Mathematischer Brückenkurs

Version vom SS 2023*

Universität Mainz
Fachbereich 08
Theorie der kondensierten Materie
Prof. Dr. Friederike Schmid†

Mathematischer Brückenkurs

- Grundlegendes
  - Schreibweisen, Einheiten, Größenordnungen
  - Zahlen
  - Elementare Rechenoperationen
  - Gleichungen und Ungleichungen
- Elemente der linearen Algebra
  - Lineare Gleichungssysteme
  - Vektoren
  - Matrizen
- Reelle Funktionen
- Infinitesimalrechnung
  - Differentialrechnung
  - Folgen, Reihen, Potenzreihen
  - Integralrechnung
- Statistik und Fehlerrechnung
  - Wahrscheinlichkeit (inklusive Kombinatorik)
  - Verteilungen
  - Stichproben, Konfidenz, Beobachtungsfehler
  - (Zusatzstoff: Parameter-Fitting)

---

*Elektronisch: Letzte Änderung am 7.04.2023

†03-523, Tel. (06131-)39-20365, <friederike.schmid@uni-mainz.de>

Literatur

K. Hefft Mathematischer Vorkurs zum Studium der Physik
(online unter `http://www.thphys.uni-heidelberg.de/~hefft/vk1/`)

E. Cramer, J. Neslehova Vorkurs Mathematik
Arbeitsbuch zum Studienbeginn in Bachelorstudiengängen

M. Klinger Vorkurs Mathematik für Nebenfach-Studierende

Inhaltsverzeichnis

1 Grundlegendes **7**

1.1 Die Sprache der Physik . . . 7

1.1.1 Zeichen für physikalische und mathematische Größen . . . 8

1.1.2 Zeichen für Verknüpfungen . . . 9

1.1.3 Einheiten . . . 10

1.2 Zahlen . . . 12

1.2.1 Vorab: Mengen, Gruppen, Ringe, Körper . . . 12

1.2.2 Natürliche Zahlen $\mathbb{N}$ . . . 13

1.2.3 Ganze Zahlen $\mathbb{Z}$ . . . 15

1.2.4 Rationale Zahlen $\mathbb{Q}$ . . . 15

1.2.5 Reelle Zahlen $\mathbb{R}$ . . . 16

1.2.6 Komplexe Zahlen $\mathbb{C}$ . . . 17

1.2.7 Zusammenfassung . . . 17

1.3 Elementare Rechenoperationen . . . 18

1.3.1 Bruchrechnung . . . 18

1.3.2 Potenzrechnung . . . 19

1.4 Gleichungen und Ungleichungen . . . 21

1.4.1 Lineare Gleichungen und Ungleichungen . . . 21

1.4.2 Quadratische Gleichungen und Ungleichungen . . . 22

1.4.3 Polynomgleichungen höherer Ordnung . . . 23

1.4.4 Kompliziertere Gleichungen . . . 23

2 Elemente der linearen Algebra **27**

2.1 Lineare Gleichungssysteme . . . 27

2.1.1 Zwei Gleichungen und zwei Unbekannte . . . 28

2.1.2 Matrixdarstellung von linearen Gleichungssystemen . . . 29

2.1.3 Gaußsches Eliminationsverfahren zum Lösen linearer Gleichungssysteme . . . 30

2.2 Vektoren . . . 32

2.2.1 Definition bzw. Begriffsklärung . . . 32

2.2.2 Koordinatensysteme und Koordinatendarstellung 32
2.2.3 Elementares Rechnen mit Vektoren, Vektorräume 33
2.2.4 Skalarprodukt (inneres Produkt) 35
2.2.4.1 Definition und mathematische Struktur 35
2.2.4.2 Koordinatendarstellung und Kronecker-Symbol . 35
2.2.5 Vektorprodukt (äußeres Produkt, Kreuzprodukt) 36
2.2.5.1 Definition und mathematische Struktur 36
2.2.5.2 Koordinatendarstellung und Levi-Civita-Symbol 37
2.2.6 Höhere Vektorprodukte 38
2.3 Matrizen 39
2.3.1 Beispiele von Matrizen 39
2.3.2 Elementare Begriffe 40
2.3.3 Rechnen mit Matrizen 41
2.3.3.1 Elementare Operationen 41
2.3.3.2 Invertieren von Matrizen 41
2.3.3.3 Charakteristische Größen quadratischer Matrizen 43
2.3.4 Determinanten 44
2.3.4.1 Einführung und Definition 44
2.3.4.2 Rechenregeln für Determinanten 45
2.3.4.3 Folgerungen und Anwendungen 46
2.3.5 Drehungen und Drehmatrizen 47

3 Reelle Funktionen 49
3.1 Elementare Funktionen 49
3.1.1 Polynome und rationale Funktionen 49
3.1.2 Algebraische Funktionen 50
3.1.3 Exponentialfunktion 51
3.1.4 Logarithmus 52
3.1.5 Trigonometrische Funktionen 53
3.1.6 Hyperbolische Funktionen 53
3.1.7 Funktionen mit Ecken und Sprüngen 54
3.1.8 Weitere wichtige abgeleitete Funktionen 54
3.2 Eigenschaften von Funktionen 55
3.2.1 Spiegelsymmetrie 55
3.2.2 Beschränktheit 55
3.2.3 Monotonie 55
3.2.4 Eineindeutigkeit 55
3.2.5 Stetigkeit 56
3.2.6 Grenzwerte 56

4 Infinitesimalrechnung **59**
4.1 Differentialrechnung 59
4.1.1 Die Ableitung 59
4.1.2 Elementare Beispiele 61
4.1.3 Differentiationsregeln 62
4.1.4 Anwendungen der Differentiationsregeln 64
4.1.5 Tabelle wichtiger Ableitungen 65
4.2 Taylor-Entwicklung 66
4.2.1 Allgemein: Folgen und Reihen 66
4.2.1.1 Folgen 66
4.2.1.2 Reihen 67
4.2.2 Kurzer Abriss über Potenzreihen 69
4.2.3 Konstruktion der Taylor-Reihe 71
4.2.4 Anwendungen 72
4.3 Integralrechnung 75
4.3.1 Das Riemannsche Integral 75
4.3.2 Hauptsatz und Stammfunktion 76
4.3.3 Integrationsmethoden 78
4.3.4 Uneigentliche Integrale 81
4.3.5 Mehrfachintegrale 83
4.3.5.1 Beispiele 83

5 Statistik und Fehlerrechnung **85**
5.1 Motivation und Einführung 85
5.2 Wahrscheinlichkeitstheorie 87
5.2.1 Konzept und Interpretation von Wahrscheinlichkeit 87
5.2.2 Klassisch-intuitive Berechnung von Wahrscheinlichkeiten 88
5.2.2.1 "Klassische" Wahrscheinlichkeit 88
5.2.2.2 Die vier Grundprobleme der Kombinatorik . . . 89
5.2.2.3 Grenzen der klassischen Beschreibung 90
5.2.3 Mathematischer Wahrscheinlichkeitsbegriff 91
5.2.4 Bedingte Wahrscheinlichkeit 92
5.2.5 Verteilungen 92
5.2.5.1 Definitionen 93
5.2.5.2 Wichtige Verteilungen 93
5.2.5.3 Charakterisierung von Verteilungen 94
5.2.5.4 Der zentrale Grenzwertsatz 95
5.3 Stichproben und Beobachtungsfehler 96
5.3.1 Charakterisierung von Stichproben 96

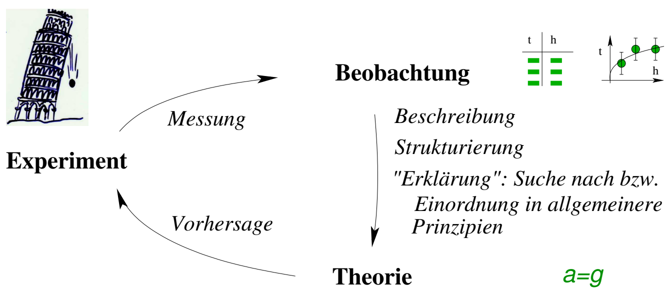
5.3.2 Schätzung von Erwartungswert und Varianz 97
5.3.3 Fehlerabschätzung und Verteilung des Mittelwertes 98
5.3.4 Fehlerfortpflanzung 98
5.4 Parameter-Fitting 100
5.4.1 Bayes-Schätzer und Maximum Likelihood (ML) 101
5.4.2 Methode der kleinsten Fehlerquadrate 102
5.4.3 Lineare Regression 103

Kapitel 1

Grundlegendes

1.1 Die Sprache der Physik

Kreislauf der empirischen Wissenschaften:



In der Physik (und großen Teilen der Chemie, zunehmend Biologie, Volkswirtschaft ...) spielt Mathematik als Sprache in diesem Kreislauf eine zentrale Rolle.

(NB: Nicht einfach nur als Hilfsmittel: Sprache bestimmt das Denken).

Historisch: Physik und Mathematik

Bis Mitte des 19. Jahrhundert (~ 1830)

- Mathematik im Wesentlichen zur Beschreibung von Experimenten
  (Funktionaler Zusammenhang zwischen Messwerten)
- Spekulationen über allgemeine Zusammenhänge i.A. umgangssprachlich
  (nicht-technisch)
  ("experimentelle Philosophie")

Beschäftigung mit der Physik war i. A. Nebenbeschäftigung.

Seit ~ 1870

- Mathematik ist die Sprache der Theorie.
  → Interpretation der Natur wird von Mathematik geleitet.
  Wichtige Rolle dabei: Entwicklung der Infinitesimalrechnung

Physik wird zum Beruf (Universitäten)

Beispiel: Christiaan Huygens: "Horologium Olscillatorium", 1673
(frühe Mechanik, 14 Jahre vor Newtons "Principia")

über fallende Körper:
"Zu gleichen Zeiten werden einem fallenden Körper gleiche Mengen Geschwindigkeit hinzugefügt, und zu gleichen Zeiten erhöhen sich die von einem Körper durchquerten Abstände sukzessive um den gleichen Betrag."

Übersetzung in die heutige Sprache:
$\frac{\mathrm{d}v}{\mathrm{d}t}$ = const. oder $\dot{v}$ = const.
$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\mathrm{d}x}{\mathrm{d}t}$ = const. oder $\ddot{x}$ = const.

Zutaten zu einer mathematischen Formulierung physikalischer Sachverhalte:

- Zahlen (2, -3.5, π, ∞)
- Einheiten (m, km, s, EUR)
- Zeichen für

  Physikalische Größen
  Manipulationen und Verknüpfungen

Im Einzelnen:

1.1.1 Zeichen für physikalische und mathematische Größen

Lateinisches Alphabet
$a, b, c, \cdots$ $A, B, C, \cdots$

Griechisches Alphabet

$\alpha\ \beta\ \gamma\ \delta\ \varepsilon\ \zeta\ \eta\ \theta\ \iota\ \kappa\ \lambda\ \mu\ \nu\ \xi\ o\ \pi\ \rho\ \sigma\ \tau\ \upsilon\ \varphi\ \chi\ \psi\ \omega$
$A\ B\ \Gamma\ \Delta\ E\ Z\ H\ \Theta\ I\ K\ \Lambda\ M\ N\ \Xi\ O\ \Pi\ P\ \Sigma\ T\ \Upsilon\ \Phi\ X\ \Psi\ \Omega$

Altdeutsches Alphabet, Fraktur oder Sütterlin (heutzutage selten)
$\mathfrak{a}, \mathfrak{b}, \mathfrak{c}\cdots$ $\mathfrak{A}, \mathfrak{B}, \mathfrak{C}\cdots$
a, b, c ··· A, B, C ···

Spezielle Zeichen, z.B. ∞ (Unendlich), $\varnothing$ (Leere Menge)

z.B. Newtonsches Kraftgesetz: Kraft = Masse mal Beschleunigung
übersetzt: $F = m \cdot a$ (F = "Force", a = "acceleration")
oder $K = m \cdot b$ (K = "Kraft", b = "Beschleunigung")

$\rightsquigarrow$ Notation wird i. A. nicht einheitlich sein
(von Buch zu Buch, von Dozent zu Dozent, sogar innerhalb eines Kurses)

(Schlimmes Beispiel: $E = mc^2$ könnte die Gleichung für die Beschleunigung auf einer Kreisbahn sein: Radius · Winkelgeschwindigkeit^2)

1.1.2 Zeichen für Verknüpfungen

Einfache mathematische Zeichen:

$=, \neq, \equiv$
$<, >, \leq, \geq$
$+, -, \cdot$(oft weggelassen), $/, \pm$
$\approx, \sim, \propto, \ll, \gg$

Zeichen der Mengenlehre und logische Zeichen:

$\in, \ni, \notin$
$\subset, \supset, \subseteq, \supseteq$
$\cup, \cap, \setminus$
$\exists, \exists!, \forall$ (sogenannte Quantoren)
$\Leftarrow, \Rightarrow, \Leftrightarrow$
$:=$
Beispiel: $\forall x : \exists! y : x + y = 0$

Kompliziertere Zeichen:

Summenzeichen $\sum$: $\sum_{k=1}^n a_k = a_1 + a_2 + \cdots + a_n$

Beispiele:

<u>Gaußsche Summenformel</u>

$\sum_{k=0}^n k = 0 + 1 + 2 + 3 + \cdots + n = n(n+1)/2$

(denn: $\sum_{k=1}^n k = \frac{1}{2}(\sum_{k=1}^n k + \sum_{k=1}^n (n+1-k)$
$= \frac{1}{2} \sum_{k=1}^n (n+1) = \frac{1}{2} n(n+1))$.

<u>Geometrische Summe</u>

$\sum_{k=0}^n x^k = 1 + x + \cdots + x^n = (x^{n+1} - 1)/(x-1)$ für $x \neq 1$

(denn: $S_n := \sum_{k=0}^n x^k \;\Rightarrow xS_n - S_n = x^{n+1} - 1 = S_n(x-1))$

<u>Geometrische Reihe</u>

Speziell: $\sum_{k=0}^\infty x^k = 1/(1-x)$ für $|x| < 1$.

Produktzeichen $\prod$: $\prod_{k=1}^n a_k = a_1 \cdot a_2 \cdots a_n$

Fakultät: $n! = 1 \cdot 2 \cdot 3 \cdots n$

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

Bedeutung des Binomialkoeffizienten (eine von vielen)

$\rightarrow$ <u>Binomischer Lehrsatz</u>: $\boxed{(a+b)^n = \sum_{k=0}^n \binom{n}{k} a^k b^{n-k}}$

Aufgaben

- Berechnen Sie $\sum_{k=1}^4 (-1)^k, \prod_{k=1}^4 (-1)^k, \binom{6}{3}$.
- Berechnen Sie $\binom{n}{0}, \binom{n}{1}, \binom{n}{n}$
- Zeigen Sie $\binom{n}{k} = \binom{n}{n-k}$
- Zeigen Sie $\binom{n}{k-1} + \binom{n}{k} = \binom{n+1}{k}$
- Zeigen Sie $\sum_{k=0}^n \binom{n}{k} = 2^n$
- Zeigen Sie $\sum_{k=1}^n k^2 = \frac{1}{6} n(n+1)(2n+1)$.
  Trick: Ausmultiplizieren von $\sum_{k=1}^n (k+1)^3$.

- Betrachten Sie folgenden "Beweis", dass Wasserstoff=Chlor. Wo steckt der Wurm?
  Aus $H_2 + Cl_2 = 2$ HCl
  folgt H H + Cl Cl - 2 H Cl = 0
  $\Rightarrow$ (H-Cl)$^2 = 0$
  $\Rightarrow$ H-Cl = 0
  $\Rightarrow$ H = Cl

1.1.3 Einheiten

Vorab: Einheiten $\leftrightarrow$ Maßstäbe
Die Festlegung von Maßstäben ist eine eigene Wissenschaft.
(Maßstäbe für Zeit, Länge, Masse, Ladung, ···)

Darüberhinaus gibt es aber noch verschiedene Einheitensysteme.
Häufig (nicht immer) benutzt man heute SI-Einheiten.
(système international d'unités)

mit Grundeinheiten
Länge: Meter (m)
Zeit: Sekunde (s)
Masse: Kilogramm (kg)
Elektrischer Strom: Ampère (A)
Lichststärke: Candela (cd)
Temperatur: Kelvin (K)
(Temperatur in K ist Temperatur in Grad Celsius + 273.14.)
Stoffmenge: Mol (mol)
(1 mol sind $6.02214076 \cdot 10^{23}$ Einzelteilchen.)
ebener Winkel: Radian (rd)
(Eine volle Umdrehung entspricht 2π rd)
Raumwinkel: Steradian (srd)
(Der Raumwinkel in eine Halbkugel ist 2π srd)

und vielen abgeleiteten Einheiten, z.B.
Frequenz: Hertz (Hz = 1/s)
Kraft: Newton (N = kg m/s^2)
Druck: Pascal (Pa = N/m^2)
Energie: Joule (J = N m)
Leistung: Watt (W = J/s)
Elektrische Spannung: Volt (V=J/As)
Elektrischer Widerstand: Ohm (Ω = V/A)

Andere auch gebräuchliche Einheiten
Winkel: Grad ^{0} = $(\pi/180)$ rd = 0.01745 rd,
Minute ' = $\frac{1}{60}{}^0$, Sekunde " = $\frac{1}{60}{}'$,
Masse: atomare Masse u=$1.66 \cdot 10^{-27}$ kg
Energie: Elektronenvolt eV = $1.6022 \cdot 10^{-19}$ J
verschiedene Einheitensysteme im elektromagnetischen Bereich

Größenordnungen

Um Nullen zu vermeiden → international standardisierte Dezimalvorsilben
(z.B. m, km, mm, μm, $\cdots$, entsprechend g, kg, mg, μg, $\cdots$

$10^{-1} \equiv$ d (Dezi-)
$10^{-2} \equiv$ c (Zenti-)
$10^{-3} \equiv$ m (Milli-)
$10^{-6} \equiv \mu$ (Mikro-)
$10^{-9} \equiv$ n (Nano-)
$10^{-12} \equiv$ p (Pico-)
$10^{-15} \equiv$ f (Femto-)
$10^{-18} \equiv$ a (Atto-)

$10^{1} \equiv$ D (Deka-)
$10^{2} \equiv$ h (Hekto-)
$10^{3} \equiv$ k (Kilo-)
$10^{6} \equiv$ M (Mega-)
$10^{9} \equiv$ G (Giga-)
$10^{12} \equiv$ T (Tera-)
$10^{15} \equiv$ P (Peta-)

Aufgaben

- Wie viele Kubikdezimeter Wasser hält ein Kanister mit Länge 80 cm, Breite 40 cm, Höhe 60 cm?
- Wie viele Stunden und Minuten sind 18600 Sekunden?
- Wieviele Platten der Größe 10 cm × 20 cm braucht man, um eine Fläche von 50 m^2 zu pflastern?
- Welchen Längen entsprechen 1 Grad, 1 Minute, 1 Sekunde, 1 rad auf der Erdoberfläche (Umfang ist 40.000 kkm).
- Welchen Flächen entspricht auf der Erde 1 Quadratgrad, 1 Quadratminute, 1 Quadratsekunde, 1 Steradian?
- Wieviele Steradian hat eine Kugel?
- Wie lange ist ein Mikrojahrhundert?

1.2 Zahlen

Dieses Kapitel: Rudimentäre Übersicht über verschiedene Arten von Zahlen und deren Eigenschaften.

NB: Obwohl Zahlen uns sehr vertraut sind, sind sie eigentlich höchst abstrakte Gebilde

1.2.1 Vorab: Mengen, Gruppen, Ringe, Körper

- **Mengen**: Verbund von Objekten (Elementen), z.B. $A = \{a, b, c\}$.
  Elemente: Objekte in einer Menge ($a \in A$)
  leere Menge: Menge ohne Elemente ($\varnothing$)
  Teilmenge: z.B. $B = \{b, c\}$ ist Teilmenge von $A = \{a, b, c\}$ ($B \subset A$)
  Obermenge: z.B. $D = \{a, b, c, d\}$ ist Obermenge von $A = \{a, b, c\}$ ($D \supset A$)
  Vereinigungsmenge, Schnittmenge: z.B. sei $C = \{c, d\}$
  $\Rightarrow A \cup C = \{a, b, c, d\}, A \cap C = \{c\}$

- **Relation**: Beziehung zwischen Elementen einer Menge M: $x \sim y$
  Mögliche Attribute von Relationen:
  Reflexiv, wenn $x \sim y \quad \forall x$.
  Symmetrisch, wenn mit $x \sim y$ immer auch $y \sim x$ gilt.
  Transitiv, wenn aus $x \sim y$ und $y \sim z$ folgt: $x \sim z$
  Äquivalenzrelation: Reflexiv, symmetrisch und transitiv.

  (Beispiel: Betrachte die Menge der Quotienten von Zahlen, z.B. $\frac{1}{2}, \frac{2}{4}, \frac{3}{5}$
  $<$: transitiv, $\le$: transitiv und reflexiv, $=$: Äquivalenzrelation)

- **Abbildung** von einer Menge N in eine Menge N: Vorschrift, die jedem Element $a \in M$ ein Element $\varphi(a) \in N$ zuordnet: $a \to \varphi(a)$.
  Dann ist $\varphi(a)$ das Bild von a, a das Urbild von $\varphi(a)$
  Mögliche Eigenschaften von Abbildungen:
  Surjektiv, wenn alle Elemente von N ein Urbild haben.
  Injektiv oder eindeutig, wenn es jeweils genau ein Urbild gibt.
  Bijektiv: Injektiv und surjektiv (eineindeutig).

- **Verknüpfung**: Abbildung von der Menge $M \times M$ (der "Produktmenge") aller Paare (a, b) mit $a, b \in M$: $(a, b) \to \varphi(a, b)$ oder $a \circ b$.
  (z.B. auf Zahlenmengen: Addition, Multiplikation)

- **Gruppe**: Menge G mit einer Verknüpfung mit den Eigenschaften:
  1) Abgeschlossenheit: $a \in G, b \in G \Rightarrow a \circ b \in G$
  2) Assoziativgesetz: $(a \circ b) \circ c = a \circ (b \circ c)$
  3) Neutrales Element: Es existiert eine "Identität" $e \in G$
  mit $e \circ a = a$ für alle $a \in G$.
  4) Inverses Element: $\forall a \in G : \ \exists a^{-1} \in G : \ a^{-1} \circ a = e$

- **Abelsche Gruppe**: Gruppe mit der Zusatzeigenschaft
  5) Kommutativgesetz: $a \circ b = b \circ a$

- **Halbgruppe**: Menge mit Verknüpfung, die die Eigenschaften 1) und 2) erfüllt (Abgeschlossenheit und Assoziativgesetz).

- **Ring**: Menge R mit zwei Verknüpfungen $+, \cdot$ (Addition, Multiplikation)
  – Abelsche Gruppe bzgl. Addition $(+)$
  – Halbgruppe bzgl. Multiplikation $(\cdot)$
  – Distributivgesetze: $a \cdot (b+c) = a \cdot b + a \cdot c$ und $(a+b) \cdot c = a \cdot c + b \cdot c$
  Identität bzgl. de Addition heißt üblicherweise Null.

  Folgerung: Zum Beispiel binomische Formeln
  Allgemein: $(a+b)(a-b) = (a^2 + ba - ab - b^2)$
  $(a \pm b)^2 = (a^2 \pm ab \pm ba + b^2)$
  Falls Multiplikation kommutativ: $(a+b)(a-b) = (a^2 - b^2)$
  $(a \pm b)^2 = (a^2 \pm 2ab + b^2)$

- **Körper**: Ein Ring K mit folgenden Eigenschaften:
  – Mindestens ein Element ist von Null verschieden.
  – $K \setminus \{0\}$ (K ohne Null) ist von Abelsche Gruppe bzgl. Multiplikation.

Bemerkung: Diese Strukturen wurden natürlich im Hinblick auf die Zahlenmengen eingeführt, die gleich besprochen werden. Sie sind aber auch von viel allgemeinerer Bedeutung: Zusammenhänge, die man nur aus Gruppen-, Ringen- oder Körpereigenschaften beweist, gelten ganz allgemein in allen derartigen Strukturen.

Aufgaben:
– Zeigen Sie: Für Gruppen folgt aus Gruppeneigenschaften auch $a \circ a^{-1} = e$, $a \circ e = a$.
– Sei R ein Ring. Zeigen Sie, daß dann die Menge der Polynome $f(x) = \sum_{k+0}^{n} a_k x^k$ mit $a_k \in R$, n eine natürliche Zahl (siehe 1.2.2), x ein Symbol, auch ein Ring ist.

1.2.2 Natürliche Zahlen $\mathbb{N}$

$\mathbb{N} = \{1, 2, 3, \cdots\}$: "Natürlich", da seit alters her zum Zählen benutzt.

Axiomatische Definition über Peano-Axiome (Giuseppe Peano, 1858-1932)

I 1 ist eine natürliche Zahl.

II Jede natürliche Zahl a hat einen Nachfolger a^+ in der Menge der natürlichen Zahlen. $(\forall a \in \mathbb{N} : \exists! a^+ \in \mathbb{N})$

III Es gibt keine Zahl mit dem Nachfolger 1.

IV Aus $a^+ = b^+$ folgt $a = b$.

V Prinzip der vollständigen Induktion: Enthält eine Teilmenge der natürlichen Zahlen die Zahl 1 und zu jeder Zahl a auch deren Nachfolger a^+, enthält sie alle natürlichen Zahlen.

Aus dem letzten Axiom folgt Beweismethode der vollständigen Induktion.

Ziel: Eine Eigenschaft E für alle natürlichen Zahlen N beweisen.

Weg: – Induktionsanfang: Beweise E für $N = 1$
– Induktionsannahme: Nimm an, E gilt für alle $N \le n \in \mathbb{N}$.
– Induktionsschritt: Beweise E für $N = n^+$.
$\Rightarrow$ Dann gilt E für alle natürlichen Zahlen.

Beispiel: Beweis von $\sum_{k=1}^{n} k = n(n+1)/2$:
– Induktionsanfang: Gilt für $N = 1$ ✓
– Induktionsannahme: Gelte für $N \le n$
– Induktionsschritt: $\Rightarrow \sum_{k=1}^{n+1} k = \underbrace{\sum_{k=1}^{n} k}_{n(n+1)/2 \text{ laut Annahme}} + (n+1) = (n+2)(n+1)/2$ ✓

Auf natürlichen Zahlen kann man Addition und Multiplikation definieren.

Addition über $x + 1 = x^+$, $x + y^+ = (x + y)^+ \ \forall x, y \in \mathbb{N}$

Multiplikation über $x \cdot 1 = x$, $x \cdot y^+ = x \cdot y + x \ \forall x, y \in \mathbb{N}$

Eigenschaften (ohne Beweis)
– Addition $+$: Assoziativ, kommutativ
– Multiplikation $\cdot$: Assoziativ, kommutativ, Identität
$\Rightarrow$ $\mathbb{N}$ ist Halbgruppe bzgl Addition und Multiplikation.
– Zusätzlich gelten Distributivgesetze.

Häufig wird noch die Null hinzugenommen: $\mathbb{N}_0 = \mathbb{N} \cup \{0\} = \{0, 1, 2, 3, \cdots\}$
Dann hat die Addition auch eine Identität

Bemerkung: Die Griechen und Römer hatten keine Null. In China gab es sie schon seit dem 4. Jhdt. vor unserer Zeitrechnung. Wurde im 12. Jhdt. nach unserer Zeitrechnung von Arabern nach Europa gebracht.

Aufgaben zum Induktionsbeweis:
Beweisen Sie mittels vollständiger Induktion:

- $\sum_{k=1}^{n} k = n(n+1)/2$
- $\sum_{k=1}^{n} k^2 = n(n+1)(2n+1)/6$
- $\sum_{k=1}^{n} k^3 = (\sum_{k=1}^{n} k)^2$
- $\sum_{k=1}^{n} k\binom{n}{k} = n2^{n-1}$
- $n/2 < \sum_{k=1}^{2^n-1} \frac{1}{k} \le n$
- Berechnen Sie $\sum_{k=1}^{n}(2k-1)$ (erst erraten, dann mit vollständiger Induktion beweisen)

Diskutieren Sie folgende zwei "Beweise", die mit der Methode der vollständigen Induktion geführt wurden. Wo steckt der Wurm?

Beweis, dass alle Zahlen interessant sind:
Die Eins ist interessant (klar, z.B. Identität der Multiplikation).
Nimm an, alle Zahlen kleiner als n sind interessant. Dann ist auch $n+1$ interessant. Anderenfalls wäre es die erste nicht interessante Zahl, und das macht sie auch wieder interessant.

Überraschungstest: Ein Lehrer kündigt für die kommende Woche an, dass er an einem Tag einen Überraschungstest schreiben wird. Die Schüler stellen folgende Überlegung an.
– Am Freitag kann er den Test nicht mehr schreiben,
da wäre es ja dann kein Überraschungstest mehr.
– Am Donnerstag ist es deshalb auch unmöglich.
– Ebenso am Mittwoch, Dienstag, Montag.
– Folglich wird der Test nicht stattfinden.
(Aber: Dann schreibt der Lehrer dann am Mittwoch völlig überraschend einen Test ...)

1.2.3 Ganze Zahlen $\mathbb{Z}$

Ausgehend von den natürlichen Zahlen

Forderung: Gleichung $\boxed{a + x = b}$ soll für alle $a, b \in \mathbb{N}$ lösbar sein.
$\rightsquigarrow$ führt zu negativen Zahlen:
Zu jedem $a \in \mathbb{N}$ gehört ein $(-a)$ mit $a + (-a) = 0$.

(Beispiele: Mathematiker im Aufzug, Schulden, Vorwärts-/Rückwärtsgehen)

Gesamtmenge: Natürliche Zahlen, Null und negative Zahlen bilden zusammen die ganzen Zahlen $\mathbb{Z} = \mathbb{N} \cup \{-a : a \in \mathbb{N}\} \cup \{0\}$

Eigenschaften:

– In $\mathbb{Z}$ ist die Gleichung $a + x = b$ immer lösbar $(x = b + (-a))$

– $\mathbb{Z}$ bildet bzgl. der Addition eine abelsche Gruppe (a hat Inverses), bzgl. der Multiplikation nach wie vor eine Halbgruppe.

– Distributivgesetze gelten

$\Rightarrow$ $\mathbb{Z}$ ist ein Ring.

1.2.4 Rationale Zahlen $\mathbb{Q}$

Ausgehend von den ganzen Zahlen

Forderung: Gleichung $\boxed{a \cdot x = b}$ soll lösbar sein für alle $\boxed{a \neq 0}$.
$\rightsquigarrow$ führt zu rationalen Zahlen:
Zu jedem $a \in \mathbb{Z} \setminus \{0\}$ gehört ein Inverses a^{-1} mit $a \cdot a^{-1} = 1$.

- Erweiterung von $\mathbb{Z}$ um Kombinationen $a^{-1}b = b/a$
  $\Rightarrow a \cdot x = b$ wird lösbar.
- Bemerkung: De facto steht eine rationale Zahl p/q für eine ganze "Äquivalenzklasse" $[p/q]$ von äquivalenten Tupeln $(a, b) = b/a$ mit $b/a = p/q$. Beispielsweise sind 2/3 und 4/6 äquivalent, 2/3=4/6.

  Formale Schreibweise: p/q ist ein Repräsentant der Äquivalenzklasse, $[p/q]$ steht für die ganze Äquivalenzklasse.

Gesamtmenge: Rationale Zahlen $\mathbb{Q} = \{[p/q] : p \in \mathbb{Z}, q \in \mathbb{N}\}$

Eigenschaften:

- In $\mathbb{Q}$ ist die Gleichung $a \cdot x = b$ lösbar.
- $\mathbb{Q}$ ist eine abelsche Gruppe bzgl. $+$, $\mathbb{Q} \setminus \{0\}$ ist abelsche Gruppe bzgl. $\cdot$

$\Rightarrow$ $\mathbb{Q}$ ist ein Körper.

Beispiele: Praktisch gesehen alle Zahlen in der Physik
(Messung liefert immer Dezimalzahl, die irgendwo abbricht.)

1.2.5 Reelle Zahlen $\mathbb{R}$

Bei rationalen Zahlen fehlen immer noch wichtige Zahlen, z.B.

- Lösung von $x^2 = 2$
  und anderen Polynomgleichungen $p(x) = \sum_{k=0}^{n} a_k x^k = 0$

  Graphisch: $x^2 - 2 = 0$ sollte Lösung haben.
  (Kurve $y = x^2 - 2$ schneidet die Nullinie).

  entsprechende Erweiterung um Zahlen, bei denen in einem Plot $p(x)$ vs. x die Nullinie geschnitten wird (Nullstellen des Polynoms $p(x)$): algebraische Zahlen.

  (NB: Unvollständige Erweiterung, da z.B. $x^2 = -1$ nach wie vor keine Lösung hat. Das wird erst mit der Einführung der komplexen Zahlen (s.u.) behoben.)
- Kreiszahl π, Eulersche Zahl e u.a.: Transzendente Zahlen.

Zusammen: Reelle Zahlen $\mathbb{R}$.

Bemerkungen

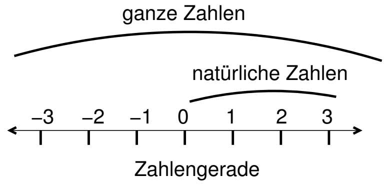
Die Erkenntnis, dass die Lösung von $x^2 = 2$, also $\sqrt{2}$ irrational ist, ist uralt. Das war schon den Pythagoräern bekannt, wurde aber unter den Tisch gekehrt, weil es nicht ins Weltbild passte.

Der Beweis geht so: Nimm an, $\sqrt{2} = p/q$ sei rational.
Dann folgt $\sqrt{2}q = p \Rightarrow 2q^2 = p^2$.
Primfaktorzerlegung von p und q: Links steht eine ungerade Zahl Faktoren, rechts eine gerade Zahl. Es gilt aber, dass die Primfaktorzerlegung eindeutig ist. (ohne Beweis: Folgt aus der Theorie der Ringe). Also kann $\sqrt{2}$ nicht rational sein.

Konstruktion der reellen Zahlen $\mathbb{R}$ aus den rationalen Zahlen $\mathbb{Q}$ basiert auf der Infinitesimalrechnung / braucht Begriff des Unendlichen.
(Vorgriff auf Begriff der Folgen): $\mathbb{R}$ enthält mit jeder konvergierenden Folge von rationalen Zahlen auch deren Grenzwert. Rationale Zahlen liegen dicht in $\mathbb{R}$, d.h. sie kommen beliebig nahe an jede reelle Zahl heran. (In jeder noch so kleinen "ϵ-Umgebung" einer reellen Zahlen liegt auch eine rationale Zahl: $\forall x \in \mathbb{R}, \epsilon \in \mathbb{R} : \exists y \in \mathbb{Q} : |x - y| < \epsilon$).

Anschaulich:

Eindeutige Abbildung zwischen Zahlen und Zahlengeraden



1.2.6 Komplexe Zahlen $\mathbb{C}$

Erweiterung so, dass alle Polynomgleichungen lösbar sind, also z.B. auch die Gleichung $x^2 + 1 =$. Dafür muss man nur *eine* neue Zahl $i = \sqrt{-1}$ einführen. Allerdings muss man das Bild der Zahlengeraden aufgeben/erweitern.

1.2.7 Zusammenfassung

<table>
<tr><th rowspan="2"></th><th rowspan="2">Symbol</th><th rowspan="2">lösbare Gleichungen</th><th rowspan="2">zusätzliche Zahlen</th><th colspan="2">Rechenregeln</th><th rowspan="2">Abbildung</th></tr>
<tr><th>Addition</th><th>Multiplikation</th></tr>
<tr><td rowspan="2">Natürliche Zahlen</td><td>$\mathbb{N}$</td><td>$a + b = x$
$a \cdot b = x$</td><td></td><td>assoziativ
kommutativ</td><td>assoziativ
kommutativ
Identität</td><td rowspan="2">Gitterpunkte auf Halbgeraden</td></tr>
<tr><td>$\mathbb{N}_0$</td><td></td><td>Null</td><td>Identität</td><td></td></tr>
<tr><td>Ganze Zahlen</td><td>$\mathbb{Z}$</td><td>$a + x = b$</td><td>Differenzen
$a - b$</td><td>abelsche Gruppe</td><td></td><td rowspan="2">Gitterpunkte auf Geraden</td></tr>
<tr><td></td><td></td><td></td><td></td><td colspan="2">Ring</td></tr>
<tr><td>Rationale Zahlen</td><td>$\mathbb{Q}$</td><td>$a \cdot x = b$
$(a \neq 0)$</td><td>Brüche
b/a</td><td>abelsche Gruppe</td><td>abelsche Gruppe</td><td rowspan="2">Dicht auf Geraden</td></tr>
<tr><td></td><td></td><td></td><td></td><td colspan="2">Körper</td></tr>
<tr><td>Reelle Zahlen</td><td>$\mathbb{R}$</td><td>$x^2 = a$
$(a \geq 0)$</td><td>Irrationale Zahlen</td><td colspan="2">Körper</td><td>Vollständige Gerade</td></tr>
<tr><td>Komplexe Zahlen</td><td>$\mathbb{C}$</td><td>$x^2 = a$
$(a < 0)$</td><td>Imaginäre Zahl</td><td colspan="2">Körper</td><td>Zahlen-Ebene</td></tr>
</table>

Aufgaben zum Abschluss:

Was ist an folgenden "Beweisen" falsch?

1=0 , denn: Betrachte $x = 1$

$$\Rightarrow x^2 = x$$
$$\Rightarrow x^2 - x = x - 1$$
$$\Rightarrow (x-1)(x+1) = x - 1$$
$$\Rightarrow x + 1 = 1$$
$$\Rightarrow x = 0 \quad \checkmark?.$$

4=5 , denn: Es gilt $16 - 36 = 25 - 45$

$$\Rightarrow 16 - 36 + 81/4 = 25 - 45 + 81/4$$
$$\Rightarrow 42 - 2 \cdot 4 \cdot \tfrac{9}{2} + (\tfrac{9}{2})^2 = 5^2 - 2 \cdot 5 \cdot \tfrac{9}{2} + (\tfrac{9}{2})^2$$
$$\Rightarrow (4 - \tfrac{9}{2})^2 = (5 - \tfrac{9}{2})^2$$
$$\Rightarrow 4 - \tfrac{9}{2} = 5 - \tfrac{9}{2}$$
$$\Rightarrow 4 = 5 \quad \checkmark?.$$

1.3 Elementare Rechenoperationen

(Für weitere Erläuterungen, Beispiele, und Übungsaufgaben siehe z.B. E. Cramer, J. Neslehova, Vorkurs Mathematik)

1.3.1 Bruchrechnung

Erinnerung an 1.2.4: Brüche lassen sich als Quotienten darstellen, z.B. $\frac{2}{4}$, $\frac{xy}{1+x}$, $\frac{\mathrm{cm}^2}{\mathrm{s}}$

Eigenart von Brüchen: Zwei Brüche können für dasselbe Objekt stehen (dieselbe rationale Zahl, derselbe algebraische Ausdruck, dieselbe Einheit), zum Beispiel $\frac{1}{2} = \frac{2}{4}$, $\frac{xy^2}{xy} = \frac{y}{x}$, $\frac{\mathrm{m}^2}{\mathrm{m\,s}} = \frac{\mathrm{m}}{\mathrm{s}}$

Formal: Sie gehören zur selben "Äquivalenzklasse", siehe 1.2.4.

Grundlage ist der **Satz der eindeutigen Primfaktorzerlegung**

(i) Für jede natürliche Zahl gibt es eine (bis auf die Reihenfolge) eindeutige Zerlegung in ein Produkt aus Primzahlen.

(Hier ist eine Primzahl eine Zahl, die nur durch sich selbst oder 1 so teilbar ist, dass der Quotient wieder eine natürliche Zahl ist).

(ii) Eine entsprechende Aussage gilt auch für Polynome. Ein Polynom $P(x) = a_0 + a_1 x + a_2 x^2 + \cdots a_n x^n$ mit Koeffizienten $a_i \in \mathbb{R}$ lässt sich in eindeutiger Weise in Primpolynome $P_i(x)$ zerlegen $P(x) = C \prod_i P_i(x)$ mit Primpolynomen $P_i(x)$, die selber nicht weiter zerlegbar sind und den führenden Koeffizienten $a_n = 1$ haben.

(De facto haben Primpolynome maximal den Grad 2).

(iii) Allgemein gilt der Satz der eindeutigen Primfaktorzerlegung in allen sogenannten "Hauptidealringen", aber das führt hier zu weit.

Wegen dieses Satzes lässt sich jeder Bruch a/b zu einem **eindeutig bestimmten** Bruch p/q kürzen, bei dem p und q keine gemeinsamen Primfaktoren haben. Die Äquivalenzklasse $[p/q]$ ist daher "wohldefiniert".

Rechnen mit Brüchen:

- Erweitern und Kürzen ist oft hilfreich: z.B. $\frac{a}{b} = \frac{ka}{kb}$
- Multiplikation: $\frac{a}{b} \cdot \frac{c}{d} = \frac{a \cdot c}{b \cdot d}$
- Division: $\frac{a}{b} \Big/ \frac{c}{d} = \frac{a \cdot d}{b \cdot c}$
- Addieren, Subtrahieren $\frac{a}{b} \pm \frac{c}{d} =?$
  - Gleicher Nenner $b = d$: $\frac{a}{b} \pm \frac{c}{b} = \frac{a \pm c}{d}$
  - Verschiedener Nenner $b \neq d$:
    Gleiche Nenner an durch Erweitern, dann wie oben.
    * Falls b und d teilerfremd sind: $\frac{a}{b} \pm \frac{c}{d} = \frac{ad}{bd} \pm \frac{bc}{bd} = \frac{ad \pm bc}{bd}$
    * Anderenfalls erweitere nur bis zum kleinsten gemeinsamen Vielfachen, um Rechenaufwand zu sparen.

Anmerkung: Der Satz der eindeutigen Primfaktorzerlegung legt folgendes systematisches Vorgehen bei komplizierten Brüchen a/b nahe:
$\longrightarrow$ Zerlege a und b in Primfaktoren
$\longrightarrow$ Kürze und vereinfache soweit möglich

Beispiele: $\frac{30}{18} = \frac{2\cdot 3\cdot 5}{2\cdot 3\cdot 3} = \frac{5}{3}$, $\quad \frac{2x-2}{x^2-1} = \frac{2(x-1)}{(x+1)(x-1)} = \frac{2}{x+1}$

(In der Praxis kürzt man oft besser erst mal drauf los, um den Bruch zu vereinfachen, und greift dann erst auf aufwändigere systematischere Verfahren zurück.)

1.3.2 Potenzrechnung

Ausgangspunkt der Potenzrechnung ist eine Operation, bei der man eine Multiplikation mit einer Zahl / einem Term mehrfach hintereinander ausführt: $a^n = \underbrace{a \cdot a \cdot a \cdots a}_{n \text{ Mal}}$ mit a: Basis, n: Exponent

Beispiele $2^3 = 8, x^5, \text{cm}^2$

$\rightarrow$ Die Basis kann eine reelle Zahl sein $a \in \mathbb{R}$, ein mathematischer Ausdruck, oder ein Einheit. Der Exponent ist in dieser Definition zunächst eine natürliche Zahl, $n \in \mathbb{N}$. Die Definition wird aber unten sukzessive verallgemeinert werden.

Rechenregeln ergeben sich aus der Definition der Potenzen

- $a^n\, a^m = a^{n+m}$
- $a^n/a^m = a^{n-m}$ (* falls $n > m$)
- $(a \cdot b)^n = a^n \cdot b^n$
- $(a^n)^m = a^{n \cdot m}$
  Beachte: $(a^n)^m \neq a^{n^m}$, z.B. $(2^3)^2 = 8^2 = 64 \neq 2^{3^2} = 2^9 = 512$
- $a^{n/m} = \sqrt[m]{a^n}$ für $a > 0$ (* falls $n/m \in \mathbb{N}$)

Die Einschränkungen (*) werden wir später fallen lassen.

Erweiterungen

(i) Negative Exponenten : Lasse in a^n alle $n \in \mathbb{Z}$ zu.
Dafür Einschränkung an a: $a \neq 0$
$\rightarrow$ Die Regel $a^n/a^m = a^{n-m}$ gilt ohne Einschränkungen
$\rightarrow$ $a^0 = 1$ (wegen $a^0 = a^{n-n} = a^n/a^n = 1$)
$\rightarrow$ $a^{-n} = 1/a^n$ (wegen $a^{-n} = a^{0-n} = a^0/a^n = 1/a^n$)
NB: Die neue Einschränkung $a \neq 0$ ist nötig, da man nicht durch Null teilen darf.

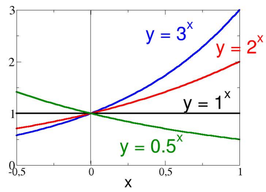
(ii) Rationale Exponenten : Lasse in a^x alle $x \in \mathbb{Q}$ zu.
Dafür Einschränkung an a: $a > 0$, falls $a \in \mathbb{R}$ (entspricht bei mathematischen Termen einer Einschränkung ihres Definitionsbereichs.)

→ Regel $a^{n/m} = \sqrt[m]{a^n}$ (mit $m \in \mathbb{N}$) gilt ohne Einschränkungen
→ $a^{1/n} = \sqrt[n]{a}$
→ $a^{n/m} = \sqrt[m]{a^n} = (\sqrt[m]{a})^n$
NB: Die neue Einschränkung $a > 0$ ist nötig, da man in $\mathbb{R}$ nicht die Wurzel einer negativen Zahl nehmen darf. (In $\mathbb{C}$ entfällt die Einschränkung.)

(iii) Reelle Exponenten : Lasse in a^x alle $x \in \mathbb{R}$ zu.
Schwierigster Schritt. Um diese Erweiterung durchzuführen, müssen Methoden der Infinitesimalrechnung eingesetzt werden. Reelle Zahlen lassen sich beliebig genau durch rationale Zahlen annähern (siehe 1.2.5). Definiert man eine reelle Zahl x als Grenzwert $\lim_{n\to\infty} x_n$ einer Folge von rationalen Zahlen x_n (zum Begriff der Folge siehe 4.2.1), dann ist a^x der Grenzwert der Folge a^{x_n}. Dieser Grenzwert existiert.

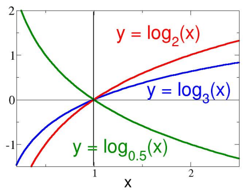
Umkehrung und Logarithmus

Wir betrachten die Gleichung $y = a^x$ und konkret den Verlauf von a^x für verschiedene Werte von $a > 0$



→ Für $a \neq 1$ deckt a^x den Wertebereich aller positiven reellen Zahlen ab ($y \in \mathbb{R} \& b > 0$) und ist eindeutig, d.h. es gibt nur ein $x \in \mathbb{R}$ mit $a^x = y$.

→ Erlaubt die Definition einer Umkehrfunktion $x = \log_a(y)$ (Logarithmus von b zur Basis a)



Es gibt einen eindeutigen Zusammenhang
$a^x = b \leftrightarrow x = \log_a(b)$

Rechenregeln zum Logarithmus

- $x = a^{\log_a(x)}$ (bilde $\log_a$ auf beiden Seiten)
- $\log_a(a) = 1$, $\log_a(1) = 0$ (direkte Folgerung daraus)
- $\log_a(x \cdot y) = \log_a(x) + \log_a(y)$ (folgt aus $a^x \cdot a^y = a^{x+y}$)
- $\log_a(x^m) = m \log_a(x)$ (folgt aus $(a^n)^m \cdot a^{n \cdot m}$)
- $\log_c(b) = \log_a(b)/\log_a(c)$ für, $a, c \neq 1$

Wichtige spezielle Basen:
– dekadischer Logarithmus: $\log_{10}(b) =: \lg(b)$
– dyadischer Logarithmus: $\log_2(b) =: \mathrm{ld}(b)$
– natürlicher Logarithmus: $\log_e(b) =: \ln(b)$
mit $e = 2.718...$ (Eulerzahl), siehe 3.1.3

Umrechnungsformel: $\log_c(b) = \log_a(b)/\log_a(c)$ für, $a, c \neq 1$

Aufgabe: Beweisen Sie diese Gleichung.

1.4 Gleichungen und Ungleichungen

Allgemeine Fragestellung:

Gegeben sei eine Gleichung oder eine Ungleichung, in der eine unbekannten Variable x vorkommt.

Gesucht ist die Menge $\mathbb{L}$ aller x, die diese Gleichung erfüllen.

Verfahren: Umformung der Gleichung, bis $\mathbb{L}$ abgelesen werden kann.

Bemerkung: Als Naturwissenschaftler steht einem oft eine einfache Möglichkeit zur Verfügung, abzuchecken, ob eine abgeleitete Gleichung oder Ungleichung so noch stimmen kann: Die Einheiten müssen konsistent sein. Wenn das nicht der Fall ist, dann hat man sich auf jeden Fall irgendwo verrechnet. Dies gilt für alle Einheiten ausser den "Winkeleinheiten" rd und srd, die in diesem Sinne keine "echten" physikalischen Einheiten sind, sondern eher Konventionen in mathematischen Notationen.

Beispiele: $v_i, \bar{v}, \hat{v}$ seien Geschwindigkeiten.

$\bar{v} = \frac{1}{2}(v_1 + v_2)$ ist konsistent, $\hat{v} = \frac{2v_1 v_2}{v_1+v_2}$ ebenfalls.

($\bar{v}$ beschreibt die mittlere Geschwindigkeit, wenn man die Hälfte der Zeit mit v_1 unterwegs ist und die andere Hälfte mit v_2. $\hat{v}$ beschreibt die mittlere Geschwindigkeit, wenn man die Hälfte der Strecke mit der Geschwindigkeit v_1 zurücklegt und die andere Hälfte mit der Geschwindigkeit v_2.)

$v_0 > \frac{\sqrt{v_1 v_2}}{v_1+v_2}$ ist nicht konsistent.

(Die linke Seite hat die Einheit m/s, die rechte Seite hat keine Einheit - ist also dimensionslos. Die Ungleichung kann nicht korrekt sein!)

$v = v_0\, 2^{\hat{v}}$ ist nicht konsistent.

(Die Gleichung sieht zwar auf den ersten Blick ganz gut aus, aber der Exponent $\hat{v}$ hat eine Einheit und das ist ein absolutes Nogo! Exponenten müssen immer dimensionslos sein!)

1.4.1 Lineare Gleichungen und Ungleichungen

In linearen (Un)gleichungen geht x höchstens linear ein.

Lineare Gleichungen: Können in die Form $ax + b = 0$ gebracht werden.

Fälle:	$a \neq 0$	$\Rightarrow$ $x = -b/a$	$\mathbb{L} = \{b/a\}$
	$a = 0, b \neq 0$	$\Rightarrow$ Keine Lösung	$\mathbb{L} = \emptyset$
	$a = 0, b = 0$	$\Rightarrow$ Gleichung ist ür alle $x \in \mathbb{R}$ erfüllt!	$\mathbb{L} = \mathbb{R}$

Lineare Ungleichungen: können in die Form $ax + b > 0$, $ax + b \geq 0$, $ax + b < 0$ oder $ax + b \leq 0$ gebracht werden.

Lösungsmengen der Gleichung $ax + b > 0$:

Fälle:	$a > 0$	$\Rightarrow$ $x > -b/a$	$\mathbb{L} = \{x \in \mathbb{R} : x > -b/a\}$
	$a < 0$	$\Rightarrow$ $x < -b/a$	$\mathbb{L} = \{x \in \mathbb{R} : x < -b/a\}$
	$a = 0, b > 0$	$\Rightarrow$ Für alle $x \in \mathbb{R}$ erfüllt.	$\mathbb{L} = \mathbb{R}$
	$a = 0, b \leq 0$	$\Rightarrow$ Keine Lösung	$\mathbb{L} = \emptyset$

Lösungsmengen der Gleichung $ax + b \geq 0$:

Fälle:	$a > 0$	$\Rightarrow$	$x > -b/a$	$\mathbb{L} = \{x \in \mathbb{R} : x > -b/a\}$
	$a < 0$	$\Rightarrow$	$x < -b/a$	$\mathbb{L} = \{x \in \mathbb{R} : x < -b/a\}$
	$a = 0, b \geq 0$	$\Rightarrow$	Für alle $x \in \mathbb{R}$ erfüllt.	$\mathbb{L} = \mathbb{R}$
	$a = 0, b < 0$	$\Rightarrow$	Keine Lösung	$\mathbb{L} = \varnothing$

Die Ungleichungen $ax+b < 0$ und $ax+b \leq 0$ werden völlig analog behandelt.

1.4.2 Quadratische Gleichungen und Ungleichungen

In linearen (Un)gleichungen geht x höchstens quadratisch ein. Gleichungen lassen sich in die Form $ax^2 + bx + c = 0$ bringen, Ungleichungen in die Form $ax^2 + bx + c > 0$, $ax^2 + bx + c \geq 0$, $ax^2 + bx + c < 0$ bzw. $ax^2 + bx + c \leq 0$.

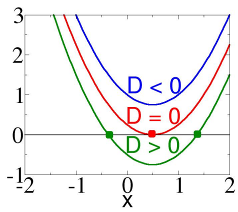
Wir nehmen im Folgenden an, dass $a \neq 0$. Anderenfalls wäre die (Un)gleichung de facto eine lineare Gleichung und diese wurden in 1.4.1 behandelt.

Quadratische Gleichungen: $ax^2 + bx + c = 0$ mit $a \neq 0$

→ Bringe sie in Normalform: $x^2 + px + q = 0$

→ Anwendung der "pq-Formel": Zwei Lösungen $x_{1,2} = -\frac{p}{2} \pm \sqrt{(\frac{p}{2})^2 - q}$
Aber: Das setzt voraus, dass die Diskriminante,
$D := (\frac{p}{2})^2 - q)$, positiv ist:

→ Fallunterscheidung:
$D > 0 \Rightarrow \mathbb{L} = \{-\frac{p}{2} + \sqrt{D}, -\frac{p}{2} - \sqrt{D}\}$ (Zwei Lösungen)
$D = 0 \Rightarrow \mathbb{L} = \{-\frac{p}{2}\}$ (Eine Lösung)
$D < 0 \Rightarrow \mathbb{L} = \varnothing$ (Keine Lösung)



Anmerkung:

– Falls $D > 0$, lässt sich $P(x) = x^2 + px + q$ zerlegen
als $P(x) = (x - x_1)(x - x_2) = (x + \frac{p}{2} + \sqrt{D})(x + \frac{p}{2} - \sqrt{D})$

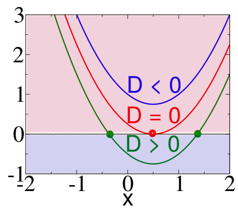
– Falls $D = 0$, gilt $P(x) = (x + \frac{p}{2})^2$ (doppelte Nullstelle)

– Falls $D < 0$, läst sich $P(x)$ nicht zerlegen: Es ist ein Primpolynom.

Quadratische Ungleichungen, z.B. $ax^2 + bx > 0$ mit $a \neq 0$

→ Bringe sie wieder in Normalform:
$a > 0: \quad x^2 + px + q > 0$
$a < 0: \quad x^2 + px + q < 0$

$\rightsquigarrow$ Lösungsmenge kann graphisch abgelesen werden
Für $x^2 + px + q > 0$ (rot schattierter Bereich)
$D > 0 \Rightarrow \mathbb{L} = \mathbb{R} \setminus [-\frac{p}{2} - \sqrt{D} : -\frac{p}{2} + \sqrt{D}]$
$D = 0 \Rightarrow \mathbb{L} = \mathbb{R} \setminus \{-\frac{p}{2}\}$
$D < 0 \Rightarrow \mathbb{L} = \mathbb{R}$



Für $x^2 + px + q < 0$ (blau schattierter Bereich)
$D > 0 \Rightarrow \mathbb{L} =]-\frac{p}{2} - \sqrt{D} : -\frac{p}{2} + \sqrt{D}[$
$D \leq 0 \Rightarrow \mathbb{L} = \varnothing$

1.4.3 Polynomgleichungen höherer Ordnung

Wir betrachten nun allgemein Polynomgleichungen vom Grad n in der Standardform $x^n + c_{n-1}x^{n-1} + c_{n-2}x^{n-2} + \cdots + c_0 = 0$ mit $c_i \in \mathbb{R}$

Die Lösungen für Polynomgleichungen vom Grad 1 und 2 wurden in 1.4.1 und 1.4.2 angegeben.

Für die allgemeinen Lösungen von Polynomgleichungen vom Grad 3 und 4 gibt es ebenfalls Ausdrücke ähnlich den *pq*-Formeln, nur viel komplizierter (hier nicht gezeigt).

Für Polynomgleichungen vom Grad 5 oder höher kann es prinzipiell keine derartigen allgemeinen Ausdrücke geben – d.h. allgemeine Lösung mit Hilfe von Wurzelausdrücken (Satz von Abel-Ruffini). Dies ist eine Folgerung aus der Galoistheorie von Évariste Galois.

Man kann die Lösungen aber immer noch numerisch finden.

Eine Polynomgleichung $n$ten Grades kann maximal n Lösungen haben. Wenn es n Lösungen $x_1, \cdots, x_n$ gibt, dann gilt die Zerlegung $x^n + c_{n-1}x^{n-1} + c_{n-2}x^{n-2} + \cdots + c_0 = (x - x_1)(x - x_2)\cdots(x - x_n)$

1.4.4 Kompliziertere Gleichungen

Im Folgenden sollen anhand von Beispielen einige weitere Strategien zum Lösen anderer Gleichungstypen erläutert werden. Es muss jedoch darauf hingewiesen werden, dass die meisten Gleichungen nicht nach x aufgelöst werden können.

Gleichungen mit Brüchen

Beispiel: $\frac{x-2}{x^2-4} = 0$

- Bruch durch Kürzen vereinfachen (wg. $(x^2 - 4) = (x - 2)(x + 2)$)
  $\Rightarrow\ 0 = \frac{1}{x+2}$
- Mit $(x + 2)$ multiplizieren
  $\Rightarrow\ 0 = 1$: Nie erfüllt!

$\Rightarrow\ \mathbb{L} = \varnothing$

Bemerkung: Wenn man die Kürzungsmöglichkeit nicht gesehen und als ersten Schritt direkt mit $(x^2 - 4)$ multipliziert hätte, dann hätte man die (falsche) Lösung $x = 2$ erhalten. Um solche Fehler auszuschließen, muss man unbedingt die Lösung hinterher einsetzen und testen. Spätestens dann hätte man gemerkt, dass sie Probleme bringt, weil man unerlaubterweise durch Null teilt.

Beispiel: $\frac{x-2}{x+1} + \frac{x}{x-1} = 1 + \frac{2x}{x^2-1}$

- Alle Terme auf eine Seite bringen
  $\Leftrightarrow\ 0 = \frac{x-2}{x+1} + \frac{x}{x-1} - 1 - \frac{2x}{x^2-1}$
- Alles auf einen Nenner bringen: $(x^2 - 1 = (x - 1)(x + 1))$.
  $\Leftrightarrow\ 0 = \frac{((x-2)(x-1)+x(x+1)-(x^2-1)-2x}{x^2-1} \overset{\text{Vereinfachen}}{=} \frac{x^2-4x+3}{x^2-1}$

- Mit $(x^2 - 1)$ multiplizieren.
  $\Rightarrow\ 0 = x^2 - 4x + 3$
- pq-Formel anwenden
  $\Leftrightarrow\ x_{1,2} = 2 \pm \sqrt{2^2 - 3} = 2 \pm 1$
- Check durch Einsetzen
  $x_1 = 3$: Passt
  $x_1 = 1$: Verboten, da (z.B. in $\frac{x}{x+1}$) durch Null dividiert wird

$\Rightarrow\ \mathbb{L} = \{3\}$

Gleichungen mit Wurzeln

Beispiel: $\sqrt{x+8} - x = 2$

- Wurzel auf eine Seite bringen
  $\Leftrightarrow\ \sqrt{x+8} = x + 2$
- Quadrieren
  $\Rightarrow\ x + 8 = (x+2)^2 \overset{\text{Ausmultiplizieren}}{=} x^2 + 4x + 4$
- Alle Terme auf eine Seite bringen
  $\Leftrightarrow\ x^2 + 3x - 4 = 0$
- pq-Formel anwenden
  $\Leftrightarrow\ x_{1,2} = -\frac{3}{2} \pm \sqrt{(\frac{3}{2})^2 + 4} \overset{\text{Vereinfachen}}{=} -\frac{3}{2} \pm \frac{5}{2}$
- Check durch Einsetzen
  $x_1 = 1$: $\sqrt{1+8} - 1 = 2$ ✓
  $x_2 = -4$: $\sqrt{-4+8} - (-4) = 6 \neq 2$: Passt nicht!
  Grund: Das Quadrieren in Schritt 2 ist keine Äquivalenzumformung. Die quadrierte Gleichung hat mehr Lösungen als die Originalgleichung.

$\Rightarrow\ \mathbb{L} = \{1\}$

Beispiel: $\sqrt{(x-2)^2} = 2$

- Quadrieren
  $\Rightarrow\ (x-2)^2 = 4$
- Alles auf eine Seite bringen und vereinfachen
  $\Leftrightarrow\ x^2 - 4x = 0$
- pq-Formel oder scharf hinschauen
  $\Leftrightarrow\ x_1 = 0,\ x_2 = 4$
- Check durch Einsetzen
  x_1 passt, x_2 passt ✓

$\Rightarrow\ \mathbb{L} = \{0, 4\}$

Bemerkung: Naiverweise hätte man als Lösungsweg auch ganz zu Anfang die Umformung $\sqrt{(x-2)^2} = x - 2$ vornehmen können und die Gleichung $x - 2 = 2$ erhalten. Die hätte jedoch nur die Lösung $x = 4$ geliefert. Die Umformung $\sqrt{(x-2)^2} = x - 2$ reduziert die Zahl der Lösungen!

Gleichungen mit Potenzen und Logarithmen

Beispiel: $3^{x-1} = 10$

- Logarithmus zur Basis 3 nehmen
  $\Leftrightarrow\ x - 1 = \log_3(10)$
  $\Leftrightarrow\ x = 1 + \log_3(10)$

- Eventuell (bei Wunsch) auf eine gängigere Basis umrechnen, z.B. den dekadischen Logarithmus
  $\Leftrightarrow\ x = 1 + \lg(10)/\lg(3) = 1 + 1/\lg(3) = 3.0959$
- Check durch Einsetzen: Passst

$\Rightarrow\ \mathbb{L} = \{3.0959\}$

Beispiel: $\lg(x) - \frac{1}{2}\lg(3x-2) = 0$

- Logarithmen-Terme zusammenfassen, $\lg(1) = 0$ verwenden
  $\Leftrightarrow\ \lg(x) - \lg(\sqrt{3x-2}) = 0$
  $\Leftrightarrow\ \lg(x/\sqrt{3x-2}) = \lg(1)$
- Ausnutzen, dass $\lg(a) = \lg(b) \Leftrightarrow a = b$ für $a, b > 0$
  $\Leftrightarrow\ x/\sqrt{3x-2} = 1$
- Weitere Umformungen
  $\Rightarrow\ x = \sqrt{3x-2}$
  $\Rightarrow\ x^2 = 3x - 2$
  $\Leftrightarrow\ x^2 - 3x + 2 = 0$
- pq-Formel
  $\Leftrightarrow\ x_{1,2} = \frac{3}{2} \pm \sqrt{(\frac{3}{2})^2 - 4} = \frac{3}{2} \pm \frac{1}{2} \Leftrightarrow x_1 = 2,\ x_2 = 1$
- Check durch Einsetzen: Beide Lösungen passen.

$\Rightarrow\ \mathbb{L} = \{1, 2\}$

Beispiel: $\lg(x^2 - 1) = \lg(2x^2)$

- Ausnutzen, dass $\lg(a) = \lg(b) \Leftrightarrow a = b$ für $a, b > 0$
  $\Leftrightarrow\ x^2 - 1 = 2x^2$
  $\Leftrightarrow\ x^2 = -1$: Nicht lösbar.

$\Rightarrow\ \mathbb{L} = \varnothing$

Kapitel 2

Elemente der linearen Algebra

Lineare Algebra: Teilgebiet der Mathematik, das in den Naturwissenschaften eine zentrale Rolle spielt.

Ursprung:

- Einerseits Entwicklung von Strategien zum Lösen von linearen Gleichungssystemen
- Andererseits Beschreibung von bestimmten geometrischen Objekten im Raum (Geraden, Ebenen, "Hyperebenen")

Bedeutung:

- Grundlegend physikalische Größen (z.B. Impuls, Kraft) und Theorien (z.B. Quantenmechanik) haben oft Strukturen, die von der linearen Algebra bereitgestellt werden.
- Näherungsverfahren zur numerischen Berechnung basieren oft auf Verfahren aus der linearen Algebra

2.1 Lineare Gleichungssysteme

Das Lösen linearer Gleichungssysteme ist ein klassisches Problem der linearen Algebra. Die Allgemeine Form eines linearen Gleichungssystems mit n Gleichungen und m Unbekannten x_i ist:

$$\begin{aligned}
a_{11}\,x_1 + a_{12}\,x_2 + \cdots + a_{1m}\,x_m &= b_1 \\
a_{21}\,x_1 + a_{22}\,x_2 + \cdots + a_{2m}\,x_m &= b_2 \\
\vdots \\
a_{n1}\,x_1 + a_{n2}\,x_2 + \cdots + a_{nm}\,x_m &= b_n
\end{aligned}$$

mit Koeffizienten $a_{ij} \in \mathbb{R}$, $b_j \in \mathbb{R}$.

Vorab:

- Wenn $n < m$ (weniger Gleichungen als Unbekannte), kann es keine eindeutige Lösung geben. Das Gleichungssystem ist unterbestimmt.
- Umgekehrt: Wenn $n > m$ (mehr Gleichungen als Unbekannte), ist das Gleichungssystem "in der Regel" überbestimmt und es gibt keine Lösung.

Es kann aber sein, dass manche der Gleichungen de facto Kombinationen von anderen Gleichungen sind (sie sind "linear abhängig"). Dann kann es doch eine Lösung geben, oder ggf. sogar eine Schar Lösungen.

Wir betrachten zuerst den Fall $n = m = 2$

2.1.1 Zwei Gleichungen und zwei Unbekannte

Beispiele zum Einstieg

⋆ Gleichungssystem: (i) $5x + 3y = -3$
(ii) $x - y = 1$

⇝ Lösungsschritte:
- Zeile $(\text{ii})_{\text{neu}} = 5\ (\text{ii})_{\text{alt}}$: $5x - 5y = 5$
- Zeile $(\text{ii})_{\text{neu}} = (\text{i})_{\text{alt}} - (\text{ii})_{\text{alt}}$: $8y = -8$
- Zeile $(\text{ii})_{\text{neu}} = (\text{ii})_{\text{alt}}/8$: $y = -1$
- Zeile $(\text{ii})_{\text{neu}}$ einsetzen in (i) : $5x - 3 \cdot (-1) = -3$
- Auflösen nach x : $x = 0$

⇒ Die Lösungsmenge für (x, y) ist $\mathbb{L} = \{(0, -1)\}$

⋆ Gleichungssystem: (i) $5x - 5y = -3$
(ii) $x - y = 1$

⇝ Lösungsschritte:
- Zeile $(\text{ii})_{\text{neu}} = 5\ (\text{ii})_{\text{alt}}$: $5x - 5y = 5$
- Zeile $(\text{ii})_{\text{neu}} = (\text{i})_{\text{alt}} - (\text{ii})_{\text{alt}}$: $0 = -8$ Nie erfüllt!

⇒ Die Lösungsmenge für (x, y) ist $\mathbb{L} = \emptyset$

⋆ Gleichungssystem: (i) $5x - 5y = 5$
(ii) $x - y = 1$

⇝ Lösungsschritte:
- Zeile $(\text{ii})_{\text{neu}} = 5\ (\text{ii})_{\text{alt}}$: $5x - 5y = 5$
- Zeile $(\text{ii})_{\text{neu}} = (\text{i})_{\text{alt}} - (\text{ii})_{\text{alt}}$: $0 = 0$ Immer erfüllt!
- (i) auflösen nach x : $x = 1 - y$

⇒ Man erhält eine unendliche Schar von Lösungen
$\mathbb{L} = \{(x, y) : x = 1 - y\}$

Fazit: Je nach Struktur des Gleichungssystems kann es entweder eine eindeutige Lösung geben, unendlich viele Lösungen, oder gar keine.

Im Folgenden soll systematischer untersucht werden, wann welcher Fall eintritt.

Allgemeines Gleichungssystem

Wir betrachten das Gleichungssystem

(i) $a_{11}\, x_1 + a_{12}\, x_2 = b_1$
(ii) $a_{21}\, x_1 + a_{22}\, x_2 = b_2$

Um die Notation zu vereinfachen, definieren wir $D := a_{11}a_{22} - a_{12}a_{21}$.

Wir wenden die gleichen Schritte an wie oben, (i) und (ii) zu lösen.

⇝ Lösungsschritte, zunächst für $a_{21} \neq 0$:
- Zeile $(\text{ii})_{\text{neu}} = \frac{a_{11}}{a_{21}}\ (\text{ii})_{\text{alt}}$: $a_{11}\, x_1 + \frac{a_{11}a_{22}}{a_{21}}\, x_2 = \frac{a_{11}}{a_{21}}\, b_2$
- Zeile $(\text{ii})_{\text{neu}} = (\text{i})_{\text{alt}} - (\text{ii})_{\text{alt}}$: $(a_{12} - \frac{a_{11}a_{22}}{a_{21}})\, x_2 = b_1 - \frac{a_{11}}{a_{21}}\, b_2$
- Zeile $(\text{ii})_{\text{neu}} = a_{21}\ (\text{ii})_{\text{alt}}$: $(a_{12}a_{21} - a_{11}a_{22})\, x_2 = a_{21}\, b_1 - a_{11}\, b_2$

$$-D\, x_2 = a_{21}\, b_1 - a_{11}\, b_2$$

Letztere Gleichung gilt auch für $a_{21} = 0$ (d.h. $D = a_{11}a_{22}$) und folgt in diesem Fall direkt aus (ii) nach Multiplikation mit a_{11}.
Völlig analog kann man herleiten (wenn man x_1 isoliert):

– Lösungschritte von oben mit (i),(ii) vertauscht: $-D\,x_1 = -a_{22}\,b_1 + a_{12}\,b_2$

$\rightsquigarrow$ Weiter mit Fallunterscheidung $D = 0$ und $D \neq 0$

Fall $D = 0$

$\Rightarrow$ Keine Lösung, falls $a_{21}b_1 - a_{11}b_2 \neq 0$ oder $a_{22}b_1 - a_{12}b_2 \neq 0$.
Anderenfalls unendlich viele Lösungen

Fall $D \neq 0$

– Zeile $(\text{ii})_{\text{neu}}$ = - $(\text{ii})_{\text{alt}}/D$: $x_2 = \frac{1}{D}(-a_{21}\,b_1 + a_{11}\,b_2)$
– Zeile $(\text{ii})_{\text{neu}}$ einsetzen in (i) : $a_{11}\,x_1 + (-a_{21}\,b_1 + a_{11}\,b_2)\frac{a_{12}}{D} = b_1$
– Auflösen nach x_1 : $x_1 = \frac{1}{D}(a_{22}\,b_1 - a_{12}\,b_2)$

Fazit: Die Lösungsmenge hängt von $D = a_{11}a_{12} - a_{12}a_{21}$ ab.
Für $D \neq 0$ ist $\mathbb{L} = \{\big(\frac{1}{D}(a_{22}\,b_1 - a_{12}b_2), \frac{1}{D}(-a_{21}\,b_1 + a_{11}\,b_2)\big)\}$.
Für $D = 0$ ist $\mathbb{L} = \varnothing$ oder umfasst eine unendliche Menge.

2.1.2 Matrixdarstellung von linearen Gleichungssystemen

Wie man in 2.1.1 sieht, kann bereits in zwei Dimensionen der Umgang mit linearen Gleichungssystemen recht unübersichtlich werden. Einen Ausweg bietet die Matrixschreibweise.

Kurzer Einschub Matrizen (Vorgriff auf 2.3)

Beispiel: 2×2-Matrizen

Gegeben "Matrix" $\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ und $\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$, $\mathbf{w} = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}$,

- Anwendung von Matrix auf $\mathbf{v}$, $\mathbf{w}$: $\mathbf{w} = \mathbf{A}\mathbf{v}$ bedeutet

$$\begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} a_{11}v_1 + a_{12}v_2 \\ a_{21}v_1 + a_{22}v_2 \end{pmatrix},$$

d.h. in Komponentenschreibweise: $\boxed{w_j = \sum_k a_{jk}v_k}$.

- Analog Matrixmultiplikation:

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{pmatrix},$$

d.h. in Komponentenschreibweise:

Mit $\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, $\mathbf{B} = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$, $\mathbf{C} = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix}$

$\Rightarrow$ $\mathbf{C} = \mathbf{AB}$ bedeutet $\boxed{c_{ik} = \sum_j a_{ij}b_{jk}}$.

Die Matrixmultiplikation definiert eine assoziative Verknüpfung.
Neutrales Element: $\exists\mathbf{E}$ mit $\mathbf{AE} = \mathbf{A} = \mathbf{EA}\ \forall\mathbf{A}$ $\quad(\mathbf{E} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix})$
Die Matrix $\mathbf{A}^{-1}$ heisst linksinvers zu $\mathbf{A}$, wenn $\mathbf{A}^{-1}\mathbf{A} = \mathbf{E}$.

Kann verallgemeinert werden auf $n \times m$ -Matrizen (mit $\mathbf{E} = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}$).

In Matrixschreibweise kann man lineare Gleichungssysteme formal ganz einfach hinschreiben und "lösen". Zur Beschreibung des allgemeinen Gleichungssystems

$$\begin{aligned} a_{11}\,x_1 + a_{12}\,x_2 + \cdots + a_{1m}\,x_m &= b_1 \\ &\vdots \\ a_{n1}\,x_1 + a_{n2}\,x_2 + \cdots + a_{nm}\,x_m &= b_n \end{aligned}$$

definiert man

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ \vdots & & & \vdots \\ a_{n1} & a_{n2} & & a_{nm} \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}.$$

Dann entspricht das Gleichungssystem der Matrixgleichung $\mathbf{Ax} = \mathbf{b}$.

- Falls $\mathbf{A}$ ein Linksinverses hat, gibt es eine eindeutige Lösung: $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$.
  Anderenfalls gibt es keine oder unendlich viele Lösungen.
- Kriterium im Fall $n = m$: "Determinante" $\det(\mathbf{A})$ (siehe 2.3.4).
  $\mathbf{A}$ ist invertierbar, wenn $\det(\mathbf{A}) \neq 0$.

Für 2×2-Matrizen ist $\det(\mathbf{A}) = a_{11}a_{22} - a_{12}a_{21}$, genau die Größe D aus 2.1.1.
In 2.1.1 haben wir de facto auch $\mathbf{A}^{-1}$ ausgerechnet, $\mathbf{A}^{-1} = \frac{1}{\det(\mathbf{A})}\begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}$.

Für allgemeine n gibt es systematische Algorithmen zum Invertieren von $n \times n$-Matrizen, mehr dazu in Kapitel 2.3.

2.1.3 Gaußsches Eliminationsverfahren zum Lösen linearer Gleichungssysteme

Das Gaußsche Verfahren ist ein systematisches Verfahren zur Lösung linearer Gleichungssysteme ohne explizite Verwendung von Matrizen.

Der Ausgangspunkt ist ein lineares $n \times m$ Gleichungssystem

$$\begin{aligned} a_{11}\,x_1 + a_{12}\,x_2 + a_{13}\,x_3 + \cdots + a_{1m}\,x_m &= b_1 \\ &\vdots \\ a_{n1}\,x_1 + a_{n2}\,x_2 + a_{n3}\,x_3 + \cdots + a_{nm}\,x_m &= b_n \end{aligned}$$

- Erster Schritt: "Eliminiere Koeffizienten a_{1j}",
  d.h. bringe das Gleichungssystem auf die Form

$$\begin{aligned} a'_{11}\,x_1 + a'_{12}\,x_2 + a'_{13}\,x_3 + \cdots + a'_{1m}\,x_m &= b'_1 \\ a'_{22}\,x_2 + a'_{23}\,x_3 + \cdots + a'_{2m}\,x_m &= b'_2 \\ &\vdots \\ a'_{n2}\,x_2 + a'_{n3}\,x_3 + \cdots + a'_{nm}\,x_m &= b'_n \end{aligned}$$

  mit $a'_{11} \neq 0$. Dazu muss gegebenenfalls (falls $a_{11} = 0$) zunächst die erste Zeile mit einer anderen Zeile vertauscht werden. Danach werden die Koeffizienten $a_{21}, \cdots, a_{n1}$, die nicht sowieso schon Null sind, mit Hilfe derselben Operationen wie in 2.1.1 eliminiert. Insgesamt braucht man also nur drei Arten von sogenannten "Zeilenoperationen":

  - Vertausche zwei Zeilen
  - Multipliziere eine Zeile mit einer Zahl
  - Subtrahiere von einer Zeile eine andere

- Zweiter Schritt: "Eliminiere Koeffizienten a_{2j}":

  Falls $a'_{2k} \neq 0$ für ein $k > 1$, dann bringe das Gleichungssystem mittels analoger Zeilenoperationen wie im ersten Schritt auf die Form

$$\begin{aligned} a'_{11}\, x_1 + a'_{12}\, x_2 + a'_{13}\, x_3 + \cdots + a'_{1m}\, x_m &= b'_1 \\ a''_{22}\, x_2 + a''_{13}\, x_3 + \cdots + a''_{2m}\, x_m &= b''_2 \\ a''_{13}\, x_3 + \cdots + a''_{2m}\, x_m &= b''_2 \\ &\vdots \\ a''_{n3}\, x_3 + \cdots + a''_{nm}\, x_m &= b'_n \end{aligned}$$

  Falls $a'_{2k} = 0$ für alle $k > 1$, dann belasse es bei

$$\begin{aligned} a'_{11}\, x_1 + a'_{12}\, x_2 + a'_{13}\, x_3 + \cdots + a'_{1m}\, x_m &= b'_1 \\ a'_{13}\, x_3 + \cdots + a''_{2m}\, x_m &= b'_2 \\ &\vdots \\ a'_{n3}\, x_3 + \cdots + a'_{nm}\, x_m &= b'_n \end{aligned}$$

- Eliminiere sukzessive weiter, bis das Gleichungssystem eine Stufenform hat:

  Für $n < m$ erhält man in der Regel die Form

$$\begin{aligned} a^*_{11}\, x_1 + a^*_{12}\, x_2 + \cdots + a^*_{1n}\, x_n + \cdots + a^*_{1m}\, x_m &= b^*_1 \\ a^*_{22}\, x_2 + \cdots + a^*_{2n}\, x_n + \cdots + a^*_{2m}\, x_m &= b^*_2 \\ &\vdots \\ a^*_{nn}\, x_n + \cdots + a^*_{mm}\, x_m &= b^*_n \end{aligned}$$

  und für $n > m$ die Form

$$\begin{aligned} a^*_{11}\, x_1 + a^*_{12}\, x_2 + \cdots + a^*_{1m}\, x_m &= b^*_1 \\ a^*_{22}\, x_2 + \cdots + a^*_{2m}\, x_m &= b^*_2 \\ &\vdots \\ a^*_{mm}\, x_m &= b^*_m \\ 0 &= b^*_{m+1} \\ &\vdots \\ 0 &= b^*_n \end{aligned}$$

  Jedes x_k ist maximal einmal führende (erste) Variable und taucht in den Zeilen danach nicht mehr auf. Wie in Schritt 2 gesehen, ist es jedoch möglich, dass manche x_j übersprungen werden und nie führend sind.

- Analysiere:

  – Falls das resultierende Gleichungssystem Gleichungen $0 = b^*_k \neq 0$ enthält, hat es keine Lösung.

  – Anderenfalls löse auf nach der untersten führenden Variablen (in der Regel x_n) und dann durch Einsetzen sukzessive nach den darüberliegenden führenden Variablen (in der Regel x_{n-1}, x_{n-2}, ... x_1).

Der Algorithmus liefert mit Sicherheit die vollständige Lösungsmenge: Entweder eine unendliche Schar von Lösungen, eine eindeutige Lösung, oder $\mathbb{L} = \varnothing$.

2.2 Vektoren

2.2.1 Definition bzw. Begriffsklärung

Vektoren: Zentrale Größen in der linearen Algebra und der Physik
Aber: Was ist ein Vektor eigentlich?
→ physikalische und mathematische Sichtweise

Mathematischer Vektor: Element eines Vektorraums (siehe 2.2.3)

Physikalischer Vektor: Gerichtete Strecke mit Anfangs- und Endpunkt, z.B. Ort $\vec{r}$, Geschwindigkeit $\vec{v}$, Kraft $\vec{F}$.
Notation: $\vec{a}, \mathbf{a}, \underline{a}, \cdots$
Charakterisierung: Durch Betrag (Länge der Strecke) und Richtung (Einheitsvektor $\vec{e}_a$).

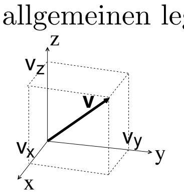
Skalare: Im Gegensatz zum Vektoren einfache "Zahlen"
Wieder physikalische und mathematische Sichtweise.

Mathematisch: Vektorraum ist "über einem Körper" definiert (siehe 2.2.3)
Skalare sind Elemente dieses Körpers.

Physikalisch: Größen, die durch einzelne Zahlen charakterisiert werden können (z.B. Temperatur, Druck, Betrag eines Vektors)

2.2.2 Koordinatensysteme und Koordinatendarstellung

Im allgemeinen legt man im Raum ein Koordinatensystem (KDS) fest.



Dann kann Vektor durch Satz von D Zahlen dargestellt werden. ($D = 3$ im Raum, $D = 2$ in der Ebene etc.)
Zahlen (Koordinaten) hängen vom KDS ab.

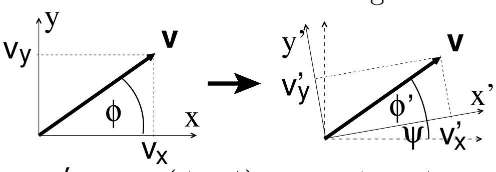
$\Longrightarrow$ Alternative Definition eines **physikalischen Vektors**:
Größe, charakterisiert durch D Zahlen (D: Dimension des Raums), die sich bei Drehung des Koordinatensystems in "bestimmter Weise" transformieren. (mehr dazu gleich).

Alternative Definition eines **physikalischen Skalars**:
Größe, charakterisiert durch eine Zahl, die vom Koordinatensystem unabhängig ist.

Frage: Was bedeutet "bestimmte Weise"?
Wie transformieren sich Vektorkoordinaten unter Drehung des KDS?
Annahme hier: rechtwinkliges Koordinatensystem.

Diskussion zunächst in 2 Dimensionen (Vektor in der Ebene)
Hier müssen trigonometrische Funktionen und deren Eigenschaften verwendet werden (Vorgriff auf 3.1.5)

Betrachte Einfluss auf Drehung um Winkel ψ auf Vektor $\vec{v}$.



Koordinaten des Vektor $\vec{v}$:
$$\begin{pmatrix} v_x = v\cos\phi \\ v_y = v\sin\phi \end{pmatrix} \to \begin{pmatrix} v_x' = v\cos\phi' \\ v_y' = v\sin\phi' \end{pmatrix}$$
mit $\phi' = \phi - \psi$

$$\Rightarrow \begin{array}{l} v_x' = v\cos(\phi-\psi) = v\cos\phi\cos\psi + v\sin\phi\sin\psi = v_x\cos\psi + v_y\sin\psi \\ v_y' = v\sin(\phi-\psi) = v\sin\phi\cos\psi - v\cos\phi\sin\psi = -v_x\sin\psi + v_y\cos\psi \end{array}$$

Matrix-Schreibweise: $\begin{pmatrix} v_x' \\ v_y' \end{pmatrix} = \underbrace{\begin{pmatrix} \cos\psi & \sin\psi \\ -\sin\psi & \cos\psi \end{pmatrix}}_{\text{Drehmatrix } \mathcal{D}(\psi)} \begin{pmatrix} v_x \\ v_y \end{pmatrix}$

$\rightsquigarrow$ Drehung des KDS kann durch **Drehmatrix** beschrieben werden.

Analog kann man Drehmatrizen auch in drei Dimensionen definieren.
(kompliziertere Form, die Matrizen hängen von drei Winkeln ab).

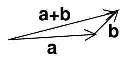
Vorteile der Beschreibung von Drehungen durch Drehmatrizen:

- Eine Drehmatrix beschreibt Einfluss der Drehung auf die Koordinaten aller Vektoren
- Kombinierte Drehungen kann man leicht durch Hintereinanderschalten von Drehmatrizen (Matrixmultiplikation, siehe 2.1.2) erzeugen.

2.2.3 Elementares Rechnen mit Vektoren, Vektorräume

1) Addition von Vektoren $\vec{a} + \vec{b}$:

- Graphisch: Hänge Pfeile aneinander



- Koordinaten: $\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} a_1 + b_2 \\ a_2 + b_2 \\ a_3 + b_3 \end{pmatrix}$.
- Rechengesetze: Vektoren mit Addition bilden "Abelsche Gruppe"

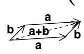
  (i) Abgeschlossen: $\vec{a} + \vec{b}$ ist wieder ein Vektor

  (ii) Assoziativ: $(\vec{a} + \vec{b}) + \vec{c} = \vec{a} + (\vec{b} + \vec{c})$.

  (iii) Es gibt eine Null $\vec{0}$ mit: $\vec{a} + \vec{0} = \vec{0} + \vec{a} = \vec{a} \qquad \forall \vec{a}$

  (iv) Inverses: $\forall \vec{a} \quad \exists \quad (-\vec{a}): \qquad \vec{a} + (-\vec{a}) = \vec{a} - \vec{a} = 0$.

  (v) Kommutativ: $\vec{a} + \vec{b} = \vec{b} + \vec{a}$.



2) Multiplikation mit einem Skalar : (i.e., einer Zahl $\lambda \in \mathbb{R}$ oder $\mathbb{C}$.)

- Graphisch: Streckung/Stauchung um Faktor λ
  (Falls $\lambda < 0$: Richtung dreht sich um.)
- Koordinaten: $\lambda \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} \lambda a_1 \\ \lambda a_2 \\ \lambda a_3 \end{pmatrix}$.

- Rechengesetze:

  (i') Immer noch abgeschlossen: $\lambda\vec{a}$ ist wieder ein Vektor.

  (vi) Kommutativ: $\lambda\vec{a} = \vec{a}\lambda$

  (vii) Assoziativ: $\beta(\alpha\vec{a}) = (\beta\alpha)\vec{a}$

  (viii) Distributivgesetze: $(\alpha + \beta)\vec{a} = \alpha\vec{a} + \beta\vec{a}$, $\alpha(\vec{a} + \vec{b}) = \alpha\vec{a} + \alpha\vec{b}$.

Bemerkung: Die Menge der Skalare $\mathbb{R}$ oder $\mathbb{C}$ für sich genommen bildet einen Körper (siehe 1.2.1). Eine Menge von Objekten, auf der eine Addition und eine Skalarmultiplikation mit den obigen Eigenschaften (i)-(viii) definiert ist bezeichnet man als linearen Vektorraum über dem Körper der Skalare.

⇒ Mathematische Definition eines Vektors: Element eines Vektorraums
Der mathematische Begriff des Vektors ist allgemeiner als der physikalische Begriff. In dieser Vorlesung werden mathematisch gesehen meistens Vektorräume $\mathbb{R}^D$ diskutiert (D ist die Raumdimension). Aber es gibt andere Möglichkeiten, z.B. bildet die Menge der Funktionen $f(x)$ auch einen Vektorraum über $\mathbb{R}$ oder $\mathbb{C}$.

3) Lineare Abhängigkeit :

Definition: Vektoren $\vec{a}_1, \vec{a}_2, \cdots \vec{a}_n$ heißen **linear unabhängig**,
falls aus $\sum \alpha_i \vec{a}_i = 0$ folgt $\alpha_i = 0 \quad \forall \quad i$.
Anderenfalls sind die Vektoren **linear abhängig**.

Anschaulich: Zwei Vektoren linear abhängig ⇒ parallel.
Drei Vektoren linear abhängig ⇒ koplanar (in einer Ebene) .

Bemerkungen:

– Zwei linear unabhängige Vektoren $\vec{a}, \vec{b}$ spannen Ebene durch Ursprung auf, d.h. alle Punkte der Ebene können in die Form $\alpha\vec{a} + \beta\vec{b}$ gebracht werden (α, β sind Skalare).

– Drei linear unabhängige Vektoren spannen gesamten Raum auf.

– In drei Dimensionen sind maximal drei Vektoren linear unabhängig.

(Mathematisch: Dimension ist definiert als die maximale Anzahl linear unabhängiger Vektoren in einem Vektorraum)

Aufgaben:

- Gegeben sei $\vec{a} = \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix}, \vec{b} = \begin{pmatrix} -1 \\ 0 \\ 2 \end{pmatrix}, \vec{c} = \begin{pmatrix} 1 \\ 1 \\ 5 \end{pmatrix}$.

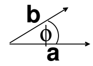
  Berechnen Sie $\vec{a} + \vec{b}$, $\vec{b} - \vec{c}$.
  Sind $(\vec{a}, \vec{b})$, $(\vec{a}, \vec{c})$, $(\vec{b}, \vec{c})$, $(\vec{a}, \vec{b}, \vec{c})$, linear unabhängig?
  Wie lautet die Gleichung für die Gerade durch $\vec{a}$, die zu $\vec{b}$ parallel ist?

- Bilden Sie in der (x, y)Ebene die Summe von 7 Vektoren der Länge a, wobei $\vec{a}_1$ in x-Richtung zeigt und $\vec{a}_j$ mit $\vec{a}_{j-1}$ die Winkeldifferenz $\pi/6$ hat.

2.2.4 Skalarprodukt (inneres Produkt)

2.2.4.1 Definition und mathematische Struktur

1) Skalarprodukt im Vektorraum $\mathbb{R}^3$ (oder $\mathbb{R}^D$ mit beliebigem $D \geq 1$):
Abbildung eines Paars von Vektoren auf einen Skalar



$\boxed{\vec{a} \cdot \vec{b} = a\, b\, \cos(\phi)}$ mit ϕ : Winkel zwischen $\vec{a}$ und $\vec{b}$.

Physikalisches Beispiel: Arbeit = Kraft mal Weg ($W = \vec{F} \cdot \vec{\Delta r}$).

2) Rechenregeln:

(i) Kommutativ: $\vec{a} \cdot \vec{b} = \vec{b} \cdot \vec{a}$

(ii) Homogen: $(\lambda \vec{a}) \cdot \vec{b} = \vec{a} \cdot (\lambda \vec{b}) = \lambda(\vec{a} \cdot \vec{b})$

(iii) Distributivgesetz: $\vec{a} \cdot (\vec{b} + \vec{c}) = \vec{a} \cdot \vec{b} + \vec{a} \cdot \vec{c}$.

(iv) $\vec{a} \cdot \vec{a} = a^2 \geq 0$
($a = \sqrt{\vec{a} \cdot \vec{a}}$ ist Betrag des Vektors).

- Kein eindeutiges Inverses
(durch Vektoren darf man nicht dividieren !)

Bemerkung: Einen Vektorraum mit einem Skalarprodukt, das die Eigenschaften (i)-(iv) erfüllt, nennt man **unitär**.

3) Folgerungen

- **Cauchy-Schwartzsche Ungleichung**: $\boxed{|\vec{a} \cdot \vec{b}| \leq ab}$
(hier konkret: klar, da $\cos\phi \leq 1$
Kann man aber auch allein aus (i)-(iv) beweisen:
Für alle λ gilt: $0 \leq (\vec{a} - \lambda \vec{b})^2 = a^2 + \lambda^2 b^2 - 2\lambda(\vec{a} \cdot \vec{b})$
Wähle $\lambda = \frac{(\vec{a} \cdot \vec{b})}{b^2} \Rightarrow 0 \leq a^2 - \frac{(\vec{a} \cdot \vec{b})}{b^2} \Rightarrow (\vec{a} \cdot \vec{b})^2 \leq a^2 b^2 \quad \checkmark$)

Falls $|\vec{a} \cdot \vec{b}| = ab$ folgt: Vektoren $\vec{a}, \vec{b}$ sind parallel.
(möglicher Test auf lineare Abhängigkeit).

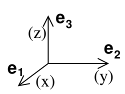
- **Dreiecks-Ungleichung**: $\boxed{|a - b| \leq |\vec{a} + \vec{b}| \leq a + b}$

$$\text{vgl.} \left\{ \begin{array}{lcl} (a-b)^2 & = & a^2 + b^2 - 2ab \\ (\vec{a} + \vec{b})^2 & = & a^2 + b^2 + 2\vec{a} \cdot \vec{b} \\ (a+b)^2 & = & a^2 + b^2 + 2ab \end{array} \right\} \begin{array}{l} \text{Aus } -ab \leq \vec{a} \cdot \vec{b} \leq ab \\ \text{(Cauchy-Schwartzsche Ungl.)} \\ \text{folgt } (a-b)^2 \leq (\vec{a} + \vec{b})^2 \leq (a+b)^2 \ \checkmark \end{array}$$

- **Einheitsvektor** in Richtung $\vec{a}$: $\vec{e}_a = \frac{\vec{a}}{a} = \frac{\vec{a}}{\sqrt{\vec{a} \cdot \vec{a}}}$

2.2.4.2 Koordinatendarstellung und Kronecker-Symbol

Ab jetzt sollen unsere Koordinatensysteme rechtwinklige Koordinatenachsen haben, die ein "Rechtssystem" bilden (Rechte-Hand Regel).



Definiere $\vec{e}_i$: Einheitsvektoren entlang der Koordinatenachse i.
$\rightarrow \vec{a} = a_1 \vec{e}_1 + a_2 \vec{e}_2 + a_3 \vec{e}_3 = \sum_i a_i \vec{e}_i$ (mit a_i: Koordinaten)

Es gilt: $\vec{e}_i \cdot \vec{e}_j = \left\{ \begin{array}{ll} 1: & i = j \\ 0: & i \neq j \end{array} \right.$

1) Motiviert Definition des **Kronecker-Symbols**

$$\boxed{\delta_{ij} := \begin{cases} 1: & i = j \\ 0: & i \neq j \end{cases}} = \vec{e}_i \cdot \vec{e}_j$$

2) Skalarprodukt in Koordinatendarstellung: $\vec{a} = \sum_i a_i \vec{e}_i$; $\vec{b} = \sum_j b_j \vec{e}_j$

$\Rightarrow \vec{a} \cdot \vec{b} = (\sum_i a_i \vec{e}_i)(\sum_j b_j \vec{e}_j) = \sum_{ij} a_i\, b_j\, (\vec{e}_i \cdot \vec{e}_j) = \sum_{ij} a_i\, b_j\, \delta_{ij}$

$\Rightarrow \boxed{\vec{a} \cdot \vec{b} = \sum_i a_i\, b_i}$

3) Anwendung: Berechnung der Koordinaten eines Vektors $\vec{a}$ in einem anderen (rechtwinkligen) Koordinatensystem mit Koordinatenachsen in Richtung $\{\vec{e}_i'\}$ (wobei die $\vec{e}_i'$ wieder Basisvektoren sein sollen, $|\vec{e}_i'| = 1 \quad \forall i$.
Koordinaten: Projektion von $\vec{a}$ auf die Achsen (Basisvektoren) $\vec{e}_i'$:
$\Rightarrow \boxed{a_i' = \vec{a} \cdot \vec{e}_i'} \Rightarrow \vec{a} = \sum_i a_i' \vec{e}_i' = \sum_i (\vec{a} \cdot \vec{e}_i')\, \vec{e}_i'$.
Gilt für alle Vektoren $\vec{a}$ und rechtwinklige Koordinatensysteme $\{\vec{e}_i'\}$.

$\rightsquigarrow$ Liefert einfache Berechnungsformel für Drehmatrizen.
Betrachte Drehung, die KDS $\{\vec{e}_i\}$ in KDS $\{\vec{e}_i'\}$ überführt.
Vektor $\vec{a}$ hat Koordinaten a_i bzw. a_i' mit $a_i = \vec{a} \cdot \vec{e}_i$, $a_i' = \vec{a} \cdot \vec{e}_i'$
Aus $\vec{a} = \sum_j a_j \vec{e}_j$ und $a_i' = \vec{a} \cdot \vec{e}_i'$ folgt $a_i' = \sum_j a_j (\vec{e}_j \cdot \vec{e}_i')$
also $\boxed{a_i' = \sum_j \mathcal{D}_{ij} a_j}$ mit Drehmatrix $\boxed{\mathcal{D}_{ij} = \vec{e}_i' \cdot \vec{e}_j}$

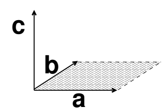
Aufgaben:

- Beweisen Sie den Kosinussatz für Dreiecke: Für Dreiecke mit Seitenlängen a, b, c und dem Winkel ϕ zwischen den Seiten a und b gilt $a^2 + b^2 - 2ab\cos\phi = c^2$.
- Betrachten Sie Vektoren mit Koordinaten $\vec{a} = \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix}, \vec{b} = \begin{pmatrix} -1 \\ 0 \\ 2 \end{pmatrix}$. Berechnen Sie $\vec{a} \cdot \vec{b}$
- Bestimmen Sie den Einheitsvektor in Richtung $\begin{pmatrix} -1 \\ 2 \\ -2 \end{pmatrix}$.

2.2.5 Vektorprodukt (äußeres Produkt, Kreuzprodukt)

2.2.5.1 Definition und mathematische Struktur

1) Vektorprodukt im Vektorraum $\mathbb{R}^3$:
Abbildung eines Paars von Vektoren auf Vektor $\vec{c} = \vec{a} \times \vec{b}$
mit Betrag $\boxed{|\vec{c}| = a\, b\, \sin(\phi)}$
(ϕ : Winkel zwischen $\vec{a}$ und $\vec{b}$.)
Richtung: $\vec{c}$ senkrecht auf $\vec{a}$, $\vec{b}$
Vorzeichen: So, dass $(\vec{a}, \vec{b}, \vec{c})$ Rechtssystem.
Geometrisch: $|\vec{c}|$ ist Fläche des von $\vec{a}$ und $\vec{b}$ aufgespannten Parallelogramms
Physikalisches Beispiel: Drehimpuls $\vec{r} \times \vec{p}$.



2) Rechengesetze

(i) Antikommutativ: $\vec{a} \times \vec{b} = -\vec{b} \times \vec{a}$

(ii) Homogen: $(\lambda\vec{a}) \times \vec{b} = \vec{a} \times (\lambda\vec{b}) = \lambda(\vec{a} \times \vec{b})$.

(iii) Distributivgesetz: $\vec{a} \times (\vec{b} + \vec{c}) = \vec{a} \times \vec{b} + \vec{a} \times \vec{c}$
(zur Herleitung siehe z.B. Nolting Bd. 1)
NB: Aus (i) und (iii) folgt auch $(\vec{a} + \vec{b}) \times \vec{c} = \vec{a} \times \vec{c} + \vec{b} \times \vec{c}$

(iv) Jacobi-Identität $\vec{a} \times (\vec{b} \times \vec{c}) + \vec{b} \times (\vec{c} \times \vec{a}) + \vec{c} \times (\vec{a} \times \vec{b}) = 0$
(vorläufig ohne Herleitung).

- Nicht assoziativ, keine Identität, kein Inverses
(durch Vektoren kann man immer noch nicht dividieren).

Bemerkung: Einen Vektorraum mit einem äußeren Produkt, das die Eigenschaften (i)-(iv) erfüllt, nennt man eine **Lie Algebra**.

2.2.5.2 Koordinatendarstellung und Levi-Civita-Symbol

Das Koordinatensystem sei wieder ein rechtwinkliges Rechtssystem, mit Einheitsvektoren $\vec{e}_i$ entlang den Koordinatenachsen.

$$\begin{pmatrix} \Rightarrow & \vec{e}_1 \times \vec{e}_2 & = & \vec{e}_3 & = & -\vec{e}_2 \times \vec{e}_1 \\ & \vec{e}_2 \times \vec{e}_3 & = & \vec{e}_1 & = & -\vec{e}_3 \times \vec{e}_2 \\ & \vec{e}_3 \times \vec{e}_1 & = & \vec{e}_2 & = & -\vec{e}_1 \times \vec{e}_3 \end{pmatrix} \qquad \vec{e}_i \times \vec{e}_i = 0 \quad \forall i.$$

1) Motiviert Definition des Epsilon-Tensor oder Levi-Civita-Symbol:

$$\boxed{\epsilon_{ijk} := \begin{cases} 1: & (ijk) = (123), (231), (312) \\ -1: & (ijk) = (213), (321), (132) \\ 0: & \text{sonst} \end{cases}} = (\vec{e}_i \times \vec{e}_j) \cdot \vec{e}_k$$

(bzw. "total antisymmetrischer Tensor dritter Stufe").

2) Vektorprodukt in Koordinatendarstellung:

Mit $\vec{a} = \sum_i a_i \vec{e}_i$; $\quad \vec{b} = \sum_j b_j \vec{e}_j$ folgt $\vec{c} = \vec{a} \times \vec{b} = \sum_{ij} a_i\, b_j\, (\vec{e}_i \times \vec{e}_j)$
$\Rightarrow c_k = \vec{c} \cdot \vec{e}_k = [\vec{a} \times \vec{b}]_k = \sum_{ij} a_i\, b_j\, (\vec{e}_i \times \vec{e}_j) \cdot \vec{e}_k = \sum_{ij} a_i\, b_j\, \epsilon_{ijk}$

$$\Rightarrow \quad \boxed{[\vec{a} \times \vec{b}]_k = \sum_{ij} a_i\, b_j\, \epsilon_{ijk}}, \qquad \text{konkret:} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} a_2 b_3 - a_3 b_2 \\ a_3 b_1 - a_1 b_3 \\ a_1 b_2 - a_2 b_1 \end{pmatrix}$$

3) Rechenregeln mit dem Epsilon-Tensor:

$$\boxed{\epsilon_{ijk} = \underbrace{\epsilon_{jki} = \epsilon_{kij}}_{\text{gerade Permutationen}} = \underbrace{-\epsilon_{jik} = -\epsilon_{kji} = -\epsilon_{ikj}}_{\text{ungerade Permutationen von } (i,j,k)}}$$

$$\boxed{\sum_m \epsilon_{klm}\epsilon_{pqm} = \delta_{kp}\delta_{lq} - \delta_{kq}\delta_{lp}}$$

(Beweis: Hausaufgabe – Buchhaltung, argumentieren ...)

$$\boxed{\sum_{lm} \epsilon_{klm}\epsilon_{plm} = 2\delta_{kp}}$$

(denn: $\sum_{lm} \epsilon_{klm}\epsilon_{plm} = \sum_l (\delta_{kp} \underbrace{\delta_{ll}}_{1} - \delta_{kq}\delta_{lp}) = 3\delta_{kp} - \delta_{kp}$ ✓)

Aufgaben:

- Berechnen Sie $\vec{a} \times \vec{b}$ mit $\vec{a} = \begin{pmatrix} -1 \\ 2 \\ -3 \end{pmatrix}$ und $\vec{b} = \begin{pmatrix} 3 \\ -1 \\ 5 \end{pmatrix}$.
- Berechnen Sie die Fläche des von $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$. und $\begin{pmatrix} -2 \\ 1 \end{pmatrix}$ aufgespannten Parallelogramms.

2.2.6 Höhere Vektorprodukte

1) Spatprodukt $\boxed{(\vec{a}\vec{b}\vec{c}) = \vec{a} \cdot (\vec{b} \times \vec{c})}$ ($\rightarrow$ Skalar)

Geometrische Interpretation:

Betrag: Volumen eines von $(\vec{a}, \vec{b}, \vec{c})$ aufgespannten Parallelepipeds.

Vorzeichen: +, falls $(\vec{a}, \vec{b}, \vec{c})$ Rechtssystem,

– für Linkssystem.

Koordinatenschreibweise:

$$(\vec{a}\vec{b}\vec{c}) = \sum_k a_k (\vec{b} \times \vec{c})_k = \sum_{ijk} a_k b_i c_j \, \epsilon_{ijk} = \sum_{ijk} \epsilon_{ijk} \, a_i b_j c_k$$

$$\Rightarrow \boxed{(\vec{a}\vec{b}\vec{c}) = \sum_{ijk} \epsilon_{ijk} \, a_i b_j c_k}.$$

2) Doppeltes Vektorprodukt: $\vec{a} \times (\vec{b} \times \vec{c})$ ($\rightarrow$ Vektor)

Dafür gilt "Entwicklungssatz": $\boxed{\vec{a} \times (\vec{b} \times \vec{c}) = \vec{b}(\vec{a} \cdot \vec{c}) - \vec{c}(\vec{a} \cdot \vec{b})}$

(Herleitung am schnellsten mit Hilfe des Epsilon-Tensors:

$$\begin{aligned} (\vec{a} \times (\vec{b} \times \vec{c}))_k &= \sum_{ij} \epsilon_{ijk} a_i (\vec{b} \times \vec{c})_j = \sum_{ij} \epsilon_{ijk} a_i \sum_{lm} \epsilon_{lmj} b_l c_m \\ &= \sum_{ijlm} \epsilon_{ijk} \epsilon_{lmj} a_i b_l c_m = \sum_{ilm} (\sum_j \epsilon_{kij} \epsilon_{lmj}) a_i b_l c_m \\ &= \sum_{ilm} (\delta_{kl} \delta_{im} - \delta_{km} \delta_{il}) a_i b_l c_m \\ &= \sum_m a_m b_k c_m - \sum_l a_l b_l c_k = b_k (\vec{a} \cdot \vec{c} - c_k (\vec{a} \cdot \vec{b}) \quad). \end{aligned}$$

3) <u>Weitere wichtige Beziehungen</u>:

$$(\vec{a} \times \vec{b}) \cdot (\vec{c} \times \vec{d}) = (\vec{a} \cdot \vec{c})(\vec{b} \cdot \vec{d}) - (\vec{a} \cdot \vec{d})(\vec{b} \cdot \vec{c})$$

$$(\vec{a} \times \vec{b})^2 = a^2 b^2 - (\vec{a} \cdot \vec{b})^2 \qquad \text{(\underline{Lagrange-Identität})}$$

$$(\vec{a} \times \vec{b}) \times (\vec{c} \times \vec{d}) = (\vec{c}\vec{d}\vec{a})\vec{b} - (\vec{c}\vec{d}\vec{b})\vec{a} = (\vec{a}\vec{b}\vec{d})\vec{c} - (\vec{a}\vec{b}\vec{c})\vec{d}$$

Aufgaben:

- Beweisen Sie die Jacobi-Identität
- Gegeben die Vektoren $\vec{a} = \begin{pmatrix} -1 \\ 2 \\ -3 \end{pmatrix}, \vec{b} = \begin{pmatrix} 3 \\ -1 \\ 5 \end{pmatrix}, \vec{c} = \begin{pmatrix} -1 \\ 0 \\ 2 \end{pmatrix}$.

  Berechnen Sie $(\vec{a}\vec{b}\vec{c})$, $|(\vec{a} \times \vec{b}) \times \vec{c}|$, $|\vec{a} \times (\vec{b} \times \vec{c})|$, $(\vec{a} \times \vec{b}) \times (\vec{b} \times \vec{c})$.
- Beweisen Sie mit Hilfe des Levi-Civita-Symbols: $(\vec{a} \times \vec{b}) \cdot (\vec{c} \times \vec{d}) = (\vec{a} \cdot \vec{c})(\vec{b} \cdot \vec{d}) - (\vec{a} \cdot \vec{d})(\vec{b} \cdot \vec{c})$
- Berechnen Sie $(\vec{a} \times \vec{b}) \cdot (\vec{c} \times \vec{d}) + (\vec{b} \times \vec{c}) \cdot (\vec{a} \times \vec{d}) + (\vec{c} \times \vec{a}) \cdot (\vec{b} \times \vec{d})$.
- Die Vektoren $\vec{a}_1, \vec{a}_2, \vec{a}_3$ seien linear unabhängig.

  Die reziproken Vektoren seien definiert durch: $\vec{b}_1 = \frac{\vec{a}_2 \times \vec{a}_3}{(\vec{a}_1 \vec{a}_2 \vec{a}_3)} \vec{b}_2 = \frac{\vec{a}_3 \times \vec{a}_1}{(\vec{a}_1 \vec{a}_2 \vec{a}_3)} \vec{b}_3 = \frac{\vec{a}_1 \times \vec{a}_2}{(\vec{a}_1 \vec{a}_2 \vec{a}_3)}$.

  Zeigen Sie $\vec{a}_i \vec{b}_j = \delta_{ij}$
- Zeigen Sie $(\vec{a} - \vec{b}) \cdot ((\vec{a} + \vec{b}) \times \vec{c}) = 2\vec{a}(\vec{b} \times \vec{c})$.

2.3 Matrizen

2.3.1 Beispiele von Matrizen

Matrizen sind hier schon einige Male verwendet worden (2.1.2, 2.2.2). Sie sind neben Vektoren offenbar weitere nützliche Konstrukte, mit denen sich Sachverhalte kurz und präzise ausdrücken lassen. Bedeutung unter anderem für

- Lineare Gleichungssysteme (2.1.2)
- Beschreibung von Drehungen im Raum (2.2.2)
- Darstellung bestimmter physikalischer Größen

Das soll im Folgenden kurz rekapituliert/illustriert werden.

1) Lineare Gleichungssysteme

Beispiel:
$$\begin{aligned} 3x \; + 7y - 2z &= 2 \\ -2x + \; y \qquad &= 1 \end{aligned}$$

Charakterisiert durch "Koeffizientenmatrix" $\begin{pmatrix} 3 & 7 & -2 \\ -2 & 1 & 0 \end{pmatrix}$

und "Spaltenvektor" $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$.

Eigenschaften des Gleichungssystems werden im Wesentlichen von der Koeffizientenmatrix bestimmt.

"Matrix" hier: Zahlenschema aus $m \times n$ Zahlen a_{ij} (m Zeilen, n Spalten)

Darstellung des Gleichungssystems in Matrixschreibweise:

$$\begin{pmatrix} 3 & 7 & -2 \\ -2 & 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

2) Drehungen

Gegeben physikalischer Vektor $\vec{a}$, zwei rechtwinklige Koordinatensysteme Σ, Σ' mit Basisvektoren $\{\vec{e}_i\}$, $\{\vec{e}_i'\}$ (Einheitsvektoren entlang Achsen i).

Darstellung von $\vec{a}$ im Koordinatensystem Σ: $\vec{a} = \sum_i a_i \vec{e}_i$ mit $a_i = (\vec{a} \cdot \vec{e}_i)$.
Darstellung von $\vec{a}$ im Koordinatensystem Σ': $\vec{a} = \sum_i a_i' \vec{e}_i'$ mit $a_i' = (\vec{a}' \cdot \vec{e}_i')$.

$\Rightarrow$ Einfache Regel für die Umrechnung von Koordinaten $\{a_i\} \to \{a_i'\}$:

$$a_i' = (\vec{a} \cdot \vec{e}_i') = (\sum_j a_j \vec{e}_j) \cdot \vec{e}_i' = \sum_j (\vec{e}_i' \cdot \vec{e}_j) a_j =: \sum_j \mathcal{D}_{ij} a_j \tag{2.1}$$

Definiert <u>Drehmatrix</u> $\mathbf{D} = (\mathcal{D}_{ij})$ mit $\mathcal{D}_{ij} = \vec{e}_i' \cdot \vec{e}_j$.

Dann folgt $a_i' = \sum_j \mathcal{D}_{ij} a_j$ bzw. in Matrixschreibweise: $\mathbf{a}' = \mathbf{D}\mathbf{a}$.

Dies gilt für <u>alle</u> physikalischen Vektoren $\vec{a}$. Die Drehung des Koordinatensystems wird durch das Zahlenfeld $\mathbf{D}$ vollständig bestimmt.

Konkret z.B. Ebene: $\mathbf{D} = \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix}$

Verallgemeinerung: Allgemeine affine Koordinatentransformation zwischen Koordinatensystemen, die nicht unbedingt rechtwinklig sind: Beschrieben durch allgemeine invertierbare 3×3-Matrix.

Bemerkung: Mit der Einführung der Drehmatrizen wird die Konkretisierung des Begriffs **"physikalischer Vektor"** möglich: Ein physikalischer Vektor ist eine Größe, charakterisiert durch drei Zahlen $\{v_i\}$ (im dreidimensionalen Raum – zwei Zahlen in der Ebene), die sich unter Drehung des Koordinatensystems in folgender Weise transformieren:

$$\{v_i\} \to \{v_i'\} \text{ mit } v_i' = \textstyle\sum_j \mathcal{D}_{ij} v_j.$$

3) Physikalische Tensoren

Es gibt physikalische Größen, die nicht durch Vektoren, sondern durch Matrizen beschrieben werden, z.B. der sogenannte "Trägheitstensor". Wie die physikalischen Vektoren haben auch diese unter Drehung des KDS ein wohldefiniertes Transformationsverhalten. Soll hier nicht weiter diskutiert werden.

2.3.2 Elementare Begriffe

1) Definition einer **Matrix**

Eine $m \times n$-Matrix ist ein rechteckiges Zahlenschema mit m Zeilen und n Spalten: $\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} = (a_{ij})$ ($a_{ij} \in \mathbb{R}$ oder $\mathbb{C}$).

Zwei Matrizen gelten als gleich, wenn jeder Eintrag a_{ij} gleich ist.

2) Spezielle Matrizen

- Nullmatrix: $\mathbf{0} = \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{pmatrix}$.
- Quadratische $n \times n$-Matrizen
  - Symmetrische Matrix: $a_{ij} = a_{ji} \ \forall\, i, j$
  - Antisymmetrische Matrix: $a_{ij} = -a_{ji} \ \forall\, i, j$
  - Diagonalmatrix: $a_i j = \delta_{ij}\, a_i$
  - Einheitsmatrix: $\mathbf{E} = \mathbb{1} = \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix}$, d.h. $E_{ij} = \delta_{ij}$
- Vektoren:

  Spaltenvektor: $n \times 1$-Matrix

  Zeilenvektor: $1 \times n$-Matrix

→ Koordinatendarstellungen von Vektoren sind spezielle Formen von Matrizen.

3) Rang einer Matrix

$m \times n$-Matrix kann man sich zusammengesetzt denken aus m Zeilenvektoren oder n Spaltenvektoren.

Zeilenrang: Maximale Zahl linear unabhängiger Zeilenvektoren
Spaltenrang: Maximale Zahl linear unabhängiger Spaltenvektoren

Es gilt: Spaltenrang = Zeilenrang.

2.3.3 Rechnen mit Matrizen

2.3.3.1 Elementare Operationen

⋆ **Addition** : Seien $\mathbf{A} = (a_{ij})$ und $\mathbf{B} = (b_{ij})$ zwei $n \times m$-Matrizen

$\mathbf{C} = (c_{ij})$, $\mathbf{C} = \mathbf{A} + \mathbf{B}$ bedeutet: $c_{ij} = a_{ij} + b_{ij}\ \forall\, i, j$

⋆ **Multiplikation mit Skalar** : Sei λ Skalar

$\mathbf{C} = \lambda A$ bedeutet: $c_{ij} = \lambda a_{ij}\ \forall\, i, j$.

Bemerkung: Für gegebenes n, m bildet die Menge der $n \times m$-Matrizen mit Addition und Skalarmultiplikation wieder einen Vektorraum über dem Körper der Skalare ($\mathbb{R}$ oder $\mathbb{C}$).

⋆ **Transposition**

$\mathbf{C} = \mathbf{A}^T$ bedeutet: $c_{ij} = a_{ji}$

⋆ **Matrixmultiplikation** : Sei $\mathbf{A} = (a_{ij})$, $\mathbf{B} = (b_{ij})$,
und: Die Anzahl Spalten von A entspricht der Anzahl Zeilen von B

$\mathbf{C} = \mathbf{AB}$ bedeutet: $c_{ij} = \sum_k a_{ik}\, b_{kj}$.

Eigenschaften der Matrixmultiplikation

- Assoziativ: $\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}$
- Neutrales Element: Sei $\mathbf{A}$ eine $m \times n$ Matrix und $\mathbb{1}_k$ die $k \times k$-Einheitsmatrix. Dann erfüllt $\mathbb{1}$: $\mathbf{A}\mathbb{1}_n = \mathbb{1}_m\mathbf{A} = \mathbf{A}\ \forall \mathbf{A}$
- Nicht kommutativ: Im allgemeinen ist $\mathbf{AB} \neq \mathbf{BA}$.

  Bemerkung: Falls $\mathbf{A}$ nicht eine quadratische $n \times n$-Matrix ist, ist $\mathbf{BA}$ noch nicht einmal definiert!
- Matrix A kann Inverses haben (siehe 5), muss aber nicht.

2.3.3.2 Invertieren von Matrizen

$\mathbf{A}$ sei eine $m \times n$-Matrix:
Die $n \times m$-Matrix $(\mathbf{A}^{-1})$ ist "Linksinverses" von $\mathbf{A}$, wenn $(\mathbf{A}^{-1})\mathbf{A} = \mathbb{1}$.
(Definition der "Rechtsinversen": Analog)

Es gilt: $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$ (leicht zu zeigen)

Für $n = m$: $\mathbf{A}\mathbf{A}^{-1} = \mathbb{1}$

(denn: Sei $\mathbf{B} = \mathbf{A}\mathbf{A}^{-1} \Rightarrow \mathbf{A}^{-1}\mathbf{B} = \underbrace{\mathbf{A}^{-1}\mathbf{A}}_{\mathbb{1}}\mathbf{A}^{-1} = \mathbf{A}^{-1} \ \forall\ \mathbf{A} \Rightarrow \mathbf{B} = \mathbb{1}$)

d.h. $\mathbf{A}^{-1}$ ist dann auch Rechtsinverses und $(\mathbf{A}^{-1})^{-1} = \mathbf{A}$

Praktische Berechnung:

Lösung eines Satzes von linearen Gleichungssystemen:

Beispiel: Gegeben $A = \begin{pmatrix} 1 & 3 \\ 2 & 1 \end{pmatrix}$. Gesucht $A^{-1} = \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix}$

mit $\begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix}\begin{pmatrix} 1 & 3 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

Entspricht zwei linearen Gleichungssystemen:

$$\begin{array}{rlclrcl} (*) & 1x_{11} + 3x_{12} & = & 1 & \qquad 1x_{21} + 3x_{22} & = & 0 \\ (**) & 2x_{11} + 1x_{12} & = & 0 & \qquad 2x_{21} + 1x_{22} & = & 1 \end{array}$$

Selbe Koeffizienten, nur rechte Seiten sind verschieden.

→ Man kann jeweils dieselben Transformationen benutzen, um nach den beiden Unbekannten aufzulösen.

Nach 2.1.3 können Gleichungssysteme über drei Arten von "Zeilentransformationen" gelöst werden. Diese lassen sich auch als Matrizen $\mathbf{Z}_j$ darstellen:

Zum Beispiel für 2×2-Matrizen:

– Vertauschung der Zeilen 1 und 3: $\mathbf{Z} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$

– Multiplikation der Zeile 2 mit λ: $\mathbf{Z} = \begin{pmatrix} 1 & 0 \\ 0 & \lambda \end{pmatrix}$

– Ersetze Zeile 2 durch (Zeile 1 - Zeile 2): $\mathbf{Z} = \begin{pmatrix} 1 & 0 \\ 1 & -1 \end{pmatrix}$

Zeilentransformationen $\mathbf{Z}_j$ werden aneinandergereiht, bis die Gleichungen nach den Unbekannten aufgelöst sind.

→ Entspricht formal Matrixmultiplikation $\mathbf{A} \to \mathbf{Z}_k \cdots \mathbf{Z}_2\mathbf{Z}_1\mathbf{A} = \mathbb{1}$.
Damit folgt $\mathbf{A}^{-1} = \mathbf{Z}_k \cdots \mathbf{Z}_2\mathbf{Z}_1$.

Das motiviert folgendes

Lösungsverfahren (für $n \times n$-Matrizen).

Schreibe Koeffizientenmatrix (für linke Seite) und Einheitsmatrix (für rechte Seiten) nebeneinander auf. Führe dann die Zeilentransformationen aus, die die Gleichungssysteme lösen.

→ Linke Matrix (Koeffizientenmatrix) wird zur Einheitsmatrix, Rechte Matrix wird zur gesuchten inversen Matrix.

Konkret: $\left(\begin{array}{cc|cc} 1 & 3 & 1 & 0 \\ 2 & 1 & 0 & 1 \end{array}\right) \xrightarrow{(**)\to(*)-2(**)} \left(\begin{array}{cc|cc} 1 & 3 & 1 & 0 \\ 0 & 5 & 2 & -1 \end{array}\right)$

$\xrightarrow{(**)\to(**)/5} \left(\begin{array}{cc|cc} 1 & 3 & 1 & 0 \\ 0 & 1 & \frac{2}{5} & -\frac{1}{5} \end{array}\right) \xrightarrow{(*)\to(*)-3(**)} \left(\begin{array}{cc|cc} 1 & 0 & -\frac{1}{5} & \frac{3}{5} \\ 0 & 1 & \frac{2}{5} & -\frac{1}{5} \end{array}\right)$

$\Rightarrow$ <u>Ergebnis:</u> $A^{-1} = \begin{pmatrix} -\frac{1}{5} & \frac{3}{5} \\ \frac{2}{5} & -\frac{1}{5} \end{pmatrix}$

(Alternativen: Cramers Regel, siehe 2.3.4 wird aber für $n > 2$ impraktikabel. Weitere numerische Verfahren, z.B. LU-Zerlegung, siehe Numerikliteratur).

2.3.3.3 Charakteristische Größen quadratischer Matrizen

Sogenannte Invarianten: Abbildung von $n \times n$ Matrizen auf Skalare, die sich nicht ändern, wenn man die Matrizen als physikalische Objekte (Tensoren) auffasst und das Koordinatensystem dreht. Bei Drehungen würden sich, wie in 2.3.1, die Einträge in den Matrizen transformieren. Die Invarianten dagegen sind unabhängig vom Koordinatensystem (ähnlich dem Skalarprodukt von Vektoren).

⋆ **Spur** : Für $n \times n$ Matrizen $\mathbf{A}$ ist $\mathrm{Sp}(\mathbf{A}) := \sum_{i=1}^{n} a_{ii}$
Es gilt: $\mathrm{Sp}(\mathbf{AB}) = \mathrm{Sp}(\mathbf{BA})$
(leicht zu sehen, wenn man es explizit in Koordinaten hinschreibt).
$\Rightarrow \mathrm{Sp}(\mathbf{A}_1 \cdots \mathbf{A}_{n-1}\mathbf{A}_n) = \mathrm{Sp}(\mathbf{A}_n\mathbf{A}_1 \cdots \mathbf{A}_{n-1})$:
Matrizen in der Spur dürfen zyklisch vertauscht werden.

⋆ **Determinante** : Siehe nächster Abschnitt

Aufgaben:

- Gegeben seien die Matrizen $\mathbf{A} = \begin{pmatrix} 0 & 1 & 2 \\ 3 & 0 & 4 \\ 0 & 0 & 5 \end{pmatrix}, \mathbf{B} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \mathbf{a} = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}$.
  Berechnen Sie $\mathbf{AB}$, $\mathbf{BA}$, $\mathbf{Aa}$, $\mathbf{a}^T\mathbf{B}$.
- Gegeben sei die Matrix $\mathbf{A} = \begin{pmatrix} a & b & c & d \\ -b & a & -d & c \\ -c & d & a & -b \\ -d & -c & b & a \end{pmatrix}$.
  Berechnen Sie $\mathbf{AA}^T$.
- Gegeben seien die Vektoren $\vec{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}, \vec{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$ und die Matrix $\mathbf{A} = \begin{pmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{pmatrix}$.
  Zeigen Sie $\mathbf{A}\vec{b} = \vec{a} \times \vec{b}$.
- Gegeben sei die Matrix $\mathbf{A} = \begin{pmatrix} 1 & 2 & 4 \\ -1 & 4 & 1 \\ 4 & 2 & 1 \end{pmatrix}$.
  Berechnen Sie die Inverse $\mathbf{A}^{-1}$. Testen Sie explizit $\mathbf{A}^{-1}\mathbf{A} = \mathbb{1}$.
- Gegeben sei die Matrix $\mathbf{B} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$.
  Berechnen Sie explizit alle Matrizen $\mathbf{X} = \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix}$, für die gilt: $\mathbf{BX} = \mathbf{XB}$.
- Zeigen Sie explizit: Für $n \times n$ Matrizen $\mathbf{A}, \mathbf{B}, \mathbf{C}$ gilt
  $\mathrm{Sp}(\mathbf{AB}) = \mathrm{Sp}(\mathbf{BA})$ und $\mathrm{Sp}(\mathbf{ABC}) = \mathrm{Sp}(\mathbf{BCA}) = \mathrm{Sp}(\mathbf{CAB})$
- In der Quantenmechanik sind die Pauli-Matrizen
  $\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ von großer Bedeutung.
  Hier ist i die sogenannte imaginäre Einheit, die im Körper $\mathbb{C}$ der komplexen Zahlen den reellen Zahlen hinzugefügt wird. Für diese Aufgabe hier müssen Sie nur wissen, dass $i^2 = -1$ und alle Skalare in $\mathbb{C}$ die Form $a + ib$ mit $a, b \in \mathbb{R}$ haben.

– Zeigen Sie, dass der Satz von Matrizen $(\mathbb{1}, \sigma_1, \sigma_2, \sigma_3)$ linear unabhängig ist im Vektorraum der 2×2-Matrizen über $\mathbb{C}$.
– Zeigen Sie $\sigma_j \sigma_k - \sigma_k \sigma_j = 2i \sum_{l=1}^{3} \epsilon_{jkl} \sigma_l$ für alle j, k.
– Berechnen Sie $\sigma_j \sigma_k + \sigma_k \sigma_j$ für alle j, k.

2.3.4 Determinanten

2.3.4.1 Einführung und Definition

Betrachte $n \times n$ Matrix $\mathbf{A} = (a_{ij})$

Die Determinante $\det(\mathbf{A})$ ist bis auf ein Vorzeichen das "Volumen", das von den n Spaltenvektoren aufgespannt wird:

Daraus folgt z.B. $\det(A) = 0 \Leftrightarrow \vec{a}_i$ linear abhängig $\Leftrightarrow$ Rang $A < n$.

- In **zwei** Dimensionen $(n = 2)$ ist dieses "Volumen" die Fläche des Parallelogramms, die von den beiden Spaltenvektoren $\begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix}$ und $\begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix}$ aufgespannt wird. Sie kann nach 2.2.5 über das Kreuzprodukt berechnet werden, wenn man eine hypothetische dritte Dimension hinzufügt: $\det(\mathbf{A}) = a_{11}a_{22} - a_{12}a_{21}$

  $$\left(\begin{pmatrix} a_{11} \\ a_{21} \\ 0 \end{pmatrix} \times \begin{pmatrix} a_{12} \\ a_{22} \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ a_{11}a_{22} - a_{12}a_{21} \end{pmatrix} \right)$$

- In **drei** Dimensionen $(n = 3)$ ist das Volumen nach 2.2.6 gegeben durch das Spatprodukt der drei Spaltenvektoren:

  $$\det(\mathbf{A}) = \begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \end{pmatrix} \cdot \left(\begin{pmatrix} a_{12} \\ a_{22} \\ a_{32} \end{pmatrix} \times \begin{pmatrix} a_{13} \\ a_{23} \\ a_{33} \end{pmatrix} \right)$$

  $$\Rightarrow \det(\mathbf{A}) = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33}$$

- **Allgemein** in n Dimensionen: Verallgemeinertes n-dimensionales "Volumen"

  Formale Berechnung:

  $$\det(A) := \sum_P (-1)^P a_{1P_1} a_{2P_2} \cdots a_{nP_n} \tag{2.2}$$

  Hier ist: $(P_1 \cdots P_n)$: Eine Anordnung von $(1, ..., n)$ ("Permutation")

  $\sum_P$: Summe über alle möglichen Permutationen

  $(-1)^P$: $\begin{cases} 1 & \text{, falls Permutation gerade} \\ -1 & \text{, falls Permutation ungerade} \end{cases}$

  wobei - 'gerade' Permutation: lässt sich aus $(1, ...n)$ durch eine gerade Anzahl paarweiser Vertauschungen von Zahlen (Transpositionen) erzeugen.
  - 'ungerade Permutation: ungerade Anzahl Transpositionen

  Es gilt: Zuordnung Permutation $\leftrightarrow$ gerade/ungerade ist eindeutig (ein Ergebnis der Gruppentheorie)

Notation: $\det(\mathbf{A}) = \det\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} = \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}$

2.3.4.2 Rechenregeln für Determinanten

- **Transposition** : $\det(\mathbf{A}^T) = \det(\mathbf{A})$.

  (denn: $\det(\mathbf{A}^T) = \sum_P (-1)^P a_{P_1 1} \cdots a_{P_n n} = \sum_P (-1)^P a_{1 P_1^{-1}} \cdots a_{n P_n^{-1}}$
  $= \sum_{\hat{P}} (-1)^{\hat{P}} a_{1\hat{P}_1^{-1}} \cdots a_{n\hat{P}_n^{-1}}$ mit $\hat{P} = P^{-1}$,
  letzter Schritt folgt aus $(-1)^P = (-1)^{P^{-1}}$.)

- **Addition** einer Zeile/Spalte

$$\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{i1} + b_{i1} & \cdots & a_{in} + b_{in} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{i1} & \cdots & a_{in} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ b_{i1} & \cdots & b_{in} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}.$$

  Spalte analog.

  (Begründung: Faktoren $(a_{ij} + b_{ij})$ in Gl. (2.2) ausmultiplizieren.)

- **Multiplikation** einer Zeile/Spalte mit einer Zahl α

$$\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ \alpha a_{i1} & \cdots & \alpha a_{in} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} = \alpha \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{i1} & \cdots & a_{in} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}.$$

  (Begründung: Faktor α in Gl. (2.2) vor die Summe ziehen.)

  Folgerung $\det(\alpha\mathbf{A}) = \alpha^n \det(\mathbf{A})$.

- **Vertauschung** zweier Zeilen/Spalten ändert Vorzeichen.

  (Begründung bei Spaltenvertauschung $i \leftrightarrow j$: Entspricht in Gl. (2.2) dem Ersetzen der Permutationen P durch $P' = T_{ij}P$, wobei T_{ij} die Transposition $(i \leftrightarrow j)$ ist. Falls P gerade, ist P' ungerade und umgekehrt.)

  Folgerung: Sind zwei Zeilen/Spalten gleich, folgt $\det(\mathbf{A}) = 0$.

  (Begründung: Vertauschung ändert $\mathbf{A}$ nicht, aber Vorzeichen von $\det \mathbf{A}$.)

  → Wenn ein Vielfaches einer Zeile/Spalte auf eine andere addiert wird, ändert sich Determinante nicht.

- **Determinanten-Multiplikationssatz** :

$$\det(\mathbf{AB}) = \det(\mathbf{A}) \det(\mathbf{B})$$

  (ohne Begründung: siehe Mathematik-Vorlesung)

- **Algebraisches Komplement** : $\mathbf{A}^{(ij)}$ sei die Matrix, die aus $\mathbf{A}$ durch Streichen der i-ten Zeile und j-ten Spalte entsteht.
  Algebraisches Komplement: $\mathbf{U}_{ij} = (-1)^{i+j} \det(\mathbf{A}^{(ij)})$.

- **Determinanten-Entwicklungssatz** :

$$\begin{array}{rcll} \det(\mathbf{A}) & = & \sum_i a_{ij}\,\mathbf{U}_{ij} & \text{Entwicklung nach Zeile } i \\ & = & \sum_j a_{ij}\,\mathbf{U}_{ij} & \text{Entwicklung nach Spalte } j \end{array}$$

(ohne Begründung: siehe Mathematik-Vorlesung)

Folgerung: $\sum_k a_{ik}\,\mathbf{U}_{jk} = 0, \quad \sum_k a_{ki}\,\mathbf{U}_{kj} = 0$ für $i \neq k$.

(Begründung: Definiere Matrix $\overline{\mathbf{A}}$, die bis auf idie j-te Zeile identisch ist mit $\mathbf{A}$, nur j-te Zeile durch i-te Zeile ersetzt. Dann ist $\det(\overline{\mathbf{A}}) = 0$.
Die erste Behauptung (Zeile) folgt aus dem Entwicklungssatz:
$\det(\overline{\mathbf{A}}) = \sum_k \overline{a}_{jk}\mathbf{U}_{jk} = \sum_k a_{ik}\mathbf{U}_{jk}$.
Die Begründung der zweiten Behauptung ist analog.)

2.3.4.3 Folgerungen und Anwendungen

- **Inverses** einer $n \times n$-Matrix: Cramers Regel
  Voraussetzung: $\det(\mathbf{A}) \neq 0$
  Dann lässt sich $\mathbf{A}$ invertieren und $[\mathbf{A}^{-1}]_{ij} = \mathbf{U}_{ji} / \det(\mathbf{A})$.
  ($[\mathbf{A}^{-1}]_{ij}$: Eintrag in der i-ten Zeile, j-ten Spalte der Matrix $\mathbf{A}^{-1}$).

- **Lineare Gleichungssysteme** Allgemeine Form: $\mathbf{A}\mathbf{x} = \mathbf{b}$
  mit Koeffizientenmatrix $\mathbf{A} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$,
  und Spaltenvektoren $\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, $\mathbf{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$.

  Lösbarkeit

  - $\mathbf{A}\mathbf{x} = \mathbf{b}$ ist lösbar $\Leftrightarrow$ $\quad$ Rang$(\mathbf{A})$ = Rang$(\mathbf{A}, \mathbf{b})$.
    (denn: Schreibe $\mathbf{A} = (\underline{a}_1 \cdots \underline{a}_n)$ mit $\underline{a}_i = \begin{pmatrix} a_{1i} \\ \vdots \\ a_{mi} \end{pmatrix}$.
    $\mathbf{A}\mathbf{x} = \mathbf{b} \Rightarrow \sum_i \underline{a}_i x_i = \mathbf{b} \Rightarrow \mathbf{b}$ ist Linearkombination der $\underline{a}_i$.)
  - Lösung ist <u>eindeutig</u> $\Leftrightarrow$ $\quad$ Rang$(A) = n$.
    (denn: Rang$(A) = n \Rightarrow$ alle $\underline{a}_i$ linear unabhängig.
    Aus $\sum_i \underline{a}_i x_i = \mathbf{b} = \sum_i \underline{a}_i \lambda_i$ folgt $x_i = \lambda_i \quad \forall\, i$)

  Lösung für $n \times n$-Systeme

  Gleichungssystem $A\underline{x} = \underline{b}$ allgemein lösbar (für alle $\mathbf{b}$),
  wenn $\det(\mathbf{A}) \neq 0 \Leftrightarrow \mathbf{A}$ invertierbar.
  Dann ist $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$.

Cramersche Regel:

Definiere $\mathbf{A}^{(k)}$: Matrix wie $\mathbf{A}$, k-te Spalte ersetzt durch $\mathbf{b}$.
Dann ist Lösung von $\mathbf{Ax} = \mathbf{b}$: $x_k = \det(\mathbf{A}^{(k)})/\det(\mathbf{A})$.
(folgt nach Einsetzen von $[A^{-1}]_{ij} = Uji/\det(\mathbf{A})$ in $x_k = [\mathbf{A}^{-1}\mathbf{b}]_k$).

Aufgaben:

- Berechnen Sie die Determinanten
$$\begin{vmatrix} 4 & 3 & 2 \\ 1 & 0 & -1 \\ 5 & 2 & 2 \end{vmatrix}, \quad \begin{vmatrix} 1 & 6 & 8 & 7 \\ -2 & 3 & 11 & 5 \\ 5 & 0 & 6 & 7 \\ -1 & 9 & 19 & 12 \end{vmatrix}, \quad \begin{vmatrix} 4 & 3 & 0 & 1 \\ 6 & 7 & 8 & -1 \\ 0 & 1 & 0 & 7 \\ 3 & -4 & 0 & 6 \end{vmatrix}$$
- Berechnen Sie $A = \begin{vmatrix} a & b & c & d \\ -b & a & -d & c \\ -c & d & a & -b \\ -d & -c & b & a \end{vmatrix}$. (Tip: Berechnen Sie zuerst AA^T.)
- Zeigen Sie: Für Inverse A^{-1} einer Matrix gilt $\det(A^{-1}) = 1/\det(A)$.
- Berechnen Sie $A = \begin{vmatrix} 1 & a_{12} & a_{13} & \cdots & a_{1n} \\ & 2 & a_{23} & \cdots & a_{2n} \\ & & 3 & \cdots & a_{3n} \\ & & & \ddots & \vdots \\ 0 & & & & n \end{vmatrix}$ für beliebige a_{ij} $(j > i)$.
- Lösen Sie folgende Gleichungssysteme:
$$\begin{array}{rcrcrcr} 2x_1 & + & x_2 & + & 5x_3 & = & -21 \\ x_1 & + & 5x_2 & + & 2x_3 & = & 19 \\ 5x_1 & + & 2x_2 & + & x_3 & = & 2 \end{array} \qquad \begin{array}{rcrcrcr} x_1 & - & x_2 & + & 2x_3 & = & 4 \\ 9x_1 & + & 3x_2 & - & 12x_3 & = & -3 \\ 3x_1 & + & x_2 & - & 4x_3 & = & -1 \end{array}$$

2.3.5 Drehungen und Drehmatrizen

Erinnerung (2.3.1): Koordinaten eines Vektors $\vec{a}$ ändern sich bei Drehung des Koordinatensystems $\Sigma \to \Sigma'$ gemäß $\mathbf{a}' = \mathbf{Da}$ mit $\mathbf{D} = (\mathcal{D}_{ij})$, $\mathcal{D}_{ij} = \vec{e}_i\,' \cdot \vec{e}_j$.

Im Folgenden: Vertiefung der Diskussion von Drehmatrizen.

1) Charakteristika von Drehmatrizen

- **Orthonormal** : $\mathbf{D}^{-1} = \mathbf{D}^T$ bzw. $\mathbf{D}^T\,\mathbf{D} = \mathbf{D}\,\mathbf{D}^T = \mathbb{1}$

  (Begründung: $[\mathbf{D}^T\mathbf{D}]_{ik} = \sum_j \mathcal{D}^T_{ij}\mathcal{D}_{jk} = \sum_j \mathcal{D}_{ji}\mathcal{D}_{jk}$
  $= \sum_j (\vec{e}_i \cdot \vec{e}_j\,')(\vec{e}_j\,' \cdot \vec{e}_k) = \vec{e}_i \cdot \sum_j \vec{e}_j\,'(\vec{e}_j\,' \cdot \vec{e}_k) = \vec{e}_i \cdot \vec{e}_k = \delta_{ik}\ \forall\, \vec{a}$)

  Folgerung für die Struktur von Drehmatrizen:
  Spaltenvektoren $\mathbf{v}_i$ in $\mathbf{D} = (\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ stehen senkrecht aufeinander und $|\mathbf{v}_i| = 1$.
- **Determinante** : $\det(\mathbf{D}) = 1$.

  (Begründung: $\det \mathbf{D} = \pm 1$ folgt aus $\mathbf{D}^T\mathbf{D} = \mathbb{1}$. Vorzeichen + folgt daraus, dass Σ, Σ' beides Rechtssysteme sind $\to$ lassen sich kontinuierlich ineinander überführen.
  Konkret: $\mathbf{D}$ ist parametrisierbar durch drei Winkel, $\mathbf{D}(\phi,\theta,\xi)$ mit $\mathbf{D}(0,0,0) = \mathbb{1} \Rightarrow \det(\mathbf{D}(0,0,0)) = 1$. Drehung $\mathbf{D}$ sind stetige Funktionen dieser Winkel, damit ist auch $\det(\mathbf{D}(\phi,\theta,\xi)$ stetig und springt nicht einfach von $+1$ nach -1 um.)

Bemerkung: Es gibt auch Transformationen $\mathbf{T}$ mit $\mathbf{T}^T\mathbf{T} = \mathbb{1}$ und $\det(\mathbf{T}) = -1$, z.B. Spiegelung am Ursprung, $\mathbf{T} = -\mathbb{1}$. In diesem Fall geht Rechtssystem in Linkssystem über.

2) Einordnung der Drehmatrizen, Drehgruppe

Klassifizierung reeller Matrizen

(i) **Invertierbare $n \times n$-Matrizen** bilden Gruppe $\mathrm{GL}(n, \mathbb{R})$

(ii) Matrizen $\mathbf{U} \in \mathrm{GL}(n, \mathbb{R})$ mit $\mathbf{U}^{-1} = \mathbf{U}^T$ bilden Gruppe: **Orthogonale Gruppe** $\mathrm{O}(n)$

(iii) Matrizen $\mathbf{D} \in \mathrm{O}(n)$ mit $\det(\mathbf{D}) = +1$ bilden Gruppe: **Spezielle orthogonale Gruppe** $\mathrm{SO}(n)$.

Aufgaben:

- Überprüfen Sie: Ist die Matrix $\mathbf{A} = \begin{pmatrix} -\sqrt{2}/2 & 0 & -\sqrt{2}/2 \\ 0 & 1 & 0 \\ \sqrt{2}/2 & 0 & -\sqrt{2}/2 \end{pmatrix}$ eine Drehung?
- Die Vektoren $\vec{a}$ und $\vec{b}$ haben in einem Koordinatensystem Σ die Koordinaten $\mathbf{a} = (0, -2, 0)$ und $\mathbf{b} = (35 - 4)$.

  Betrachten Sie eine Drehung $\mathbf{D} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi & -\sin\varphi \\ 0 & \sin\varphi & \cos\varphi \end{pmatrix}$, die das Koordinatensystem Σ in ein Koordinatensystem Σ' überführt.

  Berechnen Sie die Koordinaten von $\vec{a}$ und $\vec{b}$ im Koordinatensystem Σ'.

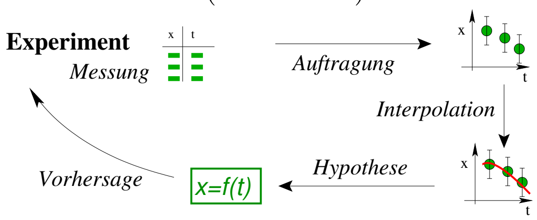
  Berechnen Sie $\vec{a} \cdot \vec{b}$ mit den Koordinaten aus Σ und aus Σ'. Das Ergebnis muss gleich sein. Warum?
- Zeigen Sie: Mit $\mathbf{U} \in O(n)$ und $\mathbf{A}' = \mathbf{U}\mathbf{A}\mathbf{U}^T$ gilt: $\det(\mathbf{A}') = \det(\mathbf{A})$ und $\mathrm{Sp}(\mathbf{A}') = \mathrm{Sp}(\mathbf{A})$.

Kapitel 3

Reelle Funktionen

Bedeutung von Funktionen

- Mathematisch – anderes Word für Abbildung
  z.B. reelle Funktion $f(x) = 1/|x|$
  charakterisiert durch:
  - Definitionsbereich D_f: Erlaubte Werte von x (hier: $\mathbb{R} \setminus \{0\}$)
  - Wertebereich W_f: Menge der Bilder (hier: $\{y \in \mathbb{R} : y > 0\}$)
  - Abbildungsvorschrift
- Physikalisch – Hilfsmittel zur Formulierung von Theorien und zur Beschreibung von Experimenten.
  Empirischer Kreislauf (vereinfacht)



Dieses Kapitel: Zusammenstellung/Erinnerung an einige wichtige reelle Funktionen und Charakterisierung.

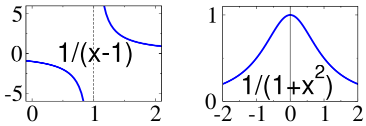
3.1 Elementare Funktionen

3.1.1 Polynome und rationale Funktionen

- Konstante Funktion: $f(x) = a$ (ebene Linie)
- Lineare Funktion: $f(x) = a + bx$ (Gerade)
- Quadratische Funktion: $f(x) = a + bx + cx^2$ (Parabel)
- Polynome $n$ten Grades: $f(x) = \sum_{k=1}^{n} a_k x^k$
- Rationale Funktion: $f(x) = p(x)/q(x)$ wobei $p(x), q(x)$ Polynome sind.

Beispiele

- $f(x) = 1/(x-1)$
- $f(x) = 1/(1+x^2)$
  (Lorentzkurve)



Bemerkungen

a) Polynome kann man ähnlich wie natürliche und Zahlen (ggf. ”mit Rest”) durcheinander teilen. Dabei geht man genauso vor wie beim schriftlichen Dividieren von natürlichen Zahlen:
z.B. $(x^2 + 2x - 1) : (x + 1) = x + 1 - 2/(x + 1)$
$(x^2 + x)$
$(x + 1)$
-2

b) Vergleiche 1.3.1: Die Menge der Polynome bildet einen Ring, und zwar sogar einen ”Hauptidealring”, in dem z.B. der Satz der eindeutigen Primfaktorzerlegung gilt.

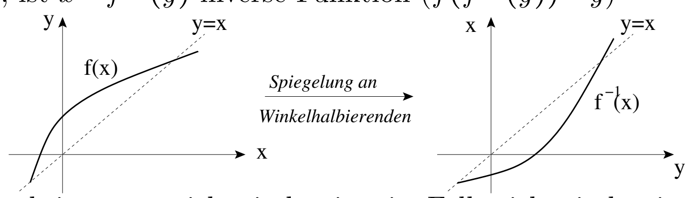
d) Von einem anderen Blickwinkel betrachtet bilden Polynome mit Koeffizienten $a_k \in \mathbb{R}$ einen Vektorraum über $\mathbb{R}$. Eine mögliche Basis von linear unabhängigen Vektoren ist die Menge $\{x^k\}$. Diese Basis hat unendlich viele Elemente, also ist der Vektorraum unendlich dimensional.

Aufgaben zu rationalen Funktionen
- Skizzieren Sie x^3, x^4, $1/x^2$, $1/x^3$, $1/(x-2)^2$, $x/(1+x^2)$
- Berechnen Sie $(x^4 - a^4)/(x - a)$

3.1.2 Algebraische Funktionen

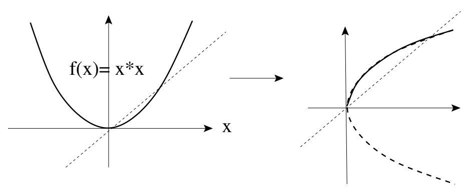
Wurzel: $g(x) = \sqrt{x} \equiv x^{1/2}$ inverse Funktion zu $f(x) = x^2$

Allgemein inverse Funktionen:
Falls $y = f(x)$, ist $x = f^{-1}(y)$ inverse Funktion $(f(f^{-1}(y)) = y)$
Graphisch



NB: Inverse Funktion muss nicht eindeutig sein. Falls nicht eindeutig, muss man sich eine Variante aussuchen.

Speziell Wurzel:
Konvention: $\sqrt{x}$ ist oberer Ast.
NB: Definitionsbereich eingeschränkt $(x \geq 0)$.



Andere rationale Potenzen

z.B. $\sqrt[3]{x} \equiv x^{1/3}$ ist inverse Funktion zu $f(x) = x^3$

$\sqrt[4]{x} \equiv x^{1/4}$ ist inverse Funktion zu $f(x) = x^4$

Zusammengesetzte Ausdrücke

z.B. $\sqrt{x^2 + a^2}/(x-3)^{3/5}$ etc. (mit $y^{3/5} := (y^{1/5})^3$)

Aufgaben zu algebraischen Funktionen / Zahlen

– Skizzieren Sie $1/\sqrt{x}$, $1/\sqrt{x+1}$, $1/\sqrt{x}+1$

– Zeigen Sie folgende Identitäten (Aufgaben aus Indien):

Apastamba, 5.-4. Jhd. v. Christus: $\frac{36}{\sqrt{3}}\,\frac{1}{2}\,(\frac{24}{\sqrt{3}} + \frac{30}{\sqrt{3}}) = 324$

Bhaskara II, 1114-1185: $\sqrt{10 + \sqrt{24} + \sqrt{40} + \sqrt{60}} = \sqrt{2} + \sqrt{3} + \sqrt{5}$.

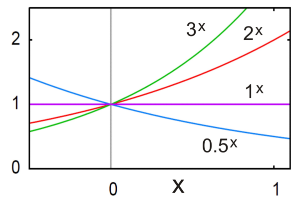
Ramanujan, 1887-1920: $\sqrt[3]{\sqrt[3]{2} - 1} = \sqrt[3]{\frac{1}{9}} - \sqrt[3]{\frac{2}{9}} + \sqrt[3]{\frac{4}{9}}$

3.1.3 Exponentialfunktion

Allgemeine Klasse der Exponentialfunktionen (vgl. 1.3.2)

Funktion b^x für $b \in \mathbb{R}$, $b > 0$ Für jedes b erhält man eigene Kurve.
Es gilt jeweils Multiplikationstheorem

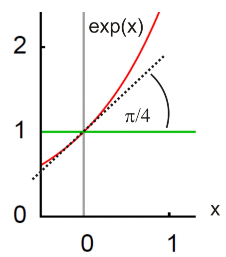
$$\boxed{\begin{array}{l} b^{x+y} = b^x \cdot b^y \\ (b^x)^y = b^{x \cdot y} \end{array}} \quad \text{für } x, y, b \in \mathbb{R},\ b > 0.$$



Natürliche Exponentialfunktion

Spezielle Wahl von b, so dass lokaler Winkel zur y-Achse bei $x = 0$ genau $\pi/4$ ist (45 Grad).
Dies gilt für $\boxed{b = 2.71828\cdots =: \mathrm{e}}$: Euler-Zahl (irrational)

Notation $\boxed{f(x) = \mathrm{e}^x = \exp(x)}$



Besondere Eigenschaften

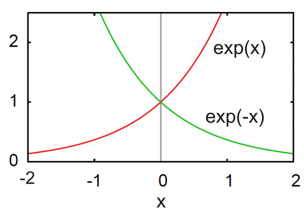
– Ableitung: $f(x) = \exp(x) \Rightarrow f'(x) = f(x)$ (Vorgriff auf Kapitel 4.1)

– Potenzreihe: $\exp(x) = \sum_{k=0}^{\infty} \frac{1}{k!} x^k$ (Vorgriff auf Kapitel 4.2)

– Wachstum: $\exp(x)$ wächst schneller an als jede Potenzfunktion x^n
$(\forall n, b > 0 : \exists x_0 : \mathrm{e}^x > b x^n\ \forall x > x_0)$

Bedeutung der Exponentialfunktion

- ”exponentielles Wachstum”:
  Charakteristisch für ungebremstes Wachstum
  (z.B. Bakterien mit genug Nahrung in Petrischale)
- ”exponentieller Zerfall”: $(1/e)^x = \exp(-x)$
  charakteristisch für freien Zerfall
  (z.B. radioaktiver Zerfall)



Aufgaben zur Exponentialfunktion

Skizzieren Sie: – $f(x) = 1 - \mathrm{e}^{-x}$ (Spannung beim Aufladen eines Kondensators)
– $f(x) = x + \mathrm{e}^{-x}$
– $f(x) = x\,\mathrm{e}^{-x}$ ('Poissonverteilung')
– $f(x) = x^2\,\mathrm{e}^{-x}$
– $f(x) = 1/(\mathrm{e}^x - 1)$ ('Bose-Einstein-Verteilung')
– $f(x) = 1/(\mathrm{e}^x + 1)$ ('Fermi-Dirac-Verteilung')

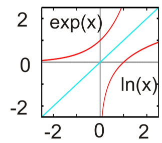
Schauen Sie sich verschiedene Funktionen auf dem Computer an, z.B. Java-Applet "Funktionenschaufenster" im Skript von K. Hefft (www.thphys.uni-heidelberg.de/ hefft/vk1) oder die Seite von Wolfram Alpha (http://www.wolframalpha.com)

3.1.4 Logarithmus

Inverse Funktion zur Exponentialfunktion

Natürlicher Logarithmus

Inverses zur natürlichen Exponentialfunktion
$y = \exp(x) \;\Leftrightarrow\; x = \ln(y) \equiv \log(y)$



Allgemeiner Logarithmus (vgl. 1.3.2)
Inverses zur Funktion $y = b^x \;\Leftrightarrow\; x = \log_b(y)$
"Logarithmus zur Basis b" (mit $b > 0$)
Besonders beliebt (siehe 1.3.2): $b = e$ (siehe oben), $b = 10$ und $b = 2$

Rechenregeln: Aus Multiplikationstheorem $b^x\,b^y = b^{x+y}$, $(b^x)^y = b^{xy}$
folgt für jede Basis:

$$\boxed{\begin{array}{l} \log_b(y\cdot z) = \log_b(y) + \log_b(z) \\ y\log_b(z) = \log_b(z^y) \end{array}}$$

Umrechnung zwischen Exponentialfunktion und Logarithmen:
Rechne als Referenz alles auf natürliche Funktionen um.

$$\boxed{b^x = \mathrm{e}^{x\ln(b)}} \qquad (\text{denn: } b = \mathrm{e}^{\ln b} \Rightarrow b^x = (\mathrm{e}^{\ln b})^x = \mathrm{e}^{x\ln b})$$

$$\boxed{\log_b(y) = \ln(y)\cdot\log_b(e)} \qquad (\ln(y)\log_b(e) = \log_b(\underbrace{\mathrm{e}^{\ln(y)}}_{y}) = \log_b(y))$$

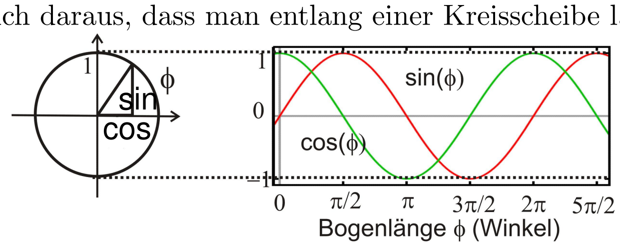
Bemerkung: So wie die Exponentialfunktion schneller als jede Potenzfunktion anwächst, wächst der Logarithmus langsamer als jede Potenzfunktion:
$\forall\epsilon > 0 : \exists x_0 :\ \ln(x) < x^\epsilon\ \forall x > x_0$.

Aufgaben zum Logarithmus

– Zeigen Sie $\log_b(e) = 1/\ln(b)$
(also ist Umrechnungsformel auch: $\log_b(y) = \ln(y)/\ln(b)$)
– Berechnen Sie $\log_b(b)$
– Berechnen Sie $\mathrm{ld}(x) := \log_2(x)$ aus $\ln(x)$.
– Wo steckt der Wurm in folgendem "Beweis" für die Behauptung $2 > 3$.
$1/4 > 1/8 \Rightarrow (1/2)^2 > (1/2)^3 \Rightarrow \ln(1/2)^2 > \ln(1/2)^3$
$\Rightarrow 2\ln(1/2) > 3\ln(1/2) \Rightarrow 2 > 3\,\checkmark$?

3.1.5 Trigonometrische Funktionen

Ergeben sich daraus, dass man entlang einer Kreisscheibe läuft



- Eigenschaften

  periodisch mit Periode 2π: $f(x+2\pi) = f(x)$

  $\cos^2(\phi) + \sin^2(\phi) = 1$ ($\leftrightarrow$ Satz des Pythagoras)
  (Notation: $\sin^2(\phi) = (\sin(\phi))^2$ etc.)

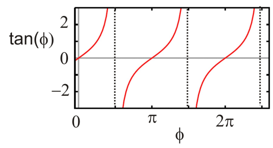
  $\cos(\phi) = \sin(\phi + \pi/2)$

- Additionstheoreme

$$\boxed{\begin{array}{l} \cos(a \pm b) = \cos(a)\ \cos(b) \mp \sin(a)\sin(b) \\ \sin(a \pm b) = \sin(a)\ \cos(b) \pm \cos(a)\sin(b) \end{array}}$$

- Abgeleitete Funktionen

  Tangens: $\tan(\phi) = \sin(\phi)/\cos(\phi)$
  Cotangens: $\cot(\phi) = \cos(\phi)/\sin(\phi)$



- Umkehrfunktionen (zyklometrische Funktionen)

  $\sin(x) = y,\ x \in [-\frac{\pi}{2}, \frac{\pi}{2}] \quad \rightarrow \quad x = \text{arc}\ \sin(y)$
  $\cos(x) = y,\ x \in [0, \pi] \quad \rightarrow \quad x = \text{arc}\ \cos(y)$
  $\tan(x) = y,\ x \in [-\frac{\pi}{2}, \frac{\pi}{2}] \quad \rightarrow \quad x = \text{arc}\ \tan(y)$
  $\cot(x) = y,\ x \in [0, \pi] \quad \rightarrow \quad x = \text{arc}\ \cot(y)$

Aufgaben zu trigonometrischen Funktionen
– Skizzieren Sie $\cot(\phi), \text{arc}\ \tan(\phi),\ \phi + \sin(\phi, \sin(\phi)/\phi$.
– Zeigen Sie: $\sin^2(\phi) = \tan^2(\phi)/(1+\tan^2(\phi))$
– Zeigen Sie: $\sin(x) + \sin(y) = 2\sin(\frac{x+y}{2})\ \cos(\frac{x-y}{2})$

3.1.6 Hyperbolische Funktionen

Aus Exponentialfunktion abgeleitet
Eigenschaften, die an trigonometrische Funktionen erinnern

- Definition

  Sinus hyperbolicus: $\sinh(x) = \frac{1}{2}(\exp(x) - \exp(-x))$
  Cosinus hyperbolicus: $\cosh(x) = \frac{1}{2}(\exp(x) + \exp(-x))$
  Tangens hyperbolicus: $\tanh(x) = \sinh(x)/\cosh(x)$
  Cotangens hyperbolicus: $\coth(x) = \cosh(x)/\sinh(x)$

- Eigenschaften

  $\cosh^2(x) - \sinh^2(x) = 1$

- Additionstheoreme

  $\cosh(x \pm y) = \cosh(x)\ \cosh(y) \pm \sinh(x)\ \sinh(y)$
  $\sinh(x \pm y) = \sinh(x)\ \cosh(y) \pm \cosh(x)\ \sinh(y)$
  (kann durch Einsetzen leicht nachgeprüft werden)

→ Große Ähnlichkeit mit trigonometrischen Funktionen
Deutet auf Zusammenhang zwischen Exponentialfunktion und Sinus/Cosinus hin – wird ersichtlich nach Einführung der komplexen Zahlen.

- Umkehrfunktionen: Area-Funktionen

  $\sinh(x) = y, \qquad \rightarrow \quad x = \operatorname{arsinh}(y)$
  $\cosh(x) = y,\ x \in [0, \infty] \quad \rightarrow \quad x = \operatorname{arcosh}(y)$
  etc.

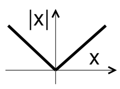
Aufgaben zu hyperbolischen Funktionen
– Skizzieren Sie $\sinh(x), \cosh(x), \tanh(x), \operatorname{arsinh}(x)$
– Beweisen Sie die Additionstheoreme.
– Zeigen Sie: $\operatorname{arcosh}(x) = \ln(x + \sqrt{x^2 - 1})$

3.1.7 Funktionen mit Ecken und Sprüngen

Es gibt vor allem zwei wichtige derartige Funktionen

- **Betragsfunktion** $y = |x|$
  Definition des Betrages: Sei $a \in \mathbb{R}$

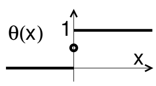
  $$|a| = \begin{cases} a & : & a > 0 \\ -a & : & a < 0 \end{cases}$$



  Rechenregeln:
  $|a \cdot b| = |a| \cdot |b|$
  $\big||a| - |b|\big| \le |a \pm b| \le |a| + |b|$ (Dreiecksungleichung)

- **Heavisidesche Stufenfunktion** $\Theta(x)$

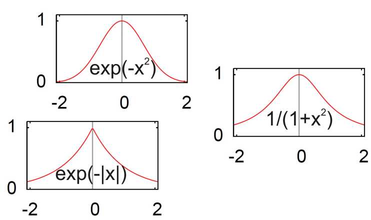
  $$\Theta(x) = \begin{cases} 1 & : & x > 0 \\ 0.5 & : & x = 0 \\ 0 & : & x < 0 \end{cases}$$



  Beispiel: $\Theta(x - a)\ \Theta(b - x)$ (mit $a < b$) gibt Kastenfunktion.

3.1.8 Weitere wichtige abgeleitete Funktionen

- Gaußfunktion: $f(x) = \exp(-x^2)$
- Lorentzfunktion: $f(x) = 1/(1 + x^2)$
- $f(x) = \exp(-|x|/a)$



3.2 Eigenschaften von Funktionen

3.2.1 Spiegelsymmetrie

$y = f(x)$ gerade, wenn $f(-x) = f(x)\ \forall x$ (z.B. x^2)

$y = f(x)$ ungerade, wenn $f(-x) = -f(x)\ \forall x$ (z.B. x^3)

Bemerkung: Man kann jede Funktion in einen geraden und einen ungeraden Anteil zerlegen: $f(x) = f_g(x) + f_u(x)$ mit
$f_g(x) = \frac{1}{2}(f(x) + f(-x),\ f_u(x) = \frac{1}{2}(f(x) - f(-x))$.

3.2.2 Beschränktheit

Funktion heisst nach oben (unten) beschränkt im Intervall $[a,b]$, wenn es eine obere (untere) Schranke für Funktionswerte gibt.

$y = f(x)$ nach$\binom{\text{oben}}{\text{unten}}$beschränkt in $[a,b] \Leftrightarrow \exists B : B\binom{\geq}{\leq}f(x)\ \forall x \in [a,b]$

3.2.3 Monotonie

Funktion heisst monoton steigend (fallend) im Intervall $[a,b]$, wenn mit wachsendem Argument die Funktionswerte steigen (fallen)

$y = f(x)$ monoton$\binom{\text{steigend}}{\text{fallend}}$in $[a,b] \Leftrightarrow \forall x_1 < x_2 \in [a,b] : f(x_1)\binom{\leq}{\geq}f(x_2)$

$y = f(x)$ streng monoton$\binom{\text{steigend}}{\text{fallend}}$in $[a,b] \Leftrightarrow \forall x_1 < x_2 \in [a,b] : f(x_1)\binom{<}{>}f(x_2)$

Anwendung auf einige der Funktionen in 3.1 (im Definitionsbereich)

Funktion	Symmetrie (gerade, ungerade)	Beschränktheit oben	Beschränktheit unten	Monotonie (steigend, fallend)
$\sin(x)$	u	✓	✓	–
$\cos(x)$	g	✓	✓	–
$\exp(x)$	–	–	✓	s
$\log(x)$	–	–	–	s
$\sinh(x)$	u	–	–	s
$\cosh(x)$	g	–	✓	–

3.2.4 Eineindeutigkeit

Der Definitionsbereich D_f sei $[a,b]$. Funktion $f(x)$ ist eineindeutig, wenn zu jedem Punkt im Wertebereich W_f genau ein Punkt im Definitionsbereich D_f gehört.

$y = f(x)$ eineindeutig in $[a,b] \Leftrightarrow \forall y \in W_f : \exists! x \in [a,b] : y = f(x)$

→ dann ist die Umkehrung von $f(x)$ eindeutig ("umkehrbar eindeutig") bzw. $f(x)$ ist surjektiv (vgl. 1.2.1).

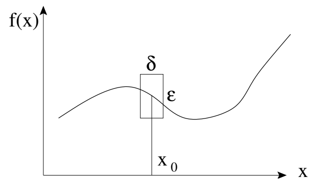
Bemerkung: Streng monotone Funktionen sind eineindeutig.

3.2.5 Stetigkeit

Anschaulich: Keine Sprünge

Mathematisch:

Stetig bei $x_0 \to$ kein Sprung
$\to f(x)$ rückt nahe an $f(x_0)$,
wenn x nahe an x_0 rückt.



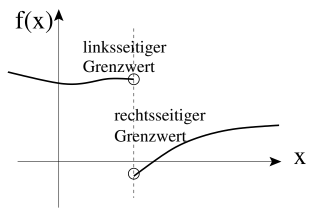
$y = f(x)$ stetig bei $x_0 \;\Rightarrow \forall \epsilon > 0 : \exists \delta > 0 : |f(x) - f(x_0)| < \epsilon \;\forall x \text{ mit } |x - x_0| < \delta.$

Beispiele:
– $\sin(x), \cos(x), \exp(x), \log(x)$ im gesamten Definitionsbereich stetig.
– $1/x$ bei $x = 0$ nicht stetig (divergiert sogar!), sonst überall stetig.
– Heaviside-Funktion bei $x = 0$ nicht stetig.

3.2.6 Grenzwerte

- Grenzwert an vorgegebenen Punkten x_0:

$f(x)$ nähert sich bei Annäherung an x_0 einem Grenzwert y_0 beliebig nahe an, z.B.



Mathematisch: Grenzwert wird durch $\lim_{x\to x_0}$ dargestellt.
Linksseitiger Grenzwert: $\lim_{x\to x_0^-} f(x) = y_0$ bedeutet:
$\forall \epsilon > 0 : \exists \delta > 0 : |f(x) - y_0| < \epsilon \;\forall x \in D_f \text{ mit } 0 < x - x_0 < \delta$
Rechtsseitiger Grenzwert: $\lim_{x\to x_0^+} f(x) = y_0$ bedeutet:
$\forall \epsilon > 0 : \exists \delta > 0 : |f(x) - y_0| < \epsilon \;\forall x \in D_f \text{ mit } 0 < x_0 - x < \delta$
Allgemeiner Grenzwert: $\lim_{x\to x_0} f(x) = y_0$ bedeutet:
$\forall \epsilon > 0 : \exists \delta > 0 : |f(x) - y_0| < \epsilon \;\forall x \in D_f \text{ mit } |x - x_0| < \delta$
(falls Definitionsbereich das Intervall um x_0 einschließt, impliziert das rechtsseitiger Grenzwert = linksseitiger Grenzwert.)

Bemerkungen:

(i) $\lim_{x\to x_0^\pm} = y_0$ bedeutet nicht automatisch $f(x_0) = y_0$
Gegenbeispiel ist Heaviside-Funktion:
$\lim_{x\to 0^-} \theta(x) = 0$, $\lim_{x\to 0^+} \theta(x) = 1$, $\theta(0) = 1/2$.

(ii) Stetigkeit bei x_0 bedeutet: Grenzwert bei x_0 existiert und ist gleich dem Funktionswert

- Grenzwert im Unendlichen ($\pm\infty$):
$f(x)$ nähert sich bei grossen x oder $-x$ beliebig nahe einem Grenzwert an (z.B. $\lim_{x\to\infty} \exp(-x) = 0$)

| $\lim_{x\to\infty} f(x) = y_0 \;\Leftrightarrow\; \forall\epsilon > 0 : \exists\Delta > 0 : |f(x) - y_0| < \epsilon \;\forall x \text{ mit } x > \Delta$
| $\lim_{x\to-\infty} f(x) = y_0 \;\Leftrightarrow\; \forall\epsilon > 0 : \exists\Delta > 0 : |f(x) - y_0| < \epsilon \;\forall x \text{ mit } x < -\Delta$

- Regel von l'Hôpital (ohne Beweis)

  Oft hilfreich zur Bestimmung schwieriger Grenzwerte

  Betrachte Grenzwert der Form $\lim_{x\to x_0} \frac{f(x)}{g(x)}$
  mit $\lim_{x\to x_0} f(x) = \lim_{x\to x_0} g(x) = 0$
  oder $\lim_{x\to x_0} f(x) = \pm\infty$, $\lim_{x\to x_0} g(x) = \pm\infty$.

  Zur Berechnung des Grenzwertes bestimme Ableitungen $f'(x), g'(x)$ von $f(x), g(x)$ (Vorgriff auf Abschnitt 4.1, hoffentlich bekannt aus der Schule!).

  Falls der Grenzwert $\lim_{x\to x_0} \frac{f'(x)}{g'(x)}$ existiert, gilt $\lim_{x\to x_0} \frac{f(x)}{g(x)} = \lim_{x\to x_0} \frac{f'(x)}{g'(x)}$.

Beispiele
– $\lim_{x\to\infty} \frac{1}{1+x^2} = 0$,
– $\lim_{x\to\infty} \frac{x}{1+x^2} = 0$,
– $\lim_{x\to\infty} \frac{x^2}{1+x^2} = 1$,
– $\lim_{x\to\infty} \frac{x^3}{1+x^2} = \infty$ (Grenzwert existiert nicht).
– $\lim_{x\to0} \frac{\sin(x)}{x} = 1$
(Berechnung nach l'Hôpital mit $f(x) = \sin x$, $g(x) = x$:
$\Rightarrow f'(x) = \cos x, f'(x) = 1 \;\Rightarrow \lim_{x\to0} \frac{f'(x)}{g'(x)} = 1 = \lim_{x\to0} \frac{f(x)}{g(x)}$)

NB: Gründe, warum Funktionen eventuell keinen Grenzwert haben:
– $\lim f(x)$ divergiert (wird $\pm\infty$)
– Rechtsseitiger und linksseitiger Grenzwert verschieden.
– Grenzwert nicht eindeutig

Aufgaben
– Berechnen Sie $\lim_{x\to\infty} \frac{1+x}{1-x}$, $\lim_{x\to1} \frac{1+x}{1-x}$, $\lim_{x\to1} \frac{1-x^2}{1-x}$
– Berechnen Sie $\lim_{x\to0}(\tan x)^2$, $\lim_{x\to\infty}(\tan x)^2$, $\lim_{x\to0} \arctan x$, $\lim_{x\to\infty} \arctan x$.
– Berechnen Sie $\lim_{x\to\infty} \tanh x$, $\lim_{x\to0} |x|$ $\lim_{x\to0} \frac{|x|}{x}$, $\lim_{x\to\pi} \frac{\sin x}{x-\pi}$.
– Betrachten Sie Grenzwerte von Funktion $\exp(-1/x)$. bei $x = 0$ und $x = \pm\infty$.
– Skizzieren Sie $\exp(-1/x)$.

Kapitel 4

Infinitesimalrechnung

Rechnen mit dem Unendlichen, Gegenstand der Analysis
Entwicklung war einerseits motiviert aus der Physik.
Macht andererseits moderne Physik überhaupt erst möglich!
Erste Begriffe schon bei Funktionen (Grenzwerte, Stetigkeit)

4.1 Differentialrechnung

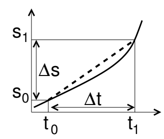
4.1.1 Die Ableitung

1) Differenzenquotion und Differentialquotient

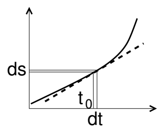
Betrachte als konkretes Beispiel einen Wagen, der entlang einer geraden Straße fährt. Zur Zeit t hat er die Strecke $s(t)$ zurückgelegt.

Frage: Was ist seine Geschwindigkeit?

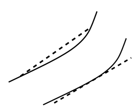
Erste Antwort: Mittlere Geschwindigkeit = Strecke/Zeit
entspricht $\bar{v} = \frac{s(t_1)-s(t_0)}{t_1-t_0} = \frac{\Delta s}{\Delta t}$: Differenzenquotient.
Das ist aber *nicht*, was der Tacho anzeigt.



Zweite Antwort: Momentangeschwindigkeit = Tacho-Wert
→ im Prinzip Differenzenquotient, aber so,
dass die beiden Zeiten t_0, t_1 sehr nahe aneinander sind.
→ Grenzwert $\bar{v} = \lim_{\Delta t \to 0} \frac{\Delta s}{\Delta t} =: \left.\frac{\mathrm{d}s}{\mathrm{d}t}\right|_{t_0}$: Differentialquotient.
Leibniz-Schreibweise



Geometrische Interpretation:
Differenzenquotient: Sekante
Differentialquotient: Tangente



Allgemein: Gegeben Funktion $f(x)$ einer reellen Variablen x.
→ **Differentialquotient:**

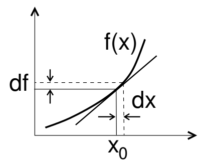
$$\boxed{\left.\frac{\mathrm{d}f}{\mathrm{d}x}\right|_{x_0} = \lim_{\Delta x \to 0} \frac{f(x_0+\Delta x)-f(x_0)}{\Delta x}}$$

heißt **Ableitung** nach x.

Alternative Schreibweisen: $\left.\frac{\mathrm{d}f}{\mathrm{d}x}\right|_{x_0} = \left.\frac{\mathrm{d}}{\mathrm{d}x}f(x)\right|_{x_0} = f'(x_0)$ (Strich)

Speziell Ableitung nach der <u>Zeit</u> t: $\left.\frac{\mathrm{d}f}{\mathrm{d}t}\right|_{t_0} = \dot{f}(t_0)$ (<u>Punkt</u>)

2) Differential: Alternative Sichtweise



Betrachte Funktion $f(x)$ am Punkt x_0.
Schätze ab $f(x_0 + \Delta x)$:
→ Zuwachs $f(x_0 + \Delta x) = f(x_0) + \Delta f(x)$
$\approx f(x_0) + f'(x_0)\,\Delta x + \text{Rest}.$

Rest verschwindet (auch relativ zum Zuwachs), für kleine $\Delta x \to 0$.

Im Grenzwert $\Delta x \to 0$ schreibt man

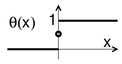
$$\boxed{f(x_0 + \mathrm{d}x) = f(x_0) + \left.\mathrm{d}f(x)\right|_{x_0}} \text{ mit } \boxed{\left.\mathrm{d}f(x)\right|_{x_0} = f'(x_0)\,\mathrm{d}x} = \left.\frac{\mathrm{d}f}{\mathrm{d}x}\right|_{x_0}\mathrm{d}x.$$

sieht formal aus wie "kürzen" ($\frac{\mathrm{d}f}{\mathrm{d}x}\mathrm{d}x = \mathrm{d}f$).

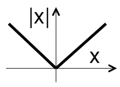
NB: Kürzen darf man bei Differentialquotienten natürlich eigentlich nicht. Trotzdem machen Physiker davon ausgiebig Gebrauch. Hintergrund ist genau dieses Denken in Differentialen, also: $\mathrm{d}f$ entspricht einem realen <u>Zuwachs</u> von f (also einer echten, sehr kleinen Zahl).

3) Differenzierbarkeit

Grenzwert $\lim_{\Delta x \to 0} \frac{\Delta f}{\Delta x}$ muß existiert nicht immer.

Gegenbeispiele: 

(unstetige Funktion)



(Funktion mit Knick)

Wenn der Grenzwert existiert, dann heißt die Funktion **differenzierbar**.

$$f(x) \text{ differenzierbar bei } x_0 \Leftrightarrow \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} \text{ existiert.}$$

4) Verallgemeinerung auf Funktionen mehrere Variablen

Gegeben sei z.B. eine Funktion von zwei Variablen $f(x, y)$.

→ Man kann ohne weiteres nach einer der Variablen, x oder y, ableiten und die andere dabei festhalten. Man muss nur festlegen, welcher. Das nennt man dann **partielle Ableitung**

<u>Notation</u>

Ableitung nach x: $\lim_{x \to x_0} \frac{f(x,y_0) - f(x_0,y_0)}{x - x_0} = \left.\frac{\partial f}{\partial x}\right|_{x_0,y_0}$

Ableitung nach y: $\lim_{y \to y_0} \frac{f(x_0,y) - f(x_0,y_0)}{y - y_0} = \left.\frac{\partial f}{\partial y}\right|_{x_0,y_0}$

Geschwungene Zeichen ∂ sagen aus:
Achtung, hier gibt es noch weitere Variablen.
→ Physiker dürfen nicht mehr ohne weiteres kürzen !!!

Abgesehen davon ist die partielle Ableitung nichts grundsätzlich anderes als eine "normale" Ableitung.

5) Verallgemeinerung: Höhere Ableitungen

Gegeben Funktion $f(x)$

Ableitung $f'(x)$: Neue Funktion, kann man evtl. wieder ableiten
→ Zweite Ableitung $f''(x) = \frac{\mathrm{d}}{\mathrm{d}x}\frac{\mathrm{d}f}{\mathrm{d}x} = \frac{\mathrm{d}^2 f}{\mathrm{d}x^2}$.

Beispiele: Ort $s(t)$ (Wagen)
Geschwindigkeit $v(t) = \frac{\mathrm{d}s}{\mathrm{d}t} = \dot{s}(t)$
Beschleunigung $a(t) = \frac{\mathrm{d}v}{\mathrm{d}t} = \dot{v}(t) = \frac{\mathrm{d}^2 s}{\mathrm{d}t^2} = \ddot{s}(t)$

Allgemein $n$te Ableitung: $f^{(n)}(x) = \frac{\mathrm{d}}{\mathrm{d}x} f^{(n-1)}(x)$

Notation: $f^{(n)}(x) = \frac{\mathrm{d}^n f}{\mathrm{d}x^n} = (\frac{\mathrm{d}}{\mathrm{d}x})^n f(x)$.

6) Mittelwertsatz der Differentialrechnung

Zum Abschluss ein nützlicher Satz für reelle Funktionen:

Ist $f(x)$ stetig im Intervall $[a,b]$ und differenzierbar in $]a,b[$, dann folgt: Es existiert ein Wert $x_0 \in]a,b[$, so dass $\frac{f(b)-f(a)}{b-a} = f'(x_0)$.

4.1.2 Elementare Beispiele

Explizite Berechnung der Ableitung einiger elementarer Funktionen. Aus diesen kann man später die meisten übrigen Ableitungen herleiten.

1) Potenzen: $f(x) = x^n \;\Rightarrow\; f'(x) = nx^{n-1}$

Rechnung:

$$f'(x) = \lim_{\Delta x\to 0} \frac{f(x+\Delta x)-f(x)}{\Delta x} = \lim_{\Delta x\to 0} \frac{1}{\Delta x}[(x+\Delta x)^n - x^n]$$

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

$$f'(x) = \lim_{\Delta x\to 0} \frac{1}{\Delta x}[x^n + nx^{n-1}\Delta x + \binom{n}{2}x^{n-2}(\Delta x)^2 + \dots - x^n]$$
$$= \lim_{\Delta x\to 0}[nx^{n-1} + \binom{n}{2}x^{n-2}\Delta x + \mathcal{O}((\Delta x)^2)] = nx^{n-1} \quad ✓$$

2) Exponentialfunktion: $f(x) = \mathrm{e}^x \;\Rightarrow\; f'(x) = \mathrm{e}^x$

Rechnung:

$$f'(x) = \lim_{\Delta x\to 0} \frac{1}{\Delta x}[\mathrm{e}^{x+\Delta x} - \mathrm{e}^x] = \lim_{\Delta x\to 0} \frac{1}{\Delta x}\mathrm{e}^x(\mathrm{e}^{\Delta x} - 1)$$
$$= \mathrm{e}^x \lim_{\Delta x\to 0} \frac{\mathrm{e}^{\Delta x}-1}{\Delta x} = \mathrm{e}^x \lim_{\Delta x\to 0} \frac{\mathrm{e}^{0+\Delta x}-\mathrm{e}^0}{\Delta x}$$
$$= \mathrm{e}^x \cdot [\text{Steigung von } \mathrm{e}^x \text{ bei } x = 0]$$

Aber: Per Konstruktion hat e^x bei $x = 0$ die Steigung 1
(e^x schneidet die y-Achse im Winkel $\frac{\pi}{4}$).

$\Rightarrow f'(x) = \mathrm{e}^x f'(0) = \mathrm{e}^x \cdot 1 = \mathrm{e}^x$ ✓

3) Trigonometrische Funktionen:

$f(x) = \sin(x) \;\Rightarrow\; f'(x) = \cos(x)$
$f(x) = \cos(x) \;\Rightarrow\; f'(x) = -\sin(x)$

Rechnung zu $f(x) = \sin(x)$

$$f'(x) = \lim_{\Delta x \to 0} \frac{1}{\Delta x}[\sin(x + \Delta x) - \sin(x)]$$

| Additionstheorem : $\sin(a+b) = \sin(a)\cos(b) + \cos(a)\sin(b)$

$$f'(x) = \lim_{\Delta x \to 0} \frac{1}{\Delta x}[\sin(x)\cos(\Delta x) + \cos(x)\sin(\Delta x) - \sin(x)]$$

$$= \sin(x) \lim_{\Delta x \to 0} \frac{\cos(\Delta x) - 1}{\Delta x} + \cos(x) \lim_{\Delta x \to 0} \frac{\sin(\Delta x)}{\Delta x}$$

| Es gilt: $\cos(\Delta x) - 1 = -2\sin^2(\frac{\Delta x}{2})$. Setze $\Delta x' = \Delta x/2$

$$= -\sin(x) \lim_{\Delta x' \to 0} \frac{\sin^2(\Delta x')}{\Delta x'} + \cos(x) \lim_{\Delta x \to 0} (\frac{\sin(\Delta x)}{\Delta x}$$

| Es gilt: $\lim_{\phi \to 0} \frac{\sin(\phi)}{\phi} = 1$, $\lim_{\phi \to 0} \frac{\sin^2(\phi)}{\phi} = 0$ (Beweis siehe unten)

$$\Rightarrow f'(x) = \cos(x) \quad \checkmark$$

Rechnung zu $f(x) = \cos(x)$

$$f'(x) = \lim_{\Delta x \to 0} \frac{1}{\Delta x}[\cos(x + \Delta x) - \cos(x)]$$

| $\cos(x) = \sin(x + \frac{\pi}{2})$

$$f'(x) = \lim_{\Delta x \to 0} \frac{1}{\Delta x}[\sin(x + \frac{\pi}{2} + \Delta x) - \sin(x + \frac{\pi}{2})]$$

| Substituiere $x' = x + \frac{\pi}{2}$

$$f'(x) = \lim_{\Delta x \to 0} \frac{1}{\Delta x}[\sin(x' + \Delta x) - \sin(x')]$$

| Setze Ergebnis der ersten Rechnung ein.

$$f'(x) = \cos(x') = \cos(x - \frac{\pi}{2})$$

| Substituiere zurück ($x = x' - \frac{\pi}{2}$)

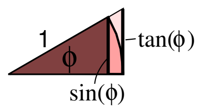
$$f'(x) = \cos(x - \frac{\pi}{2})$$

| $\cos(x - \frac{\pi}{2}) = \cos(\frac{\pi}{2} - x) = \sin(-x) = -\sin(x)$

$$\Rightarrow f'(x) = -\sin(x) \quad \checkmark$$

Nachtrag Beweis $\lim_{\phi \to 0} \frac{\sin(\phi)}{\phi} = 1$, $\lim_{\phi \to 0} \frac{\sin^2(\phi)}{\phi} = 0$
(ohne Verwendung von Ableitungen, sonst beißt sich die Katze in den Schwanz.)

Beweis nach l'Hôpital zunächst für $\phi > 0$
Betrachte Kreissegment mit einem einbeschriebenem
und einem umfassenden rechtwinkligen Dreieck



- Fläche kleines Dreieck: $\frac{1}{2}\cos(\phi)\sin(\phi)$
- Fläche Kreissegment: $\frac{\phi}{2\pi}\pi = \frac{1}{2}\phi$
- Fläche großes Dreieck: $\frac{1}{2}\tan(\phi)$

$$\Rightarrow \frac{1}{2}\cos(\phi)\sin(\phi) \leq \frac{1}{2}\phi \leq \frac{1}{2}\tan(\phi) (= \frac{1}{2}\sin(\phi)/\cos(\phi))$$

$$\Rightarrow \cos(\phi) \leq \phi/\sin(\phi) \leq 1/\cos(\phi) \Rightarrow 1/\cos(\phi) \geq \sin(\phi)/\phi \geq \cos(\phi)$$

$$\Rightarrow \underbrace{\lim_{\phi \to 0^+} 1/\cos(\phi)}_{1} \geq \lim_{\phi \to 0^+} \sin(\phi)/\phi \geq \underbrace{\lim_{\phi \to 0^+} \cos(\phi)}_{1}$$

$$\Rightarrow \lim_{\phi \to 0^+} \sin(\phi)/\phi = 1 \text{ und } \lim_{\phi \to 0^+} \sin^2(\phi)/\phi = 0$$

Beweis für $\phi < 0$ geht analog.

4.1.3 Differentiationsregeln

Von den elementaren Funktionen aus Abschnitt 3.1 haben wir die Potenzen x^n und die Exponentialfunktion e^x abgeleitet.

Es fehlen noch: Polynome, rationale Funktionen, Wurzeln, Logarithmen, etc.

Diese können jedoch aus den bekannten Ableitungen 4.1.2 unter Benutzung allgemeiner Regeln hergeleitet werden.

1) Linearität $\boxed{f(x) = a\,g(x) + b\,h(x) \;\Rightarrow\; f'(x) = a\,g'(x) + b\,h'(x).}$

Denn: $f'(x) = \lim_{\Delta x\to 0} \frac{1}{\Delta x}[ag(x+\Delta x) + bh(x+\Delta x) - ag(x) - bh(x)]$
$= a\lim_{\Delta x\to 0} \frac{1}{\Delta x}[g(x+\Delta x) - g(x)] + b\lim_{\Delta x\to 0} \frac{1}{\Delta x}[h(x+\Delta x) - h(x)]$
$= ag'(x) + bh'(x) \quad \checkmark$

2) Produktregel $\boxed{f(x) = g(x)\,h(x) \;\Rightarrow\; f'(x) = g'(x)\,h(x) + g(x)\,h'(x).}$

Denn: $f'(x) = \lim_{\Delta x\to 0} \frac{1}{\Delta x}[g(x+\Delta x)h(x+\Delta x) - g(x)h(x)]$
$= \lim_{\Delta x\to 0} \frac{1}{\Delta x}[(g(x+\Delta x) - g(x))h(x+\Delta x) + g(x)(h(x+\Delta x) - h(x))]$
$= \lim_{\Delta x\to 0} \underbrace{\frac{g(x+\Delta x)-g(x)}{\Delta x}}_{\to g'(x)} \underbrace{h(x+\Delta x)}_{\to h(x)} + g(x) \underbrace{\frac{h(x+\Delta x)-h(x)}{\Delta x}}_{\to h'(x)}]$
$= g'(x)\,h(x) + g(x)\,h'(x) \quad \checkmark$

3) Inversenregel $\boxed{f(x) = 1/g(x) \;\Rightarrow\; f'(x) = -g'(x)/g(x)^2}$

Denn: $f'(x) = \lim_{\Delta x\to 0} \frac{1}{\Delta x}\left[\frac{1}{g(x+\Delta x)} - \frac{1}{g(x)}\right] = \lim_{\Delta x\to 0} \frac{1}{\Delta x} \frac{g(x)-g(x+\Delta x)}{g(x+\Delta x)\,g(x)}$
$= -\lim_{\Delta x\to 0} \frac{1}{\Delta x} \underbrace{\frac{g(x+\Delta x)-g(x)}{\Delta x}}_{\to g'(x)} \underbrace{\frac{1}{g(x+\Delta x)g(x)}}_{\to 1/g(x)^2} = -\frac{g'(x)}{g(x)^2} \quad \checkmark$

4) Quotientenregel $\boxed{f(x) = g(x)/h(x) \;\Rightarrow\; f'(x) = \frac{g'(x)\,h(x) - g(x)\,h'(x)}{h(x)^2}}$

Ergibt sich aus Kombination von Produktregel und Quotientenregel:
$\frac{\mathrm{d}}{\mathrm{d}x}(\frac{g}{h}) = \frac{\mathrm{d}}{\mathrm{d}x}(g\cdot\frac{1}{h}) = g \underbrace{\frac{\mathrm{d}}{\mathrm{d}x}\frac{1}{h}}_{-h'/h^2} + (\frac{\mathrm{d}}{\mathrm{d}x}g)\frac{1}{h} = \frac{1}{h^2}(-gh' + g'h) \quad \checkmark$

5) Kettenregel $\boxed{f(x) = f(g(x)) \;\Rightarrow\; f'(x) = f'(g)\,g'(x).}$

Leibniz-Schreibweise: $\boxed{\frac{\mathrm{d}f}{\mathrm{d}x} = \frac{\mathrm{d}f}{\mathrm{d}g}\frac{\mathrm{d}g}{\mathrm{d}x}.}$ → Wieder ein Fall von "kürzen".

Beweisskizze mit Differentialen:
$z = f(y) \;\Rightarrow\; \Delta z = \Delta f(y) = f'(y)\Delta y + \text{Rest}$
$y = g(x) \;\Rightarrow\; \Delta y = \Delta g(x) = g'(x)\Delta x + \text{Rest} \;\Rightarrow \Delta x = \frac{\Delta y}{g'(y)} - \frac{\text{Rest}}{g'(x)}$
$\Rightarrow \frac{\Delta z}{\Delta x} \xrightarrow[\text{Reste verschwinden}]{\Delta x,\Delta y\to 0} \frac{f'(y)\Delta y}{\Delta y/g'(x)} = f'(y)g'(x) \quad \checkmark$

6) Umkehrfunktionsregel

Sei $y = f(x)$ differenzierbar und umkehrbar → $x = g(y)$.

Dann ist $g(y)$ differenzierbar und $\boxed{g'(y) = 1/f'(x)\big|_{x=g(y)}}$.

Leibniz-Schreibweise: $f'(x) = \frac{\mathrm{d}y}{\mathrm{d}x} \Rightarrow g'(y) = \frac{\mathrm{d}x}{\mathrm{d}y} = 1/(\frac{\mathrm{d}y}{\mathrm{d}x} = 1/f'(x).$

→ Wieder so ähnlich wie "Kürzen".

Beweisskizze ohne Kürzen:
$x = g(f(x)) \Rightarrow \frac{\mathrm{d}}{\mathrm{d}x} = 1 \& \frac{\mathrm{d}}{\mathrm{d}x} = g(f(x)) \underset{\text{Kettenregel}}{=} g'(f(x))\cdot f'(x)$
$\Rightarrow 1 = g'f' \;\Rightarrow\; g' = 1/f' \quad \checkmark$

4.1.4 Anwendungen der Differentiationsregeln

Mit Hilfe der Differentiationsregeln können aus den elementaren Ableitungen von Abschnitt 4.1.3 die Ableitungen fast aller übrigen Funktionen berechnet werden, zum Beispiel:

1) Polynome: $P(x) = \sum_{k=0}^{n} a_k x^k = a_0 + a_1 x + a_2 x^2 + \cdots$
über $\frac{\mathrm{d}}{\mathrm{d}x} x^k = k x^{k-1}$ und Linearitätsregel
$\Rightarrow\ P'(x) = \sum_{k=0}^{n} a_k k x^{k-1} = a_1 + 2a_2 x + \cdots$

2) Rationale Funktionen: $f(x) = P(x)/Q(x)$ mit Polynomen P und Q
über 1) und Quotientenregel

3) Wurzeln: $f(x) = \sqrt[n]{x} = x^{1/n}$
über $\frac{\mathrm{d}}{\mathrm{d}x} x^n = n x^{n-1}$ und Umkehrfunktionsregel $y = \sqrt[n]{x} \Rightarrow x = y^n$
$\Rightarrow\ f'(x) = \frac{\mathrm{d}y}{\mathrm{d}x} = 1/\frac{\mathrm{d}x}{\mathrm{d}y} = 1/(n y^{n-1}) = \frac{1}{n} \frac{1}{x^{\frac{1}{n}(n-1)}} = \frac{1}{n} x^{\frac{1}{n}-1}.$

4) Natürlicher Logarithmus: $f(x) = \ln(x)$
über $\frac{\mathrm{d}}{\mathrm{d}x} \mathrm{e}^x = \mathrm{e}^x$ und Umkehrfunktionsregel $y = \ln(x) \Rightarrow x = \mathrm{e}^y$
$\Rightarrow\ f'(x) = 1/\frac{\mathrm{d}x}{\mathrm{d}y} = 1/\mathrm{e}^y = 1/\mathrm{e}^{\ln x} = 1/x.$

5) Allgemeine Potenz: $f(x) = x^\alpha,\ \alpha \in \mathbb{R}$
über $x^\alpha = \mathrm{e}^{\alpha \ln(x)}$ und Kettenregel $f(x) = \mathrm{e}^y$ mit $y = \alpha \ln x$
$\Rightarrow\ f'(x) = \frac{\mathrm{d}f}{\mathrm{d}y} \frac{\mathrm{d}y}{\mathrm{d}x} = \mathrm{e}^y \alpha \frac{1}{x} = \alpha x^\alpha \frac{1}{x} = \alpha\, x^{\alpha-1}$

4.1.5 Tabelle wichtiger Ableitungen

$f(x)$	$f'(x)$	Einschränkungen/Anmerkungen
const.	0	
x^α	$\alpha x^{\alpha-1}$	$\alpha \in \mathbb{R}$
$\exp(x)$	$\exp(x)$	
$\ln(\lvert x\rvert)$	$1/x$	$x \neq 0$
r^x	$r^x \ln(r)$	$0 < r \in \mathbb{R}$
$\log_b(\lvert x\rvert)$	$1/(x\ln(b))$	$0 < b \in \mathbb{R}, x \neq 0, b \neq 1$
$\sin(x)$	$\cos(x)$	
$\cos(x)$	$-\sin(x)$	
$\tan(x)$	$1/\cos^2(x)$	$x \neq (z+1/2)\pi$ für $z \in \mathbb{Z}$
$\cot(x)$	$-1/\sin^2(x)$	$x \neq z\pi$ für $z \in \mathbb{Z}$
$\arcsin(x)$	$1/\sqrt{1-x^2}$	$\lvert x\rvert < 1$ (NB: $-\pi/2 < \arcsin(x) < \pi/2$, $\lvert x\rvert < 1$)
$\arccos(x)$	$-1/\sqrt{1-x^2}$	$\lvert x\rvert < 1$ (NB: $0 < \arccos(x) < \pi$, $\lvert x\rvert < 1$)
$\arctan(x)$	$1/(1+x^2)$	$-\pi/2 < \arctan(x) < \pi/2)$
$\mathrm{arccot}(x)$	$-1/(1+x^2)$	$0 < \mathrm{arccot}(x) < \pi)$
$\sinh(x)$	$\cosh(x)$	
$\cosh(x)$	$\sinh(x)$	
$\tanh(x)$	$1/\cosh^2(x)$	
$\coth(x)$	$-1/\sinh^2(x)$	
$\mathrm{arsinh}(x)$	$1/\sqrt{1+x^2}$	
$\mathrm{arcosh}(x)$	$1/\sqrt{x^2-1}$	$x > 1$ (NB: $0 < \mathrm{arcosh}(x)$, $x > 1$)
$\mathrm{artanh}(x)$	$1/(1-x^2)$	$\lvert x\rvert < 1$
$\mathrm{arcoth}(x)$	$-1/(x^2-1)$	$\lvert x\rvert > 1$

Aufgaben

Überprüfen Sie einige der Ableitungen in obiger Tabelle, z.B.
$\tan(x), \cot(x), \arcsin(x), \arccos(x), \arctan(x), \mathrm{arccot}(x),$
$\sinh(x), \cosh(x), \tanh(x), \coth(x), \mathrm{arsinh}(x), \mathrm{arcosh}(x), \mathrm{artanh}(x), \mathrm{arcoth}(x)$

Berechnen Sie die Ableitungen von $b^x, \log_b(x), x^x$.

Berechnen Sie einige der folgenden Ableitungen:
$\sin^3(4x)$, $\exp(-(x/a)^2)$, $1/\sqrt{ax^2+b}$, $\ln(3\mathrm{e}^{2x})$, $\sqrt{1+\sqrt{x}}$, $\ln(x)\ \ln(\ln(x)) - \ln(x)$.
$a\cosh(\frac{x-b}{a})$, $ax^2\mathrm{e}^{-bx}$, $1/(1+(\frac{x}{a})^2)$, $\left(\frac{\sin(x/a)}{x/a}\right)^2$

Berechnen Sie folgende Ableitungen:
$\frac{\mathrm{d}^2}{\mathrm{d}t^2}\mathrm{artanh}\left(\frac{\sqrt{\cosh^2(\omega t)-1}}{\cosh(\omega t)}\right)$, $\frac{\mathrm{d}^2}{\mathrm{d}a^2}f(g(a))$, $\frac{\mathrm{d}^2}{\mathrm{d}x^2}\left(\exp[(\alpha x^2+\beta x+\gamma)^2]-A\right)$, $\frac{\mathrm{d}}{\mathrm{d}t}\ln\sqrt{f(\omega t)-xt}$

Berechnen Sie folgende Grenzwerte mit Hilfe der Regel von l'Hôpital (siehe 3.2.6):
$\lim_{x\to 0}\frac{\sin x^2}{x}$, $\lim_{x\to 0}\frac{x}{\mathrm{e}^x-1}$, $\lim_{x\to 0}\frac{\arcsin(7x)}{\tan(x)}$, $\lim_{x\to 0}\frac{a^x-1}{x}$
$\lim_{x\to\infty}\frac{\ln x}{x^a}$, $\lim_{x\to 0}\frac{\tanh(x)}{\sin(x)}$, $\lim_{x\to\infty}\frac{\ln(x)}{x^a}$, $\lim_{x\to 2\pi}\frac{\ln(\cos x)}{1-\exp(2\pi-x)}$,

Berechnen Sie folgende partiellen Ableitungen
$\frac{\partial}{\partial x}(x+y+z)$, $\frac{\partial}{\partial y}(x^2+y^2+z^2)$, $\frac{\partial}{\partial x}(xyz)$, $\frac{\partial}{\partial x}\frac{1}{\sqrt{x^2+y^2+z^2}}$, $\lim_{x\to 0}\frac{\partial}{\partial x}\frac{2xy}{x^2+y^2}$

Betrachten Sie die sogenannte Cobb-Douglas Produktionsfunktion $q(p_1, p_2) = Ap_1^\alpha p_2^{1-\alpha}$
Zeigen Sie: Es gilt $q(x_1, x_2) = \frac{\partial q}{\partial x_1}x_1 + \frac{\partial q}{\partial x_2}x_2$.

4.2 Taylor-Entwicklung

Eine wichtige Anwendung der Differentiation:

Darstellung von Funktionen als **Potenzreihen**: $\boxed{f(x) = \sum_{k=0}^{\infty} a_k x^k}$

(Alternative Darstellung unendlich oft differenzierbarer Funktionen)

Beispiele für eine Darstellung als Potenzreihe

- Geometrische Reihe $f(x) = 1/(1-x) = \sum_k x^k$ (bereits bekannt!)
- Exponentialfunktion $f(x) = \exp(x) = \sum_k \frac{1}{k!} x^k$ (wird unten gezeigt)

Fragen:

(i) Was bringt das?

Beispiel Exponentialfunktion

– Man kann sie überhaupt erst mal ausrechnen!

– Man kann sie nähern, wenn's nicht genau sein muss.

(z.B. $\exp(x) \approx 1 + x$)

(ii) Wann geht das? (d.h., wann gibt es eine Potenzreihendarstellung?)

(iii) Wie kommt man an die Koeffizienten a_k?

(iv) Praktisch gesehen: Wie rechnet man mit solchen Reihen?

(Addition, Multiplikation, Differentiation)

Werden von rückwärts beantwortet.

Zuerst: Allgemeine Eigenschaften von (Potenz)reihen

4.2.1 Allgemein: Folgen und Reihen

4.2.1.1 Folgen

Formal: Folge ist eine unendliche Menge von durchnummerierten Zahlen

$(a_0, a_1, a_2, a_3, \cdots) \equiv (a_n)_{n\in\mathbb{N}_0}$ mit einem Bildungsgesetz

(oder: Abbildung von $\mathbb{N}_0$ (manchmal auch $\mathbb{N}$) in $\mathbb{R}$ oder $\mathbb{C}$)

Beispiele:

(a)	$(0, 1, 2, 3, \cdots)$	$(a_n = n)$	
(b)	$(1, -1, 1, -1, \cdots)$	$(a_n = (-1)^n)$	
(c)	$(1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \cdots)$	$(a_n = \frac{1}{n}\ n \in \mathbb{N})$	harmonische Folge
(d)	$(0, \frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \cdots)$	$(a_n = \frac{n}{n+1})$	
(e)	$(1, 1, \frac{1}{2}, \frac{1}{6}, \frac{1}{24}, \cdots)$	$(a_n = \frac{1}{n!})$	(Konvention: $0! = 1$)
(f)	$(1, q, q^2, q^3, q^4, \cdots)$	$(a_n = q^n)$	geometrische Folge
(g)	$(1, 1, 2, 3, 5, 8, \cdots)$	$(a_n = a_{n-1} + a_{n-2})$	Fibonacci-Folge

zu (f): Beschreibt Zinseszinsentwicklung

z.B. 3% Zins → nach 1 Jahr Vermehrung um Faktor 1.03

nach 2 Jahren Vermehrung um Faktor $(1.03)^2$

nach 3 Jahren Vermehrung um Faktor $(1.03)^3$

zu (g): Berühmte Folge. Sollte ursprünglich Kaninchenwachstum beschreiben. Heute u.a. in der Kryptographie benutzt.

Charakterisierung von Folgen (Eigenschaften) – ähnlich Funktionen (3.2)

- **Beschränktheit**
  (a_n) beschränkt $\Leftrightarrow \quad \exists B : |a_n| \le B \; \forall n \in \mathbb{N}_0$
  Analog: nach oben / nach unten beschränkt (für reelle Folgen)
- **Monotonie** (nur für reelle Folgen)
  (a_n) monoton steigend $\Leftrightarrow \; a_n \le a_{n+1} \; \forall n$
  (a_n) streng monoton steigend $\Leftrightarrow \; a_n < a_{n+1} \; \forall n$
  analog monoton / streng monoton fallend
- **Konvergenz**
  Folge ist konvergent, wenn sie einen Grenzwert bei $n \to \infty$ hat
  $\lim_{n\to\infty} a_n = a \;\Leftrightarrow\; \forall \epsilon > 0 : \; \exists N \ge 0 : \; |a_n - a| < \epsilon \; \forall n > N$
  Alternatives, äquivalentes Kriterium: **Cauchy-Kriterium**
  (a_n) konvergent $\;\Leftrightarrow\; \forall \epsilon > 0 : \; \exists N \ge 0 : \; |a_n - a_m| < \epsilon \; \forall n, m > N$
  (Vorteil: Leichter zu überprüfen, falls Grenzwert nicht bekannt.)

Nützliche Sätze (anschaulich klar: Beweis hier weggelassen)
- Jede konvergente Folge ist beschränkt.
- Eine Folge, die gleichzeitig monoton und beschränkt ist, hat einen Grenzwert.

Anwendung auf unsere Beispiele:

	beschränkt	monoton	Grenzwert
(a)	nach unten	steigend	–
(b)	✓	–	–
(c)	✓	fallend (streng)	0
(d)	✓	steigend (streng)	1
(e)	✓	fallend (nicht streng)	0
(f) $(q < 1)$	✓	fallend (streng)	0
(f) $(q = 1)$	✓	steigend/fallend (nicht streng)	1
(f) $(q > 1)$	–	steigend (streng)	–
(g)	nach unten	steigend (streng)	–

4.2.1.2 Reihen

Gegeben Folge (a_n). Konstruiere neue Folge (S_m) mit $S_m = \sum_{n=0}^m a_n$
→ diese nennt man dann eine **Reihe**. (möglich auch: $S_m = \sum_{n=1}^m a_n$)

Notation: Man schreibt dafür $S = \sum_{n=0}^\infty a_n$ (bzw. $S = \sum_{n=1}^\infty a_n$)
Dieser Ausdruck steht aber für die *Folge* (S_m) und sagt a priori noch nichts darüber aus, ob diese überhaupt einen Grenzwert S hat.

Falls der Grenzwert $\lim_{m\to\infty} \sum_{n=0}^m a_n$ existiert, heißt die Reihe **konvergent**.
Falls sogar $\lim_{m\to\infty} \sum_{n=0}^m |a_n|$ existiert, heißt sie **absolut konvergent**.
NB: Falls $\sum_{n=0}^\infty |a_n|$ konvergiert, dann auch $\sum_{n=0}^\infty a_n$.

Zum Beispiel **Geometrische Reihe** (siehe 1.1.2)

$S_m = \sum_{n=0}^m q^n = \frac{1-q^{m+1}}{1-q} \;\Rightarrow\; \lim_{m\to\infty} S_m \begin{cases} \text{divergiert für } q \ge 1 \\ \text{konvergiert } (\to \frac{1}{1-q}) \text{ für } q < 1 \end{cases}$

wird häufig benutzt, um andere Reihen abzuschätzen.

Weitere Beispiele: Reihen aus unseren Musterfolgen (4.2.1.1)

(a) $\sum_{n=0}^\infty n = 0+1+2+3+\cdots$ divergiert.

(b) $\sum_{n=0}^\infty (-1)^n = 1-1+1-\cdots$ hat keinen Grenzwert.
(S_m alterniert zwischen 0 und 1).

(c) $\sum_{n=1}^\infty \frac{1}{n} = 1+\frac{1}{2}+\frac{1}{3}+\cdots$ divergiert.
(denn $1+\frac{1}{2}+\underbrace{\frac{1}{3}+\frac{1}{4}}_{\ge 2\cdot\frac{1}{4}=\frac{1}{2}}+\underbrace{\frac{1}{5}+\frac{1}{6}+\frac{1}{7}+\frac{1}{8}}_{\ge 4\cdot\frac{1}{8}=\frac{1}{2}}+\cdots = 1+\underbrace{\frac{1}{2}+\frac{1}{2}+\frac{1}{2}}_{\text{unendlich oft}}$ divergiert.)

(c') aber: $\sum_{n=1}^\infty (-1)^{n+1}\frac{1}{n} = 1-\frac{1}{2}+\frac{1}{3}-\cdots$ konvergiert.
(denn $\underbrace{1-\frac{1}{2}}_{\frac{1}{1\cdot 2}}+\underbrace{\frac{1}{3}-\frac{1}{4}}_{\frac{1}{3\cdot 4}}+\underbrace{\frac{1}{5}-\frac{1}{6}}_{\frac{1}{5\cdot 6}}+\underbrace{\frac{1}{7}-\frac{1}{8}}_{\frac{1}{7\cdot 8}}+\cdots = \frac{1}{1\cdot 2}+\frac{1}{3\cdot 4}+\frac{1}{5\cdot 6}+\cdots$

weiterhin: $\frac{1}{1\cdot 2}+\underbrace{\frac{1}{3\cdot 4}+\frac{1}{5\cdot 6}}_{<2\cdot\frac{1}{3\cdot 4}<\frac{1}{4}}+\underbrace{\frac{1}{7\cdot 8}+\cdots+\frac{1}{13\cdot 14}}_{<4\cdot\frac{1}{7\cdot 8}<\frac{1}{8}}+\cdots < \sum_{n=0}^\infty (\frac{1}{2})^n = 2$

$\Rightarrow$ Zusammengefasste Reihe ist monoton (nur positive Elemente)
und beschränkt $\to$ konvergent!) (de facto: $\sum_{n=1}^\infty (-1)^{n+1}\frac{1}{n} = \ln(2)$)

(d) $\sum_{n=1}^\infty \frac{n-1}{n} = 0+\frac{1}{2}+\frac{2}{3}+\cdots$ divergiert (da Grenzwert $\lim_{n\to\infty}\frac{n-1}{n} = 1$).

(e) $\sum_{n=1}^\infty \frac{1}{n!} = 1+\frac{1}{2}+\frac{1}{2\cdot 3}+\frac{1}{2\cdot 3\cdot 4}+\cdots$ konvergiert. (de facto: $\sum_{n=0}^\infty \frac{1}{n!} = \mathrm{e}$)
(denn $1+1+\frac{1}{2}+\underbrace{\frac{1}{2\cdot 3}}_{<\frac{1}{2^2}}+\underbrace{\frac{1}{2\cdot 3\cdot 4}}_{<\frac{1}{2^3}}+\cdots < 1+\sum_{n=0}^\infty (\frac{1}{2})^n = 3$.

also: Folge $S_m = \sum_{n=1}^m \frac{1}{n!}$ ist beschränkt und monoton $\to$ hat Grenzwert.)

(f) : Geometrische Reihe, oben bereits diskutiert.

(g) : Divergiert.

Aufgaben

- Bestimmen Sie die Grenzwerte der Folgen $a_n = \frac{2n}{3n+7}$, $a_n = \frac{(3n-4)(n^2+1)}{7n(2n^2+10000)}$,
  $a_n = \sqrt{n+1}-\sqrt{n}$, $a_n = \frac{t^n-t^{-n}}{t^n+t^{-n}}$ für $|t|<1$, $|t|=1$, $|t|>1$.
- Zeigen Sie: $\sum_{n=1}^\infty \frac{1}{n^2}$ ist konvergent. (de facto: $\sum_{n=1}^\infty = \frac{\pi^2}{6}$)
- Berechnen Sie $\sqrt{2+\sqrt{2+\sqrt{2+\cdots}}}$
- Diskutieren Sie folgenden "Beweis" der Behauptung $2=4$:
  Betrachte die Gleichung $x^{x^{x^{\cdot^{\cdot^{\cdot}}}}} = 2$ (Notation: $x^{x^x} = x^{(x^x)}$, $x^{x^{x^x}} = x^{(x^{x^x})}$ etc.)
  Lösungsversuch: Da $x^{x^{x^{\cdot^{\cdot^{\cdot}}}}} = 2$, folgt $x^{(x^{x^{\cdot^{\cdot^{\cdot}}}})} = x^2 = 2 \;\Rightarrow x = \sqrt{2}$
  Nun die Gleichung $x^{x^{x^{\cdot^{\cdot^{\cdot}}}}} = 4$
  Lösungsversuch wie oben: Da $x^{x^{x^{\cdot^{\cdot^{\cdot}}}}} = 4$, folgt $x^{(x^{x^{\cdot^{\cdot^{\cdot}}}})} = x^4 = 4 \;\Rightarrow x = \sqrt{2}$
  Also folgt $\sqrt{2}^{\sqrt{2}^{\sqrt{2}^{\cdot^{\cdot^{\cdot}}}}} = 2$ und $\sqrt{2}^{\sqrt{2}^{\sqrt{2}^{\cdot^{\cdot^{\cdot}}}}} = 4 \;\Rightarrow 2=4$ (?)
- Hier noch ein weiterer "Beweis", diesmal für $0 > \frac{1}{2}$.
  Betrachte $S = \sum_{n=1}^\infty (-1)^{n+1}\frac{1}{n} = 1-\frac{1}{2}+\frac{1}{3}-\frac{1}{4}+\frac{1}{5}\cdots$
  Wir haben gezeigt: $S = \frac{1}{1\cdot 2}+\frac{1}{3\cdot 4}+\frac{1}{5\cdot 6}+\cdots$. Daraus folgt $S > \frac{1}{2}$.
  Betrachte nun folgende Umformung:
  $1-\frac{1}{2}+\frac{1}{3}-\frac{1}{4}+\frac{1}{5}\cdots = (1-\frac{1}{2}-\frac{1}{4}-\frac{1}{8}\cdots)+(\frac{1}{3}-\frac{1}{6}-\frac{1}{12}-\frac{1}{24}\cdots)+(\frac{1}{5}-\frac{1}{10}-\frac{1}{20}-\frac{1}{40}\cdots)+\cdots$
  $= (1+\frac{1}{3}+\frac{1}{5}\cdots)(1-\frac{1}{2}-\frac{1}{4}-\frac{1}{8}\cdots)$
  Mit $(1-\frac{1}{2}-\frac{1}{4}-\frac{1}{8}\cdots) = 1-\frac{1}{2}\sum_{n=0}^\infty(\frac{1}{2})^n = 1-\frac{1}{2}\cdot 2 = 0$ folgt: $S = 0 > \frac{1}{2}$ (?)

(Hintergrund: Vorsicht beim Vertauschen von Termen in unendlichen Reihen! Nur erlaubt für *absolut* konvergente Reihen).

4.2.2 Kurzer Abriss über Potenzreihen

Funktionen der Form $P(x) = \sum_{k=0}^{\infty} a_k x^k$ (x komplex, a_k reell oder komplex).

1) Konvergenz:

Erste Frage: Für welche x konvergiert eine solche Reihe überhaupt?

Antwort: Es gibt einen **Konvergenzradius** R, so dass gilt:

- Für $|x| < R$: Die Reihe ist **absolut konvergent** d.h. schon $\sum_{k=0}^{\infty} |a_k x^k|$ konvergiert.
- Für $|x| > R$: Die Reihe konvergiert nicht.
- Für $|x| = R$: Die Reihe konvergiert oder divergiert.

(Beweis in zwei Schritten:

(i) Wenn $P(x_0)$ konvergiert, dann ist $P(x)$ absolut konvergent für $|x| < |x_0|$
denn: $P(x_0)$ konvergent $\Rightarrow$ $a_k x_0^k$ beschränkt $\Rightarrow$ $\exists C$ mit $|a_k x_0^k| < C \; \forall k$
$\Rightarrow$ Für $q := |\frac{x}{x_0}| < 1$ gilt: $|a_k x^k| = |a_k x_0^k||\frac{x}{x_0}|^k \le C q^k$
$\Rightarrow$ $\sum_{k=0}^{\infty} |a_k x^k| \le C \sum_{k=0}^{\infty} q^k = \frac{C}{1-q}$ konvergiert.

(ii) Wenn $P(x_1)$ divergiert, dann divergiert $P(x)$ für $|x| > |x_1|$
(denn: wenn $P(x)$ konvergent wäre, dann nach (i) auch $P(x_1)$.)

(i) und (ii) zusammengenommen $\Rightarrow$ Es existiert ein Konvergenzradius. ✓)

Bemerkung: Sofern die Reihe konvergiert, ist sie bei $|x| = R$ auch stetig (**Abelscher Grenzwertsatz**).

Berechnung des Konvergenzradius oft ganz einfach:

- **Quotientenkriterium**: $\boxed{\frac{1}{R} = \lim_{k\to\infty} |\frac{a_{k+1}}{a_k}|}$
  (falls Grenzwert existiert)

(Beweis: Nimm an $\lim_{k\to\infty} |\frac{a_{k+1}}{a_k}|$ existiert.
$\xi = \lim_{k\to\infty} |\frac{a_{k+1}}{a_k}| \Leftrightarrow \forall \delta > 0 : \exists k_0 > 0 : \xi'' := \xi - \delta < |\frac{a_{k+1}}{a_k}| < \xi' := \xi + \delta \; \forall \; k > k_0$
Sei nun $\xi|x| < 1$. Wähle δ so, dass $\xi'|x| < 1 \;\Rightarrow\; \exists q$ mit $\xi'|x| < q < 1$
$\Rightarrow$ $|a_{k+1} x^{k+1}| < |a_k x^k| q < |a_{k_0}| q^{k-k_0} \; \forall \; k > k_0$. Schätze $\sum_{k=k_0}^{\infty} |a_k||x|^k$ ab.
$\Rightarrow \sum_{k=k_0}^{\infty} |a_k||x|^k < |a_{k_0}| \sum_{k=k_0}^{\infty} q^{k-k_0} = |a_{k_0}| \frac{1}{1-q}$ konvergiert.
Sei $\xi|x| > 1$. Wähle δ so, dass $\xi''|x| > 1 \;\Rightarrow\; \exists Q$ mit $\xi''|x| > Q > 1$
$\Rightarrow$ $|a_{k+1} x^{k+1}| > |a_k x^k| Q > |a_{k_0}| Q^{k-k_0} \; \forall \; k > k_0$.
$\Rightarrow \sum_{k=0}^{\infty} |a_k||x|^k > \sum_{k=k_0}^{\infty} |a_k||x|^k > |a_{k_0}| \sum_{k=k_0}^{\infty} Q^{k-k_0}$ divergiert.)

- **Wurzelkriterium**: $\boxed{\frac{1}{R} = \lim_{k\to\infty} \sqrt[k]{|a_k|}}$
  (falls Grenzwert existiert)

(Beweis: Nimm an $\lim_{k\to\infty} \sqrt[k]{|a_k|}$ existiert.
$\xi = \lim_{k\to\infty} \sqrt[k]{|a_k|} \Leftrightarrow \forall \delta > 0 : \exists k_0 > 0 : \xi'' := \xi - \delta < \sqrt[k]{|a_k|} < \xi' := \xi + \delta \; \forall \; k > k_0$
Sei nun $\xi|x| < 1$. Wähle δ so, dass $\xi'|x| < 1 \;\Rightarrow\; \exists q$ mit $\xi'|x| < q < 1$
$\Rightarrow$ $\sqrt[k]{|a_k x^k|} = \sqrt[k]{|a_k|}|x| < \xi'|x| < q < 1 \; \forall k > k_0$. Schätze $\sum_{k=k_0}^{\infty} |a_k||x|^k$ ab.
$\Rightarrow$ $|a_k||x|^k < q^k \;\Rightarrow\; \sum_{k=k_0}^{\infty} |a_k||x|^k < \sum_{k=0}^{\infty} q^k = \frac{1}{1-q}$ konvergiert.
Sei $\xi|x| > 1$. Wähle δ so, dass $\xi''|x| > 1 \;\Rightarrow\; \exists Q$ mit $\xi''|x| > Q > 1$
$\Rightarrow$ $\sqrt[k]{|a_k x^k|} = \sqrt[k]{|a_k|}|x| > \xi''|x| > Q > 1 \; \forall \; k > k_0$
$\Rightarrow$ $|a_k||x|^k > Q^k \;\Rightarrow\; \sum_{k=k_0}^{\infty} |a_k||x|^k > \sum_{k=k_0}^{\infty} Q^k$ divergiert.)

Beispiele:

- Geometrische Reihe $P(x) = \sum_{k=0}^{\infty} x^k$, d.h. $a_k = 1 \; \forall k$
  $\Rightarrow$ $R = 1$ nach beiden Kriterien.
- Exponentialfunktion: $P(x) = \sum_{k=0}^{\infty} \frac{1}{k!} x^k$, d.h. $a_k = \frac{1}{k!}$
  Quotientenkriterium: $\frac{1}{R} = \lim_{k\to\infty} |\frac{a_{k+1}}{a_k}| = \lim_{k\to\infty} |\frac{1}{k+1}| = 0$
  $\Rightarrow$ $R = \infty$
- Harmonische Reihe $P(x) = \sum_{k=0}^{\infty} \frac{1}{k} x^k$, d.h. $a_k = \frac{1}{k}$
  Quotientenkriterium: $\frac{1}{R} = \lim_{k\to\infty} |\frac{a_{k+1}}{a_k}| = \lim_{k\to\infty} |\frac{k}{k+1}| = 1$
  $\Rightarrow$ $R = 1$

Fehlerabschätzung: Wann kann man eine Reihe abbrechen?

Sei $f(x) = \sum_{k=0}^{\infty} a_k x^k$. Betrachte Restglied $R_n(x) = P(x) - \sum_{k=0}^{n} a_k x^k$.

- **Lagrangesches Abschätzung**: $|R_n(x)| \le \frac{x^{n+1}}{(n+1)!} \max_{|t|<|x|} f(t)$.
- Praktische Abschätzung mittels geometrischer Reihe
  Falls $|a_{k+1}| < |a_k|$ für $k > n$ und $|x| < 1$, folgt $|R_n(x)| \le a_{n+1} \frac{x^{n+1}}{1-x}$.

Aufgaben

Berechnen Sie den Konvergenzradius der Reihen
$\sum_{k=1}^{\infty} \frac{1}{k^n} x^k$, $\sum_{k=1}^{\infty} k^n x^k$, $\sum_{k=1}^{\infty} \frac{m^k}{(k+m)^n} x^k$ für festes $n, m \in \mathbb{N}$
$\sum_{n=1}^{\infty} n! x^n$, $\sum_{n=1}^{\infty} \sqrt{1+9^n} x^n$, $\sum_{n=1}^{\infty} n^n x^n$

2) Rechnen mit Potenzreihen

- Addieren, Subtrahieren: Gliedweise
  $\sum a_k x^k \pm \sum b_k x^k = \sum (a_k \pm b_k) x^k$
- Multiplizieren: Nur bei *absolut konvergenten* Reihen erlaubt. Dann wird ausmultipliziert wie bei Polynomen. Ebenso kann man dividieren, sofern nicht durch Null geteilt wird.
  (Im Allgemeinen wird sich dabei der Konvergenzradius ändern).
- Substitution: *Absolut konvergente* Reihen können ineinander substituiert werden.
  (Im Allgemeinen wird sich dabei der Konvergenzradius ändern).
- Differentiation: Gliedweise.
  $f(x) = \sum a_k x^k \quad \rightarrow \quad f'(x) = \sum_k a_k k x^{k-1}$ etc.
- Eindeutigkeit Wenn $\sum a_k x^k \equiv \sum b_k x^k$ auf einem Intervall um $x = 0$, dann ist $a_k = b_k \; \forall k$.
  $\Rightarrow$ erlaubt Koeffizientenvergleich.

  Äquivalent: $\sum c_k x^k \equiv 0$ auf einem Intervall $\Rightarrow c_k = 0 \; \forall k$.
  (Folgt letztlich daraus, dass alle Ableitungen Null sind.)

  NB: Interpretation innerhalb der **linearen Algebra**:
  Potenzreihen bilden unendlichdimensionalen Vektorraum,
  Funktionen x^k bilden linear unabhängige Basis.

4.2.3 Konstruktion der Taylor-Reihe

Methode, Potenzreihen für Funktionen zu konstruieren. Funktioniert für sehr viele Funktionen $f(x)$.

Beantwortet Eingangsfrage (iii) ("Wie kommt man an die Koeffizienten?")

1) Taylor-Entwicklung um $x = 0$ **(MacLaurin-Reihe)**

Funktion $f(x)$ sei unendlich oft differenzierbar.
Setze an: $f(x) = \sum_{k=0}^{\infty} a_k x^k$

$$\begin{array}{llllllllll} \Rightarrow & f(x) & = & a_0 & + \; a_1 x & + \; a_2 x^2 & + \; a_3 x^3 & + \cdots & \Rightarrow & f(0) = a_0 \\ & f'(x) & = & & a_1 & + \; 2a_2 x & + \; 3a_3 x^2 & + \cdots & \Rightarrow & f'(0) = a_1 \\ & f''(x) & = & & & 2a_2 & + \; 6a_3 x & + \cdots & \Rightarrow & f''(0) = 2a_2 \\ & \vdots & & & & & & & & \end{array}$$

Allgemein: $f^{(n)}(x) = \sum_{k=n}^{\infty} \frac{k!}{n!} x^{k-n} = n! a_n + \cdots \;\Rightarrow\; f^{(n)}(0) = n! a_n$

$\Rightarrow$ Taylor-Reihe $\boxed{f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} x^k}$.

NB: Falls $f(x)$ unendlich oft differenzierbar und nur im Reellen bekannt, kann man so $f(x)$ für komplexe Zahlen verallgemeinern ("analytische Fortsetzung").

2) Anwendungsbeispiele

- Geometrische Reihe

$$\left.\begin{array}{lllllll} f(x) & = & (1-x)^{-1} & \Rightarrow & f(0) & = & 1 \\ f'(x) & = & (1-x)^{-2} & & f'(0) & = & 1 \\ f''(x) & = & 2(1-x)^{-3} & & f''(0) & = & 2 \\ f'''(x) & = & 2\cdot 3(1-x)^{-4} & & f'''(0) & = & 2\cdot 3 \\ & \vdots & & & & & \\ f^{(n)}(x) & = & n!(1-x)^{-n} & \Rightarrow & f^{(n)}(0) & = & n! \end{array}\right\} a_n = \frac{f^{(n)}(0)}{n!} = 1 \quad \checkmark$$

$\Rightarrow$ Rekonstruiere $\dfrac{1}{1-x} = \sum\limits_{n=0}^{\infty} x^n$!

- Exponentialfunktion

$$\left.\begin{array}{lllllll} f(x) & = & e^x & \Rightarrow & f(0) & = & 1 \\ f'(x) & = & e^x & \Rightarrow & f'(0) & = & 1 \\ & \vdots & & & & & \\ f^{(n)}(x) & = & e^x & \Rightarrow & f^{(n)}(0) & = & 1 \end{array}\right\} a_n = \frac{f^{(n)}(0)}{n!} = \frac{1}{n!}$$

$\Rightarrow \exp(x) = \sum\limits_{n=0}^{\infty} \dfrac{1}{n!} x^n$ wie behauptet!

3) Gültigkeit der Taylor-Entwicklung

- Wann kann man Taylor-Reihen bilden?

  Notwendige Bedingung: $f(x)$ bei $x = 0$ unendlich oft differenzierbar.

  Achtung: Diese Bedingung ist nicht hinreichend.

  Berühmtes Gegenbeispiel $f(x) = \exp(-1/x^2)$. Bei $x = 0$ unendlich oft differenzierbar, aber $f^{(n)}(0) = 0 \quad \forall n$.

$\Rightarrow$ Taylorreihe wäre identisch Null.
(Hintergrund: $f(x)$ kann bei $x \to 0$ nicht in die komplexe Ebene erweitert werden: $\lim_{r\to 0} \exp(-1/(ir)^2) \to \infty$ divergiert.)

- In welchem Gebiet ist Taylor-Reihe gültig?
  Antwort: Innerhalb des Konvergenzradius.
  Achtung: Nicht notwendig deckungsgleich mit dem Definitionsbereich der ursprünglichen Funktion $f(x)$
  Beispiel: $f(x) = \frac{1}{1-x}$ ist definiert im zusammenhängenden Gebiet $x \in]-\infty, 1[$, aber Taylorreihe $\sum x^k$ nur für $|x| < 1$.

4) Erweiterung: Taylor-Reihe um beliebigen Punkt x_0

Nachteile der MacLaurin-Reihe bei großen $|x|$:

- Für große x muss man immer mehr Terme mitnehmen, um eine gute numerische Genauigkeit zu erzielen
- Eventuell konvergiert die Reihe gar nicht mehr.

Ausweg: Bilde Taylor-Reihe um anderen Punkt x_0.
Schreibe $f(x) = f(x_0 + h) =: \hat{f}(h)$, entwickle $\hat{f}(h)$ nach $h = x - x_0$
NB: Ableitungen $\hat{f}^{(n)}(h) = f^{(n)}(x_0)$

$\Rightarrow$ Taylor-Reihe $\boxed{f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k}$

bzw. $\boxed{f(x + x_0) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!} x^k}$

4.2.4 Anwendungen

1) Exponentialfunktion: Wie gehabt

$$e^x = \sum_{n=0}^{\infty} \frac{1}{n!} x^n = 1 + x + \frac{x^2}{2} + \frac{x^3}{3!} + \cdots$$

2) Hyperbolische Funktionen: Summen/Differenzen von $e(x)$

$$\sinh(x) = \frac{1}{2}(e^x - e^{-x}) = \sum_{n=0}^{\infty} \frac{x^{2n+1}}{(2n+1)!} = x + \frac{x^3}{3!} + \cdots$$

$$\cosh(x) = \frac{1}{2}(e^x + e^{-x}) = \sum_{n=0}^{\infty} \frac{x^{2n}}{(2n)!} = 1 + \frac{x^2}{2!} + \cdots$$

3) Trigonometrische Funktionen: über Taylor-Entwicklung

$$\sin(x) = \sum_{n=0}^{\infty} (-1)^n \frac{1}{(2n+1)!} x^{2n+1}$$

$$\cos(x) = \sum_{n=0}^{\infty} (-1)^n \frac{1}{(2n)!} x^{2n}$$

(Rechnung für $\sin(x)$:

$$\left.\begin{array}{lcl}
f(x) = \sin(x) & \Rightarrow & f(0) = 0 \\
f'(x) = \cos(x) & & f'(0) = 1 \\
f''(x) = -\sin(x) & & f''(0) = 0 \\
f'''(x) = -\cos(x) & & f'''(0) = -1 \\
f^{(4)}(x) = \sin(x) & & f^{(4)}(0) = 0 \\
 & \vdots & \\
f^{(4n)}(x) = \sin(x) & \Rightarrow & f^{(4n)}(0) = 0 \\
f^{(4n+1)}(x) = \cos(x) & & f^{(4n+1)}(0) = 1 \\
f^{(4n+2)}(x) = -\sin(x) & & f^{(4n)}(0) = 0 \\
f^{(4n+3)}(x) = -\cos(x) & & f^{(4n)}(0) = -1
\end{array}\right\} \quad \begin{array}{l} a_{2n} = \dfrac{f^{(2n)}(0)}{n!} = 0 \\ a_{2n+1} = (-1)^{n+1} \dfrac{1}{n!} \end{array}$$

Rechnung für $\cos(x)$:

$$\left.\begin{array}{lll} f(x)=\cos(x) & \Rightarrow & f(0)=1 \\ f'(x)=-\sin(x) & & f'(0)=0 \\ f''(x)=-\cos(x) & & f''(0)=-1 \\ f'''(x)=\sin(x) & & f'''(0)=0 \\ f^{(4)}(x)=\cos(x) & & f^{(4)}(0)=1 \\ & \vdots & \\ f^{(4n)}(x)=\cos(x) & \Rightarrow & f^{(4n)}(0)=1 \\ f^{(4n+1)}(x)=-\sin(x) & & f^{(4n+1)}(0)=0 \\ f^{(4n+2)}(x)=-\cos(x) & & f^{(4n)}(0)=-1 \\ f^{(4n+3)}(x)=\sin(x) & & f^{(4n)}(0)=0 \end{array}\right\} \begin{array}{ll} a_{2n} & =(-1)^n \frac{1}{n!} \\ a_{2n+1} & =0 \end{array}$$

4) Logarithmus:

Funktion $\ln(x)$ lässt sich um $x=0$ nicht Taylor-entwickeln (divergiert).
Aber: Entwicklung um $x_0=1$ möglich.

$$\ln(1+x)=\sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{n}x^n=x-\frac{x^2}{2}+\frac{x^3}{3}-\frac{x^4}{4}\cdots$$

(Rechnung
$$\left.\begin{array}{lll} f(x)=\ln(1+x) & \Rightarrow & f(0)=0 \\ f'(x)=(1+x)^{-1} & & f'(0)=1 \\ f''(x)=-(1+x)^{-2} & & f''(0)=-1 \\ f'''(x)=2(1+x)^{-3} & & f'''(0)=2 \\ & \vdots & \\ f^{(n)}(x)=(-1)^{n+1}(n-1)!(1+x)^{-n} & \Rightarrow & f^{(n)}(0)=(-1)^n\,(n-1)! \end{array}\right\}$$
$\Rightarrow a_n=\frac{f^{(n)}(0)}{n!}=(-1)^{n+1}\frac{(n-1)!}{n!}=(-1)^{n+1}\frac{1}{n}$)

Bemerkungen:

- Konvergenzradius ist $R=\lim_{n\to\infty}|\frac{a_n}{a_{n+1}}|=\lim_{n\to\infty}|\frac{n+1}{n}|=1$
  Macht Sinn, da $\ln(1+x)$ bei $x=1$ divergiert.
- Mit $x\to 1$ kann man den Wert von $1-\frac{1}{2}+\frac{1}{3}-\frac{1}{4}\cdots$ bestimmen:
  $\sum_{k=1}^{\infty}\frac{(-1)^{k+1}}{k}=1-\frac{1}{2}+\frac{1}{3}-\frac{1}{4}\cdots=\ln(1+1)=\ln(2)$

5) Potenzen: x^α (wobei $\alpha\in\mathbb{R}$ beliebig)

Analog 4): Für allgemeine α ist Entwicklung um $x=0$ i.A. nicht möglich, da x^α nicht beliebig oft differenzierbar
$\rightarrow$ Entwicklung um $x=1$

$$(1+x)^\alpha=\sum_{n=1}^{\infty}\binom{\alpha}{n}x^n=x-\frac{x^2}{2}+\frac{x^3}{3}-\frac{x^4}{4}\cdots \text{ mit } \binom{\alpha}{n}=\frac{\alpha(\alpha-1)\cdot(\alpha-n+1)}{n!}$$

(Rechnung :
$$\left.\begin{array}{lll} f(x)=(1+x)^\alpha & \Rightarrow & f(0)=0 \\ f'(x)=\alpha(1+x)^{\alpha-1} & & f'(0)=1 \\ f''(x)=\alpha(\alpha-1)(1+x)^{\alpha-2} & & f''(0)=-1 \\ & \vdots & \\ f^{(n)}(x)=\alpha(\alpha-1)\cdots(\alpha-n+1)(1+x)^{\alpha-n} & \Rightarrow & f^{(n)}(0)=\alpha(\alpha-1)\cdots(\alpha-n+1) \end{array}\right\}$$
$\Rightarrow a_n=\frac{\alpha(\alpha-1)\cdots(\alpha-n+1}{n!}=\binom{\alpha}{n}$)

Speziell $\alpha=n\in\mathbb{N}_0$ natürliche Zahl $\rightarrow\binom{N}{n}=0$ für $n>N$)
$\Rightarrow$ Potenzreihe bricht ab.
Man erhält bekannte Formel $(1+x)^N=\sum_{k=0}^{N}\binom{N}{k}x^k$.

6) Tabelle: Taylor-Reihe bis zur Ordnung x^2 für wichtige Funktionen:

- $(1+x)^\alpha\approx 1+\alpha x+\frac{\alpha(\alpha-1)}{2}x^2$
- $\sin(x)\approx x$
- $\cos(x)\approx 1-\frac{x^2}{2}$

- $\exp(x) \approx 1 + x + \frac{x^2}{2}$
- $\ln(1 + x) \approx x - \frac{x^2}{2}$

Aufgaben

Berechnen Sie die Taylorreihe um $x = 0$ von
- $f(x) = \exp(-x^2)$ (Gaußkurve)
- $f(x) = 1/(1 - x^2)$

Berechnen Sie die ersten vier Glieder der Taylor-Reihe von
- $\tan(x)$
- $e^x \sin(x)$ (multiplizieren der Reihen für e^x und $\sin(x)$
- $e^{\sin(x)}$

Berechnen Sie die Taylor-Reihe von
- $\exp(x)$ um den Punkt $x_0 = 10$
- $\sin(x)$ um den Punkt $x_0 = \pi$

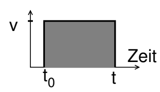
Lösen Sie näherungsweise die Gleichung $\frac{1}{\sqrt{3}} e^{-2x} = x$

4.3 Integralrechnung

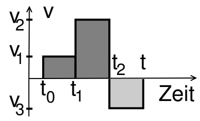
Erinnerung: Einleitendes Beispiel bei der Differentialrechnung (Kapitel 4.1.1):
Wagen, der entlang gerader Straße fährt.
Gegeben Weg $s(t)$ $\rightarrow$ Geschwindigkeit $v(t)$?

Betrachte nun umgekehrtes Problem: $v(t)$ gegeben (z.B. vom Tacho)
$\rightarrow$ Wie kann man daraus Strecke $s(t)$ berechnen (z.B. Kilometerzähler)

<u>Graphisch</u>

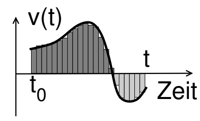
- Wenn v konstant ist: 

$\rightarrow$ $s(t) = v(t - t_0)$
$\hat{=}$ <u>Fläche</u> ■

- Wenn v sich ab und zu abrupt ändert: 

$\rightarrow$ $s(t) = v_1(t_1 - t_0) + v_2(t_2 - t_1) + v_3(t - t_2)$
$\hat{=}$ Fläche ■ – Fläche □

- Beliebiger Verlauf $v(t)$:
  $\rightarrow$ Annäherung durch viele infinitesimale Stufen
  $\hat{=}$ Fläche unter der Kurve des Teils $v(t) > 0$
  – Fläche über der Kurve des Teils $v(t) < 0$
  (Fläche mit Vorzeichen)

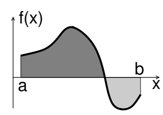


Führt zum sogenannten **Riemannschen Integral**

Allgemeine Aufgabenstellung: Berechnung der Fläche unter einer Kurve
Unser Beispiel zeigt, dass das so etwas wie eine "inverser Ableitung" ist.
$\rightarrow$ Hauptsatz der Differential- und Integralrechnung

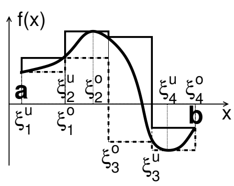
4.3.1 Das Riemannsche Integral

Gegeben reelle Funktion $f(x)$ im Intervall $[a, b]$.
Aufgabe: Berechnung der Fläche F unter der Kurve $f(x)$
(ggf. mit negativen Beiträgen für $f(x) < 0$).



1) <u>Konstruktion des Integrals</u>

- Zerlege Intervall $[a, b]$ in Teilintervalle der Breite $\Delta x = (b - a)/n$.
  $\rightarrow$ Teilintervalle ν: $[x_{\nu-1}, x_\nu]$ mit $x_\nu - x_{\nu-1} = \Delta x$.
- Lege in jedem Teilintervall x-Werte ξ_ν fest, so dass:
  $x = \xi_\nu^{(o)}$: Wert, an dem $f(x)$ maximal wird.
  $x = \xi_\nu^{(u)}$: Wert, an dem $f(x)$ minimal wird.



- Berechne $S_n = \sum_{\nu=1}^{n} f(\xi_n)\,\Delta x$ jeweils für $\xi_\nu^{(o)}$ und $\xi_\nu^{(u)}$.
  $\rightarrow$ Obersumme $S_n^{(o)} = \sum_{\nu=1}^{n} f(\xi_\nu^{(o)})\,\Delta x$
  Untersumme $S_n^{(u)} = \sum_{\nu=1}^{n} f(\xi_\nu^{(u)})\,\Delta x$
  $\Rightarrow$ $S_n^{(u)} < F$ (gesuchte Fläche) $< S_n^{(o)}$ $\quad \forall\, n$

- Bilde Grenzwert $n \to \infty$.
  $f(x)$ heißt Riemann-integrierbar, falls die Grenzwerte $\lim_{n\to\infty} S_n^{(o)}$ und $\lim_{n\to\infty} S_n^{(u)}$ existieren und $\lim_{n\to\infty} S_n^{(o)} = \lim_{n\to\infty} S_n^{(u)}$.
  Dann ist die Fläche $F = \lim_{n\to\infty} S_n^{(o,u)} = \lim_{n\to\infty} \sum_{\nu=1}^{n} f(\xi_\nu^{(o,u)})\, \Delta x$.

Notation: $\boxed{F = \int_a^b f(x)\,\mathrm{d}x = \int_a^b \mathrm{d}x\, f(x)}$.

2) Riemann-Integrierbarkeit

Es gilt: Falls $f(x)$ auf $[a,b]$ stückweise stetig ist (d.h. stetig auf endlich vielen Teilintervallen) und beschränkt, dann ist $f(x)$ Riemann-integrierbar.
(anschaulich klar. Beweis weggelassen.)

3) Eigenschaften des Integrals

- Linearität: $\int_a^b \{Af(x) + Bg(x)\}\,\mathrm{d}x = A\int_a^b f(x)\,\mathrm{d}x + B\int_a^b g(x)\,\mathrm{d}x$.
- Intervall-Addition: $\int_a^b f(x)\,\mathrm{d}x + \int_b^c f(x)\,\mathrm{d}x = \int_a^c f(x)\,\mathrm{d}x$.
- Ungleichungen:
  (i) Wenn $f(x) \le g(x)\ \forall\, x \in [a,b]$
  Dann folgt $\int_a^b f(x)\,\mathrm{d}x \le \int_a^b g(x)\,\mathrm{d}x$.
  (ii) Dreiecksungleichung: $\underbrace{\left|\int_a^b f(x)\,\mathrm{d}x\right|}_{\text{entält negative Flächen}} \le \underbrace{\int_a^b |f(x)|\,\mathrm{d}x}_{\text{alle Flächen positiv}}$.
  (iii) Wenn $m \le f(x) \le M\ \ \forall\, x \in [a,b]$
  Dann folgt: $m(b-a) \le \int_a^b f(x)\,\mathrm{d}x \le M(b-a)$
- Umkehrung $\int_a^b f(x)\,\mathrm{d}x = -\int_b^a f(x)\,\mathrm{d}x$.
  (denn: dann sind in dem Ausdruck $\lim_{nto\infty} \sum_{\nu=1}^{n} f(\xi_\nu^{(o,u)})\,\Delta x$ die Größen $\Delta x = x_\nu - x_{\nu-1}$ negativ.)

4) Mittelwertsatz der Integralrechnung

Ist $f(x)$ stetig in $[a,b]$, so existiert ein $\bar\xi \in [a,b]$ mit $\int_a^b f(x)\,\mathrm{d}x = (b-a)f(\bar\xi)$.

4.3.2 Hauptsatz und Stammfunktion

Aus dem Beispiel Geschwindigkeit war schon ersichtlich: Integrieren ist irgendwie invers zum Differenzieren. Das soll nun spezifiziert werden (in zwei Teilen).

1) Gegeben eine stetige Funktion $f(x)$. Definiere $\boxed{F_{x_0}(y) = \int_{x_0}^y f(x)\,\mathrm{d}x}$.

Dann gilt: $\boxed{\frac{\mathrm{d}}{\mathrm{d}y} F_{x_0}(y) = f(y)}$.

(Beweisskizze: $\frac{\mathrm{d}}{\mathrm{d}y}\int_{x_0}^y f(x)\mathrm{d}x = \lim_{\Delta y\to 0}\frac{1}{\Delta y}\Big[\int_{x_0}^{y+\Delta y} f(x)\mathrm{d}x - \int_{x_0}^y f(x)\mathrm{d}x\Big]$
$= \lim_{\Delta y\to 0}\frac{1}{\Delta y}\int_y^{y+\Delta y} f(x)\mathrm{d}x \underbrace{=}_{\substack{\text{Mittelwertsatz der Integralrechnung}\\ \text{Wähle } \bar\xi \in [y, y+\Delta y] \text{ so, dass } \int_y^{y+\Delta y} f(x)\mathrm{d}x = \Delta y f(\bar\xi)}} \lim_{\Delta y\to 0}\frac{1}{\Delta y}\Delta y\, f(\bar\xi) \underbrace{=}_{f \text{ stetig}} f(y))$

2) Gegeben eine differenzierbare Funktion $f(x)$

Dann gilt: $\boxed{\int_a^b f'(x)\,\mathrm{d}x = f(b) - f(a) =: f(x)\Big|_a^b}$.

(Beweisskizze: $\int_a^b f'(x)\mathrm{d}x = \lim_{n\to\infty} \sum_{\nu=1}^n \Delta x\, f'(\xi_\nu) \underbrace{=}_{\substack{\text{Mittelwertsatz der Differentialrechnung:} \\ \text{Wähle } \xi_\nu \in [x_{\nu-1}, x_\nu] \text{ so, dass } f'(\xi_\nu)\Delta x = f(x_\nu) - f(x_{\nu-1})}} \lim_{n\to\infty} \sum_{\nu=1}^n \Delta x \frac{f(x_\nu)-f(x_{\nu-1})}{\Delta x}$

$= \lim_{n\to\infty} f(x_n) - f(x_0) \underbrace{=}_{x_n = b,\ x_0 = a} f(b) - f(a)$.)

Zusammen: **Hauptsatz der Differential- und Integralrechnung**

Fazit: Um das Integral $\int f(x)\,\mathrm{d}x$ zu berechnen, muss man die **Stammfunktion** von $f(x)$ kennen, d.h. die Funktion $F(x)$ mit $\boxed{\frac{\mathrm{d}}{\mathrm{d}x}F(x) = f(x)}$.

Dann ist $\boxed{\int_a^b f(x)\,\mathrm{d}x = F(x)\Big|_a^b}$.

Bemerkungen:

- Stammfunktion ist natürlich nicht eindeutig. Mit $F(x)$ ist auch $F(x)+c$ Stammfunktion zu $f(x)$ für jede beliebige Konstante c.
  Aber: an dem Wert von $F(x)|_a^b$ ändert das nichts.
- Wegen des Hauptsatzes kann man Integrale auch dazu benutzen, Stammfunktionen zu ermitteln (über $F(y) = \int^y f(x)\,\mathrm{d}x$).
  Deshalb unterscheidet man zwischen

  unbestimmten Integralen: $\int f(x)\,\mathrm{d}x$.
  → keine expliziten Integrationsgrenzen angegeben.
  Es wird nur allgemein nach Stammfunktion gesucht.

  bestimmten Integralen $\int_a^b f(x)\,\mathrm{d}x$.
  → explizite Integrationsgrenzen.
  Gesucht wird konkreter Wert eines Integrals.

Beispiele
- $\int \sin(x)\,\mathrm{d}x = -\cos(x)$ (+ Konstante)
- $\int_a^b \sin(x)\,\mathrm{d}x = [-\cos(x)]_a^b = -\cos(b) + \cos(a)$

Aufgaben

Geben Sie die Stammfunktionen der folgenden Funktionen an:
$f(x) = x^3$, $f(x) = -1/\sqrt{1-x^2}$,
$f(x) = \sinh(x)$ $f(x) = 2^x$, $f(x) = (x+2)\sin(x^2+4x-6)$

Bestimmen Sie die folgenden Integrale
$\int x^n\,\mathrm{d}x$, $\int x\,\mathrm{e}^{-x^2}\,\mathrm{d}x$
$\int_0^1 \mathrm{d}x/(1+x^2) (\rightsquigarrow \pi/4)$, $\int_0^a \mathrm{d}x/x^{1-a}$ mit $a > 0$ ($\rightsquigarrow a^{a-1}$)

4.3.3 Integrationsmethoden

Vorab: Es gibt leider kein Patentrezept, ein Integral zu knacken.
Anders als beim Differenzieren: Man kann für (fast) jede differenzierbare Funktion einen expliziten Ausdruck für die Ableitung herleiten. Für Integrale ist das oft nicht möglich.

Hier: Ein paar Tricks, mit denen man Integrale manchmal doch knacken kann.

1) Differentiationstabelle rückwärts lesen
bzw. Stammfunktion erraten.

2) Lineare Zerlegung

Ausnutzen von $\int \{Af(x) + Bg(x)\}\,\mathrm{d}x = A\int f(x)\,\mathrm{d}x + B\int g(x)\,\mathrm{d}x$.

Beispiele

(i) $\int_0^1 (1-x^2)^2 \mathrm{d}x = \int_0^1 (1-2x^2+x^4)\mathrm{d}x$
$= \int_0^1 \mathrm{d}x - 2\int_0^1 x^2 \mathrm{d}x + \int_0^1 x^4 \mathrm{d}x = 1 - 2\cdot\frac{1}{3} + \frac{1}{5} = \frac{8}{15}$

(ii) $\int_0^{\pi/2} \cos^2(\phi)\mathrm{d}\phi \underbrace{=}_{\text{Symmetrie}} \int_0^{\pi/2} \sin^2(\phi)\mathrm{d}\phi \underbrace{=}_{\text{Trick}} \frac{1}{2}\int_0^{\pi/2} \underbrace{[\sin^2(\phi)+\cos^2(\phi)]}_{1}\mathrm{d}\phi$
$= \frac{1}{2}\int_0^{\pi/2} \mathrm{d}\phi = \frac{\pi}{4}$

(iii) Partialbruchzerlegung

Gesucht sei z.B. $\int_0^1 \frac{1}{(x+1)(x+2)}\mathrm{d}x$

$\rightsquigarrow$ Zerlege $\frac{1}{(x+1)(x+2)} = \frac{A}{x+1} + \frac{B}{x+2} \quad \left(= \frac{A(x+2)+B(x+1)}{(x+1)(x+3)}\right)$

$\Rightarrow\ A(x+2) + B(x+1) = (A+B)x + (2A+B) \stackrel{!}{=} 1\ \forall x$

$\Rightarrow\ A+B = 0,\quad 2A+B = 1\ \Rightarrow\ A = 1,\ B = -1$

$\Rightarrow\ \int_0^1 \frac{1}{(x+1)(x+2)}\mathrm{d}x = \int_0^1 \frac{1}{x+1} - \int_0^1 \frac{1}{x+2}$
$= \ln(x+1)\big|_0^1 - \ln(x+2)\big|_0^2 = \ln(4/3)$.

Aufgaben

Berechnen Sie $\int_{-1}^1 (1+2x^3)^3 \mathrm{d}x$, $\int_0^1 \frac{x}{(x+1)(x+2)}\mathrm{d}x$, $\int_0^1 \frac{x(1-x)}{(x+1)(x^2+1)}\mathrm{d}x$
$\left(\rightsquigarrow \frac{38}{7},\ (6\ln(\frac{3}{2})-1),\ (\frac{\pi}{4}-\ln(2)) \right)$

3) Partielle Integration (Umkehrung der Produktregel)

Voraussetzung: $f(x)$, $g(x)$ differenzierbar.

Dann gilt: $\boxed{\int_a^b f(x)\,g'(x)\,\mathrm{d}x = f(x)\,g(x)\big|_a^b - \int_a^b f'(x)\,g(x)\,\mathrm{d}x}$.

(denn: $f(x)g(x)\big|_a^b = \int_a^b \mathrm{d}x \frac{\mathrm{d}}{\mathrm{d}x}\big[f(x)g(x)\big] = \int_a^b \mathrm{d}x\big[f'(x)g(x)\big] + \int_a^b \mathrm{d}x\big[f(x)g'(x)\big]$)

Beispiele

(i) $\int_0^y \mathrm{d}x\ \underbrace{x}_{f}\ \underbrace{\mathrm{e}^{x}{}'}_{g} = x\,\mathrm{e}^x\big|_0^y - \int_0^y \mathrm{d}x\,\mathrm{e}^x = x\,\mathrm{e}^x\big|_0^y - \mathrm{e}^x\big|_0^y = \mathrm{e}^x(x-1)\big|_0^y$

(ii) $\int_1^y \mathrm{d}x \ln(x) = \int_1^y \mathrm{d}x \cdot \underbrace{1}_{g'} \cdot \underbrace{\ln(x)}_{f} = x\ln x\big|_1^y - \int_1^y \mathrm{d}x x \frac{1}{x} = x(\ln(x)-1)\big|_1^y$

($\rightarrow$ Stammfunktion von $\ln(x)$ ist $x(\ln(x)-1)$)

(iii) $\int_0^y \mathrm{d}x \underbrace{\sin(x)}_{f} \underbrace{\mathrm{e}^x}_{g'} = \sin(x)\mathrm{e}^x\big|_0^y - \int_0^y \mathrm{d}x \underbrace{\cos(x)}_{f} \underbrace{\mathrm{e}^x}_{g'}$

$= \sin(x)\,\mathrm{e}^x\big|_0^y - \cos(x)\,\mathrm{e}^x\big|_0^y - \int_0^y \mathrm{d}x \sin(x)\,\mathrm{e}^x$

$\Rightarrow \int_0^y \sin(x)\,\mathrm{e}^x\,\mathrm{d}x = \frac{1}{2}\mathrm{e}^x(\sin(x) - \cos(x))\big|_0^y$

Aufgaben

Berechnen Sie $\int_0^y \mathrm{d}x x\sin(x)$, $\int_0^y \mathrm{d}x x^2\cos(x)$, $\int_0^y \mathrm{d}x x\ln(x)$, $\int_0^y \mathrm{d}x x^3 \mathrm{e}^{x^2}$, $\int_0^y \mathrm{d}x x\sinh(x)$

Zeigen Sie
- $\int \mathrm{d}x\, f'(x)x^n = f(x)x^n - n\int \mathrm{d}x\, f(x)\, x^{n-1}$
- $\int_0^\infty \mathrm{d}x\, \mathrm{e}^{-x}x^n = n!$ (z.B. mittels vollständiger Induktion)

4) Substitution (Umkehrung der Kettenregel)

Idee: Austausch der Integrationsvariablen $x \rightarrow u$ in $\int f(x)\,\mathrm{d}x$.

Voraussetzung: $x \leftrightarrow u$ hängen in umkehrbarer und stetig differenzierbarer Weise voneinander ab.
($u = g(x)$, $x = g^{-1}(u) = x(u)$, $\frac{\mathrm{d}x}{\mathrm{d}u} = x'(u)$ ist stetig).

Dann gilt: $\boxed{\int_{x_a}^{x_b} f(x)\,\mathrm{d}x = \int_{g(x_a)}^{g(x_b)} x'(u)\, f(x(u))\,\mathrm{d}u = \int_{u_a}^{u_b} \frac{\mathrm{d}x}{\mathrm{d}u}\, f(x(u))\,\mathrm{d}u}$.

(denn: Sei $F(x)$ Stammfunktion von $f(x)$,
also $F(x) = \int_a^x \mathrm{d}x'\, f(x')$, $\frac{\mathrm{d}}{\mathrm{d}x}F(x) = f(x)$
$\Rightarrow \frac{\mathrm{d}F}{\mathrm{d}u} = \frac{\mathrm{d}F}{\mathrm{d}x}\frac{\mathrm{d}x}{\mathrm{d}u} = f(x(u))\cdot x'(u) \;\Rightarrow\; F(x(u)) = \int \mathrm{d}u\, f(x(u))\, g'(u)$ ✓)

Beispiele

(i) $\int_1^5 \mathrm{d}x \sqrt{2x-1}$ | Setze $u = 2x-1 \;\Rightarrow\; x(u) = \frac{1}{2}(u+1)$, $\frac{\mathrm{d}x}{\mathrm{d}u} = \frac{1}{2}$

$= \int_{2\cdot 1-1}^{2\cdot 5-1} \mathrm{d}u \frac{\mathrm{d}x}{\mathrm{d}u}\sqrt{u} = \int_1^9 \mathrm{d}u\, \frac{1}{2}\sqrt{u} = \frac{1}{2}\frac{u^{3/2}}{3/2}\Big|_1^9 = \frac{1}{3}(9^{\frac{3}{2}} - 1^{\frac{3}{2}}) = \frac{26}{3}$

(ii) $\int \mathrm{d}t\, \cos(\omega t) \underbrace{=}_{u=\omega t} \int \mathrm{d}u\, \frac{\mathrm{d}t}{\mathrm{d}u}\, \cos(u) \underbrace{=}_{\mathrm{d}t/\mathrm{d}u=1/\omega} \frac{1}{\omega}\sin(u) = \frac{1}{\omega}\sin(\omega t)$

(iii) $\int_0^b \mathrm{d}t\, t\,\mathrm{e}^{-\alpha t^2} = \frac{1}{2\alpha}\int_0^{\alpha b^2} \mathrm{e}^{-u} = -\frac{1}{2\alpha}\mathrm{e}^{-u}\Big|_0^{\alpha b^2} = \frac{1}{2\alpha}(1 - \mathrm{e}^{-\alpha b^2})$
(mit $u = \alpha t^2 \;\Rightarrow\; \frac{\mathrm{d}u}{\mathrm{d}t} = 2\alpha t$ bzw. $\mathrm{d}u = 2\alpha t \mathrm{d}t$)

(iv) $\int \mathrm{d}x \tan(x) = \int \mathrm{d}x \sin(x)\frac{1}{\cos(x)} = -\int \frac{\mathrm{d}u}{u} = -\ln(u) = -\ln(\cos(x))$
(mit $u = \cos(x) \;\Rightarrow\; \frac{\mathrm{d}u}{\mathrm{d}x} = -\sin(x)$ bzw. $\mathrm{d}u = -\sin(x)\,\mathrm{d}x$)

Man sieht: Eine geeignete Substitution zu finden erfordert Intuition.
– geht nicht "von alleine"!

Aufgaben

Zeigen Sie mit der Methode der Substitution: $\int_C^x \mathrm{d}x' \frac{1}{x'+\alpha} = \ln(x'+\alpha) + \mathrm{const}$

Berechnen Sie
$\int_0^r \mathrm{d}x \mathrm{e}^{-2x/a}$, $\int_{-a}^a \mathrm{d}x\, \cosh(\frac{x}{A})$, $\int_{-a}^a \mathrm{d}x\, \sinh(\frac{x}{A})$, $\int \mathrm{d}x \sqrt{x \pm b}$, $\int_{-1}^1 \mathrm{d}z/\sqrt{az-b}$,
$\int_0^1 \mathrm{d}x \sqrt{1-x^2}$, (Tip: Substitution $x = \sin(\phi)$, Lösung ist $\frac{\pi}{4}$), $\int_0^r \mathrm{d}x \sqrt{r^2 - x^2}$,
$\int \mathrm{d}t\, \dot{x}(t)$, $\int \mathrm{d}t\, x(t)\dot{x}(t)$

Berechnen Sie
$\int_{-\pi}^{\pi} \mathrm{d}\phi \frac{\sin(\phi)}{\cos^2(\phi)+1}$, $\int \mathrm{d}x\, x\sqrt{x^2 \pm a}$, $\int \mathrm{d}x\, \frac{x}{1+x^2}$

Zeigen Sie

- $\int_{x_a}^{x_b} g'(x)\, g^n(x) = \frac{g^{n+1}(x)}{n+1}\Big|_{x_a}^{x_b}$
- $\int_{x_a}^{x_b} g'(x)/g(x) = \ln|g(x)|\Big|_{x_a}^{x_b}$
- $\int_{x_a}^{x_b} g'(x)\, \sqrt[n]{g(x)} = \frac{n}{n+1}\, \sqrt[n]{g(x)} g(x)\Big|_{x_a}^{x_b}$
- Bestimmen Sie die analoge Formel für $\int_{x_a}^{x_b} g'(x)/g^n(x)$

5) Integralfunktionen

Bei manchen Funktionen läßt sich das Integral nicht durch bekannte Funktionen ausdrücken → Integral definiert neue Funktion

Beispiele:

- **Error-Funktion**: $\mathrm{Erf}(y) = \int_0^y \exp(-x^2)\, \mathrm{d}x$.
- **Elliptische Integrale**: $F(k,\phi) = \int_0^\phi \frac{1}{\sqrt{1-k^2\sin^2(\psi)}}\, \mathrm{d}\psi$,
  $E(k,\phi) = \int_0^\phi \sqrt{1-k^2\sin^2(\psi)}\, \mathrm{d}\psi$.
- **Integralsinus**: $\mathrm{Si}(y) = \int_0^y \frac{\sin(x)}{x}\, \mathrm{d}x$.
- **Integralcosinus**: $\mathrm{Ci}(y) = -\int_y^\infty \frac{\cos(x)}{x}\, \mathrm{d}x$.
- **Integralexponentialfunktion**: $\mathrm{Ei}(y) = -\int_{-\infty}^y \frac{\exp(x)}{x}\, \mathrm{d}x$.
- **Integrallogarithmus**: $\mathrm{li}(y) = -\int_0^y \frac{1}{\ln(x)}\, \mathrm{d}x\ (y \neq 1)$.
  $\mathrm{Li}(y) = -\int_2^y \frac{1}{\ln(x)}\, \mathrm{d}x\ (y > 1)$.

  NB: Im Falle $\mathrm{li}(y)$, $y > 1$ wird über den Pol bei $x = 1$ im Sinne eines Cauchyschen Hauptwertes integriert, siehe 4.3.4.

6) Hilfsmittel

- Formelsammlungen, Integraltafeln, z.B.
  - Bronstein, Semendjajew, Taschenbuch der Mathematik
  - Abramowitz, Stegun, Handbook of Mathematical Functions
  - Gradshteyn, Ryzhik, Tables of Integrals, Series, and Products
- Symbolische Programme, z.B. MATHEMATICA, MAPLE
  Achtung: Teilweise fehlerhaft, vor allem wenn die Lösung Integralfunktionen beinhaltet. → Ergebnisse gegenchecken, z.B. durch Vergleich mit der numerischen Lösung für ausgewählte Zahlen.

7) Numerische Lösung (intelligent " Kästchen zählen")

Aufgaben

Berechnen Sie
(i) $\int \mathrm{d}x \sqrt{1+x^2}$,
(ii) $\int \mathrm{d}x\, \frac{x}{1+x^4}$,
(iii) $\int_0^1 \mathrm{d}x\, \frac{1}{1+x+x^2+x^3}$,
(iv) $\int \mathrm{d}x\, \ln(1+x^2)$,
(v) $\int \mathrm{d}x \arcsin(x)$

(Lösungswege: (i) Substitution $x = \sinh(u)$, (ii) Substitution $u = x^2$,
(iii) Partialbruchzerlegung, (iv) Partielle Ableitung von $\int dx \cdot 1 \cdot \ln(1+x^2)$, (v) Partielle Ableitung von $\int dx \cdot 1 \cdot \arcsin(x)$

Lösungen: (i) $\frac{1}{2}(x\sqrt{1+x^2} + \mathrm{arsinh}(x))$, (ii) $\frac{1}{2}\mathrm{artanh}(x^2)$, (iii) $\frac{1}{8}(\pi + 2\ln(2))$,
(iv) $\ln(1+x^2)(x - \frac{1}{2})$, (v) $x\arcsin(x) + \sqrt{1-x^2}$)

4.3.4 Uneigentliche Integrale

Bis jetzt: Lästige Einschränkungen für das Integral:
1) Intervallgrenzen endlich
2) Integrand $f(x)$ beschränkt

Manchmal kann man Integrale im Grenzwert auch ausrechnen, wenn Einschränkungen nicht zutreffen → "Uneigentliche Integrale".

1) Uneigentliche Integrale erster Art: Integrationsgrenze unendlich

- Definition
  Obere Grenze: $\int_a^\infty dx\, f(x) = \lim_{b\to\infty} \int_a^b dx\, f(x)$
  Untere Grenze: $\int_{-\infty}^b dx\, f(x) = \lim_{a\to-\infty} \int_a^b dx\, f(x)$
  Beide Grenzen: $\int_{-\infty}^\infty dx\, f(x) = \lim_{\substack{a\to-\infty \\ b\to\infty}} \int_a^b dx\, f(x)$
  Cauchy-Hauptwert: $P\int_{-\infty}^\infty dx\, f(x) := \lim_{a\to\infty} \int_{-a}^a dx\, f(x)$
  (falls $\int_{-\infty}^\infty dx\, f(x)$ nicht existiert,
  existiert vielleicht wenigstens der Hauptwert $P\int_{-\infty}^\infty dx\, f(x)$).
- Beispiele
  (i) $\int_0^\infty dx\, e^{-x} = \lim_{b\to\infty} \int_0^b dx\, e^{-x} = \lim_{b\to\infty}(-e^{-b} - e^0) = 1$
  (ii) $\int_{a>0}^\infty \frac{dx}{x} = \lim_{b\to\infty} \int_a^b dx\, \frac{1}{x} = \lim_{b\to\infty} \ln(x)\big|_a^b = \lim_{b\to\infty} \ln(\frac{b}{a}) \to \infty$
  (iii) $\int_{a>0}^\infty \frac{dx}{x^{1+\alpha}} = \lim_{b\to\infty} \int_a^b dx\, x^{-(1+\alpha)} = \lim_{b\to\infty} \frac{1}{\alpha}(a^{-\alpha} - b^{-\alpha})$

$$\Rightarrow \boxed{\int_{a>0}^\infty \frac{dx}{x^{1+\alpha}} = \begin{cases} a^{-\alpha}/\alpha & : \alpha > 0 \\ \infty & : \alpha \le 0 \end{cases}}$$

  (iv) $\int_{-\infty}^\infty dx\, \frac{x}{1+x^2} = \lim_{\substack{a\to-\infty \\ b\to\infty}} \int_a^b dx\, \frac{x}{1+x^2} = \lim_{\substack{a\to-\infty \\ b\to\infty}} [\ln(1+b^2) - \ln(1+a^2)]$
  → divergiert auf beiden Seiten.
  Aber: Hauptwert $P\int_{-\infty}^\infty dx \frac{x}{1+x^2} = 0$ existiert.
  (v) $\int_{-\infty}^\infty dx\, x$ existiert nicht, aber $P\int_{-\infty}^\infty dx\, x = 0$ existiert.

Vergleichskriterien:

Kriterien, mit denen man abschätzen kann, ob ein Integral überhaupt konvergiert.

a) Absolute Konvergenz
Falls $\int_a^\infty dx\, |f(x)|$ konvergiert, konvergiert auch $\int_a^\infty dx\, f(x)$.

b) Konvergente Majorante
$|f(x)| \le M(x)\ \forall x$ und $\int dx\, M(x)$ konvergiert. Dann konvergiert
$\int dx\, |f(x)| \le \int dx\, M(x)$,
und damit auch $\int dx\, f(x)$.

c) Divergente Minorante

$f(x) \geq m(x) \geq 0 \ \forall\ x$ und $\int \mathrm{d}x\, m(x)$ divergiert.
Dann divergiert auch $\int \mathrm{d}x\, f(x)$.

Beispiel: Abschätzung der Konvergenz von $\int_0^\infty \mathrm{d}x\, \mathrm{e}^{-x^2}$:

$x^2 > (x-1)\ \forall x \ \Rightarrow\ \mathrm{e}^{-x^2} \leq \mathrm{e}^{-(x-1)}$

Wegen $\int_0^\infty \mathrm{d}x\, \mathrm{e}^{-(x-1)} = \frac{1}{\mathrm{e}} < \infty$ konvergiert auch $\int_0^\infty \mathrm{d}x\, \mathrm{e}^{-x^2}$.

Aufgaben

Berechnen Sie $\int_0^\infty \mathrm{d}x\, \cos(x)\, \mathrm{e}^{-x}$, $\int_{-\infty}^\infty \mathrm{d}x/(1+x^2)$, $\int_{-\infty}^{-2/\pi} \mathrm{d}x \frac{\sin(1/x)}{x^2}$, $\int_{-\infty}^\infty x^n$, $P\int_{-\infty}^\infty x^n$

2) Uneigentliche Integrale zweiter Art:

Integrand unbeschränkt an einem Punkt x_0 ("divergiert", "hat Singularität").

(Beispiel: $1/x$ divergiert am Punkt $x_0 = 0$).

Definition:

Singularität an der unteren Grenze: $\int_{x_0}^{b>x_0} \mathrm{d}x f(x) = \lim_{\eta\to 0^+} \int_{x_0+\eta}^b \mathrm{d}x f(x)$.
Singularität an der oberen Grenze: $\int_{a<x_0}^{x_0} \mathrm{d}x f(x) = \lim_{\epsilon\to 0^+} \int_a^{x_0-\epsilon} \mathrm{d}x f(x)$.
Intervall $[a,b]$ schließt Singularität ein:

$$\int_a^b \mathrm{d}x\, f(x) = \lim_{\substack{\eta\to 0^+\\ \epsilon\to 0^+}} \left[\int_a^{x_0-\epsilon} \mathrm{d}x\, f(x) + \int_{x_0+\eta}^b \mathrm{d}x\, f(x) \right]$$

Cauchyscher Hauptwert:

$$P\int_a^b \mathrm{d}x\, f(x) = \lim_{\eta\to 0^+} \left[\int_a^{x_0-\epsilon} \mathrm{d}x\, f(x) + \int_{x_0+\epsilon}^b \mathrm{d}x\, f(x) \right]$$

Beispiele:

(i) $\int_0^b \mathrm{d}x \frac{1}{x} = \lim_{\eta\to 0^+} \int_\eta^b \mathrm{d}x \frac{1}{x} = \lim_{\eta\to 0^+} \ln|x|\Big|_0^b = \lim_{\eta\to 0^+} \ln|b/\eta| \to -\infty$

(ii) $\int_{-b}^b \mathrm{d}x\, \frac{1}{x}$ existiert dann natürlich auch nicht,
aber $P\int_{-b}^b \mathrm{d}x \frac{1}{x} = 0$ existiert.

(iii) $\int_0^b \frac{\mathrm{d}x}{x^{1+\alpha}} = \lim_{\eta\to 0^+} \int_\eta^b \mathrm{d}x\, \frac{1}{x^{1+\alpha}} = \lim_{\eta\to 0^+} \frac{1}{\alpha}(\eta^{-\alpha} - b^{-\alpha})$

$$\Rightarrow \boxed{\int_0^b \frac{\mathrm{d}x}{x^{1+\alpha}} = \begin{cases} \infty & : \ \alpha \geq 0 \\ -b^{-\alpha}/\alpha & : \ \alpha < 0 \end{cases}}$$

Vergleichskriterien

Analog den uneigentlichen Integralen erster Art.
– Falls $\int \mathrm{d}x\, |f(x)|$ konvergiert, dann auch $\int \mathrm{d}x\, f(x)$.
– Majorantenkriterium/Minorantenkriterium analog

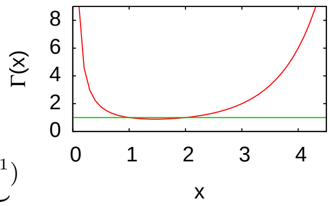
Aufgaben

Berechnen Sie $\int_0^1 \mathrm{d}x/\sqrt{1-x^2}$, $\int_0^{\pi/2} \mathrm{d}x\, \tan(x)$, $\int_0^\pi \mathrm{d}x\, \tan(x)$, $P\int_0^\pi \mathrm{d}x\, \tan(x)$,

3) Wichtige uneigentliche Integrale

Gaußintegral: $\int_{-\infty}^{\infty} \mathrm{d}x\, \mathrm{e}^{-x^2} = \sqrt{\pi}$ (Herleitung siehe 4.3.5).

Gamma-Funktion: $\Gamma(x) = \int_0^\infty \mathrm{d}t\, \mathrm{e}^{-t} t^{x-1} \quad (x > 0)$



- $\Gamma(1) = \int_0^\infty \mathrm{d}t\, \mathrm{e}^{-t} = 1$ (siehe 1) Beispiel (i))
- $\Gamma(x+1) = x\,\Gamma(x)$.
  (da $\Gamma(x+1) = \int_0^\infty \mathrm{d}t\, \mathrm{e}^{-t} t^{x} = \underbrace{-t^x \mathrm{e}^{-t}\big|_0^\infty}_{0} + x \underbrace{\int_0^\infty \mathrm{d}t\, \mathrm{e}^{-t} t^{x-1}}_{\Gamma(x)}$)

$\Rightarrow \Gamma(n+1) = n!$ für $n \in \mathbb{N}$
$\Gamma(x)$ interpoliert Fakultät.

- Weiterer spezieller Wert:
  $\Gamma(1/2) = \int_0^\infty \mathrm{d}x\, x^{-1/2} \mathrm{e}^{-x} \overset{(y=\sqrt{x})}{=} \frac{1}{2} \int_0^\infty dy \mathrm{e}^{-y^2} = \sqrt{\pi}$.

Beta-Funktion: $B(x,y) = \int_0^1 t^{x-1}(1-t)^{y-1} \mathrm{d}t \quad (x > 0, y > 0)$
Es gilt: $B(x,y) = \Gamma(x)\Gamma(y)/\Gamma(x+y)$

Elliptische Integrale:
$F(k,\phi) = \int_0^\phi \frac{1}{\sqrt{1-k^2 \sin^2(\psi)}} \mathrm{d}\psi = \int_0^{\sin\phi} \mathrm{d}x \frac{1}{\sqrt{(1-x^2)(1-k^2x^2)}}$,
$E(k,\phi) = \int_0^\phi \sqrt{1-k^2\sin^2(\psi)}\, \mathrm{d}\psi = \int_0^{\sin\phi} \mathrm{d}x \sqrt{\frac{1-k^2x^2}{1-x^2}}$,
Vollständige elliptische Integrale: $\mathrm{K} = F(k,\pi/2), \mathrm{E} = E(k,\pi/2)$.

4.3.5 Mehrfachintegrale

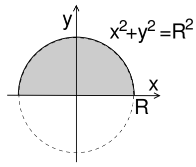
Mit den vorher behandelten Methoden kann man auch komplizierte Probleme lösen, z.B. verschachtelte Integrale.
Wichtige Anwendung: Berechnung von Volumina

4.3.5.1 Beispiele

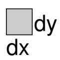
(i) Fläche A eines Kreises des Radius R

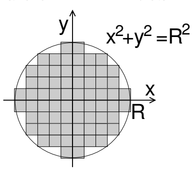
Erste Lösung: Fläche des Halbkreises über Einfachintegral

$y = \sqrt{R^2 - x^2}$
$\Rightarrow A/2 = \int_{-R}^{R} \mathrm{d}x\, \sqrt{R^2-x^2} = 2 \underbrace{\int_0^R \mathrm{d}x \sqrt{R^2-x^2}}_{R^2 \frac{\pi}{4} \text{ lt. Übungsaufgabe}} = 2R^2 \frac{\pi}{4}$
$\Rightarrow A = \pi R^2$



Zweite Lösung: Alternative Sichtweise: – Integriere (d.h. summiere) über infinitesimale Flächenelemente $\mathrm{d}A = \mathrm{d}x\, \mathrm{d}y$



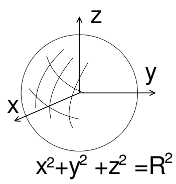


$$A = \iint_{\text{Kreis}} \mathrm{d}A = \int_{-R}^{R} \mathrm{d}x \int_{-\sqrt{R^2-x^2}}^{\sqrt{R^2-x^2}} \mathrm{d}y$$
$$= \int_{-R}^{R} \mathrm{d}x\, 2\sqrt{R^2-x^2} = 4\int_0^R \mathrm{d}x\, 2\sqrt{R^2-x^2} = \pi R^2$$

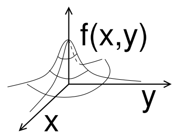
(ii) Volumen V einer Kugel des Radius R

Integriere über infinitesimale Volumenelemente $\mathrm{d}V = \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z$.

$$V = \iiint_{\text{Kugel}} \mathrm{d}V = \int_{-R}^{R} \mathrm{d}x \int_{-\sqrt{R^2-x^2}}^{\sqrt{R^2-x^2}} \mathrm{d}y \int_{-\sqrt{R^2-x^2-y^2}}^{\sqrt{R^2-x^2-y^2}} \mathrm{d}z$$
$$= \int_{-R}^{R} \mathrm{d}x \int_{-\sqrt{R^2-x^2}}^{\sqrt{R^2-x^2}} \mathrm{d}y\, 2\sqrt{R^2 - x^2 - y^2}$$
$$= \int_{-R}^{R} \mathrm{d}x 4 \underbrace{\int_{0}^{\sqrt{R^2-x^2}} \mathrm{d}y\, 2\sqrt{(R^2 - x^2) - y^2}}_{(R^2-x^2)\frac{\pi}{4} \text{ lt. Übungsaufgabe}}$$
$$= \pi \int_{-R}^{R} \mathrm{d}x\, (R^2 - x^2)$$
$$= \pi(2R^3 - R^2 \tfrac{x^3}{3}\Big|_{-R}^{R}) = \pi R^3(2 - \tfrac{2}{3}) = \tfrac{4}{3}\pi R^3$$



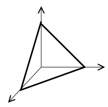
(iii) Volumen unter einer Gaussglocke



$$\iint_{\infty} \mathrm{d}A\, f(x,y) := \int_{\infty}^{\infty} \int_{\infty}^{\infty} \mathrm{d}x\mathrm{d}y\, f(x,y)$$
$$= \iint_{\infty} \mathrm{d}x\,\mathrm{d}y\, \mathrm{e}^{-x^2-y^2} = \int_{-\infty}^{\infty} \mathrm{d}x\, \mathrm{e}^{-x^2} \int_{-\infty}^{\infty} \mathrm{d}y\, \mathrm{e}^{-y^2}$$
$$= \Big[\underbrace{\int_{-\infty}^{\infty} \mathrm{d}x \mathrm{e}^{-x^2}}_{\text{Nachschlagen: } \sqrt{\pi}} \Big]^2 = \sqrt{\pi}^2 = \pi.$$

Aufgaben

- Berechnen Sie die Funktion $f(x,y)$ auf dem Gebiet G
  $f(x,y) = 1$ auf $G = \{(x,y) : 2 \le x \le 3, x^2 \le y \le x^3\}$
  $f(x,y) = y\sqrt{1-x^2}$ auf $G = \{(x,y) : x^2 + y^2 \le 1, y \ge 0\}$
- Berechnen Sie das Volumen eines rechtwinkligen Dreiecks der Seitenlänge 3cm
  $V = \int_0^1 \mathrm{d}x \int_0^{1-x} \mathrm{d}y$
- Berechnen Sie das Volumen eines dreidimensionalen Simplex der Seitenlänge a:
  $V = \int_0^1 \mathrm{d}x \int_0^{1-x} \mathrm{d}y \int_0^{1-x-y} \mathrm{d}z$
- Berechnen Sie das Volumen eines d-dimensionalen Simplex der Seitenlänge 1

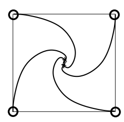


Abschlussaufgaben zum Thema Infinitesimalrechnung

1) Ein Auto fährt die Hälfte einer Strecke mit 50 km/h. Wie schnell muss es die andere Hälfte fahren, um eine Durchschnittsgeschwindigkeit von 100 km/h zu erzielen?

2) Man stelle sich vor, eine Reihe von n Ziegelsteinen sei übereinander gestapelt, ohne umzufallen. Das Zentrum des untersten Ziegelsteins sei bei $x = 0$. Wie weit kann der oberste Ziegelstein im Grenzfall $n \to \infty$ maximal von $x = 0$ weg stehen?

3) Vier Mäuse sitzen an den Ecken eines quadratischen Zimmers. Zum Zeitpunkt $t = 0$ laufen alle gleichzeitig und gleich schnell los. Dabei läuft jede Maus zu jedem Zeitpunkt in Richtung ihres rechten Nachbarn. Irgendwann treffen sie sich alle in der Mitte. Welche Strecke haben sie bis dahin zurückgelegt?



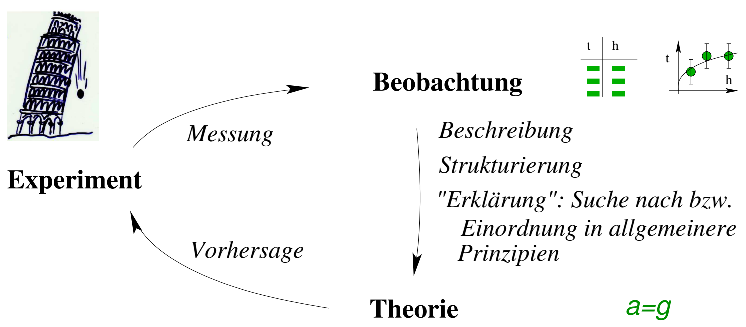
Lösungen 1): Unendlich schnell; 2): Unendlich weit (harmonische Reihe); 3): Die Seitenlänge des Zimmers (Mäuse laufen immer senkrecht zueinander und aufeinander zu, bis sie sich erreicht haben.)

Kapitel 5

Statistik und Fehlerrechnung

5.1 Motivation und Einführung

Erinnerung Kapitel 1.1: Kreislauf der empirischen Wissenschaften

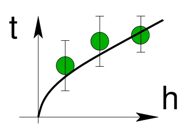


Quantifizierung der Beobachtungen:

Typischerweise Werte mit Fehlerbalken

z.B. Fallversuche: Wertetabelle

Fallhöhe h	Fallzeit t
1 m	(4.9 ± 0.9) s
2 m	(7.2 ± 0.9) s
3 m	(7.5 ± 0.5) s



Ziel: – Auffinden/Test von Gesetzmäßigkeiten, z.B. $t = \sqrt{2gh}$

– Bestimmung von Naturkonstanten, z.B. $g = (9.81 \pm 0.5)$ m/s^2

Dieses Kapitel: Diskussion dieser Fehler

1) Was versteht man unter einem Fehler?
   (Interpretation? Mathematische Beschreibung?)
2) Wie bestimmt man Fehler in Experimenten (oder Simulationen)?

Zunächst zur ersten Frage:
Was versteht man unter Fehlern und warum gibt es sie überhaupt?

Beobachtung: Führt man Experimente mehrfach durch oder misst man eine Größe (z.B. g) mehrfach auf verschiedene Weise, so kommt fast nie genau dasselbe heraus.

Ursachen

- Fehler im Versuchsaufbau, ungünstiger Versuchsaufbau, schlechte Kalibrierung, unvollständige Theorie
(z.B. Fallexperiment: Uhr geht falsch, Wind, Luftreibung)

  $\leadsto$ **Systematische Fehler**

  Können im Prinzip behoben werden (wenn auch nicht unbedingt in der Praxis), z.B. durch Verbesserung des Versuchsaufbaus, bessere Kalibrierung und/oder Verbesserung der Theorie

  Verzerren die Messergebnisse in der Regel systematisch!

- Selbst in durchoptimierten Versuchen streuen Messergebnisse immer noch. Das läßt sich nie ganz beheben.

  Die Gründe sind vielfach: Zumindest die Messapparatur ist ein nicht vollständig kontrollierbares Vielteilchensystem, thermisches Rauschen spielt immer eine Rolle (Temperatur Null ist nicht erreichbar), Umwelteinflüsse können nie ganz abgeschirmt werden, Weltall, quantenmechanische Unschärfe, ...

  $\leadsto$ **Statistische Fehler**

  Können nicht behoben werden.

  Aber: Mitteln sich weg!

In diesem Kapitel geht es um die statistischen Fehler

Interpretation: Wofür steht die Aussage " $y = y_0 \pm \Delta y$ "?

Mögliche Antworten:

- Aussage über die Wahrscheinlichkeit, in einem künftigen identischen Experiment zur Messung von y einen Wert im Intervall $y \in [y_0 - \Delta y : y_0 + \Delta y]$ zu finden (unter bestimmten Voraussetzungen 0.68 %, siehe Kapitel 5.3.3).

- oder: Aussage über die Wahrscheinlichkeit, dass der tatsächliche Wert von y im Bereich $y \in [y_0 - \Delta y : y_0 + \Delta y]$ liegt.

$\leadsto$ Aussagen über **Wahrscheinlichkeiten**
Das heisst: Als erstes müssen wir uns mit dem Begriff der Wahrscheinlichkeit befassen!

5.2 Grundbegriffe der Wahrscheinlichkeitstheorie

5.2.1 Konzept und Interpretation von Wahrscheinlichkeit

Wir beginnen mit zwei Beispielen.

Beispiel 1: Wurf eines perfekten Würfels

Aussage: "*Die Wahrscheinlichkeit, die Zahl 2 zu würfeln, ist 1/6.*"
Gemeint ist im Allgemeinen:

"*Das Ergebnis des Wurfes ist offen. Wenn ich den Wurf $N \to \infty$ mal wiederhole und die Zahl n_2 der Fälle zähle, in denen die Zahl 2 gewürfelt wurde, dann ist fast immer $p_2 = \lim_{N\to\infty} \frac{n_2}{N} = \frac{1}{6}$.*"

$\leadsto$ **Frequentistischer Wahrscheinlichkeitsbegriff**

Wahrscheinlichkeit wird interpretiert als eine objektive Größe, die die relative Häufigkeit eines Ereignisses nach einer großen Zahl gleicher, unabhängiger Zufallsexperimente beschreibt.

Beispiel 2: Fallexperiment

Aussage: "*Die Erdbeschleunigung g liegt mit einer Wahrscheinlichkeit von 68 % im Intervall $g \in [9.80 : 9.82] m/s^2$*"
Gemeint ist im Allgemeinen

"*g ist ist eine Naturkonstante und hat natürlich einen festen Wert, der unabhängig von Experimenten ist. Aber: Ich kenne den Wert von g nicht. Basierend auf meinen Experimenten wette ich 0.68:0.32, dass er im Intervall $g \in [9.80 : 9.82] m/s^2$ liegt!*"

$\leadsto$ **Bayes-scher Wahrscheinlichkeitsbegriff**

Wahrscheinlichkeit wird interpretiert als subjektive Aussage im Sinne einer Prognose für den Wert einer Größe basierend auf dem Grad des persönlichen Wissens.

$\leadsto$ Zwei verschiedene Auffassungen von Wahrscheinlichkeit, in der Vergangenheit kontrovers diskutiert.

- Aussagen über Fehlerabschätzungen in der Physik liegt oft implizit eine Bayessche Auffassung zugrunde. (Die frequentistische Alternative wäre eine Vielweltentheorie, in der Naturkonstanten viele mögliche Werte annehmen können!)
- Bei der Herleitung statistischer Aussagen und der Konzeption statistischer Tests ist es oft einfacher, mit einem frequentistischen Wahrscheinlichkeitsbegriff zu arbeiten.

Formal mathematisch kann man beide Wahrscheinlichkeitsbegriffen in ähnlicher Weise beschreiben. Wir beginnen mit dem intuitiven Zugang der "klassischen Wahrscheinlichkeiten" und führen dann einen allgemeiner einsetzbaren mathematischen Zugang ein.

5.2.2 Klassisch-intuitive Berechnung von Wahrscheinlichkeiten

Frage: Wie kann man Wahrscheinlichkeiten ermitteln?

Antwort: Leider gibt es keine allgemeingültige Regel. Man muss immer Annahmen machen.

5.2.2.1 "Klassische" Wahrscheinlichkeit

Wenn bei einem Versuch nur eine endlich große Anzahl von Ergebnissen möglich ist (z.B. Würfeln, eine Münze werfen, Roulette spielen) und man sonst keine weiteren Informationen hat (z.B., ob der Würfel gezinkt ist), dann macht man häufig die sogenannte "klassische" Annahme: Die Wahrscheinlichkeit ist gleichverteilt auf alle möglichen Ergebnisse (dann auch "Elementarereignisse" genannt).

Unter der klassischen Annahme lassen sich Wahrscheinlichkeiten für "Ereignisse" E wie z.B. "Mit einem Würfel wird eine gerade Zahl gewürfelt" einfach berechnen. Die Regel lautet:

$$\boxed{P(E) = \frac{\text{Anzahl der günstigen Ergebnisse}}{\text{Anzahl aller mögliche Ergebnisse}}}$$

wobei die "günstigen" Ergebnisse die sind, bei denen das "Ereignis" eintritt. In unserem Beispiel gibt es insgesamt 6 mögliche Ergebnisse des Würfelwurfs, und 3 davon sind günstig (gerade Zahl). Also ist $P(E) = 3/6 = 1/2$.

Beispiele:

(i) Wahrscheinlichkeit, in n Würfen nie eine Sechs zu würfeln
- Für n=1: Günstige Fälle: Zahlen (1...5): 5
  Mögliche Fälle: Zahlen (1...6): 6
  $\Rightarrow P = (\frac{5}{6})$
- Für n=2: Günstige Fälle: 2 Mal Zahlen (1...5): 5^2
  Mögliche Fälle: 2 Mal Zahlen (1...6): 6^2
  $\Rightarrow P = (\frac{5}{6})^2$
- Allgemein: $P = (\frac{5}{6})^n$ (wird sehr klein bei großen n)

(ii) Wahrscheinlichkeit, in n Würfen mindestens eine Sechs zu würfeln
Komplementär zur Wahrscheinlichkeit (i) $\Rightarrow P = 1 - (\frac{5}{6})^n$

(iii) Wahrscheinlichkeit, in n Würfen genau einmal eine Sechs zu würfeln:
- Mögliche Fälle: 6^n
- Günstige Fälle: $\underbrace{5^{n-1}}_{[(n-1)\text{ mal keine 6}]} \cdot \underbrace{n}_{[\text{welcher Wurf die 6 war}]}$

$\Rightarrow P = (\frac{5}{6})^{n-1} \frac{1}{6} n$

(iv) Wahrscheinlichkeit, nach 20 Sechsen wieder Sechs zu würfeln:
$\rightsquigarrow$ Wie (i), denn der 21. Wurf ist unabhängig von den Würfen davor!

Zur Berechnung "möglicher" und "günstiger" Anzahlen werden häufig Methoden der **Kombinatorik** eingesetzt. Deshalb folgt jetzt ein kleiner Exkurs dazu.

5.2.2.2 Die vier Grundprobleme der Kombinatorik

(Aus irgendeinem Grund werden kombinatorische Formeln immer anhand von Urnenmodellen eingeführt!)

a) Variationen: Wie viele Möglichkeiten gibt es, aus einer Urne mit N Kugeln nacheinander k Kugeln in bestimmter Reihenfolge zu ziehen?

Antwort: $N(N-1)\cdots(N-k+1) = \binom{N}{k}\, k!$

(Für die Wahl der ersten Kugel hat man N Möglichkeiten, für die zweite nur noch $N-1$, für die dritte $N-2$ etc.)

Praktisches Beispiel: Es gibt $N(N-1)(N-2)$ mögliche Konstellationen von Gold-, Silber- und Bronzemedaillenträgern in einer Gruppe von N Sportlern.

b) Kombinationen: Wie viele Möglichkeiten gibt es, aus einer Urne mit N Kugeln k-Mal eine Kugel zu ziehen, wenn die Reihenfolge egal ist?

Antwort: $\binom{N}{k} = \frac{N(N-1)\cdots(N-k+1)}{k!}$

(Es gibt $N(N-1)\cdots(N-k+1)$ Möglichkeiten, k Kugeln auszuwählen, aber da es $k!$ Möglichkeiten gibt, diese anzuordnen, und die Reihenfolge egal ist, sind davon jeweils $k!$ äquivalent.)

Praktische Beispiele: Es gibt $\binom{N}{3}$ Möglichkeiten, unter N Sportlern ein Dreierteam auszuwählen. Wenn in einer Gruppe von N Party-gästen jeder mit jedem einmal anstösst, erklingt $\binom{N}{2}$ mal ein Glas.

c) Variationen mit Wiederholungen: Wie viele Möglichkeiten gibt es, aus einer Urne mit N Kugeln k-Mal in bestimmter Reihenfolge eine Kugel zu kopieren, ohne sie zu entnehmen?

Antwort: N^k

(Für die Wahl der ersten Kugel hat man N Möglichkeiten, für die zweite auch wieder N, da man die erste ja zurückgelegt hat, für die dritte N etc.

Praktisches Beispiel: Es gibt N^k Möglichkeiten, aus N Symbolen (Ziffern) einen Code der Länge k zu generieren.

d) Kombinationen mit Wiederholungen: Wie viele Möglichkeiten gibt es, aus einer Urne mit N Kugeln k-Mal eine Kugel zu kopieren ohne sie zu entnehmen, wenn die Reihenfolge egal ist?

Antwort: $\binom{N+k-1}{k}$

(Man kann sich das Ergebnis so visualisieren, als ob k Kopien in N Fächern einsortiert seien (jedes Fach steht für einen Kugeltyp). Wir haben also Sequenzen von k Kopien und $(N-1)$ Fächerwänden, die beliebig verteilt sein können. Es gibt $(k+N-1)!$ Möglichkeiten, diese anzuordnen, allerdings sind ist die Reihenfolge der Fächerwände egal und die der Kopien auch, d.h. $(N-1)! \cdot k!$ Anordnungen sind jeweils äquivalent. Also gibt es insgesamt $(k+N-1)!/(k!(N-1)!) = \binom{N+k-1}{k}$ verschiedene Möglichkeiten.)

Praktisches Beispiel: Es gibt $\binom{3+n-1}{n}$ Möglichkeiten, n Wählerstimmen auf 3 Kandidaten zu verteilen.

Beispiele für Anwendungen kombinatorischer Formeln bei der Berechnung klassischer Wahrscheinlichkeiten

(i) Wahrscheinlichkeit, bei n Würfen einer Münze k mal Zahl zu erhalten.
- Mögliche Fälle: 2^n (Grundproblem (c))
- Günstige Fälle: $\binom{n}{k}$ (Grundproblem (b))

$\Rightarrow\ P(n) = \binom{n}{k}(\frac{1}{2})^n$.

(ii) Wahrscheinlichkeit, im Lotto genau $n \leq 6$ Richtige zu tippen
- Mögliche Fälle: $\binom{49}{6}$ (Grundproblem (b))
- Günstige Fälle: $\binom{6}{n}\binom{43}{6-n}$ (Grundproblem (b), mehrfach)

Erläuterungen:
- $\binom{49}{6}$: Möglichkeiten, 6 aus 49 Kugeln zu ziehen, 43 verbleiben in der Urne.
- $\binom{6}{n}$: mögliche Verteilungen von n richtig getippten Kugeln auf die 6 gezogenen Kugeln.
- $\binom{49-6}{6-n}$: mögliche Verteilungen der $6-n$ falsch getippten Kugeln auf die 43 Kugeln in der Urne.

$\Rightarrow\ P(n) = \frac{\binom{6}{n}\binom{43}{6-n}}{\binom{49}{6}}$

(iii) Wahrscheinlichkeit, dass in einer Gruppe von N Personen zwei am selben Tag Geburtstag haben. (Schalttage werden ignoriert).

Hier ist es einfacher, die Wahrscheinlichkeit P^* für das entgegengesetzte Ereignis zu bestimmen: Keine zwei Personen haben am selben Tag Geburtstag.
- Mögliche Fälle: 365^N (Grundproblem (c))
- Günstige Fälle: $\binom{365}{N}\,N!$ (Grundproblem (a))

$\Rightarrow\ P^* = \frac{\binom{365}{N}N!}{365^m}$ und $P = 1 - P^* = 1 - \frac{\binom{365}{N}N!}{365^N}$

5.2.2.3 Grenzen der klassischen Beschreibung

In vielen Fällen kann die klassische Annahme, dass die Wahrscheinlichkeiten auf Elementarereignisse gleichverteilt ist, nicht angewendet werden. Beispiele sind:

⋆ Es liegen zusätzliche Informationen vor, z.B. dass Würfel gezinkt sind.

⋆ Es gibt unendlich viele "Elementarereignisse". Wenn man z.B. die Wahrscheinlichkeit bestimmen will, eine bestimmte natürliche Zahl $n \in \mathbb{N}$ zu ziehen, dann kann man keine klassische Annahme machen, man brauche Zusatzinformationen.

⋆ Der Raum der Ergebnisse ist kontinuierlich. Dann ist macht die Gleichsetzung von Ergebnissen mit "Elementarereignissen" nicht einmal mehr Sinn. Zum Beispiel ist die Wahrscheinlichkeit, dass ein Schütze "genau" das Zentrum einer Scheibe trifft, immer Null, egal wie gut der Schütze ist.

Aus diesen Gründen ist eine Erweiterung und Formalisierung des Wahrscheinlichkeitsbegriffes notwendig.

5.2.3 Mathematischer Wahrscheinlichkeitsbegriff

Wichtig im Zusammenhang mit Messungen: Die physikalische Welt ist weitgehend kontinuierlich. Punkte im Kontinuum sind nicht messbar, deshalb kann man aus Messungen de facto immer nur Intervalle ablesen. Eine dazu passende mathematische Struktur muss auf Intervallen bzw. allgemeiner auf Teilmengen von Ergebnismengen aufbauen (Kolmogorov, 1903-1987)

⋆ Mathematisch wird Wahrscheinlichkeit über einen sogenannten "Wahrscheinlichkeitsraum" definiert. Die Grundidee (mathematisch nicht sauber formuliert) ist wie folgt:

- Alle möglichen Ergebnisse eines Versuchs bilden zusammen die Grundmenge Ω.
- Jedes Ereignis entspricht einer Teilmenge ω von Ω, und ihr wird eine Wahrscheinlichkeit $P(\omega)$ zugeordnet. Diese beschreibt die Wahrscheinlichkeit, dass Versuchsergebnis in der Teilmenge ω liegt.
- Die Wahrscheinlichkeiten von zwei Teilmengen, die sich nicht überschneiden, addieren sich auf. Die Wahrscheinlichkeit, dass überhaupt ein Ergebnis erzielt wird, ist $P(\Omega) = 1$. Die Wahrscheinlichkeit, dass es gar kein Ergebnis gibt, ist $P(\varnothing) = 0$.

⋆ Beispiele

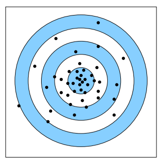
(i) Wurf einer Münze

Grundmenge der möglichen Ergebnisse: $\Omega = \{B, Z\}$ (Bild oder Zahl)
Als mögliche Ereignisse zählen

- $\varnothing$ (Weder Bild noch Zahl): Wahrscheinlichkeit $P(\varnothing) = 0$
- $\{B\}$ (Bild): Wahrscheinlichkeit $P(\{B\}) = 1/2$
- $\{B\}$ (Zahl): Wahrscheinlichkeit $P(\{Z\}) = 1/2$
- $\{B, Z\}$ (Bild oder Zahl): Wahrscheinlichkeit $P(\{B, Z\}) = 1$

(ii) Schuss auf Zielscheibe

- Grundmenge der Ergebnisse: $\Omega = \mathbb{R}^2$ (gesamte Fläche)
- Mögliche Ereignisse: ω Alle Teilflächen (z.B. blaue, weiße Ringe)
- $P(\omega)$: Wahrscheinlichkeit, die Teilfläche ω zu treffen (hängt vom Schützen ab).



⋆ Zufallsgrößen: Oft interessiert man sich nicht für das Ereignis ω selbst, sondern für eine davon abgeleitete Zahl $Z(\omega)$, z.B. die Punktzahl, die man beim Schuss auf die Zielscheibe erzielt hat.

Eine solche Zahl (bzw. Funktion) $Z(\omega)$ nennt man Zufallsvariable.

5.2.4 Bedingte Wahrscheinlichkeit

Betrachte zwei Ereignisse A, B. Die Wahrscheinlichkeit von A, falls B vorgegeben ist, wird bedingte Wahrscheinlichkeit genannt, und ist gegeben durch: $P(A|B) = P(A \cap B)/P(B)$

Beispiel: Europawahl 2019 in Deutschland
Gesucht ist die Wahrscheinlichkeit p, dass ein Jungwähler (unter 25) DIE PARTEI gewählt hat
– 7% der Wähler waren unter 25
– 0.56% der Wähler waren unter 25 und haben DIE PARTEI gewählt
$\Rightarrow$ $p = \frac{0.0056}{0.07} = 0.08$=8%.
(Zum Vergleich: Insgesamt hat DIE PARTEI 2.4% der Stimmen erhalten).

⋆ Wichtige Folgerung: **Satz von Bayes** $\boxed{P(A|B) = \dfrac{P(B|A)\,P(A)}{P(B)}}$

(Beweis: Umstellen $P(A \cap B) = P(A|B)\,P(B) = P(B|A)\,P(B)$)

Beispiel: Sie haben eine Reise in ein exotisches Land gemacht und lassen sich auf eine Krankheit testen, die 0.1% der Rückkehrer sich eingefangen haben. Der Test hat eine Spezifizität von 99% und eine Sensitivität von 98%. Er ist bei Ihnen positiv. Besteht Grund zur Panik?

Erklärung: Spezifizität ≙ correct positive, Sensitivität ≙ correct negative
Rechnung: Gesucht ist $P(A|B)$, wobei A ≙ " krank", B ≙ "Test positiv".
Gegeben ist: $P(A) = 0.001$, $P(B|A) = 0.99$, $P(\overline{B}|\overline{A}) = 0.98$
($\overline{A} := \Omega \setminus A$ ist die Komplementmenge von A)
Daraus folgt: $P(\overline{A}) = 1 - P(A) = 0.999$ und $P(B|\overline{A}) = 1 - P(\overline{B}|\overline{A}) = 0.02$
$\Rightarrow$ $P(B) = P(B|A)P(A) + P(B|\overline{A})P(\overline{A}) = 0.99 \cdot 0.001 + 0.02 \cdot 0.999 = 0.021$
$\Rightarrow$ $P(A|B) = P(B|A)P(A)/P(B) = 0.99 \cdot 0.001/0.021 = 0.047$= 4.7%

$\rightsquigarrow$ Die Wahrscheinlichkeit, dass Sie tatsächlich infiziert sind, liegt nur bei 4.7%.
Also nicht verzagen und weitere Tests machen!

⋆ **Statistische Unabhängigkeit**:

Zwei Ereignisse A, B heißen statistisch unabhängig, wenn $P(A|B) = P(A)$

5.2.5 Verteilungen

Als nächstes diskutieren wir Verteilungen von Zufallsgrößen.

Verteilungen beschreiben im Prinzip, "mit welcher Wahrscheinlichkeit eine Zufallsgröße einen bestimmten Wert hat". Im Fall von kontinuierlichen Zufallsgrößen müssen wir allerdings noch spezifizieren, was wir damit genau meinen.

Beispiele:

(i) Wurf eines Würfels: Die Zufallsvariable ist die gewürfelte Zahl n
$\rightsquigarrow$ Diskrete Verteilung p_n: Wahrscheinlichkeit, n zu würfeln.
mit Normierung: $\sum_{n=1}^{6} p_n = 1$ (Klassische Annahme: $p_n = 1/6\ \forall n$)

(ii) Radioaktiver Zerfall: Die Zufallsvariable ist die Zerfallszeit t
$\rightsquigarrow$ Jeder individuelle Zerfall hat eine andere Zerfallszeit.
$\rightsquigarrow$ Definition der Verteilung nicht so offensichtlich

Ansatz: kontinuierliche Funktion $p(t)$ mit $\int_0^\infty p(t)\mathrm{d}t = 1$ (Normierug), so dass $\int_{t_0}^{t_1} p(t)\mathrm{d}t$ die Wahrscheinlichkeit angibt, eine Zerfallszeit im Intervall $[t_0, t_1]$ zu messen.
(Hier: Exponentieller Zerfall: $p(t) = \frac{1}{\tau} e^{-t/\tau}$)

Bemerkung: Die Funktion $p(t)$ ist nicht dimensionslos (sie hat die Dimension von $1/t$) und hängt damit von der Einheit ab. Ein besserer Ausgangspunkt für eine saubere Definition von Verteilungen wäre das dimensionslose Integral $P_c(t) = \int_0^t p(t')\mathrm{d}t'$: Die Wahrscheinlichkeit, dass ein Atom zur Zeit t bereits zerfallen ist.

5.2.5.1 Definitionen

Motiviert durch die vorigen Beispiele

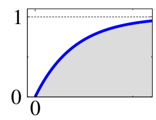
(i) **Kumulative Verteilungsfunktion**: Betrachte Zufallsvariable Z mit möglichen Werten z (bzw. z_n mit $n \in \mathbb{Z}$, falls es nur diskrete Werte gibt, aufsteigend angeordnet)

$$P_c(z) := P(Z \leq z)$$

Diskrete Ereignisse (z.B. Würfel): Diskrete Stufen $P_c(z_n) = \sum_{j=n_{\min}}^{n: z_n < z < z_{n+1}} p_j$

Kontinuierliches Spektrum von Ereignissen: Funktion $P_c(z)$

Beispiel radioaktiver Zerfall: $P_c(t) = 1 - \mathrm{e}^{-t/\tau}$
$\triangleq$ Anteil Atome, die zur Zeit t zerfallen sind.

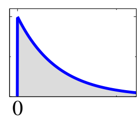


(ii) **Wahrscheinlichkeitsdichteverteilung**

Diskrete Ereignisse: p_n

Kontinuierliches Spektrum von Ereignissen: $p(z) = \frac{\mathrm{d}}{\mathrm{d}z} P_c(z)$

Beispiel radioaktiver Zerfall: $p(t) = \frac{1}{\tau}\, \mathrm{e}^{-t/\tau}\, \theta(t)$

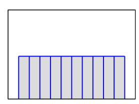


$\to$ Wohldefinierte Definition auch für kontinuierliche Zufallszahlen.
Im diskreten Fall (d.h. wenn Z nur diskrete Werte z_i annehmen kann), ist $f_Z(z) = \sum_i p_i\, \delta(Z - z_i)$

Bemerkung: $\mathrm{d}z\, p(z) = P(Z \in [z, z + \mathrm{d}z]\})$ kann man als (infinitesimale) Wahrscheinlichkeit auffassen, die Dichtefunktion $p(z)$ dagegen nicht. Zum Beispiel hat $p(z)$ im Allgemeinen eine Einheit, und der Wert hängt von der Einheit ab.

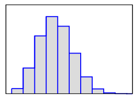
5.2.5.2 Wichtige Verteilungen

Diskrete Verteilungen



- $\star$ Gleichverteilung: $p_n = \text{const.}$

⋆ Binomialverteilung: $p_n = q^n\,(1-q)^{(N-n)}\,\binom{N}{n}$



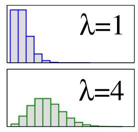
Beispiel: Lasse ein Butterbrot N Mal fallen, das jeweils mit Wahrscheinlichkeit q auf die Butterseite fällt ($q > 1/2$).
p_n: Wahrscheinlichkeit, dass es genau n Mal auf die Butterseite fällt

· Kombinatorischer Faktor $\binom{N}{n}$: Möglichkeiten, n Butterseiten-Fälle auf insgesamt N Fälle zu verteilen.

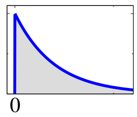
· q^n: Wahrscheinlichkeit, dass es in diesen n Fällen tatsächlich auf die Butterseite fällt.

· $(1-q)^{N-n}$: Wahrscheinlichkeit, dass es in den restlichen $(N-n)$ Fällen nicht auf die Butterseite fällt.

⋆ Poissonverteilung: Grenzfall der Binomialverteilung
für $q \to 0$, $N \to \infty$, $qN = \text{const.} = \lambda \Rightarrow p_n = \mathrm{e}^{-\lambda}\frac{\lambda^n}{n!}$
Beispiel: Wahrscheinlichkeit für n seltene Ereignisse, z.B. atomare Zerfälle pro Zeiteinheit



Kontinuierliche Verteilungen

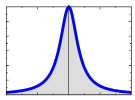


⋆ Exponentialverteilung: $p(z) = \Theta(z)\,\frac{1}{\beta}\,\mathrm{e}^{-z/\beta}$

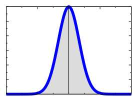
⋆ Lorentzverteilung: bzw. Cauchy-Verteilung:

$$p(z) = \frac{1}{\pi}\frac{\lambda}{\lambda^2+(z-\mu)^2}$$

Beispiel: Linienbreiten in der Physik



⋆ Gaußverteilung: $p(z) = \frac{1}{\sqrt{2\pi\sigma^2}}\mathrm{e}^{-(z-\mu)^2/(2\sigma^2)}$
Die wichtigste aller Verteilungen
Grund: Siehe Kapitel 5.2.5.3, zentraler Grenzwertsatz



5.2.5.3 Charakterisierung von Verteilungen

Gegeben eine Zufallsvariable Z mit Verteilungsfunktion $p(z)$. Wir diskutieren verschiedene Kenngrößen der Verteilung.
Am bekanntesten sind

- **Erwartungswert** $\langle Z\rangle = \int \mathrm{d}z\, p(z)\, z$ bzw. diskret $\langle n\rangle = \sum_n p_n\, n$

  Verallgemeinert $\langle g(Z)\rangle = \int \mathrm{d}z\, p(z)\, g(z)$

- **Varianz**: "Streuung" der Verteilung um $\langle Z\rangle$:

  $$\sigma_Z^2 = \langle (Z-\langle Z\rangle)^2\rangle = \langle Z^2\rangle - \langle Z\rangle^2$$

Beispiele

⋆ Binomialverteilung:	$\langle Z\rangle = Nq$,	$\sigma_Z^2 = Nq(1-q)$
⋆ Poissonverteilung:	$\langle Z\rangle = \lambda$,	$\sigma_Z^2 = \lambda$
⋆ Exponentialverteilung:	$\langle Z\rangle = \beta$,	$\sigma_Z^2 = \beta^2$
⋆ Lorentz-Verteilung:	$\langle Z\rangle = \mu$,	$\sigma_Z^2 = \infty$
⋆ Gaußverteilung:	$\langle Z\rangle = \mu$,	$\sigma_Z^2 = \sigma^2$

Der Erwartungswert und die Varianz charakterisieren den dominierenden Wert von z in der Verteilung und die Streuung darum. Bei manchen Verteilungen ist $\langle Z \rangle$ jedoch nicht so aussagekräftig, da extreme Werte zu sehr ins Gewicht fallen. Daher gibt es noch andere Kenngrößen

- **Modus**: Maximum der Verteilung
- **Median** oder **Zentralwert**:
  "Zahl z_m, so dass $Z < z_m$ gleichwahrscheinlich ist wie $Z > z_m$"
  $\rightsquigarrow$ Kontinuierliche Verteilung: $\int_{-\infty}^{z_m} \mathrm{d}z\, p(z) \stackrel{!}{=} \int_{z_m}^{\infty} \mathrm{d}z\, p(z)$
  Allgemeiner: $P(Z < z_m) \le 1/2 \le P(Z \le z_m)$
  (kann auch auf diskrete Verteilungen angewendet werden)
  (NB: Nicht immer eindeutig, kann auch für ein Intervall stehen).
- Verallgemeinerung: **Quantile**
  q-Quantil: Zahl z_q mit $\int_{-\infty}^{z_q} \mathrm{d}z\, p(z) = q$ (kontinuierliche Verteilung)
  bzw. allgemeiner: $P(Z < z_m) \le q \le P(Z \le z_m)$
  Der Median ist das 1/2-Quantil. Andere öfter gebrauchte Quantile sind die Quartile, d.h. das 1/4-Quantil und das 3/4-Quantil, sowie die Perzentile, bei denen p als Prozentzahl angegeben wird.

5.2.5.4 Der zentrale Grenzwertsatz

Grund, warum die Gaußverteilung die mit Abstand wichtigste Verteilung ist: Der Mittelwert von n gleichverteilten Zufallszahlen ist unter bestimmten Voraussetzungen im Grenzwert $n \to \infty$ gaussverteilt!

Satz (zentraler Grenzwertsatz nach Lindenberg-Lévy)

Seien n unabhängige Zufallsvariablen alle nach derselben Verteilungsdichte $p(z)$ verteilt: $p(z)$ besitze einen Erwartungswert $\langle Z \rangle$ und eine Varianz $\sigma_Z^2 < \infty$. Dann strebt die Verteilungsdichte $\tilde{p}(y)$ der Zufallsvariablen

$$\boxed{Y_n = \Big(\frac{1}{n}\sum_{i=1}^{n} Z_i - \langle Z \rangle\Big) \Big/ \frac{\sigma_Z}{\sqrt{n}}}$$

im Grenzwert $n \to \infty$ gegen eine **Normalverteilung**

$$\boxed{\lim_{n\to\infty} \tilde{p}(y) = \mathcal{N}(y) := \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-y^2/2}}$$

(Für praktische Zwecke: Normalverteilung gut ab etwa $n \sim 30$)

Beweis: Führt hier zu weit, weggelassen.

Bemerkung: Der zentrale Grenzwertsatz gilt auch für nicht identisch verteilte Zufallsvariablen. Man muss dann eine Zusatzbedingung an die dritten Momente stellen: $\lim_{n\to\infty} \frac{\left(\langle Z_1^3 \rangle + \cdots \langle Z_n^3 \rangle\right)}{\left(\sigma_1^2 + \ldots + \sigma_n^2\right)^{3/2}} = 0$

5.3 Stichproben und Beobachtungsfehler

Wir kommen nun zur ursprünglichen Fragestellung dieses Kapitels zurück: Der Auswertung von Datensätzen.

Angenommen, eine Größe x wurde N Mal unabhängig und auf identische Weise gemessen → Messreihe (Datensatz) $(x_1, ...x_N)$ mit N unabhängigen Punkten.

Wir fassen die x_i als Zufallsvariablen auf. Sie wurden aus aus einer identischen Verteilung $p(x_i)$ gezogen wurden, die jedoch nicht bekannt ist. Wir haben nur die Stichprobe $(x_1, ..., x_N)$.

Frage: Wie wertet man diese Stichprobe optimal aus?
Konkret: Wie schätzt man daraus den Erwartungswert $\langle X \rangle$ und die Varianz σ_X^2 der Verteilung ab, und wie bestimmt man den Fehler, den man dabei macht?

5.3.1 Charakterisierung von Stichproben

Für Stichproben definiert man ähnliche charakteristische Größen wie für Verteilungen (5.2.5.4):

- **Median** oder **Zentralwert**:

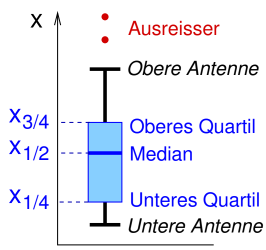
  "Zahl x_m, so dass es gleich viele Datenpunkte oberhalb und unterhalb von x_m gibt"

  bzw. genauer: Zahl x_m mit $\sum_{x_i < x_m} \le N/2 \le \sum_{x_i \le x_m}$.

- **p-Quantile**: Analog: $\sum_{x_i < x_p} \le pN \le \sum_{x_i \le x_p}$

Beliebte Auftragung bei kleinen Stichproben (z.B. in der Medizin): **Boxplot**

Schnelle Art, Datenverteilungen zu visualisieren, z.B., um Datensätze zu vergleichen
- 50 % der Datenpunkte liegen in der Box
- Mittelstrich: Median
- Antennen: charakterisieren restliche Daten
  Definition nicht einheitlich,
  z.B. 2.5%-Quantil und 97.5%-Quantil
- Punkte: Ausreißer



- **Stichprobenmittelwert**: $\bar{x} = \frac{1}{N} \sum_{i=1}^{N} x_i$

- **Stichprobenvarianz**: $\tilde{\sigma}_X^2 = \frac{1}{N-1} \sum_{i=1}^{N} (x_i - \bar{x})^2$

  (Erklärung für den Faktor $\frac{1}{N-1}$: Nächster Abschnitt 5.3.2)

5.3.2 Schätzung von Erwartungswert und Varianz

Da die Stichprobe ein endlicher Satz von Zufallsvariablen ist, sind die charakteristischen Größen (Mittelwert etc.) selbst wieder Zufallsgrößen.

Die nächste Frage lautet also: Wie kann man daraus die Parameter der Grundverteilung $p(x)$ abschätzen, insbesondere den Erwartungswert $\mu_X := \langle X \rangle$ und die Varianz σ_X? (die zwei Parameter, die eine Gaußverteilung charakterisieren)

$\rightsquigarrow$ Gesucht: **Schätzer** $\tilde{\mu}_X$ und $\tilde{\sigma}_X^2$ für μ und σ_X^2
Bedingung: Soll **erwartungstreu** sein, d.h. $\langle \tilde{\mu}_X \rangle = \mu_X$, $\langle \tilde{\sigma}_X^2 \rangle = \sigma_X^2$

⋆ Erwartungswert

Ansatz: $\tilde{\mu}_X = A_\mu \frac{1}{N} \sum_{i=1}^N x_i$ (mit A_μ: anpassbarer Parameter)

Rechnung: $\langle \tilde{\mu}_X \rangle = A_\mu \frac{1}{N} \sum_{i=1}^N \underbrace{\langle x_i \rangle}_{\mu_X} = A_\mu \frac{1}{N} \underbrace{\sum_{i=1}^N}_{N} \mu_X = A_\mu \, \mu_X \overset{!}{=} \mu_X \Rightarrow A_\mu = 1$

$\Rightarrow$ Erwartungstreuer Schätzer für μ_X: $\boxed{\tilde{\mu}_X = \frac{1}{N} \sum_{i=1}^N x_i = \bar{x}}$

⋆ Varianz

Ansatz: $\tilde{\sigma}_X^2 = A_\sigma \sum_{i=1}^N (x_i - \bar{x})^2$

Rechnung: $\langle \tilde{\sigma}_X^2 \rangle = A_\sigma \langle \sum_{i=1}^N (x_i - \bar{x})^2 \rangle$

$\quad \Big| \quad \sum_{i=1}^N (x_i - \bar{x})^2 = \sum_i x_i^2 - 2\bar{x} \underbrace{\sum_i x_i}_{N\bar{x}} + \sum_i \bar{x}^2 = \sum_i x_i^2 - N\bar{x}^2$

$= A_\sigma N \left(\langle x^2 \rangle - \langle \bar{x} \rangle^2 \right)$

$\quad \Big| \quad \langle x^2 \rangle = \mu_X^2 + \sigma_X^2, \ \langle \bar{x}^2 \rangle = \mu_X^2 + \sigma_X^2/N$ (Nebenrechnung unten)

$= A_\sigma N (\mu_X^2 + \sigma_X^2 - \mu_X^2 - \sigma_x^2/N) = A_\sigma N \sigma_X^2 (1 - \frac{1}{N} \overset{!}{=} \sigma_X^2$

$\Rightarrow A_\sigma = \frac{1}{N(1-1/N)} = \frac{1}{N-1}$

$\rightsquigarrow$ Erwartungstreuer Schätzer für σ_X^2: $\boxed{\tilde{\sigma}_X^2 = \frac{1}{(N-1)} \sum_{i=1}^N (x_i - \bar{x})^2}$

Interpretation: Warum der Faktor $\frac{1}{N-1}$?

- Die Varianz der "wahren" Verteilung $p(x)$ ergibt sich als Summe der mittleren Streuung der Messdaten um den Stichprobenmittelwert, $\Delta^2 := \langle \frac{1}{N} \sum_i (x_i - \bar{x})^2 \rangle$, und der Streuung des Stichprobenmittelwertes: $\sigma_X^2 = \Delta^2 + \sigma_{\bar{x}}^2$ mit $\Delta^2 = \frac{N-1}{N} \sigma_X^2$ und $\sigma_{\bar{x}}^2 = \sigma_X^2/N$.
- Die mittlere Streung der Messdaten um den Stichprobenmittelwert, muss mit der Zahl der Freiheitsgrade $(N-1)$ normiert werden.

Nebenrechnung: Zeige $\langle \bar{x}^2 \rangle = \mu_X^2 + \sigma_X^2/N$ bzw. $\boxed{\sigma_{\bar{x}}^2 = \langle \bar{x}^2 \rangle - \langle \bar{x} \rangle^2 = \sigma_X^2/N}$

$\langle \bar{x}^2 \rangle = \langle (\frac{1}{N} \sum_i x_i)(\frac{1}{N} \sum_j x_j) \rangle = \frac{1}{N^2} \sum_{ij} \langle x_i x_j \rangle \quad \Big| \quad$ mit $\langle x_i x_j \rangle = \begin{cases} \langle x^2 \rangle & : i = j \\ \langle x \rangle^2 & : i \neq j \text{ (unabhängig)} \end{cases}$

$= \frac{1}{N^2} \Big[\underbrace{\sum_{i \neq j}}_{N(N-1)} \underbrace{\langle x \rangle^2}_{\mu_X^2} + \underbrace{\sum_i}_{N} \underbrace{\langle x^2 \rangle}_{\mu_X^2 + \sigma_X^2} \Big] = \frac{1}{N^2} \big[N(N-1) \, \mu_X^2 + N(\sigma_X^2 + \mu_X^2) \big] = \mu_X^2 + \sigma_X^2/N$

5.3.3 Fehlerabschätzung und Verteilung des Mittelwertes

Wir betrachten wieder eine Stichprobe wie in 5.3.2.

Aus 5.2.5.3 wissen wir: Wenn die Stichprobe hinreichend groß ist, sollte der Mittelwert $\bar{x}$ gaußverteilt sein. Die Breite der Gaußverteilung gibt den Fehler $\Delta\tilde{\mu}_X$ unserer Schätzung $\tilde{\mu}_X = \bar{x}$ für μ_X an.

Aus dem zentralen Grenzwertsatz bzw. den Rechnungen in Abschnitt 5.3.2 wissen wir: $\sigma_{\bar{x}} = \sigma_X/\sqrt{N}$. Die Varianz σ_X schätzen wir mit $\tilde{\sigma}_X$ ab.

$\rightsquigarrow$ Abschätzung mit Fehler für den Erwartungswert der Verteilung $f_X(x)$:

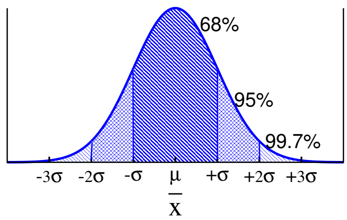
$$\boxed{\langle X \rangle = \bar{x} \pm \Delta\bar{x}} \quad \text{mit} \quad \boxed{\Delta\bar{x} = \sqrt{\frac{1}{N(N-1)}\sum_{i=1}^{N}(x_i - \bar{x})^2}}$$

Unter der Annahme, dass $\bar{x}$ tatsächlich gaußverteilt ist, können wir noch weitere statistische Aussagen machen. Zum Beispiel können wir **Konfidenzintervalle** bestimmen, d.h. Intervalle, innerhalb derer sich der "wahre" Erwartungswert μ_X mit einer bestimmten Wahrscheinlichkeit befindet (Vertrauensbereiche):

Die Wahrscheinlichkeit, dass der wahre Wert μ_X im Bereich $\bar{x} \pm \alpha\,\Delta\bar{x}$ liegt, beträgt $P_\alpha = \frac{1}{\sqrt{2\pi}}\int_{-\alpha}^{\alpha} \mathrm{d}y\, \mathrm{e}^{-y^2/2}$

(bzw. nach der frequentistischer Wahrscheinlichkeitsauffassung:
Bei unendlich häufiger Wiederholung der N Experimente liegt der Anteil der Mittelwerte $\bar{x}$, die dann im Intervall $[\bar{x} - \alpha\,\Delta\bar{x} : \bar{x} + \alpha\,\Delta\bar{x}]$ landen, bei $P_\alpha = \frac{1}{\sqrt{2\pi}}\int_{-\alpha}^{\alpha} \mathrm{d}y\, \mathrm{e}^{-y^2/2}$.)

Konkret: Der wahre Wert μ_X liegt mit der Wahrscheinlichkeit ...
- 68% im Bereich $\bar{x} \pm \Delta\bar{x}$
- 95% im Bereich $\bar{x} \pm 2\Delta\bar{x}$
- 99.7% im Bereich $\bar{x} \pm 3\Delta\bar{x}$



NB: Die Zahl 99.7% hört sich nach viel an. Aber sie besagt auch, dass immerhin in jedem 300. Datensatz der Mittelwert außerhalb des "3 σ"-Intervalls $\bar{x} \pm 3\Delta\bar{x}$ liegt! Das heißt, wenn man sehr intensiv nach einer solchen "signifikanten" Abweichung sucht, sind die Chancen, eine zu finden, doch recht hoch.

5.3.4 Fehlerfortpflanzung

Nächste Frage: Angenommen, aus Experimenten kennen wir eine oder mehrere Messgrößen x_α mit Genauigkeit (Fehler) Δx_α. De facto interessieren wir uns aber für eine abgeleitete Größe $y(x_1, ..., x_n)$

$\to$ Welchen Fehler Δy hat dann y?

Beispiel: Fallexperiment liefert fehlerbehaftete Messrößen $t = \bar{t} \pm \Delta t,\ h = \bar{h} \pm \Delta h$
Bestimme $g = 2h/t^2 \to$ Welchen Fehler Δg hat g?

Annahme im Folgenden: Δx_α sei so klein, dass man eine Taylorentwicklung nach der ersten Ordnung abbrechen darf:

$\langle y\rangle \approx y(\bar{x}_1, ...\bar{x}_n) + \mathcal{O}(\Delta x_\alpha^2)$
$y - \langle y\rangle \approx \sum_\alpha \frac{\partial y}{\partial x_\alpha}\Big|_{x_\alpha=\bar{x}_\alpha}(x_\alpha - \langle x_\alpha\rangle) + \mathcal{O}(\Delta x_\alpha^2)$

Fehlerabschätzungen:

(i) Konservative Abschätzung: **Maximalfehler** $\Delta y \lesssim \sum_\alpha \left|\frac{\partial y}{\partial x_\alpha}\right|_{x_\alpha=\bar{x}_\alpha} \Delta x_\alpha$

$$\Rightarrow \quad \boxed{\Delta y \lesssim \sum_\alpha \left|\left(\frac{\partial y}{\partial x_\alpha}\right)_{x_\alpha=\bar{x}_\alpha}\right| \sigma_{x_\alpha}}$$

(ii) **Gaußsche Fehlerfortpflanzung**:

Zusätzliche Annahme: x_α unkorreliert $\rightsquigarrow$ reduziert den Fehler!

$\Rightarrow \ \Delta y^2 = \langle (y-\langle y\rangle)^2\rangle \approx \left\langle \left(\sum_\alpha \frac{\partial y}{\partial x_\alpha}(x_\alpha - \langle x_\alpha\rangle)^2\right)\right\rangle = \sum_\alpha \left(\frac{\partial y}{\partial x_\alpha}\right)^2_{x_\alpha=\bar{x}_\alpha} \left\langle (x_\alpha - \langle x_\alpha\rangle)^2\right\rangle$
$= \sum_\alpha \left(\frac{\partial y}{\partial x_\alpha}\right)^2_{x_\alpha=\bar{x}_\alpha} \sigma_{x_\alpha}^2$

$$\Rightarrow \quad \boxed{\Delta y = \sqrt{\sum_\alpha \left(\frac{\partial y}{\partial x_\alpha}\right)^2_{x_\alpha=\bar{x}_\alpha} \sigma_{x_\alpha}^2}}$$

Konkret:

- Additive Fehler: $y = \sum_\alpha x_\alpha$
  Maximalfehler: $\Delta y = \sum_\alpha \Delta x_\alpha$
  Gaußscher Fehler: $\Delta y = \sqrt{\sum_\alpha (\Delta x_\alpha)^2}$
  $\rightsquigarrow$ Absolute Fehler addieren sich auf!
- Multiplikative Fehler: $y = \prod_\alpha x_\alpha$
  Maximalfehler: $\Delta y = \sum_\alpha \Delta x_\alpha \prod_{\beta\neq\alpha} x_\beta \ \Rightarrow \ \frac{\Delta y}{y} = \sum_\alpha \frac{\Delta x_\alpha}{x_\alpha}$
  Gaußscher Fehler: $\Delta y/y = \sqrt{\sum_\alpha (\Delta x_\alpha/x_\alpha)^2}$
  $\rightsquigarrow$ Relative Fehler addieren sich auf!

Beispiele:

- Fallexperiment: Bestimme Fehler von $g = 2h/t^2$
  aus Kenntnis der Fehler von $t \pm \Delta t$ und $h \pm \Delta h$
  Maximalfehler: $\Delta g = \left|\frac{\partial g}{\partial t}\right| \Delta t + \left|\frac{\partial g}{\partial h}\right| \Delta h = 4\frac{h}{t^3}\Delta t + 2\frac{1}{t^2}\Delta h$
  Gaußscher Fehler: $\Delta g^2 = \left(\frac{\partial g}{\partial t}\right)^2 \Delta t^2 + \left(\frac{\partial g}{\partial h}\right)^2 \Delta h^2 = 16\frac{h^2}{t^6}\Delta t^2 + 4\frac{1}{t^4}\Delta h^2$
  $\Rightarrow \ (\Delta g/g)^2 = 4(\Delta t/t)^2 + (\Delta h/h)^2$
- Kleine Differenz von großen Messgrößen
  $y = x_1 - x_2 \ll x_1 \approx x_2$
  $\Rightarrow \ y \ll \max(x_1, x_2);\ \Delta y > \max(\Delta x_1, \Delta x_2)$
  $\Rightarrow$ Relativer Fehler $\Delta y/y$ wird sehr groß !

5.4 Parameter-Fitting

Zusatzstoff zum Weiterlesen, wird im Kurs nicht besprochen.

Häufige Problemstellung:

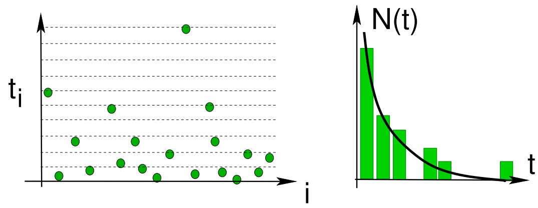
Gegeben Messreihe (Datensatz) $D = \{(y_i; \Delta y_i, \underline{x}_i\}$
mit y_i: Gemessene Größen (evtl. bereits mit Fehler Δy_i
$\underline{x}_i$: Vektor von zugehörigen Kontrollparameter
evtl. systematisch variiert

und Hypothese H dazu, wie die y_i für gegebene Kontrollparameter $\underline{x}_i$ verteilt sein sollten. H hängt von Modellparametern c_α ab

→ Aufgabe: Abschätzen der Modellparameter $\{c_\alpha\}$ aus den Daten D
"Fitten der Parameter $\{c_\alpha\}$ an die Daten D!"

Beispiele

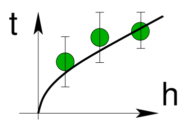
(i) Radioaktiver Zerfall; Messreihe: Satz von Zerfallszeiten $\{t_i\}$
- Daten: $\{t_i\}$ (keine Fehler, keine Kontrollparameter)
- Hypothese: Wahrscheinlichkeitsdichteverteilung $p(t) = k\mathrm{e}^{-kt}$
- Modellparameter: Zerfallsrate k



⇝ Beispiel, in dem eine Verteilung gefittet werden soll.

(ii) Fallexperiment; Messreihe: $t_i \pm \Delta t_i$ bei Fallhöhe h_i
- Daten: $\{(t_i; \Delta t_i, h_i)\}$ (Fehler Δt_i, Kontrollparameter h_i)
- Hypothese: Für gegebenes h_i ist t_i gaußverteilt mit Varianz Δt_i und Erwartungswert $\langle t_i \rangle = \sqrt{2h_i/g}$.
- Modellparameter: g

⇝ Beispiel, in dem eine Gleichung zu einem funktionalen Zusammenhang zwischen Größen gefittet werden soll.



Dieses Kapitel: Einführung in einige Methoden des Parameter-Fittings

5.4.1 Bayes-Schätzer und Maximum Likelihood (ML)

⋆ Ausgangspunkt: Fasse Fitparameter $\{c_\alpha\}$ nach Bayes als Zufallsvariablen auf, für die eine Prognose gemacht werden soll.

Wende Satz von Bayes an: $\boxed{P(\{c_\alpha\}|D) = \frac{P(D|\{c_\alpha\})\, P(\{c_\alpha\})}{P(D)}}$

mit: $P(D|\{c_\alpha\})$: **Likelihood**
Wahrscheinlichkeit, mit der bei gegebenen Parameter $\{c_\alpha\}$ unter der Hypothese H der Datensatz D eintritt
(als bekannt vorausgesetzt)

$P(\{c_\alpha\})$: **Prior**
A-Priori Wahrscheinlichkeit für die Verteilung der Parameter (Vorwissen)

$P(D)$: **Marginal** oder **Base rate**
De facto einfach Normierung
"Wahrscheinlichkeit von D für alle möglichen Parameter $\{c_\alpha\}$"

$P(\{c_\alpha\}|D)$: **Posterior**
Wahrscheinlichkeit für Parameter $\{c_\alpha\}$, gegeben dass der Datensatz D existiert und unter Berücksichtigung des Priors

Idee: Wähle als Schätzer für die Parameter $\{c_\alpha\}$ diejenigen, die laut Posterior am wahrscheinlichsten sind!
→ Maximiere $P(\{c_\alpha\}|D)$ bzgl. c_α für den gegebenen Datensatz D.
(NB: Es reicht offensichtlich, $P(D|\{c_\alpha\})\, P(\{c_\alpha\})$ zu maximieren. $P(D)$ muss dafür nicht bekannt sein!)

⋆ Speziell: Maximum Likelihood

Falls es keinen Prior (Vorwissen) gibt, dann nimmt man für $P(\{c_\alpha\})$ einfach eine Gleichverteilung an (innerhalb sinnvoller Grenzen)
→ $P(\{c_\alpha\}|D) \propto P(D|\{c_\alpha\})$

⇒ Maximiere $P(D|\{c_\alpha\})$ bzgl. der c_α,
d.h. optimiere die Modellparameter $\{c_\alpha\}$ so, dass der Datensatz D in der zugehörigen Hypothese der wahrscheinlichste Datensatz ist.

⇝ **Maximum Likelihood (ML)-Methode**

⋆ Beispiele

– Radioaktiver Zerfall:
Datensatz $\{t_i\}$ $(i = 1, .., N)$, Hypothese $p(t) = k\,\mathrm{e}^{-kt}$
$\Rightarrow\ P(\{t_i\}|k) \propto \prod_i \left(k\mathrm{e}^{-kt_i}\right) = k^N \mathrm{e}^{-k\sum_i t_i}$
Da $\ln P$ monotone in P ist, kann man statt P auch $\ln P$ maximieren.
$\Rightarrow\ \frac{\partial}{\partial k} \ln P = \frac{N}{k} - \sum_i t_i \stackrel{!}{=} 0 \ \Rightarrow\ k = N / \sum_i t_i = 1/\bar{t}$
⇝ Nach der ML-Methode ist $k = 1/\bar{t}$
(Ohne Zweifel ein sinnvolles Ergebnis!)

– Gaußsche Verteilung

Datensatz $\{y_i\}$ $(i = 1, .., N)$, Hypothese $p(t) = \frac{1}{\sqrt{2\pi\sigma^2}} \mathrm{e}^{-(y-\mu)^2/2\sigma^2}$

$\Rightarrow\ P(\{y_i\}|\{\mu,\sigma\}) \propto \prod_i \Big(\frac{1}{\sqrt{2\pi\sigma^2}} \exp(-\sum_i \frac{(y_i-\mu)^2}{2\sigma^2}) \Big)$

Maximiere $\ln P = -\frac{N}{\sigma} - \sum_i \frac{(y_i-\mu)^2}{2\sigma^2} + \text{const.}$ bzgl. μ und σ

$\Rightarrow\ \frac{\partial}{\partial\mu} \ln P = \sum_i \frac{(y_i-\mu)}{2\sigma^2} \overset{!}{=} 0 \ \Rightarrow\ \mu = \frac{1}{N}\sum_i y_i = \bar{y}$

$\frac{\partial}{\partial\sigma} \ln P = -\frac{N}{\sigma} + \sum_i \frac{(y_i-\mu)^2}{\sigma^3} \overset{!}{=} 0 \ \Rightarrow\ \sigma^2 = \frac{1}{N}\sum_i (y_i - \mu)^2$

Fasse zusammen $\rightsquigarrow$ ML-Schätzung $\mu = \bar{y}$, $\sigma^2 = \frac{1}{N}\sum_i (y_i - \bar{y})^2$

Vergleiche mit den Schätzern aus 5.3.3

→ Die Schätzer für μ stimmen überein.

→ Die Schätzer für σ^2 unterscheiden sich um den Faktor $\frac{N-1}{N}$.

Der ML-Schätzer für σ^2 ist nicht optimal, da er nicht erwartungstreu ist. Allerdings geben im Grenzfall $N \to \infty$ (unendlich viele Datenpunkte) beide wieder dasselbe Ergebnis.

5.4.2 Methode der kleinsten Fehlerquadrate

Nun kommen wir konkret zum Fitten von Gleichungen

Ziel: Datensatz $\{y_i; \Delta y_i; \underline{x}_i\}$ mit Kontrollparametern $\underline{x}_i$, und Messdaten $y_i \pm \Delta y_i$ an eine Funktion $y = \phi(\underline{x}; \{c_\alpha\})$ anfitten (z.B. Gerade, Polynom, auch kompliziertere Funktionen)

Verfahren: Ausgangspunkt ist die ML-Methode mit zusätzlicher Annahme: Für feste $\underline{x}_i$ ist y_i gaußverteilt mit Erwartungswert $\langle y_i \rangle = \phi(\underline{x}_i, \{c_\alpha\})$ und Varianz $\sigma_i = \Delta y_i$.

(Motivation: Die Daten y_i sind selber Mittelwerte, so dass der zentrale Grenzwertsatz angewendet werden kann.)

$\rightsquigarrow\ P(D|\{c_\alpha\}) \ \propto\ \exp\Big(-\sum_i \frac{1}{2\sigma_i^2}(y_i - \phi(\underline{x}_i, \{c_\alpha\})^2 \Big)$

Maximiere $\ln P(D|\{c_\alpha\})$ bzgl. c_α

$\Rightarrow$ Minimiere Funktion $\boxed{\chi^2 = \sum_i \frac{1}{\sigma_i^2} (y_i - \phi(\underline{x}_i, \{c_\alpha\}))^2}$

$\rightsquigarrow$ Methode der kleinsten Fehlerquadrate

Engl. **"Least squares fitting"**

⋆ Bemerkungen

- Die Varianzen σ_i werden nicht mitgefittet. Dort setzt man die Schätzer Δy_i ein. Falls keine Fehler bekannt sind (z.B., weil nur einzelne Datenpunkte vorliegen), setzt man einfach $\sigma_i \equiv 1$.
- Der Vektor von Kontrollparametern $\underline{x}_i$ kann auch mehrdimensional sein - d.h., man kann ohne weiteres mehrere Kurven gleichzeitig fitten.

⋆ Konkretes Vorgehen: Was ist zu tun?

- Definiere $\chi^2 = \sum_i \frac{1}{\sigma_i^2} (y_i - \phi(\underline{x}_i, \{c_\alpha\}))^2$
- Minimiere χ^2 bzgl. $\{c_\alpha\}$

  → Satz von Gleichungen $\frac{\partial}{\partial c_\alpha}\chi^2 \stackrel{!}{=} 0$

  → Schätzer $\tilde{c}_\alpha$ für den optimalen Wert von c_α
  (Explizite oder implizite Gleichungen)
- Bestimme Schätzer Δc_α für den Fehler von $\tilde{c}_\alpha$ aus den Fehlern Δy_i z.B. über Gaußsche Fehlerfortpflanzung

  (NB: Erfahrungsgemäß oft zu klein.
  Alternativ konservativere Abschätzung:
  Variiere y_i innerhalb der Fehlergrenzen und fitte neu.)
- Überprüfe Fit: Wichtig, nicht vergessen !!

  1) Wenn es kein Routine-Fit ist: Unbedingt Ergebnis anschauen!
  → Geht die gefittete Kurve auch wirklich durch die Daten?

  2) "Residuen" $(y_i - \phi(\underline{x}_i, \{\tilde{c}_\alpha\})/\sigma_i$ auftragen.
  → Sind sie wirklich normalverteilt ?
  (Normalverteilung: Gaußverteilung um 0 mit Breite 1,
  $\mathcal{N}(y) = \frac{1}{\sqrt{2\pi}} \exp(-y^2/2)$ siehe 5.2.5.3)

5.4.3 Lineare Regression

Wichtiger Spezialfall der Methode der kleinsten Fehlerquadrate:

Fitfunktion $\phi(\underline{x}; \{c_\alpha\})$ ist lineare Funktion der Parameter c_α. Dann kann das Fitproblem mit Methoden der linearen Algebra gelöst werden.

Gegeben sei ein Datensatz $\{y_i; \Delta y_i; \underline{x}_i\}$ mit n Datenpunkten und eine lineare Fitfunktion mit k Parametern c_α, d.h. $\phi(\underline{x}; \{c_\alpha\}) = \sum_{\alpha=1}^k c_\alpha\, a_\alpha(\underline{x})$

⋆ Bestimmung der Fitparameter $\tilde{c}_\alpha$

– Aufgabe: Minimiere $\chi^2 = \sum_{i=1}^n \frac{1}{\sigma_i^2} \big(y_i - \sum_\alpha c_\alpha a_\alpha(x_i)^2\big)^2$

– Lösung in Matrixnotation

- Definiere Vektoren $y = (y_1, ..., y_n)$, $c = (c_1, ..., c_k)$
  und Matrizen A, Σ mit $\boxed{A_{i\alpha} := a_\alpha(\underline{x}_i)}$ und $\boxed{\Sigma_{ij} = \delta_{ij}\, \sigma_i^2}$
  → $\chi^2 = (y - Ac)^T \Sigma^{-1} (y - Ac)$
- Minimiere $\chi^2 \;\Rightarrow\; \frac{\partial \chi^2}{\partial c_\alpha} \stackrel{!}{=} 0$

  Rechnung: $\frac{\partial \chi^2}{\partial c_\alpha} = -2\sum_i \frac{1}{\sigma_i^2}\big(y_i - \sum_\beta c_\beta a_\beta(\underline{x}_i)\big)\, a_\alpha(\underline{x}_i) = -2\big[A^T\Sigma^{-1}(y - Ac)\big]_\alpha$

  $\Rightarrow \frac{\partial \chi^2}{\partial c_\alpha} \stackrel{!}{=} 0$ führt zu Gleichungssystem $A^T\Sigma^{-1}Ac = A^T\Sigma^{-1}y$

⇒ Lösung: $\boxed{\tilde{c} = (A^T\Sigma^{-1}A)^{-1}A^T\Sigma^{-1}y}$

(Voraussetzung: $(A^T\Sigma^{-1}A)$ invertierbar.
Anderenfalls ist Lösung nicht eindeutig!)

⋆ Fehlerabschätzung

Bestimme Fehler von $\tilde{c}_\alpha$ mittels Gaußscher Fehlerfortpflanzung

Rechnung: Definiere $B = (A^T\Sigma^{-1}A)^{-1}\,A^T\Sigma^{-1} \quad\Rightarrow\quad C = By$

$\Rightarrow\ \Delta c_\alpha^2 = \sum_i \underbrace{(\partial c_\alpha/\partial y_i)}_{B_{\alpha i}}{}^2\ \underbrace{\Delta y_i^2}_{\Sigma_{ii}} = \sum_{ij} B_{\alpha i}\ \underbrace{\Sigma_{ij}}_{\Sigma_{ii}\delta_{ij}}\ B_{\alpha j}$

$\Rightarrow\ \Delta c_\alpha^2$ sind die Diagonalelemente von $\Delta c\,\Delta c^T = B\,\Sigma\,B^T$

Setze $B = (A^T\Sigma^{-1}A)^{-1}\,A^T\Sigma^{-1}$ ein und verwende:

$(A^T\Sigma^{-1}A)$ und Σ sind symmetrisch, ebenso damit auch ihre Inverse.

$$\Rightarrow\ \Delta c\,\Delta c^T = \big[(A^T\Sigma^{-1}A)^{-1}\,A^T\Sigma^{-1}\big]\Sigma\big[(A^T\Sigma^{-1}A)^{-1}\,A^T\Sigma^{-1}\big]^T$$
$$= \big[(A^T\Sigma^{-1}A)^{-1}\,A^T\Sigma^{-1}\big]\Sigma\big[\Sigma^{-1}A(A^T\Sigma^{-1}A)^{-1}\big] = (A^T\Sigma^{-1}A)^{-1}$$

$$\Rightarrow \quad \boxed{(\Delta c_\alpha)^2 = \Big[(A^T\Sigma^{-1}A)^{-1}\Big]_{\alpha\alpha}}$$

⋆ Anmerkungen

- Auch wenn die y_i unabhängig sind, sind die c_α i. A. nicht mehr unabhängig. Die Korrelation der Parameter $\tilde{c}_\alpha$ und $\tilde{c}_\beta$ wird durch die Nebendiagonalen $\big[(A^T\Sigma^{-1}A)^{-1}\big]_{\alpha\beta}$ angegeben.
- Die obigen Formeln können auch dann noch verwendet werden, wenn die y_i nicht unabhängig sind. Man ersetzt dann $\Sigma_{ij} = \delta_{ij}\,\sigma_i^2$ durch die Korrelationsmatrix $\Sigma_{ij} = \langle y_i y_j\rangle - \langle y_i\rangle\,\langle y_j\rangle$

⋆ Anwendungsbeispiel: **Regressionsgerade** (vgl. Übungsaufgabe)

Aufgabe: Fit eines Datensatzes an eine Gerade $\boxed{\phi(x) = c_1 + c_2\,x}$

⇝ Entspricht linearer Regression mit Fitfunktion $\phi(x) = \sum_{\alpha=1}^2 c_\alpha a_\alpha(x)$ mit $a_1(x) = 1$, $a_2(x) = x \Rightarrow\ A_{i1} = 1$, $A_{i2} = x_i$

bzw. $A = \begin{pmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix}$, $\Sigma = \begin{pmatrix} \sigma_1^2 & & 0 \\ & \ddots & \\ 0 & & \sigma_n^2 \end{pmatrix}$, $y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$

$\Rightarrow\ (A^T\Sigma^{-1}A) = \ldots = \begin{pmatrix} S_e & S_x \\ S_x & S_{xx} \end{pmatrix}$ und $A^T\Sigma^{-1}y = \ldots = \begin{pmatrix} S_y \\ S_{xy} \end{pmatrix}$

mit $S_e = \sum_i 1/\sigma_i^2$,
$S_x = \sum_i x_i/\sigma_i^2$,
$S_y = \sum_i y_i/\sigma_i^2$,
$S_{xx} = \sum_i x_i^2/\sigma_i^2$,
$S_{xy} = \sum_i x_i y_i/\sigma_i^2$.

$$\Rightarrow \quad \boxed{\begin{aligned} \tilde{c}_1 &= \frac{1}{S_e S_{xx} - S_x^2}\Big(S_{xx}S_y - S_x S_{xy} \pm S_{xx}\Big) \\ \tilde{c}_2 &= \frac{1}{S_e S_{xx} - S_x^2}\Big(S_{xy}S_e - S_x S_y \pm S_e\Big) \end{aligned}}$$