

Höhere Mathematik I

Robert Haller-Dintelmann

TU Darmstadt, WS 2020/2021
Version vom 12. Dezember 2020

Danksagung

Dieses Skript ist bis auf wenige Änderungen und Ergänzungen übernommen aus dem Skript *Mathematik für Chemiker*, Universität Paderborn, WS 2004/2005, von Dieter Bothe, Hermann Hembd und Norbert Köckler. Es wurde überarbeitet von Sören Kraußhar, Achim Blumensath, Christian Mönch und Steffen Roch. Ich danke ihnen allen vielmals für das freundliche Zurverfügungstellen ihres Materials.

Inhaltsverzeichnis

Bezeichnungen	vii
1 Grundbegriffe	1
1.1 Aussagenlogik	1
1.2 Mengen und Zahlen	3
1.3 Elementare Beweismethoden	5
1.3.1 Direkter Beweis	5
1.3.2 Indirekter Beweis	6
1.3.3 Vollständige Induktion	6
1.4 Binomialkoeffizienten	8
1.4.1 Fakultät und Binomialkoeffizient	8
1.4.2 Binomische Formeln und Pascalsches Dreieck	9
1.5 Ungleichungen und Beträge	9
2 Rechentechniken	13
2.1 Potenzen	13
2.1.1 Exponentialschreibweise	13
2.1.2 Multiplikation und Division	13
2.1.3 Potenzieren und Wurzelziehen	14
2.1.4 Addition und Subtraktion	14
2.2 Logarithmen	15
2.2.1 Definition und erste Eigenschaften	15
2.2.2 Graphische Darstellung des Logarithmus	16
2.2.3 Der natürliche Logarithmus	17
2.3 Algebraische Gleichungen	18
3 Elementare Funktionen	22
3.1 Der mathematische Funktionsbegriff	22
3.2 Trigonometrische Funktionen	24
3.3 Eigenschaften von Funktionen	26
3.4 Komplexe Zahlen	32

4	Folgen und Grenzwerte	36
4.1	Folgen	36
4.2	Konvergenz von Folgen	37
4.3	Grenzwertsätze	38
4.4	Reihen	41
4.5	Konvergenzkriterien für Reihen	42
4.6	Potenzreihen	44
4.7	Grenzwerte von Funktionen und Stetigkeit	47
5	Differentiation	50
5.1	Tangente und Ableitung	50
5.2	Differentiationsregeln	54
5.3	Mittelwertsatz (der Differentialrechnung)	56
5.4	Extrema und Wendepunkte	59
5.5	Der Satz von Taylor	61
5.6	Das Newtonsche Iterationsverfahren	65
6	Integration	68
6.1	Das bestimmte Integral	68
6.2	Eigenschaften des bestimmten Integrals	70
6.2.1	Rechenregeln	70
6.2.2	Flächeninhalt	70
6.2.3	Mittelwertsatz der Integralrechnung	71
6.3	Zusammenhang zwischen Differential- und Integralrechnung	71
6.4	Integrationsmethoden	73
6.4.1	Partielle Integration (Produktregel)	73
6.4.2	Substitution (Kettenregel)	73
6.4.3	Einige Standardsubstitutionen	75
6.4.4	Partialbruchzerlegung	77
6.5	Uneigentliche Integrale	80
6.5.1	Unbeschränkter Integrationsbereich	80
6.5.2	Unbeschränkter Integrand	82
6.6	Numerische Integration	83
6.6.1	Newton-Cotes-Formeln	83

7	Elemente der linearen Algebra	85
7.1	Der euklidische Raum \mathbb{R}^n	85
7.2	Lineare Unabhängigkeit	88
7.3	Lineare Abbildungen und Matrizen	90
7.4	Lineare Gleichungssysteme	97
7.5	Der Gaußsche Algorithmus	99
7.6	Determinanten	103
7.7	Anwendung der Linearen Algebra in der Geometrie	108
7.7.1	Geometrie in der Ebene	108
7.7.2	Geometrie im Raum	110
	Index	113

Bezeichnungen

$A \wedge B$	A und B
$A \vee B$	A oder B
$A \Rightarrow B$	Aus A folgt B
$A \Leftrightarrow B$	Die Aussagen A und B sind gleichwertig
\neg	nicht
$\{\cdot\cdot\cdot\}$	Mengenklammern
\in	ist Element von
\notin	ist nicht Element von
\emptyset	leere Menge
\subset	Teilmenge
\cup	vereinigt mit
\cap	geschnitten mit
$=$	gleich
\neq	ungleich
\approx	ungefähr gleich
$<$	kleiner als
\leq	kleiner als oder gleich
$>$	größer als
\geq	größer als oder gleich
$A \setminus B$	A ohne B
$A \times B$	Kartesisches Produkt von A und B
\mathbb{N}	Menge der natürlichen Zahlen
\mathbb{Z}	Menge der ganzen Zahlen
\mathbb{R}	Menge der reellen Zahlen
\mathbb{Q}	Menge der rationalen Zahlen
\mathbb{C}	Menge der komplexen Zahlen
o.E.	ohne Einschränkung
\forall	für alle
\exists	es gibt
\sum	Summenzeichen
\prod	Produktzeichen
$n!$	n Fakultät
$\binom{n}{k}$	n über k (Binomialkoeffizient)
$\sqrt{\quad}$	Quadratwurzel
$\sqrt[n]{\quad}$	n -te Wurzel

$[a, b]$	abgeschlossenes Intervall
(a, b)	geordnetes Paar oder offenes Intervall
$]a, b[$	offenes Intervall
$[a, b)$	(rechts) halb offenes Intervall
$(a, b]$	(links) halb offenes Intervall
$ x $	Betrag von x
sgn	signum (Vorzeichen)
log	Logarithmus (auch lg)
ln	Natürlicher Logarithmus
e^x	Exponentialfunktion
$P(x)$	Polynom
sinh	Sinus hyperbolicus
cosh	Cosinus hyperbolicus
$\begin{vmatrix} \dots \\ \dots \\ \dots \end{vmatrix}, \det$	Determinante
$\begin{pmatrix} \dots \\ \dots \\ \dots \end{pmatrix}$	Matrix
\vec{a}	Vektor
$\vec{a} \times \vec{b}$	Vektorprodukt
$f(x)$	Funktionswert (f von x)
$\frac{df}{dx}$	1. Ableitung (df nach dx)
D_f	Definitionsbereich von f
W_f	Wertebereich von f
D_{\max}	maximaler Definitionsbereich
gr(f)	Graph von f
f^{-1}	Umkehrfunktion von f
Re z	Realteil von z
Im z	Imaginärteil von z
i	imaginäre Einheit
\bar{z}	konjugiert komplexe Zahl
arg(z)	Argument von z
sin	Sinus
cos	Cosinus
tan	Tangens
cot	Kotangens
arcsin	Arkussinus
arccos	Arkuscosinus
arctan	Arkustangens

Kapitel 1

Grundbegriffe

1.1 Aussagenlogik

In der Mathematik betrachten wir nur solche Aussagen, die entweder wahr (w) oder falsch (f) sind. Eine dritte Möglichkeit wird nicht zugelassen. In der mathematischen Logik heißt dieses Prinzip *Satz vom ausgeschlossenen Dritten*. Wir erlauben auch nicht, dass eine Aussage gleichzeitig wahr und falsch ist; dies ist der *Satz vom ausgeschlossenen Widerspruch*.

Beispiele 1.1.1: (i) A_1 : „3 ist eine ungerade Zahl“; Wahrheitswert: w

(ii) A_2 : „4 plus 5 ist gleich 10“; Wahrheitswert: f

(iii) A_3 : „Guten Morgen“; keine Aussage i.o. Sinn

(iv) A_4 : „Die Erde ist keine Scheibe“; Wahrheitswert: w

(v) A_5 : „Die Erde ist eine Kugel“; Wahrheitswert: f

(vi) A_6 : „Die millionste Ziffer von π ist 7.“; Wahrheitswert: sicher w oder f, aber was?

Vorsicht: die Beispiele A_4 und A_5 sind zwar Aussagen, beziehen sich jedoch nicht auf mathematische Sachverhalte.

Logische Verknüpfungen von Aussagen:

1. *Konjunktion* („und“), Abkürzung: \wedge

$A_1 \wedge A_2$ bedeutet: Es gilt A_1 *und* A_2 , genauer: Die Aussage $A_1 \wedge A_2$ ist genau dann (dann und nur dann) wahr, falls A_1 wahr ist *und* A_2 wahr ist.

Beispiel 1.1.2: A_1 : 3 ist ungerade. A_7 : 3 ist eine Primzahl. Dann gilt $A_1 \wedge A_7$ ist wahr.

2. *Disjunktion* („oder“), Abkürzung: \vee

$A_1 \vee A_2$ bedeutet: Es gilt A_1 *oder* A_2 .

Beachte: $A_1 \vee A_2$ heißt nicht „entweder A_1 oder A_2 “, sondern im Sinne von „und/oder“.

Beispiel 1.1.3: $A_1 \vee A_2$: 3 ist ungerade Zahl oder 4 plus 5 ist gleich 10 (wahr).

3. *Negation* („nicht“, Verneinung), Abkürzung: \neg

Beispiel 1.1.4: $\neg A_1$: 3 ist keine ungerade Zahl (falsch).

Beachte: „ $\neg A_1$: 3 ist eine gerade Zahl“ ist keine formal korrekte Negation von A_1 . Nur wenn wir „wissen“, bzw. festlegen dass jede natürliche Zahl gerade oder ungerade sein muss, ist die Gleichsetzung von „nicht ungerade“ und „gerade“ erlaubt.

4. *Implikation* („wenn ... dann“, Folgerung), Abkürzung: \Rightarrow

Beispiel 1.1.5: F_1 : n ist eine gerade Zahl. F_2 : n^2 ist eine gerade Zahl. Dann gilt: $F_1 \Rightarrow F_2$.

Bemerkungen 1.1.6:

- (i) F_1 und F_2 sind strenggenommen keine Aussagen, sondern sogenannte Aussageformen. Sie erhalten erst dann einen festen Wahrheitswert, wenn für n eine bestimmte Zahl eingesetzt wird! Dennoch ist „ $F_1 \Rightarrow F_2$ “ stets wahr.
- (ii) Sprechweisen für „ $A \Rightarrow B$ “ sind: Aus A folgt B , A impliziert B , A ist hinreichend für B , B ist notwendig für A .
- (iii) A falsch, dann $A \Rightarrow B$ immer wahr!
„Wenn die Erde eine Scheibe ist, dann ist der Mond fünfeckig“
- (iv) Aus $A \Rightarrow B$ folgt *nicht* $B \Rightarrow A$!

Beispiel 1.1.7: A : Es regnet, B : Die Straße ist nass.

5. *Äquivalenz* („genau dann ... wenn“), Abkürzung: \Leftrightarrow

Beispiel 1.1.8: A : Die Zahlen a und b sind ungerade. B : Das Produkt $a \cdot b$ ist ungerade. Dann gilt: $A \Leftrightarrow B$.

Bemerkungen 1.1.9:

- (i) Sprechweise: A und B sind äquivalent, A ist notwendig und hinreichend für B .
- (ii) $A \Leftrightarrow B$ ist gleichbedeutend mit $(A \Rightarrow B) \wedge (B \Rightarrow A)$.

Eine strenge Definition dieser Verknüpfungen ist mittels folgender Wahrheitstafel möglich:

A	B	$\neg A$	$A \wedge B$	$A \vee B$	$A \Rightarrow B$	$A \Leftrightarrow B$
w	w	f	w	w	w	w
w	f	f	f	w	f	f
f	w	w	f	w	w	f
f	f	w	f	f	w	w

Kontrapositionsgesetz: $(A \Rightarrow B) \Leftrightarrow (\neg B \Rightarrow \neg A)$

Beispiel 1.1.10: $(x^2 \leq 1 \Rightarrow x \leq 1) \Leftrightarrow (x > 1 \Rightarrow x^2 > 1)$.

Die aussagenlogischen „Spielregeln“ in diesem Abschnitt sollen uns als Grundlage für die Formulierung mathematischer Sachverhalte ausreichen. Eine etwas weiter ausgreifende Darstellung findet sich z.B. bei Dallmann/Elster, Mathematik I, S. 17–32.

1.2 Mengen und Zahlen

Wir wollen nun die wichtigsten grundlegenden mathematischen Objekte einführen¹ und insbesondere die Notation für den Rest der Vorlesung festlegen.

Man fasst oft gewisse Objekte zu einer Gesamtheit zusammen, etwa die Einwohnerschaft Darmstadts (= Gesamtheit aller Einwohner Darmstadts). Eine solche Gesamtheit nennen wir *Menge*². Die einzelnen Objekte heißen *Elemente* der Menge. Man kann eine Menge angeben, indem man

1. alle Elemente angibt, z.B.: $M_1 = \{\alpha, \beta, \gamma, \delta, \varepsilon\}$, $M_2 = \{!, ?, @\}$.
2. alle Elemente durch eine Vorschrift erklärt, z.B.:

$$\begin{aligned} M_3 &= \{x : x \text{ ist eine ganze Zahl und } x > 10\} \\ &= \{11, 12, 13, \dots\} \\ &= \{x \mid x \text{ ist eine ganze Zahl, } x > 10\} \end{aligned}$$

Wie angedeutet erlauben wir auch Abkürzungen wie z.B. $M_4 = \{A, B, C, \dots\}$ etc., wenn diese unmissverständlich sind. Beachten Sie auch, dass in den soeben vorgenommenen Definitionen der Mengen M_1, M_2, M_3 und M_4 das Zeichen $=$ eine andere Funktion hat als im Ausdruck $2 + 2 = 4$. Im ersteren Fall wird durch das Gleichheitszeichen etwas *definiert*, im zweiten etwas *ausgesagt*. In der Literatur wird für Definitionen oft das Symbol $:=$ verwendet. Wir werden in dieser Vorlesung ebenfalls darauf zurückgreifen, wenn wir neue Schreibweisen einführen.

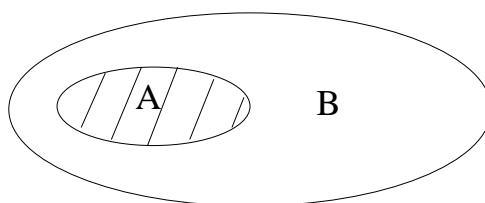
Aus formalen Gründen definiert man die *leere Menge* (kurz \emptyset) als Menge ohne Element. Ist a ein Element der Menge M , so schreibt man $a \in M$; entsprechend $a \notin M$ für „ a ist kein Element der Menge M “³.

Hier: $\gamma \in M_1$, $7 \notin M_2$.

Teilmengen und Obermengen:

A heißt *Teilmenge* von B , falls jedes $a \in A$ auch zu B gehört; B heißt dann *Obermenge* von A , kurz $A \subset B$, und $B \supset A$. Achtung: In vielen Büchern wird statt \subset auch \subseteq verwendet um zu betonen, dass die Mengen gleich sein können. Für $A \subset B$ und $A \neq B$ schreiben wir $A \subsetneq B$ und nennen A eine *echte Teilmenge* von B .

$A = B$ bedeutet $A \subset B$ und $B \subset A$.



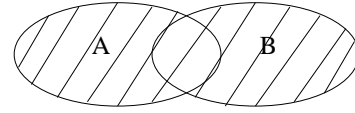
¹Strenggenommen haben wir im vorigen Abschnitt schon eine gewisse Vertrautheit mit mathematischen Begriffen (z.B. den natürlichen Zahlen) vorausgesetzt.

²Dieser *naive* Mengenbegriff reicht für unsere Zwecke aus, birgt jedoch einige Schwierigkeiten, z.B. die *Russellsche Antinomie*.

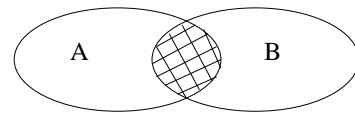
³Allgemein erlauben wir uns, Beziehungen durch $/$ zu negieren, wenn klar ist, was gemeint ist.

Mengenoperationen:

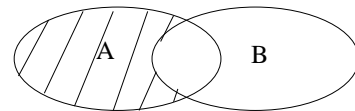
Vereinigung: $A \cup B = \{x : x \in A \text{ oder } x \in B\}$



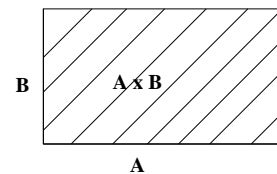
Durchschnitt: $A \cap B = \{x : x \in A \text{ und } x \in B\}$



Differenz „A ohne B“: $A \setminus B = \{x : x \in A \text{ und } x \notin B\}$



Kartesisches Produkt: $A \times B = \{(a, b) : a \in A, b \in B\}$



Dabei bezeichnet (a, b) ein geordnetes Paar, die Reihenfolge ist wichtig: $(2, 3) \neq (3, 2)$.

Wir können die Mengenoperationen natürlich auch mehrfach anwenden, z.B.

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

A und B heißen *disjunkt*, falls sie kein gemeinsames Element haben, d.h. falls $A \cap B = \emptyset$ gilt.

Die wichtigsten Mengen in der Mathematik sind die *Zahlen*. Man unterscheidet folgende Grundmengen von Zahlen:

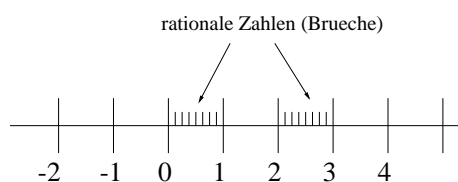
$\mathbb{N} = \{1, 2, 3, 4, \dots\}$, die *natürlichen* Zahlen,

$\mathbb{N}_0 = \{0, 1, 2, 3, \dots\}$, die *nichtnegativen ganzen* Zahlen,

$\mathbb{Z} = \{\dots, -3, -2, -1, 0, 1, 2, 3, \dots\}$, die *ganzen* Zahlen,

$\mathbb{Q} = \{\frac{p}{q} : p, q \in \mathbb{Z}, q \neq 0\}$, die *rationalen* Zahlen (Brüche).

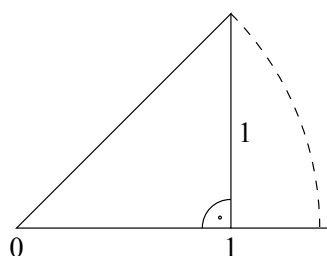
Geometrisch lassen sich diese Zahlen als Zahlengerade (oder -strahl) auffassen:



Das Rechnen mit rationalen Zahlen ist uns aus der Schule bekannt. Zur Erinnerung: Es gelten $\frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{bd}$ (und *nicht*: $\frac{a+c}{b+d}$!), sowie $\frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd}$ und $\frac{a}{b} : \frac{c}{d} = \frac{a}{b} \cdot \frac{d}{c} = \frac{ad}{bc}$.

Zwar gibt es zwischen je zwei Brüchen noch beliebig viele weitere rationale Zahlen, dennoch wird die Zahlengerade *nicht* vollständig ausgefüllt!

Beispiel 1.2.1:



Die Zahl $\sqrt{2}$ (definiert durch die Beziehung: $\sqrt{2}^2 = 2$) ist nicht Element von \mathbb{Q} ; deshalb heißt $\sqrt{2}$ eine *irrationale* Zahl.

Die Vereinigung von \mathbb{Q} mit der Menge aller irrationalen Zahlen ergibt die Menge \mathbb{R} der *reellen* Zahlen. Die Rechenregeln in \mathbb{R} sind dieselben wie in \mathbb{Q} . Es gilt

$$\mathbb{N} \subset \mathbb{N}_0 \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}.$$

Dies sind offensichtlich jeweils echte Teilmengen. Weniger offensichtlich ist: Die Menge der rationalen Zahlen \mathbb{Q} ist *abzählbar*, d.h. sie kann durchnummeriert werden. Die Menge der reellen Zahlen \mathbb{R} ist *überabzählbar*. Insbesondere liegen zwischen je zwei verschiedenen rationalen Zahlen unendlich viele irrationale Zahlen.

1.3 Elementare Beweismethoden

1.3.1 Direkter Beweis

Voraussetzung: Aussage A

Behauptung: Aussage B

Beweis: $A \Rightarrow A_1 \Rightarrow A_2 \Rightarrow \dots \Rightarrow B$

(statt \Rightarrow ist auch \Leftrightarrow zulässig)

Beispiel 1.3.1: Wenn n eine gerade natürliche Zahl ist, dann ist auch n^2 gerade.

Beweis.

$$A : n \in \mathbb{N} \text{ ist gerade} \Leftrightarrow n = 2k \text{ für ein } k \in \mathbb{N}$$

$$\Rightarrow n^2 = 4k^2 \Rightarrow n^2 = 2m \text{ mit } m = 2k^2 \text{ und } k \in \mathbb{N}$$

$$\Rightarrow n^2 = 2m \text{ mit } m \in \mathbb{N} \Leftrightarrow n^2 \text{ ist gerade.} \quad \square$$

1.3.2 Indirekter Beweis

Voraussetzung: A

Behauptung: B

Beweis: $\neg B \Rightarrow C_1 \Rightarrow C_2 \Rightarrow \dots \Rightarrow \neg A$

oder $A \wedge \neg B \Rightarrow \dots \Rightarrow C \wedge \neg C$ d.h. eine falsche Aussage,
daher auch „Beweis durch Widerspruch“

Beispiel 1.3.2: $\sqrt{2} \notin \mathbb{Q}$.

Beweis. Es ist $A : x = \sqrt{2}$ $B : x \notin \mathbb{Q}$ oder besser

$$A : x > 0 \text{ mit } x^2 = 2 \quad B : x \neq \frac{p}{q} \text{ für alle } p, q \in \mathbb{N}$$

$$\neg B : x = \frac{p}{q} \text{ mit } p, q \in \mathbb{N}$$

dann gilt ohne Einschränkung (kurz: o.E.): p und q teilerfremd und $x^2 = 2$

$$\Rightarrow \frac{p^2}{q^2} = 2 \Rightarrow p^2 = 2q^2, \text{ also ist } p^2 \text{ gerade}$$

$$\Rightarrow p \text{ gerade, also } p = 2m \text{ mit } m \in \mathbb{N}$$

$$\Rightarrow q^2 = 2m^2 \text{ also } q^2 \text{ gerade}$$

$$\Rightarrow q \text{ gerade} \Rightarrow p \text{ und } q \text{ durch } 2 \text{ teilbar; Widerspruch!} \quad \square$$

1.3.3 Vollständige Induktion

Die folgende Formel wird oft C.F. Gauß (1777-1855) zugeschrieben, der sie als 10-jähriger Schüler fand:

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

Dies ist ein Beispiel für eine Aussage, die von einem Parameter (hier: n) abhängt. Wie kann man beweisen, dass diese Aussage für jedes $n \in \mathbb{N}$ richtig ist?

Gegeben sind Aussagen $A(n)$ für alle natürlichen Zahlen $n \in \mathbb{N}$.

Behauptung: $A(n)$ ist wahr für jedes $n \in \mathbb{N}$.

Beweisschema der vollständigen Induktion:

1. Induktionsanfang: Zeige, dass $A(1)$ wahr ist.

2. Induktionsschritt: Zeige $A(n) \Rightarrow A(n+1)$ für beliebiges $n \in \mathbb{N}$.

Dann gilt $A(n)$ für jedes $n \in \mathbb{N}$.

„Domino-Effekt“: $\overset{1}{\curvearrowright} \overset{2}{\curvearrowright} \overset{3}{\curvearrowright} \overset{4}{\curvearrowright} \overset{5}{\curvearrowright} \dots$

Notation: \forall bedeutet „für alle“ (Allquantor), \exists heißt „es gibt“ (Existenzquantor).

Abkürzung für Summation von n Zahlen a_1 bis a_n :

$$a_1 + a_2 + \dots + a_n =: \sum_{k=1}^n a_k, \quad \text{etwa} \quad \sum_{k=1}^n k = 1 + 2 + \dots + n,$$

$$\text{z.B.} \quad \sum_{k=1}^3 k = 1 + 2 + 3 = 6 = \frac{3 \cdot 4}{2}$$

Mit diesen Abkürzungen können wir schreiben

Behauptung:

$$\sum_{k=1}^n k = \frac{n(n+1)}{2} \quad \forall n \in \mathbb{N}.$$

Beweis.

1. Induktionsanfang:

$$\sum_{k=1}^1 k = 1 = \frac{1(1+1)}{2} \quad \text{also ist } A(1) \text{ wahr.}$$

2. Induktionsschritt: Zeige, dass wenn für ein gewisses $n \in \mathbb{N}$

$$\sum_{k=1}^n k = \frac{n(n+1)}{2} \quad \text{gilt,}$$

dann folgt daraus

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

Es gilt:

$$\begin{aligned} \sum_{k=1}^{n+1} k &= \sum_{k=1}^n k + n + 1 = \frac{n(n+1)}{2} + n + 1 \\ &= \frac{n(n+1) + 2(n+1)}{2} = \frac{(n+2)(n+1)}{2} \end{aligned} \quad \square$$

Aus den bekannten Rechenregeln für die Addition und Multiplikation ergeben sich die folgenden Regeln für den Umgang mit der Summennotation:

$$c \sum_{k=1}^n a_k = \sum_{k=1}^n c a_k, \quad \sum_{k=1}^n a_k + \sum_{k=1}^n b_k = \sum_{k=1}^n (a_k + b_k),$$

$$\sum_{k=1}^n a_k = \sum_{k=1}^n a_{n+1-k}, \quad \sum_{k=1}^m a_k + \sum_{k=m+1}^n a_k = \sum_{k=1}^n a_k \quad (1 \leq m < n),$$

Bemerkung 1.3.3. Gelegentlich hat man Aussagen $A(n)$ für alle $n \in \mathbb{N}$ mit $n \geq n_0$ mit anderem $n_0 \in \mathbb{N}$ statt $n_0 = 1$. (Wir schreiben der Einfachheit halber $n \geq n_0$ statt $n \in \mathbb{N} \wedge n \geq n_0$.) Es ist offensichtlich, dass wir das Prinzip der Induktion auch auf andere Grundmengen anwenden können, solange sie ein „Anfangselement“ enthalten und jedes Element einen klar bestimmten „Nachfolger“ hat.

Beispiel 1.3.4: Für welche $n \in \mathbb{N}$ gilt $2^n > n^2$?

Also $A(n) : 2^n > n^2$; Vermutung: $A(n)$ ist wahr für $n = 1$ und für alle $n \geq 5$.

Beweis. Offenbar gilt die Aussage für $n = 1$. Wir beweisen durch vollständige Induktion, dass sie für alle $n \geq 5$ gilt. Dazu starten wir mit $A(5)$ und beachten im Induktionsschritt die zusätzliche Voraussetzung $n \geq 5$.

1. Induktionsanfang: Offenbar gilt

$$A(5) : 32 = 2^5 > 5^2 = 25.$$

2. Induktionsschritt: Es gelte $2^n > n^2$ für ein gewisses $n \geq 5$. Dann gilt dank einer binomischen Formel:

$$2^{n+1} = 2 \cdot 2^n > 2 \cdot n^2 = n^2 + n \cdot n > n^2 + 3n > n^2 + 2n + 1 = (n+1)^2 \quad \square$$

1.4 Binomialkoeffizienten und Binomialsatz

Wir erinnern uns an die *Binomischen Formeln*: Für $a, b \in \mathbb{R}$ gilt

$$\begin{aligned} (a+b)^2 &= a^2 + 2ab + b^2 \\ (a-b)^2 &= a^2 - 2ab + b^2 \\ (a+b)(a-b) &= a^2 - b^2 \\ (a+b)^3 &= (a+b)(a+b)^2 = a(a^2 + 2ab + b^2) + b(a^2 + 2ab + b^2) \\ &= a^3 + 2a^2b + ab^2 + a^2b + 2ab^2 + b^3 \\ &= a^3 + 3a^2b + 3ab^2 + b^3 \end{aligned}$$

1.4.1 Fakultät und Binomialkoeffizient

Definition 1.4.1 (Fakultät). Die Zahl $n! := 1 \cdot 2 \cdot 3 \cdot \dots \cdot n$ heißt „ n -Fakultät“. Aus formalen Gründen definiert man zusätzlich $0! := 1$.

Mit Hilfe des Produktzeichens

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

können wir auch schreiben $n! = \prod_{k=1}^n k$. Weiter gilt $(n+1)! = n!(n+1)$ für alle $n \in \mathbb{N}_0$.

Definition 1.4.2 (Binomialkoeffizient). Die Zahl

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

für $n \in \mathbb{N}_0$ und $0 \leq k \leq n$ heißt *Binomialkoeffizient*; gelesen als „ n über k “.

Bemerkung 1.4.3. Es gilt

$$\binom{n}{k} = \binom{n}{n-k}, \quad \binom{n}{k} = \frac{(n-k+1)(n-k+2)\cdots(n-1)n}{1 \cdot 2 \cdots (k-1)k}.$$

Beispiele 1.4.4: $\binom{4}{2} = \frac{4!}{2!2!} = \frac{24}{4} = 6, \quad \binom{14}{12} = \frac{13 \cdot 14}{1 \cdot 2} = 91, \quad \binom{10}{0} = 1.$

1.4.2 Binomische Formeln und Pascalsches Dreieck

Wie lauten die Koeffizienten nach Ausmultiplikation des Produktes $(a+b)^n$?

$$\begin{array}{lccccccc} n=0: & (a+b)^0 & = & 1 & & & & & & & 1 \\ n=1: & (a+b)^1 & = & a+b & & & & & & & 1 & 1 \\ n=2: & (a+b)^2 & = & a^2+2ab+b^2 & & & & & & & 1 & 2 & 1 \\ & & & \vdots & & & & & & & & & \vdots \\ & & & & & & & & & & 1 & 3 & 3 & 1 \\ & & & & & & & & & & & 1 & 4 & 6 & 4 & 1 \\ & & & & & & & & & & & & \nearrow_{k=0} & \nearrow_{k=1} & \nearrow_{k=2} & & & & \end{array}$$

Der k -te Eintrag in Zeile n ist $\binom{n}{k}$, und die Einträge im Pascalschen Dreieck entstehen jeweils als Summe der unmittelbar darüberstehenden Nachbarn:

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

Der durch diese Beispiele nahegelegte Zusammenhang zwischen binomischen Formeln und Binomialkoeffizienten kann bestätigt werden:

Satz 1.4.5 (Binomialsatz). Für $a, b \in \mathbb{R}$ und $n \in \mathbb{N}_0$ ist

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

Der Beweis kann durch vollständige Induktion erbracht werden.

1.5 Ungleichungen und Beträge

Die reellen Zahlen \mathbb{R} besitzen folgende *Anordnungseigenschaften*: Für alle $a, b, c \in \mathbb{R}$ gilt

1. Es gilt genau eine der Beziehungen: $a < b$ oder $a = b$ oder $b < a$.
2. $a < b \wedge b < c \Rightarrow a < c$ (Transitivität)
3. $a < b \Rightarrow a + c < b + c$ (Monotonie der Addition)
4. $a < b \Rightarrow ac < bc$ für $c > 0$
 $\Rightarrow ac > bc$ für $c < 0$ (Monotonie der Multiplikation)

Bemerkungen 1.5.1: (i) $a < b$ heißt Ungleichung („ a kleiner als b “); gleichbedeutend mit $b > a$ („ b größer als a “).

(ii) $a \leq b$ bedeutet $a < b$ oder $a = b$. Obige Regeln gelten entsprechend für \leq .

Aus den Anordnungseigenschaften ergeben sich die einige Folgerungen:

$$a < 0 \Leftrightarrow -a > 0 \quad (1.1)$$

(man verwende Eigenschaft 4 mit $c := -1$ oder Eigenschaft 3 mit $c := -a$)

$$0 < a < b \Rightarrow a^n < b^n \quad \forall n \in \mathbb{N}. \quad (1.2)$$

Beweis. Durch vollständige Induktion: Für $n = 1$ ist die Aussage sicher wahr. Sei also $0 < a < b$ und $a^n < b^n$. Dann gilt

$$a^{n+1} = a^n \cdot a < b^n a \text{ und } ab^n < b \cdot b^n = b^{n+1}, \quad \text{also auch } a^{n+1} < b^{n+1}. \quad \square$$

Definition 1.5.2 (Wurzeln). Sei $a \in \mathbb{R}$ mit $a \geq 0$. Unter $\sqrt[n]{a}$ („ n -te Wurzel aus a “) versteht man diejenige *nichtnegative* reelle Zahl x , für die $x^n = a$ gilt.

Es gilt auch

$$0 \leq a < b \Rightarrow \sqrt[n]{a} < \sqrt[n]{b}. \quad (1.3)$$

Bernoullische Ungleichung: Für $n \in \mathbb{N}$ und $x \in \mathbb{R}$ mit $x > -1$ gilt

$$(1+x)^n \geq 1+nx. \quad (1.4)$$

Beweis. Durch vollständige Induktion: Die Aussage ist wahr für $n = 1$; zu zeigen bleibt der Schritt $n \rightarrow n+1$:

$$\begin{aligned} (1+x)^{n+1} &= (1+x)(1+x)^n \geq (1+x)(1+nx) \\ &= 1+(n+1)x+nx^2 \geq 1+(n+1)x, \text{ denn } x^2 \geq 0 \quad \forall x. \end{aligned} \quad \square$$

Die Bernoullische Ungleichung kann verwendet werden, um komplizierte Terme mit Potenzen nach unten abzuschätzen. Hier ist ein einfaches Beispiel.

Beispiel 1.5.3: Wir wollen eine untere Abschätzung für $2 \cdot 2^{10}$ finden. Dazu rechnen wir

$$2 \cdot 2^{10} = (2+0, 2)^{10} = [2(1+0, 1)]^{10} = 2^{10}(1+0, 1)^{10} \geq 2^{10}(1+10 \cdot 0, 1) = 2^{10} \cdot 2 = 2048.$$

Zum Vergleich: Der reale Wert ist ungefähr 2656.

Beispiel 1.5.4 (zu Ungleichungen): Für welche $x \in \mathbb{R}$ gilt $\frac{2x+1}{x-1} < 1$? Gesucht ist also

$$L = \left\{ x \in \mathbb{R} : \frac{2x+1}{x-1} < 1 \right\},$$

die *Lösungsmenge* der Ungleichung; dabei ist die zugrunde liegende Zahlenmenge (hier: \mathbb{R}) wichtig! Eine Rechnung mit Fallunterscheidung liefert $L = \{x \in \mathbb{R} : -2 < x < 1\}$. Dies ist ein sogenanntes Intervall.

Definition 1.5.5 (Intervalle). Die folgenden Teilmengen von \mathbb{R} werden *Intervalle* genannt; dabei seien $a, b \in \mathbb{R}$ mit $a < b$.

(i) $[a, b] := \{x \in \mathbb{R} : a \leq x \leq b\}$ abgeschlossenes Intervall,

- (ii) $(a, b) := \{x \in \mathbb{R} : a < x < b\}$ offenes Intervall, (auch $]a, b[$),
- (iii) $[a, b) := \{x \in \mathbb{R} : a \leq x < b\}$ (rechts) halboffenes Intervall (auch $[a, b[$), $(a, b] := \{x \in \mathbb{R} : a < x \leq b\}$ (links) halboffenes Intervall (auch $]a, b]$),
- (iv) $[a, \infty) := \{x \in \mathbb{R} : x \geq a\}$ und $(a, \infty) := \{x \in \mathbb{R} : x > a\}$,
 $(-\infty, b] := \{x \in \mathbb{R} : x \leq b\}$ und $(-\infty, b) := \{x \in \mathbb{R} : x < b\}$.

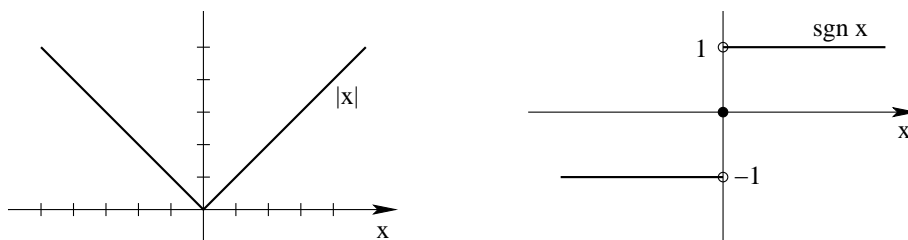
Dabei heißen jeweils a der linke Randpunkt, b der rechte Randpunkt und $(b-a)$ die Intervalllänge.

Definition 1.5.6 (Betrag und Signum). Sei $x \in \mathbb{R}$. Dann bedeutet

$$|x| := \begin{cases} x & \text{für } x \geq 0 \\ -x & \text{für } x < 0 \end{cases} \quad (\text{Absolut-}) \text{ Betrag von } x \text{ und}$$

$$\operatorname{sgn} x := \begin{cases} 1 & \text{für } x > 0 \\ 0 & \text{für } x = 0 \\ -1 & \text{für } x < 0 \end{cases} \quad \text{Signum oder Vorzeichen von } x.$$

Bemerkungen 1.5.7: (i) Wir können $|x|$ und $\operatorname{sgn} x$ als *Funktion*⁴ des Arguments x auffassen. Die Funktionsgraphen haben folgenden Verlauf:



- (ii) Für $x \in \mathbb{R}$ ist $|x|$ der Abstand von x zum Nullpunkt. Entsprechend ist $|x - y|$ der Abstand von x zu $y \in \mathbb{R}$.
- (iii) Es gilt $x = |x| \operatorname{sgn} x$ und $|x| = x \operatorname{sgn} x$ für $x \in \mathbb{R}$.
- (iv) Wichtige Eigenschaften des Betrages $|\cdot|$ sind
- (a) $|x| \geq 0 \quad \forall x \in \mathbb{R}$, und $|x| = 0 \Leftrightarrow x = 0$,
 - (b) $|\lambda x| = |\lambda| \cdot |x|$ für $\lambda, x \in \mathbb{R}$,
 - (c) $|x + y| \leq |x| + |y|$ für $x, y \in \mathbb{R}$ (Dreiecksungleichung).

Diese Eigenschaften sind auch für den (euklidischen) Abstand im \mathbb{R}^2 erfüllt (und allgemeiner im \mathbb{R}^n für $n \geq 2$).

(v) Weiterhin gilt

$$(d) \quad -|x| \leq x \leq |x| \quad \forall x \in \mathbb{R}$$

⁴In Kapitel 3 werden wir uns dem Funktionsbegriff näher zuwenden.

- (vi) Aufgrund der stückweisen Definition von $|x|$ können (Un-)Gleichungen, in denen $|x|$ vorkommt, am besten durch Fallunterscheidung aufgelöst werden.

Beispiele 1.5.8:

1. Für welche $x \in \mathbb{R}$ gilt $|x - 3| \leq 1$?

Anschaulich: Abstand von x zu 3 ist kleiner oder gleich 1, also $L = [2, 4]$.

1. Fall: $x - 3 \geq 0 \Leftrightarrow x \geq 3$.

Dann $|x - 3| \leq 1 \Leftrightarrow x - 3 \leq 1 \Leftrightarrow x \leq 4$. Dies liefert $[3, 4]$ als Lösungen.

2. Fall: $x - 3 < 0 \Leftrightarrow x < 3$.

Dann $|x - 3| \leq 1 \Leftrightarrow -(x - 3) \leq 1 \Leftrightarrow x - 3 \geq -1 \Leftrightarrow x \geq 2$. Dies liefert $[2, 3)$ als Lösungen.

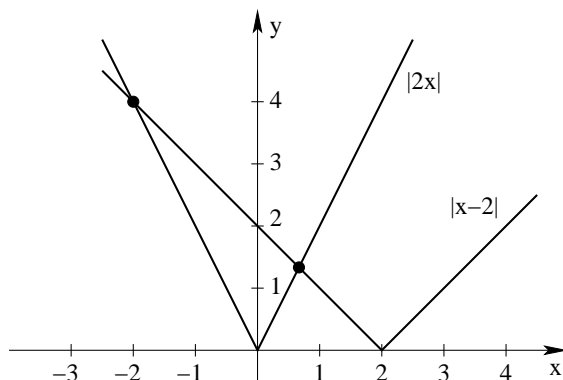
Es folgt also $L = [3, 4] \cup [2, 3) = [2, 4]$

2. Für welche $x \in \mathbb{R}$ gilt $|x - 2| = |2x|$?

1. Fall: $x \geq 2$: $|x - 2| = |2x| \Leftrightarrow x - 2 = 2x \Leftrightarrow -2 = x$; keine Lösung in diesem Fall ($x \geq 2$).

2. Fall: $0 \leq x < 2$: $|x - 2| = |2x| \Leftrightarrow 2 - x = 2x \Leftrightarrow x = \frac{2}{3}$; dies ist eine Lösung.

3. Fall: $x < 0$: $2 - x = -2x \Leftrightarrow 2 = -x \Leftrightarrow x = -2$; ist ebenfalls eine Lösung.



3. Es gilt $x^2 \leq y^2 \Leftrightarrow |x| \leq |y|$ für $x, y \in \mathbb{R}$!

Denn: $\sqrt{\cdot}$ ist monoton und $\sqrt{x^2} = |x| \quad \forall x \in \mathbb{R}$! ($\sqrt{|-2|^2} = 2 \neq -2$!)

4. Es gilt

$$|x - 2| \leq |x| \Leftrightarrow (x - 2)^2 \leq x^2 \Leftrightarrow x^2 - 4x + 4 \leq x^2 \Leftrightarrow 4 \leq 4x \Leftrightarrow x \geq 1.$$

5. Es gilt

$$x^2 + 4x \geq 1 \Leftrightarrow x^2 + 4x + 4 \geq 5 \Leftrightarrow (x + 2)^2 \geq 5 \Leftrightarrow |x + 2| \geq \sqrt{5}$$

$$\Leftrightarrow x \geq \sqrt{5} - 2 \text{ oder } x \leq -2 - \sqrt{5}, \text{ also } L = (-\infty, -2 - \sqrt{5}] \cup [\sqrt{5} - 2, \infty).$$

Kapitel 2

Elementare Rechentechniken

2.1 Potenzen

2.1.1 Exponentialschreibweise

Rechnen mit sehr großen bzw. kleinen Zahlen in Dezimalschreibweise ist unhandlich. Deshalb verwendet man oft die sogenannte exponentielle Standardform: Jede reelle Zahl $x \neq 0$ kann als Produkt aus einem *Vorzeichen*, einer Zahl zwischen 1 und 10 (der *Mantisse*) und einer Zehnerpotenz mit ganzzahligem *Exponenten* geschrieben werden:

$$x = \pm c \cdot 10^n, \text{ mit } c \in [1, 10) \text{ und } n \in \mathbb{Z}. \quad (2.1)$$

Beispiele:

$$x = +1.17 \cdot 10^0 = 1.17$$

$$x = +2.37 \cdot 10^1 = 23.7$$

$$x = -3.17 \cdot 10^{-2} = -0.0317$$

Bemerkungen 2.1.1: (i) Für die Zahldarstellung auf Rechnern wird die sogenannte *normalisierte Gleitpunktdarstellung* gewählt, dazu wird die Mantisse aus dem Intervall $[0.1, 1)$ gewählt.

(ii) Die Zahl n gibt die Anzahl der Stellen an, um die der Dezimalpunkt verschoben werden muss, um aus c die Zahl x zu erhalten. Bei positivem n ist der Punkt nach rechts, bei negativem nach links zu schieben: $1 \cdot 10^{-3} = 0.001$.

2.1.2 Multiplikation und Division

Zur Erinnerung: Die folgenden Regeln gelten für das Rechnen mit Potenzen:

$$a^n \cdot a^m = a^{n+m} \quad \text{bei gleicher Basis (= Grundzahl),}$$

$$a^n \cdot b^n = (ab)^n \quad \text{bei gleichem Exponenten (= Hochzahl),}$$

$$\frac{a^n}{a^m} = a^{n-m}, \quad \frac{a^n}{b^n} = \left(\frac{a}{b}\right)^n \quad (\text{sofern dividiert werden darf}),$$

$$(a^m)^n = a^{mn}, \quad a^{1/n} = \sqrt[n]{a}.$$

Daraus folgt für Zahlen $x = c_1 \cdot 10^n$ und $y = c_2 \cdot 10^m$ in exponentieller Darstellung:

$$xy = c_1 \cdot c_2 \cdot 10^{n+m}, \quad \frac{x}{y} = \frac{c_1}{c_2} \cdot 10^{n-m}.$$

Beachte: Das Ergebnis ist nicht automatisch in exponentieller Standardform!

Beispiel 2.1.2: Für $x = 7.2 \cdot 10^4$ und $y = 2.1 \cdot 10^{-2}$ ist

$$x \cdot y = 7.2 \cdot 2.1 \cdot 10^2 = 15.12 \cdot 10^2 = \underline{1.512 \cdot 10^3}.$$

2.1.3 Potenzieren und Wurzelziehen

Zur Erinnerung: Es ist $(a^m)^n = a^{m \cdot n}$. Damit folgt

$$(x \cdot 10^m)^n = x^n \cdot (10^m)^n = x^n \cdot 10^{m \cdot n}.$$

Das Ergebnis muss nicht exponentielle Standardform haben!

Beispiel 2.1.3: $(4.2 \cdot 10^3)^2 = 4.2^2 \cdot 10^6 = 17.64 \cdot 10^6 = 1.764 \cdot 10^7$.

Die gleiche Regel gilt auch für das Wurzelziehen:

$$\sqrt{4 \cdot 10^6} = (4 \cdot 10^6)^{1/2} = 4^{1/2} \cdot (10^6)^{1/2} = 2 \cdot 10^{6 \cdot \frac{1}{2}} = 2 \cdot 10^3.$$

Dabei kann eine Zehnerpotenz mit gebrochenem Exponenten entstehen!

Beispiel 2.1.4: Es ist $\sqrt{4.2 \cdot 10^5} = \sqrt{4.2} \cdot 10^{5/2}$. Um zur exponentiellen Darstellung des Ergebnisses zu gelangen, muß der ganzzahlige Anteil des Exponenten abgespalten werden:

$$\sqrt{4.2 \cdot 10^5} = \sqrt{4.2} \cdot 10^{2+1/2} = \sqrt{4.2} \cdot \sqrt{10} \cdot 10^2 = \sqrt{42} \cdot 10^2 \approx 6.5 \cdot 10^2.$$

Wichtig: Für $a, b \geq 0$ gilt $\sqrt{ab} = \sqrt{a}\sqrt{b}$. Im allgemeinen gilt aber $\sqrt{a+b} \neq \sqrt{a} + \sqrt{b}$. Die Gleichheit wäre fast immer **falsch**.

Für die Kubikwurzel $\sqrt[3]{a}$ schreibt man auch $a^{1/3}$. Also gilt zum Beispiel $\sqrt[3]{256} = \sqrt[3]{4^4} = 4 \cdot 4^{1/3}$.

2.1.4 Addition und Subtraktion

Es ist klar, wie zwei Zahlen in exponentieller Darstellung addiert (bzw. subtrahiert) werden, falls die Zehnerpotenzen gleiche Exponenten besitzen:

$$c_1 \cdot 10^n + c_2 \cdot 10^n = (c_1 + c_2) \cdot 10^n,$$

etwa

$$7.4 \cdot 10^{-3} + 4.2 \cdot 10^{-3} = 11.6 \cdot 10^{-3} = 1.16 \cdot 10^{-2}.$$

Sind die Exponenten verschieden, so müssen die Zahlen erst umgeformt werden:

$$c_1 \cdot 10^n + c_2 \cdot 10^m = c_1 \cdot 10^n + (c_2 \cdot 10^{m-n}) \cdot 10^n = (c_1 + c_2 \cdot 10^{m-n}) \cdot 10^n.$$

Das Ergebnis muss nicht die exponentielle Standardform haben!

Beispiele 2.1.5: (a) $6.04 \cdot 10^4 + 3.6 \cdot 10^3 = 60.4 \cdot 10^3 + 3.6 \cdot 10^3 = 64 \cdot 10^3 = 6.4 \cdot 10^4$, oder
 $= 6.04 \cdot 10^4 + 0.36 \cdot 10^4 = 6.4 \cdot 10^4$

(b) $6.04 \cdot 10^3 + 3.6 \cdot 10^5 = 6.04 \cdot 10^3 + 360 \cdot 10^3 = 366.04 \cdot 10^3 = 3,6604 \cdot 10^5$

(c) $9.82 \cdot 10^{-4} - 7,1 \cdot 10^{-5} = 9.82 \cdot 10^{-4} - 0.71 \cdot 10^{-4} = 9,11 \cdot 10^{-4}$

2.2 Logarithmen

2.2.1 Definition und erste Eigenschaften

Wir haben bisher folgende Zehnerpotenzen kennengelernt:

$$10^0 = 1, \quad 10^1 = 10, \quad 10^2 = 100, \quad \dots$$

$$10^n = 10 \cdot 10 \cdot \dots \cdot 10 \quad \text{für } n \in \mathbb{N}$$

$$10^{-n} = 0.1 \cdot 0.1 \cdot \dots \cdot 0.1 = \frac{1}{10} \cdot \dots \cdot \frac{1}{10} = \frac{1}{10^n} \quad \text{für } n \in \mathbb{N}$$

$$10^{m/n} = (\sqrt[n]{10})^m = \sqrt[n]{10^m} \quad \text{für } m \in \mathbb{Z}, n \in \mathbb{N}.$$

Man kann 10^x auch für beliebige reelle Zahlen x definieren (als Grenzwert von 10^{r_k} mit einer Folge von Brüchen $r_k \rightarrow x$; der Begriff „Grenzwert“ wird später erklärt).

Definition 2.2.1 (Logarithmus). Es sei $a \in \mathbb{R}$ mit $a > 0$, und es gelte $a = 10^b$ für ein $b \in \mathbb{R}$. Dann wird b der (dekadische) Logarithmus von a genannt und mit $\log a$ bezeichnet.

Bemerkungen 2.2.2: (i) Gelegentlich wird $\log a$ auch $\lg a$ oder $\log_{10} a$ geschrieben.

(ii) Der Logarithmus von a , also $\log a$, ist diejenige reelle Zahl, für die $a = 10^{\log a}$ gilt.

Beispiel 2.2.3:

$$\log 0,001 = -3, \quad \text{denn } 10^{-3} = 0,001$$

$$\vdots \qquad \qquad \qquad \vdots$$

$$\log 10 = 1, \quad \text{denn } 10^1 = 10.$$

Welchen Wert hat dann $\log 2$?

Nach Definition ist $\log 2$ diejenige Zahl $b \in \mathbb{R}$, für die $10^b = 2$ gilt. Wie groß ist b näherungsweise?

Es ist $10^0 = 1$, $10^1 = 10$ und $1 < 2 < 10$, d.h. es gilt

$$10^0 < 10^{\log 2} < 10^1.$$

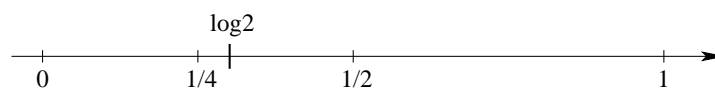
Aus der Monotonie der Potenzfunktion $x \rightarrow 10^x$ (später!) folgt

$$0 < \log 2 < 1.$$

Es ist $\sqrt{10} = 10^{1/2} = 3.16\dots$ und $10^{1/4} = 10^{1/2 \cdot 1/2} = \sqrt{\sqrt{10}} = \sqrt{3.16\dots} = 1.778\dots$. Damit gilt genauer:

$$10^{1/4} < 10^{\log 2} < 10^{1/2}, \quad \text{also } \underline{0,25 < \log 2 < 0,5}.$$

Schachtelt man $\log 2$ auf diese Weise immer weiter ein, so erhält man $\log 2 = 0.301029\dots$, eine irrationale Zahl.



Was ist dann $\log 20$, oder $\log(2 \cdot 10^n)$? Zur Erinnerung: Aus den Potenzgesetzen ergeben sich für $a, b > 0$ und $x \in \mathbb{R}$ die Regeln

$$\begin{aligned}\log(ab) &= \log a + \log b, \\ \log \frac{a}{b} &= \log a - \log b, \\ \log a^x &= x \log a.\end{aligned}$$

Damit folgt beispielsweise

$$\log 20 = \log(2 \cdot 10) = \log 2 + \log 10 = \log 2 + 1 = 1.301029 \dots$$

und

$$\log(2 \cdot 10^n) = \log 2 + n.$$

Allgemein gilt:

$$\log(c \cdot 10^n) = \log c + n.$$

2.2.2 Graphische Darstellung des Logarithmus

Zunächst einige Fakten:

- (i) $\log a$ existiert nur für $a > 0$, denn $10^x > 0$ für jedes $x \in \mathbb{R}$.
- (ii) $\log 1 = 0$, denn $10^0 = 1$
- (iii) $\log x$ ist *streng monoton wachsend*, d.h. $0 < a < b \Rightarrow \log a < \log b$.
Dies folgt aus der entsprechenden Monotonie von $x \mapsto 10^x$:

$$0 < a < b \Rightarrow 10^{\log a} < 10^{\log b}, \text{ also } \log a < \log b, \text{ sonst Widerspruch.}$$

- (iv) Für $a \in (1, 10)$ gilt $\log a \in (0, 1)$.

- (v) Für $a \in (0, 1)$ gilt $\log a < 0$, denn

$$\log a = -\log a^{-1} = -\log \frac{1}{a} \text{ und } \frac{1}{a} > 1 \Rightarrow \log \frac{1}{a} > 0 \Rightarrow -\log \frac{1}{a} < 0.$$

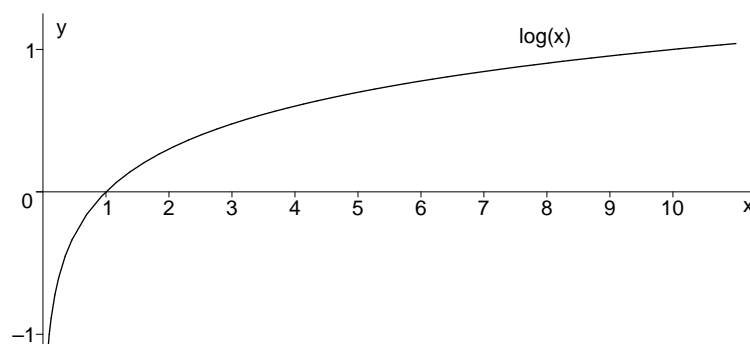
- (vi) $\log a$ nimmt beliebig große Werte für große a an:

$$a = 10^n \text{ mit } n \in \mathbb{N} \Rightarrow \log a = n.$$

Also: $\log a$ ist nach oben unbeschränkt.

- (vii) $\log a$ ist auch nach unten unbeschränkt: $a = \frac{1}{10^n} = 10^{-n} \Rightarrow \log a = -n$.

Der Graph von $\log x$ hat folgenden Verlauf:



2.2.3 Der natürliche Logarithmus

Wir haben oben $\log a$ als diejenige Zahl definiert, für die $10^{\log a} = a$ gilt. Anders ausgedrückt: Bei vorgegebenem $a > 0$ ist $x = \log a$ die Lösung der Gleichung $10^x = a$, d.h. der (dekadische) Logarithmus ist die Umkehrfunktion (genauer Begriff später) der Potenzfunktion $x \rightarrow 10^x$.

Bei vielen (Wachstums-) Prozessen in der Natur spielt eine andere Potenzfunktion eine herausragende Rolle, nämlich die sogenannte Exponentialfunktion e^x . Dabei wird als Basis die sogenannte Eulersche Zahl $e = 2.71828\dots$ verwendet. Die genaue Definition dieser irrationalen Zahl kann erst später gegeben werden. Zur Motivation:

Beispiel 2.2.4: Eine Population (Bevölkerung; auch von Tieren, Pflanzen, Bakterien etc.) bestehe zur Zeit $t = 0$ aus N_0 und zur Zeit $t > 0$ aus $N(t)$ Individuen. Angenommen, die Wachstumsrate ($\frac{dN}{dt} \approx \frac{\Delta N}{\Delta t}$) sei proportional zur Bevölkerungszahl, also $\frac{dN(t)}{dt} = kN(t)$ mit einer Wachstumskonstante $k > 0$. Wie lässt sich dann $N(t)$ zur Zeit $t > 0$ aus N_0 berechnen?

Da wir die sogenannte Differentialgleichung $\frac{dN(t)}{dt} = kN(t)$ für $t > 0$, $N(0) = N_0$ mit unserem jetzigen Wissen nicht lösen können, versuchen wir folgende Näherung. Wir wählen ein großes $n \in \mathbb{N}$ und setzen $\Delta t := t/n$. Dann gilt näherungsweise

$$\frac{N(\Delta t) - N(0)}{\Delta t} = \frac{\Delta N}{\Delta t} \approx kN(0),$$

also

$$N(\Delta t) = N(0) + kN(0)\Delta t \approx N_0(1 + k\Delta t).$$

Entsprechend gilt

$$\frac{N(2\Delta t) - N(\Delta t)}{\Delta t} \approx kN(\Delta t),$$

also $N(2\Delta t) \approx N(\Delta t)(1 + k\Delta t) \approx N_0 \cdot (1 + k\Delta t)^2$.

Nach insgesamt n Schritten erhält man so

$$N(n\Delta t) \approx N_0(1 + k\Delta t)^n, \text{ also } N(t) \approx N_0 \left(1 + \frac{kt}{n}\right)^n.$$

Für den speziellen Zeitpunkt $t = 1/k$ ergibt sich

$$N(1/k) \approx N_0(1 + 1/n)^n,$$

wobei wir mit wachsendem n eine bessere Näherung erwarten. Es gilt:

n	1	2	3	...	10	...	100	...	„∞“
$(1 + \frac{1}{n})^n$	2	2,25	2,37	...	2,594	...	2,705	...	e

D.h. $(1 + \frac{1}{n})^n$ strebt gegen die Eulersche Zahl $e = 2,71828\dots$, wenn n immer größer wird. Also gilt $N(1/k) = N_0 \cdot e$.

Allgemein ist $N(t) = N_0 \cdot e^{kt}$ für $t \geq 0$, denn wie $(1 + \frac{1}{n})^n$ strebt auch $(1 + \frac{kt}{n})^{\frac{n}{kt}}$ gegen e . Damit gilt

$$\left(1 + \frac{kt}{n}\right)^n = \left[\left(1 + \frac{kt}{n}\right)^{\frac{n}{kt}}\right]^{kt} \rightarrow e^{kt}.$$

Die Umkehrfunktion zur Exponentialfunktion e^x wird natürlicher Logarithmus genannt und mit $\ln x$ bezeichnet (Logarithmus naturalis). Genauer:

Definition 2.2.5 (Natürlicher Logarithmus). Sei $a > 0$. Der *natürliche Logarithmus* von a , kurz $\ln a$, ist diejenige reelle Zahl, für die $e^{\ln a} = a$ gilt. Dabei ist $e = 2.71828\dots$ die Eulersche Zahl.

Insbesondere gilt $e^{\ln 10} = 10$ und damit auch

$$e^{x \ln 10} = (e^{\ln 10})^x = 10^x.$$

Hieraus folgt nach Definition 2.2.1 und Definition 2.2.5:

$$\begin{aligned} 10^x &= a \Rightarrow x = \log a \\ e^{x \ln 10} &= a \Rightarrow x \ln 10 = \ln a. \end{aligned}$$

Also gilt folgende *Umrechnungsformel* zwischen $\ln a$ und $\log a$:

$$\boxed{\ln a = \ln 10 \cdot \log a} \quad \text{mit } \ln 10 \approx 2.303.$$

Beispiel 2.2.6: $\ln(9.831 \cdot 10^{-2}) \approx ?$ Es ist

$$\log(9.831 \cdot 10^{-2}) = \log 9.831 - 2 \approx 0.9926 - 2 = -1.0074$$

und damit

$$\ln(9.831 \cdot 10^{-2}) \approx 2.303 \cdot (-1.0074) \approx -2.320.$$

2.3 Algebraische Gleichungen

Unter einer algebraischen Gleichung versteht man eine Gleichung in einer oder mehreren Unbekannten (etwa x, y, z, \dots), bei der jede Variable stets als Potenz mit natürlichem Exponenten auftritt.

Beispiele 2.3.1: (a) $x^3 + 2x^2 - 4x = 7 + 2x$

$$(b) \quad x^2 + 2xy = y^2$$

Viele Anwendungen führen auf mehrere Gleichungen mit mehreren Variablen, d.h. auf ein sogenanntes Gleichungssystem, etwa (a) und (b) zusammen. Wie löst man solche Gleichungssysteme?

Eine algebraische Gleichung in einer Variablen läßt sich immer auf die Form

$$P(x) = 0$$

bringen mit einem sogenannten *Polynom*

$$P(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 = \sum_{k=0}^n a_k x^k, \quad a_n \neq 0.$$

Hier **vereinbaren** wir, dass $x^0 = 1$ für **alle** x .

Da die Gleichung $P(x) = 0$ durch a_n dividiert werden kann, reicht es aus, Polynome mit führendem Koeffizienten $a_n = 1$ zu betrachten. Der höchste Exponent n heißt der Grad des Polynoms.

Beispiel 2.3.2: $P(x) = x^3 + 2x^2 - 6x - 7$ ist ein Polynom 3. Grades (kubisches Polynom).

Bemerkung 2.3.3. Eine gegebene Gleichung in einer Variable läßt sich evtl. erst durch Umformen in diese Form bringen.

Beispiele 2.3.4: (a) $\frac{5}{2x-7} = \frac{2}{x+2} \Leftrightarrow \cdots \Leftrightarrow x+24=0,$

(b) $\frac{2x}{x-1} = x+1 \Leftrightarrow \cdots \Leftrightarrow x^2 - 2x - 1 = 0$ (quadratische Gleichung).

Die Lösungen von $P(x) = 0$ sind die sogenannten Nullstellen des Polynoms P . Das Beispiel $P(x) = x^2 + 1$ zeigt, daß ein Polynom *keine reellen Nullstellen* haben muß. Dieser Defekt läßt sich beheben, indem man sogenannte komplexe Zahlen $z = x + iy$ mit $i^2 = -1$ einführt; genaueres dazu später.

Die Berechenbarkeit der Lösungen von $P(x) = 0$ hängt vom Grad des Polynoms ab.

Der Fall Polynomgrad $n = 1$.

Dann ist $P(x) = a_1 x + a_0$ mit $a_1 \neq 0$ und daher $P(x) = 0 \Leftrightarrow x = -\frac{a_0}{a_1}$.

Der Fall Polynomgrad $n = 2$.

Nach Normierung gilt $P(x) = x^2 + px + q$. Die Lösung der Gleichung $P(x) = 0$ erfolgt durch quadratische Ergänzung:

$$x^2 + 2x \frac{p}{2} + q = 0 \Leftrightarrow \left(x + \frac{p}{2}\right)^2 - \left(\frac{p}{2}\right)^2 + q = 0 \Leftrightarrow \left(x + \frac{p}{2}\right)^2 = \frac{p^2}{4} - q.$$

Die Zahl $D := \frac{p^2}{4} - q$ heißt die *Diskriminante*, da das Vorzeichen von D über die reelle Lösbarkeit entscheidet.

Für $D = 0$ gibt es nur die Lösung $x = -p/2$. Für $D > 0$ gilt:

$$\left|x + \frac{p}{2}\right| = \sqrt{\frac{p^2}{4} - q};$$

also gibt es für $D > 0$ genau zwei Lösungen

$$x_1 = -\frac{p}{2} + \sqrt{\frac{p^2}{4} - q} \quad \text{und} \quad x_2 = -\frac{p}{2} - \sqrt{\frac{p^2}{4} - q},$$

oder kurz

$$\boxed{x_{1,2} = -\frac{p}{2} \pm \sqrt{\frac{p^2}{4} - q}} \quad (pq\text{-Formel}).$$

Für Polynome dritten und vierten Grades existieren ebenfalls Formeln zur Berechnung der Nullstellen. Diese sind sehr kompliziert und werden hier nicht angegeben. Für $n \geq 5$ gibt es keine allgemeingültigen exakten Lösungsformeln mehr!

Manchmal helfen folgende Überlegungen weiter:

- (i) Wenn das Polynom nur gerade Potenzen von x enthält, kann man x^2 durch y substituieren.

Beispiel 2.3.5: $x^4 - 3x^2 + 2 = 0$.

$y := x^2$ liefert die quadratische Gleichung $y^2 - 3y + 2 = 0$ mit den Lösungen

$$y = \frac{3}{2} \pm \sqrt{\frac{9}{4} - 2} = \frac{3}{2} \pm \sqrt{\frac{1}{4}} = \frac{3}{2} \pm \frac{1}{2},$$

also $y_1 = 1$ und $y_2 = 2$. Rücksubstitution ergibt

$$x^2 = 1 \text{ oder } x^2 = 2, \text{ also } |x| = 1 \text{ oder } |x| = \sqrt{2}.$$

Damit sind $x = \pm 1$ und $x = \pm\sqrt{2}$ alle Lösungen der Aufgabe.

- (ii) Wenn eine Nullstelle x_0 bekannt ist, führt eine Polynomdivision durch $x - x_0$ zum Abspalten des entsprechenden Linearfaktors und damit zur Verringerung des Grades des Polynoms.

Beispiel 2.3.6: Sei $P(x) = x^3 + x^2 - x - 1$. Offensichtlich ist $x = 1$ eine Lösung der Gleichung $P(x) = 0$. Eine Polynomdivision durch $x - 1$ liefert dann $P(x) = (x - 1)Q(x)$ mit einem Polynom Q , dessen Grad um 1 kleiner ist als der von P . Konkret: $P(x) = (x - 1)(x^2 + 2x + 1)$, und $Q(x) = x^2 + 2x + 1 = (x + 1)^2$ hat die doppelte Nullstelle -1 .

Beispiel 2.3.7: Gelegentlich liegt ein System von nicht algebraischen Gleichungen vor, die in algebraische Gleichungen überführbar sind, etwa

$$\begin{aligned} \sqrt{x} + \sqrt{x+1} &= y \\ 1 - 2x &= y^2. \end{aligned}$$

Wir setzen ein:

$$\begin{aligned} 1 - 2x &= (\sqrt{x} + \sqrt{x+1})^2 \Leftrightarrow 1 - 2x = x + 2\sqrt{x}\sqrt{x+1} + x + 1 \\ \Leftrightarrow -4x &= 2\sqrt{x(x+1)} \quad \Leftrightarrow -2x = \sqrt{x(x+1)} \\ \Rightarrow 4x^2 &= x^2 + x \quad \Leftrightarrow 3x^2 - x = 0 \quad \Leftrightarrow x(3x - 1) = 0 \\ \Leftrightarrow x &= 0 \text{ oder } x = \frac{1}{3}. \end{aligned}$$

Vermutung: Die Lösungen sind $(x, y) = (0, 1)$ und $(x, y) = \left(\frac{1}{3}, \sqrt{\frac{1}{3}} + \sqrt{\frac{4}{3}}\right) = \left(\frac{1}{3}, \sqrt{3}\right)$.

Achtung! Wir haben Umformungen benutzt, die *keine* Äquivalenzumformungen sind! Eine Probe ist daher erforderlich, und sie zeigt: $(0, 1)$ ist eine Lösung, $\left(\frac{1}{3}, \sqrt{3}\right)$ dagegen *nicht*.

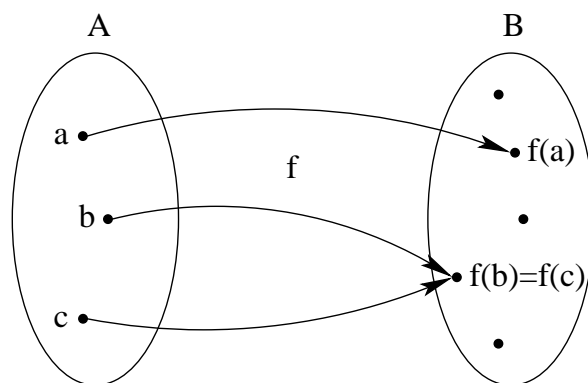
Es gibt keine vollständige Lösungstheorie für allgemeine algebraische Gleichungssysteme. Man beachte auch, dass bereits die Lösung scheinbar einfacher Gleichungen wie $x + e^x = 0$ „elementar“ nicht berechenbar ist; zur numerischen Lösung solcher „transzendenten“ Gleichungen später mehr.

Kapitel 3

Elementare Funktionen

3.1 Der mathematische Funktionsbegriff

Anschaulich versteht man unter einer Funktion oder Abbildung eine Zuordnung zwischen gewissen Objekten. Wir werden zur Formulierung des Funktionsbegriffes auf Mengen zurückgreifen, Funktionen also als Zuordnungen zwischen Elementen von Mengen verstehen.



Entscheidend für den Begriff der Funktion ist die Eindeutigkeit der Abbildung, d.h. jedem Objekt darf nur genau ein Objekt zugeordnet werden. Als Beispiel haben wir schon die Wurzelfunktion kennen gelernt: für die Wurzel lässt man nur nichtnegative Zahlen zu, obwohl z.B. ja auch $(-1)^2 = 1$ gilt. Offenbar wäre die Zuordnung $\sqrt{1} = 1$ und zugleich $\sqrt{1} = -1$ wenig sinnvoll; wie sollte man z.B. die Gleichheit $\sqrt{1} + \sqrt{1} = 2\sqrt{1}$ verstehen?

In naturwissenschaftlichen Anwendungen sind meist Abbildungen zwischen reellen Zahlen von Interesse.

Beispiele 3.1.1: Einige typische Zuordnungsvorschriften:

1. $f(x) = x^2$,
2. $g(x) = \sqrt{x+1}$,
3. $N(t) = N_0 e^{kt}$ (siehe Abschnitt 2.2),
4. $y = ax + b$.

Das letzte Beispiel ist ungenau formuliert; man kann ohne Zusammenhang nur vermuten, dass x die sogenannte unabhängige Variable ist, d.h. $f(x) = ax + b$ gemeint ist.

Definition 3.1.2 (Funktion). Gegeben seien Mengen A und B . Unter einer *Funktion* (oder auch *Abbildung*)

$$f : A \rightarrow B, \quad x \mapsto f(x)$$

versteht man eine Vorschrift, die *jedem* Element $x \in A$ *genau ein* Element $f(x) \in B$ zuordnet. Eine konkrete Funktion kann mit beispielsweise mit Hilfe einer Formel oder durch Aufzählen aller zugeordneten Paare angegeben werden.

Beispiel 3.1.3: Es seien $A = \{a, b, c\}$ und $B = \{\alpha, \beta, \gamma\}$. Dann beschreiben die Formulierungen

- $f : A \rightarrow B$ mit $f(a) = \alpha, f(b) = \beta, f(c) = \gamma$;
- $f : A \rightarrow B, \quad a \mapsto \alpha, b \mapsto \beta, c \mapsto \gamma$;
- $f : A \rightarrow B$ ist die Abbildung, die einem lateinischen Buchstaben im ABC den griechischen Buchstaben mit gleicher Position im Alphabet zuordnet;

alle die gleiche Funktion.

Bemerkung 3.1.4. Sei $f : A \rightarrow B$ eine Funktion mit der Zuordnungsvorschrift $y = f(x)$. Dann heißt x *Argument* oder *unabhängige Variable*, y *abhängige Variable*, $f(x)$ *Funktionswert* bzw. Wert von f an der Stelle x und A der *Definitionsbereich* von f . Für den Definitionsbereich schreibt man auch $D_f = A$.

Weiter unterscheidet man zwischen der *Bildmenge* (bzw. dem *Bildbereich*) B und dem *Wertebereich* $W_f \subseteq B$ von f . W_f ist die Menge aller Funktionswerte $f(x)$ mit $x \in D_f$, d.h. $W_f = \{f(x) : x \in D_f\} =: f(D_f)$.

Mitunter ist D_f nicht explizit angegeben; dann verwendet man den sogenannten „maximalen Definitionsbereich“ (gelegentlich als D_{\max} bezeichnet), d.h. die größte Menge, für welche die Zuordnungsvorschrift definiert ist.

Beispiele 3.1.5: (a) $f(x) = \frac{1}{x^2 - 1}$ hat den maximalen Definitionsbereich $D_f = \mathbb{R} \setminus \{-1, 1\}$.

(b) $g(x) = \sqrt{x+1}$ hat den maximalen Definitionsbereich $D_g = [-1, \infty)$.

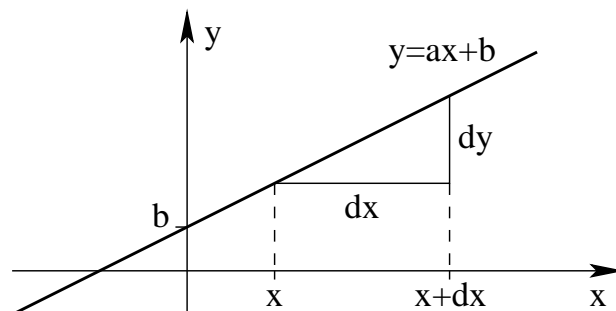
Eine Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ kann veranschaulicht werden mittels ihres (Funktions-)Graphen. Allgemein definiert man:

Definition 3.1.6 (Graph). Der *Graph* einer Funktion $f : A \rightarrow B$ ist die Menge

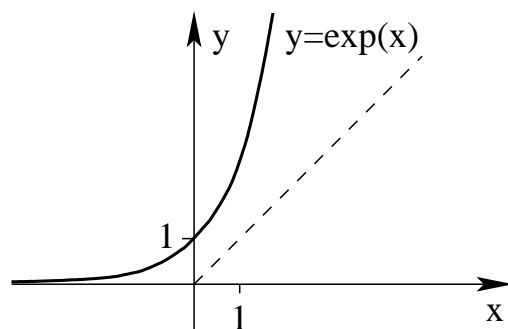
$$\text{gr}(f) := \{(x, y) \in A \times B : x \in A, y = f(x)\}.$$

Im Fall $A, B \subset \mathbb{R}$ ist der Graph von $f : A \rightarrow B$ eine Teilmenge der Ebene \mathbb{R}^2 und lässt sich daher bildlich darstellen.

Beispiele 3.1.7: Beispielhaft betrachten wir eine lineare Funktion und die Exponentialfunktion.



$$f : \mathbb{R} \rightarrow \mathbb{R} \quad \text{mit } f(x) = ax + b. \quad W_f = \mathbb{R}, \text{ falls } a \neq 0$$



$$f : \mathbb{R} \rightarrow \mathbb{R} \quad \text{mit } f(x) = e^x. \quad W_f = (0, \infty).$$

3.2 Trigonometrische Funktionen

Wir messen Winkel im *Bogenmaß*. Ist α ein Winkel in Gradmaß (also 360° für den Vollkreis), so ist seine Größe x im Bogenmaß (also 2π für den Vollkreis) gegeben durch

$$x = \alpha \frac{2\pi}{360^\circ}; \quad \text{etwa } 30^\circ = \frac{\pi}{6}.$$

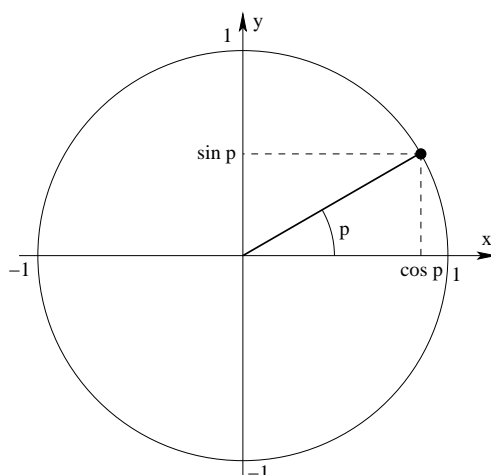
Definition 3.2.1 (Sinus, Cosinus). Es seien x, y die Koordinaten des Punktes Z auf dem Einheitskreis, auf den der Punkt $(1, 0)$ durch Drehung um den Winkel p (Winkel im positiven Drehsinn, also gegen den Uhrzeigersinn) abgebildet wird. Dann definieren wir:

$$\sin p := y \quad \text{und} \quad \cos p := x.$$

Dadurch sind $\sin p$ und $\cos p$ für Winkel $p \in [0, 2\pi)$ erklärt. Für andere Werte $p \in \mathbb{R}$ definieren wir

$$\sin p := \sin(p - 2k\pi) \quad \text{und} \quad \cos p := \cos(p - 2k\pi),$$

wobei $k \in \mathbb{Z}$ so gewählt ist, dass $p - 2k\pi \in [0, 2\pi)$ gilt.



Bemerkung 3.2.2. Die Funktionen $\sin x$ und $\cos x$, $x \in \mathbb{R}$, entstehen also durch 2π -periodische Fortsetzung der Verläufe der Funktionen im Bereich $[0, 2\pi) \subset \mathbb{R}$.

Eigenschaften von $\sin x$ und $\cos x$. Unmittelbar aus der geometrischen Definition folgen

1. $(\sin x)^2 + (\cos x)^2 = 1 \quad \forall x \in \mathbb{R}$ (Satz des Pythagoras; man schreibt oft $\sin^2 x$ bzw. $\cos^2 x$ statt $(\sin x)^2$ bzw. $(\cos x)^2$),
2. $\sin 0 = 0, \quad \cos 0 = 1,$
3. $\sin \frac{\pi}{2} = 1, \quad \cos \frac{\pi}{2} = 0,$
4. $\sin \frac{\pi}{4} = \cos \frac{\pi}{4} = \frac{1}{\sqrt{2}},$
5. $\sin \frac{\pi}{6} = \frac{1}{2}, \quad \cos \frac{\pi}{6} = \frac{\sqrt{3}}{2}.$

Desweiteren gelten die sogenannten **Additionstheoreme**¹:

$$\begin{aligned} \sin(x+y) &= \sin x \cos y + \cos x \sin y \quad \forall x, y \in \mathbb{R}, \\ \cos(x+y) &= \cos x \cos y - \sin x \sin y \quad \forall x, y \in \mathbb{R}. \end{aligned}$$

Mithilfe der Additionstheoreme (oder durch einfache geometrische Überlegungen) erhält man weiter:

$$\begin{aligned} \sin(x + \pi/2) &= \cos x, \quad \cos(x + \pi/2) = -\sin x, \\ \sin(x + \pi) &= -\sin x, \quad \cos(x + \pi) = -\cos x, \\ \sin 2x &= 2 \sin x \cos x, \\ \cos 2x &= \cos^2 x - \sin^2 x \quad (= 2 \cos^2 x - 1 = 1 - 2 \sin^2 x). \end{aligned}$$

In Anwendungen verwendet man die Winkelfunktionen oft zur Längenberechnung. Dabei benutzt man folgende Zusammenhänge²:

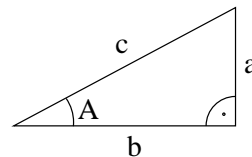
¹Der Beweis der Additionstheoreme ist eine einfache Übung, sobald man die komplexe Exponentialfunktion zu Hilfe nimmt, s. Abschnitt 3.4

²Wir leiten diese Beziehungen nicht her, eine leicht zugängliche Darstellung findet sich z.B. in RiesBinger/Walz/Zeifelder, Brückenkurs Mathematik, S. 143 ff

1. in rechtwinkligen Dreiecken gilt

$$\sin A = \frac{a}{c} = \frac{\text{Gegenkathete}}{\text{Hypotenuse}}$$

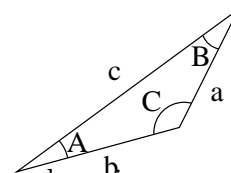
$$\cos A = \frac{b}{c} = \frac{\text{Ankathete}}{\text{Hypotenuse}}$$



2. in beliebigen Dreiecken gilt

$$\text{Sinussatz: } \frac{\sin A}{a} = \frac{\sin B}{b} = \frac{\sin C}{c}.$$

$$\text{Cosinussatz: } c^2 = a^2 + b^2 - 2ab \cos C$$



Aus den grundlegenden trigonometrische Funktionen \sin und \cos werden wir nun weitere Funktionen erzeugen.

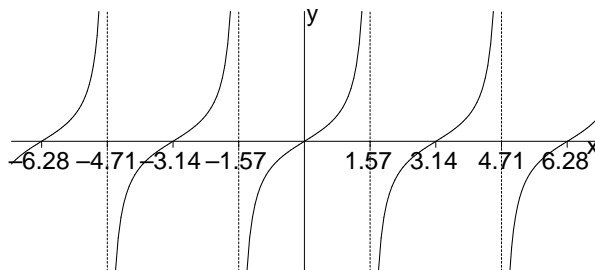
Definition 3.2.3 (Tangens, Cotangens). Mittels Sinus und Cosinus definiert man

$$\tan x := \frac{\sin x}{\cos x} \quad \text{für } x \in \mathbb{R} \setminus \left\{ \frac{\pi}{2} + n\pi : n \in \mathbb{Z} \right\} \quad (\text{Tangens})$$

und

$$\cot x := \frac{\cos x}{\sin x} \quad \text{für } x \in \mathbb{R} \setminus \{n\pi : n \in \mathbb{Z}\} \quad (\text{Cotangens}).$$

Die Tangensfunktion ist π -periodisch, ungerade und im Intervall $(-\pi/2, \pi/2)$ streng monoton wachsend. Der Tangens hat aber Polstellen in allen Punkten der Form $\pi/2 + \mathbb{Z}$; an diesen Punkten hat die Funktion Definitionslücken.



Der Graph des Cotangens ergibt sich aus dem Zusammenhang

$$\tan(x + \pi/2) = \frac{\sin(x + \pi/2)}{\cos(x + \pi/2)} = \frac{\cos x}{-\sin x} = -\cot x.$$

3.3 Umkehrbarkeit und weitere elementare Eigenschaften von Funktionen

Für viele Anwendungen benötigt man die Umkehrfunktionen elementarer Funktionen wie $\sin x$, $\cos x$ oder e^x ; etwa, wenn man bei gegebenem Sinuswert wissen möchte, wie der ursprüngliche Winkelwert war. Wir haben auch schon einige Umkehrfunktionen kennen gelernt, z.B. ist die Funktion \sqrt{x} die Umkehrfunktion zu x^2 auf $[0, \infty)$, und $\ln x$ auf $(0, \infty)$ ist die Umkehrfunktion zu e^x auf \mathbb{R} .

Eine Umkehrfunktion existiert aber nicht immer, und wenn sie existiert, dann oft nur auf eingeschränkten Gebieten. Um die Frage nach der Umkehrbarkeit zu klären, müssen wir einige vorbereitende Begriffe einführen.

Definition 3.3.1. Eine Funktion $f : D_f \rightarrow B$ heißt

- *injektiv* oder *eindeutig*, falls aus $x \neq \tilde{x}$ auch $f(x) \neq f(\tilde{x})$ folgt,
- *surjektiv* oder *Abbildung auf B* , falls $W_f = B$ gilt,
- *bijektiv*, falls f injektiv und surjektiv ist.

Bemerkungen 3.3.2: (i) $f : D_f \rightarrow B$ ist injektiv, falls gilt: $f(x) = f(\tilde{x}) \Rightarrow x = \tilde{x}$.

(ii) $f : D_f \rightarrow W_f$ ist stets surjektiv, denn $W_f = f(D_f)$ nach Definition.

Diese Begriffe sind deshalb wichtig, weil jede bijektive Funktion umkehrbar ist!

Definition 3.3.3. Sei $f : D_f \rightarrow W_f$ eine injektive Funktion (sie ist offenbar auch surjektiv). Die Funktion $g : W_f \rightarrow D_f$, die jedem $y \in W_f$ genau das $x \in D_f$ zuordnet, für welches $y = f(x)$ gilt, heißt *Umkehrfunktion* von f ; geschrieben als $g = f^{-1}$.

Bemerkungen 3.3.4: (i) Es gilt also $f(f^{-1}(y)) = y$ für alle $y \in W_f$ sowie $f^{-1}(f(x)) = x$ für alle $x \in D_f$.

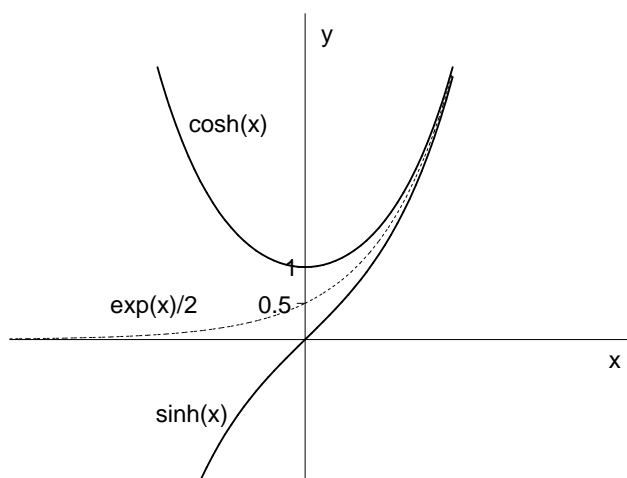
(ii) Die Zuordnungsvorschrift für f^{-1} lässt sich oft durch Auflösen der Gleichung $y = f(x)$ nach x angeben.

Beispiel 3.3.5: Die hyperbolischen Funktionen sind wie folgt erklärt:

$$\text{Sinus hyperbolicus: } \sinh x := \frac{e^x - e^{-x}}{2} \quad \text{für } x \in \mathbb{R},$$

$$\text{Cosinus hyperbolicus: } \cosh x := \frac{e^x + e^{-x}}{2} \quad \text{für } x \in \mathbb{R}.$$

Der Name kommt von der Beziehung $(\cosh x)^2 - (\sinh x)^2 = 1$ (die man leicht nachrechnet), denn $y^2 - x^2 = 1$ ist die Gleichung für eine Hyperbel. Zum Vergleich: bei den trigonometrischen Funktionen gilt stattdessen $\cos^2(x) + \sin^2(x) = 1$, und $x^2 + y^2 = 1$ ist die Gleichung eines Kreises. Daher werden die trigonometrischen Funktionen auch oft als Kreisfunktionen bezeichnet.



Wir betrachten den hyperbolischen Cosinus $f : \mathbb{R} \rightarrow \mathbb{R}$ mit $f(x) = \cosh x$. Dann gilt $W_f = [1, \infty)$. Schon am Funktionsgraphen kann man leicht sehen, dass $f : \mathbb{R} \rightarrow W_f$ *nicht* injektiv ist, denn es ist $f(x) = f(-x)$ für alle $x \in \mathbb{R}$. Ähnlich wie bei $y = x^2$ kann nur nach Einschränkung der Funktion auf den kleineren Definitionsbereich

$$D_f = \mathbb{R}_+ := [0, \infty)$$

invertiert werden. Mit anderen Worten:

$$f : D_f = \mathbb{R}_+ \rightarrow [1, \infty) = W_f, \quad x \rightarrow \cosh x \quad \text{ist bijektiv.}$$

Wir bestimmen die Abbildungsvorschrift von f^{-1} , indem wir $y = \frac{e^x + e^{-x}}{2}$ nach x auflösen. Wir setzen $z = e^x$ und erhalten $z + \frac{1}{z} = 2y$ mit $z > 0$. Dies ist der Reihe nach äquivalent zu

$$z^2 + 1 = 2zy \quad \Leftrightarrow \quad z^2 - 2zy + 1 = 0 \quad \Leftrightarrow \quad (z - y)^2 = y^2 - 1 \quad \Leftrightarrow \quad |z - y| = \sqrt{y^2 - 1}.$$

Zunächst erhalten wir also zwei Lösungen: $z = y + \sqrt{y^2 - 1}$ und $z = y - \sqrt{y^2 - 1}$. Welche dieser Lösungen liefert nun f^{-1} ? Wir haben $z = e^x$ mit $x \in D_f$, also $x \geq 0$ und damit $z \geq 1$. Nun ist $y - \sqrt{y^2 - 1} < 1$ für $y \in W_f$. Somit gilt

$$e^x = z = y + \sqrt{y^2 - 1} \quad \Rightarrow \quad f^{-1}(y) = \ln \left(y + \sqrt{y^2 - 1} \right).$$

Bezeichnung: $\operatorname{Arcosh} x := \ln \left(x + \sqrt{x^2 - 1} \right)$ für $x \geq 1$ *Areacosinushyperbolicus*.

Beispiel 3.3.6: Wie betrachten die durch $f(x) = x^2 - 4x + 1 = (x - 2)^2 - 3$ gegebene Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$.

Wegen $W_f = [-3, \infty)$ ist f *nicht* surjektiv, und auch $f : \mathbb{R} \rightarrow W_f$ ist *nicht* injektiv.

Wie bei $y = x^2$ kann f nur durch Einschränkung auf einen kleineren Definitionsbereich, etwa auf $[2, \infty)$, invertiert werden. Sei also $D_f := [2, \infty)$. Dann ist

$$f : D_f \rightarrow W_f = [-3, \infty) \text{ bijektiv.}$$

Wir berechnen f^{-1} , indem wir $y = x^2 - 4x + 1$ für $y \geq -3$ nach x auflösen. Wegen

$$y = (x - 2)^2 - 3 \Leftrightarrow (x - 2)^2 = y + 3$$

erhalten wir zunächst zwei Lösungen:

$$x = 2 - \sqrt{y + 3} \quad \text{und} \quad x = 2 + \sqrt{y + 3}.$$

Nun muss aber $x \in D_f$ (also $x \geq 2$) gelten. Daher bleibt nur

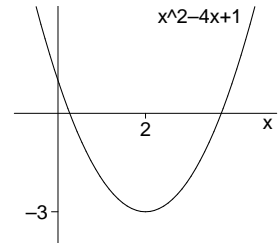
$$x = 2 + \sqrt{y + 3}$$

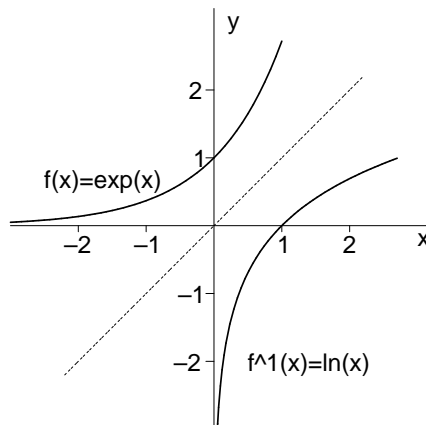
als Lösung. Fazit: Es ist $f^{-1} : [-3, \infty) \rightarrow [2, \infty)$ mit $f^{-1}(x) = 2 + \sqrt{x + 3}$.

Eine geometrische Interpretation der Umkehrfunktion liefert uns eine Betrachtung ihres Graphen:

$$\begin{aligned} \operatorname{gr}(f^{-1}) &= \{(y, x) : y \in W_f, x = f^{-1}(y)\} = \{(y, x) : y \in W_f, y = f(x)\} \\ &= \{(y, x) : (x, y) \in \operatorname{gr}(f)\}, \end{aligned}$$

d.h. $\operatorname{gr}(f^{-1})$ entsteht aus $\operatorname{gr}(f)$ durch Spiegelung an der Winkelhalbierenden zwischen den Koordinatenachsen.



Beispiel 3.3.7:

Viele der in den Anwendungen auftretenden Funktionen haben zusätzliche Eigenschaften, etwa Monotonie oder Symmetrie.

Definition 3.3.8. Eine Funktion $f : D_f \subseteq \mathbb{R} \rightarrow \mathbb{R}$ heißt

- (i) *beschränkt*, falls $|f(x)| \leq M \forall x \in D_f$ mit einer „Schranke“ $M > 0$ gilt.
- (ii) *gerade*, falls $f(-x) = f(x) \forall x \in D_f$ gilt.
- (iii) *ungerade*, falls $f(-x) = -f(x) \forall x \in D_f$ gilt.
- (iv) *(streng) monoton wachsend*, falls gilt:

$$x_1 < x_2 \Rightarrow f(x_1) \leq f(x_2) \quad (\text{bzw. } f(x_1) < f(x_2)).$$

- (v) *(streng) monoton fallend*, falls gilt:

$$x_1 < x_2 \Rightarrow f(x_1) \geq f(x_2) \quad (\text{bzw. } f(x_1) > f(x_2)).$$

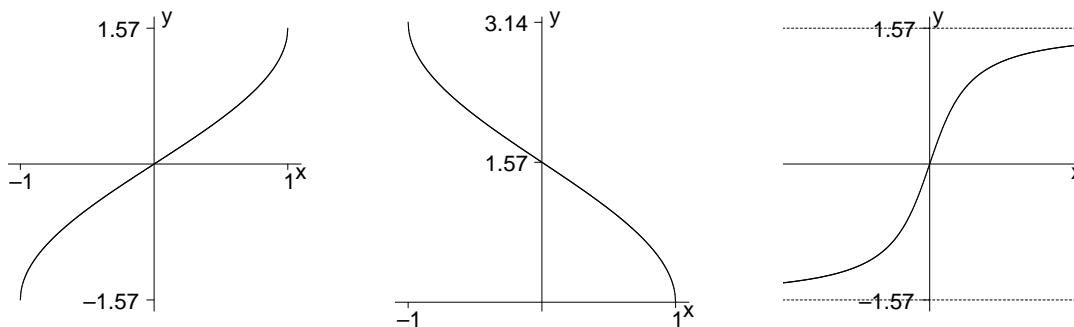
- (vi) *periodisch* (mit der Periodenlänge L), falls $f(x + L) = f(x) \forall x \in D_f = \mathbb{R}$.

Bemerkungen 3.3.9: (i) Eine Funktion $f : \mathbb{R} \supset D_f \rightarrow \mathbb{R}$ heißt (streng) monoton, falls f (streng) monoton wachsend oder (streng) monoton fallend ist.

- (ii) Jede streng monotone Funktion $f : \mathbb{R} \supset D_f \rightarrow W_f \subset \mathbb{R}$ ist automatisch injektiv, also sogar umkehrbar.

Zu Winkelberechnungen benötigt man oft Umkehrfunktionen der trigonometrischen Funktionen. Da letztere nicht injektiv sind, kann man die Inversen nur nach Einschränkung auf einen geeigneten Definitionsbereich erklären.

1. Die Arcussinusfunktion $y = \arcsin x$ ist die Umkehrfunktion von $f(x) = \sin x$ im Monotonieintervall $-\frac{\pi}{2} \leq x \leq \frac{\pi}{2}$.
2. Die Arcuscosinusfunktion $y = \arccos x$ ist die Umkehrfunktion von $f(x) = \cos x$ im Monotonieintervall $0 \leq x \leq \pi$.
3. Die Arcustangensfunktion $y = \arctan x$ ist die Umkehrfunktion von $f(x) = \tan x$ im Monotonieintervall $-\frac{\pi}{2} < x < \frac{\pi}{2}$.



Die Graphen von Arcussinus, Arcuscosinus und Arcustangens

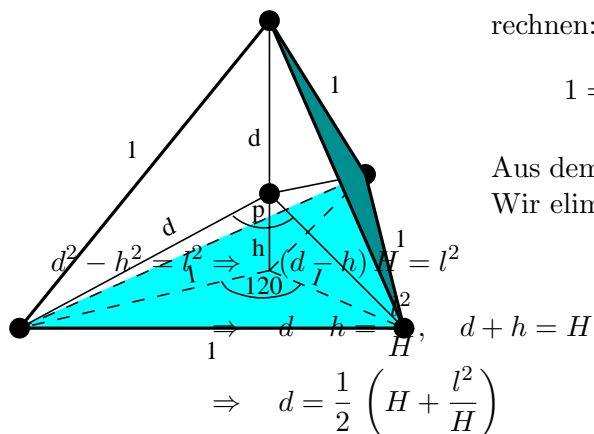
Beispiel 3.3.10: Wir illustrieren die Nützlichkeit der Winkelfunktionen an einem Beispiel, das u.a. für die Chemie³ Relevanz besitzt: der Berechnung des *Tetraederwinkels*.

Gesamthöhe des Tetraeders: $H = d + h$

Der Winkel p lässt sich aus d mit dem Cosinussatz berechnen:

$$1 = 2d^2 - 2d^2 \cos p.$$

Aus dem Satz des Pythagoras wissen wir: $d^2 = l^2 + h^2$.
Wir eliminieren h :



Berechne l : $1 = 2l^2 - 2l^2 \cos 120^\circ$ (Cosinussatz); $\cos 120^\circ = -\sin 30^\circ = -\frac{1}{2} \Rightarrow l^2 = \frac{1}{3}$.

Berechne H : $l^2 + H^2 = 1 \Rightarrow H^2 = \frac{2}{3}, H = \sqrt{\frac{2}{3}}$

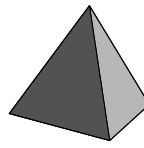
$$\Rightarrow d = \frac{1}{2} \sqrt{\frac{3}{2}}, d^2 = \frac{3}{8}, 2d^2 = \frac{3}{4} \Rightarrow 1 - \cos p = \frac{4}{3}$$

$$\cos p = -\frac{1}{3} \Rightarrow p = 109,5^\circ$$

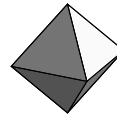
Bemerkung 3.3.11. Das Tetraeder ist der einfachste Körper, der durch gleich große regelmäßige n -Ecke (hier: $n = 3$) begrenzt wird. Solche Körper heißen reguläre Polyeder. Es gibt genau fünf verschiedene reguläre Polyeder, die sogenannten Platonischen Körper:

³Der Tetraeder ist ein gutes Modell für die räumliche Struktur gewisser Kristalle bzw. Moleküle, etwa Methan (CH_4) aber auch Wasser (verzerrter Tetraederwinkel $104,5^\circ$).

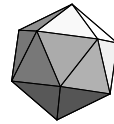
Tetraeder: 4 Dreiecke



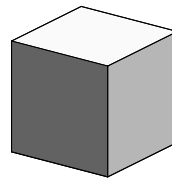
Oktaeder: 8 Dreiecke



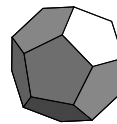
Ikosaeder: 20 Dreiecke



Würfel: 6 Quadrate



Dodekaeder: 12 Fünfecke



Diese regulären Polyeder spielen in der Chemie eine wichtige Rolle, da Kristalle und Moleküle möglichst regelmäßige Konfigurationen bilden.

3.4 Komplexe Zahlen

Wir haben bereits in Abschnitt 2.3. bemerkt, dass es algebraische Gleichungen ohne reelle Lösungen gibt, etwa $x^2 + 1 = 0$. Ähnliche Probleme haben wir beim Aufbau des Zahlensystems schon überwunden:

- $x + 1 = 0$ hat keine Lösung in \mathbb{N} ; deshalb wurde \mathbb{Z} eingeführt.
- $nx = 1$ mit $n \geq 2$ hat keine Lösung in \mathbb{Z} ; deshalb wurde \mathbb{Q} eingeführt.
- $x^2 = 2$ hat keine Lösung in \mathbb{Q} ; deshalb wurde \mathbb{R} eingeführt.

Da es keine reelle Zahl x mit $x^2 = -1$ gibt, führt man eine neue Zahl i ein, und zwar durch die Festlegung

$$i^2 = -1.$$

Es ist also i im Prinzip ein Ersatz für $\sqrt{-1}$. Nun kann man die Gleichung

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

auch im Fall $D := p^2/4 - q < 0$ (zunächst formal) lösen:

$$x_{1,2} = -\frac{p}{2} \pm \sqrt{D} = -\frac{p}{2} \pm \sqrt{q - p^2/4} \sqrt{-1} = -\frac{p}{2} \pm i \sqrt{q - p^2/4}.$$

Dabei entstehen offenbar Zahlen der Form $a + ib$ mit $a, b \in \mathbb{R}$.

Definition 3.4.1 (Komplexe Zahlen). Unter einer *komplexen* Zahl z versteht man eine Zahl der Form $z = x + iy$ mit $x, y \in \mathbb{R}$ und $i^2 = -1$. Dabei heißt x der *Realteil* von z (kurz $\operatorname{Re} z$) und y der *Imaginärteil* von z (kurz $\operatorname{Im} z$); i heißt auch imaginäre Einheit. Für die Menge aller komplexen Zahlen schreiben wir \mathbb{C} .

Für zwei komplexe Zahlen $z_1 = x_1 + iy_1$ und $z_2 = x_2 + iy_2$ definieren wir

- (i) $z_1 = z_2 \iff x_1 = x_2$ und $y_1 = y_2$ (Gleichheit),
- (ii) $z_1 + z_2 := (x_1 + x_2) + i(y_1 + y_2)$ (Addition),
- (iii) $z_1 z_2 := x_1 x_2 - y_1 y_2 + i(x_1 y_2 + x_2 y_1)$ (Multiplikation).

Die Regeln (ii) und (iii) besagen, dass mit i wie mit einer reellen Zahl gerechnet wird, wobei aber $i^2 = -1$ zu beachten ist. Dabei gelten die bekannten Rechenregeln.

Beispiel 3.4.2: Für $z_1 = 1 + 2i$ und $z_2 = -2 - i$ ist $z_1 + z_2 = -1 + i$ und $z_1 z_2 = -5i$.

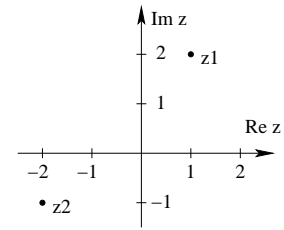
Da komplexe Zahlen aus zwei unabhängigen Anteilen bestehen, kann man sie mit Punkten der Ebene identifizieren:

$$z = x + iy \in \mathbb{C} \iff (x, y) \in \mathbb{R} \times \mathbb{R} =: \mathbb{R}^2.$$

Man erweitert dadurch den reellen Zahlenstrahl zur komplexen Zahlenebene, der sogenannten *Gauß'schen Zahlenebene*.

Eine wichtige Operation ist die Spiegelung an der reellen Achse, d.h. der Übergang von $z = x + iy$ zu $\bar{z} := x - iy$, der *zu z konjugiert komplexen Zahl*. Die Bedeutung von \bar{z} ergibt sich aus der Gleichung $z\bar{z} = (x + iy)(x - iy) = x^2 + y^2 \in \mathbb{R}$. Nach dem Satz von Pythagoras ist $\sqrt{z\bar{z}}$ der Abstand der komplexen Zahl z vom Koordinatenursprung. Analog zum Reellen heißt dieser Wert der *Betrag von z* , geschrieben als $|z|$. Also gilt $z\bar{z} = |z|^2$. Hieraus erhalten wir für $z \neq 0$, d.h. für $x \neq 0$ oder $y \neq 0$,

$$\frac{1}{z} = \frac{1}{|z|^2} \bar{z} = \frac{x - iy}{x^2 + y^2}.$$

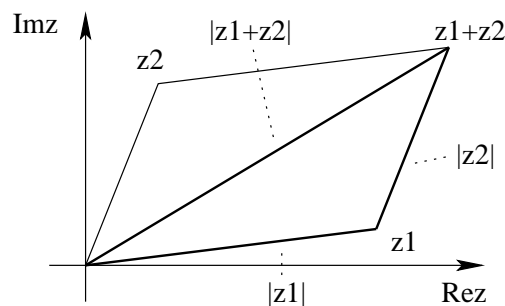


Definition 3.4.3. Für $z = x + iy \in \mathbb{C}$ definieren wir:

- (i) $\bar{z} := x - iy$ (konjugierte komplexe Zahl),
- (ii) $|z| := \sqrt{z\bar{z}} = \sqrt{x^2 + y^2}$ (Betrag von z),
- (iii) für $z \neq 0$: $z^{-1} = \frac{1}{z}$ als Lösung von $z^{-1}z = 1$ (Inverse von z bzgl. Multiplikation),
- (iv) für $z_2 \neq 0$: $\frac{z_1}{z_2} = z_1 \frac{1}{z_2}$ (Division).

Beispiel 3.4.4: Für $z = 1 + 2i$ ist $\bar{z} = 1 - 2i$, $z\bar{z} = 5$, $|z| = \sqrt{5}$ und $\frac{1}{z} = \frac{1}{5} - \frac{2}{5}i$.

Bemerkungen 3.4.5: (i) Für den Betrag gilt $|z| \geq 0$ für alle $z \in \mathbb{C}$ und $|z| = 0 \Leftrightarrow z = 0$. Weiter gilt $|z \cdot w| = |z| \cdot |w|$ und $|z + w| \leq |z| + |w|$. Der Name *Dreiecksungleichung* für die letztgenannte Ungleichung erklärt sich aus folgender Skizze:



(ii) Für die Konjugation gelten die folgenden Rechenregeln. Für alle $w, z \in \mathbb{C}$ ist

$$\overline{\bar{z}} = z, \quad \overline{z \pm w} = \bar{z} \pm \bar{w}, \quad \overline{z\bar{w}} = \bar{z} \cdot w \quad \text{und} \quad \overline{\left(\frac{z}{w}\right)} = \frac{\bar{z}}{\bar{w}}.$$

Außerdem ist eine komplexe Zahl z genau dann reell, wenn $\bar{z} = z$ gilt.

- (iii) Die komplexen Zahlen sind *nicht* angeordnet, d.h. es gibt keine Relation „ $<$ “ auf \mathbb{C} !
- (iv) Bei Polynomen mit reellen Koeffizienten treten komplexe Nullstellen immer in konjugiert komplexen Paaren (also z_0 und \bar{z}_0) auf.

Wie wir bereits gesehen haben, besitzt nun jedes quadratische Polynom in einer komplexen Variablen genau zwei komplexe Nullstellen. Wie verhält es sich mit Polynomen von allgemeinem Grad?

Satz 3.4.6 (Hauptsatz der Algebra). Sei P ein Polynom über \mathbb{C} vom Grad n , also

$$P(z) = a_n z^n + a_{n-1} z^{n-1} + \dots + a_1 z + a_0 \text{ mit Koeffizienten } a_k \in \mathbb{C} \text{ und } a_n \neq 0.$$

Dann hat P genau n komplexe Nullstellen, wobei mehrfache Nullstellen (d.h. Faktorisierung mit Faktor $(z - z_0)^k$) mit ihrer Vielfachheit (also k) zu zählen sind.

Beispiele 3.4.7: (a) Das Polynom

$$z^3 + 1 = (z + 1)(z^2 - z + 1) = (z + 1) \left(z - \frac{1}{2} - i \frac{\sqrt{3}}{2} \right) \left(z - \frac{1}{2} + i \frac{\sqrt{3}}{2} \right)$$

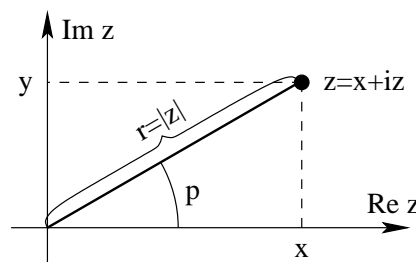
hat die Nullstellen $z_1 = -1$ und $z_{2,3} = \frac{1}{2} \pm i \frac{\sqrt{3}}{2}$; alle Nullstellen sind einfach.

(b) Das Polynom

$$\begin{aligned} z^3 + iz^2 + z + i &= z^3 + z + i(z^2 + 1) = (z + i)(z^2 + 1) \\ &= (z + i)(z + i)(z - i) \\ &= (z + i)^2 (z - i) \end{aligned}$$

hat die Nullstellen $z_1 = z_2 = -i$ (zweifach) und $z_3 = i$ (einfach).

Da komplexe Zahlen den Punkten in der Gauß'schen Zahlenebene entsprechen, lässt sich jede komplexe Zahl $z \in \mathbb{C}$ auch in *Polarkoordinaten* durch Angabe einer Länge und eines Winkels darstellen. Dazu benötigen wir wieder die trigonometrischen Funktionen Sinus und Cosinus.



Die Zahl $z = x + iy \in \mathbb{C}$ hat die *Polarkoordinatendarstellung*

$$z = r(\cos \varphi + i \sin \varphi)$$

mit dem Radius $r = |z|$ und dem Winkel φ zwischen „ \mathbb{R}_+ “ und z “ (gemessen im Bogenmaß). Man beachte, dass man der Zahl $0 \in \mathbb{C}$ *keinen* Winkel zuordnen kann.

Die Zahl φ heißt das *Argument* von z , kurz $\arg(z)$. Da es beim Winkel auf Vielfache von 2π nicht ankommt, hat man mehrere Möglichkeiten, das Intervall für $\arg(z)$ festzulegen. Oft definiert man $\arg(z)$ als den Winkel φ , für den $z = |z|(\cos \varphi + i \sin \varphi)$ und $-\pi < \varphi \leq \pi$ gilt.

Bemerkungen 3.4.8: (i) Die Exponentialfunktion e^z kann auch für komplexe $z \in \mathbb{C}$ definiert werden (als sog. Potenzreihe; dazu später mehr). Es gilt dann das *Eulersche Theorem*

$$e^{i\varphi} = \cos \varphi + i \sin \varphi \quad \text{für } \varphi \in \mathbb{R};$$

dies kann auch als Definition von $e^{i\varphi}$ verwendet werden. Damit folgt

$$e^z = e^{x+iy} = e^x(\cos y + i \sin y) \quad \text{für } z = x + iy \in \mathbb{C}.$$

Für die komplexe e -Funktion gilt wie im Reellen

$$e^{z+w} = e^z e^w \quad \forall z, w \in \mathbb{C}.$$

Setzt man hier speziell $z = ix$ und $w = iy$ ein, so erhält man mit Hilfe des Eulerschen Theorems leicht die Additionstheoreme.

(ii) Setzt man in obiger Formel $\varphi = \pi$, erhält man

$$e^{i\pi} = -1 \quad \text{bzw.} \quad e^{i\pi} + 1 = 0;$$

eine der bemerkenswertesten Formeln der Mathematik, die die 5 wichtigen Zahlen $0, 1, \pi, e$ und i vereint.

(iii) Für $z_1 = r_1 e^{i\varphi_1}$ und $z_2 = r_2 e^{i\varphi_2} \in \mathbb{C}$ gilt $z_1 z_2 = r_1 r_2 e^{i(\varphi_1 + \varphi_2)}$. Also werden bei der Multiplikation zweier komplexer Zahlen die Beträge (Längen) *multipliziert* und die Argumente (Winkel) *addiert*.

Kapitel 4

Folgen und Grenzwerte

4.1 Folgen

Definition 4.1.1 (Folge). Eine (reelle) Folge $(a_n)_{n \in I}$ ist eine Funktion $f : I \rightarrow \mathbb{R}$, die jedem $n \in I$ die Zahl $a_n = f(n) \in \mathbb{R}$ zuordnet. I ist dabei eine unendliche, abzählbare Indexmenge. Falls $I = \mathbb{N}$ kürzen wir die Notation ab zu (a_n) .

Bemerkungen 4.1.2: 1. Die Funktion f wird also durch die (unendl.) Aufzählung der Funktionswerte a_1, a_2, a_3, \dots angegeben. Die a_n heißen Folgenglieder (mit Index n).

2. Für $I = \{n_0, n_0 + 1, n_0 + 2, \dots\} \subset \mathbb{Z}$ schreiben wir $(a_n)_{n \geq n_0}$.

Beispiele 4.1.3: (a) $(a_n)_{n \geq 1}$ mit $a_n = 1/n$ ist die Folge $1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \dots$

(b) $(a_n)_{n \geq 2}$ mit $a_n = \frac{1}{n-1}$ ist die Folge $1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \dots$

(c) $(a_n)_{n \geq 0}$ mit $a_n = 2^n$ ist die Folge $1, 2, 4, 8, 16, \dots$. Die a_n wachsen über alle Grenzen hinaus.

(d) $(a_n)_{n \geq 1}$ mit $a_n = (-1)^n$ ist die Folge $-1, 1, -1, 1, -1, 1, \dots$. Diese Folge „oszilliert“ zwischen den Werten -1 und 1 .

(e) $(a_n)_{n \geq 1}$ mit $a_n = (1 + \frac{1}{n})^n$ beginnt mit $2, 2.25, 2.37, \dots$. Diese Folge haben wir bereits im Zusammenhang mit Wachstumsprozessen betrachtet. Die a_n nähern sich dem Wert $e = 2.71828 \dots$

(f) $(a_n)_{n \geq 0}$ mit $a_0 = a_1 = 1$ und $a_{n+1} = a_{n-1} + a_n$ für $n \geq 1$ ist ein Beispiel für eine rekursiv definierte Folge. Sie beginnt mit $1, 1, 2, 3, 5, 8, 13, \dots$. Die Glieder dieser Folge heißen Fibonacci-Zahlen.

(g) (a_n) mit $a_1 = 1$ und $a_{n+1} = \frac{1}{2}(a_n + \frac{2}{a_n})$ für $n \geq 1$ beginnt mit $1, 1.5, 1.4166, 1.4142 \dots$. Die a_n^2 nähern sich dem Wert 2 . Die konsequente Berechnung der Folgenglieder ist bekannt als das *Verfahren von Heron* zur näherungsweise Berechnung von $\sqrt{2}$.

Bemerkungen 4.1.4: (i) Bei (f) und (g) handelt es sich um rekursiv definierte Folgen. Dabei wird der Wert einiger Folgenglieder explizit vorgegeben, sowie eine Vorschrift, wie der Reihe nach alle weiteren Folgenglieder berechnet werden.

- (ii) Da Folgen spezielle Funktionen sind, überträgt sich der Monotoniebegriff. Z.B. heißt (a_n) (streng) monoton wachsend, falls $a_n \leq a_{n+1}$ ($a_n < a_{n+1}$) für alle $n \in \mathbb{N}$ gilt. In den Beispielen: (a_n) streng fallend in (a), streng wachsend in (c), nicht monoton in (d), wachsend in (f); die Folgen aus (e) und (g) werden wir später auf Monotonie untersuchen.

4.2 Konvergenz von Folgen

In den Beispielen 4.1.3 (e) und (g) interessiert man sich für den „Grenzwert“, der durch diese Folgen bestimmt wird.

Definition 4.2.1 (Grenzwert). Eine reelle Folge (a_n) hat den *Grenzwert* (Limes) a , falls für jedes $\varepsilon > 0$ ein $n_\varepsilon \in \mathbb{N}$ existiert, so dass

$$|a_n - a| \leq \varepsilon \quad \text{für alle } n \geq n_\varepsilon.$$

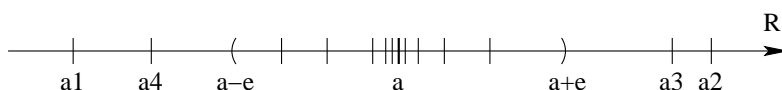
Man schreibt dann $\lim_{n \rightarrow \infty} a_n = a$ (oder auch: $a_n \rightarrow a$ für $n \rightarrow \infty$) und nennt (a_n) *konvergent* gegen a . Ist eine Folge (a_n) nicht konvergent, so heißt sie *divergent*.

Bemerkungen 4.2.2: (i) Für die Konvergenz von (a_n) sind nur Folgenglieder mit großem Index wichtig. Insbesondere gilt dieselbe Definition für Folgen $(a_n)_{n \geq n_0}$.

(ii) Ist (a_n) konvergent mit dem Grenzwert 0, so heißt (a_n) eine *Nullfolge*.

(iii) Entscheidend für die Konvergenz gegen a ist offenbar, dass alle Folgenglieder mit hinreichend großem Index *nah genug* an a liegen. Konvergente Folgen können daher überall dort definiert werden, wo man von Abständen sprechen kann (z.B. in \mathbb{C} , \mathbb{R}^2 , \mathbb{R}^3 , ...).

(iv) Man bezeichnet das offene Intervall $(a - \varepsilon, a + \varepsilon)$ auch als ε -Umgebung von a . Es gilt genau dann $\lim_{n \rightarrow \infty} a_n = a$, wenn *jede* ε -Umgebung von a alle Folgenglieder a_n bis auf endlich viele enthält.



Liegen in jeder ε -Umgebung von a unendlich viele der a_n (aber evtl. auch unendlich viele außerhalb), so bezeichnet man a als *Häufungswert* der Folge (a_n) . Ist (a_n) konvergent, so besitzt (a_n) also genau einen Häufungswert, und zwar den Grenzwert $\lim_{n \rightarrow \infty} a_n$.

Beispielsweise hat die Folge (a_n) aus 4.1.3 (d) die beiden Häufungswerte ± 1 ; sie ist also divergent.

Definition 4.2.3 (Beschränktheit). Eine Folge (a_n) heißt *beschränkt*, wenn es eine *Schranke* $S > 0$ gibt mit der Eigenschaft, dass $|a_n| \leq S$ für alle $n \in \mathbb{N}$.

Satz 4.2.4. Jede konvergente Folge ist beschränkt.

Beweis. Die Folge (a_n) sei konvergent, d.h. es gibt $a \in \mathbb{R}$ mit $\lim_{n \rightarrow \infty} a_n = a$. Dann gibt es insbesondere ein $N \in \mathbb{N}$ so, dass $|a_n - a| \leq 1$ für alle $n \geq N$ (setze $\varepsilon := 1$ in der Definition des Grenzwertes). Dann gilt für alle $n \geq N$

$$|a_n| = |a_n - a + a| \leq |a_n - a| + |a| \leq |a| + 1.$$

Setzen wir noch $M := \max\{|a_1|, |a_2|, \dots, |a_{N-1}|\}$, so gilt für alle $n \in \mathbb{N}$

$$|a_n| \leq \max\{M, |a| + 1\} =: S. \quad \square$$

Beispiele 4.2.5: (a) Wir betrachten die Folge (a_n) mit $a_n = 1/n$. Vermutung: (a_n) ist eine Nullfolge.

Zu zeigen ist also: $\forall \varepsilon > 0 \exists n_\varepsilon \in \mathbb{N} \forall n \geq n_\varepsilon : |a_n| \leq \varepsilon$.

Wir arbeiten rückwärts: Wir wollen $|\frac{1}{n}| \leq \varepsilon$ erhalten. Dies ist äquivalent zu $\frac{1}{n} \leq \varepsilon$ bzw. $n \geq \frac{1}{\varepsilon}$.

Ein Kandidat für n_ε ist also $\frac{1}{\varepsilon}$; diese Zahl ist aber i. allg. nicht ganzzahlig. Wir wählen daher für n_ε die nächstgrößere ganze Zahl. Man schreibt $\lfloor x \rfloor$ für die nächstkleinere ganze Zahl von x , etwa $\lfloor 3.14 \rfloor = 3$, $\lfloor -7.1 \rfloor = -8$ (Gaußklammerfunktion).

Damit gilt für $n_\varepsilon = \lfloor \frac{1}{\varepsilon} \rfloor + 1$ (beachte: n_ε hängt von ε ab!):

$$n \geq n_\varepsilon \quad \Rightarrow \quad n \geq \frac{1}{\varepsilon} \quad \Rightarrow \quad \left| \frac{1}{n} \right| \leq \varepsilon.$$

(b) Sei (a_n) mit $a_n = 2^n$. Die naheliegende Vermutung, dass (a_n) divergent ist, wird durch das oben bewiesene Lemma bestätigt, denn (a_n) ist unbeschränkt. Hier gilt sogar:

$$\forall S > 0 \exists N \in \mathbb{N} \forall n \geq N : a_n \geq S. \quad (*)$$

In diesem Fall heißt (a_n) *bestimmt divergent* gegen $+\infty$; man schreibt auch $\lim_{n \rightarrow \infty} a_n = \infty$ und bezeichnet $+\infty$ als *uneigentlichen Grenzwert* von (a_n) (entsprechend mit $-\infty$ falls „ $a_n \leq -S$ “ in (*)).

Bemerkung 4.2.6. Es gilt folgender Zusammenhang zwischen Nullfolgen und bestimmt divergenten Folgen:

- (i) Ist (a_n) Nullfolge mit $a_n > 0$ (bzw. mit $a_n < 0$) für alle $n \in \mathbb{N}$, so ist (b_n) mit $b_n = 1/a_n$ bestimmt divergent gegen $+\infty$ (bzw. $-\infty$).
- (ii) Ist (b_n) bestimmt divergent gegen $+\infty$ (bzw. $-\infty$), so gibt es ein $n_0 \in \mathbb{N}$ so, dass $(a_n)_{n \geq n_0}$ mit $a_n = 1/b_n$ eine Nullfolge (mit $a_n \gtrless 0$ für $n \geq n_0$) ist.

Inbesondere ist also (a_n) mit $a_n = \frac{1}{n^p}$ für $p > 0$ stets eine Nullfolge.

4.3 Grenzwertsätze

Für komplizierte Folgen möchte man die Untersuchung auf Konvergenz und die Grenzwertbestimmung auf einzelne Anteile zurückführen. Dies gelingt in folgenden Situationen.

Satz 4.3.1 (Rechnen mit Grenzwerten). Die Folgen (a_n) und (b_n) seien konvergent mit den Grenzwerten a bzw. b . Dann gilt:

- (i) $\lim_{n \rightarrow \infty} (a_n + b_n) = a + b$;
- (ii) $\lim_{n \rightarrow \infty} (a_n b_n) = a b$;
- (iii) $\lim_{n \rightarrow \infty} (a_n/b_n) = a/b$ falls $b \neq 0$;

(iv) aus $a_n \leq b_n$ für alle $n \geq n_0$ folgt $a \leq b$.

Beispiel 4.3.2: Sei (a_n) mit $a_n = \frac{n^2-n}{2n^2+1} = \frac{1-1/n}{2+1/n^2}$. Wir wissen, dass $\frac{1}{n} \rightarrow 0$ und $\frac{1}{n^2} \rightarrow 0$ für $n \rightarrow \infty$. Aus Satz 4.3.1 i und iii folgt dann

$$1 - \frac{1}{n} \rightarrow 1 \quad \text{und} \quad 2 + \frac{1}{n^2} \rightarrow 2 \quad \text{für } n \rightarrow \infty,$$

und mit Satz 4.3.1 iii erhalten wir $a_n \rightarrow \frac{1}{2}$ für $n \rightarrow \infty$.

Bemerkungen 4.3.3: (i) Teil (iv) von 4.3.1 wird oft mit $b_n = b$ für alle n angewandt. Man beachte auch: aus $a_n < b_n \forall n$ folgt *nicht* $a < b$!

(ii) Eine Folge $(a_n - b_n)$ kann konvergieren, obwohl (a_n) und (b_n) divergent sind.

Beispiel 4.3.4: Für $a_n = \sqrt{n+1}$ und $b_n = \sqrt{n}$ ist $\lim_{n \rightarrow \infty} a_n = \infty = \lim_{n \rightarrow \infty} b_n$, aber

$$a_n - b_n = \sqrt{n+1} - \sqrt{n} = \frac{n+1-n}{\sqrt{n+1} + \sqrt{n}} = \frac{1}{\sqrt{n+1} + \sqrt{n}} \rightarrow 0 \quad \text{für } n \rightarrow \infty.$$

Um Letzteres zu zeigen, verwendet man Teil (i) des folgenden Satzes.

Satz 4.3.5 (Einschachtelungssatz). Sei (a_n) eine Folge und (b_n) eine Nullfolge.

(i) Gilt $|a_n| \leq |b_n|$ ab einem $n_0 \in \mathbb{N}$, so ist auch (a_n) Nullfolge.

(ii) Ist (a_n) beschränkt, so ist $(a_n b_n)$ eine Nullfolge.

Beweis.

(i) Sei $\varepsilon > 0$. Dann existiert $n_\varepsilon \in \mathbb{N}$, so dass $|b_n| \leq \varepsilon \Rightarrow |a_n| \leq |b_n| \leq \varepsilon$ für $n \geq n_\varepsilon$.

(ii) Aus $|a_n| \leq S$ für alle $n \geq n_\varepsilon$ folgt $|a_n b_n| \leq S |b_n|$ und $S b_n \rightarrow 0$ nach Satz 4.3.1 (ii). Die Aussage ergibt sich nun durch Anwendung von (i). \square

Beispiel 4.3.6: Wir betrachten die Folge (a_n) mit $a_n = q^n$ für ein $q \in \mathbb{R}$ und unterscheiden folgende Fälle:

(a) $q = 0$: $a_n = 0 \quad \forall n \Rightarrow \lim a_n = 0$

(b) $q = 1$: $a_n = 1 \quad \forall n \Rightarrow \lim a_n = 1$

(c) $0 < q < 1$: $a_n \rightarrow 0$, denn $|q^n| \leq \varepsilon \Leftrightarrow q^n \leq \varepsilon \Leftrightarrow n \ln q \leq \ln \varepsilon \Leftrightarrow n \geq \frac{\ln \varepsilon}{\ln q}$;

wähle also $n_\varepsilon = 1 + \left\lceil \frac{\ln \varepsilon}{\ln q} \right\rceil$.

(d) $q > 1$: $a_n \rightarrow \infty$, denn $q = 1 + x$ mit $x > 0 \Rightarrow q^n = (1+x)^n \geq 1 + nx$ nach Bernoulli-Ungleichung. Daraus folgt $q^n \nearrow \infty$ für $n \rightarrow \infty$.

(e) $-1 < q < 0$: $q^n = (-1)^n (-q)^n$; $(-1)^n$ ist beschränkt, $(-q)^n$ ist nach (c) Nullfolge $\Rightarrow q^n \rightarrow 0$ nach Satz 4.3.5 (ii).

(f) $q = -1$: $q^n = (-1)^n$ divergent; zwei Häufungswerte ± 1 .

- (g) $q < -1$: $|q^n| = |q|^n \rightarrow \infty$, d.h. divergent; (q^n) ist nicht bestimmt divergent, da die Vorzeichen wechseln.

Mit den bisherigen Sätzen lässt sich Konvergenz nur zeigen, wenn man den Grenzwert der Folge kennt. Das folgende Kriterium für monotone Folgen ist daher besonders wichtig:

Satz 4.3.7 (Monotone beschränkte Konvergenz). Sei (a_n) monoton wachsend und nach oben beschränkt oder monoton fallend und nach unten beschränkt. Dann ist (a_n) konvergent.

Bemerkungen 4.3.8: Satz 4.3.7 gilt auch, wenn (a_n) erst ab einem bestimmten Index $n_0 \in \mathbb{N}$ monoton ist.

Beispiel 4.3.9: Wir betrachten die Folge $(a_n)_{n \geq 0}$ aus Beispiel 4.1.3 (g):

$$a_0 = 1, \quad a_{n+1} = \frac{1}{2} \left(a_n + \frac{2}{a_n} \right) \quad \text{für } n \geq 0.$$

Für diese Folge wollen wir zeigen:

- (1) Beschränktheit (genauer: nach unten),
- (2) Monotonie (genauer: monoton fallend),
- (3) Konvergenz gegen $a = \sqrt{2}$.

zu (1): Aus $a_0 = 1$ und $a_{n+1} = \frac{1}{2} \left(a_n + \frac{2}{a_n} \right)$ folgt induktiv, dass alle $a_n > 0$ sind. Die Folge (a_n) ist also nach unten durch 0 beschränkt.

zu (2): Wegen

$$\begin{aligned} a_{n+1}^2 - 2 &= \frac{1}{4} \left(a_n^2 + 2a_n \frac{2}{a_n} + \frac{4}{a_n^2} \right) - 2 = \frac{1}{4} \left(a_n^2 - 2a_n \frac{2}{a_n} + \frac{4}{a_n^2} \right) \\ &= \frac{1}{4} \left(a_n - \frac{2}{a_n} \right)^2 \geq 0 \quad \text{für } n \geq 0 \end{aligned}$$

ist $a_n^2 \geq 2$ für $n \geq 1$, und wegen der Äquivalenzen

$$a_{n+1} \leq a_n \Leftrightarrow a_n + \frac{2}{a_n} \leq 2a_n \Leftrightarrow \frac{2}{a_n} \leq a_n \Leftrightarrow 2 \leq a_n^2$$

für $n \geq 1$ folgt aus $a_n^2 \geq 2$, dass $a_{n+1} \leq a_n$. Also ist die Folge (a_n) monoton fallend.

zu (4): Die Folge (a_n) konvergiert nach Satz 4.3.7, und für ihren Grenzwert $a \in \mathbb{R}$ gilt $a \geq 0$ nach Satz 4.3.1 (iv). Welchen Wert hat a ? Aus der Rekursionsformel folgt:

$$2a_n a_{n+1} = a_n^2 + 2 \quad (\text{wir wissen noch nicht, ob } a \neq 0 \text{ gilt!}).$$

Außerdem ist klar: Die verschobene Folge (a_{n+1}) konvergiert ebenfalls gegen a . Mit Satz 4.3.1 erhalten wir, dass $2a_n a_{n+1} \rightarrow 2a^2$ und $a_n^2 \rightarrow a^2$ für $n \rightarrow \infty$. Damit gilt $a^2 = 2$ und $a \geq 0$, also $a = \sqrt{2}$.

Auf ähnliche Weise kann man zeigen, dass die Folge (a_n) mit $a_n = \left(1 + \frac{1}{n}\right)^n$ konvergiert.

Bemerkung 4.3.10. Es gibt ein weiteres allgemeingültiges Konvergenzkriterium, das keine Kenntnis des Grenzwertes voraussetzt: *Eine reelle Folge (a_n) ist genau dann konvergent, wenn (a_n) eine Cauchy-Folge ist, d.h. wenn*

$$\forall \varepsilon > 0 \exists n_\varepsilon \in \mathbb{N} \forall n, m \geq n_\varepsilon : |a_n - a_m| \leq \varepsilon.$$

Dieses *Cauchysche Konvergenzkriterium* ist oft schwierig anzuwenden.

4.4 Reihen

Wir betrachten nun eine spezielle Art von Folgen, die sogenannten *Reihen*. Sei $(a_k)_{k \geq 0}$ eine (reelle) Folge. Wir betrachten die *Folge der Partialsummen* (s_n) mit $s_n := \sum_{k=0}^n a_k$. Die Bezeichnung $\sum_{k=0}^{\infty} a_k$ steht sowohl für diese Folge als auch für ihren Grenzwert, falls dieser existiert.

Definition 4.4.1 (Reihe). Die (unendliche) Reihe $\sum_{k=0}^{\infty} a_k$ heißt *konvergent*, falls die Folge (s_n) der Partialsummen konvergiert. Dann nennt man $s := \lim_{n \rightarrow \infty} s_n$ den Wert der Reihe und schreibt kurz $\sum_{k=0}^{\infty} a_k = s$; andernfalls heißt die Reihe *divergent*.

Entsprechend erklärt man auch $\sum_{k=k_0}^{\infty} a_k$, etwa $\sum_{k=1}^{\infty} \frac{1}{k}$; man schreibt auch kurz $\sum_{k \geq k_0} a_k$.

Satz 4.4.2 (Geometrische Reihe). Die *geometrische Reihe*

$$\sum_{n=0}^{\infty} q^n = 1 + q + q^2 + q^3 + \dots \quad \text{mit } q \in \mathbb{R}$$

hat die Partialsummen

$$s_n = \sum_{k=0}^n q^k = \begin{cases} \frac{1-q^{n+1}}{1-q} & \text{falls } q \neq 1, \\ n+1 & \text{falls } q = 1. \end{cases}$$

Die Folge der Partialsummen und damit die geometrische Reihe ist genau dann konvergent, wenn $|q| < 1$ gilt. In diesem Fall ist

$$\sum_{n=0}^{\infty} q^n = \frac{1}{1-q}.$$

Beweis. Es ist

$$s_n = 1 + q + q^2 + \dots + q^n \quad \text{und} \quad q s_n = q + q^2 + \dots + q^n + q^{n+1},$$

also $s_n - q s_n = 1 - q^{n+1}$. Für $q \neq 1$ ist daher

$$s_n = \frac{1 - q^{n+1}}{1 - q},$$

während für $q = 1$ offenbar $s_n = n + 1$ ist; in diesem Fall ist $\sum_{n \geq 0} q^n$ divergent.

Sei also $q \neq 1$. Für Konvergenz von $\frac{1-q^{n+1}}{1-q}$ muss (q^{n+1}) konvergent sein und dies gilt nur, falls $|q| < 1$. Also ist die Bedingung $|q| < 1$ notwendig für die Konvergenz der geometrischen Reihe. Aus $|q| < 1$ folgt umgekehrt $q^{n+1} \rightarrow 0$ und damit die Konvergenz der s_n . \square

Eine *notwendige Bedingung* für die Konvergenz der Reihe $\sum_{k=0}^{\infty} a_k$ ist $\lim_{n \rightarrow \infty} a_n = 0$, denn aus $s_n \rightarrow s$ folgt $s_{n-1} \rightarrow s$, also $a_n = s_n - s_{n-1} \rightarrow 0$ für $n \rightarrow \infty$). Vorsicht ist jedoch angebracht: Die Bedingung $a_n \rightarrow 0$ ist NICHT HINREICHEND für die Konvergenz der Reihe, wie folgender Satz zeigt.

Satz 4.4.3 (Harmonische Reihe). Die *harmonische Reihe* $\sum_{n=1}^{\infty} \frac{1}{n}$ ist bestimmt divergent.

Beweis. Es ist

$$\begin{aligned} \sum_{n=1}^{\infty} \frac{1}{n} &= 1 + \frac{1}{2} + \left(\frac{1}{3} + \frac{1}{4}\right) + \left(\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}\right) + \left(\frac{1}{9} + \dots + \frac{1}{16}\right) + \dots \\ &\geq 1 + \frac{1}{2} + \left(\frac{1}{4} + \frac{1}{4}\right) + \left(\frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8}\right) + \left(\frac{1}{16} + \dots + \frac{1}{16}\right) + \dots \\ &= 1 + \frac{1}{2} + 2 \cdot \frac{1}{4} + 4 \cdot \frac{1}{8} + 8 \cdot \frac{1}{16} + \dots \\ &= 1 + \frac{1}{2} + \frac{1}{2} + \frac{1}{2} + \frac{1}{2} + \dots \end{aligned}$$

Für $n = 2^m$ gilt deshalb $s_n \geq 1 + \frac{m}{2}$, d.h. s_n ist bestimmt divergent gegen $+\infty$. \square

4.5 Konvergenzkriterien für Reihen

Aus den Grenzwertsätzen aus Abschnitt 4.3 folgt sofort

Satz 4.5.1 (Rechenregeln für Reihen). Es ist

$$(i) \sum_{n=0}^{\infty} (a_n + b_n) = a + b \quad \text{falls} \quad \sum_{n=0}^{\infty} a_n = a \quad \text{und} \quad \sum_{n=0}^{\infty} b_n = b.$$

$$(ii) \sum_{n=0}^{\infty} (\lambda a_n) = \lambda a \quad \text{falls} \quad \sum_{n=0}^{\infty} a_n = a \quad \text{und} \quad \lambda \in \mathbb{R}.$$

(iii) Ist $a_n \geq 0$ für alle $n \geq 0$, so gilt:

$$\sum_{n=0}^{\infty} a_n \quad \text{ist konvergent} \quad \Leftrightarrow \quad \text{Die Folge } (s_n) \text{ der Partialsummen ist beschränkt.}$$

Bemerkungen 4.5.2: (i) In (iii) ist die Folge (s_n) monoton wachsend.

(ii) Es gilt NICHT $\sum_{n \geq 0} a_n b_n = \sum_{n \geq 0} a_n \sum_{n \geq 0} b_n$! Falls aber die Reihen $\sum_{n \geq 0} |a_n|$ und $\sum_{n \geq 0} |b_n|$ konvergieren, dann gilt

$$\sum_{n \geq 0} a_n \sum_{n \geq 0} b_n = \sum_{n \geq 0} \sum_{k=0}^n a_k b_{n-k} \quad (\text{Cauchy-Produkt}).$$

Beispiel 4.5.3: Wir betrachten die Reihe $\sum_{k=1}^{\infty} \frac{1}{k(k+1)}$. Aus $\frac{1}{k(k+1)} = \frac{1}{k} - \frac{1}{k+1}$ folgt

$$s_n = \sum_{k=1}^n \left(\frac{1}{k} - \frac{1}{k+1} \right) = \sum_{k=1}^n \frac{1}{k} - \sum_{k=2}^{n+1} \frac{1}{k} = 1 - \frac{1}{n+1} \rightarrow 1.$$

Also ist die Reihe konvergent, und ihre Summe ist 1.

Eine Reihe der Form $\sum_{k=0}^{\infty} (-1)^k a_k$ mit $a_k \geq 0$ für alle $n \geq 0$ heißt *alternierend*.

Satz 4.5.4 (Leibniz-Kriterium). Die Reihe $\sum_{k=0}^{\infty} (-1)^k a_k$ mit $a_k \geq 0$ ist konvergent, falls (a_n) eine *monotone Nullfolge* ist.

In diesem Fall gilt

$$|s_n - s| \leq a_{n+1}, \quad \text{wobei } s_n = \sum_{k=0}^n (-1)^k a_k \text{ und } s = \sum_{k=0}^{\infty} (-1)^k a_k \text{ ist.}$$

Beispiel 4.5.5: Die *alternierende harmonische Reihe* $\sum_{n=1}^{\infty} (-1)^{n-1} \frac{1}{n}$ ist konvergent. (Ihre Summe ist $\ln 2$, was wir noch nicht begründen können.)

Dieses Beispiel zeigt, dass aus Konvergenz der Reihe $\sum_{n=0}^{\infty} a_n$ NICHT die Konvergenz der Reihe $\sum_{n=0}^{\infty} |a_n|$ folgt. Wenn die Reihe $\sum_{n=0}^{\infty} |a_n|$ konvergiert, heißt die Reihe $\sum_{n=0}^{\infty} a_n$ *absolut konvergent*. Jede absolut konvergente Reihe ist konvergent.

Satz 4.5.6 (Kriterien für absolute Konvergenz). Die Reihe $\sum_{k=0}^{\infty} a_k$ ist absolut konvergent, falls eine der folgenden Bedingungen erfüllt ist.

- (i) $|a_k| \leq b_k$ für alle $k \geq k_0$, und $\sum_{k=0}^{\infty} b_k$ ist konvergent. (**Majorantenkriterium**)
- (ii) $\sqrt[k]{|a_k|} \leq q < 1$ für alle $k \geq k_0$. (**Wurzelkriterium**)
- (iii) $\left| \frac{a_{k+1}}{a_k} \right| \leq q < 1$ für alle $k \geq k_0$. (**Quotientenkriterium**)

Bemerkung 4.5.7. Die Teile (ii) und (iii) folgen aus (i), indem man $b_k = Mq^k$ (geometrische Reihe) als Majorante verwendet. Anstelle von (ii) bzw. (iii) verwendet man oft die stärkeren Bedingungen

- (ii)' $\lim_{k \rightarrow \infty} \sqrt[k]{|a_k|} < 1$,
- (iii)' $\lim_{k \rightarrow \infty} \left| \frac{a_{k+1}}{a_k} \right| < 1$.

Beispiele 4.5.8: (a) Die Reihen $\sum_{n=1}^{\infty} \frac{1}{n^p}$ mit $p \geq 2$.

Für $a_n = \frac{1}{n^p}$ ist $|a_n| = \frac{1}{n^p} \leq \frac{1}{n^2} \leq 2 \frac{1}{n(n+1)}$ für $p \geq 2$ und $n \geq 1$.

Da die Reihe $\sum_{n=1}^{\infty} \frac{2}{n(n+1)}$ konvergiert, konvergiert nach dem Majorantenkriterium auch die Reihe $\sum_{n=1}^{\infty} \frac{1}{n^p}$. Für $p = 2$ kann man ihre Summe berechnen:

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}.$$

Man kann zeigen, dass die Reihe $\sum_{n=1}^{\infty} \frac{1}{n^p}$ bereits für alle $p > 1$ konvergiert. Die harmonische Reihe ($p = 1$) zeigt, dass diese Bedingung scharf ist.

(b) Die Reihen $\sum_{n=0}^{\infty} \frac{1}{n!}$ bzw. $\sum_{n=0}^{\infty} \frac{x^n}{n!}$ für $x \in \mathbb{R}$.

Hier ist $a_n = x^n/n!$, also

$$\left| \frac{a_{n+1}}{a_n} \right| = \left| \frac{x^{n+1}}{(n+1)!} \cdot \frac{n!}{x^n} \right| = \left| \frac{x}{n+1} \right| = \frac{|x|}{n+1}.$$

Somit ist $\lim_{n \rightarrow \infty} |a_{n+1}/a_n| = 0 < 1$. Die Reihe ist also nach dem Quotientenkriterium für jedes $x \in \mathbb{R}$ konvergent.

Es gilt $\sum_{n=0}^{\infty} \frac{x^n}{n!} = e^x$; dies ist die übliche Definition der Exponentialfunktion.

Satz 4.5.9 (Divergenzkriterien). Die Reihe $\sum_{k=0}^{\infty} a_k$ ist divergent, falls eine der folgenden Bedingungen erfüllt ist.

- (i) (a_k) ist keine Nullfolge.
- (ii) $a_k \geq b_k \geq 0$ für alle $k \geq k_0$ und $\sum_{k=0}^{\infty} b_k$ ist divergent. **(Minorantenkriterium)**
- (iii) $\sqrt[k]{|a_k|} \geq 1$ für alle $k \geq k_0$. **(Wurzelkriterium)**
- (iv) $a_k \neq 0$ und $\left| \frac{a_{k+1}}{a_k} \right| \geq 1$ für alle $k \geq k_0$. **(Quotientenkriterium)**

Bemerkungen 4.5.10: (i) Wie bei den Konvergenzkriterien zeigt man bei Verwendung des Wurzel- bzw. Quotientenkriteriums meist

$$\lim_{n \rightarrow \infty} \sqrt[n]{|a_n|} > 1 \quad \text{bzw.} \quad \lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| > 1.$$

- (ii) Man beachte, dass im Fall, dass obige Grenzwerte den Wert 1 haben, das Wurzel- und Quotientenkriterium keine Aussage über Konvergenz oder Divergenz liefern. Dann sind andere Methoden gefragt.

Beispiel 4.5.11: Sei $0 < p < 1$. Da dann $n^p \leq n^1 = n$ für alle $n \in \mathbb{N}$ gilt, haben wir

$$\frac{1}{n^p} \geq \frac{1}{n} \geq 0 \quad \text{für alle } n \in \mathbb{N}.$$

Da $\sum_{n=1}^{\infty} 1/n$ divergent ist (harmonische Reihe), folgt die Divergenz von $\sum_{n=1}^{\infty} \frac{1}{n^p}$ für alle $0 < p < 1$ nach dem Minorantenkriterium.

Bemerkungen 4.5.12: (i) Die Reihe $\sum_{n=1}^{\infty} \frac{1}{n^p}$ ist *konvergent* für $p > 1$ und *divergent* für $p \leq 1$.

- (ii) Eine kleine Liste wichtiger

konvergenter Majoranten: $\frac{1}{n^2}, \frac{1}{n\sqrt{n}}$, geometrische Reihe mit $|q| < 1, \dots$

divergenter Minoranten: $1, \frac{1}{n}, \frac{1}{\sqrt{n}}, \dots$

4.6 Potenzreihen

Definition 4.6.1. Sei $x_0 \in \mathbb{R}$ und $(a_n)_{n \geq 0}$ eine Folge.

$$\sum_{n=0}^{\infty} a_n (x - x_0)^n = a_0 + a_1(x - x_0) + a_2(x - x_0)^2 + \dots, \quad x \in \mathbb{R}$$

heißt *Potenzreihe* (zu $(a_n)_{n \geq 0}$) mit *Entwicklungsstelle* x_0 .

Potenzreihen sind – sehr lax gesprochen – unendlich lange Polynome, zählen aber natürlich im Allgemeinen nicht zu den Polynomen. Allerdings ist jedes Polynom eine (besonders einfache) Potenzreihe, bei der nur endlich viele Koeffizienten a_n ungleich 0 sind.

Beispiele 4.6.2: Beispiele für Potenzreihen sind

$$\sum_{n=0}^{\infty} x^n, \quad \sum_{n=0}^{\infty} \frac{x^n}{n!}, \quad \sum_{n=0}^{\infty} 2^n (x-2)^n, \quad \sum_{n=1}^{\infty} \frac{2^n}{n} x^{3n}, \quad \sum_{n=0}^{\infty} n^n x^n.$$

Die Konvergenz von Potenzreihen hängt natürlich davon ab, welches $x \in \mathbb{R}$ man einsetzt. Nach dem Wurzelkriterium, vgl. Satz 4.5.6 und Satz 4.5.9, konvergiert die Potenzreihe $\sum_{n=0}^{\infty} a_n (x - x_0)^n$, falls

$$\lim_{n \rightarrow \infty} \sqrt[n]{|a_n (x - x_0)^n|} = \lim_{n \rightarrow \infty} \sqrt[n]{|a_n|} |x - x_0| = |x - x_0| \lim_{n \rightarrow \infty} \sqrt[n]{|a_n|} < 1$$

ist, und sie divergiert, falls dieser Grenzwert größer als Eins ist.

1. Fall, $\lim_{n \rightarrow \infty} \sqrt[n]{|a_n|} = 0$: Dann konvergiert die Potenzreihe für alle $x \in \mathbb{R}$.
2. Fall, $\lim_{n \rightarrow \infty} \sqrt[n]{|a_n|} = \infty$: Dann konvergiert die Potenzreihe nur für $x = x_0$.
3. Fall, $0 < \lim_{n \rightarrow \infty} \sqrt[n]{|a_n|} < \infty$: Konvergenz für $|x - x_0| < \frac{1}{\lim_{n \rightarrow \infty} \sqrt[n]{|a_n|}}$ und Divergenz für $|x - x_0| > \frac{1}{\lim_{n \rightarrow \infty} \sqrt[n]{|a_n|}}$.

Definition 4.6.3 (Formel von Hadamard). Falls der Grenzwert $\lim_{n \rightarrow \infty} \sqrt[n]{|a_n|}$ existiert, so heißt

$$\varrho := \begin{cases} \frac{1}{\lim_{n \rightarrow \infty} \sqrt[n]{|a_n|}} & \text{falls } \lim_{n \rightarrow \infty} \sqrt[n]{|a_n|} \in (0, \infty), \\ 0 & \text{falls } \lim_{n \rightarrow \infty} \sqrt[n]{|a_n|} = \infty, \\ \infty & \text{falls } \lim_{n \rightarrow \infty} \sqrt[n]{|a_n|} = 0 \end{cases}$$

der *Konvergenzradius* der Potenzreihe $\sum_{n=0}^{\infty} a_n (x - x_0)^n$.

Obige Überlegungen ergeben damit sofort den folgenden Satz.

Satz 4.6.4. Sei $\sum_{n=0}^{\infty} a_n (x - x_0)^n$ eine Potenzreihe mit Konvergenzradius ϱ . Diese Reihe konvergiert für alle $x \in \mathbb{R}$ mit $|x - x_0| < \varrho$, und sie divergiert für alle $|x - x_0| > \varrho$. (Für Punkte x mit $|x - x_0| = \varrho$ ist auf diesem Wege keine Aussage möglich.)

Für die Reihen aus Beispiel 4.6.2 ergibt sich damit das folgende Konvergenzverhalten.

1. $\sum_{n=0}^{\infty} x^n$. Hier ist $a_n = 1$ für alle $n \in \mathbb{N}$, also

$$\lim_{n \rightarrow \infty} \sqrt[n]{|a_n|} = \lim_{n \rightarrow \infty} 1 = 1 \implies \varrho = 1.$$

Wir finden so das Konvergenzverhalten der geometrischen Reihe wieder, vgl. Satz 4.4.2. Dort haben wir außerdem schon

$$\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}, \quad |x| < 1$$

gesehen.

2. $\sum_{n=0}^{\infty} \frac{x^n}{n!}$. Hier ist $a_n = \frac{1}{n!}$, $n \in \mathbb{N}$. Die Anwendung der Formel von Hadamard, d.h. des Wurzelkriteriums, führt hier auf den kniffligen Grenzwert $\lim_{n \rightarrow \infty} \sqrt[n]{n!}$. Wir verwenden darum das Quotientenkriterium. Dazu betrachten wir die Folge $b_n := \frac{x^n}{n!}$, $n \in \mathbb{N}$. Dann folgt Konvergenz der Reihe, falls

$$1 > \lim_{n \rightarrow \infty} \frac{|b_{n+1}|}{|b_n|} = \lim_{n \rightarrow \infty} \left| \frac{x^{n+1}}{(n+1)!} \right| \cdot \left| \frac{n!}{x^n} \right| = \lim_{n \rightarrow \infty} \frac{|x|^{n+1}}{(n+1)!} \cdot \frac{n!}{|x|^n} = \lim_{n \rightarrow \infty} \frac{|x|}{n+1} = 0.$$

Also ist diese Reihe für alle $x \in \mathbb{R}$ konvergent, der Konvergenzradius also $\rho = \infty$.

Tatsächlich gilt

$$\sum_{n=0}^{\infty} \frac{x^n}{n!} = e^x \quad \text{für alle } x \in \mathbb{R} \text{ (Exponentialreihe).}$$

3. $\sum_{n=0}^{\infty} 2^n (x-2)^n$. Mit $a_n = 2^n$, $n \in \mathbb{N}$ gilt hier

$$\lim_{n \rightarrow \infty} \sqrt[n]{|a_n|} = \lim_{n \rightarrow \infty} 2 = 2 \implies \rho = \frac{1}{2}.$$

Diese Potenzreihe konvergiert also für $x \in (3/2, 5/2)$ und divergiert falls $x \in (-\infty, 3/2)$ oder $x \in (5/2, \infty)$.

Ist auch das Konvergenzverhalten in den Punkten $3/2$ und $5/2$ von Interesse, so muss dieses „von Hand“ untersucht werden: Für $x = 3/2$ erhält man die Reihe

$$\sum_{n=0}^{\infty} 2^n \left(-\frac{1}{2}\right)^n = \sum_{n=0}^{\infty} (-1)^n,$$

die offensichtlich divergent ist. Gleiches gilt für $x = 5/2$, das die Reihe $\sum_{n=0}^{\infty} 1$ liefert. Also ist die Reihe an beiden Randpunkten des Konvergenzintervalls divergent.

4. $\sum_{n=1}^{\infty} \frac{2^n}{n} x^{3n}$. Bei dieser Potenzreihe ist eine direkte Anwendung der Hadamard-Formel nicht möglich, da der entsprechende Grenzwert nicht existiert. Mit der Substitution $y = x^3$ haben wir $\sum_{n=1}^{\infty} \frac{2^n}{n} y^n$, also $a_n = \frac{2^n}{n}$, und wir erhalten

$$\lim_{n \rightarrow \infty} \sqrt[n]{|a_n|} = \lim_{n \rightarrow \infty} \frac{2}{\sqrt[n]{n}} = \frac{2}{\lim_{n \rightarrow \infty} \sqrt[n]{n}} = \frac{2}{1} = 2.$$

Das entspricht dem Konvergenzradius $1/2$ für die Reihe $\sum_{n=1}^{\infty} \frac{2^n}{n} y^n$: Es liegt also Konvergenz vor für alle $y \in \mathbb{R}$ mit $|y| < 1/2$ und Divergenz für $|y| > 1/2$. Mit $x = \sqrt[3]{y}$ haben wir damit Konvergenz der ursprünglichen Reihe für alle $|x| < 1/\sqrt[3]{2}$ und Divergenz für $|x| > 1/\sqrt[3]{2}$.

Die Untersuchung an den Rändern führt für $x = 1/\sqrt[3]{2}$ auf die divergente harmonische Reihe und für $x = -1/\sqrt[3]{2}$ auf die konvergente alternierende harmonische Reihe. Das Konvergenzgebiet der Reihe ist daher das halboffene Intervall $[-1/\sqrt[3]{2}, 1/\sqrt[3]{2})$.

5. $\sum_{n=0}^{\infty} n^n x^n$. Hier ist $a_n = n^n$. Wegen

$$\lim_{n \rightarrow \infty} \sqrt[n]{|a_n|} = \lim_{n \rightarrow \infty} n = \infty$$

ist der Konvergenzradius gleich 0, d.h. die Potenzreihe konvergiert nur für $x = 0$.

Bemerkungen 4.6.5: (i) Bei der Anwendung der Formel von Hadamard oft auftretende Grenzwerte:

$$\lim_{n \rightarrow \infty} \sqrt[n]{c} = 1 \quad \text{für alle } c \in (0, \infty), \quad \lim_{n \rightarrow \infty} \sqrt[n]{n} = 1, \quad \lim_{n \rightarrow \infty} \sqrt[n]{n!} = \infty.$$

(ii) Jede Potenzreihe definiert auf ihrem Konvergenzintervall I eine Funktion $f : I \rightarrow \mathbb{R}$ mit $f(x) = \sum_{n=0}^{\infty} a_n(x - x_0)^n$.

(iii) Wichtige Potenzreihen:

$$\begin{aligned} \sum_{n=0}^{\infty} x^n &= \frac{1}{1-x}, & |x| < 1 & \quad (\text{geometrische Reihe}) \\ \sum_{n=0}^{\infty} \frac{x^n}{n!} &= e^x, & x \in \mathbb{R} & \quad (\text{Exponentialreihe}) \\ \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n} &= \cos(x), & x \in \mathbb{R} & \\ \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1} &= \sin(x), & x \in \mathbb{R} & \\ \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} x^n &= \ln(1+x), & |x| < 1. & \end{aligned}$$

Diese unendlichen Reihen als Darstellung (bzw. Definition) spezieller Funktionen eignen sich insbesondere zur Berechnung von Näherungswerten für die entsprechenden Funktionswerte für betragsmäßig kleine Werte von x . Für drei der oben angegebenen Reihen liefert dies folgende Approximationen:

$$\frac{1}{1-x} \approx 1+x \quad \text{mit einem Fehler von } \approx x^2,$$

$$\sin(x) \approx x \quad \text{mit einem Fehler von } \approx x^3/6,$$

$$\cos(x) \approx 1 \quad \text{mit einem Fehler von } \approx x^2/2,$$

$$\cos(x) \approx 1 - x^2/2 \quad \text{mit einem Fehler von } \approx x^4/24.$$

4.7 Grenzwerte von Funktionen und Stetigkeit

Definition 4.7.1. Sei $D \subseteq \mathbb{R}$ und $f : D \rightarrow \mathbb{R}$ eine Funktion.

- (i) f konvergiert gegen a für $x \rightarrow x_0$, falls gilt: Für jede Folge (x_n) mit $x_n \in D \setminus \{x_0\}$ und $x_n \rightarrow x_0$ konvergiert die Folge $(f(x_n))$ gegen a . Man schreibt dann $\lim_{x \rightarrow x_0} f(x) = a$.
- (ii) Sei $x_0 \in D$. Dann heißt f stetig in x_0 , falls $\lim_{x \rightarrow x_0} f(x) = f(x_0)$ gilt. f heißt stetig auf D , falls f in jedem Punkt $x_0 \in D$ stetig ist.

Bemerkungen 4.7.2: (i) Wird die Bedingung in Definition 4.7.1(i) nur für Folgen $x_n \rightarrow x_0$ mit $x_n < x_0$ (bzw. $x_n > x_0$) verlangt, so spricht man vom linksseitigen (bzw. rechtsseitigen) Grenzwert und schreibt $\lim_{x \rightarrow x_0^-} f(x)$ (bzw. $\lim_{x \rightarrow x_0^+} f(x)$). Man beachte auch, dass in (i) nicht $x_0 \in D$ verlangt ist! Man benötigt lediglich, dass es Folgen in D gibt, die gegen x_0 konvergieren. Solche Punkte x_0 heißen auch *Häufungspunkte* von D .

- (ii) Gilt $f(x_n) \rightarrow \infty$ (bzw. $-\infty$) für jede Folge $x_n \rightarrow x_0$, so schreibt man auch $\lim_{x \rightarrow x_0} f(x) = \infty$ (bzw. $-\infty$); entsprechendes gilt für einseitige Grenzwerte.
- (iii) Man betrachtet auch $x_0 = \pm\infty$, d.h. das Verhalten von f für $x \rightarrow \pm\infty$. Gilt etwa $f(x_n) \rightarrow a$ für jede Folge $x_n \rightarrow \infty$, so schreibt man $\lim_{x \rightarrow \infty} f(x) = a$.

Als unmittelbare Folgerung aus den Grenzwertsätzen für Folgen erhält man

Satz 4.7.3. Sei $D \subseteq \mathbb{R}$ und $f, g : D \rightarrow \mathbb{R}$.

- (i) Es sei $\lim_{x \rightarrow x_0} f(x) = a$ und $\lim_{x \rightarrow x_0} g(x) = b$. Dann gilt

$$\lim_{x \rightarrow x_0} (f(x) + g(x)) = a + b, \quad \lim_{x \rightarrow x_0} (f(x)g(x)) = ab, \quad \lim_{x \rightarrow x_0} \frac{f(x)}{g(x)} = \frac{a}{b} \quad \text{falls } b \neq 0.$$

- (ii) Es seien f und g stetig in x_0 . Dann sind $f + g$, fg und f/g (für $g(x_0) \neq 0$) stetig in x_0 .
- (iii) Ist g stetig in $x_0 \in D$ und $h : \tilde{D} \rightarrow \mathbb{R}$ stetig in $y_0 = g(x_0) \in \tilde{D}$, so ist die Verkettung $f = h \circ g$ (also die durch $f(x) := h(g(x))$ definierte Funktion) stetig in x_0 .

Beispiele 4.7.4: (a) $f : \mathbb{R}_0^+ \rightarrow \mathbb{R}$ mit $f(x) = \sqrt{x}$. Ist f stetig? Sei (x_n) eine Folge mit $x_n \geq 0$, $x_n \rightarrow x_0$. Für $x_0 = 0$: $\sqrt{x_n} \rightarrow 0$, denn $x_n \rightarrow 0$ heißt: $\forall \tilde{\varepsilon} > 0 \exists n_{\tilde{\varepsilon}} \in \mathbb{N} : |x_n| \leq \tilde{\varepsilon}$ für alle $n \geq n_{\tilde{\varepsilon}}$. Sei nun $\varepsilon > 0$ gegeben; wähle $\tilde{\varepsilon} = \varepsilon^2$ und dazu das $n_{\tilde{\varepsilon}}$ als n_ε . Dann gilt:

$$n \geq n_\varepsilon \quad \Rightarrow \quad |x_n| \leq \tilde{\varepsilon} = \varepsilon^2 \quad \Rightarrow \quad |\sqrt{x_n}| = \sqrt{x_n} \leq \sqrt{\tilde{\varepsilon}} = \varepsilon, \quad \text{also } (\sqrt{x_n}) \text{ Nullfolge.}$$

Für $x_0 > 0$:

$$|\sqrt{x_n} - \sqrt{x_0}| = \left| \frac{x_n - x_0}{\sqrt{x_n} + \sqrt{x_0}} \right| \leq \frac{|x_n - x_0|}{\sqrt{x_0}} \rightarrow 0, \quad \text{also } \sqrt{x_n} \rightarrow \sqrt{x_0}.$$

Fazit: \sqrt{x} ist stetig auf \mathbb{R}_0^+ .

- (b) $f(x) = \frac{x-1}{\sqrt{x}-1}$; maximaler Definitionsbereich: $D = \mathbb{R}_0^+ \setminus \{1\}$.

Nach Satz 4.7.3 ist f dort stetig (insbesondere: $\sqrt{x} - 1 \neq 0$ für $x \neq 1$).

$$\lim_{x \rightarrow 1} f(x) ? \quad \text{Für } x \neq 1 \text{ ist } f(x) = \frac{(\sqrt{x}-1)(\sqrt{x}+1)}{\sqrt{x}-1} = \sqrt{x} + 1.$$

Also stimmt f auf D mit der *stetigen Funktion* $g : \mathbb{R}_0^+ \rightarrow \mathbb{R}$, $g(x) = \sqrt{x} + 1$ überein!

Damit ist klar: $x_n \rightarrow 1 \Rightarrow f(x_n) = g(x_n) \rightarrow g(1) = 2$. Also $\lim_{x \rightarrow 1} f(x) = 2$.

Bemerkung: f ist also „stetig ergänzbar“ auf ganz \mathbb{R}_0^+ durch $f(1) := 2$.

- (c) $f(x) = 1/x^2$ für $x \neq 0$ hat $\lim_{x \rightarrow 0} f(x) = \infty$.
- (d) $f(x) = 1/x$ für $x \neq 0$ hat $\lim_{x \rightarrow 0^+} f(x) = \infty$ und $\lim_{x \rightarrow 0^-} f(x) = -\infty$.
- (e) $f(x) = e^{\sin(x^2+1)}$ stetig, falls e^x , $\sin x$ etc. stetig sind. Nach Satz 4.7.3 sind alle Polynome stetig. Man kann auch zeigen, dass Potenzreihen im Bereich der absoluten Konvergenz stetig sind; insbesondere also e^x , $\sin x$ und $\cos x$ für $x \in \mathbb{R}$.

(f) $f(x) = \frac{e^x}{x^n}$ mit $n \in \mathbb{N}$ für $x > 0$. Welches Verhalten hat $f(x)$ für $x \rightarrow \infty$? Es ist

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!} \Rightarrow e^x \geq \frac{x^{n+1}}{(n+1)!} \text{ für } x > 0 \Rightarrow f(x) \geq \frac{x}{(n+1)!} \text{ für } x > 0.$$

Damit ist klar, dass $f(x) \rightarrow \infty$ für $x \rightarrow \infty$, also $\lim_{x \rightarrow \infty} \frac{e^x}{x^n} = \infty$ für alle $n \in \mathbb{N}$.

Fazit: „Die e -Funktion wächst stärker als jede Potenz von x “.

Satz 4.7.5. Sei $f : [a, b] \rightarrow \mathbb{R}$ stetig. Dann

(a) besitzt f ein Maximum und ein Minimum in $[a, b]$, d.h. es gibt Punkte $x_1, x_2 \in [a, b]$ mit

$$\left(\min_{a \leq y \leq b} f(y) = \right) f(x_1) \leq f(x) \leq f(x_2) \left(= \max_{a \leq y \leq b} f(y) \right).$$

für alle $x \in [a, b]$.

(b) nimmt f jeden Wert $z \in [f(x_1), f(x_2)]$ an (**Zwischenwertsatz**).

Anwendung. Wir benutzen Satz 4.7.5, um zu zeigen, dass die nichtlineare Gleichung $x = \cos x$ eine Lösung im Intervall $[0, \pi/2]$ besitzt.

Die Funktion $f : [0, \pi/2] \rightarrow \mathbb{R}$ mit $f(x) = x - \cos x$ ist stetig, und es ist $f(0) = -1 < 0$ und $f(\pi/2) = \pi/2 > 0$. Es ist also

$$\min_{0 \leq x \leq \pi/2} f(x) < 0 < \max_{0 \leq x \leq \pi/2} f(x).$$

Nach dem Zwischenwertsatz existiert ein $x^* \in [0, \pi/2]$ mit $f(x^*) = 0$, also $x^* = \cos x^*$.

Bemerkung 4.7.6. Eine Funktion $f : D \rightarrow \mathbb{R}$ heißt *Lipschitz-stetig* (mit der Lipschitz-Konstanten $L \geq 0$), falls

$$|f(x) - f(y)| \leq L|x - y| \quad \text{für alle } x, y \in D.$$

Diese Eigenschaft ist stärker als Stetigkeit.

Ist $L < 1$, heißt f eine *Kontraktion*. Diese Eigenschaft ist u.a. wichtig für die Numerik. Zum Lösen einer Gleichung $x = f(x)$ mit einer Kontraktion f kann man wie folgt vorgehen: Wähle $x_0 \in \mathbb{R}$ beliebig und definiere eine Folge (x_n) rekursiv durch $x_{n+1} := f(x_n)$ für $n \geq 0$. Die Folge (x_n) ist konvergent, und für ihren Grenzwert x^* gilt $x^* = f(x^*)$. Weiter hat man die Fehlerabschätzung:

$$|x_n - x^*| \leq \frac{L^n}{1 - L} |x_1 - x_0|.$$

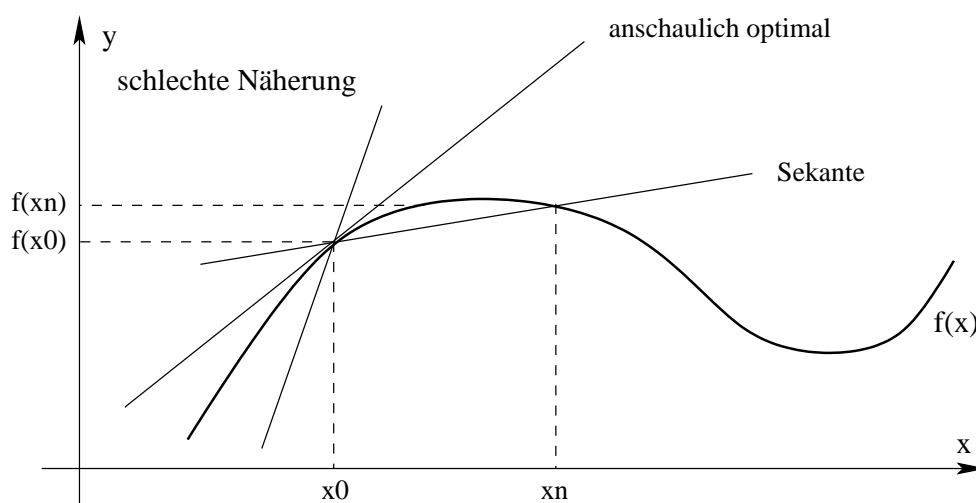
Dieser Sachverhalt ist als *Banachscher Fixpunktsatz* bekannt.

Kapitel 5

Differentiation

5.1 Tangente und Ableitung

An eine ‘glatte’ Kurve kann in jedem Punkt eine Tangente gelegt werden. Ist die Kurve eine Funktion, so wird die Steigung einer solchen Tangente im Punkt x_0 die Ableitung der Funktion im Punkt x_0 genannt. In einer kleinen Umgebung des Punktes ist die Tangente eine gute Approximation der Funktion; Sekanten durch $(x_0, f(x_0))$ stellen schlechtere Näherungen dar.



Da die Tangente durch den Punkt $(x_0, f(x_0))$ verläuft, hat sie die Darstellung $y = l(x)$ mit

$$l(x) = f(x_0) + \alpha(x - x_0).$$

Dabei ist α die Tangenten-Steigung. Der Fehler bei dieser Näherung ist $|f(x) - l(x)|$, und die Näherung ist optimal, falls

$$\lim_{x \rightarrow x_0} \left| \frac{f(x) - l(x)}{x - x_0} \right| = 0$$

(„der Fehler geht schneller gegen Null als $x - x_0$ “).

Die Tangenten-Steigung α kann wie folgt berechnet werden: es ist

$$0 = \lim_{x \rightarrow x_0} \frac{f(x) - l(x)}{x - x_0} = \lim_{x \rightarrow x_0} \left(\frac{f(x) - f(x_0)}{x - x_0} - \alpha \right)$$

und damit

$$\alpha = \lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0}.$$

Nach Definition des Grenzwertes bedeutet dies

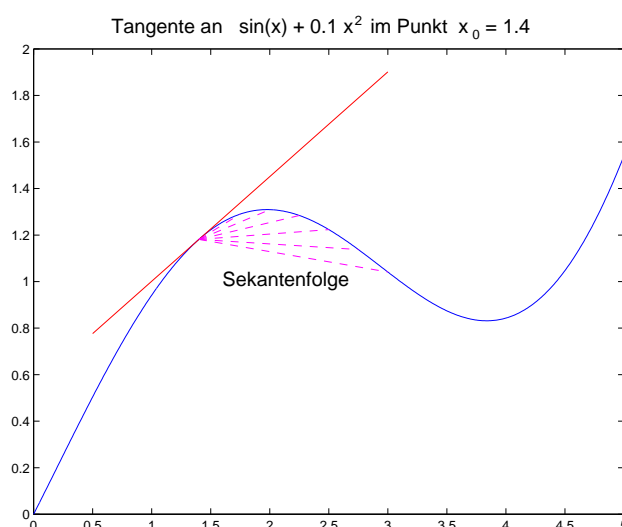
$$\alpha = \lim_{n \rightarrow \infty} \frac{f(x_n) - f(x_0)}{x_n - x_0}$$

für jede Folge (x_n) mit $x_n \rightarrow x_0$, d.h. für jede solche Folge ist α Grenzwert der Steigungen

$$\alpha_n := \frac{f(x_n) - f(x_0)}{x_n - x_0}.$$

Das sind die Steigungen der Sekanten durch den Punkt $(x_0, f(x_0))$ und durch die näher kommenden Nachbarpunkte $(x_n, f(x_n))$.

In der folgenden Zeichnung ist die Folge der Steigungen der eingezeichneten Sekanten $\alpha_n = \{-0.0877, -0.0323, 0.0382, 0.121, 0.2131, 0.3108\}$, die Tangentensteigung ist $\alpha = 0.45$.



Definition 5.1.1 (Ableitung). Die Funktion $f : I \rightarrow \mathbb{R}$ sei auf einem Intervall I definiert. Dann heißt f *differenzierbar* in $x_0 \in I$, falls der Grenzwert

$$\lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} \left(= \lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0)}{h} \right)$$

existiert. Diesen Grenzwert bezeichnet man als *Ableitung von f in x_0* , und man schreibt dafür $f'(x)$ oder $\frac{df}{dx}(x_0)$. Die Funktion f heißt *differenzierbar*, wenn f in jedem $x_0 \in I$ differenzierbar ist.

Bemerkungen 5.1.2: (i) $\frac{df}{dx}$ wird *Differentialquotient* genannt, ist aber kein *echter* Quotient, sondern Grenzwert von Differenzenquotienten $\frac{\Delta f}{\Delta x}$ für $\Delta x \rightarrow 0$.

(ii) Aufgrund der Vorüberlegungen ist es klar, dass genau dann eine eindeutige Tangente an den Graphen von f im Punkt $(x_0, f(x_0))$ existiert, wenn f in x_0 differenzierbar ist. Die Tangentengleichung lautet dann $y = f(x_0) + f'(x_0)(x - x_0)$, und es gilt

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + r(x; x_0) \quad \text{mit} \quad \lim_{x \rightarrow x_0} \frac{|r(x; x_0)|}{|x - x_0|} = 0.$$

Diese Beziehung ist äquivalent zur Differenzierbarkeit von f in x_0 . Als Folgerung erhält man sofort die Stetigkeit von f in x_0 , denn für (x_n) mit $x_n \rightarrow x_0$ folgt:

$$f(x_n) = f(x_0) + f'(x_0)(x_n - x_0) + (x_n - x_0) \frac{r(x_n; x_0)}{x_n - x_0} \rightarrow f(x_0).$$

Differenzierbarkeit ist also eine stärkere Eigenschaft als Stetigkeit.

(iii) Ist $x_0 \in I$ ein Randpunkt von I , so ist in Definition 5.1.1 der entsprechende einseitige Grenzwert zu verwenden. Man schreibt dann

$$f'_+(x_0) = \lim_{h \rightarrow 0^+} \frac{f(x_0 + h) - f(x_0)}{h} \quad \text{bzw.} \quad f'_-(x_0) = \lim_{h \rightarrow 0^-} \frac{f(x_0 + h) - f(x_0)}{h}$$

und nennt $f'_\pm(x_0)$ die *einseitigen Ableitungen* von f im Punkt x_0 .

(iv) Ableitungen treten in Anwendungen bei der Betrachtung kontinuierlicher dynamischer Prozesse auf. Ist etwa $s(t)$ der Ort eines Objekts zum Zeitpunkt t , so ist $v(t) := s'(t)$ die Geschwindigkeit zur Zeit t .

Beispiele 5.1.3: (a) Sei $f(x) = x^2$ und $x_0 \in \mathbb{R}$ beliebig. Dann ist

$$\frac{f(x_0 + h) - f(x_0)}{h} = \frac{(x_0 + h)^2 - x_0^2}{h} = \frac{x_0^2 + 2x_0h + h^2 - x_0^2}{h} = 2x_0 + h,$$

also

$$\lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0)}{h} = \lim_{h \rightarrow 0} (2x_0 + h) = 2x_0.$$

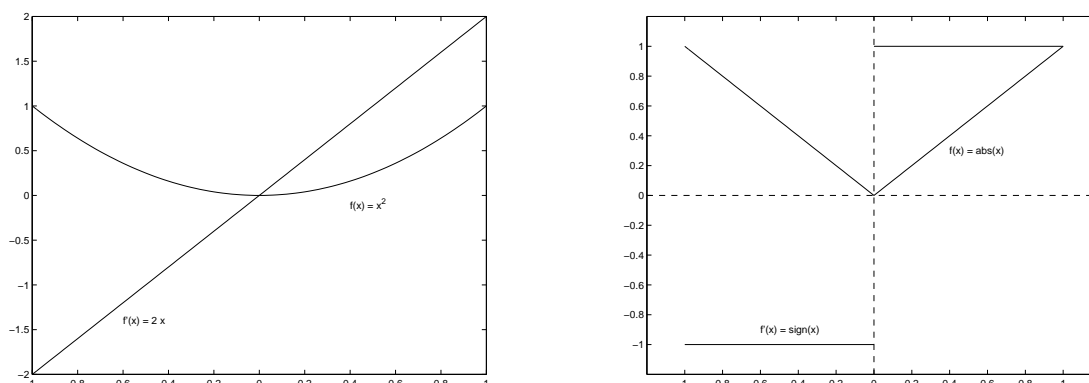
Damit ist f in jedem Punkt x_0 differenzierbar und $f'(x_0) = 2x_0$.

(b) Sei $f(x) = |x|$ und $x_0 = 0$. Dann ist $\frac{f(x_0 + h) - f(x_0)}{h} = \frac{|h|}{h}$, und wir unterscheiden zwei Fälle:

$$\text{Fall 1: Folge } h_n \text{ mit } h_n \geq 0 \text{ oder } h_n \rightarrow 0^+ \Rightarrow \frac{|h_n|}{h_n} = 1 \rightarrow 1$$

$$\text{Fall 2: Folge } h_n \text{ mit } h_n \leq 0 \text{ oder } h_n \rightarrow 0^- \Rightarrow \frac{|h_n|}{h_n} = \frac{-h_n}{h_n} = -1 \rightarrow -1.$$

Zusammen heißt das: $f'(x_0)$ existiert nicht, aber es gibt eine rechtsseitige und eine linksseitige Ableitung. Dies ist auch klar am Graphen der Funktion abzulesen.



Ableitungen wichtiger Funktionen

1. $f(x) = c$ ist differenzierbar mit $f'(x) = 0$.
2. $f(x) = x^n$ mit $n \in \mathbb{N}$ ist differenzierbar mit $f'(x) = nx^{n-1}$.
Dies ist die Verallgemeinerung von Beispiel 5.1.3 (a) und kann mit Hilfe der binomischen Formel bewiesen werden.
3. $f(x) = x^r$ für $x > 0$ und $r \in \mathbb{R}$ ist differenzierbar mit $f'(x) = rx^{r-1}$.
Zum Beweis schreibt man $x^r = e^{r \ln x}$ und benutzt die Kettenregel, die wir später kennenlernen.
4. $f(x) = \sin x$ für $x \in \mathbb{R}$ ist differenzierbar mit $f'(x) = \cos x$.
Anschaulich: Die Sinus-Funktion steigt bei Null mit einer 45-Grad-Tangente an ($\cos(0) = 1$) und flacht allmählich so ab, wie die Werte der Cosinus-Funktion kleiner werden.
5. $f(x) = \cos x$ ist für $x \in \mathbb{R}$ differenzierbar mit $f'(x) = -\sin x$.
Anschaulich: Die Cosinus-Funktion beginnt bei $x = 0$ mit einer waagerechten Tangente ($\sin(0) = 0$) und fällt dann so, wie die Werte der Sinus-Funktion ansteigen.
6. $f(x) = e^x$ ist für $x \in \mathbb{R}$ differenzierbar mit $f'(x) = e^x$.

Zur Erinnerung: Die Exponentialfunktion ist definiert als Potenzreihe

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}.$$

Diese Reihe konvergiert für jedes $x \in \mathbb{R}$ absolut. Man kann allgemein zeigen, dass Potenzreihen im Innern des Konvergenzbereiches (hier also: für jedes $x \in \mathbb{R}$) „gliedweise differenziert“ werden können, d.h.

$$(e^x)' = \left(\sum_{n=0}^{\infty} \frac{x^n}{n!} \right)' = \sum_{n=0}^{\infty} \left(\frac{x^n}{n!} \right)'$$

Wegen $(x^n)' = nx^{n-1}$ für $n \in \mathbb{N}$ und $(x^0)' = 1' = 0$ folgt

$$(e^x)' = \sum_{n=1}^{\infty} \frac{1}{n!} nx^{n-1} = \sum_{n=1}^{\infty} \frac{x^{n-1}}{(n-1)!} = \sum_{k=0}^{\infty} \frac{x^k}{k!} = e^x.$$

7. $f(x) = \ln x$ für $x > 0$ ist differenzierbar mit $f'(x) = x^{-1}$. Der Beweis erfolgt mittels der Ableitungsformel für Umkehrfunktionen, die wir später einführen werden.

5.2 Differentiationsregeln

Satz 5.2.1. Seien f, g in x_0 differenzierbar. Dann sind auch $c_1f + c_2g$ (mit $c_1, c_2 \in \mathbb{R}$), $f \cdot g$ und f/g (falls $g(x_0) \neq 0$) differenzierbar in x_0 und es gilt:

- (i) $(c_1f + c_2g)'(x_0) = c_1f'(x_0) + c_2g'(x_0)$,
 (ii) $(f \cdot g)'(x_0) = f'(x_0)g(x_0) + g'(x_0)f(x_0)$ (**Produktregel**),
 (iii) $\left(\frac{f}{g}\right)'(x_0) = \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{g(x_0)^2}$ (**Quotientenregel**).

Folgerungen 5.2.2: (a) Aus (i) und der Differenzierbarkeit von x^n für $n \in \mathbb{N}_0$ folgt, dass alle Polynome differenzierbar sind. Dann sind nach (iii) auch alle rationalen Funktionen (also $P(x)/Q(x)$ mit Polynomen P, Q) differenzierbar (auf ihrem maximalen Definitionsbereich, d.h. außerhalb der Nullstellen von Q).

- (b) Aus (iii) folgt $\left(\frac{1}{g}\right)'(x_0) = -\frac{g'(x_0)}{g(x_0)^2}$ für differenzierbares g mit $g(x_0) \neq 0$.

Beispiele 5.2.3: 1. Offenbar ist: $x' = 1$, also folgt

$$(x^2)' = (x \cdot x)' = 1 \cdot x + x \cdot 1 = 2x, \quad (x^3)' = (x \cdot x^2)' = 2x \cdot x + x^2 \cdot 1 = 3x^2 \text{ etc.}$$

Induktiv erhält man so einen Beweis für $(x^n)' = nx^{n-1}$.

2. $f(x) = e^x \sin x \Rightarrow f'(x) = e^x \sin x + e^x \cos x = e^x(\sin x + \cos x)$.

3. $f(x) = \frac{x^3 + 1}{x^2 - 1}$ für $x \in \mathbb{R} \setminus \{-1, 1\}$. Dann ist

$$\begin{aligned} f'(x) &= \frac{(3x^2)(x^2 - 1) - (x^3 + 1)2x}{(x^2 - 1)^2} = \frac{3x^4 - 3x^2 - 2x^4 - 2x}{(x^2 - 1)^2} \\ &= \frac{x^4 - 3x^2 - 2x}{(x^2 - 1)^2} = \frac{x^2 - 2x}{(x - 1)^2}. \end{aligned}$$

4. $f(x) = \tan x = \frac{\sin x}{\cos x}$ für $x \in \mathbb{R} \setminus \{\frac{\pi}{2} + k\pi : k \in \mathbb{Z}\}$. Dann ist

$$f'(x) = \frac{\cos x \cos x - \sin x(-\sin x)}{\cos^2 x} = \frac{1}{\cos^2 x}.$$

5. $f(x) = x \ln x - x \Rightarrow f'(x) = (x \ln x)' - 1 = 1 \cdot \ln x + x \cdot \frac{1}{x} - 1 = \ln x$.

Satz 5.2.4 (Kettenregel). Es seien f im Punkt x_0 und g im Punkt $y_0 = f(x_0)$ differenzierbar. Dann ist die verkettete Funktion $h(x) := g(f(x))$ im Punkt x_0 differenzierbar mit

$$\boxed{h'(x_0) = g'(f(x_0)) f'(x_0)}.$$

Bemerkung 5.2.5. Für die verkettete Funktion h schreibt man auch $g \circ f$, d.h.

$$(g \circ f)(x) := g(f(x)).$$

Man spricht auch von hintereinander geschalteten Funktionen. Damit schreibt sich die Kettenregel als

$$(g \circ f)'(x_0) = g'(f(x_0)) f'(x_0).$$

Dabei heißt $g'(f(x_0))$ die *äußere Ableitung* und der Faktor $f'(x_0)$ *innere Ableitung*.

Beispiele 5.2.6: (a) Es ist $\cos x = \sin\left(x + \frac{\pi}{2}\right)$, also $\cos x = g(f(x))$ mit $g(y) = \sin y$ und $f(x) = x + \frac{\pi}{2}$. Nach der Kettenregel folgt

$$(\cos x)' = g'(f(x)) f'(x) = \cos\left(x + \frac{\pi}{2}\right) \cdot 1 = \cos\left(x + \frac{\pi}{2}\right) = -\sin x.$$

(b) Es ist $x^r = e^{r \ln x}$, also $x^r = g(f(x))$ mit $g(y) = e^y$ und $f(x) = r \ln x$.

Nach der Kettenregel folgt

$$(x^r)' = g'(f(x)) f'(x) = e^{f(x)} r \frac{1}{x} = e^{r \ln x} \frac{r}{x} = x^r \frac{r}{x} = r x^{r-1}.$$

(c) Sei $h(x) = \sin \sqrt{x^2 + 1}$. Berechne $h'(x)$.

Es ist $h(x) = g(f(x))$ mit $g(y) = \sin y$ und $f(x) = \sqrt{x^2 + 1}$.

$$\Rightarrow h'(x) = \cos(f(x)) f'(x).$$

Wie lautet $f'(x)$? Es ist $f(x)$ wiederum eine verkettete Funktion: $f(x) = u(v(x))$ mit $u(y) = \sqrt{y}$ und $v(x) = x^2 + 1$. Also gilt (mit $(\sqrt{y})' = \frac{1}{2\sqrt{y}}$)

$$f'(x) = \frac{1}{2\sqrt{v(x)}} v'(x) = \frac{2x}{2\sqrt{x^2 + 1}} = \frac{x}{\sqrt{x^2 + 1}}$$

$$\text{Fazit: } h'(x) = \frac{x}{\sqrt{x^2 + 1}} \cos \sqrt{x^2 + 1}.$$

Ableitung von Umkehrfunktionen

Den Definitions- und Wertebereich einer Funktion f bezeichnen wir wieder mit D_f bzw. W_f . Unter einer Umgebung U um einen Punkt $x_0 \in \mathbb{R}$ versteht man ein Intervall der Form $(x_0 - \epsilon, x_0 + \epsilon)$ für ein $\epsilon > 0$.

Satz 5.2.7 (Ableitung der Umkehrfunktion). Sei $f : D_f \rightarrow W_f$ ($D_f, W_f \subseteq \mathbb{R}$) eine umkehrbare Funktion, die in einer Umgebung U um den Punkt $x_0 \in D_f$ differenzierbar ist. Es gelte $f'(x_0) \neq 0$. Dann ist f^{-1} im Punkt $y_0 = f(x_0)$ differenzierbar und

$$(f^{-1})'(y_0) = \frac{1}{f'(f^{-1}(y_0))} \quad \left(= \frac{1}{f'(x_0)} \right).$$

Beweisidee: Sei $g = f^{-1}$. Dann gilt $g(f(x)) = x$ für alle $x \in D_f$. Falls g in y_0 differenzierbar ist, liefert die Kettenregel:

$$(g \circ f)'(x_0) = g'(f(x_0)) f'(x_0) = 1,$$

also mit $y_0 = f(x_0)$:

$$(f^{-1})'(y_0) = \frac{1}{f'(x_0)} = \frac{1}{f'(f^{-1}(y_0))}.$$

Beispiele 5.2.8: 1. $f(x) = e^x$ mit $D_f = \mathbb{R}$, $W_f = (0, \infty)$. Hier ist $f^{-1}(y) = \ln y$ für alle $y > 0$. Es ist $f'(x) = e^x > 0$ für alle $x \in \mathbb{R}$ und damit nach Satz 5.2.7

$$\ln'(y) = (f^{-1})'(y) = \frac{1}{f'(\ln y)} = \frac{1}{e^{\ln y}} = \frac{1}{y} \quad \text{für } y > 0.$$

2. $f(x) = \sin x$ mit $D_f = (-\pi/2, \pi/2)$, $W_f = (-1, 1)$. Hier ist $f^{-1}(y) = \arcsin y$ für $y \in (-1, 1)$. Es ist $f'(x) = \cos x \neq 0$ in D_f , also

$$\arcsin'(y) = \frac{1}{f'(\arcsin y)} = \frac{1}{\cos(\arcsin y)} = \frac{1}{\sqrt{1 - \sin^2(\arcsin y)}} = \frac{1}{\sqrt{1 - y^2}}.$$

3. $f(x) = x + e^x$ mit $D_f = \mathbb{R}$, $W_f = \mathbb{R}$. Die Funktion f ist streng monoton wachsend, also injektiv. Damit ist f auf dem Wertebereich W_f invertierbar. Sei f^{-1} die Umkehrfunktion. Welchen Wert hat $(f^{-1})'(1)$?

Problem: Für f^{-1} kann keine geschlossene Formel angegeben werden.

Ausweg: Um Satz 5.2.7 anzuwenden, brauchen wir nur den Punkt $x_0 \in D_f$ mit $f(x_0) = 1$ ($=: y_0$) zu finden! In diesem Beispiel gilt offensichtlich $x_0 = 0$. Also folgt

$$(f^{-1})'(y_0) = \frac{1}{f'(x_0)} = \frac{1}{1 + e^{x_0}} = \frac{1}{2}.$$

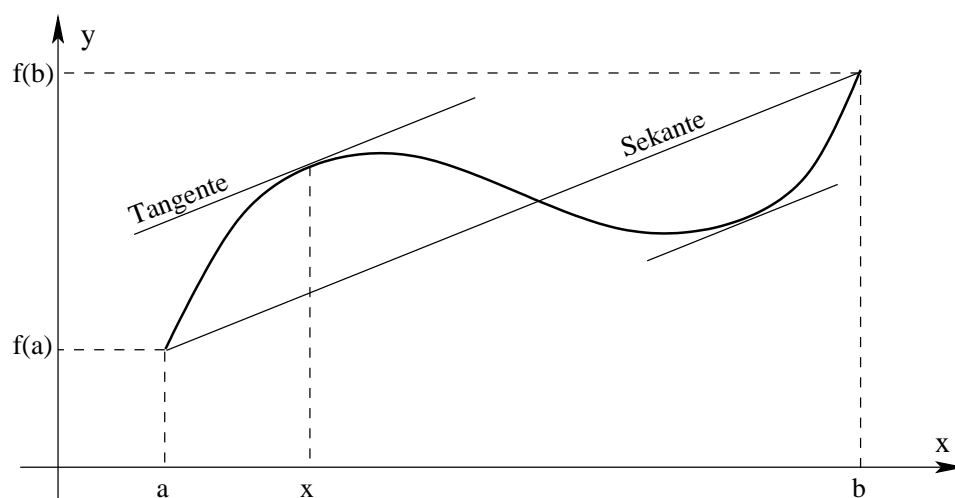
5.3 Mittelwertsatz (der Differentialrechnung)

Satz 5.3.1. Sei $f : [a, b] \rightarrow \mathbb{R}$ stetig und auf (a, b) differenzierbar. Dann gilt

(i) Ist $f(a) = f(b)$, so existiert ein $\xi \in (a, b)$ mit $f'(\xi) = 0$. (**Satz von Rolle**)

(ii) Es gibt ein $\xi \in (a, b)$ mit $f'(\xi) = \frac{f(b) - f(a)}{b - a}$. (**Mittelwertsatz, kurz MWS**)

Bemerkung 5.3.2. Der Satz von Rolle ist ein Spezialfall des Mittelwertsatzes. Anschaulich sagt der Mittelwertsatz, dass es im Intervall (a, b) mindestens eine Stelle gibt, an der die Tangente parallel zur Sekante (durch die Endpunkte) verläuft:



Man sieht an der Skizze, dass die Stelle x i.a. nicht eindeutig bestimmt ist.

Der Mittelwertsatz wird oft in einer der folgenden Formulierungen verwendet:

- $f(y) = f(x) + f'(\xi)(y - x)$ für eine Zwischenstelle ξ , d.h. $x < \xi < y$ bzw. $y < \xi < x$.

Mit $h := y - x$ wird daraus

- $f(x + h) = f(x) + f'(x + \theta h)h$ für ein $\theta \in (0, 1)$.

Korollar 5.3.3: Sei $f : [a, b] \rightarrow \mathbb{R}$ stetig und in (a, b) differenzierbar. Dann gilt

- (i) $f'(x) = 0$ für alle $x \in (a, b) \Rightarrow f$ ist konstant in $[a, b]$.
- (ii) f ist monoton wachsend (fallend) in $[a, b] \Leftrightarrow f'(x) \geq 0$ (≤ 0) in (a, b) .
- (iii) $f'(x) > 0$ (< 0) in $(a, b) \Rightarrow f$ ist streng monoton wachsend (fallend).

Wir beweisen nicht das ganze Korollar. Zur Illustration der Anwendung des Mittelwertsatz aber Teil (iii).

Beweis von (iii). Zu zeigen ist:

$$x, y \in [a, b] \text{ mit } x < y \Rightarrow f(x) < f(y).$$

Nach dem Mittelwertsatz gilt

$$f(y) = f(x) + f'(\xi)(y - x) \text{ mit einem } x < \xi < y.$$

Nach Voraussetzung ist $f'(\xi) > 0$. Wegen $y - x > 0$ folgt $f'(\xi)(y - x) > 0$, und damit

$$f(y) = f(x) + f'(\xi)(y - x) > f(x).$$

□

Beispiele 5.3.4: (a) Zeige: $\cos x \geq 1 - \frac{1}{2}x^2$ für alle $x \in \mathbb{R}$.

Beweis 1:

Die Aussage ist trivial für $|x| > 2$. Sei also $|x| \leq 2$. Wir betrachten die Potenzreihe für die Cosinus-Funktion (siehe Abschnitt 4.5, Seite 44):

$$\cos(x) = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!} = 1 - \frac{x^2}{2} + \frac{x^4}{24} - \frac{x^6}{720} \pm \dots$$

Dies ist eine alternierende Reihe, deren Glieder für alle $x \leq 2$ eine monoton fallende Nullfolge bilden. Also ist

$$\cos(x) \geq 1 - \frac{x^2}{2}.$$

Beweis 2:

Setze $f(x) := \cos x - 1 + \frac{1}{2}x^2$; zu zeigen ist dann $f(x) \geq 0$ für $x \in \mathbb{R}$. Da f eine gerade Funktion ist, reicht es, $f(x) \geq 0$ für $x \geq 0$ zu zeigen. Es gilt $f(0) = \cos(0) - 1 = 0$ und $f'(x) = -\sin x + x \geq 0$ für $x \geq 0$. Nach Korollar 5.3.3 ist dann f monoton wachsend auf \mathbb{R}_+ . Also folgt $f(x) \geq f(0) = 0$ für $x \geq 0$.

(b) Welche differenzierbaren Funktionen $f : \mathbb{R} \rightarrow \mathbb{R}$ genügen der Beziehung

$$(*) \quad f'(x) = af(x) \quad \text{mit einer Konstanten } a \in \mathbb{R}?$$

Wir wissen $(e^x)' = e^x$ und nach der Kettenregel folgt $(e^{ax})' = ae^{ax}$, d.h. die Funktionen $f(x) = ce^{ax}$ (mit beliebiger Konstante $c \in \mathbb{R}$) erfüllen die Gleichung (*). Umgekehrt gilt die folgende

Behauptung: Die einzigen differenzierbaren Funktionen $f : \mathbb{R} \rightarrow \mathbb{R}$, die (*) erfüllen sind durch $f(x) = ce^{ax}$ mit einer Konstante $c \in \mathbb{R}$ gegeben.

Beweis. Gegeben sei eine Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$, die (*) erfüllt. Setze $g(x) := f(x)e^{-ax}$. Dann gilt

$$g'(x) = f'(x)e^{-ax} + f(x)e^{-ax} \cdot (-a) = e^{-ax}(f'(x) - af(x)) = 0 \quad \text{auf } \mathbb{R}.$$

Nach Korollar 5.3.3 folgt: g ist konstant, also $g(x) = c$ für ein $c \in \mathbb{R}$.

Also ist $f(x)e^{-ax} = c$ und damit $f(x) = ce^{ax}$. □

Fazit: Sei $a \in \mathbb{R}$ gegeben, und $f : \mathbb{R} \rightarrow \mathbb{R}$ differenzierbar. Dann gilt

$$\boxed{f'(x) = af(x) \text{ auf } \mathbb{R} \quad \Leftrightarrow \quad f(x) = ce^{ax} \text{ mit } c \in \mathbb{R}}.$$

Eine weitere Folgerung aus dem Mittelwertsatz ist

Satz 5.3.5 (Regel von de L'Hôpital). Seien f und g in (a, b) differenzierbar und $g'(x) \neq 0$ für $x \in (a, b)$. Gilt

$$\lim_{x \rightarrow a+} f(x) = 0, \quad \lim_{x \rightarrow a+} g(x) = 0, \quad \text{und} \quad \lim_{x \rightarrow a+} \frac{f'(x)}{g'(x)} \quad \text{existiert,}$$

so folgt

$$\boxed{\lim_{x \rightarrow a+} \frac{f(x)}{g(x)} = \lim_{x \rightarrow a+} \frac{f'(x)}{g'(x)}}$$

Bemerkung 5.3.6. An Stellen $x_0 \in (a, b)$ gilt der entsprechende Satz mit den üblichen zweiseitigen Grenzwerten, im Punkt b entsprechend mit linksseitigen Grenzwerten. Der Satz bleibt auch richtig, wenn die Voraussetzung

$$\lim_{x \rightarrow a^+} f(x) = \lim_{x \rightarrow a^+} g(x) = 0$$

durch

$$\lim_{x \rightarrow a^+} f(x) = \lim_{x \rightarrow a^+} g(x) = \infty$$

ersetzt wird.

Beispiele 5.3.7: (a) $\lim_{x \rightarrow 0} \frac{\sin x}{x} = ?$ Es ist $\lim_{x \rightarrow 0} \sin x = 0 = \lim_{x \rightarrow 0} x$.

Prüfe, ob $\lim_{x \rightarrow 0} \frac{\sin x}{x}$ existiert: $\frac{(\sin x)'}{x'} = \frac{(\cos x)}{1} \rightarrow 1$ für $x \rightarrow 0$

Also gilt $\lim_{x \rightarrow 0} \frac{\sin x}{x} = 1$ nach Satz 5.3.5.

(b) $\lim_{x \rightarrow 0} \frac{1 - \cos x}{x^2} = ?$ Existiert $\lim_{x \rightarrow 0} \frac{(1 - \cos x)'}{(x^2)'} = \lim_{x \rightarrow 0} \frac{\sin x}{2x}$?

Nach (a) gilt $\lim_{x \rightarrow 0} \frac{\sin x}{2x} = \frac{1}{2}$, und daraus folgt $\lim_{x \rightarrow 0} \frac{1 - \cos x}{x^2} = \frac{1}{2}$.

(c) Beachten Sie, dass die Regel von de l'Hôpital falsche Ergebnisse liefern kann, wenn sie auf Grenzwerte angewendet wird, die nicht die Voraussetzungen erfüllen. Beispielsweise gilt nach den Grenzwertsätzen

$$\lim_{x \rightarrow 0} \frac{e^x - 1}{\sin(x) - 1} = \frac{e^0 - 1}{\sin(0) - 1} = \frac{0}{-1} = -1.$$

Wendet man allerdings (ungerechtfertigterweise) den Hôpital an, bekommt man mit

$$\lim_{x \rightarrow 0} \frac{e^x - 1}{\sin(x) - 1} \stackrel{???}{=} \lim_{x \rightarrow 0} \frac{e^x}{\cos(x)} = \frac{e^0}{\cos(0)} = \frac{1}{1} = 1$$

ein falsches Ergebnis.

Im Beispiel (b) haben wir eigentlich zweite Ableitungen verwendet!

Definition 5.3.8. Sei f in einer Umgebung U um den Punkt x_0 differenzierbar. Ist die Ableitung f' in dem Punkt x_0 differenzierbar, d.h. existiert

$$\lim_{h \rightarrow 0} \frac{f'(x_0 + h) - f'(x_0)}{h} =: f''(x_0),$$

so heißt f *zweimal differenzierbar in x_0* und $f''(x_0)$ die zweite Ableitung von f in x_0 . Analog definiert man höhere Ableitungen $f'''(x_0)$, $f^{(4)}(x_0)$, \dots , $f^{(n)}(x_0)$.

5.4 Extrema und Wendepunkte

Definition 5.4.1. Sei $f : (a, b) \rightarrow \mathbb{R}$ und $x_0 \in (a, b)$.

1. Dann hat f in x_0 ein *globales Maximum (Minimum)*, falls

$$f(x) \leq f(x_0) \quad (\text{bzw.} \quad f(x) \geq f(x_0)) \quad \text{für alle } x \in (a, b)$$

gilt.

2. Dann hat f in x_0 ein *lokales Maximum (Minimum)*, falls eine Umgebung U um x_0 existiert, so dass

$$f(x) \leq f(x_0) \quad (\text{bzw.} \quad f(x) \geq f(x_0)) \quad \text{für alle } x \in U$$

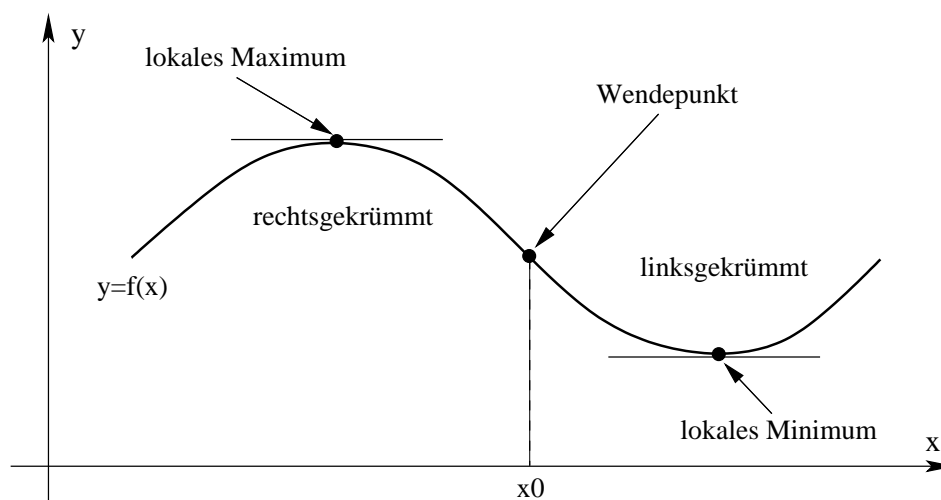
gilt. Sind die Ungleichungen für $x \neq x_0$ strikt, so spricht man von einem *strengen lokalen Maximum (Minimum)*. Unter dem Begriff *Extremum* versteht man ein Minimum oder ein Maximum.

Um Extremstellen einer differenzierbaren Funktion zu finden, verwendet man

Satz 5.4.2. Sei $f : (a, b) \rightarrow \mathbb{R}$ in einer Umgebung des Punktes $x_0 \in (a, b)$ differenzierbar.

- (i) Hat f in x_0 ein lokales Extremum, so gilt $f'(x_0) = 0$.
- (ii) Ist f in x_0 zweimal differenzierbar, gilt $f'(x_0) = 0$ und $f''(x_0) \neq 0$, so besitzt f im Punkt x_0 ein strenges lokales Extremum und zwar ein Maximum falls $f''(x_0) < 0$ bzw. ein Minimum falls $f''(x_0) > 0$.

Bemerkungen 5.4.3: (i) $f'(x_0) = 0$ ist *notwendig* für ein Extremum in x_0 , aber allein *nicht hinreichend*, d.h. x_0 muss dann keine Extremstelle sein! Z.B. hat $f(x) = x^3$ kein Extremum in $x_0 = 0$, obwohl $f'(0) = 0$ gilt.



- (ii) Der Zusammenhang zwischen dem Vorzeichen von $f''(x_0)$ und der Art des Extremums in Teil(b) erklärt sich durch die geometrische Bedeutung der zweiten Ableitung: Ist $f''(x_0) > 0$ (bzw. < 0), so ist der Graph von f an dieser Stelle linksgekrümmt (bzw. rechtsgekrümmt). Ein Punkt in dem die Krümmung wechselt heißt *Wendepunkt*.

Notwendig für einen Wendepunkt von f in x_0 ist $f''(x_0) = 0$, hinreichend ist $f''(x_0) = 0$ und $f'''(x_0) \neq 0$.

Beispiel 5.4.4: Gesucht sind die lokalen Extrema und Wendepunkte von $f(x) = xe^{-x}$.

Berechne f' : $f'(x) = e^{-x} + xe^{-x} \cdot (-1) = e^{-x}(1 - x)$

Also $f'(x_0) = 0 \Leftrightarrow x_0 = 1$; d.h. $x_0 = 1$ ist einziger Kandidat für eine Extremstelle.

Berechne f'' : $f''(x) = -e^{-x}(1 - x) + e^{-x} \cdot (-1) = e^{-x}(x - 2)$

$\Rightarrow f''(x_0) = e^{-1}(1 - 2) = -\frac{1}{e} < 0 \Rightarrow f$ hat strenges lokales Maximum in $x_0 = 1$.

Wendepunkte: $f''(x_1) = 0 \Leftrightarrow x_1 = 2$; einziger Kandidat für einen Wendepunkt.

Berechne f''' : $f'''(x) = -e^{-x}(x - 2) + e^{-x} \cdot 1 = e^{-x}(3 - x)$

$\Rightarrow f'''(x_1) = e^{-2}(3 - 2) = \frac{1}{e^2} > 0 \Rightarrow f$ hat Wendepunkt in $x_1 = 2$.

5.5 Der Satz von Taylor

Eine Motivation bei der Einführung der Ableitung am Anfang des Kapitels war es eine Gerade zu finden, die den Funktionsgraphen in der Nähe einer Stelle x_0 möglichst gut annähert. Es stellte sich heraus, dass diese Gerade die Tangente an den Funktionsgraphen ist, welche durch die Gleichung

$$g(x) = f(x_0) + f'(x_0)(x - x_0)$$

gegeben ist. Entsprechend kann man nun natürlich die Frage nach der besten Annäherung durch eine Parabel, ein Polynom 3., 4., ... Grades stellen. Diese Frage klärt der

Satz 5.5.1 (Satz von Taylor). Seien I ein Intervall, $f : I \rightarrow \mathbb{R}$ eine $(n+1)$ -mal differenzierbare Funktion und $x, x_0 \in I$. Dann gibt es ein ξ zwischen x_0 und x mit

$$f(x) = \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k + \frac{f^{(n+1)}(\xi)}{(n+1)!} (x - x_0)^{n+1}.$$

Bemerkungen 5.5.2: (i) Das Polynom

$$\sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2!} (x - x_0)^2 + \cdots + \frac{f^{(n)}(x_0)}{n!} (x - x_0)^n$$

heißt *Taylorpolynom n -ter Ordnung* von f . Es ist unter allen Polynomen vom Grade n das, das die Funktion f in der Nähe von x_0 am besten approximiert.

(ii) Für $n = 0$ ist der Satz von Taylor gerade der Mittelwertsatz.

(iii) Der Term

$$\frac{f^{(n+1)}(\xi)}{(n+1)!} (x - x_0)^{n+1},$$

der den Fehler bei der Näherung von f durch das Taylorpolynom ausdrückt, heißt *Lagrange-Restglied*.

Beispiel 5.5.3 (Näherungsweise Berechnung von $1,05^{1,02}$ mit Fehlerabschätzung):

Es sei $f : (0, \infty) \rightarrow \mathbb{R}$ gegeben durch $f(x) = x^{1,02}$. Dann ist

$$f'(x) = 1,02 \cdot x^{0,02}, \quad f''(x) = 1,02 \cdot 0,02x^{-0,98} = 0,0204 \cdot x^{-0,98}.$$

Nach dem Satz von Taylor gibt es nun für jedes $x > 0$ ein ξ zwischen x und $x_0 = 1$ mit

$$f(x) = f(1) + f'(1)(x-1) + \frac{f''(\xi)}{2!}(x-1)^2 = 1 + 1,02 \cdot (x-1) + 0,0102 \cdot \xi^{-0,98}(x-1)^2.$$

Für $x = 1,05$ erhalten wir so

$$1,05^{1,02} = f(1,05) = \underbrace{1 + 1,02 \cdot 0,05}_{\text{Näherung}} + \underbrace{0,0102 \cdot \xi^{-0,98} \cdot 0,05^2}_{\text{Fehler}}$$

Wir können nun den Fehler abschätzen, denn wir wissen, dass ξ zwischen 1 und 1,05 liegt, d.h. es ist $\xi \geq 1$ und damit $\xi^{-0,98} \leq 1^{-0,98} = 1$:

$$|0,0102 \cdot \xi^{-0,98} \cdot 0,05^2| = 2,55 \cdot 10^{-5} \cdot \xi^{-0,98} \leq 2,55 \cdot 10^{-5}.$$

Zusammengenommen haben wir also $1,05^{1,02} \approx 1,051$ mit einem Fehler, der kleiner ist als $2,55 \cdot 10^{-5}$.

Haben wir eine beliebig oft differenzierbare Funktion, so gibt uns der Satz von Taylor für jedes $n \in \mathbb{N}$ ein Taylorpolynom vom Grade n . Es ist naheliegend, den Grad immer größer werden zu lassen und damit zu einer Potenzreihe überzugehen. Dieser geben wir zunächst einmal einen Namen.

Definition 5.5.4. Es seien $I \subseteq \mathbb{R}$ ein Intervall, $f : I \rightarrow \mathbb{R}$ beliebig oft differenzierbar und $x_0 \in I$. Die Potenzreihe

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

heißt *Taylorreihe* von f .

Hat diese Reihe einen Konvergenzradius $\varrho > 0$ und gilt $f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k$ für alle $x \in (x_0 - \varrho, x_0 + \varrho)$, so sagt man, dass die Taylorreihe f *darstellt*. Funktionen, die sich durch eine Taylorreihe darstellen lassen, heißen *analytisch*.

Bemerkungen 5.5.5: (i) Es kann passieren, dass die Taylorreihe einer Funktion den Konvergenzradius Null hat und damit nutzlos ist. Genauso gibt es den Fall, dass die Taylorreihe zwar konvergiert, aber nicht gegen f (sondern gegen eine andere Funktion).

(ii) Zum Nachweis, dass die Taylorreihe wirklich gegen f konvergiert, zeigt man, dass die Restglieder des jeweils n -ten Taylorpolynoms gegen Null gehen, wenn n gegen unendlich strebt.

(iii) Ist f schon durch eine Potenzreihe gegeben, so ist die Taylorreihe gleich der Potenzreihe.

Beispiel 5.5.6: Wir berechnen die Taylorreihe von $f : (-1, \infty) \rightarrow \mathbb{R}$ mit $f(x) = \ln(1+x)$. Für diese Funktion gilt

$$f'(x) = \frac{1}{1+x}, \quad f''(x) = \frac{-1}{(1+x)^2}, \quad f'''(x) = \frac{2}{(1+x)^3}, \quad f^{(4)}(x) = \frac{-3 \cdot 2}{(1+x)^4}.$$

Dies führt uns auf die

Vermutung: $f^{(n)}(x) = \frac{(-1)^{n-1}(n-1)!}{(1+x)^n}$ für $n \geq 1$.

Beweis: (mit vollständiger Induktion) Der Induktionsanfang für $n = 1$ findet sich oben. Für den Induktionsschritt gelte also die Behauptung für ein $n \in \mathbb{N}$. Dann haben wir für $n + 1$

$$f^{(n+1)}(x) = (f^{(n)})'(x) = \frac{(-1)^{n-1}(n-1)!(-n)}{(1+x)^{n+1}} = \frac{(-1)^n n!}{(1+x)^{n+1}},$$

was die Behauptung beweist. □

Nach der Definition der Taylorreihe ist diese also für $x_0 = 0$ gegeben durch (vgl. mit Bemerkung 4.6.5 (iii))

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

Nach der Formel von Hadamard ergibt sich der Konvergenzradius dieser Potenzreihe zu 1, wir können also hoffen, dass sie auch die Funktion f in einer Umgebung von $x_0 = 0$ darstellt. Dazu betrachten wir das Verhalten der zugehörigen Lagrange-Restglieder für n gegen unendlich. Diese sind für x aus dem Konvergenzgebiet, also für $x \in (-1, 1)$, gegeben durch

$$\frac{f^{(n+1)}(\xi)}{(n+1)!} x^{n+1} = \frac{(-1)^n n!}{(1+\xi)^{n+1} (n+1)!} x^{n+1} = \frac{(-1)^n x^{n+1}}{(n+1)(1+\xi)^{n+1}}$$

mit einem ξ zwischen x und 0. Also gilt für den Betrag des Restgliedes

$$\left| \frac{f^{(n+1)}(\xi)}{(n+1)!} x^{n+1} \right| = \left| \frac{x}{1+\xi} \right|^{n+1} \frac{1}{n+1}.$$

Im Fall $x > 0$ gilt nun $0 < \xi < x < 1$, also $\left| \frac{x}{1+\xi} \right| = \frac{x}{1+\xi} < \frac{1}{1+0} = 1$. Ähnlich können wir im Fall $x \in [-1/2, 0]$ die Ungleichung $\left| \frac{x}{1+\xi} \right| \leq 1$ zeigen. Dann ist nämlich $-1/2 \leq x < \xi \leq 0$, und mit der umgekehrten Dreiecksungleichung folgt $|1+\xi| = |1 - (-\xi)| \geq 1 - |\xi| \geq 1 - 1/2$. Das liefert

$$\left| \frac{x}{1+\xi} \right| = \frac{|x|}{|1+\xi|} \leq \frac{|x|}{1-|\xi|} \leq \frac{1/2}{1-1/2} = 1.$$

Also ist für alle $x \in [-1/2, 1)$

$$\left| \frac{f^{(n+1)}(\xi)}{(n+1)!} x^{n+1} \right| \leq \frac{1}{n+1}.$$

Die Restglieder gehen also gegen Null für n gegen unendlich, und wir haben

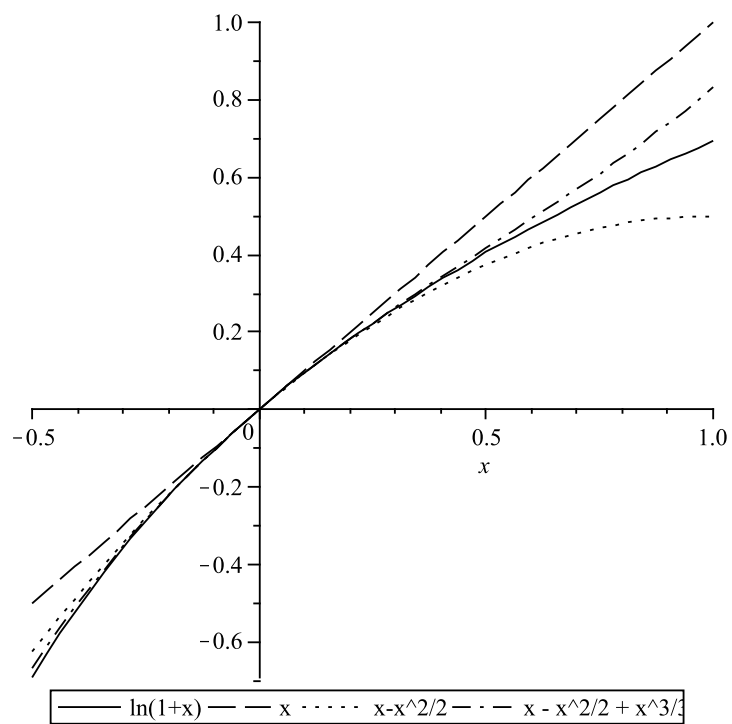
$$f(x) = \ln(1+x) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} x^n \quad \text{für } x \in [-1/2, 1).$$

Diese Gleichheit gilt tatsächlich sogar für alle $x \in (-1, 1)$, aber das bekommt man nicht durch so eine „einfache“ Abschätzung aus dem Lagrange-Restglied gekitzelt.

An der Taylorreihe von $\ln(1+x)$ können wir nun auch problemlos die Taylorpolynome dieser Funktion ablesen. Für $n = 1, 2, 3$ sind das

$$T_1(x) = x, \quad T_2(x) = x - \frac{x^2}{2}, \quad T_3(x) = x - \frac{x^2}{2} + \frac{x^3}{3}.$$

Diese sind in der folgenden Graphik zusammen mit $\ln(1+x)$ dargestellt. Man sieht, dass die Näherung in der Nähe von $x_0 = 0$ immer besser wird.



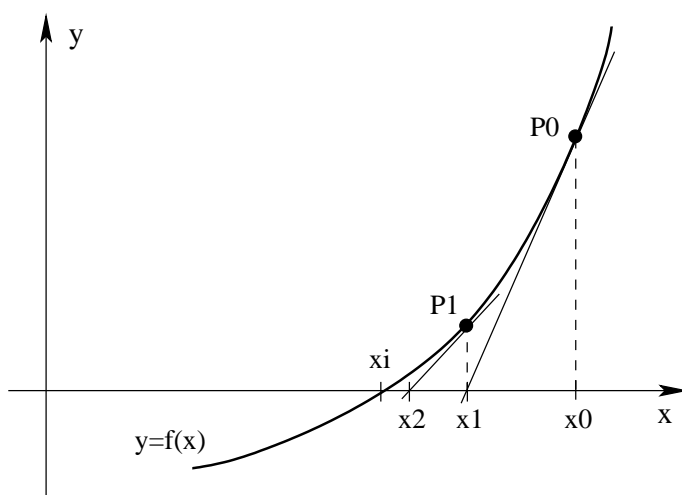
5.6 Das Newtonsche Iterationsverfahren

Die Lösung von Gleichungen der Form $f(x) = 0$ (also die Bestimmung der Nullstellen einer Funktion) ist oft *exakt* (d.h. mittels einer Formel) *nicht möglich*. Das Newton-Verfahren ist ein numerisches Verfahren um solche Gleichungen *näherungsweise* zu lösen.

Geometrische Idee: Ist x_0 eine Stelle „in der Nähe“ der gesuchten Nullstelle ξ der Funktion f , so ersetze man f durch die Tangente an f im Punkt x_0 . Der Schnittpunkt der Tangente mit der x -Achse ist dann eine Näherung für ξ .

Diese Idee wird wiederholt verwendet:

- Tangente im Punkt $P_0 = (x_0, f(x_0))$ liefert Schnittpunkt x_1 ,
- Tangente im Punkt $P_1 = (x_1, f(x_1))$ liefert Schnittpunkt x_2 ,
- Tangente im Punkt $P_2 = (x_2, f(x_2))$ liefert Schnittpunkt x_3 u.s.w.



Wie berechnet sich x_{n+1} aus x_n ? Die Tangentengleichung im Punkt $(x_n, f(x_n))$ lautet

$$y = f(x_n) + f'(x_n)(x - x_n),$$

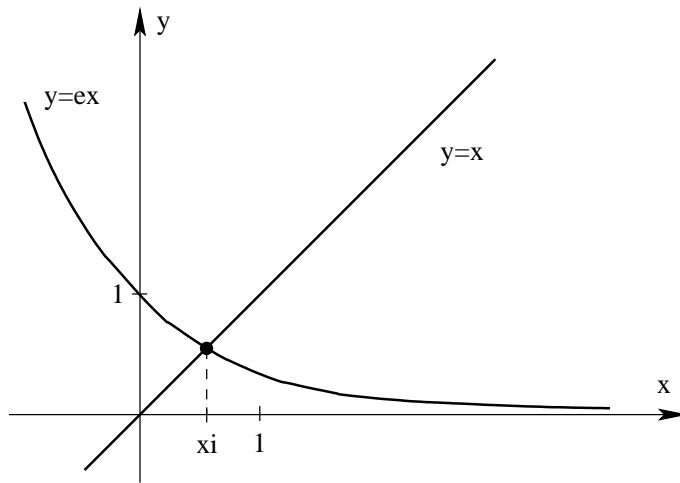
also ist $x = x_n - \frac{f(x_n)}{f'(x_n)}$ der Schnittpunkt mit der x -Achse ($y = 0$). Damit ist

$$\boxed{x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \quad (n = 0, 1, 2, \dots)} \quad (5.1)$$

die Iterationsvorschrift des Newton-Verfahrens.

Beispiel 5.6.1: Löse die Gleichung $x = e^{-x}$.

1. Ist diese Gleichung überhaupt lösbar?



Nach der Skizze sollte eine eindeutige Lösung $x > 0$ existieren!

Mathematisch exakt ergibt sich $x = e^{-x} \Leftrightarrow 0 = f(x) := x - e^{-x}$. Es ist $f(0) = -1 < 0$ und $f(1) = 1 - 1/e > 0$, also hat f einen Vorzeichenwechsel in $[0, 1]$. Außerdem ist $f'(x) = 1 + e^{-x} > 0$ für alle $x \in \mathbb{R}$, also nach Korollar 5.1: f ist streng monoton wachsend, daher hat f nur diese eine Nullstelle (d.h. $f(x) = 0$ ist eindeutig lösbar).

2. Näherungsweise Berechnung der Nullstelle von f . Für $f(x) = x - e^{-x}$ lautet das Iterationsverfahren (5.1):

$$x_{n+1} = x_n - \frac{x_n - e^{-x_n}}{1 + e^{-x_n}} \quad n = 0, 1, 2, \dots$$

Nach (a) ist z.B. $x_0 = 0$ ein sinnvoller Startwert. Ausrechnen der Iterationsvorschrift liefert die Werte:

$$\begin{aligned} x_1 &= 0,5, & x_2 &= 0,5663\dots, & x_3 &= 0,5671421\dots, \\ x_4 &= 0,5671433\dots, & x_5 &= 0,5671433\dots \end{aligned}$$

Der folgende Satz gibt Bedingungen an, die hinreichend sind für die Konvergenz der Folge (x_n) gegen die gesuchte Lösung ξ von $f(x) = 0$.

Satz 5.6.2. Sei f auf $[a, b]$ zweimal stetig differenzierbar. Es gelte $f(a) \cdot f(b) < 0$ (d.h. f hat einen Vorzeichenwechsel in $[a, b]$), $f'(x) \neq 0$ auf $[a, b]$ sowie

$$\left| \frac{f(x)f''(x)}{f'(x)^2} \right| \leq q \quad \text{für alle } x \in [a, b] \text{ mit einem } q < 1.$$

Dann konvergiert die nach der Vorschrift (5.1) gebildete Folge von Näherungswerten bei beliebigem Startwert $x_0 \in [a, b]$ gegen die Lösung ξ der Gleichung $f(x) = 0$.

- Bemerkungen 5.6.3:** (i) Wegen $f'(\xi) \neq 0$ (nach Vorschrift) und $f(\xi) = 0$ sind die Werte von $\frac{f(x)f''(x)}{f'(x)^2}$ in einer hinreichend kleinen Umgebung von ξ beliebig klein. Das Newton-Verfahren konvergiert im Fall $f'(\xi) \neq 0$ also immer, wenn nur der Startwert x_0 ausreichend nahe bei ξ liegt.
- (ii) In der Praxis verzichtet man oft auf die Nachprüfung der Abschätzung im Satz 5.6.2. Wenn Konvergenz eintritt, dann sicher gegen eine Nullstelle.
- (iii) Man kann zeigen, dass das Newton-Verfahren auch im Fall $f'(\xi) = 0$ für genügend nahe Startwerte noch konvergiert - allerdings nicht mehr so schnell.

Fortsetzung des Beispiels 5.6.1: Nach 1.) kann hier $[a, b] = [0, 1]$ gewählt werden. Dann gilt

$$\left| \frac{f(x)f''(x)}{f'(x)^2} \right| = \left| \frac{(x - e^{-x})(-e^{-x})}{(1 + e^{-x})^2} \right| \leq \frac{x + e^{-x}}{(1 + e^{-x})^2} \leq \frac{1 + e^{-x}}{(1 + e^{-x})^2} \leq \frac{1}{1 + 1/e} =: q < 1$$

in $[0, 1]$. Nach Satz 5.6.2 ist die Folge der x_n also konvergent gegen die gesuchte Lösung

$$\xi = 0,56714329\dots \text{ von } x = e^{-x}.$$

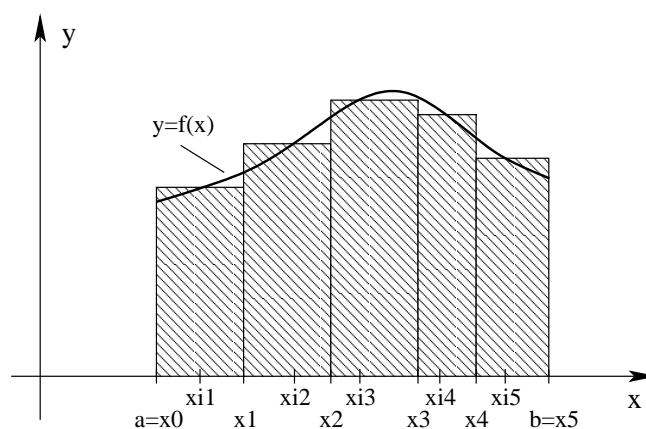
Kapitel 6

Integration

6.1 Das bestimmte Integral

Die Integration ist uns schon aus der Schule als Umkehroperation zur Differentiation geläufig. Historischer Ausgangspunkt der Entwicklung der Integrationstheorie war das Problem der Flächenberechnung.

Grundaufgabe: Berechne die Fläche zwischen einer Kurve $y = f(x)$ und der x -Achse im Bereich $a \leq x \leq b$ (hierbei sei zunächst $f(x) \geq 0$).



Idee: Approximiere die Fläche folgendermaßen durch Rechtecke:

- Zerlege $[a, b]$ in n Teilintervalle: $a = x_0 < x_1 < x_2 \dots < x_n = b$. Dann heißt

$$Z = (x_0, x_1, x_2, \dots, x_n)$$

Zerlegung von $[a, b]$.

- Wähle aus jedem Teilintervall $[x_{i-1}, x_i]$ eine Zwischenstelle $\xi_i \in [x_{i-1}, x_i]$.
- Verwende als Näherung für den Flächeninhalt über der Grundfläche $[x_{i-1}, x_i]$ die Rechteckfläche $f(\xi_i)(x_i - x_{i-1})$.

Dies liefert als Näherung für die Gesamtfläche die *Zerlegungssumme* (Riemannsumme)

$$S(Z, \xi) = \sum_{i=1}^n f(\xi_i)(x_i - x_{i-1}).$$

Man betrachtet nun immer feinere Zerlegungen, um als Grenzwert den exakten Flächeninhalt zu bekommen. Dabei versteht man unter der *Feinheit* $l(Z)$ einer Zerlegung Z die Länge des größten Teilintervalls, d.h. $l(Z) = \max_{i=1, \dots, n} (x_i - x_{i-1})$.

Definition 6.1.1. Sei $f : [a, b] \rightarrow \mathbb{R}$ beschränkt. Dann heißt f über $[a, b]$ *integrierbar*, wenn für jede Folge von Zerlegungen Z_n von $[a, b]$ mit $l(Z_n) \rightarrow 0$ und beliebiger Wahl von Zwischenstellen ξ_i^n , die Folge der Zerlegungssummen $S(Z_n, \xi^n)$ stets konvergent ist und immer denselben Grenzwert besitzt. Dieser Grenzwert heißt das *bestimmte Integral von f über $[a, b]$* , und wird mit

$$\int_a^b f(x) dx$$

bezeichnet.

Der folgende Satz gibt eine wichtige hinreichende Bedingung für die Integrierbarkeit.

Satz 6.1.2. Sei $f : [a, b] \rightarrow \mathbb{R}$ stetig. Dann ist f über $[a, b]$ integrierbar.

Bemerkungen 6.1.3: (i) Wenn f auf $[a, b]$ integrierbar ist, kann man zur Berechnung von $\int_a^b f(x) dx$ möglichst günstige Zerlegungen und Zwischenstellen verwenden.

(ii) Die Bezeichnung der Integrationsvariablen ist willkürlich! Zum Beispiel ist

$$\int_a^b f(x) dx = \int_a^b f(t) dt.$$

Beispiele 6.1.4: (a) $\int_0^b x^2 dx$. Es ist $f(x) = x^2$ stetig, also nach Satz 6.1.2 integrierbar.

Verwende hier *äquidistante* Zerlegungen: $0 = x_0 < x_1 < \dots < x_n = b$ mit $x_i = i \frac{b}{n}$.

Zwischenstellen: $\xi_i = x_i$. Dann erhält man

$$S_n (= S(Z_n, \xi^n)) = \sum_{i=1}^n f(\xi_i)(x_i - x_{i-1}) = \sum_{i=1}^n \left(i \frac{b}{n}\right)^2 \frac{b}{n} = \left(\frac{b}{n}\right)^3 \sum_{i=1}^n i^2.$$

Es ist $\sum_{i=1}^n i^2 = \frac{1}{6} n(n+1)(2n+1)$ (Beweis mit vollständiger Induktion), und damit

$$\lim_{n \rightarrow \infty} S_n = \lim_{n \rightarrow \infty} \frac{1}{6} b^3 \frac{n(n+1)(2n+1)}{n \cdot n \cdot n} = \frac{1}{3} b^3.$$

Fazit: $\int_0^b x^2 dx = \frac{1}{3} b^3.$

(b) $\int_1^b \frac{dx}{x}$ für $b > 1$. Die Funktion $f(x) = \frac{1}{x}$ ist für $x > 0$ stetig, also integrierbar.

Zerlege $[1, b]$ durch $x_i = b^{i/n}$ ($i = 0, \dots, n$) und wähle $\xi_i = x_{i-1}$. Dann erhält man

$$S_n = \sum_{i=1}^n f(\xi_i)(x_i - x_{i-1}) = \sum_{i=1}^n b^{-\frac{i-1}{n}} \left(b^{\frac{i}{n}} - b^{\frac{i-1}{n}} \right) = \sum_{i=1}^n (b^{\frac{1}{n}} - 1),$$

also $S_n = n(b^{\frac{1}{n}} - 1)$. Es gilt

$$\lim_{n \rightarrow \infty} n(b^{\frac{1}{n}} - 1) = \lim_{n \rightarrow \infty} \frac{b^{\frac{1}{n}} - 1}{1/n} = \lim_{x \rightarrow 0+} \frac{b^x - 1}{x} = \lim_{x \rightarrow 0+} \frac{e^{x \ln b} - 1}{x} = \ln b,$$

wobei im letzten Schritt die Regel von de L'Hôpital verwendet wurde.

Fazit: $\int_1^b \frac{dx}{x} = \ln b.$

6.2 Eigenschaften des bestimmten Integrals

6.2.1 Rechenregeln

Satz 6.2.1 (Linearität und Monotonie des Integrals). Seien f und g auf $[a, b]$ integrierbar. Dann gilt

(i) $\int_a^b f(x) dx = \int_a^c f(x) dx + \int_c^b f(x) dx$ für $c \in (a, b)$

(ii) $\int_a^b (\alpha f(x) + \beta g(x)) dx = \alpha \int_a^b f(x) dx + \beta \int_a^b g(x) dx$ für $\alpha, \beta \in \mathbb{R}$

(iii) Aus $f(x) \leq g(x)$ auf $[a, b]$ folgt $\int_a^b f(x) dx \leq \int_a^b g(x) dx$

(iv) $\left| \int_a^b f(x) dx \right| \leq \int_a^b |f(x)| dx$

Bemerkungen 6.2.2: (a) Teil (i) bleibt für beliebiges c richtig, falls die Teilintegrale existieren. Dabei definiert man

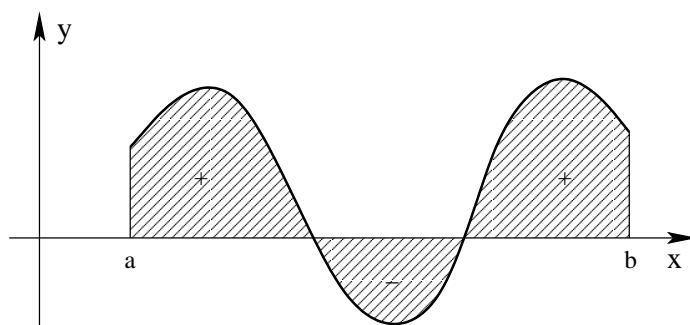
$$\int_a^a f(x) dx := 0 \quad \text{und} \quad \int_a^b f(x) dx := - \int_b^a f(x) dx \quad \text{für } a > b.$$

(b) Kombination von Satz 6.1.2 und Satz 6.2.1 (i) zeigt, dass alle stückweise stetigen Funktionen integrierbar sind.

6.2.2 Flächeninhalt

Ist $f : [a, b] \rightarrow \mathbb{R}$ integrierbar mit $f(x) \geq 0$ auf $[a, b]$ und $a \leq b$, so ist $A := \int_a^b f(x) dx$ der Inhalt der Fläche zwischen dem Graphen von f und der x -Achse. Im Fall $f(x) \leq 0$ auf $[a, b]$ ist der Flächeninhalt durch $A := - \int_a^b f(x) dx$ gegeben. Allgemein gilt

$$\begin{aligned} \int_a^b f(x) dx &= \{\text{Flächeninhalt der Bereiche oberhalb der } x\text{-Achse}\} \\ &\quad - \{\text{Flächeninhalt der Bereiche unterhalb der } x\text{-Achse}\} \end{aligned}$$

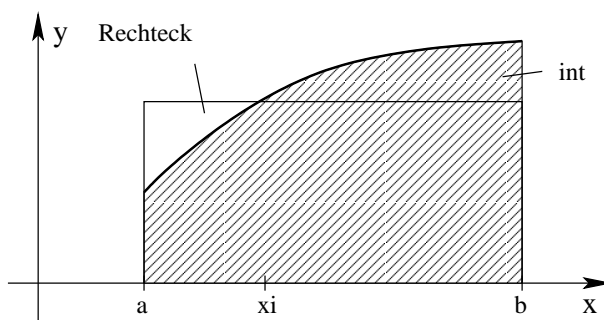


6.2.3 Mittelwertsatz der Integralrechnung

Satz 6.2.3 (Mittelwertsatz der Integralrechnung). Sei $f : [a, b] \rightarrow \mathbb{R}$ stetig. Dann gibt es ein $\xi \in (a, b)$ mit

$$\int_a^b f(x) dx = f(\xi)(b - a).$$

Geometrische Bedeutung: Die Fläche zwischen der x -Achse und der Kurve $y = f(x)$ (im Bereich $a \leq x \leq b$) ist gleich der Rechtecksfläche über $[a, b]$ mit der Höhe $f(\xi)$ für eine Stelle $\xi \in (a, b)$.



6.3 Zusammenhang zwischen Differential- und Integralrechnung

Wie kann man $\int_a^b f(x) dx$ ohne Betrachtung von Zerlegungssummen berechnen?

Satz 6.3.1 (Hauptsatz der Differential- und Integralrechnung).

(i) Sei $f : [a, b] \rightarrow \mathbb{R}$ stetig, $x_0 \in [a, b]$ beliebig und $F(x) := \int_{x_0}^x f(t) dt$ für $x \in [a, b]$.

Dann ist F differenzierbar und es gilt $F'(x) = f(x)$ für $x \in [a, b]$.

(ii) Sei $F : [a, b] \rightarrow \mathbb{R}$ eine stetig differenzierbare Funktion mit $F'(x) = f(x)$ für $x \in [a, b]$. Dann gilt

$$\int_a^b f(x) dx = F(b) - F(a) =: F(x) \Big|_a^b$$

Beweis.

(i) Es ist

$$F(x+h) - F(x) = \int_{x_0}^{x+h} f(t) dt - \int_{x_0}^x f(t) dt = \int_x^{x+h} f(t) dt.$$

Nach Satz 6.2.3 gibt es eine Stelle ξ zwischen x und $x+h$ mit

$$\int_x^{x+h} f(t) dt = f(\xi) \cdot h$$

Dabei hängt ξ von h ab, also $\xi = \xi(h)$, und wegen $x < \xi(h) < x+h$ gilt $\xi(h) \rightarrow x$ für $h \rightarrow 0$. Daher folgt

$$\lim_{h \rightarrow 0} \frac{F(x+h) - F(x)}{h} = \lim_{h \rightarrow 0} f(\xi(h)) = f(x),$$

da f nach Voraussetzung stetig ist. Also existiert $F'(x)$, und es ist $F'(x) = f(x)$.

(ii) Setze $G(x) := \int_a^x f(t) dt$. Dann ist $G(a) = 0$ und $G(b) = \int_a^b f(t) dt$.

Nach Teil (i) dieses Satzes gilt $G'(x) = f(x)$, also folgt

$$(F - G)'(x) = 0 \quad \text{für alle } x \text{ in } [a, b].$$

Nach Korollar 5.3.3 (i) ist $F - G$ konstant, insbesondere $F(a) - G(a) = F(b) - G(b)$ und damit

$$F(b) - F(a) = G(b) - G(a) = \int_a^b f(t) dt. \quad \square$$

Bemerkung 6.3.2. Eine differenzierbare Funktion F mit $F'(x) = f(x)$ auf $[a, b]$ heißt *Stammfunktion von f* . Sind F und G zwei Stammfunktionen von f auf $[a, b]$, so gilt $(F - G)' = 0$, also $G(x) = F(x) + c$ mit einer Konstante $c \in \mathbb{R}$. Man erhält also alle Stammfunktionen zu f durch Addition von Konstanten. Die Menge aller Stammfunktion von f nennt man auch *unbestimmtes Integral von f* und schreibt dafür $\int f(x) dx$. Insbesondere kann man das unbestimmte Integral also nicht wie die Ableitung ohne weiteres wieder als Funktion auffassen. Erst nach der Wahl einer Integrationskonstanten funktioniert diese Identifikation. Um *alle* Stammfunktionen von f anzugeben, wird die (frei wählbare) sogenannte Integrationskonstante mit aufgeführt, welche wir normalerweise mit c bezeichnen.

Liste einiger Stammfunktionen.

$\int x^r dx = \frac{1}{r+1} x^{r+1} + c \quad \text{für } r \neq -1$	
$\int \frac{dx}{x} = \ln x + c$	$\int e^x dx = e^x + c$
$\int \sin x dx = -\cos x + c$	$\int \cos x dx = \sin x + c$
$\int \frac{dx}{1+x^2} = \arctan x + c$	$\int \frac{dx}{\sqrt{1-x^2}} = \arcsin x + c$
$\int \frac{dx}{\cos^2 x} = \tan x + c$	$\int \frac{dx}{\sin^2 x} = -\cot x + c$

6.4 Integrationsmethoden

6.4.1 Partielle Integration (Produktregel)

Unbestimmte Integration der Produktregel $(fg)' = f'g + fg'$ liefert

$$\int f'(x)g(x) dx = f(x)g(x) - \int f(x)g'(x) dx,$$

und damit

$$\int_a^b f'(x)g(x) dx = f(x)g(x)\Big|_a^b - \int_a^b f(x)g'(x) dx.$$

Anwendung zur Berechnung von $\int h(x) dx$: Versuche h als $f'g$ zu schreiben, und zwar so, daß $\int fg' dx$ einfacher zu berechnen ist.

Beispiele 6.4.1: (a) $\int x \sin x dx = ?$ Hier: $g(x) = x, f'(x) = \sin x \Rightarrow f(x)g'(x) = -\cos x$

$$\begin{aligned} \Rightarrow \int x \sin x dx &= -x \cos x - \int 1 \cdot (-\cos x) dx = -x \cos x + \int \cos x dx \\ &= -x \cos x + \sin x + c \end{aligned}$$

(b) $\int \ln x dx = ?$ Hier: $(\ln x)' = \frac{1}{x}$; deshalb $g(x) = \ln x$, also $f'(x) = 1$.

$$\Rightarrow \int \ln x dx = x \ln x - \int x \frac{1}{x} dx = x \ln x - \int dx = x \ln x - x + c.$$

(c) $\int \sin^2 x dx = \int \sin x \sin x dx = -\cos x \sin x + \int \cos^2 x dx \quad \Big| + \int \sin^2 x dx$

$$\Rightarrow 2 \int \sin^2 x dx = -\sin x \cos x + \int dx \Rightarrow \int \sin^2 x dx = \frac{x - \sin x \cos x}{2} + c$$

6.4.2 Substitution (Kettenregel)

Für differenzierbare Funktionen F und g gilt $(F \circ g)'(t) = F'(g(t))g'(t)$ nach der Kettenregel. Also folgt mit $f = F'$:

$$\int f(g(t))g'(t) dt = \int F(g(t))' dt = F(g(t)) + c.$$

In die Stammfunktion F von f ist also $g(t)$ als Argument einzusetzen. Als Abkürzung für diese Ersetzung verwendet man die Schreibweise: $[h(x)]_{x=g(t)} := h(g(t))$.

Dann lautet die Substitutionsregel für unbestimmte Integrale

$$\int f(g(t))g'(t) dt = \left[\int f(x) dx \right]_{x=g(t)},$$

falls f stetig und g stetig differenzierbar ist.

Beispiel 6.4.2: $\int \sin t \cos^2 t \, dt = - \int \cos^2 t (\cos t)' \, dt$; hier: $g(t) = \cos t$, $f(x) = x^2$. Also folgt

$$\int \sin t \cos^2 t \, dt = \left[- \int x^2 \, dx \right]_{x=\cos t} = \left[-\frac{1}{3}x^3 + c \right]_{x=\cos t} = -\frac{1}{3} \cos^3 t + c.$$

Oft wird die Substitutionsregel „von rechts nach links“ angewendet; dazu benötigt man die Umkehrbarkeit von g . Es gilt

$$\boxed{\int f(x) \, dx = \left[\int f(g(t))g'(t) \, dt \right]_{t=g^{-1}(x)},}$$

falls f stetig und g stetig differenzierbar mit $g'(t) \neq 0$ ist.

Beispiel 6.4.3: $\int \frac{e^x}{e^{2x} + 1} \, dx = ?$

Hier: $e^x = t$ substituieren; also $x = \ln t$, d.h. $g(t) = \ln t$, $g'(t) = \frac{1}{t}$.

$$\begin{aligned} \Rightarrow \int \frac{e^x}{e^{2x} + 1} \, dx &= \left[\int \frac{e^{\ln t}}{e^{2 \ln t} + 1} \frac{1}{t} \, dt \right]_{t=e^x} = \left[\int \frac{t}{t^2 + 1} \frac{1}{t} \, dt \right]_{t=e^x} = \left[\int \frac{dt}{t^2 + 1} \right]_{t=e^x} \\ &= \left[\arctan t + c \right]_{t=e^x} = \arctan(e^x) + c. \end{aligned}$$

Beim Anwenden der Substitutionsregel benutzt man oft die Leibnizsche Differentialschreibweise, d.h. $\frac{dx}{dt}$ für $x'(t)$, und rechnet mit $\frac{dx}{dt}$ formal wie mit einem gewöhnlichen Bruch; im Beispiel von oben:

$$e^x = t \quad \Rightarrow \quad \frac{dt}{dx} = e^x, \text{ also } dt = e^x \, dx \quad \Rightarrow \quad \int \frac{e^x}{e^{2x} + 1} \, dx = \int \frac{dt}{t^2 + 1}$$

Dabei wird die Ersetzungsklammer $[\dots]_{t=g^{-1}(x)}$ meist weggelassen. Anschließend muss an die Rücksubstitution gedacht werden!

Beispiel 6.4.4: $\int \sin \sqrt{x} \, dx = ?$

Substitution: $t = \sqrt{x} \quad \Rightarrow \quad x = t^2, \quad dx = 2t \, dt$

$$\begin{aligned} \Rightarrow \int \sin \sqrt{x} \, dx &= \int (\sin t) 2t \, dt = 2 \int t \sin t \, dt \\ &= 2t(-\cos t) - 2 \int (-\cos t) \, dt = -2t \cos t + 2 \sin t + c. \end{aligned}$$

Rücksubstitution:

$$\int \sin \sqrt{x} \, dx = -2\sqrt{x} \cos \sqrt{x} + 2 \sin \sqrt{x} + c.$$

Bemerkung 6.4.5. Das Ergebnis einer unbestimmten Integration kann durch Differenzieren leicht nachgeprüft werden!

Bei Substitution in bestimmten Integralen sind die Grenzen mit zu transformieren:

$$\int_a^b f(g(t))g'(t) dt = \int_\alpha^\beta f(x) dx \quad \text{mit } \alpha = g(a), \beta = g(b),$$

falls f stetig und g stetig differenzierbar ist.

Beispiel 6.4.6: $\int_1^e \frac{dt}{t(1+\ln t)} = ?$

Substitution: $x = \ln t$, $dx = \frac{1}{t} dt$ (also $g(t) = \ln t$, $f(x) = \frac{1}{x+1}$)

$$\Rightarrow \int_1^e \frac{dt}{t(1+\ln t)} = \int_{\ln 1}^{\ln e} \frac{dx}{1+x} = \int_0^1 \frac{dx}{1+x} = \ln|1+x| \Big|_0^1 = \ln 2.$$

6.4.3 Einige Standardsubstitutionen

1. Integrale mit (ganzzahligen) *Potenzen von e^x* :

Substituiere $t = e^x \Rightarrow dt = e^x dx$, $dx = \frac{1}{t} dt$

Beispiele 6.4.7: (a) $\int \frac{1+e^{2x}}{e^x} dx = \int \frac{1+t^2}{t} \cdot \frac{1}{t} dt = \int \left(1 + \frac{1}{t^2}\right) dt = t - \frac{1}{t} + c = e^x - e^{-x} + c.$

(b) $\int \frac{\cosh x}{1+e^x} dx = ?$ Beachte: $\cosh x$, $\sinh x$ sind mittels e^x , e^{-x} definiert!

$$\begin{aligned} \int \frac{\cosh x}{1+e^x} dx &= \frac{1}{2} \int \frac{e^x + e^{-x}}{1+e^x} dx = \frac{1}{2} \int \frac{t + \frac{1}{t}}{1+t} dt = \frac{1}{2} \int \frac{t^2 + 1}{t^2 + t^3} dt \\ &= \frac{1}{2} \int \frac{1}{1+t} dt + \frac{1}{2} \int \frac{dt}{t^2 + t^3} = \frac{1}{2} \ln|1+t| + \frac{1}{2} \int \frac{dt}{t^2 + t^3}. \end{aligned}$$

Unklar ist: wie berechnet man $\int \frac{dt}{t^2 + t^3}$? Dies klären wir später im Abschnitt „Partialbruchzerlegung“.

2. Integrale mit *Potenzen von x und $\sqrt[n]{ax+b}$ ($n \in \mathbb{N}$)*:

Substituiere $t = \sqrt[n]{ax+b} \Rightarrow x = \frac{t^n - b}{a}$, $dx = \frac{n}{a} t^{n-1} dt.$

Beispiele 6.4.8: (a) $\int \frac{x}{\sqrt{x-1}} dx = ?$ Hier: $t = \sqrt{x-1}$, $x = t^2 + 1$, $dx = 2t dt$

$$\begin{aligned} \Rightarrow \int \frac{x}{\sqrt{x-1}} dx &= \int \frac{t^2 + 1}{t} 2t dt = 2 \int (t^2 + 1) dt = \frac{2}{3} t^3 + 2t + c \\ &= \frac{2}{3} (x-1)^{\frac{3}{2}} + 2(x-1)^{\frac{1}{2}} + c. \end{aligned}$$

$$(b) \int \frac{\sqrt[3]{x+1}}{x} dx = ? \text{ Hier: } t = \sqrt[3]{x+1}, x = t^3 - 1, dx = 3t^2 dt$$

$$\Rightarrow \int \frac{\sqrt[3]{x+1}}{x} dx = \int \frac{t}{t^3 - 1} 3t^2 dt = 3 \int \frac{t^3}{t^3 - 1} dt = 3t + 3 \int \frac{dt}{t^3 - 1}.$$

Unklar ist: wie berechnet man $\int \frac{dt}{t^3 - 1}$? Dazu später.

3. Integrale mit *Potenzen von x und $\sqrt{1-x^2}$:*

Substituiere $x = \sin t$. Dann ist formal $dx = \cos t dt$ und $\sqrt{1-x^2} = \cos t$.

Beispiele 6.4.9: (a) $\int x^2 \sqrt{1-x^2} dx = \int \sin^2 t \cos^2 t dt = \frac{1}{4} \int \sin^2 2t dt$

Mit der Substitution $y = 2t$, $dy = 2 dt$ erhält man weiter

$$= \frac{1}{8} \int \sin^2 y dy = \frac{y - \sin y \cos y}{16} + C = \frac{2t - \sin(2t) \cos(2t)}{16} + C.$$

Nach Rücktransformationen folgt daraus

$$\int x^2 \sqrt{1-x^2} dx = \frac{1}{8} \arcsin x - \frac{1}{4} x (1-x^2)^{\frac{3}{2}} + \frac{1}{8} x^3 (1-x^2)^{\frac{1}{2}} + C.$$

(b) $\int x \sqrt{1-x^2} dx$. Hier besser $y = 1-x^2$ substituieren, denn $\frac{dy}{dx} = -2x$ und der Integrand enthält den Faktor x .

$$\Rightarrow \int x \sqrt{1-x^2} dx = -\frac{1}{2} \int \sqrt{y} dy = -\frac{1}{3} y^{\frac{3}{2}} = -\frac{1}{3} (1-x^2)^{\frac{3}{2}}$$

4. Integrale mit *Potenzen von x und $\sqrt{x^2-1}$:*

Substituiere $x = \cosh t \Rightarrow dx = \sinh t dt$, $\sqrt{x^2-1} = \sinh t$.

Beispiel 6.4.10: $\int \sqrt{x^2-1} dx = \int \sinh^2 t dt$. Mit partieller Integration erhält man

$$\int \sqrt{x^2-1} dx = \frac{1}{2} \cosh t \sinh t - \frac{t}{2} = \frac{1}{2} x \sqrt{x^2-1} - \frac{1}{2} \operatorname{arcosh} x$$

5. Integrale mit *Potenzen von x und $\sqrt{x^2+1}$:*

Substituiere $x = \sinh t \Rightarrow dx = \cosh t dt$, $\sqrt{x^2+1} = \cosh t$

Beispiel 6.4.11: $\int \sqrt{1 + \frac{1}{x^2}} dx = \int \frac{\sqrt{x^2+1}}{x} dx$ (für $x > 0$) $= \int \frac{\cosh^2 t}{\sinh t} dt$

$$= \frac{1}{2} \int \frac{e^{2t} + 2 + e^{-2t}}{e^t - e^{-t}} dt = \frac{1}{2} \int \frac{y^2 + 2 + \frac{1}{y^2}}{y - \frac{1}{y}} dy \text{ (nach Substitution } y = e^t)$$

$$= \frac{1}{2} \int \frac{y^4 + 2y^2 + 1}{y^4 - y^2} dy = \frac{1}{2} \int dy + \frac{1}{2} \int \frac{3y^2 + 1}{y^4 - y^2} dy$$

Unklar: wie berechnet man $\int \frac{3y^2 + 1}{y^4 - y^2} dy$?

Diese Beispiele zeigen, dass man oft Stammfunktionen von rationalen Funktionen (also Funktionen der Form $P(x)/Q(x)$ mit Polynomen P, Q) bestimmen muss. Dazu benötigt man die sogenannte Partialbruchzerlegung.

6.4.4 Partialbruchzerlegung

Aufgabe: Berechne $\int \frac{P(x)}{Q(x)} dx$ für gegebene Polynome P und Q .

Grundidee: Schreibe die rationale Funktion $\frac{P(x)}{Q(x)}$ als Summe einfacher Brüche, etwa

$$\frac{1}{x^2 - x} = \frac{1}{x(x-1)} = \frac{1}{x-1} - \frac{1}{x}.$$

Beispiel 6.4.12: $P(x) = x^4 - 8x^2 - 4x + 13$, $Q(x) = x^3 - 2x^2 - 5x + 6$, also

$$\int \frac{x^4 - 8x^2 - 4x + 13}{x^3 - 2x^2 - 5x + 6} dx$$

Es sind folgende Schritte auszuführen:

1. *Polynomdivision* solange, bis Grad des Zählers kleiner als Grad des Nenners ist.

$$(x^4 - 8x^2 - 4x + 13) : (x^3 - 2x^2 - 5x + 6) = x + 2 + \frac{x^2 + 1}{x^3 - 2x^2 - 5x + 6}$$

2. *Zerlege den Nenner.*

$$x^3 - 2x^2 - 5x + 6 = (x-1)(x+2)(x-3)$$

3. *Partialbruchzerlegung.*

- (i) Wenn alle Nullstellen des Nenners reell und verschieden sind:

$$\text{Ansatz: } \frac{x^2 + 1}{(x-3)(x-1)(x+2)} = \frac{A}{x-3} + \frac{B}{x-1} + \frac{C}{x+2}$$

Ansatz auf gemeinsamen Nenner bringen; dies ergibt

$$\begin{aligned} \frac{x^2 + 1}{(x-3)(x-1)(x+2)} &= \frac{A(x-1)(x+2) + B(x-3)(x+2) + C(x-3)(x-1)}{(x-3)(x-1)(x+2)} \\ &= \frac{A(x^2 + x - 2) + B(x^2 - x - 6) + C(x^2 - 4x + 3)}{(x-3)(x-1)(x+2)} \end{aligned}$$

Koeffizientenvergleich, d.h. die Faktoren der gleichen x -Potenzen müssen links und rechts übereinstimmen. Dies liefert das Gleichungssystem

$$\begin{aligned} A + B + C &= 1 \\ A - B - 4C &= 0 \\ -2A - 6B + 3C &= 1 \end{aligned} \quad \Rightarrow \dots \Rightarrow \quad A = 1, \quad B = -\frac{1}{3}, \quad C = \frac{1}{3}.$$

Ergebnis:

$$\begin{aligned} \int \frac{x^2 + 1}{x^3 - 2x^2 - 5x + 6} dx &= \int \frac{dx}{x-3} - \frac{1}{3} \int \frac{dx}{x-1} + \frac{1}{3} \int \frac{dx}{x+2} \\ &= \ln|x-3| - \frac{1}{3} \ln|x-1| + \frac{1}{3} \ln|x+2| + c. \end{aligned}$$

(ii) Wenn der Nenner mehrfache reelle Nullstellen hat:

Beispiel 6.4.13: $P(x) = 1$, $Q(x) = (x - 1)^2(x + 2)$.

$$\text{Ansatz: } \frac{1}{(x - 1)^2(x + 2)} = \frac{A}{x - 1} + \frac{B}{(x - 1)^2} + \frac{C}{x + 2}$$

Weiter wie in (i); dies gibt die Koeffizienten $A = -\frac{1}{9}$, $B = \frac{1}{3}$, $C = \frac{1}{9}$

$$\text{Ergebnis: } \int \frac{dx}{Q(x)} = -\frac{1}{9} \ln|x - 1| - \frac{1}{3} \frac{1}{x - 1} + \frac{1}{9} \ln|x + 2|.$$

(iii) Wenn nicht alle Nullstellen des Nenners reell sind:

Beispiel 6.4.14: $P(x) = 1$, $Q(x) = (x - 1)(x^2 + 1)$.

$$\text{Ansatz: } \frac{1}{(x - 1)(x^2 + 1)} = \frac{A}{x - 1} + \frac{Bx + C}{x^2 + 1}.$$

Weiter wie in (i); dies gibt die Koeffizienten $A = \frac{1}{2}$, $B = C = -\frac{1}{2}$.

$$\text{Ergebnis: } \int \frac{dx}{Q(x)} = \frac{1}{2} \ln|x - 1| - \frac{1}{2} \int \frac{x + 1}{x^2 + 1} dx.$$

Allgemein: *Rechenschema für die Integration rationaler Funktionen.*

Gesucht ist $\int \frac{P(x)}{Q(x)} dx$ für Polynome P und Q .

1. *Schritt:* Ist Grad von $P <$ Grad von Q ? Falls ja: weiter mit Schritt 2. Falls nein: Teile $P(x)$ durch $Q(x)$ mit Rest. Dies liefert

$$\frac{P(x)}{Q(x)} = R(x) + \frac{\tilde{P}(x)}{Q(x)}$$

mit Polynomen R , \tilde{P} , wobei der Grad von \tilde{P} kleiner als der Grad von Q ist. Der Teil $\int R(x) dx$ ist leicht zu berechnen. Wende daher die folgenden Schritte auf $\frac{\tilde{P}}{Q}$ an.

2. *Schritt:* Zerlegung von $Q(x)$ in Faktoren der Form $(x - a)^m$ und $((x - a)^2 + b^2)^n$.

Dazu sind zunächst alle reellen Nullstellen von Q zu bestimmen (schwierig!).

Jede dieser Nullstellen liefert einen Faktor der Form $(x - a)$. Dies ergibt die Zerlegung

$$Q(x) = (x - x_1)^{m_1} (x - x_2)^{m_2} \dots (x - x_k)^{m_k} Q_r(x)$$

mit einem Polynom $Q_r(x)$ ohne reelle Nullstelle. Anschließend ist Q_r in quadratische Faktoren zu zerlegen.

3. *Schritt:* Aufspalten von $\frac{\tilde{P}(x)}{Q(x)}$ in Partialbrüche. Verwende dazu für jeden

Faktor in $Q(x)$	den Ansatz
$(x - a)^m$	$\frac{A_1}{x - a} + \frac{A_2}{(x - a)^2} + \dots + \frac{A_m}{(x - a)^m}$
$((x - a)^2 + b^2)^n$	$\frac{A_1x + B_1}{(x - a)^2 + b^2} + \frac{A_2x + B_2}{((x - a)^2 + b^2)^2} + \dots + \frac{A_nx + B_n}{((x - a)^2 + b^2)^n}$

Berechnung der Konstanten A_i, B_i durch Koeffizientenvergleich oder Einsetzen spezieller Werte für x oder „Grenzwertverfahren“ (siehe unten).

4. *Schritt*: Integration der Partialbrüche. Verwende dazu

$$\int \frac{dx}{x-a} = \ln|x-a| + c$$

$$\int \frac{dx}{(x-a)^m} = -\frac{1}{m-1} \frac{1}{(x-a)^{m-1}} + c \quad (\text{für } m \geq 2)$$

$$\int \frac{Ax+B}{(x-a)^2+b^2} dx = \frac{A}{2} \ln((x-a)^2+b^2) + \frac{Aa+B}{b} \arctan\left(\frac{x-a}{b}\right) + c$$

Für den Nenner $((x-a)^2+b^2)^n$ mit $n \geq 2$ lässt sich durch partielle Integration eine Rekursionsformel herleiten (\rightarrow Übung). Man kann beweisen, dass die Schritte 1–4 stets durchführbar sind. Daher sind alle rationalen Funktionen elementar integrierbar.

Beispiele 6.4.15: (a) Ansatz zur Partialbruchzerlegung von $\frac{P(x)}{Q(x)}$ mit

$$Q(x) = (x-1)(x-2)^3(x^2+1)^2((x-3)^2+9) \quad (\text{Grad von } P \leq 9)$$

$$\frac{A_1}{x-1} + \frac{A_2}{x-2} + \frac{A_3}{(x-2)^2} + \frac{A_4}{(x-2)^3} + \frac{A_5x+A_6}{x^2+1} + \frac{A_7x+A_8}{(x^2+1)^2} + \frac{A_9x+A_{10}}{(x-3)^2+9}$$

(b) Berechne $\int \frac{x^3+5x}{x^4-6x^2+8x+24} dx$

1. *Schritt*: Entfällt.

2. *Schritt*: Ausprobieren von $x = \pm 1, \pm 2, \dots$ liefert $x = -2$ als doppelte Nullstelle des Nennerpolynoms.

$$\Rightarrow Q(x) = (x+2)^2(x^2-4x+6).$$

Weitere reelle Nullstelle? $x^2-4x+6 = (x-2)^2+2$, also keine reellen Nullstellen.

3. *Schritt*: Ansatz

$$\frac{x^3+5x}{(x+2)^2((x-2)^2+2)} = \frac{A}{x+2} + \frac{B}{(x+2)^2} + \frac{Cx+D}{(x-2)^2+2}.$$

Berechnung von B durch Grenzwertverfahren: Multipliziere beide Seiten mit $(x+2)^2$. Auf der rechten Seite lautet der mittlere Summand dann B , und die anderen Summanden enthalten (mindestens einmal) den Faktor $(x+2)$. Lässt man nun x gegen -2 gehen, so bleibt auf der rechten Seite nur B stehen!

Insgesamt erhält man so $\frac{(-2)^3+5(-2)}{(-4)^2+2} = B$, also $B = -1$.

Zur Bestimmung der restlichen Koeffizienten kann man z.B. drei spezielle Werte für x einsetzen. Dies gibt drei Gleichungen für die drei Unbekannten A, C, D .

$$x=0: \quad 0 = \frac{A}{2} + \frac{-1}{4} + \frac{D}{6}$$

$$x=1: \quad \frac{6}{9 \cdot 3} = \frac{A}{3} + \frac{-1}{9} + \frac{C+D}{3}$$

$$x=3: \quad \frac{18}{16 \cdot 2} = \frac{A}{4} + \frac{-1}{16} + \frac{2C+D}{2}$$

$$\begin{aligned}
3A + D &= \frac{3}{2} \\
\Rightarrow A + C + D &= 1 \quad \Rightarrow \quad A = \frac{1}{2}, C = \frac{1}{2}, D = 0 \\
A + 4C + 2D &= \frac{5}{2}
\end{aligned}$$

4. Schritt:

$$\begin{aligned}
\int \frac{P(x)}{Q(x)} dx &= \frac{1}{2} \int \frac{dx}{x+2} - \int \frac{dx}{(x+2)^2} + \frac{1}{2} \int \frac{x}{(x-2)^2+2} dx \\
&= \frac{1}{2} \ln|x+2| + \frac{1}{x+2} + \frac{1}{4} \ln|(x-2)^2+2| + \frac{1}{\sqrt{2}} \arctan\left(\frac{x-2}{\sqrt{2}}\right) + c
\end{aligned}$$

6.5 Uneigentliche Integrale

Unter welchen Voraussetzungen und wie lassen sich bestimmte Integrale über unbeschränkte Integrationsbereiche bzw. mit unbeschränkten Integranden definieren? Etwa

$$\int_0^\infty e^{-x} dx, \quad \int_0^1 \frac{1}{\sqrt{x}} dx.$$

6.5.1 Unbeschränkter Integrationsbereich

Definition 6.5.1. Es sei $f : [a, \infty) \rightarrow \mathbb{R}$ auf jedem Intervall $[a, b]$ (für $b > a$) integrierbar. Falls der Grenzwert

$$\lim_{b \rightarrow \infty} \int_a^b f(x) dx$$

existiert, so definiert man das *uneigentliche Integral von f über $[a, \infty)$* durch

$$\int_a^\infty f(x) dx := \lim_{b \rightarrow \infty} \int_a^b f(x) dx$$

Analog definiert man die uneigentlichen Integrale

$$\int_{-\infty}^b f(x) dx := \lim_{a \rightarrow -\infty} \int_a^b f(x) dx$$

und

$$\int_{-\infty}^\infty f(x) dx := \int_{-\infty}^a f(x) dx + \int_a^\infty f(x) dx \quad (\text{mit beliebigem } a \in \mathbb{R})$$

falls die rechten Seiten existieren.

Bemerkungen 6.5.2: (i) Wenn z.B. $\lim_{b \rightarrow \infty} \int_a^b f(x) dx$ existiert, so sagt man auch $\int_a^\infty f(x) dx$ existiert oder *konvergiert*. Sonst heißt $\int_a^\infty f(x) dx$ *divergent*.

Existiert sogar $\int_a^\infty |f(x)| dx$, so heißt das Integral $\int_a^\infty f(x) dx$ *absolut konvergent*.

(ii) F sei eine Stammfunktion von f . Dann gilt

$$\int_a^\infty f(x) dx = \lim_{b \rightarrow \infty} F(b) - F(a),$$

falls der Grenzwert existiert.

(iii) Damit $\int_{-\infty}^\infty f(x) dx$ existiert, müssen beide Teilintegrale (d.h. die entsprechenden Grenzwerte) unabhängig voneinander existieren.

Beispiele 6.5.3: (a) $\int_0^\infty e^{-x} dx$; es ist $\int_0^b e^{-x} dx = -e^{-x} \Big|_0^b = -e^{-b} + 1$.

Also gilt

$$\lim_{b \rightarrow \infty} \int_0^b e^{-x} dx = \lim_{b \rightarrow \infty} (1 - e^{-b}) = 1, \text{ d.h. } \int_0^\infty e^{-x} dx = 1.$$

(b) $\int_{-\infty}^\infty \frac{dx}{1+x^2}$; berechne zunächst $\int_0^\infty \frac{dx}{1+x^2}$.

$$\int_0^\infty \frac{dx}{1+x^2} = \lim_{b \rightarrow \infty} \int_0^b \frac{dx}{1+x^2} = \lim_{b \rightarrow \infty} \left(\arctan x \Big|_0^b \right) = \lim_{b \rightarrow \infty} \arctan b = \frac{\pi}{2}$$

Man schreibt oft kurz:

$$\int_0^\infty \frac{dx}{1+x^2} = \arctan x \Big|_0^\infty = \frac{\pi}{2}$$

Entsprechend:

$$\int_{-\infty}^0 \frac{dx}{1+x^2} = \arctan x \Big|_{-\infty}^0 = \frac{\pi}{2} \Rightarrow \int_{-\infty}^\infty \frac{dx}{1+x^2} = \frac{\pi}{2} + \frac{\pi}{2} = \pi.$$

(c) $\int_{-\infty}^\infty x dx = ?$

$$\int_0^\infty x dx = \lim_{b \rightarrow \infty} \left(\frac{1}{2} x^2 \Big|_0^b \right) = \lim_{b \rightarrow \infty} \frac{1}{2} b^2 = \infty$$

Fazit: $\int_{-\infty}^\infty x dx$ existiert nicht (oder: ist divergent).

Falsch wäre: $\int_{-\infty}^\infty x dx = \lim_{b \rightarrow \infty} \int_{-b}^b x dx = \lim_{b \rightarrow \infty} \left(\frac{1}{2} x^2 \Big|_{-b}^b \right) = 0$.

Satz 6.5.4 (Majoranten-/Minorantenkriterium). 1. Es gelte $|f(x)| \leq g(x)$ auf $[a, \infty)$ und $\int_a^\infty g(x) dx$ existiere. Dann ist $\int_a^\infty f(x) dx$ (absolut) konvergent und es gilt

$$\left| \int_a^\infty f(x) dx \right| \leq \int_a^\infty |f(x)| dx \leq \int_a^\infty g(x) dx.$$

2. Es gelte $f(x) \geq g(x) \geq 0$ auf $[a, \infty)$ und $\int_a^\infty g(x) dx$ sei divergent. Dann ist auch $\int_a^\infty f(x) dx$ divergent.

Beispiel 6.5.5:

$$\int_1^\infty \frac{\sin x}{x^2} dx \text{ existiert, denn } \left| \frac{\sin x}{x^2} \right| \leq \frac{1}{x^2} \text{ und } \int_1^\infty \frac{1}{x^2} dx = -\frac{1}{x} \Big|_1^\infty = 1.$$

6.5.2 Unbeschränkter Integrand

Definition 6.5.6. Sei f stetig in $[a, b] \setminus \{x_0\}$ (typische Situation: f ist bei x_0 unbeschränkt).

Unter dem uneigentlichen Integral $\int_a^b f(t) dt$ versteht man:

(i) im Fall $x_0 = b$: $\lim_{x \rightarrow b^-} \int_a^x f(t) dt,$

(ii) im Fall $x_0 = a$: $\lim_{x \rightarrow a^+} \int_x^b f(t) dt,$

(iii) im Fall $x_0 \in (a, b)$: $\lim_{x \rightarrow x_0^-} \int_a^x f(t) dt + \lim_{x \rightarrow x_0^+} \int_x^b f(t) dt,$
falls die jeweiligen Grenzwerte existieren.

Beispiel 6.5.7: $\int_0^1 \frac{dt}{\sqrt{t}}$; hier: f ist unbeschränkt bei $t_0 = 0$. Es gilt

$$\lim_{x \rightarrow 0^+} \int_x^1 \frac{dt}{\sqrt{t}} = \lim_{x \rightarrow 0^+} \left(2\sqrt{t} \Big|_x^1 \right) = \lim_{x \rightarrow 0^+} (2 - 2\sqrt{x}) = 2.$$

allgemeiner: $\int_0^1 \frac{dt}{t^\alpha}$ mit $\alpha \in \mathbb{R}$.

$$\alpha = 1: \int_0^1 \frac{dt}{t} = \lim_{x \rightarrow 0^+} \left(\ln |t| \Big|_x^1 \right) = \infty$$

$$\alpha \neq 1: \int_0^1 \frac{dt}{t^\alpha} = \lim_{x \rightarrow 0^+} \left(\frac{1}{1-\alpha} t^{1-\alpha} \Big|_x^1 \right) = \lim_{x \rightarrow 0^+} \frac{1}{1-\alpha} (1 - x^{1-\alpha}) = \begin{cases} \frac{1}{1-\alpha} & \alpha < 1 \\ \infty & \alpha > 1 \end{cases}$$

Fazit: $\boxed{\int_0^1 \frac{dt}{t^\alpha} \text{ existiert} \Leftrightarrow \alpha < 1}.$

Auf diesem Beispiel basiert

Satz 6.5.8. Sei f stetig in $[a, b] \setminus \{x_0\}$. Gilt in einer Umgebung von x_0

(i) $|f(x)| \leq \frac{c}{|x - x_0|^\alpha}$ mit $\alpha < 1$ und $c \in \mathbb{R}$, so existiert $\int_a^b f(x) dx$

(ii) $f(x) \geq \frac{c}{|x - x_0|^\alpha}$ mit $\alpha \geq 1$ und $c \in \mathbb{R}$, so ist $\int_a^b f(x) dx$ divergent.

Der Beweis von Satz 6.5.8 (i) beruht auf dem zu Satz 6.5.4 analogen Majorantenkriterium.

Beispiele 6.5.9: (a) $\int_0^1 \frac{dx}{1-x^2}$; hier ist $x_0 = 1$ und

$$f(x) = \frac{1}{(1-x)(1+x)} \geq \frac{1}{2(x-1)} \Rightarrow \text{Integral ist divergent.}$$

(b) $\int_0^1 \frac{dx}{\sqrt{1-x^2}}$; hier ist $x_0 = 1$ und

$$|f(x)| = \frac{1}{|(1-x)(1+x)|^{\frac{1}{2}}} \leq \frac{1}{|x-1|^{\frac{1}{2}}} \Rightarrow \text{Integral existiert.}$$

6.6 Numerische Integration

Für das Folgende wollen wir voraussetzen, dass das zu berechnende bestimmte Integral $\int_a^b f(x) dx$ existiert. Trotzdem kann es sein, dass es nur näherungsweise numerisch berechnet werden kann. Das trifft z.B. zu, wenn

- f keine elementare Stammfunktion besitzt, z.B. $f(x) = e^{-x^2}$,
- die Bestimmung der Stammfunktion zu kompliziert ist,
- f nur tabellarisch gegeben ist (Messwerte).

In solchen Fällen wird der zu berechnende Integralausdruck angenähert ausgewertet durch numerische Integration (Quadratur):

$$I := \int_a^b f(x) dx \approx \tilde{I} := \sum_{i=1}^n w_i f(x_i)$$

Die Gewichte w_i , die Stützstellen x_i und die Anzahl der Stützstellen und der Funktionsauswertungen n bestimmen Methode und Genauigkeit.

Die "Integration von Tabellendaten" wird hier nicht behandelt. Durch eine Wertetabelle kann eine interpolierende oder approximierende Funktion gelegt werden, die dann exakt integriert werden kann.

6.6.1 Newton-Cotes-Formeln

Dies ist die einfachste Idee: Um das Integral I zu berechnen, wird f durch ein interpolierendes Polynom p ersetzt und dieses exakt integriert. Die zur Interpolation benötigten Funktionswerte werden an $m+1$ äquidistanten Stellen berechnet.

Für $m=1$ und $m=2$ ergeben sich die folgenden Formeln:

$$\text{Trapezregel : } \int_a^b f(x) dx \approx \frac{b-a}{2}(f(a) + f(b)),$$

$$\text{Simpsonregel : } \int_a^b f(x) dx \approx \frac{b-a}{6}(f(a) + 4f(\frac{a+b}{2}) + f(b)).$$

Soll die Genauigkeit erhöht werden, so werden diese einfachen Näherungsformeln mehrfach aneinandergesetzt. Sei zu gegebenem n

$$h = \frac{b-a}{n} \quad \text{und} \quad x_j = a + jh, \quad j = 0, 1, \dots, n.$$

Dann liefert das Aneinanderhängen von n Trapez- bzw. $n/2$ Simpsonregeln (n gerade!) die Näherungsformeln

$$\tilde{I} = T(h) = \frac{h}{2} (f(x_0) + 2f(x_1) + \cdots + 2f(x_{n-1}) + f(x_n)),$$

$$\tilde{I} = S(h) = \frac{h}{3} (f(x_0) + 4f(x_1) + 2f(x_2) + \cdots + 2f(x_{n-2}) + 4f(x_{n-1}) + f(x_n)).$$

Sind Schranken für die 2. bzw. 4. Ableitung der zu integrierenden Funktion bekannt, so lässt sich der Fehler dieser Regeln abschätzen:

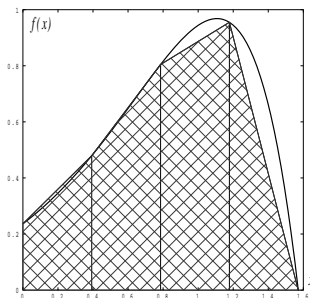
$$|I - T(h)| \leq \frac{|b-a|}{12} h^2 \max_{x \in [a,b]} |f''(x)|, \quad |I - S(h)| \leq \frac{|b-a|}{180} h^4 \max_{x \in [a,b]} |f^{(4)}(x)|.$$

Beispiel 6.6.1:

$$I = \int_0^{\pi/2} \frac{5}{e^\pi - 2} \exp(2x) \cos(x) dx = 1.$$

Die Zeichnung unten zeigt die Trapezfläche, die als Näherung für das Integral bei $n = 4$ entsteht, und den Integranden. Die Ergebnisse für Trapez- und Simpsonregel sind in der folgenden Tabelle festgehalten:

Regel	h	\tilde{I}	Fehler $I - \tilde{I}$	Fehlerabschätzung
Trapez	$\pi/8$	0,926	0,074	0,12
Simpson	$\pi/8$	0,9925	0,0075	0,018



Kapitel 7

Elemente der linearen Algebra

7.1 Der euklidische Raum \mathbb{R}^n

In diesem Kapitel wollen wir uns mit einigen Grundlagen der linearen Algebra auseinandersetzen – uns also mit der *Vektor- und Matrizenrechnung* vertraut machen. Als Motivation sollen uns vor allem zwei wichtige Anwendungen dienen: Zum einen möchten wir lineare Gleichungssysteme¹ systematisch behandeln, anstatt sie nur ad hoc durch Einsetzen oder Umformungen aufzulösen. Zum anderen erlaubt uns die lineare Algebra in effizienter Weise *geometrische* Probleme zu lösen, also Mathematik mit Größen zu betreiben, die eine räumliche Interpretation besitzen.²

Definition 7.1.1. Unter dem *Raum* \mathbb{R}^n versteht man das n -fache kartesische Produkt $\mathbb{R} \times \mathbb{R} \times \dots \times \mathbb{R}$, d.h. die Menge aller n -Tupel (x_1, x_2, \dots, x_n) reeller Zahlen; kurz:

$$\mathbb{R}^n = \{(x_1, x_2, \dots, x_n) : x_k \in \mathbb{R} \text{ für } k = 1, \dots, n\}. \quad (7.1)$$

Die Komponenten x_k heißen *Koordinaten* von x . Zwei Punkte $x, y \in \mathbb{R}^n$ sind gleich, falls $x_k = y_k$ für alle $k = 1, \dots, n$ gilt.

Addition, sowie *Multiplikation mit Skalaren* sind komponentenweise erklärt, d.h.

$$x + y = (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n) \text{ für } x, y \in \mathbb{R}^n$$

$$\lambda x = (\lambda x_1, \lambda x_2, \dots, \lambda x_n) \text{ für } x \in \mathbb{R}^n, \lambda \in \mathbb{R}.$$

Bemerkungen 7.1.2: (i) Jedes Element $x \in \mathbb{R}^n$ definiert einen Punkt des \mathbb{R}^n . Oft nennt man Elemente $x \in \mathbb{R}^n$ auch *Vektoren*, obwohl der genaue Zusammenhang etwas komplizierter ist (siehe unten); zur Unterscheidung werden reelle Zahlen als *Skalare* bezeichnet

(ii) Aus erst später ersichtlichen rechnerischen Gründen werden wir Elemente $x \in \mathbb{R}^n$ oft als Spalten schreiben. Zeilenvektoren sind dann die Transponierten dazu:

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = (x_1, x_2, \dots, x_n)^T$$

¹Diese kommen in *allen* mathematischen Anwendungen vor. Entweder explizit, wie z.B. im vorigen Kapitel bei der Partialbruchzerlegung oder implizit, z.B. bei der rechnergestützten Lösung nichtlinearer Probleme.

²In der Schule wird dies gewöhnlich als *analytische Geometrie* bezeichnet.

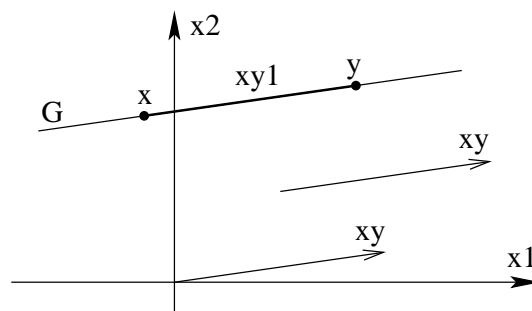
Zusammenhang zum Vektorbegriff.

- (a) *Geraden*. Es seien $x, y \in \mathbb{R}^n$ mit $x \neq y$ gegeben. Unter der durch die Punkte x und y verlaufenden Gerade versteht man die Menge

$$G = \{x + \lambda(y - x) : \lambda \in \mathbb{R}\} \quad (\text{„Zwei-Punkte-Form“}).$$

- (b) *Strecken*. Wird in der Geraden G der Skalar λ auf den Bereich $0 \leq \lambda \leq 1$ eingeschränkt, so erhält man die Strecke \overline{xy} zwischen den Punkten x und y , also

$$\overline{xy} = \{x + \lambda(y - x) : 0 \leq \lambda \leq 1\}.$$



- (c) *Vektoren*. Gibt man der Strecke \overline{xy} eine *Orientierung*, indem man etwa x als Anfangs- und y als Endpunkt festlegt, so erhält man den Vektor \overrightarrow{xy} . Dabei werden alle Vektoren als gleich angesehen, die durch Parallelverschiebung ineinander übergehen. Ein Vektor hat also eine definierte *Richtung* und *Länge*, jedoch *keine feste Position* im Raum.

Jeder Punkt $P = (x_1, x_2, \dots, x_n)^T$ des \mathbb{R}^n lässt sich mit dem Vektor \overrightarrow{OP} identifizieren, der vom Koordinatenursprung $O = (0, 0, \dots, 0)^T$ zum Punkt P führt. Der Vektor \overrightarrow{OP} heißt *Ortsvektor* des Punktes P . Man nennt x_k die Koordinaten des Vektors \overrightarrow{OP} und schreibt

$$\overrightarrow{OP} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = (x_1, x_2, \dots, x_n)^T.$$

Ein beliebiger Vektor \overrightarrow{xy} stimmt mit dem Ortsvektor des Punktes $y - x$ überein, hat also die *Koordinatendarstellung*

$$\overrightarrow{xy} = \begin{pmatrix} y_1 - x_1 \\ y_2 - x_2 \\ \vdots \\ y_n - x_n \end{pmatrix}.$$

Der Vektor $\vec{0} = (0, 0, \dots, 0)^T$ heißt *Nullvektor*.

Um Abstände und Winkel messen zu können, führen wir folgende Begriffe ein.

Definition 7.1.3. Seien $x, y \in \mathbb{R}^n$. Dann heißt

$$\langle x, y \rangle = x_1 y_1 + x_2 y_2 + \dots + x_n y_n$$

das *innere Produkt* (oder *Skalarprodukt*) von x und y und

$$\|x\| = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2}$$

die *Euklidnorm* oder der *Betrag* von x .

Bemerkungen 7.1.4: (i) Andere Schreibweisen für $\langle x, y \rangle$ sind $x \cdot y$ oder gelegentlich (x, y) .

(ii) Die Euklidnorm $\|x\|$ ist die geometrische Länge der Strecke von 0 nach x , d.h. die Länge des Ortsvektors des Punktes x ; entsprechend ist $\|x - y\|$ der Abstand zweier Punkte x und y . Ein Vektor $x \in \mathbb{R}^n$ mit $\|x\| = 1$ heißt *normiert* oder *Einheitsvektor*.

(iii) Die wichtigsten Eigenschaften der Euklidnorm sind

- $\|x\| \geq 0$ und $(\|x\| = 0 \Leftrightarrow x = 0)$
- $\|\lambda x\| = |\lambda| \|x\|$ für $\lambda \in \mathbb{R}, x \in \mathbb{R}^n$
- $\|x + y\| \leq \|x\| + \|y\|$ für $x, y \in \mathbb{R}^n$ (*Dreiecksungleichung*).

(iv) Die wichtigsten Eigenschaften des Skalarproduktes sind

- $\langle x, y \rangle = \langle y, x \rangle$
- $\langle x, x \rangle \geq 0$, und $\langle x, x \rangle = 0$ nur für $x = 0$
- $\langle \alpha x + \beta y, z \rangle = \alpha \langle x, z \rangle + \beta \langle y, z \rangle$,

sowie

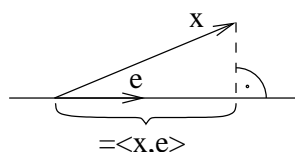
$$\langle x, y \rangle = \|x\| \|y\| \cos \varphi,$$

wobei φ der von x und y eingeschlossene Winkel ist. Insbesondere ist $\langle x, x \rangle = \|x\|^2$. Wegen $|\cos \varphi| \leq 1$ gilt weiter

$$|\langle x, y \rangle| \leq \|x\| \|y\| \quad (\text{Cauchy-Schwarzsche Ungleichung}).$$

Das Skalarprodukt hat folgende anschauliche Bedeutung:

Bildet man das Skalarprodukt zwischen $x \in \mathbb{R}^n$ und einem Einheitsvektor $e \in \mathbb{R}^n$ (also $\|e\| = 1$), so erhält man die Länge der Projektion von x auf die durch e aufgespannte Gerade bis auf das Vorzeichen. Ist der Winkel zwischen x und e größer als $\frac{\pi}{2}$ ($= 90^\circ$), so ist $\langle x, e \rangle < 0$, und $-\langle x, e \rangle$ ist die Länge der Projektion.



(Wichtig etwa bei Zerlegung von Kräften: $\vec{F}_e = \langle \vec{F}, \vec{e} \rangle \vec{e}$, \vec{e} Anteil von \vec{F} in Richtung \vec{e})

Gilt $\langle x, y \rangle = 0$, so heißen x und y zueinander senkrecht (*orthogonal*).

7.2 Lineare Unabhängigkeit, Teilräume, Basis und Dimension

Definition 7.2.1. Vektoren $a_1, a_2, \dots, a_m \in \mathbb{R}^n$ heißen *linear unabhängig*, falls gilt:

Aus dem Ansatz $\sum_{k=1}^m \lambda_k a_k = 0$ (mit $\lambda_k \in \mathbb{R}$) folgt stets $\lambda_1 = \lambda_2 = \dots = \lambda_m = 0$. Sonst heißen sie *linear abhängig*.

Bemerkung 7.2.2. Die Bedingung sagt gerade, dass die Gleichung $\lambda_1 a_1 + \lambda_2 a_2 + \dots + \lambda_m a_m = 0$ nach keinem der a_k aufgelöst werden kann. Also sind Vektoren a_1, \dots, a_m genau dann linear abhängig, wenn mindestens ein a_i als Linearkombination der übrigen a_j geschrieben werden kann.

Beispiele 7.2.3: (a)

$$a_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, a_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, a_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \in \mathbb{R}^3.$$

Anschaulich: a_1, a_2, a_3 linear unabhängig, denn die a_k zeigen in Richtung der Koordinatenachsen. Zeige:

$$\lambda_1 a_1 + \lambda_2 a_2 + \lambda_3 a_3 = 0 \quad \Rightarrow \quad \lambda_1 = \lambda_2 = \lambda_3 = 0.$$

Es gilt:

$$\lambda_1 a_1 + \lambda_2 a_2 + \lambda_3 a_3 = 0 \quad \Rightarrow \quad \begin{pmatrix} \lambda_1 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ \lambda_2 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \lambda_3 \end{pmatrix} = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix} = 0$$

$\Rightarrow \lambda_1 = \lambda_2 = \lambda_3 = 0$, also sind a_1, a_2, a_3 linear unabhängig.

(b)

$$a_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, a_2 = \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix} \in \mathbb{R}^3.$$

Ansatz:

$$\lambda_1 a_1 + \lambda_2 a_2 = 0 \quad \Rightarrow \quad \begin{array}{l} \lambda_1 = 0 \\ -\lambda_1 + 2\lambda_2 = 0 \\ \lambda_2 = 0 \end{array} \quad \Rightarrow \quad \lambda_1 = \lambda_2 = 0.$$

$\Rightarrow a_1, a_2$ sind linear unabhängig.

(c)

$$a_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, a_2 = \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix}, a_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Es ist $a_3 = a_1 + a_2$, also $1 \cdot a_1 + 1 \cdot a_2 + (-1) \cdot a_3 = 0$. Daraus folgt, dass a_1, a_2, a_3 linear abhängig sind. Geometrisch: a_3 liegt in der Ebene, die von a_1, a_2 aufgespannt wird.

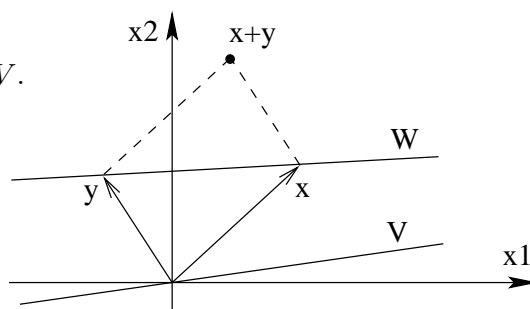
Definition 7.2.4. Eine Teilmenge $V \subseteq \mathbb{R}^n$ heißt *Teilraum* von \mathbb{R}^n (oder *Unterraum*), wenn Addition und Multiplikation mit Skalaren nicht aus V herausführt, d.h.

$$x, y \in V \text{ und } \lambda, \mu \in \mathbb{R} \Rightarrow \lambda x + \mu y \in V$$

Insbesondere ist also $0 \in V$ und $-x \in V$ für jedes $x \in V$.

Beispiele 7.2.5:

- (a) Im \mathbb{R}^2 : Die Gerade V durch den Ursprung ist Teilraum.
Die Gerade W ist *kein* Teilraum.



- (b) $V \subseteq \mathbb{R}^3$ mit $V = \{(x_1, x_2, x_3) : x_1 + x_2 + x_3 = 0\}$. V ist Teilraum des \mathbb{R}^3 , denn:
Sei $x, y, \in V, \lambda, \mu \in \mathbb{R}$ und $z := \lambda x + \mu y$. Dann gilt

$$z_1 + z_2 + z_3 = \lambda(x_1 + x_2 + x_3) + \mu(y_1 + y_2 + y_3) = 0,$$

also $z \in V$.

- (c) $V = \{0\}$ und $V = \mathbb{R}^n$ sind Teilräume des \mathbb{R}^n .

Basis und Dimension.

Jedes Element $x = (x_1, x_2) \in \mathbb{R}^2$ lässt sich als Linearkombination der Vektoren $e_1 = (1, 0)$ und $e_2 = (0, 1)$ darstellen:

$$x = (x_1, 0) + (0, x_2) = x_1 e_1 + x_2 e_2.$$

Ebenso durch die Vektoren $e'_1 = (1, 1)$ und $e'_2 = (1, -1)$:

$$x = \frac{x_1 + x_2}{2} e'_1 + \frac{x_1 - x_2}{2} e'_2.$$

Sowohl $\{e_1, e_2\}$ also auch $\{e'_1, e'_2\}$ bilden deshalb eine sogenannte Basis des \mathbb{R}^2 .

Definition 7.2.6. Sei V ein Teilraum des \mathbb{R}^n . Eine Menge $\{e_1, \dots, e_m\} \subseteq V$ heißt *Basis* von V , wenn

- (i) e_1, \dots, e_m linear unabhängig sind, und
(ii) jedes $v \in V$ Linearkombination der e_1, \dots, e_m ist.

Beispiele 7.2.7: (a) $\{(1, 0), (0, 1)\}$ ist Basis des \mathbb{R}^2 .

Allgemein ist $\{e_1, e_2, \dots, e_n\}$ mit $e_k = (0, \dots, 0, 1, 0, \dots, 0)$ die Standardbasis des \mathbb{R}^n .

zu (i):

$$\lambda_1 e_1 + \dots + \lambda_n e_n = 0 \Rightarrow (\lambda_1, \lambda_2, \dots, \lambda_n) = 0 \Rightarrow \lambda_1 = \lambda_2 = \dots = \lambda_n = 0.$$

zu (ii):

$$x = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n \Rightarrow x = x_1 e_1 + \dots + x_n e_n$$

- (b) $\{(1, 1)\}$ ist keine Basis des \mathbb{R}^2 , denn (ii) verletzt.

(c) $\{(1, 1), (1, -1), (1, 2)\}$ keine Basis des \mathbb{R}^2 , denn (i) verletzt:

$$(1, 2) = \frac{3}{2}(1, 1) - \frac{1}{2}(1, -1).$$

Die folgende Definition wirkt zunächst einigermaßen willkürlich, erweist sich jedoch als nützlich.

Definition 7.2.8. Der Teilraum $\{0\}$ von \mathbb{R}^n hat die Basis \emptyset .

Satz 7.2.9. Sei V Teilraum von \mathbb{R}^n .

- (i) $\{e_1, \dots, e_m\} \subseteq V$ ist genau dann eine Basis von V , wenn jedes $v \in V$ eine eindeutige Darstellung als Linearkombination der e_1, \dots, e_m hat.
- (ii) Jede Basis von V hat die gleiche Anzahl von Elementen.

Aufgrund der letzten Bemerkung ist folgende Definition sinnvoll:

Definition 7.2.10. Sei V ein Teilraum des \mathbb{R}^n . Die Anzahl der Elemente einer Basis von V bezeichnet man als *Dimension* von V . Schreibweise: $\dim V$.

Beispiele 7.2.11: (a) $V = \{x \in \mathbb{R}^3 : x_1 + x_2 + x_3 = 0\}$. Ermittle Basis von V :

Es ist

$$x \in V \Leftrightarrow x = \begin{pmatrix} x_1 \\ x_2 \\ -x_1 - x_2 \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}.$$

Also ist jedes $x \in V$ Linearkombination von $e_1 = (1, 0, -1)$ und $e_2 = (0, 1, -1)$.

Da e_1, e_2 linear unabhängig sind, ist $\{e_1, e_2\}$ eine Basis von V . Also gilt $\dim V = 2$.

(b) $V = \mathbb{R}^n$ hat $\dim V = n$, denn die Standardbasis des \mathbb{R}^n hat n Elemente.

7.3 Lineare Abbildungen und Matrizen

Definition 7.3.1. Eine Funktion $\Phi : \mathbb{R}^n \rightarrow \mathbb{R}^m$ heißt *lineare Abbildung*, falls gilt:

$$\Phi(\lambda x + \mu y) = \lambda \Phi(x) + \mu \Phi(y) \quad \text{für alle } x, y \in \mathbb{R}^n, \lambda, \mu \in \mathbb{R}.$$

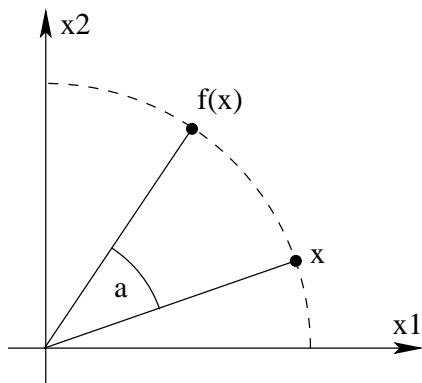
Bemerkung 7.3.2. $\Phi(x)$ ist ein Vektor im \mathbb{R}^m , also ist $\Phi(x) = (\Phi_1(x), \dots, \Phi_m(x))$ mit Komponenten $\Phi_i(x)$. Die Abbildung Φ ist genau dann linear, wenn alle $\Phi_i : \mathbb{R}^n \rightarrow \mathbb{R}$ lineare Abbildungen sind.

Beispiele 7.3.3: (a) „ k -te Koordinate“. $\Phi : \mathbb{R}^n \rightarrow \mathbb{R}$ mit $\Phi(x) = x_k$.

(b) „Projektion auf x_k -Achse“. $\Phi : \mathbb{R}^n \rightarrow \mathbb{R}^n$ mit $\Phi(x) = x_k e_k$.

(c) „Spiegelung an der Winkelhalbierenden“. $\Phi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ mit $\Phi(x_1, x_2) = (x_2, x_1)$.

(d) „Drehung um Winkel α “. $\Phi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ mit $\Phi(x_1, x_2) = \begin{pmatrix} x_1 \cos \alpha - x_2 \sin \alpha \\ x_1 \sin \alpha + x_2 \cos \alpha \end{pmatrix}$



$$(e) \Phi : \mathbb{R}^3 \rightarrow \mathbb{R}^2 \text{ mit } \Phi(x) = \begin{pmatrix} x_1 + 2x_2 - x_3 \\ x_2 + x_3 \end{pmatrix}.$$

Jede lineare Abbildung $\Phi : \mathbb{R}^n \rightarrow \mathbb{R}^m$ ist durch Angabe von $m \cdot n$ Koeffizienten eindeutig bestimmt. Um dies einzusehen, sei $\{e_1, \dots, e_n\}$ die Standardbasis des \mathbb{R}^n .

Dann ist $\Phi(x) = (\Phi_1(x), \dots, \Phi_m(x))$ und

$$\Phi_i(x) = \Phi_i\left(\sum_{k=1}^n x_k e_k\right) = \sum_{k=1}^n x_k \Phi_i(e_k) = \sum_{k=1}^n \alpha_{ik} x_k, \quad i = 1, \dots, m,$$

mit $\alpha_{ik} := \Phi_i(e_k)$ für $i = 1, \dots, m$ und $k = 1, \dots, n$.

Definition 7.3.4. Ein Schema von Elementen $\alpha_{ik} \in \mathbb{R}$ der Form

$$\begin{pmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} & \cdots & \alpha_{2n} \\ \vdots & \vdots & \vdots & & \vdots \\ \alpha_{m1} & \alpha_{m2} & \alpha_{m3} & \cdots & \alpha_{mn} \end{pmatrix}$$

heißt $m \times n$ -Matrix. Die Menge aller $m \times n$ -Matrizen (mit Elementen aus \mathbb{R}) wird mit $M_{m,n}(\mathbb{R})$ oder kurz $\mathbb{R}^{m \times n}$ bezeichnet.

Schreibweise: $A = (\alpha_{ik}) \in M_{m,n}(\mathbb{R})$ oder $A \in M_{m,n}(\mathbb{R})$ mit $A = (\alpha_{ik})$. Für $m = n$ heißt A quadratische Matrix. $(\alpha_{i1}, \dots, \alpha_{in})$ heißt i -te Zeile von A , $\begin{pmatrix} \alpha_{1j} \\ \vdots \\ \alpha_{mj} \end{pmatrix}$ heißt j -te Spalte von A .

Beispiel 7.3.5:

$$A = \begin{pmatrix} 1 & 0 & 2 \\ -1 & 3 & 7 \end{pmatrix} \in M_{2,3}(\mathbb{R}), \quad B = \begin{pmatrix} 4 \\ 1 \\ -2 \end{pmatrix} \in M_{3,1}(\mathbb{R}),$$

$$C = (0 \quad 1 \quad -3) \in M_{1,3}(\mathbb{R}), \quad D = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \in M_{2,2}(\mathbb{R}).$$

Bemerkungen 7.3.6: (i) Zu jeder Matrix $A = (\alpha_{ik}) \in M_{m,n}(\mathbb{R})$ gehört eine eindeutige lineare Abbildung $\Phi_A : \mathbb{R}^n \rightarrow \mathbb{R}^m$, definiert durch

$$(\Phi_A(x))_i := \sum_{k=1}^n \alpha_{ik} x_k, \quad i = 1, \dots, m.$$

Man nennt A die Abbildungsmatrix der linearen Abbildung Φ_A bezüglich der Standardbasis.

- (ii) Jeder Vektor $x \in \mathbb{R}^n$ kann als Matrix mit einer Zeile bzw. einer Spalte aufgefasst werden. Im Zusammenhang mit Matrizen ist es sinnvoll, $x \in \mathbb{R}^n$ als $n \times 1$ -Matrix ($\hat{=}$ Spaltenvektor) anzusehen.
- (iii) Entsprechend der Definition 7.3.4 erklärt man die Menge $M_{m,n}(\mathbb{C})$ der $m \times n$ -Matrizen A mit komplexen Einträgen $\alpha_{ik} \in \mathbb{C}$. Ist die zugrunde liegende Zahlenmenge (\mathbb{R} oder \mathbb{C}) nicht relevant, so schreiben wir kurz $M_{m,n}$.

Definition 7.3.7 (Addition und Multiplikation mit Skalaren). Seien $A, B \in M_{m,n}(\mathbb{R})$ mit $A = (\alpha_{ij}), B = (\beta_{ij})$ und $\lambda \in \mathbb{R}$ (oder \mathbb{C}). Wir definieren

$$A + B := (\alpha_{ij} + \beta_{ij}) = \begin{pmatrix} \alpha_{11} + \beta_{11} & \cdots & \alpha_{1n} + \beta_{1n} \\ \vdots & & \vdots \\ \alpha_{m1} + \beta_{m1} & \cdots & \alpha_{mn} + \beta_{mn} \end{pmatrix}$$

$$\lambda A := (\lambda \alpha_{ij}) = \begin{pmatrix} \lambda \alpha_{11} & \cdots & \lambda \alpha_{1n} \\ \vdots & & \vdots \\ \lambda \alpha_{m1} & \cdots & \lambda \alpha_{mn} \end{pmatrix}$$

Matrizenprodukt.

Seien $\Phi_B : \mathbb{R}^n \rightarrow \mathbb{R}^m$ und $\Phi_A : \mathbb{R}^m \rightarrow \mathbb{R}^p$ lineare Abbildungen mit zugehörigen Matrizen $A = (\alpha_{ij})$ und $B = (\beta_{ik})$. Die Nacheinanderausführung $\Phi_A \circ \Phi_B : \mathbb{R}^n \rightarrow \mathbb{R}^p$ ist dann wieder eine lineare Abbildung. Frage: Welche Matrix gehört zu dieser durch Verknüpfung entstandenen Abbildung?

Es ist

$$(\Phi_A(y))_i = \sum_{j=1}^m \alpha_{ij} y_j \quad \text{und} \quad (\Phi_B(x))_j = \sum_{k=1}^n \beta_{jk} x_k.$$

Daraus folgt

$$\begin{aligned} (\Phi_A(\Phi_B(x)))_i &= \sum_{j=1}^m \alpha_{ij} (\Phi_B(x))_j = \sum_{j=1}^m \alpha_{ij} \sum_{k=1}^n \beta_{jk} x_k \\ &= \sum_{j=1}^m \sum_{k=1}^n \alpha_{ij} \beta_{jk} x_k = \sum_{k=1}^n \left(\sum_{j=1}^m \alpha_{ij} \beta_{jk} \right) x_k. \end{aligned}$$

Also gilt

$$(\Phi_A(\Phi_B(x)))_i = \sum_{k=1}^n \gamma_{ik} x_k \quad \text{mit} \quad \gamma_{ik} := \sum_{j=1}^m \alpha_{ij} \beta_{jk},$$

d.h. es ist $\Phi_A \circ \Phi_B = \Phi_C$ für die Matrix $C = (\gamma_{ik})$. Das legt folgende Definition nahe:

Definition 7.3.8 (Matrizenprodukt). Sei $A = (\alpha_{ij})$ eine $m \times n$ -Matrix und $B = (\beta_{ij})$ eine $n \times p$ -Matrix. Wir definieren das *Produkt* AB als die $m \times p$ -Matrix

$$AB := (\gamma_{ik}) \quad \text{mit} \quad \gamma_{ik} = \sum_{j=1}^n \alpha_{ij} \beta_{jk}, \quad i = 1, \dots, m, \quad k = 1, \dots, p.$$

Ausgeschrieben:

$$AB = \begin{pmatrix} \sum_{j=1}^n \alpha_{1j}\beta_{j1} & \cdots & \sum_{j=1}^n \alpha_{1j}\beta_{jp} \\ \vdots & & \vdots \\ \sum_{j=1}^n \alpha_{mj}\beta_{j1} & \cdots & \sum_{j=1}^n \alpha_{mj}\beta_{jp} \end{pmatrix} = \begin{pmatrix} \gamma_{11} & \cdots & \gamma_{1p} \\ \vdots & & \vdots \\ \gamma_{m1} & \cdots & \gamma_{mp} \end{pmatrix}$$

Bemerkungen 7.3.9: (i) AB ist *nur dann* definiert, wenn die Anzahl der Spalten von A mit der Anzahl der Zeilen von B übereinstimmt!

(ii) Die Berechnung von AB merkt man sich als Schema „Zeile \times Spalte“:

$$\begin{array}{ccc}
 & & \begin{array}{c} k\text{-te Spalte} \\ \text{von } B \text{ (} m \times n \text{)} \\ \downarrow \\ \beta_{1k} \\ \beta_{2k} \\ \vdots \\ \beta_{mk} \end{array} \\
 & \begin{array}{c} \boxed{m \times m} \end{array} & \begin{pmatrix} \\ \\ \\ \end{pmatrix} \\
 & & \downarrow \\
 \begin{array}{c} i\text{-te Zeile} \\ \text{von } A \rightarrow \\ (p \times m) \end{array} & \begin{pmatrix} \alpha_{i1} & \alpha_{i2} & \cdots & \alpha_{im} \end{pmatrix} & \begin{pmatrix} \\ \\ \gamma_{ik} \\ \end{pmatrix} & \begin{array}{c} \leftarrow i\text{-te Zeile} \\ \text{von } AB \\ (p \times n) \end{array} \\
 & & \uparrow \\
 & & \begin{array}{c} k\text{-te Spalte} \\ \text{von } AB \end{array} \\
 & & \gamma_{ik} = \alpha_{i1}\beta_{1k} + \alpha_{i2}\beta_{2k} + \dots + \alpha_{im}\beta_{mk}.
 \end{array}$$

Beispiele 7.3.10: (a)

$$A = \begin{pmatrix} 1 & 2 & 3 \\ -2 & 0 & 4 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & -3 \\ 2 & 1 \\ -1 & 4 \end{pmatrix} \Rightarrow AB = \begin{pmatrix} 1 & 11 \\ -4 & 22 \end{pmatrix}$$

(b) A, B wie in (a)

$$BA = \begin{pmatrix} 0 & -3 \\ 2 & 1 \\ -1 & 4 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ -2 & 0 & 4 \end{pmatrix} = \begin{pmatrix} 6 & 0 & -12 \\ 0 & 4 & 10 \\ -9 & -2 & 13 \end{pmatrix}$$

(c) A wie in (a)

$$B = \begin{pmatrix} 1 & -1 \\ 2 & 0 \end{pmatrix} \Rightarrow AB \text{ ist nicht definiert.}$$

(d)

$$C = \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix}, \quad D = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \Rightarrow CD = \begin{pmatrix} 2 & -1 \\ 0 & 1 \end{pmatrix}, \quad DC = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$$

Beachte: auch für quadratische Matrizen gilt i.a. $CD \neq DC$.

Das Matrizenprodukt ist *nicht* kommutativ.

Rechenregeln

- (i) $A(BD) = (AB)D$ (assoziativ)
- (ii) $A(B+C) = AB+AC$ (distributiv von links)
- (iii) $(B+C)D = BD+CD$ (distributiv von rechts)
- (iv) $\lambda(AB) = (\lambda A)B = A(\lambda B)$

Dabei sei $A \in M_{n,m}$, $B, C \in M_{m,l}$ und $D \in M_{l,k}$, damit alle Produkte und Summen definiert sind.

Produkt „Matrix mal Vektor“: Für $A = (\alpha_{ij}) \in M_{m,n}$ und $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$ (oder \mathbb{C}^n) wird das Produkt Ax so definiert, als wenn x eine $n \times 1$ -Matrix ist. D.h.

$$(Ax)_i = \sum_{j=1}^n \alpha_{ij}x_j$$

Damit ist die Abbildung $x \rightarrow Ax$ gleich der durch A gegebenen linearen Abbildung Φ_A , d.h. $\Phi_A(x) = Ax$.

Ein „neutrales Element“ bzgl. der Matrizenmultiplikation ist die quadratische *Einheitsmatrix* $I_n \in M_{n,n}$, definiert durch

$$I_n = \begin{pmatrix} 1 & & 0 \\ & 1 & \\ & & \ddots \\ 0 & & & 1 \end{pmatrix} \quad (\text{außerhalb der Diagonalen stehen Nullen}).$$

Genauer: $I_n = (\delta_{ij})$ mit dem sogenannten Kroneckersymbol $\delta_{ij} = \begin{cases} 1 & \text{für } i = j \\ 0 & \text{für } i \neq j \end{cases}$

Durch Ausmultiplizieren sieht man

$$I_n A = A \quad \text{für alle } A \in M_{n,m}$$

$$B I_n = B \quad \text{für alle } B \in M_{k,n}.$$

Wir nennen eine $n \times n$ -Matrix A *invertierbar* (oder *regulär*), wenn es eine Matrix $B \in M_{n,n}$ gibt mit

$$AB = BA = I_n.$$

In diesem Fall heißt B *invers* zu A , und man schreibt A^{-1} statt B .

Bemerkung 7.3.11. Um die Inverse von $A \in M_{n,n}$ zu bestimmen, reicht es aus, eine Matrix $B \in M_{n,n}$ mit $AB = I_n$ zu finden. Dann folgt automatisch die Gleichung $BA = I_n$.

Beispiele 7.3.12: (a) In Beispiel 7.3.3 d) haben wir die lineare Abbildung „Drehung um den Winkel α “ in \mathbb{R}^2 betrachtet. Die zugehörige Abbildungsmatrix war

$$A_\alpha = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix}.$$

Anschaulich sollte A_α invertierbar sein, wobei die Inverse die Matrix $A_{-\alpha}$ ist, die zur Drehung um $-\alpha$ gehört. Das ist tatsächlich so, denn es gilt:

$$\begin{aligned} A_\alpha A_{-\alpha} &= \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix} \begin{pmatrix} \cos(-\alpha) & -\sin(-\alpha) \\ \sin(-\alpha) & \cos(-\alpha) \end{pmatrix} \\ &= \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix} \begin{pmatrix} \cos(\alpha) & \sin(\alpha) \\ -\sin(\alpha) & \cos(\alpha) \end{pmatrix} \\ &= \begin{pmatrix} \cos^2(\alpha) + \sin^2(\alpha) & 0 \\ 0 & \cos^2(\alpha) + \sin^2(\alpha) \end{pmatrix} = I_2. \end{aligned}$$

(b) Wir betrachten

$$A = \begin{pmatrix} 3 & -2 \\ -1 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix}.$$

Nachrechnen ergibt

$$AB = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I_2, \text{ also ist } A^{-1} = B.$$

(c) $A = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$ ist *nicht* invertierbar. Annahme: es gibt eine 2×2 -Matrix $B = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ mit $AB = I_2$.

Dann folgt

$$\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \Rightarrow \begin{pmatrix} a+c & b+d \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

und damit den Widerspruch $0 = 1$.

Satz 7.3.13. Sind $A, B \in M_{n,n}$ invertierbare Matrizen, so ist auch AB invertierbar und es gilt

$$\boxed{(AB)^{-1} = B^{-1}A^{-1}.}$$

Man beachte im obigen Satz die Umkehrung in der Reihenfolge der beiden Matrizen!

Bemerkung 7.3.14. Ist A eine invertierbare Matrix, und ist A^{-1} bekannt, so ist das lineare Gleichungssystem $Ax = b$ sofort lösbar:

$$\underline{Ax = b} \Leftrightarrow (A^{-1}A)x = A^{-1}b \Leftrightarrow I_n x = A^{-1}b \Leftrightarrow \underline{x = A^{-1}b}.$$

Beispiel 7.3.15:

$$\begin{aligned} 3x_1 - 2x_2 &= 4 \\ -x_1 + x_2 &= -2 \end{aligned} \Leftrightarrow Ax = \begin{pmatrix} 4 \\ -2 \end{pmatrix} \text{ mit } A \text{ aus Beispiel 7.3.12 b) von oben.}$$

Also ist

$$x = \begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} 4 \\ -2 \end{pmatrix} = \begin{pmatrix} 0 \\ -2 \end{pmatrix} \text{ Lösung.}$$

Auf die Berechnung von A^{-1} gehen wir später ein. Sie wird in der Praxis nur selten durchgeführt, da sie sehr rechenintensiv ist (s.u.).

Definition 7.3.16. Sei $A = (\alpha_{ij}) \in M_{m,n}$. Die zu A *transponierte Matrix* entsteht aus A durch Vertauschen von Zeilen mit Spalten und wird mit A^\top bezeichnet. A^\top ist eine $n \times m$ -Matrix.

$$A = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \vdots & \vdots & & \vdots \\ \alpha_{m1} & \alpha_{m2} & \cdots & \alpha_{mn} \end{pmatrix} \Rightarrow A^\top = \begin{pmatrix} \alpha_{11} & \alpha_{21} & \cdots & \alpha_{m1} \\ \alpha_{12} & \alpha_{22} & \cdots & \alpha_{m2} \\ \vdots & \vdots & & \vdots \\ \alpha_{1n} & \alpha_{2n} & \cdots & \alpha_{mn} \end{pmatrix}$$

Beispiel 7.3.17:

$$A = \begin{pmatrix} 1 & 2 & 0 \\ -1 & 0 & 3 \end{pmatrix}, \quad A^\top = \begin{pmatrix} 1 & -1 \\ 2 & 0 \\ 0 & 3 \end{pmatrix}.$$

Rechenregeln.

- (i) $(AB)^\top = B^\top A^\top$ für $A \in M_{n,m}$ und $B \in M_{m,l}$
- (ii) $(A^\top)^{-1} = (A^{-1})^\top =: A^{-\top}$ für invertierbare $A \in M_{n,n}$.

Definition 7.3.18. Sei $A \in M_{m,n}(\mathbb{R})$. Dann heißt

$$\begin{aligned} \{x \in \mathbb{R}^n : Ax = 0\} & \text{ der Kern von } A, \quad \text{kurz } \ker A. \\ \{Ax \in \mathbb{R}^m : x \in \mathbb{R}^n\} & \text{ das Bild von } A, \quad \text{kurz } \operatorname{Im} A. \end{aligned}$$

Satz 7.3.19. Sei $A \in M_{m,n}(\mathbb{R})$. Dann ist $\ker A$ ein Teilraum von \mathbb{R}^n und $\operatorname{Im} A$ ein Teilraum von \mathbb{R}^m .

Beweis: Sind $x, y \in \ker A$ und $\lambda, \mu \in \mathbb{R}$, so gilt wegen $Ax = Ay = 0$

$$A(\lambda x + \mu y) = \lambda Ax + \mu Ay = 0.$$

Also ist $\lambda x + \mu y$ ebenfalls in $\ker A$ und dieses damit ein Teilraum von \mathbb{R}^n .

Für $x, y \in \operatorname{Im} A$ gibt es nach Definition des Bildes $u, v \in \mathbb{R}^n$ mit $Au = x$ und $Av = y$. Damit gilt für $\lambda, \mu \in \mathbb{R}$ wieder wegen $A(\lambda u + \mu v) = \lambda Au + \mu Av = \lambda x + \mu y$, dass auch der Vektor $\lambda x + \mu y$ im Bild von A liegt, dieses also auch ein Teilraum ist. \square

Definition 7.3.20. Man bezeichnet die Dimension von $\ker A$ als *Defekt* von A und die Dimension von $\operatorname{Im} A$ als *Rang von A* , kurz $\operatorname{rg} A$.

Insbesondere der Rang von A ist wichtig zur Lösung von linearen Gleichungssystemen. Ein erster Schritt zur Berechnung von $\operatorname{rg} A$ liefert folgender Satz.

Satz 7.3.21. Sei $A \in M_{m,n}$. Dann gilt

$$\begin{aligned} \operatorname{rg} A &= \text{maximale Anzahl linear unabhängiger Spalten von } A \\ &= \text{maximale Anzahl linear unabhängiger Zeilen von } A. \end{aligned}$$

Insbesondere gilt also $\operatorname{rg} A = \operatorname{rg}(A^\top)$.

Beispiel 7.3.22:

$$A = \begin{pmatrix} 1 & 0 & 3 & 4 \\ 0 & 1 & 5 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Die ersten zwei Zeilen sind linear unabhängig, alle drei Zeilen nicht. Also gilt $\operatorname{rg} A = 2$.

Zwischen Rang und Defekt besteht folgender Zusammenhang:

Satz 7.3.23. Sei $A \in M_{m,n}$. Dann gilt die *Dimensionsregel*:

$$\dim(\ker A) + \dim(\operatorname{Im} A) = n.$$

Also: Rang von A + Defekt von A = Anzahl der Spalten von A .

Im Beispiel von oben: $\dim(\ker A) = 4 - \operatorname{rg} A = 4 - 2 = 2$.

Man kann am Rang einer $n \times n$ -Matrix erkennen, ob diese invertierbar ist:

Satz 7.3.24 (Invertierbarkeitskriterien). Sei $A \in M_{n,n}$. Dann sind äquivalent:

- (i) A ist invertierbar,
- (ii) $\operatorname{rg} A = n$,
- (iii) $\ker A = \{0\}$,
- (iv) $\operatorname{Im} A = \mathbb{R}^n$,
- (v) Φ_A ist umkehrbar (also bijektiv).

7.4 Lineare Gleichungssysteme

Wir betrachten lineare Gleichungssysteme (kurz LGS) der Form

$$\begin{array}{cccccc} a_{11}x_1 & +a_{12}x_2 & + \cdots & +a_{1n}x_n & = & b_1 \\ \vdots & & & & & \vdots \\ a_{m1}x_1 & +a_{m2}x_2 & + \cdots & +a_{mn}x_n & = & b_m \end{array}$$

Man fasst die a_{ij} zur *Koeffizientenmatrix* $A = (a_{ij})$ und die b_i zur *rechten Seite* $b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$

zusammen. Dann lautet das LGS in Matrixschreibweise: $Ax = b$. Zu gegebenem A und b sind alle Lösungsvektoren $x \in \mathbb{R}^n$ gesucht. Im Fall $b \neq 0$ heißt das LGS $Ax = b$ *inhomogen*, sonst *homogen*. Zu dem LGS $Ax = b$ heißt $Ax = 0$ das zugeordnete homogene System. Schließlich heißt

$$(A|b) := \begin{pmatrix} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{pmatrix}$$

die *erweiterte Koeffizientenmatrix*.

Wann ist $Ax = b$ zu gegebenem A und b lösbar?

Satz 7.4.1. Sei $A \in M_{m,n}(\mathbb{R})$ und $b \in \mathbb{R}^m$. Dann sind äquivalent

- (i) $Ax = b$ ist lösbar
- (ii) $b \in \operatorname{Im} A$

$$(iii) \quad \boxed{\operatorname{rg} A = \operatorname{rg}(A|b)}$$

Gilt $\operatorname{rg} A = m$, so ist $Ax = b$ für beliebiges $b \in \mathbb{R}^m$ lösbar.

Beispiel 7.4.2: (a)

$$A = \begin{pmatrix} 1 & 0 \\ -1 & 0 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Hier ist $\operatorname{rg} A = 1$ und $\operatorname{rg}(A|b) = 2$, also ist $Ax = b$ nicht lösbar.

(b)

$$A = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \end{pmatrix} \in M_{2,3}$$

hat $\operatorname{rg} A = 2 (= m)$, also ist $Ax = b$ für jedes $b \in \mathbb{R}^2$ lösbar.

Wie bekommt man alle Lösungen von $Ax = b$?

Satz 7.4.3. Sei $A \in M_{m,n}(\mathbb{R})$, $b \in \mathbb{R}^m$ und x_s eine spezielle Lösung von $Ax = b$. Dann ist die Lösungsmenge (=Menge aller Lösungen) von $Ax = b$ gleich

$$x_s + \ker A = \{x_s + x : Ax = 0\}.$$

Beispiel 7.4.4: Alle Lösungen von $Ax = b$ für $A = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \end{pmatrix}$, $b = \begin{pmatrix} -1 \\ 2 \end{pmatrix}$.

1. Schritt:

Bestimme eine spezielle Lösung x_s von $Ax = b$, hier z.B. $x_s = \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix}$.

2. Schritt:

Bestimme $\ker A$, d.h. alle Lösungen von $Ax = 0$.

Hier:

$$\left. \begin{array}{l} x_1 + 2x_3 = 0 \\ x_2 + x_3 = 0 \end{array} \right\} \Rightarrow \begin{array}{l} x_1 = -2x_3 \\ x_2 = -x_3 \end{array}$$

Also

$$\{x : Ax = 0\} = \left\{ x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} : x_1 = -2x_3, x_2 = -x_3 \right\} = \left\{ \begin{pmatrix} -2x_3 \\ -x_3 \\ x_3 \end{pmatrix} : x_3 \in \mathbb{R} \right\}.$$

Lösungsmenge $\ker A$ hat einen Parameter (etwa x_3). Dies ist schon zuvor klar nach Satz 7.3.23, denn $\dim(\ker A) = n - \operatorname{rg} A = 3 - 2 = 1$. Bezeichnung des Parameters als x_3 ist nicht gut. Besser ist z.B.

$$\ker A = \left\{ \lambda \begin{pmatrix} -2 \\ -1 \\ 1 \end{pmatrix} : \lambda \in \mathbb{R} \right\}.$$

3. Schritt:

Bilde $x_s + \ker A$. Hier: Lösungsmenge von $Ax = b$ ist

$$\left\{ \left(\begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix} + \lambda \begin{pmatrix} -2 \\ -1 \\ 1 \end{pmatrix} : \lambda \in \mathbb{R} \right\}.$$

Wie berechnet man den Rang von A , eine spezielle Lösung von $Ax = b$ und schließlich alle Lösungen von $Ax = b$ systematisch?

7.5 Der Gaußsche Algorithmus

Der Gaußsche Algorithmus oder das Gaußsche Eliminationsverfahren zur Lösung eines linearen Gleichungssystems beruht auf der Beobachtung, dass sich die Lösungsmenge dieses Systems nicht ändert, wenn in der *erweiterten* Systemmatrix die folgenden *elementaren Zeilenoperatoren* durchgeführt werden:

- (a) Addition einer Zeile zu einer anderen.
- (b) Multiplikation einer Zeile mit einer Zahl $\neq 0$.
- (c) Vertauschen zweier Zeilen.

Außerdem wollen wir noch zulassen:

- (d) Vertauschen zweier Spalten der Systemmatrix.

Letzteres entspricht einer Umbenennung der gesuchten Zahlen x_1, \dots, x_n . Ziel ist es, durch diese Operationen eine „Dreiecks-“ oder wenigstens „Trapez-Gestalt“ der Systemmatrix zu erreichen, aus der man die Lösung durch „Rückwärtseinsetzen“ gewinnt. Sehen wir uns zunächst einige Beispiele an.

Beispiel 1 Das lineare Gleichungssystem

$$\begin{array}{cccccc} x_1 & + & x_2 & + & x_3 & + & x_4 & = & 1 \\ x_1 & + & 2x_2 & + & 2x_3 & + & 2x_4 & = & 2 \\ x_1 & + & x_2 & + & 2x_3 & + & 2x_4 & = & 2 \\ x_1 & + & x_2 & + & x_3 & + & 2x_4 & = & 2 \end{array}$$

hat die erweiterte Systemmatrix

$$\left(\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 & 2 \\ 1 & 1 & 2 & 2 & 2 \\ 1 & 1 & 1 & 2 & 2 \end{array} \right).$$

Subtraktion der ersten Zeile von der zweiten, dritten und vierten Zeile führt auf

$$\left(\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{array} \right),$$

d.h.

$$\begin{array}{rcl} x_1 + x_2 + x_3 + x_4 & = & 1 \\ x_2 + x_3 + x_4 & = & 1 \\ x_3 + x_4 & = & 1 \\ x_4 & = & 1. \end{array}$$

Wir erhalten ein System von Dreiecksgestalt. Aus der letzten Gleichung folgt $x_4 = 1$. Wir setzen dies in die vorletzte Gleichung ein und erhalten $x_3 = 0$. In die zweite Gleichung eingesetzt liefert dies $x_2 = 0$, und aus der ersten Gleichung folgt schließlich, dass auch $x_1 = 0$. Die Lösung des Systems ist also eindeutig bestimmt und lautet $(0, 0, 0, 1)^T$, d. h. sie besteht aus einem einzigen Punkt. \square

Beispiel 2 Das Gleichungssystem werde durch die erweiterte Systemmatrix

$$\left(\begin{array}{ccc|c} 0 & 3 & 6 & 3 \\ 1 & 2 & 3 & 1 \\ -2 & -4 & -6 & -2 \end{array} \right)$$

beschrieben. Austausch von erster und zweiter Zeile liefert

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 0 & 3 & 6 & 3 \\ -2 & -4 & -6 & -2 \end{array} \right).$$

Addition des Doppelten der ersten Zeile zur letzten Zeile ergibt

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 0 & 3 & 6 & 3 \\ 0 & 0 & 0 & 0 \end{array} \right),$$

und Division der zweiten Zeile durch 3 liefert schließlich

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 0 \end{array} \right) \tag{7.2}$$

bzw. das Gleichungssystem

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= 1 \\ x_2 + 2x_3 &= 1. \end{aligned} \tag{7.3}$$

Die letzte Zeile der Matrix (7.2) entspricht der Gleichung $0 = 0$, die wir natürlich weglassen. Zum Auflösen von (7.3) wählt man $x_3 = t$ als Parameter. Aus der letzten Gleichung von (7.3) folgt dann

$$x_2 + 2t = 1 \quad \text{bzw.} \quad x_2 = 1 - 2t,$$

und durch Einsetzen in die erste Zeile findet man

$$x_1 + 2(1 - 2t) + 3t = 1 \quad \text{bzw.} \quad x_1 = -1 + t.$$

Die Lösungsmenge des Systems besteht also aus allen Vektoren der Form

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -1 + t \\ 1 - 2t \\ t \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} + t \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}, \quad t \in \mathbb{R}.$$

Hier erkennen wir die in Satz 7.4.3 beschriebene Struktur: die Lösung des inhomogenen Systems setzt sich zusammen aus einer speziellen Lösung des inhomogenen Systems, nämlich $(-1, 1, 0)^T$, und der allgemeinen Lösung des zugehörigen homogenen Systems, nämlich allen Vielfachen von $(1, -2, 1)^T$. Geometrisch gesehen ist die Lösungsmenge also eine Gerade. \square

Beispiel 3 Wir betrachten das System mit der erweiterten Systemmatrix

$$\left(\begin{array}{ccc|c} 0 & 3 & 6 & 3 \\ 1 & 2 & 3 & 1 \\ -2 & -4 & -6 & 1 \end{array} \right),$$

die sich von Beispiel 2 nur vom Eintrag in der rechten unteren Ecke unterscheidet. Die im Beispiel 2 durchgeführten Operationen führen nun auf

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 3 \end{array} \right).$$

Die letzte Gleichung des umgeformten Systems lautet also $0 = 3$. Dieser Widerspruch zeigt, dass das System keine Lösung besitzt. Die Lösungsmenge ist also leer. \square

Wir kommen nun zum allgemeinen LGS zurück. Durch die elementaren Operationen (a) - (d) kann man die erweiterte Systemmatrix dieses Systems stets in die folgende Gestalt bringen:

$$\left(\begin{array}{cccc|cccc|c} 1 & c_{12} & c_{13} & \dots & c_{1r} & c_{1,r+1} & \dots & c_{1n} & d_1 \\ & 1 & c_{23} & \dots & c_{2r} & c_{2,r+1} & \dots & c_{2n} & d_2 \\ & & 1 & \dots & c_{3r} & c_{3,r+1} & \dots & c_{3n} & d_3 \\ & & & \ddots & \vdots & \vdots & & \vdots & \vdots \\ 0 & & & & 1 & c_{r,r+1} & \dots & c_{rn} & d_r \\ \hline & & & & & & & & d_{r+1} \\ & & 0 & & & & 0 & & \vdots \\ & & & & & & & & d_m \end{array} \right)_{m,n}. \quad (7.4)$$

Aus dieser Darstellung können wir alle gewünschten Informationen über das Gleichungssystem gewinnen:

1. Das System ist genau dann lösbar, wenn

$$d_{r+1} = \dots = d_m = 0. \quad (7.5)$$

2. Der Rang der Systemmatrix ist gleich r , und die Dimension des Kerns von A ist $n - r$.
3. Ist die Bedingung (7.5) erfüllt, so können wir die Lösung des Systems wie folgt gewinnen:
Wir wählen $n - r = n - \text{rang } A$ Parameter $t_1, t_2, \dots, t_{n-r} \in \mathbb{R}$ oder \mathbb{C} und setzen

$$x_{r+1} = t_1, \quad x_{r+2} = t_2, \dots, \quad x_n = t_{n-r}.$$

Die r -te Gleichung von (7.4) lautet dann

$$x_r + c_{r,r+1}t_1 + \dots + c_{rn}t_{n-r} = d_r$$

bzw.

$$x_r = d_r - c_{r,r+1}t_1 - \dots - c_{rn}t_{n-r}.$$

Wir setzen dies in die $(r-1)$ -te Gleichung von (7.4) ein und erhalten eine Darstellung von x_{r-1} in Abhängigkeit von t_1, \dots, t_{n-r} , usw.

Wir fahren so fort bis zur ersten Zeile und erhalten die Lösung in der Gestalt

$$\begin{aligned} x_1 &= e_1 + f_{11}t_1 + \dots + f_{1,n-r}t_{n-r}, \\ &\vdots \\ x_r &= e_r + f_{r1}t_1 + \dots + f_{r,n-r}t_{n-r}, \\ x_{r+1} &= t_1 \\ &\vdots \\ x_n &= t_{n-r} \end{aligned}$$

mit gewissen Zahlen e_i und f_{ij} und Parametern t_1, \dots, t_{n-r} . In Vektorschreibweise lautet die Lösung

$$\begin{pmatrix} x_1 \\ \vdots \\ x_r \\ x_{r+1} \\ x_{r+2} \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} e_1 \\ \vdots \\ e_r \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + t_1 \begin{pmatrix} f_{11} \\ \vdots \\ f_{r1} \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \dots + t_{n-r} \begin{pmatrix} f_{1,n-r} \\ \vdots \\ f_{r,n-r} \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}, \quad (7.6)$$

in der wir wieder die Lösungsstruktur aus Satz 7.4.3 erkennen.

Man kann sich das „Rückwärtseinsetzen“ sparen, wenn man (ähnlich wie bei der Rangberechnung angedeutet) durch weitere elementare Operationen aus (7.4) eine Matrix der Gestalt

$$\left(\begin{array}{cccc|ccc|c} 1 & 0 & 0 & \dots & 0 & \hat{c}_{1,r+1} & \dots & \hat{c}_{1n} & \hat{d}_1 \\ & 1 & 0 & \dots & 0 & \hat{c}_{2,r+1} & \dots & \hat{c}_{2n} & \hat{d}_2 \\ & & 1 & \dots & 0 & \hat{c}_{3,r+1} & \dots & \hat{c}_{3n} & \hat{d}_3 \\ & & & \ddots & \vdots & \vdots & & \vdots & \vdots \\ & & & & 1 & \hat{c}_{r,r+1} & \dots & \hat{c}_{rn} & \hat{d}_r \\ \hline & & & & & & & & \hat{d}_{r+1} \\ & & & & & & & & \vdots \\ & & & & & & & & \hat{d}_n \end{array} \right)$$

erzeugt, aus der man im Falle $\hat{d}_{r+1} = \dots = \hat{d}_n = 0$ sofort die Darstellung der Lösung in der Form (7.6) ablesen kann. Wir sehen uns ein abschließendes Beispiel an.

Beispiel Die erweiterte Systemmatrix eines linearen Gleichungssystems sei

$$\left(\begin{array}{cccc|c} 1 & 2 & 1 & -1 & 2 \\ -1 & -1 & -1 & 3 & -4 \\ 2 & -1 & 2 & 2 & 8 \\ 2 & 0 & 2 & 4 & 6 \\ 4 & 2 & 4 & -2 & 14 \\ 2 & -2 & 2 & 0 & 10 \end{array} \right).$$

Durch elementare Operationen gelangen wir nacheinander zu den Matrizen

$$\begin{pmatrix} 1 & 2 & 1 & -1 & | & 2 \\ 0 & 1 & 0 & 2 & | & -2 \\ 0 & -5 & 0 & 4 & | & 4 \\ 0 & -4 & 0 & 6 & | & 2 \\ 0 & -6 & 0 & 2 & | & 6 \\ 0 & -6 & 0 & 2 & | & 6 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & 2 & 1 & -1 & | & 2 \\ 0 & 1 & 0 & 2 & | & -2 \\ 0 & 0 & 0 & 14 & | & -6 \\ 0 & 0 & 0 & 14 & | & -6 \\ 0 & 0 & 0 & 14 & | & -6 \\ 0 & 0 & 0 & 14 & | & -6 \end{pmatrix} \Rightarrow$$

$$\begin{pmatrix} 1 & 2 & 1 & -1 & | & 2 \\ 0 & 1 & 0 & 2 & | & -2 \\ 0 & 0 & 0 & 1 & | & -3/7 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & 2 & -1 & 1 & | & 2 \\ 0 & 1 & 2 & 0 & | & -2 \\ 0 & 0 & 1 & 0 & | & -3/7 \end{pmatrix},$$

wobei wir Nullzeilen weglassen und im letzten Schritt die dritte und vierte Spalte getauscht haben, wodurch x_3 und x_4 ihre Rollen tauschen. Wir erhalten daher

$$\begin{aligned} x_4 &= -3/7, \\ x_2 + 2x_4 &= -2, \quad \text{d.h.} \quad x_2 = -2 + 6/7 = -8/7, \\ x_1 + 2x_2 - x_4 + x_3 &= 2, \quad \text{d.h.} \quad x_1 = 2 + 16/7 - 3/7 - x_3. \end{aligned}$$

Mit dem Parameter $x_3 = t$ folgt schließlich

$$x_1 = 27/7 - t, \quad x_2 = -8/7, \quad x_3 = t, \quad x_4 = -3/7$$

bzw.

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 27/7 \\ -8/7 \\ 0 \\ -3/7 \end{pmatrix} + t \begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \end{pmatrix}. \quad \square$$

Möchte man mehrere LGS mit gleicher Systemmatrix (jedoch verschiedenen rechten Seiten) lösen, bietet es sich an, die Systemmatrix um alle diese Spalten zu erweitern und den Gaußalgorithmus für alle rechten Seiten gleichzeitig durchzuführen. Ein spezielles Problem dieser Art ist die *Matrixinversion*:

Gegeben sei eine invertierbare Matrix $A \in \mathbb{R}^{n \times n}$.

Gesucht ist eine Matrix $X \in \mathbb{R}^{n \times n}$ mit $AX = I_n$.

Die Spalten $x_i \in \mathbb{R}^n$ der Matrix X kann man durch Lösen der Gleichungssysteme

$$Ax_i = e_i \quad i = 1, \dots, n \tag{7.7}$$

finden. Man muss also n Gleichungssysteme mit derselben Koeffizientenmatrix A lösen und dann die Inverse $X = A^{-1}$ aus ihren Spalten x_i zusammensetzen. Das ist das einfachste und stabilste Verfahren zur Matrixinversion.

7.6 Determinanten

Das Gleichungssystem

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= b_1 \\ a_{21}x_1 + a_{22}x_2 &= b_2 \end{aligned}$$

ist für jede rechte Seite $b = (b_1, b_2)$ eindeutig lösbar, falls

$$a_{11}a_{22} - a_{12}a_{21} \neq 0$$

gilt. Den Ausdruck $a_{11}a_{22} - a_{12}a_{21}$ bezeichnet man als Determinante der 2×2 -Matrix $A = (a_{ij})$. Schreibweise ist $\det A$ (oder auch $|A|$), also

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} := a_{11}a_{22} - a_{12}a_{21}.$$

Um allgemein die Determinante einer $n \times n$ -Matrix zu definieren, benötigen wir zunächst den Begriff der *Permutation*. Unter einer Permutation der Zahlen $1, 2, \dots, n$ versteht man eine Umordnung dieser Zahlen, z.B. (für $n = 4$)

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 2 & 4 \end{pmatrix}.$$

Jeder solchen Umordnung entspricht eine bijektive Abbildung $\varphi : \{1, 2, \dots, n\} \rightarrow \{1, 2, \dots, n\}$.

Im Beispiel von oben ist

$$\varphi(1) = 3, \quad \varphi(2) = 1, \quad \varphi(3) = 2, \quad \varphi(4) = 4.$$

Zu gegebenem $n \in \mathbb{N}$ bezeichne S_n die Menge aller möglichen Permutationen von $1, 2, \dots, n$. Für $\varphi \in S_n$ setzen wir

$$\operatorname{sgn} \varphi = \prod_{\substack{i,j \\ i < j}} \frac{\varphi(i) - \varphi(j)}{i - j}$$

(Produkt über alle $i, j \in \{1, \dots, n\}$ für die $i < j$ gilt). $\operatorname{sgn} \varphi$ heißt *Signum von φ* .

Beispiel 7.6.1:

$$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}, \text{ also } \varphi \in S_3 \text{ mit } \varphi(1) = 3, \varphi(2) = 1, \varphi(3) = 2.$$

Hier ist das Produkt über die Indizes $(i, j) = (1, 2), (1, 3), (2, 3)$ zu bilden.

$$\Rightarrow \operatorname{sgn} \varphi = \frac{\varphi(1) - \varphi(2)}{1 - 2} \cdot \frac{\varphi(1) - \varphi(3)}{1 - 3} \cdot \frac{\varphi(2) - \varphi(3)}{2 - 3} = \frac{(3 - 1) \cdot (3 - 2) \cdot (1 - 2)}{(1 - 2) \cdot (1 - 3) \cdot (2 - 3)} = 1.$$

Es gilt immer entweder $\operatorname{sgn} \varphi = 1$ oder $\operatorname{sgn} \varphi = -1$. Im ersten Fall heißt die Permutation φ *gerade*, sonst *ungerade*.

Bemerkung 7.6.2. Jede Permutation kann durch mehrfaches Austauschen von jeweils zwei Elementen erzielt werden. Für eine (un-)gerade Permutation benötigt man eine (un-)gerade Anzahl von Vertauschungen.

Definition 7.6.3 (Determinanten). Sei $A \in M_{n,n}$ mit $A = (a_{ij})$. Dann heißt

$$\sum_{\varphi \in S_n} (\operatorname{sgn} \varphi) a_{1\varphi(1)} a_{2\varphi(2)} \cdots a_{n\varphi(n)}$$

die *Determinante* von A , geschrieben als $\det A$ oder

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

Beispiele 7.6.4: (a) $n = 1$, $A = (a_{11})$. Hier ist $\det A = a_{11}$.

(b) $n = 2$; mögliche Permutationen: $\begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}$ gerade $\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$ ungerade. Dies liefert

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}.$$

(c) $n = 3$; Wir notieren nur das Ergebnis:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = + a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}$$

Dies lässt sich folgendermaßen merken (*Regel von Sarrus*):

Schreibe die ersten beiden Spalten der Matrix nochmals rechts neben die Determinante. Bilde die Produkte entlang den Diagonalen von links oben nach rechts unten (+) und von rechts oben nach links unten (-):

$$\begin{array}{cccccc} a_{11} & a_{12} & a_{13} & a_{11} & a_{12} & \\ & \ddots & \diagdown & \diagdown & / & \\ a_{21} & a_{22} & a_{23} & a_{21} & a_{22} & \\ & / & \diagdown & \diagdown & \ddots & \\ a_{31} & a_{32} & a_{33} & a_{31} & a_{32} & \\ - & - & - & + & + & + \end{array}$$

Versehe die Produkte mit dem Vorzeichen laut Skizze und addiere auf.

Beachte: Die Regel von Sarrus gilt nur für 3×3 -Matrizen!

(d) A sei eine „Dreiecksmatrix“ der Form

$$A = \begin{pmatrix} a_{11} & \cdots & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{nn} \end{pmatrix}$$

Dann gilt $\boxed{\det A = a_{11} \cdot a_{22} \cdot \cdots \cdot a_{nn}}$. Insbesondere ist $\det I_n = 1$. Die konkrete *Berechnung* von $\det A$ erfolgt für $n > 3$ am zweckmäßigsten durch Anwendung des Gauß-Algorithmus. Es ist auch das Zurückführen auf die Berechnung kleinerer Determinanten möglich nach Satz 7.6.5. Das ist aber für große Matrizen zu aufwändig.

Satz 7.6.5 (Laplace'scher Entwicklungssatz). Sei $A = (a_{ij}) \in M_{n,n}$. Dann gilt:

$$\det A = \sum_{j=1}^n (-1)^{i+j} a_{ij} \det S_{ij} \quad (\text{Entwicklung nach der } i\text{-ten Zeile})$$

$$\det A = \sum_{i=1}^n (-1)^{i+j} a_{ij} \det S_{ij} \quad (\text{Entwicklung nach der } j\text{-ten Spalte})$$

Dabei ist S_{ij} diejenige Matrix, die aus A durch Streichung der i -ten Zeile und j -ten Spalte entsteht.

Beispiel 7.6.6:

$$\begin{vmatrix} -1 & 4 & 2 & 1 \\ 2 & 1 & -2 & 1 \\ 1 & 0 & 0 & 3 \\ 3 & -1 & 2 & -1 \end{vmatrix}$$

Günstig ist hier die Entwicklung nach der 3. Zeile, da diese nur zwei Einträge $\neq 0$ enthält! Dies liefert

$$= (-1)^{3+1} \cdot 1 \cdot \begin{vmatrix} 4 & 2 & 1 \\ 1 & -2 & 1 \\ -1 & 2 & -1 \end{vmatrix} + (-1)^{3+4} \cdot 3 \cdot \begin{vmatrix} -1 & 4 & 2 \\ 2 & 1 & -2 \\ 3 & -1 & 2 \end{vmatrix}$$

Jetzt kann mit der Regel von Sarrus weitergerechnet werden

$$= 8 - 2 + 2 - 2 - 8 + 2 - 3 \cdot (-2 - 24 - 4 - 6 + 2 - 16) = 150.$$

Rechenregeln für Determinanten. Für $A, B \in M_{n,n}$ gelten

1. $\det(A \cdot B) = \det A \det B$
2. $\det A^{-1} = \frac{1}{\det A}$, falls A invertierbar ist
3. $\det A^T = \det A$

Zur Formulierung der nachfolgenden Regeln bezeichnen wir die Spalten von $A = (a_{ij})$ mit a^1, \dots, a^n und schreiben $\det(a^1, \dots, a^n)$ für $\det A$. Es gilt

4. $\det(a^1, \dots, a^{i-1}, \lambda a^i, a^{i+1}, \dots, a^n) = \lambda \det(a^1, \dots, a^{i-1}, a^i, a^{i+1}, \dots, a^n)$
5. $\det(a^1, \dots, a^{i-1}, a^i + b, a^{i+1}, \dots, a^n) = \det(a^1, \dots, a^{i-1}, a^i, a^{i+1}, \dots, a^n) + \det(a^1, \dots, a^{i-1}, b, a^{i+1}, \dots, a^n)$
6. Beim Austausch zweier Spalten (oder Zeilen) von A wechselt $\det A$ das Vorzeichen.
7. Sind zwei Spalten (oder Zeilen) von A gleich, so ist $\det A = 0$.
8. Addition des λ -fachen der i -ten Spalte (Zeile) von A zur j -ten Spalte (Zeile) ändert $\det A$ nicht, wenn $i \neq j$ ist.

Die Regeln (f) und (h) werden auch bei der Lösung eines LGS mit dem Gauß-Algorithmus verwendet.

Satz 7.6.7. Sei $A \in M_{n,n}(\mathbb{R})$. Dann sind äquivalent:

1. $\det A \neq 0$.
2. A ist invertierbar.
3. $Ax = b$ ist für jedes $b \in \mathbb{R}^n$ eindeutig lösbar.

Bemerkung 7.6.8. Teil (c) bedeutet, dass $Ax = 0$ nur die Lösung $x = 0$ hat. Also folgt: $Ax = 0$ hat Lösungen $x \neq 0 \Leftrightarrow \det A = 0$.

Satz 7.6.9. Seien $a_1, \dots, a_n \in \mathbb{R}^n$. Dann sind äquivalent:

- (i) a_1, \dots, a_n sind linear unabhängig.
- (ii) $\det A \neq 0$, wenn A die Spaltenvektoren a_1, \dots, a_n hat.

Für kleine n kann $Ax = b$ sinnvoll mit folgender Determinanten-Regel formelmäßig gelöst werden:

Satz 7.6.10 (Cramer-Regel). Sei $A = (\alpha_{ij}) \in M_{n,n}(\mathbb{R})$ mit $\det A \neq 0$ und sei $b \in \mathbb{R}^n$. Dann ist die eindeutige Lösung von $Ax = b$ durch

$$x_i = \frac{1}{\det A} \det C_i \quad i = 1, \dots, n$$

gegeben. Dabei ist

$$C_i = \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1,i-1} & b_1 & \alpha_{1,i+1} & \cdots & \alpha_{1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \alpha_{n1} & \cdots & \alpha_{n,i-1} & b_n & \alpha_{n,i+1} & \cdots & \alpha_{n,n} \end{pmatrix},$$

d.h. die i -te Spalte von A wird durch b ersetzt.

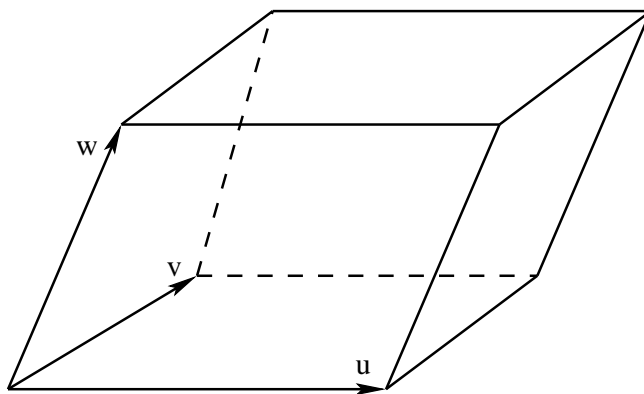
Bemerkungen 7.6.11: (i) Im Spezialfall $n = 3$ kann die Determinante als Merkregel für das *Vektorprodukt* (auch äußeres Produkt) $x \times y$ für $x, y \in \mathbb{R}^3$ formal verwendet werden. Die Einheitsvektoren in Richtung der Koordinatenachsen seien e_1, e_2, e_3 . Dann gilt:

$$x \times y = \begin{vmatrix} e_1 & e_2 & e_3 \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{vmatrix} = e_1 \begin{vmatrix} x_2 & x_3 \\ y_2 & y_3 \end{vmatrix} - e_2 \begin{vmatrix} x_1 & x_3 \\ y_1 & y_3 \end{vmatrix} + e_3 \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} = \begin{pmatrix} x_2 y_3 - x_3 y_2 \\ x_3 y_1 - x_1 y_3 \\ x_1 y_2 - x_2 y_1 \end{pmatrix}$$

(ii) Zur Berechnung des *Spatproduktes* (d.h. $\langle x \times y, z \rangle$). Es ist

$$\langle x \times y, z \rangle = \begin{vmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{vmatrix} \quad \text{für } x, y, z \in \mathbb{R}^3.$$

$|\langle x \times y, z \rangle|$ ist das Volumen des von x, y, z aufgespannten *Spates*.



7.7 Anwendung der Linearen Algebra in der Geometrie

Die Sprache der Vektorrechnung ist wie bereits zu Beginn des Kapitels erwähnt ein starkes Hilfsmittel zur Behandlung elementargeometrischer Fragen. Wir wollen dieses noch kurz im zweidimensionalen und im nächsten Abschnitt im dreidimensionalen Fall beleuchten.

7.7.1 Geometrie in der Ebene

Wir beginnen mit dem schon früher kurz angeklungenen Begriff der Gerade.

Definition 7.7.1. Seien $u, v \in \mathbb{R}^2$ mit $v \neq 0$. Dann heißt $g := \{u + \lambda v : \lambda \in \mathbb{R}\}$ eine *Gerade*. Der Vektor u wird *Aufpunkt* und v *Richtungsvektor* der Gerade genannt.

Diese Form der Darstellung der Geraden heißt *Parameterdarstellung*.

Bemerkung 7.7.2. Oft wird eine Gerade durch 2 Punkte $P, Q \in \mathbb{R}^2$ gegeben (Zwei-Punkte-Form). Dann kann man $u = \overrightarrow{OP}$ und $v = \overrightarrow{PQ}$ in der Parameterdarstellung nehmen.

Bemerkungen 7.7.3: (i) Seien $a, b, c \in \mathbb{R}$ und a und b seien nicht beide Null. Dann ist die Lösungsmenge der Gleichung $ax_1 + bx_2 = c$ immer eine Gerade (vgl. Satz 7.4.3). Eine solche Gleichung heißt darum auch *Geradengleichung*.

(ii) Umgekehrt lässt sich auch jede Gerade als Lösungsmenge einer solchen Gleichung auffassen. Um das einzusehen sei g eine Gerade mit Aufpunkt u und Richtungsvektor v . Dann ist $x = (x_1, x_2)^T \in g$, genau dann wenn es ein $\lambda \in \mathbb{R}$ gibt mit $u + \lambda v = x$, d.h. es muss $u_1 + \lambda v_1 = x_1$ und $u_2 + \lambda v_2 = x_2$ gelten.

Da $v \neq 0$ ist, muss $v_1 \neq 0$ oder $v_2 \neq 0$ gelten. Wir beschränken uns auf den Fall $v_1 \neq 0$, der andere geht analog. Dann können wir die erste Gleichung nach λ auflösen und erhalten $\lambda = (x_1 - u_1)/v_1$. Also gilt

$$\begin{aligned} x \in g &\iff u_2 + \frac{x_1 - u_1}{v_1} v_2 = x_2 \iff v_1 u_2 + x_1 v_2 - u_1 v_2 = x_2 v_1 \\ &\iff v_2 x_1 - v_1 x_2 = u_1 v_2 - v_1 u_2 \end{aligned}$$

und das ist eine Gleichung wie in (i) mit $a = v_2$, $b = -v_1$ und $c = u_1 v_2 - v_1 u_2$.

Ein und dieselbe Gerade hat viele verschiedene Darstellungen als Gleichung, so haben z.B. die beiden Geradengleichungen $x_1 - x_2 = 1$ und $2x_1 - 2x_2 = 2$ offensichtlich die selben Lösungen. Unter diesen verschiedenen Darstellungen einer Geraden als Gleichung wollen wir eine besonders praktische gesondert erwähnen.

Seien dazu $u \in \mathbb{R}^2$, $v \in \mathbb{R}^2 \setminus \{0\}$ und $g = \{u + \lambda v : \lambda \in \mathbb{R}\}$. Da $v \neq 0$ ist, gibt es einen Vektor $n \in \mathbb{R}^2$, der orthogonal zu v ist und $\|n\| = 1$ erfüllt. Dies ist ein *Normaleneinheitsvektor* von g . Mit diesem gilt nun

$$g = \{x = (x_1, x_2)^T \in \mathbb{R}^2 : \langle n, x \rangle = \langle n, u \rangle\},$$

d.h. g ist durch die Gleichung $n_1 x_1 + n_2 x_2 = n_1 u_1 + n_2 u_2$ gegeben. Diese Geradengleichung heißt *Hesse-Normalform* (HNF) von g .

Beispiel 7.7.4: Sei $g = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \lambda \begin{pmatrix} -1 \\ 1 \end{pmatrix} : \lambda \in \mathbb{R} \right\}$

Hier ist also $v = (-1, 1)^\top$ und ein dazu senkrechter Vektor ist z.B. $(1, 1)^\top$, wir nehmen also als Normaleneinheitsvektor

$$n = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

und erhalten als Hesse-Normalform von g

$$\frac{1}{\sqrt{2}}x_1 + \frac{1}{\sqrt{2}}x_2 = \left\langle \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\rangle = \frac{1}{\sqrt{2}} \cdot 2 = \sqrt{2}.$$

Ein Grund für die Wichtigkeit der Hesse-Normalform ist, dass damit Abstandsbestimmungen einfach sind, denn es gilt der folgende Satz.

Satz 7.7.5. Sei $g \subseteq \mathbb{R}^2$ eine Gerade mit HNF $ax_1 + bx_2 = c$ und sei $w \in \mathbb{R}^2$. Dann ist der Abstand von w zu g gegeben durch $|aw_1 + bw_2 - c|$. Insbesondere ist $|c|$ der Abstand von g zum Ursprung.

Für die Gerade aus Beispiel 7.7.4 bekommen wir also sofort $\sqrt{2}$ für den Abstand zum Ursprung.

Beispiel 7.7.6: Gegeben sind zwei Geraden

$$g_1 = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \lambda \begin{pmatrix} -1 \\ 1 \end{pmatrix} : \lambda \in \mathbb{R} \right\} \quad \text{und} \quad g_2 = \left\{ \mu \begin{pmatrix} -2 \\ 1 \end{pmatrix} : \mu \in \mathbb{R} \right\}$$

und gesucht ist deren Schnittpunkt, sowie der Schnittwinkel zwischen den beiden Geraden.

Schnittpunkt: Ein Punkt $x \in \mathbb{R}^2$ ist der Schnittpunkt der beiden Geraden, wenn $x \in g_1 \cap g_2$ liegt, wenn es also reelle Zahlen λ und μ gibt mit

$$x = \mu \begin{pmatrix} -2 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \lambda \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \quad \text{d.h. mit} \quad \mu \begin{pmatrix} -2 \\ 1 \end{pmatrix} + \lambda \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Dieses LGS lässt sich leicht lösen:

$$\left(\begin{array}{cc|c} -2 & 1 & 1 \\ 1 & -1 & 1 \end{array} \right) \begin{array}{l} \uparrow \\ \cdot 2 \end{array} \rightsquigarrow \left(\begin{array}{cc|c} 0 & -1 & 3 \\ 1 & -1 & 1 \end{array} \right) \begin{array}{l} -1 \\ \leftarrow \end{array} \rightsquigarrow \left(\begin{array}{cc|c} 0 & 1 & -3 \\ 1 & 0 & -2 \end{array} \right).$$

Damit ist $\mu = -2$, d.h. wir bekommen $x = -2 \cdot (-2, 1)^\top = (4, -2)^\top$ als Schnittpunkt.

Schnittwinkel: Der Schnittwinkel der beiden Geraden ist gerade der Winkel zwischen den beiden Richtungsvektoren v_1 und v_2 . Für diesen Winkel φ gilt die Beziehung $\langle v_1, v_2 \rangle = \|v_1\| \cdot \|v_2\| \cdot \cos(\varphi)$, also ist

$$\cos(\varphi) = \frac{\left\langle \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \begin{pmatrix} -2 \\ 1 \end{pmatrix} \right\rangle}{\left\| \begin{pmatrix} -1 \\ 1 \end{pmatrix} \right\| \cdot \left\| \begin{pmatrix} -2 \\ 1 \end{pmatrix} \right\|} = \frac{3}{\sqrt{2} \cdot \sqrt{5}}.$$

Das ergibt mit dem Arcuscosinus $\varphi \approx 0,32175 \approx 18,43^\circ$.

7.7.2 Geometrie im Raum

Wir wenden uns nun dem Raum \mathbb{R}^3 zu. Die Definition einer Gerade bleibt hier unverändert die selbe wie in \mathbb{R}^2 , hinzu kommt als flaches geometrisches Gebilde die Ebene.

Definition 7.7.7. Seien $u, v, w \in \mathbb{R}^3$ und v und w seien linear unabhängig. Dann ist die Parameterdarstellung der Ebene mit Aufpunkt u und Richtungsvektoren v und w gegeben durch

$$E := \{u + \lambda v + \mu w : \lambda, \mu \in \mathbb{R}\}.$$

Bemerkung 7.7.8. Eine Ebene wird eindeutig festgelegt durch die Angabe von drei Punkten $P, Q, R \in \mathbb{R}^3$, die nicht auf einer Geraden liegen, d.h. für die die Vektoren \overrightarrow{PQ} und \overrightarrow{PR} linear unabhängig sind. Dann bekommt man eine Parameterdarstellung z.B. mit der Wahl $u = \overrightarrow{OP}$, $v = \overrightarrow{PQ}$ und $w = \overrightarrow{PR}$.

Auch in \mathbb{R}^3 gibt es einen direkten Zusammenhang zwischen Geraden, bzw. Ebenen und den Lösungsmengen von linearen Gleichungssystemen. Ist $A \in M_{3,3}(\mathbb{R})$ und $b \in \mathbb{R}^3$, so ist die Lösungsmenge von $Ax = b$ genau dann eine Gerade, wenn $\text{rg } A = 2$ und $b \in \text{Im}(A)$ ist. Umgekehrt gibt es auch zu jeder Geraden wieder ein solches Gleichungssystem mit zwei Gleichungen und drei Unbekannten.

Eine Ebene erhält man als Lösung von $Ax = b$, falls $\text{rg}(A) = 1$ ist und $b \in \text{Im}(A)$ liegt und auch hier gibt es zu jeder Ebene wieder eine Gleichung, als deren Lösungsmenge sie auftritt, eine *Ebenengleichung*.

Wir betrachten zunächst ein Beispiel für eine Gerade.

Beispiel 7.7.9: Es sei

$$g := \left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + \lambda \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} : \lambda \in \mathbb{R} \right\}.$$

Die Elemente $x \in g$ erfüllen jeweils für ein $\lambda \in \mathbb{R}$ die Gleichungen $x_1 = 1 + \lambda$, $x_2 = \lambda$ und $x_3 = 1 - \lambda$. Setzt man in der ersten und dritten Gleichung die zweite ein, erhält man das LGS

$$\begin{cases} x_1 - x_2 = 1 \\ x_2 + x_3 = 1, \end{cases}$$

dessen Lösungsmenge genau g ist.

Im \mathbb{R}^3 kann es keine Hesse-Normalform für Geraden geben, denn es gibt ja nun unendlich viele verschiedene Richtungen, die senkrecht zu einer gegebenen Gerade stehen. Dafür klappt es hier bei Ebenen, denn diese haben eine ausgezeichnete orthogonale Richtung.

Sei also $E \subseteq \mathbb{R}^3$ eine Ebene mit Aufpunkt u und linear unabhängigen Richtungsvektoren v und w . Dann gibt es einen Vektor $n \in \mathbb{R}^3$, der senkrecht auf v und auf w steht und $\|n\| = 1$ erfüllt. Dieser ist bis auf ein Vorzeichen eindeutig und heißt *Normaleneinheitsvektor* von E . Es gilt dann

$$E = \{x \in \mathbb{R}^3 : \langle n, x \rangle = \langle n, u \rangle\},$$

d.h. E ist die Lösungsmenge der Gleichung $n_1x_1 + n_2x_2 + n_3x_3 = n_1u_1 + n_2u_2 + n_3u_3$. Diese Ebenengleichung heißt *Hesse-Normalform* (HNF) von E .

Beispiel 7.7.10: Wir bestimmen die HNF von

$$E := \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \lambda \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + \mu \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} : \lambda, \mu \in \mathbb{R} \right\}.$$

Hier ist also $v = (0, 1, 0)^\top$ und $w = (1, 1, 1)^\top$. Ein zu diesen beiden Vektoren senkrechter Vektor ist offensichtlich $(1, 0, -1)^\top$, wir wählen also

$$n = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}.$$

(Dieses Rateverfahren stößt natürlich bei komplizierteren Vektoren an seine Grenzen, eine rechnerische Methode den Normaleneinheitsvektor zu finden werden wir in Kürze kennenlernen.) Wegen $\langle n, x \rangle = 1/\sqrt{2}x_1 - 1/\sqrt{2}x_3$ und $\langle n, u \rangle = 1/\sqrt{2}$ ist damit die HNF von E

$$\frac{1}{\sqrt{2}}x_1 - \frac{1}{\sqrt{2}}x_3 = \frac{1}{\sqrt{2}}.$$

Auch hier gilt der Satz, dass wir Abstände direkt mit der HNF ausrechnen können:

Satz 7.7.11. Sei $E \subseteq \mathbb{R}^3$ eine Ebene mit HNF $ax_1 + bx_2 + cx_3 = d$ und $w \in \mathbb{R}^3$. Dann ist der Abstand von w zu E gegeben durch $|aw_1 + bw_2 + cw_3 - d|$. Insbesondere ist der Abstand von E zum Ursprung gleich $|d|$

Beispiel 7.7.12: Wir bestimmen den Schnittpunkt von g aus Beispiel 7.7.9 und der Ebene aus Beispiel 7.7.10. Wir hatten dort jeweils gesehen, dass

$$x \in g \iff \begin{cases} x_1 - x_2 = 1 \\ x_2 + x_3 = 1 \end{cases} \quad \text{und} \quad x \in E \iff x_1 - x_3 = 1.$$

Der Schnittpunkt muss also alle drei Gleichungen gleichzeitig erfüllen. Das führt uns auf das LGS

$$\left(\begin{array}{ccc|c} 1 & -1 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & -1 & 1 \end{array} \right) \begin{array}{l} \cdot 1 \\ \leftarrow \quad \uparrow \\ \cdot 1 \end{array} \rightsquigarrow \left(\begin{array}{ccc|c} 1 & -1 & 0 & 1 \\ 2 & 0 & 0 & 3 \\ 1 & 0 & -1 & 1 \end{array} \right) \begin{array}{l} \uparrow \\ \cdot (-1/2) \\ \leftarrow \end{array} \rightsquigarrow \left(\begin{array}{ccc|c} 0 & -1 & 0 & -1/2 \\ 1 & 0 & 0 & 3/2 \\ 0 & 0 & -1 & -1/2 \end{array} \right)$$

Also ist $x = (3/2, 1/2, 1/2)^\top$ der Schnittpunkt von g und E .

Definition 7.7.13. Seien $x, y \in \mathbb{R}^3$ und $x = (x_1, x_2, x_3)^\top$, sowie $y = (y_1, y_2, y_3)^\top$. Dann heißt

$$x \times y = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \times \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} := \begin{pmatrix} x_2y_3 - y_2x_3 \\ y_1x_3 - x_1y_3 \\ x_1y_2 - y_1x_2 \end{pmatrix}$$

Kreuzprodukt oder *Vektorprodukt* oder *äußeres Produkt* von x und y .

Beispiel 7.7.14:

$$\begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} \times \begin{pmatrix} 0 \\ 3 \\ -2 \end{pmatrix} = \begin{pmatrix} -1 \cdot (-2) - 3 \cdot 2 \\ 0 \cdot 2 - 1 \cdot (-2) \\ 1 \cdot 3 - 0 \cdot (-1) \end{pmatrix} = \begin{pmatrix} -4 \\ 2 \\ 3 \end{pmatrix}$$

- Bemerkungen 7.7.15:** (i) Zwei Vektoren $x, y \in \mathbb{R}^3$ sind genau dann linear abhängig, wenn $x \times y = 0$ gilt.
- (ii) Sind $x, y \in \mathbb{R}^3$ linear unabhängig und ist α der Winkel zwischen x und y (dieser ist dann weder Null noch π), so ist $x \times y$ der eindeutig bestimmte Vektor mit
- $x \times y$ steht sowohl auf x als auch auf y senkrecht
 - $\|x \times y\| = \|x\| \cdot \|y\| \cdot \sin(\alpha)$
 - x, y und $x \times y$ bilden ein Rechtssystem.
- (iii) Sind $v, w \in \mathbb{R}^3$ Richtungsvektoren einer Ebene E , so ist $\frac{v \times w}{\|v \times w\|}$ ein Normaleneinheitsvektor von E .

Rechenregeln: Seien $u, v, w \in \mathbb{R}^3$ und $\lambda \in \mathbb{R}$. Dann gilt

1. $u \times v = -(v \times u)$ (Anti-Kommutativität)
2. $\lambda \cdot (u \times v) = (\lambda u) \times v = u \times (\lambda v)$ (Assoziativität)
3. $u \times (v + w) = u \times v + u \times w$ (Distributivität)
4. $u \times (v \times w) = \langle u, w \rangle v - \langle u, v \rangle w$.

Das Kreuzprodukt hat noch eine weitere anschauliche Anwendung.

Definition 7.7.16. Seien $u, v, w \in \mathbb{R}^3$. Die Zahl $\langle u, v \times w \rangle$ heißt *Spatprodukt* von u, v und w .

Bemerkungen 7.7.17: (i) Der Betrag des Spatproduktes entspricht genau dem Volumen des von den Vektoren u, v und w aufgespannten *Parallelotops*. Dabei ist das Vorzeichen positiv, wenn die Vektoren u, v, w ein Rechtssystem und negativ, wenn sie ein Linkssystem bilden.

(ii) **Rechenregeln:**

- $\langle u, v \times w \rangle = \langle w, u \times v \rangle = \langle v, w \times u \rangle$
- $\langle u + v, v \times w \rangle = \langle u, (v + u) \times w \rangle = \langle u, v \times (u + w) \rangle = \langle u, v \times w \rangle$

(iii) Das Spatprodukt von $u, v, w \in \mathbb{R}^3$ ist genau dann Null, wenn diese in einer Ebene liegen.

Beispiel 7.7.18: Das Spatprodukt der drei Standardbasisvektoren ist

$$\langle e_1, e_2 \times e_3 \rangle = \left\langle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \times \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\rangle = \left\langle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\rangle = 1,$$

in Übereinstimmung damit, dass das Volumen eines Würfels mit Seitenlänge Eins Eins ist.

Index

- Abbildung, lineare, 90
- Abbildungsmatrix, 91
- abgeschlossenes Intervall, 10
- Ableitung, 51
 - einseitige, 52
 - n -te, 59
 - zweite, 59
- absolut konvergente
 - Reihe, 43
 - uneigentliche Integrale, 80
- Absolutbetrag, 11
- abzählbar, 5
- Addition
 - in \mathbb{R}^n , 85
 - von Matrizen, 92
- Additionstheoreme, 25
- Äquivalenz, 2
- äußeres Produkt, 107
- algebraische Gleichung, 18
- Allquantor, 7
- alternierende harmonische Reihe, 43
- alternierende Reihe, 42
- Anordnung der reellen Zahlen, 9
- Arcuscosinus, 29
- Arcussinus, 29
- Arcustangens, 29
- Areacosinushyperbolicus, 28
- Argument, 34
 - einer Funktion, 23
- Aufpunkt, 108, 110
- Aussage, 1
- Aussageform, 2
- äußeres Produkt, 111

- Basis, 89
- Bernoullische Ungleichung, 10
- beschränkte
 - Folge, 37
 - Funktion, 29
- bestimmt divergente Folge, 38
- bestimmtes Integral, 69
- Betrag
 - einer komplexen Zahl, 33
 - einer reellen Zahl, 11
 - eines Vektors, 87
- Beweis
 - direkter, 5
 - durch Widerspruch, 6
 - indirekter, 6
- bijektiv, 27
- Bild einer Matrix, 96
- Bildmenge, 23
- Binomialkoeffizient, 8
- Binomische Formeln, 8

- Cauchy-Folge, 40
- Cauchy-Produkt, 42
- Cauchy-Schwarzsche Ungleichung, 87
- Cosinus, 24
- Cosinus hyperbolicus, 27
- Cosinussatz, 26
- Cotangens, 26
- Cramer-Regel, 107

- Defekt, 96
- Definitionsbereich, 23
 - maximaler, 23
- dekadischer Logarithmus, 15
- Determinante, 104
- Differentialquotient, 52
- Differenz von Mengen, 4
- Differenzenquotient, 52
- Differenzierbarkeit, 51
 - mehrfache, 59
- Dimension, 90
- Dimensionsregel, 97
- direkter Beweis, 5
- disjunkt, 4
- Disjunktion, 1
- Diskriminante, 19
- divergente
 - Folge, 37
 - Reihe, 41
 - uneigentliche Integrale, 80
- Dodekaeder, 31

- Drehung, 90
 Dreiecksungleichung, 11, 33, 87
 Durchschnitt, 4
 e , 17
 Ebene, 110
 Drei-Punkte-Form, 110
 Parameterdarstellung, 110
 Ebenengleichung, 110
 Hesse-Normalform, 110
 Einheitsmatrix, 94
 Einheitsvektor, 87
 einseitige Ableitung, 52
 Element, 3
 Entwicklungsstelle, 44
 ε -Umgebung, 37
 erweiterte Koeffizientenmatrix, 97
 Euklidnorm, 87
 Euler'sche Zahl, 17
 Euler'sches Theorem, 35
 Existenzquantor, 7
 Exponentialreihe, 46
 exponentielle Standardform, 13
 Extremum
 globales, 60
 lokales, 60
 strenges lokales, 60
 Fakultät, 8
 Feinheit einer Zerlegung, 69
 Fibonacci-Zahlen, 36
 Flächeninhalt, 70
 Folge, 36
 beschränkte, 37
 bestimmt divergente, 38
 Cauchy-, 40
 der Partialsummen, 41
 divergente, 37
 konvergente, 37
 monoton fallende, 37
 monoton wachsende, 37
 monotone, 37
 Null-, 37
 rekursiv definierte, 36
 streng monoton fallende, 37
 streng monoton wachsende, 37
 streng monotone, 37
 Folgenglied, 36
 Folgenindex, 36
 Formel von Hadamard, 45
 Funktion, 23
 Ableitung einer, 51
 beschränkte, 29
 bijektive, 27
 differenzierbare, 51
 gerade, 29
 hyperbolische, 27
 injektive, 27
 integrierbare, 69
 monoton fallende, 29
 monoton wachsende, 29
 monotone, 29
 n -mal differenzierbare, 59
 periodische, 29
 -sgraph, 23
 stetige, 47
 streng monoton fallende, 29
 streng monoton wachsende, 29
 streng monotone, 29
 surjektive, 27
 -swert, 23
 Umkehr-, 27
 ungerade, 29
 zweimal differenzierbare, 59
 ganze Zahlen, 4
 Gaußklammerfunktion, 38
 Gauß'sche Zahlenebene, 33
 geometrische Reihe, 41
 Gerade, 86, 108
 in Parameterdarstellung, 108
 Zwei-Punkte-Form, 86, 108
 gerade Funktion, 29
 gerade Permutation, 104
 Geradengleichung, 108
 Hesse-Normalform, 108
 Gleichung
 algebraische, 18
 Gleichungssystem, 103
 globales Extremum, 60
 Grad eines Polynoms, 19
 Graph, 23
 der Umkehrfunktion, 28
 Grenzwert, 37
 linksseitiger, 47
 rechtsseitiger, 47
 von Funktionen, 47
 Grenzwertsätze, 38
 Grenzwertsätze
 für Funktionen, 48
 Hadamard-Formel, 45

- Häufungswert einer Folge, 37
- halboffenes Intervall, 11
- harmonische Reihe, 41
- Häufungspunkt einer Menge, 47
- Hauptsatz der Algebra, 34
- Hauptsatz der Differential- und Integralrechnung, 71
- Heron-Verfahren, 36
- Hesse-Normalform, 108, 110
- hinreichend, 2
- höhere Ableitungen, 59
- homogenes lineares Gleichungssystem, 97
- Hôpital, Regel von de L', 58
- Hyperbolische Funktionen, 27
- i*, 32
- Ikosaeder, 31
- imaginäre Einheit, 32
- Imaginärteil, 32
- Implikation, 2
- Index einer Folge, 36
- indirekter Beweis, 6
- Induktionsanfang, 6
- Induktionsschritt, 7
- inhomogenes lineares Gleichungssystem, 97
- injektiv, 27
- inneres Produkt, 87
- Integral
 - bestimmtes, 69
 - unbestimmtes, 72
 - uneigentliches, 80, 82
- integrierbare Funktion, 69
- Interpolationspolynom, 83
- Intervall, 10
 - abgeschlossenes, 10
 - halboffenes, 11
 - offenes, 11
- Intervalllänge, 11
- inverse Matrix, 94
- invertierbare Matrix, 94
- irrationale Zahlen, 5
- Iterationsverfahren, Newtonsches, 65
- Kartesisches Produkt, 4
- Kern einer Matrix, 96
- Kettenregel, 54
- komplexe Zahlen, 32
- Konjugation, 33
- konjugiert komplexe Zahl, 33
- Konjunktion, 1
- Kontrapositionsgesetz, 2
- konvergente
 - Folge, 37
 - Reihe, 41
 - uneigentliche Integrale, 80
- Konvergenzradius, 45
- Koordinaten, 85, 86
- Kreuzprodukt, 111
- Kroneckersymbol, 94
- Lagrange-Restglied, 61
- Laplace'scher Entwicklungssatz, 105
- leere Menge, 3
- Leibniz-Kriterium, 42
- Limes, 37
- linear abhängig, 88
- linear unabhängig, 88
- lineare Abbildung, 90
- lineares Gleichungssystem, 97
 - homogenes, 97
 - inhomogenes, 97
- Linearkombination, 89
- linksseitiger Grenzwert, 47
- Lösungsmenge, 10
- Logarithmus
 - dekadischer, 15
 - natürlicher, 18
- lokales Extremum, 60
- Majorantenkriterium
 - für Reihen, 43
 - für uneigentliche Integrale, 81
- Mantisse, 13
- Matrix, 91
 - Abbildungs-, 91
 - inverse, 94
 - invertierbare, 94
 - quadratische, 91
 - reguläre, 94
 - transponierte, 96
- Matrizenprodukt, 92
- maximaler Definitionsbereich, 23
- Maximum
 - globales, 60
 - lokales, 60
 - strenges lokales, 60
- Menge, 3
 - leere, 3
- Mengendifferenz, 4
- Minimum
 - globales, 60
 - lokales, 60

- strenges lokales, 60
- Minorantenkriterium
 - für Reihen, 44
 - für uneigentliche Integrale, 81
- Mittelwertsatz, 56
 - der Integralrechnung, 71
- monoton fallende
 - Folge, 37
 - Funktion, 29
- monoton w achsende
 - Funktion, 29
- monoton wachsende
 - Folge, 37
- monotone
 - Folge, 37
 - Funktion, 29
- Multiplikation mit Skalaren
 - in \mathbb{R}^n , 85
 - von Matrizen, 92
- natürliche Zahlen, 4
- natürlicher Logarithmus, 18
- Negation, 1
- Newton-Cotes-Formeln, 83
- Newton-Verfahren, 65
- Norm, 87
- Normaleneinheitsvektor
 - einer Ebene, 110
 - einer Geraden, 108
- normalisierte Gleitpunktdarstellung, 13
- normierter Vektor, 87
- notwendig, 2
- n -te Wurzel, 10
- Nullfolge, 37
- Nullstelle, 19
- Nullvektor, 86
- numerische Integration, 83
- Obermenge, 3
- offenes Intervall, 11
- Oktaeder, 31
- orthogonale Vektoren, 87
- Ortsvektor, 86
- Parallelotop, 112
- Parameterdarstellung einer Ebene, 110
- Partialbruchzerlegung, 77
- Partialsomme, 41
- partielle Integration, 73
- Pascal'sches Dreieck, 9
- periodische Funktion, 29
- Permutation, 104
 - gerade, 104
 - ungerade, 104
- Platonische Körper, 30
- Polarkoordinaten, 34
- Polynom, 19
 - grad, 19
 - interpolierendes, 83
- Potenzreihe, 44
- pq -Formel, 20
- Produktregel, 54
- Produktzeichen, 8
- Projektion, 90
- Pythagoras
 - trigonometrischer, 25
- quadratische Ergänzung, 19
- quadratische Matrix, 91
- Quadratur, 83
- Quotientenkriterium, 43, 44
- Quotientenregel, 54
- Rang, 96
- rationale Zahlen, 4
- Realteil, 32
- rechtsseitiger Grenzwert, 47
- reelle Zahlen, 5
- Regel von de L'Hôpital, 58
- Regel von Sarrus, 105
- reguläre Matrix, 94
- Reihe, 41
 - absolut konvergente, 43
 - alternierende, 42
 - alternierende harmonische, 43
 - divergente, 41
 - Exponential-, 46
 - geometrische, 41
 - harmonische, 41
 - konvergente, 41
 - nwert, 41
 - Potenz-, 44
- rekursiv definierte Folge, 36
- Restglied von Lagrange, 61
- Richtungsvektor
 - einer Ebene, 110
 - einer Gerade, 108
- Riemannsumme, 69
- \mathbb{R}^n , 85
- Rolle, Satz von, 56
- Sarrus

- Regel von, 105
- Satz
 - Laplace'scher Entwicklungs-, 105
 - von Rolle, 56
 - von Taylor, 61
- Satz vom ausgeschlossenen Dritten, 1
- Satz vom ausgeschlossenen Widerspruch, 1
- Signum, 11
 - einer Permutation, 104
- Simpsonregel, 83
- Sinus, 24
- Sinus hyperbolicus, 27
- Sinussatz, 26
- Skalar, 85
- Skalarprodukt, 87
- Spat, 107
- Spatprodukt, 107, 112
- Spiegelung, 90
- Stammfunktion, 72
- Standardbasis, 89
- stetig ergänzbar, 48
- Stetigkeit, 47
- Strecke, 86
- streng monoton fallende
 - Folge, 37
 - Funktion, 29
- streng monoton wachsende
 - Folge, 37
 - Funktion, 29
- streng monotone
 - Folge, 37
 - Funktion, 29
- strenges lokales Extremum, 60
- Substitutionsregel, 73
- Summenzeichen, 7
- surjektiv, 27

- Tangens, 26
- Taylor, Satz von, 61
- Taylorpolynom, 61
- Taylorreihe, 62
- Teilmenge, 3
- Teilraum, 89
- Tetraeder, 31
- Tetraederwinkel, 30
- Transitivität, 9
- transponierte Matrix, 96
- Trapezregel, 83
- trigonometrischer Pythagoras, 25

- überabzählbar, 5

- Umkehrfunktion, 27
 - Ableitung der, 55
- unbestimmtes Integral, 72
- uneigentliches Integral, 80, 82
 - absolut konvergentes, 80
 - divergentes, 80
 - konvergentes, 80
- ungerade Funktion, 29
- ungerade Permutation, 104
- Ungleichung, 9
 - Bernoulli'sche, 10
 - Dreiecks-, 11, 33
- Unterraum, 89

- Vektor, 86
 - Addition, 85
 - Multiplikation mit Skalar, 85
 - normierter, 87
- Vektorprodukt, 107, 111
- Vereinigung, 4
- Verkettung, 48
- vollständige Induktion, 6

- Würfel, 31
- Wahrheitstafel, 2
- Wendepunkt, 60
- Wertebereich, 23
- Widerspruchsbeweis, 6
- Wurzel, 10
 - n -te, 10
- Wurzelkriterium, 43, 44

- Zahlen
 - ganze, 4
 - irrationale, 5
 - komplexe, 32
 - natürliche, 4
 - rationale, 4
 - reelle, 5
- Zahlengerade, 5
- Zerlegung, 68
 - Feinheit einer, 69
- Zwei-Punkte-Form einer Geraden, 86, 108
- zweite Ableitung, 59
- Zwischenwertsatz, 49