

Skriptum zur Vorlesung (v7)
Analysis und Lineare Algebra

Dr. Hartwig Bosse

 Dieses Skript ist noch in Bearbeitung, sollten sich Fehler eingeschlichen haben, kontaktieren Sie bitte den Autor.

Versionskontrolle:

v2.0 Eingefügt: Kapitel zu Vektorräumen.

v3.0 Eingefügt: Kapitel zu Skalarprodukten.

v4.0 Eingefügt: Kapitel zu orthogonalen Matrizen.

v4.1 Eingefügt: (unvollständiges) Kapitel zu Determinanten.

v5 Eingefügt: Kapitel linearen Gleichungssystemen & Taylorreihen.

v6 Eingefügt: Kapitel zu komplexen Zahlen & korrektoren der Definition eines Taylorpolynoms.

v7 Eingefügt: Korollar 10.7 zu Taylorreihen.

Inhaltsverzeichnis

1	Grundbegriffe	5
1.1	Grundbegriffe	5
2	Grundlagen	7
2.1	Mathematische Logik: Aussagen und Logische Quantoren	7
2.1.1	Negieren von Aussagen	8
2.2	Mengen	9
2.2.1	Potenzmenge	13
3	Vollständige Induktion	15
3.1	Idee der vollständigen Induktion	15
3.1.1	Vollständige Induktion als Verfahren	16
3.1.2	Typische formale Fehler	19
3.1.3	Fehlerhafte Induktion	20
3.2	(Induktive Argumentation)	21
4	Gruppen	23
4.1	Gruppen: Abstraktes Rechnen mit einem Operator	23
4.1.1	Axiomatische Definition einer Gruppe	23
4.1.2	Rechenregeln in Gruppen	25
4.1.3	Isomorphe Gruppen	27
4.1.4	Aufgaben	28
4.2	Die Gruppenordnung	29
5	Vektorräume	33
5.1	Die Vektorräume \mathbb{R}^n	33
5.2	Allgemeine Vektorräume	36
5.3	Basen und Dimension	39
5.3.1	Basis	42
5.4	Unterräume	46
6	Lineare Abbildungen	49
6.1	Matrizen	49
6.2	Rechnen mit Matrizen	50
6.3	Matrix-Matrix-Multiplikation	53
6.4	Lineare Abbildungen vs. Matrix-Abbildungen	55
6.5	Dimensionssatz	57
6.6	Lineare Selbstabbildungen: Die Welt der quadratischen Matrizen	59

7	Das Skalarprodukt	61
7.1	Das Skalarprodukt und die Euklidische Norm im \mathbb{R}^n	61
7.2	Orthonormalbasen	66
8	Orthogonale Abbildungen	71
8.1	Die Determinante	74
8.1.1	Übersicht	74
8.1.2	Berechnung und Rechenregeln	77
8.1.3	Rechenregeln	79
8.1.4	Rechentricks	80
8.2	Lineare Gleichungssysteme	82
8.2.1	Interpretation eines Tableaus mit Nullzeilen	84
8.2.2	Interpretation eines Tableaus mit Sprüngen	85
9	Folgen und Reihen	87
9.0.3	Folgen	87
9.0.4	Berechnen von Grenzwerten	89
9.0.5	Reihen	91
9.0.6	Wichtige Reihen	94
10	Taylorreihen	97
10.1	Voraussetzungen	97
10.1.1	Potenzreihen	98
10.1.2	Einsetzen von Zahlen in Potenzreihen	99
10.2	Die Taylorreihe einer Funktion	99
10.2.1	Die allgemeine Form der Taylorreihe :	99
10.2.2	Eigenschaften der Taylorreihe	100
10.3	Wichtige Taylorreihen	102
10.4	Das Taylorpolynom als Approximation einer Funktion	103
10.4.1	Die allgemeine Form des Taylorpolynoms :	103
10.4.2	Restgliedabschätzung	103
11	Komplexe Zahlen	107
11.1	Warum imaginäre Zahlen so heißen wie sie heißen	107
11.2	Definition	107
11.2.1	Die imaginäre Einheit	108
11.2.2	Rechnen mit komplexen Zahlen	109
11.3	Komplexe Zahlen als kartesische Vektoren	110
11.4	Konjugiert komplexe Zahlen und Division	111
11.4.1	Verwendung	111
11.5	Polardarstellung komplexer Zahlen	112
11.6	Multiplizieren und Dividieren in Polarkoordinaten	114
11.6.1	Einheitswurzeln	116

Kapitel 1

Grundbegriffe

1.1 Grundbegriffe

Dieser Abschnitt fasst einige Konzepte zusammen, die als bereits bekannt vorausgesetzt werden:

Zahlen

Die Mengen der

- *natürlichen Zahlen* $\mathbb{N} := \{0, 1, 2, 3, \dots\}$
- der *ganzen Zahlen* $\mathbb{Z} := \{0, 1, -1, 2, -2, \dots\}$
- der *rationalen Zahlen* $\mathbb{Q} := \{\frac{a}{b} : a \in \mathbb{Z}, b \in \mathbb{N} \setminus \{0\}\}$
- der *reellen Zahlen* \mathbb{R} .

Wir verwenden die folgenden Abkürzungen für Summen und Produkte:

Summen- und Produktzeichen

Für eine Funktion $f : \mathbb{N} \rightarrow \mathbb{R}$ und Zahlen $n, m \in \mathbb{N}$ mit $m \leq n$ ist

- die Summe der Terme der Form $f(k)$ mit $k \in \mathbb{N}$ und $m \leq k \leq n$ definiert als:

$$\sum_{k=m}^n f(k) := 0 + f(m) + f(m+1) + f(m+2) + \dots + f(n)$$

- das Produkt der Terme der Form $f(k)$ mit $k \in \mathbb{N}$ und $m \leq k \leq n$ definiert als:

$$\prod_{k=m}^n f(k) := 1 \cdot f(m) \cdot f(m+1) \cdot f(m+2) \cdot \dots \cdot f(n)$$

Ist in den Summationsgrenzen (bzw. Produktgrenzen) die obere Grenze n echt kleiner als die untere Grenze m , so ist die Summe (bzw. das Produkt) leer. Die Leere Summe wird interpretiert als **0** und das leere Produkt wird interpretiert als **1**.

Beispiel 1.1

Es gelten also :

$$\sum_{k=3}^5 k^2 = 3^2 + 4^2 + 5^2 \quad \text{und} \quad \sum_{k=5}^3 k^2 = 0$$

Abbildungen

Eine Abbildung (oder auch Funktion) $f : D \rightarrow B$, $x \mapsto f(x)$ bildet Werte aus dem Definitionsbereich D in den Bildbereich B ab. Jedem Element $x \in D$ wird durch f genau ein *Bild* $f(x) \in B$ zugeordnet.

Gilt $f(x) = y$ für ein $y \in B$, so nennt man x das *Urbild* von y . Bemerke: Jedes $x \in D$ besitzt ein Bild $f(x)$, aber nicht jedes Element $y \in B$ muss ein *Urbild* besitzen.

Eine Funktion heißt

- *injektiv*, wenn je zwei verschiedene $x, x' \in D$ auch verschiedene Bilder besitzen, d.h. wenn gilt:

$$x \neq x' \quad \Rightarrow \quad f(x) \neq f(x')$$

- *surjektiv*, wenn jeder Bildpunkt $y \in B$ auch ein Urbild $x \in D$ besitzt mit $y = f(x)$.
- *bijektiv*, wenn F injektiv und surjektiv ist.

Injektiv: Die Definition für “injektiv” kann man sich anschaulich vorstellen als: Die Definitionsmenge D wird in die Bildmenge “injiziert”, d.h. man findet jedes $x \in D$ einen *eigenen* Funktionswert $y \in B$ vor, der einmal x war.

Surjektiv: Eine surjektive Funktion dagegen “deckt die ganzen Bildmenge ab”, jeder Bildpunkt wird bei einer surjektiven Funktion auch tatsächlich angenommen. Dies ist nicht selbstverständlich:

Das Wort “Bildmenge” klingt zwar wie “die Sammlung aller Bilder $f(x)$ ”. Tatsächlich kann die Bildmenge aber auch einfach nur eine “grobe Schätzung” sein, wie im Beispiel $g : \mathbb{R} \rightarrow \mathbb{R}$ mit $g(x) := x^2$ hier sind die Werte von g stets nicht-negativ, d.h. man “erreicht” mit g nicht die ganze Bildmenge \mathbb{R} .

Bijektiv: Eine bijektive Abbildung stellt eine Eins-zu-Eins-Relation zwischen Definitionsmenge D und Bildmenge B her. D.h. jedes $x \in D$ hat zum einen “sein eigenes(!)” Bild $f(x) \in B$, aber auch jeder Punkt $\tilde{y} \in B$ hat genau(!) ein Urbild $\tilde{x} \in D$. Dies liefert:

Sind D, B endliche Mengen und sind B, D über eine bijektive Abbildung $f : D \rightarrow B$ verknüpft, so haben D und B gleich viele Elemente! Der Begriff einer bijektiven Abbildung ist in der Mathematik also beim Zählen von Dingen von zentraler Bedeutung.

Kapitel 2

Grundlagen

2.1 Mathematische Logik: Aussagen und Logische Quantoren

Unter einer *mathematischen Aussage* versteht man eine mathematische Formel, der ein Wahrheitswert “wahr” oder “falsch” zugewiesen werden kann.

- Der Ausdruck “ $x^2 - 2x + 1$ ” ist *keine* mathematische Aussage sondern nur ein mathematischer Term.
- Der Ausdruck “ $x^2 - 2x + 1 = 0$ ” ist eine mathematische Aussage (die je nach Wert von x wahr oder falsch ist).
- Der Ausdruck “ $1 = 0$ ” ist eine mathematische Aussage, die falsch ist.
- Der Ausdruck “4 ist eine Quadratzahl” ist eine mathematische Aussage, die richtig ist.
- Die Goldbach-Vermutung “*Jede gerade natürliche Zahl größer als 2 kann als Summe zweier Primzahlen geschrieben werden.*” ist eine Mathematische Aussage von der bisher nicht klar ist, ob sie wahr oder falsch ist.

Wie Aussagen im “normalen” Leben, muss jede mathematische Aussage ein Verb enthalten. Diese Verben stecken oft in logischen Quantoren oder logischen Operatoren, die im Grunde Abkürzungen für Textbausteine sind. In den obigen Beispielen steckt das Verb an einigen Stellen in dem Operator “=”, den man als “... ist gleich...” liest.

Wir verwenden die folgenden Quantoren und Operatoren:

logische Operatoren

Symbol	Name	Zugehörige Formulierung	Beispiel
\neg	Negation	Es gilt nicht ...	$\neg (3 = 4)$
\vee	Oder	Es gilt ...oder ...	$(n \geq 2) \vee (n \leq 2)$
$\overset{\circ}{\vee}$	Exklusives Oder	Es gilt entweder ...oder ...	$(n \geq 2) \overset{\circ}{\vee} (n < 2)$
\wedge	Und	Es gilt ...und ...	$(0 \leq 1) \wedge (-1 \leq 0)$

Rechenregeln

Es seien $\mathcal{A}, \mathcal{B}, \mathcal{C}$ mathematische Aussagen. Dann gelten:

L1) Symmetrie	$(\mathcal{A} \vee \mathcal{B}) \Leftrightarrow (\mathcal{B} \vee \mathcal{A})$ $(\mathcal{A} \wedge \mathcal{B}) \Leftrightarrow (\mathcal{B} \wedge \mathcal{A})$
L2) Assoziativität	$(\mathcal{A} \vee \mathcal{B}) \vee \mathcal{C} \Leftrightarrow \mathcal{A} \vee (\mathcal{B} \vee \mathcal{C})$ $(\mathcal{A} \wedge \mathcal{B}) \wedge \mathcal{C} \Leftrightarrow \mathcal{A} \wedge (\mathcal{B} \wedge \mathcal{C})$
L3) Distributivgesetze	$(\mathcal{A} \vee \mathcal{B}) \wedge \mathcal{C} \Leftrightarrow (\mathcal{A} \wedge \mathcal{C}) \vee (\mathcal{B} \wedge \mathcal{C})$ $(\mathcal{A} \wedge \mathcal{B}) \vee \mathcal{C} \Leftrightarrow (\mathcal{A} \vee \mathcal{C}) \wedge (\mathcal{B} \vee \mathcal{C})$

logische Quantoren

Symbol	Name	Zugehörige Formulierung	Beispiel
\forall	All-Quantor	Für alle ... gilt	$\forall n \in \mathbb{N} : n \geq 0$
\exists	Existenz-Quantor	Es existiert (mindestens) ein ...	$\exists n \in \mathbb{N} : n \geq 5$
$\exists!$		Es existiert genau ein ...	$\exists! n \in \mathbb{N} : n^2 = 25$
\nexists		Es existiert kein ...	$\nexists n \in \mathbb{N} : n < 0$

2.1.1 Negieren von Aussagen

Negiert man eine Aussage, die mit einem All- oder Existenzquantor beginnen, so taucht in der negierten Aussage stets der andere Quantor auf.

Dies ist wichtig für den Beweis von Aussagen durch Widerspruch, hier wird in der einleitenden Widerspruchsanahme die Originalaussage negiert.

Lemma 2.1

Es sei M eine Menge und es sei $\mathcal{A}(x)$ eine Aussage, die in Abhängigkeit von einem Input $x \in M$ formuliert ist. (z.B. $M = \mathbb{N}$ und $\mathcal{A}(x) = \text{“Die Zahl } x \text{ ist eine Primzahl”}$)

Dann gelten die folgenden Regeln für die Negation:

$$\neg[\forall x \in M : \mathcal{A}(x)] \Leftrightarrow [\exists x \in M : \neg\mathcal{A}(x)]$$

$$\neg[\exists x \in M : \mathcal{A}(x)] \Leftrightarrow [\forall x \in M : \neg\mathcal{A}(x)]$$

Beispiel 2.2 (Negation einer Aussage mit Allquantor)

$$\neg[\forall n \in \mathbb{N} : n \text{ ist Primzahl}] \Leftrightarrow [\exists n \in \mathbb{N} : n \text{ ist keine Primzahl}]$$

Sprachlich:

Die Aussage *“Nicht jede Zahl in \mathbb{N} ist eine Primzahl”*

ist äquivalent zu *“Es gibt eine Zahl in \mathbb{N} , die nicht eine Primzahl ist”*

Beispiel 2.3 (Negation einer Aussage mit Existenzquantor)

$$\neg[\exists n \in \mathbb{N} : n \text{ ist negativ}] \quad \Leftrightarrow \quad [\forall n \in \mathbb{N} : n \text{ ist nicht negativ}]$$

Sprachlich:

Die Aussage “*Es gilt nicht: Es gibt eine Zahl n in \mathbb{N} , die negativ ist.*”

ist äquivalent zu “*Für jede Zahl n in \mathbb{N} gilt: n ist nicht negativ.*”

Beispiel 2.4 (Iterative Negation)

Die Aussage aus Lemma 2.1 kann man iterativ anwenden:

$$\neg[\forall x \in \mathbb{N} : \exists y \in \mathbb{N} : y < x] \quad (\star)$$

$$\Leftrightarrow [\exists x \in \mathbb{N} : \neg[\exists y \in \mathbb{N} : y < x]]$$

$$\Leftrightarrow [\exists x \in \mathbb{N} : [\forall y \in \mathbb{N} : \neg(y < x)]] \quad (\star\star)$$

$$\Leftrightarrow [\exists x \in \mathbb{N} : [\forall y \in \mathbb{N} : (y \geq x)]]$$

Sprachlich: Die Aussage (\star) ist äquivalent zu $(\star\star)$:

(\star) “*Es gilt nicht: Für alle x in \mathbb{N} gibt es ein y in \mathbb{N} , das echt kleiner als x ist.*”

$(\star\star)$ “*Es gibt ein x in \mathbb{N} , so dass für jedes y in \mathbb{N} gilt: y ist nicht kleiner als x .*”

Die Aussagen sind beide wahr, allerdings ist $(\star\star)$ leichter zu beweisen:

Das Besagte x aus $(\star\star)$ ist $x = 0$, denn für jedes $y \in \mathbb{N}$ gilt: $y \geq 0$ bzw. $\neg(y < 0)$.

2.2 Mengen

Den Begriff der Menge kann hier aufgrund einiger Komplikationen in den Details nicht im strengen mathematischen Sinne sauber definiert werden. Hier werden wir Mengen nur als nützliche Hilfsmittel verwenden und in einem solchen Umfeld ist es nützlich und üblich, die Mengendefinition von Georg Cantor (1845-1918), dem Begründer der Mengentheorie, zu zitieren:

Eine Menge ist eine Zusammenfassung bestimmter wohlunterschiedener Objekte unserer Anschauung oder unseres Denkens, welche Elemente der Menge genannt werden, zu einem Ganzen.
[Georg Cantor]

Leider führt diese harmlos scheinende Mengendefinition bereits zu Komplikationen, z.B. zum Paradoxon der Russellschen Antinomie.

Exkurs: Russellsche Antinomie

Definiert man Mengen als “Zusammenfassung unterscheidbarer Objekte” so ergibt sich das folgende Paradoxon:

Die *Menge der Mengen, welche sich nicht selbst enthalten* kann es nicht geben (und doch steht Ihre Definition am Anfang dieses Satzes).

Gäbe es diese Menge, und nennen wir sie A , so stellt sich die Frage: Enthält A sich selbst?^(*)

- Falls ja, so ergibt sich ein Widerspruch:

A enthält die Menge A (weil wir die Frage $(*)$ mit “ja” beantworten).

Damit ist A (als Menge) **keine** jener erlesenen Mengen, die wir unter dem Titel “Mengen die sich nicht selbst enthalten” in A versammelt haben.

D.h. A ist nicht dabei, ergo: A ist (doch) nicht in A enthalten. Ein Widerspruch!

- Falls nein, so ergibt sich ein Widerspruch:

A enthält die Menge A nicht (weil wir die Frage $(*)$ mit “nein” beantworten).

Damit ist A (als Menge) **eine** jener erlesenen Mengen, die wir unter dem Titel “Mengen die sich nicht selbst enthalten” in A versammelt haben.

D.h. A ist dabei, ergo: A ist (doch) in A enthalten. Ein Widerspruch!

Mengen schreibt man mit geschweiften Klammern “{” und “}”, so genannten Mengenklammern. Beispielsweise hat die Menge $\{1, 2, 3\}$ die Elemente 1, 2 und 3.

Definition 2.5

Eine Menge A heißt *endlich*, wenn A nur endlich viele Elemente besitzt. Die Anzahl der Elemente einer endlichen Menge A wird als die *Kardinalität* von A bezeichnet, und mit $|A|$ notiert.

Eine Menge ist *definiert*, wenn angegeben ist, welche Elemente in ihr enthalten sind, z.B.

$$\mathbb{N} := \text{Menge der natürlichen Zahlen}$$

Eine Menge kann definiert werden

- *deskriptiv* durch Angabe einer definierenden Eigenschaft, z.B. $A := \{n \in \mathbb{N} : n \text{ ist gerade}\}$
- *konstruktiv* durch Aufzählung aller in ihr enthaltenen Elemente, z.B. $B := \{2, 4, 6, 8, 10\}$.

Wenn bei Mengen mit unendlich vielen Elementen das Bildungsgesetz klar ist, können auch unendliche Aufzählungen verwendet werden, z.B. $A := \{2, 4, 6, 8, \dots\}$.

Definition 2.6 (Teilmengen und Gleichheit)

Eine Menge A heißt *Teilmenge* einer Menge B (notiert $A \subseteq B$), wenn jedes Element von A auch in B enthalten ist:

$$A \subseteq B \quad :\Leftrightarrow \quad (x \in A \Rightarrow x \in B).$$

Zwei Mengen A, B sind gleich, wenn sie die gleichen Elemente enthalten, d.h.,

$$A = B \quad :\Leftrightarrow \quad (A \subseteq B \text{ und } B \subseteq A).$$

Die Gleichheit zweier Mengen A, B kann entsprechend in zwei Schritten bewiesen werden: Man beweist zunächst $A \subseteq B$ und dann $B \subseteq A$ (bzw. $A \supseteq B$). Zum Beweis für $A \subseteq B$ muss man nach Definition 2.6 die Aussage $\forall x \in A : x \in B$ beweisen, eine solche Aussage beweist man meist nach dem folgenden Schema (z.B. im Beweis von Lemma 2.8):

Zu zeigen: $A \subseteq B$.

Beweis: Wähle $x \in A$ beliebig. Dann gilt für x :

[Bedingungen an x , damit es in A ist]
 \Leftrightarrow [Umformung]
 \Leftrightarrow [Umformung]
 \vdots
 \Leftrightarrow [Bedingungen an x , damit es in B ist]

Deswegen gilt: $x \in B$. Da $x \in A$ beliebig gewählt wurde, muss $x \in B$ für alle $x \in A$ gelten.

Definition 2.7 (Rechnen mit Mengen)

Für Mengen A, B ist

- i) $A \cap B := \{x : x \in A \wedge x \in B\}$ der Schnitt von A und B
- ii) $A \cup B := \{x : x \in A \vee x \in B\}$ die Vereinigung von A und B
- iii) $A \setminus B := \{x : x \in A \wedge x \notin B\}$ die Differenz von A und B
- iv) $A \times B := \{(a, b) : a \in A \wedge b \in B\}$ das kartesische Produkt von A und B

Lemma 2.8

Für beliebige Mengen A, B gilt:

- | | | |
|------|--|-------------------|
| 1.1) | $(A \cap B) = (B \cap A)$ | Kommutativität |
| 1.2) | $(A \cap B) \cap C = A \cap (B \cap C)$ | Assoziativität |
| 1.3) | $(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$ | Distributivgesetz |
| | | |
| 2.1) | $(A \cup B) = (B \cup A)$ | Kommutativität |
| 2.2) | $(A \cup B) \cup C = A \cup (B \cup C)$ | Assoziativität |
| 2.3) | $(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$ | Distributivgesetz |

Im folgenden Beweis stehen die Symbole " \subseteq " und " \supseteq " für folgende Text-überschriften:

" \subseteq " entspricht "Wir Zeigen nun: linke Menge ist enthalten in rechter Menge" bzw.

" \supseteq " entspricht "Wir Zeigen nun: rechte Menge ist enthalten in linker Menge".

Diese Symbole sind also Text-Kürzel und nicht als Mathematische Symbole zu deuten.

Beweis: Wir beweisen exemplarisch die Aussage 1.3) in Lemma 2.8. Zu zeigen ist:

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

“ \subseteq ” Es sei $x \in (A \cap B) \cup C$ beliebig gewählt. Für x gilt dann:

$$\begin{aligned} x &\in (A \cap B) \cup C \\ \stackrel{\text{Def. } \cup}{\Leftrightarrow} & [x \in (A \cap B) \quad \vee \quad x \in C] \quad \text{Das “Oder” } \vee \text{ ist } \textit{kein} \text{ “Entweder oder”.} \\ \stackrel{\text{Def. } \cap}{\Leftrightarrow} & [(x \in A \wedge x \in B) \quad \vee \quad x \in C] \quad (*) \end{aligned}$$

Es gibt nun zwei Fälle:

- Gilt $x \in C$ so folgen sofort $x \in A \cup C$ und $x \in B \cup C$.
- Gilt $x \notin C$ so folgen aus $(*)$ sofort $x \in A$ und $x \in B$. Damit gilt auch $x \in A \cup C$ und $x \in B \cup C$.

In beiden Fällen gilt also $x \in A \cup C$ und $x \in B \cup C$ und damit $x \in (A \cup C) \cap (B \cup C)$.

Da x beliebig gewählt war gilt also allgemein für alle x :

$$x \in (A \cap B) \cup C \quad \Rightarrow \quad x \in (A \cup C) \cap (B \cup C)$$

Damit gilt nach Def. von „ \subseteq “ also $(A \cap B) \cup C \subseteq (A \cup C) \cap (B \cup C)$

“ \supseteq ” Es sei $x \in (A \cup C) \cap (B \cup C)$ beliebig gewählt. Für x gilt dann:

$$\begin{aligned} x &\in (A \cup C) \quad \cap \quad (B \cup C) \\ \stackrel{\text{Def. } \cap}{\Leftrightarrow} & [x \in (A \cup C) \quad \wedge \quad x \in (B \cup C)] \\ \stackrel{\text{Def. } \cup}{\Leftrightarrow} & [(x \in A \vee x \in C) \quad \wedge \quad (x \in B \vee x \in C)] \quad (*) \end{aligned}$$

Es gibt nun zwei Fälle:

- Gilt $x \in C$ so folgt sofort $x \in (A \cap B) \cup C$.
- Gilt $x \notin C$ so folgen wegen $(*)$ dass $x \in A$ und $x \in B$ gelten müssen:
Wegen $(*)$ gelten sowohl $(x \in A \vee x \in C)$ als auch $(x \in B \vee x \in C)$, wegen $x \notin C$ müssen also $x \in A$ und $x \in B$ gelten.

In beiden Fällen gilt also $x \in (A \cap B) \cup C$.

Da x beliebig gewählt war gilt also allgemein für alle x :

$$x \in (A \cap B) \cup C \quad \Leftarrow \quad x \in (A \cup C) \cap (B \cup C)$$

Damit gilt nach Def. von „ \subseteq “ also $(A \cap B) \cup C \supseteq (A \cup C) \cap (B \cup C)$

□

2.2.1 Potenzmenge

Definition 2.10 (Potenzmenge)

Für eine Menge A ist die Potenzmenge $\mathcal{P}(A)$ die Menge aller Teilmengen von A (inkl. \emptyset).

Beispiel 2.10

Für $A := \{1, 2, 3\}$ ist

$$\mathcal{P}(A) = \left\{ \emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\} \right\}$$

Die Potenzmenge einer endlichen Menge A schreibt man auch als 2^A :

Lemma 2.11

Für eine endliche Menge A mit $|A|$ Elementen gilt: $\mathcal{P}(A)$ hat $|\mathcal{P}(A)| = 2^{|A|}$ Elemente.

Den Beweis (z.B. vollständiger Induktion) überlassen wir dem Leser.

Kapitel 3

Vollständige Induktion

Die **vollständige Induktion** ist eine mathematische Beweismethode. Man verwendet sie, um eine Aussage, die von einer natürlichen Zahl abhängt, zu beweisen. Das Beweisverfahren ist von grundlegender Bedeutung für alle Gebiete der Mathematik.

Die Bezeichnung “vollständige Induktion” leitet sich ab von lat. *inductio* (= “Herein-” oder “Hinaufführung”). Der Beiname “vollständig” signalisiert, dass es sich hier um ein anerkanntes deduktives Beweisverfahren handelt – im Gegensatz zur *philosophischen Induktion*, die aus Spezialfällen ein allgemeines Gesetz erschließt und kein exaktes Schlussverfahren ist.

3.1 Idee der vollständigen Induktion

Die Vollständige Induktion beweist Aussagen, die von einer natürlichen Zahl abhängen:

Beispiel 3.1 (Formel, die von $n \in \mathbb{N}$ abhängt)

Gauss wurde als junger Schüler von seinem Lehrer mit einer Fleißaufgabe “ruhiggestellt”, er und seine Mitschüler sollten die Zahlen von 1 bis 100 addieren (die beiden Zahlen 1 und 100 inklusive).

Neben der richtigen Lösung 5050 hatte Gauss recht bald die allgemeine Formel für solche Summen zur Hand – der Anekdote nach hat der Lehrer Gauss für seine “Frechheit” daraufhin den Hosenboden verschliffen.

Die richtige Formel lautet: Ist $n \in \mathbb{N}$ eine natürliche Zahl und $n \geq 1$, so gilt:

$$1 + 2 + 3 + \dots + (n - 1) + n = \frac{n \cdot (n + 1)}{2}$$

Um eine solche Aussage (die von einem $n \in \mathbb{N}$ abhängt) für alle $n \in \mathbb{N}$ zu beweisen, ist die *vollständige Induktion* die richtige Technik: **Statt einen Beweis für alle $n \in \mathbb{N}$ zu führen, führt per vollständiger Induktion zwei Beweise:**

1. Induktions Anfang

Man beweist: Die Aussage gilt für $n = 1$ (Startpunkt des Dominoeffektes).

2. Induktions Schritt

Dann beweist man: Die Aussage wird von k nach $k + 1$ “weitergereicht”:

Wenn die Aussage für **eine** Zahl k mit $k \geq 1$ gilt, *dann* gilt die Aussage auch für deren Nachfolger $k + 1$. (Dominoeffekt)

Nach dem Motto “Klappt’s bei k , so klappts auch bei dem Nachbarn”, haben wir mit dem Ind.Anfang nicht nur gezeigt, dass die Aussage für $n = 1$ gilt, sondern wegen des Induktions-Schrittes gilt sie auch für den Nachbarn $n = 1 + 1 = 2$.

Damit gilt die Aussage auch für $n = 2 + 1$ also für $n = 3$.

... und damit auch für $n = 3 + 1$ also für $n = 4$.

... und damit auch für $n = 4 + 1$ also für $n = 5$.

... und damit auch für $n = 5 + 1$ also für $n = 6$.

Der Induktionsanfang liefert also das folgende:

$n =$	0	1	2	3	...
Aussage gilt	?	✓ (I.Anf.)	?	?	...

Und aus dem Induktionsschritt folgert man:

$n =$	0	1	2	3	...
Aussage gilt	?	✓ (I.Anf.)	✓	✓	...

Der Beweis wird also in zwei Etappen durchgeführt: Als “*Induktionsanfang*” für die kleinste Zahl $m \in \mathbb{N}$, für die man die Aussage zeigen will, (meist 1 oder 0) und als “*Induktionsschritt*”, der aus der Aussage für eine Zahl k die entsprechende Aussage für die nächste Zahl $k + 1$ ableitet.

Damit das ganze Sinn macht, muss man k natürlich größer oder gleich m nehmen.

3.1.1 Vollständige Induktion als Verfahren

Lemma 3.2 (Induktionsprinzip)

Es sei $M \subseteq \mathbb{N}$ eine Menge mit den folgenden Eigenschaften:

1. Es gilt $0 \in M$.
2. Gilt $n \in M$, dann gilt auch $n + 1 \in M$.

Dann gilt: $M = \mathbb{N}$.

Beweis: Angenommen es gilt $M \neq \mathbb{N}$, dann muss wegen $M \subseteq \mathbb{N}$ gelten: $M \subsetneq \mathbb{N}$.

D.h. es gibt Zahlen $n \in \mathbb{N}$ mit $n \notin M$. Wir betrachten nun unter diesen Zahlen die kleinste, d.h. die kleinste Zahl $k \in \mathbb{N}$ mit $k \notin M$.

Wegen 1.) ist $k \neq 0$, und damit ist $k - 1 \in \mathbb{N}$

Weiter ist $k \in \mathbb{N}$ ist die kleinste natürliche Zahl, die nicht in M enthalten ist, also folgt: $k - 1 \in M$.

Aus 2.) und $k - 1 \in \mathbb{N}$ sowie $k - 1 \in M$ folgert man $k \in M$. ein Widerspruch zu $k \notin M$. \square

Bemerkung 3.1 Dieser Beweis ist ein typischer Vertreter vom “Typ des kleinsten Bösewichts”, sie funktioniert gut bei Aussagen der Form $\forall n \in \mathbb{N} : \mathcal{A}(n)$, die einer Zahl in $n \in \mathbb{N}$ abhängen.

Die zu beweisende Aussage $\forall n \in \mathbb{N} : \mathcal{A}(n)$ ist genau dann falsch, wenn es Gegenbeispiele $\tilde{n} \in \mathbb{N}$ gibt, für die $\mathcal{A}(\tilde{n})$ falsch ist. Die Existenz solcher Gegenbeispiele will man widerlegen und tut dies mit einem (etwas ungewöhnlichen) Widerspruch: Man zeigt, dass es unter den Gegenbeispielen kein kleinstes gibt! Gibt es kein kleinstes Gegenbeispiel $\tilde{n} \in \mathbb{N}$, so muss die Menge aller Gegenbeispiele $\tilde{N} \subseteq \mathbb{N}$ leer sein!

Korollar 3.3

Es sei $\mathcal{A}(n)$ eine Aussage, deren Wahrheitswert von einer Zahl $n \in \mathbb{N}$ abhängt. Weiter sei $N \in \mathbb{N}$. Will man beweisen dass $\mathcal{A}(n)$ wahr ist für alle $n \geq N$ in \mathbb{N} so genügt es zu zeigen:

1. $\mathcal{A}(n)$ ist wahr für $n = N$.
2. Wenn $\mathcal{A}(n)$ wahr ist für ein $n \geq N$, dann ist auch stets $\mathcal{A}(n + 1)$ wahr.

Hier ein Beispiel für den Beweis der Summenformel von Gauss mittels vollständiger Induktion.

Beispiel 3.4

Zu zeigen ist: Für jede natürliche Zahl $n \in \mathbb{N}$ mit $n \geq 1$ gilt:

$$1 + 2 + 3 + \dots + (n - 1) + n = \sum_{i=1}^n (i) = \frac{n \cdot (n + 1)}{2}$$

Beweis durch Induktion.

Induktionsanfang (*Anfang des Dominoeffekts*)

Behauptung: Die Aussage gilt für $n = 1$.

Beweis: Es sei $n = 1$.

$$\underbrace{\sum_{i=1}^n i}_{\text{Linke Seite}} = \sum_{i=1}^1 i = 1 = \frac{1 \cdot 2}{2} = \underbrace{\frac{n(n+1)}{2}}_{\text{Rechte Seite}} \quad \checkmark$$

Die Aussage gilt also für $n = 1$.

Induktionsschritt (*Dominoeffekt*)

Zeige: Gilt die Aussage für **ein** $k \in \mathbb{N}$ mit $k \geq 1$, so gilt sie auch für dessen Nachfolger $k + 1$.

Induktions-Annahme: Für **ein** k mit $k \geq 1$ gelte:

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

Induktions-Behauptung: Dann gilt die Aussage auch für $k + 1$:

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

Beweis:

$$\begin{aligned} \sum_{i=0}^{k+1} (i) &= \sum_{i=0}^k (i) + (k + 1) \quad (\text{letzten Summanden abkoppeln}) \\ &\stackrel{\text{I.A.}}{=} \underbrace{\frac{k \cdot (k + 1)}{2}}_{\text{Induktions Annahme benutzen}} + (k + 1) \\ &= \frac{k \cdot (k + 1)}{2} + \frac{2(k + 1)}{2} \quad (\text{Erweitern}) \\ &= \frac{(k + 2) \cdot (k + 1)}{2} \quad (\text{Zusammenfassen: wieviele '(k + 1)' sind es? -(k + 2)-viele!}) \end{aligned}$$

Hier ein weiteres Beispiel für den Einsatz von vollständiger Induktion:

Summiert man die ersten ungeraden Zahlen, so erhält man *anscheinend* stets eine Quadratzahl:

$$\begin{array}{rcccccccl} & & & & & & & & 1 & = & 1 & = & 1^2 \\ & & & & & & & & 3 & +1 & = & 4 & = & 2^2 \\ & & & & & & & & 5 & +3 & +1 & = & 9 & = & 3^2 \\ & & & & & & & & 7 & +5 & +3 & +1 & = & 16 & = & 4^2 \\ & & & & & & & & 9 & +7 & +5 & +3 & +1 & = & 25 & = & 5^2 \end{array}$$

Diesen Umstand wollen wir in allgemeiner Form beweisen (s.u.). Hierzu ein paar Erläuterungen:

Eine *gerade* Zahl lässt sich stets in der Form $2 \cdot n$ schreiben mit $n \in \mathbb{N}$.

Eine *ungerade* Zahl lässt sich stets in der Form $2 \cdot n + 1$ schreiben mit $n \in \mathbb{N}$.

In den obigen fünf Summen gilt: Ist der letzte Summand die Zahl $2n + 1$ so ist der Wert der Summe $(n + 1)^2$. In der letzten Beispielzeile wird zum Beispiel bis $9 = 2 \cdot 4 + 1$ summiert und die Summe ergibt $(4 + 1)^2 = 5^2$.

Achtung: Dies alles ist nur eine Erläuterung der zu beweisenden Formel. Wir haben noch keinen Beweis für den allgemeinen Fall geführt!

Beispiel 3.5

Zeige: Für alle $n \in \mathbb{N}$ gilt: $\sum_{i=0}^n (2i + 1) = (n + 1)^2$.

Die Lösung schreiben wir hier einmal in einer deutlich kürzeren Form, und zwar so so wie die Induktion in der mathematischen Literatur geführt wird:

Das “ k ” haben wir zum Beispiel durch n ersetzt (was den Beweis nicht falsch macht) und auch die Trennung der beiden Beweise (ein Bew. für $n = 1$ und ein Bew. für $k \rightarrow k + 1$) ist nicht mehr so deutlich zu sehen:

Induktionsanfang: $n = 1$

$$\begin{aligned} \sum_{i=0}^1 (2i + 1) &= \underbrace{(2 \cdot 0 + 1)}_1 + \underbrace{(2 \cdot 1 + 1)}_3 \\ &= 1 + 3 = 4 = 2^2 = (1 + 1)^2 = (n + 1)^2 \quad \checkmark \end{aligned}$$

Induktions-Annahme: Für ein n mit $n \geq 1$ gelte: $\sum_{i=0}^n (2i + 1) = (n + 1)^2$.

Induktionsschritt: $n \rightarrow n + 1$ Zu zeigen ist: $\sum_{i=0}^{n+1} (2i + 1) = ((n + 1) + 1)^2$

$$\begin{aligned} \sum_{i=0}^{n+1} (2i + 1) &= \underbrace{\sum_{i=0}^n (2i + 1)}_{\stackrel{\text{I.A.}}{=} (n+1)^2} + (2(n+1) + 1) \quad (\text{letzten Summanden abkoppeln}) \\ &\stackrel{\text{I.A.}}{=} (n+1)^2 + 2(n+1) + 1 \quad (\text{Induktions-Annahme benutzen}) \\ &= ((n+1) + 1)^2 \quad (\text{binomische Formel}) \end{aligned}$$

3.1.2 Typische formale Fehler

Beim Führen eines Beweises per vollständiger Induktion gibt es einen typischen Fehler, den prinzipiell alle Anfänger machen:

Der Beweis per Induktion wird von "echten" Mathematikern oft lax ohne vollständige Angabe der Induktionsannahme formuliert. Dies macht aber den Beweis nicht nur *formal falsch*, sondern auch widersprüchlich und damit für Laien unverständlich. Ein konkretes Beispiel, das das Weglassen der Ind. Annahme zu Problemen führen kann findet sich im folgenden Abschnitt (Seite 20 ff.).

Der folgende Beweis ist falsch geführt und damit ist zwar die Aussage weiterhin wahr, aber der hier geführte Beweis ist ungültig. Der Fehler findet sich im Anwenden der Induktionsvoraussetzung ohne eine Einschränkung auf ein bestimmtes n .

(Anti-)Beispiel 3.6

Zu zeigen:

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

Falscher(!) Beweis durch Induktion.

Induktionsanfang: " $n = 1$ "

Sei $n = 1$. Dann gilt:

$$\underbrace{1 + \dots + n}_{\text{Linke Seite}} = 1 = \frac{1 \cdot 2}{2} = \underbrace{\frac{n(n+1)}{2}}_{\text{Rechte Seite}} \quad \checkmark$$

Die Aussage gilt also für $n = 1$.

Induktionsschritt: " $n \rightarrow (n + 1)$ "

$$\begin{aligned} \text{Linke Seite:} \quad & \underbrace{1 + \dots + n}_{\text{I.A.}} + (n + 1) \quad (\text{letzten Summanden abkoppeln}) \\ & \underline{\underline{\frac{n \cdot (n + 1)}{2}}} + (n + 1) \quad (\text{Induktions Annahme benutzen}) \\ & = \frac{n \cdot (n + 1)}{2} + \frac{2(n + 1)}{2} \quad (\text{Erweitern}) \\ & = \frac{(n + 2) \cdot (n + 1)}{2} \quad (\text{Zusammenfassen}) \end{aligned}$$

Völlig zu recht wird man an diesem Beweis bemängeln, dass im Induktionsschritt die Aussage

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

verwendet wird, und *die soll im Beweis ja erst bewiesen werden!*

Was ist hier falsch gelaufen? Der Beweisführende hat die Mathematischen Quantoren vergessen!

- Im zu zeigen fehlt ein "gilt für alle $n \in \mathbb{N}$ "
- In der Induktions-Annahme fehlt die Einschränkung "gilt für ein $n \in \mathbb{N}$ mit $n \geq 0$ "

Lässt man diese Quantoren weg, so verbleibt bei "zu zeigen" und "induktions-annahme" nur die "nackte Formel" und die ist in beiden Fällen *tatsächlich* und völlig zu recht identisch!

Der obige Beweis läuft also auf einen Kreisschluss hinaus, und die Beweisführung ist deswegen *falsch*.

Eine Anschauung dafür, dass “Induktions Voraussetzung” und “zu zeigen” *mit Quantoren* sehr sehr unterschiedlich sind, kann vielleicht das folgende Beispiel liefern:

- a) Eine Aussage in “zu zeigen” enthält immer einen Allquantor, und lautet damit in etwa wie:
 “zu zeigen: *Alle* Menschen mögen sich.” .
- b) Die Aussage in der Induktions Voraussetzung enthält immer einen Hinweis auf *ein einzelnes Element* und seinen Nachbarn, also in etwa:
 “zu zeigen: Zwei Menschen, die direkt nebeneinander sitzen, mögen sich.”

Während die erste Aussage offensichtlich schwer zu beweisen sein wird, scheint die zweite “deutlich beweisbarer” und sagt auch deutlich weniger aus!

3.1.3 Fehlerhafte Induktion

Wenn man mit den Begriffen innerhalb der Induktion schludrig umgeht, so kann man sich in (scheinbaren) Widersprüchen verhäddern:

(Anti-)Beispiel 3.7

Wir “zeigen” nun, dass alle Pferde die selbe Farbe haben und führen einen Beweis durch Induktion über die Zahl n aller Pferde:

- Im **Induktionsanfang** ist $n = 1$ und die Behauptung ist offensichtlich richtig: ein Pferd hat stets mit sich selbst die selbe Farbe.
- Für den **Induktionsschritt** nehmen wir an, dass es $n + 1$ Pferde p_1, \dots, p_{n+1} gibt.
 - Nach Induktionsvoraussetzung haben die ersten n Pferde p_1, \dots, p_n die selbe Farbe, sagen wir die Farbe von Pferd p_2 .
 - Ebenso, nach Induktionsvoraussetzung haben die n -vielen Pferde p_2, \dots, p_{n+1} die Farbe von p_2 .

Somit haben alle $n + 1$ Pferde p_1, \dots, p_{n+1} die selbe Farbe und wir haben die Behauptung gezeigt.

Was haben wir hier falsch gemacht?

Der Fehler in Beispiel 3.7 liegt im Induktionsschritt – und zwar im Rückgriff auf die (schludrigerweise!) gar nicht erst hingeschriebene Induktions-Annahme:

Wir haben beim Induktions-Anfang lediglich gezeigt, dass die Behauptung für $n = 1$ gilt. Im Induktionsschritt sprechen wir aber plötzlich vom Pferd p_2 ! Auf p_2 können wir uns aber erst beziehen, wenn wir wissen, dass die Aussage auch für $n = 2$ gilt!

... dies wollen wir aber im Induktionsschritt eigentlich (unter anderem) Beweisen!

Dieser Fehler ist *nur deswegen* nicht gleich offensichtlich, weil wir die Induktions-Annahme nicht *explizit* hingeschrieben haben. Richtig lautet die Induktionsannahme:

“Annahme, für ein $k \geq 1$ gelte stets: k beliebig gewählte Pferde haben stets die selbe Farbe. ”

Jetzt können wir uns im Induktionsschritt nicht mehr auf ein zweites Pferd p_2 beziehen, denn falls $k = 1$ ist und $k+1 = 2$, dann kommt in der Induktionsannahme bei $k = 1$ nur das Pferd p_1 vor. Der Induktionsschritt von $k = 1$ auf $k = 2$ geht also schief!

Um den Beweis so zu führen wie er oben steht, müsste man also den Induktions-Anfang auf $n = 2$ legen. Und in der Tat, wenn man zeigen *könnte*, dass *zwei* beliebig gewählte Pferde stets die selbe Farbe haben, dann haben alle Pferde die selbe Farbe.

3.2 (Induktive Argumentation)

Die vollständige Induktion wird nicht nur benutzt um “bloß” Formeln zu beweisen. Man kann damit eigentlich jeden Sachverhalt in dem zählbare Dinge vorkommen beweisen:

Aufgabe 3.2.1 Zeige: Teilt man ein Rechteck durch (verschiedene) Geraden in Teilflächen, so kann man die Teilflächen mit den Farben Schwarz und Weiß *konfliktfrei* färben, d.h. so färben, dass Teilflächen, die an einer *Kante* zusammenstoßen, verschiedene Farben besitzen.

Lösung: Induktion über n , die Anzahl der teilenden Geraden.

Induktionsanfang: $n = 1$

Die Aussage gilt offensichtlich, eine Teilfläche färben wir Weiß, eine Schwarz.

Induktionsvoraussetzung: Die Behauptung gelte für ein $n \geq 1$ (d.h. für n Geraden).

Induktionsschritt: $n \rightarrow n + 1$

Wir betrachten ein Rechteck das von $n + 1$ Geraden g_1, \dots, g_{n+1} zerteilt wird.

Wenn wir die Gerade g_{n+1} entfernen, erhalten ein Rechteck, das durch n Geraden in Teilflächen zerteilt wird. Diese lassen sich laut Induktionsvoraussetzung alle konfliktfrei in Schwarz und Weiß färben.

Wenn wir nun die teilende Gerade wieder hinzufügen, wird das vorherige Rechteck durch g_{n+1} in 2 Teilflächen A und B geteilt, und einige (oder alle) der bereits bestehenden, gefärbten Teilflächen auch. Es entstehen also neue Teilflächen.

Nun “vertauschen” auf A die Färbung: Auf dieser Seite der Geraden g_{n+1} gilt dann: Teilflächen, die ehemals weiß waren, werden nun schwarz, ehemals weiße Teilflächen werden schwarz. Dadurch erhalten wir eine zulässige Färbung des Rechtecks mit $n + 1$ teilenden Geraden:

Zwei Flächen aus dem Bereich A , die an einer Kante zusammenstoßen, taten dies auch, bevor wir die $n + 1$ -te Gerade hinzufügten. Sie waren und sind also verschieden gefärbt.

Gleiches gilt für zwei Flächen aus B .

Ist E eine Fläche aus A und F eine Fläche aus B , die an einer Kante zusammenstoßen, so ist diese Kante offensichtlich ein Teil der Geraden g_{n+1} . Also waren E und F vor Einfügen von g_{n+1} eine einzige Fläche, hatten also vorher die selbe Farbe, sagen wir mal “schweiz”.

Nach Einfügen von g_{n+1} bekommt E die entgegengesetzte Farbe, F dagegen bleibt “schweiz”: E und F sind verschieden gefärbt.

Kapitel 4

Gruppen

4.1 Gruppen: Abstraktes Rechnen mit einem Operator

Eine Gruppe ist die mathematische Abstraktion von Rechnen mit einer einzelnen *umkehrbaren* Operation z.B. “Rechnen mit Plus in \mathbb{Z} ” oder “Rechnen mit Multiplikation in $\mathbb{Q} \setminus \{0\}$ ”.

Eine Gruppe (G, \circ) besteht stets aus einer Menge (oft “ G ” genannt) auf der mit einer einzelnen Operation (oft mit “ \circ ” bezeichnet) gerechnet werden kann. Ein typisches Beispiel ist die Gruppe $(\mathbb{Z}, +)$ der Ganzen Zahlen zusammen mit der Addition.

4.1.1 Axiomatische Definition einer Gruppe

Das Einführen einer Gruppe mittels der Gruppen-Axiome liefert eine sehr knappe Formulierung für eine Menge in der man lineare Gleichungen lösen kann:

Definition 4.1 (Definition: Gruppe)

Eine Gruppe (G, \circ) ist ein Tupel aus

- einer Menge G und
- einer Verknüpfung $\circ : G \times G \rightarrow G$,

so dass gelten:

Abgeschlossenheit G1) $a \circ b \in G \quad \forall a, b \in G$

Assoziativität G2) $a \circ (b \circ c) = (a \circ b) \circ c \quad \forall a, b, c \in G$

Neutrales Element G3) $\exists e \in G : (e \circ a = a \circ e = a \quad \forall a \in G)$

Inverses Element G4) $\forall a \in G : \exists \bar{a} \in G : \bar{a} \circ a = e$

Analyse: Das Neutrale Element e ist dasjenige Element $e \in G$, das jedes Element $a \in G$ unverändert lässt, wenn man e mit a verknüpft. In einer Gruppe mit einer additiven Verknüpfung übernimmt e die Rolle der Null, in einer Gruppe mit einer multiplikativen Verknüpfung übernimmt e die Rolle der 1.

Beispiel 4.1 Das Tupel $(\mathbb{Z}, +)$ aus der Menge \mathbb{Z} zusammen mit der gewöhnlichen Addition bildet eine Gruppe:

- G1) Für alle $m, n \in \mathbb{Z}$ gilt $m + n \in \mathbb{Z}$.
- G2) Für alle $m, n, k \in \mathbb{Z}$ gilt $m + (n + k) = (m + n) + k$.
- G3) Das Neutrale Element ist hier die Zahl $e = 0$, sie erfüllt: $0 + x = x$ für alle $x \in \mathbb{Z}$.
- G4) Das Inverse zu einer Zahl $m \in \mathbb{Z}$ ist die zugehörige Zahl $-m$ mit entgegengesetztem Vorzeichen: Die zu 3 inverse Zahl ist -3 , es gilt $(-3) + 3 = 0 = e$.

Beispiel 4.2 Die Menge $\mathbb{Q} \setminus \{0\}$ bildet zusammen mit der gewöhnlichen Multiplikation eine Gruppe:

- G1) Für alle $x, y \in \mathbb{Q} \setminus \{0\}$ gilt $x \cdot y \in \mathbb{Q} \setminus \{0\}$.
- G3) Das Neutrale Element ist hier die Zahl $e = 1$, sie erfüllt: $1 \cdot x = x$ für alle $x \in \mathbb{Q} \setminus \{0\}$.
- G4) Das Inverse zu einer Zahl $x \in \mathbb{Q} \setminus \{0\}$ ist die zugehörige Zahl $\frac{1}{x} \in \mathbb{Q} \setminus \{0\}$: Die zu 3 Inverse Zahl ist $\frac{1}{3}$, es gilt $3 \cdot \frac{1}{3} = 1 = e$.

Beispiel 4.3 Die Menge $(\{1, 2, 3, 4\}, \odot_5)$ bildet eine Gruppe:

- G1) Für alle $a, b \in \{1, 2, 3, 4\}$ gilt $a \odot_5 b \in \{1, 2, 3, 4\}$, dies kann man der unten angefügten Verknüpfungstabelle entnehmen.
- G3) Das Neutrale Element ist hier die Zahl $e = 1$, sie erfüllt: $1 \odot_5 b = b$ für alle $b \in \{1, 2, 3, 4\}$.
- G4) Das Inverse zu den Zahlen $b \in \{1, 2, 3, 4\}$ liest man aus der Verknüpfungstabelle ab:

$b =$	1	2	3	4
Inverses $\bar{b} =$	1	3	2	4

Verknüpfungstabelle

$a \odot_5 b$	$b =$			
$a =$	1	2	3	4
	1	2	3	4
	2	4	1	3
	3	3	1	4
	4	4	3	2
	4	3	2	1

Auslesen:
Neutrales Element

$a \odot_5 b$	$b =$			
$a =$	1	2	3	4
	1	2	3	4
	2	2	4	1
	3	3	1	4
	4	4	3	2
	4	3	2	1

Auslesen:
Inverse Elemente

$a \odot_5 b$	$b =$			
$a =$	1	2	3	4
	1	1	•	•
	2	•	•	1
	3	•	1	•
	4	•	•	4

4.1.2 Rechenregeln in Gruppen

In jeder Gruppe gelten die folgenden Regeln:

Lemma 4.2

Es sei (G, \circ) eine Gruppe.

1. a) Es gibt in (G, \circ) genau ein neutrales Element e .
b) Es gilt $e \circ a = a \circ e$ für alle $a \in G$.
2. a) Für jedes $a \in G$ gibt es genau ein inverses \bar{a} .
b) Es gilt $a \circ \bar{a} = \bar{a} \circ a$
3. Das inverse zu \bar{a} ist a , d.h. es gilt $\overline{(\bar{a})} = a$.

Beweis: Exemplarisch führen wir hier die Beweise für 1 a) und 1 b) sowie 2 b).

- **Zu 2 b)** Es sei $a \in G$ beliebig.

Nach Axiom G2 hat a ein Inverses \bar{a} mit $\bar{a} \circ a = e$. Das Element \bar{a} hat wiederum ein Inverses $\tilde{a} := \overline{(\bar{a})}$.
Es gilt dann für $a \circ \bar{a}$:

$$a \circ \bar{a} \stackrel{\text{G3}}{=} e \circ (a \circ \bar{a}) \stackrel{\text{G4}}{=} (\tilde{a} \circ \bar{a}) \circ (a \circ \bar{a}) \stackrel{\text{G2}}{=} \tilde{a} \circ \underbrace{(\bar{a} \circ a)}_e \circ \bar{a} \stackrel{\text{G4}}{=} \tilde{a} \circ (e \circ \bar{a}) \stackrel{\text{G3}}{=} \tilde{a} \circ \bar{a} \stackrel{\text{G4}}{=} e$$

Es gilt also $a \circ \bar{a} = e$. Wegen $\bar{a} \circ a = e$ gilt also $\bar{a} \circ a = a \circ \bar{a}$.

- **Zu 1 b)** Es sei $a \in G$ beliebig.

Nach Axiom G2 hat a ein Inverses \bar{a} mit $\bar{a} \circ a = e$. Es gilt dann:

$$a \circ e \stackrel{\text{G3}}{=} a \circ (\bar{a} \circ a) \stackrel{\text{G2}}{=} \underbrace{(a \circ \bar{a})}_{=e} \circ a \stackrel{\text{2b)}}{=} e \circ a$$

- **Zu 1 a)** Es seien $e, \tilde{e} \in G$ zwei neutrale Elemente, d.h. sie erfüllen $e \circ a = a$ und $\boxed{\tilde{e} \circ a = a}^*$ für alle $a \in G$. Dann gilt $\tilde{e} \circ e = e$ (Aussage \star für $a = e$), sowie $\tilde{e} \circ e \stackrel{\text{1b)}}{=} e \circ \tilde{e} \stackrel{\text{G4}}{=} \tilde{e}$. Insgesamt gilt also $\tilde{e} = e$.

□

Lösen von Gleichungen

Die Gruppe ist die kleinste Recheneinheit der Mathematik, in der lineare Gleichungen stets lösbar sind: In einer Gruppe (G, \circ) sind Gleichungen der Form $a \circ x = b$ bei gegebenem $a, b \in G$ lösbar mit einem $x \in G$. Um dies garantieren zu können, muss sicher sein, dass man den Vorgang "a mit x mittels \circ verknüpfen" rückgängig machen kann. Rechnet man zum Beispiel in $\mathbb{Q} \setminus \{0\}$ mit der Verknüpfung \cdot so kann man die Gleichung $2 \cdot x = 7$ durch Multiplikation mit $1/2$ bzw. $0,5 \in \mathbb{Q}$ wie folgt lösen:

$$\begin{aligned} 2 \cdot x &= 7 \\ \Leftrightarrow (0,5) \cdot 2 \cdot x &= 0,5 \cdot 7 \\ \Leftrightarrow x &= 3,5 \end{aligned}$$

Die Zahl $0,5 \in \mathbb{Q} \setminus \{0\}$ nennt man das “multiplikative Inverse zu 2”. Die Zahl $0,5$ kann also das ungeschehen machen, was die Multiplikation mit 2 “angerichtet hat”, denn es gilt: $0,5 \cdot 2 = 1$ und die Zahl 1 benimmt sich bei der Multiplikation neutral.

Lemma 4.3

Es sei (G, \circ) eine Gruppe und $a, b \in G$.

Die Gleichung der Form $a \circ x = b$ hat stets eine Lösung $x \in G$, nämlich $x = \bar{a} \circ b$.

Beweis: In der Gruppe G gibt es ein zu a inverses Element \bar{a} , und Verknüpfung mit \bar{a} “entfernt” a auf der linken Seite der Gleichung. Beachten Sie, dass beim Umformen der Gleichung $A \circ x = b$ alle Gruppenaxiome $G1)$ bis $G4)$ verwendet werden müssen:

$$\begin{aligned} a \circ x &= b \\ \Leftrightarrow \bar{a} \circ (a \circ x) &= \bar{a} \circ b \\ \stackrel{G2}{\Leftrightarrow} (\bar{a} \circ a) \circ x &= \bar{a} \circ b && \text{Assoziativität: G2)} \\ \stackrel{G4}{\Leftrightarrow} e \circ x &= \bar{a} \circ b && \text{Eigenschaften des Inversen: G4)} \\ \stackrel{G3}{\Leftrightarrow} x &= \bar{a} \circ b && \text{Eigenschaften des Neutralen: G3)} \end{aligned}$$

Dass Die Lösung $x = \bar{a} \circ b$ wieder in G ist, liegt an Axiom $G1)$. □

Beispiel 4.4 (Rechnen in der Gruppe $(\mathbb{Q} \setminus \{0\}, \cdot)$)

Eine Gleichung der Form $a \cdot x = b$ mit $a, b \in \mathbb{Q} \setminus \{0\}$ hat immer eine Lösung $x = \frac{1}{a} \cdot b$.

Hier ist $\frac{1}{a}$ das Multiplikativ-Inverse zu $a \in \mathbb{Q} \setminus \{0\}$, das Inverse $\frac{1}{a}$ zu bilden ist immer möglich, da $a \neq 0$ gilt.

Beispiel 4.5 (Rechnen in der Gruppe $(\mathbb{Z}, +)$)

Eine Gleichung der Form $a + x = b$ mit $a, b \in \mathbb{Z}$ hat immer eine Lösung $x = (-a) + b$.

Hier ist $-a$ das Additiv-Inverse zu $a \in \mathbb{Z}$.

(Anti-)Beispiel 4.6 (Rechnen in $(\mathbb{N}, +)$)

Die Menge $(\mathbb{N}, +)$ ist keine Gruppe, und deswegen gibt es Gleichungen der Form $a + x = b$ mit $a, b \in \mathbb{N}$, die keine Lösung in \mathbb{N} besitzen:

Ein Beispiel für eine solche Gleichung ist $5 + x = 0$ die “Lösung” $x = -5$ liegt nicht in \mathbb{N} , d.h. man verlässt beim Lösen der Gleichung die vorgegebene Menge.

(Anti-)Beispiel 4.7 (Rechnen in (\mathbb{Q}, \cdot))

Die Menge (\mathbb{Q}, \cdot) ist keine Gruppe, und deswegen gibt es Gleichungen der Form $a \cdot x = b$ mit $a, b \in \mathbb{Q}$, die keine Lösung in \mathbb{N} besitzen:

Ein Beispiel für eine solche Gleichung ist $0 \cdot x = 5$, diese Gleichung besitzt keine Lösung.

Abelsche Gruppen

Für die meisten bisher betrachteten Gruppen ist die Verknüpfungsreihenfolge in $a \circ b$ gleichgültig, dies ist aber etwas besonderes für Gruppen:

Definition 4.4

Eine Gruppe (G, \circ) heißt abelsch, wenn zusätzlich zu den Gruppenaxiomen G1) bis G4) gilt:

Symmetrie: $G_{\text{symm}} \quad a \circ b = b \circ a \quad \forall a, b \in G$

Abelsche Gruppen sind beispielsweise $(\mathbb{Z}, +)$ und $(\mathbb{Q} \setminus \{0\}, \cdot)$ (s. Beisp.4.1 bis 4.3).

(Anti-)Beispiel 4.5

Die Menge $O(n)$ der orthogonalen Matrizen (s. Kapitel “Lineare Abbildungen”) bildet zusammen mit der Matrix-Multiplikation eine nicht-abelsche Gruppe.

4.1.3 Isomorphe Gruppen

Es ist möglich, ein und die selbe Gruppe auf verschiedene Arten und Weisen zu notieren:

Die Gruppe $(\{0, 1\}, \oplus_2)$ mit $a \oplus_2 b := \text{Rest}(a + b, 2)$ kann man abstrakt auffassen als eine Gruppe mit zwei Elementen, dem neutralen Element $e = 0$ und einem weiteren Element $x = 1$ mit den Rechenregeln $e \oplus_2 x = x \oplus_2 e = x$ und $e \oplus_2 e = x \oplus_2 x = e$.

Beispiel 4.6

Augenscheinlich sind die Gruppen $(\{0, 1\}, \oplus_2)$, $(\{1, 2\}, \odot_3)$ und $(\{\text{Falsch}, \text{Wahr}\}, \overset{\circ}{\vee})$ “strukturell gleich”, wenn man die jeweiligen Verknüpfungstabelle anschaut:

		$b =$	
$a \oplus_2 b$	0	1	
$a =$	0	0	1
	1	1	0

		$b =$	
$a \odot_3 b$	1	2	
$a =$	1	1	2
	2	2	1

		$b =$	
$a \overset{\circ}{\vee} b$	F	W	
$a =$	F	F	W
	F	W	F

Ersetzt man in einer Gleichung $a \oplus_2 b = c$ in $(\{0, 1\}, \oplus_2) \dots$

- jede 0 durch F,
- jede 1 durch W
- und \oplus_2 durch $\overset{\circ}{\vee}$

so erhält man wieder eine korrekte Gleichung.

Diesen Umstand wollen wir mathematisch formalisieren. Um einem Leser mitteilen zu können, dass zwei Gruppen (G, \circ) und $(H, *)$ prinzipiell gleich (also “isomorph¹”) sind, müssen wir dem Leser mitteilen, welche Elemente der beiden Gruppen einander entsprechen:

Hierzu verwenden Mathematiker als Zuordnung eine bijektive Abbildung, d.h. eine Abbildung $f : G \rightarrow H$, die jedem $g \in G$ ein $h \in H$ zuordnet, so dass ...

- je zwei verschiedene $g, \tilde{g} \in G$ auch verschiedene Bilder $f(g) \neq f(\tilde{g}) \in H$ haben.
- es für jedes $h \in H$ ein $g \in G$ gibt mit $f(g) = h$.

¹ “gleich geformt”

Mit dieser Zuordnung f wird nun einem $g \in G$ sein Gegenstück $h := f(g)$ in H zugeordnet, so dass sich h in $(H, *)$ genauso verhält wie g sich in (G, \circ) verhält.

Definition: Isomorphe Gruppen

Zwei Gruppen (G, \circ) und $(H, *)$ heißen *isomorph*, wenn es eine bijektive Abbildung $f : G \rightarrow H$ gibt, so dass gilt:

$$f(g) * f(\tilde{g}) = f(g \circ \tilde{g}) \quad \forall g, \tilde{g} \in G$$

Die Gleichung in der Definition lässt sich wie folgt lesen:

Ergibt $g \circ \tilde{g}$ das Element $c \in G$, so muss für die zu g, \tilde{g} und c zugeordneten Elemente $f(g), f(\tilde{g}), f(c) \in H$ gelten: $f(g) * f(\tilde{g}) = f(c)$. Kurz gesagt, $f(g), f(\tilde{g})$ benehmen sich innerhalb $H = f(G)$ wie g, \tilde{g} innerhalb von G . Es gilt:

$$g \circ \tilde{g} = c \quad \Rightarrow \quad f(g) * f(\tilde{g}) = f(c)$$

Beispiel 4.7

Die Gruppen $(\{0, 1\}, \oplus_2)$ und $(\{1, 2\}, \odot_3)$ sind isomorph.

Mit $f : \{0, 1\} \rightarrow \{1, 2\}$ definiert durch $f(0) := 1$ und $f(1) := 2$ erhält man:

$$\begin{array}{llll} a \oplus_2 b = c \text{ für } a, b, c \in \{0, 1\} & \Rightarrow & f(a) \odot_3 f(b) = f(c) \\ 0 \oplus_2 0 = 0 & \Rightarrow & 1 \odot_3 1 = 1 \\ 0 \oplus_2 1 = 1 & \Rightarrow & 1 \odot_3 2 = 2 \\ 1 \oplus_2 0 = 1 & \Rightarrow & 2 \odot_3 1 = 2 \\ 1 \oplus_2 1 = 0 & \Rightarrow & \underbrace{2}_{f(1)} \odot_3 \underbrace{2}_{f(1)} = \underbrace{1}_{f(0)} \end{array}$$

4.1.4 Aufgaben

Aufgabe 4.1.1 Es sei

$$G := \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \right\}$$

Zeigen Sie: Das Tupel (G, \cdot) ist eine Gruppe, wobei “ \cdot ” für die gewöhnliche Matrix-Multiplikation steht. Hinweis: Die Menge von Matrizen hat die Form $\{I, A, -A, -I\}$

Aufgabe 4.1.2 Es sei $G := \{a, b, c, d, g, h\}$ und $\circ : G \times G \rightarrow G$ eine Verknüpfung mit der Verknüpfungstabelle:

$\alpha \circ \beta$	$\beta =$					
	a	b	c	d	g	h
$\alpha = a$	c	g	b	a	h	d
b	g	d	h	b	a	c
c	b	h	g	c	d	a
d	a	b	c	d	g	h
g	h	a	d	g	c	b
h	d	c	a	h	b	g

- a) Nennen Sie das neutrale Element e in (G, \circ) .
- b) Bestimmen Sie für alle Elemente in G jeweils das Inverse Element.
- c) Ist (G, \circ) eine abelsche Gruppe?
- d) Zeigen Sie: $(a \circ b) \circ c = a \circ (b \circ c)$.
- e) Nennen Sie alle Quadrate in (G, \circ) .
- f) Bestimmen Sie das kleinste $k \in \mathbb{N}$ mit $a^k = e$.
- g) Bestimmen Sie das kleinste $\ell \in \mathbb{N}$ mit $c^\ell = e$.

4.2 Die Gruppenordnung

Im Folgenden werden wir zeigen, was beim mehrfachen Verknüpfen ein und desselben Gruppenelements passiert. Interessanterweise kommt man in einer abelschen Gruppe immer wieder am neutralen Element "vorbei". Um dies zu zeigen benötigen wir die Eigenschaften einer Funktion $f_a : G \rightarrow G$, die ein Element $x \in G$ einfach mit einem (festen!) Element a verknüpft: $f_a(x) = a \circ x$.

Für eine Gruppe (G, \circ) kann man die Mehrfach-Verknüpfung eines Elementes mit a^n abkürzen:

Definition 4.8

Für ein Element $a \in G$ einer Gruppe (G, \circ) und eine natürliche Zahl $n \in \mathbb{N}$ ist definiert:

$$a^n = \underbrace{a \circ a \circ \dots \circ a}_{n\text{-oft}} \circ e$$

Es gilt also $a^0 = e$, $a^1 = a$, $a^2 = a \circ a$ etc.

Die Gruppenordnung zählt die Elemente der Gruppe:

Definition 4.9 (Ordnung einer Gruppe)

Für eine Gruppe (G, \circ) ist die Anzahl $|G|$ der Elemente in G die *Ordnung* von G .

Hat G endlich viele Elemente, so nennt man (G, \circ) eine *endliche* Gruppe.

Hat G unendlich viele Elemente, so sagt man: Die Gruppenordnung von (G, \circ) ist ∞ .

Satz 4.10

Es sei (G, \circ) eine endliche abelsche Gruppe mit neutralem Element e .

Für alle $a \in G$ gilt dann: $a^{|G|} = e$ ($|G| = \text{Anzahl der Elemente von } G$).

Um diesen Satz zu beweisen benötigen wir das folgende Lemma:

Lemma 4.11

Es sei (G, \circ) eine endliche abelsche Gruppe und $a \in G$ sei ein fest gewähltes Element.

Dann ist die Abbildung $f_a : G \rightarrow G$ mit $f_a(x) := a \circ x$ bijektiv.

- **Surjektivität:** Zu zeigen ist: für jedes $y \in G$ gibt es ein $x \in G$ mit $f_a(x) = y$.
Sei $y \in G$ beliebig. Wähle $x := \bar{a} \circ y$, dann gilt: $f_a(x) = a \circ (\bar{a} \circ y) \stackrel{G1}{=} (a \circ \bar{a}) \circ y = y$.
- **Injektivität:** Zu zeigen ist: für jedes Paar $x, \tilde{x} \in G$ mit $x \neq \tilde{x}$ gilt: $f_a(x) \neq f_a(\tilde{x})$.
Es seien $x, \tilde{x} \in G$ beliebig mit $x \neq \tilde{x}$. Annahme es gelte: $f_a(x) = f_a(\tilde{x})$. Dann gilt:

$$\begin{aligned} f_a(x) = f_a(\tilde{x}) &\Leftrightarrow a \circ x = a \circ \tilde{x} && \quad | \bar{a} \circ \\ &\Rightarrow \bar{a} \circ (a \circ x) = \bar{a} \circ (a \circ \tilde{x}) \\ &\Leftrightarrow (\bar{a} \circ a) \circ x = (\bar{a} \circ a) \circ \tilde{x} && \Leftrightarrow x = \tilde{x} \end{aligned}$$

Die letzte Gleichung ist ein Widerspruch zu $x \neq \tilde{x}$. □

Nun können wir mit Lemma 4.11 den Satz 4.10 beweisen:

Beweis: (Satz 4.10) Es sei (G, \circ) eine endliche abelsche Gruppe mit $n := |G|$ Elementen $g_1, \dots, g_n \in G$. Weiter sei $a \in G$ beliebig (d.h. es gilt $a = g_i$ für ein i).

Da $f_a : G \rightarrow G$ mit $f_a(x) = a \circ x$ nach Lemma 4.11 eine bijektive Abbildung ist, gilt:

$$G = \{g_1, g_2, \dots, g_n\} = \{a \circ g_1, a \circ g_2, \dots, a \circ g_n\}$$

Die Gruppe (G, \circ) ist abelsch. Bildet man also die Verknüpfung aller Elemente in G , so spielt die Reihenfolge keine Rolle. Es gilt also

$$\underbrace{g_1 \circ \dots \circ g_n}_{\text{Verknüpfung aller } g_j \text{ sortiert}} = \underbrace{(a \circ g_1) \circ (a \circ g_2) \circ \dots \circ (a \circ g_n)}_{\text{Verknüpfung aller } g_j \text{ "durcheinander"}}$$

Da (G, \circ) abelsch ist, dürfen wir umsortieren, es gilt zum Beispiel $a \circ g_1 \circ a \circ g_2 = g_1 \circ a \circ a \circ g_2$. Wiederholt man dies immer wieder, so erhält man aus der letzten Gleichung schließlich:

$$g_1 \circ \dots \circ g_n = g_1 \circ g_2 \circ \dots \circ g_n \circ \underbrace{a \circ a \circ \dots \circ a}_{n \text{ Stück}}$$

Nun "kürzt" man durch sukzessives Multiplizieren auf beiden Gleichungsseiten mit $(g_1)^{-1}, (g_2)^{-1}, \dots$. Dieses Kürzen mit $g_1 \circ \dots \circ g_n$ liefert: $e = a^n$ bzw. $a^{|G|} = e$. □

Beispiel 4.12 (Lemma 4.11)

Dass die Abbildung $f_a : x \mapsto a \circ x$ bijektiv ist, hat zur Folge, dass die Verknüpfungstabelle einer Gruppe (G, \circ) immer ein kleines "Sudoku" ist:

In jeder Zeile (und in jeder Spalte) kommt jedes Element aus G *genau einmal* vor.

Dies veanschaulichen wir am Beispiel der Gruppe $(\mathbb{Z}_9^*, \odot_9)$. Es gilt für $a := 5$:

Element	g	1	2	4	5	7	8	
Bild	$f_5(g) = 5 \odot_9 g$	5	1	2	7	8	4	← Jedes g taucht genau einmal auf.

Hier ist die zweite Zeile " $f_5(g)$ " eine Zeile aus der Verknüpfungstabelle von \odot_9 :

Verknüpfungstabelle ist Sudoku.

		$b =$					
$a \odot_9 b$		1	2	4	5	7	8
$a = 1$	1	1	2	4	5	7	8
2	2	4	8	1	5	7	
4	4	8	7	2	1	5	
5	5	1	2	7	8	4	← Jedes g taucht genau einmal auf.
7	7	5	1	8	4	2	
8	8	7	5	4	2	1	

↑ Jedes g taucht genau einmal auf.

Beispiel 4.13 (Satz 4.10)

Wir untersuchen die Aussage $a^{|G|} = e$ am Beispiel der Gruppe $(\mathbb{Z}_5^*, \odot_5)$.

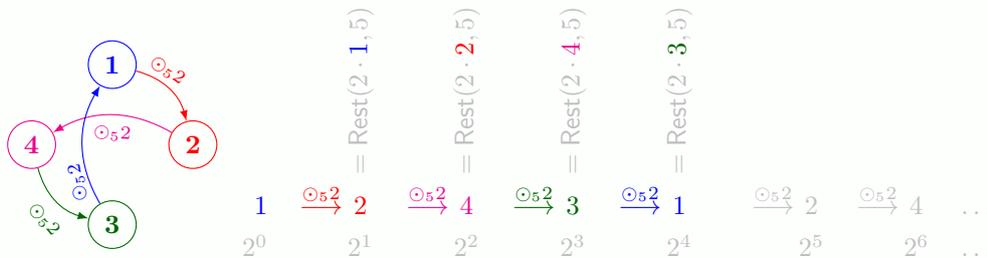
Hier gelten $\mathbb{Z}_5^* = \{1, 2, 3, 4\}$ und für zwei Elemente $a, b \in \mathbb{Z}_5^*$ ist $a \odot_5 b := \text{Rest}(a \cdot b, 5)$, d.h. $a \odot_5 b$ ist der Rest der beim Teilen von $a \cdot b$ durch 5 entsteht.

Das Paar $(\mathbb{Z}_5^*, \odot_5)$ bildet eine Gruppe, dies wurde in Beispiel 4.3 (S. 24) gezeigt.

Hier ist $e = 1$. Weiter gilt $|G| = 4$ wegen $\mathbb{Z}_5^* = \{1, 2, 3, 4\}$. Wir untersuchen also nun a^4 für $a \in \{1, 2, 3, 4\}$:

$$\begin{aligned} 1^4 &= \text{Rest}(1^4, 5) = \text{Rest}(1, 5) = 1 \\ 2^4 &= \text{Rest}(2^4, 5) = \text{Rest}(16, 5) = 1 \\ 3^4 &= \text{Rest}(3^4, 5) = \text{Rest}(81, 5) = 1 \\ 4^4 &= \text{Rest}(4^4, 5) = \text{Rest}(256, 5) = 1 \end{aligned}$$

Berechnet man alle Werte von 2^k in \mathbb{Z}_5^* durch sukzessive Multiplikation " $\odot_5 2$ ", so erhält man nach und nach alle Elemente aus \mathbb{Z}_5^* : Man startet bei $1 = 2^0$ und nach einem Zyklus von $4 = |\mathbb{Z}_5^*|$ -oft "mal-zwei-nehmen" erreicht man zwangsläufig wieder die 1:



Kapitel 5

Vektorräume

5.1 Die Vektorräume \mathbb{R}^n

In diesem Skript erinnern wir *zunächst* an die aus der Schule bekannten Vektoren aus dem \mathbb{R}^n . Eine genaue Definition dafür, was ein allgemeiner Vektor ist findet sich im nachfolgenden Kapitel. Ein Vektor in der Schulmathematik ist zunächst einmal ein Vektor aus \mathbb{R}^3 oder \mathbb{R}^2 , d.h. eine Spalte mit Zahleneinträgen. Diese Vektoren haben eine geometrische Bedeutung, die sich auf zwei verschiedene Weisen verstehen lässt:

- Ein Vektor kann als Punkt in einem Raum aufgefasst werden. Wählt man beispielsweise einen festen Bezugspunkt im uns umgebenden dreidimensionalen Raum, so lässt sich jeder Punkt in unserem Universum durch einen Vektor mit drei Einträgen (Höhe, Breite, Länge) relativ zu diesem Punkt beschreiben.
- Andererseits repräsentieren Vektoren in der Physik Kräfte, also eine Messgröße die mit einer Richtung einhergeht: Im Gegensatz zu *skalaren* Messgrößen wie Temperatur oder Masse, muss man um eine Kraft vollständig zu beschreiben nicht nur angeben wie groß die Kraft ist, sondern auch in welche Richtung sie wirkt.

Wdh: Vektoren des \mathbb{R}^n

Für ein festes $n \in \mathbb{N}$ ist \mathbb{R}^n ein n -dimensionaler Vektorraum.

Ein Vektor $\vec{x} \in \mathbb{R}^n$ ist ein Tupel mit n reellen Zahleneinträgen $x_1, x_2, \dots, x_n \in \mathbb{R}$. Die allgemeine Form eines solchen Vektors \vec{x} lautet:

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

Die Zahlen x_1, \dots, x_n heißen die Komponenten des Vektors.

Wir führen zwei Rechenregeln für Vektoren die folgenden Rechenregeln ein:

Definition 5.1

Es sei $n \in \mathbb{N} \setminus \{0\}$. Für zwei Vektoren $\vec{a}, \vec{b} \in \mathbb{R}^n$ und $\lambda \in \mathbb{R}$ gelten:

$$\vec{a} + \vec{b} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} + \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} := \begin{pmatrix} a_1 + b_1 \\ \vdots \\ a_n + b_n \end{pmatrix} \quad \lambda \cdot \vec{a} = \lambda \cdot \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} := \begin{pmatrix} \lambda \cdot a_1 \\ \vdots \\ \lambda \cdot a_n \end{pmatrix}$$

Bemerke: Man kann Vektoren mit *gleich vielen* Einträgen addieren oder von einander abziehen (Dies geht mit Vektoren mit *verschieden vielen* Einträgen nicht!). Für die Addition von Vektoren und die Multiplikation mit einer Zahl gelten die selben Rechenregeln, die man schon von “normalen Zahlen” kennt:

Korollar 5.2

Für $\vec{a}, \vec{b}, \vec{c} \in \mathbb{R}^n$ und $\lambda, \mu \in \mathbb{R}$ gelten:

Kommutativgesetz: $\vec{a} + \vec{b} = \vec{b} + \vec{a}$

Assoziativgesetz: $(\vec{a} + \vec{b}) + \vec{c} = \vec{a} + (\vec{b} + \vec{c})$

Distributivgesetze: $(\lambda + \mu) \cdot \vec{a} = \lambda \cdot \vec{a} + \mu \cdot \vec{a}$ $\lambda \cdot (\vec{a} + \vec{b}) = \lambda \cdot \vec{a} + \lambda \cdot \vec{b}$

Den Beweis überlassen wir dem Leser.

Addition

Die *Addition* zweier Vektoren \vec{a} und \vec{b} entspricht geometrisch dem Aneinanderhängen der entsprechenden Vektoren (siehe Abb.5.1).

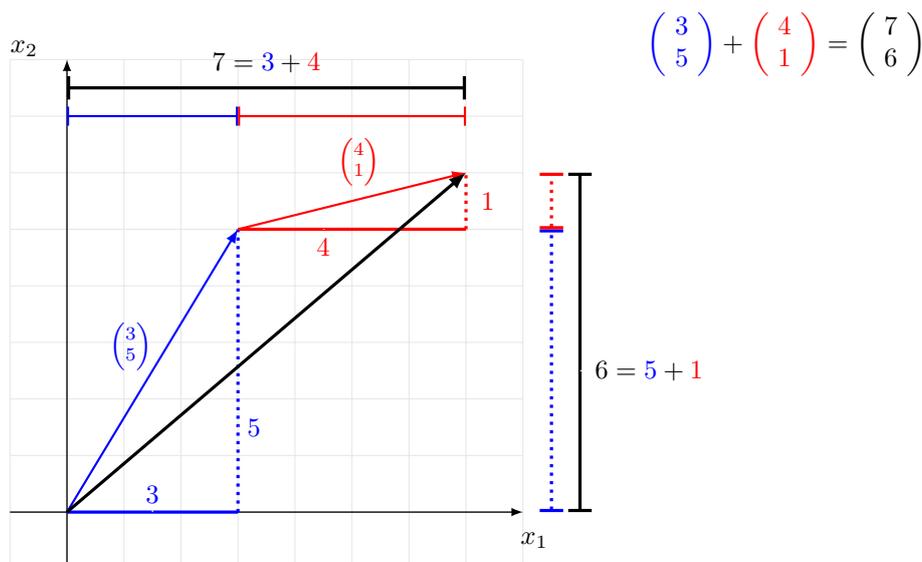


Abbildung 5.1: Addition von Vektoren

Die *Subtraktion* zweier Vektoren, $\vec{a} - \vec{b}$, wird einfach als Addition von \vec{a} und $-\vec{b}$ aufgefasst. Geometrisch kann man dies verstehen als dass man den Vektor \vec{b} umkehrt und dann an den Vektor \vec{a} anhängt.

Eine deutlich bessere Anschauung erhält man jedoch, wenn man $\vec{c} = \vec{a} - \vec{b}$ liest als “ \vec{c} ist derjenige Vektor, der von \vec{b} zu \vec{a} führt. Denn es gilt:

$$\vec{c} = \vec{a} - \vec{b} \quad \Leftrightarrow \quad \vec{b} + \vec{c} = \vec{a}$$

Dies liefert dann das folgende Bild:

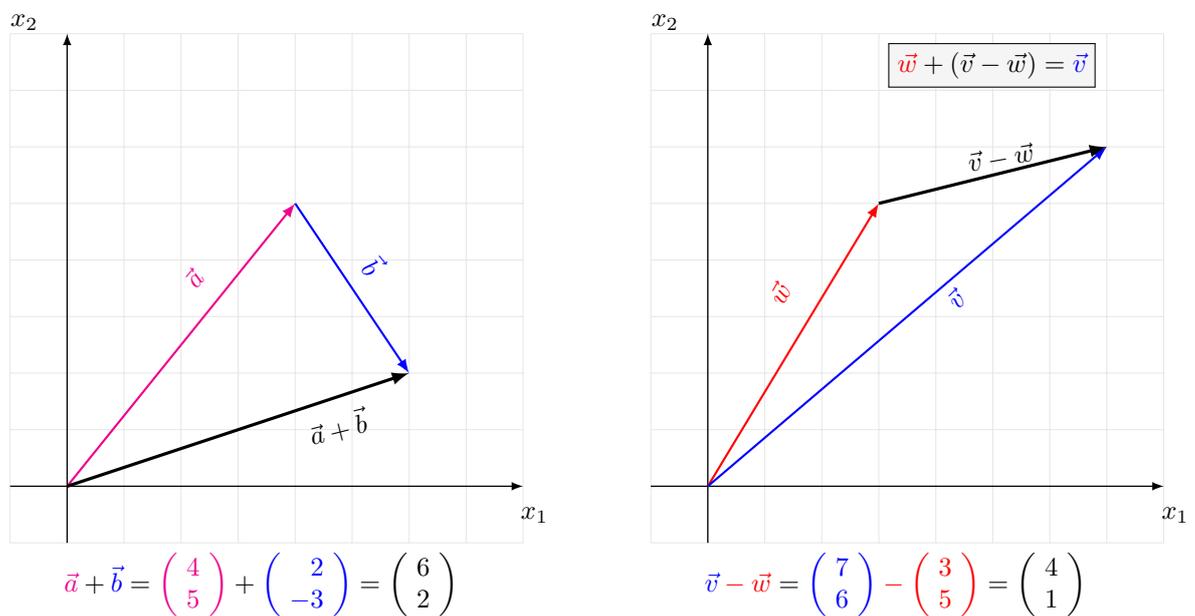


Abbildung 5.2: Addition und Subtraktion von Vektoren

In Koordinatenschreibweise sieht dies dann naheliegenderweise so aus.

$$\vec{a} - \vec{b} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} - \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} a_1 - b_1 \\ a_2 - b_2 \\ \vdots \\ a_n - b_n \end{pmatrix}$$

Multiplikation mit einer Zahl

Die *Multiplikation* eines Vektors \vec{x} mit einer *Zahl* ist *verträglich* mit der Vektor-Addition. Dies bedeutet dass beispielsweise $2 \cdot \vec{a} = \vec{a} + \vec{a}$ gilt:

$$2 \cdot \vec{a} = 2 \cdot \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} 2 \cdot a_1 \\ \vdots \\ 2 \cdot a_n \end{pmatrix} = \begin{pmatrix} a_1 + a_1 \\ \vdots \\ a_n + a_n \end{pmatrix} = \vec{a} + \vec{a}.$$

Geometrisch entspricht die Multiplikation eines Vektors mit einer Zahl λ also einer Streckung bzw. einer Stauchung von \vec{a} um den Faktor λ ,

für $0 < |\lambda| < 1$ ist $\lambda \cdot \vec{a}$ *kürzer* als \vec{a} . für $1 < |\lambda|$ ist $\lambda \cdot \vec{a}$ *länger* als \vec{a} .

Ist λ negativ, so kehrt sich die Richtung eines Vektors \vec{a} beim Multiplizieren mit λ um, der Vektor $-\vec{a} = (-1) \cdot \vec{a}$ zeigt also genau entgegengesetzt zu \vec{a} .

5.2 Allgemeine Vektorräume

Im weiteren Verlauf werden wir uns mit abstrakteren Vektorräumen beschäftigen zum Beispiel dem Vektorraum der Polynome. Die bereits bekannten Rechengesetze aus dem \mathbb{R}^n wollen wir dabei “mitnehmen” also auf ein abstrakteres Niveau anheben. Im Wesentlichen verlangen wir also einfach, dass die beiden Rechenoperationen “Addition von Vektoren” und “Multiplikation eines Vektors mit einer Zahl” *sinnvoll* funktionieren:

Bei genauerem Hinsehen entdeckt man, dass $(\mathbb{R}^n, +)$ eine Gruppe ist, dass also die Elemente aus \mathbb{R}^n beliebig addiert und subtrahiert werden können, und dass es eine Null (den Nullvektor) gibt. Die Multiplikation mit Zahlen $\lambda \in \mathbb{R}$ kommt dann als “Extra” hinzu. Die Addition innerhalb der Gruppe und die Multiplikation mit Elementen “von außen” aus \mathbb{R} müssen bestimmte Verträglichkeitsregeln erfüllen.

Wdh: Eigenschaften des \mathbb{R}^n

In der Menge $V := \mathbb{R}^n$ sind die folgenden eigenschaften erfüllt:

- 1) Man kann **in V addieren**, es gilt sogar: $(V, +)$ ist eine *abelsche* Gruppe.

Man kann Vektoren aus V **mit einer Zahl λ Multiplizieren**: $\lambda \cdot \vec{v} \in V$ gilt für alle $\lambda \in \mathbb{R}$, $\vec{v} \in V$.

- 2) Es gelten Verträglichkeitsregeln für $\vec{v}, \vec{w} \in V$ und $\lambda, \mu \in \mathbb{R}$:

$$\begin{array}{ll} \mathbf{2.1)} & (\lambda + \mu) \cdot \vec{v} = \lambda \cdot \vec{v} + \mu \cdot \vec{v} & \mathbf{2.3)} & \lambda \cdot (\mu \cdot \vec{v}) = (\lambda \cdot \mu) \cdot \vec{v} \\ \mathbf{2.2)} & \lambda \cdot (\vec{v} + \vec{w}) = \lambda \cdot \vec{v} + \lambda \cdot \vec{w} & \mathbf{2.4)} & 1 \cdot v = v \end{array}$$

Diese Eigenschaften wollen wir verallgemeinern.

Um dabei die Definition korrekt und so allgemein wie möglich zu halten, verwenden wir für die “neuen” Operationen (Vektoraddition & skalare Multiplikation) hervorgehobene Symbole \oplus und \odot . Im normalen mathematischen Alltag schreibt man aber statt dessen einfach “+” und “.”. Dies werden wir nach diesem Abschnitt auch so tun, im Prinzip sind aber die beiden Additionen $\begin{pmatrix} 1 \\ 1 \end{pmatrix} \oplus \begin{pmatrix} 2 \\ 2 \end{pmatrix}$ und $1 + 2$ tatsächlich unterschiedliche Operationen.

Ein Vektorraum ist eine abelsche, additive Gruppe (V, \oplus) zusammen mit einer *äußeren* Multiplikation “ \odot ” mit Elementen aus einem Körper. Die Addition innerhalb der Gruppe und die Multiplikation mit

Elementen “von außen” aus \mathbb{R} müssen bestimmte *Verträglichkeitsregeln* erfüllen.

Definition 5.3

Eine Menge V zusammen mit

- einer inneren Verknüpfung $+$: $V \times V \rightarrow V$ (“Vektor-Addition”)
- einer äußeren Verknüpfung \cdot : $\mathbb{R} \times V \rightarrow V$ (“Multiplikation mit einer Zahl”)

nennt man einen *Vektorraum über \mathbb{R}* , falls gelten: die folgenden Bedingungen erfüllt sind.

V1) $(V, +)$ ist eine *abelsche* Gruppe.

V2) Es gelten für $v, w \in V$ und $\lambda, \mu \in \mathbb{R}$:

$$\begin{array}{ll} \mathbf{V2.1)} & (\lambda + \mu) \cdot \vec{v} = \lambda \cdot \vec{v} + \mu \cdot \vec{v} \\ \mathbf{V2.2)} & \lambda \cdot (\vec{v} + \vec{w}) = \lambda \cdot \vec{v} + \lambda \cdot \vec{w} \end{array} \quad \begin{array}{ll} \mathbf{V2.3)} & \lambda \cdot (\mu \cdot \vec{v}) = (\lambda \cdot \mu) \cdot \vec{v} \\ \mathbf{V2.4)} & 1 \cdot v = v \end{array}$$

In der Gruppe $(V, +)$ bezeichnet man ...

- das neutrale Element mit $\mathbf{0}$ (im bekannten \mathbb{R}^2 gilt $\mathbf{0} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$).
- das zu $v \in V$ inverse Element mit $-v$.

Aus der Definition des Vektorraumes ergeben sich sofort Konsequenzen:

Korollar 5.4

In einem Vektorraum V gelten:

1. Die Menge V ist nicht leer (Definition einer Gruppe, Def. 4.1).
2. Für $\lambda \in \mathbb{R}$ und $\vec{v} \in V$ gelten die Rechenregeln:

$$\begin{array}{ll} \mathbf{2.1)} & 0 \cdot \vec{v} = \mathbf{0} \\ \mathbf{2.2)} & \lambda \cdot \mathbf{0} = \mathbf{0} \end{array} \quad \begin{array}{ll} \mathbf{2.3)} & (-1) \cdot \vec{v} = -\vec{v} \quad (-\vec{v} \text{ ist das Inverse zu } \vec{v} \text{ in der Gruppe } (V, +)) \\ \mathbf{2.4)} & \lambda \cdot \vec{v} = \mathbf{0} \iff (\lambda = 0) \vee (\vec{v} = \mathbf{0}) \end{array}$$

Beweis:

2.1) Wegen $\vec{v} \stackrel{\text{Trick}}{=} (0 + 1) \cdot \vec{v} \stackrel{\text{V2.1)}}{=} 0 \cdot \vec{v} + \vec{v}$ folgt: $0 \cdot \vec{v}$ ist das neutrale Element $\mathbf{0}$ in $(V, +)$.

2.2) Für $\lambda = 0$ gilt wegen 2.1) sofort $\lambda \cdot \mathbf{0} = \mathbf{0}$.

Es sei nun $\lambda \neq 0$ und $\vec{v} \in V$ beliebig. Dann gilt:

$$\begin{aligned} \vec{v} &= 1 \cdot \vec{v} \stackrel{\text{Trick}}{=} (\lambda \cdot \frac{1}{\lambda}) \cdot \vec{v} \stackrel{\text{V2.3)}}{=} \lambda \cdot (\frac{1}{\lambda} \cdot \vec{v}) \\ &\stackrel{\text{Trick}}{=} \lambda \cdot (\frac{1}{\lambda} \cdot \vec{v} + \mathbf{0}) \\ &\stackrel{\text{V2.2)}}{=} (\lambda \cdot \frac{1}{\lambda}) \cdot \vec{v} + \lambda \cdot \mathbf{0} = \vec{v} + \lambda \cdot \mathbf{0} \end{aligned}$$

Wegen $\vec{v} = \lambda \cdot \mathbf{0} + \vec{v}$ folgt: $\lambda \cdot \mathbf{0}$ ist das neutrale Element $\mathbf{0}$ der Gruppe $(V, +)$.

2.3) Den Beweis für 2.3) überlassen wir dem Leser.

2.4) Die Richtung “ \Leftarrow ” folgt direkt aus 2.1) und 2.2).

“ \Rightarrow ”: Sei $\lambda \cdot \vec{v} = \vec{0}$ und gelte $\lambda \neq 0$ zu zeigen ist: Dann gilt $\vec{v} = \vec{0}$.

$$\vec{v} = 1 \cdot \vec{v} \stackrel{\text{Trick}}{=} \left(\frac{1}{\lambda} \cdot \lambda\right) \cdot \vec{v} \stackrel{\text{V2.3)}}{=} \frac{1}{\lambda} \cdot (\lambda \cdot \vec{v}) = \frac{1}{\lambda} \cdot \vec{0} = \vec{0}$$

□

Wieso “ $1 \cdot \vec{v} = \vec{v}$ ”?

Wieso steht in V3.3) die scheinbar offensichtliche Bedingung “ $1 \cdot \vec{v} = \vec{v}$ ”?

Lässt man V3.3) weg, so bilden **Vektor-Addition** und **Multiplikation mit einem Skalar** auf V zwei völlig voneinander losgelöste Operationen: Denkbar wäre, den uns bekannten \mathbb{R}^2 mit einer “ y -Weglassen-Multiplikation” der Form $\lambda \odot \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \lambda \cdot x \\ 0 \end{pmatrix}$ zu versehen. Diese Form der Multiplikation erfüllt alle Regeln in V1) und V2) außer V2.4). Diese Definition liefert aber

$$2 \odot \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2x \\ 0 \end{pmatrix} \neq \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} x \\ y \end{pmatrix}$$

Die Bedingung V2.4) stellt also sicher, dass **Vektoraddition** $+$ in V und äußere Multiplikation “sinnvoll” mit einander funktionieren: Dank $1 \cdot \vec{v} = \vec{v}$ kann man zum Beispiel auf $2 \cdot \vec{v} = \vec{v} + \vec{v}$ schließen und entdeckt, dass die Multiplikation $\lambda \cdot \vec{v}$ *tatsächlich* das tut was man erwartet:

$$2 \cdot \vec{v} = (1 + 1) \cdot \vec{v} \stackrel{\text{V3.1)}}{=} 1 \cdot \vec{v} + 1 \cdot \vec{v} \stackrel{\text{V3.3)}}{=} \vec{v} + \vec{v}$$

Beispiel 5.5

- a) Die Bekannte Menge $\mathbb{R}^3 := \left\{ \begin{pmatrix} a \\ b \\ c \end{pmatrix} : a, b, c \in \mathbb{R} \right\}$ bildet mit der üblichen Vektor-Addition und Skalarmultiplikation einen Vektorraum.
- b) Die Menge $\mathbb{R}[x]_2 := \{ax^2 + bx + c : a, b, c \in \mathbb{R}\}$ aller Polynome vom Grad höchstens 2 bildet einen Vektorraum zusammen mit der üblichen Addition von Polynomen und der üblichen Multiplikation mit einer Zahl. Der Nullvektor ist hier da Polynom 0 (bzw. das Polynom $0x^2 + 0x^1 + 0$).

Die Beispiele in 5.5 sehen sich überraschend ähnlich und auch die Operationen (Addition und skalare Multiplikation) laufen weitestgehend gleich ab:

$$\begin{array}{r} (ax^2 \quad +bx \quad +c) \\ + \quad (\alpha x^2 \quad +\beta x \quad +\gamma) \\ \hline + \quad ((a+\alpha)x^2 \quad +(b+\beta)x \quad +(c+\gamma)) \end{array} \quad \left| \quad \begin{pmatrix} a \\ b \\ c \end{pmatrix} + \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = \begin{pmatrix} a+\alpha \\ b+\beta \\ c+\gamma \end{pmatrix} \right.$$

Der Vektorraum $\mathbb{R}[x]_2$ ist also quasi nichts anderes als ein “hingelegter” \mathbb{R}^3 , die Koeffizienten der Polynome benehmen sich identisch zu den Koeffizienten der Vektoren des \mathbb{R}^3 . Den $\mathbb{R}[x]_2$ als Vektorraum zu betrachten bringt also auf im Sinne von Vektorräumen nichts wirklich neues.

Den Umstand das diese Vektorräume die gleiche (iso) Form (morph) haben werden wir in Def. 5.7 abstrakt fassen. Vorher wollen wir jedoch mit zwei “unendlich dimensionalen Räumen” ein neues abstraktes Konzept entdecken. Diese Vektorräume lassen sich dann nicht mehr ohne weiteres mit den bekannten

\mathbb{R}^n -Vektorräumen vergleichen:

Beispiel 5.6

- Die Menge *aller* reellen Polynome $\mathbb{R}[x]$ bildet einen Vektorraum mit der üblichen Addition von Polynomen und der üblichen Multiplikation mit einer Zahl. Der Nullvektor ist hier das Polynom $p(x) := 0$.
- Die Menge aller reellen Funktionen $\mathcal{C} := \{f : \mathbb{R} \rightarrow \mathbb{R}\}$ ist ein Vektorraum mit der üblichen Addition von Funktionen und der üblichen Multiplikation mit einer Zahl.

Definition 5.7

Zwei \mathbb{R} -Vektorräume $(V, +, \cdot)$ und $(W, +, \cdot)$ heißen *isomorph*^a, wenn es eine *bijektive* Abbildung $f : V \rightarrow W$ gibt mit:

- $f(v_1 + v_2) = f(v_1) + f(v_2)$ gilt für alle $v_1, v_2 \in V$
- $f(\lambda \cdot v) = \lambda \cdot f(v)$ gilt für $v \in V$ und $\lambda \in \mathbb{R}$.

Die Abbildung f nennt man dann einen Isomorphismus.

^agleich-geformt

Beispiel 5.8

Die Vektorräume $(\mathbb{R}^3, +, \cdot)$ und $(\mathcal{R}[x]_2, +, \cdot)$ sind isomorph, der Isomorphismus lautet:

$$f : \begin{pmatrix} a \\ b \\ c \end{pmatrix} \mapsto ax^2 + bx + c$$

5.3 Basen und Dimension

Wir führen ein Maß für die Größe eines Vektorraums ein, die *Dimension*. Die Dimension ist grob gesagt ein Maß für die Bewegungsmöglichkeiten in einem Vektorraum, sogenannte Freiheitsgrade.

Es ist leicht zu sehen, dass man für die Wahl eines Punktes im \mathbb{R}^3 drei Freiheitsgrade hat, während es für einen Punkt im \mathbb{R}^2 nur zwei Freiheitsgrade sind. Die hat direkte Konsequenzen beim Bau von ferngesteuertem Spielzeug:

- Eine Spielzeugeisenbahn muss nur vor oder zurück fahren. Mathematisch gesehen fährt die Lok auf einer Geraden, hat also nur einen Freiheitsgrad und benötigt deswegen auch nur einen Motor.
- Ein Modellauto dagegen benötigt mindestens *zwei Kontrollmöglichkeiten*, weil es sich abstrakt gesehen im \mathbb{R}^2 (alias “der Fußboden”) frei bewegen können muss.
- Ein Modellflugzeug muss mindestens *drei Kontrollmöglichkeiten* besitzen, weil es sich frei im \mathbb{R}^3 bewegen können muss.

Um die Dimension definieren zu können müssen wir zunächst definieren, was es bedeutet “mit bestimmten

Laufrichtungen Punkte im Vektorraum zu erreichen". Dazu definieren wir:

Definition 5.9

Es seien $\vec{v}_1, \dots, \vec{v}_k \in V$ Vektoren in einem Vektorraum V .

Eine Summe der Form $\mu_1 \cdot \vec{v}_1 + \dots + \mu_k \cdot \vec{v}_k$ mit $\mu_1, \dots, \mu_k \in \mathbb{R}$ nennt man *eine Linearkombination* der Vektoren \vec{v}_i . Die Menge aller Linearkombinationen der \vec{v}_i nennt man den Spann von $\vec{v}_1, \dots, \vec{v}_k$:

$$\text{spann}(\vec{v}_1, \dots, \vec{v}_k) := \left\{ \sum_{i=1}^k \mu_i \vec{v}_i : \mu_1, \dots, \mu_k \in \mathbb{R} \right\}$$

Die Addition von Vektoren entspricht dem "Hinterinanderhängen" der entsprechenden Vektoren. Geometrisch ist eine Linearkombination $\mu_1 \cdot \vec{v}_1 + \dots + \mu_k \cdot \vec{v}_k$ also eine "Laufanweisung" der Form "Folge dem Vektor v_1 , dann v_2 etc.". Dabei geben die μ_i jeweils (grob gesagt) an, wie weit jeweils zu laufen ist.

Der Spann ist dann die Menge an Punkten, die über solche Laufwege erreichbar ist.

Beispiel 5.10

1) Für die Vektoren $\vec{v}_1 := \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, $\vec{v}_2 := \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ ist der Spann:

$$\text{spann}(\vec{v}_1, \vec{v}_2) := \left\{ \mu_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \mu_2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} : \mu_1, \mu_2 \in \mathbb{R} \right\} = \left\{ \begin{pmatrix} \mu_1 \\ \mu_2 \\ 0 \end{pmatrix} : \mu_1, \mu_2 \in \mathbb{R} \right\}$$

2) Fügt man den Vektor $\vec{v}_3 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$ hinzu, ändert sich der Spann nicht, da sich \vec{v}_3 bereits als Summe von \vec{v}_1 und \vec{v}_2 schreiben lässt (Die "Zutat" \vec{v}_3 liefert also keine "echt neue" Laufrichtung, das Verwenden von \vec{v}_3 lässt sich durch das Verwenden von v_1 und v_2 ersetzen:

$$\begin{aligned} \text{spann}(\vec{v}_1, \vec{v}_2, \vec{v}_3) &:= \left\{ \mu_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \mu_2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + \mu_3 \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} : \mu_1, \mu_2, \mu_3 \in \mathbb{R} \right\} \\ &= \left\{ \begin{pmatrix} \mu_1 + \mu_3 \\ \mu_2 + \mu_3 \\ 0 \end{pmatrix} : \mu_1, \mu_2, \mu_3 \in \mathbb{R} \right\} \\ &= \left\{ \begin{pmatrix} x \\ y \\ 0 \end{pmatrix} : x, y \in \mathbb{R} \right\} \end{aligned}$$

Um das letzte Gleichheitszeichen einzusehen macht man sich klar, dass hier \subseteq stets gilt: Die letzte Menge enthält *alle* Vektoren mit drittem Eintrag 0 (und jeder Vektor der vorletzten Menge ist ein solcher).

Umgekehrt gilt: Ist $\vec{w} := \begin{pmatrix} x \\ y \\ 0 \end{pmatrix}$ ein beliebiger Vektor mit $x, y \in \mathbb{R}$, so ist \vec{w} auch ein Element der vorletzten

Menge: Setze im Vektor $\begin{pmatrix} \mu_1 + \mu_3 \\ \mu_2 + \mu_3 \\ 0 \end{pmatrix}$ einfach $\mu_1 = x$, $\mu_2 = y$, $\mu_3 = 0$ und erhalte \vec{w} .

Das letzte Beispiel zeigt: Geht es darum eine Menge als Spann von Vektoren zu beschreiben, so ist es nicht hilfreich "überflüssige" Vektoren zu verwenden. Um solche Mengen von Vektoren zu vermeiden, führen

wir den folgenden Begriff ein (das Lemma 5.15 klärt dann, was dieser für den Spann bedeutet):

Definition 5.11

Eine Menge von Vektoren $\{\vec{v}_1, \dots, \vec{v}_k\} \subseteq V$ in einem Vektorraum V heißt *linear unabhängig* falls gilt:

$$\mu_1 \cdot \vec{v}_1 + \dots + \mu_k \cdot \vec{v}_k = \mathbf{0} \quad \Rightarrow \quad \mu_1 = \mu_2 = \dots = \mu_k = 0$$

Gilt dies nicht nennt man die Vektoren *linear abhängig* (oder auch kurz abhängig).

Man Bemerke: Für die Gleichung $\mu_1 \cdot \vec{v}_1 + \dots + \mu_k \cdot \vec{v}_k = \mathbf{0}$ ist $\mu_1 = \dots = \mu_k = 0$ *immer* eine Lösung, es kann aber noch weitere Lösungen geben. Die Vektoren $\{\vec{v}_1, \dots, \vec{v}_k\}$ sind also linear unabhängig, wenn $\mu_1 = \mu_2 = \dots = \mu_k = 0$ die *einzigste* Lösung der Gleichung $\mu_1 \cdot \vec{v}_1 + \dots + \mu_k \cdot \vec{v}_k = \mathbf{0}$ ist.

Beispiel 5.12

Die Vektoren $\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\} \subset \mathbb{R}^3$ sind linear **un**abhängig, denn:

$$\mu_1 \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \mu_2 \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + \mu_3 \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \mathbf{0} \quad \Leftrightarrow \quad \begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad \Leftrightarrow \quad \mu_1 = \mu_2 = \mu_3 = 0$$

Beispiel 5.13

Die Vektoren $\vec{v}_1 := \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \vec{v}_2 := \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \vec{v}_3 := \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \in \mathbb{R}^3$ sind linear **ab**hängig, denn es gilt:

$$1 \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + 1 \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + (-1) \cdot \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Beispiel 5.14

In jedem Vektorraum V gilt:

Die Menge $\{\mathbf{0}\} \subseteq V$ ist linear **ab**hängig, denn für $\lambda_1 \cdot \mathbf{0} = \mathbf{0}$ ist jedes $\lambda_1 \in \mathbb{R}$ eine Lösung.

Ganz genauso ist jede andere Menge der Form $\{\mathbf{0}, \vec{v}_1, \dots, \vec{v}_k\} \subset V$ linear abhängig.

In einer Menge von linear **ab**hängigen Vektoren ist beim Bilden des Spanns immer (mindestens) ein Vektor überflüssig, weil er sich aus den anderen herstellen lässt:

Lemma 5.15

Die Vektoren $\{\vec{v}_1, \dots, \vec{v}_k\}$ sind genau dann linear **ab**hängig, wenn es einen Vektor $\vec{v}_j \in \{\vec{v}_1, \dots, \vec{v}_k\}$ gibt, der sich als Linearkombination der anderen Vektoren schreiben lässt, wenn also gilt:

$$\exists \mu_1, \dots, \mu_{j-1}, \mu_{j+1}, \dots, \mu_k \in \mathbb{R} \quad : \quad \vec{v}_j = \sum_{\substack{i=1 \\ i \neq j}}^n \mu_i \vec{v}_i \quad (5.1)$$

Beweis:

“ \Rightarrow ” Annahme: Die Vektoren $\vec{v}_1, \dots, \vec{v}_k$ seien linear **abhängig**. Dann hat $\sum_{i=1}^k \lambda_i \vec{v}_i = \mathbf{0}$ mehr als nur die triviale Lösung $\lambda_1 = \dots = \lambda_k = 0$, d.h. eine Lösung mit mindestens einem $\lambda_j \neq 0$. Es gilt also:

$$\lambda_j \cdot \vec{v}_j + \sum_{\substack{i=1 \\ i \neq j}}^k \lambda_i \vec{v}_i = \mathbf{0} \quad \Longrightarrow \quad \vec{v}_j = \sum_{\substack{i=1 \\ i \neq j}}^k \frac{-\lambda_i}{\lambda_j} \cdot \vec{v}_i$$

Setzt man nun $\mu_i := \frac{-\lambda_i}{\lambda_j}$ für alle $i \in \{1, \dots, k\} \setminus \{j\}$, so hat man die gewünschte Darstellung von v_j .

“ \Leftarrow ” Annahme es gilt (5.1), d.h. $\vec{v}_j = \sum_{\substack{i=1 \\ i \neq j}}^k \mu_i \cdot \vec{v}_i$ mit $\mu_i \in \mathbb{R}$. Setzt man zusätzlich $\mu_j := -1$ so gilt:

$$v_j = \sum_{\substack{i=1 \\ i \neq j}}^k \mu_i \cdot \vec{v}_i \quad \Longrightarrow \quad \mathbf{0} = -v_j + \sum_{\substack{i=1 \\ i \neq j}}^k \mu_i \cdot \vec{v}_i \quad \Longrightarrow \quad \mathbf{0} = \sum_{i=1}^k \mu_i \cdot \vec{v}_i \quad \text{mit } \mu_j \neq 0$$

Die Gleichung $\sum_{i=1}^k \mu_i \vec{v}_i = \mathbf{0}$ hat also mehr als nur die Lösung $\mu_1 = \dots = \mu_k = 0$, die Vektoren sind also linear **abhängig**. □

Beispiel 5.16

Die Vektoren $\vec{v}_1 := \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, $\vec{v}_2 := \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$, $\vec{v}_3 := \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$, $\vec{v}_4 := \begin{pmatrix} 5 \\ 7 \\ 0 \end{pmatrix}$ aus \mathbb{R}^3 sind linear **abhängig**, denn es gilt $\vec{v}_4 = 5 \cdot \vec{v}_1 + 7 \cdot \vec{v}_2$.

Beispiel 5.17

Es sei $\mathbb{R}[x]$ der Vektorraum aller Polynome. Die Vektoren $p_1 := x$, $p_2 := x^2 + x$, $p_3 := x^2 + 5x$ aus $\mathbb{R}[x]$ sind linear abhängig, denn es gilt $\vec{p}_3 = 4 \cdot \vec{p}_1 + \vec{p}_2$.

Exkurs: Linear unabhängig

Betrachtet man Lemma 5.15 so bedeutet “ v_1, \dots, v_k sind linear unabhängig”, dass *kein* Vektor aus der Menge $\{v_1, \dots, v_k\}$ durch die anderen “hergestellt” werden kann.

Frage: Wieso verwendet man dies nicht als Definition für den Begriff “linear unabhängig”?

Antwort: Man verwendet dies nicht als Definition, weil diese Aussage kein praktisches Prüf-Kriterium bietet: Um zu Zeigen, dass v_1, \dots, v_k linear **unabhängig** sind, müsste man für jeden der k Vektoren ein eigenes ineares Gleichungssystem aufstellen und zeigen, dass *jedes* davon unlösbar ist.

5.3.1 Basis

Der Kernbegriff den wir benötigen, um die Dimension eines Vektorraums zu definieren ist der einer Basis.

Definition 5.18

Seien $v_1, \dots, v_n \in V$ Vektoren in einem Vektorraum V . Wir nennen $\{v_1, \dots, v_n\}$ eine **Basis** von V , wenn erfüllt sind:

B1) Die Vektoren $\{v_1, \dots, v_n\}$ sind linear unabhängig.

B2) $V = \text{spann}(v_1, \dots, v_n)$

Geometrisch bedeutet dies: Mit den ‐Laufrichtungen‐ \vec{v}_i kann man jeden Punkt in V erreichen, und die ‐Wegbeschreibung‐ ist eindeutig:

Korollar 5.19

Ist $\{v_1, \dots, v_n\}$ eine Basis des Vektorraums V ,
so gibt es zu jedem Vektor $\vec{w} \in V$ genau ein n -Tupel $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ mit $\vec{w} = \sum_{i=1}^n \lambda_i \cdot \vec{v}_i$.

Beweis: Es sei $\{v_1, \dots, v_n\}$ eine Basis des Vektorraums V und sei $\vec{w} \in V$ beliebig.
Wegen $V = \text{spann}(v_1, \dots, v_n)$ gibt es (mindestens) ein Tupel $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ mit $\vec{w} = \sum_{i=1}^n \lambda_i \cdot \vec{v}_i$.
Nehmen wir nun an, es gäbe noch ein anderes n -Tupel $\mu_1, \dots, \mu_n \in \mathbb{R}$ mit $\vec{w} = \sum_{i=1}^n \mu_i \cdot \vec{v}_i$ ist, dann gilt:

$$\vec{0} = \vec{w} - \vec{w} = \sum_{i=1}^n (\lambda_i - \mu_i) \cdot \vec{v}_i$$

Weil nach B1) die Vektoren v_1, \dots, v_n linear **un**abhängig sind, muss $\lambda_i - \mu_i = 0$ für alle $i \in \{1, \dots, n\}$ gelten, d.h. die beiden n -Tupel sind identisch, ein Widerspruch zur Annahme μ_1, \dots, μ_n sei ein anderes n -Tupel als $\lambda_1, \dots, \lambda_n$. \square

Dimension

Satz 5.20

Hat ein Vektorraum V eine Basis $\{\vec{v}_1, \dots, \vec{v}_n\}$ mit $n \in \mathbb{N}$, so gilt: jede Basis von V hat genau n Elemente.

Satz 5.20 folgt direkt aus Lemma 5.25, dessen Beweis ist allerdings etwas technisch, weswegen wir ihn am Ende dieses Abschnittes führen. Satz 5.20 motiviert die folgende Definition:

Definition 5.21

Hat ein Vektorraum V eine Basis $\mathcal{B} = \{\vec{v}_1, \dots, \vec{v}_n\}$ mit $n \in \mathbb{N}$ so sagt man die Dimension von V ist n .
Ist es nicht möglich, eine (endliche) Basis von V zu finden, so sagt man die Dimension von V ist unendlich.

Beispiel 5.22

- Die Menge $\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$ ist eine Basis des $\mathbb{R}^3 = \left\{ \begin{pmatrix} a \\ b \\ c \end{pmatrix} : a, b, c \in \mathbb{R} \right\}$,
d.h. der Vektorraum \mathbb{R}^3 hat die Dimension 3.
- Die Menge $\{1, x, x^2\}$ ist eine Basis von $\mathbb{R}[x]_2 := \{ax^2 + bx + c : a, b, c \in \mathbb{R}\}$,
d.h. der Vektorraum $\mathbb{R}[x]_2$ hat die Dimension 3.
- Die Menge $\mathbb{R}[x]$ aller Polynome ist ein Vektorraum, der *keine* endliche Basis besitzt,
d.h. der Vektorraum $\mathbb{R}[x]$ hat die Dimension unendlich.

Satz 5.23

Jeder n -dimensionale Vektorraum ist isomorph zu \mathbb{R}^n (wobei $n \in \mathbb{N} \setminus \{0\}$).

Beweis: Es sei V ein n -dimensionaler Vektorraum, dann besitzt V eine Basis $\{v_1, \dots, v_n\}$.
Zu jedem Vektor $w \in V$ gibt es genau ein n -Tupel $(a_{w,1}, \dots, a_{w,n}) \in \mathbb{R}^n$ mit $w = a_{w,1}v_1 + \dots + a_{w,n}v_n$.

Die Abbildung $f : V \rightarrow \mathbb{R}^n$ mit $w \mapsto f(w) := (a_{w,1}, \dots, a_{w,n})$ ist bijektiv:

Injektiv. : Klar ist: Für $w \neq \tilde{w}$ ist $(a_{w,1}, \dots, a_{w,n}) \neq (a_{\tilde{w},1}, \dots, a_{\tilde{w},n})$.

Surjektiv. : Umgekehrt gilt: Für jeden Vektor $\vec{a} \in \mathbb{R}^n$ ist $u := a_1 v_1 + \dots + a_n v_n$ ein Vektor mit $u \in V$ und $f(u) = \vec{a}$.

□

Basis-Austausch Sätze

Die folgenden zwei Lemmata sind von Ihren Aussagen nur scheinbar rein technische Hilfs-Sätze. Das Lemma 5.25 zeigt insbesondere:

- dass es Sinn macht von “ n -dimensionalen” Vektorräumen zu sprechen und
- dass im n -dimensionalen Vektorraum eine Menge von linear unabhängige Vektoren höchstens n Elemente haben kann.

Insgesamt beweisen Lemma 5.24 und 5.25 den Satz 5.20.

Lemma 5.24

Sei $\{v_1, \dots, v_k\}$ eine Basis des Vektorraumes V , und $z = a_1 v_1 + \dots + a_k v_k$ eine Linearkombination mit $a_j \neq 0$. Dann ist auch $\{v_1, \dots, v_{j-1}, z, v_{j+1}, \dots, v_k\}$ eine Basis von V .

Beweis: Nach Umnummerierung der v_i dürfen wir annehmen, dass in $z = a_1 v_1 + \dots + a_k v_k$ gilt: $a_1 \neq 0$. Wir zeigen, dass in diesem Fall $\{z, v_2, \dots, v_k\}$ eine Basis von V ist:

Lineare Unabhängigkeit

Angenommen es gäbe Zahlen $\lambda_1, \dots, \lambda_k$ so dass gilt:

$$\lambda_1 z + \lambda_2 v_2 + \dots + \lambda_k v_k = \mathbf{0}.$$

Setzt man hier $\lambda_1 z = \lambda_1 a_1 v_1 + \dots + \lambda_1 a_k v_k$ ein, so erhält man:

$$\lambda_1 a_1 v_1 + (\lambda_2 + \lambda_1 a_2) v_2 + \dots + (\lambda_k + \lambda_1 a_k) v_k = \mathbf{0}.$$

Da $\{v_1, \dots, v_k\}$ linear unabhängig sind, sind hier alle Koeffizienten Null, d.h. insbesondere gilt $\lambda_1 a_1 = 0$. Wegen der Annahme $a_1 \neq 0$ folgt sofort $\lambda_1 = 0$, und dies liefert

$$\lambda_2 v_2 + \dots + \lambda_k v_k = \mathbf{0}.$$

und damit $\lambda_2 = \dots = \lambda_k = 0$, da $\{v_1, \dots, v_k\}$ linear unabhängig sind. Folglich sind $\{z, v_1, \dots, v_k\}$ linear unabhängig.

Spann ist ganz V

Es gibt Zahlen $c_1, \dots, c_k \in \mathbb{R}$ mit $v_1 = c_1 z + c_2 v_2 + \dots + c_k v_k$, denn löst man $z = a_1 v_1 + a_2 v_2 + \dots + a_k v_k$ nach v_1 auf, so erhält man:

$$v_1 = \frac{1}{a_1} z - \frac{a_2}{a_1} v_2 - \dots - \frac{a_k}{a_1} v_k$$

Da $\{v_1, \dots, v_k\}$ eine Basis von V ist, lässt sich jedes $w \in V$ darstellen als Linearkombination:

$$w = \mu_1 v_1 + \mu_2 v_2 + \dots + \mu_k v_k \quad \text{mit } \mu_1, \dots, \mu_k \in \mathbb{R}.$$

Ersetzt man hier $v_1 = c_1 z + c_2 v_2 + \dots + c_k v_k$, so erhält man für w eine Linearkombination aus z, v_2, \dots, v_k , nämlich:

$$w = \mu_1 (c_1 z + c_2 v_2 + \dots + c_k v_k) + \mu_2 v_2 + \dots + \mu_k v_k$$

Es gilt also $w \in \text{spann}(z, v_2, \dots, v_k)$. □

Lemma 5.25 (Basisaustauschsatz von Steinitz)

Es seien $\{w_1, \dots, w_k\} \subseteq V$ linear unabhängig, und es sei $\{v_1, \dots, v_n\}$ eine Basis von V .

Dann gilt $k \leq n$, und es gibt $n - k$ paarweise verschiedene $\tilde{v}_{k+1}, \dots, \tilde{v}_n \in \{v_1, \dots, v_n\}$, so dass gilt:

$$\{w_1, \dots, w_k, \tilde{v}_{k+1}, \dots, \tilde{v}_n\} \quad \text{ist eine Basis von } V$$

Beweis: Wir führen eine Induktion über k , beginnend bei $k = 1$.

I. Anf: Es sei $\{w_1\}$ linear unabhängig. Der Vektor w_1 lässt sich darstellen als $w_1 = a_1 v_1 + \dots + a_n v_n$ mit $a_1, \dots, a_n \in \mathbb{R}$. Es gilt $w_1 \neq 0$, weil $\{w_1\}$ linear unabhängig ist. Also gibt es ein $a_i \neq 0$. Lemma 5.24 beweist: $\{v_1, \dots, v_{i-1}, w_1, v_{i+1}, \dots, v_n\}$ ist eine Basis.

I. Vor: Die Aussage gelte für ein ℓ mit $1 \leq \ell$.

I. Schritt: Es sei $\{w_1, \dots, w_\ell, w_{\ell+1}\}$ linear unabhängig, entsprechend sind $\{w_1, \dots, w_\ell\}$ linear unabhängig. Nach I. Vor. gibt es $u_i \in \{v_1, \dots, v_n\}$, so dass $\{w_1, \dots, w_\ell, u_{\ell+1}, \dots, u_n\}$ eine Basis von V ist. Insbesondere lässt sich der Vektor $w_{\ell+1}$ darstellen als

$$w_{\ell+1} = a_1 w_1 + \dots + a_\ell w_\ell + b_{\ell+1} u_{\ell+1} + \dots + b_n u_n$$

Da $\{w_1, \dots, w_\ell, w_{\ell+1}\}$ linear unabhängig sind, ist $w_{\ell+1}$ keine Linearkombination der restlichen w_i (s. Lemma 5.15). D.h. es gibt ein $j \geq \ell + 1$ mit $b_j \neq 0$.

Nummeriert man die Vektoren u_i so, dass $b_{k+1} \neq 0$ gilt, so zeigt Lemma 5.24, dass $\{w_1, \dots, w_\ell, w_{\ell+1}, u_{\ell+2}, \dots, u_n\}$ eine Basis von V ist.

Noch zu zeigen: $k \leq n$ (s. folgendes FAQ)

Annahme: Es seien $W := \{w_1, \dots, w_k\} \subset V$ linear unabhängig. Per Induktion haben wir bewiesen: Weil die Vektoren in W linear unabhängig sind, lässt sich W zu einer Basis der Länge n auffüllen^a. Es folgt also insbesondere $|W| \leq n$, d.h. es folgt $k \leq n$.

^a(mit Vektoren aus $\mathcal{B} := \{v_1, \dots, v_n\}$) □

FAQ: Eine Induktion über k und am Schluss wird k eingeschränkt?!

Frage: Eine Induktion über k beweist doch eine Aussage für *alle* $k \in \mathbb{N}$.

Hier kommt im zweiten Teil des Beweises aber eine Einschränkung $k < n$. Geht das überhaupt?

Antwort: Die Induktion beweist eine “Wenn-Dann-Aussage”, deren “Wenn-Teil” für $k > n$ einfach nicht mehr eintritt:

“Wenn $W := \{w_1, \dots, w_k\}$ linear unabhängig sind,
dann lässt sich W zu einer Basis der Länge n ergänzen.”

Diese “Wenn-Dann-Aussage” ist *tatsächlich* richtig für *alle* $k \in \mathbb{N} \setminus \{0\}$, und das beweist die Induktion.

Allerdings gibt es einen kleinen “Twist”: Die Aussage gilt ab $k > n$ *trivialerweise*, weil die Voraussetzung “ $\{w_1, \dots, w_k\}$ ist linear unabhängig” unerfüllbar (also immer falsch) ist. Dies ist dann die Aussage des zweiten Teils des Beweises:

Ab $k > n$ gilt für eine Menge mit k -vielen Vektoren: Die Vektoren sind *nicht* linear unabhängig.

Lemma 5.25 beweist also:

- für $1 \leq k \leq n - 1$:

Eine (tatsächlich) linear unabhängige Menge $W = \{w_1, \dots, w_k\}$ lässt sich *tatsächlich* zu einer Basis der Länge n auffüllen.

- für $k = n$:

Eine (tatsächlich) linear unabhängige Menge $W = \{w_1, \dots, w_n\}$ ist *bereits* eine Basis der Länge n , denn “zu einer Basis der Länge n auffüllen” bedeutet hier: *keinen* Vektor aus $\{v_1, \dots, v_n\}$ dazufügen, weil die Länge n bereits erreicht ist.

- für $k > n$:

Jede *hypothetisch* linear unabhängige Menge $W = \{w_1, \dots, w_k\}$, *ließe* sich zu einer Basis der Länge n auffüllen. Da W aber schon mehr als n Elemente hat, kann dies nicht gehen, d.h. es gibt solche Mengen W nicht.

5.4 Unterräume

Definition 5.27

Es sei V ein Vektorraum. Eine Teilmenge $W \subseteq V$ ist ein *Unter-Vektorraum* von V (kurz *Unterraum*), wenn W ein Vektorraum ist.

Korollar 5.27

Für einen Unterraum $W \subseteq V$ des Vektorraums V gelten: $\mathbf{0} \in W$ und $x, y \in W \Rightarrow x + y \in W$.

Eine Teilmenge $W \subseteq V$ “erbt” vom Vektorraum V (automatisch) die Rechenregeln:

Dies bedeutet, dass in W zum Beispiel weiterhin das Assoziativgesetz gilt, d.h. $(x + y) + z = x + (y + z)$ gilt für alle $x, y, z \in W$. Ebenso gelten auch weiterhin alle Verträglichkeitsregeln V2.1) bis V2.4) aus den Vektorraumaxiomen in W .

Um also zu zeigen, dass eine Teilmenge $W \subseteq V$ ein (Unter-)Vektorraum ist, genügt es zu zeigen, dass W abgeschlossen ist bezgl. Addition und die skalarer Multiplikation:

Lemma 5.28 (Charakterisierung eines Unterraumes)

Es sei V ein Vektorraum mit Vektoraddition $+$ und skalarer Multiplikation \cdot .
Eine Teilmenge $W \subseteq V$ ist ein Unterraum von V genau dann wenn gelten:

- a) Für alle $x, y \in W$ gilt $x + y \in W$
- b) Für alle $\lambda \in \mathbb{R}, x \in W$ gilt $\lambda \cdot x \in W$

Beweis: Es sei $W \subseteq V$ eine Teilmenge des Vektorraums V .

“ \Rightarrow ” Ist W ein Vektorraum, so gelten nach Def. 5.3 sofort a) und b).

“ \Leftarrow ” Es seien a) und b) erfüllt. Wir zeigen nun: W ist zusammen mit $+$ und \cdot ein Vektorraum. D.h. wir zeigen: $(W, +)$ ist eine abelsche Gruppe, und die Anforderungen an die skalare Multiplikation sind erfüllt (Abgeschlossenheit & Verträglichkeit).

$(W, +)$ ist abelsche Gruppe

Da alle Elemente von $v, w \in W$ auch Elemente von V sind folgt sofort:

Die Vektoraddition ist assoziativ und symmetrisch, d.h. die Gruppenaxiome **G2)** und **G_{symm})** sind auch in W erfüllt.

G1) Wegen a) gilt: W ist abgeschlossen unter Vektoraddition, d.h. Gruppenaxiom **G1)** ist in W erfüllt.

G3) Wegen b) gilt: für $w \in W$ ist auch $0 = 0 \cdot w \in W$. D.h. das neutrale Element von $(W, +)$ ist in W .

G4) Wegen b) gilt: für jedes $w \in W$ ist auch $-w = (-1) \cdot w \in W$. D.h. zu jedem $w \in W$ gibt es auch das additiv Inverse $-w \in W$.

Voraussetzungen an die skalare Multiplikation

- **Abgeschlossenheit.** Wegen b) wissen wir: W ist abgeschlossen unter skalarer Multiplikation, d.h. Vektorraum-Voraussetzungen an “ \cdot ” sind erfüllt.
- **Verträglichkeit V2).** Da alle Elemente von $w \in W$ auch Elemente von V sind folgt sofort:
Wie in V gelten auch in W die Verträglichkeitsregeln V2.1) bis V2.4) aus den Vektorraumaxiomen.

□

Zu erkennen, ob eine Teilmenge $W \subseteq V$ ein Untervektorraum eines Vektorraumes V ist, erfordert also “legiglich” zu prüfen, ob W abgeschlossen ist unter Addition und skalarer Multiplikation. Ein erster Anhaltspunkt ist das “Enthaltensein der Null”:

Gilt $0 \notin W$ so ist W kein Vektorraum und damit auch kein Untervektorraum.

Beispiel 5.29

Es sei $\{v_1, \dots, v_k\} \subset V$ eine endliche Teilmenge eines Vektorraumes V , dann ist $W := \text{spann}(v_1, \dots, v_k)$ ein Untervektorraum von V .

Beispiel 5.30

1. Der Vektorraum $\mathbb{R} = \mathbb{R}^1$ hat nur zwei Untervektorräume: sich selbst und die Menge $\{0\}$, die nur den Nullvektor enthält.
2. Die Untervektorräume von \mathbb{R}^2 sind genau
 - die Menge $\{0\}$,
 - alle Mengen der Form $V_a = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 : y = a \cdot x \right\}$ mit $a \in \mathbb{R}$,
 - die Menge $V_\infty = \left\{ \begin{pmatrix} 0 \\ y \end{pmatrix} : y \in \mathbb{R} \right\}$ und
 - der gesamte Vektorraum \mathbb{R}^2 selbst.

Geometrisch ist eine Menge V_a mit festem a eine Gerade durch den Ursprung $0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ (genauso auch V_∞).

3. Die Untervektorräume des \mathbb{R}^3 sind entsprechend die Mengen $\{0\}$ und \mathbb{R}^3 selbst sowie die Geraden und Ebenen durch 0 .

(Anti-)Beispiel 5.31

Die Menge $W := \left\{ \begin{pmatrix} a \\ b \end{pmatrix} \in \mathbb{R}^2 : a, b \geq 0 \right\}$ ist abgeschlossen unter Vektor-Addition, aber trotzdem kein Untervektorraum von \mathbb{R}^2 :

Für $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ ist $\begin{pmatrix} -1 \\ -1 \end{pmatrix} \notin W$, das additiv Inverse zu $\begin{pmatrix} 1 \\ 1 \end{pmatrix} \in W$ ist also nicht Teil der Menge W .

Kapitel 6

Lineare Abbildungen

6.1 Matrizen

Im den vorherigen Abschnitten haben wir gezeigt, dass ein endlich-dimensionaler Vektorraum V stets isomorph ist zu einem \mathbb{R}^n . Im folgenden werden wir uns daher vorerst mit linearen Abbildungen der Form $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ beschäftigen. Diese linearen Abbildungen lassen sich sehr kurz und knapp durch eine Matrix beschreiben.

Matrizen (Einzahl: *Matrix*) sind ein Schlüsselkonzept der linearen Algebra und tauchen in fast allen Gebieten der Mathematik auf. Sie stellen Zusammenhänge, in denen Linearkombinationen eine Rolle spielen, übersichtlich dar und werden insbesondere benutzt, um lineare Abbildungen darzustellen.

Eine *Matrix* $A \in \mathbb{R}^{m \times n}$ ist grob gesagt eine Tabelle von m mal n Zahlen, die in einem rechteckigen Schema von m Zeilen und n Spalten angeordnet sind. In $A \in \mathbb{R}^{m \times n}$ steht die **erste Dimensions-Variable** „ m “ für die **Höhe** der Matrix. Dies kann man sich merken, indem man sich vorstellt, dass ein (virtueller) Leser, der von links nach rechts „angelaufen kommt“ immer zuerst die Höhe der Matrix wahrnimmt.

Definition 6.1 (Allgemeine Form einer Matrix)

Eine $m \times n$ -Matrix $A \in \mathbb{R}^{m \times n}$ ist ein Zahlenschema mit m Zeilen und n Spalten. Dabei ist $A_{ij} \in \mathbb{R}$ jeweils den Eintrag in der i -ten Zeile und der j -ten Spalte:

Der Leser nimmt immer zuerst die Höhe der Matrix wahr, deswegen steht m vorne.



Höhe m

Breite n

$$A = \begin{pmatrix} A_{1,1} & A_{1,2} & A_{1,3} & \cdots & A_{1,n} \\ A_{2,1} & A_{2,2} & A_{2,3} & \cdots & A_{2,n} \\ A_{3,1} & A_{3,2} & A_{3,3} & \cdots & A_{3,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ A_{m,1} & A_{m,2} & A_{m,3} & \cdots & A_{m,n} \end{pmatrix}$$

Die Einträge $A_{ij} \in \mathbb{R}$ einer Matrix können beliebige Zahlen aus \mathbb{R} sein, es ist aber unzulässig eine Position leer zu lassen.

Eine Matrix mit m Zeilen und n Spalten nennt man $m \times n$ -Matrix (sprich „ m -Kreuz- n -Matrix“).

Beispiel 6.2

Sei $A = \begin{pmatrix} 1 & 2 \\ 5 & 3 \\ 7 & 4 \end{pmatrix}$. Dann ist A eine 3×2 -Matrix (d.h. $A \in \mathbb{R}^{3 \times 2}$) und es gilt

$A_{1,1} = 1$	$A_{1,2} = 2$
$A_{2,1} = 5$	$A_{2,2} = 3$
$A_{3,1} = 7$	$A_{3,2} = 4$

Matrizen sind sehr nützliche Hilfsmittel in einer Vielzahl von Anwendungen. Sie eignen sich als Kurzschreibweise für größere Mengen von Daten. Die wahrscheinlich wichtigste solcher Anwendungen sind lineare Gleichungssysteme.

Beispiel 6.3

Betrachten wir die folgenden beiden linearen Gleichungen:

$$\begin{aligned} 4x_1 + 6x_2 - 8x_3 &= 0 \\ -2x_2 - 8x_3 &= 0 \end{aligned}$$

Die wichtige Information dieses Systems steckt lediglich in den Koeffizienten der beiden Gleichungen. Wir können diese in einer Matrix A zusammenfassen, indem wir im Eintrag A_{ij} den Koeffizienten von x_j in der i -ten Gleichung schreiben. Taucht x_j in der i -ten Gleichung nicht auf, so setzen wir $A_{ij} = 0$. Hier lautet die Matrix A also

$$A = \begin{pmatrix} 4 & 6 & -8 \\ 0 & -2 & -8 \end{pmatrix}.$$

Mit den Rechenregeln, die wir in Kürze lernen werden, können wir das Gleichungssystem dann folgendermaßen schreiben:

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 4 & 6 & -8 \\ 0 & -2 & -8 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

6.2 Rechnen mit Matrizen

Definition 6.5 (Transposition)

Es sei $A \in \mathbb{R}^{m \times n}$ eine Matrix.

Die zu A *transponierte* Matrix $A^T \in \mathbb{R}^{n \times m}$ ist die Matrix, mit den Einträgen $(A^T)_{k\ell} = A_{\ell k}$.

Die *Spalten* der Matrix A (von oben nach unten gelesen) werden zu *Zeilen* der Matrix A^T (von links nach rechts gelesen)

Beispiel 6.5

Aus $A \in \mathbb{R}^{3 \times 2}$ wird wie folgt eine Matrix $A^T \in \mathbb{R}^{2 \times 3}$

$$A = \begin{pmatrix} 1 & 7 \\ 2 & 8 \\ 3 & 9 \end{pmatrix} \Rightarrow A^T = \begin{pmatrix} 1 & 2 & 3 \\ 7 & 8 & 9 \end{pmatrix}$$

Liest man einen Vektor als eine $\mathbb{R}^{n \times 1}$ -Matrix so kann man diesen auch Transponieren. Aus einem "stehenden" Vektor $v \in \mathbb{R}^3$ wird so ein "liegender Vektor":

$$v = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \Rightarrow v^T = (1 \ 2 \ 3)$$

Addition

Matrizen mit gleichen Dimensionen lassen sich addieren. Die Addition funktioniert komponentenweise:

Definition 6.8

Es seien $A, B \in \mathbb{R}^{m \times n}$. Die Matrix $C := A + B$ ist die Matrix mit den Einträgen $C_{ij} = A_{ij} + B_{ij}$.

Beispiel 6.7

$$\begin{pmatrix} 0 & 6 \\ 2 & 8 \\ 4 & 10 \end{pmatrix} + \begin{pmatrix} 1 & 7 \\ 3 & 9 \\ 5 & 11 \end{pmatrix} = \begin{pmatrix} 0+1 & 6+7 \\ 2+3 & 8+9 \\ 4+5 & 10+11 \end{pmatrix} = \begin{pmatrix} 1 & 13 \\ 5 & 17 \\ 9 & 21 \end{pmatrix}$$

(Anti-)Beispiel 6.8

$$\begin{pmatrix} 3 & 5 & 1 \\ 2 & 1 & 7 \end{pmatrix} + \begin{pmatrix} 2 & 6 \\ 1 & 2 \end{pmatrix} \text{ ist nicht definiert.}$$

Ganz analog definieren wir natürlich die Subtraktion $A - B$ ganz einfach als die komponentenweise Subtraktion aller Einträge. Die Addition von zwei Matrizen ist nur definiert, wenn sie beide die gleiche Anzahl von Zeilen und auch die gleiche Anzahl von Spalten haben.

Multiplikation mit Skalaren

Die Multiplikation einer Matrix A mit einem Skalar $\lambda \in \mathbb{R}$ funktioniert genauso wie bei Vektoren: Man multipliziert jeden Eintrag von A mit λ .

Definition 6.10

Es sei $A \in \mathbb{R}^{m \times n}$ und $\lambda \in \mathbb{R}$. Dann ist $B := \lambda \cdot A \in \mathbb{R}^{m \times n}$ die Matrix mit den Einträgen $B_{ij} = \lambda \cdot a_{ij}$.

Beispiel 6.10

$$5 \cdot \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} = \begin{pmatrix} 5 \cdot 1 & 5 \cdot 2 & 5 \cdot 3 \\ 5 \cdot 4 & 5 \cdot 5 & 5 \cdot 6 \end{pmatrix} = \begin{pmatrix} 5 & 10 & 15 \\ 20 & 25 & 30 \end{pmatrix}$$

Bei der Multiplikation einer Matrix mit einem Skalar müssen wir uns keine Gedanken um passende Zeilen- und Spaltenanzahl machen. Diese Multiplikation ist immer definiert. Es gelten die folgenden Rechenregeln:

Rechenregeln für Addition und Multiplikation mit Skalaren

Seien $A, B, C \in \mathbb{R}^{m \times n}$ drei $m \times n$ -Matrizen und seien $\lambda, \mu \in \mathbb{R}$ Skalare. Dann gelten:

- $A + B = B + A$ (Kommutativgesetz der Addition)
 - $(A + B) + C = A + (B + C)$ (Assoziativgesetz der Addition)
-
- $\lambda(\mu A) = (\lambda\mu)A = \mu(\lambda A)$ (Assoziativgesetz der Multiplikation)
 - $(\lambda + \mu) \cdot A = \lambda A + \mu A$
 - $\lambda \cdot (A + B) = \lambda A + \lambda B$

Die Menge $\mathbb{R}^{n \times m}$ bildet zusammen mit der Addition und der skalaren Multiplikation einen Vektorraum.

Multiplikation mit Vektoren

Die Multiplikation einer Matrix A mit einem Vektor \vec{v} ist so definiert, dass die in A gespeicherten Koeffizienten wieder an die entsprechenden Einträge von \vec{v} multipliziert werden. Das Berechnen von $A \cdot \vec{v}$ erfolgt also zeilenweise, für jede Zeile von A wird eine Summe berechnet:

Matrix-Vektor-Multiplikation

Für eine Matrix $A \in \mathbb{R}^{m \times n}$ mit n Spalten und einen Vektor $\vec{v} \in \mathbb{R}^n$ mit n Einträgen gilt:

$$A \cdot \vec{v} = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{m1} & A_{m2} & \cdots & A_{mn} \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} A_{11} \cdot v_1 + A_{12} \cdot v_2 + \cdots + A_{1n} \cdot v_n \\ A_{21} \cdot v_1 + A_{22} \cdot v_2 + \cdots + A_{2n} \cdot v_n \\ \vdots \\ A_{m1} \cdot v_1 + A_{m2} \cdot v_2 + \cdots + A_{mn} \cdot v_n \end{pmatrix}$$

Das Ergebnis der Multiplikation $A \cdot \vec{v}$ ist also ein Vektor aus \mathbb{R}^m .

In Beispiel 6.3 haben wir bereits eine Multiplikation von einer Matrix A mit einem Vektor \vec{x} gesehen. Dort wurde die Multiplikation verwendet, um die linke Seite eines Gleichungssystems in Kurzschreibweise zu notieren.

Beispiel 6.11

$$\begin{pmatrix} 9 & 7 & 5 \\ 8 & 6 & 4 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 9 \cdot 1 + 7 \cdot 2 + 5 \cdot 3 \\ 8 \cdot 1 + 6 \cdot 2 + 4 \cdot 3 \end{pmatrix} = \begin{pmatrix} 38 \\ 32 \end{pmatrix}$$

Interpretation der Matrixmultiplikation

Die Multiplikation eines Vektors \vec{v} mit einer Matrix A lässt sich auf zwei Weisen verstehen:

a) Skalarprodukt mit den Zeilen von A :

Es wird *zeilenweise* das Skalarprodukt der i -ten Zeile von A mit \vec{v} berechnet.

$$\begin{pmatrix} a & b & c \\ 1 & 2 & 3 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} a \cdot x + b \cdot y + c \cdot z \\ 1 \cdot x + 2 \cdot y + 3 \cdot z \end{pmatrix} = \begin{pmatrix} \left\langle \begin{pmatrix} a \\ b \\ c \end{pmatrix}, \begin{pmatrix} x \\ y \\ z \end{pmatrix} \right\rangle \\ \left\langle \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} x \\ y \\ z \end{pmatrix} \right\rangle \end{pmatrix}$$

b) Linearkombination der Spalten von A :

Es werden Vielfache der *Spalten* von A addiert – mit Vorfaktoren aus \vec{v} . Der Vektor \vec{v} ist also eine *Linearkombinations-Anweisung* für die Spalten von A .

$$\begin{pmatrix} a & b & c \\ 1 & 2 & 3 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} a \cdot x + b \cdot y + c \cdot z \\ 1 \cdot x + 2 \cdot y + 3 \cdot z \end{pmatrix} = x \cdot \begin{pmatrix} a \\ 1 \end{pmatrix} + y \cdot \begin{pmatrix} b \\ 2 \end{pmatrix} + z \cdot \begin{pmatrix} c \\ 3 \end{pmatrix}$$

6.3 Matrix-Matrix-Multiplikation

Üblicherweise wird die Multiplikation $A \cdot B$ einer Matrix $A \in \mathbb{R}^{k \times n}$ mit einer Matrix $B \in \mathbb{R}^{n \times \ell}$ „von rechts nach links“ durchgeführt. D.h. die Matrix A wird als Liste von Spalten aufgefasst, die jeweils mit B multipliziert werden:

Definition 6.12

Es sei $A \in \mathbb{R}^{k \times m}$ mit „Inputdimension“ m
und $B \in \mathbb{R}^{m \times n}$ mit n Spalten $\vec{b}_1, \vec{b}_2, \dots, \vec{b}_n \in \mathbb{R}^m$:

$$A = \begin{pmatrix} \overbrace{A_{11} \dots A_{1m}}^{\text{Breite } m} \\ \vdots \\ A_{k1} \dots A_{km} \end{pmatrix} \quad \begin{matrix} \text{Höhe } m \end{matrix} \quad B = \begin{pmatrix} \begin{matrix} \vec{b}_1 & \vec{b}_2 & \dots & \vec{b}_n \end{matrix} \end{pmatrix}$$

Die Matrix $C := A \cdot B$ ist die $k \times n$ -Matrix mit den Einträgen $C_{i,j} := \langle \vec{a}_i, \vec{b}_j \rangle$.

Das heißt die Matrix C hat n -viele Spalten der Form $A \cdot \vec{b}_j \in \mathbb{R}^k$.

$$A \cdot B = \begin{pmatrix} A \cdot \vec{b}_1 & A \cdot \vec{b}_2 & \dots & A \cdot \vec{b}_n \end{pmatrix}$$



Das Produkt $A \cdot B$ zweier Matrizen A und B kann nur dann gebildet werden, wenn die **Spaltenanzahl von A** gleich der **Zeilenanzahl von B** ist.

Die Matrix $A \cdot B$ „erbt“ von A die „Output-Dimension“ und von B die „Input-Dimension“.

Die Zwischen-Dimension m geht bei der Matrixmultiplikation verloren:

Ist $A \in \mathbb{R}^{k \times m}$ mit „Output-Dimension“ k und „Input-Dimension“ m

und $B \in \mathbb{R}^{m \times n}$ mit „Output-Dimension“ m und „Input-Dimension“ n

So ist $A \cdot B \in \mathbb{R}^{k \times n}$ mit „Output-Dimension“ k und „Input-Dimension“ n

Beispiel 6.13

Sei

$$A := \begin{pmatrix} 1 & 5 & 9 \\ 2 & 6 & 9 \end{pmatrix} \quad \text{und} \quad B := \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Die Matrix A hat 3 Spalten und B hat 3 Zeilen.

Da diese Zahlen gleich sind, können wir das Produkt $A \cdot B$ berechnen:

$$A \cdot B = \left(A \cdot \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, A \cdot \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \right) = \begin{pmatrix} 10 & 6 \\ 11 & 8 \end{pmatrix}$$

Korollar 6.14

Es seien $A \in \mathbb{R}^{k \times m}$, $B \in \mathbb{R}^{m \times n}$ und $x \in \mathbb{R}^n$. Dann gilt: $A \cdot (B \cdot x) = (A \cdot B) \cdot x$

Beweis:

Es sei $A \in \mathbb{R}^{k \times m}$ mit "Inputdimension" m
 und $B \in \mathbb{R}^{m \times n}$ mit n Spalten $\vec{b}_1, \vec{b}_2, \dots, \vec{b}_n \in \mathbb{R}^m$. Weiter sei $\vec{x} \in \mathbb{R}^n$.
 Dann gilt $B \cdot \vec{x} = x_1 \cdot \vec{b}_1 + \dots + x_n \cdot \vec{b}_n \in \mathbb{R}^m$.

$$\begin{aligned} A \cdot (B \cdot \vec{x}) &= A \cdot (x_1 \cdot \vec{b}_1 + \dots + x_n \cdot \vec{b}_n) \\ &\stackrel{*}{=} x_1 \cdot A \cdot \vec{b}_1 + \dots + x_n \cdot A \cdot \vec{b}_1 && (\star \text{ Linearitat von } A \cdot y) \\ &= \underbrace{(A \cdot \vec{b}_1 \ \dots \ A \cdot \vec{b}_n)}_{\text{Matrix mit Spalten } A \cdot \vec{b}_i} \cdot \vec{x} = (A \cdot B) \cdot \vec{x} \end{aligned}$$

□

Die Rechenregeln fur die Multiplikation (Lemma 6.15) von Matrizen sehen im Pinzip aus, wie Rechnen mit Zahlen aus \mathbb{R} . Allerdings gilt fur Matrizen das Kommutativgesetz nicht! D.h. im Allgemeinen ist $A \cdot B \neq B \cdot A$. Selbst wenn beide Produkte definiert sind, gilt nicht immer die Gleichheit.

Lemma 6.15 (Rechenregeln fur die Matrix-Multiplikation)

Es seien $A, \tilde{A}, B, \tilde{B}, C$ Matrizen mit passenden Dimensionen, d.h.

$$A, \tilde{A} \in \mathbb{R}^{m \times n}, \quad B, \tilde{B} \in \mathbb{R}^{n \times k} \quad \text{und} \quad C \in \mathbb{R}^{k \times \ell}$$

Dann gelten:

- 1) $(A \cdot B) \cdot C = A \cdot (B \cdot C)$ (Assoziativgesetz)
- 2) $A \cdot (B + \tilde{B}) = A \cdot B + A \cdot \tilde{B}$ (Distributivgesetz)
 $(A + \tilde{A}) \cdot B = A \cdot B + \tilde{A} \cdot B$
- 3) $A \cdot (\lambda \cdot B) = \lambda \cdot A \cdot B$ gilt fur alle $\lambda \in \mathbb{R}$

Im Allgemeinen gilt:

$$A \cdot B \neq B \cdot A \quad (\text{Kommutativgesetz gilt im Allgemeinen nicht})$$

Beweis: Die Aussagen in 2) und 3) lassen sich direkt aus der Definition der Matrixmultiplikation ableiten.

Zu 1) Es sei $C = (\vec{c}_1, \dots, \vec{c}_\ell)$ mit Spalten $c_i \in \mathbb{R}^k$.

$$\begin{aligned} A \cdot (B \cdot C) &= A \cdot (B \cdot \vec{c}_1 \ \dots \ B \cdot \vec{c}_\ell) = \begin{pmatrix} A \cdot (B \cdot \vec{c}_1) & \dots & A \cdot (B \cdot \vec{c}_\ell) \end{pmatrix} \\ &\stackrel{*}{=} \begin{pmatrix} (A \cdot B) \cdot \vec{c}_1 & \dots & (A \cdot B) \cdot \vec{c}_\ell \end{pmatrix} = (A \cdot B) \cdot C \end{aligned}$$

Denn bei (\star) gilt nach Korollar 6.14 Assoziativitat: $A \cdot (B \cdot \vec{x}) = (A \cdot B) \cdot \vec{x}$ gilt fur alle $\vec{x} \in \mathbb{R}^k$.

□

Beispiel 6.16 (Kommutativgesetz gilt nicht)

Fur die Matrizen $A := \begin{pmatrix} 3 & 5 \\ 0 & 0 \end{pmatrix}$ und $B := \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ gilt $A \cdot B \neq B \cdot A$, denn:

$$A \cdot B = \begin{pmatrix} A \cdot \begin{pmatrix} 0 \\ 0 \end{pmatrix} & A \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} \end{pmatrix} = \begin{pmatrix} 0 & 3 \\ 0 & 0 \end{pmatrix} \quad \text{und} \quad B \cdot A = \begin{pmatrix} B \cdot \begin{pmatrix} 3 \\ 0 \end{pmatrix} & B \cdot \begin{pmatrix} 5 \\ 0 \end{pmatrix} \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Interpretieren wir Matrizen wiederum als lineare Abbildungen, so entspricht das Produkt zweier Matrizen der Hintereinanderschaltung zweier linearer Abbildungen:

6.4 Lineare Abbildungen vs. Matrix-Abbildungen

Satz 6.17

Jede lineare Abbildung $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ lässt sich als eine Matrix-Multiplikation $f(x) = A \cdot x$ mit $A \in \mathbb{R}^{m \times n}$ beschreiben. Genauer gilt:

Die Matrix A hat die Spalten $\vec{a}_1 := f(e_1), \vec{a}_2 := f(e_2), \dots, \vec{a}_n := f(e_n)$ (wobei $\vec{a}_i \in \mathbb{R}^m$), d.h.

$$A = \begin{pmatrix} \begin{array}{|c|} \hline \vec{a}_1 \\ \hline \end{array} & \begin{array}{|c|} \hline \vec{a}_2 \\ \hline \end{array} & \dots & \begin{array}{|c|} \hline \vec{a}_n \\ \hline \end{array} \end{pmatrix}$$

Die Matrix A heißt *darstellende Matrix* von f , man schreibt kurz $M(f) := (f(e_1), \dots, f(e_n))$.

Umgekehrt gilt: Ist $B \in \mathbb{R}^{m \times n}$, so ist die Abbildung $g : \mathbb{R}^n \rightarrow \mathbb{R}^m, x \mapsto B \cdot x$ linear.

Beweis: Es sei $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ und $A \in \mathbb{R}^{m \times n}$ die Matrix mit Spalten $\vec{a}_1 := f(e_1), \dots, \vec{a}_n := f(e_n)$. Für $x \in \mathbb{R}^n$ gilt dann wegen $x = x_1 \cdot e_1 + \dots + x_n e_n$ und wegen der Linearität von f :

$$\begin{aligned} f(\vec{x}) = f(x_1 \cdot e_1 + \dots + x_n \cdot e_n) &= x_1 \cdot f(e_1) + \dots + x_n \cdot f(e_n) \\ &= x_1 \cdot \vec{a}_1 + \dots + x_n \cdot \vec{a}_n = A \cdot \vec{x} \end{aligned}$$

□

Beispiel 6.18

Es sei $f : \mathbb{R}^2 \Rightarrow \mathbb{R}^2$ mit $\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} y \\ x \end{pmatrix}$. Diese Abbildung ist linear und es gilt:

$$f(e_1) = f\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad f(e_2) = f\left(\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

Die darstellende Matrix lautet also $A := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

Es gilt tatsächlich $f(\vec{v}) = A \cdot \vec{v}$, denn:

$$f\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} y \\ x \end{pmatrix} = x \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} + y \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix}$$

Beispiel 6.19

Es sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ mit $\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x+y \\ x-y \end{pmatrix}$. Diese Abbildung ist linear und es gilt:

$$f(e_1) = f\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad f(e_2) = f\left(\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right) = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \quad \text{d.h.} \quad M(f) = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

Wir berechnen nun den Funktionswert $f\left(\begin{pmatrix} 3 \\ 2 \end{pmatrix}\right) = \begin{pmatrix} 3+2 \\ 3-2 \end{pmatrix} = \begin{pmatrix} 5 \\ 1 \end{pmatrix}$ auf zwei Weisen:

- Berechnung aus $f(e_1), f(e_2)$ unter Nutzung der Linearität von f .

$$\begin{aligned} f\left(\begin{pmatrix} 3 \\ 2 \end{pmatrix}\right) &= f\left(3 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 2 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right) = 3 \cdot f\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) + 2 \cdot f\left(\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right) \\ &= 3 \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} + 2 \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 3+2 \\ 3-2 \end{pmatrix} = \begin{pmatrix} 5 \\ 1 \end{pmatrix} \end{aligned}$$

- Per Matrixmultiplikation

$$M(f) \cdot \begin{pmatrix} 3 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \cdot \begin{pmatrix} 3 \\ 2 \end{pmatrix} = 3 \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} + 2 \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 3+2 \\ 3-2 \end{pmatrix} = \begin{pmatrix} 5 \\ 1 \end{pmatrix}$$

Lemma 6.20

Es seien $f, g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ und $h : \mathbb{R}^m \rightarrow \mathbb{R}^k$ lineare Funktionen mit den folgenden Matrixdarstellungen:
 $f(x) = M(f) \cdot x$, $g(x) := M(g) \cdot x$ und $h(y) := M(h) \cdot y$.

Dann haben die linearen Funktionen $f + g$, λf und $h \circ f$ die folgenden Matrixdarstellungen:

- 1) $M(f + g) = M(f) + M(g)$
- 2) $M(\lambda \cdot f) = \lambda \cdot M(f)$
- 3) $M(h \circ f) = M(h) \cdot M(f)$

Beweis: 1) und 2) sind klar bei 3) gilt für alle $\vec{x} \in \mathbb{R}^n$:

$$h \circ f(\vec{x}) = h(f(\vec{x})) = M(h) \cdot (M(f) \cdot \vec{x}) = \underbrace{(M(h) \cdot M(f))}_{=M(h \circ f)} \cdot \vec{x}$$

□

Beispiel 6.21

Die Abbildung $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, $\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} y \\ x \end{pmatrix}$ entspricht der Spiegelung des \mathbb{R}^2 an der Geraden $x = y$.

Die darstellende Matrix hat die Spalten $f(e_1) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, $f(e_2) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, d.h.

$$M(f) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

Die Abbildung $h : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, $\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} -x \\ y \end{pmatrix}$ entspricht einer Spiegelung des \mathbb{R}^2 an der y -Achse.

Die darstellende Matrix hat die Spalten $h(e_1) = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$, $h(e_2) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, d.h.

$$M(h) = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$$

Die Abbildung $h \circ f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, entspricht der Hintereinanderausführung beider Spiegelungen:

$$h \circ f \left(\begin{pmatrix} x \\ y \end{pmatrix} \right) = h \left(\begin{pmatrix} y \\ x \end{pmatrix} \right) = \begin{pmatrix} -y \\ x \end{pmatrix}$$

Die darstellende Matrix hat also Spalten $h \circ f(e_1) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, $h \circ f(e_2) = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$, d.h. $M(h \circ f) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$

Es gilt hier:

$$h \circ f(\vec{v}) = h(f(\vec{v})) = M(h) \cdot (M(f) \cdot \vec{v}) = (M(h) \cdot M(f)) \cdot \vec{v}$$

Das heißt $M(h \circ f)$ (die darstellende Matrix von $h \circ f$) sollte von der Form $M(h) \cdot M(f)$ sein. Dies ist auch so:

$$M(h) \cdot M(f) = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

6.5 Dimensionssatz**Definition 6.22**

Für eine Matrix $A \in \mathbb{R}^{m \times n}$ definieren wir den Kern der Matrix und das Bild der Matrix als:

$$\text{Kern}(A) := \{ \vec{x} \in \mathbb{R}^n : A \cdot \vec{x} = \mathbf{0} \} \quad \text{Bild}(A) := \{ A \cdot \vec{x} : \vec{x} \in \mathbb{R}^n \}$$

Bemerke: Für $A = (\vec{a}_1 \dots \vec{a}_n)$ mit Spalten $\vec{a}_i \in \mathbb{R}^m$ gilt: $\text{Bild}(A) = \text{spann}(\{ \vec{a}_1, \dots, \vec{a}_n \})$

Bemerke: Für die Abbildung $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$, $x \mapsto A \cdot x$ gilt:

- Der Kern ist Teilmenge der Urbildmenge \mathbb{R}^n der Abbildung F
- Das Bild ist Teilmenge des Bildmenge \mathbb{R}^m der Abbildung F

Lemma 6.23

Für jede Matrix $A \in \mathbb{R}^{m \times n}$ sind $\text{Kern}(A) \subseteq \mathbb{R}^n$ und $\text{Bild}(A) \subseteq \mathbb{R}^m$ Vektorräume.

Beweis: Hat $A \in \mathbb{R}^{m \times n}$ die Spalten $\vec{a}_1, \dots, \vec{a}_n \in \mathbb{R}^m$, d.h. $A = (\vec{a}_1, \dots, \vec{a}_n)$ so gilt $\text{Bild}(A) = \text{spann}(\{\vec{a}_1, \dots, \vec{a}_n\})$:

$$\begin{aligned} \text{Bild}(A) &= \{A \cdot \vec{x} : \vec{x} \in \mathbb{R}^n\} \\ &= \{x_1 \cdot \vec{a}_1 + \dots + x_n \vec{a}_n : x_1, \dots, x_n \in \mathbb{R}\} = \text{spann}(\{\vec{a}_1, \dots, \vec{a}_n\}) \end{aligned}$$

Der Spann von endlich vielen Vektoren ist stets ein Vektorraum (s. Beispiel 5.29), d.h. $\text{Bild}(A)$ ist ein Vektorraum.

$\text{Kern}(A) \subseteq \mathbb{R}^n$, "erbt" als Teilmenge von \mathbb{R}^n fast alle Eigenschaften des \mathbb{R}^n . Deswegen genügt es nach Lemma 5.28 zu zeigen: $\text{Kern}(A)$ ist abgeschlossen unter Addition und skalarer Multiplikation.

- z.z.: $\forall \vec{x}, \vec{y} \in \text{Kern}(A) : \vec{x} + \vec{y} \in \text{Kern}(A)$.

Seien $x, y \in \text{Kern}(A)$, dann gilt $A \cdot (\vec{x} + \vec{y}) = \underbrace{A \cdot \vec{x}}_{=0} + \underbrace{A \cdot \vec{y}}_{=0} = \mathbf{0}$ und das heißt: $\vec{x} + \vec{y} \in \text{Kern}(A)$.

- z.z.: $\forall \vec{x} \in \text{Kern}(A), \lambda \in \mathbb{R} : \lambda \cdot \vec{x} \in \text{Kern}(A)$.

Seien $x \in \text{Kern}(A)$ und $\lambda \in \mathbb{R}$, dann gilt $A \cdot (\lambda \vec{x}) = \lambda \cdot \underbrace{A \cdot \vec{x}}_{=0} = \mathbf{0}$ und das heißt: $\lambda \cdot \vec{x} \in \text{Kern}(A)$.

□

Lemma 6.24

Es sei $A \in \mathbb{R}^{m \times n}$ dann gilt für die Vektorräume Kern und Bild:

$$\dim(\text{Kern}(A)) + \dim(\text{Bild}(A)) = n$$

Beweis: Wir konstruieren aus einer Basis von $\text{Kern}(A)$ und (den Urbildern von) einer Basis von $\text{Bild}(A)$ eine Basis von \mathbb{R}^n :

Es sei $\{u_1, \dots, u_k\} \subset \mathbb{R}^n$ eine Basis von $\text{Kern}(A)$, d.h. $\dim(\text{Kern}(A)) = k$.

Es sei $\{w_1, \dots, w_\ell\} \subset \mathbb{R}^m$ eine Basis von $\text{Bild}(A)$, d.h. $\dim(\text{Bild}(A)) = \ell$.

Nach Definition von $\text{Bild}(A)$ ist jeder der Vektoren w_i von der Form $w_i = A \cdot v_i$.

Es sei $\{v_1, \dots, v_\ell\} \subset \mathbb{R}^n$ so dass für jedes $i = \{1, \dots, \ell\}$ gilt: $A \cdot v_i = w_i$.

Wir zeigen nun, dass $u_1, \dots, u_k, v_1, \dots, v_\ell$ linear **unabhängig** sind: Es seien $\mu_1, \dots, \mu_k, \lambda_1, \dots, \lambda_\ell \in \mathbb{R}$

$$\begin{aligned}
 (*) \quad & \mu_1 \cdot u_1 + \dots + \mu_k \cdot u_k + \lambda_1 \cdot v_1 + \dots + \lambda_\ell \cdot v_\ell = \mathbf{0} \\
 \implies & A \cdot (\mu_1 \cdot u_1 + \dots + \mu_k \cdot u_k + \lambda_1 \cdot v_1 + \dots + \lambda_\ell \cdot v_\ell) = A \cdot \mathbf{0} \\
 \implies & \mu_1 \underbrace{A \cdot u_1}_{=0} + \dots + \mu_k \underbrace{A \cdot u_k}_{=0} + \lambda_1 \cdot \underbrace{A \cdot v_1}_{=w_1} + \dots + \lambda_\ell \cdot \underbrace{A \cdot v_\ell}_{=w_\ell} = \mathbf{0} \\
 & \stackrel{*}{\implies} \lambda_1 = \lambda_2 = \dots = \lambda_\ell = 0 \quad \star w_1, \dots, w_\ell \text{ sind linear unabhängig}
 \end{aligned}$$

Setzt man $\lambda_1 = \lambda_2 = \dots = \lambda_\ell = 0$ in (*) ein, so erhält man

$$\begin{aligned}
 & \mu_1 \cdot u_1 + \dots + \mu_k \cdot u_k + \mathbf{0} \cdot v_1 + \dots + \mathbf{0} \cdot v_\ell = \mathbf{0} \\
 \implies & \mu_1 = \dots = \mu_k = 0 \quad \text{weil } u_1, \dots, u_k \text{ linear unabhängig sind}
 \end{aligned}$$

Aus (*) folgt also $\mu_1 = \dots = \mu_k = 0$ und $\lambda_1 = \lambda_2 = \dots = \lambda_\ell = 0$, d.h. die Vektoren sind linear **unabhängig**.

Wir haben $k + \ell$ -viele linear **unabhängige** Vektoren im \mathbb{R}^n gefunden. Damit gilt nun: $k + \ell \leq \dim(\mathbb{R}^n) = n$

Es bleibt zu zeigen: Für $V := \text{spann}(\{u_1, \dots, u_k, v_1, \dots, v_\ell\})$ gilt $\mathbb{R}^n \subseteq V$.

Es sei $x \in \mathbb{R}^n$ beliebig und $\tilde{y} := A \cdot x$. Wegen $\tilde{y} \in \text{Bild}(A)$ gilt $\tilde{y} = \lambda_1 w_1 + \dots + \lambda_\ell w_\ell$.

Es gibt damit also $\tilde{x} \in \text{spann}(v_1, \dots, v_\ell)$ mit $A\tilde{x} = \tilde{y}$ nämlich $\tilde{x} := \lambda_1 v_1 + \dots + \lambda_\ell v_\ell$.

Wegen $Ax = A\tilde{x}$ folgt $A(x - \tilde{x}) = \mathbf{0}$, d.h. $\tilde{z} := x - \tilde{x} \in \text{Kern}(A)$.

Insgesamt folgt: $x = \tilde{z} + \tilde{x}$ mit $\tilde{z} \in \text{spann}(u_1, \dots, u_k)$ und $\tilde{x} \in \text{spann}(v_1, \dots, v_\ell)$.

Weil $x \in \mathbb{R}^n$ beliebig gewählt wurde, gilt also $\mathbb{R}^n \subseteq V$. □

6.6 Lineare Selbstabbildungen: Die Welt der quadratischen Matrizen

Eine Matrix $A \in \mathbb{R}^{n \times n}$ bei der die Anzahl der Zeilen mit der Anzahl der Spalten übereinstimmt heißt *quadratisch* ("Input-" gleich "Output-Dimension").

Quadratische Matrizen stehen für *lineare Selbstabbildungen*, d.h. eine Matrix $A \in \mathbb{R}^{n \times n}$ liefert eine Abbildung $f(x) := A(x)$ mit $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$. Für solche "Automorphismen"¹ gibt es viele Anwendungen und

¹Automorphismus= Strukturerhaltende Selbstabbildungen, von 'auto' selbst & 'morphismus' Verformung

deswegen eine reichhaltige Theorie.

Definition 6.26 (Identitätsmatrix)

Für $n \in \mathbb{N}$ ist $I_n \in \mathbb{R}^{n \times n}$ mit $I_n := (\vec{e}_1 \dots \vec{e}_n)$ die *Einheitsmatrix* der Dimension n . Die Matrix I_n hat also in allen Diagonaleinträgen eine 1, und 0 sonst:

$$(I_n)_{i,j} = \begin{cases} 1 & \text{wenn } i = j \\ 0 & \text{sonst} \end{cases}$$

Wenn klar ist, welche Dimension n gemeint ist, so schreibt man für I_n auch kurz I .

Beispiel 6.26

$$I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Korollar 6.27

Für jede Matrix $A \in \mathbb{R}^{n \times n}$ gilt $I_n \cdot A = A \cdot I_n = A$. Für jeden Vektor $\vec{x} \in \mathbb{R}^n$ gilt $I_n \cdot \vec{x} = \vec{x}$

Definition 6.28

Eine Matrix $A \in \mathbb{R}^{n \times n}$ heißt **invertierbar** mit *inverser Matrix* $B \in \mathbb{R}^{n \times n}$, falls gilt: $B \cdot A = I$. Man sagt in diesem Falle “ A ist **regulär**” und schreibt $A^{-1} := B$. Ist A nicht invertierbar, so sagt man “ A ist **singulär**”.

Lemma 6.29

Eine Matrix $A \in \mathbb{R}^{n \times n}$ ist genau dann invertierbar, wenn die Abbildung $f_A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ mit $x \mapsto A \cdot x$ bijektiv ist.

Beispiel 6.30

Für $A := \begin{pmatrix} 3 & 7 \\ 2 & 5 \end{pmatrix}$ gilt $A^{-1} = \begin{pmatrix} 5 & -7 \\ -2 & 3 \end{pmatrix}$

$$A^{-1} \cdot A = \begin{pmatrix} 5 & -7 \\ -2 & 3 \end{pmatrix} \cdot \begin{pmatrix} 3 & 7 \\ 2 & 5 \end{pmatrix} = \begin{pmatrix} 5 \cdot 3 - 7 \cdot 2 & 0 \\ 0 & -2 \cdot 7 + 3 \cdot 5 \end{pmatrix} = I$$

Beispiel 6.31 (Formel für das Invertieren in $\mathbb{R}^{2 \times 2}$)

Es sei $A := \begin{pmatrix} a & c \\ b & d \end{pmatrix}$ mit Einträgen $a, b, c, d \in \mathbb{R}$ so dass $ad - cb \neq 0$ gilt.

Es gilt dann:

$$A^{-1} = \frac{1}{ad - cb} \cdot \begin{pmatrix} d & -c \\ -b & a \end{pmatrix}$$

$$A^{-1} \cdot A = \frac{1}{ad - cb} \cdot \begin{pmatrix} d & -c \\ -b & a \end{pmatrix} \cdot \begin{pmatrix} a & c \\ b & d \end{pmatrix} = \frac{1}{ad - cb} \cdot \begin{pmatrix} ad - cb & 0 \\ 0 & ad - cb \end{pmatrix} = I$$

Kapitel 7

Das Skalarprodukt

7.1 Das Skalarprodukt und die Euklidische Norm im \mathbb{R}^n

Definition 7.2

Es seien $\vec{x}, \vec{y} \in \mathbb{R}^m$ dann ist das **Skalarprodukt** von \vec{x} der beiden \vec{y} definiert als:

$$\langle \vec{x}, \vec{y} \rangle := x_1 \cdot y_1 + \dots + x_n \cdot y_n$$

eine Weitere Notation für das Skalarprodukt ist $\langle \vec{x}, \vec{y} \rangle = (\vec{x})^T \cdot \vec{y} = (x_1, \dots, x_n) \cdot \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$.

Beispiel 7.2

$$\text{Es gilt: } \left\langle \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} \right\rangle = 1 \cdot 4 + 2 \cdot 5 + 3 \cdot 6 = 32$$



Man beachte, dass das **Skalarprodukt eine Zahl berechnet**.

Für das Skalarprodukt gelten die folgenden Rechenregeln:

Korollar 7.3

Es seien $\vec{x}, \vec{y}, \vec{z} \in \mathbb{R}^m$ und $\lambda \in \mathbb{R}$ dann gelten:

$$1) \quad \langle \vec{x}, \vec{z} \rangle = \langle \vec{z}, \vec{x} \rangle \quad (\text{Symmetrie})$$

$$2.a) \quad \langle \vec{x} + \vec{y}, \vec{z} \rangle = \langle \vec{x}, \vec{z} \rangle + \langle \vec{y}, \vec{z} \rangle$$

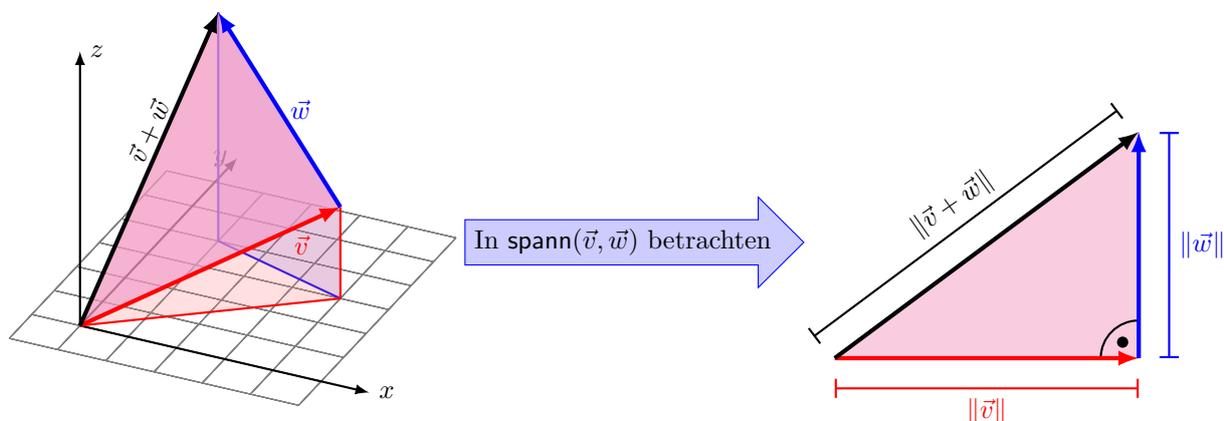
$$2.b) \quad \langle \lambda \vec{x}, \vec{z} \rangle = \lambda \cdot \langle \vec{x}, \vec{z} \rangle$$

Wegen der Symmetrie ergeben sich sofort:

$$\langle \vec{z}, \vec{y} + \vec{x} \rangle = \langle \vec{z}, \vec{x} \rangle + \langle \vec{z}, \vec{y} \rangle$$

$$\langle \vec{z}, \lambda \cdot \vec{x} \rangle = \lambda \cdot \langle \vec{z}, \vec{x} \rangle$$

(Linearität in beiden Einträgen)



Phytagoras: $\|\vec{v}\|^2 + \|\vec{w}\|^2 = \|\vec{v} + \vec{w}\|^2$

Abbildung 7.1: Pythagoras für Dreiecke im \mathbb{R}^3

Definition 7.5

Für $\vec{x} \in \mathbb{R}^m$ ist die **euklidische Norm** definiert als $\|\vec{x}\| := \sqrt{\langle \vec{x}, \vec{x} \rangle} = \sqrt{x_1^2 + \dots + x_n^2}$. Die euklidische Norm eines Vektors entspricht geometrisch seiner Länge.

Beispiel 7.5

Es gilt: $\left\| \begin{pmatrix} -3 \\ 4 \end{pmatrix} \right\| = \sqrt{(-3)^2 + 4^2} = \sqrt{25} = 5$

FAQ: Wieso Doppelstriche?

F: Warum wird die Länge eines Vektors mit Doppelstrichen angegeben?

A: Um den Unterschied zum Betrag einer Zahl klar darzustellen (z.B. $|-3|$)

Definition 7.6

Zwei Vektoren $\vec{v}, \vec{w} \in \mathbb{R}^n$ heißen *orthogonal*, wenn gilt: $\langle \vec{v}, \vec{w} \rangle = 0$.

Geometrisch stehen zwei orthogonale Vektoren \vec{v} und \vec{w} senkrecht aufeinander.

Die Vektoren \vec{v} , \vec{w} und $\vec{v} + \vec{w}$ bilden die Seiten eines Dreiecks¹ (s. Abb. 7.1). Nach Satz des Pythagoras ist das Dreieck rechtwinklig genau dann wenn gilt: $\|\vec{v} + \vec{w}\|^2 = \|\vec{v}\|^2 + \|\vec{w}\|^2$.

Nach der Definition des Betrags gilt jedoch: $\|\vec{v} + \vec{w}\|^2 = \|\vec{v}\|^2 + \|\vec{w}\|^2 + 2\langle \vec{v}, \vec{w} \rangle$ (s. Bew. Lem. 7.7).

Das Dreieck ist also genau dann rechtwinklig wenn gilt: $\langle \vec{v}, \vec{w} \rangle = 0$.

Aus der Gleichung oben lassen sich zwei Formeln ableiten, die helfen, die geometrischen Eigenschaften des Skalarprodukts zu klären:

Lemma 7.7

Für $\vec{v}, \vec{w} \in \mathbb{R}^n$ gilt: $\langle \vec{v}, \vec{w} \rangle = \frac{\|\vec{v} + \vec{w}\|^2 - \|\vec{v}\|^2 - \|\vec{w}\|^2}{2}$

¹Dieses Dreieck sieht man sich relativ zur 2-dimensionalen Ebene $\text{spann}(\vec{v}, \vec{w})$ an.

Beweis: Es seien $\vec{v}, \vec{w} \in \mathbb{R}^n$. Dann gilt:

$$\|\vec{v} + \vec{w}\|^2 = \langle \vec{v} + \vec{w}, \vec{v} + \vec{w} \rangle \stackrel{*}{=} \underbrace{\langle \vec{v}, \vec{v} \rangle}_{=\|\vec{v}\|^2} + \langle \vec{v}, \vec{w} \rangle + \langle \vec{w}, \vec{v} \rangle + \underbrace{\langle \vec{w}, \vec{w} \rangle}_{=\|\vec{w}\|^2} \stackrel{**}{=} \|\vec{v}\|^2 + \|\vec{w}\|^2 + 2\langle \vec{v}, \vec{w} \rangle$$

Hier haben wir verwendet: * Das Skalarprodukt ist bilinear, ** Symmetrie $\langle \vec{v}, \vec{w} \rangle = \langle \vec{w}, \vec{v} \rangle$

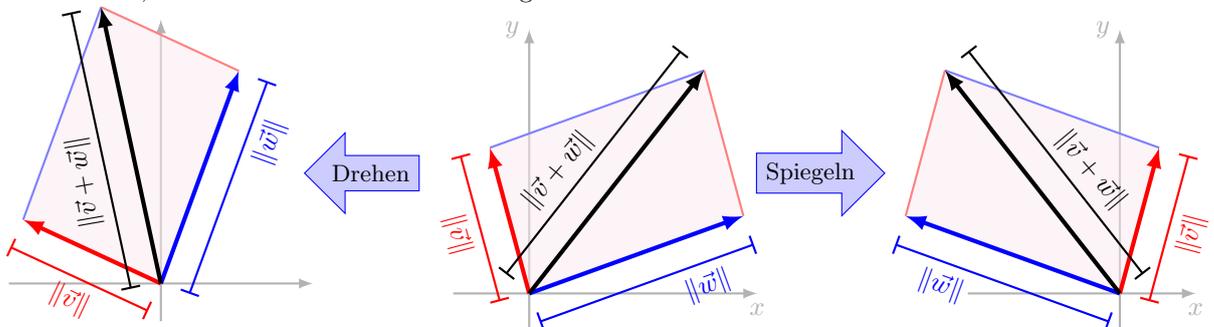
Es gilt also $\langle \vec{v}, \vec{w} \rangle = \frac{1}{2} \left(\|\vec{v} + \vec{w}\|^2 - \|\vec{v}\|^2 - \|\vec{w}\|^2 \right)$. \square

Die Formel in Lemma 7.7 verrät:

Korollar 7.8

Das Skalarprodukt ist Invariant unter Drehungen und Spiegelungen des \mathbb{R}^n .

Denn: Drehungen und Spiegelungen verändern die Länge der abgebildeten Vektoren nicht, und sie erhalten den Winkel, der zwischen zwei Vektoren eingeschlossen wird.



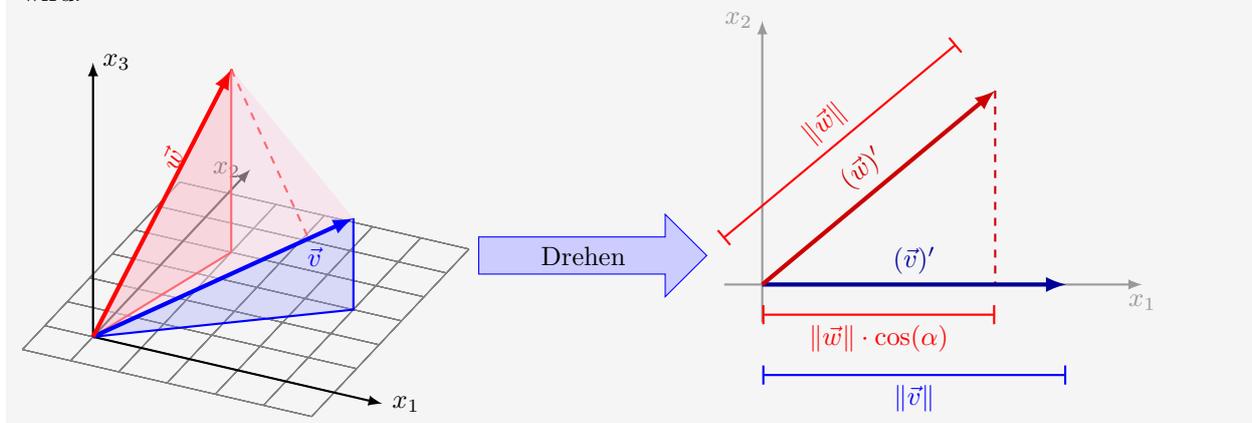
Wendet man also eine Drehung oder eine Spiegelung also gleichzeitig auf \vec{v} und \vec{w} so ändern sich deren Längen nicht, und der *eingeschlossene Winkel* ändert sich ebenfalls nicht. Dies führt dazu, dass sich die Länge der Summe $\vec{v} + \vec{w}$ ebenfalls nicht ändert: Ist \vec{v}' das Bild von \vec{v} und \vec{w}' das Bild von \vec{w} , so gilt: $\|\vec{v} + \vec{w}\| = \|\vec{v}' + \vec{w}'\|$.

Lemma 7.9

Es seien $\vec{v}, \vec{w} \in \mathbb{R}^n$ und $\alpha \in [0, \pi]$ der Winkel, der zwischen \vec{v} und \vec{w} eingeschlossen wird, dann gilt:

$$\langle \vec{v}, \vec{w} \rangle = \|\vec{v}\| \cdot \|\vec{w}\| \cdot \cos(\alpha)$$

Beweis: Um die Gleichung $\langle \vec{v}, \vec{w} \rangle = \|\vec{v}\| \cdot \|\vec{w}\| \cdot \cos(\alpha)$ zu zeigen bedienen wir uns einer geeigneten Drehung: Man drehe den gesamten \mathbb{R}^n , so dass \vec{v} auf die x_1 -Achse gedreht wird, und \vec{w} in die x_1 - x_2 Ebene gedreht wird.



Ist $(\vec{v})'$ das Bild von \vec{v} unter der Drehung und $(\vec{w})'$ das Bild von \vec{w} , so gibt es $\lambda, a, b \in \mathbb{R}$ mit:

$$(\vec{v})' = \begin{pmatrix} \lambda \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad (\vec{w})' = \begin{pmatrix} a \\ b \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad \text{und} \quad \langle \vec{v}, \vec{w} \rangle = \langle (\vec{v})', (\vec{w})' \rangle = \lambda \cdot a$$

Um die Rechnungen kurz zu halten, streichen wir die letzten $n - 2$ Nullen aus $(\vec{v})'$, $(\vec{w})'$ und erhalten:

$$(\vec{v})'' = \begin{pmatrix} \lambda \\ 0 \end{pmatrix} \quad (\vec{w})'' = \begin{pmatrix} a \\ b \end{pmatrix} \quad \text{mit weiterhin} \quad \langle \vec{v}, \vec{w} \rangle = \langle (\vec{v})'', (\vec{w})'' \rangle = \lambda \cdot a$$

Wir zeigen nun $\lambda = \|\vec{v}\|$ und $a = \|\vec{w}\| \cdot \cos(\alpha)$. Daraus folgt dann die Behauptung $\langle \vec{v}, \vec{w} \rangle = \|\vec{v}\| \cdot \|\vec{w}\| \cdot \cos(\alpha)$

Berechnung von $\lambda = \|\vec{v}\|$

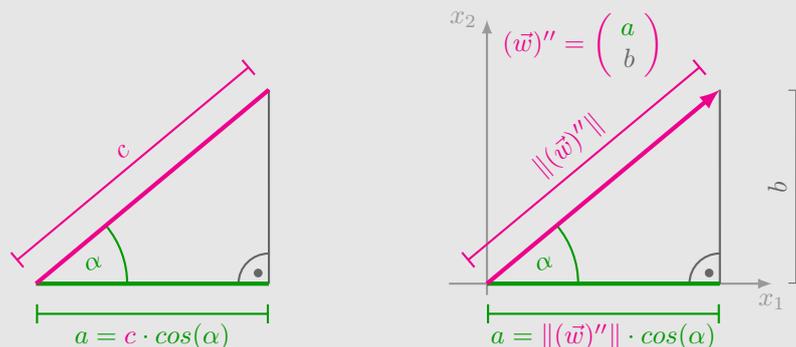
Es gilt $\|\vec{v}\| = \|(\vec{v})'\|$ weil das Drehen von \vec{v} auf $(\vec{v})'$ die Längen der Vektoren nicht ändert.

Es muss also $\|\vec{v}\| = \lambda$ gelten, weil $\|\vec{v}\| = \|(\vec{v})'\|$ (drehen) und $\|(\vec{v})'\| = \|(\vec{v})''\| = \lambda$ (nachrechnen!) gelten.

Berechnung von $a = \|\vec{w}\| \cdot \cos(\alpha)$

Auch für \vec{w} ist Drehen Längenerhaltend, d.h. $\|\vec{w}\| = \|(\vec{w})'\|$, und es gilt $\|(\vec{w})'\| = \|(\vec{w})''\| = \sqrt{a^2 + b^2}$.

Weiter gilt $a = \|(\vec{w})''\| \cdot \cos(\alpha)$, denn das Dreieck mit den Ecken $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$, $\begin{pmatrix} a \\ 0 \end{pmatrix}$ und $(\vec{w})'' = \begin{pmatrix} a \\ b \end{pmatrix}$ ist rechtwinklig:



Im Rechtwinkligen Dreieck mit Hypotenuse c und Ankathete^b a gilt:

$$\cos(\alpha) = \frac{a}{c} = \frac{\text{Länge der Ankathete}}{\text{Länge der Hypotenuse}} \quad \text{bzw.} \quad a = c \cdot \cos(\alpha)$$

Angewendet auf das Dreieck mit Hypotenuse $(\vec{w})''$ liefert dies: $a = \|(\vec{w})''\| \cdot \cos(\alpha)$, und wegen $\|(\vec{w})''\| = \|\vec{w}\|$ folgt die Behauptung.

^b(zum Winkel α , der zwischen c und a eingeschlossen wird)

□

Korollar 7.10 (Cauchy-Schwarz)

Für Vektoren $\vec{x}, \vec{y} \in \mathbb{R}^n$ gilt stets $|\langle \vec{x}, \vec{y} \rangle| \leq \|\vec{x}\| \cdot \|\vec{y}\|$

Lemma 7.11 (Rechenregeln für die Norm)

Für alle $x, y \in \mathbb{R}^n$ und $\lambda \in \mathbb{R}$ gelten:

1. Es gilt $\|x\| = 0$ genau dann wenn $x = \mathbf{0}$.
2. $\|\lambda x\| = |\lambda| \cdot \|x\|$
3. $\|x + y\| \leq \|x\| + \|y\|$

Beweis: Aussagen 1) und 2) folgen direkt aus der Definition der Norm. Für 3) berechnen wir:

$$\begin{aligned} \|x + y\|^2 &= \langle x + y, x + y \rangle = \langle x, x \rangle + 2\langle x, y \rangle + \langle y, y \rangle \\ &\leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2 \quad (\text{nach Cauchy Schwarz}) \end{aligned}$$

Wurzelziehen auf beiden Seiten von $\|x + y\|^2 \leq (\|x\| + \|y\|)^2$ liefert die Aussage. \square

Definition 7.12

Für einen Vektor $\vec{x} \in \mathbb{R}^n$ ist $\frac{\vec{x}}{\|\vec{x}\|}$ der zu \vec{x} gehörige *normierte* Vektor.

Für $\vec{y} := \frac{\vec{x}}{\|\vec{x}\|}$ gelten: $\|\vec{y}\| = 1$ und \vec{y} hat die selbe Richtung wie \vec{x} .

7.2 Orthonormalbasen

Im folgenden betrachten wir besonders “schöne” Basen des \mathbb{R}^n , die sogenannten Orthonormalbasen. Diese sind deshalb sehr beliebt, weil sich die Darstellung eines Vektors bezüglich einer solchen Basis leicht berechnen lässt.

Die Notation $\vec{x} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ steht “eigentlich” für eine Linearkombination der kanonischen Einheitsvektoren:

$$\vec{x} = 1 \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + 2 \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + 3 \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = 1 \cdot \vec{e}_1 + 2 \cdot \vec{e}_2 + 3 \cdot \vec{e}_3$$

Dabei sind 1, 2, 3 die “Mengenangaben” für die “Zutatenliste” $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$. Natürlich lassen sich Vektoren aber auch aus beliebigen anderen “Zutatenlisten” (d.h. Basen) “erstellen” (d.h. linearkombinieren). Den

zu einem Vektor gehörigen Koeffizienten-Vektor definiert man wie folgt:

Definition 7.13

Es sei $\mathcal{B} := \{\vec{b}_1, \dots, \vec{b}_n\} \subset \mathbb{R}^n$ eine Basis des \mathbb{R}^n .

- Die Linearkombination der \vec{b}_i mit Vorfaktoren $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ schreibt man kurz als

$$\begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}_{\mathcal{B}} := \lambda_1 \cdot \vec{b}_1 + \dots + \lambda_n \cdot \vec{b}_n$$

- Ist umgekehrt $\vec{x} \in \mathbb{R}^n$, so ist \vec{x} eine eindeutige Linearkombination der \vec{b}_i . Es gibt also μ_1, \dots, μ_n mit

$$\vec{x} = \mu_1 \cdot \vec{b}_1 + \dots + \mu_n \cdot \vec{b}_n = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix}_{\mathcal{B}}$$

Hier heißt $\begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix}_{\mathcal{B}}$ die Darstellung von \vec{x} bezüglich der Basis \mathcal{B} .

Die Darstellung eines Vektors in einer Basis lässt sich besonders leicht berechnen (s. Lemma 7.20), wenn die gewählte Basis nur Vektoren der Länge 1 enthält, die paarweise senkrecht zueinander stehen. Deswegen gibt es eine eigene Bezeichnung für solche Vektor-Mengen/Basen:

Definition 7.14

Eine Menge von Vektoren $\{\vec{b}_1, \dots, \vec{b}_k\}$ heißt

- *orthogonal*, wenn die Vektoren paarweise **orthogonal** sind, wenn also $\langle \vec{b}_i, \vec{b}_j \rangle = 0$ für alle $i \neq j$ mit $i, j \in \{1, \dots, k\}$ gilt.
- *orthonormal*, wenn die Vektoren paarweise **orthogonal** sind und alle Länge 1 haben, wenn also für alle $i, j \in \{1, \dots, k\}$ gilt

$$\langle \vec{b}_i, \vec{b}_j \rangle = \begin{cases} 0 & \text{falls } i \neq j \\ 1 & \text{falls } i = j \end{cases}$$

Beispiel 7.15

Die Vektoren $\{\vec{v}, \vec{w}\}$ mit $\vec{v} = \frac{1}{5} \begin{pmatrix} 3 \\ 4 \end{pmatrix}$, $\vec{w} = \frac{1}{5} \begin{pmatrix} -4 \\ 3 \end{pmatrix}$ sind **orthonormal**, denn :

$$\text{orthogonal} \quad \langle \vec{v}, \vec{w} \rangle = \left\langle \begin{pmatrix} 3/5 \\ 4/5 \end{pmatrix}, \begin{pmatrix} -4/5 \\ 3/5 \end{pmatrix} \right\rangle = \frac{3}{5} \cdot \frac{(-4)}{5} + \frac{4}{5} \cdot \frac{3}{5} = 0$$

$$\text{Länge} \quad \|\vec{v}\|^2 = \langle \vec{v}, \vec{v} \rangle = \left\langle \begin{pmatrix} 3/5 \\ 4/5 \end{pmatrix}, \begin{pmatrix} 3/5 \\ 4/5 \end{pmatrix} \right\rangle = \frac{3^2}{5^2} + \frac{4^2}{5^2} = 1$$

$$\|\vec{w}\|^2 = \langle \vec{w}, \vec{w} \rangle = \left\langle \begin{pmatrix} -4/5 \\ 3/5 \end{pmatrix}, \begin{pmatrix} -4/5 \\ 3/5 \end{pmatrix} \right\rangle = \frac{(-4)^2}{5^2} + \frac{3^2}{5^2} = 1$$

(Anti-)Beispiel 7.16

Die Vektoren $\left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$ sind **orthogonal** aber nicht **orthonormal**, denn $\left\| \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \right\| = \sqrt{2} \neq 1$

Die Vektoren $\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \frac{1}{5} \begin{pmatrix} 3 \\ 4 \\ 0 \end{pmatrix} \right\}$ sind *weder orthogonal noch orthonormal*.

(Die Vektoren sind zwar normiert (d.h. haben Länge 1), sind aber **nicht orthogonal**.)

Genau so wie die **n**-vielen Vektoren $\vec{e}_1, \dots, \vec{e}_n$ eine Basis des \mathbb{R}^n bilden, bilden auch beliebige andere **n-viele** paarweise orthogonale Vektoren stets eine Basis von \mathbb{R}^n :

Lemma 7.17

Ist $\mathcal{B} := \{\vec{b}_1, \dots, \vec{b}_n\} \subset \mathbb{R}^n$ eine **orthogonale** Menge mit **n-vielen** Vektoren, so ist \mathcal{B} eine Basis des \mathbb{R}^n .

Beispiel 7.18

Die Menge $\mathcal{B} := \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} \right\}$ bildet eine Basis des \mathbb{R}^2 , denn $\left\langle \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} \right\rangle = 0$.

Die Menge $\tilde{\mathcal{B}} := \left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$ ist eine Basis des \mathbb{R}^3 .

Die Menge $\left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \right\}$ ist **keine** Basis des \mathbb{R}^3 (denn jede Basis des \mathbb{R}^3 hat genau 3 Elemente).

Definition 7.19

Eine **orthonormale** Menge $\mathcal{B} := \{\vec{b}_1, \dots, \vec{b}_n\} \subset \mathbb{R}^n$ aus **n-vielen** Vektoren, nennt man *Orthonormalbasis* des \mathbb{R}^n .

Dass eine **Orthonormalbasis** tatsächlich auch eine Basis ist, zeigt Lemma 7.17.

Lemma 7.20

Es sei $\mathcal{B} := \{\vec{b}_1, \dots, \vec{b}_n\} \subset \mathbb{R}^n$ eine Orthonormalbasis des \mathbb{R}^n . Dann gilt für jedes $\vec{x} \in \mathbb{R}^n$:

$$\vec{x} = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}_{\mathcal{B}} \quad \text{mit} \quad \lambda_i := \langle \vec{x}, \vec{b}_i \rangle \quad \text{für } i \in \{1, \dots, n\}$$

Eine geometrische Anschauung dieses Sachverhaltes findet sich in Abbildung 7.2. Der Beweis erfordert Wissen aus dem folgenden Abschnitt über orthogonale Matrizen.

Beweis: Es sei $\vec{x} \in \mathbb{R}^n$ und $\mathcal{B} := \{\vec{b}_1, \dots, \vec{b}_n\} \subset \mathbb{R}^n$ eine Orthonormalbasis des \mathbb{R}^n .

Weiter Sei $B := (\vec{b}_1 \dots \vec{b}_n)$ die Matrix mit den Spalten $\vec{b}_1, \dots, \vec{b}_n$. Nach Korollar 8.4 ist B eine orthogonale Matrix und es gilt deswegen $B^T \cdot B = B \cdot B^T = I$.

Der Vektor $\vec{\lambda}$ lässt sich mittels B^T berechnen:

$$B^T \cdot \vec{x} = \begin{pmatrix} \vec{b}_1^T \\ \vdots \\ \vec{b}_n^T \end{pmatrix} \cdot \vec{x} = \begin{pmatrix} \langle \vec{b}_1, \vec{x} \rangle \\ \vdots \\ \langle \vec{b}_n, \vec{x} \rangle \end{pmatrix} = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}$$

Und es gilt:

$$\begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}_{\mathcal{B}} := \lambda_1 \cdot \vec{b}_1 + \dots + \lambda_n \cdot \vec{b}_n = B \cdot \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} = B \cdot (B^T \cdot \vec{x}) \stackrel{\star}{=} (B \cdot B^T) \vec{x} = \vec{x}$$

Bei \star verwendet man die Assoziativität der Matrix-Multiplikation. □

Beispiel 7.21

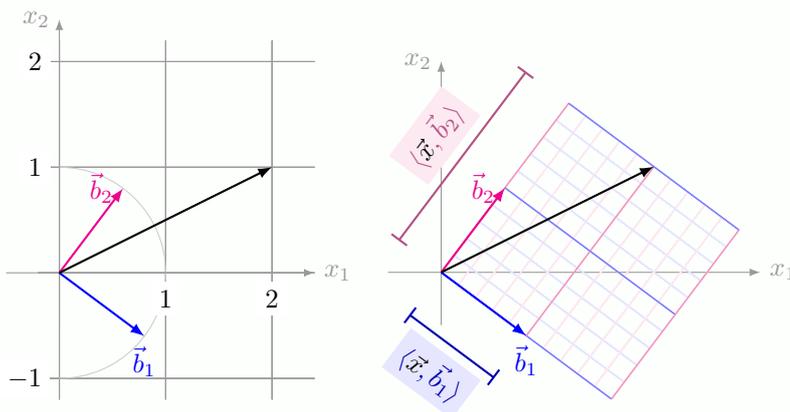
Es sei $\mathcal{B} := \{\vec{b}_1, \vec{b}_2\}$ mit $\vec{b}_1 = \frac{1}{5} \begin{pmatrix} 4 \\ -3 \end{pmatrix}$, $\vec{b}_2 = \frac{1}{5} \begin{pmatrix} 3 \\ 4 \end{pmatrix}$ und $\vec{x} := \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ dann gelten:

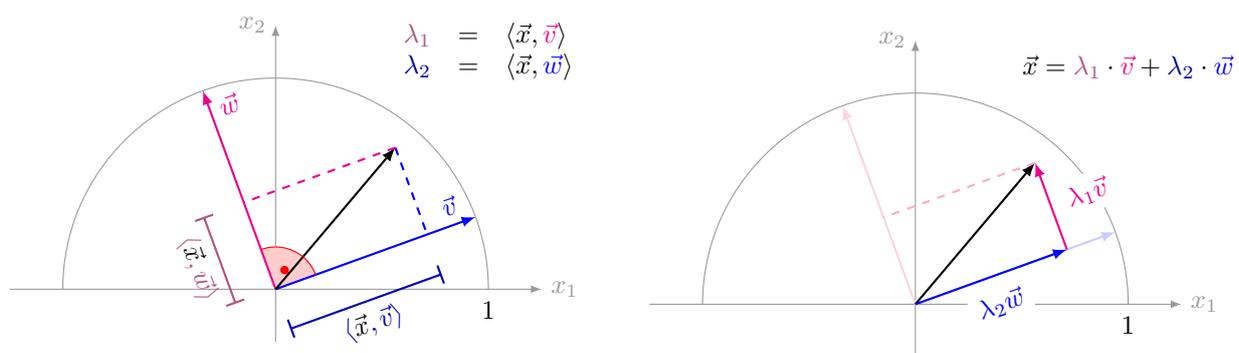
$$\lambda_1 = \langle \vec{x}, \vec{b}_1 \rangle = \left\langle \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \frac{1}{5} \begin{pmatrix} 4 \\ -3 \end{pmatrix} \right\rangle = \frac{1}{5} (2 \cdot 4 + 1 \cdot (-3)) = 1$$

$$\lambda_2 = \langle \vec{x}, \vec{b}_2 \rangle = \left\langle \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \frac{1}{5} \begin{pmatrix} 3 \\ 4 \end{pmatrix} \right\rangle = \frac{1}{5} (2 \cdot 3 + 1 \cdot 4) = 2$$

Probe: Man erhält tatsächlich $\vec{x} = \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix}_{\mathcal{B}} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}_{\mathcal{B}}$ denn:

$$\begin{pmatrix} 1 \\ 2 \end{pmatrix}_{\mathcal{B}} = 1 \cdot \frac{1}{5} \begin{pmatrix} 4 \\ -3 \end{pmatrix} + 2 \cdot \frac{1}{5} \begin{pmatrix} 3 \\ 4 \end{pmatrix} = \frac{1}{5} \cdot \begin{pmatrix} 4+6 \\ -3+8 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$





Berechnen von λ_1, λ_2 per Skalarprodukt.

Zusammensetzen von \vec{x} aus \vec{v}, \vec{w} .

Abbildung 7.2: Darstellung eines Vektors \vec{x} in einer Orthonormalbasis $\{\vec{v}, \vec{w}\} \subset \mathbb{R}^2$

Kapitel 8

Orthogonale Abbildungen

Lemma 8.1

Für Matrizen $A, B \in \mathbb{R}^{n \times n}$ und Vektoren $\vec{x} \in \mathbb{R}^n$ gilt:

$$(A \cdot \vec{x})^T = \vec{x}^T \cdot A^T \quad \text{und} \quad (A \cdot B)^T = B^T \cdot A^T$$

Beweis: Sind $\vec{a}_1^T, \vec{a}_2^T, \dots, \vec{a}_m^T$ die Zeilen der Matrix A , so gilt:

$$\begin{aligned} A \cdot \vec{x} &= \begin{pmatrix} \vec{a}_1^T \\ \vec{a}_2^T \\ \vdots \\ \vec{a}_m^T \end{pmatrix} \cdot \begin{pmatrix} \vec{x} \end{pmatrix} = \begin{pmatrix} \vec{a}_1^T \cdot \vec{x} \\ \vec{a}_2^T \cdot \vec{x} \\ \vdots \\ \vec{a}_m^T \cdot \vec{x} \end{pmatrix} \\ \vec{x}^T \cdot A^T &= \vec{x}^T \cdot \begin{pmatrix} \vec{a}_1 \\ \vec{a}_2 & \dots & \vec{a}_n \end{pmatrix} = \begin{pmatrix} \vec{x}^T \cdot \vec{a}_1 & \vec{x}^T \cdot \vec{a}_2 & \dots & \vec{x}^T \cdot \vec{a}_n \end{pmatrix} \end{aligned}$$

↙ Transponieren ↘

In den Ergebnissen von $A \cdot \vec{x}$ und $\vec{x}^T \cdot A^T$ sind die Einträge $\vec{a}_i^T \cdot \vec{x} = \langle \vec{a}_i, \vec{x} \rangle = \vec{x}^T \cdot \vec{a}_i$ Zahlen. Transponiert man also $A \cdot \vec{x}$ so erhält man $\vec{x}^T \cdot A^T$.

Für Matrix-Matrix Multiplikation ergibt sich:

- Hat $B = (\vec{b}_1 \dots \vec{b}_n)$ die Spalten \vec{b}_i und A hat (wie oben) die Zeilen $\vec{a}_1^T, \dots, \vec{a}_n^T$, so sind die Einträge von $C := A \cdot B$ die Skalarprodukte $C_{i,j} := \vec{a}_i^T \cdot \vec{b}_j$ der Zeilen von A mit den Spalten von B .
- Die Matrix $A^T = (\vec{a}_1 \dots \vec{a}_n)$ hat dann die Spalten $\vec{a}_1, \dots, \vec{a}_n$ und B^T hat die Zeilen $\vec{b}_1^T, \dots, \vec{b}_n^T$. die Einträge von $D := A^T \cdot B^T$ die Skalarprodukte $D_{k,\ell} := \vec{b}_k^T \cdot \vec{a}_\ell$ der Zeilen von A^T mit Spalten von B^T .
- Es gilt also: $C_{i,j} = D_{j,i}$ für alle $i, j \in \{1, \dots, n\}$.

Skizze:

$$A \cdot B = \begin{pmatrix} \vec{a}_1^T \\ \vdots \\ \vec{a}_n^T \end{pmatrix} \cdot \begin{pmatrix} \vec{b}_1 & \dots & \vec{b}_n \end{pmatrix} = \begin{pmatrix} \vec{a}_1^T \cdot \vec{b}_1 & \dots & \vec{a}_1^T \cdot \vec{b}_n \\ \vdots & \dots & \vdots \\ \vec{a}_n^T \cdot \vec{b}_1 & \dots & \vec{a}_n^T \cdot \vec{b}_n \end{pmatrix}$$

$$B^T \cdot A^T = \begin{pmatrix} \vec{b}_1^T \\ \vdots \\ \vec{b}_n^T \end{pmatrix} \cdot \begin{pmatrix} \vec{a}_1 & \dots & \vec{a}_n \end{pmatrix} = \begin{pmatrix} \vec{b}_1^T \cdot \vec{a}_1 & \dots & \vec{b}_1^T \cdot \vec{a}_n \\ \vdots & \dots & \vdots \\ \vec{b}_n^T \cdot \vec{a}_1 & \dots & \vec{b}_n^T \cdot \vec{a}_n \end{pmatrix}$$

Transponieren

Definition 8.2

Eine Matrix $A \in \mathbb{R}^{n \times n}$ heißt *orthogonal* wenn gilt: $A^T \cdot A = I$.

Korollar 8.3

Ist $A \in \mathbb{R}^{n \times n}$ orthogonal so ist A invertierbar mit Inverser Matrix $A^{-1} = A^T$.

Es gilt $A \cdot A^{-1} = I$ bzw. $A \cdot A^T = I$ und damit ist auch A^T orthogonal, denn $(A^T)^T = A$.

Korollar 8.4

Eine Matrix $A = (\vec{a}_1 \dots \vec{a}_n) \in \mathbb{R}^{n \times n}$ ist orthogonal genau dann wenn ihre Spalten $\vec{a}_1, \dots, \vec{a}_n \in \mathbb{R}^n$ eine Orthonormalbasis bilden.

Beweis: Es sei $A = (\vec{a}_1 \dots \vec{a}_n) \in \mathbb{R}^{n \times n}$ mit Spalten $\vec{a}_1, \dots, \vec{a}_n \in \mathbb{R}^n$.

Die Einträge der Matrix $A^T A$ sind die Skalarprodukte der \vec{a}_i , es gilt: $(A^T A)_{k,\ell} = \langle \vec{a}_k, \vec{a}_\ell \rangle = \vec{a}_k^T \cdot \vec{a}_\ell$:

$$A^T \cdot A = \begin{pmatrix} \vec{a}_1^T \\ \vec{a}_2^T \\ \vdots \\ \vec{a}_n^T \end{pmatrix} \cdot \begin{pmatrix} \vec{a}_1 & \vec{a}_2 & \dots & \vec{a}_n \end{pmatrix} = \begin{pmatrix} \vec{a}_1^T \cdot \vec{a}_1 & \vec{a}_1^T \cdot \vec{a}_2 & \dots & \vec{a}_1^T \cdot \vec{a}_n \\ \vec{a}_2^T \cdot \vec{a}_1 & \vec{a}_2^T \cdot \vec{a}_2 & \dots & \vec{a}_2^T \cdot \vec{a}_n \\ \vdots & \vdots & \ddots & \vdots \\ \vec{a}_n^T \cdot \vec{a}_1 & \vec{a}_n^T \cdot \vec{a}_2 & \dots & \vec{a}_n^T \cdot \vec{a}_n \end{pmatrix}$$

Es gilt $A^T A = I$ genau dann wenn,

- für die Diagonal-Einträge $(A^T A)_{k,k} = 1$ gilt (oben hervorgehoben) und
- für alle *anderen* Einträge $(A^T A)_{k,\ell} = 0$ gilt.

D.h. $A^T A$ ist orthogonal genau dann wenn gilt:

$$(A^T A)_{k,\ell} = \langle a_k, a_\ell \rangle = \begin{cases} 1 & \text{falls } k = \ell \\ 0 & \text{falls } k \neq \ell \end{cases}$$

Dies ist äquivalent zur Aussage, dass $\{\vec{a}_1, \dots, \vec{a}_n\}$ eine Orthonormalbasis ist. □

Beispiel 8.5

Die folgenden Matrizen sind orthogonale Matrizen

$$A := \begin{pmatrix} 3/5 & -4/5 \\ 4/5 & 3/5 \end{pmatrix} \quad B := \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \quad C := \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{3} & 1/2 \\ -1/\sqrt{2} & 1/\sqrt{3} & 1/2 \\ 0 & 1/\sqrt{3} & -2/2 \end{pmatrix}$$

Lemma 8.6

Es sei $A \in \mathbb{R}^{n \times n}$ eine orthogonale Matrix. Dann gilt für alle $\vec{x}, \vec{y} \in \mathbb{R}^n$:

$$\langle \vec{x}, \vec{y} \rangle = \langle A \cdot \vec{x}, A \cdot \vec{y} \rangle$$

Die Abbildung $\vec{v} \mapsto A \cdot \vec{v}$ ist also Längen und Winkelerhaltend:

- $\|\vec{v}\| = \|A \cdot \vec{v}\|$
- Ist $\alpha \in [0, \pi]$ der zwischen \vec{v} und \vec{w} eingeschlossene Winkel, so ist der zwischen $A \cdot \vec{v}$ und $A \cdot \vec{w}$ eingeschlossene Winkel ebenfalls α .

Für eine geometrische Anschauung: Siehe Abbildung 8.1.

Beweis: Für eine Matrix $B \in \mathbb{R}^{m \times n}$ und einen Vektor $\vec{x} \in \mathbb{R}^n$ gilt $(B \cdot \vec{x})^T = \vec{x}^T \cdot B^T$. (s. Lemma 8.1).
Es seien $\vec{x}, \vec{y} \in \mathbb{R}^n$ und $A \in \mathbb{R}^{n \times n}$ eine orthogonale Matrix. Dann gilt:

$$\begin{aligned} \langle A \cdot \vec{x}, A \cdot \vec{y} \rangle &= (A \cdot \vec{x})^T \cdot (A \cdot \vec{y}) \\ &= (\vec{x}^T \cdot A^T) \cdot (A \cdot \vec{y}) \stackrel{*}{=} \vec{x}^T \cdot \underbrace{\left(\overbrace{(A^T \cdot A)}^{=I} \cdot \vec{y} \right)}_{=\vec{y}} = \vec{x}^T \cdot \vec{y} = \langle \vec{x}, \vec{y} \rangle \end{aligned}$$

In \star haben wir die Assoziativität der Matrixmultiplikation verwendet.

Die Zusatzaussage a) folgt dann durch Wurzelziehen in $\|\vec{v}\|^2 = \langle \vec{v}, \vec{v} \rangle = \langle A\vec{v}, A\vec{v} \rangle = \|A\vec{v}\|^2$.

Für die Zusatzaussage b) sei α der von \vec{v} und \vec{w} eingeschlossene Winkel, und β der Winkel zwischen $A\vec{v}$ und $A\vec{w}$. Dann gilt $\cos(\alpha) = \cos(\beta)$ wegen:

$$\cos(\alpha) = \frac{\langle \vec{v}, \vec{w} \rangle}{\|\vec{v}\| \cdot \|\vec{w}\|} = \frac{\langle A\vec{v}, A\vec{w} \rangle}{\|A\vec{v}\| \cdot \|A\vec{w}\|} = \cos(\beta)$$

Da für die Winkel $0 \leq \alpha, \beta \leq \pi$ gilt¹ muss $\alpha = \beta$ gelten, denn auf dem Intervall $[0, \pi]$ ist die Funktion $\cos(t)$ injektiv. \square

Lemma 8.7

Es sei $\mathcal{B} := \{\vec{b}_1, \dots, \vec{b}_n\}$ eine Orthonormalbasis des \mathbb{R}^n und $B = (\vec{b}_1 \dots \vec{b}_n)$ die zugehörige orthogonale Matrix. Dann lässt sich die Darstellung von \vec{x} bzgl. \mathcal{B} wie folgt berechnen:

$$\text{Für } \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} := B^T \vec{x} \quad \text{gilt} \quad \vec{x} = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix}_{\mathcal{B}} \quad \text{bzw.} \quad \vec{x} = \beta_1 \cdot \vec{b}_1 + \dots + \beta_n \cdot \vec{b}_n$$

Der Beweis ist eine direkte Konsequenz von Lemma 7.20, und der Beweis läuft völlig analog zum Beweis dort:

Beweis: Es sei $\mathcal{B} := \{\vec{b}_1, \dots, \vec{b}_n\}$ eine Orthonormalbasis des \mathbb{R}^n und $B = (\vec{b}_1 \dots \vec{b}_n)$ die zugehörige orthogonale Matrix.

Dann gilt:

$$\begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} := B^T \cdot \vec{x} = \begin{pmatrix} b_1^T \\ \vdots \\ b_n^T \end{pmatrix} \cdot \vec{x} = \begin{pmatrix} b_1^T \cdot \vec{x} \\ \vdots \\ b_n^T \cdot \vec{x} \end{pmatrix} = \begin{pmatrix} \langle b_1, \vec{x} \rangle \\ \vdots \\ \langle b_n, \vec{x} \rangle \end{pmatrix}$$

Es gilt also $\beta_i = \langle \vec{b}_i, \vec{x} \rangle$ und damit gilt nach Lemma 7.20 (weil \mathcal{B} eine Orthonormalbasis ist):

$$\vec{x} = \beta_1 \cdot \vec{b}_1 + \dots + \beta_n \cdot \vec{b}_n$$

\square

8.1 Die Determinante

8.1.1 Übersicht

Die Determinante ist ein Maß für eine quadratische Matrix A , mit dem sich entscheiden lässt, ob die Spalten von A linear abhängig sind.

Genauer gilt: Die Determinante misst das (orientierte) Volumen des von den Spalten von A aufgespannten Parallelotops (s. Def 8.8, Bsp. 8.9). Dieses Volumen hat den Wert 0 genau dann wenn die Spalten von A linear abhängig sind. Diese geometrische Interpretation als Volumen der Determinante ist jedoch nur eine *Hilfe*, um sich die Regeln bei der Determinantenberechnung dauerhaft merken zu können. Wichtig sind die drei folgenden Einsichten:

- Es gilt $\det(A) = 0$ genau dann wenn die Spalten von A linear abhängig sind.
- Die Determinante ist multilinear (dies hilft beim Berechnen per Gauß-Verfahren).
- Die Determinante lässt sich...

¹Winkel werden hier im Bogenmaß angegeben, im Gradmaß entspricht dies $\alpha, \beta \in [0^\circ, 180^\circ]$.

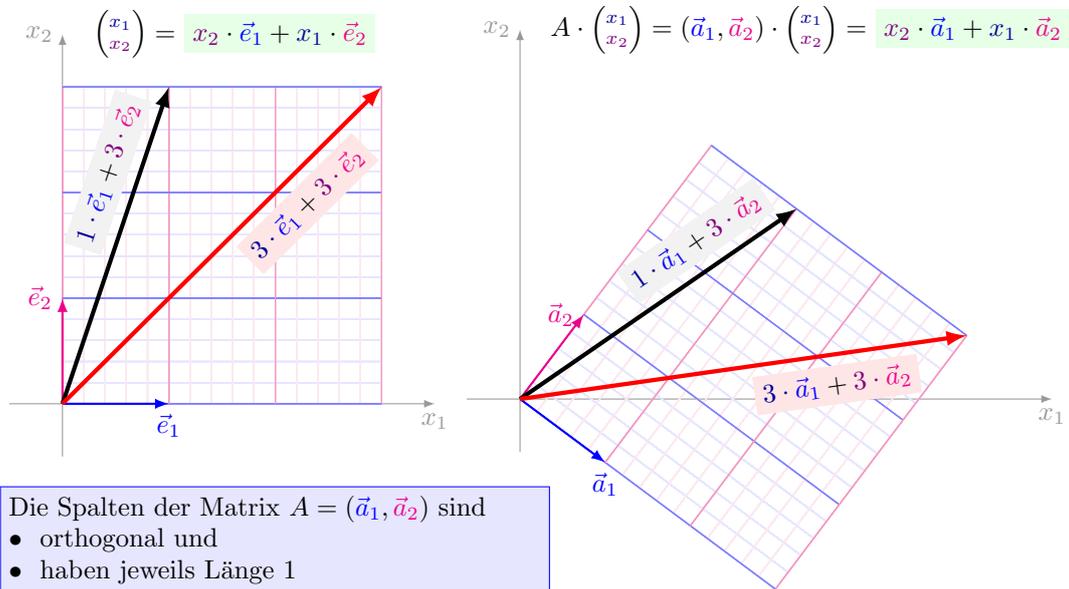


Abbildung 8.1: Orthogonale Abbildungen sind Winkel- und Längenerhaltend

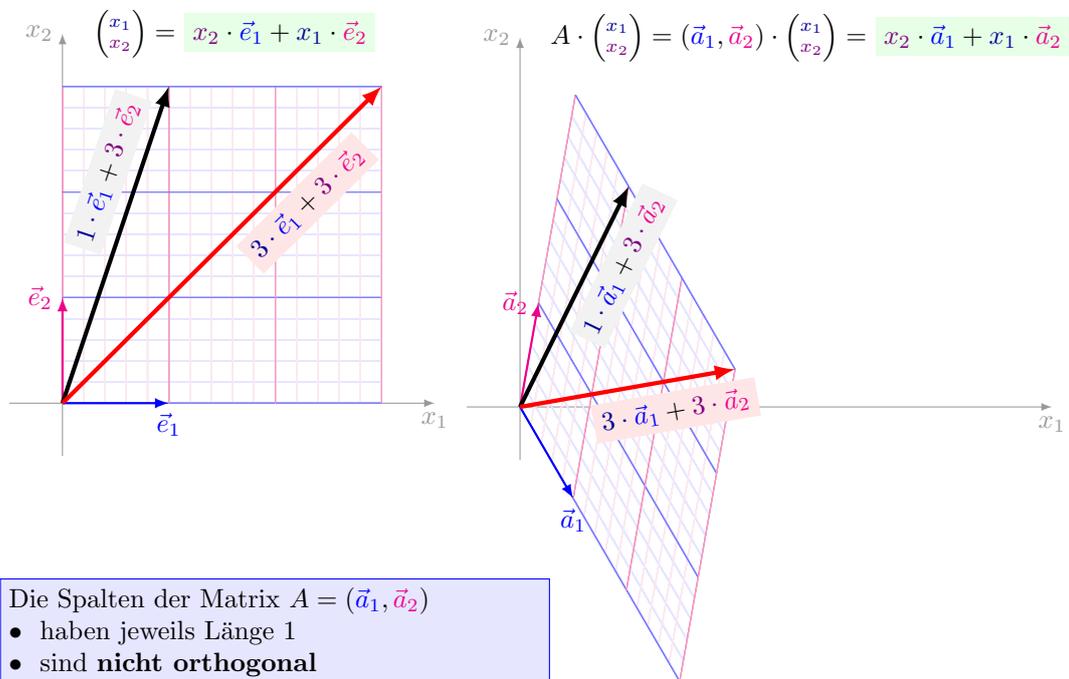


Abbildung 8.2: Gewöhnliche lineare Abbildungen sind nicht Winkel- und Längenerhaltend

- per Zeilen- oder Spaltenentwicklung berechnen
(Verfahren hat theoretische Bedeutung, in der Praxis nur für kleine Matrizen geeignet)
- per Gauss-Verfahren berechnen
(Verfahren hat große praktische Bedeutung).

Hat die *quadratische* Matrix $A \in \mathbb{R}^{n \times n}$ die Spalten $\vec{a}_1, \dots, \vec{a}_n \in \mathbb{R}^n$, so ist die Determinante von A das *orientierte* Volumen, des von $\vec{a}_1, \dots, \vec{a}_n$ aufgespannten Parallelotops².

Die Orientierung bedeutet hier ein Vorzeichen, dass die Reihenfolge der Vektoren widerspiegelt.

Definition 8.8

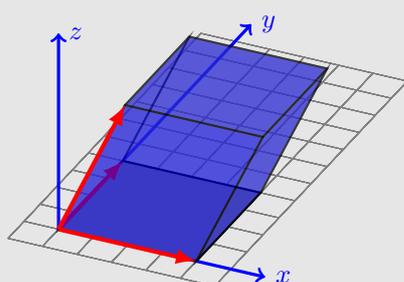
Für n Vektoren $\vec{a}_1, \dots, \vec{a}_n \in \mathbb{R}^n$ sei

$$P(\{\vec{a}_1, \dots, \vec{a}_n\}) := \{\lambda_1 \cdot \vec{a}_1 + \dots + \lambda_n \cdot \vec{a}_n : 0 \leq \lambda_1, \dots, \lambda_n \leq 1\}.$$

Für die Matrix $A \in \mathbb{R}^{n \times n}$ mit Spalten \vec{a}_i schreibt man auch kurz $P(A)$ für $P(\{\vec{a}_1, \dots, \vec{a}_n\})$

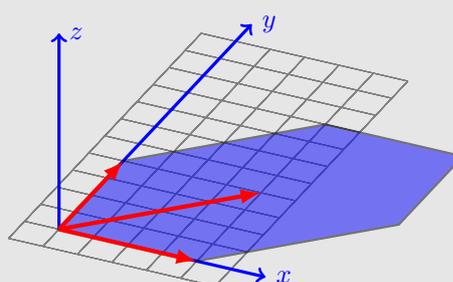
Beispiel 8.9

Sind die $\vec{a}_1, \dots, \vec{a}_n$ linear **un**abhängig so ist $P(A)$ das Parallelotop mit Kanten parallel zu den \vec{a}_i .
Sind die $\vec{a}_1, \dots, \vec{a}_n$ linear **ab**hängig so ist $P(A)$ *flach*.



Das Parallelotop $P\left(\begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}\right)$

Volumen ist $8 = \text{Höhe } 2 \text{ mal Grundfläche } 4$



Das Parallelotop $P\left(\begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix}\right)$ ist **flach**.

Volumen ist 0

Die Determinante $\det(A)$ einer Matrix $A \in \mathbb{R}^{n \times n}$ ist das orientierte Volumen von $P(A)$ (die Genaue Definition folgt in Def. XX). Das *tatsächliche* Volumen ist in den seltensten Fällen interessant! Was wirklich interessiert ist der Fall $\det(A) = 0$:

Lemma 8.10

In einer quadratischen Matrix A sind die **Spalten linear abhängig** genau dann wenn $\det(A) = 0$ gilt.

Eine Richtung dieser Aussage ist leicht zu sehen: Sind die Spalten von A linear **abhängig**, so ist der Körper $P(A)$ relativ zum Raum \mathbb{R}^n *flach*. Das Volumen hat in diesem Falle den Wert 0 .

$$\text{Spalten von } A \text{ sind linear abhängig} \Rightarrow P(A) \text{ ist flach} \Rightarrow \text{vol}(P(A)) = 0 \Rightarrow \det(A) = 0$$

Das die Umkehrungen “ \Leftarrow ” hier auch gelten werden wir später zeigen.

Die Determinante ist also ein (schneller!) Weg zu entscheiden ob Vektoren linear Abhängig sind.

²das Parallelotop ist die mehrdimensionale Verallgemeinerung des Parallelogramms im \mathbb{R}^2 .

8.1.2 Berechnung und Rechenregeln

Die Formel zur Berechnung des orientierten Volumens von P lässt sich in allgemeiner Dimension am besten durch ein Verfahren ausdrücken. Wir werden zwei solcher Verfahren kennenlernen:

- Das Entwickeln der Determinante – dieses Verfahren hat theoretische Bedeutung, in der Praxis nur für kleine Matrizen geeignet.
- Berechnen per Gauss-Verfahren – dieses Verfahren hat große praktische Bedeutung.

Definition 8.11 (Streichungsmatrix)

Für eine Matrix $A \in \mathbb{R}^{n \times n}$ und zwei Indices $k, \ell \in \{1, \dots, n\}$ ist die *Streichungsmatrix* $A^{k,\ell} \in \mathbb{R}^{(n-1) \times (n-1)}$ diejenige Matrix, die aus A hervorgeht durch Streichen der k -ten Zeile und Streichen der ℓ -ten Spalte :

$$A = \begin{pmatrix} A_{1,1} & \cdots & A_{1,\ell-1} & A_{1,\ell} & A_{1,\ell+1} & \cdots & A_{1,n} \\ \vdots & & & \vdots & & & \vdots \\ A_{k-1,1} & \cdots & A_{k-1,\ell-1} & A_{k-1,\ell} & A_{k-1,\ell+1} & \cdots & A_{k-1,n} \\ A_{k,1} & \cdots & A_{k,\ell-1} & A_{k,\ell} & A_{k,\ell+1} & \cdots & A_{k,n} \\ A_{k+1,1} & \cdots & A_{k+1,\ell-1} & A_{k+1,\ell} & A_{k+1,\ell+1} & \cdots & A_{k+1,n} \\ \vdots & & & \vdots & & & \vdots \\ A_{n,1} & \cdots & A_{n,\ell-1} & A_{n,\ell} & A_{n,\ell+1} & \cdots & A_{n,n} \end{pmatrix}$$

$$A^{k,\ell} = \begin{pmatrix} A_{1,1} & \cdots & A_{1,\ell-1} & A_{1,\ell+1} & \cdots & A_{1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ A_{k-1,1} & \cdots & A_{k-1,\ell-1} & A_{k-1,\ell+1} & \cdots & A_{k-1,n} \\ A_{k+1,1} & \cdots & A_{k+1,\ell-1} & A_{k+1,\ell+1} & \cdots & A_{k+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ A_{n,1} & \cdots & A_{n,\ell-1} & A_{n,\ell+1} & \cdots & A_{n,n} \end{pmatrix}$$

Beispiel 8.12

$$A := \begin{pmatrix} 1 & a & \alpha \\ 2 & b & \beta \\ 3 & c & \gamma \end{pmatrix} \Rightarrow A^{2,2} = \begin{pmatrix} 1 & a & \alpha \\ 2 & b & \beta \\ 3 & c & \gamma \end{pmatrix} = \begin{pmatrix} 1 & \alpha \\ 3 & \gamma \end{pmatrix} \quad A^{2,3} = \begin{pmatrix} 1 & a & \alpha \\ 2 & b & \beta \\ 3 & c & \gamma \end{pmatrix} = \begin{pmatrix} 1 & a \\ 3 & c \end{pmatrix}$$

Definition 8.13

Für eine $A \in \mathbb{R}^{n \times n}$ mit $n \in \mathbb{N} \setminus \{0\}$ gilt:

- Falls $n = 1$ gilt, so hat $A = (A_{1,1})$ genau einen Eintrag $A_{1,1} \in \mathbb{R}$ und es gilt $\det(A) = A_{1,1}$
- Falls $n > 1$ gilt, so können wir eine Spalte $\ell \in \{1, \dots, n\}$ wählen und nach dieser Spalte die Determinante *entwickeln*

$$\det(A) = \sum_{i=1}^n (-1)^{i+\ell} A_{i,\ell} \cdot \det(A^{i,\ell})$$

Bemerkung 8.1 Das Vorzeichen $(-1)^{i+\ell}$ für das Element $A_{i,\ell}$ ist nur bestimmt durch die Zahlen i und ℓ , d.h. bestimmt durch die Position in der i -ten Zeile und ℓ -ten Spalte. Die Zahl $(-1)^{i+\ell}$ ist also unabhängig von der Dimension n (und auch unabhängig vom Wert $A_{i,\ell}$).

Hilfreich zum Ermitteln des korrekten Vorzeichens ist eine "Vorzeichen-Matrix": Man trägt in einer Matrix in Zeile i und Spalte ℓ nur das Vorzeichen von $(-1)^{i+\ell}$ ein, dies sieht dann wie folgt aus:

$$\begin{pmatrix} + & - & + \\ - & + & - \\ + & - & + \end{pmatrix} \quad \text{für } \mathbb{R}^{3 \times 3} \qquad \begin{pmatrix} + & - & + & - \\ - & + & - & + \\ + & - & + & - \\ - & + & - & + \end{pmatrix} \quad \text{für } \mathbb{R}^{4 \times 4}$$



Aufwand $O(n!)$

Diese rekursive Definition der Determinante ist *nicht* für große Inputdaten geeignet:

Um $\det(A)$ für $A \in \mathbb{R}^{k \times k}$ zu berechnen, muss man die Determinanten von k -vielen *verschiedenen* $(k-1) \times (k-1)$ -Matrizen berechnen. Ist $R(n)$ der Aufwand beim Berechnen der Determinante einer $n \times n$ -Matrix, dann gilt also (grob) die Rekursionsformel:

$$R(n) = n \cdot R(n-1).$$

Mit dem Wissen $R(1) = 1$ führt dies zu $R(n) = n \cdot (n-1) \cdot \dots \cdot 2 \cdot 1 = n!$.

Beispiel 8.14

Für $A := \begin{pmatrix} a & c \\ b & d \end{pmatrix} \in \mathbb{R}^{2 \times 2}$ gilt stets: $\det(A) = a \cdot d - c \cdot b$.

Entwickelt man die Determinante nach der ersten Spalte so erhält man:

$$\begin{aligned} \det(A) &= (1) \cdot a \cdot \det \begin{pmatrix} c & \\ & d \end{pmatrix} + (-1) \cdot b \cdot \det \begin{pmatrix} a & c \\ & d \end{pmatrix} \\ &= a \cdot \det(d) - b \cdot \det(c) \end{aligned}$$

Beispiel 8.15

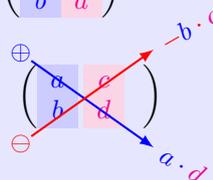
Es sei $A := \begin{pmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{pmatrix}$. Entwickelt man die Determinante nach der zweiten Spalte so erhält man:

$$\begin{aligned} \det(A) &= (-1) \cdot 4 \cdot \det \begin{pmatrix} 1 & 7 \\ 3 & 9 \end{pmatrix} + (1) \cdot 5 \cdot \det \begin{pmatrix} 1 & 7 \\ 3 & 9 \end{pmatrix} + (-1) \cdot 6 \cdot \det \begin{pmatrix} 1 & 7 \\ 2 & 8 \end{pmatrix} \\ &= -4 \cdot \det \begin{pmatrix} 2 & 8 \\ 3 & 9 \end{pmatrix} + 5 \cdot \det \begin{pmatrix} 1 & 7 \\ 3 & 9 \end{pmatrix} - 6 \cdot \det \begin{pmatrix} 1 & 7 \\ 2 & 8 \end{pmatrix} \\ &= -4 \cdot (2 \cdot 9 - 3 \cdot 8) + 5 \cdot (1 \cdot 9 - 3 \cdot 7) - 6 \cdot (1 \cdot 8 - 2 \cdot 7) \end{aligned}$$

Für die Fälle $\det(A)$ mit $A \in \mathbb{R}^{2 \times 2}$ bzw. $A \in \mathbb{R}^{3 \times 3}$ gibt es Sonderformeln:

Lemma 8.16

Es sei $A \in \mathbb{R}^{2 \times 2}$ mit $A = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$ dann lässt sich $\det(A)$ über das Bilden von 2 Summanden berechnen:

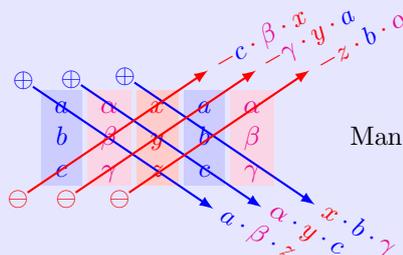


Man erhält $\det(A) = a \cdot d - b \cdot c$

Lemma 8.17 (Sarrusregel)

Es sei $A \in \mathbb{R}^{3 \times 3}$ mit $A = \begin{pmatrix} a & \alpha & x \\ b & \beta & y \\ c & \gamma & z \end{pmatrix}$ dann lässt sich $\det(A)$ über das Bilden von 6 Summanden berechnen.

Für das Ermitteln der Summanden schreibt man die drei Spalten von A in ein Schema und wiederholt die ersten beiden Spalten. Dann bildet man drei “Abwärts-Produkte” mit positivem Vorzeichen, und drei “Aufwärts-Produkte” mit negativem Vorzeichen. Die Determinante ist dann die Summe dieser 6 Produkte:



Man erhält: $\det(A) = \begin{cases} a \cdot \beta \cdot z & + \alpha \cdot y \cdot c & + x \cdot b \cdot \gamma \\ -c \cdot \beta \cdot x & - \gamma \cdot y \cdot a & - z \cdot b \cdot \alpha \end{cases}$

8.1.3 Rechenregeln

Für die Determinante gelten die folgenden Rechenregeln:

Lemma 8.18

Für $A, B \in \mathbb{R}^{n \times n}$ gelten stets

- $\det(A \cdot B) = \det(A) \cdot \det(B)$ (Multiplikativität)
- $\det(A) = \det(A^T)$

Lemma 8.19 (alternierend)

Tauscht man in $\vec{a}_1, \dots, \vec{a}_n \in \mathbb{R}^n$ zwei Vektoren, so ändert die Determinante ihr Vorzeichen:

$$= -\det \left(\vec{a}_1, \dots, \vec{a}_{k-1}, \vec{a}_\ell, \vec{a}_{k+1}, \dots, \vec{a}_{\ell-1}, \vec{a}_k, \vec{a}_{\ell+1}, \dots, \vec{a}_n \right)$$

Lemma 8.20 (Multi-Linearität)

Für $\vec{a}_1, \dots, \vec{a}_n \in \mathbb{R}^n$ sowie $\vec{b} \in \mathbb{R}^n$ und $\lambda \in \mathbb{R}$ gilt stets :

$$\det(\vec{a}_1, \dots, \vec{a}_{\ell-1}, \lambda \cdot \vec{a}_\ell, \vec{a}_{\ell+1}, \dots, \vec{a}_n) = \lambda \cdot \det(\vec{a}_1, \dots, \vec{a}_{\ell-1}, \vec{a}_\ell, \vec{a}_{\ell+1}, \dots, \vec{a}_n)$$

$$\det(\vec{a}_1, \dots, \vec{a}_{\ell-1}, \vec{a}_\ell + \vec{b}, \vec{a}_{\ell+1}, \dots, \vec{a}_n) = \det(\vec{a}_1, \dots, \vec{a}_{\ell-1}, \vec{a}_\ell, \vec{a}_{\ell+1}, \dots, \vec{a}_n) + \det(\vec{a}_1, \dots, \vec{a}_{\ell-1}, \vec{b}, \vec{a}_{\ell+1}, \dots, \vec{a}_n)$$

Beispiel 8.21

Es gilt

$$\det \begin{pmatrix} 2 & 3 \\ 4 & 15 \end{pmatrix} = 2 \cdot \det \begin{pmatrix} 1 & 3 \\ 2 & 15 \end{pmatrix} = 2 \cdot 3 \cdot \det \begin{pmatrix} 1 & 1 \\ 2 & 5 \end{pmatrix} = 2 \cdot 3 \cdot 3 = 18$$

$2 \cdot 15 - 3 \cdot 4 = 18$ $1 \cdot 5 - 2 \cdot 1 = 3$

Weiter gilt:

$$\det \begin{pmatrix} 1 & 1 \\ 2 & 5 \end{pmatrix} = \det \begin{pmatrix} 1 & 1 \\ 0 & 5 \end{pmatrix} + \det \begin{pmatrix} 0 & 1 \\ 2 & 5 \end{pmatrix} \quad \text{wegen } \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 2 \end{pmatrix}$$

$5 - 2 = 3$ $5 - 0$ $0 - 2$

**Die Determinante ist nicht Additiv**

Im Allgemeinen gilt *nicht* $\det(A + B) = \det(A) + \det(B)$, wie das folgende (einfache) Beispiel zeigt.

$$\det \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 1 \neq 0 = \underbrace{\det \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}}_{=0} + \underbrace{\det \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}}_{=0}$$

8.1.4 Rechenricks

Das rekursive Entwickeln einer Determinante nach einer Zeile oder Spalte ist für "große" Matrizen im Allgemeinen schlicht *nicht* möglich.

Für Matrizen mit vielen Null-Einträgen ist dies aber weiterhin möglich:

Beispiel 8.22

Beim Entwickeln der Determinante nach einer Spalte (bzw. Zeile) sollte man eine Spalte (bzw. Zeile) mit vielen Nulleinträgen wählen. Hierdurch verringert sich der Rechenaufwand erheblich!

Es sei $A = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 6 & 7 & 8 & 9 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix}$. Entwickeln nach Zeile 3, dann nach Original-Zeile 4 und dann nach

Original-Zeile 5 liefert:

$$\det(A) = +1 \cdot \det \begin{pmatrix} 2 & 3 & 4 & 5 \\ 7 & 8 & 9 & 1 \\ 2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} = +1 \cdot 2 \det \begin{pmatrix} 3 & 4 & 5 \\ 8 & 9 & 1 \\ 1 & 0 & 0 \end{pmatrix} = +1 \cdot 2 \cdot 1 \cdot \det \begin{pmatrix} 4 & 5 \\ 9 & 1 \end{pmatrix}$$

Eine Matrix heißt obere Dreiecksmatrix, wenn alle Einträgen unterhalb der Diagonale Null sind:

Definition 8.23 (obere Dreiecksmatrix)

Eine Matrix $A \in \mathbb{R}^{n \times n}$ heißt obere Dreiecksmatrix, wenn gilt: $A_{z,s} = 0$ für alle Paare $1 \leq s < z \leq n$. Ist die "Zeilennummer" ist größer als die "Spaltennummer", so steht $A_{z,s}$ unterhalb der Diagonale. In einer oberen Dreiecksmatrix sind also alle Einträge unterhalb der Diagonale Null.

$$A = \begin{pmatrix} A_{1,1} & A_{1,2} & A_{1,3} & \dots & A_{1,n} \\ 0 & A_{2,2} & A_{2,3} & \dots & A_{2,n} \\ 0 & 0 & A_{3,3} & \dots & A_{3,n} \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & 0 & A_{n,n} \end{pmatrix} \quad (\text{allgemeine Form einer oberen Dreiecksmatrix})$$

Lemma 8.24

Für eine obere Dreiecksmatrix $A \in \mathbb{R}^{n \times n}$ ist die Determinante das Produkt der Diagonaleinträge: $\det(A) = A_{1,1} \cdot A_{2,2} \cdot \dots \cdot A_{n,n}$

Beweis: Der Beweis erfolgt durch eine Induktion über die Dimension n der Matrix $A \in \mathbb{R}^{n \times n}$.

Induktions-Anfang: Man bemerke: Jede 1×1 -Matrix $A \in \mathbb{R}^{1 \times 1}$ ist eine obere Dreiecksmatrix³.

Es sei $n = 1$ und $A \in \mathbb{R}^{1 \times 1}$ eine obere Dreiecksmatrix. Die Matrix hat die Form $A = (A_{1,1})$ mit Diagonalelement $A_{1,1}$, und es gilt $\det(A) = A_{1,1}$.

Induktions-Voraussetzung: Für ein $n \in \mathbb{N}$ mit $n \geq 1$ gelte:

Für jede obere Dreiecksmatrix $A \in \mathbb{R}^{n \times n}$ ist die Determinante das Produkt der Diagonaleinträge.

Induktions-Behauptung: Dann gilt auch:

Für eine obere Dreiecksmatrix $A \in \mathbb{R}^{(n+1) \times (n+1)}$ ist die Determinante das Produkt der Diagonaleinträge.

Beweis: Es sei $A \in \mathbb{R}^{(n+1) \times (n+1)}$ eine obere Dreiecksmatrix. Entwickelt man die Determinante nach der ersten Spalte so erhält man wegen $A_{2,1} = A_{3,1} = \dots = A_{n+1,1} = 0$:

$$\det(A) = A_{1,1} \cdot \det(A^{1 \times 1}) - 0 \cdot \det(A^{2 \times 1}) \pm \dots \pm 0 \cdot \det(A^{n \times 1})$$

Die Streichungsmatrix $A^{1 \times 1}$ ist eine obere Dreiecksmatrix aus $\mathbb{R}^{n \times n}$ mit Diagonalelementen $A_{2,2}, \dots, A_{n+1,n+1}$. Entsprechend gilt nach I.Vor. $\det(A^{1 \times 1}) = A_{2,2} \cdot \dots \cdot A_{n+1,n+1}$:

$$\det \begin{pmatrix} A_{1,1} & A_{1,2} & A_{1,3} & \dots & A_{1,n+1} \\ 0 & A_{2,2} & A_{2,3} & \dots & A_{2,n+1} \\ 0 & 0 & A_{3,3} & \dots & A_{3,n+1} \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & 0 & A_{n+1,n+1} \end{pmatrix} = A_{1,1} \cdot \det \begin{pmatrix} A_{2,2} & A_{2,3} & \dots & A_{2,n+1} \\ 0 & A_{3,3} & \dots & A_{3,n+1} \\ \vdots & & \ddots & \vdots \\ 0 & \dots & 0 & A_{n+1,n+1} \end{pmatrix} \\ = A_{1,1} \cdot \underbrace{A_{2,2} \cdot \dots \cdot A_{n+1,n+1}}$$

Insgesamt gilt also $\det(A) = A_{1,1} \cdot A_{2,2} \cdot \dots \cdot A_{n+1,n+1}$. □

³Um eine Dreiecksmatrix zu sein müssen in A Elemente unterhalb der Diagonalen 0 sein (wenn es solche Elemente gibt). Die Matrix $(A_{1,1})$ erfüllt dies, denn sie hat keine solchen "nicht-diagonal-Elemente".

8.2 Lineare Gleichungssysteme

Definition 8.25

Ein lineares Gleichungssystem ist ein System von Gleichungen der Form

$$\begin{array}{cccccc} A_{1,1} \cdot x_1 & + A_{1,2} \cdot x_2 & + \dots + & A_{1,n} \cdot x_n & = & b_1 \\ A_{2,1} \cdot x_1 & + A_{2,2} \cdot x_2 & + \dots + & A_{2,n} \cdot x_n & = & b_2 \\ \vdots & \vdots & & \vdots & & \\ A_{m,1} \cdot x_1 & + A_{m,2} \cdot x_2 & + \dots + & A_{m,n} \cdot x_n & = & b_m \end{array}$$

mit $A_{1,1}, \dots, A_{m,n} \in \mathbb{R}$ und $b_1, \dots, b_m \in \mathbb{R}$.

Jedes Lineare Gleichungssystem lässt sich äquivalent schreiben als $A \cdot \vec{x} = \vec{b}$ mit $A \in \mathbb{R}^{m \times n}$ und $\vec{b} \in \mathbb{R}^m$.

Das Lösen eines Linearen Gleichungssystems kann per Gauß'schem Eliminationsverfahren erfolgen. Dabei wird eine Gleichung der Form $A \cdot \vec{x} = \vec{b}$ durch ein Tableau der Form $A|b$ ersetzt, auf dem Umformungen durchgeführt werden (sog. Gaußschritte), bis ein Tableau entsteht, das eine Matrix in *Zeilen-Stufen-Form* enthält:

Eine Matrix $A \in \mathbb{R}^{m \times n}$ hat Zeilen-Stufen-Form, wenn in jeder Zeile $z > 2$ mehr führende Nullen stehen als in der vorherigen Zeile $z - 1$ (es sei denn die Zeile $z - 1$ bestand schon komplett aus Nullen).

Definition 8.27

Eine Matrix $A \in \mathbb{R}^{m \times n}$ hat Zeilen-Stufen-Form, wenn es einen Zeilen-Index $\ell \in \{1, \dots, m + 1\}$ gibt mit:

- Jede Zeile mit Index in $\{2, \dots, \ell - 1\}$ enthält mindesten eine führende Null mehr als die Zeile davor.
- Jede Zeile mit Index in $\{\ell, \dots, m\}$ enthält nur Nullen (falls $\ell = m + 1$ gibt es keine solchen Zeilen).

Beispiel 8.27

Die folgenden Matrizen haben jeweils Zeilen-Stufen-Form mit $\ell = 4$:

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 3 & 5 & 7 & 9 \\ 0 & 0 & 0 & 2 & 4 \end{pmatrix} \quad B = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 3 & 5 & 7 & 9 \\ 0 & 0 & 0 & 2 & 4 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Anzahl führender Nullen steigt bis Zeile 3 = $\ell - 1$
 Anzahl führender Nullen ist maximal ab Zeile 4 = ℓ

(Anti-)Beispiel 8.28

Die folgende Matrix ist nicht in Zeilen-Stufen-Form, denn Zeile 2 enthält genauso viele führende Nullen wie Zeile 3, obwohl beide nicht maximal viele Nullen enthalten:

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 0 & 5 & 7 & 9 \\ 0 & 0 & 2 & 4 & 6 \end{pmatrix}$$

Ein Gleichungssystem $A \cdot \vec{x} = \vec{b}$ mit einer Matrix in Zeilen Stufen-Form zu lösen ist leicht rekursiv möglich.

Definition 8.29

Für eine Gleichung $A \cdot \vec{x} = \vec{b}$ mit $A \in \mathbb{R}^{m \times n}$ und $\vec{b} \in \mathbb{R}^m$ ist das zugehörige Tableau

$$\begin{array}{ccc|c} A_{1,1} & \dots & A_{1,n} & b_1 \\ \vdots & & \vdots & \vdots \\ A_{m,1} & \dots & A_{m,n} & b_m \end{array}$$

Die Lösungen eines Tableaus sind die Lösungen der zugehörigen Gleichung $A \cdot \vec{x} = \vec{b}$. Zwei Tableaus heißen *äquivalent*, wenn sie dieselben Lösungen haben.

Im Gaußverfahren sind folgende Schritte Zulässig auf einem Tableau:

Lemma 8.30

Die Lösungen eines Tableaus verändern sich nicht, wenn man einen der folgenden folgenden Schritte anwendet:

- Typ I) Addiere eine das Vielfache einer Zeile zu einer *anderen* Zeile.
- Typ II) Multipliziere eine Zeile des Tableaus mit einer Zahl.
- Typ III) Vertausche zwei Zeilen des Tableaus.
- Typ IV) Streiche eine Nullzeile (eine Zeile deren Einträge alle Null sind).

... wobei jeweils die anderen Zeilen des Tableaus unberührt bleiben.

Der Beweis für Schritte vom Typ I)-III) ergibt sich direkt aus den Eigenschaften von Gleichungen.

Beispiel 8.31

Gegeben sei $A := \begin{pmatrix} 2 & 1 & 1 \\ 4 & 3 & 3 \\ -6 & -5 & -5 \end{pmatrix}$ und $b := \begin{pmatrix} 1 \\ 3 \\ -5 \end{pmatrix}$ gesucht ist die Lösung \vec{x} von $A\vec{x} = \vec{b}$

$$\begin{array}{ccc|c} 2 & 1 & 1 & 1 \\ 4 & 3 & 3 & 3 \\ -6 & -5 & -5 & -5 \end{array} \quad \begin{array}{l} \text{II} - (2) \cdot \text{I} \\ \text{III} + (3) \cdot \text{I} \end{array} \quad \begin{array}{ccc|c} 2 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & -2 & -2 & -2 \end{array} \quad \begin{array}{l} \text{III} + (2) \cdot \text{II} \end{array} \quad \begin{array}{ccc|c} 2 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{array}$$

Das letzte Tableau hat bereits Zeilen-Stufen-Form. Die letzte Zeile übersetzt sich zu der Gleichung $0 = 0$, die die Lösungsmenge nicht einschränkt. Diese Gleichung kann also gestrichen werden. Nach Streichen der letzten Nullzeile erhält man aus dem Resultierenden Tableau die Gleichungen:

$$\begin{array}{l} \text{I} \quad 2x_1 + x_2 + x_3 = 1 \\ \text{II} \quad 0x_1 + x_2 + x_3 = 1 \end{array}$$

Setzt man $x_3 := \lambda$ mit $\lambda \in \mathbb{R}$ so erhält man nach umstellen:

$$\left. \begin{array}{l} \text{I} \quad 2 \cdot x_1 = 1 - x_2 - x_3 \stackrel{\text{II}}{=} 1 - (1 - \lambda) - \lambda = 0 \\ \text{II} \quad x_2 = 1 - x_3 = 1 - \lambda \end{array} \right\} \Rightarrow \vec{x} = \begin{pmatrix} 0 \\ 1 - \lambda \\ \lambda \end{pmatrix} \text{ mit } \lambda \in \mathbb{R}$$

Die Lösungsmege lautet also: $\mathbb{L} := \left\{ \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + \lambda \cdot \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} : \lambda \in \mathbb{R} \right\}$

8.2.1 Interpretation eines Tableaus mit Nullzeilen

Das Interpretieren eines Tableaus mit Nullzeile ist üblicherweise anfänglich etwas schwierig, aber im Grunde genommen einfach: Eine Nullzeile entspricht einer Gleichung $0 = 0$, und weil diese Gleichung immer wahr ist, schränkt sie die Startmenge \mathbb{R}^n nicht ein.

- “Fragt” man alle $x \in \mathbb{R}$, ob sie $x^2 = 4$ erfüllen, so “antworten” nur 2 und -2 mit ja:

$$\{x \in \mathbb{R} : x^2 = 4\} = \{-2, 2\}$$

- “Fragt” man alle $x \in \mathbb{R}$, ob sie $0 = 0$ erfüllen, so “antworten” *alle* mit ja, denn diese Gleichung ist stets erfüllt! Es gilt also:

$$\{x \in \mathbb{R} : 0 = 0\} = \mathbb{R}$$

- “Fragt” man alle $x \in \mathbb{R}$, ob sie $0 = 1$ erfüllen, so “antworten” *alle* mit nein, denn diese Gleichung ist immer falsch! Es gilt also:

$$\{x \in \mathbb{R} : 0 = 1\} = \emptyset$$

Nullzeilen

Hat man ein Tableau, das (am Ende) eine Nullzeile enthält, so entspricht dies der folgenden Situation: Das Tableau hat einen “Vorspann” aus einer Matrix $A \in \mathbb{R}^{m \times n}$ und $\vec{b} \in \mathbb{R}^m$, gefolgt von einer Nullzeile. Dieses Tableau hat dann die selben Lösungen wie $A \cdot \vec{x} = \vec{b}$:

Tableau:	$\begin{array}{ccc c} \vec{A}_{1,1} & \dots & A_{1,n} & b_1 \\ \vdots & & \vdots & \vdots \\ \vec{A}_{m,1} & \dots & A_{m,n} & b_m \\ 0 & \dots & 0 & 0 \end{array}$	Gleichungssystem	$\begin{cases} \vec{A}_{1,1}x_1 + \dots + A_{1,n}x_n = b_1 \\ \vdots \\ \vec{A}_{m,1}x_1 + \dots + A_{m,n}x_n = b_{m-1} \\ 0 = 0 \end{cases}$
----------	--	------------------	---

Die Lösungen dieses Systems sind:

$$\{ \vec{x} \in \mathbb{R}^n : A\vec{x} = \vec{b} \} \cap \underbrace{\{ \vec{x} \in \mathbb{R}^n : 0 = 0 \}}_{=\mathbb{R}^n} = \{ \vec{x} \in \mathbb{R}^n : A\vec{x} = \vec{b} \}$$



Fast-Nullzeilen

Hat man ein Tableau, das (am Ende) eine Nullzeile mit **Rechter Seite ungleich Null** enthält, so hat das zugehörige Gleichungssystem **keine Lösung**

Allgemein sieht dies wie folgt aus: Nach umsordieren der Zeilen, hat das Tableau einen “Vorspann” aus einer Matrix $A \in \mathbb{R}^{m \times n}$ und $\vec{b} \in \mathbb{R}^m$, gefolgt von einer weiteren Zeile:

Tableau:	$\begin{array}{ccc c} \vec{A}_{1,1} & \dots & A_{1,n} & b_1 \\ \vdots & & \vdots & \vdots \\ \vec{A}_{m,1} & \dots & A_{m,n} & b_m \\ 0 & \dots & 0 & 1 \end{array}$	Gleichungssystem	$\begin{cases} \vec{A}_{1,1}x_1 + \dots + A_{1,n}x_n = b_1 \\ \vdots \\ \vec{A}_{m,1}x_1 + \dots + A_{m,n}x_n = b_{m-1} \\ 0 = 1 \end{cases}$
----------	--	------------------	---

Die Lösungen dieses Systems sind:

$$\{ \vec{x} \in \mathbb{R}^n : A\vec{x} = \vec{b} \} \cap \underbrace{\{ \vec{x} \in \mathbb{R}^n : 0 = 1 \}}_{=\emptyset} = \emptyset$$

8.2.2 Interpretation eines Tableaus mit Sprüngen

Hat in einem Tableau eine Zeile $z + 1$ mehr als eine führende Nullen mehr als die vorhergehende Zeile so bezeichnet man dies als einen “Sprung”. Die folgende Matrix hat zwei Sprünge:

- einen Sprung der Länge **3** von Zeile 2 zu Zeile 3
- einen Sprung der Länge **2** von Zeile 4 zu Zeile 5

1	2	3	4	5	6	7
0	2	3	4	5	6	7
0	0	0	0	5	6	7
0	0	0	0	0	6	7
0	0	0	0	0	0	0

┌──────────┐ ┌────────┐
Länge 3 Länge 2

Hat ein Tableau in Zeilen-Stufen-Form einen Sprung der Länge k , so bedeutet dies, dass sich $k - 1$ der gegebenen Variablen *nicht konkret* festlegen lassen (s. Beispiel 8.31). Diese Variablen bleiben als Variablen in der Lösung erhalten, man drückt dies durch ersetzen der Variable mit einer weiteren, neuen Variable (meist Griechische Lettern) aus.

Beispiel 8.32

1	0	1	0	1	9	→	$x_1 = 9$	$-x_3$	$-x_5$
0	1	2	0	4	7		$x_2 = 7$	$-2x_3$	$-4x_5$
0	0	0	1	1	5		$x_4 = 5$	$-x_5$	

Hier lässt sich

- x_4 nur ausdrücken in Abhängigkeit von x_5 (oder umgekehrt).
- x_2 nur ausdrücken in Abhängigkeit von x_3 , x_4 und x_5 (oder umgekehrt).
- x_1 nur ausdrücken in Abhängigkeit von x_2 , x_3 , x_4 und x_5 (oder umgekehrt).

Hat man dies getan (s. Gleichungen oben), und legt man x_5 und x_3 fest, so lassen sich *alle* weiteren Variablen aus x_5 und x_3 berechnen. D.h. für *alle* Werte von x_5 und x_3 gibt es eine Lösungsvektor \vec{x} . Die Lösung hat also 2 Freiheitsgrade, bzw. ist 2 dimensional. Konkret erhält man mit $x_3 := \lambda$ und $x_5 := \mu$ ist die Lösungsmege:

$$\mathbb{L} = \left\{ \begin{pmatrix} 7 & -\lambda & -\mu \\ 7 & -2\lambda & -4\mu \\ \lambda & & \\ 5 & -\mu & \\ & & \mu \end{pmatrix} : \lambda, \mu \in \mathbb{R} \right\}$$

Kapitel 9

Folgen und Reihen

9.0.3 Folgen

Definition 9.1

Eine Folge $(a_n)_{n \in \mathbb{N}}$ reeller Zahlen ist eine Abbildung $\mathbb{N} \rightarrow \mathbb{R}$, der Form $n \mapsto a_n$.
Man schreibt die Folge in der Form $(a_n)_{n \in \mathbb{N}} = (a_0, a_1, a_2, a_3, \dots)$.

Eine Folge kann angegeben werden:

- aufzählend, z.B. $(a_n)_{n \in \mathbb{N}} = (1, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \dots)$
- definierend, z.B. $a_n := \frac{1}{2^n}$
- rekursiv durch einen Startwert und eine Rekursionsgleichung: z.B. $a_0 = 1$ und $a_{n+1} = \frac{1}{2} \cdot a_n$.

Beispiel 9.2

Die Folge $(a_n)_{n \in \mathbb{N}} = (0, 1, 2, 3, 4, \dots)$ der natürlichen Zahlen lässt sich wie folgt angeben:

- durch die Definition $a_n := n$
- rekursiv durch einen Startwert und eine Rekursionsgleichung: z.B. $a_0 = 0$ und $a_{n+1} = a_n + 1$.

Definition 9.3

Eine Zahl $a \in \mathbb{R}$ heißt Grenzwert (oder Limes) der Folge $(a_n)_{n \in \mathbb{N}}$, wenn folgende Bedingung erfüllt ist:

Zu jedem reellen Abstand $\varepsilon > 0$ existiert ein Index $N \in \mathbb{N}$, ab dem für alle $n \geq N$ gilt:

a_n ist dichter an a als der Abstand ε .

Eine Folge $(a_n)_{n \in \mathbb{N}}$ konvergiert also gegen a wenn gilt:

$$\forall \varepsilon \in \mathbb{R}, \varepsilon > 0 : \exists N \in \mathbb{N} : |a_n - a| \leq \varepsilon \quad \forall n \in \mathbb{N}, n \geq N$$

In diesem Fall schreibt man $a = \lim_{n \rightarrow \infty} a_n$ an und sagt, dass " $(a_n)_{n \in \mathbb{N}}$ gegen a konvergiert".

Beispiel 9.4

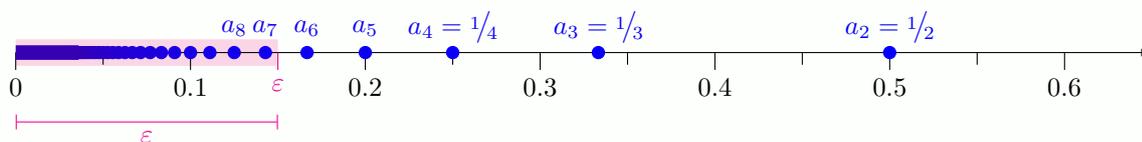
Die Folge $(a_n)_{n \in \mathbb{N}}$ mit $a_n = \frac{1}{n}$ hat den Grenzwert $a = 0$:

Zu gegebenem $\varepsilon > 0$ definiert man $N_\varepsilon := \lceil \frac{1}{\varepsilon} \rceil$ (dies ist Aufrunden von $\frac{1}{\varepsilon}$). Für alle $n \geq N$ gilt dann zunächst $|a_n - a| = |a_n - 0| = |a_n| = \frac{1}{n}$ und dies lässt sich wie folgt gegen ε abschätzen:

$$|a_n - a| = \frac{1}{n} \leq \frac{1}{N} = \frac{1}{\lceil \frac{1}{\varepsilon} \rceil} \leq \frac{1}{1/\varepsilon} = \varepsilon$$

Ein Beispiel: Für $\varepsilon = 0.15$ ist $N_\varepsilon := \lceil \frac{1}{0.15} \rceil = \lceil \frac{100}{15} \rceil = 7$.

Es gilt also: Ab $n = 7$ ist der Abstand $|a_n - 0|$ kleiner als 0.16, d.h. $|a_7 - 0|, |a_8 - 0|, |a_9 - 0|, \dots \leq \varepsilon$

**Definition 9.5**

Eine Folge $(a_n)_{n \in \mathbb{N}}$ heißt:

- nach *oben* beschränkt, falls es eine obere Schranke $A \in \mathbb{R}$ gibt, so dass $a_n \leq A$ für alle $n \in \mathbb{N}$ gilt.
- nach *unten* beschränkt, falls es eine untere Schranke $\tilde{A} \in \mathbb{R}$ gibt, so dass $a_n \geq \tilde{A}$ für alle $n \in \mathbb{N}$ gilt.
- *monoton wachsend*, falls für alle $n \in \mathbb{N}$ gilt: $a_n \leq a_{n+1}$
- *streng monoton wachsend*, falls für alle $n \in \mathbb{N}$ gilt: $a_n < a_{n+1}$
- *monoton fallend*, falls für alle $n \in \mathbb{N}$ gilt: $a_n \geq a_{n+1}$
- *streng monoton fallend*, falls für alle $n \in \mathbb{N}$ gilt: $a_n > a_{n+1}$

Beispiel 9.6

Die Folge $(a_n)_{n \in \mathbb{N}} = (0, 1, 2, 3, 4, \dots)$ ist

- *streng monoton wachsend* denn es gilt $a_n < a_{n+1} = a_n + 1$ für alle $n \in \mathbb{N}$
- *nach unten beschränkt*, denn es gilt $a_n \geq 0$ für alle $n \in \mathbb{N}$

Hinweise: Die Folge ist *nicht* nach oben beschränkt.

Die Folge ist auch “monoton wachsend” (ohne den Zusatz “streng”) denn die schwächere Forderung “ $a_n \leq a_{n+1}$ ” ist auch erfüllt!

Beispiel 9.7

Die Folge $(a_n)_{n \in \mathbb{N}} = (1, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \dots)$ mit $a_n := \frac{1}{2^n}$ ist

- **streng monoton fallend** denn es gilt $a_n > a_{n+1} = a_n \cdot \frac{1}{2}$ für alle $n \in \mathbb{N}$
- **nach unten beschränkt**, denn es gilt $a_n \geq 0$ für alle $n \in \mathbb{N}$
- **nach oben beschränkt**, denn es gilt $a_n \leq 1$ für alle $n \in \mathbb{N}$

Hinweis:

Die Folge ist auch “monoton fallend” (ohne den Zusatz “streng”) denn die schwächere Forderung “ $a_n \geq a_{n+1}$ ” ist auch erfüllt!

Das folgende Konvergenzkriterium liefert eine *qualitative* Aussage über das Konvergenzverhalten einer Folge, d.h. wir erfahren *ob* eine Folge konvergiert oder nicht. Das Kriterium liefert aber *keine quantitative* Aussage, d.h. *wogegen* die Folge konvergiert bleibt unklar.

Lemma 9.8

Jede nach *oben beschränkte* und *monoton wachsende* Folge ist konvergent.
Jede nach *unten beschränkte* und *monoton fallende* Folge ist konvergent.

Beispiel 9.9

Die Folge $(a_n)_{n \in \mathbb{N}} = (0, \frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \dots)$ mit $a_n := \frac{n}{n+1}$ ist konvergent, denn:

- Offensichtlich gilt $\frac{n}{n+1} \leq \frac{n}{n+1} + \frac{1}{n+1} = 1$ und damit ist die Folge **nach oben beschränkt**.
- Weiter ist die Folge **monoton wachsend**, d.h. es gilt stets $a_n \leq a_{n+1}$ wegen:

$$\begin{aligned} \frac{n}{n+1} &\leq \frac{n+1}{n+2} && | \cdot (n+1) \cdot (n+2) \\ \Leftrightarrow n \cdot (n+2) &\leq (n+1) \cdot (n+1) \\ \Leftrightarrow n^2 + 2n &\leq n^2 + 2n + 1 && \text{(eine wahre Aussage)} \end{aligned}$$

Beispiel 9.10

Sei $0 \leq q < 1$. Die Folge $(a_n)_{n \in \mathbb{N}}$ mit $a_n := q^n$ konvergiert, denn diese Folge ist monoton fallend und nach unten beschränkt durch 0.

9.0.4 Berechnen von Grenzwerten

Das folgende Lemma enthält implizit ein Rechenverfahren zum Berechnen von Grenzwerten von zusammengesetzten Funktionen. Im Prinzip darf man das Bilden des Grenzwerts an allen bekannten Rechenzeichen “zerteilen”.



Um formal korrekt zu bleiben ist es wichtig, dass in der entsprechenden Berechnung **zuerst** in einer Nebenrechnung **die Grenzwerte der Teilfunktionen berechnet werden**, um diese *erst danach* zu einem

Ganzen zusammensetzen:

Lemma 9.11

Es seien $(a_n)_{n \in \mathbb{N}}$ und $(b_n)_{n \in \mathbb{N}}$ zwei konvergente Folgen mit Grenzwerten

$$\lim_{n \rightarrow \infty} a_n = a \quad \text{und} \quad \lim_{n \rightarrow \infty} b_n = b$$

Dann gelten für die zusammengesetzten Folgen:

- **Addition** $\lim_{n \rightarrow \infty} (a_n + b_n) = a + b$ sowie
 $\lim_{n \rightarrow \infty} (a_n - b_n) = a - b$
- **Multiplikation** $\lim_{n \rightarrow \infty} (a_n \cdot b_n) = a \cdot b$

Gilt zusätzlich $b \neq 0$ so folgt

- **Division** $\lim_{n \rightarrow \infty} \frac{a_n}{b_n} = \frac{a}{b}$

Beispiel 9.12

Aus Beispiel 9.4 wissen wir bereits, dass $\lim_{n \rightarrow \infty} \frac{1}{n} = 0$ gilt.

Entsprechend folgern wir für $\frac{1}{n^3}$:

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{n^2} = 0 \quad \text{wegen} \quad \lim_{n \rightarrow \infty} \frac{1}{n^2} &= \lim_{n \rightarrow \infty} \frac{1}{n} \cdot \frac{1}{n} = 0 \cdot 0 \\ \lim_{n \rightarrow \infty} \frac{1}{n^3} = 0 \quad \text{wegen} \quad \lim_{n \rightarrow \infty} \frac{1}{n^3} &= \lim_{n \rightarrow \infty} \frac{1}{n} \cdot \frac{1}{n^2} = 0 \cdot 0 \end{aligned}$$

Induktiv lässt sich analog $\lim_{n \rightarrow \infty} \frac{1}{n^k} = 0$ für alle $k \in \mathbb{N} \setminus \{0\}$ zeigen.

Für rationale Funktionen (Brüche aus Polynomen) in $n \in \mathbb{N}$ gibt es ein einfaches Verfahren zum Berechnen des Grenzwertes:

1. Finde die größte im Nenner auftretende Potenz n^k und kürze den Bruch mit dieser.
2. Berechne den Grenzwert unter Verwendung von $\lim_{n \rightarrow \infty} 1/n^k = 0$ und $\lim_{n \rightarrow \infty} n^k = \infty$ und

Beispiel 9.13

Es sei $a_n := \frac{8 \cdot n^2 + 3 \cdot n^3}{3 \cdot n + 2 \cdot n^4}$. Dann gilt:

$$\lim_{n \rightarrow \infty} \frac{8 \cdot n^2 + 3 \cdot n^3}{3 \cdot n + 2 \cdot n^4} = \lim_{n \rightarrow \infty} \frac{(8 \cdot \frac{1}{n^2} + 3 \cdot \frac{1}{n}) \cdot n^4}{(3 \cdot \frac{1}{n^3} + 2 \cdot 1) \cdot n^4} = \lim_{n \rightarrow \infty} \frac{(8 \cdot \overset{0}{\frac{1}{n^2}} + 3 \cdot \overset{0}{\frac{1}{n}})}{(3 \cdot \underset{0}{\frac{1}{n^3}} + 2 \cdot \underset{2}{1})} = \frac{0 + 0}{0 + 2} = 0$$

Denn es gelten: $\lim_{n \rightarrow \infty} \frac{1}{n} = 0$, $\lim_{n \rightarrow \infty} \frac{1}{n^2} = 0$ und $\lim_{n \rightarrow \infty} \frac{1}{n^3} = 0$

Beispiel 9.14

Es sei $a_n := \frac{8 \cdot n^2 + 3 \cdot n^4}{3 \cdot n + 2 \cdot n^4}$. Dann gilt:

$$\lim_{n \rightarrow \infty} \frac{8 \cdot n^2 + 3 \cdot n^4}{3 \cdot n + 2 \cdot n^4} = \lim_{n \rightarrow \infty} \frac{(8 \cdot \frac{1}{n^2} + 3 \cdot 1) \cdot n^4}{(3 \cdot \frac{1}{n^3} + 2 \cdot 1) \cdot n^4} = \lim_{n \rightarrow \infty} \frac{(8 \cdot \frac{1}{n^2} + 3 \cdot 1)}{(3 \cdot \frac{1}{n^3} + 2 \cdot 1)} = \frac{0 + 3}{0 + 2} = 1.5$$

Denn es gelten: $\lim_{n \rightarrow \infty} \frac{1}{n^3} = 0$ und $\lim_{n \rightarrow \infty} \frac{1}{n^2} = 0$.

Ist der Grad des Zählers größer als der Grad des Nenners, so kann das Konvergenzverhalten durch Abschätzungen bestimmt werden:

Beispiel 9.15

Es sei $a_n := \frac{8 \cdot n^2 + 3 \cdot n^5}{3 \cdot n + 2 \cdot n^4}$. Dann gilt:

$$a_n = \frac{8 \cdot n^2 + 3 \cdot n^5}{3 \cdot n + 2 \cdot n^4} = \frac{(8 \cdot \frac{1}{n^2} + 3 \cdot n) \cdot n^4}{(3 \cdot \frac{1}{n^3} + 2 \cdot 1) \cdot n^4} = \frac{(8 \cdot \frac{1}{n^2} + 3 \cdot n)}{(3 \cdot \frac{1}{n^3} + 2 \cdot 1)} \geq \frac{8 \cdot 0 + 3 \cdot n}{3 \cdot 1 + 2 \cdot 1} = \frac{3}{5} \cdot n$$

Es gilt also $\lim_{n \rightarrow \infty} a_n = \infty$, denn die Folge ist monoton wachsend und unbeschränkt.

**Rechnen mit ∞ ist nicht möglich!**

Das Symbol ∞ steht nicht für eine Zahl, entsprechend versagen bei ∞ die üblichen Rechengesetze. Insbesondere der Wert von " $\frac{\infty}{\infty}$ " ist nicht bestimmt.

a) " $\frac{\infty}{\infty}$ " ist 0? $\lim_{n \rightarrow \infty} \frac{42 \cdot n^2}{n^3} = \lim_{n \rightarrow \infty} \frac{42}{n} = 0$

b) " $\frac{\infty}{\infty}$ " ist ∞ ? $\lim_{n \rightarrow \infty} \frac{42 \cdot n^3}{n^2} = \lim_{n \rightarrow \infty} \frac{42 \cdot n}{1} = \infty$

c) " $\frac{\infty}{\infty}$ " ist 42? $\lim_{n \rightarrow \infty} \frac{42 \cdot n^2}{n^2} = \lim_{n \rightarrow \infty} \frac{42 \cdot 1}{1} = 42$

9.0.5 Reihen

Vereinfacht ausgedrückt, ist eine *Reihe* eine "Summe unendlich vieler Summanden". Wenn aber unendlich viele Werte addiert werden, kann der Wert dieser Summe auch unendlich groß werden und dadurch gar nicht existieren. Bei manchen Summen ist ihr Wert, oder sogar die Existenz des Wertes, unklar:

$$1 - 1 + 1 - 1 + 1 \pm \dots = ?$$

$$1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \dots = ?$$

$$1 + \frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{4^2} + \frac{1}{5^2} + \dots = ?$$

Bei solchen “unendlichen Summen” müssen wir also offensichtlich Grenzwerte betrachten, wie wir sie von Folgen bereits kennen:

Definition 9.16

Es sei $(a_n)_{n \in \mathbb{N}}$ eine Folge. Die *Partialsomme* der Folge bis zum N -ten Glied ist $\sum_{k=0}^N a_k$. Die *Reihe* über die a_n der Grenzwert der Partialsommen, d.h.

$$\sum_{k=0}^{\infty} a_k := \lim_{N \rightarrow \infty} \sum_{k=0}^N a_k$$

Nicht konvergente Reihen heißen divergent.

Bemerkung 9.1 Die Reihe muss nicht bei $k = 0$ oder $k = 1$ beginnen, sondern kann bei jeder *ganzen* Zahl $m \in \mathbb{Z}$ beginnen: Analog zu obiger Definition definiert man dann $\sum_{k=m}^{\infty} a_k$.

Ändert man in einer Reihe (nur) *endlich viele* Summanden, so ändert sich das *Konvergenzverhalten* (Konvergenz oder Divergenz) nicht. Natürlich kann sich hierdurch der Wert der Reihe ändern (falls die die Reihe konvergiert, d.h. falls sie überhaupt “einen Wert hat”). Gleiches gilt für das Weglassen oder Hinzufügen von *endlich vielen* Summanden.

Beispiel 9.17

Es gilt

$$1 = \sum_{k=2}^{\infty} \frac{1}{(k-1) \cdot k} = \frac{1}{2} + \frac{1}{6} + \frac{1}{12} + \frac{1}{20} + \dots$$

Diese Reihe ist also *konvergent*. Denn wegen $\frac{1}{k-1} - \frac{1}{k} = \frac{k}{(k-1)k} - \frac{k-1}{(k-1)k} = \frac{k-k+1}{(k-1) \cdot k}$ (ein Trick!) folgt für die Partialsomme bis $N \in \mathbb{N}$:

$$\begin{aligned} \sum_{k=2}^N \frac{1}{(k-1)k} &= \sum_{k=2}^N \left(\frac{1}{(k-1)} - \frac{1}{k} \right) \\ &= \underbrace{\left(\frac{1}{1} - \frac{1}{2} + \frac{1}{2} - \frac{1}{3} + \frac{1}{3} - \frac{1}{4} + \dots + \frac{1}{N-1} - \frac{1}{N} \right)}_{=0} = 1 - \frac{1}{N} \end{aligned}$$

Es folgt also

$$\sum_{k=2}^{\infty} \frac{1}{(k-1)k} = \lim_{N \rightarrow \infty} \left(\sum_{k=2}^N \frac{1}{(k-1)k} \right) = \lim_{N \rightarrow \infty} \left(1 - \frac{1}{N} \right) = 1$$

(Anti-)Beispiel 9.18

Die Reihe $\sum_{k=0}^{\infty} k = 0 + 1 + 2 + 3 + 4 + 5 + \dots$ ist *divergent*, denn die Folge der Partialsummen $0, 1, 3, 6, 10, 15, \dots$ ist divergent. Genauer gilt $\sum_{k=0}^{\infty} k = \infty$ wegen

$$\sum_{k=0}^N k = \frac{N(N+1)}{2} \quad \text{und} \quad \lim_{N \rightarrow \infty} \left(\sum_{k=0}^N k \right) = \lim_{N \rightarrow \infty} \frac{N(N+1)}{2} = \infty$$

Die Reihe $\sum_{k=0}^{\infty} (-1)^k = 1 - 1 + 1 - 1 \pm \dots$ ist *divergent*, denn die Folge der Partialsummen $1, 0, 1, 0, 1, \dots$ ist divergent, weil sie zwei Häufungspunkte hat (nämlich 1 und 0). Hier lässt sich als "Wert" für die Reihe nichts angeben, im Gegensatz zur ersten Reihe in diesem Beispiel.

Lemma 9.19 (Majorantenkriterium)

Es sei $\sum_{n=0}^{\infty} b_n = b$ eine konvergente Reihe mit Grenzwert $b \in \mathbb{R}$.

Gilt für eine Folge $(a_n)_{n \in \mathbb{N}}$ die Abschätzung $0 \leq a_n \leq b_n$ für alle $n \in \mathbb{N}$, so ist die Reihe über die a_n **konvergent** und es gilt:

$$0 \leq \sum_{n=0}^{\infty} a_n \leq \sum_{n=0}^{\infty} b_n = b$$

Beweis: Es seien $(a_n)_{n \in \mathbb{N}}$ und $(b_n)_{n \in \mathbb{N}}$ zwei Folgen mit $0 \leq a_n \leq b_n$ und die Reihe über die b_n sei konvergent mit Grenzwert $b = \sum_{n=0}^{\infty} b_n$. Dann gilt für die Folge der Partialsummen $S_N := \sum_{n=0}^N a_n$:

- Die Folge S_N ist monoton wachsend, es gilt also $S_{N+1} \geq S_N$ für alle $N \in \mathbb{N}$:

$$S_{N+1} = \sum_{n=0}^{N+1} a_n = \underbrace{a_{N+1}}_{\geq 0} + \underbrace{\sum_{n=0}^N a_n}_{=S_N} \geq 0 + S_N$$

- Die Folge S_N ist nach oben beschränkt, denn es gilt $S_N \leq b$ für alle $N \in \mathbb{N}$:

$$S_N = \sum_{n=0}^N a_n \leq \sum_{n=0}^N b_n \leq \sum_{n=0}^N b_n + \underbrace{\sum_{n=N+1}^{\infty} b_n}_{\geq 0} = b$$

Die Folge $(S_N)_{N \in \mathbb{N}}$ ist also monoton wachsend und nach oben beschränkt und damit konvergent. \square

Beispiel 9.20

Die Reihe $\sum_{i=1}^{\infty} \frac{1}{n^2}$ ist konvergent, denn die konvergente Reihe $1 = \sum_{n=2}^{\infty} \frac{1}{(n-1) \cdot n}$ aus 9.17 lässt sich als Majorante nutzen.

Um dies zu erreichen müssen wir die Folge $\frac{1}{(n-1) \cdot n}$ "vorne" ergänzen, weil dieser Bruch für $n = 0$ nicht definiert ist. Wir setzen:

$$b_n := \begin{cases} 1 & \text{falls } n = 1 \\ \frac{1}{(n-1) \cdot n} & \text{sonst} \end{cases}$$

Für die Folge $a_n := \frac{1}{n^2}$ und b_n gilt dann $0 \leq a_n \leq b_n$ für alle $n \in \mathbb{N}$, denn

$$a_1 = 1 = b_1 \quad \text{und} \quad a_n = \frac{1}{n \cdot n} \leq \frac{1}{(n-1) \cdot n} = b_n \quad \text{für } n \in \mathbb{N} \setminus \{0\}$$

Es gilt also

$$0 \leq \sum_{n=1}^{\infty} a_n \leq \sum_{n=1}^{\infty} b_n = 1 + \underbrace{\sum_{n=2}^{\infty} \frac{1}{(n-1) \cdot n}}_{=1 \text{ s. Bsp. 9.17}} = 2$$

9.0.6 Wichtige Reihen

Die folgenden Reihen sind in der Mathematik und Informatik wichtig:

Lemma 9.21 (geometrische Reihe)

Die *geometrische Reihe* hat die Form $\sum_{k=0}^{\infty} q^k$ mit $q \in \mathbb{R}$

Für die Partialsummen der geometrischen Reihe gilt für jedes $q \in \mathbb{R}$:

$$\sum_{k=0}^n q^k = \frac{1 - q^{n+1}}{1 - q}$$

Für $-1 < q < 1$ konvergiert die geometrische Reihe, es gilt:

$$\sum_{k=0}^{\infty} q^k = \frac{1}{1 - q} \quad \text{für } -1 < q < 1$$

Für $1 \leq q$ gilt $\sum_{k=0}^{\infty} q^k = \infty$ und für $q \leq -1$ ist die geometrische Reihe unbestimmt divergent.

Beweis: Für die Partialsumme $S_n := \sum_{k=0}^n q^k$ gilt:

$$\begin{array}{rcl}
 1 \cdot \sum_{k=0}^n q^k & = & 1 + q + q^2 + \dots + q^n \\
 -q \cdot \sum_{k=0}^n q^k & = & -q - q^2 + \dots - q^n - q^{n+1} \\
 \hline
 1 \cdot (\sum_{k=0}^n q^k) - q \cdot (\sum_{k=0}^n q^k) & = & 1 - q^{n+1}
 \end{array}$$

Es gilt also $(1 - q) \cdot S_n = 1 - q^{n+1}$. Teilen durch $(1 - q)$ auf beiden Seiten liefert:

$$S_n = \sum_{k=0}^n q^k = \frac{1 - q^{n+1}}{1 - q}$$

Für $-1 < q < 1$ gilt $\lim_{n \rightarrow \infty} q^{n+1} = 0$ und es folgt:

$$\sum_{k=0}^{\infty} q^k = \lim_{n \rightarrow \infty} \frac{1 - q^{n+1}}{1 - q} = \frac{1 - 0}{1 - q}$$

□

Kapitel 10

Taylorreihen

Dieser Abschnitt beschäftigt sich mit Taylorreihen, Taylorpolynomen und der Restgliedabschätzung für Taylorpolynome.

Die *Taylorreihe* einer reellen Funktion ist eine spezielle Potenzreihe, die sich recht leicht aus den Ableitungen der Funktion berechnen lässt. Unter bestimmten Voraussetzungen ist die berechnete Taylorreihe identisch mit der gegebenen Funktion. Dies macht Taylorreihen zu einem äußerst wichtigen mathematischen Hilfsmittel im Umgang mit Funktionen, sie sind sowohl beim Modellieren von Problemen als auch beim Lösen von Gleichungen äußerst Hilfreich.

Taylorpolynome, die “endliche Version” und quasi der “kleine Bruder” der Taylorreihe, werden häufig als Approximation für Funktionen benutzt: Wenn das Rechnen mit der tatsächlichen Funktion nicht oder nur schwer möglich ist, bietet ein Taylorpolynom eine leicht handhabbare Approximation der Funktion – mit einer Gütegarantie.

10.1 Voraussetzungen

Im folgenden benötigen wir die folgenden Namenskonventionen:

Für ganze Zahlen $n \in \mathbb{N}$ ist **die Fakultät** $n!$ das Produkt aller ganzen Zahlen bis zu n :

$$\begin{aligned}n! &:= 1 \cdot 2 \cdot 3 \cdot 4 \cdots (n-1) \cdot n \\4! &= 1 \cdot 2 \cdot 3 \cdot 4 = 24\end{aligned}$$

Als besondere Definition setzt man:

$$0! := 1$$

Ein Monom hat die Form $a \cdot x^k$.

Hierbei ist

- x die *Variable*
- $k \in \mathbb{N}$ der ganzzahlige *Exponent*
- $a \in \mathbb{R}$ der reelle *Koeffizient*

Null als Exponent:

Für alle reellen Zahlen $x \in \mathbb{R}$ gilt $x^0 = 1$

10.1.1 Potenzreihen

Potenzreihen sind zum einen Funktionen, d.h. zu einer Variable x wird ein Funktionswert errechnet. Gleichzeitig ist jede Potenzreihe eine *Reihe*, also die Summe unendlich vieler Summanden, nämlich die Summe *unendlich vieler* Monome $a_n x^n$.

Entsprechend ist beim Umgang mit Potenzreihen Vorsicht geboten: Nicht immer ergibt sich beim Einsetzen von Werten für x ein sinnvoller Wert (s. Beispiel 10.3).

Definition 10.1

Die **allgemeine Form einer Potenzreihe** lautet

$$f(x) = \sum_{k=0}^{\infty} a_k (x - m)^k = a_0 + a_1(x - m)^1 + a_2(x - m)^2 + a_3(x - m)^3 + \dots$$

- Hier ist “ x ” die *Variable*, für die spezielle Werte eingesetzt werden können.
- Die Zahl “ m ” nennt man den *Mittelpunkt* der Potenzreihe.
- Die Folge aus **reellen** Zahlen $(a_n)_{n \in \mathbb{N}} = a_0, a_1, a_2, a_3, \dots$ heißen die *Koeffizienten* der Potenzreihe.

Beispiel 10.2

Die Reihe

$$\underbrace{\sum_{k=0}^{\infty} \frac{1}{k!} x^k}_{\text{Summen-Schreibweise}} = \underbrace{\frac{1}{0!} + \frac{1}{1!} x^1 + \frac{1}{2!} x^2 + \frac{1}{3!} x^3 + \dots}_{\text{“Pünktchen-Schreibweise”}}$$

nennt man die *Eulersche Funktion* oder auch *e-Funktion*, abgekürzt e^x .

- Hier ist der Mittelpunkt $m = 0$ und
- die allgemeine Formel für die Koeffizienten lautet $a_k = \frac{1}{k!}$.

10.1.2 Einsetzen von Zahlen in Potenzreihen

Das Auswerten einer Potenzreihe $f(x)$ muss nicht immer “gut gehen”: Es kann Werte für x geben, für die $f(x)$ divergiert:

Beispiel 10.3

Die bereits bekannte geometrische Reihe

$$\sum_{k=0}^{\infty} 1 \cdot x^k = 1 + x^1 + x^2 + x^3 + \dots$$

kann für $x = 1$ keinen endlichen Wert annehmen, weil eine Summe von unendlich vielen Einsen größer sein muß als jede bekannte Zahl. Man sagt entsprechend “diese Potenzreihe divergiert am Punkt $x = 1$ ”. Setzt man jedoch $\frac{1}{2}$ für x ein, so ergibt sich der Wert 2:

$$\begin{aligned} \sum_{k=0}^{\infty} (1)^k &= 1 + 1^1 + 1^2 + 1^3 + \dots = \infty \\ \sum_{k=0}^{\infty} \left(\frac{1}{2}\right)^k &= 1 + \left(\frac{1}{2}\right)^1 + \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^3 + \dots = 2 \end{aligned}$$

Allgemein gilt: Die geometrische Reihe konvergiert für alle $x \in \mathbb{R}$ mit $-1 < x < 1$.

Für weitergehende Informationen über das Konvergenzverhalten von Reihen, insbesondere über den sogenannten *Konvergenzradius*, empfehlen wir einen Blick in die entsprechende Literatur.

10.2 Die Taylorreihe einer Funktion

Die Taylorreihe einer Funktion f ist eine spezielle Potenzreihe, die sich recht leicht aus den Ableitungen von f berechnen lässt. Die Taylorreihe stimmt unter bestimmten Umständen(!) rund um den *Entwicklungspunkt* mit der gegebenen Funktion überein.

Entsprechend benutzt man die Taylorreihe (bzw. das Taylorpolynom) einer Funktion häufig in komplexeren Rechnungen, weil sich mit ihr häufig leichter rechnen lässt als mit der eigentlichen Funktion selbst. Ein Taschenrechner verwendet beispielsweise zum Berechnen (d.h. eigentlich: zum Annähern) von Wurzelausdrücken ein Taylorpolynom der Funktion \sqrt{x} .

Der anstrengende und fehleranfällige Teil der Berechnung einer Taylorreihe ist das Bilden der zahlreichen benötigten Ableitungen. Hierzu verwendet man heute meistens spezielle Software.

10.2.1 Die allgemeine Form der Taylorreihe:

In der nachfolgenden Definition ist $f^{(k)}(m)$ die k -te Ableitung von f an der Stelle m , beispielsweise gelten $f^{(0)}(m) = f(m)$, $f^{(1)}(m) = f'(m)$, $f^{(2)}(m) = f''(m)$ etc.

Definition 10.4

Sei $f : \mathbb{R} \rightarrow \mathbb{R}$ eine am Punkt $m \in \mathbb{R}$ unendlich oft differenzierbare Funktion, dann heißt

$$\mathbb{T}^m f(x) := \sum_{k=0}^{\infty} f^{(k)}(m) \cdot \frac{(x-m)^k}{k!}$$

die *Taylorreihe der Funktion f am Entwicklungspunkt m* .

Für die Taylorreihe am Entwicklungspunkt $m = 0$ schreibt man auch kurz $\mathbb{T}f$, d.h. es gilt:

$$\mathbb{T}f(x) := \sum_{k=0}^{\infty} f^{(k)}(m) \cdot \frac{x^k}{k!}$$

Beispiel 10.5

Um die Taylorreihe von $f(x) := e^x$ am Entwicklungspunkt $m = 0$ zu berechnen, benötigen wir zunächst alle Ableitungen an diesem Punkt.

$$\begin{aligned} f(x) &= e^x, & f'(x) &= e^x, & f''(x) &= e^x, & f^{(3)}(x) &= e^x, & \dots \\ f(0) &= e^0 = 1, & f'(0) &= 1, & f''(0) &= 1, & f^{(3)}(0) &= 1, & \dots \end{aligned}$$

Entsprechend ergibt sich die Taylorreihe wie folgt:

$$\begin{aligned} \mathbb{T}f(x) &:= \sum_{k=0}^{\infty} f^{(k)}(0) \cdot \frac{(x-0)^k}{k!} \\ &= \sum_{i=0}^{\infty} 1 \cdot \frac{x^k}{k!} &= 1 + \frac{x^1}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots \end{aligned}$$



Man beachte, dass beim Berechnen einer Taylorreihe zunächst stets *Zahlen* berechnet werden, d.h. die Werte für $f^{(k)}(0)$ sind vom Typ “reelle Zahl”. Ein typischer Anfängerfehler sind “Reihen” der Form $e^x + e^x \cdot x + e^x \cdot x^2/2! + \dots$. Dies sind dann aber *keine* Potenzreihen und schon gar keine Taylorreihen!

10.2.2 Eigenschaften der Taylorreihe

Die Taylorreihe muss nicht für jeden Wert von x konvergieren. Wenn die Taylorreihe jedoch in einem Intervall konvergiert, so ist sie auf dem Intervall identisch mit der Ausgangsfunktion:

Satz 10.6

Sei $\mathbb{T}f(x) := \sum_{k=0}^{\infty} f^{(k)}(0) \frac{x^k}{k!}$ die Taylorreihe von f am Entwicklungspunkt $m = 0$. Falls $\mathbb{T}f(x)$ für alle x in einem Intervall $[-a, a]$ um die 0 konvergiert, gilt:

$$f(x) = \mathbb{T}f(x) \quad \text{für alle } x \in (-a, a).$$

Grob gesagt folgt aus diesem Theorem: Wenn man bereits “irgendeine” Reihendarstellung von f gefunden hat, so ist diese Reihe die Taylorreihe von f .

Korollar 10.7

Es sei $f : \mathbb{R} \rightarrow \mathbb{R}$ eine beliebig oft differenzierbare Funktion und $g(x) := \sum_{k=0}^{\infty} a_k x^k$ eine Potenzreihe. Falls auf einem Intervall $(-a, a)$ mit $a \in \mathbb{R}$ und $a > 0$ gilt:

- i) $g(x)$ konvergiert für alle $x \in (-a, a)$ ii) $f(x) = g(x)$ gilt für alle $x \in (-a, a)$

dann gilt $\mathbb{T}f(x) = \sum_{k=0}^{\infty} a_k x^k$

Beispiel 10.8

Für $f(x) = \frac{1}{1-x}$ lautet die Taylorreihe $\mathbb{T}f(x) = \sum_{k=0}^{\infty} x^k$.

Denn nach Lemma 9.21 gilt für die *geometrische Reihe*:

- i) Die Reihe konvergiert für alle $x \in (-1, 1)$
- ii) es gilt $\sum_{k=0}^{\infty} x^k = \frac{1}{1-x}$ für alle $x \in (-1, 1)$.

Probe: Rechnet man tatsächlich nach, so erhält man für die k -te Ableitung die allgemeine Formel

$$f^{(k)}(x) = k! \cdot (1-x)^{k+1} \quad \text{bzw.} \quad f^{(k)}(0) = k!$$

und daraus ergibt sich

$$\mathbb{T}f(x) = \sum_{k=0}^{\infty} k! \frac{x^k}{k!} = \sum_{k=0}^{\infty} x^k$$

Diese Berechnung von $\mathbb{T}f(x)$ kann man sich durch Korollar 10.7 sparen.

Die in Satz 10.6 genannte Eigenschaft der Taylorreihe $\mathbb{T}f$ rührt daher, dass die Taylorreihe am Punkt Null die selben Ableitungswerte besitzt wie f . Wir verdeutlichen dies mit einem Blick auf den Baustein: $\frac{x^k}{k!}$.

Der Baustein $\frac{x^k}{k!}$

Der Baustein $\frac{x^k}{k!}$, ein normiertes Monom, hat für $k \neq 0$ zwei bemerkenswerte Eigenschaften:

1) Der k -te Baustein hat als Ableitung stets den $(k-1)$ -ten Baustein, hier einmal am Beispiel demonstriert:

$$\left(\frac{x^4}{4!}\right)' = \left(\frac{\cancel{4} \cdot x^2}{\cancel{4} \cdot 3 \cdot 2 \cdot 1}\right) = \left(\frac{x^2}{3!}\right) \quad \left(\frac{x^3}{3!}\right)' = \left(\frac{\cancel{3} \cdot x^2}{\cancel{3} \cdot 2 \cdot 1}\right) = \left(\frac{x^2}{2!}\right) \quad \dots$$

2) Für fast alle Bausteine ergibt sich der Wert 0 beim Einsetzen von $x = 0$, mit Ausnahme von $\frac{x^0}{0!} = 1$.

		Ableitung	Ableitung	Ableitung	
Baustein	$\frac{x^0}{0!}$	$\frac{x^1}{1!}$	$\frac{x^2}{2!}$	$\frac{x^3}{3!}$...
Wert bei $x = 0$	1	0	0	0	...

Mit diesen Eigenschaften ergibt sich die folgende Eigenschaft der Taylorreihe:

Korollar 10.9

Sei $g(x) := \sum_{k=0}^{\infty} f^{(k)}(0) \frac{x^k}{k!}$ die Taylorreihe am Entwicklungspunkt $m = 0$. Falls $g(x)$ für alle x in einem Intervall $[-a, a]$ um die 0 konvergiert, gilt:

$$g^{(k)}(0) = f^{(k)}(0) \quad \text{für alle } k = 0, 1, 2, 3, \dots$$

Mit anderen Worten: Die Funktion $\mathbb{T}f$ hat an der Stelle $x = 0$ die selben Ableitungen wie f . Dies wird deutlich, wenn man die Ableitungen der Taylorreihe $\mathbb{T}f(x)$ berechnet und $x = 0$ einsetzt.

Die folgende Tabelle verdeutlicht, wie die beiden zentralen Eigenschaften der $\frac{x^k}{k!}$ -Bausteine die Ableitung

der Taylorreihe bestimmen:

Die Ableitungen der Taylorreihe						
$g(x)$	$=$	$f(0) \cdot 1$	$+ f'(0) \cdot \frac{x^1}{1!}$	$+ f''(0) \cdot \frac{x^2}{2!}$	$+ f'''(0) \cdot \frac{x^3}{3!}$	$+ \dots$
$g'(x)$	$=$	$f(0) \cdot 0$	$+ f'(0) \cdot 1$	$+ f''(0) \cdot \frac{x^1}{1!}$	$+ f'''(0) \cdot \frac{x^2}{2!}$	$+ \dots$
$g''(x)$	$=$	$f(0) \cdot 0$	$+ f'(0) \cdot 0$	$+ f''(0) \cdot 1$	$+ f'''(0) \cdot \frac{x^1}{1!}$	$+ \dots$
\vdots				\dots		
Einsetzen des Entwicklungspunktes $x = 0$						
$g(0)$	$=$	$f(0) \cdot 1$	$+ f'(0) \cdot 0$	$+ f''(0) \cdot 0$	$+ f'''(0) \cdot 0$	$+ \dots$
$g'(0)$	$=$	$f(0) \cdot 0$	$+ f'(0) \cdot 1$	$+ f''(0) \cdot 0$	$+ f'''(0) \cdot 0$	$+ \dots$
$g''(0)$	$=$	$f(0) \cdot 0$	$+ f'(0) \cdot 0$	$+ f''(0) \cdot 1$	$+ f'''(0) \cdot 0$	$+ \dots$
\vdots				\dots		

10.3 Wichtige Taylorreihen

Die folgenden Taylorreihen sind wichtig zu kennen:

Selbst wenn man in mühevoller Arbeit die Taylorreihe eines Polynoms berechnet, so erhält man am Ende doch nur wieder das Polynom selbst:

Lemma 10.10

Ist $f(x) := a_0 + a_1x + \dots + a_nx^n$ ein Polynom mit Koeffizienten $a_0, \dots, a_n \in \mathbb{R}$, so gilt für jeden Entwicklungspunkt m :

$$\mathbb{T}^m f(x) = a_0 + a_1x + \dots + a_nx^n$$

Für die Trigonometrischen Funktionen gilt:

Lemma 10.11

Für $f(x) := \cos(x)$ und $g(x) := \sin(x)$ gelten:

$$\begin{aligned} \cos(x) = \mathbb{T}f(x) &= \sum_{k=0}^{\infty} (-1)^k \cdot \frac{x^{2k}}{(2k)!} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} \pm \dots \\ \sin(x) = \mathbb{T}g(x) &= \sum_{k=0}^{\infty} (-1)^k \cdot \frac{x^{2k+1}}{(2k+1)!} = \frac{x^1}{1!} - \frac{x^3}{3!} \pm \dots \end{aligned}$$

Diese beiden Reihen konvergieren für alle $x \in \mathbb{R}$.

Die geometrische Reihe liefert nach Korollar 10.7 eine Taylorreihe für $\frac{1}{1-x}$, es gilt:

Lemma 10.12

Für $f(x) := \frac{1}{1-x}$ gilt

$$\mathbb{T}f(x) = \sum_{k=0}^{\infty} x^k = 1 + x + x^2 + \dots$$

Die Reihe konvergiert für alle $x \in (-1, 1)$ und divergiert sonst.

10.4 Das Taylorpolynom als Approximation einer Funktion

10.4.1 Die allgemeine Form des Taylorpolynoms:

Definition 10.13

Sei $f : \mathbb{R} \rightarrow \mathbb{R}$ eine am Punkt $a \in \mathbb{R}$ n -mal differenzierbare Funktion, dann heißt

$$\mathbb{T}_n^a f(x) := \sum_{k=0}^n f^{(k)}(a) \cdot \frac{(x-a)^k}{k!}$$

das *Taylorpolynom vom Grad n der Funktion f am Entwicklungspunkt a* .

Ist $a = 0$ so schreibt man für $\mathbb{T}_n^0 f$ kurz $\mathbb{T}_n f$. Es gilt dann:

$$\mathbb{T}_n f(x) := \sum_{k=0}^n f^{(k)}(0) \cdot \frac{x^k}{k!} = f(0) + f^{(1)}(0) \frac{x^1}{1!} + f^{(2)}(0) \frac{x^2}{2!} + \dots + f^{(n)}(0) \frac{x^n}{n!}$$

Das Taylorpolynom vom Grad n ist eine "endliche Version" der Taylorreihe. Dies hat den entscheidenden Vorteil, dass das Taylorpolynom –als endliche Summe– für alle Werte von x einen endlichen Wert ergibt. Die *Taylorreihe* ist –auf dem Konvergenzintervall– identisch mit der ursprünglichen Funktion (s. Satz 10.6). Im Gegensatz dazu ist das *Taylorpolynom* allerdings "nur noch" eine Approximation der ursprünglichen Funktion. Je höher der Grad des Taylorpolynoms, um so besser stimmt das Taylorpolynom rund um den Entwicklungspunkt mit der gegebenen Funktion überein. Dies sieht man zum Beispiel in Abbildung 10.1, dort werden die ersten vier Taylorpolynome der Eulerfunktion e^x mit der Funktion selbst verglichen.

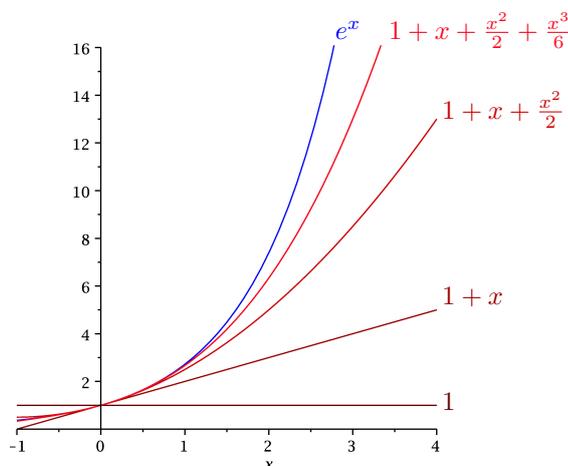
Beispiel 10.14 (Anwendungsbeispiel)

Will man den Wert der Eulerzahl e näherungsweise berechnen, so kann man das Taylorpolynom der Funktion e^x an der Stelle $a = 0$ entwickeln und den Wert $x = 1$ einsetzen. Wir tun dies mittels des Taylorpolynoms vom Grad 3:

$$\begin{aligned} \mathbb{T}_3 f(x) &= 1 + \frac{x^1}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} \\ &= 1 + x + \frac{x^2}{2} + \frac{x^3}{6} \\ \mathbb{T}_3 f(1) &= 1 + 1 + \frac{1}{2} + \frac{1}{6} = \frac{8}{3} = 2,\bar{6} \end{aligned}$$

10.4.2 Restgliedabschätzung

Wie genau ist die Approximation von e aus Beispiel 10.14? Die Eulerzahl hat ungefähr den Wert 2,71828, den Fehler zum berechneten Wert $2,\bar{6}$ kann man jedoch auch ohne dies zu wissen mit dem Satz von Taylor

Abbildung 10.1: Die ersten vier Taylorpolynome der Eulerfunktion e^x .

abschätzen:

Satz 10.15 (Satz von Taylor)

Die Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ sei auf dem abgeschlossenen Intervall $[0, \bar{x}]$ n -mal stetig differenzierbar und die Ableitung $f^{(n+1)}(x)$ sei auf dem offenen Intervall $(0, \bar{x})$ definiert.

Dann gibt es eine Zahl $\xi \in (0, \bar{x})$ so dass gilt:

$$f(\bar{x}) = \underbrace{f(0) + f^{(1)}(0)\frac{\bar{x}^1}{1!} + f^{(2)}(0)\frac{\bar{x}^2}{2!} + \cdots + f^{(n)}(0)\frac{\bar{x}^n}{n!}}_{\text{das Taylorpolynom vom Grad } n} + \underbrace{f^{(n+1)}(\xi)\frac{\bar{x}^{n+1}}{(n+1)!}}_{\text{Fehlerterm}} \quad (10.1)$$

Definition 10.16

Der Term $R_n(x, \xi) := f^{(n+1)}(\xi)\frac{x^{n+1}}{(n+1)!}$ in Gleichung (10.1) heißt *Lagrangesches Restglied*^a.

^a Es gibt noch weitere mögliche Formen für das Restglied, die hier nicht erwähnt werden.

Wenn man die Größe des Lagrangeschen Restgliedes abschätzen kann, so kann man über das Umstellen der Gleichung (10.1) den Fehler abschätzen, den man macht, wenn man statt $f(\bar{x})$ den Wert des Taylorpolynoms $T_n(\bar{x})$ berechnet:

Beispiel 10.17 (Fortsetzung Anwendungsbeispiel 10.14)

In Beispiel 10.14 wird der Wert der Eulerzahl e näherungsweise berechnet:

$$T_3 f(1) = 1 + 1 + \frac{1}{2} + \frac{1}{6} = \frac{8}{3} = 2,\bar{6}$$

Nach Satz 10.1 gibt es nun ein $\xi \in (0, 1)$ so dass gilt:

$$f(1) - T_3 f(1) = f^{(4)}(\xi) \frac{1}{4!} = \frac{e^\xi}{4!}$$

Mit dem zusätzlichen Wissen, dass $e \leq 3$ gilt, und dass $e^\xi \leq e^1$ gilt folgt:

$$f(1) - T_3 f(1) = \frac{e^\xi}{4!} \leq \frac{3}{4 \cdot 3 \cdot 2 \cdot 1} = \frac{1}{8}$$

Mit dieser Rechnung können wir zeigen, dass der errechnete Wert $2,\bar{6}$ höchstens um $1/8 = 0,125$ kleiner ist als die Eulerzahl e .

Dies ist natürlich nur eine obere Schranke mit gehöriger "Sicherheitsreserve", der tatsächliche Fehler ist deutlich kleiner. Zum Vergleich: Der tatsächliche Fehler liegt etwa bei $2,71828 - 2,\bar{6} \sim 0.051615161$.

Kapitel 11

Komplexe Zahlen

11.1 Warum imaginäre Zahlen so heißen wie sie heißen

Dieses Kapitel widmet sich den *komplexen Zahlen*, einer Zahlenmenge, die zusammen mit Addition und Multiplikation einen *Zahlkörper* bildet. Die folgenden Zahlenmengen sollten bereits bekannt sein:

- die *natürlichen* Zahlen $\mathbb{N} = \{1, 2, 3, 4, \dots\}$
- die *ganzen* Zahlen $\mathbb{Z} = \{0, 1, -1, 2, -2, 3, -3, \dots\}$
- die *rationalen* Zahlen $\mathbb{Q} = \left\{ \frac{p}{q} : p, q \in \mathbb{Z}, q \neq 0 \right\}$
- die *reellen* Zahlen $\mathbb{R} = \left\{ a + 0, d_1 d_2 d_3 \dots : a \in \mathbb{Z} \text{ und } \underbrace{(d_n)_{n \in \mathbb{N}} \text{ mit } d_k \in \{0, 1, \dots, 9\}}_{\text{Nachkomma-Ziffern}} \right\}$

Diese Mengen von Zahlen haben Entsprechungen in der echten Welt:

- Die natürlichen Zahlen entsprechen der Anzahl zählbarer Gegenstände.
- Brüche (rationale Zahlen) stehen für “Anteile eines geteilten Ganzen” (Pizza).
- Auch reellen Zahlen treten im Alltags auf: zum Beispiel ist $\sqrt{2}$ die Länge der Diagonale in einem 1-mal-1 großen Quadrat (Fliese). Auch das Verhältnis der Seiten eines Blatt Papiers im DinA4-Format ist 1 zu $\sqrt{2}$.

Alle diese Zahlen bis hin zu den reellen Zahlen lassen sich also in der Umwelt anschaulich darstellen. Nicht so die imaginäre Einheit bzw. die komplexen Zahlen: Sie entsprechen den nicht ganz so greifbaren, abstrakten Lösungen von Polynomgleichungen. Man kann sich diese Zahlen zwar veranschaulichen, also auf ein Blatt Papier zeichnen, man findet Sie jedoch in der Natur nur hinter den Kulissen, im Wirken von physikalischen Gesetzen.

11.2 Definition

Die Klasse der reellen Zahlen ist sehr groß, trotzdem lernt man in der Schule, dass $a \cdot x^2 + b \cdot x + c = 0$ nur dann lösbar ist, falls ihre Diskriminante $\Delta = b^2 - 4ac$ nicht-negativ ist, da aus einer negativen Zahl keine Wurzel gezogen werden kann.

Beispielsweise hat die Gleichung $x^2 = -1$ keine *reelle* Lösung, was etwas unbefriedigend ist. Lässt sich nicht vielleicht doch irgendwie eine Zahl $\sqrt{-1}$ definieren? Die Antwort ist *jein*. Tatsächlich gibt es keine reelle Zahl $x \in \mathbb{R}$ mit $x^2 = -1$, hier wurde Ihnen also nichts vorenthalten, trotzdem konstruieren wir nun eine Lösung, indem wir uns eine entsprechende Zahl definieren.

11.2.1 Die imaginäre Einheit

Definition 11.1

Wir definieren eine Zahl i (genannt die imaginäre Einheit) so, dass $i^2 = -1$ gilt.

Die Zahl i ist in sofern imaginär, als dass sie keine *reelle* Zahl ist, und keine direkt greifbare Entsprechung in der “echten” Welt besitzt. Mit dieser neuen, zusätzlichen Zahl können nun die Lösungen für $x^2 = -1$ angeben, dies sind nämlich i und $-i$.

Aus der Regel $i^2 = -1$ lässt sich folgern, dass sich Potenzen von i wieder als ± 1 oder $\pm i$ schreiben lassen:

k	0	1	2	3	4	
i^k	i^0	i^1	i^2	i^3	i^4	...
Wert	1	i	-1	$-i$	1	...

Definition 11.2

Eine Komplexe Zahl z hat die Form $z = a + i \cdot b$ dabei sind $a, b \in \mathbb{R}$ und i ist die *imaginäre Einheit*.

Für $z = a + i \cdot b$ ist $\operatorname{Re}(z) := a$ der *Realteil* von z und $\operatorname{Im}(z) := b$ der *Imaginärteil* von z .

Die Menge aller Komplexen Zahlen kürzt man ab mit $\mathbb{C} := \{a + i \cdot b : a, b \in \mathbb{R}\}$



Typische Fehler

Der Imaginärteil von $z = a + i \cdot b$ ist die **reelle** Zahl b und **nicht** $i \cdot b$.

Mit i können wir nun Lösungen für *alle* Polynomgleichungen angeben – nicht nur eine Lösung für $x^2 = -1$.

Satz 11.3

Für das Polynom $x^2 + px + q$ mit $p, q \in \mathbb{R}$ sei $(p/2)^2 - q < 0$.

Dann hat die Gleichung $x^2 + px + q = 0$ die komplexen Lösungen

$$z_1 := -\frac{p}{2} + i \cdot \sqrt{\left(\frac{p}{2}\right)^2 - q} \quad \text{und} \quad z_2 = \bar{z}_1 := -\frac{p}{2} - i \cdot \sqrt{\left(\frac{p}{2}\right)^2 - q}$$

Den Beweis führt man durch einfaches Einsetzen und Nachrechnen, unter Beachtung, dass hier gilt:

$$\left(\frac{p}{2}\right)^2 - q < 0 \quad \Rightarrow \quad \left|\left(\frac{p}{2}\right)^2 - q\right| = q - \left(\frac{p}{2}\right)^2$$

Beispiel 11.4

Gesucht sind die Lösungen von $x^2 + 2x + 10 = 0$ mit der p - q -Formel ergibt sich:

$$x_{1,2} = -1 \pm \sqrt{1 - 10} = -1 \pm \sqrt{(-1) \cdot 9} = -1 \pm 3 \cdot \sqrt{-1}.$$

In reellen Zahlen gibt es hier keine Lösung, in komplexen Zahlen lauten die Lösungen $x_{1/2} = -1 \pm 3 \cdot i$.



i ist nicht “ $\sqrt{-1}$ ”

Obwohl es verführerisch azussieht, ist es **nie richtig** “ $i = \sqrt{-1}$ ” zu schreiben. In Beispiel 11.4 *sieht es zwar so aus* als hätte man “ $\sqrt{-1}$ ” durch i ersetzt, **trotzdem gilt nicht** “ $i = \sqrt{-1}$ ”.

Warum nicht $\sqrt{-1}$

Die Wurzelfunktion \sqrt{x} liefert für $x \in \mathbb{R}$ mit $x \geq 0$ das *nicht-negative* y , für das $y^2 = x$ gilt. Wegen der Einschränkung “*nicht-negativ*” funktioniert die Wurzelfunktion nicht bei -1 :

- Die Lösungen für $y^2 = 16$ sind zum Beispiel 4 und -4 ,
die Wurzelfunktion liefert aber nur die positive Lösung $\sqrt{16} = 4$.
- Die Lösungen für $y^2 = -1$ sind i und $-i$,
aber hier ist (und bleibt!) unklar, wer “die positive” Lösung ist.

Sowohl i als auch $-i$ sind *weder* positiv *noch* negativ!

Auch das Argument “ i hat kein Vorzeichen” zieht hier nicht, denn: Bei der Definition von \mathbb{C} hätte man (statt die Zahl i zu wählen) ebenso gut die komplexen Zahlen über $j := (-i)$ definieren können! Dann hätte die Zahl j (also eigentlich $-i$) “kein Vorzeichen”.

Ein anderes, etwas schwierigeres Argument für $i \neq \sqrt{-1}$ ist das Folgende:

Wäre $\sqrt{-1}$ eine echte Zahl, so dürfte man also aus -1 die Wurzel ziehen, d.h. für die entstehende Zahl $\sqrt{-1}$ müsste dann die übliche Wurzelrechenregel $\sqrt{a \cdot b} = \sqrt{a} \cdot \sqrt{b}$ gelten.

(\star) Annahme: Es gelte $i = \sqrt{-1}$.

($\star\star$) Dann gilt die Wurzelrechenregel $\sqrt{a \cdot b} = \sqrt{a} \cdot \sqrt{b}$ auch für $a = b = -1$.

$$\begin{array}{ll} \text{Es gilt dann also:} & 1) \quad 1 = \sqrt{(-1) \cdot (-1)} \\ & 2) \quad \overset{\star\star}{\Leftrightarrow} 1 = \sqrt{(-1)} \cdot \sqrt{(-1)} \\ & 3) \quad \overset{\star}{\Leftrightarrow} 1 = i \cdot i \\ & 4) \quad \Leftrightarrow 1 = -1 \end{array}$$

Während hier 1) unstrittig wahr ist, ist 4) unstrittig falsch, d.h. die Annahme $i = \sqrt{-1}$ führt zu einem Widerspruch (und zwar durch ($\star\star$), das direkt aus $i = \sqrt{-1}$ folgt).

11.2.2 Rechnen mit komplexen Zahlen

Das Rechnen mit komplexen Zahlen folgt im Wesentlichen den Rechenregeln für reelle Zahlen, d.h. auch hier gilt beispielsweise die Regel “Punkt- vor Strichrechnung”. Für das Multiplizieren von komplexen Zahlen benötigt man zum einen ein gutes Verständnis der Binomischen Formeln, und zum anderen die einfache Einsicht, dass sich Potenzen von i wieder als ± 1 oder $\pm i$ schreiben lassen:

Addition in \mathbb{C}		
	$a +$	$b \cdot i$
$+$	$c +$	$d \cdot i$
<hr/>		
	$= (a + c) +$	$(b + d) \cdot i$

Addition von Polynomen		
	$a +$	$b \cdot x$
$+$	$c +$	$d \cdot x$
<hr/>		
	$= (a + c) +$	$(b + d) \cdot x$

Multiplikation in \mathbb{C}		
	$(a + b \cdot i) \cdot$	$(c + d \cdot i)$
<hr/>		
$=$	$a \cdot c + (a \cdot d + c \cdot b) i +$	$b \cdot d \cdot \underbrace{i^2}_{=-1}$
$=$	$a \cdot c - b \cdot d +$	$(a \cdot d + c \cdot b) i$

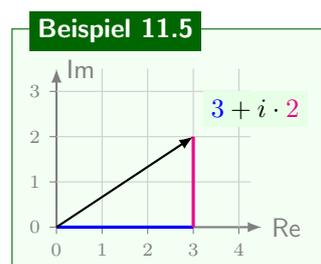
Multiplikation von Polynomen		
	$(a + b \cdot x) \cdot$	$(c + d \cdot x)$
<hr/>		
$=$	$a \cdot c + (a \cdot d + c \cdot b) x +$	$b \cdot d \cdot x^2$

11.3 Komplexe Zahlen als kartesische Vektoren

Die Menge der komplexen Zahlen \mathbb{C} lässt sich auffassen als ein zwei-dimensionaler reeller Vektorraum.

Jeder Zahl $z = a + i \cdot b$ kann man einen Vektor $\begin{pmatrix} \operatorname{Re}(z) \\ \operatorname{Im}(z) \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix}$ zuordnen. Im zugehörigen \mathbb{R}^2 bezeichnet man die x_1 -Achse mit Re und die x_2 -Achse mit Im . Den entstehenden besonderen \mathbb{R}^2 nennt man die “komplexe Zahlenebene” oder auch die “Gaussche Zahlenebene”.

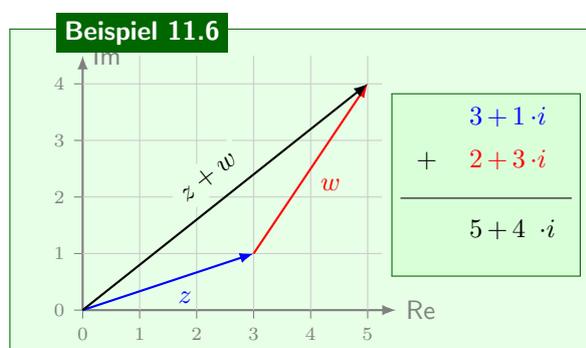
Beide Koordinaten-Achsen (auch die Im -Achse!) werden Zahlen aus \mathbb{R} beschriftet. So liegt die Zahl $z := 2i$ als Vektor $\begin{pmatrix} 0 \\ 2 \end{pmatrix}$ auf der Im -Achse, und zwar beim Achsenabschnitt “2” (d.h. bei $\operatorname{Im}(2i) = 2$). Der Wert der komplexen Zahl z ist aber weiterhin $2i$ und damit komplex.



Die Addition von zwei komplexen Zahlen durch *getrenntes* Addieren von jeweils zwei Realteilen und zwei Imaginärteilen läuft ganz analog zur Addition von Vektoren im \mathbb{R}^2 ab. Wie dort so entspricht auch in \mathbb{C} die Addition dem *Aneinanderhängen* der Zahlen bzw. Vektoren:

Addition in \mathbb{C}	
	$a + b \cdot i$
+	$c + d \cdot i$
<hr/>	
=	$(a + c) + (b + d) \cdot i$

Addition von Vektoren	
$\begin{pmatrix} a \\ b \end{pmatrix}$	+ $\begin{pmatrix} c \\ d \end{pmatrix}$
<hr/>	
	$\begin{pmatrix} a + c \\ b + d \end{pmatrix}$



Jede Zahl $z \in \mathbb{C}$ hat als Vektor eine Länge und einen Winkel (zur Re -Achse):

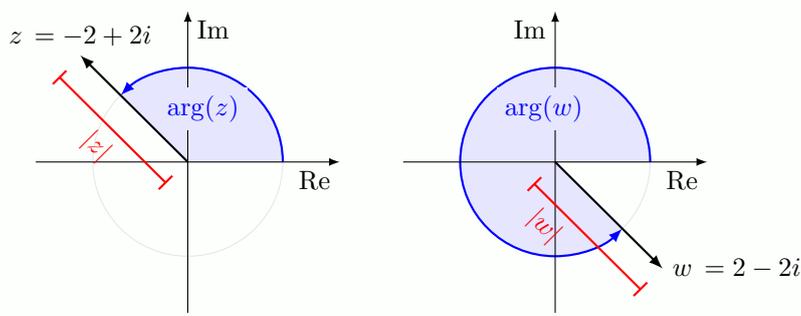
Definition 11.7

Der *Betrag* einer komplexen Zahl $z = a + b \cdot i$ ist die Länge des Vektors $\begin{pmatrix} a \\ b \end{pmatrix}$, es gilt $|z| := \sqrt{a^2 + b^2}$.

Das *Argument* $\arg(z) \in [0, 2\pi)$ einer komplexen Zahl $z \in \mathbb{C}$ ist der Winkel, der zwischen der Re -Achse und der Strecke von 0 nach z eingeschlossen wird.

Beispiel 11.8

- Für die Zahl $z = -2 + 2i$ ist $|z| = \sqrt{2^2 + 2^2} = 2 \cdot \sqrt{2}$ und $\arg(z) = \frac{\pi}{4}$.
- Für die Zahl $w = 2 - 2i$ ist $|z| = \sqrt{2^2 + 2^2} = 2 \cdot \sqrt{2}$ und $\arg(z) = 2\pi - \frac{\pi}{4} = \frac{7}{4}\pi$.

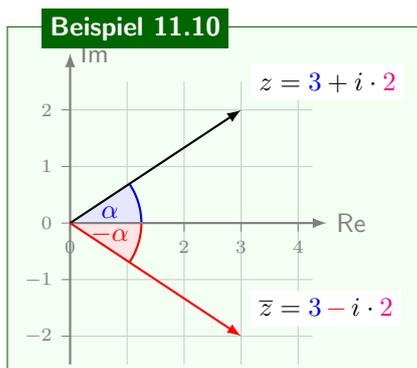
**Winkel vs. Argument**

Man beachte, dass nach obiger Definition für $\arg(z)$ stets $0 \leq \arg(z) < 2\pi$ gilt. Gibt man jedoch Winkel an, so sind zwei Winkel äquivalent, wenn sie sich nur um ein Vielfaches von 2π unterscheiden.

11.4 Konjugiert komplexe Zahlen und Division

Definition 11.9

Für eine komplexe Zahl $z = a + b \cdot i$, bezeichnet man mit $\bar{z} = a - b \cdot i$ die zu z *konjugiert komplexe Zahl*, die aus z durch Spiegelung an der reellen Achse hervorgeht.



Offensichtlich haben z und \bar{z} die selbe Länge, aber entgegengesetzte Winkel. Es gelten:

$$|\bar{z}| = |z| \quad \text{und} \quad \arg(\bar{z}) = 2\pi - \arg(z).$$

Bei Summen und Produkten kann das Konjugieren vor oder nach dem Addieren bzw. Multiplizieren geschehen. Genauer gelten für $z, w \in \mathbb{C}$:

$$\overline{(z + w)} = \bar{z} + \bar{w} \quad \text{und} \quad \overline{(z \cdot w)} = \bar{z} \cdot \bar{w}$$

11.4.1 Verwendung

Die zu z konjugierte Zahl wird beim Berechnen von reellen Zahlen aus z genutzt. Einer der dabei genutzten Umstände ist, dass für $z = a + b \cdot i$ gilt:

$$z \cdot \bar{z} = (a + b \cdot i) \cdot (a - b \cdot i) = a^2 - (i \cdot b)^2 = a^2 + b^2$$

Dies kann beim Teilen durch komplexe Zahlen verwendet werden: Um auf einen reellen Nenner zu kommen, erweitert man einen komplexwertigen Bruch w/z mit \bar{z} , dem komplex-konjugierten des Nenners.

Beispiel 11.11

Für $z = 4 + 3i$ und $w = 1 + i$ gilt:

$$\frac{w}{z} = \frac{w \cdot \bar{z}}{z \cdot \bar{z}} = \frac{(1+i) \cdot (4-3i)}{(4+3i) \cdot (4-3i)} = \frac{7+i}{4^2+3^2} = \frac{7}{25} + \frac{1}{25} \cdot i$$

Lemma 11.12

Für $z = a + b \cdot i$ und $w = c + d \cdot i$ gelten:

$$\frac{w}{z} = \frac{w \cdot \bar{z}}{z \cdot \bar{z}} = \frac{(c+d \cdot i) \cdot (a-b \cdot i)}{(a+b \cdot i) \cdot (a-b \cdot i)} = \frac{ca+db}{a^2+b^2} + \frac{ad-bc}{a^2+b^2} \cdot i$$

Das Ergebnis der Division $\frac{w}{z}$ ist also stets wieder eine komplexe Zahl.

Die zu z konjugierte Zahl wird benötigt, um aus z die Größen $\operatorname{Re}(z)$, $\operatorname{Im}(z)$ und $|z|$ zu berechnen.

Lemma 11.13 (Berechnungen mittels \bar{z})

Für $z = a + ib$ gelten:

$$\operatorname{Re}(z) = \frac{z + \bar{z}}{2} = \frac{(a+i \cdot b) + (a-i \cdot b)}{2} = \frac{2a}{2} = a$$

$$\operatorname{Im}(z) = \frac{z - \bar{z}}{2i} = \frac{(a+i \cdot b) - (a-i \cdot b)}{2i} = \frac{2bi}{2i} = b$$

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

11.5 Polardarstellung komplexer Zahlen

Anstatt Real- und Imaginärteil einer komplexen Zahl z zu kennen, reicht es auch ihren Betrag r , d.h. den Abstand vom Ursprung, und den Winkel α , den die Strecke vom Ursprung zu z mit der reellen Achse einschließt, zu kennen.

Definition 11.14

Die *Polarkoordinaten* einer komplexen Zahl $z \in \mathbb{C} \setminus \{0\}$ bestehen aus dem Paar $(r, \alpha) \in \mathbb{R}^2$ mit

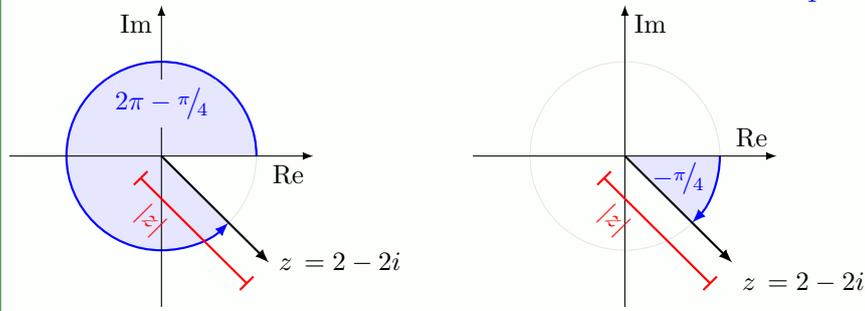
- Länge $r \geq 0$, d.h. $r := |z|$ und
- Winkel $\alpha \in \mathbb{R}$ so dass α äquivalent zu $\arg(z)$ ist, d.h. $\alpha = \arg(z) + k \cdot 2\pi$ mit $k \in \mathbb{Z}$.

Für $z = 0$ sind die Polarkoordinaten von der Form $(0, \alpha)$, dabei ist der Winkel α beliebig.

Beispiel 11.15

Für die Zahl $z = 2 - 2i$ ist $|z| = 2 \cdot \sqrt{2}$ und $\arg(z) = 2\pi - \frac{\pi}{4}$.

Als Polarkoordinaten kommen beispielsweise in Frage: $(2\sqrt{2}, 2\pi - \frac{\pi}{4})$ aber auch $(2\sqrt{2}, -\frac{\pi}{4})$

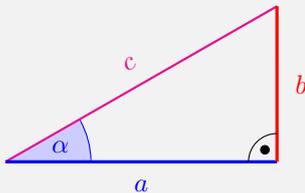
**Lemma 11.16 (Umrechnen von Polarkoordinaten in kartesischen Koordinaten)**

Liegt z in Polarkoordinaten $z \hat{=} (r, \alpha)$ vor, so gilt $z = r \cdot \cos(\alpha) + r \cdot \sin(\alpha) \cdot i$, d.h.

$$\operatorname{Re}(z) = r \cdot \cos(\alpha)$$

$$\operatorname{Im}(z) = r \cdot \sin(\alpha)$$

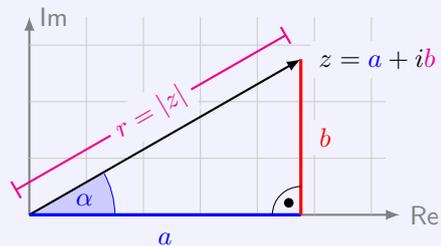
Schulwissen: rechtwinkliges Dreieck



$$\cos(\alpha) = \frac{\text{Ankathete}}{\text{Hypotenuse}} = \frac{a}{c} \quad \Rightarrow \quad a = c \cdot \cos(\alpha)$$

$$\sin(\alpha) = \frac{\text{Gegenkathete}}{\text{Hypotenuse}} = \frac{b}{c} \quad \Rightarrow \quad b = c \cdot \sin(\alpha)$$

Komplexe Zahlen



$$\operatorname{Re}(z) = a = r \cdot \cos(\alpha)$$

$$\operatorname{Im}(z) = b = r \cdot \sin(\alpha)$$

Lemma 11.17 (Umrechnen von kartesischen Koordinaten in Polarkoordinaten)

Sind die kartesischen Koordinaten $z = a + b \cdot i \neq 0$ bekannt, so können die zugehörigen Polarkoordinaten (r, α) wie folgt berechnet werden:

$$r = |z| = \sqrt{a^2 + b^2} \quad \alpha = \begin{cases} \arccos(a/r) & \text{falls } b \geq 0 \\ -\arccos(a/r) & \text{falls } b < 0 \end{cases}$$

dabei ist $\arccos(x)$ die Umkehrfunktion des Cosinus (auf dem Taschenrechner \cos^{-1}).

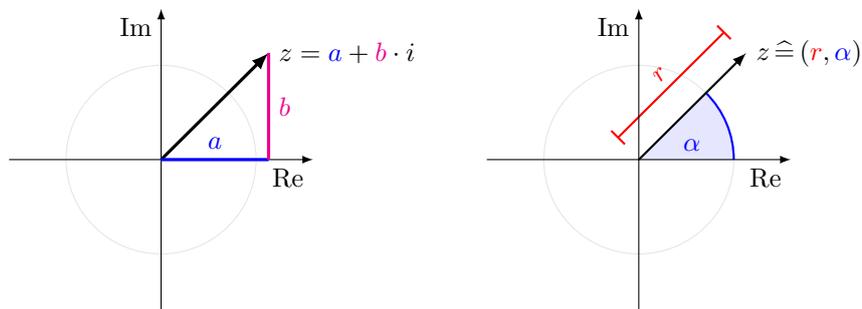


Abbildung 11.1: Kartesische vs. Polar-Darstellung einer komplexen Zahl

11.6 Multiplizieren und Dividieren in Polarkoordinaten

Für zwei komplexe Zahlen ist die Multiplikation in Polarkoordinatendarstellung erheblich einfacher als in kartesischen koordinaten. Kurz gesagt gilt:

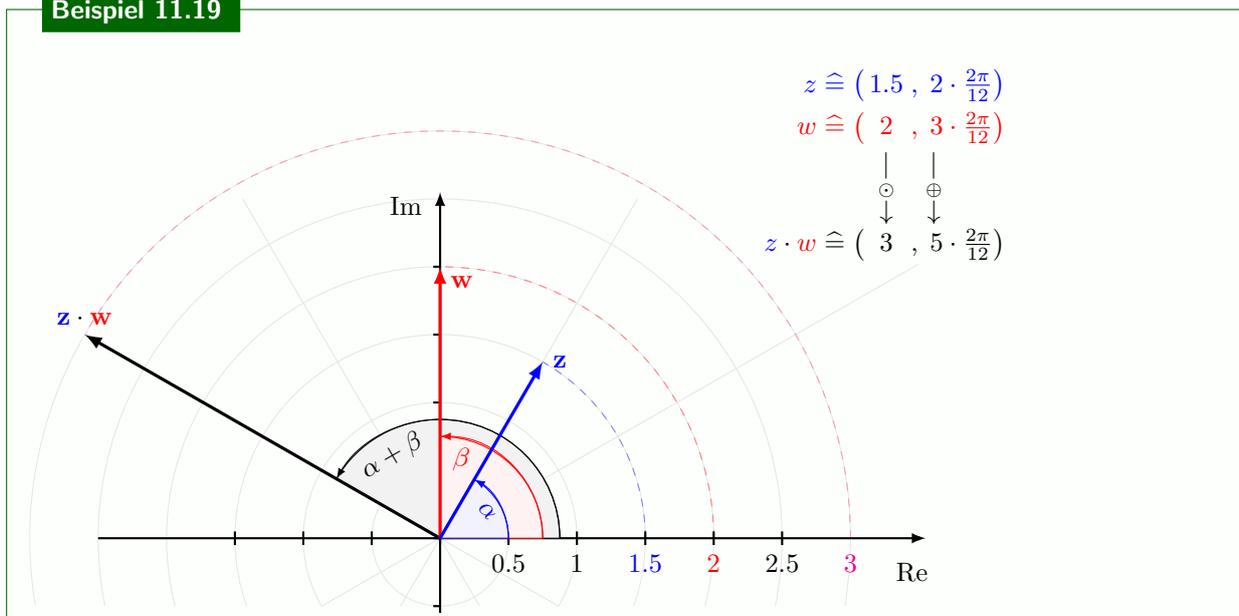
Längen werden multipliziert, Winkel werden addiert.

Lemma 11.18

Es seien $z, w \in \mathbb{C}$ mit $z \hat{=} (r, \alpha)$ und $w \hat{=} (R, \beta)$. Dann gilt für das Produkt und den Quotienten:

$$z \cdot w \hat{=} (r \cdot R, \alpha + \beta) \quad \text{und} \quad \frac{w}{z} \hat{=} \left(\frac{R}{r}, \beta - \alpha \right)$$

Beispiel 11.19

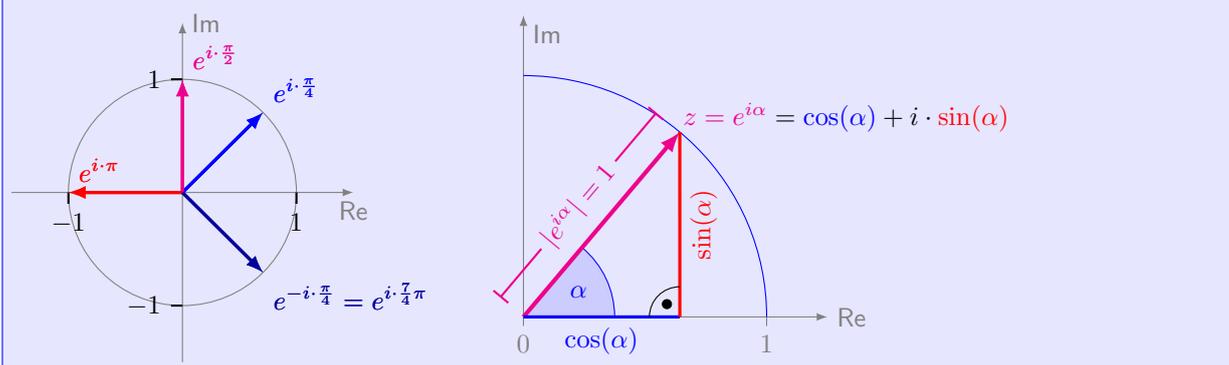


Den Beweis werden wir mit der Eulerschen Darstellung einer komplexen Zahl führen, dafür benötigen wir das Verhalten der Funktion e^x beim Einsetzen von rein komplexen Zahlen $i \cdot \alpha$:

Lemma 11.20

Für $\alpha \in \mathbb{R}$ gilt stets $e^{i \cdot \alpha} = \cos(\alpha) + i \cdot \sin(\alpha)$. Insbesondere gelten: $e^{i \cdot \pi} = -1$ und $e^{i \cdot \frac{\pi}{2}} = i$

Die Zahl $e^{i\alpha}$ ist die Zahl auf dem Einheitskreis mit Winkel α :



Beweis: Die Taylorreihenentwicklungen von e^x , $\cos(x)$ und $\sin(x)$ lauten:

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \frac{x^5}{5!} + \dots$$

$$\cos(x) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} \pm \dots$$

$$\sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} \pm \dots$$

Setzt man in e^x die Zahl $i \cdot \alpha$ ein so erhält man (wegen $i^1 = i$, $i^2 = -1$, $i^3 = -i$, $i^4 = 1$, $i^5 = i, \dots$):

$$e^{i\alpha} = 1 + i \cdot \alpha + i^2 \cdot \frac{\alpha^2}{2!} + i^3 \cdot \frac{\alpha^3}{3!} + i^4 \cdot \frac{\alpha^4}{4!} + i^5 \cdot \frac{\alpha^5}{5!} + \dots$$

$$= 1 + i \cdot \alpha - \frac{\alpha^2}{2!} - i \cdot \frac{\alpha^3}{3!} + \frac{\alpha^4}{4!} + i \cdot \frac{\alpha^5}{5!} \pm \dots \left. \vphantom{e^{i\alpha}} \right\} = \left\{ \begin{array}{l} \cos(\alpha) \\ + i \cdot \sin(\alpha) \end{array} \right.$$

Es gilt also $e^{i \cdot \alpha} = \cos(\alpha) + i \cdot \sin(\alpha)$ □

Satz 11.21 (Eulerdarstellung)

Für eine komplexe Zahl $z \in \mathbb{C}$ mit Polarkoordinaten $z \hat{=} (r, \alpha)$ gilt: $z = r \cdot e^{i \cdot \alpha}$

Umgekehrt gilt $r \cdot e^{i \cdot \alpha} \hat{=} (r, \alpha)$, d.h. $r \cdot e^{i \cdot \alpha}$ hat die Länge r und den Winkel α .

Beweis: Die Zahl $z \hat{=} (r, \alpha)$ lässt sich nach Lemma 11.16 schreiben als

$$z = r \cdot \cos(\alpha) + i \cdot r \cdot \sin(\alpha) = r \cdot \underbrace{(\cos(\alpha) + i \cdot \sin(\alpha))}_{e^{i\alpha}},$$

d.h. es gilt $z = r \cdot e^{i\alpha}$.

Für die Zahl $r \cdot e^{i\alpha}$ gilt:

$$|r \cdot e^{i\alpha}| = r \cdot |\cos(\alpha) + i \sin(\alpha)| = r \cdot \underbrace{\sqrt{\cos(\alpha)^2 + \sin(\alpha)^2}}_{=1}$$

Darüberhinaus hat $e^{i\alpha}$ nach Lemma 11.20 den Winkel α , und entsprechend hat $r \cdot e^{i\alpha}$ den Winkel α . \square

Jetzt können wir Lemma 11.18 beweisen: es lautet:

Lemma 11.18 (Wiederholung)

Es seien $z, w \in \mathbb{C}$ mit $z \hat{=} (r, \alpha)$ und $w \hat{=} (R, \beta)$. Dann gilt für das Produkt und den Quotienten:

$$z \cdot w \hat{=} (r \cdot R, \alpha + \beta) \quad \text{und} \quad \frac{w}{z} \hat{=} \left(\frac{R}{r}, \beta - \alpha \right)$$

Beweis: Es gelten $z = r \cdot e^{i\alpha}$ und $w = R \cdot e^{i\beta}$ und es folgt:

$$z \cdot w = r \cdot e^{i\alpha} \cdot R \cdot e^{i\beta} = r \cdot R \cdot e^{i(\alpha+\beta)}$$

Die letzte Zahl hat offensichtlich den Betrag (Länge) $r \cdot R$ und den Winkel $\alpha + \beta$. Das heißt es gilt $z \cdot w \hat{=} (r \cdot R, \alpha + \beta)$

Für den Quotienten gilt:

$$\frac{w}{z} = \frac{R \cdot e^{i\beta}}{r \cdot e^{i\alpha}} = \frac{R \cdot e^{i\beta}}{r} \cdot e^{-i\alpha} = \frac{R}{r} \cdot e^{i(\beta-\alpha)}$$

Die letzte Zahl hat offensichtlich den Betrag (Länge) R/r und den Winkel $\beta - \alpha$. Das heißt es gilt $w/z \hat{=} (R/r, \beta - \alpha)$ \square

11.6.1 Einheitswurzeln

Definition 11.22 (Einheitswurzeln)

Es sei $n \in \mathbb{N} \setminus \{0\}$. Die Lösungen von $z^n = 1$ mit $z \in \mathbb{C}$ heißen die n -ten Einheitswurzeln.

Satz 11.23

Für $n \in \mathbb{N} \setminus \{0\}$ sind die Lösungen von $z^n = 1$ die folgenden Punkte auf dem Einheitskreis:

$$\mathbb{L} := \left\{ e^{i\alpha} \in \mathbb{C} : \alpha = k \cdot \frac{2\pi}{n} \text{ mit } k = 0, 1, 2, \dots, n-1 \right\}$$

Eine Lösung ist (bei $k = 0$) die Zahl $z = 1$, alle Lösungen liegen auf einem symmetrischen n -Eck.

Beispiel 11.24

Für $n = 3$ sind die Lösungen von $z^3 = 1$ die **drei** komplexen Zahlen

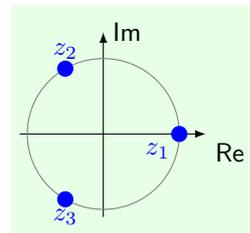
$$z_1 = e^{i \cdot 0 \cdot \frac{2\pi}{3}} \quad z_2 = e^{i \cdot 1 \cdot \frac{2\pi}{3}} \quad z_3 = e^{i \cdot 2 \cdot \frac{2\pi}{3}}$$

Diese Zahlen liegen *gleichmäßig* über den Einheitskreis verteilt mit Winkeln

$$0 \hat{=} 0^\circ \quad 1 \cdot \frac{2\pi}{3} \hat{=} 120^\circ \quad \text{und} \quad 2 \cdot \frac{2\pi}{3} \hat{=} 240^\circ$$

Mit dem Taschenrechner erhält man per “ $z = r \cos(\alpha) + i \cdot r \sin(\alpha)$ ” die kartesischen Koordinaten:

$$z_1 = 1 \quad z_2 = -\frac{1}{2} + i \cdot \frac{\sqrt{3}}{2} \quad z_3 = -\frac{1}{2} - i \cdot \frac{\sqrt{3}}{2}$$



Satz 11.23 lässt sich mit den Multiplikationsregeln (Längen multiplizieren, Winkel addieren) für Polarkoordinaten veranschaulichen. Wir betrachten das Beispiel $z^5 = 1$:

Die Zahl $z = r \cdot e^{i\alpha}$ hat Polarkoordinaten (r, α) , d.h. $z^5 = r^5 \cdot e^{i5\alpha}$ hat die Polarkoordinaten $(r^5, 5 \cdot \alpha)$.

Die Zahl 1 hat Polarkoordinaten $(1, \beta)$ mit möglichen Winkeln $\beta \in \{\dots, -2\pi, 0, 2\pi, 2 \cdot 2\pi, \dots\}$

Damit $z^5 = 1$ gelten kann, muss also gelten: $r^5 = 1$ und $5 \cdot \alpha = k \cdot 2\pi$ mit $k \in \mathbb{Z}$.

- **Radius:** Aus $r^5 = 1$ folgt sofort $r = 1$, denn es gelten $r \in \mathbb{R}$ und $r \geq 0$ (r ist eine Längenangabe), und in $\mathbb{R}_{\geq 0}$ gibt es nur eine Lösung von $x^5 = 1$, nämlich $x = 1$
- **Winkel:** Wegen den Multiplikationsregeln (Winkel addieren!) folgt: $z \hat{=} (1, \alpha) \Rightarrow z^5 \hat{=} (1, 5\alpha)$
Die Zahl 1 hat Polarkoordinaten $(1, \beta)$ mit möglichen Winkeln $\beta \in \{\dots, -2\pi, 0, 2\pi, 2 \cdot 2\pi, \dots\}$.
Es folgt also

$$\dots \quad 5\alpha = -1 \cdot 2\pi \quad \text{oder} \quad 5\alpha = 0 \cdot 2\pi \quad \text{oder} \quad 5\alpha = 1 \cdot 2\pi \quad \text{oder} \quad 5\alpha = 2 \cdot 2\pi \quad \text{oder} \quad \dots$$

d.h. es gilt $5\alpha = k \cdot 2\pi$ mit $k \in \mathbb{Z}$ beziehungsweise $\alpha = k \cdot \frac{2\pi}{5}$ mit $k \in \mathbb{Z}$

- Für $k = 0, 1, 2, 3, 4$ ergeben sich Punkte auf dem Einheitskreis mit Winkelabstand $2\pi/5$ zueinander (die Ecken eines symmetrischen 5-Ecks!), z.B. ergibt $k = 0$ die Lösung $z = 1$, d.h. die Zahl mit Winkel $\alpha = 0 \cdot \frac{2\pi}{5} = 0$.
- Für $k \in \mathbb{Z} \setminus \{0, 1, 2, 3, 4\}$ ergeben sich zwar *neue Zahlen* α , aber **nicht neue Lösungen** $z \in \mathbb{C}$, denn ab $k = 5$ wiederholen sich die Ergebniszahlen z , weil sich immer Winkel der Form “echtes- 2π -Fünftel plus Vielfaches-von- 2π ” ergeben. Die Punkte mit Winkel “echten-Fünftel- 2π ” sind mit $k = 0, 1, 2, 3, 4$ schon entdeckt, während “Vielfaches-von- 2π ” für Winkelangaben unbedeutend ist! Zum Beispiel gilt wegen $5 \cdot 2\pi/5 = 2\pi$:

$$\begin{array}{lll} k = 5 & \text{heißt} & \alpha = 5 \cdot 2\pi/5 = 2\pi + 0 \quad \text{ergibt selbes } z \text{ wie bei } k = 0. \\ k = 6 & \text{heißt} & \alpha = (5 + 1) \cdot 2\pi/5 = 2\pi + 1 \cdot 2\pi/5 \quad \text{ergibt selbes } z \text{ wie bei } k = 1. \\ k = 7 & \text{heißt} & \alpha = (5 + 2) \cdot 2\pi/5 = 2\pi + 2 \cdot 2\pi/5 \quad \text{ergibt selbes } z \text{ wie bei } k = 2. \end{array}$$

Literaturverzeichnis