - 3 \cdot m) + 3 \cdot m &= (9 - 3 \cdot m) + 7 \cdot m - 9 \\ \Leftrightarrow 9 &= 4 \cdot m \end{aligned}$$

<sup>15</sup>Mehr zu linearen Gleichungssystemen im zugehörigen späteren Kapitel der Vorlesung



**Addition von Ungleichungen:** Dies ist ebenfalls möglich, falls die Richtung der Ungleichheit dieselbe ist (falls sie dies nicht ist, müsste man eine der Ungleichungen nur “umgedreht” notieren). Wir betrachten ohne Beschränkung der Allgemeinheit<sup>16</sup> den Fall “kleiner(-oder-gleich)”. Die beiden möglichen Szenarien sind (hier abstrakt notiert):

$$\begin{aligned} \bullet \quad & \underbrace{(a \leq b)}_{(1)} \wedge \underbrace{(c \leq d)}_{(2)} \stackrel{(1)}{\Rightarrow} a + c \leq b + c \quad \stackrel{(2)}{\Rightarrow} a + c \leq b + d \\ \bullet \quad & \underbrace{(a \leq b)}_{(1)} \wedge \underbrace{(c < d)}_{(2')} \stackrel{(1)}{\Rightarrow} a + c \leq b + c \quad \stackrel{(2')}{\Rightarrow} a + c < b + d \end{aligned}$$

Falls bei (mindestens) einer der beiden Ungleichungen die Gleichheit explizit ausgeschlossen ist, so ist sie das dann auch für die Summe der Ungleichungen.

(Man betrachte zur Übung auch den Fall, dass statt Ungleichung (1) eine Gleichung (1') mit  $a = b$  angenommen wird. Es stellt sich heraus, dass dann in jeweils mittleren Schritt Gleichheit vorliegt.)

Auch hier gilt: Sind beide Ungleichungen keine Tautologien, dann muss eine der beiden “mitgenommen” werden, um eine Äquivalenzumformung zu erreichen, also zu ermöglichen, dass von rechts nach links zurück gerechnet werden kann.

Wir halten das bisher gezeigte mit folgendem Satz fest:

**Satz 1.30** (Skalierung und Addition von (Un-)Gleichungen). *Seien  $\alpha, \beta, \gamma, \delta$  reelle Ausdrücke (ggf. mit Unbekannten), und  $c \in \mathbb{R}$ ,  $c \neq 0$ .*

*Dann gilt:*

$$(\alpha = \beta) \wedge (\gamma = \delta) \quad \Rightarrow \quad \alpha + \gamma = \beta + \delta \quad \text{und} \quad c \cdot \alpha = c \cdot \beta$$

*Für  $c > 0$  gilt auch:*

$$(\alpha \leq \beta) \wedge (\gamma = \delta) \quad \Rightarrow \quad \alpha + \gamma \leq \beta + \delta \quad \text{und} \quad c \cdot \alpha \leq c \cdot \beta$$

*Für  $c > 0$  gilt auch:*

$$(\alpha < \beta) \wedge (\gamma = \delta) \quad \Rightarrow \quad \alpha + \gamma < \beta + \delta \quad \text{und} \quad c \cdot \alpha < c \cdot \beta$$

*Sowie:*

$$(\alpha \leq \beta) \wedge (\gamma \leq \delta) \quad \Rightarrow \quad \alpha + \gamma \leq \beta + \delta$$

*Und:*

$$(\alpha < \beta) \wedge (\gamma \leq \delta) \quad \Rightarrow \quad \alpha + \gamma < \beta + \delta$$

*Und:*

$$(\alpha < \beta) \wedge (\gamma < \delta) \quad \Rightarrow \quad \alpha + \gamma < \beta + \delta$$

### 1.4.3 Lineare Gleichungen in einer reellen Veränderlichen

Alle linearen Gleichungen mit einer reellen Variablen (hier o.B.d.A.  $x$  genannt) lassen sich durch Zusammenfassen der Terme äquivalent umformen zu

$$a \cdot x + b = 0$$

Ob diese Gleichung lösbar ist, hängt von den reellen Zahlen  $a$  und  $b$  ab:

- Falls  $a \neq 0$ , liegt immer noch eine lineare Gleichung vor, die genau eine Lösung besitzt:

$$a \cdot x + b = 0 \quad \Leftrightarrow \quad a \cdot x = -b \quad \Leftrightarrow \quad x = -\frac{b}{a}$$

(Falls  $x$  ganzzahlig oder anderweitig (z.B. auf ein bestimmtes reelles Intervall) eingeschränkt ist, kann es trotzdem sein, dass keine Lösung existiert!)

- Falls  $a = 0$  und  $b = 0$ , war die ursprüngliche Gleichung tautologisch, da sie äquivalent zu  $0 = 0$  ist. Dann ist jedes zulässige  $x$  eine Lösung.
- Falls  $a = 0$ , aber  $b \neq 0$ , ist die Gleichung unlösbar, da sie äquivalent zu  $b = 0$  ist (logischer Widerspruch  $\text{f!}$ ).

<sup>16</sup>ein feststehender Ausdruck, auch “o.B.d.A.” abgekürzt

### Beispiele:

- Für die Gleichung  $3 \cdot m = 7 \cdot m - 9$  hatten wir oben schon die (eindeutige) Lösung  $m = \frac{9}{4}$  ermittelt.
- Wir lösen folgende Gleichung in  $x$  (den Multiplikationspunkt lassen wir ab jetzt meist aus; er muss aber immer mitgedacht werden):

$$5x - 2 = 3x + 4 - x + 3x$$

Zum Lösen vereinfachen wir zunächst die rechte Seite durch Zusammenfassen der Terme (lineare Terme mit Faktor  $x^1 = x$  und absolute Terme mit Faktor  $x^0 = 1$ ). Dann bringen wir durch Addition mit einer tautologischen Gleichung die Terme mit Faktor  $x$  auf die linke Seite des Gleichheitszeichens, die anderen nach rechts:

$$\begin{aligned} 5x - 2 &= 3x + 4 - x + 3x = 5x + 4 \\ \Leftrightarrow 0x &= 0 = 6 \quad \text{!} \end{aligned}$$

Offenbar ist die ursprüngliche Gleichung also unlösbar.

(Den Ausdruck “ $0x$ ” müsste man nicht explizit anschreiben)

- Analog verfahren wir für die Gleichung

$$3x + 2 = 5x - 4 - 2x + 6$$

Also ergibt sich:

$$\begin{aligned} 3x + 2 &= 5x - 4 - 2x + 6 = 3x + 2 \\ \Leftrightarrow 0 &= 0 \end{aligned}$$

Diese Gleichung ist offenbar stets richtig, unabhängig von  $x$ . Also besteht die Lösungsmenge aus ganz  $\mathbb{R}$ .

### 1.4.4 Lineare Ungleichungen in einer reellen Veränderlichen

Hier gelten die analogen Zusammenhänge wie für lineare Gleichungen, wobei beim Skalieren noch auf die Richtung der Ungleichheit zu achten ist.

Man beachte außerdem, dass zwar die Aussage  $0 \leq 0$  tautologisch ist – aber die Aussage  $0 < 0$  ist eine Kontradiktion!

Außerdem liegen mit einer Ungleichung für  $x$  meist unendlich viele Lösungen vor.

### Beispiele:

- Wir betrachten folgende Ungleichung:

$$3m \leq 7m - 9$$

Auch hier können wir tautologisch eine Gleichung addieren (s.o.). Danach stellen wir die resultierende Ungleichung so um, dass die Variable  $m$  auf der linken Seite steht; man beachte, dass sich dabei die Richtung des Größenvergleichs ändert:

$$\begin{aligned} 3m &\leq 7m - 9 \\ \Leftrightarrow 9 &\leq 4m \\ \Leftrightarrow 4m &\geq 9 \\ \Leftrightarrow m &\geq \frac{9}{4} \end{aligned}$$

Die Lösungsmenge ist unendlich mächtig und lässt sich als Intervall (siehe Definition 1.25) schreiben:

$$\left\{ m \in \mathbb{R} \mid m \geq \frac{9}{4} \right\} = \left[ \frac{9}{4}, \infty \right)$$

Man beachte, dass das Symbol  $\infty$ , das für “Unendlich” steht, *keine* reelle Zahl ist und also niemals zu einem Intervall reeller Zahlen dazu gehören kann – daher die Schreibweise als halboffenes Intervall.

- Betrachtet man die selbe Ungleichung wie eben, aber nicht auf den kompletten reellen Zahlen, sondern nur auf dem Intervall  $[-2, \frac{9}{4}]$ , so ergibt sich die Lösung durch nachträglichen Schnitt mit diesem Intervall, und wir würden, trotz Ungleichung, nur einen Wert erhalten, nämlich genau für den Fall  $m = \frac{9}{4}$ , bei der Gleichheit.

### Bemerkungen:

- Die Operationen, die wir bisher bei den Äquivalenzumformungen benötigt hatten (Tauschen von Seiten, Addition von tautologischen Gleichungen, Skalierung von (Un-)Gleichungen), dürfen stillschweigend (also ohne Kommentar) ausgeführt werden.

Wichtig sind allerdings die Äquivalenzzeichen, da diese die logische Verbindung zwischen der (Un-)Gleichung und der Lösungsmenge herstellen.

- Speziell bei der Umformung von Gleichungen, und für den Fall, dass sich eine Seite der Gleichung durch die Umformung gar nicht verändert, lohnt es sich oft, statt der Kette von Äquivalenzumformungen eine Gleichungskette anzuschreiben – dann entfällt das mehrfache Notieren der unveränderten Seite.

Das haben wir oben im vorigen Unterabschnitt in den beiden ausgeführten Beispielen jeweils schon so gemacht (hier wurde die rechte Seite der Gleichung weiter vereinfacht und die linke hatte sich im ersten Schritt nicht geändert)

### 1.4.5 Binomische Formeln und quadratische Gleichungen in einer reellen Variablen

Erinnerung: Mit den Distributiv-, Kommutativ- und Assoziativgesetzen für reelle (oder rationale oder ganze/natürliche) Zahlen<sup>17</sup> lässt sich schnell das folgende zeigen:

**Satz 1.31** (Binomische Formeln). *Für zwei reelle Zahlen  $a, b \in \mathbb{R}$ , gelten die folgenden drei Gleichungen:*

$$\begin{aligned}(a+b)^2 &= a^2 + 2ab + b^2 \\ (a-b)^2 &= a^2 - 2ab + b^2 \\ (a+b)(a-b) &= a^2 - b^2\end{aligned}$$

(Der Beweis der ersten binomischen Formel findet sich auf Seite 312; man versuche nachzuvollziehen wie die Rechenregeln für Addition und Multiplikation jeweils verwendet werden.)

### Beispiele:

- $(x+9)^2 = x^2 + 18x + 81$
- $(x - \frac{3}{2})^2 = x^2 - 3x + \frac{9}{4}$
- $(x + \frac{p}{2})^2 = x^2 + px + \frac{p^2}{4}$
- $g^2 - 7 = (g + \sqrt{7})(g - \sqrt{7})$

---

Um zu Lösung einer allgemeinen quadratischen Gleichung zu gelangen, betrachten wir zunächst den folgenden

**Satz 1.32** (Quadratische Gleichung ohne Linearteil). *Für  $s \in \mathbb{R}$  hat die quadratische Gleichung*

$$x^2 = s$$

*abhängig von  $s$  folgende Lösungen:*

- Falls  $s < 0$ , wäre  $x^2 < 0$ , was für reelles  $x$  nicht lösbar ist.
- Falls  $s = 0$ , wäre in den reellen Zahlen die einzige Lösung  $x = 0$ .
- Falls  $s > 0$ , existiert  $\sqrt{s}$  mit  $s = (\sqrt{s})^2 = (-\sqrt{s})^2 = (\pm\sqrt{s})^2$ . Dann gibt es genau zwei Lösungen, nämlich

$$x = \pm\sqrt{s} \quad \Leftrightarrow \quad x = -\sqrt{s} \quad \vee \quad x = \sqrt{s}$$

---

<sup>17</sup>Eine genauere Aufstellung der Rechenregeln folgt im Algebra-Kapitel

### Bemerkungen:

- Für  $s > 0$  ist stets  $\sqrt{s} > 0$ . Die Quadratwurzel einer positiven reellen Zahl kann niemals negativ sein.
- Beim Quadrieren spielt das Vorzeichen des Arguments dagegen keine Rolle, d.h. für  $s > 0$  gilt:  $(-\sqrt{s})^2 = (\sqrt{s})^2 = s$
- Im Fall  $s > 0$  würde die Gleichung also  $x^2 = (\sqrt{s})^2$  lauten, was eigentlich vier Kombinationen von Vorzeichen erlauben würde. Man überzeugt sich aber schnell, dass zwei davon redundant sind, denn:

$$-x = +\sqrt{s} \quad \Leftrightarrow \quad x = -\sqrt{s}$$

Und:

$$-x = -\sqrt{s} \quad \Leftrightarrow \quad x = \sqrt{s}$$

Es reicht also, sich auf die beiden Fälle  $x = \pm\sqrt{s}$  zu beschränken.

- Die Schreibweise mit “ $\pm$ ” ist zulässig, aber nur dann sinnvoll, wenn zwei Zahlen symmetrisch zur Null auf dem Zahlenstrahl liegen. Gemeint ist eigentlich die Disjunktion, die im obigen Satz äquivalent notiert ist.

---

Nun lösen wir eine allgemeine quadratische Gleichung mit beliebigem Linearteil:

**Definition 1.33** (Quadratische Gleichung). Eine reelle quadratische Gleichung ist mit Zahlen  $a, b, c \in \mathbb{R}$  und  $a \neq 0$  gegeben per

$$ax^2 + bx + c = 0$$

Mit  $p := \frac{b}{a}$  und  $q := \frac{c}{a}$  liegt die Gleichung normiert vor:

$$\dots \Leftrightarrow x^2 + px + q = 0$$

### Beispiele:

- Die Gleichung  $x^2 - 7x + 19 = 0$  ist quadratisch und normiert.
- Die Gleichung  $4x^2 - 12x + 1 = 0$  ist quadratisch, aber nicht normiert.
- Die Gleichung  $0x^2 - 13x + 12 = 0 \Leftrightarrow -13x + 12 = 0$  ist nicht quadratisch.
- Die Gleichung  $2x - 5 = 3x^2 + 17x - 1$  kann äquivalent (und zwar durch Addition von  $(-2x + 5)$  auf beiden Seiten und anschließendem Vertauschen beider Seiten) zu einer quadratischen Gleichung umgeformt werden, nämlich zu  $3x^2 + 15x + 4 = 0$ . Man kann diese quadratische Gleichung natürlich außerdem noch normieren (hier durch eine Skalierung mit dem Faktor  $\frac{1}{3}$ ).

---

Solch eine normierte quadratische Gleichung wollen wir nun lösen. Anders als vorhin, wo die Unbekannte  $x$  nur quadriert vorlag, gibt es hier auch einen Linearteil  $p \cdot x$  – also lässt sich die Gleichung nicht direkt durch Wurzelziehen lösen. Wir können jedoch das dritte Beispiel zu Satz 1.31 (Binomische Formeln) zu Hilfe nehmen:

$$\left(x + \frac{p}{2}\right)^2 = x^2 + px + \left(\frac{p}{2}\right)^2$$

Es fällt auf dass die rechte Seite dieser binomischen Formel beide Terme mit der Unbekannten  $x$  enthält, die auch in der normierten quadratischen Gleichung vorkommen. Wir müssen also nur im Absolutglied (also dem Term ohne  $x$ -Abhängigkeit) eine additive Anpassung vornehmen und erhalten:

$$\begin{aligned} x^2 + px + q &= 0 \\ \Leftrightarrow x^2 + px + \underbrace{\left(\frac{p}{2}\right)^2 - \left(\frac{p}{2}\right)^2 + q}_0 &= 0 \\ \Leftrightarrow \left(x + \frac{p}{2}\right)^2 - \left(\frac{p}{2}\right)^2 + q &= 0 \\ \Leftrightarrow \left(x + \frac{p}{2}\right)^2 &= \underbrace{\left(\frac{p}{2}\right)^2 - q}_{=:D} \end{aligned}$$

Nun führen wir in Gedanken kurzzeitig statt  $x$  eine neue Unbekannte  $y$  ein, die die Rechnung vereinfacht (dieses Vorgehen heißt *Variablensubstitution*):

$$y := x + \frac{p}{2}$$

Damit, und mit der oben schon notierten Abkürzung  $D := \dots$  gilt:

$$\dots \Leftrightarrow y^2 = D$$

Aber dies hat nun genau die richtige Form, um Satz 1.32 für quadratische Gleichungen *ohne* Linearteil verwenden zu können. Wir erhalten also abhängig vom Wert der Zahl  $D$ , die *Diskriminante der quadratischen Gleichung* genannt wird, einen von drei Fällen (dort, wo Lösungen existieren, führen wir die *Rücksubstitution*  $x = y - \frac{p}{2}$  aus, um wieder zur ursprünglichen Unbekannten  $x$  zu gelangen):

- Es gibt kein  $y \in \mathbb{R}$ , und also auch kein  $x \in \mathbb{R}$ , das die Gleichung löst, falls  $D < 0$
- Falls  $D = 0$  gilt, gibt es eine („doppelte“) Lösung  $y = 0$ , also  $x = -\frac{p}{2}$
- Falls  $D > 0$  gilt, gibt es zwei verschiedene reelle Lösungen, nämlich  $y = \pm\sqrt{D}$ . Mit der Rücksubstitution führt das auf  $x = -\frac{p}{2} \pm \sqrt{D}$

Wir fassen dies zusammen im

**Satz 1.34** (Allgemeine quadratische Gleichung). *Mit  $p, q \in \mathbb{R}$  und der Diskriminanten*

$$D := \left(\frac{p}{2}\right)^2 - q$$

*hat die quadratische Gleichung*

$$x^2 + px + q = 0$$

*das folgende Lösungsverhalten:*

- Falls  $D < 0$ , existiert keine reelle Lösung
- Falls  $D = 0$ , existiert eine (doppelte) Lösung  $x = -\frac{p}{2}$
- Falls  $D > 0$ , existieren zwei reelle Lösungen:

$$x = -\frac{p}{2} \pm \sqrt{D} = -\frac{p}{2} \pm \sqrt{\left(\frac{p}{2}\right)^2 - q}$$

**Bemerkungen:**

- Durch Einführung *imaginärer Zahlen* erhält man (im Gegensatz zu reellen Zahlen) auch negative Quadrate; dann gibt es für  $D < 0$  zwei verschiedene Lösungen – allerdings innerhalb der *komplexen Zahlen*. Nicht prüfungsrelevant für Mathematik 1, aber Stoff von Mathematik 2.
- Das Verfahren, das wir zur Lösung verwendet haben, heißt *quadratische Ergänzung*: Wir haben zum Linearteil der Gleichung noch einen passenden Absolutbeitrag addiert/ergänzt (und gleich wieder subtrahiert), sodass wir die erste binomische Formel rückwärts verwenden konnten.
- Falls nicht mit normierten Gleichungen gerechnet wird, lauten die Lösungen:

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

Die Diskriminante wäre dann  $b^2 - 4ac$ , was bis auf einen festen positiven Faktor der Zahl  $D$  entspricht und daher gleichwertig verwendet werden kann.

- Quadratische Gleichungen ohne Linearteil lassen sich ebenfalls mit obiger Formel lösen; hier ist dann  $p = 0$  einzusetzen.

**Beispiele:**

- Die Gleichung  $x^2 + 2x - 15 = 0$  hat die Diskriminante

$$D = 1^2 - (-15) = 1 + 15 = 16 > 0$$

Also bestehen die Lösungen in den Zahlen

$$x = -1 \pm \sqrt{16} = -1 \pm 4 \Leftrightarrow x = -5 \vee x = 3$$

Proben:

Für  $x = -5$  erhalten wir:

$$x^2 + 2x - 15 = 25 - 10 - 15 = 0 \quad \checkmark$$

Für  $x = 3$  analog:

$$x^2 + 2x - 15 = 9 + 6 - 15 = 0 \quad \checkmark$$

- Die Gleichung  $x^2 + 6x + 9 = 0$  hat die (doppelte) Lösung  $x = -3$ . Entweder rechnet man dies mit der allgemeinen Formel (die Diskriminante wäre genau null), oder man sieht, dass die linke Seite der Gleichung mit binomischer Formel gerade  $(x + 3)^2$  entspricht.

Probe:  $x^2 + 6x + 9 = 9 - 18 + 9 = 0 \quad \checkmark$

- Die Gleichung  $x^2 + 2x + 17 = 0$  hat die Diskriminante

$$D = 1^2 - 17 = -16 < 0$$

und also keine reelle Lösung.

## 1.5 Betrag, Summen- und Produktzeichen

Dieser Abschnitt als Vorbereitung noch das Konzept des Absolutbetrags ein, sowie die Kurznotation bestimmter regelmäßiger Summen und Produkte.

**Definition 1.35** ((Absolut-)Betrag einer reellen Zahl). *Der Absolutbetrag (kurz: Betrag) einer Zahl  $x \in \mathbb{R}$  wird notiert mit  $|x|$  und ist gegeben als*

$$|x| := \begin{cases} x, & \text{falls } x \geq 0 \\ -x, & \text{falls } x < 0 \end{cases}$$

**Bemerkungen:**

- Es handelt sich um die Länge des Abstandes einer Zahl zum Nullpunkt des Zahlenstrahls.
- Insbesondere ist der Betrag einer reellen Zahl niemals negativ.
- Der Betrag eines Produkts zweier Zahlen entspricht dem Produkt der einzelnen Beträge. Weitere Rechenregeln für Beträge folgen in Mathematik 2.

**Beispiele:**

- $|42| = |-42| = 42$

- Für den zusammen gesetzten reellen Ausdruck  $(x - 7)$  erhalten wir:

$$|x - 7| = \begin{cases} x - 7, & \text{falls } (x - 7) \geq 0 \\ -(x - 7), & \text{falls } (x - 7) < 0 \end{cases} = \begin{cases} x - 7, & \text{falls } x \geq 7 \\ 7 - x, & \text{falls } x < 7 \end{cases}$$

- Mit vorigem Beispiel wäre z.B. für  $x := 42$ :

$$|x - 7| = |42 - 7| = 42 - 7 = 35$$

Aber für  $x := -5$ :

$$|x - 7| = |-5 - 7| = 7 - (-5) = 7 + 5 = 12$$

Hier sind verschiedene Rechenwege möglich, u.a. je nach dem, ob die obige Vereinfachung verwendet wird (so wie hier) oder nicht – aber das Ergebnis bleibt dasselbe. Beispielsweise gilt (durch direktes Verrechnen von  $x$  genauso:

$$|42 - 7| = |35| = 35$$

Und:

$$|-5 - 7| = |-12| = 12$$


---

Wichtig ist im Zusammenhang mit Beträgen noch der folgende

**Satz 1.36** (Dreiecksungleichung). *Für  $x, y \in \mathbb{R}$  gilt stets:*

$$|x + y| \leq |x| + |y|$$

(Beweis: S. 312.)

**Bemerkung:** Der namensgebende Bezug zum Dreieck ist für die hier angegebene Form noch nicht offensichtlich (mit Vektoren – siehe Kapitel 6 – ergibt sich eine anschaulichere Erklärung). Eine verwandte Gleichung kann man tatsächlich für ein Dreieck mit den Seitenlängen  $a, b, c$  zeigen – nämlich, dass stets gilt:

$$|a - b| \leq c$$

Dies folgt daraus, dass im ebenen Dreieck eine Seitenlänge stets höchstens so groß sein kann wie die Summe der beiden anderen Seitenlängen (Gleichheit gilt nur für ein “entartetes” Dreieck, das gar keinen Flächeninhalt hat).

**Beispiel:** Wir betrachten die Zahlen  $x = \pm 7$  und  $y = \pm 4$ :

- $|7 + 4| = |11| = 11 \leq |7| + |4| = 7 + 4 = 11 \quad \checkmark$
  - $|7 + (-4)| = |7 - 4| = |3| = 3 \leq |7| + |-4| = 7 + 4 = 11 \quad \checkmark$
  - $|-7 + 4| = |4 - 7| = |-3| = 3 \leq |-7| + |4| = 7 + 4 = 11 \quad \checkmark$
  - $|-7 + (-4)| = |-7 - 4| = |-11| = 11 \leq |-7| + |-4| = 7 + 4 = 11 \quad \checkmark$
- 

Für das folgende benötigen wir noch den Begriff der indizierten Objekte:

**Definition 1.37** (Indizierte Objekte). *Eine Menge von Objekten heißt indiziert mit dem Index  $j$  (wobei  $j$  aus einer Indexmenge  $J$  stammt), falls die Objekte durch Angabe von  $j$  eindeutig zugeordnet werden können.*

*Der Index wird meist als Subskript notiert.*

**Bemerkung:** Oft sind die Indices<sup>18</sup> natürliche Zahlen.

**Beispiele:**

- Die Menge aller Quadratzahlen  $Q := \{1, 4, 9, 16, \dots\}$  lässt sich mit  $j \in \mathbb{N}$  indizieren als

$$q_j := j \cdot j = j^2$$

Die  $q_j$  sind dann die Quadratzahlen, und  $j$  wäre der Index. Z.B. wäre dann  $q_7 = 49$  und  $q_1 = 1$ .

Die Menge  $Q$  lässt sich also notieren als

$$Q = \{q_j \mid j \in \mathbb{N}\} = \{j^2 \mid j \in \mathbb{N}\}$$

- Wenn wir für  $j$  nur die Indexmenge  $J := \{1, 4, 7\}$  betrachten, dann würden wir über die Indices aus  $J$  die Objekte einer Menge  $\{q_1, q_4, q_7\}$  adressieren können.
- 

<sup>18</sup>Plural von “Index”

Für indizierte Objekte, die summiert und multipliziert werden können, kann auch eine Gesamtsumme und ein Gesamtprodukt definiert werden:

**Definition 1.38** (Summe und Produkt indizierter Zahlenobjekte). *Es sei  $J$  eine Indexmenge und  $M$ , eine Menge von mit  $j \in J$  indizierten Zahlen, also*

$$M = \{m_j \mid j \in J\}$$

*Dann sind die Gesamtsumme und das Gesamtprodukt der Elemente von  $M$  notiert als*

$$\sum_{j \in J} m_j \quad \text{bzw.} \quad \prod_{j \in J} m_j$$

*Falls die Indexmenge  $J$  leer ist, evaluiert die Gesamtsumme zu 0 und das Gesamtprodukt zu 1. Falls  $J$  ein zusammen hängendes Intervall ganzer Zahlen von  $a$  bis  $b$  (mit  $b \geq a$ ) ist, d.h.*

$$J = [a, b] \cap \mathbb{Z},$$

*schreibt man für die Gesamtsumme auch*

$$\sum_{j=a}^b m_j = m_a + m_{a+1} + m_{a+2} + \cdots + m_{b-1} + m_b$$

*Und für das Gesamtprodukt analog*

$$\prod_{j=a}^b m_j = m_a \cdot m_{a+1} \cdot m_{a+2} \cdot \cdots \cdot m_{b-1} \cdot m_b$$

#### Bemerkungen:

- Die Symbole entsprechen dem großen griechischen Sigma (für Summe) und Pi (für Produkt).
- Obige Definition ist nur dann sinnvoll, wenn es bei Addieren (bzw. Multiplizieren) nicht auf die Reihenfolge oder Klammerung ankommt, d.h. wenn diese Operationen kommutativ und assoziativ sind (dazu mehr im Algebra-Kapitel s.u.).

#### Beispiele:

- Die Summe der ersten fünf natürlichen Zahlen (auch als fünfte *Dreieckszahl* bekannt), ist

$$\sum_{j=1}^5 j = 1 + 2 + 3 + 4 + 5 = 15$$

- Die Summe der ersten fünf Quadratzahlen ist

$$\sum_{j=1}^5 j^2 = 1^2 + 2^2 + 3^2 + 4^2 + 5^2 = 1 + 4 + 9 + 16 + 25 = 55$$

- Die Summe der ersten  $n$  natürlichen Zahlen lässt sich mit einem geschlossenen Ausdruck notieren ("Kleiner Gauß<sup>19</sup>-Trick"; Beweis auf S. 312, und später nochmal formal sauberer in Mathematik 2):

$$\sum_{j=1}^n j = \frac{n(n+1)}{2}$$

Für  $n = 5$  wie im ersten Beispiel erhalten wir tatsächlich:

$$\sum_{j=1}^5 j = \frac{5 \cdot 6}{2} = \frac{30}{2} = 15$$

---

<sup>19</sup>C.F. Gauß, dt. Mathematiker



- Die Potenz  $2^n$  einer Zahl  $n \in \mathbb{N}_0$  lässt sich schreiben als

$$2^n = \prod_{j=1}^n 2$$

Speziell ist für  $n = 0$  die Indexmenge leer, da der obere Index 0 hier kleiner wäre als 1. Dann erhalten wir nach der obigen Definition ein Gesamtprodukt von 1, also  $2^0 = 1$ , wie erwartet.

Für  $n = 3$  würden wir dagegen erhalten:

$$2^3 = \prod_{j=1}^3 2 = 2 \cdot 2 \cdot 2 = 8$$

Hier wurde für jeden der drei Indices genau einmal das Argument (hier ein fester Wert) der Produktfunktion multipliziert.

- Für  $n \in \mathbb{N}_0$  ist das  $n$ -Fache einer Zahl  $x$  schreibbar als

$$n \cdot x = \sum_{j=1}^n x$$

Mit der gleichen Begründung wie im vorigen Beispiel erhalten wir speziell  $0 \cdot x = 0$ , wie erwartet.

- Eine Zahl im Stellenwertsystem zur Basis  $b \in \mathbb{N}$  mit den Ziffern  $x_n, \dots, x_1, x_0$ , jeweils aus der Menge  $\{0, 1, \dots, (b-1)\}$ , hat den Zahlenwert

$$x = \sum_{j=0}^n x_j \cdot b^j$$

Falls wir die Zahl 3142 aus dem Fünfer-System betrachten, gilt z.B.  $x_3 = 3$ ,  $x_2 = 1$ ,  $x_1 = 4$  und  $x_0 = 2$  mit  $b = 5$ . Dann ist der Zahlenwert (dezimal):

$$(3142)_5 = 3 \cdot 5^3 + 1 \cdot 5^2 + 4 \cdot 5^1 + 2 \cdot 5^0 = 3 \cdot 125 + 1 \cdot 25 + 4 \cdot 5 + 2 \cdot 1 = 375 + 25 + 20 + 2 = 422$$

## 1.6 Fakultätsfunktion und Binomialkoeffizienten

Dieser kurze Abschnitt dient der Einführung der Fakultät einer natürlichen Zahl – ein Begriff, den wir später im Algebra-Kapitel noch benötigen werden, wenn wir uns mit der symmetrischen Gruppe befassen. Als direkteres Beispiel sollen hier die Binomialkoeffizienten dienen. Beide Konzepte sind auch für Mathematik 2 noch wichtig.

**Definition 1.39** (Fakultätsfunktion). Für  $n \in \mathbb{N}_0$  ist der Ausdruck  $n!$ , gelesen “ $n$  Fakultät”, definiert als

$$n! := \prod_{j=1}^n j$$

### Bemerkungen:

- Für  $n = 0$  ist speziell:  $0! = 1$
- Die Fakultät lässt sich für  $n \geq 1$  auch *rekursiv*, also über die Vorgängerwerte definieren:

$$n! = n \cdot (n-1)!$$

**Beispiel:** Für  $n = 4$  gilt:  $4! = 1 \cdot 2 \cdot 3 \cdot 4 = 24 = 4 \cdot 3!$

Eine wichtige Anwendung der Fakultätsfunktion findet sich in

**Satz 1.40** (Anordnungen einer endlichen Menge). *Sei  $M$  eine endliche Menge mit  $n := |M|$  Elementen. Dann lassen sich aus den  $n$  Elementen genau  $n!$  verschiedene  $n$ -Tupel konstruieren.*

(Beweis: S.312)

**Bemerkung:** Erinnerung: Mengen sind ungeordnet, aber Tupel sind geordnet!

**Beispiele:**

- Für die Menge  $\{a, x, 3\}$  mit drei Elementen erhalten wir die folgenden  $3! = 6$  Tripel:

$$(a, x, 3), \quad (a, 3, x), \quad (x, a, 3), \quad (x, 3, a), \quad (3, a, x), \quad (3, x, a)$$

- Für die Menge  $\{47, 11\}$  mit Mächtigkeit 2 erhalten wir  $2! = 2$  Paare:

$$(47, 11), \quad (11, 47)$$

- Für die Menge  $\{x\}$  mit einem Element gibt es auch nur ein 1-Tupel, nämlich  $(x)$ .

---

**Definition 1.41** (Binomialkoeffizient). *Für  $n, k \in \mathbb{Z}$  definieren wir den Binomialkoeffizient, gelesen „ $n$  über  $k$ “ als*

$$\binom{n}{k} := \begin{cases} \frac{n!}{k! \cdot (n-k)!}, & \text{falls } 0 \leq k \leq n \\ 0, & \text{sonst} \end{cases}$$

**Beispiele:**

- Für  $k = 0$  und  $k = n$  ergibt sich jeweils:

$$\binom{n}{0} = \binom{n}{n} = \frac{n!}{n!} = 1$$

- Für  $k = 1$  oder  $k = (n - 1)$  ergibt sich:

$$\binom{n}{1} = \binom{n}{n-1} = \frac{n!}{(n-1)!} = \frac{(n-1)! \cdot n}{(n-1)!} = n$$

**Bemerkungen:**

- Die obige Definition enthält, da  $n - (n - k) = k$  gilt, folgende Symmetrie:

$$\binom{n}{k} = \binom{n}{n-k}$$

- Die Binomialkoeffizienten, die nicht null betragen, lassen sich in ein dreieckiges Schema notieren; das ist das Pascalsche Dreieck<sup>20</sup> – siehe Abbildung 1.6 (Zeile:  $n$ , ab 0 gezählt; Spalte:  $k$ , diagonal von links ab 0 gezählt).

---

<sup>20</sup>B. Pascal, frz. Mathematiker



**Satz 1.42** (Addition benachbarter Binomialkoeffizienten). *Für  $n, k \in \mathbb{N}_0$  und  $n < k$  gilt:*

$$\binom{n}{k} + \binom{n}{k+1} = \binom{n+1}{k+1}$$

Eine wichtige Anwendung der Binomialkoeffizienten findet sich in

$$\binom{n}{k}$$

(Beweis: S. 313)

$$\binom{5}{2} = 10$$
$$\{a, b\}, \{a, c\}, \{a, d\}, \{a, e\}, \{b, c\}, \{b, d\}, \{b, e\}, \{c, d\}, \{c, e\}, \{d, e\}$$

# Kapitel 2

## Teilbarkeit und Primzahlen

Wir befassen uns in diesem Kapitel mit der Teilbarkeit natürlicher Zahlen, einem Teilgebiet der *diskreten Mathematik* (also der Mathematik abzählbarer (oder sogar endlicher) Mengen). Wir erarbeiten zunächst die Konzepte der Division mit Rest und des größten gemeinsamen Teilers – für dessen Berechnung wir im folgenden Abschnitt eine effiziente Methode kennen lernen.

In Kombination mit den Konzepten der Äquivalenzrelationen und -klassen, die aus der Vorlesung Theoretische Informatik 1 voraus gesetzt werden, folgt dann ein eigenes Kapitel (3) zum Thema Restklassen.

### 2.1 Konzepte

#### 2.1.1 Teiler

**Definition 2.1** (Teiler). *Eine Zahl  $a \in \mathbb{N}_0$  heißt Teiler (auch: Faktor) von  $n \in \mathbb{Z}$ , falls es eine Zahl  $k \in \mathbb{Z}$  gibt, sodass*

$$n = k \cdot a$$

*Man schreibt dann auch:  $a \mid n$  (lies: “a teilt n”).*

*Falls ein solches  $k$  nicht existiert, schreibt man die Verneinung als  $a \nmid n$ .*

*Die Menge aller Teiler von  $n$  sei bezeichnet als*

$$T_n := \left\{ a \in \mathbb{N}_0 \mid a \mid n \right\}$$

#### Bemerkungen:

- 1 ist Teiler jeder natürlichen Zahl, da  $1 \cdot n = n$ .
- Für  $n \in \mathbb{N}$  ist auch stets  $n$  ein Teiler von sich selbst, also  $n \mid n$ .
- Jede natürliche Zahl ist Teiler der Null, da  $a \cdot 0 = 0$ .
- 0 ist nur ein Teiler von 0, da  $0 \cdot k = 0$  für alle  $k \in \mathbb{Z}$  gilt.
- Wenn  $a, n, k$  alle aus  $\mathbb{N}$  genommen werden (also ohne die Null), dann sind  $a, k$  bei  $n = ka$  komplementäre Teiler von  $n$  (denn auch  $k$  ist ein Teiler von  $n$ ).
- Falls  $n < 0$  und  $a$  ein Teiler von  $n$  ist, müssten wir akzeptieren, dass mit  $n = ka$  der komplementäre Teiler  $k$  negativ ist. Zur Teilmengen wollen wir jedoch dann  $a$  und  $(-k)$  hinzufügen, damit  $T_n$  keine negativen Zahlen enthält. Dies hilft uns, Redundanzen zu vermeiden, die rechnerisch keine neuen Erkenntnisse brächten. Immer dann, wenn  $k \mid n$  gilt, ist auch  $(-k) \mid n$ .

#### Beispiele:

- Es gilt z.B.:  $4 \nmid 2$ ;  $4 \nmid 7$ ;  $4 \nmid 6$ ;  $4 \mid 4$ ;  $4 \mid 0$ ;  $4 \mid 124$
- Wir finden systematisch die Teiler von 60, indem wir mit  $a = 1$  starten und dann die Paare aus  $a$  und  $k$  mit  $60 = ka$ , von außen nach innen anordnen, wobei  $a$  schrittweise erhöht wird (das muss bedeuten, dass das komplementäre  $k$  in jedem Schritt sinkt!), solange es kleiner ist als  $k$ . Das liefert uns nacheinander:

- $a = 1, k = 60$
- $a = 2, k = 30$
- $a = 3, k = 20$
- $a = 4, k = 15$
- $a = 5, k = 12$
- $a = 6, k = 10$

Zwischen 6 und 10 liegen keine weiteren Teiler mehr, wie man leicht durch Ausprobieren bestätigt. Also sind die Teiler von 60 wie folgt:

$$T_{60} = \{1, 2, 3, 4, 5, 6, 10, 12, 15, 20, 30, 60\}$$


---

**Satz 2.2** (Anzahl der Teiler von natürlichen Zahlen). *Für  $n \in \mathbb{N}$  gilt stets: Entweder ist die Anzahl der Teiler gerade, oder  $n$  ist eine Quadratzahl.*

(Beweis auf S. 313.)

**Beispiele:**

- Die Zahl 60 ist keine Quadratzahl und hat, wie oben gesehen, zwölf Teiler – eine gerade Anzahl.
- Hier die Teiler der Quadratzahl 36:

$$T_{36} = \{1, 2, 3, 4, 6, 9, 12, 18, 36\}$$

Es gibt vier Paare komplementärer Teiler sowie den Teiler  $6 = \sqrt{36}$ ; insgesamt also neun Teiler – eine ungerade Anzahl.

---

**Satz 2.3** (Teiler eines Produkts). *Für  $a, b \in \mathbb{Z}$  mit Teiler-Mengen  $T_a, T_b$  ist die Menge der Teiler des Produkts  $a \cdot b$  gegeben als*

$$T_{a \cdot b} = \{j \cdot k \mid j \in T_a \wedge k \in T_b\}$$

(Beweis: S. 314.)

**Beispiele:**

- Für  $a = 12$  und  $b = 15$  erhalten wir:

$$T_{12} = \{1, 2, 3, 4, 6, 12\} \quad \text{und} \quad T_{15} = \{1, 3, 5, 15\}$$

Wir notieren nun alle möglichen Produkte von Teilern in systematischer Weise (zeilenweise  $T_{12}$  mit den Elementen von  $T_{15}$  multipliziert) und erhalten:

$$\begin{array}{l} 1, 2, 3, 4, 6, 12 \\ 3, 6, 9, 12, 18, 36 \\ 5, 10, 15, 20, 30, 60 \\ 15, 30, 45, 60, 90, 180 \end{array}$$

Wenn wir nun die Dubletten entfernen, erhalten wir:

$$T_{180} = T_{12 \cdot 15} = \{1, 2, 3, 4, 5, 6, 9, 10, 12, 15, 18, 20, 30, 36, 45, 60, 90, 180\}$$

Dies lässt sich direkt durch systematische Suche aller Teiler von 180 bestätigen.

- Mit  $a = 60$  und  $b = 3$ , also  $T_{60}$  von oben und  $T_3 = \{1, 3\}$  bekommen wir mit selbiger Methode:

$$\begin{array}{l} 1, 2, 3, 4, 5, 6, 10, 12, 15, 20, 30, 60 \\ 3, 6, 9, 12, 15, 18, 30, 36, 45, 60, 90, 180 \end{array}$$

Auch das führt nach Dublettenentfernung wieder auf  $T_{180}$ .

## 2.1.2 Primzahlen

**Definition 2.4** (Primzahlen). Eine Zahl  $n \in \mathbb{N}$  heißt Primzahl, falls die Menge ihrer Teiler genau zwei Elemente enthält. Diese beiden Teiler heißen triviale Teiler.

Falls  $n > 3$ , aber  $n$  nicht prim ist, heißt  $n$  zusammen gesetzt.

### Bemerkungen:

- Wir hatten oben schon gesehen, dass stets  $1 \mid n$  und  $n \mid n$  gelten.  $n = 1$  hat nur den Teiler 1 und ist deswegen nicht prim. Für  $n > 1$  sind also  $1, n$  die trivialen Teiler.
- Eine Methode, Primzahlen in einem bestimmten Intervall  $[2, n] \cap \mathbb{N}$  zu finden, ist das *Sieb des Eratosthenes*: Man beginnt bei  $j = 2$  und streicht sämtliche Vielfachen von  $j$  aus – denn diese können keine Primzahlen sein, da sie  $j > 1$  als Teiler enthalten. Danach erhöht man  $j$ , bis man eine Zahl erreicht, die noch nicht gestrichen wurde; dann verfährt man analog. Man muss dies nur für  $j \leq \sqrt{n}$  durchführen, siehe die Bemerkungen zum Satz 2.6 weiter unten.

**Beispiel:** Wir geben für die Zahlen 1 bis 20 jeweils die Teiler an und notieren, ob sie prim sind. Auch die Quadratzahlen mit ungerade vielen Teilern sind angegeben:

$n$	Teiler	Bemerkung
1	1	Quadratzahl
2	1, 2	Primzahl
3	1, 3	Primzahl
4	1, 2, 4	Quadratzahl
5	1, 5	Primzahl
6	1, 2, 3, 6	
7	1, 7	Primzahl
8	1, 2, 4, 8	
9	1, 3, 9	Quadratzahl
10	1, 2, 5, 10	
11	1, 11	Primzahl
12	1, 2, 3, 4, 6, 12	
13	1, 13	Primzahl
14	1, 2, 7, 14	
15	1, 3, 5, 15	
16	1, 2, 4, 8, 16	Quadratzahl
17	1, 17	Primzahl
18	1, 2, 3, 6, 9, 18	
19	1, 19	Primzahl
20	1, 2, 4, 5, 10, 20	

**Definition 2.5** (Primteiler). Falls ein Teiler  $a$  von  $n \in \mathbb{N}$  prim ist, heißt  $a$  Primteiler (oder: Primfaktor) von  $n$ .

### Beispiele:

- Die Primteiler von 60 sind (s.o.):  $\{2, 3, 5\}$ .
- Die Primteiler von 36 sind (s.o.):  $\{2, 3\}$ .

**Satz 2.6** (Primteiler von zusammen gesetzten Zahlen). Ist eine Zahl  $n \in \mathbb{N}$  zusammen gesetzt, so besitzt sie stets einen Primteiler  $p$  mit  $p \leq \sqrt{n}$ .

(Beweis auf S. 314)

### Bemerkungen:

- Kontrapositorisch gilt also auch: Hat eine Zahl  $n \in \mathbb{N}$  keinen Primteiler, der höchstens  $\sqrt{n}$  beträgt, so ist  $n$  nicht zusammen gesetzt, also prim.
- Beim Sieb des Eratosthenes (s.o.) reicht es also aus, die Primzahlen  $j$ , deren Vielfache gestrichen werden sollen, nur bis zur Wurzel aus  $n$  zu erhöhen. Alle Zahlen zwischen  $\sqrt{n}$  und  $n$  haben nämlich, falls sie zusammen gesetzt sind, einen Primteiler, der höchstens  $\sqrt{n}$  beträgt – und werden daher mit  $j \leq \sqrt{n}$  bereits sicher gestrichen. Alle bis dahin nicht gestrichenen Zahlen zwischen  $\sqrt{n}$  und  $n$  sind dann entsprechend Primzahlen.

### 2.1.3 Division mit Rest

Wir bereiten uns nun auf das Rechnen mit Restklassen (“Modulo-Rechnen”) vor, das im Kapitel 3 weiter vertieft werden soll. Wichtig ist hierbei besonders die Division mit Rest.

Wir sind das Rechnen mit Divisionsresten aus dem Alltag von der Zeitrechnung her gewohnt – die physikalische Zeit läuft immer weiter, aber es ist für uns sinnvoll, sie in wiederholende Abschnitte zu unterteilen. Daher rechnen wir oft modulo 12 (Stunden, Monate), modulo 24 (Stunden), modulo 28/29/30/31 (Tage) oder modulo 60 (Sekunden, Minuten). Um z.B. die Uhrzeit abzulesen, wird dann vom Beginn des aktuellen Tages aus gerechnet – dabei spielt es keine Rolle, wie viel Zeit seit der Fabrikation der Uhr vergangen ist.

Bezogen auf die ganzen Zahlen, die sich mit gleichförmigen Abständen in beiden Richtungen von null Richtung Unendlich erstrecken, können wir ähnlich wie bei der Uhr- oder Kalenderrechnung allgemein eine Unterteilung in Abschnitte von jeweils  $m$  Zahlen vornehmen. So wie es unendlich viele ganze Zahlen gibt, gibt es dann auch unendlich viele solche Abschnitte. Es leuchtet ein, dass jede beliebige ganze Zahl in genau einem dieser Abschnitte zu finden ist. Die Unterteilungsbreite  $m$  wird auch (der) *Modul* genannt.

Abbildung 2.1 zeigt eine solche Unterteilung für den Modul  $m = 4$  und einen Ausschnitt des Zahlenstrahls von -12 bis 11. Die einzelnen Abschnitte beginnen jeweils mit einem Vielfachen von  $m$  (hier: 4) und erstrecken sich bis zum Vorgänger des nächstgrößeren Vielfachen. Der Faktor  $q \in \mathbb{Z}$  dient zum Identifizieren der Abschnitte. Innerhalb eines Abschnitts ist dann noch interessant, wie weit man vom linken Rand (nämlich dem Vielfachen  $q \cdot m$  noch gehen muss, um eine konkrete ganze Zahl  $n$  zu erreichen. Das ist genau der Divisionsrest  $r$ , wenn wir diese Zahl  $n$  durch den Modul  $m$  (hier 4) dividieren; er kann, weil er jeweils pro Abschnitt gezählt wird, nur Werte zwischen null und  $(m - 1)$  (hier 3) annehmen.

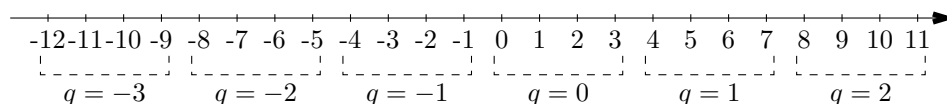


Abbildung 2.1: Ganze Zahlen auf dem Zahlenstrahl, unterteilt in Abschnitte zu jeweils  $m = 4$  Zahlen. Die Zahlen  $q \in \mathbb{Z}$  identifizieren die Abschnitte.

Es leuchtet ein, dass wir  $n$  und das Paar  $(q, r)$  stets eindeutig ineinander umrechnen können. Wir halten daher diese Tatsache fest als

**Satz 2.7** (Division mit Rest). *Gegeben sei eine Zahl  $m \in \mathbb{N}$  (der Modul).*

*Für jede Zahl  $n \in \mathbb{Z}$  gibt es dann Zahlen  $q \in \mathbb{Z}$  und  $r \in \mathbb{N}_0$  mit  $0 \leq r < m$ , sodass*

$$n = q \cdot m + r$$

*Dabei sind  $q, r$  eindeutig bestimmt.*

### Bemerkungen:

- Wenn man sich wie oben beschrieben den Zahlenstrahl in Abschnitte von jeweils  $m$  Zahlen unterteilt denkt, dann ist der Abschnitt  $q = 0$  heraus gehoben dadurch, dass die ganzen Zahlen dort nichtnegativ sind, aber kleiner als  $m$  – das sind genau die möglichen Divisionsreste  $r$ , die pro Abschnitt auftreten können.
- Falls ein Divisionsrest null ist, dann ist  $n$  ohne Rest durch  $m$  dividierbar – also ist  $n$  dann ein *Vielfaches* von  $m$ .

- Falls der Modul  $m$  den Wert 1 hat, enthält jeder Abschnitt genau eine ganze Zahl. Der einzige mögliche Divisionsrest ist dann immer null, da sich jede ganze Zahl ohne Rest durch 1 dividieren lässt. Im Vergleich zu den ganzen Zahlen  $\mathbb{Z}$  selbst brächte dies nichts neues, weswegen wir später nur noch Moduln von mindestens 2 betrachten werden.

#### Beispiele:

- $16 = 3 \cdot 5 + 1$ , d.h. für  $m = 5$  befindet sich  $n = 16$  im Abschnitt  $q = 3$ , mit einem Rest von  $r = 1$ .
- Analog ist auch:  $16 = 5 \cdot 3 + 1$ , d.h. für einen Modul  $m = 3$  befindet sich  $n = 16$  im Abschnitt  $q = 5$ , ebenfalls mit Rest  $r = 1$ .
- $16 = 4 \cdot 4 + 0$ , d.h.  $n = 16$  lässt sich ohne Rest durch  $m = 4$  dividieren,  $n$  ist ein Vielfaches von  $m$ .
- $-7 = (-2) \cdot 4 + 1$ , d.h. für  $m = 4$  liegt  $n = (-7)$  im Abschnitt  $q = (-2)$ , mit einem Rest von  $r = 1$ .

---

Hat man  $q, r$  gegeben und kennt den Modul  $m$ , so ist die ganze Zahl  $n$  direkt über obige Formel eindeutig berechenbar. Um die Rechnung in der anderen Richtung etwas formaler zu beschreiben, führen wir folgende Notation ein:

**Definition 2.8** (Operationen bei Division mit Rest). *Es seien  $n, q, m, r$  gegeben wie in Satz 2.7 beschrieben, und:*

$$n = q \cdot m + r$$

*Dann schreibt man für das Ergebnis der ganzzahligen Division von  $n$  mit  $m$ :*

$$q = \left\lfloor \frac{n}{m} \right\rfloor$$

*(gelesen: “ $n/m$  abgerundet”), und für den (eindeutig bestimmten) Divisionsrest:*

$$r = n \bmod m := n - m \cdot \left\lfloor \frac{n}{m} \right\rfloor$$

*(gelesen: “ $n$  modulo  $m$ ”).*

#### Bemerkungen:

- Die halben eckigen Klammern “[”, “]” sind auch allgemein benutzbar und dienen stets dazu, den Inhalt zur nächsten ganzen Zahl abzurunden. Analog dazu existieren Klammern zum Aufrunden (“[”, “[”).
- Das Symbol “mod” wird hier als Operator verwendet. Im Kapitel 3 taucht das selbe Symbol außerdem auf, um eine “Äquivalenz modulo  $m$ ” auszudrücken; siehe dort. Die beiden Konzepte sind verwandt, aber nicht gleich; es sollte aber aus dem Kontext klar sein, in welcher Weise das Symbol zu lesen ist.  
(Solche Überladung von Notation geschieht auch in der Mathematik, um ein Ausufern des Symbolvorrats zu vermeiden.)

Wir gehen obige Beispiele mit den neuen Symbolen nochmals durch:

#### Beispiele:

- Für  $16 = 3 \cdot 5 + 1$  erhalten wir:

$$\left\lfloor \frac{16}{5} \right\rfloor = \left\lfloor 3 + \frac{1}{5} \right\rfloor = 3 \quad \text{und} \quad 16 \bmod 5 = 1$$

- Für  $16 = 5 \cdot 3 + 1$  erhalten wir:

$$\left\lfloor \frac{16}{3} \right\rfloor = \left\lfloor 5 + \frac{1}{3} \right\rfloor = 5 \quad \text{und} \quad 16 \bmod 3 = 1$$



- Für  $16 = 4 \cdot 4 + 0$  erhalten wir:

$$\left\lfloor \frac{16}{4} \right\rfloor = \lfloor 4 \rfloor = 4 \quad \text{und} \quad 16 \bmod 4 = 0$$

- Für  $-7 = (-2) \cdot 4 + 1$  erhalten wir:

$$\left\lfloor -\frac{7}{4} \right\rfloor = \left\lfloor -\frac{8}{4} + \frac{1}{4} \right\rfloor = \left\lfloor -2 + \frac{1}{4} \right\rfloor = -2 \quad \text{und} \quad -7 \bmod 4 = 1$$

## 2.2 Größter gemeinsamer Teiler

In diesem Abschnitt befassen wir uns mit größten gemeinsamen Teilern (von zwei oder mehr natürlichen Zahlen). Wir greifen den Teilbarkeitsbegriff von oben auf und vertiefen ihn weiter. Dann lernen wir mit dem euklidischen<sup>1</sup> Algorithmus einen effizienten Weg kennen, um den größten gemeinsamen Teiler “ggT” zweier Zahlen zu berechnen. Anschließend halten wir einige Eigenschaften mit Bezug zum ggT fest, die für das Kapitel 3 wichtig sind.

Als Vorbereitung halten wir noch eine kleine, aber wichtige Tatsache fest:

**Satz 2.9** (Teiler bei Summe oder Differenz). *Für  $a \in \mathbb{N}$  und  $m, n \in \mathbb{Z}$  gilt: Ist  $a$  sowohl Teiler von  $m$  als auch von  $n$ , so teilt  $a$  auch deren Summe und Differenz:*

$$((a \mid m) \wedge (a \mid n)) \Rightarrow ((a \mid (m+n)) \wedge (a \mid (m-n)))$$

**Beweis:** Mit  $a \mid m$  und  $a \mid n$  gibt es  $s, t \in \mathbb{Z}$ , sodass  $m = sa$  und  $n = ta$ . Dann gilt nach dem Distributivgesetz auch:  $(m \pm n) = sa \pm ta = (s \pm t)a$ . Da die Zahl  $(s \pm t)$  aus  $\mathbb{Z}$  stammt, folgt die Behauptung nach Definition 2.1. ■

### 2.2.1 Gemeinsame Teiler

**Definition 2.10** (Gemeinsame Teiler). *Sei  $M \subseteq \mathbb{Z}$  eine Menge von ganzen Zahlen, und sei für jedes  $n \in M$  die Menge der Teiler von  $n$  mit  $T_n$  bezeichnet.*

*Dann heißt die Menge*

$$T_M := \bigcap_{n \in M} T_n$$

*die Menge der gemeinsamen Teiler der Zahlen aus  $M$ .*

**Bemerkung:** Das Schnittsymbol mit dem Selektor  $n \in M$  ist genau so zu verstehen wie das allgemeine Summen- oder Produktsymbol aus Definition 1.38: Sämtliche Mengen  $T_n$  für alle  $n$  aus der Menge  $M$  werden geschnitten.

**Beispiel:** Wir hatten oben im Unterabschnitt zu Teilern bereits die Mengen  $T_{36}$  und  $T_{60}$  bestimmt:

$$T_{36} = \{1, 2, 3, 4, 6, 9, 12, 18, 36\} \quad \text{und} \quad T_{60} = \{1, 2, 3, 4, 5, 6, 10, 12, 15, 20, 30, 60\}$$

Damit lauten die gemeinsamen Teiler von 36 und 60 also:

$$T_{\{36, 60\}} = T_{36} \cap T_{60} = \{1, 2, 3, 4, 6, 12\}$$

Damit führen wir nun auch den größten gemeinsamen Teiler ein:

**Definition 2.11** (Größter gemeinsamer Teiler). *Sei  $M \subseteq \mathbb{Z}$  eine Menge von mindestens zwei natürlichen Zahlen, und  $T_M$  die Menge der gemeinsamen Teiler der Zahlen aus  $M$ .*

*Dann heißt die Zahl*

$$\text{ggT}(M) := \max T_M$$

*der größte gemeinsame Teiler (abgekürzt ggT) der Zahlen aus  $M$ .*

---

<sup>1</sup>Euklid von Alexandria, gr. Mathematiker

### Bemerkungen:

- Der Operator “max” selektiert das größte Element einer Menge. Das ist immer dann möglich, falls die Elemente der Größe nach geordnet werden können – das ist für alle Teilmengen von reellen Zahlen (insbesondere also auch für Mengen ganzer oder natürlicher Zahlen) gegeben.
- Sehr hilfreich ist der ggT auch beim Kürzen von Brüchen. Dividiert man Zähler und Nenner durch ihren gemeinsamen ggT, so erhält man einen vollständig gekürzten Bruch.

### Beispiele:

- Mit den gemeinsamen Teilern wie im vorigen Beispiel ist:  $\text{ggT}(36, 60) = 12$ . Damit wäre z.B. auch folgende Rechnung in den rationalen Zahlen möglich:

$$\frac{36}{60} = \frac{36/12}{60/12} = \frac{3}{5}$$

- Die Teiler von 8, 12 und 24 sind:

$$T_8 = \{1, 2, 4, 8\}, \quad T_{12} = \{1, 2, 3, 4, 6, 12\} \quad \text{und} \quad T_{24} = \{1, 2, 3, 4, 6, 8, 12, 24\}$$

Damit sind die gemeinsamen Teiler:

$$T_{\{8,12,24\}} = T_8 \cap T_{12} \cap T_{24} = \{1, 2, 4\}$$

Und damit:

$$\text{ggT}(8, 12, 24) = 4$$

(Die Mengenklammern dürfen wir innerhalb des ggT-Ausdrucks weglassen.)

---

Wir halten folgende (nicht erschöpfende) Eigenschaften des ggT fest, die sich aus den Bemerkungen zu Definition 2.1 sowie aus Definition 2.4 ergeben:

**Satz 2.12** (Eigenschaften des ggT zweier Zahlen).

Gegeben  $n \in \mathbb{Z}$ , dann ist:

- $\text{ggT}(n, n) = |n|$
- $\text{ggT}(1, n) = 1$
- $\text{ggT}(0, n) = |n|$

Falls  $p \in \mathbb{N}$  prim ist und  $m \in \mathbb{Z}$ , dann ist

$$\text{ggT}(p, m) = \begin{cases} p, & \text{falls } p \mid m \\ 1, & \text{sonst} \end{cases}$$

---

Die Eins ist Teiler jeder ganzen Zahl, wird also auch immer Element der gemeinsamen Teiler mehrerer Zahlen sein. Interessant ist aber der Fall, dass dies auch der einzige gemeinsame Teiler ist:

**Definition 2.13** (Teilerfremdheit). Ist der ggT einer Menge von ganzen Zahlen genau 1, so heißen diese Zahlen teilerfremd (oder: relativ prim) untereinander).

### Beispiele:

- Alle Primzahlen sind paarweise zueinander teilerfremd, z.B.

$$\text{ggT}(2, 7) = \text{ggT}(2, 3) = \text{ggT}(3, 7) = 1$$

Also sind auch alle Primzahlen untereinander teilerfremd:

$$\text{ggT}(2, 3, 5, 7, 11, 13, \dots) = 1$$

- Nach dem obigen Satz sind Primzahlen  $p$  stets teilerfremd zu allen ganzen Zahlen, die keine Vielfachen von  $p$  sind, z.B.

$$\text{ggT}(3, 14) = \text{ggT}(128, 17) = \text{ggT}(42, 11) = 1$$

- Es ist

$$\text{ggT}(3, 14, 128) = 1$$

Also sind die drei Zahlen untereinander teilerfremd. Allerdings ist  $\text{ggT}(14, 128) = 2 \neq 1$ , also sind nicht alle drei Zahlen paarweise teilerfremd zueinander.

- Die Zahlen 6, 14 und 21 sind untereinander teilerfremd (also  $\text{ggT}(6, 14, 21) = 1$ ), obwohl

$$\text{ggT}(6, 14) = 2 \neq 1, \quad \text{ggT}(6, 21) = 3 \neq 1 \quad \text{und} \quad \text{ggT}(14, 21) = 7 \neq 1$$

---

Eine kurze Bemerkung zur Beziehung zwischen paarweiser und kollektiver Teilerfremdheit:

**Satz 2.14** (Teilerfremdheit von mehr als zwei Zahlen). *Finden sich in einer Menge  $M \subsetneq \mathbb{Z}$  zwei Zahlen, die zueinander paarweise teilerfremd sind, so ist  $\text{ggT}(M) = 1$ .*

(Beweis: S. 315.)

**Bemerkung:** Die Umkehrung gilt jedoch nicht – wie oben schon an zwei Beispielen demonstriert.

---

**Satz 2.15** (Skalierung und ggT). *Sei  $M := \{n_1, n_2, \dots, n_k\} \subsetneq \mathbb{Z}$  mit  $k \in \mathbb{N}$ . Für  $q \in \mathbb{N}$  gilt dann:*

$$\text{ggT}(q \cdot n_1, q \cdot n_2, \dots, q \cdot n_k) = q \cdot \text{ggT}(n_1, n_2, \dots, n_k)$$

(Beweis: S. 315.)

**Beispiele:**

- Wir betrachten die Zahlen 6, 14 und 21 aus den Beispielen zu Definition 2.13 und skalieren alle Zahlen mit  $q = 4$ , zu 24, 56 und 84. Wir erwarten also:

$$\text{ggT}(4 \cdot 6, 4 \cdot 14) = 4 \cdot 2 = 8, \quad \text{ggT}(4 \cdot 6, 4 \cdot 21) = 4 \cdot 3 = 12 \quad \text{und} \quad \text{ggT}(4 \cdot 14, 4 \cdot 21) = 4 \cdot 7 = 28$$

Die Teiler-Mengen der skalierten Zahlen lauten:

$$T_{24} = \{1, 2, 3, 4, 6, 8, 12, 24\}$$

$$T_{56} = \{1, 2, 4, 7, 8, 14, 28, 56\}$$

$$T_{84} = \{1, 2, 3, 4, 6, 7, 12, 14, 21, 28, 42, 84\}$$

Damit ergeben sich die oben erwarteten größten gemeinsamen Teiler 8, 12 und 28.

- Weiterhin ergibt sich aus dem vorigen Beispiel:

$$\text{ggT}(4 \cdot 6, 4 \cdot 14, 4 \cdot 21) = 4 = 4 \cdot 1 = 4 \cdot \text{ggT}(6, 14, 21)$$

---

**Satz 2.16** (Gemeinsame Teiler und ggT). *Sei  $M := \{n_1, n_2, \dots, n_k\} \subsetneq \mathbb{Z}$  mit  $k \in \mathbb{N}$ . Dann enthält die Menge  $T_M$  der gemeinsamen Teiler genau die Teiler des ggT von  $M$ .*

(Beweis auf S. 315.)

**Bemerkung:** Anders ausgedrückt gibt es also keine gemeinsamen Teiler, die nicht Teiler des ggT sind.

**Beispiele:**

- Im Beispiel zur Definition 2.10 hatten wir ermittelt:

$$T_{\{36,60\}} = \{1, 2, 3, 4, 6, 12\}$$

Tatsächlich sind dies genau die Teiler von  $\text{ggT}(36, 60) = 12$ .

- Analog hatten wir beim zweiten Beispiel zur Definition 2.11 erhalten:

$$T_{\{8,12,24\}} = \{1, 2, 4\}$$

Auch hier liegen genau die Teiler von  $\text{ggT}(8, 12, 24) = 4$  vor.

**Satz 2.17** (Assoziativität des ggT). *Sei  $M := \{n_1, n_2, \dots, n_k\} \subsetneq \mathbb{Z}$  mit  $k \geq 2$ . Dann gilt für beliebige nichtleere Teilmengen  $P \subsetneq M$  von  $M$ , und mit  $Q := M \setminus P$ :*

$$\text{ggT}(M) = \text{ggT}(\text{ggT}(P), \text{ggT}(Q))$$

(Beweis: S. 316.)

**Bemerkungen:**

- Falls  $P$  eine zweielementige Teilmenge von  $M$  mit  $\text{ggT}(P) = 1$  ist, folgt hieraus direkt (und alternativ) die Aussage von Satz 2.14
- Insbesondere erlaubt uns dieser Satz, den ggT von mehr als zwei Zahlen sukzessive auf mehrere ggT-Berechnungen von jeweils zwei Zahlen zu reduzieren, indem man  $P$  einelementig wählt. Die gemeinsamen Teiler von  $P$  sind dann trivial die Teiler der einen Zahl in  $P = \{p\}$ , und entsprechend ist  $\text{ggT}(P)$  damit gerade diese eine Zahl  $p$ . Die Menge  $Q$  enthält dann ein Element weniger als  $M$ . Für die Berechnung von  $\text{ggT}(Q)$  wiederholt man diesen Ansatz, etc., bis  $Q$  selbst einelementig ist.

Mit dem *Euklidischen Algorithmus* lernen wir im nächsten Unterabschnitt eine effiziente<sup>2</sup> Methode kennen, um gerade den ggT von zwei Zahlen zu berechnen.

**Beispiele:**

- Wir berechnen auf drei Arten den  $\text{ggT}(8, 12, 24)$  neu (siehe das Beispiel zu Definition 2.11; Ergebnis war 4).

Laut dem obigen Satz dürfen wir also rechnen:

$$\begin{aligned} \text{ggT}(8, 12, 24) &= \text{ggT}(\text{ggT}(\{8\}), \text{ggT}(\{12, 24\})) = \text{ggT}(8, \text{ggT}(\{12, 24\})) \\ &= \text{ggT}(8, 12) = 4 \end{aligned}$$

Oder alternativ:

$$\begin{aligned} \text{ggT}(8, 12, 24) &= \text{ggT}(\text{ggT}(\{8, 12\}), \text{ggT}(\{24\})) = \text{ggT}(\text{ggT}(\{8, 12\}), 24) \\ &= \text{ggT}(4, 24) = 4 \end{aligned}$$

Oder:

$$\begin{aligned} \text{ggT}(8, 12, 24) &= \text{ggT}(\text{ggT}(\{8, 24\}), \text{ggT}(\{12\})) = \text{ggT}(\text{ggT}(\{8, 24\}), 12) \\ &= \text{ggT}(8, 12) = 4 \end{aligned}$$

<sup>2</sup>Das Ermitteln aller Teiler und der Schnitt der Teiler-Mengen ist *nicht* effizient

- Auch für die Zahlen 6, 14 und 21 rechnen wir nochmal. Hier lassen wir, wie oben bereits erlaubt, einige der Mengenklammern weg:

$$\begin{aligned}\text{ggT}(6, 14, 21) &= \text{ggT}(\text{ggT}(6), \text{ggT}(14, 21)) = \text{ggT}(6, \text{ggT}(14, 21)) \\ &= \text{ggT}(6, 7) = 1\end{aligned}$$

Oder alternativ:

$$\begin{aligned}\text{ggT}(6, 14, 21) &= \text{ggT}(\text{ggT}(6, 14), \text{ggT}(21)) = \text{ggT}(\text{ggT}(6, 14), 21) \\ &= \text{ggT}(2, 21) = 1\end{aligned}$$

Oder:

$$\begin{aligned}\text{ggT}(6, 14, 21) &= \text{ggT}(\text{ggT}(6, 21), \text{ggT}(14)) = \text{ggT}(\text{ggT}(6, 21), 14) \\ &= \text{ggT}(3, 14) = 1\end{aligned}$$

### 2.2.2 Euklidischer Algorithmus

Wir wollen hier nur noch natürliche Zahlen für die ggT-Berechnung betrachten, da Teiler ohnehin mit positivem Vorzeichen notiert werden. Außerdem beschränken wir uns hier nur noch auf die ggT-Berechnung von je genau zwei Zahlen.

Zunächst soll motiviert werden, warum die Rechnungen, wie sie später im euklidischen Algorithmus auftreten, richtig sind. Im letzten Unterabschnitt verwenden wir den Algorithmus dann, um einige Zusammenhänge abzuleiten, die für das Kapitel 3 wichtig sein werden.

Zunächst halten wir folgende wichtige Tatsache fest:

**Satz 2.18** (Invarianz des ggT bei Differenzbildung). *Für natürliche Zahlen  $m, M \in \mathbb{N}$  mit  $m \leq M$  gilt:*

$$\text{ggT}(m, M) = \text{ggT}(m, (M - m))$$

(Beweis: S. 316.)

**Beispiele:**

- Wir ermitteln  $\text{ggT}(4, 24) = 4$ , indem wir wiederholt den Satz anwenden und am Schluss einmal Satz 2.12 (Eigenschaften des ggT zweier Zahlen) nutzen.

$$\begin{aligned}\text{ggT}(4, 24) &= \text{ggT}(4, (24 - 4)) \\ &= \text{ggT}(4, 20) = \text{ggT}(4, 16) = \text{ggT}(4, 12) = \text{ggT}(4, 8) = \text{ggT}(4, 4) \\ &= \text{ggT}(4, 0) = 4\end{aligned}$$

- Und das gleiche für den  $\text{ggT}(13, 21)$ . Man beachte, dass die Reihenfolge der Argumente in der ggT-Funktion keine Bedeutung hat – die beiden Namen  $m$  und  $M$  dienen hauptsächlich dazu, die Argumente nach der Größe zu ordnen.

Wie im vorigen Beispiel schreiben wir die Klammer, aus der die Differenz berechnet wird, nur in der ersten Zeile mit an:

$$\begin{aligned}\text{ggT}(13, 21) &= \text{ggT}(13, (21 - 13)) \\ &= \text{ggT}(13, 8) = \text{ggT}(8, 5) = \text{ggT}(5, 3) = \text{ggT}(3, 2) \\ &= \text{ggT}(2, 1) = \text{ggT}(1, 1) = \text{ggT}(1, 0) = 1\end{aligned}$$

(Wie im vorigen Beispiel hätten wir das Ergebnis ohne diesen Satz auch schneller erhalten können: Oben war 4 ein Teiler von 24, also muss es auch der ggT beider Zahlen sein; und hier ist 13 eine Primzahl, aber 21 kein Vielfaches von 13 – die Teilerfremdheit folgt dann direkt nach Satz 2.12)

Beim ersten der beiden vorigen Beispiele haben wir gesehen, dass wir mit Satz 2.18 mehrfach die Zahl 4 vom größeren (oder gleichen) Argument subtrahieren konnten, und zwar so lange, bis ein Wert kleiner als 4 (das anfänglich kleinere der beiden Argumente) heraus kam. Diese Schritte lassen sich in der Tat zusammen fassen, indem man die kleinere (genauer: nicht-größere) Zahl  $m$  so oft von  $M$  subtrahiert, bis das Ergebnis kleiner ist als  $m$ . Danach liegt die Ausgangssituation des Satzes in umgekehrter Weise vor: Das anfänglich nicht-größere der beiden Argumente ist nach den Subtraktionen (und das stets) das größere. Falls das andere Argument (das anfänglich nicht-kleinere) nach der Subtraktion nicht null ist, ist es auf jeden Fall kleiner als das andere geworden.

Danach kann der Satz erneut verwendet werden.

Betrachtet man die Subtraktionen wie im obigen Beispiel zusammen gefasst, dann erhält man eine bereits bekannte Technik, nämlich die *Division mit Rest* von Satz 2.7. Denn mit

$$M = q \cdot m + r$$

wird durch die Subtraktion von  $q \cdot m$  genau der Divisionsrest ( $r = M - q \cdot m$ ) errechnet – und der ist (siehe den Satz 2.7) echt kleiner als  $m$ .

Aber dann können wir den obigen Satz zur Invarianz bei Differenzbildung verallgemeinern, indem wir die  $q$ -fache Anwendung zusammen fassen. Wir halten diese Tatsache fest im folgenden

**Satz 2.19** (Invarianz des ggT bei Division mit Rest). *Für natürliche Zahlen  $m, M \in \mathbb{N}$  mit  $m \leq M$  gilt:*

$$\text{ggT}(m, M) = \text{ggT}(m, (M \bmod m))$$

**Bemerkung:** Dies kürzt die ggT-Berechnung vor allem dann ab, wenn  $m$  und  $M$  recht weit auseinander liegen. Im zweiten Beispiel zu Satz 2.18 war dies gerade nicht der Fall – sogar “maximal” nicht, da  $q$  in jedem Schritt 1 betrug. (*Mindestens* einmal lässt sich  $m$  wegen  $M \geq m$  stets von  $M$  subtrahieren.)

(Das war kein Zufall, denn die beiden ggT-Argumente waren aufeinander folgende *Fibonacci-zahlen*<sup>3</sup> – bei der ggT-Berechnung haben wir die rekursive Summenbildung rückwärts verfolgt.)

### Beispiele:

- Das erste Beispiel zu Satz 2.18 vereinfacht sich zu:

$$\text{ggT}(4, 24) = \text{ggT}(4, (24 \bmod 4)) = \text{ggT}(4, 0) = 4$$

- Und für das zweite Beispiel erhalten wir (wie angekündigt hier ohne Beschleunigung):

$$\begin{aligned} \text{ggT}(13, 21) &= \text{ggT}(13, 8) = \text{ggT}(8, 5) = \text{ggT}(5, 3) = \text{ggT}(3, 2) = \text{ggT}(2, 1) = \text{ggT}(1, 0) \\ &= 1 \end{aligned}$$

- Für die Zahlen 36 und 60 erhalten wir:

$$\text{ggT}(36, 60) = \text{ggT}(36, 24) = \text{ggT}(24, 12) = \text{ggT}(12, 0) = 12$$

- Für die Zahlen 3 und 14 erhalten wir:

$$\text{ggT}(3, 14) = \text{ggT}(3, 2) = \text{ggT}(2, 1) = \text{ggT}(1, 0) = 1$$

---

Die nun entscheidende Beobachtung, wenn wir den obigen Satz über die Division mit Rest verwenden, ist, dass die ggT-Argumente systematisch ihre Rollen wechseln:

- Das aktuell größere (oder nicht-kleinere) Argument  $M$  wird im nächsten Schritt ersetzt durch das aktuell kleinere (oder nicht-größere).
- Das aktuell kleinere (oder nicht-größere) Argument  $m$  wird im nächsten Schritt ersetzt durch den Divisionsrest  $r = M \bmod m$ .
- Im darauf folgenden Schritt ist das neue  $M$  das vorige  $m$ , und das neue  $m$  das vorige  $r$ .

---

<sup>3</sup> $F_0 := 0, F_1 := 1$ . Für  $n > 1$  gilt:  $F_n := F_{n-1} + F_{n-2}$ . Die ersten Fibonaccizahlen sind: 0, 1, 1, 2, 3, 5, 8, 13, 21, 34, ...

- Insgesamt ergibt sich also ein Muster, das (bis auf den Anfang, den wir gleich noch extra besprechen) für alle Argumente der ggT-Funktion gilt: nämlich der Rollenwechsel

$$\text{Divisionsrest } r \rightsquigarrow \text{ nicht-größeres Argument } m \rightsquigarrow \text{ nicht-kleineres Argument } M$$

- Wenn wir diesen Rollenwechsel Schritt für Schritt (jeder Schritt ist eine Anwendung von Satz 2.19) nachvollziehen, dann ergibt sich für Schritt  $j$ :

$$r_j = M_j \bmod m_j = M_j - q_j \cdot m_j \quad \text{mit} \quad m_j = r_{j-1} \quad \text{und} \quad M_j = m_{j-1} = r_{j-2}$$

Bevor wir das formal zusammen fassen und auch auf die Besonderheiten der ersten beiden Schritte eingehen, zunächst ein

**Beispiel:**  $\text{ggT}(34, 12) = 2$

$j$	$M_j$	$=$	$q_j$	$\cdot$	$m_j$	$+$	$r_j$	Bemerkung
1	34	=	2	$\cdot$	12	+	10	(also $\text{ggT}(34, 12) = \text{ggT}(12, 10)$ )
2	12	=	1	$\cdot$	10	+	2	(also $\text{ggT}(12, 10) = \text{ggT}(10, 2)$ )
3	10	=	5	$\cdot$	2	+	0	(also $\text{ggT}(10, 2) = \text{ggT}(2, 0) = 2$ )

**Bemerkungen:**

- Man beachte, wie die Zahl 10 im ersten Schritt als Rest  $r_1$  auftritt; im zweiten als  $m_2$ , und im dritten als  $M_3$ . Man sieht auch, wie die Zahl 2 im zweiten Schritt als  $r_2$  auftritt und im dritten als  $m_3$ . Gäbe es noch einen vierten Schritt, würde diese 2 dann als  $M_4$  fungieren – aber dies ist hier nicht mehr nötig, da der ggT trivial abgelesen werden kann, wenn eines seiner Argumente null ist.
- Schreibt man die in den Bemerkungen angegebenen Zusammenhänge auf, so ergibt sich, ähnlich wie in den obigen Beispielen, eine Gleichungskette, die die ursprünglichen Argumente der ggT-Funktion lückenlos mit dem Endergebnis verknüpft. Hier:

$$\text{ggT}(34, 12) = \text{ggT}(12, 10) = \text{ggT}(10, 2) = \text{ggT}(2, 0) = 2$$

Zwei Probleme bleiben zu lösen:

- Zunächst stellt sich die Frage, ob das Verfahren für jedes Beispiel funktioniert. Dass es grundsätzlich richtig ist, wissen wir bereits von Satz 2.19 – aber bricht es auch stets nach endlich vielen Schritten ab?

Tatsächlich tut es das – denn nach jedem Schritt  $j$  gilt im nächsten Schritt:

$$r_{j+1} = M_{j+1} \bmod m_{j+1} = M_{j+1} \bmod r_j$$

Diese Zahl liegt zwischen 0 und  $(r_j - 1)$ , ist aber auf jeden Fall echt kleiner als  $r_j$ .

Wenn aber in jedem Schritt der neu berechnete Divisionsrest kleiner ist als der aus dem vorigen Schritt, aber trotzdem nicht negativ wird, muss sich dieser Rest zwangsläufig in Richtung null bewegen. Und da es einen ggT von zwei natürlichen Zahlen stets gibt, wird der Rest null auch tatsächlich erreicht – dann lässt sich das Verfahren abbrechen.

- Wenn wir  $m_j$  und  $M_j$  in Schritt  $j$  auf die Divisionsreste  $r_{j-1}$  und  $r_{j-2}$  zurück führen können – was ist dann mit Schritt 1 und Schritt 2? Dort müsste es dann Reste aus hypothetischen Schritten 0 und (-1) geben, die gar nicht vorliegen.

Diese Situation können wir auf eine von zwei Arten auflösen (wir erinnern uns, dass  $0 < m_1 \leq M_1$  gilt):

- Falls die Argumente der ggT-Funktion gleich sind (oder falls  $M_1$  ein echtes Vielfaches von  $m_1$  ist), ist  $r_1 = 0$  und wir wissen sofort, dass  $\text{ggT}(M_1, m_1) = m_1$  gilt.

- Falls die Argumente  $M_1$  und  $m_1$  anderweitig ungleich sind, können wir in Gedanken zwei vorige Schritte einführen, die mit dem oben besprochenen Muster harmonisieren. Für unser obiges Beispiel wäre etwa möglich:

$$\text{ggT}(80, 46) = \text{ggT}(46, 34) = \text{ggT}(34, 12) = \dots$$

Dann tritt die 34 (also  $M_1$ ) im Schritt 0 als  $m_0$  und im Schritt (-1) als  $r_{-1}$  auf, und die 12 (also  $m_1$ ) im Schritt 0 als  $r_0$ . (Die jeweils neuen Zahlenwerte  $M_0, M_{-1}$  für die hypothetischen zusätzlichen Schritte haben wir durch Summieren der Argumente vom jeweiligen Nachfolgeschritt erhalten.)

Die Tabelle würde mit den hypothetischen Schritten dann so aussehen:

$j$	$M_j$	$=$	$q_j$	$\cdot$	$m_j$	$+$	$r_j$	Bemerkung
-1	80	=	1	·	46	+	34	(also $\text{ggT}(80, 46) = \text{ggT}(46, 34)$ )
0	46	=	1	·	34	+	12	(also $\text{ggT}(46, 34) = \text{ggT}(34, 12)$ )
1	34	=	2	·	12	+	10	(also $\text{ggT}(34, 12) = \text{ggT}(12, 10)$ )
2	12	=	1	·	10	+	2	(also $\text{ggT}(12, 10) = \text{ggT}(10, 2)$ )
3	10	=	5	·	2	+	0	(also $\text{ggT}(10, 2) = \text{ggT}(2, 0) = 2$ )

So lassen sich auch die beiden ursprünglichen ggT-Argumente als Divisionsreste begreifen, die (unter Einbeziehung der beiden hypothetischen Schritte) das obige Muster des Rollenwechsels durchlaufen haben.

In *beiden* Fällen erreichen wir das Gewünschte, wenn wir zu Beginn folgendes fest legen:

$$r_0 := m_1 \quad \text{und} \quad r_{-1} := M_1$$

Die Schrittfolge können wir abbrechen, sobald im Schritt  $k$  für den Divisionsrest gilt:  $r_k = 0$ . Denn dann ist:

$$\underbrace{\text{ggT}(r_{-1}, r_0)}_{\text{ggT}(M_1, m_1)} = \dots = \text{ggT}(\underbrace{r_{k-1}}_{m_k}, \underbrace{r_k}_0) = r_{k-1}$$

Das ist auch dann richtig, wenn sofort im ersten Schritt  $r_1 = 0$  errechnet wird, denn dann ist der ggT entsprechend:  $r_{1-1} = r_0 = m_1$ .

Wir fassen das oben beschriebene Verfahren formal zusammen, um iterativ den ggT von zwei Zahlen zu berechnen:

**Satz 2.20** (Euklidischer Algorithmus). *Für  $a, b \in \mathbb{N}$  berechnet sich  $\text{ggT}(a, b)$  mit den folgenden Schritten:*

- Initialisierung:*
  - Definiere  $r_{-1} := \max\{a, b\}$  und  $r_0 := \min\{a, b\}$
  - Setze Schrittzähler  $j \in \mathbb{N}$  auf  $j := 1$
- Rechenschritt  $j$  für  $j \in \mathbb{N}$ :*
  - Definiere  $M_j := r_{j-2}$  und  $m_j := r_{j-1}$
  - Berechne  $r_j := M_j \bmod m_j$

Dann ist

$$\text{ggT}(\underbrace{r_{j-2}}_{M_j}, \underbrace{r_{j-1}}_{m_j}) = \text{ggT}(\underbrace{r_{j-1}}_{m_j}, r_j)$$

- Update des Schrittzählers:  $j := j + 1$
  - Abbruch, sobald  $r_k = 0$
- Dann ist

$$\text{ggT}(a, b) \left[ = \text{ggT}(r_{-1}, r_0) = \dots = \text{ggT}(r_{k-2}, r_{k-1}) = \text{ggT}(r_{k-1}, 0) \right] = r_{k-1} = m_k$$



### Bemerkungen:

- (Der historische Algorithmus von Euklid verzichtete auf das Modulo und betrachtete nur die Invarianz des ggT unter Differenzbildung – bei der Modulo-Variante werden dagegen möglichst viele solche Differenzen in einem Schritt zusammen behandelt.)
- Die Anwendung des Algorithmus ist am besten nachvollziehbar, wenn die Bestandteile der einzelnen Schritte tabellarisch notiert werden, siehe die Beispiele.

Die Kommentarspalte rechts ist nicht zwingend erforderlich und wird dort nur notiert, um die Gleichungskette explizit verfolgbar zu machen. Auch auf die explizite Nummerierung der Schritte wird meist verzichtet.

Die hypothetischen Schritte 0 und (-1) werden hier nicht mit angegeben.

### Beispiele:

- Wir behandeln zunächst das erste Beispiel von Satz 2.18 und berechnen  $\text{ggT}(24, 4) = 4$ :

$j$	$M_j$	$=$	$q_j$	$\cdot$	$m_j$	$+$	$r_j$	Bemerkung
1	24	$=$	6	$\cdot$	4	$+$	0	(also $\text{ggT}(24, 4) = \text{ggT}(4, 0) = 4$ )

- Nun das zweite (ungünstige) Beispiel vom selben Satz (s.o):  $\text{ggT}(21, 13) = 1$

$j$	$M_j$	$=$	$q_j$	$\cdot$	$m_j$	$+$	$r_j$	Bemerkung
1	21	$=$	1	$\cdot$	13	$+$	8	(also $\text{ggT}(21, 13) = \text{ggT}(13, 8)$ )
2	13	$=$	1	$\cdot$	8	$+$	5	(also $\text{ggT}(13, 8) = \text{ggT}(8, 5)$ )
3	8	$=$	1	$\cdot$	5	$+$	3	(also $\text{ggT}(8, 5) = \text{ggT}(5, 3)$ )
4	5	$=$	1	$\cdot$	3	$+$	2	(also $\text{ggT}(5, 3) = \text{ggT}(3, 2)$ )
5	3	$=$	1	$\cdot$	2	$+$	1	(also $\text{ggT}(3, 2) = \text{ggT}(2, 1)$ )
6	2	$=$	2	$\cdot$	1	$+$	0	(also $\text{ggT}(2, 1) = \text{ggT}(1, 0) = 1$ )

- Und noch ein Beispiel mit etwas größeren Zahlen, wo das Auffinden der Teiler-Mengen länger dauern könnte:  $\text{ggT}(969, 627)$ .

Hier in der minimalen (*mindestens so auch in der Prüfung erwarteten*) Form, ohne Schrittzählung oder Bemerkungen (wegen der fehlenden Bemerkungen ist dann ein Antwortsatz erforderlich):

$M$	$=$	$q$	$\cdot$	$m$	$+$	$r$
969	$=$	1	$\cdot$	627	$+$	342
627	$=$	1	$\cdot$	342	$+$	285
342	$=$	1	$\cdot$	285	$+$	57
285	$=$	5	$\cdot$	57	$+$	0

Also:  $\text{ggT}(969, 627) = 57$

## 2.3 Anwendungen

Wir befassen uns zum Abschluss des Kapitels noch mit drei wichtigen Resultaten: hauptsächlich mit dem Lemma von Bezout<sup>4</sup> (das uns erlaubt, den ggT von zwei Zahlen als ganzzahlige Linearkombination dieser Zahlen darzustellen), dem Lemma von Euklid (das eine Aussage zur Teilbarkeit von Faktoren eines Produkts liefert), und einer Betrachtung der Teilerfremdheit bei Produkten aus mehreren Zahlen. Alle diese Tatsachen benötigen wir später im Kapitel 3.

### 2.3.1 Lemma von Bezout

Wir wollen versuchen, den ggT als ganzzahlige Linearkombination seiner beiden Argumente darzustellen, d.h. wir suchen  $\alpha, \beta \in \mathbb{Z}$ , sodass gilt:

$$\text{ggT}(M_1, m_1) = \alpha \cdot M_1 + \beta \cdot m_1$$

Wir haben hierbei  $M_1$  und  $m_1$  so verwendet wie oben beim euklidischen Algorithmus vereinbart – dies ist hier hilfreich, weil sich über  $M_1 \geq m_1$  klar ergibt, welche der beiden Zahlen die kleinere (oder wenigstens: nicht-größere) ist.

---

<sup>4</sup>E. Bezout, frz. Mathematiker

Wir werden jetzt konstruktiv zeigen, dass sich solche Zahlen  $\alpha, \beta$  strukturiert finden lassen. Zum besseren Verständnis werden wir aber auch das erste Beispiel ausführlich durchrechnen; erst danach geben wir die schematisch berechnete Tabelle des *erweiterten euklidischen Algorithmus* an.

Hier lässt es sich nicht vermeiden, erst einmal “in Buchstaben” (also mit Unbekannten) zu rechnen – allerdings hat das auch den Vorteil, dass wir nicht in die Versuchung kommen, zwischendurch schon Zahlenwerte einzusetzen, weil sie (vom konkreten Beispiel her) “so günstig” wirken.

---

Falls nun  $M_1$  ein Vielfaches von  $m_1$  ist, terminiert der Algorithmus sofort, und dann wäre  $\text{ggT}(M_1, m_1) = m_1$ . In diesem Fall wäre die Linearkombination (zur besseren Lesbarkeit sind die Faktoren unterstrichen):

$$\text{ggT}(M_1, m_1) = m_1 = \underline{0} \cdot M_1 + \underline{1} \cdot m_1$$

Falls aber  $M_1$  nur mit nichtverschwindendem Rest  $r_1$  durch  $m_1$  dividierbar ist, benötigt der Algorithmus mindestens einen weiteren Schritt. Wie oben betrachten wir hier die Divisionsreste  $r_j$ , die in solchen Schritten jeweils berechnet werden.

Außerdem ist für uns jetzt noch wichtig, wie sich die  $r_j$  in den Zahlen  $M_1$  und  $m_1$  ausdrücken lassen. Für den ersten Schritt geht dies noch recht einfach:

$$r_1 = M_1 \bmod m_1 = M_1 - q_1 \cdot m_1 = \underline{1} \cdot M_1 + \underline{(-q_1)} \cdot m_1 \quad (*)$$

---

Falls weitere Schritte nötig sind, wird es etwas komplizierter. Wir wissen allerdings schon, dass wir die  $r_j$  über den euklidischen Algorithmus einheitlich berechnen können – es liegt also nahe, dasselbe parallel auch mit den Faktoren  $\alpha_j, \beta_j$  zu versuchen. Hier zahlt sich aus, dass wir weiter oben schon die hypothetischen Schritte 0 und (-1) eingeführt hatten, es ist nämlich:

$$r_j = M_j \bmod m_j = r_{j-2} \bmod r_{j-1} = r_{j-2} - q_j \cdot r_{j-1} \quad (**)$$

Falls in den beiden vorigen Schritten also die Reste bereits als Linearkombinationen ausgedrückt waren, lässt sich damit direkt die Linearkombination von  $r_j$  berechnen, wobei

$$r_j = \alpha_j \cdot M_1 + \beta_j \cdot m_1 \quad (***)$$

Mit der rechten Gleichung von (\*\*) können wir nun die Gleichung (\*\*\*) durch die in den beiden vorigen Schritten aufgesammelten Zahlenwerte ausdrücken. Wir setzen also ein (\*\*) sowohl für  $r_{j-2}$  als auch für  $r_{j-1}$  in (\*\*) ein. Danach stellen wir um, sodass wir wieder eine Linearkombination von  $M_1$  und  $m_1$  erhalten (durch Ausklammern und Gruppieren der Vorfaktoren nach dem Distributivgesetz):

$$\begin{aligned} r_j &= r_{j-2} - q_j \cdot r_{j-1} \\ &= (\alpha_{j-2} \cdot M_1 + \beta_{j-2} \cdot m_1) - q_j \cdot (\alpha_{j-1} \cdot M_1 + \beta_{j-1} \cdot m_1) \\ &= \underbrace{(\alpha_{j-2} - q_j \cdot \alpha_{j-1})}_{=:\alpha_j} \cdot M_1 + \underbrace{(\beta_{j-2} - q_j \cdot \beta_{j-1})}_{=:\beta_j} \cdot m_1 \\ &= \alpha_j \cdot M_1 + \beta_j \cdot m_1 \end{aligned}$$

---

Damit haben wir über den euklidischen Algorithmus schon die Methode gefunden, die Vorfaktoren für Schritt  $j$  aus den beiden vorigen Schritten zu berechnen – und dies geht offenbar stets nach dem gleichen Schema (s.o.). Nimmt man also in der Tabelle noch die Spalten für  $\alpha_j$  und  $\beta_j$  pro Schritt auf, so handelt es sich um den *erweiterten euklidischen Algorithmus*.

---

Genau wie beim einfachen euklidischen Algorithmus brauchen wir aber noch den “Ansatzpunkt”, um beginnen zu können – denn auch Schritt 1 benötigt für die obige Formel zwei vorige Schritte. Wie eben schon erwähnt, sind dies aber gerade die schon bekannten Schritte 0 und (-1). Wir stellen für diese beiden Schritte also noch die gesuchten Linearkombinationen auf – sie gelten in dieser Form stets:

$$\begin{aligned} r_{-1} &= M_1 = \underline{1} \cdot M_1 + \underline{0} \cdot m_1 &=: \alpha_{-1} \cdot M_1 + \beta_{-1} \cdot m_1 \\ r_0 &= m_1 = \underline{0} \cdot M_1 + \underline{1} \cdot m_1 &=: \alpha_0 \cdot M_1 + \beta_0 \cdot m_1 \end{aligned}$$

Für die ersten beiden (fest definierten) Reste erhalten wir also  $\alpha_{-1} = 1$ ,  $\beta_{-1} = 0$ ,  $\alpha_0 = 0$  und  $\beta_0 = 1$ .

Nun berechnen wir zur Probe die Werte  $\alpha_1$  und  $\beta_1$  über obige Formel und gleichen dies mit dem Ergebnis aus (\*) ab:

$$\begin{aligned}\alpha_1 &= \alpha_{-1} - q_1 \cdot \alpha_0 = 1 - q_1 \cdot 0 = 1 \quad \checkmark \\ \beta_1 &= \beta_{-1} - q_1 \cdot \beta_0 = 0 - q_1 \cdot 1 = (-q_1) \quad \checkmark\end{aligned}$$

Wir sehen also: Der Divisionsrest von Schritt  $j$  (mit  $j \in \mathbb{N}$ ) ergibt sich als Linearkombination aus  $M_1$  und  $m_1$  immer über die Linearfaktoren der vorigen beiden Schritte; außerdem geht dabei der Quotient aus der Ganzzahldivision  $q_j$  mit ein, der im einfachen euklidischen Algorithmus (bei dem es nur um den Wert des ggT geht) streng genommen nicht wichtig war.

Wir erhalten allgemein:

$$r_j = \underbrace{(\alpha_{j-2} - q_j \cdot \alpha_{j-1})}_{=: \alpha_j} \cdot M_1 + \underbrace{(\beta_{j-2} - q_j \cdot \beta_{j-1})}_{=: \beta_j} \cdot m_1$$

Dies gilt auch für  $r_k = 0$  beim Abbruch des euklidischen Algorithmus, sowie für  $r_{k-1}$ , den Wert des ggT. Die Anwendung des erweiterten euklidischen Algorithmus in Kombination mit den dabei ermittelten Quotienten  $q_j$  liefert uns also die Linearkombination des ggT( $M_1, m_1$ ) direkt mit. Daher gilt der folgende

**Satz 2.21** (Lemma von Bezout). *Für  $a, b \in \mathbb{Z}$  gibt es stets Zahlen  $\alpha, \beta \in \mathbb{Z}$ , sodass gilt:*

$$\text{ggT}(a, b) = \alpha \cdot a + \beta \cdot b$$

**Beispiele:**

- Wir berechnen den ggT von 34 und 15 als Linearkombination dieser Zahlen zunächst ohne Kenntnis des erweiterten euklidischen Algorithmus (sondern nur mit der einfachen Variante).

Für den euklidischen Algorithmus ergibt sich:

$M$	$=$	$q$	$\cdot$	$m$	$+$	$r$
34	$=$	2	$\cdot$	15	$+$	4
15	$=$	3	$\cdot$	4	$+$	3
4	$=$	1	$\cdot$	3	$+$	1
3	$=$	3	$\cdot$	1	$+$	0

Nun sammeln wir die Linearfaktoren auf:

- Der erste Rest ist:

$$4 = 34 - 2 \cdot 15 = \underline{1} \cdot 34 + \underline{(-2)} \cdot 15$$

- Dann als nächstes:

$$3 = 15 - 3 \cdot 4$$

Die auftretende 4 ersetzen wir durch die eben gefundene Linearkombination:

$$3 = 15 - 3 \cdot 4 = 15 - 3 \cdot (34 - 2 \cdot 15) = \underline{(-3)} \cdot 34 + \underline{7} \cdot 15$$

- Zum Schluss noch:

$$\text{ggT}(34, 15) = 1 = 4 - 1 \cdot 3$$

Hier ist die 4 mit der Linearkombination vom vorvorigen Schritt zu ersetzen, die 3 jedoch mit der vom vorigen Schritt. Damit:

$$\begin{aligned}\text{ggT}(34, 15) &= 1 = 4 - 1 \cdot 3 \\ &= (34 - 2 \cdot 15) - 1 \cdot ((-3) \cdot 34 + 7 \cdot 15) \\ &= \underline{4} \cdot 34 + \underline{(-9)} \cdot 15\end{aligned}$$

Man erkennt hoffentlich schon für dieses Beispiel, dass einige Sorgfalt nötig war, um im jeweiligen Schritt die richtigen Ersetzungen vorzunehmen. Eine ähnliche Variante dieser Rechnung, das *Rückwärtseinsetzen*, ist ähnlich fehleranfällig – wenn man mit konkreten Zahlen hantiert, ist es schwierig, zu erkennen, ob man nun ein  $m$  oder ein  $q$  durch einen vorigen Divisionsrest ersetzt – nur ersteres wäre richtig. In der zweiten Zeile bemerken wir, dass  $q_2 = r_2 = 3$  gilt, aber der Wert  $q_2$  sollte während der Rechnung *nicht* durch  $r_2$  ersetzt werden – das würde potentiell zu quadratischen Termen der ursprünglichen ggT-Argumente führen, obwohl wir nur eine lineare Kombination suchen.

- Wir berechnen den selben ggT also nochmals mit dem erweiterten euklidischen Algorithmus (tabellarisch). Die künstlichen Schritte 0 und (-1) sind nur in sofern notiert, als sie für die Berechnung der Vorfaktoren nötig sind. In den (optionalen) Kommentarspalten sind die Rechnungen für  $\alpha, \beta$  aufgeschlüsselt.

Die eckig eingeklammerten Einträge in der letzten Zeile sind überflüssig, wenn nur die Linearkombination des ggT gesucht ist, aber als Proberechnung (Linearkombination der Null) nicht völlig sinnfrei.

$M = q \cdot m + r$	$\alpha$	$\beta$	Rechnung $\alpha$	Rechnung $\beta$
34	1	0		
15	0	1		
34 = 2 · 15 + 4	1	-2	1 = 1 - 2 · 0	-2 = 0 - 2 · 1
15 = 3 · 4 + 3	-3	7	-3 = 0 - 3 · 1	7 = 1 - 3 · (-2)
4 = 1 · 3 + 1	4	-9	4 = 1 - 1 · (-3)	-9 = -2 - 1 · 7
3 = 3 · 1 + 0	[-15]	[34]	[-15 = -3 - 3 · 4]	[34 = 7 - 3 · (-9)]

Wie man an den Spalten  $\alpha, \beta$  entnehmen kann, werden durch die systematische Rechnung die gleichen Koeffizienten der Linearkombinationen der jeweiligen Reste (aus der Spalte  $r$ ) gefunden wie im vorigen händisch durchgerechneten Beispiel. Insbesondere:

$$1 = 4 \cdot 34 - 9 \cdot 15 = 136 - 135 \quad \checkmark$$

Auch der letzte (verschwindende) Rest lässt sich schreiben als

$$0 = (-15) \cdot 34 + 34 \cdot 15 \quad \checkmark$$

Das ist zwar wenig überraschend – aber man bedenke, dass die Vorfaktoren sich hier nicht durch direkte Anschauung, sondern durch systematische Rechnung ergeben haben.

Man beachte in den Kommentarspalten noch, wie die Rechnungen den  $q_j$ -Wert der jeweiligen Zeile verwenden (nach den Minuszeichen finden sich jeweils die Werte 2,3,1,3).

- Wir berechnen ggT(927, 104):

$M = q \cdot m + r$	$\alpha$	$\beta$
927	1	0
104	0	1
927 = 8 · 104 + 95	1	-8
104 = 1 · 95 + 9	-1	9
95 = 10 · 9 + 5	11	-98
9 = 1 · 5 + 4	-12	107
5 = 1 · 4 + 1	23	-205
4 = 4 · 1 + 0	[-104]	[927]

Also ist ggT(927, 104) = 1. Zur Probe die Linearkombination:

$$23 \cdot 927 + (-205) \cdot 104 = 21321 - 21320 = 1 \quad \checkmark$$

Zum Vergleich die ggT-Berechnung nochmal über die Teilmengen:

$$T_{927} = \{1, 3, 9, 103, 309, 927\}$$

$$T_{104} = \{1, 2, 4, 8, 13, 26, 52, 104\}$$

Auch hier ergibt sich der ggT als 1; eine Aussage zur Linearkombination des ggT gibt es hier natürlich nicht.

### 2.3.2 Lemma von Euklid

**Satz 2.22** (Lemma von Euklid). Für  $a, b \in \mathbb{Z}$  und  $n \in \mathbb{N}$  gilt:

$$\left( (n \mid (a \cdot b)) \wedge (\text{ggT}(a, n) = 1) \right) \Rightarrow (n \mid b)$$

(Beweis: S. 316.)

#### Bemerkungen:

- Die ursprünglich von Euklid formulierte Fassung war spezieller und lautet:

Für  $p$  prim,  $p \mid (a \cdot b)$  und  $p \nmid a$  gilt:  $p \mid b$ .

Teilt also eine Primzahl ein Produkt zweier Zahlen, aber nicht den einen Faktor, so muss diese Primzahl den anderen Faktor teilen.

Dies ist ein Spezialfall der obigen allgemeineren Form, die nicht nur Primzahlen  $p$ , sondern beliebige natürliche Zahlen  $n$  zulässt. Denn für  $n = p$  prim gilt, falls  $p \nmid a$ , nach Satz 2.12 (Eigenschaften des ggT zweier Zahlen) auch  $\text{ggT}(a, n) = \text{ggT}(a, p) = 1$ .

- Der Beweis der allgemeineren Fassung ist mit dem Lemma von Bezout sehr leicht machbar.

**Beispiel:** Für  $a = 17$  und  $b = 6$ , also  $a \cdot b = 102$  gilt:  $2 \mid (a \cdot b)$ , aber  $\text{ggT}(a, 2) = \text{ggT}(17, 2) = 1$ . Somit teilt 2 den anderen Faktor, also  $2 \mid b$  ✓

### 2.3.3 Teilerfremdheit bei Produkten ganzer Zahlen

**Satz 2.23** (Teilerfremdheit eines Produktes mit einer Zahl).  $m$  sei ein Produkt aus  $k \in \mathbb{N}$  ganzen Zahlen  $a_1, a_2, \dots, a_k$ , also

$$m = \prod_{j=1}^k a_j$$

Dann gilt für  $n \in \mathbb{N}$ :

$$\left( (\text{ggT}(a_1, n) = 1) \wedge (\text{ggT}(a_2, n) = 1) \wedge \dots \wedge (\text{ggT}(a_k, n) = 1) \right) \Leftrightarrow (\text{ggT}(m, n) = 1)$$

(Beweis: S. 316.)

#### Bemerkungen:

- Ein Produkt von Zahlen ist also genau dann teilerfremd mit  $n$ , wenn sämtliche seine Faktoren ebenfalls teilerfremd mit  $n$  sind.
- Speziell für  $k = 2$  gilt: Ein Produkt zweier Zahlen ist also genau dann teilerfremd mit  $n$ , wenn beide Faktoren jeweils teilerfremd mit  $n$  sind.

Wegen der Äquivalenz gilt auch: Sobald ein Produkt zweier Zahlen mit  $n$  einen gemeinsamen Teiler größer als 1 besitzt, muss mindestens einer der beiden Faktoren ebenfalls einen mit  $n$  gemeinsamen Teiler größer als 1 haben.

- Betrachtet man weiterhin beliebige ganze Zahlen  $m$  als Produkt zweier komplementärer Teiler, so ist  $m$  genau dann teilerfremd zu  $n$ , wenn beide diese Teiler ebenfalls teilerfremd zu  $n$  sind. Und da die Auswahl der beiden Teiler nicht weiter eingeschränkt war, gilt die Aussage demnach für sämtliche Teiler von  $m$ .

#### Beispiele:

- Wir betrachten  $a_1 := 103$  und  $a_2 := 9$  sowie  $n := 104$ . Das Produkt  $m := a_1 \cdot a_2$  ist 927. Wir wissen von oben (Beispiele zum Lemma von Bezout) schon, dass  $\text{ggT}(927, 104) = 1$ . Dann haben wir noch nachzurechnen, dass die beiden Faktoren  $a_1, a_2$  jeweils teilerfremd mit 104 sind.

Zunächst ist nach Satz 2.18:

$$\text{ggT}(103, 104) = \text{ggT}(103, (104 - 103)) = \text{ggT}(103, 1) = 1 \quad \checkmark$$

Für die andere Rechnung ( $\text{ggT}(104, 9)$ ) verwenden wir Euklid:

$$\begin{array}{rclclcl}
 M & = & q & \cdot & m & + & r \\
 \hline
 104 & = & 11 & \cdot & 9 & + & 5 \\
 9 & = & 1 & \cdot & 5 & + & 4 \\
 5 & = & 1 & \cdot & 4 & + & 1 \\
 4 & = & 4 & \cdot & 1 & + & 0
 \end{array}$$

Also auch  $\text{ggT}(104, 9) = 1$  ✓

- Wir betrachten  $a_1 := 2$  und  $a_2 := 26$  sowie  $n := 28$ . Das Produkt  $m := a_1 \cdot a_2$  ist 52. Nun ist  $\text{ggT}(52, 28) = 4 > 1$ . Nach Aussage des Satzes muss kann  $n = 28$  also nicht mit beiden Faktoren  $a_1$  und  $a_2$  teilerfremd sein. Und tatsächlich ist  $\text{ggT}(2, 28) = 2 > 1$ , und auch  $\text{ggT}(26, 28) = 2 > 1$ .
- Wir betrachten eine Primzahl  $p$  und das Produkt

$$m := (p-1)! = \prod_{j=1}^{p-1} j$$

Da  $p$  prim ist, sind alle Faktoren  $j$  des Produkts  $m$  teilerfremd zu  $p$  (denn zwischen  $p$  und 1 gibt es keine weiteren Teiler von  $p$ ). Nach dem obigen Satz ist (mit  $n := p$ ) damit auch:

$$\text{ggT}((p-1)!, p) = 1$$

# Relationen

Dieses Kapitel wird in der Vorlesung Theoretische Informatik 1 behandelt und hier voraus gesetzt. Erwartet werden folgende Konzepte:

- Zweistellige Relationen allgemein; speziell Relationen auf einer Menge
- Symmetrie, Reflexivität, Transitivität bei Relationen auf einer Menge
- Äquivalenzrelationen auf einer Menge
- Äquivalenzklassen
- Partitionierung einer Grundmenge in disjunkte Äquivalenzklassen mit einer Äquivalenzrelation (Konzept)

Möglicherweise werden die obigen Inhalte später noch in diesem Skript ergänzt, aber nicht während der Vorlesung Mathematik 1 behandelt.

# Kapitel 3

## Restklassen

Wir betrachten in diesem Kapitel eine wichtige Äquivalenzrelation auf der Menge der ganzen Zahlen  $\mathbb{Z}$ : Die Äquivalenz modulo  $m$  (mit  $m \in \mathbb{N}$ ). Zur Erinnerung verweisen wir auf Satz 2.7 (Division mit Rest). Wenn der Modul  $m$  fest gewählt wird, hat jede ganze Zahl  $n$  einen eindeutig bestimmten Divisionsrest  $r$  modulo  $m$ , der zwischen 0 und  $(m - 1)$  liegt.

Zwei Zahlen sind genau dann äquivalent modulo  $m$ , wenn ihre eindeutigen Divisionsreste modulo  $m$  gleich sind. Und das sind sie genau dann, wenn sich ihre Differenz als Vielfaches des Moduls schreiben lässt.

### 3.1 Äquivalenz modulo $m$

**Definition 3.1** (Äquivalenz modulo  $m$ ). Sei  $m \in \mathbb{N}$  gegeben und fest. Zwei Zahlen  $a, b \in \mathbb{Z}$  heißen äquivalent modulo  $m$ , geschrieben

$$a \equiv b \pmod{m}$$

genau dann, wenn  $m \mid (b - a)$ .

#### Bemerkungen:

- Das Symbol  $\equiv$  bezeichnet hier etwas völlig anderes als die logische Äquivalenz von Aussagen – es gilt allerdings auch nur für ganze Zahlen (oder Ausdrücke, die einen Zahlenwert aus  $\mathbb{Z}$  besitzen). Trotz der überladenen Notation sollte eine Verwechslung dadurch ausgeschlossen sein.
- Zur Vorbemerkung dieses Kapitels: Die Relation  $\equiv$  lässt sich auch gleichwertig mit den Divisionsresten formulieren. Sei  $a = q_a \cdot m + r_a$  und  $b = q_b \cdot m + r_b$ . Dann gilt:

$$\begin{aligned} a &\equiv b \pmod{m} \\ \Leftrightarrow m &\mid (b - a) \\ \Leftrightarrow m &\mid (q_b \cdot m + r_b - (q_a \cdot m + r_a)) \\ \Leftrightarrow m &\mid ((q_b - q_a) \cdot m + (r_b - r_a)) \\ \Leftrightarrow m &\mid (r_b - r_a) \end{aligned}$$

Da aber  $0 \leq r_a, r_b < m$ , kann die Differenz  $(r_b - r_a)$  nur zwischen den Werten  $-(m - 1)$  und  $(m - 1)$  liegen. Das einzige Vielfache von  $m$  in dieser Zahlenmenge ist jedoch  $0 = 0 \cdot m$ . Also können wir obige (logische) Äquivalenzkette fortführen:

$$\begin{aligned} \cdots &\Leftrightarrow (r_b - r_a) = 0 \\ &\Leftrightarrow r_a = r_b \end{aligned}$$

Die beiden Kriterien für die Äquivalenz modulo  $m$  sind also gleichwertig.

#### Beispiele:

- $9 \equiv 3 \pmod{6}$ , da  $(9 - 3) = 6 = 1 \cdot 6$
- $14 \equiv 0 \pmod{7}$ , da  $(14 - 0) = 14 = 2 \cdot 7$



- $-12 \equiv 4 \pmod{16}$ , da  $(-12 - 4) = -16 = (-1) \cdot 16$
- $3 \equiv 3 \pmod{8}$ , da  $(3 - 3) = 0 = 0 \cdot 8$
- Jede Zahl  $n \in \mathbb{Z}$  ist äquivalent zu  $0 \pmod{1}$ , da sie ohne Rest durch 1 teilbar ist:

$$(n - 0) = n = n \cdot 1$$


---

Wir haben allerdings noch zu zeigen, dass die Relation " $\equiv \pmod{m}$ " eine Äquivalenzrelation ist, wie oben in Definition 3.1 schon suggeriert wurde:

**Satz 3.2** (Äquivalenzrelation " $\equiv \pmod{m}$ "). *Sei  $m \in \mathbb{N}$  gegeben und fest. Dann ist die Relation  $\equiv \pmod{m}$  eine Äquivalenzrelation auf  $\mathbb{Z}$ .*

(Beweis: S. 317)

**Bemerkung:** Erinnerung:

- Unter Äquivalenzrelationen zerfällt die Grundmenge vollständig in disjunkte Teilmengen von jeweils zueinander in Relation stehenden (und damit hier auch äquivalenten) Elementen.
  - Elemente verschiedener Äquivalenzklassen stehen jedoch niemals in Relation zueinander.
  - Jedes Element einer Äquivalenzklasse darf als gleichwertiger Repräsentant der gesamten Klasse angesehen werden.
- 

**Definition 3.3** (Restklassen und Restsystem modulo  $m$ ). *Für gegebenes festes  $m \in \mathbb{N}$  heißen die Äquivalenzklassen der Relation  $\equiv \pmod{m}$  Restklassen modulo  $m$ . Die Menge aller Äquivalenzklassen heißt Rest(klassen-)system modulo  $m$ , geschrieben als*

$$\mathbb{Z}_m$$


---

Wir wissen schon, dass es genau  $m$  verschiedene Divisionsreste modulo  $m$  gibt, nämlich die Zahlen 0 bis  $(m - 1)$ . Da äquivalente ganze Zahlen alle jeweils den gleichen Divisionsrest aufweisen, liegen damit alle verschiedenen Divisionsreste in jeweils eigenen Restklassen, nämlich in

$$[0], [1], \dots, [m - 1]$$

Wir halten diese Tatsache fest im folgenden

**Satz 3.4** (Restsystem modulo  $m$ ). *Das Restsystem modulo  $m$  enthält für ein beliebiges festes  $m \in \mathbb{N}$  genau  $m$  Restklassen. Die eindeutigen Repräsentanten dieser Klassen sind die Divisionsreste*

$$0, 1, \dots, (m - 1)$$

**Beispiele:**

- Für den Modul  $m = 7$  ist

$$[3] = \{x \in \mathbb{Z} \mid x \equiv 3 \pmod{7}\} = \{x \in \mathbb{Z} \mid x \bmod 7 = 3\} = \{\dots, -11, -4, 3, 10, 17, \dots\}$$

Damit ist übrigens auch  $[3] = [-4] = [10]$  usw.

- Für den selben Modul ist die Menge der Vielfachen von 7 gerade die Restklasse

$$[0] = \{\dots, -14, -7, 0, 7, 14, 21, \dots\}$$

## 3.2 Rechenoperationen mit Restklassen

Wir betrachten zunächst die Addition und Subtraktion von Restklassen. Dabei stellen wir fest, dass man mit den Restklassen quasi genauso rechnen kann wie mit gewöhnlichen ganzen Zahlen, solange nur stets bedacht wird, dass die Rechenergebnisse modulo  $m$  genommen werden.

Streng genommen entspricht die Addition von zwei Restklassen aber einer anderen Operation als die Addition zweier ganzer Zahlen. Wir führen deswegen dafür zunächst eigene Notation ein – wegen der sehr einfachen Korrespondenz zwischen ganzen Zahlen und ihren Restklassen gestatten wir uns danach aber eine gewisse Ungenauigkeit, indem wir für die Restklassen die gleichen Operatoren wie für ganze Zahlen benutzen.

### 3.2.1 Addition und Subtraktion modulo $m$

**Definition 3.5** (Addition und Subtraktion von Restklassen). Sei  $m \in \mathbb{N}$  gegeben und fest. Für Restklassen  $[a], [b] \in \mathbb{Z}_m$  ist eine Addition  $\oplus$  definiert per

$$[a] \oplus [b] := [a + b]$$

Analog eine Subtraktion  $\ominus$  per

$$[a] \ominus [b] := [a] \oplus [-b] = [a - b]$$

#### Bemerkungen:

- Man beachte, dass die Operationen  $\oplus$  und  $\ominus$  die Addition und Subtraktion von Restklassen bezeichnen, also von ganzen Mengen mit jeweils unendlich vielen Elementen. Die Operationen  $+$  und  $-$  beziehen sich hingegen auf einzelne ganze Zahlen.
- Allerdings zeigt die Definition deutlich, wie die neuen Operationen direkt (und in erwarteter Weise) auf die bekannten Operationen für ganze Zahlen zurück geführt werden – lediglich mit dem Unterschied, dass am Ende der Rechenoperation noch die jeweilige Restklasse des Rechenergebnisses zu nehmen ist.
- Auch über die Division mit Rest ergibt sich die gleiche Verträglichkeit. Für  $a = q_a \cdot m + r_a$  und  $b = q_b \cdot m + r_b$  erhalten wir:

$$a + b = (q_a + q_b) \cdot m + (r_a + r_b) \equiv (r_a + r_b) \pmod{m}$$

Oder anders formuliert:

$$(a + b) \bmod m = (r_a + r_b) \bmod m$$

(Hier wurde  $\bmod$  jedoch als Rechenoperator benutzt, nicht um die Äquivalenzrelation zu spezifizieren!)

#### Beispiele:

- Für den Modul  $m = 7$ :
  - $[3] \oplus [2] = [3 + 2] = [5]$
  - $[3] \oplus [4] = [3 + 4] = [7] = [0]$
  - $[3] \oplus [5] = [3 + 5] = [8] = [1]$
  - $[3] \ominus [4] = [3 - 4] = [-1] = [6]$

Wir können gleichwertig so formulieren:

- $3 + 2 = 5 \equiv 5 \pmod{7}$
- $3 + 4 = 7 \equiv 0 \pmod{7}$
- $3 + 5 = 8 \equiv 1 \pmod{7}$
- $3 - 4 = (-1) \equiv 6 \pmod{7}$

Dabei haben wir jeweils die Äquivalenz mit den eindeutigen Repräsentanten notiert, also den Divisionsresten zwischen 0 und 6.

- Für den gleichen Modul  $m = 7$ :

$$17 + 25 = 42 \equiv 0 \pmod{7}$$

Beziehungsweise, mit Modulo-Operator:

$$(17 + 25) \bmod 7 = 42 \bmod 7 = 0$$

Wir erlauben nun die synonyme Verwendung von Restklassen und Divisionsresten. Dadurch wird auch folgende Schreibweise zulässig (ohne die eckigen Klammern, die die Äquivalenzklassen angeben):

$$\mathbb{Z}_m = \{0, 1, \dots, m-1\}$$

Wenn wir jedoch  $\mathbb{Z}_m$  so auffassen, wollen wir uns eindeutig auf die Divisionsreste zwischen 0 und  $(m-1)$  beschränken, nicht auf andere dazu äquivalente ganze Zahlen.

### 3.2.2 Multiplikation modulo $m$

Für die Multiplikation können wir ähnlich verfahren wie für die Addition:

**Definition 3.6** (Multiplikation von Restklassen). *Sei  $m \in \mathbb{N}$  gegeben und fest. Für Restklassen  $[a], [b] \in \mathbb{Z}_m$  ist ein Produkt  $\odot$  definiert per*

$$[a] \odot [b] := [a \cdot b]$$

**Bemerkungen:**

- Auch hier liegt streng genommen nicht die gleiche Operation vor wie bei der Multiplikation ganzer Zahlen, aber die Verträglichkeit ist über die Definition wie oben ersichtlich.
- Über die Division mit Rest mit  $a = q_a \cdot m + r_a$  und  $b = q_b \cdot m + r_b$  erhalten wir:

$$a \cdot b = (q_a \cdot m + r_a) \cdot (q_b \cdot m + r_b) = (q_a \cdot q_b \cdot m + r_a + r_b) \cdot m + (r_a \cdot r_b) \equiv (r_a \cdot r_b) \pmod{m}$$

- Daraus ergibt sich auch die folgende Formel mit Modulo-Operatoren:

$$(a \cdot b) \bmod m = (r_a \cdot r_b) \bmod m = ((a \bmod m) \cdot (b \bmod m)) \bmod m$$

- Zusammen mit der Addition gelten für Restklassen analog die Kommutativ- und Assoziativgesetze (jeweils für beide Rechenarten; d.h. innerhalb von Summen können die Summanden beliebig vertauscht oder geklammert werden; innerhalb von Produkten gilt das gleiche für die Faktoren) sowie das Distributivgesetz. Diese Zusammenhänge beleuchten wir nochmal genauer im Algebra-Kapitel 5.

**Beispiele:**

- Für den Modul  $m := 5$ :
  - $4 \cdot 7 = 28 \equiv 3 \pmod{5}$
  - $4 \cdot 7 \equiv 4 \cdot 2 = 8 \equiv 3 \pmod{5}$
  - $13 \cdot 29 = 377 = 375 + 2 \equiv 2 \pmod{5}$
  - $13 \cdot 29 \equiv 3 \cdot 4 = 12 \equiv 2 \pmod{5}$

Man beachte, dass in einer einzigen Zeile mehrere (modulo-)Äquivalenzzeichen und Gleichheitszeichen vorkommen dürfen. Der Modul muss für solch eine Kette nur einmal am Ende spezifiziert werden (das setzt aber voraus, dass die letzte verkettete Aussage eine modulo-Äquivalenzaussage ist). Die Umformungen der Terme, die mit Gleichheitszeichen notiert sind, sind allerdings nur vorgesehen für Operationen auf ganzen Zahlen – so vermeiden wir Missverständnisse zwischen ganzen Zahlen und ihren Restklassen.

Richtig wäre also z.B. “ $7 \cdot 5 \equiv 3 \pmod{8}$ ”, aber *nicht* “ $7 \cdot 5 = 3 \equiv 3 \pmod{8}$ ”.

- Wir dürfen die Äquivalenzbeziehung auch mit ganzen Zahlen skalieren, ähnlich wie mit Gleichungen. *Allerdings gilt die Umkehrung nicht notwendigerweise (s.u.)* – daher ist hier nur eine Implikation richtig:

$$7 \equiv 2 \pmod{5} \quad \Rightarrow \quad 3 \cdot 7 \equiv 3 \cdot 2 \pmod{5}$$

- Achtung beim Skalieren allerdings bei Faktoren aus der Restklasse [0]: Für den Modul  $m = 5$  würde z.B. die Gleichung

$$x \equiv 2 \pmod{5}$$

als Lösungsmenge die gesamte Restklasse [2] besitzen. Skaliert man nun aber z.B. mit  $15 \in [0]$ , ergibt sich:

$$(15x \equiv 30 \pmod{5}) \quad \Leftrightarrow \quad (0 \equiv 0 \pmod{5})$$

Hier ist die Information über  $x$  verloren gegangen, da

$$[15 \cdot x] = [15] \odot [x] = [0] \odot [x] = [0 \cdot x] = [0]$$

Diese Gleichung ist für alle Restklassen modulo 5 lösbar, nicht nur für [2].

Wir formulieren die Bemerkung zum Skalieren von modulo-Äquivalenzbeziehungen noch als

**Satz 3.7** (Erweitern von Äquivalenzbeziehungen modulo  $m$ ). *Für  $m \in \mathbb{N}$  und  $x, y \in \mathbb{Z}$  gilt für  $j \in \mathbb{Z}$ :*

$$(x \equiv y \pmod{m}) \quad \Rightarrow \quad (j \cdot x \equiv j \cdot y \pmod{m})$$

(Beweis: S. 318.)

**Bemerkung:** Achtung – die Umkehrung gilt nicht immer! Siehe dazu den nächsten Unterabschnitt.

**Beispiel:** Wir betrachten die Beziehung

$$15 \equiv 27 \pmod{12}$$

Die Erweiterung mit 3 führt auf:

$$45 \equiv 81 \pmod{12}$$

Das stimmt. In der ursprünglichen Äquivalenz stammten beide Zahlen aus der Restklasse [3]; nach der Erweiterung stammen beide aus der Klasse [9].

### 3.2.3 Zur Division modulo $m$

Die Subtraktion (als Gegenstück zur Addition) ist für ganze Zahlen problemlos möglich (wir erinnern uns, dass diese gerade hierfür eingeführt wurden, da dies in den natürlichen Zahlen so nicht der Fall ist) – das gilt auch für Restklassen.

Produkte ganzer Zahlen sind stets ganze Zahlen. Die Division als Gegenstück zur Multiplikation ist jedoch für ganze Zahlen ähnlich problematisch wie die Subtraktion in  $\mathbb{N}$  – dies war die Motivation, die ganzen Zahlen auf die rationalen Zahlen zu erweitern. In  $\mathbb{Q}$  hat ein  $a \neq 0$  stets ein multiplikatives Gegenstück (*Inverses*; mehr dazu im Algebra-Kapitel 5), nämlich die Zahl  $a^{-1} = \frac{1}{a}$ , die dem Kehrwert des Bruchs  $a$  entspricht.

Das geht für ganze Zahlen also nicht. Möglich ist natürlich die Division dann, wenn sie durch den Teiler einer Zahl erfolgt – das war auch für die natürlichen Zahlen der Fall.

Beim Rechnen modulo  $m$  ist das Kürzen allerdings nicht immer möglich – auch dann nicht, wenn es von den Zahlenwerten her zunächst so scheint.

**Beispiel:** Wir betrachten nochmals die Beziehung

$$15 \equiv 27 \pmod{12}$$

Diese lässt sich, obwohl beide Zahlen Vielfache von 3 sind, *nicht* mit 3 kürzen – denn

$$5 \not\equiv 9 \pmod{12}$$

**Beispiel:** Kürzen kann auch möglich sein. Wir betrachten die Beziehung

$$10 \equiv 65 \pmod{11}$$

Diese lässt sich mit 5 kürzen zu

$$2 \equiv 13 \pmod{11}$$

Die Frage, ob sich eine modulo-Äquivalenz kürzen lässt, hängt also von den beteiligten Zahlen ab – und zwar über folgenden

**Satz 3.8** (Kürzungsregel modulo  $m$ ). Für  $m \in \mathbb{N}$  und  $x, y \in \mathbb{Z}$  gilt für  $j \in \mathbb{Z}$  und  $\text{ggT}(j, m) = 1$ :

$$(j \cdot x \equiv j \cdot y \pmod{m}) \quad \Rightarrow \quad (x \equiv y \pmod{m})$$

(Beweis: S. 318.)

### 3.2.4 Potenzieren modulo $m$

Das Potenzieren mit natürlichen Zahlen lässt sich direkt auf das mehrfache Multiplizieren zurück führen. Aus Definition 3.6 folgt dann analog für  $[x] \in \mathbb{Z}_m, a \in \mathbb{N}$ :

$$[x]^a = \underbrace{[x] \odot [x] \odot \cdots \odot [x]}_{a \text{ mal}} = [x^a]$$

Offenbar kann man also, wenn man die Restklasse eines Potenzausdrucks sucht, den Exponenten aus der Restklasse heraus ziehen. Wir behandeln gleich mehrere Beispiele mit vermischten Umformungen per modulo-Äquivalenz und per Gleichheit.

Bei den Rechnungen sind stets die Potenzgesetze zu beachten. Z.B. ist

$$[-1]^{512} = [(-1)^{512}] = [(-1)^{2 \cdot 256}] = [((-1)^2)^{256}] = [1^{256}] = [1]$$

**Beispiele:**

- Berechnung von  $4^{37} \pmod{5}$ : Mit einem Langarithmetikrechner erhält man:

$$4^{37} = 18889465931478580854784 \equiv 4 \pmod{5}$$

Mit den Rechenregeln erhält man die Restklasse aber auch leichter:

$$4^{37} = 4 \cdot 4^{2 \cdot 18} = 4 \cdot (4^2)^{18} = 4 \cdot 16^{18} = 4 \cdot (3 \cdot 5 + 1)^{18} \equiv 4 \cdot 1^{18} \equiv 4 \pmod{5}$$

- Berechnung von  $3^{12} \pmod{5}$ : Der Rechner liefert:

$$3^{12} = 531441 \equiv 1 \pmod{5}$$

Mit den Rechenregeln ergibt sich:

$$3^{12} = 3^{3 \cdot 4} = (3^3)^4 = 27^4 = (5 \cdot 5 + 2)^4 \equiv 2^4 = 16 \equiv 1 \pmod{5}$$

Oder alternativ:

$$3^{12} = 3^{3 \cdot 4} = (3^4)^3 = 81^3 = (16 \cdot 5 + 1)^3 \equiv 1^3 \equiv 1 \pmod{5}$$

- Berechnung von  $47^{11} \pmod{3}$ :

$$47^{11} = (48 - 1)^{11} = (16 \cdot 3 - 1)^{11} \equiv (-1)^{11} = (-1) \equiv 2 \pmod{3}$$

Im folgenden führen wir Aufschlüsselungen wie in den ersten beiden Gleichheiten dieser Rechnung meist stillschweigend durch.

- Berechnung von  $11^{47} \pmod{8}$ :

$$11^{47} \equiv 3^{47} = 3^{2 \cdot 23 + 1} = 9^{23} \cdot 3 \equiv 1^{23} \cdot 3 = 3 \equiv 3 \pmod{8}$$

- Berechnung von  $42^{39} \bmod 38$ . Mit etwas länglicher Rechnung erhalten wir:

$$\begin{aligned}
 42^{39} &\equiv 4^{39} = 2^{2 \cdot 39} = 2^{78} = 2^{5 \cdot 15 + 3} = 32^{15} \cdot 8 \\
 &\equiv (-6)^{15} \cdot 8 = -6^{15} \cdot 8 = -6^{2 \cdot 7 + 1} \cdot 8 = -36^7 \cdot 6 \cdot 8 \\
 &\equiv -(-2)^7 \cdot 6 \cdot 8 = 2^{10} \cdot 6 = 2^{5 \cdot 2} \cdot 6 = 32^2 \cdot 6 \\
 &\equiv (-6)^2 \cdot 6 = 36 \cdot 6 \equiv (-2) \cdot 6 = -12 \equiv 26 \pmod{38}
 \end{aligned}$$

- Berechnung von  $4711^{123} \bmod 17$ . Hier ist es sinnvoll, zunächst den Rest von 4711 modulo 17 zu ermitteln. Also dividieren wir schriftlich:

$$\begin{array}{r}
 4711 \quad = \quad 17 \cdot 277 + 2 \\
 \underline{-34} \\
 1311 \\
 \underline{-119} \\
 121 \\
 \underline{-119} \\
 2
 \end{array}$$

Nun können wir die eigentliche Rechnung beginnen:

$$4711^{123} \equiv 2^{123} = 2^{4 \cdot 30 + 3} = 16^{30} \cdot 8 \equiv (-1)^{30} \cdot 8 = 8 \equiv 8 \pmod{17}$$

### 3.3 Multiplikation und Division mit Restklassen

Wir befassen uns in diesem Abschnitt genauer mit der Multiplikation von Restklassen; hierbei verwenden wir überwiegend eindeutige Repräsentanten (d.h. die Zahlen zwischen 0 und  $(m-1)$  für den Modul  $m$ ). Außerdem werden wir Restklassen und ihre Vertreter synonym gebrauchen – das ist dann vertretbar, wenn aus dem Kontext klar hervor geht, wie die Notation zu verstehen ist. Analog verfahren wir mit den Operationen “ $\odot$ ” und “ $\cdot$ ”.

#### 3.3.1 Multiplikationstabellen modulo $m$ und Nullteiler

Zunächst betrachten wir *Multiplikationstabellen* für die Moduln  $m = 5$  und  $m = 10$ . Dabei lassen wir die Zeile/Spalte für die Faktoren Null aus, da diese trivial nur Nullen enthalten. Zu lesen sind die Einträge jeweils: “Zeilenindex  $j$  multipliziert mit Spaltenindex  $k$  ergibt Eintrag in Zeile  $j$ , Spalte  $k$ ”<sup>1</sup>

**Beispiele:**

- Für  $m = 5$ :

$\cdot$	1	2	3	4
1	1	2	3	4
2	2	4	1	3
3	3	1	4	2
4	4	3	2	1

- Für  $m = 10$ :

$\cdot$	1	2	3	4	5	6	7	8	9
1	1	2	3	4	5	6	7	8	9
2	2	4	6	8	0	2	4	6	8
3	3	6	9	2	5	8	1	4	7
4	4	8	2	6	0	4	8	2	6
5	5	0	5	0	5	0	5	0	5
6	6	2	8	4	0	6	2	8	4
7	7	4	1	8	5	2	9	6	3
8	8	6	4	2	0	8	6	4	2
9	9	8	7	6	5	4	3	2	1

<sup>1</sup>Bei Operationen, die nicht kommutativ sind – bei denen es also auf die Reihenfolge der Operanden ankommt – gibt es hier ein Risiko für Flüchtigkeitsfehler.

### Bemerkungen:

- Wir beobachten, dass modulo 5 die Tabelle in jeder Zeile/Spalte die Zahlen 1 bis 4 je genau einmal enthält. Die Gleichung  $j \cdot x = k$  für feste  $j, k \in \mathbb{Z}_5$  ist also eindeutig lösbar. Beispiel:  $3 \cdot x = 4$  hat die Lösung<sup>2</sup>  $x = 3$ . All dies ist natürlich jeweils modulo 5 zu verstehen.
- Beim Modul  $m = 10$  findet man solche Zeilen/Spalten für die Faktoren 1, 3, 7 und 9. Alle anderen Zeilen/Spalten enthalten mindestens einmal die Null, und außerdem andere Zahlenwerte doppelt. Die Zahlen 1 bis 9 können dort also nicht genau je einmal auftreten. Beispiel: Die Gleichung  $8 \cdot x = 4$  hat nicht eine, sondern zwei Lösungen – die Elemente 3 und 8. Die Gleichung  $8 \cdot x = 5$  hat hingegen keine einzige Lösung.  
Dieses Verhalten passt nicht zum Lösen linearer Gleichungen, wie wir es von  $\mathbb{Q}$  oder  $\mathbb{R}$  kennen.
- Wir lesen z.B. ab, dass  $2 \cdot 5 \equiv 0 \pmod{10}$ . Auch dies ist vom Rechnen mit rationalen oder reellen Zahlen (oder auch auf  $\mathbb{N}_0$  oder  $\mathbb{Z}$ ) so nicht möglich – dort erhält man die Null als Produkt nur, wenn mindestens einer der beiden Faktoren null ist.

---

Das führt uns zur folgenden

**Definition 3.9** (Nullteiler modulo  $m$ ). *Sei  $m$  gegeben und fest. Die Zahl  $j \in \mathbb{Z}_m \setminus \{0\}$ , bzw. ihr eindeutiger Repräsentant, heißt Nullteiler modulo  $m$ , falls es ein  $k \in \mathbb{Z}_m \setminus \{0\}$  gibt, sodass*

$$j \cdot k \equiv 0 \pmod{m}$$

**Beispiel:** Die Zahlen 2 und 5 sind Nullteiler modulo 10. Das gleiche gilt für die Zahlen 4, 6 und 8.

---

Für  $n = 10$  erstellen wir noch eine Multiplikationstafel für die Zahlen aus  $\mathbb{Z}_{10}$ , die keine Nullteiler sind:

$\cdot$	1	3	7	9
1	1	3	7	9
3	3	9	1	7
7	7	1	9	3
9	9	7	3	1

### Bemerkungen:

- (Die Ähnlichkeit mit der obigen Tabelle für  $\mathbb{Z}_5$  ist kein Zufall; die Ursache dafür (Stichwort *Galoiskörper*) führt für diese Vorlesung aber zu weit)
- Wir halten aber fest, dass hier wieder in jeder Zeile/Spalte jede der betrachteten Zahlen genau einmal vorkommt.

### 3.3.2 Multiplikatives Inverses modulo $m$

**Definition 3.10** (Multiplikatives Inverses). *Für eine Menge  $M$ , die eine Eins enthält und für deren Objekte eine kommutative (d.h. von der Reihenfolge der Operanden unabhängige) Multiplikation „ $\cdot$ “ definiert ist, heißt das Element  $b \in M$  Multiplikatives Inverses von  $a \in M$ , für das*

$$a \cdot b = b \cdot a = 1$$

Man schreibt dann das Inverse  $b$  auch als  $a^{-1}$

**Bemerkung:** (Bei nicht-kommutativer Multiplikation kann es vorkommen, dass es keine Inversen gibt, aber immerhin linksseitige (oder rechtsseitige) Inverse, für die nur jeweils eine der beiden Gleichungen in der obigen Definition gilt – dies führt aber über unsere Vorlesung hinaus. Wenn wir hier von Inversen sprechen, sind stets die beidseitigen gemeint.)

---

<sup>2</sup>Spezifischer:  $[3] \odot [x] = [4]$  hat die Lösungen  $x \in [3]$ ; oder:  $[3] \odot x = [4]$  hat die Lösung  $x = [3]$ .

**Beispiele:** Wir geben als Vorbereitung die multiplikativen Inversen für alle Zahlen aus  $\mathbb{Z}_5$  und  $\mathbb{Z}_{10}$  an, für die jeweils eines existiert:

- Für den Modul  $m = 5$ : Die Zahlen 1, 2, 3 und 4 besitzen alle je ein multiplikatives Inverses; dies sind die Zahlen 1, 3, 2 und 4 (in dieser Reihenfolge). Man kann diese in der Multiplikationstafel nachschlagen, indem man zur Zeile  $j$  die entsprechende Spalte  $k$  mit dem Eintrag 1 sucht.

Ausgeschrieben:

$$1^{-1} = 1, \quad 2^{-1} = 3, \quad 3^{-1} = 2, \quad 4^{-1} = 4$$

- Für den Modul  $m = 10$  beobachten wir, dass die einzigen Zahlen, die multiplikative Inverse besitzen, genau die Einträge aus reduzierten Multiplikationstafel ohne Nullteiler sind:

$$1^{-1} = 1, \quad 3^{-1} = 7, \quad 7^{-1} = 3, \quad 9^{-1} = 9$$


---

Mit den multiplikativen Inversen können wir nun auch formal die Division von Restklassen erklären:

**Definition 3.11** (Division von Restklassen). *Sei  $m \in \mathbb{N}$  gegeben und fest. Für Restklassen  $[a], [b] \in \mathbb{Z}_m$ , wobei  $[b]$  ein multiplikatives Inverses besitzt, ist die Division von  $[a]$  durch  $[b]$  definiert per*

$$\frac{[a]}{[b]} := [a] \odot [b]^{-1}$$

**Beispiele:**

- Für den Modul  $m = 5$  ist (mit eindeutigen Vertretern der Restklassen gerechnet!):

$$\frac{2}{4} = 2 \cdot 4^{-1} = 2 \cdot 4 = 3$$

Und:

$$\frac{1}{2} = 1 \cdot 2^{-1} = 2^{-1} = 3$$

Hier hat sich der “Bruch” also kürzen lassen, mit gleich bleibendem Zahlenwert 3.

Die “Hälfte” des obigen Bruchs wäre:

$$\frac{1}{2} \cdot \frac{2}{4} = 3 \cdot 3 = 4 = 4^{-1} = 1 \cdot 4^{-1} = \frac{1}{4}$$

Hier funktioniert also die Bruchrechnung in gewohnter Weise.

- Für  $m = 10$  ist (auch hier auf  $\mathbb{Z}_{10}$  gerechnet):

$$\frac{2}{9} = 2 \cdot 9^{-1} = 2 \cdot 9 = 8$$

Die Hälfte dieses Zahlenwerts wäre 4, allerdings ist

$$\frac{1}{9} = 9^{-1} = 9$$

Und auch die Multiplikation mit einem Bruch  $\frac{1}{2}$  können wir hier nicht anschreiben, da die Zahl 2 in  $\mathbb{Z}_{10}$  kein multiplikatives Inverses hat.

---

Das Konzept der Division (und damit: Bruchrechnung) mit Restklassen ist gleichbedeutend mit der lückenlosen Existenz von multiplikativen Inversen. Wir halten eine wichtige Tatsache fest, die uns die obigen Beobachtungen zusammen fasst:

**Satz 3.12** (Nullteiler und multiplikativ Inverse in  $\mathbb{Z}_m$ ). *Sei  $m \in \mathbb{N}$  mit  $m > 1$  gegeben und fest, und  $a \in \mathbb{Z}_m \setminus \{0\}$ . Dann ist  $a$  stets entweder ein Nullteiler modulo  $m$ , oder  $a$  besitzt ein multiplikatives Inverses modulo  $m$ :*

$$a \text{ ist Nullteiler von } \mathbb{Z}_m \quad \Leftrightarrow \quad \neg \left( \exists b \in \mathbb{Z}_m \setminus \{0\} : a \cdot b = 1 \right)$$



(Beweis: S. 318.)

---

Falls eine Zahl aus  $\mathbb{Z}_m$  also kein Nullteiler ist, so besitzt sie ein multiplikatives Inverses. Wir finden aber auch noch eine weitere wichtige Eigenschaft, die sich auf die Multiplikationstafel modulo  $m$  bezieht:

**Satz 3.13** (Invertierbare Elemente von  $\mathbb{Z}_m$ ). *Sei  $m \in \mathbb{N}$  mit  $m > 1$  gegeben und fest. Dann enthält die Multiplikationstafel modulo  $m$  für ein invertierbares  $a \in \mathbb{Z}_m \setminus \{0\}$  in ihrer Zeile/Spalte  $a$  jeden Zahlenwert aus  $\mathbb{Z}_m$  genau einmal.*

(Beweis: S. 319.)

---

Nun fehlt uns noch eine Eigenschaft, mit der wir direkt berechnen können, ob  $a \in \mathbb{Z}_m$  Nullteiler ist oder nicht. Tatsächlich ist dies möglich, und zwar mit folgendem

**Satz 3.14** (Multiplikative Invertierbarkeit modulo  $m$ ). *Sei  $m \in \mathbb{N}$  mit  $m > 1$  gegeben und fest. Eine Zahl  $a$  aus  $\mathbb{Z}_m \setminus \{0\}$  besitzt genau dann ein multiplikativ Inverses  $a^{-1}$ , wenn sie teilerfremd zu  $m$  ist.*

(Beweis: S. 319.)

**Bemerkung:** Dies ist genau verträglich mit der Kürzungsregel für modulo-Äquivalenzen aus Satz 3.8. Alternativ: Falls  $j$  ein multiplikatives Inverses modulo  $m$  besitzt, lässt sich die Äquivalenzbeziehung

$$j \cdot x \equiv j \cdot y \pmod{m}$$

mit dem Faktor  $j^{-1}$  erweitern (nach Satz 3.7). Es folgt (mit den Assoziativ- und Kommutativgesetzen der Multiplikation):

$$\underbrace{j^{-1} \cdot j}_1 \cdot x \equiv \underbrace{j^{-1} \cdot j}_1 \cdot y \pmod{m}$$

Das ist genau die Aussage der Kürzungsregel.

**Beispiel:** Für den Modul  $m = 12$  betrachten wir die Multiplikationstafel:

$\cdot$	1	2	3	4	5	6	7	8	9	10	11
1	1	2	3	4	5	6	7	8	9	10	11
2	2	4	6	8	10	0	2	4	6	8	10
3	3	6	9	0	3	6	9	0	3	6	9
4	4	8	0	4	8	0	4	8	0	4	8
5	5	10	3	8	1	6	11	4	9	2	7
6	6	0	6	0	6	0	6	0	6	0	6
7	7	2	9	4	11	6	1	8	3	10	5
8	8	4	0	8	4	0	8	4	0	8	4
9	9	6	3	0	9	6	3	0	9	6	3
10	10	8	6	4	2	0	10	8	6	4	2
11	11	10	9	8	7	6	5	4	3	2	1

Wir erkennen wieder die Nullteiler; das sind 2, 3, 4, 6, 8, 9 und 10. Die Zeilen/Spalten 1, 5, 7 und 11 enthalten hingegen genau jede positive Zahl aus  $\mathbb{Z}_{12}$  genau je einmal. Letzteres sind gerade die Zahlen, die teilerfremd zu 12 sind.

---

Wir fassen also zusammen:

Eine positive Zahl  $a$  aus  $\mathbb{Z}_m$  ist genau dann invertierbar modulo  $m$ , wenn sie teilerfremd zu  $m$  ist; den Wert  $a^{-1}$  bestimmt man (z.B.) mit dem erweiterten Euklidischen Algorithmus. Ist dagegen  $\text{ggT}(a, m) > 1$ , so ist  $a$  ein Nullteiler modulo  $m$ .

### 3.3.3 Prime Restklassensysteme

Nach den Erkenntnissen von oben führen wir noch einen neuen Begriff ein:

**Definition 3.15** (Primes Restklassensystem modulo  $m$ ). *Sei  $m \in \mathbb{N}$  mit  $m > 1$  gegeben und fest. Dann heißt die Menge*

$$\mathbb{Z}_m^* := \{j \in \mathbb{Z}_m \mid \text{ggT}(j, m) = 1\}$$

primes Restklassensystem modulo  $m$ .

#### Bemerkungen:

- Die Definition lässt sich sowohl für Restklassen als auch für deren eindeutigen Repräsentanten lesen. Genau genommen ist  $\mathbb{Z}_m = \{[0], \dots, [m-1]\}$ . Wählen wir nun aber  $j$  als den eindeutigen Repräsentanten einer dieser Restklassen aus  $\mathbb{Z}_m$ , so haben sämtliche Elemente von  $[j]$  genau den Divisionsrest  $j$  modulo  $m$ . Und dann gilt für jedes  $k \in [j]$  nach Satz 2.19:

$$\text{ggT}(k, m) = \text{ggT}((k \bmod m), m) = \text{ggT}(j, m)$$

Daher ist es möglich, den ggT einer gesamten Restklasse mit  $m$  zu definieren, und die Formel in obiger Definition ist sowohl für eindeutige Repräsentanten als auch für Restklassen sinnvoll.

- (Die Notation mit hoch gestelltem Stern rührt daher, dass die Menge  $\mathbb{Z}_m^*$  genau die Elemente aus  $\mathbb{Z}_m$  enthält, die multiplikative Inverse besitzen. Im Algebra-Kapitel 5 werden wir den Begriff der *Gruppe* kennen lernen, und  $\mathbb{Z}_m^*$  ist genau die multiplikative Gruppe von  $\mathbb{Z}_m$ . Dazu später mehr – jedenfalls soll der Stern eine Multiplikation andeuten.)

#### Beispiele:

- $\mathbb{Z}_2 = \{0, 1\}; \quad \mathbb{Z}_2^* = \{1\}$
- $\mathbb{Z}_3 = \{0, 1, 2\}; \quad \mathbb{Z}_3^* = \{1, 2\}$
- $\mathbb{Z}_5 = \{0, 1, 2, 3, 4\}; \quad \mathbb{Z}_5^* = \{1, 2, 3, 4\}$
- $\mathbb{Z}_6 = \{0, 1, \dots, 5\}; \quad \mathbb{Z}_6^* = \{1, 5\}$
- $\mathbb{Z}_9 = \{0, 1, \dots, 8\}; \quad \mathbb{Z}_9^* = \{1, 2, 4, 5, 7, 8\}$
- $\mathbb{Z}_{10} = \{0, 1, \dots, 9\}; \quad \mathbb{Z}_{10}^* = \{1, 3, 7, 9\}$
- $\mathbb{Z}_{12} = \{0, 1, \dots, 11\}; \quad \mathbb{Z}_{12}^* = \{1, 5, 7, 11\}$

Damit können wir eine Folgerung aus Satz 2.12 (Eigenschaften des ggT zweier Zahlen) ableiten, nämlich folgenden

**Satz 3.16** (Restklassensysteme modulo Primzahlen). *Für  $p \in \mathbb{N}$  mit  $p > 1$  ist  $\mathbb{Z}_p$  nullteilerfrei genau dann, wenn  $p$  prim ist.*

(Beweis: S. 319.)

#### Bemerkungen:

- Für  $p$  prim ist damit auch:

$$\mathbb{Z}_p^* = \mathbb{Z}_p \setminus \{0\}$$

- Und genau für  $p$  prim ist dann für  $\mathbb{Z}_p$  eine sinnvolle Division erklärt, da alle Elemente bis auf 0 ein multiplikatives Inverses haben.
- Ansonsten gilt für zusammen gesetzte Zahlen  $m$ :  $\mathbb{Z}_m^* \neq \mathbb{Z}_m \setminus \{0\}$ . Hier ist die Division nur teilweise erklärt, denn dividieren dürfen wir nur durch die Elemente aus  $\mathbb{Z}_m^*$  – und das sind hier *nicht* alle positiven Zahlen aus  $\mathbb{Z}_m$ .

### 3.3.4 Kleiner Satz von Fermat

Zum Abschluss dieses Kapitels geben wir noch eine Eigenschaft an, die manche Rechnungen stark vereinfachen kann:

**Satz 3.17** (Kleiner Satz von Fermat). <sup>3</sup> Für  $p$  prim und  $a \in \mathbb{Z}$  mit  $a \not\equiv 0 \pmod{p}$  gilt:

$$a^{p-1} \equiv 1 \pmod{p}$$

(Beweis: S. 320.)

#### Bemerkungen:

- Falls wir  $a$  aus dem Restklassensystem  $\mathbb{Z}_p$  als eindeutigen Vertreter wählen, gilt die Aussage für alle  $a \neq 0$ .
- Eine alternative Formulierung des Satzes findet sich durch Erweitern nach Satz 3.7:

$$a^p \equiv a \pmod{p}$$

Falls  $\text{ggT}(a, p) = 1$ , also wenn  $p \nmid a$ , lässt sich hierauf die Kürzungsregel aus Satz 3.8 anwenden, und es folgt wieder die obige Formulierung.

---

Anwendung des Satzes z.B. beim Bestimmen des Inversen, falls modulo einer Primzahl  $p$  gerechnet wird:

$$1 \equiv a^{p-1} \equiv a^{p-2} \cdot a \pmod{p}$$

Also ist das Inverse von  $a$  gerade äquivalent zu  $a^{p-2} \pmod{p}$ . Hier ist keine ggT-Berechnung mehr nötig.

**Beispiel:** Wir ermitteln das Inverse von 5 (modulo 7):

$$5^{-1} \equiv 5^{7-2} = 5^5 = 5 \cdot 5^4 = 5 \cdot (5^2)^2 = 5 \cdot (25)^2 \equiv 5 \cdot 4^2 = 5 \cdot 16 \equiv 5 \cdot 2 = 10 \equiv 3 \pmod{7}$$

---

Bei hohen Potenzen lässt sich Fermat auch manchmal verwenden, wenn der Modul  $p$  prim ist. Dann versucht man, Faktoren  $a^{p-1}$  zu bilden, die äquivalent zu 1 sind (oder erweitert  $a^p \equiv a$ ).

**Bemerkung:** Achtung: Für die Anwendung von Fermat ist es erforderlich, dass der Modul prim ist – das ist für eigene Rechnungen explizit zu erwähnen, wenn der Satz zitiert wird.

#### Beispiele:

- Beispiel mit Modul 11:

$$32^{2022} = 32^2 \cdot 32^{2020} = 32^2 \cdot (32^{10})^{202} \equiv 32^2 \cdot 1^{202} = 32^2 \equiv (-1)^2 = 1 \equiv 1 \pmod{11}$$

Tatsächlich funktioniert diese Rechnung wegen  $32 = 33 - 1$  sogar kürzer ohne Fermat per  $32^{2022} \equiv (-1)^{2022} = 1 \equiv 1 \pmod{11}$ , da wir eine gerade Potenz von  $(-1)$  sofort als 1 erhalten.

- Beispiel mit Modul 13:

$$18^{2425} = 18 \cdot 18^{2424} = 18 \cdot (18^{12})^{202} \equiv 18 \cdot 1^{202} = 18 \equiv 5 \pmod{13}$$

- Beispiel mit Modul 11:

$$5^{93} = 5^3 \cdot 5^{90} = 5^3 \cdot (5^{10})^9 \equiv 5^3 \cdot 1^9 = 5^3 = 5 \cdot 5^2 = 5 \cdot 25 \equiv 5 \cdot 3 = 15 \equiv 4 \pmod{11}$$

---

<sup>3</sup>P. de Fermat, frz. Mathematiker

### 3.3.5 Auffinden der multiplikativen Inversen

Wir haben verschiedene Strategien kennen gelernt, um das Inverse einer Zahl  $a$  modulo  $m$ , also die Zahl  $b \in \mathbb{Z}_m$  mit  $ab = 1$  zu bestimmen. Zusammen gefasst:

- Ausprobieren (für kleine Moduln  $m$  durchaus möglich)
- Etwas systematischer: Aufstellen der Multiplikationstafel. Findet man in der betreffenden Zeile  $a$  der Tafel den Eintrag 1, so ist  $a$  kein Nullteiler, und das Inverse  $b$  ist die entsprechende Spalte für den Eintrag 1 in Zeile  $a$ . Nach Satz 3.13 wird der Eintrag 1 in Zeile  $a$ , falls er auftritt, genau einmal auftreten – das Inverse ist damit eindeutig bestimmt.
- Falls der Modul  $m$  prim ist, hilft unter Umständen der kleine Satz von Fermat weiter (s.o.).
- Falls der Modul zu groß ist, um (ohne Taschenrechner oder Computer) zu beurteilen, ob  $m$  prim ist, fällt der kleine Fermat also als Methode aus. Dann ist oft auch das Anlegen der Multiplikationstafel unpraktisch.

Es bleibt aber noch das *Lemma von Bezout* (Satz 2.21): Denn danach gibt es Zahlen  $\alpha, \beta \in \mathbb{Z}$ , sodass

$$\text{ggT}(a, m) = \alpha \cdot a + \beta \cdot m$$

Wenn wir nun mit Euklid (Satz 2.20) den ggT berechnen und heraus finden, dass er größer ist als 1, so ist klar, dass es kein multiplikatives Inverses  $a^{-1}$  gibt. Falls aber ein ggT mit Wert 1 heraus kommt, können wir (falls wir für die Tabelle der Euklid-Schritte rechts noch Platz für die Spalten  $\alpha$  und  $\beta$  gelassen haben, und falls auch oben noch Platz für die beiden virtuellen Schritte  $(-1)$  und  $0$  vor der ersten Division mit Rest ist) das Inverse mit dem *erweiterten euklidischen Algorithmus* finden. Denn dann gilt ja nach Bezout, dass es  $\alpha, \beta$  gibt, sodass

$$1 = \alpha \cdot a + \beta \cdot m \equiv \alpha \cdot a \pmod{m}$$

Also haben wir mit dem korrekten Wert  $\alpha$  aus der Linearkombination des  $\text{ggT}(a, m)$  auch das gesuchte Inverse gefunden, da für  $\text{ggT}(a, m) = 1$  dann gilt:  $\alpha \equiv a^{-1} \pmod{m}$

In der Regel muss  $\alpha$  dann noch modulo  $m$  gerechnet werden, um in den Bereich der eindeutigen Divisionsreste  $\{0, \dots, m-1\}$  zu gelangen.

**Beispiel:** Wir hatten in den Beispielen zum Lemma von Bezout (Satz 2.21) bereits mit dem erweiterten euklidischen Algorithmus berechnet, dass

$$1 = \text{ggT}(927, 104) = 23 \cdot 927 + (-205) \cdot 104$$

Nehmen wir diese Gleichung modulo 927, so verschwindet der linke Beitrag (da er ein Vielfaches von 927 ist), und wir erhalten für das Inverse von 104 modulo 927:

$$104^{-1} \equiv -205 \equiv 722 \pmod{927}$$

Nehmen wir die Gleichung statt dessen modulo 104, so verschwindet der rechte Beitrag (ein Vielfaches von 104), und wir erhalten für das Inverse von 927 modulo 104:

$$927^{-1} \equiv 95^{-1} \equiv 23 \pmod{104}$$

# Kapitel 4

## Funktionen

### 4.1 Allgemeine Eigenschaften

#### 4.1.1 Funktionsbegriff

Wir starten mit einer sehr umfangreichen Definition, die mehrere zusammen hängende Begriffe einführt:

**Definition 4.1** (Funktion). *Falls es eine Relation  $R$  zwischen Mengen  $D$  und  $B$  gibt, sodass jedes Element  $x \in D$  mit genau einem Element  $y \in B$  in Relation steht, sagt man,  $D$  werde eindeutig auf  $B$  abgebildet.*

*Dann ist die Funktion  $f_R$  das Objekt, das die beiden Mengen spezifiziert und die Abbildungsvorschrift ( $x$  wird abgebildet auf  $f_R(x)$ ) angibt. Geschrieben:*

$$f_R : D \rightarrow B; \quad x \mapsto f_R(x)$$

*Dabei ist  $f_R(x)$  das Element aus  $B$ , auf welches  $x \in D$  abgebildet wird, gelesen “ $f_R$  von  $x$ ”. Dieses Objekt heißt Funktionswert (oder: Bild) von  $x$ . Das Objekt  $x$  selbst ist ein Urbild des Objekts  $f_R(x)$ .*

*Die Menge  $D$  heißt Definitionsbereich der Funktion; die Menge  $B$  heißt Bildbereich der Funktion.*

*Die Relation  $R$  selbst, also die Menge aller Paare von Elementen aus  $D$  und den ihnen jeweils eindeutig zugeordneten Elementen aus  $B$ , nämlich*

$$R = \left\{ (x, f_R(x)) \mid x \in D \right\} \subseteq D \times B,$$

*heißt Graph der Funktion.*

*Die Menge  $W \subseteq B$  mit*

$$W := \{ f_R(x) \mid x \in D \}$$

*heißt Wertebereich der Funktion*

#### Bemerkungen:

- Der Pfeil  $\mapsto$  bedeutet “wird abgebildet auf”.
- Meist verzichtet man auf die Angabe der Relation  $R$ ; dann schreibt man  $f$  statt  $f_R$ . Statt dessen nennt man den Graph der Funktion dann z.B.  $G_f$ .
- Den Funktionsgraph kann man in einem Mengendiagramm mit Pfeilen visualisieren. Falls Bild- und Definitionsbereich identisch sind, sollte die Menge auch nur einmal skizziert werden.
- Falls die Funktion reelle Zahlen auf reelle Zahlen abbildet (oder Teilmengen davon), visualisiert man den Funktionsgraph oft auch in einem kartesischen Koordinatensystem; mehr dazu s.u.
- Es kann durchaus sein, dass mehrere verschiedene Elemente aus  $D$  auf das gleiche Element von  $B$  abgebildet werden. Unmöglich ist aber, dass eine Funktion einem  $x \in D$  mehr als ein Element aus  $B$  zuordnet.

- Die Funktionsdefinition wird oft (und etwas ungenau, da die Angabe der beiden Mengen fehlt) abgekürzt als

$$f(x) := \dots,$$

wobei auf der rechten Seite dann die konkrete Abbildungsvorschrift steht. Diese Notation ist aber ein gleichwertiger Ersatz für die Schreibweise “ $x \mapsto \dots$ ”

### Beispiele:

- Für  $D := \{1, 2, 3\}$  und  $B := \{0, 1, 2, 3, 4, 8\}$  betrachten wir die Funktion  $f : D \rightarrow B$  mit der Abbildungsvorschrift

$$x \mapsto x^2 - 1$$

Abbildung 4.1 zeigt die beiden Mengen mit eingezeichnetem Funktionsgraphen.

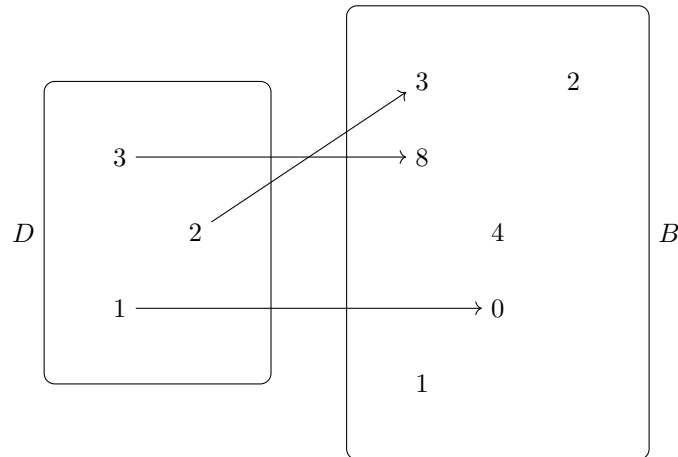


Abbildung 4.1: Funktion  $f : D \rightarrow B$  mit  $f(x) = x^2 - 1$

Wir lesen die konkreten Zuordnungen ab:

$$1 \mapsto 0; \quad 2 \mapsto 3; \quad 3 \mapsto 8$$

Die alternative Notation für die Abbildungsvorschrift für das obige Beispiel wäre:

$$f(x) := x^2 - 1$$

Dann ist:  $f(1) = 0$ ;  $f(2) = 3$ ;  $f(3) = 8$ .

Der Graph der Funktion ist dann:

$$G_f = \{(1, 0), (2, 3), (3, 8)\}$$

Wir zeigen den Funktionsgraph in Abbildung 4.2 alternativ noch in einem *kartesischen Koordinatensystem*. Die mit  $x$  bezeichnete Achse zeigt einen Teil des reellen Zahlenstrahls mit abgeteilten Beispielementen. Für die Elemente, die in  $D$  liegen, wird im rechten Winkel dazu auf der  $y$ -Achse der jeweilige Funktionswert abgetragen. Der eigentliche Graph besteht hier aus den drei markierten Punkten; die Hilfslinien dienen nur zur Orientierung.

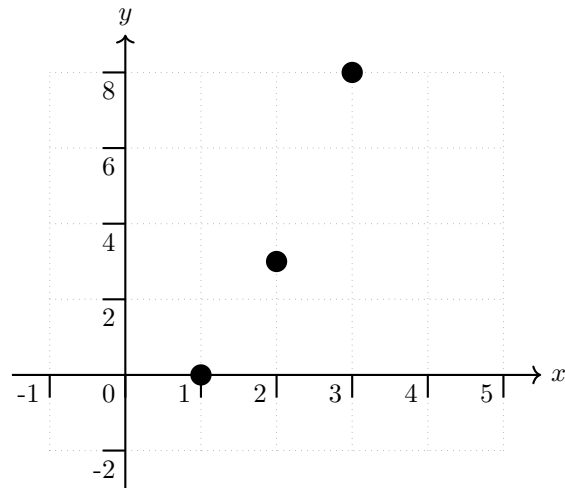


Abbildung 4.2: Graph der Funktion  $f : D \rightarrow B$  mit  $f(x) = x^2 - 1$

- Mit der selben Funktionsvorschrift, aber den Mengen  $D := [-1, 3]$  sowie  $B := \mathbb{R}$  erhält man den Funktionsgraphen aus Abbildung 4.3. Dieser ist nun nicht länger diskret, sondern enthält unendlich viele Punkte, die hier als durchgezogene Linie symbolisiert werden.

(Die drei Punkte aus dem vorigen Beispiel gehören natürlich auch hier zum Funktionsgraph dazu.)

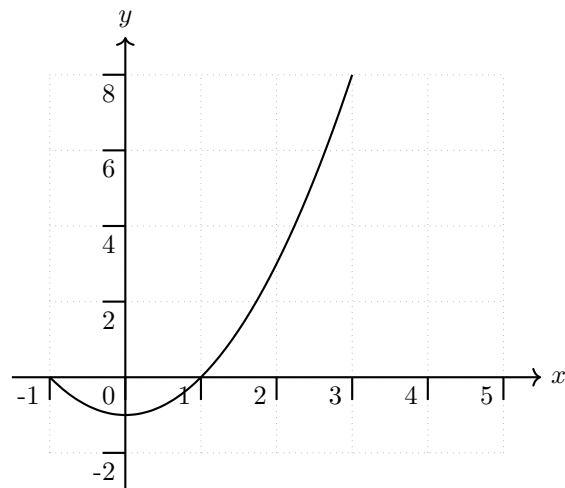


Abbildung 4.3: Graph der Funktion  $f : [-1, 3] \rightarrow \mathbb{R}$  mit  $f(x) = x^2 - 1$

- Jedes Prädikat (siehe Definition 1.8) ist eine Funktion mit dem Bildbereich  $\{\mathcal{W}, \mathcal{F}\}$ .
- Die Funktion  $f : \mathbb{R}^2 \rightarrow \mathbb{R}^3$  mit

$$(x_1, x_2) \mapsto (4x_1, 3x_1^4 - 2x_2, 42x_2)$$

ordnet jedem Paar von reellen Zahlen ein Tripel aus  $\mathbb{R}^3$  zu. Der Funktionsgraph ist eine Fläche in drei Dimensionen.

- Die Funktion  $f : \mathbb{R}^2 \rightarrow [0, \infty)$  mit

$$(x_1, x_2) \mapsto \sqrt{x_1^2 + x_2^2}$$

ordnet jedem Paar von reellen Zahlen eine nichtnegative reelle Zahl zu. Im kartesischen Koordinatensystem mit rechtwinkligen Achsen für  $x_1, x_2$  ist das der Abstand des Punktes  $(x_1, x_2)$  vom Koordinatenursprung (siehe Satz des Pythagoras).

### 4.1.2 Abbildungseigenschaften

Eine Funktion ordnet jedem Element  $x$  des Definitionsbereichs genau ein Element  $y$  aus dem Bildbereich zu. Aber wie sieht es mit der umgekehrten Richtung aus – lässt sich stets von  $y$  auf dessen Urbild  $x$  schließen? Wir können Abbildungen an diesem Verhalten klassifizieren:

**Definition 4.2** (Abbildungseigenschaften einer Funktion). *Eine Funktion  $f : D \rightarrow B$  mit Wertebereich  $W \subseteq B$  heißt*

- surjektiv, falls es zu jedem  $y \in B$  (mindestens) ein  $x \in D$  gibt, sodass  $y = f(x)$
- injektiv (oder: umkehrbar), falls es zu jedem  $y \in B$  höchstens ein Urbild  $x \in D$  gibt, d.h. falls

$$(f(x_1) = f(x_2)) \Rightarrow (x_1 = x_2)$$

- bijektiv (oder: umkehrbar eindeutig), falls  $f$  sowohl surjektiv als auch injektiv ist

#### Bemerkungen:

- Bijektive Funktionen besitzen also zu jedem  $y \in B$  genau ein Urbild (das ist die einzige Möglichkeit, um die Anforderungen “mindestens ein Urbild” und “höchstens ein Urbild” gleichzeitig zu erfüllen).
- Bei surjektiven Funktionen gilt für den Wertebereich:  $W = B$ ; es gibt also keine Elemente aus  $B$ , die nicht Funktionswert irgendeines  $x \in D$  sind.
- Indem man alle Elemente aus  $B \setminus W$  aus dem Bildbereich entfernt, lässt sich eine neue Funktion mit selber Abbildungsvorschrift konstruieren, die surjektiv ist.
- Bei injektiven Funktionen lässt sich im Prinzip eine Funktion konstruieren, die von den Bildern aus  $B$  zurück auf deren Urbilder aus  $D$  abbildet. Je nach Bildbereich (also abhängig davon, ob  $f$  surjektiv ist oder nicht) kann es aber noch sein, dass es in  $B$  Elemente gibt, für die kein Urbild existiert.
- Für injektive Funktionen kann man auch kontrapositorisch formulieren:

$$(x_1 \neq x_2) \Rightarrow (f(x_1) \neq f(x_2))$$

#### Beispiele:

- Die Funktion

$$f : \{1, 2, 3\} \rightarrow \{0, 1, 2, 3, 4, 8\}; \quad x \mapsto x^2 - 1$$

aus dem ersten obigen Beispiel ist nicht surjektiv, denn der Bildbereich enthält drei Elemente, die keine Urbilder besitzen (die Zahlen 1, 2 und 4).

Sie ist aber injektiv, da die Elemente des Bildbereichs jeweils höchstens ein Urbild besitzen.

Würden wir den Wertebereich  $\{0, 3, 8\}$  als neuen Bildbereich vereinbaren, wäre die neue Funktion auch surjektiv, und damit dann auch bijektiv.

- Die Funktion

$$f : [-1, 3] \rightarrow \mathbb{R}; \quad x \mapsto x^2 - 1$$

aus dem zweiten obigen Beispiel ist nicht surjektiv, da z.B. die Zahl  $42 \in \mathbb{R}$  kein Urbild in  $[-1, 3]$  besitzt.

Sie ist auch nicht injektiv, da z.B. die Zahl  $0 \in \mathbb{R}$  mehr als ein Urbild besitzt (nämlich die beiden Zahlen  $\pm 1 \in [-1, 3]$ ).

Würde man den Bildbereich auf  $[-1, 8]$  einschränken, so wäre die Funktion immerhin surjektiv – an der fehlenden Injektivität ändert das jedoch nichts.

Man könnte Injektivität herstellen, indem man den Definitionsbereich auf  $[0, 3]$  einschränkt. Diese neue Funktion hat allerdings dann einen anderen Funktionsgraphen!



- Die Funktion

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}^3; \quad (x_1, x_2) \mapsto (4x_1, 3x_1^4 - 2x_2, 42x_2)$$

ist nicht surjektiv. In diesem Fall sieht man dies allerdings vielleicht nicht sofort (ein gut sichtbares Gegenbeispiel reicht aus, um eine allgemeine Aussage zu widerlegen; siehe Satz 1.10) – hier müssen wir rechnen. Die Frage ist, ob ein beliebiges Tripel  $(a, b, c) \in \mathbb{R}^3$  ein Urbild in  $\mathbb{R}^2$  besitzt. Dazu betrachten wir die Gleichung

$$f(x_1, x_2) = (a, b, c)$$

und setzen die Funktionsvorschrift ein. Gleichheit von Tupeln besteht nach Definition 1.23 genau dann, wenn alle Komponenten gleich zueinander sind; hier bedeutet dies:

$$\begin{aligned} 4x_1 &= a \\ \wedge \quad 3x_1^4 - 2x_2 &= b \\ \wedge \quad 42x_2 &= c \end{aligned}$$

Falls wir nun  $x_1$  und  $x_2$  so in Abhängigkeit von  $a, b, c$  ausdrücken können, dass  $(a, b, c)$  beliebig gewählt werden kann, ist die Surjektivität von  $f$  gezeigt.

Wir lesen direkt ab:  $x_1 = \frac{a}{4}$ , sowie  $x_2 = \frac{c}{42}$ .

Diese Ausdrücke eingesetzt in die zweite Bedingung (die ja auch erfüllt sein muss), ergibt:

$$3 \cdot \frac{a^4}{256} - 2 \cdot \frac{c}{42} = b$$

Also ist  $b$  nicht frei wählbar, sondern über eben diese Bedingung von  $a$  und  $c$  abhängig. Als einfachstes Beispiel probieren wir  $a := 0$  und  $c := 0$ . Das ergibt nach obiger Bedingung  $b = 0$ . Wählen wir dann also ein  $b \neq 0$ , so wäre diese Bedingung verletzt – und dann lässt sich zu  $(a, b, c)$  kein Urbild finden.

Konkret: Das Tripel  $(0, 7, 0) \in \mathbb{R}^3$  besitzt unter  $f$  kein Urbild in  $\mathbb{R}^2$ . Und damit ist  $f$  nicht surjektiv.

Allerdings ist die Funktion injektiv. Auch dies ist nicht direkt offensichtlich, denn der Ausdruck  $3x_1^4$  ergibt für  $\pm x_1$  jeweils den gleichen Zahlenwert. Allerdings unterscheiden sich dann die jeweiligen ersten Komponenten der Tripel, da  $x_1$  dort mit ungerader Potenz eingeht. Um sicher fest zu stellen, ob tatsächlich Injektivität vorliegt, müssen wir also formal rechnen, ob gilt:

$$(f(x_1, x_2) = f(y_1, y_2)) \Rightarrow ((x_1, x_2) = (y_1, y_2))$$

Wir setzen für die linke Seite der Folgerung die Funktionsvorschrift ein und erhalten:

$$\begin{aligned} 4x_1 &= 4y_1 \\ \wedge \quad 3x_1^4 - 2x_2 &= 3y_1^4 - 2y_2 \\ \wedge \quad 42x_2 &= 42y_2 \end{aligned}$$

Wenn dies gilt, so sind insbesondere die erste und die dritte Gleichung erfüllt. Wir dividieren diese durch 4 bzw. durch 42, und erhalten direkt:  $x_1 = y_1 \wedge x_2 = y_2$ . Damit ist die Injektivität gezeigt.

Durch Anpassung des Bildbereichs auf die Menge

$$\left\{ (a, b, c) \in \mathbb{R}^3 \mid b = 3 \cdot \frac{a^4}{256} - 2 \cdot \frac{c}{42} \right\}$$

ließe sich aus  $f$  eine surjektive, und damit auch bijektive, Funktion mit gleicher Abbildungsvorschrift konstruieren.

- Ändern wir die Funktion des vorigen Beispiels ab zu

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}^3; \quad (x_1, x_2) \mapsto (4x_1^2, 3x_1^4 - 2x_2, 42x_2),$$

so ist  $f$  hier nicht länger injektiv (ließe sich also auch nicht durch Einschränken des Bildbereichs bijektiv machen). Denn hier wäre z.B.  $f(-1, 0) = f(1, 0) = (4, 3, 0)$ , d.h. es gibt in  $\mathbb{R}^3$  ein Element mit mehr als einem Urbild.

- Für  $a, b \in \mathbb{R}$  und  $a \neq 0$  ist die Funktion

$$f : \mathbb{R} \rightarrow \mathbb{R}; \quad x \mapsto a \cdot x + b$$

surjektiv, denn wir finden für jedes  $y \in \mathbb{R}$  ein Urbild:

$$y = ax + b \quad \Leftrightarrow \quad x = \frac{y - b}{a}$$

Sie ist außerdem injektiv, denn es gilt:

$$\begin{aligned} y_1 &= y_2 \\ \Leftrightarrow ax_1 + b &= ax_2 + b \\ \Leftrightarrow ax_1 &= ax_2 \\ \Leftrightarrow x_1 &= x_2 \end{aligned}$$

Hierbei durften wir die Gleichung mit  $a$  dividieren, da nach Voraussetzung  $a \neq 0$  gilt.

Damit ist die Funktion  $f$  (die allgemeine Darstellung einer *affin-linearen Funktion*) bijektiv.

- Dagegen ist die Funktion

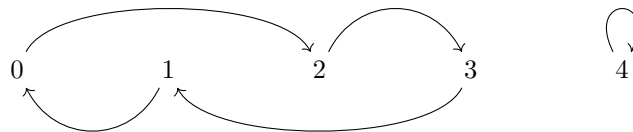
$$f : \mathbb{R} \rightarrow \mathbb{R}; \quad x \mapsto 7$$

weder surjektiv (z.B. hat  $42 \in \mathbb{R}$  kein Urbild) noch injektiv (da z.B.  $f(1) = f(2) = 7$ ). Eine Einschränkung des Bildbereichs auf  $\{7\}$  würde eine immerhin surjektive Funktion ergeben.

- Für  $M := \{0, 1, 2, 3, 4\}$  sei

$$f : M \rightarrow M; \quad x \mapsto (3x + 2) \bmod 5$$

Bei endlichen Mengen mit wenigen Elementen ist die grafische Darstellung oft hilfreich. Hier stimmen Definitions- und Bildbereich überein, d.h. wir skizzieren die Menge nur einmal:



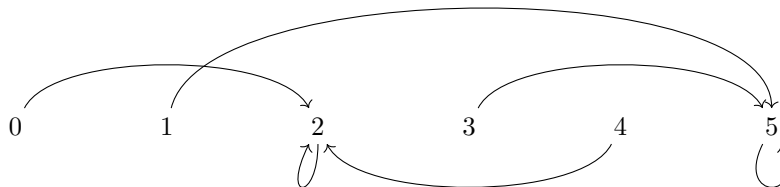
Für die Frage nach Injektivität und Surjektivität reicht in der Skizze die Untersuchung, wie viele Pfeilspitzen an den einzelnen Elementen von  $M$  anliegen. Jede Pfeilspitze gehört nämlich zu genau einem Pfeil, der von genau einem Urbild ausgehen muss.

$f$  ist injektiv, da jeder Funktionswert  $y = f(x)$  nur genau (also auch höchstens) ein Urbild besitzt.  $f$  ist außerdem surjektiv, da jedes  $y \in M$  genau (also auch mindestens) ein Urbild  $x$  mit  $f(x) = y$  besitzt. Also ist  $f$  sogar bijektiv.

- Für  $N := \{0, 1, 2, 3, 4, 5\}$  sei

$$g : N \rightarrow N; \quad x \mapsto (3x + 2) \bmod 6$$

Skizze:



$g$  ist nicht injektiv, da z.B. der Wert  $y = 2$  drei verschiedene Urbilder besitzt (nämlich 0, 2 und 4).  $g$  ist außerdem nicht surjektiv, da die Zahlen  $\{0, 1, 3, 4\} \subseteq N$  keine Urbilder besitzen. Man könnte  $g$  surjektiv machen, indem man den Bildbereich auf  $\{2, 5\}$  einschränkt, aber  $g$  lässt sich ohne Änderung der Abbildungsvorschrift nicht injektiv (umkehrbar) machen.

(Der Grund für dieses Verhalten ist, dass 3 ein Nullteiler modulo 6 ist. Wäre die Abbildungsvorschrift statt dessen  $f(x) := (5x + 2) \bmod 6$ , so läge wieder eine bijektive Funktion vor, denn  $5 \in \mathbb{Z}_6^* = \{1, 5\}$ . Nach Satz 3.13 erhält man für verschiedene  $x \in N$  auch verschiedene Restklassen  $[5x]$  modulo 6.)

Für bijektive Funktionen  $f$  lässt sich die Zuordnung zwischen  $D$  und  $B$  auch in umgekehrter Richtung definieren; es gibt also eine Abbildung, die jedem Element aus  $B$  eindeutig auf sein Urbild aus  $D$  abbildet. Falls man den Funktionsgraphen von  $f$  (also die Menge der Abbildungspfeile) vorliegen hat, reicht es zum Auffinden dieser umgekehrten Abbildung, bei sämtlichen Pfeilen die Richtung umzukehren.

**Definition 4.3** (Umkehrfunktion). *Sei  $f : D \rightarrow B$  eine bijektive Funktion. Dann heißt die Abbildung*

$$f^{-1} : B \rightarrow D$$

Umkehrfunktion (oder: Umkehrabbildung, Inverse) von  $f$ , falls für alle  $x \in D, y \in B$  gilt

$$y = f(x) \quad \Leftrightarrow \quad x = f^{-1}(y)$$

**Bemerkung:** Die Umkehrfunktion  $f^{-1}$  ist dann selbst bijektiv, und ihre eigene Umkehrfunktion ist wieder die ursprüngliche Funktion  $f$ :

$$(f^{-1})^{-1} = f$$

**Beispiele:**

- Für die Funktion

$$f : \{1, 2, 3\} \rightarrow \{0, 3, 8\}; \quad x \mapsto x^2 - 1$$

lautet die Umkehrabbildung:

$$f^{-1} : \{0, 3, 8\} \rightarrow \{1, 2, 3\}; \quad x \mapsto \sqrt{x+1}$$

Dass  $f$  in dieser Form bijektiv ist, hatten wir oben bereits ermittelt. Für die Funktionsvorschrift können wir schreiben:

$$\begin{aligned} y = f(x) &= x^2 - 1 \\ \Leftrightarrow y + 1 &= x^2 \\ \stackrel{x \geq 0}{\Leftrightarrow} \sqrt{y+1} &= x \end{aligned}$$

Die zweite Äquivalenzumformung ist nur deswegen richtig, weil  $x \geq 0$  für alle Elemente des Definitionsbereichs von  $f$  gilt. Allgemein hätten wir nämlich  $x = \pm\sqrt{\dots}$  schreiben müssen. Daher die Annotation am Äquivalenzpfeil.

Aber da  $x = f^{-1}(y)$ , ist damit

$$f^{-1}(y) = \sqrt{y+1}$$

Die Umbenennung der Veränderlichen in  $x$  lässt sich hierbei durch einfaches Pattern Matching durchführen: Die beiden Vorschriften

$$x \mapsto \sqrt{x+1} \quad \text{und} \quad y \mapsto \sqrt{y+1}$$

sind absolut gleichwertig und unterscheiden sich nur im Namen der Variablen.

- Für die affin-lineare Funktion

$$f : \mathbb{R} \rightarrow \mathbb{R}; \quad x \mapsto a \cdot x + b$$

mit  $a, b \in \mathbb{R}, a \neq 0$ , ist die Umkehrfunktion:

$$f^{-1} : \mathbb{R} \rightarrow \mathbb{R}; \quad x \mapsto \frac{x-b}{a}$$

Die nötigen Überlegungen hierzu hatten wir oben schon angestellt; hier wurde nur noch die Umbenennung der Veränderlichen für  $f^{-1}$  durchgeführt.

- Für  $M := \{0, 1, 2, 3, 4\}$  und

$$f : M \rightarrow M; \quad x \mapsto (3x + 2) \bmod 5$$

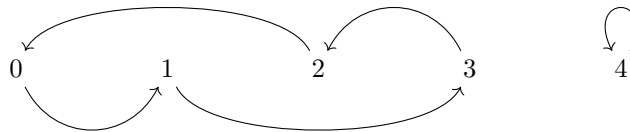
finden wir die Umkehrfunktion ( $f$  ist bijektiv, s.o.) durch Rechnen mit Restklassen. Dafür benötigen wir das multiplikative Inverse von 3, das wir oben schon als 2 bestimmt hatten. Also gilt:

$$\begin{aligned} y = f(x) &\equiv 3x + 2 \pmod{5} \\ \Leftrightarrow 2y &\equiv x + 4 \pmod{5} \\ \Leftrightarrow 2y + 1 &\equiv x \pmod{5} \end{aligned}$$

Damit ist also die Umkehrfunktion gegeben per

$$f^{-1} : M \rightarrow M; \quad x \mapsto (2x + 1) \bmod 5$$

Hier das Schaubild von  $f^{-1}$  mit den Abbildungspfeilen: Im Vergleich zum Graphen von  $f$



erkennt man, dass alle Abbildungspfeile die Richtung gewechselt haben.

### 4.1.3 Verkettung (Hintereinanderausführung) von Funktionen

Falls die Definitions- und Wertebereiche miteinander verträglich sind, lässt sich die Hintereinanderausführung von zwei (oder mehr) Funktionen insgesamt als eine Gesamtfunktion beschreiben:

**Definition 4.4** (Komposition von Funktionen). *Mit den Funktionen  $f : D_f \rightarrow B_f$  mit Wertebereich  $W_f \subseteq B_f$ , und  $g : D_g \rightarrow B_g$  ist für den Fall  $W_f \subseteq D_g$  die Komposition (oder: Hintereinanderausführung, Verkettung) von  $g$  mit  $f$  als eine Funktion  $g \circ f : D_f \rightarrow B_g$  (lies: “ $g$  Kringel  $f$ ”, “ $g$  angewendet auf  $f$ ”) definiert per*

$$(g \circ f)(x) := g(f(x))$$

**Bemerkungen:**

- Achtung: Oft ist  $g \circ f \neq f \circ g$ , d.h. die Hintereinanderausführung ist *nicht kommutativ*!
- Als Eselsbrücke: Die Funktion, die in  $(g \circ f)(x)$  “näher bei  $x$ ” notiert ist, wird zuerst angewandt.
- Falls  $f, g$  bijektiv sind, ist auch  $g \circ f$  bijektiv, und die Umkehrfunktion lautet:

$$(g \circ f)^{-1}(x) = (g(f(x)))^{-1} = f^{-1}(g^{-1}(x)) = (f^{-1} \circ g^{-1})(x)$$

Wie beim Ablegen von zuvor übereinander angezogenen Kleidungsstücken kehrt sich die Reihenfolge um. Zuerst ist die Anwendung von  $g$  rückgängig zu machen, danach die von  $f$ .

**Beispiele:**

- Wir betrachten  $f, g : \mathbb{R} \rightarrow \mathbb{R}$  mit  $f(x) := x^2$  und  $g(x) := 7x$ . Wir definieren  $h(x) := (g \circ f)(x) = g(f(x))$  und erhalten:

$$h(x) = g(f(x)) = g(x^2) = 7x^2$$

Oder alternativ:

$$h(x) = g(f(x)) = 7f(x) = 7x^2$$

Allerdings ist  $\tilde{h} := f \circ g$  nicht die gleiche Funktion:

$$\tilde{h}(x) = (f \circ g)(x) = f(g(x)) = f(7x) = (7x)^2 = 49x^2$$

Alternativ:

$$\tilde{h}(x) = f(g(x)) = (g(x))^2 = (7x)^2 = 49x^2$$

- Die Komposition zweier affin-linearer reeller Funktionen ist stets wieder solch eine affin-lineare Funktion (Beweis: Übung!). Wir betrachten als Beispiel:  $f, g: \mathbb{R} \rightarrow \mathbb{R}$  mit

$$f(x) := 3x + 4 \quad \text{und} \quad g(x) := -2x + 1$$

Die Umkehrfunktionen sind:

$$f^{-1}(x) = \frac{x-4}{3} \quad \text{und} \quad g^{-1}(x) = -\frac{x-1}{2}$$

Wir bestimmen die Komposition  $g \circ f$  und ihre Umkehrfunktion (denn affin-lineare Funktionen sind bijektiv):

$$(g \circ f)(x) = g(3x + 4) = -2 \cdot (3x + 4) + 1 = -6x - 8 + 1 = -6x - 7$$

Deren Umkehrfunktion ist (direkt bestimmt):

$$(g \circ f)^{-1}(x) = -\frac{x+7}{6}$$

Aber es ergibt sich auch:

$$(f^{-1} \circ g^{-1})(x) = f^{-1}\left(-\frac{x-1}{2}\right) = \frac{\left(-\frac{x-1}{2}\right) - 4}{3} = -\frac{\left(\frac{x-1}{2}\right) + 4}{3} = -\frac{(x-1) + 8}{6} = -\frac{x+7}{6}$$

Die Verknüpfung ist also nicht kommutativ, aber es wir haben eben schon exemplarisch gesehen, dass es keine Auswirkungen auf das Endergebnis hat, ob wir in  $g \circ f(x)$  zunächst den Ausdruck  $f(x)$  berechnen und in  $g$  einsetzen, oder ob wir zuerst  $g$  abhängig von  $f(x)$  (als Ganzes) aufschreiben und dann erst  $f(x)$  einsetzen – an der Hierarchie der Schachtelung ändert sich dadurch nichts.

Allgemeiner gilt:

**Satz 4.5** (Assoziativität der Komposition). *Für Funktionen  $f, g, h$  mit passenden Werte- und Definitionsbereichen, sodass  $h(g(f(x)))$  für alle  $x \in D_f$  wohldefiniert ist, gilt:*

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

**Beweis:** Wir zeigen die Gleichheit der ersten und zweiten Schreibweise:

$$\begin{aligned} [h \circ (g \circ f)](x) &= h((g \circ f)(x)) = h(g(f(x))) = h(g(f(x))) = (h \circ g)(f(x)) \\ &= [(h \circ g) \circ f](x) \quad \blacksquare \end{aligned}$$

**Beispiel:** Wir betrachten die reellen Funktionen  $f, g, h$  mit den Abbildungsvorschriften

$$f(x) := 3x + 4 \quad \text{und} \quad g(x) := \frac{2}{x} \quad \text{und} \quad h(x) := 5 - x^2$$

Wir berechnen  $h(g(f(x)))$  auf beide Arten:

Zunächst ist

$$(g \circ f)(x) = \frac{2}{3x+4}$$

Hierbei stellen wir noch fest, dass wir den Definitionsbereich von  $f$  (und damit auch von allen Kompositionen, die mit  $f$  beginnen) auf  $\mathbb{R} \setminus \{-\frac{4}{3}\}$  einschränken müssen, damit bei Anwendung von  $g$  nicht durch null dividiert wird.

Dann ist

$$(h \circ (g \circ f))(x) = 5 - ((g \circ f)(x))^2 = 5 - \frac{4}{(3x+4)^2}$$

Alternativ: Zunächst ist

$$(h \circ g)(x) = 5 - (g(x))^2 = 5 - \frac{4}{x^2}$$

Und dann ist

$$((h \circ g) \circ f)(x) = 5 - \frac{4}{(f(x))^2} = 5 - \frac{4}{(3x+4)^2}$$

#### 4.1.4 Addition, Multiplikation, Division

**Definition 4.6** (Rechenoperationen mit Funktionen). Für zwei Funktionen  $f, g$  mit geeigneten und verträglichen Definitions- und Bildbereichen werden elementweise erklärt:

- die Summenfunktion  $(f + g)$  per

$$(f + g)(x) := f(x) + g(x)$$

- die Produktfunktion  $(f \cdot g)$  per

$$(f \cdot g)(x) := f(x) \cdot g(x)$$

- die Quotientenfunktion  $\left(\frac{f}{g}\right)$  per

$$\left(\frac{f}{g}\right)(x) := \frac{f(x)}{g(x)}$$

überall dort, wo  $g(x) \neq 0$ .

**Beispiele:**

- Die affin-lineare Funktion  $f$  mit  $f(x) := 4x - 5$  lässt sich als Summenfunktion  $f = f_1 + f_2$  schreiben, mit

$$f_1(x) := 4x \quad \text{und} \quad f_2(x) := \text{const} = -5$$

- Die Funktion  $f_1$  aus vorigem Beispiel ist eine Produktfunktion  $f_1 = g_1 \cdot g_2$  mit

$$g_1(x) := \text{const} = 4 \quad \text{und} \quad g_2(x) := x$$

- Der Quotient aus den Funktionen  $p$  und  $q$  mit  $p(x) := 2x^2 - 8$  und  $q(x) := 6x + 12$  ist, für  $x \neq (-2)$ , wieder eine Funktion; diese lässt sich sogar noch vereinfachen (dritte binomische Formel siehe Satz 1.31):

$$\left(\frac{p}{q}\right)(x) = \frac{2x^2 - 8}{6x + 12} = \frac{2(x^2 - 4)}{6(x + 2)} = \frac{2(x + 2)(x - 2)}{6(x + 2)} = \frac{(x + 2)(x - 2)}{3(x + 2)}$$

Wenn wir nun die Funktion  $(x + 2)$  kürzen, entfällt sogar die *Definitionslücke*, und wir erhalten:

$$\left(\frac{p}{q}\right)(x) = \dots = \frac{x - 2}{3}$$

Diese Funktion ist auf ganz  $\mathbb{R}$  definiert.

Definitionslücken, die sich auf diese Art eliminieren lassen, heißen *hebbar* (weil man sie durch Umformen "aufheben" kann).

## 4.2 Basisfunktionen

Wir behandeln hier in aller Kürze einige grundlegende reelle Funktionen, die später in Mathematik 2 noch genauer vorgestellt werden. Zuvor noch eine

**Definition 4.7** (Parität reeller Funktionen).  $f$  sei eine reelle Funktion in einer Veränderlichen. Dann heißt  $f$

- gerade, falls für alle  $x \in \mathbb{R}$  gilt:  $f(-x) = f(x)$
- ungerade, falls für alle  $x \in \mathbb{R}$  gilt:  $f(-x) = -f(x)$

Ist  $f$  weder gerade noch ungerade, so hat  $f$  keine bestimmte Parität.

### 4.2.1 Potenzfunktionen

**Definition 4.8** (Potenzfunktionen). Für  $j \in \mathbb{N}$  ist die Potenzfunktion  $j$ -ten Grades gegeben durch die  $j$ -fache Multiplikation der identischen Funktion ( $x \mapsto x$ ) mit sich selbst:

$$x^j = \underbrace{x \cdot x \cdot \dots \cdot x}_{j\text{-mal}}$$

Zusätzlich vereinbaren<sup>1</sup> wir, auch für  $x = 0$ :

$$x^0 := 1$$

#### Bemerkungen:

- Definitionsbereich ist ganz  $\mathbb{R}$ .
- Falls  $j$  gerade ist, ist die Potenzfunktion *gerade*, sonst *ungerade*.
- Gerade Potenzfunktionen ( $j \neq 0$ ) haben den Wertebereich  $[0, \infty)$ , da sie sich stets als Quadrat einer Potenzfunktion schreiben lassen.  
Ungerade Potenzfunktionen bilden ab in den Wertebereich  $\mathbb{R}$ .
- Ungerade reelle Potenzfunktionen sind injektiv, gerade dagegen nicht.

### 4.2.2 Rationale Funktionen

**Definition 4.9** (Ganzrationale Funktionen). Summen aus reellen Potenzfunktionen, die jeweils noch mit reellen Faktoren skaliert werden können, bezeichnet man als ganzrationale Funktionen (oder: Polynomfunktionen). Der Exponent der Potenzfunktion mit höchstem Grad heißt Grad der Funktion.

Eine ganzrationale Funktion  $f$  vom Grad  $n \in \mathbb{N}$  lässt sich stets schreiben als

$$f(x) = \sum_{j=0}^n c_j \cdot x^j = c_n x^n + \dots + c_2 x^2 + c_1 x + c_0$$

Die Zahlen  $c_j$  (null ist erlaubt, falls die Potenz  $j$  nicht vorkommen soll; aber  $c_n \neq 0$ ) heißen Koeffizienten.

#### Bemerkungen:

- Sind alle Potenzfunktionen gerade, so ist auch die Polynomfunktion gerade; sind alle Potenzfunktionen ungerade, so ist auch die Polynomfunktion ungerade. Anderenfalls hat die Polynomfunktion keine definierte Parität.
- Ganzrationale Funktionen vom Grad 1 heißen *linear* oder (s.o., und genauer) *affin-linear*. Ihre Graphen sind im kartesischen Koordinatensystem *Geraden*.
- Ganzrationale Funktionen vom Grad 2 heißen *quadratisch*. Ihre Graphen sind im kartesischen Koordinatensystem *Parabeln*.
- (Ganzrationale Funktionen der Grade 3 und 4 werden auch als *kubische* bzw. *quartische* Funktionen bezeichnet.)

#### Beispiele:

- Eine quadratische Funktion (also vom Grad 2) findet sich in Abbildung 4.3; diese ist gerade.
- Eine weitere quadratische Funktion zeigt Abbildung 4.4; die Funktionsvorschrift ist

$$f(x) := x^2 - 4x + 3$$

Diese Funktion hat keine definierte Parität.

- Eine Funktion fünften Grades mit der Vorschrift

$$f(x) := \frac{x^5}{100} - \frac{x^4}{50} - \frac{2x^3}{25} - \frac{21x^2}{100} + \frac{x}{20} + 2$$

(und ohne definierte Parität) ist in Abbildung 4.5 zu sehen.

---

<sup>1</sup>Der Ausdruck  $0^0$  ist im allgemeinen nicht definiert, da es sinnvolle Argumente gibt, ihn als 1 zu erklären (so wie hier), aber auch Argumente für den Wert 0.

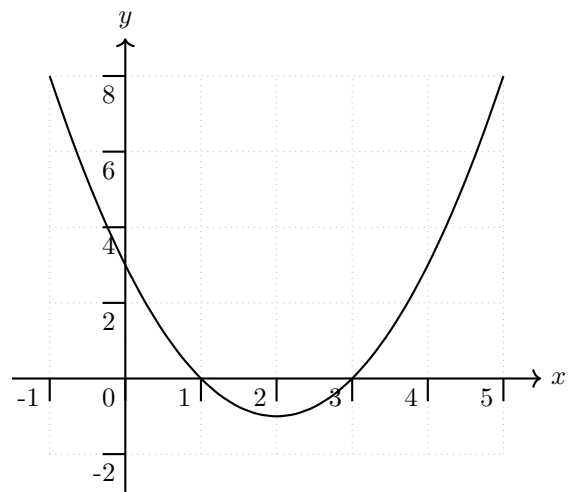


Abbildung 4.4: Quadratische Funktion  $f(x) = x^2 - 4x + 3$

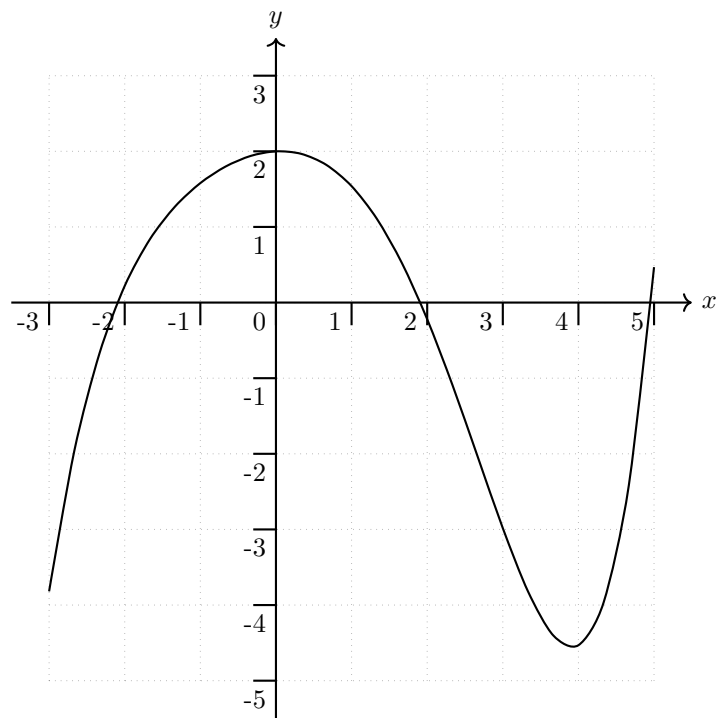


Abbildung 4.5: Eine ganzrationale Funktion fünften Grades



**Definition 4.10** (Gebrochenrationale Funktionen). Der Quotient zweier ganzrationaler Funktionen  $p$  und  $q$  heißt gebrochenrationale Funktion. Der Definitionsbereich von  $\frac{p}{q}$  entspricht den reellen Zahlen, jedoch ohne die Nullstellen der Nennerfunktion, also die Stellen  $x$  für die  $q(x) = 0$ .

**Beispiele:**

- Wir hatten weiter oben bereits eine gebrochenrationale Funktion mit hebbarer Definitionslücke gesehen; diese ließ sich zur ganzrationalen Funktion mit der Vorschrift

$$x \mapsto \frac{1}{3}x - \frac{2}{3}$$

vereinfachen.

- Die Funktion  $f$  mit der Vorschrift

$$f(x) := \frac{x+1}{x+3}$$

ist gebrochenrational. Die Definitionslücke bei  $x = (-3)$  ist *nicht* hebbar; es handelt sich vielmehr um eine *Polstelle mit Vorzeichenwechsel* – bei Annäherung an den Wert  $x = (-3)$  laufen die Funktionswerte nach positiv Unendlich (von links) bzw. nach negativ Unendlich (von rechts).

Der Graph ist in Abbildung 4.6 zu sehen. Die Polstelle ist mit einer senkrechten gestrichelten Linie markiert. Die waagerechte gestrichelte Linie bei  $y = 1$  heißt *Asymptote*; sie beschreibt das Verhalten der Funktion, wenn  $x$  in Richtung positiv oder negativ Unendlich läuft – mehr hierzu in Mathematik 2.

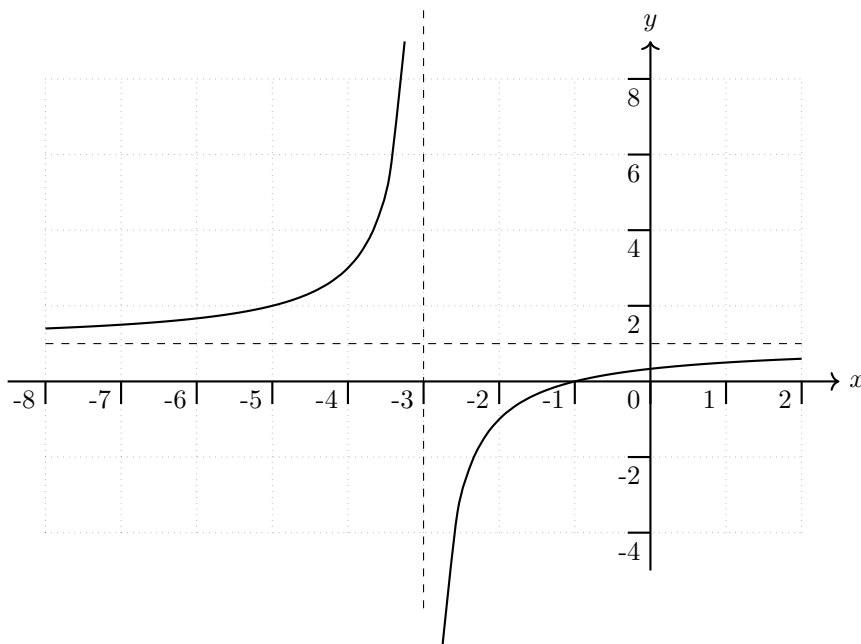


Abbildung 4.6: Gebrochenrationale Funktion  $f(x)$  (Polstelle und Asymptoten markiert)

Die Funktion ist injektiv – dies zeigen wir wie oben, indem wir  $f(a) = f(b)$  auf  $(a = b)$  folgern:

$$\begin{aligned} f(a) &= f(b) \\ \Leftrightarrow \frac{a+1}{a+3} &= \frac{b+1}{b+3} \\ \Leftrightarrow (a+1)(b+3) &= (b+1)(a+3) \\ \Leftrightarrow ab + 3a + b + 3 &= ab + 3b + a + 3 \\ \Leftrightarrow 2a &= 2b \\ \Leftrightarrow a &= b \end{aligned}$$

Für die Surjektivität prüfen wir, ob jedes  $y \in \mathbb{R}$  ein Urbild besitzt, indem wir  $y = f(x)$  anschreiben und versuchen, nach  $x$  umzustellen. Gelingt das überall, so ist  $f$  surjektiv. Gelingt

das sogar eindeutig, dann ist auch direkt die Umkehrfunktion gefunden.

$$\begin{aligned}
 y = f(x) &= \frac{x+1}{x+3} \\
 \Leftrightarrow y(x+3) &= x+1 \\
 \Leftrightarrow xy+3y &= x+1 \\
 \Leftrightarrow xy-x+3y &= 1 \\
 \Leftrightarrow xy-x &= 1-3y \\
 \Leftrightarrow x(y-1) &= 1-3y \\
 \Leftrightarrow x &= \frac{1-3y}{y-1}
 \end{aligned}$$

Hierbei wurde zunächst durch Multiplikation mit  $(x+3)$  der Nenner auf die linke Seite gebracht. Das war zulässig, da  $x \neq (-3)$ , d.h. wir haben die Gleichung nicht mit Null multipliziert. Dann wurde die Klammer ausmultipliziert, und Terme mit  $x$  links gesammelt.

Im vorletzten Schritt lässt sich  $x$  ausklammern; danach wird noch spekulativ durch  $(y-1)$  dividiert – dies ist aber nur dann zulässig, wenn  $y \neq 1$ , denn sonst würden wir durch Null dividieren.

Für den Fall  $y = 1$  ist die vorletzte obige Gleichung gar nicht lösbar (sie würde “ $0 = (-2)$ ” lauten und obendrein keinen Wert für  $x$  ergeben).

Also ist  $f$  surjektiv, falls wir den Bildbereich auf den Wertebereich  $\mathbb{R} \setminus \{1\}$  anpassen.

Insgesamt ist also folgende Funktion bijektiv:

$$f : \mathbb{R} \setminus \{-3\} \rightarrow \mathbb{R} \setminus \{1\}; \quad x \mapsto \frac{x+1}{x+3}$$

Die Umkehrfunktion lautet:

$$f^{-1} : \mathbb{R} \setminus \{1\} \rightarrow \mathbb{R} \setminus \{-3\}; \quad x \mapsto \frac{1-3x}{x-1}$$

#### 4.2.3 Trigonometrische Funktionen (mehr dazu in Mathematik 2/Analysis)

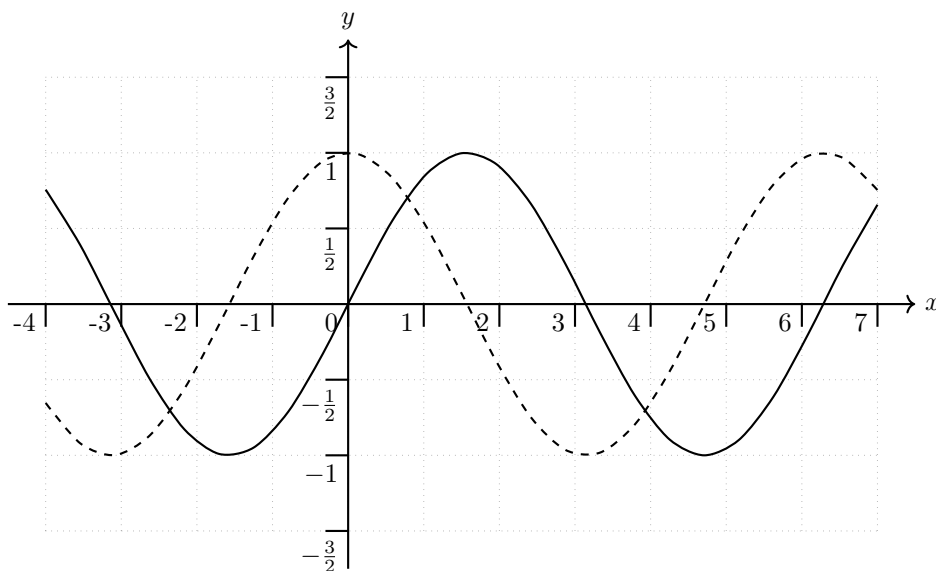


Abbildung 4.7: Funktionen  $\sin(x)$  (durchgezogen) und  $\cos(x)$  (gestrichelt);  $x$  im Bogenmaß

- surjektiv, falls Bildbereich als  $[-1, 1]$  gewählt wird
- injektiv nur auf Abschnitten, z.B. für  $D := [-\pi/2, \pi/2]$  für  $\sin(x)$ , oder  $D := [0, \pi]$  für  $\cos(x)$
- Falls  $D$  und  $B$  geeignet gewählt sind, lauten die Umkehrfunktionen  $\sin^{-1}(x) =: \arcsin(x)$  bzw.  $\cos^{-1}(x) =: \arccos(x)$  (Arcussinus, Arcuscosinus).

- Die *Kreiszahl*  $\pi \approx 3.14159265$  beschreibt das Verhältnis von Umfang und Durchmesser jedes beliebigen Kreises.
- Das *Bogenmaß* beschreibt Winkel als Länge eines Kreisbogens mit Radius 1. Der volle Kreis hat dann den Winkel  $2\pi$ , was 360 Grad entspräche.
- Die Funktionen Sinus und Cosinus beschreiben auch die Seitenverhältnisse im rechtwinkligen Dreieck:

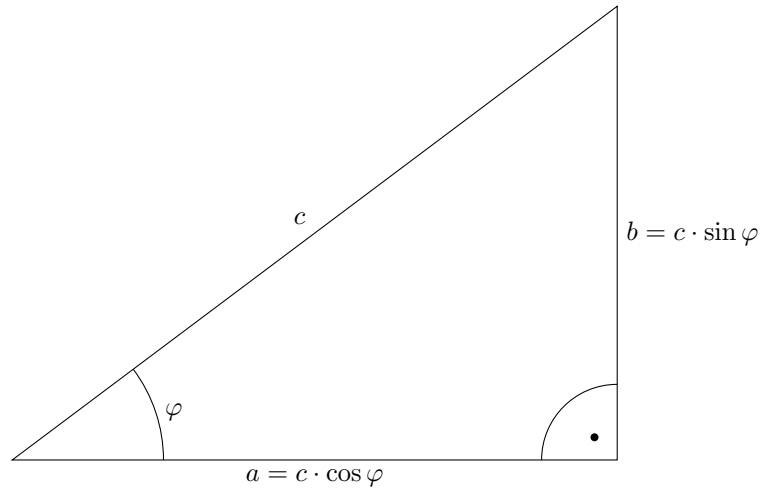


Abbildung 4.8: Längenverhältnisse im rechtwinkligen Dreieck

#### 4.2.4 Exponentialfunktion (mehr dazu in Mathematik 2/Analysis)

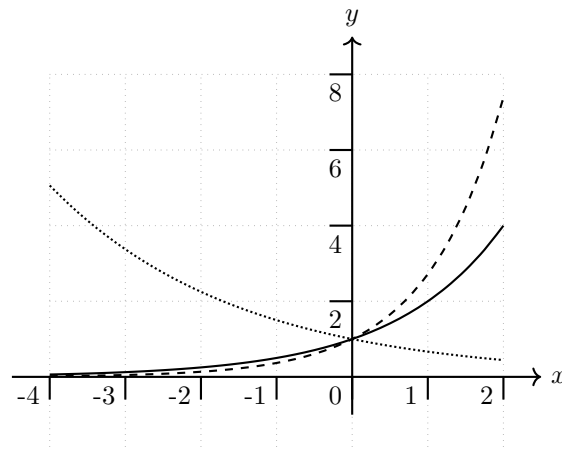


Abbildung 4.9: Funktionen  $2^x$  (durchgezogen),  $e^x$  (gestrichelt) und  $\left(\frac{3}{2}\right)^{-x}$  (gepunktet)

- Definitionsbereich  $D = \mathbb{R}$
- injektiv
- surjektiv, falls Bildbereich als  $(0, \infty)$  gewählt wird
- Die Umkehrfunktion heißt *Logarithmus*; die Basis des Logarithmus wird als Subskript notiert. Für die Beispielfunktionen aus Abbildung 4.9 würden die entsprechenden Umkehrfunktionen lauten:  $\log_2(x)$ ,  $\log_e(x) = \ln(x)$  und  $-\log_{3/2}(x)$
- Die *Eulersche Zahl*  $e \approx 2.7182818$  ist die Basis, für die die Tangente an die zugehörige Exponentialfunktion an der Stelle  $x = 0$  genau die Steigung 1 besitzt. Sauberere und geschlossenere Definitionen dann in Mathematik 2.

## 4.3 Permutationen (Einführung)

Wir führen in diesem Abschnitt den Begriff der Permutationsfunktion ein – das ist eine Funktion, die eine endliche Menge bijektiv auf sich selbst abbildet.

Wir betrachten zunächst nur einige grundsätzliche Eigenschaften. Die Permutationen begegnen uns später nochmal im Algebra-Kapitel 5, wenn es um die *Symmetrische Gruppe* geht, und im Kapitel 8, wenn wir uns mit *Determinanten* beschäftigen.

---

Für diesen Abschnitt wollen wir aus praktischen Gründen noch eine parametrische Menge  $M_n$  definieren, für  $n \in \mathbb{N}$ :

$$M_n := \{j \in \mathbb{N} \mid 1 \leq j \leq n\} = \{1, 2, \dots, n\}$$

### 4.3.1 Permutationsfunktionen

**Definition 4.11** (Permutation). Für  $n \in \mathbb{N}$  heißt eine Abbildung  $\sigma : M_n \rightarrow M_n$  Permutation auf  $M_n$  genau dann, wenn sie bijektiv ist.

Die Abbildung

$$\text{id}_n : M_n \rightarrow M_n; \quad j \mapsto j$$

heißt dabei identische Permutation.

In Tabellenschreibweise wird die Abbildungsvorschrift von  $\sigma$  notiert, indem in der ersten Zeile die Elemente  $j$  aus  $M_n$  geordnet aufgelistet werden; in der zweiten Zeile jeweils ihre Bilder  $\sigma(j)$ :

$$\begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{pmatrix}$$

**Bemerkungen:**

- Jede Zahl aus  $M_n$  hat dann genau ein Bild und genau ein Urbild.
- Oft ist ein Schaubild mit den Abbildungspfeilen (also: der Funktionsgraph) sehr hilfreich, um das Verhalten einer Permutation zu erfassen.

**Beispiele:**

- Die Abbildung  $\sigma$  auf  $M_5$  mit der folgenden Vorschrift ist eine Permutation:

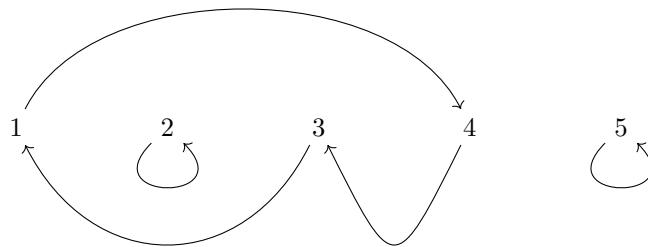
$$1 \mapsto 4; \quad 2 \mapsto 2; \quad 3 \mapsto 1; \quad 4 \mapsto 3; \quad 5 \mapsto 5$$

Die Tabellenschreibweise lautet:

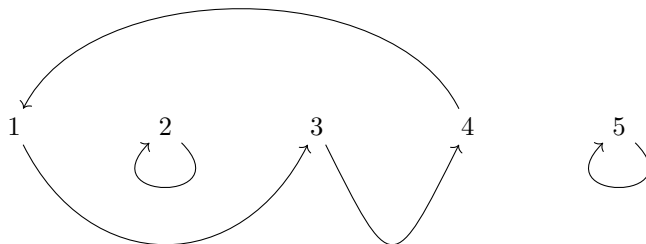
$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 2 & 1 & 3 & 5 \end{pmatrix}$$

(jede Zahl aus  $M_5$  kommt in beiden Zeilen jeweils genau einmal vor).

Schaubild:



- Die Umkehrabbildung (eine solche existiert bei Permutationen immer; s.u.) zum vorigen Beispiel können wir aus dem Schaubild direkt konstruieren, indem alle Pfeilrichtungen umgedreht werden: Daraus lässt sich wiederum die Tabellenschreibweise gut ablesen:



$$\sigma^{-1} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 2 & 4 & 1 & 5 \end{pmatrix}$$

- Für die Abbildung  $\text{id}_5$  lautet die Tabellenschreibweise:

$$\text{id}_5 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 3 & 4 & 5 \end{pmatrix}$$

Und das Schaubild:



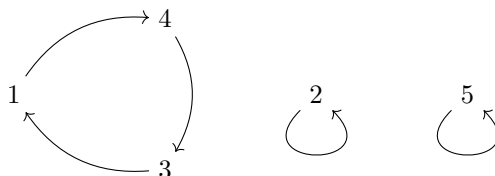
Für die Umkehrabbildung gilt hier natürlich:  $\text{id}_5^{-1} = \text{id}_5$  (das gilt für jede identische Permutation analog).

Die Frage, wie viele verschiedene Permutationen von  $M_n$  es geben kann, wurde übrigens schon beantwortet. Bedenkt man, dass für festes  $n$  die Tabellenschreibweisen der verschiedenen Permutationen sich nur in den jeweiligen unteren Zeilen unterscheiden können, führt das zur Frage, wie viele verschiedene solche Zeilen es geben kann.

Diese Zahl entspricht genau der Zahl möglicher  $n$ -Tupel aus  $n$  Elementen, und nach Satz 1.40 sind dies genau  $n!$ .

### 4.3.2 Zyklen und Transpositionen

Zur Motivation dieses Unterabschnitts geben wir nochmal das Schaubild der obigen Beispielpermutation auf  $M_5$  an, jedoch in “entwirrter” Form:



Wir sehen im Funktionsgraph drei disjunkte *Zyklen*: Einmal von 1 über 4 und 3 zurück zu 1; einmal von 2 zu 2; und einmal von 5 zu 5.

Wählt man eine Zahl aus  $M_5$  beliebig aus, befindet man sich in genau einem dieser Zyklen an einer gewissen Stelle. Wendet man nun immer wieder die Abbildungsvorschrift an, so bewegt man sich entlang der Pfeilrichtung innerhalb genau dieses Zyklus jeweils um eine Position weiter und gelangt irgendwann zur ursprünglichen Zahl zurück. Zum Beispiel ist

$$\sigma(\sigma(\sigma(3))) = \sigma(\sigma(1)) = \sigma(4) = 3$$

Auch für die beiden Elemente 1 und 4 führt das dreimalige Anwenden von  $\sigma$  auf den Ausgangswert zurück. Bei 2 und 5 reicht schon die einmalige Anwendung von  $\sigma$  (die Permutation  $\text{id}_5$  bestand (s.o.) aus fünf solcher einelementigen Zyklen).

Wir wollen motivieren, dass solch eine entwirrte Darstellung des Permutationsgraphs für jede beliebige Permutation möglich ist. Hierzu brauchen wir zunächst folgende

**Definition 4.12** (Zyklus (Permutation)). Für  $n \in \mathbb{N}$  und  $k \in \mathbb{N}$  mit  $1 \leq k \leq n$  sei eine  $k$ -elementige Teilmenge von  $M_n$  gegeben als

$$A_k := \{a_1, a_2, \dots, a_k\}$$

Eine Permutation  $\zeta$  auf  $M_n$ , die alle Elemente aus  $M_n \setminus A_k$  auf sich selbst abbildet, und für die Elemente in  $A_k$  das folgende Verhalten zeigt

$$\zeta(a_1) = a_2; \quad \zeta(a_2) = a_3; \quad \dots \quad \zeta(a_{k-1}) = a_k; \quad \zeta(a_k) = a_1,$$

heißt Zyklus der Länge  $k$ . Die Länge wird mit Betragsstrichen ausgedrückt:  $|\zeta| = k$ .

Man notiert den Zyklus durch Angabe der  $a_j$  per

$$(a_1 \ a_2 \ \dots \ a_k)$$

Zwei Zyklen heißen *disjunkt*, wenn ihre  $A_k$ -Mengen (die Mengen der zyklisch durchgetauschten Elemente aus  $M_n$ ) disjunkt sind.

Ein Zyklus der Länge 2 heißt Transposition. Man notiert die Transposition, die die Elemente  $a$  und  $b$  aus  $M_n$  vertauscht, als

$$\tau_{a,b} := (a \ b) = (b \ a) = \tau_{b,a}$$

#### Bemerkungen:

- Der Ausdruck “ $A_k$ -Menge” ist *kein* formaler Begriff! Wir verwenden ihn hier nur als Kurzschreibweise für “die  $k$  vom Zyklus betroffenen Elemente”. Dass wir diese explizit auch als Menge notiert haben, soll betonen, dass die betroffenen Elemente alle genau einmal gezählt werden ( $A_k$  enthält keine Dubletten).
- Jeder Zyklus der Länge 1 entspricht der identischen Permutation  $\text{id}_n$ .
- Ein Zyklus der Länge  $k$  kann auf  $k$  verschiedene Arten notiert werden:

$$\begin{aligned} (a_1 \ a_2 \ \dots \ a_{k-1} \ a_k) &= (a_2 \ a_3 \ \dots \ a_k \ a_1) \\ &= (a_3 \ a_4 \ \dots \ a_1 \ a_2) \\ &= \dots \\ &= (a_k \ a_1 \ \dots \ a_{k-2} \ a_{k-1}) \end{aligned}$$

Das Abbildungsverhalten ist stets das gleiche.

- Jede Transposition entspricht ihrer eigenen Umkehrabbildung.
- Wenn man einen Zyklus der Länge  $k$  genau  $k$ -mal hintereinander anwendet, erhält man effektiv die identische Permutation (dann sind die  $k$  betroffenen Elemente  $k$ -mal zyklisch durchgetauscht worden).

#### Beispiele:

- In obiger Beispielpermutation  $\sigma$  auf  $M_5$  gibt es einen Zyklus der Länge 3 und zwei Zyklen der Länge 1; diese sind:

$$(1 \ 4 \ 3); \quad (2); \quad (5)$$

- In der Umkehrpermutation liegt dieselbe Situation vor, mit den Zyklen

$$(3 \ 4 \ 1); \quad (2); \quad (5)$$

- *Kein Zyklus* ist folgender Ausdruck:

$$(1 \quad 2 \quad 3 \quad 2 \quad 3)$$

Nicht nur enthält er die Elemente 2 und 3 jeweils doppelt, was wir ausgeschlossen hatten. Sondern wenn man den Ausdruck nach den Regeln von oben liest, erhält man u.a. die beiden Abbildungsvorschriften

$$3 \mapsto 2 \quad \text{und} \quad 3 \mapsto 1$$

Die zweite Vorschrift geht aus dem letzten Element der Zyklus-Sequenz hervor, das durch die Permutation wieder auf deren erstes Element abgebildet werden müsste.

Hier liegt somit noch nicht einmal eine Funktion vor, denn das Element 3 müsste eindeutig zugeordnet werden!

Für eine beliebige Permutation  $\sigma$  auf  $M_n$  ist die Angabe der Zyklen, die sich aus dem Schaubild direkt ablesen lassen, eine Alternative zur Tabellenschreibweise:

**Definition 4.13** (Kanonische Zyklen einer Permutation). *Für  $n \in \mathbb{N}$  und  $\sigma$  eine Permutation auf  $M_n$  ist eine Dekomposition/Zerlegung in kanonische Zyklen (oder kurz: Zykelschreibweise) gegeben, falls die Mengen der von den Zyklen betroffenen Elemente eine Zerlegung von  $M_n$  bilden (d.h. falls diese Mengen alle nichtleer und paarweise disjunkt sind, und falls ihre Vereinigung genau  $M_n$  ergibt).*

*Man notiert die kanonischen Zyklen als Komposition, schreibt aber i.A. den Kringel-Operator nicht mit an.*

#### Bemerkungen:

- Zur Komposition von Permutationen siehe den nächsten Unterabschnitt.
- Eine Komposition von Zyklen auf  $M_n$  ist also immer genau dann kanonisch, wenn jedes Element von  $n$  in genau einem der Zyklen vorkommt – in allen anderen notierten Zyklen jedoch nicht.

#### Beispiele:

- Für die obige Beispielpermutation  $\sigma$  auf  $M_5$  ist z.B. folgende Zerlegung kanonisch:

$$\sigma = (1 \quad 4 \quad 3)(2)(5)$$

- Für die Permutation  $\text{id}_5$  ist eine kanonische Zerlegung durch

$$\text{id}_5 = (1)(2)(3)(4)(5)$$

gegeben.

Zwei wichtige Fragen sind allerdings noch ungeklärt:

- Ist die Reihenfolge der kanonischen Zyklen von Bedeutung?
- Findet sich überhaupt für jede beliebige Permutation eine kanonische Zerlegung?

Dies wollen wir jetzt behandeln.

**Satz 4.14** (Kommutativität disjunkter Zyklen). *Für  $n \in \mathbb{N}$  und  $\zeta_1, \zeta_2$  disjunkte Zyklen auf  $M_n$  ist die Komposition kommutativ, d.h.*

$$\zeta_1 \circ \zeta_2 = \zeta_2 \circ \zeta_1$$

**Beweis:** Da die beiden Zyklen disjunkt sind, lässt  $\zeta_1$  alle Elemente, die sich unter  $\zeta_2$  ändern, fest; analog lässt  $\zeta_2$  alle Elemente aus  $M_n$  unverändert, die von  $\zeta_1$  betroffen sind. ■

**Bemerkung:** Demnach ist es bei kanonischen Zyklen unwesentlich, in welcher Reihenfolge sie notiert werden.

**Beispiel:** Also wäre z.B. auf  $M_6$  sicher:

$$(1 \ 4 \ 5)(2 \ 6 \ 3) = (2 \ 6 \ 3)(1 \ 4 \ 5)$$


---

Für den zweiten ausstehenden Punkt bedenken wir, dass eine beliebige Permutation  $\sigma$  auf  $M_n$  stets eine bijektive Abbildung ist. Jedes Element aus  $M_n$  hat also genau einen Nachfolger (der durch Anwenden von  $\sigma$  erreicht wird) und einen Vorgänger (aus dem es durch Anwenden von  $\sigma$  hervor geht).

Es können zwei Fälle auftreten:

- Falls für  $j \in M_n$  der Nachfolger  $\sigma(j)$  identisch mit  $j$  ist, muss auch der Vorgänger von  $j$  genau  $j$  entsprechen. Dann ist  $j$  Bestandteil des kanonischen Einerzyklus  $(j)$ .
- Anderenfalls ist  $\sigma(j) \neq j$ . Wenden wir nun  $n$ -mal die Permutation auf  $j$  an, so können wir höchstens  $(n - 1)$ -mal neue, noch nicht besuchte Elemente aus  $M_n$  erreichen, da  $M_n$  nur genau  $n$  verschiedene Zahlen enthält. Spätestens nach der  $n$ -ten Anwendung von  $\sigma$  müssen wir also einen Wert aus  $M_n$  in der Abfolge

$$j, \sigma(j), \sigma(\sigma(j)), \dots$$

doppelt antreffen. Wir verkürzen nun die Abfolge so, dass wir beim ersten solchen doppelt vorkommenden Wert  $x$  mit der Anwendung von  $\sigma$  aufhören. Dabei kann eine der beiden folgenden Situationen eintreten:

- Entweder  $x = j$  – dann hätten wir einen Zyklus von  $j$  nach  $j$  gefunden.
- Oder  $x \neq j$ . Dann gibt es aber immerhin einen Zyklus von  $x$  nach  $x$ . Wenn wir von  $j$  zum “ersten”  $x$  gelangen, muss man in Gegenrichtung wegen der Bijektivität auch von  $x$  zu  $j$  zurück gelangen (durch Anwenden von  $\sigma^{-1}$ ).

Natürlich kann es aber keinen Unterschied machen, ob wir vom “ersten”  $x$  oder von “zweiten”  $x$  aus rückwärts zu  $j$  laufen. Falls wir aber letzteres tun, so müssen wir  $j$  auf dem Zyklus von  $x$  nach  $x$  erreichen.

Folglich muss  $j$  aber schon während des Zyklus von  $x$  nach  $x$  ein zweites Mal aufgetreten sein; also war  $x$  falsch gewählt! ✗

Also gelangen wir von  $j$  aus durch genügend häufiges Anwenden von  $\sigma$  auch in diesem Fall stets zu  $j$  zurück. Dabei wird keine Zahl aus  $M_n$  doppelt besucht, denn  $j$  ist der erste doppelt auftretende Wert in der obigen Abfolge. Es liegt also auch hier ein Zyklus vor.

Alle Elemente dieses Zyklus sind über die Permutation  $\sigma$  durch diesen Zyklus bereits beschrieben.

Zusammenfassend finden wir also eine kanonische Zerlegung für gegebenes  $\sigma$  folgendermaßen:

1. (Initialisierung:) Konstruiere die Menge  $N := M_n$ .
2. (Finde einen Zyklus:) Solange  $N \neq \emptyset$ : Wähle aus  $N$  ein Element  $j$  aus. Verfolge durch wiederholtes Anwenden von  $\sigma$ , welche Elemente von  $N$  aufgefunden werden, bis  $j$  ein zweites Mal erreicht wird. Dadurch ergibt sich der Zyklus  $(j \ \dots)$  mit allen unterwegs erreichten Elementen. Entferne alle von diesem Zyklus betroffenen Elemente aus  $N$  (denn ihr Abbildungsverhalten ist mit diesem Zyklus bereits beschrieben).

Da mit jedem Zyklus mindestens ein Element aus  $N$  entfernt wird, schließt der Algorithmus nach spätestens  $n$  Durchläufen ab (letzteres, falls  $\sigma$  die identische Permutation war). Zusammen fassend gilt also folgender

**Satz 4.15** (Zerlegbarkeit einer Permutation in kanonische Zyklen). *Jede Permutation  $\sigma$  lässt sich in kanonische Zyklen zerlegen.*



### Bemerkungen:

- Die kanonischen Zyklen sind genau die, welche sich aus dem Funktionsgraph einer Permutation ablesen lassen. Solch ein Graph besteht nach dem oben gezeigten stets aus disjunkten Zyklen. Fertigt man also zu einer in Tabellenschreibweise gegebenen Permutation das Schaubild an, erhält man die kanonische Zerlegung direkt mit.
- Umgekehrt lässt sich aus jeder kanonischen Zerlegung leicht die Tabellenschreibweise einer Permutation rekonstruieren, indem man die Abbildungsregeln einzeln aus den Zyklen abliest.

### Beispiele:

- Wir betrachten die folgende Permutation auf  $M_4$ :

$$\sigma := \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix}$$

Wir beginnen mit dem Wert 3. Nun verfolgen wir die Anwendung von  $\sigma$ , bis wir wieder bei 3 ankommen, und finden damit den kanonischen Zyklus, der 3 enthält. Dabei finden wir:

$$3 \mapsto 4 \mapsto 1 \mapsto 2 \mapsto 3$$

Damit ist der Zyklus gefunden als  $(3 \ 4 \ 1 \ 2)$ . Und das ist auch schon die gesamte Permutation  $\sigma$ , da das Abbildungsverhalten aller vier Elemente von  $M_4$  damit vollständig beschrieben ist. Wir können die Auflistung der Zykelselemente an einem beliebigen Punkt beginnen, z.B. auch bei 1. Damit erhalten wir:

$$\sigma = (1 \ 2 \ 3 \ 4)$$

Hier ist also die gesamte Permutation ein Zyklus der Länge 4.

Schaubild:



- Wir betrachten die folgende Permutation auf  $M_6$ :

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

Wiederum suchen wir die kanonischen Zyklen. Wir beginnen dieses Mal mit 1:

$$1 \mapsto 4 \mapsto 6 \mapsto 1$$

Dies liefert uns den Zyklus  $(1 \ 4 \ 6)$ .

Von den restlichen verbleibenden Elementen aus  $M_6$  wählen wir wieder eines aus, z.B. die 2:

$$2 \mapsto 3 \mapsto 2$$

Das ergibt den Zyklus  $(2 \ 3)$ , eine Transposition (nämlich  $\tau_{2,3}$ ).

Es bleibt nur ein Element aus  $M_6$  übrig, und tatsächlich liefert  $5 \mapsto 5$  den Zyklus  $(5)$ .

Insgesamt also:

$$\sigma = (1 \ 4 \ 6)(2 \ 3)(5)$$

Schaubild in Abbildung 4.10.

Abbildung 4.11 zeigt eine gleichwertige Darstellung des Graphen, bei der die disjunkten Zyklen voneinander abgesetzt dargestellt sind.

- Wir rekonstruieren für  $M_7$  aus einer kanonischen Zerlegung die Tabellenschreibweise. Sei also

$$\sigma := (2 \ 1 \ 5)(3 \ 7)(4 \ 6)$$

Dann lesen wir aus den drei Zyklen ab:

$$2 \mapsto 1 \mapsto 5 \mapsto 2; \quad 3 \mapsto 7 \mapsto 3; \quad 4 \mapsto 6 \mapsto 4$$

Also:

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 5 & 1 & 7 & 6 & 2 & 4 & 3 \end{pmatrix}$$

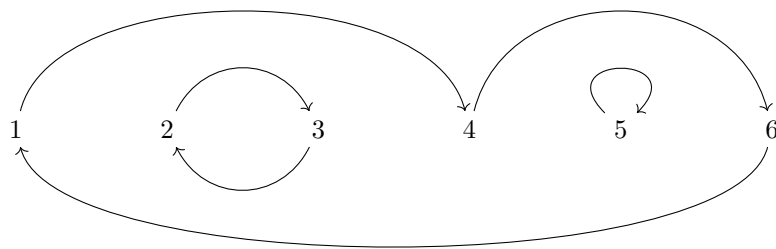


Abbildung 4.10: Eine Permutation mit drei kanonischen Zyklen auf  $M_6$

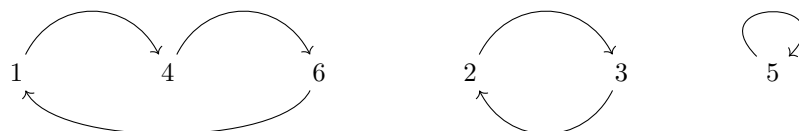


Abbildung 4.11: Dieselbe Permutation; im Funktionsgraph sind die kanonischen Zyklen voneinander abgesetzt

### 4.3.3 Verkettung von Permutationen

Zwei Permutationen  $\sigma_1$  und  $\sigma_2$  auf  $M_n$  lassen sich stets verketteten (hintereinander ausführen); das Resultat ist wiederum eine Permutation auf  $M_n$ , denn es entsteht erneut eine bijektive Funktion auf der gleichen Menge; siehe auch die entsprechende Bemerkung zu Definition 4.4 (Komposition von Funktionen).

Insbesondere ist die identische Permutation  $\text{id}_n$  *neutral* bezüglich der Verkettung (genauerer hierzu später im Kapitel 5), denn es gilt für jede beliebige Permutation  $\sigma$  auf  $M_n$ :

$$\sigma \circ \text{id}_n = \text{id}_n \circ \sigma = \sigma$$

#### Bemerkungen:

- Für das Ausführen der Komposition ist die *Tabellenschreibweise* die nützlichere. Bei der Zyklenschreibweise wird schnell übersehen, dass das rechte Element der Sequenz auch noch auf das linke Ende abgebildet wird; außerdem kommen die Elemente meist nicht aufsteigend geordnet in den Sequenzen vor, sodass die Suche nach den relevanten Abbildungsvorschriften etwas unhandlicher ist.
- Man beachte, dass einzelne disjunkte Zyklen zwar kommutieren – aber für beliebige Permutationen gilt dies meist *nicht*.

#### Beispiele:

- Wir betrachten auf  $M_3$  die beiden Permutationen

$$\sigma := \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \quad \text{und} \quad \pi := \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$$

(Hierbei ist  $\sigma = (1)(2 \ 3) = \tau_{2,3}$  eine Transposition und  $\pi = (1 \ 2 \ 3)$  ein Dreierzyklus.)

Wir berechnen die vier Kompositionen aus jeweils zwei Permutationen, die hier möglich sind:

- Zunächst  $\pi \circ \sigma$ . Zuerst ist also  $\sigma$  anzuwenden, danach  $\pi$ . Wir notieren als Hilfe (und zur Übung) die Zuordnungen von  $\sigma$ , darunter die von  $\pi$ , und verfolgen dann die Abbildung der Elemente aus  $M_3$  von oben nach unten, wie in Abbildung 4.12 gezeigt: Die von  $\sigma$  abgebildeten Werte sind die Argumente für  $\pi$ .

Insgesamt lesen wir also ab:

$$\pi \circ \sigma = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$$

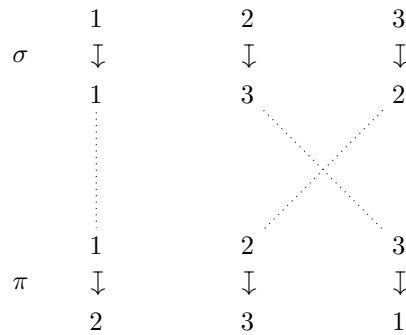


Abbildung 4.12: Zwei Permutationen; Nachverfolgung der Abbildung für  $\pi \circ \sigma$

- Für die umgekehrte Komposition geben wir auch noch die Nachverfolgung an; siehe Abbildung 4.13.

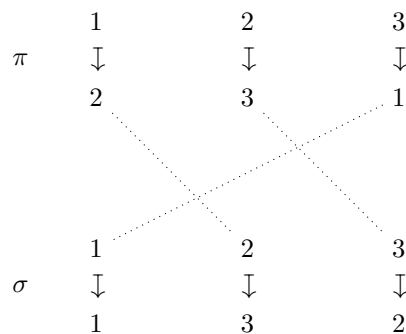


Abbildung 4.13: Zwei Permutationen; Nachverfolgung der Abbildung für  $\sigma \circ \pi$

Hier ergibt sich:

$$\sigma \circ \pi = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \neq \pi \circ \sigma$$

Wir haben hier also auch direkt ein Beispiel für zwei Permutationen, die *nicht* kommutieren!

- Nach etwas Übung benötigt man die Skizzen mit Hilfslinien nicht mehr und kann die komponierten Transpositionen über die beiden jeweiligen Tabellenschreibweisen ablesen. Man achte jedoch darauf, die Reihenfolge nicht zu verwechseln.

Wir geben ohne Hilfs-Skizze noch die beiden anderen Kompositionen an:

$$\sigma \circ \sigma = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} = \tau_{2,3} \circ \tau_{2,3} = \text{id}_3 \quad \text{und} \quad \pi \circ \pi = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$$

Wir sehen hier auch, dass sich eine Transposition durch wiederholtes Anwenden rückgängig machen lässt.

- Als letztes wollen wir noch die Bemerkung zu Definition 4.12 demonstrieren. Da  $\pi$  ein Dreierzyklus ist, müsste  $\pi^3 := \pi \circ \pi \circ \pi$  wieder die identische Permutation ergeben. Und tatsächlich findet man mit dem oben berechneten  $\pi^2$ :

$$\pi^3 = \pi \circ (\pi \circ \pi) = (\pi \circ \pi) \circ \pi = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} = \text{id}_3$$

- Wir betrachten auf  $M_6$  die Kompositionen der beiden Zyklen

$$\sigma := (1 \ 4) \quad \text{und} \quad \pi := (1 \ 2 \ 3 \ 4 \ 5)$$

- Zunächst  $\pi \circ \sigma = (1 \ 2 \ 3 \ 4 \ 5)(1 \ 4)$ . Wir haben offensichtlich hier keine kanonische Zerlegung einer  $M_6$ -Permutation, denn die Zahlen 1 und 4 treten doppelt auf. Der Verdacht liegt also nahe, dass  $\sigma$  und  $\pi$  nicht kommutieren.

Wir rekonstruieren die komponierte Permutation hier auf die oben *nicht* empfohlene Weise, ohne Umweg über die Tabellenschreibweise, um zu demonstrieren, dass dies etwas aufwändiger ist.

Die 1 tritt in beiden Zyklen auf; wir müssen also zunächst den rechten Zyklus auswerten, der uns die Vorschrift  $1 \mapsto 4$  liefert. Dann folgt der linke Zyklus, der diese 4 auf 5 abbildet.

Die Zahlen 2 und 3 treten nur im linken Zyklus auf, sodass wir dort direkt ihre Bilder (3 und 4) ablesen können.

Die 4 wird im rechten Zyklus zunächst auf 1 abgebildet, welche vom linken Zyklus danach noch auf 2 abgebildet wird.

Für 5 ergibt sich ausschließlich aus dem linken Zyklus die Abbildung auf 1.

Die Zahl 6 taucht in keinem der Zyklen auf und wird daher auf sich selbst abgebildet.

Insgesamt (die kanonischen Zyklen lesen wir aus der Tabellenschreibweise direkt ab):

$$\pi \circ \sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 5 & 3 & 4 & 2 & 1 & 6 \end{pmatrix} = (1 \ 5)(2 \ 3 \ 4)(6)$$

- Nun  $\sigma \circ \pi = (1 \ 4)(1 \ 2 \ 3 \ 4 \ 5)$ . Hier geben wir zunächst die Tabellenschreibweisen an:

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 4 & 2 & 3 & 1 & 5 & 6 \end{pmatrix} \quad \text{und} \quad \pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 3 & 4 & 5 & 1 & 6 \end{pmatrix}$$

Nun konstruieren wir die Komposition durch Nachverfolgen wie oben:

$$\sigma \circ \pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 3 & 1 & 5 & 4 & 6 \end{pmatrix} = (1 \ 2 \ 3)(4 \ 5)(6)$$

Auch hier haben wir wieder einen Fall, für den  $\sigma$  und  $\pi$  nicht kommutieren.

- Übrigens ist, da  $\sigma$  eine Transposition ist (nämlich  $\tau_{1,4}$ ),  $\sigma \circ \pi$  auch direkt aus der Tabellenschreibweise von  $\pi$  konstruierbar, indem man in der unteren Zeile (also nach der  $\pi$ -Abbildung) noch die Zahlen 1 und 4 vertauscht.

Achtung: Das funktioniert in Zykelschreibweise so *nicht*! Denn angenommen, wir würden in

$$\pi = (1 \ 2 \ 3 \ 4 \ 5)$$

nachträglich 1 und 4 vertauschen, bekämen wir  $(4 \ 2 \ 3 \ 1 \ 5)$ . Das ist ein Fünferzyklus, und zusammen mit  $(6)$  eine kanonische Zerlegung von

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

was eine völlig andere Permutation ist.

- Zur Übung rechne man gerne noch nach, dass  $\pi^5 = \text{id}_6$ .

#### 4.3.4 Inverse Permutation

Da Permutationen gerade die bijektiven Funktionen auf  $M_n$  sind, existiert selbstverständlich zu jeder Permutation  $\sigma$  eine eindeutige Umkehrpermutation  $\sigma^{-1}$ , sodass

$$\boxed{\sigma \circ \sigma^{-1} = \sigma^{-1} \circ \sigma = \text{id}_n}$$

Hier sind die Tabellenschreibweise und die Zykelschreibweise beide ähnlich im Aufwand:

- Bei der Tabellenschreibweise vertauscht man die beiden Zeilen und ordnet danach die Spalten wieder so an, dass eine Tabellenschreibweise entsteht (d.h. dass in der oberen Zeile die Zahlen 1 bis  $n$  in aufsteigender Reihenfolge angeordnet sind).
- Bei der Schreibweise in kanonischen Zyklen schreibt man die Sequenzen der einzelnen Zyklen in umgekehrter Reihenfolge und erhält den gewünschten Effekt. Da die einzelnen Zyklen disjunkt sind und kommutieren, ist damit alles getan.
- *Vorsicht allerdings* bei Kompositionen von Zyklen, die nicht kanonisch sind! Auch, wenn die Kompositions-Operatoren “ $\circ$ ” nicht mit notiert sind, müssen sie trotzdem mit gedacht werden. Anders als bei den kanonischen Zyklen ist hierbei auch die Reihenfolge der Zyklen (mit umgekehrten Sequenzen) umzukehren, damit das richtige Verhalten erzielt wird – siehe die Bemerkung zur Inversen von komponierten Funktionen bei Definition 4.4.

### Beispiele:

- Wir betrachten das entsprechende Beispiel auf  $M_3$  vom vorigen Unterabschnitt erneut:

$$\sigma := \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} = (1)(2 \ 3) = \tau_{2,3} \quad \text{und} \quad \pi := \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} = (1 \ 2 \ 3)$$

Die inversen Permutationen sind, von den Zyklen abgelesen:

$$\sigma^{-1} = (1)(3 \ 2) \quad \text{und} \quad \pi^{-1} = (3 \ 2 \ 1)$$

Alternativ über die Tabellenschreibweise – zuerst für  $\sigma$ :

$$\sigma^{-1} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}^{-1} = \underbrace{\begin{pmatrix} 1 & 3 & 2 \\ 1 & 2 & 3 \end{pmatrix}}_{\text{Zwischenform}} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} = \sigma$$

Transpositionen sind selbst-invers<sup>2</sup>, d.h. die zweifache Ausführung ergibt stets die identische Permutation.

Und für  $\pi$ :

$$\pi^{-1} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}^{-1} = \underbrace{\begin{pmatrix} 2 & 3 & 1 \\ 1 & 2 & 3 \end{pmatrix}}_{\text{Zwischenform}} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} = \pi^2$$

Die Gleichheit mit  $\pi^{-1} = \pi^2$  kann man anhand der oben ausgeführten Rechnung erkennen – oder man erinnert sich, dass  $\pi$  ein Dreierzyklus ist, d.h.  $\pi^3 = \pi^2 \circ \pi = \text{id}_3$ . Da die Inverse von  $\pi$  aber eindeutig bestimmt ist, muss sie auch  $\pi^2$  entsprechen.

Das ist allgemein richtig: Für einen Zyklus  $\zeta$  mit Länge  $k$  ist  $\zeta^{k-1}$  die Inverse von  $\zeta$ .

Die Zwischenform mit vertauschten Zeilen wird nach etwas Übung meist weggelassen – die Endform der Inversen in Tabellenschreibweise lässt sich auch direkt sukzessive konstruieren, indem man die erste Zeile wie gewöhnlich ausfüllt, und dann die Einträge der unteren Zeile durch Nachschlagen aus der Tabellenschreibweise der ursprünglichen Permutation gewinnt. In  $\pi^{-1}$  ist etwa das Bild von 1 die Zahl, die in  $\pi$  oberhalb der 1 notiert ist – also die 3, das Urbild von 1 in  $\pi$ .

Man beachte, dass die in Tabellenschreibweise ermittelten Inversen genau den in kanonischen Zyklen notierten entsprechen. Liest man aus den Tabellenschreibweisen die kanonischen Zyklen erneut ab, so erhält man die gleichen Darstellungen wie oben, mit zyklisch vertauschten Elementen in den Sequenzen – solche zyklischen Vertauschungen ergeben sich aber nur aus der Frage, an welcher Stelle man in den Zyklus einsteigt; der Zyklus selbst bleibt der gleiche (siehe dazu die entsprechende Bemerkung bei Definition 4.12).

- Nun zu den Kompositionen: Zunächst hatten wir oben  $\pi \circ \sigma$  berechnet als

$$\pi \circ \sigma = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} = (1 \ 2)(3)$$

Die Inverse hiervon ist in Tabellenschreibweise:

$$(\pi \circ \sigma)^{-1} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$$

Aber wir erhalten dies mit den oben berechneten einzelnen Inversen auch per

$$(\pi \circ \sigma)^{-1} = \sigma^{-1} \circ \pi^{-1} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$$

Für die Inverse der Komposition  $\sigma \circ \pi$  lässt sich (Übung) ähnliches nachrechnen.

---

<sup>2</sup>Solche Funktionen nennt man auch *Involutionen*.

- Wir betrachten noch das obige  $M_6$ -Beispiel:

$$\sigma := (1 \ 4) \quad \text{und} \quad \pi := (1 \ 2 \ 3 \ 4 \ 5)$$

Dazu hatten wir schon zwei kanonische Zykelschreibweisen ermittelt:

$$\pi \circ \sigma = (1 \ 5)(2 \ 3 \ 4)(6) \quad \text{und} \quad \sigma \circ \pi = (1 \ 2 \ 3)(4 \ 5)(6)$$

Wir bestimmen die Inversen der beiden Kompositionen hier direkt aus den kanonischen Zyklen (die dabei nicht ungeordnet werden müssen):

$$(\pi \circ \sigma)^{-1} = (5 \ 1)(4 \ 3 \ 2)(6) \quad \text{und} \quad (\sigma \circ \pi)^{-1} = (3 \ 2 \ 1)(5 \ 4)(6)$$

Nun rekonstruieren wir dieses Ergebnis, indem wir die Inversen der Einzelpermutationen komponieren. Zur Berechnung der Komposition setzen wir zeitweise die Tabellenschreibweisen der Zyklen ein; am Schluss leiten wir aus der Tabellenschreibweise wieder die kanonischen Zyklen ab:

$$\begin{aligned} (\pi \circ \sigma)^{-1} &= \sigma^{-1} \circ \pi^{-1} \\ &= (4 \ 1) \circ (5 \ 4 \ 3 \ 2 \ 1) \\ &= \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 4 & 2 & 3 & 1 & 5 & 6 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 5 & 1 & 2 & 3 & 4 & 6 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 5 & 4 & 2 & 3 & 1 & 6 \end{pmatrix} \\ &= (1 \ 5)(2 \ 4 \ 3)(6) \end{aligned}$$

Und analog:

$$\begin{aligned} (\sigma \circ \pi)^{-1} &= \pi^{-1} \circ \sigma^{-1} \\ &= (5 \ 4 \ 3 \ 2 \ 1) \circ (4 \ 1) \\ &= \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 5 & 1 & 2 & 3 & 4 & 6 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 4 & 2 & 3 & 1 & 5 & 6 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 1 & 2 & 5 & 4 & 6 \end{pmatrix} \\ &= (1 \ 3 \ 2)(4 \ 5)(6) \end{aligned}$$

Wir sehen also, dass es sich lohnt, die kanonischen Zyklen einer Permutation zu kennen, und dass wir beim Invertieren von verketteten Zyklen, falls diese nicht kanonisch sind, auch die Zyklen selbst umordnen müssen (hier hatten wir das über die Formeln in den jeweiligen ersten Zeilen der Rechnung automatisch richtig gemacht).

Obwohl wir in den Bemerkungen zu Definition 4.4 bereits für beliebige bijektive Funktionen ähnliches bemerkt hatten, hier speziell für Permutationen nochmals der Hinweis auf die Inverse einer Komposition von Permutationen. Für  $\pi, \sigma$  Permutationen auf  $M_n$  ist

$$\boxed{(\pi \circ \sigma)^{-1} = \sigma^{-1} \circ \pi^{-1}}$$

Falls die Inverse einer Komposition von mehreren *Zyklen* bestimmt werden soll, benötigen wir hierfür zunächst die Inversen der einzelnen Zyklen (jeweils direkt zu erhalten durch Notieren der Zykelselemente in umgekehrter Reihenfolge). Diese invertierten Zyklen sind dann nach obiger eingerahmter Formel wiederum in umgekehrter Reihenfolge zu komponieren. Die Reihenfolge der invertierten Zyklen darf nur dann gleich bleiben, wenn die Zyklen alle paarweise disjunkt zueinander sind.

**Beispiel:** Für die Zyklen  $\zeta_1, \zeta_2$  auf  $M_7$  mit

$$\zeta_1 := (4 \ 7) \quad \text{und} \quad \zeta_2 := (3 \ 5 \ 1 \ 4)$$

stellen wir fest, dass sie *nicht disjunkt* sind, da das Element 4 in beiden Zyklen auftaucht. Hier ist die Reihenfolge der Zyklen bei Komposition also entscheidend, und es ist

$$\begin{aligned} (\zeta_1 \circ \zeta_2)^{-1} &= \left( (4 \ 7) \circ (3 \ 5 \ 1 \ 4) \right)^{-1} = (3 \ 5 \ 1 \ 4)^{-1} \circ (4 \ 7)^{-1} = (4 \ 1 \ 5 \ 3) \circ (7 \ 4) \\ (\zeta_2 \circ \zeta_1)^{-1} &= \dots = (7 \ 4) \circ (4 \ 1 \ 5 \ 3) \end{aligned}$$

# Kapitel 5

## Algebra

### 5.1 Motivation/Grundlagen

In der Algebra wird versucht, Zahlenmengen (oder andere typisierte Mengen mathematischer Objekte) mit den numerischen *Operationen* zusammen zu fassen, die für die Mengenelemente möglich sind. Daraus ergeben sich *algebraische Strukturen*. Der Vorteil solcher Strukturen ist ein Gewinn an Ordnung. Bildet eine Menge mit zwei bestimmten Operationen z.B. einen *Ring*, so ist damit klar, dass die Mengenelemente sich unter diesen Operationen auf eine gewisse (und einheitliche) Weise verhalten, die für alle Ringe gleich ist und damit nicht mehr jedes Mal einzeln nachgewiesen werden muss.

Wir haben teilweise solches Verhalten schon benötigt (vor allem in den Beweisen), wenn z.B. von der Kommutativität der Multiplikation die Rede war, oder von der Assoziativität bei der Komposition von Funktionen.

---

Wir betrachten im folgenden zunächst einige gruppenartige Strukturen, die sich hierarchisch gliedern lassen. Hier werden Mengen jeweils mit einer Operation zusammen betrachtet. Danach befassen wir uns mit ringartigen Strukturen – diese bauen auf dem Konzept der Gruppen auf und erweitern dieses, indem Mengen mit zwei statt mit einer Operation betrachtet werden. Dabei begegnen uns als wichtigste Beispiele die Zahlenmengen  $\mathbb{N}$ ,  $\mathbb{Z}$  und  $\mathbb{Q}$  sowie  $\mathbb{R}$  wieder – aber auch die Restklassensysteme aus Kapitel 3 und die Permutationen aus Kapitel 4 treten wieder auf.

Auch *Vektorräume* sind algebraische Strukturen; diese sind Gegenstand der *linearen Algebra* und werden in den letzten Kapiteln der Vorlesung behandelt. Hierfür dienen die Inhalte dieses Kapitels als Voraussetzung.

---

Zum Schluss noch eine wichtige

**Definition 5.1** (Zweistellige Operation). *Für Mengen  $M, N$  heißt eine Funktion  $*$  :  $M \times M \rightarrow N$  zweistellige (oder: binäre) Operation. Mit Blick auf  $*$  heißen die Objekte aus  $M$  dann Operanden, und die Funktion  $*$  wird Operator genannt.*

*Falls  $N \subseteq M$ , heißt die Operation  $*$  abgeschlossen auf  $M$ .*

*Die Operation heißt kommutativ, falls für alle  $a, b \in M$  gilt:*

$$a * b = b * a$$

*Die Operation heißt assoziativ, falls für alle  $a, b, c \in M$  gilt:*

$$a * (b * c) = (a * b) * c$$

#### Bemerkungen:

- Das Symbol  $*$  steht hier als Platzhalter für eine allgemeine Funktion; dies muss nicht zwingend eine Multiplikation sein!
- Bei zweistelligen Operationen wird der Operator oft (aber nicht immer) *infix* notiert, also zwischen die beiden Operanden geschrieben. Daher ist

$$a + b \quad \text{gleichbedeutend mit} \quad +(a, b)$$

- Neben zweistelligen gibt es auch allgemein  $k$ -stellige Operationen, mit  $k \in \mathbb{N}$ . Operationen mit einem Operanden heißen *unär*.

**Beispiele:** (Die Assoziativität klammern wir als Eigenschaft zunächst aus, da sie im Unterabschnitt zu Halbgruppen explizit behandelt wird)

- Die Addition “+” ist eine abgeschlossene und kommutative Operation auf  $\mathbb{N}$ . Jede Summe von natürlichen Zahlen ist wieder eine natürliche Zahl.
- Die Addition ist auch abgeschlossen und kommutativ auf  $\mathbb{Z}$ ,  $\mathbb{Q}$  und  $\mathbb{R}$ .
- Die Subtraktion “−” ist auch eine Operation auf  $\mathbb{N}$ ; allerdings ist sie *nicht* abgeschlossen. Zum Beispiel ist  $(4 - 7) \notin \mathbb{N}$ . Da aber z.B.  $(7 - 4) \in \mathbb{N}$ , ist die Subtraktion sicher nicht kommutativ.
- Die Subtraktion ist jedoch abgeschlossen auf der Menge  $\mathbb{Z}$ , und auch auf  $\mathbb{Q}$  und  $\mathbb{R}$ .
- Die Subtraktion, nachträglich komponiert mit der Betragsfunktion (siehe Definition 1.35), ist eine abgeschlossene Operation auf  $\mathbb{N}_0$  (jedoch nicht auf  $\mathbb{N}$ , da gleiche natürliche Zahlen eine Differenz von 0 besitzen). Beispiel:

$$|4 - 7| = 3 \in \mathbb{N}_0$$

Wege  $|a - b| = |b - a|$  ist diese kombinierte Operation auch kommutativ.

- Addition und Subtraktion sind abgeschlossen auf  $\mathbb{Z}_m$ , den Restklassensystemen modulo  $m$ . Die Addition ist außerdem kommutativ.
- Die Funktion ggT ist, wenn sie auf genau zwei Argumente eingeschränkt wird, eine abgeschlossene und kommutative Operation auf  $\mathbb{N}$ . Hier wird jedoch der Operator *nicht* infix notiert, sondern in der üblichen Funktionsnotation. Der ggT von zwei beliebigen natürlichen Zahlen ist stets wiederum eine natürliche Zahl und beträgt mindestens 1.
- Die Multiplikation “.” ist eine abgeschlossene und kommutative Operation auf  $\mathbb{N}$ ,  $\mathbb{Z}$ ,  $\mathbb{Q}$  und  $\mathbb{R}$  und auf  $\mathbb{Z}_m$ .
- Dagegen ist die Division (als Gegenstück zur Multiplikation) nicht abgeschlossen auf  $\mathbb{N}$  oder  $\mathbb{Z}$ . Entfernt man jedoch die Null, so ist “/” eine abgeschlossene Operation auf  $\mathbb{Q} \setminus \{0\}$  und  $\mathbb{R} \setminus \{0\}$ .

Die Division ist nicht kommutativ.

- Die Division modulo  $m$  durch  $x$  ist nur möglich, wenn  $x$  ein multiplikatives Inverses besitzt. Die Mengen, für die dies möglich ist, sind gerade die primen Restklassensysteme (siehe Definition 3.15). Wegen Satz 2.23 sind die primen Restklassensysteme auch abgeschlossen unter der Multiplikation: Falls  $a, b$  beide teilerfremd mit dem Modul  $m$  sind, ist auch ihr Produkt teilerfremd mit  $m$  und also ein Element von  $\mathbb{Z}_m^*$ .

## 5.2 Gruppenartige Strukturen

Wir bauen nun eine Hierarchie von algebraischen Strukturen auf, die jeweils auf einer zusätzlichen Eigenschaft beruhen. Betrachtet werden hier Mengen sowie jeweils eine Verknüpfung. Am oberen Ende der Hierarchie steht die *Gruppe*. Hier liegen vier günstige Eigenschaften vor, die es unter anderem erlauben, die Verknüpfung rückgängig zu machen – in den darunter liegenden Strukturen ist dies letzteres nicht gefordert.

### 5.2.1 Abgeschlossenheit $\rightsquigarrow$ Magma

**Definition 5.2** (Magma). *Für eine Operation  $*$  und eine Menge  $M$  heißt die Struktur*

$$(M, *)$$

*Magma, falls  $*$  auf  $M$  abgeschlossen ist.*

*Falls  $*$  außerdem kommutativ ist, ist  $(M, *)$  ein kommutatives Magma.*



**Bemerkung:** Die Eigenschaft der Kommutativität ist eine zusätzliche Spezialisierung; sie überträgt sich, da sie an die Operation  $*$  geknüpft ist, auch auf höhere Stufen der Hierarchie (s.u.).

**Beispiele:** Nach den oben betrachteten Beispielen zur Abgeschlossenheit gilt:

- $(\mathbb{N}, +)$ ,  $(\mathbb{Z}, +)$ ,  $(\mathbb{Q}, +)$ ,  $(\mathbb{R}, +)$  und  $(\mathbb{Z}_m, +)$  sind kommutative Magmata (oder: “Magmen”).
- Analog für die Multiplikation. Man beachte, dass “+” und “.” jeweils für die passenden Operation stehen. Das Multiplizieren oder Addieren von Brüchen in  $\mathbb{Q}$  benötigt “intern” die Multiplikation (bzw. die Multiplikation und Addition) von ganzen und natürlichen Zahlen.
- $(\mathbb{N}, \text{ggT})$  ist ein kommutatives Magma.

## 5.2.2 Assoziativität $\rightsquigarrow$ Halbgruppe

**Definition 5.3** (Halbgruppe). *Ein Magma  $(M, *)$  heißt Halbgruppe, falls die Operation  $*$  auf  $M$  assoziativ ist.*

**Bemerkungen:**

- Hier befinden wir uns hierarchisch oberhalb der Magmata: Jede Halbgruppe ist immer auch ein Magma; die Umkehrung gilt jedoch nicht immer.
- Für alle Halbgruppen ist die Operation  $*$  also jeweils abgeschlossen.
- Eine kommutative Halbgruppe liegt vor, falls das Magma  $(M, *)$  ein kommutatives Magma ist.
- Der Name “Halbgruppe” kommt daher, dass zwei der vier Eigenschaften die eine Gruppe ausmachen (s.u.) erfüllt sind.

**Beispiele:**

- Addition und Multiplikation sind assoziativ für natürliche, ganze, rationale und reelle Zahlen. Mit den entsprechenden Operationen bilden diese Zahlenmengen also kommutative Halbgruppen.
- Subtraktion und Division sind auf diesen Mengen aber *nicht* assoziativ. Z.B. ist
  - $1 - (1 - 1) = 1 - 0 = 1$ , aber  $(1 - 1) - 1 = 0 - 1 = (-1)$
  - $5/(5/5) = 5/1 = 5$ , aber  $(5/5)/5 = 1/5$
- Die Verkettung (Hintereinanderausführung) von Funktionen ist assoziativ (bereits gezeigt in Satz 4.5). Eine Menge von Funktionen, die bezüglich der Komposition abgeschlossen ist, bildet also mit der Komposition “ $\circ$ ” nicht nur ein Magma, sondern eine Halbgruppe.
- Nach Satz 2.17 ist die Funktion ggT assoziativ, d.h.  $(\mathbb{N}, \text{ggT})$  ist eine kommutative Halbgruppe.

## 5.2.3 Neutrales Element $\rightsquigarrow$ Monoid

**Definition 5.4** (Monoid). *Falls eine Menge  $(M, *)$  eine Halbgruppe ist und  $M$  außerdem ein Element  $e$  enthält, sodass für alle  $x \in M$  gilt, dass*

$$x * e = e * x = x,$$

*so nennt man  $e$  das neutrale Element von  $*$ . Die Struktur  $(M, *)$  heißt dann Monoid mit neutralem Element  $e$ .*

### Bemerkungen:

- Es gibt Fälle, für die nur eine der Bestimmungsgleichungen des neutralen Elements erfüllt sind; dann spricht man von rechts- bzw. linksneutralen Elementen. Bei Monoiden muss jedoch beides gelten.
- Manchmal wird die Struktur auch als  $(M, *, e)$  notiert. Dies wäre eigentlich die korrektere Notation; in der Regel ist das neutrale Element aber aus dem Kontext klar – oder es reicht aus, es einmal zu erwähnen.
- Bei Additionen heißt das neutrale Element oft “0”, bei Multiplikationen oft “1”. Es gibt allerdings Fälle, wo gesonderte Symbole vereinbart werden (z.B. bei der Multiplikation von Matrizen, die wir im Vorlesungsteil zur linearen Algebra behandeln; s.u.).

### Beispiele:

- Da die Zahl 1 Teil der oben bereits betrachteten Zahlenmengen  $\mathbb{N}$ ,  $\mathbb{Z}$ ,  $\mathbb{Q}$ ,  $\mathbb{R}$  ist, sind diese Mengen mit dem neutralen Element 1 multiplikative Monoide. Das gleiche gilt für  $\mathbb{Z}_m$  und  $\mathbb{Z}_m^*$  mit  $m \geq 2$ ; hierbei wird 1 synonym zur Restklasse [1] verwendet. Die erwähnten Monoide sind kommutativ.
- Wo die Zahl 0 Teil der Zahlenmenge ist (bzw. die Restklasse [0] Teil des Restsystems), bilden die Zahlenmengen auch additive Monoide – hier  $\mathbb{Z}$ ,  $\mathbb{Q}$ ,  $\mathbb{R}$  und  $\mathbb{Z}_m$ . Auch diese erwähnten Monoide sind kommutativ.
- Die Verkettung von Funktionen mit der identischen Funktion  $\text{id}$  ist neutral.
- Nimmt man zu den natürlichen Zahlen noch die Null hinzu, so ist  $(\mathbb{N}_0, \text{ggT}, 0)$  ein kommutatives Monoid, da  $\text{ggT}(0, n) = n$  nach Satz 2.12.
- Für eine Menge  $M$  sind kommutative Monoide (hier der besseren Unterscheidbarkeit wegen ausführlich mit neutralem Element notiert):

$$(\mathcal{P}(M), \cap, M) \quad \text{und} \quad (\mathcal{P}(M), \cup, \emptyset)$$

---

Für Monoide gilt noch eine wichtige Eigenschaft:

**Satz 5.5** (Neutrales Element eines Monoids). *In einem Monoid  $(M, *)$  ist das neutrale Element eindeutig bestimmt.*

(Beweis: S. 320.)

**Bemerkung:** Daher ist es auch sinnvoll, bei Monoiden von *dem* neutralen Element zu sprechen anstatt von *einem* neutralen Element.

---

**Beispiel:** Zum Abschluss noch eine größere Aufgabe, die die oben besprochenen Eigenschaften beleuchtet. Wir betrachten eine binäre Operation “\*” auf den reellen Zahlen per

$$(x, y) \mapsto (x + y) + (x \cdot y),$$

wobei rechts die gewöhnlichen Addition “+” und Multiplikation “·” für reelle Zahlen verwendet wird. Nun soll  $(\mathbb{R}, *)$  hinsichtlich der Hierarchie eingeordnet werden.

- Offensichtlich ist  $x * y$  eine reelle Zahl, da Addition und Multiplikation reeller Zahlen abgeschlossene Operationen sind;  $(\mathbb{R}, *)$  ist also mindestens ein Magma.
- Für die Kommutativität betrachten wir:

$$y * x = (y + x) + (y \cdot x) = (x + y) + (x \cdot y) = x * y$$

Da die Addition und Multiplikation auf  $\mathbb{R}$  kommutativ sind, ist die obige Gleichung richtig, und damit auch “\*” kommutativ. Das Magma ist demnach ein kommutatives Magma.

- Für die Assoziativität:

$$\begin{aligned} x * (y * z) &= x * ((y + z) + (y \cdot z)) = (x + (y + z) + (y \cdot z)) + (x \cdot (y + z) + x \cdot (y \cdot z)) \\ &= (x + y + z) + (x \cdot y + x \cdot z + y \cdot z) + x \cdot y \cdot z, \text{ aber auch} \\ (x * y) * z &= ((x + y) + (x \cdot y)) * z = ((x + y) + (x \cdot y) + z) + ((x + y) \cdot z + (x \cdot y) \cdot z) \\ &= (x + y + z) + (x \cdot y + x \cdot z + y \cdot z) + x \cdot y \cdot z \end{aligned}$$

Also ist “ $*$ ” assoziativ, und  $(\mathbb{R}, *)$  ist eine kommutative Halbgruppe.

- Für die Suche nach einem neutralen Element reicht es wegen der Kommutativität,  $e$  so zu suchen, dass  $x * e = x$ , die Linksneutralität muss nicht zusätzlich überprüft werden. Es lohnt sich zunächst, die neutralen Elemente von Addition und Multiplikation einmal zu probieren:

$$(x, 1) \mapsto (x + 1) + (x \cdot 1) = 2x + 1 \quad \text{und} \quad (x, 0) \mapsto (x + 0) + (x \cdot 0) = x$$

Also ist für alle  $x \in \mathbb{R}$ :  $x * 0 = 0 * x = x$ , und damit ist  $(\mathbb{R}, *, 0)$  ein Monoid.

### 5.2.4 Inverse Elemente $\rightsquigarrow$ Gruppe

**Definition 5.6.** Ein Monoid  $(M, *)$  mit neutralem Element  $e$  heißt Gruppe, falls zu jedem  $x \in M$  ein inverses Element  $x^{-1} \in M$  existiert, sodass

$$x * x^{-1} = x^{-1} * x = e$$

Ist das Monoid kommutativ, so heißt die Gruppe abelsche<sup>1</sup> Gruppe.

#### Bemerkungen:

- Auch hier kann das neutrale Element per  $(M, *, e)$  mit notiert werden.
- Kommutative Strukturen unterhalb der Gruppe werden jedoch nicht “abelsch” genannt. Bei Gruppen sind die Begriffe “kommutative Gruppe” und “abelsche Gruppe” synonym.
- Die Verknüpfung mit dem inversen Element führt auf das (eindeutig bestimmte) neutrale Element, macht also für beliebiges  $x \in M$  die Verknüpfung mit  $x$  rückgängig. Denn von einem beliebigen  $y \in M$  gelangt man durch Verknüpfung mit  $x$  zu  $y * x$ . Verknüpfen wir dies nun (von rechts) mit  $x^{-1}$ , so ergibt sich (die Assoziativität der Verknüpfung ist gegeben, da  $(M, *)$  auch Halbgruppe ist):

$$(y * x) * x^{-1} = y * (x * x^{-1}) = y * e = y$$

- Subtraktionen lassen sich in additiven Gruppen durch Additionen mit dem inversen Element erklären:

$$x - y = x + (-y)$$

Also  $y^{-1} = (-y)$

- Divisionen lassen sich multiplikativen Gruppen durch Multiplikation mit dem inversen Element erklären:

$$x/y = \frac{x}{y} = x \cdot \frac{1}{y}$$

Also  $y^{-1} = \frac{1}{y}$

Man beachte aber, dass die beiden Schreibweisen ohne Multiplikationspunkt nur eindeutig sind, wenn die Multiplikation kommutativ ist! In nichtkommutativen Gruppen müssen linke und rechte Operanden immer klar als solche erkennbar notiert sein.

- Da aber die Addition bzw. Multiplikation “häufiger” erklärbar ist als ihre inversen Operationen (die erst bei Gruppen gefordert sind), werden Addition und Multiplikation als grundsätzlichere Operationen verstanden. Subtraktion und Division sind hingegen bei Vorliegen einer Gruppeneigenschaft auf Addition bzw. Multiplikation zurückführbar (s.o.).

---

<sup>1</sup>N. H. Abel, norwegischer Mathematiker

### Beispiele:

- $(\mathbb{Z}, +)$  ist eine abelsche Gruppe mit neutralem Element 0. Analog für die Mengen  $\mathbb{Q}$  und  $\mathbb{R}$ , die auch jeweils alle negativen Zahlen enthalten.
- $(\mathbb{N}_0, +, 0)$  ist *keine* Gruppe, da hier die Addition nicht invertierbar ist.
- Die rationalen und reellen Zahlen sind zwar keine multiplikativen Gruppen (also Gruppen mit der Verknüpfung “ $\cdot$ ”) mit neutralem Element 1, aber die jeweiligen Mengen ohne die Null sind es.
- $(\mathbb{N}_0, \text{ggT}, 0)$  ist keine Gruppe, denn der ggT von einer Zahl  $x \in \mathbb{N}$  mit irgendeinem  $y \in \mathbb{N}_0$  beträgt stets mindestens 1, kann also nicht 0 werden. Also hat  $x$  kein inverses Element bzgl. der ggT-Operation.
- Die Menge der bijektiven reellen Funktionen mit der Verkettung  $\circ$  ist eine Gruppe mit neutralem Element  $\text{id}$ . Beweis: Übung.
- Die Restklassensysteme  $(\mathbb{Z}_m, +)$  sind abelsche Gruppen für die Addition modulo  $m$ . Die neutralen Elemente heißen je nach Kontext 0 oder, strenger,  $[0]$ .
- Die primen Restklassensysteme  $(\mathbb{Z}_m^*, \cdot)$  für  $m \geq 2$  sind abelsche Gruppen (mit neutralem Element 1 (bzw.  $[1]$ ) für die Multiplikation modulo  $m$ , da sie genau die multiplikativ invertierbaren Elemente aus  $\mathbb{Z}_m$  enthalten. Die Abgeschlossenheit hatten wir oben bereits diskutiert. Daher erklärt sich auch die Bezeichnung als “multiplikative Gruppe von  $\mathbb{Z}_m$ ”.

---

Für die inversen Elemente gilt noch eine wichtige Eigenschaft:

**Satz 5.7** (Eindeutigkeit der inversen Elemente). *In einer Gruppe  $(M, *, e)$  ist für jedes  $x \in M$  das Inverse Element  $x^{-1}$  eindeutig bestimmt.*

(Beweis: S. 320.)

**Bemerkung:** Daher rechtfertigt sich auch die Schreibweise  $x^{-1}$  für *das* inverse Element zu  $x$ .

---

Wir geben noch zwei wichtigen Eigenschaften von Gruppen:

**Satz 5.8** (Inverse einer Verknüpfung). *Sei  $(M, *)$  eine Gruppe mit neutralem Element  $e$ , und  $a, b \in M$ . Dann gilt:*

$$(a * b)^{-1} = b^{-1} * a^{-1}$$

(Beweis: S. 320.)

**Satz 5.9** (Kürzungsregel bei Gruppen). *Sei  $(M, *)$  eine Gruppe mit neutralem Element  $e$ , und  $a, b, x \in M$ . Dann gilt:*

$$(a * x = b * x) \Rightarrow (a = b) \quad \text{und} \quad (x * a = x * b) \Rightarrow (a = b)$$

(Beweis: S. 320.)

---

**Beispiel:** Wir hatten oben (S. 114) eine Verknüpfung “ $*$ ” auf  $\mathbb{R}$  kennen untersucht, wobei  $(\mathbb{R}, *)$  ein kommutatives Monoid mit neutralem Element 0 bildet. Es war:  $x * y = (x + y) + (x \cdot y)$ . Nun wäre noch interessant, ob sich nicht sogar inverse Elemente finden lassen, sodass wir eine Gruppe erhalten. Wegen der Kommutativität reicht es, zu beliebigem  $x \in \mathbb{R}$  ein  $y \in \mathbb{R}$  zu finden, sodass  $x * y = 0$ .

Leider wird in der Verknüpfung “ $*$ ” die Addition und Multiplikation auf  $\mathbb{R}$  kombiniert. Wir können aber trotzdem versuchen, die Gleichung  $x * y = 0$  nach  $y$  aufzulösen; falls das in eindeutiger Weise gelingt, hätten wir unser Inverses  $y = x^{-1}$  gefunden.

$$\begin{aligned}
x * y &= 0 \\
\Leftrightarrow x + y + x \cdot y &= 0 \\
\Leftrightarrow x + y \cdot (1 + x) &= 0 \\
\Leftrightarrow y \cdot (1 + x) &= -x
\end{aligned}$$

Für den Fall  $x \neq (-1)$  lässt sich tatsächlich weiter rechnen; dann dürfen wir die Gleichung durch  $(1 + x)$  dividieren und erhalten:

$$\dots \Leftrightarrow y = -\frac{x}{1+x}$$

Dies ist zunächst eine eindeutige Zuordnung für jedes  $x$ , was für die Gruppeneigenschaft auch so erforderlich wäre. Bevor wir dies weiter vertiefen, rechnen wir noch zur Probe nach:

$$\begin{aligned}
x * \left(-\frac{x}{1+x}\right) &= x - \frac{x}{1+x} + x \cdot \left(-\frac{x}{1+x}\right) \\
&= \frac{1}{1+x} \cdot (x \cdot (1+x) - x - x^2) \\
&= \frac{1}{1+x} \cdot (x + x^2 - x - x^2) \\
&= 0 \quad \checkmark
\end{aligned}$$

Alle Elemente in  $\mathbb{R} \setminus \{-1\}$  besitzen also ein eindeutig bestimmtes Inverses bezüglich der Operation “ $*$ ”. Nun wäre für die Gruppeneigenschaft allerdings nochmals die Abgeschlossenheit zu untersuchen. Die Assoziativität gilt weiterhin, und auch das neutrale Element 0 befindet sich noch in der neuen Menge. Ist aber jede Verknüpfung von Elementen aus  $\mathbb{R} \setminus \{-1\}$  wieder ein Element von  $\mathbb{R} \setminus \{-1\}$ ?

Zunächst betrachten wir die Inversen von  $x$ . Wegen  $x \neq (-1)$  dürfen wir äquivalent umformen:

$$\begin{aligned}
-\frac{x}{1+x} &= -1 \\
\Leftrightarrow \frac{x}{1+x} &= 1 \\
\Leftrightarrow x &= 1+x \\
\Leftrightarrow 0 &= 1 \quad \text{!}
\end{aligned}$$

Wegen der Äquivalenzumformungen ist somit gesichert, dass keines der Inversen von  $x \in \mathbb{R} \setminus \{-1\}$  den Wert  $(-1)$  hat.

Nun untersuchen wir allgemein, wann die Verknüpfung  $x * y$  den Wert  $(-1)$  annimmt:

$$\begin{aligned}
x * y &= -1 \\
\Leftrightarrow x + y + x \cdot y &= -1 \\
\Leftrightarrow x + y \cdot (1 + x) &= -1 \\
\Leftrightarrow y \cdot (1 + x) &= -1 - x = -(1 + x) \\
\stackrel{x \neq -1}{\Leftrightarrow} y &= -1
\end{aligned}$$

In der letzten Umformung haben wir beide Seiten der Gleichung durch  $(1 + x)$  dividiert, was (siehe Kommentar am Äquivalenzpfeil) hier zulässig war.

Für  $x \neq (-1)$  nimmt also die Verknüpfung  $x * y$  genau dann den Wert  $(-1)$  an, wenn  $y$  den Wert  $(-1)$  hat. Dies ist jedoch für  $y$  aus  $\mathbb{R} \setminus \{-1\}$  ausgeschlossen. Somit ist die Operation  $*$  in der Tat auf  $\mathbb{R} \setminus \{-1\}$  abgeschlossen, und damit ist

$$(\mathbb{R} \setminus \{-1\}, *)$$

eine abelsche (denn die Kommutativität von  $*$  gilt natürlich auch weiterhin) Gruppe mit neutralem Element 0.

## 5.2.5 Untergruppen

Jede Gruppe ist auch ein Monoid. Falls wir das neutrale Element  $e \in M$  eines Monoids  $(M, *)$  betrachten, so ist die Struktur

$$(\{e\}, *)$$

stets eine Gruppe (und sogar eine abelsche!). Denn wegen  $e * e = e$  ist  $e$  selbstinvers – das gilt sogar dann, wenn für andere Elemente aus  $M$  vielleicht gar keine Inversen existieren.

Dies ist ein trivialer Spezialfall folgender

**Definition 5.10** (Untergruppe). *Die Struktur  $(U, *, e)$  heißt Untergruppe der Gruppe  $(M, *, e)$ , falls  $U \subseteq M$  und falls  $(U, *, e)$  eine Gruppe ist.*

### Bemerkungen:

- Im Fall  $(U, *, e)$  wird eine auf  $U$  eingeschränkte Variante der Operation  $*$  verwendet.
- Das neutrale Element  $e$  ist für  $U$  und  $M$  das gleiche, denn jedes Element aus  $U$  ist auch Element aus  $M$ , und in  $M$  ist das neutrale Element eindeutig bestimmt. Da aber auch die Menge  $U$  mit  $*$  eine Gruppe mit eindeutig bestimmtem neutralem Element bildet, muss es sich hierbei um das gleiche  $e$  handeln, und damit muss  $e \in U$  gelten.

### Beispiele:

- Für jede Gruppe  $(M, *)$  mit neutralem Element  $e$  ist (s.o.)  $(\{e\}, *)$  eine *triviale Untergruppe*. (Noch trivialer ist die Gruppe  $(M, *)$  auch Untergruppe von sich selbst, da die Teilmengenbeziehung Gleichheit nicht ausschließt).
- Für  $k \in \mathbb{N}$  ist die Menge  $k\mathbb{Z} = \{k \cdot x \mid x \in \mathbb{Z}\}$  eine Untergruppe von  $\mathbb{Z}$  bezüglich der Addition. Für z.B.  $k := 7$  wäre nämlich:

$$7\mathbb{Z} = \{7x \mid x \in \mathbb{Z}\} \subseteq \mathbb{Z}$$

Nun könnten wir hier wieder die vier Gruppeneigenschaften zeigen. Die Abgeschlossenheit ist erfüllt, da die Summe zweier Vielfachen von 7 wieder ein Vielfaches von 7 ist (Faktor 7 ausklammern nach Distributivgesetz). Die Null ist neutrales Element der Addition, und zu jedem  $7x$  ist  $(-7x)$  jeweils das Inverse, welches ebenfalls in  $7\mathbb{Z}$  liegt.

Oder wir erkennen, dass die Menge  $7\mathbb{Z}$  nichts anderes ist als die Restklasse  $[0]$  modulo 7. Nun hatten wir oben schon angemerkt, dass die Restklassensysteme modulo  $m$  additive abelsche Gruppen bilden. Bei  $(\{[0]\}, +, [0])$  handelt es sich um die triviale abelsche Untergruppe von (hier)  $(\mathbb{Z}_7, +, [0])$ .

- Die geraden ganzen Zahlen  $2\mathbb{Z}$  bilden (siehe voriges Beispiel) eine additive Untergruppe von  $\mathbb{Z}$ . Die ungeraden Zahlen  $\mathbb{Z} \setminus 2\mathbb{Z}$  jedoch nicht (denn die Summe aus zwei solchen Zahlen wäre gerade, d.h. die Addition wäre dort nicht abgeschlossen).

## 5.2.6 Symmetrische Gruppe $S_n$

Wir hatten oben schon angemerkt, dass die bijektiven Funktionen auf einer Menge mit der Komposition  $\circ$  und dem neutralen Element  $\text{id}$  eine Gruppe bilden. Das gilt z.B. für reelle bijektive Funktionen, aber auch für diskrete Mengen; insbesondere:

**Definition 5.11** (Symmetrische Gruppe). Für  $n \in \mathbb{N}$  sei  $M_n := \{1, 2, \dots, n\}$ , und  $S_n$  die Menge sämtlicher Permutationen auf  $M_n$ . Die Struktur

$$(S_n, \circ, \text{id}_n)$$

heißt symmetrische Gruppe

Um den Namen zu rechtfertigen, haben wir die Gruppeneigenschaften zu zeigen:

- Da  $S_n$  sämtliche Permutationen von  $M_n$  enthält, ist die Verknüpfung abgeschlossen. Jede Komposition von Permutationen auf  $M_n$  ergibt erneut eine Permutation auf  $M_n$ .
- Die Komposition von Permutationen auf  $M_n$  ist, wie jede Komposition von Funktionen, assoziativ.
- Das neutrale Element ist  $\text{id}_n \in S_n$ , mit  $\text{id}_n(j) = j$
- Jede Permutation auf  $M_n$  besitzt ein Inverses, das ebenfalls eine Permutation auf  $M_n$  ist, und  $\sigma^{-1} \circ \sigma = \sigma \circ \sigma^{-1} = \text{id}_n$ . Siehe dazu Abschnitt 4.3.

Wir fassen dies zusammen im

**Satz 5.12** (Symmetrische Gruppe). Die Struktur  $(S_n, \circ, \text{id}_n)$  aus Definition 5.11 ist eine Gruppe.

### Bemerkungen:

- Es gibt  $n!$  verschiedene Elemente in  $S_n$ ; siehe Satz 1.40.
- Ein Beispiel für eine Untergruppe der  $S_5$  findet sich als Exkurs im Anhang A.1.
- Bis auf  $S_1$  und  $S_2$  sind die Gruppen<sup>2</sup>  $S_n$  nicht abelsch. Wir hatten in Abschnitt 4.3 schon ein Beispiel für zwei nicht kommutierende Permutationen auf  $M_3$  gegeben. Diese lassen sich problemlos in  $S_n$  mit  $n \geq 3$  einbetten (alle anderen Elemente von  $M_n$  auf sich selbst abbilden). Daher gibt es für  $n \geq 3$  stets Permutationen aus  $S_n$ , die nicht kommutieren, sodass  $S_n$  nicht abelsch sein kann.

### Beispiele:

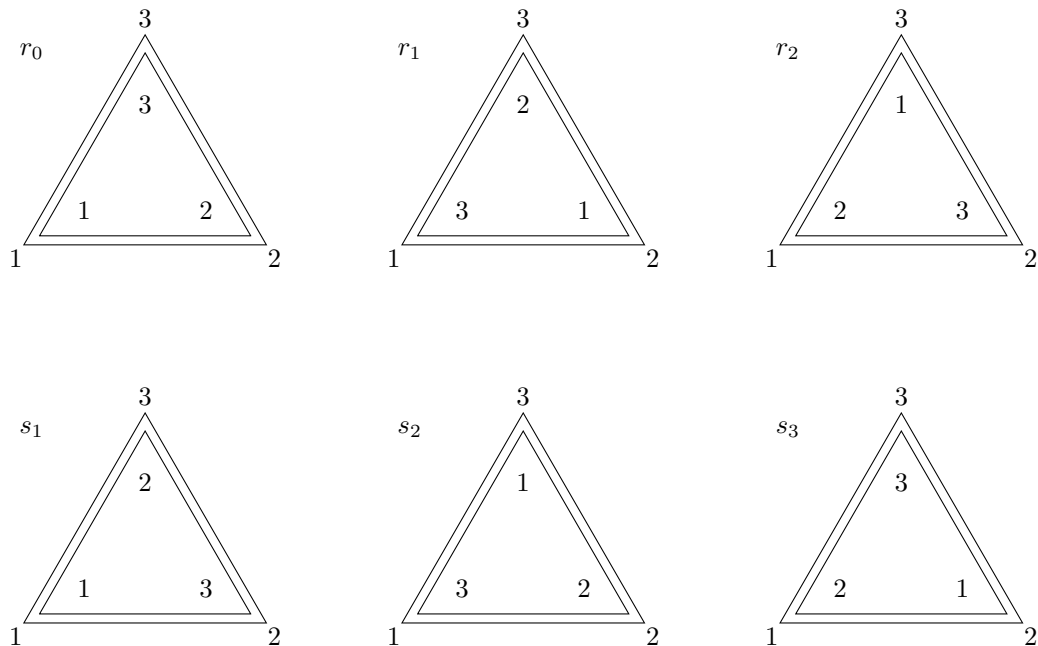
- Die Symmetrische Gruppe  $S_3$  besitzt sechs Elemente, nämlich die Permutationen

$$r_0 := \text{id}_3 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \quad r_1 := \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}, \quad r_2 := \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix},$$

$$s_1 := \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \quad s_2 := \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}, \quad s_3 := \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$$

Dabei steht “r” für Rotation; es handelt sich hierbei um Drehungen von 0, 1 oder 2 mal  $\frac{2\pi}{3}$  – das sind die Drehungen, die ein gleichseitiges Dreieck mit nummerierten Ecken auf sich selbst abbilden (mit ggf. rotierter Eckenfolge). Die anderen drei Permutationen (“s”) sind Spiegelungen des Dreiecks an den Symmetrieachsen; der Index gibt dabei die Ecke an, die fest bleibt; die anderen beiden vertauschen sich beim Spiegeln jeweils.

Wir zeigen die Permutationen grafisch: Das äußere Dreieck bleibt fest, während die Ecken des inneren Dreiecks entsprechend der beschriebenen Operationen angeordnet werden:



Das führt auf folgende Gruppentafel für die Verknüpfung  $\circ$  (“Zeile kringel Spalte”):

$\circ$	$r_0$	$r_1$	$r_2$	$s_1$	$s_2$	$s_3$
$r_0$	$r_0$	$r_1$	$r_2$	$s_1$	$s_2$	$s_3$
$r_1$	$r_1$	$r_2$	$r_0$	$s_2$	$s_3$	$s_1$
$r_2$	$r_2$	$r_0$	$r_1$	$s_3$	$s_1$	$s_2$
$s_1$	$s_1$	$s_3$	$s_2$	$r_0$	$r_2$	$r_1$
$s_2$	$s_2$	$s_1$	$s_3$	$r_1$	$r_0$	$r_2$
$s_3$	$s_3$	$s_2$	$s_1$	$r_2$	$r_1$	$r_0$

<sup>2</sup>Falls die Operation und das neutrale Element aus dem Kontext hervor gehen, werden die Mengen und die Gruppen gerne synonym verwendet

Wir beobachten, dass jede der Permutationen in jeder Zeile und jeder Spalte je einmal auftritt. Außerdem lässt sich zu jeder Permutation das Inverse nachschlagen, indem man die Zeile (oder Spalte) sucht, für die zur gegebenen Spalte (Zeile) der Eintrag  $r_0$  entspricht (das ist die identische Permutation).

Die Gruppe ist weiterhin *nicht* abelsch, da z.B.  $r_1 \circ s_1 = s_2$ , aber  $s_1 \circ r_1 = s_3$ .

- Wir finden einige Untergruppen:

- Zunächst die triviale Untergruppe  $\{r_0\}$
- Dann zweielementige Untergruppen bestehend aus der Identität und jeweils einer der Spiegelungen (diese sind nämlich zu sich selbst invers, d.h.  $s_j \circ s_j = r_0 = \text{id}_3$ ):

$$\{r_0, s_1\}, \quad \{r_0, s_2\}, \quad \{r_0, s_3\}$$

Zum Beispiel lautet die Gruppentafel der letzten notierten Untergruppe (die Elemente lassen sich aus der obigen großen Gruppentafel entnehmen):

$\circ$	$r_0$	$s_3$
$r_0$	$r_0$	$s_3$
$s_3$	$s_3$	$r_0$

Diese drei Untergruppen sind abelsch.

- Die drei Rotationen  $\{r_0, r_1, r_2\}$  bilden ebenfalls eine Untergruppe; diese ist abelsch. Ihre Gruppentafel kann man oben in der großen Tafel ablesen, wenn man nur das linke obere Viertel betrachtet.

Wir sehen, dass (wie erwartet bzw. gefordert) das neutrale Element in jeder der angegebenen Untergruppen vorhanden (und das gleiche) ist.

- Noch eine Bemerkung zu den oben motivierten Bezeichnern: Die Gruppe  $S_3$  entspricht tatsächlich den Symmetrieoperationen des gleichseitigen Dreiecks. Für höheres  $n$  ist das bei  $S_n$  *nicht mehr* der Fall. Die Symmetrieoperationen des regulären  $n$ -Ecks bilden dann jeweils eine Untergruppe von  $S_n$ , die so genannte Dieder-Gruppe  $D_n$  (hierzu ein Beispiel für  $n = 5$  im Exkurs A.1). Wir haben hier also streng genommen eher  $D_3$  als  $S_3$  betrachtet, was aber in diesem Fall auf dasselbe hinaus läuft<sup>3</sup>.

Es gibt aber dann in  $S_n$  darüber hinaus noch weitere Permutationen, die sich nicht (direkt) in obiger Weise benennen lassen; dies führt aber über den Rahmen der Vorlesung hinaus.

## 5.3 Ringartige Strukturen

Bisher haben wir Mengen mit je einer Verknüpfung kombiniert; das ergab die gruppenartigen Strukturen. Wir kennen aber schon von der Arithmetik der natürlichen Zahlen das *Distributivgesetz*, das Addition und Multiplikation verbindet. Diese Zusammenhänge lassen sich über Gruppen nicht ausdrücken. Wir führen hier deswegen die *ringartigen*<sup>4</sup> Strukturen ein, die Mengen mit zwei Verknüpfungen kombinieren. Dies geschieht mit dem Ziel, die Arithmetik auf natürlichen, ganzen und rationalen Zahlen formal zusammen zu fassen. Auch hier ergibt sich dann eine Hierarchie, auf deren oberster Ebene sich die *Körper* befinden. In Körpern sind die vier Grundrechenarten möglich (solange nicht durch null dividiert wird) – inklusive allen höheren Operationen, die diesen basieren; dazu an geeigneter Stelle unten noch weitere Bemerkungen.

Bei allen ringartigen Strukturen setzen wir in Gedanken voraus, dass sich die beiden Operationen wie Addition bzw. Multiplikation verhalten, und benennen diese denn auch meist mit den dafür üblichen Symbolen. Falls die Operationen (und ggf. ihre neutralen Elemente) aus dem Kontext bekannt sind, werden die ringartigen (genau wie bei den gruppenartigen) Strukturen teilweise synonym mit den beteiligten Mengen notiert.

<sup>3</sup>Wegen der noch halbwegs übersichtlichen Gruppentafel ist die  $S_3$  ein beliebtes Beispiel für die symmetrische Gruppe; auch, weil es hier schon nichttriviale Untergruppen gibt. Der Unterschied zur Dieder-Gruppe für höheres  $n$  ist aber wesentlich.

<sup>4</sup>“Ring” soll hier nicht rundes oder gar schwimmreifenförmiges bedeuten, sondern ähnlich wie die “Gruppe” ein Verband von gewissen Elementen mit bestimmtem Verhalten sein. Abseits der Mathematik kennt man den Ausdruck in dieser Form noch von Wörtern wie “Schmuggler-Ring” oder “Drogenhändler-Ring”



### 5.3.1 Halbringe

Wir fassen zunächst die Arithmetik der natürlichen Zahlen  $\mathbb{N}$  (die wir *axiomatisch*, also ohne Beweis, akzeptieren) zusammen:

**Definition 5.13** (Halbring). *Für eine Menge  $M$  und die Operationen Addition  $(+)$  sowie Multiplikation  $(\cdot)$  auf  $M$  heißt die Struktur*

$$(M, +, \cdot)$$

*Halbring, falls folgende Eigenschaften gelten:*

- $(M, +)$  ist eine kommutative Halbgruppe
- $(M, \cdot)$  ist eine Halbgruppe
- Für  $x, y, z \in M$  “verteilt sich” die Multiplikation auf die Addition mit den beiden Distributivgesetzen

$$x \cdot (y + z) = (x \cdot y) + (x \cdot z)$$

$$(x + y) \cdot z = (x \cdot z) + (y \cdot z)$$

*Ein kommutativer Halbring liegt dann vor, wenn  $(M, \cdot)$  eine kommutative Halbgruppe ist.*

*Ein Halbring mit Eins (oder: unitärer Halbring) liegt dann vor, wenn die Multiplikation ein neutrales Element besitzt und  $(M, \cdot)$  damit ein Monoid bildet.*

#### Bemerkungen:

- Die Kommutativität der Addition ist für alle Halbringe voraus gesetzt. In der Bezeichnung “kommutativer Halbring” bezieht sich die explizit erwähnte Kommutativität also auf die Multiplikation.
- Die Zusatzeigenschaften “kommutativer Halbring” und “Halbring mit Eins” übertragen sich ähnlich wie bei den gruppenartigen Strukturen in der Hierarchie nach oben.
- Für kommutative Halbringe sind die beiden Distributivgesetze gleichwertig; es lässt sich also das eine aus dem anderen durch Äquivalenzumformung herleiten (Beweis: Übung).
- Genau wie bei multiplikativen Monoiden wird das neutrale Element der Multiplikation bei Halbringen mit Eins meist als “1” bezeichnet. Es gibt aber Ausnahmen, z.B. bei der Matrizenmultiplikation (siehe den Vorlesungsteil zur linearen Algebra)!

**Beispiel:** Die natürlichen Zahlen  $\mathbb{N}$  bilden mit den gebräuchlichen Operationen einen kommutativen Halbring mit Eins:

$$(\mathbb{N}, +, \cdot)$$

Es sind also Addition und Multiplikation kommutativ. Während die Addition kein neutrales Element in  $\mathbb{N}$  besitzt (die 0 gehört nicht zu  $\mathbb{N}$ ), ist dies für die Multiplikation in Gestalt der 1 (der wichtigsten natürlichen Zahl!) gegeben. Auch die Distributivgesetze (bzw. das eine von beiden – sie sind ja hier gleichwertig) sind verträglich mit der gewohnten Arithmetik natürlicher Zahlen.

### 5.3.2 Ringe

Die Arithmetik natürlicher Zahlen lässt sich auch in den ganzen Zahlen  $\mathbb{Z}$  einbetten und mit gleichem Effekt verwenden. Weiterhin gibt es dort aber noch die negativen Zahlen und die Null, sodass die Addition eine Gruppeneigenschaft erhält. Das führt uns zu einer neuen Struktur:

**Definition 5.14** (Ring). *Ein Halbring  $(M, +, \cdot)$  heißt Ring, falls  $(M, +)$  eine abelsche Gruppe ist.*

#### Bemerkungen:

- Zusätzlich zu Halbringen ist hier die Addition in beiden Richtungen uneingeschränkt möglich, also insbesondere auch die Subtraktion.
- Das neutrale Element der Addition wird meist als “0” bezeichnet.
- Kommutative Ringe oder Ringe mit Eins sind genauso definiert wie beim Halbring.

### Beispiele:

- Die ganzen Zahlen  $\mathbb{Z}$  bilden mit den gebräuchlichen Operationen einen kommutativen Ring mit Eins:

$$(\mathbb{Z}, +, \cdot)$$

(Die Arithmetik ganzer Zahlen müssen wir nicht als gegeben hinnehmen.  $\mathbb{Z}$  lässt sich über eine Äquivalenzrelation aus  $\mathbb{N}$  erzeugen, und man kann unter Rückgriff auf die axiomatische Arithmetik von  $\mathbb{N}$  zeigen, dass mit den dafür zu definierenden Operationen für die Addition und Multiplikation in der Tat ein Ring gegeben ist.)

- Die Restklassensysteme  $\mathbb{Z}_m$  mit  $m \in \mathbb{N}$  bilden kommutative Ringe mit Eins. Die neutralen Elemente sind hier jedoch (bei strenger Notation) Restklassen. Falls die neutralen Elemente zu den Operationen mit notiert werden, lauten die Strukturen folgendermaßen:

$$(\mathbb{Z}_m, +, [0], \cdot, [1])$$

Diese (endlichen) Ringe heißen entsprechend auch *Restklassenringe*.

- Es gibt durchaus auch Ringe ohne die beiden Zusatzeigenschaften. Wir werden (s.u.) zum Beispiel finden, dass bestimmte Matrizen einen Ring mit Eins bilden, der jedoch (in der Multiplikation!) nicht kommutativ ist. Ein Beispiel für kommutative Ringe, die keine Ringe mit Eins sind, folgt als nächstes.

---

**Beispiel:** Zum Abschluss des Unterabschnitts über Ringe betrachten wir als größeres Beispiel noch eine Familie von endlichen Ringen. Für fest gewählte  $k, n \in \mathbb{N}$  bildet nämlich die Menge

$$R_{k,n} := \{0, k, 2k, 3k, \dots, nk\} = \bigcup_{j=0}^n \{j \cdot k\}$$

einen Ring, wobei Addition und Multiplikation jeweils modulo  $(nk + k)$  auszuführen sind. Zunächst einige konkrete Ausprägungen:

- $R_{2,5} = \{0, 2, 4, 6, 8, 10\}$ , Rechenoperationen modulo  $12 = 2 \cdot (5 + 1)$
- $R_{3,3} = \{0, 3, 6, 9\}$ , Rechenoperationen modulo  $12 = 3 \cdot (3 + 1)$
- $R_{5,3} = \{0, 5, 10, 15\}$ , Rechenoperationen modulo  $20 = 5 \cdot (3 + 1)$
- $R_{1,6} = \{0, 1, 2, 3, 4, 5, 6\}$ , Rechenoperationen modulo  $7 = 1 \cdot (6 + 1)$

Wenn man die Elemente der Größe nach ordnet, ist  $k$  die Differenz aufeinander folgender Elemente, und  $(n + 1)$  ist die Anzahl der Elemente. Es handelt sich jeweils um endliche Teilmengen der Restklassen  $[0]$  modulo  $k$ .

Im Fall  $k = 1$  beschreibt die Menge isomorph den Restklassenring  $\mathbb{Z}_{n+1}$ , von dessen Ringeigenschaft wir uns (mit Vorarbeit aus Kapitel 3) oben schon überzeugt hatten. Wir betrachten also ab jetzt den Fall  $k > 1$ . Für letzteren Fall ist 1 *kein* Element der Menge, d.h. es liegt dann ein kommutativer Ring vor, jedoch kein Ring mit Eins!

Für die oben gegebenen Beispiele zeigt man (Übung) durch Auflisten aller möglichen Ergebnisse für Addition und Multiplikation zweier Elemente, dass die Mengen abgeschlossen sind; die restlichen Kriterien der Ringeigenschaft folgen analog aus den Rechenregeln für Restklassen wie oben.

Wir untersuchen hier die Abgeschlossenheit in der allgemeinen Form. Zunächst sind alle Elemente von  $R_{k,n}$  Vielfache von  $k$ . Das gleiche gilt dann auch für alle denkbaren Summen und Produkte solcher Elemente. Das Rechenergebnis vor der Division mit Rest ist also ein Wert  $x \cdot k$  mit  $x \in \mathbb{N}_0$ .

Nun gibt es nach Satz 2.7 eindeutig bestimmte  $q, r$  mit  $0 \leq r < nk + k = k \cdot (n + 1)$ , sodass gilt:

$$x \cdot k = q \cdot (k \cdot (n + 1)) + r$$

(der Modul wurde hier in Klammern notiert). Da  $x \geq 0$  ist hier auch  $q \in \mathbb{N}_0$ , nicht wie allgemein aus den ganzen Zahlen.

Wir vereinfachen nun den Divisionsrest:

$$r = x \cdot k - q \cdot (k \cdot (n+1)) = k \cdot (x - q \cdot (n+1))$$

Aber damit ist  $r$  ein Vielfaches von  $k$ , und wegen  $0 \leq r < k \cdot (n+1)$  muss also gelten:  $r \in R_{k,n}$ . Daher sind die beiden Rechenoperationen abgeschlossen auf  $R_{k,n}$ ; und es liegt die Eigenschaft eines endlichen kommutativen Rings vor, der für  $k > 1$  kein Ring mit Eins ist.

Im Übrigen sind die  $R_{k,n}$  nicht nur Ringe, sondern *Unterringe* (Konzept analog zu Untergruppen; hier kein Vorlesungsstoff) zu den Ringen  $\mathbb{Z}_{nk+k}$ . Denn diese letzteren Ringe enthalten sämtliche  $k \cdot (n+1)$  Restklassen modulo  $k \cdot (n+1)$ . Die  $R_{k,n}$  enthalten statt dessen  $(n+1)$  Restklassen, also den Anteil  $\frac{1}{k}$  aller jeweils möglichen Restklassen.

### 5.3.3 Körper

Um die Grundrechenarten zu vervollständigen, benötigen wir noch die Division, also die Umkehrbarkeit der Multiplikation. Ähnlich wie beim Übergang von  $\mathbb{N}$  auf  $\mathbb{Z}$  können wir die Arithmetik von  $\mathbb{Z}$  verträglich in die rationalen Zahlen  $\mathbb{Q}$  einbetten, die (bis auf die Null) außerdem noch multiplikative Inverse enthalten. Das führt uns auf

**Definition 5.15** (Körper). *Ein kommutativer Ring mit Eins  $(M, +, \cdot)$  heißt Körper, falls die Struktur  $(M \setminus \{e_+\}, \cdot)$  eine abelsche Gruppe ist ( $e_+$  sei hierbei das neutrale Element der Gruppe  $(M, +)$ ).*

**Bemerkungen:**

- Da  $e_+$  in der Regel als Null bezeichnet wird, gilt in Körpern: Alle Elemente ungleich null besitzen ein multiplikatives Inverses.
- (Manchmal wird in der Literatur keine Kommutativität in der Multiplikation verlangt; dann müsste  $(M \setminus \{e_+\}, \cdot)$  lediglich eine Gruppe sein. Solche Ringe mit nicht-kommutativer Multiplikation sind anderweitig als *Schiefkörper* bekannt (kein Vorlesungsstoff). Wir halten uns jedoch an obige Definition!)
- Im Körper kann also zusätzlich zum Ring auch dividiert werden, solange man die Null auslässt. In manchen Ringen ohne Körpereigenschaft ist aber zumindest eine Division mit Rest<sup>5</sup> möglich (siehe Kapitel 2).

**Beispiele:**

- Die rationalen Zahlen  $\mathbb{Q}$  bilden mit den gebräuchlichen Operationen (Bruchrechnung) und den neutralen Elementen 0 und 1 einen Körper:

$$(\mathbb{Q}, +, \cdot)$$

(Auch die rationalen Zahlen lassen sich über eine Äquivalenzrelation aus  $\mathbb{Z}$  und  $\mathbb{N}$  konstruieren. Hier lässt sich wiederum schlüssig zeigen, dass die Arithmetik rationaler Zahlen genau den Körpereigenschaften genügt.)

- Bei den Restklassenringen  $\mathbb{Z}_m$  (hier wieder synonym für die Ringstruktur verwendet) gibt es Fälle, für die  $\mathbb{Z}_m \setminus \{[0]\}$  mit der Multiplikation tatsächlich eine abelsche Gruppe bildet. Diese entspricht dann genau der multiplikativen Gruppe  $\mathbb{Z}_m^*$ . Das sind genau die Restklassenringe mit primem Modul! Für  $p$  prim ist also (streng notiert)

$$(\mathbb{Z}_p, +, [0], \cdot, [1])$$

ein Körper, d.h. auch hier sind, unter Benutzung der modulo-Arithmetik, alle Grundrechenarten erklärt. Anders als die rationalen Zahlen enthält  $\mathbb{Z}_p$  jedoch endlich viele Elemente! Diese (endlichen) Körper sind entsprechend auch als *Restklassenkörper* bekannt.

<sup>5</sup>Solche Ringe heißen *Euklidische Ringe* (kein Vorlesungsstoff), da die Division mit Rest auch "euklidische Division" genannt wird.

- Für  $m \in \mathbb{N}$ ,  $m > 3$  und  $m$  nicht prim ist der Restklassenring  $\mathbb{Z}_m$  jedoch *kein* Körper. Denn dann unterscheiden sich die multiplikative Gruppe  $\mathbb{Z}_m^*$  und die Menge  $\mathbb{Z}_m \setminus \{[0]\}$ . Zwar ist innerhalb der multiplikativen Gruppe dann weiter uneingeschränkt Multiplikation und Division möglich – aber die Addition ist nicht länger abgeschlossen!

Denn da  $[1]$  stets zur multiplikativen Gruppe gehört und nicht gleich  $[0]$  ist, müsste diese Restklasse bei einem Körper auch zur additiven Gruppe gehören. Aber dann ließen sich durch fortgesetztes Addieren sämtliche Restklassen erzeugen, inklusive derer, die *nicht* in  $\mathbb{Z}_m^*$  liegen. Entfernt man von der additiven Gruppe dann die Restklasse  $[0]$ , so liegt hier dann *nicht* die multiplikative Gruppe vor.

- Es kann (kein Vorlesungsstoff!) gezeigt werden, dass alle endlichen Körper  $p^k$  viele Elemente besitzen, mit  $p$  prim und  $k \in \mathbb{N}$ ; bei gleichen  $p, k$  sind diese untereinander isomorph (d.h. bis auf Benennung der Elemente gleich).

Für  $k = 1$  sind diese isomorph zu den Restklassenkörpern  $\mathbb{Z}_p$ . Für größere  $k$  lassen sich endliche Körper über Polynome auf  $\mathbb{Z}_p$  konstruieren (Polynome behandeln wir im nächsten Abschnitt).

Die endlichen Körper werden auch *Galoiskörper*<sup>6</sup> genannt und oft mit  $GF(p^k)$  bezeichnet (“GF” für “Galois Field”); auch die Notation  $\mathbb{F}_{p^k}$  ist üblich.

- Für die Informatik ist wegen der Schaltalgebra besonders

$$\mathbb{F}_2 = GF(2) = \mathbb{Z}_2 = \{0, 1\}$$

für das Rechnen mit Bits von Interesse; ebenso auch die Zweierpotenzen, z.B.

$$GF(2^8) = \mathbb{F}_{256}$$

für das Rechnen mit Bytes. (Für  $k > 1$  liegt allerdings kein Restklassenkörper vor.)

In  $\mathbb{F}_2$  entsprechen Addition und Subtraktion (modulo 2) der gleichen Operation, nämlich einem bitweisen **xor**. Dazu im Abschnitt zu Polynomen noch mehr.

In Körpern  $\mathbb{K}$  gibt es jeweils genau ein additives neutrales Element – die “Null”. Weiterhin ist, wenn wir diese Null als 0 bezeichnen, die multiplikative Gruppe gerade die Menge  $\mathbb{K} \setminus \{0\}$ . Als abelsche Gruppe ist diese insbesondere ein Magma, d.h. abgeschlossen. Somit kann innerhalb der multiplikativen Gruppe die Null *nicht* durch Multiplikation erzeugt werden!

Dies fassen wir zusammen mit folgendem

**Satz 5.16** (Satz vom Nullprodukt). *Sei  $\mathbb{K}$  ein Körper und  $a, b \in \mathbb{K}$ . Dann gilt:*

$$(a \cdot b = 0) \Leftrightarrow ((a = 0) \vee (b = 0))$$

#### Bemerkungen:

- Ein Produkt zweier Körperelemente kann also nur genau dann null sein, wenn mindestens einer der beiden Faktoren selbst null ist!
- Alternative Formulierung des Satzes:

*Körper sind nullteilerfrei*

Denn Nullteiler wären sämtliche Elemente ungleich null, die mit einem geeigneten komplementären Teiler im Produkt null ergeben. Und diese müssten sich innerhalb der multiplikativen Gruppe eines Körpers befinden – aber dann wäre selbige Gruppe nicht länger abgeschlossen, da es Produkte gäbe, die außerhalb der Gruppe lägen (nämlich genau die mit Produktwert 0). Entsprechend wäre diese Menge nicht nur keine Gruppe mehr, sondern nicht einmal ein Magma, und die Struktur insgesamt könnte kein Körper sein.

<sup>6</sup>E. Galois, französischer Mathematiker

- Die Richtung “ $\Leftarrow$ ” in obigem Satz gilt hingegen in allen Ringen, denn für beliebige Ringelemente  $a, b$  gilt:

$$a \cdot b = a \cdot (b + 0) = (a \cdot b) + (a \cdot 0)$$

Subtrahiert man von linker und rechter Seite der Gleichung das Produkt  $(a \cdot b)$  – und das ist immer möglich, da die Multiplikation im Ring abgeschlossen ist; das Produkt ist also ein Ringelement, und hat als solches auch ein additives Inverses –, so gilt:

$$\dots \Leftrightarrow 0 = (a \cdot b) - (a \cdot b) = a \cdot 0$$

Und weil  $a$  beliebig gewählt war, führt die Multiplikation jedes Ringelements mit Null auf die Null.

### Beispiele:

- Im Ring  $\mathbb{Z}_6$ , der *kein* Körper ist, lässt sich die Null z.B. durch das Produkt  $2 \cdot 3$  erzeugen. Dieser Ring ist demnach nicht nullteilerfrei (denn sowohl 2 als auch 3 liegen innerhalb  $\mathbb{Z}_6 \setminus \{0\}$ ).
- Der Ring  $\mathbb{Z}_7$  ist jedoch ein Körper. Hier ist es nicht möglich, die Null durch Multiplikation der Körperelemente aus der multiplikativen Gruppe zu erzeugen.

Neben  $\mathbb{Q}$  und  $\mathbb{Z}_p$  (mit  $p$  prim) haben auch die *reellen Zahlen*  $\mathbb{R}$  die Körpereigenschaft. Das können wir mit den Mitteln dieser Vorlesung nicht zeigen, aber immerhin (leicht vorgreifend auf Mathematik 2) motivieren:

Jede reelle Zahl lässt sich als *Grenzwert* (oder: *Limes*) einer Folge rationaler Zahlen ausdrücken. Dabei übertragen sich die Rechenregeln in verträglicher Weise. Falls also ein Produkt von zwei irrationalen Zahlen  $a, b$  zu berechnen ist, wird dessen Zahlenwert auch sehr dicht bei dem Produkt von zwei rationalen Zahlen liegen, von denen die eine nahe bei  $a$  und die andere nahe bei  $b$  liegt. Für Summen, Differenzen und Quotienten gilt entsprechendes (Stichwort *Limessatz*). Und da die rationalen Zahlen eine Körperstruktur bilden, gilt dies deswegen auch für die reellen Zahlen.

(Der Grenzwertbegriff ist hier kein Vorlesungsstoff; siehe dazu den Analysis-Teil von Mathematik 2.)

Alle rationalen Zahlen lassen sich in  $\mathbb{R}$  einbetten. Insgesamt gilt, und dies akzeptieren wir hier ohne Beweis:

$\mathbb{R}$  ist ein Körper

## 5.4 Polynome in einer Veränderlichen

Wir haben in Definition 4.9 bereits die ganzrationalen Funktionen kennen gelernt, die sich über endliche Summen von skalierten Potenzen einer Variablen definieren. Solche Funktionen sind eng verwandt mit den Polynomen, die wir in diesem Abschnitt betrachten: sie werden *Polynomfunktionen* genannt. Tatsächlich findet man zu jedem Polynom eindeutig eine Polynomfunktion; die Umkehrung gilt aber (s.u.) nicht immer – es gibt Fälle, für die verschiedene Polynome auf die gleiche Polynomfunktion führen.

Polynome sind gewissermaßen die algebraischen Objekte hinter den ganzrationalen Funktionen. Um den Unterschied zwischen beiden Konzepten heraus zu stellen, schreiben wir die Veränderlichen von Polynomen groß, während sie bei numerischen Funktionen (auch den ganzrationalen) meist eher klein geschrieben sind).

### 5.4.1 Definitionen und Beispiele

**Definition 5.17** (Polynom). Sei  $\mathbb{K}$  ein kommutativer Ring mit Eins. Dann heißt für  $n \in \mathbb{N}_0$  der Summenausdruck

$$p(X) = \sum_{j=0}^n c_j \cdot X^j = c_n X^n + \cdots + c_2 X^2 + c_1 X + c_0$$

Polynom über  $\mathbb{K}$  in der Veränderlichen  $X$ .

Die Zahlen  $c_j \in \mathbb{K}$  heißen Koeffizienten des Polynoms; dabei gilt  $c_n \neq 0$ .

Die Ausdrücke  $X^j$  (das sind  $j$ -fache Produkte von  $X$  mit sich selbst) heißen Monome.

Die Zahl  $n =: \deg(p)$  heißt Grad des Polynoms<sup>7</sup>.

Das Polynom  $c_0 = 0$  heißt Nullpolynom; für dieses vereinbaren wir keinen Grad.

Die Menge aller Polynome über  $\mathbb{K}$  in der Veränderlichen  $X$  wird mit  $\mathbb{K}[X]$  bezeichnet.

Die das Polynom definierenden Koeffizienten können auch als  $(n+1)$ -Tupel  $p$  mit  $p_j := c_j$  notiert werden (Tupel-Index hier von rechts ab 0 gezählt):

$$p = (c_n, c_{n-1}, \dots, c_1, c_0) \in \mathbb{K}^{n+1}$$

Zwei Polynome in  $X$  sind gleich, falls ihre Koeffiziententupel gleich sind.

Jedes Polynom, dessen sämtliche Koeffizienten 0 betragen, ist gleich dem Nullpolynom (0).

#### Bemerkungen:

- Die Veränderliche  $X$  wird ebenfalls als ein (unbekanntes) Element aus  $\mathbb{K}$  angenommen. Für das Polynom selbst dient sie allerdings nur als Platzhalter. In  $\mathbb{K}[X]$  ist  $X$  also zunächst nur ein Name; genauso dürften wir die Veränderliche auch  $Z$  nennen – ein Polynom in letzterer Veränderlichen wäre dann aus  $\mathbb{K}[Z]$ .
- (Polynome können auch mehrere Veränderliche besitzen (kein Vorlesungsstoff). Dann würde man z.B.  $\mathbb{K}[X, Y]$  für Polynome mit den Veränderlichen  $X$  und  $Y$  notieren.)
- Der Absolutbeitrag  $c_0$  lässt sich mit der Konvention  $X^0 := 1$  auch als  $c_0 X^0$  schreiben; das wurde oben bereits so notiert.
- Das Polynom ist eine Summe von mehreren (daher “poly”, gr. für “viele”) skalierten Monomen<sup>8</sup> (“monos” gr. für “einzeln”).
- Wird  $p$  als Tupel notiert, aber  $p \in \mathbb{K}[X]$  vereinbart, ist somit klar, dass die Veränderliche in  $p(X)$  genau  $X$  heißt. Die Tupelnotation  $p$  und die ausgeschriebene Summe mit Potenzen in  $p(X)$  sind somit gleichwertig.
- Oft werden Polynome über *Körpern* betrachtet (daher das Symbol  $\mathbb{K}$ ) – aber kommutative Ringe mit Eins ohne Körpereigenschaft sind (siehe die Definition) ebenfalls möglich. Insbesondere gelten aber stets die Kommutativ- und Assoziativgesetze für Addition und Multiplikation; außerdem sind stets die beiden neutralen Elemente 0 und 1 Elemente von  $\mathbb{K}$ .
- Wegen der Ringeigenschaft aller Beiträge im Polynom werden (s.u.) problemlos Summen, Differenzen und Produkte von Polynomen definierbar sein, mit dem erwarteten Verhalten dieser Operationen. Die Division von Polynomen erfordert mehr Voraussetzungen und wird danach gesondert behandelt; siehe dazu die Bemerkungen dort.

#### Beispiele:

- $4, \frac{4}{3}$  und  $\sqrt{7}$  sind reelle Polynome (also Polynome aus  $\mathbb{R}[X]$ ) vom Grad 0.  
In der Tupel-Notation würde man diese Zahlen noch in Klammern setzen.
- $X^3 - 9X + 12$  ist ein reelles Polynom vom Grad 3. Als Tupel:  $(1, 0, -9, 12)$ . *Man beachte, dass in der Tupel-Notation auch die Koeffizienten mit Wert 0 aufzuführen sind!*  
(Falls die Koeffizienten reelle Zahlen sind, wird meist nicht weiter eingeschränkt. Die ersten beiden Polynome wären eigentlich auch rationale Polynome, aber wir werden später sehen, dass schon auf  $\mathbb{R}$  bestimmte Rechnungen nicht immer möglich sind – die Verwendung von  $\mathbb{Q}[X]$ )

<sup>7</sup>engl.: “degree”

<sup>8</sup>Im Exkurs A.5 und mit dem Wissen von Kapitel 7 sehen wir, dass ein Polynom eine *Linearkombination* von Monomen ist; siehe dort.

oder gar  $\mathbb{N}[X]$  ist daher nur dann sinnvoll, wenn die Koeffizienten absichtlich eingeschränkt werden sollen.

Im übrigen sind von diesen drei Möglichkeiten nur Polynome aus  $\mathbb{R}[X]$  verträglich mit ganzrationalen Funktionen (welche *stetig* sind, d.h. sich mit einem Stift ohne abzusetzen zeichnen lassen<sup>9</sup>).)

- Genau wie bei ganzrationalen Funktionen heißen Polynome vom Grad 1 *linear*, solche vom Grad 2 *quadratisch* und solche mit Grad 3 *kubisch*. Das eben gezeigte Polynom war also ein kubisches reelles Polynom.
- Ein Polynom über  $\mathbb{F}_2$  vom Grad 8 wäre:

$$X^8 + X^4 + X^3 + X + 1$$

(dieses Polynom spielt eine wichtige Rolle in der Kryptographie (AES-Standard).)

Wie oben schon bemerkt, gibt es bei  $\mathbb{F}_2$  keinen Unterschied zwischen Addition und Subtraktion (modulo 2), sodass es hier reicht, die Monome zu addieren. Außerdem lässt man die Skalierungsfaktoren 1 gewohnheitsmäßig weg – das auch bei anderen Polynomen; aber hier besonders, da 1 ohnehin der einzige mögliche Wert ungleich 0 für Koeffizienten ist.

Als Tupel:

$$(1, 0, 0, 0, 1, 1, 0, 1, 1)$$

Speziell für  $\mathbb{F}_2$  ist es außerdem üblich (da die Koeffizienten stets einzelne Ziffern sind), die Tupelnotation noch weiter abzukürzen, indem man die Bits direkt aneinander schreibt. Hier:

$$100011011$$

- Weil Koeffizienten aus Restklassenringen stets als nichtnegative Zahlen ausgedrückt werden können, kann man sich für Polynome aus  $\mathbb{Z}_m[X]$  auf solche beschränken. Denn das Polynom

$$3X^4 - 4X + 2$$

auf  $\mathbb{Z}_6[X]$  ist wegen  $(-4) \equiv 2 \pmod{6}$  äquivalent zu

$$3X^4 + 2X + 2$$

Letzteres als Tupel:  $(3, 0, 0, 2, 2)$

### 5.4.2 Polynomfunktion vs. Polynom

**Definition 5.18** (Polynomfunktion). *Für einen kommutativen Ring mit Eins  $\mathbb{K}$  existiert zu jedem Polynom  $p(X) \in \mathbb{K}[X]$  eine Funktion  $f_p: \mathbb{K} \rightarrow \mathbb{K}$ , sodass für*

$$p(X) = \sum_{j=0}^n c_j X^j$$

*gilt:*

$$f_p(x) = \sum_{j=0}^n c_j x^j$$

*Diese Funktion heißt Polynomfunktion.*

#### Bemerkungen:

- Während die Polynome in  $\mathbb{K}[X]$  algebraische Objekte sind, bei denen die Unbekannte  $X$  frei gelassen wird, beschreibt die Polynomfunktion die Abbildung, falls für  $X$  ein konkreter Wert  $x$  eingesetzt wird. Das Ergebnis ist dann ein Element von  $\mathbb{K}$ . Polynomfunktionen besitzen daher auch einen Funktionsgraph (der ggf. gezeichnet werden kann) – Polynome nicht.

<sup>9</sup>Eine mathematisch saubere Definition der Stetigkeit folgt in Mathematik 2/Analysis

- Die Zuordnung  $p \mapsto f_p$  ist eindeutig (also als Funktion begreifbar) – aber sie ist *nicht* immer injektiv. Als Beispiel betrachten wir  $\mathbb{F}_2[X]$ : Die beiden Polynome

$$X^3 + X^2 + 1 \quad \text{und} \quad X^4 + X + 1$$

sind offensichtlich verschieden, da sie nicht einmal gleichen Grad besitzen. Ihre Polynomfunktionen sind aber im Abbildungsverhalten identisch, entsprechen also der gleichen Abbildung – nämlich

$$x \mapsto 1$$

Dies bestätigt man leicht durch Einsetzen der beiden möglichen Werte für  $x$ . Im Fall  $x = 1$  ergibt sich sowohl für  $x \mapsto (x^3 + x^2 + 1) \bmod 2$  als auch für  $x \mapsto (x^4 + x + 1) \bmod 2$  der Wert 1:  $1^3 + 1^2 + 1 = 1 + 1 + 1 \equiv 1$  sowie  $1^4 + 1 + 1 = 1 + 1 + 1 \equiv 1$ . Für  $x = 0$  bilden ebenfalls beide Funktionen auf 1 ab.

### 5.4.3 Rechenregeln, Polynomring

Die Polynome in einer Veränderlichen lassen sich, wie oben schon bemerkt, problemlos addieren, subtrahieren und multiplizieren – dies liegt daran, dass sowohl für die Koeffizienten als auch für die Variable die Ringeigenschaft vorliegt (genauer: kommutativer Ring mit Eins, d.h. die Multiplikation ist kommutativ und hat das neutrale Element 1).

Da die Veränderliche aber unbekannt ist, dürfen wir sie beim Rechnen mit Polynomen nicht mit zusätzlichen Annahmen versehen: auch nach einer Rechenoperation mit zwei Polynomen aus  $\mathbb{K}[X]$  bleibt  $X$  unbekannt. Die Ringstruktur erlaubt uns aber, das Distributivgesetz anzuwenden. Zum Beispiel ist es in  $\mathbb{R}[X]$  richtig zu schreiben:

$$2X \cdot (4X^3 - 13) = 8X^4 - 26X$$

Denn egal welchen konkreten Wert  $x$  die Unbekannte  $X$  annimmt, wäre diese Gleichung stets korrekt. Neben dem Distributivgesetz wurde hier auch die Assoziativität und Kommutativität der Multiplikation verwendet, um auf der rechten Seite wieder ein Polynom mit der üblichen Struktur (Koeffizienten links von den Potenzen in  $X$ ) herzustellen.

Zum besseren Verständnis der Unterschiede zwischen den Operationen auf Polynomen und denen auf dem Ring der Koeffizienten/Unbekannten führen wir (wie oben bei den Restklassen) temporär neue Symbole (“ $\oplus$ ” und “ $\odot$ ”) ein, auf die dann später wegen der guten Verträglichkeit wieder verzichten können.

---

Wir stellen fest, dass schon in obiger Definition 5.17 eine Art Ungenauigkeit vorliegt – denn da jedes Monom  $X^j$  ein Polynom vom Grad  $j$  ist, und jeder Skalierungsfaktor  $c_j$  ein Polynom vom Grad 0, könnte man auch schreiben:

$$p(X) = ((c_n) \odot X^n) \oplus \cdots \oplus ((c_1) \odot X^1) \oplus (c_0)$$

Die Inneren Klammern um die  $c_j$  bezeichnen hierbei Polynome in Tupelnotation. Die beiden hier notierten äußeren Klammern drücken die in Halbringen übliche Regel “Punktrechnung vor Strichrechnung” explizit aus.

Wenn das Polynom aber später in Gestalt seiner zugehörigen Polynomfunktion ausgewertet wird, müssen ohnehin die Ringoperationen von  $\mathbb{K}$  benutzt werden, so wie sie oben in der Definition schon angeschrieben waren. Für jeden konkreten Wert  $x \in \mathbb{K}$  hat die Polynomfunktion  $f_p$  einen Funktionswert aus  $\mathbb{K}$ . Es liegt also nahe, die Rechenoperationen für Polynome so zu definieren, dass sie selbst die Ringeigenschaft besitzen – dann ist die (oben schon angenommene) Verträglichkeit leicht nachzuweisen.

---

**Definition 5.19** (Addition von Polynomen). *Für einen kommutativen Ring mit Eins  $(\mathbb{K}, +, \cdot)$  und Polynome  $p, q \in \mathbb{K}[X]$  mit den Graden  $m := \deg(p)$  und  $n := \deg(q)$  hat das Summenpolynom*

$$p \oplus q$$

die Koeffizienten

$$(p \oplus q)_j := p_j + q_j$$

Falls  $\text{oBdA}^{10}$   $m < n$ , sind die Koeffizienten  $p_{m+1}$  bis  $p_n$  jeweils mit 0 zu ergänzen.

---

<sup>10</sup>“ohne Beschränkung der Allgemeinheit”



### Bemerkungen:

- Es handelt sich also um eine *komponentenweise* Addition der Koeffizienten. Diese ist mit dem Distributivgesetz von  $\mathbb{K}$  verträglich, denn auf  $\mathbb{K}$  muss für jeden Wert von  $X$  und alle relevanten Indices  $j$  gelten:

$$(p_j \cdot X^j) + (q_j \cdot X^j) = (p_j + q_j) \cdot X^j$$

- Da  $\mathbb{K}$  bezüglich der Addition abgeschlossen ist, ist auch jeder Koeffizient des Summenpolynoms aus  $\mathbb{K}$ , und damit liegt wieder ein Polynom aus  $\mathbb{K}[X]$  vor. Also ist die Addition von Polynomen abgeschlossen auf  $\mathbb{K}[X]$
- Über die komponentenweise Definition ist auch klar, dass die Addition von Polynomen assoziativ ist.
- Weiterhin gibt es ein neutrales Element der Addition – nämlich  $(0)$ . Wird das Nullpolynom  $(0)$  auf ein Polynom  $p$  addiert, so erhält man nach Definition als Summe wieder  $p$ .
- Die Addition ist wegen der Kommutativität der Addition auf  $\mathbb{K}$  ebenfalls kommutativ.
- Da  $\mathbb{K}$  ein Ring ist, besitzt jedes Element ein additives Inverses. Also existiert auch ein additiv inverses Polynom zu  $p$ , nämlich das Polynom  $(-p)$ , in welchem alle Koeffizienten  $p_j$  durch ihre Inversen  $(-p_j)$  ersetzt sind. Die Summe der beiden Polynome ist dann das Nullpolynom.
- Also ist  $(\mathbb{K}[X], \oplus)$  eine abelsche Gruppe, so wie es auch  $(\mathbb{K}, +)$  ist.

---

Wir betrachten noch den Grad des Summenpolynoms. Sicher kann dieser die Grade der Summandenpolynome nicht übersteigen. Haben die Summanden verschiedene Grade, so muss der Grad der Summe dem größeren der beiden Grade entsprechen, denn der zugehörige Koeffizient entspricht dann dem höchsten Koeffizient des höhergradigen Polynoms, der dabei unverändert bleibt.

Bei Polynomen gleichen Grads kann die Summe ebenfalls diesen Grad annehmen, muss es aber nicht. Denn es wäre möglich, dass die beiden höchsten Koeffizienten sich gegenseitig neutralisieren. Im Extremfall (siehe obige Bemerkungen) kann die Summe zweier Polynome sogar das Nullpolynom ergeben – und wenn man diesem einen Grad zuordnet, so kann dieser nicht größer sein als  $0^{11}$ .

Es gilt also folgender

**Satz 5.20** (Grad des Summenpolynoms). Für  $p, q \in \mathbb{K}[X]$  gilt:

$$\deg(p \oplus q) \leq \max\{\deg(p), \deg(q)\}$$

**Beispiel:** Wir betrachten die folgenden drei Polynome  $p, q, s$  über verschiedenen Ringen und bilden jeweils die Summen  $p \oplus q$  sowie  $q \oplus s$ :

$$p(X) := 3X^4 - 2X^2 + 7X + 4$$

$$q(X) := 5X^3 + 8X^2 - 5X + 3$$

$$s(X) := 2X^3 + 12X - 9$$

- Auf  $\mathbb{R}$  steht schon alles richtig da – wir haben nur noch die komponentenweise Addition der Koeffizienten auszuführen:

$$(p \oplus q)(X) = 3X^4 + 5X^3 + 6X^2 + 2X + 7$$

$$(q \oplus s)(X) = 7X^3 + 8X^2 + 7X - 6$$

- Auf  $\mathbb{F}_7 = \mathbb{Z}_7$  empfiehlt es sich, die Polynome zunächst auf eindeutige Repräsentanten umzuschreiben, denn statt der Zahlen stehen oben streng genommen die Restklassen modulo 7 dieser Zahlen. Wir erhalten:

$$p(X) = 3X^4 + 5X^2 + 4$$

$$q(X) = 5X^3 + X^2 + 2X + 3$$

$$s(X) = 2X^3 + 5X + 5$$

---

<sup>11</sup>Oft wird für das Nullpolynom auch ein Grad von  $-\infty$  vereinbart

Bei der komponentenweisen Addition gelten jetzt die Rechenregeln von  $\mathbb{Z}_7$ , d.h. wir erhalten:

$$\begin{aligned}(p \oplus q)(X) &= 3X^4 + 5X^3 + 6X^2 + 2X \\ (q \oplus s)(X) &= X^2 + 1\end{aligned}$$

In der ersten Summe ist das Absolutglied 7 verschwunden, da  $4 + 3 \equiv 0 \pmod{7}$ . In der zweiten Summe hatten sich die kubischen und die linearen Koeffizienten zu 7 addiert und sind daher verschwunden; im Absolutglied steht  $8 \equiv 1 \pmod{7}$ .

- Für den Ring  $\mathbb{Z}_6$  (der kein Körper ist), verfahren wir analog:

$$\begin{aligned}p(X) &= 3X^4 + 4X^2 + X + 4 \\ q(X) &= 5X^3 + 2X^2 + X + 3 \\ s(X) &= 2X^3 + 3 \\ (p \oplus q)(X) &= 3X^4 + 5X^3 + 2X + 1 \\ (q \oplus s)(X) &= X^3 + 2X^2 + X\end{aligned}$$

- Und für den Körper  $\mathbb{F}_3 = \mathbb{Z}_3$ :

$$\begin{aligned}p(X) &= X^2 + X + 1 \\ q(X) &= 2X^3 + 2X^2 + X \\ s(X) &= 2X^3 \\ (p \oplus q)(X) &= 2X^3 + 2X + 1 \\ (q \oplus s)(X) &= X^3 + 2X^2 + X\end{aligned}$$

- Und für den Körper  $\mathbb{F}_2 = \mathbb{Z}_2$ :

$$\begin{aligned}p(X) &= X^4 + X \\ q(X) &= X^3 + X + 1 \\ s(X) &= 1 \\ (p \oplus q)(X) &= X^4 + X^3 + 1 \\ (q \oplus s)(X) &= X^3 + X\end{aligned}$$

Da die Koeffizienten der drei reellen Polynome alle aus  $\mathbb{Z}$  sind (sonst hätte die Übertragung auf Restklassenringe so auch nicht funktioniert!), hätte es im übrigen gereicht, die (ebenfalls ganzzahligen) Koeffizienten der reellen Summenpolynome für  $\mathbb{Z}_m$  jeweils modulo  $m$  zu nehmen (das prüfe man als Übung gerne nach).

Die Übertragung von  $\mathbb{Z}[X]$  auf  $\mathbb{Z}_m[X]$  bedeutet übrigens einen Informationsverlust und lässt sich nicht rückgängig machen. Ein deutliches Beispiel ist  $s(X)$ , das in  $\mathbb{Z}_2[X]$  auf das Polynom (1) "schrumpft" – daraus lässt sich das ursprüngliche  $s(X)$  aus  $\mathbb{R}[X]$  sicher nicht rekonstruieren!

Wir führen die vorigen Rechnungen auch nochmal in Tupelnotation aus (diese ist gleichwertig). Wir lassen die Kommata weg und ersetzen sie statt dessen durch lesbar große Abstände (wie bei der Zyklenschreibweise von Permutationen). Es lohnt sich, die Tupel wie bei der schriftlichen Addition anzuschreiben – man beachte aber, dass es hier *keine* Überträge gibt! Denn die Addition geschieht durchgängig komponentenweise, und eine Summe von  $a \cdot X^7$  und  $b \cdot X^7$  kann z.B. nicht irgendeinen Beitrag  $c \cdot X^8$  ergeben. Für vereinzelte Werte  $x$  der Variablen  $X$  mag dies teilweise numerisch richtig sein, aber der Wert von  $X$  ist hier eben nicht bekannt.

- In  $\mathbb{R}$ :

$$\begin{array}{rcl}p & = & ( \quad 3 \quad 0 \quad -2 \quad 7 \quad 4 \quad ) \\ q & = & ( \quad \quad 5 \quad 8 \quad -5 \quad 3 \quad ) \\ s & = & ( \quad \quad 2 \quad 0 \quad 12 \quad -9 \quad ) \\ \hline p \oplus q & = & ( \quad 3 \quad 5 \quad 6 \quad 2 \quad 7 \quad ) \\ q \oplus s & = & ( \quad \quad 7 \quad 8 \quad 7 \quad -6 \quad )\end{array}$$

- In  $\mathbb{Z}_7$ :

$$\begin{array}{rcl}p & = & ( \quad 3 \quad 0 \quad 5 \quad 0 \quad 4 \quad ) \\ q & = & ( \quad \quad 5 \quad 1 \quad 2 \quad 3 \quad ) \\ s & = & ( \quad \quad 2 \quad 0 \quad 5 \quad 5 \quad ) \\ \hline p \oplus q & = & ( \quad 3 \quad 5 \quad 6 \quad 2 \quad 0 \quad ) \\ q \oplus s & = & ( \quad \quad \quad 1 \quad 0 \quad 1 \quad )\end{array}$$

- In  $\mathbb{Z}_6$ :

$$\begin{array}{rcl} p & = & ( \quad 3 \quad 0 \quad 4 \quad 1 \quad 4 \quad ) \\ q & = & ( \quad \quad 5 \quad 2 \quad 1 \quad 3 \quad ) \\ s & = & ( \quad \quad 2 \quad 0 \quad 0 \quad 3 \quad ) \\ \hline p \oplus q & = & ( \quad 3 \quad 5 \quad 0 \quad 2 \quad 1 \quad ) \\ q \oplus s & = & ( \quad \quad 1 \quad 2 \quad 1 \quad 0 \quad ) \end{array}$$

- In  $\mathbb{Z}_3$ :

$$\begin{array}{rcl} p & = & ( \quad \quad \quad 1 \quad 1 \quad 1 \quad ) \\ q & = & ( \quad \quad 2 \quad 2 \quad 1 \quad 0 \quad ) \\ s & = & ( \quad \quad 2 \quad 0 \quad 0 \quad 0 \quad ) \\ \hline p \oplus q & = & ( \quad \quad 2 \quad 0 \quad 2 \quad 1 \quad ) \\ q \oplus s & = & ( \quad \quad 1 \quad 2 \quad 1 \quad 0 \quad ) \end{array}$$

- In  $\mathbb{Z}_2$  in Kurzschreibweise:

$$\begin{array}{rcl} p & = & 10010 \\ q & = & 1011 \\ s & = & 1 \\ \hline p \oplus q & = & 11001 \\ q \oplus s & = & 1010 \end{array}$$

Für die Multiplikation zweier Polynome motivieren wir die Definition, indem wir zunächst das Produkt zweier skaliert Monome  $p_j X^j$  und  $q_k X^k$  betrachten – dies sind ja auch spezielle Polynome mit je nur einem Glied.

Wenn wir auf  $\mathbb{K}$  rechnen und erwarten, dass für die Unbekannte später ein beliebiger Wert  $x \in \mathbb{K}$  eingesetzt werden kann, sollte sich nach den Potenzgesetzen und den Regeln der Ringarithmetik das Produkt als

$$p_j X^j \cdot q_k X^k = (p_j \cdot q_k) X^{j+k}$$

darstellen. Der Skalierungsfaktor des Monoms  $X^{j+k}$  entspricht also dem Produkt von  $p_j$  und  $q_k$ , das wir auf dem Ring  $\mathbb{K}$  stets ausführen können.

Wenn aber die Summe der Exponenten  $(j+k)$  nicht gerade 0 ist, so lässt sie sich auf verschiedene Weise aus zwei Summanden zusammen setzen. Sei  $l := j + k$ , dann gilt:

$$\begin{aligned} l &= 0 + l \\ &= 1 + (l - 1) \\ &= 2 + (l - 2) \\ &\dots \\ &= (l - 1) + 1 \\ &= l + 0 \end{aligned}$$

Es gibt also bis zu  $(l + 1)$  Möglichkeiten, den Exponent  $l$  aus verschiedenen Paaren von  $j$  und  $k$  aus  $\mathbb{N}_0$  zu kombinieren. Nicht alle davon sind stets realisierbar, da die Indices  $j$  und  $k$  nach oben durch die jeweiligen Polynomgrade  $m := \deg(p)$  und  $n := \deg(q)$  beschränkt sind.

Es gilt aber, dass zum Exponenten  $l$  des Produktpolynoms  $p \odot q$  alle die Indexpaare aus  $j$  und  $k$  beitragen, für die  $j + k = l$  gilt. Dies betrachten wir zunächst an einem Beispiel, dann allgemein-formal.

---

**Beispiel:** Wir berechnen (auf  $\mathbb{R}$ ) das Produkt der Polynome  $p$  und  $s$  aus dem vorigen Beispiel. Dabei verwenden wir mehrfach das Distributivgesetz, um die einzelnen Monom-Terme nach der Art “jedes mit jedem” zu multiplizieren. Zunächst spalten wir die linke Klammer auf; dann in jedem Beitrag jeweils die rechte. Danach sammeln wir die Terme mit gleichem Exponenten auf,

wobei wir wieder das Distributivgesetz anwenden.

$$\begin{aligned}
(p \odot s)(X) &= p(X) \cdot s(X) = (3X^4 - 2X^2 + 7X + 4) \cdot (2X^3 + 12X - 9) \\
&= \begin{array}{l} (3X^4) \cdot (2X^3 + 12X - 9) \\ + (-2X^2) \cdot (2X^3 + 12X - 9) \\ + (7X) \cdot (2X^3 + 12X - 9) \\ + (4) \cdot (2X^3 + 12X - 9) \end{array} \\
&= \begin{array}{l} (3X^4) \cdot (2X^3) + (3X^4) \cdot (12X) + (3X^4) \cdot (-9) \\ + (-2X^2) \cdot (2X^3) + (-2X^2) \cdot (12X) + (-2X^2) \cdot (-9) \\ + (7X) \cdot (2X^3) + (7X) \cdot (12X) + (7X) \cdot (-9) \\ + (4) \cdot (2X^3) + (4) \cdot (12X) + (4) \cdot (-9) \end{array} \\
&= \begin{array}{l} (6X^7 + 36X^5 - 27X^4) \\ + (-4X^5 - 24X^3 + 18X^2) \\ + (14X^4 + 84X^2 - 63X) \\ + (8X^3 + 48X - 36) \end{array} \\
&= 6X^7 + (36 - 4)X^5 + (-27 + 14)X^4 \\
&\quad + (-24 + 8)X^3 + (18 + 84)X^2 + (-63 + 48)X - 36 \\
&= 6X^7 + 32X^5 - 13X^4 - 16X^3 + 102X^2 - 15X - 36
\end{aligned}$$

Mit etwas Übung lässt man nach dem Einsetzen der beiden Faktorpolynome meist die nächsten beiden Schritte (explizites Ausschreiben des Distributivgesetzes) weg und rechnet die relevanten Produkte direkt aus – es empfiehlt sich aber hier, strukturiert vorzugehen, um keine Terme auszulassen. Auch der vorletzte Schritt zum Sammeln der Terme wird oft ausgelassen – hier ist die Gefahr für Flüchtigkeitsfehler allerdings noch größer, weil nicht immer gleich viele Terme zu sammeln sind.

Die gleiche Rechnung führen wir nochmal in Tupelschreibweise aus, und zwar analog zu oben im Stil der schriftlichen Multiplikation. Um hier die gleichen Teilbeiträge wie oben zu erhalten, drehen wir die beiden Faktoren hier um, was aber wegen der Kommutativität in Addition und Multiplikation auf  $\mathbb{K}$  grundsätzlich kein Problem darstellt. Im Mittelteil verzichten wir dabei nicht nur auf die Kommata, sondern auch auf die Tupelklammern:

$$\begin{array}{rcccccccc}
(2 & 0 & 12 & -9) & \odot & (3 & 0 & -2 & 7 & 4) \\
& & & & & & 8 & 0 & 48 & -36 \\
& & & & & 14 & 0 & 84 & -63 & \\
& & & -4 & 0 & -24 & 18 & & & \\
& & 0 & 0 & 0 & 0 & & & & \\
& 6 & 0 & 36 & -27 & & & & & \\
\hline
(6 & 0 & 32 & -13 & -16 & 102 & -15 & -36)
\end{array}$$

Auch hier ist darauf zu achten, dass keine Überträge beim Addieren der Terme aus dem Mittelteil anfallen. Die Nullzeile im Mittelteil kann ausgelassen werden, ebenso die Additionszeichen für die Zeilen aus dem Mittelteil.

Der Grund, warum die Notation so funktioniert, liegt darin, dass in den Tupeln jeweils alle Potenzen von  $X$  aufgeführt sind. Von rechts nach links steigen diese Potenzen also gleichmäßig in Einerschritten an; der Versatz der Zeilen im Mittelteil entspricht gerade dem Ausklammern der Potenzen. Üblicherweise beginnt man bei der niedrigstwertigen Potenz – sind die Zeilen gegenüber der ausgeschriebenen Rechnung oben umgekehrt sortiert.

Jede aufgeführte Zeile entsteht dadurch, dass das links stehende Polynom mit dem jeweiligen Faktor aus dem rechten Polynom skaliert wird. Die Potenz  $X^k$  vom rechten Polynom bildet sich in dem Versatz nach links hin ab. Die ausgelassenen Null-Einträge rechts von den versetzten Zeilen im Mittelteil würden streng genommen dazu gehören, wenn auch im Mittelteil korrekt als Tupel notierte Polynome stehen sollten. Im Interesse besserer Lesbarkeit können diese allerdings ausgelassen werden, wie hier schon geschehen – in der Summierung zum Ergebnispolynom tragen sie nicht bei. Im Tupel des Ergebnispolynoms müssen hingegen wieder alle Potenzen berücksichtigt werden. So erklärt sich auch der Koeffizient 0 für das Monom  $X^6$ .

Als Übung empfiehlt es sich, die Produkte sowohl ausgeschrieben als auch in Tupelnotation mit schriftlicher Multiplikation auch für die oben betrachteten Ringe  $\mathbb{Z}_7$ ,  $\mathbb{Z}_6$ ,  $\mathbb{Z}_3$  und  $\mathbb{Z}_2$  zu berechnen<sup>12</sup>. Man achte hierbei stets auf die Bildung der (eindeutigen) Divisionsreste modulo  $m$  – besonders in der Tupelschreibweise sind eindeutige Repräsentanten üblich.

Mit dem bereits berechneten reellen Ergebnispolynom erlauben wir uns hier eine Abkürzung und nehmen direkt dieses modulo  $m$ ; für die erwähnten endlichen Ringe erhalten wir dann:

<sup>12</sup>Solch eine Rechnung kann in der Prüfung verlangt sein

- In  $\mathbb{Z}_7$ :

$$(p \odot s)(X) = 6X^7 + 4X^5 + X^4 + 5X^3 + 4X^2 + 6X + 6 = (6 \ 0 \ 4 \ 1 \ 5 \ 4 \ 6 \ 6)$$

- In  $\mathbb{Z}_6$ :

$$(p \odot s)(X) = 2X^5 + 5X^4 + 2X^3 + 3X = (2 \ 5 \ 2 \ 0 \ 3 \ 0)$$

- In  $\mathbb{Z}_3$ :

$$(p \odot s)(X) = 2X^5 + 2X^4 + 2X^3 = (2 \ 2 \ 2 \ 0 \ 0 \ 0)$$

- In  $\mathbb{Z}_2$ :

$$(p \odot s)(X) = X^4 + X = (1 \ 0 \ 0 \ 1 \ 0)$$

Bevor wir formal das Produkt ausrechnen, beobachten wir noch, dass sich bei allen Beispielen, für die die Koeffizientenringe auch Körper waren, die Grade der Polynome beim Multiplizieren genau addiert haben – bei  $\mathbb{Z}_6$  jedoch nicht. Das liegt dort daran, dass die Vorfaktoren der beiden höchsten Monome komplementäre Nullteiler modulo 6 sind; im Produkt ergeben sie  $6 \equiv 0 \pmod{6}$ . Bei Körpern wäre dies nach Satz 5.16 nicht möglich. Es kann also sein, dass der Term mit höchster Potenz (der sich als das Produkt der beiden höchsten Terme der Faktoren errechnet) im Produktpolynom verschwindet, falls der Ring der Koeffizienten nicht nullteilerfrei (hier also: ein Körper) ist.

---

**Definition 5.21** (Multiplikation von Polynomen). *Für einen kommutativen Ring mit Eins  $(\mathbb{K}, +, \cdot)$  und Polynome  $p, q \in \mathbb{K}[X]$  mit den Graden  $m := \deg(p)$  und  $n := \deg(q)$  hat das Produktpolynom*

$$p \odot q$$

*die Koeffizienten*

$$(p \odot q)_l := \sum_{\substack{0 \leq s \leq \min\{l, m\} \\ 0 \leq t \leq \min\{l, n\} \\ s+t=l}} p_s \cdot q_t$$

**Bemerkungen:**

- Die drei Kriterien unterhalb des Summensymbols schränken  $s$  auf den Bereich der Indices im Polynom  $p$  ein,  $t$  auf den Bereich der Indices in  $q$ , und stellen sicher, dass die Summe beider Indices genau  $l$  beträgt. Dies hatten wir in der Vorbereitung zum Produkt bereits so motiviert, ohne es jedoch formal zu notieren.

*In der Praxis wird jedoch so gerechnet wie im obigen Beispiel; die etwas umständliche Notation in der Definition wird dann erfüllt durch das sorgfältige Aufsammeln aller Terme mit gleicher Potenz.*

- Falls  $\mathbb{K}$  ein Körper ist, so addieren sich die Polynomgrade der Faktoren zu dem des Produkts, d.h. das maximale  $l$  hat den Wert  $(m+n)$ . Der Koeffizient von  $X^{m+n}$  entspricht dann genau  $p_m \cdot q_n$ . Analog entspricht das Absolutglied (zur Potenz  $X^0$ ) des Produkts stets  $p_0 \cdot q_0$ . Für alle anderen Koeffizienten sind in der Summe aus der Definition mehrere Beiträge möglich.
- Falls  $\mathbb{K}$  kein Körper ist, kann der Grad des Produkts niedriger ausfallen als  $(m+n)$ , obwohl  $m, n \neq 0$ .
- Wie in der Vorbereitung bemerkt, wird zum Multiplizieren mehrfach das Distributivgesetz von  $\mathbb{K}$  benutzt.
- Falls eines der Polynome das Nullpolynom  $(0)$  ist, so sind alle Koeffizienten des Produktpolynoms null; damit ist auch das Produktpolynom gleich  $(0)$ .  
Für Ringe mit Nullteilern (aber nur für solche!) ist es auch möglich, das Nullpolynom anderweitig zu erzeugen, z.B. auf  $\mathbb{Z}_6$  per

$$(2X^4 + 2X + 4) \odot (3X^2 + 3X) = \dots = (0)$$

- Mit obiger Definition erhält man, da die Multiplikation auf  $\mathbb{K}$  abgeschlossen ist, stets ein Polynom aus  $\mathbb{K}[X]$ , also ist  $\odot$  auf dieser Menge abgeschlossen.

- Die Assoziativität des Polynomprodukts lässt sich mit einiger Schreiarbeit zeigen.
- Auch ein neutrales Element der Multiplikation existiert; dabei handelt es sich um das Polynom

$$(1)$$

Dieses ist, da  $\mathbb{K}$  ein kommutativer Ring mit Eins ist, stets vorhanden. Dass 1 das neutrale Element von  $(\mathbb{K}, \cdot)$  ist, sorgt in obiger Definition dafür, dass bei Multiplikation eines Polynoms mit (1) sämtliche Koeffizienten des beliebigen Polynoms unverändert bleiben.

- Weil die Multiplikation auf  $\mathbb{K}$  kommutativ ist (es war ja ein kommutativer Ring gefordert), folgt aus obiger Definition, dass auch das Produkt von Polynomen kommutativ ist.
- Also ist  $(\mathbb{K}[X], \odot)$  ein kommutatives Monoid, so wie es auch  $(\mathbb{K}, \cdot)$  ist.

**Beispiel:** Wir berechnen auf  $\mathbb{R}$  die Koeffizienten von  $q \odot s$  mit  $q, s$  wie in obigen Beispielen unter Verwendung der Definition.

- Für  $l = 0$  gibt es nur den einen Beitrag:

$$(q \odot s)_0 = q_0 \cdot s_0 = 3 \cdot (-9) = -27$$

- Der Koeffizient für  $l = 1$  kann zwei Beiträge haben:

$$(q \odot s)_1 = q_0 \cdot s_1 + q_1 \cdot s_0 = 3 \cdot 12 + (-5) \cdot (-9) = 36 + 45 = 81$$

- Für  $l = 2$  gibt es bis zu drei Beiträge:

$$(q \odot s)_2 = q_0 \cdot s_2 + q_1 \cdot s_1 + q_2 \cdot s_0 = 3 \cdot 0 + (-5) \cdot 12 + 8 \cdot (-9) = 0 - 60 - 72 = -132$$

- Für  $l = 3$  gibt es bis zu vier Beiträge:

$$(q \odot s)_3 = q_0 \cdot s_3 + q_1 \cdot s_2 + q_2 \cdot s_1 + q_3 \cdot s_0 = 3 \cdot 2 + (-5) \cdot 0 + 8 \cdot 12 + 5 \cdot (-9) = 6 + 0 + 96 - 45 = 57$$

- Für  $l = 4$  könnte es bis zu fünf Beiträge geben, jedoch hat keines der Faktorpolynome einen Koeffizient mit Index 4. Ab hier werden nicht mehr alle denkbaren Kombinationen von Indices realisiert, welche die Summe 4 haben könnten:

$$(q \odot s)_4 = q_1 \cdot s_3 + q_2 \cdot s_2 + q_3 \cdot s_1 = (-5) \cdot 2 + 8 \cdot 0 + 5 \cdot 12 = -10 + 0 + 60 = 50$$

- Für  $l = 5$  sind nur zwei der theoretisch sechs möglichen Indexkombinationen realisiert:

$$(q \odot s)_5 = q_2 \cdot s_3 + q_3 \cdot s_2 = 8 \cdot 2 + 5 \cdot 0 = 16 + 0 = 16$$

- Für  $l = 6 = \deg(q) + \deg(s)$  erhalten wir genau einen Beitrag:

$$(q \odot s)_6 = q_3 \cdot s_3 = 5 \cdot 2 = 10$$

Also erhalten wir insgesamt:

$$(q \odot s)(X) = 10X^6 + 16X^5 + 50X^4 + 57X^3 - 132X^2 + 81X - 27 = (10 \quad 16 \quad 50 \quad 57 \quad -132 \quad 81 \quad -27)$$

Für den Grad des Produktpolynoms fassen wir noch zusammen:

**Satz 5.22** (Grad des Produktpolynoms). Für  $p, q \in \mathbb{K}[X]$  gilt:

$$\deg(p \odot q) \leq \deg(p) + \deg(q)$$

Da wir wissen, dass die Multiplikation von Polynomen kommutiert, reicht es, eines der beiden Distributivgesetze zu prüfen, die für eine Ringstruktur erforderlich sind. Wir betrachten für beliebige Polynome  $p, q, r \in \mathbb{K}[X]$  (also nicht nur für die obigen Beispiele!) den Ausdruck

$$p \odot (q \oplus r)$$

Seine Koeffizienten lassen sich nach obiger Definition berechnen. Ein Beitrag in der Summe hat stets die Struktur

$$p_s \cdot (q \oplus r)_t = p_s \cdot (q_t + r_t) = (p_s \cdot q_t) + (p_s \cdot r_t)$$

Hierbei haben wir im ersten Schritt die Definition 5.19 des Summenpolynoms eingesetzt, und danach das Distributivgesetz von  $\mathbb{K}$  verwendet.

Wegen der Assoziativität und Kommutativität der Addition auf  $\mathbb{K}$  lässt sich die Summe über die linken Beiträge auch als zwei Summen über jeweils einen der rechten Beiträge denken. Aber da  $s, t$  rechts genau die gleichen Werte haben wie auf der linken Seite, ist damit klar, dass die Summen über die rechten Seiten gerade die Addition von zwei Produktpolynomen ergibt, nämlich von

$$p \odot q \quad \text{und} \quad p \odot r$$

Damit ist das Distributivgesetz gezeigt.

Wir haben nun alle Kriterien gesammelt, um diesen Unterabschnitt abzuschließen mit folgendem

**Satz 5.23** (Polynomring). *Die Struktur  $(\mathbb{K}[X], \oplus, \odot)$  ist ein kommutativer Ring mit Eins. Die neutralen Elemente sind  $(0)$  für  $\oplus$ , und  $(1)$  für  $\odot$ .*

#### Bemerkungen:

- Ab sofort werden wir wegen der verträglichen Übertragung der Ringeigenschaft von Polynomkoeffizienten und Unbekannten auf die Polynome an sich wieder auf die speziellen Operatoren  $\oplus$  und  $\odot$  verzichten.
- Man beachte, dass der Polynomring *nicht* zu einem Körper wird, falls die Koeffizienten und Unbekannten aus einem Körper stammen! Denn dann müsste  $\mathbb{K}[X]$  für jedes Polynom ein multiplikatives Inverses besitzen, sodass das Produkt der beiden Polynome  $(1)$  ergibt – ein Polynom vom Grad 0.

Aber schon das einfachste Polynom vom Grad 1, nämlich  $p(X) = X$ , lässt sich nicht durch Multiplikation im Grad reduzieren. Jedes Polynom  $q(X) \in \mathbb{K}[X]$  ergibt im Produkt mit  $p(X)$  entweder  $(0)$  (falls  $q(X) = (0)$ ), oder ein Polynom vom Grad mindestens 1. Da 1 kein Nullteiler ist, gibt es keine Chance, diesen Grad über ein Nullprodukt zu reduzieren.

(Analog dazu gibt es außer  $\pm 1$  keine ganzen Zahlen, die mit irgendeiner ganzen Zahl ein Produkt mit dem Wert 1 bilden.)

#### 5.4.4 Polynomdivision

Die Polynome bilden wie die ganzen Zahlen einen kommutativen Ring mit Eins. In  $\mathbb{Z}$  lassen sich zwei Zahlen dividieren, falls die erste Zahl Vielfaches der zweiten ist. Für sämtliche ganzen Zahlen ist aber die Division mit Rest nach Satz 2.7 möglich, solange der Modul aus  $\mathbb{N}$  stammt und insbesondere nicht 0 ist.

Für Polynome  $\mathbb{K}[X]$  lässt sich (gewisse Eigenschaften des Rings  $\mathbb{K}$  voraus gesetzt) ebenfalls eine Division mit Rest vereinbaren – das werden wir nun motivieren. Wir betrachten hier allerdings für  $\mathbb{K}$  ausschließlich *Körper*; für diese sind die nötigen Eigenschaften erfüllt.

Die Vielfachen eines Polynoms  $m(X) \neq (0)$  sind sämtliche Polynome  $q(X) \cdot m(X)$  mit beliebigem  $q(X) \in \mathbb{K}[X]$ . Alle anderen Polynome aus  $\mathbb{K}[X]$  – also die, die sich nicht als ein solches Produkt  $q \cdot m$  schreiben lassen – sollten bei der Division durch  $m$  einen nichtverschwindenden Rest besitzen. Wir motivieren nun, wie sich dieser Rest (auch der Rest ist dann ein Polynom!) ähnlich zum Satz 2.7 eindeutig bestimmen lässt.

Für gegebenes  $p$  und  $m \neq (0)$  suchen wir also  $q, r$ , sodass diese eindeutig bestimmt sind und folgender Gleichung genügen:

$$p(X) = q(X) \cdot m(X) + r(X) \quad (*)$$

In Satz 2.7 hatten wir für die Eindeutigkeit verlangt, dass sich der Rest zwischen 0 und dem Vorgänger des Moduls befinden soll – wir sehen gleich, dass wir hier etwas ähnliches fordern müssen. Allerdings gibt es für Polynome an sich zunächst keine Größenabschätzung.

Wir kommen aber mit der Betrachtung der Grade weiter. Und zwar gilt in Körpern (die nullteilerfrei sind) und für  $q, m \neq 0$ , dass die Beziehung aus Satz 5.22 eine Gleichheit ist, und dass also

$$\deg(q \cdot m) = \deg(q) + \deg(m)$$

Weiterhin gilt nach Definition 5.17, dass die beiden Seiten der Gleichung (\*) das gleiche Polynom sein müssen und also insbesondere gleichen Grad haben.  $q, r$  müssen also entsprechend gewählt werden. Wir betrachten zwei Fälle für die möglichen Grade von  $p, m$  von denen einer uns die erwähnte Einschränkung für die Eindeutigkeit liefert.

Wir vereinbaren für die Grade die Bezeichner

$$a := \deg(p), \quad b := \deg(q), \quad c := \deg(m), \quad d := \deg(r)$$

Für den Fall  $a < c$  hat das Polynom  $q \cdot m$ , solange  $q \neq 0$  erfüllt ist, schon einen Grad

$$b + c > b + a \geq a,$$

denn das Produkt  $q_b \cdot m_c$  ist ungleich null: Wegen  $q, m \neq (0)$  und der Nullteilerfreiheit des Körpers  $\mathbb{K}$  erhalten wir somit ein Produktpolynom vom Grad  $(b + c)$ .

Potenzen größer als  $a$  im Produkt  $q \cdot m$  müssen hier ausgeglichen werden, damit die rechte Seite von (\*) insgesamt den Grad  $a$  bekommt – das geschieht durch Subtraktion mit den Komponenten aus  $r$ . Allerdings ist dies *nicht eindeutig*, sondern auf beliebig viele Weisen möglich. Für die gewünschte Eindeutigkeit von  $q, r$  müssen wir also solche Beliebigkeiten stets verhindern; das gelingt, indem wir den Grad von  $r$  so einschränken, dass das Restpolynom nie zum Kompensieren höherer Potenzen aus dem Produkt  $q \cdot m$  taugt. Da das Produkt  $q \cdot m$  mindestens Grad  $c$  hat, fordern wir also generell, dass  $d < c$  gelten soll, bzw.:

$$\deg(r) < \deg(m)$$

Dann bleibt aber für unseren betrachteten Fall nur die Möglichkeit  $q := (0)$ , und damit  $q \cdot m = (0)$ , und  $r := p$ . Das kennen wir von ganzen Zahlen auch schon: Die Zahlen zwischen 0 (inklusive) und dem Modul (exklusiv) sind die möglichen Divisionsreste; jeder von ihnen behält bei Division mit Rest den ursprünglichen Wert bei.

Für den Fall  $a < c$  lässt sich also mit obiger Einschränkung die Polynomdivision ohne Rechnung ausführen.

**Beispiel:** Für die Polynome  $p(X) = 3X^2 - 5$  und  $m(X) = 12X^4 - 6X^3 + 4X^2 + 14X$  aus  $\mathbb{R}[X]$  ist

$$p(X) = (0) \cdot m(X) + r(X) \quad \text{mit} \quad r(X) = p(X) = 3X^2 - 5$$

Der (arithmetisch interessantere) Fall  $a \geq c$  erfordert, dass das Quotientenpolynom  $q$  mindestens Grad 0 hat (bei  $a > c$  sogar mindestens Grad 1) und nicht dem Nullpolynom  $(0)$  entsprechen kann. Denn mit der Einschränkung  $\deg(r) < c$  reicht der Grad des Restpolynoms nicht aus, um auf der rechten Seite von (\*) ein Polynom mit Grad  $a$  zu erreichen.

Wir schreiben in Tupelnotation an:

$$(p_a \ p_{a-1} \ \cdots \ p_0) = (q_b \ \cdots \ q_0) \cdot (m_c \ \cdots \ m_0) + (r_d \ \cdots \ r_0)$$

Nun darf das Produktpolynom  $q \cdot m$  auch keinen höheren Grad als  $a$  besitzen (sonst lässt sich die Gleichheit in (\*) wiederum nicht herstellen); es bleibt also nur, dass es den gleichen Grad wie  $p$  hat. Für den Fall  $a > c$  gilt also, dass  $a = b + c$ , bzw.:

$$\deg(p) = \deg(q) + \deg(m)$$



Für das weitere Vorgehen erinnern wir uns, dass im Produktpolynom  $q \cdot m$  der höchste Term eindeutig gegeben ist als (hier dann:)

$$(q_b \cdot m_c)X^a$$

Niedrigere Terme haben dagegen meist mehrere Beiträge (siehe die Beispiele oben). Wegen Definition 5.17 müssen aber nun die Koeffizienten der Polynome auf beiden Seiten der Gleichung (\*) einander entsprechen. Es gilt also:

$$p_a = q_b \cdot m_c$$

(Das Produkt ist hier auf dem Körper  $\mathbb{K}$  auszuführen, kein Polynomprodukt.) Da  $p_a$  und  $m_c$  bereits gegeben waren, ist hiermit  $q_b$  eindeutig bestimmt!

Bevor wir dies formal weiter fortsetzen, zunächst ein Zahlenbeispiel. Wir betrachten die beiden Polynome aus  $\mathbb{R}[X]$  von oben mit vertauschten Rollen, also

$$p(X) = 12X^4 - 6X^3 + 4X^2 + 14X \quad \text{und} \quad m(X) = 3X^2 - 5$$

Wir können nun bereits voraus sehen, dass  $q$  ein Polynom vom Grad 2 sein wird (die Differenz von  $a = 4$  und  $c = 2$ ). Weiterhin wird der Grad des Restpolynoms höchstens 1 betragen, da  $d < c$ . Also gilt:

$$\underbrace{12X^4 - 6X^3 + 4X^2 + 14X}_{p(X)} = \underbrace{(q_2X^2 + q_1X + q_0)}_{q(X)} \cdot \underbrace{(3X^2 - 5)}_{m(X)} + \underbrace{r_1X + r_0}_{r(X)} \quad (**)$$

Wie eben beschrieben ist der höchste Beitrag von  $q \cdot m$  ist  $(q_2 \cdot 3)X^4$ , und durch *Koeffizientenvergleich* mit  $p$  muss gelten:

$$12 = q_2 \cdot 3 \Leftrightarrow q_2 = \frac{12}{3} = 4$$

Wir verwenden nun das Distributivgesetz von  $\mathbb{K}[X]$  und schreiben (hier:)

$$q(X) \cdot m(X) = (q_2X^2 + (q_1X + q_0)) \cdot m(X) = q_2X^2 \cdot m(X) + (q_1X + q_0) \cdot m(X)$$

Bei dieser Aufspaltung hat das rechte Produkt einen Grad kleiner als 4, das linke hat Grad 4.

Nun subtrahieren wir von beiden Seiten in (\*\*) jeweils dieses linke Produkt vom Grad 4 und erhalten:

$$12X^4 - 6X^3 + 4X^2 + 14X - q_2X^2 \cdot (3X^2 - 5) = (q_1X + q_0) \cdot (3X^2 - 5) + r_1X + r_0$$

Wir setzen links noch  $q_2 = 4$  ein und erhalten links damit:

$$\begin{aligned} 12X^4 - 6X^3 + 4X^2 + 14X - 4X^2 \cdot (3X^2 - 5) &= 12X^4 - 6X^3 + 4X^2 + 14X - (12X^4 - 20X^2) \\ &= -6X^3 + 24X^2 + 14X \end{aligned}$$

Insgesamt gilt nun mit  $\tilde{p}(X) := p(X) - q_2 \cdot m(X)$  und  $\tilde{q}(X) := q(X) - q_2X^2$ :

$$\underbrace{-6X^3 + 24X^2 + 14X}_{\tilde{p}(X)} = \underbrace{(q_1X + q_0)}_{\tilde{q}(X)} \cdot \underbrace{(3X^2 - 5)}_{m(X)} + \underbrace{r_1X + r_0}_{r(X)} \quad (***)$$

Wir haben hier nun eine Polynomgleichung vom Grad 3, mit selbem  $m$  und  $r$ , für die wieder der Grad des linken Polynoms mindestens dem Grad von  $m$  entspricht. Wir können also mit dieser neuen Gleichung analog verfahren!

Bleibt die Frage, warum wir nicht nur  $(q_2 \cdot 3)X^4$  subtrahiert haben, sondern  $q_2X^2 \cdot m(X)$ : So verschwindet auf der linken Seite von (\*\*) die höchste Potenz, und das neue Quotientenpolynom vereinfacht sich derart, dass seine Koeffizienten  $q_1, q_0$  unverändert bleiben, also keine zusätzlichen Beiträge erhalten. Da diese Koeffizienten weiterhin gesucht sind, ist das zielführend.

Wir erkaufen uns dies damit, dass sich auf der linken Seite weitere Terme geändert haben (hier  $4X^2$  in  $24X^2$ ). In dieser Form ist unser Vorgehen aber analog zur Division mit Rest auf  $\mathbb{Z}$ , denn dort werden auch Vielfache des ganzen Moduls subtrahiert.

Nun wieder ad rem: Der Höchste Beitrag von  $\tilde{q} \cdot m$  ist  $(q_1 \cdot 3)X^3$  und führt so zu dem Koeffizientenvergleich

$$-6 = q_1 \cdot 3 \Leftrightarrow q_1 = \frac{-6}{3} = -2$$

Damit haben wir einen weiteren Koeffizienten von  $q$  gewonnen.

Für die neue linke Seite von (\*\*\*) errechnen wir also:

$$-6X^3 + 24X^2 + 14X - (-2)X \cdot (3X^2 - 5) = -6X^3 + 24X^2 + 14X - (-6X^3 + 10X) = 24X^2 + 4X$$

Wir subtrahieren wieder das passende Polynom, nämlich  $q_1 \cdot m(X)$  und erhalten mit  $\tilde{p}(X) := \tilde{p}(X) - q_1 \cdot m(X)$  sowie  $\tilde{q}(X) := \tilde{q}(X) - q_1 X$ :

$$\underbrace{24X^2 + 4X}_{\tilde{p}(X)} = \underbrace{(q_0)}_{\tilde{q}(X)} \cdot \underbrace{(3X^2 - 5)}_{m(X)} + \underbrace{r_1 X + r_0}_{r(X)} \quad (****)$$

Auch hier entspricht der Grad des Polynoms  $\tilde{p}$  (gerade noch) mindestens dem von  $m$ , sodass wir ein letztes Mal das obige Vorgehen ausführen können – dies liefert uns  $q_0$ . Der Koeffizientenvergleich ergibt:

$$24 = q_0 \cdot 3 \Leftrightarrow q_0 = \frac{24}{3} = 8$$

Damit errechnen wir die neue linke Seite von (\*\*\*\*):

$$24X^2 + 4X - 8 \cdot (3X^2 - 5) = 24X^2 + 4X - (24X^2 - 40) = 4X + 40$$

Und damit erhalten wir:

$$\underbrace{4X + 40}_{\tilde{\tilde{p}}(X)} = \underbrace{r_1 X + r_0}_{r(X)} \quad (*****)$$

Auf der rechten Seite ist der Term mit  $m$  verschwunden, und die linke Seite hat nun einen Grad kleiner als der des Moduls  $m(X)$ . Falls wir diese linke Seite also modulo  $m(X)$  rechnen würden, käme diese unverändert als Divisionsrest heraus (s.o.). Da hier aber auf der rechten Seite das Polynomprodukt mit  $m(X)$  ohnehin bereits verschwunden ist lassen sich sofort die beiden Koeffizienten des Restpolynoms ablesen:

$$r_1 = 4 \quad \text{und} \quad r_0 = 40$$

Damit haben wir alle unbekannten Koeffizienten berechnet, und es gilt insgesamt:

$$\underbrace{12X^4 - 6X^3 + 4X^2 + 14X}_{p(X)} = \underbrace{(4X^2 - 2X + 8)}_{q(X)} \cdot \underbrace{(3X^2 - 5)}_{m(X)} + \underbrace{4X + 40}_{r(X)}$$

Somit sind Quotientenpolynom  $q(X)$  und Restpolynom  $r(X)$  eindeutig gefunden.

Wir rechnen zur Probe (in Tupelschreibweise) nach, dass das Produkt  $q \cdot m$  und die nachträgliche Addition von  $r$  tatsächlich  $p$  ergibt:

$$\begin{array}{rrrrr} (4 & -2 & 8) & \cdot & (3 & 0 & -5) \\ & & & & -20 & 10 & -40 \\ & 12 & -6 & 24 & & & \\ \hline & (12 & -6 & 4 & 10 & -40) \\ & + & & & (4 & 40) \\ \hline & (12 & -6 & 4 & 14 & 0) & \checkmark \end{array}$$

Zur Übung führe man die Probe auch in Notation mit Potenzen von  $X$  aus.

Nun zurück zum allgemeinen Fall, also zu

$$(p_a \ p_{a-1} \ \cdots \ p_0) = (q_b \ \cdots \ q_0) \cdot (m_c \ \cdots \ m_0) + (r_d \ \cdots \ r_0)$$

Die Subtraktion von  $q_b \cdot m(X)$  mit der Bedingung  $q_b = p_a/m_c$  ergibt dann:

$$(\tilde{p}_{a-1} \ \cdots \ \tilde{p}_0) = (q_{b-1} \ \cdots \ q_0) \cdot (m_c \ \cdots \ m_0) + (r_d \ \cdots \ r_0)$$

Wie im obigen Beispiel motiviert, führen wir dies solange analog aus, bis  $q_0$  bestimmt ist. Dann bleibt auf der linken Seite ein Polynom mit geringerem Grad als  $c$ ; rechts ist dagegen das Produktpolynom mit  $m(X)$  verschwunden, und der Koeffizientenvergleich ergibt direkt das Restpolynom  $r(X)$ .

Wir fassen dies zusammen im

**Satz 5.24** (Polynomdivision). *Über dem Körper  $\mathbb{K}$  sei ein Polynom  $m(X) \in \mathbb{K}[X] \setminus \{(0)\}$  als Modul gegeben. Für jedes Polynom  $p(X) \in \mathbb{K}[X]$  gibt es nun Polynome  $q(X), r(X) \in \mathbb{K}[X]$  mit*

$$\deg(r) < \deg(m),$$

sodass

$$p(X) = q(X) \cdot m(X) + r(X)$$

Dabei sind die Polynome  $q(X), r(X)$  eindeutig bestimmt.

### Bemerkungen:

- Für den oben schon erwähnten Fall dass  $\deg(p) < \deg(m)$  gilt  $q(X) = (0)$  und  $r(X) = p(X)$ .
- Ansonsten lässt sich  $q$  bestimmen, indem sukzessive die höchsten Terme von linker und rechter Seite der Gleichung verglichen werden; danach wird das Polynom  $q_{\bullet} \cdot m(X)$  von beiden Seiten subtrahiert, was auf eine Gleichung niedrigeren Grades führt. Nachdem sämtliche Koeffizienten von  $q$  so bestimmt wurden, liefert der Koeffizientenvergleich das Restpolynom  $r(X)$ .
- In einer Rechnung wird die Division üblicherweise angeschrieben wie eine schriftliche Division. Für das obige Beispiel ergibt sich:

$$\begin{array}{r}
 \begin{array}{rrrr}
 12X^4 & -6X^3 & +4X^2 & +14X \\
 -(12X^4 & -20X^2) & & \\
 \hline
 & -6X^3 & +24X^2 & +14X \\
 & -(-6X^3 & +10X) & \\
 \hline
 & & 24X^2 & +4X \\
 & & -(24X^2 & -40) \\
 \hline
 & & 4X & +40
 \end{array}
 & = &
 \begin{pmatrix} 3X^2 & -5 \\ & + \end{pmatrix}
 \cdot
 \begin{pmatrix} 4X^2 & -2X & +8 \\ 4X & +40 \end{pmatrix}
 \end{array}$$

Es empfiehlt sich, nach jeder Subtraktion links das gesamte verbleibende Polynom anzuschreiben und von dort aus weiter zu rechnen. Anders als bei der schriftlichen Division muss nicht für jede Potenz von  $X$  eine separate Spalte vorgehalten werden. Aber es wäre unübersichtlich, wenn das Polynom des nächsten Schritts zu Teilen aus Termen unterhalb des Subtraktionsstrichs und zu Teilen aus ursprünglichen Termen von oben bestünde.

Man achte sauber darauf, bei der Subtraktion das Minuszeichen vor der Klammer auf alle Terme im Inneren anzuwenden, und danach korrekt mit den Termen vom vorigen Schritt zu bilanzieren. Dabei sind Fehler schnell passiert – es empfiehlt sich bei Abgaben oder in der Prüfung also eine Probe!

Hinter dem Gleichheitszeichen wurde hier zuerst der (bereits bekannte) Modul  $m(X)$  notiert. Darauf folgt der Faktor  $q(X)$ , dessen Koeffizienten erst während der Rechnung ermittelt werden; die zugehörige Klammer ist also zu Beginn noch leer, genau wie jene für das Restpolynom, das hier aus Platzgründen in der zweiten Zeile steht.

- Oder man benutzt die Tupelnotation. Hier sind separate Spalten für alle Potenzen kein Problem, da ohnehin alle Potenzen in der Notation mitzunehmen sind. Das Holen der nächsten Terme von oben geschieht hier systematisch Schritt für Schritt. Dadurch braucht man nicht das ganze Zwischenpolynom zu notieren:

$$\begin{array}{r}
 \begin{array}{rrrrrr}
 (12 & -6 & 4 & 14 & 0) \\
 -(12 & 0 & -20) \\
 \hline
 & (-6 & 24 & 14) \\
 & -(-6 & 0 & 10) \\
 \hline
 & & (24 & 4 & 0) \\
 & & -(24 & 0 & -40) \\
 \hline
 & & & (4 & 40)
 \end{array}
 & = &
 \begin{pmatrix} 3 & 0 & -5 \\ & + & (4 & 40) \end{pmatrix}
 \cdot
 \begin{pmatrix} 4 & -2 & 8 \end{pmatrix}
 \end{array}$$

- Mit den Graden  $\deg(p)$  und  $\deg(m)$  lassen sich (Motivation siehe oben) direkt (vor der Rechnung) die Grade von  $q(X)$  und  $r(X)$  ermitteln (wobei für das Restpolynom ein geringerer Grad noch möglich ist, siehe Beispiele unten).
- Falls der Modul  $m(X)$  den Grad 1 besitzt, handelt es sich um eine Polynomdivision durch einen *Linearfaktor*. Dann kann der Rest (wenn nicht  $\deg(p) = 0$  gilt) nur eine Zahl aus  $\mathbb{K}$  sein (ein Polynom vom Grad 0).
- Bei Restklassenkörpern  $\mathbb{Z}_n$  (mit  $n$  prim) sind die Subtraktionen und Multiplikationen modulo  $n$  auszuführen (auch dazu folgen Beispiele).  
Beim Koeffizientenvergleich für  $q(X)$  werden multiplikative Inverse benötigt, d.h. es lohnt sich, die Multiplikationstafel modulo  $n$  parat zu haben.
- Am einfachsten ist hier das Rechnen mit dem Körper  $\mathbb{Z}_2 = \mathbb{F}_2$ . Multiplikation mit 1 entspricht einer Kopie der Bitfolge; Subtraktion und Addition entsprechen beide dem exklusiven Oder, und es genügt in der Tupelnotation, die Bits ohne Abstände direkt hintereinander zu schreiben. Für die Subtraktion kann man den Operator “−” weglassen, nur der Bilanz-Strich ist notwendig.

### Beispiele:

- Wir haben oben bereits ein Beispiel gesehen, für das der Grad des Dividenden höher war als der des Moduls. Hier betrachten wir noch ein Beispiel mit gleichen Graden auf  $\mathbb{R}[X]$ :

$$p(X) := 3X^2 - 2X + 5 \quad \text{und} \quad m(X) := X^2 + 7$$

Offenbar wird das Quotientenpolynom  $q(X)$  Grad 0 besitzen, und der Rest  $r(X)$  kann höchstens Grad 1 haben. Rechnung:

$$\begin{array}{r} 3X^2 \quad -2X \quad +5 \\ -(3X^2 \quad +21) \\ \hline \quad -2X \quad -16 \end{array} = (X^2 + 7) \cdot (3) + (-2X - 16)$$

- Falls der Grad des Modulpolynoms  $m(X)$  null ist, liegt nur eine *Skalierung* vor. Der Rest verschwindet dann (die Formel würde sowieso nur einen maximalen Grad von  $(-1)$  zulassen). Für  $p(X)$  wie im vorigen Beispiel und  $m(X) = 7$  erhalten wir in  $\mathbb{R}[X]$ :

$$\underbrace{3X^2 - 2X + 5}_{p(X)} = \underbrace{(7)}_{m(X)} \cdot \underbrace{\left(\frac{3}{7}X^2 - \frac{2}{7}X + \frac{5}{7}\right)}_{q(X)}$$

- Wir dividieren in  $\mathbb{R}[X]$  nochmal das Polynom  $p(X)$  wie eben, aber dieses Mal durch den Linearfaktor  $m(X) := X - 7$ . Als Rest können wir dann nur eine einzelne Zahl erhalten:

$$\begin{array}{r} 3X^2 \quad -2X \quad +5 \\ -(3X^2 \quad -21X) \\ \hline \quad 19X \quad +5 \\ \quad -(19X \quad -133) \\ \hline \quad \quad 138 \end{array} = (X - 7) \cdot (3X + 19) + (138)$$

- Wir dividieren in  $\mathbb{R}[X]$  das Polynom  $p(X) := X^2 - 16$  durch  $m(X) := X + 4$ :

$$\begin{array}{r} X^2 \quad -16 \\ -(X^2 \quad +4X) \\ \hline \quad -4X \quad -16 \\ \quad -(-4X \quad -16) \\ \hline \quad \quad 0 \end{array} = (X + 4) \cdot (X - 4)$$

(Das gleiche Ergebnis hätten wir durch scharfes Hinsehen (dritte binomische Formel!) auch direkt erhalten.) Hier ist der Modul ein Teiler von  $p(X)$ , sodass beim Dividieren kein Rest anfällt.

- Wir berechnen eine Polynomdivision im endlichen Körper  $\mathbb{Z}_5$ . Dazu zunächst die Gruppenoperationen:

+	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

·	1	2	3	4
1	1	2	3	4
2	2	4	1	3
3	3	1	4	2
4	4	3	2	1

Tabelle 5.1: Gruppenoperationen in  $\mathbb{Z}_5$

Nun betrachten wir die Polynome

$$p(X) := X^4 + 3X^3 + X^2 + 2X + 2 \quad \text{und} \quad m(x) := 2X^2 + 3X + 4$$

Das Quotientenpolynom  $q(X)$  hat also Grad 2, und der Rest höchstens Grad 1. Wie oben wollen wir uns bei endlichen Körpern auf nichtnegative eindeutige Repräsentanten für die Koeffizienten beschränken.

Die Koeffizienten des Polynoms  $q$  erfordern jeweils Divisionen, die wir per

$$\frac{a}{b} = a \cdot (b^{-1})$$

durch Multiplikation und Nachschlagen der inversen Elemente ausführen.

Wir erhalten folgende Rechnung:

$$\begin{array}{r}
 \begin{array}{cccccc}
 (1 & 3 & 1 & 2 & 2) & = & (2 & 3 & 4) & \cdot & (3 & 2 & 4) \\
 -(1 & 4 & 2) & & & & + & (2 & 1) \\
 \hline
 (4 & 4 & 2) & & & & & & & & & & \\
 -(4 & 1 & 3) & & & & & & & & & & \\
 \hline
 (3 & 4 & 2) & & & & & & & & & & \\
 -(3 & 2 & 1) & & & & & & & & & & \\
 \hline
 (2 & 1) & & & & & & & & & & & 
 \end{array}
 \end{array}$$

Also:  $q(X) = 3X^2 + 2X + 4$  und  $r(X) = 2X + 1$ .

Zur Erklärung: den ersten Koeffizienten von  $q$  bestimmen wir als

$$q_2 = \frac{p_4}{m_2} = \frac{1}{2} = 1 \cdot 2^{-1} = 1 \cdot 3 = 3$$

Damit ermitteln wir das Polynom, das links zu subtrahieren ist; es handelt sich um

$$q_2 X^2 \cdot m(X) = 3X^2 \cdot (2 \ 3 \ 4) = X^2 \cdot (1 \ 4 \ 2) = (1 \ 4 \ 2 \ 0 \ 0),$$

denn  $3 \cdot 2 = 6 \equiv 1$ ,  $3 \cdot 3 = 9 \equiv 4$  und  $3 \cdot 4 = 12 \equiv 2$ . Die zwei angehängten Nullen durch Multiplikation mit  $X^2$  werden im Divisionsschema üblicherweise nicht mitnotiert, da die Subtraktion einer Null stets neutral ist. Hier würde man also den Faktor  $X^2$  stillschweigend ignorieren und ihn durch korrektes Positionieren des Tripels ausdrücken – wie wir das übrigens weiter oben schon entsprechend getan haben. Beim Faktor  $X^2$  entspricht dies zwei Shifts des Tripels nach links.

Die Subtraktion von  $(1 \ 4 \ 2)$  entspricht einer Addition von  $(4 \ 1 \ 3)$ . Dadurch ergibt sich mit der gleichen Einrückung ein Tripel  $(0 \ 4 \ 4)$ . Die führende Null kommt durch die richtige Wahl von  $q_2$ . Im Divisionsschema ist an dieser Stelle (dritte Zeile) bereits die 2 vom Term  $2X$  des Polynoms  $p$  nach unten geholt, um wieder ein Tripel zu erhalten, das mit dem Koeffiziententripel von  $m$  verrechnet werden kann – dieses Mal, um  $q_1 = 2$  zu erhalten.

Das Divisionsschema ist komplett, wenn nach der letzten Subtraktion ein (nicht nach links geschiftetes) Tupel entsteht, das kürzer ist als  $m$  – was hier mit einer Länge von 2 (also dem Polynomgrad 1) der Fall ist.

Also:

$$X^4 + 3X^3 + X^2 + 2X + 2 = (2X^2 + 3X + 4) \cdot (3X^2 + 2X + 4) + (2X + 1)$$

Wir führen noch die Probe aus, indem wir den Ausdruck auf der rechten Seite in Tupel-schreibweise nachrechnen:

$$\begin{array}{r}
 (2 \quad 3 \quad 4) \cdot (3 \quad 2 \quad 4) \\
 \hline
 \phantom{(2 \quad 3 \quad 4)} \phantom{\cdot} \phantom{(3 \quad 2 \quad 4)} 3 \quad 2 \quad 1 \\
 \phantom{(2 \quad 3 \quad 4)} \phantom{\cdot} \phantom{(3 \quad 2 \quad 4)} 4 \quad 1 \quad 3 \\
 \phantom{(2 \quad 3 \quad 4)} \phantom{\cdot} \phantom{(3 \quad 2 \quad 4)} 1 \quad 4 \quad 2 \\
 \hline
 (1 \quad 3 \quad 1 \quad 0 \quad 1) \\
 + \phantom{(1 \quad 3 \quad 1 \quad 0 \quad 1)} \phantom{(2 \quad 1)} (2 \quad 1) \\
 \hline
 (1 \quad 3 \quad 1 \quad 2 \quad 2) \quad \checkmark
 \end{array}$$

- In  $\mathbb{Z}_2 = \mathbb{F}_2$  berechnen wir eine Polynomdivision mit Potenzen aus

$$p(X) := X^7 + X^4 + X^3 + X + 1 \quad \text{und} \quad m(X) := X^2 + X$$

Wir erinnern uns daran, dass in  $\mathbb{F}_2$  Addition und Subtraktion gleichwertig sind. Wir deuten die Subtraktionen hier nur durch die Bilanzierungsstriche an:

$$\begin{array}{r}
 \begin{array}{r}
 X^7 + X^4 + X^3 + X + 1 \\
 X^7 + X^6 \\
 \hline
 X^6 + X^4 + X^3 + X + 1 \\
 X^6 + X^5 \\
 \hline
 X^5 + X^4 + X^3 + X + 1 \\
 X^5 + X^4 \\
 \hline
 X^3 + X + 1 \\
 X^3 + X^2 \\
 \hline
 X^2 + X + 1 \\
 X^2 + X \\
 \hline
 1
 \end{array}
 = (X^2 + X) \cdot (X^5 + X^4 + X^3 + X + 1) + 1
 \end{array}$$

- Und in  $\mathbb{F}_2$  das gleiche Beispiel nochmal mit reduzierter Tupelnotation (nur Binär-Strings):

$$\begin{array}{r}
 10011011 = 110 \cdot 111011 + 1 \\
 110 \\
 \hline
 101 \\
 110 \\
 \hline
 111 \\
 110 \\
 \hline
 101 \\
 110 \\
 \hline
 111 \\
 110 \\
 \hline
 1
 \end{array}$$

Man beachte, dass nach der dritten Subtraktion in  $q(X)$  eine Null entsteht, weil nicht ein, sondern zwei Bits nach unten geholt werden müssen. Als Zwischenschritt hätte man natürlich hier auch zunächst nur das Teilpolynom 10 betrachten können, von dem nur 0 subtrahierbar ist (das gäbe die 0 in  $q(X)$ ), um dann im nächsten Schritt die 1 nach unten zu holen.

Zur Übung rechne man gerne die Probe nach (eine Umkehrung der Faktoren, also die Berechnung von  $111011 \cdot 110$ , spart etwas Platz beim Multiplizieren, führt aber auf das gleiche Ergebnis).

---

Für endliche Körper ist, da die Koeffizienten jeweils nur endlich viele (nämlich  $|\mathbb{K}|$  viele) Werte annehmen können, auch die Zahl der Polynome mit beschränktem Grad endlich:

**Beispiel:** Für  $\mathbb{Z}_3$  gibt es drei verschiedene Werte für Koeffizienten. Wir listen alle Polynome bis zum Grad 2 (exklusiv):

$$0, \quad 1, \quad 2, \quad X, \quad X+1, \quad X+2, \quad 2X, \quad 2X+1, \quad 2X+2$$

Oder in Tupelnotation, hier zielorientiert mit führenden Nullen angeschrieben:

$$(0\ 0), \quad (0\ 1), \quad (0\ 2), \quad (1\ 0), \quad (1\ 1), \quad (1\ 2), \quad (2\ 0), \quad (2\ 1), \quad (2\ 2)$$

Man kann in der Tupelnotation erkennen, wie sukzessive alle  $3^2$  Kombinationen aufgelistet werden (zur Herleitung dieser Zahl siehe die Bemerkung bei Definition 1.24; hier enthält die Menge  $\mathbb{Z}_3$  drei Elemente; es gibt also  $3^2 = 9$  verschiedene 2-Tupel ("Paare") in  $\mathbb{Z}_3 \times \mathbb{Z}_3$ ).

Generell sind führende Nullen in der Tupelnotation nicht zulässig, da die Tupel-Länge über den Grad des jeweiligen Polynoms fest liegt (sie beträgt eins mehr als der Grad). Korrekt lautet die Liste also:

$$(0), \quad (1), \quad (2), \quad (1\ 0), \quad (1\ 1), \quad (1\ 2), \quad (2\ 0), \quad (2\ 1), \quad (2\ 2)$$

Es handelt sich um sämtliche Polynome vom Grad kleiner als 2. Man erkennt hier auch, dass es genau  $3^1$ , also drei, Polynome vom Grad kleiner als 1 gibt.

Allgemein gilt dann folgender

**Satz 5.25** (Anzahl der Polynome über endlichen Körpern). *Für einen endlichen Körper  $\mathbb{K}$  mit  $k := |\mathbb{K}|$  beträgt die Anzahl der verschiedenen Polynome vom Grad kleiner als  $n \in \mathbb{N}$  genau  $k^n$ .*

**Beispiele:**

- Es gäbe also über  $\mathbb{Z}_3$  genau ( $3^3 = 27$ ) verschiedene Polynome, die höchstens quadratisch sind (neun davon, nämlich die ohne quadratischen Term, sind oben schon gelistet; man finde als Übung (und möglichst systematisch) die anderen achtzehn Polynome mit quadratischen Termen). Analog gibt es ( $3^4 = 81$ ) verschiedene Polynome über  $\mathbb{Z}_3$ , die höchstens kubisch sind, also Grad kleiner als 4 besitzen.
- Für  $p$  prim gibt es  $p^k$  verschiedene Polynome vom Grad kleiner als  $k$  über  $\mathbb{Z}_p$ . Das spielt (kein Vorlesungsstoff!) eine entscheidende Rolle bei der Konstruktion endlicher Körper mit  $p^k$  Elementen: Diese Elemente entsprechen nämlich gerade diesen Polynomen.

### 5.4.5 Nullstellen

**Definition 5.26** (Nullstelle). *Sei  $\mathbb{K}$  ein Körper und  $p(X) \in \mathbb{K}[X]$  ein Polynom. Dann heißt eine Zahl  $\tilde{x} \in \mathbb{K}$  Nullstelle von  $p$ , falls die zugehörige Polynomfunktion  $f_p : \mathbb{K} \rightarrow \mathbb{K}$  dort verschwindet, d.h. falls*

$$f_p(\tilde{x}) = 0$$

**Beispiele:**

- Jedes lineare Polynom  $p(X) = aX + b$  mit  $a \neq 0$  besitzt die Nullstelle

$$\tilde{x} = -\frac{b}{a} = -b \cdot (a^{-1})$$

Denn mit  $f_p(x) = ax + b$  ist:

$$f_p(\tilde{x}) = a \cdot \left(-\frac{b}{a}\right) + b = -b + b = 0 \quad \checkmark$$

Tatsächlich ist dies sogar schon jeweils die einzige Nullstelle.

- Das konstante Polynom  $p(X) = a$  mit  $a \neq 0$  hat dagegen keine Nullstelle: Die Polynomfunktion  $f_p$  würde jedes  $x \in \mathbb{K}$  auf den Wert  $a$  abbilden, sodass die Bedingung aus der Definition auf ganz  $\mathbb{K}$  unerfüllbar ist.
- Die Zahl  $\tilde{x} = 7$  ist Nullstelle des reellen Polynoms

$$p(X) = 4X^5 - 30X^4 + 17X^3 - 20X^2 - 2X - 35$$

Zur Auswertung der Polynomfunktion  $f_p$  an der Stelle 7 hilft z.B. das Horner Schema aus Exkurs A.2.

---

Wie findet man nun die Nullstellen eines Polynoms? Es stellt sich heraus, dass die Gleichung  $f_p(\tilde{x}) = 0$  schon für niedrige Polynomgrade schwierig zu lösen ist. Für den linearen Fall hatten wir oben schon die Lösung angegeben. Quadratische Polynome haben in  $\mathbb{R}$  bis zu zwei Nullstellen, die sich als Lösung der zugehörigen quadratischen Gleichung berechnen lassen (siehe Satz 1.34).

Schon hier fällt allerdings auf, dass es nicht für sämtliche reellen quadratischen Polynome zwei Nullstellen gibt. Das gilt in dieser Form auch für alle höheren Grade!

Weiterhin wurde im Rahmen der *Galoistheorie* (kein Vorlesungsstoff!) gezeigt, dass Polynomgleichungen nur bis einschließlich zum Grad 4 geschlossen lösbar sind, wenn nur die Körperoperationen und das Ziehen ganzzahliger Wurzeln erlaubt sind.

---

Man findet allerdings (kein Vorlesungsstoff), dass mit einer Erweiterung von reellen auf *komplexe Zahlen*<sup>13</sup>  $\mathbb{C} \supset \mathbb{R}$  sämtliche Polynomgleichungen lösbar werden. Man nennt daher die komplexen Zahlen *algebraisch abgeschlossen*.

Da sich jede Polynomgleichung durch Umstellen der Terme auf die Form  $p(X) = 0$  bringen lässt, hängt die Frage der Lösbarkeit identisch mit der Frage nach der Existenz und Anzahl von Nullstellen zusammen; dazu weitere Bemerkungen s.u.

---

Für die weiteren Überlegungen zunächst folgende

**Definition 5.27** (Linearfaktor). *Für einen Körper  $\mathbb{K}$  und ein  $\tilde{x} \in \mathbb{K}$  heißt das lineare Polynom*

$$X - \tilde{x}$$

*Linearfaktor eines Polynoms  $p(X) \in \mathbb{K}[X]$ , falls es ein Polynom  $q(X) \in \mathbb{K}[X]$  gibt, sodass*

$$p(X) = (X - \tilde{x}) \cdot q(X)$$

**Bemerkung:** Das lineare Polynom ist also ein Teiler (oder eben: Faktor) von  $p(X)$

---

Damit können wir folgenden Satz formulieren:

**Satz 5.28** (Nullstellen und Linearfaktoren). *Für einen Körper  $\mathbb{K}$  ist  $\tilde{x} \in \mathbb{K}$  genau dann Nullstelle eines Polynoms  $p(X) \in \mathbb{K}[X]$ , wenn  $(X - \tilde{x})$  Linearfaktor von  $p(X)$  ist.*

(Beweis: S. 321.)

**Bemerkungen:**

- Findet man also eine Nullstelle  $\tilde{x}$  für ein Polynom  $p(X)$ , so enthält  $p(X)$  den Linearfaktor  $(X - \tilde{x})$ , durch den es ohne Rest dividierbar ist. Für das resultierende Polynom  $q(X)$  mit  $p(X) = (X - \tilde{x}) \cdot q(X)$  gilt nach der Polynomdivision die folgende Grad-Abschätzung:

$$\deg(q) = \deg(p) - 1$$

Die Nullstellen von  $p(X)$  entsprechen dann den Nullstellen von  $q(X)$  und  $\tilde{x}$ , also

$$\{x \in \mathbb{K} \mid f_p(x) = 0\} = \{x \in \mathbb{K} \mid f_q(x) = 0\} \cup \{\tilde{x}\}$$

Damit lässt sich die Nullstellensuche auf ein Polynom mit niedrigerem Grad zurück führen!

---

<sup>13</sup>siehe Mathematik 2/Analysis



- Das Polynom fünften Grades im Beispiel zu Definition 5.26 wurde übrigens gefunden, indem der Linearfaktor  $(X - 7)$  an ein willkürliches Polynom, hier

$$(4 \quad -2 \quad 3 \quad 1 \quad 5),$$

multipliziert wurde. Dieses Polynom erhalten wir zurück, wenn wir die Polynomdivision ausführen:

$$\begin{array}{r}
 (4 \quad -30 \quad 17 \quad -20 \quad -2 \quad -35) = (1 \quad -7) \cdot (4 \quad -2 \quad 3 \quad 1 \quad 5) \\
 \underline{-(4 \quad -28)} \\
 (-2 \quad 17) \\
 \underline{-(-2 \quad 14)} \\
 (3 \quad -20) \\
 \underline{-(3 \quad -21)} \\
 (1 \quad -2) \\
 \underline{-(1 \quad -7)} \\
 (5 \quad -35) \\
 \underline{-(5 \quad -35)} \\
 (0)
 \end{array}$$


---

Weiterhin gilt (ohne Beweis!) auf den komplexen Zahlen  $\mathbb{C}$  der folgende

**Satz 5.29** (Fundamentalsatz der Algebra). *Jedes nichtkonstante Polynom aus  $\mathbb{C}[X]$  besitzt eine Nullstelle.*

#### Bemerkungen:

- Zusammen mit der Beobachtung nach dem Satz 5.28 ergibt sich, dass man in  $\mathbb{C}$  die Nullstellen sukzessive findet, indem man immer weiter Linearfaktoren durch Polynomdivision abspaltet. Insbesondere zerfällt dann jedes Polynom vom Grad  $n \in \mathbb{N}$  in genau  $n$  Linearfaktoren (und möglicherweise einen globalen Faktor  $c \in \mathbb{K}$ ) und besitzt entsprechend  $n$  Nullstellen.
- Eine Nullstelle  $\tilde{x}$  kann auch *mehrfach* sein, falls der Linearfaktor  $(X - \tilde{x})$  mehrfach auftritt. Zum Beispiel hat das quadratische Polynom  $X^2 - 14X + 49 = (X - 7)^2 = (X - 7) \cdot (X - 7)$  die doppelte Nullstelle 7.
- Reelle Polynome vom Grad  $n$  können also höchstens  $n$  reelle Nullstellen besitzen; alle anderen Nullstellen liegen in  $\mathbb{C} \setminus \mathbb{R}$ . Zum Beispiel hat das kubische Polynom

$$p(X) = X^3 - 3X^2 + 2X - 6$$

nur eine reelle Nullstelle bei  $\tilde{x} = 3$ , und es ist  $p(X) = (X - 3) \cdot (X^2 + 2)$ . Nach Abspaltung des Linearfaktors liegt ein Polynom ohne reelle Nullstellen vor, da die Gleichung

$$x^2 + 2 = 0 \quad \Leftrightarrow \quad x^2 = -2$$

in  $\mathbb{R}$  nicht lösbar ist.

**Beispiel:** Wir betrachten das reelle Polynom  $p(X) := 2X^3 - X^2 - 15X + 18$ . Falls wir durch Raten (oder extra scharfes Hinsehen; die Grenzen sind fließend) eine Nullstelle finden, könnten wir diese abspalten und hätten nur noch eine quadratische Gleichung zu lösen. Es lohnt sich im akademischen Kontext oft, die Zahlen 0,  $\pm 1$  und  $\pm 2$  auszuprobieren – für reale Probleme sind dagegen numerische Verfahren sinnvoll.

Hier finden wir, dass 0 keine Nullstelle sein kann (sonst ließe sich der Faktor  $X$  ausklammern!). Auch  $\pm 1$  führen nicht weiter. Aber es ist:

$$f_p(2) = 2 \cdot 8 - 4 - 15 \cdot 2 + 18 = 16 - 4 - 30 + 18 = 0$$

Aber dann können wir den Faktor  $(X - 2)$  abdividieren:

$$\begin{array}{r}
 \begin{array}{cccc} 2 & -1 & -15 & 18 \end{array} & = & \begin{array}{ccc} (1 & -2) \cdot (2 & 3 & -9) \\ -2 & -4 \end{array} \\
 \hline
 \begin{array}{ccc} (3 & -15) \\ -3 & -6 \end{array} \\
 \hline
 \begin{array}{cc} (-9 & 18) \\ -(-9 & 18) \end{array} \\
 \hline
 (0)
 \end{array}$$

Also ist  $p(X) = (X - 2) \cdot (2X^2 + 3X - 9)$ . Für die Nullstellen des quadratischen Polynoms verwenden wir Satz 1.34 und erhalten für die Polynomfunktion mit

$$2x^2 + 3x - 9 = 0 \quad \Leftrightarrow \quad x^2 + \frac{3}{2}x - \frac{9}{2} = 0$$

die beiden Lösungen

$$\begin{aligned}
 x &= -\frac{3}{4} \pm \sqrt{\frac{9}{16} + \frac{9}{2}} = -\frac{3}{4} \pm \sqrt{\frac{9}{16} + \frac{72}{16}} = -\frac{3}{4} \pm \sqrt{\frac{81}{16}} = -\frac{3}{4} \pm \frac{9}{4} \\
 &\Leftrightarrow \left( x = -\frac{12}{4} = -3 \right) \vee \left( x = \frac{6}{4} = \frac{3}{2} \right)
 \end{aligned}$$

Würden wir also das quadratische Polynom  $(2X^2 + 3X - 9)$  durch den Faktor

$$(X - (-3)) = (X + 3)$$

dividieren, so erhielten wir das lineare Polynom  $X - \frac{3}{2}$  als Quotienten (als Übung empfohlen).

Insgesamt erhalten wir:

$$p(X) = (X - 2) \cdot (X + 3) \cdot \left( X - \frac{3}{2} \right)$$

Damit ist das reelle Polynom komplett in Linearfaktoren zerlegt. Zur Übung empfiehlt sich auch die Probe, d.h. die Berechnung des Produkts der drei Linearfaktoren und der Vergleich mit  $p(X)$  wie anfangs gegeben.

# Kapitel 6

## Lineare Algebra: Vektoren

### 6.1 Vektorräume

Wir betrachten zunächst *Ortsvektoren*, mit denen sich Positionen in *Koordinatensystemen* beschreiben und verrechnen lassen. Eine geeignete – wenn auch allgemeinere – algebraische Struktur, mit der dieses Verhalten zusammen gefasst werden kann, ist der *Vektorraum*, der (ähnlich wie, aber doch strukturell anders als die ringartigen Strukturen) eine Menge von Objekten (nämlich die Vektoren) mit einer Addition und einer Multiplikation kombiniert.

Während die Körper (und die ihnen zu Grunde liegenden Strukturen) für die Arithmetik sehr wichtig sind (in Körpern sind alle “Grundrechenarten” erklärt), erlauben Vektorräume darüber hinaus, geometrische (oder physikalische) *Struktur* zu vereinbaren, welche nicht als einfachen Zahlen (den Körperelementen) ausdrückbar wäre. Dies vereinfacht diverse Rechnungen mit Geraden, Kreisen, Dreiecken, etc; siehe dazu auch den Exkurs A.3 im Anhang.

Entscheidend beim Umgang mit Vektoren ist auch die Frage der *linearen Abhängigkeit*, welche wir nach der Einführung des Vektorbegriffs beleuchten.

#### 6.1.1 Koordinatensysteme und Dimensionen

Koordinatensysteme sind Verallgemeinerungen des reellen *Zahlenstrahls* (das ist eine gerichtete Achse, auf der sämtliche reellen Zahlen abgetragen werden können). Legen wir nun durch den Nullpunkt des reellen Zahlenstrahls eine weitere gerichtete Achse, die senkrecht auf ersterem steht, erhalten wir ein *zweidimensionales kartesisches Koordinatensystem*. Dieses kann man sich wie ein ebenes und unendlich ausgedehntes Blatt Kästchenpapier vorstellen<sup>1</sup>.

Sämtliche Positionen von Punkten auf solch einem gedachten Kästchenpapier können, wenn man den Rasterlinien folgt, eindeutig einem Paar reeller Zahlen zugeordnet werden; dies sind die *Koordinaten* der Punkte. Vom Nullpunkt (der im Koordinatensystem *Ursprung* heißt) aus erreicht man einen beliebigen Punkt, indem man zunächst entlang der ersten Achse läuft, bis die zugehörige Koordinate erreicht ist; danach läuft man senkrecht dazu (und parallel zur zweiten Achse), bis auf der zweiten Achse die zweite Koordinate erreicht ist – währenddessen bleibt die erste Koordinate unverändert.

---

Auch Koordinatensysteme höherer Dimension sind möglich. Eine dritte Dimension können wir uns noch bildlich vorstellen – also eine Achse, die senkrecht auf den beiden vorigen Achsen steht. Die Vorstellung des Kästchenpapiers ändert sich dadurch in ein raumfüllendes Gitter. Unsere ganze räumliche Umwelt lässt sich mit drei Dimensionen beschreiben.

Physikalisch oder mathematisch können auch noch mehr Dimensionen hilfreich bzw. nötig sein. In der Relativitätstheorie benötigt man z.B. vier Dimensionen (diese lassen sich mit unserem Gehirn aber nicht mehr vorstellen); in der Teilchenphysik teilweise zehn oder 26; in der Mathematik sind ohnehin keine Grenzen gesetzt – auch unendlichdimensionale Vektorräume gibt es!

---

<sup>1</sup>“unendlich” kann man sich natürlich nicht bildlich vorstellen – für unsere Begriffe reicht zunächst: “beliebig groß”

Koordinatensysteme müssen übrigens nicht immer rechtwinklig sein. Für die Beschreibung von Kristallen werden auch Systeme mit anderen Winkeln als  $\frac{\pi}{2}$  verwendet, die die Symmetrie besser ausdrücken können.

Darüber hinaus gibt es Koordinatensysteme, die zwar rechtwinklig sind, aber nur *lokal* gelten, d.h. die Achsen zeigen an verschiedenen Punkten ggf. in verschiedene Richtungen. Auch dies kann sinnvoll sein, um bestimmte Symmetrien (z.B. sphärisch oder zylindrisch) eleganter abzubilden.

Wir betrachten in der Vorlesung allerdings praktisch nur globale rechtwinklige Koordinatensysteme im Reellen, die schon erwähnten kartesischen Koordinatensysteme.

### 6.1.2 Motivation: Ortsvektoren

**Beispiel:** Wir betrachten (siehe Abbildung 6.1) ein zweidimensionales kartesisches Koordinatensystem; eingezeichnet sind die vier Punkte  $P, Q, R, Q'$

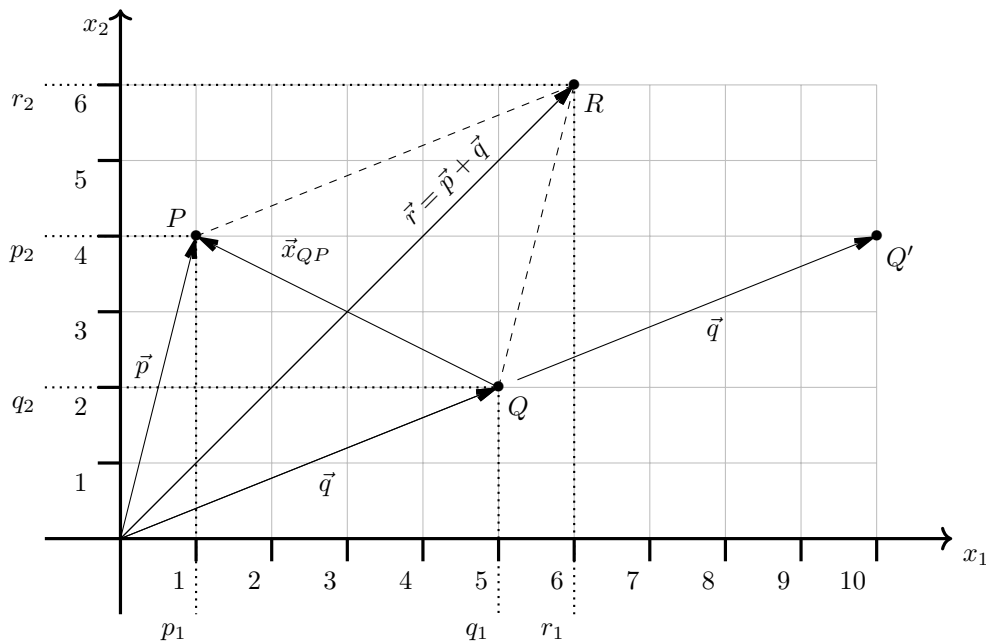


Abbildung 6.1: Kartesisches Koordinatensystem in 2D mit Ortsvektoren

Eingezeichnet sind neben den Punkten noch Pfeile, die vom Ursprung des Koordinatensystems zu den jeweiligen Punkten führen. Diese Pfeile heißen *Ortsvektoren*. Wir wollen Ortsvektoren, wie in den Natur- und Ingenieurwissenschaften üblich<sup>2</sup>, mit darüber notierten Pfeilen kennzeichnen. Dazu folgende Beobachtungen:

- Die Koordinaten der Punkte werden meist als Tupel (hier: Paare) angegeben; sie lauten, inklusive dem *Ursprung*  $O$ :

$$O(0, 0); \quad P(1, 4); \quad Q(5, 2); \quad R(6, 6); \quad Q'(10, 4)$$

Für die Punkte  $P, Q, R$  sind jeweils auch die Koordinaten an den Achsen abgetragen.

- Die Ortsvektoren werden hingegen in Spaltenform notiert. Das ist wichtig, da es (wie wir später noch sehen) auch *Zeilenvektoren* gibt<sup>3</sup>. Die zugehörigen Ortsvektoren lauten also (wobei der Ortsvektor des Ursprungs *Nullvektor* heißt; siehe unten):

$$\vec{0} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}; \quad \vec{p} = \begin{pmatrix} p_1 \\ p_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 4 \end{pmatrix}; \quad \vec{q} = \begin{pmatrix} 5 \\ 2 \end{pmatrix}; \quad \vec{r} = \begin{pmatrix} 6 \\ 6 \end{pmatrix}; \quad \vec{q'} = \begin{pmatrix} 10 \\ 4 \end{pmatrix}$$

Die Komponenten der Ortsvektoren entsprechen den Koordinaten der Punkte (bei  $\vec{p}$  exemplarisch notiert).

<sup>2</sup>In der (theoretischen) Mathematik ist das nicht verbreitet. Manchmal gibt es keine gesonderte Notation. Oder man findet Vektoren fett gedruckt oder in Fraktur gesetzt – beides aber eher unhilfreich beim Schreiben mit der Hand

<sup>3</sup>Mit der mathematischen Tatsache, dass Spaltenvektoren aus dem Vektorraum stammen, Zeilenvektoren jedoch aus dessen zugehörigem *Dualraum*, befassen wir uns in dieser Vorlesung nicht; entscheidender ist für uns in der Anwendung, dass wir mit den Objekten korrekt umgehen.

- Zwei Ortsvektoren lassen sich addieren, indem man den zweiten Pfeil derart parallel verschiebt, dass sein Startpunkt am Endpunkt des ersten Pfeils liegt. In Abbildung 6.1 ist z.B.  $\vec{r}$ , der Ortsvektor des Punkts  $R$ , die Addition von  $\vec{p}$  und  $\vec{q}$ . Die gestrichelten Linien zeigen an, wo die verschobenen Vektoren jeweils liegen würden. Dabei ist es offenbar unwesentlich, ob  $\vec{q}$  an  $\vec{p}$  angehängt wird (obere gestrichelte Linie) oder  $\vec{p}$  an  $\vec{q}$  (rechte gestrichelte Linie); in beiden Fällen wird der gleiche Punkt  $R$  erreicht. Die Addition von Vektoren sollte also auch allgemein kommutativ sein!

Betrachten wir die Koordinaten der Vektoren, so fällt auf, dass die Koordinaten von  $\vec{r}$  jeweils genau der Summe der entsprechenden Koordinaten aus  $\vec{p}$  und  $\vec{q}$  entsprechen:

$$\vec{r} = \vec{p} + \vec{q} = \begin{pmatrix} 1 \\ 4 \end{pmatrix} + \begin{pmatrix} 5 \\ 2 \end{pmatrix} = \begin{pmatrix} 1+5 \\ 4+2 \end{pmatrix} = \begin{pmatrix} 6 \\ 6 \end{pmatrix}$$

Wir sehen bald, dass das kein Zufall ist.

- Auch die Differenz zweier Ortsvektoren kann gebildet werden: Vom Punkt  $Q$  in Abbildung 6.1 gelangt man zu  $P$ , indem man zunächst den Ortsvektor  $\vec{q}$  von  $Q$  *rückwärts* läuft, also von  $Q$  bis zum Ursprung. Danach mit Vektor  $\vec{p}$  nach  $P$ . Insgesamt:  $\vec{x}_{QP} = (-\vec{q}) + \vec{p} = \vec{p} - \vec{q}$ . Das Minuszeichen drückt also aus, dass der Vektor in Gegenrichtung gelesen wird.

Auch hier stellen wir fest, dass die Subtraktion der Vektoren sich komponentenweise berechnen lässt per

$$\vec{x}_{QP} = \vec{p} - \vec{q} = \begin{pmatrix} 1 \\ 4 \end{pmatrix} - \begin{pmatrix} 5 \\ 2 \end{pmatrix} = \begin{pmatrix} 1-5 \\ 4-2 \end{pmatrix} = \begin{pmatrix} -4 \\ 2 \end{pmatrix}$$

Verschieben wir den Pfeil von  $Q$  nach  $P$  an  $\vec{q}$  entlang zum Ursprung, so erhalten wir einen Ortsvektor zum Punkt  $(-4, 2)$ , der vier Einheiten links und zwei Einheiten oberhalb vom Ursprung liegt. So wie im Bild eingezeichnet liegt mit  $\vec{x}_{QP}$  jedoch kein Ortsvektor vor. Wir sehen in der folgenden Definition aber ein, dass auch Differenzvektoren Vektoren sind.

- Wird ein Ortsvektor des Punkts  $Q$  mit dem Faktor 2 *skaliert*, so gelangt man zu einem Punkt  $Q'$ , der, vom Ursprung aus gesehen, in gleicher Richtung liegt, aber mit doppeltem Abstand. Wie oben ersichtlich, haben sich auch die Komponenten des Ortsvektors dabei jeweils genau verdoppelt.

Diesen speziellen Gesamtvektor  $2\vec{q}$  erhielte man auch, wenn man den Vektor  $\vec{q}$  nochmal zu  $\vec{q}$  addieren würde. Also:  $2 \cdot \vec{q} = \vec{q} + \vec{q}$ .

Für die Addition und Skalierung von Vektoren führen wir nun eine algebraische Struktur ein, die sicher stellt, dass sich bei Ortsvektoren das im Beispiel motivierte Verhalten ergibt.

### 6.1.3 Vektorraumbegriff

**Definition 6.1** (Vektorraum). Für eine Menge  $V$ , einen Körper  $(\mathbb{K}, +, \cdot)$ , und mit Operationen

$$\begin{aligned} \oplus & : V \times V \rightarrow V & (\text{Vektoraddition}) \\ \odot & : \mathbb{K} \times V \rightarrow V & (\text{Skalierung bzw. skalare Multiplikation}) \end{aligned}$$

heißt die Struktur  $(V, \mathbb{K}, \oplus, \odot)$  Vektorraum über  $\mathbb{K}$  (oder:  $\mathbb{K}$ -Vektorraum), falls gilt:

- $(V, \oplus)$  ist abelsche Gruppe (deren neutrales Element heißt Nullvektor)
- Mit  $a, b \in \mathbb{K}$  und  $\vec{v}, \vec{w} \in V$  beliebig gilt:

1.  $(a + b) \odot \vec{v} = (a \odot \vec{v}) \oplus (b \odot \vec{v})$
2.  $a \odot (\vec{v} \oplus \vec{w}) = (a \odot \vec{v}) \oplus (a \odot \vec{w})$
3.  $(a \cdot b) \odot \vec{v} = a \odot (b \odot \vec{v})$
4.  $1 \odot \vec{v} = \vec{v}$

Die Elemente von  $\mathbb{K}$  heißen dann Skalare; die Elemente von  $V$  heißen Vektoren.

### Bemerkungen:

- Die Rechenregeln 1 und 2 sind Distributivgesetze, die das Skalieren und die Vektoraddition in “erwarteter” Weise verbinden.
- Die Vorschriften 3 und 4 spezifizieren außerdem das Skalieren von Vektoren.
- Wenn für Vektoren Pfeile zur Notation verwendet werden, heißt der Nullvektor üblicherweise

$$\vec{0}$$

- Wenn vom Kontext her klar ist, welche Rechenoperation auszuführen ist (speziell, ob es sich um eine auf  $\mathbb{K}$  oder eine auf  $V$  handelt), werden meist die Symbole der Körperoperationen auch für den Vektorraum verwendet. Speziell lässt man den Punkt beim Skalieren oft aus, sodass folgendes gleichwertig wäre:

$$c \odot \vec{v} = c \cdot \vec{v} = c\vec{v}$$

Das ist hier problemlos möglich, da  $c$  keinen Vektorpfeil trägt – also muss es sich um einen Skalar aus  $\mathbb{K}$  handeln.

- Jeder Körper  $\mathbb{K}$  ist ein (trivialer) Vektorraum über sich selbst (Beweis: Übung; hierzu wären noch die beiden Operationen  $\oplus$  und  $\odot$  zu definieren) – hier entsteht aber keine Struktur, die sich von einem Körper unterscheidet.
- Wie bei den gruppen- und ringartigen Strukturen wird, sofern der Kontext klar ist, die Menge  $V$  der Vektoren synonym mit dem zugehörigen Vektorraum verwendet.

---

Beispiele folgen in Kürze! Wichtig ist (mit ähnlicher Motivation wie beim Konzept der Untergruppe) auch noch folgende

**Definition 6.2** (Unter(vektor)raum). *Für einen Vektorraum  $(V, \mathbb{K}, \oplus, \odot)$  heißt  $(U, \mathbb{K}, \oplus, \odot)$  Untervektorraum (oder: Unterraum), falls  $\emptyset \neq U \subseteq V$  und falls für alle  $a \in \mathbb{K}$  und  $\vec{v}, \vec{w} \in U$  gilt:*

$$\vec{v} \oplus \vec{w} \in U \quad \text{und} \quad a \odot \vec{v} \in U$$

### Bemerkungen:

- Jeder Vektorraum enthält den trivialen Unterraum  $\{\vec{0}\}$ .
- Jeder Unterraum von  $V$  enthält den Nullvektor  $\vec{0} \in V$ .
- Mit “interessanteren” (weil nicht-trivialen) Untervektorräumen befassen wir uns später noch.

---

Aus den Definitionen des Vektorraums lassen sich noch zwei wichtige zusätzliche Eigenschaften in Bezug auf den Nullvektor ableiten:

**Satz 6.3** (Skalierung und Nullvektor). *Für einen Vektorraum  $(V, \mathbb{K}, \oplus, \odot)$  und mit  $a \in \mathbb{K}$  sowie  $\vec{v} \in V$  beliebig gilt:*

$$0 \odot \vec{v} = \vec{0} = a \odot \vec{0}$$

(Beweis: S. 321.)

## 6.1.4 Kartesische Produkträume

Wie bereits oben im Beispiel gezeigt, besteht ein direkter Zusammenhang zwischen den Koordinaten von Punkten im kartesischen Koordinatensystem und den Komponenten ihrer Ortsvektoren. Die Rechenoperationen für diese Vektoren hatten wir allerdings bisher nur motiviert, nicht formal definiert. Darum:

**Satz 6.4** (Kartesischer Produktraum). Für einen Körper  $(\mathbb{K}, +, \cdot)$ , und  $n \in \mathbb{N}$  beliebig und fest, bilden die Elemente des kartesischen Produkts

$$\mathbb{K}^n = \underbrace{\mathbb{K} \times \mathbb{K} \times \cdots \times \mathbb{K}}_{n\text{-mal}}$$

einen  $\mathbb{K}$ -Vektorraum, wenn die Addition und Skalierung jeweils komponentenweise ausgeführt werden, d.h. für  $a \in \mathbb{K}$  und  $\vec{v}, \vec{w} \in \mathbb{K}^n$ :

$$\vec{v} \oplus \vec{w} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} \oplus \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} := \begin{pmatrix} v_1 + w_1 \\ v_2 + w_2 \\ \vdots \\ v_n + w_n \end{pmatrix} \quad \text{und} \quad a \odot \vec{v} = a \odot \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} := \begin{pmatrix} a \cdot v_1 \\ a \cdot v_2 \\ \vdots \\ a \cdot v_n \end{pmatrix}$$

(Beweis: S. 322.)

#### Bemerkungen:

- Wir übergehen hier die Tatsache, dass  $n$ -Tupel strukturell nicht dasselbe sind wie Spaltenvektoren; beide sind isomorph ineinander überführbar. Weil später der Unterschied zwischen Spalten- und Zeilenvektoren noch eine Rolle spielt, notieren wir, wenn wir  $\mathbb{K}^n$  als Vektorraum begreifen, die Elemente aber stets als Spaltenvektoren.
- Der Nullvektor im Vektorraum  $\mathbb{K}^n$  hat die Darstellung

$$\vec{0} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

- Im Vektorraum  $\mathbb{K}^n$  ist das additive Komplement eines Vektors  $\vec{v}$  gegeben durch

$$-\vec{v} = (-1) \cdot \vec{v} = \begin{pmatrix} -v_1 \\ -v_2 \\ \vdots \\ -v_n \end{pmatrix}$$

- Die Gleichheit zweier solcher Vektoren ist (analog zu den  $n$ -Tupeln) durch Gleichheit der Komponenten erklärt:

$$(\vec{v} = \vec{w}) \Leftrightarrow ((v_1 = w_1) \wedge (v_2 = w_2) \wedge \cdots \wedge (v_n = w_n))$$

- Wir verzichten ab jetzt auf gesonderte Operatoren für die Vektoraddition und -skalierung.

#### Beispiele:

- Offenbar haben wir in Abbildung 6.1 (S. 148) bereits den Vektorraum  $\mathbb{R}^2$  kennen gelernt.
- Analog liegen die Ortsvektoren von Punkten in einem dreidimensionalen kartesischen Koordinatensystem im Vektorraum  $\mathbb{R}^3$ .
- Die Vektoren

$$V_1 := \left\{ \begin{pmatrix} x_1 \\ x_2 \\ 0 \end{pmatrix} \mid x_1, x_2 \in \mathbb{R} \right\} \subseteq \mathbb{R}^3$$

bilden einen (zweidimensionalen<sup>4</sup>) Unterraum von  $\mathbb{R}^3$ , da  $V_1 \subseteq \mathbb{R}^3$  und für  $\vec{v}, \vec{w} \in V_1$  sowie  $a \in \mathbb{R}$  gilt:

$$\vec{v} + \vec{w} = \begin{pmatrix} v_1 \\ v_2 \\ 0 \end{pmatrix} + \begin{pmatrix} w_1 \\ w_2 \\ 0 \end{pmatrix} = \begin{pmatrix} v_1 + w_1 \\ v_2 + w_2 \\ 0 \end{pmatrix} \in V_1 \quad \text{und} \quad a\vec{v} = a \begin{pmatrix} v_1 \\ v_2 \\ 0 \end{pmatrix} = \begin{pmatrix} a \cdot v_1 \\ a \cdot v_2 \\ 0 \end{pmatrix} \in V_1$$

Dieser Unterraum ist eine Einbettung des Raums  $\mathbb{R}^2$  in den Raum  $\mathbb{R}^3$

<sup>4</sup>Zum Begriff der Dimension kommen wir in Kürze

- Die Vektoren

$$V_2 := \left\{ \begin{pmatrix} x_1 \\ x_2 \\ 42 \end{pmatrix} \mid x_1, x_2 \in \mathbb{R} \right\} \subseteq \mathbb{R}^3$$

bilden jedoch *keinen* Unterraum von  $\mathbb{R}^3$ , denn sie enthalten den Nullvektor von  $\mathbb{R}^3$  nicht!

- Man könnte aber mit neuen Operationen auf  $V_2$  einen Vektorraum erklären, wenn gilt:

$$\vec{v} + \vec{w} = \begin{pmatrix} v_1 \\ v_2 \\ 42 \end{pmatrix} + \begin{pmatrix} w_1 \\ w_2 \\ 42 \end{pmatrix} := \begin{pmatrix} v_1 + w_1 \\ v_2 + w_2 \\ 42 \end{pmatrix} \in V_2 \quad \text{und} \quad a\vec{v} = a \begin{pmatrix} v_1 \\ v_2 \\ 42 \end{pmatrix} := \begin{pmatrix} a \cdot v_1 \\ a \cdot v_2 \\ 42 \end{pmatrix} \in V_2$$

Hier wirken Addition und Skalierung nur auf die beiden ersten Komponenten; die dritte bleibt fest.

Der Nullvektor dieses Vektorraums wäre dann:

$$\begin{pmatrix} 0 \\ 0 \\ 42 \end{pmatrix} \in V_2$$

(Es handelt sich hierbei um die Einbettung einer zweidimensionalen Ebene in den dreidimensionalen Raum. Der Schnittpunkt der dritten Koordinatenachse mit dieser Ebene (die 42 Einheiten oberhalb der  $x_1, x_2$ -Ebene des  $\mathbb{R}^3$  liegt) entspricht dem eben angegebenen Nullvektor; von dort aus sind mit den angegebenen Operationen sämtliche Punkte der Ebene erreichbar – aber nur diese.  $V_2$  ist allerdings immer noch kein Unterraum von  $\mathbb{R}^3$ ; hier sind ja sogar die Rechenoperationen andere als in  $\mathbb{R}^3$ .)

- Die Vektoren

$$V_3 := \left\{ \begin{pmatrix} 4x_1 \\ 0 \\ \frac{1}{7}x_3 \end{pmatrix} \mid x_1, x_3 \in \mathbb{R} \right\} \subseteq \mathbb{R}^3$$

bilden einen Unterraum von  $\mathbb{R}^3$ . Da  $4x_1$  und  $\frac{1}{7}x_3$  beliebige reelle Zahlen sind, ergibt diese Struktur allerdings eher dann Sinn, wenn sie zusammen mit den unskalierten Vektoren aus  $x_1$ , 0 und  $x_3$  betrachtet wird. In dem Fall findet nämlich eine Spreizung der  $x_1$ -Komponente und eine Stauchung der  $x_3$ -Komponente statt – so eine Verzerrung könnte z.B. in der Computergrafik von Interesse sein.

- Wir hatten im Beispiel von Abbildung 6.1 schon gesehen, dass der Punkt  $Q'$  durch Verdopplung des Ortsvektors von  $Q$  erreicht werden kann. Tatsächlich bildet die Menge

$$V_4 := \left\{ \begin{pmatrix} 5x \\ 2x \end{pmatrix} = x \begin{pmatrix} 5 \\ 2 \end{pmatrix} \mid x \in \mathbb{R} \right\} \subseteq \mathbb{R}^2$$

einen Untervektorraum von  $\mathbb{R}^2$ , da

$$\begin{pmatrix} 5x \\ 2x \end{pmatrix} + \begin{pmatrix} 5y \\ 2y \end{pmatrix} = \begin{pmatrix} 5(x+y) \\ 2(x+y) \end{pmatrix} = (x+y) \begin{pmatrix} 5 \\ 2 \end{pmatrix} \in V_4 \quad \text{und} \quad a \begin{pmatrix} 5x \\ 2x \end{pmatrix} = \begin{pmatrix} 5ax \\ 2ax \end{pmatrix} = (ax) \begin{pmatrix} 5 \\ 2 \end{pmatrix} \in V_4$$

Dieser Untervektorraum ist eindimensional; es handelt sich um die *Ursprungsgerade*, die den Punkt  $Q(5, 2)$  enthält. Für diesen Punkt ist  $x = 1$  einzusetzen; für  $Q'$  wäre  $x = 2$ .

- Neben den reellen Vektorräumen (die wir gleich noch im Detail behandeln) sind auch kartesische Produkträume über endlichen Körpern möglich, z.B:

$$\mathbb{Z}_3^2 = \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \end{pmatrix} \right\}$$

Die einzelnen Komponenten können nur hier nur drei verschiedene Werte annehmen, daher ergeben sich  $3^2 = 9$  verschiedene Vektoren – auch der Vektorraum ist also endlich! Die komponentenweise Addition und Multiplikation sind hier modulo 3 auszuführen (denn wir erinnern uns, dass  $\mathbb{Z}_3$  eigentlich die drei Restklassen modulo 3 enthält). Es wäre also z.B.

$$\begin{pmatrix} 2 \\ 2 \end{pmatrix} + \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{und} \quad 2 \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$



**Bemerkung:** Zum Schluss dieses Unterabschnitts noch zur Vorsicht: In kartesischen Produkträumen stammen die Vektorkomponenten und die Skalierungsfaktoren für Vektoren aus der gleichen Menge  $\mathbb{K}$ .

Es gibt aber auch Vektorräume, die völlig anders strukturiert sind, z.B. die auf dem Intervall  $[a, b]$  stetigen reellen Funktionen<sup>5</sup>. Hier besteht die Menge  $V$  der Vektoren aus den einzelnen Funktionen (von Vektorkomponenten zu sprechen hat hier keinen Sinn), und die Skalare sind aus  $\mathbb{R}$  – es handelt sich also um einen  $\mathbb{R}$ -Vektorraum (Addition und Skalierung der Funktionen sind ähnlich wie in Abschnitt 4.1.4 erklärt). In Mathematik 1 kein Vorlesungsstoff.

Die Notation mit Vektorpfeilen (die z.B. für Funktionenräume nicht üblich ist) beschränkt sich streng genommen auf die Produkträume  $\mathbb{K}^n$ . Wir haben sie bisher durchgängig verwendet, weil wir in dieser Vorlesung nur mit solchen zu tun haben werden.

### 6.1.5 Skalarprodukt bei $\mathbb{R}$ -Vektorräumen

Bevor wir weitere Beispiele betrachten, führen wir zunächst den Begriff des *Skalarproduktes* ein. Ein Skalarprodukt bildet ein Paar von Vektoren eines  $\mathbb{R}$ -Vektorraums auf eine Zahl aus  $\mathbb{R}$  ab. Allgemein (kein Vorlesungsstoff) heißen solche Produkte bei  $\mathbb{K}$ -Vektorräumen, falls sie definiert sind, *innere Produkte* – und die Räume, für die ein inneres Produkt definierbar ist, heißen *Innenprodukträume*.

Für uns wird maßgeblich das Skalarprodukt auf  $\mathbb{R}^n$  wichtig sein, das wir im nächsten Abschnitt nochmal gezielt aufgreifen.

**Definition 6.5** (Skalarprodukt). Für einen  $\mathbb{R}$ -Vektorraum  $V$  heißt eine Funktion, welche von  $V \times V$  nach  $\mathbb{R}$  abbildet per

$$(\vec{v}, \vec{w}) \mapsto \langle \vec{v}, \vec{w} \rangle$$

Skalarprodukt auf  $V$  (auch: inneres Produkt), falls sie folgenden Eigenschaften genügt ( $\vec{u}, \vec{v}, \vec{w} \in V$  beliebig):

1. *Symmetrie:*  $\langle \vec{v}, \vec{w} \rangle = \langle \vec{w}, \vec{v} \rangle$
2. *Bilinearität:*  $\langle a\vec{u} + b\vec{v}, \vec{w} \rangle = a \langle \vec{u}, \vec{w} \rangle + b \langle \vec{v}, \vec{w} \rangle$
3. *Positive Definitheit:* Für alle  $\vec{v} \neq \vec{0}$  ist  $\langle \vec{v}, \vec{v} \rangle > 0$

**Bemerkungen:**

- Das Produkt bildet ab in den Körper  $\mathbb{R}$  der *Skalare*, über dem der Vektorraum aufgespannt ist; daher der Name.
- Beim Kriterium der Bilinearität wurde schon verwendet, dass das Skalarprodukt symmetrisch ist. Sonst hätte man auch noch  $\langle \vec{u}, a\vec{v} + b\vec{w} \rangle = a \langle \vec{u}, \vec{v} \rangle + b \langle \vec{u}, \vec{w} \rangle$  zu prüfen, denn “bilinear” bedeutet “linear” in beiden Argumenten.

(Linearität einer Funktion bedeutet (Vorgriff auf das nächste Kapitel, hier nur zum besseren Verständnis), dass die Anwendung der Funktion mit der Addition und Skalierung (also den Vektorraumoperationen) *vertauscht*; es gibt also dann keinen Unterschied, ob man die Funktion einer Summe von z.B. drei Vektoren betrachtet oder die Summe der drei Werte, wenn man die Funktion jeweils einzeln auf die Vektoren anwendet.)

- Aus der Definitheit folgt, dass es sich, wenn das Skalarprodukt  $\langle \vec{v}, \vec{v} \rangle$  null ergibt, bei  $\vec{v}$  um den Nullvektor  $\vec{0}$  handeln muss (kontrapositionische Betrachtung der Implikation).

---

Bei der Definitheit werden allerdings nur Skalarprodukte von Vektoren mit sich selbst betrachtet – es kann auch andere Fälle geben, für die das Skalarprodukt verschwindet:

**Definition 6.6** (Orthogonalität von Vektoren). Zwei Vektoren  $\vec{v}, \vec{w}$  eines reellen Vektorraums  $V$  mit Skalarprodukt heißen orthogonal (zueinander), falls

$$\langle \vec{v}, \vec{w} \rangle = 0$$

Falls  $\vec{v}, \vec{w} \neq \vec{0}$ , notiert man dann auch:

$$\vec{v} \perp \vec{w}$$

---

<sup>5</sup>zum Stetigkeitsbegriff mehr in Mathematik 2/Analysis; hier erklären wir vorläufig (und ohne Anspruch auf Exaktheit!), dass “stetig” bedeutet, dass der Funktionsgraph sich in einem Zug mit einem Stift zeichnen ließe – er besitzt also keine Lücken und keine Sprünge (Knicke sind erlaubt); insbesondere verschwinden die Funktionswerte auf  $[a, b]$  nicht nach  $\pm\infty$ .

### Bemerkungen:

- Damit ist der Nullvektor stets (und trivial) orthogonal zu allen Vektoren  $\vec{v} \in V$ , denn nach Satz 6.3 ist für beliebiges  $\vec{w} \in V$ :

$$\langle \vec{v}, \vec{0} \rangle = \langle \vec{v}, (0 \cdot \vec{w}) \rangle = 0 \langle \vec{v}, \vec{w} \rangle = 0$$

Hierbei wurde im zweiten Schritt die Bilinearität des Skalarprodukts ausgenutzt.

- Wie der Name allerdings vermuten lässt, werden Vektoren, die nicht  $\vec{0}$  sind, in  $\mathbb{R}^n$  in geometrischer Anschauung *senkrecht* aufeinander stehen.
- Auch das Konzept der Orthogonalität gilt für sämtliche Vektorräume mit innerem Produkt.

---

Zum Abschluss dieses sehr allgemein gehaltenen Unterabschnitts betrachten wir noch einen Zusammenhang, der ebenfalls mit der Orientierung von Vektoren zusammen hängt (hier nur für reelle Vektorräume formuliert):

**Satz 6.7** (Cauchy-Schwarz-Ungleichung). <sup>6</sup> Für einen  $\mathbb{R}$ -Vektorraum  $V$  gilt mit  $\vec{v}, \vec{w} \in V$ :

$$|\langle \vec{v}, \vec{w} \rangle| \leq \sqrt{\langle \vec{v}, \vec{v} \rangle} \cdot \sqrt{\langle \vec{w}, \vec{w} \rangle}$$

(Beweis: S. 322.)

### Bemerkungen:

- Man findet auch oft die quadrierte Form der Ungleichung. Für reelle Vektorräume wäre dies:

$$\langle \vec{v}, \vec{w} \rangle^2 \leq \langle \vec{v}, \vec{v} \rangle \langle \vec{w}, \vec{w} \rangle$$

- Eine weitere Schreibweise der Ungleichung ist

$$-1 \leq \frac{\langle \vec{v}, \vec{w} \rangle}{\sqrt{\langle \vec{v}, \vec{v} \rangle} \cdot \sqrt{\langle \vec{w}, \vec{w} \rangle}} \leq 1$$

- (Mit dieser Ungleichung folgt in der Quantenmechanik die Heisenbergsche Unschärferelation.)
- Insbesondere gilt die Gleichheit für den Betrag des Skalarprodukts, falls  $\vec{w} = a\vec{v}$ . Tatsächlich gilt sie auch *nur* für diesen Fall, was wir hier aber ohne Beweis akzeptieren (in der Motivation des kanonischen Skalarprodukts für  $\mathbb{R}^n$  im nächsten Abschnitt lässt sich dies einsehen).

## 6.1.6 Norm

Wir sehen gleich, dass mit dem Skalarprodukt auf reellen Vektorräumen auch eine *Norm* definiert werden kann – darunter können wir uns ein *Längenmaß* für Vektoren vorstellen. Wie beim Skalarprodukt definieren wir zunächst, was wir unter dem Begriff der Norm verstehen wollen.

Es stellt sich heraus, dass es diverse Abbildungen gibt, die den Eigenschaften einer Norm genügen; für reelle Vektorräume mit Skalarprodukt gibt es jedoch eine ganz heraus gehobene Norm – nur mit dieser wollen wir uns im weiteren Verlauf dieser Vorlesung beschäftigen<sup>7</sup>

**Definition 6.8** (Norm eines reellen Vektorraums). Für einen  $\mathbb{R}$ -Vektorraum  $V$  heißt eine Abbildung von  $V$  nach  $\mathbb{R}$  per

$$\vec{v} \mapsto \|\vec{v}\|$$

Norm auf  $V$ , falls sie folgende Eigenschaften erfüllt ( $\vec{v}, \vec{w} \in V$  und  $a \in \mathbb{R}$  beliebig):

1. Positive Definitheit:

$$\|\vec{v}\| \geq 0 \quad \text{und} \quad (\|\vec{v}\| = 0) \Rightarrow (\vec{v} = \vec{0})$$

2. Dreiecksungleichung:

$$\|\vec{v} + \vec{w}\| \leq \|\vec{v}\| + \|\vec{w}\|$$

3. Skalierbarkeit:

$$\|a\vec{v}\| = |a| \cdot \|\vec{v}\|$$

Besitzt  $V$  eine Norm auf  $V$ , so heißt  $V$  auch normierter Vektorraum.

---

<sup>6</sup>A.-L. Cauchy, frz. Mathematiker; H. A. Schwarz, dt. Mathematiker

<sup>7</sup>Allerdings kann der Normbegriff in späteren Vorlesungen durchaus noch einmal wichtig werden!

**Bemerkung:** Diese Eigenschaften decken sich mit der geometrischen Beobachtung, wenn man die Länge der Vektorpfeile im kartesischen Koordinatensystem als Norm interpretiert (wir sehen später noch, dass das legitim ist). Zur Dreiecksungleichung betrachte man das Parallelogramm in Abbildung 6.1 (S. 148), das von den Vektoren  $\vec{p}, \vec{q}$  aufgespannt wird. Die Summe der Vektoren entspricht einer der Diagonalen des Parallelogramms; es ist aber klar, dass solch eine Diagonale nicht länger sein kann als die Summe der beiden Seitenlängen.

Wie angekündigt, gehen wir in dieser Vorlesung nicht weiter auf die verschiedenen Normen für diverse Vektorräume ein – aber wir stellen allgemein fest, dass für jeden Vektorraum mit Skalarprodukt eine solche Norm existiert (und das wird für uns in der Folge wichtig sein):

**Satz 6.9** (Skalarprodukt-Norm). *Für jeden  $\mathbb{R}$ -Vektorraum  $V$  mit Skalarprodukt ist die folgende Abbildung (mit  $\vec{v} \in V$ )*

$$\|\vec{v}\| := \sqrt{\langle \vec{v}, \vec{v} \rangle}$$

*eine Norm.*

(Beweis: S. 322.)

**Bemerkungen:**

- Jedes Skalarprodukt induziert also direkt auch eine Norm!
- Aus der Definitheit des Skalarprodukts folgt, dass nur der Nullvektor  $\vec{0}$  eine Länge von 0 hat. Das deckt sich mit der geometrischen Anschauung, wenn wir bedenken, dass z.B. in  $\mathbb{R}^n$  alle Vektoren, die sich als Pfeile zeichnen lassen (und das gelingt für den Nullvektor als einzigem Fall nicht), die Eigenschaft einer geometrischen (und positiven) Länge besitzen.
- Wegen der Definitheit des Skalarprodukts ist die obige Abbildung wohldefiniert, da das Skalarprodukts eines Vektors mit sich selbst eine nichtnegative Zahl ergibt; aus jeder solchen lässt sich die Quadratwurzel ziehen.

### 6.1.7 Metrik

Um diesen Abschnitt vollständig zu machen, führen wir noch einen *Abstandsbegriff* ein, der formal als *Metrik* bezeichnet wird. Wie bei der Norm gibt es diverse Abbildungen, die den Eigenschaften einer Metrik genügen; für diese Vorlesung wird aber nur eine konkrete Abbildung davon nötig sein<sup>8</sup>.

Wir hatten bei den Ortsvektoren aus Abbildung 6.1 bereits einen Abstandsvektor  $\vec{x}_{QP}$  gesehen, den wir hinterher mit dem Ausdruck  $\vec{p} - \vec{q}$  identifiziert hatten. Es ist intuitiv ersichtlich, dass die Länge dieses Vektors den Abstand zwischen den Punkten  $P$  und  $Q$  angibt. Für eine solche Vektorlänge haben wir aber bereits den Begriff der Norm, und wir haben mit der Skalarproduktsnorm bereits eine wohl definierte Norm zur Verfügung.

Nachdem wir nun den Abstandsbegriff formal definieren, stellen wir fest, dass jeder normierte Vektorraum auch direkt über eine Metrik verfügt.

**Definition 6.10** (Metrik eines reellen Vektorraums). *Für einen  $\mathbb{R}$ -Vektorraum  $V$  heißt eine Abbildung  $d : V \times V \rightarrow \mathbb{R}$  Metrik (oder: Distanzmaß, Abstandsfunktion) auf  $V$ , falls sie folgende Eigenschaften erfüllt ( $\vec{u}, \vec{v}, \vec{w} \in V$  beliebig):*

1. *Definitheit:*  $(d(\vec{v}, \vec{w}) = 0) \Leftrightarrow (\vec{v} = \vec{w})$
2. *Symmetrie:*  $d(\vec{v}, \vec{w}) = d(\vec{w}, \vec{v})$
3. *Dreiecksungleichung:*  $d(\vec{v}, \vec{w}) \leq d(\vec{v}, \vec{u}) + d(\vec{u}, \vec{w})$

*Besitzt  $V$  eine Metrik auf  $V$ , so heißt  $V$  auch metrischer Raum.*

<sup>8</sup>Andere Metriken findet man z.B. im Bereich Machine Learning, etwa die Manhattan-Distanz oder die Mahalanobis-Distanz. Hier kein Vorlesungsstoff.

**Bemerkung:** Der Begriff der Metrik ist nicht nur auf Vektorräume beschränkt, wird aber im Kontext der Vorlesung nur für solche benötigt.

Wir zeigen nun, dass jeder normierte reelle Vektorraum auch ein metrischer Raum ist:

**Satz 6.11** (Norminduzierte Metrik). *Jeder normierte  $\mathbb{R}$ -Vektorraum  $V$  besitzt mit der Abbildung*

$$d(\vec{v}, \vec{w}) := \|\vec{v} - \vec{w}\|$$

*eine Metrik.*

(Beweis: S. 323.)

**Bemerkung:** Also ist jeder Vektorraum mit einem Skalarprodukt ein normierter und damit auch ein metrischer Raum.

## 6.2 Die reellen Vektorräume $\mathbb{R}^n$

Mit dem Körper  $\mathbb{K} := \mathbb{R}$  liegt in den reellen Vektorräumen  $\mathbb{R}^n$ ,  $n \in \mathbb{N}$ , die “gewöhnlichste” Klasse kartesischer Produkträume vor; falls Vektorrechnung in der Schule Thema war, wurden genau diese Strukturen und ihre Rechenoperationen dort schon geübt.

Für diese Vektorräume gibt es ein spezielles Skalarprodukt, für das der oben schon eingeführte Betrag eines Vektors gerade der geometrischen Länge des Vektorpfeils im kartesischen Koordinatensystem entspricht.

Weiterhin existiert im  $\mathbb{R}^3$  noch ein Produkt von Vektoren, das zwei Vektoren auf einen Vektor aus  $\mathbb{R}^3$  abbildet – das *Kreuzprodukt*. Zu beiden Produkten werden wir jeweils danach auch die geometrische Bedeutung (im kartesischen Koordinatensystem) untersuchen.

### 6.2.1 Kanonisches Skalarprodukt in $\mathbb{R}^n$

**Satz 6.12** (Kanonisches Skalarprodukt in  $\mathbb{R}^n$ ). *Für  $n \in \mathbb{N}$  ist für  $\vec{v}, \vec{w} \in \mathbb{R}^n$  ist die Abbildung von  $\mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$  mit*

$$(\vec{v}, \vec{w}) \mapsto \vec{v} \bullet \vec{w} := \sum_{j=1}^n v_j w_j$$

*ein Skalarprodukt, das kanonisches Skalarprodukt der Vektoren  $\vec{v}$  und  $\vec{w}$ .*

(Beweis: S. 324.)

#### Bemerkungen:

- Die Schreibweise mit dem dicken Multiplikationspunkt wird speziell für das kanonische Skalarprodukt in  $\mathbb{R}^n$  verwendet. Man achte bei handschriftlicher Notation darauf, den Punkt dick genug zu machen, dass keine Verwechslungsgefahr mit dem gewöhnlichen Multiplikationspunkt “.” besteht.

Die “mathematischere” Schreibweise ist allerdings diejenige mit Winkelklammern.

- Für das Skalarprodukt eines Vektors aus  $\mathbb{R}^n$  mit sich selbst schreibt man auch:

$$\vec{v} \bullet \vec{v} = \vec{v}^2$$

Vorsicht aber mit anderen Potenzen! Denn  $\vec{v}^3$  als dreifaches Skalarprodukt wäre z.B. gar nicht definiert, und der Ausdruck  $\vec{v}^2 \vec{v}$  entspricht

$$(\vec{v} \bullet \vec{v}) \cdot \vec{v}$$

Der Term in Klammern ist eine reelle Zahl; mit dieser wird der Vektor  $\vec{v}$  skaliert – Letzteres ist aber eine skalare Multiplikation, die *nicht* dem Skalarprodukt entspricht (und für die wir eigentlich oben ein eigenes Symbol  $\odot$  vereinbart hatten).

Wir merken uns für den  $\mathbb{K}$ -Vektorraum  $V$ :

Skalare Multiplikation bildet von  $\mathbb{K} \times V$  nach  $V$  ab;  
 Skalarprodukte dagegen von  $V \times V$  nach  $\mathbb{K}$

- Bei Summen in  $\mathbb{R}^n$  gehen wir ab sofort davon aus, dass die Indices von 1 bis  $n$  laufen und schreiben in Rechnungen oder Beweisen meist abkürzend

$$\sum_j \cdots \quad \text{statt} \quad \sum_{j=1}^n \cdots$$

**Beispiele:**

- Wir betrachten alle (kanonischen) Skalarprodukte der drei Vektoren

$$\vec{u} := \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \quad \vec{v} := \begin{pmatrix} 3 \\ -7 \end{pmatrix} \quad \text{und} \quad \vec{w} := \begin{pmatrix} -2 \\ -3 \end{pmatrix}$$

Dabei erhalten wir:

$$\begin{aligned} \vec{u} \bullet \vec{u} &= 2 \cdot 2 + 1 \cdot 1 = 4 + 1 = 5 \\ \vec{u} \bullet \vec{v} &= 2 \cdot 3 + 1 \cdot (-7) = 6 - 7 = -1 \\ \vec{u} \bullet \vec{w} &= 2 \cdot (-2) + 1 \cdot (-3) = -4 - 3 = -7 \\ \vec{v} \bullet \vec{v} &= 3 \cdot 3 + (-7) \cdot (-7) = 9 + 49 = 58 \\ \vec{v} \bullet \vec{w} &= 3 \cdot (-2) + (-7) \cdot (-3) = -6 + 21 = 15 \\ \vec{w} \bullet \vec{w} &= (-2) \cdot (-2) + (-3) \cdot (-3) = 4 + 9 = 13 \end{aligned}$$

- Wir betrachten alle (kanonischen) Skalarprodukte der drei Vektoren

$$\vec{u} := \begin{pmatrix} -2 \\ 2 \\ 1 \end{pmatrix}, \quad \vec{v} := \begin{pmatrix} 1 \\ 2 \\ -2 \end{pmatrix} \quad \text{und} \quad \vec{w} := \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}$$

Dabei erhalten wir:

$$\begin{aligned} \vec{u} \bullet \vec{u} &= (-2) \cdot (-2) + 2 \cdot 2 + 1 \cdot 1 = 4 + 4 + 1 = 9 \\ \vec{u} \bullet \vec{v} &= (-2) \cdot 1 + 2 \cdot 2 + 1 \cdot (-2) = -2 + 4 - 2 = 0 \\ \vec{u} \bullet \vec{w} &= (-2) \cdot 2 + 2 \cdot 1 + 1 \cdot 2 = -4 + 2 + 2 = 0 \\ \vec{v} \bullet \vec{v} &= 1 \cdot 1 + 2 \cdot 2 + (-2) \cdot (-2) = 1 + 4 + 4 = 9 \\ \vec{v} \bullet \vec{w} &= 1 \cdot 2 + 2 \cdot 1 + (-2) \cdot 2 = 2 + 2 + (-4) = 0 \\ \vec{w} \bullet \vec{w} &= 2 \cdot 2 + 1 \cdot 1 + 2 \cdot 2 = 4 + 1 + 4 = 9 \end{aligned}$$

Offenbar sind die drei Vektoren paarweise orthogonal zueinander nach Definition 6.6, denn ihre jeweiligen Skalarprodukte verschwinden.

Nun haben wir oben allgemein erkannt, dass jeder Vektorraum mit Skalarprodukt auch normiert und metrisch ist. Für unser kanonisches Skalarprodukt definieren wir daher:

**Definition 6.13** (Euklidische Norm). *Mit dem kanonischen Skalarprodukt aus der Definition in Satz 6.12 ist die euklidische Norm auf dem Vektorraum  $\mathbb{R}^n$  gegeben per*

$$\|\vec{v}\| := \sqrt{\vec{v} \bullet \vec{v}} = \sqrt{\sum_{j=1}^n v_j^2}$$

In  $\mathbb{R}^n$  schreiben wir statt  $\|\vec{v}\|$  auch  $|\vec{v}|$  und nennen dies die Länge des Vektors  $\vec{v}$ .

### Bemerkungen:

- Nach Satz 6.12 handelt es sich bei “•” um ein Skalarprodukt. Dann ist nach Satz 6.9 die obige Abbildung auch eine Norm. Weiterhin lässt sich dann über die euklidische Metrik

$$d(\vec{v}, \vec{w}) := \|\vec{v} - \vec{w}\| = \sqrt{\sum_{j=1}^n (v_j - w_j)^2}$$

auch ein Abstandsmaß definieren.

Zu einem Punkt  $P$  mit Ortsvektor  $\vec{p}$  im kartesischen Koordinatensystem ist dann der Abstand von  $P$  zum Koordinatenursprung gegeben durch  $d(\vec{p}, \vec{0}) = |\vec{p}|$ .

- Die euklidische Norm ergibt für Pfeile im kartesischen Koordinatensystem genau die geometrische Länge dieser Pfeile. Für  $n = 2$  folgt dies aus dem Satz des Pythagoras (die Länge der Diagonalen eines Rechtecks lässt sich über ein rechtwinkliges Dreieck berechnen); für höheres  $n$  lässt sich dies beliebig verallgemeinern. Für  $n = 3$  passt die Formel zur Länge der Diagonalen eines Quaders mit den Seitenlängen  $v_1, v_2, v_3$ .

**Beispiele:** Wir betrachten die Vektoren von den vorigen beiden Beispielen:

- Mit

$$\vec{u} := \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \quad \vec{v} := \begin{pmatrix} 3 \\ -7 \end{pmatrix} \quad \text{und} \quad \vec{w} := \begin{pmatrix} -2 \\ -3 \end{pmatrix}$$

gilt:

$$\begin{aligned} |\vec{u}| &= \sqrt{\vec{u} \bullet \vec{u}} = \sqrt{5} \\ |\vec{v}| &= \sqrt{\vec{v} \bullet \vec{v}} = \sqrt{58} \\ |\vec{w}| &= \sqrt{\vec{w} \bullet \vec{w}} = \sqrt{13} \end{aligned}$$

- Mit

$$\vec{u} := \begin{pmatrix} -2 \\ 2 \\ 1 \end{pmatrix}, \quad \vec{v} := \begin{pmatrix} 1 \\ 2 \\ -2 \end{pmatrix} \quad \text{und} \quad \vec{w} := \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}$$

gilt:

$$|\vec{u}| = |\vec{v}| = |\vec{w}| = \sqrt{9} = 3$$

- In  $\mathbb{R}^2$  bezeichnet für  $r \geq 0$  die Menge

$$K_r := \{ (x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = r^2 \}$$

einen *Kreis* um den Koordinatenursprung. Alle (und nur!) diese Punkte haben genau den Abstand  $r$  zum Ursprung. Das gleiche gilt mit der euklidischen Metrik für die Vektoren aus

$$\left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \mid x^2 + y^2 = r^2 \right\}$$

Alle diese Vektoren  $\vec{v}$  erfüllen  $|\vec{v}| = r$ .

- Weil die euklidische Norm über das kanonische Skalarprodukt vollständig verträglich ist mit dem geometrischen Abstand im kartesischen Koordinatensystem, können wir folgern, dass es auf einem Kreis  $K_r$  wie im vorigen Beispiel zu jedem Punkt  $P(x, y) \in K_r$  mit Ortsvektor  $\vec{p}$  genau zwei Punkte aus  $K_r$  gibt, deren Ortsvektoren orthogonal zu  $\vec{p}$  sind.

Es handelt sich dabei um die beiden Punkte, die beim Schnitt des Kreises  $K_r$  mit der Ursprungsgeraden entstehen, die senkrecht zu der Ursprungsgeraden  $OP$  steht (also orthogonal zu dieser ist!).

Für

$$p = \begin{pmatrix} x \\ y \end{pmatrix}$$

sind die beiden fraglichen Punkte dann:

$$\vec{v}_1 := \begin{pmatrix} -y \\ x \end{pmatrix} \quad \text{und} \quad \vec{v}_2 := \begin{pmatrix} y \\ -x \end{pmatrix}$$

Man rechnet sofort nach, dass  $\vec{v}_1 \bullet \vec{p} = -yx + xy = 0$  und  $\vec{v}_2 \bullet \vec{p} = yx - xy = 0$ . Wegen der geometrischen Verträglichkeit akzeptieren wir, dass dies auch die beiden einzigen zu  $\vec{p}$  orthogonalen Ortsvektoren mit Länge  $r = \sqrt{x^2 + y^2}$  sind.

Bevor wir das kanonische Skalarprodukt weiter geometrisch untersuchen, vereinbaren wir noch diese

**Definition 6.14** (Einheitsvektor). *Ein Vektor  $\vec{v} \in \mathbb{R}^n$  heißt normiert oder Einheitsvektor, falls  $|\vec{v}| = 1$ .*

**Bemerkung:** Jeder Vektor  $\vec{v}$ , der nicht der Nullvektor ist, lässt sich normieren, indem man ihn mit dem Kehrwert seiner Länge skaliert. Wir notieren hier mit der Norm-Schreibweise, um die Vektornorm und die Beträge reeller Zahlen besser auseinander halten zu können:

$$\left\| \frac{1}{|\vec{v}|} \vec{v} \right\| = \left| \frac{1}{|\vec{v}|} \right| \cdot \|\vec{v}\| = \frac{1}{|\vec{v}|} \cdot \|\vec{v}\| = \frac{\|\vec{v}\|}{|\vec{v}|} = 1$$

**Beispiele:**

- Mit

$$\vec{u} := \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \quad \vec{v} := \begin{pmatrix} 3 \\ -7 \end{pmatrix} \quad \text{und} \quad \vec{w} := \begin{pmatrix} -2 \\ -3 \end{pmatrix}$$

sind die folgenden Vektoren Einheitsvektoren:

$$\frac{1}{\sqrt{5}} \vec{u}, \quad \frac{1}{\sqrt{58}} \vec{v} \quad \text{und} \quad \frac{1}{\sqrt{13}} \vec{w}$$

- Mit

$$\vec{u} := \begin{pmatrix} -2 \\ 2 \\ 1 \end{pmatrix}, \quad \vec{v} := \begin{pmatrix} 1 \\ 2 \\ -2 \end{pmatrix} \quad \text{und} \quad \vec{w} := \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}$$

sind die jeweiligen mit  $\frac{1}{3}$  skalierten Vektoren normiert.

- In  $\mathbb{R}^2$  sind die Vektoren

$$\left\{ \begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix} \mid \varphi \in [0, 2\pi) \right\}$$

normiert. Es handelt sich um die Ortsvektoren der Punkte auf dem Kreis  $K_1$  (s.o.), in *Polarardarstellung*. Lässt man  $\varphi$  von 0 bis  $2\pi$  laufen, so erreicht man in *mathematisch positivem Umlaufsinn* (also gegen den Uhrzeigersinn) nacheinander alle Punkte von  $K_1$ .

Man rechnet nach, dass jeder der Vektoren genau die Länge 1 besitzt, denn

$$\cos^2 \varphi + \sin^2 \varphi = 1$$

Dieser Zusammenhang folgt direkt aus dem Satz des Pythagoras für ein rechtwinkliges Dreieck (siehe Abbildung 4.2.3, S. 99, setze dort  $c := 1$ ) mit Hypothenusenlänge 1 und  $\varphi$  als dem Winkel zwischen Hypothenuse und einer Kathete; er ist auch als *trigonometrischer (Satz des) Pythagoras* bekannt.

Für  $\varphi := \frac{3\pi}{4}$  (also 135 Grad) erhalten wir den Ortsvektor

$$\begin{pmatrix} \cos \frac{3\pi}{4} \\ \sin \frac{3\pi}{4} \end{pmatrix} = \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} \approx \begin{pmatrix} -0.707 \\ 0.707 \end{pmatrix}$$

Für  $\varphi := \frac{\pi}{6}$  (also 30 Grad) erhalten wir den Ortsvektor

$$\begin{pmatrix} \cos \frac{\pi}{6} \\ \sin \frac{\pi}{6} \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{3}}{2} \\ \frac{1}{2} \end{pmatrix} \approx \begin{pmatrix} 0.866 \\ 0.5 \end{pmatrix}$$

### 6.2.2 Geometrische Bedeutung des kanonischen Skalarprodukts

Gegeben seien zwei beliebige Vektoren  $\vec{v}, \vec{w} \in \mathbb{R}^n \setminus \{\vec{0}\}$ . Nach dem oben gesagten haben beide Vektoren eine positive Länge. Nun wissen wir bereits, dass das Skalarprodukt der beiden Vektoren damit genau dann null ist, wenn die Vektoren orthogonal zueinander sind.

Im kartesischen Koordinatensystem stehen orthogonalen Vektoren senkrecht aufeinander; hiervon können wir uns natürlich nur in  $\mathbb{R}^2$  und  $\mathbb{R}^3$  durch Anschauung überzeugen. Falls aber die Vektoren nicht parallel sind (das wären sie, wenn es ein  $a \in \mathbb{R}$  gäbe, sodass  $\vec{w} = a\vec{v}$ ), so spannen sie eine Ebene auf (mehr dazu im nächsten Abschnitt). Wir wollen  $\vec{v}, \vec{w}$  beide vom Koordinatenursprung abtragen. Der Verbindungsvektor  $\vec{w} - \vec{v}$  liegt dann ebenfalls in dieser Ebene.

Nun zerlegen wir  $\vec{w}$  in einen Anteil parallel zu  $\vec{v}$  und einen senkrecht zu  $\vec{v}$ , wie in Abbildung 6.2 gezeigt.

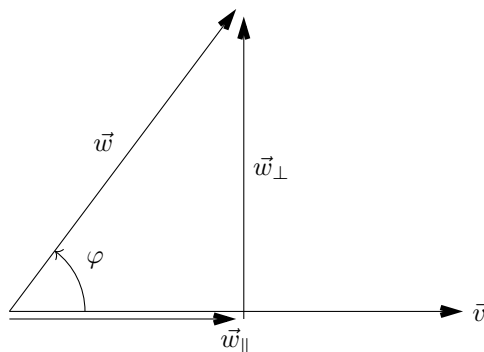


Abbildung 6.2: Parallele Projektion  $\vec{w}$  entlang  $\vec{v}$  (Teilvektoren leicht versetzt gezeichnet)

Auch der Vektor  $\vec{w}_\perp$  liegt in der Ebene von  $\vec{v}, \vec{w}$ ; der Parallelanteil  $\vec{w}_\parallel$  ohnehin.

Diese Beobachtungen gelten so in jedem  $\mathbb{R}^n$  mit  $n \geq 2$ . Dann reicht es aber intuitiv aus, sich auf die Situation in  $\mathbb{R}^2$  zu beschränken – hier gibt es nur genau eine Ebene (nämlich die  $x_1, x_2$ -Ebene), was die Rechnung erleichtert.

Zunächst halten wir fest, dass  $\vec{w}_\perp \bullet \vec{v} = 0$  gilt (wir erinnern an die Beispiele von Definition 6.13). Weiterhin ist  $\vec{w} = \vec{w}_\parallel + \vec{w}_\perp$ .

Aber dann ist wegen der Bilinearität des Skalarprodukts auch:

$$\vec{w} \bullet \vec{v} = (\vec{w}_\parallel + \vec{w}_\perp) \bullet \vec{v} = \vec{w}_\parallel \bullet \vec{v}$$

Für den Wert des Skalarproduktes ist also nur der zu  $\vec{v}$  parallele Anteil von  $\vec{w}$  entscheidend.

Die Länge von  $\vec{w}_\parallel$  können wir aus Abbildung 6.2 ermitteln; es handelt sich nämlich um ein rechtwinkliges Dreieck, sodass

$$|\vec{w}_\parallel| = \cos \varphi \cdot |\vec{w}|$$

Und damit ist  $\vec{w}_\parallel$  der Einheitsvektor in Richtung  $\vec{v}$ , aber skaliert mit dieser Länge, und es gilt:

$$\vec{w}_\parallel = \cos \varphi \cdot |\vec{w}| \cdot \frac{\vec{v}}{|\vec{v}|}$$

Dann erhalten wir für das Skalarprodukt von  $\vec{v}, \vec{w}$ :

$$\vec{v} \bullet \vec{w} = \vec{v} \bullet \vec{w}_\parallel = \vec{v} \bullet \left( \cos \varphi \cdot |\vec{w}| \cdot \frac{\vec{v}}{|\vec{v}|} \right) = \cos \varphi \cdot |\vec{w}| \cdot \frac{\vec{v} \bullet \vec{v}}{|\vec{v}|} = \cos \varphi \cdot |\vec{w}| \cdot \frac{|\vec{v}|^2}{|\vec{v}|} = \cos \varphi \cdot |\vec{v}| \cdot |\vec{w}|$$

Oder auch:

$$\cos \varphi = \frac{\vec{v} \bullet \vec{w}}{|\vec{v}| \cdot |\vec{w}|}$$

Falls übrigens der Winkel  $\varphi$  größer wird als  $\frac{\pi}{2}$ , wird die Parallelkomponente von  $\vec{w}$  entgegen gerichtet zu  $\vec{v}$  sein – dann wäre auch das Skalarprodukt von  $\vec{w}_\parallel$  mit  $\vec{v}$  negativ. Die Cosinusfunktion bildet dies bereits korrekt ab, da  $\cos \varphi < 0$  für  $\frac{\pi}{2} < \varphi \leq \pi$  (einen größeren Winkel als  $\pi$  müssen wir nicht betrachten, da wir sonst einfach von der anderen Seite aus messen könnten. Beim maximalen Winkel  $\varphi = \pi$  stehen  $\vec{v}$  und  $\vec{w}$  antiparallel zueinander).

Erwartungsgemäß verschwindet das Skalarprodukt für einen Winkel  $\varphi = \frac{\pi}{2}$ , wenn die Vektoren genau orthogonal zueinander sind.



Übrigens verträgt sich die Tatsache  $-1 \leq \cos \varphi \leq 1$  auch mit der Cauchy-Schwarz-Ungleichung aus Satz 6.7 (siehe dort die Bemerkungen).

Damit haben wir eine Verbindung zwischen dem kanonischen Skalarprodukt und der euklidischen Geometrie im kartesischen Koordinatensystem gefunden. Wir halten fest:

**Satz 6.15** (Winkel zwischen zwei Vektoren in  $\mathbb{R}^n$ ). *Für den Winkel  $\varphi$  zwischen zwei Vektoren  $\vec{v}, \vec{w} \in \mathbb{R}^n$ ,  $n > 1$ , gilt:*

$$\cos \varphi = \frac{\vec{v} \bullet \vec{w}}{|\vec{v}| \cdot |\vec{w}|}$$

**Bemerkung:** Allgemein lässt sich für alle Vektorräume mit Skalarprodukt über diese Gleichung ein Winkel definieren, nur würde man dann allgemeiner schreiben:

$$\cos \varphi := \frac{\langle \vec{v}, \vec{w} \rangle}{\sqrt{\langle \vec{v}, \vec{v} \rangle} \cdot \sqrt{\langle \vec{w}, \vec{w} \rangle}}$$

Den Wert des Winkels berechnet man hier wie dort über die Umkehrfunktion des Cosinus, die auf dem Intervall  $[0, \pi)$  sogar umkehrbar eindeutig ist.

**Beispiele:**

- Die Vektoren

$$\vec{v} := \begin{pmatrix} 4 \\ 2 \\ -3 \\ 0 \end{pmatrix} \quad \text{und} \quad \vec{w} := \begin{pmatrix} -1 \\ 2 \\ 2 \\ 42 \end{pmatrix}$$

haben das Skalarprodukt

$$\vec{v} \bullet \vec{w} = -4 + 4 - 6 + 0 = -6$$

und die Beträge

$$|\vec{v}| = \sqrt{16 + 4 + 9 + 0} = \sqrt{29} \quad \text{und} \quad |\vec{w}| = \sqrt{1 + 4 + 4 + 1764} = \sqrt{1773}$$

Damit gilt für den eingeschlossenen Winkel:

$$\varphi = \cos^{-1} \left( \frac{-6}{\sqrt{29} \cdot \sqrt{1773}} \right) = \cos^{-1} \left( \frac{-6}{\sqrt{51417}} \right) \approx \cos^{-1}(-0.02646) \approx 1.59726 \approx 0.508\pi$$

Der Winkel entspricht ungefähr 91.516 Grad. Das ist übrigens kein Zufall, denn die vierte Komponente von  $\vec{v}$  ist null – also steht  $\vec{v}$  senkrecht zur vierten Koordinatenachse. Dagegen hat der Vektor  $\vec{w}$  wegen seiner hohen vierten Komponente 42 nur eine geringe Winkelabweichung zur vierten Koordinatenachse. Am Vorzeichen des Skalarprodukts lesen wir ab, dass diese Abweichung von  $\vec{v}$  weg gerichtet ist. Insgesamt erhalten wir also nahezu einen rechten Winkel.

- Eine andere Situation hätten wir mit selbem  $\vec{v}$  und dem neuen

$$\vec{w} := \begin{pmatrix} -1 \\ 2 \\ 2 \\ 0 \end{pmatrix}$$

Dann ergäbe sich nach wie vor das gleiche Skalarprodukt, aber mit  $|\vec{w}| = 3$  bekämen wir für den eingeschlossenen Winkel:

$$\varphi = \cos^{-1} \left( \frac{-6}{\sqrt{29} \cdot \sqrt{9}} \right) = \cos^{-1} \left( \frac{-6}{\sqrt{261}} \right) \approx \cos^{-1}(-0.3714) \approx 1.951 \approx 0.621\pi$$

Der Winkel entspricht hier ungefähr 111.8 Grad.

Eine wichtige Anwendung des Skalarprodukts und der Projektion findet sich bei der Gram-Schmidt-Orthogonalisierung; siehe dazu den Exkurs A.4 – hierfür werden allerdings noch die Konzepte der Basen und der linearen Abhängigkeit aus dem nächsten Abschnitt benötigt.

### 6.2.3 Kreuzprodukt in $\mathbb{R}^3$

Für Vektoren in  $\mathbb{R}^3$  führen wir ein weiteres Produkt ein:

**Definition 6.16** (Kreuzprodukt). Für  $\vec{v}, \vec{w} \in \mathbb{R}^3$  wird die Abbildung von  $\mathbb{R}^3 \times \mathbb{R}^3$  nach  $\mathbb{R}^3$  mit

$$(\vec{v}, \vec{w}) \mapsto \vec{v} \times \vec{w} := \begin{pmatrix} v_2 w_3 - v_3 w_2 \\ v_3 w_1 - v_1 w_3 \\ v_1 w_2 - v_2 w_1 \end{pmatrix}$$

als Kreuzprodukt (oder: Vektorprodukt, äußeres Produkt) der Vektoren  $\vec{v}$  und  $\vec{w}$  bezeichnet.

**Beispiele:**

- Für

$$\vec{v} := \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix} \quad \text{und} \quad \vec{w} := \begin{pmatrix} 1 \\ 2 \\ -2 \end{pmatrix}$$

ist das Kreuzprodukt:

$$\vec{v} \times \vec{w} = \begin{pmatrix} 1 \cdot (-2) - 2 \cdot 2 \\ 2 \cdot 1 - 2 \cdot (-2) \\ 2 \cdot 2 - 1 \cdot 1 \end{pmatrix} = \begin{pmatrix} -6 \\ 6 \\ 3 \end{pmatrix}$$

Achtung: das Produkt ist nicht symmetrisch, sondern sogar antisymmetrisch:

$$\vec{w} \times \vec{v} = \begin{pmatrix} 2 \cdot 2 - (-2) \cdot 1 \\ (-2) \cdot 2 - 1 \cdot 2 \\ 1 \cdot 1 - 2 \cdot 2 \end{pmatrix} = \begin{pmatrix} 6 \\ -6 \\ -3 \end{pmatrix} = -\vec{v} \times \vec{w}$$

- Für

$$\vec{v} := \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix} \quad \text{und} \quad \vec{w} := \begin{pmatrix} -2 \\ 1 \\ -3 \end{pmatrix}$$

erhalten wir ein Kreuzprodukt

$$\vec{v} \times \vec{w} = \begin{pmatrix} 2 \cdot (-3) - 4 \cdot 1 \\ 4 \cdot (-2) - 1 \cdot (-3) \\ 1 \cdot 1 - 2 \cdot (-2) \end{pmatrix} = \begin{pmatrix} -10 \\ -5 \\ 5 \end{pmatrix}$$

Man vergewissere sich zur Übung, dass  $\vec{w} \times \vec{v}$  auch hier auf genau den negativen Vektor führt.

- Für

$$\vec{v} := \begin{pmatrix} 4 \\ -2 \\ 6 \end{pmatrix} \quad \text{und} \quad \vec{w} := \begin{pmatrix} -2 \\ 1 \\ -3 \end{pmatrix}$$

erhalten wir ein Kreuzprodukt

$$\vec{v} \times \vec{w} = \begin{pmatrix} (-2) \cdot (-3) - 6 \cdot 1 \\ 6 \cdot (-2) - 4 \cdot (-3) \\ 4 \cdot 1 - (-2) \cdot (-2) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = \vec{0}$$

- Vektoren aus  $\mathbb{R}^2$  lassen sich als  $\mathbb{R}^3$ -Vektoren ausdrücken, falls man deren dritte Komponente konstant auf null setzt. Dann gilt für

$$\vec{v} := \begin{pmatrix} a \\ b \\ 0 \end{pmatrix} \quad \text{und} \quad \vec{w} := \begin{pmatrix} c \\ d \\ 0 \end{pmatrix}$$

das Kreuzprodukt

$$\vec{v} \times \vec{w} = \begin{pmatrix} b \cdot 0 - 0 \cdot d \\ 0 \cdot c - a \cdot 0 \\ a \cdot d - b \cdot c \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ ad - bc \end{pmatrix}$$

Dieser Vektor liegt offenbar vollständig entlang der dritten Koordinatenachse.

Wir formulieren nun einige Eigenschaften zum Kreuzprodukt:

**Satz 6.17** (Eigenschaften des Kreuzprodukts). *Das Kreuzprodukt aus Definition 6.16 hat folgende Eigenschaften (mit  $a, b \in \mathbb{R}$  und  $\vec{u}, \vec{v}, \vec{w} \in \mathbb{R}^3$  beliebig):*

1. *Bilinearität in den Faktoren:*

$$(a\vec{u} + b\vec{v}) \times \vec{w} = a \cdot (\vec{u} \times \vec{w}) + b \cdot (\vec{v} \times \vec{w})$$

$$\vec{u} \times (a\vec{v} + b\vec{w}) = a \cdot (\vec{u} \times \vec{v}) + b \cdot (\vec{u} \times \vec{w})$$

2. *Verswinden des Kreuzprodukts eines Vektors mit sich selbst:  $\vec{v} \times \vec{v} = \vec{0}$*

3. *Antisymmetrie:  $\vec{v} \times \vec{w} = -\vec{w} \times \vec{v}$*

4. *Orthogonalität mit den Faktoren:  $\vec{v} \perp (\vec{v} \times \vec{w}) \perp \vec{w}$*

(Beweis: S. 324.)

**Bemerkungen:**

- Wegen  $\vec{0} = 0 \cdot \vec{w}$  (Satz 6.3) folgt aus der Linearität direkt:

$$\vec{v} \times \vec{0} = \vec{0} \times \vec{v} = \vec{0}$$

- Die Antisymmetrie konnten wir oben bei den Beispielen bereits beobachten.
- Im dritten Beispiel war  $\vec{v} = -2\vec{w}$ . Mit der Bilinearität und der alternierenden Eigenschaft  $\vec{v} \times \vec{v} = \vec{0}$  musste dann auch  $\vec{v} \times \vec{w}$  den Nullvektor ergeben.
- Von der Orthogonalität mit den Faktoren überzeuge man sich in obigen Beispielen zur Übung.
- Zur Orientierung:  $\vec{v}$ ,  $\vec{w}$  und  $\vec{v} \times \vec{w}$  bilden, falls  $\vec{v}, \vec{w}$  weder den Nullvektor enthalten noch parallel sind, in dieser Reihenfolge ein *Rechtssystem*: Das ist ein Koordinatensystem, dessen Achsen nach den Fingern der rechten Hand ausgerichtet sind: Die erste Achse zeigt entlang des abgespreizten Daumens; die zweite entlang des ausgestreckten Zeigefingers, und die dritte entlang des Mittelfingers, der senkrecht zur Handfläche gestreckt ist.

Das kartesische Koordinatensystem des  $\mathbb{R}^3$  ist selbst ein Rechtssystem, denn es gilt für die Koordinaten-Einheitsvektoren (mehr dazu in Kürze), wie man leicht nachrechnen kann:

$$\vec{e}_1 \times \vec{e}_2 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \times \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \vec{e}_3$$

- (In der Physik findet sich das Kreuzprodukt bei mechanischen Größen wieder, z.B. beim Drehmoment, bei der Corioliskraft (die für die Drehrichtungen der Hoch- und Tiefdruckgebiete in der Atmosphäre verantwortlich ist) oder bei der Lorentzkraft (die Kraft auf elektrische Ströme in einem Magnetfeld).)

## 6.2.4 Geometrische Bedeutung des Kreuzprodukts

Beim kanonischen Skalarprodukt hatten wir fest gestellt, dass es einen Vektor  $\vec{w}$  parallel auf einen Vektor  $\vec{v}$  projiziert, d.h. seinen senkrechten Anteil  $\vec{w}_\perp$  eliminiert. Der Wert des Skalarprodukts hängt also eng mit dem Winkel zwischen den beiden Vektoren zusammen, falls man beide vom selben Punkt aus abträgt.

Beim Kreuzprodukt hingegen verschwindet der parallele Anteil  $\vec{w}_\parallel$ , und maßgeblich bleibt der senkrechte Anteil  $\vec{w}_\perp$ . Abbildung 6.3 zeigt die Situation – eingezeichnet sind die beiden Vektoren, der senkrechte Anteil von  $\vec{w}$  sowie das Kreuzprodukt. Ähnlich wie bei den Überlegungen zum Skalarprodukt nehmen wir an, dass  $\vec{v}, \vec{w}$  in der  $x_1, x_2$ -Ebene des  $\mathbb{R}^3$  liegen. Nach Satz 6.17 (Eigenschaft 4) steht das Kreuzprodukt senkrecht auf beiden Faktoren und ist in diesem Fall parallel zur  $x_3$ -Achse; die rechten Winkel sind im Bild eingezeichnet. Wir gehen außerdem davon aus, dass  $\vec{v}, \vec{w}$  nicht parallel sind (sonst würde deren Kreuzprodukt verschwinden) und beide nicht dem Nullvektor entsprechen (dito).

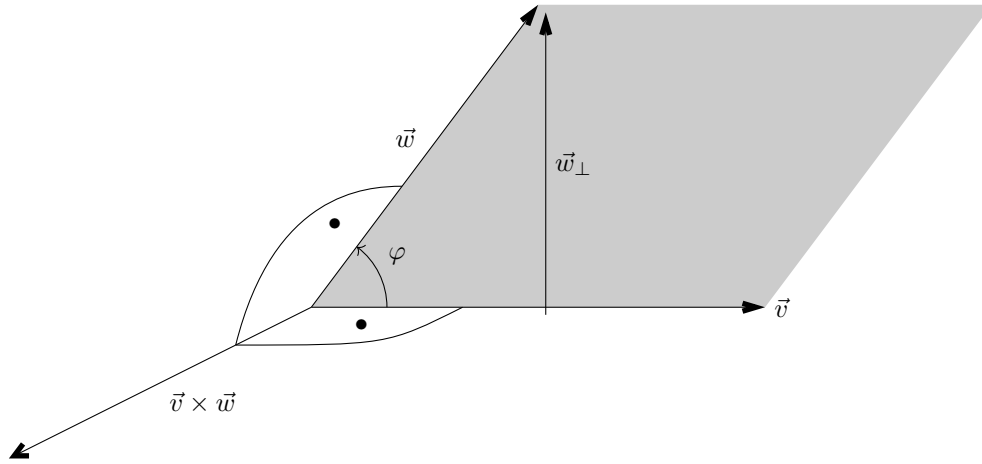


Abbildung 6.3: Kreuzprodukt und aufgespanntes Parallelogramm von  $\vec{v}$  und  $\vec{w}$  (Anteil  $\vec{w}_\perp$  leicht versetzt gezeichnet)

Weiterhin enthält die Abbildung 6.3 das Parallelogramm (schraffiert eingezeichnet), das von den Vektoren  $\vec{v}, \vec{w}$  aufgespannt wird (vgl. dazu auch Abbildung 6.1, S. 148). Sein Flächeninhalt beträgt gerade  $|\vec{w}_\perp| \cdot |\vec{v}|$  – zur Begründung denke man sich das Dreieck aus Koordinatenursprung, Ausgangspunkt von  $\vec{w}_\perp$  und Endpunkt von  $\vec{w}$  um  $|\vec{v}|$  nach rechts verschoben. Dort füllt es genau die dreieckige Lücke rechts vom Parallelogramm aus, um mit der verbliebenen Fläche ein *Rechteck* zu bilden; dessen Fläche entspricht dem Produkt der beiden Seitenlängen.

Nun sei also

$$\vec{v} = \begin{pmatrix} v_1 \\ v_2 \\ 0 \end{pmatrix} \quad \text{und} \quad \vec{w} = \begin{pmatrix} w_1 \\ w_2 \\ 0 \end{pmatrix}$$

Wir wissen außerdem (für den letzten Schritt siehe das entsprechende Beispiel bei Definition 6.16):

$$\vec{v} \times \vec{w} = \vec{v} \times (\vec{w}_\parallel + \vec{w}_\perp) = \vec{v} \times \vec{w}_\perp =: \begin{pmatrix} 0 \\ 0 \\ x \end{pmatrix}$$

Dann gilt für die von null (potentiell) verschiedene dritte Komponente des Kreuzprodukts:

$$x = v_1 w_2 - v_2 w_1 = v_1 (\vec{w}_\perp)_2 - v_2 (\vec{w}_\perp)_1$$

Nun erinnern wir uns daran, dass es in der  $x_1, x_2$ -Ebene nur genau zwei Vektoren mit Länge  $|\vec{v}|$  gibt, die senkrecht auf  $\vec{v}$  stehen (siehe dazu das Beispiel bei Definition 6.13 zur euklidischen Norm). Der Vektor, der in *mathematisch positivem Umlaufsinn*, d.h. gegen den Uhrzeigersinn, mit Drehung um  $\frac{\pi}{2}$  erreicht wird, ist

$$\vec{n}_v := \begin{pmatrix} -v_2 \\ v_1 \\ 0 \end{pmatrix}$$

Nun findet man durch Vergleich mit obiger Formel für  $x$ :

$$x = \vec{n}_v \bullet \vec{w} = \vec{n}_v \bullet \vec{w}_\perp$$

Die dritte Komponente des Kreuzprodukts ist also das kanonische Skalarprodukt aus dem *Normalenvektor* (bzgl.  $\vec{v}$ )  $\vec{n}_v$  und dem Vektor  $\vec{w}$ , beziehungsweise mit dessen Anteil  $\vec{w}_\perp$ . Die Gleichheit ist nicht verwunderlich, da  $\vec{w}_\perp$  gerade die zu  $\vec{n}_v$  parallele Komponente des Vektors  $\vec{w}$  ist. Die Situation ist in Abbildung 6.4 gezeigt.

Wenn aber nun die Komponente  $x$  des Kreuzprodukts als (kanonisches!) Skalarprodukt beschreibbar ist, so gelten alle Überlegungen, die wir weiter oben hierzu schon angestellt hatten. Und dann ist:

$$x = |\vec{n}_v| \cdot |\vec{w}_\perp| = |\vec{v}| \cdot |\vec{w}_\perp| = |\vec{v}| \cdot |\vec{w}| \cdot \sin \varphi$$

Denn die Längen von  $\vec{n}_v$  und  $\vec{v}$  sind gleich, und die Länge von  $\vec{w}_\perp$  liest man mit dem Sinus aus dem rechtwinkligen Dreieck ab (vgl. dazu auch Abbildung 4.2.3, S. 99).

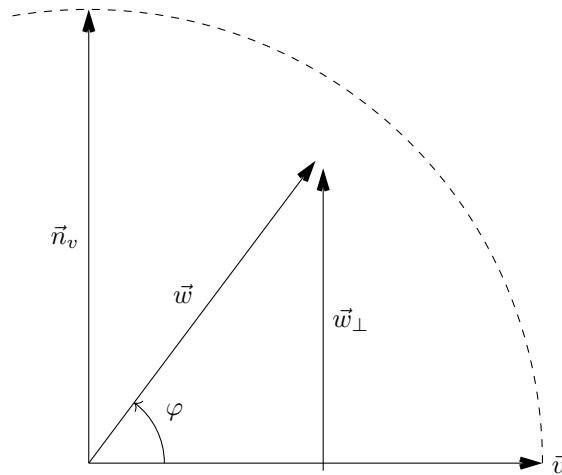


Abbildung 6.4: Zur Herleitung des Kreuzprodukts von  $\vec{v}$  und  $\vec{w}$

Weiterhin entspricht  $x$  offenbar genau der *Fläche des Parallelogramms*, das von  $\vec{v}$  und  $\vec{w}$  aufgespannt wird (s.o.).

Wir haben nun fast die gesuchte geometrische Bedeutung des Kreuzprodukts gefunden – aber noch nicht ganz. Wie verhält es sich nämlich, falls der Winkel  $\varphi$  größer ist als  $\pi$  (180 Grad)? Beim Skalarprodukt, für das der parallele Anteil von  $\vec{w}$  maßgeblich war, konnten wir uns auf einen Winkel von höchstens  $\pi$  beschränken und zur Not in der anderen Umlaufrichtung zählen (der Cosinus ist als gerade Funktion nicht abhängig vom Vorzeichen seines Arguments).

Hier ist jedoch der orthogonale Anteil von  $\vec{w}$  wichtig; dieser wechselt für  $\varphi \in (\pi, 2\pi)$  die Richtung. Ein Beispiel ist in Abbildung 6.5 gezeigt.

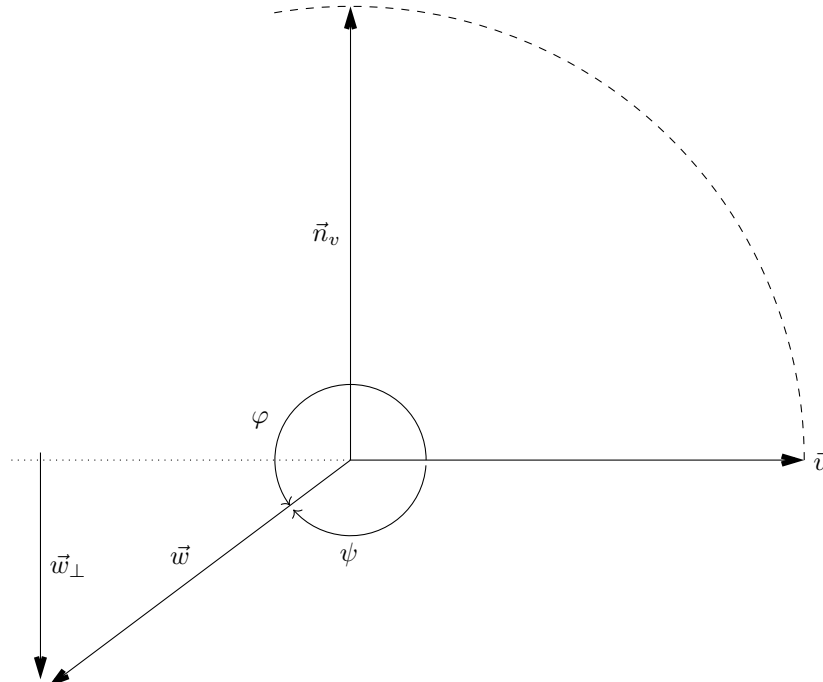


Abbildung 6.5: Zur Herleitung des Kreuzprodukts von  $\vec{v}$  und  $\vec{w}$ ;  $\varphi > \pi$ ; negativer Ergänzungswinkel  $\psi$  eingezeichnet

Es zeigt sich aber, dass die Formel für  $x$  trotzdem stimmt, denn das Skalarprodukt aus  $\vec{w}_\perp$  und  $\vec{n}_v$  ist nun durch die Antiparallelität negativ – aber das gleiche gilt auch für den Sinus des Winkels  $\varphi$ .

Wir wollen aber konzeptionell “den” Winkel zwischen zwei Vektoren immer so wählen, dass sein Betrag höchstens  $\pi$  ist. Daher müssten wir in dieser Situation mit dem im negativen Umlaufsinn

abgetragenen Ergänzungswinkel  $\psi$  (in Abbildung 6.5 abgesetzt eingezeichnet) arbeiten; dessen Vorzeichen ist negativ, aber sein Betrag passt zu dieser Forderung. Da der Sinus eine ungerade Funktion ist, liefert er für negative Winkel aus  $(-\pi, 0)$  auch das richtige (negative) Vorzeichen.

Ohnehin lässt sich auch nur so ein Parallelogramm aufspannen, denn jedes ebene Viereck hat eine Innenwinkelsumme von 360 Grad ( $2\pi$ );  $\varphi$  könnte also gar kein Innenwinkel eines Parallelogramms sein<sup>9</sup>.

Das Vorzeichen von  $x$  orientiert das Kreuzprodukt also stets so, dass die drei Vektoren  $\vec{v}, \vec{w}, \vec{v} \times \vec{w}$  in dieser Reihenfolge ein Rechtssystem bilden (siehe die Bemerkung bei Satz 6.17).

---

Soweit zur Situation der  $x_1, x_2$ -Ebene im  $\mathbb{R}^3$ . Falls die Vektoren  $\vec{v}, \vec{w}$  jedoch nicht in dieser Ebene liegen, müssen wir die Beobachtungen so formulieren, dass sie davon nicht konkret abhängig sind. Das Kreuzprodukt wird dann meist nicht die einfache Struktur von oben besitzen, in der nur eine Komponente von null verschieden sein konnte. Es gilt aber weiterhin:

**Satz 6.18** (Parallelogramm zweier Vektoren in  $\mathbb{R}^3$ ). *Für den Winkel  $\varphi$  zwischen zwei Vektoren  $\vec{v}, \vec{w} \in \mathbb{R}^3$  mit  $|\varphi| \in [0, \pi]$  gilt*

$$|\vec{v} \times \vec{w}| = |\vec{v}| \cdot |\vec{w}| \cdot |\sin \varphi|$$

*Der Betrag des Kreuzprodukts (euklidische Norm) entspricht hierbei genau dem Flächeninhalt des von  $\vec{v}, \vec{w}$  aufgespannten Parallelogramms.*

---

Eine wichtige Anwendung des Kreuzprodukts findet sich beim Rechnen mit Ebenen in  $\mathbb{R}^3$ ; siehe dazu der Exkurs A.3.

## 6.3 Lineare Abhängigkeit

Die Bemerkungen aus diesem Abschnitt gelten nicht nur für die euklidischen Vektorräume  $\mathbb{R}^n$ , sondern allgemein für alle  $\mathbb{K}$ -Vektorräume (nur für die Begriffe der Orthogonal- und Orthonormalbasis sind Räume mit Skalarprodukt voraus gesetzt). Als Beispiele betrachten wir allerdings ausschließlich  $\mathbb{R}^n$  mit dem kanonischen Skalarprodukt. Wir werden daher die Vektoren auch weiterhin allgemein mit Pfeilen notieren.

### 6.3.1 Definitionen

**Definition 6.19** (Linearkombination). *Gegeben  $n$  Vektoren  $\vec{v}_1, \dots, \vec{v}_n$  aus einem  $\mathbb{K}$ -Vektorraum  $V$  und  $n$  Zahlen  $c_1, \dots, c_n \in \mathbb{K}$ . Dann heißt die Summe der jeweils mit  $c_j$  skalierten Vektoren  $\vec{v}_j$  Linearkombination dieser Vektoren mit den Koeffizienten  $c_1, \dots, c_n$ , geschrieben*

$$\sum_{j=1}^n c_j \vec{v}_j = c_1 \cdot \vec{v}_1 + c_2 \cdot \vec{v}_2 + \dots + c_n \cdot \vec{v}_n$$

**Bemerkungen:**

- Die Operationen “+” und “ $\cdot$ ” sind die Vektorraumoperationen von  $V$ ; jede Linearkombination ist also wiederum ein Vektor aus  $V$ .
- In einer Linearkombination soll jeder der Vektoren höchstens einmal notiert sein. In Summen, bei denen ein Vektor mehrfach auftritt sind nach dem Distributivgesetz die zugehörigen Skalierungsfaktoren zu einem Koeffizienten zusammen zu addieren.
- Vektoren mit Koeffizient  $c_j = 0$  müssen nicht explizit angeschrieben werden, da sie durch Skalierung nach Satz 6.3 zu  $\vec{0}$  werden und damit zur beteiligten Summe neutral sind.

---

<sup>9</sup>Man bedenke, dass der gegenüber liegende Innenwinkel im Parallelogramm gleich groß wäre, und zweimal  $\varphi$  würde schon einen Wert über  $2\pi$  ergeben.

### Beispiele:

- Wir betrachten die folgenden Vektoren aus  $\mathbb{R}^2$ :

$$\vec{v}_1 := \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad \vec{v}_2 := \begin{pmatrix} 2 \\ 3 \end{pmatrix} \quad \text{und} \quad \vec{v}_3 := \begin{pmatrix} -1 \\ -1 \end{pmatrix}$$

Dann ist der Vektor

$$\vec{x} := \begin{pmatrix} 2 \\ 4 \end{pmatrix}$$

als aus den  $\vec{v}_j$  linear kombinierbar als

$$\vec{v}_1 + \vec{v}_2 + \vec{v}_3 = 1 \cdot \vec{v}_1 + 1 \cdot \vec{v}_2 + 1 \cdot \vec{v}_3 = \begin{pmatrix} 1+2-1 \\ 2+3-1 \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \end{pmatrix} = \vec{x}$$

- Die Linearkombination von  $\vec{x}$  aus dem vorigen Beispiel ist aber *nicht eindeutig*, denn man sieht, dass

$$2\vec{v}_1 = 2 \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \end{pmatrix} = \vec{x}$$

Hier handelt es sich um eine andere Linearkombination des gleichen Vektors  $\vec{x}$ . Die Koeffizienten für  $\vec{v}_2, \vec{v}_3$  sind hier beide 0; daher wurden diese beiden Vektoren nicht angeschrieben.

- Falls wir (mit gleichem  $\vec{x}$  wie oben) nur die beiden Vektoren

$$\vec{v}_1 := \begin{pmatrix} 1 \\ 2 \end{pmatrix} \quad \text{und} \quad \vec{v}_2 := \begin{pmatrix} 2 \\ 3 \end{pmatrix}$$

betrachten, ist dann  $\vec{x} = 2\vec{v}_1$  die einzige mögliche Linearkombination? Zur Beantwortung dieser Frage stellen wir eine *Vektorgleichung* auf, die uns sämtliche Linearkombinationen von  $\vec{x}$  liefert:

$$c_1 \begin{pmatrix} 1 \\ 2 \end{pmatrix} + c_2 \begin{pmatrix} 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \end{pmatrix}$$

Die Gleichung hat die beiden Unbekannten  $c_1, c_2 \in \mathbb{R}$ . Da  $\mathbb{R}^2$  ein kartesischer Produktraum ist, lässt sich die Vektorgleichung komponentenweise in (hier) zwei reelle Gleichungen zerlegen. Dabei erhalten wir ein *lineares Gleichungssystem* (mehr dazu im übernächsten Kapitel):

$$\cdots \Leftrightarrow \begin{array}{rcl} c_1 + 2c_2 & = & 2 \\ \wedge & 2c_1 + 3c_2 & = & 4 \end{array}$$

Wir erkennen an den Koeffizienten der beiden Gleichungen (die 1 bei  $c_1$  in der oberen Gleichung wurde nicht notiert) gerade die Komponenten der beiden Vektoren  $\vec{v}_1, \vec{v}_2$  wieder.

Nach Satz 1.30 dürfen wir Gleichungen addieren und skalieren – wir bilden also zunächst (in Gedanken) das  $(-2)$ -fache der ersten Gleichung und addieren dies dann zur zweiten hinzu. Dann verschwindet wegen  $(-2c_1 + 2c_1 = 0)$  der Koeffizient  $c_1$  in der zweiten Gleichung, und wir erhalten:

$$\cdots \Leftrightarrow \begin{array}{rcl} c_1 + 2c_2 & = & 2 \\ \wedge & 0 - c_2 & = & 0 \end{array}$$

Damit ist aber über die zweite Gleichung  $c_2 = 0$  eindeutig fest gelegt, und zusammen mit der oberen Gleichung folgt dann:

$$c_1 = 2 - 2c_2 = 2 - 2 \cdot 0 = 2$$

Es gibt nun also nur genau die eine Linearkombination  $\vec{x} = 2\vec{v}_1 (+0\vec{v}_2)$ , die wir bereits gefunden hatten.

- Wir können sogar zeigen, dass sich *jeder* Vektor aus  $\mathbb{R}^2$  aus

$$\vec{v}_1 := \begin{pmatrix} 1 \\ 2 \end{pmatrix} \quad \text{und} \quad \vec{v}_2 := \begin{pmatrix} 2 \\ 3 \end{pmatrix}$$

linear kombinieren lässt (und dies eindeutig). Hierzu betrachten wir die gleiche Rechnung wie im vorigen Beispiel, aber für

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

Das Gleichungssystem für die Koeffizienten  $c_1, c_2$  lautet dann:

$$\cdots \Leftrightarrow \quad \begin{array}{rcl} c_1 + 2c_2 & = & x_1 \\ \wedge & & 2c_1 + 3c_2 = x_2 \end{array}$$

Nach wie vor sind  $c_1, c_2$  die Unbekannten, die wir zu berechnen haben –  $x_1, x_2$  betrachten wir hingegen als feste gegebene Zahlenwerte. Wir gehen genau so vor wie eben und addieren das  $(-2)$ -fache der ersten Gleichung zur zweiten. Dadurch erhalten wir:

$$\cdots \Leftrightarrow \quad \begin{array}{rcl} c_1 + 2c_2 & = & x_1 \\ \wedge & & 0 - c_2 = -2x_1 + x_2 \end{array}$$

Hier lesen wir ab, dass  $c_2 = 2x_1 - x_2$  sein muss. Eingesetzt in die erste Gleichung folgt dann:

$$c_1 = x_1 - 2c_2 = x_1 - 2(2x_1 - x_2) = x_1 - 4x_1 + 2x_2 = -3x_1 + 2x_2$$

Also zusammen gefasst:

$$\cdots \Leftrightarrow \quad c_1 = -3x_1 + 2x_2 \quad \wedge \quad c_2 = 2x_1 - x_2$$

Für jeden Vektor  $\vec{x} \in \mathbb{R}^2$  erhalten wir somit *eindeutig* die Koeffizienten  $c_1, c_2$  der Linearkombination.

Probe:

$$\begin{aligned} c_1 \begin{pmatrix} 1 \\ 2 \end{pmatrix} + c_2 \begin{pmatrix} 2 \\ 3 \end{pmatrix} &= \begin{pmatrix} c_1 + 2c_2 \\ 2c_1 + 3c_2 \end{pmatrix} = \begin{pmatrix} (-3x_1 + 2x_2) + 2(2x_1 - x_2) \\ 2(-3x_1 + 2x_2) + 3(2x_1 - x_2) \end{pmatrix} \\ &= \begin{pmatrix} -3x_1 + 2x_2 + 4x_1 - 2x_2 \\ -6x_1 + 4x_2 + 6x_1 - 3x_2 \end{pmatrix} \\ &= \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad \checkmark \end{aligned}$$

Man überprüft leicht, dass für  $x_1 := 2$  und  $x_2 := 4$  genau die beiden Koeffizienten ermittelt werden wie im vorigen Beispiel.

Wir suchen nun noch die Linearkombination von

$$\vec{x} := \begin{pmatrix} 13 \\ 42 \end{pmatrix}$$

Nach obigen Gleichungen berechnen wir:

$$c_1 = -39 + 84 = 45 \quad \text{und} \quad c_2 = 26 - 42 = -16$$

Zur Probe:

$$45 \begin{pmatrix} 1 \\ 2 \end{pmatrix} - 16 \begin{pmatrix} 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 45 - 32 \\ 90 - 48 \end{pmatrix} = \begin{pmatrix} 13 \\ 42 \end{pmatrix} \quad \checkmark$$

Zum Abschluss erkennen wir noch, dass hier auch der Nullvektor  $\vec{0}$  aus  $\mathbb{R}^2$  eindeutig linear kombinierbar ist. Die beiden Koeffizienten sind dann:  $c_1 = c_2 = 0$ .

- Mit Blick auf die ersten beiden Beispiele mit

$$\vec{v}_1 := \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad \vec{v}_2 := \begin{pmatrix} 2 \\ 3 \end{pmatrix} \quad \text{und} \quad \vec{v}_3 := \begin{pmatrix} -1 \\ -1 \end{pmatrix}$$

fällt nun auf, dass der Nullvektor hier auf mindestens zwei Weisen linear kombiniert werden kann. Zum einen natürlich per  $c_1 = c_2 = c_3 = 0$  (die *triviale* Linearkombination von  $\vec{0}$ ). Aber auch durch Differenzbildung der beiden oben angegebenen Linearkombinationen, also per

$$-\vec{v}_1 + \vec{v}_2 + \vec{v}_3 = \vec{0} \quad (*)$$

Denn, die Zahlenwerte eingesetzt:

$$-\begin{pmatrix} 1 \\ 2 \end{pmatrix} + \begin{pmatrix} 2 \\ 3 \end{pmatrix} + \begin{pmatrix} -1 \\ -1 \end{pmatrix} = \begin{pmatrix} -1 + 2 - 1 \\ -2 + 3 - 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \vec{0}$$

Hier gibt es also auch eine *nichttriviale* Linearkombination des Nullvektors. Und damit übrigens unendlich viele weitere, denn die Vektorgleichung  $(*)$  lässt sich noch mit beliebigen Faktoren  $a \in \mathbb{R} \setminus \{0\}$  skalieren und führt dann auf den skalierten Nullvektor  $a\vec{0} = \vec{0}$ .



---

**Definition 6.20** (Spann). *Unter dem Spann der Vektoren  $\vec{v}_1, \dots, \vec{v}_n$  aus dem  $\mathbb{K}$ -Vektorraum  $V$  versteht man die Menge sämtlicher Linearkombinationen der Vektoren, notiert als*

$$\text{span}(\vec{v}_1, \dots, \vec{v}_n) := \left\{ \sum_{j=1}^n c_j \vec{v}_j \mid c_1, \dots, c_n \in \mathbb{R} \right\}$$

*Alternative Bezeichnung:* Lineare Hülle

**Bemerkungen:**

- Der Nullvektor ist stets ein Element des Spanns.
- Nach Definition 6.2 ist der Spann stets ein Unterraum von  $V$ .

**Beispiele:**

- Nach den Beispielen zu Linearkombinationen ist schon geklärt, dass

$$\text{span} \left( \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ 3 \end{pmatrix} \right) = \mathbb{R}^2$$

- Daran ändert sich auch nichts, wenn wir den dritten Vektor der obigen Beispiele mit dazu nehmen:

$$\text{span} \left( \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ 3 \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \end{pmatrix} \right) = \mathbb{R}^2$$

- Für einen gegebenen Vektor  $\vec{x} \in \mathbb{R}^n \setminus \{\vec{0}\}$  ist

$$\text{span}(\vec{x}) = \{c\vec{x} \mid c \in \mathbb{R}\}$$

die Darstellung einer *Geraden*, die durch den Ursprung  $O$  und den Punkt mit Ortsvektor  $\vec{x}$  verläuft. Denn denken wir uns den Vektor als Pfeil, so sind im Spann gerade sämtliche Skalierungen dieses Pfeils enthalten (positive und negative Richtung). Das sind genau die Punkte, die auf der erwähnten Ursprungsgeraden liegen.

- Es gilt in  $\mathbb{R}^3$ :

$$\text{span} \left( \begin{pmatrix} 1 \\ 3 \\ -2 \end{pmatrix}, \begin{pmatrix} -4 \\ -12 \\ 8 \end{pmatrix} \right) = \text{span} \left( \begin{pmatrix} 1 \\ 3 \\ -2 \end{pmatrix} \right)$$

Denn:

$$c_1 \begin{pmatrix} 1 \\ 3 \\ -2 \end{pmatrix} + c_2 \begin{pmatrix} -4 \\ -12 \\ 8 \end{pmatrix} = c_1 \begin{pmatrix} 1 \\ 3 \\ -2 \end{pmatrix} - 4c_2 \begin{pmatrix} 1 \\ 3 \\ -2 \end{pmatrix} = (c_1 - 4c_2) \begin{pmatrix} 1 \\ 3 \\ -2 \end{pmatrix} =: \tilde{c} \begin{pmatrix} 1 \\ 3 \\ -2 \end{pmatrix}$$

Nun findet man per  $\tilde{c} := c_1 - 4c_2$  zu jeder Kombination aus  $c_1, c_2 \in \mathbb{R}$  ein  $\tilde{c} \in \mathbb{R}$ , sodass die Linearkombination nur mit dem ersten Vektor schreibbar ist. Und umgekehrt findet sich zu jedem  $\tilde{c} \in \mathbb{R}$  per  $c_1 := \tilde{c}$  und  $c_2 := 0$  eine Linearkombination aus beiden Vektoren. Die linearen Hüllen sind also gleich.

---

Wir betrachten nun das Konzept der linearen Abhängigkeit von Vektoren, das zur praktischen Arbeit mit Vektorräumen fundamental wichtig ist.

**Definition 6.21** (Lineare Abhängigkeit). *Gegeben ein  $\mathbb{K}$ -Vektorraum  $V$  und eine nichtleere Menge  $A \subseteq V$  von Vektoren. Dann heißt  $A$  (bzw. es heißen die Vektoren aus  $A$ ) linear unabhängig ("l.u.") genau dann, wenn sich der Nullvektor  $\vec{0}$  nur trivial linear aus  $A$  kombinieren lässt per*

$$\left( \sum_j c_j \vec{a}_j = \vec{0} \right) \Rightarrow (\forall j : c_j = 0)$$

*Falls  $A$  nicht linear unabhängig ist, so heißt  $A$  linear abhängig ("l.a.").*

### Bemerkungen:

- Die triviale Linearkombination des Nullvektors ist für jede Menge  $A \subseteq V$  möglich, da  $\vec{0}$  stets Element des Spans von  $A$  ist. Für die lineare Unabhängigkeit ist jedoch entscheidend, dass es *nur* auf diese triviale Weise möglich ist, den Nullvektor linear zu kombinieren.
- Die Abkürzungen “l.a.” und “l.u.” gelten für uns als offiziell vereinbart und dürfen ab hier verwendet werden.

### Beispiele:

- In den Beispielen zur Definition 6.19 hatten wir schon erkannt, dass die Vektoren

$$\vec{a}_1 := \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad \vec{a}_2 := \begin{pmatrix} 2 \\ 3 \end{pmatrix} \quad \text{und} \quad \vec{a}_3 := \begin{pmatrix} -1 \\ -1 \end{pmatrix}$$

linear abhängig sind: der Nullvektor lässt sich nämlich (s.o.) *nichttrivial* linear kombinieren als

$$-\vec{a}_1 + \vec{a}_2 + \vec{a}_3 = \vec{0}$$

- Wir hatten außerdem bereits gesehen, dass sich aus den Vektoren

$$\vec{a}_1 := \begin{pmatrix} 1 \\ 2 \end{pmatrix} \quad \text{und} \quad \vec{a}_2 := \begin{pmatrix} 2 \\ 3 \end{pmatrix}$$

(also ohne  $\vec{a}_3$  von eben) jeder Vektor aus  $\mathbb{R}^2$  *eindeutig* linear kombinieren lässt, mit

$$c_1 = -3x_1 + 2x_2 \quad \wedge \quad c_2 = 2x_1 - x_2,$$

wenn  $x_1, x_2$  die beiden Komponenten des beliebigen Vektors  $\vec{x}$  sind. Dann gilt insbesondere, dass auch der Nullvektor eindeutig linear kombinierbar ist – wir hatten dies oben schon bemerkt –, nämlich per  $c_1 = c_2 = 0$ , also mit der trivialen Linearkombination. Somit sind  $\vec{a}_1, \vec{a}_2$  linear unabhängig.

- Die Vektoren

$$\vec{a}_1 := \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \quad \text{und} \quad \vec{a}_2 := \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}$$

sind linear unabhängig. Wir zeigen dies per Gleichungssystem wie gewohnt:

$$c_1 \vec{a}_1 + c_2 \vec{a}_2 = \vec{0} \Leftrightarrow \begin{array}{lcl} c_1 + c_2 & = & 0 \\ \wedge & c_1 + c_2 & = & 0 \\ \wedge & c_1 + 2c_2 & = & 0 \end{array}$$

Nun können wir eine der beiden oberen Gleichungen direkt auslassen (siehe die Rechenregeln zur Aussagenlogik in Satz 1.7, speziell das Idempotenzgesetz). Außerdem addieren wir die verbleibende obere Gleichung, mit  $(-1)$  skaliert, zur dritten:

$$\dots \Leftrightarrow \begin{array}{lcl} c_1 + c_2 & = & 0 \\ \wedge & 0 + c_2 & = & 0 \end{array}$$

Hieraus ergibt sich direkt:  $c_1 = -c_2 = 0$ ; damit sind  $\vec{a}_1, \vec{a}_2$  linear unabhängig.

- Wir betrachten die beiden Vektoren aus  $\mathbb{R}^3$  von eben, aber noch einen zusätzlichen Vektor, also

$$\vec{a}_1 := \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad \vec{a}_2 := \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} \quad \text{und} \quad \vec{a}_3 := \begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix}$$

Für den Test auf lineare Abhängigkeit erhalten wir die folgenden Gleichungen:

$$\begin{array}{lcl} c_1 + c_2 + c_3 & = & 0 \\ \wedge & c_1 + c_2 + c_3 & = & 0 \\ \wedge & c_1 + 2c_2 + 3c_3 & = & 0 \end{array}$$

Wieder verzichten wir auf eine der beiden oberen Gleichungen und addieren das Negative der verbleibenden oberen Gleichung zur dritten:

$$\cdots \Leftrightarrow \begin{array}{lcl} c_1 + c_2 + c_3 & = & 0 \\ \wedge & & \\ 0 + c_2 + 2c_3 & = & 0 \end{array}$$

Wir können oben noch  $c_2$  eliminieren, indem wir nun das Negative der zweiten Gleichung zur ersten addieren:

$$\cdots \Leftrightarrow \begin{array}{lcl} c_1 + 0 - c_3 & = & 0 \\ \wedge & & \\ 0 + c_2 + 2c_3 & = & 0 \end{array}$$

Hier lässt sich nichts weiter eliminieren, denn jede skalierte Addition von einer Gleichung zur anderen kann im besten Fall eine Variable dort eliminieren, bringt aber dafür zusätzlich wieder eine neue ein – es bleibt also eine Variable frei. Falls wir die freie Variable als  $c_3$  wählen, können wir aber die Gleichungen so umstellen, dass die beiden *abhängigen* Variablen  $c_1, c_2$  hierüber definierbar sind:

$$\cdots \Leftrightarrow c_1 = c_3 \quad \wedge \quad c_2 = -2c_3$$

Dann existieren aber nichttriviale Linearkombinationen des Nullvektors, denn für  $c_3 := 1$  folgt für unser Gleichungssystem:  $c_1 = 1$  und  $c_2 = -2$ . Probe:

$$\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - 2 \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 1 - 2 + 1 \\ 1 - 2 + 1 \\ 1 - 4 + 3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = \vec{0} \quad \checkmark$$

Der Nullvektor ist also nichttrivial linear kombinierbar, und somit sind  $\vec{a}_1, \vec{a}_2, \vec{a}_3$  linear abhängig.

- Die drei Vektoren

$$\vec{a}_1 := \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \vec{a}_2 := \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad \text{und} \quad \vec{a}_3 := \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

sind dagegen linear unabhängig. Man überprüft dies leicht, indem man wieder das lineare Gleichungssystem aufstellt, um  $\vec{0}$  zu kombinieren – es führt ohne Umformungen direkt auf  $c_1 = c_2 = c_3 = 0$ .

Besonders reizvoll ist hier auch, dass sich jeder Vektor

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

auf eindeutige Weise aus den  $\vec{a}_j$  linear kombinieren lässt, nämlich per

$$c_1 := x_1 \quad \text{und} \quad c_2 := x_2 \quad \text{und} \quad c_3 := x_3$$

Auch dies ergibt sich direkt aus der Vektorgleichung der Linearkombination. Diese Vektoren sind auch unter dem Namen *Standardbasisvektoren* bekannt; dazu mehr in Kürze.

Wir halten in zwei Sätzen noch einige wichtige Eigenschaften zur linearen Abhängigkeit fest.

**Satz 6.22** (Eigenschaften linear (un-)abhängiger Vektoren). *Gegeben ein  $\mathbb{K}$ -Vektorraum  $V$ . Dann gilt:*

1. Der Nullvektor  $\vec{0} \in V$  ist linear abhängig.
2. Jede Menge  $A \subseteq V$ , die  $\vec{0}$  enthält, ist ebenfalls linear abhängig.
3. Einzelne Vektoren  $\vec{a} \in V \setminus \{\vec{0}\}$  sind stets linear unabhängig.
4. Zwei Vektoren  $\vec{a}_1 \neq \vec{a}_2$  aus  $V \setminus \{\vec{0}\}$  sind linear abhängig genau dann, wenn es ein  $c \in \mathbb{K} \setminus \{0\}$  gibt, sodass  $\vec{a}_2 = c\vec{a}_1$ .
5. Fügt man zu einer linear abhängigen Menge  $A \subseteq V$  einen weiteren Vektor  $\vec{a} \in V$  hinzu, so ist auch  $A \cup \{\vec{a}\}$  linear abhängig.
6. Entfernt man aus einer linear unabhängigen Menge  $A \subseteq V$  einen Vektor  $\vec{a} \in A$ , so ist auch  $A \setminus \{\vec{a}\}$  linear unabhängig, sofern diese Menge nicht leer ist.

(Beweis: S. 325.)

**Bemerkung:** Eine linear unabhängige Menge von Vektoren kann durch Hinzufügen von weiteren Vektoren allerdings linear abhängig werden (in den Beispielen weiter oben haben wir dies schon beobachtet); genauso kann eine linear abhängige Menge von Vektoren durch Entfernen von Vektoren linear unabhängig werden (spätestens dann, wenn nur noch ein Vektor ungleich  $\vec{0}$  in der Menge verbleibt).

Außerdem stellen wir gesondert fest:

**Satz 6.23** (Eindeutigkeit von Linearkombinationen). *Gegeben ein  $\mathbb{K}$ -Vektorraum  $V$  und eine Menge  $A \subseteq V$  von Vektoren. Dann ist ein beliebiger Vektor  $\vec{x} \in \text{span}(A)$  genau dann eindeutig linear aus  $A$  kombinierbar, wenn  $A$  linear unabhängig ist.*

(Beweis: S. 326.)

**Bemerkung:** Falls also  $A$  linear unabhängig ist (und nur dann), ist jeder Vektor aus dem Spann von  $A$  eindeutig linear kombinierbar – nicht nur der Nullvektor. Das ist eine besonders wünschenswerte Eigenschaft, da jedem Vektor damit auch ein Tupel von Koeffizienten zugeordnet werden kann, das genau seine Linearkombination innerhalb des Spanns von  $A$  ergibt (und für keinen anderen Vektor aus  $\text{span}(A)$  steht).

### 6.3.2 Basis und Dimension

Offenbar ist es also besonders sinnvoll, Linearkombinationen aus linear unabhängigen Vektoren zu betrachten. Dabei gibt es bestimmte linear unabhängige Mengen, die obendrein noch den ganzen Vektorraum aufspannen:

**Definition 6.24** (Basis eines Vektorraums). *Sei  $V$  ein  $\mathbb{K}$ -Vektorraum. Eine Menge  $A \subseteq V$  heißt dann Basis von  $V$ , falls*

- $A$  linear unabhängig ist und
- $\text{span}(A) = V$

**Bemerkung:** Schon aus der linearen Unabhängigkeit folgt mit Satz 6.23, dass alle Linearkombinationen aus  $A$  eindeutig sind. Diese Eindeutigkeit ist, falls  $A$  Basis von  $V$  ist, dann für sämtliche Vektoren aus  $V$  gegeben.

Ist also eine Basis von  $V$  gegeben, so lässt sich jeder Vektor aus  $V$  über ein eindeutiges Koeffiziententupel als Linearkombination der Basisvektoren ausdrücken.

**Beispiele:**

- Wir hatten oben bereits gesehen:

$$\text{span} \left( \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ 3 \end{pmatrix} \right) = \mathbb{R}^2$$

Da die beiden Vektoren (ebenfalls oben schon nachgerechnet) linear unabhängig sind, bilden sie damit auch eine Basis von  $\mathbb{R}^2$ .

- Dagegen ist zwar auch

$$\text{span} \left( \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ 3 \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \end{pmatrix} \right) = \mathbb{R}^2,$$

aber wir hatten schon nachgerechnet, dass die drei Vektoren linear abhängig sind. Sie bilden also *keine* Basis von  $\mathbb{R}^2$ .

- Für den  $\mathbb{R}^3$  hatten wir in den Beispielen zu Definition 6.21 bereits gezeigt, dass

$$\vec{a}_1 := \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \vec{a}_2 := \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad \text{und} \quad \vec{a}_3 := \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

sowohl ganz  $\mathbb{R}^3$  aufspannen (da jeder beliebige  $\vec{x} \in \mathbb{R}^3$  daraus linear kombinierbar ist) als auch l.u. sind. Also bilden sie eine Basis von  $\mathbb{R}^3$ .

Wegen der Eindeutigkeit der Linearkombinationen ist es besonders lohnend, eine Basis für einen Vektorraum zu kennen. Dazu noch folgende

**Definition 6.25** (Koordinaten eines Vektors). *Sei  $A \subseteq V$  eine Basis des  $\mathbb{K}$ -Vektorraums  $V$ . Dann heißen für  $\vec{x} \in V$  die Koeffizienten  $c_j$  der Linearkombination*

$$\vec{x} = \sum_j c_j \vec{a}_j$$

Koordinaten von  $\vec{x}$  bezüglich der Basis  $A$ .

**Bemerkung:** Ein Beispiel aus  $\mathbb{R}^2$  mit der Entwicklung eines Vektors nach zwei verschiedenen Basen folgt in Kürze.

Hierzu gibt es noch einen grundlegenden Zusammenhang, den wir ohne Beweis akzeptieren:

**Satz 6.26** (Dimension eines Vektorraums). *Sei  $V$  ein  $\mathbb{K}$ -Vektorraum und  $A \subseteq V$  eine Basis von  $V$ . Dann hat jede andere Basis  $\tilde{A} \subseteq V$  von  $V$  die gleiche Mächtigkeit wie  $A$ , und der Raum  $V$  heißt*

- $n$ -dimensional, falls  $|A| = n \in \mathbb{N}_0$  oder
- unendlichdimensional, falls  $|A| = \infty$ .

**Bemerkungen:**

- Ein null-dimensionaler Vektorraum besteht nur aus dem Nullvektor. Dann gibt es nämlich gar keine Vektoren ungleich  $\vec{0}$ , die den Raum aufspannen könnten. Der Nullvektor selbst ist jedoch linear abhängig, sodass der Raum  $\{\vec{0}\}$  keine Basis haben kann. Sobald aber mindestens ein Vektor ungleich  $\vec{0}$  in  $V$  enthalten ist, muss eine Basis von  $V$  mindestens Dimension 1 besitzen.
- Wir befassen uns in dieser Vorlesung nicht mit unendlichdimensionalen Vektorräumen. Solche gibt es aber durchaus – z.B. die *Schwartz-Räume*, die für die *Fourier-Transformation* wichtig sind, oder auch viele *Hilbert-Räume* aus der Quantenmechanik.
- Siehe auch den Exkurs A.5 über den Vektorraum der Polynome.

Für einen endlichdimensionalen Vektorraum halten wir noch eine wichtige Tatsache fest:

**Satz 6.27** (Lineare Abhängigkeit von Vektoren aufgrund ihrer Anzahl). *Für einen  $n$ -dimensionalen Vektorraum  $V$  sind Teilmengen  $W \subseteq V$  mit mehr als  $n$  Elementen stets linear abhängig.*

(Beweis: S. 327.)

**Beispiel:** Wir verweisen auf das erste Beispiel zu Definition 6.21: Dort hatten wir drei Vektoren aus  $\mathbb{R}^2$  betrachtet; diese waren linear abhängig.

### 6.3.3 Kronecker-Delta und Standardbasis von $\mathbb{R}^n$

Wir betrachten nun noch eine spezielle Familie von Basen, die für die kartesischen Produkträume  $\mathbb{R}^n$  immer zur Verfügung steht – nämlich die *Standardbasen*. Zur einfacheren Definition dieser Basisvektoren zuvor noch diese

**Definition 6.28** (Kronecker-Symbol). <sup>10</sup> *Seien  $j, k$  Elemente einer Indexmenge. Dann ist das Kronecker-Symbol (oder: Kronecker-Delta) definiert per*

$$\delta_{jk} := \begin{cases} 1, & j = k \\ 0, & j \neq k \end{cases}$$

<sup>10</sup>L. Kronecker, dt. Mathematiker

Hiermit gelingt eine besonders kompakte Definition von Standardbasen in  $\mathbb{R}^n$ :

**Definition 6.29** (Standardbasis von  $\mathbb{R}^n$ ). Für die reellen Vektorräume  $\mathbb{R}^n$  bilden die Vektoren  $\vec{e}_j$  aus der Menge  $E_n \subseteq \mathbb{R}^n$  die Standardbasis von  $\mathbb{R}^n$  per

$$(\vec{e}_j)_k := \delta_{jk}$$

**Beispiele:**

- Für  $\mathbb{R}^2$  ist die Standardbasis also:

$$E_2 = \{\vec{e}_1, \vec{e}_2\} \quad \text{mit} \quad \vec{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{und} \quad \vec{e}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

- Für  $\mathbb{R}^3$  hatten wir die drei Vektoren der Standardbasis bereits kennen gelernt:

$$E_3 = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$$

Sie sind (in dieser Reihenfolge) als  $\vec{e}_1, \vec{e}_2, \vec{e}_3$  bezeichnet.

**Bemerkungen:**

- Wenn wir also weiter oben von den Komponenten eines Vektors  $\vec{x} \in \mathbb{R}^n$  gesprochen haben, war dies immer gleichbedeutend mit den Koordinaten von  $\vec{x}$  bezüglich der Standardbasis  $E_n$ .
- Dass es sich bei den so beschriebenen Vektoren tatsächlich um Basen handelt, lässt sich genau so zeigen, wie wir das oben für  $\mathbb{R}^3$  bereits ausgeführt hatten – wir verzichten deswegen hier auf einen formalen Beweis.

Nun reichen wir noch das schon angekündigte Beispiel nach:

**Beispiel:** In  $\mathbb{R}^2$  betrachten wir den Vektor

$$\vec{x} := \begin{pmatrix} 6 \\ 4 \end{pmatrix}$$

Seine Komponenten (also die Koordinaten bezüglich der Standardbasis) lauten also  $x_1 = 6$  und  $x_2 = 4$ , und es gilt:

$$\vec{x} = 6\vec{e}_1 + 4\vec{e}_2 = 6 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 4 \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 6 \\ 4 \end{pmatrix}$$

Dazu seien zwei Basen  $A := \{\vec{a}_1, \vec{a}_2\}$  und  $B := \{\vec{b}_1, \vec{b}_2\}$  gegeben mit

$$\vec{a}_1 := \begin{pmatrix} 4 \\ -1 \end{pmatrix}; \quad \vec{a}_2 := \begin{pmatrix} -1 \\ 3 \end{pmatrix}; \quad \vec{b}_1 := \begin{pmatrix} 3 \\ 1 \end{pmatrix}; \quad \vec{b}_2 := \begin{pmatrix} 2 \\ 4 \end{pmatrix}$$

(Man überprüfe zur Übung, dass es sich hierbei tatsächlich um Basen handelt, d.h. dass  $A$  und  $B$  l.u. sind und jeweils ganz  $\mathbb{R}^2$  aufspannen.)

Wir berechnen nun die Koordinaten von  $\vec{x}$  bezüglich der Basen  $A$ . Dazu stellen wir die Vektorgleichung auf:

$$\alpha_1 \begin{pmatrix} 4 \\ -1 \end{pmatrix} + \alpha_2 \begin{pmatrix} -1 \\ 3 \end{pmatrix} = \begin{pmatrix} 6 \\ 4 \end{pmatrix} \quad \Leftrightarrow \quad \begin{array}{rcl} 4\alpha_1 - \alpha_2 & = & 6 \\ -\alpha_1 + 3\alpha_2 & = & 4 \end{array}$$

Wir addieren das 4-fache der zweiten Gleichung zur ersten:

$$\dots \Leftrightarrow \quad \begin{array}{rcl} 0 + 11\alpha_2 & = & 22 \\ -\alpha_1 + 3\alpha_2 & = & 4 \end{array}$$

Daraus erhalten wir  $\alpha_2 = 2$  sowie  $\alpha_1 = 3\alpha_2 - 4 = 6 - 4 = 2$ . Somit lauten die Koordinaten von  $\vec{x}$  bezüglich  $A$ :  $\alpha_1 = 2$  und  $\alpha_2 = 2$ . Probe:

$$2\vec{a}_1 + 2\vec{a}_2 = 2 \begin{pmatrix} 4 \\ -1 \end{pmatrix} + 2 \begin{pmatrix} -1 \\ 3 \end{pmatrix} = \begin{pmatrix} 8-2 \\ -2+6 \end{pmatrix} = \begin{pmatrix} 6 \\ 4 \end{pmatrix} \quad \checkmark$$

Analog rechnet man die Koordinaten bezüglich  $B$  aus. Zu lösen ist das System

$$\begin{aligned} 3\beta_1 + 2\beta_2 &= 6 \\ \wedge \quad \beta_1 + 4\beta_2 &= 4 \end{aligned}$$

Addieren wir das  $(-3)$ -fache der zweiten Gleichung zur ersten, so erhalten wir:

$$\dots \Leftrightarrow \begin{aligned} 0 - 10\beta_2 &= -6 \\ \wedge \quad \beta_1 + 4\beta_2 &= 4 \end{aligned}$$

Damit erhalten wir für die Koordinaten von  $\vec{x}$  bezüglich  $B$ :

$$\beta_2 = \frac{6}{10} = \frac{3}{5} \quad \text{und} \quad \beta_1 = 4 - 4\beta_2 = \frac{20}{5} - \frac{12}{5} = \frac{8}{5}$$

Probe:

$$\frac{8}{5} \begin{pmatrix} 3 \\ 1 \end{pmatrix} + \frac{3}{5} \begin{pmatrix} 2 \\ 4 \end{pmatrix} = \frac{1}{5} \begin{pmatrix} 24+6 \\ 8+12 \end{pmatrix} = \frac{1}{5} \begin{pmatrix} 30 \\ 20 \end{pmatrix} = \begin{pmatrix} 6 \\ 4 \end{pmatrix} \quad \checkmark$$

Abbildung 6.6 zeigt die Basen, den Vektor  $\vec{x}$  sowie die Eckpunkte der beiden Parallelogramme, die die Koordinaten bezüglich der jeweiligen Basis definieren, in einem kartesischen Koordinatensystem. Die Basen  $A, B$  bilden jeweils nicht-kartesische Koordinatensysteme, da ihre Achsen nicht senkrecht aufeinander stehen.

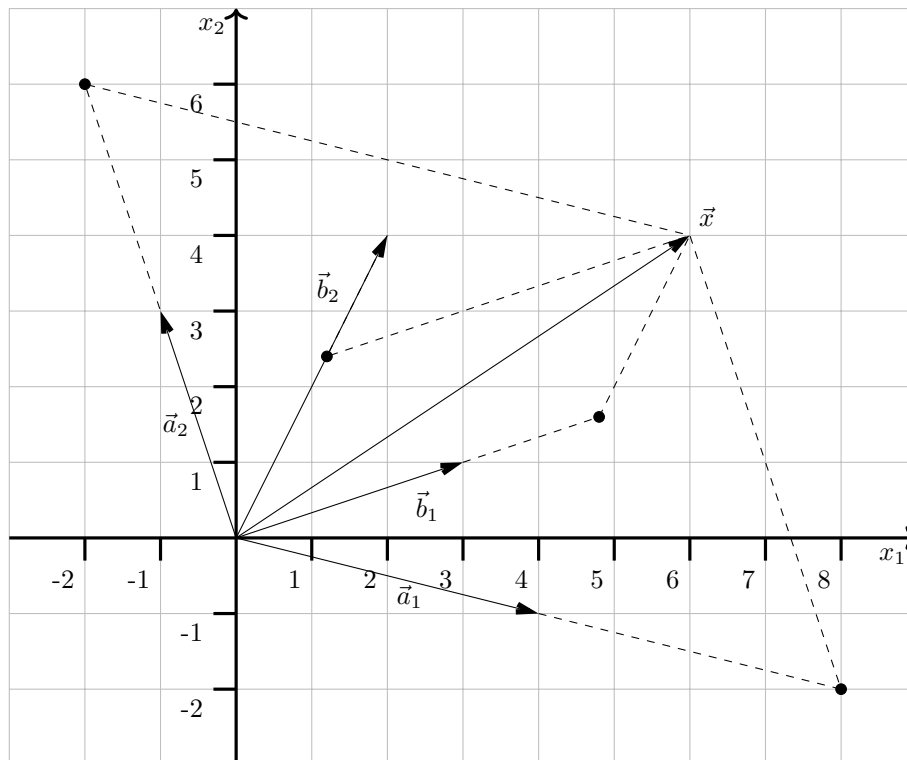


Abbildung 6.6: Vektor  $\vec{x}$  in zwei verschiedenen Basen  $\{\vec{a}_1, \vec{a}_2\}$  und  $\{\vec{b}_1, \vec{b}_2\}$

Es gilt noch folgendes Detail, das wir mit den Standardbasen und Satz 6.26 nun formulieren können:

**Satz 6.30** (Dimension des Vektorräume  $\mathbb{R}^n$ ). *Die kartesischen Produkträume  $\mathbb{R}^n$  mit  $n \in \mathbb{N}$  haben Dimension  $n$ .*

### 6.3.4 Orthogonale und Orthonormale Basen bei Innenprodukträumen

Falls ein  $\mathbb{K}$ -Vektorraum ein Skalarprodukt besitzt, gibt es spezielle Basen, in denen sich die Koordinaten von Vektoren besonders leicht berechnen lassen. Wir beschränken uns hier auf die reellen Vektorräume  $\mathbb{R}^n$  mit dem kanonischen Skalarprodukt (und der darüber induzierten euklidischen Norm) – aber die Beobachtungen sind auch für andere Vektorräume möglich<sup>11</sup>.

**Definition 6.31** (Orthogonalbasen und Orthonormalbasen von  $\mathbb{R}^n$ ). *Eine Basis  $A \subseteq \mathbb{R}^n$  heißt Orthogonalbasis, falls alle ihre  $n$  Basisvektoren paarweise orthogonal zueinander sind, d.h. falls für  $j, k \in \{1, \dots, n\}$  und  $j \neq k$  stets gilt:*

$$\vec{a}_j \bullet \vec{a}_k = 0$$

*$A$  heißt außerdem Orthonormalbasis (“ONB”), falls alle Basisvektoren normiert sind, d.h.*

$$|\vec{a}_j| = 1$$

**Bemerkungen:**

- Die Abkürzung “ONB” für Orthonormalbasen ist ab hier offiziell eingeführt.
- Für ONB gilt also insgesamt (unter Zuhilfenahme des Kroneckersymbols):

$$\vec{a}_j \bullet \vec{a}_k = \delta_{jk}$$

**Bemerkung:** Aus jeder Basis eines endlichdimensionalen Vektorraums lässt sich eine ONB konstruieren; siehe dazu den Exkurs A.4.

**Beispiele:**

- Die Basen  $A, B$  aus dem vorigen Beispiel (siehe Abbildung 6.6) sind weder ONB noch orthogonal.
- Die Standardbasen  $E_n$  von  $\mathbb{R}^n$  sind ONB, denn für beliebiges, aber festes  $j$  gilt:

$$\vec{e}_j \bullet \vec{e}_j = \sum_{k=1}^n (\vec{e}_j)_k \cdot (\vec{e}_j)_k = \sum_{k=1}^n \delta_{jk} \cdot \delta_{jk}$$

Nun ist aber  $\delta_{jk}$  nur für genau eines der  $n$  Werte von  $k$  ungleich 0, nämlich für  $k = j$ . Also gibt es in der rechten Summe genau einen Beitrag mit dem Wert  $1 \cdot 1 = 1$ ; die anderen Beiträge verschwinden. Entsprechend:

$$\vec{e}_j \bullet \vec{e}_j = 1$$

Außerdem ist für  $k \neq j$  (beide fest):

$$\vec{e}_j \bullet \vec{e}_k = \sum_{l=1}^n \delta_{jl} \cdot \delta_{kl}$$

Nun ist  $\delta_{jl}$  nur für  $l = j$  ungleich 0. Da allerdings  $k \neq j$ , ist in diesem Fall  $\delta_{kl} = 0$ . Analog verhält es sich für  $l = k$ . Alle anderen Beiträge verschwinden ohnehin, sodass insgesamt (und mit der Normiertheit von oben) gilt:

$$\vec{e}_j \bullet \vec{e}_k = \delta_{jk}$$

- Für  $\mathbb{R}^3$  rechnen wir explizit nach, dass  $\vec{e}_3$  normiert ist:

$$|\vec{e}_3| = \sqrt{\vec{e}_3 \bullet \vec{e}_3} = \sqrt{\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \bullet \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}} = \sqrt{0 + 0 + 1} = \sqrt{1} = 1$$

Und, dass  $\vec{e}_1 \perp \vec{e}_3$ :

$$\vec{e}_1 \bullet \vec{e}_3 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \bullet \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = 0 + 0 + 0 = 0$$

---

<sup>11</sup>Stichwort z.B. *Fourier-Entwicklung*



### 6.3.5 Koordinaten eines Vektors in einer Orthonormalbasis

Angenommen, wir hätten zu einem Vektorraum eine ONB  $A$  gegeben. Jeder Vektor  $\vec{x}$  besitzt also eindeutige Koordinaten bezüglich  $A$ . Außerdem gibt es ein Skalarprodukt (sonst könnte von einer ONB nicht die Rede sein). Wir behandeln hier wieder  $\mathbb{R}^n$  mit dem kanonischen Skalarprodukt “ $\bullet$ ” und bilden das Skalarprodukt von  $\vec{x}$  mit einem beliebigen Basisvektor  $\vec{a}_k \in A$ . Dazu verwenden wir noch die *eindeutige* Linearkombination von  $\vec{x}$  in  $A$ . Somit erhalten wir:

$$\vec{x} \bullet \vec{a}_k = \left( \sum_j c_j \vec{a}_j \right) \bullet \vec{a}_k = \left( \sum_j c_j (\vec{a}_j \bullet \vec{a}_k) \right)$$

Da aber  $A$  eine ONB ist, gilt für die Klammer im rechten Ausdruck:

$$\dots = \left( \sum_j c_j \delta_{jk} \right) = c_k$$

Von der Summe bleibt aufgrund des Kroneckerdeltas nur genau ein Beitrag übrig, nämlich der für  $j = k$ ; sein Zahlenwert ist  $c_k \cdot 1 = c_k$ .

Wir halten dies fest im

**Satz 6.32** (Koordinaten bezüglich ONBs). *Für eine ONB  $A = \{\vec{a}_1, \dots, \vec{a}_n\}$  des reellen Vektorraums  $\mathbb{R}^n$  ( $n \in \mathbb{N}$ ) berechnen sich die Koordinaten eines Vektors  $\vec{x} \in \mathbb{R}^n$  bezüglich  $A$  mit dem kanonischen Skalarprodukt per*

$$\vec{x} = \sum_j c_j \vec{a}_j \quad \text{mit} \quad \forall j : c_j = \vec{x} \bullet \vec{a}_j$$

**Bemerkungen:**

- Hier muss also kein Gleichungssystem gelöst werden – die Berechnung des kanonischen Skalarprodukts ist meistens leichter.
- Da die Standardbasis eine ONB ist, gilt natürlich auch  $x_j = (\vec{x})_j = \vec{x} \bullet \vec{e}_j$ .

---

Darüberhinaus gilt noch eine Eigenschaft, die wir im kanonischen Skalarprodukt bereits kennen gelernt hatten, für sämtliche ONB:

**Satz 6.33** (Skalarprodukt von Vektoren bezüglich einer ONB). *Seien  $c_1, \dots, c_n$  die Koordinaten eines Vektors  $\vec{x} \in \mathbb{R}^n$  bezüglich einer ONB  $A$ ; außerdem  $d_1, \dots, d_n$  die Koordinaten von  $\vec{y} \in \mathbb{R}^n$  bezüglich  $A$ . Dann gilt für das Skalarprodukt:*

$$\vec{x} \bullet \vec{y} = \sum_j c_j d_j$$

(Beweis: S. 327.)

**Bemerkung:** Also ist es unerheblich, bezüglich welcher ONB wir die Vektoren betrachten; die Formel für das Skalarprodukt hat die gleiche Struktur wie in der Standardbasis. (Die Koordinaten müssen natürlich bezüglich der aktuellen ONB genommen werden!)

Wir begegnen dem Skalarprodukt unter anderem im Kapitel 9 noch einmal und werden dort fest stellen, dass sogar der Zahlenwert des Skalarprodukts unabhängig von der konkreten ONB ist.<sup>12</sup>

---

<sup>12</sup>Die *orthogonalen Transformationen* sind nämlich genau die, welche den Wert des Skalarprodukts erhalten – und alle ONB lassen sich durch Drehungen (ggf. Drehspiegelungen) aus der Standardbasis erzeugen. Alle Dreh(spiegel)ungen sind orthogonale Transformationen. Hierzu benötigen wir aber noch die Methoden der Matrizenrechnung.

# Kapitel 7

## Lineare Algebra: Lineare Abbildungen und Matrizen

Wir führen zunächst den Begriff der *linearen Abbildung* zwischen reellen Vektorräumen ein. Auch hier beschränken wir uns auf die kartesischen Produkträume  $\mathbb{R}^n$ .

Das Konzept der reellen *Matrix* zur Beschreibung von linearen Abbildungen führt uns auf ein neues Produkt, nämlich das zwischen einer Matrix und einem Vektor. Eine Verallgemeinerung hiervon ist das *Matrizenprodukt* – dies ist im Vergleich zum Skalarprodukt und zur skalaren Multiplikation eine arithmetische Neuheit; hier werden Paare von Matrizen auf Matrizen abgebildet. Dadurch lassen sich dann wiederum algebraische Strukturen definieren – auch die *inversen Matrizen* spielen hier eine wichtige Rolle; letztere führen wir in diesem Kapitel allerdings nur ein und geben erst in Kapitel 9 mit den Mitteln von Kapitel 8 eine Methode, inverse Matrizen systematisch zu bestimmen.

Eine weitere wichtige Operation für Matrizen ist die *Transposition*. Wir werden über die Konzepte der transponierten Matrix und der Matrizenmultiplikation eine alternative Definition für das kanonische Skalarprodukt finden.

### 7.1 Lineare Abbildungen

Wir betrachten nun lineare Abbildungen von  $\mathbb{R}^n$  nach  $\mathbb{R}^m$  mit  $m, n \in \mathbb{N}$ , also Abbildungen zwischen  $\mathbb{R}$ -Vektorräumen.

**Definition 7.1** (Lineare Abbildung). Sei  $\vec{f}: \mathbb{R}^n \rightarrow \mathbb{R}^m$  eine Abbildung.  $\vec{f}$  heißt lineare Abbildung, falls für alle  $a, b \in \mathbb{R}$  sowie  $\vec{x}, \vec{y} \in \mathbb{R}^n$  gilt:

$$\vec{f}(a\vec{x} + b\vec{y}) = a\vec{f}(\vec{x}) + b\vec{f}(\vec{y})$$

**Bemerkungen:**

- Da der Bildbereich dieser Funktionen ein Vektorraum ist, notieren wir auch am Funktionsnamen einen Vektorpfeil.
- Linear ist eine Abbildung also dann, wenn die Vektorraum-Operationen (Addition und Skalierung von Vektoren) mit der Anwendung der Funktion vertauschbar sind.

**Beispiele:**

- Wir betrachten die Abbildung  $\vec{f}: \mathbb{R}^2 \rightarrow \mathbb{R}^3$  mit

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \begin{pmatrix} x_1 + 3x_2 \\ x_1 - x_2 \\ 2x_1 \end{pmatrix}$$

Zum Test auf Linearität berechnen wir

$$\begin{aligned}
 \vec{f}(a\vec{x} + b\vec{y}) &= \vec{f}\left(a \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + b \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) \\
 &= \vec{f}\left(\begin{pmatrix} ax_1 + by_1 \\ ax_2 + by_2 \end{pmatrix}\right) \\
 &= \begin{pmatrix} (ax_1 + by_1) + 3(ax_2 + by_2) \\ (ax_1 + by_1) - (ax_2 + by_2) \\ 2(ax_1 + by_1) \end{pmatrix} \\
 &= \begin{pmatrix} a(x_1 + 3x_2) + b(y_1 + 3y_2) \\ a(x_1 - x_2) + b(y_1 - y_2) \\ a(2x_1) + b(2y_1) \end{pmatrix} \\
 &= \begin{pmatrix} a(x_1 + 3x_2) \\ a(x_1 - x_2) \\ a(2x_1) \end{pmatrix} + \begin{pmatrix} b(y_1 + 3y_2) \\ b(y_1 - y_2) \\ b(2y_1) \end{pmatrix} \\
 &= a \begin{pmatrix} x_1 + 3x_2 \\ x_1 - x_2 \\ 2x_1 \end{pmatrix} + b \begin{pmatrix} y_1 + 3y_2 \\ y_1 - y_2 \\ 2y_1 \end{pmatrix} \\
 &= a\vec{f}(\vec{x}) + b\vec{f}(\vec{y})
 \end{aligned}$$

Die Funktion  $\vec{f}$  ist also tatsächlich eine lineare Abbildung.

(Das Ausklammern von  $a, b$  im vorletzten Schritt wird üblicherweise direkt zusammen mit dem Aufspalten des Bildvektors in die Summe aus zwei Vektoren im Schritt davor ausgeführt. Hier nur der Übersichtlichkeit wegen nacheinander notiert.)

- Wir testen die eben nachgerechnete Linearität der Beispielfunktion noch mit den Vektoren

$$\vec{x}_1 := \begin{pmatrix} 2 \\ -3 \end{pmatrix}, \quad \vec{x}_2 := 3\vec{x}_1 = \begin{pmatrix} 6 \\ -9 \end{pmatrix} \quad \text{und} \quad \vec{x}_3 := \vec{x}_1 - \vec{x}_2 = \begin{pmatrix} -4 \\ 6 \end{pmatrix}$$

Dabei erwarten wir, dass  $\vec{f}(\vec{x}_2) = 3\vec{f}(\vec{x}_1)$  und  $\vec{f}(\vec{x}_3) = \vec{f}(\vec{x}_1) - \vec{f}(\vec{x}_2)$ .

$$\vec{f}(\vec{x}_1) = \vec{f}\left(\begin{pmatrix} 2 \\ -3 \end{pmatrix}\right) = \begin{pmatrix} 2 + 3 \cdot (-3) \\ 2 - (-3) \\ 2 \cdot 2 \end{pmatrix} = \begin{pmatrix} -7 \\ 5 \\ 4 \end{pmatrix}$$

$$\vec{f}(\vec{x}_2) = \vec{f}\left(\begin{pmatrix} 6 \\ -9 \end{pmatrix}\right) = \begin{pmatrix} 6 + 3 \cdot (-9) \\ 6 - (-9) \\ 2 \cdot 6 \end{pmatrix} = \begin{pmatrix} -21 \\ 15 \\ 12 \end{pmatrix} = 3 \begin{pmatrix} -7 \\ 5 \\ 4 \end{pmatrix} \quad \checkmark$$

$$\vec{f}(\vec{x}_3) = \vec{f}\left(\begin{pmatrix} -4 \\ 6 \end{pmatrix}\right) = \begin{pmatrix} -4 + 3 \cdot 6 \\ -4 - 6 \\ 2 \cdot (-4) \end{pmatrix} = \begin{pmatrix} 14 \\ -10 \\ -8 \end{pmatrix} = \begin{pmatrix} -7 - (-21) \\ 5 - 15 \\ 4 - 12 \end{pmatrix} = \begin{pmatrix} -7 \\ 5 \\ 4 \end{pmatrix} - \begin{pmatrix} -21 \\ 15 \\ 12 \end{pmatrix} \quad \checkmark$$

Übrigens ist wegen  $\vec{x}_3 = -2\vec{x}_1$  auch  $\vec{f}(\vec{x}_3) = -2\vec{f}(\vec{x}_1)$ .

- Wir betrachten die Abbildung  $\vec{f}: \mathbb{R}^4 \rightarrow \mathbb{R}^1$  per

$$\vec{x} \mapsto \begin{pmatrix} 2 \\ -3 \\ 1 \\ 7 \end{pmatrix} \bullet \vec{x} = 2x_1 - 3x_2 + x_3 + 7x_4$$

Also ein Skalarprodukt mit einem konstanten Vektor. Wir wissen aufgrund der Bilinearität des Skalarprodukts eigentlich bereits, dass die Abbildung linear ist, rechnen aber trotzdem nochmal nach (hier ohne Vektorpfeil über  $f$ , da nach  $\mathbb{R}$  abgebildet wird – falsch wäre der

Pfeil aber natürlich in diesem Kontext nicht):

$$\begin{aligned}
 f(a\vec{x} + b\vec{y}) &= f\left(\begin{pmatrix} ax_1 + by_1 \\ ax_2 + by_2 \\ ax_3 + by_3 \\ ax_4 + by_4 \end{pmatrix}\right) \\
 &= 2(ax_1 + by_1) - 3(ax_2 + by_2) + (ax_3 + by_3) + 7(ax_4 + by_4) \\
 &= a(2x_1 - 3x_2 + x_3 + 7x_4) + b(2y_1 - 3y_2 + y_3 + 7y_4) \\
 &= af(\vec{x}) + bf(\vec{y})
 \end{aligned}$$

Auch diese Abbildung ist also linear. Zur Übung setze man gerne einige Vektoren  $\vec{x}$  ein und überprüfe das lineare Verhalten wie im vorigen Beispiel.

- Die Abbildung  $\vec{f}: \mathbb{R}^3 \rightarrow \mathbb{R}^5$  per

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mapsto \begin{pmatrix} 3x_1 + 2x_2x_3 \\ \sqrt{|x_2 - x_1|} \\ \cos(2x_3) - x_1 \\ x_2^2 \\ 3 \end{pmatrix}$$

ist *nicht* linear. Es reicht, dies an einer der Komponenten zu zeigen; wir nehmen beispielsweise die erste:

$$\begin{aligned}
 (\vec{f}(a\vec{x} + b\vec{y}))_1 &= 3(ax_1 + by_1) + 2(ax_2 + by_2)(ax_3 + by_3) \\
 &= 3ax_1 + 3by_1 + 2a^2x_2x_3 + 2abx_2y_3 + 2bay_2x_3 + 2b^2y_2y_3 \\
 &\neq 3ax_1 + 3by_1 + 2ax_2x_3 + 2by_2y_3 \\
 &= a(3x_1 + 2x_2x_3) + b(3y_1 + 2y_2y_3) \\
 &= a\vec{f}(\vec{x})_1 + b\vec{f}(\vec{y})_1
 \end{aligned}$$

Man bestätige dies gerne mit Zahlenbeispielen zur Übung.

Für die vierte Komponente sieht man aus der binomischen Formel (Satz 1.31) direkt, dass im allgemeinen  $(x_2 + y_2)^2 \neq x_2^2 + y_2^2$ . Und selbst wenn hier eine Gleichheit erfüllt wäre, gäbe es noch das Problem der Skalierungsfaktoren, die sich mit quadrieren und danach – eben – quadriert vorlägen, nicht mehr linear.

Für die fünfte Komponente mit dem konstanten Wert 3 lässt sich die Bedingung für lineare Abbildungen ebenfalls nicht erfüllen: Sie müsste sich, wenn das Funktionsargument z.B. verdoppelt würde, auf 6 verdoppeln, bleibt aber konstant beim Wert 3.

Auch die anderen Operationen in den Komponenten des Bildvektors sind nicht linear. Als Faustregel halten wir für  $\mathbb{R}^n$  fest:

Linear ist eine Abbildung nur dann, wenn im Bildvektor die Komponenten des Urbildvektors rein linearen Operationen unterliegen, d.h. mit konstanten Faktoren skaliert und/oder summiert werden.

Man beachte, dass es in einer Prüfungsaufgabe allerdings durchaus trotzdem gefordert sein kann, die Linearität explizit nachzurechnen, selbst wenn diese Faustregel erfüllt ist!

## 7.2 Matrizen

Nun führen wir den Begriff der reellen *Matrix* ein. Zunächst wirkt dieser eher künstlich, aber wir stellen direkt danach den Zusammenhang zu linearen Abbildungen her.

### 7.2.1 Definition

**Definition 7.2** (Reelle Matrix). Für  $m, n \in \mathbb{N}$  heißt ein rechteckiges Schema  $A$  aus  $m$  Zeilen und  $n$  Spalten, welches  $m \cdot n$  Einträge aus  $\mathbb{R}$  enthält, reelle  $(m \times n)$ -Matrix:

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

Der erste Index einer Komponente (auch: Matrixelement)  $A_{j,k}$  gibt die Zeile an, der zweite die Spalte.

Die Menge aller reellen  $(m \times n)$ -Matrizen heißt  $\mathbb{R}^{(m,n)}$ .

Die Spaltendarstellung von  $A$  besteht aus  $n$  Spaltenvektoren und lautet

$$A = (\vec{a}_1 \quad \vec{a}_2 \quad \cdots \quad \vec{a}_n)$$

Dabei ist die  $k$ -te Komponente des  $j$ -ten Spaltenvektors das Matrixelement aus der  $k$ -ten Zeile und  $j$ -ten Spalte:

$$(\vec{a}_j)_k = A_{k,j}$$

Zwei Matrizen sind gleich, wenn sie die gleiche Dimensionierung (Zeilen- und Spalten-Anzahl) besitzen und alle korrespondierenden Matrixelemente (mit gleichen Indexpaaren) gleich sind.

**Bemerkung:** Die Bezeichnung “ $(m \times n)$ -Matrix” verwendet absichtlich dasselbe Zeichen wie beim Kreuzprodukt von Vektoren (man spricht auch wörtlich von einer “ $m$ -kreuz- $n$ -Matrix”), da auch das Kreuzprodukt nicht kommutativ ist. Wie dort kommt es hier auf die Reihenfolge an – eine Matrix mit drei Zeilen und sieben Spalten ist strukturell etwas ganz anderes als eine solche mit sieben Zeilen und drei Spalten.

**Beispiele:** Wir listen einige Matrizen auf und geben einige einzelne Matrixelemente explizit an.

- Die Matrix

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

ist aus  $\mathbb{R}^{(3,4)}$ . Einige Elemente sind:

$$A_{2,1} = 2, \quad A_{3,2} = 2, \quad A_{2,2} = 1, \quad A_{2,3} = -3, \quad \text{und} \quad A_{3,4} = 7$$

- Die quadratische Matrix

$$B := \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 4 \\ 4 & 2 & 1 \end{pmatrix}$$

ist aus  $\mathbb{R}^{(3,3)}$ . Ihre *Diagonalelemente*  $B_{1,1}, B_{2,2}, B_{3,3}$  haben alle den Wert 1.

- Das Objekt

$$C := (2 \quad -3 \quad 4 \quad 1)$$

ist ebenfalls eine Matrix, wobei  $C \in \mathbb{R}^{(1,4)}$ . Solche Matrizen heißen (s.u.) auch *Zeilenvektoren*. Beispielsweise ist  $C_{1,3} = 4$ .

- Ebenso ist das Objekt

$$D := \begin{pmatrix} 3 \\ 2 \\ 1 \\ 7 \end{pmatrix}$$

eine Matrix aus  $\mathbb{R}^{(4,1)}$ . Sie enthält genau einen Spaltenvektor aus  $\mathbb{R}^4$ , der identisch notiert wird. Streng genommen sind dies aber zwei verschiedene Dinge – ähnlich wie ein Vektor nicht das Gleiche ist wie ein Tupel, obwohl  $\mathbb{R}^4$  auch als kartesisches Produkt reeller Zahlen aufgefasst werden könnte (dann hätten wir aber keine zugehörigen Vektorraumoperationen!).

Diese Ungenauigkeit riskieren wir (wie bei vorigen solchen Situationen) in der Erwartung, dass aus dem Kontext jeweils klar ist, ob von Vektoren, Matrizen, Vektormengen oder Vektorräumen die Rede ist.

Wenn wir  $D$  als Matrix begreifen, ist z.B.  $D_{3,1} = 1$ . Das entspricht, falls  $\vec{d}_1$  der (eine) Spaltenvektor von  $D$  ist, der Komponente  $(\vec{d}_1)_3$ .

- Sogar dieses Objekt ist eine Matrix:

$$E := (42)$$

Es handelt sich um die  $(1 \times 1)$ -Matrix mit dem einzigen Element  $E_{1,1} = 42$ . Solche  $\mathbb{R}^{(1,1)}$ -Matrizen werden uns in Zukunft nicht oft begegnen, weil es nur wenig Sinn macht, um eine

reelle Zahl herum noch Struktur zu deklarieren, die dann keine sichtbaren Effekte hat. Aber es ist möglich, solche Matrizen als Rechenergebnis bei Matrixprodukten zu erhalten (das wird bei der Wiederentdeckung des kanonischen Skalarprodukts geschehen, s.u.). In solchen Fällen werden wir diese Matrizen direkt mit den reellen Zahlen in ihrem (einzigen) Element identifizieren.

## 7.2.2 Zusammenhang mit linearen Abbildungen

Wir verbinden nun das eben eingeführte Konzept der reellen Matrizen mit den linearen Abbildungen. Dazu betrachten wir eine allgemeine lineare Abbildung  $\vec{f}: \mathbb{R}^n \rightarrow \mathbb{R}^m$ ; als Basen für die beiden Vektorräume wählen wir die jeweiligen Standardbasen  $E_n, E_m$ .

Für den Vektor  $\vec{x} \in \mathbb{R}^n$  gilt also:

$$\vec{x} = \sum_{k=1}^n x_k \vec{e}_k$$

Dann setzen wir diese Entwicklung von  $\vec{x}$  nach den Basisvektoren in die lineare Abbildung ein und benutzen Definition 7.1 – denn  $\vec{x}$  lässt sich als Linearkombination von (jeweils mit  $x_k$  skalierten) Basisvektoren schreiben:

$$\vec{f}(\vec{x}) = \vec{f}\left(\sum_{k=1}^n x_k \vec{e}_k\right) = \sum_{k=1}^n x_k \vec{f}(\vec{e}_k)$$

Weil also die Abbildung  $\vec{f}$  linear ist, lassen sich die konkreten Koordinaten (die reelle Zahlen sind) als lineare Skalierungsfaktoren aus der Funktionsanwendung heraus ziehen, und es bleibt eine Linearkombination von den Bildern der Einheitsvektoren aus  $E_n$  übrig (man beachte nochmals das obige Negativbeispiel zu Definition 7.1, um einzusehen, dass das so nur für lineare Abbildungen möglich ist).

Die  $j$ -te Komponente vom Bildvektor,  $f_j$ , ist dann (in der Standardbasis  $E_m$  von  $\mathbb{R}^m$ ) gegeben durch

$$f_j(\vec{x}) = \left(\vec{f}(\vec{x})\right)_j = \sum_{k=1}^n x_k \left(\vec{f}(\vec{e}_k)\right)_j$$

Die eingeklammerten Ausdrücke  $f_j(\vec{e}_k)$  haben zwei Indices, wobei  $j$  zwischen 1 und  $m$  wählbar ist und  $k$  jeweils von 1 bis  $n$  läuft. Das erlaubt uns, diese (reellen) Zahlen als Matrix wie in Definition 7.2 zu schreiben, per

$$F_{j,k} := \left(\vec{f}(\vec{e}_k)\right)_j = f_j(\vec{e}_k)$$

Damit wird dann:

$$f_j(\vec{x}) = \sum_{k=1}^n F_{j,k} \cdot x_k$$

Offenbar genügt es also, die Abbildungen der Basisvektoren aus  $E_n$  (dem Urbildraum) zu bestimmen und komponentenweise in der Matrix  $F$  zu sammeln. Die Eigenschaften der konkreten Abbildung  $\vec{f}$  sind dann in den Matrixelementen von  $F$  fest gehalten – daher spricht man auch von der *zugehörigen Abbildungsmatrix*.

Die Abbildung eines *beliebigen* Vektors aus  $\mathbb{R}^n$  ist danach durch die Linearkombination der Koordinaten von  $\vec{x}$ , skaliert mit den Matrixelementen, berechenbar. Dieses Verfahren funktioniert für *jede* lineare Abbildung  $\vec{f}$ .

Man beachte die Spaltendarstellung der Abbildungsmatrix nach Definition 7.2 – in Spalte  $k$  stehen die Komponenten  $f_j(\vec{e}_k)$ . Insgesamt entspricht also die  $k$ -te Spalte der Matrix genau dem Bild des  $k$ -ten Basisvektors aus  $E_n$ :

$$F = \begin{pmatrix} \vec{f}(\vec{e}_1) & \vec{f}(\vec{e}_2) & \cdots & \vec{f}(\vec{e}_n) \end{pmatrix} \in \mathbb{R}^{(m,n)}$$

Wir fassen zusammen:

**Definition 7.3** (Abbildungsmatrix). *Für jede lineare Abbildung  $\vec{f}: \mathbb{R}^n \rightarrow \mathbb{R}^m$ , betrachtet in den Standardbasen  $E_n, E_m$ , ist eine eindeutige Abbildungsmatrix  $F \in \mathbb{R}^{(m,n)}$  gegeben, welche  $\vec{f}$  beschreibt. Dabei ist das Matrixelement  $F_{j,k}$  gegeben durch die  $j$ -te Komponente des Bildes von  $\vec{e}_k$  unter der Abbildung  $\vec{f}$ :*

$$F_{j,k} = f_j(\vec{e}_k)$$

Für die  $j$ -te Komponente ( $j \in \{1, \dots, m\}$ ) des Bildvektors gilt dann:

$$f_j(\vec{x}) = \sum_{k=1}^n F_{j,k} x_k$$

### 7.2.3 Bestimmung der Abbildungsmatrix in der Praxis

Bevor wir die mathematischen Beobachtungen zu linearen Abbildungen weiter führen, wollen wir anhand der beiden Beispiele zu Definition 7.1 zwei (gleichwertige) Verfahren betrachten, wie die Abbildungsmatrix einer beliebigen linearen Abbildung bestimmt werden kann.

Wir beginnen mit der formaleren Methode: Und zwar hatten wir gesehen, dass die Spalten der Abbildungsmatrix gerade den Bildern der jeweiligen Einheitsvektoren entsprechen (von links nach rechts gezählt). Nun sind die Einheitsvektoren der Standardbasis an dieser Stelle besonders dankbar, da sie jeweils bis auf genau eine Komponente aus Nullen bestehen. Wie sich das auswirkt, sehen wir nun:

**Beispiele:**

- Wir betrachten erneut die Abbildung  $\vec{f}: \mathbb{R}^2 \rightarrow \mathbb{R}^3$  mit

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \begin{pmatrix} x_1 + 3x_2 \\ x_1 - x_2 \\ 2x_1 \end{pmatrix}$$

Wir erwarten eine Abbildungsmatrix  $F \in \mathbb{R}^{(3,2)}$ , also mit drei Zeilen (so viele, wie der Bildvektor hat) und zwei Spalten (so viele, wie der Urbildvektor Komponenten besitzt):

$$F = \begin{pmatrix} F_{1,1} & F_{1,2} \\ F_{2,1} & F_{2,2} \\ F_{3,1} & F_{3,2} \end{pmatrix} = \begin{pmatrix} \vec{f}(\vec{e}_1) & \vec{f}(\vec{e}_2) \end{pmatrix}$$

Wir haben also für jede der beiden Spalten von  $F$  das Bild je eines Einheitsvektors zu berechnen. Für die erste Spalte ist relevant:

$$\vec{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \in E_2$$

Um das Bild zu bestimmen, reicht es, in obiger Abbildungsvorschrift überall den Ausdruck  $x_1$  durch 1 zu ersetzen und alle Beiträge mit anderen  $x_j$  (hier nur  $x_2$ ) zu ignorieren (d.h. auf 0 zu setzen). Es ergibt sich damit:

$$\vec{f}(\vec{e}_1) = \begin{pmatrix} 1 + 3 \cdot 0 \\ 1 - 0 \\ 2 \cdot 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}$$

Hierbei spart man sich in der Praxis den Zwischenschritt und schreibt das Ergebnis direkt an. Entsprechend verfahren wir für das Bild des zweiten Basisvektors:

$$\vec{f}(\vec{e}_2) = \begin{pmatrix} 3 \\ -1 \\ 0 \end{pmatrix}$$

(In der dritten Komponente des Bildvektors kommt  $x_2$  als Variable gar nicht vor, sodass wir in diesem Fall dort auch keine Beiträge bekommen.)

Damit sind die beiden Spalten von  $F$  gefunden:

$$F = \begin{pmatrix} 1 & 3 \\ 1 & -1 \\ 2 & 0 \end{pmatrix}$$

- Wir betrachten die Abbildung  $\vec{f}: \mathbb{R}^4 \rightarrow \mathbb{R}^1$  per

$$\vec{x} \mapsto \begin{pmatrix} 2 \\ -3 \\ 1 \\ 7 \end{pmatrix} \bullet \vec{x} = 2x_1 - 3x_2 + x_3 + 7x_4$$

Wir erwarten eine Abbildungsmatrix mit einer Zeile und vier Spalten,  $F \in \mathbb{R}^{(1,4)}$ . In diesem Fall ist für jede Spalte nur eine einzelne Zahl zu berechnen.

Das Bild des Vektors

$$\vec{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \in E_4$$

erhalten wir analog wie im vorigen Beispiel dadurch, indem  $x_2, x_3, x_4$  auf 0 gesetzt (also ihre Beiträge ignoriert) werden und  $x_1$  den Wert 1 bekommt. Es ergibt sich:

$$\vec{f}(\vec{e}_1) = (2)$$

Wie oben bereits bemerkt, werden wir einkomponentige Vektoren (oder Matrizen) direkt mit ihrer Komponente identifizieren, also ist auch richtig:

$$f(\vec{e}_1) = 2$$

Analog berechnet man:  $f(\vec{e}_2) = -3$ ,  $f(\vec{e}_3) = 1$  und  $f(\vec{e}_4) = 7$ .

Also erhalten wir die folgende Abbildungsmatrix:

$$F = \begin{pmatrix} 2 & -3 & 1 & 7 \end{pmatrix}$$

Wir werden später noch sehen, dass es absolut kein Zufall ist, dass diese Matrix hier gerade dem um 90 Grad gedrehten konstanten Vektor aus der Abbildungsvorschrift (Skalarprodukt) entspricht.

Nun betrachten wir die beiden Beispiele nochmals, mit einer leicht variierten Methode. Und zwar gilt nach der weiter oben eingerahmten Faustregel für lineare Abbildungen, dass bei einer Abbildung von  $\mathbb{R}^n$  nach  $\mathbb{R}^m$  die Komponenten des Urbildvektors  $x_1, \dots, x_n$  jeweils in jeder der  $m$  Komponenten des Bildvektors als Linearkombination mit konstanten Skalierungsfaktoren auftreten (wo der Faktor 0 beträgt, wird die jeweilige Komponente nicht notiert, da sie dann nichts zur Linearkombination beiträgt).

Das heißt, dass wir den Bildvektor jederzeit als eine Addition von  $n$  einzelnen Vektoren schreiben können, von denen jeder nur jeweils die Beiträge einer der Komponenten  $x_1$  bis  $x_n$  enthält. Wir sehen das weitere Vorgehen am besten wieder am Beispiel:

- Für  $\vec{f}: \mathbb{R}^2 \rightarrow \mathbb{R}^3$  mit

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \begin{pmatrix} x_1 + 3x_2 \\ x_1 - x_2 \\ 2x_1 \end{pmatrix}$$

notieren wir den Bildvektor entsprechend als Addition von zwei Vektoren:

$$\vec{f}(\vec{x}) = \begin{pmatrix} x_1 + 3x_2 \\ x_1 - x_2 \\ 2x_1 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_1 \\ 2x_1 \end{pmatrix} + \begin{pmatrix} 3x_2 \\ -x_2 \\ 0 \end{pmatrix}$$

Da nun der  $k$ -te solche Vektor jeweils genau den Skalierungsfaktor  $x_k$  enthält, können wir diesen heraus ziehen:

$$\dots = x_1 \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} + x_2 \begin{pmatrix} 3 \\ -1 \\ 0 \end{pmatrix}$$

Hier erkennen wir mit dem Vorwissen von oben direkt die beiden Spalten der Abbildungsmatrix  $F$  wieder – denn wenn wir nun wieder die Bilder der kartesischen Einheitsvektoren bestimmen, läuft das jeweils darauf hinaus, genau einen der Skalierungsfaktoren  $x_k$  auf 1 zu setzen und alle anderen auf 0: das selektiert uns genau einen der Spaltenvektoren! Es ergibt sich also genau die gleiche Abbildungsmatrix  $F$  wie oben.



- Für  $\vec{f}: \mathbb{R}^4 \rightarrow \mathbb{R}^1$  per

$$\vec{x} \mapsto \begin{pmatrix} 2 \\ -3 \\ 1 \\ 7 \end{pmatrix} \bullet \vec{x} = 2x_1 - 3x_2 + x_3 + 7x_4$$

ist mit der rechten Seite der Abbildungsvorschrift schon fast alles vorbereitet. Wir schreiben (nur zum Trainieren) diesen Ausdruck nochmal so um, wie wir ihn dann mit dem Verfahren verarbeiten würden:

$$f(\vec{x}) = x_1 (2) + x_2 (-3) + x_3 (1) + x_4 (7)$$

In Klammern sind die (einkomponentigen Vektoren der) Spalten der Abbildungsmatrix  $F$  wie oben.

Beide Verfahren zur Gewinnung der Abbildungsmatrix sind gleichwertig. Beim ersten können wir direkt mit dem Bildvektor aus der Abbildungsvorschrift arbeiten, müssen dort jedoch sorgfältig darauf achten, welche Variablen wir auf 0 setzen und welche jeweils 1 beträgt. Bei der zweiten Methode bekommen wir allein durch die Notation die Spalten der Matrix geliefert, haben aber die zusätzliche Arbeit, den Bildvektor in die  $n$  Spalten zu zerlegen.

## 7.2.4 Produkt aus Matrix und Vektor

Nun zurück zur mathematischen Betrachtung der linearen Abbildung – was wir in Definition 7.3 bereits fest gehalten hatten, ist tatsächlich noch mehr als dort angedeutet war – hier hilft uns, dass wir uns gestattet hatten, in Notation und Konzept keinen Unterschied zwischen einem Vektor und einer 1-spaltigen Matrix zu machen.

Werten wir nämlich Urbild- und Bildvektoren kurzfristig als Matrizen, dann liegt mit der linearen Abbildung eine Verknüpfung vor, die zwei Matrizen (die Abbildungsmatrix und den Urbildvektor) zum Bildvektor (hier: eine Matrix) verknüpft. Die Formel in obiger Definition macht dabei deutlich, dass es sich dabei um ein *Produkt* handelt.

Wir werden im nächsten Abschnitt in der Tat das Produkt von Matrizen allgemeiner ausdrücken; es wird sich aber lediglich um eine Erweiterung des bis hierhin bekannten Konzepts handeln, das wir daher hier schon separat vorstellen:

**Definition 7.4** (Produkt aus Matrix und Vektor). *Für eine Matrix  $F \in \mathbb{R}^{(m,n)}$  und einen Vektor  $\vec{x} \in \mathbb{R}^n$  ist das Produkt aus  $F$  und  $\vec{x}$  gegeben als Vektor  $\vec{y} \in \mathbb{R}^m$ , mit*

$$\vec{y} = F \cdot \vec{x} \quad \text{mit} \quad \forall j \in \{1, \dots, m\} : y_j = (F \cdot \vec{x})_j := \sum_{k=1}^n F_{j,k} x_k$$

### Bemerkungen:

- Der Multiplikationspunkt zwischen Matrix und Vektor wird oft nicht mit angeschrieben.
- Die Matrix muss stets genau so viele Spalten besitzen, wie der Vektor (an den sie von links multipliziert wird) Komponenten (Zeilen) besitzt.

Die Zeilenzahl der Matrix bestimmt dabei die Anzahl der Komponenten (Zeilen) des Ergebnisvektors.

**Beispiel:** Wir behandeln hier ein allgemeines Beispiel, um die Rechenmethode für das obige Produkt zu erfassen; die konkreten Zahlenbeispiele folgen dann nach dem nächsten Satz, in Verbindung mit linearen Abbildungen.

Exemplarisch betrachten wir eine Matrix  $F \in \mathbb{R}^{(3,4)}$  und einen Vektor  $\vec{x} \in \mathbb{R}^4$ . Das Ergebnis des Produkts aus  $F$  und  $\vec{x}$  wird ein Vektor mit drei Komponenten  $\vec{y} \in \mathbb{R}^3$  sein. Insgesamt:

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} F_{1,1} & F_{1,2} & F_{1,3} & F_{1,4} \\ F_{2,1} & F_{2,2} & F_{2,3} & F_{2,4} \\ F_{3,1} & F_{3,2} & F_{3,3} & F_{3,4} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}$$

Nun betrachten wir die einzelnen Komponenten des Ergebnisvektors nacheinander. Für die erste Komponente liegt folgende Situation vor (nur die relevanten Komponenten sind notiert):

$$\begin{pmatrix} y_1 \\ \dots \\ \dots \end{pmatrix} = \begin{pmatrix} F_{1,1} & F_{1,2} & F_{1,3} & F_{1,4} \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}$$

Die Summenformel aus der obigen Definition hält bei den Matrixelementen den linken Index fest auf 1 (die erste Zeile, passend zur ersten Komponente von  $\vec{y}$ ); die rechten Indices laufen von 1 bis 4 durch. Dem linken Element  $F_{1,1}$  wird dabei  $x_1$  anmultipliziert, dem nächsten ( $F_{1,2}$ ) dann  $x_2$  usw. Falls wir die einzelnen Komponenten in Gedanken mit den Fingern antippen, um zu bestimmen, welche mit welcher zu multiplizieren ist, starten wir in der Matrixzeile links und im Urbildvektor oben. Dann rückt der Finger der linken Hand eine Position nach rechts, und der der rechten Hand eine nach unten. Jeweils wird das Produkt der beiden selektierten Komponenten berechnet und auf das Zwischenergebnis addiert. Am Schluss sind alle Spalten der selektierten Matrixzeile abgetastet; ebenso der Urbildvektor einmal komplett von oben nach unten. Hier die einzelnen Schritte (die zu multiplizierenden Komponenten sind eingerahmt):

$$\begin{pmatrix} \boxed{F_{1,1}} & F_{1,2} & F_{1,3} & F_{1,4} \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \end{pmatrix} \begin{pmatrix} \boxed{x_1} \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}$$

$$\begin{pmatrix} F_{1,1} & \boxed{F_{1,2}} & F_{1,3} & F_{1,4} \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \end{pmatrix} \begin{pmatrix} x_1 \\ \boxed{x_2} \\ x_3 \\ x_4 \end{pmatrix}$$

$$\begin{pmatrix} F_{1,1} & F_{1,2} & \boxed{F_{1,3}} & F_{1,4} \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \boxed{x_3} \\ x_4 \end{pmatrix}$$

$$\begin{pmatrix} F_{1,1} & F_{1,2} & F_{1,3} & \boxed{F_{1,4}} \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \boxed{x_4} \end{pmatrix}$$

Die Summe aus diesen vier Produkten ergibt dann den Wert  $y_1 = (F \cdot \vec{x})_1$ .

Analog für die anderen beiden Komponenten von  $\vec{y}$ , mit den Situationen

$$\begin{pmatrix} \dots \\ y_2 \\ \dots \end{pmatrix} = \begin{pmatrix} \dots & \dots & \dots & \dots \\ F_{2,1} & F_{2,2} & F_{2,3} & F_{2,4} \\ \dots & \dots & \dots & \dots \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}$$

und

$$\begin{pmatrix} \dots \\ \dots \\ y_3 \end{pmatrix} = \begin{pmatrix} \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ F_{3,1} & F_{3,2} & F_{3,3} & F_{3,4} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}$$

Jeweils sind wieder vier einzelne Komponentenprodukte zu berechnen und zu summieren.

### 7.2.5 Zusammenfassung zu linearen Abbildungen

Mit dem eben eingeführten Produkt aus Matrix und Vektor gehen wir zurück zur Definition 7.3 und halten folgendes fest:

**Satz 7.5** (Lineare Abbildung als Produkt). *Jede lineare Abbildung  $\vec{f} : \mathbb{R}^n \rightarrow \mathbb{R}^m$  lässt sich als Produkt der zugehörigen Abbildungsmatrix  $F \in \mathbb{R}^{(m,n)}$  mit dem Urbildvektor  $\vec{x} \in \mathbb{R}^n$  ausdrücken:*

$$\vec{f}(\vec{x}) = F \cdot \vec{x}$$

Weiterhin sind alle Produkte von Matrizen aus  $\mathbb{R}^{(m,n)}$  mit Vektoren  $\vec{x} \in \mathbb{R}^n$  lineare Abbildungen.

(Beweis: S. 327.)

**Bemerkung:** Mit den beiden Teilen des Satzes gilt also zusammen:

Lineare Abbildungen von  $\mathbb{R}^n$  nach  $\mathbb{R}^m$  sind genau die Funktionen, die sich durch Produkte von Matrizen aus  $\mathbb{R}^{(m,n)}$  mit Vektoren aus  $\mathbb{R}^n$  beschreiben lassen.

**Beispiele:**

- Für  $\vec{f}: \mathbb{R}^2 \rightarrow \mathbb{R}^3$  mit

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \begin{pmatrix} x_1 + 3x_2 \\ x_1 - x_2 \\ 2x_1 \end{pmatrix}$$

hatten wir oben bereits die Abbildungsmatrix

$$F = \begin{pmatrix} 1 & 3 \\ 1 & -1 \\ 2 & 0 \end{pmatrix}$$

ermittelt. Wir berechnen für die Vektoren

$$\vec{x}_1 := \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad \vec{x}_2 := \begin{pmatrix} 2 \\ 3 \end{pmatrix} \quad \text{und} \quad \vec{x}_3 := \begin{pmatrix} -1 \\ -1 \end{pmatrix}$$

die Bilder jeweils zunächst durch Einsetzen in die Abbildungsvorschrift, dann durch Produktbildung mit  $F$ .

$$\vec{f}(\vec{x}_1) = \vec{f}\left(\begin{pmatrix} 1 \\ 2 \end{pmatrix}\right) = \begin{pmatrix} 1 + 3 \cdot 2 \\ 1 - 2 \\ 2 \cdot 1 \end{pmatrix} = \begin{pmatrix} 7 \\ -1 \\ 2 \end{pmatrix}$$

$$\vec{f}(\vec{x}_2) = \vec{f}\left(\begin{pmatrix} 2 \\ 3 \end{pmatrix}\right) = \begin{pmatrix} 2 + 3 \cdot 3 \\ 2 - 3 \\ 2 \cdot 2 \end{pmatrix} = \begin{pmatrix} 11 \\ -1 \\ 4 \end{pmatrix}$$

$$\vec{f}(\vec{x}_3) = \vec{f}\left(\begin{pmatrix} -1 \\ -1 \end{pmatrix}\right) = \begin{pmatrix} -1 + 3 \cdot (-1) \\ -1 - (-1) \\ 2 \cdot (-1) \end{pmatrix} = \begin{pmatrix} -4 \\ 0 \\ -2 \end{pmatrix}$$

Nun zu den Produkten (Zwischenschritte mit zeilenweiser Berechnung eingefügt):

$$F \cdot \vec{x}_1 = \begin{pmatrix} 1 & 3 \\ 1 & -1 \\ 2 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 \cdot 1 + 3 \cdot 2 \\ 1 \cdot 1 + (-1) \cdot 2 \\ 2 \cdot 1 + 0 \cdot 2 \end{pmatrix} = \begin{pmatrix} 7 \\ -1 \\ 2 \end{pmatrix}$$

$$F \cdot \vec{x}_2 = \begin{pmatrix} 1 & 3 \\ 1 & -1 \\ 2 & 0 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 1 \cdot 2 + 3 \cdot 3 \\ 1 \cdot 2 + (-1) \cdot 3 \\ 2 \cdot 2 + 0 \cdot 3 \end{pmatrix} = \begin{pmatrix} 11 \\ -1 \\ 4 \end{pmatrix}$$

$$F \cdot \vec{x}_3 = \begin{pmatrix} 1 & 3 \\ 1 & -1 \\ 2 & 0 \end{pmatrix} \cdot \begin{pmatrix} -1 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \cdot (-1) + 3 \cdot (-1) \\ 1 \cdot (-1) + (-1) \cdot (-1) \\ 2 \cdot (-1) + 0 \cdot (-1) \end{pmatrix} = \begin{pmatrix} -4 \\ 0 \\ -2 \end{pmatrix}$$

- Für  $\vec{f}: \mathbb{R}^4 \rightarrow \mathbb{R}^1$  per

$$\vec{x} \mapsto \begin{pmatrix} 2 \\ -3 \\ 1 \\ 7 \end{pmatrix} \bullet \vec{x} = 2x_1 - 3x_2 + x_3 + 7x_4$$

mit der Abbildungsmatrix

$$F = (2 \quad -3 \quad 1 \quad 7)$$

und dem Urbildvektor

$$\vec{x} := \begin{pmatrix} 2 \\ 1 \\ 3 \\ 5 \end{pmatrix} :$$

erhalten wir:

$$f(\vec{x}) = \begin{pmatrix} 2 & -3 & 1 & 7 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ 1 \\ 3 \\ 5 \end{pmatrix} = (2 \cdot 2 + (-3) \cdot 1 + 1 \cdot 3 + 7 \cdot 5) = (39)$$

Wenn wir diese einkomponentige Matrix mit ihrem Element identifizieren, haben wir genau das (kanonische) Skalarprodukt erhalten, das in der Abbildungsvorschrift bereits ausgedrückt war.

- Im zweiten Teil des Satzes wurde behauptet, dass jedes Produkt aus Matrix und Vektor (so lange die Dimensionierung verträglich ist, also die Spaltenzahl der Matrix mit der Dimension des Vektors überein stimmt) eine lineare Abbildung ist. Wir wollen daher für die Matrix

$$F := \begin{pmatrix} 2 & 3 & -1 & 12 \\ -5 & 1 & 0 & 0 \\ 42 & 0 & -2 & -6 \end{pmatrix} \in \mathbb{R}^{(3,4)}$$

die Abbildungsvorschrift aufstellen. Dazu bestimmen wir das Produkt mit

$$\vec{x} := \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}$$

und ignorieren Beiträge mit Wert 0:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \mapsto F \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 2x_1 + 3x_2 - x_3 + 12x_4 \\ -5x_1 + x_2 \\ 42x_1 - 2x_3 - 6x_4 \end{pmatrix}$$

Man überprüfe zur Übung, dass es sich nach Definition 7.1 um eine lineare Abbildung handelt, und dass sich (wahlweise mit einer der beiden oben besprochenen Methoden) aus der Abbildungsvorschrift wieder die Abbildungsmatrix  $F$  rekonstruieren lässt.

Wir halten für spätere Überlegungen noch fest, dass *quadratische* Matrizen, also solche aus  $\mathbb{R}^{(n,n)}$  – bzw. die mit ihnen assoziierten linearen Abbildungen –, Vektorräume (hier  $\mathbb{R}^n$ ) auf sich selbst abbilden. Das folgt direkt aus obigen Überlegungen, falls  $m := n$  gesetzt wird.

## 7.2.6 Matrizenräume

Bisher haben wir uns noch nicht weiter mit den algebraischen Eigenschaften der Mengen  $\mathbb{R}^{(m,n)}$  befasst. Wir hatten allerdings bereits ein Produkt zwischen Matrizen und Vektoren kennen gelernt (das sich, wie angedeutet, als ein Spezialfall eines allgemeineren Produkts zwischen Matrizen herausstellen wird).

Wir hatten aber schon bemerkt, dass die Matrizen aus solchen Mengen genau mit den linearen Abbildungen von  $\mathbb{R}^n$  nach  $\mathbb{R}^m$  identifizierbar sind. Nun haben lineare Abbildungen gerade die Eigenschaft, dass sie mit den Vektorraum-Operationen (Skalierung und Addition) vertauschen.

Dies, und die Tatsache, dass wir oben schon von den “Spaltenvektoren” von Matrizen gesprochen hatten, lässt vermuten, dass Matrizen selbst mit Vektorraumoperationen verträglich sind. Tatsächlich gilt folgender

**Satz 7.6** (Matrizenräume). *Die Matrizen aus  $\mathbb{R}^{(m,n)}$  ( $m, n \in \mathbb{N}$ ) bilden mit der komponentenweisen Addition und Skalierung einen  $\mathbb{R}$ -Vektorraum. Speziell ist für  $F, G \in \mathbb{R}^{(m,n)}$  und  $a \in \mathbb{R}$ :*

$$(F + G)_{j,k} := F_{j,k} + G_{j,k} \quad \text{und} \quad (aF)_{j,k} = aF_{j,k}$$

(Beweis: S. 328.)

### Bemerkungen:

- Wir verzichten hier auf die Einführung gesonderter Operatorsymbole für die Addition und Skalierung von Matrizen.
- Das neutrale Element der Addition ist die *Nullmatrix*, die  $m$  Zeilen und  $n$  Spalten besitzt und in jeder Komponente 0 enthält.
- Wenn  $\vec{f}, \vec{g}$  lineare Abbildungen von  $\mathbb{R}^n$  nach  $\mathbb{R}^m$  sind, und  $F, G$  ihre jeweiligen Abbildungsmatrizen, dann gilt für  $a, b \in \mathbb{R}$ :

$$(a\vec{f} + b\vec{g})(\vec{x}) = a\vec{f}(\vec{x}) + b\vec{g}(\vec{x}) = a(F \cdot \vec{x}) + b(G \cdot \vec{x}) = (aF + bG) \cdot \vec{x}$$

Links steht hierbei eine zusammen gesetzte Abbildung. Diese ist wohldefiniert, da sowohl  $\vec{f}$  als auch  $\vec{g}$  nach  $\mathbb{R}^m$  abbilden – aber dies ist ein reeller Vektorraum, sodass die Addition und Skalierung der Bildvektoren von  $\vec{f}$  und  $\vec{g}$  möglich ist.

Die erste Gleichheit ist richtig, da lineare Abbildungen Funktionen sind; siehe dazu Definition 4.6 (Rechenoperationen mit Funktionen).

Für die zweite Gleichheit wurde Satz 7.5 verwendet, der die linearen Abbildungen mit den Abbildungsmatrizen verbindet.

Die dritte (letzte) Gleichheit nutzt nun die Vektorraum-Eigenschaft von  $\mathbb{R}^{(m,n)}$  in Verbindung mit der Darstellung der zusammen gesetzten Abbildung von ganz links. Damit wird auch die linksseitige Linearität des Matrixprodukts bestätigt.<sup>1</sup>

Linearkombinationen von linearen Abbildungen sind also ebenfalls wieder lineare Abbildungen, und ihre Abbildungsmatrizen entsprechen den Linearkombinationen der einzelnen Teilabbildungs-Matrizen.

**Beispiel:** In  $\mathbb{R}^{(2,3)}$  ist folgende Rechnung korrekt:

$$4 \cdot \begin{pmatrix} 1 & 3 & -2 \\ 2 & 0 & 3 \end{pmatrix} + \begin{pmatrix} 5 & 1 & -12 \\ 3 & 2 & 2 \end{pmatrix} = \begin{pmatrix} 4 & 12 & -8 \\ 8 & 0 & 12 \end{pmatrix} + \begin{pmatrix} 5 & 1 & -12 \\ 3 & 2 & 2 \end{pmatrix} = \begin{pmatrix} 9 & 13 & -20 \\ 11 & 2 & 14 \end{pmatrix}$$

## 7.3 Matrixprodukt

Wir hatten oben bereits die Produkte von Matrizen mit Vektoren kennen gelernt, die gerade der Ausführung von linearen Abbildungen entsprachen. Tatsächlich werden wir aber fest stellen, dass ein allgemeineres Produkt von Matrizen existieren muss, denn die Komposition (Hintereinanderausführung) von linearen Abbildungen ist wiederum eine lineare Abbildung und muss daher eine korrespondierende Abbildungsmatrix besitzen. Das Matrixprodukt wird gerade die Methode sein, um aus den zwei Matrizen der komponierten Abbildungen die Matrix der Komposition zu bestimmen (auch letztere ist ja nach Satz 7.5 eindeutig definiert).

### 7.3.1 Komposition linearer Abbildungen

Es gilt folgender

**Satz 7.7** (Komposition linearer Abbildungen). *Seien  $\vec{f} : \mathbb{R}^n \rightarrow \mathbb{R}^m$  und  $\vec{g} : \mathbb{R}^m \rightarrow \mathbb{R}^p$  lineare Abbildungen. Dann ist auch die Komposition  $\vec{h} : \mathbb{R}^n \rightarrow \mathbb{R}^p$  mit  $\vec{h} := \vec{g} \circ \vec{f}$  eine lineare Abbildung.*

(Beweis: S. 328.)

**Beispiel:** Wir betrachten die beiden linearen Abbildungen  $\vec{f} : \mathbb{R}^2 \rightarrow \mathbb{R}^3$  sowie  $\vec{g} : \mathbb{R}^3 \rightarrow \mathbb{R}^4$  mit

$$\vec{y} := \vec{f}(\vec{x}) := \begin{pmatrix} x_1 + 3x_2 \\ x_1 - x_2 \\ 2x_1 \end{pmatrix} \quad \text{und} \quad \vec{z} := \vec{g}(\vec{y}) := \begin{pmatrix} 2y_1 + y_2 + y_3 \\ 3y_2 \\ 4y_1 - 2y_2 + y_3 \\ y_1 + 2y_3 \end{pmatrix}$$

<sup>1</sup>für die rechtsseitige Linearität, also das lineare Verhalten bei  $F \cdot (a\vec{x} + b\vec{y})$  reicht die Tatsache, dass  $\vec{f}$  eine lineare Abbildung ist – s.o.

Die zugehörigen Abbildungsmatrizen sind:

$$F = \begin{pmatrix} 1 & 3 \\ 1 & -1 \\ 2 & 0 \end{pmatrix} \quad \text{und} \quad G = \begin{pmatrix} 2 & 1 & 1 \\ 0 & 3 & 0 \\ 4 & -2 & 1 \\ 1 & 0 & 2 \end{pmatrix}$$

Nun drücken wir  $\vec{z}$  als Funktion von  $\vec{x}$  aus, indem wir die Komponenten von  $\vec{y}$  einsetzen und dann zusammen fassen:

$$\begin{aligned} z_1 &= 2y_1 + y_2 + y_3 = 2(x_1 + 3x_2) + (x_1 - x_2) + (2x_1) = 2x_1 + 6x_2 + x_1 - x_2 + 2x_1 \\ &= 5x_1 + 5x_2 \\ z_2 &= 3y_2 = 3(x_1 - x_2) \\ &= 3x_1 - 3x_2 \\ z_3 &= 4y_1 - 2y_2 + y_3 = 4(x_1 + 3x_2) - 2(x_1 - x_2) + (2x_1) = 4x_1 + 12x_2 - 2x_1 + 2x_2 + 2x_1 \\ &= 4x_1 + 14x_2 \\ z_4 &= y_1 + 2y_3 = (x_1 + 3x_2) + 2(2x_1) = x_1 + 3x_2 + 4x_1 \\ &= 5x_1 + 3x_2 \end{aligned}$$

Somit ist

$$\vec{z} = \vec{h}(\vec{x}) = \begin{pmatrix} 5x_1 + 5x_2 \\ 3x_1 - 3x_2 \\ 4x_1 + 14x_2 \\ 5x_1 + 3x_2 \end{pmatrix} \quad \text{mit} \quad H = \begin{pmatrix} 5 & 5 \\ 3 & -3 \\ 4 & 14 \\ 5 & 3 \end{pmatrix}$$

Da wir für  $\vec{h}$  eine Abbildungsmatrix  $H$  angeben können, ist nach Satz 7.5 klar, dass es sich um eine lineare Abbildung handeln muss.

### 7.3.2 Herleitung des Matrixprodukts

Nun wollen wir systematisch versuchen, die Abbildungsmatrix der Komposition zu bestimmen (von der wir nach Satz 7.7 wissen, dass sie existieren muss, und welche Dimensionierung sie besitzt).

Mit den selben Dimensionsangaben wie im eben zitierten Satz ist also:

$$\vec{y} := \vec{f}(\vec{x}) = F\vec{x} \in \mathbb{R}^m \quad \text{und} \quad \vec{z} := \vec{g}(\vec{y}) = G\vec{y} \in \mathbb{R}^p$$

Dann erhalten wir für die  $j$ -te Komponente von  $\vec{z}$ :

$$z_j = (G\vec{y})_j = \sum_{k=1}^m G_{j,k} \cdot y_k$$

mit

$$y_k = \sum_{r=1}^n F_{k,r} \cdot x_r$$

Also ist insgesamt:

$$z_j = \sum_{k=1}^m G_{j,k} \cdot \left( \sum_{r=1}^n F_{k,r} \cdot x_r \right)$$

Wegen des Distributivgesetzes und der Kommutativität bei Addition und Multiplikation für reelle Zahlen können wir die Summe über  $r$  nach außen ziehen und erhalten (bei weiterhin insgesamt  $m \cdot n$  Produkten, die am Ende zu  $z_j$  summiert werden):

$$z_j = \sum_{r=1}^n \sum_{k=1}^m G_{j,k} \cdot F_{k,r} \cdot x_r$$

Nun wissen wir bereits, dass es eine Abbildungsmatrix  $H \in \mathbb{R}^{(p,n)}$  geben muss, sodass  $\vec{z} = H\vec{x}$  gilt, also

$$z_j = \sum_{r=1}^n H_{j,r} x_r$$

Aber dann können wir die beiden Formeln für  $z_j$  vergleichen und erhalten:

$$H_{j,r} = \sum_{k=1}^m G_{j,k} \cdot F_{k,r}$$

Wie beim Produkt aus Matrix und Vektor haben wir hier also eine Summe auszuführen, zu der Produkte von Komponenten beitragen. Der linke Index  $j$  der Matrix  $H$  fixiert die Zeile von  $G$ ; der rechte Index  $r$  von  $H$  fixiert die Spalte von  $F$ .

Wir fassen zusammen:

**Definition 7.8** (Matrixprodukt). Für  $G \in \mathbb{R}^{(p,m)}$  und  $F \in \mathbb{R}^{(m,n)}$  ist das Matrixprodukt aus  $G$  und  $F$ , die Matrix  $G \cdot F \in \mathbb{R}^{(p,n)}$ , definiert per

$$(G \cdot F)_{j,r} := \sum_{k=1}^m G_{j,k} \cdot F_{k,r}$$

#### Bemerkungen:

- Man beachte, wie die Dimensionierungen zusammen passen müssen: Die Zahl der Spalten des linken Faktors  $G$  muss der Zahl der Zeilen des rechten Faktors  $F$  entsprechen. Der Summationsindex bei der Berechnung der Komponenten von  $G \cdot F$  läuft genau von 1 bis zu dieser Zahl.

Dadurch wird (s.u.) ein ganz analoger Prozess nötig, wie wir ihn schon beim Produkt aus Matrix und Vektor gesehen hatten.

- In der Summe von obiger Vorschrift ist der linke (Zeilen-)Index von  $G$  (dem linken Faktor) fixiert; ebenso der rechte (Spalten-)Index von  $F$  (dem rechten Faktor). Der Summationsindex ist, wenn man die Faktoren in der richtigen Anordnung anschreibt, genau der *innen liegende*.
- Auch hier verzichtet man oft auf den Multiplikationspunkt (solange aus dem Kontext klar ist, dass es sich um ein Produkt von Matrizen handelt).

Speziell für die Komposition von linearen Abbildungen gilt dann folgender

**Satz 7.9** (Abbildungsmatrizen bei Komposition). Seien  $\vec{f} : \mathbb{R}^n \rightarrow \mathbb{R}^m$  und  $\vec{g} : \mathbb{R}^m \rightarrow \mathbb{R}^p$  lineare Abbildungen mit den zugehörigen Matrizen  $F \in \mathbb{R}^{(m,n)}$  und  $G \in \mathbb{R}^{(p,m)}$ . Dann gilt für die Abbildungsmatrix  $H \in \mathbb{R}^{(p,n)}$  der Komposition  $\vec{h} : \mathbb{R}^n \rightarrow \mathbb{R}^p$  mit  $\vec{h} := \vec{g} \circ \vec{f}$ :

$$H = G \cdot F$$

**Bemerkung:** Dann ist auch:

$$\vec{h}(\vec{x}) = H\vec{x} = (GF)\vec{x} = \vec{g}(\vec{f}(\vec{x})) = \vec{g}(F\vec{x}) = G(F\vec{x})$$

**Beispiel:** Wir berechnen nun nach der Formel die komponierte Abbildungsmatrix aus dem Beispiel zu Satz 7.7 nochmal nach. Es war

$$F = \begin{pmatrix} 1 & 3 \\ 1 & -1 \\ 2 & 0 \end{pmatrix} \quad \text{und} \quad G = \begin{pmatrix} 2 & 1 & 1 \\ 0 & 3 & 0 \\ 4 & -2 & 1 \\ 1 & 0 & 2 \end{pmatrix}$$

Wir wissen bereits, dass  $H$  eine Matrix mit vier Zeilen und zwei Spalten sein wird. Nach der Formel gilt (hier ist  $m = 3$ ) z.B.:

$$H_{1,1} = \sum_{k=1}^3 G_{1,k} F_{k,1} = G_{1,1} F_{1,1} + G_{1,2} F_{2,1} + G_{1,3} F_{3,1} = 2 \cdot 1 + 1 \cdot 1 + 1 \cdot 2 = 5$$

$$H_{1,2} = \sum_{k=1}^3 G_{1,k} F_{k,2} = G_{1,1} F_{1,2} + G_{1,2} F_{2,2} + G_{1,3} F_{3,2} = 2 \cdot 3 + 1 \cdot (-1) + 1 \cdot 0 = 5$$

Für  $H_{3,1}$  stellen wir die Situation wie oben beim Produkt aus Matrix und Vektor dar:

$$\begin{pmatrix} \dots & \dots \\ \dots & \dots \\ H_{3,1} & \dots \\ \dots & \dots \end{pmatrix} = \begin{pmatrix} \dots & \dots & \dots \\ \dots & \dots & \dots \\ 4 & -2 & 1 \\ \dots & \dots & \dots \end{pmatrix} \cdot \begin{pmatrix} 1 & \dots \\ 1 & \dots \\ 2 & \dots \end{pmatrix}$$

Im Prinzip kann man zur Berechnung von  $H_{3,1}$  die zweite Spalte von  $H$  und die zweite Spalte von  $F$  ignorieren und hat dann wieder die selben Rechnungen auszuführen wie oben schon einmal gezeigt. Hier die drei Produkte, die zu summieren sind (die zusammen gehörigen Faktoren jeweils umrahmt; auch hier kann es hilfreich sein, die jeweiligen Komponenten mit den Fingern abzutasten, um die Bewegungsrichtungen zu verinnerlichen):

$$\begin{pmatrix} \dots & \dots & \dots \\ \dots & \dots & \dots \\ \boxed{4} & -2 & 1 \\ \dots & \dots & \dots \end{pmatrix} \begin{pmatrix} \boxed{1} & \dots \\ 1 & \dots \\ 2 & \dots \end{pmatrix}$$

$$\begin{pmatrix} \dots & \dots & \dots \\ \dots & \dots & \dots \\ 4 & \boxed{-2} & 1 \\ \dots & \dots & \dots \end{pmatrix} \begin{pmatrix} 1 & \dots \\ \boxed{1} & \dots \\ 2 & \dots \end{pmatrix}$$

$$\begin{pmatrix} \dots & \dots & \dots \\ \dots & \dots & \dots \\ 4 & -2 & \boxed{1} \\ \dots & \dots & \dots \end{pmatrix} \begin{pmatrix} 1 & \dots \\ 1 & \dots \\ \boxed{2} & \dots \end{pmatrix}$$

Das Ergebnis ist also  $H_{3,1} = 4 \cdot 1 + (-2) \cdot 1 + 1 \cdot 2 = 4$ .

Man rechne die anderen Komponenten zur Übung nach; es ergibt sich wie erwartet:

$$H = \begin{pmatrix} 5 & 5 \\ 3 & -3 \\ 4 & 14 \\ 5 & 3 \end{pmatrix}$$

### 7.3.3 Eigenschaften des Matrixprodukts

**Bemerkungen:** Wir sammeln hier einige algebraische Eigenschaften des Matrizenprodukts auf:

- Zunächst halten wir also fest, dass das Produkt aus Matrix und Vektor ein Spezialfall des Matrixprodukts ist, falls wir den Vektor mit einer einspaltigen Matrix identifizieren. Es ergibt sich dann nach Definition 7.8 auch (wie erwartet) eine einspaltige Matrix als Resultat, die wir dann wieder als Vektor begreifen können.

- Das Produkt ist *nicht kommutativ*. Zum Teil ergibt sich das schon aus der Dimensionalität. Für  $A \in \mathbb{R}^{(m,n)}$  und  $B \in \mathbb{R}^{(p,m)}$  ist zwar  $BA$  definiert, aber  $AB$  kann nur existieren, wenn  $n = p$ . Falls das nicht der Fall ist, wäre  $AB$  gar nicht definiert.

Falls aber mit  $BA$  auch  $AB$  definiert sein sollte, müsste somit  $A \in \mathbb{R}^{(m,n)}$  und  $B \in \mathbb{R}^{(n,m)}$  gelten. Dann wäre allerdings  $BA$  aus  $\mathbb{R}^{(n,n)}$  und  $AB$  aus  $\mathbb{R}^{(m,m)}$ . Für  $m \neq n$  ist auch das noch strukturell verschieden.

Für  $m = p = n$  wären tatsächlich beide Produkte definiert:  $A, B, AB, BA$  stammen dann alle aus  $\mathbb{R}^{(n,n)}$ . Die Formel in Definition 7.8 zeigt aber, dass auch dann noch keine Kommutativität gegeben ist, da es eindeutig auf die Reihenfolge der Faktoren ankommt (beim linken Faktor läuft die Summation über den rechten Index, beim rechten Faktor über den linken).

- Immerhin sehen wir aber, dass das Matrixprodukt auf den Räumen der *quadratischen Matrizen*  $\mathbb{R}^{(n,n)}$  *abgeschlossen* ist.
- Das Matrixprodukt ist auch *assoziativ*. Das rechnet man entweder nach durch Einsetzen der Formel und Anwendung von Distributiv- und Assoziativgesetzen für reelle Zahlen.

Oder man erinnert sich, dass das Matrixprodukt direkt mit der Komposition linearer Abbildungen verbunden ist. Und nach Satz 4.5 ist die Komposition von Funktionen assoziativ. Dies muss speziell auch für die Komposition linearer Abbildungen gelten, und damit auch für das Matrixprodukt.



Wir stellen noch eine weitere Formulierung für die Spalten eines Matrizenprodukts auf. Für  $A \in \mathbb{R}^{(m,n)}$  und  $B \in \mathbb{R}^{(p,m)}$  sei  $C := BA \in \mathbb{R}^{(p,n)}$ . Die Spaltendarstellungen von  $A$  und  $C$  besitzen jeweils  $n$  Vektoren:

$$A = (\vec{a}_1 \quad \cdots \quad \vec{a}_n) \quad \text{und} \quad C = (\vec{c}_1 \quad \cdots \quad \vec{c}_n)$$

Nun gilt für den  $r$ -ten Spaltenvektor von  $C$  (komponentenweise mit  $j \in \{1, \dots, p\}$ ):

$$(\vec{c}_r)_j = C_{j,r} = (BA)_{j,r} = \sum_{k=1}^m B_{j,k} A_{k,r} = \sum_{k=1}^m B_{j,k} (\vec{a}_r)_k = (B \cdot \vec{a}_r)_j$$

Für den letzten Schritt haben wir die Formel aus Definition 7.4 verwendet.

Da  $j$  beliebig war, gilt also folgender

**Satz 7.10** (Spalten des Matrixprodukts). Für  $A \in \mathbb{R}^{(m,n)}$  und  $B \in \mathbb{R}^{(p,m)}$  sei  $C := BA \in \mathbb{R}^{(p,n)}$ . Falls  $A$  die Spaltendarstellung

$$A = (\vec{a}_1 \quad \cdots \quad \vec{a}_n)$$

besitzt, hat  $C$  die Spaltendarstellung

$$C = (B\vec{a}_1 \quad \cdots \quad B\vec{a}_n)$$

### 7.3.4 Quadratische Matrizen

Eben haben wir gesehen, dass das Matrixprodukt für quadratische Matrizen eine abgeschlossene Operation ist – außerdem ist es assoziativ. Wir zeigen nun, dass es sogar ein neutrales Element der Multiplikation gibt:

**Definition 7.11** (Einheitsmatrix). Die quadratische Matrix  $\mathbb{1}_n \in \mathbb{R}^{(n,n)}$  mit

$$(\mathbb{1}_n)_{j,k} := \delta_{jk}$$

heißt ( $n$ -reihige) Einheitsmatrix. Ihre Spaltendarstellung lautet mit den Basisvektoren der Standardbasis  $E_n$ :

$$\mathbb{1}_n = (\vec{e}_1 \quad \vec{e}_2 \quad \cdots \quad \vec{e}_n)$$

**Beispiel:** Für  $n = 4$  lautet die Einheitsmatrix:

$$\mathbb{1}_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Wir bilden nun die rechts- und linksseitigen (dimensionsmäßig korrekten) Produkte einer beliebigen quadratischen Matrix  $A \in \mathbb{R}^{(m,n)}$ :

$$(A \cdot \mathbb{1}_n)_{j,r} = \sum_{k=1}^n A_{j,k} \cdot (\mathbb{1}_n)_{k,r} = \sum_{k=1}^n A_{j,k} \cdot \delta_{kr} = A_{j,r}$$

(Hierbei eliminiert das Kroneckerdelta wieder die Summe, indem es den Summationsindex  $k$  auf den Wert  $r$  fixiert – alle anderen Beiträge der Summe verschwinden). Analog:

$$(\mathbb{1}_m \cdot A)_{j,r} = \sum_{k=1}^m (\mathbb{1}_m)_{j,k} \cdot A_{k,r} = \sum_{k=1}^m \delta_{jk} \cdot A_{k,r} = A_{j,r}$$

Also ist  $\mathbb{1}_n$  rechtsneutral zu  $A$ , und  $\mathbb{1}_m$  linksneutral zu  $A$ . Bei quadratischen Matrizen fallen diese beiden Eigenschaften zusammen; damit haben quadratische Matrizen aus  $\mathbb{R}^{(n,n)}$  die Einheitsmatrix  $\mathbb{1}_n$  als *neutrales Element*.

Bis hierhin haben wir ermittelt, dass die quadratischen Matrizen mit der passenden Einheitsmatrix als neutralem Element bezüglich der Matrixmultiplikation ein *Monoid* bilden – jedoch ein nicht-kommutatives.

Da die  $\mathbb{R}^{(n,n)}$  auch ein  $\mathbb{R}$ -Vektorraum ist, gilt nach Definition 6.1, dass

$$(\mathbb{R}^{(n,n)}, +)$$

eine abelsche Gruppe ist (also bezüglich der komponentenweisen Addition von Matrixelementen).

Man rechnet außerdem leicht nach (durch Einsetzen der Definition des Matrizenprodukts), dass für quadratische Matrizen  $A, B, C \in \mathbb{R}^{(n,n)}$  die beiden Distributivgesetze gelten:

$$(A + B)C = AC + BC \quad \text{und} \quad A(B + C) = AB + AC$$

Alle beteiligten Matrizen sind aus  $\mathbb{R}^{(n,n)}$ .

Insgesamt haben wir also mit den quadratischen Matrizen, der komponentenweisen Addition und dem Matrizenprodukt eine Ringstruktur (siehe Definition 5.14):

**Satz 7.12** (Ring der quadratischen Matrizen). *Für  $n \in \mathbb{N}$  bilden die quadratischen Matrizen  $\mathbb{R}^{(n,n)}$  mit der Matrizenaddition und dem Matrixprodukt einen (nicht-kommutativen) Ring mit Eins; die Einheitsmatrix  $\mathbb{1}_n$  ist das neutrale Element der Multiplikation.*

**Beispiel:** Wir zeigen die Neutralität beim Multiplizieren mit der Einheitsmatrix am Beispiel der (von oben schon bekannten) Matrix

$$H := \begin{pmatrix} 5 & 5 \\ 3 & -3 \\ 4 & 14 \\ 5 & 3 \end{pmatrix}$$

Es ist:

$$H \cdot \mathbb{1}_2 = \begin{pmatrix} 5 & 5 \\ 3 & -3 \\ 4 & 14 \\ 5 & 3 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 5 & 5 \\ 3 & -3 \\ 4 & 14 \\ 5 & 3 \end{pmatrix}$$

Dies ist in sofern wenig überraschend, als die Einheitsmatrix aus den beiden Basisvektoren von  $E_2$  konstruierbar ist. Betrachtet man ihre erste Spalte und führt  $H \cdot \vec{e}_1$  aus, so erhält man das Bild des ersten Basisvektors, also die erste Spalte der Abbildungsmatrix. Analog führt  $H \cdot \vec{e}_2$  auf die zweite Spalte der Abbildungsmatrix. Siehe dazu die Verfahren, um aus linearen Abbildungsvorschriften die Abbildungsmatrizen zu ermitteln.

Das andere Produkt lautet:

$$\mathbb{1}_4 \cdot H = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 5 & 5 \\ 3 & -3 \\ 4 & 14 \\ 5 & 3 \end{pmatrix} = \begin{pmatrix} 5 & 5 \\ 3 & -3 \\ 4 & 14 \\ 5 & 3 \end{pmatrix}$$

Hier greift die Erklärung “Bilder der Basisvektoren” nicht (obwohl man mit Hilfe der Transposition wieder solch einen Fall herstellen könnte; siehe dazu den nächsten Abschnitt). Aber man sieht beim Rechnen ein, dass die Multiplikation der Einheitsmatrix von links dafür sorgt, dass jeweils nur genau ein Element der rechten Matrix selektiert wird und (dank des Faktors 1 von links) unverändert als Matrixelement des Produkts stehen bleibt.

## 7.4 Matrixtransposition

### 7.4.1 Definition

**Definition 7.13** (Transponierte Matrix). *Für eine Matrix  $A \in \mathbb{R}^{(m,n)}$  ist die transponierte Matrix  $A^T$  diejenige Matrix aus  $\mathbb{R}^{(n,m)}$ , für die für alle  $j \in \{1, \dots, n\}$  und  $k \in \{1, \dots, m\}$  gilt:*

$$(A^T)_{j,k} = A_{k,j}$$

### Bemerkungen:

- Durch Transposition werden also Zeilen zu Spalten und umgekehrt. Die Matrix  $A$  wird an ihrer *Diagonalen* (das ist die gedachte Linie, auf denen die Matricelemente  $A_{1,1}, A_{2,2}, \dots$  liegen) gespiegelt.
- Ist  $A$  die Abbildungsmatrix einer linearen Abbildung von  $\mathbb{R}^n$  nach  $\mathbb{R}^m$ , so ist  $A^T$  die Matrix einer linearen Abbildung von  $\mathbb{R}^m$  nach  $\mathbb{R}^n$ . *Dabei handelt es sich aber i.A. nicht um die Umkehrabbildung!*<sup>2</sup>
- Da wir uns gestatten, Vektoren und einspaltige Matrizen synonym zu verwenden, können wir durch Transposition einen (Spalten-)Vektor zu einem *Zeilenvektor* machen. Analog können wir gewöhnliche Vektoren (also solche in Spaltenform) als Transponierte von Zeilenvektoren notieren – das spart im Fließtext oft Platz.

### Beispiele:

- Für die von oben schon bekannten Matrizen

$$F := \begin{pmatrix} 1 & 3 \\ 1 & -1 \\ 2 & 0 \end{pmatrix}, \quad G := \begin{pmatrix} 2 & 1 & 1 \\ 0 & 3 & 0 \\ 4 & -2 & 1 \\ 1 & 0 & 2 \end{pmatrix} \quad \text{und} \quad H := \begin{pmatrix} 5 & 5 \\ 3 & -3 \\ 4 & 14 \\ 5 & 3 \end{pmatrix}$$

lauten die Transponierten:

$$F^T = \begin{pmatrix} 1 & 1 & 2 \\ 3 & -1 & 0 \end{pmatrix}, \quad G^T = \begin{pmatrix} 2 & 0 & 4 & 1 \\ 1 & 3 & -2 & 0 \\ 1 & 0 & 1 & 2 \end{pmatrix} \quad \text{und} \quad H^T = \begin{pmatrix} 5 & 3 & 4 & 5 \\ 5 & -3 & 14 & 3 \end{pmatrix}$$

- Noch zwei Beispiele mit Zeilen- und Spaltenvektoren:

$$\begin{pmatrix} 1 \\ 2 \\ 4 \\ 3 \end{pmatrix}^T = (1 \quad 2 \quad 4 \quad 3) \quad \text{und} \quad (42 \quad 47 \quad 11)^T = \begin{pmatrix} 42 \\ 47 \\ 11 \end{pmatrix}$$

### 7.4.2 Eigenschaften

Wir halten einige zentrale Eigenschaften transponierter Matrizen fest:

**Satz 7.14** (Eigenschaften transponierter Matrizen). *Es seien  $A, B$  reelle Matrizen mit jeweils geeigneter Dimensionierung, und  $c \in \mathbb{R}$ . Für die Transposition gilt:*

1.  $(A^T)^T = A$  (die Transposition ist eine Involution)
2.  $(A + B)^T = A^T + B^T$  sowie  $(cA)^T = c(A^T)$  (die Transposition vertauscht mit linearen Operationen)
3.  $(AB)^T = B^T \cdot A^T$  (Transponierte des Produkts über die Transponierten der Faktoren ausgedrückt)
4. Die Produkte  $A \cdot A^T$  und  $A^T \cdot A$  sind stets wohldefiniert; es sind quadratische Matrizen.

(Beweis: S. 328.)

### Beispiele:

- Für die Matrizen

$$F := \begin{pmatrix} 1 & 3 \\ 1 & -1 \\ 2 & 0 \end{pmatrix}, \quad G := \begin{pmatrix} 2 & 1 & 1 \\ 0 & 3 & 0 \\ 4 & -2 & 1 \\ 1 & 0 & 2 \end{pmatrix} \quad \text{und} \quad H := \begin{pmatrix} 5 & 5 \\ 3 & -3 \\ 4 & 14 \\ 5 & 3 \end{pmatrix}$$

---

<sup>2</sup>Die Ausnahme hiervon bilden genau die *orthogonalen Matrizen*, die im Kapitel 9 behandelt werden

war (siehe das Beispiel zu Satz 7.9 (Abbildungsmatrizen bei Komposition)):

$$H = G \cdot F$$

Wir rechnen nach, dass  $F^T \cdot G^T = H^T$ . Die Transponierten wurden im Beispiel zu Definition 7.13 schon angegeben – damit ergibt sich:

$$F^T \cdot G^T = \begin{pmatrix} 1 & 1 & 2 \\ 3 & -1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 2 & 0 & 4 & 1 \\ 1 & 3 & -2 & 0 \\ 1 & 0 & 1 & 2 \end{pmatrix} = \begin{pmatrix} 5 & 3 & 4 & 5 \\ 5 & -3 & 14 & 3 \end{pmatrix} = H^T \quad \checkmark$$

- Wir rechnen die beiden Produkte  $H \cdot H^T$  und  $H^T \cdot H$  mit gleichem  $H$  wie eben:

$$H \cdot H^T = \begin{pmatrix} 5 & 5 \\ 3 & -3 \\ 4 & 14 \\ 5 & 3 \end{pmatrix} \cdot \begin{pmatrix} 5 & 3 & 4 & 5 \\ 5 & -3 & 14 & 3 \end{pmatrix} = \begin{pmatrix} 50 & 0 & 90 & 40 \\ 0 & 18 & -30 & 6 \\ 90 & -30 & 212 & 62 \\ 40 & 6 & 62 & 34 \end{pmatrix}$$

$$H^T \cdot H = \begin{pmatrix} 5 & 3 & 4 & 5 \\ 5 & -3 & 14 & 3 \end{pmatrix} \cdot \begin{pmatrix} 5 & 5 \\ 3 & -3 \\ 4 & 14 \\ 5 & 3 \end{pmatrix} = \begin{pmatrix} 75 & 87 \\ 87 & 239 \end{pmatrix}$$

Wir beobachten, dass beide Produktmatrizen spiegelbildlich bzgl. ihrer jeweiligen Diagonalen sind – das ist kein Zufall, denn solche Produkte (Matrix mit ihrer eigenen Transponierten) sind stets *symmetrische* Matrizen – dazu mehr in Kürze.

### 7.4.3 Kanonisches Skalarprodukt

Wenn wir Vektoren als einspaltige Matrizen auffassen und einelementige Matrizen mit ihrem Element identifizieren, erlaubt uns die Transposition noch eine Neuformulierung des kanonischen Skalarprodukts:

$$\vec{x} \bullet \vec{y} = \vec{x}^T \cdot \vec{y}$$

**Beispiel:** Wir berechnen mit dieser Formel das Skalarprodukt von

$$\vec{x} := \begin{pmatrix} 4 \\ 2 \\ -3 \\ 0 \end{pmatrix} \quad \text{und} \quad \vec{y} := \begin{pmatrix} -1 \\ 2 \\ 2 \\ 42 \end{pmatrix}$$

Wir erhalten:

$$\vec{x}^T \cdot \vec{y} = \begin{pmatrix} 4 & 2 & -3 & 0 \end{pmatrix} \cdot \begin{pmatrix} -1 \\ 2 \\ 2 \\ 42 \end{pmatrix} = -6 = \vec{x} \bullet \vec{y}$$

Mit den Zeilenvektoren lässt sich auch ein beliebiges Matrixprodukt (falls es existiert) alternativ beschreiben. Sei also  $A \in \mathbb{R}^{(p,m)}$  und  $B \in \mathbb{R}^{(m,n)}$ . Dann ist das Produkt  $AB$  aus  $\mathbb{R}^{(p,n)}$ . Wenn wir nun die Matrizen wie folgt notieren:

$$A = \begin{pmatrix} \vec{v}_1^T \\ \vdots \\ \vec{v}_p^T \end{pmatrix} \quad \text{und} \quad B = \begin{pmatrix} \vec{b}_1 & \dots & \vec{b}_n \end{pmatrix},$$

dann gilt:

$$AB = \begin{pmatrix} \vec{v}_1^T \vec{b}_1 & \dots & \vec{v}_1^T \vec{b}_n \\ \vdots & & \vdots \\ \vec{v}_p^T \vec{b}_1 & \dots & \vec{v}_p^T \vec{b}_n \end{pmatrix}$$

Und allgemein:

$$(AB)_{j,k} = \vec{v}_j^T \cdot \vec{b}_k = \vec{v}_j \bullet \vec{b}_k$$

Man beachte aber, dass die Vektoren  $\vec{v}_j^T$  Zeilenvektoren sind – die korrespondierenden Spaltenvektoren ergeben die Spaltendarstellung der Transponierten

$$A^T = (\vec{v}_1 \quad \cdots \quad \vec{v}_p)$$

Das ist aber meist nicht identisch mit  $A$  – daher wurden die Zeilenvektoren von  $A$  auch nicht mit  $\vec{a}_1^T$  etc. bezeichnet.  $A$  besitzt natürlich auch eine Spaltendarstellung – nämlich

$$(A = \vec{a}_1 \quad \cdots \quad \vec{a}_m)$$

#### 7.4.4 Symmetrische Matrizen

**Definition 7.15** (Symmetrische Matrix). Eine quadratische Matrix  $A \in \mathbb{R}^{(n,n)}$  heißt symmetrisch, falls gilt:

$$A^T = A$$

Sie heißt dagegen schiefsymmetrisch, falls  $A^T = -A$ .

**Bemerkungen:**

- Für  $B \in \mathbb{R}^{(m,n)}$  sind die Produkte  $BB^T$  sowie  $B^TB$  symmetrische Matrizen, denn mit Satz 7.14 gilt:

$$\begin{aligned}(BB^T)^T &= (B^T)^T \cdot B^T = BB^T \\ (B^TB)^T &= B^T \cdot (B^T)^T = B^TB\end{aligned}$$

(Das hatten wir oben im Beispiel zu Satz 7.14 bereits beobachtet.)

- Jede quadratische Matrix  $C$  kann als Summe einer symmetrischen und einer schiefsymmetrischen Matrix geschrieben werden, denn es gilt:

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

(Die Transponierte von  $C$  wurde hier einmal addiert und einmal subtrahiert.)

Aber nun ist die linke Klammer symmetrisch, die rechte schiefsymmetrisch, denn

$$\begin{aligned}(C + C^T)^T &= C^T + C = (C + C^T) \\ (C - C^T)^T &= C^T - C = -(C - C^T)\end{aligned}$$

Eine besondere Klasse von symmetrischen Matrizen sind solche, die nur auf ihrer Diagonalen Elemente besitzen, die (ggf.) von null verschieden sind:

**Definition 7.16** (Diagonalmatrix). Eine quadratische Matrix  $A \in \mathbb{R}^{(n,n)}$  heißt Diagonalmatrix (oder: diagonal), wenn alle ihre off-Diagonal-Elemente verschwinden, d.h.

$$(j \neq k) \Rightarrow (A_{j,k} = 0)$$

Man notiert  $A$  dann auch durch Auflistung ihrer Diagonalelemente per

$$A = \text{diag}(a_1, \dots, a_n),$$

wobei für  $j \in \{1, \dots, n\}$  gilt:  $A_{j,j} = a_j$ .

**Bemerkungen:**

- Auch auf der Diagonale dürfen sich natürlich Nullen befinden – insbesondere ist die quadratische Nullmatrix auch diagonal.
- Jede Einheitsmatrix  $\mathbb{1}_n$  ist diagonal (und damit auch symmetrisch):

$$\mathbb{1}_n = \text{diag}(\underbrace{1, \dots, 1}_{n\text{-mal}})$$

**Beispiel:**

$$\text{diag}(3, 4, -7, 199) = \begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & -7 & 0 \\ 0 & 0 & 0 & 199 \end{pmatrix}$$

## 7.5 Inverse Matrizen (Einführung)

Wir erinnern uns an Satz 7.12 über den Ring der quadratischen Matrizen  $\mathbb{R}^{(n,n)}$ . Damit dieser Ring ein Körper sein könnte, müsste das Matrixprodukt kommutativ sein – das ist nicht der Fall. Aber  $\mathbb{R}^{(n,n)}$  ist auch kein Schiefkörper, denn schon für solche sind multiplikative Inverse nötig. Diese existieren aber nicht immer.

**Beispiel:** Für Nullteiler existiert kein Inverses der Multiplikation. Man betrachte das folgende Produkt:

$$\begin{pmatrix} 2 & 7 \\ -4 & -14 \end{pmatrix} \cdot \begin{pmatrix} 21 & -7 \\ -6 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Offenbar haben die beiden linken Matrizen also kein Inverses bzgl. des Matrixprodukts, denn ihr Produkt ist die Nullmatrix aus  $\mathbb{R}^{(2,2)}$  – das neutrale Element der Addition.

---

Wir werden (erst) im Kapitel 9 eine Methode entwickeln, das Inverse einer quadratischen Matrix eindeutig zu bestimmen (falls es existiert), welche die Erkenntnisse aus Kapitel 8 benötigt. Dort lernen wir auch Kriterien kennen, um die Existenz eines Inversen im Voraus zu entscheiden.<sup>3</sup>

---

Wir können aber an dieser Stelle schon eine Definition für inverse Matrizen geben und einige ihrer Eigenschaften untersuchen.

**Definition 7.17** (Inverse Matrix). Für  $n \in \mathbb{N}$  seien  $A, B \in \mathbb{R}^{(n,n)}$ . Dann heißt  $B$  Inverse von  $A$ , geschrieben  $A^{-1}$ , wenn gilt:

$$AB = BA = \mathbb{1}_n$$

Existiert die Inverse einer Matrix  $A$ , so heißt  $A$  invertierbar.

**Bemerkungen:**

- Nach Satz 5.7 über die Eindeutigkeit inverser Elemente dürfen wir hier sogar von *der* Inversen von  $A$  sprechen – es gibt dann keine andere quadratische Matrix, die obige Bedingung erfüllt. Daher schreibt man die obige Gleichung üblicherweise so:

$$A \cdot A^{-1} = A^{-1} \cdot A = \mathbb{1}_n$$

- Die Tatsache, dass sowohl  $AB$  als auch  $BA$  das neutrale Element  $\mathbb{1}_n$  ergeben, ist hier nicht trivial, da das Matrixprodukt nicht kommutativ ist.
- Gilt obige Gleichung, so ist auch  $A$  das Inverse von  $B$ .

**Beispiele:**

- Alle Einheitsmatrizen sind ihre eigenen Inversen. Das ist nicht verwunderlich, denn diese Eigenschaft haben allgemein alle neutralen Elemente von Monoiden.
- Die Matrizen

$$A := \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix} \quad \text{und} \quad B := \begin{pmatrix} -1 & 2 \\ 1 & -1 \end{pmatrix}$$

---

<sup>3</sup>Teaser: Im obigen Beispiel sind die Zeilen- bzw. Spaltenvektoren beider Matrizen linear abhängig; genau solche Matrizen sind nicht invertierbar.

sind zueinander invers, denn:

$$AB = \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} -1 & 2 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \mathbb{1}_2$$

$$BA = \begin{pmatrix} -1 & 2 \\ 1 & -1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \mathbb{1}_2$$

Also gilt:  $A^{-1} = B$  sowie  $B^{-1} = A$ .

- Diagonalmatrizen (siehe Definition 7.16) sind genau dann invertierbar, wenn alle Elemente auf der Diagonalen von null verschieden sind. Dann gilt:

$$(\text{diag}(a_1, \dots, a_n))^{-1} = \text{diag}\left(\frac{1}{a_1}, \dots, \frac{1}{a_n}\right)$$

Man vergewissert sich nämlich leicht, dass das Produkt zweier Diagonalmatrizen wieder diagonal ist, und dass

$$\text{diag}(a_1, \dots, a_n) \cdot \text{diag}(b_1, \dots, b_n) = \text{diag}(a_1 b_1, \dots, a_n b_n)$$

Wenn dieses Produkt also der Einheitsmatrix  $\mathbb{1}_n$  entsprechen soll, müssen alle Diagonalelemente aus  $\text{diag}(a_1, \dots, a_n)$  dazu mit ihrem jeweiligen Kehrwert multipliziert werden. Damit dieser existiert, müssen die Elemente ungleich 0 sein.

Exemplarisch:

$$\text{diag}(1, 4, -7) \cdot \text{diag}\left(1, \frac{1}{4}, -\frac{1}{7}\right) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & -7 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{4} & 0 \\ 0 & 0 & -\frac{1}{7} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \mathbb{1}_3$$

- Nur quadratische Matrizen können invertierbar sein. Für nichtquadratische Matrizen ergeben sich, selbst wenn  $AB$  und  $BA$  wohldefiniert sind, quadratische Matrizen *mit verschiedener Dimensionierung* – diese können nicht gleich sein!

Als Beispiel betrachten wir die Matrizen

$$A := \begin{pmatrix} 1 & 2 & 2 \\ 2 & 1 & -2 \end{pmatrix} \quad \text{und} \quad B := \frac{1}{9} \begin{pmatrix} 1 & 2 \\ 2 & 1 \\ 2 & -2 \end{pmatrix}$$

Zwar ist

$$AB = \frac{1}{9} \cdot \begin{pmatrix} 1 & 2 & 2 \\ 2 & 1 & -2 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 \\ 2 & 1 \\ 2 & -2 \end{pmatrix} = \frac{1}{9} \begin{pmatrix} 9 & 0 \\ 0 & 9 \end{pmatrix} = \mathbb{1}_2,$$

aber das umgekehrte Produkt ergibt eine  $(3 \times 3)$ -Matrix, die nicht einmal  $\mathbb{1}_3$  entspricht:

$$BA = \frac{1}{9} \begin{pmatrix} 1 & 2 \\ 2 & 1 \\ 2 & -2 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 & 2 \\ 2 & 1 & -2 \end{pmatrix} = \frac{1}{9} \begin{pmatrix} 5 & 4 & -2 \\ 4 & 5 & 2 \\ -2 & 2 & 8 \end{pmatrix}$$

Wir halten noch fünf wichtige Eigenschaften inverser Matrizen fest:

**Satz 7.18** (Eigenschaften inverser Matrizen). *A sei eine invertierbare Matrix aus  $\mathbb{R}^{(n,n)}$ . Dann gilt:*

1.  $(A^{-1})^{-1} = A$  (das Invertieren ist eine Involution)

2. Für  $c \in \mathbb{R} \setminus \{0\}$  ist  $(cA)^{-1} = \frac{1}{c} \cdot A^{-1}$

3. Auch  $A^T$  ist invertierbar, und es ist

$$(A^T)^{-1} = (A^{-1})^T$$

4. Ist  $A$  außerdem symmetrisch, so gilt dies auch für  $A^{-1}$

5. Falls  $B$  eine invertierbare Matrix aus  $\mathbb{R}^{(n,n)}$  ist, so ist auch das Produkt  $AB$  invertierbar, und es ist

$$(AB)^{-1} = B^{-1} \cdot A^{-1}$$

(Beweis: S. 329.)

### Beispiele:

- Wir hatten oben schon nachgerechnet, dass die Matrizen

$$A := \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix} \quad \text{und} \quad B := \begin{pmatrix} -1 & 2 \\ 1 & -1 \end{pmatrix}$$

zueinander invers sind, dass also  $B = A^{-1}$  gilt.

Wir rechnen nach, dass die Inverse von  $A^T$  durch  $B^T$  gegeben ist:

$$\begin{aligned} A^T B^T &= \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix} \cdot \begin{pmatrix} -1 & 1 \\ 2 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \mathbb{1}_2 \quad \checkmark \\ B^T A^T &= \begin{pmatrix} -1 & 1 \\ 2 & -1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \mathbb{1}_2 \quad \checkmark \end{aligned}$$

- Wir betrachten die dreireihigen Matrizen

$$A := \begin{pmatrix} 1 & 2 & 1 \\ 2 & -2 & 1 \\ 2 & 2 & 2 \end{pmatrix} \quad \text{und} \quad B := \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 2 \\ 1 & 2 & 1 \end{pmatrix}$$

Ihre Inversen sind (man überprüfe dies zur Übung durch Bilden der jeweiligen Matrixprodukte):

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

Zunächst bemerken wir, dass  $B$  und  $B^{-1}$  beide symmetrisch sind, wie in der vierten Aussage behauptet.

Wir betrachten nun noch die Produkte: Es gilt (auch das überprüfe man gerne durch Nachrechnen):

$$\begin{aligned} AB &= \begin{pmatrix} 1 & 2 & 1 \\ 2 & -2 & 1 \\ 2 & 2 & 2 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 2 \\ 1 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 4 & 5 & 6 \\ 1 & 2 & -1 \\ 6 & 8 & 8 \end{pmatrix} \\ BA &= \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 2 \\ 1 & 2 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 & 1 \\ 2 & -2 & 1 \\ 2 & 2 & 2 \end{pmatrix} = \begin{pmatrix} 5 & 2 & 4 \\ 7 & 4 & 6 \\ 7 & 0 & 5 \end{pmatrix} \\ A^{-1}B^{-1} &= \begin{pmatrix} 3 & 1 & -2 \\ 1 & 0 & -\frac{1}{2} \\ -4 & -1 & 3 \end{pmatrix} \cdot \begin{pmatrix} 3 & -1 & -1 \\ -1 & 0 & 1 \\ -1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 10 & -5 & -2 \\ \frac{7}{2} & -\frac{3}{2} & -1 \\ -14 & 7 & 3 \end{pmatrix} \\ B^{-1}A^{-1} &= \begin{pmatrix} 3 & -1 & -1 \\ -1 & 0 & 1 \\ -1 & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 3 & 1 & -2 \\ 1 & 0 & -\frac{1}{2} \\ -4 & -1 & 3 \end{pmatrix} = \begin{pmatrix} 12 & 4 & -\frac{17}{2} \\ -7 & -2 & 5 \\ -2 & -1 & \frac{3}{2} \end{pmatrix} \end{aligned}$$

Nun lässt sich nachrechnen, dass in der Tat

$$\begin{aligned} (AB) \cdot (B^{-1}A^{-1}) &= (B^{-1}A^{-1}) \cdot (AB) = \mathbb{1}_3 \\ \text{und} \quad (BA) \cdot (A^{-1}B^{-1}) &= (A^{-1}B^{-1}) \cdot (BA) = \mathbb{1}_3 \end{aligned}$$

Wir halten noch ein Detail fest, das uns etwas Rechenarbeit spart:

**Satz 7.19** (Links- und Rechtsinversität). *Für reelle Matrizen  $A, B \in \mathbb{R}^{(n,n)}$  sind die beiden Inversitätsbedingungen in Definition 7.17 gleichwertig, d.h.*

$$(AB = \mathbb{1}_n) \Leftrightarrow (BA = \mathbb{1}_n)$$

(Der Beweis ist etwas umfangreich und benötigt Resultate aus Kapitel 8, weswegen er auch dort zu finden ist – auf S. 332.)



**Bemerkung:** Die linke Gleichung bedeutet, dass  $B$  eine *Rechtsinverse* zu  $A$  ist – die rechte jedoch, dass  $B$  eine *Linksinverse* zu  $A$  ist.

Wegen der Äquivalenz reicht es also, eine der beiden Bedingungen zu zeigen.

Zum Abschluss dieses Abschnitts formulieren wir folgenden Satz:

**Satz 7.20** (Bijektive lineare Abbildungen). *Eine lineare Abbildung  $\vec{f}: \mathbb{R}^n \rightarrow \mathbb{R}^m$  ist genau dann bijektiv, wenn ihre Abbildungsmatrix  $F$  quadratisch und invertierbar ist. Die Abbildungsmatrix der Umkehrabbildung entspricht dann  $F^{-1}$ .*

(Beweis: S. 329.)

**Bemerkungen:**

- Insbesondere muss also  $n = m$  gelten; es gibt keine bijektiven Abbildungen zwischen reellen kartesischen Produkträumen mit verschiedenen Dimensionen.
- Ein Gegenbeispiel hatten wir bei Definition 7.17 schon gesehen: Mit den zwei nichtquadratischen Matrizen

$$A := \begin{pmatrix} 1 & 2 & 2 \\ 2 & 1 & -2 \end{pmatrix} \quad \text{und} \quad B := \frac{1}{9} \begin{pmatrix} 1 & 2 \\ 2 & 1 \\ 2 & -2 \end{pmatrix}$$

ließ sich zwar  $AB = \mathbb{1}_2$  realisieren, aber es war  $BA \neq \mathbb{1}_3$ .  $B$  ist also rechtsinvers zu  $A$ , aber nicht linksinvers. Dieser Satz zeigt, dass das kein Zufall ist, denn die zu  $A, B$  korrespondierenden linearen Abbildungen sind aufgrund ihrer nicht-quadratischen Matrizen nicht bijektiv.

## 7.6 Orthogonale Matrizen (Einführung)

Um die Klasse der orthogonalen Matrizen zu motivieren (die Bedeutung des Namens klären wir am Ende des Abschnitts weiter auf), untersuchen wir die Frage, welche bijektiven linearen Abbildungen das Skalarprodukt konstant lassen.

Wir beschränken uns auf  $\mathbb{R}^n$  und das kanonische Skalarprodukt und suchen also invertierbare Matrizen  $A \in \mathbb{R}^{(n,n)}$ , sodass für  $\vec{x}, \vec{y} \in \mathbb{R}^n$  stets gilt:

$$\vec{x} \bullet \vec{y} = (A\vec{x}) \bullet (A\vec{y})$$

Schon dies könnten wir komponentenweise anschreiben und kämen zum Ziel (das führe man gerne zur Übung aus!) – noch etwas kürzer geht es aber, wenn wir die Matrizenschreibweise des kanonischen Skalarprodukts aus Unterabschnitt 7.4.3 verwenden. Dann lautet obige Gleichheit:

$$\vec{x}^T \vec{y} = (A\vec{x})^T (A\vec{y})$$

Nun schreiben wir die rechte Seite um, indem wir die Transponierte des Produkts verwenden; danach klammern wir per Assoziativgesetz um:

$$(A\vec{x})^T (A\vec{y}) = (\vec{x}^T \cdot A^T)(A\vec{y}) = \vec{x}^T \cdot (A^T \cdot A) \cdot \vec{y}$$

Wenn dies aber dem Ausdruck  $\vec{x}^T \vec{y}$  entsprechen soll, muss gelten:

$$A^T \cdot A = \mathbb{1}_n$$

**Definition 7.21** (Orthogonale Matrizen). *Eine Matrix  $A \in \mathbb{R}^{(n,n)}$  heißt orthogonal, falls*

$$A^T = A^{-1}$$

**Bemerkung:** Mit Satz 7.19 ergeben sich daraus die äquivalenten Bestimmungsgleichungen

$$A^T \cdot A = A \cdot A^T = \mathbb{1}_n$$

### Beispiele:

- Die Matrix

$$A := \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$$

ist orthogonal, denn

$$A^T \cdot A = \frac{1}{\sqrt{2}} \cdot \frac{1}{\sqrt{2}} \cdot \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} = \mathbb{1}_2$$

- Auch die Matrix

$$B := \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} & 0 \\ \frac{\sqrt{3}}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

ist orthogonal, wie man durch Nachrechnen leicht prüfen kann. Bei  $A, B$  handelt es sich, wie wir in Kapitel 9 sehen werden, um *Drehmatrizen*.

- Ebenso lässt sich nachrechnen, dass

$$C := \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 & 1 \\ 0 & \sqrt{2} & 0 \\ 1 & 0 & -1 \end{pmatrix}$$

orthogonal ist. Hier handelt es sich um eine *Drehspiegelung*, also um eine Kombination aus Drehung und Spiegelung.

- Eine weitere orthogonale Matrix ist

$$D := \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

Solche Matrizen behandeln wir im nächsten Abschnitt noch weiter;  $D$  ist eine *Permutationsmatrix*.

---

Wir fassen die Motivation von oben zusammen als

**Satz 7.22** (Invarianz des Skalarprodukts bei orthogonalen Abbildungen). *Orthogonale Abbildungen  $\mathbb{R}^n \rightarrow \mathbb{R}^n$  erhalten das Skalarprodukt auf  $\mathbb{R}^n$ .*

**Bemerkung:** Da wir in Kapitel 6 schon gezeigt hatten, dass das kanonische Skalarprodukt in  $\mathbb{R}^n$  sowohl die Winkel zwischen Vektoren als auch mit der euklidischen Norm die Länge von Vektoren im kartesischen Koordinatensystem erklärt, sind orthogonale Abbildungen *winkeltreu* und *längentreu*.

**Beispiel:** Wir betrachten in  $\mathbb{R}^3$  die Vektoren

$$\vec{x} := \begin{pmatrix} 3 \\ 1 \\ 5 \end{pmatrix} \quad \text{und} \quad \vec{y} := \begin{pmatrix} -2 \\ 2 \\ -1 \end{pmatrix}$$

Ihr (kanonisches) Skalarprodukt beträgt

$$\vec{x} \bullet \vec{y} = -6 + 2 - 5 = -9$$

Ihre Längen mit euklidischer Norm sind

$$|\vec{x}| = \sqrt{9 + 1 + 25} = \sqrt{35} \quad \text{und} \quad |\vec{y}| = \sqrt{4 + 4 + 1} = 3$$

Mit der von eben bekannten orthogonalen Matrix

$$C := \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 & 1 \\ 0 & \sqrt{2} & 0 \\ 1 & 0 & -1 \end{pmatrix}$$

ergeben sich die Bilder

$$C\vec{x} = \frac{1}{\sqrt{2}} \begin{pmatrix} 8 \\ \sqrt{2} \\ -2 \end{pmatrix} \quad \text{und} \quad C\vec{y} = \frac{1}{\sqrt{2}} \begin{pmatrix} -3 \\ 2\sqrt{2} \\ -1 \end{pmatrix}$$

Deren Skalarprodukt beträgt

$$(C\vec{x}) \bullet (C\vec{y}) = \frac{1}{2}(-24 + 4 + 2) = -\frac{18}{2} = -9 \quad \checkmark$$

Und ihre Längen sind

$$|C\vec{x}| = \frac{1}{\sqrt{2}} \sqrt{64 + 2 + 4} = \sqrt{\frac{70}{2}} = \sqrt{35} \quad \checkmark$$

$$|C\vec{y}| = \frac{1}{\sqrt{2}} \sqrt{9 + 8 + 1} = \sqrt{\frac{18}{2}} = 3 \quad \checkmark$$

Da sowohl die Längen als auch das Skalarprodukt erhalten bleiben, gilt das gleiche auch für den Winkel zwischen den beiden Vektoren (bzw. ihren Bildern), nach Satz 6.15.

---

Wir halten noch fest:

**Satz 7.23** (Eigenschaften orthogonaler Matrizen). *Falls  $A \in \mathbb{R}^{(n,n)}$  orthogonal ist, so gilt:*

1.  $A^T$  ist ebenfalls orthogonal.
2. Falls  $B \in \mathbb{R}^{(n,n)}$  ebenfalls orthogonal ist, so sind auch  $AB$  und  $BA$  orthogonal.

(Beweis: S. 330. Beispiel folgt am Ende des Abschnitts.)

---

Nun noch zum Hintergrund der Bezeichnung “orthogonal”. Wir hatten gesehen, dass eine quadratische Matrix  $A \in \mathbb{R}^{(n,n)}$  genau dann orthogonal ist, wenn sie

$$A^T \cdot A = \mathbb{1}_n$$

erfüllt.

Wenn wir uns an die Bemerkungen im Unterabschnitt 7.4.3 erinnern, können wir dies noch umschreiben. Denn die Zeilenvektoren der linken Matrix  $A^T$  im obigen Produkt entsprechen aufgrund der Transposition genau den Spaltenvektoren von  $A$ . Und damit folgt für die Matrixelemente des Produkts, falls  $A$  die Spaltendarstellung  $A = (\vec{a}_1 \ \cdots \ \vec{a}_n)$  besitzt:

$$(A^T \cdot A)_{j,k} = \vec{a}_j \bullet \vec{a}_k = (\mathbb{1}_n)_{j,k} = \delta_{jk}$$

(Auch ohne den referenzierten Unterabschnitt lässt sich das komponentenweise über die Definition des Matrixprodukts nachrechnen.)

Daraus erkennen wir zwei Dinge: Je zwei verschiedene Spaltenvektoren von  $A$  sind *orthogonal* zueinander (denn ihr kanonisches Skalarprodukt verschwindet; daher die Bezeichnung) – aber sie sind darüber hinaus noch *normiert*, da für jedes  $j$  dann gilt, dass  $\vec{a}_j \bullet \vec{a}_j = 1$ .

Wir halten dies als gleichwertige Definition für orthogonale Matrizen fest:

**Satz 7.24** (Alternatives Kriterium für orthogonale Matrizen). *Eine quadratische Matrix*

$$A := (\vec{a}_1 \ \cdots \ \vec{a}_n) \in \mathbb{R}^{(n,n)}$$

*ist genau dann orthogonal, wenn ihre Spaltenvektoren für  $j, k \in \{1, \dots, n\}$  die Bedingung*

$$\vec{a}_j \bullet \vec{a}_k = \delta_{jk}$$

*erfüllen.*

### Bemerkungen:

- Jede Matrix, die aus  $n$  normierten und paarweise orthogonalen Vektoren gebildet wird, ist also orthogonal. Man beachte, dass die Normiertheit ebenfalls gefordert ist.
- Die Spaltenvektoren einer orthogonalen Matrix bilden damit stets eine *Orthonormalbasis* (ONB) des  $\mathbb{R}^n$ . Die Standardbasis  $E_n$  ist aus den kartesischen Einheitsvektoren konstruiert, und  $\mathbb{1}_n$  ist symmetrisch, selbst-invers und orthogonal.
- Nach Satz 7.23 ist für eine orthogonale Matrix  $A$  auch immer  $A^T$  orthogonal. Für deren Spaltenvektoren gilt entsprechendes. Und damit gelten die obigen Beobachtungen *genauso auch für die Zeilenvektoren von  $A$* , wenn man sie (etwa zum Berechnen des Skalarprodukts) transponiert.

**Beispiele:** Wir betrachten nochmal die Matrizen  $A, B, C, D$  von den Beispielen zur Definition 7.21.

- Für

$$A := \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$$

sind beide Spalten (und beide Zeilen, wenn man sie transponiert) Vektoren der Länge 1, und ihr Skalarprodukt verschwindet.

- Für

$$C := \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 & 1 \\ 0 & \sqrt{2} & 0 \\ 1 & 0 & -1 \end{pmatrix}$$

sind alle drei Spalten (Zeilen) normiert. Die zweite Spalte (Zeile) ist orthogonal zu den beiden anderen, da in den Skalarprodukten jeweils eine Summe aus drei Nullen vorliegt. Die erste und dritte Spalte sind analog zu  $A$  ebenfalls orthogonal zueinander.

- Für

$$B := \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} & 0 \\ \frac{\sqrt{3}}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

beobachten wir aufgrund ebenso vieler günstig platzierter Nullen wie bei  $C$ , dass die Spalten (Zeilen) paarweise orthogonal zueinander sind. Wir rechnen noch die Länge der zweiten Spalte aus:

$$\sqrt{\frac{3}{4} + \frac{1}{4} + 0} = \sqrt{\frac{4}{4}} = 1$$

- Für

$$D := \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

stellen wir fest, dass die Spalten eine *Permutation* der kartesischen Basisvektoren aus  $E_4$  darstellen. Sie sind offensichtlich normiert und auch paarweise orthogonal, denn jede Zeile und jede Spalte enthält jeweils nur genau eine Komponente mit Wert 1; alle anderen sind 0. Damit verschwinden alle Skalarprodukte von Spalten, die nicht identisch sind (denn die beiden Komponenten mit Wert 1 würden dann in unterschiedlichen Zeilen liegen, sodass sie nicht zur Summe des kanonischen Skalarprodukts beitragen können).

**Beispiel:** (Zum Satz 7.23): Für die von oben bekannten orthogonalen Matrizen

$$B := \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} & 0 \\ \frac{\sqrt{3}}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{und} \quad C := \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 & 1 \\ 0 & \sqrt{2} & 0 \\ 1 & 0 & -1 \end{pmatrix}$$

berechnen wir die beiden Produkte und prüfen sie mit Satz 7.24 auf Orthogonalität (alternativ könnten wir die Matrizen mit ihren Transponierten multiplizieren!). Wir erhalten:

$$BC = \frac{1}{\sqrt{2}} \cdot \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} & 0 \\ \frac{\sqrt{3}}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 1 \\ 0 & \sqrt{2} & 0 \\ 1 & 0 & -1 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{6}}{2} & \frac{1}{2} \\ \frac{\sqrt{3}}{2} & \frac{\sqrt{2}}{2} & \frac{\sqrt{3}}{2} \\ 1 & 0 & -1 \end{pmatrix}$$

$$CB = \frac{1}{\sqrt{2}} \cdot \begin{pmatrix} 1 & 0 & 1 \\ 0 & \sqrt{2} & 0 \\ 1 & 0 & -1 \end{pmatrix} \cdot \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} & 0 \\ \frac{\sqrt{3}}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} & 1 \\ \frac{\sqrt{6}}{2} & \frac{\sqrt{2}}{2} & 0 \\ \frac{1}{2} & -\frac{\sqrt{3}}{2} & -1 \end{pmatrix}$$

Wir rechnen jeweils die Orthogonalität der Spaltenvektoren nach (den Vorfaktor an der Matrix können wir hier ignorieren, da wir mit dem Ergebnis 0 rechnen). Es sind aufgrund der Symmetrie des Skalarprodukts je drei Rechnungen auszuführen; wir rechnen zuerst das Produkt von erster und zweiter Spalte, danach das von erster und dritter, dann das von zweiter und dritter.

Für  $BC$ :

$$-\frac{\sqrt{6}}{4} + \frac{\sqrt{6}}{4} + 0 = 0; \quad \frac{1}{4} + \frac{3}{4} - 1 = 0; \quad -\frac{\sqrt{6}}{4} + \frac{\sqrt{6}}{4} + 0 = 0 \quad \checkmark$$

Für  $CB$ :

$$-\frac{\sqrt{3}}{4} + \frac{\sqrt{12}}{4} - \frac{\sqrt{3}}{4} = -2\frac{\sqrt{3}}{4} + \frac{2\sqrt{3}}{4} = 0; \quad \frac{1}{2} + 0 - \frac{1}{2} = 0; \quad -\frac{\sqrt{3}}{2} + 0 + \frac{\sqrt{3}}{2} = 0 \quad \checkmark$$

Nun ist noch die Normiertheit der Spalten zu prüfen. Wegen des Vorfaktors erwarten wir, dass alle Spalten (ohne den Vorfaktor) jeweils eine Länge von  $\sqrt{2}$  besitzen, d.h. dass ihr Skalarprodukt mit sich selbst genau 2 beträgt – letzteres rechnen wir nach.

Für  $BC$ :

$$\frac{1}{4} + \frac{3}{4} + 1 = 2; \quad \frac{6}{4} + \frac{2}{4} + 0 = 2; \quad \frac{1}{4} + \frac{3}{4} + 1 = 2 \quad \checkmark$$

Für  $CB$ :

$$\frac{1}{4} + \frac{6}{4} + \frac{1}{4} = 2; \quad \frac{3}{4} + \frac{2}{4} + \frac{3}{4} = 2; \quad 1 + 0 + 1 = 2 \quad \checkmark$$

## 7.7 Permutationsmatrizen

**Definition 7.25** (Permutationsmatrix). Für  $n \in \mathbb{N}$  und eine Permutation  $\sigma \in S_n$  heißt die Matrix  $P_\sigma \in \mathbb{R}^{(n,n)}$  mit der Spaltendarstellung

$$P_\sigma = (\vec{e}_{\sigma(1)} \quad \vec{e}_{\sigma(2)} \quad \cdots \quad \vec{e}_{\sigma(n)})$$

Permutationsmatrix von  $\sigma$ . Die Vektoren  $\vec{e}_j$  sind die kartesischen Einheitsvektoren der Standardbasis  $E_n$ .

**Bemerkungen:**

- Wegen Satz 1.40 gibt es  $n!$  verschiedene Permutationsmatrizen in  $\mathbb{R}^{(n,n)}$
- Die Standardbasis ist eine ONB (ihre Basisvektoren sind orthonormiert). Diese Eigenschaft überträgt sich auch auf die Spaltenvektoren einer beliebigen Permutationsmatrix – egal in welcher Reihenfolge die Basisvektoren dort vorliegen, solange jeder Basisvektor aus  $E_n$  genau einer der Spalten entspricht.

Also sind alle Permutationsmatrizen nach Satz 7.24 *orthogonal*.

- Insbesondere ist  $\mathbb{1}_n$  die Permutationsmatrix der identischen Permutation  $\text{id}_n \in S_n$ .
- Man beachte, dass in der Literatur auch oft die Permutationsmatrizen über ihre *Zeilen* statt über die Spalten definiert werden. Dadurch ändern sich nachgelagerte Formeln. Wir verwenden in dieser Vorlesung die obige Definition über die Spaltenvektoren.

**Beispiel:** Die Matrix

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} = (\vec{e}_3 \quad \vec{e}_1 \quad \vec{e}_4 \quad \vec{e}_2) =: P_\sigma$$

aus dem vorigen Abschnitt (dort hieß sie  $D$ ) ist eine Permutationsmatrix, und zwar zur Permutation

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 4 & 2 \end{pmatrix} = (1 \quad 3 \quad 4 \quad 2) \in S_4$$

Wir untersuchen nun beispielhaft das Verhalten der Permutationsmatrizen bei Multiplikation.

**Beispiel:** Zunächst betrachten wir eine beliebige vierspaltige Matrix  $A$  und die eben gefundene Permutationsmatrix  $P_\sigma$ , die wir von rechts multiplizieren können. Die Zeilenzahl von  $A$  ist dafür nicht wichtig; wir wählen eine Matrix mit zwei Zeilen:

$$A \cdot P_\sigma = \begin{pmatrix} A_{1,1} & A_{1,2} & A_{1,3} & A_{1,4} \\ A_{2,1} & A_{2,2} & A_{2,3} & A_{2,4} \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} A_{1,3} & A_{1,1} & A_{1,4} & A_{1,2} \\ A_{2,3} & A_{2,1} & A_{2,4} & A_{2,2} \end{pmatrix}$$

Offenbar permutiert dies die *Spalten* von  $A$  gemäß  $\sigma$ .

Wollen wir dagegen  $P_\sigma$  von links an eine Matrix  $B$  multiplizieren, so muss diese vier Zeilen besitzen; für die Spaltenzahl wählen wir hier 3:

$$P_\sigma \cdot B = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} B_{1,1} & B_{1,2} & B_{1,3} \\ B_{2,1} & B_{2,2} & B_{2,3} \\ B_{3,1} & B_{3,2} & B_{3,3} \\ B_{4,1} & B_{4,2} & B_{4,3} \end{pmatrix} = \begin{pmatrix} B_{2,1} & B_{2,2} & B_{2,3} \\ B_{4,1} & B_{4,2} & B_{4,3} \\ B_{1,1} & B_{1,2} & B_{1,3} \\ B_{3,1} & B_{3,2} & B_{3,3} \end{pmatrix}$$

Offenbar werden hier die *Zeilen* permutiert – allerdings nicht gemäß  $\sigma$ , sondern gemäß  $\sigma^{-1}$ , der inversen Permutation. Würden wir die Zeilen von  $B$  gemäß  $\sigma$  permutieren wollen, so müssten wir von links die Matrix der inversen Permutation multiplizieren; das ist:

$$P_{\sigma^{-1}}$$

Da aber Permutationen offenbar lineare Abbildungen sind, und obendrein bijektiv (das wissen wir schon aus Kapitel 4), können wir Satz 7.20 (Bijektive lineare Abbildungen) verwenden und erhalten, dass die Matrix gegeben ist durch

$$P_{\sigma^{-1}} = (P_\sigma)^{-1} = P_\sigma^T$$

(Für die zweite Gleichheit haben wir die Orthogonalität der Permutationsmatrizen verwendet.)

Also können wir die Zeilen von  $B$  gemäß  $\sigma$  permutieren, wenn wir von links die Transponierte  $P_\sigma^T$  multiplizieren. Wir rechnen nach:

$$P_\sigma^T \cdot B = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} B_{1,1} & B_{1,2} & B_{1,3} \\ B_{2,1} & B_{2,2} & B_{2,3} \\ B_{3,1} & B_{3,2} & B_{3,3} \\ B_{4,1} & B_{4,2} & B_{4,3} \end{pmatrix} = \begin{pmatrix} B_{3,1} & B_{3,2} & B_{3,3} \\ B_{1,1} & B_{1,2} & B_{1,3} \\ B_{4,1} & B_{4,2} & B_{4,3} \\ B_{2,1} & B_{2,2} & B_{2,3} \end{pmatrix} \quad \checkmark$$

Falls wir nun Zeilen *und* Spalten permutieren möchten, können wir von links und von rechts multiplizieren. Dann muss die mittlere Matrix allerdings quadratisch sein; in unserem Beispiel aus  $\mathbb{R}^{(4,4)}$ . Für solch eine Matrix  $C$  rechnen wir allgemein nach (Wir rechnen zuerst das rechte Produkt – wegen der Assoziativität des Produkts dürfte man natürlich auch zuerst das linke rechnen!):

$$\begin{aligned} P_\sigma^T \cdot C \cdot P_\sigma &= \begin{pmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} C_{1,1} & C_{1,2} & C_{1,3} & C_{1,4} \\ C_{2,1} & C_{2,2} & C_{2,3} & C_{2,4} \\ C_{3,1} & C_{3,2} & C_{3,3} & C_{3,4} \\ C_{4,1} & C_{4,2} & C_{4,3} & C_{4,4} \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} C_{1,3} & C_{1,1} & C_{1,4} & C_{1,2} \\ C_{2,3} & C_{2,1} & C_{2,4} & C_{2,2} \\ C_{3,3} & C_{3,1} & C_{3,4} & C_{3,2} \\ C_{4,3} & C_{4,1} & C_{4,4} & C_{4,2} \end{pmatrix} \\ &= \begin{pmatrix} C_{3,3} & C_{3,1} & C_{3,4} & C_{3,2} \\ C_{1,3} & C_{1,1} & C_{1,4} & C_{1,2} \\ C_{4,3} & C_{4,1} & C_{4,4} & C_{4,2} \\ C_{2,3} & C_{2,1} & C_{2,4} & C_{2,2} \end{pmatrix} \end{aligned}$$

Hier haben wir eine Matrix erhalten, die in Zeile  $j$  und Spalte  $k$  das Element  $C_{\sigma(j),\sigma(k)}$  besitzt. Man kann in der allgemeinen Schreibweise die Permutation auf den Diagonalelementen wieder erkennen.

Wir halten die Erkenntnisse, die sich auch allgemein beweisen lassen, fest im

**Satz 7.26** (Wirkung von Permutationsmatrizen). *Für  $n \in \mathbb{N}$  und einer Permutation  $\sigma \in S_n$  mit Permutationsmatrix  $P_\sigma$  gilt (jeweils komponentenweise für gültige Indexkombinationen):*

- Für  $A \in \mathbb{R}^{(m,n)}$  ist  $A \cdot P_\sigma$  in  $\mathbb{R}^{(m,n)}$  mit

$$(A \cdot P_\sigma)_{j,k} = A_{j,\sigma(k)}$$

(Multiplikation mit  $P_\sigma$  von rechts permutiert die Spalten)

- Für  $B \in \mathbb{R}^{(n,p)}$  ist  $P_\sigma^T \cdot B$  in  $\mathbb{R}^{(n,p)}$  mit

$$(P_\sigma^T \cdot B)_{j,k} = B_{\sigma(j),k}$$

(Multiplikation mit  $P_\sigma^T$  von links permutiert die Zeilen)

- Für  $C \in \mathbb{R}^{(n,n)}$  ist  $P_\sigma^T \cdot C \cdot P_\sigma$  in  $\mathbb{R}^{(n,n)}$  mit

$$(P_\sigma^T \cdot C \cdot P_\sigma)_{j,k} = C_{\sigma(j),\sigma(k)}$$

(Permutation sowohl der Spalten als auch der Zeilen)

(Beweis: S. 331.)

Neben der Inversion von Permutationsmatrizen, die wir eben schon benötigt hatten, ist auch die *Komposition* von Permutationen interessant. Wir untersuchen dies wieder zunächst an einem Beispiel; die allgemeine Formel wird dann im Anhang bewiesen.

**Beispiel:** Wir betrachten die beiden  $S_5$ -Permutationen

$$\sigma := \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 1 & 5 & 2 & 3 \end{pmatrix} \quad \text{und} \quad \pi := \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 2 & 1 & 3 & 4 \end{pmatrix}$$

Die zugehörigen Matrizen sind:

$$P_\sigma = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix} \quad \text{und} \quad P_\pi = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Nun hatten wir über die Komposition von linearen Abbildungen das Matrixprodukt motiviert – es liegt also nahe, die beiden Produkte der Matrizen zu untersuchen. Vor der Rechnung können wir schon sagen, dass es sich dabei wieder um Permutationsmatrizen handeln wird – denn jede Zeile und jede Spalte der beiden Faktoren  $P_\sigma, P_\pi$  enthält je genau eine Komponente 1; alle anderen verschwinden. Es wird also zu jeder Zeile des linken Faktors genau eine Spalte des rechten Faktors geben, die besagte Komponente 1 an der “richtigen” Stelle besitzt, sodass sich ein Skalarprodukt von 1 ergibt; sonst wird es keine Beiträge geben. Dann ist aber auch klar, dass in allen anderen Spalten dieser Zeile in der Produktmatrix 0 stehen muss, denn dort steht die 1 in den Spaltenvektoren des rechten Faktors dann nicht an einer der anderen Stellen – also versetzt. Man überprüfe das durch Nachrechnen und wird dabei genau diesen Effekt bemerken.

Wir erhalten:

$$P_\pi \cdot P_\sigma = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{pmatrix} = (\vec{e}_3 \quad \vec{e}_5 \quad \vec{e}_4 \quad \vec{e}_2 \quad \vec{e}_1)$$

$$P_\sigma \cdot P_\pi = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix} = (\vec{e}_3 \quad \vec{e}_1 \quad \vec{e}_4 \quad \vec{e}_5 \quad \vec{e}_2)$$

Also bekommen wir tatsächlich Permutationsmatrizen heraus, und zwar zu (in dieser Reihenfolge)

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 5 & 4 & 2 & 1 \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 1 & 4 & 5 & 2 \end{pmatrix}$$

Man überprüft leicht, dass es sich dabei um die Permutationen

$$\pi \circ \sigma \quad \text{und} \quad \sigma \circ \pi$$

handelt.

Wir halten dieses Ergebnis (das sich auch allgemein beweisen lässt) wieder fest als

**Satz 7.27** (Komposition von Permutationsmatrizen). *Für  $n \in \mathbb{N}$  und Permutationen  $\sigma, \pi \in S_n$  mit Permutationsmatrizen  $P_\sigma, P_\pi$  gilt:*

$$P_{\pi \circ \sigma} = P_\pi \cdot P_\sigma$$

(Beweis: S. 331.)

### Bemerkungen:

- Man beachte, dass bei der Konvention, welche Permutationsmatrizen über Zeilen definiert, diese Eigenschaft nicht in gleicher Form gilt – dort sind die beiden Faktoren im Matrixprodukt vertauscht.
- Offenbar bilden die Permutationsmatrizen aus  $\mathbb{R}^{(n,n)}$  mit der Matrixmultiplikation eine nicht-abelsche Gruppe – diese ist aber isomorph zur bereits bekannten symmetrischen Gruppe  $S_n$  (das waren die  $n$ -stelligen Permutationen bezüglich der Komposition) und bringt deswegen keine neuen Erkenntnisse.



# Kapitel 8

## Lineare Algebra: Gleichungssysteme und Determinanten

Wir betrachten in diesem Kapitel ein Lösungsverfahren für *lineare Gleichungssysteme* (das sind gekoppelte, also gleichzeitig zu erfüllende, affin-lineare Gleichungen mit einer oder mehreren Unbekannten). Während unsere Methode mathematisch zum Ziel führt, werden solche Systeme in der Praxis meist eher *numerisch* gelöst – also näherungsweise. Das ist am Computer schon deswegen erforderlich, weil nicht-ganze Zahlen nur mit beschränkter Genauigkeit speicherbar sind. Eine analytisch genaue Rechnung ist nur mit *Computeralgebrasystemen* wie Maple oder Mathematica möglich (dort werden Zahlen wie  $\sqrt{2}$  symbolisch behandelt und nach den Rechenregeln richtig weiter verarbeitet; auch rationale Zahlen sind als Brüche genau darstellbar), dies ist aber zeitintensiver – und oft zählt in der Praxis auch eher ein genügend genaues Ergebnis.

Mit unserem Verfahren stellen wir aber mathematisch sicher, dass lineare Gleichungssysteme überhaupt rechenbar sind. Wir stellen zu Anfang einen Zusammenhang her zwischen diesen Systemen, dem Produkt von Matrix und Vektor sowie der Linearkombination von Vektoren – diese drei Perspektiven stellen sich als gleichwertig heraus. Auch das Konzept der linearen Abhängigkeit (siehe Definition 6.21) wird eine bedeutende Rolle spielen.

Der zweite Teil des Kapitels behandelt die *Determinanten* quadratischer Matrizen: Das sind Funktionen, die jeder reellen quadratischen Matrix eine reelle Zahl zuordnen, die wiederum eng verknüpft ist mit der Frage der linearen Abhängigkeit, bzw. der Lösbarkeit von linearen Gleichungssystemen.

Mit den Erkenntnissen dieses Kapitels wird es danach (Kapitel 9) möglich sein, die Inversen quadratischer Matrizen (sofern existent) geschlossen zu bestimmen.

### 8.1 Lineare Gleichungssysteme

#### 8.1.1 Definition und Einordnung

**Definition 8.1** (Lineares Gleichungssystem). *Ein (reelles) lineares Gleichungssystem (LGS) mit  $m$  Gleichungen und  $n$  Unbekannten  $x_1, \dots, x_n \in \mathbb{R}$  liegt vor, falls die Gleichungen folgende Struktur aufweisen:*

$$\begin{array}{ccccccccc} & A_{1,1}x_1 & + & A_{1,2}x_2 & + & \cdots & + & A_{1,n}x_n & = & y_1 \\ \wedge & A_{2,1}x_1 & + & A_{2,2}x_2 & + & \cdots & + & A_{2,n}x_n & = & y_2 \\ & \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ \wedge & A_{m,1}x_1 & + & A_{m,2}x_2 & + & \cdots & + & A_{m,n}x_n & = & y_m \end{array}$$

Die Zahlen  $A_{j,k} \in \mathbb{R}$  heißen Koeffizienten des LGS.

Die Zahlen  $y_1, \dots, y_m$  beschreiben die Inhomogenität des LGS. Betragen sie alle 0, so heißt das LGS homogen.

Die Menge aller Tupel  $(x_1, \dots, x_n) \in \mathbb{R}^n$ , für die das LGS bei gegebenen Koeffizienten und gegebener Inhomogenität den Wahrheitswert  $\mathcal{W}$  annimmt, heißt Lösung des LGS.

Jedem LGS mit nicht-verschwindender Inhomogenität ist (über die Koeffizienten) genau ein homogenes LGS zugeordnet.

Die Lösung eines homogenen LGS heißt Kern des LGS.

### Bemerkungen:

- Die Abkürzung “LGS” gilt für uns offiziell.
- Jede der  $m$  Gleichungen ist für sich eine logische Aussage, die genau dann den Wahrheitswert  $\mathcal{W}$  hat, wenn die Gleichung erfüllt ist. Das LGS ist per “ $\wedge$ ” eine Und-Verknüpfung von Gleichungen. Der Wahrheitswert des LGS kann also nur genau dann  $\mathcal{W}$  sein, wenn *alle* Gleichungen gleichzeitig den Wert  $\mathcal{W}$  annehmen.
- Ziel dieses Abschnitts ist es, durch *Äquivalenzumformungen* das LGS so zu manipulieren, dass ein leichter lesbares LGS entsteht, aus dem idealerweise die einzelnen Variablen  $x_1, \dots, x_n$  *entkoppelt* sind – dann liegen nämlich jeweils lineare reelle Gleichungen vor, aus denen man die Lösungen direkt berechnen kann (siehe Kapitel 1).

Entscheidend wird aber sein, bei diesen Umformungen stets logische Äquivalenz sicher zu stellen, damit das vereinfachte LGS genau den gleichen Wahrheitswert besitzt wie das ursprüngliche. Das werden wir, wie bei Gleichungen üblich, mit dem Symbol “ $\Leftrightarrow$ ” ausdrücken.

**Beispiele:** Wir geben drei LGS an, die wir später bei der Lösungssuche weiter bearbeiten:

- Das folgende LGS besitzt drei Gleichungen für drei Veränderliche:

$$\begin{array}{rclcl} & 2x_1 & + & x_2 & + & 2x_3 & = & 3 \\ \wedge & x_1 & & & + & 2x_3 & = & 1 \\ \wedge & 3x_1 & + & 2x_2 & + & x_3 & = & 4 \end{array} \quad (*)$$

Wir notieren aus Gründen der Lesbarkeit die Veränderlichen spaltenweise, sodass alle  $x_j$  mit ihren zugehörigen Koeffizienten übereinander stehen; das erleichtert spätere Rechnungen.

Wo ein Koeffizient 0 beträgt, notieren wir in dieser Schreibweise die Veränderliche nicht, sondern lassen in der betreffenden Spalte eine Lücke (hier in der zweiten Gleichung bei  $x_2$ ).

- Hier ein LGS mit drei Gleichungen in vier Variablen:

$$\begin{array}{rclclclcl} & 2x_1 & + & 4x_2 & + & x_3 & + & x_4 & = & 2 \\ \wedge & 3x_1 & + & x_2 & + & 2x_3 & + & x_4 & = & 5 \\ \wedge & x_1 & - & 3x_2 & + & x_3 & & & = & 3 \end{array} \quad (**)$$

- Und noch ein weiteres LGS mit drei Gleichungen für drei Unbekannte:

$$\begin{array}{rclclcl} & 2x_1 & + & 2x_2 & + & x_3 & = & 3 \\ \wedge & x_1 & - & x_2 & + & 2x_3 & = & 4 \\ \wedge & x_1 & + & 3x_2 & - & x_3 & = & 1 \end{array} \quad (***)$$

Nun wollen wir den Zusammenhang zur linearen Algebra herstellen. Wenn wir uns an Definition 7.4 (Produkt aus Matrix und Vektor) erinnern, so ist klar, dass ein LGS wie in obiger Definition äquivalent mit einer vektoriellen Gleichung ausgedrückt werden kann, nämlich per

$$\boxed{A\vec{x} = \vec{y}} \quad \text{mit} \quad y_j = \sum_{k=1}^n A_{j,k} x_k \quad \text{für } y \in \{1, \dots, m\}$$

mit der *Koeffizientenmatrix*  $A \in \mathbb{R}^{(m,n)}$ , dem (unbekannten) *Lösungsvektor*  $\vec{x} \in \mathbb{R}^n$  sowie der *Inhomogenität*  $\vec{y} \in \mathbb{R}^m$ . Die Koeffizienten des LGS sind dabei (in gleicher Notation) die Koeffizienten der Matrix  $A$ . Es handelt sich um eine *lineare Abbildung* vom Urbild  $\vec{x}$  auf den Bildvektor  $\vec{y}$ .

Über die vektorielle Notation des LGS kommen wir noch zu einem weiteren wichtigen Aspekt: Wenn wir nämlich die Spaltendarstellung der Koeffizientenmatrix  $A$  betrachten, sehen wir, dass das LGS (bzw. die eingerahmte Vektorgleichung) ebenfalls äquivalent ist zu

$$\boxed{x_1 \vec{a}_1 + x_2 \vec{a}_2 + \dots + x_n \vec{a}_n = \vec{y}} \quad \text{mit} \quad A = (\vec{a}_1 \quad \vec{a}_2 \quad \dots \quad \vec{a}_n)$$

Dabei gilt wie üblich  $A_{j,k} = (\vec{a}_k)_j$ . Hier liegt eine *Linearkombination* der Inhomogenität  $\vec{y}$  aus den Spaltenvektoren von  $A$  vor. Vorsicht: Bei Linearkombinationen heißen die Skalierungsfaktoren  $x_j$  Koeffizienten – obwohl es sich um die Unbekannten des LGS handelt.

Abgesehen von dem kleinen Fallstrick bei der Bezeichnung “Koeffizient” sehen wir, dass wir LGS auf drei äquivalente Arten betrachten können:

- als System gekoppelter affin-linearer Gleichungen
- als Vektorgleichung mit Koeffizientenmatrix  $A$
- als Linearkombination der Inhomogenität  $\vec{y}$  aus den Spalten von  $A$  mit unbekannten Koeffizienten

Wird zu einer dieser drei Beschreibungen die Lösung für  $x_1, \dots, x_n$  gefunden, so ist sie damit auch für die anderen beiden gefunden. Insbesondere lassen sich die Koeffizienten von Linearkombinationen, die wir im Kapitel 6 bereits durch Manipulation der einzelnen Gleichungen berechnen konnten, hier systematisch ermitteln, indem das zugehörige LGS gelöst wird.

Wir fassen zusammen:

**Satz 8.2** (Äquivalente Beschreibungen von LGS). *Ein reelles LGS aus  $m$  Gleichungen mit  $n$  Veränderlichen  $x_1, \dots, x_n \in \mathbb{R}$ , Koeffizienten  $A_{j,k} \in \mathbb{R}$  für  $(j,k) \in \{1, \dots, m\} \times \{1, \dots, n\}$  und der Inhomogenität aus  $y_1, \dots, y_m \in \mathbb{R}$  ist äquivalent beschrieben durch die Gleichung*

$$A \cdot \vec{x} = \vec{y}$$

mit den Matrixelementen  $A_{j,k}$  wie oben, dem Urbildvektor  $\vec{x} := (x_1 \ \dots \ x_n)^T$  und dem Bildvektor  $\vec{y} := (y_1 \ \dots \ y_m)^T$ .

Betrachtet man die Spalten der Matrix  $A = (\vec{a}_1 \ \dots \ \vec{a}_n)$ , so ist die Linearkombination des Bildvektors aus den Spalten von  $A$  mit Koeffizienten  $x_1, \dots, x_n$  per

$$\sum_{k=1}^n x_k \vec{a}_k = \vec{y}$$

eine weitere äquivalente Beschreibung des LGS.

**Bemerkung:** Die Beschreibung als lineare Abbildung mit Koeffizientenmatrix  $A$  wird für uns von entscheidender Bedeutung beim Errechnen der Lösung eines LGS sein; dazu gleich mehr. Anders als im vorigen Kapitel, wo wir meist aus gegebenen Urbildern die Bildvektoren berechnet hatten, interessiert hier vielmehr, *welches Urbild (bzw. welche Urbilder) einer linearen Abbildung zu einem gegebenen Bild  $\vec{y}$  existieren.*

Hier stellen sich erneut die Fragen nach *Injektivität* und *Surjektivität*:

- Eine Abbildung ist z.B. nicht surjektiv, falls zur gegebenen Inhomogenität  $\vec{y}$  kein Urbild  $\vec{x}$  gefunden wird.
- Sie ist nicht injektiv, falls zu gegebenem  $\vec{y}$  mehr als ein Urbild gefunden wird.
- Sie ist genau dann bijektiv, wenn zu jedem gegebenen  $\vec{y}$  genau ein Urbild gefunden wird.

### 8.1.2 Lösung von LGS: Überlegungen

Wir hatten uns in Kapitel 1 bereits mit äquivalenten Umformungen von Gleichungen und Gleichungssystemen beschäftigt. Dabei hatten wir entdeckt, dass sich Gleichungen *addieren* lassen (indem man die linke Seite einer Gleichung zur linken Seite einer anderen addiert, dito für die rechten Seiten), und mit reellen Faktoren ungleich 0 *skaliert* werden können. Ist die linke Seite einer reellen Gleichung gleich der rechten Seite, so gilt dies auch für das  $c$ -fache von jeweils linker und rechter Seite.

---

Zur Übung versuche man sich an der Lösung der oben als Beispiele notierten LGSe mit diesen Operationen; solche Rechnungen haben wir bereits im vorigen Kapitel ausgeführt. Beispielsweise bietet es sich in (\*) an, die zweite Gleichung mit  $(-2)$  skaliert zur ersten, und mit  $(-3)$  skaliert zur dritten zu addieren; dadurch verschwindet dort jeweils die Unbekannte  $x_1$ . Mit der ersten und dritten Gleichung kann man analog verfahren, um in einer der beiden  $x_2$  zu eliminieren – das gibt die Lösung für  $x_3$ . Diese setzt man dann in die anderen beiden Gleichungen ein und erhält  $x_1, x_2$ . Für (\*) geht dies eindeutig.

Bei (\*\*) werden Sie fest stellen, dass zwei freie Variablen übrig bleiben; und (\*\*\*) ist sogar unlösbar.

---

Als besonders hilfreich erweist sich hier die Notation als Vektorgleichung (lineare Abbildung):

$$A\vec{x} = \vec{y}$$

Gesucht ist zu gegebenem  $A$  und  $\vec{y}$  das Urbild  $\vec{x}$ . Angenommen, es gibt genau eine Lösung  $\vec{x}$ , dann finden wir diese durch Multiplikation der Inversen von  $A$ :

$$\dots \Leftrightarrow A^{-1}A\vec{x} = A^{-1}\vec{y} \Leftrightarrow \vec{x} = A^{-1}\vec{y}$$

Bis hierher haben wir noch kein Verfahren, um die Inverse zu bestimmen, wenn sie nicht aufgrund besonderer Eigenschaften (orthogonale Matrix, diagonale Matrix) leicht erkennbar ist. Tatsächlich aber entspricht die Lösung des LGS, also das Finden des Urbilds von  $\vec{y}$ , gerade genau der Umkehrabbildung (falls dies eindeutig möglich ist) – das ist also implizit die Bestimmung der inversen Matrix!<sup>1</sup>

---

Nun ist leider nicht jedes LGS eindeutig lösbar, d.h. wir werden manchmal gar keine Lösung finden – oft aber auch *unendlich viele* – nämlich immer dann, wenn sich das LGS nicht bis zum Schluss in lineare Gleichungen auflösen lässt, die jeweils nur noch eine Unbekannte betreffen (und einfach lösbar sind; siehe Kapitel 1).

Für solche letzteren *mehrdeutig* lösbaren LGS lässt sich durch unsere oben beschriebenen Umformungen immerhin eine deutliche Vereinfachung erreichen, die darin besteht, möglichst viele *abhängige Variablen* zu finden. Die verbleibenden Variablen sind dann *freie Variablen* und deuten qua Name bereits an, dass sie beliebig belegt werden können. Jede solche Belegung führt dann, abhängig davon, zu konkreten Werten für die abhängigen Variablen, und damit zu einer einzelnen Lösung. Wir werden hierfür später noch geeignete Begriffe definieren, um dies strukturiert anzuschreiben.

---

Was aber trotzdem gilt, ist, dass die Umformungen des LGS so statt zu finden haben, dass in jedem Schritt stets ein logisch äquivalentes LGS entsteht. Das ist nur möglich mit *Äquivalenzumformungen*. Im Exkurs A.7 werden wir die drei Operationen (neben Skalierung und Addition ist auch das Vertauschen ganzer Gleichungen möglich, da die Und-Verknüpfung kommutativ ist) genauer analysieren und nachträglich rechtfertigen.

Da eine Äquivalenz beidseitig als Implikation verstanden werden kann, ist das Ziel also, das LGS *auf umkehrbar eindeutige Weise* zu manipulieren. Bezogen auf die vektorielle Notation (und deswegen ist diese hier besonders anschaulich) bedeutet das, eine *invertierbare* Matrix  $M \in \mathbb{R}^{(m,m)}$  zu finden, sodass

$$A\vec{x} = \vec{y} \Leftrightarrow MA\vec{x} = M\vec{y}$$

Da  $M$  invertierbar ist, kann dann in der rechten Gleichung jederzeit  $M^{-1}$  von links multipliziert werden, um die Manipulation rückgängig zu machen – dies führt dann eindeutig wieder zurück auf die ursprüngliche Vektorgleichung, also auf das äquivalente ursprüngliche LGS. Im Exkurs A.7 werden wir demonstrieren, dass die drei erwähnten Operationen durch elementare *und invertierbare* Matrizen ausgedrückt werden können – und nach Satz 7.18 (Eigenschaften inverser Matrizen) ist auch jedes Produkt solcher Matrizen wiederum invertierbar. Die fiktive Matrix  $M$  von oben kann also über ein (ggf. recht längliches) Produkt von elementaren invertierbaren Matrizen konstruiert werden, welches genau die Reihenfolge und Art von Operationen widerspiegelt, die wir an den Gleichungen des LGS vornehmen.

---

Für den Moment verlassen wir uns darauf, dass die erwähnten drei Operationen vollständig rückabwickelbar sind. Die rechte Gleichung von oben ist übrigens wieder ein LGS mit  $m$  Gleichungen für (die ursprünglichen)  $n$  Variablen, denn

$$MA\vec{x} = (MA)\vec{x} \quad \text{und} \quad M\vec{y} = (M\vec{y})$$

Die Matrix  $MA$  ist, da  $M \in \mathbb{R}^{(m,m)}$  und  $A \in \mathbb{R}^{(m,n)}$ , wieder aus  $\mathbb{R}^{(m,n)}$ , und der Vektor  $M\vec{y}$  ist weiterhin aus  $\mathbb{R}^m$ .

Wegen der Äquivalenz ist dann die Lösung  $\vec{x}$  stets auch eine Lösung der ursprünglichen Gleichung  $A\vec{x} = \vec{y}$ .

---

<sup>1</sup>Wir formalisieren dies speziell für Matrizen nochmal genauer in Kapitel 9.

### 8.1.3 Lösung von LGS: Vorgehen

Folgende Operationen für die Manipulation der Gleichungen im LGS stehen zur Verfügung:

- Skalieren einer Zeile (Gleichung)  $j$  des LGS mit konstantem reellen Faktor  $c \neq 0$ .  
Hierbei wird jeder Koeffizient  $A_{j,k}$  durch  $cA_{j,k}$  ersetzt (für alle  $k$  von 1 bis  $n$ ), und  $y_j$  ebenfalls durch  $cy_j$ .  
Wir dokumentieren dies, indem wir im *ursprünglichen* LGS (also in dem LGS vor der ausgeführten Umformung) an der betreffenden Gleichung rechts “ $c$ ” notieren.
- Addition einer (ggf. mit  $c \neq 0$  skalierten) Zeile  $r$  zu einer anderen Zeile  $s \neq r$ .  
Hier ersetzt man jeden Koeffizient  $A_{s,k}$  durch  $A_{s,k} + cA_{r,k}$ , und  $y_s$  durch  $y_s + cy_r$ .  
Wir dokumentieren dies, indem wir im ursprünglichen LGS rechts einen Pfeil von der Gleichung  $r$  zur Gleichung  $s$  zeichnen; bei Zeile  $r$  notieren wir noch den Skalierungsfaktor  $c$  (falls er nicht 1 ist). Die Pfeilrichtung gibt an, was wohin zu addieren ist.
- Vertauschen von Zeilen.  
Dies dokumentieren wir rechts durch einen Pfeil mit zwei Spitzen zwischen den betreffenden Zeilen.

Achtung:

Es lassen sich teilweise mehrere Operationen in einem Schritt kombinieren – aber *nur dann, wenn sie unabhängig voneinander durchführbar sind!*

Möglich ist z.B.:

- mehrere Gleichungen im selben Schritt skalieren
- (skalierte) Addition einer Gleichung zu verschiedenen anderen Gleichungen (je nach Bedarf mit unterschiedlichen Skalierungsfaktoren)
- mehrere Vertauschungen, solange jede der Gleichungen jeweils höchstens von einer Vertauschung betroffen ist – man achte aber darauf, dass die Pfeile gut auseinander zu halten sind; im Zweifel lieber mehrere Umformungsschritte verwenden

Nicht erlauben wollen wir z.B. einen Ringtausch, da die Dokumentation hierzu dann unübersichtlich wird. Ebenfalls verbieten wir das Addieren zu Gleichungen, die selbst wiederum getauscht oder anderswohin addiert werden.

Man beachte, dass für Prüfungsabgaben sämtliche Umformungsschritte in der o.g. Weise zu dokumentieren sind! Entsprechend empfiehlt es sich, die Dokumentation auch vorher schon stets mit anzuschreiben.

Wegen der vielen Rechenoperationen ist auch das Potential für Flüchtigkeitsfehler durchaus gegeben. Eine vollständige Dokumentation hilft dabei, solche Fehler nachträglich zu berichtigen.

Zur effizienteren Notation der LGS-Umformungen führen wir noch eine neue Schreibweise ein, die die Verbindung zwischen dem LGS und seiner vektoriellen Schreibweise deutlich macht. Uns fällt nämlich auf, dass wir beim Umformen der Gleichungen die Unbekannten  $x_k$  gar nicht antasten – die Manipulationen finden ausschließlich an den Koeffizienten  $A_{j,k}$  sowie an den Komponenten  $y_j$  der Inhomogenität statt. Da dies dort jeweils in der gleichen Weise geschieht, können wir die Koeffizientenmatrix und die Inhomogenität zu einer neuen Struktur zusammen fassen:

**Satz 8.3** (LGS-Darstellung mit (erweiterter) Koeffizientenmatrix). *Ein gegebenes reelles LGS mit  $m$  Gleichungen und  $n$  Unbekannten ist vollständig beschrieben durch die Matrix der Koeffizienten  $A \in \mathbb{R}^{(m,n)}$  sowie die Inhomogenität  $\vec{y}$  und lässt sich notieren als*

$$(A \mid \vec{y}) = \left( \begin{array}{cccc|c} A_{1,1} & A_{1,2} & \cdots & A_{1,n} & y_1 \\ A_{2,1} & A_{2,2} & \cdots & A_{2,n} & y_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ A_{m,1} & A_{m,2} & \cdots & A_{m,n} & y_m \end{array} \right)$$

Diese Struktur nennt man erweiterte Koeffizientenmatrix (oder: Gauß-Matrix). Falls das LGS homogen ist, wird der Trennstrich und die Inhomogenität nicht notiert.

### Bemerkungen:

- Die oben besprochene Dokumentation der Umformungsschritte geschieht analog rechts von der Gauß-Matrix an den jeweils relevanten Zeilen.
- Man beachte, dass die Gauß-Matrix keine echte Matrix ist, sondern eine Darstellung eines LGS. Daher hat solch eine Struktur einen Wahrheitswert. Genau wie bei LGS ist beim Umformen der Koeffizientenmatrizen daher das *Äquivalenzzeichen* zu notieren!

Anders herum: Wenn Matrizen *nicht* Gauß-Matrizen von LGSen sind, haben Äquivalenzzeichen bei ihnen nichts verloren; dann ist nur die Gleichheit von Matrizen relevant.

**Beispiel:** Die erweiterten Koeffizientenmatrizen der oben als Beispiel gegebenen LGSen lauten (und sind per “ $\Leftrightarrow$ ” jeweils äquivalent zu diesen):

$$(*) \Leftrightarrow \left( \begin{array}{ccc|c} 2 & 1 & 2 & 3 \\ 1 & 0 & 2 & 1 \\ 3 & 2 & 1 & 4 \end{array} \right); \quad (**) \Leftrightarrow \left( \begin{array}{cccc|c} 2 & 4 & 1 & 1 & 2 \\ 3 & 1 & 2 & 1 & 5 \\ 1 & -3 & 1 & 0 & 3 \end{array} \right); \quad (***) \Leftrightarrow \left( \begin{array}{ccc|c} 2 & 2 & 1 & 3 \\ 1 & -1 & 2 & 4 \\ 1 & 3 & -1 & 1 \end{array} \right)$$

---

Bevor wir nun (endlich!) zum konkreten Lösen der Beispiel-LGSen kommen, vereinbaren wir noch die *Lösungsstrategie*:

- Freie Variablen sollen, falls sie nötig sind, möglichst hohen Index besitzen. Wir gehen (für den Verlauf dieser Vorlesung) davon aus, dass die LGSen so formuliert sind, dass dies dem größeren Kontext, in welchem ein LGS zu lösen ist (also einer realen Problemsituation mit linearem Verhalten), nicht widerspricht. Zur Not lassen sich die Spalten der Koeffizientenmatrix im Voraus entsprechend vertauschen – davon sehen wir hier in der Vorlesung allerdings ab.
- Für die abhängigen Variablen (die also möglichst niedrigen Index besitzen), wollen wir die Gleichungen so umformen, dass sie mit Einträgen 1 in der Koeffizientenmatrix auftreten.
- Die Gleichungen sollten am Schluss möglichst knapp sein – d.h. die Koeffizientenmatrix sollte möglichst viele Nullen als Einträge besitzen.
- Ganze *Null-Zeilen* (also Zeilen, für die sowohl alle Koeffizienten als auch die jeweilige Inhomogenität 0 betragen) werden im jeweils nächsten Umformungsschritt kommentarlos ausgelassen.

Mit dieser Strategie wollen wir ein bestimmtes Aussehen des LGS erreichen:

**Definition 8.4** (Zeilen-Stufen-Form). *Die Gauß-Matrix eines LGS ist in Zeilen-Stufen-Form, falls in jeder Zeile der Koeffizient mit niedrigstem Spaltenindex, der nicht 0 ist, den Wert 1 hat. Unterhalb und/oder links davon befinden sich ausschließlich Nullen.*

**Bemerkung:** Verfolgt man die jeweils führenden Eins-Elemente der Zeilen, so ergibt sich eine stufenförmige Anordnung. Als Linie gedacht, befinden sich links unterhalb dieser Linie ausschließlich Nullen.

### 8.1.4 Schematisches Beispiel

Schematisch ist das Vorgehen in Abbildung 8.1 beschrieben: Für ein gegebenes LGS (hier mit sieben Variablen  $x_1, \dots, x_7$  und vier Gleichungen, inhomogen, zu sehen im oberen Teil der Abbildung) wird zunächst die Zeilen-Stufen-Form hergestellt.

Die oben erwähnten Eigenschaften sind im mittleren Teil der Abbildung erfüllt; die Stufenlinie ist zur Illustration mit eingezeichnet. Hier stellt sich heraus, dass das LGS eine Null-Zeile enthält, die noch entfernt werden kann.

Bis hierhin ist das Verfahren als *Gauß-Eliminationsverfahren* bekannt. Oberhalb der führenden Einsen in den Stufen lassen sich jedoch noch Nullen erzeugen – mindestens die im unteren Teil der Abbildung zusätzlich eingezeichneten, aber ggf. auch mehr. Danach liegt eine *reduzierte Zeilen-Stufen-Form* vor. Führt man diese nachträgliche Anpassung noch durch, so ist das Verfahren als *Gauß-Jordan-Verfahren* bekannt.

$$\begin{aligned}
& \left( \begin{array}{ccccccc|c} ? & ? & ? & ? & ? & ? & ? & ? \\ ? & ? & ? & ? & ? & ? & ? & ? \\ ? & ? & ? & ? & ? & ? & ? & ? \\ ? & ? & ? & ? & ? & ? & ? & ? \end{array} \right) \\
& \Leftrightarrow \dots \Leftrightarrow \left( \begin{array}{ccccccc|c} 1 & ? & ? & ? & ? & ? & ? & ? \\ 0 & 0 & 0 & 1 & ? & ? & ? & ? \\ 0 & 0 & 0 & 0 & 0 & 1 & ? & ? \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right) \\
& \Leftrightarrow \dots \Leftrightarrow \left( \begin{array}{ccccccc|c} 1 & a & b & 0 & c & 0 & d & h_1 \\ 0 & 0 & 0 & 1 & e & 0 & f & h_2 \\ 0 & 0 & 0 & 0 & 0 & 1 & g & h_3 \end{array} \right)
\end{aligned}$$

Abbildung 8.1: Stationen der LGS-Lösung mit Zeilen-Stufen-Form

Der untere Teil der Abbildung zeigt die besprochenen Anpassungen. Nun liegen klar erkennbar drei *abhängige* Variablen vor ( $x_1$ ,  $x_4$  und  $x_6$ ). Die Gleichungen mit den verbliebenen Zahlenwerten lauten:

$$\begin{array}{rcll}
\dots \Leftrightarrow & \wedge & x_1 + ax_2 + bx_3 & + cx_5 & + dx_7 & = & h_1 \\
& \wedge & & x_4 + ex_5 & + fx_7 & = & h_2 \\
& \wedge & & & x_6 + gx_7 & = & h_3
\end{array}$$

Die *freien* Variablen (in welchen die abhängigen auszudrücken sind) lauten  $x_2$ ,  $x_3$ ,  $x_5$  sowie  $x_7$ . Wir formulieren die Gleichungen entsprechend um:

$$\begin{array}{rcll}
\dots \Leftrightarrow & \wedge & x_1 & = & (-a)x_2 & + & (-b)x_3 & + & (-c)x_5 & + & (-d)x_7 & + & h_1 \\
& \wedge & x_4 & = & & & & & (-e)x_5 & + & (-f)x_7 & + & h_2 \\
& \wedge & x_6 & = & & & & & & & (-g)x_7 & + & h_3
\end{array}$$

Möchte man nun das Ergebnis wieder vollständig vektoriell anschreiben, so lassen sich die freien Variablen mit trivialen Gleichungen ergänzen:

$$\begin{array}{rcll}
& & x_1 & = & (-a)x_2 & + & (-b)x_3 & + & (-c)x_5 & + & (-d)x_7 & + & h_1 \\
& \wedge & x_2 & = & x_2 & & & & & & & & 0 \\
& \wedge & x_3 & = & & & x_3 & & & & & & 0 \\
\dots \Leftrightarrow & \wedge & x_4 & = & & & & & (-e)x_5 & + & (-f)x_7 & + & h_2 \\
& \wedge & x_5 & = & & & & & x_5 & & & & 0 \\
& \wedge & x_6 & = & & & & & & & (-g)x_7 & + & h_3 \\
& \wedge & x_7 & = & & & & & & & x_7 & + & 0
\end{array}$$

Vektoriell geschrieben:

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \end{pmatrix} = x_2 \begin{pmatrix} -a \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} -b \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + x_5 \begin{pmatrix} -c \\ 0 \\ 0 \\ -e \\ 1 \\ 0 \\ 0 \end{pmatrix} + x_7 \begin{pmatrix} -d \\ 0 \\ 0 \\ -f \\ 0 \\ -g \\ 1 \end{pmatrix} + \begin{pmatrix} h_1 \\ 0 \\ 0 \\ h_2 \\ 0 \\ h_3 \\ 0 \end{pmatrix}$$

Die vier freien Variablen können beliebig aus  $\mathbb{R}$  gewählt werden. Da die jeweiligen Vektoren, die von ihnen skaliert werden, linear unabhängig sind (um aus ihnen den Nullvektor  $\vec{0}_{(7)}$  linear zu kombinieren, müssen alle vier freien Variablen, bedingt durch die trivialen Gleichungen der Form  $x_j = x_j$ , auf 0 gesetzt werden), bilden diese eine Basis ihres Spanns – ein vierdimensionaler Unterraum von  $\mathbb{R}^7$ .

Ignoriert man den letzten Vektor (mit den Komponenten  $h_1$ ,  $h_2$  und  $h_3$ ), der nicht skaliert wird, so liegt die Lösung des zugehörigen homogenen LGS vor; der eben erwähnte Spann ist also gleichzeitig auch der *Kern* des LGS. Die allgemeine Lösung  $\vec{x}$  ist also die Summe aus einem beliebigen Vektor des Kerns,  $\vec{x}_h$  (einer *homogenen* Lösung) und einer eindeutigen *speziellen* Lösung  $\vec{x}_s$ . *Eindeutig lösbar* ist das LGS genau dann, wenn es keine freien Variablen gibt; dann ist  $\vec{x}_h = \vec{0}$  und  $\vec{x} = \vec{x}_s$ . Ein trivialer Kern enthält lediglich den Nullvektor und ist null-dimensional.

Wir definieren die hier erwähnten Begriffe später nochmals formal.

### 8.1.5 Beispiel: Eindeutig lösbares LGS

Wir lösen das LGS (\*) gemäß obiger Strategie. (Zur Übung führe man die selben Operationen gerne mit der gewöhnlichen LGS-Notation (Und-verknüpfte Gleichungen, wie in den Beispielen bei Definition 8.1) aus)

Um Nullen zu erzeugen, aber gleichzeitig Bruchrechnung zu vermeiden, kann man den kleinsten Eintrag einer Spalte suchen und dann die kleinsten gemeinsamen Vielfachen mit den anderen Einträgen der Spalte ermitteln. Dadurch ergeben sich dann passende Skalierungen für die Addition.

Wir erinnern uns, dass die erste Spalte der Koeffizientenmatrix der Variablen  $x_1$  vom LGS entspricht. Wir wollen also versuchen, dort eine 1 zu erzeugen. Wenn dies nicht in der ersten Zeile passiert, lässt sich das über Zeilenvertauschungen noch anpassen.

Unsere ursprüngliche Gaußmatrix lautet:

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

Wir bringen nun zunächst die zweite Zeile nach oben (durch Vertauschung mit der ersten). Danach addieren wir ihr  $(-2)$ -faches zur (neuen) zweiten Zeile, um dort in der ersten Spalte eine Null zu erzeugen. Analog verfahren wir danach mit der dritten Zeile. Aus Gründen der Übersichtlichkeit notieren wir dies in zwei Schritten, es wäre aber (und handschriftlich lässt sich die Dokumentation dazu auch leichter notieren) erlaubt, diese beiden Additionen in einem Schritt auszuführen.

Anschließend verfahren wir analog weiter, um in der zweiten Spalte noch eine Null zu erzeugen. Dann skalieren wir noch die dritte Zeile; dann ist eine Zeilen-Stufen-Form hergestellt.

$$\begin{aligned} & \left( \begin{array}{ccc|c} 2 & 1 & 2 & 3 \\ 1 & 0 & 2 & 1 \\ 3 & 2 & 1 & 4 \end{array} \right) \begin{array}{l} \leftarrow \\ \leftarrow \end{array} \Leftrightarrow \left( \begin{array}{ccc|c} 1 & 0 & 2 & 1 \\ 2 & 1 & 2 & 3 \\ 3 & 2 & 1 & 4 \end{array} \right) \begin{array}{l} \leftarrow -2 \\ \leftarrow \end{array} \Leftrightarrow \left( \begin{array}{ccc|c} 1 & 0 & 2 & 1 \\ 0 & 1 & -2 & 1 \\ 3 & 2 & 1 & 4 \end{array} \right) \begin{array}{l} \leftarrow \\ \leftarrow -3 \end{array} \\ & \Leftrightarrow \left( \begin{array}{ccc|c} 1 & 0 & 2 & 1 \\ 0 & 1 & -2 & 1 \\ 0 & 2 & -5 & 1 \end{array} \right) \begin{array}{l} \leftarrow \\ \leftarrow -2 \end{array} \Leftrightarrow \left( \begin{array}{ccc|c} 1 & 0 & 2 & 1 \\ 0 & 1 & -2 & 1 \\ 0 & 0 & -1 & -1 \end{array} \right) \begin{array}{l} \leftarrow \\ \leftarrow -1 \end{array} \Leftrightarrow \left( \begin{array}{ccc|c} 1 & 0 & 2 & 1 \\ 0 & 1 & -2 & 1 \\ 0 & 0 & 1 & 1 \end{array} \right) \begin{array}{l} \leftarrow \\ \leftarrow \end{array} \end{aligned}$$

Die Stufen laufen hier über die Diagonale der Koeffizientenmatrix – genau dann ist das LGS eindeutig lösbar, denn für jede Unbekannte kann nun sukzessive der Wert bestimmt werden. Die eine Möglichkeit ist, hier schon in ein gewöhnliches LGS zurück zu übersetzen und dann nacheinander durch Einsetzen die Lösung zu bestimmen.



Zwischenschritt (nur zum besseren Verständnis; würde man in der Praxis nicht anschreiben):

$$\dots \Leftrightarrow \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \Leftrightarrow \begin{array}{rcl} x_1 & +2x_3 & = 1 \\ x_2 & -2x_3 & = 1 \\ x_3 & & = 1 \end{array}$$

Das ergibt hier:

$$\begin{aligned} x_3 &= 1 \\ x_2 - 2x_3 &= x_2 - 2 = 1 \Leftrightarrow x_2 = 3 \\ x_1 + 2x_3 &= x_1 + 2 = 1 \Leftrightarrow x_1 = -1 \end{aligned}$$

Alternativ können wir (und für eindeutig lösbare LGS geht das immer) oberhalb der Stufenlinie noch Nullen erzeugen, um das Gauß-Jordan-Verfahren zu vervollständigen. Hier ist das nur in der dritten Spalte nötig. Wir setzen nochmals an (und kopieren die letzte Gauß-Matrix von oben):

$$\dots \Leftrightarrow \left( \begin{array}{ccc|c} 1 & 0 & 2 & 1 \\ 0 & 1 & -2 & 1 \\ 0 & 0 & 1 & 1 \end{array} \right) \begin{array}{l} \swarrow \\ \searrow \end{array} \begin{array}{l} -2 \\ 2 \end{array} \Leftrightarrow \left( \begin{array}{ccc|c} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 1 \end{array} \right)$$

Hier lesen wir die Lösung des LGS direkt ab:

$$x_1 = -1 \quad \wedge \quad x_2 = 3 \quad \wedge \quad x_3 = 1$$

Zur Probe rechnen wir mit Matrixmultiplikation nach, dass tatsächlich eine Lösung des LGS vorliegt (und wegen der Äquivalenzumformungen muss es auch die einzige sein):

$$\begin{pmatrix} 2 & 1 & 2 \\ 1 & 0 & 2 \\ 3 & 2 & 1 \end{pmatrix} \cdot \begin{pmatrix} -1 \\ 3 \\ 1 \end{pmatrix} = \begin{pmatrix} -2 + 3 + 2 \\ -1 + 0 + 2 \\ -3 + 6 + 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \\ 4 \end{pmatrix} \quad \checkmark$$

Getrennt notiert erhalten wir, wie in Satz 8.2 beschrieben, auch direkt die (hier eindeutige) *Linearkombination* der Inhomogenität aus den Spaltenvektoren der Matrix:

$$-\begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} + 3\begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} + \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \\ 4 \end{pmatrix}$$

### 8.1.6 Beispiel: Mehrdeutig lösbares LGS

Wir lösen das LGS (\*\*) von oben in Gauß-Notation. Zunächst stellen wir die Zeilen-Stufen-Form her:

$$\begin{aligned} &\left( \begin{array}{cccc|c} 2 & 4 & 1 & 1 & 2 \\ 3 & 1 & 2 & 1 & 5 \\ 1 & -3 & 1 & 0 & 3 \end{array} \right) \begin{array}{l} \swarrow \\ \searrow \end{array} \begin{array}{l} -2 \\ -3 \end{array} \Leftrightarrow \left( \begin{array}{cccc|c} 0 & 10 & -1 & 1 & -4 \\ 0 & 10 & -1 & 1 & -4 \\ 1 & -3 & 1 & 0 & 3 \end{array} \right) \begin{array}{l} \swarrow \\ \searrow \end{array} \begin{array}{l} -1 \\ -1 \end{array} \Leftrightarrow \left( \begin{array}{cccc|c} 0 & 0 & 0 & 0 & 0 \\ 0 & 10 & -1 & 1 & -4 \\ 1 & -3 & 1 & 0 & 3 \end{array} \right) \begin{array}{l} \swarrow \\ \searrow \end{array} \begin{array}{l} -1 \\ -1 \end{array} \\ &\Leftrightarrow \left( \begin{array}{cccc|c} 1 & -3 & 1 & 0 & 3 \\ 0 & 10 & -1 & 1 & -4 \end{array} \right) \frac{1}{10} \Leftrightarrow \left( \begin{array}{cccc|c} 1 & -3 & 1 & 0 & 3 \\ 0 & 1 & -\frac{1}{10} & \frac{1}{10} & -\frac{2}{5} \end{array} \right) \begin{array}{l} \swarrow \\ \searrow \end{array} \begin{array}{l} 3 \\ 3 \end{array} \Leftrightarrow \left( \begin{array}{cccc|c} 1 & 0 & \frac{7}{10} & \frac{3}{10} & \frac{9}{5} \\ 0 & 1 & -\frac{1}{10} & \frac{1}{10} & -\frac{2}{5} \end{array} \right) \end{aligned}$$

Damit haben wir die Zeilen-Stufen-Form erreicht. Weitere Nullen lassen sich oberhalb der Stufenlinie nur noch auf Kosten anderer bereits vorhandener Nullen erzeugen. Offenbar gibt es zwei abhängige Variablen (nach unserer Strategie  $x_1, x_2$ ) und zwei freie ( $x_3, x_4$ ).

Hier ist die Stufenlinie keine “gleichmäßige Treppe”, sondern wird nach der zweiten Zeile waagrecht – dies passt dazu, dass nur noch zwei Gleichungen für die vier Variablen verblieben sind – das reicht nicht, um die vier Variablen eindeutig zu bestimmen; das LGS ist also *unterbestimmt* und *mehrdeutig lösbar*.

Für die abhängigen Variablen gelten die Gleichungen, die sich aus der Gauß-Matrix ablesen lassen (man beachte dabei die Vorzeichen der Koeffizienten für die freien Variablen, die wir jeweils

auf die rechte Seite der Gleichungen bringen!). Damit weiterhin ein LGS für alle vier Veränderlichen vorliegt, ergänzen wir die beiden Gleichungen aus der Gauß-Matrix um die trivialen Gleichungen für die freien Variablen und erhalten insgesamt:

$$\begin{aligned} \dots \Leftrightarrow \quad & \begin{aligned} x_1 &= -\frac{7}{10}x_3 - \frac{3}{10}x_4 + \frac{9}{5} \\ \wedge \quad x_2 &= \frac{1}{10}x_3 - \frac{1}{10}x_4 - \frac{2}{5} \\ \wedge \quad x_3 &= x_3 \\ \wedge \quad x_4 &= x_4 \end{aligned} \end{aligned}$$

Nun können wir das gleiche auch noch vektoriell anschreiben:

$$\dots \Leftrightarrow \vec{x} = x_3 \begin{pmatrix} -\frac{7}{10} \\ \frac{1}{10} \\ 1 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} -\frac{3}{10} \\ -\frac{1}{10} \\ 0 \\ 1 \end{pmatrix} + \begin{pmatrix} \frac{9}{5} \\ -\frac{2}{5} \\ 0 \\ 0 \end{pmatrix} \quad \text{mit} \quad x_3, x_4 \in \mathbb{R}$$

Wir sehen, dass es *unendlich viele* Lösungen für das LGS (\*\*) gibt – zu jedem Paar  $(x_3, x_4)$  reeller Zahlen gibt es genau eine Lösung.

Falls  $x_3 = x_4 = 0$  gewählt wird, erhalten wir eine spezielle Lösung, nämlich

$$\vec{x} = \begin{pmatrix} \frac{9}{5} \\ -\frac{2}{5} \\ 0 \\ 0 \end{pmatrix}$$

Wir rechnen zur Probe nach, dass die drei Gleichungen erfüllt sind:

$$\frac{1}{5}(2 \cdot 9 + 4 \cdot (-2) + 0 + 0) = \frac{10}{5} = 2 \quad \checkmark$$

$$\frac{1}{5}(3 \cdot 9 - 2 + 0 + 0) = \frac{25}{5} = 5 \quad \checkmark$$

$$\frac{1}{5}(9 - 3 \cdot (-2) + 0 + 0) = \frac{15}{5} = 3 \quad \checkmark$$


---

Wir rechnen die Gleichungen auch noch für die beiden anderen Vektoren aus unserer allgemeinen Lösung  $\vec{x}$  nach. Zunächst der Vektor mit Skalierungsfaktor  $x_3$  (wir klammern  $\frac{1}{10}$  aus):

$$\frac{1}{10}(2 \cdot (-7) + 4 \cdot 1 + 10 + 0) = 0$$

$$\frac{1}{10}(3 \cdot (-7) + 1 \cdot 1 + 20 + 0) = 0$$

$$\frac{1}{10}(1 \cdot (-7) - 3 \cdot 1 + 10 + 0) = 0$$

Und nun analog für den Vektor mit Skalierungsfaktor  $x_4$ :

$$\frac{1}{10}(2 \cdot (-3) + 4 \cdot (-1) + 0 + 10) = 0$$

$$\frac{1}{10}(3 \cdot (-3) + 1 \cdot (-1) + 0 + 10) = 0$$

$$\frac{1}{10}(1 \cdot (-3) - 3 \cdot (-1) + 0 + 0) = 0$$

Diese beiden Vektoren (und damit auch sämtliche Linearkombinationen aus ihnen!) ergeben offensichtlich unter der linearen Abbildung genau den Nullvektor  $\vec{0}_{(3)}$ . Gemäß Definition 8.1 handelt es sich hier also um Lösungen des zugeordneten *homogenen* LGS von (\*\*) – also des LGS mit der gleichen Koeffizientenmatrix, das eine Inhomogenität von  $\vec{0}_{(3)}$  besitzt.

Der Kern des LGS (\*\*) ist also:

$$\text{span} \left( \begin{pmatrix} -\frac{7}{10} \\ \frac{1}{10} \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -\frac{3}{10} \\ -\frac{1}{10} \\ 0 \\ 1 \end{pmatrix} \right) = \left\{ x_3 \begin{pmatrix} -\frac{7}{10} \\ \frac{1}{10} \\ 1 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} -\frac{3}{10} \\ -\frac{1}{10} \\ 0 \\ 1 \end{pmatrix} \mid x_3, x_4 \in \mathbb{R} \right\}$$

Es handelt sich um einen zweidimensionalen Unterraum des  $\mathbb{R}^4$ , da die beiden Vektoren linear unabhängig sind (wenn wir die Zeilen-Stufen-Form nach der vereinbarten Strategie aufstellen, werden die Vektoren, die mit freien Variablen skaliert werden und daher den Kern des LGS ausmachen, stets linear unabhängig sein).

Zur Übung wähle man einige willkürliche reelle Werte für  $x_3, x_4$ , berechne die allgemeine Lösung  $\vec{x}$  und überprüfe mit dem LGS (entweder wie eben mit einzelnen Gleichungen, oder über die Matrixmultiplikation  $A\vec{x}$  wie im vorigen Beispiel), dass sich jeweils genau die Inhomogenität

$$\begin{pmatrix} 2 \\ 5 \\ 3 \end{pmatrix}$$

des LGS ergibt. Es empfehlen sich für  $x_3, x_4$  Vielfache von 10, um Bruchrechnungen zu vermeiden.

### 8.1.7 Beispiel: Unlösbares LGS

Wir lösen das LGS (\*\*\*) mit dem Gauß-Verfahren.

$$\left( \begin{array}{ccc|c} 2 & 2 & 1 & 3 \\ 1 & -1 & 2 & 4 \\ 1 & 3 & -1 & 1 \end{array} \right) \begin{array}{l} \leftarrow -2 \\ \leftarrow -1 \end{array} \Leftrightarrow \left( \begin{array}{ccc|c} 0 & 4 & -3 & -5 \\ 1 & -1 & 2 & 4 \\ 0 & 4 & -3 & -3 \end{array} \right) \leftarrow -1 \Leftrightarrow \left( \begin{array}{ccc|c} 0 & 4 & -3 & -5 \\ 1 & -1 & 2 & 4 \\ 0 & 0 & 0 & 2 \end{array} \right)$$

An dieser Stelle können wir die Rechnung bereits abbrechen, da ein unlösbares LGS haben: Die dritte Zeile führt nämlich auf die Gleichung  $0x_1 + 0x_2 + 0x_3 = 0 = 2$ , die stets unerfüllbar ist.

Wir erinnern uns, dass die Gauß-Matrix nur eine platz sparende Notation eines LGS ist; diese Gleichungen im LGS sind mit Konjunktionen ( $\wedge$ ) verknüpft. Nach äquivalenter Umformung haben wir eine Gleichung erhalten, die stets logisch falsch ist; das zieht wegen der Konjunktionen das gesamte LGS auf den Wahrheitswert  $\mathcal{F}$  herunter: Somit ist das LGS nicht lösbar, für keine Kombination von Unbekannten  $\vec{x} \in \mathbb{R}^3$ .

Wir können uns also auch ersparen, die Zeilen-Stufen-Form herzustellen.

### 8.1.8 Bestimmtheit und Lösbarkeit von LGS

Wir spezifizieren hier noch den Begriff der Bestimmtheit:

**Definition 8.5** (Bestimmtheit eines LGS). *Ein LGS mit mehr Gleichungen als Variablen (also mit einer Koeffizientenmatrix, die mehr Zeilen als Spalten besitzt) heißt überbestimmt.*

*Hat das LGS weniger Gleichungen als Variablen, so heißt es unterbestimmt.*

*Für eine quadratische Koeffizientenmatrix heißt das LGS eindeutig bestimmt.*

#### Bemerkungen:

- Wegen der Null-Zeilen, die während dem Lösen eines LGS entstehen können, kann sich die Bestimmtheit eines LGS ändern:
  - Aus einem eindeutig bestimmten LGS kann sich (durch die Reduktion der Anzahl von Gleichungen beim Streichen einer Null-Zeile) nachträglich noch ein unterbestimmtes LGS ergeben.
  - Aus einem überbestimmten LGS kann sich nachträglich noch ein eindeutig bestimmtes oder ein unterbestimmtes LGS ergeben.
- Man beachte, dass die eindeutige Bestimmtheit eines LGS *noch nichts* über die eindeutige Lösbarkeit aussagt. Wie im Beispiel mit dem LGS (\*\*) gesehen, kann ein eindeutig bestimmtes LGS sich als unlösbar heraus stellen; das gleiche kann auch für über- oder unterbestimmte LGS geschehen.

- Ist das LGS schon zu Anfang unterbestimmt, kann es jedoch niemals eindeutig lösbar sein, denn höchstens kann sich die Zahl der Gleichungen beim Lösen weiter reduzieren. Entweder es ist mehrdeutig lösbar, oder unlösbar.

---

Noch eine Beobachtung zu unlösbaren LGS: Die beim LGS (\*\*\*) beobachtete Situation gilt für sämtliche unlösbaren LGS: Beim Umformen ergibt sich irgendwann eine Gleichung der Art  $0 = c$  mit  $c \neq 0$ , die nicht erfüllbar ist. Dies kann nur dann überhaupt passieren, wenn, die Inhomogenität des LGS nicht verschwindet. Denn die rechte Seite der Gleichungen in  $(A | \vec{y})$  werden (nur!) über die Zeilenoperationen der Inhomogenität  $\vec{y}$  berechnet. Beträgt diese schon zu Anfang  $\vec{0}$ , bleibt dies auch nach beliebigen Zeilenoperationen so.

Für homogene LGS besteht also nur die Wahl zwischen mehrdeutiger und eindeutiger Lösbarkeit. Und eine eindeutige Lösbarkeit bedeutet, dass der Nullvektor des Bildraums genau ein Urbild besitzt – nämlich den Nullvektor des Urbildraums (denn dieses Urbild besitzt er ohnehin stets)

Wir halten fest:

**Satz 8.6** (Lösbarkeit homogener LGS). *Homogene LGS sind niemals unlösbar. Ist ein homogenes LGS  $(A | \vec{0})$  mit  $A \in \mathbb{R}^{(m,n)}$  eindeutig lösbar, so ist die Lösung  $\vec{0}_{(n)}$ , das triviale Urbild von  $\vec{0}_{(m)}$ .*

**Bemerkung:** Inhomogene LGS können sich dagegen unabhängig von ihrer anfänglichen Bestimmtheit als unlösbar heraus stellen.

### 8.1.9 Kern einer Matrix/Linearen Abbildung

Wir formalisieren unsere Beobachtungen vom zweiten obigen Beispiel mit unterbestimmtem LGS.

**Definition 8.7** (Homogene und Spezielle Lösung, Kern). *Jede Lösung  $\vec{x} \in \mathbb{R}^n$  eines LGS  $(A | \vec{y})$  mit  $A \in \mathbb{R}^{(m,n)}$  hat die Struktur*

$$\vec{x} = \vec{x}_h + \vec{x}_s \quad \text{mit} \quad A\vec{x}_h = \vec{0}$$

*Dabei heißt der Teil  $\vec{x}_h$  homogene Lösung, der Teil  $\vec{x}_s$  spezielle Lösung des LGS.*

*Die Menge aller homogenen Lösungen, also die Lösungen von  $(A | \vec{0})$ , heißt Kern der linearen Abbildung (oder: Kern der Matrix  $A$ ), geschrieben*

$$\ker A := \left\{ \vec{x} \in \mathbb{R}^n \mid A\vec{x} = \vec{0} \right\}$$

*Der Nullvektor  $\vec{0} \in \mathbb{R}^n$  ist stets Teil von  $\ker A$ .*

*Enthält der Kern nur den Nullvektor, so ist er nulldimensional und heißt trivialer Kern.*

**Bemerkungen:**

- Falls ein lösbares LGS nur den trivialen Kern besitzt, gibt es in der Lösung keine freien Variablen; dann hat die Lösung die Struktur:

$$\vec{x} = \vec{x}_h + \vec{x}_s = \vec{0} + \vec{x}_s$$

Es gibt nur ein Element in  $\ker A$ , nämlich den Nullvektor; das LGS  $(A | \vec{y})$  ist dann eindeutig lösbar.

Übrigens gibt es dann auch nur genau eine spezielle Lösung, nämlich  $\vec{0}$ , die auf den Nullvektor aus  $\mathbb{R}^m$  abbildet. Nach Satz 8.2 bedeutet dies, dass die Spalten der Koeffizientenmatrix linear unabhängig sind!

- Der Kern der Matrix  $A$  ist ein Unterraum von  $\mathbb{R}^n$ . Wir hatten oben schon gesehen, dass wir beim Umformen des LGS auf Zeilen-Stufen-Form eine gewisse Anzahl freier Variablen erhalten (null bei eindeutig lösbarem LGS). Diese Zahl entspricht genau der Dimension des Kerns, geschrieben

$$\dim \ker A$$

### Beispiele:

- Unser LGS (\*) von oben war eindeutig lösbar und hatte  $\dim \ker A = 0$ , also einen trivialen Kern.
- Das LGS (\*\*) dagegen hatte zwei freie Variablen und damit  $\dim \ker A = 2$ .

### 8.1.10 Rang einer Matrix/Linearen Abbildung

Die neben dem Kern einer linearen Abbildung andere wichtige Größe ist die Zahl der abhängigen Variablen nach Umformung des LGS auf Zeilen-Stufen-Form. Diese Anzahl ist genau wie  $\ker A$  unabhängig von der konkreten Inhomogenität und also eine Eigenschaft der Koeffizientenmatrix  $A$ :

**Definition 8.8** (Rang). *Der Rang eines LGS bzw. seiner Koeffizientenmatrix bzw. der korrespondierenden linearen Abbildung ist die Zahl der übrig gebliebenen Zeilen eines LGS, nachdem dessen Koeffizientenmatrix  $A$  auf Zeilen-Stufen-Form gebracht wurde (Null-Zeilen werden nicht mitgezählt), geschrieben*

$$\operatorname{rg} A$$

### Beispiele:

- Unser LGS (\*) von oben hatte  $\operatorname{rg} A = 3$ , da sich die Zeilen-Stufen-Form der Matrix über drei Zeilen erstreckt.
- Das LGS (\*\*) hatte dagegen  $\operatorname{rg} A = 2$ , da von den drei Zeilen eine beim Umformen zu einer Nullzeile wurde. Es gab zwei abhängige Variablen.

---

Es gilt (ohne Beweis) folgender

**Satz 8.9** (Rangsatz). *Sei  $A \in \mathbb{R}^{(m,n)}$  die Koeffizientenmatrix einer linearen Abbildung von  $\mathbb{R}^n$  nach  $\mathbb{R}^m$ . Dann gilt:*

$$n = \operatorname{rg} A + \dim \ker A$$

### Bemerkungen:

- Wir hatten schon  $\dim \ker A$  als die Zahl der Variablen ausgemacht, die nach der Umformung eines LGS auf Zeilen-Stufen-Form frei ist. Zu diesen Variablen gibt es jeweils linear unabhängige Vektoren, deren Spann genau den Kern  $\ker A$  bildet.

Die anderen  $n - (\dim \ker A)$  Variablen waren abhängig – ihre Zahl entsprach genau der Zahl der (nicht-Null-)Zeilen in der Zeilen-Stufen-Form; dies ist der Rang. Insgesamt gibt es  $n$  Variablen, um Vektoren im Urbildraum zu beschreiben.

- Wir hatten oben schon gesehen, dass bei trivialem Kern die Spalten von  $A$  linear unabhängig sind. Die Zeilen sind es allerdings auch – denn sonst wäre es möglich, durch nichttriviale Linearkombination der Zeilen (und genau dies ist es, was wir mit den Zeilenoperationen berechnen) mindestens eine Null-Zeile zu erzeugen.

Versucht man, so lange Null-Zeilen zu erzeugen wie möglich, so sind die danach verbleibenden Zeilen der Koeffizientenmatrix linear unabhängig. Der Rang der Matrix  $A$  entspricht also genau der Zahl linear unabhängiger Zeilen von  $A$ .

Falls wir nun eine Zeilen-Stufen-Form einer Matrix hergestellt haben, könnten wir auch (in Gedanken) mit Spaltenoperationen weiter Nullen erzeugen – und zwar für jede der  $(\operatorname{rg} A)$  Zeilen sämtliche bis auf die führende 1. Der Rang von  $A$  entspricht also auch genau der Zahl linear unabhängiger Spalten von  $A$ .

### Beispiele:

- Das LGS (\*) hatte Rang 3 und trivialen Kern; die Summe aus  $\text{rg } A$  und  $\dim \ker A$  beträgt also  $n = 3$ .
- Das LGS (\*\*) hatte Rang 2 und einen zweidimensionalen Kern; die Summe aus  $\text{rg } A$  und  $\dim \ker A$  beträgt also  $n = 4$ .

---

Die Aussage der zweiten obigen Bemerkung halten wir noch fest als

**Satz 8.10** (Rang einer Matrix). *Für  $A \in \mathbb{R}^{(m,n)}$  ist  $\text{rg } A$  genau die Anzahl der linear unabhängigen Zeilen von  $A$  und die Anzahl der linear unabhängigen Spalten von  $A$ . Es gilt:*

$$\text{rg } A \leq \min\{m, n\}$$

*Ein LGS  $(A \mid \vec{y})$  ist genau dann eindeutig lösbar, wenn  $\text{rg } A = n$ .*

**Bemerkung:** Wenn die Matrix nicht quadratisch ist, kann ihr Rang dann nur höchstens so groß wie die kleinere der Dimensionszahlen sein. Ist der Rang tatsächlich gleich dieser Zahl, so spricht man von *vollem Rang*. War das LGS aber schon zu Anfang unterbestimmt, so reicht auch voller Rang nicht für eindeutige Lösbarkeit aus.

## 8.2 Permutationen Revisited

### 8.2.1 Recap

Bevor wir uns dem wichtigen Thema der *Determinanten quadratischer Matrizen* zuwenden, müssen wir uns nochmal mit den Permutationen befassen, die in Kapitel 4 bereits eingeführt wurden. Als wichtigste Erinnerung:

- Es handelt sich um bijektive Abbildungen von  $M_n := \{1, 2, \dots, n\}$  auf  $M_n$ .
- Zwei gleichwertige Notationen sind die Tabellenschreibweise sowie die Zykelschreibweise.
- Permutationen vertauschen im allgemeinen *nicht*; dies gilt jedoch für Zyklen, falls diese sich auf disjunkte Teilmengen von  $M_n$  auswirken.
- Zu jeder Permutation gibt es eine Zerlegung in kommutierende “kanonische” Zyklen, die sich aus der Tabellenschreibweise schnell ableiten lassen. Bis auf die Reihenfolge der Zyklen und jeweils die Einstiegspunkte (die auf jedem der Zyklus-Elemente liegen dürfen) sind diese Zyklen eindeutig.
- Eine Transposition ist ein Zyklus der Länge 2; sie vertauscht genau zwei Elemente aus  $M_n$ .
- Die Permutationen auf  $M_n$  bilden mit der Hintereinanderausführung eine (nicht abelsche) Gruppe, die symmetrische Gruppe  $S_n$ .

### 8.2.2 Zyklen und Transpositionen

Wir haben im Kapitel 7 bereits die Permutationsmatrizen kennen gelernt, die für jede Permutation  $\sigma \in S_n$  als  $P_\sigma$  eindeutig definiert sind. So entspricht auch jeder Permutationsmatrix aus  $\mathbb{R}^{(n,n)}$  genau eine Permutation  $\sigma \in S_n$ .

Die Frage ist nun, wie man eine Permutationsmatrix auf systematische Art aus der Einheitsmatrix  $\mathbb{1}_n$  erzeugen kann, wenn die Permutation  $\sigma \in S_n$  gegeben ist. Anschaulich ist aber klar, dass dies gehen muss, da die kartesischen Einheitsvektoren der Standardbasis  $E_n$  nur richtig als Spalten von  $P_\sigma$  anzuordnen sind.

Dies geschieht durch Vertauschungen. Zum Beispiel soll  $\vec{e}_{\sigma(1)}$  die erste Spalte von  $P_\sigma$  bilden. Geht man von  $1_n$  aus, so befindet sich dieser Vektor an der Spaltenposition  $\sigma(1)$ . Falls dies nicht die erste Spalte ist, können wir die beiden Spalten tauschen; danach ist die erste Spalte der Permutationsmatrix hergestellt.

Allerdings liegt danach in den Spalten 2 bis  $n$  keine Einheitsmatrix mehr vor; es kann nun sein, dass der richtige Vektor für die zweite Spalte,  $\vec{e}_{\sigma(2)}$ , an einer Position ungleich  $\sigma(2)$  liegt. In kleinen überschaubaren Beispielen findet man den passenden Vektor natürlich trotzdem schnell und kann wieder tauschen. Mit insgesamt höchstens  $(n - 1)$  solchen Tauschvorgängen ist dann die Matrix  $P_\sigma$  hergestellt (ein  $n$ -ter Schritt ist nicht nötig, da dann schon alle anderen Spaltenvektoren richtig sind – also muss dies auch für die Spalte  $n$  gelten).

Wir geben zwar nachher ein Beispiel dafür an, weisen aber darauf hin, dass die Strategie “den richtigen Vektor heraus suchen” für eine weitere Analyse nicht systematisch genug ist.

Satz 7.26 zeigt aber, wie sich die Spalten von Matrizen permutieren lassen; das obige Vorgehen ließe sich durch die sukzessive Multiplikation (von rechts) mit den Permutationsmatrizen von Transpositionen realisieren.

Eine andere Möglichkeit ist, die *kanonischen Zyklen* von  $\sigma$  zu verwenden. Wir motivieren nun, wie sich diese Zyklen systematisch in Verkettungen von Transpositionen auswerten lassen.

**Beispiel:** Wir betrachten den Zyklus

$$\zeta := (1 \quad 2 \quad 4 \quad 5 \quad 3)$$

Wir möchten also folgendes Abbildungsverhalten erreichen:

$$1 \mapsto 2; \quad 2 \mapsto 4; \quad 4 \mapsto 5; \quad 5 \mapsto 3; \quad 3 \mapsto 1$$

Starten wir mit der ersten Zuordnung, so bietet es sich an, mit der Transposition

$$\tau_{1,2} = (1 \quad 2)$$

zu starten. Damit wird die 1 schon einmal richtig abgebildet. Allerdings wird die 2 nicht wie gewünscht auf 4, sondern auf 1 abgebildet – dies müssen wir nun kompensieren! Man erreicht dies, indem man direkt danach eine weitere Transposition einführt, die genau dies bewirkt. Diese muss von links an  $\tau_{1,2}$  komponiert werden, und wir erhalten:

$$\tau_{1,4} \circ \tau_{1,2} = (1 \quad 4)(1 \quad 2)$$

So wird die 1 richtig auf 2 abgebildet, und die 2 richtig auf 4. Genau wie vorhin ist nun das Problem, dass die 4 nicht richtig abgebildet wird. Die rechte Transposition wirkt auf die 4 gar nicht, aber die linke bildet sie auf 1 ab. Das kompensieren wir erneut:

$$\tau_{1,5} \circ \tau_{1,4} \circ \tau_{1,2} = (1 \quad 5)(1 \quad 4)(1 \quad 2)$$

Nun wird die 5 jedoch auf 1 abgebildet; also ergänzen wir noch:

$$\zeta = \tau_{1,3} \circ \tau_{1,5} \circ \tau_{1,4} \circ \tau_{1,2} = (1 \quad 3)(1 \quad 5)(1 \quad 4)(1 \quad 2)$$

Jetzt sind wir allerdings fertig, denn jetzt bildet 5 zunächst auf 1 ab, und danach (über die linke äußere Transposition, die als letztes ausgeführt wird) auf 3, wie gewünscht. Und 3 soll ja (s.o.) auf 1 abbilden; es ist also alles erreicht.

Sucht man nach dem Abbildungsverhalten  $\zeta(j)$  für eine der Zahlen  $j$ , so startet man rechts (innen). Falls  $j$  die 1 ist, wird direkt in der ersten (rechten) Permutation richtig abgebildet. Die weiteren Transpositionen enthalten den Wert 2 alle nicht und beeinflussen ihn also nicht mehr weiter.

Für alle anderen Zahlen  $j$  gilt, dass von rechts zunächst eine gewisse Anzahl von unbeteiligten Transpositionen durchlaufen werden (null für 2, für andere  $j$  mindestens eine) – und zwar, bis die Transposition  $\tau_{1,j} = (1 \quad j)$  erreicht ist. Dann wird  $j$  auf 1 abgebildet. Falls dies nicht das richtige Abbildungsverhalten war, wird dies *direkt in der nächsten* Transposition korrigiert. Weiter links davon liegende Transpositionen betreffen dann diese beiden Werte ( $j$  und  $\zeta(j)$ ) nicht weiter.

Von dem konkreten Beispiel können wir mit gleicher Logik abstrahieren:

**Satz 8.11** (Dekomposition eines Zyklus in Transpositionen). *Jeder Zyklus  $\zeta \in S_n$  mit Länge  $k := |\zeta|$  lässt sich als Komposition von  $k - 1$  Transpositionen darstellen per*

$$(a_1 \quad a_2 \quad \cdots \quad a_k) = (a_1 \quad a_k)(a_1 \quad a_{k-1}) \cdots (a_1 \quad a_2)$$

### Bemerkungen:

- Jedes Element aus  $\zeta$  ist von maximal zwei dieser Transpositionen betroffen. Die Abbildung

$$a_j \mapsto a_{j+1}$$

wird durch die Kette  $a_j \mapsto a_1 \mapsto a_{j+1}$  hergestellt, also durch die Transpositionen

$$(a_1 \ a_{j+1})(a_1 \ a_j)$$

ausgeführt. Sowohl  $a_{j+1}$  als auch  $a_j$  kommen für  $j > 1$  nur genau in diesem Paar vor und werden weder von den Transpositionen rechts davon (also vorher) noch links davon (also nachher) betroffen.

- Zyklen der Länge 2 sind bereits Transpositionen und müssen daher nicht aufgespalten werden. Zyklen der Länge 1 sind trivial und erfordern gar keine Transpositionen, da hier nur ein Element von  $M_n$  auf sich selbst abgebildet wird – im Sinne der Vertauschungen ist also nichts zu tun.

### Beispiele:

- Wir spalten die folgende Permutation aus  $S_7$  in Transpositionen auf:

$$\sigma := (3 \ 7 \ 1 \ 4)(2 \ 6 \ 1 \ 3 \ 5)$$

Man beachte, dass dies keine Aufspaltung in kanonische Zyklen ist, da die Werte 1 und 3 in beiden Teilzyklen vorkommen! Hier ist also die Reihenfolge der beiden Zyklen nicht beliebig.

Wir ermitteln mit dem obigen Satz die Transpositionen

$$\sigma = \underbrace{(3 \ 4)(3 \ 1)(3 \ 7)}_{\text{linker Zyklus}} \underbrace{(2 \ 5)(2 \ 3)(2 \ 1)(2 \ 6)}_{\text{rechter Zyklus}}$$

- Nun wollen wir auch noch die kanonischen Zyklen von  $\sigma$  ermitteln und verschaffen uns dazu zunächst die Wertetabelle. Hierzu sind die ursprünglichen (nicht kommutierenden) beiden Zyklen übersichtlicher. Bei den Abbildungen mit zwei Schritten wird zunächst auf einen der beiden Werte abgebildet, die in beiden Zyklen vorkommen, sodass im linken Zyklus nachträglich noch Änderung bewirkt wird:

$$1 \mapsto 3 \mapsto 7; \quad 2 \mapsto 6; \quad 3 \mapsto 5; \quad 4 \mapsto 3; \quad 5 \mapsto 2; \quad 6 \mapsto 1 \mapsto 4; \quad 7 \mapsto 1$$

Also erhalten wir (die kanonischen Zyklen ermitteln wir direkt aus der Tabellenschreibweise):

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 7 & 6 & 5 & 3 & 2 & 4 & 1 \end{pmatrix} = (1 \ 7)(2 \ 6 \ 4 \ 3 \ 5)$$

Zerlegen wir nun die kanonischen Zyklen in Transpositionen, so fällt weniger Arbeit an (da diese Zyklen miteinander kommutieren und daher keine Zwischenwerte erreicht werden, die sich später noch ändern). Wir erhalten:

$$\sigma = (1 \ 7)(2 \ 5)(2 \ 3)(2 \ 4)(2 \ 6)$$

Man vergewissere sich zur Übung, dass mit diesen Transpositionen das gleiche Verhalten erzielt wird wie mit obiger Aufspaltung.

- Wir wollen hier noch das eingangs erwähnte (und dann wieder verworfene) Vorgehen zeigen; dazu notieren wir untereinander die Elemente geordnet als Tupel (oder Liste – *nicht zu verwechseln mit Permutationszyklen*) und lesen dann die Tauschvorgänge ab. Wir rücken von links ausgehend die nach der Wertetabelle gegebenen Elemente an ihre richtigen Positionen – dabei bedenken wir aber, dass die Vertauschungen sich nicht auf Listenpositionen, sondern auf die Werte der Einträge beziehen. Unterstrichene Elemente sind jeweils für den Tausch markiert und im nächsten Schritt überstrichen notiert.

$$(\underline{1}, 2, 3, 4, 5, 6, \underline{7})$$

$$(\overline{7}, \underline{2}, 3, 4, 5, 6, \overline{1})$$

$$(7, \overline{6}, \underline{3}, 4, \underline{5}, \overline{2}, 1)$$

$$(7, 6, \overline{5}, \underline{4}, \underline{3}, \overline{2}, 1)$$

$$(7, 6, 5, \overline{3}, \underline{4}, \underline{2}, 1)$$

$$(7, 6, 5, 3, \overline{2}, \overline{4}, 1)$$



Hier waren nur fünf Vertauschungen nötig, da das Element 1 direkt nach dem ersten Tausch schon an richtiger Stelle liegt. Insgesamt erhalten wir als Permutation:

$$\sigma = \tau_{2,4} \circ \tau_{3,4} \circ \tau_{5,3} \circ \tau_{6,2} \circ \tau_{7,1} = (2\ 4)(3\ 4)(5\ 3)(6\ 2)(7\ 1)$$

Zwar hatten wir auch in der kanonischen Zerlegung fünf Transpositionen erhalten, allerdings dort auf strukturierte Art. Hier sieht das Abbildungsverhalten etwas anders aus:

$$1 \mapsto 7; \quad 2 \mapsto 6; \quad 3 \mapsto 5; \quad 4 \mapsto 3; \quad 5 \mapsto 3 \mapsto 4 \mapsto 2; \quad 6 \mapsto 2 \mapsto 4; \quad 7 \mapsto 1$$

Wir beobachten, dass die Abbildung der 5 zwei Zwischenschritte erfordert. Die Abbildung der 6 erfordert zwar nur einen Zwischenschritt; dafür erfolgt die Kompensation des Zwischenwerts aber erst nach dem Überspringen von zwei dazwischen liegenden unbeteiligten Transpositionen.

Wir geben dieses letzte Beispiel nur deswegen an, weil es, wie man sieht, auch funktioniert – wir haben die gleiche Permutation  $\sigma$  aus  $S_7$  nun auf diverse verschiedene Weisen notiert, die alle richtig sind – zielführend für unser weiteres Vorgehen ist aber die Zerlegung in kanonische Zyklen.

---

Man achte darauf, komponierte Permutationen stets von innen nach außen (also von rechts nach links) auszuwerten.

---

Ist eine Permutation vollständig in Transpositionen aufgespalten, so lässt sich ihr Inverses direkt angeben – es entspricht gerade der umgekehrten Abfolge dieser Transpositionen (die als Involutionen jeweils alle selbstinvers sind). Siehe dazu auch die Ausführungen in Abschnitt 4.3.

### 8.2.3 Vorzeichen einer Permutation

Hier behandeln wir die wesentliche Vorbereitung für das Thema Determinanten. Denn die Zahl von Vertauschungen, die eine in Transpositionen aufgespaltene Permutation  $\sigma \in S_n$  besitzt ist stets entweder gerade oder ungerade.

Bevor wir das formal fest halten (und im Beweis begründen), ein

**Beispiel:** Wir betrachten in  $S_9$  die Permutation

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

Zunächst verschaffen wir uns die kanonischen Zyklen durch Auswerten der Tabellenschreibweise; wir erhalten (siehe auch Abbildung 8.2):

$$\sigma = (1\ 4\ 5)(2\ 6\ 3\ 8)(7\ 9)$$

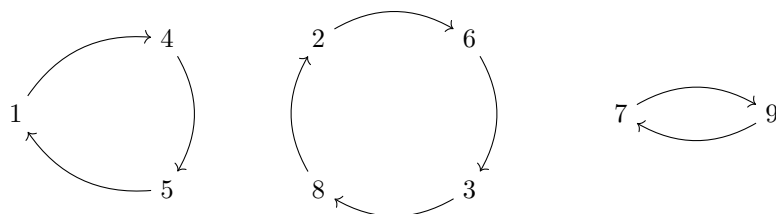


Abbildung 8.2: Graph der Beispielpermutation  $\sigma \in S_9$

Wir spalten die kanonischen Zyklen in Transpositionen auf; danach ermitteln wir deren Gesamtzahl  $t$ :

$$\sigma = \underbrace{(1\ 5)(1\ 4)}_{(1\ 4\ 5)} \underbrace{(2\ 8)(2\ 3)(2\ 6)}_{(2\ 6\ 3\ 8)} (7\ 9)$$

Zum Addieren der Transpositions-Anzahlen (das mag für dieses kleine Beispiel umständlich wirken, soll aber die Systematik verdeutlichen) schreiben wir für jeden Zyklus einzeln die Aussage von Satz 8.11 an:

$$t = (3 - 1) + (4 - 1) + (2 - 1) = (3 + 4 + 2) - (1 + 1 + 1) = 9 - 3 = 6$$

Unsere Beispielpermutation hat  $r = 3$  kanonische Zyklen (deren Anzahl entspricht der Zahl der Zyklen im Funktionsgraph; vgl. Abschnitt 4.3). Es addieren sich insgesamt die Längen der  $r$  Zyklen; außerdem wird für jeden Zyklus jeweils 1 subtrahiert, da für einen Zyklus der Länge  $k$  genau  $(k - 1)$  Transpositionen benötigt werden.

Allgemein lässt sich die Zahl der Permutationen, da die Summe der Längen der kanonischen Zyklen genau der Elementzahl der betrachteten Menge entspricht, auch schreiben als

$$t = n - r$$

Wir beachten, dass die Zyklenzahl stets  $r$  zwischen 1 und  $n$  liegt – das sind die beiden Extreme, falls die gesamte Permutation ein großer Zyklus ist, bzw. falls es sich um die identische Permutation handelt, bei der jedes Element auf sich selbst abgebildet wird und dadurch einen (trivialen) Zyklus der Länge 1 beiträgt.

Nun halten wir unsere oben bereits formulierte Behauptung fest. Der Beweis im Anhang ist länglich, aber nicht allzu schwer nachzuvollziehen – da es überhaupt nicht offensichtlich ist, warum die Aussage stimmen muss (wir hatten oben schon gesehen, dass es viele verschiedene Weisen gibt, die gleiche Permutation in Transpositionen auszudrücken), wird dringend empfohlen, ihn zumindest zu überfliegen.

**Satz 8.12** (Anzahl Transpositionen einer zerlegten Permutation). *Für eine beliebige Permutation  $\sigma \in S_n$  mit  $r$  kanonischen Zyklen besitzt jede vollständige Zerlegung von  $\sigma$  in  $t$  Transpositionen die Eigenschaft*

$$t \equiv (n - r) \pmod{2}$$

(Beweis: S. 334.)

### Bemerkungen:

- Die Zahl  $t$  ist also für eine gegebene Permutation stets gerade oder stets ungerade – egal auf welche Weise  $\sigma$  in Transpositionen zerlegt wird.
- Die Restklasse von  $t$  ist wohldefiniert, da  $n$  und  $r$  für jede Permutation  $\sigma \in S_n$  eindeutig bestimmt sind.

### Beispiele:

- In obiger Beispielpermutation (siehe Abbildung 8.2) hatten wir  $t = 6$  Transpositionen aus den kanonischen Zyklen ermittelt. Eine weitere Zerlegung von  $\sigma$  in (nicht-kanonische!) Zyklen wäre z.B.

$$\sigma = (1 \ 4 \ 3 \ 5)(2 \ 6 \ 4 \ 3 \ 8)(7 \ 9)$$

Die Dekomposition dieser Darstellung nach Satz 8.11 lautet dann:

$$\sigma = (1 \ 5)(1 \ 3)(1 \ 4)(2 \ 8)(2 \ 3)(2 \ 4)(2 \ 6)(7 \ 9)$$

Dies sind nicht 6, sondern 8 Transpositionen. Man vergewissere sich zur Übung, dass die obige Zerlegung in Zyklen richtig ist, also auf die gleiche Tabellenschreibweise wie  $\sigma$  führt.

- Wir hatten oben in den Beispielen zu Satz 8.11 Zerlegungen einer Permutation aus  $S_7$  ermittelt, darunter eine mit sieben Transpositionen und zwei mit je fünf Transpositionen – beide Anzahlen sind ungerade.

Nun wissen wir, dass die Zahl von Transpositionen einer vollständig zerlegten Permutation stets entweder gerade oder ungerade ist, und können die ausstehenden Begriffe einführen:

**Definition 8.13** (Vorzeichen einer Permutation). Für  $n \in \mathbb{N}$  heißt die Funktion

$$\text{sign} : S_n \rightarrow \{-1, +1\}$$

Vorzeichen (oder: Signum), falls sie die folgenden Eigenschaften erfüllt:

1.  $\text{sign}(\text{id}_n) = 1$  (Normiertheit)
2. Für eine beliebige Transposition  $\tau \in S_n$  gilt für alle  $\sigma \in S_n$ , dass die Komposition mit  $\tau$  das Vorzeichen umkehrt:

$$\text{sign}(\tau \circ \sigma) = \text{sign}(\sigma \circ \tau) = -\text{sign}(\sigma)$$

Eine Permutation  $\sigma \in S_n$  heißt gerade, falls  $\text{sign}(\sigma) = 1$ , und ungerade, falls  $\text{sign}(\sigma) = -1$ .

**Bemerkungen:**

- Die identische Permutation ist also qua Definition stets gerade.
- Transpositionen sind dagegen stets ungerade, denn setzt man  $\sigma := \text{id}_n$  in die zweite Bedingung ein, so folgt

$$\text{sign}(\tau) = \text{sign}(\tau \circ \text{id}_n) = -\text{sign}(\text{id}_n) = -1$$

Da nun nach Satz 4.15 jede Permutation  $\sigma \in S_n$  in  $r$  kanonische Zyklen zerlegbar ist und weiterhin in  $t$  Transpositionen zerlegbar ist, wobei  $t$  nach Satz 8.12 eindeutig gerade oder ungerade ist, gilt auch folgender

**Satz 8.14** (Vorzeichen einer Permutation). Für  $n \in \mathbb{N}$  sei  $\sigma \in S_n$  eine Permutation mit  $r$  kanonischen Zyklen. Dann gilt:

$$\text{sign}(\sigma) = (-1)^{n-r}$$

Für jede Zerlegung von  $\sigma$  in  $t$  Transpositionen gilt weiterhin:

$$\text{sign}(\sigma) = (-1)^t$$

**Bemerkungen:**

- Die beiden Definitionen sind gleichwertig. Nach Satz 8.12 ist  $t$  kongruent mit  $(n-r)$  modulo 2. Falls  $(n-r)$  gerade ist, so ist dies also auch  $t$ , und das Vorzeichen ist nach beiden Formeln  $+1$ .
- Da sowohl die Zahl der kanonischen Zyklen als auch die Parität der Anzahl von Transpositionen einer beliebigen Zerlegung für jede Permutation eindeutig definiert sind, ist die Funktion aus Definition 8.13 über die Formeln im obigen Satz wohldefiniert.
- Mit Satz 8.11 folgt, dass ein Zyklus der Länge  $k$  das Vorzeichen  $(-1)^{k-1}$  besitzt. Zyklen mit einer geraden Zahl von Elementen sind also ungerade Permutationen; Zyklen mit ungerader Elementzahl sind gerade Permutationen.

Es gelten weiterhin die Eigenschaften aus folgendem

**Satz 8.15** (Vorzeichen von Permutationen bei Gruppenoperationen). Seien  $\sigma, \pi \in S_n$  Permutationen, dann gilt:

1.  $\text{sign}(\sigma \circ \pi) = \text{sign}(\pi \circ \sigma) = \text{sign}(\sigma) \cdot \text{sign}(\pi)$
2.  $\text{sign}(\sigma^{-1}) = \text{sign}(\sigma)$

(Beweis: S. 337.)

**Bemerkung:** Mit der letzten Bemerkung zu Satz 8.14 und der ersten Aussage dieses Satzes folgt, dass sich das Vorzeichen einer beliebig in Zyklen zerlegten Permutation direkt als Produkt der Vorzeichen der Zyklen schreiben lässt – auch wenn es sich nicht um die kanonischen Zyklen handelt.

**Beispiel:** In den Beispielen zu Satz 8.11 hatten wir diverse Zyklenzerlegungen einer Permutation kennen gelernt. Da Transpositionen ebenfalls Zyklen sind, listen wir sie mit auf:

$$\begin{aligned}\sigma &= \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 7 & 6 & 5 & 3 & 2 & 4 & 1 \end{pmatrix} = (1 \ 7)(2 \ 6 \ 4 \ 3 \ 5) \\ &= (1 \ 7)(2 \ 5)(2 \ 3)(2 \ 4)(2 \ 6) \\ &= (3 \ 7 \ 1 \ 4)(2 \ 6 \ 1 \ 3 \ 5) \\ &= (3 \ 4)(3 \ 1)(3 \ 7)(2 \ 5)(2 \ 3)(2 \ 1)(2 \ 6)\end{aligned}$$

Aus der kanonischen Zerlegung können wir schon ablesen, dass  $(n - r) = (7 - 2) = 5$ ; es handelt sich also um eine ungerade Permutation aus  $S_7$ .

Ohne zu bedenken, dass die vier angegebenen Zerlegungen der gleichen Permutation  $\sigma$  entsprechen, können wir mit der obigen Bemerkung die Vorzeichen der Gesamtpermutationen ausrechnen (in gleicher Reihenfolge):

$$(-1)^1 \cdot (-1)^4 = -1; \quad ((-1)^1)^5 = (-1)^5 = -1; \quad (-1)^3 \cdot (-1)^4 = -1; \quad ((-1)^1)^7 = -1$$

## 8.3 Determinanten

Nun wenden wir uns wieder den Matrizen zu; genauer gesagt, den quadratischen. Wir erinnern uns, dass lineare Gleichungssysteme genau dann eindeutig lösbar waren, wenn sie sich mit den eingeführten Zeilenoperationen auf eine Zeilen-Stufen-Form mit quadratischer Koeffizientenmatrix bringen lassen (nach eventueller Entfernung von Nullzeilen im Fall von anfänglich überbestimmten LGS).

Die entscheidende Größe hierbei war der *Rang* der Koeffizientenmatrix, der der Anzahl von linear unabhängigen Zeilen bzw. Spalten der Matrix entspricht. Für quadratische Matrizen  $A \in \mathbb{R}^{(n,n)}$  ist nach Satz 8.10:

$$\text{rg } A \leq n$$

Nun wissen wir schon, wie wir den Rang berechnen können. Manchmal ist aber nicht dessen genauer Wert interessant, sondern zunächst die Frage, ob er echt kleiner als  $n$  ist, oder gleich  $n$ . Dies entscheidet die eindeutige Lösbarkeit. Im Kapitel 9 werden wir noch sehen, dass damit gleichzeitig auch die Invertierbarkeit der quadratischen Koeffizientenmatrix entschieden ist.

Es stellt sich heraus, dass sich die Frage am Zahlenwert einer besonderen Funktion ablesen lässt, der *Determinanten* von  $A$ .

### 8.3.1 Allgemeine Eigenschaften

Wir stellen die Determinantenfunktion zunächst axiomatisch vor und leiten dann einige fundamentale Eigenschaften ab, die das Rechnen mit Determinanten ermöglichen/vereinfachen.

Zu Beginn die axiomatische Definition der Determinante nach Weierstraß<sup>2</sup>:

**Definition 8.16** (Weierstraß-Axiome). *Die Funktion  $\det : \mathbb{R}^{(n,n)} \rightarrow \mathbb{R}$  heißt Determinante, falls für jede Matrix  $A = (\vec{a}_1 \ \cdots \ \vec{a}_n) \in \mathbb{R}^{(n,n)}$  die folgenden Eigenschaften gelten:*

1. *Normiertheit:*  $\det(\mathbb{1}_n) = 1$
2. *Alterniertheit:* Falls zwei Spalten  $\vec{a}_j, \vec{a}_k$  ( $j \neq k$ ) gleich sind, gilt  $\det A = 0$
3. *Multilinearität in den Spalten:* Für  $r, s \in \mathbb{R}$  und zwei beliebige Spalten  $\vec{a}_j, \vec{a}_k$  gilt an einer beliebigen, aber festen Spaltenposition:

$$\det(\cdots \ (r\vec{a}_j + s\vec{a}_k) \ \cdots) = r \cdot \det(\cdots \ \vec{a}_j \ \cdots) + s \cdot \det(\cdots \ \vec{a}_k \ \cdots)$$

(die Spaltenposition ist in allen drei Matrizen die gleiche)

---

<sup>2</sup>K. Weierstraß, dt. Mathematiker

Einstweilen lässt sich für konkrete quadratische Matrizen hier noch nicht viel aussagen, daher halten wir zunächst weitere Eigenschaften der Determinante fest, die aus obigen Axiomen ableitbar sind. Es wird empfohlen, den Beweis ebenfalls durchzugehen; dies sollte die Rechenregeln nachvollziehbarer machen.

**Satz 8.17** (Rechenregeln für Determinanten). *Alle Determinanten nach Definition 8.16 besitzen für Matrizen  $A = (\vec{a}_1 \ \cdots \ \vec{a}_n) \in \mathbb{R}^{(n,n)}$  die Eigenschaften*

1. *Skalierung von Spalten – für  $r \in \mathbb{R}$  gilt:*

$$\det(\cdots \ r\vec{a}_j \ \cdots) = r \cdot \det(\cdots \ \vec{a}_j \ \cdots) = r \cdot \det A$$

2. *Wert der Determinante bei einer Null-Spalte:*

$$\det(\cdots \ \vec{0} \ \cdots) = 0$$

3. *Invarianz bei Addition einer skalierten Spalte  $s\vec{a}_k$  zur Spalte  $j$  mit  $j \neq k$*

$$\det(\cdots \ (\vec{a}_j + s\vec{a}_k) \ \cdots) = \det(\cdots \ \vec{a}_j \ \cdots) = \det A$$

4. *Falls die Spalten von  $A$  linear abhängig sind, ist  $\det A = 0$*

5. *Antisymmetrie bei Vertauschung der Spalten – für  $j < k$  gilt:*

$$\det(\cdots \ \vec{a}_k \ \cdots \ \vec{a}_j \ \cdots) = -\det(\cdots \ \vec{a}_j \ \cdots \ \vec{a}_k \ \cdots) = -\det A$$

(Beweis: S. 338.)

**Bemerkung:** Manchmal wird auch die Antisymmetrie anstatt der Alterniertheit als axiomatische Eigenschaft betrachtet. Wenn zwei Spalten von  $A$  gleich sind, die Determinante bei deren Vertauschen das Vorzeichen wechseln muss, dabei aber offensichtlich identisch bleibt, kann ihr Wert nur null betragen.

Diese Argumentation ist allerdings nur dann richtig, wenn der Körper  $\mathbb{K}$ , in den die Determinantenfunktion abbildet, außer der 0 keine weiteren Elemente besitzt, die ihr eigenes additives Inverses sind. Im Körper  $GF(2^3)$  wäre es z.B. möglich, die Null auch durch Addition von 4 und 4 zu erreichen.

Für die reellen Zahlen sind beide Kriterien gleichwertig.

**Beispiel:** Mit diesen Regeln können wir schon problemlos die Determinante einer allgemeinen  $2 \times 2$ -Matrix bestimmen:

$$A := \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

Nun verwenden wir die Einheitsvektoren aus der Standardbasis  $E_2 = \{\vec{e}_1, \vec{e}_2\}$ , um die Spaltendarstellung von  $A$  geschickt auszudrücken:

$$A = \left( \left( a \begin{pmatrix} 1 \\ 0 \end{pmatrix} + c \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right) \ \left( b \begin{pmatrix} 1 \\ 0 \end{pmatrix} + d \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right) \right) = ((a\vec{e}_1 + c\vec{e}_2) \ (b\vec{e}_1 + d\vec{e}_2))$$

Wir verwenden die Multilinearität der Determinante:

$$\begin{aligned} \det A &= \det((a\vec{e}_1 + c\vec{e}_2) \ (b\vec{e}_1 + d\vec{e}_2)) \\ &= a \cdot \det(\vec{e}_1 \ (b\vec{e}_1 + d\vec{e}_2)) + c \cdot \det(\vec{e}_2 \ (b\vec{e}_1 + d\vec{e}_2)) \\ &= a \cdot [b \cdot \det(\vec{e}_1 \ \vec{e}_1) + d \cdot \det(\vec{e}_1 \ \vec{e}_2)] \\ &\quad + c \cdot [b \cdot \det(\vec{e}_2 \ \vec{e}_1) + d \cdot \det(\vec{e}_2 \ \vec{e}_2)] \end{aligned}$$

Nun verschwinden zwei dieser vier Determinanten, da jeweils beide Spaltenvektoren gleich sind, aufgrund der Alterniertheit. Es bleibt:

$$\cdots = ad \cdot \det(\vec{e}_1 \ \vec{e}_2) + bc \cdot \det(\vec{e}_2 \ \vec{e}_1)$$

Jetzt ist die erste Determinante aber genau die Determinante der Einheitsmatrix  $\mathbb{1}_2$  – diese beträgt wegen der Normiertheit also 1; die zweite Matrix geht aus  $\mathbb{1}_2$  durch *Vertauschen* der beiden Spalten hervor und muss wegen der Antisymmetrie also umgekehrtes Vorzeichen in der Determinanten haben. Damit erhalten wir:

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc$$

Oft merkt man sich diese Determinante als “Produkt der Hauptdiagonalelemente minus Produkt der Nebendiagonalelemente”. Man beachte aber, dass dies nur für  $2 \times 2$ -Matrizen richtig ist.

Anwendungsbeispiele:

•

$$\det \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 1 \cdot 1 - 0 \cdot 0 = 1 = \det \mathbb{1}_2 \quad \checkmark$$

•

$$\det \begin{pmatrix} 2 & 3 \\ -4 & 2 \end{pmatrix} = 2 \cdot 2 - 3 \cdot (-4) = 4 + 12 = 16$$

•

$$\det \begin{pmatrix} 3 & -4 \\ -12 & 16 \end{pmatrix} = 3 \cdot 16 - (-4) \cdot (-12) = 48 - 48 = 0$$

Im Prinzip lassen sich auch die Determinanten größerer Matrizen nach der gleichen Methode berechnen – allerdings gibt es dabei zwei Schwierigkeiten:

- Zum einen erhalten wir für jede der  $n$  Spalten eine Linearkombination von  $n$  Einheitsvektoren. Das gibt insgesamt  $n^n$  Beiträge – nur für kleine  $n$  ist das praktikabel.
- Weiterhin verschwinden die meisten dieser Beiträge wieder, weil mindestens zwei der betrachteten Spalten aus Einheitsvektoren identisch sind. Übrig bleiben dann nur Kombinationen von Einheitsvektoren, die alle paarweise verschieden sind – das sind genau die Permutationsmatrizen. Dieses Vorgehen führt uns im übernächsten Unterabschnitt zur *Leibnizformel*.

### 8.3.2 Permutationsmatrizen

Wir hatten für die Permutation  $\sigma \in S_n$  die zugehörige Permutationsmatrix über ihre Spaltendarstellung definiert als

$$P_\sigma = (\vec{e}_{\sigma(1)} \quad \cdots \quad \vec{e}_{\sigma(n)})$$

Nun wissen wir aus dem vorigen Abschnitt, dass Permutationen sich stets als Kompositionen von Transpositionen (Vertauschungen) schreiben lassen. Nach Satz 8.17 bewirkt jede Vertauschung von zwei Spalten einen Vorzeichenwechsel der Determinante. Starten wir also mit der Einheitsmatrix  $\mathbb{1}_n$ , so ist es möglich, jede Permutationsmatrix  $P_\sigma$  durch entsprechende Vertauschungen herzustellen (dazu kann man sich z.B. der kanonischen Zyklen von  $\sigma$  bedienen).

Das Vorzeichen der identischen Permutation  $\text{id}_n$  ist dabei stets 1 (siehe Definition 8.13). Und da  $\mathbb{1}_n$  genau die Permutationsmatrix zur identischen Permutation ist, und nach Definition 8.16 Determinante 1 besitzt, geschieht nun beim Vertauschen von zwei Spalten jeweils ein Vorzeichenwechsel sowohl im Signum der Permutation als auch in der Determinanten (nach Satz 8.17). Also gilt folgender

**Satz 8.18** (Determinante einer Permutationsmatrix). *Mit der Permutation  $\sigma \in S_n$  gilt für die zugehörige Permutationsmatrix  $P_\sigma$ :*

$$\det P_\sigma = \text{sign}(\sigma) \in \{-1, +1\}$$

**Beispiel:** Wir ermitteln für  $n = 5$  zur Permutationsmatrix

$$P_\sigma := \begin{pmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix}$$

zunächst die zugehörige Permutation: Wegen  $P_\sigma = (\vec{e}_4 \quad \vec{e}_2 \quad \vec{e}_5 \quad \vec{e}_1 \quad \vec{e}_3)$  ist

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 2 & 5 & 1 & 3 \end{pmatrix} = (1 \ 4)(2)(3 \ 5)$$

Das Vorzeichen von  $\sigma$  ist  $(-1) \cdot 1 \cdot (-1) = 1$ ; damit ist auch die Determinante von  $P_\sigma$  gefunden.

Werden nun in einer Matrix die Spaltenvektoren permutiert (dies lässt sich nach Satz 7.26 durch Multiplikation einer Permutationsmatrix von rechts bewirken), so lässt sich diese Permutation analog über eine Abfolge von Vertauschungen der Spalten ermitteln. Die Anzahl der Vertauschungen ist gerade für gerade Permutationen und ungerade für ungerade Permutationen. Da die Determinante aufgrund der Antisymmetrie bei jedem Tausch das Vorzeichen wechselt (siehe Satz 8.17), gilt entsprechend auch folgender

**Satz 8.19** (Determinante einer Matrix mit permutierten Spalten). *Für eine quadratische Matrix  $A = (\vec{a}_1 \quad \cdots \quad \vec{a}_n) \in \mathbb{R}^{(n,n)}$  und  $\sigma \in S_n$  gilt:*

$$\det(A \cdot P_\sigma) = \det(\vec{a}_{\sigma(1)} \quad \cdots \quad \vec{a}_{\sigma(n)}) = \text{sign}(\sigma) \cdot \det(\vec{a}_1 \quad \cdots \quad \vec{a}_n) = \text{sign}(\sigma) \cdot \det A$$

**Bemerkung:** Diese wichtige Tatsache benötigen wir zum Ende des Kapitels im Beweis zum *Determinanten-Produktsatz* (s.u.).

### 8.3.3 Leibniz-Formel

Wie nach dem Beispiel zu Satz 8.17 bereits angemerkt, lässt sich die Determinante einer Matrix  $A \in \mathbb{R}^{(n,n)}$  berechnen, indem man die Spalten nach den kartesischen Einheitsvektoren der Standardbasis  $E_n$  entwickelt und mannigfach die Multilinearitätseigenschaft verwendet. Wir hatten bereits eingesehen, dass dabei viele Beiträge verschwinden und dass nur solche zu  $\det A$  beitragen, für die in jeder Spalte der betreffenden Matrix ein anderer Einheitsvektor steht; das sind genau die  $n!$  verschiedenen Permutationsmatrizen  $P_\sigma$  mit  $\sigma \in S_n$ .

Im vorigen Unterabschnitt haben wir fest gestellt, dass die Determinanten solcher Permutationsmatrizen stets genau dem Signum der zugehörigen Permutationen entsprechen. Wir formulieren dies allgemein (und auf zwei Arten) in folgendem

**Satz 8.20** (Leibnizformel). <sup>3</sup> *Die Determinante einer quadratischen Matrix  $A \in \mathbb{R}^{(n,n)}$  berechnet sich als*

$$\det A = \sum_{\sigma \in S_n} \text{sign}(\sigma) \cdot A_{1,\sigma(1)} \cdot \cdots \cdot A_{n,\sigma(n)} = \sum_{\sigma \in S_n} \text{sign}(\sigma) \cdot A_{\sigma(1),1} \cdot \cdots \cdot A_{\sigma(n),n}$$

(Beweis: S. 339.)

**Bemerkung:** Die Leibniz-Formel ist wegen der nötigen Vorarbeit (Auffinden aller Permutationen in  $S_n$  und Ermittlung ihrer jeweiligen Vorzeichen) für die Praxis eher wenig hilfreich; sie kann aber verwendet werden, um einige weitere grundsätzliche Zusammenhänge zu zeigen.

Eine praktikablere Methode zur Determinantenberechnung lernen wir in Kürze mit der *Laplace-Entwicklung* kennen (s.u.).

#### Beispiele:

- Wir reproduzieren die weiter oben bereits berechnete Formel für Determinanten von  $2 \times 2$ -Matrizen. Dazu listen wir zunächst die Permutationen in  $S_2$  auf, zusammen mit ihrer kanonischen Zyklendarstellung und ihrem Vorzeichen:

Name	Darstellungen	Vorzeichen
$\text{id}_2 =: \sigma_1$	$\begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} = (1)(2)$	+1
$\sigma_2$	$\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} = (1 \ 2)$	-1

<sup>3</sup>G.W. Leibniz, dt. Mathematiker

Damit folgt für  $A \in \mathbb{R}^{(2,2)}$ :

$$\det A = (+1) \cdot A_{1,1}A_{2,2} + (-1) \cdot A_{1,2}A_{2,1}$$

Dies ist genau die gleiche Formel wie oben bereits eingerahmt.

- Für dreireihige Matrizen verfahren wir analog. Zunächst die sechs Permutationen aus  $S_3$

Name	Darstellungen	Vorzeichen
$\text{id}_3 =: \sigma_1$	$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} = (1)(2)(3)$	+1
$\sigma_2$	$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} = (1)(2 \ 3)$	-1
$\sigma_3$	$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} = (1 \ 2)(3)$	-1
$\sigma_4$	$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} = (1 \ 2 \ 3)$	+1
$\sigma_5$	$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} = (1 \ 3 \ 2)$	+1
$\sigma_6$	$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} = (1 \ 3)(2)$	-1

Damit folgt für  $A \in \mathbb{R}^{(3,3)}$ :

$$\begin{aligned} \det A = & A_{1,1}A_{2,2}A_{3,3} + A_{1,2}A_{2,3}A_{3,1} + A_{1,3}A_{2,1}A_{3,2} \\ & - A_{1,1}A_{2,3}A_{3,2} - A_{1,2}A_{2,1}A_{3,3} - A_{1,3}A_{2,2}A_{3,1} \end{aligned}$$

Als Merkhilfe kann auch (aber nur für  $n = 3$ ) die *Regel von Sarrus* verwendet werden. Hier setzt man die Matrix in beiden Richtungen periodisch fort. Dann ergeben sich die drei positiven Beiträge durch Produkte von Matrixelementen entlang der positiven Diagonalen (nach rechts unten), und die drei negativen durch Produkte von Elementen entlang der negativen Diagonalen (nach links unten). Siehe dazu Abbildung 8.3 (S. 232).

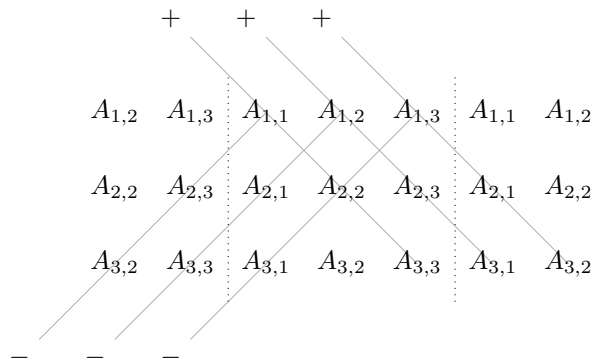


Abbildung 8.3: Regel von Sarrus

Man beachte, dass die Sarrus-Regel nur für dreireihige Matrizen funktioniert. Für höherdimensionale Matrizen aus  $\mathbb{R}^{(n,n)}$  mit  $n > 3$  gilt nicht länger, dass  $n! = 2n$ . Folglich werden mit den  $2n$  Diagonalen nur noch Bruchteile sämtlicher Permutationen gefunden – für  $n = 4$  z.B. ein Drittel. Weiterhin entfällt für die Diagonalen das bei Sarrus leicht zu merkende einheitliche Schema der zugehörigen Permutations-Vorzeichen – beispielsweise hätte bei  $n = 4$  die Diagonale “rechts neben” der Hauptdiagonalen der Matrix die Permutation

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix} = (1 \ 2 \ 3 \ 4)$$

mit Vorzeichen  $(-1)$ .



- Ein Zahlenbeispiel für die Sarrus-Regel:

$$\det \begin{pmatrix} 2 & 1 & 4 \\ 3 & -2 & 1 \\ 2 & 2 & 1 \end{pmatrix} = (-4) + 2 + 24 - (-16) - 3 - 4 = 31$$


---

Wir haben im Beweis zur Leibniz-Formel gezeigt, dass jede Determinante nach Weierstraß genau die im Satz angegebene Form besitzt. Noch nicht klar ist, dass jede Funktion, die eine Matrix  $A \in \mathbb{R}^{(n,n)}$  auf eine Zahl nach der Vorschrift im Satz abbildet, auch eine Determinante ist. Es könnte also noch sein, dass die Gleichung im obigen Satz nicht immer von rechts nach links lesbar ist<sup>4</sup>. Dem ist allerdings nicht so, wie wir nun formulieren. Der Beweis des Satzes ist beigelegt, allerdings etwas länglich und natürlich nicht prüfungsrelevant (Neugierige mögen natürlich trotzdem gerne hinein schauen!).

**Satz 8.21** (Existenz der Determinante). *Für  $n \in \mathbb{N}$  ist jede Abbildung  $f : \mathbb{R}^{(n,n)} \rightarrow \mathbb{R}$  mit*

$$f(A) = \sum_{\sigma \in S_n} \text{sign}(\sigma) \cdot A_{1,\sigma(1)} \cdot \cdots \cdot A_{n,\sigma(n)}$$

*eine Determinante nach Weierstraß gemäß Definition 8.16.*

(Beweis: S. 341.)

**Bemerkung:** Damit ist geklärt, dass Determinanten genau die Funktionen auf quadratischen Matrizen sind, deren Abbildungseigenschaften in der Leibnizformel ausgedrückt sind.

### 8.3.4 Transponierte Matrizen

Im Beweis zur Leibnizformel (Satz 8.20) wurde gezeigt, dass die beiden Schreibweisen (permutierte Indices für die Zeilen oder für die Spalten einer Matrix) gleichwertig sind. Wir berechnen nun für eine gegebene Matrix  $A \in \mathbb{R}^{(n,n)}$  die Determinante der Transponierten. Mit Leibniz gilt:

$$\begin{aligned} \det A^T &= \sum_{\sigma \in S_n} \text{sign}(\sigma) \cdot A_{1,\sigma(1)}^T \cdot \cdots \cdot A_{n,\sigma(n)}^T \\ &= \sum_{\sigma \in S_n} \text{sign}(\sigma) \cdot A_{\sigma(1),1} \cdot \cdots \cdot A_{\sigma(n),n} \\ &= \det A \end{aligned}$$

Hierbei haben wir für die Transponierte die eine Schreibweise, für die ursprüngliche Matrix jedoch die andere Schreibweise der Leibnizformel verwendet. Wir halten fest:

**Satz 8.22** (Determinante der Transponierten). *Für  $A \in \mathbb{R}^{(n,n)}$  gilt:*

$$\det A^T = \det A$$

**Bemerkung:** Damit gelten im Übrigen auch sämtliche Aussagen, die wir bisher im Zusammenhang mit Determinanten über Spalten getätigt haben, in gleicher Weise für die *Zeilen* einer Matrix – inklusive der Rechenregeln und der Axiome: Denn die Zeilen von  $A$  sind die Spalten von  $A^T$ ; von dort aus kann wie oben über Spalten weiter argumentiert werden.

---

Wir werden weiter unten in einem separaten Unterabschnitt einige Beispiele zeigen, in denen Determinanten vor dem Anwenden allgemeiner Berechnungsformeln zunächst mit Zeilen- und Spaltenoperationen manipuliert werden. Auf diese Weise lassen sich nämlich meist zusätzliche Matrixelemente auf 0 bringen (bei weiterhin gleichem Wert der Determinanten), sodass in den Berechnungsformeln weniger nichttriviale Beiträge zu berücksichtigen sind.

---

<sup>4</sup>Dies ist eine ähnliche Frage wie die, ob jede lineare Abbildung durch ein Produkt aus Matrix und Vektor beschrieben werden kann; siehe Kapitel 7

### 8.3.5 Dreiecksmatrizen (und Diagonalmatrizen)

**Definition 8.23** (Dreiecksmatrix). Eine quadratische Matrix  $A \in \mathbb{R}^{(n,n)}$  heißt obere Dreiecksmatrix, falls für  $j > k$  stets

$$A_{j,k} = 0$$

gilt. Falls die Matrixelemente  $A_{j,k}$  für  $j < k$  verschwinden, handelt es sich dagegen um eine untere Dreiecksmatrix.

**Bemerkungen:**

- In den jeweils nicht weiter eingeschränkten Bereichen (incl. der Diagonalen) dürfen selbstverständlich ebenfalls Nullen stehen.
- Die Matrizen, die sowohl obere als auch untere Dreiecksmatrizen sind, sind genau die *Diagonalmatrizen* aus Definition 7.16.

**Beispiel:** Die Matrizen

$$\begin{pmatrix} 1 & 4 & 7 & 2 \\ 0 & -2 & 1 & 4 \\ 0 & 0 & 0 & -7 \\ 0 & 0 & 0 & 3 \end{pmatrix}, \quad \begin{pmatrix} -2 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & -1 & 2 & 0 & 0 \\ 3 & 0 & 3 & 1 & 0 \\ 2 & 3 & 1 & 3 & 2 \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} 2 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

sind Dreiecksmatrizen. Dabei ist die linke eine obere und die mittlere eine untere Dreiecksmatrix; die rechte Matrix ist sogar diagonal.

Für die Determinanten solcher Matrizen gilt folgender

**Satz 8.24** (Determinanten von Dreiecksmatrizen). Für jede Dreiecksmatrix  $A \in \mathbb{R}^{(n,n)}$  gilt:

$$\det A = \prod_{j=1}^n A_{j,j} = A_{1,1} \cdot \dots \cdot A_{n,n}$$

(Beweis: S. 343.)

**Bemerkung:** Es sind für die Determinante hier also nur die Diagonalelemente der Matrix zu beachten.

**Beispiel:** Für die drei Matrizen aus dem Beispiel zu Definition 8.23 ergeben sich die Determinanten

$$1 \cdot (-2) \cdot 0 \cdot 3 = 0; \quad (-2) \cdot 1 \cdot 2 \cdot 1 \cdot 2 = -8 \quad \text{und} \quad 2 \cdot (-3) \cdot 0 = 0$$

Im ersten und dritten Fall befand sich auf der Diagonalen eine 0, sodass der Wert der gesamten Determinanten jeweils auf 0 gezogen wurde.

Mit diesen Mitteln können wir nun eine wichtige Tatsache beweisen, die später nochmals aufgegriffen wird:

**Satz 8.25** (Lineare Abhängigkeit und Determinante). Genau dann, wenn die Spalten (oder die Zeilen) einer Matrix  $A \in \mathbb{R}^{(n,n)}$  linear abhängig sind, gilt

$$\det A = 0$$

(Beweis: S. 343.)

**Bemerkungen:**

- Eine Richtung der obigen Äquivalenz hatten wir für die Spalten bereits in Satz 8.17 gezeigt: Falls diese linear abhängig sind, verschwindet die Determinante. Der Beweis behandelt also hier nur die umgekehrte Richtung.
- Äquivalent dazu gilt auch: Die Spalten (und Zeilen) einer quadratischen Matrix sind genau dann linear *unabhängig*, wenn die Determinante der Matrix *ungleich* 0 ist.

### 8.3.6 Laplace-Entwicklung

Wir haben mit der Leibniz-Formel bereits die allgemeine Methode zur Berechnung von Determinanten kennen gelernt; diese funktioniert auch immer – allerdings wird die Zahl von Permutationen schnell unübersichtlich und ist schon bei  $n = 4$  mit  $4! = 24$  Beiträgen für handschriftliche Lösungen nicht mehr praktikabel.

Für uns bieten sich zwei Methoden an, um auch höher-reihige Determinanten zu berechnen:

- Mit Zeilen- und Spaltenoperationen eine Struktur herstellen, die möglichst einer Dreiecksmatrix entspricht – also systematisch viele Nullen erzeugen. Dabei ist aber zu bedenken, dass das Skalieren einzelner Zeilen oder Spalten jeweils den Wert der Determinante skaliert – das ist zu kompensieren. Auch kehrt das Tauschen von Zeilen oder Spalten das Vorzeichen der Determinante um.

Bei LGS war hingegen das Skalieren oder Tauschen von Gleichungen (Zeilen) ohne weitere Konsequenzen möglich!

- Die Berechnung einer Determinanten für  $\mathbb{R}^{(n,n)}$  auf die Berechnung von Determinanten in  $\mathbb{R}^{(n-1,n-1)}$  zurück führen. Dies ist auf einheitliche Art möglich und führt nach endlicher Zeit zu Determinanten in  $\mathbb{R}^{(3,3)}$  oder  $\mathbb{R}^{(2,2)}$ , die mit den bereits bekannten Formeln direkt berechenbar sind.

Die erste Methode ist als Vorbereitung oft sinnvoll – auch als Vorarbeit für die zweite Methode, die *Laplace-Entwicklung*<sup>5</sup>.

**Satz 8.26** (Laplace-Entwicklung). *Für eine Matrix  $A \in \mathbb{R}^{(n,n)}$  und  $M_n := \{1, \dots, n\}$  mit  $n > 1$  ist die Determinante von  $A$  gegeben als*

- Entwicklung nach Spalte  $k$ : Für  $k \in M_n$  beliebig, aber fest, gilt:

$$\det A = \sum_{j=1}^n (-1)^{j+k} \cdot A_{j,k} \cdot \det A_{\setminus(j,k)}$$

- Entwicklung nach Zeile  $j$ : Für  $j \in M_n$  beliebig, aber fest, gilt:

$$\det A = \sum_{k=1}^n (-1)^{j+k} \cdot A_{j,k} \cdot \det A_{\setminus(j,k)}$$

Dabei sei  $A_{\setminus(j,k)}$  die Matrix aus  $\mathbb{R}^{(n-1,n-1)}$ , die entsteht, wenn aus  $A$  die  $j$ -te Zeile sowie die  $k$ -te Spalte entfernt werden.

(Beweis: S. 344.)

#### Bemerkungen:

- Die Berechnung einer  $n$ -reihigen Determinante wird also durch die Berechnung von  $n$  *Unterdeterminanten* ersetzt, die jeweils  $(n-1)$ -reihig sind<sup>6</sup>. Wendet man diese Formeln wiederholt an, so gelangt man stets zur Berechnung von drei- oder zwei-reihigen Unterdeterminanten nach den bereits bekannten Formeln.
- Skaliert werden die Unterdeterminanten jeweils mit dem Matrixelement, dessen Koordinaten die Unterdeterminante definieren, und mit einem Vorzeichen, das sich (s.u.) leicht ermitteln lässt. Siehe dazu den Beweis. Da ein Element aus  $\{j, k\}$  beim Entwickeln stets fest ist und das andere von 1 bis  $n$  läuft, alterniert das Vorzeichen in regelmäßiger Weise.
- Je weniger Unterdeterminanten zu berechnen sind, desto günstiger ist das Verfahren – das ist dann der Fall, wenn die Spalte oder Zeile, nach welcher entwickelt wird, viele Nullen enthält. Diese tauchen multiplikativ in den oben angeschriebenen Summen auf, womit sich das Berechnen der zugehörigen Unterdeterminante erübrigt.

Falls gewünscht (oder gefordert), können zusätzliche Nullen im Voraus durch Spalten- und/oder Zeilenoperationen erzeugt werden. Hier empfehlen sich insbesondere die skalierten Additionen, welche den Wert der Determinanten unverändert lassen (ansonsten müssen die Operationen durch geeignete Faktoren außerhalb kompensiert werden – etwa  $(-1)$  beim Vertauschen von zwei Zeilen, oder  $\frac{1}{c}$  beim Skalieren einer Spalte mit  $c \neq 0$ ).

<sup>5</sup>P.-S. Laplace, frz. Mathematiker

<sup>6</sup>Stichwort: Rekursion!

- Die Vorzeichen für die Skalierung der Unterdeterminanten sind schachbrettartig angeordnet. Als Beispiel die Anordnung für  $n = 7$ , nicht vollständig ausgefüllt:

$$\begin{pmatrix} + & - & + & - & + & - & + \\ - & + & - & & & & - \\ + & - & + & - & & & + \\ - & & - & + & - & & - \\ + & & & - & + & - & + \\ - & & & & - & + & - \\ + & - & + & - & + & - & + \end{pmatrix}$$

Die Einträge auf der Diagonalen ( $j = k$ ) haben stets positives Vorzeichen, da

$$(-1)^{j+j} = (-1)^{2j} = 1$$

### Beispiele:

- Wir zeigen zunächst symbolisch die Entwicklung nach der zweiten Spalte für eine vierreihige Determinante. Die Matrix ist:

$$\begin{pmatrix} \cdots & a & \cdots & \cdots \\ \cdots & b & \cdots & \cdots \\ \cdots & c & \cdots & \cdots \\ \cdots & d & \cdots & \cdots \end{pmatrix}$$

Nun zur Ermittlung der vier Unterdeterminanten (die zugehörigen Elemente umrahmt, die zu streichenden Elemente der Ursprungsmatrix entfernt bis auf den Skalierungsfaktor):

$$\left( \begin{array}{c} \cdots \cdots \cdots \\ \boxed{\begin{array}{ccc} \cdots & \cdots & \cdots \end{array}} \\ \cdots \cdots \cdots \end{array} \right) \quad \left( \begin{array}{c} \boxed{\begin{array}{ccc} \cdots & \cdots & \cdots \end{array}} \\ \cdots \cdots \cdots \\ \boxed{\begin{array}{ccc} \cdots & \cdots & \cdots \end{array}} \\ \cdots \cdots \cdots \end{array} \right) \quad \left( \begin{array}{c} \cdots \cdots \cdots \\ \cdots \cdots \cdots \\ \boxed{\begin{array}{ccc} \cdots & \cdots & \cdots \end{array}} \\ \cdots \cdots \cdots \end{array} \right) \quad \left( \begin{array}{c} \cdots \cdots \cdots \\ \cdots \cdots \cdots \\ \cdots \cdots \cdots \\ \boxed{\begin{array}{ccc} \cdots & \cdots & \cdots \end{array}} \end{array} \right)$$

Die eingerahmten Matricelemente bilden jeweils  $(3 \times 3)$ -Matrizen, deren Determinanten zu berechnen wären. Die Skalierungsfaktoren (mit positionsabhängigem Vorzeichen aus dem Schachbrettmuster) für diese Determinanten sind:

$$-a, \quad b, \quad -c \quad \text{und} \quad d$$

- Wir wiederholen die Berechnung der zweireihigen Determinante, indem wir nach der ersten Spalte entwickeln (dies annotieren wir als "S.1"). Nach Streichen der jeweiligen Zeilen und Spalten bleibt als Unterdeterminante nur jeweils eine einzige Zahl übrig:

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} \stackrel{\text{S.1}}{=} (+1) \cdot a \cdot \det(d) + (-1) \cdot c \cdot \det(b) = ad - bc$$

- Auch die Regel von Sarrus erhalten wir direkt wieder. Wir entwickeln nach der ersten Zeile (und annotieren das als "Z.1"):

$$\begin{aligned} \det \begin{pmatrix} A_{1,1} & A_{1,2} & A_{1,3} \\ A_{2,1} & A_{2,2} & A_{2,3} \\ A_{3,1} & A_{3,2} & A_{3,3} \end{pmatrix} &\stackrel{\text{Z.1}}{=} (+1) \cdot A_{1,1} \cdot \det \begin{pmatrix} A_{2,2} & A_{2,3} \\ A_{3,2} & A_{3,3} \end{pmatrix} \\ &\quad + (-1) \cdot A_{1,2} \cdot \det \begin{pmatrix} A_{2,1} & A_{2,3} \\ A_{3,1} & A_{3,3} \end{pmatrix} \\ &\quad + (+1) \cdot A_{1,3} \cdot \det \begin{pmatrix} A_{2,1} & A_{2,2} \\ A_{3,1} & A_{3,2} \end{pmatrix} \\ &= A_{1,1}(A_{2,2}A_{3,3} - A_{2,3}A_{3,2}) \\ &\quad - A_{1,2}(A_{2,1}A_{3,3} - A_{2,3}A_{3,1}) \\ &\quad + A_{1,3}(A_{2,1}A_{3,2} - A_{2,2}A_{3,1}) \end{aligned}$$

- Wir berechnen die Determinante der folgenden vierreihigen Matrix  $A$  mit einer Laplace-Entwicklung nach der dritten Zeile.

$$A := \begin{pmatrix} 1 & 2 & 2 & 3 \\ 4 & 1 & 4 & 2 \\ 2 & 1 & 1 & 2 \\ 1 & 3 & 3 & 2 \end{pmatrix}$$

Wir erhalten (die Vorzeichen der Skalierungsfaktoren ergeben sich aus dem Schachbrettmuster):

$$\det A \stackrel{\text{Z.3}}{=} 2 \cdot \underbrace{\det \begin{pmatrix} 2 & 2 & 3 \\ 1 & 4 & 2 \\ 3 & 3 & 2 \end{pmatrix}}_{=:B} - 1 \cdot \underbrace{\det \begin{pmatrix} 1 & 2 & 3 \\ 4 & 4 & 2 \\ 1 & 3 & 2 \end{pmatrix}}_{=:C} + 1 \cdot \underbrace{\det \begin{pmatrix} 1 & 2 & 3 \\ 4 & 1 & 2 \\ 1 & 3 & 2 \end{pmatrix}}_{=:D} - 2 \cdot \underbrace{\det \begin{pmatrix} 1 & 2 & 2 \\ 4 & 1 & 4 \\ 1 & 3 & 3 \end{pmatrix}}_{=:E}$$

Wir berechnen die Determinante von  $B$  mit einer Entwicklung nach der dritten Spalte:

$$\begin{aligned} \det B &= \det \begin{pmatrix} 2 & 2 & 3 \\ 1 & 4 & 2 \\ 3 & 3 & 2 \end{pmatrix} \stackrel{\text{S.3}}{=} 3 \cdot \det \begin{pmatrix} 1 & 4 \\ 3 & 3 \end{pmatrix} - 2 \cdot \det \begin{pmatrix} 2 & 2 \\ 3 & 3 \end{pmatrix} + 2 \cdot \det \begin{pmatrix} 2 & 2 \\ 1 & 4 \end{pmatrix} \\ &= 3(1 \cdot 3 - 4 \cdot 3) - 2(2 \cdot 3 - 2 \cdot 3) + 2(2 \cdot 4 - 2 \cdot 1) \\ &= 3 \cdot (-9) - 2 \cdot 0 + 2 \cdot 6 \\ &= -15 \end{aligned}$$

Die Determinante von  $C$  berechnen wir als Entwicklung nach der zweiten Zeile:

$$\begin{aligned} \det C &= \det \begin{pmatrix} 1 & 2 & 3 \\ 4 & 4 & 2 \\ 1 & 3 & 2 \end{pmatrix} \stackrel{\text{Z.2}}{=} -4 \cdot \det \begin{pmatrix} 2 & 3 \\ 3 & 2 \end{pmatrix} + 4 \cdot \det \begin{pmatrix} 1 & 3 \\ 1 & 2 \end{pmatrix} - 2 \cdot \det \begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix} \\ &= -4(2 \cdot 2 - 3 \cdot 3) + 4(1 \cdot 2 - 3 \cdot 1) - 2(1 \cdot 3 - 2 \cdot 1) \\ &= -4 \cdot (-5) + 4 \cdot (-1) - 2 \cdot 1 \\ &= 14 \end{aligned}$$

Die Determinante von  $D$  berechnen wir mit Laplace nach der ersten Spalte:

$$\begin{aligned} \det D &= \det \begin{pmatrix} 1 & 2 & 3 \\ 4 & 1 & 2 \\ 1 & 3 & 2 \end{pmatrix} \stackrel{\text{S.1}}{=} 1 \cdot \det \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix} - 4 \cdot \det \begin{pmatrix} 2 & 3 \\ 3 & 2 \end{pmatrix} + 1 \cdot \det \begin{pmatrix} 2 & 3 \\ 1 & 2 \end{pmatrix} \\ &= (1 \cdot 2 - 2 \cdot 3) - 4(2 \cdot 2 - 3 \cdot 3) + (2 \cdot 2 - 3 \cdot 1) \\ &= -4 - 4 \cdot (-5) + 1 \\ &= 17 \end{aligned}$$

Die Determinante von  $E$  berechnen wir direkt mit der Regel von Sarrus:

$$\begin{aligned} \det E &= \det \begin{pmatrix} 1 & 2 & 2 \\ 4 & 1 & 4 \\ 1 & 3 & 3 \end{pmatrix} \stackrel{\text{Sarrus}}{=} 1 \cdot 1 \cdot 3 + 2 \cdot 4 \cdot 1 + 2 \cdot 4 \cdot 3 - 2 \cdot 1 \cdot 1 - 2 \cdot 4 \cdot 3 - 1 \cdot 4 \cdot 3 \\ &= 3 + 8 + 24 - 2 - 24 - 12 \\ &= -3 \end{aligned}$$

Nun können wir die Determinante von  $A$  angeben:

$$\det A = 2 \det B - \det C + \det D - 2 \det E = 2 \cdot (-15) - 14 + 17 - 2 \cdot (-3) = -21$$

### 8.3.7 Vereinfachung von Determinantenberechnungen mit Zeilen- und Spaltenoperationen

Wir haben im letzten Beispiel ganz bewusst “rein mechanisch” gerechnet, um die Anwendung der Laplace-Entwicklung zu demonstrieren. Wie im Voraus schon bemerkt, ist es aber hilfreich,

möglichst wenige Unterdeterminanten berechnen zu müssen – also, möglichst viele Nullen in der ursprünglichen Matrix zu erzeugen. Nutzt man dann Laplace, reicht es, nur die Unterdeterminanten mit nichtverschwindenden Skalierungsfaktoren anzuschreiben.

Da mit Laplace nach beliebigen Zeilen oder Spalten entwickelt werden kann, wird es dazu nicht nötig sein, Zeilen oder Spalten in einer gegebenen Matrix zu tauschen. Das Skalieren von Zeilen oder Spalten kann dagegen sinnvoll sein – man achte aber darauf, dies jeweils mit einem Kompensationsfaktor zu berücksichtigen. Die wichtigsten Operationen zum Erzeugen von Nullen sind jedoch die Additionen skalierteter Zeilen (Spalten) zu anderen Zeilen (Spalten) – diese lassen den Wert der Determinanten gleich.

Wir zeigen nun für die im vorigen Beispiel eingeführten Matrizen  $A$  bis  $E$  einige Beispiele, wie die Laplace-Entwicklung verkürzt werden kann. Die Zeilen- und Spalten-Operationen notieren wir dabei genau so wie bei den Zeilen-Operationen von LGS. Wir beginnen mit den dreireihigen Beispielen und zeigen am Schluss, wie auch in der vierreihigen Matrix  $A$  schon deutliche Vereinfachungen möglich sind.

### Beispiele:

- Wir betrachten die Matrix

$$B = \begin{pmatrix} 2 & 2 & 3 \\ 1 & 4 & 2 \\ 3 & 3 & 2 \end{pmatrix}$$

Es fällt auf, dass in der ersten Zeile die ersten beiden Elemente gleich sind – genau wie in der dritten Zeile. Wir skalieren nun die dritte Zeile, addieren eine geeignet skalierte erste Zeile dazu und entwickeln erst dann mit Laplace (und zwar nach der dritten Zeile):

$$\begin{aligned} \det \begin{pmatrix} 2 & 2 & 3 \\ 1 & 4 & 2 \\ 3 & 3 & 2 \end{pmatrix} &= \frac{1}{2} \det \begin{pmatrix} 2 & 2 & 3 \\ 1 & 4 & 2 \\ 6 & 6 & 4 \end{pmatrix} \begin{matrix} \uparrow \\ -3 \end{matrix} = \frac{1}{2} \det \begin{pmatrix} 2 & 2 & 3 \\ 1 & 4 & 2 \\ 0 & 0 & -5 \end{pmatrix} \\ &\stackrel{\text{Z.3}}{=} \frac{1}{2} \cdot (-5) \cdot \det \begin{pmatrix} 2 & 2 \\ 1 & 4 \end{pmatrix} = -\frac{5}{2} \cdot (2 \cdot 4 - 2 \cdot 1) = -\frac{5}{2} \cdot 6 = -\frac{30}{2} = -15 \end{aligned}$$

- Für die selbe Matrix  $B$  könnte man auch erkennen, dass die Einträge in erster und dritter Zeile für die beiden ersten Spalten identisch sind. Wir rechnen also erneut, aber dieses Mal mit einer Spaltenoperation; danach entwickeln wir nach der zweiten Spalte:

$$\det \begin{pmatrix} 2 & 2 & 3 \\ 1 & 4 & 2 \\ 3 & 3 & 2 \end{pmatrix} = \det \begin{pmatrix} 2 & 0 & 3 \\ 1 & 3 & 2 \\ 3 & 0 & 2 \end{pmatrix} \begin{matrix} \text{S.2} \\ -1 \end{matrix} = 3 \cdot \det \begin{pmatrix} 2 & 3 \\ 3 & 2 \end{pmatrix} = 3(4 - 9) = 3 \cdot (-5) = -15$$

- Die Matrix

$$C = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 4 & 2 \\ 1 & 3 & 2 \end{pmatrix}$$

enthält solche Regelmäßigkeiten nicht, aber dafür sind Elemente mit Wert 1 enthalten, die wir z.B. direkt für Zeilenoperationen verwenden können (um danach nach der ersten Spalte zu entwickeln):

$$\det \begin{pmatrix} 1 & 2 & 3 \\ 4 & 4 & 2 \\ 1 & 3 & 2 \end{pmatrix} \begin{matrix} \uparrow \\ -4 \end{matrix} \begin{matrix} \uparrow \\ -1 \end{matrix} = \det \begin{pmatrix} 1 & 2 & 3 \\ 0 & -4 & -10 \\ 0 & 1 & -1 \end{pmatrix} \stackrel{\text{S.1}}{=} \det \begin{pmatrix} -4 & -10 \\ 1 & -1 \end{pmatrix} = 4 - (-10) = 14$$

- Für die Matrix  $D$  verfahren wir ähnlich, aber mit Spaltenoperationen:

$$\det \begin{pmatrix} 1 & 2 & 3 \\ 4 & 1 & 2 \\ 1 & 3 & 2 \end{pmatrix} = \det \begin{pmatrix} 1 & 0 & 0 \\ 4 & -7 & -10 \\ 1 & 1 & -1 \end{pmatrix} \stackrel{\text{Z.1}}{=} \det \begin{pmatrix} -7 & -10 \\ 1 & -1 \end{pmatrix} = 7 - (-10) = 17$$

$\begin{matrix} \curvearrowright \\ -2 \\ \curvearrowright \\ -3 \end{matrix}$

- In Matrix  $E$  gibt es wieder eine günstige Situation wie bei  $B$ . Wir könnten das  $(-1)$ -fache der zweiten Spalte zur dritten addieren, bekämen dort dadurch zwei Nullen und könnten nach der dritten Spalte entwickeln (mit nur einem Beitrag). Das führe man gerne zur Übung aus. Wir zeigen hier statt dessen eine etwas ungeschicktere Methode mit einem zusätzlichen Schritt, wobei eine Spalten- und eine Zeilenoperation auftritt:

$$\det \begin{pmatrix} 1 & 2 & 2 \\ 4 & 1 & 4 \\ 1 & 3 & 3 \end{pmatrix} = \det \begin{pmatrix} 1 & 2 & 1 \\ 4 & 1 & 0 \\ 1 & 3 & 2 \end{pmatrix} \begin{matrix} \curvearrowright \\ -2 \\ \curvearrowleft \end{matrix} = \det \begin{pmatrix} 1 & 2 & 1 \\ 4 & 1 & 0 \\ -1 & -1 & 0 \end{pmatrix}$$

$\begin{matrix} \curvearrowright \\ -1 \end{matrix}$

$$\stackrel{\text{S.3}}{=} \det \begin{pmatrix} 4 & 1 \\ -1 & -1 \end{pmatrix} = -4 - (-1) = -3$$

- Die Matrizen  $B$  bis  $E$  kamen oben als Zwischenergebnisse für die Berechnung der Determinanten von  $A$  auf. Wir zeigen nun, dass auch für  $A$  schon Vereinfachungen möglich sind:

$$\det \begin{pmatrix} 1 & 2 & 2 & 3 \\ 4 & 1 & 4 & 2 \\ 2 & 1 & 1 & 2 \\ 1 & 3 & 3 & 2 \end{pmatrix} = \det \begin{pmatrix} 1 & 0 & 0 & 3 \\ 4 & -7 & -4 & 2 \\ 2 & -3 & -3 & 2 \\ 1 & 1 & 1 & 2 \end{pmatrix} \begin{matrix} \curvearrowright \\ -1 \\ \curvearrowleft \end{matrix} = \det \begin{pmatrix} 1 & 0 & 0 & 3 \\ 4 & -7 & -4 & 2 \\ 2 & -3 & -3 & 2 \\ 0 & 1 & 1 & -1 \end{pmatrix} \begin{matrix} \curvearrowright \\ 3 \end{matrix}$$

$\begin{matrix} \curvearrowright \\ -2 \\ \curvearrowright \\ -2 \end{matrix}$

$$= \det \begin{pmatrix} 1 & 0 & 0 & 3 \\ 4 & -7 & -4 & 2 \\ 2 & 0 & 0 & -1 \\ 0 & 1 & 1 & -1 \end{pmatrix} \begin{matrix} \curvearrowright \\ -2 \\ \curvearrowleft \end{matrix} = \det \begin{pmatrix} 1 & 0 & 0 & 3 \\ 4 & -7 & -4 & 2 \\ 0 & 0 & 0 & -7 \\ 0 & 1 & 1 & -1 \end{pmatrix} \stackrel{\text{Z.3}}{=} -(-7) \cdot \det \begin{pmatrix} 1 & 0 & 0 \\ 4 & -7 & -4 \\ 0 & 1 & 1 \end{pmatrix}$$

$$\stackrel{\text{Z.1}}{=} 7 \cdot \det \begin{pmatrix} -7 & -4 \\ 1 & 1 \end{pmatrix} = 7 \cdot (-7 - (-4)) = 7 \cdot (-3) = -21$$

Im Vergleich zur ursprünglichen Berechnung von  $\det A$  haben wir hier nur jeweils eine dreireihige und eine zweireihige Unterdeterminante berechnen müssen – durchaus eine gewisse Ersparnis.

Mit den Zeilen- und Spalten-Umformungen lassen sich Determinanten auch bestimmen, ohne die Formeln von Leibniz oder Laplace zu benutzen – man verfährt ähnlich wie bei LGS per Gauß-Jordan und erzeugt möglichst viele Nullen. Da hier auch Spaltenoperationen zugelassen sind, gibt es entsprechend mehr Möglichkeiten. Zeilen oder Spalten, die nur einen Eintrag ungleich 0 besitzen, erlauben es über die Skalierungsregel, den Eintrag durch Ausklammern auf 1 zu setzen (das ist für weitere Operationen oft hilfreich).

Die Rechnung bricht entweder ab, wenn sich eine Null-Zeile oder Null-Spalte erzeugen lässt (dann ist die Determinante offenbar 0), oder wenn nur noch die Determinante der Einheitsmatrix zu berechnen wäre – diese ist aber axiomatisch schon als 1 bekannt.

**Beispiel:** Wir berechnen die Determinante der folgenden Matrix:

$$\begin{pmatrix} 2 & 0 & 5 \\ 1 & -4 & 2 \\ 3 & 2 & 4 \end{pmatrix}$$

$$\begin{aligned} \det \begin{pmatrix} 2 & 0 & 5 \\ 1 & -4 & 2 \\ 3 & 2 & 4 \end{pmatrix} &\overset{\substack{\leftarrow -2 \\ \leftarrow -3}}{=} \det \begin{pmatrix} 0 & 8 & 1 \\ 1 & -4 & 2 \\ 0 & 14 & -2 \end{pmatrix} = \det \begin{pmatrix} 0 & 0 & 1 \\ 1 & -20 & 2 \\ 0 & 30 & -2 \end{pmatrix} \\ &\overset{\substack{\leftarrow -8 \\ \leftarrow 20 \\ \leftarrow -2}}{=} \det \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 30 & 0 \end{pmatrix} = \det \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 30 & 0 \end{pmatrix} \overset{\leftarrow}{=} -\det \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 30 & 0 \end{pmatrix} \\ &= \det \begin{pmatrix} 1 & 0 & 0 \\ 0 & 30 & 0 \\ 0 & 0 & 1 \end{pmatrix} = 30 \cdot \det \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = 30 \cdot 1 = 30 \end{aligned}$$

(Mit der Sarrusregel ließe sich der Wert 30 natürlich hier schneller berechnen)

**Beispiel:** Noch zur Anwendung der Determinanteneigenschaften: Anstatt die Summe der beiden folgenden Determinanten (mit Sarrus oder Laplace) direkt zu berechnen, erkennen wir, dass sie sich nur in einer Spalte unterscheiden (der mittleren). Daher können wir die Multilinearität nach Definition 8.16 (in Leserichtung von rechts nach links) ausnutzen und fallen auf die Berechnung nur einer Determinanten zurück:

$$\det \begin{pmatrix} 4 & 1 & 2 \\ 2 & 3 & 1 \\ 2 & 7 & 2 \end{pmatrix} + \det \begin{pmatrix} 4 & 1 & 2 \\ 2 & -2 & 1 \\ 2 & -5 & 2 \end{pmatrix} = \det \begin{pmatrix} 4 & (1+1) & 2 \\ 2 & (3-2) & 1 \\ 2 & (7-5) & 2 \end{pmatrix} = \det \begin{pmatrix} 4 & 2 & 2 \\ 2 & 1 & 1 \\ 2 & 2 & 2 \end{pmatrix}$$

Und diese letzte Determinante ist null, da zwei verschiedene Spaltenvektoren (Mitte und rechts) gleich sind (Alternierungseigenschaft).



### 8.3.8 Determinanten und LGS

Nach Satz 8.25 gilt, dass (nur genau) für quadratische Matrizen  $A \in \mathbb{R}^{(n,n)}$  mit nichtverschwindender Determinante deren Zeilen und Spalten linear unabhängig sind.

Für solche Matrizen folgt mit Satz 8.10, dass sie vollen Rang  $n$  besitzen. Damit ist jedes LGS  $(A \mid \vec{y})$  eindeutig lösbar; der Kern von  $A$  ist dann nach dem Rangsatz (8.9) trivial, enthält also nur den Nullvektor  $\vec{0}$ .

Insbesondere ist dann auch das homogene LGS  $(A \mid \vec{0})$  eindeutig lösbar.

---

Dagegen sind bei verschwindender Determinante die Zeilen und Spalten von  $A$  linear abhängig. Dann ist  $\text{rg } A < n$ , und das homogene LGS  $(A \mid \vec{0})$  ist mehrdeutig lösbar; der Kern von  $A$  ist nichttrivial, und seine Dimension ergibt sich nach dem Rangsatz als  $n - \text{rg } A$ .

---

Für  $\det A \neq 0$  (also eindeutig lösbare LGS mit quadratischer Koeffizientenmatrix) lässt sich die Lösung  $\vec{x}$  für das LGS  $(A \mid \vec{y})$  mit einer geschlossenen Formel (*Regel von Cramer*) durch Determinantenberechnungen bestimmen; siehe dazu Exkurs A.8.

### 8.3.9 Determinanten-Produktsatz

Zum Abschluss des Abschnitts über Determinanten stellen wir noch eine wichtige Formel auf, die die Determinante eines Produkts zweier quadratischer Matrizen angibt:

**Satz 8.27** (Determinanten-Produktsatz). Für  $A, B \in \mathbb{R}^{(n,n)}$  ist die Determinante der Produkte  $AB$  und  $BA$  gegeben durch

$$\det(AB) = \det(BA) = (\det A) \cdot (\det B)$$

(Beweis: S. 346.)

#### Bemerkungen:

- Da auf der rechten Seite das Produkt aus den Determinanten der beiden Faktoren auftritt, und da (nach Satz 8.22)  $\det A^T = \det A$  (für  $B$  entsprechend), haben auch sämtliche Produkte aus  $A$  und  $B$ , bei denen einer oder beide der Faktoren transponiert sind, die gleiche Determinante.
- Falls wir zwei Permutationsmatrizen zu  $\pi, \sigma \in S_n$  betrachten, bestätigt die Formel den Satz 8.15 (Vorzeichen von Permutationen bei Gruppenoperationen). Mit Satz 7.27 (Komposition von Permutationsmatrizen) gilt:

$$\text{sign}(\pi \circ \sigma) = \det P_{\pi \circ \sigma} = \det (P_\pi \cdot P_\sigma) = (\det P_\pi) \cdot (\det P_\sigma) = \text{sign}(\pi) \cdot \text{sign}(\sigma) \quad \checkmark$$

**Beispiel:** Wir betrachten die dreireihigen quadratischen Matrizen

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

Für die Matrix  $A$  berechnen wir die Determinante, indem wir zunächst die mittlere Zeile (mit  $(-1)$  skaliert) jeweils auf die obere und die untere Zeile addieren; danach lässt sich mit Laplace nach der zweiten Spalte entwickeln:

$$\det \begin{pmatrix} 2 & 1 & 3 \\ 1 & 1 & 0 \\ -2 & 1 & -2 \end{pmatrix} \begin{matrix} \nwarrow -1 \\ \nearrow -1 \end{matrix} = \det \begin{pmatrix} 1 & 0 & 3 \\ 1 & 1 & 0 \\ -3 & 0 & -2 \end{pmatrix} \stackrel{\text{S.2}}{=} \det \begin{pmatrix} 1 & 3 \\ -3 & -2 \end{pmatrix} = -2 - (-9) = 7$$

Bei Matrix  $B$  lässt sich direkt nach der dritten Spalte entwickeln:

$$\det B \stackrel{\text{S.3}}{=} \det \begin{pmatrix} 4 & -1 \\ 2 & 1 \end{pmatrix} = 4 - (-2) = 6$$

Das Produkt  $AB$  berechnet sich zu

$$AB = \begin{pmatrix} 13 & 5 & 3 \\ 6 & 0 & 0 \\ -8 & -1 & -2 \end{pmatrix}$$

Die Determinante davon ist (man beachte den Vorfaktor  $(-1)$  bei der Laplace-Entwicklung):

$$\det(AB) \stackrel{\text{Z.2}}{=} -6 \cdot \det \begin{pmatrix} 5 & 3 \\ -1 & -2 \end{pmatrix} = -6 \cdot ((-10) - (-3)) = -6 \cdot (-7) = 42$$

Das Produkt  $BA$  berechnet sich zu

$$BA = \begin{pmatrix} 7 & 3 & 12 \\ 5 & 3 & 6 \\ 2 & 4 & 1 \end{pmatrix}$$

Für die Berechnung von  $\det(BA)$  verwenden wir die Sarrus-Regel:

$$\det(BA) \stackrel{\text{Sarrus}}{=} 21 + 36 + 240 - 72 - 15 - 168 = 42$$

Wir erhalten also in der Tat für beide Matrizenprodukte eine Determinante von

$$(\det A) \cdot (\det B) = 7 \cdot 6 = 42$$

## Kapitel 9

# Lineare Algebra: Matrizen Revisited

Hier besprechen wir zunächst einige Themen, die sich mit dem Wissen über LGS und Determinanten leichter handhaben lassen bzw. die von oben noch ausstehen (speziell zu den inversen Matrizen und den orthogonalen Matrizen, die aus Kapitel 7 bereits bekannt sind).

Danach befassen wir uns mit dem *Eigenwertproblem*, gewissermaßen der “Krönung” der linearen Algebra – bei dessen Lösung kommen (abgesehen von den Restklassen) alle gelernten Konzepte zusammen: Vektoren, Matrizen, LGS, Determinanten, Polynome und das Lösen algebraischer Gleichungen.

Eine wichtige Anwendung des Eigenwertproblems findet sich beim Diagonalisieren von Matrizen.

### 9.1 Invertierung reeller quadratischer Matrizen

Wie im vorigen Kapitel bereits fest gehalten, dass eindeutig lösbare LGS mit quadratischer Koeffizientenmatrix schreibbar sind (bei überbestimmten LGS stellt sich, falls sie eindeutig lösbar sind, eine solche Form immerhin während der Rechnung ein – darauf gehen wir hier nicht weiter ein); die Spalten und Zeilen solcher Matrizen sind jeweils linear unabhängig, die Matrizen haben vollen Rang und ihre Determinanten sind ungleich 0.

Für solche Matrizen  $A \in \mathbb{R}^{(n,n)}$  ist also die Gleichung

$$A\vec{x} = \vec{y}$$

eindeutig lösbar – sie beschreibt die Abbildung eines Urbildvektors  $\vec{x}$  auf einen Bildvektor  $\vec{y}$ . Dann existiert aber auch die Umkehrabbildung, über die aus  $\vec{y}$  wieder das Urbild  $\vec{x}$  berechnet werden kann. Nach Satz 7.20 (Bijektive lineare Abbildungen) entspricht die Matrix der fraglichen Umkehrabbildung gerade der Inversen  $A^{-1}$ . Also:

$$\dots \Leftrightarrow \vec{x} = A^{-1}\vec{y}$$

Beim Lösen eines eindeutig lösbaren LGS mit Koeffizientenmatrix  $A$  berechnen wir also *implizit* die Inverse  $A^{-1}$  – jedoch direkt gekoppelt mit ihrer Anwendung auf den Inhomogenitätsvektor  $\vec{y}$ . Wir lernen gleich, wie man die Inverse von  $A$  *explizit* berechnet. Zunächst betrachten wir aber in Ergänzung zu Kapitel 7 noch einige Eigenschaften invertierbarer Matrizen.

#### 9.1.1 Eigenschaften inverser Matrizen (Vervollständigung)

**Definition 9.1** (Reguläre und Singuläre Matrizen). *Eine quadratische Matrix  $A \in \mathbb{R}^{(n,n)}$  heißt regulär, falls sie vollen Rang  $n$  besitzt. Falls  $\text{rg } A < n$ , heißt  $A$  singulär.*

Damit halten wir die bisherigen Überlegungen zu invertierbaren Matrizen fest:

**Satz 9.2** (Invertierbare Matrizen). *Folgende Aussagen sind für quadratische Matrizen  $A \in \mathbb{R}^{(n,n)}$  äquivalent:*

- $A$  ist invertierbar nach Definition 7.17
- $\det A \neq 0$
- Die Spalten von  $A$  sind linear unabhängig (und bilden eine Basis von  $\mathbb{R}^n$ )
- Die Zeilen von  $A$  sind linear unabhängig (ihre Transponierten bilden eine Basis von  $\mathbb{R}^n$ )
- $\operatorname{rg} A = n$
- Der Kern von  $A$  ist trivial (enthält also nur  $\vec{0}$ )
- Das LGS  $(A \mid \vec{y})$  ist eindeutig lösbar für beliebiges  $\vec{y}$ ; die Zeilen-Stufen-Form enthält  $n$  Zeilen

**Bemerkung:** In Exkurs A.6 findet sich weiter führendes Material zum Basiswechsel mit Matrizen; dabei spielt auch eine Rolle, dass die Spalten einer invertierbaren Matrix eine Basis bilden.

Wir hatten in Satz 7.18 (Eigenschaften inverser Matrizen) bereits einige Eigenschaften invertierbarer Matrizen kennen gelernt. Mit dem Determinanten-Produktsatz (Satz 8.27) können wir noch ein wichtiges Detail hinzu fügen. Und zwar gilt für eine invertierbare Matrix  $A$ :

$$(A^{-1}A = \mathbb{1}_n) \Rightarrow (\det(A^{-1}A) = (\det A^{-1}) \cdot (\det A) = \det \mathbb{1}_n = 1)$$

Auch dies halten wir fest:

**Satz 9.3** (Determinante der Inversen). *Für eine invertierbare Matrix  $A \in \mathbb{R}^{(n,n)}$  gilt:*

$$\det A^{-1} = \frac{1}{\det A}$$

**Beispiel:** Wir betrachten die dreireihigen Matrizen aus den Beispielen zu Satz 7.18

$$A := \begin{pmatrix} 1 & 2 & 1 \\ 2 & -2 & 1 \\ 2 & 2 & 2 \end{pmatrix} \quad \text{und} \quad B := \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 2 \\ 1 & 2 & 1 \end{pmatrix}$$

mit den Inversen

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

Die Determinanten berechnen wir hier alle mit der Sarrus-Regel. Zur Übung rechne man auch nochmal mit Zeilen-/Spalten-Operationen und Laplace-Entwicklung (man kann bei allen vier Rechnungen mit je einer Zeilen-/Spalten-Operation eine Zeile oder Spalte mit zwei Nullen erzeugen, sodass dann nur eine zweireihige Unterdeterminante zu berechnen ist).

$$\begin{aligned} \det A &\stackrel{\text{Sarrus}}{=} -4 + 4 + 4 - (-4) - 8 - 2 = -2 \\ \det B &\stackrel{\text{Sarrus}}{=} 1 + 2 + 2 - 1 - 1 - 4 = -1 \\ \det A^{-1} &\stackrel{\text{Sarrus}}{=} 0 + 2 + 2 - 0 - 3 - \frac{3}{2} = -\frac{1}{2} = \frac{1}{-2} \quad \checkmark \\ \det B^{-1} &\stackrel{\text{Sarrus}}{=} 0 + 1 + 1 - 0 - 0 - 3 = -1 = \frac{1}{-1} \quad \checkmark \end{aligned}$$

### 9.1.2 Berechnung der Inversen mit Gauß

Statt der Vektorgleichung  $A\vec{x} = \vec{y}$  haben wir nun die (immerhin ähnlich aussehende) Matrixgleichung

$$A \cdot X = \mathbb{1}_n$$

zu lösen – wir erhalten dann als Lösung  $X$  die Rechtsinverse von  $A$ , die aber nach Satz 7.19 (Links- und Rechtsinvertierbarkeit) auch der Linksinversen entspricht, also insgesamt die Inverse von  $A$  ist.

Zunächst erinnern wir uns an die Spaltendarstellung des Matrixprodukts. Wenn die Spalten von  $X$  gegeben sind als

$$X = (\vec{x}_1 \quad \cdots \quad \vec{x}_n),$$

dann ist nach Satz 7.10:

$$A \cdot X = A \cdot (\vec{x}_1 \quad \cdots \quad \vec{x}_n) = (A\vec{x}_1 \quad \cdots \quad A\vec{x}_n)$$

Und nun setzen wir die rechte Seite der obigen Matrixgleichung ein – es ist also:

$$A \cdot X = (A\vec{x}_1 \quad \cdots \quad A\vec{x}_n) = \mathbb{1}_n = (\vec{e}_1 \quad \cdots \quad \vec{e}_n)$$

Somit zerfällt die Matrixgleichung in  $n$  Vektorgleichungen der Form

$$A\vec{x}_j = \vec{e}_j$$

Diese  $n$  Gleichungen sind simultan zu lösen für alle  $j \in \{1, \dots, n\}$ . Es handelt sich um  $n$  inhomogene LGS, die wir mit den bekannten Methoden aus Kapitel 8, im Speziellen mit Gauß-Umformungen, lösen können.

---

Da aber die Koeffizientenmatrix all dieser LGS gleich ist, lohnt es sich natürlich nicht, die  $n$  LGS hintereinander zu lösen – die Operationen, um die Zeilen-Stufen-Form herzustellen, sind ja stets die gleichen. Vielmehr erweitern wir das Gauß-Schema derart, dass wir alle  $n$  verschiedenen Inhomogenitäten (das sind Spalten von  $\mathbb{1}_n$ , also die Einheitsvektoren der Standardbasis  $E_n$ ) nebeneinander notieren und gleichzeitig den selben Zeilenoperationen unterwerfen. Dadurch brauchen wir nur genau einmal die Zeilen-Stufen-Form herzustellen. Beispiele folgen in Kürze!

So, wie bei einer gewöhnlichen Vektor-Gleichung für ein einzelnes LGS rechts vom Trennstrich anfangs die Inhomogenität steht, am Ende der Rechnung aber die spezielle Lösung, so erhalten wir mit dieser Methode alle  $n$  speziellen Lösungen simultan, wenn wir durch Zeilen-Operationen im (linken) Koeffiziententeil des Gauß-Schemas die Einheitsmatrix  $\mathbb{1}_n$  hergestellt haben.

---

Um zu beurteilen, ob solch ein Inverses  $A^{-1}$  existiert, könnte man nun im Voraus die Determinante von  $A$  berechnen. Die hierfür nötigen Operationen sind aber nicht ohne Aufwand ausführbar, und mit der Determinanten ist dann eben auch nur eine Aussage über die Existenz getroffen.

Wenn also nach der Inversen gesucht ist, lohnt es sich meistens, direkt (und optimistisch) anzunehmen, dass diese existiert, und die entsprechenden Umformungen des erweiterten Gauß-Schemas zu rechnen. Falls sich dabei eine Zeilen-Stufen-Form mit vermindertem Rang einstellen sollte, braucht man nicht weiter zu rechnen – falls jedoch voller Rang vorliegt, wird entsprechend der Kern von  $A$  trivial sein, sodass die spezielle Lösung (welche am Ende rechts vom Trennstrich steht) auch die einzige Lösung darstellt. Mit dieser Methode hätte man Eindeutigkeit und Existenz zusammen gezeigt.

Dazu ist es allerdings nötig, den linken Teil des Gauß-Schemas nicht nur auf Zeilen-Stufen-Form zu bringen, sondern voll bis zur Einheitsmatrix umzuformen, d.h.

$$(A \mid \mathbb{1}_n) \Leftrightarrow \cdots \Leftrightarrow (\mathbb{1}_n \mid A^{-1})$$

Auch bei diesem erweiterten Gauß-Schema sieht man, dass während der Rechnung effektiv  $A^{-1}$  berechnet wird – nur dass diese Matrix aufgrund der besonderen Inhomogenitäten danach tatsächlich spaltenweise ablesbar ist, also explizit vorliegt. Effektiv hat man dann während der Umformungen beide Teile des Gauß-Schemas von links mit  $A^{-1}$  multipliziert

---

**Beispiele:** (Erinnerung: Bei Gauß-LGS sind zum Lösen nur Zeilenoperationen erlaubt!)

- Wir berechnen zu den beiden Matrizen aus den Beispielen zu Satz 9.3 die Inversen; zunächst für  $A$ :

$$\begin{aligned}
& \left( \begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 2 & -2 & 1 & 0 & 1 & 0 \\ 2 & 2 & 2 & 0 & 0 & 1 \end{array} \right) \begin{array}{l} \curvearrowright -2 \\ \curvearrowleft \\ \curvearrowleft \end{array} \Leftrightarrow \left( \begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & -6 & -1 & -2 & 1 & 0 \\ 0 & -2 & 0 & -2 & 0 & 1 \end{array} \right) \begin{array}{l} \curvearrowleft -3 \\ \curvearrowleft \\ \curvearrowleft 1 \end{array} \\
& \Leftrightarrow \left( \begin{array}{ccc|ccc} 1 & 0 & 1 & -1 & 0 & 1 \\ 0 & 0 & -1 & 4 & 1 & -3 \\ 0 & -2 & 0 & -2 & 0 & 1 \end{array} \right) \begin{array}{l} \curvearrowleft 1 \\ \curvearrowleft \\ -\frac{1}{2} \end{array} \Leftrightarrow \left( \begin{array}{ccc|ccc} 1 & 0 & 0 & 3 & 1 & -2 \\ 0 & 0 & -1 & 4 & 1 & -3 \\ 0 & 1 & 0 & 1 & 0 & -\frac{1}{2} \end{array} \right) -1 \\
& \Leftrightarrow \left( \begin{array}{ccc|ccc} 1 & 0 & 0 & 3 & 1 & -2 \\ 0 & 0 & 1 & -4 & -1 & 3 \\ 0 & 1 & 0 & 1 & 0 & -\frac{1}{2} \end{array} \right) \begin{array}{l} \curvearrowleft \\ \curvearrowleft \\ \curvearrowleft \end{array} \Leftrightarrow \left( \begin{array}{ccc|ccc} 1 & 0 & 0 & 3 & 1 & -2 \\ 0 & 1 & 0 & 1 & 0 & -\frac{1}{2} \\ 0 & 0 & 1 & -4 & -1 & 3 \end{array} \right)
\end{aligned}$$

Da im linken Teil des Gauß-Schemas genau die Einheitsmatrix  $\mathbb{1}_3$  steht, entspricht der rechte Teil gerade der Inversen  $A^{-1}$ .

- Für  $B$ :

$$\begin{aligned}
& \left( \begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 2 & 0 & 1 & 0 \\ 1 & 2 & 1 & 0 & 0 & 1 \end{array} \right) \begin{array}{l} \curvearrowright -1 \\ \curvearrowleft \\ \curvearrowleft \end{array} -1 \Leftrightarrow \left( \begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 & 1 & 0 \\ 0 & 1 & 0 & -1 & 0 & 1 \end{array} \right) \begin{array}{l} \curvearrowleft \\ \curvearrowleft \\ \curvearrowleft -1 \end{array} \\
& \Leftrightarrow \left( \begin{array}{ccc|ccc} 1 & 0 & 1 & 2 & 0 & -1 \\ 0 & 0 & 1 & -1 & 1 & 0 \\ 0 & 1 & 0 & -1 & 0 & 1 \end{array} \right) \begin{array}{l} \curvearrowleft -1 \\ \curvearrowleft \\ \curvearrowleft \end{array} \Leftrightarrow \left( \begin{array}{ccc|ccc} 1 & 0 & 0 & 3 & -1 & -1 \\ 0 & 0 & 1 & -1 & 1 & 0 \\ 0 & 1 & 0 & -1 & 0 & 1 \end{array} \right) \begin{array}{l} \curvearrowleft \\ \curvearrowleft \\ \curvearrowleft \end{array} \\
& \Leftrightarrow \left( \begin{array}{ccc|ccc} 1 & 0 & 0 & 3 & -1 & -1 \\ 0 & 1 & 0 & -1 & 0 & 1 \\ 0 & 0 & 1 & -1 & 1 & 0 \end{array} \right)
\end{aligned}$$

Auch hier ist der rechte Teil des letzten Gauß-Schemas gerade die Inverse  $B^{-1}$ .

- Für zweireihige Matrizen wollen wir noch allgemein die Inverse berechnen, genau wie oben für die Determinante im Beispiel bei Satz 8.17.

Hier wäre die Regel von Cramer aus Exkurs A.8 sehr hilfreich; sie liefert uns die Komponenten der inversen Matrix direkt und ohne weitere Fallunterscheidung. Diese Rechnung ist im Exkurs als Beispiel ausgeführt.

Wir rechnen hier statt dessen gemäß Überschrift mit Gauß. Zu lösen ist also für

$$M := \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

die Matrixgleichung

$$(M \mid \mathbb{1}_2) \Leftrightarrow \left( \begin{array}{cc|cc} a & b & 1 & 0 \\ c & d & 0 & 1 \end{array} \right),$$

wobei  $\det M = ad - bc \neq 0$  voraus gesetzt wird.

Problematisch für die allgemeine Rechnung ist, dass beim Skalieren von Gleichungen nur Faktoren ungleich 0 zulässig sind.

Wir führen unsere Rechnung zunächst unter der Annahme  $c \neq 0$  durch. Die Skalierung der ersten Zeile mit  $c$  liefert dann (die Kommentare zum Umformen notieren wir hier als Text, da teilweise längere Begründungen nötig sind):

$$\left( \begin{array}{cc|cc} a & b & 1 & 0 \\ c & d & 0 & 1 \end{array} \right) \Leftrightarrow \left( \begin{array}{cc|cc} ac & bc & c & 0 \\ c & d & 0 & 1 \end{array} \right)$$

Nun addieren wir das  $(-a)$ -fache der zweiten Zeile zur ersten:

$$\dots \Leftrightarrow \left( \begin{array}{cc|cc} 0 & bc-ad & c & -a \\ c & d & 0 & 1 \end{array} \right)$$

Dieses Schema ist auch verträglich mit dem Fall  $a = 0$ , denn dann wäre einfach das vorige Schema numerisch unverändert geblieben (das Addieren einer mit 0 skalierten Gleichung zu einer anderen ändert ein LGS nicht).

Wir skalieren nun die zweite Zeile mit  $\det M = ad - bc \neq 0$ :

$$\dots \Leftrightarrow \left( \begin{array}{cc|cc} 0 & bc-ad & c & -a \\ c(ad-bc) & d(ad-bc) & 0 & ad-bc \end{array} \right)$$

Nun addieren wir das  $d$ -fache der ersten Zeile zur zweiten:

$$\dots \Leftrightarrow \left( \begin{array}{cc|cc} 0 & bc-ad & c & -a \\ c(ad-bc) & 0 & cd & -bc \end{array} \right)$$

Dieses Schema ist (analog zu eben) verträglich mit dem Fall  $d = 0$ .

Nun können wir die erste Zeile mit  $(-1)$  skalieren und die zweite mit  $\frac{1}{c} \neq 0$ . Danach vertauschen wir noch beide Zeilen:

$$\dots \Leftrightarrow \left( \begin{array}{cc|cc} 0 & ad-bc & -c & a \\ ad-bc & 0 & d & -b \end{array} \right) \Leftrightarrow \left( \begin{array}{cc|cc} ad-bc & 0 & d & -b \\ 0 & ad-bc & -c & a \end{array} \right)$$

Wir haben nun im linken Teil des Gauß-Schemas das  $(\det M)$ -fache der Einheitsmatrix  $\mathbb{1}_2$ . Entsprechend folgt für die inverse Matrix:

$$\dots \Leftrightarrow \boxed{M^{-1} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \frac{1}{\det M} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}}$$

Nun könnten wir noch gesondert den Fall  $c = 0$  behandeln, den wir vorhin ausgeschlossen hatten. Man führe dies zur Übung gerne selbst durch (dabei wäre wichtig, dass dann  $\det M = ad$ , und da dies ungleich 0 sein muss, damit  $M$  invertierbar ist, müssen auch  $a$  und  $d$  ungleich 0 sein – dies benötigt man zum Skalieren der Gleichungen).

Es stellt sich aber heraus, dass die hier gefundene Form von  $M^{-1}$  auch für  $c = 0$  richtig ist, denn wir können allgemein (also für beliebiges  $c$ ) die Probe rechnen:

$$M^{-1}M = \frac{1}{\det M} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \cdot \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \frac{1}{\det M} \begin{pmatrix} ad-bc & bd-bd \\ -ac+ac & -bc+ad \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \mathbb{1}_2$$

Also ist schon die allgemeine Inverse zu  $M$  gefunden, unabhängig vom Wert von  $c$ .

- Wir überprüfen die Formel für ein Zahlenbeispiel, dessen  $(2, 1)$ -Komponente tatsächlich null ist:

$$N := \begin{pmatrix} 3 & 4 \\ 0 & -2 \end{pmatrix}$$

Dann ist  $\det N = -6$ , und nach Formel erhalten wir:

$$N^{-1} = -\frac{1}{6} \begin{pmatrix} -2 & -4 \\ 0 & 3 \end{pmatrix}$$

Wir rechnen die beiden Produkte:

$$N^{-1}N = -\frac{1}{6} \begin{pmatrix} -2 & -4 \\ 0 & 3 \end{pmatrix} \cdot \begin{pmatrix} 3 & 4 \\ 0 & -2 \end{pmatrix} = -\frac{1}{6} \begin{pmatrix} -6 & 0 \\ 0 & -6 \end{pmatrix} = \mathbb{1}_2 \quad \checkmark$$

$$NN^{-1} = -\frac{1}{6} \begin{pmatrix} 3 & 4 \\ 0 & -2 \end{pmatrix} \cdot \begin{pmatrix} -2 & -4 \\ 0 & 3 \end{pmatrix} = -\frac{1}{6} \begin{pmatrix} -6 & 0 \\ 0 & -6 \end{pmatrix} = \mathbb{1}_2 \quad \checkmark$$

### 9.1.3 Die lineare Gruppe

Wir hatten in Satz 7.12 schon fest gehalten, dass die quadratischen Matrizen  $\mathbb{R}^{(n,n)}$  jeweils einen (nicht-kommutativen) Ring mit Eins bilden, mit dem Matrixprodukt als Multiplikation. Die Körper-eigenschaft liegt allerdings nicht vor – selbst wenn wir uns nur auf die invertierbaren Matrizen beschränken –, denn das Produkt ist nicht kommutativ.

Allerdings können wir für die invertierbaren Matrizen zumindest eine multiplikative Gruppenstruktur erkennen. Denn nach Satz 9.2 haben alle invertierbaren Matrizen jeweils eine Determinante ungleich 0, und nach dem Determinanten-Produktsatz (8.27) und dem Satz (5.16) vom Nullprodukt hat dann auch ein Produkt von invertierbaren Matrizen eine Determinante ungleich 0 und ist somit wieder eine invertierbare Matrix; damit ist die Abgeschlossenheit des Produkts sicher gestellt:

**Satz 9.4** (Lineare Gruppe). *Die Menge*

$$GL(n) := \left\{ A \in \mathbb{R}^{(n,n)} \mid \det A \neq 0 \right\}$$

*bildet mit dem Matrixprodukt eine Gruppe, die lineare Gruppe.*

**Bemerkungen:**

- Die Abkürzung kommt vom englischen Begriff “general linear group”.
- $GL(n)$  ist nicht abelsch.

---

Eine wichtige Untergruppe von  $GL(n)$  ist dabei die Menge der invertierbaren Matrizen, die Determinante 1 haben. Da  $1 \cdot 1 = 1$  überträgt sich diese Eigenschaft beim Produkt zweier solcher Matrizen per Determinanten-Produktsatz auch auf das Produkt, sodass auch hier eine abgeschlossene Operation vorliegt:

**Satz 9.5** (Spezielle lineare Gruppe). *Die Menge*

$$SL(n) := \{ A \in GL(n) \mid \det A = 1 \}$$

*bildet mit dem Matrixprodukt eine Untergruppe von  $GL(n)$ , die spezielle lineare Gruppe.*

## 9.2 Orthogonale Matrizen (Fortsetzung)

### 9.2.1 Eigenschaften

Wir hatten im Kapitel 7 bereits die meisten Eigenschaften orthogonaler Matrizen kennen gelernt. Dies sind die quadratischen Matrizen, deren Inverse durch ihre jeweiligen Transponierten gegeben sind. Alternativ (Satz 7.24) ist eine Matrix genau dann orthogonal, wenn ihre Spaltenvektoren normiert sind und paarweise orthogonal zueinander sind (für die transponierten Zeilen gilt das gleiche).

Da die Inversen orthogonaler Matrizen stets existieren, müssen deren Determinanten nach Satz 9.2 alle ungleich null sein; ihre Spalten (und auch ihre transponierten Zeilen) bilden *Orthonormalbasen* (ONB) des  $\mathbb{R}^n$ . Auch die Umkehrung gilt, denn jede Matrix mit  $n$  orthonormalen Vektoren aus  $\mathbb{R}^n$  als Spalten ist nach Satz 7.24 orthogonal.

---

Was die Determinanten orthogonaler Matrizen angeht, können wir aber noch nachschärfen. Denn nach Satz 8.22 ist  $\det A^T = \det A$ , und für  $A^T = A^{-1}$  gilt dann mit dem Determinanten-Produktsatz für orthogonale Matrizen  $A$  auch:

$$(\det A)^2 = (\det A) \cdot (\det A) = (\det A^T) \cdot (\det A) = \det (A^T \cdot A) = \det (A^{-1} \cdot A) = \det (\mathbb{1}_n) = 1$$

Also gilt folgender

**Satz 9.6** (Determinanten orthogonaler Matrizen). *Für alle orthogonalen Matrizen  $A \in \mathbb{R}^{(n,n)}$  gilt:*

$$\det A \in \{-1, 1\}$$



**Bemerkung:** Man beachte, dass die Umkehrung *nicht* gilt. Eine quadratische Matrix mit Determinante  $\pm 1$  ist nicht zwingend orthogonal! Beispiel: Die Matrix

$$M := \begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix}$$

hat Determinante  $\det M = 3 - 2 = 1$ , aber

$$M^T \cdot M = \begin{pmatrix} 1 & 1 \\ 2 & 3 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix} = \begin{pmatrix} 2 & 5 \\ 5 & 13 \end{pmatrix} \neq \mathbb{1}_2$$

Der obige Satz ist eine Implikation, aber keine Äquivalenz; die Eigenschaft  $\det A = \pm 1$  ist also *notwendig, jedoch nicht hinreichend*, für orthogonale Matrizen.

### 9.2.2 Drehmatrizen

Wir hatten in Kapitel 7 bereits gesehen, dass Permutationsmatrizen (Definition 7.25) orthogonal sind. Eine andere sehr wichtige Klasse orthogonaler Matrizen bilden die *Drehmatrizen* – das sind die Abbildungsmatrizen für Drehungen um beliebig orientierte Ursprungsgeraden (hier als Drehachsen bezeichnet).

### 9.2.3 Aktive Drehungen in $\mathbb{R}^2$

Der einfachste Fall einer Drehung ist in  $\mathbb{R}^2$  darstellbar – die einzig sinnvolle Drehachse steht senkrecht auf der  $x_1, x_2$ -Ebene. Wir denken uns also eine  $x_3$ -Achse mit dazu, die aus der Zeichenebene heraus ragt (und auf die betrachtende Person gerichtet ist); dies soll die Drehachse sein.

Zunächst betrachten wir die *aktive* Drehung. Bei solchen Drehungen bleiben die Koordinatenachsen fest; durch Multiplikation mit der Abbildungsmatrix ändern sich jedoch für beliebige Vektoren die Koordinaten bezüglich diesem festen Koordinatensystem.

---

Die Spalten der Abbildungsmatrix  $D_\alpha \in \mathbb{R}^{(2,2)}$ , welche die Drehung um den Drehwinkel  $\alpha$  in mathematisch positivem Umlaufsinn (gegen den Uhrzeigersinn) um die Drehachse  $x_3$  beschreibt, ergeben sich nach Definition 7.3 (Abbildungsmatrix) als die Abbildungen der kartesischen Einheitsvektoren. Abbildung 9.1 zeigt die Situation. Eine Drehung mit positivem Winkel lässt sich mit der *Rechte-Hand-Regel* beschreiben: Formt man mit rechts eine Faust und spreizt danach den Daumen ab, so zeigt letzterer in Richtung der Drehachse. Die übrigen Finger geben den (positiven) Drehsinn an.

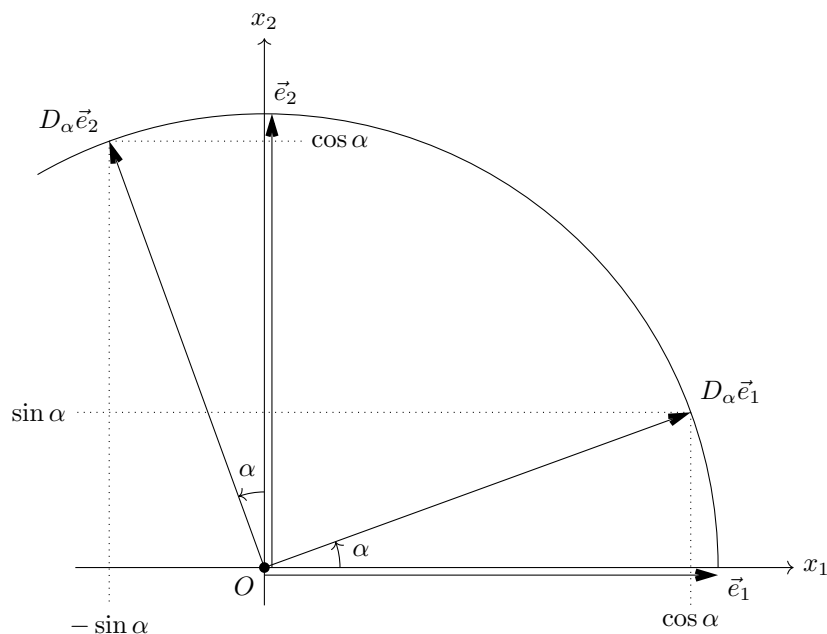


Abbildung 9.1: Herleitung der aktiven Drehung um den Winkel  $\alpha$  in  $\mathbb{R}^2$

Mit den Koordinaten, die in Abbildung 9.1 abgetragen sind (und die sich aus rechtwinkligen Dreiecken wie in Abbildung 4.2.3 (S. 99) ergeben, wobei die Hypotenuse jeweils dem Kreistradius mit Länge 1 entspricht), folgt für die Spalten der Drehmatrix:

$$\vec{d}_1 = D_\alpha \vec{e}_1 = \begin{pmatrix} \cos \alpha \\ \sin \alpha \end{pmatrix}; \quad \vec{d}_2 = D_\alpha \vec{e}_2 = \begin{pmatrix} -\sin \alpha \\ \cos \alpha \end{pmatrix}$$

Damit ergibt sich die gesuchte Matrix als:

$$D_\alpha := \begin{pmatrix} \vec{d}_1 & \vec{d}_2 \end{pmatrix} = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}$$


---

Wir überprüfen, ob es sich bei  $D_\alpha$  tatsächlich um eine orthogonale Matrix handelt: Dass die beiden Spalten orthogonal zueinander sind, lässt sich direkt ablesen (kanonisches Skalarprodukt). Sie sind jedoch zusätzlich auch normiert, denn mit der euklidischen Norm haben beide Spalten die Länge

$$\sqrt{\cos^2 \alpha + \sin^2 \alpha}$$

Der Vergleich mit Abbildung 4.2.3 und die Anwendung des *Satzes von Pythagoras* ergeben, dass stets

$$\boxed{\cos^2 \alpha + \sin^2 \alpha = 1}$$

gilt (dieser Zusammenhang ist deswegen auch als *trigonometrischer Pythagoras* bekannt).

Damit ist  $D_\alpha$  nach Satz 7.24 also in der Tat (und für jeden beliebigen Drehwinkel  $\alpha$ ) eine orthogonale Matrix.

---

Die Determinante von  $D_\alpha$  ist, wie man direkt nachrechnen kann, über den trigonometrischen Pythagoras ebenfalls als +1 gegeben, unabhängig vom Drehwinkel.

---

Die Inverse der Drehmatrix  $D_\alpha$  entspricht nach Satz 7.20 (Bijektive lineare Abbildungen) genau der Umkehrabbildung – das ist hier eine Drehung um den Winkel  $(-\alpha)$ . Ihre Abbildungsmatrix ist aufgrund der Orthogonalität gegeben als

$$D_{-\alpha} = \begin{pmatrix} \cos(-\alpha) & -\sin(-\alpha) \\ \sin(-\alpha) & \cos(-\alpha) \end{pmatrix} = D_\alpha^{-1} = D_\alpha^T = \begin{pmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{pmatrix}$$

Auch hier handelt es sich um eine (orthogonale) Drehmatrix mit Determinante 1. Der Vergleich der Matrixelemente in obiger Gleichung zeigt nochmals (siehe Kapitel 4), dass der Cosinus eine gerade und der Sinus eine ungerade Funktion ist.

---

Auch eine weitere Eigenschaft der trigonometrischen Funktionen können wir hier zeigen. Wir betrachten die Hintereinanderausführung von zwei Drehungen um die gleiche Achse, mit Winkeln  $\alpha$  und  $\beta$ . Die Gesamtdrehung ist dann eine Drehung um den Winkel  $(\alpha + \beta)$  mit der Drehmatrix

$$D_{\alpha+\beta} = \begin{pmatrix} \cos(\alpha + \beta) & -\sin(\alpha + \beta) \\ \sin(\alpha + \beta) & \cos(\alpha + \beta) \end{pmatrix}$$

Wir können aber nach Satz 7.9 (Abbildungsmatrizen bei Komposition) die gleiche Matrix auch berechnen als

$$\begin{aligned} D_{\alpha+\beta} &= D_\alpha \cdot D_\beta = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \cdot \begin{pmatrix} \cos \beta & -\sin \beta \\ \sin \beta & \cos \beta \end{pmatrix} \\ &= \begin{pmatrix} \cos \alpha \cdot \cos \beta - \sin \alpha \cdot \sin \beta & -(\cos \alpha \cdot \sin \beta + \sin \alpha \cdot \cos \beta) \\ \cos \alpha \cdot \sin \beta + \sin \alpha \cdot \cos \beta & \cos \alpha \cdot \cos \beta - \sin \alpha \cdot \sin \beta \end{pmatrix} \end{aligned}$$

Der Vergleich der Matrixelemente ergibt die *Additionstheoreme* für Sinus und Cosinus:

$$\boxed{\begin{aligned} \cos(\alpha + \beta) &= \cos \alpha \cdot \cos \beta - \sin \alpha \cdot \sin \beta \\ \sin(\alpha + \beta) &= \sin \alpha \cdot \cos \beta + \cos \alpha \cdot \sin \beta \end{aligned}}$$

### 9.2.4 Passive Drehungen

Unter einer *passiven Drehung* versteht man eine *Drehung des Koordinatensystems*. Dadurch ändern sich für beliebige Vektoren (die hier unverändert bleiben) die Koordinaten.

Dabei führt die aktive Drehung um eine Achse mit Winkel  $\alpha$  genau auf die gleiche Situation wie eine Drehung des Koordinatensystems mit Winkel  $(-\alpha)$  um die selbe Achse: In beiden Fällen erscheint der Vektor nach der Abbildung um den Winkel  $\alpha$  gedreht.

Also lassen sich passive Drehungen gerade durch die Inversen (also Transponierten) der aktiven Drehmatrizen beschreiben.

### 9.2.5 Drehungen in $\mathbb{R}^3$

Im  $\mathbb{R}^3$  lassen sich drei fundamentale Drehungen um die Koordinatenachsen definieren. Zunächst stellt sich heraus, dass unsere  $(2 \times 2)$ -Drehmatrix für  $\mathbb{R}^2$  von oben direkt zu einer (*aktiven*) Drehmatrix um die  $x_3$ -Achse erweitert werden kann (indem man die  $x_3$ -Koordinate fest hält), und man erhält:

$$D_{z,\alpha} = \begin{pmatrix} \cos \alpha & -\sin \alpha & 0 \\ \sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Mit ähnlichen Überlegungen erhält man für die beiden anderen Achsen:

$$D_{x,\alpha} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha & -\sin \alpha \\ 0 & \sin \alpha & \cos \alpha \end{pmatrix} \quad \text{und} \quad D_{y,\alpha} = \begin{pmatrix} \cos \alpha & 0 & \sin \alpha \\ 0 & 1 & 0 \\ -\sin \alpha & 0 & \cos \alpha \end{pmatrix}$$

Man rechnet leicht nach, dass alle diese Matrizen Determinante  $(+1)$  besitzen und orthogonal sind.

Aus den fundamentalen Drehmatrizen lassen sich durch Multiplikation Drehungen um beliebige Achsen konstruieren; das ist z.B. in der Computergraphik oder in der Robotik extrem wichtig.<sup>1</sup>

Dabei ist zu beachten, dass Drehungen *um die gleiche Achse* kommutieren; d.h. man kann sie in beliebiger Reihenfolge ausführen.

*Das gilt jedoch nicht für das Produkt zweier Drehungen um verschiedene Achsen!* (Zum Beweis berechne man die beiden Produkte von  $D_{z,\alpha}$  und  $D_{x,\beta}$ ; je nach Reihenfolge ergibt sich eine unterschiedliche Gesamtmatrix.)

Zum Abschluss noch ein

**Beispiel:** In  $\mathbb{R}^2$  betrachten wir die (aktive) Drehung des Vektors  $\vec{x} := \begin{pmatrix} 4 & -3 \end{pmatrix}^T$  um einen Winkel von  $\alpha := \frac{\pi}{2}$  (das sind 90 Grad):

$$\vec{y} := D_{\pi/2} \cdot \vec{x} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 4 \\ -3 \end{pmatrix} = \begin{pmatrix} 3 \\ 4 \end{pmatrix}$$

Und tatsächlich ist  $\vec{y} \perp \vec{x}$ , denn  $\vec{y} \bullet \vec{x} = 12 - 12 = 0$ .

### 9.2.6 Orthogonale Gruppe

Produkte orthogonaler Matrizen sind nach Satz 7.23 (Eigenschaften orthogonaler Matrizen) wiederum orthogonal. Da sie auch stets invertierbar sind (durch Transposition), ist auch hier das Matrixprodukt eine abgeschlossene Operation mit inversen Elementen, sodass gilt:

**Satz 9.7** (Orthogonale Gruppe). *Die Menge*

$$O(n) := \{ M \in GL(n) \mid M^T \cdot M = M \cdot M^T = \mathbb{1}_n \}$$

*bildet mit dem Matrixprodukt eine Untergruppe von  $GL(n)$ , die orthogonale Gruppe.*

<sup>1</sup>Stichwort: Euler-Winkel. Jede Drehung in  $\mathbb{R}^3$  lässt sich mit drei Drehwinkeln und den Drehmatrizen um die Koordinatenachsen beschreiben.

**Bemerkung:** Auch diese Gruppe ist nicht abelsch, wie man an den Drehmatrizen in  $\mathbb{R}^3$  einsehen kann.

Genau wie bei der linearen Gruppe (siehe die Sätze 9.4 und 9.5) können wir auch hier noch eine Untergruppe definieren, die aus den unimodularen Matrizen von  $O(n)$  besteht (also denen, die Determinante 1 haben):

**Satz 9.8** (Spezielle orthogonale Gruppe). *Die Menge*

$$SO(n) := \{M \in O(n) \mid \det M = 1\}$$

*bildet mit dem Matrixprodukt eine Untergruppe von  $O(n)$ , die spezielle orthogonale Gruppe.*

**Bemerkungen:**

- $SO(n)$  ist zusätzlich eine Untergruppe von  $SL(n)$ ; allerdings hatten wir bereits bei Satz 9.6 (Determinanten orthogonaler Matrizen) gesehen, dass nicht alle unimodularen Matrizen orthogonal sind.
- Die Drehungen im  $\mathbb{R}^2$  und  $\mathbb{R}^3$  gehören zu  $SO(2)$  bzw.  $SO(3)$ . Eine Komposition von zwei Drehungen ist also stets wieder eine Drehung, und die kombinierte Abbildungsmatrix hat Determinante 1.
- Die  $SO(n)$  ist auch als *Drehgruppe* bekannt. Sie enthält genau die Drehmatrizen zu  $\mathbb{R}^n$ . Die Matrizen aus  $O(n) \setminus SO(n)$  beschreiben dagegen *Drehspiegelungen*.
- Für  $n = 2$  ist der Drehpunkt der Koordinatenursprung. Für  $n = 3$  existiert eine eindeutige Drehachse zu jeder Drehmatrix (mehr dazu im nächsten Abschnitt). Für höhere Dimensionen wird es möglich, mit einer einzigen Drehmatrix in mehreren Ebenen (und um mehrere verschiedene Winkel) zu drehen; dies verfolgen wir hier nicht im Detail weiter.

## 9.3 Eigenwertproblem

### 9.3.1 Motivation und Formulierung

Jede lineare Abbildung von  $\mathbb{R}^n$  nach  $\mathbb{R}^n$  mit einer Abbildungsmatrix ungleich der Nullmatrix bildet Ursprungsgeraden auf Ursprungsgeraden ab, denn für solche Matrizen  $A \in \mathbb{R}^{(n,n)}$  und für beliebiges festes  $\vec{x} \neq \vec{0}$  gilt:

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

Dabei ist  $c\vec{x}$  mit  $c \in \mathbb{R}$  die Parametrisierung einer Ursprungsgeraden. Da auch  $\vec{y} := A\vec{x}$  fest und aus  $\mathbb{R}^n$  ist, ist auch  $c\vec{y}$  die Parametrisierung einer Ursprungsgeraden.

Üblicherweise ändert eine Ursprungsgerade unter der Abbildung aber ihre Orientierung:

**Beispiel:** Wir betrachten in  $\mathbb{R}^3$  die Abbildungsmatrix

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

sowie den Richtungsvektor der ursprünglichen Geraden  $\vec{x} := (3 \ 2 \ -1)^T$ . Dann ist

$$\vec{y} := A\vec{x} = \begin{pmatrix} 1 & -2 & 2 \\ 2 & 1 & -1 \\ 3 & 0 & 2 \end{pmatrix} \cdot \begin{pmatrix} 3 \\ 2 \\ -1 \end{pmatrix} = \begin{pmatrix} -3 \\ 9 \\ 7 \end{pmatrix}$$

Wenn diese Vektoren parallel wären, dürften sie sich nur um einen Skalierungsfaktor unterscheiden. Der Vergleich der ersten Komponenten führt auf den Faktor  $(-1)$ ; dieser findet sich für die anderen beiden Komponenten aber nicht wieder. (Alternativ, und gerne zur Übung, berechne man das Kreuzprodukt und verifiziere, dass es nicht dem Nullvektor entspricht; allgemein (nicht nur in  $\mathbb{R}^3$ ) könnte man auch das Skalarprodukt und die Cauchy-Schwarz-Ungleichung (Satz 6.7) verwenden.)

Uns interessieren hier die Ursprungsgeraden, die bei einer gegebenen Abbildung *auf sich selbst* abgebildet werden:

**Definition 9.9** (Fixgerade). Eine Gerade  $G \subseteq \mathbb{R}^n$  heißt Fixgerade einer linearen Abbildung mit Abbildungsmatrix  $A \in \mathbb{R}^{(n,n)}$ , falls

$$\forall \vec{x} \in G : A\vec{x} \in G$$

**Bemerkung:** Falls die Fixgerade eine Ursprungsgerade mit Richtungsvektor  $\vec{x}$  ist, werden ihre Vektoren unter der Abbildung lediglich skaliert, denn aufgrund der Fixgeraden-Eigenschaft muss für ein  $\lambda \neq 0$  gelten:

$$\vec{y} := A\vec{x} = \lambda\vec{x}$$

Und dann ist auch  $c\vec{y} = c\lambda\vec{x} = \lambda \cdot c\vec{x}$ . Jeder Ortsvektor eines Punktes auf der Geraden wird also durch  $A$  mit einem festen Faktor  $\lambda$  skaliert.

**Beispiel:** Jede Gerade (ob Ursprungsgerade oder nicht) ist Fixgerade zur identischen Abbildung mit der Matrix  $\mathbb{1}_n$ .

Wie wir gleich noch sehen werden, lohnt es sich nicht nur, einzelne Geraden zu betrachten, sondern sogar ganze Untervektorräume. Zunächst formulieren wir aber nun das Eigenwertproblem:

**Definition 9.10** (Eigenwertproblem).  $\lambda$  heißt Eigenwert einer quadratischen Matrix  $A \in \mathbb{R}^{(n,n)}$ , falls es Vektoren  $\vec{v} \in \mathbb{R}^n$  gibt, die der Gleichung

$$A \cdot \vec{v} = \lambda \cdot \vec{v}$$

genügen. Jeder solche Vektor heißt Eigenvektor von  $A$  zum Eigenwert  $\lambda$ .

**Bemerkungen:**

- Das titelgebende “Problem” besteht darin, zu gegebenem  $A$  die Eigenwerte und jeweils die zugehörigen Eigenvektoren zu finden; dazu später mehr.
- Der Nullvektor  $\vec{0}$  erfüllt die Eigenwertgleichung stets – aber er beschreibt keine Fixgerade für  $A$ , und es ist nicht klar, zu welchem Eigenwert er Eigenvektor wäre – denn die Gleichung wäre für jeden Wert  $\lambda$  richtig.

Wir bezeichnen  $\vec{0}$  daher als *trivialen Eigenvektor*, lassen ihn aber bei der Suche nach Lösungen des Eigenwertproblems außer acht.

**Beispiel:** Wir rechnen nach, dass der Vektor  $\vec{v} := (1 \ 1 \ 1)^T$  ein Eigenvektor zur folgenden Matrix  $A \in \mathbb{R}^{(3,3)}$  ist:

$$A := \begin{pmatrix} 2 & 1 & 2 \\ 0 & 2 & 3 \\ 3 & 2 & 0 \end{pmatrix}$$

Dazu setzen wir den Vektor in die obige Eigenwertgleichung ein:

$$A \cdot \vec{v} = \begin{pmatrix} 2 & 1 & 2 \\ 0 & 2 & 3 \\ 3 & 2 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 5 \\ 5 \\ 5 \end{pmatrix} = 5 \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = 5 \cdot \vec{v} \quad \checkmark$$

Obendrein haben wir so auch den Eigenwert gefunden:  $\vec{v}$  ist Eigenvektor von  $A$  zum Eigenwert 5.

### 9.3.2 Eigenräume und geometrische Vielfachheit

Die Eigenwertgleichung ist linear; daher können wir statt einzelner Eigenvektoren auch *Linearkombinationen* von Eigenvektoren (jedoch zu einem gemeinsamen festen Eigenwert) betrachten:

**Satz 9.11** (Linearkombination von Eigenvektoren). Für gegebene Matrix  $A \in \mathbb{R}^{(n,n)}$  sowie einen Eigenwert  $\lambda$  von  $A$  sind alle Linearkombinationen von Eigenvektoren zu  $\lambda$  ebenfalls Eigenvektoren von  $A$  zum Eigenwert  $\lambda$ .

(Beweis: S. 347.)

**Bemerkung:** Insbesondere ist also jede Skalierung eines Eigenvektors wiederum ein Eigenvektor zum selben Eigenwert.

**Beispiele:**

- Für  $A, \vec{v}$  wie im Beispiel zu Definition 9.10 rechnen wir nach, dass  $2\vec{v}$  ebenfalls Eigenvektor von  $A$  zum Eigenwert 5 ist:

$$A \cdot (2\vec{v}) = \begin{pmatrix} 2 & 1 & 2 \\ 0 & 2 & 3 \\ 3 & 2 & 0 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix} = \begin{pmatrix} 10 \\ 10 \\ 10 \end{pmatrix} = 5 \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix} = 5 \cdot (2\vec{v}) \quad \checkmark$$

- Man findet (man überprüfe das zur Übung durch Nachrechnen mit der Eigenwertgleichung), dass die Vektoren  $\vec{v} := (1 \ 0 \ 1)^T$  und  $\vec{w} := (0 \ 1 \ 0)^T$  beide Eigenvektoren der Matrix

$$A := \begin{pmatrix} 1 & 0 & 2 \\ 0 & 3 & 0 \\ 2 & 0 & 1 \end{pmatrix}$$

sind, jeweils zum gemeinsamen Eigenwert  $\lambda = 3$ ; also sind  $\vec{v}, \vec{w} \in E_{A,3}$ .

Die Vektoren sind linear unabhängig, spannen also eine Ebene auf. Wir zeigen, dass jeder Vektor in dieser Ebene ebenfalls Eigenvektor zum Eigenwert 3 ist. Dazu betrachten wir eine Linearkombination  $\vec{x} := c\vec{v} + d\vec{w}$ , also  $\vec{x} = (c \ d \ c)^T$ . Dann ist

$$A \cdot \vec{x} = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 3 & 0 \\ 2 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} c \\ d \\ c \end{pmatrix} = \begin{pmatrix} c + 0 + 2c \\ 0 + 3d + 0 \\ 2c + 0 + c \end{pmatrix} = \begin{pmatrix} 3c \\ 3d \\ 3c \end{pmatrix} = 3 \begin{pmatrix} c \\ d \\ c \end{pmatrix} = 3 \cdot \vec{x} \quad \checkmark$$

Dies erlaubt uns die Definition nicht nur von Fixgeraden, sondern ganzer Unterräume von  $\mathbb{R}^n$ :

**Definition 9.12** (Eigenraum). Die Menge aller Eigenvektoren einer Matrix  $A \in \mathbb{R}^{(n,n)}$  zum Eigenwert  $\lambda$  heißt Eigenraum von  $A$  zu  $\lambda$ :

$$E_{A,\lambda} := \{ \vec{v} \in \mathbb{R}^n \mid A \cdot \vec{v} = \lambda \cdot \vec{v} \}$$

Dessen Dimension wird geometrische Vielfachheit des Eigenwerts  $\lambda$  genannt:

$$\gamma(\lambda) := \dim E_{A,\lambda}$$

**Bemerkung:** Für den Fall, dass die Eigenvektoren alle auf einer Fixgeraden liegen, ist die Dimension des Eigenraums (und damit die geometrische Vielfachheit des Eigenwerts) genau 1.

Weiterhin gilt mit Satz 9.11 und der Tatsache, dass auch der triviale Eigenvektor  $\vec{0}$  stets als Linearkombination von Eigenvektoren zu einem bestimmten  $\lambda$  möglich ist, folgender

**Satz 9.13** (Eigenräume sind Unterräume). Der Eigenraum  $E_{A,\lambda}$  einer Matrix  $A$  zum Eigenwert  $\lambda$  ist ein Untervektorraum von  $\mathbb{R}^n$

### 9.3.3 Lösung des Eigenwertproblems

Nun wollen wir uns die Eigenwerte und Eigenvektoren von  $A \in \mathbb{R}^{(n,n)}$  systematisch verschaffen. Dazu formulieren wir die Eigenwertgleichung folgendermaßen um:

$$\begin{aligned} A \cdot \vec{v} &= \lambda \cdot \vec{v} = \lambda \cdot \mathbb{1}_n \cdot \vec{v} \\ \Leftrightarrow \quad &\boxed{(A - \lambda \cdot \mathbb{1}_n) \cdot \vec{v} = \vec{0}} \end{aligned}$$

Dies ist ein *homogenes LGS* mit der Unbekannten  $\vec{v}$  – Die Lösungen sind Elemente des Kerns von  $(A - \lambda \mathbb{1}_n)$ . Eine alternative Definition der Eigenräume ist demnach:

$$\boxed{E_{A,\lambda} = \ker(A - \lambda \cdot \mathbb{1}_n)}$$

Also können wir die (genauer: sämtliche) Eigenvektoren zu gegebenem  $\lambda$  mit den bereits bekannten Methoden bestimmen (Lösen des homogenen LGS mit Gauß-Verfahren).

Doch wie finden sich die hierfür nötigen Werte  $\lambda$ ? Nun ist der Nullvektor stets trivial im Kern jeder Matrix enthalten – aber diesen suchen wir explizit nicht. Für uns ist sind also gerade genau die anderen Vektoren aus dem Kern von Interesse. Damit solche existieren können, muss der Kern nichttrivial sein – und dies ist nur möglich, wenn die Matrix

$$A - \lambda \mathbb{1}_n$$

singulär ist (siehe Satz 9.2 (Invertierbare Matrizen)). Für alle Eigenwerte  $\lambda$  gilt daher:

$$\boxed{\det(A - \lambda \mathbb{1}_n) = 0}$$

Nun ist dies eine (skalare) Gleichung in  $\lambda$ . Wenn wir gewisse begriffliche Subtilitäten außer acht lassen (siehe z.B. [1] für eine genauere Behandlung) und uns auf das Rechnen im unendlichen Körper  $\mathbb{R}$  beschränken<sup>2</sup>, sehen wir auf der Diagonalen der Matrix  $n$  Linearfaktoren  $(A_{j,j} - \lambda)$ , sowie auf den off-Diagonal-Positionen weitere Zahlen.

Die Determinante solch einer Matrix ist dann eine Polynomfunktion:

**Definition 9.14** (Charakteristisches Polynom). Für  $A \in \mathbb{R}^{(n,n)}$  ist mit

$$\chi_A(\lambda) := \det(A - \lambda \mathbb{1}_n) \in \mathbb{R}[\lambda]$$

ein reelles Polynom in  $\lambda$  gegeben – das charakteristische Polynom von  $A$ .

#### Bemerkungen:

- Das Polynom hat Grad  $n$ , denn bei Benutzung der Leibniz-Formel (Satz 8.20) trägt stets auch die identische Permutation  $\text{id}_n$  bei, die das Produkt aller Diagonalelemente der Matrix liefert – dadurch ergibt sich unter anderem stets der Beitrag

$$(-1)^n \lambda^n$$

- Das Absolutglied des Polynoms erhalten wir, wenn wir als Argument 0 einsetzen:

$$\chi_A(0) = \det(A - 0 \mathbb{1}_n) = \det A$$

**Beispiel:** Für die von oben bekannte Matrix

$$A := \begin{pmatrix} 2 & 1 & 2 \\ 0 & 2 & 3 \\ 3 & 2 & 0 \end{pmatrix}$$

berechnen wir das charakteristische Polynom mit der Sarrus-Regel:

$$\begin{aligned} \chi_A(\lambda) &= \det \begin{pmatrix} 2-\lambda & 1 & 2 \\ 0 & 2-\lambda & 3 \\ 3 & 2 & -\lambda \end{pmatrix} \\ &\stackrel{\text{Sarrus}}{=} -\lambda(2-\lambda)^2 + 9 + 0 - 6(2-\lambda) - 0 - 6(2-\lambda) \\ &= -\lambda(\lambda^2 - 4\lambda + 4) + 9 - 12(2-\lambda) \\ &= -\lambda^3 + 4\lambda^2 - 4\lambda + 9 - 24 + 12\lambda \\ &= -\lambda^3 + 4\lambda^2 + 8\lambda - 15 \end{aligned}$$

Wir berechnen die selbe Determinante später noch einmal mit einem Trick, um ein teilweise faktorisiertes Polynom zu erhalten. Wir können hier jedoch schon ablesen, dass  $\det A = -15$ .

Dann gilt für die Lösung auch:

**Satz 9.15** (Eigenwerte einer Matrix). Die Eigenwerte einer Matrix  $A \in \mathbb{R}^{(n,n)}$  sind die Nullstellen ihres charakteristischen Polynoms, also die Werte  $\lambda$ , für die

$$\chi_A(\lambda) = 0$$

gilt.

<sup>2</sup>Dort kann man nämlich Polynome und Polynomfunktionen quasi-identisch betrachten.

### Bemerkungen:

- Nach dem Fundamentalsatz der Algebra (Satz 5.29) wird das Polynom  $n$  Nullstellen in  $\mathbb{C}$  besitzen (ggf. einige davon mehrfach); das Polynom ist ein Produkt der entsprechenden Linearfaktoren.

Diese sind jedoch oft nicht leicht zu finden, wenn die Matrix nicht eine günstige Struktur besaß, die es erlaubt hat, beim Berechnen der fraglichen Determinante bereits Linearfaktoren auszuklammern.

- Man beachte, dass die reellen Zahlen  $\mathbb{R}$  algebraisch nicht abgeschlossen sind. Die Nullstellen des charakteristischen Polynoms müssen also nicht alle reell sein, selbst wenn dessen Koeffizienten alle reell sind.

### Beispiele:

- Wir berechnen zunächst die Eigenwerte der von oben bekannten Matrix

$$A := \begin{pmatrix} 1 & 0 & 2 \\ 0 & 3 & 0 \\ 2 & 0 & 1 \end{pmatrix}$$

Aufgrund der günstigen Struktur können wir direkt eine Laplace-Entwicklung (z.B. nach der zweiten Spalte) ausführen:

$$\begin{aligned} \chi_A(\lambda) &= \det \begin{pmatrix} 1-\lambda & 0 & 2 \\ 0 & 3-\lambda & 0 \\ 2 & 0 & 1-\lambda \end{pmatrix} \\ &\stackrel{\text{S.2}}{=} (3-\lambda) \det \begin{pmatrix} 1-\lambda & 2 \\ 2 & 1-\lambda \end{pmatrix} \\ &= (3-\lambda)((1-\lambda)^2 - 4) = (3-\lambda)(\lambda^2 - 2\lambda + 1 - 4) \\ &= (3-\lambda)(\lambda^2 - 2\lambda - 3) = (3-\lambda)(\lambda+1)(\lambda-3) \\ &= -(\lambda-3)^2(\lambda+1) \end{aligned}$$

Offenbar gibt es einen doppelten Eigenwert 3 sowie einen Eigenwert  $(-1)$ . Die Faktorisierung gelang hier gut, weil ein Linearfaktor schon ganz zu Anfang per Laplace ausgeklammert war, und weil dann nur noch ein quadratisches Polynom zurück blieb (hier mit dem Satz von Vieta gelöst; mit Satz 1.34 hätte man dieselbe Lösung berechnet).

- Für die andere Beispiel-Matrix (wir nennen sie hier  $B$ ) hatten wir schon einmal das charakteristische Polynom berechnet – jedoch ließ sich nicht ohne weiteres eine Nullstelle ablesen. Entweder man erhält den Tipp, dass  $\lambda = 5$  eine Nullstelle ist – dann könnte man den entsprechenden Linearfaktor mit Polynomdivision abspalten und bekäme auch hier ein quadratisches Polynom.

Oder man erkennt, dass die Matrix

$$B := \begin{pmatrix} 2 & 1 & 2 \\ 0 & 2 & 3 \\ 3 & 2 & 0 \end{pmatrix}$$

in jeder Zeile Einträge der Gesamtsumme 5 besitzt. Wenn dies der Fall ist, können wir durch Addition der Spalten (hier erste und zweite jeweils zur dritten) folgende Situation erreichen:

$$\chi_B(\lambda) = \det \begin{pmatrix} 2-\lambda & 1 & 2 \\ 0 & 2-\lambda & 3 \\ 3 & 2 & -\lambda \end{pmatrix} = \det \begin{pmatrix} 2-\lambda & 1 & 5-\lambda \\ 0 & 2-\lambda & 5-\lambda \\ 3 & 2 & 5-\lambda \end{pmatrix}$$

Nach der Multilinearität der Determinanten (vgl. die Weierstraß-Axiome in Definition 8.16) können wir den Faktor  $(5-\lambda)$  ausklammern. Danach könnte man noch weitere skalierte Additionen durchführen und mit Laplace fertig rechnen – wir wenden statt dessen Sarrus an:

$$\begin{aligned} \dots &= (5-\lambda) \det \begin{pmatrix} 2-\lambda & 1 & 1 \\ 0 & 2-\lambda & 1 \\ 3 & 2 & 1 \end{pmatrix} \\ &\stackrel{\text{Sarrus}}{=} (5-\lambda)((2-\lambda)^2 + 3 + 0 - 3(2-\lambda) - 0 - 2(2-\lambda)) \\ &= (5-\lambda)(\lambda^2 - 4\lambda + 4 + 3 - 6 + 3\lambda - 4 + 2\lambda) \\ &= (5-\lambda)(\lambda^2 + \lambda - 3) \end{aligned}$$



Offenbar ist  $\lambda = 5$  also ein Eigenwert. Für die beiden anderen Eigenwerte rechnen wir nach:

$$\lambda = -\frac{1}{2} \pm \sqrt{\frac{1}{4} + 3} = -\frac{1}{2} \pm \sqrt{\frac{13}{4}} = \frac{-1 \pm \sqrt{13}}{2}$$

Hier liegen also drei verschiedene (und jeweils einfache) Eigenwerte vor.

Bevor wir Beispiele zum Berechnen von Eigenvektoren betrachten, hier noch zwei Definitionen:

**Definition 9.16** (Spektrum einer Matrix). *Das Spektrum einer quadratischen Matrix  $A \in \mathbb{R}^{(n,n)}$  bezeichnet die Menge aller ihrer Eigenwerte.*

**Beispiel:** Die Matrix  $A$  von oben hat das Spektrum  $\{-1, 3\}$ , die Matrix  $B$  dagegen

$$\left\{ \frac{-1 - \sqrt{13}}{2}, \frac{-1 + \sqrt{13}}{2}, 5 \right\}$$

**Definition 9.17** (Algebraische Vielfachheit). *Die algebraische Vielfachheit eines Eigenwerts  $\tilde{\lambda}$  einer Matrix  $A \in \mathbb{R}^{(n,n)}$  entspricht der Anzahl der Linearfaktoren  $(\lambda - \tilde{\lambda})$  im charakteristischen Polynom  $\chi_A(\lambda)$ . Wir notieren sie mit*

$$\alpha(\lambda)$$

**Beispiel:** Für die Matrix  $A$  von oben ist  $\alpha(3) = 2$  und  $\alpha(-1) = 1$ . Die drei Eigenwerte von  $B$  haben jeweils algebraische Vielfachheit 1.

### 9.3.4 Berechnung von Eigenräumen

**Beispiele:**

- Zunächst eine volle Rechnung in  $\mathbb{R}^2$ . Wir betrachten die Matrix

$$A := \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

Zunächst bestimmen wir die Eigenwerte; hierzu suchen wir nach Satz 9.15 die Nullstellen des charakteristischen Polynoms:

$$\begin{aligned} 0 &= \det(A - \lambda \mathbb{1}_2) = \det \begin{pmatrix} 2 - \lambda & 1 \\ 1 & 2 - \lambda \end{pmatrix} = (2 - \lambda)^2 - 1 = \lambda^2 - 4\lambda + 4 - 1 = \lambda^2 - 4\lambda + 3 \\ &= (\lambda - 1)(\lambda - 3) \end{aligned}$$

Die Zerlegung des charakteristischen Polynoms  $(\lambda^2 - 4\lambda + 3)$  in Linearfaktoren gelingt entweder durch “scharfes Hinsehen” (Satz von Vieta) oder durch Bestimmen der Lösungen der quadratischen Gleichung (das ergäbe  $2 \pm \sqrt{4 - 3} = 2 \pm 1$ ).

Wir wählen  $\lambda_1 := 1$  und  $\lambda_2 := 3$ . Beide Linearfaktoren kommen jeweils mit Potenz 1 im charakteristischen Polynom vor, also haben beide Eigenwerte jeweils algebraische Vielfachheit 1.

Da die Eigenwerte gerade über die Bedingung berechnet wurden, dass der Kern der Matrix  $(A - \lambda \mathbb{1})$  nichttrivial sein muss, werden wir für jeden Eigenwert beim Lösen des homogenen LGS  $(A - \lambda \mathbb{1} \mid \vec{0})$  mindestens eine Null-Zeile erhalten, die zu einer freien Variablen führt und damit auch bedingt, dass der zugehörige Eigenraum mindestens die Dimension 1 besitzt.

Für  $\lambda_1 = 1$  erhalten wir:

$$A - \lambda_1 \mathbb{1}_2 = \begin{pmatrix} 2 - 1 & 1 \\ 1 & 2 - 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

Das entsprechende homogene LGS hat zwei identische Zeilen, sodass wir direkt durch Addition einer mit  $(-1)$  skalierten Zeile zur anderen die erwähnte Null-Zeile erzeugen können. Das Resultat ist das LGS (da homogen, hier ohne Trennstrich und Inhomogenität)

$$\begin{pmatrix} 1 & 1 \end{pmatrix} \Leftrightarrow x_1 = -x_2$$

Wenn wir nun die zweite Komponente des Eigenvektors als freie Variable wählen, erhalten wir:

$$\vec{v}_1 = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -x_2 \\ x_2 \end{pmatrix} = x_2 \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

Damit gilt für den Eigenraum:

$$E_{A,\lambda_1} = \left\{ x_2 \begin{pmatrix} -1 \\ 1 \end{pmatrix} \mid x_2 \in \mathbb{R} \right\}$$

Also hat  $\lambda_1$  die geometrische Vielfachheit 1.

Mit dem Eigenraum haben wir *sämtliche* Eigenvektoren von  $A$  zum Eigenwert  $\lambda_1$  angegeben. Falls jedoch nur nach *einem* Eigenvektor gefragt ist, dürfen wir die Angabe des Eigenraums überspringen, und es reicht, ein *konkretes*  $x_2$  zu wählen<sup>3</sup> – z.B.  $x_2 := 1$ . Damit würden wir erhalten:

$$\vec{v}_1 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

Für  $\lambda_2 = 3$  entsprechend:

$$A - \lambda_2 \mathbb{1}_2 = \begin{pmatrix} 2-3 & 1 \\ 1 & 2-3 \end{pmatrix} = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}$$

Wenn wir die zweite Zeile zur ersten addieren, erhalten wir direkt eine Null-Zeile. Das resultierende LGS ist dann:

$$(1 \quad -1) \Leftrightarrow x_1 = x_2$$

Damit ergibt sich der zweite Eigenvektor mit freier Variable  $x_2$  als

$$\vec{v}_2 = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_2 \\ x_2 \end{pmatrix} = x_2 \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Und für den Eigenraum gilt:

$$E_{A,\lambda_2} = \left\{ x_2 \begin{pmatrix} 1 \\ 1 \end{pmatrix} \mid x_2 \in \mathbb{R} \right\}$$

Also hat  $\lambda_2$  ebenfalls geometrische Vielfachheit 1.

Man rechnet zur Probe leicht nach, dass es sich jeweils um Eigenvektoren zum betreffenden Eigenwert handelt.

- Dreireihiges Beispiel mit reduzierter geometrischer Vielfachheit. Wir betrachten die Matrix

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

Wir berechnen das charakteristische Polynom mit der Sarrus-Regel:

$$\begin{aligned} 0 &= \det \begin{pmatrix} -\lambda & 2 & -1 \\ 2 & -1-\lambda & 1 \\ 2 & -1 & 3-\lambda \end{pmatrix} \\ &= \lambda(1+\lambda)(3-\lambda) + 4 + 2 - 2(1+\lambda) - 4(3-\lambda) - \lambda \\ &= \lambda[(1+\lambda)(3-\lambda)] + 6 - 2 - 2\lambda - 12 + 4\lambda - \lambda \\ &= \lambda(3-\lambda+3\lambda-\lambda^2) - 8 + \lambda \\ &= -\lambda(\lambda^2 - 2\lambda - 3) - 8 + \lambda \\ &= -\lambda^3 + 2\lambda^2 + 3\lambda - 8 + \lambda \\ &= -\lambda^3 + 2\lambda^2 + 4\lambda - 8 \end{aligned}$$

---

<sup>3</sup>Meist wählt man hier besonders einfache Werte, z.B. 1, oder ggf. größere Zahlen, falls sich dadurch Brüche noch zu ganzen Zahlen multiplizieren lassen. Jeder Wert ungleich 0 wäre aber zulässig. Auch die Angabe *normierter* Eigenvektoren kann nützlich sein – dann wäre in diesem Fall  $x_2 := \frac{1}{\sqrt{2}}$  zu wählen.

Der Wert  $\lambda = 2$  ist eine Nullstelle des Polynoms – daher können wir den entsprechenden Linearfaktor abdividieren:

$$\begin{array}{r} (-1 \quad 2 \quad 4 \quad -8) \\ -(-1 \quad 2) \\ \hline (0 \quad 4) \\ (4 \quad -8) \\ -(4 \quad -8) \\ \hline (0) \end{array}$$

Also (unter Benutzung der dritten binomischen Formel):

$$-\lambda^3 + 2\lambda^2 + 4\lambda - 8 = (\lambda - 2) \cdot (-\lambda^2 + 4) = (\lambda - 2) \cdot (2 - \lambda)(2 + \lambda) = -(\lambda - 2)^2(\lambda + 2)$$

Also ist  $\lambda_1 := 2$  eine doppelte Nullstelle des charakteristischen Polynoms (d.h. der Eigenwert hat algebraische Vielfachheit 2), und  $\lambda_2 := -2$  ist einfache Nullstelle.

Wir berechnen die Eigenvektoren für  $\lambda_1 = 2$ . Zu lösen ist hier das homogene LGS (als Gauß-Schema notiert):

$$\begin{pmatrix} -2 & 2 & -1 \\ 2 & -3 & 1 \\ 2 & -1 & 1 \end{pmatrix}$$

Das Gauß-Schema lösen wir mit den bekannten Methoden:

$$\left( \begin{array}{ccc} -2 & 2 & -1 \\ 2 & -3 & 1 \\ 2 & -1 & 1 \end{array} \right) \begin{array}{l} \leftarrow 1 \\ \leftarrow -1 \end{array} \Leftrightarrow \left( \begin{array}{ccc} 0 & -5 & 0 \\ 2 & -3 & 1 \\ 0 & 2 & 0 \end{array} \right) \begin{array}{l} \frac{1}{5} \\ \\ \frac{1}{2} \end{array} \Leftrightarrow \left( \begin{array}{ccc} 0 & -1 & 0 \\ 2 & -3 & 1 \\ 0 & 1 & 0 \end{array} \right) \begin{array}{l} \leftarrow 3 \\ \leftarrow 1 \end{array} \Leftrightarrow \left( \begin{array}{ccc} 0 & 0 & 0 \\ 2 & 0 & 1 \\ 0 & 1 & 0 \end{array} \right)$$

$$\Leftrightarrow (2x_1 + x_3 = 0) \wedge (x_2 = 0) \Leftrightarrow (x_3 = -2x_1) \wedge (x_2 = 0)$$

Hier haben wir uns für  $x_1$  als freie Variable entschieden. Der Eigenraum ist:

$$E_{A, \lambda_1} = \left\{ x_1 \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix} \mid x_1 \in \mathbb{R} \right\}$$

Offensichtlich hat dieser Eigenraum nur die Dimension 1; entsprechend ist  $\gamma(\lambda_1) = 1 < \alpha(\lambda_1)$ . Hier liegt also ein Fall vor, bei dem geometrische und algebraische Vielfachheit eines Eigenwerts *nicht* übereinstimmen.

Wir bestimmen noch einen Eigenvektor für  $\lambda_2 = -2$ . Zu lösen ist das System

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

Also:

$$\left( \begin{array}{ccc} 2 & 2 & -1 \\ 2 & 1 & 1 \\ 2 & -1 & 5 \end{array} \right) \begin{array}{l} \leftarrow -1 \\ \leftarrow -1 \end{array} \Leftrightarrow \left( \begin{array}{ccc} 0 & 1 & -2 \\ 2 & 1 & 1 \\ 0 & -2 & 4 \end{array} \right) \begin{array}{l} \leftarrow -1 \\ \leftarrow 2 \end{array} \Leftrightarrow \left( \begin{array}{ccc} 0 & 1 & -2 \\ 2 & 0 & 3 \\ 0 & 0 & 0 \end{array} \right)$$

$$\Leftrightarrow (2x_1 + 3x_3 = 0) \wedge (x_2 - 2x_3 = 0) \Leftrightarrow \left( x_1 = -\frac{3}{2}x_3 \right) \wedge (x_2 = 2x_3)$$

Wenn wir  $x_3 := 2\tilde{x}_3$  vereinbaren, kompensieren wir den Nenner des auftretenden Bruchs und erhalten:

$$E_{A, \lambda_2} = \left\{ \tilde{x}_3 \begin{pmatrix} -3 \\ 4 \\ 2 \end{pmatrix} \mid \tilde{x}_3 \in \mathbb{R} \right\}$$

- Dreireihiges Beispiel mit vollen geometrischen Vielfachheiten. Wir betrachten die Matrix

$$A := \begin{pmatrix} 1 & 0 & 2 \\ 0 & 3 & 0 \\ 2 & 0 & 1 \end{pmatrix}$$

Hier ist die Eigenwertberechnung leicht durchführbar, weil sich sofort ein Linearfaktor des charakteristischen Polynoms abspalten lässt, wenn wir die Determinante nach der zweiten Zeile entwickeln (Laplace):

$$\begin{aligned} 0 &= \det \begin{pmatrix} 1-\lambda & 0 & 2 \\ 0 & 3-\lambda & 0 \\ 2 & 0 & 1-\lambda \end{pmatrix} = (3-\lambda) \det \begin{pmatrix} 1-\lambda & 2 \\ 2 & 1-\lambda \end{pmatrix} \\ &= (3-\lambda)[(1-\lambda)^2 - 4] = (3-\lambda)[\lambda^2 - 2\lambda + 1 - 4] \\ &= (3-\lambda)[\lambda^2 - 2\lambda - 3] = (3-\lambda) \cdot (\lambda+1)(\lambda-3) \\ &= -(\lambda-3)^2 \cdot (\lambda+1) \end{aligned}$$

Wir erhalten also  $\lambda_1 := -1$ , einen einfachen Eigenwert; und  $\lambda_2 = 3$  mit algebraischer Vielfachheit 2.

Für den Eigenvektor zu  $\lambda_1 = -1$  ist folgendes System zu lösen:

$$\begin{pmatrix} 2 & 0 & 2 \\ 0 & 4 & 0 \\ 2 & 0 & 2 \end{pmatrix}$$

Wir addieren das  $(-1)$ -fache der ersten Zeile auf die dritte Zeile, um direkt eine Null-Zeile zu erhalten; danach müssen wir nur noch die verbleibenden Zeilen skalieren. Es ergeben sich die Gleichungen

$$\Leftrightarrow (x_2 = 0) \wedge (x_1 = -x_3)$$

Damit erhalten wir ( $x_3$  als freie Variable) den Eigenraum

$$E_{A,\lambda_1} = \left\{ x_3 \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \mid x_3 \in \mathbb{R} \right\}$$

Für den Eigenraum zu  $\lambda_2 = 3$  lösen wir das folgende System:

$$\begin{pmatrix} -2 & 0 & 2 \\ 0 & 0 & 0 \\ 2 & 0 & -2 \end{pmatrix}$$

Hier liegt schon zu Anfang eine Nullzeile vor, und wir können direkt eine weitere erzeugen, indem wir die dritte Zeile auf die erste addieren. Die Skalierung der übrig gebliebenen dritten Zeile führen wir hier im Kopf aus und erhalten:

$$\cdots \Leftrightarrow (x_1 = x_3)$$

Nun haben wir zwei freie Variablen,  $x_2$  und (hier gewählt:)  $x_3$ . Dann gilt für den allgemeinen Eigenvektor:

$$\vec{v}_2 = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

Also gilt für den Eigenraum:

$$E_{A,\lambda_2} = \left\{ x_2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \mid x_2, x_3 \in \mathbb{R} \right\}$$

Die beiden Spannvektoren des Eigenraums sind linear unabhängig (hier sogar orthogonal!) und spannen also eine Ebene durch den Ursprung auf. Die geometrische Vielfachheit von  $\lambda_2$  beträgt 2, genau wie die algebraische Vielfachheit.

### 9.3.5 Weitere Eigenschaften von Eigenvektoren und Eigenwerten

Zunächst gilt für die Eigenwerte folgender

**Satz 9.18** (Produkt und Summe der Eigenwerte). *Für eine Matrix  $A \in \mathbb{R}^{(n,n)}$  mit den Eigenwerten  $\lambda_1, \dots, \lambda_n$  (Mehrfache Eigenwerte hier mit verschiedenen Indices benannt) gilt:*

$$\prod_{j=1}^n \lambda_j = \det A \quad \text{und} \quad \sum_{j=1}^n \lambda_j = \sum_{j=1}^n A_{jj}$$

(Beweis: S. 347.)

**Bemerkung:** Diese Formeln können praktisch sein, um das berechnete Eigenwertspektrum einer Matrix auf Rechenfehler zu prüfen.

**Beispiele:**

- Für die Matrix

$$A := \begin{pmatrix} 1 & 0 & 2 \\ 0 & 3 & 0 \\ 2 & 0 & 1 \end{pmatrix}$$

hatten wir oben die Eigenwerte 3, 3 und  $(-1)$  ermittelt. Ihr Produkt ist  $(-9)$  und ihre Summe beträgt 5. Man sieht sofort, dass die Summe auch der Summe der Diagonalelemente von  $A$  entspricht, da  $1+3+1 = 5$ . Die Determinante von  $A$  können wir anhand des charakteristischen Polynoms bestimmen. Es war:

$$\chi_A(\lambda) = -(\lambda - 3)^2(\lambda + 1)$$

Werten wir dieses Polynom an der Stelle 0 aus, so erhalten wir in der Tat den Wert

$$-(-3)^2 = -9$$

- Für die Matrix

$$B := \begin{pmatrix} 2 & 1 & 2 \\ 0 & 2 & 3 \\ 3 & 2 & 0 \end{pmatrix}$$

hatten wir als Determinante oben bereits  $\det B = -15$  ermittelt. Die Summe der Diagonalelemente beträgt 4. Das Spektrum von  $B$  war durch drei einfache Eigenwerte gegeben als

$$\left\{ \frac{-1 - \sqrt{13}}{2}, \frac{-1 + \sqrt{13}}{2}, 5 \right\}$$

Die Summe der Eigenwerte beträgt also

$$\frac{-1 - \sqrt{13}}{2} + \frac{-1 + \sqrt{13}}{2} + 5 = -\frac{1}{2} - \frac{1}{2} + 5 = 4$$

Und das Produkt der Eigenwerte ist (dritte binomische Formel; siehe Satz 1.31):

$$\left( \frac{-1 - \sqrt{13}}{2} \right) \left( \frac{-1 + \sqrt{13}}{2} \right) \cdot 5 = \frac{1}{4} \cdot ((-1)^2 - (\sqrt{13})^2) \cdot 5 = \frac{1}{4} \cdot (-12) \cdot 5 = -3 \cdot 5 = -15$$

- Für die Matrix

$$C := \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

hatten wir oben im Unterabschnitt zur Eigenraumberechnung bereits die Eigenwerte 1 und 3 ermittelt. Die Determinante beträgt  $\det C = 4 - 1 = 3 = 1 \cdot 3$ , und die Summe der Diagonalelemente ist  $2 + 2 = 4 = 1 + 3$ .

Aus dem vorigen Satz leiten wir noch eine wichtige Folgerung ab: Denn da das Produkt aller Eigenwerte gerade die Determinante einer Matrix ergibt, ist letztere nach dem Satz vom Nullprodukt (5.16) genau dann 0, wenn einer der Eigenwerte 0 beträgt:

**Satz 9.19** (Eigenwerte invertierbarer Matrizen).  $A \in \mathbb{R}^{(n,n)}$  ist genau dann invertierbar (d.h.  $\det A \neq 0$ ), wenn alle Eigenwerte von  $A$  ungleich 0 sind.

Für invertierbare Matrizen  $A \in \mathbb{R}^{(n,n)}$  sind also sämtliche Eigenwerte ungleich 0. Es gilt weiterhin folgender

**Satz 9.20** (Eigenvektoren und Eigenwerte der inversen Matrix). Ist  $A \in \mathbb{R}^{(n,n)}$  invertierbar und ist  $\vec{v}$  ein Eigenvektor zum Eigenwert  $\lambda$  von  $A$ , so ist  $\vec{v}$  auch ein Eigenvektor von  $A^{-1}$ , und zwar zum Eigenwert

$$\frac{1}{\lambda}$$

(Beweis: S. 348.)

**Bemerkung:** Das Spektrum der Eigenwerte einer inversen Matrix  $A^{-1}$  lässt sich also aus dem Spektrum von  $A$  direkt durch Bilden der Kehrwerte ermitteln; dabei gilt für die Eigenräume:

$$E_{A^{-1}, \frac{1}{\lambda}} = E_{A, \lambda}$$

**Beispiel:** Für die Matrix

$$C := \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

mit den Eigenwerten 1 und 3 bestimmen wir die Inverse nach der Formel aus den Beispielen in Abschnitt 9.1.2 (S. 245ff.):

$$C^{-1} = \frac{1}{3} \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} \frac{2}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} \end{pmatrix}$$

Für die Eigenwertberechnung wollen wir uns das Bruchrechnen sparen und stellen daher die Eigenwertgleichung um:

$$C^{-1}\vec{v} = \lambda\vec{v} \Leftrightarrow (3C^{-1})\vec{v} = (3\lambda)\vec{v}$$

Die Eigenwerte von  $C^{-1}$  entsprechen also den gedrittelten Eigenwerten von  $(3C^{-1})$ . Mit  $\tilde{\lambda} := 3\lambda$  bestimmen wir also die Eigenwerte  $\tilde{\lambda}$  von  $(3C^{-1})$ :

$$\chi_{3C^{-1}}(\tilde{\lambda}) = \det \begin{pmatrix} 2 - \tilde{\lambda} & -1 \\ -1 & 2 - \tilde{\lambda} \end{pmatrix} = 4 - 4\tilde{\lambda} + \tilde{\lambda}^2 - 1 = \tilde{\lambda}^2 - 4\tilde{\lambda} + 3 = (\tilde{\lambda} - 3)(\tilde{\lambda} - 1)$$

Genau wie  $C$  hat also  $(3C^{-1})$  die Eigenwerte 1 und 3. Damit sind die Eigenwerte von  $C^{-1}$  also per  $\lambda = \frac{\tilde{\lambda}}{3}$  gegeben als

$$\frac{1}{3} \quad \text{und} \quad \frac{3}{3} = 1$$

Beispielhafte Eigenvektoren von  $C$  waren (in dieser Reihenfolge) für die Eigenwerte 1 und 3 gegeben mit

$$\begin{pmatrix} -1 \\ 1 \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Wir prüfen für diese die Eigenwertgleichungen von  $C^{-1}$  nach und erwarten, dass dies auch Eigenvektoren zu den durch Kehrwertbildung erhaltenen Eigenwerten 1 und  $\frac{1}{3}$  (in dieser Reihenfolge) sind:

$$\begin{aligned} C^{-1} \begin{pmatrix} -1 \\ 1 \end{pmatrix} &= \frac{1}{3} \begin{pmatrix} -2-1 \\ 1+2 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} -3 \\ 3 \end{pmatrix} = 1 \cdot \begin{pmatrix} -1 \\ 1 \end{pmatrix} \quad \checkmark \\ C^{-1} \begin{pmatrix} 1 \\ 1 \end{pmatrix} &= \frac{1}{3} \begin{pmatrix} 2-1 \\ -1+2 \end{pmatrix} = \frac{1}{3} \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \checkmark \end{aligned}$$

(Übrigens hätten wir mit diesen Gleichungen direkt auch die beiden Eigenwerte von  $C^{-1}$  erhalten.)

Noch ein wichtiger Zusammenhang über die Vielfachheiten der Eigenwerte, den wir ohne vollständigen Beweis akzeptieren:

**Satz 9.21** (Algebraische und geometrische Vielfachheiten). *Sei  $\lambda$  ein Eigenwert der Matrix  $A \in \mathbb{R}^{(n,n)}$ . Dann gilt für dessen algebraische und geometrische Vielfachheit die Abschätzung*

$$1 \leq \gamma(\lambda) \leq \alpha(\lambda) \leq n$$

**Bemerkungen:**

- Die linke Abschätzung ist einzusehen, da die geometrische Vielfachheit  $\gamma(\lambda)$  der Dimension des Kerns von  $(A - \lambda \mathbb{1}_n)$  entspricht – diese ist mindestens 1, denn die Eigenwerte wurden gerade so bestimmt, dass dieser Kern nichttrivial ist.

Die rechte Abschätzung folgt daraus, dass das charakteristische Polynom  $\chi_A(\lambda)$  Grad  $n$  besitzt und nach dem Fundamentalsatz der Algebra (Satz 5.29) in  $n$  Linearfaktoren zerfällt. Selbst wenn ein Eigenwert (und diese bilden gerade die Nullstellen des Polynoms, definieren also die erwähnten Linearfaktoren) mehrfach vorkommt, so kann der korrespondierende Linearfaktor doch höchstens  $n$ -fach auftreten.

Die mittlere Abschätzung kann man mit den Mitteln des nächsten Abschnitts zur Diagonalisierung (s.u.) beweisen, was aber über diese Vorlesung hinaus führt. Es ist möglich, aus  $A$  durch teilweise Diagonalisierung eine ähnliche Matrix (s.u.) mit gleichem charakteristischen Polynom zu erzeugen, für die jeder der  $\gamma(\lambda)$  Basisvektoren von  $E_{A,\lambda}$  einen Diagonaleintrag mit Wert  $\lambda$  erzeugt. Bildet man nun formal das neue charakteristische Polynom, so muss der Linearfaktor für den Eigenwert  $\lambda$  dort mindestens  $\gamma(\lambda)$ -fach auftreten.

Die Lücken in letzterer Erklärung akzeptieren wir für die Vorlesung Mathematik 1.

**Beispiele:** Von der Richtigkeit der Abschätzung kann man sich an obigen Beispielen, z.B. in Unterabschnitt 9.3.4, überzeugen – dort hatten wir auch einen Fall gesehen, für den die geometrische Vielfachheit *nicht* der algebraischen entsprach, sondern echt kleiner war.

Wir schließen diesen Unterabschnitt ab mit einer sehr wichtigen Tatsache zu Eigenräumen:

**Satz 9.22** (Eigenvektoren zu verschiedenen Eigenwerten). *Eigenvektoren zu verschiedenen Eigenwerten einer Matrix  $A \in \mathbb{R}^{(n,n)}$  sind linear unabhängig.*

(Beweis: S. 348.)

**Bemerkung:** Insbesondere folgt daraus, dass beliebige Vektoren aus verschiedenen Eigenräumen stets linear unabhängig sind.

**Beispiele:** Auch hier verweisen wir auf die weiter oben durchgerechneten Beispiele zu Eigenräumen.

### 9.3.6 Dreiecks- und Diagonalmatrizen

#### Eigenwerte von Dreiecksmatrizen

Nach Definition 9.14 und Satz 9.15 erhalten wir die Eigenwerte von  $A \in \mathbb{R}^{(n,n)}$  Durch Lösen der Gleichung

$$0 = \det(A - \lambda \mathbb{1}_n)$$

Im Vergleich zu  $A$  wird die Matrix  $(A - \lambda \mathbb{1}_n)$  nur auf der Diagonalen verändert, bevor die Determinante gebildet wird. Insbesondere bleibt eine Dreiecksmatrix hierbei eine Dreiecksmatrix, sodass mit Satz 8.24 für eine solche Dreiecksmatrix  $A$  direkt folgt:

$$\det(A - \lambda \mathbb{1}_n) = \prod_{j=1}^n (A_{j,j} - \lambda)$$

Das charakteristische Polynom ist also bereits vollständig faktorisiert, und es gilt folgender

**Satz 9.23** (Eigenwerte von Dreiecksmatrizen). *Ist  $A \in \mathbb{R}^{(n,n)}$  eine Dreiecksmatrix, so ist das Spektrum von  $A$  durch die  $n$  Diagonalelemente von  $A$  gegeben.*

## Eigenvektoren von Diagonalmatrizen

Mit Satz 9.23 folgt, dass wir Diagonalmatrizen (die Dreiecksgestalt besitzen) aus  $\mathbb{R}^{(n,n)}$  stets schreiben können als

$$A = \text{diag}(\lambda_1, \dots, \lambda_n),$$

wobei die Zahlen  $\lambda_1, \dots, \lambda_n$  die Eigenwerte von  $A$  sind (mit Mehrfachnennung!).

Für die Bestimmung der Eigenräume zu einem Eigenwert  $\lambda$  mit algebraischer Vielfachheit  $\alpha(\lambda)$  ermittelt man wie gewohnt den Kern von

$$A_\lambda := A - \lambda \mathbb{1}_n$$

Die Matrix  $A_\lambda$  enthält an genau  $\alpha(\lambda)$  verschiedenen Positionen auf der Diagonalen eine 0; dies bedeutet hier auch direkt eine Null-Zeile im zu lösenden Gauß-Schema. Alle anderen Diagonaleinträge von  $A_\lambda$  sind ungleich 0 und lassen sich somit im Gauß-Schema auf 1 skalieren. Die zugehörigen Komponenten des Eigenvektors zu diesen letzteren Einträgen sind jeweils alle null, da das Gauß-Schema keine Inhomogenität besitzt. Die Komponenten des Eigenvektors, die zu den Null-Zeilen korrespondieren, sind beliebig wählbar.

Sei also  $I$  die Indexmenge der Nullzeilen in  $A_\lambda$ , und  $N := \{1, \dots, n\} \setminus I$  die Indexmenge der nichtverschwindenden Diagonaleinträge in  $A_\lambda$ . Dann ist

$$\ker A_\lambda = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \middle| (\forall j \in I : x_j \in \mathbb{R}) \wedge (\forall j \in N : x_j = 0) \right\} = \left\{ \sum_{j \in I} x_j \vec{e}_j \middle| \forall j \in I : x_j \in \mathbb{R} \right\}$$

Dies ist eine Linearkombination von  $\alpha(\lambda)$  verschiedenen kartesischen Einheitsvektoren, die alle linear unabhängig sind; damit ist die Dimension des Eigenraums  $\gamma(\lambda) = \alpha(\lambda)$ .

**Beispiel:** Für

$$A := \text{diag}(2, 2, 3) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

betrachten wir den Eigenwert  $\lambda = 2$ . Für den Eigenraum  $E_{A,2}$  und den Eigenvektor  $(x_1 \ x_2 \ x_3)^T$  rechnen wir:

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \Leftrightarrow x_2 = 0$$

Also ist

$$E_{A,2} = \left\{ x_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \middle| x_1, x_3 \in \mathbb{R} \right\}$$

Für den anderen Eigenwert  $\lambda = 3$  ergibt sich analog:

$$\begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} \Leftrightarrow \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \Leftrightarrow x_1 = 0 \wedge x_3 = 0$$

Also ist

$$E_{A,3} = \left\{ x_2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \middle| x_2 \in \mathbb{R} \right\}$$

Man überzeuge sich zur Übung, dass das gleiche Vorgehen richtig ist, wenn die diagonale Matrix  $A$  den Eigenwert 0 besitzt, d.h. wenn auf der Diagonalen von  $A$  schon mindestens eine 0 zu finden ist.

Insgesamt gilt folgender

**Satz 9.24** (Eigenräume von Diagonalmatrizen). *Für eine Diagonalmatrix  $A \in \mathbb{R}^{(n,n)}$  und einen Eigenwert  $\lambda$  von  $A$  ist der zugehörige Eigenraum gegeben als*

$$E_{A,\lambda} = \text{span}(\{ \vec{e}_j \mid A_{j,j} = \lambda \})$$



**Bemerkung:** Da alle Eigenwerte volle geometrische Vielfachheit haben, spannen die Eigenvektoren von  $A$  den gesamten Raum  $\mathbb{R}^n$  auf, und es ist möglich, die Eigenvektoren als kartesische Einheitsvektoren aus der Standardbasis  $E_n$  zu wählen.

### 9.3.7 Drehmatrizen für $\mathbb{R}^3$

Jede Drehung in  $\mathbb{R}^3$  mit einer Drehachse durch den Koordinatenursprung lässt sich durch einen Vektor  $\vec{n} \neq \vec{0}$  (welcher die Drehachse angibt) und einen Drehwinkel beschreiben. Die Transformation ist linear und wird durch eine Drehmatrix  $D \in SO(3)$  vermittelt.

Nun werden alle Punkte auf der Drehachse auf sich selbst abgebildet. Diese Punkte sind gerade die Vektoren  $k\vec{n}$  für  $k \in \mathbb{R}$ . Es gilt also:

$$D\vec{n} = \vec{n} = 1\vec{n}$$

Demnach ist der Richtungsvektor der Drehachse ein Eigenvektor von  $D$  zum Eigenwert 1; die Drehachse stellt den Eigenraum zum Eigenwert 1 dar.

**Bemerkung:** Die anderen beiden Eigenwerte von  $D$  sind üblicherweise nicht reell; wir wollen diese daher nicht weiter betrachten.

## 9.4 Diagonalisierung

Wir führen zunächst das Konzept der *Ähnlichkeit* von quadratischen Matrizen ein. Wie wir am Ende des vorigen Abschnitts gesehen haben, sind Diagonalmatrizen in Bezug auf das Eigenwertproblem besonders gutartig, da ihre Eigenwerte direkt ablesbar sind und ihre Eigenvektoren als die kartesischen Einheitsvektoren aus der Standardbasis  $E_n$  wählbar sind. Da sich für ähnliche Matrizen einige Eigenschaften übertragen – insbesondere das Eigenwertspektrum –, untersuchen wir danach, welche Matrizen mit Diagonalmatrizen ähnlich sein können.

Da die Eigenwerte reeller quadratischer Matrizen durchaus nicht alle reell sein müssen, wollen wir in diesem Abschnitt allgemeiner von quadratischen Matrizen sprechen. Die Beispiele, Übungs- und eventuelle Prüfungsaufgaben sind jedoch reell ausgeführt.

### 9.4.1 Ähnliche Matrizen

**Definition 9.25** (Ähnlichkeit von Matrizen). *Eine  $n \times n$ -Matrix  $B$  heißt ähnlich zu einer  $n \times n$ -Matrix  $A$ , falls es eine invertierbare  $n \times n$ -Matrix  $S$  gibt, sodass*

$$B = S^{-1}AS$$

**Bemerkung:** Da nach Satz 7.18 auch die Matrix  $S^{-1}$  invertierbar ist, ist dann auch  $A$  ähnlich zu  $B$ , denn

$$SBS^{-1} = A$$

Wir werden gleich zeigen, dass ähnliche Matrizen das selbe Eigenwertspektrum besitzen. Mit Satz 9.18 würde daraus bereits folgen, dass Determinante und *Spur* (das ist die Summe der Diagonalelemente einer Matrix) bei der Ähnlichkeitstransformation invariant bleiben. Es lohnt sich aber, dies schon zuvor separat fest zu stellen:

**Satz 9.26** (Determinante und Spur ähnlicher Matrizen). *Seien  $A, B$  ähnliche  $n \times n$ -Matrizen mit einer Transformationsmatrix  $S$  per  $B = S^{-1}AS$ . Dann gilt:*

$$\det B = \det A \quad \text{und} \quad \sum_{j=1}^n B_{j,j} = \sum_{j=1}^n A_{j,j}$$

(Beweis: S. 349.)

**Beispiel:** Wir betrachten für  $n = 2$  die Matrix

$$A := \begin{pmatrix} 1 & -2 \\ 1 & 7 \end{pmatrix} \quad \text{und dazu die Transformationsmatrix} \quad S := \begin{pmatrix} 3 & 1 \\ 5 & 2 \end{pmatrix}$$

Das Inverse von  $S$  bestimmen wir mit der Regel von Seite 246f. aus Unterabschnitt 9.1.2 (die Determinante von  $S$  beträgt 1):

$$S^{-1} = \begin{pmatrix} 2 & -1 \\ -5 & 3 \end{pmatrix}$$

Damit erhalten wir für  $B$ :

$$B := S^{-1}AS = \begin{pmatrix} 2 & -1 \\ -5 & 3 \end{pmatrix} \cdot \begin{pmatrix} 1 & -2 \\ 1 & 7 \end{pmatrix} \cdot \begin{pmatrix} 3 & 1 \\ 5 & 2 \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ -5 & 3 \end{pmatrix} \cdot \begin{pmatrix} -7 & -3 \\ 38 & 15 \end{pmatrix} = \begin{pmatrix} -52 & -21 \\ 149 & 60 \end{pmatrix}$$

Schon jetzt sehen wir, dass die Spuren beider Matrizen gleich sind:  $1 + 7 = 8 = -52 + 60$ . Die Determinanten betragen:

$$\det A = 7 - (-2) = 9 \quad \text{und} \quad \det B = -3120 - (-3129) = 9$$

---

Weiterhin gilt, wie oben schon erwähnt, folgender

**Satz 9.27** (Spektren ähnlicher Matrizen). *Zwei ähnliche  $n \times n$ -Matrizen  $A, B$  mit Transformationsmatrix  $S$  und  $B = S^{-1}AS$  besitzen gleiche Eigenwertspektren und gleiche charakteristische Polynome.*

*Ist  $\vec{v}$  ein Eigenvektor von  $A$  zum Eigenwert  $\lambda$ , so ist  $S^{-1}\vec{v}$  ein Eigenvektor von  $B$  zum selben Eigenwert  $\lambda$ .*

(Beweis: S. 350.)

**Bemerkung:** Man beachte, dass zwar die Eigenwerte invariant bleiben – die Eigenvektoren hingegen sind meistens nicht die gleichen – viel mehr berechnen sich die Eigenräume von  $B$  dadurch, dass man in den Eigenräumen von  $A$  alle Vektoren (von links) mit  $S^{-1}$  multipliziert.

**Beispiel:** Wir setzen unser Beispiel von Satz 9.26 fort und bestimmen das charakteristische Polynom von  $A$ :

$$\chi_A(\lambda) = \det \begin{pmatrix} 1-\lambda & -2 \\ 1 & 7-\lambda \end{pmatrix} = (1-\lambda)(7-\lambda) - (-2) = 7 - 8\lambda + \lambda^2 + 2 = \lambda^2 - 8\lambda + 9$$

Und für  $B$ :

$$\begin{aligned} \chi_B(\lambda) &= \det \begin{pmatrix} -52-\lambda & -21 \\ 149 & 60-\lambda \end{pmatrix} = (-52-\lambda)(60-\lambda) - (-3129) \\ &= -3120 + 52\lambda - 60\lambda + \lambda^2 + 3129 = \lambda^2 - 8\lambda + 9 \end{aligned}$$

Damit haben beide Matrizen die Eigenwerte  $4 \pm \sqrt{7}$

Wir berechnen noch einen Eigenvektor zum Eigenwert  $4 + \sqrt{7}$ . Zu lösen ist das homogene LGS

$$\begin{pmatrix} -3 - \sqrt{7} & -2 \\ 1 & 3 - \sqrt{7} \end{pmatrix}$$

Wenn wir die untere Zeile, mit  $(3 + \sqrt{7})$  skaliert, zur oberen addieren, erhalten wir oben eine Nullzeile. Für die linke Spalte ist dies klar; für die rechte ergibt sich mit der dritten binomischen Formel (Satz 1.31) ein Eintrag von

$$-2 + (3 - \sqrt{7})(3 + \sqrt{7}) = -2 + (9 - 7) = 0$$

Demnach ist

$$\vec{v} := \begin{pmatrix} -3 + \sqrt{7} \\ 1 \end{pmatrix}$$

ein Eigenvektor von  $A$  zum betrachteten Eigenwert. Probe:

$$A\vec{v} = \begin{pmatrix} 1 & -2 \\ 1 & 7 \end{pmatrix} \begin{pmatrix} -3 + \sqrt{7} \\ 1 \end{pmatrix} = \begin{pmatrix} -5 + \sqrt{7} \\ 4 + \sqrt{7} \end{pmatrix}$$

Weiterhin ist

$$(4 + \sqrt{7})\vec{v} = (4 + \sqrt{7}) \begin{pmatrix} -3 + \sqrt{7} \\ 1 \end{pmatrix} = \begin{pmatrix} -12 + 4\sqrt{7} - 3\sqrt{7} + 7 \\ 4 + \sqrt{7} \end{pmatrix} = \begin{pmatrix} -5 + \sqrt{7} \\ 4 + \sqrt{7} \end{pmatrix} \quad \checkmark$$

Der korrespondierende Eigenvektor von  $B$  ist:

$$\vec{w} := S^{-1}\vec{v} = \begin{pmatrix} 2 & -1 \\ -5 & 3 \end{pmatrix} \begin{pmatrix} -3 + \sqrt{7} \\ 1 \end{pmatrix} = \begin{pmatrix} -7 + 2\sqrt{7} \\ 18 - 5\sqrt{7} \end{pmatrix}$$

Probe:

$$B\vec{w} = \begin{pmatrix} -52 & -21 \\ 149 & 60 \end{pmatrix} \begin{pmatrix} -7 + 2\sqrt{7} \\ 18 - 5\sqrt{7} \end{pmatrix} = \begin{pmatrix} 364 - 104\sqrt{7} - 378 + 105\sqrt{7} \\ -1043 + 298\sqrt{7} + 1080 - 300\sqrt{7} \end{pmatrix} = \begin{pmatrix} -14 + \sqrt{7} \\ 37 - 2\sqrt{7} \end{pmatrix}$$

Weiterhin ist

$$(4 + \sqrt{7})\vec{w} = (4 + \sqrt{7}) \begin{pmatrix} -7 + 2\sqrt{7} \\ 18 - 5\sqrt{7} \end{pmatrix} = \begin{pmatrix} -28 + 8\sqrt{7} - 7\sqrt{7} + 14 \\ 72 - 20\sqrt{7} + 18\sqrt{7} - 35 \end{pmatrix} = \begin{pmatrix} -14 + \sqrt{7} \\ 37 - 2\sqrt{7} \end{pmatrix} \quad \checkmark$$

Für den anderen Eigenwert rechnet man gerne zur Übung nach, dass die zweite Aussage des Satzes ebenfalls zutrifft.

### 9.4.2 Diagonalisierbarkeit

**Definition 9.28** (Diagonalisierbarkeit). Eine  $n \times n$ -Matrix  $A$  heißt diagonalisierbar, falls  $A$  mit einer  $n \times n$ -Diagonalmatrix  $B$  ähnlich ist.

Wir suchen nun das Kriterium für die Diagonalisierbarkeit von  $A$ . Zunächst muss es also eine invertierbare Matrix  $S$  geben, sodass  $B = S^{-1}AS$ .

Wir wissen von Satz 9.27, dass  $A$  und  $B$  das gleiche Eigenwertspektrum besitzen. Da  $B$  jedoch eine Diagonalmatrix ist, ist damit klar, dass diese gemeinsamen Eigenwerte auch die Diagonalelemente von  $B$  darstellen (siehe Satz 9.23). Wir schreiben also (mit Mehrfachnennung der Eigenwerte):

$$B = \text{diag}(\lambda_1, \dots, \lambda_n)$$

Die Eigenwerte sind identisch mit den Eigenwerten von  $A$ .

Nun lässt sich die Ähnlichkeitsbedingung umschreiben als

$$SB = AS$$

Falls  $S$  die Spaltendarstellung  $S = (\vec{v}_1 \ \dots \ \vec{v}_n)$  besitzt, so ist die rechte Seite dieser Gleichung mit Satz 7.10 (Spalten des Matrixprodukts) gegeben als

$$AS = A(\vec{v}_1 \ \dots \ \vec{v}_n) = (A\vec{v}_1 \ \dots \ A\vec{v}_n)$$

Die  $k$ -te Spalte dieses Produkts ist also  $A\vec{v}_k$ . Wir berechnen die  $k$ -te Spalte des linken Matrixprodukts (komponentenweise), wobei wir verwenden, dass  $B$  eine Diagonalmatrix ist:

$$(SB)_{j,k} = \sum_{r=1}^n S_{j,r} \cdot B_{r,k} = \sum_{r=1}^n S_{j,r} \cdot \delta_{r,k} \lambda_k = \lambda_k S_{j,k} = \lambda_k (\vec{v}_k)_j$$

Somit entspricht die  $k$ -te Spalte des Produkts  $SB$  gerade  $\lambda_k \vec{v}_k$ . Der Vergleich ergibt:

$$\boxed{A\vec{v}_k = \lambda_k \vec{v}_k}$$

Also sind die Spalten der Transformationsmatrix  $S$  genau die Eigenvektoren von  $A$  zu den entsprechenden Eigenwerten, in gleicher Reihenfolge, wie sie in  $B$  als Diagonalelemente auftreten!

Jedoch muss  $S$  außerdem noch invertierbar sein, also regulär – nach Satz 9.2 (Invertierbare Matrizen) bedeutet dies, dass die Spalten von  $S$ , also die Eigenvektoren von  $A$ , linear unabhängig sein müssen. Genau dann bilden sie auch eine Basis des  $n$ -dimensionalen Vektorraums.

Dies ist dann der Fall, wenn alle Eigenräume von  $A$  maximale geometrische Vielfachheit besitzen, wenn also für jeden Eigenwert  $\lambda$  gilt, dass  $\gamma(\lambda) = \alpha(\lambda)$ , mit der algebraischen Vielfachheit  $\alpha(\lambda)$ , die sich aus dem charakteristischen Polynom ergibt. Somit gilt folgender

**Satz 9.29** (Diagonalisierbarkeit). *Eine  $n \times n$ -Matrix  $A$  ist genau dann diagonalisierbar, wenn die Eigenvektoren  $\vec{v}_1, \dots, \vec{v}_n$  zu den Eigenwerten  $\lambda_1, \dots, \lambda_n$  (mit Mehrfachnennung) von  $A$  eine Basis des Vektorraums bilden (bzw. linear unabhängig sind). Die Transformationsmatrix  $S$  hat die Spaltendarstellung*

$$S = (\vec{v}_1 \quad \dots \quad \vec{v}_n)$$

Die zu  $A$  ähnliche Diagonalmatrix ist

$$S^{-1}AS = \text{diag}(\lambda_1, \dots, \lambda_n)$$

**Bemerkung:** (kein Vorlesungsstoff!) Wie im Exkurs A.6 beschrieben, ist  $S^{-1}$  die Matrix für den Basiswechsel von  $E_n$  nach  $\mathcal{S} := \{\vec{v}_1, \dots, \vec{v}_n\}$ . Wir hatten in Satz 9.27 bereits gesehen, dass bei ähnlichen Matrizen die Eigenvektoren von  $S^{-1}AS$  aus den Eigenvektoren von  $A$  hervorgehen, indem man letztere von links mit  $S^{-1}$  multipliziert – dies ist genau solch ein Basiswechsel.

Hier ist  $S^{-1}AS$  diagonal, und nach Satz 9.24 sind die Eigenvektoren der Diagonalmatrix gerade die kartesischen Einheitsvektoren. Dies ist verträglich mit unserer Beobachtung, denn die Eigenvektoren von  $A$  haben, ausgedrückt in der Basis  $\mathcal{S}$  gerade die Struktur kartesischer Einheitsvektoren! Beispielsweise hat  $\vec{v}_1$  in  $\mathcal{S}$  die Form  $(1 \ 0 \ 0 \ \dots)^T$ .

#### Beispiele:

- Wir haben bereits in Unterabschnitt 9.3.4 (S. 257ff.) zur Berechnung von Eigenräumen ein Beispiel kennen gelernt, bei dem die volle geometrische Vielfachheit nicht für alle Eigenwerte gegeben war: Für

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

hatten wir zum doppelten Eigenwert  $\lambda = 2$  nur einen eindimensionalen Eigenraum gefunden. Diese Matrix ist also offenbar *nicht* diagonalisierbar.

- Wir hatten aber auch ein Beispiel für eine dreireihige Matrix, deren Eigenwerte volle geometrische Vielfachheiten hatten. Für

$$A := \begin{pmatrix} 1 & 0 & 2 \\ 0 & 3 & 0 \\ 2 & 0 & 1 \end{pmatrix}$$

hatten wir die Eigenwerte  $\lambda_1 := -1$  und  $\lambda_2 := \lambda_3 := 3$  gefunden, dazu Eigenvektoren

$$\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

Wir stellen nochmals fest, dass diese Eigenvektoren zueinander paarweise orthogonal sind, also sind sie auch sicher linear unabhängig.

Damit ist schon klar, dass  $A$  diagonalisierbar ist. Für die Probe wollen wir  $S^{-1}AS$  berechnen. Da die Eigenvektoren orthogonal zueinander sind, können wir die Berechnung der Inversen  $S^{-1}$  abkürzen, indem wir die Vektoren noch zusätzlich normieren – dann nämlich ist die Matrix  $S$  eine *orthogonale* Matrix (siehe Definition 7.21), und ihr Inverses ist direkt durch die Transponierte gegeben.

Also wählen wir:

$$\vec{v}_1 := \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}; \quad \vec{v}_2 := \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad \text{und} \quad \vec{v}_3 := \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

und insgesamt:

$$S := \frac{1}{\sqrt{2}} \begin{pmatrix} -1 & 0 & 1 \\ 0 & \sqrt{2} & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

(Zur Probe überzeuge man sich, dass  $S^T \cdot S = \mathbb{1}_3$ )

Damit gilt also für die zu  $A$  ähnliche Diagonalmatrix  $B$ :

$$\begin{aligned} B = S^{-1}AS &= S^TAS = \frac{1}{2} \begin{pmatrix} -1 & 0 & 1 \\ 0 & \sqrt{2} & 0 \\ 1 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 2 \\ 0 & 3 & 0 \\ 2 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} -1 & 0 & 1 \\ 0 & \sqrt{2} & 0 \\ 1 & 0 & 1 \end{pmatrix} \\ &= \frac{1}{2} \begin{pmatrix} -1 & 0 & 1 \\ 0 & \sqrt{2} & 0 \\ 1 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 3 \\ 0 & 3\sqrt{2} & 0 \\ -1 & 0 & 3 \end{pmatrix} \\ &= \frac{1}{2} \begin{pmatrix} -2 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 6 \end{pmatrix} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix} \\ &= \text{diag}(-1, 3, 3) \quad \checkmark \end{aligned}$$

### 9.4.3 Zusammenfassung Diagonalisierbarkeit

- Für eine gegebene quadratische  $n \times n$ -Matrix  $A$  bestimmen wir die Eigenwerte.
- Zu jedem Eigenwert  $\lambda$  ermitteln wir den jeweiligen Eigenraum  $E_{A,\lambda}$ .
- Falls die Eigenräume alle maximale Dimension haben (die durch die algebraische Vielfachheit der jeweiligen Eigenwerte gegeben ist), finden wir  $n$  linear unabhängige Eigenvektoren, die eine Basis des Vektorraums bilden. Dann ist die Matrix  $S$ , die diese Eigenvektoren als Spalten enthält (mit gleicher Nummerierung wie die Eigenwerte), regulär, d.h. sie hat vollen Rang und ist invertierbar.
- Und damit ist  $A$  mit einer Diagonalmatrix  $B$  ähnlich (via  $S$ ), die die Eigenwerte von  $A$  in gleicher Nummerierung als Diagonalelemente enthält; also ist  $A$  dann auch diagonalisierbar, und es ist

$$B = S^{-1}AS = \text{diag}(\lambda_1, \dots, \lambda_n)$$

### 9.4.4 Symmetrische Matrizen und Spektralsatz

Für symmetrische reelle Matrizen (siehe Definition 7.15)<sup>4</sup> gelten zwei wichtige Eigenschaften:

**Satz 9.30** (Spektralsatz). *Sei  $A \in \mathbb{R}^{(n,n)}$  symmetrisch, d.h.  $A^T = A$ . Dann gilt:*

- *Sämtliche Eigenwerte von  $A$  sind reell.*
- *Sämtliche Eigenwerte von  $A$  haben volle geometrische Vielfachheit; außerdem sind Eigenvektoren zu verschiedenen Eigenwerten orthogonal.*

(Beweis der zweiten Aussage: S. 350.)

#### Bemerkungen:

- Die erste Aussage lässt sich leicht beweisen; dazu benötigt man jedoch die Rechenregeln für komplexe Zahlen; wir verzichten daher hier auf einen Beweis.
- Die zweite Aussage enthält eine Verschärfung von Satz 9.22 (Eigenvektoren zu verschiedenen Eigenwerten). Dort hieß es nur, dass die Eigenvektoren zu verschiedenen Eigenwerten stets linear unabhängig sind. Für symmetrische Matrizen sind die Eigenräume zu verschiedenen Eigenwerten sogar paarweise orthogonal zueinander.

<sup>4</sup>in den komplexen Zahlen  $\mathbb{C}$  werden hier *hermitesche Matrizen* betrachtet (kein Vorlesungsstoff)

- Aus der zweiten Aussage folgt direkt, dass  $A$  dann auch stets diagonalisierbar ist. Weiterhin lässt sich für jeden Eigenraum mit den Methoden aus Exkurs A.4 (Gram-Schmidt-Orthogonalisierung) eine Orthonormalbasis konstruieren. Insgesamt ist es dann möglich, die Spalten der Transformationsmatrix aus paarweise orthogonalen normierten Eigenvektoren zu bilden, sodass  $S$  eine orthogonale Matrix ist. Dann vereinfacht sich die Ähnlichkeitsbedingung zu

$$B = S^{-1}AS = S^TAS$$

- (kein Vorlesungsstoff in Mathematik 1!) Es ist für verschiedene Situationen sehr hilfreich, eine Basis aus orthonormalen Eigenvektoren zu kennen, z.B. bei der Analyse gekoppelter Schwingungen, in der Quantenmechanik oder bei der Hauptachsentransformation stochastischer Probleme (die Kovarianzmatrix ist symmetrisch, also diagonalisierbar; in der Eigenbasis sind die transformierten Zufallsvariablen entkoppelt; mehr hierzu in späteren Vorlesungen).

**Beispiel:** Wir hatten oben bereits die Diagonalisierung der symmetrischen Matrix

$$\begin{pmatrix} 1 & 0 & 2 \\ 0 & 3 & 0 \\ 2 & 0 & 1 \end{pmatrix}$$

durchgeführt. Dass wir dort die Matrix  $S$  als orthogonale Matrix konstruieren konnten, war in sofern kein Zufall!

# Anhang A

## Exkurse

### A.1 Dieder-Gruppe $D_5$

#### A.1.1 Intro

- “Di-eder” bedeutet Zwei-Flächner – hier (wir betrachten die Gruppe  $D_5$ ) vorstellbar als ein ebenes Stück Pappe in Form eines regulären Fünfecks (Vorder- und Rückseite sind die beiden Flächen), mit nummerierten Ecken.
- Wir stellen die Transformationen auf, die solch ein Fünfeck wieder auf ein deckungsgleiches Fünfeck abbilden. Bezogen auf die nummerierten Ecken sind das Permutationen der Eckenfolge  $(1 - 2 - 3 - 4 - 5)$ . Diese *Symmetrioperationen* sind die Elemente der Gruppe  $D_5$ .
- Achtung: Die *Gruppen-Operation* ist die Hintereinanderausführung  $\circ$ , die die Elemente der Gruppe verknüpft. Bitte nicht mit den Symmetrie-Operationen (also den Elementen) verwechseln!
- Es gibt zehn Symmetrie-Operationen, die ein regelmäßiges Fünfeck auf sich selbst abbilden:
  - 5 Drehungen  $r_0$  bis  $r_4$ , wobei  $r_j$  die Drehung um den Winkel  $j \cdot \frac{2\pi}{5}$  ist (gegen den Uhrzeigersinn)
  - 5 Spiegelungen  $s_1$  bis  $s_5$ , wobei die Spiegelachse für  $s_j$  durch den Punkt  $j$  verläuft und dessen gegenüber liegende Seite senkrecht schneidet

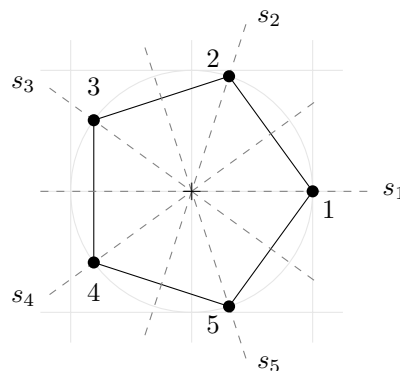


Abbildung A.1: Spiegelachsen für die Diedergruppe  $D_5$

- (Im Fall eines regulären  $n$ -Ecks mit geradem  $n$  verlaufen  $n/2$  Spiegelachsen durch Paare gegenüber liegender Ecken, und  $n/2$  Achsen sind Mittelsenkrechten von gegenüber liegenden Seiten. Auch hier ergeben sich insgesamt  $n$  Symmetrieoperationen für  $D_n$ )

### A.1.2 Permutationen für die Symmetrieoperationen

Zunächst zwei Beispiele für die transformierten Fünfecke nach  $r_3$  und nach  $s_2$ :

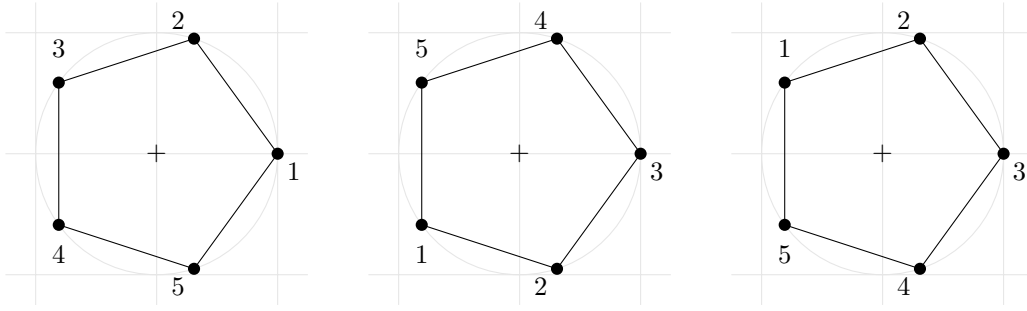


Abbildung A.2: Diedergruppe  $D_5$ : Original (links), Rotation  $r_3$  (Mitte) und Spiegelung  $s_2$  (rechts)

Wir geben die Permutationen für die zehn Operationen an – rechts in Zykelschreibweise; die Tabellenschreibweisen links daneben sind nur verkürzt als Wertetabellen notiert:

$j$	1	2	3	4	5	
$r_0(j)$	1	2	3	4	5	$(1)(2)(3)(4)(5)$
$r_1(j)$	5	1	2	3	4	$(1\ 5\ 4\ 3\ 2)$
$r_2(j)$	4	5	1	2	3	$(1\ 4\ 2\ 5\ 3)$
$r_3(j)$	3	4	5	1	2	$(1\ 3\ 5\ 2\ 4)$
$r_4(j)$	2	3	4	5	1	$(1\ 2\ 3\ 4\ 5)$
$s_1(j)$	1	5	4	3	2	$(1)(2\ 5)(3\ 4)$
$s_2(j)$	3	2	1	5	4	$(1\ 3)(2)(4\ 5)$
$s_3(j)$	5	4	3	2	1	$(1\ 5)(2\ 4)(3)$
$s_4(j)$	2	1	5	4	3	$(1\ 2)(3\ 5)(4)$
$s_5(j)$	4	3	2	1	5	$(1\ 4)(2\ 3)(5)$

Offenbar sind dies alle Permutationen auf der Menge  $M_5$ , d.h.  $D_5 \subseteq S_5$ .

### A.1.3 Nachweis der Gruppeneigenschaft

Es gibt  $5! = 120$  verschiedene Permutationen in der Gruppe  $S_5$ . Wir zeigen nun, dass  $D_5 := \{r_0, r_1, r_2, r_3, r_4, s_1, s_2, s_3, s_4, s_5\}$  eine Untergruppe von  $S_5$  ist; in Klammern notieren wir die Hierarchiestufen, die unterwegs erreicht werden.

- Die Hintereinanderausführung  $\circ$  ist in  $D_5$  abgeschlossen (Magma; siehe nächster Unterabschnitt).
- Die Hintereinanderausführung  $\circ$  von Permutationen ist assoziativ (Halbgruppe).
- Die Symmetrieoperation  $r_0$  entspricht dem neutralen Element  $\text{id}_5 \in S_5$  (Monoid).
- Zu jeder Symmetrieoperation in  $D_5$  gibt es ein Inverses aus  $D_5$  (Gruppe):
  - Für die Rotationen ist  $r_j^{-1} = r_{(5-j) \bmod 5}$ , also z.B.  $r_3^{-1} = r_2$
  - Die Spiegelungen sind selbst-invers:  $s_j^{-1} = s_j$ .

Damit ist  $D_5$  eine Gruppe, und wegen  $D_5 \subseteq S_5$  auch eine Untergruppe von  $S_5$ .

### A.1.4 Abgeschlossenheit

Scharfes Hinsehen zeigt im Beispiel von Abbildung A.2, dass man die Rotation  $r_3$  aus der Spiegelung  $s_2$  erhalten kann, indem man auf diese nochmals  $s_3$  anwendet, d.h.  $r_3 = s_3 \circ s_2$ . Gleichfalls ergibt die Spiegelung  $s_3$ , angewendet auf  $r_3$ , die Spiegelung  $s_2$ , also  $s_2 = s_3 \circ r_3$ . Man erhält die zusätzliche Anwendung von  $s_3$ , indem man die Diagramme in der Mitte und rechts jeweils in Gedanken an der Horizontalen spiegelt.



Für die Drehungen gilt:  $r_j \circ r_k = r_k \circ r_j = r_{(j+k) \bmod 5}$ , d.h. die Drehungen sind untereinander kommutativ.

Spiegelungen an verschiedenen Achsen kommutieren allerdings nicht.

Wenn man zwei Spiegelungen verknüpft, ändert sich zwei Mal der Umlaufsinn der Eckenfolge  $(1 - 2 - 3 - 4 - 5)$ , d.h. das Fünfeck wird zwei Mal geflippt. Danach liegt wieder die ursprüngliche Seite oben, bzw. der ursprüngliche Umlaufsinn der Eckenfolge vor.

Damit  $D_5$  eine Gruppe sein kann, muss also die Kombination von zwei Spiegelungen eine Drehung ergeben (weil außer den Spiegelungen nur Drehungen zur Verfügung stehen, wenn  $\circ$  eine abgeschlossene Operation in  $D_5$  sein soll). Wir hatten bereits oben beobachtet, dass  $s_3 \circ s_2 = r_3$ .

Wir rechnen exemplarisch noch die Verknüpfungen zwischen  $s_1$  und  $s_4$  nach:

$j$	1	2	3	4	5	
$s_1(j)$	1	5	4	3	2	
$s_4(j)$	2	1	5	4	3	
$(s_1 \circ s_4)(j)$	5	1	2	3	4	$s_1 \circ s_4 = r_1$
$(s_4 \circ s_1)(j)$	2	3	4	5	1	$s_4 \circ s_1 = r_4$

Kombiniert man eine Drehung und eine Spiegelung, so wird der Umlaufsinn der Eckenfolge einmal umgekehrt. Das Resultat einer solchen Kombination muss also eine Spiegelung sein.

Wir rechnen dies für die Verknüpfungen zwischen  $s_2$  und  $r_4$  nach:

$j$	1	2	3	4	5	
$s_2(j)$	3	2	1	5	4	
$r_4(j)$	2	3	4	5	1	
$(s_2 \circ r_4)(j)$	2	1	5	4	3	$s_2 \circ r_4 = s_4$
$(r_4 \circ s_2)(j)$	4	3	2	1	5	$r_4 \circ s_2 = s_5$

### A.1.5 Gruppentafel von $D_5$

Die restlichen Rechnungen erfolgen analog; es ergibt sich die folgende Verknüpfungstafel:

$\circ$	$r_0$	$r_1$	$r_2$	$r_3$	$r_4$	$s_1$	$s_2$	$s_3$	$s_4$	$s_5$
$r_0$	$r_0$	$r_1$	$r_2$	$r_3$	$r_4$	$s_1$	$s_2$	$s_3$	$s_4$	$s_5$
$r_1$	$r_1$	$r_2$	$r_3$	$r_4$	$r_0$	$s_3$	$s_4$	$s_5$	$s_1$	$s_2$
$r_2$	$r_2$	$r_3$	$r_4$	$r_0$	$r_1$	$s_5$	$s_1$	$s_2$	$s_3$	$s_4$
$r_3$	$r_3$	$r_4$	$r_0$	$r_1$	$r_2$	$s_2$	$s_3$	$s_4$	$s_5$	$s_1$
$r_4$	$r_4$	$r_0$	$r_1$	$r_2$	$r_3$	$s_4$	$s_5$	$s_1$	$s_2$	$s_3$
$s_1$	$s_1$	$s_4$	$s_2$	$s_5$	$s_3$	$r_0$	$r_2$	$r_4$	$r_1$	$r_3$
$s_2$	$s_2$	$s_5$	$s_3$	$s_1$	$s_4$	$r_3$	$r_0$	$r_2$	$r_4$	$r_1$
$s_3$	$s_3$	$s_1$	$s_4$	$s_2$	$s_5$	$r_1$	$r_3$	$r_0$	$r_2$	$r_4$
$s_4$	$s_4$	$s_2$	$s_5$	$s_3$	$s_1$	$r_4$	$r_1$	$r_3$	$r_0$	$r_2$
$s_5$	$s_5$	$s_3$	$s_1$	$s_4$	$s_2$	$r_2$	$r_4$	$r_1$	$r_3$	$r_0$

Man stellt fest, dass in der Tat die Hintereinanderausführungen der Symmetrieoperationen in  $D_5$  abgeschlossen ist. Also ist  $D_5$  eine Gruppe, und damit auch eine nichttriviale Untergruppe von  $S_5$ . Sie ist nicht abelsch, wie oben bereits im Beispiel nachgewiesen.

Wir finden außerdem einige Untergruppen:

- Zunächst die triviale Untergruppe  $\{r_0\}$
- Dann fünf (abelsche) Untergruppen mit je zwei Elementen, bestehend aus der Identität und einer der Spiegelungen:  $\{r_0, s_j\}$  für  $j \in \{1, 2, 3, 4, 5\}$
- Und eine Untergruppe mit fünf Elementen, nämlich die (abelsche) Untergruppe der fünf Rotationen:  $\{r_0, r_1, r_2, r_3, r_4\}$

### A.1.6 Nebenbemerkung: Die restlichen Permutationen in $S_5$

Die Gruppe  $S_5$  enthält 120 Permutationen, davon die Untergruppe  $D_5$  nur zehn. Dies liegt daran, dass alle Operationen in  $D_5$  die Eckenfolge  $(1-2-3-4-5)$  beibehalten (höchstens der Umlaufsinn ändert sich, falls eine Spiegelung erfolgt). Eine Eckenfolge wie  $(1-4-2-5-3)$  ließe sich so beispielsweise nicht herstellen.

Erlaubt man aber allgemeine Permutationen, so lässt sich die Eckenfolge auf vielfältige Art aufbrechen. Dies ist (in der Vorstellung) mit einem festen Stück Pappe nicht mehr zu machen. Hier stellt man sich besser eine elastische Schnur vor, die alle fünf Ecken erreicht. Damit lassen sich Teile der Figur gegenüber dem Rest verdrillen. Es ergeben sich unter anderem die folgenden möglichen Objekte:

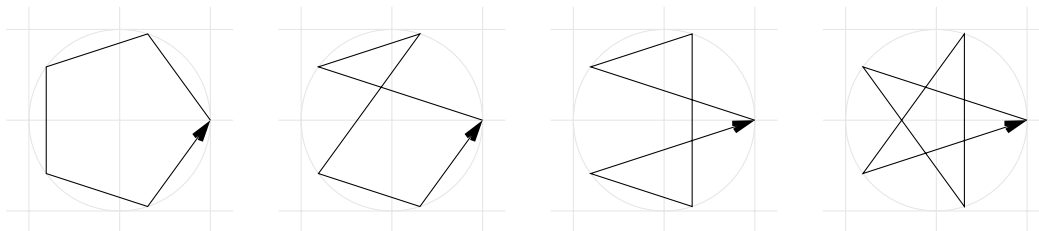


Abbildung A.3: Vier Grundfiguren des Pentagons in  $S_5$ , mit je null, einer, zwei und drei Verdrillungen

Für  $S_5$  ergeben sich zwölf verschiedene Grundfiguren (abhängig davon, wo die Eckenfolge beginnt – in der oben gezeigten Auswahl gibt es für die mittleren beiden Figuren noch jeweils vier Alternativen). Alle diese Figuren können im Umlaufsinn umgekehrt und/oder auf fünf Arten rotiert werden – das sind jeweils wieder die  $D_5$ -Symmetrieoperationen. So erhält man insgesamt 120 verschiedene Permutationen.

So lässt sich die Bezeichnung “Symmetrische Gruppe” für  $S_n$  motivieren. Im Sonderfall  $n = 3$  gibt es jedoch keine Verdrillungen, die nicht als Spiegelungen ausdrückbar sind – deswegen sind  $S_3$  und  $D_3$  quasi identisch.

## A.2 Horner-Schema

### A.2.1 Polynom-Auswertung

Wir berechnen für die reelle Polynomfunktion  $f(x) := 1x^5 + 4x^4 + 2x^3 + 1x^2 + 9x + 4$  den Wert  $f(2)$  auf drei Arten:

- (naive Variante:)

$$f(2) = 1 \cdot (2 \cdot 2 \cdot 2 \cdot 2 \cdot 2) + 4 \cdot (2 \cdot 2 \cdot 2 \cdot 2) + 2 \cdot (2 \cdot 2 \cdot 2) + 1 \cdot (2 \cdot 2) + 9 \cdot 2 + 4 = 138$$

Dafür wurden benötigt:  $2^2 = 4$ ,  $2^3 = 2 \cdot 4 = 8$ , etc.; das sind 10 Multiplikationen für die Potenzen von  $x = 2$ . Dazu noch fünf, um die Koeffizienten einzurechnen.

- (geschickter, mit Zwischenergebnissen:)  $x^2 = 4$ ,  $x^3 = x^2 \cdot x = 4 \cdot 2 = 8$ ,  $x^4 = x^2 \cdot x^2 = 4 \cdot 4 = 16$ ,  $x^5 = x^4 \cdot x = 16 \cdot 2 = 32$ . Dazu die fünf Koeffizienten-Multiplikationen, ergibt insgesamt neun Multiplikationen (anstatt 15):

$$f(2) = 2 \cdot (16 \cdot 2) + 4 \cdot (4 \cdot 4) + 2 \cdot (4 \cdot 2) + 1 \cdot (2 \cdot 2) + 9 \cdot 2 + 4 = 138$$

- (Horner-Schema:) Wir schreiben das Polynom um, sodass jede Potenz nur genau einmal berechnet werden muss; das ergibt fünf Multiplikationen.

$$\begin{aligned} f(x) &= 1x^5 + 4x^4 + 2x^3 + 1x^2 + 9x + 4 \\ &= (((((1 \cdot x + 4) \cdot x + 2) \cdot x + 1) \cdot x + 9) \cdot x + 4 \end{aligned}$$

Setzt man nun einen konkreten Wert  $x$  ein (hier also 2), so starten wir mit dem höchsten Koeffizienten. Ab dann wird abwechselnd mit  $x$  multipliziert und der nächstkleinere Koeffizient addiert, wobei das Zwischenergebnis gespeichert wird. Man bewegt sich von links (innerste Klammer) bis nach rechts.

Wir erhalten für  $x = 2$ :

$$\begin{aligned}
 f(2) &= (((((1 \cdot 2 + 4) \cdot 2 + 2) \cdot 2 + 1) \cdot 2 + 9) \cdot 2 + 4 \\
 &= (((((2 + 4) \cdot 2 + 2) \cdot 2 + 1) \cdot 2 + 9) \cdot 2 + 4 \\
 &= (((6 \cdot 2 + 2) \cdot 2 + 1) \cdot 2 + 9) \cdot 2 + 4 \\
 &= (((12 + 2) \cdot 2 + 1) \cdot 2 + 9) \cdot 2 + 4 \\
 &= ((14 \cdot 2 + 1) \cdot 2 + 9) \cdot 2 + 4 \\
 &= ((28 + 1) \cdot 2 + 9) \cdot 2 + 4 \\
 &= (29 \cdot 2 + 9) \cdot 2 + 4 \\
 &= (58 + 9) \cdot 2 + 4 \\
 &= 67 \cdot 2 + 4 \\
 &= 134 + 4 \\
 &= 138
 \end{aligned}$$

Diese “mechanische” Abfolge von Multiplikationen und Addition lässt sich gut in einem Tabellenschema ausführen. In der ersten Zeile trägt man die Koeffizienten der Polynomfunktion ein (pro Spalte einen). Die zweite Zeile bleibt zunächst leer. Dann folgt ein Bilanzstrich; in der dritten Zeile wird in linker Spalte der erste Koeffizient von Zeile 1 übernommen. Das Schema für unser Beispiel sieht zu Anfang also so aus:

$$\begin{array}{cccccc}
 1 & 4 & 2 & 1 & 9 & 4 \\
 \hline
 & & & & & \\
 1 & & & & & 
 \end{array}$$

Nun multiplizieren wir den rechten (momentan: einzigen) Eintrag der dritten Zeile mit  $x$  (hier mit 2) und tragen ihn eine Spalte weiter rechts in die zweite Zeile ein:

$$\begin{array}{cccccc}
 1 & 4 & 2 & 1 & 9 & 4 \\
 & 2 & & & & \\
 \hline
 1 & & & & & 
 \end{array}$$

Dann addieren wir zu diesem Produktwert den Koeffizienten aus Zeile 1 der entsprechenden Spalte; die Summe ergibt (in gleicher Spalte) den neuen Zwischenwert unter dem Bilanzstrich:

$$\begin{array}{cccccc}
 1 & 4 & 2 & 1 & 9 & 4 \\
 & 2 & & & & \\
 \hline
 1 & 6 & & & & 
 \end{array}$$

Für den nächsten Schritt fassen wir beide Operationen direkt zusammen und erhalten:

$$\begin{array}{cccccc}
 1 & 4 & 2 & 1 & 9 & 4 \\
 & 2 & 12 & & & \\
 \hline
 1 & 6 & 14 & & & 
 \end{array}$$

Und in dieser Art weiter. Am Schluss steht in der rechten Spalte unter dem Bilanzstrich der gesuchte Funktionswert:

$$\begin{array}{cccccc}
 1 & 4 & 2 & 1 & 9 & 4 \\
 & 2 & 12 & 28 & 58 & 134 \\
 \hline
 1 & 6 & 14 & 29 & 67 & 138
 \end{array}$$

Wir erkennen in der Tabelle (in der Reihenfolge, wie wir sie ausgefüllt haben) genau die Zahlen wieder, die in obiger ausführlicher Berechnung jeweils ganz links notiert waren

Achtung: Bei diesen Rechnungen schleichen sich oft Flüchtigkeitsfehler ein; es lohnt sich also, die Einträge nochmal zu kontrollieren.

## A.2.2 Stellenwertsysteme

Eine besondere Rolle spielt das Horner-Schema bei der Auswertung von Zahlen in Stellenwertsystemen, also der Transformation vom  $b$ -er-System ins Dezimalsystem.

Jede Zahl  $n \in \mathbb{N}_0$  zur Basis  $b$  hat die Struktur

$$n = a_k b^k + a_{k-1} b^{k-1} + \dots + a_2 b^2 + a_1 b + a_0 = \sum_{j=0}^k a_j b^j$$

Dabei sind die *Ziffern*  $a_j$  aus  $\{0, 1, \dots, b-1\}$ , und das Ziffern-Tupel (und mit dessen Länge auch der Wert von  $k$ ) ist für jedes  $n$  eindeutig bestimmt (dies akzeptieren wir hier ohne Beweis<sup>1</sup>).

Nun steht aber oben nichts anderes als ein spezieller Wert einer Polynomfunktion mit den Koeffizienten  $a_j$ , ausgewertet an der Stelle  $b$ .

### Beispiele:

- Zur Basis  $b := 2$  berechnen wir  $n := 1101_{(2)}$ :

$$1101_{(2)} = 1 \cdot 2^3 + 1 \cdot 2^2 + 0 \cdot 2^1 + 1 \cdot 2^0 = 8 + 4 + 1 = 13$$

Dann gilt aber mit  $f(x) := 1 \cdot x^3 + 1 \cdot x^2 + 0 \cdot x + 1$ :

$$f(2) = 1 \cdot 2^3 + 1 \cdot 2^2 + 0 \cdot 2 + 1 = 8 + 4 + 1 = 13$$

- Zur Basis  $b := 7$  werten wir  $n := 3152_{(7)}$  aus:

$$3152_{(7)} = 3 \cdot 7^3 + 1 \cdot 7^2 + 5 \cdot 7^1 + 2 \cdot 7^0 = \dots = 1115$$

Die zugehörige Polynomfunktion lautet (etwas kürzer notiert):  $f(x) := 3x^3 + x^2 + 5x + 2$ . Dann ist:

$$f(7) = \dots = 1115$$

Zu den ausgelassenen Rechnungen siehe unten.

Wir werten nun die Polynomfunktionen der obigen beiden Beispiele mit dem Hornerschema aus. Die Koeffizienten entsprechen gerade den Ziffern der Zahlen zur Basis  $b$ ; und der Funktionswert wird an der Stelle  $b$  genommen (im Dezimalsystem  $b = 10$  funktioniert all dies natürlich ebenfalls!):

- Berechnung von  $1101_{(2)}$ :

$$\begin{array}{r} 1 \quad 1 \quad 0 \quad 1 \\ \quad 2 \quad 6 \quad 12 \\ \hline 1 \quad 3 \quad 6 \quad 13 \end{array}$$

Also:  $1101_{(2)} = 13$

- Berechnung von  $3152_{(7)}$ :

$$\begin{array}{r} 3 \quad 1 \quad 5 \quad 2 \\ \quad 21 \quad 154 \quad 1113 \\ \hline 3 \quad 22 \quad 159 \quad 1115 \end{array}$$

Also:  $3152_{(7)} = 1115$

## A.3 Geraden und Ebenen

Wir befassen uns hier mit verschiedenen Darstellungen von Geraden und Ebenen und gehen in wachsender Komplexität vor.

<sup>1</sup>Die Existenz einer Entwicklung lässt sich mit vollständiger Induktion beweisen; siehe Mathematik 2/Analysis. Die Eindeutigkeit kann man mit den Mitteln von Mathematik 1 zeigen: Man nimmt an, eine Zahl  $n$  hätte zwei verschiedene Entwicklungen zur Basis  $b$ . Dann bildet man die Differenz der zugehörigen Polynomfunktionen, wertet die resultierende Funktion an der Stelle  $b$  aus und erhält einen Wert ungleich 0 (indirekter Beweis).

### A.3.1 Geraden in $\mathbb{R}^2$

Aus der Schule ist bekannt, dass die Graphen der affin-linearen Funktionen

$$x \mapsto y := mx + n$$

Geraden sind, wobei  $m$  die Steigung einer solchen Gerade ist, und  $n$  die Stelle, an der die Gerade die  $y$ -Achse schneidet. Fast alle Geraden in  $\mathbb{R}^2$  lassen sich so ausdrücken – bis auf die Geraden  $x = \text{const} = c$  mit  $c \in \mathbb{R}$ . Diese stehen senkrecht auf der  $x$ -Achse und hätten damit unendlich große (oder kleine) Steigung.

Wir können Geraden aber auch vektoriell ausdrücken und dabei dieses Problem umgehen – damit nehmen wir aber in Kauf, dass nicht länger alle Geraden auch Funktionsgraphen sind.

Wir starten mit der *Parameterform*. Hierbei werden alle Punkte  $\vec{x}$  einer Geraden auf einen bestimmten Punkt mit dem Ortsvektor  $\vec{x}_0$  bezogen. Die Verbindungsvektoren  $\vec{x} - \vec{x}_0$  liegen dann alle auf dieser Geraden, wie in Abbildung A.4 gezeigt.

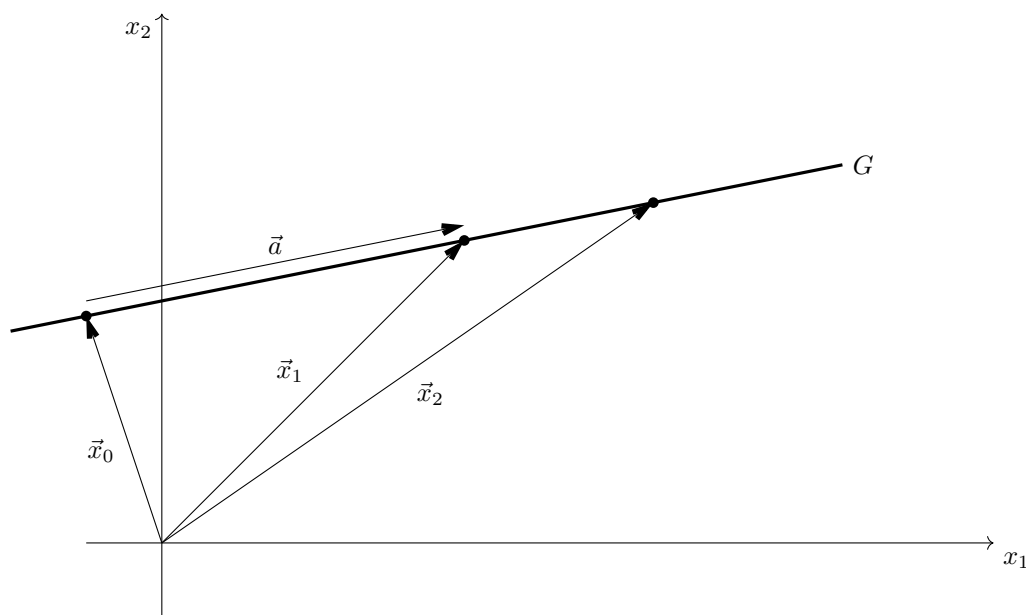


Abbildung A.4: Gerade in Parameterform (Richtungsvektor  $\vec{a}$  leicht versetzt eingezeichnet)

Nun ist z.B.  $\vec{x}_1 = \vec{x}_0 + \vec{a}$ . Dann erreichen wir alle weiteren Punkte auf der Geraden  $G$ , indem wir von  $\vec{x}_0$  aus sämtliche Skalierungen von  $\vec{a}$  addieren. Dann ist auch  $\vec{x}_2 = \vec{x}_0 + t\vec{a}$  für ein bestimmtes  $t \in \mathbb{R}$ .

Mit der Kenntnis von zwei Punkten auf  $G$  erhalten wir also den Richtungsvektor durch Subtraktion. Dann können wir einen beliebigen Punkt von  $G$  als *Aufpunkt* (oder: *Stützpunkt*) wählen, durchaus auch einen der beiden bereits verwendeten. Die Ortsvektoren der Punkte auf der Geraden lassen sich dann durch Variation des *Parameters*  $t$  erhalten, mit der Gleichung

$$\vec{x} - \vec{x}_0 = t\vec{a}$$

Das führt uns direkt auf folgende

**Definition A.1** (Parameterform von Geraden). *Gegeben ein Vektor  $\vec{a} \in \mathbb{R}^2 \setminus \{\vec{0}\}$  sowie ein spezieller Punkt  $\vec{x}_0 \in \mathbb{R}^2$ . Dann ist die Gerade  $G(\vec{x}_0, \vec{a})$  durch den Punkt mit Ortsvektor  $\vec{x}_0$  mit Richtungsvektor  $\vec{a}$  die Punktmenge*

$$G(\vec{x}_0, \vec{a}) := \{\vec{x}_0 + t\vec{a} \mid t \in \mathbb{R}\}$$

Mit dieser Darstellung lassen sich sämtliche Geraden in  $\mathbb{R}^2$  beschreiben, inklusive derer, die keine Funktionsgraphen  $x_2 = f(x_1)$  sein können – letztere hätten Richtungsvektoren mit 0 als erster Komponente. Erkauft haben wir uns das damit, dass wir den Schnittpunkt mit der  $x_2$ -Achse nicht mehr direkt ablesen können, sondern explizit berechnen müssten – und auch das Konzept der Steigung fehlt.

Man kann Steigung und Achsenabschnitt jedoch aus der Punkt-Richtungsform wieder errechnen, falls denn die Gerade ein Funktionsgraph ist (also  $a_1 \neq 0$  gilt). Dazu schreiben wir die Vektorgleichung als ein System von zwei reellen Gleichungen um und eliminieren den Parameter  $t$  durch Einsetzen:

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} := \vec{x} = \vec{x}_0 + t\vec{a} =: \begin{pmatrix} p_1 \\ p_2 \end{pmatrix} + t \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$$

$$\Leftrightarrow \quad \wedge \quad \begin{array}{rcl} x_1 & = & p_1 + ta_1 \\ x_2 & = & p_2 + ta_2 \end{array}$$

$$\begin{array}{l} \xrightarrow{a_1 \neq 0} \quad t = \frac{p_1 - x_1}{a_1} \\ \wedge \quad x_2 = p_2 + ta_2 = p_2 + a_2 \frac{p_1 - x_1}{a_1} = \frac{a_2}{a_1} x_1 + \left( p_2 - \frac{p_1 a_2}{a_1} \right) \end{array}$$

In der letzten Umformung steht ganz rechts wieder die affin-lineare Gleichung mit Steigung und Achsenabschnitt.

**Beispiel:** Für die Punkte

$$\vec{x}_0 := \begin{pmatrix} -1 \\ 3 \end{pmatrix} \quad \text{und} \quad \vec{x}_1 := \begin{pmatrix} 4 \\ 4 \end{pmatrix}$$

erhalten wir einen Richtungsvektor (es gibt beliebig viele andere, die gleichwertig wären) von

$$\vec{a} := \vec{x}_1 - \vec{x}_0 = \begin{pmatrix} 5 \\ 1 \end{pmatrix}$$

Mit  $\vec{x}_0$  als Aufpunkt gilt für die Gerade  $G$  also mit reellem Parameter  $t$ :

$$\vec{x}(t) = \begin{pmatrix} -1 \\ 3 \end{pmatrix} + t \begin{pmatrix} 5 \\ 1 \end{pmatrix}$$

Dann finden wir  $\vec{x}_0 = \vec{x}(0)$  und  $\vec{x}_1 = \vec{x}(1)$  mit den Werten 0 und 1 für den Parameter  $t$ . Ein weiterer Punkt auf der Geraden hätte z.B. mit  $t = \frac{3}{2}$  den Ortsvektor

$$\vec{x}_2 := \vec{x}\left(\frac{3}{2}\right) = \begin{pmatrix} -1 \\ 3 \end{pmatrix} + \frac{3}{2} \begin{pmatrix} 5 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 + \frac{15}{2} \\ 3 + \frac{3}{2} \end{pmatrix} = \begin{pmatrix} \frac{13}{2} \\ \frac{9}{2} \end{pmatrix}$$

Die Steigung dieser Geraden errechnet sich nach obiger Umformung per

$$m := \frac{a_2}{a_1} = \frac{1}{5}$$

Und der Achsenabschnitt bei  $x_1 = 0$  wäre

$$n := p_2 - \frac{p_1 a_2}{a_1} = 3 - \frac{-1}{5} = \frac{16}{5}$$

Also lautet die affin-lineare Geradengleichung:

$$x_2 = \frac{1}{5}x_1 + \frac{16}{5}$$

Zur Probe rechnet man nach, dass der Punkt mit Ortsvektor

$$\begin{pmatrix} 0 \\ \frac{16}{5} \end{pmatrix}$$

tatsächlich auf der Geraden liegt, indem man den Ortsvektor des Aufpunkts subtrahiert: der resultierende Vektor ist dann ein Vielfaches von  $\vec{a}$ , nämlich mit dem Faktor  $t = \frac{1}{5}$ .

Im übrigen hätten wir diesen Achsenabschnitt auch direkt erhalten, wenn wir auf die allgemeine Umformung oben verzichtet hätten – man kann nämlich gleichwertig den Punkt mit Ortsvektor

$$\begin{pmatrix} 0 \\ y \end{pmatrix}$$

auf der Geraden bestimmen, indem man ihn einsetzt und dann den Wert  $t$  ermittelt:

$$\begin{pmatrix} 0 \\ y \end{pmatrix} = \begin{pmatrix} -1 \\ 3 \end{pmatrix} + t \begin{pmatrix} 5 \\ 1 \end{pmatrix} \Rightarrow 0 = -1 + 5t \Leftrightarrow 1 = 5t \Leftrightarrow t = \frac{1}{5}$$

Mit diesem  $t$  ist dann (hier benutzen wir die Gleichung der zweiten Komponente):

$$y = 3 + t = 3 + \frac{1}{5} = \frac{16}{5}$$

Führt dieses Verfahren zu keiner Lösung, dann existiert ein Achsenabschnitt für die  $x_2$ -Achse nicht.

Hat eine Gerade  $G$  in Parameterform bestimmt, so lässt sich die Vektorgleichung von  $G$  (welche wir oben schon als ein System von zwei reellen Gleichungen umgeschrieben hatten) in eine Gleichung für die beiden Koordinaten  $x_1, x_2$  umformen. Die erste Hälfte dieser Umformung hatten wir oben bereits ausgeführt, als wir die  $x_1$ -Komponente der Vektorgleichung nach  $t$  aufgelöst hatten. Das gleiche können wir auch mit der  $x_2$ -Komponente tun und erhalten, da  $t = t$  gilt, mit den selben Bezeichnern wie oben:

$$(t =) \quad \frac{p_1 - x_1}{a_1} = \frac{p_2 - x_2}{a_2}$$

Wir multiplizieren die Gleichung mit  $a_1 a_2$  (zunächst gehen wir davon aus, dass keine der beiden Komponenten von  $\vec{a}$  null ist):

$$\dots \Leftrightarrow a_2 p_1 - a_2 x_1 = a_1 p_2 - a_1 x_2 \Leftrightarrow (-a_2)x_1 + a_1 x_2 = a_1 p_2 - a_2 p_1$$

Wir erinnern uns, dass in  $\mathbb{R}^2$  nur zwei Normalenvektoren zu  $\vec{a}$  mit gleicher Länge existieren – in obiger Gleichung taucht gerade einer davon auf (derjenige, der durch Drehung in positivem Umlaufsinn um  $\frac{\pi}{2}$  erreicht wird):

$$\dots \Leftrightarrow \begin{pmatrix} -a_2 \\ a_1 \end{pmatrix} \bullet \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -a_2 \\ a_1 \end{pmatrix} \bullet \begin{pmatrix} p_1 \\ p_2 \end{pmatrix}$$

Dies ist die

**Satz A.2** (Normalenform von Geraden in  $\mathbb{R}^2$ ). Eine Gerade  $G \subset \mathbb{R}^2$  mit Aufpunkt  $\vec{x}_0$  und Richtungsvektor  $\vec{a} \neq \vec{0}$  erfüllt die Gleichung

$$\vec{n}_a \bullet \vec{x} = \vec{n}_a \bullet \vec{x}_0,$$

wobei der Normalenvektor definiert ist als

$$\vec{n}_a := \begin{pmatrix} -a_2 \\ a_1 \end{pmatrix}$$

### Bemerkungen:

- Der Aufpunkt ist beliebig wählbar. Genau für Ursprungsgeraden ist auch  $\vec{x}_0 = \vec{0}$  möglich. In diesem Fall lautet die Normalenform dann  $\vec{n}_a \bullet \vec{x} = 0$ .
- Äquivalent zur Normalenform ist die *Koordinatenform*

$$kx_1 + lx_2 = d$$

Hierbei sind  $k, l$  die Komponenten des obigen Normalenvektors; es liegt eine lineare Gleichung mit zwei Unbekannten vor.

Auch hier lassen sich sämtliche Geraden aus  $\mathbb{R}^2$  ausdrücken.

(Es ist  $k = (-a_2)$ ,  $l = a_1$  und  $d = \vec{n}_a \bullet \vec{x}_0$  mit den Bezeichnern aus der Parameterform.)

- Die Koordinaten- und Normalenform der Geradengleichung bleiben auch richtig, wenn eine der Komponenten von  $\vec{a}$  null ist – beide gleichzeitig können nicht null werden, da sonst die Gerade gar keine Richtung hätte. Die hier angesprochenen Fälle beschreiben Geraden, die parallel zu einer der beiden Koordinatenachsen sind.

Ist  $a_1 = 0$ , so ist die Gerade parallel zur  $x_2$ -Achse, und die Normalenform lautet

$$(-a_2)x_1 = (-a_2)p_1 \Leftrightarrow x_1 = p_1$$

für (irgend)einen Aufpunkt mit erster Koordinate  $p_1$ . (Diese Geraden lassen sich nicht als Funktionsgraphen ausdrücken.) Analog bei  $a_2 = 0$ : Diese Geraden sind parallel zur  $x_1$ -Achse, mit der Normalenform

$$a_1x_2 = a_1p_2 \Leftrightarrow x_2 = p_2$$

für (irgend)einen Aufpunkt mit zweiter Koordinate  $p_2$ .

- Aus einer Koordinatenform (bzw. Normalenform) lässt sich die Punkt-Richtungsform wieder extrahieren. Falls der Normalenvektor gegeben ist als

$$\vec{n} = \begin{pmatrix} k \\ l \end{pmatrix},$$

ergibt sich der Richtungsvektor direkt durch Drehung um  $-\frac{\pi}{2}$  per

$$\vec{a} := \begin{pmatrix} l \\ -k \end{pmatrix}$$

Für die Koordinaten des Aufpunkts kann man in die Koordinatengleichung irgendeinen Punkt der Geraden wählen. Legt man die eine Koordinate fest, so folgt aus der Gleichung eindeutig die andere (oder es ist, falls eine Komponente des Normalenvektors null war, gar keine weitere Rechnung nötig).

**Beispiel:** Wir betrachten die Gerade mit der Koordinatengleichung  $3x_1 + 4x_2 = 5$ . Ihr Normalenvektor lautet:

$$\vec{n} = \begin{pmatrix} 3 \\ 4 \end{pmatrix}$$

Damit bestimmen wir den Richtungsvektor der Geraden als

$$\vec{a} = \begin{pmatrix} 4 \\ -3 \end{pmatrix}$$

Für den Aufpunkt könnten wir willkürlich eine der Koordinaten in der Gleichung auf null setzen. Mit  $x_1 := 0$  erhalten wir  $x_2 = \frac{5}{4}$ . Alternativ folgt mit  $x_2 := 0$ , dass  $x_1 = \frac{5}{3}$ . Zwei mögliche Aufpunkte sind also:

$$\vec{x}_0 := \begin{pmatrix} 0 \\ \frac{5}{4} \end{pmatrix} \quad \text{oder} \quad \vec{x}_0 := \begin{pmatrix} \frac{5}{3} \\ 0 \end{pmatrix}$$

Eine mögliche Parameterform der Geraden ist also:

$$\vec{x}(t) = \begin{pmatrix} 0 \\ \frac{5}{4} \end{pmatrix} + t \begin{pmatrix} 4 \\ -3 \end{pmatrix}$$

Natürlich dürfte man auch Aufpunkte wählen, die nicht auf den Koordinatenachsen liegen. Für  $x_1 := 7$  würde z.B. folgen, dass  $x_2 = (5 - 21)/4 = (-4)$ . Eine alternative Parameterform der Geraden wäre damit:

$$\vec{x}(t) = \begin{pmatrix} 7 \\ -4 \end{pmatrix} + t \begin{pmatrix} 4 \\ -3 \end{pmatrix}$$

Wir betrachten nochmals die Normalenform einer Geraden:

$$\vec{n} \bullet \vec{x} = \vec{n} \bullet \vec{x}_0 = d$$

Die Situation ist in Abbildung A.5 gezeigt.

Wir erinnern uns, dass das Skalarprodukt (und das kanonische Skalarprodukt “ $\bullet$ ” ist ja ein solches) mit dem Vektor  $\vec{n}$  die zu  $\vec{n}$  orthogonale Komponente eines Vektors heraus projiziert – es verschwinden also alle Anteile proportional zum Richtungsvektor der Geraden!



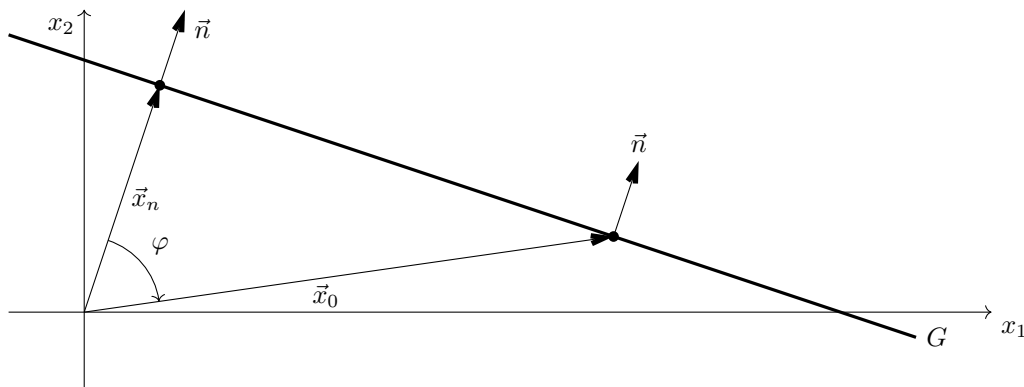


Abbildung A.5: Zur Hesse-Normalform einer Geraden

Der Punkt auf der Geraden, der den geringsten Abstand zum Koordinatenursprung hat, ist der Fußpunkt des Lotes vom Ursprung zur Geraden  $G$ . In der Abbildung ist dessen Ortsvektor mit  $\vec{x}_n$  bezeichnet, weil genau dieser Vektor parallel zu  $\vec{n}$  ist.

Falls der (euklidische) Abstand der Geraden vom Nullpunkt interessant ist, wäre die Länge  $|\vec{x}_n|$  zu berechnen. Wegen der Parallelität von  $\vec{x}_n$  mit  $\vec{n}$  taugt hierfür das Skalarprodukt:

$$\vec{n} \bullet \vec{x}_n = |\vec{n}| \cdot |\vec{x}_n|$$

Und wegen der Geradengleichung gilt weiterhin, da der Punkt mit Ortsvektor  $\vec{x}_n$  auf  $G$  liegt:

$$|\vec{n}| \cdot |\vec{x}_n| = \vec{n} \bullet \vec{x}_0 = |\vec{n}| \cdot |\vec{x}_0| \cdot \cos \varphi$$

Nun können wir beide Seiten der Gleichung durch  $|\vec{n}|$  dividieren (der Normalenvektor ist schließlich nicht der Nullvektor) und erhalten:

$$|\vec{x}_n| = |\vec{x}_0| \cdot \cos \varphi$$

Das ist mit der Situation im rechtwinkligen Dreieck verträglich (siehe Abbildung 4.2.3, S. 99). Wenn aber die besondere Situation vorliegt, dass der Normalenvektor schon Betrag 1 hat (also ein *Normaleneinheitsvektor* ist), gilt sogar:

$$|\vec{x}_n| = \vec{n} \bullet \vec{x}_0 = d$$

Hier muss der Winkel  $\varphi$ , um den der Aufpunkt vom Lotfußpunkt abweicht, nicht extra bestimmt werden, sondern der Abstand  $d$  liegt direkt in dem Skalarprodukt  $\vec{n} \bullet \vec{x}_0$  von Normalen- und Aufpunktvektor der Geraden vor, das (kanonisch) leicht über deren Komponenten berechenbar ist.

Eine kleine Ungenauigkeit bleibt, denn der Abstand der Geraden vom Nullpunkt sollte keine negative Zahl sein. Der Abstand ist dann mit dem (richtigen) Wert angegeben, wenn  $d \geq 0$ . Dann ist der Normalenvektor  $\vec{n}$  gleichgerichtet mit  $\vec{x}_n$ . Im anderen Fall ist (und dies hatten wir oben unterschlagen)

$$\vec{n} \bullet \vec{x}_n = -|\vec{n}| \cdot |\vec{x}_n|$$

Dann kompensiert sich das Minuszeichen dieses Ausdrucks mit dem negativen Wert von  $d$ , und  $|\vec{x}_n|$  ist wieder (richtigerweise) positiv. Entscheidend ist also das Vorzeichen von  $d$ .

Wir fassen dies zusammen als

**Satz A.3** (Hesse-Normalform einer Geraden in  $\mathbb{R}^2$ ). <sup>2</sup> Eine Gerade  $G \subset \mathbb{R}^2$  mit Aufpunkt  $\vec{x}_0$  ist genau dann in Hesse-Normalform gegeben, wenn sie in Normalenform gegeben ist und ihr Normalenvektor  $\vec{e}_n$  auf die Länge 1 normiert ist. Dann gilt für alle  $\vec{x} \in G$ :

$$\vec{e}_n \bullet \vec{x} = \vec{e}_n \bullet \vec{x}_0 = d$$

Falls  $d > 0$ , liegt die Gerade vom Ursprung aus betrachtet mit Abstand  $d$  in Richtung von  $\vec{e}_n$ .

Falls  $d = 0$ , handelt es sich um eine Ursprungsgerade.

Falls  $d < 0$ , liegt die Gerade vom Ursprung aus betrachtet mit Abstand  $|d| = (-d)$  entgegen der Richtung von  $\vec{e}_n$ .

<sup>2</sup>L.O. Hesse, dt. Mathematiker

### Bemerkungen:

- Ist eine Geradengleichung in Koordinatenform gegeben, so erhält man eine äquivalente Gleichung, die auf der rechten Seite den Abstand  $d$  angibt, indem man sie mit dem Kehrwert der Länge des impliziten Normalenvektors skaliert.
- Die Hesse-Normalform ist besonders günstig, um Abstandsrechnungen auszuführen. Dazu zählen:
  - lotrechte Abstände von Punkten zu Geraden in  $\mathbb{R}^2$  oder Ebenen in  $\mathbb{R}^3$
  - Abstände paralleler Geraden in  $\mathbb{R}^2$ , paralleler Ebenen oder windschiefer Geraden in  $\mathbb{R}^3$ .

Wir gehen hier auf solche Rechnungen allerdings nicht weiter ein.

**Beispiel:** Für die Geradengleichung  $G$  mit  $x_1 + 3x_2 = 7$  erhalten wir einen Normalenvektor

$$\vec{n} = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$$

mit Länge  $|\vec{n}| = \sqrt{10}$ . Also beträgt der Abstand der Geraden vom Nullpunkt:

$$d = \frac{7}{\sqrt{10}}$$

Und da  $d > 0$  wissen wir, dass der Normalenvektor  $\vec{n}$  vom Nullpunkt aus zu  $G$  hin zeigt. Die Hesse-Normalform der Gerade wäre:

$$\frac{1}{\sqrt{10}} \begin{pmatrix} 1 \\ 3 \end{pmatrix} \bullet \vec{x} = \frac{7}{\sqrt{10}}$$

### A.3.2 Geraden in $\mathbb{R}^n$ , $n > 2$

In höherdimensionalen Räumen als  $\mathbb{R}^2$  lassen sich Geraden nicht in (Hesse-)Normalform angeben. Dahinter steckt (was wir aber nicht weiter im Detail diskutieren) die Frage, wie sich die Dimensionen des Raums mit der Einschränkung auf eine Punktmenge in Gestalt einer Gerade verhalten. Bei  $n$  Dimensionen sind, grob formuliert,  $(n-1)$  so genannte *Zwangsbedingungen* (realisiert durch voneinander unabhängige einzelne Gleichungen) nötig, um eine eindimensionale Struktur zu erhalten (und eine Gerade ist genau solch eine). Daher reicht in  $\mathbb{R}^2$  schon eine Gleichung aus, um eine Gerade zu definieren. In  $\mathbb{R}^3$  würde man dagegen zwei Gleichungen benötigen (die nicht äquivalent zueinander sind).

Die Gleichung der (Hesse-)Normalform ist jedenfalls eine skalare Gleichung und reduziert die Dimensionalität von  $n$  auf  $(n-1)$  – das reicht nur in  $\mathbb{R}^2$  aus, um eine Gerade zu definieren.

---

Was aber stets in  $\mathbb{R}^n$  möglich ist, ist die Parameterform! Geraden haben für  $n \geq 2$  stets die Form

$$\vec{x}(t) = \vec{x}_0 + t\vec{a}$$

Hierbei sind die Aufpunkt- und Richtungsvektoren dann aus  $\mathbb{R}^n$  zu ziehen, aber es bleibt bei einem freien Parameter, der damit auch eine eindimensionale Struktur definiert. Die Verträglichkeit mit obigen Bemerkungen ist auch hier gegeben: Im  $\mathbb{R}^n$  handelt es sich nämlich bei der Vektorgleichung der Parameterform um genau  $n$  skalare Gleichungen. Durch den Parameter  $t$  kommt aber wieder ein Freiheitsgrad hinzu, sodass die Dimensionsreduktion hier genau die erforderlichen  $(n-1)$  beträgt.

---

Man macht sich in Gedanken leicht klar, dass es für  $n > 2$  auch nicht mehr möglich ist von *dem* Normalenvektor einer Geraden  $G$  zu sprechen (in  $\mathbb{R}^2$  gab es effektiv nur einen, weil die Richtung jeder Geraden, die  $G$  im rechten Winkel schneidet, durch diesen Normalenvektor formulierbar ist (natürlich mit reellem Skalierungsfaktor). Ab  $n = 3$  können aber die Geraden, die  $G$  orthogonal schneiden, unendlich viele Richtungen annehmen: Man stelle sich einen festen Punkt auf der Gerade  $G$  und einen senkrecht abstehenden Vektorpfeil vor – dies ist offenbar ein Normalenvektor für  $G$ . Dann kann man die Basis des Vektors fest halten und die Spitze des Pfeils kreisförmig um die Gerade  $G$  herum rotieren lassen. Dabei entstehen jeweils wieder Vektoren, die orthogonal zu  $G$  sind. Insgesamt definieren alle diese Normalenvektoren eine Ebene durch den fest gehaltenen Punkt, für die  $G$  eine Normale ist.

---

Die Bemerkungen zu Dimensionalität, Zwangsbedingungen und Freiheitsgraden sind natürlich kein Vorlesungsstoff!

---

Wir halten aber noch fest (Beweis: Übung):

**Satz A.4** (Geraden als Unterräume). *Geraden durch den Koordinatenursprung von  $\mathbb{R}^n$  ( $n \geq 2$ ) sind eindimensionale Unterräume von  $\mathbb{R}^n$ .*

**Bemerkung:** In den Beispielen zu Satz 6.4 (S. 151) wurde schon angedeutet, wie man die Vektorraum-Operationen anpassen kann, sodass Geraden, die *nicht* durch den Ursprung verlaufen, (eindimensionale) Vektorräume sind. Unterräume wären sie dann trotzdem nicht, da sie den Nullvektor des Raumes  $\mathbb{R}^n$ , in den sie eingebettet sind, nicht enthalten<sup>3</sup>.

### A.3.3 Ebenen in $\mathbb{R}^n$ , $n > 2$

Eine Ebene ist eine zweidimensionale Punktmenge. In  $\mathbb{R}^2$  gibt es nur genau eine Ebene, nämlich diejenige, die von den beiden Koordinatenachsen aufgespannt wird – die  $x_1, x_2$ -Ebene. Für höherdimensionale Räume sind dann allerdings beliebig viele Ebenen möglich.

---

Wir beginnen wieder mit einer vektoriellen Beschreibung, die in  $\mathbb{R}^n$  ( $n > 2$ ) unabhängig von  $n$  richtig bleibt:

**Definition A.5** (Parameterform von Ebenen). *Gegeben ein spezieller Punkt  $\vec{x}_0 \in \mathbb{R}^n$ ,  $n \geq 2$  sowie zwei linear unabhängige Vektoren  $\vec{a}, \vec{b} \in \mathbb{R}^n \setminus \{\vec{0}\}$ . Dann ist die Ebene  $E(\vec{x}_0, \vec{a}, \vec{b})$  durch den Punkt  $\vec{x}_0$  und mit den Spannvektoren  $\vec{a}, \vec{b}$  die Punktmenge*

$$E(\vec{x}_0, \vec{a}, \vec{b}) := \left\{ \vec{x}_0 + s\vec{a} + t\vec{b} \mid s, t \in \mathbb{R} \right\}$$

**Bemerkungen:**

- Die  $x_1, x_2$ -Ebene in  $\mathbb{R}^2$  ist auch eine Ebene und lässt sich mit zwei beliebigen linear unabhängigen Spannvektoren aus  $\mathbb{R}^2$  realisieren; der Aufpunkt ist hier irrelevant und kann als  $\vec{0}$  gewählt werden. Die Definition lässt  $n = 2$  zu, aber die “interessanteren” Situationen treten natürlich erst für  $n > 2$  auf.
- Es handelt sich hier um eine Bestimmungsgleichung mit zwei freien (und voneinander unabhängigen Parametern).

**Beispiel:** Eine Ebene in  $\mathbb{R}^5$  wäre gegeben durch

$$\vec{x}(s, t) := \begin{pmatrix} 2 \\ 1 \\ 0 \\ 1 \\ -3 \end{pmatrix} + s \begin{pmatrix} 1 \\ 0 \\ 0 \\ 2 \\ 1 \end{pmatrix} + t \begin{pmatrix} 2 \\ 1 \\ 1 \\ 0 \\ 1 \end{pmatrix}$$

Jeder Punkt der Ebene ist mit einem eindeutigen Paar  $(s, t)$  vom Aufpunkt aus erreichbar. Zwei Punkte der Ebene wären etwa

$$\vec{x}(1, 0) = \begin{pmatrix} 3 \\ 1 \\ 0 \\ 3 \\ -2 \end{pmatrix} \quad \text{und} \quad \vec{x}(-1, 2) = \begin{pmatrix} 5 \\ 3 \\ 2 \\ -1 \\ -2 \end{pmatrix}$$

---

<sup>3</sup>Sie sind (kein Vorlesungsstoff) aber *affine Unterräume* – das sind Mengen von Vektoren, die sich durch Subtraktion eines globalen Offset-Vektors in Unterräume des  $\mathbb{R}^n$  verwandeln. Die Subtraktion des Aufpunkts  $\vec{x}_0$  führt hier zum Ziel.

Wir rechnen noch nach, dass der Differenzvektor dieser beiden Punkte innerhalb der Ebene liegt, sich also linear aus den beiden Spannvektoren kombinieren lässt:

$$\vec{x}(-1, 2) - \vec{x}(1, 0) = \begin{pmatrix} 2 \\ 2 \\ 2 \\ -4 \\ 0 \end{pmatrix} \stackrel{!}{=} \begin{pmatrix} s \\ 0 \\ 0 \\ 2s \\ s \end{pmatrix} + \begin{pmatrix} 2t \\ t \\ t \\ 0 \\ t \end{pmatrix} = \begin{pmatrix} s + 2t \\ t \\ t \\ 2s \\ s + t \end{pmatrix}$$

Hieraus liest man sofort ab, dass  $s = (-2)$  und  $t = 2$ . Einsetzen dieser Werte bestätigt die Aussage.

Mit dem Wissen darüber, wie die beiden Punkte berechnet wurden, hätten wir die Linearkombination natürlich auch direkt erhalten, denn (wir nennen die beiden Spannvektoren hier  $\vec{a}, \vec{b}$  wie in der Definition):

$$\vec{x}(-1, 2) - \vec{x}(1, 0) = (\vec{x}_0 - \vec{a} + 2\vec{b}) - (\vec{x}_0 + \vec{a}) = -2\vec{a} + 2\vec{b}$$

Jede Ebene  $E \subset \mathbb{R}^3$  besitzt einen (bis auf Länge und Vorzeichen eindeutigen) Normalenvektor  $\vec{n}$ . Dieser ist orthogonal zu allen Vektoren, die (als Pfeil betrachtet) innerhalb von  $E$  liegen, sich also linear aus den Spannvektoren kombinieren lassen. Für

$$\vec{x} = \vec{x}_0 + s\vec{a} + t\vec{b}$$

ist also insbesondere:

$$\vec{n} \bullet \vec{a} = \vec{n} \bullet \vec{b} = 0$$

Wir bilden nun das Skalarprodukt aus  $\vec{n}$  und der Ebenengleichung und erhalten eine skalare Bestimmungsgleichung für die Ebene:

$$\vec{n} \bullet \vec{x} = \vec{n} \bullet \vec{x}_0$$

Offensichtlich haben wir hier eine *Normalenform* der Ebene, genau analog wie für die Geraden in  $\mathbb{R}^2$ .

Für  $\mathbb{R}^3$  ist durch die Normalenform (Falls  $|\vec{n}| = 1$ , ist es sogar eine *Hesse-Normalform*) die Ebene bereits vollständig (und gleichwertig zur Darstellung mit Spannvektoren) beschrieben.

Hat man eine Darstellung der Ebene mit Spannvektoren (eine solche lässt sich leicht durch Differenzbildung aus *drei* nicht-kollinearen Punkten der Ebene gewinnen) gegeben, so lässt sich der Normalenvektor direkt als *Kreuzprodukt der beiden Spannvektoren* berechnen.

Wir betrachten daher zunächst ein solches

**Beispiel:** Die drei Punkte  $P(2, 1, 7)$ ,  $Q(3, -1, 1)$  und  $R(-2, 1, 0)$  seien gegeben. Wir wählen zunächst den Ortsvektor von  $P$  als Aufpunkt und die Abstandsvektoren  $\vec{x}_{PQ}$  sowie  $\vec{x}_{PR}$  als Spannvektoren. Damit erhalten wir die vektorielle Ebenengleichung:

$$\vec{x}(s, t) = \begin{pmatrix} 2 \\ 1 \\ 7 \end{pmatrix} + s \begin{pmatrix} 1 \\ -2 \\ -6 \end{pmatrix} + t \begin{pmatrix} -4 \\ 0 \\ -7 \end{pmatrix}$$

Wir berechnen den Normalenvektor:

$$\vec{n} := \vec{a} \times \vec{b} = \begin{pmatrix} 14 \\ 31 \\ -8 \end{pmatrix}$$

(Wäre hier der Nullvektor heraus gekommen, dann wären entweder die Punkte kollinear, oder wir hätten die Richtungsvektoren falsch berechnet!)

Nun geben wir die Normalenform der Ebene an (inkl. der Koordinatenform):

$$\vec{n} \bullet \vec{x} = 14x_1 + 31x_2 - 8x_3 = \vec{n} \bullet \vec{x}_0 = 28 + 31 - 56 = 3$$

Die Hesse-Normalform würden wir erhalten, wenn wir diese Gleichung noch mit  $|\vec{n}| = \sqrt{1221}$  dividieren würden. Da das Skalarprodukt aus  $\vec{n}$  und  $\vec{x}_0$  positiv ist, weist  $\vec{n}$  vom Ursprung aus zur

Ebene hin. Die Schnittpunkte mit den Koordinatenachsen erhält man leicht, indem man jeweils die anderen beiden Variablen in der Koordinatengleichung auf null setzt.

Zur Probe rechnen wir noch nach, dass die Punkte  $Q$  und  $R$  der Ebenengleichung in Normalform genügen:

$$\vec{n} \bullet \begin{pmatrix} 3 \\ -1 \\ 1 \end{pmatrix} = 42 - 31 - 8 = 3 \quad \checkmark \quad \text{und} \quad \vec{n} \bullet \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} = -28 + 31 = 3 \quad \checkmark$$

Wie verhält es sich aber nun mit der Ebenengleichung vom vorigen Beispiel aus  $\mathbb{R}^5$ :

$$\vec{x}(s, t) := \begin{pmatrix} 2 \\ 1 \\ 0 \\ 1 \\ -3 \end{pmatrix} + s \begin{pmatrix} 1 \\ 0 \\ 0 \\ 2 \\ 1 \end{pmatrix} + t \begin{pmatrix} 2 \\ 1 \\ 1 \\ 0 \\ 1 \end{pmatrix} ?$$

Zunächst ist klar, dass wir aus den Richtungsvektoren nicht per Kreuzprodukt auf den Normalenvektor der Ebene kommen, denn dieses Produkt ist nur für  $\mathbb{R}^3$  definiert. Wir können aber für

$$\vec{n} := \begin{pmatrix} n_1 \\ n_2 \\ n_3 \\ n_4 \\ n_5 \end{pmatrix}$$

zumindest die beiden Bedingungen notieren, die sich aus der Orthogonalität von  $\vec{n}$  zu den beiden Spannvektoren (wir nennen sie wieder  $\vec{a}, \vec{b}$ ) ergeben:

$$\vec{n} \bullet \vec{a} = n_1 + 2n_4 + n_5 = 0 \quad \text{und} \quad \vec{n} \bullet \vec{b} = 2n_1 + n_2 + n_3 + n_5 = 0$$

Das sind zwei Gleichungen für fünf Unbekannte – wir werden also nicht *den* Normalenvektor bestimmen können! Durch Addition des  $(-2)$ -fachen der ersten Gleichung zur zweiten erhalten wir das Gleichungssystem

$$\begin{aligned} n_1 &= -2n_4 - n_5 \\ \wedge \quad n_2 &= -n_3 + 4n_4 + n_5 \end{aligned}$$

Es gibt also drei freie Parameter, und die Vektoren  $\vec{n}$  bilden einen *dreidimensionalen* Unterraum von  $\mathbb{R}^5$ . Beispiele sind (systematisch je einen der freien Parameter auf 1 gesetzt, die anderen beiden auf 0):

$$\begin{pmatrix} 0 \\ -1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} -2 \\ 4 \\ 0 \\ 1 \\ 0 \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

Man überprüft leicht, dass die drei Vektoren linear unabhängig sind (also als Basis des Normalen-Unterraums dienen können) und orthogonal zu  $\vec{a}, \vec{b}$  sind. Wir rechnen für den zweiten Vektor nach:

$$\begin{pmatrix} -2 \\ 4 \\ 0 \\ 1 \\ 0 \end{pmatrix} \bullet \begin{pmatrix} 1 \\ 0 \\ 0 \\ 2 \\ 1 \end{pmatrix} = -2 + 2 = 0 \quad \text{und} \quad \begin{pmatrix} -2 \\ 4 \\ 0 \\ 1 \\ 0 \end{pmatrix} \bullet \begin{pmatrix} 2 \\ 1 \\ 1 \\ 0 \\ 1 \end{pmatrix} = -4 + 4 = 0 \quad \checkmark$$

Eine geschlossene Ebenendarstellung haben wir aber hiermit nicht! Offenbar ist das Gebilde, das normal zur Ebene liegt, ein ganzer Vektorraum und keine einzelne Gerade. Hier versagt auch die räumliche Vorstellungskraft, die uns oben noch geholfen hatte einzusehen, dass es in  $\mathbb{R}^3$  nicht *die* Normalenform für Geraden geben kann (dort besteht das gleiche Problem: der Normalenraum einer Geraden in  $\mathbb{R}^3$  ist zweidimensional, was wir aber immerhin mit dem Gedankenexperiment bestätigen konnten).

Natürlich lässt sich auch hier formulieren, dass  $\vec{n} \bullet \vec{x} = \vec{n} \bullet \vec{x}_0$ , aber dies ist dann eine Gleichung mit acht Unbekannten (nach obigem Vorgehen:  $x_1, \dots, x_5, n_3, n_4, n_5$ ). Hier hätten wir unser Problem *nicht* einfacher gemacht.

Was allerdings möglich wäre (darauf gehen wir im letzten Unterabschnitt nochmal kurz ein), wäre, für die drei schon ermittelten linear unabhängigen Normalenvektoren *jeweils* eigene Gleichungen aufzustellen. Nennen wir die oben gefundenen Vektoren  $\vec{n}_1, \vec{n}_2, \vec{n}_3$  (hier dienen die Indices zum Nummerieren der Vektoren, nicht um Vektorkomponenten zu kennzeichnen!), so gelten die drei Gleichungen

$$\begin{aligned} \vec{n}_1 \bullet \vec{x} &= \vec{n}_1 \bullet \vec{x}_0 = -1 \\ \wedge \quad \vec{n}_2 \bullet \vec{x} &= \vec{n}_2 \bullet \vec{x}_0 = 1 \\ \wedge \quad \vec{n}_3 \bullet \vec{x} &= \vec{n}_3 \bullet \vec{x}_0 = -4 \end{aligned}$$

Dies sind drei Gleichungen für fünf Variablen  $(x_1, \dots, x_5)$ , die insgesamt zwei Freiheitsgrade zulassen – genau das, was wir für eine zweidimensionale Ebene benötigen.

---

Als Vorlesungsstoff behandeln wir nur Ebenen in  $\mathbb{R}^3$ .

---

Analog zu den Geraden halten wir aber noch fest:

**Satz A.6** (Ebenen als Unterräume). *Ebenen durch den Koordinatenursprung von  $\mathbb{R}^n$  ( $n \geq 2$ ) sind zweidimensionale Unterräume von  $\mathbb{R}^n$ .*

**Bemerkung:** Auch hier bilden die anderen Ebenen immerhin affine Unterräume von  $\mathbb{R}^n$ ; siehe die Fußnote in der Bemerkung zu Satz A.4.

### A.3.4 Abstandsprobleme

Zwei Punkte mit Ortsvektoren  $\vec{x}, \vec{y} \in \mathbb{R}^n$  haben im kartesischen Koordinatensystem den Abstand

$$|\vec{x} - \vec{y}| = \sqrt{(\vec{x} - \vec{y}) \bullet (\vec{x} - \vec{y})} = \sqrt{(x_1 - y_1)^2 + \dots + (x_n - y_n)^2}$$

Für  $n = 2$  lässt sich dies leicht mit dem Satz des Pythagoras überprüfen.

---

Gegeben eine Gerade  $G$  in  $\mathbb{R}^2$  und ein Punkt  $P$  mit Ortsvektor  $\vec{p} \in \mathbb{R}^2$ , lässt sich der *lotrechte* von  $P$  zu  $G$  bestimmen, indem man die Hesse-Normalform von  $G$  betrachtet. Mit dem Normalen-Einheitsvektor von  $G$  konstruiert man in Gedanken eine parallele Gerade  $G_P$  durch  $P$ ; diese besitzt selben Normaleneinheitsvektor.

Mit

$$G: \quad \vec{e}_n \bullet \vec{x} = d \quad \text{und} \quad G_P: \quad \vec{e}_n \bullet \vec{x} = \vec{e}_n \bullet \vec{p}$$

ergibt sich der Abstand dann als Betrag der Differenz beider rechter Seiten. Hier ist besonders bei verschiedenen Vorzeichen der rechten Seiten Sorgfalt empfohlen.

Mit der selben Methode rechnet man auch die Abstände paralleler Geraden in  $\mathbb{R}^2$  aus.

---

Analog lässt sich über die Hesse-Normalform einer Ebene  $E$  in  $\mathbb{R}^3$  der Abstand eines beliebigen Punktes  $P$  zu  $E$  berechnen. Ebenfalls sind darüber in  $\mathbb{R}^3$  die Abstände paralleler Ebenen sowie die Abstände zwischen  $E$  und dazu parallelen Geraden (deren Richtungsvektor in  $E$  liegt, bzw. orthogonal zum Normalenvektor der Ebene ist) berechenbar.

---

Für den Abstand eines Punktes zu einer Geraden in  $\mathbb{R}^n$  mit  $n > 2$  können wir nicht auf die Hesse-Normalform der Geraden zurück greifen. Trotzdem lässt sich aber das Lot von  $P$  auf solch eine Gerade fallen. Es sei die Situation wie in Abbildung A.6 gegeben.

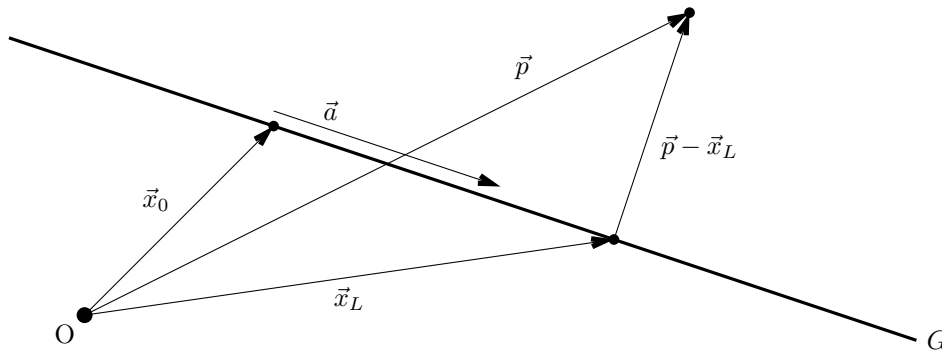


Abbildung A.6: Lotrechter Abstand Punkt-Gerade

Der Fußpunkt des Lotes von  $P$  auf  $G$  habe den Ortsvektor  $\vec{x}_L$ . Dann gilt:

$$\vec{x}_L = \vec{x}_0 + t_L \vec{a}$$

Gleichzeitig steht der Lotvektor von  $G$  zu  $P$  senkrecht auf  $\vec{a}$ , sodass

$$0 = \vec{a} \bullet (\vec{p} - \vec{x}_L) = \vec{a} \bullet (\vec{p} - \vec{x}_0 - t_L \vec{a}) = \vec{a} \bullet (\vec{p} - \vec{x}_0) - t_L \vec{a}^2$$

Also:

$$t_L = \frac{\vec{a} \bullet (\vec{p} - \vec{x}_0)}{|\vec{a}|^2}$$

Damit lässt sich also über die Geradengleichung der Fußpunkt  $\vec{x}_L$  berechnen. Und dann ist der lotrechte Abstand (als Abstand zweier Punkte siehe oben) gegeben per

$$|\vec{p} - \vec{x}_L|$$

### A.3.5 Schnittprobleme

Geraden  $G_1, G_2$  können sich in  $\mathbb{R}^2$  auf drei Weisen schneiden:

- Ist der Schnitt  $G_1 \cap G_2$  leer, so sind die Geraden *parallel* mit verschiedenen Abständen vom Ursprung.
- Ist der Schnitt ein einzelner Vektor, so ist dies der Ortsvektor des *Schnittpunkts*. *Alle nicht parallelen Geraden in  $\mathbb{R}^2$  besitzen einen Schnittpunkt.*
- Ist der Schnitt eine Gerade, so sind  $G_1, G_2$  *identisch* (und damit implizit auch parallel).

In  $\mathbb{R}^3$  sind alle diese Fälle für Geraden ebenfalls möglich. Zusätzlich kann es noch sein, dass es keine gemeinsame Ebene gibt, die beide Geraden enthält – solche Geraden heißen *windschief*. Sie liegen dann in zwei parallelen Ebenen mit verschiedenen Abständen vom Ursprung. *Alle Geraden in  $\mathbb{R}^3$ , die weder parallel sind noch einen Schnittpunkt besitzen, sind zueinander windschief.*

Parallele nicht-identische Geraden und Geraden mit Schnittpunkt in  $\mathbb{R}^3$  definieren jeweils eine gemeinsame Ebene:

- Für parallele nicht-identische Geraden wählt man den Aufpunkt einer der Geraden als Aufpunkt der Ebene. Die Spannvektoren ergeben sich dann aus dem Richtungsvektor dieser Geraden sowie dem Verbindungsvektor beider Aufpunkte.
- Für Geraden mit Schnittpunkt wählt man diesen als Aufpunkt der Ebene, dazu die beiden Richtungsvektoren als Spannvektoren der Ebene.

Ebenen  $E_1, E_2$  können sich in  $\mathbb{R}^3$  auf drei Weisen schneiden:

- Ist der Schnitt  $E_1 \cap E_2$  leer, so sind die Ebenen *parallel* mit verschiedenen Abständen vom Ursprung.
- Ist der Schnitt eine Gerade, so ist dies die *Schnittgerade*. Analog zu den Geraden in  $\mathbb{R}^2$  gilt: *Alle nicht parallelen Ebenen in  $\mathbb{R}^3$  besitzen eine Schnittgerade.*

- Ist der Schnitt eine Ebene, so sind  $E_1, E_2$  *identisch* (und damit implizit auch parallel).
- 

Betrachten wir eine Gerade  $G$  und eine Ebene  $E$  in  $\mathbb{R}^3$ , so können sie sich auf drei Weisen schneiden:

- Ist der Schnitt  $G \cap E$  leer, so ist die Gerade *parallel* zu  $E$  mit verschiedenem Abstand vom Ursprung.
  - Ist der Schnitt ein einzelner Vektor, so ist dies der Ortsvektor des *Schnittpunkts*. Sind Gerade und Ebene nicht parallel, so schneiden sie sich stets in einem Punkt.
  - Ist der Schnitt eine Gerade (nämlich  $G$ ), so liegt  $G$  (ganz) in  $E$  (und ist damit implizit auch parallel zu  $E$ ).
- 

Aus den obigen Beobachtungen (die sich mit der Vorstellungskraft bestätigen lassen), ergeben sich die Lösungsideen für die hier angegebenen Schnittprobleme (Beispiele folgen):

- $G_1, G_2 \subset \mathbb{R}^2$ : Vergleich der Richtungs- oder Normalenvektoren. Sind diese kollinear (hier identisch mit: linear abhängig), so sind die Geraden parallel. Die Frage der Gleichheit lässt sich dann mit den Hesseschen Normalenformen klären. Ergeben sich verschiedene Abstände, so ist  $G_1 \cap G_2 = \emptyset$ ; ansonsten ist  $G_1 \cap G_2 = G_1 = G_2$ . Einfacher ist aber (je nach gegebener Darstellung), zu prüfen, ob der Aufpunkt der einen Gerade die andere Geradengleichung erfüllt. Falls ja, müssen die parallelen Geraden identisch sein.

Für nicht parallele Geraden ist hingegen noch der Schnittpunkt zu bestimmen.

- $G_1, G_2 \subset \mathbb{R}^3$ : Der Vergleich der Richtungsvektoren beantwortet die Frage der Parallelität (lineare Abhängigkeit/Kollinearität prüfen). Bei parallelen Geraden ist noch zu prüfen, ob Gleichheit vorliegt, indem der Aufpunkt der einen Gerade in die Gleichung der anderen eingesetzt wird. Im Erfolgsfall sind beide Geraden identisch.

Für nicht parallele Geraden muss der Schnitt bestimmt werden. Falls er leer ist, liegen windschiefe Geraden vor.

Falls für windschiefe Geraden nach den beiden parallelen Ebenen gefragt ist, die jeweils  $G_1$  bzw.  $G_2$  enthalten, liefert das Kreuzprodukt der beiden Richtungsvektoren der Geraden einen (gemeinsamen) Normalenvektor dieser Ebenen. Hieraus und aus den beiden Aufpunkten der Geraden lassen sich die Normalformen der Ebenen leicht bestimmen. In Hessescher Normalform ist damit auch der Abstand der beiden windschiefen Geraden (nämlich als Abstand der beiden Ebenen) erklärt.

- $E_1, E_2 \subset \mathbb{R}^3$ : Die Parallelität der Ebenen klärt sich durch Vergleich der Normalenvektoren. Falls parallel, lässt sich über die Hesse-Normalformen oder das Einsetzen des jeweils anderen Stützvektors auf Gleichheit prüfen. Der Abstand zweier paralleler Ebenen ergibt sich über die Hesse-Normalform.

Für nicht parallele Ebenen ist der Schnitt zu bestimmen.

- Für  $G, E \subset \mathbb{R}^3$ : Die Gerade  $G$  ist parallel zur Ebene  $E$ , falls das Skalarprodukt aus Richtungsvektor von  $G$  und Normalenvektor von  $E$  verschwindet. Für den Abstand lässt sich die Hesse-Normalform der Ebene verwenden. Setzt man dort den Aufpunkt der Geraden ein, ergibt sich der Abstand wie bei parallelen Ebenen. Damit lässt sich auch die Frage klären, ob  $G \subset E$ .

Falls das erwähnte Skalarprodukt nicht verschwindet, ist ein Schnittpunkt zu berechnen.

---



### Beispiele:

- Wir betrachten zwei Geraden aus  $\mathbb{R}^2$  in Normalenform:

$$\begin{aligned} G_1 &: 3x_1 + 4x_2 = 7 \\ G_2 &: 2x_1 - x_2 = 3 \end{aligned}$$

Für den Schnitt nehmen wir an, dass beide Gleichungen gleichzeitig erfüllt sind, d.h. wir konstruieren ein lineares Gleichungssystem

$$\begin{aligned} 3x_1 + 4x_2 &= 7 \\ \wedge \quad 2x_1 - x_2 &= 3 \end{aligned}$$

Multiplizieren wir die zweite Gleichung mit 4, so können wir danach durch Addition mit der ersten Gleichung die Variable  $x_2$  eliminieren; das liefert uns  $x_1$ :

$$\begin{aligned} \dots \Leftrightarrow \quad & \begin{aligned} 3x_1 + 4x_2 &= 7 \\ \wedge \quad 8x_1 - 4x_2 &= 12 \end{aligned} \\ \Rightarrow \quad & 11x_1 = 19 \Leftrightarrow x_1 = \frac{19}{11} \end{aligned}$$

Danach können wir die ursprüngliche zweite Gleichung verwenden, um  $x_2$  zu bestimmen:

$$x_2 = 2x_1 - 3 = 2 \cdot \frac{19}{11} - 3 = \frac{38 - 33}{11} = \frac{5}{11}$$

Also schneiden sich beide Geraden genau in einem Punkt, nämlich in

$$\frac{1}{11} \begin{pmatrix} 19 \\ 5 \end{pmatrix}$$

Zur Probe rechne man nach, dass dieser Punkt beiden oben angegebenen Geradengleichungen genügt.

- Wir rechnen das gleiche Beispiel nochmal in Parameterform nach. Zunächst bestimmen wir die Richtungsvektoren aus den Normalenvektoren wie oben beschrieben. Für die Aufpunkte verwenden wir die o.g. Geradengleichungen und setzen  $x_2 := 1$  ein. Dies führt uns auf  $x_1 = 1$  für  $G_1$  sowie auf  $x_1 = 2$  für  $G_2$ . Die Geradengleichungen sind dann äquivalent:

$$\begin{aligned} G_1 &: \vec{x} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} + s \begin{pmatrix} 4 \\ -3 \end{pmatrix} \\ G_2 &: \vec{x} = \begin{pmatrix} 2 \\ 1 \end{pmatrix} + t \begin{pmatrix} -1 \\ -2 \end{pmatrix} \end{aligned}$$

Der Schnitt liefert uns eine Vektorgleichung, die wir direkt als System zweier einzelner Gleichungen anschreiben und dann lösen:

$$\begin{aligned} 1 + 4s &= 2 - t \\ \wedge \quad 1 - 3s &= 1 - 2t \\ \Leftrightarrow \quad & \begin{aligned} t &= 1 - 4s \\ \wedge \quad 3s &= 2t = 2 - 8s \end{aligned} \\ \Leftrightarrow \quad & \begin{aligned} t &= 1 - 4s \\ \wedge \quad s &= \frac{2}{11} \end{aligned} \\ \Leftrightarrow \quad & \begin{aligned} t &= 1 - \frac{8}{11} = \frac{3}{11} \\ \wedge \quad s &= \frac{2}{11} \end{aligned} \end{aligned}$$

Auch hier erhalten wir eine eindeutige Lösung. Wir setzen  $s$  in die Gleichung für  $G_1$  ein und erhalten

$$\begin{aligned} G_1 &: 3x_1 + 4x_2 = 7 \\ G_2 &: 2x_1 - x_2 = 3 \\ \vec{x} &= \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \frac{2}{11} \begin{pmatrix} 4 \\ -3 \end{pmatrix} = \frac{1}{11} \begin{pmatrix} 11 + 4 \cdot 2 \\ 11 - 3 \cdot 2 \end{pmatrix} = \frac{1}{11} \begin{pmatrix} 19 \\ 5 \end{pmatrix} \end{aligned}$$

Zur Probe rechnen wir noch mit der anderen Gleichung nach:

$$\vec{x} = \begin{pmatrix} 2 \\ 1 \end{pmatrix} + \frac{3}{11} \begin{pmatrix} -1 \\ -2 \end{pmatrix} = \frac{1}{11} \begin{pmatrix} 22 - 3 \\ 11 - 2 \cdot 3 \end{pmatrix} = \frac{1}{11} \begin{pmatrix} 19 \\ 5 \end{pmatrix} \quad \checkmark$$

- Wir betrachten zwei Geraden aus  $\mathbb{R}^3$ :

$$G_1 : \vec{x} = \begin{pmatrix} 2 \\ 1 \\ -3 \end{pmatrix} + s \begin{pmatrix} 2 \\ 0 \\ 5 \end{pmatrix}$$

$$G_2 : \vec{x} = \begin{pmatrix} 3 \\ -1 \\ 0 \end{pmatrix} + t \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}$$

Durch scharfes Hinsehen erkennt man bereits, dass die Richtungsvektoren (wir nennen sie  $\vec{a}, \vec{b}$ ) linear unabhängig sind; es kann sich also nicht um parallele Geraden handeln. Wir berechnen den Schnitt:

$$\begin{array}{rcl} 2 + 2s & = & 3 + t \\ \wedge \quad 1 & = & -1 - t \\ \wedge \quad -3 + 5s & = & 2t \end{array}$$

$$\Leftrightarrow \begin{array}{rcl} 2s & = & 1 + t \\ \wedge \quad t & = & 0 \\ \wedge \quad 5s & = & 2t + 3 \end{array}$$

$$\Leftrightarrow \begin{array}{rcl} s & = & \frac{1}{2} \\ \wedge \quad t & = & 0 \\ \wedge \quad s & = & \frac{3}{5} \end{array}$$

Dies ist offensichtlich ein Widerspruch, also gibt es keinen Schnittpunkt:  $G_1 \cap G_2 = \emptyset$ . Da außerdem die beiden Richtungsvektoren nicht linear abhängig sind (die Komponente  $a_2 = 0$  lässt sich nicht mit einem Faktor ungleich 0 aus  $b_2$  erzeugen!), sind die beiden Geraden *windschief*.

Wir bestimmen hierzu noch die beiden parallelen Ebenen, von denen jeweils eine  $G_1$  bzw.  $G_2$  enthält. Ein (gemeinsamer) Normalenvektor der Ebenen berechnet sich über das Kreuzprodukt:

$$\vec{n} := \vec{a} \times \vec{b} = \begin{pmatrix} 2 \\ 0 \\ 5 \end{pmatrix} \times \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} = \begin{pmatrix} 5 \\ 1 \\ -2 \end{pmatrix}; \quad |\vec{n}| = \sqrt{30}$$

Damit formulieren wir die Ebenengleichungen in Normalenform (für die rechten Seiten setzen wir jeweils den Aufpunkt der entsprechenden Geraden ein):

$$E_1 : \begin{pmatrix} 5 \\ 1 \\ -2 \end{pmatrix} \cdot \vec{x} = \begin{pmatrix} 5 \\ 1 \\ -2 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ 1 \\ -3 \end{pmatrix} = 17$$

$$E_2 : \begin{pmatrix} 5 \\ 1 \\ -2 \end{pmatrix} \cdot \vec{x} = \begin{pmatrix} 5 \\ 1 \\ -2 \end{pmatrix} \cdot \begin{pmatrix} 3 \\ -1 \\ 0 \end{pmatrix} = 14$$

Die beiden Ebenen (und damit auch die beiden windschiefen Geraden) haben dann (Normierung des Normalenvektors eingerechnet) einen Abstand von

$$\frac{17 - 13}{\sqrt{30}} = \frac{3}{\sqrt{30}}$$

- Wir betrachten zwei Geraden aus  $\mathbb{R}^3$ :

$$G_1 : \vec{x} = \begin{pmatrix} -2 \\ -2 \\ 3 \end{pmatrix} + s \begin{pmatrix} 4 \\ 3 \\ 1 \end{pmatrix}$$

$$G_2 : \vec{x} = \begin{pmatrix} 8 \\ -2 \\ 7 \end{pmatrix} + t \begin{pmatrix} -2 \\ 1 \\ -1 \end{pmatrix}$$

Offenbar sind die Richtungsvektoren (wir nennen sie  $\vec{a}, \vec{b}$ ) nicht parallel, denn sonst müsste mit  $a_1 = -2b_1$  auch  $a_2 = -2b_2$  gelten – was aber nicht zutrifft.

Wir berechnen den Schnitt:

$$\begin{array}{rcl} -2 + 4s & = & 8 - 2t \\ \wedge \quad -2 + 3s & = & -2 + t \\ \wedge \quad 3 + s & = & 7 - t \end{array}$$

$$\Leftrightarrow \begin{array}{rcl} 4s + 2t & = & 10 \\ \wedge \quad 3s - t & = & 0 \\ \wedge \quad s + t & = & 4 \end{array}$$

Aus der dritten Gleichung setzen wir  $t = 4 - s$  in die zweite ein und erhalten:

$$3s = t = 4 - s \Leftrightarrow 4s = 4 \Leftrightarrow s = 1$$

Damit ist  $t = 4 - s = 3$ . Zur Probe setzen wir in die erste Gleichung ein:

$$4s + 2t = 4 \cdot 1 + 2 \cdot 3 = 4 + 6 = 10 \quad \checkmark$$

Mit der eindeutigen Lösung des LGS ist damit ein Schnittpunkt gefunden. Mit  $s = 1$  rechnen wir für  $G_1$ :

$$\vec{x} = \begin{pmatrix} -2 \\ -2 \\ 3 \end{pmatrix} + \begin{pmatrix} 4 \\ 3 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 4 \end{pmatrix}$$

Zur Probe auch mit  $G_2$ :

$$\vec{x} = \begin{pmatrix} 8 \\ -2 \\ 7 \end{pmatrix} + 3 \begin{pmatrix} -2 \\ 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 8 - 6 \\ -2 + 3 \\ 7 - 3 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 4 \end{pmatrix} \quad \checkmark$$

Da die beiden Geraden sich in genau einem Punkt schneiden, ist damit nochmal bestätigt, dass ihre beiden Richtungsvektoren nicht parallel sein können. Die gemeinsame Ebene  $E$ , in der  $G_1, G_2$  liegen, lässt sich dann aus dem Schnittpunkt und den beiden Richtungsvektoren definieren:

$$\vec{x} = \begin{pmatrix} 2 \\ 1 \\ 4 \end{pmatrix} + s \begin{pmatrix} 4 \\ 3 \\ 1 \end{pmatrix} + t \begin{pmatrix} -2 \\ 1 \\ -1 \end{pmatrix}$$

Ein Normalenvektor von  $E$  ist

$$\begin{pmatrix} 4 \\ 3 \\ 1 \end{pmatrix} \times \begin{pmatrix} -2 \\ 1 \\ -1 \end{pmatrix} = \begin{pmatrix} -4 \\ 2 \\ 10 \end{pmatrix}$$

oder skaliert:

$$\vec{n} := \begin{pmatrix} -2 \\ 1 \\ 5 \end{pmatrix}; \quad |\vec{n}| = \sqrt{30}$$

Damit erhalten wir die Normalenform von  $E$  als

$$\begin{pmatrix} -2 \\ 1 \\ 5 \end{pmatrix} \cdot \vec{x} = \begin{pmatrix} -2 \\ 1 \\ 5 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ 1 \\ 4 \end{pmatrix} = 17$$

Der Abstand dieser Ebene vom Nullpunkt beträgt (Hesse-Normalform):

$$\frac{17}{\sqrt{30}} \approx 3.1038$$

Wir verwenden die Formel vom vorigen Unterabschnitt, um die Abstände der beiden Geraden vom Nullpunkt zu bestimmen. Mit  $\vec{p} = \vec{0}$  ergibt sich bei  $G_1$ :

$$s = \frac{-\begin{pmatrix} 4 \\ 3 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} -2 \\ -2 \\ 3 \end{pmatrix}}{16 + 9 + 1} = \frac{11}{26}$$

Damit ergibt sich der Lotfußpunkt

$$\vec{x}_1 = \begin{pmatrix} -2 \\ -2 \\ 3 \end{pmatrix} + \frac{11}{26} \begin{pmatrix} 4 \\ 3 \\ 1 \end{pmatrix} = \frac{1}{26} \begin{pmatrix} -52 + 44 \\ -52 + 33 \\ 78 + 11 \end{pmatrix} = \frac{1}{26} \begin{pmatrix} -8 \\ -19 \\ 89 \end{pmatrix}$$

Wir kontrollieren, dass  $\vec{x}_1$  in  $E$  liegt, indem wir die Normalenform benutzen:

$$\begin{pmatrix} -2 \\ 1 \\ 5 \end{pmatrix} \bullet \vec{x}_1 = \frac{1}{26}(16 - 19 + 445) = \frac{442}{26} = 17 \quad \checkmark$$

Außerdem ist  $\vec{x}_1 \perp \vec{a}$ , denn

$$\begin{pmatrix} 4 \\ 3 \\ 1 \end{pmatrix} \bullet \begin{pmatrix} -8 \\ -19 \\ 89 \end{pmatrix} = -32 - 57 + 89 = 0 \quad \checkmark$$

Der Abstand von  $G_1$  zum Nullpunkt beträgt dann:

$$|\vec{x}_1| = \frac{1}{26} \sqrt{64 + 361 + 7921} = \frac{\sqrt{8346}}{26} \approx 3.5137$$

Analog für  $G_2$ :

$$t = \frac{-\begin{pmatrix} -2 \\ 1 \\ -1 \end{pmatrix} \bullet \begin{pmatrix} 8 \\ -2 \\ 7 \end{pmatrix}}{4 + 1 + 1} = \frac{25}{6}$$

Damit erhalten wir den Lotfußpunkt

$$\vec{x}_2 = \begin{pmatrix} 8 \\ -2 \\ 7 \end{pmatrix} + \frac{25}{6} \begin{pmatrix} -2 \\ 1 \\ -1 \end{pmatrix} = \frac{1}{6} \begin{pmatrix} 48 - 50 \\ -12 + 25 \\ 42 - 25 \end{pmatrix} = \frac{1}{6} \begin{pmatrix} -2 \\ 13 \\ 17 \end{pmatrix}$$

Auch  $\vec{x}_2$  liegt in  $E$ , denn

$$\begin{pmatrix} -2 \\ 1 \\ 5 \end{pmatrix} \bullet \vec{x}_2 = \frac{1}{6}(4 + 13 + 85) = \frac{102}{6} = 17 \quad \checkmark$$

Außerdem ist  $\vec{x}_2 \perp \vec{b}$ , denn

$$\begin{pmatrix} -2 \\ 1 \\ -1 \end{pmatrix} \bullet \begin{pmatrix} -2 \\ 13 \\ 17 \end{pmatrix} = 4 + 13 - 17 = 0 \quad \checkmark$$

Der Abstand von  $G_2$  zum Nullpunkt beträgt

$$|\vec{x}_2| = \frac{1}{6} \sqrt{4 + 169 + 289} = \frac{\sqrt{462}}{6} \approx 3.582$$

Es stellt sich vielleicht die Frage, warum die beiden Abstände der Geraden vom Nullpunkt verschieden sind zum Abstand der Ebene vom Nullpunkt (daher haben wir gerundete Zahlenwerte angegeben).

Tatsächlich entspricht der Abstand der Ebene vom Nullpunkt gerade der Länge des kürzesten Vektors von der Ebene zu  $O$ ; das ist ein Vielfaches des Normalenvektors. Sämtliche anderen Punkte aus  $E$  sind damit *weiter* vom Koordinatenursprung entfernt.

Nun sind  $G_1, G_2$  beides echte Teilmengen von  $E$ , enthalten also jeweils gewisse Punkte aus  $E$ , aber viele andere Punkte nicht. Der Abstand dieser Geraden vom Nullpunkt entspricht wiederum der Länge des kürzesten Vektors von der jeweiligen Gerade zum Nullpunkt – das sind Normalenvektoren auf den Geraden (siehe Abbildung A.6). Wenn aber diese Lotfußpunkte nicht dem Lotfußpunkt von  $E$  entsprechen, sind sie (da sie ebenfalls in  $E$  liegen) von  $O$  auch weiter entfernt als letzterer!

Insbesondere sehen wir also, dass mit dem Abstand der gemeinsamen Ebene zu  $O$  *nicht* die Abstände der jeweiligen Geraden zu  $O$  gegeben sind – nur eine untere Schranke für diese Abstände.

- Wir betrachten zwei Geraden aus  $\mathbb{R}^3$ :

$$G_1 : \vec{x} = \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix} + s \begin{pmatrix} -2 \\ 2 \\ 3 \end{pmatrix}$$

$$G_2 : \vec{x} = \begin{pmatrix} -6 \\ 9 \\ 14 \end{pmatrix} + t \begin{pmatrix} 4 \\ -4 \\ -6 \end{pmatrix}$$

Zunächst stellen wir fest, dass die beiden Richtungsvektoren (wir nennen sie  $\vec{a}, \vec{b}$ ) linear abhängig sind, denn  $b_j = -2a_j$  für  $j \in \{1, 2, 3\}$ , also  $\vec{b} = -2\vec{a}$ .

Die Geraden sind also mindestens parallel. Nun bleibt zu prüfen, ob sie sich auch schneiden (falls ja, müssen sie identisch sein). Dazu setzen wir den Aufpunkt der ersten Gerade in die Gleichung von  $G_2$  ein. Falls wir eine Lösung bekommen, sind die Geraden identisch:

$$\begin{aligned} 2 &= -6 + 4t \\ \wedge \quad 1 &= 9 - 4t \\ \wedge \quad 2 &= 14 - 6t \end{aligned}$$

$$\Leftrightarrow \begin{aligned} 4t &= 8 \\ \wedge \quad 4t &= 8 \\ \wedge \quad 6t &= 12 \end{aligned}$$

$$\Leftrightarrow t = 2$$

Somit ist der Aufpunkt von  $G_1$  Teil von  $G_2$ , und damit  $G_1 = G_2$ .

- Wir betrachten zwei Geraden aus  $\mathbb{R}^3$ :

$$G_1 : \vec{x} = \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix} + s \begin{pmatrix} -2 \\ 2 \\ 3 \end{pmatrix}$$

$$G_2 : \vec{x} = \begin{pmatrix} -3 \\ 2 \\ 1 \end{pmatrix} + t \begin{pmatrix} 4 \\ -4 \\ -6 \end{pmatrix}$$

Wie im vorigen Beispiel sind beide Geraden parallel. Für den Test auf Gleichheit setzen wir den Aufpunkt von  $G_2$  in die Gleichung von  $G_1$  ein:

$$\begin{aligned} -3 &= 2 - 2s \\ \wedge \quad 2 &= 1 + 2s \\ \wedge \quad 1 &= 2 + 3s \end{aligned}$$

$$\Leftrightarrow \begin{aligned} 2s &= 5 \\ \wedge \quad 2s &= 1 \\ \wedge \quad 3s &= -1 \end{aligned}$$

Wegen  $5 \neq 1$  ist das LGS nicht lösbar, also ist hier  $G_1 \cap G_2 = \emptyset$ . Es liegen also parallele, aber nicht identische Geraden vor.

Die gemeinsame Ebene  $E$ , die hierdurch definiert ist, lässt sich aus der ersten Geradengleichung konstruieren, indem noch ein Abstandsvektor hinzu gefügt wird, der der Differenz der beiden Aufpunkte entspricht (da die Geraden nicht identisch sind, muss diese Differenz linear unabhängig zum Richtungsvektor von  $G_1$  (bzw.  $G_2$ ) sein:

$$\vec{x} = \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix} + s \begin{pmatrix} -2 \\ 2 \\ 3 \end{pmatrix} + t \begin{pmatrix} -5 \\ 1 \\ -1 \end{pmatrix}$$

Ein Normalenvektor dieser Ebene ist

$$\vec{n} := \begin{pmatrix} -2 \\ 2 \\ 3 \end{pmatrix} \times \begin{pmatrix} -5 \\ 1 \\ -1 \end{pmatrix} = \begin{pmatrix} -5 \\ -17 \\ 8 \end{pmatrix}$$

Hiermit lässt sich auch eine Normalenform von  $E$  angeben. Ähnlich wie im vorvorigen Beispiel könnte man dann die Abstände von  $E, G_1, G_2$  zum Nullpunkt  $O$  bestimmen, wobei wiederum der Abstand von  $E$  eine untere Schranke darstellt, da beide Geraden in  $E$  liegen.

- Wir betrachten zwei (von oben bereits bekannte) Ebenen aus  $\mathbb{R}^3$ :

$$E_1 : \vec{x} = \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix} + s \begin{pmatrix} -2 \\ 2 \\ 3 \end{pmatrix} + t \begin{pmatrix} -5 \\ 1 \\ -1 \end{pmatrix}$$

$$E_2 : \vec{x} = \begin{pmatrix} 2 \\ 1 \\ 4 \end{pmatrix} + u \begin{pmatrix} 4 \\ 3 \\ 1 \end{pmatrix} + v \begin{pmatrix} -2 \\ 1 \\ -1 \end{pmatrix}$$

Der Schnitt in Parameterform führt auf drei Gleichungen mit vier Unbekannten ( $s, t, u, v$ ) und sei nachträglich zur Übung empfohlen. Im Fall eines Widerspruchs sind die Ebenen parallel und nicht identisch. Falls das LGS immer lösbar ist, sind die Ebenen identisch. Im Fall einer Schnittgeraden bleibt ein Parameter frei wählbar; die anderen drei hängen davon funktional ab.

Ebenfalls ist es aber auch möglich, die Normalenformen der beiden Ebenen in ein LGS zu kombinieren – dieses besitzt zwei Gleichungen und drei Variablen und ist also etwas “handlicher”.

Zunächst bestimmen wir also die Normalenformen. Dazu erinnern wir uns an die oben schon berechneten Kreuzprodukte, mit denen wir Normalenvektoren erhalten:

$$\vec{n}_1 := \begin{pmatrix} -5 \\ -17 \\ 8 \end{pmatrix} \quad \text{und} \quad \vec{n}_2 := \begin{pmatrix} -4 \\ 2 \\ 10 \end{pmatrix}$$

Die beiden Normalenvektoren sind nicht parallel, wie sich durch ein Kreuzprodukt bestätigen lässt:

$$\vec{n}_1 \times \vec{n}_2 = \begin{pmatrix} -186 \\ 18 \\ -78 \end{pmatrix} \neq \vec{0}$$

Dann sind aber auch die Ebenen nicht parallel, und damit schneiden sie sich in einer Geraden  $G$ , die wir noch zu bestimmen haben.

(Der Richtungsvektor der Geraden ist übrigens eigentlich hier schon gegeben, denn er entspricht (bis auf Skalierung) dem eben berechneten  $\vec{n}_1 \times \vec{n}_2$ . Aufgrund der Eigenschaften des Kreuzprodukts (Satz 6.17) steht dieser Vektor nämlich senkrecht auf beiden Faktoren.)

Die Normalenformen der Ebenen lauten also (für die rechten Seiten jeweils die Aufpunkte der Ebenen eingesetzt):

$$\begin{aligned} E_1 : -5x_1 - 17x_2 + 8x_3 &= -11 \\ E_2 : -4x_1 + 2x_2 + 10x_3 &= 34 \end{aligned}$$

Nun lösen wir das entsprechende LGS für den Schnitt, mit dem zielführenden Gedanken, dass eine Variable frei bleibt und die anderen beiden davon abhängen werden. Das System ist:

$$\begin{aligned} -5x_1 - 17x_2 + 8x_3 &= -11 \\ \wedge \quad -4x_1 + 2x_2 + 10x_3 &= 34 \end{aligned}$$

Wir eliminieren in der zweiten Gleichung die Variable  $x_1$ , indem wir das  $(-4)$ -fache der ersten Gleichung zum 5-fachen der zweiten Gleichung addieren:

$$\dots \Leftrightarrow \begin{aligned} -5x_1 - 17x_2 + 8x_3 &= -11 \\ \wedge \quad 0 + 78x_2 + 18x_3 &= 214 \end{aligned}$$

Nun ist ungünstigerweise  $17 \nmid 78$ . Aber 8 und 18 haben als gemeinsames Vielfaches die Zahl 72. Wir können also die zweite Gleichung nutzen, um in der ersten die Variable  $x_3$  zu eliminieren; hierzu addieren wir das  $(-4)$ -fache der zweiten zum 9-fachen der ersten Gleichung:

$$\dots \Leftrightarrow \begin{array}{rcl} -45x_1 - 465x_2 & = & -955 \\ \wedge & & 78x_2 + 18x_3 & = & 214 \end{array}$$

Wenn wir  $x_2$  als freie Variable nehmen, lassen sich nun  $x_1, x_3$  jeweils durch  $x_2$  ausdrücken. Es lässt sich die erste Gleichung noch durch  $(-5)$  dividieren, die zweite hingegen durch 2. Nach Umstellung der Terme erhalten wir:

$$\dots \Leftrightarrow \begin{array}{rcl} 9x_1 & = & 191 - 93x_2 \\ \wedge & & 9x_3 & = & 107 - 39x_2 \end{array}$$

Um hieraus die Geradengleichung des Schnitts in Parameterform zu ermitteln, ergänzen wir noch die stets wahre Gleichung  $9x_2 = 9x_2$  als neue zweite Gleichung. Dann definieren die drei linken Seiten einen Vektor  $9\vec{x}$ :

$$\dots \Leftrightarrow \begin{array}{rcl} 9x_1 & = & 191 - 93x_2 \\ \wedge & & 9x_2 & = & 0 + 9x_2 \\ \wedge & & 9x_3 & = & 107 - 39x_2 \end{array} \Rightarrow 9 \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 191 \\ 0 \\ 107 \end{pmatrix} + x_2 \begin{pmatrix} -93 \\ 9 \\ -39 \end{pmatrix}$$

Nun müssen wir nur noch skalieren. Wir setzen  $w := x_2/3$ . Dann erhalten wir die Gleichung für  $G$  als

$$\vec{x} = \frac{1}{9} \begin{pmatrix} 191 \\ 0 \\ 107 \end{pmatrix} + w \begin{pmatrix} -31 \\ 3 \\ -13 \end{pmatrix}$$

Wir erkennen (oder prüfen per Rechnung nach), dass der Richtungsvektor von  $G$  tatsächlich genau einem Viertel des oben bereits berechneten Kreuzprodukts  $\vec{n}_1 \times \vec{n}_2$  entspricht. Zur Probe setzen wir noch den Aufpunkt von  $G$  in beide Ebenengleichungen ein:

$$\begin{aligned} -5 \cdot \frac{191}{9} + 8 \frac{107}{9} &= \frac{-99}{9} = -11 \quad \checkmark \\ -4 \cdot \frac{191}{9} + 10 \frac{107}{9} &= \frac{306}{9} = 34 \quad \checkmark \end{aligned}$$

### A.3.6 Hyperebenen (kein Vorlesungsstoff)

**Definition A.7** (Hyperebene). Gegeben ein spezieller Punkt  $\vec{x}_0 \in \mathbb{R}^n$ ,  $n \in \mathbb{N}$ , sowie  $(n-1)$  linear unabhängige Vektoren  $\vec{a}_1, \dots, \vec{a}_{n-1} \in \mathbb{R}^n \setminus \{\vec{0}\}$ . Dann ist die Hyperebene  $H(\vec{x}_0, \vec{a}_1, \dots, \vec{a}_{n-1})$  durch den Punkt  $\vec{x}_0$  und mit den Spannvektoren  $\vec{a}_j$  die Punktmenge

$$H(\vec{x}_0, \vec{a}_1, \dots, \vec{a}_{n-1}) := \{ \vec{x}_0 + \lambda_1 \vec{a}_1 + \dots + \lambda_{n-1} \vec{a}_{n-1} \mid \lambda_1, \dots, \lambda_{n-1} \in \mathbb{R} \}$$

#### Bemerkungen:

- Eine Hyperebene teilt den Raum  $\mathbb{R}^n$  stets in zwei Hälften; sie ist ein  $(n-1)$ -dimensionaler affiner Unterraum von  $\mathbb{R}^n$ .
- Für  $n = 3$  sind Hyperebenen zweidimensionale affine Unterräume – das sind gerade die Ebenen!
- Für  $n = 2$  sind Hyperebenen durch Geraden gegeben.
- Für  $n = 1$  sind einzelne Punkte Hyperebenen – jeder davon teilt den reellen Zahlenstrahl in zwei Teile.
- Jede Hyperebene besitzt einen (bis auf Skalierung) eindeutig bestimmten *Normalenvektor*  $\vec{n}$ , der orthogonal zu sämtlichen Spannvektoren ist.

Bilden wir das Skalarprodukt von  $\vec{n}$  mit den Vektoren aus  $H$ , so erhalten wir:

$$\vec{n} \bullet \vec{x} = \vec{n} \bullet \vec{x}_0$$

Jede Hyperebene lässt sich also in (Hessescher) *Normalenform* schreiben, und damit gleichwertig als Koordinatengleichung mit  $n$  Unbekannten ausdrücken.

Das ist auch der Grund, wieso Geraden in  $\mathbb{R}^2$  sowie Ebenen in  $\mathbb{R}^3$  solche Darstellungen besitzen; in beiden Situationen handelt es sich um Hyperebenen. Die Geraden in  $\mathbb{R}^3$  haben hingegen (s.o.) einen zweidimensionalen Normalenraum; für sie existiert solch eine Normalenform nicht.

- In obigem Beispiel einer Ebene in  $\mathbb{R}^5$  hatten wir einen dreidimensionalen Unterraum für die Normalenvektoren gefunden. Mit drei linear unabhängigen Normalenvektoren sind auch drei Hyperebenen (in Normalenform) gegeben.

Stellt man ein LGS der drei Normalengleichungen auf, so schneidet man effektiv drei Hyperebenen des  $\mathbb{R}^5$ . Das Ergebnis ist dann tatsächlich eine zweidimensionale Struktur – die Ebene.

(Analog kann man durch Schnitt zweier Ebenen mit linear unabhängigen Normalenvektoren in  $\mathbb{R}^3$ , wie oben gesehen, eine Schnittgerade definieren – eine eindimensionale Struktur).

## A.4 Gram-Schmidt-Orthogonalisierung

Gegeben seien  $n$  linear unabhängige Vektoren  $\mathcal{V} = \{\vec{v}_1, \dots, \vec{v}_n\}$ . Diese bilden eine Basis für den  $n$ -dimensionalen Raum

$$V := \text{span}(\vec{v}_1, \dots, \vec{v}_n)$$

Wir finden nun durch Linearkombination dieser Vektoren eine Orthonormalbasis  $\mathcal{U} := \{\vec{u}_1, \dots, \vec{u}_n\}$  von  $V$ .

---

Zunächst sei

$$\vec{u}_1 := \frac{\vec{v}_1}{\sqrt{\vec{v}_1 \bullet \vec{v}_1}}$$

Dieser Vektor ist normiert und gleichgerichtet zu  $\vec{v}_1$ .

---

Nun sei  $\vec{v}_2'$  der Vektor, der entsteht, wenn man von  $\vec{v}_2$  den Anteil parallel zu  $\vec{u}_1$  entfernt:

$$\vec{v}_2' := \vec{v}_2 - c_{2,1}\vec{u}_1$$

Wir fordern, dass  $\vec{v}_2' \perp \vec{u}_1$ , also  $\vec{v}_2' \bullet \vec{u}_1 = 0$ . Dies liefert uns den Koeffizienten  $c_{2,1}$ :

$$0 = \vec{v}_2' \bullet \vec{u}_1 = (\vec{v}_2 - c_{2,1}\vec{u}_1) \bullet \vec{u}_1 = \vec{v}_2 \bullet \vec{u}_1 - c_{2,1} \underbrace{(\vec{u}_1 \bullet \vec{u}_1)}_{=1} = \vec{v}_2 \bullet \vec{u}_1 - c_{2,1}$$

Also setzen wir

$$c_{2,1} := \vec{v}_2 \bullet \vec{u}_1$$

(Dies ist gerade der Anteil von  $\vec{v}_2$  parallel zu  $\vec{u}_1$ ; siehe dazu Unterabschnitt 6.2.2 zur geometrischen Bedeutung des Skalarprodukts, S. 160ff.)

Offensichtlich ist  $\vec{v}_2' \in V$ . Und da  $\vec{v}_2, \vec{v}_1$  als Teilmenge von  $\mathcal{V}$  nach Satz 6.22 (Eigenschaften linear (un-)abhängiger Vektoren) linear unabhängig sind und  $\vec{v}_2'$  als eine Linearkombination von  $\vec{v}_1, \vec{v}_2$  ausgedrückt werden kann, ist insbesondere  $\vec{v}_2' \neq \vec{0}$ , denn der Koeffizient von  $\vec{v}_2$  in der Linearkombination ist  $1 \neq 0$ .

Wir normieren per

$$\vec{u}_2 := \frac{\vec{v}_2'}{\sqrt{\vec{v}_2' \bullet \vec{v}_2'}}$$

und haben so einen zu  $\vec{u}_1$  orthogonalen normierten Vektor in  $V$  erhalten.

---

Für  $\vec{u}_3$  gehen wir analog vor und projizieren die Anteile parallel zu  $\vec{u}_1$  und  $\vec{u}_2$  heraus per

$$\vec{v}_3' := \vec{v}_3 - c_{3,1}\vec{u}_1 - c_{3,2}\vec{u}_2$$

Mit den Forderungen  $\vec{v}_3' \perp \vec{u}_1$  sowie  $\vec{v}_3' \perp \vec{u}_2$ , und da  $\vec{u}_1 \perp \vec{u}_2$ , erhält man analog:

$$c_{3,1} := \vec{v}_3 \bullet \vec{u}_1 \quad \text{und} \quad c_{3,2} := \vec{v}_3 \bullet \vec{u}_2$$



Wieder liegt eine Linearkombination aus  $V$  vor, und  $\vec{v}_3' \neq \vec{0}$ , da  $\vec{v}_3, \vec{v}_2, \vec{v}_1$  linear unabhängig sind und  $\vec{u}_1, \vec{u}_2$  aus  $\vec{v}_1, \vec{v}_2$  kombiniert wurden, sodass der Koeffizient von  $\vec{v}_3$  in dieser Linearkombination von  $\vec{v}_3'$  sicher  $1 \neq 0$  ist.

Wir erhalten  $\vec{u}_3$  durch Normieren.

---

Analog und allgemein ist für  $1 \leq k \leq n$ :

$$\vec{v}_k' := \vec{v}_k - \sum_{j=1}^{k-1} c_{k,j} \vec{u}_j$$

mit

$$c_{k,j} := \vec{v}_k \bullet \vec{u}_j,$$

sodass

$$\vec{v}_k' \perp \vec{u}_j$$

Da die Vektoren  $\vec{u}_1$  bis  $\vec{u}_{k-1}$  Linearkombinationen aus den Vektoren  $\vec{v}_1$  bis  $\vec{v}_{k-1}$  sind ist der Koeffizient für  $\vec{v}_k$  in der Linearkombination  $\vec{v}_k'$  gleich  $1 \neq 0$ , und es liegt keine Linearkombination des Nullvektors vor – denn diese wäre aufgrund der linearen Unabhängigkeit von  $\mathcal{V}$  nur möglich, wenn alle Koeffizienten verschwinden. Man erhält dann den  $k$ -ten Vektor der Orthonormalbasis  $\mathcal{U}$  durch Normieren von  $\vec{v}_k'$ .

Die Formel für  $\vec{v}_k'$  ist auch richtig für  $k = 1$ , da dann die Summe gar keinen Beitrag liefert und  $\vec{v}_1' = \vec{v}_1$  als einzelner Vektor ungleich  $\vec{0}$  ohnehin linear unabhängig ist.

---

Am Ende der Prozedur liegen  $n$  paarweise zueinander orthogonale (und damit linear unabhängige) Vektoren in  $\mathcal{U}$  vor. Da diese sämtlich Linearkombinationen von Vektoren aus  $\mathcal{V}$  sind, ist  $\text{span}(\mathcal{U}) \subseteq V$ . Weiterhin wissen wir nach Satz 6.26 (Dimension eines Vektorraums), dass eine Basis von  $V$  aus  $n$  l.u. Vektoren bestehen muss, da  $V$   $n$ -dimensional ist.

Es gilt auch, dass  $V \subseteq \text{span}(\mathcal{U})$ . Denn wäre dies nicht so, so gäbe es in  $V$  einen Vektor  $\vec{x}$ , der nicht in  $\text{span}(\mathcal{U})$  liegt und also aus den  $\vec{u}_k$  nicht linear kombinierbar ist. Dieser hätte dann die Gestalt

$$\vec{x} = \left( \sum_k d_k \vec{u}_k \right) + \vec{y}$$

mit  $\vec{y} \in V \setminus \text{span}(\mathcal{U})$ .

Wenn aber  $\vec{y}$  nicht als Linearkombination aus  $\mathcal{U}$  beschreibbar ist, muss  $\mathcal{U} \cup \{\vec{y}\}$  l.u. sein. Denn wäre diese Menge l.a., so gäbe es eine nichttriviale Linearkombination des Nullvektors per

$$\vec{0} = \left( \sum_k a_k \vec{u}_k \right) + a \vec{y}$$

Dabei muss  $a \neq 0$  gelten, denn  $\mathcal{U}$  ist nach Voraussetzung l.u. – für  $a = 0$  würde also nur die triviale Linearkombination des Nullvektors existieren. Aber mit  $a \neq 0$  folgt auch:

$$\vec{y} = -\frac{1}{a} \left( \sum_k a_k \vec{u}_k \right)$$

Damit wäre  $\vec{y}$  allerdings eine Linearkombination aus  $\text{span}(\mathcal{U})$ , im Widerspruch zur Annahme.  $\nexists$

Dann gibt es also keine Vektoren, die in  $V$  liegen, aber nicht in  $\text{span}(\mathcal{U})$  enthalten sind, sodass  $V \subseteq \text{span}(\mathcal{U})$ . Somit gilt insgesamt:

$$V = \text{span}(\mathcal{V}) = \text{span}(\mathcal{U}),$$

und  $\mathcal{U}$  ist damit eine Basis von  $V$ .

---

Wir fassen zusammen:

**Satz A.8** (Gram-Schmidt-Orthogonalisierung). *Für eine Menge  $n$  linear unabhängiger Vektoren  $\mathcal{V} = \{\vec{v}_1, \dots, \vec{v}_n\}$  existiert eine Menge orthonormierter Vektoren  $\mathcal{U} = \{\vec{u}_1, \dots, \vec{u}_n\}$  mit*

$$\text{span}(\mathcal{U}) = \text{span}(\mathcal{V})$$

*Dabei ist für  $k \in \{1, \dots, n\}$  und in aufsteigender Reihenfolge:*

$$\vec{v}'_k := \vec{v}_k - \sum_{j=1}^{k-1} (\vec{v}_k \bullet \vec{u}_j) \vec{u}_j$$

und

$$\vec{u}_k := \frac{\vec{v}'_k}{\sqrt{\vec{v}'_k \bullet \vec{v}'_k}}$$

*Es ist  $\mathcal{U}$  eine Basis von  $\text{span}(\mathcal{V})$ .*

**Beispiele:**

- Wir betrachten in  $\mathbb{R}^4$  die Vektoren

$$\vec{v}_1 := \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \end{pmatrix}, \quad \vec{v}_2 := \begin{pmatrix} 1 \\ 0 \\ 2 \\ 0 \end{pmatrix}, \quad \vec{v}_3 := \begin{pmatrix} -1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \quad \text{und} \quad \vec{v}_4 := \begin{pmatrix} 2 \\ 1 \\ 0 \\ 1 \end{pmatrix}$$

Dann erhalten wir direkt:

$$\vec{u}_1 := \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \end{pmatrix}$$

Damit:

$$\vec{v}'_2 := \vec{v}_2 - (\vec{v}_2 \bullet \vec{u}_1) \vec{u}_1 = \begin{pmatrix} 1 \\ 0 \\ 2 \\ 0 \end{pmatrix} - \frac{1}{3} \cdot 3 \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix}$$

Hier haben wir für den Beitrag  $(\vec{v}_2 \bullet \vec{u}_1) \vec{u}_1$  den Faktor  $\frac{1}{3}$  erhalten durch Multiplikation der beiden Skalierungsfaktoren für  $\vec{u}_1$  (das erspart uns das Weiterrechnen mit Wurzelausdrücken). Der Faktor 3 stammt aus dem Skalarprodukt von  $\vec{v}_2$  und  $\vec{u}_1$ , jedoch ohne Beachtung des bereits eingerechneten Skalierungsfaktors. Wir erhalten also:

$$\vec{u}_2 := \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix}$$

Man überprüft leicht, dass  $\vec{u}_2 \perp \vec{u}_1$ .

Analog weiter:

$$\begin{aligned} \vec{v}'_3 &:= \vec{v}_3 - (\vec{v}_3 \bullet \vec{u}_1) \vec{u}_1 - (\vec{v}_3 \bullet \vec{u}_2) \vec{u}_2 \\ &= \begin{pmatrix} -1 \\ 0 \\ 0 \\ 1 \end{pmatrix} - \underbrace{\frac{1}{3} \cdot 0}_{=\vec{0}} \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \end{pmatrix} - \frac{1}{2} \cdot (-1) \begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \\ \frac{1}{2} \\ \frac{1}{2} \end{pmatrix} = \frac{1}{2} \begin{pmatrix} -2 \\ 0 \\ 1 \\ 1 \end{pmatrix} \end{aligned}$$

Den Vorfaktor  $\frac{1}{2}$  können wir beim Normieren ignorieren und erhalten:

$$\vec{u}_3 := \frac{1}{\sqrt{6}} \begin{pmatrix} -2 \\ 0 \\ 1 \\ 1 \end{pmatrix}$$

Auch hier sieht man schnell, dass  $\vec{u}_3 \perp \vec{u}_1$  und  $\vec{u}_3 \perp \vec{u}_2$ . (Beim Bilden der Skalarprodukte kann man die Skalierungsfaktoren in Gedanken ignorieren, da wir zum Kontrollieren nur am Endergebnis 0 interessiert sind.)

Nun zum letzten Vektor:

$$\begin{aligned}\vec{v}'_4 &:= \vec{v}_4 - (\vec{v}_4 \bullet \vec{u}_1)\vec{u}_1 - (\vec{v}_4 \bullet \vec{u}_2)\vec{u}_2 - (\vec{v}_4 \bullet \vec{u}_3)\vec{u}_3 \\ &= \begin{pmatrix} 2 \\ 1 \\ 0 \\ 1 \end{pmatrix} - \frac{1}{3} \cdot 3 \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \end{pmatrix} - \frac{1}{2} \cdot (-1) \begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix} - \frac{1}{6} \cdot (-3) \begin{pmatrix} -2 \\ 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}\end{aligned}$$

Damit:

$$\vec{u}_4 := \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

Da die anderen drei  $\vec{u}$ -Vektoren in der zweiten Komponente jeweils eine 0 haben, ist klar, dass  $\vec{u}_4$  zu ihnen orthogonal ist. Damit ist die ONB gefunden:

$$\mathcal{U} = \left\{ \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix}, \frac{1}{\sqrt{6}} \begin{pmatrix} -2 \\ 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \right\}$$

Als Spalten einer Matrix interpretiert, wäre diese Matrix orthogonal.

- Wir betrachten noch ein Beispiel in  $\mathbb{R}^3$ :

$$\vec{v}_1 := \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad \vec{v}_2 := \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}, \quad \text{und} \quad \vec{v}_3 := \begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix}$$

Wieder erhalten wir direkt:

$$\vec{u}_1 := \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

Dann weiter:

$$\vec{v}'_2 = \vec{v}_2 - (\vec{v}_2 \bullet \vec{u}_1)\vec{u}_1 = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix} - \frac{1}{2} \cdot 3 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -\frac{1}{2} \\ 2 \\ \frac{1}{2} \end{pmatrix} = \frac{1}{2} \begin{pmatrix} -1 \\ 4 \\ 1 \end{pmatrix}$$

Also:

$$\vec{u}_2 := \frac{1}{\sqrt{18}} \begin{pmatrix} -1 \\ 4 \\ 1 \end{pmatrix}$$

Und für den dritten Vektor:

$$\vec{v}'_3 := \vec{v}_3 - (\vec{v}_3 \bullet \vec{u}_1)\vec{u}_1 - (\vec{v}_3 \bullet \vec{u}_2)\vec{u}_2 = \begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix} - \frac{1}{2} \cdot 1 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} - \frac{1}{18} \cdot (-9) \begin{pmatrix} -1 \\ 4 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = \vec{0}$$

Hier schlägt die Orthogonalisierung fehl – der Grund ist, dass die drei  $\vec{v}$ -Vektoren gar nicht linear unabhängig waren. Vielmehr liegen sie alle in einer gemeinsamen Ebene durch den Koordinatenursprung und spannen damit nur einen zwei-dimensionalen Unterraum von  $\mathbb{R}^3$  auf. Dieser wird auch durch die orthonormalen Vektoren  $\vec{u}_1, \vec{u}_2$  aufgespannt. Projiziert man nun die Komponenten von  $\vec{v}_3$  entlang dieser beiden Spannvektoren heraus, so bleibt, da  $\vec{v}_3$  in der gleichen Ebene liegt, nichts übrig. Wären die drei  $\vec{v}$ -Vektoren linear unabhängig gewesen, so wäre  $\vec{u}_3$  ein Normalen-Einheitsvektor der Ebene  $\text{span}(\vec{u}_1, \vec{u}_2)$ .

Wir können von diesem Negativbeispiel folgendes mitnehmen: Die Orthogonalisierung nach Gram-Schmidt<sup>4</sup> funktioniert immer so lange, bis einer der gegebenen Vektoren mit den bereits

<sup>4</sup>J. P. Gram, dänischer Mathematiker, E. Schmidt, dt. Mathematiker

zuvor berechneten Vektoren der gesuchten ONB linear abhängig wird. Insbesondere wird also möglicherweise keine maximale Basis des Unterraums gefunden, wenn das Verfahren schon fehlschlägt, während von den gegebenen (und bisher nicht betrachteten) Vektoren noch solche vorhanden wären, die linear unabhängig zu den bereits berechneten ONB-Vektoren sind.

## A.5 Vektorraum der Polynome

Wir betrachten Polynome in  $\mathbb{K}[X]$  mit Grad kleiner als ein festes  $n \in \mathbb{N}$ . Alle diese Polynome enthalten in Standardnotation maximal  $n$  Terme und lassen sich also durch  $n$ -Tupel darstellen.

Wir hatten in Satz 5.23 bereits fest gehalten, dass die Polynome mit der Polynom-Addition und der Polynom-Multiplikation einen *kommutativen Ring mit Eins* bilden, wobei die neutralen Elemente dem Null- und dem Eins-Polynom (beide vom Grad 0) entsprechen.

Tatsächlich lässt sich mit Polynomen auch ein *Vektorraum* konstruieren (siehe Definition 6.1, S. 149). Die Tupel-Schreibweise für Polynome suggeriert dies bereits – und wenn wir führende Nullen in den Tupeln akzeptieren, dann lassen sich alle oben erwähnten Polynome mit Grad kleiner als  $n$  mit  $n$ -Tupeln aus dem kartesischen Produktraum (siehe Satz 6.4)  $\mathbb{K}^n$  identifizieren.

Dabei ist die Addition komponentenweise auszuführen wie bei kartesischen Produkträumen üblich. Für die skalare Multiplikation nutzen wir aus, dass Skalierungsfaktoren  $c$  als Polynome ( $c$ ) vom Grad 0 schreibbar sind, sodass auch die Polynom-Multiplikation den Grad des Produktpolynoms nicht erhöht – der Vektorraum wird also durch Skalieren seiner Elemente nicht verlassen. Siehe dazu Definition 5.21 und Satz 5.22 zum Polynom-Produkt.

Die Distributivgesetze und weiteren Vorschriften für Vektorräume sind aufgrund der Ringstruktur alle erfüllt, sodass die Polynome  $\mathbb{K}[X]$  vom Grad kleiner als  $n$  in der Tat einen Vektorraum bilden.

---

Innerhalb dieses Vektorraums existiert dann auch das Konzept der linearen (Un-)Abhängigkeit, wie wir dies in Kapitel 6 schon diskutiert hatten. In diesem Fall benötigen wir noch eine *Basis* – wir wollen nun zeigen, dass die intuitive Übertragung der kartesischen Einheitsvektoren auf die Polynomtupel eine solche Basis liefert.

Dazu betrachten wir die  $n$  verschiedenen Monome

$$\begin{aligned} X^0 &= (0, \dots, 0, 0, 1) \\ X^1 &= (0, \dots, 0, 1, 0) \\ X^2 &= (0, \dots, 1, 0, 0) \\ &\dots = \dots \\ X^{n-1} &= (1, \dots, 0, 0, 0) \end{aligned}$$

Es gilt dann folgender

**Satz A.9** (Monome im Vektorraum der Polynome). *Im Vektorraum der Polynome aus  $\mathbb{K}[X]$  mit Grad kleiner als  $n \in \mathbb{N}$  bilden die Monome*

$$\mathcal{M}_n := \{X^0, X^1, \dots, X^{n-1}\}$$

*eine Basis.*

**Beweis:** Wir schreiben die Linearkombination des Nullvektors (also des Nullpolynoms) der Vektoren aus  $\mathcal{M}_n$  an:

$$(0, \dots, 0) = \sum_{j=0}^{n-1} c_j X^j$$

Dies ist aber direkt die Schreibweise eines allgemeinen Polynoms  $(n-1)$ -ten Grades (siehe Definition 5.17). Das bedeutet direkt, dass alle  $c_j$ , nämlich die Elemente des links notierten Tupels, null betragen müssen.

Weiterhin ist jedes Polynom  $p(X) \in \mathbb{K}[X]$  vom Grad kleiner als  $n$  nach den Monomen aus  $\mathcal{M}_n$  entwickelbar per

$$p(X) = (p_{n-1}, \dots, p_1, p_0) = \sum_{j=0}^{n-1} p_j X^j$$

Dann spannen die Monome in  $\mathcal{M}_n$  den ganzen Vektorraum auf, und

$$\text{span}(\mathcal{M}_n) = \{p(X) \in \mathbb{K}[X] \mid \deg(p) < n\}$$

Nach Definition 6.24 ist  $\mathcal{M}_n$  damit eine Basis des betrachteten Vektorraums, und nach Satz 6.26 ist der Raum  $n$ -dimensional. ■

#### Bemerkungen:

- Dies ist verträglich mit der intuitiv hergestellten Verbindung mit dem kartesischen Produkt und seinen kartesischen Einheitsvektoren.
- Insbesondere sind Monome verschiedenen Grades paarweise linear unabhängig zueinander, da für  $j \neq k$  die Ausdrücke

$$X^j \quad \text{und} \quad X^k$$

nicht durch Skalieren mit Körperelementen ineinander überführbar sind.

## A.6 Basiswechsel mit Matrizen

Wir erinnern uns, dass quadratische Matrizen  $A \in \mathbb{R}^{(n,n)}$  genau dann invertierbar sind, wenn sie vollen Rang  $n$  haben, also ihre Zeilen und Spalten jeweils linear unabhängig sind. Insbesondere ist genau dann auch jedes LGS  $(A \mid \vec{y})$  eindeutig lösbar zu  $\vec{x}$  per

$$A\vec{x} = \vec{y},$$

und nach Satz 8.2 (Äquivalente Beschreibungen von LGS) ist die Inhomogenität  $\vec{y}$  eine (eindeutige) Linearkombination der Spalten von  $A$ , deren Koeffizienten den Komponenten des Lösungsvektors  $\vec{x}$  entsprechen.

Für das folgende wollen wir nur noch solche invertierbaren quadratischen Matrizen aus  $\mathbb{R}^{(n,n)}$  betrachten, ohne dies hier explizit zu erwähnen.

### A.6.1 Invertierbare Matrizen als Basis von $\mathbb{R}^n$

Da die Spalten von  $A = (\vec{a}_1 \ \cdots \ \vec{a}_n)$  linear unabhängig sind, bilden sie nach Satz 6.26 (Dimension eines Vektorraums) eine *Basis* von  $\mathbb{R}^n$  – es lässt sich also jeder Vektor  $\vec{y}$  nach den  $\vec{a}_j$  entwickeln – genau dies passiert implizit beim Berechnen der Lösung

$$\vec{x} = A^{-1}\vec{y}$$

Denn dann ist auch:

$$\vec{y} = \sum_j x_j \vec{a}_j$$

Fasst man nun  $\mathcal{A} := \{\vec{a}_1, \dots, \vec{a}_n\}$  als Basis von  $\mathbb{R}^n$  auf, dann sind die  $x_j$  gerade die Koordinaten von  $\vec{y}$  bezüglich  $\mathcal{A}$ .

Die Koordinaten von  $\vec{y}$  bezüglich der Standardbasis  $E_n$  sind hingegen die Komponenten von  $\vec{y}$  selbst.

---

Nun sei  $B = (\vec{b}_1 \ \cdots \ \vec{b}_n)$  eine weitere invertierbare Matrix. Dann definieren auch deren Spalten eine Basis; diese heiße  $\mathcal{B} := \{\vec{b}_1, \dots, \vec{b}_n\}$ .

Für den gleichen Vektor  $\vec{y}$  gibt es dann ein  $\vec{z}$ , sodass

$$B\vec{z} = \vec{y} \Leftrightarrow \vec{y} = \sum_k z_k \vec{b}_k \Leftrightarrow \vec{z} = B^{-1}\vec{y}$$

Dann sind die  $z_k$  die Koordinaten von  $\vec{y}$  bezüglich  $\mathcal{B}$ .

Insgesamt gilt dann:

$$\boxed{A\vec{x} = \vec{y} = B\vec{z}} \quad (*)$$

### A.6.2 Basiswechsel zwischen Standardbasis und $\mathcal{A}$ und $\mathcal{B}$

Die Vektoren  $\vec{x}$  und  $\vec{z}$  enthalten die Koordinaten von  $\vec{y}$  in den Basen  $\mathcal{A}$  bzw.  $\mathcal{B}$ . Oben wurde schon notiert, wie man sie durch Benutzung der inversen Matrizen ermittelt. Wir notieren also eine allgemeine Basis-Wechsel-Matrix, die von der Standardbasis  $E_n$  nach  $\mathcal{A}$  führt, als

$$M_{E_n \rightarrow \mathcal{A}} := A^{-1}$$

Dann ist

$$\vec{x} = (\vec{y})_{\mathcal{A}} = M_{E_n \rightarrow \mathcal{A}} \cdot \vec{y} \quad \left( = M_{E_n \rightarrow \mathcal{A}} \cdot \vec{y}_{E_n} \right)$$

(Wenn der Vektor nicht mit der Basis gekennzeichnet ist, wird davon ausgegangen, dass die Standardbasis gemeint ist)

Analog gilt dann natürlich, dass  $A$  aus  $\mathcal{A}$  zurück zur Standardbasis führt:

$$M_{\mathcal{A} \rightarrow E_n} := A$$

Für  $\mathcal{B}$  und  $\mathcal{B}$  entsprechend.

### A.6.3 Basiswechsel allgemein

Wir betrachten nun die linke und rechte Seite der oben eingerahmten Gleichung (\*) und nutzen die Invertierbarkeit:

$$\vec{x} = A^{-1} \cdot B \cdot \vec{z} \quad \text{und} \quad \vec{z} = B^{-1} \cdot A \cdot \vec{x}$$

Nun war  $\vec{x}$  die Darstellung von  $\vec{y}$  in der Basis  $\mathcal{A}$  und  $\vec{z}$  die Darstellung von  $\vec{y}$  in der Basis  $\mathcal{B}$ . Die obigen beiden Gleichungen definieren also den Wechsel zwischen den Basen  $\mathcal{A}$  und  $\mathcal{B}$ :

$$M_{\mathcal{B} \rightarrow \mathcal{A}} = A^{-1} \cdot B = M_{E_n \rightarrow \mathcal{A}} \cdot M_{\mathcal{B} \rightarrow E_n}$$

Die direkte Umrechnung von  $\mathcal{B}$  nach  $\mathcal{A}$  kann man also auf dem Umweg über die Standardbasis ausdrücken.

In umgekehrter Richtung gilt entsprechend:

$$M_{\mathcal{A} \rightarrow \mathcal{B}} = (M_{\mathcal{B} \rightarrow \mathcal{A}})^{-1} = (A^{-1} \cdot B)^{-1} = B^{-1} \cdot A = M_{E_n \rightarrow \mathcal{B}} \cdot M_{\mathcal{A} \rightarrow E_n}$$

Mit dieser Methode kann also beliebig zwischen Basen des  $\mathbb{R}^n$  hin und her gewechselt werden.

Auch allgemein gilt für eine Basis  $\mathcal{C}$ , ausgedrückt mit der Matrix  $C$ :

$$M_{\mathcal{A} \rightarrow \mathcal{B}} = M_{\mathcal{C} \rightarrow \mathcal{B}} \cdot M_{\mathcal{A} \rightarrow \mathcal{C}} \quad \text{und} \quad M_{\mathcal{B} \rightarrow \mathcal{A}} = M_{\mathcal{C} \rightarrow \mathcal{A}} \cdot M_{\mathcal{B} \rightarrow \mathcal{C}}$$

### A.6.4 Zahlenbeispiel in $\mathbb{R}^3$

Wir betrachten zwei Basen  $\mathcal{A}$  und  $\mathcal{B}$ , deren Vektoren die Spalten der Matrizen  $A$  und  $B$  bilden, per

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

Also sind die Basisvektoren gegeben durch:

$$\vec{a}_1 = \begin{pmatrix} 1 \\ 0 \\ -3 \end{pmatrix}, \quad \vec{a}_2 = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}, \quad \vec{a}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad \vec{b}_1 = \begin{pmatrix} 4 \\ 1 \\ -1 \end{pmatrix}, \quad \vec{b}_2 = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}, \quad \vec{b}_3 = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}$$

Diese Komponenten beziehen sich auf die Standardbasis  $E_3$ .

Für die weiteren Rechnungen brauchen wir zunächst die Inversen von  $A$  und  $B$ . Wir lösen dafür die entsprechenden erweiterten LGS mit den üblichen Zeilenoperationen im Gauß-Verfahren:

- Für  $A$ :

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ -3 & 0 & 1 & 0 & 0 & 1 \end{array}\right) \begin{array}{l} \leftarrow -2 \\ \\ \end{array} \Leftrightarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & -2 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ -3 & 0 & 1 & 0 & 0 & 1 \end{array}\right) \begin{array}{l} \\ \leftarrow 3 \\ \leftarrow \end{array} \Leftrightarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & -2 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 3 & -6 & 1 \end{array}\right)$$

Damit ist also:

$$A^{-1} = \begin{pmatrix} 1 & -2 & 0 \\ 0 & 1 & 0 \\ 3 & -6 & 1 \end{pmatrix}$$

- Für  $B$ :

$$\begin{aligned} \left(\begin{array}{ccc|ccc} 4 & 2 & 1 & 1 & 0 & 0 \\ 1 & 0 & 2 & 0 & 1 & 0 \\ -1 & 1 & 2 & 0 & 0 & 1 \end{array}\right) \begin{array}{l} \leftarrow -4 \\ \\ \leftarrow 1 \end{array} &\Leftrightarrow \left(\begin{array}{ccc|ccc} 0 & 2 & -7 & 1 & -4 & 0 \\ 1 & 0 & 2 & 0 & 1 & 0 \\ 0 & 1 & 4 & 0 & 1 & 1 \end{array}\right) \begin{array}{l} \leftarrow -2 \\ \\ \leftarrow \end{array} \Leftrightarrow \left(\begin{array}{ccc|ccc} 0 & 0 & -15 & 1 & -6 & -2 \\ 1 & 0 & 2 & 0 & 1 & 0 \\ 0 & 1 & 4 & 0 & 1 & 1 \end{array}\right) \begin{array}{l} -1 \\ 15 \\ 15 \end{array} \\ \\ \Leftrightarrow \left(\begin{array}{ccc|ccc} 0 & 0 & 15 & -1 & 6 & 2 \\ 15 & 0 & 30 & 0 & 15 & 0 \\ 0 & 15 & 60 & 0 & 15 & 15 \end{array}\right) \begin{array}{l} \leftarrow -2 \\ \\ \leftarrow \end{array} -4 &\Leftrightarrow \left(\begin{array}{ccc|ccc} 0 & 0 & 15 & -1 & 6 & 2 \\ 15 & 0 & 0 & 2 & 3 & -4 \\ 0 & 15 & 0 & 4 & -9 & 7 \end{array}\right) \end{aligned}$$

Jetzt sind nur noch die Zeilen zu permutieren und der globale Faktor auszuklammern, und die Inverse ist gefunden:

$$B^{-1} = \frac{1}{15} \begin{pmatrix} 2 & 3 & -4 \\ 4 & -9 & 7 \\ -1 & 6 & 2 \end{pmatrix}$$

Jetzt sei ein Vektor  $\vec{y}$  in der Standardbasis gegeben als

$$\vec{y} := \begin{pmatrix} 3 \\ 4 \\ -1 \end{pmatrix} = 3\vec{e}_1 + 4\vec{e}_2 - 1\vec{e}_3$$

Wir berechnen die Komponenten von  $\vec{y}$  in der Basis  $\mathcal{A}$ :

$$\vec{x} = M_{E_3 \rightarrow \mathcal{A}} \cdot \vec{y} = A^{-1} \cdot \vec{y} = 3 \begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix} + 4 \begin{pmatrix} -2 \\ 1 \\ -6 \end{pmatrix} - 1 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 - 8 + 0 \\ 0 + 4 + 0 \\ 9 - 24 - 1 \end{pmatrix} = \begin{pmatrix} -5 \\ 4 \\ -16 \end{pmatrix}$$

Zur Probe rechnen wir nach, dass die Linearkombination der Spalten von  $A$  mit diesen Koeffizienten wieder (in der Standardbasis) den Vektor  $\vec{y}$  ergibt:

$$\vec{y} = M_{\mathcal{A} \rightarrow E_3} \cdot \vec{x} = A \cdot \vec{x} = (-5) \begin{pmatrix} 1 \\ 0 \\ -3 \end{pmatrix} + 4 \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} - 16 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -5 + 8 + 0 \\ 0 + 4 + 0 \\ 15 + 0 - 16 \end{pmatrix} = \begin{pmatrix} 3 \\ 4 \\ -1 \end{pmatrix} \quad \checkmark$$

Das Gleiche für die Basis  $\mathcal{B}$ :

$$\begin{aligned} \vec{z} = M_{E_3 \rightarrow \mathcal{B}} \cdot \vec{y} &= B^{-1} \cdot \vec{y} = \frac{1}{15} \left[ 3 \begin{pmatrix} 2 \\ 4 \\ -1 \end{pmatrix} + 4 \begin{pmatrix} 3 \\ -9 \\ 6 \end{pmatrix} - 1 \begin{pmatrix} -4 \\ 7 \\ 2 \end{pmatrix} \right] = \frac{1}{15} \begin{pmatrix} 6 + 12 + 4 \\ 12 - 36 - 7 \\ -3 + 24 - 2 \end{pmatrix} \\ &= \frac{1}{15} \begin{pmatrix} 22 \\ -31 \\ 19 \end{pmatrix} \end{aligned}$$

Und die Probe (Linearkombination der Spalten von  $B$  mit diesen Koeffizienten):

$$\begin{aligned}\vec{y} &= M_{\mathcal{B} \rightarrow E_3} \cdot \vec{z} = B \cdot \vec{z} = \frac{1}{15} \left[ 22 \begin{pmatrix} 4 \\ 1 \\ -1 \end{pmatrix} - 31 \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} + 19 \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix} \right] = \frac{1}{15} \begin{pmatrix} 88 - 62 + 19 \\ 22 + 0 + 38 \\ -22 - 31 + 38 \end{pmatrix} \\ &= \frac{1}{15} \begin{pmatrix} 45 \\ 60 \\ -15 \end{pmatrix} = \begin{pmatrix} 3 \\ 4 \\ -1 \end{pmatrix} \quad \checkmark\end{aligned}$$


---

Jetzt wollen wir noch die Vektoren  $\vec{x}$  und  $\vec{z}$  direkt ineinander umrechnen, indem wir die direkten Basiswechsel-Matrizen ermitteln: Für den Wechsel von  $\mathcal{A}$  nach  $\mathcal{B}$  bekommen wir:

$$\begin{aligned}M_{\mathcal{A} \rightarrow \mathcal{B}} &= M_{E_n \rightarrow \mathcal{B}} \cdot M_{\mathcal{A} \rightarrow E_n} = B^{-1} \cdot A \\ &= \frac{1}{15} \begin{pmatrix} 2 & 3 & -4 \\ 4 & -9 & 7 \\ -1 & 6 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ -3 & 0 & 1 \end{pmatrix} = \frac{1}{15} \begin{pmatrix} 14 & 7 & -4 \\ -17 & -1 & 7 \\ -7 & 4 & 2 \end{pmatrix}\end{aligned}$$

Und in umgekehrter Richtung:

$$\begin{aligned}M_{\mathcal{B} \rightarrow \mathcal{A}} &= M_{E_n \rightarrow \mathcal{A}} \cdot M_{\mathcal{B} \rightarrow E_n} = A^{-1} \cdot B \\ &= \begin{pmatrix} 1 & -2 & 0 \\ 0 & 1 & 0 \\ 3 & -6 & 1 \end{pmatrix} \begin{pmatrix} 4 & 2 & 1 \\ 1 & 0 & 2 \\ -1 & 1 & 2 \end{pmatrix} = \begin{pmatrix} 2 & 2 & -3 \\ 1 & 0 & 2 \\ 5 & 7 & -7 \end{pmatrix}\end{aligned}$$

Man kann nachrechnen, dass in der Tat  $M_{\mathcal{A} \rightarrow \mathcal{B}} \cdot M_{\mathcal{B} \rightarrow \mathcal{A}} = \mathbb{1}_3$

---

Nun noch zur Umrechnung der Darstellungen:

Von  $\mathcal{A}$  nach  $\mathcal{B}$ : Hier setzen wir die Darstellung  $\vec{x}$  aus der Basis  $\mathcal{A}$  ein und multiplizieren von links die Basiswechselmatrix  $M_{\mathcal{A} \rightarrow \mathcal{B}}$ :

$$\begin{aligned}M_{\mathcal{A} \rightarrow \mathcal{B}} \cdot \vec{x} &= \frac{1}{15} \begin{pmatrix} 14 & 7 & -4 \\ -17 & -1 & 7 \\ -7 & 4 & 2 \end{pmatrix} \cdot \begin{pmatrix} -5 \\ 4 \\ -16 \end{pmatrix} = \frac{1}{15} \begin{pmatrix} -70 + 28 + 64 \\ 85 - 4 - 112 \\ 35 + 16 - 32 \end{pmatrix} \\ &= \frac{1}{15} \begin{pmatrix} 22 \\ -31 \\ 19 \end{pmatrix} = \vec{z} \quad \checkmark\end{aligned}$$

Und in umgekehrter Richtung:

$$\begin{aligned}M_{\mathcal{B} \rightarrow \mathcal{A}} \cdot \vec{z} &= \begin{pmatrix} 2 & 2 & -3 \\ 1 & 0 & 2 \\ 5 & 7 & -7 \end{pmatrix} \cdot \frac{1}{15} \begin{pmatrix} 22 \\ -31 \\ 19 \end{pmatrix} = \frac{1}{15} \begin{pmatrix} 44 - 62 - 57 \\ 22 + 0 + 38 \\ 110 - 217 - 133 \end{pmatrix} \\ &= \frac{1}{15} \begin{pmatrix} -75 \\ 60 \\ -240 \end{pmatrix} = \begin{pmatrix} -5 \\ 4 \\ -16 \end{pmatrix} = \vec{x} \quad \checkmark\end{aligned}$$

## A.7 Matrizenoperationen für das Gaußverfahren

In diesem Abschnitt wollen wir rechtfertigen, dass die Operationen bei Manipulation eines Gauß-Schemas (für ein LGS) durch invertierbare Matrizen dargestellt werden können, was die Äquivalenzzeichen zwischen den Umformungen erklärt. Beim Konstruieren der Zeilen-Stufen-Form dürfen Gleichungen (Zeilen des Gauß-Schemas) vertauscht, skaliert oder skaliert zu anderen Gleichungen/Zeilen addiert werden. Insgesamt entsteht aus dem Schema  $(A \mid \vec{y})$  dann ein Schema  $(B \mid \vec{z})$ , welches sich in Zeilen-Stufen-Form befindet. Falls  $A, B \in \mathbb{R}^{(m,n)}$ , werden wir nun zeigen, dass diese drei Operationen durch Multiplikation bestimmter Matrizen aus  $\mathbb{R}^{(m,m)}$  von links vermittelbar sind.

Hierzu wollen wir darauf verzichten, eventuelle Null-Zeilen aus der Koeffizientenmatrix zu streichen, sodass auch das Schema  $(B \mid \vec{z})$  nach wie vor  $m$  Zeilen hat.



Insgesamt wird also gelten:

$$B = M \cdot A \quad \text{und} \quad \vec{z} = M \cdot \vec{y}$$

Die beiden LGS (bzw. ihre Gauß-Schemata) sind genau dann äquivalent zueinander, wenn  $M$  eine reguläre (invertierbare) Matrix ist. Wir zeigen, dass  $M$  sich als Produkt einfacher invertierbarer Matrizen schreiben lässt; die Invertierbarkeit von  $M$  folgt dann mit Satz 7.18.

### A.7.1 Vertauschen von Zeilen

Das Vertauschen der Zeilen  $j$  und  $k \neq j$  einer Matrix lässt sich durch Multiplikation mit einer *Permutationsmatrix* umsetzen. Nach Satz 7.26 (Wirkung von Permutationsmatrizen) benötigen wir hierfür die Matrix

$$P_{\tau_{j,k}}^T$$

Nun sind Permutationsmatrizen aber nach der Bemerkung bei Definition 7.25 *orthogonal*; daher handelt es sich bei unserer Matrix um die Inverse von  $P_{\tau_{j,k}}$ , beziehungsweise um die Matrix der inversen Permutation  $P_{\tau_{j,k}^{-1}}$ . Da aber die Transposition  $\tau_{j,k} \in S_m$  selbstinvers ist, reicht es,  $P_{\tau_{j,k}}$  zu verwenden:

**Satz A.10** (Vertauschen von Zeilen eines Gauß-Schemas). *Für ein Gauß-Schema mit Koeffizientenmatrix  $A \in \mathbb{R}^{(m,n)}$  lassen sich die Zeilen  $j$  und (o.B.d.A.)  $k > j$  vertauschen, indem von links die Matrix*

$$P_{\tau_{j,k}} = (\cdots \quad \vec{e}_{j-1} \quad \vec{e}_k \quad \vec{e}_{j+1} \quad \cdots \quad \vec{e}_{k-1} \quad \vec{e}_j \quad \vec{e}_{k+1} \quad \cdots)$$

*multipliziert wird.*

*Die Matrix ist (wie  $\tau_{j,k}$ ) selbstinvers.*

**Bemerkung:** Die Matrix erhält man aus  $\mathbb{1}_n$ , indem man die Spalten  $j$  und  $k$  vertauscht.

**Beispiel:** Wir vertauschen im LGS

$$(A \mid \vec{y}) := \left( \begin{array}{ccccc|c} 1 & 4 & 2 & 0 & 1 & -2 \\ 2 & 1 & 1 & 0 & 3 & 1 \\ 1 & 3 & -2 & 1 & 0 & 2 \\ -1 & 1 & 0 & 0 & 2 & 1 \\ 1 & 0 & 0 & 1 & 2 & 0 \\ 3 & -4 & 1 & 2 & 0 & 3 \end{array} \right)$$

die Zeilen  $j = 3$  und  $k = 5$ . Die Transformationsmatrix ist nach obigem Satz

$$M := \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Dann ist für das System  $(B \mid \vec{z})$ :

$$B := MA = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 4 & 2 & 0 & 1 \\ 2 & 1 & 1 & 0 & 3 \\ 1 & 3 & -2 & 1 & 0 \\ -1 & 1 & 0 & 0 & 2 \\ 1 & 0 & 0 & 1 & 2 \\ 3 & -4 & 1 & 2 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 4 & 2 & 0 & 1 \\ 2 & 1 & 1 & 0 & 3 \\ 1 & 0 & 0 & 1 & 2 \\ -1 & 1 & 0 & 0 & 2 \\ 1 & 3 & -2 & 1 & 0 \\ 3 & -4 & 1 & 2 & 0 \end{pmatrix}$$

$$\vec{z} := M\vec{y} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -2 \\ 1 \\ 2 \\ 1 \\ 0 \\ 3 \end{pmatrix} = \begin{pmatrix} -2 \\ 1 \\ 0 \\ 1 \\ 2 \\ 3 \end{pmatrix}$$

Also

$$(B \mid \vec{z}) = (MA \mid M\vec{y}) = \left( \begin{array}{ccccc|c} 1 & 4 & 2 & 0 & 1 & -2 \\ 2 & 1 & 1 & 0 & 3 & 1 \\ 1 & 0 & 0 & 1 & 2 & 0 \\ -1 & 1 & 0 & 0 & 2 & 1 \\ 1 & 3 & -2 & 1 & 0 & 2 \\ 3 & -4 & 1 & 2 & 0 & 3 \end{array} \right)$$

### A.7.2 Skalieren von Zeilen mit $c \neq 0$

Um die Zeile  $j$  mit dem Faktor  $c \in \mathbb{R} \setminus \{0\}$  zu skalieren, reicht es, die Transformationsmatrix  $M$  als  $\mathbb{1}_m$  zu wählen und dann an  $j$ -ter Position auf der Diagonalen statt 1 den Faktor  $c$  einzutragen. Da  $c$  verschieden von 0 ist, ist diese Operation invertierbar; die inverse Matrix enthält an gleicher Position den Eintrag  $\frac{1}{c}$ :

**Satz A.11** (Skalieren von Zeilen eines Gauß-Schemas). *Für ein Gauß-Schema mit Koeffizientenmatrix  $A \in \mathbb{R}^{(m,n)}$  lässt sich die Zeile  $j$  mit dem Faktor  $c \in \mathbb{R} \setminus \{0\}$  skalieren, indem von links die Matrix*

$$(\cdots \quad \vec{e}_{j-1} \quad c \cdot \vec{e}_j \quad \vec{e}_{j+1} \quad \cdots)$$

multipliziert wird.

Die inverse Operation hat in Spalte  $j$  den Spaltenvektor  $\frac{1}{c} \cdot \vec{e}_j$ .

**Bemerkung:** Die Determinante der Matrix ist nach Satz 8.24 (Determinanten von Dreiecksmatrizen) genau  $c \neq 0$ , also ist auch formal geklärt, dass sie invertierbar ist.

**Beispiel:** Für das System  $(A \mid \vec{y})$  von oben skalieren wir die dritte Zeile mit  $c := -3$ . Die Transformationsmatrix ist dann:

$$M := \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Dann ist für das System  $(B \mid \vec{z})$ :

$$\begin{aligned} B := MA &= \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 4 & 2 & 0 & 1 \\ 2 & 1 & 1 & 0 & 3 \\ 1 & 3 & -2 & 1 & 0 \\ -1 & 1 & 0 & 0 & 2 \\ 1 & 0 & 0 & 1 & 2 \\ 3 & -4 & 1 & 2 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 4 & 2 & 0 & 1 \\ 2 & 1 & 1 & 0 & 3 \\ -3 & -9 & 6 & -3 & 0 \\ -1 & 1 & 0 & 0 & 2 \\ 1 & 0 & 0 & 1 & 2 \\ 3 & -4 & 1 & 2 & 0 \end{pmatrix} \\ \vec{z} := M\vec{y} &= \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -2 \\ 1 \\ 2 \\ 1 \\ 0 \\ 3 \end{pmatrix} = \begin{pmatrix} -2 \\ 1 \\ -6 \\ 1 \\ 0 \\ 3 \end{pmatrix} \end{aligned}$$

Also

$$(B \mid \vec{z}) = (MA \mid M\vec{y}) = \left( \begin{array}{ccccc|c} 1 & 4 & 2 & 0 & 1 & -2 \\ 2 & 1 & 1 & 0 & 3 & 1 \\ -3 & -9 & 6 & -3 & 0 & -6 \\ -1 & 1 & 0 & 0 & 2 & 1 \\ 1 & 0 & 0 & 1 & 2 & 0 \\ 3 & -4 & 1 & 2 & 0 & 3 \end{array} \right)$$

### A.7.3 Addition der mit $c$ skalierten Zeile $j$ zu Zeile $k \neq j$

Auch hier können wir für die Transformationsmatrix von  $\mathbb{1}_m$  ausgehen. In der Zeile  $k$  befindet sich dann zunächst der Eintrag 1 bei Spalte  $k$ . Weiterhin ergänzen wir dort in Spalte  $j$  noch einen Eintrag  $c$  – dies liefert genau das Gewünschte.

Für die inverse Operation reicht es, das  $(-c)$ -Fache von Zeile  $j$  auf Zeile  $k$  zu addieren. Entsprechend ergänzt man zur Einheitsmatrix  $\mathbb{1}_m$  in Spalte  $j$  und Zeile  $k$  den Eintrag  $(-c)$ . Bevor wir dies als Satz zusammen fassen und formal die Invertierbarkeit (die hier nicht mehr völlig offensichtlich ist) zeigen, betrachten wir zunächst ein

**Beispiel:** In unserem obigen Beispielsystem  $(A \mid \vec{y})$  addieren wir das  $(-3)$ -fache von Zeile 3 auf Zeile 5. Die entsprechende Transformationsmatrix ist:

$$M := \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & -3 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Dann ist für das System  $(B \mid \vec{z})$ :

$$B := MA = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & -3 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 4 & 2 & 0 & 1 \\ 2 & 1 & 1 & 0 & 3 \\ 1 & 3 & -2 & 1 & 0 \\ -1 & 1 & 0 & 0 & 2 \\ -2 & -9 & 6 & -2 & 2 \\ 3 & -4 & 1 & 2 & 0 \end{pmatrix}$$

$$\vec{z} := M\vec{y} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & -3 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -2 \\ 1 \\ 2 \\ 1 \\ 0 \\ 3 \end{pmatrix} = \begin{pmatrix} -2 \\ 1 \\ 2 \\ 1 \\ -6 \\ 3 \end{pmatrix}$$

Also

$$(B \mid \vec{z}) = (MA \mid M\vec{y}) = \left( \begin{array}{cccccc|c} 1 & 4 & 2 & 0 & 1 & -2 \\ 2 & 1 & 1 & 0 & 3 & 1 \\ 1 & 3 & -2 & 1 & 0 & 2 \\ -1 & 1 & 0 & 0 & 2 & 1 \\ -2 & -9 & 6 & -2 & 2 & -6 \\ 3 & -4 & 1 & 2 & 0 & 3 \end{array} \right)$$

Wir halten fest:

**Satz A.12** (Addition skalierten Zeilen im Gauß-Schema). *Für ein Gauß-Schema mit Koeffizientenmatrix  $A \in \mathbb{R}^{(m,n)}$  lässt sich die mit  $c \in \mathbb{R}$  Zeile  $j$  zur Zeile  $k \neq j$  addieren, indem von links die Matrix*

$$(\cdots \quad \vec{e}_{j-1} \quad (\vec{e}_j + c \cdot \vec{e}_k) \quad \vec{e}_{j+1} \quad \cdots)$$

*multipliziert wird.*

*Die inverse Operation hat in Spalte  $j$  den Spaltenvektor  $(\vec{e}_j - c \cdot \vec{e}_k)$*

**Beweis für die Invertierbarkeit:** Wir schreiben die Matrizen komponentenweise. Falls  $M$  die ursprüngliche Transformationsmatrix und  $M^{-1}$  deren Inverse ist, gilt:

$$M_{r,s} = \delta_{r,s} + c\delta_{k,r}\delta_{j,s} \quad \text{und} \quad (M^{-1})_{r,s} = \delta_{r,s} - c\delta_{k,r}\delta_{j,s}$$

So wird für  $M$  im Vergleich zur Einheitsmatrix  $\mathbb{1}_m$  noch in Zeile  $k$  (also für  $r = k$ ) und Spalte  $j$  (also für  $s = j$ ) das Element  $c$  addiert.

Wir berechnen nun das Produkt der beiden Matrizen komponentenweise. Alle Summen sind auszuführen von 1 bis  $m$ :

$$\begin{aligned}
(M \cdot M^{-1})_{r,s} &= \sum_t M_{r,t} \cdot M_{t,s}^{-1} = \sum_t \left( (\delta_{r,t} + c\delta_{k,r}\delta_{j,t})(\delta_{t,s} - c\delta_{k,t}\delta_{j,s}) \right) \\
&= \sum_t \left( \delta_{r,t}\delta_{t,s} - c\delta_{r,t}\delta_{k,t}\delta_{j,s} + c\delta_{k,r}\delta_{j,t}\delta_{t,s} - c^2\delta_{k,r}\delta_{j,t}\delta_{k,t}\delta_{j,s} \right) \\
&= \delta_{r,s} - c\delta_{r,k}\delta_{j,s} + c\delta_{k,r}\delta_{j,s} - c^2\delta_{k,r}\delta_{j,k}\delta_{j,s} \\
&= \delta_{r,s} - c^2\delta_{k,r}\delta_{j,k}\delta_{j,s} \\
&= \delta_{r,s} = (\mathbb{1}_m)_{r,s}
\end{aligned}$$

Hierbei haben wir verwendet, dass das Kronecker-Delta symmetrisch in seinen zwei Operanden ist. Außerdem verschwindet der Beitrag  $c^2$ , da  $k \neq j$  und daher  $\delta_{j,k} = 0$ . Die Rechnung für  $M^{-1} \cdot M = \mathbb{1}_m$  funktioniert genau analog und mag gerne zur Übung ausgeführt werden. Die beiden Matrizen wie oben definiert sind also tatsächlich invers zueinander. ■

#### A.7.4 Zusammenfassung

Wir haben mit den Sätzen A.10, A.11 und A.12 gezeigt, dass die Zeilenoperationen für Gauß-Schemata mit invertierbaren Matrizen durchführbar sind, die jeweils von links an Koeffizientenmatrix bzw. Inhomogenität multipliziert werden.

Die kombinierte Transformationsmatrix  $M$ , die ein System  $(A \mid \vec{y})$  in  $p$  Schritten in die Zeilen-Stufen-Form (oder einen beliebigen Zwischenstand)  $(B \mid \vec{z}) = (MA \mid M\vec{y})$  überführt, ist mit den einzelnen elementaren Transformationsmatrizen  $M_1, \dots, M_p$  für die einzelnen Schritte schreibbar als Produkt

$$M := M_p \cdot M_{p-1} \cdot \dots \cdot M_2 \cdot M_1$$

und ist insgesamt invertierbar per

$$M^{-1} = M_1^{-1} \cdot M_2^{-1} \cdot \dots \cdot M_{p-1}^{-1} \cdot M_p^{-1}$$

Daher sind die Schemata  $(A \mid \vec{y})$  und  $(B \mid \vec{z})$  (also die durch sie beschriebenen LGSs) logisch äquivalent zueinander: alle einzelnen Schritte können “zurück abgewickelt” werden, sodass aus dem Schema  $(B \mid \vec{z})$  per  $M^{-1}$  wieder das Ausgangsschema  $(A \mid \vec{y})$  erzeugbar ist. Dies rechtfertigt die Verknüpfung der einzelnen Gauß-Schemata mit Äquivalenzzeichen.

### A.8 Regel von Cramer für LGS

Angenommen, eine Matrix  $A \in \mathbb{R}^{(n,n)}$  erfülle  $\det A \neq 0$ . Sie hat also vollen Rang  $n$ , und jedes LGS  $(A \mid \vec{y})$  ist eindeutig lösbar.

Satz 8.2 (Äquivalente Beschreibungen von LGS) besagt, dass zu solch einem LGS mit Lösung  $\vec{x}$  dann die Inhomogenität  $\vec{y}$  als Linearkombination der Spalten von  $A = (\vec{a}_1 \ \dots \ \vec{a}_n)$  schreibbar ist:

$$\vec{y} = \sum_{j=1}^n x_j \vec{a}_j$$

Nun setzen wir diese Linearkombination *anstatt* Spalte  $k$  in die Matrix  $A$  ein und berechnen die Determinante dieser neuen Matrix

$$(\dots \ \vec{a}_{k-1} \ \vec{y} \ \vec{a}_{k+1} \ \dots)$$

Unter Beachtung der Multilinearität der Determinante (siehe Definition 8.16) erhalten wir:

$$\begin{aligned}
\det(\dots \ \vec{a}_{k-1} \ \vec{y} \ \vec{a}_{k+1} \ \dots) &= \det\left(\dots \ \vec{a}_{k-1} \ \sum_{j=1}^n x_j \vec{a}_j \ \vec{a}_{k+1} \ \dots\right) \\
&= \sum_{j=1}^n x_j \cdot \det(\dots \ \vec{a}_{k-1} \ \vec{a}_j \ \vec{a}_{k+1} \ \dots)
\end{aligned}$$

Nun verschwinden in dieser Summe jedoch aufgrund der Alterniertheit der Determinanten alle Beiträge  $j \neq k$  – denn dann steht in Spalte  $k$  jeweils ein Vektor  $\vec{a}_j$ , der genau der Spalte  $j$  entspricht. Es bleibt also nur:

$$\dots = x_k \cdot \det(\dots \ \vec{a}_{k-1} \ \vec{a}_k \ \vec{a}_{k+1} \ \dots) = x_k \cdot \det A$$

Aber dann haben wir eine (eindeutige!) Methode gefunden, um  $x_k$  zu berechnen, wenn wir nur noch durch  $\det A \neq 0$  kürzen:

**Satz A.13** (Regel von Cramer).<sup>5</sup> Für ein eindeutig lösbares LGS  $(A \mid \vec{y})$  mit  $A \in \mathbb{R}^{(n,n)}$  und  $\det A \neq 0$  ist die Lösung der Vektorgleichung  $A\vec{x} = \vec{y}$  komponentenweise gegeben durch:

$$x_k = \frac{\det(\cdots \vec{a}_{k-1} \quad \vec{y} \quad \vec{a}_{k+1} \cdots)}{\det A}$$

**Bemerkung:** Die Formel ist genau für die Matrizen definiert, die nichtverschwindende Determinante besitzen.

**Beispiele:**

- Wir betrachten die dreireihige Matrix

$$A := \begin{pmatrix} 2 & 1 & 4 \\ 3 & -2 & 1 \\ 2 & 2 & 1 \end{pmatrix}$$

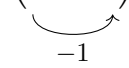
mit  $\det A = 31 \neq 0$ , die wir bereits im Beispiel zur Leibniz-Formel (Satz 8.20) kennen gelernt hatten.

Wir lösen nun das Gleichungssystem für den Ergebnisvektor

$$\vec{y} := \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}$$

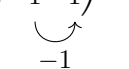
Wir betrachten die erste Komponente  $x_1$ :

$$x_1 = \frac{1}{31} \cdot \det \begin{pmatrix} 2 & 1 & 4 \\ 1 & -2 & 1 \\ 1 & 2 & 1 \end{pmatrix} = \frac{1}{31} \cdot \det \begin{pmatrix} 2 & 1 & 2 \\ 1 & -2 & 0 \\ 1 & 2 & 0 \end{pmatrix} \stackrel{\text{S.3}}{=} \frac{2}{31} \cdot \det \begin{pmatrix} 1 & -2 \\ 1 & 2 \end{pmatrix} = \frac{2}{31} \cdot (2 - (-2)) = \frac{8}{31}$$

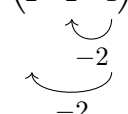


Die anderen beiden Komponenten von  $\vec{x}$  errechnen sich analog:

$$x_2 = \frac{1}{31} \cdot \det \begin{pmatrix} 2 & 2 & 4 \\ 3 & 1 & 1 \\ 2 & 1 & 1 \end{pmatrix} = \frac{1}{31} \cdot \det \begin{pmatrix} 2 & 2 & 2 \\ 3 & 1 & 0 \\ 2 & 1 & 0 \end{pmatrix} \stackrel{\text{S.3}}{=} \frac{2}{31} \cdot \det \begin{pmatrix} 3 & 1 \\ 2 & 1 \end{pmatrix} = \frac{2}{31} \cdot (3 - 2) = \frac{2}{31}$$



$$x_3 = \frac{1}{31} \cdot \det \begin{pmatrix} 2 & 1 & 2 \\ 3 & -2 & 1 \\ 2 & 2 & 1 \end{pmatrix} = \frac{1}{31} \cdot \det \begin{pmatrix} -2 & -3 & 2 \\ 1 & -4 & 1 \\ 0 & 0 & 1 \end{pmatrix} \stackrel{\text{Z.3}}{=} \frac{1}{31} \cdot \det \begin{pmatrix} -2 & -3 \\ 1 & -4 \end{pmatrix} = \frac{1}{31} \cdot (8 - (-3)) = \frac{11}{31}$$



Zur Probe stellen wir die Linearkombination der Spalten von  $A$  mit den drei gefundenen Koeffizienten auf:

$$\frac{1}{31} \left[ 8 \begin{pmatrix} 2 \\ 3 \\ 2 \end{pmatrix} + 2 \begin{pmatrix} 1 \\ -2 \\ 2 \end{pmatrix} + 11 \begin{pmatrix} 4 \\ 1 \\ 1 \end{pmatrix} \right] = \frac{1}{31} \begin{pmatrix} 16 + 2 + 44 \\ 24 - 4 + 11 \\ 16 + 4 + 11 \end{pmatrix} = \frac{1}{31} \begin{pmatrix} 62 \\ 31 \\ 31 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} \quad \checkmark$$

---

<sup>5</sup>G. Cramer, schweiz. Mathematiker

- Wir berechnen die Inverse der allgemeinen zweireihigen Matrix

$$A := \begin{pmatrix} a & b \\ c & d \end{pmatrix} =: (\vec{a}_1 \quad \vec{a}_2)$$

wobei  $\det A \neq 0$  voraus gesetzt wird.

Dabei haben wir für  $A^{-1} := (\vec{x}_1 \quad \vec{x}_2)$  zwei Vektorgleichungen zu lösen:

$$A\vec{x}_1 = \vec{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{und} \quad A\vec{x}_2 = \vec{e}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Für die Determinante von  $A$  gilt:

$$\det A = ad - bc$$

Dann gilt für die Komponenten der beiden Spalten von  $A^{-1}$ :

$$\begin{aligned} (\vec{x}_1)_1 &= \frac{\det(\vec{e}_1 \quad \vec{a}_2)}{\det A} = \frac{\det \begin{pmatrix} 1 & b \\ 0 & d \end{pmatrix}}{\det A} = \frac{d}{\det A} \\ (\vec{x}_1)_2 &= \frac{\det(\vec{a}_1 \quad \vec{e}_1)}{\det A} = \frac{\det \begin{pmatrix} a & 1 \\ c & 0 \end{pmatrix}}{\det A} = \frac{-c}{\det A} \\ (\vec{x}_2)_1 &= \frac{\det(\vec{e}_2 \quad \vec{a}_2)}{\det A} = \frac{\det \begin{pmatrix} 0 & b \\ 1 & d \end{pmatrix}}{\det A} = \frac{-b}{\det A} \\ (\vec{x}_2)_2 &= \frac{\det(\vec{a}_1 \quad \vec{e}_2)}{\det A} = \frac{\det \begin{pmatrix} a & 0 \\ c & 1 \end{pmatrix}}{\det A} = \frac{a}{\det A} \end{aligned}$$

Also insgesamt:

$$\boxed{\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = A^{-1} = (\vec{x}_1 \quad \vec{x}_2) = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}}$$

# Anhang B

## Ausgewählte Beweise

### B.1 Für Kapitel 1

**Beweis zu Satz 1.10 (Verneinung prädikatenlogischer Aussagen) auf Seite 18:**

Wir betrachten zunächst die Aussage  $A$  und führen die Verneinung unter Benutzung textueller Beschreibungen durch:

$$\begin{aligned}\neg A &\equiv \neg(\forall x : a(x)) \\ &\equiv \neg(\text{Für sämtliche } x \text{ gilt: } a(x) \equiv \mathcal{W}) \\ &\equiv \neg(\text{Für kein } x \text{ gilt: } a(x) \equiv \mathcal{F}) \\ &\equiv (\text{Für (mindestens) ein } x \text{ gilt: } a(x) \equiv \mathcal{F}) \\ &\equiv (\text{Für (mindestens) ein } x \text{ gilt: } \neg a(x) \equiv \mathcal{W}) \\ &\equiv (\exists x : \neg a(x))\end{aligned}$$

Analog für  $B$ :

$$\begin{aligned}\neg B &\equiv \neg(\exists x : a(x)) \\ &\equiv \neg(\text{Für (mindestens) ein } x \text{ gilt: } a(x) \equiv \mathcal{W}) \\ &\equiv \neg(\text{Für (mindestens) ein } x \text{ gilt: } \neg a(x) \equiv \mathcal{F}) \\ &\equiv (\text{Für kein } x \text{ gilt: } \neg a(x) \equiv \mathcal{F}) \\ &\equiv (\text{Für sämtliche } x \text{ gilt: } \neg a(x) \equiv \mathcal{W}) \\ &\equiv (\forall x : \neg a(x))\end{aligned}$$

Damit ist alles gezeigt. ■

(Die vielleicht etwas unintuitivere Rechnung für  $\neg B$  hätte man auch direkt aus der leichter verständlichen für  $A$  gewinnen können, indem ein Prädikat  $b(x) := \neg a(x)$  eingeführt wird. Setzt man in der Rechnung für  $\neg A$  überall das Komplement  $a(x) \equiv \neg b(x)$  ein, verneint die gesamte Aussage (Gesetz zur doppelten Negation benutzen) und liest sie in umgekehrter Schrittreihenfolge, so erhält man eine analoge Rechnung wie oben für  $\neg B$ .)

**Beweis zu Satz 1.31 (Binomische Formeln) auf Seite 43:**

Wir wenden das Distributivgesetz an (auch einmal rückwärts für den vorletzten Schritt), dazu das Assoziativgesetz für die Addition und das Kommutativgesetz für die Multiplikation. Für die erste binomische Formel ergibt sich dann:

$$\begin{aligned}
 (a+b)^2 &= (a+b) \cdot (a+b) \\
 &= ((a+b) \cdot a) + ((a+b) \cdot b) \\
 &= (a \cdot a + b \cdot a) + (a \cdot b + b \cdot b) \\
 &= a \cdot a + b \cdot a + a \cdot b + b \cdot b \\
 &= a^2 + b \cdot a + a \cdot b + b^2 \\
 &= a^2 + a \cdot b + a \cdot b + b^2 \\
 &= a^2 + 1 \cdot (ab) + 1 \cdot (ab) + b^2 \\
 &= a^2 + (1+1) \cdot (ab) + b^2 \\
 &= a^2 + 2 \cdot ab + b^2 \quad \blacksquare
 \end{aligned}$$


---

**Beweis zu Satz 1.36 (Dreiecksungleichung) auf Seite 47:**

Wir unterscheiden zwei Fälle:

- Falls  $x + y \geq 0$ , so gilt  $|x + y| = x + y$ . Da aber stets  $x \leq |x|$  und  $y \leq |y|$  gilt, ist auch

$$|x + y| = x + y \leq |x| + |y|$$

- Falls dagegen  $x + y < 0$ , so gilt  $|x + y| = -(x + y) = -x - y = (-x) + (-y)$ . Auch die negativen Werte von  $x$  und  $y$  können höchstens den jeweiligen Beträgen entsprechen; also auch hier:

$$|x + y| = (-x) + (-y) \leq |x| + |y|$$

Damit ist alles gezeigt.  $\blacksquare$

---

**Beweis zum Kleinen Gauß-Trick auf Seite 48:**

Wir nennen die Summe der ersten  $n$  natürlichen Zahlen  $S_n$ . Dann schreiben wir die Summanden einmal in aufsteigender, und darunter in absteigender Anordnung an:

$$\begin{array}{rcccccc}
 S_n & = & 1 & + & 2 & + & \cdots & + & (n-1) & + & n \\
 \wedge & S_n & = & n & + & (n-1) & + & \cdots & + & 2 & + & 1
 \end{array}$$

Wir addieren beide Gleichungen. Wenn wir die Addition der rechten Seite spaltenweise geordnet durchführen, dann haben wir genau  $n$  Spalten mit jeweils der Summe  $(n+1)$ .

Damit erhalten wir:  $2S_n = n \cdot (n+1)$ , und nach Skalierung mit dem Faktor  $\frac{1}{2}$  die behauptete Gleichung.  $\blacksquare$

---

**Beweis zu Satz 1.40 (Anordnungen einer endlichen Menge) auf Seite 50:**

Wir bauen für eine  $n$ -elementige Menge  $M$  das  $n$ -Tupel Stück für Stück auf. Zunächst füllen wir die  $n$  Elemente der Menge in Gedanken in ein Gefäß; dort ziehen wir nun nacheinander die Elemente wie aus einer Lostrommel.

Für die erste Komponente des Tupels gibt es  $n$  Möglichkeiten, da das gedachte Gefäß noch ganz gefüllt ist. Danach ist die erste Tupel-Komponente festgelegt.

Für die nächste (zweite) Komponente enthält das Gefäß dann nur noch  $(n-1)$  Elemente aus  $M$ . Entsprechend gibt es für die Wahl der zweiten Tupelkomponente auch nur genau  $(n-1)$  verschiedene Möglichkeiten.

Und so weiter: für die vorletzte Tupel-Komponente haben wir nur zwei verschiedene Möglichkeiten – das eine oder das andere der verbliebenen zwei Elemente aus dem Gefäß. Ist auch diese vorletzte Komponente gewählt, dann ergibt sich die letzte Tupel-Komponente direkt von selbst, da nur noch ein Element im Gefäß verblieben ist.



Die Gesamtzahl möglicher verschiedener Tupel erhalten wir, indem wir diese Möglichkeiten miteinander multiplizieren. Da die obige Argumentation unabhängig von den konkreten Elementen von  $M$  war, gilt sie allgemein und führt auf

$$n \cdot (n-1) \cdot (n-2) \cdot \dots \cdot 2 \cdot 1 = n!$$

Möglichkeiten, ein  $n$ -Tupel aus den Elementen von  $M$  zu konstruieren. ■

**Beweis zu Satz 1.42 (Addition benachbarter Binomialkoeffizienten) auf Seite 51:**

Wir schreiben die Definitionen für die linke Seite der Gleichung an und klammern geeignet aus, wobei wir in beiden Brüchen jeweils einen Faktor erweitern:

$$\begin{aligned} \binom{n}{k} + \binom{n}{k+1} &= \frac{n!}{k! \cdot (n-k)!} + \frac{n!}{(k+1)! \cdot \underbrace{(n-(k+1))!}_{n-k-1}} \\ &= \frac{n! \cdot (k+1)}{(k+1) \cdot k! \cdot (n-k)!} + \frac{n! \cdot (n-k)}{(k+1)! \cdot (n-k) \cdot (n-k-1)!} \\ &= \frac{n! \cdot (k+1)}{(k+1)! \cdot (n-k)!} + \frac{n! \cdot (n-k)}{(k+1)! \cdot (n-k)!} \\ &= \frac{n!}{(k+1)! \cdot (n-k)!} \cdot ((k+1) + (n-k)) \\ &= \frac{n!}{(k+1)! \cdot (n-k)!} \cdot (n+1) = \frac{n! \cdot (n+1)}{(k+1)! \cdot \underbrace{(n-k) \cdot (n-k-1) \cdot \dots \cdot 1}_{(n+1)-(k+1)}} \\ &= \frac{(n+1)!}{(k+1)! \cdot ((n+1) - (k+1))!} = \binom{n+1}{k+1} \quad \blacksquare \end{aligned}$$

**Beweis zu Satz 1.43 (Teilmengen einer endlichen Menge) auf Seite 51:**

Wir gehen genauso vor wie im Beweis von Satz 1.40 und füllen die Elemente von  $M$  zunächst in ein gedachtes Gefäß.

Nun konstruieren wir aber kein  $n$ -Tupel (bei dem wir immer sämtliche Elemente von  $M$ ) nacheinander ziehen, sondern ein  $k$ -Tupel mit  $0 \leq k \leq n$ . Mit analoger Argumentation wie oben ist dies auf

$$n \cdot (n-1) \cdot \dots \cdot (n-k+1)$$

Weisen möglich. Somit ist die Zahl möglicher  $k$ -Tupel gegeben als:

$$n \cdot (n-1) \cdot \dots \cdot (n-k+1) = n \cdot (n-1) \cdot \dots \cdot (n-k+1) \cdot \underbrace{\frac{(n-k)!}{(n-k)!}}_1 = \frac{n!}{(n-k)!}$$

Wenn nun allerdings nur nach der Zahl möglicher  $k$ -elementiger Teilmengen von  $M$  gefragt ist, ist diese Zahl zu groß, da die  $k$ -Tupel geordnet sind: Für jede Teilmenge mit  $k$  Elementen haben wir nach Satz 1.40 die Möglichkeit, diese  $k$  Elemente auf  $k!$  Weisen zu Tupeln anzuordnen – und jede dieser Anordnungen wurde oben einzeln in Betracht gezogen. Also erhalten wir die Anzahl möglicher  $k$ -elementiger Teilmengen (die ungeordnet sind), indem wir obige Tupel-Anzahl noch durch  $k!$  dividieren. Dann ergeben sich insgesamt

$$\frac{n!}{(n-k)!} \cdot \frac{1}{k!} = \frac{n!}{k! \cdot (n-k)!} = \binom{n}{k}$$

Möglichkeiten für  $k$ -elementige Teilmengen von  $M$ . ■

## B.2 Für Kapitel 2

**Beweis zu Satz 2.2 (Anzahl der Teiler von natürlichen Zahlen) auf Seite 53:**

Wenn wir bei der Suche der Teiler vorgehen wie im zweiten Beispiel von Definition 2.1 beschrieben,

müssen die Werte  $a, k$  mit  $n = ka$  sich in verschiedene Richtungen bewegen;  $a$  steigt mit jedem Schritt an und  $k$  sinkt mit jedem Schritt ab (denn sonst ergäbe sich ein Produkt größer als  $n$ , wenn  $a$  weiter erhöht wird).

Für  $n \in \mathbb{N}$  sind  $a = 1$  und  $k = n$  stets Teiler, und für  $n > 1$  ist in diesem Fall auch  $a \neq k$ . Der Fall  $n = 1$  führt auf genau einen Teiler, aber wegen  $1 = 1 \cdot 1$  ist die Eins auch eine Quadratzahl.

Wenn wir nun  $a$  stets weiter erhöhen und  $k$  stets weiter absenken, solange  $a < k$ , können wir im Extremfall die Situation  $a = k = \sqrt{n}$  erreichen. Das ist aber nur genau für die Quadratzahlen möglich; für genau diese gilt nämlich  $\sqrt{n} \in \mathbb{N}$ .

$a$  braucht nicht über  $\sqrt{n}$  hinaus erhöht zu werden, da die Multiplikation kommutativ ist (symmetrisch bzgl. der beiden Argumente) und ab dann keine neuen komplementären Teilerpaare mehr gefunden werden.

Für  $a < \sqrt{n}$  muss aber  $k > \sqrt{n}$  gelten (sonst wäre  $ka < n$ ) – somit sind dann  $a, k$  stets verschieden. Bei der Suche nach den Teilern gewinnen wir also für nicht-Quadratzahlen stets pro Schritt zwei Teiler hinzu, falls  $n = ka$  erfüllt ist – somit bleibt deren Anzahl gerade. Weil wir bei Quadratzahlen den Teiler  $\sqrt{n}$  nicht doppelt zählen, sind letztere die einzigen Fälle mit  $n \in \mathbb{N}$ , für die eine ungerade Anzahl von Teilern möglich ist. ■

### Beweis zu Satz 2.3 (Teiler eines Produkts) auf Seite 53:

Die Gleichheit von Mengen zeigen wir über beidseitige Teilmengenbeziehung:

- Zunächst sind alle  $j \cdot k$  mit  $j \in T_a$  und  $k \in T_b$  sicher Teiler von  $a \cdot b$ , also Elemente von  $T_{a \cdot b}$ .  
Denn dann gibt es  $r, s$ , sodass  $a = j \cdot r$  und  $b = k \cdot s$ . Dann ist aber wegen der Kommutativität und Assoziativität der Multiplikation in  $\mathbb{Z}$  auch:

$$a \cdot b = (j \cdot r) \cdot (k \cdot s) = (j \cdot k) \cdot (r \cdot s)$$

Also ist  $j \cdot k$  ein Teiler von  $a \cdot b$ .

- Die andere Teilmengenbeziehung gilt ebenfalls, denn  $a$  und  $b$  können stets als Produkte von zwei Teilern notiert werden:

$$a = p \cdot q \quad \text{und} \quad b = r \cdot s$$

mit  $p, q \in T_a$  und  $r, s \in T_b$ . Und das ist auch die einzige Art, auf die solche Produkte schreibbar sind. Aber dann ist auch:

$$a \cdot b = (p \cdot q) \cdot (r \cdot s) = (p \cdot r) \cdot (q \cdot s)$$

Aber dann ist  $p \cdot r$  ein Produkt eines Teilers aus  $T_a$  mit einem aus  $T_b$ , und ebenso  $q \cdot s$

Damit ist alles gezeigt. ■

### Beweis zu Satz 2.6 (Primteiler von zusammen gesetzten Zahlen) auf Seite 54:

Indirekter Beweis: Angenommen, es gebe zusammen gesetzte Zahlen, die die behauptete Eigenschaft nicht haben. Wir wählen unter allen solchen Zahlen die kleinste aus – dies ist immer möglich, da die natürlichen Zahlen sich ordnen lassen. Sei  $n$  also die kleinste solche Zahl.

Da  $n$  zusammen gesetzt ist, gibt es nichttriviale Teiler  $a, k$  mit  $n = ka$ . O.B.d.A. sei  $a \leq k$ , also auch  $a \leq \sqrt{n}$  (siehe das Argument dazu im Beweis zu Satz 2.2). Falls nun  $a$  prim wäre, hätten wir einen Widerspruch zur Behauptung.

Falls  $a$  hingegen nicht prim ist, muss  $a$  aber zusammen gesetzt sein, da  $a > 1$  ( $a$  ist ja nicht-trivialer Teiler von  $n$ ). Da  $n$  jedoch als die kleinste zusammen gesetzte Zahl gewählt wurde, die keinen Primteiler besitzt, der höchstens ihrer Wurzel entspricht, muss  $a$  wegen  $a < n$  also einen Primteiler  $j$  besitzen, für den  $j \leq \sqrt{a}$  gilt. Damit gilt aber wegen  $a < n$  auch direkt  $j \leq \sqrt{n}$ . Und da  $j$  prim ist und  $j \mid a$ , gilt wegen  $a \mid n$  auch  $j \mid n$ . Auch hier erhalten wir einen Widerspruch zur Behauptung.

Es kann also gar keine kleinste zusammen gesetzte Zahl geben, die nicht die im Satz behauptete Eigenschaft hat – folglich müssen alle zusammen gesetzten Zahlen die Eigenschaft erfüllen, und damit gilt die Behauptung. ■

**Beweis zu Satz 2.14 (Teilerfremdheit von mehr als zwei Zahlen) auf Seite 59:**

Wir finden in der Menge  $M$  also zwei Zahlen  $m, n$  mit  $\text{ggT}(m, n) = 1$ . Dann gilt für die beiden zugehörigen Teiler-Mengen von  $m$  und  $n$ :

$$T_m \cap T_n = \{1\}$$

Ab jetzt ist es unwesentlich, ob andere Zahlen aus  $M$  möglicherweise mehr Teiler besitzen, denn die Menge  $T_M$  der gemeinsamen Teiler entsteht durch Schnitt aller einzelnen Teiler-Mengen. Wegen der Assoziativität der Schnittoperation können wir immer  $T_m$  und  $T_n$  zuerst schneiden; ab dann kann der Schnitt nicht mehr größer werden als  $\{1\}$ . Kleiner wird er allerdings auch nicht, da 1 stets gemeinsamer Teiler ist. ■

**Beweis zu Satz 2.15 (Skalierung und ggT) auf Seite 59:**

Sei  $m := \text{ggT}(n_1, n_2, \dots, n_k)$ . Skaliert man sämtliche Elemente von  $M$  mit dem Faktor  $q \in \mathbb{N}$ , so ist klar, dass  $q \cdot m$  auch ein gemeinsamer Teiler der Menge

$$\{q \cdot n_1, q \cdot n_2, \dots, q \cdot n_k\}$$

ist.

Es ist weiterhin auch der größte gemeinsame Teiler dieser Menge, denn nach Satz 2.3 werden die Teiler der Produkte  $q \cdot n_j$  gebildet aus den Produkten der Teiler von  $q$  und denen von  $n_j$ . Die höchsten solchen Teiler entstehen durch Multiplikation der Elemente von  $T_{n_j}$  mit  $q$  selbst. Wenn also die obige Menge einen gemeinsamen Teiler größer als  $q \cdot m$  enthalten würde, hätte es zuvor schon einen gemeinsamen Teiler der unskalierten Menge  $M$  gegeben, der größer als  $m$  wäre, im Widerspruch zur Annahme. ■

**Beweis zu Satz 2.16 (Gemeinsame Teiler und ggT) auf Seite 59:**

Sei  $m := \text{ggT}(M)$ . Dann ist zu zeigen, dass Menge  $T_M$  der gemeinsamen Teiler identisch ist mit der Menge der Teiler von  $m$ , also  $T_m$ . Wir zeigen die Gleichheit der beiden Mengen, indem wir die Teilmengenbeziehung in beiden Richtungen fest stellen:

- Zunächst ist jeder Teiler von  $m$  auch ein gemeinsamer Teiler von  $M$ . Denn da jede der Zahlen aus  $M$  die Zahl  $m$  als Teiler enthält, und da für einen Teiler  $a$  von  $m$  ein  $b$  existiert, sodass  $m = a \cdot b$ , so ist für eine Zahl  $n_j \in M$ :

$$n_j = m \cdot \frac{n_j}{m} = (a \cdot b) \cdot \frac{n_j}{m} = a \cdot \left(b \cdot \frac{n_j}{m}\right)$$

Der Bruch  $\frac{n_j}{m}$  ist aus  $\mathbb{Z}$ , da  $m$  ein Teiler von  $n_j$  ist. Wegen der Assoziativität der Multiplikation auf  $\mathbb{Z}$  dürfen wir die Klammerung umstellen wie oben ausgeführt, und erhalten, dass  $a$  auch Teiler jedes  $n_j \in M$  ist, also in der Menge  $T_M$  der gemeinsamen Teiler enthalten ist.

- Umgekehrt bleibt zu zeigen, dass jeder gemeinsame Teiler, also jedes Element aus  $T_M$  auch ein Teiler von  $m$  ist. Sei  $c \in T_M$  ein solcher gemeinsamer Teiler. Dann betrachten wir die Zahlen  $\nu_1, \nu_2, \dots, \nu_k$ , die jeweils durch Division der Elemente von  $M$  mit  $c$  entstehen:

$$\nu_j := \frac{n_j}{c}$$

Alle  $\nu_j$  sind ganze Zahlen, da  $c$  jeweils Teiler von  $n_j$  ist.

Sei nun  $r := \text{ggT}(\nu_1, \nu_2, \dots, \nu_k)$  der ggT dieser mit  $\frac{1}{c}$  skalierten Zahlen. Nach Satz 2.15 gilt:

$$\text{ggT}(c \cdot \nu_1, c \cdot \nu_2, \dots, c \cdot \nu_k) = c \cdot \text{ggT}(\nu_1, \nu_2, \dots, \nu_k) = c \cdot r$$

Die linke Seite dieser Gleichungskette entspricht aber genau dem  $\text{ggT}(n_1, n_2, \dots, n_k)$ , also gilt:

$$m = c \cdot r$$

Aber damit ist  $c$  dann ein Teiler des ggT  $m$ .

Damit ist alles gezeigt. ■

**Beweis zu Satz 2.17 (Assoziativität des ggT) auf Seite 60:**

Nach Definition 2.11 ist der ggT von  $M$  das maximale Element der Menge  $T_M$  der gemeinsamen Teiler der Zahlen aus  $M$ . Die Menge  $T_M$  hingegen entsteht nach Definition 2.10 durch Schnitt aller einzelnen Teiler-Mengen  $T_{n_1}, T_{n_2}, \dots, T_{n_k}$ .

Die Schnitt-Operation ist jedoch assoziativ (siehe Satz 1.22 (Rechenregeln für Mengen)) – also kommt es beim Schneiden der Teiler-Mengen von  $n_1, n_2, \dots, n_k$  nicht auf die Klammerung an. Somit können die Mengen  $P$  und  $Q = M \setminus P$  beliebig gewählt werden; der Schnitt der dort gruppierten Teilmengen muss am Ende auf die Menge  $T_M$  führen. Aber dann ist auch das Maximum dieser Menge unabhängig von der konkreten Klammerung. ■

**Beweis zu Satz 2.18 (Invarianz des ggT bei Differenzbildung) auf Seite 61:**

Der größte gemeinsame Teiler von  $m$  und  $n$  ist ein gemeinsamer Teiler. Wir definieren  $a := \text{ggT}(m, n)$ . Nach dem Satz 2.9 (Teiler bei Summe oder Differenz) ist  $a$  auch Teiler von  $n - m$ .

Es bleibt noch zu zeigen, dass es keinen *größeren* Teiler von  $m$  und  $(n - m)$  gibt. Beweis durch Widerspruch (“indirekter Beweis”): Wir nehmen an, es gäbe solch einen größeren Teiler  $b > a$ , sodass mit  $r, s \in \mathbb{N}$  gilt:

$$m = b \cdot r \quad \text{und} \quad (n - m) = b \cdot s \quad (*)$$

Dann addieren wir (nach Satz 1.30 (Skalierung und Addition von (Un-)Gleichungen)) die beiden Gleichungen in  $(*)$  und erhalten mit dem Distributivgesetz:

$$m + (n - m) = (b \cdot r) + (b \cdot s) = b \cdot (r + s)$$

Die linke Seite dieser Gleichung vereinfacht sich noch zu  $m + (n - m) = n$ .

Aber dann ist  $b$  nicht nur ein Teiler von  $m$  (linke Gleichung in  $(*)$ ), sondern auch ein Teiler von  $n$  – und dann ist  $a$  wegen  $a < b$  nicht der größte gemeinsame Teiler von  $m$  und  $n$ . Das steht im Widerspruch ( $\neq$ ) zur Wahl von  $a$  am Beginn des Beweises. Die Annahme, es gebe einen größeren Teiler  $b$  von  $m$  und  $(n - m)$  muss also falsch sein. Und da  $a$  wie oben erwähnt, Teiler von  $m$  und  $(n - m)$  ist, ist es mithin auch der größte solche Teiler. ■

**Beweis zu Satz 2.22 (Lemma von Euklid) auf Seite 69:**

Da  $\text{ggT}(a, n) = 1$ , gibt es nach Satz 2.21 (Lemma von Bezout) ganze Zahlen  $\alpha, \beta$ , sodass

$$1 = \alpha \cdot a + \beta \cdot n$$

Wir multiplizieren diese Gleichung in  $\mathbb{Z}$  noch mit der ganzen Zahl  $b$  und erhalten (unter Benutzung des Assoziativgesetzes für die Multiplikation):

$$b = \alpha \cdot (a \cdot b) + \beta \cdot n \cdot b$$

Nun ist aber nach Voraussetzung  $n$  ein Teiler von  $(a \cdot b)$ , also dem linken Term der rechten Seite von obiger Gleichung. Weiterhin ist  $n$  offensichtlich ein Teiler des Produkts  $\beta \cdot n \cdot b$ . Folglich kann  $n$  nach dem Distributivgesetz aus beiden Termen ausgeklammert werden und somit gilt:

$$b = n \cdot \left( \alpha \cdot \frac{a \cdot b}{n} + \beta \cdot b \right)$$

Der Bruch in obiger Klammer ist eine ganze Zahl, und also ist der Zahlenwert der gesamten Klammer aus  $\mathbb{Z}$ . Damit ist aber  $n$  ein Teiler von  $b$ . ■

**Beweis zu Satz 2.23 (Teilerfremdheit eines Produktes mit einer Zahl) auf Seite 69:**

Für den Beweis nutzen wir aus, dass die Äquivalenzbeziehung auch zwischen den logisch gegenteiligen Aussagen besteht, und dass allgemein der ggT von zwei Zahlen genau dann größer ist als 1, wenn es irgendeinen gemeinsamen Teiler größer als 1 gibt:

Ein Produkt  $m$  von  $k$  Faktoren  $a_1, a_2, \dots, a_k$  hat genau dann einen gemeinsamen Teiler größer als 1 mit  $n$ , wenn mindestens ein Faktor  $a_j$  (mit  $1 \leq j \leq k$ ) einen gemeinsamen Teiler größer als 1 mit  $n$  besitzt.

Wir zeigen beide Richtungen:

“ $\Leftarrow$ ”: Falls der Faktor  $a_j$  einen gemeinsamen Teiler  $b > 1$  mit  $n$  besitzt, dann teilt  $b$  als Teiler von  $a_j$  auch das gesamte Produkt  $m$ . Folglich ist  $b > 1$  ein gemeinsamer Teiler von  $m$  und  $n$ .

“ $\Rightarrow$ ”: Sei  $b > 1$  ein gemeinsamer Teiler von  $m$  und  $n$ . Wir gehen nun die Faktoren  $a_1$  bis  $a_k$  von  $m$  der Reihe nach durch.

Sei also zunächst  $m = a_1 \cdot m_1$  mit

$$m_1 := \prod_{j=2}^k a_j$$

Nun gilt sicher  $b \mid m$ , also  $b \mid (a_1 \cdot m_1)$ . Wir betrachten  $b_1 := \text{ggT}(b, a_1)$  unterscheiden zwei Fälle:

- Falls  $b_1 > 1$ , ist schon alles gezeigt, denn  $b_1$  ist damit ein Teiler von  $a_1$ . Da  $b_1$  ein Teiler von  $b$  ist, ist  $b_1$  auch ein Teiler von  $n$  (denn  $b$  ist ein gemeinsamer Teiler von  $m$  und  $n$ ). Aber damit ist  $b_1 > 1$  gemeinsamer Teiler von  $a_1$  und von  $n$ .
- Falls jedoch  $b_1 = 1$ , also falls  $b$  und der Faktor  $a_1$  teilerfremd sind, verwenden wir das Lemma von Euklid (Satz 2.22) – dieses liefert uns, dass  $b$  dann ein Teiler von  $m_1$  sein muss.

$m_1$  ist ein Produkt mit  $(k-1)$  Faktoren. Wir verfahren wieder analog (rekursiv!) und spalten auf:  $m_1 = a_2 \cdot m_2$ , mit

$$m_2 := \prod_{j=3}^k a_j$$

Betrachten wir nun  $b_2 := \text{ggT}(b, a_2)$ , so können wir wiederum die Fälle  $b_2 > 1$  und  $b_2 = 1$  unterscheiden. Für  $b_2 > 1$  argumentieren wir wie oben und haben einen gemeinsamen Teiler von  $a_2$  und  $n$  gefunden, der größer ist als 1. Ansonsten teilt  $b$  nach Euklid das Produkt  $m_2$ .

Und so weiter, bis ein günstiger Fall (also  $b_j = \text{ggT}(b, a_j) > 1$ ) eintritt. Falls  $b$  teilerfremd mit  $a_1$  bis  $a_{k-1}$  ist, folgt im vorletzten Schritt nach Euklid, dass  $b$  ein Teiler von  $a_k$  sein muss (denn  $m_{k-1} = a_k$ ). Nach spätestens  $k$  Schritten bricht die Kaskade also erfolgreich ab.

Damit ist alles gezeigt. ■

## B.3 Für Kapitel 3

**Beweis zu Satz 3.2 (Äquivalenzrelation “ $\equiv \pmod{m}$ ”) auf Seite 73:**

Sei  $m \in \mathbb{N}$  gegeben und fest. Es sind drei Kriterien zu prüfen, um zu zeigen, dass die Relation aus Definition 3.1 eine Äquivalenzrelation ist:

- Reflexivität: Für jedes  $a \in \mathbb{Z}$  gilt:  $a \equiv a \pmod{m}$ , da  $(a - a) = 0 = 0 \cdot m$ .
- Symmetrie: Für alle  $a, b \in \mathbb{Z}$  gilt:

$$\begin{aligned} a \equiv b \pmod{m} &\Leftrightarrow m \mid (b - a) \\ &\Leftrightarrow \exists k \in \mathbb{Z} : (b - a) = k \cdot m \\ &\Leftrightarrow \exists k \in \mathbb{Z} : -(b - a) = (a - b) = (-k) \cdot m \\ &\Leftrightarrow \exists \tilde{k} \in \mathbb{Z} : (a - b) = \tilde{k} \cdot m \\ &\Leftrightarrow m \mid (a - b) \\ &\Leftrightarrow b \equiv a \pmod{m} \end{aligned}$$

Hierbei wurde  $\tilde{k} := (-k)$  gewählt.

- Transitivität: Für alle  $a, b, c \in \mathbb{Z}$  ist zu zeigen:

$$(a \equiv b \pmod{m}) \wedge (b \equiv c \pmod{m}) \Rightarrow (a \equiv c \pmod{m})$$

Wenn also die Äquivalenzen modulo  $m$  zwischen  $a$  und  $b$  sowie zwischen  $b$  und  $c$  bestehen, gilt:

$$m \mid (b - a) \quad \text{und} \quad m \mid (c - b)$$

Dann gibt es ganze Zahlen  $k, l$ , sodass  $(b - a) = k \cdot m$  und  $(c - b) = l \cdot m$ . Also gilt:

$$b = a + k \cdot m \quad \text{und} \quad c = b + l \cdot m$$

Setzen wir die erste Gleichung in die zweite ein, so erhalten wir:

$$c = b + l \cdot m = (a + k \cdot m) + l \cdot m = a + (k + l) \cdot m$$

Aber damit gilt auch:  $(c - a) = (k + l) \cdot m$ , also  $m \mid (c - a)$ , und damit  $a \equiv c \pmod{m}$

Alle drei Kriterien treffen zu; somit ist die Äquivalenz modulo  $m$  tatsächlich eine Äquivalenzrelation. ■

### **Beweis zu Satz 3.7 (Erweitern von Äquivalenzbeziehungen modulo $m$ ) auf Seite 76:**

Wir setzen Definition 3.1 ein und erhalten folgende Schlusskette:

$$\begin{aligned} x \equiv y \pmod{m} &\Leftrightarrow m \mid (x - y) \\ &\Rightarrow m \mid j \cdot (x - y) \\ &\Leftrightarrow m \mid (j \cdot x - j \cdot y) \\ &\Leftrightarrow j \cdot x \equiv j \cdot y \pmod{m} \end{aligned}$$

Hierbei haben wir zunächst ausgenutzt, dass der Teiler einer Zahl auch stets Teiler aller Vielfachen dieser Zahl ist, und dann das Distributivgesetz für  $\mathbb{Z}$  verwendet. ■

### **Beweis zu Satz 3.8 (Kürzungsregel modulo $m$ ) auf Seite 77:**

Wir verfolgen die Beweiskette von Satz 3.7 in umgekehrter Richtung. Die entscheidende Stelle ist:

$$m \mid (x - y) \quad \Rightarrow \quad m \mid j \cdot (x - y)$$

Tatsächlich gilt die Implikation für  $\text{ggT}(j, m) = 1$  auch in der umgekehrten Richtung, und zwar nach Satz 2.22 (Lemma von Euklid). Unter dieser Voraussetzung lässt sich also die Schlusskette im Beweis von Satz 3.7 auch umgekehrt als Implikation lesen, und es folgt die Behauptung. ■

### **Beweis zu Satz 3.12 (Nullteiler und multiplikativ Inverse in $\mathbb{Z}_m$ ) auf Seite 80:**

Wir zeigen beide Richtungen der Implikation:

“ $\Rightarrow$ ”: Sei also  $a$  Nullteiler von  $\mathbb{Z}_m$ , dann gibt es ein  $x \in \mathbb{Z}_m \setminus \{0\}$ , sodass  $a \cdot x = 0$ .

Wir beweisen indirekt, dass es dann kein Inverses von  $a$  geben kann. Angenommen nämlich, es gäbe ein  $b \in \mathbb{Z}_m \setminus \{0\}$ , sodass  $a \cdot b = 1$ , dann dürfen wir auch schreiben:

$$x = x \cdot 1 = x \cdot (a \cdot b)$$

Wegen des Assoziativgesetzes für die Multiplikation dürfen wir umklammern:

$$\dots \Rightarrow x = (x \cdot a) \cdot b$$

Aber mit der Nullteiler-Eigenschaft von  $a$  gilt dann:

$$\dots \Rightarrow x = 0 \cdot b = 0 \quad \text{!}$$

Dies ist aber ein Widerspruch zur obigen Wahl von  $x \in \mathbb{Z}_m \setminus \{0\}$ ; also war die Annahme falsch, und es kann kein Inverses zu  $a$  geben.

“ $\Leftarrow$ ”: Wir finden also unter den positiven Zahlen aus  $\mathbb{Z}_m$  kein Inverses von  $a$ . Dann haben alle  $m$  Produkte  $a \cdot 0, a \cdot 1, a \cdot 2, \dots, a \cdot (m-1)$  Zahlenwerte verschieden von 1. Da wir uns auf eindeutige Vertreter der Restklassen beschränken, so haben wir für die  $m$  Produkte nur höchstens  $(m-1)$  verschiedene Zahlenwerte in  $\mathbb{Z}_m$  zur Verfügung. Das kann nur erfüllbar sein, wenn mindestens zwei dieser Produkte den gleichen Zahlenwert modulo  $m$  haben (*Schubfachprinzip*: Verteilt

man  $m$  Objekte auf  $n$  Kategorien mit  $n < m$ , dann muss wenigstens eine Kategorie mehr als ein Objekt abbekommen).

Sei also  $y \in \mathbb{Z}_n$  ein solcher (mindestens) doppelt auftretender Zahlenwert, und seien  $b, c \in \mathbb{Z}_m$  mit  $b \neq c$  die zu  $a$  passenden Faktoren, sodass gilt:

$$a \cdot b = a \cdot c = y$$

Damit ergibt sich (das Distributivgesetz verwendet):

$$0 \equiv y - y = a \cdot b - a \cdot c = a \cdot (b - c)$$

Da aber, wegen  $b \neq c$  auch  $(b - c) \neq 0$  gilt, und da ohnehin  $a \neq 0$ , sind beide Zahlen Nullteiler von  $\mathbb{Z}_m$ ; insbesondere ist also  $a$  ein Nullteiler von  $\mathbb{Z}_m$ .

Damit ist alles gezeigt. ■

### **Beweis zu Satz 3.13 (Invertierbare Elemente von $\mathbb{Z}_m$ ) auf Seite 81:**

Mit dem gleichen Argument wie im Beweisteil “ $\Leftarrow$ ” von Satz 3.12 ergibt sich (als indirekter Beweis): Wenn nicht jeder Zahlenwert genau einmal vorkäme, müsste einer der Werte mindestens doppelt vorkommen, für  $a \cdot b$  und  $a \cdot c$ , also in Zeile  $a$  für die Spalten  $b$  und  $c$ , mit  $b \neq c$  in  $\mathbb{Z}_n$ , also  $b \neq c \pmod n$ .

Dann wäre die Differenz der Produkte aber äquivalent zu Null, und,  $a$  ausgeklammert, ergäbe sich, dass  $a$  Nullteiler wäre – aber dann wäre  $a$  nach Satz 3.12 nicht invertierbar. ✗

Also enthält jede Zeile/Spalte der Multiplikationstabelle für invertierbare Elemente (also nicht-Nullteiler) genau die Zahlen aus  $\mathbb{Z}_n$  je einmal (die Null tritt in der Regel nicht auf, da diese Zeile/Spalte meist weggelassen wird). ■

### **Beweis zu Satz 3.14 (Multiplikative Invertierbarkeit modulo $m$ ) auf Seite 81:**

Hier verwenden wir die beiden Eigenschaften, die wir in den Anwendungen des ggT gezeigt hatten.

“ $\Rightarrow$ ”:  $a$  besitze also ein Inverses  $a^{-1}$ , sodass  $a \cdot a^{-1} \equiv 1 \pmod m$ . Nun ist aber jede Zahl aus der Restklasse  $[1]$  teilerfremd zu  $m$ , denn alle diese Zahlen lassen sich schreiben als  $k \cdot m + 1$  mit einem  $k \in \mathbb{Z}$ . Für den Nachweis der Teilerfremdheit berufen wir uns auf den Satz 2.19 über die Invarianz des ggT bei Division mit Rest. Also ist auch  $a \cdot a^{-1}$  teilerfremd zu  $m$ .

Nun verwenden wir Satz 2.23 über die Teilerfremdheit von Faktoren und ihrem Produkt (von rechts nach links) und erhalten, dass sowohl  $a$  als auch  $a^{-1}$  teilerfremd zu  $n$  sind.

“ $\Leftarrow$ ”:  $a$  sei also teilerfremd zu  $m$ , also  $\text{ggT}(a, m) = 1$ . Nach dem Lemma von Bezout wissen wir, dass es ganze Zahlen  $\alpha, \beta \in \mathbb{Z}$  gibt, sodass

$$\alpha \cdot a + \beta \cdot m = 1$$

Da  $\beta \cdot m$  ein Vielfaches von  $m$  ist, können wir entsprechend schreiben (von rechts nach links):

$$1 \equiv \alpha \cdot a \pmod m$$

Aber damit ist das Inverse von  $a$  gefunden, nämlich  $a^{-1} \equiv \alpha \pmod m$ , und wir müssen nur den eindeutigen Repräsentanten von  $[\alpha]$  modulo  $n$  suchen; dieser liegt in  $\mathbb{Z}_m$ .

Damit ist alles gezeigt. ■

**Beweis zu Satz 3.16 (Restklassensysteme modulo Primzahlen) auf Seite 82:** Wenn  $p$  eine Primzahl ist, sind alle positiven Zahlen in  $\mathbb{Z}_p$  teilerfremd zu  $p$  (siehe Satz 2.12). Folglich haben alle diese Zahlen nach Satz 3.14 auch multiplikative Inverse, und damit ist nach Satz 3.12 keine dieser Zahlen Nullteiler.

Ist  $p$  dagegen keine Primzahl, so ist  $p$  zusammen gesetzt und besitzt also Teiler ungleich null in  $\mathbb{Z}_p$ . Diese Teiler sind Nullteiler modulo  $p$ . ■

**Beweis zu Satz 3.17 (Kleiner Satz von Fermat) auf Seite 83:**

Da  $a$  nicht äquivalent zu 0 ist, liegt  $a$  nach Satz 3.16 in einer der Restklassen aus  $\mathbb{Z}_p^*$ . Wir betrachten deren eindeutigen Repräsentanten  $j$  der Restklasse  $[a]$ .

Wir betrachten die  $j$ -te Zeile der Multiplikationstabelle von  $\mathbb{Z}_p^*$ . Nach Satz 3.13 über die Einträge in der Multiplikationstabelle enthält jede Zeile der Tabelle für  $\mathbb{Z}_p^*$  jeweils genau einmal die Zahlen  $1, 2, \dots, (p-1)$ . Das Produkt aller dieser Zahlen ist der Wert der Fakultätsfunktion  $(p-1)!$

Andererseits enthält die zu  $a$  gehörige Zeile die Einträge  $a \cdot 1, a \cdot 2, \dots, a \cdot (p-1)$  (gerechnet modulo  $p$ ). Deren Produkt ist also äquivalent zu  $a^{p-1} \cdot (p-1)!$

Offenbar müssen beide Produkte gleich sein, d.h.

$$a^{p-1} \cdot (p-1)! \equiv (p-1)! \pmod{p}$$

Hier dürfen wir aber noch die Kürzungsregel modulo  $p$  (Satz 3.8) verwenden, denn die Zahl  $(p-1)!$  ist teilerfremd zum Modul  $p$  – dies ergibt sich aus Satz 2.23 (Teilerfremdheit eines Produktes mit einer Zahl); genau dieser Zusammenhang ist dort als Beispiel angeführt.

Nach dem Kürzen dieser Äquivalenzbeziehung mit  $(p-1)!$  ergibt sich dann genau der behauptete Zusammenhang. ■

## B.4 Für Kapitel 5

**Beweis zu Satz 5.5 (Neutrales Element eines Monoids) auf Seite 114:**

Indirekter Beweis: Angenommen,  $e, f$  seien zwei verschiedene neutrale Elemente des Monoids aus der Menge  $M$  mit der Operation  $*$ . Dann gilt, da  $e, f$  neutrale Elemente sind, auch folgende Gleichung (die sich von links nach rechts sowie von rechts nach links lesen lässt):

$$e = e * f = f$$

Von links nach rechts multiplizieren wir  $e$  mit  $f$ , das ja neutrales Element ist; also gilt die linke Gleichheit ( $f$  ist neutral, also auch rechtsneutral). Da  $e$  aber selbst neutral, also auch linksneutral ist, gilt auch die rechte Gleichheit; also  $e = f$  im Widerspruch zur Annahme. ✗

Also gibt es genau ein neutrales Element in einem Monoid. ■

**Beweis zu Satz 5.7 (Eindeutigkeit der inversen Elemente) auf Seite 116:**

Indirekter Beweis. Angenommen, ein Element  $x \in M$  hätte zwei verschiedene inverse Elemente  $a, b \in M$  bezüglich der Operation  $*$ . Dann gilt:

$$x * a = a * x = e \quad \text{und} \quad x * b = b * x = e$$

Wir benutzen die Assoziativität der Verknüpfung  $*$  und erhalten:

$$a = a * e = a * (x * b) = (a * x) * b = e * b = b$$

Damit wäre  $a = b$  im Widerspruch zur Annahme. ✗

Also gibt es zu jedem  $x \in M$  in einer Gruppe genau ein Inverses  $x^{-1}$ . ■

**Beweis zu Satz 5.8 (Inverse einer Verknüpfung) auf Seite 116:**

Wir verknüpfen das behauptete Inverse mit der Verknüpfung  $(a * b)$  und nutzen die Assoziativität:

$$(b^{-1} * a^{-1}) * (a * b) = b^{-1} * (a^{-1} * a) * b = b^{-1} * e * b = b^{-1} * b = e$$

Wegen der Eindeutigkeit des Inversen nach Satz 5.7 folgt die Behauptung. ■

**Beweis zu Satz 5.9 (Kürzungsregel bei Gruppen) auf Seite 116:**

Wir verknüpfen beide Seiten der Gleichung  $a * x = b * x$  von rechts mit  $x^{-1}$ . Das ist immer möglich, da  $x \in M$  und das Inverse  $x^{-1} \in M$  eindeutig bestimmt ist. Also:

$$a * x = b * x \Rightarrow a * x * x^{-1} = b * x * x^{-1} \Leftrightarrow a * e = b * e \Leftrightarrow a = b$$

Für die andere Gleichung verfahren wir analog und multiplizieren  $x^{-1}$  von links; auch hier folgt die Behauptung aufgrund der Eigenschaft des neutralen Elements  $e$ . ■



**Beweis zu Satz 5.28 (Nullstellen und Linearfaktoren) auf Seite 144:**

Wir zeigen beide Implikationsrichtungen der Äquivalenz

$$(\tilde{x} \text{ ist Nullstelle von } p(X)) \Leftrightarrow ((X - \tilde{x}) \text{ ist Linearfaktor von } p(X))$$

“ $\Leftarrow$ ”: Falls  $(X - \tilde{x})$  Linearfaktor ist, so gibt es also ein  $q(X) \in \mathbb{K}[X]$ , sodass  $p(X) = (X - \tilde{x}) \cdot q(X)$ .  
Dann gilt für die Polynomfunktion allerdings auch:

$$f_p(x) = (x - \tilde{x}) \cdot f_q(x)$$

Setzen wir hier den Wert  $\tilde{x}$  ein, so erhalten wir mit dem Satz 5.16 vom Nullprodukt:

$$f_p(\tilde{x}) = (\tilde{x} - \tilde{x}) \cdot f_q(\tilde{x}) = 0 \cdot f_q(\tilde{x}) = 0$$

Also ist  $\tilde{x}$  in der Tat Nullstelle von  $p(X)$ .

“ $\Rightarrow$ ”: Wir wissen, dass  $f_p(\tilde{x}) = 0$  gilt. Nun führen wir nach Satz 5.24 eine Polynomdivision mit dem Modul  $m(X) := (X - \tilde{x})$  aus. Es gibt also eindeutig bestimmte  $q(X), r(X)$ , sodass:

$$p(X) = q(X) \cdot (X - \tilde{x}) + r(X)$$

Wir stellen nach dem Divisionsrest um:

$$\dots \Leftrightarrow r(X) = p(X) - q(X) \cdot (X - \tilde{x})$$

Weiter gilt aber nach Satz 5.24, dass der Grad des Restpolynoms  $r(X)$  kleiner sein muss als der des Moduls. Da der Modul jedoch ein lineares Polynom ist, muss  $r(X)$  demnach eine Zahl  $c \in \mathbb{K}$  (also ein Polynom nullten Grades) sein.

Da  $c$  ein konstantes Polynom ist, muss die Gleichung  $c = p(X) - q(X) \cdot (X - \tilde{x})$  auf ganz  $\mathbb{K}$  gelten. Für die Polynomfunktion folgt entsprechendes. Werten wir diese jedoch an der Stelle  $\tilde{x}$  aus, so erhalten wir:

$$c = f_p(\tilde{x}) - f_q(\tilde{x}) \cdot (\tilde{x} - \tilde{x}) = f_p(\tilde{x})$$

Da aber nach Voraussetzung  $\tilde{x}$  eine Nullstelle war, folgt:  $c = 0$ . Aber dann ist  $p(X)$  *ohne Rest* durch  $m(X)$  teilbar – und folglich ist  $m(X) = (X - \tilde{x})$  nach Definition 5.27 ein Linearfaktor von  $p(X)$ .

Damit ist alles gezeigt. ■

## B.5 Für Kapitel 6

**Beweis zu Satz 6.3 (Skalierung und Nullvektor) auf Seite 150:**

Wir verwenden für die erste Gleichheit die Rechenregeln für Vektorräume:

$$\begin{aligned} 0 \odot \vec{v} &= (0 \cdot 2) \odot \vec{v} \\ &\stackrel{(3)}{=} 0 \odot (2 \odot \vec{v}) \\ &= 0 \cdot \left( (1 + 1) \odot \vec{v} \right) \\ &\stackrel{(1)}{=} 0 \odot \left( (1 \odot \vec{v}) \oplus (1 \odot \vec{v}) \right) \\ &\stackrel{(4)}{=} 0 \odot (\vec{v} \oplus \vec{v}) \\ &\stackrel{(2)}{=} (0 \odot \vec{v}) \oplus (0 \odot \vec{v}) \end{aligned}$$

Diese Gleichung kann nur erfüllt sein, wenn einer der Terme  $(0 \odot \vec{v})$  in der letzten Zeile (und damit natürlich auch der andere) dem neutralen Element  $\vec{0}$  der Vektoraddition entspricht. Also:

$$0 \odot \vec{v} = \vec{0}$$

Für die zweite Gleichheit verwenden wir ebenfalls die Rechenregeln, und die folgende Tatsache, die aus dem eben Gezeigten als Spezialfall direkt folgt:

$$0 \odot \vec{0} = \vec{0}$$

Damit:

$$a \odot \vec{0} = a \odot (0 \odot \vec{0}) \stackrel{(3)}{=} (a \cdot 0) \odot \vec{0} = 0 \odot \vec{v} = \vec{0}$$

Damit sind beide behaupteten Gleichheiten gezeigt. ■

**Beweis zu Satz 6.4 (Kartesischer Produktraum) auf Seite 151:**

Die komponentenweise Addition von Vektorkomponenten aus  $\mathbb{K}$  ergibt wieder Körperelemente; ebenso die Skalierung mit  $a \in \mathbb{K}$ .

Da  $(\mathbb{K}, +)$  eine abelsche Gruppe ist, gilt dies auch für  $(\mathbb{K}^n, \oplus)$ .

Die beiden Distributivgesetze von  $\mathbb{K}$  übertragen sich direkt (und komponentenweise) auf die Rechenregeln 1 und 2 für Vektorräume; die komponentenweise Skalierung führt, da  $(\mathbb{K} \setminus \{0\}, \cdot)$  eine abelsche Gruppe mit neutralem Element 1 ist, direkt auf die Rechenregeln 3 und 4 für Vektorräume. Damit ist die Struktur  $(\mathbb{K}^n, \mathbb{K}, \oplus, \odot)$  ein  $\mathbb{K}$ -Vektorraum. ■

**Beweis zu Satz 6.7 (Cauchy-Schwarz-Ungleichung) auf Seite 154:**

Wir zeigen die quadrierte Form, aus der sich durch Ziehen der Quadratwurzel die im Satz notierte Form ergibt. Das ist zulässig da die Wurzelfunktion auf den nichtnegativen reellen Zahlen *monoton* ist, d.h. für  $x, y \geq 0$  gilt:

$$x \leq y \Leftrightarrow x^2 \leq y^2$$

Also wollen wir folgendes zeigen:

$$\langle \vec{v}, \vec{w} \rangle^2 \leq \langle \vec{v}, \vec{v} \rangle \cdot \langle \vec{w}, \vec{w} \rangle$$

Zunächst stimmt die Gleichung für  $\vec{w} = \vec{0}$ , da dann beide Seiten der Ungleichung den Wert 0 annehmen. Im folgenden sei also  $\vec{w} \neq \vec{0}$ .

Wir nutzen die positive Definitheit des Skalarprodukts aus, um mit einer zunächst beliebigen Zahl  $\lambda \in \mathbb{R}$  die folgende Abschätzung zu erhalten:

$$\begin{aligned} 0 &\leq \langle \vec{v} - \lambda \vec{w}, \vec{v} - \lambda \vec{w} \rangle \\ &= \langle \vec{v}, \vec{v} \rangle - 2\lambda \langle \vec{v}, \vec{w} \rangle + \lambda^2 \langle \vec{w}, \vec{w} \rangle \end{aligned}$$

Für den zweiten Schritt haben wir die Bilinearität und die Symmetrie (zum erhalten des Vorfaktors 2) des Skalarprodukts verwendet.

Nun setzen wir für  $\lambda$  einen speziellen Wert ein; auch für diesen gilt die obige Ungleichung dann weiter:

$$\lambda := \frac{\langle \vec{v}, \vec{w} \rangle}{\langle \vec{w}, \vec{w} \rangle}$$

Der Zahlenwert ist wohldefiniert, da  $\vec{w} \neq \vec{0}$  angenommen wurde und also das Skalarprodukt im Nenner nicht verschwindet.

Weiter:

$$\begin{aligned} \dots \Rightarrow 0 &\leq \langle \vec{v}, \vec{v} \rangle - 2 \frac{\langle \vec{v}, \vec{w} \rangle}{\langle \vec{w}, \vec{w} \rangle} \cdot \langle \vec{v}, \vec{w} \rangle + \left( \frac{\langle \vec{v}, \vec{w} \rangle}{\langle \vec{w}, \vec{w} \rangle} \right)^2 \cdot \langle \vec{w}, \vec{w} \rangle \\ &= \langle \vec{v}, \vec{v} \rangle - 2 \frac{\langle \vec{v}, \vec{w} \rangle^2}{\langle \vec{w}, \vec{w} \rangle} + \frac{\langle \vec{v}, \vec{w} \rangle^2}{\langle \vec{w}, \vec{w} \rangle} \\ &= \langle \vec{v}, \vec{v} \rangle - \frac{\langle \vec{v}, \vec{w} \rangle^2}{\langle \vec{w}, \vec{w} \rangle} \end{aligned}$$

Nun addieren wir den Bruch auf beide Seiten der Ungleichung – dadurch ändert sich die Richtung des Größenvergleichs nicht. Auch das nachträgliche Skalieren mit  $\langle \vec{w}, \vec{w} \rangle$  lässt die Richtung gleich, da dieser Wert positiv ist (positive Definitheit des Skalarprodukts!). Also:

$$\begin{aligned} \dots \Leftrightarrow \frac{\langle \vec{v}, \vec{w} \rangle^2}{\langle \vec{w}, \vec{w} \rangle} &\leq \langle \vec{v}, \vec{v} \rangle \\ \Leftrightarrow \langle \vec{v}, \vec{w} \rangle^2 &\leq \langle \vec{v}, \vec{v} \rangle \cdot \langle \vec{w}, \vec{w} \rangle \quad \blacksquare \end{aligned}$$

**Beweis zu Satz 6.9 (Skalarprodukt-Norm) auf Seite 155:**

Wir zeigen, dass die Abbildung die drei Eigenschaften aus Definition 6.8 erfüllt:

1. Positive Definitheit: Jedes Skalarprodukt ist positiv definit nach Definition 6.5. Damit gilt:  $\langle \vec{v}, \vec{v} \rangle \geq 0$ . Aus nichtnegativen reellen Zahlen lässt sich stets die Wurzel ziehen; diese ist wiederum nichtnegativ. Und sie ist nur dann gleich 0, wenn auch ihr Argument null ist. Aber das Skalarprodukt eines Vektors mit sich selbst ergibt nur für  $\vec{v} = \vec{0}$  den Wert 0 (siehe die Bemerkungen zu Definition 6.5). Entsprechend gilt also auch:

$$(\|\vec{v}\| = 0) \Leftrightarrow (\sqrt{\langle \vec{v}, \vec{v} \rangle} = 0) \Leftrightarrow (\langle \vec{v}, \vec{v} \rangle = 0) \Rightarrow (\vec{v} = \vec{0})$$

2. Dreiecksungleichung: Wir betrachten zunächst das Quadrat der Abbildung und nutzen die Symmetrie des Skalarprodukts aus:

$$\|\vec{v} + \vec{w}\|^2 = \langle \vec{v} + \vec{w}, \vec{v} + \vec{w} \rangle = \langle \vec{v}, \vec{v} \rangle + 2\langle \vec{v}, \vec{w} \rangle + \langle \vec{w}, \vec{w} \rangle$$

Nun vereinbaren wir die (nichtnegativen) reellen Variablen

$$s := \sqrt{\langle \vec{v}, \vec{v} \rangle} \quad \text{und} \quad t := \sqrt{\langle \vec{w}, \vec{w} \rangle}$$

Weiterhin gilt für das Skalarprodukt die Cauchy-Schwarz-Ungleichung (Satz 6.7), sodass folgende Abschätzung richtig ist:

$$\dots = s^2 + 2\langle \vec{v}, \vec{w} \rangle + t^2 \leq s^2 + 2\sqrt{\langle \vec{v}, \vec{v} \rangle} \sqrt{\langle \vec{w}, \vec{w} \rangle} + t^2 = s^2 + 2st + t^2 = (s + t)^2$$

Dann können wir aber durch Wurzelziehen diese Ungleichung für die Beträge folgern:

$$\dots \Rightarrow \|\vec{v} + \vec{w}\| \leq |s + t|$$

Und für die reellen Zahlen auf der rechten Seite gilt nun die gewöhnliche Dreiecksungleichung aus Satz 1.36! Damit ist also

$$\|\vec{v} + \vec{w}\| \leq |s| + |t| = |\sqrt{\langle \vec{v}, \vec{v} \rangle}| + |\sqrt{\langle \vec{w}, \vec{w} \rangle}|$$

Und weil der Betrag einer reellen Wurzel der Wurzel selbst entspricht, ist also

$$\|\vec{v} + \vec{w}\| \leq \|\vec{v}\| + \|\vec{w}\|$$

3. Skalierbarkeit: Das Skalarprodukt ist bilinear. Also gilt:

$$\|a\vec{v}\| = \sqrt{\langle a\vec{v}, a\vec{v} \rangle} = \sqrt{a^2 \cdot \langle \vec{v}, \vec{v} \rangle} = |a| \cdot \sqrt{\langle \vec{v}, \vec{v} \rangle} = |a| \cdot \|\vec{v}\|$$

Damit ist alles gezeigt. ■

### **Beweis zu Satz 6.11 (Norminduzierte Metrik) auf Seite 156:**

Wir zeigen, dass die Norm eines Vektorraums die drei Eigenschaften aus Definition 6.10 erfüllt (unter Benutzung der Eigenschaften aus Definition 6.8):

1. Definitheit: Es gilt folgende Kette von äquivalenten Umformungen (dabei wird die positive Definitheit der Norm benutzt):

$$(d(\vec{v}, \vec{w}) = 0) \Leftrightarrow (\|\vec{v} - \vec{w}\| = 0) \Leftrightarrow (\vec{v} - \vec{w} = \vec{0}) \Leftrightarrow (\vec{v} = \vec{w})$$

2. Symmetrie: Wir verwenden die Skalierungseigenschaft der Norm:

$$d(\vec{v}, \vec{w}) = \|\vec{v} - \vec{w}\| = \|(-1) \cdot (\vec{w} - \vec{v})\| = |-1| \cdot \|\vec{w} - \vec{v}\| = \|\vec{w} - \vec{v}\| = d(\vec{w}, \vec{v})$$

3. Dreiecksungleichung: Wir ergänzen innerhalb der Norm-Abbildung einen Nullvektor geeignet; danach verwenden wir die Dreiecksungleichung der Norm:

$$\begin{aligned} d(\vec{v}, \vec{w}) &= \|\vec{v} - \vec{w}\| = \|\vec{v} - \underbrace{\vec{u} + \vec{u}}_{\vec{0}} - \vec{w}\| = \|(\vec{v} - \vec{u}) + (\vec{u} - \vec{w})\| \\ &\leq \|\vec{v} - \vec{u}\| + \|\vec{u} - \vec{w}\| = d(\vec{v}, \vec{u}) + d(\vec{u}, \vec{w}) \end{aligned}$$

Damit ist alles gezeigt. ■

**Beweis zu Satz 6.12 (Kanonisches Skalarprodukt in  $\mathbb{R}^n$ ) auf Seite 156:**

Wir zeigen, dass die Abbildung die drei Eigenschaften aus Definition 6.5 erfüllt:

1. Symmetrie: Wegen der Kommutativität der Multiplikation auf  $\mathbb{R}$  ist

$$\langle \vec{v}, \vec{w} \rangle = \sum_j v_j w_j = \sum_j w_j v_j = \langle \vec{w}, \vec{v} \rangle$$

2. Bilinearität: Wir nutzen die Assoziativität der Addition und die Distributivgesetze von  $\mathbb{R}$  (das ist ein Körper!), die beide auch für endliche Summen gelten:

$$\begin{aligned} \langle a\vec{u} + b\vec{v}, \vec{w} \rangle &= \sum_j (a\vec{u} + b\vec{v})_j w_j \\ &= \sum_j (au_j + bv_j) w_j \\ &= \sum_j ((au_j w_j) + (bv_j w_j)) \\ &= \left( \sum_j au_j w_j \right) + \sum_j bv_j w_j \\ &= a \left( \sum_j u_j w_j \right) + b \sum_j v_j w_j \\ &= a \langle \vec{u}, \vec{w} \rangle + b \langle \vec{v}, \vec{w} \rangle \end{aligned}$$

3. Positive Definitheit: Für  $\vec{v} \neq \vec{0}$  können nicht sämtliche  $n$  Vektorkomponenten 0 betragen. Sei  $v_k$  eine solche Komponente, also  $v_k \neq 0$ . Dann ist

$$\langle \vec{v}, \vec{v} \rangle = \sum_j (v_j)^2 = v_k^2 + \sum_{j \neq k} (v_j)^2 > \sum_{j \neq k} (v_j)^2 \geq 0$$

Die verbleibende Summe beträgt mindestens 0, da alle Quadrate von reellen Zahlen mindestens 0 betragen. Mit  $v_k$  ist auch  $v_k^2$  größer als 0, sodass sich die geforderte Ungleichung ergibt.

(Die Summe von  $n$  Quadratzahlen kann auch nur dann 0 betragen, wenn sämtliche einzelnen Quadrate null sind – das bedeutet wiederum, dass nur für den Nullvektor  $\vec{0}$  das Skalarprodukt mit sich selbst 0 ergibt.)

Damit ist alles gezeigt. ■

**Beweis zu Satz 6.17 (Eigenschaften des Kreuzprodukts) auf Seite 163:**

Wir zeigen die einzelnen Eigenschaften:

1. Bilinearität in den Faktoren: Wir zeigen die erste Beziehung durch Einsetzen in die Definition 6.16:

$$\begin{aligned} (a\vec{u} + b\vec{v}) \times \vec{w} &= \begin{pmatrix} (a\vec{u} + b\vec{v})_2 \cdot w_3 - (a\vec{u} + b\vec{v})_3 \cdot w_2 \\ (a\vec{u} + b\vec{v})_3 \cdot w_1 - (a\vec{u} + b\vec{v})_1 \cdot w_3 \\ (a\vec{u} + b\vec{v})_1 \cdot w_2 - (a\vec{u} + b\vec{v})_2 \cdot w_1 \end{pmatrix} \\ &= \begin{pmatrix} au_2 w_3 + bv_2 w_3 - au_3 w_2 - bv_3 w_2 \\ au_3 w_1 + bv_3 w_1 - au_1 w_3 - bv_1 w_3 \\ au_1 w_2 + bv_1 w_2 - au_2 w_1 - bv_2 w_1 \end{pmatrix} \\ &= \begin{pmatrix} au_2 w_3 - au_3 w_2 \\ au_3 w_1 - au_1 w_3 \\ au_1 w_2 - au_2 w_1 \end{pmatrix} - \begin{pmatrix} bv_2 w_3 - bv_3 w_2 \\ +bv_3 w_1 - bv_1 w_3 \\ bv_1 w_2 - bv_2 w_1 \end{pmatrix} \\ &= a \begin{pmatrix} u_2 w_3 - u_3 w_2 \\ u_3 w_1 - u_1 w_3 \\ u_1 w_2 - u_2 w_1 \end{pmatrix} + b \begin{pmatrix} v_2 w_3 - v_3 w_2 \\ +v_3 w_1 - v_1 w_3 \\ v_1 w_2 - v_2 w_1 \end{pmatrix} \\ &= a(\vec{u} \times \vec{w}) + b(\vec{v} \times \vec{w}) \end{aligned}$$

Die zweite Beziehung rechnet man analog nach.

2. Verschwinden des Kreuzprodukts eines Vektors mit sich selbst: Einsetzen in die Definition ergibt direkt:

$$\vec{v} \times \vec{v} = \begin{pmatrix} v_2 v_3 - v_3 v_2 \\ v_3 v_1 - v_1 v_3 \\ v_1 v_2 - v_2 v_1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = \vec{0}$$

3. Antisymmetrie: Entweder erkennt man direkt, dass die Definition beim Vertauschen der Rollen von  $\vec{v}$  und  $\vec{w}$  auch das Vorzeichen im Produkt umkehrt (Ausklammern von  $(-1)$  und Umordnen der einzelnen Faktoren in den sechs Produkten auf  $\mathbb{R}$  in den Komponenten). Oder man verwendet die hier schon gezeigten Eigenschaften 1 und 2:

$$\vec{0} = (\vec{v} + \vec{w}) \times (\vec{v} + \vec{w}) = (\vec{v} \times \vec{v}) + (\vec{v} \times \vec{w}) + (\vec{w} \times \vec{v}) + (\vec{w} \times \vec{w})$$

Nun verschwinden der erste und der vierte Beitrag dieser Summe wegen Eigenschaft 2, aber dann bleibt:

$$\vec{0} = (\vec{v} \times \vec{w}) + (\vec{w} \times \vec{v}) \quad \Leftrightarrow \quad (\vec{v} \times \vec{w}) = -(\vec{w} \times \vec{v})$$

4. Orthogonalität mit den Faktoren: Wir rechnen nach:

$$\begin{aligned} \vec{v} \bullet (\vec{v} \times \vec{w}) &= \vec{v} \bullet \begin{pmatrix} v_2 w_3 - v_3 w_2 \\ v_3 w_1 - v_1 w_3 \\ v_1 w_2 - v_2 w_1 \end{pmatrix} = v_1(v_2 w_3 - v_3 w_2) + v_2(v_3 w_1 - v_1 w_3) + v_3(v_1 w_2 - v_2 w_1) \\ &= v_1 v_2 w_3 - v_1 v_3 w_2 + v_2 v_3 w_1 - v_2 v_1 w_3 + v_3 v_1 w_2 - v_3 v_2 w_1 \\ &= (v_1 v_2 - v_2 v_1) w_3 + (-v_1 v_3 + v_3 v_1) w_2 + (v_2 v_3 - v_3 v_2) w_1 \\ &= 0 + 0 + 0 = 0 \end{aligned}$$

(Dabei haben wir die aus den sechs Termen diejenigen mit  $w_j$  zusammen gefasst.)

Analog rechnet man, dass  $\vec{w} \bullet (\vec{v} \times \vec{w}) = 0$ .

Damit ist alles gezeigt. ■

### **Beweis zu Satz 6.22 (Eigenschaften linear (un-)abhängiger Vektoren) auf Seite 171:**

Wir zeigen die einzelnen Eigenschaften nacheinander:

1. Der Nullvektor ist linear abhängig, da mit  $c \neq 0$  nach Satz 6.3 stets gilt:

$$c\vec{0} = \vec{0}$$

Dies ist eine nichttriviale Linearkombination des Nullvektors.

2. Falls eine Menge  $A \subseteq V$  den Nullvektor enthält, so wählt man  $c_j := 0$  für alle  $\vec{a}_j \neq \vec{0}$  und für den Nullvektor einen Skalierungsfaktor  $c \neq 0$ . Damit hat man wieder eine nichttriviale Linearkombination des Nullvektors.
3. Wenn die Menge  $A \subseteq V$  nur einen einzigen Vektor  $\vec{a}$  enthält, der *nicht* der Nullvektor ist, so sind alle Skalierungen  $c\vec{a}$  mit  $c \neq 0$  ungleich dem Nullvektor. Denn nach Definition 6.1 ist  $(V, +)$  eine abelsche Gruppe und damit auch ein kommutatives Monoid. Die neutralen Elemente sind in Monoiden jedoch nach Satz 5.5 eindeutig bestimmt – es kann also nur einen Nullvektor in  $V$  geben. Diesen erhalten wir jedoch, da  $\mathbb{K}$  als Körper nullteilerfrei ist, nur per  $0 \cdot \vec{a}$ .
4. Für zwei Vektoren  $\vec{a}_1 \neq \vec{a}_2$ , die beide nicht  $\vec{0}$  sind, stellen wir die Linearkombination des Nullvektors um:

$$c_1 \vec{a}_1 + c_2 \vec{a}_2 = \vec{0} \quad \Leftrightarrow \quad c_2 \vec{a}_2 = -c_1 \vec{a}_1$$

In der Implikation von links nach rechts (" $\Rightarrow$ ") nehmen wir an, dass  $\vec{a}_1, \vec{a}_2$  linear abhängig sind.

Falls nun  $c_2 = 0$  gilt, so stünde in der umgeformten Gleichung links der Nullvektor, woraus, da  $\vec{a}_1 \neq \vec{0}$ , direkt  $c_1 = 0$  folgen würde – das ist die triviale Linearkombination des Nullvektors, die stets möglich ist – und diesen Fall hatten wir ausgeschlossen, da  $\vec{a}_1, \vec{a}_2$  als linear abhängig angenommen sind.

Also muss  $c_2 \neq 0$  gelten. Dann können wir die Gleichung mit  $(1/c_2)$  skalieren und erhalten:

$$\vec{a}_2 = -\frac{c_1}{c_2}\vec{a}_1 \quad (*)$$

Da  $\vec{a}_2 \neq \vec{0}$ , muss somit auch  $c_1 \neq 0$  gelten. Sind nun  $\vec{a}_1, \vec{a}_2$  linear abhängig, so ist eben diese Gleichung (\*) mit  $c_1, c_2 \in \mathbb{K} \setminus \{0\}$  erfüllbar; dann lässt sich  $\vec{0}$  nichttrivial linear kombinieren. Aber mit

$$c := -\frac{c_1}{c_2}$$

ist dann auch der Faktor gefunden, sodass gilt  $\vec{a}_2 = c\vec{a}_1$ .

Für die andere Implikationsrichtung (" $\Leftarrow$ ") gehen wir davon aus, dass  $\vec{a}_2 = c\vec{a}_1$  mit einem  $c \neq 0$ . Dann ist aber auch folgende Umformung richtig:

$$\vec{a}_2 = c\vec{a}_1 \quad \Leftrightarrow \quad c\vec{a}_1 - \vec{a}_2 = \vec{0}$$

Damit liegt eine nichttriviale Linearkombination des Nullvektors vor, mit den Koeffizienten  $c_1 := c$  und  $c_2 := (-1)$ ; und somit sind die Vektoren linear abhängig.

5. Wenn eine Menge  $A \subseteq V$  bereits linear abhängig ist, so gibt es eine nichttriviale Linearkombination des Nullvektors aus den  $\vec{a}_j \in A$ . Addiert man nun auf eine solche Linearkombination noch den Beitrag

$$0 \cdot \vec{a} = \vec{0}$$

hinzu, so liegt immer noch eine nichttriviale Linearkombination des Nullvektors vor. Also ist auch

$$A \cup \{\vec{a}\}$$

linear abhängig.

6. Sei  $A \subseteq V$  linear unabhängig.

Nehmen wir nun an, dass  $\vec{a}_k$  der Vektor ist, der aus  $A$  entfernt werden soll. Dann sei

$$\tilde{A} := A \setminus \{\vec{a}\}$$

Wir beweisen indirekt, dass auch  $\tilde{A}$  linear unabhängig ist. Wäre dem nämlich nicht so, dann wäre  $\tilde{A}$  linear abhängig, und es gäbe eine nichttriviale Linearkombination von  $\vec{0}$  mit den Vektoren aus  $\tilde{A}$ . Nach der vorigen Eigenschaft wäre dann aber auch  $\tilde{A} \cup \{\vec{a}\} = A$  linear abhängig, im Widerspruch zur Voraussetzung.  $\nexists$

Also muss auch  $\tilde{A}$  linear unabhängig sein.

Damit ist alles gezeigt.  $\blacksquare$

### **Beweis zu Satz 6.23 (Eindeutigkeit von Linearkombinationen) auf Seite 172:**

Wir zeigen beide Implikationsrichtungen der Äquivalenz:

" $\Rightarrow$ ": Jeder Vektor  $\vec{x} \in \text{span}(A)$  ist also eindeutig aus den Vektoren in  $A$  linear kombinierbar. Dann gilt dies insbesondere auch für  $\vec{x} = \vec{0}$ , denn der Nullvektor ist im Spann von beliebigen Vektoren stets enthalten. Damit ist  $A$  aber linear unabhängig nach Definition 6.21 – denn die triviale Linearkombination für  $\vec{0}$  (welche immer möglich ist) ist dann aufgrund der Eindeutigkeit auch die einzig mögliche.

" $\Leftarrow$ ": Sei  $A$  linear unabhängig. Wir beweisen indirekt, dass dann auch jeder Vektor  $\vec{x} \in \text{span}(A)$  eindeutig linear aus  $A$  kombinierbar sein muss. Wäre dem nämlich nicht so, dann gäbe es irgendein  $\vec{x} \in \text{span}(A)$ , das auf mindestens zwei verschiedene Weisen linear kombinierbar wäre. Es würde also gelten:

$$\vec{x} = \sum_j c_j \vec{a}_j = \sum_j d_j \vec{a}_j,$$

wobei für wenigstens ein  $k$  gilt, dass  $c_k \neq d_k$ .

Da aber  $\vec{x}$  jeder dieser beiden Summen entspricht, gilt auch:

$$\cdots \Rightarrow \quad \vec{0} = \vec{x} - \vec{x} = \left( \sum_j c_j \vec{a}_j \right) - \left( \sum_j d_j \vec{a}_j \right) = \sum_j (c_j - d_j) \vec{a}_j$$

Da aber  $A$  linear unabhängig ist, gibt es nur die triviale Linearkombination des Nullvektors. Dann müssen aber alle Differenzen  $(c_j - d_j)$  genau 0 entsprechen, und das geht nur, wenn für alle  $j$  auch  $c_j = d_j$  gilt. Aber dann sind die beiden Linearkombinationen von  $\vec{x}$  identisch, im Widerspruch zur Annahme.  $\nexists$

Also muss jedes  $\vec{x}$  aus dem Spann von  $A$  eindeutig linear kombinierbar sein.

Damit ist alles gezeigt. ■

### **Beweis zu Satz 6.27 (Lineare Abhängigkeit von Vektoren aufgrund ihrer Anzahl) auf Seite 173:**

Wir beweisen die Behauptung indirekt: Angenommen,  $W$  enthielte wäre linear unabhängig. Dann gäbe es in  $W$  mindestens  $(n + 1)$  linear unabhängige Vektoren. Jedoch würden diese dann auch einen mindestens  $(n + 1)$ -dimensionalen Vektorraum aufspannen, im Widerspruch zur Dimension  $n$  von  $V$  nach Satz 6.26.  $\nexists$

Also muss  $W$  linear abhängig sein. ■

### **Beweis zu Satz 6.33 (Skalarprodukt von Vektoren bezüglich einer ONB) auf Seite 177:**

Wir setzen die (eindeutigen) Entwicklungen der beiden Vektoren ein und nutzen unterwegs, dass  $\vec{a}_j \bullet \vec{a}_k = \delta_{jk}$  für alle Basisvektoren  $\vec{a}_j, \vec{a}_k$  gilt:

$$\vec{x} \bullet \vec{y} = \left( \sum_j c_j \vec{a}_j \right) \bullet \left( \sum_k d_k \vec{a}_k \right) = \sum_{j,k} c_j d_k (\vec{a}_j \bullet \vec{a}_k) = \sum_{j,k} c_j d_k \delta_{jk} = \sum_j c_j d_j \quad \blacksquare$$

## **B.6 Für Kapitel 7**

### **Beweis zu Satz 7.5 (Lineare Abbildung als Produkt) auf Seite 186:**

Der erste Teil des Satzes wurde bereits oben motiviert: Zu jeder linearen Abbildung findet man die Abbildungsmatrix, und die Abbildung entspricht dann einem Produkt aus dieser Matrix  $F$  mit dem Urbildvektor  $\vec{x}$ .

Der zweite Teil des Satzes ist noch zu zeigen: Sei  $F \in \mathbb{R}^{(m,n)}$  und  $\vec{x} \in \mathbb{R}^n$ . Um zu zeigen, dass das Produkt  $F \cdot \vec{x}$  stets eine lineare Abbildung ist, müssen wir mit Definition 7.1 nachrechnen, dass für beliebige  $a, b \in \mathbb{R}$  und  $\vec{x}, \vec{y} \in \mathbb{R}^n$  gilt:

$$F \cdot (a\vec{x} + b\vec{y}) = a(F\vec{x}) + b(F\vec{y})$$

Wir setzen die Definition 7.4 ein und erhalten für die  $j$ -te Komponente ( $j \in \{1, \dots, m\}$ ):

$$\begin{aligned} (F \cdot (a\vec{x} + b\vec{y}))_j &= \sum_{k=1}^n F_{j,k} \cdot (a\vec{x} + b\vec{y})_k = \sum_{k=1}^n F_{j,k} \cdot (ax_k + by_k) \\ &= \left( \sum_{k=1}^n F_{j,k} \cdot ax_k \right) + \left( \sum_{k=1}^n F_{j,k} \cdot by_k \right) = a \left( \sum_{k=1}^n F_{j,k} \cdot x_k \right) + b \left( \sum_{k=1}^n F_{j,k} \cdot y_k \right) \\ &= a(F \cdot \vec{x})_j + b(F \cdot \vec{y})_j \\ &= (a(F \cdot \vec{x}) + b(F \cdot \vec{y}))_j \end{aligned}$$

(Hierbei haben wir unter anderem mehrfach das Distributivgesetz der reellen Zahlen verwendet.)

Da die Formel für alle Komponenten  $j \in \{1, \dots, m\}$  richtig ist, gilt sie auch in der oben notierten vektoriellen Form; damit ist alles gezeigt. ■

**Beweis zu Satz 7.6 (Matrizenräume) auf Seite 188:**

Es gelten die gleichen Argumente wie im Beweis zu Satz 6.4 (Kartesischer Produktraum). ■

---

**Beweis zu Satz 7.7 (Komposition linearer Abbildungen) auf Seite 189:**

Wir nutzen nacheinander die Linearität von  $\vec{f}$  und  $\vec{g}$  aus, um daraus die Linearität von  $\vec{h}$  herzuleiten. Für  $a, b \in \mathbb{R}$  und  $\vec{x}, \vec{y} \in \mathbb{R}^n$  gilt:

$$\begin{aligned}
 \vec{h}(a\vec{x} + b\vec{y}) &= (\vec{g} \circ \vec{f})(a\vec{x} + b\vec{y}) \\
 &= \vec{g}(\vec{f}(a\vec{x} + b\vec{y})) \\
 &= \vec{g}(a\vec{f}(\vec{x}) + b\vec{f}(\vec{y})) \\
 &= a\vec{g}(\vec{f}(\vec{x})) + b\vec{g}(\vec{f}(\vec{y})) \\
 &= a(\vec{g} \circ \vec{f})(\vec{x}) + b(\vec{g} \circ \vec{f})(\vec{y}) \\
 &= a\vec{h}(\vec{x}) + b\vec{h}(\vec{y}) \quad \blacksquare
 \end{aligned}$$


---

**Beweis zu Satz 7.14 (Eigenschaften transponierter Matrizen) auf Seite 195:**

Wir zeigen die Eigenschaften der Reihe nach, unter Benutzung von Definition 7.13:

1. Sei  $B := A^T$ . Dann gilt für alle Komponenten von  $B$ :

$$B_{j,k} = (A^T)_{j,k} = A_{k,j}$$

Dann ist aber auch  $(A^T)^T = B^T$ , und damit:

$$((A^T)^T)_{j,k} = (B^T)_{j,k} = B_{k,j} = A_{j,k}$$

Also gilt, da alle korrespondierenden Komponenten überein stimmen:

$$(A^T)^T = A$$

2. Falls  $A, B \in \mathbb{R}^{(m,n)}$ , so gilt dies auch für ihre Summe  $C := A + B$ , und auch für die skalierte Matrix  $cA$ . Die Transponierten von  $C$  und  $cA$  ermitteln wir mit der Definition komponentenweise:

$$\begin{aligned}
 (C^T)_{j,k} &= C_{k,j} = (A + B)_{k,j} = A_{k,j} + B_{k,j} = (A^T)_{j,k} + (B^T)_{j,k} \\
 ((cA)^T)_{j,k} &= (cA)_{k,j} = c \cdot A_{k,j} = c(A^T)_{j,k}
 \end{aligned}$$

3. Sei  $A \in \mathbb{R}^{(p,m)}$  und  $B \in \mathbb{R}^{(m,n)}$ . Dann ist das Produkt  $C := AB$  wohldefiniert und aus  $\mathbb{R}^{(k,n)}$ . Für dessen Transponierte gilt:

$$\begin{aligned}
 (C^T)_{j,k} &= C_{k,j} = (AB)_{k,j} = \sum_{r=1}^m A_{k,r} B_{r,j} = \sum_{r=1}^m B_{r,j} A_{k,r} = \sum_{r=1}^m (B^T)_{j,r} (A^T)_{r,k} \\
 &= (B^T \cdot A^T)_{j,k}
 \end{aligned}$$

4. Falls  $A \in \mathbb{R}^{(m,n)}$ , so ist  $A^T \in \mathbb{R}^{(n,m)}$ . Dann sind beide Produkte dieser Matrizen wohldefiniert, da jeweils die Spaltenzahl der einen Matrix der Zeilenzahl der anderen entspricht.

Dabei ist  $A \cdot A^T$  aus  $\mathbb{R}^{(m,m)}$  sowie  $A^T \cdot A$  aus  $\mathbb{R}^{(n,n)}$ , wie sich aus Definition 7.8 (Matrixprodukt) ergibt. Dies sind beides quadratische Matrizen, wie behauptet.

Damit ist alles gezeigt. ■

---



**Beweis zu Satz 7.18 (Eigenschaften inverser Matrizen) auf Seite 199:**

Für die vier Eigenschaften gilt:

1. Wegen

$$A \cdot A^{-1} = A^{-1} \cdot A = \mathbb{1}_n$$

sind  $A$  und  $A^{-1}$  zueinander invers. Die Inverse zu  $A^{-1}$  ist somit  $A$ , wie behauptet.

2. Wir rechnen nach:

$$\frac{1}{c} \cdot A^{-1} \cdot (cA) = \frac{1}{c} \cdot c \cdot (A^{-1} \cdot A) = 1 \cdot \mathbb{1}_n = \mathbb{1}_n$$

Da die Inverse eindeutig bestimmt ist, ist sie hiermit gefunden, und die Behauptung ist richtig.

3. Wir transponieren die Gleichung

$$A \cdot A^{-1} = A^{-1} \cdot A = \mathbb{1}_n$$

Da  $\mathbb{1}_n$  symmetrisch ist, gilt dann:

$$(A^{-1})^T \cdot A^T = A^T \cdot (A^{-1})^T = \mathbb{1}_n^T = \mathbb{1}_n$$

4. Wegen der vorigen Aussage und wegen der Symmetrie von  $A$  gilt folgende Gleichungskette:

$$(A^{-1})^T = (A^T)^{-1} = A^{-1}$$

Also ist mit  $A$  auch immer die Inverse symmetrisch.

5. Die fünfte Behauptung ergibt sich aus Satz 5.8 (Inverse einer Verknüpfung), denn die invertierbaren Matrizen bilden mit  $\mathbb{1}_n$  als neutralem Element eine (nicht-kommutative) Gruppe; die Voraussetzungen des Monoids sind mit Satz 7.12 über die Matrizenringe bereits gegeben.

Damit ist alles gezeigt. ■

**Beweis zu Satz 7.20 (Bijektive lineare Abbildungen) auf Seite 201:**

Zunächst erinnern wir uns, dass die Basis eines Vektorraums  $V$  den gesamten Raum  $V$  aufspannt. Die Basisvektoren sind außerdem linear unabhängig. Alle endlichdimensionalen Vektorräume haben Basen mit jeweils gleich vielen Vektoren; für  $\mathbb{R}^n$  sind dies  $n$  Vektoren.

Wir betrachten nun eine beliebige Abbildung  $\vec{f} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ ; diese ist durch eine eindeutig bestimmte Matrix  $F \in \mathbb{R}^{(m,n)}$  repräsentiert. Zunächst zeigen wir, dass  $F$  quadratisch sein muss, damit  $\vec{f}$  bijektiv sein kann. Danach ist es leicht, zu zeigen, dass  $F$  auch invertierbar sein muss.

- Sei also zunächst  $m > n$ . Wir wählen eine beliebige Basis  $\{\vec{v}_1, \dots, \vec{v}_n\}$  von  $\mathbb{R}^n$ ; dann ist jeder Vektor  $\vec{x} \in \mathbb{R}^n$  eindeutig nach dieser Basis entwickelbar:

$$\vec{x} = c_1 \vec{v}_1 + \dots + c_n \vec{v}_n$$

Nun betrachten wir das Bild von  $\vec{x}$  und verwenden, dass  $\vec{f}$  eine lineare Abbildung ist:

$$F\vec{x} = F \cdot \sum_{j=1}^n c_j \vec{v}_j = \sum_{j=1}^n c_j (F \cdot \vec{v}_j)$$

Das Bild liegt also in

$$\text{span}(F\vec{v}_1, \dots, F\vec{v}_n)$$

Dies ist aber ein Unterraum von  $\mathbb{R}^m$ , dessen Dimension *höchstens*  $n$  betragen kann (dies, falls die Bilder der Basisvektoren linear unabhängig sind) – jedoch sicher nicht  $m$ . Also spannen die Bilder der Basisvektoren von  $\mathbb{R}^n$  nicht den Raum  $\mathbb{R}^m$  auf und sind somit keine Basis davon. Aber dann gibt es Vektoren  $\vec{y} \in \mathbb{R}^m$ , die *nicht* im obigen Spann liegen – diese haben jedoch kein Urbild in  $\mathbb{R}^n$  unter der Abbildung  $\vec{f}$ . Somit kann  $\vec{f}$  *nicht surjektiv* sein und ist also auch nicht bijektiv.

- Sei nun  $m < n$ . Auch hier wählen wir eine Basis aus  $n$  Vektoren von  $\mathbb{R}^n$  und betrachten die Bilder  $F\vec{v}_j$ . Dies sind  $n$  Vektoren aus  $\mathbb{R}^m$ , und nach Satz 6.27 sind diese Bildvektoren *linear abhängig* (denn es kann in  $\mathbb{R}^m$  nur linear unabhängige Mengen mit höchstens  $m$  Vektoren geben).

Dann finden wir allerdings reelle Koeffizienten  $c_1, \dots, c_n$ , die nicht alle 0 sind, sodass

$$\sum_{j=1}^n c_j (F \cdot \vec{v}_j) = \vec{0}_{(m)}$$

Es gibt also eine nichttriviale Linearkombination des Nullvektors von  $\mathbb{R}^m$  – diese entspricht jedoch dem Bild eines Vektors

$$\vec{x} = \sum_{j=1}^n c_j \vec{v}_j$$

aus  $\mathbb{R}^n$ , der, da die  $\vec{v}_j$  als Basisvektoren linear unabhängig sind und da die  $c_j$  nicht alle 0 entsprechen, *nicht* der Nullvektor  $\vec{0}_{(n)}$  sein kann.

Zusätzlich gilt aber (trivialerweise) auch noch:

$$F \cdot \vec{0}_{(n)} = \vec{0}_{(m)}$$

Dann sind aber  $\vec{x}$  und  $\vec{0}_{(n)}$  Urbilder des Nullvektors  $\vec{0}_{(m)}$ , und die Abbildung  $\vec{f}$  ist somit *nicht injektiv*. Aber dann kann sie auch nicht bijektiv sein.

Nun ist also klar, dass eine bijektive lineare Abbildung  $\vec{f}$  nur zwischen  $\mathbb{R}^n$  und  $\mathbb{R}^n$  existieren kann und demnach eine quadratische Abbildungsmatrix  $F \in \mathbb{R}^{(n,n)}$  besitzen muss.

Aber dann gilt, falls  $\vec{y} = F\vec{x}$  das Bild von  $\vec{x}$  unter der Abbildung  $\vec{f}$  ist, dass es auch eine Umkehrabbildung  $\vec{g}: \mathbb{R}^n \rightarrow \mathbb{R}^n$  mit Abbildungsmatrix  $G \in \mathbb{R}^{(n,n)}$  gibt, sodass  $\vec{x} = G\vec{y}$ .

Dann können wir folgende vektorielle Gleichung in  $\mathbb{R}^n$  anschreiben:

$$\vec{x} = G\vec{y} = G(F\vec{x}) = (GF)\vec{x}$$

Damit folgt aber direkt, dass  $GF = \mathbb{1}_n$  gilt, und dass, da  $F, G$  quadratisch sind, wegen Satz 7.19 direkt, dass

$$G = F^{-1}$$

Also muss die bijektive Abbildung eine invertierbare quadratische Abbildungsmatrix besitzen. ■

### **Beweis zu Satz 7.23 (Eigenschaften orthogonaler Matrizen) auf Seite 203:**

Wir beweisen die Eigenschaften nacheinander:

1. Nach Satz 7.14 ist  $(A^T)^T = A$ . Also gilt für eine orthogonale Matrix  $A \in \mathbb{R}^{(n,n)}$  auch

$$(A^T)^T \cdot A^T = A \cdot A^T = \mathbb{1}_n$$

2. Mit den Voraussetzungen  $A^T \cdot A = B^T \cdot B = \mathbb{1}_n$  gilt, die Transponierte des Produkts eingesetzt:

$$(AB)^T \cdot (AB) = (B^T \cdot A^T) \cdot (A \cdot B) = B^T \cdot (A^T \cdot A) \cdot B = B^T \cdot \mathbb{1}_n \cdot B = B^T \cdot B = \mathbb{1}_n$$

Für das Produkt  $BA$  rechnet man ganz analog.

Damit ist alles gezeigt. ■

**Beweis zu Satz 7.26 (Wirkung von Permutationsmatrizen) auf Seite 207:**

Wir nutzen hier die Erkenntnisse aus Unterabschnitt 7.4.3 über die Komponenten einer Produktmatrix. Hierfür benötigen wir die Zeilen des linken Faktors und die Spalten des rechten.

- $A \in \mathbb{R}^{(m,n)}$  besitzt  $n$  Spalten und  $m$  Zeilen; wir wählen folgende Bezeichner:

$$A = (\vec{a}_1 \quad \cdots \quad \vec{a}_n) = \begin{pmatrix} \vec{v}_1^T \\ \vdots \\ \vec{v}_m^T \end{pmatrix}$$

(Die Vektoren  $\vec{v}_j$  sind die Spaltenvektoren von  $A^T$ )

Nun entspricht die  $k$ -te Spalte von  $P_\sigma$  genau dem  $\sigma(k)$ -ten kartesischen Einheitsvektor aus der Standardbasis  $E_n$ ; also ist

$$(A \cdot P_\sigma)_{j,k} = \vec{v}_j \bullet \vec{e}_{\sigma(k)}$$

Wir erinnern uns, dass in Orthonormalbasen die Komponenten eines Vektors gerade durch Skalarprodukt mit den jeweiligen Basisvektoren ermittelbar sind – also liegt hier folgendes vor:

$$\cdots = (\vec{v}_j)_{\sigma(k)}$$

Bei  $\vec{v}_j$  handelt es sich um die  $j$ -te Spalte von  $A^T$ , daher:

$$\cdots = (A^T)_{\sigma(k),j} = A_{j,\sigma(k)}$$

- $B \in \mathbb{R}^{(n,p)}$  besitzt  $n$  Zeilen und  $p$  Spalten. Da wir  $P_\sigma^T \cdot B$  berechnen wollen, kommt uns hier zugute, dass die Zeilen von  $P_\sigma^T$  gerade den Spalten von  $P_\sigma$  entsprechen. Analog zu oben gilt für

$$B = \begin{pmatrix} \vec{b}_1 & \cdots & \vec{b}_p \end{pmatrix}$$

dann folgende Gleichungskette:

$$(P_\sigma^T \cdot B)_{j,k} = \vec{e}_{\sigma(j)} \bullet \vec{b}_k = \left( \vec{b}_k \right)_{\sigma(j)} = B_{\sigma(j),k}$$

- Wir kombinieren die beiden obigen Formeln und erhalten:

$$(P_\sigma^T \cdot C \cdot P_\sigma)_{j,k} = (P_\sigma^T \cdot (C \cdot P_\sigma))_{j,k} = (C \cdot P_\sigma)_{\sigma(j),k} = C_{\sigma(j),\sigma(k)}$$

Damit ist alles gezeigt. ■

**Beweis zu Satz 7.27 (Komposition von Permutationsmatrizen) auf Seite 208:**

Wir verwenden, wie im Beweis zu Satz 7.26, wiederum Unterabschnitt 7.4.3.

Für die Berechnung der Komponenten von  $P_\pi \cdot P_\sigma$  wäre es praktisch, wenn wir den linken Faktor als transponierte Matrix schreiben könnten – deren Zeilen entsprechen nämlich den Spalten der un-transponierten Matrix. Tatsächlich geht das, da wir oben schon fest gestellt hatten, dass

$$P_\pi = \left( (P_\pi)^{-1} \right)^T = (P_{\pi^{-1}})^T$$

Also erhalten wir, wenn wir dies und die Spaltenvektoren der beteiligten Permutationsmatrizen verwenden:

$$(P_\pi \cdot P_\sigma)_{j,k} = \vec{e}_{\pi^{-1}(j)} \bullet \vec{e}_{\sigma(k)} = \delta_{\pi^{-1}(j),\sigma(k)}$$

Für ein bestimmtes  $k$  (also die Spalte  $k$  in der Produktmatrix) Gibt es also nur einen Eintrag 1, falls obiges Kronecker-Delta den Wert 1 annimmt. Genau in diesem Fall gilt:

$$\pi^{-1}(j) = \sigma(k)$$

Diese zunächst unhandlich wirkende Gleichung können wir vereinfachen, indem wir auf beide Seiten die Permutation  $\pi$  anwenden – das ist wegen der Bijektivität der Permutationen stets wohldefiniert:

$$\cdots \Leftrightarrow j = \pi(\sigma(k)) = (\pi \circ \sigma)(k)$$

Nun war aber  $j$  gerade der Zeilenindex im Matrixprodukt. Das heißt, dass der Eintrag 1 in Spalte  $k$  der Produktmatrix eben in genau dieser Zeile  $j$  steht. Der korrespondierende  $k$ -te Spaltenvektor des Matrixprodukts ist dann:

$$\vec{e}_{(\pi \circ \sigma)(k)}$$

Das gilt für alle Spalten des Matrixproduktes, und da  $\pi \circ \sigma$  aufgrund der Gruppenstruktur (spezieller hier: der Abgeschlossenheit) von  $S_n$  wiederum eine Permutation ist, liegt abermals eine Permutationsmatrix vor; es ist also

$$P_\pi \cdot P_\sigma = P_{\pi \circ \sigma} \quad \blacksquare$$

## B.7 Für Kapitel 8

### Beweis zu Satz 7.19 (Links- und Rechtsinversität) auf Seite 200:

Wir beweisen die Behauptung durch Betrachtung der Injektivität und Surjektivität, angelehnt an den Beweis zu Satz 7.20 (S. 201).

Sei zunächst vorausgesetzt, dass

$$A \cdot B = \mathbb{1}_n$$

Also muss  $A \in \mathbb{R}^{(n,m)}$  und  $B \in \mathbb{R}^{(m,n)}$  gelten (In der Voraussetzung des Satzes war zwar schon  $m = n$  gefordert, aber es lohnt sich, zunächst auch nichtquadratische Matrizen zu betrachten).

Wenn wir die drei obigen Matrizen mit linearen Abbildungen identifizieren, so gilt entsprechend für alle  $\vec{x} \in \mathbb{R}^n$ :

$$\vec{a}(\vec{b}(\vec{x})) = \text{id}(\vec{x}) = \vec{x}$$

Da nun  $\vec{a}$  die äußere Abbildung ist (als zweites angewendet), die jedoch jedes  $\vec{x}$  aus  $\mathbb{R}^n$  erreicht, muss  $\vec{a}$  eine *surjektive* Abbildung sein.

Sei nun  $\vec{b}(\vec{x}) = B\vec{x} =: \vec{y}$ . Angenommen, für  $\vec{x}_1 \neq \vec{x}_2$  ergäbe sich  $\vec{y}_1 = \vec{y}_2$ . Dann wäre jedoch auch  $\vec{a}(\vec{y}_1) = \vec{a}(\vec{y}_2)$ , und somit  $\vec{x}_1 = \vec{x}_2$ , im Widerspruch zur Annahme. Offenbar muss also für  $\vec{x}_1 \neq \vec{x}_2$  auch  $\vec{y}_1 \neq \vec{y}_2$  gelten, somit muss  $\vec{b}$  eine *injektive* Abbildung sein.

Für die Zwischenwerte  $\vec{y} \in \mathbb{R}^m$  gilt also mit den Argumenten aus dem Beweis von Satz 7.20, dass  $m \geq n$  gelten muss. Der Fall  $m < n$  ist ausgeschlossen.

Wir betrachten nacheinander die Fälle  $m = n$  und  $m > n$ :

- Für  $m = n$  sind die beiden Abbildungen  $\vec{a}, \vec{b}$  von  $\mathbb{R}^n$  nach  $\mathbb{R}^n$ . Die Matrix  $B$  ist hier quadratisch, und wegen der Injektivität von  $\vec{b}$  ist die Gleichung

$$B\vec{x} = \vec{y}$$

niemals mehrdeutig lösbar. Da mehrdeutige Lösbarkeit beim assoziierten LGS stets mit einem nichttrivialen Kern verbunden ist, muss hier gelten:  $\dim \ker B = 0$ . Damit wissen wir aber nach dem Rangsatz (8.9), dass  $B$  vollen Rang hat:  $\text{rg } B = n$ . Dann sind die Spalten von  $B$  linear unabhängig, und  $\text{span}(\vec{b}_1, \dots, \vec{b}_n)$  ist ein  $n$ -dimensionaler Unterraum von  $\mathbb{R}^n$ .

Allerdings sind auch die Zeilen von  $B$  linear unabhängig. Für beliebiges  $\vec{y} \in \mathbb{R}^n$  ist es nicht möglich, beim Lösen des LGS ( $B \mid \vec{y}$ ) im Koeffiziententeil mit den üblichen Zeilen-Operationen eine Nullzeile zu erzeugen. Das LGS kann also nicht unlösbar sein. Folglich gibt es bei der Berechnung des Spans der Spalten von  $B$  auch keine Zwangsbedingung an die Komponenten von  $\vec{y}$ , sodass jedes  $\vec{y} \in \mathbb{R}^n$  aus den Spalten von  $B$  linear kombinierbar ist. Damit ist  $\vec{b}$  nicht nur injektiv, sondern auch eine surjektive Abbildung.

Sei nun  $\vec{y} \in \mathbb{R}^n$  beliebig gewählt und fest, dann existiert aufgrund der Surjektivität von  $\vec{b}$  auch stets ein  $\vec{x} \in \mathbb{R}^n$ , sodass  $\vec{y} = \vec{b}(\vec{x})$ .

Wir betrachten nun die Komposition  $\vec{b} \circ \vec{a}$ , deren Abbildungsmatrix durch  $BA$  gegeben ist. Dann gilt:

$$(\vec{b} \circ \vec{a})(\vec{y}) = (\vec{b} \circ \vec{a})(\vec{b}(\vec{x})) = (\vec{b} \circ \vec{a} \circ \vec{b})(\vec{x}) = (\vec{b} \circ (\vec{a} \circ \vec{b}))(\vec{x})$$

Aber da nach Voraussetzung  $\vec{a} \circ \vec{b} = \text{id}$  gilt, ergibt sich weiter:

$$\dots = (\vec{b} \circ \text{id})(\vec{x}) = \vec{b}(\vec{x}) = \vec{y}$$

Offenbar handelt es sich also bei  $\vec{b} \circ \vec{a}$  um die identische Abbildung, somit ist  $B \cdot A = \mathbb{1}_n$ .

- Für  $m > n$  ist die Abbildung  $\vec{b}$  zwar injektiv, aber nicht mehr surjektiv (siehe die Begründung von Satz 7.20). Dann gibt es in  $\mathbb{R}^m$  Vektoren  $\vec{y}$  ohne Urbild – es handelt sich um solche, die gerade nicht im Spann der Spaltenvektoren von  $B$  liegen. Letzter Spann ist also hier ein echter Unterraum von  $\mathbb{R}^m$ .

Weiterhin kann die Abbildung  $\vec{a}$  nicht injektiv sein. Zu mindestens einem Vektor  $\vec{x} \in \mathbb{R}^n$  existieren also mindestens zwei verschiedene Urbilder in  $\mathbb{R}^m$ ; höchstens eines davon (wir wollen es  $\vec{y}_1$  nennen) kann im Spann der Spalten von  $B$  liegen, mindestens eines muss außerhalb liegen. Sei  $\vec{y}_2$  ein solches Urbild. Es ist also möglich von zwei verschiedenen  $\vec{y}_1 \neq \vec{y}_2$  per  $\vec{a}$  zum gleichen  $\vec{x}$  zu gelangen. Von dort aus gelangen wir mit  $\vec{b}$  jedoch nur noch zu  $\vec{y}_1$ .

Immerhin ist  $\text{rg } A = \text{rg } B = n$ ,  $A$  ist linksinvers zu  $B$  und  $B$  ist rechtsinvers zu  $A$ . Damit  $A$  auch rechtsinvers bzw.  $B$  auch linksinvers sein kann, müsste die Gleichung

$$B \cdot A = \mathbb{1}_m$$

erfüllbar sein. Das ist jedoch nicht möglich, da die Matrizenränge hierfür jeweils  $m$  betragen müssten. Außerdem müsste nach obiger Argumentation dann  $\vec{a}$  injektiv sowie  $\vec{b}$  surjektiv sein; keines davon trifft zu.

$A$  hat also gar kein Linksinverses, und  $B$  kein Rechtsinverses, und die Gleichung  $BA = \mathbb{1}_m$  folgt gerade *nicht*.

**Beispiel:** Wir betrachten zwei Matrizen mit  $n = 2$  und  $m = 3$ :

$$B := \begin{pmatrix} -2 & 1 \\ 1 & -2 \\ 2 & 2 \end{pmatrix} \quad \text{und} \quad A := \frac{1}{9} \begin{pmatrix} -2 & 1 & 2 \\ 1 & -2 & 2 \end{pmatrix}$$

Tatsächlich gilt:

$$A \cdot B = \frac{1}{9} \begin{pmatrix} 9 & 0 \\ 0 & 9 \end{pmatrix} = \mathbb{1}_2$$

Der Spann der Spalten von  $B$  ist eine Ebene in  $\mathbb{R}^3$ ; deren Normalenvektor lässt sich hier durch ein Kreuzprodukt berechnen und ist kollinear zu

$$\begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}$$

Für den Spann folgt also:

$$\text{span}(\vec{b}_1, \vec{b}_2) = \left\{ \vec{y} \in \mathbb{R}^3 \mid \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix} \bullet \vec{y} = 0 \right\}$$

Wir wählen nun einen Vektor  $\vec{y}_2$ , der nicht in diesem Spann enthalten ist, z.B.

$$\vec{y}_2 := \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

Dann berechnen wir dessen Bild unter  $\vec{a}$ :

$$\vec{a}(\vec{y}_2) = A\vec{y}_2 = \frac{1}{9} \begin{pmatrix} -2 & 1 & 2 \\ 1 & -2 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \frac{1}{9} \begin{pmatrix} 1 \\ 1 \end{pmatrix} =: \vec{x} \in \mathbb{R}^2$$

Bilden wir diesen Vektor jedoch mit  $\vec{b}$  ab, so erhalten wir:

$$\vec{y}_1 := \vec{b}(\vec{x}) = B\vec{x} = \frac{1}{9} \begin{pmatrix} -2 & 1 \\ 1 & -2 \\ 2 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{1}{9} \begin{pmatrix} -1 \\ -1 \\ 4 \end{pmatrix}$$

Dieser Vektor  $\vec{y}_1$  liegt im Spann der Spalten von  $B$ , und es ist weiter:

$$\vec{a}(\vec{y}_1) = \frac{1}{81} \begin{pmatrix} -2 & 1 & 2 \\ 1 & -2 & 2 \end{pmatrix} \begin{pmatrix} -1 \\ -1 \\ 4 \end{pmatrix} = \frac{1}{81} \begin{pmatrix} 9 \\ 9 \end{pmatrix} = \frac{1}{9} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \vec{x}$$

Somit ist klar, dass  $\vec{a}$  nicht injektiv sein kann (denn wir erreichen das Bild  $\vec{x}$  von zwei verschiedenen Urbildern aus); auch ist  $\vec{b}$  nicht surjektiv.

Wir bestimmen noch  $B \cdot A$ :

$$B \cdot A = \frac{1}{9} \begin{pmatrix} -2 & 1 \\ 1 & -2 \\ 2 & 2 \end{pmatrix} \begin{pmatrix} -2 & 1 & 2 \\ 1 & -2 & 2 \end{pmatrix} = \frac{1}{9} \begin{pmatrix} 5 & -4 & -2 \\ -4 & 5 & -2 \\ -2 & -2 & 8 \end{pmatrix}$$

Diese Matrix ist nicht nur nicht die Einheitsmatrix  $\mathbb{1}_3$ , sondern sie ist obendrein singulär, d.h. ihre Determinante verschwindet, da  $\text{rg}(BA) = 2 < 3$  – wovon man sich durch Nachrechnen (Zeilen- bzw. Spaltenoperationen für die Determinantenberechnung) schnell überzeugen kann.

Zusammen gefasst: Eine Matrix kann nur dann Rechts- *und* Links-Inverse besitzen, wenn sie quadratisch ist. In diesem Fall folgt aus  $AB = \mathbb{1}_n$  aber auch  $BA = \mathbb{1}_n$ . Da wir die Namen  $A, B$  nicht weiter eingeschränkt hatten, sind die beiden Gleichungen dann äquivalent. Wenn wir also (siehe Kapitel 9) für eine gegebene quadratische Matrix  $A$  die Gleichung

$$A \cdot X = \mathbb{1}_n$$

mit Gauß-Jordan nach  $X$  auflösen, bestimmen wir rechnerisch zwar nur das Rechtsinverse von  $A$  – wegen der Äquivalenz ist dies jedoch gleichzeitig auch das Linksinverse von  $A$ , somit ist  $X$  dann auch das Inverse von  $A$ . Damit ist alles gezeigt. ■

### **Beweis zu Satz 8.12 (Anzahl Transpositionen einer zerlegten Permutation) auf Seite 226:**

Wir betrachten eine beliebige Permutation  $\sigma \in S_n$ , die  $r$  kanonische Zyklen besitzt. Eine Zerlegung in  $t$  Transpositionen können wir dann schreiben als

$$\sigma = \tau_t \circ \tau_{t-1} \circ \cdots \circ \tau_1 \circ \text{id}_n$$

Die Indices an den Transpositionen dienen hier zum Nummerieren.

Es stellt sich heraus, dass sich die Gesamtzahl der kanonischen Zyklen jeweils in einheitlicher Weise ändert, wenn wir die Transpositionen nach und nach wirken lassen (beginnend mit  $\text{id}_n$ , der Permutation mit  $n$  Zyklen). Diese wichtige (und für sich genommen schon interessante) Tatsache formulieren wir als Lemma (Hilfssatz):

**Lemma:** Gegeben eine beliebige Permutation  $\pi \in S_n$  mit  $s$  kanonischen Zyklen, sowie eine beliebige Transposition  $\tau \in S_n$ . Dann hat die Permutation

$$\tau \circ \pi$$

entweder  $(s+1)$  oder  $(s-1)$  kanonische Zyklen, je nachdem, ob die beiden Elemente in  $\tau$  Teil desselben kanonischen Zyklus von  $\pi$  sind (erster Fall) oder zu zwei verschiedenen kanonischen Zyklen von  $\pi$  gehören (zweiter Fall).

Bevor wir das Lemma beweisen, zunächst ein

**Beispiel:** Sei  $\pi \in S_9$  gegeben als

$$\pi := \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 4 & 6 & 8 & 5 & 1 & 3 & 9 & 2 & 7 \end{pmatrix} = (1 \ 4 \ 5)(2 \ 6 \ 3 \ 8)(7 \ 9)$$

mit den kanonischen Zyklen wie in Abbildung B.1 gezeigt.

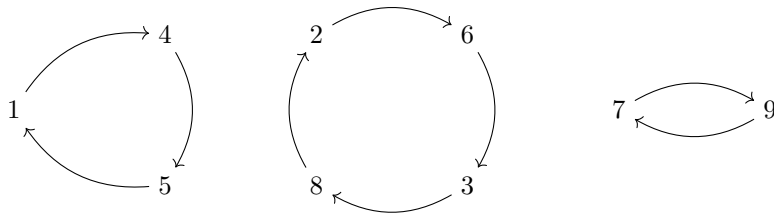


Abbildung B.1: Graph der Beispielpermutation  $\pi \in S_9$

Nun betrachten wir die Transposition  $\tau_{2,3}$ , deren Elemente sich beide auf dem Viererzyklus von  $\pi$  befinden. Wir rechnen die Komposition in Tabellenschreibweise und ermitteln direkt danach wieder die kanonischen Zyklen:

$$\tau_{2,3} \circ \pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 4 & 6 & 8 & 5 & 1 & 2 & 9 & 3 & 7 \end{pmatrix} = (1 \ 4 \ 5)(2 \ 6)(3 \ 8)(7 \ 9)$$

(Graph in Abbildung B.2.)

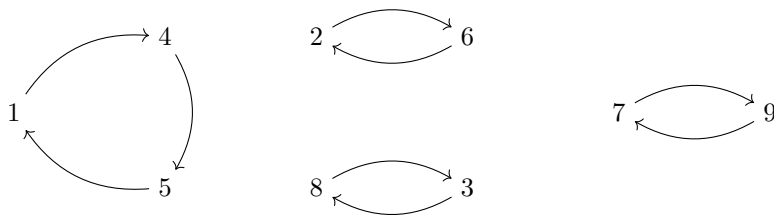


Abbildung B.2: Graph von  $\tau_{2,3} \circ \pi$

Offensichtlich wurde der Viererzyklus durch die Transposition aufgebrochen, und es liegen nun nicht mehr drei, sondern vier kanonische Zyklen vor.

Außerdem betrachten wir die Wirkung der Transposition  $\tau_{2,5}$ , deren Elemente zu zwei verschiedenen Zyklen von  $\pi$  gehören: Die 2 gehört zum Viererzyklus, die 5 zum Dreierzyklus. Selbes Vorgehen wie eben ergibt:

$$\pi := \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 4 & 6 & 8 & 2 & 1 & 3 & 9 & 5 & 7 \end{pmatrix} = (1 \ 4 \ 2 \ 6 \ 3 \ 8 \ 5)(7 \ 9)$$

(Graph in Abbildung B.3.)

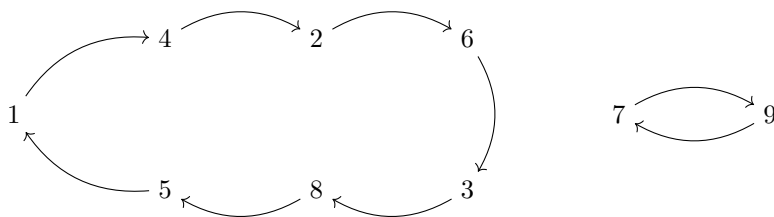


Abbildung B.3: Graph von  $\tau_{2,5} \circ \pi$

Dies hat offenbar die beiden betroffenen Zyklen zu einem Siebenerzyklus vereinigt; nun liegen nicht mehr drei, sondern zwei kanonische Zyklen vor.

(Zur Übung verifiziere man gerne noch, dass sich das gleiche Verhalten auch dann einstellt, wenn beteiligte kanonische Zyklen trivial sind, also nur aus einem Element bestehen.)

Nun begründen wir, warum die Aussage des Lemmas allgemein richtig sein muss. Wir betrachten dazu die beiden genannten Fälle (implizit ist übrigens klar, dass wir nur  $n \geq 2$  betrachten müssen, da sonst gar keine Transposition  $\tau$  definierbar wäre).

Da  $\tau := (a \ b)$  nachträglich auf  $\pi$  wirkt, werden die Bilder  $a$  und  $b$  in  $\pi$  nachträglich vertauscht; also ist

$$(\tau \circ \pi)(j) = \begin{cases} b & , \text{ falls } \pi(j) = a \\ a & , \text{ falls } \pi(j) = b \\ \pi(j) & \text{sonst} \end{cases}$$

Stellen wir uns dies grafisch vor, so müssten wir im Graph von  $\pi$  dann die beiden Pfeile abändern, die zu  $a$  und zu  $b$  hin weisen. Es ändert sich also das Abbildungsverhalten der Urbilder von  $a, b$ :

$$(\tau_{a,b} \circ \pi)(\pi^{-1}(a)) = (\tau_{a,b} \circ (\pi \circ \pi^{-1}))(a) = (\tau_{a,b} \circ \text{id}_n)(a) = \tau_{a,b}(a) = b$$

Und analog:

$$(\tau_{a,b} \circ \pi)(\pi^{-1}(b)) = a$$

Also bildet das Urbild von  $a$  nun nicht mehr auf  $a$ , sondern auf  $b$  ab, und das Urbild von  $b$  bildet nun auf  $a$  ab. Die beiden *Bilder* von  $a, b$  sind dagegen nicht betroffen (außer falls es sich gleichzeitig um deren Urbilder handeln sollte) – etwas, das zunächst vielleicht unintuitiv erscheint.

Wir unterscheiden nun die beiden im Lemma angegebenen Fälle:

- Falls die Elemente von  $\tau_{a,b}$  auf dem selben kanonischen Zyklus von  $\pi$  liegen, hat der Zyklus die Struktur

$$(a \quad \underbrace{\quad \cdots \quad}_{(1)} \quad b \quad \underbrace{\quad \cdots \quad}_{(2)})$$

Nun wirkt  $\tau_{a,b}$  – danach führt der Pfad “(1)” nicht länger zu  $b$ , sondern zu  $a$ . Aber damit haben wir einen Zyklus

$$(a \quad \underbrace{\quad \cdots \quad}_{(1)})$$

erhalten, der  $b$  nicht länger enthält (da Elemente in Zyklen nicht doppelt auftauchen können; siehe Definition 4.12).

Außerdem führt der Pfad “(2)” nun nicht länger zu  $a$ , sondern zu  $b$ . Dies ergibt den Zyklus

$$(b \quad \underbrace{\quad \cdots \quad}_{(2)}),$$

der wiederum  $a$  nicht mehr enthält.

Also haben wir zwei disjunkte Zyklen enthalten (denn wären sie dies nicht, so wären  $a, b$  weiterhin auf dem gleichen Zyklus, und das hatten wir bereits ausgeschlossen) – der ursprüngliche Zyklus wurde aufgespalten.

**Bemerkung:** Das gilt auch für den Extremfall, dass einer der erwähnten Pfade (oder beide) leer sein sollte, d.h. falls es im Graph von  $\pi$  eine direkte Kante von  $a$  nach  $b$  (oder umgekehrt) gibt. Falls “(1)” leer ist, entsteht mit  $\tau_{a,b}$  ein trivialer Zyklus  $(a)$ . Das Urbild von  $b$  (und das ist  $a$ ) wird danach auf  $a$  abgebildet. Falls “(2)” leer ist entsteht nach  $\tau_{a,b}$  entsprechend ein trivialer Zyklus  $(b)$ .

- Falls die Elemente von  $\tau_{a,b}$  auf verschiedenen (disjunkten!) kanonischen Zyklen von  $\pi$  liegen, so geschieht genau das Umgekehrte. Diese beiden Zyklen sind nämlich

$$(a \quad \underbrace{\quad \cdots \quad}_{(1)}) \quad \text{und} \quad (b \quad \underbrace{\quad \cdots \quad}_{(2)})$$

Wenn nun  $\tau_{a,b}$  wirkt, dann führt Pfad “(1)” nicht länger zu  $a$  zurück, sondern zu  $b$ . Und Pfad “(2)” führt nicht zu  $b$  zurück, sondern zu  $a$ . Aber das bedeutet nichts anderes, als das wir aus den beiden ursprünglichen Zyklen nun einen gemeinsamen Zyklus

$$(a \quad \underbrace{\quad \cdots \quad}_{(1)} \quad b \quad \underbrace{\quad \cdots \quad}_{(2)})$$

erhalten haben.

**Bemerkung:** Das gilt auch für den Extremfall, dass einer der erwähnten Pfade (oder beide) leer sein sollte, d.h. falls Einer-Zyklen beteiligt sind. Ein leerer Pfad im Zyklus von  $a$  führt nach  $\tau$  zu einer direkten Kante von  $a$  nach  $b$  (denn  $a$  war sein eigenes Urbild); ein leerer Pfad im Zyklus von  $b$  würde nach  $\tau$  eine direkte Kante von  $b$  nach  $a$  bewirken. Falls beide Pfade leer wären, würde aus den trivialen Zyklen  $(a)(b)$  danach der Zweierzyklus  $(a \quad b) = \tau_{a,b}$ .



Damit ist das Lemma bewiesen. ■

Nun befassen wir uns wieder mit dem eigentlich zu beweisenden Satz. Mit

$$\sigma = \tau_t \circ \tau_{t-1} \circ \cdots \circ \tau_1 \circ \text{id}_n$$

haben wir  $t$ -Mal die Situation des Lemmas vorliegen, da wir die Hintereinanderausführung der Permutationen beliebig klammern dürfen. Wenn wir uns von innen nach außen durch die Komposition bewegen, starten wir zunächst mit der Permutation  $\text{id}_n$ , welche aus  $n$  trivialen und disjunkten Einerzyklen besteht. Von da an wird jede der  $t$  Transpositionen die Zyklenzahl entweder um 1 reduzieren oder um 1 erhöhen.

Ohne genau wissen zu müssen, welche der Transpositionen genau welche der beiden Wirkungen hat, können wir  $k_+$  als die Anzahl der Transpositionen definieren, die die Zyklenzahl (um jeweils 1) erhöhen, und analog  $k_-$  als die Zahl von Transpositionen, die die Zyklenzahl (um jeweils 1) erniedrigen.

Dann gilt offensichtlich:

$$t = k_+ + k_-$$

Nun wissen wir aber auch, dass  $\sigma$  am Ende, d.h. nach Anwendung der  $t$  Transpositionen auf  $\text{id}_n$  genau  $r$  Zyklen besitzt. Da wir mit  $n$  Zyklen starten, ergibt sich damit folgende Bilanzgleichung:

$$r = n + k_+ - k_- \quad \Leftrightarrow \quad n - r = k_- - k_+$$

Nun formulieren wir den obigen Zusammenhang für  $t$  um, indem wir künstlich die Differenz  $(k_- - k_+)$  einführen:

$$t = k_- + k_+ = (k_- - k_+) + 2k_+$$

Dann setzen wir die Gleichung für  $(n - r)$  ein und erhalten:

$$\cdots \Rightarrow t = (n - r) + 2k_+$$

Nehmen wir nun den Rest modulo 2, so verschwindet sicher der Beitrag  $2k_+$ , und es ist

$$t \equiv (n - r) \pmod{2} \quad \blacksquare$$

(Die Ideen für diesen Beweis stammen von

<https://math.stackexchange.com/questions/46403/alternative-proof-that-the-parity-of-permutation-is-well-defined#answer-3685376>

alternative-proof-that-the-parity-of-permutation-is-well-defined#answer-3685376

In der Hauptfrage geht es um das Vorzeichen von Permutationen, für das es verschiedene Definitionen gibt – siehe dazu auch [https://de.wikipedia.org/wiki/Vorzeichen\\_\(Permutation\)](https://de.wikipedia.org/wiki/Vorzeichen_(Permutation)). Die typische Lehrbuch-Definition über Fehlstände ist zwar wohldefiniert, aber begrifflich nicht sehr leicht zu fassen, weswegen wir eine gleichwertige alternative Definition verwenden, bei der eben bewiesene Satz das entscheidende Argument ist.)

## Beweis zu Satz 8.15 (Vorzeichen von Permutationen bei Gruppenoperationen) auf Seite 227:

Wir zeigen beide Eigenschaften:

1. Für die Komposition zerlegen wir beide Permutationen vollständig in Transpositionen; ihre Anzahlen seien  $t_\sigma$  und  $t_\pi$ .

Damit lassen sich beide Kompositionen von  $\sigma$  und  $\pi$  jeweils als Kompositionen dieser Zerlegungen ausdrücken. Es handelt sich in beiden Fällen um insgesamt genau  $(t_\sigma + t_\pi)$  Transpositionen.

Damit ist nach Satz 8.14 das Vorzeichen der Komposition jeweils gegeben durch

$$(-1)^{t_\sigma + t_\pi} = (-1)^{t_\sigma} \cdot (-1)^{t_\pi} = \text{sign}(\sigma) \cdot \text{sign}(\pi)$$

2. Die inverse Permutation  $\sigma^{-1}$  besitzt eben so viele kanonische Zyklen wie  $\sigma$  (der Funktionsgraph ist identisch, nur dass sämtliche Abbildungspfeile die umgekehrte Richtung haben). Damit müssen nach Satz 8.14 auch die Vorzeichen von  $\sigma$  und  $\sigma^{-1}$  identisch sein.

Alternativ kann man  $\sigma$  wieder in  $t$  Transpositionen zerlegen. Die Inverse besteht aus genau diesen Transpositionen, in gespiegelter Reihenfolge. Da aber deren Anzahl ebenfalls gleich  $t$  ist bleibt auch das Signum beim Invertieren gleich.

Damit ist alles gezeigt. ■

**Beweis zu Satz 8.17 (Rechenregeln für Determinanten) auf Seite 229:**

Wir zeigen die Rechenregeln mit den Axiomen von Weierstraß aus Definition 8.16:

1. Skalierung von Spalten: Diese Eigenschaft folgt direkt aus der Multilinearität: Wir verwenden

$$\det(\cdots (r\vec{a}_j + s\vec{a}_k) \cdots) = r \cdot \det(\cdots \vec{a}_j \cdots) + s \cdot \det(\cdots \vec{a}_k \cdots)$$

mit  $s := 0$  und erhalten:

$$\begin{aligned} \det(\cdots r\vec{a}_j \cdots) &= \det(\cdots (r\vec{a}_j + 0\vec{a}_k) \cdots) \\ &= r \cdot \det(\cdots \vec{a}_j \cdots) + 0 \cdot \det(\cdots \vec{a}_k \cdots) \\ &= r \cdot \det(\cdots \vec{a}_j \cdots) \\ &= r \cdot \det A \end{aligned}$$

2. Null-Spalte: Falls  $\vec{a}_j = \vec{0}$ , denken wir uns an Stelle dieser Spalte einen beliebigen anderen Vektor  $\vec{a} \neq \vec{0}$ . Mit Satz 6.3 (Skalierung und Nullvektor) können wir nun  $\vec{a}$  mit dem Faktor 0 skalieren, was  $\vec{a}_j = \vec{0}$  ergibt. Aber dann folgt mit der Skalierungsregel direkt, dass

$$\begin{aligned} \det A &= \det(\cdots \vec{0} \cdots) \\ &= \det(\cdots 0 \cdot \vec{a} \cdots) \\ &= 0 \cdot \det(\cdots \vec{a} \cdots) \\ &= 0 \end{aligned}$$

3. Addition einer skalierten Spalte zu einer anderen Spalte: Auch hier verwenden wir zunächst die Multilinearität:

$$\begin{aligned} \det(\cdots (\vec{a}_j + s\vec{a}_k) \cdots \vec{a}_k \cdots) &= \det(\cdots (1 \cdot \vec{a}_j + s\vec{a}_k) \cdots \vec{a}_k \cdots) \\ &= 1 \cdot \det(\cdots \vec{a}_j \cdots \vec{a}_k \cdots) \\ &\quad + s \cdot \det(\cdots \vec{a}_k \cdots \vec{a}_k \cdots) \end{aligned}$$

Nun entspricht der erste Summand gerade  $\det A$ ; der zweite hingegen *verschwindet*, da die Determinante alterniert – hier steht an zwei verschiedenen Positionen  $j \neq k$  der gleiche Spaltenvektor  $\vec{a}_k$ . Insgesamt also:

$$\cdots = \det A$$

4. Linear abhängige Spalten: Wir unterscheiden zwei Fälle:

- Falls eine der Spalten der Nullvektor ist sind die Spalten linear abhängig. Dann folgt aber (s.o.) direkt, dass  $\det A = 0$  gilt.
- Wir nehmen nun an, dass keine der Spalten  $\vec{0}$  entspricht. Da die Spalten aber linear abhängig sind, gibt es eine nichttriviale Linearkombination des Nullvektors mit Koeffizienten  $x_1, \dots, x_n$

$$\sum_{j=1}^n x_j \vec{a}_j = \vec{0},$$

wobei *nicht alle*  $x_j$  null sind.

Wir wählen ein (ab hier festes)  $j$ , das  $x_j \neq 0$  erfüllt, und skalieren alle  $x_j$  mit dem Kehrwert  $\frac{1}{x_j}$ . Dann liegt immer noch eine Linearkombination des Nullvektors vor, und mit

$$y_k := \frac{x_k}{x_j}$$

ist  $x_j = 1$  und

$$\sum_{k=1}^n y_k \vec{a}_k = \vec{a}_j + \sum_{k \neq j} y_k \vec{a}_k = \vec{0}$$

Damit gilt insbesondere:

$$\vec{a}_j = - \sum_{k \neq j} y_k \vec{a}_k$$

Also lässt sich einer der Spaltenvektoren linear aus den anderen kombinieren. Die Vektoren auf der rechten Seite der letzten Gleichung sind jedoch sämtlich skalierte Spalten

$k \neq j$ . Addieren wir nun zu Spalte  $j$  alle anderen Spalten  $\vec{a}_k$ , skaliert mit den Faktoren  $y_k$ , so bleibt die Determinante  $\det A$  dabei unverändert (siehe die obige Invarianzeigenschaft).

Danach steht aber in Spalte  $j$  der Nullvektor  $\vec{0}$ , und es folgt wiederum, dass dann  $\det A = 0$  gelten muss.

5. Antisymmetrie in den Spalten: Wir zeigen, dass die Determinante beim Vertauschen zweier Spalten das Vorzeichen wechselt, indem wir mehrfach die Multilinearität ausnutzen; dazu die Skalierungsregel. Wir beginnen für  $j \neq k$  mit der ursprünglichen Spalten-Aufteilung von  $A$ , mit (oBdA)  $j < k$ :

$$\det A = \det \left( \cdots \quad \vec{a}_j \quad \cdots \quad \vec{a}_k \quad \cdots \right)$$

Nun addieren wir die Spalte  $k$  zur Spalte  $j$ , was wegen der Multilinearität die Determinante nicht verändert:

$$\cdots = \det \left( \cdots \quad (\vec{a}_j + \vec{a}_k) \quad \cdots \quad \vec{a}_k \quad \cdots \right)$$

Jetzt addieren wir das  $(-1)$ -fache der aktuellen Spalte  $j$  zur Spalte  $k$ :

$$\begin{aligned} \cdots &= \det \left( \cdots \quad (\vec{a}_j + \vec{a}_k) \quad \cdots \quad \vec{a}_k - (\vec{a}_j + \vec{a}_k) \quad \cdots \right) \\ &= \det \left( \cdots \quad (\vec{a}_j + \vec{a}_k) \quad \cdots \quad -\vec{a}_j \quad \cdots \right) \end{aligned}$$

Nun addieren wir nochmals die aktuelle Spalte  $k$  auf Spalte  $j$ :

$$\begin{aligned} \cdots &= \det \left( \cdots \quad (\vec{a}_j + \vec{a}_k) - \vec{a}_j \quad \cdots \quad -\vec{a}_j \quad \cdots \right) \\ &= \det \left( \cdots \quad \vec{a}_k \quad \cdots \quad -\vec{a}_j \quad \cdots \right) \end{aligned}$$

Nun skalieren wir die Spalte  $k$  mit  $(-1)$  und erhalten:

$$\cdots = -\det \left( \cdots \quad \vec{a}_k \quad \cdots \quad \vec{a}_j \quad \cdots \right)$$

Also ist

$$\det \left( \cdots \quad \vec{a}_k \quad \cdots \quad \vec{a}_j \quad \cdots \right) = -\det \left( \cdots \quad \vec{a}_j \quad \cdots \quad \vec{a}_k \quad \cdots \right) = -\det A$$

Damit ist alles gezeigt. ■

### Beweis zu Satz 8.20 (Leibnizformel) auf Seite 231:

Sei  $A = (\vec{a}_1 \quad \cdots \quad \vec{a}_n) \in \mathbb{R}^{(n,n)}$  gegeben. Dabei ist das Matrixelement  $A_{j,k}$  durch die  $j$ -te Komponente des Spaltenvektors  $\vec{a}_k$  gegeben. Dann ist

$$\vec{a}_k = \sum_{j=1}^n A_{j,k} \vec{e}_j$$

Wegen der Multilinearität gilt dann für die erste Spalte:

$$\det A = \det \left( \vec{a}_1 \quad \vec{a}_2 \quad \cdots \quad \vec{a}_n \right) = \sum_{j_1} A_{j_1,1} \cdot \det \left( \vec{e}_{j_1} \quad \vec{a}_2 \quad \cdots \quad \vec{a}_n \right)$$

Das gleiche Vorgehen lässt sich auf die anderen Spalten anwenden. Es ergibt sich die folgende Mehrfachsumme:

$$\cdots = \sum_{j_1, \dots, j_n} A_{j_1,1} \cdot \cdots \cdot A_{j_n,n} \det \left( \vec{e}_{j_1} \quad \cdots \quad \vec{e}_{j_n} \right)$$

Die Summe hat  $n^n$  Beiträge, da jeder Index unabhängig von den anderen einen Wert aus  $\{1, \dots, n\}$  annehmen kann.

Es verschwinden jedoch alle die Beiträge, für die zwei Indices (oder mehr) den selben Wert haben – dann sind nämlich zwei (oder mehr) Spalten der verbliebenen Matrix identisch, und wegen der Alterniertheit erhalten wir eine Determinante 0.

Die verbleibenden Beiträge sind genau die, für die in der verbliebenen Matrix sämtlich verschiedene kartesische Einheitsvektoren als Spalten auftreten – also für die die verbliebene Matrix eine *Permutationsmatrix* ist. Da die Mehrfachsumme alle Indexkombinationen durchläuft, tragen auch alle  $n!$  Permutationsmatrizen  $P_\sigma$  für  $\sigma \in S_n$  bei. Die Summationsindices  $j_1, \dots, j_n$  entsprechen

dann genau den Werten  $\sigma(1), \dots, \sigma(n)$ . Da alle Permutationen aus  $S_n$  gleichberechtigt beitragen, dürfen wir  $j_1$  mit  $\sigma(1)$  belegen, und so weiter bis  $j_n$ , das mit  $\sigma(n)$  belegt wird. Die Mehrfachsumme über alle Kombinationen von  $j_1, \dots, j_n$  wird dadurch zu einer (einfachen) Summe über sämtliche Permutationen aus  $S_n$ .

Die Determinanten der Permutationsmatrizen entsprechen nach Satz 8.18 jeweils dem Vorzeichen der zugehörigen Permutationen, sodass wir für die Determinante von  $A$  erhalten:

$$\dots = \sum_{\sigma \in S_n} \text{sign}(\sigma) \cdot A_{\sigma(1),1} \cdot \dots \cdot A_{\sigma(n),n}$$

Somit haben wir die eine der zwei Schreibweisen der Leibnizformel gefunden. Um die andere Schreibweise (in welcher die rechten Indices der Matrixelemente durch Permutationen ausgedrückt sind) zu erhalten, bedenken wir, dass in dem Produkt

$$A_{\sigma(1),1} \cdot \dots \cdot A_{\sigma(n),n}$$

aus obiger Formel sowohl die linken als auch die rechten Indices der Matrixelemente sämtliche Werte aus  $\{1, \dots, n\}$  jeweils in genau einem der Faktoren annehmen. Zu einem Matrixelement

$$A_{\sigma(j),j}$$

finden wir also stets ein Matrixelement

$$A_{j,k},$$

welches ebenfalls in dem Produkt auftritt, den linken Index mit Wert  $j$  besitzt. Da dies gerade dem Matrixelement

$$A_{\sigma(k),k}$$

entsprechen muss, gilt:

$$j = \sigma(k) \Leftrightarrow k = \sigma^{-1}(j)$$

Das gesuchte korrespondierende Matrixelement ist also

$$A_{j,\sigma^{-1}(j)}$$

Da Permutationen bijektive Abbildungen sind, besitzt jedes Element aus  $\{1, \dots, n\}$  genau ein Bild und auch genau ein Urbild; alle diese Ausdrücke sind also wohldefiniert – und zu jedem  $A_{\sigma(j),j}$  finden wir genau ein  $A_{j,\sigma^{-1}(j)}$ .

Die Produkte sämtlicher  $n$  Matrixelemente sind für festes  $\sigma$  gleich, da die Faktoren lediglich umgeordnet werden müssen. Also gilt:

$$\det A = \sum_{\sigma \in S_n} \text{sign}(\sigma) \cdot A_{1,\sigma^{-1}(1)} \cdot \dots \cdot A_{n,\sigma^{-1}(n)}$$

Nun verwenden wir noch Satz 8.15 (Vorzeichen von Permutationen bei Gruppenoperationen) und nutzen, dass sowohl  $\sigma$  als auch das Inverse  $\sigma^{-1}$  gleiches Vorzeichen haben:

$$\dots = \sum_{\sigma \in S_n} \text{sign}(\sigma^{-1}) \cdot A_{1,\sigma^{-1}(1)} \cdot \dots \cdot A_{n,\sigma^{-1}(n)}$$

Wir nutzen abermals aus, dass die Permutationen aus  $S_n$  bijektiv sind. Wenn wir über sämtliche Permutationen aus  $S_n$  summieren, summieren wir damit auch über sämtliche inversen Permutationen, also:

$$\dots = \sum_{\sigma^{-1} \in S_n} \text{sign}(\sigma^{-1}) \cdot A_{1,\sigma^{-1}(1)} \cdot \dots \cdot A_{n,\sigma^{-1}(n)}$$

Aber nun können wir, da in dem Ausdruck nur noch  $\sigma^{-1}$  auftritt, auch überall  $\sigma^{-1}$  durch  $\sigma$  umbenennen. Wir gehen die  $n!$  Permutationen aus  $S_n$  dann lediglich in einer anderen Reihenfolge durch; es werden jedoch weiterhin alle erreicht. Also:

$$\dots = \sum_{\sigma \in S_n} \text{sign}(\sigma) \cdot A_{1,\sigma(1)} \cdot \dots \cdot A_{n,\sigma(n)}$$

Hiermit ist die zweite Schreibweise der Leibniz-Formel gefunden. ■

**Beweis zu Satz 8.21 (Existenz der Determinante) auf Seite 233:**

Wir zeigen nun, dass die Funktion  $f$  mit

$$f(A) = \sum_{\sigma \in S_n} \text{sign}(\sigma) \cdot A_{1,\sigma(1)} \cdot \cdots \cdot A_{n,\sigma(n)} = \sum_{\sigma \in S_n} \text{sign}(\sigma) \cdot A_{\sigma(1),1} \cdot \cdots \cdot A_{\sigma(n),n}$$

die drei Weierstraß-Axiome erfüllt. Dass beide Schreibweisen gleichwertig sind, hatten wir bereits im Beweis zur Leibnizformel (Satz 8.20) gezeigt.

Die Spaltendarstellung von  $A$  sei

$$A = (\vec{a}_1 \quad \cdots \quad \vec{a}_n)$$

1. Normiertheit: Wir haben zu zeigen, dass  $f(\mathbb{1}_n) = 1$ . Nun sind für  $\mathbb{1}_n$  die einzigen Elemente ungleich 0 die Einsen auf der Diagonalen, da

$$(\mathbb{1}_n)_{j,k} = \delta_{jk}$$

Daher ist:

$$f(\mathbb{1}_n) = \sum_{\sigma \in S_n} \text{sign}(\sigma) \cdot \delta_{1,\sigma(1)} \cdot \cdots \cdot \delta_{n,\sigma(n)}$$

Diese Formel hat nur einen einzigen von 0 verschiedenen Beitrag, nämlich wenn sämtliche Kronecker-Deltas den Wert 1 besitzen. Aber dann ist für jedes  $j \in \{1, \dots, n\}$ :

$$j = \sigma(j)$$

Das ist die identische Permutation  $\text{id}_n$  mit Vorzeichen 1. Daher ist  $f(\mathbb{1}_n) = 1$ .

2. Alterniertheit: Wir nehmen an, dass die Spalten  $j$  und  $k$  mit (oBdA)  $j < k$  gleich sind. Wir wollen nun die Summe in  $f(A)$  geschickt so zerlegen, dass wir direkt zeigen können, dass sie 0 ergibt. Hierzu benötigen wir einige Vorarbeit.

Zunächst betrachten wir eine Untergruppe von  $S_n$ , nämlich die *alternierende Gruppe*  $A_n$ , die aus sämtlichen geraden Permutationen aus  $S_n$  besteht. Dies ist in der Tat eine Untergruppe, da die Komposition zweier gerader Permutationen nach Satz 8.15 (Vorzeichen von Permutationen bei Gruppenoperationen) erneut eine gerade Permutation ergibt (für die Elemente aus  $S_n \setminus A_n$  ist die Abgeschlossenheit dagegen nicht erfüllt).

Nun zeigen wir folgendes

**Lemma:** Für  $n \in \mathbb{N} \setminus \{1\}$  haben  $A_n$  und  $S_n \setminus A_n$  jeweils genau  $\frac{n!}{2}$  Elemente.

Wir wählen eine beliebige Permutation  $\pi \in S_n$  mit ungeradem Vorzeichen aus. Solch eine existiert stets, da  $n > 1$  – wir könnten z.B. eine beliebige Transposition für  $\pi$  wählen.

Dann definieren wir eine neue Funktion auf  $\{1, \dots, n\}$  per

$$g_\pi(\sigma) := \sigma \circ \pi$$

Für  $\sigma \in A_n$  ist  $g_\pi(\sigma)$  stets in  $S_n \setminus A_n$  und umgekehrt, da sich das Vorzeichen von  $\sigma$  beim Anwenden von  $\pi$  nach Satz 8.15 wegen  $\text{sign}(\pi) = -1$  stets umkehrt.

Nun ist  $S_n$  aber eine Gruppe; daher gilt Satz 5.9 (Kürzungsregel bei Gruppen). Dann ist aber die Abbildung  $g_\pi$  *bijektiv*, denn es gilt sowohl (und trivial)

$$(\sigma_1 = \sigma_2) \Rightarrow (g_\pi(\sigma_1) = g_\pi(\sigma_2))$$

als auch (Implikationsschritt per Kürzungsregel)

$$(g_\pi(\sigma_1) = g_\pi(\sigma_2)) \Leftrightarrow (\sigma_1 \circ \pi = \sigma_2 \circ \pi) \Rightarrow (\sigma_1 = \sigma_2)$$

Es können aber bijektive Funktionen zwischen endlichen Mengen nur dann bestehen, wenn die beiden Mengen gleich groß sind. Wir können weiterhin  $g_\pi$  nicht nur als Funktion auf  $S_n$  begreifen (das gilt ohnehin), sondern wegen der obigen Argumente auch einschränken:

$$g_\pi : A_n \rightarrow S_n \setminus A_n$$

An der Bijektivität von  $g_\pi$  ändert dies nichts – also müssen  $A_n$  und  $S_n \setminus A_n$  jeweils gleich groß sein. Sie sind außerdem beide nichtleer, da  $\text{id}_n \in A_n$  und  $\pi \in S_n \setminus A_n$ .

Da  $S_n$  nach Satz 1.40 (Anordnungen einer endlichen Menge) genau  $n!$  Permutationen enthält, gilt damit die Behauptung des Lemmas. ■

Wegen der Bijektivität von  $g_\pi$  können wir nun die Summe

$$\sum_{\sigma \in S_n} \cdots$$

aufsplitten als

$$\sum_{\sigma \in S_n} \cdots = \left( \sum_{\sigma \in A_n} \cdots \right) + \sum_{\tilde{\sigma} \in S_n \setminus A_n} \cdots$$

Beide Summen auf der rechten Seite haben  $\frac{n!}{2}$  Terme, und wenn wir die Permutation  $\tilde{\sigma}$  genau als  $g_\pi(\sigma)$  wählen (mit  $\text{sign}(\pi) = -1$ ), reicht es wegen der Bijektivität, nur die Summe über  $A_n$  zu betrachten und jeweils zwei Terme zu berücksichtigen – den linken mit  $\sigma \in A_n$ , und den korrespondierenden rechten mit  $\tilde{\sigma}$ : Denn sämtliche  $\sigma \in A_n$  besitzen aufgrund der Bijektivität von  $g_\pi$  paarweise verschiedene Bilder in  $S_n \setminus A_n$ .

Jetzt wählen wir als ungerade Permutation  $\pi$  genau die Transposition, die uns die Werte  $j$  und  $k$  vertauscht, also

$$\pi := \tau_{j,k}$$

Dann ist zu  $\sigma \in A_n$  stets:  $\tilde{\sigma} = \sigma \circ \tau_{j,k}$ .

Wir verwenden nun die zweite Schreibweise von  $f(A)$  (bei der die Spalten-Indices der Matrixelemente feste Zahlen sind) und bedenken, dass die Spalten  $j$  und  $k$  von  $A$  gleich sind. Dann erhalten wir:

$$\begin{aligned} f(A) &= \sum_{\sigma \in S_n} \text{sign}(\sigma) \cdot A_{\sigma(1),1} \cdot \cdots \cdot A_{\sigma(n),n} \\ &= \sum_{\sigma \in A_n} \underbrace{\text{sign}(\sigma)}_{=+1} \cdot A_{\sigma(1),1} \cdot \cdots \cdot A_{\sigma(n),n} \\ &\quad + \sum_{\tilde{\sigma} \in S_n \setminus A_n} \underbrace{\text{sign}(\tilde{\sigma})}_{=-1} \cdot A_{\tilde{\sigma}(1),1} \cdot \cdots \cdot A_{\tilde{\sigma}(n),n} \\ &= \sum_{\sigma \in A_n} (A_{\sigma(1),1} \cdot \cdots \cdot A_{\sigma(n),n}) - (A_{(\sigma \circ \tau_{j,k})(1),1} \cdot \cdots \cdot A_{(\sigma \circ \tau_{j,k})(n),n}) \end{aligned}$$

Nun enthält der linke Beitrag der Summe für jedes  $\sigma \in A_n$  an den Positionen  $j$  und  $k$  die Faktoren

$$A_{\sigma(j),j} \quad \text{und} \quad A_{\sigma(k),k}$$

Alle anderen Matrixelemente tauchen jeweils unverändert auch im rechten Beitrag der Summe auf – nur für die obigen beiden werden durch die Transposition zuvor noch die Argumente getauscht; es kommen dort also vor:

$$A_{(\sigma \circ \tau_{j,k})(j),j} = A_{\sigma(\tau_{j,k}(j)),j} = A_{\sigma(k),j} \quad \text{und} \quad A_{(\sigma \circ \tau_{j,k})(k),k} = A_{\sigma(\tau_{j,k}(k)),k} = A_{\sigma(j),k}$$

Da aber die Spalten  $j$  und  $k$  identisch angenommen waren, lässt sich für diese beiden Matrixelemente der Spaltenindex tauschen; also kommen im rechten Beitrag der Summe an den Positionen  $j$  und  $k$  die beiden Matrixelemente

$$A_{\sigma(k),k} \quad \text{und} \quad A_{\sigma(j),j}$$

vor. Wegen der Kommutativität der Multiplikation entspricht daher der rechte Beitrag der Summe genau dem linken – aber da ersterer ein negatives Vorzeichen hat, ergibt der Gesamtbeitrag der Summe für jedes  $\sigma \in A_n$  stets 0, und dann ist damit auf  $f(A) = 0$ .

3. Multilinearität in den Spalten: Zu zeigen ist für  $r, s \in \mathbb{R}$  sowie beliebige Spaltenvektoren  $\vec{a}_j, \vec{a}_k$  an einer beliebigen aber festen Spaltenposition  $p$ :

$$f(\cdots \quad (r\vec{a}_j + s\vec{a}_k) \quad \cdots) = r \cdot f(\cdots \quad \vec{a}_j \quad \cdots) + s \cdot f(\cdots \quad \vec{a}_k \quad \cdots)$$

Wir verwenden für  $f(A)$  die zweite Schreibweise (bei der die Spalten-Indices der Matrixelemente feste Zahlen sind) und bedenken, dass die Komponente mit Zeilenindex  $\sigma(p)$  und Spaltenindex  $p$  gegeben ist durch die  $\sigma(p)$ -te Komponente des  $p$ -ten Spaltenvektors:

$$f(\cdots (r\vec{a}_j + s\vec{a}_k) \cdots) = \sum_{\sigma \in S_n} \text{sign}(\sigma) \cdot \cdots \cdot (r\vec{a}_j + s\vec{a}_k)_{\sigma(p)} \cdot \cdots$$

Nun können wir die komponentenweise Erklärung der Operationen in  $\mathbb{R}^n$  verwenden, um zusammen mit dem Distributivgesetz zu folgern:

$$\begin{aligned} \cdots &= r \cdot \sum_{\sigma \in S_n} \text{sign}(\sigma) \cdot \cdots \cdot (\vec{a}_j)_{\sigma(p)} \cdot \cdots \\ &\quad + s \cdot \sum_{\sigma \in S_n} \text{sign}(\sigma) \cdot \cdots \cdot (\vec{a}_k)_{\sigma(p)} \cdot \cdots \\ &= r \cdot f(\cdots \vec{a}_j \cdots) + s \cdot f(\cdots \vec{a}_k \cdots) \end{aligned}$$

Es gelten also die drei Eigenschaften aus den Axiomen von Weierstraß, und damit ist  $f(A) = \det A$  eine Determinante. ■

**Beweis zu Satz 8.24 (Determinanten von Dreiecksmatrizen) auf Seite 234:**

Sei  $A \in \mathbb{R}^{(n,n)}$  eine Dreiecksmatrix. Wir berechnen die Determinante mit der Leibnizformel aus Satz 8.20. Hierzu beweisen wir folgendes

**Lemma:** Für alle Permutationen  $\sigma \in S_n \setminus \{\text{id}_n\}$  existieren  $j, k \in \{1, \dots, n\}$ , sodass

$$\sigma(j) > j \quad \text{und} \quad \sigma(k) < k$$

Nur die identische Permutation besitzt genau  $n$  kanonische Zyklen, die alle trivial sind, da jedes Element auf sich selbst abgebildet wird. Daher besitzen alle anderen Permutationen aus  $S_n$  mindestens einen nichttrivialen kanonischen Zyklus mit mindestens zwei Elementen. Da diese beiden Elemente verschieden sind, ist eine der beiden Behauptungen damit schon gezeigt. Wir verfolgen solch einen Zyklus  $\zeta$  nun durch wiederholtes Anwenden von  $\sigma$  (bzw.  $\zeta$ , da es sich um einen kanonischen Zyklus handelt) weiter. Klar ist, dass der Zyklus, von unserem Ursprungselement  $l$  ausgehend, nach  $|\zeta|$  Schritten wieder bei  $l$  ankommen muss. Aber dann muss die Änderung  $(\zeta(l) - l)$ , ob sie nun positiv oder negativ war, im Verlauf dieser Schritte wieder ausgeglichen werden. Somit sind auf  $\zeta$  stets beide behaupteten Eigenschaften irgendwo realisiert. ■

Nun betrachten wir die Determinante aus der Leibnizformel:

$$\det A = \sum_{\sigma \in S_n} \text{sign}(\sigma) \cdot A_{1,\sigma(1)} \cdot \cdots \cdot A_{n,\sigma(n)}$$

Falls  $\sigma \neq \text{id}_n$  gibt es nach obigem Lemma stets, und damit auch für obere Dreiecksmatrizen, ein  $k$ , sodass  $\sigma(k) < k$ . Aber dann ist  $A_{k,\sigma(k)}$  nach Definition 8.23 sicher 0, und der zu  $\sigma$  gehörige Beitrag in der Summe verschwindet. Für untere Dreiecksmatrizen findet sich dagegen stets ein  $j$  mit  $\sigma(j) > j$ , sodass  $A_{j,\sigma(j)} = 0$  gilt.

Dann bleibt jedoch aus der Summe über  $S_n$  nur ein potentiell nichtverschwindender Beitrag für die Determinante übrig, nämlich  $\sigma = \text{id}_n$ : Dieser Beitrag liefert uns genau das Produkt der Diagonalelemente von  $A$ . ■

**Beweis zu Satz 8.25 (Lineare Abhängigkeit und Determinante) auf Seite 234:**

Zunächst gilt mit den Sätzen 8.17 und 8.22, dass die Determinante verschwindet, falls die Spalten oder die Zeilen einer Matrix  $A \in \mathbb{R}^{(n,n)}$  linear abhängig sind. Wir zeigen hier noch die umgekehrte Richtung der Implikation, nämlich, dass bei  $\det A = 0$  sowohl die Spalten als auch die Zeilen von  $A$  linear abhängig sind.

Wir betrachten das homogene Gleichungssystem  $(A \mid \vec{0})$ . Mit den bekannten (und linearen!) Zeilenoperationen können wir eine Zeilen-Stufen-Form mit Koeffizientenmatrix  $B$  ermitteln. Wir wollen hier die entstehenden Null-Zeilen nicht löschen, damit die Koeffizientenmatrix während der Operationen quadratisch bleibt.

Die Zeilen-Stufen-Form einer quadratischen Matrix (ohne Löschen von Null-Zeilen) ist stets eine obere Dreiecksmatrix.

Nun hat sich durch die Zeilen-Operationen allerdings der Wert der Determinanten nicht geändert. Denn entweder bleibt der Wert der Determinanten gleich (beim Addieren skalierten Zeilen), oder er ändert sich multiplikativ (beim Vertauschen oder Skalieren von Zeilen). Da aber  $\det A = 0$  vorausgesetzt ist, kann sich auch durch multiplikative Änderung kein anderer Wert als 0 einstellen; somit ist auch  $\det B = 0$ .

Da aber nach Satz 8.24 die Determinante einer Dreiecksmatrix als Produkt aller Diagonalelemente gegeben ist, muss mindestens eines der Diagonalelemente von  $B$  nach Satz 5.16 (Satz vom Nullprodukt) null betragen.

Hätte  $A$  nun aber vollen Rang  $n$ , so wären sämtliche Elemente auf der Diagonalen von  $B$  den Wert 1 einstellbar. So jedoch muss die Stufenlinie ungleichmäßig verlaufen (denn eine der Zeilen enthält auf der Diagonalen eine Null, und die 1 der zugehörigen Zeile tritt, wenn überhaupt, erst weiter rechts auf) und kann daher die unterste Zeile von  $B$  nicht mehr erreichen; es gibt in  $B$  also mindestens eine Null-Zeile.

Damit ist aber klar, dass der Kern von  $A$  nichttrivial ist; das LGS  $(A \mid \vec{0})$  ist also nach Satz 8.10 (Rang einer Matrix) nicht eindeutig lösbar. Da es sich um ein homogenes LGS handelt, muss es also unendlich viele Lösungen geben, und damit sicher auch eine nichttriviale. Solch eine Lösung ist aber nach Satz 8.2 (Äquivalente Beschreibungen von LGS) auch eine nichttriviale Linearkombination des Nullvektors aus den Spalten von  $A$ . Also folgt, dass diese Spalten linear abhängig sind.

Da mit  $\det A = 0$  nach Satz 8.22 immer auch  $\det A^T = 0$  gilt, können wir das Gleiche für die Spalten von  $A^T$  (also für die Zeilen von  $A$ ) argumentieren – auch diese sind linear abhängig. ■

#### **Beweis zu Satz 8.26 (Laplace-Entwicklung) auf Seite 235:**

Wir leiten die behaupteten Formeln aus der Leibniz-Formel (Satz 8.20) her, die für alle Determinanten gültig ist. Zunächst betrachten wir die erste Spalte – danach folgt das Vorgehen für eine beliebige Spalte, dann für beliebige Zeilen.

Gegeben sei die Matrix

$$A = (\vec{a}_1 \quad \vec{a}_2 \quad \cdots \quad \vec{a}_n) \in \mathbb{R}^{(n,n)}$$

Nun kann man über die Multilinearität der Determinanten (siehe Definition 8.16) argumentieren, indem man die erste Spalte nach den Einheitsvektoren der Standardbasis entwickelt:

$$\vec{a}_1 = \sum_{j=1}^n A_{j,1} \vec{e}_j$$

Es ergibt sich also direkt das Summenzeichen, das auch in der behaupteten Formel zu finden ist, sowie die jeweiligen Vorfaktoren  $A_{j,1}$ , die sich aus der Determinantenberechnung heraus ziehen lassen. Die verbleibenden  $n$  Determinanten enthalten als erste Spalte jeweils einen der kartesischen Einheitsvektoren, von denen jeweils  $(n-1)$  Komponenten 0 sind. Setzt man nun die Leibniz-Formel ein, so werden dadurch die meisten Beiträge überflüssig, und man gelangt zur behaupteten Formel.

Wir verfolgen hier einen etwas anderen Ansatz, der aber effektiv das Gleiche bewirkt. Und zwar gehen wir zunächst von der Leibniz-Formel (in der hier für uns günstigen Schreibweise) aus:

$$\det A = \sum_{\sigma \in S_n} \text{sign}(\sigma) \cdot A_{\sigma(1),1} \cdot \cdots \cdot A_{\sigma(n),n}$$

Da wir zunächst nach der ersten Spalte entwickeln wollen, sind die Matricelemente  $A_{\sigma(1),1}$  hier besonders wichtig.

Wir erinnern uns, dass jeder Beitrag zur Summe in der Leibniz-Formel ein Produkt aus Matricelementen derart enthält, dass sich in jeder Zeile und in jeder Spalte jeweils genau eines dieser  $n$  Elemente befindet. Ist also ein Matricelement  $A_{j,k}$  im Beitrag enthalten, werden die  $(n-1)$  anderen Matricelemente aus Zeilen ungleich  $j$  und Spalten ungleich  $k$  stammen.

Das Matricelement  $A_{1,1}$  ist hierbei in sämtlichen Beiträgen enthalten, für die  $\sigma(1) = 1$  erfüllt ist, d.h.

$$\sigma = (1) \circ \tilde{\sigma},$$

wobei  $\tilde{\sigma}$  effektiv als eine bijektive Abbildung der Menge  $\{2, 3, \dots, n\} = M_n \setminus \{1\}$  auf sich selbst aufgefasst werden kann. Hierbei kommt die Zahl 1 weder als Urbild noch als Bild vor. Die erwähnte Menge enthält  $(n-1)$  Elemente und ist isomorph zu  $M_{n-1}$ , d.h. wir finden eine umkehrbar



eindeutige Abbildung zwischen den beiden Mengen. Wir können, diese entsprechende Abbildung mit einbezogen, also  $\tilde{\sigma}$  effektiv als eine der Permutationen aus  $S_{n-1}$  betrachten; wir notieren dies als  $S_{n-1}(M_n \setminus \{1\})$ .

Das Vorzeichen von  $\tilde{\sigma}$  entspricht nach Satz 8.15 genau dem Vorzeichen von  $\sigma$ , denn der triviale Zyklus (1) ist eine gerade Permutation.

Also lassen alle Beiträge mit dem Matricelement  $A_{1,1}$  zusammen fassen als

$$A_{1,1} \cdot \sum_{\tilde{\sigma} \in S_{n-1}(M_n \setminus \{1\})} \text{sign}(\tilde{\sigma}) A_{\tilde{\sigma}(2),2} \cdot \cdots \cdot A_{\tilde{\sigma}(n),n}$$

Wenn wir mit  $A_{\setminus(j,k)}$  wie in der Behauptung definiert die  $(n-1) \times (n-1)$ -Matrix verstehen, die durch Streichen der Zeile  $j$  und der Spalte  $k$  entsteht, ist  $A_{\setminus(1,1)}$  genau die Matrix, für die in  $A$  erste Zeile und erste Spalte gestrichen wurden. Es bleiben von  $A$  dann nur die Zeilen und Spalten 2 bis  $n$ . Damit, und mit Satz 8.21 (Existenz der Determinante), erkennen wir in obigem Ausdruck eine Unterdeterminante von  $A$  wieder und können die Beiträge zu  $\det A$  mit Matricelement  $A_{1,1}$  umschreiben als

$$\cdots = A_{1,1} \cdot \det A_{\setminus(1,1)}$$

Nun reicht es für die Berechnung von  $\det A$ , weiterhin sämtliche Beiträge mit den Matricelementen  $A_{j,1}$  mit  $j = 2, 3, \dots, n$  zu betrachten. Denn jeder Beitrag von  $\det A$  nach der Leibnizformel enthält genau ein Matricelement aus der ersten Spalte. Also tun wir nichts anderes, als dieser Beiträge so zu ordnen, dass wir die Zeilen  $j$  systematisch der Reihe nach durchgehen – in der allgemeinen Leibnizformel dürfen hingegen die Permutationen  $\sigma$  in beliebiger Reihenfolge ausgewertet werden.

Wir können die weiteren Überlegungen nun wesentlich vereinfachen, indem wir den Fall  $A_{j,1}$  auf unseren Einstiegsfall  $A_{1,1}$  zurück führen. Und zwar wissen wir, dass wir auch hier Beiträge betrachten, die neben dem Element  $A_{j,1}$  keine Matricelemente aus Zeile  $j$  und Spalte 1 enthalten; wir streichen diese also ähnlich wie oben und gelangen zu einer Unterdeterminanten.

Nun können wir einen zyklischen Tausch der ersten  $j$  Zeilen von  $A$  so durchführen, dass in der neuen Matrix  $B(j)$  die  $j$ -te Zeile von  $A$  an erster Stelle steht. Das ist nach Satz 8.11 (Dekomposition eines Zyklus in Transpositionen) mit genau  $(j-1)$  Transpositionen möglich – die dabei entstehenden Vorzeichenwechsel in der Determinanten kompensieren wir demnach mit dem Faktor  $(-1)^{j-1}$ .

Dann ist jedoch auch:

$$A_{j,1} = (B(j))_{1,1}$$

Und damit sind die Beiträge zu  $\det A$ , die  $A_{j,1}$  enthalten, gerade die mit dem Kompensationsfaktor versehenen Beiträge, die bei der Determinantenberechnung von  $B(j)$  das Element  $(B(j))_{1,1}$  enthalten. Dies sind also:

$$(-1)^{j-1} \cdot (B(j))_{1,1} \cdot \det B(j)_{\setminus(1,1)}$$

Nun können wir wieder zurück einsetzen:

$$\cdots = (-1)^{j-1} \cdot A_{j,1} \cdot \det A_{\setminus(j,1)}$$

Die Formel stimmt im Übrigen auch für  $j = 1$ , also für die erste Zeile von  $A$ , denn dann sind  $(j-1) = 0$  Zeilenvertauschungen nötig, und der Kompensationsfaktor ist  $(-1)^0 = 1$ .

Insgesamt gilt also für die Determinante von  $A$  (nach der ersten Spalte entwickelt):

$$\det A = \sum_{j=1}^n (-1)^{j-1} \cdot A_{j,1} \cdot \det A_{\setminus(j,1)}$$

Die Argumentation für die Entwicklung nach einer beliebigen Spalte  $k$  ist nun nicht mehr schwer: Wir führen sie auf die Entwicklung nach Spalte 1 zurück, indem wir im Voraus  $(k-1)$  Transpositionen durchführen, die uns Spalte  $k$  zyklisch an Position 1 bringen. Die Determinante von  $A$  bekommt dann analog zu eben einen globalen Kompensationsfaktor  $(-1)^{k-1}$ .

Nun berücksichtigen wir noch, dass  $(-1)^{j-1} \cdot (-1)^{k-1} = (-1)^{j+k-2} = (-1)^{j+k} \cdot 1$ , und erhalten die behauptete Formel für die Entwicklung nach einer beliebigen (aber festen) Spalte  $k$ :

$$\boxed{\det A = \sum_{j=1}^n (-1)^{j+k} \cdot A_{j,k} \cdot \det A_{\setminus(j,k)}}$$

Zum Schluss leiten wir daraus die Formel für die Entwicklung nach einer beliebigen festen Zeile  $j$  her. Wir benötigen dafür Satz 8.22 über die Determinante der transponierten Matrix. Denn

es läuft auf das Gleiche hinaus, ob wir die Determinante von  $A$  nach Zeile  $j$  entwickeln, oder die von  $A^T$  nach Spalte  $j$  (den Summenindex hier auf  $k$  geändert):

$$\det A = \det A^T = \sum_{k=1}^n (-1)^{k+j} \cdot A_{k,j}^T \cdot \det A_{\setminus(k,j)}^T$$

erhalten wir aber, wenn wir aus  $A^T$  die Zeile  $k$  und Spalte  $j$  streichen, die gleiche Matrix, die sich ergibt, wenn wir aus  $A$  die Zeile  $j$  und die Spalte  $k$  streichen und dann transponieren. Weiterhin ist  $A_{k,j}^T = A_{j,k}$  sowie  $(-1)^{k+j} = (-1)^{j+k}$ . Also gilt auch, wie behauptet:

$$\det A = \sum_{k=1}^n (-1)^{j+k} \cdot A_{j,k} \cdot \det A_{\setminus(j,k)}$$

Damit sind beide behaupteten Formeln gezeigt. ■

**Beweis zu Satz 8.27 (Determinanten-Produktsatz) auf Seite 241:**

Für  $A, B \in \mathbb{R}^{(n,n)}$  sei  $C := AB$ . Die Spaltendarstellungen von  $A$  und  $C$  sind:

$$A = (\vec{a}_1 \quad \cdots \quad \vec{a}_n) \quad \text{und} \quad C = (\vec{c}_1 \quad \cdots \quad \vec{c}_n)$$

Wir drücken zunächst die Spalten von  $C$  als Linearkombinationen der Spalten von  $A$  aus (die Summenindices verstehen sich hier sämtlich von 1 bis  $n$ ):

$$(\vec{c}_r)_j = C_{j,r} = (AB)_{j,r} = \sum_s A_{j,s} B_{s,r} = \sum_s B_{s,r} (\vec{a}_s)_j \quad \Leftrightarrow \quad \vec{c}_r = \sum_s B_{s,r} \cdot \vec{a}_s$$

Damit wird dann die Determinante von  $C = AB$  zu

$$\det(AB) = \det C = \det (\vec{c}_1 \quad \cdots \quad \vec{c}_n) = \det \left( \sum_{s_1} B_{s_1,1} \vec{a}_{s_1} \quad \cdots \quad \sum_{s_n} B_{s_n,n} \vec{a}_{s_n} \right)$$

Nun verwenden wir die Multilinearität (ähnlich wie im Beweis der Leibnizformel, Satz 8.20) und extrahieren eine Mehrfachsumme:

$$\cdots = \sum_{s_1, \dots, s_n} B_{s_1,1} \cdot \cdots \cdot B_{s_n,n} \cdot \underbrace{\det (\vec{a}_{s_1} \quad \cdots \quad \vec{a}_{s_n})}_*$$

Aufgrund der Alterniertheit der Determinante kann  $*$  jedoch nur von 0 verschieden sein, wenn keine der Spalten von  $A$  mehrfach auftritt – sämtliche Summationsindices müssen also paarweise verschieden sein.

Aber dann sind  $s_1, \dots, s_n$  als die Werte einer der Permutationen aus  $S_n$  gegeben. Also bleibt die einfache Summe über  $S_n$ :

$$\cdots = \sum_{\sigma \in S_n} B_{\sigma(1),1} \cdot \cdots \cdot B_{\sigma(n),n} \cdot \underbrace{\det (\vec{a}_{\sigma(1)} \quad \cdots \quad \vec{a}_{\sigma(n)})}_{**}$$

Für den Ausdruck  $**$  wenden wir Satz 8.19 (Determinante einer Matrix mit permutierten Spalten) an:

$$\cdots = \sum_{\sigma \in S_n} B_{\sigma(1),1} \cdot \cdots \cdot B_{\sigma(n),n} \cdot \det(A \cdot P_\sigma) = \sum_{\sigma \in S_n} B_{\sigma(1),1} \cdot \cdots \cdot B_{\sigma(n),n} \cdot \text{sign}(\sigma) \cdot \det A$$

Nun ziehen wir den Faktor  $\det A$  aus der Summe und erhalten:

$$\cdots = (\det A) \cdot \sum_{\sigma \in S_n} \text{sign}(\sigma) \cdot B_{\sigma(1),1} \cdot \cdots \cdot B_{\sigma(n),n}$$

Aber nun entspricht die Summe gerade einer der beiden Schreibweisen der Leibniz-Formel (Satz 8.20), sodass insgesamt gilt:

$$\cdots = (\det A) \cdot (\det B)$$

Da diese Gleichungskette auch von rechts nach links lesbar ist, und da

$$(\det A) \cdot (\det B) = (\det B) \cdot (\det A),$$

gilt auch die Gleichheit mit  $\det(BA)$ , wie behauptet. ■

## Für Kapitel 9

### Beweis zu Satz 9.11 (Linearkombination von Eigenvektoren) auf Seite 253:

Seien  $k$  Eigenvektoren  $\vec{v}_1, \dots, \vec{v}_k$  von  $A$  zum (festen) Eigenwert  $\lambda$  gegeben. Es gilt dann für alle  $j \in \{1, \dots, k\}$  die Eigenwertgleichung:

$$A\vec{v}_j = \lambda\vec{v}_j$$

Dann betrachten wir eine beliebige Linearkombination und setzen diese in die Eigenwertgleichung ein. Hierbei benutzen wir die Linearität der Matrixmultiplikation:

$$A \left( \sum_j c_j \vec{v}_j \right) = \sum_j c_j (A\vec{v}_j) = \sum_j c_j (\lambda \vec{v}_j) = \lambda \cdot \left( \sum_j c_j \vec{v}_j \right) \quad \blacksquare$$

### Beweis zu Satz 9.18 (Produkt und Summe der Eigenwerte) auf Seite 261:

In den Bemerkungen zu Definition 9.14 (Charakteristisches Polynom) hatten wir schon fest gestellt, dass

$$\chi_A(0) = \det A$$

Nun schreiben wir das charakteristische Polynom für  $A_\lambda := A - \lambda \mathbb{1}_n$  nochmals an und benutzen für die Determinante die Leibniz-Formel aus Satz 8.20:

$$\chi_A(\lambda) = \det A_\lambda = \sum_{\sigma \in S_n} \text{sign}(\sigma) (A_\lambda)_{1,\sigma(1)} \cdot \dots \cdot (A_\lambda)_{n,\sigma(n)}$$

Nun gibt es genau eine Permutation in  $S_n$ , die sämtliche Diagonalelemente von  $A_\lambda$  beitrugen lässt, nämlich  $\text{id}_n$ . Alle anderen Permutationen aus  $S_n \setminus \{\text{id}_n\}$  können höchstens  $n-2$  Diagonalelemente von  $A_\lambda$  zur Summe beisteuern, denn falls  $j = \sigma(j)$  für insgesamt  $n-1$  Elemente gelten würde, dann bliebe aufgrund der Bijektivität von  $\sigma$  für das letzte Element ebenfalls nur noch die identische Abbildung übrig.

Da aber genau die Diagonalelemente von  $A_\lambda$  Linearfaktoren mit  $\lambda$  enthalten, kann mit den Permutationen aus  $S_n \setminus \{\text{id}_n\}$  nur ein Polynom  $(n-2)$ -ten Grades gebildet werden. Die Beiträge mit Grad  $n$  und mit Grad  $n-1$  stammen daher ausschließlich vom Beitrag  $\sigma = \text{id}_n$ . Wir spalten also das charakteristische Polynom auf:

$$\chi_A(\lambda) = \underbrace{\prod_{j=1}^n (A_{j,j} - \lambda)}_{=: p(\lambda)} + \sum_{\sigma \in S_n \setminus \{\text{id}_n\}} \text{sign}(\sigma) (A_\lambda)_{1,\sigma(1)} \cdot \dots \cdot (A_\lambda)_{n,\sigma(n)}$$

Wir schreiben nun das fragliche Polynom  $p(\lambda)$ , dessen Koeffizienten zu  $\lambda^n$  und  $\lambda^{n-1}$  den entsprechenden Koeffizienten in  $\chi_A(\lambda)$  entsprechen, alternativ an:

$$\begin{aligned} p(\lambda) &= \prod_{j=1}^n (A_{j,j} - \lambda) = (-1)^n \cdot \prod_{j=1}^n (\lambda - A_{j,j}) \\ &= (-1)^n \cdot (\lambda - A_{1,1})(\lambda - A_{2,2}) \cdots (\lambda - A_{n,n}) \end{aligned}$$

Wenn wir in Gedanken dieses Produkt ausmultiplizieren, dann entstehen Beiträge mit Faktor  $\lambda^{n-1}$  genau dann, wenn  $(n-1)$  der Klammern zu  $\lambda$  evaluieren und eine der Klammern zu  $(-A_{j,j})$ . Dies ist auf genau  $n$  verschiedene Weisen möglich, und jede trägt additiv bei. Daher ist der Koeffizient  $p_{n-1}$  gegeben als

$$p_{n-1} = (-1)^n \cdot \sum_{j=1}^n (-A_{j,j}) = -(-1)^n \cdot \sum_{j=1}^n A_{j,j}$$

Dies ist auch genau der Koeffizient für das Monom  $\lambda^{n-1}$  in  $\chi_A(\lambda)$ .

Weiterhin gilt nach dem Fundamentalsatz der Algebra (Satz 5.29) und den dortigen Bemerkungen, dass das Polynom  $\chi_A(\lambda)$  in  $n$  Linearfaktoren zerfällt, die durch die Nullstellen des Polynoms bestimmt sind – also hier durch die Eigenwerte (siehe Satz 9.15):

$$\chi_A(\lambda) = (-1)^n \cdot (\lambda - \lambda_1)(\lambda - \lambda_2) \cdots (\lambda - \lambda_n) \quad (*)$$

Nun ist der Koeffizient für das Monom  $\lambda^{n-1}$  hier mit gleicher Begründung gegeben als

$$-(-1)^n \cdot \sum_{j=1}^n \lambda_j$$

Aber dann folgt durch Koeffizientenvergleich, dass die Summe der Eigenwerte (mehrfache Eigenwerte entsprechend mehrfach gezählt) gerade der Summe der Diagonalelemente von  $A$  entspricht.

Weiterhin folgt für den Wert  $\det A = \chi_A(0)$  mit der Formel aus (\*):

$$\det A = \chi_A(0) = (-1)^n \cdot (-\lambda_1)(-\lambda_2) \cdots (-\lambda_n) = (-1)^n \cdot \prod_{j=1}^n (-\lambda_j) = \prod_{j=1}^n \lambda_j$$

Hiermit sind beide behaupteten Eigenschaften gezeigt. ■

**Beweis zu Satz 9.20 (Eigenvektoren und Eigenwerte der inversen Matrix) auf Seite 262:**

Wir schreiben die Eigenwertgleichung für  $\vec{v}$  an:

$$A\vec{v} = \lambda\vec{v}$$

Nun multiplizieren wir von links die inverse Matrix:

$$\cdots \Leftrightarrow A^{-1}A\vec{v} = \mathbb{1}_n\vec{v} = \vec{v} = A^{-1}(\lambda\vec{v}) = \lambda A^{-1}\vec{v}$$

Also gilt mit  $\vec{v} = \lambda A^{-1}\vec{v}$  auch:

$$\cdots \Leftrightarrow \frac{1}{\lambda}\vec{v} = A^{-1}\vec{v}$$

Die Skalierung mit dem Kehrwert von  $\lambda$  ist möglich, da nach Satz 9.19  $\lambda \neq 0$  gilt.

Aber nun steht dort genau die Eigenwertgleichung für den selben Vektor  $\vec{v}$  zur Matrix  $A^{-1}$ , mit Eigenwert  $\frac{1}{\lambda}$ . ■

**Beweis zu Satz 9.22 (Eigenvektoren zu verschiedenen Eigenwerten) auf Seite 263:**

Der Beweis benötigt eigentlich die Technik der *vollständigen Induktion*, die in Mathematik 1 kein Stoff ist. Wir können unsere Argumente allerdings auch ohne Kenntnis der formalen Begriffe eines Induktionsbeweises begründen.

Für  $A \in \mathbb{R}^{(n,n)}$  seien  $m \leq n$  paarweise verschiedene Eigenwerte  $\lambda_1, \dots, \lambda_m$  ermittelt worden (hier ohne Mehrfachzählung); zu jedem  $\lambda_j$  dieser  $m$  verschiedenen Eigenwerte liege außerdem ein Eigenvektor  $\vec{v}_j$  vor.

Wir zeigen nun rekursiv, dass die  $m$  Eigenvektoren  $\vec{v}_1, \dots, \vec{v}_m$  linear unabhängig sind. Zunächst die Bedingung für den Rekursionsabbruch: Falls  $m = 1$ , so ist  $\vec{v}_1 \neq \vec{0}$  der einzige Eigenvektor und ist nach Satz 6.22 linear unabhängig.

Für  $m > 1$  führen wir das Problem von  $k \in \mathbb{N}$  verschiedenen Eigenwerten auf das für  $(k-1)$  zurück. Wenn dies allgemein gelingt, so können wir von  $k = 1$  auf  $k = 2$  schließen, danach von  $k = 2$  auf  $k = 3$ , und so weiter bis zum letzten Schluss von  $k = (m-1)$  auf  $k = m$ .

Sei also für ein beliebiges  $1 \leq k < m$  gezeigt, dass die Vektoren  $\vec{v}_1, \dots, \vec{v}_{k-1}$  linear unabhängig sind.

Dann gilt nach Definition 6.21, dass die Linearkombination des Nullvektors

$$\sum_{j=1}^{k-1} \tilde{c}_j \vec{v}_j = \vec{0}$$

trivial ist – d.h.  $\tilde{c}_1, \dots, \tilde{c}_{k-1}$  betragen alle 0.

Nun betrachten wir eine Linearkombination des Nullvektors aus den ersten  $k$  Eigenvektoren, d.h. wir fügen  $\vec{v}_k$  noch hinzu. Dann betrachten wir die Gleichung

$$c_1 \vec{v}_1 + \cdots + c_{k-1} \vec{v}_{k-1} + c_k \vec{v}_k = \vec{0} \quad (*)$$

Wir multiplizieren die Gleichung (\*) nun mit der Matrix  $A$  (von links) – dabei verwenden wir die jeweiligen Eigenwertgleichungen.

$$\begin{aligned} (*) &\Rightarrow c_1 A \vec{v}_1 + \cdots + c_{k-1} A \vec{v}_{k-1} + c_k A \vec{v}_k = A \vec{0} \\ &\Rightarrow c_1 \lambda_1 \vec{v}_1 + \cdots + c_{k-1} \lambda_{k-1} \vec{v}_{k-1} + c_k \lambda_k \vec{v}_k = \vec{0} \end{aligned} \quad (**)$$

Außerdem multiplizieren wir die Gleichung (\*) auch einmal mit dem Eigenwert  $\lambda_k$ :

$$(*) \Rightarrow c_1 \lambda_k \vec{v}_1 + \cdots + c_{k-1} \lambda_k \vec{v}_{k-1} + c_k \lambda_k \vec{v}_k = \vec{0} \quad (***)$$

Die Gleichungen (\*\*) und (\*\*\*) sehen recht ähnlich aus, unterscheiden sich jedoch darin, dass in (\*\*\*) jeder Vektor mit  $\lambda_k$  skaliert ist, in (\*\*) dagegen jeder Vektor mit seinem zugehörigen Eigenwert. Wir bilden die Differenz der beiden Gleichungen. Wegen der Distributivgesetze in Vektorräumen (siehe Definition 6.1) können wir die Skalierungsfaktoren der einzelnen Eigenvektoren jeweils isoliert voneinander verrechnen:

$$(*) \Rightarrow c_1 (\lambda_1 - \lambda_k) \vec{v}_1 + \cdots + c_{k-1} (\lambda_{k-1} - \lambda_k) \vec{v}_{k-1} + c_k (\lambda_k - \lambda_k) \vec{v}_k = \vec{0}$$

Nun ist der Klammerterm bei Vektor  $\vec{v}_k$  genau 0, sodass sich ergibt:

$$\cdots \Rightarrow c_1 (\lambda_1 - \lambda_k) \vec{v}_1 + \cdots + c_{k-1} (\lambda_{k-1} - \lambda_k) \vec{v}_{k-1} = \vec{0}$$

Hier liegt wieder eine Linearkombination des Nullvektors aus den ersten  $(k-1)$  Eigenvektoren vor, die wir bereits als linear unabhängig angenommen hatten. Wenn wir nun also

$$\tilde{c}_j := c_j (\lambda_j - \lambda_k)$$

setzen, so gilt (s.o.), dass sämtliche  $(k-1)$  Faktoren  $\tilde{c}_j$  gleich 0 sein müssen:

$$\forall j \in \{1, \dots, k-1\} : \tilde{c}_j = c_j (\lambda_j - \lambda_k) = 0$$

Da nun aber die Eigenwerte  $\lambda_1, \dots, \lambda_m$  alle paarweise verschieden sind, müssen die Differenzen  $(\lambda_j - \lambda_k)$  für jedes  $j$  zwischen 1 und  $(k-1)$  *ungleich* 0 sein. Nach dem Satz vom Nullprodukt (5.16) impliziert dies allerdings, dass  $c_j = 0$  für alle betreffenden  $j$  gilt.

Dies setzen wir in die ursprüngliche Linearkombination (\*) der ersten  $k$  Eigenvektoren ein und erhalten:

$$\cdots \Rightarrow c_k \vec{v}_k = \vec{0}$$

Da nun aber  $\vec{v}_k \neq \vec{0}$  gilt, folgt auch, dass  $c_k = 0$ .

Somit sind in (\*) sämtliche Koeffizienten  $c_1, \dots, c_k$  gleich 0, und damit sind die Vektoren  $\vec{v}_1, \dots, \vec{v}_k$  ebenfalls linear unabhängig.

Da der Vektor  $\vec{v}_1$  für sich betrachtet linear unabhängig ist, so sind dies dann auch  $\vec{v}_1$  und  $\vec{v}_2$ . Dann sind aber auch  $\vec{v}_1, \vec{v}_2, \vec{v}_3$  linear unabhängig. In dieser Weise folgt nach  $(m-1)$  Schritten, dass auch die Vektoren  $\vec{v}_1, \dots, \vec{v}_m$  linear unabhängig sind. ■

### Beweis zu Satz 9.26 (Determinante und Spur ähnlicher Matrizen) auf Seite 265:

Mit  $B = S^{-1}AS$  gilt nach dem Determinantenproduktsatz (8.27) und Satz 9.3 über die Determinante der inversen Matrix:

$$\det B = \det (S^{-1}AS) = (\det S^{-1})(\det A)(\det S) = \frac{1}{\det S^{-1}}(\det A)(\det S) = \det A$$

Für die Invarianz der Spur verwenden wir das Matrixprodukt (Definition 7.8). Alle Summen sind auszuführen von 1 bis  $n$ :

$$\begin{aligned} \sum_j B_{j,j} &= \sum_j (S^{-1}AS)_{j,j} = \sum_{j,k,l} (S^{-1})_{j,k} A_{k,l} S_{l,j} = \sum_{j,k,l} S_{l,j} (S^{-1})_{j,k} A_{k,l} \\ &= \sum_{k,l} \underbrace{\left( \sum_j S_{l,j} (S^{-1})_{j,k} \right)}_{(S^{-1}S)_{l,k}} A_{k,l} = \sum_{k,l} (S^{-1}S)_{l,k} A_{k,l} = \sum_l \underbrace{\left( \sum_k (\mathbb{1}_n)_{l,k} A_{k,l} \right)}_{(\mathbb{1}_n \cdot A)_{l,l}} = \sum_l A_{l,l} \end{aligned}$$

Damit ist alles gezeigt. ■

**Beweis zu Satz 9.27 (Spektren ähnlicher Matrizen) auf Seite 266:**

Sei  $\lambda$  ein Eigenwert von  $A$  und  $\vec{v}$  ein Eigenvektor von  $A$  zum Eigenwert  $\lambda$ . Dann gilt die Eigenwertgleichung

$$A\vec{v} = \lambda\vec{v}$$

Wir setzen nun ein, dass  $A = SBS^{-1}$ :

$$\dots \Leftrightarrow SBS^{-1}\vec{v} = \lambda\vec{v}$$

Da die Matrix  $S$  invertierbar ist (und ebenso ihre Inverse  $S^{-1}$ ), dürfen wir diese Gleichung mit  $S^{-1}$  von links multiplizieren:

$$\dots \Leftrightarrow S^{-1} \cdot SBS^{-1}\vec{v} = S^{-1} \cdot (\lambda\vec{v}) \Leftrightarrow BS^{-1}\vec{v} = \lambda S^{-1}\vec{v} \Leftrightarrow B(S^{-1}\vec{v}) = \lambda(S^{-1}\vec{v})$$

Aber dann ist die Eigenwertgleichung für  $A$  äquivalent zu einer Eigenwertgleichung für  $B$ , und  $S^{-1}\vec{v}$  ist ein Eigenvektor von  $B$  zum selben Eigenwert  $\lambda$ .

Auch die charakteristischen Polynome von  $A$  und  $B$  sind gleich. Es ist

$$\chi_B(\lambda) = \det(B - \lambda \mathbb{1}_n) = \det(S^{-1}AS - \lambda \mathbb{1}_n)$$

Wir verwenden den Determinanten-Produktsatz (8.27) in umgekehrter Leserichtung, um einen Faktor

$$1 = (\det S) \cdot (\det S^{-1})$$

zu ergänzen:

$$\begin{aligned} \chi_B(\lambda) &= \dots = (\det S) \cdot \left( \det(S^{-1}AS - \lambda \mathbb{1}_n) \right) \cdot (\det S^{-1}) \\ &= \det(S \cdot S^{-1}AS \cdot S^{-1} - \lambda S \cdot \mathbb{1}_n \cdot S^{-1}) = \det(A - \lambda \mathbb{1}_n) \\ &= \chi_A(\lambda) \end{aligned}$$

Damit ist alles gezeigt. ■

**Beweis zu Satz 9.30 (Spektralsatz) auf Seite 269:**

Wir benötigen zunächst die Tatsache, dass wir das kanonische Skalarprodukt zweier Vektoren als Matrixprodukt schreiben können (siehe Unterabschnitt 7.4.3 und Satz 7.14 (Eigenschaften transponierter Matrizen)). Speziell:

$$\vec{v} \bullet (A\vec{w}) = \vec{v}^T \cdot A \cdot \vec{w} = (A^T \vec{v})^T \cdot \vec{w}$$

Ist außerdem  $A$  eine symmetrische Matrix, so gilt weiterhin:

$$\vec{v} \bullet (A\vec{w}) = \dots = (A\vec{v})^T \cdot \vec{w} = (A\vec{v}) \bullet \vec{w}$$

Bei symmetrischer Matrix  $A$  ist es also möglich, die Multiplikation von  $A$  vom einen auf den anderen Vektor umzuwälzen, ohne dabei den Wert des Skalarprodukts zu ändern.

Seien  $\lambda_1, \dots, \lambda_n$  die Eigenwerte von  $A$  (mit Mehrfachnennung), und  $\vec{v}_1, \dots, \vec{v}_n$  eine Auswahl zugehöriger Eigenvektoren (in gleicher Reihenfolge). Es gilt also für  $j \in \{1, \dots, n\}$ :

$$A\vec{v}_j = \lambda_j \vec{v}_j$$

Wir konstruieren nun schrittweise eine Orthonormalbasis des  $\mathbb{R}^n$  mit Vektoren  $\vec{u}_1, \dots, \vec{u}_n$ , die ebenfalls (in dieser Reihenfolge) Eigenvektoren von  $A$  zu den Eigenwerten  $\lambda_1, \dots, \lambda_n$  sind.

Der erste solche Vektor ist schnell konstruiert:

$$\vec{u}_1 := \frac{\vec{v}_1}{\sqrt{\vec{v}_1 \bullet \vec{v}_1}}$$

Es handelt sich um eine skalierte (nämlich: normierte) Version von  $\vec{v}_1$  und ist also ein Eigenvektor zum Eigenwert  $\lambda_1$ .

Nun erinnern wir uns, dass das Skalarprodukt von  $\vec{v}_2$  mit  $\vec{u}_1$  den zu  $\vec{u}_1$  parallelen Anteil von  $\vec{v}_2$  beschreibt, da  $\vec{u}_1$  normiert ist (siehe dazu Unterabschnitt 6.2.2, S. 160ff.). Aber dann gilt mit

$$\vec{v}'_2 := \vec{v}_2 - \underbrace{(\vec{v}_2 \bullet \vec{u}_1)}_{=: c_{2,1}} \vec{u}_1,$$

dass  $\vec{v}'_2$  orthogonal zu  $\vec{u}_1$  ist:

$$\vec{v}'_2 \bullet \vec{u}_1 = (\vec{v}_2 - c_{2,1}\vec{u}_1) \bullet \vec{u}_1 = \vec{v}_2 \bullet \vec{u}_1 - c_{2,1} \underbrace{\vec{u}_1 \bullet \vec{u}_1}_{=1} = c_{2,1} - c_{2,1} = 0$$

Weiterhin sei

$$W_1 := \{\vec{w} \in \mathbb{R}^n \mid \vec{u}_1 \bullet \vec{w} = 0\}$$

der *Orthogonalraum* von  $\vec{u}_1$ .

Für sämtliche  $\vec{w} \in W_1$  gilt mit obiger Tatsache für das Skalarprodukt:

$$\vec{u}_1 \bullet (A\vec{w}) = (A\vec{u}_1) \bullet \vec{w} = \lambda_1 \cdot \vec{u}_1 \bullet \vec{w} = 0$$

Falls also  $\vec{w}$  aus  $W_1$  ist, so gilt dies auch für  $A\vec{w}$ .

Nun hatten wir  $\vec{v}'_2$  absichtlich so konstruiert, dass der zu  $\vec{u}_1$  parallele Anteil entfernt wird, und es ist  $\vec{v}'_2 \in W_1$ . Aber dann ist, wie gerade gezeigt, auch  $A\vec{v}'_2$  aus  $W_1$ . Wir schreiben dies an und verwenden die Eigenwertgleichungen für  $\vec{v}_2$  sowie  $\vec{u}_1$ :

$$A\vec{v}'_2 = A(\vec{v}_2 - c_{2,1}\vec{u}_1) = \lambda_2\vec{v}_2 - c_{2,1}\lambda_1\vec{u}_1 = \lambda_2 \underbrace{(\vec{v}_2 + c_{2,1}\vec{u}_1)}_{=\vec{v}_2} - c_{2,1}\lambda_1\vec{u}_1 = \lambda_2\vec{v}'_2 + c_{2,1}(\lambda_2 - \lambda_1)\vec{u}_1$$

Da aber  $A\vec{v}'_2$  aus  $W_1$  ist, muss der zu  $\vec{u}_1$  parallele Anteil verschwinden; dies bedeutet:

$$c_{2,1}(\lambda_2 - \lambda_1) = 0$$

Falls nun  $\lambda_2 = \lambda_1$ , so sind  $\vec{v}_2$  und  $\vec{u}_1$  aus dem gleichen Eigenraum. Dann ist  $\vec{v}'_2$  eine Linearkombination *innerhalb* dieses Eigenraums; es ist orthogonal zu  $\vec{u}_1$  und ein Eigenvektor von  $A$  zum betreffenden Eigenwert.

Falls aber  $\lambda_2 \neq \lambda_1$  muss  $c_{2,1} = 0$  gelten. Aber dann war zuvor bei der Konstruktion von  $\vec{v}'_2$  nichts zu tun, da  $\vec{v}_2$  bereits orthogonal zu  $\vec{u}_1$  war! Hier wäre also  $\vec{v}'_2 = \vec{v}_2$ .

In beiden Fällen gilt jedoch:

$$A\vec{v}'_2 = \lambda_2\vec{v}'_2$$

Wir skalieren  $\vec{v}'_2$  und erhalten einen normierten Eigenvektor  $\vec{u}_2$ , welcher orthogonal zu  $\vec{u}_1$  ist:

$$\vec{u}_2 := \frac{\vec{v}'_2}{\sqrt{\vec{v}'_2 \bullet \vec{v}'_2}}$$

Falls  $n > 2$ , gehen wir nun analog vor, indem wir aus  $\vec{v}_3$  (dem Eigenvektor zu  $\lambda_3$ ) die Anteile parallel zu  $\vec{u}_1, \vec{u}_2$  heraus projizieren:

$$\vec{v}'_3 := \vec{v}_3 - \underbrace{(\vec{v}_3 \bullet \vec{u}_1)}_{=:c_{3,1}} \vec{u}_1 - \underbrace{(\vec{v}_3 \bullet \vec{u}_2)}_{=:c_{3,2}} \vec{u}_2$$

Auch hier ist  $\vec{v}'_3 \bullet \vec{u}_1 = 0$ , aber auch  $\vec{v}'_3 \bullet \vec{u}_2 = 0$ . Wir definieren einen weiteren Orthogonalraum:

$$W_2 := \{\vec{w} \in W_1 \mid \vec{u}_2 \bullet \vec{w} = 0\}$$

Dieser Raum enthält alle Vektoren, die sowohl zu  $\vec{u}_1$  als auch zu  $\vec{u}_2$  orthogonal sind; der neu definierte Vektor  $\vec{v}'_3$  ist aus  $W_2$ .

Wir können auch hier zeigen, dass für alle  $\vec{w} \in W_2$  gilt, dass auch  $A\vec{w}$  in  $W_2$  liegt. Dann schreiben wir  $A\vec{v}'_3$  an, nutzen das Eigenwertproblem für  $\vec{v}_3, \vec{u}_1$  und  $\vec{u}_2$  und erhalten drei Beiträge: zum einen  $\lambda_3\vec{v}'_3$ , und je einen Beitrag parallel zu  $\vec{u}_1$  und  $\vec{u}_2$ . Für die letzteren beiden Beiträge müssen beide Skalierungsfaktoren verschwinden, da  $A\vec{v}'_3 \in W_2$ . Also ist  $\vec{v}'_3$  ein Eigenvektor von  $A$  zu  $\lambda_3$ .

Die beiden resultierenden Bedingungen lassen sich wie eben erfüllen, indem  $\vec{v}'_3$  zum Eigenraum für  $\lambda_1$  und/oder zum Eigenraum für  $\lambda_2$  gehört. Oder aber der zugehörige Projektionsfaktor ist null; falls  $\lambda_3 \neq \lambda_1$  und  $\lambda_3 \neq \lambda_2$ , war der ursprüngliche Vektor  $\vec{v}_3$  schon Teil des entsprechenden Orthogonalraums  $W_2$ .

In jedem Fall ist jedoch  $\vec{u}_3$  durch Normieren von  $\vec{v}'_3$  konstruierbar und orthogonal zu sowohl  $\vec{u}_1$  und  $\vec{u}_2$ .

Setzen wir diesen Prozess iterativ fort, so erhalten wir nach und nach die gesamte Liste von orthonormierten Eigenvektoren  $\vec{u}_1, \dots, \vec{u}_n$  zu den Eigenwerten  $\lambda_1, \dots, \lambda_n$  in dieser Reihenfolge.

Da sich so  $n$  linear unabhängige Eigenvektoren ergeben, muss jeder Eigenraum maximale Dimension besessen haben, denn sonst könnte er nicht die nötige Zahl an orthogonalen Eigenvektoren beisteuern, die zum Erreichen der vollen Menge  $\{\vec{u}_1, \dots, \vec{u}_n\}$  nötig wäre (die geometrischen Vielfachheiten der Eigenwerte sind ja nach Satz 9.21 nach oben durch die algebraischen Vielfachheiten beschränkt).

Damit ist die ganze zweite Aussage des Spektralsatzes gezeigt. ■

# Literaturverzeichnis

- [1] G. Fischer and B. Springborn. *Lineare Algebra*. Springer, Heidelberg, 2020.
- [2] D. W. Hoffmann. *Grundlagen der technischen Informatik*. Hanser, München, 2010.
- [3] D. W. Hoffmann. *Einführung in die Informations- und Codierungstheorie*. Springer, Heidelberg, 2014.
- [4] C. Karpfinger and K. Meyberg. *Algebra*. Springer, Heidelberg, 2021.
- [5] G. Teschl and S. Teschl. *Mathematik für Informatiker, Band 1: Diskrete Mathematik und lineare Algebra*. Springer, Berlin, 2008.