# Mathematik für ChemikerInnen I

Prof. Dr. Ansgar Jüngel
Institut für Mathematik
Johannes Gutenberg-Universität Mainz

Winter 2006

*unkorrigiertes Vorlesungsskript*

## Inhaltsverzeichnis

# 1 Motivation

In diesem Kapitel verdeutlichen wir anhand von Beispielen, daß die Sprache und die Methoden der Mathematik für das Studium chemischer Prozesse und Fragestellungen ein unentbehrliches Handwerkszeug darstellen.

**Beispiel 1.1** (Zerfall von Kaliumdichromat)
Wird Kaliumdichromat $K_2Cr_2O_7$ auf über 500° C erhitzt, zerfällt es in Kaliumchromat $K_2CrO_4$, Chromoxid $Cr_2O_3$ und Sauerstoff $O_2$. Die Reaktionsgleichung lautet

$$x_1\,\mathrm{K_2Cr_2O_7} \to x_2\,\mathrm{K_2CrO_4} + x_3\,\mathrm{Cr_2O_3} + x_4\,\mathrm{O_2}.$$

Wir suchen Zahlen $x_1, x_2, x_3, x_4$, die diese Gleichung erfüllen. Natürlich sollte die Lösung ganzzahlig (und positiv) sein. Die linke Seite benötigt $2 \cdot x_1$ Kaliumatome ($x_1\mathrm{K}_2$), die rechte Seite $2 \cdot x_2$ Kaliumatome ($x_2\mathrm{K}_2$), d.h. $2x_1 = 2x_2$. Analog gilt $2x_1 = x_2 + 2x_3$ für die Bilanz der Chromatome und $7x_1 = 4x_2 + 3x_3 + 2x_4$ für die der Sauerstoffatome. Die Zahlen $x_1, x_2, x_3, x_4$ erfüllen also das Gleichungssystem

$$\begin{aligned} 2x_1 - 2x_2 \qquad\qquad &= 0, \\ 2x_1 - x_2 - 2x_3 \qquad &= 0, \\ 7x_1 - 4x_2 - 3x_3 - 2x_4 &= 0. \end{aligned} \tag{1.1}$$

Man kann dieses System auch abkürzend durch Auflistung der Koeffizienten der linken Seite und der Komponenten der rechten Seite durch

$$\begin{pmatrix} 2 & -2 & 0 & 0 \\ 2 & -1 & -2 & 0 \\ 7 & -4 & -3 & -2 \end{pmatrix} \qquad \text{und} \qquad \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \tag{1.2}$$

formulieren. Wir nennen das erste Objekt *Matrix* und das zweite Objekt *Vektor*. In Kapitel 4 werden wir uns mit der Matrizen- und Vektorrechnung befassen und den Gaußschen Algorithmus zum Lösen linearer Gleichungssysteme kennenlernen. □

**Beispiel 1.2** (Radioaktiver Zerfall)
Eine radioaktive Substanz zerfalle im Laufe der Zeit. Wieviel Substanz ist zur Zeit $t$ vorhanden?

Sei $N(t)$ die Menge der radioaktiven Substanz zur Zeit $t$. Aus der Physik ist bekannt, daß die Änderung $\triangle N$ der Stoffmenge proportional ist zur Stoffmenge $N(t)$ und zur Zeitspanne $\triangle t$. Nennen wir die Proportionalitätskonstante $-\alpha < 0$ (sie ist negativ, da die Substanz zerfällt), so folgt

$$\triangle N = -\alpha N(t) \triangle t. \tag{1.3}$$

Schreiben wir $\triangle N = N(t + \triangle t) - N(t)$, erhalten wir

$$N(t + \triangle t) = (1 - \alpha \triangle t) N(t) \qquad \text{für alle } t.$$

Für $t = 0$ ist

$$N(\triangle t) = (1 - \alpha \triangle t) N(0),$$

für $t = \triangle t$

$$N(2\triangle t) = (1 - \alpha \triangle t) N(\triangle t) = (1 - \alpha \triangle t)^2 N(0),$$

und allgemein für $t = n\triangle t$ ($n \in \mathbb{N}$)

$$N(n\triangle t) = (1 - \alpha \triangle t)^n N(0) = \left(1 - \frac{\alpha t}{n}\right)^n N(0).$$

Was geschieht nun, wenn $n$ immer größer wird? Für größer werdendes $n$ nähert sich die Differenz $1 - \frac{\alpha t}{n}$ immer mehr der Zahl Eins an, und man könnte wegen $1^n = 1$ denken, daß $\left(1 - \frac{\alpha t}{n}\right)^n$ sich ebenfalls immer mehr der Eins nähert. Dies ist jedoch falsch! Es gilt

$$\left(1 - \frac{\alpha t}{n}\right)^n \quad \text{nähert sich für großes } n \quad e^{-\alpha t}, \tag{1.4}$$

wobei $e = 2.7182818\ldots$ die *Eulersche Zahl* ist. Mathematisch formulieren wir die Aussage (1.4) als

$$\left(1 - \frac{\alpha t}{n}\right)^n \quad \text{strebt für } n \to \infty \text{ gegen} \quad e^{-\alpha t}$$

oder in Zeichen

$$\lim_{n \to \infty} \left(1 - \frac{\alpha t}{n}\right)^n = e^{-\alpha t}.$$

Diesen *Grenzwert* ("limes") definieren wir in Kapitel 5. In Kapitel 3 untersuchen wir die Exponentialfunktion $e^x$. Die Stoffmenge zur Zeit $t$ lautet also

$$N(t) = e^{-\alpha t} N(0).$$

In diesem Beispiel sind verschiedene mathematische Begriffe aufgetreten, die wir in den folgenden Kapiteln genauer definieren und erläutern werden, nämlich *Grenzwerte* und *Funktionen.* Außerdem werden wir Funktionen *differenzieren* und *integrieren* (siehe Kapitel 6 und 7).

# 2 Grundbegriffe

## 2.1 Mengen

Die Gesamtheit von Objekten heißt *Menge*. Die einzelnen Objekte heißen *Elemente* der Menge. Beispiele von Mengen sind

$$\begin{aligned} M_1 &= \{1,2,3,4\} \\ M_2 &= \{\text{Kohlenstoff, Wasserstoff, Sauerstoff}\}, \\ M_1 &= \{5,6,7,\dots\} = \{x : x \text{ ist natürliche Zahl und } x \geq 5.\} \end{aligned}$$

Die Reihenfolge der Elemente spielt bei Mengen keine Rolle, d.h.

$$\{1,2,3,4\} = \{3,2,4,1\}.$$

Die *leere Menge* wird mit $\{\ \}$ oder $\emptyset$ bezeichnet.

Es gibt die folgenden Mengenoperationen.

**Definition 2.1** *Seien $A, B$ Mengen.*

*(1) $A$ heißt* Teilmenge *von $B$, kurz $A \subset B$, wenn für alle $x \in A$ auch $x \in B$ folgt.*

*(2)* Gleichheit*: $A = B$ bedeutet $A \subset B$ und $B \subset A$.*

*(3)* Vereinigung*: $A \cup B = \{x : x \in A \text{ oder } x \in B\}$.*

*(4)* Durchschnitt*: $A \cap B = \{x : x \in A \text{ und } x \in B\}$.*

*(5)* Differenz*: $A \backslash B = \{x : x \in A \text{ und } x \notin B\}$. Man sagt "A ohne B".*

*(6)* Kartesisches Produkt*: $A \times B = \{(x,y), x \in A, y \in B\}$. Beachten Sie, daß $(x,y)$ ein geordnetes Paar ist, d.h., die Reihenfolge ist wichtig. Beispielsweise gilt $(1,2) \neq (2,1)$. (Aber $\{1,2\} = \{2,1\}$!)*

**Beispiel 2.2** Seien $A = \{1,2,3\}$, $B = \{3,4\}$. Dann folgt

$$\begin{aligned} A \cap B &= \{3\}, \\ A \cup B &= \{1,2,3,4\} \qquad \text{(Elemente werden nur einmal aufgezählt)}, \\ A \backslash B &= \{1,2\}. \\ A \times B &= \{(1,3),(1,4),(2,3),(2,4),(3,3),(3,4)\}. \end{aligned}$$

□

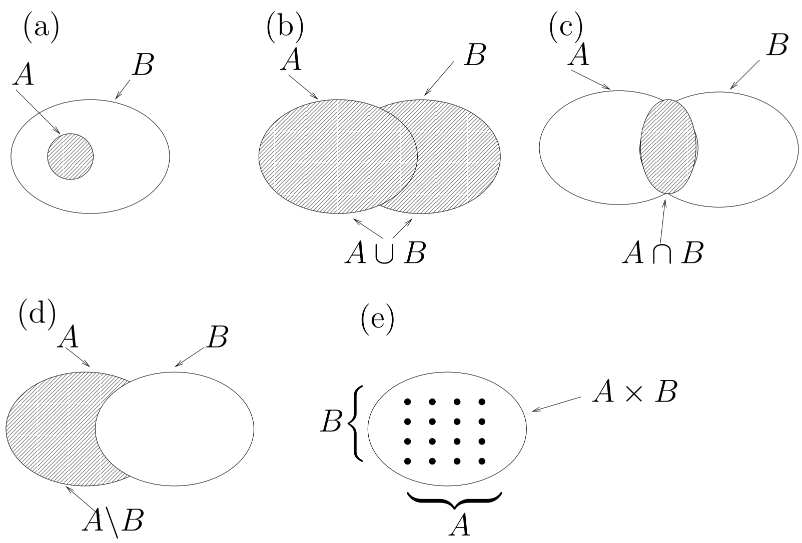


Abbildung 2.1: (a) Teilmenge; (b) Vereinigung; (c) Durchschnitt; (d) Differenz; (e) kartesisches Produkt.

## 2.2 Reelle Zahlen

Die wichtigsten Mengen sind Zahlenmengen.

**Definition 2.3** *(1) Menge der natürlichen Zahlen:*
$\mathbb{N} = \{1, 2, 3, \dots\}$.
*Die Menge der natürlichen Zahlen mit Null wird geschrieben als:*
$\mathbb{N}_0 = \mathbb{N} \cup \{0\} = \{0, 1, 2, 3, \dots\}$.

*(2) Menge der ganzen Zahlen:*
$\mathbb{Z} = \{0, 1, -1, 2, -2, 3, -3 \dots\}$.

*(3) Menge der rationalen Zahlen:*
$\mathbb{Q} = \{\frac{p}{q} : p, q \in \mathbb{Z}, q \neq 0\}$.
*Die Brüche $\frac{p}{0}$ und $\frac{0}{0}$ sind also* nicht *definiert.*

*(4) Menge der reellen Zahlen:*

$$\mathbb{R} = \{g+r : g \in \mathbb{Z}, r = 0, a_1a_2a_3 \dots = \frac{a_1}{10}+\frac{a_2}{100}+\frac{a_3}{1000}+\dots, a_1, a_2, a_3, \dots \in \{0, 1, \dots, 9\}\}.$$

*Reelle Zahlen sind also Zahlen mit unendlicher Dezimalentwicklung.*

**Beispiel 2.4** (1) Sei $x = 0,101010\ldots \in \mathbb{R}$. Wir wollen $x$ als Bruch schreiben. Dafür schreiben wir

$$\begin{aligned} 100 \cdot x &= 10,101010\ldots, \\ x &= 0,101010\ldots \end{aligned}$$

Die Differenz lautet

$$99 \cdot x = 10, \qquad \text{also} \quad x = \frac{10}{99} \in \mathbb{Q}.$$

(2) Es gibt auch Zahlen, die nicht als Bruch geschrieben werden können. Derartige Zahlen werden *irrational* genannt. Beispiele sind $\pi$ und $e$ (die Eulersche Zahl). Eine reelle Zahl ist also entweder rational oder irrational. $\square$

Es gelten die folgenden Inklusionen: $\mathbb{N} \subset \mathbb{N}_0 \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}$.

**Definition 2.5** *Seien $a, b \in \mathbb{R}$. Die folgenden Mengen werden* Intervalle *genannt:*

*(1) Abgeschlossenes Intervall:* $[a, b] = \{x \in \mathbb{R} : a \leq x \leq b\}$.

*(2) Rechts halboffenes Intervall:* $[a, b) = \{x \in \mathbb{R} : a \leq x < b\}$.

*(3) Links halboffenes Intervall:* $(a, b] = \{x \in \mathbb{R} : a < x \leq b\}$.

*(4) Offenes Intervall:* $(a, b) = \{x \in \mathbb{R} : a < x < b\}$.

*Man schreibt auch*

$$\begin{aligned}
[a, \infty) &= \{x \in \mathbb{R} : x \geq a\}, \\
(-\infty, b) &= \{x \in \mathbb{R} : x < b\}, \\
(-\infty, \infty) &= \mathbb{R}.
\end{aligned}$$

*Definiert man den* Betrag $|x|$ *einer reellen Zahl durch*

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

*so ist*

$$[-a, a] = \{x \in \mathbb{R} : -a \leq x \leq a\} = \{x \in \mathbb{R} : |x| \leq a\}.$$

Wir benötigen noch folgende *Notationen:*

- Das Zeichen “$\forall$” bedeutet “für alle”.
- Das Zeichen “$\exists$” bedeutet “es existiert”.
- Summe:
$$\sum_{k=1}^{n} a_k = a_1 + a_2 + \cdots + a_n.$$
- Produkt:
$$\prod_{k=1}^{n} a_k = a_1 \cdot a_2 \cdot \ldots \cdot a_n = a_1 a_2 \ldots a_n.$$

**Beispiel 2.6** Die Schreibweise

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

bedeutet ausgeschrieben:

$$\text{Für alle } n \in \mathbb{N} \text{ gilt: } 1+2+3+\cdots+n = \frac{n(n+1)}{2}.$$

Daß die Summe der ersten $n$ natürlichen Zahlen gleich $n(n+1)/2$ ist, können wir übrigens einfach einsehen, indem wir die beiden Zeilen

$$\begin{aligned} a &= 1 + \quad 2 \quad + \quad 3 \quad + \cdots + n, \\ a &= n + (n-1) + (n-2) + \cdots + 1 \end{aligned}$$

addieren, denn dies ergibt

$$2a = (n+1) + (n+1) + (n+1) + \cdots + (n+1) = n \cdot (n+1),$$

also $a = n(n+1)/2$. □

## 2.3 Binomialkoeffizienten

In diesem Abschnitt beantworten wir die Frage: Wieviel Möglichkeiten gibt es, $k$ Elemente von insgesamt $n$ Elementen als Mengen zu schreiben? Sei etwa $n = 5$ und $k = 2$. Die Elemente seien die Ziffern $1, 2, 3, 4, 5$. Dann gibt es 20 Möglichkeiten:

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 12 | 13 | 14 | 15 | 23 | 24 | 25 | 34 | 35 | 45 |
| 21 | 31 | 41 | 51 | 32 | 42 | 52 | 43 | 53 | 54 |

Allgemein: Wir haben $n$ Möglichkeiten, ein 1. Element zu wählen. Greifen wir das 1. Element heraus, verbleiben $n-1$ Möglichkeiten, ein 2. Element auszuwählen, $n-2$ Möglichkeiten für das 3. Element usw.:

$$\text{Anzahl:} = n \cdot (n-1) \cdot (n-2) \cdot \ldots \cdot (n-(k-1)).$$

Definieren wir die *Fakultät:*

$$n! = n(n-1)(n-2) \cdot \ldots \cdot 3 \cdot 2 \cdot 1 \qquad \text{für } n \in \mathbb{N} \quad \text{und} \quad 0! = 1,$$

so können wir schreiben:

$$\begin{aligned} \text{Anzahl} &= \frac{n(n-1)(n-2) \cdot \ldots \cdot (n-(k-1))(n-k)(n-(k+1)) \cdot \ldots \cdot 3 \cdot 2 \cdot 1}{(n-k)(n-(k+1)) \cdots 3 \cdot 2 \cdot 1} \\ &= \frac{n!}{(n-k)!}. \end{aligned}$$

Nun wird die Reihenfolge der Elemente in einer Menge nicht unterschieden, d.h. $\{1,2\} = \{2,1\}$. Die obigen 20 Möglichkeiten reduzieren sich auf die 10 Möglichkeiten

$$\{1,2\}, \{1,3\}, \{1,4\}, \{1,5\}, \{2,3\}, \{2,4\}, \{2,5\}, \{3,4\}, \{3,5\}, \{4,5\}.$$

Allgemein: Die $k!$ Möglichkeiten, $k$ Elemente in einer Liste zu schreiben, müssen herausdividiert werden. Wir erhalten:

$$\text{Anzahl} = \frac{n!}{(n-k)!k!}.$$

Man nennt diesen Quotienten *Binomialkoeffizienten* und definiert

$$\binom{n}{k} := \frac{n!}{(n-k)!k!} \qquad \text{für } n, k \in \mathbb{N}_0, \quad n \geq k.$$

In unserem Beispiel gilt etwa

$$\binom{5}{2} = \frac{5!}{3!2!} = \frac{5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}{3 \cdot 2 \cdot 1 \cdot 2 \cdot 1} = 10.$$

Die Binomialkoeffizienten können übrigens rekursiv über das sogenannte *Pascalsche Dreieck* berechnet werden (siehe Abb. 2.2).

| | | k | | | | | |
|---|---|---|---|---|---|---|---|
| | | 0 | 1 | 2 | 3 | 4 | 5 |
| n | 0 | 1 | | | | | |
| | 1 | 1 | 1 | | | | |
| | 2 | 1 | 2 | 1 | | | |
| | 3 | 1 | 3 | 3 | 1 | | |
| | 4 | 1 | 4 | 6 | 4 | 1 | |
| | 5 | 1 | 5 | 10 | 10 | 5 | 1 |

Abbildung 2.2: Pascalsches Dreieck.

Die Zahlen in der Tabelle erhält man durch Addition der direkt und links darüber liegenden Zahlen. Beispielsweise ist

$$\binom{5}{3} = 10 = 4 + 6.$$

Interessanterweise sind die Binomialkoeffizienten gerade die Koeffizienten nach Ausmultiplizieren des Produktes $(a+b)^n$:

$$\begin{aligned}(a+b)^1 &= 1\cdot a + 1\cdot b,\\ (a+b)^2 &= 1\cdot a^2 + 2\cdot ab + 1\cdot b^2,\\ (a+b)^3 &= 1\cdot a^3 + 3\cdot a^2b + 3\cdot ab^2 + 1\cdot b^3, \qquad \text{usw.}\end{aligned}$$

Dies motiviert das folgende Resultat.

**Satz 2.7** (Binomialsatz)
*Für $a, b \in \mathbb{R}$ und $n \in \mathbb{N}_0$ gilt:*

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k = \binom{n}{0} a^n + \binom{n}{1} a^{n-1} b + \cdots + \binom{n}{n-1} a b^{n-1} + \binom{n}{n} b^n.$$

## 2.4 Komplexe Zahlen

Die Gleichung

$$x^2 + 1 = 0$$

besitzt keine reelle Lösung. Um diesem Umstand Abhilfe zu schaffen, erweitert man den Zahlenbegriff. Wir definieren eine neue Zahl $i$ mit der Eigenschaft $i^2 = -1$ (auch geschrieben als $i = \sqrt{-1}$).

**Definition 2.8** *Die Zahl $i$ mit $i^2 = -1$ heißt* komplexe Einheit. *Die* Menge der komplexen Zahlen *ist definiert durch*

$$\mathbb{C} = \{z : z = a + ib,\ a, b \in \mathbb{R}\}.$$

*Wir definieren weiter für eine Zahl $z = a + ib \in \mathbb{C}$:*

- *Realteil von $z$:* $\mathrm{Re}\,(z) = a$,
- *Imaginärteil von $z$:* $\mathrm{Im}\,(z) = b$,
- *konjugiert komplexe Zahl von $z$:* $\overline{z} = a - ib$,
- *Betrag von $z$:* $|z| = \sqrt{a^2 + b^2}$.

Diese Definitionen können mit Hilfe der *Gaußschen Zahlenebene* veranschaulicht werden (Abbbildung 2.3). Die komplexe Zahl $z$ wird als ein Ortsvektor dargestellt. Die konjugiert komplexe Zahl $\overline{z}$ ist der an der $x$-Achse (oder Re-Achse) gespiegelte Vektor; der Betrag $|z|$ ist die Länge des Vektors $z$.

Benutzt man die Beziehung $i^2 = -1$, so kann man mit komplexen Zahlen so rechnen, wie man es von den reellen Zahlen gewohnt ist.

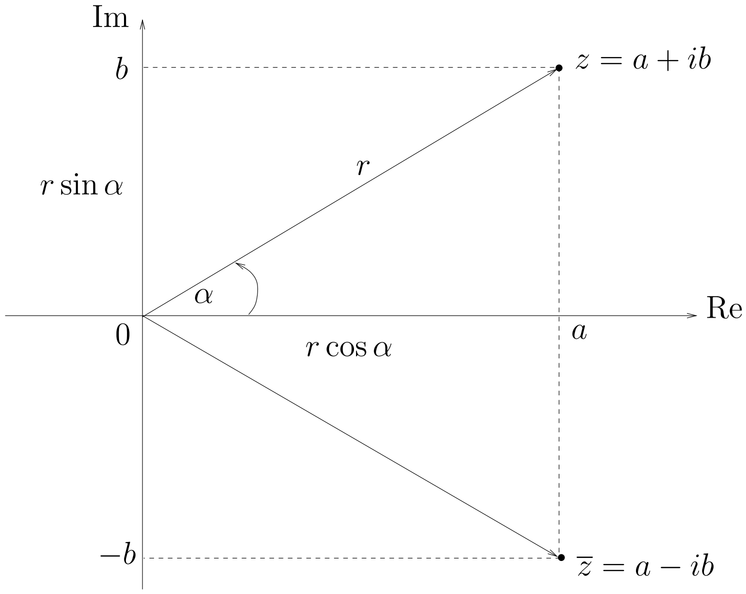


Abbildung 2.3: Gaußsche Zahlenebene.

**Beispiel 2.9** Seien $z_1 = 2 + 3i$, $z_2 = 1 - 2i \in \mathbb{C}$. Dann ist

- Addition: $z_1 + z_2 = 2 + 3i + 1 - 2i = 3 + i$,
- Subtraktion: $z_1 - z_2 = 2 + 3i - 1 + 2i = 1 + 5i$,
- Multiplikation: $z_1 \cdot z_2 = (2 + 3i)(1 - 2i) = 2 + 3i - 4i - 6i^2 = 8 - i$,
- Division: $\dfrac{z_1}{z_2} = \dfrac{2+3i}{1-2i} = \dfrac{(2+3i)(1+2i)}{(1-2i)(1+2i)} = \dfrac{2+3i+4i+6i^2}{1-4i^2} = \dfrac{-4+7i}{1+4} = -\dfrac{4}{5}+\dfrac{7}{5}i.$

□

So wie man zweidimensionale Vektoren durch Angabe der Länge und des Winkels in Polarkoordinaten darstellen kann, ist dies auch für komplexe Zahlen möglich (Abbildung 2.3). Man schreibt für $z = a + ib \in \mathbb{C}$

$$a = r\cos\alpha, \qquad b = r\sin\alpha,$$

wobei $r = |z| = \sqrt{a^2 + b^2}$, also

$$z = a + ib = r(\cos\alpha + i\sin\alpha), \qquad r > 0, \quad \alpha \in \mathbb{R}.$$

Eine dritte Darstellungsform komplexer Zahlen erhält man durch die *Eulersche Formel*

$$e^{i\alpha} := \cos\alpha + i\sin\alpha.$$

Dann können wir für $z \in \mathbb{C}$ schreiben

$$z = r(\cos\alpha + i\sin\alpha) = re^{i\alpha}.$$

Der Vorteil dieser Formulierung ist, daß komplexe Zahlen einfach potenziert werden können:

$$z^n = (re^{i\alpha})^n = r^n e^{i\alpha n} = r^n\Big(\cos(n\alpha) + i\sin(n\alpha)\Big).$$

Die drei Darstellungsformen sind in Abbildung 2.4 zusammengefaßt.

| Komplexe Zahl | Kartesische Form | Polarkoordinaten | Eulersche Form |
|---|---|---|---|
| $z$ | $a + ib$ | $r(\cos\alpha + i\sin\alpha)$ | $re^{i\alpha}$ |

Abbildung 2.4: Darstellung komplexer Zahlen.

Die trigonometrischen Funktionen sind periodisch; daher ist

$$1 = e^{i\cdot 0} = \cos 0 + i\sin 0 = \cos 2\pi + i\sin 2\pi = e^{2\pi i}. \tag{2.1}$$

Diese Beobachtung motiviert das *Wurzelziehen* komplexer Zahlen, denn

$$w_1 = \sqrt{z} = (re^{i\alpha})^{1/2} = \sqrt{r}e^{i\alpha/2}.$$

Allerdings gibt es noch eine zweite Wurzel $w_2 = \sqrt{r}e^{i(\alpha+2\pi)/2}$, da aus (2.1) folgt

$$w_2^2 = \left(\sqrt{r}e^{i(\alpha+2\pi)/2}\right)^2 = re^{i(\alpha+2\pi)} = re^{i\alpha}e^{2\pi i} = re^{i\alpha}.$$

Allgemein gilt:

**Satz 2.10** *Sei $z = re^{i\alpha} \in \mathbb{C}$, $z \neq 0$. Dann existieren genau $n$ verschiedene komplexe Zahlen $w_0, \ldots, w_{n-1} \in \mathbb{C}$, die die Gleichung $w^n = z$ lösen. Sie lauten*

$$w_k = \sqrt[n]{r}e^{i(\alpha+2\pi k)/n}, \qquad k = 0, \ldots, n-1.$$

**Beispiel 2.11** (1) Sei $z = -1 = e^{-i\pi}$. Dann lauten die beiden Quadratwurzeln

$$w_1 = e^{-i\pi/2} = -i \qquad \text{und} \qquad w_2 = e^{i(-\pi+2\pi)/2} = e^{i\pi/2} = i.$$

Die Bezeichnung $\sqrt{-1} = \sqrt{z} = i$ ist also etwas mißverständlich, da es ja zwei Wurzeln gibt. In beiden Fällen gilt $i^2 = w_1^2 = -1$, $(-i)^2 = w_2^2 = -1$.

(2) Wir suchen alle komplexen Wurzeln von $z^3 = 1 = e^{i\cdot 0}$. Sie lauten nach Satz 2.10 (siehe Abbildung 2.5)

$$w_1 = e^{0\cdot\pi i/3} = 1, \qquad w_2 = e^{2\pi i/3}, \qquad w_3 = e^{4\pi i/3}.$$

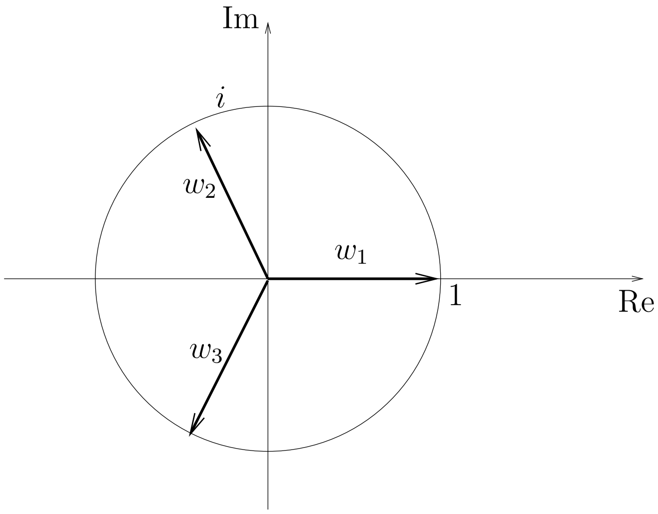


Abbildung 2.5: Die dritten Einheitswurzeln.

Wofür braucht man komplexe Zahlen? Wir geben zwei Antworten. Zum einen erlauben komplexe Zahlen die Lösung von Gleichungen der Form

$$f(x) := a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 = 0. \tag{2.2}$$

**Satz 2.12** (Fundamentalsatz der Algebra)
*Seien $n \in \mathbb{N}$ und $a_0, \ldots, a_n \in \mathbb{C}$ mit $a_n \neq 0$. Dann existieren $k \in \mathbb{N}$, $n_1, \ldots, n_k \in \mathbb{N}$ mit $n_1 + \cdots + n_k = n$ und paarweise verschiedene Zahlen $x_1, \ldots, x_k \in \mathbb{C}$, so daß*

$$f(x) = (x - x_1)^{n_1}(x - x_2)^{n_2} \cdot \ldots \cdot (x - x_k)^{n_k}.$$

*Die Zahlen $x_1, \ldots, x_k$ sind also Lösungen von (2.2). Die Zahl $n_j$ heißt* Vielfachheit *der Lösung $x_j$ $(j = 1, \ldots, k)$.*

Die zweite Antwort geben wir im nächsten Beispiel.

**Beispiel 2.13** (Freies Elektron)
Der Zustand eines freien Elektrons wird durch die sogenannte *Wellenfunktion* $\psi(x) \in \mathbb{C}$ beschrieben. In der Darstellung $\psi(x) = re^{i\alpha}$ wird das Elektron als eine Welle mit Amplitude $r$ und Phase $\alpha$ interpretiert. Außerdem interpretiert man das Betragsquadrat $|\psi(x)|^2$ als die Wahrscheinlichkeitsdichte, das Elektron "in der Nähe" des Ortes $x$ zu finden.

Ein freies Teilchen im Raum mit maximaler Aufenthaltswahrscheinlichkeit an der Stelle $x = 0$ wird z.B. dargestellt durch den stationären Zustand

$$\psi(x) = e^{ikx}e^{-x^2/2}.$$

Hierbei ist $k$ die Wellenzahl des Teilchens, die ein Maß für dessen Energie ist. Wegen $|\psi(x)| = |e^{ikx}| \cdot |e^{-x^2/2}| = e^{-x^2/2}$ ist tatsächlich die Wahrscheinlichkeitsdichte an $x = 0$ maximal. Für $x \to \pm\infty$ wird $|\psi(x)|$ immer kleiner; dies bedeutet, daß die Wahrscheinlichkeit, das Teilchen weit weg von $x = 0$ zu finden, sehr klein ist. $\square$

# 3 Funktionen

Eine *Abbildung* ist eine Vorschrift, die jedem Element $x$ aus einer (nichtleeren) Menge $M$ genau ein Element $f(x)$ aus einer Menge $N$ zuordnet; kurz:

$$f : M \to N, \qquad x \mapsto f(x).$$

Die Menge $M$ heißt *Definitionsbereich* von $f$, $N$ heißt *Wertebereich* (Abbildung 3.1). Wir schreiben meist $D_f$ für den Definitionsbereich von $f$.

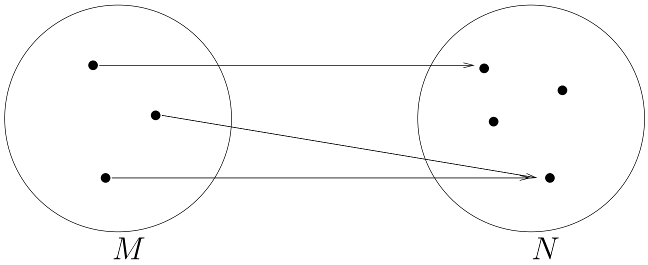

Abbildung 3.1: Zur Definition einer Abbildung.

Ist $M = \mathbb{R}^n$ und $N = \mathbb{R}$, so nennen wir eine Abbildung eine *Funktion*. In diesem Kapitel untersuchen wir einige elementare Funktionen wie Polynome, exp, ln, sin, cos usw.

## 3.1 Funktionen und ihre Inverse

Wir definieren zunächst die Inverse einer Funktion.

**Definition 3.1** *Sei $f : D_f \to \mathbb{R}$ eine Funktion. Man nennt die Funktion $g : D_g \to \mathbb{R}$ mit der Eigenschaft*

$$\begin{aligned} g(f(x)) &= x \qquad \textit{für alle } x \in D_f \textit{ bzw.} \\ f(g(y)) &= y \qquad \textit{für alle } y \in D_g \end{aligned}$$

*die* Inverse *von $f$ und schreibt $g = f^{-1}$.*

**Beispiel 3.2** Sei $f(x) = x^2 + 1$, $x \in D_f = \mathbb{R}$. Wir setzen $y = x^2 + 1$ und lösen nach $x$ auf:

$$x = \pm\sqrt{y-1}.$$

Setzen wir $g(y) = \pm\sqrt{y-1}$, so ist $g$ ein Kandidat für die Inverse von $f$ (siehe Abbildung 3.2), denn

$$f(g(y)) = (\pm\sqrt{y-1})^2 + 1 = (y-1) + 1 = y.$$

Es stellen sich jedoch zwei Fragen:

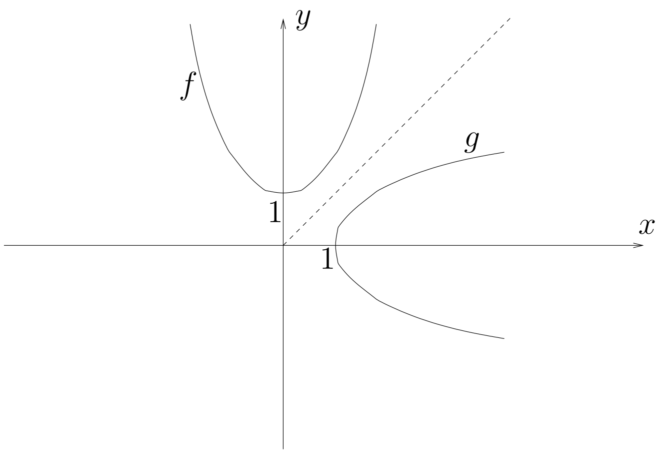


Abbildung 3.2: Die Funktion $f(x) = x^2+1$ und deren Spiegelung an der Diagonalen $x = y$.

- Wie lautet der Definitionsbereich von $g$?
- Warum gibt es zwei Funktionen $g(y) = \sqrt{y-1}$ und $g(y) = -\sqrt{y-1}$?

Die erste Frage kann leicht beantwortet werden: Damit es eine reelle Wurzel gibt, sollte $y \geq 1$ sein, also $D_g = [1, \infty)$. Dies ist übrigens gerade das *Bild* von $f$:

$$f(\mathbb{R}) = \{f(x) : x \in D_f\} = [1, \infty) = D_g.$$

Die Antwort der zweiten Frage basiert auf der Feststellung, daß zu jedem $x \neq 0$ die beiden Werte $f(x)$ und $f(-x)$ übereinstimmen. Daher gibt es zu jedem $y > 1$ zwei Werte für die "Inverse". Damit die Inverse eindeutig bestimmt ist, sollte zu jedem $y \in f(\mathbb{R})$ *genau ein* $x \in D_f$ existieren, so daß $f(x) = y$. Eine Funktion $f : D_f \to f(\mathbb{R})$ (beachte den Wertebereich!) mit dieser Eigenschaft heißt *bijektiv*. □

**Definition 3.3** *Sei $f : D_f \to W_f$ eine Funktion. Dann heißt $f$* bijektiv *genau dann, wenn für alle $y \in W_f$* genau ein *$x \in D_f$ existiert mit $f(x) = y$.*

Bijektive Funktionen sind invertierbar. Beispielsweise ist für $f : [0, \infty) \to [1, \infty)$, $f(x) = x^2 + 1$, die Inverse gegeben durch $f^{-1} : [1, \infty) \to [0, \infty)$, $f^{-1}(y) = \sqrt{y-1}$. Hätten wir $f : (-\infty, 0] \to [1, \infty)$, $f(x) = x^2 + 1$, definiert, so würde die Inverse $f^{-1} : [1, \infty) \to (-\infty, 0]$, $f^{-1}(y) = -\sqrt{y-1}$, lauten.

## 3.2 Rationale Funktionen

**Definition 3.4** *Sei $n \in \mathbb{N}$ und $a_0, \ldots, a_n \in \mathbb{R}$, $a_n \neq 0$. Dann heißt die Funktion $p : \mathbb{R} \to \mathbb{R}$, definiert durch*

$$p(x) = \sum_{k=0}^{n} a_k x^k = a_0 + a_1 x + \cdots + a_n x^n,$$

Polynom *vom Grade* $n$.

**Definition 3.5** *Seien* $p_m(x)$ *bzw.* $q_n(x)$ *Polynome vom Grade* $m$ *bzw.* $n$*. Dann heißt*

$$R(x) = \frac{p_m(x)}{q_n(x)}$$

rationale Funktion. *Sie ist für alle* $x \in \mathbb{R}$ *definiert, die keine Nullstellen von* $q_n$ *sind. Ist* $m < n$*, so heißt* $R$ echt gebrochene *rationale Funktion, anderenfalls heißt sie* unecht gebrochen.

**Beispiel 3.6** (1) Die Funktion

$$R(x) = \frac{x^2 + 2x - 3}{x + 3} = \frac{(x-1)(x+3)}{x+3} = x - 1$$

ist eigentlich keine rationale Funktion, sondern ein Polynom vom Grade 1 und damit definiert für alle $x \in \mathbb{R}$.

(2) Die Funktion

$$R(x) = \frac{x^2 + 2x - 3}{x + 4}$$

ist eine unecht gebrochene rationale Funktion, definiert für alle $x \in \mathbb{R}$, $x \neq -4$. Sie kann allerdings mit Hilfe der *Polynomdivison* geschrieben werden als die Summe eines Polynoms und einer echt gebrochenen rationalen Funktion:

$$\begin{array}{l}
\quad(x^2 + 2x - 3) : (x + 4) = x - 2 + \dfrac{5}{x+4} \\
-\underline{(x^2 + 4x)} \\
\qquad\quad -2x - 3 \\
\qquad -\underline{(-2x - 8)} \\
\qquad\qquad\quad 5
\end{array}$$

Wir erhalten:

$$R(x) = x - 2 + \frac{5}{x+4}. \qquad \square$$

**Beispiel 3.7** Für ideale Gase gilt die Zustandsgleichung

$$p \cdot V = R \cdot T,$$

wobei $p$ der Druck, $V$ das Volumen, $T$ die Temperatur des Gases und $R$ die sogenannte Gaskonstante sind. Die Aussage dieser Gleichung ist, daß das Produkt von Druck und Volumen eines Gases bei gleich gehaltener Temperatur konstant ist. Dies gilt nur annähernd. Eine bessere Approximation liefert die *Van-der-Waals-Gleichung*

$$\left(p + \frac{a}{V^2}\right)(V - b) = RT,$$

wobei $a$ die intermolekularen Kräfte und $b$ die Größe der Gasmoleküle modelliert. Bei konstanter Temperatur ist der Druck eine Funktion des Volumens:

$$p(V) = p = \frac{RT}{V-b} - \frac{a}{V^2} = \frac{RTV^2 - aV + ab}{V^2(V-b)}.$$

Der Druck $p(V)$ ist also eine echt gebrochene rationale Funktion. □

In Kapitel 7 benötigen wir eine Formulierung rationaler Funktion als Summe einfacher rationaler Funktionen, etwa von der Art

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

Dies führt auf das Konzept der *Partialbruchzerlegung*.

**Beispiel 3.8** (1) Sei

$$R(x) = \frac{x^2+1}{x^3-2x^2-x+2}.$$

Wir suchen zuerst alle Nullstellen des Nenners. Wir raten die Nullstelle $x_1 = 1$. Mit Polynomdivision erhalten wir:

$$\begin{array}{l} \quad (x^3 - 2x^2 - x + 2) : (x-1) = x^2 - x - 2 \\ -\underline{(x^3 - x^2)} \\ \qquad\quad -x^2 - x \\ \qquad -\underline{(-x^2 + x)} \\ \qquad\qquad\quad -2x + 2 \\ \qquad\qquad -\underline{(-2x + 2)} \\ \qquad\qquad\qquad\qquad 0 \end{array}$$

Die Nullstellen von $x^2 - x - 2$ ergeben sich aus der $p$-$q$-Formel:

$$x_{2/3} = \frac{1}{2} \pm \sqrt{\frac{1}{4} + 2} = \frac{1}{2} \pm \frac{3}{2}, \qquad \text{also} \quad x_2 = 2,\ x_3 = -1.$$

Daher ist

$$R(x) = \frac{x^2+1}{(x+1)(x-2)(x-1)}, \qquad x \in \mathbb{R}\backslash\{1, 2, -1\}.$$

Wir machen nun den Ansatz

$$R(x) = \frac{A}{x+1} + \frac{B}{x-2} + \frac{C}{x-1} \tag{3.1}$$

und wollen die Konstanten $A$, $B$ und $C$ bestimmen. Wir multiplizieren (3.1) mit dem Nenner:

$$x^2 + 1 = A(x-2)(x-1) + B(x+1)(x-1) + C(x+1)(x-2).$$

Eigentlich gilt diese Gleichung nur für $x \notin \{1, 2, -1\}$. Es ist dennoch möglich, diese Werte einzusetzen (dies folgt aus der Stetigkeit, die wir erst in Kapitel 6 einführen):

$$\begin{array}{llll} x = 1: & 2 = -2C & \Longrightarrow & C = -1, \\ x = 2: & 5 = 3B & \Longrightarrow & B = \frac{5}{3}, \\ x = -1: & 2 = 6A & \Longrightarrow & A = \frac{1}{3}. \end{array}$$

Wir folgern aus (3.1):

$$R(x) = \frac{1}{3(x+1)} + \frac{5}{3(x-2)} - \frac{1}{x-1}.$$

(2) Sei

$$R(x) = \frac{x^3}{x^3 + x^2 + x + 1}.$$

Diese Funktion ist unecht gebrochen rational, so daß wir zuerst Polynomdivision durchführen:

$$\begin{array}{l} x^3 : (x^3 + x^2 + x + 1) = 1 - \dfrac{x^2 + x + 1}{x^3 + x^2 + x + 1}. \\ -\underline{(x^3 + x^2 + x + 1)} \\ \qquad -x^2 - x - 1 \end{array}$$

Eine Nullstelle erraten wir: $x_1 = -1$. Eine weitere Polynomdivision führt auf

$$\begin{array}{l} (x^3 + x^2 + x + 1) : (x + 1) = x^2 + 1 \\ -\underline{(x^3 + x^2)} \\ \qquad 0 + x + 1 \\ \qquad -\underline{(x + 1)} \\ \qquad\qquad 0 \end{array}$$

Es gibt keine weiteren Nullstellen. Wir erweitern daher unseren Ansatz der Partialbruchzerlegung zu

$$\frac{x^2 + x + 1}{x^3 + x^2 + x + 1} = \frac{A}{x+1} + \frac{Bx + C}{x^2 + 1}.$$

Wir multiplizieren mit dem Nenner

$$x^2 + x + 1 = A(x^2 + 1) + (Bx + C)(x + 1)$$

und setzen spezielle Werte ein:

$$\begin{array}{rl} x = -1: & 1 = 2A \Longrightarrow A = \dfrac{1}{2}, \\ x = 0: & 1 = A + C \Longrightarrow C = 1 - A = \dfrac{1}{2}, \\ x = 1: & 3 = 2A + 2(B + C) \Longrightarrow B = \dfrac{3}{2} - A - C = \dfrac{1}{2}. \end{array}$$

Damit erhalten wir die Zerlegung

$$R(x) = 1 - \frac{x^2 + x + 1}{x^3 + x^2 + x + 1} = 1 - \frac{1}{2(x+1)} - \frac{x+1}{2(x^2+1)}. \qquad \square$$

## 3.3 Elementare Funktionen

**Exponentialfunktionen.** Betrachte für $a \in (1, \infty)$ die Funktion

$$f(x) = a^x, \qquad x \in \mathbb{R}.$$

Sie hat die Eigenschaften

$$a^0 = 1, \qquad a^1 = a, \qquad a^{x_1 + x_2} = a^{x_1} \cdot a^{x_2}.$$

Die Inverse von $f$ ist der *Logarithmus* zur Basis $a$:

$$f^{-1} : (0, \infty) \to \mathbb{R}, \qquad f^{-1}(x) = \log_a x.$$

Die Inverse hat die Eigenschaften

$$\log_a 1 = 0, \qquad \log_a a = 1, \qquad \log_a (x_1 \cdot x_2) = \log_a x_1 + \log_a x_2.$$

Im Falle $a = e$ (Eulersche Zahl) nennt man $\log_e = \ln$ den *natürlichen Logarithmus*; im Falle $a = 10$ wird $\log_{10} = \log$ *dekadischer Logarithmus* genannt. Falls $a = e$, heißt $f(x) = e^x$ *Exponentialfunktion.*

**Trigonometrische Funktionen.** Sinus und Cosinus sind Funktionen

$$\sin : \mathbb{R} \to \mathbb{R}, \qquad \cos : \mathbb{R} \to \mathbb{R}$$

mit den Eigenschaften

$$\begin{aligned} \cos^2 x + \sin^2 x &= 1, \quad x \in \mathbb{R}, \\ \sin(\pi k) &= 0, & \sin\left(\tfrac{\pi}{2}(2k+1)\right) &= (-1)^k, \\ \cos(\pi k) &= (-1)^k, & \cos\left(\tfrac{\pi}{2}(2k+1)\right) &= 0, \quad k \in \mathbb{N}_0. \end{aligned}$$

Insbesondere sind sin und cos $2\pi$-*periodisch*, d.h.

$$\sin(x + 2\pi) = \sin x, \qquad \cos(x + 2\pi) = \cos x, \qquad x \in \mathbb{R}.$$

Daher sind sin und cos auf $\mathbb{R}$ *nicht* invertierbar, sehr wohl aber die eingeschränkten Funktionen

$$\sin : \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \to [-1, 1], \qquad \cos : [0, \pi] \to [-1, 1].$$

Die Inversen heißen *Arcussinus* und *Arcuscosinus*:

$$\arcsin : [-1, 1] \to \left[-\frac{\pi}{2}, \frac{\pi}{2}\right], \qquad \arccos : [-1, 1] \to [0, \pi].$$

Der *Tangens*

$$\tan : \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \to \mathbb{R}, \qquad \tan x = \frac{\sin x}{\cos x},$$

ist invertierbar; die Inverse heißt *Arcustangens*:

$$\arctan : \mathbb{R} \to \left[-\frac{\pi}{2}, \frac{\pi}{2}\right].$$

**Hyperbolische Funktionen.** Sinus und Cosinus hyperbolicus sind definiert durch

$$\sinh x = \frac{1}{2}(e^x - e^{-x}), \qquad x \in \mathbb{R},$$
$$\cosh x = \frac{1}{2}(e^x + e^{-x}), \qquad x \in \mathbb{R}.$$

Der Name rührt daher, daß

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

gilt und $y^2 - x^2 = 1$ die Gleichung einer Hyperbel ist. Analog zum Tangens gibt es einen Tangens hyperbolicus:

$$\tanh x = \frac{\sinh x}{\cosh x} = \frac{e^x - e^{-x}}{e^x + e^{-x}}, \qquad x \in \mathbb{R}.$$

Die Inversen werden auch Area-Funktionen genannt:

$$\begin{array}{ll} \operatorname{arsinh} : \mathbb{R} \to \mathbb{R}, & \operatorname{arsinh} x = \ln(x + \sqrt{x^2+1}), \\ \operatorname{arcosh} : [1,\infty) \to [0,\infty), & \operatorname{arcosh} x = \ln(x + \sqrt{x^2-1}), \\ \operatorname{artanh} : (-1,1) \to \mathbb{R}, & \operatorname{artanh} x = \frac{1}{2}\ln\frac{1+x}{1-x}. \end{array}$$

Übrigens kann man in diese Funktionen auch komplexe Argumente einsetzen und erhält damit einen Zusammenhang zwischen den trigonometrischen und den hyperbolischen Funktionen. Benutzen wir nämlich die Eigenschaft, daß $\sin(-x) = -\sin x$ und $\cos(-x) = \cos x$, so folgt aus der Eulerschen Formel

$$e^{iz} = \cos z + i \sin z, \qquad z \in \mathbb{C},$$

für $z = -ix$ und $z = ix$, $x \in \mathbb{R}$:

$$\begin{array}{rcl} e^x & = & \cos(ix) - i\sin(ix), \\ e^{-x} & = & \cos(ix) + i\sin(ix). \end{array}$$

Addition und Subtraktion der beiden Gleichungen ergibt

$$\cosh x = \frac{1}{2}(e^x + e^{-x}) = \cos(ix),$$
$$\sinh x = \frac{1}{2}(e^x - e^{-x}) = -i\sin(ix).$$

# 4 Vektoren und Matrizen

## 4.1 Vektoren

**Definition 4.1** *Der Raum $\mathbb{R}^n$ ist das n-fache kartesische Produkt $\mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R}$ oder*

$$\mathbb{R}^n = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} : x_1, \ldots, x_n \in \mathbb{R} \right\}.$$

*Die Elemente des $\mathbb{R}^n$ heißen* Vektoren*.*

**Definition 4.2** *(1) Sei $x \in \mathbb{R}^n$. Der* transponierte Vektor *$x^\top$ ist definiert durch*

$$x^\top = (x_1, x_2, \ldots, x_n),$$

*d.h., $x^\top$ ist ein Zeilenvektor, während $x$ ein Spaltenvektor ist. Es gilt umgekehrt $x = (x_1, x_2, \ldots, x_n)^\top$.*

*(2) Seien $x = (x_1, \ldots, x_n)^\top$, $y = (y_1, \ldots, y_n)^\top \in \mathbb{R}^n$ und $\lambda \in \mathbb{R}$. Die Addition bzw. Subtraktion von Vektoren und die Multiplikation mit dem Skalar sind definiert durch*

$$\begin{aligned} x + y &= (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n)^\top, \\ x - y &= (x_1 - y_1, x_2 - y_2, \ldots, x_n - y_n)^\top, \\ \lambda x &= (\lambda x_1, \lambda x_2, \ldots, \lambda x_n)^\top. \end{aligned}$$

Die Operationen mit Vektoren sind in Abbildung 4.1 veranschaulicht.

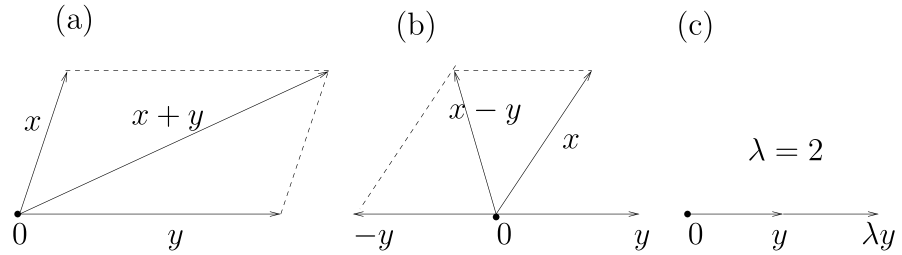


Abbildung 4.1: (a) Addition, (b) Subtraktion von Vektoren und (c) Multiplikation mit dem Skalar $\lambda$.

**Definition 4.3** *Das* Skalarprodukt *im $\mathbb{R}^n$ lautet*

$$x \cdot y = \sum_{i=1}^{n} x_i y_i \qquad \text{für} \quad x = (x_1, \ldots, x_n)^\top,\ y = (y_1, \ldots, y_n)^\top \in \mathbb{R}^n.$$

**Beispiel 4.4** Seien $x = (2, 1)^\top$, $y = (-1, 2)^\top \in \mathbb{R}^2$. Dann folgt $x \cdot y = 2 \cdot (-1) + 1 \cdot 2 = 0$. Nun steht $x$ senkrecht auf $y$ (siehe Abbildung 4.2); man erhält allgemein $x \cdot y = 0$ genau dann, wenn $x$ senkrecht auf $y$ steht. □

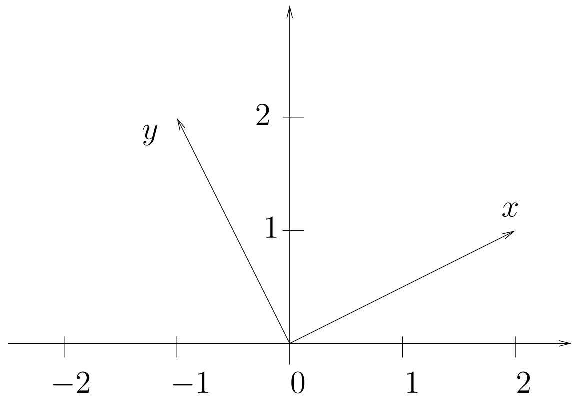


Abbildung 4.2: Die Vektoren $x$ und $y$ stehen senkrecht aufeinander.

Das Skalarprodukt hat die folgenden Eigenschaften. Für alle $x, y, z \in \mathbb{R}^n$ gilt:

(1) $(x + y) \cdot z = x \cdot z + y \cdot z$,

(2) $x \cdot (y + z) = x \cdot y + x \cdot z$,

(3) $x \cdot y = y \cdot x$,

(4) $x \cdot x \geq 0$ und $x \cdot x = 0 \iff x = 0$.

**Definition 4.5** *Die* Länge *oder* Norm *eines Vektors* $x = (x_1, \ldots, x_n)^\top \in \mathbb{R}^n$ *ist definiert als*

$$\|x\| = \sqrt{x \cdot x} = \sqrt{\sum_{i=1}^{n} x_i^2}.$$

Diese Definition ist in Abbildung 4.3 veranschaulicht. Die Norm hat die folgenden Eigenschaften. Für alle $x \in \mathbb{R}^n$ und $\lambda \in \mathbb{R}$ gilt:

(1) $\|x\| \geq 0$ und $\|x\| = 0 \iff x = 0$ (Positivität),

(2) $\|\lambda x\| = |\lambda| \cdot \|x\|$ (Homogenität),

(3) $\|x + y\| \leq \|x\| + \|y\|$ (Dreiecksungleichung).

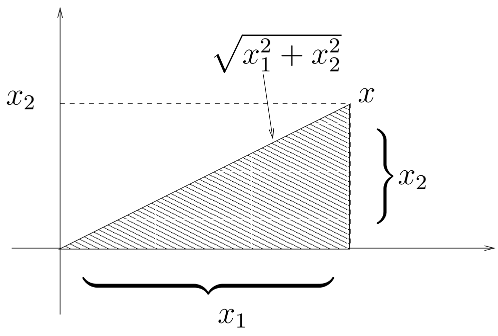


Abbildung 4.3: Zur Definition der Norm eines Vektors im $\mathbb{R}^2$.

**Definition 4.6** *Der* Winkel $\alpha \in [0, \pi]$ *zwischen zwei Vektoren* $x, y \in \mathbb{R}^n$ *ist definiert durch*

$$\cos \alpha = \frac{x \cdot y}{\|x\| \cdot \|y\|}.$$

*Wir sagen,* $x$ steht senkrecht auf $y$ *bzw.* $x$ *ist* orthogonal *zu* $y$ *(kurz* $x \perp y$*) genau dann, wenn* $x \cdot y = 0$ *oder* $\alpha = \frac{\pi}{2}$.

**Beispiel 4.7** Seien $x = (2, 1)^\top$, $y = (-1, 2)^\top$, $z \in (0, 2)^\top \in \mathbb{R}^2$ (siehe Abbildung 4.2). Dann ist

$$\begin{aligned}
\cos \alpha_{xy} &= \frac{x \cdot y}{\|x\| \cdot \|y\|} = 0, \qquad \text{also} \quad \alpha_{xy} = \frac{\pi}{2}, \\
\cos \alpha_{xz} &= \frac{x \cdot z}{\|x\| \cdot \|z\|} = \frac{2}{\sqrt{5} \cdot \sqrt{4}} = \frac{1}{\sqrt{5}}, \qquad \text{also} \quad \alpha_{xz} \approx 1.1071 = 63.43°, \\
\cos \alpha_{zy} &= \frac{z \cdot y}{\|z\| \cdot \|y\|} = \frac{4}{\sqrt{4} \cdot \sqrt{5}} = \frac{2}{\sqrt{5}}, \qquad \text{also} \quad \alpha_{zy} \approx 0.4636 = 26.57°.
\end{aligned}$$

Es gilt klarerweise $\alpha_{xz} + \alpha_{zy} = 90°$ oder $\alpha_{xz} + \alpha_{zy} \approx 1.5707 \approx \frac{\pi}{2}$. □

**Satz 4.8** (Pythagoras)
*Seien* $x, y \in \mathbb{R}^n$ *mit* $x \perp y$. *Dann ist*

$$\|x\|^2 + \|y\|^2 = \|x + y\|^2.$$

*Beweis:* Wegen $x \cdot y = 0$ nach Definition 4.6 folgt

$$\begin{aligned}
\|x + y\|^2 &= (x + y) \cdot (x + y) = x \cdot x + x \cdot y + y \cdot x + y \cdot y \\
&= \|x\|^2 + 2x \cdot y + \|y\|^2 = \|x\|^2 + \|y\|^2.
\end{aligned}$$
□

Zwischen Vektoren im $\mathbb{R}^3$ (und *nur* im $\mathbb{R}^3$!) gibt es noch eine andere "Multiplikation".

**Definition 4.9** *Seien* $x = (x_1, x_2, x_3)^\top$, $y = (y_1, y_2, y_3)^\top \in \mathbb{R}^3$. *Dann lautet das* Kreuzprodukt *(auch* Vektorprodukt *genannt)*

$$x \times y = \begin{pmatrix} x_2y_3 - x_3y_2 \\ x_3y_1 - x_1y_3 \\ x_1y_2 - x_2y_1 \end{pmatrix}.$$

**Beispiel 4.10** Seien

$$x = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}, \qquad y = \begin{pmatrix} -2 \\ 1 \\ 3 \end{pmatrix}.$$

Dann ist

$$x \times y = \begin{pmatrix} 0 \cdot 3 - 2 \cdot 1 \\ 2 \cdot (-2) - 1 \cdot 3 \\ 1 \cdot 1 - 0 \cdot (-2) \end{pmatrix} = \begin{pmatrix} -2 \\ -7 \\ 1 \end{pmatrix}.$$

Es gilt interessanterweise

$$x \cdot (x \times y) = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} \cdot \begin{pmatrix} -2 \\ -7 \\ 1 \end{pmatrix} = 0, \qquad y \cdot (x \times y) = \begin{pmatrix} -2 \\ 1 \\ 3 \end{pmatrix} \cdot \begin{pmatrix} -2 \\ -7 \\ 1 \end{pmatrix} = 0,$$

d.h., sowohl $x$ als auch $y$ sind orthogonal zu $x \times y$. Diese Eigenschaft gilt allgemein. □

Das Kreuzprodukt hat die folgenden Eigenschaften. Seien $x, y \in \mathbb{R}^3$.

(1) $x \times y$ steht senkrecht auf $x$ und $y$,

(2) $\|x \times y\|^2 = \|x\|^2\|y\|^2 - (x, y)^2$,

(3) $x \times y = 0 \implies \exists c \in \mathbb{R} : x = cy$.

Diese Eigenschaften können durch Nachrechnen bestätigt werden (für den Beweis von (3) verwendet man (2) und die sogenannte Cauchy-Schwarz-Ungleichung $|x \cdot y| \leq \|x\| \cdot \|y\|$).

## 4.2 Matrizen

**Definition 4.11** *Seien* $n, m \in \mathbb{N}$. *Ein rechteckiges Zahlenschema der Form*

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

*mit Elementen* $a_{ij} \in \mathbb{R}$ *heißt* $(m \times n)$*-Matrix. Wir schreiben kurz:* $A \in \mathbb{R}^{m \times n}$.

Es gibt einige spezielle Matrizen. Eine $(n \times 1)$-Matrix können wir als Vektor interpretieren.

**Definition 4.12** *(1) Eine Matrix $A \in \mathbb{R}^{m\times n}$ heißt* quadratisch, *wenn $m = n$.*

*(2) Eine* Diagonalmatrix *ist eine quadratische Matrix mit verschwindenden Nichtdiagonalelementen:*

$$D = \begin{pmatrix} d_1 & & & 0 \\ & d_2 & & \\ & & \ddots & \\ 0 & & & d_n \end{pmatrix} \in \mathbb{R}^{n\times n}.$$

*(3) Die* Einheitsmatrix *ist eine Diagonalmatrix, deren Diagonalelemente alle Eins sind:*

$$E = \begin{pmatrix} 1 & & & 0 \\ & 1 & & \\ & & \ddots & \\ 0 & & & 1 \end{pmatrix} \in \mathbb{R}^{n\times n}.$$

**Definition 4.13** *(1) Zwei Matrizen $A = (a_{ij})$, $B = (b_{ij}) \in \mathbb{R}^{m\times n}$ sind* gleich, *wenn alle Elemente übereinstimmen:*

$$A = B \quad \Longleftrightarrow \quad \forall\, i = 1, \ldots, m,\ j = 1, \ldots, n: \quad a_{ij} = b_{ij}.$$

*(2) Die* transponierte Matrix *$A^\top = (\widetilde{a}_{ij})$ für $A = (a_{ij}) \in \mathbb{R}^{n\times n}$ ist definiert durch*

$$\widetilde{a}_{ij} = a_{ji},$$

*d.h., man erhält $A^\top$ durch Vertauschen der Zeilen und Spalten von $A$.*

*(3) Eine quadratische Matrix $A \in \mathbb{R}^{n\times n}$ heißt* symmetrisch, *wenn $A^\top = A$.*

**Beispiel 4.14** Seien

$$A = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Dann ist $A$ quadratisch, aber nicht symmetrisch; $B$ ist dagegen symmetrisch. □

Matrizen werden elementweise addiert, subtrahiert bzw. mit Skalaren multipliziert. Seien $A = (a_{ij})$, $B = (b_{ij}) \in \mathbb{R}^{n\times m}$ und $\lambda \in \mathbb{R}$. Dann ist

$$\begin{aligned} (A+B)_{ij} &= a_{ij} + b_{ij}, \\ (A-B)_{ij} &= a_{ij} - b_{ij}, \\ (\lambda A)_{ij} &= \lambda a_{ij}. \end{aligned}$$

Die Multiplikation von Matrizen ist dagegen etwas aufwendiger.

**Definition 4.15** *Seien $A \in \mathbb{R}^{m\times n}$ und $B \in \mathbb{R}^{n\times \ell}$. Das* Matrixprodukt $C = AB \in \mathbb{R}^{m\times \ell}$ *ist definiert durch*

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj},$$

*wobei $C = (c_{ij})$.*

**Beispiel 4.16** (1) Seien

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 0 \end{pmatrix}, \qquad B = \begin{pmatrix} 3 & 2 \\ 0 & 1 \end{pmatrix}.$$

Man kann sich die Matrixmultiplikation durch das Schema "Zeile $\times$ Spalte" merken:

$$\begin{array}{cc|ll} \multicolumn{2}{c|}{\multirow{2}{*}{$AB$}} & 3 & 2 \\ & & 0 & 1 \\ \hline 1 & 2 & 1\cdot 3 + 2 \cdot 0 = 3 & 1 \cdot 2 + 2 \cdot 1 = 4 \\ 3 & 0 & 3 \cdot 3 + 0 \cdot 0 = 9 & 3 \cdot 2 + 0 \cdot 1 = 6 \end{array} \qquad \begin{array}{cc|cc} \multicolumn{2}{c|}{\multirow{2}{*}{$BA$}} & 1 & 2 \\ & & 3 & 0 \\ \hline 3 & 2 & 9 & 6 \\ 0 & 1 & 3 & 0 \end{array}$$

Daher ist

$$AB = \begin{pmatrix} 3 & 4 \\ 9 & 6 \end{pmatrix} \neq \begin{pmatrix} 9 & 6 \\ 3 & 0 \end{pmatrix} = BA.$$

(2) Seien

$$A = \begin{pmatrix} 2 & -1 & 2 \\ 4 & -2 & 4 \end{pmatrix}, \qquad B = \begin{pmatrix} -1 & 0 \\ 2 & 2 \\ 2 & 1 \end{pmatrix}, \qquad C = B^\top = \begin{pmatrix} -1 & 2 & 2 \\ 0 & 2 & 1 \end{pmatrix}.$$

Dann ist

$$\begin{array}{ccc|cc} & & & -1 & 0 \\ & AB & & 2 & 2 \\ & & & 2 & 1 \\ \hline 2 & -1 & 2 & 0 & 0 \\ 4 & -2 & 4 & 0 & 0 \end{array} \qquad \begin{array}{ccc|ccc} & \multirow{2}{*}{$AC$} & & -1 & 2 & 2 \\ & & & 0 & 2 & 1 \\ \hline 2 & -1 & 2 & \multicolumn{3}{c}{\text{nicht}} \\ 4 & -2 & 4 & \multicolumn{3}{c}{\text{definiert}} \end{array},$$

also

$$AB = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \qquad AC \text{ ist nicht definiert.}$$

□

Die Matrixmultiplikation hat also einige Besonderheiten:

(1) Das Produkt $AB$ existiert *nur*, wenn die Spaltenzahl von $A$ gleich der Zeilenzahl von $B$ ist.

(2) Falls $AB$ und $BA$ definiert sind, gilt im allgemeinen $AB \neq BA$.

(3) Falls $AB = 0$, muß *nicht* $A = 0$ oder $B = 0$ folgen.

Die Gleichung $ab = 1$ für gegebenes $a \in \mathbb{R}$, $a \neq 0$, besitzt die eindeutige Lösung $b = 1/a$, und man nennt $b$ die Inverse von $a$. Wir wollen dieses Konzept auf Matrizen erweitern.

**Definition 4.17** *Eine Matrix $A \in \mathbb{R}^{n\times n}$ heißt* invertierbar, *wenn eine Matrix $B \in \mathbb{R}^{n\times n}$ existiert, so daß*

$$AB = E \qquad \text{oder} \qquad BA = E,$$

*wobei $E \in \mathbb{R}^{n\times n}$ die Einheitsmatrix sei (siehe Definition 4.12). Wir schreiben $B = A^{-1}$ und nennen $A^{-1}$ die* Inverse *von $A$. Nicht invertierbare Matrizen heißen* singulär.

**Beispiel 4.18** (1) Sei

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}.$$

Dann ist $A$ invertierbar und

$$A^{-1} = \begin{pmatrix} -2 & 1 \\ 3/2 & -1/2 \end{pmatrix},$$

denn

$$AA^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = E.$$

(2) Die Matrix

$$A = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}$$

ist nicht invertierbar. Angenommen, es gäbe eine Matrix

$$B = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

mit $AB = E$. Dann folgt aus

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = A \cdot B = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a+c & b+d \\ 2(a+c) & 2(b+d) \end{pmatrix}$$

die Gleichungen

$$1 = a + c, \qquad 0 = 2(a+c),$$

also $0 = a + c = 1$ und damit ein Widerspruch. Solch eine Matrix $B$ kann also nicht existieren, d.h., $A$ ist singulär. $\square$

Für $(2 \times 2)$-Matrizen gibt es eine Formel für die Inverse. Für $(n \times n)$-Matrizen ist die Invertierung komplizierter und wird erst in Abschnitt 4.6 behandelt.

**Satz 4.19** (Cramersche Regel)
*Sei*

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbb{R}^{2\times 2}. \tag{4.1}$$

*Dann ist $A$ invertierbar genau dann, wenn $ad - bc \neq 0$. In diesem Fall gilt*

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

*Beweis:* Nachrechnen. □

**Satz 4.20** *Seien $A, B \in \mathbb{R}^{n\times n}$ invertierbar. Dann ist auch $AB$ invertierbar und*

$$(AB)^{-1} = B^{-1}A^{-1}.$$

*Beweis:* $(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AEA^{-1} = AA^{-1} = E$. □

Wofür braucht man die Inverse einer Matrix? Betrachte das Gleichungssystem

$$\begin{aligned} x_1 + 2x_2 &= -1, \\ 3x_1 + 4x_2 &= 3. \end{aligned} \tag{4.2}$$

Schreiben wir

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \qquad x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \qquad b = \begin{pmatrix} -1 \\ 3 \end{pmatrix},$$

so können wir (4.2) kompakt als

$$Ax = b$$

formulieren. Multiplizieren wir $A^{-1}$ von links auf beiden Seiten, so folgt

$$x = Ex = A^{-1}Ax = A^{-1}b$$

und wegen Beispiel 4.18 (1)

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = A^{-1}b = \begin{pmatrix} -2 & 1 \\ 3/2 & -1/2 \end{pmatrix} \begin{pmatrix} -1 \\ 3 \end{pmatrix} = \begin{pmatrix} 5 \\ -3 \end{pmatrix}.$$

Die Lösung von (4.2) lautet folglich $x_1 = 5$ und $x_2 = -3$. Mittels der Inversen kann also ein lineares Gleichungssystem bequem gelöst werden, sofern die Koeffizientenmatrix invertierbar ist (anderenfalls siehe Abschnitt 4.6).

## 4.3 Der Gauß-Algorithmus

In der Chemie tritt vielfach das Problem auf, unbekannte Größen $x_1, x_2, \ldots, x_n$ aus einem System linearer Gleichungen zu bestimmen. In Kapitel 1 haben wir gezeigt, daß die Bestimmung der Reaktionsgleichung für den Zerfall von Kaliumdichromat auf das lineare Gleichungssystem

$$\begin{aligned} 2x_1 - 2x_2 &= 0, \\ 2x_1 - x_2 - 2x_3 &= 0, \\ 7x_1 - 4x_2 - 3x_3 - 2x_4 &= 0 \end{aligned} \tag{4.3}$$

führt. Um dieses System von Gleichungen zu lösen, könnte man zuerst die erste Gleichung nach $x_2$ auflösen; dies führt auf $x_2 = x_1$. Setzen wir dieses Ergebnis in die zweite Gleichung ein, erhalten wir die Gleichung $x_1 - 2x_3 = 0$, also $x_3 = x_1/2$. Dies können wir in die letzte Gleichung einsetzen usw. Obwohl diese Einsetzmethode zum Ziel führt, wird sie bei *großen* Gleichungssystemen mit zehn oder mehr Gleichungen sehr umständlich. Es ist zweckmäßiger, einen Algorithmus zu entwickeln, mit dem Gleichungssysteme allgemein aufgelöst werden können. Der Vorteil eines Algorithmus ist, daß die Lösung von Gleichungssystemen mit Hilfe eines Computers ermöglicht wird. Ein solches Verfahren ist der *Gauß-Algorithmus* (auch Gaußsches Eliminationsverfahren genannt), den wir in diesem Abschnitt betrachten wollen.

Allgemein schreiben wir ein lineares Gleichungssystem mit $n$ Variablen $x_1, x_2, \ldots, x_n$ und $m$ Gleichungen als

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2, \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m. \end{aligned} \tag{4.4}$$

Die obige Schreibweise kann vereinfacht werden. Dazu schreiben wir die Koeffizienten als die *Koeffizientenmatrix*

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

und die Variablen $x_1, \ldots, x_n$ und die Werte $b_1, \ldots, b_m$ auf den rechten Seiten von (4.4) als

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \quad \text{und} \quad b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.$$

Damit können wir (4.4) kompakt formulieren als

$$Ax = b.$$

Die Matrix, die dadurch entsteht, wenn zur Matrix $A$ die Spalte $b$ hinzugefügt wird, nennen wir die *erweiterte Koeffizientenmatrix* und bezeichnen sie mit $(A|b)$. Sie besitzt $m$ Zeilen und $n+1$ Spalten.

Die Idee des Gauß-Algorithmus ist es, durch geschicktes Kombinieren von Gleichungen eine Matrix zu erhalten, deren Elemente unterhalb der Hauptdiagonalen alle gleich null sind. Dann kann das Gleichungssystem einfach rekursiv aufgelöst werden. Das Verfahren wird anhand der folgenden Beispiele deutlich.

**Beispiel 4.21** (1) Löse $Ax = b$ mit

$$A = \begin{pmatrix} 2 & 1 & -4 \\ 6 & 0 & 2 \\ -2 & 3 & 4 \end{pmatrix} \qquad \text{und} \qquad b = \begin{pmatrix} 7 \\ 0 \\ -7 \end{pmatrix}$$

Wir schreiben

$$\begin{array}{c} \mathrm{I} \\ \mathrm{II} \\ \mathrm{III} \end{array} \left(\begin{array}{ccc|c} 2 & 1 & -4 & 7 \\ 6 & 0 & 2 & 0 \\ -2 & 3 & 4 & -7 \end{array}\right) \begin{array}{c} \\ -3 \cdot \mathrm{I} \\ +\mathrm{I} \end{array}$$

$$\left(\begin{array}{ccc|c} 2 & 1 & -4 & 7 \\ 0 & -3 & 14 & -21 \\ 0 & 4 & 0 & 0 \end{array}\right) \begin{array}{c} \\ \\ +\frac{4}{3} \cdot \mathrm{II} \end{array}$$

$$\left(\begin{array}{ccc|c} 2 & 1 & -4 & 7 \\ 0 & -3 & 14 & -21 \\ 0 & 0 & \frac{56}{3} & -\frac{84}{3} \end{array}\right)$$

Aus der dritten Gleichung folgt

$$\begin{array}{rcl} \frac{56}{3}x_3 & = & -\frac{92}{3} & \Longrightarrow & x_3 & = & -\frac{3}{2}, \\ -3x_2 + 14x_3 & = & -21 & \Longrightarrow & x_2 & = & -\frac{1}{3}(-21 + 14 \cdot \frac{3}{2}) = 0, \\ 2x_1 + x_2 - 4x_3 & = & 7 & \Longrightarrow & x_1 & = & \frac{1}{2}(7 - 4 \cdot \frac{3}{2}) = \frac{1}{2}. \end{array}$$

Die Lösung lautet $x = (\frac{1}{2}, 0, -\frac{3}{2})$.

(2) Wir haben in Kapitel 1 gezeigt, daß die (ganzzahligen) Werte $x_1$,k $x_2$, $x_3$, $x_4$ der Reaktionsgleichung

$$x_1\,\mathrm{K_2Cr_2O_7} \to x_2\,\mathrm{K_2CrO_4} + x_3\,\mathrm{Cr_2O_3} + x_4\,\mathrm{O_2}$$

das lineare Gleichungssystem $Ax = b$ mit

$$A = \begin{pmatrix} 2 & -2 & 0 & 0 \\ 2 & -1 & -2 & 0 \\ 7 & -4 & -3 & -2 \end{pmatrix} \qquad \text{und} \qquad b = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

lösen müssen (siehe (1.2)). Wir schreiben

$$\left(\begin{array}{cccc|c} 2 & -2 & 0 & 0 & 0 \\ 2 & -1 & -2 & 0 & 0 \\ 7 & -4 & -3 & -2 & 0 \end{array}\right) \quad \begin{array}{c} \\ -\mathrm{II} \\ -\frac{7}{2}\cdot\mathrm{II} \end{array}$$

$$\left(\begin{array}{cccc|c} 2 & -2 & 0 & 0 & 0 \\ 0 & 1 & -2 & 0 & 0 \\ 0 & 3 & -3 & -2 & 0 \end{array}\right) \quad \begin{array}{c} \\ \\ -3\cdot\mathrm{II} \end{array}$$

$$\left(\begin{array}{cccc|c} 2 & -2 & 0 & 0 & 0 \\ 0 & 1 & -2 & 0 & 0 \\ 0 & 0 & 3 & -2 & 0 \end{array}\right).$$

Wir erhalten drei Gleichungen für vier Variable, also einen freien Parameter. Wir setzen $x_4 = \lambda$ und

$$\begin{array}{rcccccccl} 3x_3 & = & 2x_4 & = & 2\lambda & \Longrightarrow & x_3 & = & \frac{2}{3}\lambda, \\ x_2 & = & 2x_3 & = & \frac{4}{3}\lambda & \Longrightarrow & x_2 & = & \frac{4}{3}\lambda, \\ 2x_1 & = & 2x_2 & = & \frac{8}{3}\lambda & \Longrightarrow & x_1 & = & \frac{4}{3}\lambda. \end{array}$$

Die Lösungen lauten also

$$\frac{\lambda}{3}\begin{pmatrix} 4 \\ 4 \\ 2 \\ 3 \end{pmatrix}, \qquad \lambda \in \mathbb{R}.$$

Wir suchen nur ganzzahlige Lösungen und wählen daher (zum Beispiel) $\lambda = 3$. Damit lautet die obige Reaktionsgleichung

$$4\mathrm{K_2Cr_2O_7} \to 4\mathrm{K_2CrO_4} + 2\mathrm{Cr_2O_3} + 3\mathrm{O_2}.$$

(3) Löse das lineare Gleichungssystem $Ax = b$ mit

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 2 & 1 & 1 \\ -1 & -2 & -3 & -4 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Wir wenden den Gauß-Algorithmus an:

$$\begin{array}{c} \mathrm{I} \\ \mathrm{II} \\ \mathrm{III} \end{array} \left( \begin{array}{rrrr|r} 1 & 2 & 3 & 4 & 1 \\ 2 & 2 & 1 & 1 & 1 \\ -1 & -2 & -3 & -4 & 1 \end{array} \right) \begin{array}{l} \\ -2 \cdot \mathrm{I} \\ +\mathrm{I} \end{array} \qquad \begin{array}{c} \mathrm{I} \\ \mathrm{II} \\ \mathrm{III} \end{array} \left( \begin{array}{rrrr|r} 1 & 2 & 3 & 4 & 1 \\ 0 & -2 & -5 & -7 & -1 \\ 0 & 0 & 0 & 0 & 2 \end{array} \right).$$

Die letzte Gleichung lautet $0 \cdot x_1 + 0 \cdot x_2 + 0 \cdot x_3 + 0 \cdot x_4 = 2$ bzw. $0 = 2$. Diese Gleichung ist widersprüchlich. Was ist schiefgegangen? Das Resultat zeigt, daß das lineare Gleichungssystem keine Lösung besitzt, da die Umformungen des Gauß-Algorithmus auf eine widersprüchliche Gleichung führen. □

Die obigen Beispiele zeigen, daß ein lineares Gleichungssystem entweder keine, genau eine oder unendlich viele Lösungen besitzt. Welche Eigenschaften der Matrix bzw. der rechten Seite legen fest, wie viele Lösungen ein Gleichungssystem besitzt? Um diese Frage beantworten zu können, benötigen wir die Begriffe der Determinante und des Rangs einer Matrix.

## 4.4 Determinanten

Wir haben in Satz 4.19 gesehen, daß die $(2 \times 2)$-Matrix (4.1) genau dann invertierbar ist, wenn $ad - bc \neq 0$. Gibt es ein ähnliches Kriterium für allgemeine $(n \times n)$-Matrizen? Die Antwort lautet Ja und führt auf das Konzept der Determinante einer Matrix.

**Definition 4.22** *Die* Determinante *einer Matrix $A \in \mathbb{R}^{n \times n}$ ist rekursiv definiert durch*

$$\begin{aligned} \det A &= a_{11} \qquad \textit{für} \quad n = 1, \\ \det A &= \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det A^{ij} \qquad \textit{für} \quad n > 1, \end{aligned}$$

*wobei $j \in \{1, \ldots, n\}$ beliebig ist und $A^{ij}$ diejenige Matrix ist, die aus $A$ durch Streichen der $i$-ten Zeile und $j$-ten Spalte entsteht. In der obigen Formel können $i$ und $j$ vertauscht werden:*

$$\det A = \sum_{i=1}^{n} (-1)^{j+i} a_{ji} \det A^{ji}.$$

**Beispiel 4.23** (1) Für $(2 \times 2)$-Matrizen gilt

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = (-1)^{1+1} \cdot a \cdot \det(d) + (-1)^{1+2} \cdot b \cdot \det(c) = ad - bc.$$

(2) Sei

$$A = \begin{pmatrix} 3 & 0 & 2 \\ 1 & -1 & 3 \\ 2 & 4 & -2 \end{pmatrix}.$$

Dann folgt

$$\begin{aligned} \det A &= (-1)^{1+1} \cdot 3 \cdot \det \begin{pmatrix} -1 & 3 \\ 4 & -2 \end{pmatrix} + (-1)^{1+2} \cdot 0 \cdot \det \begin{pmatrix} 1 & 3 \\ 2 & -2 \end{pmatrix} \\ &\quad + (-1)^{1+3} \cdot 2 \cdot \det \begin{pmatrix} 1 & -1 \\ 2 & 4 \end{pmatrix} \\ &= 3(2-12) + 2(4+2) = -18. \end{aligned}$$

Es ist also sinnvoll, nach einer Zeile oder Spalte zu entwickeln, die möglichst viele Nullen enthält.

(3) Sei wie ein Beispiel 4.18 (1)

$$A = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}.$$

Dann ist $\det A = 1 \cdot 2 - 1 \cdot 2 = 0$. Wir wissen, daß $A$ nicht invertierbar ist. Gibt es einen Zusammenhang? Ja, siehe Satz 4.24 (3) unten. $\square$

Auch für $(3 \times 3)$- Matrizen gibt es eine Formel für die Determinante, nämlich die *Regel von Sarrus*:

$$\begin{aligned} \det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} &= \begin{array}{ccc|cc} a_{11} & a_{12} & a_{13} & a_{11} & a_{12} \\ a_{21} & a_{22} & a_{23} & a_{21} & a_{22} \\ a_{31} & a_{32} & a_{33} & a_{31} & a_{32} \end{array} \\ &= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ &\quad - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}. \end{aligned}$$

Die Regel gilt jedoch nicht für $(4 \times 4)$-Matrizen oder höherdimensionale Matrizen!

**Satz 4.24** *Seien $A, b \in \mathbb{R}^{n \times n}$. Dann gilt*

*(1)* $\det(AB) = \det A \cdot \det B$.

*(2)* $\det(A^\top) = \det A$.

*(3) A ist invertierbar genau dann, wenn* $\det A \neq 0$. *In diesem Fall gilt* $\det(A^{-1}) = 1/\det A$.

*(4) Ist* $\widetilde{A}$ *die Matrix, die entsteht, wenn man bei der Matrix A zwei Zeilen oder zwei Spalten miteinander vertauscht, so gilt* $\det \widetilde{A} = -\det A$.

Beachten Sie jedoch, daß im allgemeinen

$$\det(A+B) \neq \det A + \det B,$$

denn beispielsweise

$$\det\left[\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} + \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\right] = \det\begin{pmatrix}2 & 0\\ 0 & 2\end{pmatrix} = 4 \neq 1+1 = \det\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} + \det\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}.$$

Die Entwicklung nach einer Zeile oder Spalte erlaubt eine sehr einfache Berechnung von Matrizen, deren Elemente unterhalb der Hauptdiagonalen alle gleich null sind:

$$\det\begin{pmatrix}a_{11} & a_{12} & \cdots & a_{1n}\\ 0 & a_{22} & \cdots & a_{2n}\\ \vdots & 0 & \ddots & \vdots\\ 0 & 0 & & a_{nn}\end{pmatrix} = a_{11}a_{22}\cdots a_{nn}.$$

Solche Matrizen werden gerade beim Gauß-Algorithmus erzeugt! Da der Gauß-Algorithmus den Wert der Determinante einer Matrix nicht ändert, kann die Determinante grundsätzlich mit diesem Algorithmus berechnet werden.

**Beispiel 4.25** Sei wie in Beispiel 4.23 die Matrix

$$A = \begin{pmatrix}3 & 0 & 2\\ 1 & -1 & 3\\ 2 & 4 & -2\end{pmatrix}$$

gegeben. Der Gauß-Algorithmus ergibt:

$$\begin{pmatrix}3 & 0 & 2\\ 1 & -1 & 3\\ 2 & 4 & -2\end{pmatrix}\begin{matrix}\\ -\frac{1}{3}\cdot\mathrm{I}\\ -\frac{2}{3}\cdot\mathrm{I}\end{matrix} \qquad \begin{pmatrix}3 & 0 & 2\\ 0 & -1 & 7/3\\ 0 & 4 & -10/3\end{pmatrix}\begin{matrix}\\ \\ +4\cdot\mathrm{II}\end{matrix} \qquad \begin{pmatrix}3 & 0 & 2\\ 0 & -1 & 7/3\\ 0 & 0 & 6\end{pmatrix}.$$

Die Determinante ist dann das Produkt der Hauptdiagonalelemente: $\det A = 3\cdot(-1)\cdot 6 = -18$ in Übereinstimmung mit dem obigen Ergebnis. □

## 4.5 Lineare Unabhängigkeit

Kann man einer Matrix "ansehen", ob sie invertierbar ist bzw. (nach Satz 4.24 (3)) eine verschwindende Determinante hat? In einem gewissen Sinne gibt das Konzept der linearen Unabhängigkeit eine Antwort.

**Definition 4.26** *Die Vektoren* $a_1, \ldots, a_k \in \mathbb{R}^n$ *heißen* linear unabhängig, *wenn aus dem Ansatz*

$$\sum_{i=1}^{k} \lambda_i a_i = 0 \qquad \textit{mit} \quad \lambda_i \in \mathbb{R}$$

*folgt, daß* $\lambda_1 = \ldots = \lambda_k = 0$. *Anderenfalls heißen sie* linear abhängig.

**Beispiel 4.27** (1) Seien

$$a_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \qquad a_2 = \begin{pmatrix} 2 \\ 2 \end{pmatrix}.$$

Dann folgt aus

$$\lambda_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \lambda_2 \begin{pmatrix} 2 \\ 2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

*nicht* $\lambda_1 = \lambda_2 = 0$, denn $\lambda_1 = 2$, $\lambda_2 = -1$ ist auch eine Lösung. Es folgt also $2a_1 - a_2 = 0$ oder $a_2 = 2a_1$. Die Vektoren $a_1$ und $a_2$ sind linear abhängig. Dies können wir auch sehen, indem wir den Gauß-Algorithmus für das lineare Gleichungssystem

$$\begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = 0$$

durchführen:

$$\left(\begin{array}{cc|c} 1 & 2 & 0 \\ 1 & 2 & 0 \end{array}\right) \begin{array}{c} \\ -\mathrm{I} \end{array} \qquad \left(\begin{array}{cc|c} 1 & 2 & 0 \\ 0 & 0 & 0 \end{array}\right).$$

Wir erhalten eine Gleichung $1 \cdot \lambda_1 + 2 \cdot \lambda_2 = 0$ für zwei Unbekannte $\lambda_1$ und $\lambda_2$, d.h., es gibt unendlich viele Lösungen.

(2) Seien

$$a_1 = \begin{pmatrix} 1 \\ 2 \\ 0 \\ -1 \end{pmatrix}, \qquad a_2 = \begin{pmatrix} 2 \\ 1 \\ 0 \\ -1 \end{pmatrix}, \qquad a_3 = \begin{pmatrix} 1 \\ 2 \\ -1 \\ 0. \end{pmatrix}$$

Aus dem Ansatz

$$\lambda_1 \begin{pmatrix} 1 \\ 2 \\ 0 \\ -1 \end{pmatrix} + \lambda_2 \begin{pmatrix} 2 \\ 1 \\ 0 \\ -1 \end{pmatrix} + \lambda_3 \begin{pmatrix} 1 \\ 2 \\ -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

ergibt sich das folgende Gleichungssystem, das wir mit dem Gauß-Algorithmus lösen:

$$\left(\begin{array}{rrr|r} 1 & 2 & 1 & 0 \\ 2 & 1 & 2 & 0 \\ 0 & 0 & -1 & 0 \\ -1 & -1 & 0 & 0 \end{array}\right) \begin{array}{l} \\ -2\cdot\mathrm{I} \\ \\ +\mathrm{I} \end{array} \qquad \left(\begin{array}{rrr|r} 1 & 2 & 1 & 0 \\ 0 & -3 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 1 & 0 \end{array}\right) \begin{array}{l} \\ \\ \\ +\frac{1}{3}\cdot\mathrm{II} \end{array} \qquad \left(\begin{array}{rrr|r} 1 & 2 & 1 & 0 \\ 0 & -3 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right).$$

Es folgt aus der dritten und vierten Gleichung, daß $\lambda_3 = 0$, aus der zweiten Gleichung, daß $\lambda_2 = 0$, und schließlich aus der ersten Gleichung, daß $\lambda_1 = 0$. Die Vektoren sind folglich linear unabhängig. □

**Definition 4.28** *Der* Rang *einer Matrix* $A \in \mathbb{R}^{m\times n}$ *ist die (maximale) Anzahl der linear unabhängigen Zeilen oder Spalten. Wir schreiben* $\operatorname{rg} A$ *für den Rang von* $A$*.*

**Beispiel 4.29** (1) Sei wie in Beispiel 4.18 (2)

$$A = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}.$$

Wir haben in Beispiel 4.27 (1) gezeigt, daß die Zeilen von $A$ linear abhängig sind, d.h. $\operatorname{rg} A = 1$ (eine linear unabhängige Zeile).

(2) Sei

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 1 & 2 \\ 0 & 0 & -1 \\ -1 & -1 & 0 \end{pmatrix}.$$

Wir haben in Beispiel 4.27 (2) gezeigt, daß alle Spalten von $A$ linear unabhängig sind, d.h. $\operatorname{rg} A = 3$. □

Es gibt folgenden Zusammenhang zwischen Invertierbarkeit, Determinante und Rang einer Matrix.

**Satz 4.30** *Sei* $A \in \mathbb{R}^{n\times n}$*. Dann gilt*

$$A \textit{ invertierbar} \quad \Longleftrightarrow \quad \det A \neq 0 \quad \Longleftrightarrow \quad \operatorname{rg} A = n.$$

Der Rang einer Matrix gibt Informationen über die Lösbarkeit von linearen Gleichungssystemen. Betrachte

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= 0, \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= 0, \\ \vdots \qquad\qquad & \quad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n &= 0. \end{aligned}$$

Schreiben wir $\lambda_i$ anstatt $x_i$, so können wir dies auch formulieren als

$$\lambda_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} + \lambda_2 \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix} + \cdots + \lambda_n \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Sind die Spalten von $A = (a_{ij})$ linear unabhängig (bzw. gilt $\det A \neq 0$), so folgt $\lambda_1 = \ldots = \lambda_n = 0$. Das Gleichungssystem $Ax = 0$ besitzt also *nur* die Lösung $x = 0$. Das ist eigentlich klar, denn in diesem Fall ist $A$ invertierbar und

$$x = A^{-1}Ax = A^{-1}0 = 0.$$

Wir können also Satz 4.30 erweitern zu:

$$Ax = 0 \text{ besitzt die eindeutige Lösung } x = 0 \quad \Longleftrightarrow \quad \det A \neq 0.$$

## 4.6 Lösungstheorie linearer Gleichungssysteme

In diesem Abschnitt wollen wir mittels des Rangbegriffes die Lösbarkeit linearer Gleichungssysteme $Ax = b$ genauer charakterisieren. Ein lineares Gleichungssystem hat keine, genau eine oder unendlich viele Lösungen. Genauer gilt:

**Satz 4.31** *Seien $A \in \mathbb{R}^{m\times n}$ und $b \in \mathbb{R}^m$. Dann folgt*

*(1)* $\operatorname{rg} A \neq \operatorname{rg}(A|b) \implies$ *$Ax = b$ hat keine Lösung.*

*(2)* $\operatorname{rg} A = \operatorname{rg}(A|b) = n \implies$ *$Ax = b$ hat genau eine Lösung.*

*(3)* $\operatorname{rg} A = \operatorname{rg}(A|b) < n \implies$ *$Ax = b$ besitzt eine unendliche Schar von Lösungen, die durch $n - \operatorname{rg} A$ Parameter beschrieben wird.*

**Beispiel 4.32** (1) Seien

$$A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{pmatrix}, \qquad b_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \qquad b_2 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

Dann gilt $\operatorname{rg} A = 2$, $\operatorname{rg}(A|b_1) = 2$ (die erste und dritte Spalte sind gleich, also linear abhängig) und $\operatorname{rg}(A|b_2) = 3$. Daher besitzt

- $Ax = b_1$ wegen $n = 2 = \operatorname{rg} A = \operatorname{rg}(A|b)$ genau eine Lösung;
- $Ax = b_2$ keine Lösung.

Wir überprüfen diese Aussagen; $Ax = b_2$ bedeutet ausgeschrieben

$$\begin{aligned} 1 \cdot x_1 + 0 \cdot x_2 &= 1, \\ 1 \cdot x_1 + 1 \cdot x_2 &= 0, \\ 0 \cdot x_1 + 1 \cdot x_2 &= 0. \end{aligned}$$

Die erste und dritte Gleichung implizieren $x_1 = 1$ und $x_2 = 0$. Die zweite Gleichung ergibt

$$0 = x_1 + x_2 = 1,$$

also einen Widerspruch. Das Gleichungssystem kann keine Lösung besitzen. Das System $Ax = b_1$ bedeutet

$$x_1 = 1, \qquad x_1 + x_2 = 1, \qquad x_2 = 0.$$

Die eindeutig bestimmte Lösung lautet $x = (1, 0)^\top$.

(2) Seien

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}, \qquad b = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Dann existiert wegen $\operatorname{rg} A = 2 = \operatorname{rg}(A|b) < 3$ eine linear unabhängige Lösung. Wir können sie aus

$$x_1 + x_2 = 0, \qquad x_2 + x_3 = 0$$

berechnen, indem wir $x_2$ als Parameter betrachten:

$$x_1 = -\lambda, \qquad x_3 = -\lambda, \qquad x_2 = \lambda,$$

d.h., alle Vektoren

$$x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \lambda \begin{pmatrix} -1 \\ 1 \\ -1 \end{pmatrix}$$

sind Lösungen von $Ax = b$. Eine linear unabhängige Lösung lautet $(-1, 1, -1)^\top$. □

Der Gauß-Algorithmus erlaubt auch die Bestimmung der Inverse einer Matrix. Das geht wie folgt.

**Beispiel 4.33** Sei

$$A = \begin{pmatrix} 1 & -2 & 2 \\ 1 & 1 & -1 \\ 2 & 3 & 1 \end{pmatrix}.$$

Wir schreiben für den Gauß-Algorithmus

$$(A|E),$$

wobei $E$ die $(3 \times 3)$-Einheitsmatrix ist. Das Ziel lautet, die linke Seite so umzuformen, daß wir die Einheitsmatrix erhalten. Dann erhalten wir rechts die Inverse:

$$(E|A^{-1}).$$

Genauer ist:

$$\left(\begin{array}{ccc|ccc} 1 & -2 & 2 & 1 & 0 & 0 \\ 1 & 1 & -1 & 0 & 1 & 0 \\ 2 & 3 & 1 & 0 & 0 & 1 \end{array}\right) \qquad \begin{array}{c} \\ -\mathrm{I} \\ -2\cdot\mathrm{I} \end{array}$$

$$\left(\begin{array}{ccc|ccc} 1 & -2 & 2 & 1 & 0 & 0 \\ 0 & 3 & -3 & -1 & 1 & 0 \\ 0 & 7 & -3 & -2 & 0 & 1 \end{array}\right) \qquad \begin{array}{c} +\frac{2}{3}\cdot\mathrm{II} \\ :3 \\ -\frac{7}{3}\cdot\mathrm{II} \end{array}$$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{1}{3} & \frac{2}{3} & 0 \\ 0 & 1 & -1 & -\frac{1}{3} & \frac{1}{3} & 0 \\ 0 & 0 & 4 & \frac{1}{3} & -\frac{7}{3} & 1 \end{array}\right) \qquad \begin{array}{c} \\ +\frac{1}{4}\cdot\mathrm{III} \\ :4 \end{array}$$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{1}{3} & \frac{2}{3} & 0 \\ 0 & 1 & 0 & -\frac{1}{4} & -\frac{1}{4} & \frac{1}{4} \\ 0 & 0 & 1 & \frac{1}{12} & -\frac{7}{12} & \frac{1}{4} \end{array}\right)$$

Wir erhalten daher

$$A^{-1} = \frac{1}{12}\begin{pmatrix} 4 & 8 & 0 \\ -3 & -3 & 3 \\ 1 & -7 & 3 \end{pmatrix}.$$

□

# 5 Folgen und Reihen

## 5.1 Folgen

**Definition 5.1** *Wir nennen die Abfolge von (reellen oder komplexen) Zahlen* $a_1, a_2, a_3, \ldots$ *eine* Folge. *Die Elemente* $a_n$ *heißen* Folgenglieder, *und wir schreiben kurz* $(a_n)$ *für die Folge.*[*]

**Beispiel 5.2** (1) Die Vermehrung von Kaninchen wird häufig durch die *Fibonacci-Folge* beschrieben. Ein Kaninchenpaar werfe von jedem zweiten Monat an ein junges Paar und in jedem weiteren Monat ein weiteres Paar. Ist $a_n$ die Anzahl der Kaninchenpaare im Monat $n$, so wird $a_{n+1}$ rekursiv definiert durch

$$a_{n+1} = a_n + a_{n-1}, \qquad n \in \mathbb{N}, \quad a_0 = 1, \quad a_1 = 1.$$

Die ersten Folgenglieder sind

$$a_2 = 2, \quad a_3 = 3, \quad a_4 = 5, \quad a_5 = 8, \quad a_6 = 13, \quad a_7 = 21, \quad \text{usw.}$$

Fibonacci-Zahlen treten in einer Vielzahl von Anwendungen auf. Einige Beispiele sind:

- Für immer größere $n \in \mathbb{N}$ nähert sich das Verhältnis $a_{n+1}/a_n$ dem *Goldenen Schnitt* an, nämlich der Zahl $(1+\sqrt{5})/2 \approx 1.6180$. Der Name rührt daher, daß das Verhältnis 1 zu $(1+\sqrt{5})/2$ als besonders ästhetisch empfunden wird.

- Die Summen der Diagonalelemente des Pascalschen Dreiecks (siehe Abbildung 5.1) sind Fibonacci-Zahlen.

(2) Wir haben in Kapitel 1 die Menge $N_n$ einer radioaktiven Substanz zur Zeit $n \cdot \triangle t$ bestimmt:

$$N_n = \left(1 - \frac{1}{n}\right)^n, \qquad n \geq 2$$

(Falls $\alpha t = 1$; siehe Kapitel 1). Dann ist $N_2, N_3, N_4, \ldots$ eine Folge. Die ersten Folgenglieder lauten

$$N_2 = \frac{1}{4} = 0.25, \qquad N_3 = \left(\frac{2}{3}\right)^3 = 0.2963\ldots, \qquad N_4 = \left(\frac{3}{4}\right)^4 = 0.3164\ldots,$$

und weiter

$$N_{10} = \left(\frac{9}{10}\right)^{10} = 0.3487\ldots, \qquad N_{100} = \left(\frac{99}{100}\right)^{100} = 0.3660\ldots, \qquad N_{1000} = 0.3677\ldots$$

[*]Wir können Folgen mit dem Funktionenbegriff aus Kapitel 3 genauer als Funktionen $f : \mathbb{N} \to \mathbb{R}$ mit $f(n) = a_n$ definieren.

Es scheint, daß sich $N_n$ für immer größeres $n$ einem Wert nähert.

(3) Definiere die Folge

$$a_{n+1} = \frac{1}{2}\left(a_n + \frac{2}{a_n}\right), \qquad n \in \mathbb{N}, \quad a_1 = 1.$$

Die ersten Folgenglieder lauten

$$a_2 = 1.5, \qquad a_3 = 1.4166\ldots, \qquad a_{10} = 1.4142\ldots, \qquad a_{100} = 1.4142\ldots$$

$$\begin{array}{cccccccc}
1 & & & & & & & \\
1 & 1 & & & & & & \\
1 & 2 & 1 & & & & & \\
1 & 3 & 3 & 1 & & & & \\
1 & 4 & 6 & 4 & 1 & & & \\
1 & 5 & 10 & 10 & 5 & 1 & & \\
1 & 6 & 15 & 20 & 15 & 6 & 1 & \\
1 & 7 & 21 & 35 & 35 & 21 & 7 & 1
\end{array}$$

Abbildung 5.1: Die Summe der Diagonalelemente im Pascalschen Dreieck ergibt die Fibonacci-Zahlen, z.B. $1 + 4 + 3 = 0$, $1 + 5 + 6 + 1 = 13$,$1 + 6 + 10 + 4 = 21$ usw.

Es scheint, als strebe die Folge $a_n$ für große $n$ gegen den Wert $1.4142\ldots$. Wir vermuten, daß $a_n$ den Wert

$$\sqrt{2} = 1.4142\ldots$$

annähert. □

Die obigen Beispiele legen nahe, daß es Folgen gibt, die für immer größere $n$ einen Grenzwert besitzen. Dieser Begriff wird in der folgenden Definition präzisiert.

**Definition 5.3** *(1) Eine Folge* $(a_n)$ *besitzt den* Grenzwert *(oder* Limes*)* $a \in \mathbb{R}$ *genau dann, wenn*

$$\forall\, \varepsilon > 0 : \exists\, n_0 \in \mathbb{N} : \forall\, n \geq n_0 : \quad |a_n - a| \leq \varepsilon.$$

*Wir schreiben in diesem Fall*

$$\lim_{n\to\infty} a_n = a \qquad \textit{oder} \qquad a_n \to a \quad (n \to \infty).$$

*Eine nicht konvergente Folge heißt* divergent. *Eine konvergente Folge mit Grenzwert* $a = 0$ *heißt* Nullfolge.

*(2) Die Folge* $(a_n)$ *heißt* beschränkt *genau dann, wenn*

$$\exists\, M > 0 : \forall\, n \in \mathbb{N} : \quad |a_n| \le M.$$

**Beispiel 5.4** Sei $a_n = 1/n$, $n \in \mathbb{N}$. Wir vermuten, daß $(a_n)$ eine Nullfolge ist. Sei dazu $\varepsilon > 0$. Wir wählen ein $n_0 \in \mathbb{N}$, so daß $n_0 \ge 1/\varepsilon$. Dann folgt für beliebiges $n \ge n_0$

$$|a_n - 0| = |a_n| = \frac{1}{n} \le \frac{1}{n_0} \le \varepsilon.$$

Also ist $\lim_{n\to\infty}(1/n) = 0$. □

**Satz 5.5** *Konvergente Folgen sind beschränkt.*

**Beispiel 5.6** Klarerweise ist $a_n = 1/n$ beschränkt mit Schranke $M = 1$, denn $a_n = 1/n \le 1$. Die Folge der Fibonacci-Zahlen ist unbeschränkt und muß daher divergent sein, denn wäre sie konvergent, wäre sie ja nach Stz 5.5 beschränkt, was nicht zutrifft. □

Wichtig sind die folgenden Grenzwertsätze.

**Satz 5.7** (Grenzwertsätze)
*Seien* $(a_n)$, $(b_n)$ *konvergente Folgen mit Grenzwerten* $a$ *bzw.* $b$. *Dann gilt:*

*(1)* $\lim_{n\to\infty}(a_n \pm b_n) = a \pm b$;

*(2)* $\lim_{n\to\infty}(a_n \cdot b_n) = a \cdot b$;

*(3)* $\lim_{n\to\infty}(a_n/b_n) = a/b$, *falls* $b_n \ne 0$ *und* $b \ne 0$;

*(4)* $\forall\, n \in \mathbb{N} : a_n \le b_n \implies a \le b$.

Beachte, daß aus $a_n < b_n$ für alle $n \in \mathbb{N}$ *nicht* $a < b$ folgt. Ein Gegenbeispiel ist $a_n = 1/(n+1)$ und $b_n = 1/n$ (denn $a = b = 0$).

**Beispiel 5.8** (1) Sei

$$a_n = \frac{n^2 - n}{3n^2 + 1}, \qquad n \in \mathbb{N}.$$

Aus $\lim_{n\to\infty}(1/n) = 0$ und $\lim_{n\to\infty}(1/n^2) = 0$ folgt mit Satz 5.7

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} \frac{1 - \frac{1}{n}}{3 + \frac{1}{n^2}} = \frac{1 - \lim_{n\to\infty} \frac{1}{n}}{3 + \lim_{n\to\infty} \frac{1}{n^2}} = \frac{1}{3}.$$

(2) Sei $b_n = a_{n+1}/a_n$ der Verhältnis zweier aufeinander folgender Fibonacci-Zahlen. Aus $a_{n+1} = a_n + a_{n-1}$ folgt dann

$$b_n = \frac{a_n + a_{n-1}}{a_n} = 1 + \frac{1}{b_{n-1}}.$$

Angenommen, $(b_n)$ konvergiere gegen $b$. Dann ergibt Satz 5.7

$$b = \lim_{n\to\infty} b_n = 1 + \frac{1}{\lim_{n\to\infty} b_{n-1}} = 1 + \frac{1}{b}$$

oder $b^2 - b - 1 = 0$. Diese Gleichung hat die zwei Lösungen

$$b_{1/2} = \frac{1}{2} \pm \sqrt{\frac{1}{4} + 1} = \frac{1}{2}(1 + \sqrt{5}),$$

von denen aber nur eine positiv (und damit für uns relevant ist), nämlich $b_1 = (1+\sqrt{5})/2$. Dieses ist gerade der Goldene Schnitt, d.h.

$$\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = \frac{1}{2}(1 + \sqrt{5}).$$

(3) Sei $(a_n)$ rekursiv definiert durch

$$a_{n+1} = \frac{1}{4} + a_n^2, \qquad n \in \mathbb{N}, \quad a_1 = 1.$$

Angenommen, $(a_n)$ konvergiere gegen $a$. Wir erhalten dann mit Satz 5.7

$$a = \lim_{n\to\infty} a_{n+1} = \frac{1}{4} + \lim_{n\to\infty} a_n^2 = \frac{1}{4} + a^2$$

oder $a^2 - a + \frac{1}{4} = 0$. Diese Gleichung besitzt genau eine Lösung: $a = \frac{1}{2}$. Wir schließen

$$\lim_{n\to\infty} a_n = \frac{1}{2}.$$

Die ersten Folgenglieder von $(a_n)$ sind

$$a_1 = 1, \quad a_2 = \frac{5}{4} > 1, \quad a_3 = \frac{29}{16} > 1, \quad a_4 = \frac{84100}{256} > 1, \quad \text{usw.}$$

Man kann zeigen, daß $a_n > 1$ für alle $n \geq 2$. Aus Satz 5.7 (4) folgt dann

$$\frac{1}{2} = a = \lim_{n\to\infty} a_n \geq 1, \qquad \text{Widerspruch!}$$

Was ist schiefgelaufen? Die Folge $(a_n)$ ist divergent! Bevor man die Grenzwertsätze anwendet, muß man also prüfen, ob die entsprechenden Folgen wirklich konvergieren. $\square$

**Beispiel 5.9** Ein wichtiges Beispiel ist die Folge $a_n = (1 + 1/n)^n$, $n \in \mathbb{N}$. Eine naive Anwendung der Grenzwertsätze würde auf

$$\lim_{n\to\infty} a_n = \left(1 + \lim_{n\to\infty} \frac{1}{n}\right)^{\lim_{n\to\infty} n} = (1+0)^\infty = 1^\infty = 1$$

führen. Diese Argumentation ist jedoch *falsch*, denn der Grenzwert $n \to \infty$ muß simultan für den Quotienten $1/n$ und den Exponenten $n$ durchgeführt werden und nicht separat. Einige Folgenglieder lauten

$$a_1 = 1, \quad a_{10} = 2.5937\ldots, \quad a_{100} = 2.7048\ldots, \quad a_{1000} = 2.7169\ldots, \quad a_{10000} = 2.7181\ldots$$

Es scheint also, daß die Folge $(a_n)$ gegen einen Wert $2.718\ldots$ konvergiert. Tatsächlich kann man zeigen, daß $(a_n)$ gegen die Eulersche Zahl e konvergiert:

$$\lim_{n\to\infty}\left(1 + \frac{1}{n}\right)^n = 2.7182818\ldots = e.$$

$\square$

## 5.2 Reihen

**Definition 5.10** *(1) Sei $(a_n)$ eine Folge. Man nennt die Folge $(s_n)$ der Partialsummen,*

$$s_n = \sum_{k=1}^{n} a_k,$$

*eine (unendliche) Reihe und schreibt kurz $\sum_{k=1}^\infty a_k$.*

*(2) Die Reihe $\sum_{k=1}^\infty a_k$ heißt* konvergent *genau dann, wenn $(s_n)$ konvergiert. Man schreibt $s = \lim_{n\to\infty} s_n$ und schreibt kurz*

$$s = \sum_{k=1}^{\infty} a_k.$$

Das Symbol $\sum_{k=1}^\infty a_k$ hat also eine Doppelbedeutung:

- Es bezeichnet die *Folge der Partialsummen* $(s_n)$ und
- es bezeichnet den *Grenzwert* von $(s_n)$, falls er existiert.

**Beispiel 5.11** (1) Reihen treten in zahlreichen physikalischen und chemischen Anwendungen auf. Ein Beispiel ist die Gesamtanzahl entstandener Moleküle in einer chemischen Reaktion. Wir nehmen an, daß im Zeitraum $[n, n+1]$ etwa $a_n = e^{-n}$ $(n \in \mathbb{N}_0)$ Moleküle erzeugt werden. Dann liefert die Reaktion nach einiger Zeit zwar nur sehr wenige Moleküle, aber theoretisch läuft die Reaktion im Zeitintervall $[0, \infty)$ ab. Die Anzahl aller erzeugten Moleküle lautet also

$$\sum_{n=0}^{\infty} a_n = \sum_{n=0}^{\infty} e^{-n}.$$

(2) Reihen werden auch benötigt, um einige elementare Funktionen zu definieren. So ist die Exponentialfunktion $e^x$ etwa *definiert* durch die Reihe

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}, \quad x \in \mathbb{R}.$$

Man kann zeigen, daß diese Reihe für alle $x \in \mathbb{R}$ konvergiert. Auch die trigonometrischen Funktionen sin und cos sind als Reihen definiert, etwa

$$\sin x = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1}, \quad x \in \mathbb{R}.$$

**Beispiel 5.12** (1) Man nennt

$$\sum_{k=1}^{\infty} q^k \qquad \text{mit } q \in \mathbb{R}$$

die *geometrische Reihe*. Wegen

$$(1-q)\sum_{k=0}^{n} q^k = (1+q+\ldots+q^n) - (q+q^2+\ldots+q^{n+1}) = 1-q^{n+1}$$

lauten die Partialsummen

$$\sum_{k=0}^{n} q^k = \frac{1-q^{n+1}}{1-q}, \qquad \text{falls } q \neq 1.$$

Aus Satz 5.7 ergibt sich

$$\sum_{k=0}^{\infty} q^k = \lim_{n\to\infty} \sum_{k=0}^{n} q^k = \frac{1-\lim_{n\to\infty} q^{n+1}}{1-q}.$$

Der Grenzwert auf der rechten Seite existiert genau dann, wenn $|q| < 1$, nämlich $\lim_{n\to\infty} q^{n+1} = 0$. Folglich ist

$$\sum_{k=1}^{\infty} q^k = \frac{1}{1-q} \qquad \text{für } |q| < 1.$$

(2) Die *harmonische Reihe*

$$\sum_{k=1}^{\infty} \frac{1}{k}$$

ist divergent. Der Beweis ist nicht einfach, folgt aber aus der Ungleichung

$$\sum_{k=1}^{2^m} \frac{1}{k} \geq 1 + \frac{m}{2}.$$

(3) Betrachte die Reihe

$$s = \sum_{k=0}^{\infty} (-1)^k 2^k = 1 - 2 + 4 - 8 \pm \ldots$$

Dann erhalten wir

$$1 - 2s = 1 - \sum_{k=0}^{\infty} (-1)^k 2^{k+1} = 1 + \sum_{k=0}^{\infty} (-1)^{k+1} 2^{k+1} = 1 + \sum_{j=1}^{\infty} (-1)^j 2^j = \sum_{j=0}^{\infty} (-1)^j 2^j = s$$

und damit $s = 1/3$. Da die Reihenglieder alles natürliche Zahlen sind, sollte auch der Grenzwert eine natürliche Zahl sein. Was ist schiefgelaufen? Ganz einfach: Die Reihe konvergiert nicht, d.h., $s$ existiert nicht! $\square$

Aus den Sätzen 5.5 und 5.7 erhalten wir sofort die Grenzwertsätze für Reihen.

**Satz 5.13** *Seien $\sum_{k=1}^{\infty} a_k$ und $\sum_{k=1}^{\infty} b_k$ konvergente Reihen und sei $\lambda \in \mathbb{R}$. Dann gilt:*

*(1)* $\sum_{k=1}^{\infty} (a_k \pm b_k) = \sum_{k=1}^{\infty} a_k \pm \sum_{k=1}^{\infty} b_k$;

*(2)* $\sum_{k=1}^{\infty} \lambda a_k = \lambda \sum_{k=1}^{\infty} a_k$;

*(3)* $\sum_{k=1}^{\infty} a_k$ *ist konvergent* $\Longleftrightarrow$ $(s_n)$ *ist beschränkt.*

**Beispiel 5.14** Wir zeigen, daß die Reihe $\sum_{k=1}^{\infty} \frac{1}{k(k+1)}$ konvergiert. Es gilt nämlich

$$s_n = \sum_{k=1}^{n} \frac{1}{k(k+1)} = \sum_{k=1}^{n} \left( \frac{1}{k} - \frac{1}{k+1} \right) = \left( 1 + \frac{1}{2} + \ldots + \frac{1}{n} \right) - \left( \frac{1}{2} + \ldots + \frac{1}{n} + \frac{1}{n+1} \right)$$
$$= 1 - \frac{1}{n+1} \leq 1,$$

d.h., $(s_n)$ ist beschränkt. Aus Satz 5.13 (3) folgt die Behauptung. Beachte, daß wir den Wert der Reihe nicht bestimmt haben! $\square$

Folgende Kriterien sind nützlich, um eine Reihe auf Konvergenz zu testen.

**Satz 5.15** *Seien $(a_n)$ und $(b_n)$ Folgen.*

*(1) Notwendiges Kriterium:* $\sum_{k=1}^{\infty} a_k$ *ist konvergent* $\Longrightarrow$ $(a_n)$ *ist eine Nullfolge.*

(2) *Leibnizkriterium: Wenn* $(a_n)$ *eine monoton fallende Nullfolge ist, d.h.*

$$a_{n+1} \leq a_n \qquad \text{für alle } n \in \mathbb{N}, \qquad \lim_{n\to\infty} a_n = 0,$$

*dann ist* $\sum_{k=1}^{\infty}(-1)^k a_k$ *konvergent.*

(3) *Majorantenkriterium: Gilt* $|a_k| \leq b_k$ *für alle* $k$ *und ist* $\sum_{k=1}^{\infty} b_k$ *konvergent, so konvergiert auch* $\sum_{k=1}^{\infty} a_k$.

(4) *Minorantenkriterium: Gilt* $|a_k| \geq b_k$ *für alle* $k$ *und ist* $\sum_{k=1}^{\infty} b_k$ *divergent, so auch* $\sum_{k=1}^{\infty} a_k$.

(5) *Wurzelkriterium: Wenn* $\lim_{k\to\infty} \sqrt[k]{|a_k|} \leq q < 1$*, so ist* $\sum_{k=1}^{\infty} a_k$ *konvergent.*

(6) *Quotientenkriterium: Wenn* $\lim_{k\to\infty} |a_{k+1}/a_k| \leq q < 1$*, so ist* $\sum_{k=1}^{\infty} a_k$ *konvergent.*

**Beispiel 5.16** (1) Betrachte die Reihe $\sum_{n=0}^{\infty} \frac{n!}{n^n}$. Um zu prüfen, ob diese Reihe konvergiert oder nicht, müssen wir eines der obigen Konvergenzkritierien anwenden. Einen schnellen Nachweis erlauben meist das Wurzel- oder Quotientenkriterium (falls ansonsten nicht klar ist, welches Kriterium angewendet werden sollte). Das Wurzelkriterium führt mit $a_k = k!/k^k$ auf

$$\sqrt[k]{|a_k|} = \frac{\sqrt[k]{k!}}{\sqrt[k]{k^k}} = \frac{\sqrt[k]{k!}}{k}.$$

Die Schwierigkeit ist hier die Berechnung von $\sqrt[k]{k!}$, die nicht ohne weiteres möglich ist. Daher versuchen wir das Quotientenkriterium:

$$\left|\frac{a_{k+1}}{a_k}\right| = \frac{(k+1)!}{(k+1)^{k+1}} \, \frac{k^k}{k!} = \frac{(k+1)k^k}{(k+1)^k \cdot (k+1)} = \frac{k^k}{(k+1)^k} = \frac{1}{(1+1/k)^k}.$$

Die Folge $(1+1/k)^k$ haben wir bereits im vorigen Abschnitt kennengelernt. Es gilt: $\lim_{k\to\infty}(1+1/k)^k = e$. Damit ist

$$\lim_{k\to\infty}\left|\frac{a_{k+1}}{a_k}\right| = \frac{1}{e} < 1.$$

Also ist die Reihe nach dem Quotientenkriterium konvergent.

(2) Die Folge $\sum_{n=1}^{\infty}(-1)^n \frac{1}{n}$ ist nach dem Leibniz-Kriterium konvergent, aber *nicht*

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

denn dies ist die harmonische Reihe, die divergiert (siehe Beispiel 5.12 (2)).

(3) Die Reihe $\sum_{k=1}^{\infty} \frac{1}{k^2}$ ist konvergent, denn

$$|a_k| = \frac{1}{k^2} \leq \frac{1}{k} \cdot \frac{2}{k+1} \qquad \text{für } k \geq 1$$

und die Reihe

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

ist nach Beispiel 5.14 konvergent. Die Behauptung folgt also aus dem Majorantenkriterium.

(4) Die Reihe $\sum_{n=0}^{\infty} \frac{1}{n!}$ ist konvergent nach dem Quotientenkriterium, denn mit $a_n = 1/n!$ folgt

$$\left|\frac{a_{n+1}}{a_n}\right| = \frac{n!}{(n+1)n!} = \frac{1}{n+1} \leq \frac{1}{2} \qquad \text{für alle } n \geq 1.$$

Übrigens ist $\sum_{n=0}^{\infty} \frac{1}{n!} = e$.

□

Es ist wichtig, für das Wurzel- und Quotientenkriterium eine Zahl $q < 1$ und *nicht* nur $q \leq 1$ zu finden, da sonst die Aussage nicht gilt. Ein Beispiel ist die harmonische Reihe, für die zwar gilt

$$\left|\frac{a_{n+1}}{a_n}\right| = \frac{n}{n+1} < 1,$$

aber man kann kein $q < 1$ finden, so daß $n/(n+1) \leq q$ für *alle* $n \geq 1$ erfüllt ist!

Beachte außerdem, daß die Umkehrung des notwendigen Kriteriums i.a. *nicht* gilt: Ist die Folge $(a_n)$ konvergent, so folgt daraus nicht unbedingt, daß die Reihe $\sum_{k=1}^{\infty} a_n$ ebenfalls konvergent ist. Ein Gegenbeispiel ist wieder die harmonische Reihe $\sum_{k=1}^{\infty} 1/k$, denn die Folge $(1/n)$ konvergiert (gegen null), aber die Reihe ist divergent.

# 6 Differentiation

## 6.1 Stetigkeit

Aus den Grenzwertsätzen folgt für eine gegen $x \in \mathbb{R}$ konvergierende Folge $(x_n)$, daß auch

$$\lim_{n\to\infty} x_n^2 = \left(\lim_{n\to\infty} x_n\right)^2 = x^2$$

gilt. Folgt dann auch allgemein

$$\lim_{n\to\infty} f(x_n) = f\left(\lim_{n\to\infty} x_n\right) = f(x)?$$

Die Frage kann bejaht werden, sofern $f$ *stetig* ist. Dieser Begriff wird im folgenden definiert.

**Definition 6.1** *(1) Sei $f : D_f \subset \mathbb{R} \to \mathbb{R}$ eine Funktion. Sie* konvergiert *gegen $a \in \mathbb{R}$ für $x \to x_0$ genau dann, wenn* für alle *Folgen $(x_n) \subset D_f \setminus \{x_0\}$ mit $x_n \to x_0$ $(n \to \infty)$ gilt $f(x_n) \to a$ $(n \to \infty)$. In diesem Fall schreibt man*

$$\lim_{x\to x_0} f(x) = a.$$

*(2) Die Funktion $f : D_f \subset \mathbb{R} \to \mathbb{R}$ heißt* stetig *in $x_0 \in D_f$ genau dann, wenn*

$$\lim_{x\to x_0} f(x) = f(\lim_{x\to x_0} x) = f(x_0) \tag{6.1}$$

*gilt. Sie heißt* stetig, *falls sie in jedem $x_0 \in D_f$ stetig ist.*

**Beispiel 6.2** (1) Sei $f(x) = x^3 + 2x^2$, $x \in \mathbb{R}$. Ist $f$ stetig in $x_0 \in \mathbb{R}$? Sei $(x_n)$ eine beliebige Folge mit $x_n \to x_0$ $(n \to \infty)$. Aus den Grenzwertsätzen folgt

$$\lim_{n\to\infty} f(x_n) = \lim_{n\to\infty}(x_n^3 + 2x_n^2) = (\lim_{n\to\infty} x_n)^3 + 2(\lim_{n\to\infty} x_n)^2 = x_0^3 + 2x_0^2 = f(x_0).$$

Da $x_0$ und $(x_n)$ beliebig war, ist $f$ stetig in jedem $x_0 \in \mathbb{R}$, also ist $f$ stetig.

(2) Sei

$$f(x) = \begin{cases} -2x - 2 & \text{für } x \le 0, \\ 2 + 2x & \text{für } x > 0. \end{cases}$$

Ist $f$ stetig in $x_0 = 0$? Sei $x_n = 1/n$, $n \in \mathbb{N}$. Dann folgt

$$f(x_n) = 2 - 2x_n = 2 - \frac{2}{n} \to 2 \qquad (n \to \infty),$$

also ist

$$\lim_{n\to\infty} f(x_n) = 2 \ne -2 = f(0).$$

Die Funktion ist nicht stetig in $x_0 = 0$ (Abbildung 6.1). □

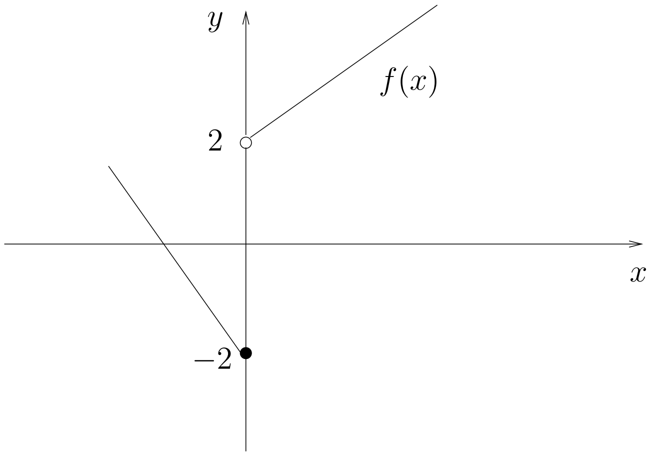


Abbildung 6.1: Graph der Funktion $f$ aus Beispiel 6.2 (2).

Der folgende Satz erweitert die Grenzwertsätze.

**Satz 6.3** *Seien $f$, $g$ und $h$ Funktionen, so daß $f$ und $g$ stetig in $x_0$ (mit $g(x_0) \neq 0$) und $h$ stetig in $y_0 = f(x_0)$ sind. Dann gilt:*

*(1) $f \pm g$, $f \cdot g$ und $f/g$ sind stetig in $x_0$;*

*(2) die Verkettung $h \circ f$, definiert durch $(h \circ f)(x) = h(f(x))$, ist stetig in $x_0$.*

Satz 6.3 zeigt, daß alle betrachteten elementaren Funktionen stetig sind, beispielsweise alle Polynome, exp, ln, sin, cos, sinh, cosh, usw.

## 6.2 Die Ableitung

Die Ableitung einer Funktion in einem Punkt ist die Steigung der Tangente in diesem Punkte (siehe Abbildung 6.2). Diese Konstruktion ist möglich, wenn der Graph der Funktion "glatt" ist. Schwierigkeiten gibt es etwa bei der Funktion in Abbildung 6.1: Die Tangente an $x_0 = 0$ hat die Steigung $-2$ (am Kurventeil für $x < 0$) bzw. die Steigung $+2$ (am Kurventeil für $x > 0$). Für die Ableitung ist die Stetigkeit ein notwendiges Kriterium. Sie ist nicht ausreichend, wie Abbildung 6.2 rechts zeigt: An der Stelle $x_0 = 0$ gibt es keine eindeutige Tangente.

Die Gleichung der Tangente $t(x)$ lautet

$$t(x) = f(x_0) + \alpha(x - x_0),$$

wobei $\alpha$ die Tangentensteigung ist. Falls $x \approx x_0$, gilt $f(x) \approx t(x)$ und

$$\alpha \approx \frac{f(x) - f(x_0)}{x - x_0}.$$

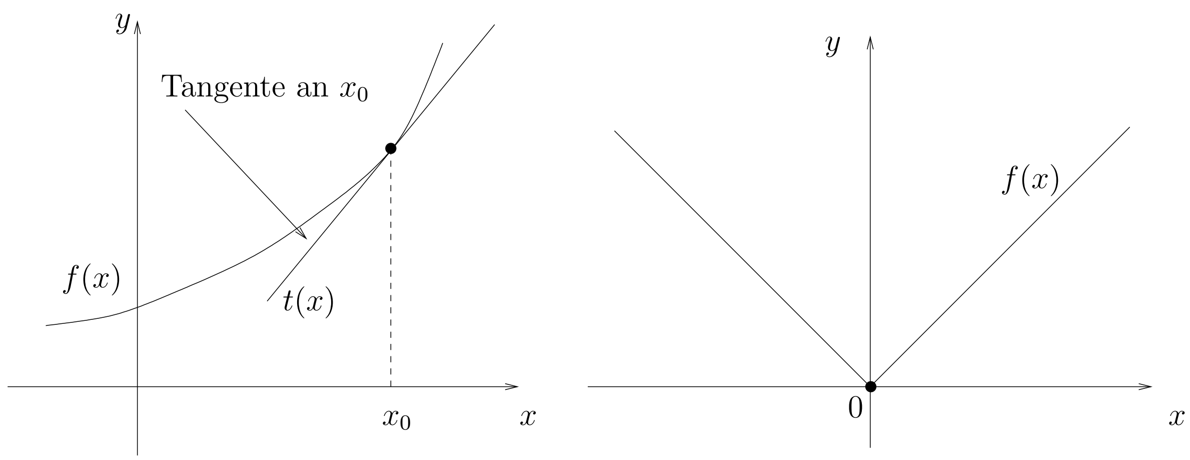


Abbildung 6.2: Illustration der Tangente am Punkt $x_0$ (links). Die rechte Funktion besitzt an $x_0 = 0$ keine eindeutige Tangente.

Die Tangentensteigung ist also näherungsweise gleich der Sekantensteigung. Wir definieren sie als den Grenzwert der Sekantensteigungen.

**Definition 6.4** *Sei $I \subset \mathbb{R}$ ein Intervall und $f : I \to \mathbb{R}$ eine Funktion. Dann heißt $f$* differenzierbar *in $x_0 \in I$, falls der Grenzwert*

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

*existiert. In diesem Fall schreibt man $f'(x_0)$ oder $\frac{df}{dx}(x_0)$ für diesen Grenzwert und nennt ihn die* Ableitung *von $f$ an $x_0$. Eine Funktion heißt* differenzierbar*, wenn sie in allen $x_0 \in I$ differenzierbar ist. Sie heißt* stetig differenzierbar*, wenn sie differenzierbar und die Ableitung stetig ist.*

**Bemerkung 6.5** (1) Schreibt man $x_0 + h = x$, so kann die Definition der Ableitung auch formuliert werden als

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

(2) Das Symbol $\frac{df}{dx}$ wird auch *Differentialquotient* genannt; es ist allerdings kein Quotient!

(3) Wir können $x \mapsto f'(x)$ wieder als Funktion interpretieren. Analog können *höhere Ableitungen* definiert werden:

$$f''(x_0) = \lim_{x \to x_0} \frac{f'(x) - f'(x_0)}{x - x_0} \qquad \text{usw.}$$

**Beispiel 6.6** (1) Sei $f(x) = x^2$, $x \in \mathbb{R}$. Sei $x_0 \in \mathbb{R}$. Dann gilt

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

und daher

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

Die Funktion $f$ ist also an jeder Stelle differenzierbar und $f'(x) = 2x$. Allgemein gilt:

$$f(x) = x^\alpha \quad \Rightarrow \quad f'(x) = \alpha x^{\alpha-1}, \quad \alpha \in \mathbb{R}. \tag{6.2}$$

(2) Sei $f(x) = |x|$, $x \in \mathbb{R}$ (siehe Abbildung 6.2 rechts). Sei $x_0 = 0$. Wegen

$$\frac{f(x_0+h) - f(x_0)}{h} = \frac{|x_0+h| - |x_0|}{h} = \frac{|h|}{h}$$

unterscheiden wir zwei Fälle:

- $h > 0:\quad \frac{1}{h}(f(x_0+h) - f(x_0)) = 1;$
- $h < 0:\quad \frac{1}{h}(f(x_0+h) - f(x_0)) = -1.$

Dies bedeutet, daß $f'(x_0)$ *nicht* existiert. Übrigens ist $f(x)$ an jeder anderen Stelle differenzierbar mit

$$f'(x) = \begin{cases} 1 & \text{für } x > 0, \\ -1 & \text{für } x < 0. \end{cases}$$

Die Ableitung $f'$ ist nur für alle $x \in \mathbb{R}$, $x \neq 0$, definiert. $\square$

Die elementaren Funktionen aus Abschnitt 3.3 sind differenzierbar; ihre Ableitungen lauten:

| $f(x)$ | $x^\alpha$ | $e^x$ | $\ln x$ | $\sin x$ | $\cos x$ | $\tan x$ | $\sinh$ | $\cosh$ |
|---|---|---|---|---|---|---|---|---|
| $f'(x)$ | $\alpha x^{\alpha-1}$ | $e^x$ | $\frac{1}{x}$ | $\cos x$ | $-\sin x$ | $\frac{1}{\cos^2 x}$ | $\cosh$ | $\sinh$ |

## 6.3 Rechenregeln

**Satz 6.7** *Seien $f$ und $g$ differenzierbare Funktionen und $a, b \in \mathbb{R}$. Dann sind auch $af+bg$, $f \cdot g$ und $f/g$ (falls $g(x) \neq 0$) differenzierbar, und es gilt:*

*(1) Linearität: $(af + bg)' = af' + bg'$,*

*(2) Produktregel: $(f \cdot g)' = f' \cdot g + f \cdot g'$;*

*(3) Quotientenregel: $\left(\frac{f}{g}\right)' = \frac{f' \cdot g - f \cdot g'}{g^2}$.*

**Beispiel 6.8** (1) Sei

$$P(x) = \sum_{k=0}^{n} a_k x^k \qquad \text{mit } n \in \mathbb{N},\ a_k \in \mathbb{R},$$

ein Polynom. Dann ist $P$ differenzierbar, und es gilt nach (6.2):

$$P'(x) = \sum_{k=1}^{n} a_k k x^{k-1}.$$

(2) Sei $f(x) = \tan x$, $-\pi/2 < x < \pi/2$. Dann ist nach der Quotientenregel

$$f'(x) = \left(\frac{\sin x}{\cos x}\right)' = \frac{\sin' x \cos x - \sin x \cos' x}{\cos^2 x} = \frac{\cos^2 x + \sin^2 x}{\cos^2 x} = \frac{1}{\cos^2 x}. \qquad \square$$

Die Produktregel kann für höhere Ableitungen verallgemeinert werden. Für genügend oft differenzierbare Funktionen $f$ und $g$ folgt

$$\begin{aligned} (f \cdot g)'' &= (f'g + fg')' = (f''g + f'g') + (f'g' + fg'') \\ &= f''g + 2f'g' + fg'', \\ (f \cdot g)''' &= (f'''g + f''g') + 2(f''g' + f'g'') + (f'g'' + fg''') \\ &= f'''g + 3f''g' + 3f'g'' + fg''' \qquad \text{usw.} \end{aligned}$$

Die Koeffizienten lauten also:

1 Ableitung: 1, 1;

2 Ableitungen: 1, 2, 1;

3 Ableitungen: 1, 3, 3, 1.

Diese Folge kommt uns bekannt vor: siehe das Pascalsche Dreieck aus Abschnitt 2.3! Wir schließen, daß der Koeffizient vor dem Produkt $f^{(k)}$, $g^{(n-k)}$ (d.h. $k$-te Ableitung von $f$ und $(n-k)$-te Ableitung von $g$) gerade der Binomialkoeffizient $\binom{n}{k}$ ist. Dies motiviert den folgenden Satz.

**Satz 6.9** (Leibniz-Formel)
*Seien $f$ und $g$ $n$-mal differenzierbar. Dann ist auch $f \cdot g$ $n$-mal differenzierbar, und es gilt*

$$(f \cdot g)^{(n)} = \sum_{k=0}^{n} \binom{n}{k} f^{(k)} g^{(n-k)}.$$

**Satz 6.10** (Kettenregel)
*Seien $f$ eine in $x_0$ differenzierbare Funktion und $g$ eine in $f(x_0)$ differenzierbare Funktion. Dann ist $h(x) = (g \circ f)(x) = g(f(x))$ in $x_0$ differenzierbar, und es gilt:*

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

*Man nennt $g'(f(x_0))$ die* äußere Ableitung *und $f'(x_0)$ die* innere Ableitung.

Wir können die Kettenregel mit Hilfe des Differentialquotienten auch einprägsam wie folgt formulieren. Seien $g(y)$ und $y(x)$ differenzierbar. Dann ist

$$\frac{dg}{dx} = \frac{dg}{dy}\frac{dy}{dx}.$$

Die linke Seite entspricht der Ableitung $\frac{d}{dx}g(y(x))$, die rechte $g'(y(x)) \cdot y'(x)$.

**Beispiel 6.11** (1) Sei $h(x) = \sin(x^2)$, $x \in \mathbb{R}$. Definiere $f(x) = x^2$ und $g(y) = \sin y$. Dann ist $h = g \circ f$, $f'(x) = 2x$, $g'(y) = \cos y$, also

$$h'(x) = g'(f(x))f'(x) = 2x\cos(x^2).$$

(2) Die Menge einer radioaktiven Substanz sei gegeben durch

$$N(t) = N_0 e^{-\alpha t}, \qquad t > 0, \quad \alpha > 0.$$

Definiere $f(t) = -\alpha t$ und $g(x) = N_0 e^x$. Dann ist $N = g \circ f$, und aus $f'(t) = -\alpha$ und $g'(x) = N_0 e^x$ folgt:

$$N'(t) = g'(f(t))f'(t) = -\alpha N_0^{-\alpha t} = -\alpha N(t).$$

Dies ist die Differentialgleichung aus Beispiel 1.2.

(3) Sei $h(x) = a^x$ mit $a > 0$. Man ist vielleicht versucht, für die Ableitung $h'(x) = xa^{x-1}$ zu schreiben; dies ist jedoch *falsch*! Schreiben wir nämlich $a^x = e^{\ln(a^x)} = e^{x \ln a}$ und definieren $f(x) = x \ln a$, $g(y) = e^y$, so folgt $h = g \circ f$ und $f'(x) = \ln a$, $g'(y) = e^y$, also

$$h'(x) = g'(f(x)) \cdot f'(x) = e^{x \ln a} \cdot \ln a = \ln a \cdot a^x.$$

(4) Die Ableitungen von sinh und cosh lauten

$$\sinh' = \cosh, \qquad \cosh' = \sinh,$$

denn

$$\begin{aligned}
\sinh'(x) &= \frac{1}{2}(e^x - e^{-x})' = \frac{1}{2}[(e^x)' - (e^{-x})'] = \frac{1}{2}[e^x - (-e^{-x})] = \cosh x, \\
\cosh'(x) &= \frac{1}{2}(e^x + e^{-x})' = \frac{1}{2}(e^x - e^{-x}) = \sinh x.
\end{aligned}$$

□

Die Ableitungen von arcsin, arsinh usw. können wir mit Hilfe des folgenden Satzes berechnen.

**Satz 6.12** (Ableitung der Umkehrfunktion)
*Sei $f : D_f \to W_f$ mit $D_f$, $W_f \subset \mathbb{R}$ eine invertierbare und differenzierbare Funktion mit $f'(x) \neq 0$ für alle $x \in D_f$. Dann ist die Inverse $f^{-1}$ auch differenzierbar, und es gilt für alle $y = f(x) \in W_f$:*

$$(f^{-1})'(y) = \frac{1}{f'(x)} = \frac{1}{f'(f^{-1}(y))}.$$

Wir können diese Regel wieder symbolisch für eine differenzierbare Funktion $y = y(x)$ mit Inverse $x = x(y)$ formulieren:

$$\frac{dx}{dy} = \frac{1}{\frac{dy}{dx}}.$$

Die linke Seite entspricht $x'(y)$, die rechte $1/y'(x)$.

**Beispiel 6.13** (1) Sei $f(x) = e^x$, $x \in \mathbb{R}$. Dann ist $D_f = \mathbb{R}$, $W_f = (0, \infty)$. Die Inverse lautet $f^{-1}(y) = \ln y$, $y \in (0, \infty)$, und

$$\ln'(y) = \frac{1}{f'(\ln y)} = \frac{1}{e^{\ln y}} = \frac{1}{y}.$$

(2) Sei $f(x) = \sin x$, $x \in D_f = (-\pi/2, \pi/2)$. Dann ist $W_f = (-1, 1)$ und $f^{-1}(y) = \arcsin y$, $y \in W_f$. Aus Satz 6.12 folgt

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

für alle $y \in (-1, 1)$. □

Wir fassen die Ableitungen einiger Umkehrfunktionen (Inversen) zusammen:

| $f(x)$ | $\arcsin x$ | $\arccos x$ | $\arctan x$ | $\operatorname{arsinh} x$ | $\operatorname{arcosh} x$ | $\operatorname{artanh} x$ |
|---|---|---|---|---|---|---|
| $f'(x)$ | $\frac{1}{\sqrt{1-x^2}}$ | $-\frac{1}{\sqrt{1-x^2}}$ | $\frac{1}{1+x^2}$ | $\frac{1}{\sqrt{x^2+1}}$ | $\frac{1}{\sqrt{x^2-1}}$ | $\frac{1}{1-x^2}$ |

Die Ableitung hilft bei der Bestimmung von Grenzwerten der Form

$$„\frac{0}{0}" \qquad \text{oder} \qquad „\frac{\infty}{\infty}",$$

nämlich beispielsweise bei $\lim_{x\to 0}(\sin x)/x$.

**Satz 6.14** (Regel von de l'Hospital)
*Seien $f, g : (a, b) \to \mathbb{R}$ differenzierbar und $g'(x) \neq 0$ für alle $x \in (a, b)$. Es gelte*

$$\lim_{x\to a} f(x) = \lim_{x\to a} g(x) = 0 \qquad \textit{oder} \qquad \lim_{x\to a} f(x) = \lim_{x\to a} g(x) = +\infty$$

*und der Grenzwert*

$$\lim_{x\to a} \frac{f'(x)}{g'(x)} \quad \textit{existiert.}$$

*Dann folgt*

$$\lim_{x\to a} \frac{f(x)}{g(x)} = \lim_{x\to a} \frac{f'(x)}{g'(x)}.$$

**Beispiel 6.15** (1) Der Grenzwert
$$\lim_{x\to 0} \frac{\sin x}{x}$$
ist vom Typ "$\frac{0}{0}$", da $\lim_{x\to 0} \sin x = \lim_{x\to 0} x = 0$. Der Grenzwert
$$\lim_{x\to 0} \frac{\sin' x}{(x)'} = \lim_{x\to 0} \frac{\cos x}{1} = 1$$
existiert, und daher ist
$$\lim_{x\to 0} \frac{\sin x}{x} = \lim_{x\to 0} \frac{\cos x}{1} = 1.$$

(2) Der Grenzwert
$$\lim_{x\to 0} \frac{\sin x - x}{x^3}$$
ist wieder vom Typ "$\frac{0}{0}$". Existiert der Grenzwert
$$\lim_{x\to 0} \frac{(\sin x - x)'}{(x^3)'} = \lim_{x\to 0} \frac{\cos x - 1}{3x^2}?$$
Er ist wieder vom Typ "$\frac{0}{0}$". Wir wenden die Regel von de l'Hospital auf den letzten Grenzwert an, und untersuchen
$$\lim_{x\to 0} \frac{(\cos x - 1)'}{(3x^2)'} = \lim_{x\to 0} \frac{-\sin x}{6x} = -\frac{1}{6} \lim_{x\to 0} \frac{\sin x}{x} = -\frac{1}{6}$$
nach (1). Wir erhalten also
$$\lim_{x\to 0} \frac{\sin x - x}{x^3} = \lim_{x\to 0} \frac{\cos x - 1}{3x^2} = \lim_{x\to 0} \frac{-\sin x}{6x} = -\frac{1}{6}.$$

(3) Der Grenzwert
$$\lim_{x\to 0} x \ln x$$
ist vom Typ "$0 \cdot \infty$", so daß die Regel von de l'Hospital nicht unmittelbar anwendbar ist. Wir schreiben daher
$$\lim_{x\to 0} x \ln x = \lim_{x\to 0} \frac{\ln x}{\frac{1}{x}} = \lim_{x\to 0} \frac{\frac{1}{x}}{-\frac{1}{x^2}} = -\lim_{x\to 0} x = 0.$$

(4) Der Grenzwert
$$\lim_{x\to 0} x^x$$
ist vom Typ "$0^0$", und die Regel von de l'Hospital ist wieder nicht anwendbar. Wir schreiben
$$\lim_{x\to 0} x^x = \lim_{x\to 0} e^{\ln(x^x)} = \lim_{x\to 0} e^{x \ln x}.$$

Vertauschen wir den Grenzwert und die Funktionsauswertung $e^y$, so folgt nach (3)

$$\lim_{x\to 0} e^{x\ln x} = e^{\lim_{x\to 0} x\ln x} = e^0 = 1.$$

Dürfen wir den Grenzwert und $e^y$ vertauschen? Ja, denn die $e$-Funktion ist stetig, und für solche Funktionen ist die Vertauschung möglich (siehe (6.1)).

(5) Die Regel von de l'Hospital führt auf

$$\lim_{x\to 0} \frac{\cos x}{x} = \lim_{x\to 0} \frac{-\sin x}{1} = 0,$$

wobei der Grenzwert vom Typ "$\frac{1}{0}$" ist. Die folgende Tabelle gibt einige Werte des Quotienten $(\cos x)/x$ an.

| $x$ | 0.1 | 0.01 | 0.001 | $-0.1$ | $-0.01$ | $-0.001$ |
|---|---|---|---|---|---|---|
| $\frac{\cos x}{x}$ | 9.95 | 99.995 | 999.9995 | $-9.95$ | $-99.995$ | $-999.9995$ |

Die Tabelle läßt vermuten, daß der Grenzwert gar nicht existiert! Tatsächlich gilt

$$\lim_{x\to 0, x>0} \frac{\cos x}{x} = +\infty, \qquad \lim_{x\to 0, x<0} \frac{\cos x}{x} = -\infty.$$

Im vorliegenden Fall sind die Voraussetzungen von Satz 6.14 nicht erfüllt. Die unkritische Anwendung der Regel von de l'Hospital kann also zu falschen Resultaten führen. □

Wir können die folgenden Fälle mit der Regel von de l'Hospital untersuchen:

- "$\infty - \infty$": Schreibe
$$f(x) - g(x) = \frac{\frac{1}{g(x)} - \frac{1}{f(x)}}{\frac{1}{f(x)g(x)}}.$$
  Der Quotient ist von der Form "$\frac{0}{0}$".

- "$0 \cdot \infty$": Schreibe
$$f(x) \cdot g(x) = \frac{f(x)}{\frac{1}{g(x)}}.$$
  Der Quotient ist wieder von der Form "$\frac{0}{0}$".

- "$0^0$": Schreibe $f(x)^{g(x)} = e^{g(x)\cdot \ln f(x)}$. Der Exponent ist vom Typ "$0 \cdot \infty$".

- "$1^\infty$": Schreibe $f(x)^{g(x)} = e^{g(x)\cdot \ln f(x)}$, und der Exponent hat die Form "$\infty \cdot 0$".

- "$\infty^0$": Schreibe wieder $f(x)^{g(x)} = e^{g(x)\cdot \ln f(x)}$ und erhalte den Typ "$0 \cdot \infty$".

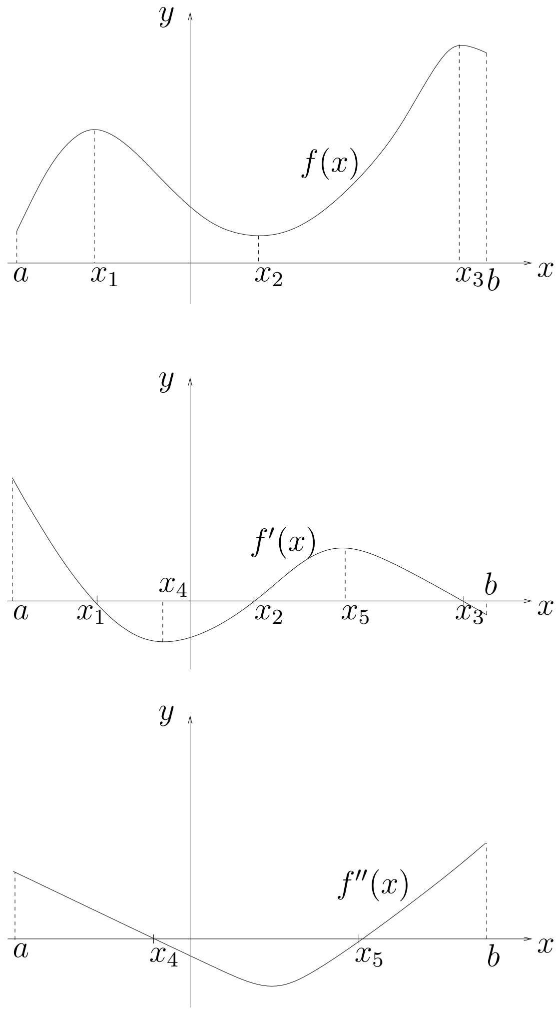


Abbildung 6.3: Die Funktion $f$ besitzt ein lokales Maximum in $x_1$ und ein lokales Minimum in $x_2$. Das *globale* Maximum wird dagegen in $x_3$ erreicht. Die Stellen $x_4$ und $x_5$ sind kritische Punkte von $f'$.

## 6.4 Kurvendiskussion

Ist die Steigung der Tangente an den Graphen einer Funktion stets positiv (bzw. negativ), so ist die Funktion monoton wachsend (bzw. fallend). Ist die Tangentensteigung zunächst positiv (bzw. negativ) und dann negativ (bzw. positiv), so erwarten wir, daß ein lokales Maximum (bzw. Minimum) vorliegt (siehe Abbildung 6.3).

**Definition 6.16** *Sei $f : [a,b] \to \mathbb{R}$ eine Funktion.*

*(1) $f$ heißt* monoton wachsend *auf $[a,b]$, wenn*

$$f(x_1) \le f(x_2) \qquad \textit{für alle} \quad x_1 \le x_2.$$

*(2) $f$ heißt* monoton fallend *auf $[a, b]$, wenn*

$$f(x_1) \geq f(x_2) \qquad \text{für alle} \quad x_1 \leq x_2.$$

*(3) $f$ besitzt in $x_0 \in (a, b)$ ein* lokales Maximum*, wenn*

$$\exists\, \varepsilon > 0 : \quad \forall\, x \in (x_0 - \varepsilon, x_0 + \varepsilon) : \quad f(x) \leq f(x_0). \tag{6.3}$$

*(4) $f$ besitzt in $x_0 \in (a, b)$ ein* lokales Minimum*, wenn*

$$\exists\, \varepsilon > 0 : \quad \forall\, x \in (x_0 - \varepsilon, x_0 + \varepsilon) : \quad f(x) \geq f(x_0). \tag{6.4}$$

*(5) Ein* Extremum *ist ein Minimum oder Maximum. Ein Minimum oder Maximum wird* global *genannt, wenn die Ungleichung (6.3) bzw. (6.4) für alle $x \in [a, b]$ gilt.*

*(6) Ist $f$ differenzierbar, so heißt ein Punkt $x_0 \in (a, b)$ mit $f'(x_0) = 0$* kritischer Punkt.

**Satz 6.17** *Sei $f(a, b) \to \mathbb{R}$ zweimal differenzierbar und $x_0 \in (a, b)$.*

*(1) Gilt $f'(x) \geq 0$ (bzw. $f'(x) \leq 0$) für alle $x \in (a, b)$, so ist $f$ monoton wachsend (bzw. monoton fallend).*

*(2) Gilt $f'(x_0) = 0$ und $f''(x_0) < 0$ (oder $f''(x_0) > 0$), so hat $f$ in $x_0$ ein lokales Maximum (bzw. lokales Minimum).*

*(3) Die Funktion $f$ sei $n$-mal differenzierbar, und es gelte $f^{(i)}(x_0) = 0$ für alle $i = 1, \ldots, n-1$ und $f^{(n)}(x_0) \neq 0$. Dann gilt:*

- *$n$ gerade: $f$ hat in $x_0$ ein Extremum, und zwar ein lokales Maximum, wenn $f^{(n)}(x_0) < 0$, bzw. ein lokales Minimum, wenn $f^{(n)}(x_0) > 0$;*
- *$n$ ungerade: $f$ hat in $x_0$ kein Extremum.*

**Beispiel 6.18** Sei $f(x) = x^4 - 2x^2 + 1 = (x^2 - 1)^2$, $x \in \mathbb{R}$ (siehe Abbildung 6.4 links). Wir bestimmen alle Extrema von $f$. Aus

$$0 = f'(x) = 4x^3 - 4x = 4x(x^2 - 1) = 4x(x - 1)(x + 1)$$

folgt, daß $x_1 = 0$, $x_2 = 1$ und $x_3 = -1$ Kandidaten für Extrema sind. Wir untersuchen diese Punkte einzeln:

- $x_1 = 0$: $f''(x_1) = 12x_1^2 - 4 = -4 < 0 \implies f$ hat an $x_1 = 0$ ein lokales Maximum.
- $x_{2/3} = \pm 1$: $f''(x_{2/3}) = 12x_{2/3}^2 - 4 = 8 > 0 \implies f$ hat an $x_{2/3} = \pm 1$ ein lokales Minimum.

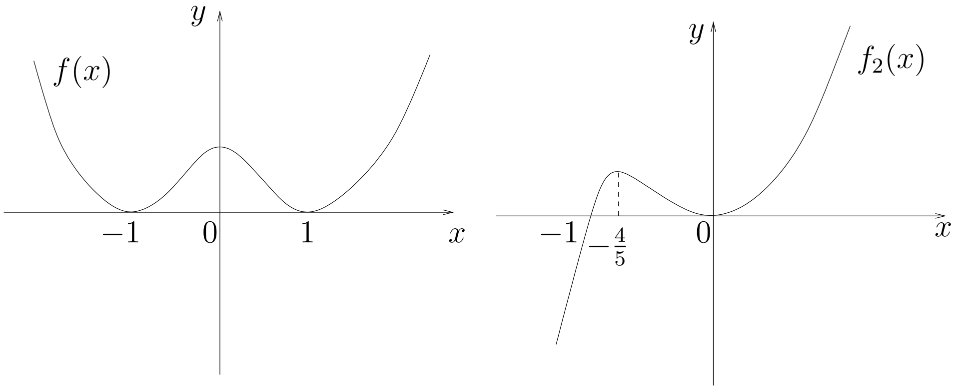


Abbildung 6.4: Graph der Funktion $f(x) = x^4 - 2x^2 + 1$ (links) und Graph der Funktion $f_2(x) = x^5 + x^4$ (rechts).

(2) Sei $f(x) = x^{-12} - ax^{-6}$ mit $a > 0$, $x > 0$. Die Funktion $f$ ist ein Modell für das Potential von Teilchen in der Nähe des Atomkerns mit Abstand $x$; sie wird auch *Lenard-Jones-Potential* genannt. Der erste Summand $x^{-12}$ beschreibt eine extrem kurzreichweitige Abstoßung, während der zweite Summand $ax^{-6}$ eine langreichweitige Anziehung modelliert. Alle kritischen Punkte von $f$ sind gegeben durch

$$0 = f'(x) = -\frac{12}{x^{13}} + \frac{6a}{x^7} = \frac{6}{x^7}\left(-\frac{2}{x^6} + a\right) \implies x = \pm\sqrt[6]{\frac{2}{a}}.$$

Da der Abstand $x$ positiv sein soll, kommt nur $x_0 = \sqrt[6]{2/a}$ in Betracht. Wegen

$$f''(x_0) = \frac{12 \cdot 13}{x_0^{14}} - \frac{6 \cdot 7a}{x_0^8} = \frac{6}{x_0^8}\left(\frac{2 \cdot 13}{x_0^6} - 7a\right) = \frac{6}{x_0^8}\,(13a - 7a) > 0$$

liegt an der Stelle $x_0$ ein lokales Minimum vor. Es folgt außerdem, daß $f$ für $x < x_0$ monoton fallend und für $x > x_0$ monoton wachsend ist (siehe Abbildung 6.5).

(4) Die Van-der-Waals-Gleichung aus Beispiel 3.7 für den Druck $p(V)$ in Abhängigkeit des Volumens $V$

$$\left(p(V) + \frac{a}{V^2}\right)(V - b) = RT, \qquad V > b \quad \text{und} \quad a, b > 0,$$

kann nach $p(V)$ aufgelöst werden:

$$p(V) = \frac{RT}{V - b} - \frac{a}{V^2}, \qquad V > b.$$

Kritische Punkte erhalten wir aus der Gleichung

$$0 = p'(V) = -\frac{RT}{(V - b)^2} + \frac{2a}{V^3} \quad \Longleftrightarrow \quad 2a(V - b)^2 = RTV^3.$$

Abbildung 6.6 zeigt, daß diese Gleichung

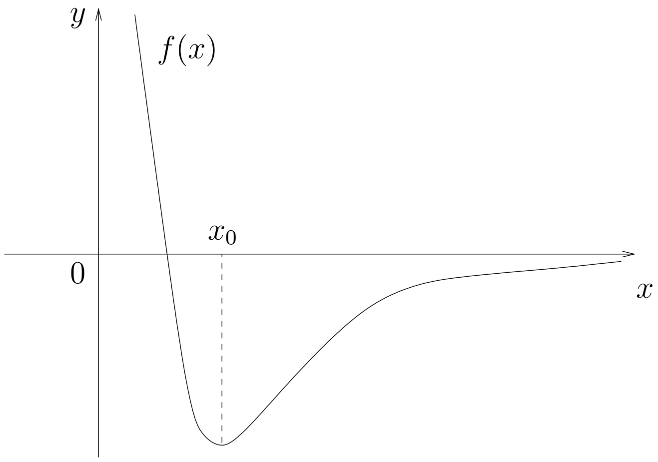


Abbildung 6.5: Graph der Funktion $f(x) = x^{-12} - ax^{-6}$, $x > 0$.

- keine Lösung besitzt, wenn $T > T_{\text{krit}}$;
- genau eine Lösung $V_0$ besitzt, wenn $T = T_{\text{krit}}$;
- genau zwei Lösungen $V_1$ und $V_2$ besitzt, wenn $T < T_{\text{krit}}$.

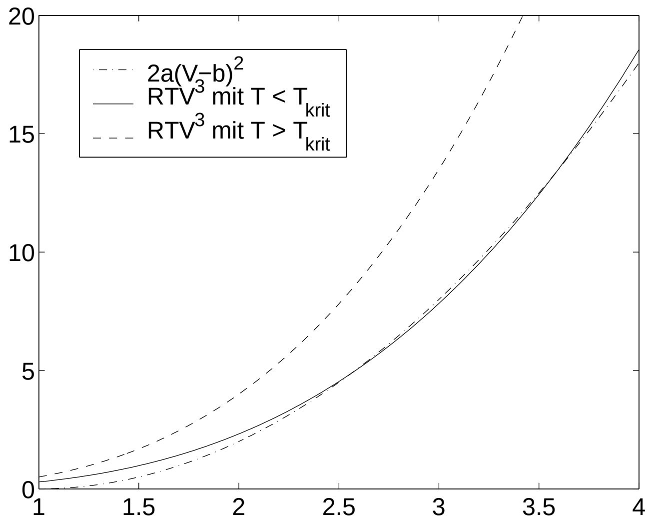


Abbildung 6.6: Graphische Lösung der Gleichung $2a(V - b)^2 = RTV^3$ für den Fall $a = b = R = 1$. Man kann berechnen, daß hier $T_{\text{krit}} \approx 0.29606$.

Diese Beobachtung erlaubt es, das Monotonieverhalten von $p(V)$ zu beschreiben:

- $T \geq T_{\text{krit}}$: $p'(V) \leq 0$ für alle $V > b \implies p$ monoton fallend in $(b, \infty)$;

- $T < T_{\text{krit}}$: $p$ ist monoton fallend in $(b, V_1)$ und $(V_2, \infty)$ und monoton wachsend in $(V_1, V_2)$.

Der Graph von $p(V)$ ist in Abbildung 6.7 skizziert. Es gibt eine Region, in der für gegebenen Druck $p$ *drei* verschiedene Volumina existieren, so daß die Van-der-Waals-Gleichung erfüllt ist. Diese Region beschreibt die Koexistenz der flüssigen und gasförmigen Phase. $\square$

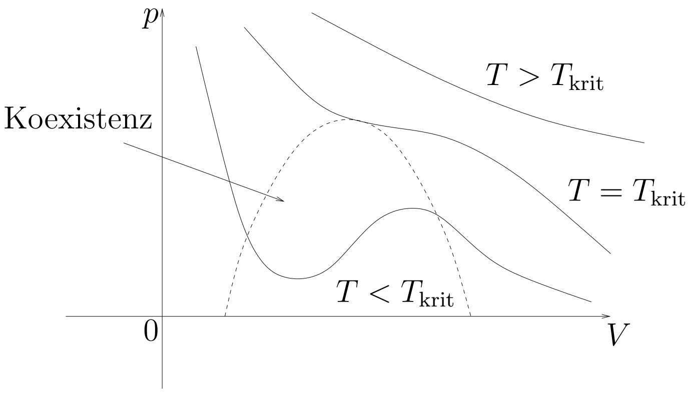


Abbildung 6.7: Graphen der Funktion $p(V)$ für verschiedene Temperaturen.

Eine Kurvendiskussion einer Funktion $f$ sollte typischerweise die folgenden Punkte enthalten:

(1) Bestimmung der Nullstellen von $f$,

(2) Bestimmung der Monotonie von $f$,

(3) Bestimmung der lokalen Minima und Maxima von $f$,

(4) Bestimmung des Verhaltens für $x \to +\infty$ bzw. $x \to -\infty$, falls zutreffend,

(5) eine Skizze des Graphen von $f$.

## 6.5 Satz von Taylor

Differenzierbare Funktionen können oft durch Polynome angenähert werden. Diese Aussage wird im Satz von Taylor präzisiert.

**Definition 6.19** *Sei $f : (a,b) \to \mathbb{R}$ $(n+1)$-mal differenzierbar, und sei $x_0 \in (a,b)$. Dann heißt das Polynom*

$$T_n(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k, \qquad x \in \mathbb{R},$$

*das* Taylor-Polynom *$n$-ten Grades von $f$ um $x_0$.*

**Satz 6.20** (Taylor)
*Sei $f : (a,b) \to \mathbb{R}$ $(n+1)$-mal stetig differenzierbar (d.h., $f$ ist $(n+1)$-mal differenzierbar und $f^{(n+1)}$ ist stetig) und seien $x_0, x \in (a,b)$. Dann existiert ein $\xi$ zwischen $x_0$ und $x$, so daß*

$$f(x) = T_n(x) + R_{n+1}(x), \tag{6.5}$$

*wobei*

$$R_{n+1}(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-x_0)^{n+1}. \tag{6.6}$$

Man nennt (6.5) die *Taylor-Formel* und (6.6) das *Lagrange-Restglied*. Der Punkt $x_0$ wird auch *Entwicklungspunkt* genannt. Im Falle $n = 0$ erhalten wir:

$$f(x) = T_0(x) + R_1(x) = f(x_0) + f'(\xi)(x - x_0). \tag{6.7}$$

Man nennt diese Beziehung *Mittelwertsatz*. Er hat eine einfache geometrische Interpretation (siehe Abbildung 6.8). Schreiben wir (6.7) als

$$\frac{f(x) - f(x_0)}{x - x_0} = f'(\xi)$$

für $x \neq x_0$, so lautet die Aussage des Mittelwertsatzes, daß es einen Punkt $\xi$ gibt, an dem die Steigung der Tangente an den Graphen von $f$ (nämlich $f'(\xi)$) gleich der Steigung der Sekante durch $x$ und $x_0$ ist.

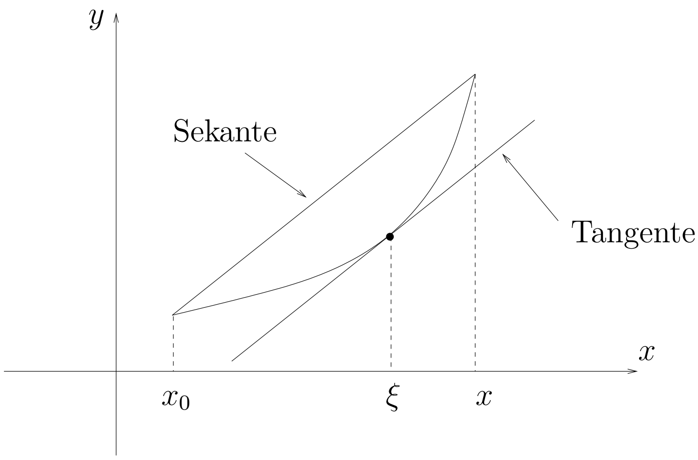


Abbildung 6.8: Geometrische Veranschaulichung des Mittelwertsatzes (6.7).

**Beispiel 6.21** Sei $f(x) = e^x$, $x \in \mathbb{R}$. Wir wählen den Entwicklungspunkt $x_0 = 0$. Die Funktion $f$ ist unendlich oft differenzierbar mit

$$f(0) = 1, \qquad f'(0) = 1, \qquad f''(0) = 1 \qquad \text{usw.}$$

Aus dem Satz 6.20 von Taylor folgt, daß es zu $x \in \mathbb{R}$ ein $\xi$ zwischen 0 und $x$ gibt, so daß

$$f(x) = e^x = \sum_{k=0}^{n} \frac{1}{k!} x^k + \frac{e^\xi}{(n+1)!} x^{n+1}.$$

Die Exponentialfunktion wird also durch das Polynom

$$\sum_{k=0}^{n} \frac{1}{k!} x^k = 1 + x + \frac{1}{2} x^2 + \frac{1}{6} x^3 + \ldots + \frac{1}{n!} x^n$$

approximiert. Der Fehler der Approximation wird gerade durch das Lagrange-Restglied gegeben. Für $|x| \leq 1$ erhalten wir wegen $|\xi| \leq |x| \leq 1$

$$|R_{n+1}(x)| = \frac{e^\xi}{(n+1)!} |x|^{n+1} \leq \frac{e}{(n+1)!}.$$

Wollen wir $e^x$ auf dem Intervall $[-1, 1]$ durch das Polynom bis auf einen Fehler von $10^{-3}$ approximieren, so muß $n \in \mathbb{N}$ so groß gewählt werden, daß

$$\frac{e}{(n+1)!} \leq 10^{-3}.$$

Wir erhalten $n = 6$, denn $e/7! \approx 5 \cdot 10^{-4}$, aber $e/6! \approx 4 \cdot 10^{-3}$. Wir haben gezeigt:

$$\left| e^x - \sum_{k=0}^{6} \frac{1}{k!} x^k \right| \leq 10^{-3} \qquad \text{für alle} \quad |x| \leq 1.$$

□

Die obigen Beispiele legen nahe, den Grenzwert $n \to \infty$ zu versuchen und die Gleichung

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k \tag{6.8}$$

für eine unendlich oft differenzierbare Funktion zu folgern. Gilt die Formel (6.8)?

Im allgemeinen nein! Sei beispielsweise $f(x) = e^{-1/x^2}$ für $x \neq 0$ und $f(x) = 0$ für $x = 0$. Die Funktion $f$ ist für $x \neq 0$ unendlich oft differenzierbar:

$$\begin{aligned} f'(x) &= \frac{2}{x^3} e^{-1/x^2}, \\ f''(x) &= \left( -\frac{6}{x^4} + \frac{4}{x^6} \right) e^{-1/x^2} \\ f'''(x) &= \left( \frac{24}{x^5} - \frac{24}{x^7} - \frac{12}{x^7} + \frac{8}{x^9} \right) e^{-1/x^2} \qquad \text{usw.} \end{aligned}$$

Man kann allgemein zeigen, daß

$$f^{(k)}(x) = p\left(\frac{1}{x}\right) e^{-1/x^2} \qquad \text{für ein geeignetes Polynom } p \text{ und } x \neq 0$$

gilt. Es gilt nach der Regel von de l'Hospital für alle Polynome $p$:

$$\lim_{x \to 0} p\left(\frac{1}{x}\right) e^{-1/x^2} = 0.$$

Wir können also die $k$-te Ableitung von $f$ in $x = 0$ fortsetzen, so daß $f^{(k)}$ stetig ist:

$$f^{(k)}(x) = \begin{cases} p(\frac{1}{x})e^{-1/x^2} & : x \neq 0 \\ 0 & : x = 0. \end{cases}$$

Aus dem Satz von Taylor folgt

$$f(x) = T_n(x) + R_{n+1}(x) \qquad \text{mit} \quad T_n(x) = \sum_{k=0}^{n} \frac{f^{(k)}(0)}{k!} x^k = 0.$$

Also ist

$$f(x) \neq \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} = 0 \qquad \text{für alle} \quad x \neq 0.$$

Unter welchen Bedingungen gilt nun die Formel (6.8) dennoch? Wir vermuten wegen (6.5): wenn $R_{n+1}(x) \to 0$ für $n \to \infty$. Die Vermutung ist richtig, wie der folgende Satz zeigt.

**Satz 6.22** *Sei $f : (a, b) \to \mathbb{R}$ unendlich oft differenzierbar und sei $x_0 \in (a, b)$. Dann gilt*

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k \qquad \textit{für alle} \quad x \in (a, b)$$

*genau dann, wenn $R_n(x) \to 0$ $(n \to \infty)$. Dies ist der Fall, wenn*

$$\exists\, \alpha, C > 0 : \quad \forall\, x \in (a, b),\ k \in \mathbb{N} : \quad |f^{(k)}(x)| \leq \alpha C^k.$$

Man nennt die unendliche Reihe aus Satz 6.22 die *Taylor-Reihe* von $f$. Beachten Sie:

- Die Taylor-Reihe muß im allgemeinen nicht konvergieren!
- Falls die Taylor-Reihe von $f$ konvergiert, so konvergiert sie nicht immer gegen $f$!

Unter den Voraussetzungen von Satz 6.22 konvergiert natürlich die Taylor-Reihe von $f$, und zwar gegen $f$.

**Beispiel 6.23** Sei wie in Beispiel 6.21 (1) $f(x) = e^x$, $x \in \mathbb{R}$, und $x_0 = 0$. Dann gilt

$$|R_{n+1}(x)| = \frac{e^\xi}{(n+1)!}|x|^{n+1} \to 0 \qquad \text{für} \quad n \to \infty,$$

d.h.

$$e^x = \sum_{k=0}^{\infty} \frac{1}{k!} x^k, \qquad x \in \mathbb{R}.$$

Insbesondere erhalten wir die Formel

$$e = \sum_{k=0}^{\infty} \frac{1}{k!} = 1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \frac{1}{4!} + \ldots, \tag{6.9}$$

die schneller konvergiert als die Folge $(1 + 1/n)^n$ mit

$$e = \lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^n.$$

Eine andere interessante Folgerung aus (6.9) ist

$$\left(\sum_{k=0}^{\infty} \frac{1}{k!}\right)^\alpha = e^\alpha = \sum_{k=0}^{\infty} \frac{1}{k!}\alpha^k \qquad \text{für} \quad \alpha \in \mathbb{R}.\square$$

## 6.6 Lineare Regression

In diesem Abschnitt präsentieren wir eine Anwendung der Differentiation, nämlich die lineare Regression oder Methode der kleinsten Quadrate.

Der Druck $p$ eines in einem Behälter mit Volumen $V$ befindlichen Gases werde für verschiedene Temperaturen gemessen. Nach dem idealen Gasgesetz $pV = nRT$ ($n$ Zahl an Molen eines Gases) erwarten wir einen linearen Zusammenhang zwischen Druck und Temperatur. Die Berechnung der Steigung der Geraden $p = p(T)$ erlaubt die Bestimmung der Gaskonstanten $R$. Allerdings werden die Meßwerte nicht alle auf einer Geraden liegen; das Problem lautet, die mit Fehlern behafteten Meßdaten durch eine Gerade zu approximieren. Eine Idee ist durch die *Methode der kleinsten Quadrate* gegeben.

Seien

$$(x_1, y_1), \ldots, (x_m, y_m)$$

$m$ Meßwertpaare (z.B. $x_i$ = Temperatur, $y_i$ = Druck) gegeben. Wir suchen die Koeffizienten $a_0$ und $a_1$ der Geraden

$$g(x) = a_0 + a_1 x,$$

so daß der quadratische Abstand zwischen den Meßwerten $y_i$ und den Punkten $g(x_i)$ auf der Geraden minimiert wird (siehe Abbildung 6.9): Minimiere

$$F(a_0, a_1) = \sum_{i=1}^{m} (g(x_i) - y_i)^2 = \sum_{i=1}^{m} (a_0 + a_1 x_i - y_i)^2.$$

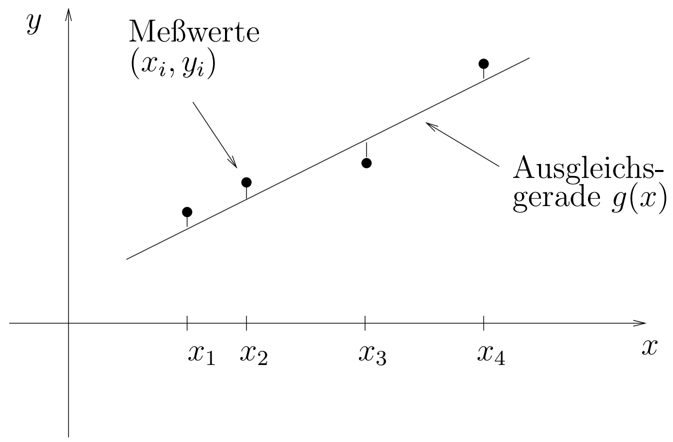


Abbildung 6.9: Meßwertpaare und Ausgleichsgerade.

Zur Minimierung von $F : \mathbb{R}^2 \to \mathbb{R}$ bestimmen wir gemäß Abschnitt 6.4 die kritischen Punkte, d.h., wir differenzieren $F$ nach $a_0$ und nach $a_1$ und setzen die Ableitungen gleich Null. Wir führen hierfür die Funktionen

$$\begin{aligned} F_1(a_0) &= F(a_0, a_1) \qquad (a_1 \text{ fest gehalten}) \text{ und} \\ F_2(a_1) &= F(a_0, a_1) \qquad (a_0 \text{ fest gehalten}) \end{aligned}$$

ein. Wir erhalten:

$$\begin{aligned} 0 &= F_1'(a_0) = 2\sum_{i=1}^{m}(a_0 + a_1 x_i - y_i), \\ 0 &= F_2'(a_1) = 2\sum_{i=1}^{m}(a_0 + a_1 x_i - y_i)x_i, \end{aligned}$$

also

$$\begin{aligned} m a_0 + \left(\sum_{i=1}^{m} x_i\right) a_1 &= \sum_{i=1}^{m} y_i, \\ \left(\sum_{i=1}^{m} x_i\right) a_0 + \left(\sum_{i=1}^{m} x_i^2\right) a_1 &= \sum_{i=1}^{m} x_i y_i. \end{aligned}$$

Dies ist ein lineares Gleichungssystem für die beiden Variablen $a_0$ und $a_1$ mit Lösung

$$a_0 = \frac{\sum y_i \sum x_i^2 - \sum x_i \sum x_i y_i}{m\sum x_i^2 - (\sum x_i)^2}, \qquad a_1 = \frac{m\sum x_i y_i - \sum x_i \sum y_i}{m\sum x_i^2 - (\sum x_i)^2}. \tag{6.10}$$

Die Lösung existiert genau dann, wenn

$$m\sum_{i=1}^{m} x_i^2 \neq \left(\sum_{i=1}^{m} x_i\right)^2.$$

Dies ist der Fall, wenn $m \geq 2$ (mindestens zwei Meßwerte) und es mindestens zwei Werte $x_i \neq x_j$ gibt.

**Beispiel 6.24** Der Druck eines idealen Gases in einem Behälter mit festem Volumen werde für verschiedene Temperaturen gemessen. Die $m = 5$ Meßwerte lauten:

| Temperatur $x_i$ in $K$ | 300 | 350 | 400 | 450 | 500 |
|---|---|---|---|---|---|
| Druck $y_i$ in $\mathrm{N/m^2}$ | 235 | 290 | 335 | 360 | 415 |

Wir nehmen weiter an, daß der Behälter $n = 10^{-4}$ mol des Gases enthalte und daß das Volumen $V = 10^{-3}\mathrm{m}^3$ betrage. Dann ist nach dem idealen Gasgesetz

$$p(T) = \frac{nR}{V}T = \left(0.1R\frac{\mathrm{mol}}{\mathrm{m}^3}\right)T.$$

Der Achsenabschnitt $a_0$ und die Steigung der Ausgleichsgeraden lauten gemäß (6.10)

$$a_0 = -17\ [\mathrm{N/m^2}], \qquad a_1 = 0.860\ [\mathrm{N/m^2}K]$$

(siehe Abbildung 6.10). Damit erhalten wir aus

$$p(300\,\mathrm{K}) = 30R\frac{\mathrm{mol\,K}}{\mathrm{m}^3} \approx 0,86 \cdot 300\,\frac{\mathrm{N}}{\mathrm{m}^2} - 17\,\frac{\mathrm{N}}{\mathrm{m}^2} = 241\,\frac{\mathrm{N}}{\mathrm{m}^2}$$

eine Approximation der Gaskonstanten

$$R \approx \frac{241}{30}\,\frac{\mathrm{Nm}}{\mathrm{mol\,K}} \approx 8,03\frac{\mathrm{Nm}}{\mathrm{mol\,K}}.$$

Der Literaturwert lautet $R = 8,314\,\mathrm{Nm/(mol\,K)}$. Die Abweichung liegt mit einem relativen Fehler von $(8,03 - 8,314)/8,314 \approx 3,4\%$ innerhalb der zu erwartenden Meßgenauigkeit. $\square$

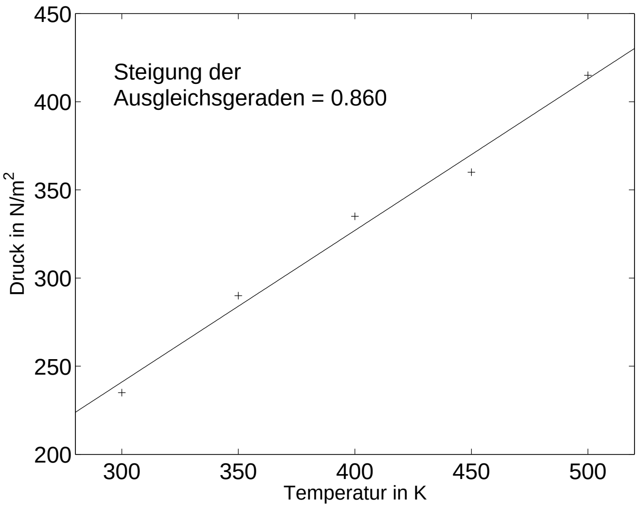


Abbildung 6.10: Meßwerte aus Beispiel 6.24 und zugehörige Ausgleichswerte.

# 7 Integration

## 7.1 Das bestimmte Integral

In Beispiel 1.2 haben wir gesehen, daß die Menge $N(t)$ einer radioaktiven Substanz die Gleichung

$$N'(t) = -\alpha N(t) \qquad \text{mit} \quad \alpha > 0,\ t > 0$$

erfüllt. Um derartige Differentialgleichungen nach $N(t)$ "auflösen" zu können, benötigen wir die "Umkehroperation" zur Differentiation. Interessanterweise wird diese durch das *Integral* gegeben, das mit der Fläche unter einer Kurve zusammenhängt.

Wir wollen also die Fläche zwischen einer Kurve $y = f(x)$ und der $x$-Achse in $[a, b]$ berechnen. Dazu approximieren wir die Fläche durch Rechtecke (siehe Abbildung 7.1). Wir zerlegen das Intervall $[a, b]$ in $n$ Teilintervalle, indem wir die Punkte $a = x_0 < x_1 < \ldots < x_n = b$ bzw. die *Zerlegung* $Z = \{x_0, x_1, \ldots, x_n\}$ einführen, und wählen in jedem Teilintervall eine Zwischenstelle $\xi_i \in [x_{i-1}, x_i]$, $i = 1, \ldots, n$. Die *Riemann-Summe*

$$S(Z) = \sum_{i=1}^{n} f(\xi_i)(x_i - x_{i-i})$$

ist dann eine Näherung an die Fläche zwischen der Kurve und der $x$-Achse. Wir erwarten, daß die Näherung umso besser wird, je feiner die Zerlegung gewählt wird.

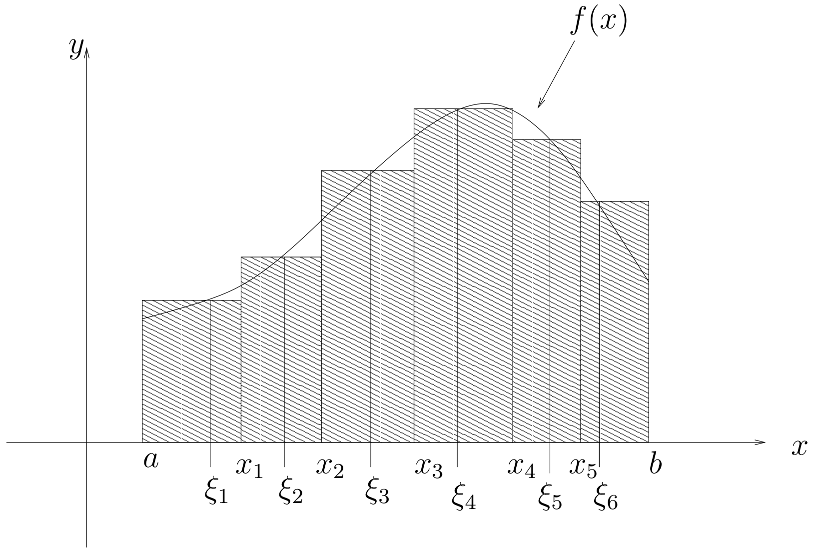


Abbildung 7.1: Zur Definition der Riemann-Summe.

**Definition 7.1** *Sei $f : [a, b] \to \mathbb{R}$ eine Funktion. Wenn*

*(1) für* jede *Zerlegung $Z$ von $[a,b]$ mit* $\max\{x_i - x_{i-1} : i = 1, \ldots, n\} \to 0$ *($n \to \infty$) und*

*(2) für* jede *Wahl von Zwischenstellen $\xi_i \in [x_{i-1}, x_i]$, $i = 1, \ldots, n$,*

*der Grenzwert $F = \lim_{n\to\infty} S(Z)$ existiert und stets denselben Wert liefert, so heißt $f$* (Riemann-)integrierbar *in $[a,b]$ und $F$ heißt das* bestimmte Integral *von $f$ in $[a,b]$. Wir schreiben kurz:*

$$\lim_{n\to\infty} \sum_{i=1}^{n} f(\xi_i)(x_i - x_{i-1}) = \int_a^b f(x)\,dx.$$

**Satz 7.2** *(1) Sei $f : [a,b] \to \mathbb{R}$ stetig. Dann ist $f$ integrierbar.*

*(2) Sei $f : [a,b] \to \mathbb{R}$ integrierbar. Dann ist*

$$\int_a^b f(x)\,dx = \text{(Fläche oberhalb } x\text{-Achse)} - \text{(Fläche unterhalb } x\text{-Achse)},$$

*siehe Abbildung 7.2.*

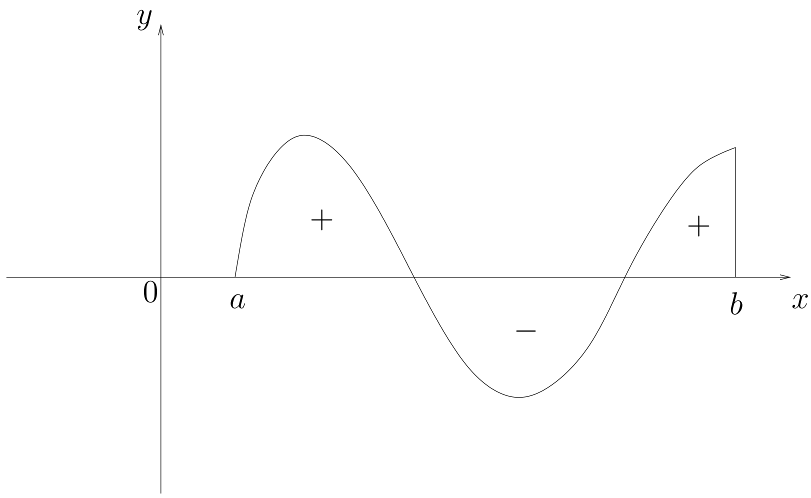


Abbildung 7.2: Veranschaulichung von Satz 7.2 (2).

**Beispiel 7.3** (1) Die Funktion $f(x) = 1/x$ ist für $x > 0$ stetig, also insbesondere auf $[1,2]$ integrierbar. Um das Integral

$$\int_1^2 \frac{dx}{x}$$

zu berechnen, wählen wir

- Zerlegung $Z$: $x_i = 2^{i/n}$, $i = 0, \ldots, n$,

- Zwischenstellen: $\xi_i = x_{i-1} = 2^{(i-1)/n}$, $i = 1, \ldots, n$.

Die Riemann-Summe lautet

$$S(Z) = \sum_{i=1}^{n} f(\xi_i)(x_i - x_{i-1}) = \sum_{i=1}^{n} 2^{-(i-1)/n}\left(2^{i/n} - 2^{(i-1)/n}\right) = \sum_{i=1}^{n}\left(2^{1/n} - 1\right) = n(2^{1/n} - 1).$$

Wir erhalten mit der Regel von de l'Hospital

$$\lim_{n\to\infty} S(Z) = \lim_{n\to\infty} \frac{2^{1/n} - 1}{1/n} = \lim_{x\to 0} \frac{2^x - 1}{x} = \lim_{x\to 0} \frac{e^{x\ln 2} - 1}{x} = \lim_{x\to 0} \frac{\ln 2 \cdot e^{x \ln 2}}{1} = \ln 2.$$

Man kann zeigen, daß man diesen Grenzwert für jede Zerlegung und jede Wahl von Zwischenstellen erhält. Daher gilt

$$\int_1^2 \frac{dx}{x} = \ln 2.$$

(2) Wir behaupten, daß die Funktion

$$f(x) = \begin{cases} 1 & : \; x \in \mathbb{Q} \cap [0,1] \\ 0 & : \; x \in (\mathbb{R}\backslash\mathbb{Q}) \cap [0,1] \end{cases}$$

nicht integrierbar ist. Sei $Z = \{x_0, \ldots, x_n\}$ eine beliebige Zerlegung und wähle

- $\xi_i \in [x_{i-1}, x_i]$ mit $\xi_i \in \mathbb{Q}$ $(i = 1, \ldots, n)$,
- $\eta_i \in [x_{i-1}, x_i]$ mit $\eta_i \in \mathbb{R}\backslash\mathbb{Q}$ $(i = 1, \ldots, n)$.

Wir erhalten

$$S_\xi(Z) = \sum_{i=1}^{n} f(\xi_i)(x_i - x_{i-1}) = \sum_{i=1}^{n}(x_i - x_{i-1}) = 1,$$
$$S_\eta(Z) = \sum_{i=1}^{n} f(\eta_i)(x_i - x_{i-1}) = 0.$$

Wählen wir die Zerlegung immer feiner und die Zwischenwerte wie oben, so folgt

$$\lim_{n\to\infty} S_\xi(Z) = 1 \neq 0 = \lim_{n\to\infty} S_\eta(Z),$$

d.h., $f$ ist nicht integrierbar. □

Es gelten die folgenden Rechenregeln.

**Satz 7.4** *Seien $f$ und $g$ auf $[a,b]$ integrierbar. Dann gilt:*

*(1)* $\int_a^b (\alpha f(x) + \beta g(x))\, dx = \alpha \int_a^b f(x)\, dx + \beta \int_a^b g(x)\, dx$ *für alle* $\alpha, \beta \in \mathbb{R}$.

*(2)* $\int_a^b f(x)\, dx = \int_a^c f(x)\, dx + \int_c^b f(x)\, dx$ *für alle* $c \in [a,b]$.

*(3)* $\int_a^a f(x)\, dx = 0$ *und* $\int_a^b f(x)\, dx = -\int_b^a f(x)\, dx$.

*(4)* $|\int_a^b f(x)\, dx| \leq \int_a^b |f(x)|\, dx$.

## 7.2 Zusammenhang zwischen Differentiation und Integration

**Definition 7.5** *Sei $f : [a,b] \to \mathbb{R}$ eine Funktion. Eine Funktion $F : [a,b] \to \mathbb{R}$ heißt* Stammfunktion *von $f$ genau dann, wenn $F$ differenzierbar ist und*

$$F'(x) = f(x) \qquad \text{für alle } x \in [a,b]$$

*gilt. Man nennt $F$ auch ein* unbestimmtes Integral *von $f$ und schreibt*

$$F(x) = \int f(x)\,dx. \tag{7.1}$$

Die Notation (7.1) ist etwas mißverständlich: Sind nämlich $F$ und $G$ zwei Stammfunktionen von $f$, so folgt aus $(F-G)' = F' - G' = f - f = 0$ in $[a,b]$, daß $F - G = \text{const.}$ in $[a,b]$ gelten muß. In der Notation (7.1) ist also die Integrationskonstante stets mit aufgeführt!

Einige Stammfunktionen lauten (mit $c \in \mathbb{R}$):

| $f(x)$ | $x^\alpha$ $(\alpha \neq -1)$ | $\frac{1}{x}$ | $e^x$ | $\sin x$ | $\cos x$ |
|---|---|---|---|---|---|
| $\int f(x)\,dx$ | $\frac{1}{\alpha+1}x^{\alpha+1} + c$ | $\ln\lvert x\rvert + c$ | $e^x + c$ | $-\cos x + c$ | $\sin x + c$ |

Ist $f : [a,b] \to \mathbb{R}$ eine stetige Funktion, so können wir für beliebiges $x_0 \in [a,b]$ mit

$$F(x) = \int_{x_0}^{x} f(t)\,dt, \qquad x \in [a,b], \tag{7.2}$$

eine Stammfunktion konstruieren, was den Zusammenhang zwischen bestimmten und unbestimmten Integralen aufweist. Es gilt:

**Satz 7.6** *Sei $f : [a,b] \to \mathbb{R}$ stetig, $x_0 \in [a,b]$. Dann ist $F$, definiert in (7.2), differenzierbar und $F' = f$, d.h., $F$ ist eine Stammfunktion von $f$.*

Den Zusammenhang zwischen der Differentiation und Integration stellt der folgende Satz her:

**Satz 7.7** (Hauptsatz der Differential- und Integralrechnung)
*Sei $f : [a,b] \to \mathbb{R}$ stetig und $F$ eine Stammfunktion von $f$. Dann gilt*

$$\int_a^b f(x)\,dx = [F(x)]_a^b = F(b) - F(a).$$

Der Hauptsatz zeigt, daß Differentiation und Integration sich im folgenden Sinne wie “Operation” und “Umkehroperation” verhalten. Sei $f : [a,b] \to \mathbb{R}$ stetig und $F$ eine (differenzierbare) Stammfunktion von $f$. Dann gilt nach Satz 7.6

$$\frac{d}{dx}\int f(x)\,dx = \frac{dF}{dx}(x) = f(x) \qquad \text{oder} \qquad \frac{d}{dx}\int_{x_0}^{x} f(t)\,dt = \frac{dF}{dx}(x) = f(x)$$

und nach Satz 7.7

$$\int_a^b \frac{dF}{dx}(x) = F(b) - F(a).$$

**Beispiel 7.8** Die Fläche zwischen der Sinuskurve und der $x$-Achse in $[0, \pi]$ lautet

$$\int_0^\pi \sin x\,dx = -\cos\pi - (-\cos 0) = 1 + 1 = 2,$$

denn $-\cos x$ ist eine Stammfunktion zu $\sin x$. Das Integral

$$\int_0^{2\pi} \sin x\,dx = -\cos(2\pi) - (-\cos 0) = -1 + 1 = 0$$

verschwindet dagegen, da die Fläche unterhalb der $x$-Achse mit negativem Vorzeichen versehen wird (siehe Abbildung 7.2). □

**Beispiel 7.9** Mit Hilfe von Satz 7.7 können wir auch Werte von unendlichen Reihen berechnen. Wir zeigen, daß

$$\sum_{n=1}^\infty \frac{\alpha^n}{n} = -\ln(1-\alpha) \qquad \text{für alle } |\alpha| < 1 \tag{7.3}$$

gilt. Definiere dazu

$$f(x) = \sum_{n=1}^\infty \frac{x^n}{n} = \sum_{k=0}^\infty \frac{x^{k+1}}{k+1}.$$

Man kann zeigen, daß $f$ für $|x| < 1$ differenzierbar ist und daß folgt

$$f'(x) = \sum_{k=0}^\infty \frac{d}{dx}\frac{x^{k+1}}{k+1} = \sum_{k=0}^\infty x^k, \qquad |x| < 1.$$

Das ist aber gerade die geometrische Reihe, so daß $f'(x) = 1/(1-x)$. Aus Satz 7.7 erhalten wir

$$f(x) = f(x) - f(0) = \int_0^x f'(t)\,dt = \int_0^x \frac{dt}{1-t} = \Big[-\ln(1-t)\Big]_0^x = -\ln(1-x).$$

Setzen wir $x = \alpha$, so ist (7.3) bewiesen. □

## 7.3 Integrationsmethoden

Wir stellen zwei Integrationsmethoden vor:

- Partielle Integration: Seien $f, g : [a, b] \to \mathbb{R}$ stetig differenzierbar. Dann ist

$$\int_a^b f(x)g'(x)\,dx = \big[f(x)g(x)\big]_a^b - \int_a^b f'(x)g(x)\,dx.$$

- Substitution: Sei $f : [a,b] \to \mathbb{R}$ stetig und $g : [c,d] \to [a,b]$ stetig differenzierbar. Dann ist
$$\int_c^d f(g(t))g'(t)\,dt = \int_{g(c)}^{g(d)} f(x)\,dx.$$

**Beispiel 7.10** (Partielle Integration)
(1) Berechne $\int x \sin x\,dx$. Wir setzen $f(x) = x$ und $g'(x) = \sin x$. Dann ist $f'(x) = 1$, $g(x) = -\cos x$ und
$$\int x \sin x\,dx = x \cdot (-\cos x) - \int 1 \cdot (-\cos x)\,dx = -x\cos x + \sin x + c.$$

(2) Berechne $\int_0^\pi \sin^2 x\,dx$. Wir setzen $f(x) = \sin x$ und $g'(x) = \sin x$, also
$$\int_0^\pi \sin^2 x\,dx = \Big[\sin x(-\cos x)\Big]_0^\pi - \int_0^\pi \cos x(-\cos x)\,dx = \int_0^\pi \cos^2 x\,dx.$$
Wir könnten noch einmal partiell integrieren:
$$\int_0^\pi \sin^2 x\,dx = \int_0^\pi \cos^2 x\,dx = \Big[\cos x \cdot \sin x\Big]_0^\pi + \int_0^\pi \sin^2 x\,dx = \int_0^\pi \sin^2 x\,dx,$$
doch dies führt nicht weiter. Wir benutzen anstelle dessen die Relation $\cos^2 x = 1 - \sin^2 x$. Dann folgt
$$\int_0^\pi \sin^2 x\,dx = \int_0^\pi dx - \int_0^\pi \sin^2 x\,dx$$
und daher
$$\int_0^\pi \sin^2 x\,dx = \frac{\pi}{2}.$$

(3) Wie in (2) können wir das unbestimmte Integral $\int \sin^2 x\,dx$ berechnen:
$$\begin{aligned}\int \sin^2 x\,dx &= -\sin x \cos x + \int \cos^2 x\,dx = -\sin x\cos x + \int (1 - \sin^2 x)\,dx \\ &= -\sin x \cos x + x - \int \sin^2 x\,dx,\end{aligned}$$
also
$$\int \sin^2 x\,dx = \frac{1}{2}(x - \sin x \cos x). \tag{7.4}$$
□

Die Substitutionsregel kann in der Praxis folgendermaßen verwendet werden. In dem Integral $\int_a^b f(x)\,dx$ substituiert man $x = x(t)$. Dann ist $\frac{dx}{dt} = x'(t)$, also kann man "$dx$" durch "$x'(t)dt$" beim Übergang von der $x$- zur $t$-Variablen ersetzen. Falls $x(\alpha) = a$ und $x(\beta) = b$, so erhält man
$$\int_a^b f(x)\,dx = \int_\alpha^\beta f(x(t))x'(t)\,dt.$$

**Beispiel 7.11** (Substitution)
(1) Berechne $\int_0^1 (1+2x)^{10}\,dx$. Wir setzen $t = 1+2x$. Dann ist $x(t) = x = \frac{1}{2}(t-1)$, also $\frac{dx}{dt} = \frac{1}{2}$, und aus $0 \le x \le 1$ folgt $1 \le t \le 3$. Damit ergibt sich

$$\int_0^1 (1+2x)^{10}\,dx = \int_1^3 t^{10}\frac{1}{2}\,dt = \frac{1}{2}\frac{1}{11}\left[t^{11}\right]_1^3 = \frac{1}{22}(3^{11}-1) = \frac{177146}{22}.$$

(2) Berechne $\int \sqrt{1-x^2}\,dx$. Wir substituieren $x = \cos t$. Dann erhalten wir $\frac{dx}{dt} = -\sin t$ und wegen (7.4)

$$\int \sqrt{1-x^2}\,dx = \int \sqrt{1-\cos^2 t}\,(-\sin t)\,dt = -\int \sin^2 t\,dt = -\frac{1}{2}(t - \sin t\cos t) + c.$$

Rücksubstitution $t = \arccos x$ und $\sin t = \sqrt{1-\cos^2 t} = \sqrt{1-x^2}$ ergeben

$$\int \sqrt{1-x^2}\,dx = \frac{1}{2}\left(x\sqrt{1-x^2} - \arccos x\right) + c.$$

Das bestimmte Integral

$$\int_0^1 \sqrt{1-x^2}\,dx$$

ist gerade der Flächeninhalt des Viertelkreises mit Radius $r = 1$. Wir erhalten

$$\begin{aligned}\int_0^1 \sqrt{1-x^2}\,dx &= \left[\frac{1}{2}\left(x\sqrt{1-x^2} - \arccos x\right)\right]_0^1 = -\frac{1}{2}\arccos 1 + \frac{1}{2}\arccos 0 \\ &= -\frac{1}{2}\cdot 0 + \frac{1}{2}\cdot\frac{\pi}{2} = \frac{\pi}{4}.\end{aligned}$$

Dies stimmt mit der Formel für den Flächeninhalt $\pi r^2/4 = \pi/4$ überein. □

Für einige Integrale gibt es Standardsubstitutionen, die zuweilen auf einfachere Integrale führen. Im folgenden geben wir einige Beispiele. Sei hierfür $R(\alpha, v, w, \ldots)$ eine rationale Funktion in den angegebenen Variablen.

- Typ $\int R(\cos x)\sin x\,dx$. Substituiere $t = \cos x$;
- Typ $\int R(\sin x)\cos x\,dx$. Substituiere $t = \sin x$;
- Typ $\int R(\sin x, \cos x)\,dx$. Substituiere $t = \tan\frac{x}{2}$;
- Typ $\int R(x, \sqrt{x^2+a^2})\,dx$. Substituiere $x = a\sinh t$;
- Typ $\int R(x, \sqrt{x^2-a^2})\,dx$. Substituiere $x = a\cosh t$;
- Typ $\int R(x, \sqrt{a^2-x^2})\,dx$. Substituiere $x = a\sin t$;

**Beispiel 7.12** (Standardsubstitutionen)
(1) Berechne $\int_0^{\pi/2} \frac{\cos^3 x}{1-\sin x}\,dx$. Wegen

$$\int_0^{\pi/2} \frac{\cos^3 x}{1-\sin x}\,dx = \int_0^{\pi/2} \frac{1-\sin^2 x}{1-\sin x}\cos x\,dx$$

ist das Integral vom Typ $\int R(\sin x)\cos x\,dx$ mit $R(u) = \frac{1-u^2}{1-u}$. Wir substituieren also $t = \sin x$, Dann ist $\frac{dt}{dx} = \cos x$, und $0 \le x \le \frac{\pi}{2}$ impliziert $0 \le t \le 1$:

$$\int_0^{\pi/2} \frac{\cos^3 x}{1-\sin x}\,dx = \int_0^1 \frac{1-t^2}{1-t}\,dt = \int_0^1 \frac{(1-t)(1+t)}{1-t}\,dt = \int_0^1 (1+t)\,dt = \frac{3}{2}.$$

(2) Das Integral

$$\int \frac{\sin^2 x}{(1+\sin x+\cos x)^3}\,dx$$

ist vom Typ $\int R(\sin x, \cos x)\,dx$, so daß wir $t = \tan(x/2)$ substituieren. Dann folgt aus $x = 2\arctan t$ zum einen $\frac{dx}{dt} = \frac{2}{1+t^2}$ und mit den Additionstheoremen $\cos(2\alpha) = \cos^2\alpha - \sin^2\alpha$ und $\sin(2\alpha) = 2\sin\alpha\cos\alpha$ zum anderen

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

Wir erhalten

$$1+\sin x+\cos x = \frac{1+t^2+2t+1-t^2}{1+t^2} = \frac{2(1+t)}{1+t^2}$$

und daher

$$\int \frac{\sin^2 x}{(1+\sin x+\cos x)^3}\,dx = \int \left(\frac{2t}{1+t^2}\right)^2 \left(\frac{1+t^2}{2(1+t)}\right)^3 \frac{2dt}{1+t^2} = \int \frac{t^2}{(1+t)^3}\,dt. \qquad (7.5)$$

Wie kann man nun das Integral über eine rationale Funktion berechnen? Wir verschieben die Integration von (7.5) auf Beispiel 7.13 und betrachten allgemein Integrale rationaler Funktionen. $\square$

Die Aufgabe lautet, für gegebene Polynome $P(x)$ und $Q(x)$ das Integral

$$\int \frac{P(x)}{Q(x)}\,dx$$

zu berechnen. Dabei gehen wir wie folgt vor:

- *1. Schritt:* Führe *Polynomdivision* durch, falls Grad $P \geq$ Grad $Q$ ("Grad" sei der Grad eines Polynoms):
$$\frac{P(x)}{Q(x)} = P_1(x) + \frac{R(x)}{Q(x)},$$
wobei $P_1$ und $R$ Polynome sind mit Grad $R <$ Grad $Q$. Das Integral über $P_1$ ist leicht zu berechnen.

- *2. Schritt:* Bestimme alle *Nullstellen* von $Q$ und zerlege $Q$ in Faktoren der Form $(x-a)^m$ und $((x-a)^2+b^2)^n$.

- *3. Schritt:* Führe *Partialbruchzerlegung* durch. Verwende dazu für den Faktor in $Q$:
$$(x-a)^m \qquad \text{den Ansatz} \quad \frac{A_1}{x-a} + \cdots + \frac{A_m}{(x-a)^m}$$
und für
$$((x-a)^2+b^2)^n \qquad \text{den Ansatz} \quad \frac{A_1 x + B_1}{(x-a)^2+b^2} + \cdots + \frac{A_n x + B_n}{((x-a)^2+b^2)^n}.$$
Die Koeffizienten $A_i$ und $B_i$ werden wie in Abschnitt 3.2 durch Koeffizientenvergleich bestimmt.

- *4. Schritt:* Schließlich können die Partialbrüche *integriert* werden. Verwende hierzu die Formeln
$$\begin{aligned}
\int \frac{dx}{x-a} &= \ln|x-a| + c, \\
\int \frac{dx}{(x-a)^m} &= -\frac{1}{m-1}\,\frac{1}{(x-a)^{m-1}} + c \qquad (m \geq 2), \\
\int \frac{Ax+B}{(x-a)^2+b^2}\,dx &= \frac{A}{2}\ln|(x-a)^2+b^2| + \frac{Aa+b}{2}\arctan\frac{x-a}{b} + c.
\end{aligned}$$
Für Integrale mit dem Nenner $((x-a^2)+b^2)^n$ mit $n \geq 2$ kann man mit partieller Integration eine Rekursionsformel herleiten, die wir hier jedoch nicht angeben.

**Beispiel 7.13** (Partialbruchzerlegung)
Wir setzen Beispiel 7.12 (2) fort und berechnen
$$\int \frac{t^2}{(1+t)^3}\,dt.$$
Der erste Schritt entfällt, da Grad $(t^2) = 2 < 3 =$ Grad $((1+t)^3)$. Das Nennerpolynom ist bereits in Faktoren aufgelöst. Wir führen nun eine Partialbruchzerlegung (3. Schritt) durch und machen den Ansatz
$$\frac{t^2}{(1+t)^3} = \frac{A_1}{1+t} + \frac{A_2}{(1+t)^2} + \frac{A_3}{(1+t)^3}.$$

Multiplikation mit $(1+t)^3$ führt auf

$$t^2 = (1+t)^2 A_1 + (1+t)A_2 + A_3 = A_1 t^2 + (2A_1 + A_2)t + (A_1 + A_2 + A_3).$$

Koeffizientenvergleich ergibt

$$A_1 = 1, \quad 2A_1 + A_2 = 0, \quad A_1 + A_2 + A_3 = 0,$$

Also $A_2 = -2$, $A_3 = 1$. Folglich ist

$$\frac{t^2}{(1+t)^3} = \frac{1}{1+t} - \frac{2}{(1+t)^2} + \frac{1}{(1+t)^3}.$$

Damit ist (4. Schritt)

$$\begin{aligned}\int \frac{t^2}{(1+t)^3}\,dt &= \int \frac{dt}{1+t} - 2\int \frac{dt}{(1+t)^2} + \int \frac{dt}{(1+t)^3} = \ln|1+t| + \frac{2}{1+t} - \frac{1}{2(1+t)^2} \\ &= \ln|1+t| + \frac{4t+3}{2(1+t)^2}.\end{aligned}$$

Das Integral aus Beispiel 7.12 (2) ist also gelöst:

$$\int \frac{\sin^2 x}{(1+\sin x+\cos x)^3}\,dx = \ln|1+t| + \frac{4t+3}{2(1+t)^2} = \ln|1+\tan\frac{x}{2}| + \frac{4\tan(x/2)+3}{2(1+\tan(x/2))^2}.$$

□

Das Integral

$$\int e^{-x^2/2}\,dx$$

besitzt keine durch elementare Funktionen ausdrückbare Stammfunktion. Das bestimmte Integral

$$\int_{-\infty}^{\infty} e^{-x^2/2}\,dx$$

muß also numerisch bestimmt werden. Doch was bedeutet der unbeschränkte Integrationsbereich? Wir definieren einfach

$$\int_{-\infty}^{\infty} e^{-x^2/2}\,dx = \lim_{R\to\infty} \int_{-R}^{R} e^{-x^2/2}\,dx$$

und betrachten zunächst das Integral über $[-R,R]$. Eine Möglichkeit ist es, das Integral wie in Abbildung 7.3 durch die eingezeichneten Trapeze zu approximieren. Sei dafür $-R = x_{-n} < x_{-n+1} < \ldots < x_n = R$ eine Zerlegung des Intervalls $[-R,R]$. Dann ist

$$\int_{-R}^{R} e^{-x^2/2}\,dx \approx \sum_{k=-n+1}^{n} \frac{1}{2}\left(e^{-x_k^2/2} + e^{-x_{k-1}^2/2}\right)(x_k - x_{k-1}).$$

Man nennt dies die Trapezregel. Wir wählen nun speziell $x_k = k$ und bezeichnen mit $a_n$ die obige Summe. Dann erhalten wir die folgende Tabelle:

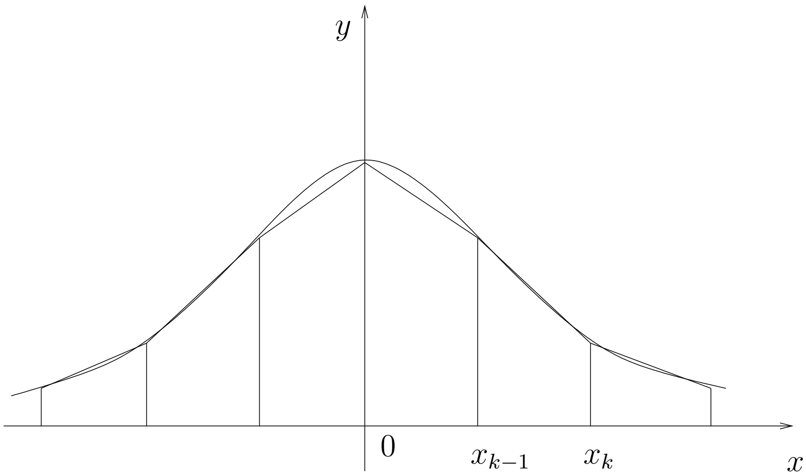


Abbildung 7.3: Zur Approximation von $\int_{-\infty}^{\infty} e^{-x^2/2}\,dx$.

| $n$ | $a_n$ |
|---|---|
| 3 | 2.49484088243673 |
| 5 | 2.50662453088395 |
| 7 | 2.50662828801998 |
| 9 | 2.50662828804291 |
| 11 | 2.50662828804291 |

Wir vermuten, daß

$$\int_{-\infty}^{\infty} e^{-x^2/2}\,dx \approx 2.50662828804291.$$

Wie genau ist dieser Wert? Wir haben zwei Arten von Fehlern zu berücksichtigen. Zum einen haben wir das Integral über $(-\infty, -n)$ und $(n, \infty)$ vernachlässigt; dieser Fehler wird immer kleiner, wenn wir $n$ groß genug wählen. Allerdings approximieren die Trapeze über $(-1, 0)$ und $(0, 1)$ das Integral mit einem Fehler, der *nicht* kleiner wird, wenn $n \to \infty$. Tatsächlich ist der oben gerechnete Wert nur bis auf sechs Nachkommastellen genau; man kann nämlich zeigen, daß

$$\int_{-\infty}^{\infty} e^{-x^2/2}\,dx = \sqrt{2\pi} \approx 2.50662827463100.$$