104.265 Algebra und Diskrete Mathematik für Informatik und Wirtschaftsinformatik

Jakob Kellner

2023S. Vorläufige End-Version vom
2023-07-02 18:47

Wer? Was? Warum?

- Wozu brauchen Sie Mathematik? Für alles mögliche.
- Wieso unterrichte *ich* diese LVA? Zufall / historische Gründe.
- Unterschiede der LVA zu bisherigen?
  Anderer Inhalt (i.A. weniger Stoff),
  andere Literatur (keine),
  (milde) Computerunterstützung.
  Erhoffte Vorteile: Interessanter (nützlicher?) für Sie.
  Offensichtliche Nachteile: Work in progress. Chaos.
- Warum einen Kurs entwickeln und nicht einfach existierende Goldstandard-LVA verwendne? (MIT OpenCourseWare oder dergl?) ...?

Kontaktmöglichkeiten

Ich will mich beschweren!
Ich habe Fragen bzw. Kommentare.

- Organisatorische und Inhaltliche Fragen: In den tuwel Foren ("normal" und anonym).
- Nur wenn vertraulich/sehr individuell/sehr dringend: email (oder tuwel Direktnachricht) an mich.
- Alle emails aussschließlich von Ihrem TU StudentInnen-Account senden! (Spamfilter, Zuordenbarkeit, "Datenschutz")
- Fragen zu den UE nicht an mich und nicht in die VO tuwel foren.
- Nach Ende: LVA-Bewertung via TISS.
  (Bei Kommentaren zur Übung: Bitte Übungsgruppe oder Leiter dazuschreiben.)

Organisatorisches zur VO

- Mo 14:10–15:50 und Do 13:10–14:50. Es werden 5 Einheiten ausfallen (noch nicht fix welche, wird rechtzeitig über tuwel bekannt gegeben).
- Prüfungstermine und -anmeldung im TISS.
- Andere Informationen über moodle=tuwel (nicht TISS).
- Prüfung: Multiple Choice Fragen die zufällig aus Fragepool ausgewählt werden. Falls ich rechtzeitig damit fertig werde können Sie davor (beliebig oft) unbenotete Prüfungs-Simulationen online (moodle) durchführen.
- Nur die VO-Prüfungen mit Prüfer *Kellner* prüfen den Stoff dieser VO. Andere Prüfer prüfen den Stoff ihrer eigenen Vorlesungen.
- Keine Anwesenheitspflicht.
- Lecturetube Aufzeichnung (keine Garantie).
- Unterlagen auf moodle.
- Sprechstunde nach Vereinbarung.

Organisatorisches zu den UE

- Wozu eine Übung?
- 5 Gruppen: Anmeldung bis 3.3. in TISS, 1. Einheit 13.3.
- Pro Übungseinheit: 4 Aufgaben Hausübung, 1 Aufgabe live, Lösungen werden von Studierenden an Tafel präsentiert
- Tafelmeldungen freiwillig, keine Anwesenheitspflicht, aber: **dringende** Empfehlung an **allen** Übungseinheiten teilzunehmen!
- 3 Tests (17.4., 22.5., 26.6.), jeweils max. 36 Punkte, die beiden besten Ergebnisse zählen für die Gesamtnote
- 1. Tafelmeldung (TM) 8 Punkte, 2. TM 4 Punkte, 3. TM 2 Punkte
- Notenschlüssel auf Basis von 80 Punkten (positiv ab 40), also: ab 2. TM Bonuspunkte
- bei Fragen: Leiter der UE-Gruppe (persönlich), TUWEL-Foren, `stefan.hetzl@tuwien.ac.at`

Computereinsatz

- Wir verwenden für Vorlesung jupyter, mit python3 und den python Paketen numpy, sympy (möglicherweise weiteren).
- Installation am einfachsten via anaconda: Installieren Sie sich `Anaconda 2022-10`, dann starten Sie `jupyter` (z.B. aus dem anaconda navigator).
- Bei der VO Prüfung wird (wenn überhaupt) nur minimales Wissen darüber verlangt, es muss nicht programmiert werden etc.
- Die zur Verfügung gestellten notebooks sind KEINE Beispiele für gelungene python Programmierung!

Logik und Sprache

Mathematische Sprache

- Die mathematische Sprache unterscheidet sich in einigen Punkten von der Alltagssprache (auch dort wo sie dieselben Wörter benutzt).
- Etwas tiefsinniger: Mathematische (und informatische) Logik unterscheidet sich wesentlich von Alltagslogik. (Bsp: monoton.)
- Die Logik ist "exakt". Die in der mathematischen Praxis verwendete Sprache aber nicht, man muss (wie in der Alltagssprache) den Kontext beachten.
  Es gibt zwar auch formale mathematische Sprachen (mathematische Logik), die werden aber nicht in der normalen mathematischen Praxis verwendet.
- Es ist wichtig dass Sie sowohl die mathematische Sprache als auch die Logik etwas (informell) kennnen- und verstehen lernen.
  Bsp: Def-Satz-Beweis
- Wir werden im Verlauf der Vorlesungen immer wieder einige Aspekte (informell) beschreiben.

Aussagen

- Aussage (engl. *proposition*): Ein (mathematischer) Satz dem man wahr oder falsch zuordnet/zuordnen kann/zuordnen will.
  - Bsp.: "7 ist eine Primzahl" ist wahr, $3 > 5$ ist falsch.
  - $x > 3$ ist eine Aussage (erst) wenn man die Variable x "belegt" hat. Für $x = 5$ wahr, für $x = 3$ falsch.
  - Keine Aussage ist z.B.: 2^3 (das ist ein math. Objekt, "Term").
  - Uns geht es um Mathematik. Ob "Komm her!" Aussage ist (wohl nicht), oder "Die Welt ist groß." (und ob wahr/falsch), ist hier egal.
- Im Folgenden stehen ϕ und ψ für Aussagen.
- Aussagenlogische Verknüpfungen:
  - Und: $\phi \wedge \psi$. Oder: $\phi \vee \psi$. Nicht: $\neg\phi$.
    Hat in Mathematik und Informatik präzise Bedeutung, anders als Alltagssprache ("Du kommst jetzt oder es gibt Probleme." eher xor).
    $\neg\neg\phi$ ist dasselbe wie ϕ.
  - Implikation: $\phi \to \psi$. ϕ impliziert ψ. Wenn ϕ dann ψ.
    $\phi \to \psi$ ist dasselbe wie $\neg\phi \vee \psi$
    Keine "kausale" Folgerung; "Wenn $2 > 3$ dann $2 > 7$" ist wahr.
  - Äquivalenz: $\phi \leftrightarrow \psi$. ϕ und ψ sind äquivalent. ϕ gdw. ψ ("genau dann wenn", engl. *iff*).

Wahrheitstabellen

- (Aussagen)variablen (hier: A,B, ...): Platzhalter für beliebige wahre/falsche Aussagen.
- Aussagenlogische Verknüpfung von (Aussagen)variablen nennt man (aussagenlogische) Formel. Bsp: $(A \to B) \leftrightarrow (B \to A)$
- Wahrheitswert einer Formel lässt sich einfach aus den Wahrheitswerten der Variablen berechnen, z.B. in Form einer "Wahrheitstabelle". Insbesondere:

A	$\neg A$
W	F
F	W

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

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

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

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

Allgemeingültige Aussagen

- Eine Formel heißt allgemeingültig, wenn sich für jede Belegung der Variablen der Wahrheitswert Wahr ergibt.
- Beispiel: $(A \to B) \leftrightarrow (\neg B \to \neg A)$ ist allgemeingültig.
  Überprüfung Mithilfe einer Wahrheitstabelle:

A	B	$A \to B$	$\neg A$	$\neg B$	$\neg B \to \neg A$	$(A \to B) \leftrightarrow (\neg B \to \neg A)$
W	W	W	F	F	W	W
W	F	F	F	W	F	W
F	W	W	W	F	W	W
F	F	W	W	W	W	W

  Anderes Beispiel: $(A \to B) \leftrightarrow (B \to A)$ ist nicht allgemeingültig:
- Bemerkung: P/NP-Problem.
  - Bei 2 Variablen (A und B) sind 4 Fälle zu überprüfen.
  - Bei n Variablen sind es 2^n Fälle.
  - Unser Algorithmus der überprüft ob eine Formel mit n Variablen allgemeingültig ist, benötigt also 2^n Schritte.
  - Fundamentale ungelöste Frage (P-NP Problem): Gibt es einen schnelleren (bzw sogar: polynomiellen) Algorithums?

Wichtige allgemeingültige Aussagen

In er Mathematik sind folgende einfache Tatsachen wichtig:

- Wenn ϕ gilt und $\phi \to \psi$, dann gilt ψ.
- $\phi \leftrightarrow \psi$ gdw: $\phi \to \psi$ und $\psi \to \phi$.
- $\phi \to \psi$ gdw. $\neg\psi \to \neg\phi$
- Aus $\phi \to F$ folgt $\neg\phi$. (Indirekter Beweis.)
- $F \to \phi$ gilt immer.

Zur Erinnerung/Bemerkung:

- ϕ is dasselbe wie "ϕ gilt".
- $\phi \wedge \psi$ ist dasselbe wie "ϕ und ψ".
- $\phi \leftrightarrow \psi$ ist dasselbe wie "ϕ gdw ψ".
- F ist "Falsch" bzw. eine beliebige falsche Aussage (wie $\neg 0 = 0$).

Quantoren

Allein mit Aussagenlogik kommt man in der Mathematik nicht weit.

Quantoren: Notation:

- $(\forall x)$ heißt: "Für all x" (aus dem Bereich über den wir reden).
- $(\exists x)$ heißt: "Es gibt ein x."

Mehr dazu später.

Zahlen

Natürliche, ganze, rationale Zahlen

Definition

- Natürliche Zahlen $\mathbb{N} = \{1, 2, \dots\}$ bzw. $\mathbb{N}_0 := \{0, 1, 2, \dots\}$. (Manchmal auch: $\mathbb{N} := \{0, 1, \dots\}$. Notation ist nicht einheitlich.)
- Ganze Zahlen $\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}$.
- Rationale Zahlen (Bruchzahlen) $\mathbb{Q}$ sind die Zahlen die sich als $\frac{n}{m}$ schreiben lassen, mit $n \in \mathbb{Z}$ und $m \in \mathbb{N}$.

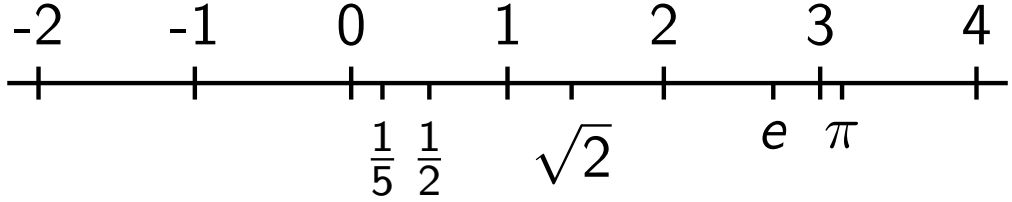
- "$\mathbb{Q}$ ist die Menge der Brüche" ist nicht ganz korrekt: $\frac{1}{2}$ und $\frac{2}{4}$ sind verschiedene Brüche, aber dieselbe rationale Zahl.
- Statt $\frac{z}{1}$ schreiben wir z, damit ist $\mathbb{Z}$ Teilmenge von $\mathbb{Q}$, dh. $\mathbb{Z} \subseteq \mathbb{Q}$.
- Brüche kann man "kürzen". Gekürzte Darstellung ist eindeutig. Bsp: $\frac{2}{3}$ ist die gekürzte Darstellung von $\frac{10}{15}$.
- Wir setzten voraus dass Sie Bruchrechnen beherrschen, zB $\frac{\frac{x}{y}}{\frac{a}{b}} = \frac{xb}{ya}$; $\frac{ax}{ay} = \frac{x}{y}$, aber i.A. $\frac{a+x}{a+y} \neq \frac{x}{y}$, etc.

Mengen-Notation 1

- Menge: Ein mathematisches Objekt das andere mathematische Objekte enthält. Ordnung egal, keine Wiederholungen.
- $x \in A$ heißt: x ist in A, x ist Element von A.
  $x \notin A$ heißt $\neg x \in A$ (gemeint natürlich: $\neg(x \in A)$, nicht $(\neg x) \in A$, was keinen Sinn ergibt).
  Bsp: $x \in \mathbb{N}$ heißt: x ist eine natürliche Zahl.
- $A \subseteq B$ heißt: A ist Teilmenge von B, d.h.: $x \in A \to x \in B$.
  Bsp: $\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q}$.
- {x,y,z} ist die Menge die genau die Elemente x, y und z hat.
  $\{x \in A : \phi(x)\}$ ist die Menge der Elemente in A die ϕ erfüllen.
  Bsp: $\{x \in \mathbb{N} : x \text{ gerade}\} = \{x \in \mathbb{N} : x\,(\exists y \in \mathbb{N})\, 2y = x\}$ ist die Menge der geraden natürlichen Zahlen.
  Alternative Schreibweise: $\{x \in A | \phi(x)\}$
- $\emptyset$ oder $\{\}$ ist die leere Menge.
- $A \cup B$ Vereinigung, $A \cap B$ Schnitt, $A \setminus B = \{x \in A : x \notin B\}$ Differenz.
- Der entsprechende python-type ist `set`

Reelle Zahlen

Die reelle Zahlen $\mathbb{R}$ sind "alle Zahlen auf dem Zahlenstrahl":



"Arbeitsdefinition"

Die reellen Zahlen $\mathbb{R}$ sind die Dezimalzahlen, d.h. Zahlen der Form $d = \pm r_n r_{n-1} \dots r_0 , r_{-1} r_{-2} \dots$ mit $n \in \mathbb{N}$ und jedes $r_i \in \{0, 1, \dots, 9\}$.

Es gilt dann $d = r_n 10^n + r_{n-1} 10^{n-1} + \dots = \sum_{i=n}^{-\infty} r_i 10^i$.
Bsp: $\pi = 3.1415 \dots$. Hier: $n = 0$, $r_0 = 3$, $r_{-1} = 1$, etc.

Zwei "Dezimalentwicklungen" können dieselbe reelle Zahl ergeben, z.B. $12,37000 \cdots = 12,36999 \cdots$ (nämlich $12,37$).
(Auch $-0 = +0$ oder $01,000 \cdots = 1,000 \dots$.)

Bemerkung: Mathematisch keine gute Definition für $\mathbb{R}$. Analog Binärentwicklung mit Basis 2 statt Basis 10, etc.

Nicht-rationale reelle Zahlen

Nicht alle "nützlichen" Zahlen sind in $\mathbb{Q}$.
Beispiele: π, oder $\sqrt{2} = 1.4142\ldots$, die (Quadrat)wurzel von 2, d.h., die (einzige) positive Zahl x die $x^2 = 2$ erfüllt.

Theorem

$\sqrt{2}$ ist keine rationale Zahl.

Der Beweis ist ca 2500 Jahre alt, (soweit ich weiss) der erste bekannte indirekte Beweis.

Was wir für den Beweis brauchen 1

Definition

n teilt m (für n, m in $\mathbb{Z}$) wenn es ein $\ell \in \mathbb{Z}$ gibt mit $m = \ell \cdot n$.
Auch geschrieben: $n|m$; $n \nmid m$ heißt $\neg n|m$.

Bsp: $3|12$, aber $\neg(3|13)$, auch geschriben als $3 \nmid 13$.

Z.B. aus dem Divisionsalgorithmus folgt:

Theorem

Für alle n, m in $\mathbb{Z}$ gibt es (eindeutige) ℓ, r in $\mathbb{Z}$ mit $0 \leq r < n$ und $m = \ell \cdot n + r$.

Bsp: Für $n = 20$ und $n = 7$ bekommen wir $20 = 2 \cdot 7 + 6$,
für $n = -17$ und $n = 3$ bekommen wir $-17 = -6 \cdot 3 + 1$, etc.
Der "Rest" r wird auch Modulus (mod) genannt, in python: `m % n`.

Definition

m heißt gerade, wenn $2|m$, sonst ungerade.

Was wir für den Beweis brauchen 2

Es gilt:

1. Die ungeraden Zahlen sind genau die der Form $2\ell + 1$.
   Beweis: Jede Zahl hat die Form $2\ell + r$ mit eindeutigem $0 \leq r < 2$, d.h. $r = 0$ oder $r = 1$.
2. Wenn n und m ungerade sind, dann ist $n \cdot m$ ungerade.
   Beweis: $n = 2\ell + 1$, $m = 2k + 1$, dann
   $nm = 4\ell k + 2(\ell + k) + 1 = 2(2\ell k + \ell + k) + 1$.
3. Wenn n^2 gerade ist, ist also auch n gerade.
   Beweis: Wäre n ungerade, dann auch $n \cdot n = n^2$.
4. Wenn $2 \cdot n$ durch 4 teilbar ist, dann ist n gerade.
   Beweis: $2 \cdot n = 4 \cdot k = 2 \cdot (2 \cdot k)$, d.h. $n = 2 \cdot k$.

Beweis von $\sqrt{2} \notin \mathbb{Q}$

Wir nehmen an $\sqrt{2} \in \mathbb{Q}$, expliziter:

$$\sqrt{2} = \frac{n}{m} \text{ (gekürzter Bruch)} \qquad (*)$$

(Jeder Bruch lässt sich ja kürzen).

- Quadrieren von $(*)$ gibt:
$$2m^2 = n^2 \qquad (**)$$
- n^2 ist also gerade, und daher auch n.
- Also hat n die Form $2 \cdot \ell$, und $n^2 = 4 \cdot \ell^2$ ist durch 4 teilbar.
- Laut $(**)$ ist $2m^2 = n^2$, d.h. $2m^2$ ist durch 4 teilbar, und damit ist m^2 gerade.
- Wie wir gesehen haben ist daher auch m gerade.
- Sowohl n als auch m sind also durch 2 teilbar, was der Annahme $(*)$ widerspricht. □

Definitionen

- In einer Defintion legen wir fest, welche Eigenschaft bzw. welches Objekt ein Buchstabe, ein Wort oder ein Symbol bedeuten soll.
- Fiktives Beispiel: Definition des Objekts (der Zahl) M (Million):
  M ist 10^6.
  Andere Schreibweise: $M := 10^6$ Oder schlicht: $M = 10^6$
  Andere Sprechweise: Sei $M := 10^6$. Oder: Setze $M := 10^6$.
- Beispiel: Definition der Eigenschaft Teilbarkeit bzw $|$:
  $n|m$ wenn es ein $\ell \in \mathbb{Z}$ gibt mit $m = \ell \cdot n$.
  Andere Schreibweise: $n|m$ wenn $(\exists \ell \in \mathbb{Z}) : \ m = \ell \cdot n$.
  Andere Schreibweise: $n|m \ :\leftrightarrow \ (\exists \ell \in \mathbb{Z}) : \ m = \ell \cdot n$ (oder nur $\leftrightarrow$).
- Achtung: Normalerweise ist "wenn" ungleich "genau dann wenn"!
  "Wenn $2 > 5$ dann $2 > 1$" ist wahr; aber "$2 > 5$ gdw $2 > 1$" nicht.
  Das "wenn" in der Definition oben heißt aber natürlich "gdw".

Definitionen (Forts.)

- Anderes Bsp: p heißt Primzahl, wenn gilt:
  $$p > 1, \text{ und wenn } p = x \cdot y, \text{ dann } x = 1 \text{ oder } y = 1.$$
  Hier ist das erste wenn ein gdw, das zweite nicht!
- Beachte den Kontext: Offenbar ist x, y in $\mathbb{N}$ gemeint, nicht in $\mathbb{Q}$.
  Wie schon betont: Mathematische (Alltags)sprache ist nicht formal exakt (aber die zugrundeliegende Logik schon).

Wohldefiniert

- Im "reinen" Fall haben Definitionen keine "Aussage" und können nicht wahr oder falsch sein: "p ist Primzahle wenn ..." legt einfach fest was wir mit dem Wort Primzahl meinen, sonst nichts.
- In der Praxis haben Definitionen aber sehr oft implizite Aussagen/Annahmen:
  - "Def.: A ist die kleinste natürliche Zahl" macht implizit die (korrekte) Annahme das es eine solche Zahl gibt (nämlich einfach 0).
  - "Def.: B ist die größte natürliche Zahl". macht die inkorrete Annahme dass es eine größte natürliche Zahl gibt. Diese "Definition" ist nicht zulässig!
- Wenn man betonen will dass die impliziten Voraussagen einer Definition erfüllt sind, sagt man oft "die Definition ist wohldefiniert".

Beweise!

Beweise sind ein zentraler Bestandteil der Mathematik.

- Was? Ein mathematischer Beweis ist eine Folge von "logischen" Ableitungsschritten, ausgehend von zugrundeliegenden Axiomen.

  Wir werden nicht exakt definieren was ein korrekter mathematischer Beweis bzw. was ein zulässiger Ableitungsschritt ist (das wäre aber durchaus möglich, aber nicht nützlich). Stattdessen vermitteln wir den Begriff anhand von Beispielen.

- Beweis != Beispiele.
  Behauptung: $(\forall x, y, z > 0)\ 313 \cdot (x^3 + y^3) \neq z^3$
  (Solche Fragen sind in der Kryptographie bedeutend.)
  Ausprobieren scheint die Behauptung zu bestätigen, aber sie ist falsch! Die ersten Gegenbeispiele treten erst bei >1000-stelligen Zahlen auf. (Man würde so ein Gegenbeispiel also nicht einmal mit den leistungsfähigsten Computern durch ausprobieren finden können.)

Beweise (Forts)

- Wozu? (In der Informatik)
  Theoretische Informatik hat viel mit Mathematik gemein; Beweise dass Algorithmen optimal sind etc etc.
  Und in der Praxis: Beim Programmieren ist es essentiell dass Sie ihren Code testen (unit-test etc). Ein Test ist aber nur so etwas wie ein Beispiel (für einen bestimmten Input). Besonders für kritische Systeme / grundlegende Hardware etc ist es manchaml wichtig, nicht nur anhand von Beispielen plausibel zu machen dass Code funktioniert, sondern es (für alle Inputs) zu beweisen.
- Wie?
  Was ist ein "richtiger" Beweis? Werden wir nicht definieren.
  Was ein "guter" Beweis? Noch weniger klar (elegant, informativ, . . .)
- welche Axiome? Mittelschulmathematik.
  (Kontext. Wenn in Übung gefragt ist: Leiten Sie aus
  $(a + b) + c = a + (b + c)$ ab dass
  $((a + b) + c) + d = (a + b) + (c + d)$, dann ist Lösung nicht "ist so".)

Beweise: Bsp

Beispiel $\phi \to \psi$.

- Methode 1:
  "Wir nehmen an, ϕ gilt. . . . (Arbeit). . . Es gilt also auch ψ"
  Beispiel: Wenn n und m ungerade sind, dann sind $m \cdot n$ ungerade.
  Beweis: Wir nehmen an, n und m sind ungerade. Dann ist $n = 2\ell + 1$ und $m = 2k + 1$. Daher ist $nm = 2(2\ell k + \ell + k) + 1$. Daher ist $n \cdot m$ ungerade. □
  Beweis vorbei (die Annahme gilt danach nicht mehr).
- Methode 2: Entspricht $\neg\psi \to \neg\phi$
  "Angenommen ψ gilt nicht. . . . (Arbeit). . . Dann gilt auch ϕ nicht."
  Beispiel: Wenn n^2 gerade ist, dann ist n gerade.
  Beispiel: Angenommen n ist nicht gerade (d.h. ungerade). Dann ist nach dem vorigen Satz $n \cdot n$ ungerade, d.h. n^2 ist nicht gerade. □
  Beweis vorbei (die Annahme gilt danach nicht mehr).

Nachtrag: Dezimaldarstellung

- Natürliche Zahlen werden üblicherweise in Dezimaldarstellung geschrieben (Basis 10). Zum Beispiel ist $2023 = 2 \cdot 10^3 + 0 \cdot 10^2 + 2 \cdot 10 + 3$.
- Allgemeine Darstellung: $a_n a_{n-1} \dots a_0$ bezeichnet $a_n 10^n + \cdots a_0 = \sum_{i=0}^{n} a_i 10^i$
  (Beachte: $10^0 = 1$. Allgemein: $x^0 = 1$.)
- Die natürliche Zahl "ist" aber nicht ihre Dezimaldarstellung. Stattdessen koennte man 2023 als 2023 viele Striche notieren. Oder:
- Binärdarstellung: 2023 (dezimal) ist in Binärdarstellung 11111100111, das ist $1 \cdot 2^{10} + 1 \cdot 2^9 + \cdots + 1$.

Nachtrag: Dezimaldarstellungen

- 126.3 bezeichnet $1 \cdot 10^2 + 2 \cdot 10^1 + 6 \cdot 10^0 + 3 \cdot 10^{-1}$.
- Auf diese Weise können wir Dezimalzahlen definieren: Eine reelle Zahl (Dezimaldarstellung) ist eine unendliche Folge

$$d = \pm r_n r_{n-1} \dots r_0 \, , r_{-1} r_{-2} \dots$$

mit $n \in \mathbb{N}$ und jedes $r_i \in \{0, 1, \dots, 9\}$, und bezeichnet die Zahl $d = r_n 10^n + r_{n-1} 10^{n-1} + \dots + r_0 10^0 + r_{-1} 10^{-1} + \dots = \sum_{i=n}^{-\infty} r_i 10^i$.

Maschinenzahlen (1/3)

- Computer können nicht mit unendlichen vielen Stellen operieren.
- Eine Möglichkeit: Fixe maximale Zahl von Vor- und von Nachkommastellen.
- Besser: Reelle Zahlen werden i.A. als (Varianten von) "floats" approximiert. Hier eine Dezimal-Variante ("wissenschaftliche Notation"):

Wiss. Notation

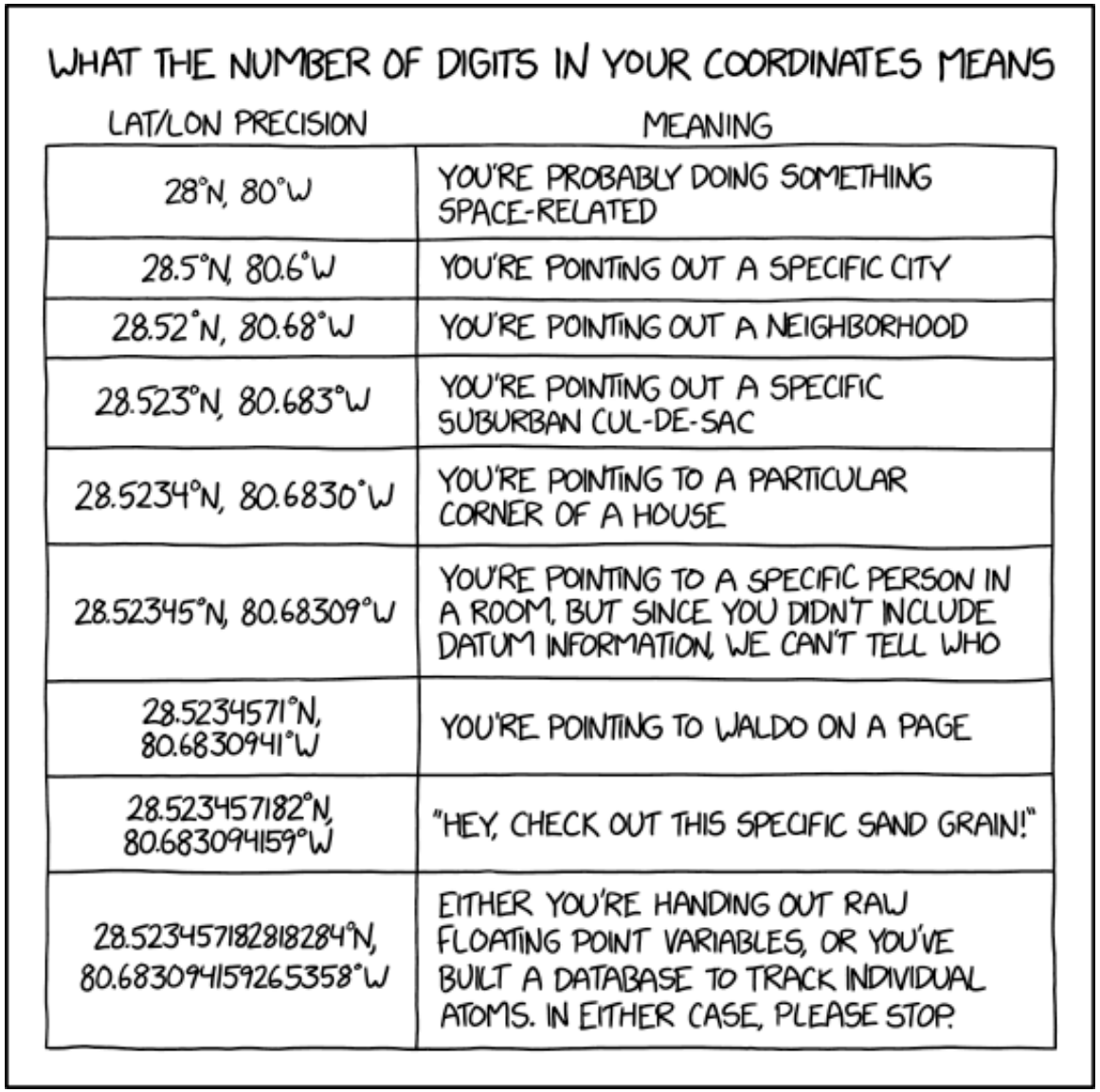
Eine Zahl (ungleich 0) wird dargestellt als: $\pm r_0, r_{-1} r_{-2} \cdots \cdot 10^e$ mit $e \in \mathbb{Z}$, $r_i \in 0, \ldots, 9$ und $1 \leq r_0 \leq 9$.

Bsp: Elektronenmasse $9.1093837 \cdot 10^{-31}$ kg

Bei Maschinenzahlen (als Bsp: double) gibt es feste Grenzen

- für e (ca. $-1000 < e < 1000$), und
- für die Anzahl der Nachkommastellen (bei "double" ca. 15.)

Maschinenzahlen (2/3)



WHAT THE NUMBER OF DIGITS IN YOUR COORDINATES MEANS

LAT/LON PRECISION	MEANING
28°N, 80°W	YOU'RE PROBABLY DOING SOMETHING SPACE-RELATED
28.5°N, 80.6°W	YOU'RE POINTING OUT A SPECIFIC CITY
28.52°N, 80.68°W	YOU'RE POINTING OUT A NEIGHBORHOOD
28.523°N, 80.683°W	YOU'RE POINTING OUT A SPECIFIC SUBURBAN CUL-DE-SAC
28.5234°N, 80.6830°W	YOU'RE POINTING TO A PARTICULAR CORNER OF A HOUSE
28.52345°N, 80.68309°W	YOU'RE POINTING TO A SPECIFIC PERSON IN A ROOM. BUT SINCE YOU DIDN'T INCLUDE DATUM INFORMATION, WE CAN'T TELL WHO
28.5234571°N, 80.6830941°W	YOU'RE POINTING TO WALDO ON A PAGE
28.523457182°N, 80.683094159°W	"HEY, CHECK OUT THIS SPECIFIC SAND GRAIN!"
28.5234571828284°N, 80.6830941592653 58°W	EITHER YOU'RE HANDING OUT RAW FLOATING POINT VARIABLES, OR YOU'VE BUILT A DATABASE TO TRACK INDIVIDUAL ATOMS. IN EITHER CASE, PLEASE STOP.

https://xkcd.com/2170/

Maschinenzahlen (3/3)

- Floats bieten für die meisten Fälle phantastische Genauigkeit.
- Rundungs"fehler" sind aber unvermeidlich und "unvorhersehbar". Die Spezifikationen des Verhaltens der floating point Arithmetik ist kompliziert und uneinheitlich; regelmäßig Verwirrungen, zB `https://stackoverflow.com/questions/56820/`.

```
> n = 5.59
> round(n, 1) # rounds to 1 dezimal place
5.599999999999996
```

- Moral:
  - Anzahl der Stellen != Genauigkeit
  - (Erst) bei der Ausgabe floats auf sinnvolle Zahl der (relevanten) Stellen runden.
  - Floats nie auf Gleichheit testen.
  - In der Informatik üblicherweise egal ob eine reele Zahl in $\mathbb{Q}$ ist oder nicht.

Komplexe Zahlen, Kartesisch

- Formal: Die komplexen Zahlen $\mathbb{C}$ sind Paare (a, b) reeller Zahlen, zusammen mit den folgenden Rechenregeln:
  $(a_1, b_1) + (a_2, b_2) := (a_1 + a_2, b_1 + b_2)$,
  $(a_1, b_1) \cdot (a_2, b_2) := (a_1 \cdot a_2 - b_1 \cdot b_2, a_1 b_2 + a_2 b_1)$.
  Das Paar $(0, 1)$ bezeichnen wir als i;
  die reelle Zahl a identifizieren wir mit $(a, 0)$.
- Schlampiger und besser: Eine komplexe Zahl hat die Form $z = a + ib$, wobei i eine "imaginäre" Zahl ist mit $i^2 = -1$.
- Notation: $\mathrm{Re}(a + ib) := a$, $\mathrm{Im}(a + ib) := b$ (Real- und Imaginärteil).
- Wenn $z = a + ib$, dann ist $z^* := a - ib = \mathrm{Re}\, z - i\, \mathrm{Im}\, z$ (Konjugierte).
  Oft auch $\bar{z}$ geschrieben.
  Es gilt $z_1^* + z_2^* = (z_1 + z_2)^*$ und $z_1^* \cdot z_2^* = (z_1 \cdot z_2)^*$.
- $|z| := \sqrt{a^2 + b^2} \geq 0$.
  Es gilt: $z \cdot z^* = |z|^2$; $|z| = 0$ gdw $z = 0$; $|z_1 \cdot z_2| = |z_1| \cdot |z_2|$.

Ein paar Beweise

Wir beweisen beispielhaft:

- Behauptung: $z_1^* \cdot z_2^* = (z_1 \cdot z_2)^*$.
  Beweis: Sei $z_1 = a_1 + ib_1$ und $z_2 = a_2 + ib_2$.
  $z_1 \cdot z_2 = (a_1 + ib_1)(a_2 + ib_2) = a_1a_2 - b_1b_2 + i(a_1b_2 + a_2b_1)$.
  Daher: $(z_1 \cdot z_2)^* = a_1a_2 - b_1b_2 - i(a_1b_2 + a_2b_1)$.
  $z_1^* \cdot z_2^* = a_1a_2 - (-b_1)(-b_2) + i\big(a_1(-b_2) + a_2(-b_1)\big) =$
  $a_1a_2 - b_1b_2 - i(a_1b_2 + a_2b_1)$. □
- Behauptung: $z \cdot z^* = |z|^2$.
  Beweis: $(a + ib)(a - ib) = a^2 - (ib)^2 = a^2 + b^2$. □
  (Verwendet: $(x + y)(x - y) = x^2 - y^2$)
- Behauptung: $|z| = 0$ gdw $z = 0$.
  Beweis: $a + ib \neq 0$ gdw ($a \neq 0$ oder $b \neq 0$) gdw $a^2 + b^2 \neq 0$. □
- Behauptung: $|z_1 \cdot z_2| = |z_1| \cdot |z_2|$.
  Beweis:
  $|z_1 \cdot z_2|^2 = (z_1 \cdot z_2) \cdot (z_1 \cdot z_2)^* = z_1 \cdot z_2 \cdot z_1^* \cdot z_2^* = |z_1|^2 \cdot |z_2|^2$. □

Division

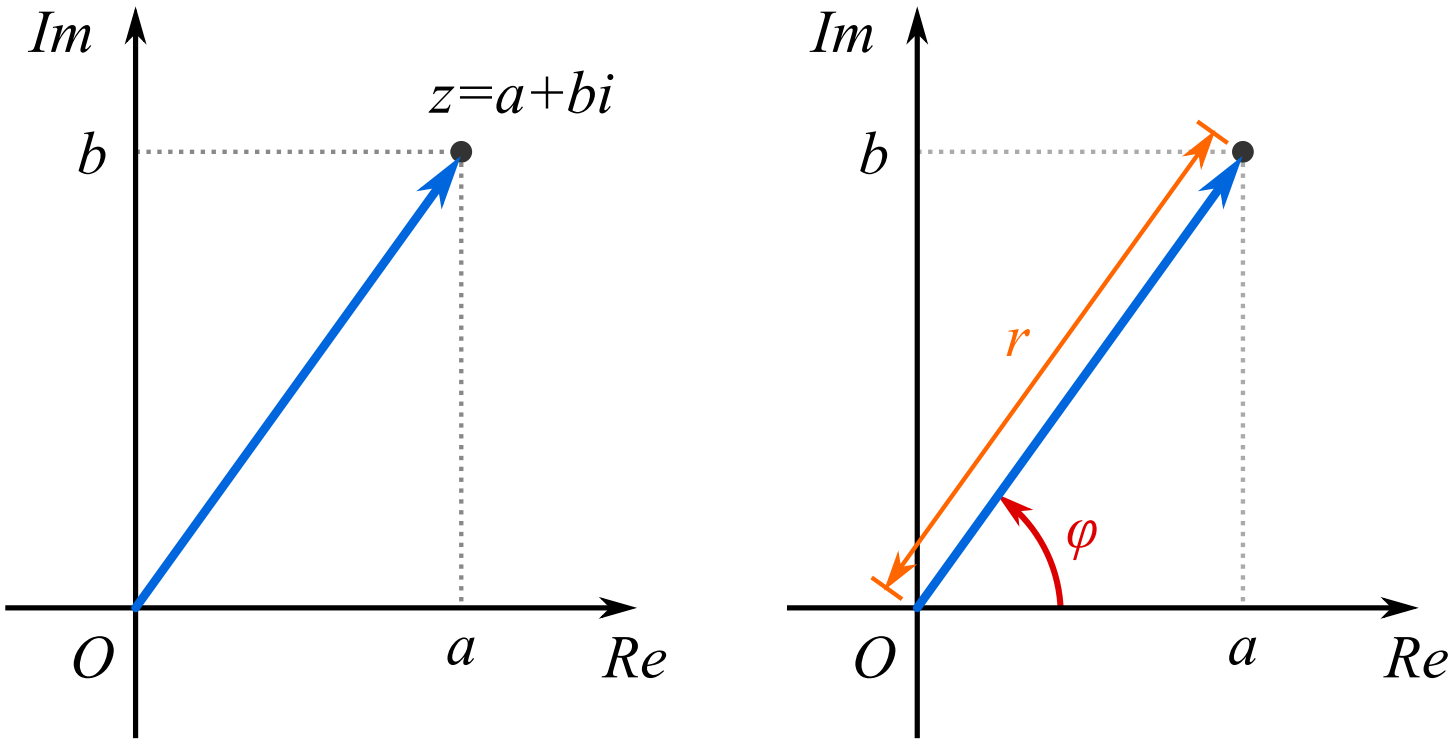
Wie können $\frac{z_1}{z_2}$ (für $z_2 \neq 0$) wie folgt berechnen:

$$\frac{z_1}{z_2} = \frac{z_1 z_2^*}{z_2 z_2^*} = \frac{z_1 z_2^*}{|z_2|^2}$$

Bsp: $\frac{2+3i}{1-2i} = \frac{(2+3i)(1+2i))}{5} = -\frac{4}{5} + \frac{7}{5}i$

Graphische Darstellung

Komplexe Ebene:

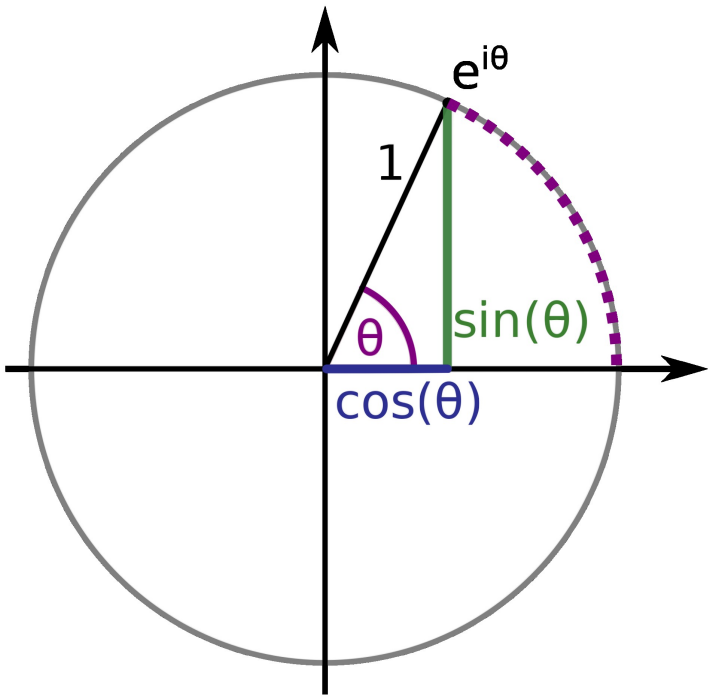


Ein Punkt z kann auch durch "Länge" $r \geq 0$ und Winkel φ beschrieben werden (Polardarstellung).
Schreibweise (nur hier): $z = (r, \varphi)$. ("Richtig": $re^{i\varphi}$.)
Dabei ist $r = |z|$, und $\varphi = \tan^{-1}(\frac{b}{a})$ falls $a \neq 0$ (sonst $\pm\frac{\pi}{2}$).

Polardarstellung

- D.h. mit (r, φ) ist (nur hier) $r\big(\cos(\varphi) + i\sin(\varphi)\big)$ gemeint.
  (Bessere/üblichere Schreibweise: $re^{i\phi}$.)
  Insbesondere ist $a = (a, 0)$ für $a \geq 0$, und $-a = (a, \pi)$.
- Nicht eindeutig: $(r, \varphi + 2k\pi) = (r, \varphi)$ für alle $k \in \mathbb{Z}$.
  (Und $(0, \varphi) = 0$ für alle φ.)
- Es gilt: $(r_1, \varphi_1) \cdot (r_2, \varphi_2) = (r_1 r_2, \varphi_1 + \varphi_2)$.
  Polardarstellung ist also nützlicher für Multiplikation (kartesische besser für Addition).
  Beweis: Additionstheoreme für sin und cos: $\mathrm{Re}(z_1 z_2) = a_1 a_2 - b_1 b_2 = r_1 \cos(\varphi_1) r_2 \cos(\varphi_2) - r_1 \sin(\varphi_1) r_2 \sin(\varphi_2) = r_1 r_2 \cos(\varphi_1 + \varphi_2)$;
  analog für $\mathrm{Im}(z_1 z_2)$. □
- Daraus folgt auch: $\frac{1}{(r,\varphi)} = (\frac{1}{r}, -\varphi)$, $\frac{(r_1,\varphi_1)}{(r_2,\varphi_2)} = (\frac{r_1}{r_2}, \varphi_1 - \varphi_2)$, und $(r, \varphi)^* = z^* = \frac{|z|^2}{z} = \frac{(r^2,0)}{(r,\varphi)} = (r, -\varphi)$.

Winkelfunktionen, Additionstheoreme



- Bem: griechische Kleinbuchstaben
  ϕ oder φ: "phi", ψ: "psi" und
  θ oder ϑ: "theta"
  bezeichnen oft Winkel, Formeln, ...
- Winkel üblicherweise gegen den Uhrzeigersinn; oft von x-Achse aus.
- Winkel sind "dimensionslos":
  θ ist "Kreisbogenlänge durch Radius"
- Insbes: 2π = ganzer Kreis = $360°$
  π = Halbreis = $180°$ $\quad \frac{\pi}{2} = 90°$
- $\cos(\theta) = \sin(\frac{\pi}{2} - \theta)$
- $\sin^2(\theta) + \cos^2(\theta) = 1$

Additionstheoreme (ohne Beweis):

- $\cos(x + y) = \cos(x)\cos(y) - \sin(x)\sin(y)$
- $\sin(x + y) = \sin(x)\cos(y) + \sin(y)\cos(x)$

Produkt komplexer Zahlen in Polarkoordinaten

Theorem

(In Polarkoordinaten) $(r_1, \phi_1)(r_2, \phi_2) = (r_1 r_2, \phi_1 + \phi_2)$

Beweis.

(r_i, ϕ_i) ist $r(\cos(\phi) + i\sin(\phi))$. Daher ist

$(r_1, \phi_1)(r_2, \phi_2) = r_1(\cos(\phi_1) + i\sin(\phi_1)) \cdot r_2(\cos(\phi_2) + i\sin(\phi_2)) =$

$$r_1 r_2 \Big(\cos(\phi_1)\cos(\phi_2) - \sin(\phi_1)\sin(\phi_2) + i\big(\cos(\phi_1)\sin(\phi_2) + \cos(\phi_2)\sin(\phi_1)\big)\Big) =$$

$r_1 r_2\big(\cos(\phi_1 + \phi_2) + i\sin(\phi_1 + \phi_2)\big)$, d.h. $(r_1 r_2, \phi_1 + \phi_2)$ □

Bemerkung: Wenn man weiss dass $(r, \phi) = re^{i\phi}$ (und dass sich die Exponentialfunktion im Komplexen anständig verhält) bekommt man das sofort so: $r_1 e^{i\phi_1} \cdot r_2 e^{i\phi_2} = r_1 r_2 e^{i\phi_1} e^{i\phi_2} = r_1 r_2 e^{i\phi_1 + i\phi_2} = r_1 r_2 e^{i(\phi_1 + \phi_2)}$.

Python

- In python sind komplexe Zahlen (type `complex`) als paar von floats (Kartesisch) implementiert.
- i kann im Code durch `1j` eingegeben werden: `2+3j` ergibt $2 + 3i$, etc. (j alleine ist nicht definiert bzw steht als (Variablen)-Name zur Verfügung.)
- Natürlich haben `numpy` und `sympy` ausführliche Unterstützung von Komplexen Berechnungen.
- Wenn nicht als reell spezifiziert dann geht `sympy` üblicherweise davon aus dass eine Variable/Funktion komplex ist:

```
In [45]: y = sp.Symbol('y')
In [46]: sp.simplify(sp.log(sp.exp(y)))
Out[46]: log(exp(y))

In [47]: y = sp.Symbol('y', real=True)
In [48]: sp.simplify(sp.log(sp.exp(y)))
Out[48]: y
```

z^n

- Wenn $z = (r, \varphi)$, dann ist also $z^2 = (r \cdot r, \varphi + \varphi) = (r^2, 2\varphi)$.
  Und $z^3 = z(z^2) = (r \cdot r^2, \varphi + 2\varphi) = (r^3, 3\varphi)$. Allgemein:
  $z^n = (r^n, n \cdot \varphi)$.
- Für $n \geq 1$ gibt es für jedes $u = (r, \phi)$ eine n-te Wurzel $z_0 = (\sqrt[n]{r}, \frac{\phi}{n})$.
  (n-te Wurzel heißt: $z_0^n = u$).
- Für $u \neq 0$ gibt genau n solcher Wurzeln:
  $z_j = (\sqrt[n]{r}, \frac{\varphi + 2j\pi}{n})$ für $j = 0, \ldots, n-1$.
  ($\frac{\varphi+2n\pi}{n} = \frac{\varphi}{n} + 2\pi$ liefert wieder z_0, etc.)
- Bsp: Quadratwurzeln von $i = (1, \frac{\pi}{2})$ sind $(1, \frac{\pi}{4})$ und $(1, \frac{3\pi}{4})$.
  Die 4-ten Wurzeln von $1 = (1, 0)$ sind 1, i, -1, $-i$.
- Die n vielen n-ten Wurzeln von 1 heißen n-te Einheitswurzeln.
- Bemerkung: In $\mathbb{C}$ ganz anders als in $\mathbb{R}$: In $\mathbb{C}$ gibt es immer Wurzeln (in $\mathbb{R}$ gibt es $\sqrt{-1}$ nicht); und für $n > 1$ ist $z \mapsto z^n$ ist in $\mathbb{C}$ nicht "injektiv": $z_1 \neq z_2$ impliziert nicht $z_1^n \neq z_2^n$. (In $\mathbb{R}$ ist z.B. x^3 injektiv.)

Fundamentalsatz

- $f(z) = a_n z^n + a_{n-1} z^{n-1} + \cdots + a_0$ mit $a_n \neq 0$ und $a_i \in \mathbb{C}$ heißt (komplexes) Polynom vom Grad n; die a_i heißen Koeffizienten.
- Fundamentalsatz der Algebra: Jedes Polynom von Grad ≥ 1 hat (mindestens) eine Nullstelle in $\mathbb{C}$, das ist ein z_0 mit $f(z_0) = 0$.
- Bemerkung: I.A. lässt sich die Nullstelle nicht "als Wurzel schreiben", selbst wenn die Koeffizienten rational sind.
  (Als Wurzeln schreiben heißt so wie $\sqrt[17]{(3+\sqrt[5]{7})\sqrt{5}}$ etc.)
  Für nicht-rationale Koeffizienten wenig überraschend: Jede Zahl z_0 ist NS eines Polynoms, nämlich $z - z_0$, und nicht jede Zahl lässt sich als Wurzel schreiben (z.B. π).
- Die NS von Polynomen von Grad $\leq$4 mit (komplex) rationalen Koeeffizienzen lassen sich immer als Wurzeln schreiben.
  Keine Nullstelle von $z^5 - z + 1$ lässt sich so schreiben.
- In $\mathbb{R}$ hat jedes Polynom ungeraden Grades ein Nullstelle, aber zB $x^2 + 1$ hat keine.

Fundamentalsatz (Forts.)

- Es gilt: Das Produkt $g(z) := f_1(z) \cdot f_2(z)$ von Polynomen f_1, f_2 ist ein Polynom.
- Es gilt: $g(z_0) = 0 \leftrightarrow (f_1(z_0) = 0 \vee f_2(z_0) = 0)$.
  D.h., die NS von g ist die Vereinigung der NS von f_1 und f_2.
- Falls $f_1(z) = (z - z_0)$, dann ist z_0 NS von f_1 (und von g).
- Wir werden hoffentlich später noch eine "Umkehrung" sehen:
  Wenn z_0 Nullstelle eines Polynoms $g(z)$ von Grad n ist, dann ist $g(z) = (z - z_0) \cdot f(z)$ für ein Polynom $f(z)$ mit Grad $n - 1$.
- Der Fundamentalsatz sagt dann: Jedes Polynom von Grad $n \geq 1$ läßt sich als $(z - z_1) \cdot (z - z_2) \cdots (z - z_n)$ schreiben, für $z_i \in \mathbb{C}$.
  Das Polynom "zerfällt" in lineare Terme.

Fundamentalsatz Beispiele

- Bsp: $2x - 1 = 0$ hat die reelle Nullstelle $\frac{1}{2}$, aber keine NS in $\mathbb{Z}$.
- Bsp: $x^2 + 1$ hat komplexe Nullstellen $\pm i$ (aber keine reellen).
- Bsp: $x^3 - 2$ hat eine NS in $\mathbb{R}$, und zwar $\sqrt[3]{2}$, zwei weitere Nullstellen in $\mathbb{C}$ (Polar: $(\sqrt[3]{2}, \frac{2\pi}{3})$ und $(\sqrt[3]{2}, \frac{4\pi}{3})$), und gar keine in $\mathbb{Q}$, weil $\sqrt[3]{2} \notin Q$.
- Bsp: $z^5 - z + 1$ hat eine reelle $(-1.1673\ldots)$ und 4 weitere komplexe NS. Keine NS läßt sich "als Wurzel schreiben".

Bedeutung der Komplexen Zahlen

- Ursprünglich eingeführt um Nullstellen von Polynomen (auch reelle) zu bestimmen/finden.
- Sehr nützlich um alle möglichen reellen Funktionen effektiver zu beschreiben (Bsp: $\mathrm{Re}(e^{i\omega t})$ statt $\cos(\omega t)$.)
- Die Struktur von $\mathbb{C}$ ist ein sehr natürliche / fundamentale; und taucht daher in verschiedensten Zusammenhängen auf. Sie entspricht zB fundamentalen Eigenschaften der Quantenmechanik (Wellenfunktion komplexwertig).
- ...

Relationen und Funktionen

Notation: Folgen

- Eine Folge ist eine Aneinanderreihung von Objekten, wobei Reihenfolge und Wiederholungen wichtig ist.
  Die Objekte heißen Elemente der Folgen oder Folgenglieder.
  Schreibweise: $(a_0, a_1, \dots)$, manchmal auch: $\langle a_1, a_2 \dots \rangle$.
- Wenn alle Elemente 0 oder 1 sind, spricht man von einer 0-1-Folge oder binären Folge.
- Beispiel: $(0, 1)$, $(1, 0)$, $(1, 0, 0)$ sind paarweise verschiedene binäre Folgen.
  ("Paarweise verschieden": je zwei sind verschieden.)
- Beispiel: Die Elemente von Folgen müssen keine Zahlen sein:
  $(\sin, \cos, \tan)$ ist eine Folge (der Länge 3) von Funktionen.
- Es gibt eine (eindeutige) "leere Folge"; manchmal mit $\langle\rangle$ bezeichnet, oder λ ("Lamda") oder ε ("Epsilon").
- Eine endliche Folge $(a_0, a_1, \dots a_{n-1})$ der Länge n nennt man auch n-Tupel. Ein 2-Tupel heißt (geordnetes) Paar.

Notation: Kartesisches Produkt

Seien A, B.

- $A \times B$ ist die Menge der geordneten Paare (a, b) mit $a \in A$ und $b \in B$.
  Bsp: $(i, 2)$ ist sowohl in $\mathbb{C} \times \mathbb{C}$ als auch in $\mathbb{C} \times \mathbb{N}$
  (aber $(i, 2) \notin \mathbb{N} \times \mathbb{C}$, weil $i \notin \mathbb{N}$).
- $A^2 := A \times A$.
  Bsp: $\mathbb{R}^2$ ist die (reelle) Ebene (kartesische Koordinaten/Vektoren).
  Statt $(12, 3)$ schreibt man in dem Fall manchmal auch $\begin{pmatrix} 12 \\ 3 \end{pmatrix}$
- Analog für n statt 2, für $A_1, \ldots, A_n$ Mengen:
  $A_1 \times \cdots \times A_n$ ist die Menge von n-Tupeln mit $a_i \in A_i$.
- $A^n = A \times \cdots \times A$.
  Z.B. $\mathbb{R}^3$ ist der 3-dimensionale Raum.

Funktionen: Informelle Definition

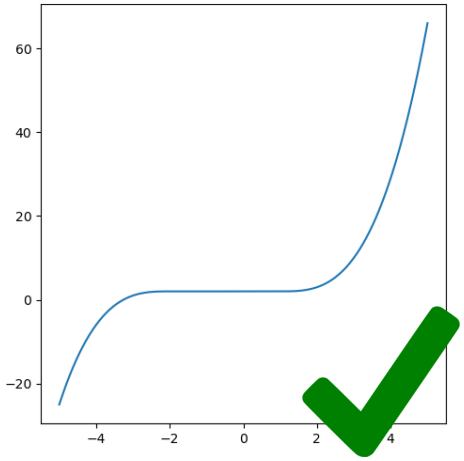
- Eine Funktion $f : A \to B$ ordnet jedem "input" aus der Menge A (Definitionsbereich) genau einen "output" aus B (Wertebereich) zu.
- Notation zur Definition/Beschreibung von Funktionen, am Bsp der Quadrat-Funktion:
  $f : \mathbb{R} \to \mathbb{R} : \; x \mapsto x^2$; oder: "so daß $f(x) = x^2$" oder: "definiert durch" oder: "mit".
  "Schlampig" oft auch (die Funktion) x^2 oder $f = x^2$ oder $f(x) = x^2$.
  (Warum schlampig? Es ist hier nicht klar won welcehm Definitionsbereich wir ausgehen: $\mathbb{N}$? Oder $\mathbb{C}$? Bei "die lineare Funktion $ax + b$" ist nur implizit angedeutet dass x die Variable ist und a, b Konstanten, etc.)

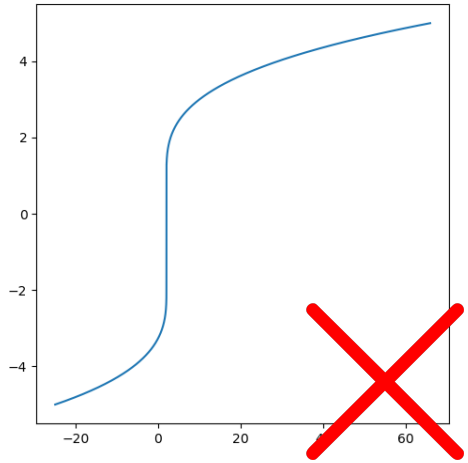
Funktionen: Formale Definition

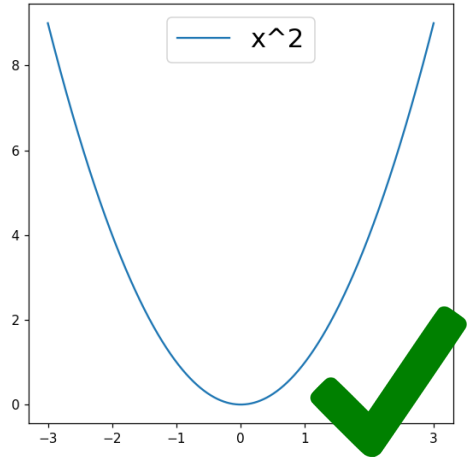
- Formal: Eine Funktion $f : A \to B$ ist eine Teilmenge von $A \times B$ so dass $(\forall a \in A)(\exists! b \in B)(a, b) \in f$.
  Statt $(a, b) \in f$ schreibt man $f(a) = b$.
- Notation: $\exists! x$ heißt: "Es gibt genau ein x".
  $\exists x$ heißt ja: Es gibt ein x (wobei es auch mehrere geben kann).
- Bsp: In diesem Sinn ist die Funktion $f : \mathbb{N} \to \mathbb{N}$ $n \mapsto 2^n$ die folgende Teilmenge von $\mathbb{N}^2$: $\{(0, 1), (1, 2), (2, 4), (3, 8), (4, 16), \dots\}$.
- Allgemein: $f : A \to B$ ist identisch mit $\{(x, f(x)) : x \in A\}$.
- Insbesondere:
  - $x^2 : \mathbb{N} \to \mathbb{N}$ ist identisch mit $x^2 : \mathbb{N} \to \mathbb{R}$.
    Wenn man den Wertebereich vergrößert ändert das die Funktion nicht.
  - (Bemerkung: Alternative formale Definition von Funktion inkludiert Wertebereich; dann wären die Funktionen unterschiedlich.)
  - $x^2 : \mathbb{N} \to \mathbb{R}$ ist nicht identisch mit $x^2 : \mathbb{R} \to \mathbb{R}$. Es kommen ja neue Paare dazu, z.b. $(\frac{1}{2}, \frac{1}{4})$.
  - Bemerkung: $x^2 : \mathbb{R} \to \mathbb{N}$ ist dagegen keine zulässige Funktion:
    $\frac{1}{2}$ hat ja keinen passenden Funktionswert im Wertebereich.

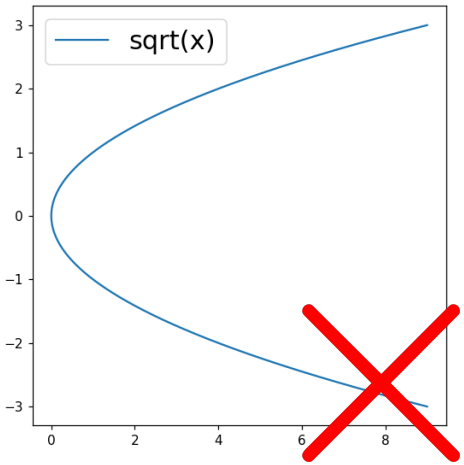
Graphen

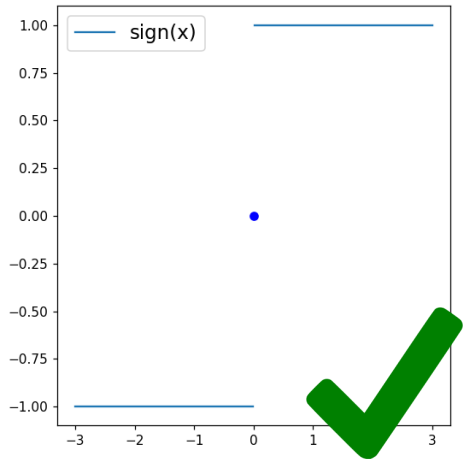
Für $f : A \to \mathbb{R}$ mit $A \subseteq \mathbb{R}$ entpricht f als Menge, d.h. $\{(x, f(x)) : x \in A\}$ einer Teilmenge der Ebene, genannt Funktionsgraph.
Nicht jede Kurve im $\mathbb{R}^2$ ist ein Graph!

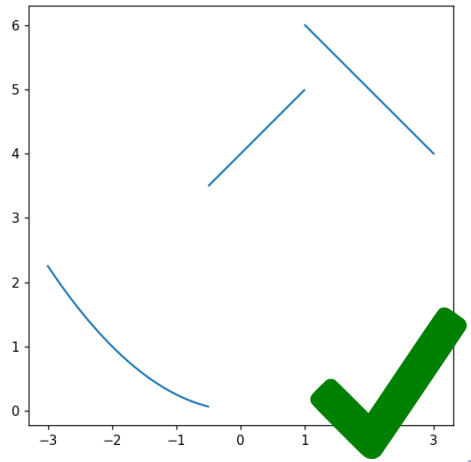












Totale und partielle Funktionen

- In der Mathematik (und in dieser Vorlesung) ist eine Funktion $f : A \to B$ auf ganz A definiert ("totale Funktion").
- In der Informatik ist es oft natürlicher, mit "Funktion" etwas zu nennen dass auch für manche inputs undefiniert sein kann ("partielle Funktion").
- Beispiel: $\frac{1}{x} : \mathbb{R} \to^{\text{part}} \mathbb{R}$ ist eine partielle Funktion (undefiniert auf 0). $\frac{1}{x} : \mathbb{R} \setminus \{0\} \to \mathbb{R}$ ist eine Funktion.
- Beispiel: Sei f^0 die "Position der ersten 0"-Funktion: Sie bildet eine (endliche) binäre Folge $(a_0, \ldots, a_n)$ auf das erste ℓ ab mit $a_\ell = 0$. Bsp.: $f^0(\langle 1, 0, 0\rangle) = 1$, $f^0(\langle 0, 0, 0\rangle) = 0$ aber $f^0(\langle 1, 1\rangle)$ ist undefiniert. f^0 ist ein partielle Funktion von den Binärfolgen nach $\mathbb{N}$. f^0 ist (totale) Funktion von "Binärfolgen die eine 0 enthalten" nach $\mathbb{N}$.
- Allgemein: $f : A \to^{\text{part}} B$ ist eine partielle Funktion gdw: $\exists A' \subseteq A : f : A' \to B$ (eine totale Funktion).

Injektiv

- $f : A \to B$ ist injektiv, wenn $(\forall a_1, a_2 \in A)\, a_1 \neq a_2 \to f(a_1) \neq f(a_2)$.
- Beispiele für injektive Funktionen:
  $x + 1$, $x^2 : \mathbb{N} \to \mathbb{N}$, $\frac{x}{2} : \mathbb{N} \to \mathbb{Q} \ldots$
- Beispiele für nicht injektive Funktionen:
  Konstante Funktionen, $x^2 : \mathbb{R} \to \mathbb{R}$ (weil $-1^2 = 1^2$), $\sin(x) : \mathbb{R} \to \mathbb{R}$, ...

Bildmengen

Sei $f : A \to B$.

- Für $X \subseteq A$ ist $f''X := \{f(x) : x \in X\}$. ("Bild" von X, "image")
  (Manchmal auch $f[X]$ geschrieben; manchmal sogar einfach $f(X)$.)
- $f''A$ heißt auch "Bild von f"
- Es gilt: $f''(X_1 \cup X_2) = (f''X_1) \cup (f''X_2)$.
  Es gilt aber i.A. nicht $f''(X_1 \cap X_2) = (f''X_1) \cap (f''X_2)$, nur
  $f''(X_1 \cap X_2) \subseteq (f''X_1) \cap (f''X_2)$.

Surjektiv und bijektiv

- $f : A \to B$ heißt surjektiv, wenn $f''A = B$.
  (Äquivalent: wenn $(\forall b \in B)\, |f^{-1}(b)| \geq 1$, oder:
  $(\forall b \in B)\, (\exists a \in A)\, f(a) = b$.)
- Achtung: Surjektivität hängt von "Schreibweise" ab: Sei $f(x) = \frac{1}{x}$.
  Dann ist $f : \mathbb{R} \setminus \{0\} \to \mathbb{R}$ nicht surjektiv; aber $f : \mathbb{R} \setminus \{0\} \to \mathbb{R} \setminus \{0\}$ schon.
- Insbesondere: $f : A \to f''A$ ist immer surjektiv.
- Bijektiv heißt: Injektiv und surjektiv.
- Für $f : A \to B$ gilt also: f injektiv gdw. $f : A \to f''B$ bijektiv.

Verknüpfung, Einschränkung

- Wenn $f : A \to B$ und $C \subseteq A$, dann ist $f \restriction C : C \to B$ die Einschränkung auf die kleinere Definitionsmenge (dh. $(f \restriction C)(x) = f(x)$ für alle $c \in C$.)
- Wenn $C \neq A$, dann ist $f \restriction C$ eine andere Funktion als f.
- Bsp: $f = x$ und $g = x^2$ und $h = 2^x - 1$, alle mit Definitionsmenge $\mathbb{N}$. Das sind paarweise verschiedene Funktionen.
  Für $C = \{0\}$ gilt: $f \restriction \{0\}$ und $g \restriction \{0\}$ und $h \restriction \{0\}$ sind dieselbe Funktion $i : \{0\} \to \mathbb{N}$ mit $0 \mapsto 0$.
- Wenn $f : A \to B$ und $g : B \to C$, dann ist die "Verknüpfung" $g \circ f : A \to C$ definiert durch $x \mapsto g(f(x))$.
  Gesprochen: "f vor g", "g nach f", "g Kringel f".
- Bsp: $f = x + 1$, $g = x^2$.
  Dann $g \circ f = (x + 1)^2$ und $f \circ g = x^2 + 1$.

Unendlichkeiten

Abzählbarkeit

Gegeben eine Funktion f mit Definitionsbereich $\mathbb{N}$. Dann können wir $A := f''\mathbb{N}$ schreiben als $\{f(0), f(1), \dots\}$, und wir sagen "f zählt A ab".

Definition

A heißt abzählbar, wenn
$A = f''\mathbb{N}$ für ein f mit Definitionsbereich $\mathbb{N}$, oder wenn $A = \emptyset$.

Eine unendliche, abzählbare Menge heißt "abzählbar unendlich."

(Warum muss man den Fall $\emptyset$ der leeren Menge extra anführen?)

Theorem

Jede endliche Menge ist abzähbar.

Beweis.

Wenn $A = \emptyset$ dann ist nichts zu tun. Sei $A = \{a_0, a_1, \dots a_n\}$ mit $n \in \mathbb{N}$.

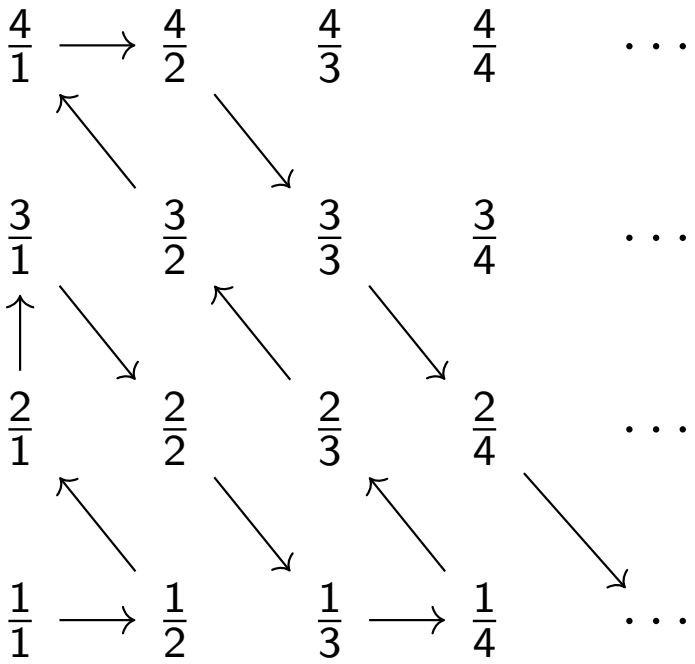
Definiere $f(\ell) := \begin{cases} a_\ell & \text{wenn } \ell \leq n \\ a_0 & \text{sonst.} \end{cases}$ □

Abzählbarkeit von $\mathbb{Q}$

Theorem (Cantor)

$\mathbb{Q}$ *ist abzählbar.*

Der übliche Beweis:



D.h., $(1, \frac{1}{2}, 2, 3, \cancel{\frac{2}{2}}, \frac{1}{3}, \frac{1}{4}, \frac{3}{2}, 3, \cancel{\frac{4}{2}}, \cancel{\frac{3}{3}}, \cancel{\frac{2}{4}}, \dots)$ zählt $\mathbb{Q}^+$ ab (ohne Widerholungen, wenn gewünscht)

Wie sehen uns das nun etwas abstrakter an.

Mengen-Notation

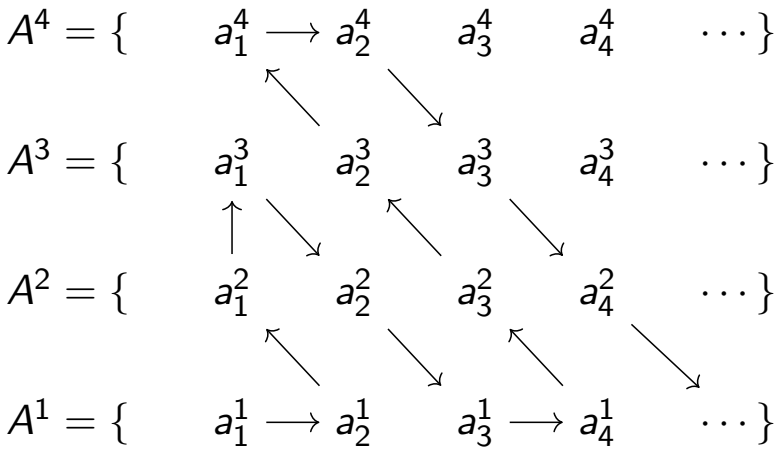
- Sie I eine Menge ("Indexmenge") und für jedes $i \in I$ sei A_i eine Menge. Dann ist $\bigcup_{i \in I} A_i := \{x : (\exists i \in I) x \in A_i\}$
- Bsp: Für $I = \{12, 13, 37\}$ ist $\bigcup_{i \in I} A_i = A_{12} \cup A_{13} \cup A_{27}$. (Endliche Vereinigung.)
- Bsp: Für $I = \mathbb{N}$ ist $\bigcup_{i \in \mathbb{N}} A_i = A_0 \cup A_1 \cup A_2 \cup \ldots$. (Abzählbare Vereinigung.)
- Wenn I endlich ist nennt man das eine endliche Vereinigung.
- Wenn I abzählbar ist, nennt man das eine abzählbare Vereinigung. (Unabhängig davon ob $\bigcup_{i \in I} A_i$ abzählbar ist oder nicht.)
- Bsp: Wenn $A_n = [0, n] := \{x \in \mathbb{R} : 0 \leq x \leq n\}$ dann ist $\bigcup_{n \in \mathbb{N}} A_n = [0, \infty) := \{x \in \mathbb{R} : 0 \leq x\}$.
  Abzählbare Vereinigung, keines der A_i ist abzählbar (wie wir später sehen werden).
- Analog bekommen wir für den Schnitt $\bigcap_{i \in I} A_i$.

Einfache Eigenschaften der Abzählbarkeit 1

Theorem

Die abzählbare Vereinigung abzählbarer Mengen ist abzählbar.
D.h.: Wenn A^n abzählbar für jedes $n \in \mathbb{N}$, dann ist die Vereinigung $\bigcup_{n\in\mathbb{N}} A^n$ abzählbar.

Beweis: Jedes A^n ist abzählbar, läßt sich also als $\{a_1^n, a_2^n, \dots\}$ schreiben.



Dann ist $\bigcup_{n\in\mathbb{N}} A^n = \{a_1^1, a_2^1, a_1^2, \dots\}$.
In anderen Worten: Wenn wir $f(n)$ definieren als: "der n-te Eintrag wenn wir der gezeichneten Schlangenlinie folgen", dann ist $f : \mathbb{N} \to \bigcup_{n\in\mathbb{N}} A_n$ surjektiv.

Einfache Eigenschaften der Abzählbarkeit 2

Theorem

Wenn $A \neq \emptyset$ abzählbar ist, dann gibt es eine Funktion $f : \mathbb{N} \to \mathbb{A}$ so dass $(\forall a \in A)\,(\exists^{\infty} n \in \mathbb{N})\, f(n) = a$.

Dabei heißt $\exists^{\infty}$ "es gibt unendlich viele".

D.h. $A = \{a_0, a_1, \dots\}$, wobei jedes Element von A unendlich oft vorkommt.

Beweis: Setze $A^n := A$ für jedes $n \in \mathbb{N}$ und verwende denselben Beweis des vorigen Theorems. □

(Bemerkung: Es nützt nicht das vorige Theorem selbst zu verwenden! Der Beweis des vorigen Theorems konstruiert aber genaue ein f so wie wir es brauchen.)

Einfache Eigenschaften der Abzählbarkeit 3

Wir haben Definiert: "A ist abzählbar" wenn "$(\exists f)\, f : \mathbb{N} \to A$ surjektiv."
Wir haben bereits gesehen:

- Endliche Mengen sind abzählbar.
- Die Vereinigung abzählbar vieler abzählbarer Mengen ist abzählbar.

Theorem

Wenn A abzählbar ist und $B \subseteq A$, dann ist B abzählbar.

Beweis: Sei $B \neq 0$ (ansonsten ist nichts zu tun), fixiere $b_0 \in B$.
Angenommen $f : \mathbb{N} \to A$ surjektiv.
Definiere g durch $g(n) = \begin{cases} f(n) & \text{wenn } f(n) \in B \\ b_0 & \text{sonst.} \end{cases}$
Dann $g : \mathbb{N} \to B$ surjektiv: Wenn $b \in B \subseteq A$ dann gibt es $n \in \mathbb{N}$ mit $f(n) = b$, und dann $g(n) = b$. □

Einfache Eigenschaften der Abzählbarkeit 4

Theorem

Wenn A abzählbar unendlich ist, dann gibt es ein $f : \mathbb{N} \to A$ bijektiv.

D.h. $A = \{a_0, a_1, \dots\}$ wobei jedes $a \in A$ genau einmal vorkommt.

Beweis (Teil 1):
Wissen: Es gibt $g : \mathbb{N} \to A$ surjektiv. Z.Z: Es gibt ein $f : \mathbb{N} \to A$ bijektiv.

- Wir konstruieren f induktiv. Das heißt: Wir konstruieren erst $f(0)$, dann $f(1)$ wobei wir $f(0)$ verwenden können, etc.
- Setze $f(0) := g(0)$.
- Für $f(1)$: Es gibt ein kleinstes/erstes ℓ so dass $g(\ell) \neq f(0)$.
  (A ist ja unendlich, und habt insbesondere nicht nur ein Element.)
  Setze $f(1) := g(\ell)$.
- Allgemein: Sei ℓ minimal mit $g(\ell) \notin \{f(0), f(1), \dots, f(n-1)\}$.
  Setze $f(n) := g(\ell)$.
- Das definiert $f : \mathbb{N} \to A$. Z.z.: f bijektiv, d.h. injektiv und surjektiv.

Der Beweis (Forts)

Wissen: Es gibt $g : \mathbb{N} \to A$ surjektiv. Z.Z: Es gibt ein $f : \mathbb{N} \to A$ bijektiv.
Haben def $f(n) := g(\ell)$ mit ℓ minimal mit $g(\ell) \notin \{f(0), f(1), \ldots, f(n-1)\}$.

- f ist injektiv: Zu zeigen: $m \neq n$ impliziert $f(m) \neq f(n)$.

  Beweis: Wir können annehmen dass $m < n$. Als wir $f(n)$ definiert haben, haben wir verlangt dass $f(n) = g(\ell) \notin \{f(0), \ldots, f(n-1)\}$, d.h. insbesondere dass $f(n) \neq f(m)$.

- Es gilt: $G_n := \{g(0), g(1), \ldots, g(n)\} \subseteq \{f(0), f(1), \ldots, f(n)\} =: F_n$.

  Beweis mit Induktion: Es gilt $G_0 = F_0 = \{f(0)\}$.

  Angenommen es gilt $G_n \subseteq F_n$.

  Fall 1: $g(n+1) \notin F_n$. Dann ist $f(n+1) = g(n+1)$ und $G_{n+1} = G_n \cup \{g(n+1)\} \subseteq F_n \cup \{g(n+1)\} = F_{n+1}$.

  Fall 2: $g(n+1) \in F_n$. $G_{n+1} = G_n \cup \{g(n+1)\} \subseteq F_n \subseteq F_{n+1}$.

- f ist surjektiv, d.h. $f''\mathbb{N} = A$.

  Sei $a \in A$. Dann $a = g(n)$ für irgendein $n \in \mathbb{N}$, weil g ja A abzählt. Daher ist $a \in G_n \subseteq F_n$. □

Unendlichkeiten

Definition

Wir sagen A und B haben dieselbe Kardinalität ("sind gleich gross", $|A| = |B|$) wenn es ein $f : A \to B$ bijektiv gibt.
Für endliche Mengen mit n Elementen schreibe auch $|A| = n$.

(Was entspricht das im Endlichen?)

Haben gesehen:

- A abzählbar unendlich $\leftrightarrow |A| = |\mathbb{N}|$
- $|A| = |\mathbb{N}| \wedge B \subseteq A \wedge B$ unendlich $\to |B| = |\mathbb{N}|$
- $|\mathbb{Q}| = |\mathbb{N}|$.
- Insbesondere:
  $|\mathbb{N}| = |\mathbb{Z}| = |\mathbb{Q}| = |\{p \in \mathbb{N} : p \text{ gerade}\}| = |\{p \in \mathbb{N} : p \text{ ungerade}\}| = \ldots$

Es gibt aber Mengen die noch größer sind!
D.h.: es gibt verschiedene Unendlichkeiten:

Überabzählbarkeit von $\mathbb{R}$

Theorem (Cantor)

$\mathbb{R}$ ist nicht abzählbar ($\mathbb{R}$ ist "überabzählbar").

Beweis: Angenommen $\mathbb{R} \cap [0,1]$ wäre abzählbar als $\{r^1, r^2, \dots\}$.
Sei $r^\ell = 0.r_1^\ell r_2^\ell r_3^\ell \dots$ dezimal.

$$
\begin{array}{lcl}
r^1 & = & 0.\ 1\ 7\ 6 \quad \cdots \quad 1 \mapsto 3 \\
r^2 & = & 0.\ 5\ 0\ 0 \quad \cdots \quad 0 \mapsto 2 \\
r^3 & = & 0.\ 1\ 9\ 9 \quad \cdots \quad 9 \mapsto 1 \\
\vdots & & \\
r^\ell & = & 0.\ r_1^\ell\ r_2^\ell\ r_3^\ell\ \cdots\ r_\ell^\ell\ \cdots \quad n \mapsto n+2 \mod 10 \\
d & = & 0.\ 3\ 2\ 1 \quad \cdots
\end{array}
$$

Setze $d = 0.d_1 d_2 d_3 \dots$ dezimal, mit $d_i = r_i^i + 2 \mod 10$.
(Frage: Warum $+2$ und nicht $+1$?)
Dann ist d ("wirklich") ungleich jedem r^ℓ, Widerspruch.

Unlösbarkeit des Halteproblems (kein Prüfungsstoff) 1

Diese Beweismethode heißt "Diagonalisierung". Sie hat alle möglichen Anwendungen, ein Beispiel aus der Informatik:

Theorem

Das Halteproblem ist unlösbar.

Formulieren wir es für Python (Computermodell aber egal):

- Halteproblem: Gegeben ein Python Programm s und einen Input n, hält s auf Input n (oder "hängt" das Programm).
- Wir können strings als natürliche Zahl interpretieren, d.h. $s \in \mathbb{N}$.
- Der Einfachheit halber: Identifiziere jede natürliche Zahl mit einem Sourcecode. D.h.: $1, 2, 3, \ldots$ sind alle möglichen Python Programme.
- Halteproblem ist also Funktion, H die dem Input (s, n) (natürl. Zahlen) den Output `True` (hält) oder `False` (hält nicht) zuordnet.
- "Halteproblem ist unlösbar" ist dann der folgende Satz: H ist nicht (in Python) programmierbar.

Unlösbarkeit des Halteproblems (kein Prüfungsstoff) 2

- Sei U ein Python-Interpreter (in Python):
  Der Input: Python-Sourcecode $s \in \mathbb{N}$, und $n \in \mathbb{N}$.
  Der Output: Was s auf Input n ausgeben würde: "$U(s,n) = s(n)$".
- Wir können nun diagonalisieren: $f(n) := U(n,n) + 1$.
  Diese Funktion läßt sich in Python programmieren:
  Wir simulieren $U(n,n)$, addierern 1 zum Ergebnis, und geben das aus.
- f hat also einen sourceode m, und es gilt: $m(m) = U(m,m) + 1$.
- Aber $U(m,m) = m(m)$. Widerspuch??
- Nein! Programm terminiert i.A. nicht für jeden Input. Es gilt also einfach $U(m,m) = m(m) = \text{undefiniert} = U(m,m) + 1$.
- Das zeigt: Python-Halteproblem in Python unlösbar. Ansonsten wäre

$$g(n) := \begin{cases} U(n,n)+1 & \text{wenn } U(n,n) \text{ terminiert} \\ 27 & \text{sonst} \end{cases}$$

(in Python) programmierbar, nun wirklich Widerspruch!

Gruppen

Notation: Funktionen mit mehreren Inputs

- Bsp: Addition: $x, y \mapsto x + y$
- In der Mathematik wird üblicherweise keine eigene Notation/Definition dafür eingeführt:
- "Zwei inputs": $f(x, y)$ mit $x \in X$ und $y \in Y$ und Wertebereich Z ist äquivalent zu einem Input "geordnetes Paar (x, y)"; d.h. $f : X \times Y \to Z$.
- Wir schreiben aber trotzdem $f(x, y)$ und nicht etwa $f(\langle x, y \rangle)$.
- Analog $f : X_1 \times \cdots \times X_n \to Z$ und $f(x_1, \ldots, x_n)$.
- Die Addition $+$ ist also eine Funktion $+ : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$. (Eine andere Funktion bekommt man für $\mathbb{Z}$, etc.)
- Für zweistellige Funktionen wird oft $a\, f\, b$ statt $f(a, b)$ geschrieben wenn das Funktionssymbol z.B. $+$, $\cdot$, $\circ$, $\times$, $\otimes$, ... ist.

Halbgruppen

Definition

Eine Halbgruppe ist ein Paar $(A, \circ)$ mit $A \neq \emptyset$ und einer Funktion $\circ : A \times A \to A$ für die das Assoziativgesetz gilt:
$(\forall a, b, c \in A)\,(a \circ b) \circ c = a \circ (b \circ c)$

Beispiele: $(\mathbb{N}, +)$, $(\mathbb{Z}, +)$, $(\mathbb{Q}, +)$, $(\mathbb{R}, +)$, $(\mathbb{C}, +)$,
$(\mathbb{N}, \cdot)$, $(\mathbb{Z}, \cdot)$, $(\mathbb{Q}, \cdot)$, $(\mathbb{R}, \cdot)$, $(\mathbb{C}, \cdot)$.

Keine Halbgruppen sind z.B.:
$(\mathbb{N}, x^y)$ (nicht assoziativ),
$(\mathbb{Z}, -)$ (auch nicht assoziativ);
$(\mathbb{R}, /)$ schon deshalb weil die Division ja nicht auf ganz $\mathbb{R}^2$ definiert ist.

Zusätzliche Features

Definition

Sei $(A, \circ)$ eine Halbgruppe.

- $(A, \circ)$ ist kommutativ, wenn gilt: $(\forall a, b \in A)\, a \circ b = b \circ a$
- $(A, \circ)$ hat ein neutrales Element $e \in A$, wenn $(\forall a \in A) e \circ a = a \circ e = a$.
- Wenn e ein neutrales Element ist, dann nennt man b Inverses von a wenn $a \circ b = b \circ a = e$.

Welche der Beispiele sind kommutativ, assoziativ, haben neutrale Element, und welche Elemente haben Inverse?

Konventionen

- Wenn $\cdot$ oder $\circ$ als Symbol für die Funktion verwendet wird, dann wird ein Inverses von a oft mit a^{-1} bezeichnet, und das neutrale Element e manchmal mit Einselement oder 1.
- Analog: Bei $+$ heißt das Inverse $-a$ und das neutrale Element e manchmal Nullelement oder 0.
- Das Symbol $+$ verwendet man üblicherweise nur für kommutative Halbgruppen.

Eindeutigkeit von neutralem Element

Theorem

Das neutrale Element ist eindeutig.

Damit ist eine der folgenden äquivalenten Formulierungen gemeint:
Sei $(A, \circ)$ eine Halbgruppe.

- Das neutrale Element von $(A, \circ)$ (wenn es eines gibt) ist eindeutig.
- $(A, \circ)$ hat höchstens ein neutrale Element.
- $(\forall e_1, e_2)\big(\forall a(a \circ e_1 = e_1 \circ a = a \wedge a \circ e_2 = e_2 \circ a = a) \rightarrow e_1 = e_2\big)$

Beweis.

Sei e_1, e_2 neutral. Dann ist $e_1 \circ e_2 = e_1$ (weil e_2 neutral) und $e_1 \circ e_2 = e_2$ (weil e_1 neutral), d.h. $e_1 \circ e_2 = e_1 = e_2$. □

Man spricht daher von "dem" (nicht "einem") neutralen Element.

Eindeutigkeit von inversem Element

Definition

Eine Halbgruppe mit neutralem Element e heißt Moniod.

Theorem

Inverse (in einem Monoid) sind eindeutig.

Damit ist eine der folgenden äquivalenten Formulierungen gemeint:
Sei $(A, \circ)$ ein Monoid, e das neutralem Element, und $a \in A$.

- Das Inverse von a (wenn es eines gibt) ist eindeutig.
- a hat höchstens ein Inverses.
- $(\forall b_1, b_2)\big(b_1 \circ a = a \circ b_1 = e \wedge b_2 \circ a = a \circ b_2 = e\big) \to b_1 = b_2\big)$

Beweis.

Seien b_1 und b_2 Inverse von a. Dann
$b_1 = b_1 \circ e = b_1 \circ (a \circ b_2) = (b_1 \circ a) \circ b_2 = b_2$. □

Gruppen

Definition (wichtig)

Eine Gruppe ist ein Monoid in dem es für jedes Element ein Inverses gibt. Kommutative Gruppen heißen auch abelsche Gruppen.

Beispiele:

- $(\{1, 2, \dots\}, +)$ ist eine (kommutative) Halbgruppe, aber kein Monoid.
- $(\mathbb{N}, +)$ ist ein Monoid, aber keine Gruppe.
- $(\mathbb{Z}, +)$, $(\mathbb{Q}, +)$, $(\mathbb{R}, +)$ und $(\mathbb{C}, +)$ sind (abelsche) Gruppen.
- $(\mathbb{Z}, \cdot)$, $\dots$, $(\mathbb{C}, \cdot)$ (kommutative) Monoide aber keine Gruppen (0 hat kein multiplikatives Inverses.)
- $(\mathbb{Z} \setminus \{0\}, \cdot)$ ist keine Gruppe, aber $(\mathbb{Q} \setminus \{0\}, \cdot)$, $(\mathbb{R}, \setminus\{0\}, \cdot)$ und $(\mathbb{C} \setminus \{0\}, \cdot)$ schon.

Rechnen und Kürzen

In Gruppen gilt:

- Wenn $a \circ b = c$, dann ist $b = a^{-1} \circ c$ (weil $a^{-1} \circ (a \circ b) = b$).
  Analog ist dann $a = c \circ b^{-1}$.
- Man kann also "kürzen":
  Wenn $a \circ b = a \circ c$, dann ist $b = c$.
  Analog andere Seite: Wenn $b \circ a = c \circ a$, dann ist $b = c$.
- Insbesondere: Wenn $a \circ b = a$ dann ist $b = e$.
  Äquivalent: Wenn $b \neq e$, dann $a \circ b \neq a$. Nur das Einselement läßt a unverändert.
  Analog: Wenn $b \circ a = a$ dann ist $b = e$.
- Wir definieren a^z für $z \in \mathbb{Z}$:
  $a^0 := e$, für $n > 0$ ist $a^n := \underbrace{a \circ \cdots \circ a}_{n \text{ mal}}$, und $a^{-n} := (a^{-1})^n$.

  Dann gilt: $a^n \circ a^m = a^{n+m}$ für n, m in $\mathbb{Z}$.
  Beweis: Fall: $n > 0$ und $m = -k < 0$: Kürze inneres $a \circ a^{-1}$ in $\underbrace{a \circ \cdots a}_{n \text{ mal}} \circ \underbrace{a^{-1} \circ \cdots a^{-1}}_{k \text{ mal}}$ solange bis $\underbrace{a \circ \cdots a}_{n-k \text{ mal}}$ oder $\underbrace{a^{-1} \circ \cdots a^{-1}}_{k-n \text{ mal}}$ übrigbleibt.

Symmetrische Gruppen

Ein wichtiges Beispiel: Fixiere eine Grundmenge A.

Definition

Eine Permutation von A ist ein $f : A \to A$ bijektiv.

- Die Identität Id ist eine Permutation.
- Die Verknüpfung $g \circ f$ zweier Permutationen f, g ist eine Permutation.
- Die Verknüpfung von Funktionen ist assoziativ.
- $f \circ \mathrm{Id} = \mathrm{Id} \circ f = f$.
- Eine Permutation f hat eine Inverse g (d.h., $f \circ g = g \circ f = \mathrm{Id}$).

Theorem

Die Permutationen mit der Verknüpfung bilden eine Gruppe, genannt Symmetrische Gruppe.

Symmetriegruppen

Definition

S_n ist die symmetrische Gruppe von $\{1, 2, \ldots, n\}$.

Sollte bekannt sein: $|S_n| = n! := 1 \cdot 2 \cdots (n-1) \cdot n$.
Im Folgenden verwenden wir als Beispiel S_3.
Wir notieren $f \in S_3$ als $\begin{pmatrix} 1 & 2 & 3 \\ f(1) & f(2) & f(3) \end{pmatrix}$

Sei $f_1 := \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$ und $f_2 := \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}$.

Dann ist $f_2 \circ f_1 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$ und ungleich zu $f_1 \circ f_2 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$.

Theorem

S_n ist nicht kommutativ für $n \geq 3$.

Beweis: Wir können in S_n die obigen f_i verwenden, wobei wir $f_i(m) := m$ setzen für $m > 3$.

Untergruppen

Definition

Sei $(G, \circ)$ Gruppe, und $U \subseteq G$.
U ist Untergruppe von G, oder: $U \leq G$, wenn $(U, \circ)$ Gruppe ist.

Bem.: Formal korrekter wäre: $(U, \circ \restriction U^2)$.

Es gilt: $U \leq G$ gdw:

1. U ist unter $\circ$ abgeschlossen.
2. Wenn $b \in U$ dann auch b^{-1} (und damit auch $b \cdot b^{-1} = e$).

Beweis:
Sei $U \leq G$. Sei e neutrales Element von G und e von U. Dann $e = e'$: (Sei $b \in U$. Dann $b \circ e' = b \rightarrow e' = e$.) Damit ist auch das Inverse-in-U von $b \in U$ gleich dem Inversen-in-G.
Sei nun U abgeschlossen. Z.z.: U ist Gruppe. Assoziativgesetz ist klar. Sei $b \in U$, dann nach Voraussetzung b^{-1} und $b \circ b^{-1} = e$. Dann ist e auch (ein, daher das) neutrale Element von U, und b^{-1} ist auch in U das Inverse von b.

Verknüpfung von Bijektionen ist Bijektion

- Wenn $f : A \to B$ injektiv und $g : B \to C$ injektiv, dann ist $g \circ f : A \to C$ injektiv.

  Beweis: $a_1 \neq a_2 \xrightarrow[\text{injektiv}]{\text{weil } f} f(a_1) \neq f(a_2) \xrightarrow[\text{injektiv}]{\text{weil } g} g(f(a_1)) \neq g(f(a_2))$.

- Wenn $f : A \to B$ surjektiv und $g : B \to C$ surjektiv, dann ist $g \circ f : A \to C$ surjektiv.

  Beweis: Sei $c \in C$. Dann gibt es ein $b \in B$ mit $g(b) = c$, weil g surjektiv. Für dieses b gibt es ein $a \in A$ mit $g(a) = b$, weil g surjektiv. Dann gilt: $g \circ f(a) = g(f(a)) = g(b) = c$.

- Wenn $f : A \to B$ bijektiv und $g : B \to C$ bijektiv, dann ist $g \circ f : A \to C$ bijektiv.

  Beweis: bijektiv ist äquivalent zu: (injektiv und surjektiv).

Assoziativität und neutrales Element

- Die Verknüpfung ist assoziativ: Sei $f : A \to B$, $g : B \to C$ und $h : C \to D$. Dann sind $(h \circ g) \circ f$ und $h \circ (g \circ f)$ dieselbe Funktion $A \to D$.

  Beweis: $\big((h \circ g) \circ f\big)(a) = (h \circ g)(f(a)) = h(g(f(a))$, und $\big(h \circ (g \circ f)\big)(a) = h\big(g \circ f)(a)\big) = h(g(f(a))$.

Definition

Die "Identität auf A", bezeichnet mit Id_A, ist die Funktion $\mathrm{Id}_A : A \to A$ definiert durch $a \mapsto a$.

(Oft nur Id geschrieben.)
Es gilt:

- Id ist neutrales Element der Verknüpfung:
  Sei $f : A \to B$. Dann $\mathrm{Id}_B \circ f = f \circ \mathrm{Id}_A = f$.
  Beweis: $(\mathrm{Id}_B \circ f)(a) = \mathrm{Id}_B(f(a)) = f(a)$.
  $(f \circ \mathrm{Id}_A)(a) = f(\mathrm{Id}_A(a)) = f(a)$.

Inverse Funktionen

- Wenn $f : A \to B$ bijektiv, dann gibt es eine Umkehrfunktion (Inverses) $g : B \to A$ mit:
  $g(f(a)) = a$ für alle $a \in A$, und $f(g(b)) = b$ für alle $b \in B$.

  Beweis: Sei $b \in B$. Wir definieren $g(b)$: Es gibt (mindestens) ein $a \in A$ mit $f(a) = b$ (weil f surjektiv), und es gibt höchstens eines (weil f injektiv). Setze $g(b) := a$.
  Sei $a \in A$. Dann ist $g(f(a)) = a$: Sei $b := f(a)$. Dann ist $g(b) = a$ nach Definition. Insbesondere ist g also surjektiv. g ist injektiv: Wenn $b_1 \neq b_2$, und $f(a_i) = b_i$, dann ist $g(b_1) = a_1 \neq a_2 = b(b_2)$; ansonsten wäre ja $b_1 = f(a_1) = f(a_2) = b_2$.
- Bemerkung: Wenn $f : A \to B$ injektiv, dann ist $f : A \to f''A$ bijektiv, daher gibt es ein Inverses $g : f''A \to A$.

Wiederholung: Die symmetrische Gruppe

Definition

- Eine Permutation von A ist ein $f : A \to A$ bijektiv.
- Die symmetrische Gruppe $S(A)$ ist die Gruppe der Permutationen mit der Verknüpfung als Gruppenoperation.

Diese Definition beinhaltet das implizite Theorem: $(S(A), \circ)$ ist eine Gruppe. Das können wir nun leicht überprüfen:

- $\circ$ ist tatsächlich eine Funktion $S(A) \times S(A) \to S(A)$, weil die Verknüpfung von Bijektionen eine Bijektion ist.
- Die Verknüpfung $\circ$ ist assoziativ.
- Id ist neutrales Element.
- Für $f \in S(A)$ erfüllt die inverse Funktion $g \in S(A)$ dass $f \circ g = g \circ f = \mathrm{Id}$.

Wiederholung: S_n

$S_n = S(\{1, 2, \ldots, n\})$ sind die Permutationen von $\{1, 2, \ldots, n\}$.
Wir sehen uns konkret S_3 an:

- Ein Element von S_3 ist eine Bijektion von $A := \{1, 2, 3\}$ nach A.
- Z.B.: a definiert durch: $a(1) := 2$, $a(2) := 3$, $a(3) := 1$.
  b definiert durch: $b(1) := 2$, $b(2) := 1$, $b(3) := 3$.
- Wird auch so geschrieben: $a = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$; $\quad b = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$
- Eine Permutation (Bijektion) ist äquivalent zu einer "Umordnung" der Folge $(1, 2, 3)$: a entspricht der Folge/Anordnung 231. Es gibt 6

  Permutationen, "lexikographisch" geordnet sind das:
  123(=e), 132, 213(=b), 231(=a), 312, 321.
- Verknüpfung: $b \circ a = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \quad \neq \quad a \circ b = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$
- D.h.: S_3 ist nicht kommutativ.

Wiederholung: Untergruppen

Definition

Sei $(G, \circ)$ Gruppe, und $U \subseteq G$.
U ist Untergruppe von G, oder: $U \leq G$, wenn $(U, \circ)$ Gruppe ist.

Theorem

$U \leq G$ gdw.: U ist abgeschlossen unter $\circ$ und $^{-1}$.

(D.h.: Wenn a, b in U, dann auch $a \circ b$ und a^{-1}.)

Bsp.:

- $(\mathbb{Z}, +)$ ist Untergruppe von $(\mathbb{Q}, +)$
- $(\mathbb{Q} \setminus \{0\}, \cdot)$ ist Untergruppe von $(\mathbb{C} \setminus \{0\}, \cdot)$
- etc.

Zyklische Untergruppe

Definition

Sei G eine Gruppe, $a \in G$. Dann ist $\langle a \rangle := \{a^z : z \in \mathbb{Z}\}$.

Es gilt:

- $\langle a \rangle$ ist eine Untergruppe von G.
  Beweis: Wir haben bereits gesehen dass $a^k \circ a^\ell = a^{k+\ell}$ für k, ℓ in $\mathbb{Z}$; insbesondere ist $(a^\ell)^{-1} = a^{-\ell}$.
- $\langle a \rangle$ ist die kleinste Untergruppe von G die a enthält.
  Beweis: Jede Untergruppe die a enthält muss auch alle a^z enthalten.
- $\langle a \rangle$ ist kommutativ (selbst wenn G es nicht ist).

Solche Gruppen heißen "zyklische Untergruppen."

Ein Beispiel in S_3

- Sei $a := \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$. Dann ist $a^2 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$ und $a^3 = e$.
- $e = a^{-3} \circ a^3 = a^{-3} \circ e = a^{-3}$.
- Daraus folgt: $a^{3k} = e$ für $k \in \mathbb{Z}$.
  Beweis: Für $k > 0$ $a^{3k} = a^3 \circ \cdots \circ a^3 = e \circ \cdots e = e$; für $k < 0$ verwende a^{-3} statt a^3.
- Daher $a^{3k+\ell} = a^{3k} \circ a^\ell = a^\ell$.
- In anderern Worten: $a^n = a^m$ wenn $n \equiv m(3)$.
- Daher ist die Untergruppe $\langle a \rangle := \{a^z : z \in \mathbb{Z}\}$ gleich $\{e, a, a^2\}$.
- Beachte:
  $|S_3| = 3! = 6$, und $|\langle a \rangle| = 3$.
- Analog erhalten wir für $b = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$: $b^2 = e$, und $\langle b \rangle = \{e, b\}$.
  In dem Fall ist $|S_3| = 3! = 6$, und $|\langle b \rangle| = 3$

Beispiel in $\mathbb{Z}$

Zur Erinnerung: In additiver Schreibweise $(G, \oplus)$ schreibt man $z\,a$ statt a^z für $a \oplus \cdots \oplus a$; $-a$ statt a^{-1}, und $\langle a \rangle := \{z\,a : z \in \mathbb{Z}\}$.

- In $(\mathbb{Z}, +)$ ist $n\,a = a + \cdots + a$ gleich $n \cdot a$ (Multiplikation).
- D.h. $\langle m \rangle = \{z \cdot m : z \in \mathbb{Z}\}$.
- Das wird naheliegenderweise auch $m\mathbb{Z}$ geschrieben, und ist gleich $\{k \in \mathbb{Z} : m|k\}$.
- In dem Fall ist also $|\langle m \rangle|$ (abzählbar) unendlich.

Nebenklassen

Definition

Sei $(G, \circ)$ Gruppe und $U \leq G$, $a \in G$.
$a \circ U := \{a \circ b : b \in U\} \subseteq G$ heißt Linksnebenklasse, und
$U \circ a := \{b \circ a : b \in U\} \subseteq G$ heißt Rechtsnebenklasse.

Es gilt:

- In abelschen Gruppen sind Links- und Rechtsnebenklassen dasselbe.
- Wenn $b \in U$, dann ist $b \circ U = U \circ b = U$.

  Beweis: U ist unter $\circ$ abgeschlossen, daher $b \circ U \subseteq U$.
  Wenn $b' \in U$, dann $b \circ (b^{-1} \circ b') \in b \circ U$ (weil U auch unter $^{-1}$ abgeschlossen), und daher $b \circ U \supseteq U$.
  Wir haben damit gezeigt: $b \circ U \subseteq U$ und $U \subseteq b \circ U$, und daraus folgt $U = b \circ U$. (Analog für $U \circ b$.) □

Nebenklassen (Forts.)

Theorem

Für $U < G$ ist G die disjunkte Vereinigung von Linksnebenklassen. (Analog Rechtsnebenklassen.)

Damit meinen wir:

1. Jedes $a \in G$ ist in einer Nebenklasse $b \circ U$, d.h. die Vereinigung der Nebenklassen sind ganz G.
2. Wenn $a_1 \circ U \neq a_2 \circ U$, dann ist $a_1 \circ U \cap a_2 \circ U = \emptyset$, d.h. zwei (verschiedene) Nebenklassen sind disjunkt.

Beweis: Der erste Punkt ist klar: $a \in a \circ U$, weil $e \in U$ und $a = a \circ e$.
Zweiter Punkt: Angenommen $c \in a_1 \circ U \cap a_2 \circ U$. D.h.,
$c = a_1 \circ b_1 = a_2 \circ b_2$ mit b_1, b_2 in U. Dann $a_1 = a_2 \circ b_2 \circ b_1^{-1} \in a_2 \circ U$.
Daraus folgt $a_1 \circ U \subseteq a_2 \circ U$. Analog bekommen wir $a_2 \circ U \subseteq a_1 \circ U$ und daher $a_1 \circ U = a_2 \circ U$. □

Nebenklassen: Beispiele in S_3

Sei $a := \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$ und $b := \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$.

Für $U := \langle a \rangle$:

- Eine (triviale) Nebenklasse (NK) ist U selbst; und $U = c \circ U = U \circ c$ für jedes beliebige $c \in U$.
- $b \notin U$, daher ist $b \circ U$ disjunkt zu U, und besteht aus den drei Elementen $\{b, a \circ b, a^2 \circ b\}$.
- $S_3 = e \circ U \,\dot\cup\, b \circ U$. ($\dot\cup$: disjunkte Vereinigung)
- Es gibt nur zwei Linksnebenklassen (LNK), mit je 3 Elementen.
- Die LNK sind genau die Rechtsnebenklassen (RNK): $U \circ b = b \circ U$. (Es ist zwar $a \circ b \neq b \circ a$, aber $a \circ b = b \circ a^2$.)

For $U' := \langle b \rangle$

- Es gibt drei LNK mit je 2 Elementen: $U' = e \circ U'$, $a \circ U'$ und $a^2 \circ U'$.
- Es gibt drei RNK mit je 2 Elementen: $U = U' \circ e$, $U' \circ a$ und $U' \circ a^2$.
- RNK und LNK sind verschieden. Nur U' ist sowohl LNK als auch RNK.

Nebenklassen: Beispiele in $\mathbb{Z}$

Sei $U := n\mathbb{Z}$, d.h. die Untergruppe $\langle n \rangle$.

Zur Erinnerung: Nun additive Schreibweise. In multiplikativer Schreibweise sähe $n\mathbb{Z}$ aus wie eine Nebenklasse; es ist aber hier die Untergruppe gemeint, Nebenklassen werden mit nun mit $\ell + U$ notiert:

- $\ell + U = \{\ldots, \ell - 2n, \ell - n, \ell, \ell + n, \ell + 2n, \ldots\}$.
- Das sind genau die $m \in \mathbb{Z}$ die bei Division mit n den Rest ℓ haben.
  Das schreibt man $m \equiv \ell\,(n)$
  (oder, in vielen Programmiersprachen, `m % n == l`).
  D.h.: $\ell + U = \{m \in \mathbb{Z} : \ m \equiv \ell\,(n)\}$.
- RNK und LNK sind trivialerweise gleich, weil $(\mathbb{Z}, +)$ abelsch ist.
- $\mathbb{Z}$ zerfällt also in n viele Nebenklassen.

Satz von Lagrange

Theorem

Sei G eine endliche Gruppe, $U \leq G$.

- *Größe $|G|$ der Gruppe heißt Ordnung der Gruppe.*
- *Anzahl der Linksnebenklassen heißt Index, $|G : U|$.*
- *Es gilt: $|U|$ teilt $|G|$, und $\frac{|G|}{|U|} = |G : U|$*
- *(Analog: Anzahl der RNK ist ebenfalls frac$|G||U| = |G : U|$.)*

Beweis:

- Wir wissen bereits dass G die disjunkte Vereinigung der $|G : U|$ vielen Nebenklassen ist. Es reicht zu zeigen dass jede Nebenklasse die Größe $|U|$ hat. (Dann ist $|G| = |U| \cdot |G : U|$ und wir sind fertig.)
- Nach Definition ist die Nebenklasser $a \circ U = \{a \circ b : \ b \in U\}$. Wenn $b \neq b'$ in U, dann ist $a \circ b \neq a \circ b'$. Daher ist $|a \circ U| = |U|$.
- Derselbe Beweis funktioniert für Rechtsnebenklassen.

Satz von Lagrange: Bsp

In unseren Beispielen:

- S_3 hat $3! = 6$ Elemente $\rightarrow$ Untergruppen können 1, 2, 3, oder 6 Element haben.
- (Nur) die triviale Untergruppen $\{e\}$ hat Ordnung 1 (und 6 LNK).
- (Nur) S_3 selbst hat Ordnung 6 (und 1 LNK).
- $\langle a \rangle$ hat Ordnung 3 und 2 LNK.
- $\langle b \rangle$ hat Ordnung 2 und 3 LNK.

Notation: Minimum und Maximum

- Wen A eine Menge von Zahlen (oder anderer geordneter Objekte ist), dann ist $\min(A)$ das kleinste Element von A.
- Das muss natürlich nicht immer existieren: $\min(\mathbb{Z})$ und $\min\{q \in \mathbb{Q} : q > 0\}$ existieren nicht.
- Wenn $A \subseteq \mathbb{N}$ nicht-leer ist, dann existiert $\min(A)$ (das ist eine Form des Induktionsprinzips).
- $\max(A)$ ist das größte Element von A.
- Auch hier gilt: $\max(A)$ nuss nicht existieren.
- Auch ein $A \subseteq \mathbb{N}$ hat i.A. kein maximales Element. (Es gibt eine kleinste Primzahl, aber keine größte, etc.)
  $A \subseteq \mathbb{N}$ hat ein maximales Element gdw A endlich ist.

Ordnung von Elementen (Forts.)

Definition (Ordnung eines Gruppenelements)

Die Ordnung von $a \in G$, $\mathrm{ord}_G(a)$, ist $|\langle a \rangle|$ (falls endlich, sonst ∞).

Theorem

$\mathrm{ord}_G(a) = \min\{n > 0 : a^n = e\}$ *(falls eine der beiden Terme endlich ist).*

Beweis: Sei $k := \min\{n > 0 : a^n = e\}$ endlich. Dann ist $\langle a \rangle = \{e, a, a^2, \ldots a^{k-1}\}$, weil $a^{k-1} = a^{-1}$ und $a^{\ell k+m} = a^m$. Es bleibt zu zeigen dass all diese Element verschieden sind: Wenn $a^n = a^m$ für irgendein $0 \leq n < m < k$, dann ist $a^{n-m} = e$, und dann wäre k nicht minimal. Daraus folgt: $k = |\langle a \rangle|$.

Sei umgekehrt $m := |\langle a \rangle|$ endlich. Dann muss es $n \neq m$ geben mit $a^n = a^m$, d.h., $a^{n-m} = e$, daher ist $k := \min\{n > 0 : a^n = e\}$ endlich; und aus dem vorigen Argument folgt $k = m$. □

Ordnung von Elementen

Corollary

Sei G eine endliche Gruppe ist, $a \in G$:

- $\mathrm{ord}_G(a)$ *teilt* $|G|$.
- *Insbesondere ist* $a^{|G|} = e$. *(Kleiner Satz von Fermat.)*

Beweis: $\langle a \rangle$ ist Untergruppe, daher gilt: $k := |\langle a \rangle|$ teilt $|G|$; $k \cdot \ell = G$.
Außerdem ist $a^k = e$, und daher ist $a^{|G|} = a^{k \cdot \ell} = e^\ell = e$. □

Sylow

Wir haben gesehen: $U < G$ impliziert $|U|$ teilt $|G|$.
Umgekehrt gilt (ohne Beweis):

Theorem (Einfache Form der Sylow Sätze)

Wenn eine Primzahl p die Gruppenordnung $|G|$ teilt, dann gibt es eine Untergruppe U mit $|U| = p$.

Das gilt sogar für Primzahlpotenzen p^ℓ, aber nicht für allgemeine Teiler von $|G|$.

Erzeugende von Gruppen

- Wir haben gesehen: $\langle a \rangle$ ist kleinste Untergruppe von G die a enthält.
- Allgemein: Wenn $A \subseteq G$, dann definiere $\langle A \rangle$ als die Menge aller $g \in G$ von der Form $g = a_1 \circ a_2 \circ \cdots a_n$, wobei jedes a_i entweder aus A ist, oder Inverses eines Elements aus A ist.
- Dann ist $g^{-1} = a_n^{-1} \circ \cdots a_2^{-1} \circ a_1^{-1}$ wieder aus $\langle A \rangle$, und wenn h ebenfalls in $\langle A \rangle$ ist, dann auch $g \circ h$.
- $\langle A \rangle$ ist also eine Untergruppe von G, und es ist die kleinste Untergruppe ist die A enthält (d.h.: die Obermenge von A ist): Jede Untergruppe $U \supseteq A$ muss ja jedes g der obigen Form enthalten.
- Betrachte $\mathcal{F}$, die Menge aller Untergruppen von G.
  Sei $\mathcal{F}^* \subseteq \mathcal{F}$ die Menge derjenigen Untergruppen U, die $U \supseteq A$ erfüllen. Dann gilt: $\langle A \rangle = \bigcap_{U \in \mathcal{F}^*} U$.
  Beweis: $\subseteq$: Jedes $U \in \mathcal{F}^*$ enthält $\langle A \rangle$; $\supseteq$: $\langle A \rangle \in F^*$.
- Wichtiges / häufiges Phänomen in der Mathematik: “Kleinstes . . . das . . . enthält” “von unten” als Erzeuignis und “von oben” als Schnitt.

Bemerkungen

- Gruppen sind wichtig, in Theorie und vielen Anwendungen (Physik, Chemie, . . .).
- Wichtige Gruppen die wir später kennen lernen werden: Matrizengruppen.
- Gruppentheorie ist eine tiefe und schwierige Mathematische Disziplin.
- Auch die Theorie der endlichen Gruppen ist sehr kompliziert.
- Die Theorie der abelschen (=kommutativen) Gruppen ist wesentlich einfacher.
- Noch einfacher ist die Theorie der endlich erzeugten abelschen Gruppen.
  Solche Gruppen kann man sehr einfach Charakterisieren. Mehr dazu eventuell später.

Ringe

Ringe

Definition

$(R, +, \cdot)$ ist ein Ring, wenn gilt:

- $+ : R \times R \to R$ und $\cdot : R \times R \to R$
- $(R, +)$ ist abelsche Gruppe mit neutralem Element (0, "Nullelement")
- $(R, \cdot)$ ist Halbgruppe mit neutralem Element (1, "Einselement").
- Es gelten die Distributivgesetze:
  $a \cdot (b + c) = (a \cdot b) + (a \cdot c)$ und $(b + c) \cdot a = (b \cdot a) + (c \cdot a)$.

R heißt "kommutativ", wenn $(R, \cdot)$ kommutativ ist.

- Ein kommutativer Ring heißt auch "KRE" (komm. Ring mit 1).
- Achtung: Eine andere übliche Definition verlangt nicht Existenz von 1.

Beispiele:

- $(\mathbb{Z}, +, \cdot)$, $(\mathbb{Q}, +, \cdot)$, $(\mathbb{R}, +, \cdot)$, $(\mathbb{C}, +, \cdot)$ sind KREs.
- Kindisches Beispiel: $\{e\}$ ist KRE (mit $0 = 1 = e$).
- Später: Die $n \times n$ Matrizen sind ein (nicht-kommutativer) Ring.

Rechnen in Ringen

Lemma

Sei R ein Ring, $a \in R$. Dann gilt: $a \cdot 0 = 0 \cdot a = 0$.

Beweis:

- Per Definition (Neutrales Element bzgl $+$) gilt: $0 + x = x$ für alle x.
- Insbesondere: $0 = 0 + 0$.
- Daher: $a \cdot 0 = a \cdot (0 + 0)$.
- Aus dem DG folgt: $a \cdot 0 = a \cdot 0 + a \cdot 0$.
- Es gibt $-a$ (mit: $a + (-a) = 0$). Addiere $-a \cdot 0$ zu obiger Gleichung:
- $a \cdot 0 + (-a \cdot 0) = (a \cdot 0 + a \cdot 0) + (-a \cdot 0)$.
- wegen AG $\ldots = a \cdot 0 + (a \cdot 0 + (-a \cdot 0))$,
- Aus $x + y = x + z$ folgt $y = z$ (Kürzen: addiere $-x$ links und rechts.)
- Kürzen von $a \cdot 0$ ergibt: $-a \cdot 0 = a \cdot 0 + (-a \cdot 0)$.
- Nochmals kürzen von $-a \cdot 0$ gibt: $0 = a \cdot 0$.

Analog für $0 = 0 \cdot a$.

Rechnen in Ringen (Forts.)

Lemma

Sei R ein Ring, $a, b \in R$. Dann gilt: $(-a) \cdot b = -(a \cdot b)$ und $b \cdot (-a) = -(b \cdot a)$.

Beweis:

- $0 = 0 \cdot b = (a - a) \cdot b = a \cdot b + (-a) \cdot b$.
- Analog: $0 = b \cdot 0 = b \cdot (a - a) = b \cdot a + b \cdot (-a)$. □

Nullteiler

Definition

a, b im Ring R heißen Nullteiler, wenn gilt: $a \neq 0$ und $b \neq 0$ und $a \cdot b = 0$.

Ein Integritätsbereich ist ein KRE ohne Nullteiler, der $0 \neq 1$ erfüllt.

- $\{0\}$ ist ein nullteilerfreier KRE.
- $(\mathbb{Z}, +, \cdot)$, $(\mathbb{Q}, +, \cdot)$, $(\mathbb{R}, +, \cdot)$, $(\mathbb{C}, +, \cdot)$ sind Integritätsbereiche.

$\mathbb{Z}_m$ (Für $m > 1$.)

- Setzen $\mathbb{Z}_m = \{\overline{0}, \overline{1}, \ldots, \overline{m-1}\}$, mit den Operationen "modulo m:"
- Bsp: in $\mathbb{Z}_{10} = \{\overline{0}, \overline{1}, \ldots, \overline{9}\}$ ist $\overline{6} + \overline{5} = \overline{1}$, und $\overline{2} \cdot \overline{5} = \overline{0}$.
- "Rechnen modulo m" verträgt sich mit Addition und Multiplikation: $\overline{a+b} = \overline{a} + \overline{b}$, und $\overline{a \cdot b} = \overline{a} \cdot \overline{b}$.
  Beweis: $(\ell_1 m) + r_1 + (\ell_2 m) + r_2 = (\ell_1 + \ell_2)m + (r_1 + r_2)$, analog für $\cdot$.
- $\mathbb{Z}_m$ ist KRE mit $0 \neq 1$.
  Beweis: Alle Eigenschaften "vererben" sich vom $\mathbb{Z}$.
  Bsp: KG: $\overline{a} + \overline{b} = \overline{a+b} = \overline{b+a} = \overline{b} + \overline{a}$.
- Wir werden (bald) sehen: $\mathbb{Z}_m$ ist Quotient des KRE $\mathbb{Z}$ mit dem sogenannten "Ideal" $m\mathbb{Z}$, und als socher KRE.

$\mathbb{Z}_m$: Nullteiler

- $\mathbb{Z}_{10}$ hat Nullteiler: $\overline{2} \neq 0$ und $\overline{5} \neq 0$ und $\overline{5} \cdot \overline{10} = 0$.
- Allgemein: $\mathbb{Z}_m$ ist Integritätsbereich gdw m Primzahl.

  Beweis:

  Falls m keine Primzahl dann gibt es $0 < i < m$ und $0 < j < m$ mit $i \cdot j = m$. Dann ist $\overline{i} \neq \overline{0}$ und $\overline{j} \neq \overline{0}$ und $\overline{i} \cdot \overline{j} = \overline{0}$.

  Sei nun m Primzahl und $\overline{i} \cdot \overline{j} = \overline{0}$ mit $i, j > 1$. Das heißt: $ij = \ell m$. D.h. m kommt in der Primfaktorenzerlegung von ij vor. D.h. m ist Primfaktor von i oder von j, damit ist $\overline{i} = \overline{0}$ oder $\overline{j} = \overline{0}$.

Additions- und Multiplikationstabellen

Endliche Ringe (und Gruppen) kann man auch durch ihre Operations-Tabellen festlegen. Z.B.:

$\mathbb{Z}_2$ festgelegt durch:

$+$	$\overline{0}$	$\overline{1}$
$\overline{0}$	$\overline{0}$	$\overline{1}$
$\overline{1}$	$\overline{1}$	$\overline{0}$

$\cdot$	$\overline{0}$	$\overline{1}$
$\overline{0}$	$\overline{0}$	$\overline{0}$
$\overline{1}$	$\overline{0}$	$\overline{1}$

$\mathbb{Z}_4$ festgelegt durch:

$+$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$
$\overline{0}$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$
$\overline{1}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{0}$
$\overline{2}$	$\overline{2}$	$\overline{3}$	$\overline{0}$	$\overline{1}$
$\overline{3}$	$\overline{3}$	$\overline{0}$	$\overline{1}$	$\overline{2}$

$\cdot$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$
$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$
$\overline{1}$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$
$\overline{2}$	$\overline{0}$	$\overline{2}$	$\overline{0}$	$\overline{2}$
$\overline{3}$	$\overline{0}$	$\overline{3}$	$\overline{2}$	$\overline{1}$

- Eine Operation ist kommutativ gdw die Tabellen symmetrisch gegen Spiegelung an der Diagonalen ist. (Was hier der Fall ist.)
- Bei Ringen ist $\cdot 0$, $0\cdot$ und $\cdot 1$, $1\cdot$ sowie $+0$ festgelegt, die entsprechenden Spalten / Reihen enthalten keine zusätzliche Information.

Relationen

- $<$ ist eine sogenannte zweistellige Relation auf (z.B.) $\mathbb{Q}$.
- Formal können wir eine zweistellige Relation auf A als Teilmenge von A^2 definieren. (Und eine n-stellige Relation als Teilmenge von A^n.)
- In dieser Definition ist $<$ auf $\mathbb{Q}$ die Menge $\{(r, s) \in \mathbb{Q}^2 : r < s\}$.
- $<$ auf $\mathbb{N}$ die Menge $\{(r, s) \in \mathbb{N}^2 : r < s\}$, eine andere Relation als die auf $\mathbb{Q}$.
  Vgl. Funktionen: $+$ auf $\mathbb{N}$ ist eine anderer Funktion als (nämlich eine Einschränkung von) $+$ auf $\mathbb{Q}$.
- Wenn $R \subseteq A^2$, dann sagt man "a steht in Relation R zu b, oder: aRb, wenn $(a, b) \in R$.
- Man könnte eine Relation auch als Abbildung $A^2 \to \{\texttt{True}, \texttt{False}\}$ auffassen.

Äquivalenzrelationen

Definition

Sei $R \subseteq A^2$

- R ist symmetrisch, wenn $aRb \rightarrow bRa$ (für alle $a, b \in A$),
- R ist transitiv, wenn $aRb \wedge bRc \rightarrow aRc$,
- R ist reflexiv, wenn aRa (wie immer: für alle $a \in A$).
- R ist Äquivalenzrelation, wenn: symmetrisch, transitiv und reflexiv.

- Bsp: Für fixes $n > 2$ ist $x \equiv y \mod n$ eine Äquivalenzrelation auf $\mathbb{N}$.
- Bsp: $|A| = |B|$ ist Äquivalenzrelation auf allen Mengen.
  (Inverse und Verknüpfung von Bijektionen, und Id, sind Bijektionen.)
- Äquivalenzrelationen bezeichnet man oft mit dem Symbol $\sim$.

Quotienten

- Wenn $\sim$ Äquivalenzrelation auf A ist, und $a \in A$, dann setzen wir $[a]_\sim := \{b \in A : \ b \sim a\}$, genannt die Äquivalenzklasse von a (oft nur $[a]$ geschrieben).
- Der Quotient $A/\sim$ ist die Menge aller Äquivalenzklassen, $\{[a] : \ a \in A\}$.
- Es gilt: A ist die disjunkte Vereinigung der Äquivalenzklassen (d.h., A wird durch die Äquivalenzklassen partitioniert).
- Wichtiges/allgegenwärtiges Konzept in der Mathematik. Dabei ist oft wichtig ob/dass man Struktur (z.B. Gruppenoperation) von A auf $A/\sim$ übertragen kann.
- Klassisches Beispiel: Für $m > 1$, sei $a \sim b$ wenn $a \equiv b \mod m$.
  Die Nebenklasse $[a]$ wird auch $\overline{a}$ geschrieben.
  Der Quotient $\mathbb{Z}/\sim$ ist dann $\mathbb{Z}_m = \{\overline{0}, \ldots, \overline{m-1}\}$.
- In diesem Fall kann man $+$ und $\cdot$ von $\mathbb{Z}$ auf den Quotienten $\mathbb{Z}_m$ übertragen (mehr dazu gleich).

Beispiel: $\mathbb{Q}$

Formale Konstruktion von $\mathbb{Q}$:

- Sei $\mathbb{P} = \{(n, m) : \ n, m \in \mathbb{Z}, m \neq 0\}$.
  (Dabei soll (n, m) der Zahl $\frac{n}{m}$ entsprechen.)
- Definiere $(n, m) \cdot (a, b) := (n \cdot a, m \cdot b)$, und
  $(n, m) + (a, b) := (n \cdot b + a \cdot m, m \cdot b)$.
- Setze $(a, b) \sim (n, m) \ :\leftrightarrow \ a \cdot m = n \cdot b$.
- Das ist eine Äquivalenzrelation.
  Transitivität: Sei $(a, b) \sim (n, m)$ und $(n, m) \sim (x, y)$. Dann ist $am = bn$ und $ny = mx$, daher $amny = bnmx$. Wenn $n \neq 0$ folgt $ay = bx$. Wenn $n = 0$, dann $a = 0$ (weil $m \neq 0$), und genauso $x = 0$, und daher ebenfalls $ay = bx$.
- $\mathbb{Q} := \mathbb{P}/\sim$.
  D.h.: $(n, m) \in [(n, m)]$, $(2n, 2m) \in [(n, m)]$, etc.

Beispiel: $\mathbb{Q}$ (Forst.)

- Es gilt nun(!): $[(n,m)] + [(a,b)] := [(n,m) + (a,b)]$ ist wohldefiniert. D.h.: Wann immer $(n',m') \sim (n,m)$ und $(a',b') \sim (a,b)$, dann ist $(n',m') + (n,m) \sim (a',b') + (a,b)$.
- Analog für $\cdot$.
- Damit ist $+$, $\cdot$ auf $\mathbb{Q}$ definiert.
- Es gilt: $\mathbb{Q}$ Integritätsbereich.
  (Übung.)

Ideale

In KREs sind Äquivalenzrelationen besonders wichtig die von Idealen kommen:

Definition

Sei R ein KRE. $I \subseteq R$, $I \neq \emptyset$ ist ein Ideal, wenn

- $a \in I \wedge b \in I \ \rightarrow \ a + b \in I$.
- $a \in I \wedge b \in R \ \rightarrow \ ab \in I$.

Wir setzen $a \sim_I b$, wenn $a - b \in I$.

- Bemerkung $\langle a \rangle := \{a \cdot b : \ b \in R\}$ ist immer ein Ideal ("Hauptideal").
- Bemerkung: In $\mathbb{Z}$ sind alle Ideale Hauptideale.
  (So einen Ring nennt man Hauptidealring.)

Ideale (Forts.)

Theorem

$\sim_I$ ist eine Äquivalenzrelation auf R, und der Quotient $R/\sim_I$ wird auch R/I geschrieben und ist mit $[a]+[b] := [a+b]$, $[a]\cdot[b] := [a]\cdot b]$ wieder ein KRE.

Beweis:

- I enthält 0 (weil $0 \cdot x$ für ein $x \in I$). Daher ist $a \sim_I a$.
- Wenn $a \in I$ dann auch $-a$ (weil $-a = (-1) \cdot a$.)
  Wenn $a \sim b$, d.h. $a - b \in I$, dann auch $b - a \in I$, d.h. $b \sim a$.
- Wenn $a \sim b$ und $b \sim c$, dann ist $a - b \in I$ und $c - b \in I$, daher $a - b - (c - b) = a - c \in I$ (weil I unter $+$ abgeschlossen), d.h. $a \sim c$.
- $\sim$ ist also Äquivalenzrelation.
- Zu zeigen: $+$ ist wohldefiniert, d.h.: $a \sim a'$ und $b \sim b'$ impliziert $a + b \sim a' + b'$. Das gilt weil $a + b - (a' + b') = a - a' + b - b' \in I$.
- Analog für $\cdot$: $ab - a'b' = a(b - b') + (a - a')b'$. (Hier braucht man dass I abgeschlossen unter Multiplikation mit beliebigen Ringelementen.)

Beweis (Forts.)

- Wir wissen also dass $+$ und $\cdot$ auf $R/\sim$ wohdefiniert sind, durch $[a]+[b]=[a+b]$, analog für $\cdot$.
- Dann sieht man leicht dass dich die KRE Eigenschaften auf den Quotienten "vererben":
- Bsp. Kommutativität der Multiplikation: $[a]\cdot[b] := [a\cdot b] = [b\cdot a] =: [b]\cdot[a]$, etc. □

Bemerkung: Nullteilerfreiheit / Integritätsbereich zu sein vererbt sich i.A. nicht, wie wir gleich sehen werden.

Nochmal $\mathbb{Z}_m$

- $m\mathbb{Z} = \langle m \rangle$ ist ein (Haupt)ideal in $\mathbb{Z}$.
- Der Quotient $\mathbb{Z}_m := \mathbb{Z}/m\mathbb{Z}$ ist daher wieder ein KRE, mit $[x + y] := [x] + [y]$ etc.
- Wenn m keine Primzahl ist, dann ist $\mathbb{Z}_m$ kein Integritätsbereich.
- Bsp: $\overline{2} \cdot \overline{5} = \overline{0}$ in $\mathbb{Z}_{10}$.
- Wenn m Primzahl ist, dann ist $\mathbb{Z}_m$ Integritätsbereich.

  Es gilt dann sogar:
  $\mathbb{Z}_m$ hat zu jedem $x \neq 0$ ein multiplikatives Inverses y. (D.h. $xy = 1$.)

Formale Polynome

- Sei R ein KRE. Ein Polynom $f \in R[X]$ hat die Form $a_0 + a_1 \cdot X + a_2 \cdot X^2 + \cdots + a_n \cdot X^n$.
- Das Produkt von Polynomen ist ein Polynom: $(a_0 + a_1 \cdot X + a_2 \cdot X^2 + \cdots + a_n \cdot X^n) \cdot (b_0 + b_1 \cdot X + b_2 \cdot X^2 + \cdots + b_m \cdot X^m) = a_0 b_0 + (a_1 b_0 + a_0 b_1)X + \cdots + c_i X^i + \cdots + a_n b_m X^{n+m}$, mit $c_i := \sum_{j=0}^{i} a_j b_{i-j}$.
- Wir fassen das nicht als Funktion, sondern als formales Objekt auf.

  Formalere Definition: Ein Polynom ist eine endliche Folge $(a_0, a_1, \ldots, a_n)$ mit Elementen aus R; um Notation zu vereinfachen setzen wir $a_m := 0$ für $m > n$ und schreiben die Folge als (a_i). Wie definieren $(a_i) + (b_i) := (a_i + b_i)$, und $(a_i) \cdot (b_i) = (c_i)$, mit $c_i := \sum_{j=0}^{i} a_j b_{i-j}$.
- Es gilt: Die Polynome bilden einen KRE.

Polynome vs Polynomfunktionen

- Ein Polynom $f(X)$ definiert auf natürliche Weise eine Funktion $R \to R$, $r \mapsto f(r)$, genannt Polynomfunktion.
- Verschiedene Polynome können i.A. dieselbe Polynomfunktion definieren.
- Bsp: Sei $R = \mathbb{Z}_2$. Dann sind $f(X) = X^2 + X$ und $g(X) = 0$ verschiedene Polynome.
  Ihre Polynomfunktionen $R \to R$ sind aber gleich:
  Für alle $x \in R$ (d.h. für $x = \overline{0}$ und $x = \overline{1}$, mehr Element gibt es ja nicht) gilt $f(x) = 0 = g(x)$.

Produkte von algebraischen Strukturen

- Wenn $(G_1, \circ_1)$ und $(G_2, \circ_2)$ Halbgruppen sind, dann ist $(G_1 \times G_2, \circ)$ eine Halbgruppe, mit $(a_1, a_2) \circ (b_1, b_2) := (a_1 \circ_1 b_1, a_2 \circ_2 b_2)$.
  Beweis: AG direkt nachrechnen.
- Analog: Wenn G_1 und G_2 [kommutative] Gruppen, dann auch $G_1 \times G_2$.
- Analog: Wenn R_1 und R_2 [kommutative] Ringe, dann auch $R_1 \times R_2$ (natürlich mit $(a_1, a_2) + (b_1, b_2) := (a_1 +_1 b_1, a_2 +_2 b_2)$ und $(a_1, a_2) \cdot (b_1, b_2) := (a_1 \cdot_1 b_1, a_2 \cdot_2 b_2)$.)
- Wenn R_1 und R_2 nicht-trivial, dann hat $R_1 \times R_2$ Nullteiler:
  Sei $r_i \in R_i$ ungleich Null. Dann $(r_0, 0) \cdot (0, r_1) = (0, 0)$.

  Insbesondere: Produkt von Integritätsbereichen ist kein Integritätsbereich mehr.
- Bsp: $\mathbb{Z}_3 \times \mathbb{Z}$ (nur als Gruppe, oder als Ring), $\mathbb{Z}_2 \times \mathbb{Z}_4$

Euklidischer Algorithmus

Teilbarkeit in Ringen

Die Definition der Teilbarkeit funktioniert in allen kommutativen Ringen R.

- $m|n$ oder "m teilt n" heißt $(\exists \ell \in R)\, n = \ell \cdot m$.
- Es gilt: $m|n$ und $m|k$ impliziert $m|(n \pm k)$;
  $m|n$ und $n|k$ impliziert $m|k$.
- $x \in R$ heißt Einheit, wenn x ein (multiplikatives) Inverses hat.
  Dann gilt: $x|y$ für alle y.
- Bsp:
  - ± 1 sind immer Einheiten.
  - In $\mathbb{Z}$ sind das die einzigen,
    ebenso in $\mathbb{Z}_2$ (d.h. nur $\overline{1}$) oder $\mathbb{Z}_6$ (d.h. nur $\overline{1}$ und $\overline{5}$).
  - In $\mathbb{Z}_5$ sind zusätzlich zu $\overline{1}$ und $-\overline{1} = \overline{4}$ auch $\overline{2}$ und $-\overline{2} = \overline{3}$ Einheiten.
    Z.B.: $\overline{2} \cdot \overline{3} = \overline{1}$.
  - In $\mathbb{Q}[X]$, den rationalen Polynomen, sind die konstanten Polynome $\neq 0$ die Einheiten.

ggTs in Integralbereichen

Der Einfachheit halber nehmen wir nun an dass R Integralbereich ist.
Dann gilt:

- $n|m$ und $m|n$ impliziert dass $m = e \cdot n$ für eine Einheit e.

  Beweis: $n = a \cdot m$ und $m = b \cdot n$, d.h. $n = a \cdot b \cdot n$ bzw. $n(a \cdot b - 1) = 0$. Wegen Integralbereich folgt $a \cdot b = 1$, d.h. a ist Einheit.
- Ein gemeinsame Teiler von n und m ist ein d mit: $d|n$ und $d|m$.
- Ein größter gemeinsamer Teiler (ggT) ist ein gemeinsamer Teiler k so dass $d|k$ für alle gemeinsamen Teiler d.

  Das muß i.A. nicht existieren. (Gegenbsp. nicht völlig trivial.)
- Wenn k ein ggT ist, und e eine Einheit, dann ist auch $e \cdot k$ ein ggT; umgekehrt hat jeder ggT die Form $e \cdot k$.

  Beweis: Wenn k, ℓ zwei ggTs sind, dann $k|\ell$ und $\ell|k$.

ggT in $\mathbb{Z}$

Bis auf weiteres arbeiten wir nun in $\mathbb{Z}$.

- In $\mathbb{Z}$ gilt für $n, m \geq 1$: $n|m$ impliziert $n \leq m$.
- Insbesondere ist $\mathrm{ggT}(n, m) := \max\{k \in \mathbb{N} : \ k|n \wedge k|m\}$ wohldefiniert, mit $1 \leq \mathrm{ggT}(n, m) \leq \min(m, n)$,
  und ist, so wie auch $-\mathrm{ggT}(n, m)$, größter gemeinsamer Teiler.
- Trivialer Algorithmus um $\mathrm{ggT}(n, m)$ zu finden:
  Probiere $1 \leq k \leq \min(m, n)$ durch und bestimme größten Teiler.
- Braucht $\sim \min(m, n)$ Schritte.
- Bemerkung: Alternativer Algorithmus: Bestimme Primfaktorzerlegungen von m, n, etc. Aber alle bekannten Algorithmen zur Primfaktorenzerlegung brauchen ebenfalls sehr lange.
- Besserer Algorithmus: Euklidischer Algorithmus.

Euklidischer Algorithmus: Beispiel

Wir berechnen ggT$(3412, 34)$:

$$\begin{aligned} 3412 &= 100 \cdot 34 + 12 \\ 34 &= 2 \cdot 12 + 10 \\ 12 &= 1 \cdot 10 + 2 \\ 10 &= 5 \cdot 2 \end{aligned}$$

Daraus folgt: ggT$(3412, 34) = 2$:

1. $2|3412$ und $2|34$,
2. Wenn $d|3412$ und $d|34$, dann $d|2$.

Das funktioniert für alle Zahlen:

Euklidischer Algorithmus

Sei $a_0 > a_1 \geq 1$. Wir können ggT(a_0, a_1) wie folgt berechnen:

- $a_0 = \ell_1 \cdot a_1 + a_2$ mit $a_2 < a_1$.
- $a_1 = \ell_2 \cdot a_2 + a_3$ mit $a_3 < a_2$.
- ...
- $a_{n-2} = \ell_{n-1} \cdot a_{n-1} + a_n$ mit $a_n < a_{n-1}$.
- $a_{n-1} = \ell_n \cdot a_n$.

Dann ist $a_n = \text{ggT}(a_0, a_1)$
Beweis:

- $a_n | a_{n-1}$, $a_n | a_{n-2}$, ..., $a_n | a_1$, $a_n | a_0$. D.h. a_n ist gemeinsamer Teiler.
- a_n ist kleinster gemeinsamer Teiler:
  Angenommen $d | a_0$ und $d | a_1$. Zu zeigen: $d | a_n$.
  $d | a_2$, $d | a_3$, ..., $d | a_n$,
- Der Algorithmus bricht nach spätestens a_1 Schritten ab, weil $a_1 > a_2 > \cdots > a_n \geq 1$.

Fibonacci-Zahlen

Es gilt sogar: Wir brauchen nur $\sim \ln(a_0)$ Schritte.

Für diese Abschätzung definieren wir die Fibonacci Zahlen/Reihe F_n und beweisen eine einfache Abschätzung:

- $F_0 := 0$, $F_1 := 1$, und für $n \geq 2$: $F_n := F_{n-2} + F_{n-1}$
- D.h. wir bekommen die Folge der Fibonacci-Zahlen

0	1	2	3	4	5	6	7	8	9	10	11	12	...
0	1	1	2	3	5	8	13	21	34	55	89	144	...

- Es gilt: Für $n \geq 11$ ist $F_n \geq (\frac{3}{2})^n$.

  Beweis: Für $n = 11$ und $n = 12$ nachrechnen.
  Für $n > 12$:
  $F(n) = F(n-2) + F(n-1) \geq (\frac{3}{2})^{n-2} + (\frac{3}{2})^{n-1} = (1 + \frac{3}{2})(\frac{3}{2})^{n-2}$.
  Mit $1 + \frac{3}{2} = \frac{5}{2} > \frac{9}{4} = (\frac{3}{2})^2$ bekommen wir:
  $F(n) > (\frac{3}{2})^2(\frac{3}{2})^{n-2} = (\frac{3}{2})^n$. □

Euklidischer Algorithmus: Laufzeit

Sei $a_0 > a_1 \geq 1$. Wir können ggT(a_0, a_1) als das folgende a_n berechnen:

- $a_0 = \ell_1 \cdot a_1 + a_2$ mit $a_2 < a_1$.
- $a_1 = \ell_2 \cdot a_2 + a_3$ mit $a_3 < a_2$.
- ...
- $a_{n-2} = \ell_{n-1} \cdot a_{n-1} + a_n$ mit $a_n < a_{n-1}$.
- $a_{n-1} = \ell_n \cdot a_n$.

Zur Abschätzung der Laufzeit ($\sim$Anzahl der Schritte, d.h. n):

- $a_n \geq 1 = F_2$
- $a_{n-1} \geq 2 = F_3$ (weil $\ell_n \geq 2$, weil $a_n < a_{n-1}$).
- $a_{n-2} \geq a_{n-1} + a_n \geq 2 + 1 = 3 = F_4$ (weil $\ell_{n-1} \geq 1$, weil $a_{n-1} < a_{n-2}$).
- Allgemein: $a_{n-\ell} \geq F_{\ell+2}$.
  Beweis: $a_{n-\ell} \geq a_{n-(\ell-1)} + a_{n-(\ell-2)} \geq F_{\ell+1} + F_\ell = F_{\ell+2}$
- Insbesondere: $a_0 \geq F_{n+2}$.

Euklidischer Algorithmus: Laufzeit

ggT(a_0, a_1), $a_0 > a_1 > 1$.
Zur Abschätzung der Laufzeit ($\sim$Anzahl der Schritte, d.h. n):

- Haben gesehen: $a_0 \geq F_{n+2}$.
- Wir setzen voraus $n \geq 9$. Dann ist $F_{n+2} > (\frac{3}{2})^{n+2}$.
- Es ist also $a_0 > (\frac{3}{2})^n$, d.h. $\ln_{\frac{3}{2}}(a_0) > n$, d.h.
  $n < \frac{\ln(a_0)}{\ln(\frac{3}{2})} \sim 2.47 \ln(a_0) \sim 1.7 \ln_2(a_0) \sim 5.7 \ln_{10}(a_0)$.
- Wir benötigen also $\sim \ln(a_0)$ viele Schritte.
  Das ist $\sim b$, wobei $b = \ln_2(a_0)$ die Größe von a_0 in Bits ist.
- Der Euklidische Algorithmus hat lineare Laufzeit (die Größe des Inputs wird in Bits gerechnet).
- Der triviale Algorithmus dagegen benötigt $\sim a_0 = 2^b$ Schritte, hat also exponentielle Laufzeit.

Exponentielles Wachstum

Exponentielles Wachstum, z.b. $n \mapsto 2^n$, ist schnell.
Beispiel:

- Wir wollen alle Möglichkeiten durchprobieren/rechnen, einen Speicher mit B Bits zu befüllen. Das sind 2^B viele Möglichkeiten.
- a Instructions per second ("Geschwindigkeit"), t benötigte Sekunden. Dann $2^B = a \cdot t$.
- Aktuell schnellste Computer: $a \sim 10^{14}$.
  Alter des Universums angeblich $t = 4.4 \cdot 10^{17}$.
  Das entspricht $B \sim \ln_2(4.4 \cdot 10^{31}) \sim 106$ Bits (13 Bytes)
- Doppelt so schneller Computer $\rightarrow$ doppelte viel Möglichkeiten $\rightarrow$ nicht doppelt so viele Bits, sondern 1 Bit mehr!
- Angenommen Computer führt 1 Berechnung pro Planck-Zeit durch: $a = 10^{44}$. Dann bekommen wir ~ 204 Bits (25 Bytes)
- Alle Möglichen Konfigurationen eines 26 Byte-Arrays durchzuprobieren benötigt demnach 8 Universen-Lebensalter.
- Bei 35 Bytes sind es $\sim 3.7 \cdot 10^{22} = 37.000.000.000.000.000.000$.

Quantencomputer

- Darin liegt auch das Potential von Quantencomputern, die (möglicherweise, potentiell, für bestimmte Probleme, unter bestimmten Umständen) bei q qubits "in einem Schritt", "simultan" alle 2^q Möglichkeiten testen können.
- Für das 35 Byte array benötigt man dann (statt 37.000.000.000.000.000.000 Universums-Lebenszeiten mit dem hypothetischen, nach aktuellem physikalischem Verständnis schnellstmöglichen Computer) nur "einen Schritt" mit einen Quantencomputer mit "nur" 280 qubit.
- Aktueller Stand ist mir unklar. Qubits müssen auch "robust" gespeichert und manipuliert werden können etc. Mit 2 qubits(??) geht das, da kann man simultan $2^2 = 4$ Zustände testen. Laut Medienberichten hat IBM einen 433 qubits Computer gebaut.

Laufzeit und Kryptographie

- Moral: Lineare Laufzeiten gut. Exponentielle katastrophal. Dazwischen unklar. ($n \log n$ OK. Quadratisch OK? ...)
- Für manche Probleme gibt es offensichtliche aber schlechte Algorithmen, und intelligentere, effizienter.
- Typischer Effekt: Berechnen schwierig, aber überprüfen einfach.
- Bsp: Angenommen wir wissen dass $n = p_1 \cdot p_2$ Produkt zweier Primzahlen ist.
  Wenn n gegeben ist dann ist es enorm schwierig, p_1 und p_2 zu finden.
  Wenn aber p_1, p_2 gegeben ist kann man (relativ) leicht prüfen dass sie Primzahlen sind (und sehr leich) dass ihr Produkt n ist.
- Normalerweise: Lange Laufzeit schlecht. Aber: Diese Asymmetrie Grundlage der Kryptographie.
- $n = p_1 \cdot p_2$ tatsächlich wichtiger Effekt für Kryptographie (n entspricht public key, (p_1, p_2) private Key).

Laufzeit und Kryptographie (Fort.)

Solche Kryptographischen Protokolle/Programme können ganz verschiedene Arten von protentiellen Problemen haben:

- Es gibt momentan keinen Beweis dass es keinen schnellen Algorithmus gibt um p_1 und p_2 zu finden. (Zahlentheoretisches Problem)
- Eventuell könnte Quantencomputer das schnell lösen. (Technisches / physikalisches / futuristisches Problem)
- Die (üblicherweise benötigten) Pseudo-Random-Zahlen sind vielleicht nicht random genug? (Mathematisch / technisches Problem.)
- Die Implementierung kann fehlerhaft sein (realistisches Problem).
- Die Anwendung kann fehlerhaft sein (wahrscheinlichstes Problem).

Euklidischer Algorithmus: Bonus

Sei $a_0 > a_1 \geq 1$. Wir können $\mathrm{ggT}(a_0, a_1)$ als das folgende a_n berechnen:

- $a_0 = \ell_1 \cdot a_1 + a_2$ mit $a_2 < a_1$.
- ...
- $a_{n-2} = \ell_{n-1} \cdot a_{n-1} + a_n$ mit $a_n < a_{n-1}$.
- $a_{n-1} = \ell_n \cdot a_n$.

Theorem

$\mathrm{ggT}(n, m) = a \cdot n + b \cdot m$ *für* $a, b \in \mathbb{Z}$.

Wir sagen: $\mathrm{ggT}(n, m)$ ist Linearkombination (LK) von n und m.
Beweis:

- $a_n = a_{n-2} - \ell_{n-1} \cdot a_{n-1}$, d.h. a_n ist LK von a_{n-1} und a_{n-2}.
- $a_{n-1} = a_{n-3} - \ell_{n-2} \cdot a_{n-2}$. D.h. a_{n-1} ist LK von a_{n-2} und a_{n-3}.
- Damit ist aber a_n auch LK von a_{n-2} und a_{n-3}.
- Allgemein ist a_ℓ Linearkombination von $a_{\ell-1}$ und $a_{\ell-2}$, und damit a_n Linearkombination von $a_{\ell-1}$ und $a_{\ell-2}$.
- Insbesondere ist a_n Linerarkobination von a_0 und a_1.

Wieder mal $\mathbb{Z}_m$

n und m heißen teilerfremd, wenn $\mathrm{ggT}(n, m) = 1$.

Corollary

$a, b \geq 1$ sind teilerfremd, gdw $(\exists \ell, k \in \mathbb{Z})\, 1 = k \cdot a + \ell \cdot b$.

Beweis: Wenn $\mathrm{ggT}(a, b) = 1$, dann ist 1 LK. Sei umgekehrt 1 LK, und d gemeinsamer Teiler. Dann $d|1$. □

Zur Erinnerung: $\mathbb{Z}_m := \mathbb{Z}/m\mathbb{Z} = \{\overline{0}, \overline{1}, \ldots, \overline{m-1}\}$, das ist der Quotient des KREs $\mathbb{Z}$ mit dem Ideal $\langle m \rangle = m\mathbb{Z}$.

Theorem

$\overline{a} \in \mathbb{Z}_m$ ist invertierbar gdw.: a und m sind teilerfremd.

Beweis: Nach Definition ist $\overline{a} \in \mathbb{Z}_m$ ist (multiplikativ) Invertierbar, gdw. $\overline{a} \cdot \overline{k} = \overline{1}$ für ein $\overline{k} \in \mathbb{Z}_m$. Gdw wenn: $a \cdot k = 1 + \ell \cdot m$ für ein $k, \ell \in \mathbb{Z}$. Gdw: 1 ist Linearkombination von a, m. Gdw: a und m teilerfremd. □

Körper

Körper

Definition

Ein Integritätsbereich, in dem jedes $x \neq 0$ ein Multiplikatives Inverses hat, heißt Körper.

- Bsp: $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ sind Körper.
- $\mathbb{Z}$ nicht.
- In einem Körper K ist der Begriff der Teilbarkeit trivial / uninteressant: Für jedes $n, m \in K \setminus \{0\}$ gibt es ℓ mit $m = \ell \cdot n$ und daher gilt $m|n$ (und $n|m$).

Körper $\mathbb{Z}_p$

Wenn p Primzahl ist, dann ist jedes $1 \leq n \leq p$ teilerfremd zu p.
Daher gilt:

Theorem

$\mathbb{Z}_m$ ist ein Körper gdw. m Primzahl.

Beweis:

- Wir wissen bereits: Wenn $m = a \cdot b$ keine Primzahl ist, dann ist $\overline{a} \cdot \overline{b} = \overline{0}$ ein Nullteiler in $\mathbb{Z}_m$.
- Sei p Primzahl. Dann ist jedes $1 < a < p$ teilerfremd zu p, und damit ist $\overline{a}$ invertierbar.
  $\mathbb{Z}_p$ ist also ein Integritätsbereich in dem jedes Element $\neq 0$ invertierbar ist, und damit per Definition ein Körper. □

Charakteristik von Körpern

Definition

Die Charakteristik eines Körpers K ist das minimale n mit $\underbrace{1 + \cdots + 1}_{n\text{-mal}} = 0$, wenn so ein n existiert ansonsten 0.

- Wenn die Characteristik $n \neq 0$ ist, dann ist n Primzahl.

  Beweis: Wegen des DG ist
  $(\underbrace{1 + \cdots + 1}_{n\text{-mal}}) \cdot (\underbrace{1 + \cdots + 1}_{m\text{-mal}}) = \underbrace{1 + \cdots + 1}_{n \cdot m\text{-mal}}$. □
- Ein Körper mit Charakteristik 0 ist unendlich (Bsp: $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$).
- Körpern mit Charakteristik $p \neq 0$ können endlich sein ($\mathbb{Z}_p$, aber auch andere) oder unendlich.
- Endliche Körper spielen eine zentrale Rolle in der Kryptographie.

Polynomdivision

Polynomdivision

- Zur Erinnerung: Wenn R ein Integritätsbereich ist, dann bilden die (formalen) Polynome ebenfalls einen Integritätsbereich, genannt $R[X]$.
- $0 \neq f(X) \in R[X]$ hat die Form $a_0 + a_1 \cdot X + \cdots + a_n X^n$, mit $a_n \neq 0$. n heißt Grad von f oder $\deg(f)$, und a_n ist der "führende Koeffizient". $\deg(0) := -1$

Theorem

Wenn $f(X), g(X) \in R[X]$ und g führenden Koeffizienten 1 hat, dann gibt es h, r mit $f(X) = g(X) \cdot h(X) + r(X)$, mit $\deg(r) < \deg(g)$.

Äquivalent: Es gibt h mit $\deg(f - h \cdot g) < \deg(g)$.

Beweis: Wenn $\deg(g) > \deg(f)$, setze $h := 0$ und $r := f$.
Sei also $m := \deg(g) \leq \deg(f) =: n$, sei a_n führender Koeffizient von f.
Setze $h_0(X) := a_n \cdot X^{n-m}$. Dann $i := \deg(f - h_0 \cdot g) < n$.
Wenn $i \geq m$ ist, finde h_1 mit $\deg(f - h_0 \cdot g - h_1 \cdot g) < i$, etc.
Nach endlich vielen Schritten ist $\deg(f - h_0 \cdot g - h_1 \cdot g - \cdots - h_\ell \cdot g) < m$,
d.h. $h := h_1 + \cdots + h_\ell$ funktioniert.

Polynomdivision: Bsp

- Beispiel in $\mathbb{Z}[X]$: $f(X) = X^2$ und $g(X) = 2 \cdot X$.
  Hier gibt es kein h mit $\deg(f - g \cdot h) < \deg(g) = 1$.
  (Kein Widerspruch zu vorigem Satz, g hat nicht führenden Koeffizienten 1).
- Beispiel in $\mathbb{Z}[X]$: $f = 3X^3 + 2X^2 + X + 1$ und $g = X^2 + 2$.
  Wir bekommen:
  $f - 3X \cdot g = 2X^2 - 5X + 1$,
  $f - 3X \cdot g - 2 \cdot g = -5X - 3$.
  D.h.: $f = g \cdot (3X + 2) + (-5X - 3)$.
- Bemerkung: Die Rechnung funktioniert genauso in $\mathbb{Q}[X]$, $\mathbb{R}[X]$ oder $\mathbb{C}[X]$; weil $\mathbb{Z}$ Teilring von $\mathbb{Q}$, $\mathbb{R}$ und $\mathbb{C}$ ist.
- In $\mathbb{Z}_2[X]$: $f = X^3 + X$ und $g = X^2 + X$
  $f - g \cdot X = -X^2 + X = X^2 + X$
  $f - g \cdot X - g = 0$
  D.h. $f = g \cdot (X + 1)$.
- Bemerkung: In $\mathbb{Z}[X]$ teilt $X^2 + X$ aber $X^3 + X$ nicht.
  $\mathbb{Z}_2$ ist kein Teilring von $\mathbb{Z}$.

Abspalten von Nullstellen

Als Folgerung bekommen wir:

Theorem

Wenn $f(a_0) = 0$ *gilt für* $f \in R[X]$ *und* $a_0 \in R$, *dann* $f(X) = (X - a_0) \cdot h(X)$ *mit* $h \in R[X]$.

Beweis: $f(X) = (X - a_0) \cdot h(X) + r(X)$ mit $\deg(r(X)) < 1$, d.h. $r(X) \equiv b$ konstant. Dann ist $f(a_0) = (a_0 - a_0) \cdot h(a_0) + b = b = 0$, d.h. $r = 0$. □

Beispiel: In $\mathbb{Z}_2[X]$: $f = X^2 + X$, $a_0 = 1$. Dann ist $f(a_0) = 1^2 + 1 = 2 = 0$, und $f = (X - 1)(X) = (X + 1)X$.

Polynomdivision in $K[X]$

- Wenn K ein Körper ist, dann ist es nicht nötig zu fordern dass der führende Koeffizient b_n von g gleich 1 ist, es reicht dass $g \neq 0$. (Wir können dann in der Konstruktion $\frac{a_n}{b_n} X^{n-m}$ verwenden, statt $a_n X^{n-m}$, etc.)
- Daher gilt in $K[X]$ "Division mit Rest":

Theorem

Wenn $f, g \in K[X]$, *und* $g \neq 0$, *dann gibt es* $h, r \in K[X]$ *mit* $f = g \cdot h + r$ *und* $\deg(r) < \deg(g)$.

- Bemerkung: Ringe in denen eine derartige "Division mit Rest" (mit einem sinnvollen Grad-Begriff) gilt, heißen Euklidische Ringe.
- Beispiele: $\mathbb{Z}$, $\mathbb{Q}[X]$.
- In Euklidischen Ringen funktioniert der Euklidische Algorithmus. Insbesonders ist $\mathrm{ggT}(x, y)$ Linearkombination von x und y.

Homomorphismen und Isomorphismen

Nachtrag

Wir haben definiert was es heißt dass $f : A \to B$ injektiv, surjektiv und bijektiv ist.

Es gilt nun:

Theorem

Wenn $|A| = |B| = n$ endlich, dann ist surjektiv äquivalent zu injektiv (und damit auch zu bijektiv).

Beweis: Wenn $|A| = n$, $f : A \to B$ dann ist $|f''A| \leq n$, und $|f''A| = n$ gdw f injektiv.
(Das sollte klar sein, wenn man das formal beweisen will kann man es zB mit Induktion für Mengen A der Größe n tun: Sei $A = A_0 \cup \{a\}$, $|A| = n + 1$, $|A| = n$. Wenn $f : A \to B$ injektiv, dann auch $f \restriction A_0$, daher $|f''A_0| = n$, und $f(a) \notin f''A_0$ (weil f injektiv), daher $|f''A| = n + 1$. Ähnlich für nicht injektiv.)

Wenn also $f : A \to B$ und $|A| = |B| = n$, dann ist $f''A = B$ gdw. f injektiv.

Gilt nicht für unendlich!

Zur Erinnerung: Für unendliche Mengen gilt das nicht.

$f : \mathbb{N} \to \mathbb{N} \quad n \mapsto 2 \cdot n$ ist injektiv aber nicht surjektiv,

$g : \mathbb{N} \to \mathbb{N} \quad n \mapsto \begin{cases} \frac{n}{2} & \text{wenn } n \text{ gerade} \\ 0 & \text{sonst,} \end{cases}$ surjektiv aber nicht injektiv.

Homomorphismen

- Ein Homomorphismus einer algebraischen Struktur (Gruppe, Ringe, Körper, ...) ist eine Abbildung die die jeweiligen Operationen erhält.
- Konkret: Seien $(G_1, \circ_1)$, $(G_2, \circ_2)$ Gruppen. Dann ist $g : G_1 \to G_2$ ein Homomorphismus, wenn $f(a \circ_1 b) = f(a) \circ_2 f(b)$ für alle $a, b \in G_1$.
- Für Ringe $(R_1, +_1, \cdot_1)$ und $(R_2, +_2, \cdot_2)$ ist $g : R_1 \to R_2$ ein Homomorphismus, wenn $g(a +_1 b) = g(a) +_2 g(b)$ und $g(a \cdot_1 b) = g(a) \cdot_2 g(b)$ für alle $a, b \in R_1$, und $g(1_1) = 1_2$.
- Die Verknüpfung von Homomorphismen ist ein Homomorphismus. (Folgt trivialerweise aus Definition.)
- $f : A \to \{0\}$ $a \mapsto 0$ ist Homomorphismus.

  $\{0\}$ ist ja (die kleinste) Gruppe und (der kleinste) Ring (sogar KRE) (aber per Definition kein Integritätsbereich).

Bilder von Homomorphismen

- Wenn $f : A \to B$ ein Gruppen-Homomorphismus ist, dann ist $f''A$ Teil-Gruppe von B, und $f(e_A) = e_B$.
  Wenn $f : A \to B$ ein Ring-Homomorphismus ist, dann ist $f''A$ Teil-Ring von B, und $f(0_A) = 0_B$.

  Beweis: Für Gruppen:
  $f(a) = f(e_1 \circ_1 a) = f(e_1) \circ_2 f(a) \xrightarrow{\text{kürzen}} e_2 = f(e_1)$.
  $f(x^{-}1) \circ_2 f(x) = f(e_1) = e_2$, d.h. $f''A$ ist abgeschlossen unter Inversen (und natürlich Verknüpfung) und damit Untergruppe.

  Für Ringe: Eingeschränkt auf $+$ ist f Gruppenhomomorphismus, d.h. $(f''A, +)$ ist Gruppe. Nach Definition ist $1 \in f''A$; und $f''A$ ist abgeschlossen unter $\cdot$; damit ist $f''A$ Teilring von B.

Isomorphismen

- $f : A \to B$ ist ein Isomorphismus, wenn f ein bijektiver Homomorphismus ist. Dann heißen A, B isomorph.

  Isomorph $\equiv$ "identisch abgesehen von Umbenennungen".
- Identität: $\mathrm{Id} : A \to A$ ist klarerweise ein Isomorphismus.
- Die Verknüpfung von Isomorphismen ist ein Isomorphismus.
- Wenn $f : A \to B$ ein Isomorphismus ist, dann ist $f^{-1} : B \to A$ ebenfalls Isomorphismus.

  Beweis für Gruppen (Ring analog):
  Zu zeigen: $f^{-1}(x \circ_2 y) = f^{-1}(x) \circ_1 f^{-1}(y)$. Sei $x = f(a)$ und $y = f(b)$. Dann ist $f^{-1}(x \circ_2 y) = f^{-1}(f(a) \circ_2 f(b)) = f^{-1}(f(a \circ_1 b)) = a \circ_1 b = f^{-1}(x) \circ_1 f^{-1}(y)$. □

Einbettungen

- Einen injektiven Homomorphismus $f : A \to B$ nennt man auch Einbettung.
- Wenn A eine Unter-Gruppe (oder -Ring) von B ist, dann ist $\mathrm{Id} : A \to B$ eine Einbettung.
- Umgekehrt: Wenn $f : A \to B$ eine Einbettung ist, dann sind A und $f''A$ isomorph. Wir können also A als Unter-Gruppe (bzw. -Ring) von B finden.

  (Wir wissen bereits dass $f''A$ eine Unter-Gruppe oder -Ring ist, und $f : A \to f''A$ ist trivialerweise bijektiv und daher ein Isomorphismus.)

Gruppen-Homomorphismen und -Isomorphismen

- Für zwei Gruppen G_1, G_2 gibt es den trivialen Homomorphismus: $f : G_1 \to G_2,\ \ x \mapsto e_2$.
  Wenn G_1, G_2 nicht-trivial sind, ist f weder injektiv noch surjektiv.

  (Zwischen je zwei Gruppen gibt es also immer mindestens einen Homomorphismus.)
- Für $a \in G$ ist $f : \mathbb{Z} \to G$, $k \mapsto a^k$ ein Homomorphismus (mit Bild $\langle a \rangle$).

  Zur Wiederholung: $\langle a \rangle = \{a^k : k \in \mathbb{Z}\}$ ist die "Zyklische Untergruppe" von G, mit den Operationen $a^k \circ a^\ell = a^{k+\ell}$ ist (und $a^0 = 1$, $a^{-k} = (a^k)^{-1}$).
- Zwei Isomorphe Strukturen sind "gleich bis auf Umbenennung". Insbesondere gilt: Wenn G_1 und G_2 isomorph sind, dann ist G_1 kommutativ gdw. G_2 kommutativ, etc.
- Beispiel: $f : (\mathbb{R}, +) \to (\mathbb{R}^+, \cdot)$ $x \mapsto e^x$ ist ein Gruppenisomorphismus. (In dem Sinn sind Addition und Multiplikation Isomorph).

Gruppen-Einbettungen

- Beispiel: $f_n : (\mathbb{Z}, +) \to (\mathbb{Z}, +)$ $k \mapsto k \cdot n$ ist eine Einbettung, mit Bild $n\mathbb{Z}$.
- Beispiel: $f_0 : S_3 \to S_7$ $\begin{pmatrix} 1 & 2 & 3 \\ a & b & c \end{pmatrix} \mapsto \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ a & b & c & 4 & 5 & 6 & 7 \end{pmatrix}$

  f_0 ist eine Einbettung (nicht surjektiv).
  Via f_0 ist S_3 ist Teilgruppe von S_7; es gibt auch viele andere Arten S_3 als Teilgruppe in S_7 einzubetten.
- Beispiel: Es kann keine Einbettung $f : S_3 \to (\mathbb{Z}_n, +)$ geben, weil S_3 nicht kommutativ ist:

  Beweis: Sei $a \circ b \neq b \circ a$. Dann
  $f(a) \circ f(b) \overset{\text{Hom.}}{=} f(a \circ b) \overset{\text{inj.}}{\neq} f(b \circ a) \overset{\text{Hom.}}{=} f(b) \circ f(a)$,
  ein Widerspruch zur Kommutativität von $\mathbb{Z}_n$.
- Bsp: $G_1 \to G_1 \times G_2 \quad a \mapsto (a, e_2)$ ist eine Einbettung.

  (Zur Erinnerung: Das kartesischen Produkt $G_1 \times G_2$ zweier Gruppen ist mit $(a, x) \circ (b, y) := (a \circ_1 b, x \circ_2 y)$ eine Gruppe.)

Zyklische Gruppen

- Erinnerung: Die Ordnung von a, $\mathrm{ord}(a)$, ist das kleinste n mit $a^n = e$ (oder unendlich); das ist auch gleich $|\langle a \rangle|$.
- Beispiel: Wir haben $a \in S_3$ mit Ordnung 3 definiert (shift). Dann ist $f_1 : (\mathbb{Z}_3, +) \mapsto S_3 \quad k \mapsto a^k$ eine Einbettung mit Bild $\langle a \rangle$.
- Allgemein: Wenn $\mathrm{ord}(a) = n$ (oder: unendlich), dann ist die Funktion $f : (\mathbb{Z}_n, +) \to G \quad \overline{k} \mapsto a^k$, (oder: $f : (\mathbb{Z}, +) \to G \quad k \mapsto a^k$) eine Einbettung mit Bild $\langle a \rangle$.

  Beweis: Die Ordnung von a ist unendlich gdw. alle a^k verschieden ist, dann ist $f : \mathbb{Z} \to G$ ein injektiver Homomorphismus. Wenn $\mathrm{ord}(a) = n$, dann ist $a^k \neq a^\ell$ für $0 \leq k, l < n$, $k \neq l$; damit ist f injektiv. f ist ein wohldefinierter Homomorphismus: Wenn $\overline{k} + \overline{\ell} = \overline{m}$, dann $k + \ell = m - i \cdot n$ für ein $i \in \mathbb{Z}$, und dann $f(\overline{m}) = a^m = a^{k+\ell} \circ a^{i \cdot n} = a^{k+\ell} \circ a^n \circ \cdots \circ a^n = a^{k+\ell} = f(\overline{k}) \circ f(\overline{\ell})$. □

Ordnungen von Gruppenelementen

- Wenn $f : G_1 \to G_2$ Homomorphismus ist, $a \in G_1$, mit endlicher Ordnung $\mathrm{ord}(a)$, dann $\mathrm{ord}(f(a)) | \mathrm{ord}(a)$.

  Beweis: $a^{\mathrm{ord}(a)} = e_1$, d.h. $(f(a))^{\mathrm{ord}(a)} = e_2$, damit ist $\mathrm{ord}(a)$ ein Vielfaches von $\mathrm{ord}(f(a))$.

- Wenn $f : G_1 \to G_2$ eine Einbettung ist, dann ist $\mathrm{ord}(a) = \mathrm{ord}(f(a))$.

  Beweis: Wenn $U < G$ und $a \in U$, dann ist $\mathrm{ord}(a)$ für a als Element von U gleich (weil gleich berechnet) wie für $a \in G$.
  Und $f''G_1$ ist eine Untergruppe von G_2 isomorph zu G_1.

  Alternativer Beweis: $\deg(a)$ ist minimales n mit $a^n = e_1$, d.h. wenn $m < n$, dann $a^m \neq e_1$, daher $f(a^m) = f(a)^m \neq e_2$. Und $f(a)^n = f(a^n) = f(e_1) = e_2$.

Beispiele

- Es gibt keinen nicht-trivialen Gruppenhomom. $f : G \to \mathbb{Z}$ wenn G eine endliche Gruppe ist.

  Beweis: Jedes $a \in G$ hat endliche Ordnung n, und es gilt $\text{ord}(f(a))|n$ aber nur 0 hat endliche Ordnung (nämlich 1) in $\mathbb{Z}$.
- Es ist aber nicht immer offensichtlich ob es eine nicht-triviale Homomorphismen, oder Einbettungen, für Gruppen G_1, G_2 gibt.
- Fall $|G_1| = |G_2|$ endlich ist, dann ist f injektiv gdw surjektiv gdw bijektiv; d.h. jede Einbettung ist ein Isomorphismus.
- Bsp: $\mathbb{Z}_2 \times \mathbb{Z}_2$ ist nicht isomorph (als Gruppe) zu $\mathbb{Z}_4$.
  Beweis: $\mathbb{Z}_4$ ist eine zyklische Gruppe der Ordnung 4. Aber jedes Element von $\mathbb{Z}_2 \times \mathbb{Z}_2$ hat Ordnung 1 oder 2.

Ring-Homomorphismen: $1 \mapsto 1$

- Bemerkung: Bei der Definition von Ring-Homomorphismus kann man die Forderung $g(1_1) = 1_2$ nicht weglassen.

  Bsp: $f_3 : \mathbb{Z}_6 \to \mathbb{Z}_6$ $x \mapsto 3x$ erhält $+$ und $\cdot$, aber $f_3(\overline{1}) \neq \overline{1}$.
  Für $+$ klar, weil $3(x + y) = 3x + 3y$, Für $\cdot$ deswegen, weil in $\mathbb{Z}_6$
  $3^2 = 9 = 3$, d.h. $3(x \cdot y) = (3 \cdot x) \cdot (3 \cdot y)$.
  f_3 ist also (nach unserer Definition) kein Ring-Homomorphismus.
- Bsp: $R_1 \to R_1 \times R_2 \quad a \mapsto (a, 0)$ ist (nach unserer Definition) keine Einbettung falls $0 \neq 1$ in R_2 , weil $1 \mapsto (1, 0) \neq (1, 1)$.

Ring-Homomorphismen

- Beispiel: $f : \mathbb{Z} \to \mathbb{Z}_m \quad n \mapsto \overline{n}$ ist ein Ring-Homomorphismus (surjektiv, nicht injektiv).
- $f : \mathbb{Z}_m \to \mathbb{Z}_n$ definiert durch $f(\overline{n}) \mapsto \overline{n}$ ist wohldefiniert (und Homorphismus) gdw $n|m$.
- $\mathbb{Z}_2 \times \mathbb{Z}_3$ ist isomorph zu $\mathbb{Z}_6$.
  Beweis: Es gibt nur einen potentiellen Isomorphismus f:
  $(\overline{0}, \overline{0}) \mapsto 0$ und $(\overline{1}, \overline{1}) \mapsto \overline{1}$ (weil f Null und Eins erhält.)
  f erhält $+$ impliziert dann:
  $(\overline{0}, \overline{2}) \mapsto \overline{2}$, $(\overline{1}, \overline{0}) \mapsto \overline{3}$, $(\overline{0}, \overline{1}) \mapsto \overline{4}$, $(\overline{1}, \overline{2}) \mapsto \overline{5}$.
  Diese Funktion ist bijektiv und erhält $+$ und $\cdot$ (nachrechnen).

Körper-Homomorphismen

- Wenn $f : K_1 \to K_2$ ein Ring-Homomorphismus ist, wobei K_1 ein Körper ist und $0_2 \neq 1_2$, dann ist f injektiv.

  Beweis: Sei $a \neq b$ mit $f(a) = f(b)$, oder äquivalent $f(c) = 0_2$ mit $c := b - a \neq 0_1$. Sei $d = c^{-1}$. Dann ist
  $1_2 = f(1_1) = f(d \cdot c) = f(d) \cdot_2 f(c) = f(d) \cdot_2 0_2 = 0_2$,
  Widerspruch. □

- Wenn $f : K_1 \to K_2$ ein Homomorphismus ist, dann ist die Charakteristik von K_1 gleich der von K_2.

  Beweis: Die Charakteristik eines Körpers K ist deg(1) in $(K, +)$. Einbettungen erhalten die Ordnung eines Elements.

Endlich erzeugte Abelsche Gruppen

- Wichtige Mathematische Frage: Klassifiziere ... bis auf Isomorphie.
- Funktioniert für bestimmte Strukturen sehr einfach (werden sehen: Vektorräume über gegebenen Körper), für andere hoffnungslos.
- Ein Beispiel bei dem das gut funktioniert: Endlich erzeugte abelsche Gruppen.
- Sei G eine Gruppe. $A \subseteq G$ erzeugt G, wenn sich jedes $g \in G$ schreiben lässt als $b_1 \circ b_2 \circ \cdots \circ b_m$, wobei jedes $b_i \in A$ oder $b_i^{-1} \in A$.
- G ist endlich erzeugt, wenn A endlich ist.

Theorem (Ohne Beweis)

Jede endlich erzeugte abelsche Gruppe ist isomorph zu $\mathbb{Z}_{n_1} \times \mathbb{Z}_{n_2} \times \cdots \times \mathbb{Z}_{n_i} \times (\mathbb{Z})^k$, *für ein* $k \geq 0$ *und* $n_j \geq 2$.

Darüberhinaus kann man verlangen dass n_j *eine Primzahlpotenz* $p_j^{\ell_j}$ *ist, und dann ist die Darstellung eindeutig.*

Kern (Ringe)

- Der Kern eines Homomorphismus $f : A \to B$ ist die Menge aller $a \in A$ die auf das neutrale Element von B abgebildet werden.
- Konkret für Ringe: $f : R_1 \to R_2$ Homomorphismus, ist $\ker(f) := \{a \in R_1 : f(a) = 0_2\}$.
- $0 \in \ker(f)$. Wenn $a, b \in \ker(f)$ und $c \in R_1$, dann sind auch in $\ker(f)$: $a \pm b$ und $a \cdot c$ und $c \cdot a$.

  Beweis: $f(a \pm b) = f(a) \pm f(b) = 0$, $f(a \cdot c) = f(a) \cdot f(c) = 0$, etc.
- (Für KREs R) haben wir so etwas "Ideal" genannt: $\ker(f)$ ist Ideal. Zur Erinnerung: R/I (die Restklassen modulo $a \sim b :\leftrightarrow b - a \in I$) ist wieder ein KRE, mit $[a + b] := [a] + [b]$ und $[a] \cdot [b] := [a \cdot b]$.
- (Bemerkung: Ein Ideal, insbesondere $\ker(f)$, ist nach unserer Definition i.A. kein Unter-Ring von R_1, weil i.A. $1 \notin \ker(f)$.)

Homomorphiesatz (Ringe)

Theorem

Wenn $f : R_1 \to R_2$ Homomorphismus, dann ist $g : R_1/\ker(f) \to f''R_1$, $[a] \mapsto f(a)$ ist ein Isomorphismus.

Das ist eine Verallgemeinerung von:
Wenn f injektiv, dann ist R_1 isomorph zu $f''R_1$.

Beweis:

- g ist wohldefiniert: Wenn $[a] = [b]$, d.h. $a - b \in \ker(f)$, d.h. $f(a - b) = 0$, dann $f(a) - f(b) = 0$, d.h. $a(a) = f(b)$.
- $g([a] + [b]) = g([a + b]) = f(a + b) = f(a) + f(b) = g([a]) + g([b])$. Analog für $\cdot$.
  $g([0]) = f(0) = 0$. □

Das übliche Beispiel: $f : \mathbb{Z} \to \mathbb{Z}_m$ mit $m \mapsto \overline{m}$. Dann ist $\ker(f) = m\mathbb{Z} = \langle m \rangle$, und $\mathbb{Z}/m\mathbb{Z}$ ist isomorph zu $\mathbb{Z}_m$.

Normalteiler

Definition

Eine Untergruppe $U < G$ heißt Normalteiler, wenn gilt:
Wenn $a \in G$ und $b \in U$, dann $a^{-1}ba \in U$.

- U Normalteiler ist äquivalent zu: $Ua = aU$ (Linksnebenklassen = Rechtsnebenklassen).

  Beweis: $c \in aU$, gdw $(\exists b \in U)\, c = a \circ b$. Dann $c = b' \circ a$, mit $b' = a \circ b \circ a^{-1}$, Nach Def. Normalteiler ist $b' \in U$, d.h. $c \in Ua$. Das impliziert $aU \subseteq Ua$, und die andere Teilmengenbeziehung bekommen wir analog. Damit ist $aU = Ua$.

Normalteiler (Forts.)

- $aU = bU$ ist äquivalent zu $b \in aU$ und zu $a^{-1}b \in U$, und wenn U Normalteiler ist, auch zu $b \in Ua$ und zu $ba^{-1} \in U$.
- Wenn U Normalteiler, dann setzte $a \sim b := ab^{-1} \in U$, und setze G/U die Menge der Restklassen $\{[a] : a \in G\}$.

Theorem

Wenn U Normalteiler ist, dann ist G/U Gruppe, mit $[a] \circ [b] := [a \circ b]$.

Beweis: Zu zeigen: Wohldefiniert. (Rest folgt wie üblich.)
Angenommen $aU = bU$ und $xU = yU$. Zu zeigen: $axU = byU$.
Aber $axU = ayU = aUy = bUy = Uby = byU$. □

Homomorphiesatz (Gruppen)

Theorem

Wenn $f : G_1 \to G_2$ *Homomorphismus, dann ist*
$\ker(f) := \{a \in G_1 : f(a) = e\}$ *ein Normalteiler, und*
$g : G_1/\ker(f) \to f''G_1$ $[a] \mapsto f(a)$ *ist ein Isomorphismus.*

Beweis:

- Zu zeigen: g Wohldefiniert. (Rest wie üblich.)
- Sei $U := \ker(f)$ und $aU = bU$. Z.Z.: $f(a) = f(b)$. $a \circ u_1 = b \circ u_2$ für $u_1, u_2 \in \ker(f)$, d.h. $b = u_2^{-1} \circ a \circ u_1$ und $f(b) = f(u_2)^{-1} \circ f(a) \circ f(u_1) = a$.

Nachtrag

- In Gruppen gilt: $a \neq b$ gdw $a \circ b^{-1} \neq e$.
  (Beweis: Wenn $a = b$ dann $a \circ b^{-1} = b \circ b^{-1} = e$.
  Wenn $e = a \circ b^{-1}$, dann $b = a \circ b^{-1} \circ b = a$.
  Das zeigt: $a = b$ gdw $a \circ b^{-1} = e$.
  $A \leftrightarrow B$ ist äquivalent zu $\neg A \leftrightarrow B$
  D.h. wir haben gezeigt: $a \neq b$ gdw. $a \circ b^{-1} \neq e$.)
- Analog für Ringe: $a \neq b$ gdw $a - b \neq 0$.
  (Beweis: Wenn $(R, +, \cdot)$ ein Ring ist, dann ist nach Definition von Ring $(R, +)$ eine (kommutative) Gruppe.)
- Sei n das neutrale Element (e in Gruppen, 0 in Ringen.)
  Für einen Homomorphismus $f : A \to B$ haben wir definiert:
  $\ker(f) = \{a \in A : f(a) = n\}$.
  D.h.: a ist in $\ker(f)$ gdw: $a \in A$ und $f(a) = n$.
- $n \in \ker(f)$, weil $f(n) = n$ nach Def. Homomorphismus.

Nachtrag (Forts.)

Theorem (Für f (Gruppen- oder Ring)homomorphismus)

f ist injektiv gdw $\ker(f) = \{n\}$.

Beweis (Gruppennotation, insbes. $n := e$):

- $e \in \ker(f)$ gilt immer. Daher gilt: $\ker(f) \neq \{n\}$ impliziert: Es gibt $a \neq e$ mit $a \in \ker(f)$. Dann ist $f(a) = e = f(e)$, d.h. f nicht injektiv. Das zeigt: $\ker(f) \neq \{n\}$ impliziert f nicht injektiv.
- Sei f nicht injektiv ist, d.h. es gibt $a \neq b$ mit $f(a) = f(b)$. Dann $f(a \circ b^{-1}) = f(a) \circ f(b)^{-1} = e$ d.h. $a \circ b^{-1} \in \ker(f)$. Wir haben gesehen dass $a \neq b \rightarrow a \circ b^{-1} \neq e$. D.h. wir haben bewiesen: f nicht injektiv impliziert $\ker(f) \neq \{n\}$.
- Sei A die Aussage "$\ker(f) = \{n\}$" und B die Aussage "f ist injektiv." Wir haben also gezeigt: $\neg A \rightarrow \neg B$ und $\neg B \rightarrow \neg A$ bewiesen, und daher $\neg A \leftrightarrow \neg B$, und damit $A \leftrightarrow B$.

(Für Ringe genauso, nur schreiben wir:
$+$ statt $\circ$, $-a$ statt a^{-1} und 0 statt e.)

Wieder einmal S_n

- Für $a \in S_n$ $(n > 1)$ definiere $\operatorname{sgn}(a) := \prod_{1 \le i < j \le n} \frac{a(j)-a(i)}{j-i} = \pm 1$.
- Behauptung: Im Zähler und Nenner treten, bis auf Vorzeichen, dieselben Zahlen auf, weil a eine Bijektion ist.
  Daher ist das Produkt ± 1.
- Statt eines Beweises, die alten Bsp aus S_3:
  $a = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$, $b = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$.
  $$\operatorname{sgn}(b) = \operatorname{sgn}\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} = \frac{3-2}{3-1} \cdot \frac{3-1}{3-2} \cdot \frac{1-2}{2-1} = 1 \cdot 1 \cdot (-1) = -1$$
  $$\operatorname{sgn}(a) = \operatorname{sgn}\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} = \frac{1-2}{3-1} \cdot \frac{1-3}{3-2} \cdot \frac{3-2}{2-1} = (-1) \cdot (-1) \cdot 1 = 1$$
- s heißt gerade Permutation, wenn $\operatorname{sgn}(s) = 1$, sonst ungerade.
- b Vertauscht zwei Positionen, solche Permutationen sind immer ungerade.

Die alternierende Gruppe

Theorem

$f : S_n \to (\{-1, 1\}, \cdot)$ $a \mapsto \operatorname{sgn}(a)$ *ist ein Gruppenhomomorphismus.*

Bem.: $(\{-1, 1\}, \cdot)$ ist eine Gruppe,
und isomorph zu $(\mathbb{Z}_2, +)$; mittels $-1 \mapsto 1$ und $1 \mapsto 0$.
Übung: Es gibt bis auf Isomorphie nur eine Gruppe der Größe 2.

Beweis:

$$\begin{aligned}\operatorname{sgn}(b \circ a) = \prod_{1 \le i \le j \le n} \frac{b(a(j)) - b(a(i))}{i - j} &= \prod_{1 \le i \le j \le n} \frac{b(a(j)) - b(a(i))}{i - j} \cdot \prod_{1 \le i \le j \le n} \frac{a(j) - a(i)}{a(j) - a(i)} \\ &= \prod_{1 \le i \le j \le n} \frac{b(a(j)) - b(a(i))}{a(j) - a(i)} \cdot \prod_{1 \le i \le j \le n} \frac{a(j) - a(i)}{i - j} \\ &= \operatorname{sgn}(b) \cdot \operatorname{sgn}(a)\end{aligned}$$

Definition

$A_n := \ker(f)$ ist die "alternierende Gruppe" (d.h., die Gruppe der geraden Permutationen.

Die alternierende Gruppe als Normalteiler

Wiederholung: (Eine Untergruppe $U < G$ heißt Normalteiler, wenn gilt: Wenn $a \in G$ und $b \in U$, dann $a^{-1}ba \in U$.)

Theorem

Wenn $f : G_1 \to G_2$ Gruppenhomomorphismus, dann ist
$\ker(f) := \{a \in G_1 : f(a) = e\}$ *ein Normalteiler, und*
$g : G_1/\ker(f) \to f''G_1$ *mit* $[a] \mapsto f(a)$ *ist ein (wohldef.) Isomorphismus.*

- $f : S_n \to (\{-1, 1\}, \cdot)$ $a \mapsto \operatorname{sgn}(a)$ ist ein Gruppenhomomorphismus.
- Die alternierende Gruppe $A_n = \ker(f)$ ist also ein Normalteiler von S_n.
- Nach dem Homomorphiesatz ist $S_n/A_n \to \{-1, 1\}$ $[s] \mapsto \operatorname{sgn}(s)$ ein Isomorphismus. S_n zerfällt also in zwei Nebenklassen (gerade und ungerade Permutationen), jeweils der Größe $\frac{n!}{2}$.

Noch ein bisschen S_n

- In S_3 war $A_3 = \langle a \rangle$.

  Beweis: $a \in A_3 = \ker(f)$, daher ist $\langle a \rangle \subseteq \ker(f)$, und $|\langle a \rangle| = |A_3| = 3$.
- Für $n \geq 4$ ist A_n aber nicht zyklisch, und nicht einmal kommutativ. (Beipiel später.)
- Wenn b das Vertauschen von 1 mit 2 ist, dann ist $b \notin A_n$.
  Das heißt: Die Nebenklasse bA_n sind die ungeraden Permutationen, und $bA_n = A_n b$ (weil A_n ein Normalteiler ist).
  Jede ungerade Permutation läßt sich also schreiben als $b \circ a$, mit a gerade, und auch als $a' \circ b$, mit a' gerade.

Satz von Caley

Erinnerung: Die symmetrische Gruppe S_A zu A ist die Gruppe aller Permutationen von A, d.h. der Bijektionen $A \to A$.

Theorem

Jede Gruppe G lässt sich in die symmetrische Gruppe S_G einbetten.

Beweis:

- Sei $\pi : G \to S_G$ definiert durch $\pi(g) : G \to G,\ a \mapsto g \circ a$.
- $f := \pi(g)$ ist tatsächlich in S_G:
  f ist injektiv: Wenn $a \neq b$, dann $f(a) = g \circ a \overset{\text{wg. kürzen}}{\neq} g \circ b = f(b)$.
  f ist surjektiv: Wenn $a \in G$, dann $f(g^{-1} \circ a) = g \circ g^{-1} \circ a = a$.
- π ist Homomorphismus: Zu zeigen:
  $\pi(g \circ h) = \pi(g) \circ \pi(h)$ (Verknüpfung von Funktionen).
  $\pi(g \circ h)(a) = \pi(g)(\pi(h)(a))$; aber $\pi(g \circ h)(a) := (g \circ h) \circ a$, und $\pi(g)(\pi(h)(a)) := \pi(g)(h \circ a) := g \circ (h \circ a)$.
- π ist injektiv. Z.z.: $a \neq e \to \pi(a) \neq \mathrm{Id}_G$. Aber $\pi(a)(a^{-1}) = e \neq a^{-1}$. □

Satz von Caley: Beispiel

- Betrachte die Gruppe $G := \mathbb{Z}_2 \times \mathbb{Z}_2$.
- $|G| = 4$, d.h. S_G ist isomorph zu S_4.
  Schreibe $G = \{a_1, a_2, a_3, a_4\}$, mit
  $a_1 = (0,0)$, $a_2 = (1,0)$, $a_3 = (0,1)$, $a_4 = (1,1)$.
  Zu jedem $b \in G$ gibt es also ein $i(b) \in \{1,2,3,4\}$ mit $b = a_{i(b)}$.
- $\pi : G \to S_4$ ist so definiert:
  $\pi(b)$ ist die Permutation in S_4 die i auf j abbildet, mit $a_j = b \circ a_i$.
- Z.B. $\pi(a_2)$: $a_2 \circ a_1 = a_2$, $a_2 \circ a_2 = a_1$, $a_2 \circ a_3 = a_4$, $a_2 \circ a_4 = a_3$.
  D.h. $\pi(a_2) = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix}$
- Auf diese Weise wird G in S_4 als (kommutative) Untergruppe eingebettet.

Satz von Caley: Bedeutung

- Gruppen können sehr kompliziert sein.
- Jede Gruppe ist (isomorph zu) einer Untergruppe einer S_A. Um Gruppen zu verstehen "genügt" es S_A (und alle Untergruppen) zu verstehen.
- Wir können eine Gruppe der Größe n immer in S_n einbetten.
- Dadurch bekommen wir aber nichts geschenkt:
  S_n ist kompliziert und riesig:
  $|S_n| = n! \sim (\frac{n}{e})^n \gg 2^n$.

Zyklen

- Eine Permutation $f \in S_A$ heißt Zyklus (der Länge m) wenn es paarweise verschiedene $a_1, a_2, \ldots, a_m$ in A gibt so dass: $f(a_i) = a_{i+1}$ falls $1 \leq i < m$, $f(a_m) = a_1$, und $f(b) = b$ falls $b \notin \{a_1, \ldots, a_m\}$. Dabei heißt $\{a_1, \ldots, a_m\}$ Träger von f, oder $\mathrm{Tr}(f)$.
- In der bekannten Schreibweise sieht ein Zyklus in S_n etwa so aus:
$$f = \begin{pmatrix} a_2 & b & a_1 & b' & \ldots & a_m & a_{m-1} & b'' \\ a_3 & b & a_2 & b' & \ldots & a_1 & a_m & b'' \end{pmatrix}$$
- Neue Schreibweise: ("Zyklenschreibweise"): $f = (a_1 \ldots a_m)$.
- Bsp in S_3: Wir üblich $a := \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$ und $b := \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$.
Dann gilt: $a = (123) = (231) = (312)$ und $b = (12) = (21)$.
- Es gilt: $(a_1, a_2, \ldots, a_{m-1}, a_m) = (a_2, a_3, \ldots, a_m, a_1) = \ldots$;
und $(m) = \mathrm{Id}$ für alle $1 \leq m \leq n$.

Träger/Support

Für $f \in S_A$, setze $\mathrm{supp}(f) := \{a \in A : f(a) \neq a\}$.
Es gilt:

- $a \in \mathrm{supp}(f)$ gdw $f(a) \in \mathrm{supp}(f)$.
  Beweis: $f(a) \neq a$ (d.h. $a \in \mathrm{supp}(f)$) impliziert $f(f(a)) \neq f(a)$ (d.h. $f(a) \in \mathrm{supp}(f)$), weil f injektiv ist.
  Und wenn $a \notin \mathrm{supp}(f)$, d.h. $f(a) = a$, dann ist offenbar $f(a) = a \notin \mathrm{supp}(f)$. □
- Für einen Zyklus $f \in S_A$ der Länge $m > 1$ ist $\mathrm{supp}(f) = \mathrm{Tr}(f)$.
  Beweis: Für $a_\ell \in \mathrm{Tr}(f)$ ist $f(a_\ell) = a_{\ell+1} \neq a_\ell$ wenn $\ell < m$, und $a_1 \neq a_\ell$ wenn $\ell = m$ (weil $m > 1$). D.h. $a_\ell \in \mathrm{supp}(f)$,
  Wenn $a \notin \mathrm{Tr}(f)$, dann $f(a) = a$ nach Definition Zyklus. D.h. $a \notin \mathrm{supp}(f)$. □
- Für einen Zyklus $f \in S_A$ gilt: $a \in \mathrm{Tr}(f)$ gdw. $f(a) \in \mathrm{Tr}(f)$.
  Beweis: Falls der Zyklus Länge > 1 hat, dann $a \in \mathrm{Tr}(f)$ gdw $a \in \mathrm{supp}(f)$ gdw $f(a) \in \mathrm{supp}(f)$ gdw $g(a) \in \mathrm{Tr}(f)$. Zyklen der Länge 1 sind die Identität, d.h. $f(a) = a$. □

S_n: Noch mehr Zyklen

Theorem

Wenn $f, g \in S_A$ *mit* $\mathrm{supp}(f) \cap \mathrm{supp}(g) = \emptyset$, *dann* $f \circ g = g \circ f$.

Beweis:

- Sei $X = \mathrm{supp}(f)$ und $Y = \mathrm{supp}(g)$. Nach Voraussetzung ist $X \cap Y = \emptyset$.
- Wir wollen zeigen: $f \circ g = g \circ f$. D.h.: Für alle $m \in A$ gilt: $f(g(m)) = g(f(m))$.
- Fall 1: $m \in X$. Damit ist $m \notin Y$, d.h. $g(m) = m$, und damit $f(g(m)) = f(m)$. Außerdem ist mit m auch $f(m) \in X$ und daher $f(m) \notin Y$, daher ist $g(f(m)) = f(m)$.
- Fall 2: $m \in Y$: Analog (tausche g und f aus).
- Fall 3: $m \notin A \cup B$: Dann $f(m) = g(m) = m$, und damit auch $f(g(m)) = g(f(m)) = m$.

Disjkunkte Zyklen kommutieren

- Sei $f \in S_n$ und $1 \leq m \leq n$. Betrachte die Folge
$$m,\ f(m),\ f(f(m)),\ \ldots$$
Wir setzen $f^{(k)}(m) := f(f(\cdots(m)\cdots))$ (k mal), und $f^{(0)}(m) := m$.
- Nach spätestens n Schritten muss ein Folgenelement erneut in der Folge auftauchen. Sei $1 \leq k \leq n-1$ die kleinste Zahl mit $f^{(k)}(m) = f^{(l)}(m)$ für ein $l < k$.
- Weil f injektiv ist, muss $l = 0$ sein, d.h. $f^{(k)} = m$: Ansonsten wäre $f(f^{(\ell-1)}(m)) = f(f^{(k-1)}(m))$, d.h. $f^{(\ell-1)}(m) = f^{(k-1)}(m)$ und k wäre nicht minimal.
- f "enthält" also den Zyklus $g := (m\, f(m) \cdots f^{(k-1)}(m))$, und $f = g \circ h = h \circ g$, wobei $\mathrm{supp}(h) \cap \mathrm{Tr}(g) = \emptyset$.
Beweis: $\ell \in \mathrm{Tr}(g)$ gdw $f(\ell) \in \mathrm{Tr}(g)$, nach Konstruktion von g (und weil f injektiv ist).
Sei $h := g^{-1} \circ f$. (In der Gruppe S_n existiert das ja.) Wenn $\ell \in \mathrm{Tr}(g)$, dann ist $f(\ell) = g(\ell)$, und $h(\ell) = \ell$; d.h. $\ell \notin \mathrm{supp}(h)$. □

Zyklendarstellung

Theorem

Jedes $f \in S_n$ lässt sich (eindeutig bis auf Reihenfolge) als Produkt $f = f_1 \circ f_2 \circ \cdots \circ f_m$ von Zyklen f_i mit disjunkten Trägern schreiben, so dass $\bigcup_{1 \leq i \leq m} \mathrm{Tr}(f_i) = \{1, \ldots, n\}$.

Zyklendarstellung: Beweis 1/2

Existenz:

- Setze $m_0 := 1$. Sei g_1 der Zyklus in f der m_0 enthält; $f = g_1 \circ h_1$ mit $h_1 := g_1^{-1} \circ f$, $\mathrm{supp}(h_1) \cap \mathrm{Tr}(g_1) = \emptyset$ und $\mathrm{supp}(h_1) \subseteq \mathrm{supp}(f)$.
- Wenn g_1 bereits ganz f ist (oder äquivalent: $h_1 = \mathrm{Id}$), dann sind wir fertig.
- Nehmen wir also an $h_1 \neq \mathrm{Id}$, d.h. es gibt ein $1 \leq m_1 \leq n$ in $\mathrm{supp}(h_1)$. Sei g_2 der Zyklus der m_1 enthält, und schreibe $h_1 = g_2 \circ h_2$, d.h. $f = g_1 \circ g_2 \circ h_2$, mit $\mathrm{supp}(h_2) \cap (\mathrm{Tr}(g_1) \cup \mathrm{Tr}(g_2)) = \emptyset$. $\mathrm{supp}(h_2) \subseteq \mathrm{supp}(h_1)$ und $|\mathrm{supp}(h_2)| < |\mathrm{supp}(h_1)|$,weil $m_1 \notin \mathrm{supp}(h_2)$.
- Nach ℓ Schritten haben wir also $f = g_1 \circ g_2 \circ \cdots \circ g_\ell \circ h_\ell$, und $\mathrm{supp}(h_1) \cap \bigcup_{1 \leq i \leq \ell} \mathrm{Tr}(g_i)$=0. Wenn $h_\ell = \mathrm{Id}$ ist sind wir fertig, sonst machen wir weiter wie zuvor.
- Nach endlich vielen Schritten muss $h_\ell = \mathrm{Id}$ sein, weil $|\mathrm{supp}(h_\ell)|$ in jedem Schritt echt kleiner wird.

Beweis (Forts.)

- Es ist also $f = g_1 \circ g_2 \circ \cdots \circ g_\ell$ ein Produkt von Zyklen mit disjunktem Träger. Damit die Träger $\{1, \ldots, n\}$ überdecken, schreiben wir $f = g_1 \circ g_2 \circ \cdots \circ g_\ell \circ (x_0) \circ \cdots \circ (x_i)$, wobei $\{x_0, \ldots, x_i\} = 1, \ldots, n \setminus \bigcup_{1 \leq i \leq \ell} \mathrm{Tr}(g_i)$.

Eindeutigkeit:

- Seien $f = g_1 \circ \cdots g_m = g'_1 \circ \cdots \circ g'_{m'}$ zwei Zyklendarstellungen.
- Sei $\ell \in n$ beliebig. Dann ist $\ell \in \mathrm{Tr}(g_a)$ und $\ell \in \mathrm{Tr}(g'_b)$ für (eindeutige) $1 \leq a \leq m$ und $1 \leq b \leq m'$.
- Dann ist $g_a = g'_b =: \tilde{g}$, weil $f(\ell) = g_a(\ell) = g'_b(\ell)$, etc.
  (Die Träger aller g_i sind ja disjunkt, dasselbe gilt für die g'_i.)
- Wir sehen also: Für jedes j gibt es (genau) ein $\ell(j)$ so dass g_j und $g'_{\ell(j)}$ dieselben Zyklen sind. Und jedes g'_k ist gleich einem g_j. Daher sind beide Darstellungen gleich bis auf Umordnung der Zyklen. □

Noch mehr S_n

- Ein Zyklus der Länge 2 heißt auch Transposition, er ist das vertauschen zweier Positionen.
- Eine Transposition f ist eine ungerade Permutation, d.h. hat Vorzeichen $\text{sign}(f) = -1$.
- Der Zyklus $(a_1 \cdots a_i\, a_{i+1} \cdots a_m)$ mit $1 < i < m$ ist gleich $(a_1 \cdots a_i) \circ (a_i \cdots a_m)$.
  Beweis: Fallunterscheidungen.
- Ein Zyklus $f = (a_1 \cdots a_m)$ der Länge m läßt sich daher in $m-1$ viele Transpositionen zerlegen: $f = (a_1\, a_2) \circ (a_2\, a_3) \circ \cdots \circ (a_{m-1}\, a_m)$.
  Bemerkung: Diese Transpositionen haben nicht disjunkten Träger und kommutieren nicht.
- Das Vorzeichen ist ein Homomorphismus, ein Zyklus f der Länge m hat also Vorzeichen $\text{sign}(f) = (-1)^{m-1}$.
  f ist also eine gerade Permutation gdw. m ungerade ist.
- Jede Permutation kann also Produkt von Transpositionen geschrieben werden. (Natürlich nicht mit disjunktem Träger.)

Und noch ein bisschen S_n

Sei $f = g_1 \circ \cdots \circ g_m$ das Produkt disjunkter Zyklen g_i der Länge l_i. Dann gilt:

- $\operatorname{sgn}(f) = \prod_{i=1}^{m} (-1)^{l_1 - 1}$, d.h.
  f ist ungerade gdw. eine ungerade Anzahl von l_1 gerade Zahlen sind.
- $f^k = g_1^k \circ g_m^k$ (aber i.A. sind die g_i^k keine Zyklen mehr).

  Beweis: $f^2 := f \circ f = (g_1 \circ \cdots \circ g_m) \circ (g_1 \circ \cdots \circ g_m) = g_1^2 \circ \cdots \circ c_m^2$, weil die g_i disjunkten Support haben und deshalb kommutieren.
  Für jede Permutation h gilt: $\operatorname{supp}(h^k) \subseteq \operatorname{supp}(h)$. Daher ist $f^3 = f^2 \circ f = (g_1^2 \circ \cdots \circ g_m^2) \circ (g_1 \circ \cdots \circ g_m) = g_1^3 \circ \cdots \circ c_m^3$, etc. □
- Bsp in S_4: $(1\,2\,3\,4)^2 = (1\,3)(2\,4)$ ist kein Zyklus mehr.
- Die Ordnung des Gruppenelements $f \in S_n$ ist $\operatorname{kgV}(l_1, \ldots, l_m)$.

  Beweis: $g_i^k = \mathrm{Id}$ gdw k Vielfaches von l_i ist. $f^k = g_1^k \circ \cdots \circ g_m^k$ ist die Identität gdw jedes g_i^k die Identität ist, d.h. gdw wenn k Vielfaches von jedem l_i ist. □

S_4

- Sei $x = (1\,2)(3\,4)$ und $y = (1\,2\,3)$ (in S_4).
- In Zwei-Zeilen-Schreibweise: $x = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix}$, $y = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 1 & 4 \end{pmatrix}$.
- x ist gerade, weil es das Produkt zweier Transpositionen ist: $\operatorname{sgn}(x) = (-1) \cdot (-1) = 1$.
- y ist gerade weil es Zyklus der Länge 3 ist: $\operatorname{sgn}(y) = (-1)^{3-1} = 1$.
- $x \circ y = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 4 & 2 & 3 \end{pmatrix}$, $y \circ x = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 3 & 1 \end{pmatrix}$.
- A_4 ist also nicht kommutativ (und insbesondere nicht zyklisch, im Unterschied zu $A_3 = \langle a \rangle$):
  x und y sind in A_4, und $x \circ y \neq y \circ x$.

Automorphismen (Gruppen)

- Ein Isomorphismus $A \to A$ heißt Automorphismus.
- Manche Strukturen haben viele Automorphismen, machen nur die Identität (sie sind "starr').
- Bsp: Die Gruppe $(\mathbb{Z}, +)$ hat nur die beiden Automorphismen $x \mapsto x$ und $x \mapsto -x$.

  Beweis: $f(k) = f(1) + \cdots + f(1) = k \cdot f(1)$. Und wenn f surjektiv ist, dann ist $1 = k \cdot f(1)$ für ein k, das geht nur mit $f(1) = \pm 1$.
- Bsp: Die Gruppe $(\mathbb{Q}, +)$ hat viele Automorphismen,

  z.B. $f(r) := q \cdot r$ für $q \neq 0$.

Automorphismen (Ringe/Körper)

- $\mathbb{Z}$ als Ring ist starr.

  Beweis: Ein Ring-Homomorphismus bildet per Definition 1 auf 1 ab, und jedes $n \in \mathbb{Z} \setminus \{0\}$ hat die Form $\pm(1 + \cdots + 1)$.
- $\mathbb{Q}$ ist starr (als Ring, oder äquivalent: als Körper).
  Beweis: Jedes $r \in Q$ hat die Form $\pm(1 + \cdots + 1)/(1 + \cdots + 1)$.
- $\mathbb{R}$ ist starr. (Beweis nächste Folie)
- $\mathbb{C}$ hat den (wichtigen) Automorphismus $z \mapsto \bar{z}$ (und noch viele weitere "wilde" Automorphismen, die aber alle nicht definierbar bzw. konstruierbar sind).

$\mathbb{R}$ ist starr

Theorem

Wenn $f : \mathbb{R} \to \mathbb{R}$ ein Automorphismus ist, dann ist $f = \mathrm{Id}$.

Beweis:

- Wir wissen bereits: $f(q) = q$ für $q \in \mathbb{Q}$
- Wenn $x \geq 0$, dann $x = y^2$ für ein y^2, d.h. $f(x) = f(y)^2$, daher $f(x) \geq 0$.
- Wenn $x \leq y$, dann $y - x \geq 0$, d.h. $f(y - x) \geq 0$, d.h. $f(x) \leq f(y)$.
- $f : \mathbb{R} \to \mathbb{R}$ ist also monoton steigend, und $f \restriction \mathbb{Q} = \mathrm{Id}$.
- Angenommen $f(r) \neq r$ für ein $r \in \mathbb{R}$.
  OBdA $r < f(r)$ (anderer Fall analog).
  Wähle $r < q < f(r)$ mit q rational.
  Dann $f(r) \leq f(q_1) = q_1$, Widerspruch. □

Vektorräume

Vektoren in Ebene und Raum

- Ein klassisches Beispiel für einen Vektorraum: Die Ebene.
  Genauer: Nicht die Punkte der Ebene, sondern Vektoren $\vec{v}$ in der Ebene, das sind "Pfeile" (mit Länge und Richtung) bzw. (gerichtete) "Strecken zwischen zwei Punkten".
- Diese Ebene können wir (via kartesische Koordinaten) mit $\mathbb{R}^2$ identifizieren: $\vec{v} = (x, y)$, mit $x, y \in \mathbb{R}$.
  Es gibt aber verschiedene Arten diese Identifizierung vorzunehmen ($\sim$verschiedene Koordinatensysteme), je nachdem wie wir die Achsen legen und welche Skala wir wählen.
- Bsp (siehe Tafel): Derselbe Vektor $\vec{v}_1$ kann in einem Koordinatensystem (KS1) $(1, 0)$ sein, in einem anderen (KS2) $(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}})$.
- Analog können wir die Vektoren im (3-dimensionalen) Raum mit den Elementen von $\mathbb{R}^3$ identifizieren.
- Notation: Manchmal schreibt man Vektor als $\vec{v}$ oder $\mathbf{v}$.
  Wir werden nur manchmal z.B. $\vec{e}$ verwenden um zu betonen dass $\vec{e}$ ein Vektor ist, meist aber auch einfach nur e.

Bsp: Drei Koordinatensysteme der Ebene

Koordinaten eines Vektors sind abhängig vom gewählten Koordinatensystem:

- KS1:
  e_1 hat Koordinaten $(1, 0)$ und e_2 hat $(0, 1)$.
  Vektor a hat Koordinaten $a = (3, 7)$, d.h. $a = 3e_1 + 7e_2$.
- KS2: 45° verdreht.
  e_1 hat in diesem Koordinatensystem die Koordinaten $(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}) = \frac{1}{\sqrt{2}}(1, -1)$, und e_2 hat Koordinaten $(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}) = \frac{1}{\sqrt{2}}(1, 1)$.
  $a = 3e_1 + 7e_2$ hat Koordinaten $\frac{1}{\sqrt{2}}\big((3(1, -1) + 7(1, 1))\big) = \frac{1}{\sqrt{2}}(10, 4)$.
- KS3: Ausgehend von KS1: x-Skala um Faktor 2 "gestaucht":
  e_1 hat Koordinaten $(2, 0)$ und e_2 hat $(0, 1)$.
  $a = 3e_1 + 7e_2$ hat Koordinaten $3(2, 0) + 7(0, 1) = (6, 7)$.

Vektoren in Ebene und Raum (Forts.)

- Vektoren lassen sich addieren und mit reellen Zahlen multiplizieren: Addition ist das "Aneinanderfügen" von Vektoren; Multiplikation ist das "Strecken" mit einem reellen Faktor.
- Im Bsp der Ebene: Nach Identifikation mit $\mathbb{R}^2$ (nach Wahl eines Koordinatensystems) entspricht die Vektor-Addition der komponentenweisen Addition, und die Multiplikation der komponentenweisen Multiplikation (mit einer reellen Zahl).
- Mit der Addition bilden die Vektoren eine Gruppe; zusammen mit der "Multiplikation mit reellen Zahlen" ergibt das eine neue Struktur, die wir jetzt allgemeiner definieren werden.

Vektorräume und Basen

Definition

Sei K ein Körper. $(V, +, \cdot)$ ist ein Vektorraum über K, wenn $(V, +)$ eine abelsche Gruppe ist, und wenn für $\lambda, \mu \in V$ und $v, v_1, v_2 \in V$ gilt:

1. $\cdot : K \times V \to V$
2. $\lambda \cdot (v_1 + v_2) = \lambda \cdot v_1 + \lambda \cdot v_2$,
3. $(\lambda + \mu) \cdot v = \lambda \cdot v + \mu \cdot v$,
4. $(\lambda\mu) \cdot v = \lambda \cdot (\mu \cdot v)$,
5. $1 \cdot v = v$.

(Statt $\lambda \cdot v$ schreiben wir auch einfach λv.)

Ein wichtiger Vektorraum ist $K^n = K \times \cdots \times K$, mit der punktweisen Addition und Multiplikation, d.h., $\lambda \cdot (a_1, \ldots, a_n) := (\lambda a_1, \ldots, \lambda a_n)$ und $(a_1, \ldots, a_n) + (b_1, \ldots, b_n) := (a_1 + b_1, \ldots, a_n + b_n)$.

Basen

Im K^n setzten wir $e_i := (0, \ldots, 0, 1, 0, \ldots 0) \in K^n$ (ein 1 an i-ter Stelle).

Dann ist $(a_1, \ldots, a_n) = \sum_{i=1}^{n} a_i \cdot e_i$. Jedes Element von K^n lässt sich (auf eindeutige Weise) als eine solche Summe schreiben.

Definition

$(b_1, \ldots, b_n)$ ist eine (endliche) Basis von V, wenn $b_i \in V$ und wenn sich jedes $v \in V$ auf eindeutige Weise als $v = \sum_{i=1}^{n} \lambda_i b_i$ schreiben lässt (mit $\lambda_i \in K$). ("Basisdarstellung")
Wenn V so eine Basis hat nennen wir V endlichdimensional bzw. n-dimensional.

(Funktioniert genauso im unendlichen; um Notation zu vereinfachen hier nur endlichdimenionsle VR.)

- $(e_i)_{1 \leq i \leq n}$ ist also ein Basis in K^n; und K^n ist n-dimensional.
- Es gib aber auch viele andere Basen ($\sim$Koordinatensysteme) von K^n.

Beispiel von Basen in der Ebene

Wir haben drei Koordinatensysteme der Ebene erwähnt, KS1, KS2 und KS3. Das entspricht drei Basen in $\mathbb{R}^2$:

- (KS1) Sei $e_1 = (1, 0)$ und $e_2 = (0, 1)$ ("Standardbasis").
- (KS2, 45° gedreht): $e_1' := \frac{1}{\sqrt{2}}(1, 1)$, und $e_2' := \frac{1}{\sqrt{2}}(-1, 1)$.

  (e_1', e_2') ist ebenfalls Basis.

  Z.B. ist $e_1 = \frac{1}{\sqrt{2}} e_1' - \frac{1}{\sqrt{2}} e_2'$, und $e_2 = \frac{1}{\sqrt{2}} e_1' + \frac{1}{\sqrt{2}} e_2'$.
- (KS3, x-Skala gestaucht) $e_1'' := \frac{1}{2} e_1$, $e_2'' := e_2$

  (e_1'', e_2'') ist ebenfalls Basis.

  Z.B. ist $e_1 = 2e_1''$.

Vektorraumoperationen in Basisdarstellung

Wir haben K^n so definiert dass $+$ und $\cdot\lambda$ Komponentenweise wirkt.
Dasselbe gilt für jede Basis eines beliebigen Vektorraums:
Sei $(b_1, \ldots, b_n)$ Basis von V.

- Wenn v die Basissdarstellung $v = \sum_{i=1}^n \lambda_i b_i$ hat, dann hat μv die Basissdarstellung $\mu v = \sum_{i=1}^n (\mu\lambda_i) b_i$.

  Beweis: Es gilt
  $\sum_{i=1}^n (\mu\lambda_i) b_i \overset{\text{VR 4}}{=} \sum_{i=1}^n \mu(\lambda_i b_i) \overset{\text{VR 2}}{=} \mu \sum_{i=1}^n \lambda_i b_i = \mu v$ (und die Basisdarstellung ist eindeutig). □

- Wenn zusätzlich $v' = \sum_{i=1}^n \lambda'_i b_i$, dann hat $v + v'$ die Darstellung $v + v' = \sum_{i=1}^n (\lambda_i + \lambda'_i) b_i$.

  Beweis: $\sum_{i=1}^n (\lambda_i + \lambda'_i) b_i \overset{\text{VR 3}}{=} \sum_{i=1}^n (\lambda_i b_i + \lambda'_i b_i) \overset{\text{abel. Gr.}}{=}$
  $\sum_{i=1}^n (\lambda_i b_i) + \sum_{i=1}^n (\lambda'_i b_i) = v + v'$. □

Spaltenvektoren

- Wenn wir eine Basis $e_1, \ldots, e_n$ von V fixieren, entspricht also jeder Vektor v genau einer Folge $(\lambda_i)_{1 \leq i \leq n}$,
  und zwar über $v = \sum_{i=1}^{n} \lambda_i e_i$.
- Wir schreiben dann v oft als "Spaltenvektor": $v = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}$
- Der Vektor μv hat dann Koordinatendarstellung

$$\mu v = \mu \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} = \begin{pmatrix} \mu\lambda_1 \\ \vdots \\ \mu\lambda_n \end{pmatrix}$$

Beispiel

- Die "kanonische Basis" des $\mathbb{R}^2$ ist e_1, e_2.
- Dieser Basis gilt (per definition) $e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ und $e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.
- Die Basis, die dem 45° gedrehten KS2 entspricht, ist e_1', e_2', mit $e_1' = \frac{1}{\sqrt{2}} e_1 + \frac{1}{\sqrt{2}} e_2$ und $e_2' = -\frac{1}{\sqrt{2}} e_1 + \frac{1}{\sqrt{2}} e_2$.
- In der e_1, e_2-Basisdarstellung: $e_1' = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, und $e_2' = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \end{pmatrix}$.
- Es gilt: $e_1' + e_2' = 2\frac{1}{\sqrt{2}} e_2 = \sqrt{2} e_2$, und $e_1' - e_2' = \sqrt{2} e_1$. Daher:
- $e_1 = \frac{1}{\sqrt{2}}(e_1' - e_2')$ und $e_2 = \frac{1}{\sqrt{2}}(e_2' + e_2')$
- In der e_1', e_2'-Basisdarstellung ist also

  $e_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ und $e_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$; und $e_1' = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ und $e_2' = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

Vektorräume sind keine Ringe

Bemerkung: Auf einem Vektorraum haben wir i.A. keine Ringstruktur!
Wir multiplizieren keine Vektoren um neue Vektoren zu erhalten!

- In $V = \mathbb{R}^n$ (und anderen Räumen) kann man zwar das wichtige "Innenprodukt" $\vec{a} \cdot \vec{b}$ definieren, aber das ist eine Zahl (in $\mathbb{R}$) und kein Vektor. D.h.: $\cdot : V \times V \to \mathbb{R}$.
- In $V = \mathbb{R}^3$ kann man $\vec{a} \times \vec{b} \in V$ definieren, aber $\times$ ist nicht assoziativ etc. (und das funktioniert auch nur in 3 Dimensionen).
- Wir wissen bereits dass wir auf $\mathbb{R}^n$ (oder jedem K^n) eine Ringstruktur definieren könnten, indem wir komponentenweise multiplizieren.
  Das ist aber für Vektoren kein sinnvolle/interessante Operation.
  Insbesondere ist das abhängig von Wahl der Basis. Bsp:

  Setze $1 = e_1$ und $b = e_2$.
  In KS1: $a = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $b = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, das Produkt ist $\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \vec{0}$.
  In KS2: $a = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$, $a = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, das Produkt ist $\frac{1}{2} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \neq \vec{0}$.

Lineare Abbildungen und Matrizen

Lineare Abbildungen

Definition

Seien V_1, V_2 Vektorräume über K. Eine Lineare Abbildung $f : V_1 \to V_2$ ist ein Vektorraum-Homomorphismus, d.h.:
$f(v_1 + v_2) = f(v_1) + f(v_2)$, und $f(\lambda \cdot v) = \lambda \cdot f(v)$.

Wenn $(b_1, \ldots, b_n)$ eine Basis für V_1 ist, dann gilt:

- Ein lineare Abbildung $f : V_1 \to V_2$ ist festgelegt durch die Werte $w_i := f(b_i)$ für $1 \leq i \leq n$; und
- für jede Folge $(w_i)_{1 \leq i \leq n}$ in V_2 gibt es eine lineare Abbildung $f : V_1 \to V_2$ mit $f(b_i) = w_i$.

Beweis

- Z.Z.: Ein lineare Abbildung $f : V_1 \to V_2$ ist festgelegt durch die Werte $w_i := f(b_i)$ für $1 \leq i \leq n$.

  Beweis: Jedes $v \in V_1$ hat die Form $v = \sum_{i=1}^{n}(\lambda_i b_i)$, und daher $f(v) = f(\sum_{i=1}^{n} \lambda_i b_i) = \sum_{i=1}^{n} f(\lambda_i \cdot b_i) = \sum_{i=1}^{n} \lambda_i \cdot f(b_i)$. □

- Z.Z.: Für jede Folge $(w_i)_{1 \leq i \leq n}$ in V_2 gibt es eine lineare Abbildung $f : V_1 \to V_2$ mit $f(b_i) = w_i$.

  Beweis: Definiere $f(v) := \sum_{i=1}^{n} \lambda_i w_i$ wobei $v = \sum_{i=1}^{n} \lambda_i b_i$.
  Wir müssen zeigen: f ist linear.
  Wir wissen $\mu v = \sum_{i=1}^{n} \mu\lambda_i b_i$, d.h.
  $f(\mu v) := \sum_{i=1}^{n} \mu\lambda_i w_i = \mu(\sum_{i=1}^{n} \lambda_i w_i) = \mu f(v)$.
  Und wenn $v' = \sum_{i=1}^{n} \lambda'_i b_i$, dann ist $f(v + v') := \sum_{i=1}^{n}(\lambda_i + \lambda'_i)w_i = \sum_{i=1}^{n} \lambda_i w_i + \sum_{i=1}^{n} \lambda'_i w_i = f(v) + f(v')$. □

Matrizen

- Sei V_1 n-dimensional mit Basis $e_1, \ldots, e_n$ und V_2 m-dimensional mit Basis $e'_1, \ldots, e'_m$; und $f : V_1 \to V_2$ linear.
- f ist dann festgelegt durch $f(e_1), \ldots, f(e_n)$.
- Jedes $f(e_j)$ hat die Basisdarstellung $f(e_j) = \sum_{i=1}^{m} a_{i,j} e'_i$.
- Daher ist f festgelegt durch die $m \times n$-Matrix $A := (a_{i,j})_{1 \leq i \leq m, 1 \leq j \leq n}$:

$$f\left(\sum_{j=1}^{n} \lambda_j e_j\right) \overset{f \text{ lin.}}{=} \sum_{j=1}^{n} \lambda_j f(e_j) = \sum_{j=1}^{n} \lambda_j \sum_{i=1}^{m} a_{i,j} e'_i \overset{\text{VR } 2}{=}$$

$$\sum_{j=1}^{n} \sum_{i=1}^{m} \lambda_j a_{i,j} e'_i \overset{Abel.Gr.}{=} \sum_{i=1}^{m} \sum_{j=1}^{n} \left(\lambda_j a_{i,j} e'_i\right) \overset{\text{VR } 3}{=}$$

$$\sum_{i=1}^{m} \left(\sum_{j=1}^{n} \lambda_j a_{i,j}\right) e'_i.$$

Matrizen

- Wir haben also: Wenn wir Basen fixieren, können wir die lineare Abbildung f festlegen durch die $m \times n$-Matrix $A = (a_{i,j})_{1 \leq i \leq m, 1 \leq j \leq n}$ bzw

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} \end{pmatrix}, \text{ mit } f(e_i) = \begin{pmatrix} a_{1,i} \\ a_{2,i} \\ \vdots \\ a_{m,i} \end{pmatrix} \text{ (in } e'\text{-Basis).}$$

- Für $v = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$ (in e-Basis) ist dann

$$f(v) = A \cdot v = \begin{pmatrix} a_{1,1} \cdot b_1 & + & a_{1,2} \cdot b_2 & + \dots + & a_{1,n} \cdot b_n \\ a_{2,1} \cdot b_1 & + & a_{2,2} \cdot b_2 & + \dots + & a_{2,n} \cdot b_n \\ \vdots & & \vdots & \ddots & \vdots \\ a_{m,1} \cdot b_1 & + & a_{m,2} \cdot b_2 & + \dots + & a_{m,n} \cdot b_n \end{pmatrix} \text{ (in } e'\text{-Basis)}$$

Beispiel: Projektion

Sei f Projektion der Ebene auf die (KS1)-Diagonale, d.h. $f(e_1) = f(e_2) = \frac{1}{2}e_1 + \frac{1}{2}e_2$.

Arbeite mit KS1 (im Definitions- und Zielbereich):

- $f(e_1) = f(e_2) = \frac{1}{2}\begin{pmatrix}1\\1\end{pmatrix}$.
- Daher hat f die Darstellung $A = \frac{1}{2}\begin{pmatrix}1 & 1\\1 & 1\end{pmatrix}$.
- Es ist also z.B. $f(7e_1 - 2e_2)$ gleich

  $A \cdot \begin{pmatrix}7\\-2\end{pmatrix} = \frac{1}{2}\begin{pmatrix}1 & 1\\1 & 1\end{pmatrix} \cdot \begin{pmatrix}7\\-2\end{pmatrix} = \frac{5}{2}\begin{pmatrix}1\\1\end{pmatrix}$, und $A \cdot \begin{pmatrix}3\\3\end{pmatrix} = \begin{pmatrix}3\\3\end{pmatrix}$.
- Wir werden später folgende Größen untersuchen:

  Determinante $\det(A)$, bei 2×2 Matrix A ist das $a_{0,0}a_{1,1} - a_{1,0}a_{0,1}$.
  Hier also $\frac{1}{2} \cdot \frac{1}{2} - \frac{1}{2} \cdot \frac{1}{2} = 0$

  Spur $\mathrm{Tr}(A)$, bei 2×2 Matrix A ist das $a_{0,0} + a_{1,1}$.
  Hier also $\frac{1}{2} + \frac{1}{2} = 1$.

Beispiel: Projektion in KS2

Betrachte nun dieselbe Projektion f in KS2, d.h. verwende die e'-Basis (45° gedreht zu KS1).

- Es gilt $f(e'_1) = e'_1$ und $f(e'_2) = 0$.
- Daher hat f im KS2 Darstellung $A' = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$.
- Die Determinante und die Spur bleiben aber gleich.
  (Wir werden sehen: Determinante und die Spur sind Eigenschaften der Abbildung, nicht ihrer Koordinatendarstellung):

  $\det(A') = 1 \cdot 0 - 0 \cdot 0 = 0$, und
  $\mathrm{Tr}(A) = 1 + 0 = 1$.

Eigenvektoren (Am Bsp. Projektion)

Definition

Ein Vektor $v \neq 0$ heißt Eigenvektor der linearen Abbildung f, wenn $f(v) = \lambda v$ für ein $\lambda \in K$ (dieses λ heißt Eigenwert).

In unserem Beispiel der Projektion auf die Diagonale:

- $f(e_1') = e_1'$ und $f(e_2') = 0$; d.h.: e_1' ist Eigenvektor zum Eigenwert 1 und e_2' ist Eigenvektor zum Eigenwert 0.
- Das entspricht der Darstellung $A' = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ in (e_1', e_2')-Koordinaten.
- (e_1', e_2') ist eine Basis aus Eigenvektoren der Abbildung f.
- Man kann sehen: $v \neq 0$ ist Eigenvektor gdw v Vielfaches von e_1' oder von e_2' ist; 0 und 1 sind die einzigen Eigenwerte.
- Oft sehr nützlich eine Basis aus Eigenvektoren zu finden (Bsp: Quantenmechanik).

Beispiel: Drehung

Sei f Drehung der Ebene um $90° = \frac{\pi}{2}$. D.h., $f(e_1) = e_2$ und $f(e_2) = -e_1$.

Verwende (in Definitions- und Zielbereich) Standard e_1, e_2-Basis (KS1):

- $f\begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}0\\1\end{pmatrix}$, $f\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}-1\\0\end{pmatrix}$, d.h. f entspricht $A = \begin{pmatrix}0 & -1\\1 & 0\end{pmatrix}$.
- Z.B. ist $f(7e_1 - 2e_2)$ gleich $A \cdot \begin{pmatrix}7\\-2\end{pmatrix} = \begin{pmatrix}0 & -1\\1 & 0\end{pmatrix} \cdot \begin{pmatrix}7\\-2\end{pmatrix} = \begin{pmatrix}2\\7\end{pmatrix}$.
- Determinante $\det(A) = a_{0,0}a_{1,1} - a_{1,0}a_{0,1}$ ist hier $0 \cdot 0 - (-1) \cdot 1 = 1$.
- Spur $\mathrm{Tr}(A) = a_{0,0} + a_{1,1}$ ist hier $0 + 0 = 0$.
- f hat keine Eigenvektoren (in $\mathbb{R}^2$).

  Aber: In $\mathbb{C}^2$ gibt es Eigenvektoren zu den Eigenwerten $\pm i$:

$$\begin{pmatrix}0 & -1\\1 & 0\end{pmatrix}\begin{pmatrix}1\\\pm i\end{pmatrix} = \begin{pmatrix}\mp i\\1\end{pmatrix} = \mp i\begin{pmatrix}1\\\pm i\end{pmatrix}$$

Wiederholung

Sei $f : V_1 \to V_2$ linear, V_1 n-dimensional, V_2 m-dimensional. Seien Basen fixiert. Dann entspricht f der $m \times n$-Matrix $A = (a_{i,j})_{1 \leq i \leq m, 1 \leq j \leq n}$ bzw

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} \end{pmatrix}, \text{ mit } f(e_i) = \begin{pmatrix} a_{1,i} \\ a_{2,i} \\ \vdots \\ a_{m,i} \end{pmatrix}$$

$$\text{Für } v = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} \text{ ist dann}$$

$$f(v) = A \cdot v = \begin{pmatrix} a_{1,1} \cdot b_1 & + & a_{1,2} \cdot b_2 & + \dots + & a_{1,n} \cdot b_n \\ a_{2,1} \cdot b_1 & + & a_{2,2} \cdot b_2 & + \dots + & a_{2,n} \cdot b_n \\ \vdots & & \vdots & \ddots & \vdots \\ a_{m,1} \cdot b_1 & + & a_{m,2} \cdot b_2 & + \dots + & a_{m,n} \cdot b_n \end{pmatrix}$$

Wiederholung 2

- Bsp: $f : \mathbb{R}^2 \to \mathbb{R}^2$ Projektion auf Diagonale hat in Standardbasis KS1 Darstellung $\begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$
- In gedrehter Basis KS2 Darstellung $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$
- In jedem Fall: Determinante 0, Spur 1, Eigenwerte 0 und 1.

Wiederholung 3

- Bsp: $f : \mathbb{R}^2 \to \mathbb{R}^2$ Rotation um $90° = \frac{\pi}{2}$ hat in Standardbasis Darstellung $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$.
- Determinante ist 1, Spur 0.
- f hat keine Eigenwerte (in $\mathbb{R}$. Die Rotation $f : \mathbb{C}^2 \to \mathbb{C}^2$ mit derselben Matrix hat Eigenwerte $\pm i$.)

Beispiel: Drehung in anderen Koordinatensystemen

- Drehung f um $\frac{\pi}{2}$ hat also in KS1 Form $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$.
- f hat dieselbe Form in KS2 (warum?).
- In KS3 (mit $e_1'' = \frac{1}{2}e_1$ und $e_2'' = e_2$):
  $f(e_1'') = \frac{1}{2}f(e_1) = \frac{1}{2}f(e_2) = \frac{1}{2}f(e_2'')$, und
  $f(e_2'') = f(e_2) = -e_1 = -2e_1''$.
- In KS3 gilt also:
  $f\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ \frac{1}{2} \end{pmatrix}$, $f\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -2 \\ 0 \end{pmatrix}$, d.h. f entspricht $A'' = \begin{pmatrix} 0 & -2 \\ \frac{1}{2} & 0 \end{pmatrix}$.
- Determinante und Spur ändern sich nicht:
  $\det(A'') = 0 \cdot 0 - \frac{1}{2} \cdot (-2) = -1$,
  $\mathrm{Tr}(A) = 0 + 0 = 0$.

Summe und Vielfache linearer Abbildungen

Seien f, f_1, f_2 lineare Funktionen $V_1 \to V_2$.

- Die Summe $f_1 + f_2 : V_1 \to V_2$, d.h. $v \mapsto f_1(v) + f_2(v)$ ist linear.

  Beweis:
  $(f_1 + f_2)(v + w) := f_1(v + w) + f_2(v + w) \overset{f_i \text{ lin.}}{=} f_1(v) + f_1(w) + f_2(v) + f_2(w) \overset{\text{komm.}}{=} f_1(v) + f_2(v) + f_1(w) + f_2(w) \overset{f_i \text{ lin.}}{=} (f_1 + f_2)(v) + (f_1 + f_2)(w)$.

  $(f_1 + f_2)(\lambda v) := f_1(\lambda v) + f_2(\lambda v) \overset{f_i \text{ lin.}}{=} \lambda f_1(v) + \lambda f_2(v) \overset{\text{VR } 2}{=} \lambda(f_1(v) + f_2(v)) =: \lambda(f_1 + f_2)(v)$. □

- Das lineare Vielfache $\lambda f : V_1 \to V_2$, d.h. $v \mapsto \lambda f(v)$, für $\lambda \in K$, ist linear.

  Beweis: Analog.

- Damit bilden die Linearen Funktionen von $V_1 \to V_2$ einen Vektorraum.

Der Vektorraum der linearen Funktionen

Seien V_1, V_2 zwei K-Vektorräume.

Theorem

Die Menge $L(V_1, V_2)$ aller linearen Funktionen $f : V_1 \to V_2$ bilden einen K-Vektorraum mit:
$(f_1 + f_2)(v) := f_1(v) + f_2(v)$ *und* $(\lambda f)(v) := \lambda(f(v))$.

Beweis:

- $(L, +)$ ist eine abelsche Gruppe: $f_1 + f_2 = f_2 + f_1$, $(-1)f$ ist das additive inverse, etc.

(2) $\lambda \cdot (f_1 + f_2) = \lambda \cdot f_1 + \lambda \cdot f_2$:
$(\lambda \cdot (f_1 + f_2))(v) := \lambda \cdot ((f_1 + f_2)(v)) = \lambda \cdot (f_1(v) + f_2(v)) =$
$(\lambda \cdot f_1(v)) + (\lambda \cdot f_2(v)) = (\lambda \cdot f_1)(v) + (\lambda \cdot f_2)(v) = (\lambda \cdot f_1 \lambda \cdot f_2)(v)$.

(3) $(\lambda + \mu) \cdot f = \lambda \cdot v + \mu \cdot f$ ähnlich,

(4) $(\lambda\mu) \cdot f = \lambda \cdot (\mu \cdot f)$ ähnlich,

(5) $1 \cdot f = f$, weil $(1 \cdot f)(v) := 1 \cdot (f(v)) = f(v)$.

Summe und Vielfache von Matrizen

Seien f_1, f_2 lineare Funktionen $V_1 \to V_2$, und fixiere (endliche) Basen für V_1 und V_2. Sei $A = (a_{i,j})$, $B = (b_{i,j})$die Matrixdarstellung von f_1, f_2.

- λf_1 hat Matrixdarstellung $\lambda A := (\lambda a_{i,j})$.

  Beweis: $f(e_i) = \begin{pmatrix} a_{1,i} \\ \vdots \\ a_{m,i} \end{pmatrix}$, dann ist $(\lambda f)(e_i) = \lambda(f(e_i)) = \begin{pmatrix} \lambda a_{1,i} \\ \vdots \\ \lambda a_{m,i} \end{pmatrix}$

- $f_1 + f_2$ hat Matrixdarstellung $A + B := (a_{i,j} + b_{i,j})$.

  Beweis: $f_1(e_i) = \begin{pmatrix} b_{1,i} \\ \vdots \\ b_{m,i} \end{pmatrix}$ und $f_2(e_i) = \begin{pmatrix} c_{1,i} \\ \vdots \\ c_{m,i} \end{pmatrix}$ dann ist

  $(f_1 + f_2)(e_i) = f_1(e_i) + f_2(e_i) = \begin{pmatrix} b_{1,i} + c_{1,i} \\ \vdots \\ b_{m,i} + c_{m,i} \end{pmatrix}$

Dimension von L

Theorem

Wenn V_1 n-dimensional ist und V_2 m-dimensional, dann ist $L(V_1, V_2)$ $n \cdot m$-dimensional.

Beweis: Fixiere Basen $(e_i)_{1 \leq i \leq n}$ für V_1 und $(e'_i)_{1 \leq i \leq m}$ für V_2. Dann bilden folgende $f^{i,j}$ eine Basis von $L(V_1, V_2)$, für $1 \leq i \leq n$ und $1 \leq j \leq m$:

$$f^{i,j}(e_\ell) = \begin{cases} e'_j & \text{für } \ell = i \\ 0 & \text{sonst.} \end{cases}$$

Äquivalent: Jedes $f \in L$ hat eine eindeutige Darstellung als $m \times n$-Matrix. Die folgenden Matrizen $A^{i,j} = (a^{i,j}_{\ell,k})$ bilden eine Basis der $m \times n$-Matrizen, mit

$$a^{i,j}_{\ell,m} = \begin{cases} 1 & \text{wenn } \ell = j \text{ und } m = i \\ 0 & \text{sonst.} \end{cases}$$

□

Verknüpfung linearer Abbildungen

- Die Verknüpfung von Homomorpismen ist Homomorphismus, insbes: die Verknüpfung linearer Funktionen ist eine lineare Funktion.
- In mehr Detail: Wenn $f_1 : V_1 \to V_2$ linear, und $f_2 : V_2 \to V_3$ linear, dann ist $f_2 \circ f_1 : V_1 \to V_3$ linear.

  Beweis: $(f_2 \circ f_1)(v + w) = f_2(f_1(v + w)) = f_2(f_1(v) + f_1(w)) = f_2(f_1(v)) + f_2(f_1(w)) = (f_2 \circ f_1)(v) + (f_2 \circ f_1)(w)$.
  $(f_2 \circ f_1)(\lambda v) = f_2(f_1(\lambda v)) = f_2(\lambda f_1(v)) = \lambda f_2(f_1(v)) = \lambda(f_2 \circ f_1)(v)$.
- Damit ist $L(V) := L(V, V)$, der Verktorraum der linearen Funktionen $V \to V$, auch ein Ring, mit den Operationen Addition und Verknüpfung und dem Einselement der Identität.

Der Ring $L(V, V)$

Theorem

$(L(V), +, \circ)$ ist ein Ring; die Identität ist das Einselement.

Beweis:

- Wir haben bereits gezeigt (bzw behauptet) dass $(L(V), +)$ eine abelsche Gruppe ist, weil $(L(V), +, \cdot_K)$ ja ein Vektorraum ist.
- Wir wissen bereits: Die Verknüpfung von Funktionen ist assoziativ.
- Es fehlen die Distributivgesetze: $f \circ (g + h) = (f \circ g) + (f \circ h)$. Beweis: $f \circ (g + h)(v) = f(g + h(v)) = f(g(v) + h(v)) = f(g(v) + f(h(v)) = (f \circ g)(v) + (f \circ h)(v)$.
- (rechte Distributivität analog) □

Die Identität, Diagonalmatrizen

- Die Identität $\mathrm{Id} : V \to V$ eines n-dimesionalen Raums ist linear und hat Matixdarstellung $I = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & 1 \end{pmatrix}$
- Determinante: $\det(I) = 1$, Spur: $\mathrm{Tr}(I) = n$, Jeder Vektor ist Eigenvektor zum Eigenwert 1.
- Eine $n \times n$-Matrix $A = (a_{i,j})$ heißt "Diagonalmatrix", wenn alle Einträge außerhalb der Diagonale Null sind; d.h. $a_{i,j} \neq 0$ impliziert $i = j$.
- Wenn $f : V \to V$ also durch die Diagonalmatrix A dargestellt wird, dann ist $f(e_i) = a_{i,i} e_i$. Die Vielfachen von e_i sind also Eigenvektoren zum Eigenwert $a_{i,i}$, die Determinante ist $\prod_{i=1}^{n} a_{i,i}$ und die Spur $\sum_{i=1}^{n} a_{i,i}$.

Matrizenmultiplikation

- Fixiere Basen $(e_i)_{1\leq i\leq n}$ für V_1, $(e'_i)_{1\leq i\leq m}$ für V_2 und $(e''_i)_{1\leq i\leq \ell}$ für V_3.
- Sei A die Matrixdarstellung von $f : V_1 \to V_2$ und B die von $g : V_2 \to V_3$. Setze $h := g \circ f$ mit Matrixdarstellung C.
- Nach der Definition der Matrixdarstellung ist $f(e_i)) = \sum_{j=1}^{m} a_{j,i} e'_j$, und $g(e'_j) = \sum_{k=1}^{\ell} b_{k,j} e''_k$.
- Daher ist $h(e_i) = g(f(e_i)) = g(\sum_{j=1}^{m} a_{j,i} e'_j) = \sum_{j=1}^{m} (a_{j,i} g(e'_j)) = \sum_{j=1}^{m} (a_{j,i} \sum_{k=1}^{\ell} b_{k,j} e''_k) = \sum_{k=1}^{\ell} (\sum_{j=1}^{m} a_{j,i} b_{k,j}) e''_k = \sum_{k=1}^{\ell} c_{k,i} e''_k$.
- D.h.: $c_{k,i} = \sum_{j=1}^{m} a_{j,i} b_{k,j}$.
  Auf diese Weise können wir C aus A und B berechnen.

$L(V)$ nicht kommutativ

- Der Ring $L(V)$ ist i.A. nicht kommutativ.
- Bsp im $\mathbb{R}^2$:
  (A) Projektion auf x-Achse gefolgt von Drehung ist ungleich
  (B) Drehung gefolgt von Projektion auf x-Achse:

  (A): e_1 bleibt unter Projektion e_1 und wird dann auf e_2 abgebildet.
  (B): e_1 wird zu e_2 rotiert und dann auf 0 projiziert.
- Als Matrizen (in Standardbasis):

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

ungleich

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

$L(V)$ hat Nullteiler

- Der Ring $L(V)$ ist i.A. nicht nullteilerfrei.
- Bsp im $\mathbb{R}^2$:
  Projektion auf x-Achse verknüpft mit Projektion auf y-Achse ist Null.
- Als Matrizen (in Standardbasis):

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

GL

- $f \in L(V)$ heißt invertierbar, wenn es ein $g \in L(V)$ gibt mit $f \circ g = g \circ f = \mathrm{Id}$.
  In dem Fall ist f eine Permutation (sogar ein Isomorphismus) von V, und g die (eindeutige) inverse Abbildung, geschrieben f^{-1}.
- In Matrixschreibweise bei n-dimensionalen Räumen:
  Eine $n \times n$-Matrix A ist invertierbar, wenn es eine Matrix B gibt mit $A \cdot B = B \cdot A = I$. Wir schreiben dieses (eindeutige) B als A^{-1}.
- Es gilt: Die Invertierbaren linearen Abbildungen bilden eine Gruppe, genennt $\mathrm{GL}(V)$.
  Beweis: Das gilt für jeden Ring $(R, +, \cdot)$: $(R, \cdot)$ ist assoziativ mit neutralem Element 1, die Teilmenge der invertierbaren Elemente ist abgeschlossen unter $\cdot$ und damit eine Gruppe: Wenn $r_1 s_1 = s_1 r_1 = 1$ und $r_2 s_2 = s_2 r_2 = 1$, dann ist $(r_1 r_2)(s_2 s_1) = (s_2 s_1)(r_1 r_2) = 1$.
- Für $f, g \in \mathrm{GL}(V)$ ist $(f \circ g)^{-1} = g^{-1} \circ f^{-1}$
- Die invertierbaren $n \times n$-Matrizen nennt man $\mathrm{GL}(n)$.
  Für $A, B \in \mathrm{GL}(n)$ gilt: $(A \cdot B)^{-1} = B^{-1} \cdot A^{-1}$.

Basiswechsel

Basiswechsel

- Seien $(e_i)_{i \in n}$ und $(e'_i)_{i \in n}$ zwei n-dimensionale Basen von V.
- Sei f die lineare Abbildung die jedes e_i auf e'_i abbildet, und U die entsprechende Matrix in e_i-Basis.
- Sei $v \in V$ in der e'_i-Basis durch $\vec{a} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$ dargestellt.
- D.h. $v = a_1 \cdot e'_1 + \cdots + a_n \cdot e'_n$.
- Daher wird v in der e_i-Basis durch $U\vec{a}$ dargestellt.
- U transformiert also e'_i-Koordinaten in e_i-Koordinaten.

Basiswechsel

- D.h.: Wenn $\vec{a}$ die Darstellung von v in e'_i-Koordinaten ist, dann ist $U\vec{a}$ die Darstellung in e_i-Koordinaten.
- f hat eine Umkehrabbildung g, mit $g(e'_i) = e_i$, mit Darstellung U^{-1} in der e_i-Basis.
- Habe $w \in V$ in der e_i-Basis die Darstellung $\vec{b}$.
- $U(U^{-1}\vec{b}) = \vec{b}$, d.h. der Vektor mit der e'_i-Darstellung $U^{-1}\vec{b}$ hat die e_i-Darstellung w.
- U^{-1} transformiert also e_i-Koordinaten nach e'_i-Koordinaten.

Beispiel

- Sei (e_1, e_2) die Standardbasis von $\mathbb{R}^2$ (wir haben das KS1) genannt, und (e_1', e_2') die um 45° gedrehte Basis:
  $e_1' = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $e_2' = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \end{pmatrix}$.
- $e_i \mapsto e_i'$ hat also e-Darstellung $U = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$
- $U^{-1} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$
- Ein Vektor mit KS1-Koordinaten $\vec{a} = \begin{pmatrix} x \\ y \end{pmatrix}$ hat KS2-Koordinaten
  $U^{-1}\vec{a} = \frac{1}{\sqrt{2}} \begin{pmatrix} x+y \\ -x+y \end{pmatrix}$

Notation Basisabhängige Darstellung

- Die Koordinaten (=Zeilenvektor) eines Vektors $v \in V$ hängt von der gewählten Basis $E = (e_1, \ldots, e_n)$ von V ab.
  Wir nennen diesen Zeilenvektor nun "E-Darstellung von v."
- Die Matrixdarstellung A einer linearen Funktion $f : V \to W$ hängt von den gewählten Basen $E = (e_1, \ldots, e_n)$ von V und $B = (b_1, \ldots, b_m)$ von W ab.
  Wir nennen nun die Matrix A die "(E, B)-Darstellung von f."
- Ein Sonderfall ist eine lineare Abbildung $f : V \to V$.
  Hier interessieren wir uns üblicherweise für den Fall $E = B$ (d.h. wir verwenden dieselbe Besis für den Definitions- und den Zielbereich), und nennen A dann einfach die "E-Darstellung."

Wiederholung: Basiswechsel

- Seien $E := (e_1, \ldots, e_n)$ und $E' = (e'_1, \ldots, e'_n)$ zwei Basen von V.
- Sei $\phi : V \to V$ linear mit $\phi(e_i) := e'_i$.
  Dann ist ϕ invertierbar, mit $\phi^{-1}(e'_i) = e_i$.
- Sei U und U^{-1} die E-Darstellungen von ϕ und ϕ^{-1}.
- Fixiere $v \in V$.
  Wenn $\vec{a}$ die E-Darstellung von v ist, dann ist $(U^{-1})\vec{a}$ die E'-Darstellung von v.
  Wenn $\vec{b}$ die E'-Darstellung von v ist, dann ist $U\vec{b}$ die E-Darstellung.

Beispiel

- Sei $E = (e_1, e_2)$ die Standardbasis von $\mathbb{R}^2$, und $E' = (e_1', e_2')$ die um $45°$ gedrehte Basis:
  $e_1' = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $e_2' = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \end{pmatrix}$.
- $e_i \mapsto e_i'$ hat also E-Darstellung $U = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$
- $U^{-1} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$
- Der Vektor mit E-Darstellung $\vec{a} = \begin{pmatrix} x \\ y \end{pmatrix}$ hat also E'-Darstellung
  $U^{-1}\vec{a} = \frac{1}{\sqrt{2}} \begin{pmatrix} x + y \\ -x + y \end{pmatrix}$

Ein 3D Beispiel

- First Person (Raumschiff oder dergl).
  Koordinatensystem $E = (e_1, e_2, e_3)$ entsprechend der Ausgangsposition: e_1 rechts, e_2 vorne, e_3 oben.
- r_z: Raumschiff dreht sich nach Links (um z-Achse) um $90°$.
  r_x: Raumschiff dreht sich nach oben (um x-Achse) um $90°$.
- Genauer:
  r_z definert durch $e_1 \mapsto e_2$, $e_2 \mapsto -e_1$ und $e_3 \mapsto e_3$;
  r_x durch $e_1 \mapsto e_1$, $e_2 \mapsto e_3$ und $e_3 \mapsto -e_2$.
- E-Darstellungen:

$$L = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} (\text{"left"}), \quad U = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} (\text{"up"})$$

Ein 3D Beispiel: Verknüpfung von Drehungen

- Achtung: $r_x \circ r_z$ ist "erst nach links, dann nach oben um "alte" x-Achse.
- In B-Darstellung:

$$L = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad U = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \quad LU = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \quad UL = \begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & -1 \\ 1 & 0 & 0 \end{pmatrix}$$

- Beachte: $r_z \circ r_x \neq r_x \circ r_z$. Im $\mathbb{R}^3$ kommutieren Drehungen i.A. nicht (im Gegensatz zu $\mathbb{R}^2$).

Ein 3D Beispiel: Vektoren unter Basiswechsel

- Nach Rotation "left" kann man Raumpunkten neue Koordinaten, entsprechend der neuen Orientierung, zuordnen. r_z^{-1} ist die Rotation zurück, d.h. nach rechts:
- $L = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, $L^{-1} = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$
- Ein Punkt der alte Koordinaten $\vec{a}$ hat, hat nach der Drehung neue Koordinaten $L^{-1}\vec{a}$:

  $$L^{-1} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} y \\ -x \\ z \end{pmatrix}.$$

- Drehung des Bezugsystems/des Raumschiffs/der first person nach links entspricht Drehung der Welt nach rechts.

Transformation von Matrizen: Basiswechsel Def.bereich

- Sei V ein Vektorraum mit Basis $E = (e_1, \ldots, e_n)$, und W ein Vektorraum mit Basis $B = (b_1, \ldots, b_m)$, und $f : V \to W$ mit (B, E)-Darstellung A.
- Sei $E' = (e'_1, \ldots, e'_n)$ eine andere Basis von V, sei $\phi : V \to V$ die lineare Abbildung $e_i \mapsto e'_i$ und sei U die E-Darstellung von ϕ.
- Wenn $\vec{a}$ die E'-Darstellung von v ist, dann ist $U\vec{a}$ die E-Darstellung von $v \in V$, und $AU\vec{a}$ daher die B-Darstellung von $f(v)$.
- Die (E', B)-Darstellung von f ist daher AU.

Transformation von Matrizen: Basiswechsel Ziel.bereich

- Ganz analog: Sei V ein Vektorraum mit Basis $E = (e_1, \ldots, e_n)$, und sei $B = (b_1, \ldots, b_m)$ Basis von W, und $g : V \to W$ mit (E, B)-Darstellung B.
- Sei $(b'_1, \ldots, b'_m)$ eine andere Basis von W, und $\psi : W \to W$ die Abbildung $b_i \mapsto b'_i$ mit B-Darstellung T.
- Wenn $\vec{a}$ die E-Darstellung von $v \in V$ ist, dann ist $B\vec{a}$ dir B-Darstellung von $g(v) \in W$, und daher $T^{-1}(B\vec{a})$ die B'-Darstellung von $g(v)$.
- Die (E, B')-Darstellung von g ist daher $T^{-1}B$.

Transformation von Matrizen: Basiswechsel $L(V)$

- Für $L(V)$, d.h. $f : V \to V$ ergibt das:
- Sei $E = (e_1, \ldots, e_n)$ Basis von V, und $f : V \to V$ mit Matrixdarstellung A.
- Sei $E' = (e'_1, \ldots, e'_n)$ eine andere Basis von V, und $\phi : V \to V$ die Abbildung $e_i \mapsto e'_i$ mit E-Darstellung U.
- Dann ist $U^{-1}AU$ die E'-Darstellung von f.
- Umgekehrt: Wenn B die E'-Darstellung von $g \in L(V)$ ist, dann ist UBU^{-1} die E-Darstellung.

Zusamenfassung

Seien E, E' eine Basis von V, und sei U die E-Darstellung der Basiswechsel-Abbildung $e_i \mapsto e'_i$.

- Wenn $\vec{a}$ die E-Darstellung von $v \in V$ ist, dann ist $U^{-1}\vec{a}$ die E'-Darstellung.
- Wenn $\vec{b}$ die E'-Darstellung von $v \in V$ ist, dann ist $U\vec{a}$ die E-Darstellung.
- Wenn A die E-Darstellung von $f \in L(V)$ ist, dann ist $U^{-1}AU$ die E'-Darstellung.
- Wenn B die E'-Darstellung von $g \in L(V)$ ist, dann ist UBU^{-1} die E-Darstellung.

Wieder das 3D Beispiel

- Wie zuvor: Raumschiff oder dergl, $E = (e_1, e_2, e_3)$ entspricht ursprünglichem rechts, vorne, oben.
- Wir wollen nun beschreiben: Erst Drehung nach links, und danach Drehung nach oben um neue x-Achse.
- Links: r_z, mit E-Darstellung L.
  Resultiert in neuer Basis $E' = (r_z(e_1), z_z(e_2), r_z(e_3))$.
  Basiswechsel-Abbildung ist r_z.
- Die zweite Drehung soll nun die E'-Darstellung U haben, und daher die E-Darstellung LUL^{-1}.
- Die Verknüpfung der E-Abbildungen ist daher $LUL^{-1}L = LU$.
- Vergleiche: "Erst links, danach oben um alte x-Achse" ergibt UL.
- Wir bekommen also:
  "Erst links dann oben alt": UL, "Erst oben dann links alt": LU,
  "Erst links dann oben neu": LU, "Erst oben dann links neu": UL.

Wieder das 3D Beispiel

- Allgemein gilt: "Erst Basiswechsel A, dann Basiswechsel der im neuen System Darstellung B hat", hat im alten System Darstellung $ABA^{-1}A = BA$.
- Weiteres Bsp: "Schrauben": Rotation um y-Achse, d.h.: $e_1 \mapsto -e_3$, $e_2 \mapsto e_2$, $e_3 \mapsto e_1$, mit Darstellung
$$R = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{pmatrix} \text{ ("roll")}$$
- Erst links dann oben dann schrauben (jeweils vom neuen Koordinatensystem) kann man wie folgt berechnen:
Links dann oben ist wie wir schon wissen LU.
Im neuen System ist schrauben $(LU)R(LU)^{-1}$.
Insgesamt bekommen wir $(LU)R(LU)^{-1}LU = LUR$.
- Im Übrigen: $LU = UR$, etc.
- Bemerkung: $(LU)^{-1} = U^{-1}L^{-1}$

Unterräume, Erzeugende, l.u., Basen

Unterräume

Definition

$W \subseteq V$ ist ein Unterraum, wenn $W \neq \emptyset$ und W abgeschlossen ist unter Addition $w_1 + w_2$ und unter $w \mapsto \lambda \cdot w$ für alle $\lambda \in K$.

- W enthält dann 0 (weil $0 = 0 \cdot v$ für jedes $v \in W$).
- $(W, +, \cdot)$ ist selbst ein Vektorraum.
  (Alle erforderlichen Eigenschaften gelten in V und daher auch in W.)

Beispiele:

- $\{0\}$ und V sind immer Unterräume von V ("triviale Unterräume").
- Die nicht-trivialen Unterräume von $\mathbb{R}^2$ sind die Geraden durch 0.
- Die nicht-trivialen Unterräume von $\mathbb{R}^2$ sind die Geraden und Ebenen durch 0.

Erzeugende

- Seien $v_j \in V$ und $\lambda_j \in K$ für $1 \leq j \leq \ell$.
  $\lambda_1 v_1 + \cdots + \lambda_\ell v_\ell = \sum_{j=1}^{\ell} \lambda_j v_j$ heißt eine Linearkombination der v_j.
- Für $X \subseteq V$ gibt es einen kleinsten Unterraum W, der X enthält, nämlich die Menge aller Linearkombinationen mit $v_j \in X$.
  Wir schreiben $W = \langle X \rangle$ (auch lineare Hülle von X genannt).

  D.h.: $W = \{0\} \cup \{\sum_{j=1}^{\ell} \lambda_j v_j : \ell \in \mathbb{N}, \lambda_j \in K, v_j \in X\}$.

  Beweis: W ist Teilraum, weil abgeschlossen unter $+$ und $\cdot_\lambda$. Wenn W' Teilraum ist mit $X \subseteq W'$, dann folgt $W \subseteq W'$, daher ist W kleinstmöglich. □
- Es gilt genauso:
  W ist der Schnitt aller Teilräume von V, die X enthalten ($X \subseteq V$).

  Beweis: Jeder Teilraum der X enthält, enthält auch W; der Schnitt enthält daher W. W selbst ist Teilraum der X enthält, es wird also u.A. mit W geschnitten, der Schnitt kann daher nicht größer sein. □

Erzeugende (Forts.)

- Wenn $v \in \langle A \rangle$, dann ist $\langle A \cup \{v\} \rangle = \langle A \rangle$.
- Allgemeiner: Wenn $B \subseteq \langle A \rangle$, dann ist $\langle A \cup B \rangle = \langle A \rangle$.
- Wir haben definiert: $E = (e_1, \ldots, e_n)$ ist eine Basis von V, wenn jedes $v \in V$ eine eindeutige Darstellung $v = \sum_{i=1}^{n} \lambda_i e_i$ hat. Insbesondere gilt dann: $\langle E \rangle = V$.
- Wenn E eine Basis ist, und $\langle A \rangle \neq V$, dann gibt es e_k mit $e_k \notin \langle A \rangle$.

Linear (Un)abhängig

- Eine Linearkombination $\sum_{j=1}^{\ell} \lambda_j v_j$ der Folge $(v_1, \ldots, v_\ell)$ heißt nicht-trivial, wenn mindestens ein $\lambda_j \neq 0$.
- Eine triviale Linearkombination ergibt immer 0.
- Eine Folge $(v_1, \ldots, v_\ell)$ heißt linear abhängig (l.a.), wenn 0 eine nicht-triviale Linearkombination $\sum_{j=1}^{\ell} \lambda_j v_j = 0$ ist (für irgendwelche $\lambda_j \in K$, mindestens einer davon $\neq 0$).
  Ansonsten heißt $(v_1, \ldots, v_\ell)$ linear unabhängig (l.u.).

  Eine (endliche) Menge A heißt l.a./l.u., wenn eine Aufzählung $(v_1, \ldots, v_\ell)$ von A l.a./l.u. ist.
- Es gilt: $(v_1, \ldots, v_\ell)$ l.a. gdw: $v_k \in \langle (v_j)_{1 \leq j \leq \ell, j \neq k} \rangle$ für ein $q \leq k \leq \ell$.

  Beweis: Wenn $v_k = \sum_{1 \leq j \leq \ell, j \neq k} \lambda_j v_j$, dann ist $0 = \sum_{1 \leq j \leq \ell} \lambda_j v_j$ mit $\lambda_k := -1$; damit ist $(v_1, \ldots, v_\ell)$ l.a.
  Wenn umgekehrt $0 = \sum_{1 \leq j \leq \ell} \lambda_j v_j$ mit $\lambda_k \neq 0$, dann $v_k = -\sum_{1 \leq j \leq \ell, j \neq k} \frac{\lambda_j}{\lambda_k} v_j$, d.h. $v_k \in \langle (v_j)_{1 \leq j \leq \ell, j \neq k} \rangle$. □

l.u. (Forts.)

- Es gilt: $(v_1, \ldots, v_\ell)$ l.a. gdw: Es gibt λ_i, λ'_i s.d. $\sum_{1 \leq j \leq \ell} \lambda_j v_j = \sum_{1 \leq j \leq \ell} \lambda'_j v_j$ mit $\lambda_j \neq \lambda'_j$ für (mindestens) ein j.

  Beweis: Wenn es solche λ, λ' gibt, dann $0 = \sum_{1 \leq j \leq \ell} (\lambda_j - \lambda'_j) v_j$, d.h. $(v_1, \ldots, v_\ell)$ l.a.
  Wenn umgekehrt $(v_1, \ldots, v_\ell)$ l.a., dann läßt sich 0 sowohl trivial als auch nicht-trivial darstellen. □

- $(v_1, \ldots, v_\ell)$ ist l.u., gdw jedes $w \in \langle v_1, \ldots, v_\ell \rangle$ eine eindeutige Darstellung $w = \sum_{1 \leq j \leq \ell} \lambda_j v_j$ hat.

  Beweis: Vorige Punkt: $A \leftrightarrow B$, dieser Punkt: $\neg A \leftrightarrow \neg B$. □

l.u. (Forts.)

- Es gilt: Wenn und A l.u. ist, dann ist $A \cup \{v\}$ l.u. gdw $v \notin \langle A \rangle$.
  Beweis: Wisen schon: $(v_1, \ldots, v_\ell)$ l.a. gdw: $v_k \in \langle (v_j)_{1 \leq j \leq \ell, j \neq k} \rangle$ für ein $q \leq k \leq \ell$.
  Wenn $v \in \langle A \rangle$, dann ist daher $A \cup \{v\}$ l.a.
  Sei nun $A \cup \{v\}$ l.a., d.h. sei $0 = \mu v + \sum \lambda_1 w_i$ eine nicht-triviale Darstellung mit $w_i \in A$. Dann muss $\mu \neq 0$ gelten: Ansonsten wäre $0 = \sum \lambda_1 w_i$ nicht-tivial, ein Widerspruch zu A l.u. Also ist $v = \sum -\frac{\lambda_i}{\mu} w_i$, d.h. $v \in \langle A \rangle$. □
- Es gilt: Wenn W Teilraum von V ist, und $A \subseteq W$, dann ist A l.u. in W genau dann wenn A l.u. in V ist.
  Beweis: In V und W gibt es genau dieselben Linearkombinationen von Elementen von A. □

Basen als l.u. Erzeugende

Wir haben definiert: $E = (e_1, \ldots, e_n)$ ist eine Basis von V, wenn jedes $v \in V$ eine eindeutige Darstellung $v = \sum_{i=1}^{n} \lambda_i e_i$ hat.

Theorem

$E = (e_1, \ldots, e_n)$ ist eine Basis gdw.: linear E ist l.u. und $\langle E \rangle = V$.

Beweis: Sei E Basis. Jedes $v \in V$ ist Linearkombination, daher ist $V = \langle E \rangle$. Linearkombinationen sind eindeutig, daher ist E l.u.

Wenn nun $\langle E \rangle = V$, dann lässt sich jedes $v \in V$ als Linearkombination darstellen, und weil E l.u. ist, ist diese Darstellung eindeutig, d.h. E ist Basis. □

Konstruktion von Basen

Wir können nun Basen folgendermaßen konstruieren:

- Wähle $e_1 \neq 0$ beliebig. Setzte $E_1 := (e_1)$ (eine l.u. Folge der Länge 1).
- Wenn wir bereits ein l.u. E_ℓ konstruiert haben:
  Wenn $\langle E_\ell \rangle = V$, dann sind wir fertig: E_ℓ ist Basis, V ist ℓ-dim.
  Ansonsten wähle $e_{\ell+1} \notin \langle E_\ell \rangle$ und setze $E_{\ell+1} := (e_1, \ldots, e_{\ell+1})$.
  Weil E_ℓ l.u. und $e_{\ell+1} \notin \langle E_\ell \rangle$ ist $E_{\ell+1}$ l.u.
- Wenn wir nach endlich vielen Schritten fertig werden, haben wir eine (endliche) Basis gefunden. Ansonsten heißt V unendlich-dimensional (wir beschäftigen uns in dieser VO nicht mit solchen Räumen).

(Bemerkung: Auch unendlich-dimensionale Vektorräume haben eine Basis, die obige Konstruktion wird dafür einfach "uber unendlich hinaus" weitergeführt. Dazu braucht man das sogenannte Auswahlaxiom.)

Basen in einem Erzeugendensystem

Sei $A \subseteq V$ endlich.

Theorem

Wenn $\langle A \rangle = V$ ist, dann gibt es Teilfolge von A die Basis ist.

Beweis:

- Sei $A = (a_1, \ldots, a_k)$. Wir können annehmen dass alle $a_i \neq 0$ (Nullvektoren können wir weglassen).
- Setze $E_1 := \langle a_1 \rangle$.
- Angenommen wir haben bereits E_ℓ l.u. für ein $\ell < k$, mit $\langle E_\ell \rangle = \langle (a_1, \ldots, a_\ell) \rangle$.
  Wenn $a_{\ell+1} \in \langle E_\ell \rangle$, setzte $E_{\ell+1} = E_\ell$.
  Ansonsten sei $E_{\ell+1}$ E_ℓ erweitert um $a_{\ell+1}$. Dann ist $E_{\ell+1}$ wieder l.u.
- Dann ist E_k Basis aus A-Elementen. □

Austauschlemma

Lemma

Sei $A = (v_1, \ldots, v_n)$ und $w = \sum_{i=1}^{n} \lambda_i v_i$ mit $\lambda_k \neq 0$.
Setze $A' := (v_1, \ldots, v_{k-1}, w, v_{k+1}, \ldots, v_n)$.
Dann ist $\langle A \rangle = \langle A' \rangle$, und A ist l.u. gdw A' l.u.

Beweis Teil 1 von 3: $\langle A \rangle = \langle A' \rangle$:

- $v_k = \frac{1}{\lambda_k} w - \sum_{1 \leq i \leq n, i \neq k} \frac{\lambda_i}{\lambda_k} v_i$
- Daher ist $v_k \in \langle A' \rangle$, damit ist $A \subseteq \langle A' \rangle$ und damit $\langle A \rangle \subseteq \langle A' \rangle$.
- Umgekehrt ist $w \in \langle A \rangle$, daher $A' \subseteq \langle A \rangle$, und damit $\langle A' \rangle \subset \langle A \rangle$.
- Insgesamt ist also $\langle A \rangle = \langle A' \rangle$.

Austauschlemma (Forts.)

Lemma

Sei $A = (v_1, \ldots, v_n)$ und $w = \sum_{i=1}^{n} \lambda_i v_i$ mit $\lambda_k \neq 0$.
Sei $A' = (v_1, \ldots, v_{k-1}, w, v_{k+1}, \ldots, v_n)$.
Dann ist $\langle A \rangle = \langle A' \rangle$, und A ist l.u. gdw A' l.u.

Beweis Teil 2 von 3: A' l.u. impliziert A l.u., oder äquivalent:
A l.a. impliziert A' l.a.

- $v_k = \frac{1}{\lambda_k} w - \sum_{1 \leq i \leq n, i \neq k} \frac{\lambda_i}{\lambda_k} v_i$
- Sei A l.a., mit nicht-trivialer Darstellung $0 = \sum_{i=1}^{n} \mu_i v_i$.
- Fall 1: $\mu_k = 0$. Dann ist $0 = \sum_{1 \leq i \leq n, i \neq k} \mu_i v_i$ eine nicht-triviale Darstellung von Elementen von A', d.h. A' ist l.a.
- Fall 2: $\mu_k \neq 0$. Dann bekommen wir die nicht-triviale Darstellung $0 = \sum_{1 \leq i \leq n, i \neq k} \mu_i v_i + \mu_k \frac{1}{\lambda_k} w - \mu_k \sum_{1 \leq i \leq n, i \neq k} \frac{\lambda_i}{\lambda_k} v_i$:
  nicht-trivial, weil der Koeffizient von w gleich $\frac{\mu_k}{\lambda_k} \neq 0$ ist.
  Daher ist A' l.a.

Austauschlemma (Forts.)

Lemma

Sei $A = (v_1, \ldots, v_n)$ und $w = \sum_{i=1}^{n} \lambda_i v_i$ mit $\lambda_k \neq 0$.
Sei $A' = (v_1, \ldots, v_{k-1}, w, v_{k+1}, \ldots, v_n)$.
Dann ist $\langle A \rangle = \langle A' \rangle$, und A ist l.u. gdw A' l.u.

Beweis Teil 3 von 3: A l.u. impliziert A' l.u., oder äquivalent:
A' l.a. impliziert A l.a.

- $v_k = \frac{1}{\lambda_k} w - \sum_{1 \leq i \leq n, i \neq k} \frac{\lambda_i}{\lambda_k} v_i$
- Sei A' l.a., mit nicht-trivialer Darstellung $0 = \mu_k w + \sum_{1 \leq i \leq n, i \neq k} \mu_i v_i$.
- Fall 1: $\mu_k = 0$. Das gibt die nicht-triviale Darstellung $0 = \sum_{1 \leq i \leq n, i \neq k} \mu_i v_i$ in A, d.h. A ist l.a..
- Fall 2: $\mu_k \neq 0$. Dann ist $0 = \mu_k \sum_{i=1}^{n} \lambda_i v_i + \sum_{1 \leq i \leq n, i \neq k} \mu_i v_i$ nicht-trivial, weil der Koeffizient für v_k gleich $\mu_k \lambda_k \neq 0$ ist. □

Corollary

Wenn A, A' wie im Lemma, dann ist A Basis gdw A' Basis.

Erweiterung zu einer Basis

Theorem

Sei $E = (e_1, \ldots, e_n)$ Basis von V, und $A \subseteq V$ endlich.
Wenn A l.u. ist, dann gibt es eine Erweiterung von A durch Elemente aus E die eine Basis der Länge n ist.
Insbesondere ist $|A| \leq n$, und wenn $|A| = n$ dann ist A eine Basis.

In anderen Worten: Wenn $A = (a_1, \ldots, a_k)$ l.u. ist, dann gibt es eine Basis $(a_1, \ldots, a_k, b_{k+1}, \ldots, b_n)$ so dass jedes b_ℓ gleich einem e_i ist.

Beweis:

- Wir ersetzen der Reihe nach Elemente von E durch Elemente aus A:
- a_1 hat eine E-Darstellung $a_1 = \sum \lambda_i e_i$, mit zumindest einem $\lambda_\ell \neq 0$ (weil ja $a_1 \neq 0$).
- Wir können a_1 gegen dieses e_ℓ austauschen, und erhalten eine neue Basis E_1 (weiterhin mit n Elementen).
- Durch Umordnung können wir schreiben: $E_1 = (e_1, b_2, \ldots, b_n)$, mit b_i aus E.

Beweis (Forts.)

- Angenommen wir haben, für ein $j < k$, bereits eine Basis $E_j = (a_1, \ldots, a_j, b_{j+1}, \ldots, b_n)$ mit b_i aus E.

  Weil $\langle E_j \rangle = V$, ist $a_{j+1} = \sum_{i=1}^{k} \lambda_i a_i + \sum_{\ell=j+1}^{n} \mu_\ell b_\ell$.

  Es können nicht alle $\mu_\ell = 0$ sein: Ansonsten wäre $0 = \sum_{i=1}^{j} \lambda_i a_i - a_{j+1}$ eine nichttriviale Darstellung in A, aber A ist l.u.

  Sei also $\mu_\ell \neq 0$. Dann können wir a_{j+1} gegen b_ℓ austauschen, und erhalten (nach Umordnung) die Basis $E_{j+1} = (a_1, \ldots, a_{j+1}, c_{j+2}, \ldots, c_n)$, mit c_i aus E.
- Beachte: Daraus folgt insbesondere $n > j$.
- Nach k Schritten erhalten wir die gesuchte Basis $E_k = (a_1, \ldots, a_k, b_{k+1}, \ldots, b_n)$. □

Insbesondere ist $n \geq k$, und wenn $n = k$ dann war A selbst schon Basis.

Dimension

Theorem

Sei $E = (e_1, \ldots, e_n)$ Basis von V.
Wenn E' eine Basis ist, dann ist E' ebenfalls n-dimensional.
Die Dimension eines (endl.) Vektorraums ist eindeutig.
Wir schreiben $\dim(V) = n$.

Beweis: E' ist l.u., wir können es also zu mit Elementen aus E zu einer Basis der Länge $n = |E|$ erweitern. Damit ist $|E'| \leq |E|$. Wenn man die Rollen von E und E' umdreht, bekommt man $|E| \leq |E'|$, und insgesamt daher $|E| = |E'|$. □

Mehr Basen

Theorem

Sei V n-dimensional, und $A \subseteq V$ endlich.

- *Wenn $\langle A \rangle = V$, gilt: $|A| \geq n$, und $|A| = n$ gdw A Basis.*
- *Wenn A l.u. ist, dann gilt: $|A| \leq n$, und $|A| = n$ gdw A Basis.*
- *Wenn $|A| = n$, gilt also: $\langle A \rangle = V$ gdw. A l.u. gdw. A Basis.*

Beweis:

- Wir wissen: Wenn $\langle A \rangle = V$, dann gibt es Teilfolge von A die Basis ist, und daher Größe n hat. Daher ist $|A| \geq n$.
  Wenn $|A| = n$ dann ist die Teilfolge ganz A und daher ist A eine Basis. (Und wenn A Basis dann ist $|A| = n$ wegen der eindeutigen Dimension.)
- Wir wissen: Wenn A l.u., dann lässt sich A zu Basis der Größe n erweitern. Daher gilt: $|A| \leq n$, und wenn $|A| = n$ dann ist A bereits Basis.

□

(Bemerkung: dieses Theorem gilt nicht im unendlichdimensionalen.)

Basen und linere Abbildungen

Theorem

Sei $E = (e_1, \ldots, e_n)$ *Basis von* V, $f : V \to W$ *linear, und* $A := (f(e_1), \ldots, f(e_n))$.

- $\langle A \rangle = W$ *gdw* f *surjektiv.*
- A *ist l.u. gdw* f *injektiv.*
- *Wenn* $\dim(W) = n$, *dann:* $\langle A \rangle = W \leftrightarrow A$ *l.u.* $\leftrightarrow A$ *Basis*
  Bzw: f *injektiv* $\leftrightarrow$ f *surjektiv* $\leftrightarrow$ f *invertierbar*

Beweis 1/3: $\langle A \rangle = W$ gdw f surjektiv.
Die folgenden Aussagen sind äquivalent:

- f ist surjektiv.
- Jedes $w \in W$ hat Form $w = f(v) = f(\sum \lambda_i e_i)$ für ein $v \in V$ bzw für bestimmte $\lambda_i \in K$. (Nach Def. von surjektiv, und weil $V = \langle E \rangle$).)
- Jedes w hat Form $\sum \lambda_i f(e_i)$ für bestimmte λ_i.
  (Weil f linear ist.)
- $\langle A \rangle = W$.

Basen und linere Abbildungen

Theorem

Sei $E = (e_1, \ldots, e_n)$ Basis von V, $f : V \to W$ linear, und $A := (f(e_1), \ldots, f(e_n))$.

- $\langle A \rangle = W$ *gdw* f *surjektiv.*
- A *ist l.u. gdw* f *injektiv.*
- *Wenn* $\dim(W) = n$, *dann:* $\langle A \rangle = W \leftrightarrow A$ *l.u.* $\leftrightarrow A$ *Basis*
  Bzw: f *injektiv* $\leftrightarrow$ f *surjektiv* $\leftrightarrow$ f *invertierbar*

Beweis 2/3: A ist l.u. gdw f injektiv.
Die folgenden Aussagen sind äquivalent:

- f ist injektiv.
- $f(v) = 0 \to v = 0$. (Weil $f(a) = f(b)$ gdw $f(a - b) = 0$.)
- $f(\sum \lambda_i e_i) = 0 \to (\forall i)\lambda_i = 0$. (Weil E l.u. ist.)
- $\sum \lambda_i f(e_i) = 0 \to (\forall i)\lambda_i = 0$. (Weil f linear ist.)
- A l.u.

Basen und linere Abbildungen

Theorem

Sei $E = (e_1, \ldots, e_n)$ *Basis von* V, $f : V \to W$ *linear, und* $A := (f(e_1), \ldots, f(e_n))$.

- $\langle A \rangle = W$ *gdw* f *surjektiv.*
- A *ist l.u. gdw* f *injektiv.*
- *Wenn* $\dim(W) = n$, *dann:* $\langle A \rangle = W \leftrightarrow$ *A l.u.* $\leftrightarrow$ *A Basis*
  Bzw: f *injektiv* $\leftrightarrow$ f *surjektiv* $\leftrightarrow$ f *invertierbar*

Beweis 3/3: Sei $\dim(W) = n$.
$|A| = n$, daher ist "A l.u." äquivalent zu "$\langle A \rangle = W$" und auch zu 'A ist Basis".
Und wir wissen bereits: Wenn $(f(e_1), \ldots, f(e_n))$ Basis ist, können wir die lineare Abbildung $d : W \to V$ definieren durch $f(e_i) \mapsto e_i$, und es gilt: $f \circ g = g \circ f = \mathrm{Id}$. □

Bemerkung: Für unendlich-dimensionale Räume gilt der Satz nicht.

Ein unendlichdimensinaler Vektorraum

- Zur Erinnerung: $K[X]$ ist der Rind der (formalen) Polynome mit Koeffizienten aus K.
- Betrachten wir der Einfachheit halber $V := \mathbb{R}[X]$. (In dem Fall entsprechen formale Polynome eindeutig den Polynomfunktionen.)
- V ist ein Vektorraum (über $\mathbb{R}$). Dabei setzen wir natürlich $\lambda \cdot (a_n X^n + a_{n-1} X^{n-1} + \cdots + a_0) := \lambda a_n X^n + \lambda a_{n-1} X^{n-1} + \cdots + \lambda a_0$.
- V ist unendlich-dimensional. Eine Basis ist z.B. $E = (X^0, X^1, X^2, X^3, \dots)$:
  Jedes $v \in V$ lässt sich auf eindeutige Weise als (endlich) Summe $\sum_{i=1}^{N} \lambda_i b_i$ schreiben, mit $N \in \mathbb{N}$ und jedes $b_i \in E$.
- Auch im unendlich-dimensionalen gilte: E Basis gdw $\langle E \rangle = V$ und E l.u.; jede l.u. Menge läßt sich zu Basis erweitern; jede Erzeugendenmenge enthält Basis.
- ABER: f definiert durch $X^n \mapsto X^{n+1}$ ist injektiv aber nicht surjektiv, g mit $X^0 \mapsto 0$, $X^{n+1} \mapsto X^n$ surjektiv aber nicht injektiv.

Basen und Unterräume

Theorem

Sei $\dim(V) = n$ *und* W *ein Teilraum von* V*.*
Dann ist $\dim(W) \leq \dim(V)$*, und jede Basis von* W *läßt sich zu einer Basis von* V *fortsetzen.*

Beweis:
Jede Basis in W ist auch l.u. in V, und läßt sich daher zu einer Basis von V fortsetzen. □

Komplemente und Projektionen

Komplemente

Seien W_1, W_2 Unterräume von V.

Definition

W_1 und W_2 sind Komplementär-Räume (kurz: Komplemente), wenn $W_1 \cap W_2 = \{\emptyset\}$ und $\langle W_1 \cup W_2 \rangle = V$.

Die folgenden sind äquivalent:

1. W_1 und W_2 sind komplementär.
2. Für jede Basis $E = (e_1, \ldots, e_n)$ von W_1 und $B = (b_1, \ldots, b_\ell)$ von W_2 gilt: $C = (e_1, \ldots, e_n, b_1, \ldots, b_\ell)$ ist Basis von V.
3. Für jedes $v \in V$ gibt es eindeutige $w_1 \in W_1$ und $w_2 \in W_2$ mit $v = w_1 + w_2$.

Beweis $1 \to 2$: Sei $v_1 = \sum_{i=1}^{m} \lambda_i e_i \in W_1$ und $v_2 = \sum_{j=1}^{\ell} \mu_j b_j \in W_2$. Wenn $v_1 + v_2 = 0$, dann $v_1 = -v_2$, und daher $v_1 \in W_1 \cap W_2$, daher $v_1 = 0$; weil E Basis ist sind alle $\lambda_i = 0$. Genauso folgt: Alle $\mu_j = 0$. Daher ist C l.u. Und $\langle C \rangle = \langle W_1 \cup W_2 \rangle = V$. Also ist C Basis.

Komplemente 2

Bew. (Fort.) Die folgenden sind äquivalent:

1. W_1 und W_2 sind komplementär.
2. Für jede Basis $E = (e_1, \ldots, e_n)$ von W_1 und $B = (b_1, \ldots, b_\ell)$ von W_2 gilt: $C = (e_1, \ldots, e_n, b_1, \ldots, b_\ell)$ ist Basis von V.
3. Für jedes $v \in V$ gibt es eindeutige $w_1 \in W_1$ und $w_2 \in W_2$ mit $v = w_1 + w_2$.

- $2 \to 3$: Fixiere Basen E und B. Dann ist C Basis, und jedes v hat Eindeutige Darstellung $v = \sum_{i=1}^{m} \lambda_i e_i + \sum_{j=1}^{\ell} \mu_j b_j$. Angenommen $v = w_1 + w_2$ mit $w_1 = \sum_{i=1}^{m} \lambda'_i e_i$ und $w_2 = \sum_{j=1}^{\ell} \mu'_j b_j$. Dann ist $0 = v - v = \sum_{i=1}^{m} (\lambda_i - \lambda'_i) e_i + \sum_{j=1}^{\ell} (\mu_j - \mu'_j) b_j$; also sind alle $\lambda_i = \lambda'_i$ und $\mu_j = \mu'_j$, und w_1 und w_2 sind eindeutig.
- $3 \to 1$: Weil jedes v Darstellung $v = w_1 + w_2$ hat, ist $\langle W_1 \cup W_2 \rangle = V$. Angenommen $v \in W_1 \cap W_2$ mit $v \neq 0$. Dann wären $w_1 = v, w_2 = 0$ und $w_1 = 0, w_2 = v$ zwei verschiedene Darstellungen. □

Beispiele

Beispiele:
Sei (e_1, e_2) die Standardbasis von $\mathbb{R}^2$.

- $W_1 := \langle e_1 \rangle$ und $W_2 := \langle e_1 \rangle$ sind komplementär.
- Sei $W_2' := \langle b \rangle$ für $b = e_1 + e_2$.
  Dann sein W_1 und W_2' ebenfalls komplementär.

Es gilt:

- Zu jedem Unterraum W_1 von V gibt es einen (nicht eindeutigen) Komplementärraum W_2.
  (Beweis: Eine Basis von W_1 kann zu einer von V erweitert werden.)
- Wenn W_1 und W_2 komplementär sind, dann ist $\dim(W_1) + \dim(W_2) = \dim(V)$.

Projektionen

Definition

$f \in L(V)$ (d.h.: $f : V \to V$ linear) heißt Projektion, wenn es komplementäre W_1, W_2 gibt mit $f \restriction W_1 = \mathrm{Id}$ und $f \restriction W_2 = 0$.

Wir sagen auch "Projektion auf W_1". (Aber f hängt auch von W_2 ab!)

Beispiele: Sei $E = (e_1, e_2)$ die Standardbasis von $\mathbb{R}^2$.

- Für $W_1 := \langle e_1 \rangle$ und $W_2 := \langle e_1 \rangle$ ist die Projektion definiert durch $e_1 \mapsto e_1$ und $e_2 \mapsto 0$. Die E-Darstellung ist $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$.
- Sei $W_2' := \langle b \rangle$ für $b = e_1 + e_2$. Für W_1 und W_2 ist die Projektion f definiert durch $f(e_1) = e_1$ und $f(e_1 + e_2) = 0$. Daher ist $f(e_2) = f(e_1 + e_2 - e_1) = 0 - e_1$, und die E-Darstellung ist $\begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix}$.

Projektionen 2

In geeigneter Basis haben Projektion eine besonders einfache Darstellung:

Definition (nicht Standard)

Die kanonische $n \times m$-Matrix von Rang ℓ ist die $m \times n$-Matrix $A = (a_{i,j})$ definiert durch $a_{i,j} = \begin{cases} 1 & \text{wenn } i = j \text{ und } 1 \leq i \leq \ell \\ 0 & \text{sonst.} \end{cases}$

Wenn $(e_1, \ldots e_m)$ Basis von W_1, und $(e_{m+1}, \ldots, e_n)$ von W_2 ist, dann ist $E = (e_1, \ldots, e_n)$ eine V-Basis, und die E-Darstellung der Projektion ist die kanonische $n \times n$-Matrix von Rang m:

$$\begin{pmatrix} 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix}$$

Ähnlichkeit

Definition

Zwei $n \times n$-Matrizen A und B heißen ähnlich, wenn A die E-Darstellung und B die E'-Darstellung derselben linearen Funktion $f : V \to V$ ist (für irgendeinen n-dimensionalen Vektorraum mit Basen E, E').

Mit anderen Worte, wenn es ein $S \in \mathrm{GL}(n)$ gibt mit $B = S^{-1}AS$.

Wir haben gesehen: f ist Projektion, genau dann wenn f ähnlich ist einer kanonischen Matrix.

Projektionen 3

Es gilt:

1. Zu jedem komplementären W_1, W_2 gibt es genau eine Projektion.
2. Wenn $f \in L(V)$, dann ist f Projektion gdw $f^2 = f$. ($f^2 := f \circ f$)

Beweis:

- Wähle Basen E von W_1 und B von W_2. Dann ist $E \cup B$ Basis von V, und das lineare f das (eindeutig) definiert ist durch $f(e) = e$ für $e \in E$ und $f(b) = 0$ für $b \in B$ ist die entsprechende Projektion.
- Sei f Projektion bzgl W_1, W_2, und $v \in V$. Dann ist $v = w_1 + w_2$, und $f(w_1 + w_2) = w_1$. Daher $f(f(w_1 + w_2)) = f(w_1) = f(w_1 + w_2)$.
- Sei $f^2 = f$. Sei $W_2 := \{v \in V : f(v) = 0\}$ und $W_1 := f''V := \{f(v) : v \in V\}$.
  W_1 und W_2 sind komplementär: $f(v - f(v)) = f(v) - f(f(v)) = 0$, daher lässt sich jedes v schreiben als $w_1 + w_2$ mit $w_1 := f(v) \in W_1$ und $w_2 := v - f(v) \in W_2$. Und wenn $v \in W_1 \cap W_2$, dann ist $v = f(v')$ (weil $v \in W_1$) und $f(v) = 0$ (weil $v \in W_2$), und daher $0 = f(v) = f(f(v')) = f(v') = v$.

Beispiele

- $A = \begin{pmatrix} 4 & 3 \\ -3 & -3 \end{pmatrix}$ ist keine Projektion, weil $A^2 \neq A$.
- $B = \begin{pmatrix} -4 & -10 \\ 2 & 5 \end{pmatrix}$ ist eine Projektion.

  Eine andere Frage ist: Was ist das dazugehörige W_1 und W_2? Werden wir später (und allgemeiner) behandeln.

Kern und Rang

Kern

- Sei $f : V \to W$ linear, $\dim(V) = n$, $\ker(f) := \{v \in V : f(v) = 0\}$.
- $\ker(f)$ ist Unterraum von V (weil f linear ist).
- $m := \dim(\ker(f))$ wird manchmal auch "Defekt von f" genannt.
- Sei V' ein Komplementärraum von $\ker(f)$. Es gilt $\dim(V') = n - m$.
- Das Bild $f''V := \{f(v) : v \in V\} =: \mathrm{img}(f)$ ist Unterraum von W.
- $f \restriction V'$ ist eine injektive lineare Abbildung $V' \to W$, mit demselben Bild wie f (d.h. $f''V' = \mathrm{img}(f)$).
  Beweis: Jedes $v \in V$ hat eindeutige Darstellung $v = w_1 + w_2$ mit $w_1 \in \ker(f)$ und $w_2 \in V'$. Dann ist $f(v) = f(w_2) \in f''V'$.
  Und wenn $a \neq b$ in V', dann ist $a - b \neq 0$ in V', daher ist $a - b \notin \ker(f)$, daher ist $f(a - b) \neq 0$, daher ist $f(a) \neq f(b)$. □
- $\dim(V') = n - m$, und $f : V' \to W$ injektiv, also ist $\dim(\mathrm{img}(f)) = \dim(f''V') = n - w$.
  Beweis: Wähle Basis on V', injektives Bild ist l.u. □
- Es gilt: $\dim(\ker(f)) + \dim(\mathrm{img}(f)) = n$.
- $\dim(\mathrm{img}(f))$ wird auch Rang von f oder $\mathrm{rk}(f)$ genannt.

Zusammenfassung

Definition

Sei $f \in L(V, W)$.

- $\ker(f) := \{v \in V : f(v) = 0\}$, ein Unterraum von V.
- $\text{img}(f) := \{f(v) : v \in V\}$, ein Unterraum von W.
- $\text{rk}(f) := \dim(\text{img}(f))$

Dann gilt:

Theorem

$\text{rk}(f) + \dim(\ker(f)) = \dim(V)$

Isomorphismen

- Ein Isomorphismus ist ein invertierbarer Homomorphismus.
- Ein Vektorraum-Isomorphismus $f : V \to W$ ist eine invertierbare (d.h. bijektive) lineare Abbildung.
- Wenn V und W isomorph sind, dann ist $\dim(V) = \dim(W)$.
  Beweis: $\dim(W) = \dim(\text{img}(f)) \leq \dim(V)$; und dasselbe für die Inverse von f zeigt $\dim(V) \leq \dim(W)$. □
- Wenn $\dim(V) = n$, dann ist V isomorph zu K^n.
  Beweis: Fixiere Basis E von V. Sei $\phi : V \to K^n$ die Abbildung die $v \in V$ auf die E-Darstellung von v (ein Zeilenvektor in K^n) abbildet. Dann ist ϕ ein Isomorphismus. □
- Insgesamt bekommen wir also:

Theorem

Zwei K-Vektorräume V und W sind isomorph gdw. $\dim(V) = \dim(W)$.

(Bem.: Gilt auch für unendlichdimensionale Vektor-Räume)

Rang von Matrizen

Zur Wiederholung:
Gegeben Basen von V und W, $dim(V) = n$, $\dim(W) = m$, entsprechen die linearen Funktionen $L(V, W)$ genau den $m \times n$-Matrizen $L(m, n)$, die wir als Funktionen von $K^n \to K^n$ auffassen können.
Wir bekommen dann:

- Die Spalten von A spannen $\text{img}(A)$ (einen Teilraum von K^m) auf. Die Dimension von $\text{img}(A)$ ist die maximale Größe einer l.u. Folge von Spalten. (Auch Spaltenrang oder nur $\text{rk}(A)$ genannt.)
- Es gilt (ohne Beweis): $\text{rk}(A)$ ist auch die Dimension des Teilraums von K^n, der durch die Zeilen von A aufgespannt wird. (Zeilenrang=Spaltenrang.)
- Wir wissen bereits: $\dim(\ker(A)) + \text{rk}(A) = n$.

Sei $f : V \to W$ linear, $\dim(V) = n$, $\dim(W) = m$.

- Sei W ein Komplement von $\ker(f)$.
- Sei $(e_1, \ldots, e_\ell)$ Basis von W und $(e_{\ell+1}, \ldots, e_n)$ von $\ker(f)$, d.h. $E = (e_1, \ldots, e_n)$ ist Basis von V.
- Dann ist $(f(e_1), \ldots, f(e_\ell))$ l.u. und erzeugt $\mathrm{img}(f)$.
- Erweitere $(f(e_1), \ldots, f(e_\ell))$ zu einer Basis B von W.
- Die (E, B)-Darstellung A von f ist also die kanonische $n \times m$ Matrix von Rang ℓ: $a_{i,i} = 1$ wenn $1 \leq i \leq \ell$, und $a_{i,j} = 0$ sonst.

Äquivalenz von Matrizen

Definition

Zwei $m \times n$-Matrizen A und B heißen äquivalent, wenn sie Darstellungen derselben Funktion $f : V \to W$ (für verschiedene Basen) sind.

In anderen Worten: A und B sind äquivalent, wenn es $S \in \mathrm{GL}(m)$ und $T \in GL(n)$ gibt mit $B = S^{-1}AT$.

- Zwei äquivalente Matrizen haben denselben Rang.
- Eine $m \times n$ Matrix A mit Rang ℓ ist äquivalent zu der kanonischen $m \times n$-Matrix von Rang ℓ.
- Daher gilt: Zwei $m \times n$ Matrizen A und B sind äquivalent gdw $\mathrm{rk}(A) = \mathrm{rk}(B)$.

Äquivalent vs ähnlich

- Erinnerung: A und B sind ähnlich, wenn sie dieselbe Funktion beschreiben, wobei man für A eine Basis wählt und für B eine andere.
  D.h. $B = T^{-1}AT$ mit $T \in \mathrm{GL}(n)$.
- A und B sind äquivalent, wenn sie dieselbe Funktion beschreiben, wobei man sowohl für A als auch für B jeweils zwei Basen wählt (für den Definitions- und den Zielraum).
  D.h.: $B = S^{-1}AT$ mit $S, T \in \mathrm{GL}(n)$.
- Für $n \times n$ Matrizen gilt impliziert ähnlich daher äquivalent.
- Die Umkehrung gilt nicht: Wir wissen: f hat E-Darstellung als $n \times n$ kanonische Matrix gdw f Projektion ist.
  D.h. nur Projektionen sind ähnlich zu einer kanonischen Matrix;
  aber jede Matrix ist äquivalent zu kanonischen Matrix.
- Zwei ähnliche Matrizen haben dieselben Eigenwerte (EW),
  zwei äquivalente aber i.A. nicht. (Bsp: Identität I und Drehung sind äquivalent.)
  Erinnerung: λ ist EW, wenn es ein $v \neq 0$ gibt mit $f(v) = \lambda v$. Das ist von der E-Darstellung von f unabhängig.

Normalformen (nicht Prüfungsstoff)

- Wichtige Frage in der Mathematik: Gegeben eine Klasse von Objekten (Gruppen, lineare Abbildungen, ...) und einen Begriff von Äquivalenz, finde eine "Normalform" oder eine "kanonische Darstellung" für jeden Repräsentanten.
- Für $m \times n$ Matrizen und "äquivalent" ist das einfach/langweilig und wenig nützlich: Kanonische Matrix von Rang ℓ.
  (Eine Andere Frage ist: Gegeben A, wie finde ich Basen von V und W in der A diese Darstellung hat?)
- Für $n \times n$ Matrizen und "ähnlich" interessanter und nützlicher.
  Es gibt verschiedene Normalformen (zB Frobenius-Normalform)
- Für bestimmte Matrizen mit bestimmten Eigenschaften gibt es oft besonders schöne Normalformen.
  Bsp für reelle Matrizen: Jordanschen Normalform für triagonalisierbare Matrizen.
  Bsp für komplexe Matrizen: ("Spektralsatz") Normale Matrizen sind diagonalisierbar.

Elementare Umformungen, Dreiecksform

Reguläre Matrizen

Zusammenfassung: Die folgenden Aussagen sind äquivalent für $n \times n$-Matrizen A:

- Spalten von A sind Basis von K^n.
- Spalten von A spannen K^n auf.
- Spalten von A sind l.u.
- A ist injektiv.
- A ist surjektiv.
- Zeilen von A sind Basis von K^n.
- Zeilen von A spannen K^n auf.
- Zeilen von A sind l.u.
- A ist bijektiv.
- A ist invertierbar.

(Invertierbar heißt: $\exists B \in L(n)\, BA = AB = I$.) So ein A heißt "regulär", nicht-invertierbare Matrizen heissen "singulär".

Elementare Spaltenumformungen: Die Matrizen

Sei A eine $m \times n$-Matrix.
Folgende Modifikationen von A erhalten $\mathrm{img}(A)$ (und daher auch $\mathrm{rk}(A) = \dim(\mathrm{img}(A))$):

- Vertausche zwei Spalten.
- Multipliziere ein Spalte mit einem $\lambda \neq 0$.
- Addiere das Vielfache einer Spalte zu einer anderen Spalte.

All diese Umformungen transformieren A auf AU mit $U \in \mathrm{GL}(n)$.

Elementare Zeilenumformungen

Konkreter: Wir bekommen Umformungen $A \mapsto AU$ für die folgenden U: (die Abweichungen von I sind rot markiert):

Vertausche Spalte i und j

$a_{i,i} = a_{j,j} = 0, a_{i,j} = a_{j,i} = 1$

$$E^{\mathrm{sw}}_{i,j} := \begin{pmatrix} 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & & \vdots & & \vdots & & \vdots \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 & \cdots & 0 \\ \vdots & & \vdots & \ddots & \vdots & & \vdots & & \vdots \\ 0 & \cdots & 0 & \cdots & 1 & \cdots & 0 & \cdots & 0 \\ \vdots & & \vdots & & \vdots & \ddots & \vdots & & \vdots \\ 0 & \cdots & 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & & \vdots & & \vdots & & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{pmatrix}$$

$$(E^{\mathrm{sw}}_{i,j})^{-1} = E^{\mathrm{sw}}_{i,j}$$

$$\det(E^{\mathrm{sw}}_{i,j})^{-1} = -1$$

Spalte i mal $\lambda \neq 0$

$a_{i,i} = \lambda$

$$E^{\mathrm{m}}_{i,\lambda} := \begin{pmatrix} 1 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda & \cdots & 0 \\ \vdots & \vdots & & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 1 \end{pmatrix}$$

$$(E^{\mathrm{m}}_{i,\lambda})^{-1} = E^{\mathrm{m}}_{i,1/\lambda}$$

$$\det(E^{\mathrm{m}}_{i,\lambda})^{-1} = \lambda$$

Addiere λ mal Spalte i zu Spalte j

$a_{i,j} = \lambda$

$$E^{\mathrm{a}}_{i,j,\lambda} := \begin{pmatrix} 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & & \vdots & & \vdots \\ 0 & \cdots & 1 & \cdots & \lambda & \cdots & 0 \\ \vdots & & \vdots & \ddots & \vdots & & \vdots \\ 0 & \cdots & 0 & \cdots & 1 & \cdots & 0 \\ \vdots & & \vdots & & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{pmatrix}$$

$$(E^{\mathrm{a}}_{i,j,\lambda})^{-1} = E^{\mathrm{a}}_{i,j,-\lambda}$$

$$\det(E^{\mathrm{a}}_{i,j,\lambda}) = 1$$

Elementare Zeilenumformungen

Analog verändern die entsprechenden Zeilen-Operationen nicht $\ker(A)$ und damit ebenfalls nicht $\mathrm{rk}(A) = n - \dim(\ker(A))$:

- Vertausche zwei Zeilen.
- Multipliziere ein Zeile mit einem $\lambda \neq 0$.
- Addiere das Vielfache einer Zeile zu einer anderen Zeile.

All diese Umformungen transformieren A auf UA mit $U \in \mathrm{GL}(m)$, konkreter die folgenden U (die Abweichungen sind rot markiert):

Vertausche Zeile i und j	Zeile i mal $\lambda \neq 0$	Addiere λ mal Zeile i zu Reihe j
$a_{i,i} = a_{j,j} = 0, a_{i,j} = a_{j,i} = 1$	$a_{i,i} = \lambda$	$a_{j,i} = \lambda$
$E^{\mathrm{sw}}_{i,j}$	$E^{\mathrm{m}}_{i,\lambda}$	$E^{\mathrm{a}}_{j,i,\lambda}$
(wie zuvor)	(wie zuvor)	(Indices vertauscht)

Dreiecksmatrizen

Elementare Zeilen- und Spaltenumformungen nennt man “elementare Umformungen”.

Um den Rang einer Matrix zu bestimmen, kann man also versuchen elementare Umformungen anzuwenden bis man eine einfachere Form bekommt.
Es reicht z.B. auf folgende Gestalt zu kommen
(nennen wir sie “ℓ-Dreiecksform”, Name nicht Standard):

$$\begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,\ell} & a_{1,\ell+1} & \cdots & a_{1,n} \\ 0 & a_{2,2} & \cdots & a_{2,\ell} & a_{2,\ell+1} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{\ell,\ell} & a_{\ell,\ell+1} & \cdots & a_{\ell,n} \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix}$$

d.h. $a_{i,i} \neq 0$ für $1 \leq i \leq \ell$, und $a_{i,j} = 0$ wenn $i > \ell$ oder $j < i$.

Theorem

Der Rang einer Matrix in ℓ-Dreiecksform ist ℓ.

Beweis:

- Der (Zeilen)rang ist $\leq \ell$, weil nur die ersten ℓ Zeilen ungleich Null sind.
- Alternativ kann man argumentieren: Der Spaltenrang ist $\leq \ell$, weil alle Spalten bereits im Raum K^ℓ leben, und jede l.u. Menge in K^ℓ Größe höchstens ℓ hat.
- Der Rang ist $\geq \ell$, weil die ersten ℓ Spalten $(s_1, s_2, \ldots, s_l)$ l.u. sind:
  Beweis: Sei $\sum_{i=i}^{\ell} \lambda_i s_i = \vec{0}$. Z.Z.: Alle $\lambda_i = 0$.
  Betrachte den ℓ-ten Eintrag dieses Spaltenvektors. Der ist
  $\lambda_1 0 + \cdots + \lambda_{\ell-1} 0 + \lambda_\ell a_{\ell,\ell} = 0$, und $a_{\ell,\ell} \neq 0$, daraus folgt $\lambda_\ell = 0$.
  Betrachte den $(\ell - 1)$-ten Eintrag: Der ist
  $\lambda_1 0 + \cdots + \lambda_{\ell-1} a_{\ell-1,\ell-1} + \lambda_\ell a_{\ell-1,\ell} = 0$, mit $\lambda_\ell = 0$ und
  $a_{\ell-1,\ell-1} \neq 0$, daraus folgt $\lambda_{\ell-1} = 0$.
  Auf dieselbe Weise sieht man $\lambda_i = 0$ für alle i. □

Beispiele

Beginne mit	Vertausche Zeilen 1 und 2:	Ziehe Zeilen ab: 2 mal 1ste von 2ter, 3 mal 1ste von 3ter	Ziehe Zeilen ab: 2 mal 2te von 3ter
$A = \begin{pmatrix} 2 & 3 & 1 \\ 1 & 3 & 2 \\ 3 & 3 & 7 \end{pmatrix}$	$\begin{pmatrix} 1 & 3 & 2 \\ 2 & 3 & 1 \\ 3 & 3 & 7 \end{pmatrix}$	$\begin{pmatrix} 1 & 3 & 2 \\ 0 & -3 & -3 \\ 0 & -6 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & 3 & 2 \\ 0 & -3 & -3 \\ 0 & 0 & 7 \end{pmatrix}$

A hat also Rang 3, dh ist regulär.

Man kann sehen dass man sogar immer mit den Zeilenoperationen plus Spalten-Vertauschen auskommt:

Beginne mit	Ziehe Zeilen ab: 2 mal 1ste von 2ter, 3 mal 1ste von 3ter	Vertausche Spalten 2 und 3	Ziehe Zeilen ab: 2 mal 2te von 3ter
$A' = \begin{pmatrix} 1 & 1 & 2 \\ 2 & 2 & 5 \\ 3 & 3 & 8 \end{pmatrix}$	$\begin{pmatrix} 1 & 1 & 2 \\ 0 & 0 & 1 \\ 0 & 0 & 2 \end{pmatrix}$	$\begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 0 \\ 0 & 2 & 0 \end{pmatrix}$	$\begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$

A' hat Rang 2 (und ist insbesondere Singulär).

Gaußsches Eliminationsverfahren

Gaußsches Eliminationsverfahren

Mit (Varianten des) Gaußsche Eliminationsverfahren kann man eine Matrix in Dreiecksform bringen und alles mögliche berechnen:

- Feststellen des Rangs einer Matrix.
- Invertieren einer Matrix.
- Berechnung der Determinante einer Matrix.
- Lösen linearer Gleichungssysteme.

Als erstes und einfachstes Beispiel haben wir bereits ein Verfahren zur Bestimmung des Rangs gesehen:

Algorithmus zur Rang-Bestimmung

Theorem

Wir können mit folgendem Algorithmus eine Matrix auf Dreiecksform bringen, unter Verwendung von elementaren Zeilenumformungen sowie Spaltenvertauschungen.

Wenn A eine $m \times n$ Matrix ist, dann ist die resultierende Matrix UAT, wobei $U \in \mathrm{GL}(m)$ eine Verknüpfung elementarer Umformungen ist, und $T \in \mathrm{GL}(n)$ eine Verknüpfung von (nur) Vertauschungen.

Weil U und T invertierbar sind gilt $\mathrm{rk}(UAT) = \mathrm{rk}(A)$, und der Rang der ℓ-Dreiecksmatrix UAT ist ℓ.

Corollary

Der Algorithmus berechnet den Rang der Matrix.

Der Algorithmus

Wir gehen davon aus dass $A \neq 0$ (sonst ist A bereits in 0-Dreiecksform).

Im folgenden ändern wir die Matrix in mehreren Schritten, und bezeichnen mit $a_{i,j}$ den Eintrag in i-ter Zeile und j-ter Spalte der "jeweils aktuellen" Matrix.

- Schritt 1: Vertausche Zeilen (und, wenn nötig, Spalten) so dass $a_{1,1} \neq 0$.
- Ziehe das $\frac{a_{i,1}}{a_{1,1}}$-fache der 1-ten Zeile von der i-ten Zeile an, für alle $2 \leq i \leq m$.
- Wenn $a_{i,j} = 0$ für alle $i, j > 2$, sind wir fertig.
- Von nun an tauschen wir nur mehr Zeilen bzw Spalten i, j mit $i, j > 1$, und von Zeile 1 wird keine andere Zeile mehr abgezogen.
- Schritt 2: Tausche Zeilen/Spalten (ungleich 1), so dass $a_{2,2} \neq 0$.
- Ziehe das $\frac{a_{i,2}}{a_{2,2}}$-fache der 2-ten Zeile von der i-ten an, für alle $3 \leq i \leq m$.
- Von nun an tauschen wir nur mehr Zeilen/Spalten (ungleich 1,2) und von Zeilen 1 und 2 wird keine andere Zeile mehr abgezogen.
- ...
- Wiederhole so lange, bis wir die letzte Zeile m bearbeitet haben, oder bis nach dem k-ten Schritt $a_{i,j} = 0$ für alle $i, j > \ell$.

Analyse des Algorithmus

Warum tut der Algorithmus was wir wollen?

- Am Beginn der k-ten Schritt des Algorithmus (wenn wir Zeile k bearbeiten) sind Spalten $1, \ldots, k-1$ bereits in Dreiecksform. Nach dem k-ten Schritt sind Spalten $1, \ldots, k-1, k$ in Dreiecksform.
  Beweis: Mit Induktion. Stimmt für $k = 1$. Angenommen es stimmt für k. Es werden im k-ten Schritt nur Zeilen/Spalten mit Index $\geq k$ geändert, und dann nur Vielfache der (neuen) Zeile k zu Zeilen darunter addiert. Die $a_{i,j}$ mit $i, j < k$ ändern sich nicht. Weiteres ist nach dem Schritt $a_{k,k} \neq 0$, und $a_{k,\ell} = 0$ für $\ell > k$.
- Wenn der Algorithmus nach k Schritten terminiert, dann ist die resultierende Matrix in Dreiecksform.
  Grund: Der Algortihmus terminiert nach Schritt k wenn entweder alle m Zeilen abgearbeitet sind, oder wenn $a_{i,j} = 0$ für alle $i, j > k$. Und die Spalten $1, \ldots, k$ sind in Dreiecksform, daher ist die gesamte Matrix in Dreiecksform.
- Der Algorithmus terminiert bei einer $m \times n$-Matrix nach höchstens $\min(m, n)$ Schritten.

Lineare Gleichungssysteme

Derselbe Algorithmus kann verwendet werden, um lineare Gleichungssysteme zu lösen. Was ist ein lineares Gleichungssystem?

- Ein Bsp: $x + y = 3$, $2x - y = 7$, $7x - 5y = 0$
  In dem Fall: $m := 3$ Gleichungen, $n := 2$ Variablen.
- Wir können das in Matrix-Notation mit einer $m \times n$ Matrix schreiben:
$$\begin{pmatrix} 1 & 1 \\ 2 & -1 \\ 5 & -5 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 3 \\ 7 \\ 0 \end{pmatrix}.$$
- Allgemein: Ein lineares Gleichungssystem ist ein Paar $A, \vec{b}$ mit A $m \times n$ Matrix und $\vec{b} \in K^m$; kurz $A\vec{x} = \vec{b}$ geschrieben.
  Ein $\vec{c} \in \mathbb{K}^n$ ist Lösung wenn $A\vec{c} = \vec{b}$.
- Spezialfall: $\vec{b} = \vec{0}$, oder äquivalent: $\vec{0}$ ist Lösung des Systems.
  So ein Gleichungssystem heißt "homogen".
- $\vec{c}$ ist Lösung des homogenen Systems $A\vec{x} = 0$ ist also äquivalent zu: $\vec{c} \in \ker(A)$.

Eindeutige und nichteindeutige Lösungen

- Natürlich müssen (inhomogene) Gleichungssysteme nicht lösbar sein. Bsp:
$$\begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$
- Wenn A eine invertierbare $n \times n$-Matrix ist, dann hat $A\vec{x} = \vec{b}$ die eindeutige Lösung $\vec{c} = A^{-1}\vec{b}$.
- Wenn A nicht injektiv ist, und eine Lösung hat, dann hat A mehrere Lösungen:

  Wenn $\vec{c}_0$ Lösung von $A\vec{x} = \vec{b}$ ist, dann ist $\vec{c}$ Lösung gdw $\vec{c} = \vec{c}_0 + \vec{v}$ mit $v \in \ker(A)$.

  (Beweis: $A\vec{c} = A\vec{c}_0$ gdw $A(\vec{c} - \vec{c}_0) = A\vec{c} - A\vec{c}_0 = \vec{0}$ gdw $\vec{c} - \vec{c}_0 \in \ker(A)$ gdw $\vec{c} = \vec{c}_0 + \vec{v}$ mit $\vec{v} \in \ker(A)$.)

  In anderer Notation: Die Lösungsmenge ist $\{\vec{c}_0 + \vec{v} : \vec{v} \in \ker(A)\}$.

  Achtung: Anders als $\ker(A)$ ist das ist kein Unterraum von V (wenn $\vec{c}_0 \neq 0$).

Notation

Theorem

$A\vec{x} = \vec{b}$ ist lösbar genau dann wenn $\vec{b}$ Linearkombination der Spalten von A ist.

Beweis: Nach Definition von Linearkombination: Sei $\vec{S}_1, \ldots, \vec{S}_n$ die Spalten von A. Dann ist $\vec{c}$ Lösung genau dann wenn $\sum_{i=1}^{n} c_i \vec{S}_i = \vec{b}$ □

Das ist also genau dann
der Fall wenn $\mathrm{rk}(A) = \mathrm{rk}(A|b)$, wobei $A|b$ die folgende $m \times (n+1)$-Matrix ist:

$$\begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} & b_1 \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} & b_m \end{pmatrix} \text{ auch geschrieben als } \left(\begin{array}{cccc|c} a_{1,1} & a_{1,2} & \cdots & a_{1,n} & b_1 \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} & b_m \end{array}\right)$$

Dreiecksform

Im Fall $A\vec{x} = \vec{b}$ mit A in Dreiecksform ($a_{i,i} \neq 0$ für $i \leq \ell$) ist es einfach, die Lösungsmenge zu finden:

$$A = \left(\begin{array}{cccccccc|c} a_{1,1} & a_{1,2} & \cdots & a_{1,\ell} & a_{1,\ell+1} & \cdots & a_{1,n} & b_1 \\ 0 & a_{2,2} & \cdots & a_{2,\ell} & a_{2,\ell+1} & \cdots & a_{2,n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & a_{\ell,\ell} & a_{\ell,\ell+1} & \cdots & a_{\ell,n} & b_\ell \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & b_\ell \\ \vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & b_m \end{array}\right)$$

- Fall 1: $\ell < m$ und ein $b_k \neq 0$ für $\ell + 1 \leq k \leq m$:
  Dann ist das Gleichungssystem unlösbar.
- Fall 2: $\ell = n$:
  Es gibt eine eindeutige Lösung, die direkt "von unten nach oben" abgelesen werden kann.
- Fall 3: $\ell < n$:
  Die Lösungen sind ein $n - \ell$-dim Teilraum von K^n:
  $x_{\ell+1}, \ldots, x_n$ sind beliebig ("Parameter"), Rest wie Fall 2.

Dreiecksform: Fall 1

$$\left(\begin{array}{cccccccc|c} a_{1,1} & a_{1,2} & \cdots & a_{1,\ell} & a_{1,\ell+1} & \cdots & a_{1,n} & b_1 \\ 0 & a_{2,2} & \cdots & a_{2,\ell} & a_{2,\ell+1} & \cdots & a_{2,n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & a_{\ell,\ell} & a_{\ell,\ell+1} & \cdots & a_{\ell,n} & b_\ell \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & b_\ell \\ \vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & b_m \end{array}\right)$$

Fall 1: $\ell < m$ und ein $b_k \neq 0$ für $\ell + 1 \leq k \leq m$:
Dann ist das System unlösbar.
Beweis: Sei $b_k \neq 0$. Sei $\vec{S}_i$ die i-te Spalte von A. Der k-te Eintrag in $\vec{S}_i$ ist 0 für alle i. Daher ist der k-te Eintrag von $A\vec{c} = c_1 \cdot \vec{S}_1 + \cdots c_n \cdot \vec{S}_n$ ebenfalls 0, und damit $A\vec{c} \neq \vec{b}$. □

Dreiecksform: Fall 2

$$\left(\begin{array}{cccc|c} a_{1,1} & a_{1,2} & \cdots & a_{1,n} & b_1 \\ 0 & a_{2,2} & \cdots & a_{2,n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & a_{n,n} & b_n \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 \end{array}\right)$$

Fall 2: $\ell = n$: Es gibt eine eindeutige Lösung:

- $x_n \cdot a_{n,n} = b_n$, das gibt und das eindeutige x_n (weil $a_{n,n} \neq 0$).
- $x_{n-1} \cdot a_{n-1,n-1} + x_n a_{n-1,n} = b_{n-1}$, das bestimmt eindeutig x_{n-1} (weil $a_{n-1,n-1} \neq 0$).
- Etc

Dreiecksform: Fall 3

Der Fall $A\vec{x} = \vec{b}$ mit A in Dreiecksform ($a_{i,i} \neq 0$ für $i \leq \ell$) ist einfach:

$$\left(\begin{array}{cccccccc|c} a_{1,1} & a_{1,2} & \cdots & a_{1,\ell} & a_{1,\ell+1} & \cdots & a_{1,n} & b_1 \\ 0 & a_{2,2} & \cdots & a_{2,\ell} & a_{2,\ell+1} & \cdots & a_{2,n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & a_{\ell,\ell} & a_{\ell,\ell+1} & \cdots & a_{\ell,n} & b_\ell \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \end{array}\right)$$

Fall 3: $\ell < n$: Für jedes $(x_{\ell+1}, \ldots, x_n) \in K^n$ gibt es eindeutige $x_0, \ldots, x_\ell$ so dass $(x_0, \ldots, x_n)$ Lösung ist (und das sind alle Lösungen).

- Wähle beliebige $x_{\ell+1}, \ldots, x_n$.
- Aus $a_{\ell,\ell} \cdot x_\ell + a_{\ell,\ell+1} \cdot x_{\ell+1} \cdots a_{\ell,n} \cdot x_n = b_\ell$ folgt der eindeutige Wert für x_ℓ (weil $a_{\ell\ell} \neq 0$).
- Aus $a_{\ell-1,\ell-1} \cdot x_{\ell-1} + a_{\ell-1,\ell} \cdot x_\ell \cdots a_{\ell-1,n} \cdot x_n$ folgt der eindeutige Wert für $x_{\ell-1}$ (weil $a_{\ell-1,\ell-1} \neq 0$).
- Auf diese Weise bestimmt man alle $x_\ell, x_{\ell-1}, \ldots, x_1$ (in Abhängigkeit von $(x_{\ell+1}, \ldots, x_n)$).

Gauß Elimination

Wir führen nun dasselbe Verfahren wie zuvor für die Matrix A durch, wobei wir aber alle Zeilenoperation auch auf $\vec{b}$ anwenden.
Dann gilt:

- Eine elementare Zeilenumformung der $m \times (n+1)$-Matrix $A|\vec{b}$ ergibt $U(A|\vec{b})$, und das ist dasselbe wie $(UA)|(U\vec{b})$.
  (Zwei Zeilen der erweiterten Matrix $A|\vec{b}$ zu vertauschen ist dasselbe wie die zwei Zeilen separat in A und in $\vec{b}$ zu tauschen und dann die Ergebnisse zusammenzukleben, und analog für die anderen Spaltenoperationen.)
- $A\vec{x} = \vec{b}$ gdw $UA\vec{x} = U\vec{b}$, weil U invertierbar.
- Zeilenumformungen ändern also die Lösungsmenge nicht.
- Vertauschung der n Spalten von A entspricht "Umordnung/Umbenennung" der Unbekannten $x_1, \ldots, x_n$.
  Wenn wir zum Beispiel Spalten 1 und 3 vertauschen, und für die neue Matrix die Lösung $(8, 12, 5, 7)$ bekommen, dann hat die ursprüngliche Matrix die Lösung $(5, 12, 8, 7)$.

(Einfache) Beispiele

- Fall 1:
$$\left(\begin{array}{ccc|c} 1 & 0 & 1 & 1 \\ 2 & 3 & -2 & 0 \\ 3 & 3 & -1 & -1 \end{array}\right) \text{ wird zu } \left(\begin{array}{ccc|c} 1 & 0 & 1 & 1 \\ 0 & 3 & -4 & -2 \\ 0 & 0 & 0 & -2 \end{array}\right)$$
Daher: Keine Lösung.
- Fall 2:
$$\left(\begin{array}{ccc|c} 1 & 0 & 1 & 1 \\ 2 & 3 & -2 & 0 \\ 3 & 3 & 3 & -1 \end{array}\right) \text{ wird zu } \left(\begin{array}{ccc|c} 1 & 0 & 1 & 1 \\ 0 & 3 & -4 & -2 \\ 0 & 0 & 4 & -2 \end{array}\right)$$
Daher: Eindeutige Lösung.
$4x_3 = -2 \to x_3 = -\frac{1}{2}$.
$3x_2 - 4x_3 = -2 \to x_2 = -\frac{4}{3}$.
$x_1 + x_3 = 1 \to x_1 = \frac{3}{2}$.
Lösung: $\begin{pmatrix} 3/2 \\ -4/3 \\ -1/2 \end{pmatrix}$. Probe: $\begin{pmatrix} 1 & 0 & 1 \\ 2 & 3 & -2 \\ 3 & 3 & 3 \end{pmatrix} \begin{pmatrix} 3/2 \\ -4/3 \\ -1/2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$

Mehr (einfache) Beispiele

- Fall 3:
  $\left(\begin{array}{ccc|c} 1 & 0 & 1 & 1 \\ 2 & 1 & -2 & 0 \\ 3 & 1 & -1 & 1 \end{array}\right)$ wird zu $\left(\begin{array}{ccc|c} 1 & 0 & 1 & 1 \\ 0 & 1 & -4 & -2 \\ 0 & 0 & 0 & 0 \end{array}\right)$

  Die Lösungen sind daher parametrisiert mit einem Parameter x_3.

  $x_2 - 4x_3 = -2$, d.h. $x_2 = 4x_3 - 2$.

  $x_1 + x_3 = 1$, d.h. $x_1 = 1 - x_3$.

  Daher: Lösungen sind alle Vektoren der Form $\begin{pmatrix} 1 - x_3 \\ 4x_3 - 2 \\ x_3 \end{pmatrix}$ für $x_3 \in K$.

  Probe: $\begin{pmatrix} 1 & 0 & 1 \\ 2 & 1 & -2 \\ 3 & 1 & -1 \end{pmatrix} \begin{pmatrix} 1 - x_3 \\ 4x_3 - 2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$

Noch mehr (einfache) Beispiele

- Beispiel Fall 3 mit Spaltenwechsel:
$$\left(\begin{array}{ccc|c} 1 & 1 & 2 & 3 \\ 2 & 2 & 5 & 2 \\ 3 & 3 & 8 & 1 \end{array}\right) \text{ wird zu } \left(\begin{array}{ccc|c} 1 & 2 & 1 & 3 \\ 0 & 1 & 0 & -4 \\ 0 & 0 & 0 & 0 \end{array}\right)$$
wobei aber Spalten 2 und 3 vertauscht wurden.

  Nenne Variablen nach Vertauschung y_1, y_2, y_3.
  Lösungen: Parameter y_3, $y_2 = -4$ und $y_1 = 11 - y_3$.
  In zurückgetauschten Variablen ergibt das:
  Lösung: $\begin{pmatrix} 11 - x_2 \\ x_2 \\ -4 \end{pmatrix}$
  Probe: $\begin{pmatrix} 1 & 1 & 2 \\ 2 & 2 & 5 \\ 3 & 3 & 8 \end{pmatrix} \begin{pmatrix} 11 - x_2 \\ x_2 \\ -4 \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}$

Spalten-Vertauschung

Der Gauss-Algorithmus benötigt Spalten-Vertauschung (im Schritt k) nur, wenn (in Schritt k) die folgenden Situation auftritt:

$$\left(\begin{array}{cccc|cccc} a_{1,1} & a_{1,2} & \cdots & a_{1,k} & a_{1,k+1} & a_{1,k+2} & \cdots & a_{1,n} \\ 0 & a_{2,2} & \cdots & a_{2,k} & a_{2,k+1} & a_{2,k+2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \\ 0 & 0 & \cdots & a_{k,k} & a_{k,k+1} & a_{k,k+2} & \cdots & a_{k,n} \\ 0 & 0 & \cdots & 0 & 0 & a_{k+1,k+2} & \cdots & a_{k+1,n} \\ 0 & 0 & \cdots & 0 & 0 & a_{k+2,k+2} & \cdots & a_{k+2,n} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & 0 & a_{m,k+2} & \cdots & a_{m,n} \end{array}\right)$$

(Die ersten k Spalten sind bereits in Dreiecksform, aber unter k stehen in Spalte $k+1$ nur 0, daher kann man nicht durch Vertauschen von Zeilen $a_{k+1,k+1} \neq 0$ erreichen sondern muss Spalten vertauschen.)

In dem Fall sind die ersten $k+1$ Spalten linear abhängig (weil alle in K^k).

Im Fall einer invertierbaren $n \times n$-Matrix verwendet der Algorithmus also ausschließlich elementare Zeilenumformungen.

Was der Algorithmus benötigt

Der grundlegende Gauss-Algorithmus um eine Matrix in Dreiecksform zu bringen benötigt nur "Addition des Vielfachen einer Zeile zu einer anderen" und "Vertauschung von Zeilen", nicht "Multipliziere Zeile mit λ". (Und nur im Fall von singulären Matrizen "Vertauschung von Spalten")

Berechnen vom Kern und Eigenvektoren

Kern:

- $\ker A$ ist der Unterraum der Lösungen von $A \cdot \vec{x} = \vec{0}$.
- Wir können also $\ker A$ berechnen indem wir den Algorithmus auf $A|\vec{0}$ anwenden.

Eigenvektoren (für eine $n \times n$ Matrix A):

- Wir wollen die Eigenvektoren von A zum (vermuteten) Eigenwert λ berechnen.
- Dazu setzen wir $B := A - \lambda I$.
  Dann ist $A\vec{v} = \lambda\vec{v}$ gdw $B\vec{v} = A\vec{v} - \lambda\vec{v} = 0$.
  D.h. den Vektorraum der Eigenvektoren zu λ ist der Kern von B.
- Es wird jedenfalls die Lösung $\vec{0}$ geben, und wenn $\vec{0}$ die einzige/eindeutige Lösung ist dann war λ gar kein Eigenwert.
- Bemerkung: Die Eigenwerte zu finden ist ein anderes Problem.

Beispiel

- Wir wissen dass $A = \begin{pmatrix} -4 & -10 \\ 2 & 5 \end{pmatrix}$ eine Projektion ist (weil $A^2 = A$), zu unbekanntem W_1, W_2.
- Es gibt also die Eigenwerte 0 und 1,
  W_1 ist der Raum der Eigenvektoren zum Eigenwert 1,
  $W_2 = \ker(A)$ der Raum der Eigenvektoren zum Eigenwert 0.
- Der Algorithmus für $A|\vec{0}$ liefert:
  $$W_2 = \left\{ \begin{pmatrix} -\frac{10}{4} x_2 \\ x_2 \end{pmatrix} : x_2 \in \mathbb{R} \right\} = \left\{ \begin{pmatrix} -10t \\ 4t \end{pmatrix} : t \in \mathbb{R} \right\}$$
- Für $B = A - 1 \cdot I = \begin{pmatrix} -5 & -10 \\ 2 & 4 \end{pmatrix}$ liefert der Algorithmus für $B|\vec{0}$:
  $$W_1 = \left\{ \begin{pmatrix} -2x_2 \\ x_2 \end{pmatrix} : x_2 \in \mathbb{R} \right\}$$
- Proben: $A \cdot \begin{pmatrix} -10 \\ 4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, und $A \cdot \begin{pmatrix} -2 \\ 1 \end{pmatrix} = \begin{pmatrix} -2 \\ 1 \end{pmatrix}$.

Invertieren von Matrizen

Eine Variante des Algorithmus berechnet das Inverse einer Matrix:

- Die $n \times n$ Matrix A ist invertierbar gdw der Algorithmus eine Dreiecksform der folgenden Form liefert:

$$\begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ 0 & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{n,n} \end{pmatrix}, \text{ mit } a_{ii} \neq 0.$$

- In diesem Fall verwendet der Algorithmus nur Zeilenumformungen.
- Mit weiteren Zeilenumformungen erhalten wir die Diagonalform: Ziehe das $\frac{a_{1,2}}{a_{2,2}}$-fache der 2. Spalte von der 1. ab, dann das (neue) $\frac{a_{1,3}}{a_{3,3}}$-fache der 3. Spalte von der 1. und das $\frac{a_{2,3}}{a_{3,3}}$-fache der 3. von der 2., etc Das ergibt

$$\begin{pmatrix} a_{1,1} & 0 & \dots & 0 \\ 0 & a_{2,2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{n,n} \end{pmatrix}$$

bzw nach "i-te Spalte mal $\frac{1}{a_{ii}}$" für alle i gibt das die Einheitsmatrix I.

Invertieren von Matrizen (Forts.)

- Wir bekommen also $TA = I$ für eine bestimmte Verknüpfungen T von elementaren Spaltenoperationen.
- Daher ist $A^{-1} = T$.
- Wenn wir uns also beim Durchführen des Algorithmus merken, welche Operation wir verwenden haben dann können wir damit auch A^{-1} bestimmen.

  Genauer: Wir müssen nur das jeweils aktuelle T aktualisieren.
- Vergleiche: Beim linearen Gleichungssystem müssen wir uns merken wie wir Spalten vertauscht haben.
- T ist Verknüpfung $T_k T_{k-1} \cdots T_2 T_1$ elementarer Zeilenumformungen T_i. Mit T_i ist auch T_i^{-1} eine elementare Zeilenumformung. Daher ist $A = T^{-1} = T_1 - 1 T_2^{-1} \cdots T_k^{-1}$ eine Verknüpfung elementarere Zeilenumformungen.
- Wie bekommen daher: Jede invertierbare $n \times n$-Matrix ist eine Verknüpfung elementarer Zeilenumformungen. (Dasselbe gilt für Spaltenumformungen.)

Invertieren von Matrizen: Ein Beispiel

Suchen T mit $TA = I$, für $A = \begin{pmatrix} 1 & 3 & 2 \\ 2 & 3 & 1 \\ 3 & 3 & 7 \end{pmatrix}$.

	Beginne mit	Z2: -2x Z1, Z3: -3x Z1	Z3: -2x Z2	Z1: +1x Z2	Z1: $+1/7$x Z3 Z2: $+3/7$x Z3	Z2 x$-1/3$ Z3 x$1/7$
Aktuelles A	$\begin{pmatrix} 1 & 3 & 2 \\ 2 & 3 & 1 \\ 3 & 3 & 7 \end{pmatrix}$	$\begin{pmatrix} 1 & 3 & 2 \\ 0 & -3 & -3 \\ 0 & -6 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & 3 & 2 \\ 0 & -3 & -3 \\ 0 & 0 & 7 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 & -1 \\ 0 & -3 & -3 \\ 0 & 0 & 7 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 7 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$
Neues T		$\begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -3 & 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 & \frac{1}{7} \\ 0 & 1 & \frac{3}{7} \\ 0 & 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{3} & 0 \\ 0 & 1 & \frac{1}{7} \end{pmatrix}$
Gesamt T	$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -3 & 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 1 & -2 & 1 \end{pmatrix}$	$\begin{pmatrix} -1 & 1 & 0 \\ -2 & 1 & 0 \\ 1 & -2 & 1 \end{pmatrix}$	$\begin{pmatrix} -\frac{6}{7} & \frac{5}{7} & \frac{1}{7} \\ -\frac{11}{7} & \frac{1}{7} & \frac{3}{7} \\ 1 & -2 & 1 \end{pmatrix}$	$\begin{pmatrix} -\frac{6}{7} & \frac{5}{7} & \frac{1}{7} \\ \frac{11}{21} & -\frac{1}{21} & -\frac{1}{7} \\ \frac{1}{7} & -\frac{2}{7} & \frac{1}{7} \end{pmatrix}$

Probe: $\begin{pmatrix} 1 & 3 & 2 \\ 2 & 3 & 1 \\ 3 & 3 & 7 \end{pmatrix} \cdot \begin{pmatrix} -\frac{6}{7} & \frac{5}{7} & \frac{1}{7} \\ \frac{11}{21} & -\frac{1}{21} & -\frac{1}{7} \\ \frac{1}{7} & -\frac{2}{7} & \frac{1}{7} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

Andere Körper

Andere Körper

- In den Beispielen haben wir bis jetzt meist stillschweigend in $\mathbb{R}$ (oder $\mathbb{Q}$ oder $\mathbb{C}$) gerechnet.
- Die Sätze die wir bewiesen haben, und insbesondere die Varianten des Gauss-Verfahrens, funktionieren aber für Vektorräume über beliebigen Körpern K.

Teilkörper

- Wenn K_1 ein Teilkörper von K_2 ist (Bsp: $K_1 = \mathbb{Q}$ und $K_2 = \mathbb{C}$), dann haben Rechnungen mit Zahlen in K_1 dasselbe Ergebnis in K_1 und in K_2.
- Insbesondere: Wenn A Einträge in K_1 hat, dann liefer die diversen Gauss-Verfahren dieselben Ergebnisse in K_1 und K_2.
- Insbesonder: Der Rang von A ist derselbe in K_1 und K_2.
- Insbesondere: Es gibt für ein lineares Gleichungssystem mit Koeffizienten in K_1 dieselben Lösungen in K_1 und K_2. Expliziter:
  - Es gibt Lösung in K_1 gdw es gibt Lösung in K_2.
  - Lösung ist eindeutig in K_1 gdw Lösung ist eindeutig ist in K_2.
  - Nicht-eindeutige Lösungen sind durch dieselben Vektoren erzeugt. Bsp: Wenn die Lösungen in $\mathbb{Q}$ als $(2 \cdot x_2 - 1, x_2)$ parametrisiert sind, dann auch in $\mathbb{C}$ (aber in $\mathbb{C}$ gibt es dann auch die Lösung $(2 \cdot i - 1, i)$, die es in $\mathbb{Q}$ nicht gibt).

Teilkörper (Forts.)

(K_1 Teilkörper von K_2, A Matrix mit Einträgen in K_1.)
Weil Lösung von Gleichungssystemen absolut zwischen K_1 und K_2, gilt insbesondere:

- Der Kern von A wird in K_1 und K_2 durch dieselben Vektoren erzeugt (und hat dieselbe Dimension).
- Für $\lambda \in K_1$ ist λ EW von A in K_1 (d.h. es gibt $\vec{v} \in K_1^n$ mit $\vec{v} \neq \vec{0}$ und $A\vec{v} = \lambda\vec{v}$) gdw λ EW von A in K_2; der Teilraums der λ-Eigenvektoren wird und K_1 und K_2 durch dieselben Vektoren erzeugt und hat dieselbe Dimension.
- Aber: "Es gibt Eigenwert" ist z.B. nicht dasselbe in K_1 und K_2. Als Bsp haben wir gesehen: Rotation $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ hat keine Eigenwerte in $\mathbb{R}$, sehr wohl aber in $\mathbb{C}$.
- Dagegen ist A^{-1} wieder absolut zwischen K_1 und K_2.

Ganz andere Körper

- Es gibt natürlich Körper K_1 und K_2 bei denen keiner Teilkörper des anderen ist.
- Bsp: $K_1 = \mathbb{Z}_p$ und $K_2 = \mathbb{Z}_q$ für Primzahlen $p \neq q$. Weil K_1 und K_2 unterschiedliche Ordnungen haben, kann man weder K_1 in K_2 einbetten noch umgekehrt.
- Bsp: $K_1 = \mathbb{Z}_p$ und $K_2 = \mathbb{Q}$. (Aus demselben Grund.)
- Konkret für dieses K_1, K_2: Ein $\overline{n} \in \mathbb{Z}_p$ sieht zwar aus wie ein $n \in \mathbb{Q}$, und $\overline{n+m} = \overline{n} + \overline{m}$ und $\overline{n \cdot m} = \overline{n} \cdot \overline{m}$, aber $\overline{n} = \overline{0}$ impliziert nicht $n = 0$ (Bsp: $\overline{p} = \overline{0}$). Daher kann der Gauss Algorithmus ganz unterschiedliche Ergebnisse liefern.

Ein Bsp in $K = \mathbb{Z}_7$

Ein Beispiel im endlichen Körper $K = \mathbb{Z}_7$.
Der besseren Lesbarkeit schreiben wir $0, 1, \ldots, 6$ statt $\bar{0}, \bar{1}, \ldots, \bar{6}$ für die Elemente von $\mathbb{Z}_7$.
In $\mathbb{R}$ hatten wir bereits das folgende Beispiel für eine Rang 3 Matrix A:

Beginne mit	Vertausche Zeilen 1 und 2:	Ziehe Zeilen ab: 2 mal 1ste von 2ter, 3 mal 1ste von 3ter	Ziehe Zeilen ab: 2 mal 2te von 3ter
$A = \begin{pmatrix} 2 & 3 & 1 \\ 1 & 3 & 2 \\ 3 & 3 & 7 \end{pmatrix}$	$\begin{pmatrix} 1 & 3 & 2 \\ 2 & 3 & 1 \\ 3 & 3 & 7 \end{pmatrix}$	$\begin{pmatrix} 1 & 3 & 2 \\ 0 & -3 & -3 \\ 0 & -6 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & 3 & 2 \\ 0 & -3 & -3 \\ 0 & 0 & 7 \end{pmatrix}$

Weil wir hier zufälligerweise nur Addition und Multiplikation verwendet haben, und nicht Division, können wir in $\mathbb{Z}_7$ dasselbe einfach "mod 7" rechnen:

Beginne mit	Vertausche Zeilen 1 und 2:	Ziehe Zeilen ab: 2 mal 1ste von 2ter, 3 mal 1ste von 3ter	Ziehe Zeilen ab: 2 mal 2te von 3ter
$A' = \begin{pmatrix} 2 & 3 & 1 \\ 1 & 3 & 2 \\ 3 & 3 & 0 \end{pmatrix}$	$\begin{pmatrix} 1 & 3 & 2 \\ 2 & 3 & 1 \\ 3 & 3 & 0 \end{pmatrix}$	$\begin{pmatrix} 1 & 3 & 2 \\ 0 & 4 & 4 \\ 0 & 1 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & 3 & 2 \\ 0 & 4 & 4 \\ 0 & 0 & 0 \end{pmatrix}$

A' hat also Rang 2, dh ist singulär.

Rechnen im Körper $\mathbb{Z}_p$

- Plus, Minus und Mal sind klar (und im Bsp hat das ausgereicht): $\overline{a} + \overline{b} = \overline{a+b}$, $\overline{a} - \overline{b} = \overline{a-b}$, $\overline{a} \cdot \overline{b} = \overline{a \cdot b}$.
- Division ist weniger klar: Z.b. gilt $\mathbb{Z}_7$: $\frac{\overline{5}}{\overline{4}} = \overline{3}$.
- Ausprobieren, oder (bei größerem n), Euklidischer Algorithmus:
- Um $x = \frac{\overline{m}}{\overline{n}}$ zu berechnen, finden wir erst $y = \frac{\overline{1}}{\overline{n}}$, dann ist $x = y \cdot \overline{m}$. Und $\mathrm{ggT}(n,p) = 1$ (weil $\bar{n} \neq \bar{0}$), daher gilt: $an + bp = 1$, damit ist $\overline{a} \cdot \overline{n} = \overline{1}$, und $\overline{a}$ ist das gesuchte y.
- Beispiel: $\frac{\overline{1}}{\overline{12}}$ in $\mathbb{Z}_{41}$:

$$\begin{array}{rcl|rcl} 41 &=& 3\cdot 12+5 & 1 &=& -2\cdot 12+5\cdot(41-3\cdot 12) = 5\cdot 41-17\cdot 12 \\ 12 &=& 2\cdot 5+2 & 1 &=& 5-2\cdot(12-2\cdot 5) = -2\cdot 12+5\cdot 5 \\ 5 &=& 2\cdot 2+1 & 1 &=& 5-2\cdot 2 \\ 2 &=& 2\cdot 1 & & & \end{array}$$

D.h. $\overline{1} = -\overline{17}\cdot\overline{12}$, d.h. $\frac{\overline{1}}{\overline{12}} = -\overline{17} = \overline{24}$. Probe: $41 \mid (24\cdot 12 - 1)$

Ein Bsp mit Division, in $\mathbb{Z}_5$

Löse $\overline{2}x_0 + x_1 = \overline{3}$, $\overline{3}x_0 + \overline{2}x_1 = 1$.

(In $\mathbb{Z}_5$ ist $\overline{2} \cdot \overline{3} = \overline{1}$. Wir schreiben wieder n statt $\overline{n}$.)

$$\left(\begin{array}{cc|c} 2 & 1 & 3 \\ 3 & 2 & 1 \end{array}\right) \quad \begin{array}{c}\text{Z1: x3} \\ \left(\begin{array}{cc|c} 1 & 3 & 4 \\ 3 & 2 & 1 \end{array}\right)\end{array} \quad \begin{array}{c}\text{Z2: -3x Z1} \\ \left(\begin{array}{cc|c} 1 & 3 & 4 \\ 0 & 3 & 4 \end{array}\right)\end{array}$$

Daher: $\overline{3}x_2 = \overline{4}$, d.h. $x_2 = \overline{3}$, und $x_1 + \overline{3} \cdot \overline{3} = \overline{4}$, d.h. $x_1 = \overline{0}$.

Determinante

Determinante

Sei $A = (a_{i,j})_{1 \leq i \leq n, 1 \leq j \leq n}$ eine $n \times n$-Matrix.

Definition

$\det(A) = \sum_{\pi \in S_n} \operatorname{sgn}(\pi) a_{\pi(1),1} a_{\pi(2),2} \cdots a_{\pi(n),n}$

$\det\left(\begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,n} \end{pmatrix}\right)$ wird auch geschrieben als $\begin{vmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,n} \end{vmatrix}$

Für 2x2 Matrizen:

S_2 besteht aus $\begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}$ (gerade), und $\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$ (ungerade), daher:

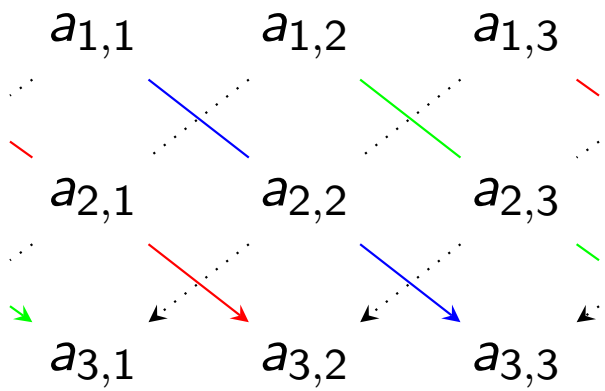
$$\begin{vmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{vmatrix} = a_{1,1} \cdot a_{2,2} - a_{1,2} \cdot a_{2,1}$$

Determinante 3x3

S_3 besteht aus den geraden $\begin{pmatrix}1&2&3\\1&2&3\end{pmatrix}, \begin{pmatrix}1&2&3\\2&3&1\end{pmatrix}, \begin{pmatrix}1&2&3\\3&1&1\end{pmatrix}$

und den ungeraden $\begin{pmatrix}1&2&3\\3&2&1\end{pmatrix}, \begin{pmatrix}1&2&3\\1&3&2\end{pmatrix}, \begin{pmatrix}1&2&3\\2&1&3\end{pmatrix}$. Daher:

$$\begin{vmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{vmatrix} = \begin{array}{l} \underline{a_{1,1} \cdot a_{2,2} \cdot a_{3,3}} + \underline{a_{1,2} \cdot a_{2,3} \cdot a_{3,1}} + \underline{a_{1,3} \cdot a_{2,1} \cdot a_{3,2}} \\ - a_{1,3} \cdot a_{2,2} \cdot a_{3,1} - a_{1,1} \cdot a_{2,3} \cdot a_{3,2} - a_{1,2} \cdot a_{2,1} \cdot a_{3,3} \end{array}$$



Bemerkungen

Bemerkungen:

- Im $\mathbb{R}^n$ ist $|\det(A)|$ das Volmen des durch die Zeilen (oder: Spalten) aufgespannten "Parallelotops".
- Die Definition nützlich für die Theorie, aber für Berechnung für größere n sehr ineffizent.

Spezielle Matrizen:

- $\det(I_n) = 1$
- Allgemeiner: Wenn $a_{i,j} = 0$ für $i > j$, dann $\det(A) = \prod_{i=1}^{n} a_{i,i}$

$$\begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ 0 & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{n,n} \end{pmatrix}$$

Beweis: Für $a_{\pi(1),1} a_{\pi(2),2} \cdots a_{\pi(n),n} \neq 0$ muss gelten: $\pi(1) = 1$ (weil alle $a_{j,1} = 0$ für $j > 1$), daher $\pi(2) = 2$ (1 ist nicht mehr möglich, weil π injektiv ist, und $a_{j,2} = 0$ für $j > 2$), etc.

Determinante von Produkten

Die zentrale Eigenschaft von Determinanten ist:

Theorem (ohne Beweis)

$\det(AB) = \det(A)\det(B)$

Daraus folgt:

- Wenn A invertierbar, dann $\det(A) \neq 0$.
  Beweis: $1 = \det(I) = \det(AA^{-1}) = \det(A)\det(A^{-1})$.
- Wenn T, S invertierbar, dann $\det(A) = 0$ gdw $\det(TAS) = 0$.
  Beweis: $\det(TAS) = \det(T)\det(A)\det(S)$, und $\det(T), \det(S) \neq 0$.
- $\det(A) \neq 0$ gdw A invertierbar.
  Beweis: A nicht invertierbar $\rightarrow \mathrm{rk}(A) = \ell < n \rightarrow$ es gibt $T, S \in \mathrm{GL}(n)$ s.d. $B := TAS$ "kanonische Matrix vom Rang ℓ" ist. (D.h. $b_{i,i} = 1$ wenn $1 \leq i \leq \ell$, und $b_{i,j} = 0$ sonst.) Dann ist $\det(B) = 0$.

Corollary

$\det(A) = 0$ *gdw* A *singulär.*

Determinante der elementaren Umformungen

- $\det(E^{\mathrm{sw}}_{i,j}) = -1$, d.h. $\det(E^{\mathrm{sw}}_{i,j}A) = \det(E^{\mathrm{sw}}_{i,j}A) = -\det(A)$.
  Vertauschen von zwei Zeilen bzw zwei Spalten ändert also Vorzeichen der Determinante.
- $\det(E^{\mathrm{m}}_{i,\lambda}) = \lambda$, d.h. $\det(E^{\mathrm{m}}_{i,\lambda}A) = \det(AE^{\mathrm{m}}_{i,\lambda}) = \lambda\det(A)$
  Multiplizieren einer Zeile bzw Spalte mit λ multipliziert also Determinante mit λ.
- $\det(E^{\mathrm{a}}_{i,j,\lambda}) = 1$, d.h., $\det(E^{\mathrm{a}}_{i,j,\lambda}A) = \det(AE^{\mathrm{a}}_{i,j,\lambda}) = \det(A)$.
  Addieren des vielfachen einer Zeile/Spalte zu einer anderen ändert die Determinante nicht.

  Beweis:

$$\begin{vmatrix} 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & & \vdots & & \vdots \\ 0 & \cdots & 1 & \cdots & \lambda & \cdots & 0 \\ \vdots & & \vdots & \ddots & \vdots & & \vdots \\ 0 & \cdots & 0 & \cdots & 1 & \cdots & 0 \\ \vdots & & \vdots & & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{vmatrix} = 1:$$

Wenn $a_{\pi(1),1}a_{\pi(2),2}\cdots a_{\pi(n),n} \neq 0$ dann:
$a_{\pi(j),j} \neq 0$, d.h. $\pi(j) = j$ oder $\pi(j) = i$.
Aber $\pi(j) = i$ impliziert $\pi(i) \neq i$ und daher $a_{\pi(i),i} = 0$.
Es ist also $\pi(j) = j$ und, wie bei Dreiecksmatrix folgt: $\det = \prod_{i=1}^{n} a_{i,i} = 1$. □

Determinante berechnen

- Wir können also den Gauss Algorithmus auch zum Berechnen von Determinanten verwenden:
  (Effizienter als direkte Anwendung der Definition, die $n!$ Schritte benötigen würde.)
- Wende Zeilenumformungen an bis TA eine Dreiecksform hat (wenn eine Spaltenumformung notwendig wäre, ist die Determinante 0).
  Dabei brauchen wir nur "Vertauschen" und "Addiere vielfaches von Zeile i zu Zeile j".
- Merke dabei das aktuelle "Vorzeichen" a, der das aktuelle $\det(T)$ angibt: Beginne mit $a = 1$. Für jede Zeilenvertauschung multipliziere mit -1. (Zeilen addieren ändert Faktor nicht.)
- Die Determinante von TA ist das Produkt P der Diagonalelement, und die Determinante von A ist $\det(TA)/\det(T)$, d.h. $a \cdot P$.
  (Weil $a = \pm 1$, d.h. $\frac{1}{a} = a$.)

Beispiel

Berechne $\det(A)$, für $A = \begin{pmatrix} 1 & 3 & 2 & 1 \\ 2 & 6 & 1 & 3 \\ 3 & 3 & 7 & 2 \\ 4 & 15 & 12 & 7 \end{pmatrix}$.

	Beginne mit	Z2: -2x Z1, Z3: -3x Z1, Z4: -4x Z1	Z2 $\leftrightarrow$ Z4	Z3: +2x Z2	Z4: $+1/9$Z3
Aktuelles A	$\begin{pmatrix} 1 & 3 & 2 & 1 \\ 2 & 6 & 1 & 3 \\ 3 & 3 & 7 & 2 \\ 4 & 15 & 12 & 7 \end{pmatrix}$	$\begin{pmatrix} 1 & 3 & 2 & 1 \\ 0 & 0 & -3 & 1 \\ 0 & -6 & 1 & -1 \\ 0 & 3 & 4 & 3 \end{pmatrix}$	$\begin{pmatrix} 1 & 3 & 2 & 1 \\ 0 & 3 & 4 & 3 \\ 0 & -6 & 1 & -1 \\ 0 & 0 & -3 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & 3 & 2 & 1 \\ 0 & 3 & 4 & 3 \\ 0 & 0 & 9 & 5 \\ 0 & 0 & -3 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & 3 & 2 & 1 \\ 0 & 3 & 4 & 3 \\ 0 & 0 & 9 & 5 \\ 0 & 0 & 0 & \frac{8}{3} \end{pmatrix}$
Vorzeichen a	1	1	-1	-1	-1

- Die Determinante der letzten Matrix in Dreiecksform, d.h. $\det(TA)$, ist $1 \cdot 3 \cdot 9 \cdot \frac{8}{3} = 72$.
- Das Vorzeichen a, d.h. $\det(T)$, ist -1.
- Daher ist $\det(A) = -72$.

SL(n)

Eine $n \times n$ Matrix A ist in SL(n) (der "speziellen linearen Gruppe") wenn $\det(A) = 1$.
SL(n) ist Untergruppe von GL(n): Wenn $\det(A) = \det(B) = 1$, dann ist $\det(AB) = 1$ und $\det(A^{-1}) = 1$.

Das charakteristische Polynom

Eine Anwendung der Determinante: Eigenwerte

- Determinanten sehr nützlich. Ein Beispiel:
- Fixiere $\lambda \in K$. λ ist Eigenwert der $n \times n$ Matrix A gdw.:
  Es gibt $\vec{v} \neq \vec{0}$ s.d. $A\vec{v} = \lambda\vec{v}$.
  Gdw: Es gibt $\vec{v} \neq \vec{0}$ s.d. $(\lambda I - A)\vec{v} = \vec{0}$.
  Gdw: $\ker(\lambda I - A) \neq \{\vec{0}\}$.
  Gdw: $\lambda I - A$ singulär.
  Gdw: $\det(\lambda I - A) = 0$
- $\chi_A(X) := \det(XI - A)$ (mit X Variable) heißt charakteristisches Polynom von A.
  $\chi_A(\lambda) = 0$ gdw. λ ist EW von A.
- Bsp: Für $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ ist $\chi_A(X) = \begin{vmatrix} X & 1 \\ -1 & X \end{vmatrix} = X^2 + 1$.
  In $\mathbb{R}$ hat $\chi_A(X)$ keine Nullstellen, d.h. A keine EW.
  In $\mathbb{C}$ hat $\chi_A(X)$ die Nullstellen (und A die EW) $\pm i$.

Zusammenfassung

Definition

$\chi_A(X) := \det(XI - A)$ (ein Polymon in der Variable X) heißt charakteristisches Polynom von A.

Theorem

λ ist EW von A gdw $\chi_A(\lambda) = 0$.

Sobald man einen Eigenwert von A gefunden hat, kann man die Eigenvektoren (bzw die Erzeugenden des Eigenraums) mit dem Gauss Algorithmus bestimmen. Dieser letzte Schritt ist unabhängig ob man zB in $\mathbb{Q}$, $\mathbb{R}$ oder $\mathbb{C}$ rechnet.

Diagonalisierung

- Eine $n \times n$-Matrix $D = (d_{i,j})_{1 \leq i,j \leq n}$ ist eine Diagonalmatrix, wenn sie nur auf der Diagonale nicht-Null Einträge hat, d.h., wenn $d_{i,j} = 0$ für alle $i \neq j$.
- Eine $n \times n$-Matrix A heißt diagonalisierbar, wenn es eine Basis E gibt so dass die E-Darstellung von A eine Diagonalmatrix D ist. In anderen Worten: Wenn es ein $U \in \mathrm{GL}(n)$ gibt mit $U^{-1}AU = D$.
- Es ist sehr nützlich wenn man eine Matrix diagonalisieren kann; und es ist wichtig zu wissen ob eine Matrix diagonalisierbar ist. ("Spektralsatz" u.Ä.)
- Bsp: Für $D = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix}$ ist $D^m = \begin{pmatrix} \lambda_1^m & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n^m \end{pmatrix}$.
  Und $A^m = (UDU^{-1})^m = UD^mU^{-1}$.
  (Weil: $UDU^{-1}UDU^{-1} = UD^2U^{-1}$ etc.
  Alternatives Argument: D^m ist die E-Darstellung von A^m.)
- A^m lässt sich also leicht ausrechnen.

Diagonalisierung 2

- Wenn A Diagonalisierbar ist, d.h. die E-Darstellung von A ist die Diagonalmatrix $\begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix}$, dann ist $E = (e_1, \ldots, e_n)$ eine Basis aus Eigenvektoren e_i zum Eigenwert λ_i.
- Umgekehrt: Wenn es eine Basis $E = (e_1, \ldots, e_n)$ aus Eigenvektoren von A gibt, dann ist A diagonalisierbar (die E-Darstellung ist diagonal).
- Um eine Matrix zu diagonalisieren "reicht" es also die Eigenwerte und die dazugehörigen Eigenvektoren zu finden. (Der Eigenwerte-Teil ist dabei der problematischere.)

Eine Anwendung: Eine Formel für Fibionacci-Zahlen

Anwendung: Fibionacci

- Zur Erinnerung: Fibonacci Zahlen definiert durch $F_0 := 0$, $F_1 := 1$, $F_{n+2} := F_n + F_{n+1}$.
- Können wir schreiben als:
  $\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} F_{n+1} \\ F_n \end{pmatrix} = \begin{pmatrix} F_{n+2} \\ F_{n+1} \end{pmatrix}$, mit $\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} F_2 \\ F_1 \end{pmatrix}$.
- Setze $A := \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$ und $v_0 := \begin{pmatrix} 1 \\ 0 \end{pmatrix}$.
- D.h. $Av_0 = \begin{pmatrix} F_2 \\ F_1 \end{pmatrix}$, $A(Av_0) = \begin{pmatrix} F_3 \\ F_2 \end{pmatrix}$, und allgemein: $A^n v_0 = \begin{pmatrix} F_{n+1} \\ F_n \end{pmatrix}$.

Anwendung: Fibonacci (2)

- $A := \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$. Wir wollen A^n ausrechnen.
- Eigenwerte: $\chi_A(X) = X(X-1) - 1 = X^2 - X - 1$, hat Lösung $\lambda_1 = \frac{1+\sqrt{5}}{2}$ und $\lambda_2 = \frac{1-\sqrt{5}}{2}$.
- Berechne Eigenvektor zum Eigenwert λ_i (d.h. Kern von $A - \lambda_i I$):
  Verwende: $1 - \lambda_1 = 1 - \frac{1\pm\sqrt{5}}{2} = \frac{1\mp\sqrt{5}}{2} = \lambda_2$.
  Dann entspricht $A - \lambda_1 I = \vec{0}$ der erweiterte Matrix $\left(\begin{array}{cc|c} \lambda_2 & 1 & 0 \\ 1 & \lambda_1 & 0 \end{array}\right)$
  Die erste Zeile ergibt: $x_2 = -\lambda_2 x_1$ (weiters müssen wir nicht testen, weil wir bereits wissen dass λ_2 EW ist).
  Ein Eigenvektor v_1 zu Eigenwert λ_1 ist also $\begin{pmatrix} 1 \\ -\lambda_2 \end{pmatrix}$
- Analog: Ein EV v_2 zu EW λ_2 ist $\begin{pmatrix} 1 \\ -\lambda_1 \end{pmatrix}$

Anwendung: Fibonacci (3)

- Für Basis $E = (v_1, v_2)$ ist die Basiswechselabbildung $U = \begin{pmatrix} 1 & 1 \\ -\lambda_2 & -\lambda_1 \end{pmatrix}$ und E-Darst von A ist $D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$
- $U^{-1} = \begin{pmatrix} 1 + \frac{\lambda_2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ -\frac{\lambda_2}{\sqrt{5}} & -\frac{1}{\sqrt{5}} \end{pmatrix}$ (verwende: $\lambda_2 - \lambda_1 = -\sqrt{5}$
- $U^{-1} v_0 = \begin{pmatrix} 1 + \frac{\lambda_2}{\sqrt{5}} \\ -\frac{\lambda_2}{\sqrt{5}} \end{pmatrix}$
- $\begin{pmatrix} F_{n+1} \\ F_n \end{pmatrix} = A^n v_0 = U D^n U^{-1} v_0 = \begin{pmatrix} 1 & 1 \\ -\lambda_2 & -\lambda_1 \end{pmatrix} \begin{pmatrix} \lambda_1^n & 0 \\ 0 & \lambda_2^n \end{pmatrix} \begin{pmatrix} 1 + \frac{\lambda_2}{\sqrt{5}} \\ -\frac{\lambda_2}{\sqrt{5}} \end{pmatrix} =$
  $\begin{pmatrix} \lambda_1^n & \lambda_2^n \\ -\lambda_1^n \lambda_2 & -\lambda_1 \lambda_2^n \end{pmatrix} \begin{pmatrix} 1 + \frac{\lambda_2}{\sqrt{5}} \\ -\frac{\lambda_2}{\sqrt{5}} \end{pmatrix}$, d.h.
  $F_n = -\frac{\lambda_1^n}{\sqrt{5}} \lambda_2 (\sqrt{5} + \lambda_2) + \frac{\lambda_2^n}{\sqrt{5}} (\lambda_1 \lambda_2)$. Weil $\lambda_1 \lambda_2 = -1$ und $\sqrt{5} + \lambda_2 = \lambda_1$, bekommen wir: $F_n = \frac{1}{\sqrt{5}} (\lambda_1^n - \lambda_2^n)$.

Fibonacci: Zusammenfassung

Wir haben also gezeigt:

Theorem

$$F_n = \frac{1}{\sqrt{5}}\left(\left(\frac{1+\sqrt{5}}{2}\right)^n + \left(\frac{1-\sqrt{5}}{2}\right)^n\right)$$

Bemerkung:

- Wenn wir die Formel bereits gefunden haben, dann ist es recht einfach zu beweisen dass sie die Fibonacci Zahlen ausrehchnet, und zwar mit Induktion:
- Angenommen die Formel stimmt für $n-2$ und $n-1$. Dann ist $F_n := F_{n-1} + F_{n-2} =$
  $\frac{1}{\sqrt{5}}\left(\left(\frac{1+\sqrt{5}}{2}\right)^{n-2} + \left(\frac{1-\sqrt{5}}{2}\right)^{n-2} + \left(\frac{1+\sqrt{5}}{2}\right)^{n-1} + \left(\frac{1-\sqrt{5}}{2}\right)^{n-1}\right) =$
  $\frac{1}{\sqrt{5}}\left(\left(\frac{1+\sqrt{5}}{2}\right)^{n-2}\left(1 + \frac{1+\sqrt{5}}{2}\right) + \left(\frac{1-\sqrt{5}}{2}\right)^{n-2} + \left(1 + \frac{1-\sqrt{5}}{2}\right)\right) =$
  $\frac{1}{\sqrt{5}}\left(\left(\frac{1+\sqrt{5}}{2}\right)^n + \left(\frac{1-\sqrt{5}}{2}\right)^n\right)$, weil $\left(1 + \frac{1\pm\sqrt{5}}{2}\right) = \left(\frac{1\pm\sqrt{5}}{2}\right)^2$. □

Dualraum und transponierte Matrizen

Dualraum

Definition

Der Dualraum V^* von V ist $L(V, K)$, der Vektorraum der linearen Abbildungen von V in den Grundkörper.

Wenn V die Basis $E = (e_1, \ldots, e_n)$ hat, bekommen wir daraus eine "duale Basis" $F = (f_1, \ldots, f_n)$ des Dualraums, durch $f_i(e_j) := \begin{cases} 1 & \text{wenn } i = j, \\ 0 & \text{sonst.} \end{cases}$.

V^* ist also ebenfalls n-dimensional (das funktioniert nur im endlichen), und die (E, F)-Darstellung eines $g \in V$ ist eine $1 \times n$ Matrix (bzw ein "Zeilenvektor") $(a_i)_{1 \leq 1 \leq n}$, mit $(a_1 \quad \cdots \quad a_n) \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = a_1 b_1 + \cdots + a_n b_n$.

Wir bekommen dann (wie immer zwischen zwei n-dimensionalen Räumen) einen Isomorphismus $i : V \to V^*$ definiert durch $e_i \mapsto f_i$.

Transponierte Matrix

- Zu der $m \times n$-Matrix $A = (a_{i,j})_{1 \leq i \leq m, 1 \leq j \leq m}$ können wir die transponierte Matrix $A^\top$ definieren: Das ist eine $n \times m$-Matrix $(b_{i,j})_{1 \leq i \leq n, 1 \leq j \leq m}$ mit $b_{i,j} = a_{j,i}$.
- In dieser Notation gilt: $i(v)(w) = v^\top w$ (wir wenden den Zeilenvektor $v^\top$ durch "normale" Matrixmultiplikation auf den Spaltenvekor w an.)
- Klarerweise gilt $(A^\top)^\top = A$
- Es gilt: $(AB)^\top = B^\top A^\top$.
  Beweis: Der (i,j)-te Eintrag $x_{i,j}$ von AB ist $\sum_{k=1}^n a_{i,k} b_{k,j}$. Der (i,j)-te Eintrag von $B^\top A^\top$ ist $\sum_{k=1}^n b_{k,j} a_{j,k} = \sum_{k=1}^n a_{j,k} b_{k,j} = x_{j,i}$. □
- Daraus folgt für $A \in \mathrm{GL}(n)$: $(A^{-1})^\top = (A^\top)^{-1}$.
  Beweis: $I = I^\top = (AA^{-1})^\top = (A^{-1})^\top A^\top$. □
- Für $n \times n$-Matrizen gilt: $\det(A^\top) = \det(A)$.
  Beweis: Ein Summand von $\det(A)$ hat die Form $\prod_{i=1}^n a_{\pi(i),i}$. Ein Summand von $\det(A^\top)$ hat die Form $\prod_{i=1}^n a_{i,\zeta(i)}$ bzw nach Umordnung $\prod_{i=1}^n a_{\zeta^{-1}(i),i}$. Es macht keinen Unterschied ob wir über $\pi \in S(n)$ oder über π^{-1} für $\pi \in S(n)$ summieren. □

Dualraum: Bemerkungen (kein Prüfungsstoff)

- Es gibt aber i.A. keinen "natürlichen", basisunabhängigen Isomorphismus zwischen V und V^*.
- Es gibt aber (im endlichdimensionalen) einen natürlichn Isomorphismus zwischen V und V^{**}: $v \mapsto (f \mapsto f(v))$.
- Im unendlichdimensionalen gibt sind V und V^* nicht isomorph.
- Es gibt aber "topologische" Varianten des Dualraums bei denen es einen Isomorphismus geben kann.
- Gegeben den Isomorphismus $i : K^n \to (K^n)^*$, könnten wir $v \cdot w$ definieren als $i(v)(w) = v_1 w_1 + \cdots + v_n w_n$.
  Das funktioniert aber nur bei reellen Vektorräumen gut, und ergibt das sogenannte Innenprodukt:

Innenprodukt und Norm

Innenprodukt

Definition

Für $\vec{v}$, $\vec{w}$ in $\mathbb{R}^n$ definiere das Innenprodukt $\vec{v} \cdot \vec{w} := v_1 w_1 + \cdots + v_n w_n$.

Oft auch $\langle v, w \rangle$ geschrieben.
(Achtung: Wir haben $\langle v, w \rangle$ auch für "den von v und w erzeugte Unterraum" verwendet. Aus dem Kontext ist klar was gemeint ist.)

Es gilt:

- $\vec{v} \cdot \vec{w} = \vec{w} \cdot \vec{v}$
- $(\lambda \vec{v}) \cdot \vec{w} = \lambda(\vec{v} \cdot \vec{w})$
- $(\vec{v}_1 + \vec{v}_2) \cdot \vec{w} = \vec{v}_1 \cdot \vec{w} + \vec{v}_2 \cdot \vec{w}$

(So etwas heißt "symmetrische Bilinearform"). Weiters gilt:

- $\vec{v} \cdot \vec{v} \geq 0$, und $\vec{v} \cdot \vec{v} = 0$ gdw $\vec{v} = 0$.

Für die Standardbasis $(e_1, \ldots, e_n)$ gilt:

- $\langle e_i, e_j \rangle = \delta_{ij} := \begin{cases} 1 & \text{wenn } i = j \\ 0 & \text{sonst.} \end{cases}$

So eine Basis nennt man "Orthonormalbasis".

Die Norm

Definition

$\|\vec{v}\| := \sqrt{\vec{v} \cdot \vec{v}}$ ist die euklidische Länge von $\vec{v}$ im $\mathbb{R}^n$.

Es gilt:

- $\|\lambda \vec{v}\| = \lambda \|\vec{v}\|$
- $|\vec{v} \cdot \vec{w}| \leq \|\vec{v}\| \cdot \|\vec{w}\|$ ("Cauchy-Schwarz", ohne Bew.)
- $\|\vec{v} + \vec{w}\| \leq \|\vec{v}\| + \|\vec{w}\|$ ("Dreiecksungl.")

  Beweis: $\|\vec{v} + \vec{w}\|^2 = (\vec{v} + \vec{w}) \cdot (\vec{v} + \vec{w}) = \|\vec{v}\|^2 + 2\vec{v} \cdot \vec{w} + \|\vec{w}\|^2 \overset{\text{C.S.}}{\leq}$
  $\|\vec{v}\|^2 + 2\|\vec{v}\| \cdot \|\vec{w}\| + \|\vec{w}\|^2 = (\|\vec{v}\| + \|\vec{w}\|)^2$ □

Innenprodukt und transponierte Matrix

- Es gilt: $v \cdot w = v^\top w$.
  (Folgt dirkt aus der Def von Transposition und Matrizenmultiplikation.)
- Daher gilt: $\langle Av, w\rangle = (Av)^\top w = v^\top A^\top w = \langle v, A^\top w\rangle$.
- Bsp: $\left\langle \begin{pmatrix} 3 & 7 \\ 4 & 1 \end{pmatrix} \begin{pmatrix} 9 \\ 12 \end{pmatrix}, \begin{pmatrix} 3 \\ 99 \end{pmatrix} \right\rangle = \left\langle \begin{pmatrix} 9 \\ 12 \end{pmatrix}, \begin{pmatrix} 3 & 4 \\ 7 & 1 \end{pmatrix} \begin{pmatrix} 3 \\ 99 \end{pmatrix} \right\rangle$

Theorem

$\langle Av, w\rangle = \langle v, A^\top w\rangle$

Weiters gilt:

Theorem

Wenn $\langle Av, w\rangle = \langle v, w\rangle$ *für alle* v, w, *dann ist* $A = I$.

Beweis: Wenn $A = (a_{i,j})_{1 \le i,j \le n}$, dann ist $\langle Ae_j, e_i\rangle = a_{i,j}$. □

Basisabhängigkeit

- Das Innenprodukt ist nicht "basisunabhängig", in dem folgenden Sinn:
- Bsp: Sei (e_1, e_2) die Standardbasis des $\mathbb{R}^2$.

  D.h. $e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $e_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, und $\langle e_1, e_2 \rangle = 1 \cdot 0 + 0 \cdot 1 = 0$.

  Für $B = (b_1, b_2) = (e_1 + \frac{1}{2} e_2, \frac{1}{2} e_2)$ gilt $e_1 = b_1 - b_2$, $e_2 = 2b_2$, d.h. e_1, e_2 haben die B-Darstellungen $e_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$, $e_1 = \begin{pmatrix} 0 \\ 2 \end{pmatrix}$, und "in dieser Basis gerechnet" ergäbe sich $1 \cdot 0 - 1 \cdot 2 = -2$.
- Wir werden sehen: Das Innenprodukt ist unabhängig von der Wahl einer "Orthonormalbasis".

Orthogonal und orthonormal

- $\vec{v}$ und $\vec{w}$ heißen orthogonal, wenn $\vec{v} \cdot \vec{w} = 0$.
- Dann gilt: $\|\vec{v} + \vec{w}\|^2 = \vec{v} \cdot \vec{v} + 2 \cdot \vec{v} \cdot \vec{w} + \vec{w} \cdot \vec{w} = \|\vec{v}\|^2 + \|\vec{w}\|^2$
  (Satz des Pythagoras)
- $(v_1, \ldots, v_m)$ heißt orthornomal, wenn $\langle v_i, v_j \rangle = \delta_i, j$, bzw äquivalent dazu: Wenn $\|v_i\| = 1$ für alle i und v_i und v_j orthogonal für alle $i \neq j$.
- Wenn $(v_1, \ldots, v_m)$ orthornomal, und $w = \sum_{i=1}^{m} \lambda_i v_i$, dann ist $\langle w, v_i \rangle = \lambda_i$.
- Wenn $(v_1, \ldots, v_m)$ orthornomal ist, dann auch l.u.
  Beweis: Sonst wäre $v_k = \sum_{1 \leq i \leq m, i \neq k} \lambda_i v_i$, und damit $\langle v_k, v_k \rangle = 0$. □
- $E = (e_1, \ldots, e_n)$ ist Orthonormalbasis (ONB), wenn E Basis ist und orthonormal.
- Dann gilt: $v = \sum_{i=1}^{n} \langle v, e_i \rangle e_i$ für alle $v \in V$.
  Beweis: Weil E Basis ist, ist $v = \sum_{i=1}^{n} \lambda_i e_i$. Dann ist $\langle v, e_i \rangle = \sum_{i=1}^{n} \lambda_i \langle e_i, e_i \rangle = \lambda_i$, weil E orthonormal ist. □

Die Orthogonale Gruppe

Orthonormale Abbildungen

Definition

Eine $n \times n$ Matrix A heißt orthogonal, $A \in \mathrm{O}(n)$ wenn die Spalten eine ONB bilden.
(Äquivalent: Wenn O die Basiswechselabbildung zu einer ONB ist.)

Theorem

Die folgenden sind äquivalent:

- $A \in \mathrm{O}(n)$
- $A^{-1} = A^{\top}$
- $\langle Av, Aw \rangle = v\dot{w}$ *für alle* v, w *(in etwa: A erhält Längen und Winkel).*
- *Wenn* $\vec{x}$ *die E-Darstellung von* v *und* $\vec{y}$ *von* w *ist* ($E = (e_1, \ldots, e_n)$ *die Spalten von O), dann ist* $\langle v, w \rangle = \sum_{i=1}^{n} x_i y_i$

Beweis

Die folgenden sind äquivalent:

1. $A \in \mathrm{O}(n)$
2. $A^{-1} = A^\top$
3. $\langle Av, Aw \rangle = v\dot{w}$ für alle v, w (in etwa: A erhält Längen und Winkel).
4. Wenn $\vec{x}$ die E-Darstellung von v und $\vec{y}$ von w ist ($E = (e_1, \ldots, e_n)$ die Spalten von O), dann ist $\langle v, w \rangle = \sum_{i=1}^{n} x_i y_i$

Beweis:

- $1 \to 4$: $\langle v, w \rangle = \langle x_1 e_1 + \cdots + x_n e_n, y_1 e_1 + \cdots + y_n e_n \rangle = \sum_{x_i y_i}$, weil $\langle e_i, e_i \rangle = 1$ und $\langle e_i, e_j \rangle = 0$ für $i \neq j$.
- $4 \to 2$: $v := A\vec{x}$ und $w := A\vec{y}$, d.h. $\langle A\vec{x}, A\vec{y} \rangle = \langle v, w \rangle = \sum v_i w_i$. Nach Annahme 4 ist das gleich $\sum x_i y_i = \langle \vec{x}, \vec{y} \rangle$. D.h. $\langle \vec{x}, \vec{y} \rangle = \langle A\vec{x}, A\vec{y} \rangle = \langle \vec{x}, A^\top A\vec{y}$ für alle $\vec{x}, \vec{y}$, daher ist $A^\top A = I$.
- $2 \to 3$: $\langle Av, Aw \rangle = \langle v, A^\top Aw \rangle = \langle v, Iw \rangle = \langle v, w \rangle$
- $3 \to 1$: Die i-te Spalte von A ist Ae_i, und $\langle Ae_i, Ae_j \rangle = \langle e_i, e_j \rangle$. □

Beispiel

- Für $A := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ gilt $AA^\top = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

  Daher ist A orthonormal, und erhält alle Winkel und Längen.

- Für $A := \begin{pmatrix} \frac{1}{\sqrt{2}} & 1 \\ \frac{1}{\sqrt{2}} & 0 \end{pmatrix}$ gilt $AA^\top = \begin{pmatrix} \frac{1}{\sqrt{2}} & 1 \\ \frac{1}{\sqrt{2}} & 0 \end{pmatrix}\begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 1 & 0 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 3 & 1 \\ 1 & 1 \end{pmatrix}$.

  A ist nicht orthonormal (es erhält zwar die Länge 1 von e_1 und e_2, aber nicht den rechten Winkel).

Einige Eigenschaften

Es gilt:

- $A \in \mathrm{O}(n) \to \det(A) = \pm 1$.
  Beweis:
  $1 = \det(I) = \det(AA^{-1}) = \det(AA^\top) = \det(A)\det(A^\top) = \det(A)^2$. □
- $\mathrm{O}(n)$ ist eine Untergruppe von $\mathrm{GL}(n)$.
  Beweis:
  Jedes $A \in \mathrm{O}(n)$ ist invertierbar, weil $A^\top A = I$. Daher ist $A \in \mathrm{GL}(n)$.
  Angenommen A, B in $\mathrm{O}(n)$. Dann
  $(AB)^\top = B^\top A^\top = B^{-1}A^{-1} = (AB)^{-1}$, und
  $(A^{-1})^\top = (A^\top)^{-1} = (A^{-1})^{-1}$. □

Finden von Orthonormalbasen

Es gilt:

- Wenn $(v_1, \ldots, v_n)$ orthonormal, und $w \notin \langle v_1, \ldots, v_n \rangle$ (der erzeugte Unterraum), dann gibt es w' s.d. $\langle v_1, \ldots, v_n, w \rangle = \langle v_1, \ldots, v_n, w' \rangle$ und $\langle v_1, \ldots, v_n, w' \rangle$ orthonormal.
  Beweis: $w_1 := w - \langle w, v_1 \rangle v_1 - \cdots - \langle w, v_n \rangle v_n$. Dann ist $\langle w, v_i \rangle = 0$ für alle i, und $\langle v_1, \ldots, v_n, w \rangle = \langle v_1, \ldots, v_n, w_1 \rangle$ (insbes: $w_1 \neq \vec{0}$).
  Dann setze $w' := \frac{1}{\|w_1\|} w_1$. □

Damit kann man dann analog zu Erzeugendensystemen zeigen:

- Jeder Teilraum von V hat eine ONB
- Jede orthonormale Menge kann zu einer ONB erweitert werden.

Orthogonale Komplemente und Orthogonalprojektionen

- Zu jedem Unterraum W_1 von V gibt es einen (nicht eindeutigen) Komplementärraum W_2.
- Wenn wir mit einer ONB $(v_1, \ldots, v_m)$ von W_1 starten, und zu einer ONB $(v_1, \ldots, v_n)$ von V erweitern, dann ist $W_2 := \langle v_{m+1}, \ldots, v_n \rangle$ ein Komplementärraum mit der folgenden zusätzlichen Eigenschaft: $w_1 \cdot w_2 = 0$ für $w_1 \in W_1$ und $w_2 \in W_2$.

Definition

W_2 heißt orthogonales Komplement zu W_1, wenn W_2 und W_1 komplementär sind, und $w_1 \cdot w_2 = 0$ für $w_1 \in W_1$ und $w_2 \in W_2$.
Die entsprechende Projektion heißt orthogonale Projektion.

- Achtung: Eine (nicht-triviale) Orthogonalprojektion ist nicht orthogonal. Eine orthogonale Matrix ist ja immer invertierbar!

Orthogonales Komplement: Eindeutigkeit

Theorem

Zu jedem Unterraum W_1 von V gibt es ein eindeutiges orthogonales Komplement W_2.

Beweis: Wähle wie zuvor ONB $(v_1, \ldots, v_n)$ so dass $(v_1, \ldots, v_m)$ ONB von W_1 ist und $(v_{m+1}, \ldots, v_n)$ von W_2.
Sei W_2' ebenfalls orthogonales Komplement von W_1, und $w \in W_2'$. Dann $v \cdot w = 0$ für alle $v \in W_1$; weiters ist $w = \sum_{i=1}^{n} \langle w, v_i \rangle v_i$; insgesamt haben wir also $w = \sum_{i=m+1}^{n} \langle w, v_i \rangle v_i \in W_2$.
Es ist daher $W_2' \subseteq W_2$; und die symmetrische Argumentation (mit W_2 und W_2 vertauscht) gibt $W_2 \subseteq W_2'$; insgesamt also $W_2 = W_2'$. □

Drehungen und Drehspiegelungen

- $A \in \mathrm{O}(n)$ hat entweder Determinante 1 oder -1.
- $\mathrm{SO}(n)$ ist $\mathrm{SL}(n) \cap \mathrm{O}(n)$, d.h. die Gruppe der orthogonalen Matrizen mit Determinante 1.

  (Gruppe, weil der Schnitt zweier Untergruppen (in dem Fall von $\mathrm{GL}(n)$) wieder Untergruppe ist.)
- Die Matrizen in $\mathrm{SO}(n)$ heißen auch Drehungen.
  Diese Bezeichnung ist besonders naheliegend im im $\mathbb{R}^2$ und $\mathbb{R}^3$: Hier kann man beweisen dass diese Matrizen tatsächlich genau die "klassische" Drehungen um den Nullpunkt ($\mathbb{R}^2$) oder eine Drehachse ($\mathbb{R}^3$) um einen Winkel $0 \leq \phi < 2\pi$ sind.
- Die orthogonalen Matrizen mit Determinante -1 heißen (vor allem in $\mathbb{R}^3$) auch "Drehspiegelungen".
  Sie bilden keine Gruppe (zB weil sie nicht das neutrale Element I enthalten).

Innenprodukt: Andere Körper

- Man könnte zwar $\vec{v} \cdot \vec{w} := v_1 w_1 + \cdots + v_n w_n$ für Vektoren in K^n für beliebige Körper K definieren. Das führt aber i.A. nicht zu einem nützlichen Begriff.
- Z.B. ist in $\mathbb{Z}_2^2$ $\begin{pmatrix} 1 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 1 + 1 = 0$.
- Oder: In $\mathbb{C}^2$ wäre $\begin{pmatrix} i \\ 1 \end{pmatrix} \cdot \begin{pmatrix} i \\ 1 \end{pmatrix} = -1 + 1 = 0$
- In Vektorräumen über $\mathbb{C}$ verwendet man stattdessen die folgende wichtige Form:

Innenprodukt in $\mathbb{C}$

- In $\mathbb{C}$-VR: $\vec{v} \cdot \vec{w} := v_1 w_1^* + \cdots + v_n w_n^*$, wobei $(a + ib)^* = a - ib$ (komplex konjugiert).
- Dann gilt: $\vec{v} \cdot \vec{v}$ ist der "richtige" kartesiche Abstand in der komplexen Ebene hoch n; insbesondere immer ≥ 0 und 0 gdw $\vec{v} = 0$.
- Jetzt ist aber $\vec{v} \cdot \vec{w} = (\vec{w} \cdot \vec{v})^*$.
- Statt $A^\top$ ("transponiert") verwenden wir nun A^* "transponiert und komplex konjuguert"; Statt othogonaler Matrizen (d.h. $A^\top A = I$) betrachten wir unitäre Matrizen, definiert durch $A^* A = A A^* = I$ (das sind genau die die das komplexe Innenprodukt erhalten)
- Faustrgel: Wenn Sie über $\mathbb{C}$ rechnen und "orthogonal" oder "transponiert" verwenden, machen Sie etws falsch.

Spur

Spur

Ab hier ist der Grundkörper K wieder beliebig.

Definition

$\mathrm{Tr}(A)$, die Spur der $n \times n$ Matrix $A = (a_{i,j})_{1 \le i,j \le n}$ ist die Summe der Diagonalelemente: $\mathrm{Tr}(A) = \sum_{i=1}^{n} a_{i,i}$.

Theorem

$\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$

Beweis: $\mathrm{Tr}(AB) = \sum_{i=1}^{n} (\sum_{k=1}^{n} a_{i,k} b_{k,i}) = \sum_{k=1}^{n} (\sum_{i=1}^{n} a_{i,k} b_{k,i}) = \sum_{k=1}^{n} (\sum_{i=1}^{n} b_{k,i} a_{i,k}) = \mathrm{Tr}(BA)$. □

Theorem

$\mathrm{Tr}(A)$ *ist unabhängig von der Basis in der A dargestellt wird, d.h.* $\mathrm{Tr}(U^{-1}AU) = \mathrm{Tr}(A)$.

Beweis: $\mathrm{Tr}(U^{-1}AU) = \mathrm{Tr}(U^{-1}(AU)) = \mathrm{Tr}((AU)U^{-1}) = \mathrm{Tr}(A)$. □

EV zu verschiedenen EW sind l.u.

Theorem

Wenn $(v_1, \ldots, v_k)$ EV zu verschiedenen EW $\lambda_1, \ldots, \lambda_k$ der $n \times n$ Matrix A sind, dann sind $(v_1, \ldots, v_k)$ l.u.

Beweis:

- Sei $1 \leq j \leq k$ maximal mit $(v_1, \ldots, v_j)$ l.u. Wir wollen zeigen $j = k$. Nehmen wir an $j < k$ (und leiten daraus einen Widerspruch ab).
- $(v_1, \ldots, v_j)$ ist also l.u., und $(v_1, \ldots, v_j, v_{j+1})$ ist l.a.
  Daraus folgt: $(*)$ $v_{j+1} = \sum_{i=1}^{j} \mu_i v_i$ (für irgendwelche μ_i aus K).
- A auf $(*)$ angewendet gibt:
  $(*_2)$ $\lambda_{j+1} v_{j+1} = A v_{j+1} = \sum_{i=1}^{j} \mu_i A v_i = \sum_{i=1}^{j} \mu_i \lambda_i v_i$.
- $(*)$ mit λ_{j+1} multipliziert gibt:
  $(*_3)$ $\lambda_{j+1} v_{j+1} = \sum_{i=1}^{j} \mu_i \lambda_{j+1} v_i$.
- Subtraktion von $(*_2)$ und $(*_3)$ ergibt:
  $0 = \sum_{i=1}^{j} \mu_i (\lambda_{j+1} - \lambda_i) v_i$, wobei nach Voraussetzung (alle λ_i sind verschieden) $\lambda_{j+1} - \lambda_i \neq 0$, ein Widerspruch zu $(v_1, \ldots, v_j)$ l.u.

Eigenraum

- Zur Erinnerung: Eine $n \times n$ Matrix A ist diagonalisierbar, wenn es eine Basis gibt in der A eine Darstellung als Diagonalmatrix hat; d.h.: wenn es ein $U \in \mathrm{GL}(n)$ gibt mit $U^{-1}AU$ diagonal.
  Das ist genau dann der Fall wenn es eine Basis von Eigenwerten von A gibt.
- Zur Erinnerung: Die EV zu λ bilden einen Unterraum von V, genannt Eigenraum.

Definition

Die Dimension des Eigenraums zu λ heißt "geometrische Vielfachheit" des EW λ.

- Wenn es eine Basis E aus Eigenvektoren von A gibt, dann ist die geometrische Vielfachheit von λ die Anzahl der Elemente von E die λ-EV sind.
  In anderen Worten: Wenn A Diagonal ist, dann ist die geometrische Vielfachheit von λ die Anzahl der Diagonal-Einträge von A die gleich λ sind.

Eigenwerte und Spur

Bezeichne die geometrischen Vielfachheit von λ mit $\mathrm{gv}(\lambda)$.

Theorem

Wenn A diagonalisierbar ist, dann ist $\mathrm{Tr}(A)$ gleich der Summe aller EW, wobei der EW λ $\mathrm{gv}(\lambda)$ oft in die Summe eingeht.

Beweis: Stelle A in einer Basis aus EW dar:

$$A = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix}$$

Dann ist $\mathrm{Tr}(A) = \lambda_1 + \cdots + \lambda_n$,
und $\mathrm{gV}(\lambda)$ ist die Anzahl der i mit $\lambda_i = \lambda$. □

Beispiel: Projektion

Sei A eine Projektion ($A^2 = A$), d.h. A ist Projektion bzgl komplentärer $W_1 = \mathrm{img}(A)$, $W_2 = \ker(A)$. Wenn $(e_1, \ldots, e_m)$ Basis von W_1 ist, und $(e_{m+1}, \ldots, e_n)$ von W_2, dann ist die $(e_1, \ldots, e_n)$-Darstellung von A gleich

$$\begin{pmatrix} 1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & \cdots & 1 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix}$$

Daher ist $\mathrm{Tr}(A) = \dim(W_1) = \mathrm{rk}(A)$.

(Insbesondere ist für eine nichttriviale Projektion $\mathrm{Tr}(A)$ eine ganze Zahl zwischen 1 und $n-1$.)

Bsp:

$$A = \begin{pmatrix} -1 & -3 & -2 & -1 \\ 3 & 6 & 3 & 2 \\ -2 & -3 & -1 & -1 \\ -3 & -6 & -3 & -2 \end{pmatrix}$$

Wenn wir bereits wissen dass A eine Projektion ist, dann folgt daraus $\mathrm{rk}(A) = \mathrm{Tr}(A) = -1 + 6 - 1 - 2 = 2$

Diagonalisierbarkeit, Spektralsätze

Diagonalisierbarkeit

- Eine $n \times n$ Matrix A ist diagonalisierbar, wenn $(\exists U \in \mathrm{GL}(n))\, U^{-1}AU = D$ mit D diagonal (d.h. $d_{i,j} = 0$ für $i \neq j$).
  In andren Worten: A ist ähnlich zu einer Diagonalmatrix.
- Äquivalent: Es gibt eine Basis aus A-Eigenvektoren.
- Wir kennen schon eine Klassen von Matrizen die immer diagonalisierbar sind, nämlich die Projektionen.

Diagonalisierbarkeit und charakteristisches Polynom

- Zur Erinnerung: λ ist EW von A gdw $\chi_A(\lambda) = 0$ für das charakteristisches Polynom $\chi_A(X)$ von A.
- Es gilt: Wenn $\chi_A(X)$ n viele verschiedene Nullstellen hat, dann ist A diagonalisierbar.
  Beweis: Jede Nullstelle λ_i von $\chi_A(X)$ ist Eigenwert, d.h. es gibt n Eigenvektoren $(v_1, \ldots, v_n)$ zu den verschiedenen EW $\lambda_1, \ldots, \lambda_n$; daher sind $(v_1, \ldots, v_n)$ l.u. und bilden eine Basis. □
- Wenn $\chi_A(X)$ keine Nullstellen hat, ist A natürlich nicht diagonalisierbar.
  (Es gibt dann gar keine EV von A, und insbesondere keine Basis von EV.)

Diagonalisierbarkeit und Grundkörper

Diagonalisierbarkeit hängt, wie auch die Nullstellen von $\chi_A(X)$, vom Grundkörper ab.
Beispiele:

- $A_1 := \begin{pmatrix} 0 & 1 \\ 2 & 0 \end{pmatrix}$ ist reell diagonalisierbar, aber nicht über $\mathbb{Q}$:
  $\chi_{A_1}(X) = X^2 - 2$ hat die zwei versch. NS $\pm\sqrt{2}$ in $\mathbb{R}$, aber keine in $\mathbb{Q}$.
- $A_2 := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ ist komplex diag.bar, aber nicht über $\mathbb{R}$ oder $\mathbb{Q}$.
  $\chi_{A_2}(X) = X^2 + 1$ hat die zwei versch. NS $\pm i$ in $\mathbb{C}$, aber keine in $\mathbb{R}$.
- $I := \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ ist diagonal(isierbar).
  Das sieht man aber nicht an $\chi_{A_3}(X) = (X-1)^2$, weil:
- $A_4 := \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ ist in überhaupt keinem Körper diagonalisierbar.
  Es gilt $\chi_{A_4}(X) = \chi_{A_3}(X) = (X-1)^2$.

Diagonalisierbarkeit und Dreiecksmatrizen

- Behauptung $A_4 := \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ ist nicht diagonalisierbar.
- Allgemein gilt: Wenn A eine Dreiecksmatrix ist mit $a_{i,i} = \lambda$ für alle $1 \leq i \leq n$, dann ist entweder $A = \lambda I$ (und damit ist A bereits diagonal), oder A ist nicht diagonalisierbar. (Für beliebiges K.)

  Beweis: $\chi_A(X)$ ist die Determinante der Dreiecksmatrix $XI - A$, und daher das Produkt der Diagonalelemente, $(X - \lambda)^n$. Daher ist λ der einzige EW. Wenn es eine Basis $E = (e_1, \ldots, e_n)$ aus EV gibt, folgt daraus dass jedes $v \in K^n$ EV (zum EW λ) ist, dh. $Av = \lambda v$.
  (Der λ-Eigenraum ist ja ein Unterraum; expliziter: Jedes v hat Darstellung $v = \sum \mu_i e_i$, dann ist $Av = \sum \mu_i Ae_i = \sum \mu_i \lambda e_i = \lambda v$.)
  Dann ist aber $A = \lambda I$ und damit bereits diagonal. □
- Wenn A eine Dreiecksmatrix ist mit lauter verschiedenen Einträgen in der Diagonale, dann ist A diagonalisierbar.
  In dem Fall ist ja $\chi_A(X) = \prod_{i=1}^{n}(X - \lambda_i)$, d.h. es gibt n verschiedene EW.

Diagonalisierbarkeit und $\mathbb{Z}_p$

Beispiele in $\mathbb{Z}_7$:

Listen wir erst die Elemente von $\mathbb{Z}_7$ und ihre Quadrate:

x	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{4}$	$\overline{5}$	$\overline{6}$
x^2	$\overline{0}$	$\overline{1}$	$\overline{4}$	$\overline{2}$	$\overline{2}$	$\overline{4}$	$\overline{1}$

- $B_1 := \begin{pmatrix} 0 & 1 \\ \overline{2} & 0 \end{pmatrix}$ ist diagonalisierbar über $\mathbb{Z}_7$:
  $\chi_{B_1}(X) = X^2 - \overline{2}$ hat die zwei versch. NS $\overline{3}$ und $-\overline{3} = \overline{4}$.
- $B_2 := \begin{pmatrix} 0 & 1 \\ \overline{3} & 0 \end{pmatrix}$ ist nicht diagonalisierbar über $\mathbb{Z}_7$:
  $\chi_{B_2}(X) = X^2 - \overline{3}$ hat keine NS.
- $A_4 := \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ ist nicht über $\mathbb{Z}_7$ diagonalisierbar (wir wissen ja: über überhaupt keinem Körper).
- etc

Ein reeller Spektralsatz

Wir betrachten nun reelle $n \times n$ Matrizen.

- Definition: A heißt alertorthogonal diagonalisierbar, wenn $A = U^{-1}DU$ für ein $U \in \mathrm{O}(n)$.
- Äquivalen: A hat ONB aus EV.
- Definition: A heißt (im reellen) normal, wenn $A^\top A = AA^\top$.

Theorem (Spektralsatz über $\mathbb{R}$, ohne Beweis)

A ist genau dann orthogonal diagonalisierbar, wenn A normal ist und alle (komplexen) Nullstellen von $\chi_A(X)$ bereits in $\mathbb{R}$ liegen.

Wenn A symmetrische ist ($A^\top = A$) dann ist A offenbar normal, weiters hat A dann nur reelle Eigenwerte (ohne Beweis). Daher folgt:

Theorem

Symmetrische reelle Matrizen sind orthogonal diagonalisierbar.

Beispiele

- $A_4 := \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ hat zwar reelle EW, aber ist nicht normal, ist also nicht orthogonal diagonalisierbar. (Wir wissen aus anderen Gründen dass A_4 überhaupt nicht diagonalisierbar ist.)
- $A_6 := \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}$ hat die reelle EW 1 und 2. Wir wissen also dass A_6 diagonalisierbar ist. A_6 ist aber nicht normal, und daher nicht orthogonal diagonalisierbar (die EV zu den beiden EW stehen nicht normal aufeinander).
- $A_2 := \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$ ist symmetrisch, daher orthogonal diagonalisierbar.

Ein komplexer Spektralsatz (kein Prüfungsstoff)

Im Komplexen funktioniert alles besser.

- A^*, die Adjungierte von A, ist A transponiert und komplex konjugiert.
- A heißt normal, wenn $A^*A = AA^*$.
- U ist unitär, wenn $U^*U = UU^* = I$. (Entspricht orthogonal für komplexes Innenprodukt; ist äquivalent zu orthogonal bei reellen Einträgen.)
- Spektralsatz (über $\mathbb{C}$): A ist genau dann "unitär diagonalisierbar", d.h. $A = U^{-1}DU$ für D diagonal und U unitär, wenn A normal ist.
- Unitäre Matrizen sind normal.
  (Wenn $A^* = A^{-1}$, dann $A^*A = AA^* = I$.)
- A ist hermitesch wenn $A^* = A$. (Das ist analogon zu symmetrisch).
  Hermitesche Matrizen sind normal.
  (Wenn $A^* = A$, dann $A^*A = A^2 = AA^*$.)

Induktion und Rekursion

Wohlordnungsprinzip

- Induktion ist eine wichtige Beweismethode.
- Zugrundeliegend ist immer eine "Wohlordnung": Hier für $\mathbb{N}$

Theorem

Jedes nichtleere $A \subseteq \mathbb{N}$ hat ein kleinstes Element.

Eine sehr spezielle Eigenschaft von $\mathbb{N}$, gilt nicht zB für $\mathbb{Z}$, oder die rationalen zahlen ≥ 0, etc.

Induktion

- Verwendung im Beweis: Für jeden Bruch $q = \frac{n}{m} > 0$ gibt es eine "gekürzte" Darstellung.
  Beweis: Sei A die Menge aller $k > 0$ so dass für ein (eindeutiges) ℓ gilt: $q = \frac{\ell}{k}$. A ist nichtleer, hat also ein kleinstes Element: $a = \frac{\ell}{k}$ mit k mininmal. Dann kann ℓ und k nicht weiter gekürzt werden. □
- Alternative Intuition: Beginne mit $\frac{n}{m}$. Wenn kürzbar, dann kürze. Wiederhole so lange wie möglich, muss nach endlich vielen Schritten abbrechen, weil sich sonst eine unendliche absteigende Folge von Zählern ergäbe. Und es gilt:

Theorem

Es gibt keine unendliche (strikt) absteigende Folge in $\mathbb{N}$.

Beweis: Sei $(a_1, a_2, \dots)$ so eine Folge. Dann hätte $A = \{a_1, a_2, \dots\}$ kein kleinstes Element: Jedes Element von A kommt als ein a_ℓ vor, dann ist $a_{\ell+1}$ ein kleineres Element von A. □

Wohlordnungsbeweise

Im allgemeinen funktioniert ein Wohlordnungs-Beweis so:

- Wir wollen zeigen dass $\varphi(n)$ für alle $n \in \mathbb{N}$ gilt.
- Angenommen das ist nicht der Fall, dann gibt es also mindestens ein Gegenbeispiel. Sei A die Menge aller Gegenbeispiele, und n_0 das kleinste Element von A.
- Jetzt verwenden wir die Tatsache dass $\varphi(n)$ gilt für alle $n < n_0$, und beweisen damit [hier braucht man natürlich jeweils andere Argumente] dass $\varphi(n_0)$ gilt.
- Wie haben also einen Widerspruch abgeleitet. Die Annahme dass es ein Gegenbeispiel gibt muss also falsch sein, und wir haben damit $\varphi(n)$ für alle $n \in \mathbb{N}$ bewiesen.

Beipiel

Es gilt: $1 + 2 + \cdots + n = \frac{n(n+1)}{2}$

- Für $n = 0$ interpretieren wir das als $0 = 0$.
- Beweis: Angenommen es gibt ein n bei dem das fehlschlägt. Sei n_0 das kleinste Gegenbeispiel.
- Für $n = 0$ ist der Satz wahr; daher kann n_0 nicht 0 sein.
- Daher ist $k := n_0 - 1 \geq 0$ in $\mathbb{N}$ kein Gegenbeispiel, d.h. für k gilt der Satz.
  Dann ist
  $1 + 2 + \cdots + k + n_0 = \frac{k(k+1)}{2} + n_0 = \frac{(n_0-1)(n_0)}{2} + n_0 = \frac{(n_0+1)(n_0)}{2}$.
  (Beachte: gilt auch für den Fall $k = 0$.)
- Der Satz hat also eh auch für n_0 gegolten, Widerspruch.

Noch ein Beispiel

Es gilt: Jedes $n > 1$ läßt sich als Produkt von Primzahlen schreiben.

- Beweis: Angenommen es gibt ein n bei dem das fehlschlägt.
  Sei n_0 das kleinste Gegenbeispiel.
- Eine Zahl ≤ 1 kann schon mal kein Gegenbeispiel sein, d.h. $n_o \geq 2$.
- n_0 kann selbst keine Primzahl sein. (Sonst könnte man n_0 ja als Produkt von Primzahlen schreiben: Als das Produkt mit dem einzigen Faktor n_0.)
- Daher lässt sich n_0 als $n_0 = \ell \cdot k$ schreiben mit $1 < \ell, k < n_0$.
  Weil n_0 das kleinste Gegenbeispiel war, sind $\ell = p_1 \cdot p_2 \cdots p_i$ und $k = p'_1 \cdot p'_2 \cdots p'_j$ Produkte von Primzahlen, und damit auch $n_0 = \ell \cdot k = p_1 \cdot p_2 \cdots p_i \cdot p'_1 \cdot p'_2 \cdots p'_j$, Widerspruch.

Nur in Wohlordnungen

Wohlordnungsbeweis funktioniert nur bei Wohlordnungen wie z.B. $\mathbb{N}$.

- Bsp: Das folgende ist kein gültiger Beweis der (falschen) Aussage "Alle rationalen Zahlen sind ≤ 0"
- (Falscher) "Beweis:" Angenommen es gibt ein Gegenbeispiel, d.h. ein $q > 0$. Sei A die Menge der Gegenbeispiele, und q_0 minimal in A. Dann ist $q_0 > 0$, und damit $\frac{q_0}{2} > 0$ ein kleiners Gegenbeispiel, Widerspruch.
- Was schlägt hier fehl? $\mathbb{Q}$ ist nicht wohlgeordnet, also folge aus "die Menge A der Gegenbeispiele ist nichtleer" nicht "es gibt ein kleinstes Gegenbeispiel".
  (Vielmehr gibt es Gegenbeispiele, alle $q > 0$, aber kein kleinstes.)

Wiederholung: Wohlordnung

Wir haben gesehen:

- Jede nichtleere Teilmenge von $\mathbb{N}$ hat ein kleinstes Element.

Das liefert eine wichtige Methode für Beweise:

Wir wollen beweisen $(\forall n \in \mathbb{N})\, \varphi(n)$ für eine bestimmte Aussage $\varphi(x)$.

- (Widerspruchsbeweis:)
  Wir nehmen an dass es ein "Gegenbeispiel" gibt, d.h. ein $n \in \mathbb{N}$ mit $\neg\varphi(n)$.
- Die Menge der Gegenbeispiele, $A = \{n \in \mathbb{N} : \neg\varphi(n)\}$, ist also nichtleer.
- Daher hat A ein kleinstes Element, nennen wir es n_0.
- Jetzt kommt der "individuelle" Beweisteil: Leite einen Widerspruch daraus ab dass n_0 das kleinste Gegenbeispiel ist.
- Die Annahme $A \neq \emptyset$ führt also zu einem Widerspruch, daher muss $A = \emptyset$ gelten, d.h. $\varphi(n)$ für alle $n \in \mathbb{N}$ gelten.

Wiederholung: Beispiel

Es gilt: $1 + 2 + \cdots + n = \frac{n(n+1)}{2}$ für alle natürlichen Zahlen n.

- Wir behaupten also: $(\forall n \in \mathbb{N})\, \varphi(n)$, wobei $\varphi(n)$ die Aussage $1 + 2 + \cdots + n = \frac{n(n+1)}{2}$ ist. ("Linke Seite=Rechte Seite")
- Für $n = 0$: Linke Seite $1 + 2 + \cdots + n$ interpretieren wir als 0.
- Für $n = 1$: Linke Seite $1 + 2 + \cdots + n$ interpretieren wir als 1.
- Für alle $n \geq 0$ gilt:
  (Linke Seite für $n + 1$) ist gleich ((linke Seite für n) $+(n + 1)$).
- Beweis der Behauptung: Angenommen es gibt ein n mit $\neg\varphi(n)$.
  Sei n_0 das kleinste solche Gegenbeispiel.
- Für $n = 0$ ist der Satz wahr; daher kann n_0 nicht 0 sein.
- Daher ist $k := n_0 - 1 \geq 0$ in $\mathbb{N}$, und $\varphi(k)$.
  Dann ist
  $1 + 2 + \cdots + k + (k + 1) = \frac{k(k+1)}{2} + (k + 1) = \frac{(n_0-1)(n_0)}{2} + n_0 = \frac{(n_0+1)(n_0+2)}{2}$
- $\varphi(n_0)$ gilt also doch, Widerspruch.

Gewöhnliche Induktion

Oft werden Induktionsbeweise in folgender spezieller Form (gewöhnliche Induktion) geführt:

Theorem (Prinzip der gewöhnlichen Induktion)

Wir können $(\forall n \in \mathbb{N})\varphi(n)$ *beweisen, indem wir zeigen:*

- *"Induktionsanfang":* $\varphi(0)$
- *"Induktionsschritt" Für beliebiges* $n \in \mathbb{N}$ *gilt:*
  Aus $\varphi(n)$ *folgt* $\varphi(n+1)$.
  (Das $\varphi(n)$ *nennt man dabei manchmal "Induktionsannahme".)*

Das ist ein gültiges Beweisprinzip: Angenommen es gäbe trotzdem ein Gegenbeispiel $k \in \mathbb{N}$, d.h. $\neg\varphi(k)$ Sei n_0 das kleinste Gegenbeispiel.
$n_0 \neq 0$ wegen der Induktionsannahme.
Daher ist $k := n_0 - 1$ in $\mathbb{N}$, und $\varphi(k)$ gilt (n_0 war das kleinste Gegenbeispiel). Es gilt also $\varphi(k)$ und $\neg\varphi(k+1)$, das widerspricht dem Induktionsschritt: $\varphi(k) \to \varphi(k+1)$.

Gewöhnliche Induktion: Bsp

Die gewöhnliche Induktion ist also eine spezielle Form des Induktionsprinzips/Wohlordnungsbeweises.

Für viele Beweise reicht sie aus, z.B. für das vorige Beispiel $1 + \cdots + n = \frac{n(n+1)}{2}$:

- Induktionsbeginn: $\varphi(0)$ ist $0 = 0$ und gilt daher.
- Inudktionsschritt: Wenn $\varphi(n)$ gilt (Induktionsannahme), dann ist (Linke Seite für $n+1$) gleich (Linke Seite für n plus $n+1$) und daher $\frac{n(n+1)}{2} + (n+1)$, was $\frac{(n+1)(n+2)}{2}$ ist. Es gilt also $\varphi(n+1)$

Vollständige Induktion

Für andere Induktionsbeweise reicht die aber die gewöhnliche Induktion nicht aus. (zB für unseren Beweis "alle natürlichen Zahlen sind Produkt von Primzahlen". Hier können wir "vollständige Induktion" verwenden:

Theorem (Prinzip der vollständigen Induktion)

Wir können $(\forall n \in \mathbb{N})\varphi(n)$ *beweisen, indem wir zeigen:*

- *"Induktionsanfang":* $\varphi(0)$
- *"Induktionsschritt" Für beliebiges* $n \in \mathbb{N}$ *gilt:*
  Aus "$\varphi(n)$ *gilt für alle* $0 \leq k \leq n$*" (Induktionsannahme) folgt* $\varphi(n+1)$

(Hier ist die Induktionsvoraussetzung stärker, daher ist die Implikation im Induktionsschritt im Prinzip einfacher zu beweisen.)
Auch dieses Beweisprinzip ist gültig; das sieht man genauso wie bei der gewöhnlichen Induktion.

Beginn bei anderen n als 0

Genauso kann man "für alle ganzen Zahlen größer-gleich m_0 gilt $\varphi(n)$" mit Induktion beweisen. Dazu reicht:

- (Induktionsanfang) zeige $\varphi(m_0)$
- (Induktionsschritt): Für $n \geq m_0$ gilt: Aus $\varphi(n)$ folgt $\varphi(n+1)$.
  (Bzw, in der Variante "vollsändigen Induktion": Aus "$\varphi(k)$ für alle $m_0 \leq k \leq n$" folgt $\varphi(n+1)$

Bemerkung: Wenn man wollte könnte man das formal auf folgendes zurückführen:
Definiere $\psi(n)$ als $\varphi(n+m_0)$.
Dann ist $(\forall n \geq m_0)\varphi(n)$ äquivalent zu $(\forall n \in \mathbb{N})\psi(n)$, und die Annahmen zeigen dass letzteres mit Induktion "von 0 an" beweisbar ist.

Ein Beispiel für vollständige Induktion

Es gilt: Jede natürliche Zahl größer als 1 ist Produkt von Primzahlen.

- Sei $\varphi(n)$ die Aussage "n ist Produkt von Primzahlen."
- Wir beweisen mit Induktion $(\forall n \geq 2)\varphi(n)$.
- Induktionsanfang: $\varphi(2)$ gilt, weil $2 = 2$ das "Produkt" (mit einem einzigen Faktor) von Primzahlen ist.
- Insuktiosnschritt: Angenommen $\varphi(2)$, $\varphi(3)$, ..., $\varphi(n)$ gilt bereits. Wir müssen zeigen: $\varphi(n+1)$. Fallunterscheidung:
  - $n+1$ ist Primzahl. Dann gilt $\varphi(n+1)$ (genau wie im Fall $n = 2$).
  - Ansonsten ist $n+1 = k \cdot \ell$ für irgendwelche $2 \leq k, \ell \leq n$. Es gilt also $\varphi(k)$ und $\varphi(\ell)$, d.h. $k = p_1 \cdot p_2 \cdots p_i$ und $\ell = p'_1 \cdot p'_2 \cdots p'_j$ für Primzahlen p_m, p'_h. Daher ist $n+1 = k \cdot \ell = p_1 \ldots p_i \cdot p'_1 \cdots p'_j$ ebenfalls Produkt von Primzahlen, d.h. $\varphi(n+1)$ gilt.

Fallstricke 1

Induktion funktioniert nicht für nicht-wohlgeordnete Strukturen:

- Bsp $\mathbb{Z}$:
  "Alle ganzen Zahlen sind ≥ 0" ist falsch, obwohl
  $0 \geq 0$ (Induktionsanfang) wahr ist, und $n \geq 0 \rightarrow n+1 \geq 0$
  (Induktionsschritt) gilt.
- Bsp $\{r \in \mathbb{R} : \ r \geq 0\}$
  "Jede nicht-negative reelle Zahl ist eine ganze Zahl" ist falsch, obwohl
  "0 ist ganz" (Induktionsanfang) wahr ist, und
  "n ganz impliziert $n+1$ ganz" (Induktionsschritt) gilt.

Offensichtlich darf man nicht den Induktionsanfang vergessen:

- "Alle natürlichen Zahlen sind irrational" (d.h. $\notin \mathbb{Q}$) ist falsch, obwohl
  $n \notin \mathbb{Q} \rightarrow n+1 \notin \mathbb{Q}$ (Induktionsschritt) stimmt.
  (Es fehlt eben der Induktionsafang $0 \notin \mathbb{Q}$, der offenbar nicht stimmt.)

Fallstricke 2

Beim Beweis von $\varphi(n) \to \varphi(n+1)$ muss man aufpassen, dass man keinen "Sonderfall" übersieht. Manchmal hat zB $n = 0$ ein Sonder-Rolle. Beipiel:

- Sei $\varphi(n)$ die (falsche!) Aussage:
  $1 + 2 + \cdots + n = 1 + \frac{n(n+1)}{2}$,
  wobei wir die linke Seite für $n = 0$ $(1 + \cdots + 0?)$ als 1 interpretieren.
- Offenbar ist $\varphi(0)$ gültig (Induktionsanfang): $1 = 1$
- (Induktionsschritt, falsch) Wenn $\varphi(n)$ gilt dann ist
  $1 + 2 + \cdots + n + (n+1) = 1 + \frac{n(n+1)}{2} = 1 + \frac{(n+1)(n+2)}{2}$, d.h.
  $\varphi(n)$ impliziert $\varphi(n+1)$.
- Das Argument ist aber falsch, weil die Implikation $\varphi(0) \to \varphi(1)$ eben nicht gilt
  (die Implikation $\varphi(n) \to \varphi(n+1)$ für $n \geq 1$ wäre dagegen OK):
  $\varphi(0)$ ist wahr, $\varphi(1)$ ist die (falsche) Aussage $1 = 1 + \frac{1 \cdot 2}{2} = 2$.
  Hier gilt eben für $n = 0$ nicht
  (Linke Seite für $n+1$) ist gleich (Linke Seite für n)+$(n+1)$.

Fallstricke 3: Pferde

Ein Beispiel in dem $\varphi(1) \to \varphi(2)$ eine "Sonderrolle" hat:

- Behauptung (falsch): Alle Pferde haben dieselbe Farbe.
- Genauer formuliert: Sei $\varphi(n)$ die (falsche!) Aussage:
  Wenn A eine (endliche) Menge von Pferden der Größe n ist, und $P, Q \in A$, dann haben P und Q dieselbe Farbe.
- Beweis, Induktionsanfang: Das stimmt offenbar für $n = 0$:
  Da gibt es gar keine $P, Q \in A$, daher ist die Implikation "wenn $P, Q \in A$, dann ..." trivialerweise erfüllt.
  (Und es gilt übrigens auch für $n = 1$: Für $n = 1$ folgt aus $P, Q \in A$ dass $P = Q$, und daher hat P und Q dieselbe Farbe.)
- Beweis, Induktionsschritt (Falsch!):
  Angenommen $\varphi(n)$ gilt, und $A = \{P_1, P_2, \cdots, P_n, P_{n+1}\}$ ist eine Menge von Pferden. Dann haben laut Induktionsannahme die n-vielen Pferde $P_1, P_2, \ldots, P_n$ dieselbe Farbe, und auch die n vielen $P_2, P_3, \ldots, P_n, P_{n+1}$. D.h. P_{n+1} hat dieselbe Farbe wie P_n, und damit haben $P_1, \ldots, P_n, P_{n+1}$ dieselbe Farbe!

Fallstricke 3: Pferde (Forts.)

- (Wiederh.) Wir haben argumentiert: ... $A = \{P_1, P_2, \cdots, P_n, P_{n+1}\}$, dann haben $\{P_1, P_2, \ldots, P_n\}$ und $\{P_2, P_3, \ldots, P_n, P_{n+1}\}$ das gemeinsame Element (Pferd) P_n.
- Was stimmt hier nicht?
  Der Induktionsschritt schlägt fehl bei $\varphi(1) \to \varphi(2)$:
  Wenn $A = \{P_1, P_2\}$, dann haben alle Pferde in $\{P_1\}$ dieselbe Farbe, und dasselbe für $\{P_2\}$; aber diese Mengen haben keine Überschneidung.

Moral: In der Mathematik ist es üblich dass Schreibweisen wie zB $a_1 + a_2 + \cdots + a_n$ oder auch $\{a_0, a_1, \ldots, a_n\}$ auch die Fälle $n = 1$ abdecken (und oft sogar $n = 0$, in dem Fall mit Resultat 0 bzw $\emptyset$); aber man muss aufpassen aus der Schreibweise nicht mehr zu folgern als erlaubt ist (zB dass es a_1 oder a_{n-1} gibt, etc).

Bem.: $a_1 \cdot a_2 \cdots a_n$ für $n = 0$ würde man als 1 interpretieren, u.A. damit $a_1 \cdots a_n = (a_1 \cdots a_m) \cdot (a_{m+1} \cdots a_n)$ dann auch für $m = 0$ funktioniert.
Insbesondere: $x^0 = 1$

Invarianten

In der Informatik beweist man oft mit Induktion über "die Zeit" (genauer: Schritte einer Berechnung, d.h. diskret und mit Anfang), dass bestimmte Dinge nicht passieren können (oder Eigenschaften erhalten bleiben).

Sehr einfaches Beispiel: Roboter auf (beliebig grossem) "Schachbrett"

- Felder sind mit kartesischen Koordinaten bezeichnet: $(0,0)$, $(1,0)$, $(4,-7)$ etc.
- Roboter startet zum Zeitpunkt $t = 0$ auf $(0,0)$
- Die Position zum Zeitpunk t bezeichnen wir mit $P(t)$. Es gilt also: $P(0) = (0,0)$.
- Robotor bewegt sich Diagonal: Also zB von $(0,0)$ nach $(1,1)$ oder nach $(1,-1)$ oder nach $(-1,1)$ oder nach $(-1,-1)$.
  D.h.: Wenn $P(t) = (x,y)$, dann ist $P(t+1) = (x',y')$ möglich gdw $|x - x'| = |y - y'| = 1$.

Wir können nun beweisen dass der Roboter niemals $(0,1)$ erreichen kann.

Invarianten: Bsp

Behauptung: der Roboter erreicht niemals $(0, 1)$.

- Man kann aber nicht "direkt" mit Induktion beweisen $(\forall t) P(t) \neq (0, 1)$

  (Induktionsschritt ist unklar: $P(t) \neq (0, 1)$ impliziert ja nicht $P(t + 1) \neq (0, 1)$, wenn zB $P(t) = (1, 2)$ wäre.)
- Daher: Wir beweisen eine stärkere Aussage (die sehr wohl induktiv gezeigt werden kann).
- Das ist ein sehr üblicher Effekt / ein zentrales Thema / ein wichtiges Phänomen in der Mathematik:
  Oft ist es einfacher, eine (richtig gewählte!) stärkere Aussage zu beweisen.
  Analog: Oft ist es einfacher einen Satz allgemeiner, für mehr Strukturen, unter weniger Vorraussetzungen zu beweisen; für eine konkrete Frage abstrakte/allgemeine Methoden zu entwickeln.

Invarianten: Bsp

Behauptung: der Roboter erreicht niemals $(0, 1)$.
Wir zeigen eine stärkere Aussage: $(\forall t)P(t) = (x, y) \rightarrow 2|(x + y)$.
D.h.: Wenn der Roboter auf Position (x, y) kommt, dann ist $x + y$ gerade.

- Beweis mit (gewöhnlicher) Induktion:
- Induktionsanfang: Für $t = 0$ ist die Position $(0, 0)$, und $0 + 0$ ist gerade.
- Induktionsschritt: $P(t) = (x, y)$, mit $x + y$ gerade (Induktionsannahme). Dann gibt es für die Position zu $t + 1$ die Möglichkeiten:
  - $(x + 1, y + 1)$ mit $x + 1 + y + 1 = x + y + 2$ wieder gerade,
  - $(x + 1, y - 1)$ mit $x + 1 + y - 1 = x + y$ wieder gerade,
  - $(x - 1, y + 1)$ mit $x - 1 + y + 1 = x + y$ wieder gerade,
  - $(x - 1, y - 1)$ mit $x - 1 + y - 1 = x + y - 2$ wieder gerade.

  In jedem Fall gilt also ist für $(x', y') := P(t + 1)$: $x' + y'$ gerade.

"`x+y mod 2`" ist also eine "Invariante" dieses Systems.

Rekursive Datentypen: Ein Beispiel

In der Informatik werden Strukturen oft induktiv (oder: rekursiv) aufgebaut/definiert; dementsprechend kann man dann Aussagen über alle Instanzen dieser Struktur mit Induktion beweisen.

Ein Beispiel: Sei `Br` alle strings aus dem Alphabet $\{, \}$. Den leeren String nennen wir hier λ.
Bsp: λ, $\{\{$, $\{\{\}\}$ sind in `Br`
Wie können `Br` "rekursiv" definieren:

- Der leere String λ ist in `Br`
- Wenn x `Br` ist, dann auch $x\{$ und $x\}$.
- Nur die so konstruierten strings sind in `Br`.
  (Diesen Punkt werde ich in Zukunft weglassen, er ist in solchen Definitionen immer implizit mitgemeint.)

Induktion nach Aufbau von Strukturen: Ein Beispiel

Definieren wir nun `BBr` als solche strings mit "balancierten Klammern:"

- Der leere String λ ist in `BBr`
- Wenn X und Y in `BBr` sind, dann auch {X}Y

Beispiel: {}{} und {{}} sind in `BBr`; { oder }{ oder {{} aber nicht.
Es ist leicht zu sehen dass jeder string in `Bbr` gleich viele { wie } haben muss.
Wir können das auch einfach ordentlich beweisen, mit "Induktion nach Aufbau" des strings:

- (eine Variante des Induktionsbeginns)
  λ, hat gleichviele (nämlich 0) { wie }.
- (eine Variante des Induktionsschritts)
  Angenommen X und Y aus `Bbr` erhalten jeweils gleich viele { wie }, nämlich n für X und m für Y. Dann enthält {X}Y gleichviele { wie }, nämlich $n + m + 1$.

Rekursive Datentypen

Zusammenfassung:
Oft wird eine Klasse K von Objekten wie folgt definiert ("rekursiv" bzw "induktiv"):

- Es gibt bestimmte (endlich oder unendlich viele) Startobjekte, die zu unser Klasse gehören. Nennen wir die Menge dieser Statobjekte S. Dann können wir das schreiben als $S \subseteq K$.
- Es gibt bestimmte Operationen/Methoden, um aus Objekten der Klasse ein neues Objekt zu generieren.
  Nennen wir so eine Methode M, und nehmen wir an dass die Methode n viele Objekte verwendet um ein neues zu erzeugen. Dann können wir das so schreiben:
  $x_1, x_2, \ldots, x_n \in K \ \rightarrow M(x_1, x_2, \ldots, x_n) \in K$.
- Nur Objekte die auf diese Weise zustande kommen sind in der Klasse K, nicht mehr.

Induktiver Beweis nach Aufbau der Struktur

In dem Fall ist es zulässig, einen Satz $(\forall x \in K)\varphi(x)$ wie folgt mit Induktion nach Aufbau von x zu beweisen:

- Beweise dass φ für alle Startobjekte gilt, d.h. $(\forall x \in S)\varphi(x)$.
- Beweise dass wenn φ für bestimmte Objekte gilt, dann auch für die daraus konstruierten. D.h.:
  Für alle in der Definition verwendeten Methoden M gilt:
  $\varphi(x_1) \wedge \cdots \wedge \varphi(x_n) \;\rightarrow\; \varphi(M(x_1, \ldots, x_n))$

Bsp: eine Implementation von Polynomen

`Pol` kann man ebenfalls induktiv aufbauen:

- Die Startobjekte sind die strings `x` sowie folgen der Ziffern 0-9, also zB `027614`.
- Gegeben p und q in `Pol`, sind auch die Strings $(p{+}q)$ und $(p{*}q)$ und $-p$ in `Pol`.

Z.B. ist $((((-3)*(\texttt{x}*\texttt{x}))+-(3*\texttt{x}))+(-5))$ in `Pol` (und soll $-3x^2-3x-5$ entsprechen).

Wir könnten nun mittels Induktion beweisen z.B. dass ein Element p von `Pol` auf nur genau eine Weise erzeugt werden kann (und dass wir diesen Aufbau leicht aus p ablesen können. D.h. wir können `Pol` "parsen".)

Bemerkung: Die Halbgruppe der strings

Sei `Str` die Klasse aller strings über einem nicht-leeren Alphabet A; `Br` war ein Beispiel dafür.
Sei $\circ$ die Konkatenation (Hintereinangerschreiben zweier Strings).
Dann ist $(\texttt{Str}, \circ)$ ein Halbgruppe mit dem leeren String als neutrales Element. (D.h. $\circ$ ist assoziativ.)
`Str` kann man ebenfalls induktiv aufbauen, z.B. so:

- $S = A \cup \{\lambda\}$, d.h. jedes Symbol in A und auch der leere string ist in `Str`.
- Gegeben x und y in `Str`, ist auch $x \circ y$ (d.h. xy, x konkateniert mit y) in `Str`.

D.h. `Str` die aus S generierte Halbgruppe (so etwas haben wir immer als $\texttt{Str} = \langle S \rangle$ geschrieben).

Jetzt ist der Aufbau eines Strings aber nicht eindeutig: Wir können den String `abcd` aufbauen als: $((\texttt{a} \circ \texttt{b}) \circ \texttt{c}) \circ \texttt{d}$ oder als $(\texttt{a} \circ \texttt{b}) \circ (\texttt{c} \circ \texttt{d})$ (oder auch anders).

Rekursiv definierte Funktionen: Bsp 1

Auf rekursiv definierten Strukturen können wir rekursiv Funktionen definieren.

Bsp: Definiere die Auswertung $\texttt{eval} : \texttt{Pol} \times \mathbb{Z} \to \mathbb{Z}$ wie folgt:

- (Startobjekte): $\texttt{eval}(\texttt{x}, n) := n$, und wenn c eine String aus Ziffern ist die als Dezimalzahl $k \in \mathbb{N}$ ergibt, dann ist $\texttt{eval}(c, n) := k$.
- $\texttt{eval}((p{+}q), n) := \texttt{eval}(p, n) + \texttt{eval}(q, n)$;
  $\texttt{eval}((p{*}q), n) := \texttt{eval}(p, n) \cdot \texttt{eval}(q, n)$ und
  $\texttt{eval}(-p, n) := -\texttt{eval}(p, n)$.

Im Unterschied zum "induktiven Beweis einer Aussage nach Aufbau" muss man bei der "induktiven Definition" aufpassen dass das Ergebnis wohldefiniert ist, insbesondere wenn der Aufbau nicht eindeutig ist.

Im Fall von `Pol` ist der Aufbau eindeutig und alles funktioniert gut.

Rekursiv definierte Funktionen: Bsp 2

Weil der Aufbau eindeutig ist, können wir auch die "Tiefe" definieren:
Bsp: Definiere die Tiele $f : \texttt{Pol} \to \mathbb{N}$ wie folgt:

- (Startobjekte): $f(p) = 0$ wenn $p \in S$ (d.h. p ist `x` oder ein String aus Ziffen).
- $f((p{+}q)) := \max(f(p), f(q)) + 1$, und genauso
  $f((p{*}q)) := \max(f(p), f(q)) + 1$, und
  $f(-p) := \max(f(p)) + 1$.
  (D.h. jeder "Konstruktionsschritt" erhoeht die Tiefe um 1)

Den (eindeutigen) Aufbau kann man sich wie einen Baum vorstellen, die "Tiefe" ist dann die Höhe des Baums (nicht die Breite).

Ein Problem

Rekursive Definitionen können fehlschlagen, insbesondere wenn der Aufbau nicht eindeutig ist:
Definiere die "Tiefe" $f : \mathtt{Str} \to \mathbb{N}$ eines Strings wie folgt:

- Der leere String λ hat Tiefe 0. $f(\lambda) = 0$.
  Wenn $a \in A$ ist, dann ist $f(1) = 1$.
- Wenn $z = x \circ y$, dann ist $f(z) = \max(f(x), f(y)) + 1$.

Hier ist f nicht wohldefiniert:

- Sei $x = \mathtt{abcd}$.
- Wenn wir x aufbauen als $x = ((\mathtt{a} \circ \mathtt{b}) \circ \mathtt{c}) \circ \mathtt{d}$, dann wäre $f(x) = 4$.
- Beim Aufbau $x = (\mathtt{a} \circ \mathtt{b}) \circ (\mathtt{c} \circ \mathtt{d})$ wäre $f(x) = 3$.

Eindeutigkeit und Uneindeutigkeit

Wir könnten natürlich dieselbe Klasse `Str` (strings über dem Alphabet A) auch "eindeutig" aufbauen:

- Der leere String λ ist in `Str`.
- Wenn $a \in A$ und x in `Str`, dann ist ach $x \circ a$ in `Str`.

In dem Fall ist der Aufbau eindeutig. $x =$ `abcd` wäre z.B. aufgebaut als $(((a \circ b) \circ c) \circ d$.
Die analoge "Tiefen-Fuinktion" ist nun einfach die Länge des strings.

Rekursiv definierte Funktionen über $\mathbb{N}$

Wir können uns auch $\mathbb{N}$ auf triviale Weise rekursiv aufgebaut denken, durch 0 und die Funktion $+1$ (hier ist der Aufbau eindeutig).

Beispiele für rekursive Definitionen in $\mathbb{N}$:

- $f(0) := 1$ und $f(n+1) := 2^{f(n)}$. Definiert eindeutige Funktion. (Weil wir für $n+1$ nur auf n zugreifen entspricht das in etwa der "gewöhnlichen Induktion".)
- $f(0) := 0$, $f(1) := 1$, und $f(n+2) := f(n) + f(n+1)$ ist wohldefiniert (Fibinacci Folge). (Entspricht eher "vollständiger Induktion".)
- $f(0) := 1$ und $f(n+1) := f(n+2)$ ist offenbar nicht wohldefiniert ($0 \mapsto 1$ und $n \mapsto c$ für $n > 0$ ist Lösung, für jede Konstante $c \in \mathbb{N}$)
- Bei $f(n+1) := 2f(n)$ fehlt ein "Anfang", die Gleichung wird durch jede Funktion $n \mapsto c \cdot 2^n$ erfüllt (für eine Konstant $c \in \mathbb{N}$).

Bsp

Wir können uns $\mathbb{N}$ auch aufgebaut denken durch $0, 1$ und die Funktionen $+, \cdot$ (nicht eindeutig). Bei entsprechenden Induktiven Definitionen muss man überprüfen ob die Definition konsistent ist.

Ein Beispiel bei der die Definition fehlschlägt:
Setze $f(0) = f(1) = 0$, $f(p) = p$ für p Primzahl, und $f(n) = 2f(x) + f(y)$ für $n = x \cdot y$ mit $1 < x \leq y < n$.

Das ist inkonsistent:

- $f(12) = f(2 \cdot 6) = 2f(2) + f(2 \cdot 3) = 4f(2) + f(3) = 4 \cdot 2 + 3 = 11$, und
- $f(12) = f(3 \cdot 4) = 2 \cdot f(3) + f(2 \cdot 2) = 6 + 2 \cdot f(2) + f(2) = 6 + 6 = 12$.

Es gibt also kein f dass die rekursive Definition erfüllt.

Induktionsbeweis vs rekursice Definition

Moral:

- Für den Induktionsbeweis einer Aussage ist Nicht-Eindeutigkeit des Aufbaus kein Problem.
- Z.B: Wenn $\varphi(p)$ gilt für jede Primzahl p und für $p = 0, 1$, und $\varphi(x) \wedge \varphi(y) \to \varphi(x \cdot y)$, dann gilt $\varphi(n)$ für alle $n \in \mathbb{N}$ (vollständige Induktion). Dabei ist es egal dass wir auf verschiedenem Weg zu einem x (z.B. 12) kommen könenn, sobald es einen weg gibt ist $\varphi(12)$ bewiesen.
- Wenn wir jedoch $f(x \cdot y)$ durch rekursiv durch die Werte von $f(x)$ und $f(y)$ definieren, dann können wir (trotz Wohlfundiertheit) Inkonsistenzen bekommen, weil wir auf verschiedenen Wefgen zu z.B. 12 kommen können.

Noch etwas Zahlentheorie

Wiederholung: Teilbarkeit, Primzahlen

Im folgenden betrachten wir Zahlen in $\mathbb{Z}$ (oft nur in $\mathbb{N}$ oder $\mathbb{N} \setminus \{0\}$, sollte aus Kontext klar sein wenn nicht explizit erwähnt.)

- $a|b$ heißt "a teilt b', d.h., $(\exists c) a \cdot c = b$.
  Bsp: $2|3$, $1|3$, $2|0$, $-10|20$, $2|-6$; aber $3 \nmid 5$, $0 \nmid 2$, $2 \nmid 1$

Ein paar einfache Eigenschaften der Teilbarkeit:

- $a|b$ und $b|c$ impliziert $a|c$
- $a|b$ und $a|c$ impliziert $a|xb + yc$ für alle $x, y \in \mathbb{Z}$.
  Def.: Wir nennen $xb + yc$ eine "Linearkombination von b und c".
- Wenn $c \neq 0$, dann $a|b$ gdw $ac|bc$.

Die wesentlichen Bausteine der natürlichen Zahlen (was die multiplikative Struktur anbelangt) sind die Primzahlen (PZ):

- p ist Primzahl, wenn $p \geq 2$ und p nur 1 und p als (positive) Teiler hat.

Zahlentheorie (kein Prüfungsstoff)

Zahlentheorie beschäftigt sich mit den natürlichen Zahlen, Teilbarkeit, etc.
Ein Gebiet in dem "ein Narr mehr Fragen stellen kann als tausend Weise beantworten können". Ein seit 2500 Jahren ungelöstes Problem:

- n heißt "perfekt" wenn n die Summe aller Teiler $\neq n$ ist.
  Bsp: $6 = 1 + 2 + 3$ ist perfekt, $8 \neq 1 + 2 + 4$ nicht.
  Man kann die geraden perfekten Zahlen recht gut charakterisieren. (Euklid, ca 300 v.u.Z.)
  Frage: Gibt es eine ungerade perfekte Zahl?

Besonders wichtig / interessant sind Fragen über Primzahlen.
Hier ein paar ungelöste:

- (Goldbachsche Vermutung) Ist jede gerade Zahl > 2 Summe zweier Primzahlen?
  (Was man weiss: Gilt für $n \leq 10^{16}$. Jede gerade Zahl ist Summe von höchstens 6 PZ.)
- Gibt es unendlich viele Primzahlzwillinge, d.h., p PZ mit $p + 2$ PZ?
  (Was man weiss: Es gibt unendlich viele PZ p s.d. $p + 2$ Produkt von höchstend zwei PZ ist.)

Wiederholung: ggT

ggT(a, b) ist der größte gemeinsame Teiler von a und b.
Für $a, b \geq 1$ gilt:

- ggT(a, b) existiert,
- ggT(a, b) is Linearkombination von a und b.
  (D.h.: Es gibt $x, y \in \mathbb{Z}$ s.d. ggT$(a, b) = xa + yb$.)
- Allgemeiner: c ist Linearkombination von a und b gdw c Vielfaches von ggT(a, b)
- Insbes.: ggT(a, b) ist die kleinste positive Linearkombination von a und b.

Mit dem Euklidischen Algorithmus kann man ggT(a, b), und das x und y der Linearkombination, effizient berechnen. Diesem Algorithmus liegt Division mit Rest zugrunde:

Für $a, b \in \mathbb{Z}$, $b > 0$ gibt es eindeutige $s \in \mathbb{Z}$ und $0 \leq r < b$ mit $a = s \cdot b + r$.

Fundamentalsatz

Theorem (Fundamentalsatz der Arithmetik)

Jedes $n > 1$ hat eine eindeutige Primfaktorenzerlegung (PFZ)
$n = p_1 \cdot p_2 \cdots p_i$, *mit* $p_1 \leq p_2 \leq \cdots \leq p_i$.

($p_1 \leq \cdots$ ist für die Eindeutigkeit notwendig; sonst wären $2 \cdot 3$ und $3 \cdot 2$ verschiedene PFZ von 6.)

Wir wissen bereits dass jedes n eine PFZ hat (der Beweis war ein Beispiel für vollsändige Induktion).

Es bleibt die Eindeutigkeit zu zeigen. Zuerst zwei Lemmata:

Beweis Fundamentalsatz: Zwei Lemmata

- Lemma 1:
  Wenn p Primzahl ist und $p|ab$, dann $p|a$ oder $p|b$.
  Beweis: ggT(a,p) kann nur 1 oder p sein. Wenn ggT$(a,p) = p$, dann $p|a$ und wir sind fertig. Wenn ggT$(a,p) = 1$ dann ist $1 = px + ay$ und damit $b = bpx + bay$, und weil $p|p$ und $p|ab$, gilt $p|pbx$ und $p|aby$, und daher $p|b$. □

  (Bemerkung: Diese Eigenschaft ist eine alternative Definition von Primzahl.)

- Lemma 2: Wenn $p|a_1 a_2 \cdots a_i$, dann $p|a_j$ für ein $1 \leq j \leq i$.
  Beweis: Induktion über $i \geq 1$.
  Induktionbeginn: Stimmt für $i = 1$.
  Induktionsschritt: Wir setzten den Satz für i voraus.
  Nun sei $p|a_1 a_2 \cdots a_i a_{i+1}$. Setze $b := a_1 a_2 \cdots a_i$. Es gilt also $p|ba_{i+1}$.
  Aus Lemma 1 folgt: $p|a_{i+1}$ oder $p|b$. Wenn $p|b$ dann folgt aus der Induktionsannahme dass $p|a_j$ für ein $j \leq i$. □

Beweis Fundamentalsatz

- Wir wissen bereit: Jede Zahl $n \geq 2$ ist das Produkt von Primzahlen, d.h. hat eine PFZ.
- Es bleibt zu zeigen: Diese PFZ ist eindeutig (abgesehen allenfalls von der Ordnung)
- Angenommen es gibt ein n das zwei verschiedene PFZ hat.
  Dann sei n_0 die kleinste solche Zahl.
- Seien $n_0 = p_1 \cdots p_i = q_1 \cdots q_j$ zwei verschiedene PFZ.
- $p_1 | n_0$, daher $p_1 | q_1 \cdots q_j$ daher $p_1 | q_\ell$ für ein ℓ.
  Weil q_ℓ PZ, folgt $p_1 = q_\ell$ (q_ℓ hat ja nur die Teiler 1 und p_ℓ).
- Wenn wir p_1 vom ersten Produkt entfernen und $p_1 = q_\ell$ vom zweiten, bekommen wir also zwei PFZ $\frac{n_0}{p_1} = p_2 \cdots p_n = q_1 \cdots q_{\ell-1} q_{\ell+1} \cdots q_j$.
  Das sind zwei tatsächlich verschiedene PFZ (sonst wären die ursprünglichen PFZ von n_0 ja auch gleich).
  Das ist ein Widerspruch zur Minimalität von n_0. □