Skriptum

Mathematik A (EEE)

Institut für Diskrete Mathematik
Technische Universität Graz

505.039

Version 20. Januar 2025

Inhaltsverzeichnis

Kapitel 1

Grundlagen

1.1 Mengen und Zahlen

Im Zuge dieser Lehrveranstaltung werden verschiedene *Zahlenbereiche* (ganze Zahlen, rationale Zahlen, reelle Zahlen, komplexe Zahlen) relevant sein. Zahlen und andere Objekte werden häufig zu *Mengen* zusammengefasst. In diesem ersten Kapitel sollen einige grundlegende Begriffe und Konzepte in diesem Zusammenhang besprochen werden.

1.1.1 Mengen und Elemente

Eine *Menge* ist eine Zusammenfassung von Objekten, die die *Elemente* der Menge genannt werden. Für uns werden hauptsächlich Zahlenmengen oder geometrisch definierte Mengen von Bedeutung sein. Die Elemente einer Menge können jedoch grundsätzlich recht beliebiger Natur sein. Mengen können durch Auflistung der Elemente in geschwungenen Klammern spezifiziert werden, zum Beispiel

$$A = \{1, 3, 5, 7\}, \qquad B = \{\text{Esel, Hund, Katze, Hahn}\}, \qquad C = \{\text{Rot, Grün, Blau}\}.$$

Eine andere Möglichkeit ist die Schreibweise

$$\{x : \cdots\} \text{ oder } \{x \mid \cdots\},$$

sprich: „die Menge aller x für die gilt, dass ...“. Man schreibt $x \in A$ für „x ist Element von A“, und $x \notin A$ für „x ist nicht Element von A“. Also etwa (mit obigen Mengen):

$$3 \in A, \qquad \text{Esel} \in B, \qquad 6 \notin A, \qquad \text{Gelb} \notin C.$$

Beispiel 1.1. Die Menge $A = \{1, 3, 5, 7\}$ ließe sich auch (zum Beispiel) wie folgt ausdrücken:

$$A = \{x : x \text{ ist eine ungerade ganze Zahl und } 0 < x < 8\}.$$

Zwei Mengen sind gleich, wenn sie die gleichen Elemente haben. Die Reihenfolge spielt dabei keine Rolle, es ist also etwa $\{1, 3, 5, 7\} = \{7, 3, 1, 5\}$. Die Menge ohne Elemente heißt *leere Menge* und wird auch als $\{\}$ oder $\emptyset$ geschrieben. Wenn eine Menge A alle Elemente von B enthält (und möglicherweise noch weitere), so sagt man, dass B eine *Teilmenge* von A ist, geschrieben $B \subseteq A$. So gilt etwa $\{1, 5\} \subseteq \{1, 3, 5, 7\}$. Will man ausdrücken, dass zudem $B \neq A$ (echte Teilmenge), so kommt auch die Schreibweise $B \subsetneq A$ oder $B \subset A$ zum Einsatz[1].

[1] Verwirrenderweise gibt es aber auch die Konvention, bei der $B \subset A$ auch $B = A$ erlaubt.

1.1.2 Vereinigung, Durchschnitt und Komplement

Es gibt einige wichtige Operationen, die man mit zwei oder mehreren Mengen durchführen kann. Zu diesen gehören *Vereinigung* und *Durchschnitt*.

Definition 1.2. Die Vereinigung von zwei oder mehreren Mengen besteht aus allen Elementen, die in zumindest einer der Mengen vorkommen. Man schreibt $A \cup B$ für die Vereinigung von A und B, und allgemeiner $A_1 \cup A_2 \cup \cdots \cup A_k$ für die Vereinigung von Mengen $A_1, A_2, \ldots, A_k$.

Die Schnittmenge (Durchschnitt) von zwei oder mehreren Mengen besteht aus allen Elementen, die in allen Mengen gleichzeitig vorkommen. Man schreibt $A \cap B$ für die Schnittmenge von A und B, und allgemeiner $A_1 \cap A_2 \cap \cdots \cap A_k$ für die Schnittmenge von Mengen $A_1, A_2, \ldots, A_k$.

Beispiel 1.3. Für $A = \{1, 2, 3, 4\}$, $B = \{1, 3, 5, 7, 9\}$ und $C = \{3, 6, 9\}$ gilt etwa $A \cup B = \{1, 2, 3, 4, 5, 7, 9\}$, $A \cap B = \{1, 3\}$, sowie

$$A \cap B \cap C = \{3\}, \qquad A \cup B \cup C = \{1, 2, 3, 4, 5, 6, 7, 9\}.$$

Die *Differenzmenge* $B \setminus A$ besteht aus allen Elementen, die in B, aber nicht in A liegen. Beispielsweise ist $\{1, 3, 5, 7\} \setminus \{1, 2, 3, 4\} = \{5, 7\}$. Gilt zudem $A \subseteq B$, so spricht man auch vom *Komplement*.

Für die Mengenoperationen gelten einige Regeln:

- Kommutativgesetz: $A \cup B = B \cup A$ und $A \cap B = B \cap A$
- Assoziativgesetz: $(A \cup B) \cup C = A \cup (B \cup C) = A \cup B \cup C$ und $(A \cap B) \cap C = A \cap (B \cap C) = A \cap B \cap C$
- Distributivgesetze: $(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$ und $(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$

Analoge Regeln werden uns auch beim Rechnen mit Zahlen begegnen.

1.1.3 Zahlenbereiche

Die wichtigsten Zahlenbereiche, mit denen wir arbeiten werden, sind

- die *natürlichen Zahlen* $\mathbb{N} = \{0, 1, 2, 3, \ldots\}$[2],
- die *ganzen Zahlen* $\mathbb{Z} = \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$,
- die *rationalen Zahlen* $\mathbb{Q}$, die sich als Brüche ganzer Zahlen (Nenner $\neq 0$) schreiben lassen, etwa $\frac{7}{3}$ oder $-\frac{2}{5}$,
- die *reellen Zahlen* $\mathbb{R}$, die man etwa als (endliche oder unendliche) Dezimalzahlen ausdrücken kann, wie zum Beispiel $\frac{11}{8} = 1.375$, $\frac{1}{7} = 0.14285714\ldots$ oder $\pi = 3.14159265\ldots$.

Diese Mengen sind Teilmengen voneinander: $\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R}$. Es ist nicht ganz offensichtlich, dass es *irrationale* reelle Zahlen gibt, also solche, die sich nicht als Bruch ganzer Zahlen schreiben lassen. Ein solche ist π (siehe oben). Ein weiteres Beispiel ist die Zahl $\sqrt{2} = 1.41421356\ldots$; dies sieht man wie folgt:

[2] Es gibt auch die Konvention $\mathbb{N} = \{1, 2, 3, \ldots\}$. Zur Unterscheidung spricht man auch von „nichtnegativen ganzen Zahlen“ und „positiven ganzen Zahlen“. Es werden auch Notationen wie $\mathbb{N}_0, \mathbb{N}^+, \mathbb{Z}^+$ verwendet.

Beispiel 1.4. Es gibt keine rationale Zahl, deren Quadrat gleich 2 ist, also $\sqrt{2} \notin \mathbb{Q}$. Um dies zu verstehen, nehmen wir an, dass die rationale Zahl $q = \frac{a}{b}$ die Gleichung $q^2 = 2$ erfüllt. Dabei können wir davon ausgehen, dass der Bruch $\frac{a}{b}$ bereits gekürzt ist, also a und b keinen gemeinsamen Teiler außer 1 haben.

Dann gilt $q^2 = \frac{a^2}{b^2} = 2$, also $a^2 = 2b^2$. Da die rechte Seite eine gerade Zahl ist, ist es auch die linke: a^2 ist gerade, daher ist auch a gerade. Wir können also $a = 2c$ schreiben, wobei c eine ganze Zahl ist. Damit folgt

$$q^2 = \frac{(2c)^2}{b^2} = \frac{4c^2}{b^2} = 2,$$

also $\frac{2c^2}{b^2} = 1$ oder $2c^2 = b^2$. Mit der gleichen Begründung wie zuvor sehen wir, dass b gerade sein muss. Damit haben a und b aber den gemeinsamen Teiler 2, was einen Widerspruch darstellt. Eine solche Darstellung $\frac{a}{b}$ für $\sqrt{2}$ kann es nicht geben.

Die Menge der irrationalen Zahlen $\mathbb{R} \setminus \mathbb{Q}$ wird auch als $\mathbb{I}$ geschrieben.

1.2 Rechnen und Vergleichen

1.2.1 Rechenregeln

Die bekannten Grundrechnungsarten $+$ und $\cdot$ (Addition und Multiplikation) erfüllen dieselben Gesetze, die wir schon bei den Mengenoperationen $\cup$ und $\cap$ gesehen haben:

- Kommutativgesetz: $a + b = b + a$ und $a \cdot b = b \cdot a$
- Assoziativgesetz: $(a + b) + c = a + (b + c) = a + b + c$ und $(a \cdot b) \cdot c = a \cdot (b \cdot c) = a \cdot b \cdot c$
- Distributivgesetz: $(a + b) \cdot c = (a \cdot c) + (b \cdot c)$

Die Zahlen 0 und 1 sind die *neutralen Elemente*: $a + 0 = 0 + a = a$ und $a \cdot 1 = 1 \cdot a = a$. Zu jeder reellen Zahl a gibt es ein negatives Element $-a$, sodass $a + (-a) = 0$. Zu jeder reellen Zahl $a \neq 0$ gibt es ein inverses Element (Kehrwert) $a^{-1} = \frac{1}{a}$, sodass $a \cdot a^{-1} = 1$. Die Operationen $-$ und $\div$ (Subtraktion und Division) können als

$$a - b = a + (-b) \quad \text{und} \quad \frac{a}{b} = a \cdot b^{-1}$$

geschrieben werden.

1.2.2 Gleichungen und Ungleichungen

Im weiteren Verlauf werden des Öfteren Gleichungen und Ungleichungen (üblicherweise in den reellen Zahlen) zu lösen sein. Dazu sind die folgenden Umformungsregeln wichtig, in denen a, b, c, d für beliebige reelle Zahlen stehen:

- Gilt $a = b$, dann auch $a + c = b + c$ und $a - c = b - c$, und umgekehrt.
- Gilt $a < b$, dann auch $a + c < b + c$ und $a - c < b - c$, und umgekehrt.
- Gilt $a < b$ und $c < d$, dann auch $a + c < b + d$ (nicht notwendigerweise umgekehrt!).
- Gilt $a = b$, dann auch $a \cdot c = b \cdot c$. Ist zudem $c \neq 0$, dann gilt auch die Umkehrung.

- Gilt $a < b$ und $c > 0$, dann auch $ac < bc$, und umgekehrt.
- Gilt $a < b$ und $c < 0$, dann auch $ac > bc$, und umgekehrt.
- Gilt $a < b$ und $c < d$, und sind a, b, c, d alle positiv, dann gilt auch $ac < bd$.

Vorsicht ist also vor allem bei der Multiplikation und Division (die der Multiplikation mit dem Kehrwert entspricht) notwendig, da es das Vorzeichen zu beachten gilt! Auch beim Kombinieren von mehreren Ungleichungen sollte man aufpassen: etwa folgt aus $a < b$ und $c < d$ *nicht* $a - c < b - d$. Als Gegenbeispiel kann man etwa $a = 3$, $b = 4$, $c = 1$, $d = 3$ betrachten.

Beispiel 1.5. Man bestimme alle reellen Zahlen, die die Ungleichung

$$\frac{x+3}{x+1} \geq 2x$$

erfüllen.

Lösung. Wir subtrahieren zunächst $2x$ auf beiden Seiten:

$$\frac{x+3}{x+1} - 2x = \frac{x+3}{x+1} - \frac{2x^2+2x}{x+1} = -\frac{2x^2+x-3}{x+1} \geq 0.$$

Der Zähler lässt sich faktorisieren, indem man die Nullstellen $-\frac{3}{2}$ und 1 des Zählerpolynoms bestimmt (etwa mit Hilfe der quadratischen Lösungsformel):

$$-\frac{(2x+3)(x-1)}{x+1} \geq 0 \quad \text{oder} \quad \frac{(2x+3)(x-1)}{x+1} \leq 0.$$

Je nachdem, wo x liegt, sind die einzelnen Faktoren entweder positiv oder negativ:

- $x \leq -\frac{3}{2}$: es gilt $2x + 3 \leq 0$, $x - 1 < 0$ und $x + 1 < 0$, daher ist der Quotient ≤ 0.
- $-\frac{3}{2} < x < -1$: es gilt $2x + 3 > 0$, $x - 1 < 0$ und $x + 1 < 0$, daher ist der Quotient > 0.
- $-1 < x \leq 1$:[3] es gilt $2x + 3 > 0$, $x - 1 \leq 0$ und $x + 1 > 0$, daher ist der Quotient ≤ 0.
- $x > 1$: $2x + 3$, $x - 1$ und $x + 1$ sind allesamt positiv, daher auch der Quotient.

Die Ungleichung wird also von allen x erfüllt, für die entweder $x \leq -\frac{3}{2}$ oder $-1 < x \leq 1$ gilt.

Bemerkung 1.6. Eine alternative Lösung besteht darin, dass man zwei Fälle unterscheidet: $x < -1$ und $x > -1$. Dann kann man zur Vereinfachung mit dem Nenner multiplizieren und erhält

- für $x < -1$: $x + 3 \leq 2x(x + 1)$ (die Ungleichung dreht sich um, da $x + 1$ in diesem Fall negativ ist!). Daraus ergibt sich durch Vereinfachung $2x^2 + x - 3 = (2x + 3)(x - 1) \geq 0$, was für $x \leq -\frac{3}{2}$ oder $x \geq 1$ gilt. Letzteres kommt für $x < -1$ nicht in Frage, womit nur $x \leq -\frac{3}{2}$ übrig bleibt.
- für $x > -1$: $x + 3 \geq 2x(x + 1)$. Daraus ergibt sich durch Vereinfachung $2x^2 + x - 3 = (2x + 3)(x - 1) \leq 0$, was für $-\frac{3}{2} \leq x \leq 1$ gilt. Da zudem $x > -1$ vorausgesetzt wurde, bleibt nur $-1 < x \leq 1$.

[3] Man beachte, dass $x = -1$ unmöglich ist, da hier der Nenner 0 wäre!

Die Lösungen von Gleichungen und Ungleichungen lassen sich oft zweckmäßig durch *Intervalle* ausdrücken. Man unterscheidet dabei *offene* und *abgeschlossene* Intervalle (sowie halboffene Intervalle, die beides kombinieren). Im Folgenden seien a und b beliebige reelle Zahlen mit $a < b$.

- offenes Intervall: $(a, b) = \{x \in \mathbb{R} : a < x < b\}$.
- abgeschlossenes Intervall: $[a, b] = \{x \in \mathbb{R} : a \leq x \leq b\}$.
- halboffenes Intervall: $(a, b] = \{x \in \mathbb{R} : a < x \leq b\}$.
- halboffenes Intervall: $[a, b) = \{x \in \mathbb{R} : a \leq x < b\}$.

Die linke Grenze in der Intervallnotation kann auch $-\infty$ sein, die rechte ∞. So ist etwa

$$(-\infty, b) = \{x \in \mathbb{R} : x < b\} \quad \text{und} \quad [a, \infty) = \{x \in \mathbb{R} : x \geq a\}.$$

Bemerkung 1.7. $\pm\infty$ gelten nicht als reelle Zahlen! Intervalle wie $[-\infty, 2)$ oder $[3, \infty]$ sind daher in unserem Kontext normalerweise nicht sinnvoll. Die Menge $\mathbb{R}$ der reellen Zahlen kann in Intervallnotation auch als $(-\infty, \infty)$ geschrieben werden.

Mit Hilfe der Intervallnotation ist etwa die Lösungsmenge der Ungleichung in Beispiel 1.5 $(-\infty, -\frac{3}{2}] \cup (-1, 1]$.

Definition 1.8. Der *Absolutbetrag* (oder auch nur *Betrag*) $|x|$ einer reellen Zahl x ist die Zahl ohne Vorzeichen, also

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

Also zum Beispiel $|3| = 3$, $|0| = 0$, $|-4| = 4$, oder auch $|2 - \pi| = \pi - 2$ (da $2 - \pi < 0$). Man beachte hier: ist x eine negative Zahl, dann ist $-x$ positiv!

Es gelten die folgenden Regeln für Absolutbeträge (wobei $a \geq 0$ sein soll):

- $|x| = a$ genau dann, wenn $x = a$ oder $x = -a$.
- $|x| \leq a$ genau dann, wenn $-a \leq x \leq a$ bzw. $x \in [-a, a]$ gilt.
- $|x| \geq a$ genau dann, wenn $x \leq -a$ oder $x \geq a$ bzw. $x \in (-\infty, -a] \cup [a, \infty)$ gilt.
- $|xy| = |x||y|$ und $\left|\frac{x}{y}\right| = \frac{|x|}{|y|}$ für alle reellen Zahlen x und y (letztere Gleichung, falls $y \neq 0$).
- $|x + y| \leq |x| + |y|$ für alle reellen Zahlen x und y (sogenannte *Dreiecksungleichung*). Dies kann etwa dadurch verifiziert werden, dass man alle vier möglichen Varianten für die Vorzeichen von x und y durchgeht.

Beispiel 1.9. Man bestimme alle reellen Zahlen x, für die

$$\left|\frac{3x + 1}{2x}\right| = 2$$

ist.

Lösung. Es gilt entweder $\frac{3x+1}{2x} = 2$, was äquivalent zu $3x + 1 = 4x$ oder $x = 1$ ist, oder $\frac{3x+1}{2x} = -2$, was äquivalent zu $3x + 1 = -4x$ oder $x = -\frac{1}{7}$ ist. Die einzigen Lösungen sind somit 1 und $-\frac{1}{7}$.

1.2.3 Schranken, Minima, Maxima, Infima und Suprema

Definition 1.10. Es sei M eine Menge reeller Zahlen. Eine Zahl a heißt *untere Schranke* für M, falls $a \leq x$ für alle $x \in M$ gilt. Analog heißt a *obere Schranke* für M, wenn $a \geq x$ für alle $x \in M$ gilt.

Beispielsweise sind $-\pi$, 0 und 2 allesamt untere Schranken für das Intervall $[2, 5)$. Die Zahlen 5, 8 und 10000 sind dagegen obere Schranken.

Definition 1.11. Es sei M eine Menge reeller Zahlen. Ist ein Element $a \in M$ eine untere Schranke (also $a \leq x$ für alle $x \in M$), dann wird a *Minimum* von M (geschrieben $\min M$) genannt. Ist ein Element eine obere Schranke, so wird es *Maximum* (geschrieben $\max M$) genannt.

Minima und Maxima müssen nicht immer existieren. So haben etwa die folgenden Mengen kein Minimum:

- $(-\infty, 1)$, da es gar keine untere Schranke gibt.
- $(0, 1)$, da es zu jedem Element a der Menge ein noch kleineres gibt (z.B. $\frac{a}{2}$).
- $\{1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots\}$, aus dem gleichen Grund.

Definition 1.12. Die größte untere Schranke einer Menge M reeller Zahlen heißt *Infimum* (geschrieben $\inf M$). Die kleinste obere Schranke heißt *Supremum* (geschrieben $\sup M$).

So ist beispielsweise $\sup\,[-1, 3) = 3$ oder $\inf \{1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots\} = 0$. Gibt es keine untere Schranke, so definiert man $\inf M = -\infty$. Gibt es keine obere Schranke, so setzt man $\sup M = \infty$. Daher gilt beispielsweise $\sup \mathbb{N} = \infty$ (aber $\inf \mathbb{N} = \min \mathbb{N} = 0$).
Gibt es ein Minimum, so ist dieses auch Infimum. Gibt es ein Maximum, so ist dieses auch Supremum. Wir wissen jedoch, dass es Minimum und Maximum nicht geben muss. Die reellen Zahlen haben aber die folgende *Vollständigkeitseigenschaft*:

- Jede nicht leere, nach oben beschränkte Menge hat ein Supremum.
- Jede nicht leere, nach unten beschränkte Menge hat ein Infimum.

Man beachte, dass die rationalen Zahlen diese Eigenschaft nicht haben. Etwa hat die Menge $M = \{x \in \mathbb{Q} : x^2 \leq 2\}$ kein rationales Supremum: in den reellen Zahlen ist $\sup M = \sqrt{2}$, wir haben aber gezeigt, dass dies keine rationale Zahl ist.

1.3 Komplexe Zahlen

Es gibt Polynomgleichungen, die in den reellen Zahlen keine Lösung haben. Die einfachste solche Gleichung ist

$$x^2 + 1 = 0 \text{ oder } x^2 = -1.$$

Diese Tatsache war die ursprüngliche Motivation, *komplexe Zahlen* zu definieren: dazu wird die *imaginäre Einheit* i eingeführt[4], die definitionsgemäß die Gleichung $i^2 = -1$ erfüllen soll. Eine komplexe Zahl ist dann eine Zahl der Form

$$z = a + bi, \qquad a, b \in \mathbb{R}.$$

[4]In der Elektrotechnik wird der Buchstabe i gerne für die Stromstärke verwendet. Daher wird dort die imaginäre Einheit oft auch mit j bezeichnet.

Die Zahlen a und b heißen *Realteil* und *Imaginärteil* von z, und man schreibt $a = \operatorname{Re} z$ und $b = \operatorname{Im} z$. Die Menge aller komplexen Zahlen wird mit $\mathbb{C}$ bezeichnet.

1.3.1 Grundrechnungsarten

Die vier Grundrechnungsarten Addition, Subtraktion, Multiplikation und Division lassen sich auf die komplexen Zahlen ausweiten. Für Addition und Subtraktion werden Real- und Imaginärteil getrennt voneinander behandelt:

$$(a+bi)+(c+di)=(a+c)+(b+d)i, \qquad (a+bi)-(c+di)=(a-c)+(b-d)i.$$

Für die Multiplikation wird ausmultipliziert und dann zusammengefasst:

$$(a+bi)(c+di)=ac+bci+adi+bdi^2=ac+bci+adi-bd=(ac-bd)+(bc+ad)i. \tag{1.1}$$

Die Division wird schließlich durchgeführt, indem man zunächst mit der sogenannten *Konjugierten* des Nenners erweitert. Durch Verwendung der bekannten Identität $(x+y)(x-y) = x^2-y^2$ wird dabei der Nenner reell:

$$\begin{aligned}\frac{a+bi}{c+di} &= \frac{(a+bi)(c-di)}{(c+di)(c-di)} = \frac{ac+bci-adi-bdi^2}{c^2-d^2i^2} = \frac{ac+bci-adi+bd}{c^2+d^2} \\ &= \frac{(ac+bd)+(bc-ad)i}{c^2+d^2} = \frac{ac+bd}{c^2+d^2} + \frac{bc-ad}{c^2+d^2}i.\end{aligned}$$

Beispiel 1.13. Für die beiden komplexen Zahlen $6+5i$ und $2-i$ erhalten wir:

$$(6+5i)+(2-i)=8+4i, \qquad (6+5i)-(2-i)=4+6i,$$

sowie

$$(6+5i)(2-i)=12+10i-6i-5i^2=12+4i+5=17+4i$$

und

$$\frac{6+5i}{2-i}=\frac{(6+5i)(2+i)}{(2-i)(2+i)}=\frac{12+10i+6i+5i^2}{2^2-i^2}=\frac{12+16i-5}{4+1}=\frac{7+16i}{5}=\frac{7}{5}+\frac{16}{5}i.$$

Bemerkung 1.14. Für komplexe Zahlen gibt es keine sinnvolle Definition der Relationen $<$, $>$, $\leq$ und $\geq$, die mit den Grundrechnungsarten kompatibel ist.

Die Konjugierte oder konjugiert komplexe Zahl ist bereits im Zuge der Division vorgekommen. Es gibt dafür auch eine eigene Notation.

Definition 1.15. Die zu $z = a+bi$ *konjugiert komplexe* Zahl ist $\overline{z} = a - bi$. Weiters ist der *Absolutbetrag* von z durch

$$|z| = \sqrt{z\overline{z}} = \sqrt{a^2+b^2}$$

definiert.

Zum Beispiel ist $\overline{2+3i} = 2-3i$, und $|2+3i| = \sqrt{2^2+3^2} = \sqrt{13}$. Es gelten folgende Eigenschaften:

- $\overline{\overline{z}} = z$,
- $\operatorname{Re} z = \frac{z+\overline{z}}{2}$, $\operatorname{Im} z = \frac{z-\overline{z}}{2}$,

- $\overline{z_1 \pm z_2} = \overline{z_1} \pm \overline{z_2}$,
- $\overline{z_1 z_2} = \overline{z_1} \cdot \overline{z_2}$,
- $|z_1 z_2| = |z_1| \cdot |z_2|$,
- $|z_1 + z_2| \leq |z_1| + |z_2|$ (Dreiecksungleichung).

Diese Eigenschaften kann man durch Einsetzen der jeweiligen Definition nachprüfen. So gilt etwa für $z_1 = a + bi$ und $z_2 = c + di$, dass

$$\begin{aligned}\overline{z_1}\,\overline{z_2} &= (a - bi)(c - di) = ac - bci - adi + bdi^2 \\ &= (ac - bd) - (bc + ad)i = \overline{(ac - bd) + (bc + ad)i} \\ &= \overline{z_1 z_2}.\end{aligned}$$

Hier wurde der Ausdruck aus (1.1) für $z_1 z_2$ verwendet. Weiters ist

$$\begin{aligned}|z_1 z_2| &= \sqrt{(ac - bd)^2 + (bc + ad)^2} = \sqrt{a^2c^2 - 2abcd + b^2d^2 + b^2c^2 + 2abcd + a^2d^2} \\ &= \sqrt{a^2c^2 + b^2d^2 + b^2c^2 + a^2d^2} = \sqrt{(a^2 + b^2)(c^2 + d^2)} \\ &= \sqrt{a^2 + b^2} \cdot \sqrt{c^2 + d^2} = |z_1| \cdot |z_2|.\end{aligned}$$

1.3.2 Geometrische Interpretation, Polarform

Man kann komplexe Zahlen als Punkte in der Ebene deuten, wobei der komplexen Zahl $z = a + bi$ der Punkt mit x-Koordinate a und y-Koordinate b zugeordnet wird. Die Koordinatenachsen werden in diesem Zusammenhang als *reelle* und *imaginäre Achse* bezeichnet. Der Absolutbetrag $|z| = \sqrt{a^2 + b^2}$ ist nach dem Satz von Pythagoras genau der Abstand von z zum Ursprung.

Das Addieren komplexer Zahlen entspricht geometrisch einem „Aneinanderfügen“, wie in Abbildung 1.1 dargestellt. Die Dreiecksungleichung kann man tatsächlich als Ungleichung in einem Dreieck interpretieren: die zwei Seiten mit den Längen $|z_1|$ und $|z_2|$ sind zusammen länger als die dritte Seite mit Seitenlänge $|z_1 + z_2|$.

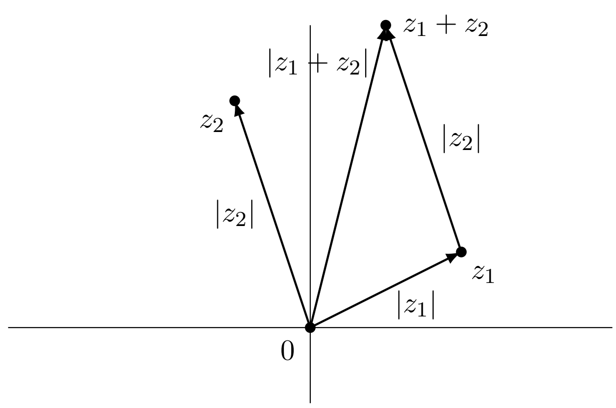

Abbildung 1.1: Addition von komplexen Zahlen.

Es gibt zu jeder positiven Zahl r unendlich viele komplexe Zahlen, deren Absolutbetrag r ist (diese bilden einen Kreis mit Radius r um den Ursprung). Man benötigt daher zusätzliche Information zum Absolutbetrag um eine komplexe Zahl z eindeutig zu bestimmen. Diese Information ist durch das *Argument* gegeben, also den Winkel zwischen der reellen Achse und der Verbindung zwischen z und dem Ursprung. Man schreibt für dieses auch $\arg z$. Das Argument ist allerdings nicht eindeutig: ist φ Argument von z, dann auch jede andere Zahl der Form $\varphi + 2k\pi$. Der sogenannte *Hauptwert* des Arguments ist dabei jener Winkel, der im Intervall $(-\pi, \pi]$ liegt. Siehe Abbildung 1.2

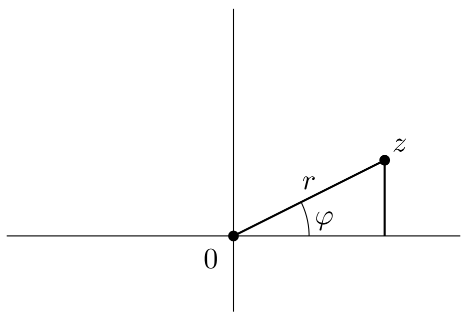

Abbildung 1.2: Polardarstellung von komplexen Zahlen.

Aus dem in der Abbildung eingezeichneten rechtwinkligen Dreieck ergeben sich für eine komplexe Zahl z mit Absolutbetrag $r = |z|$ und Argument $\varphi = \arg z$ die Ausdrücke

$$\operatorname{Re} z = r\cos\varphi, \qquad \operatorname{Im} z = r\sin\varphi.$$

Zusammengefasst ergibt sich damit die *Polarform* von z:

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

Die konjugiert komplexe Zahl $\overline{z}$ hat denselben Absolutbetrag und Argument $-\varphi$, denn die Konjugation entspricht einer Spiegelung an der reellen Achse:

$$\overline{z} = r\cos\varphi - ir\sin\varphi = r(\cos(-\varphi) + i\sin(-\varphi)).$$

Die Multiplikation von komplexen Zahlen lässt sich in Polarkoordinaten ebenfalls sehr einfach ausdrücken. Dazu benötigen wir die *Additionstheoreme* für die Winkelfunktionen Sinus und Cosinus:

$$\cos(x \pm y) = \cos x\cos y \mp \sin x\sin y, \qquad \sin(x \pm y) = \sin x\cos y \pm \cos x\sin y. \tag{1.2}$$

Die Additionstheoreme für die Summe $x + y$ sind in Abbildung 1.3 illustriert: hier betrachten wir zwei aufeinander aufgesetzte rechtwinklige Dreiecke OAB (mit Hypotenuse OB und Winkel $\angle AOB = x$) und OBC (mit Hypotenuse OC und Winkel $\angle BOC = y$). Von C aus werden zudem die Parallelen zu OA und AB eingezeichnet, die AB und OA in X bzw. Y treffen.

Wir nehmen an, dass die Länge von OC gleich 1 ist. Im rechtwinkligen Dreieck OBC ergibt sich dann, dass die Längen von OB und BC gleich $\cos y$ bzw. $\sin y$ sind. Damit haben wir die Hypotenuse von OAB bestimmt und erhalten für die Katheten OA und AB die Längen $\cos x\cos y$ bzw. $\sin x\cos y$. Außerdem ist im rechtwinkligen Dreieck BXC der Winkel $\angle XBC$ gleich x, und daher die Längen der Katheten BX und XC gleich $\cos x\sin y$ bzw. $\sin x\sin y$.

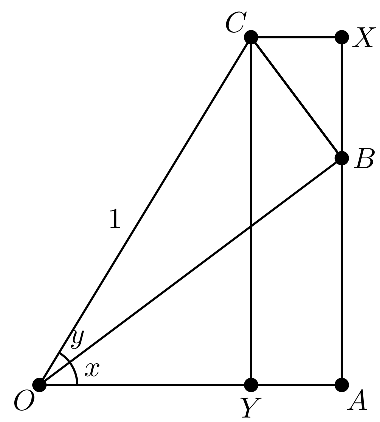

Abbildung 1.3: Illustration der Additionstheoreme.

Damit erhalten wir nun die Länge von OY als Differenz von OA und XC, also $\cos x \cos y - \sin x \sin y$, und die Länge von AX (die gleich der Länge von YC ist) als Summe von AB und BX, also $\sin x \cos y + \cos x \sin y$. Diese beiden sind Katheten im rechtwinkligen Dreieck OYC mit Hypotenuse 1 und Winkel $\angle YOC = x + y$, also folgen die Formeln in (1.2) für die Winkelfunktionen $\cos(x + y)$ und $\sin(x + y)$.

Damit ergibt sich nun für zwei komplexe Zahlen z, w in Polarform, also

$$z = r(\cos\varphi + i \sin\varphi), \qquad w = s(\cos\psi + i \sin\psi),$$

das Produkt

$$\begin{aligned} zw &= rs(\cos\varphi\cos\psi + i\sin\varphi\cos\psi + i\cos\varphi\sin\psi + i^2\sin\varphi\sin\psi) \\ &= rs\big((\cos\varphi\cos\psi - \sin\varphi\sin\psi) + i(\sin\varphi\cos\psi + \cos\varphi\sin\psi)\big) \\ &= rs\big(\cos(\varphi + \psi) + i\sin(\varphi + \psi)\big). \end{aligned}$$

Es gilt also: beim Multiplizieren von komplexen Zahlen werden

- die Absolutbeträge multipliziert,
- die Argumente addiert.

Beim Dividieren werden dementsprechend die Absolutbeträge dividiert und die Argumente subtrahiert. Das Potenzieren komplexer Zahlen lässt sich in Polarform viel einfacher bewerkstelligen, als es sonst durch wiederholtes Multiplizieren der Fall wäre: es gilt für eine komplexe Zahl $z = r(\cos\varphi + i\sin\varphi)$ und eine ganze Zahl n

$$z^n = r^n\big(\cos(n\varphi) + i\sin(n\varphi)\big).$$

Wenn wir umgekehrt eine n-te Wurzel suchen, also eine Lösung der Gleichung

$$w^n = z = r(\cos\varphi + i\sin\varphi),$$

dann muss die Polardarstellung von w, also $w = s(\cos\psi + i\sin\psi)$, zwei Bedingungen erfüllen:

- s^n ist Absolutbetrag von z, also $s^n = r$,

- $n\psi$ ist Argument von z, also $n\psi = \varphi + 2k\pi$ für ein $k \in \mathbb{Z}$.

Erstere Bedingung führt auf $s = \sqrt[n]{r} = r^{1/n}$. Aus der zweiteren Bedingung erhalten wir

$$\psi = \frac{\varphi + 2k\pi}{n} = \frac{\varphi}{n} + \frac{2k\pi}{n}.$$

Wenn wir für k der Reihe nach $0, 1, \ldots, n-1$ einsetzen, erhalten wir

$$w_k = r^{1/n}\Big(\cos\Big(\frac{\varphi}{n} + \frac{2k\pi}{n}\Big) + i\sin\Big(\frac{\varphi}{n} + \frac{2k\pi}{n}\Big)\Big), \quad k = 0, 1, \ldots, n-1.$$

Diese sind die n verschiedenen n-ten Wurzeln[5]. Für $k = n$ würden wir wegen der Periodizität von Sinus und Cosinus wieder denselben Wert wie für $k = 0$ erhalten, für $k = n+1$ denselben wie für $k = 1$, für $k = -1$ denselben wie für $k = n-1$, und so weiter.

Beispiel 1.16. Wir suchen die sechsten Wurzeln der Zahl $z = -8$. Da die Polardarstellung

$$-8 = 8 \cdot (\cos\pi + i\sin\pi)$$

ist, ergeben sich die sechs Wurzeln

$$w_k = 8^{1/6}\Big(\cos\Big(\frac{\pi}{6} + \frac{k\pi}{3}\Big) + i\sin\Big(\frac{\pi}{6} + \frac{k\pi}{3}\Big)\Big), \quad k = 0, 1, \ldots, 5.$$

Das sind (da $8^{1/6} = 2^{3/6} = \sqrt{2}$)

$$\begin{aligned}
w_0 &= \sqrt{2}\Big(\cos\frac{\pi}{6} + i\sin\frac{\pi}{6}\Big) = \frac{\sqrt{6}}{2} + \frac{\sqrt{2}}{2}i,\\
w_1 &= \sqrt{2}\Big(\cos\frac{\pi}{2} + i\sin\frac{\pi}{2}\Big) = \sqrt{2}i,\\
w_2 &= \sqrt{2}\Big(\cos\frac{5\pi}{6} + i\sin\frac{5\pi}{6}\Big) = -\frac{\sqrt{6}}{2} + \frac{\sqrt{2}}{2}i,\\
w_3 &= \sqrt{2}\Big(\cos\frac{7\pi}{6} + i\sin\frac{7\pi}{6}\Big) = -\frac{\sqrt{6}}{2} - \frac{\sqrt{2}}{2}i,\\
w_4 &= \sqrt{2}\Big(\cos\frac{3\pi}{2} + i\sin\frac{3\pi}{2}\Big) = -\sqrt{2}i,\\
w_5 &= \sqrt{2}\Big(\cos\frac{11\pi}{6} + i\sin\frac{11\pi}{6}\Big) = \frac{\sqrt{6}}{2} - \frac{\sqrt{2}}{2}i.
\end{aligned}$$

Die sechs gefundenen Lösungen sind in Abbildung 1.4 illustriert. Sie bilden ein regelmäßiges Sechseck mit dem Mittelpunkt im Ursprung. Allgemein bilden die n-ten Wurzeln einer komplexen Zahl ein regelmäßiges n-Eck.

Im Gegensatz zu den reellen Zahlen gibt es in den komplexen Zahlen also immer n-te Wurzeln zu einer gegebenen komplexen Zahl. Dies gilt auch für allgemeinere Gleichungen. So können etwa quadratische Gleichungen der Form

$$ax^2 + bx + c = 0$$

in komplexen Zahlen mit Hilfe der Lösungsformel

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

gelöst werden und haben immer zwei komplexe Lösungen (die möglicherweise zusammenfallen).

[5] Wenn $z = 0$, fallen alle Wurzeln zusammen.

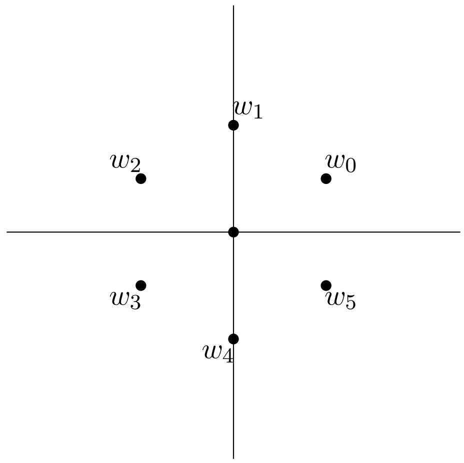

Abbildung 1.4: Die sechsten Wurzeln von -8.

Beispiel 1.17. Gegeben sei die quadratische Gleichung

$$z^2 - 6iz + 9i = 0.$$

Die Lösungsformel liefert

$$z = 3i \pm \sqrt{-9 - 9i}.$$

Nun hat $-9 - 9i$ die Polardarstellung

$$-9 - 9i = 9\sqrt{2}\big(\cos(-3\pi/4) + i\sin(-3\pi/4)\big).$$

Daher sind die beiden Wurzeln von $-9 - 9i$

$$w_0 = 3 \cdot 2^{1/4}\big(\cos(-3\pi/8) + i\sin(-3\pi/8)\big)$$

und

$$w_1 = 3 \cdot 2^{1/4}\big(\cos(5\pi/8) + i\sin(5\pi/8)\big),$$

womit sich die beiden Lösungen der quadratischen Gleichung ergeben:

$$z_1 = 3i + 3 \cdot 2^{1/4}\big(\cos(-3\pi/8) + i\sin(-3\pi/8)\big) \approx 1.36527 - 0.29605i$$

und

$$\begin{aligned} z_2 &= 3i - 3 \cdot 2^{1/4}\big(\cos(-3\pi/8) + i\sin(-3\pi/8)\big) \\ &= 3i + 3 \cdot 2^{1/4}\big(\cos(5\pi/8) + i\sin(5\pi/8)\big) \approx -1.36527 + 6.29605i. \end{aligned}$$

Die Tatsache, dass quadratische Gleichungen in komplexen Zahlen immer gelöst werden können, ist ein Spezialfall des folgenden Satzes:

Satz 1.18 (Fundamentalsatz der Algebra)**.** *Eine polynomielle Gleichung n-ten Grades ($n > 0$), also eine Gleichung der Form*

$$a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 = 0,$$

wobei die Koeffizienten a_i komplexe Zahlen mit $a_n \neq 0$ sind, hat immer eine Lösung in komplexen Zahlen.

1.4 Vollständige Induktion

1.4.1 Das Prinzip der vollständigen Induktion und einige Beispiele

Die Methode der vollständigen Induktion wird verwendet um Aussagen zu beweisen, die für alle natürlichen Zahlen (oder alle natürlichen Zahlen ab einem bestimmten Punkt) gelten sollen. Gegeben sei eine Aussage $P(n)$, die von einer natürlichen Zahl n abhängt, und die wir beweisen wollen. Die vollständige Induktion besteht aus zwei Schritten:

- *Induktionsanfang*: man beweist die Aussage für den ersten Wert von n, für den sie gelten soll (meistens $n = 0$ oder $n = 1$).

- *Induktionsschritt*: unter der Annahme, dass $P(n)$ gilt (*Induktionsannahme* oder *Induktionshypothese*), wird gezeigt, dass auch $P(n + 1)$ gelten muss.

Damit wird eine Implikationskette in Gang gesetzt: gilt $P(0)$, dann auch $P(1)$; damit auch $P(2)$; damit auch $P(3)$, und so weiter, wie in der Abbildung illustriert:

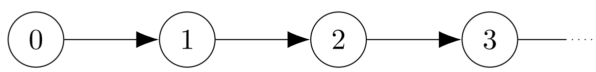

Abbildung 1.5: Zum Prinzip der Induktion

Betrachten wir zunächst ein einfaches Beispiel.

Beispiel 1.19. Wenn wir der Reihe nach die ungeraden natürlichen Zahlen aufaddieren, erkennen wir ein Muster:

$$\begin{aligned} 1 &= 1 = 1^2, \\ 1 + 3 &= 4 = 2^2, \\ 1 + 3 + 5 &= 9 = 3^2, \\ 1 + 3 + 5 + 7 &= 16 = 4^2. \end{aligned}$$

Wir vermuten also allgemein

$$1 + 3 + 5 + \cdots + (2n - 1) = \sum_{i=1}^{n} (2i - 1) = n^2. \tag{1.3}$$

Wie können wir uns sicher sein, dass das Muster immer so weitergeht? Eine Möglichkeit ist ein Beweis durch Induktion.

- Induktionsanfang: dies wurde bereits oben erledigt – die Gleichung (1.3) gilt offensichtlich für $n = 1$.

- Wenn wir annehmen, dass

  $$1 + 3 + 5 + \cdots + (2n - 1) = n^2$$

  für ein gewisses n gilt, dann können wir den nächsten Summanden $2n + 1$ hinzufügen und erhalten

  $$1 + 3 + 5 + \cdots + (2n - 1) + (2n + 1) = n^2 + (2n + 1) = n^2 + 2n + 1 = (n + 1)^2.$$

  Also haben wir gezeigt, dass (1.3) auch für $n + 1$ anstelle von n gültig ist, womit die Induktion abgeschlossen ist.

Wichtig ist immer, dass beide Teile der Induktion vorhanden sind. Der Induktionsanfang ist meistens (wie im Beispiel) der einfachere Teil. Im Induktionsschritt versucht man, die Annahme, dass $P(n)$ gilt, mit der Aussage $P(n+1)$ in Verbindung zu bringen. Dies wird auch im folgenden Beispiel illustriert.

Beispiel 1.20. Man zeige, dass $a_n = 6^n + 9$ für jede natürliche Zahl n durch 5 teilbar ist.

Lösung. Wir sehen zunächst, dass $a_0 = 10$ tatsächlich durch 5 teilbar ist. Damit ist der Induktionsanfang bereits erledigt. Für den Induktionsschritt nehmen wir nun an, dass $a_n = 6^n + 9$ ein Vielfaches von 5 ist. Wir können also $a_n = 5K$ schreiben, wobei K eine natürliche Zahl ist.

Nun wollen wir mit Hilfe dieser Annahme zeigen, dass auch $a_{n+1} = 6^{n+1} + 9$ ein Vielfaches von 5 ist. Dazu drücken wir erst a_{n+1} durch a_n aus:

$$a_{n+1} = 6^{n+1} + 9 = 6 \cdot 6^n + 9 = 6 \cdot (6^n + 9 - 9) + 9 = 6 \cdot (a_n - 9) + 9 = 6a_n - 45.$$

Aus dieser Identität und der Induktionsannahme folgt nun, dass

$$a_{n+1} = 6 \cdot 5K - 45 = 5(6K - 9),$$

d.h., a_{n+1} ist in der Tat durch 5 teilbar, womit die Induktion vollständig ist.

Das letzte Beispiel illustriert, dass Induktion ein sehr flexibles Werkzeug ist, das sich auf unterschiedlichen mathematischen Gebieten einsetzen lässt.

Beispiel 1.21. Es werden n Geraden in der Ebene gezeichnet, von denen keine zwei parallel sind und keine drei durch einen gemeinsamen Punkt gehen (siehe Abbildung 1.6, wo der Spezialfall $n = 4$ gezeigt wird). Diese Geraden teilen die Ebene in mehrere Gebiete auf. Man zeige, dass diese Anzahl gleich $\frac{n^2+n+2}{2}$ ist.

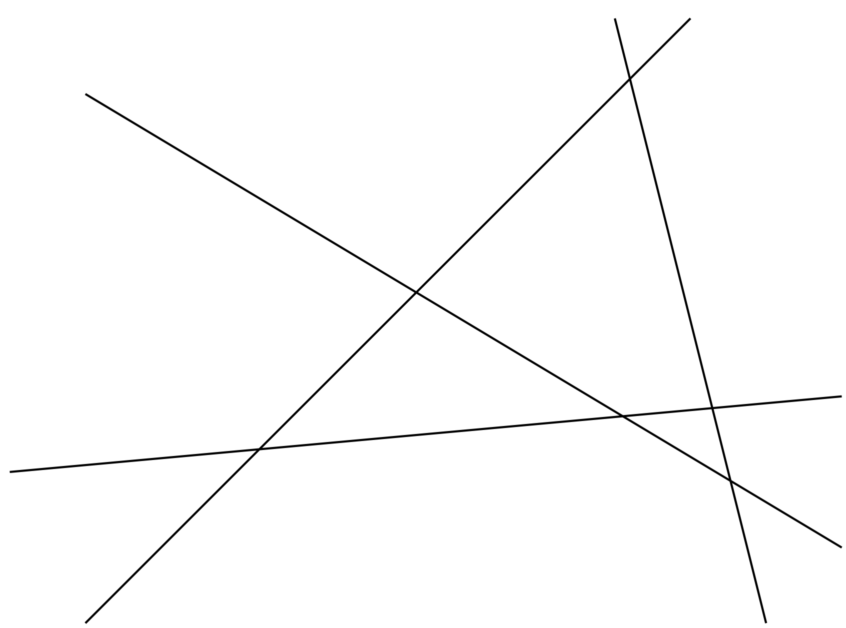

Abbildung 1.6: Vier Geraden teilen die Ebene in elf Gebiete.

Lösung. Die Formel gilt offenbar für $n = 0$: hier ist die gesamte Ebene ein Gebiet, die Anzahl folglich 1. Nun nehmen wir an, dass n Geraden $\frac{n^2+n+2}{2}$ Gebiete erzeugen. Wenn nun eine weitere Gerade (Gerade $n+1$) eingezeichnet wird, so muss diese alle anderen schneiden (weil keine zwei Geraden parallel sind). Die n Schnittpunkte, die sich ergeben, sind alle verschieden (weil keine drei Geraden durch einen gemeinsamen Punkt gehen). Es liegen auf der neuen Geraden also n Schnittpunkte, die diese in $n+1$ Stücke unterteilen. Jedes von diesen Stücken teilt eines der „alten" Gebiete in zwei „neue" Gebiete. Die Anzahl der Gebiete wächst also um $n+1$ und ist nun

$$\frac{n^2+n+2}{2} + n + 1 = \frac{n^2+3n+4}{2} = \frac{(n+1)^2+(n+1)+2}{2}.$$

Damit ist die Induktion abgeschlossen.

1.4.2 Der binomische Lehrsatz

Wir wollen nun die Methode der vollständigen Induktion auf einen wichtigen Satz anwenden, den sogenannten *binomischen Lehrsatz*. Die Motivation dafür ist das Auspotenzieren von Summen:

$$(x+y)^0 = 1, \quad (x+y)^1 = x+y, \quad (x+y)^2 = x^2+2xy+y^2, \quad (x+y)^3 = x^3+3x^2y+3xy^2+y^3, \ldots$$

Wir wollen eine allgemeine Formel für die n-te Potenz. Dazu sind zunächst einige Definitionen nötig.

Definition 1.22. Die *Faktorielle* oder *Fakultät* einer positiven ganzen Zahl n ist das Produkt der Zahlen von 1 bis n:

$$n! = 1 \cdot 2 \cdot 3 \cdots (n-1) \cdot n.$$

Beispielsweise ist $5! = 1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 = 120$. Darüber hinaus definieren wir $0! = 1$. Der *Binomialkoeffizient* $\binom{n}{k}$ („n über k") ist nun für $0 \leq k \leq n$ wie folgt definiert:

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

Zum Beispiel ist

$$\binom{7}{3} = \frac{7!}{3! \cdot 4!} = \frac{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6 \cdot 7}{(1 \cdot 2 \cdot 3) \cdot (1 \cdot 2 \cdot 3 \cdot 4)} = \frac{5 \cdot 6 \cdot 7}{1 \cdot 2 \cdot 3} = 35.$$

Bemerkung 1.23. Man beachte, dass $n! = (n-1)! \cdot n$ gilt, denn $n!$ ergibt sich aus $(n-1)!$ durch Hinzufügen eines weiteren Faktors. Dies ist auch ein Grund, warum $0!$ als 1 definiert wird, und nicht etwa 0: damit bleibt die Identität richtig, denn in der Tat ist $1! = 0! \cdot 1$. Die Zahl $0!$ ist gewissermaßen ein „Produkt mit 0 Faktoren" und damit vergleichbar mit $a^0 = 1$.

Die Binomialkoeffizienten erfüllen folgende wichtige Eigenschaften:

$$\binom{n}{0} = \binom{n}{n} = 1, \qquad \binom{n}{k} = \binom{n}{n-k}, \tag{1.4}$$

und schließlich

$$\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}. \tag{1.5}$$

Die ersten beiden Eigenschaften ergeben sich unmittelbar aus der Definition. Die dritte erhält man durch Nachrechnen: zunächst ist laut Definition

$$\binom{n-1}{k} + \binom{n-1}{k-1} = \frac{(n-1)!}{k!(n-1-k)!} + \frac{(n-1)!}{(k-1)!(n-k)!}.$$

Auf den gemeinsamen Nenner $k!(n-k)!$ fehlt im ersten Bruch ein Faktor $n-k$ (denn $(n-k)! = (n-1-k)! \cdot (n-k)$), und im zweiten Bruch analog dazu ein Faktor k. Wir erhalten also

$$\begin{aligned}\binom{n-1}{k} + \binom{n-1}{k-1} &= \frac{(n-1)! \cdot (n-k)}{k!(n-k)!} + \frac{(n-1)! \cdot k}{k!(n-k)!} \\ &= \frac{(n-1)! \cdot (n-k+k)}{k!(n-k)!} = \frac{(n-1)! \cdot n}{k!(n-k)!} \\ &= \frac{n!}{k!(n-k)!} = \binom{n}{k}.\end{aligned}$$

Die Binomialkoeffizienten lassen sich gut im sogenannten *Pascalschen Dreieck* veranschaulichen. Dabei steht $\binom{n}{k}$ an der k-ten Stelle in der n-ten Zeile (wobei wir beim Zählen mit 0 beginnen):

$$\begin{array}{ccccccccccc} &&&&& 1 &&&&& \\ &&&& 1 && 1 &&&& \\ &&& 1 && 2 && 1 &&& \\ && 1 && 3 && 3 && 1 && \\ & 1 && 4 && 6 && 4 && 1 & \\ 1 && 5 && 10 && 10 && 5 && 1 \end{array}$$

Die Zeilen sind symmetrisch, und die erste und letzte Zahl ist jeweils 1 (laut (1.4)). Die Identität (1.5) kann so interpretiert werden, dass jede Zahl die Summe der beiden direkt über ihr ist. Die Zahlen in der n-ten Zeile des Pascalschen Dreiecks sind genau jene, die sich beim Ausmultiplizieren von $(x+y)^n$ ergeben. Das ist der Inhalt des *binomischen Lehrsatzes*:

Satz 1.24. *Für jede natürliche Zahl n gilt*

$$\begin{aligned}(x+y)^n &= \sum_{k=0}^{n} \binom{n}{k} x^{n-k} y^k \\ &= \binom{n}{0} x^n + \binom{n}{1} x^{n-1} y + \binom{n}{2} x^{n-2} y^2 + \cdots + \binom{n}{n-1} x y^{n-1} + \binom{n}{n} y^n.\end{aligned}$$

Beweis. Wir beweisen die Identität durch vollständige Induktion. Für $n = 0$ ergibt sich

$$(x+y)^0 = 1 = \binom{0}{0},$$

was offensichtlich richtig ist. Wir nehmen daher nun an, dass

$$(x+y)^n = \binom{n}{0} x^n + \binom{n}{1} x^{n-1} y + \binom{n}{2} x^{n-2} y^2 + \cdots + \binom{n}{n-1} x y^{n-1} + \binom{n}{n} y^n$$

für eine gewisse natürliche Zahl n gilt. Dann folgt

$$
\begin{aligned}
(x+y)^{n+1} &= (x+y)^n(x+y) \\
&= \left(\binom{n}{0}x^n + \binom{n}{1}x^{n-1}y + \binom{n}{2}x^{n-2}y^2 + \cdots + \binom{n}{n-1}xy^{n-1} + \binom{n}{n}y^n\right)(x+y) \\
&= \binom{n}{0}x^{n+1} + \binom{n}{1}x^n y + \binom{n}{2}x^{n-1}y^2 + \cdots + \binom{n}{n-1}x^2y^{n-1} + \binom{n}{n}xy^n \\
&\quad + \binom{n}{0}x^n y + \binom{n}{1}x^{n-1}y^2 + \cdots + \binom{n}{n-2}x^2y^{n-1} + \binom{n}{n-1}xy^n + \binom{n}{n}y^{n+1}.
\end{aligned}
$$

Nun kombinieren wir Terme mit den gleichen x- und y-Potenzen:

$$
\begin{aligned}
\binom{n}{1}x^n y + \binom{n}{0}x^n y &= \binom{n+1}{1}x^n y, \\
\binom{n}{2}x^{n-1}y^2 + \binom{n}{1}x^{n-1}y^2 &= \binom{n+1}{2}x^{n-1}y^2,
\end{aligned}
$$

und allgemein

$$
\binom{n}{k+1}x^{n-k}y^{k+1} + \binom{n}{k}x^{n-k}y^{k+1} = \binom{n+1}{k+1}x^{n-k}y^{k+1},
$$

was aus der Identität (1.5) für die Binomialkoeffizienten folgt (dabei wurden n und k durch $n+1$ und $k+1$ ersetzt). Der erste Term ist zudem

$$
\binom{n}{0}x^{n+1} = x^{n+1} = \binom{n+1}{0}x^{n+1},
$$

und der letzte ist

$$
\binom{n}{n}y^{n+1} = y^{n+1} = \binom{n+1}{n+1}y^{n+1}.
$$

Zusammengefasst ergibt sich also

$$
(x+y)^{n+1} = \binom{n+1}{0}x^{n+1} + \binom{n+1}{1}x^n y + \binom{n+1}{2}x^{n-1}y^2 + \cdots + \binom{n+1}{n}xy^n + \binom{n+1}{n+1}y^{n+1},
$$

die Formel gilt also auch für $n+1$. Damit ist der binomische Lehrsatz bewiesen.

Beispiel 1.25. Man bestimme den Ausdruck, der beim Ausmultiplizieren von $(2x+3y)^4$ entsteht!

Lösung. Nach dem binomischen Lehrsatz gilt

$$
\begin{aligned}
(2x+3y)^4 &= \binom{4}{0}(2x)^4 + \binom{4}{1}(2x)^3(3y) + \binom{4}{2}(2x)^2(3y)^2 + \binom{4}{3}(2x)(3y)^3 + \binom{4}{4}(3y)^4 \\
&= 16x^4 + 96x^3y + 216x^2y^2 + 216xy^3 + 81y^4.
\end{aligned}
$$

Beispiel 1.26. Wie lautet der Koeffizient von x^7, wenn $(x+2)^{10}$ ausmultipliziert wird?

Lösung. Nach dem binomischen Lehrsatz gilt

$$(x+2)^{10} = \sum_{k=0}^{n} \binom{n}{k} x^{10-k} 2^k.$$

Der Term mit x^7 ergibt sich also für $k = 3$, und der entsprechende Koeffizient ist demnach

$$\binom{10}{7} \cdot 2^3 = 960.$$

Aus dem binomischen Lehrsatz ergibt sich durch Einsetzen von $x = y = 1$ sofort die folgende Beobachtung:

Satz 1.27. *Es gilt*

$$\sum_{k=0}^{n} \binom{n}{k} = 2^n,$$

die Summe der Zahlen in der n-ten Reihe des Pascalschen Dreiecks ist also gleich 2^n.

Wir schließen dieses Kapitel mit einer Ungleichung für die zentralen Binomialkoeffizienten $\binom{2n}{n}$ ab (also die Zahlen $1, 2, 6, 20, \ldots$, die genau in der Mitte des Pascalschen Dreiecks stehen).

Beispiel 1.28. Man zeige, dass

$$\frac{4^n}{n+1} \leq \binom{2n}{n}$$

für alle natürlichen Zahlen n gilt.

Lösung. Wir wenden wieder vollständige Induktion an. Dabei ist die Ungleichung für $n = 0$ offensichtlich richtig, denn sie lautet in diesem Fall $1 \leq \binom{0}{0} = 1$.

Wir nehmen also an, die Ungleichung gelte für ein gewisses n. Unser Ziel ist es, $\frac{4^{n+1}}{(n+1)+1} = \frac{4^{n+1}}{n+2} \leq \binom{2(n+1)}{n+1}$ zu zeigen. Dazu müssen wir zunächst $\binom{2(n+1)}{n+1} = \binom{2n+2}{n+1}$ mit $\binom{2n}{n}$ in Verbindung bringen. Es gilt

$$\begin{aligned}
\binom{2n+2}{n+1} &= \frac{(2n+2)!}{(n+1)! \cdot (n+1)!} = \frac{(2n+1)! \cdot (2n+2)}{n! \cdot (n+1) \cdot n! \cdot (n+1)} = \frac{(2n)! \cdot (2n+1) \cdot (2n+2)}{n! \cdot n! \cdot (n+1)^2} \\
&= \frac{(2n)!}{n! \cdot n!} \cdot \frac{(2n+1)(2n+2)}{(n+1)^2} = \binom{2n}{n} \cdot \frac{2(2n+1)}{n+1}.
\end{aligned}$$

Zusammen mit der Induktionsannahme erhalten wir daraus

$$\binom{2n+2}{n+1} = \binom{2n}{n} \cdot \frac{2(2n+1)}{n+1} \geq \frac{4^n}{(n+1)} \cdot \frac{2(2n+1)}{n+1} = \frac{2(2n+1)4^n}{(n+1)^2}.$$

Wenn wir zeigen können, dass letzterer Ausdruck größer oder gleich $\frac{4^{n+1}}{n+2}$ ist, dann ist die Induktion vollständig. Die Ungleichung

$$\frac{2(2n+1)4^n}{(n+1)^2} \geq \frac{4^{n+1}}{n+2}$$

ist (indem man auf beiden Seiten durch 4^n dividiert und mit $(n+2)(n+1)^2$ multipliziert; beides ist positiv und ändert daher die Richtung der Ungleichung nicht) äquivalent zu

$$2(2n+1)(n+2) = 4n^2 + 10n + 4 \geq 4(n+1)^2 = 4n^2 + 8n + 4,$$

was offensichtlich richtig ist. Damit ist der Beweis abgeschlossen.

Kapitel 2

Folgen und Reihen

2.1 Folgen

Eine (unendliche) *Folge* $a_0, a_1, a_2, \ldots$ besteht aus sogenannten *Folgengliedern* a_n für jede natürliche Zahl n[1]. Sind die Folgenglieder reelle Zahlen, so spricht man von einer reellen Folge; sind sie komplexe Zahlen, von einer komplexen Folge.

Folgen können explizit durch eine Formel definiert werden: zum Beispiel wird durch $a_n = (n+2) \cdot 3^n$ eine Folge definiert, die mit den Gliedern $a_0 = 2$, $a_1 = 9$ und $a_2 = 36$ beginnt. Bekannte und wichtige Spezialfälle sind etwa die

- *arithmetische Folge*: $a_n = a_0 + dn$, und die
- *geometrische Folge*: $a_n = a_0 q^n$.

Eine andere Möglichkeit um Folgen zu definieren ist durch eine *Rekursion*, bei der jedes Folgenglied durch die vorigen Folgenglieder ausgedrückt wird: wir haben etwa in Beispiel 1.20 gezeigt, dass die Folge $a_n = 6^n + 9$ auch durch $a_0 = 10$ und $a_{n+1} = 6a_n - 45$ beschrieben werden kann. Ein anderes bekanntes Beispiel ist die sogenannte *Fibonaccifolge*, bei der jedes Folgenglied die Summe der beiden vorhergehenden ist:

$$F_0 = 0,\ F_1 = 1,\ F_{n+2} = F_{n+1} + F_n \text{ für } n \geq 0.$$

Damit ergibt sich $F_2 = 1 + 0 = 1$, $F_3 = 1 + 1 = 2$, $F_4 = 2 + 1 = 3$, $F_5 = 3 + 2 = 5$, und so weiter.

Arithmetische und geometrische Folgen lassen sich ebenfalls rekursiv beschreiben: so ist für eine arithmetische Folge $a_{n+1} = a_n + d$, für eine geometrische Folge $a_{n+1} = qa_n$. Wir betrachten als Nächstes einige Eigenschaften, die Folgen haben können.

Definition 2.1. Eine reelle Folge a_n heißt

- *nach oben beschränkt*, wenn es eine Zahl M gibt, sodass $a_n \leq M$ für alle $n \geq 0$.
- *nach unten beschränkt*, wenn es eine Zahl m gibt, sodass $a_n \geq m$ für alle $n \geq 0$.
- *monoton steigend*, wenn $a_n \geq a_{n-1}$ für alle $n \geq 1$.

[1] Der Anfang muss nicht immer bei 0 sein. Manche Folgen beginnen sinnvollerweise erst mit a_1 oder noch später.

- *monoton fallend*, wenn $a_n \leq a_{n-1}$ für alle $n \geq 1$.

Eine Folge, die sowohl nach oben als auch nach unten beschränkt ist, wird auch nur *beschränkt* genannt. Gelten sogar die Ungleichungen $a_n > a_{n-1}$ oder $a_n < a_{n-1}$, dann spricht man von einer *streng* monoton wachsenden bzw. fallenden Folge. Wir betrachten zunächst einige einfache Beispiele.

- Die Folge $a_n = n^2$ ($a_0 = 0$, $a_1 = 1$, $a_2 = 4$, $\ldots$) ist streng monoton steigend und nach unten, aber nicht nach oben beschränkt.

- Die Folge $a_n = (-1)^n$ ($a_0 = 1$, $a_1 = -1$, $a_2 = 1$, $\ldots$) ist nicht monoton, aber nach oben und unten beschränkt.

- Es sei $\lfloor x \rfloor$ die größte ganze Zahl, die $\leq x$ ist (also „x abgerundet“). Die Folge $a_n = \lfloor \frac{n}{2} \rfloor$ ($a_0 = a_1 = 0$, $a_2 = a_3 = 1$, $a_4 = a_5 = 2$, $\ldots$) ist monoton, aber nicht streng monoton steigend. Sie ist zudem nach unten, aber nicht nach oben beschränkt.

Ein etwas komplexeres Beispiel ist das folgende.

Beispiel 2.2. Man zeige, dass die Folge $a_n = \frac{n^2+1}{n^2+2n+4}$ beschränkt und streng monoton wachsend ist.

Lösung. Zunächst gilt offenkundig $a_n > 0$ für alle n, die Folge ist also nach unten beschränkt. Andererseits ist $n^2 + 1 < n^2 + 2n + 4$, also $a_n < 1$. Damit ist sie auch nach oben beschränkt. Schließlich prüfen wir Monotonie nach, indem wir die Differenz $a_n - a_{n-1}$ betrachten:

$$\begin{aligned} a_n - a_{n-1} &= \frac{n^2+1}{n^2+2n+4} - \frac{(n-1)^2+1}{(n-1)^2+2(n-1)+4} = \frac{n^2+1}{n^2+2n+4} - \frac{n^2-2n+2}{n^2+3} \\ &= \frac{(n^2+1)(n^2+3)-(n^2+2n+4)(n^2-2n+2)}{(n^2+2n+4)(n^2+3)} = \frac{2n^2+4n-5}{(n^2+2n+4)(n^2+3)}. \end{aligned}$$

Nun gilt für $n \geq 1$, dass der Zähler positiv ist: $2n^2 + 4n - 5 \geq 2 + 4 - 5 = 1$. Der Nenner ist das Produkt positiver Faktoren und damit ebenfalls positiv. Also gilt $a_{n+1} - a_n > 0$, was bedeutet, dass die Folge streng monoton steigend ist.

2.2 Grenzwerte

Informell gesprochen ist ein Grenzwert einer Folge ein Wert, dem sie sich annähert. Dies kann wie folgt definiert werden:

Definition 2.3. Eine reelle Zahl A heißt *Grenzwert* einer reellen Folge a_n, wenn es für jedes $\epsilon > 0$ ein N_ϵ gibt, sodass

$$|a_n - A| < \epsilon \text{ für alle } n \geq N_\epsilon.$$

In diesem Fall schreibt man $\lim_{n\to\infty} a_n = A$.

Egal wie klein man also die positive Zahl ϵ wählt: ab einem Punkt N_ϵ sind alle Glieder der Folge weniger als ϵ von A entfernt und liegen also zwischen $A-\epsilon$ und $A+\epsilon$, wie Abbildung 2.1 illustriert. Wir betrachten einige einfache Beispiele dazu.

- Die konstante Folge $a_n = 5$ hat den Grenzwert 5: $\lim_{n\to\infty} 5 = 5$.

- Die Folge $a_n = \frac{n-1}{n}$ ($a_1 = 0$, $a_2 = \frac{1}{2}$, $a_3 = \frac{2}{3}$, ...) nähert sich 1 an: $\lim_{n\to\infty} \frac{n-1}{n} = 1$.
- Die geometrische Folge 2^{-n} ($a_1 = \frac{1}{2}$, $a_2 = \frac{1}{4}$, $a_3 = \frac{1}{8}$, ...) hat den Grenzwert 0: $\lim_{n\to\infty} 2^{-n} = 0$.
- Die Folge $a_n = \frac{(-1)^n}{n^2}$ ($a_1 = -1$, $a_2 = \frac{1}{4}$, $a_3 = -\frac{1}{9}$, ...) hat ebenfalls den Grenzwert 0. Die Folgenglieder nähern sich allerdings im Gegensatz zum vorigen Beispiel von beiden Seiten an.

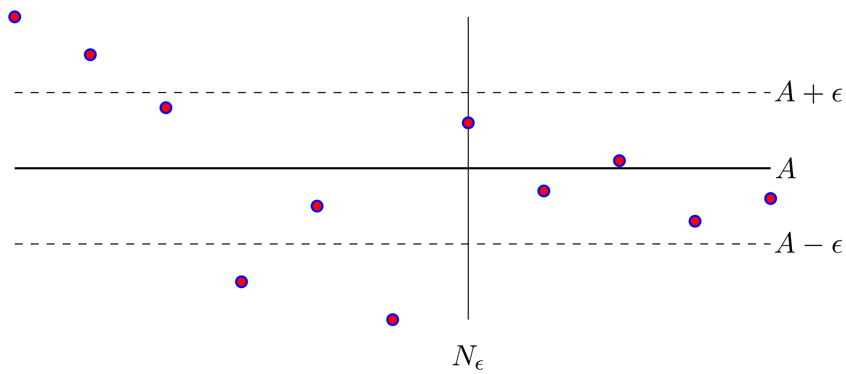

Abbildung 2.1: Zur Konvergenz von Folgen.

Eine Folge, die einen Grenzwert A hat, heißt *konvergent*. Eine Folge, die nicht konvergent ist, heißt dagegen *divergent*. Auch hier seien einige einfache Beispiele erwähnt:

- Die Folge $a_n = (-1)^n$ ist nicht konvergent: sie springt zwischen -1 und 1 hin und her, nähert sich aber nicht einem einzigen Wert an.
- Die Folge $a_n = n^2$ ist nicht konvergent, denn sie wächst über alle Grenzen.

Das zweite Beispiel ist ein Fall einer *bedingt divergenten* Folge: ihr Grenzwert ist ∞.

Definition 2.4. Eine reelle Folge a_n hat Grenzwert ∞, wenn es für jedes $M > 0$ ein N_M gibt, sodass

$$a_n > M \text{ für alle } n \geq N_M.$$

Eine reelle Folge a_n hat Grenzwert $-\infty$, wenn es für jedes $M < 0$ ein N_M gibt, sodass

$$a_n < M \text{ für alle } n \geq N_M.$$

Wenn ein Grenzwert existiert, so ist er eindeutig: eine Folge kann sich nicht gleichzeitig verschiedenen reellen Zahlen A_1 und A_2 annähern. Wir wählen zum Beweis in der Definition $\epsilon < \frac{|A_1 - A_2|}{2}$, sodass also der Abstand $|A_1 - A_2|$ größer als 2ϵ ist. Da die Intervalle $(A_1 - \epsilon, A_1 + \epsilon)$ und $(A_2 - \epsilon, A_2 + \epsilon)$ sich dann nicht überlappen, kann nicht gleichzeitig $|A_1 - a_n| < \epsilon$ und $|A_2 - a_n| < \epsilon$ gelten. Also kann nur höchstens einer der Werte A_1 und A_2 ein Grenzwert sein.

Die folgenden elementaren Grenzwerte werden im Folgenden nützlich sein:

- Für konstante Folgen ($a_n = c$ für alle n) gilt: $\lim_{n\to\infty} c = c$.

- Für Potenzen ($a_n = n^\alpha$) gilt:

$$\lim_{n\to\infty} n^\alpha = \begin{cases} \infty & \alpha > 0, \\ 1 & \alpha = 0, \\ 0 & \alpha < 0. \end{cases}$$

- Für geometrische Folgen der Form $a_n = q^n$ gilt:

$$\lim_{n\to\infty} q^n = \begin{cases} \infty & q > 1, \\ 1 & q = 1, \\ 0 & |q| < 1, \\ \text{divergent} & q \leq -1. \end{cases}$$

Mit Hilfe der folgenden Rechenregeln können Folgen kombiniert werden. Es seien a_n und b_n Folgen mit Grenzwerten $\lim_{n\to\infty} a_n = a$ und $\lim_{n\to\infty} b_n = b$. Dann gilt auch

- $\lim_{n\to\infty}(a_n \pm b_n) = a \pm b$.
- $\lim_{n\to\infty}(a_n \cdot b_n) = a \cdot b$; insbesondere gilt für eine Konstante k, dass $\lim_{n\to\infty}(ka_n) = ka$.
- $\lim_{n\to\infty} \frac{a_n}{b_n} = \frac{a}{b}$, sofern $b \neq 0$.
- falls $a_n \leq b_n$ für alle n, dann auch $a \leq b$.

Beispiel 2.5. Der Grenzwert der Folge $a_n = \frac{2n^3-n+5}{4n^3+2n^2-1}$ lässt sich bestimmen, indem man zunächst etwas umformt. Wir dividieren Zähler und Nenner durch die höchste vorkommende Potenz von n, also n^3. Dann gilt

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} \frac{2 - n^{-2} + 5n^{-3}}{4 + 2n^{-1} - n^{-3}}.$$

Da n^{-1}, n^{-2} und n^{-3} Grenzwert 0 haben, folgt aus den Rechenregeln

$$\lim_{n\to\infty} a_n = \frac{2 - 0 + 0}{4 + 0 - 0} = \frac{1}{2}.$$

Unendliche Grenzwerte können wie folgt charakterisiert werden:

Satz 2.6. *Für eine Folge a_n positiver reeller Zahlen gilt $\lim_{n\to\infty} a_n = \infty$ genau dann, wenn $\lim_{n\to\infty} \frac{1}{a_n} = 0$.*

So ist zum Beispiel $\lim_{n\to\infty} \sqrt{n} = \infty$ und $\lim_{n\to\infty} \frac{1}{\sqrt{n}} = 0$. Gleiches gilt auch für Folgen negativer reeller Zahlen: $\lim_{n\to\infty} a_n = -\infty$ genau dann, wenn $\lim_{n\to\infty} \frac{1}{a_n} = 0$.

Vorsicht ist mit „unbestimmten Ausdrücken" wie $\frac{0}{0}$, $\frac{\infty}{\infty}$ oder $\infty - \infty$ geboten (wie auch in Beispiel 2.5). Es folgt etwa aus $\lim_{n\to\infty} a_n = \infty$ und $\lim_{n\to\infty} b_n = \infty$ *nicht*, dass $\lim_{n\to\infty}(a_n - b_n) = 0$. Tatsächlich kann der Grenzwert in diesem Fall jede reelle Zahl oder auch $\pm\infty$ sein. Das folgende Beispiel illustriert dies.

Beispiel 2.7. Man bestimme den Grenzwert

$$\lim_{n\to\infty} \left(\sqrt{n^2 + 6n - 3} - n\right).$$

Lösung. Wir schreiben zunächst die Folge um, indem wir mit dem „konjugierten" Ausdruck $\sqrt{n^2+6n-3}+n$ multiplizieren und dividieren:

$$\begin{aligned}\sqrt{n^2+6n-3}-n &= \frac{(\sqrt{n^2+6n-3}-n)(\sqrt{n^2+6n-3}+n)}{\sqrt{n^2+6n-3}+n} = \frac{n^2+6n-3-n^2}{\sqrt{n^2+6n-3}+n} \\ &= \frac{6n-3}{\sqrt{n^2+6n-3}+n} = \frac{6-\frac{3}{n}}{\frac{1}{n}\sqrt{n^2+6n-3}+1} = \frac{6-\frac{3}{n}}{\sqrt{1+\frac{6}{n}-\frac{3}{n^2}}+1}.\end{aligned}$$

Nunmehr können wir von den Rechenregeln Gebrauch machen:

$$\lim_{n\to\infty}\sqrt{n^2+6n-3}-n = \lim_{n\to\infty}\frac{6-\frac{3}{n}}{\sqrt{1+\frac{6}{n}-\frac{3}{n^2}}+1} = \frac{6-0}{\sqrt{1}+1} = 3.$$

Oft kann man Grenzwerte von Folgen bestimmen, indem man sie zwischen zwei anderen Folgen einschließt. Das beruht auf dem folgenden Satz:

Satz 2.8. *Es seien a_n, b_n und c_n Folgen, für die $a_n \le c_n \le b_n$ (zumindest ab einem gewissen Punkt, also für $n \ge n_0$) gilt, und für die $\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n = A$ gilt. Dann ist auch*

$$\lim_{n\to\infty} c_n = A.$$

Beweis. Für jedes $\epsilon > 0$ gibt es Indices N_ϵ und N'_ϵ sodass $|a_n - A| < \epsilon$ für $n \ge N_\epsilon$ und $|b_n - A| < \epsilon$ für $n \ge N'_\epsilon$ gilt. Damit folgt für $n \ge \max(N_\epsilon, N'_\epsilon, n_0)$ die Ungleichungskette

$$A-\epsilon < a_n \le c_n \le b_n < A+\epsilon,$$

und damit $|c_n - A| < \epsilon$. Damit folgt aus der Grenzwertdefinition $\lim_{n\to\infty} c_n = A$.

Ein wichtiger Spezialfall ist jener, für den $a_n = -b_n$: $-b_n \le c_n \le b_n$ ist dann äquivalent zu $|c_n| \le b_n$.

Satz 2.9. *Gilt für zwei Folgen b_n und c_n, dass $|c_n| \le b_n$ (zumindest ab einem gewissen Punkt) und $\lim_{n\to\infty} b_n = 0$, dann auch $\lim_{n\to\infty} c_n = 0$.*

Beispiel 2.10. Für die Folge $a_n = \frac{2^n}{n!}$ gilt

$$0 < \frac{2^n}{n!} = \frac{2\cdot 2\cdots 2}{1\cdot 2\cdots n} = 2\cdot\frac{2}{2}\cdot\frac{2}{3}\cdots\frac{2}{n} \le 2\cdot\frac{2}{n} = \frac{4}{n}.$$

Also $|a_n| \le \frac{4}{n}$, und da der Grenzwert von $\frac{4}{n}$ gleich 0 ist, folgt $\lim_{n\to\infty}\frac{2^n}{n!} = 0$.

Beispiel 2.11. Für die Folge $a_n = \frac{\cos n}{\sqrt{n}}$ gilt

$$|a_n| = \frac{|\cos n|}{\sqrt{n}} \le \frac{1}{\sqrt{n}},$$

und da der Grenzwert von $\frac{1}{\sqrt{n}}$ gleich 0 ist, folgt wieder $\lim_{n\to\infty}\frac{\cos n}{\sqrt{n}} = 0$.

Eine weitere wichtige Methode um die Konvergenz von Folgen zu zeigen liegt im folgenden Satz begründet.

Satz 2.12. *Eine Folge a_n reeller Zahlen, die monoton wachsend und nach oben beschränkt ist, konvergiert. Ihr Grenzwert ist*

$$\lim_{n\to\infty} a_n = \sup_n a_n.$$

Ebenso ist eine Folge a_n reeller Zahlen, die monoton fallend und nach unten beschränkt ist, konvergent. Ihr Grenzwert ist

$$\lim_{n\to\infty} a_n = \inf_n a_n.$$

Beweis. Es sei a_n monoton wachsend und nach oben beschränkt, und es sei $A = \sup_n a_n$ das Supremum (das aufgrund der Beschränktheit endlich sein muss). Dann gilt $A \geq a_n$ für alle n. Für jedes $\epsilon > 0$ muss es ein N_ϵ geben, sodass $a_{N_\epsilon} > A - \epsilon$: andernfalls müsste $a_n \leq A - \epsilon$ für alle n gelten, womit $A - \epsilon$ eine kleinere obere Schranke als A wäre, im Widerspruch zur Definition des Supremums. Aufgrund der Monotonie der Folge folgt nun $A \geq a_n \geq a_{N_\epsilon} > A - \epsilon$ für alle $n \geq N_\epsilon$. Damit ist $|A - a_n| < \epsilon$ für alle $n \geq N_\epsilon$. Da dieses Argument für alle $\epsilon > 0$ funktioniert, folgt die Konvergenz der Folge gegen A. Der Beweis für monoton fallende Folgen, die nach unten beschränkt sind, ist analog.

Beispiel 2.13. Gegeben sei die Folge

$$a_0 = \sqrt{1},\ a_1 = \sqrt{1+\sqrt{1}},\ a_2 = \sqrt{1+\sqrt{1+\sqrt{1}}},\ a_3 = \sqrt{1+\sqrt{1+\sqrt{1+\sqrt{1}}}}, \ldots$$

Man zeige, dass sie konvergiert, und bestimme ihren Grenzwert.

Lösung. Die ersten Glieder dieser Folge sind $a_0 = 1$, $a_1 \approx 1.41421$, $a_2 \approx 1.55377$, $a_3 \approx 1.59805$. Man vermutet also, dass die Folge monoton steigend und nach oben durch 2 beschränkt ist. Wir zeigen beides durch vollständige Induktion.

Offenbar gilt $a_0 < a_1 < 2$. Damit ist der Induktionsanfang bereits erledigt, und wir setzen mit dem Induktionsschritt fort. Wir nehmen also an, dass $a_{n-1} < a_n$ und $a_n < 2$. Aus der Konstruktion der Folge ergibt sich, dass $a_n = \sqrt{1 + a_{n-1}}$ und $a_{n+1} = \sqrt{1 + a_n}$. Damit folgt

$$a_n = \sqrt{1 + a_{n-1}} < \sqrt{1 + a_n} = a_{n+1}$$

und

$$a_{n+1} = \sqrt{1 + a_n} < \sqrt{1 + 2} = \sqrt{3} < 2,$$

womit die Induktion vollständig ist.

Da die Folge monoton steigend und beschränkt ist, hat sie also einen Grenzwert A: $\lim_{n\to\infty} a_n = A$. Um diesen zu bestimmen, verwenden wir erneut die Gleichung

$$a_{n+1} = \sqrt{1 + a_n}.$$

Für $n \to \infty$ strebt die linke Seite nach A, die rechte nach $\sqrt{1 + A}$ (da a_{n+1} und a_n den Grenzwert A haben). Also erhalten wir die quadratische Gleichung

$$A = \sqrt{1 + A} \text{ bzw. } A^2 = 1 + A.$$

Die Lösungen sind $\frac{1\pm\sqrt{5}}{2}$. Die negative Lösung $\frac{1-\sqrt{5}}{2} < 0$ kommt nicht in Frage, also $A = \frac{1+\sqrt{5}}{2} \approx 1.61803$.

Zuletzt sei noch ein Kriterium für Konvergenz genannt, das vor allem theoretisch bedeutsam ist.

Satz 2.14 (Cauchysches Konvergenzkriterium)**.** *Eine reelle Folge a_n ist genau dann konvergent, wenn es für jedes $\epsilon > 0$ ein N_ϵ gibt, sodass*

$$|a_n - a_m| < \epsilon \text{ für alle } n, m \geq N_\epsilon.$$

Eine Folge, die das Kriterium erfüllt, wird auch *Cauchyfolge* genannt. Ein Vorteil des Kriteriums besteht darin, dass der Grenzwert dabei nicht bekannt sein muss. Es reicht zu zeigen, dass die Folgenglieder immer näher aneinander liegen.

2.3 Teilfolgen und Häufungswerte

Gegeben seien eine Folge a_n und eine streng monoton wachsende Folge natürlicher Zahlen $n_1, n_2, \ldots$. Die neue Folge $a_{n_1}, a_{n_2}, \ldots$ wird eine *Teilfolge* von a_n genannt. Einige Teilfolgen der Folge $a_n = 2n + 1$ der ungeraden Zahlen sind etwa:

- $a_{3n} = 6n + 1$ mit den Folgengliedern $1, 7, 13, 19, \ldots$;
- $a_{n^2} = 2n^2 + 1$ mit den Folgengliedern $1, 3, 9, 19, \ldots$;
- die Folge $3, 5, 7, 11, 13, 17, \ldots$ der ungeraden Primzahlen (hier ist $n_1 = 1$, $n_2 = 2$, $n_3 = 3$, $n_4 = 5$, und so weiter);

Ein *Häufungswert* oder *Häufungspunkt* einer Folge ist ein Wert, dem die Folge immer wieder beliebig nahe kommt. Formal werden Häufungswerte wie folgt definiert:

Definition 2.15. Eine reelle Zahl a heißt Häufungswert einer Folge a_n, wenn es für jedes $\epsilon > 0$ und jede natürliche Zahl n ein $N \geq n$ exisitert, sodass $|a_n - a| < \epsilon$.

Äquivalent zu dieser Definition ist: a ist Häufungswert einer Folge a_n, wenn es eine Teilfolge gibt, die gegen a konvergiert. Zunächst einige Beispiele.

- $a_n = \frac{(-1)^n n+3}{2n}$ kann in zwei Teilfolgen aufgeteilt werden: $a_n = \frac{n+3}{2n}$ (für gerades n) hat Grenzwert $\frac{1}{2}$, $a_n = \frac{-n+3}{2n}$ (für ungerades n) hat Grenzwert $-\frac{1}{2}$. Damit hat diese Folge die Häufungswerte $\frac{1}{2}$ und $-\frac{1}{2}$.
- $a_n = n^2$ strebt nach ∞ und hat keine endlichen Häufungswerte.
- Die Folge $1, 1, 2, 1, 2, 3, 1, 2, 3, 4, 1, 2, 3, 4, 5, 1, 2, \ldots$ hat unendlich viele Häufungswerte, nämlich alle positiven ganzen Zahlen. Da die Folge immer wieder zu 1 zurückkehrt und von Neuem mit $1, 2, 3, \ldots$ beginnt, wird jede positive ganze Zahl unendlich oft erreicht.
- Die Folge $a_n = \cos n$ hat alle Werte zwischen -1 und 1 als Häufungswerte (was aber nicht leicht zu beweisen ist!), wenn Winkel in Radianten gemessen werden. Messen wir die Winkel dagegen in Grad, so begint die Folge bei $a_{360} = a_0 = 1$ wieder von vorne. Damit hat sie als Häufungswerte die Zahlen $a_0 = 1$, $a_1 = a_{359} = \cos 1°$, $a_2 = a_{358} = \cos 2°$, $\ldots$, $a_{180} = -1$.

Satz 2.16 (Satz von Bolzano–Weierstraß)**.** *Jede beschränkte reelle Folge hat mindestens einen Häufungswert.*

Wenn man $\pm\infty$ hinzunimmt, dann hat sogar jede reelle Folge einen Häufungswert (unbeschränkte Folgen haben dann ∞ oder $-\infty$, oder auch beide, als Häufungswert). Konvergente Folgen sind genau diejenigen, die nur einen Häufungswert (nämlich den Grenzwert) haben:

Satz 2.17. *Eine reelle Folge ist genau dann konvergent, wenn es genau einen Häufungswert gibt.*

Definition 2.18. Der größte Häufungswert einer Folge wird *Limes superior* genannt, abgekürzt lim sup oder $\overline{\lim}$. Der kleinste Häufungswert einer Folge wird *Limes inferior* genannt, abgekürzt lim inf oder $\underline{\lim}$.

Wir betrachten wieder dieselben Beispiele wie zuvor.

- Für $a_n = \frac{(-1)^n n+3}{2n}$ ist $\limsup_{n\to\infty} a_n = \frac{1}{2}$ und $\liminf_{n\to\infty} a_n = -\frac{1}{2}$.
- Für $a_n = n^2$ ist $\limsup_{n\to\infty} a_n = \liminf_{n\to\infty} a_n = \infty$.
- Die Folge $1, 1, 2, 1, 2, 3, 1, 2, 3, 4, 1, 2, 3, 4, 5, 1, 2, \ldots$ hat Limes inferior 1 und Limes superior ∞.
- Die Folge $a_n = \cos n$ hat (sowohl in Radianten als auch in Grad) Limes inferior -1 und Limes superior 1.

Ein Vorteil von Limes superior und Limes inferior ist, dass sie im Gegensatz zum Grenzwert für alle Folgen existieren. Aus der Definition folgt unmittelbar, dass

$$\liminf_{n\to\infty} a_n \leq \limsup_{n\to\infty} a_n.$$

Aus Satz 2.17 erhält man zudem, dass die beiden genau dann gleich sind, wenn die Folge konvergiert. In diesem Fall ist

$$\liminf_{n\to\infty} a_n = \limsup_{n\to\infty} a_n = \lim_{n\to\infty} a_n.$$

2.4 Reihen

Unter einer *Reihe* versteht man eine unendliche Summe[2]

$$\sum_{n=0}^{\infty} a_n = a_0 + a_1 + a_2 + \cdots .$$

Um einem solchen Ausdruck Sinn zu geben, betrachten wir die Folge der *Partialsummen*

$$s_n = a_0 + a_1 + a_2 + \cdots + a_n.$$

Also ist $s_0 = a_0$, $s_1 = a_0 + a_1$, $s_2 = a_0 + a_1 + a_2$, und so weiter.

Definition 2.19. Eine Reihe $\sum_{n=0}^{\infty} a_n$ heißt *konvergent*, wenn die Folge der Partialsummen s_n konvergiert. Der Grenzwert $S = \lim_{n\to\infty} s_n$ wird in diesem Fall der *Wert* der Reihe genannt, und wir schreiben

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

Andernfalls heißt die Reihe *divergent*.

[2] So wie Folgen müssen Reihen nicht immer bei 0 beginnen.

Als erstes Beispiel sei die *geometrische Reihe* genannt.

Satz 2.20. *Für eine reelle Zahl q ist die Reihe $\sum_{n=0}^{\infty} q^n$*

- *konvergent falls $|q| < 1$. In diesem Fall ist der Wert der Reihe $\frac{1}{1-q}$.*
- *divergent falls $|q| \geq 1$.*

Beweis. Wir betrachten die Partialsummen

$$s_n = 1 + q + q^2 + \cdots + q^{n-1} + q^n.$$

Man beachte nun, dass sich

$$qs_n = q + q^2 + q^3 + \cdots + q^n + q^{n+1}$$

von s_n nur durch den ersten und letzten Term unterscheidet: subtrahieren wir die beiden, so ergibt sich

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

also $s_n = \frac{1-q^{n+1}}{1-q}$ (sofern $q \neq 1$). Falls nun $|q| < 1$, so gilt $\lim_{n\to\infty} q^{n+1} = 0$ und damit $\lim_{n\to\infty} s_n = \frac{1}{1-q}$. Für $q = 1$ ist $s_n = n + 1$, was gegen ∞ strebt. In allen anderen Fällen konvergiert q^{n+1} und damit auch s_n nicht.

Es gilt also zum Beispiel

$$\sum_{n=0}^{\infty} \frac{1}{2^n} = 1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \cdots = \frac{1}{1 - \frac{1}{2}} = 2.$$

Diese Identität wird in Abbildung 2.2 illustriert.

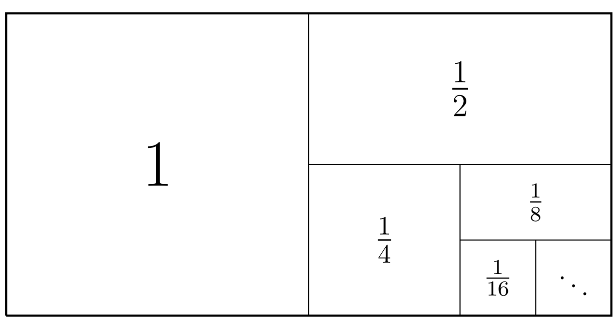

Abbildung 2.2: Zur geometrischen Reihe $1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \cdots$.

Ein weiteres Beispiel, bei dem die Partialsummen explizit bestimmt werden können, ist das folgende.

Beispiel 2.21. Man bestimme den Wert der Reihe

$$\sum_{n=1}^{\infty} \frac{1}{n(n+1)} = \frac{1}{1 \cdot 2} + \frac{1}{2 \cdot 3} + \frac{1}{3 \cdot 4} + \cdots .$$

Lösung. Wir betrachten die ersten Partialsummen:

$$s_1 = \frac{1}{2},\ s_2 = \frac{1}{2} + \frac{1}{6} = \frac{2}{3},\ s_3 = \frac{1}{2} + \frac{1}{6} + \frac{1}{12} = \frac{3}{4}.$$

Damit ergibt sich die Vermutung $s_n = \frac{n}{n+1}$, die sich durch Induktion beweisen lässt: der Induktionsanfang ist bereits gemacht (für $n = 1$), wir nehmen also an, dass $s_n = \frac{n}{n+1}$ ist. Die Partialsumme s_{n+1} ergibt sich, indem man $a_{n+1} = \frac{1}{(n+1)(n+2)}$ hinzufügt:

$$\begin{aligned} s_{n+1} = s_n + a_{n+1} &= \frac{n}{n+1} + \frac{1}{(n+1)(n+2)} = \frac{n(n+2)+1}{(n+1)(n+2)} \\ &= \frac{n^2+2n+1}{(n+1)(n+2)} = \frac{(n+1)^2}{(n+1)(n+2)} = \frac{n+1}{n+2}. \end{aligned}$$

Damit ist die Induktion abgeschlossen, und wir können den Grenzwert der Partialsummen bestimmen:

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

Eine andere Möglichkeit zur Bestimmung der Partialsummen ist die Beobachtung, dass

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

Damit ergibt sich

$$s_n = a_1 + a_2 + \cdots + a_n = \left(\frac{1}{1} - \frac{1}{2}\right) + \left(\frac{1}{2} - \frac{1}{3}\right) + \left(\frac{1}{3} - \frac{1}{4}\right) + \cdots + \left(\frac{1}{n} - \frac{1}{n+1}\right).$$

Die Terme kürzen sich paarweise (man spricht von einer sogenannten *Teleskopsumme*), und es bleibt nur

$$s_n = 1 - \frac{1}{n+1}$$

übrig. Also erhalten wir wieder

$$\sum_{n=1}^{\infty} \frac{1}{n(n+1)} = 1,$$

die Reihe ist also konvergent.

Zwei einfache Beispiele divergenter Reihen sind

- $\sum_{n=1}^{\infty} 1 = 1 + 1 + \cdots$: hier ist $s_n = n$, was gegen ∞ strebt;
- $\sum_{n=0}^{\infty} (-1)^n = 1 - 1 + 1 - 1 + 1 - \cdots$: hier sind die Partialsummen abwechselnd 1 und 0, konvergieren also wiederum nicht.

Ein etwas komplizierteres Beispiel ist die *harmonische Reihe*

$$\sum_{n=1}^{\infty} \frac{1}{n} = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots.$$

Um zu zeigen, dass diese Reihe divergiert, bilden wir die Partialsummen s_{2^n} für Zweierpotenzen. Dabei fassen wir für jedes $k \geq 1$ die Terme von $\frac{1}{2^k+1}$ bis $\frac{1}{2^{k+1}}$ zusammen. So ist

$$\frac{1}{3}+\frac{1}{4} \geq \frac{1}{4}+\frac{1}{4}=\frac{1}{2},$$
$$\frac{1}{5}+\frac{1}{6}+\frac{1}{7}+\frac{1}{8} \geq \frac{1}{8}+\frac{1}{8}+\frac{1}{8}+\frac{1}{8}=\frac{1}{2},$$
$$\frac{1}{9}+\frac{1}{10}+\cdots+\frac{1}{16} \geq \frac{1}{16}+\frac{1}{16}+\cdots+\frac{1}{16}=\frac{1}{2},$$

und so weiter. Damit erhalten wir

$$\begin{aligned} a_{2^n} &= 1+\frac{1}{2}+\left(\frac{1}{3}+\frac{1}{4}\right)+\left(\frac{1}{5}+\frac{1}{6}+\frac{1}{7}+\frac{1}{8}\right)+\cdots+\left(\frac{1}{2^{n-1}+1}+\cdots+\frac{1}{2^n}\right) \\ &\geq 1+\frac{1}{2}+\frac{1}{2}+\cdots+\frac{1}{2}=1+\frac{n}{2}. \end{aligned}$$

Da $1+\frac{n}{2}$ gegen ∞ strebt, muss dies damit auch für die Partialsummen der Reihe gelten. Daher ist die harmonische Reihe divergent.

2.5 Konvergenzkriterien

Nicht immer kann man für eine gegebene Reihe direkt mit den Partialsummen arbeiten. Wir sind daher an allgemeinen Kriterien interessiert, mit denen man Konvergenz oder Divergenz einer Reihe bestimmen kann. Ein erstes Kriterium ist, dass die Terme einer konvergenten Reihe den Grenzwert 0 haben müssen.

Satz 2.22. *Ist die Reihe* $\sum_{n=0}^{\infty} a_n$ *konvergent, so gilt* $\lim_{n\to\infty} a_n = 0$.

Beweis. Ist die Reihe konvergent, so müssen die Partialsummen $s_n = a_0 + a_1 + \cdots + a_n$ einen Grenzwert S haben. Nun bemerken wir, dass sich aufeinanderfolgende Partialsummen s_n und s_{n-1} nur durch den Term a_n unterscheiden:

$$s_n = a_0 + a_1 + \cdots + a_{n-1} + a_n = (a_0 + a_1 + \cdots + a_{n-1}) + a_n = s_{n-1} + a_n,$$

daher $a_n = s_n - s_{n-1}$. Für $n \to \infty$ streben s_n und s_{n-1} gegen S, und daher ist

$$\lim_{n\to\infty} a_n = \left(\lim_{n\to\infty} s_n\right) - \left(\lim_{n\to\infty} s_{n-1}\right) = S - S = 0.$$

Bemerkung 2.23. Das Beispiel der harmonischen Reihe zeigt, dass nicht jede Reihe, deren Terme gegen 0 gehen, konvergent ist. Die Bedingung ist also notwendig, aber nicht hinreichend.

Beispiel 2.24. Wir betrachten die Reihe $\sum_{n=0}^{\infty} \frac{2n+3}{4n+1}$. Da in diesem Fall der Grenzwert

$$\lim_{n\to\infty} \frac{2n+3}{4n+1} = \lim_{n\to\infty} \frac{2+\frac{3}{n}}{4+\frac{1}{n}} = \frac{2+0}{4+0} = \frac{1}{2}$$

nicht 0 ist, muss die Reihe divergent sein.

Definition 2.25. Eine Reihe $\sum_{n=0}^{\infty} a_n$ wird *absolut konvergent* genannt, wenn $\sum_{n=0}^{\infty} |a_n|$ konvergiert.

Absolute Konvergenz impliziert Konvergenz, wie der folgende Satz zeigt.

Satz 2.26. *Eine absolut konvergente Reihe ist konvergent.*

Beweis. Es sei $\sum_{n=0}^{\infty} |a_n|$ konvergent. Dann sind die Partialsummen dieser Reihe beschränkt (nämlich durch den Wert der Reihe):

$$|a_0| + |a_1| + \cdots + |a_n| \leq M \tag{2.1}$$

für eine Konstante M. Nun betrachten wir die Partialsummen der Reihe $\sum_{n=0}^{\infty} a_n$:

$$a_0 + a_1 + \cdots + a_n = |a_0| + |a_1| + \cdots + |a_n| - (|a_0| - a_0) - (|a_1| - a_1) - \cdots - (|a_n| - a_n).$$

Wir können $s_n = a_0 + a_1 + \cdots + a_n$ also als Differenz zweier Folgen schreiben: eine davon ist die Folge der Partialsummen von $\sum_{n=0}^{\infty} |a_n|$, deren Konvergenz vorausgesetzt wird. Es reicht also nachzuweisen, dass auch die Folge

$$t_n = (|a_0| - a_0) + (|a_1| - a_1) + \cdots + (|a_n| - a_n)$$

konvergiert. Dazu stellen wir zunächst fest, dass die Summanden

$$|a_k| - a_k = \begin{cases} 0 & a_k \geq 0, \\ 2|a_k| & a_k < 0, \end{cases}$$

alle ≥ 0 sind. Daher ist t_n monoton steigend. Andererseits ist wegen (2.1)

$$t_n \leq 2|a_0| + 2|a_1| + \cdots + 2|a_n| \leq 2M,$$

also ist t_n auch beschränkt. Daher konvergiert diese Folge, und damit auch s_n.

Beispiel 2.27. Da bereits gezeigt wurde, dass $\sum_{n=1}^{\infty} \frac{1}{n(n+1)}$ eine konvergente Reihe ist, muss $\sum_{n=1}^{\infty} \frac{(-1)^n}{n(n+1)}$ ebenfalls konvergieren, denn es ist

$$\left|\frac{(-1)^n}{n(n+1)}\right| = \left| \pm \frac{1}{n(n+1)}\right| = \frac{1}{n(n+1)}.$$

Eine häufig anwendbare Methode um Konvergenz oder Divergenz von Reihen zu beweisen besteht im Vergleich mit anderen Reihen.

- Eine Reihe $\sum_{n=0}^{\infty} b_n$ ist *Majorante* einer anderen Reihe $\sum_{n=0}^{\infty} a_n$, wenn $|a_n| \leq b_n$ für alle n (zumindest ab einem bestimmten Punkt) gilt.

- Eine Reihe $\sum_{n=0}^{\infty} b_n$ ist *Minorante* einer anderen Reihe $\sum_{n=0}^{\infty} a_n$, wenn $0 \leq b_n \leq a_n$ für alle n (zumindest ab einem bestimmten Punkt) gilt.

Satz 2.28 (Majoranten- und Minorantenkriterium)**.** *Wenn eine Reihe eine konvergente Majorante hat, dann ist sie selbst (absolut) konvergent. Wenn eine Reihe eine divergente Minorante hat, dann ist sie selbst divergent.*

Beweis. Angenommen, $\sum_{n=0}^{\infty} b_n$ ist konvergent, und es gilt $|a_n| \leq b_n$ für alle n. Da $|a_n| \geq 0$ für alle n gilt, sind die Partialsummen von $\sum_{n=0}^{\infty} |a_n|$ monoton wachsend. Andererseits sind sie durch die Partialsummen von $\sum_{n=0}^{\infty} b_n$ (und damit deren Grenzwert) beschränkt. Daher konvergeren die Partialsummen von $\sum_{n=0}^{\infty} |a_n|$, und die Reihe $\sum_{n=0}^{\infty} a_n$ ist damit absolut konvergent.

Sei andererseits $\sum_{n=0}^{\infty} b_n$ eine divergente Reihe nichtnegativer Zahlen b_n. Die Partialsummen sind monoton wachsend und können daher nur dann divergieren, wenn sie unbeschränkt sind. Gilt $a_n \geq b_n$, dann sind die Partialsummen von $\sum_{n=0}^{\infty} a_n$ mindestens so groß wie jene von $\sum_{n=0}^{\infty} b_n$ und damit ebenfalls unbeschränkt. Damit divergiert die Reihe $\sum_{n=0}^{\infty} a_n$.

In beiden Fällen ist es ausreichend, wenn die Ungleichung für $n \geq n_0$ gilt. Ändert man nämlich die Werte von a_n für $n < n_0$, so ändern sich die Partialsummen nur um eine Konstante, wodurch sich aber nichts an Konvergenz oder Divergenz ändert.

Typische Reihen, mit denen das Majoranten- oder Minorantenkriterium angewandt wird, sind

1. die Reihe $\sum_{n=1}^{\infty} \frac{1}{n^\alpha}$, die für $\alpha > 1$ konvergiert und für $\alpha \leq 1$ divergiert. Dies wird später mit Hilfe des sogenannten *Integralkriteriums* verifiziert. Die harmonische Reihe ist ein Spezialfall.
2. die geometrische Reihe $\sum_{n=0}^{\infty} q^n$, die für $0 < q < 1$ konvergiert und für $q \geq 1$ divergiert.

Oft verwendet man auch konstante Vielfache dieser Reihen.

Beispiel 2.29. Man bestimme, ob die folgenden Reihen konvergieren oder divergieren:

- $\sum_{n=0}^{\infty} \frac{\cos n}{2^n}$,
- $\sum_{n=0}^{\infty} \frac{3\cdot(-2)^n}{(n^2+1)3^n}$,
- $\sum_{n=0}^{\infty} \frac{n+3}{n^{3/2}+2}$.

Lösung.

- Für die erste Reihe gilt
$$\left|\frac{\cos n}{2^n}\right| \leq \frac{1}{2^n}.$$
Daher ist die konvergente geometrische Reihe $\sum_{n=0}^{\infty} \frac{1}{2^n}$ eine Majorante, womit die Reihe absolut konvergiert.
- Für die zweite Reihe haben wir ebenfalls eine geometrische Reihe als Majorante:
$$\left|\frac{3\cdot(-2)^n}{(n^2+1)3^n}\right| = \frac{3\cdot 2^n}{(n^2+1)3^n} \leq 3\Big(\frac{2}{3}\Big)^n.$$
Damit ist die geometrische Reihe $\sum_{n=0}^{\infty} 3 \cdot (\frac{2}{3})^n$ eine konvergente Majorante, und auch die zweite Reihe ist damit absolut konvergent.
- Die dritte Reihe hat eine divergente Minorante: für $n \geq 1$ gilt
$$\frac{n+3}{n^{3/2}+2} \geq \frac{n}{n^{3/2}+2n^{3/2}} = \frac{1}{3\sqrt{n}}.$$
Da die Reihe $\sum_{n=1}^{\infty} \frac{1}{3\sqrt{n}}$ divergiert, divergiert auch die ursprüngliche Reihe.

Bemerkung 2.30. Der exakte Wert der ersten Reihe wird später in Beispiel 3.40 berechnet.

Als nächstes behandeln wir das *Quotientenkriterium* und das *Wurzelkriterium*, die sich durch den Vergleich von Reihen mit geometrischen Reihen ergeben.

Satz 2.31 (Quotientenkriterium). *Gegeben sei eine Reihe* $\sum_{n=0}^{\infty} a_n$.

- *Wenn* $\limsup_{n\to\infty} |\frac{a_{n+1}}{a_n}| < 1$, *dann ist die Reihe absolut konvergent.*
- *Wenn* $\liminf_{n\to\infty} |\frac{a_{n+1}}{a_n}| > 1$, *dann ist die Reihe divergent.*

Beweis. Es gelte zunächst $Q = \limsup_{n\to\infty} |\frac{a_{n+1}}{a_n}| < 1$. Wähle ein $\epsilon > 0$ mit $\epsilon < 1 - Q$. Es muss ein n_0 geben, sodass für $n \geq n_0$ die Ungleichung $|\frac{a_{n+1}}{a_n}| \leq Q + \epsilon$ gilt, weil sonst $\limsup_{n\to\infty} |\frac{a_{n+1}}{a_n}| \geq Q + \epsilon$ wäre. Aus $|\frac{a_{n_0+1}}{a_{n_0}}| \leq Q + \epsilon$ folgt nun

$$|a_{n_0+1}| \leq (Q + \epsilon)|a_{n_0}|,$$

aus $|\frac{a_{n_0+2}}{a_{n_0+1}}| \leq Q + \epsilon$ folgt

$$|a_{n_0+2}| \leq (Q + \epsilon)|a_{n_0+1}| \leq (Q + \epsilon)^2|a_{n_0}|,$$

und so weiter: allgemein ist $|a_{n_0+k}| \leq (Q + \epsilon)^k|a_{n_0}|$. Damit ist aber die geometrische Reihe $\sum_{k=0}^{\infty}(Q+\epsilon)^k|a_{n_0}|$ eine Majorante von $\sum_{k=0}^{\infty} a_{n_0+k} = \sum_{n=n_0}^{\infty} a_n$. Aufgrund der Wahl von ϵ ist $Q+\epsilon < 1$, es handelt sich also um eine konvergente Majorante. Damit muss $\sum_{n=n_0}^{\infty} a_n$ absolut konvergent sein, und die endliche Summe $\sum_{n=0}^{n_0-1} a_n$ ändert an der Konvergenz nichts.

Es gelte nun $q = \liminf_{n\to\infty} |\frac{a_{n+1}}{a_n}| > 1$. Wähle ein $\epsilon > 0$ mit $\epsilon > q - 1$. Es muss ein n_0 geben, sodass für $n \geq n_0$ die Ungleichung $|\frac{a_{n+1}}{a_n}| \geq q - \epsilon$ gilt, weil sonst $\liminf_{n\to\infty} |\frac{a_{n+1}}{a_n}| \leq q - \epsilon$ wäre. Wie im ersten Fall erhalten wir damit $|a_{n_0+k}| \geq (q - \epsilon)^k|a_{n_0}|$. Damit müsste (wegen $q - \epsilon > 1$) $\lim_{k\to\infty} |a_{n_0+k}| = \infty$ gelten, die Terme der Reihe haben also insbesondere nicht Grenzwert 0. Nach Satz 2.22 ist die Reihe daher divergent.

Bemerkung 2.32. Wenn der Grenzwert $Q = \lim_{n\to\infty} |\frac{a_{n+1}}{a_n}|$ existiert, dann lautet das Kriterium:

- $Q < 1$: die Reihe konvergiert.
- $Q > 1$: die Reihe divergiert.
- $Q = 1$: beide Fälle sind möglich – das Quotientenkriterium liefert kein eindeutiges Ergebnis.

Beispiel 2.33. Man bestimme, ob die folgenden Reihen konvergieren oder divergieren:

- $\sum_{n=0}^{\infty} \frac{n!}{10^n}$,
- $\sum_{n=0}^{\infty} \binom{2n}{n}(-5)^{-n}$.

Lösung.

- Für die erste Reihe ist $a_n = \frac{n!}{10^n}$. Wir berechnen den Grenzwert von $|\frac{a_{n+1}}{a_n}|$:

$$\lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| = \lim_{n\to\infty} \frac{\frac{(n+1)!}{10^{n+1}}}{\frac{n!}{10^n}} = \lim_{n\to\infty} \frac{n+1}{10} = \infty.$$

  Damit ist die Reihe divergent.

- Für die zweite Reihe ist $a_n = \binom{2n}{n}(-5)^{-n}$, also $|a_n| = \binom{2n}{n}5^{-n} = \frac{(2n)!}{5^n \cdot n! \cdot n!}$ und damit

$$\begin{aligned}|a_{n+1}| &= \frac{(2n+2)!}{5^{n+1}\cdot(n+1)!\cdot(n+1)!} = \frac{(2n)!\cdot(2n+1)\cdot(2n+2)}{5^{n+1}\cdot n!\cdot(n+1)\cdot n!\cdot(n+1)} \\ &= \frac{(2n)!}{5^n\cdot n!\cdot n!}\cdot\frac{(2n+1)\cdot(2n+2)}{5\cdot(n+1)^2} = |a_n|\cdot\frac{2(2n+1)(n+1)}{5\cdot(n+1)^2} \\ &= |a_n|\cdot\frac{2(2n+1)}{5(n+1)} = |a_n|\cdot\frac{4n+2}{5n+5}.\end{aligned}$$

  Es folgt, dass

$$\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| = \lim_{n\to\infty}\frac{4n+2}{5n+5} = \lim_{n\to\infty}\frac{4+\frac{2}{n}}{5+\frac{5}{n}} = \frac{4}{5} < 1.$$

  Daher konvergiert die Reihe.

Satz 2.34 (Wurzelkriterium)**.** *Gegeben sei eine Reihe* $\sum_{n=0}^{\infty} a_n$.

- *Wenn* $\limsup_{n\to\infty}\sqrt[n]{|a_n|} < 1$, *dann ist die Reihe absolut konvergent.*
- *Wenn* $\limsup_{n\to\infty}\sqrt[n]{|a_n|} > 1$, *dann ist die Reihe divergent.*

Beweis. Es gelte zunächst $Q = \limsup_{n\to\infty}\sqrt[n]{|a_n|} < 1$. Wähle ein $\epsilon > 0$ mit $\epsilon < 1 - Q$. Wie im Beweis des Quotientenkriteriums muss es ein n_0 geben, sodass für $n \geq n_0$ die Ungleichung $\sqrt[n]{|a_n|} \leq Q + \epsilon < 1$ gilt. Damit ist $|a_n| \leq (Q+\epsilon)^n$ für $n \geq n_0$, womit die geometrische Reihe $\sum_{n=n_0}^{\infty}(Q+\epsilon)^n$ eine konvergente Majorante von $\sum_{n=n_0}^{\infty} a_n$ ist. Wie im Beweis des Quotientenkriteriums folgt die absolute Konvergenz.

Es gelte nun $Q = \limsup_{n\to\infty}\sqrt[n]{|a_n|} > 1$. Dazu muss $\sqrt[n]{|a_n|} \geq 1$ für unendlich viele Werte von n gelten, weil sonst $\limsup_{n\to\infty}\sqrt[n]{|a_n|} \leq 1$ wäre. Daher ist $|a_n| \geq 1$ für unendlich viele n, womit der Grenzwert von a_n nicht 0 sein kann. Wieder folgt die Divergenz aus Satz 2.22.

Bemerkung 2.35. Wenn $\limsup_{n\to\infty}\sqrt[n]{|a_n|} = 1$, dann liefert das Wurzelkriterium kein eindeutiges Ergebnis.

Beispiel 2.36. Man bestimme, ob die folgende Reihe konvergiert oder divergiert:

$$\sum_{n=0}^{\infty}(-1)^n\left(\frac{4n+3}{3n+2}\right)^{n/2}.$$

Lösung. Hier ist $a_n = (-1)^n(\frac{4n+3}{3n+2})^{n/2}$, daher $|a_n| = (\frac{4n+3}{3n+2})^{n/2}$ und schließlich $\sqrt[n]{|a_n|} = (\frac{4n+3}{3n+2})^{n/(2n)} = \sqrt{\frac{4n+3}{3n+2}}$. Der Grenzwert ist

$$\lim_{n\to\infty}\sqrt[n]{|a_n|} = \lim_{n\to\infty}\sqrt{\frac{4n+3}{3n+2}} = \lim_{n\to\infty}\sqrt{\frac{4+3/n}{3+2/n}} = \sqrt{\frac{4+0}{3+0}} = \frac{2}{\sqrt{3}} > 1.$$

Die Reihe divergiert also.

Zuletzt wollen wir noch eine spezielle Art von Reihen betrachten: sogenannte *alternierende* Reihen, also Reihen der Form

$$\sum_{n=0}^{\infty}(-1)^n a_n = a_0 - a_1 + a_2 - a_3 + a_4 - \cdots.$$

Das folgende Kriterium kann häufig zum Beweis der Konvergenz von alternierenden Reihen verwendet werden.

Satz 2.37 (Leibniz-Kriterium). *Ist a_n eine monoton fallende Folge positiver reeller Zahlen, und gilt zudem $\lim_{n\to\infty} a_n = 0$, dann konvergiert die alternierende Reihe $\sum_{n=0}^{\infty}(-1)^n a_n$. Zudem gilt folgende Ungleichung zwischen dem Wert S der Reihe und den Partialsummen s_n:*

$$|S - s_n| \le a_{n+1}.$$

Beweis. Wir betrachten die Partialsummen $s_n = a_0 - a_1 + a_2 - \cdots \pm a_n$. Für gerades n ist $s_n = s_{n-1} + a_n > s_{n-1}$, während für ungerades n stattdessen $s_n = s_{n-1} - a_n < s_{n-1}$ gilt. Darüber hinaus ist für gerades n aufgrund der Monotonie

$$s_n = s_{n-2} - a_{n-1} + a_n \le s_{n-2}$$

und für ungerades n

$$s_n = s_{n-2} + a_{n-1} - a_n \ge s_{n-2}.$$

Die Partialsummen folgen daher folgendem Muster:

$$s_0 \ge s_2 \ge s_4 \ge \cdots \ge s_5 \ge s_3 \ge s_1.$$

Die Partialsummen mit geradem Index bilden also eine monoton fallende, nach unten beschränkte Folge, und die Partialsummen mit ungeradem Index bilden eine monoton steigende, nach oben beschränkte Folge. Beide haben damit Grenzwerte: S_{gerade} und S_{ungerade}. Weil die Differenz $|s_{n-1} - s_n| = a_n$ nach Voraussetzung nach 0 strebt, müssen die beiden Grenzwerte gleich sein:

$$S_{\text{gerade}} = S_{\text{ungerade}} = S.$$

Der Grenzwert S muss im Intervall zwischen s_n und s_{n+1} liegen, und es gilt somit

$$|S - s_n| \le |s_{n+1} - s_n| = a_{n+1},$$

womit der Beweis vollständig ist.

Bemerkung 2.38. Ähnlich wie im Majoranten- und Minorantenkriterium ist es ausreichend, dass die Monotoniebedingung ab einem gewissen Punkt (also für $n \ge n_0$) gilt, denn wenn man die Terme für $n < n_0$ entfernt, dann ändern sich die Partialsummen nur um eine Konstante, womit sich nichts an der Konvergenz ändert.

Beispiel 2.39. Man zeige, dass die folgende Reihe konvergiert:

$$\sum_{n=0}^{\infty} \frac{(-1)^n n}{n^2+3}.$$

Lösung. Wir weisen nach, dass die Folge $a_n = \frac{n}{n^2+3}$ die Bedingungen erfüllt.

- Die Differenz aufeinanderfolgender Folgenglieder ist

$$\begin{aligned} a_n - a_{n+1} &= \frac{n}{n^2+3} - \frac{n+1}{(n+1)^2+3} = \frac{n}{n^2+3} - \frac{n+1}{n^2+2n+4} \\ &= \frac{n(n^2+2n+4) - (n+1)(n^2+3)}{(n^2+3)(n^2+2n+4)} = \frac{n^2+n-3}{(n^2+3)(n^2+2n+4)}. \end{aligned}$$

  Da für $n \ge 2$ der Bruch positiv ist, folgt, dass die Folge a_n ab $n = 2$ monoton fallend ist.

- Außerdem gilt

$$\lim_{n\to\infty} \frac{n}{n^2+3} = \lim_{n\to\infty} \frac{\frac{1}{n}}{1+\frac{3}{n^2}} = \frac{0}{1+0} = 0.$$

Damit sind die Bedingungen des Leibniz-Kriteriums erfüllt, und die Reihe konvergiert.

2.6 Rechenregeln

Es gelten die folgenden Rechenregeln, die sich aus den entsprechenden Regeln für Folgen ableiten lassen.

Satz 2.40. *Sind* $\sum_{n=0}^{\infty} a_n = A$ *und* $\sum_{n=0}^{\infty} b_n = B$ *konvergente Reihen, dann sind auch die folgenden Reihen konvergent:*

- $\sum_{n=0}^{\infty}(a_n \pm b_n) = A \pm B$,
- $\sum_{n=0}^{\infty} ka_n = kA$ *für einen konstanten Faktor* k.

Gilt $a_n \leq b_n$ *für alle* n, *dann auch* $A \leq B$.

Beispiel 2.41. Aus den bekannten Reihen $\sum_{n=1}^{\infty} \frac{1}{3^n} = \frac{1}{2}$ (geometrische Reihe[3]) und $\sum_{n=1}^{\infty} \frac{1}{n(n+1)} = 1$ (siehe Beispiel 2.21) folgt etwa

$$\sum_{n=1}^{\infty} \left(\frac{4}{3^n} - \frac{5}{n(n+1)} \right) = \frac{4}{2} - 5 = -3.$$

Es ist jedoch *nicht* $\sum_{n=0}^{\infty} a_n \cdot \sum_{n=0}^{\infty} b_n = \sum_{n=0}^{\infty} a_n b_n$! Für Produkte von Reihen gilt eine andere Regel. Die folgende Tabelle illustriert dies, indem alle Produkte aufgelistet werden, die beim Ausmultiplizieren entstehen.

	b_0	b_1	b_2	b_3	b_4	$\cdots$
a_0	a_0b_0	a_0b_1	a_0b_2	a_0b_3	a_0b_4	$\cdots$
a_1	a_1b_0	a_1b_1	a_1b_2	a_1b_3	a_1b_4	$\cdots$
a_2	a_2b_0	a_2b_1	a_2b_2	a_2b_3	a_2b_4	$\cdots$
a_3	a_3b_0	a_3b_1	a_3b_2	a_3b_3	a_3b_4	$\cdots$
a_4	a_4b_0	a_4b_1	a_4b_2	a_4b_3	a_4b_4	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$

Fasst man Produkte $a_i b_j$ nach der Summe $i + j$ der Indices zusammen, erhält man $c_0 = a_0 b_0$, $c_1 = a_0b_1 + a_1b_0$, $c_2 = a_0b_2 + a_1b_1 + a_2b_0$, und so weiter. Diese Summen entsprechen Diagonalen in der Tabelle (von links unten nach rechts oben). Man erhält damit den folgenden Satz.

Satz 2.42. *Es seien* $\sum_{n=0}^{\infty} a_n$ *und* $\sum_{n=0}^{\infty} b_n$ *absolut konvergente Reihen. Wenn wir*

$$c_n = \sum_{k=0}^{n} a_k b_{n-k}$$

definieren, dann ist $\sum_{n=0}^{\infty} c_n$ *ebenfalls absolut konvergent, und es gilt*

$$\sum_{n=0}^{\infty} a_n \cdot \sum_{n=0}^{\infty} b_n = \sum_{n=0}^{\infty} c_n.$$

[3]Man beachte, dass im Vergleich zur Summe in Satz 2.20 der erste Term $n = 0$ fehlt.

Beispiel 2.43. Wenn $a_n = b_n = 2^{-n}$, dann sind sowohl $\sum_{n=0}^{\infty} a_n$ als auch $\sum_{n=0}^{\infty} b_n$ geometrische Reihen mit Wert 2. Das Produkt der beiden ist also 4. Andererseits ist

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

in Satz 2.42. Damit folgt

$$\sum_{n=0}^{\infty} c_n = \sum_{n=0}^{\infty} (n+1)2^{-n} = 1 + \frac{2}{2} + \frac{3}{4} + \frac{4}{8} + \frac{5}{16} + \cdots = 4.$$

2.7 Komplexe Folgen und Reihen

Komplexe Folgen und Reihen können auf reelle Folgen und Reihen zurückgeführt werden, indem man sie in Realteil und Imaginärteil aufteilt: so ist eine komplexe Folge $z_n = a_n + b_n i$ genau dann konvergent, wenn sowohl a_n als auch b_n konvergent ist. Gilt $\lim_{n\to\infty} a_n = a$ und $\lim_{n\to\infty} b_n = b$, so ist

$$\lim_{n\to\infty} (a_n + b_n i) = a + bi.$$

Ähnliches gilt auch für Reihen:

$$\sum_{n=0}^{\infty} (a_n + b_n i) = \Big(\sum_{n=0}^{\infty} a_n\Big) + \Big(\sum_{n=0}^{\infty} b_n\Big) i,$$

sofern die Reihen konvergieren. Die Rechenregeln, die für reelle Folgen und Reihen gelten, bleiben auch für komplexe Zahlen erhalten.

Beispiel 2.44. Der Grenzwert der Folge $z_n = \frac{n^2 i - 3n + 2i}{2n^2 + in^2 - n}$ ist

$$\lim_{n\to\infty} z_n = \lim_{n\to\infty} \frac{n^2 i - 3n + 2i}{2n^2 + in^2 - n} = \lim_{n\to\infty} \frac{i - \frac{3}{n} + \frac{2i}{n^2}}{2 + i - \frac{1}{n}} = \frac{i - 0 + 0}{2 + i - 0} = \frac{i(2-i)}{(2+i)(2-i)} = \frac{1+2i}{5}.$$

Beispiel 2.45. Die Reihe $\sum_{n=0}^{\infty} \left(\frac{1+i}{3}\right)^n$ ist wegen $|\frac{1+i}{3}| = \frac{\sqrt{2}}{3} < 1$ eine konvergente geometrische Reihe, und es gilt

$$\sum_{n=0}^{\infty} \Big(\frac{1+i}{3}\Big)^n = \frac{1}{1 - \frac{1+i}{3}} = \frac{3}{2-i} = \frac{3(2+i)}{(2+i)(2-i)} = \frac{6+3i}{5}.$$

Kapitel 3

Funktionen

3.1 Grundbegriffe

Eine *Funktion* $f : A \to B$ ordnet jedem Element a einer Menge A einen Wert $f(a)$ in einer Menge B zu. Die Mengen A und B sind dabei im Prinzip beliebig, wir betrachten im Folgenden aber vor allem den Fall, dass A und B (Teilmengen von) $\mathbb{R}$ oder $\mathbb{C}$ sind.

Man nennt die Menge A die *Definitionsmenge* oder den *Definitionsbereich* der Funktion f, B die *Zielmenge*[1].

Während es zu jedem $a \in A$ einen Funktionswert $f(a)$ geben muss, damit die Funktion wohldefiniert ist, müssen nicht alle Elemente von B tatsächlich als Werte von f vorkommen. Die Menge jener Elemente, die tatsächlich Werte von f sind, also

$$f(A) = \{f(a) : a \in A\} = \{b \in B : \text{es gibt ein } a \text{ mit } f(a) = b\},$$

wird *Wertemenge* oder *Wertebereich* genannt. Es gilt nach Definition $f(A) \subseteq B$.

Funktionen können auf verschiedene Arten definiert werden; hier wird es meist eine explizite Formel sein, wie etwa $f(x) = x^2$. Funktionen (besonders reelle Funktionen) werden häufig durch ihren *Funktionsgraph* representiert. Dieser besteht aus allen Paaren der Form $(a, f(a))$, $a \in A$. Sind A und B Teilmengen von $\mathbb{R}$, so entspricht jedes solche Paar einem Punkt in der Ebene.

Funktionen werden mitunter *stückweise* definiert, wobei der Definitionsbereich in Teilstücke unterteilt wird und auf jedem der Stücke ein anderer Ausdruck zur Definition der Funktion verwendet wird. Ein typisches Beispiel ist der Absolutbetrag

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

Ein etwas komplizierteres Beispiel ist die Funktion

$$f(x) = \begin{cases} 3 & x < -2, \\ x^2 & -2 \leq x < 1, \\ \sqrt{x} & 1 \leq x < 4, \\ 5 - x & x \geq 4, \end{cases}$$

[1] Auch andere Bezeichnungen sind üblich, etwa *Bildmenge*/*Bildbereich* oder *Wertemenge*/*Wertebereich*. Letzteres hat hier aber eine andere Bedeutung.

deren Funktionsgraph in Abbildung 3.1 dargestellt ist.

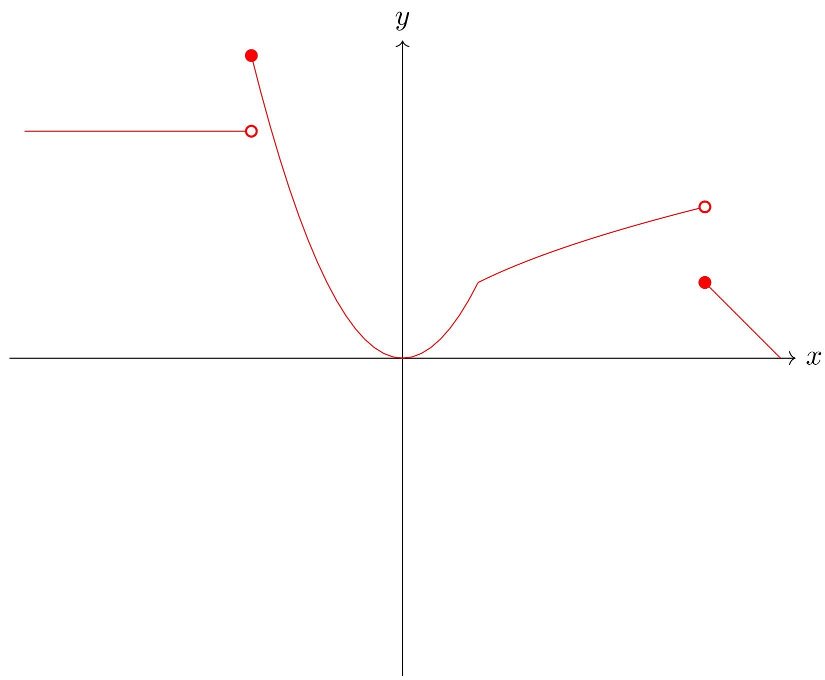

Abbildung 3.1: Eine stückweise definierte Funktion.

Definition 3.1.

- Eine Funktion $f : A \to B$ heißt *injektiv*, wenn für $a_1 \neq a_2$ auch immer $f(a_1) \neq f(a_2)$ gilt, wenn also verschiedenen Elementen von A auch verschiedene Werte zugeordnet werden.
- Eine Funktion $f : A \to B$ heißt *surjektiv*, wenn es zu jedem $b \in B$ ein $a \in A$ derart gibt, dass $f(a) = b$, wenn also jedes Element von B als Wert angenommen wird. Es gilt dann $f(A) = B$.
- Eine Funktion $f : A \to B$ heißt *bijektiv*, wenn sie sowohl injektiv als auch surjektiv ist.

Zusammengefasst kann man also sagen:

- *injektiv* bedeutet, dass es immer *höchstens* ein a mit $f(a) = b$ gibt,
- *surjektiv* bedeutet, dass es immer *mindestens* ein a mit $f(a) = b$ gibt,
- *bijektiv* bedeutet, dass es immer *genau* ein a mit $f(a) = b$ gibt.

3.2 Elementare Funktionen

In Arbeit.

3.3 Operationen

In Arbeit.

3.4 Eigenschaften von Funktionen

In diesem Abschnitt betrachten wir einige wichtige Eigenschaften, die reelle Funktionen haben können. Die erste davon ist *Monotonie*.

Definition 3.2. Es sei f eine Funktion von $A \subseteq \mathbb{R}$ nach $B \subseteq \mathbb{R}$, und es sei $A' \subseteq A$ eine Teilmenge des Definitionsbereichs (typischerweise ein Intervall). Die Funktion heißt auf A'

- *streng monoton steigend*, wenn für alle $x, y \in A'$ mit $x < y$ gilt, dass $f(x) < f(y)$,
- *monoton steigend*, wenn für alle $x, y \in A'$ mit $x < y$ gilt, dass $f(x) \leq f(y)$,
- *streng monoton fallend*, wenn für alle $x, y \in A'$ mit $x < y$ gilt, dass $f(x) > f(y)$,
- *monoton fallend*, wenn für alle $x, y \in A'$ mit $x < y$ gilt, dass $f(x) \geq f(y)$.

Ist eine Funktion entweder (streng) monoton steigend oder (streng) monoton fallend, dann wird sie auch nur (streng) *monoton* genannt.

Man beachte, dass gemäß dieser Definition konstante Funktionen ($f(x) = c$ für alle x, wobei c eine feste reelle Zahl ist) sowohl monoton steigend als auch monoton fallend sind (aber nicht streng).

Beispiele streng monoton steigender Funktionen sind etwa $f(x) = x^3$ oder $f(x) = e^x$ (jeweils auf ganz $\mathbb{R}$). Die Sinusfunktion ist auf $\mathbb{R}$ weder monoton steigend noch fallend, wird es aber wenn man sie geeignet einschränkt: $f(x) = \sin x$ ist streng monoton steigend auf $[-\frac{\pi}{2}, \frac{\pi}{2}]$, anschließend streng monoton fallend auf $[\frac{\pi}{2}, \frac{3\pi}{2}]$, dann wieder streng monoton steigend, und so weiter.

Es gilt der folgende Satz:

Satz 3.3. *Eine streng monotone Funktion ist injektiv.*

Dies folgt daraus, dass für zwei Werte $x < y$ aufgrund der strengen Monotonie immer $f(x) < f(y)$ (für steigendes f) bzw. $f(x) > f(y)$ (für fallendes f) gelten muss, insbesondere also nicht $f(x) = f(y)$. Die Umkehrung des Satzes gilt jedoch nicht: eine injektive Funktion muss auf ihrem Definitionsbereich nicht monoton sein, wie das Beispiel in Abbildung 3.2 zeigt.

Eine weitere wichtige Eigenschaft ist *Beschränktheit*.

Definition 3.4. Es sei f eine Funktion von $A \subseteq \mathbb{R}$ nach $B \subseteq \mathbb{R}$. Die Funktion heißt *nach oben beschränkt*, wenn es eine reelle Zahl M gibt, sodass $f(x) \leq M$ für alle $x \in A$ gilt. Sie heißt *nach unten beschränkt*, wenn es eine reelle Zahl m gibt, sodass $f(x) \geq m$ für alle $x \in A$ gilt. Ist f sowohl nach oben als auch nach unten beschränkt, so wird f *beschränkt* genannt.

Beispielsweise ist $f(x) = \frac{1}{x}$ auf dem Intervall $(0, \infty)$ nach unten (durch 0) beschränkt, aber nicht nach oben. Sinus und Cosinus sind beschränkte Funktionen (mit den Schranken -1 und 1), Tangens und Cotangens dagegen sind unbeschränkt.

Beispiel 3.5. Sind die Funktionen

$$f(x) = \frac{2x^2 + 3}{x^2 + 1} \text{ und } g(x) = \frac{2x^3 + 3}{x^2 + 1}$$

beschränkt?

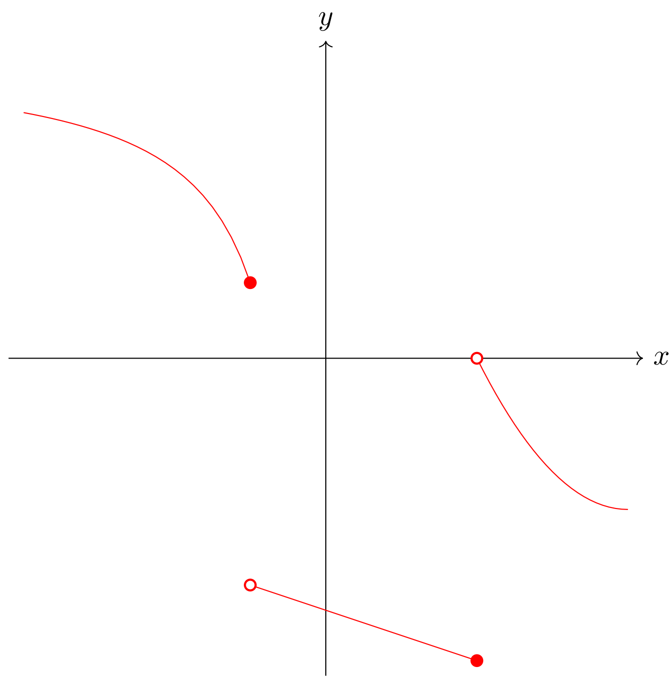

Abbildung 3.2: Eine Funktion, die injektiv, aber nicht monoton ist.

Lösung. Die erste Funktion ist in der Tat beschränkt, denn es gilt

$$0 \leq \frac{2x^2+3}{x^2+1} = \frac{2x^2+2+1}{x^2+1} = 2 + \frac{1}{x^2+1} \leq 3.$$

Die zweite Funktion dagegen ist unbeschränkt, denn es gilt etwa für $x \geq 1$

$$\frac{2x^3+3}{x^2+1} \geq \frac{2x^3}{x^2+1} \geq \frac{2x^3}{2x^2} = x,$$

was beliebig groß werden kann. Somit gibt es keine obere Schranke M.

Als Nächstes betrachten wir *Symmetrie* von Funktionen. Hier sind zwei Typen von Funktionen bedeutsam.

Definition 3.6. Eine Funktion $f : \mathbb{R} \to \mathbb{R}$ heißt *gerade* oder *symmetrisch*, wenn für alle $x \in \mathbb{R}$ die Identität $f(-x) = f(x)$ gilt. Sie heißt *ungerade* oder *schiefsymmetrisch*, wenn für alle $x \in \mathbb{R}$ die Identität $f(-x) = -f(x)$ gilt.

Geometrisch interpretiert bedeuten diese Eigenschaften Folgendes: der Graph einer geraden Funktion ist (spiegel-)symmetrisch bezüglich der y-Achse. Der Graph einer ungeraden Funktion ist (spiegel-)symmetrisch bezüglich des Ursprungs. Dies wird in Abbildung 3.3 illustriert. Beispiele gerader Funktionen sind etwa $f(x) = x^2$ (allgemein $f(x) = x^n$ für gerade ganze Zahlen n, daher auch der Name), $f(x) = |x|$ oder $f(x) = \cos x$. Beispiele ungerader Funktionen sind etwa $f(x) = x^3$ (allgemein $f(x) = x^n$ für ungerade ganze Zahlen n), $f(x) = \sin x$ oder $f(x) = \tan x$. In letzterem Fall ist der Definitionsbereich nicht ganz $\mathbb{R}$, aber er ist selbst symmetrisch bezüglich 0. Man spricht daher dennoch beim Tangens von einer ungeraden Funktion.

Man beachte, dass es auch Funktionen gibt, die weder gerade noch ungerade sind. Zum Beispiel gilt für die Exponentialfunktion $f(x) = e^x$ weder $f(-x) = f(x)$ noch $f(-x) = -f(x)$. Es

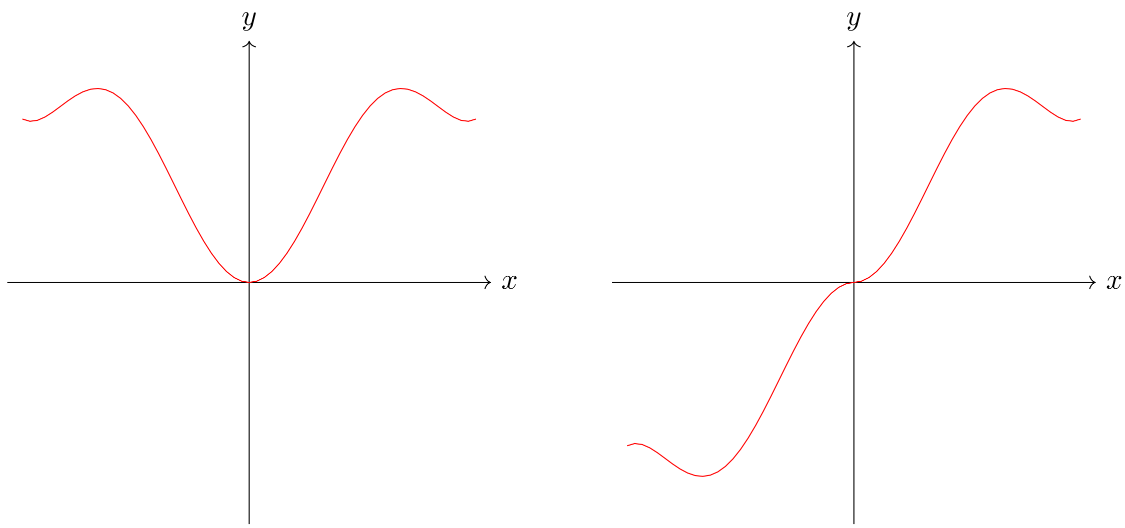

Abbildung 3.3: Eine gerade und eine ungerade Funktion.

gibt sogar eine Funktion, die sowohl gerade als auch ungerade ist, nämlich die Nullfunktion ($f(x) = 0$ für alle x). Sie ist allerdings die einzige solche Funktion, denn $f(-x) = f(x)$ und $f(-x) = -f(x)$ implizieren zusammen $f(x) = 0$.

Beispiel 3.7. Wir wollen entscheiden, welche der folgenden Funktionen gerade bzw. ungerade sind:

$$f_1(x) = \frac{x^5}{x^4 + x^3 + 1}, \qquad f_2(x) = \frac{x^5}{x^4 + x^2 + 1}, \qquad f_3(x) = \left|\frac{x^5}{x^4 + x^2 + 1}\right|.$$

Lösung. Wir ersetzen jeweils x durch $-x$ und vergleichen mit $f(x)$. So ist

$$f_1(-x) = \frac{(-x)^5}{(-x)^4 + (-x)^3 + 1} = \frac{-x^5}{x^4 - x^3 + 1} \neq \pm f_1(x),$$

die Funktion f_1 ist also weder gerade noch ungerade (so ist etwa $f(-1) = -1$ weder gleich $f(1) = \frac{1}{3}$ noch gleich $-f(1) = -\frac{1}{3}$). Dagegen ist

$$f_2(-x) = \frac{(-x)^5}{(-x)^4 + (-x)^2 + 1} = \frac{-x^5}{x^4 + x^2 + 1} = -f_2(x),$$

was bedeutet, dass f_2 eine ungerade Funktion ist. Schließlich ist

$$f_3(-x) = \big|f_2(-x)\big| = \big|-f_2(x)\big| = \big|f_2(x)\big| = f_3(x).$$

Also ist f_3 eine gerade Funktion.

Das letzte Konzept in diesem Abschnitt ist *Periodizität*.

Definition 3.8. Eine Funktion $f : A \to \mathbb{R}$ heißt *periodisch* mit *Periode* P, wenn für alle $x \in A$ gilt, dass $x + P \in A$ und

$$f(x + P) = f(x).$$

Die trigonometrischen Funktionen (Sinus, Cosinus, Tangens, Cotangens) sind typische Beispiele: es gilt

$$\sin(x \pm 2\pi) = \sin x, \quad \cos(x \pm 2\pi) = \cos x,$$

sowie

$$\tan(x \pm \pi) = \tan x, \quad \cot(x \pm \pi) = \cot x.$$

Ist f periodisch mit Periode P, dann gilt auch

$$f(x + nP) = f(x + (n-1)P) = \cdots = f(x + P) = f(x)$$

für jede natürliche Zahl n. Also ist f automatisch auch periodisch mit Periode nP. Wenn man von „der Periode“ einer periodischen Funktion spricht, dann ist üblicherweise die kleinstmögliche positive Periode gemeint (sofern eine existiert: für konstante Funktionen ist etwa jedes P eine Periode, es gibt daher keine kleinste). In diesem Sinn ist z.B. 2π die Periode von Sinus und Cosinus, während π die Periode von Tangens und Cotangens ist.

Als weiteres Beispiel betrachten wir die Funktion $\{x\}$, den *gebrochenen Anteil.* Es sei dazu wieder $\lfloor x \rfloor$ die größte ganze Zahl, die $\leq x$ ist. Wir setzen nun

$$\{x\} = x - \lfloor x \rfloor.$$

So ist etwa $\{3\} = 0$, $\{2.8\} = 0.8$ oder $\{-1.4\} = 0.6$. Der Funktionsgraph ist in Abbildung 3.4 dargestellt. Diese Funktion ist periodisch mit Periode 1: es gilt offensichtlich $f(x+1) = f(x)$.

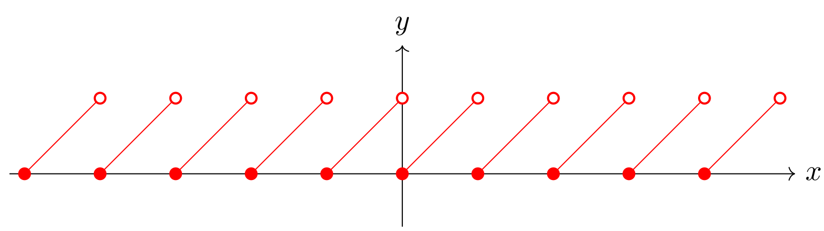

Abbildung 3.4: Die Funktion $f(x) = \{x\}$.

3.5 Grenzwert und Stetigkeit

Wir kommen nun zum fundamentalen Begriff des Grenzwertes einer Funktion an einem Punkt.

Definition 3.9. Es seien $f : A \to \mathbb{R}$ eine reelle Funktion, und a ein Häufungspunkt von A. Die Funktion f hat im Punkt a den *Grenzwert* b, geschrieben

$$\lim_{x \to a} f(x) = b,$$

wenn für jede Folge von Zahlen $x_n \in A \setminus \{a\}$ mit $\lim_{n\to\infty} x_n = a$ gilt, dass $\lim_{n\to\infty} f(x_n) = b$.

Intuitiv gesprochen bedeutet dies: wenn sich x an a annähert (ohne dass $x = a$), dann nähert sich $f(x)$ immer stärker an b. Wichtig ist in diesem Zusammenhang, dass der Wert von $f(a)$ für den Grenzwert *keine Rolle* spielt, insbesondere ist der Grenzwert nicht immer gleich $f(a)$. Der Punkt a muss a priori nicht einmal zum Definitionsbereich von f gehören.

Mit Hilfe der Definition des Grenzwertes von Folgen kann die Definition auf folgende äquivalente Art umformuliert werden.

Definition 3.10. Es sei $f : A \to \mathbb{R}$ eine reelle Funktion, und a ein Häufungspunkt von A. Die Funktion f hat im Punkt a den Grenzwert b, wenn zu jedem $\epsilon > 0$ ein $\delta > 0$ existiert (das von ϵ abhängen darf), sodass für alle $x \in A \setminus \{a\}$ mit $|x - a| < \delta$ stets $|f(x) - b| < \epsilon$ gilt.

Wir betrachten zunächst einige Beispiele:

- Für konstante Funktionen ($f(x) = c$ für alle x) gilt $\lim_{x \to a} f(x) = c$ an jedem Punkt a.
- Für die Funktion $f(x) = x$ gilt $\lim_{x \to a} f(x) = a$.
- Es sei $f(x)$ wie folgt definiert:

$$f(x) = \begin{cases} 0 & x \neq 0, \\ 1 & x = 0. \end{cases} \tag{3.1}$$

  Für jede mögliche Folge x_n mit $x_n \neq 0$ und $\lim_{n \to \infty} x_n = 0$ gilt, dass $f(x_n) = 0$ für alle n, daher $\lim_{n \to \infty} f(x_n) = 0$ und somit

$$\lim_{x \to 0} f(x) = 0.$$

  In diesem Beispiel ist der Grenzwert also nicht gleich dem Funktionswert.

Wenn wir uns auf Folgen beschränken, die von einer Seite gegen a streben, so erhalten wir sogenannte *einseitige Grenzwerte*.

Definition 3.11.

- Es sei $f : A \to \mathbb{R}$ eine reelle Funktion, und a ein Häufungspunkt von $A \cap (-\infty, a)$. Die Funktion f hat im Punkt a den *linksseitigen Grenzwert* b, geschrieben

$$\lim_{x \to a^-} f(x) = b,$$

  wenn für jede Folge von Zahlen $x_n \in A$, für die $x_n < a$ und $\lim_{n \to \infty} x_n = a$ gilt, dass $\lim_{n \to \infty} f(x_n) = b$.

- Es sei $f : A \to \mathbb{R}$ eine reelle Funktion, und a ein Häufungspunkt von $A \cap (a, \infty)$. Die Funktion f hat im Punkt a den *rechtsseitigen Grenzwert* b, geschrieben

$$\lim_{x \to a^+} f(x) = b,$$

  wenn für jede Folge von Zahlen $x_n \in A$, für die $x_n > a$ und $\lim_{n \to \infty} x_n = a$ gilt, dass $\lim_{n \to \infty} f(x_n) = b$.

Diese Definition lässt sich auch in derselben Form wie in Definition 3.17 umschreiben, wobei man sich auf Werte $x < a$ bzw. $x > a$ beschränkt. Als Beispiel betrachten wir die *Vorzeichenfunktion*

$$\operatorname{sgn} x = \begin{cases} 1 & x > 0, \\ 0 & x = 0, \\ -1 & x < 0. \end{cases}$$

Diese Funktion hat bei 0 einen links- und einen rechtsseitigen Grenzwert:

$$\lim_{x \to 0^-} f(x) = -1, \qquad \lim_{x \to 0^+} f(x) = 1.$$

Der Grenzwert $\lim_{x\to 0} f(x)$ existiert jedoch nicht, denn die Funktion nähert sich bei 0 nicht einem einzigen Wert an. Wenn wir zum Beispiel $x_n = \frac{(-1)^n}{n}$ setzen, dann gilt $\lim_{n\to\infty} x_n = 0$. Allerdings ist $f(x_n) = 1$ für gerade n und $f(x_n) = -1$ für ungerade n, die Folge $f(x_n)$ hat also keinen Grenzwert.

Für die Funktion $\{x\}$ (gebrochener Anteil), die uns im Zusammenhang mit Periodizität begegnet ist, gilt etwa

$$\lim_{x\to 1^-} \{x\} = 1, \text{ aber } \lim_{x\to 1^+} \{x\} = 0.$$

Der Grenzwert $\lim_{x\to 1}\{x\}$ existiert also nicht. In nichtganzzahligen Punkten existiert er jedoch: so ist zum Beispiel

$$\lim_{x\to 3/2^-} \{x\} = \lim_{x\to 3/2^+} \{x\} = \lim_{x\to 3/2} \{x\} = \frac{1}{2}.$$

Bemerkung 3.12. Damit die Grenzwertdefinitionen sinnvoll sind, muss a immer ein Häufungspunkt des Definitionsbereichs sein, da es sonst keine Folgen x_n wie in der Definition gibt. In der Praxis bedeutet das normalerweise, dass die Funktion zumindest auf einem Intervall I definiert ist, das a enthält, außer vielleicht in a selbst. Daher nehmen wir dies im Folgenden der Einfachheit halber üblicherweise an und schreiben nur „Funktion auf einem Intervall I“.

Es gilt folgender Zusammenhang zwischen einseitigen Grenzwerten und Grenzwerten.

Satz 3.13. *Es sei f eine reelle Funktion auf einem Intervall I, und a ein Punkt im Inneren des Intervalls (nicht am Rand). Dann existiert der Grenzwert $\lim_{x\to a} f(x)$ genau dann, wenn die einseitigen Grenzwerte $\lim_{x\to a^-} f(x)$ und $\lim_{x\to a^+} f(x)$ existieren und gleich sind. In diesem Fall ist*

$$\lim_{x\to a} f(x) = \lim_{x\to a^-} f(x) = \lim_{x\to a^+} f(x).$$

Es gelten für Grenzwerte von Funktionen dieselben Rechenregeln wie für das Rechnen mit Folgen.

Satz 3.14. *Es seien f_1 und f_2 reelle Funktionen auf einem Intervall I, das a enthält. Wenn die Grenzwerte $\lim_{x\to a} f_1(x) = b_1$ und $\lim_{x\to a} f_2(x) = b_2$ beide existieren, dann gilt*

$$\lim_{x\to a}(f_1(x) + f_2(x)) = b_1 + b_2, \quad \lim_{x\to a}(f_1(x) - f_2(x)) = b_1 - b_2, \quad \lim_{x\to a}(f_1(x) f_2(x)) = b_1 b_2,$$

und sofern $b_2 \neq 0$ auch

$$\lim_{x\to a} \frac{f_1(x)}{f_2(x)} = \frac{b_1}{b_2}.$$

Ein Spezialfall ist die Multiplikation (oder Division) mit Konstanten: es gilt

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

für jeden konstanten Faktor k (sofern der Grenzwert existiert). Wir betrachten nun einige weitere Beispiele.

- Die Funktion $f(x) = \frac{x+\{x\}}{2x+3}$ setzt sich aus Funktionen zusammen, deren Grenzwert bereits betrachtet wurde. Es gilt etwa

$$\lim_{x\to 3/2} f(x) = \frac{\lim_{x\to 3/2} x + \lim_{x\to 3/2}\{x\}}{2\lim_{x\to 3/2} x + 3} = \frac{\frac{3}{2} + \frac{1}{2}}{2\cdot\frac{3}{2} + 3} = \frac{2}{6} = \frac{1}{3}.$$

- Der Grenzwert von $f(x) = \frac{\sqrt{x+4}-2}{x}$ für $x \to 0$ lässt sich nicht direkt aus den Rechenregeln bestimmen, denn Zähler und Nenner sind im Grenzwert beide 0 (die Funktion ist bei 0 auch nicht definiert). Stattdessen schreiben wir die Funktion ein wenig um:

$$f(x) = \frac{(\sqrt{x+4}-2)(\sqrt{x+4}+2)}{x(\sqrt{x+4}+2)} = \frac{x+4-2^2}{x(\sqrt{x+4}+2)} = \frac{1}{\sqrt{x+4}+2}.$$

  Wenn wir nun x nach 0 streben lassen, dann nähert sich der Nenner $\sqrt{4}+2 = 4$ an. Es gilt also

$$\lim_{x \to 0} f(x) = \frac{1}{4}.$$

Ein nützliches Werkzeug ist auch der folgende Satz („Sandwichsatz“) über Funktionen, die zwischen zwei anderen eingeschachtelt sind. Er folgt aus dem analogen Satz für Folgen (Satz 2.8).

Satz 3.15. *Es seien f, g_1, g_2 reelle Funktionen auf einem Intervall I, das a enthält. Wenn die Ungleichungskette*

$$g_1(x) \leq f(x) \leq g_2(x)$$

gilt, und es ist $\lim_{x \to a} g_1(x) = \lim_{x \to a} g_2(x) = b$, dann auch

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

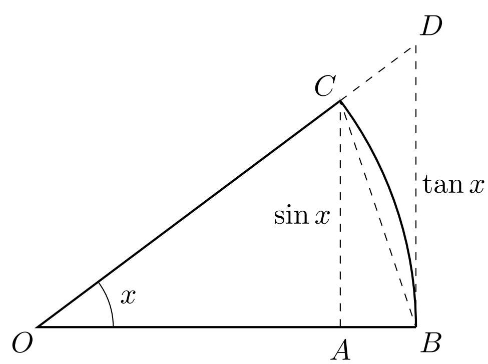

Abbildung 3.5: Zur Ungleichung $\sin x \leq x \leq \tan x$.

Beispiel 3.16. Die Funktion

$$\mathrm{si}(x) = \frac{\sin x}{x}$$

spielt in der Signalverarbeitung eine wichtige Rolle. Wir wollen ihren Grenzwert bei 0 berechnen. Dazu verwenden wir die Ungleichungskette

$$\sin x \leq x \leq \tan x,$$

die sich durch Abbildung 3.5 veranschaulichen lässt: hier ist BC ein Kreisbogen mit Mittelpunkt O und Radius 1, der den Winkel x (in Radianten) einschließt. Dann haben also OB und OC Länge 1, und aus den rechtwinkligen Dreiecken OAC und OBD ergibt sich, dass AC und BD Länge $\sin x$ bzw. $\tan x$ haben. Damit ist der Flächeninhalt des Dreiecks OBC

gleich $\frac{\sin x}{2}$ (Seitenlänge 1, Höhe $\sin x$) und der Flächeninhalt des Dreiecks OBD gleich $\frac{\tan x}{2}$ (Seitenlänge 1, Höhe $\tan x$). Zwischen den beiden liegt der Kreissektor, der von OB, OC und dem Kreisbogen BC eingeschlossen wird. Sein Flächeninhalt ist $\pi \cdot \frac{x}{2\pi} = \frac{x}{2}$ (der volle Kreis hat Flächeninhalt π, der Sektor einen Anteil von $\frac{x}{2\pi}$). Also muss in der Tat

$$\sin x \leq x \leq \tan x$$

gelten. Wir dividieren die Ungleichungskette durch x und erhalten

$$\frac{\sin x}{x} \leq 1 \leq \frac{\tan x}{x} = \frac{\sin x}{x \cos x}.$$

Daraus ergibt sich wiederum

$$\cos x \leq \frac{\sin x}{x} \leq 1.$$

Für $x \to 0$ strebt $\cos x$ gegen 1, daher können wir Satz 3.15 verwenden. Es gilt somit

$$\lim_{x \to 0} \frac{\sin x}{x} = 1.$$

Der Grenzwert b in Definition 3.9 kann auch $\pm\infty$ sein. Eine äquivalente Definition lautet dann folgendermaßen:

Definition 3.17. Es sei $f : A \to \mathbb{R}$ eine reelle Funktion, und a ein Häufungspunkt von A. Die Funktion f hat im Punkt a den Grenzwert ∞, wenn zu jedem $M > 0$ ein $\delta > 0$ existiert (das von M abhängen darf), sodass für alle $x \in A \setminus \{a\}$ mit $|x - a| < \delta$ stets $f(x) > M$ gilt.

Die Funktion wächst also in der Nähe von a über jede Grenze M. Eine analoge Definition gilt auch für den Grenzwert $-\infty$. Typische Beispiele sind alle Potenzfunktionen mit negativem Exponenten (also Funktionen der Form $f(x) = x^{-n} = \frac{1}{x^n}$ für eine positive ganze Zahl n): es gilt in diesem Fall

$$\lim_{x \to 0^+} x^{-n} = \infty, \qquad \lim_{x \to 0^-} x^{-n} = \begin{cases} \infty & n \text{ gerade}, \\ -\infty & n \text{ ungerade}. \end{cases}$$

Für gerade n ist damit auch

$$\lim_{x \to \infty} x^{-n} = \infty.$$

Hat eine Funktion bei a einen unendlichen Grenzwert, so spricht man davon, dass die Gerade mit der Gleichung $x = a$ eine *vertikale Asymptote* ist.

Umgekehrt kann man auch Grenzwerte bei $\pm\infty$ betrachten. Die Definition ist analog zu jener bei endlichen Punkten (man ersetze a in Definition 3.9 durch $\pm\infty$). Grenzwerte bei $\pm\infty$ können dabei durchaus endlich sein. So ist etwa

$$\lim_{x \to \infty} \frac{4x + 3}{2x - 1} = \lim_{x \to \infty} \frac{4 + \frac{3}{x}}{2 - \frac{1}{x}} = \frac{4 + 0}{2 - 0} = 2.$$

Gilt $\lim_{x \to \infty} f(x) = a$ oder $\lim_{x \to -\infty} f(x) = a$, so wird $y = a$ eine *horizontale Asymptote* genannt.

Schließlich gibt es auch Geraden der Form $y = kx + d$, die als Asymptoten bezeichnet werden, wenn sie sich an $f(x)$ annähern, wenn also gilt dass

$$\lim_{x \to \infty} (f(x) - (kx + d)) = 0 \text{ oder } \lim_{x \to -\infty} (f(x) - (kx + d)) = 0.$$

Diese kann man in zwei Schritten berechnen, falls sie existieren: erst k als Grenzwert von $\frac{f(x)}{x}$, dann d als Grenzwert von $f(x) - kx$.

Beispiel 3.18. Wir wollen eine solche schiefe Asymptote für die Funktion

$$f(x) = \frac{2x^2 + 5x - 1}{x + 3}$$

berechnen. Dazu bestimmen wir zunächst

$$\lim_{x\to\infty} \frac{f(x)}{x} = \lim_{x\to\infty} \frac{2x^2 + 5x - 1}{x^2 + 3x} = \lim_{x\to\infty} \frac{2 + \frac{5}{x} - \frac{1}{x^2}}{1 + \frac{3}{x}} = \frac{2 + 0 - 0}{1 + 0} = 2$$

und in weiterer Folge

$$\begin{aligned}\lim_{x\to\infty} (f(x) - 2x) &= \lim_{x\to\infty} \frac{2x^2 + 5x - 1 - 2x(x+3)}{x + 3} = \lim_{x\to\infty} \frac{-x - 1}{x + 3} \\ &= \lim_{x\to\infty} \frac{-1 - \frac{1}{x}}{1 + \frac{3}{x}} = \frac{-1 - 0}{1 + 0} = -1.\end{aligned}$$

Somit ist die Gerade $y = 2x - 1$ eine Asymptote (und zwar sowohl bei ∞ als auch bei $-\infty$, wie aus einer analogen Rechnung folgt), siehe auch Abbildung 3.6.

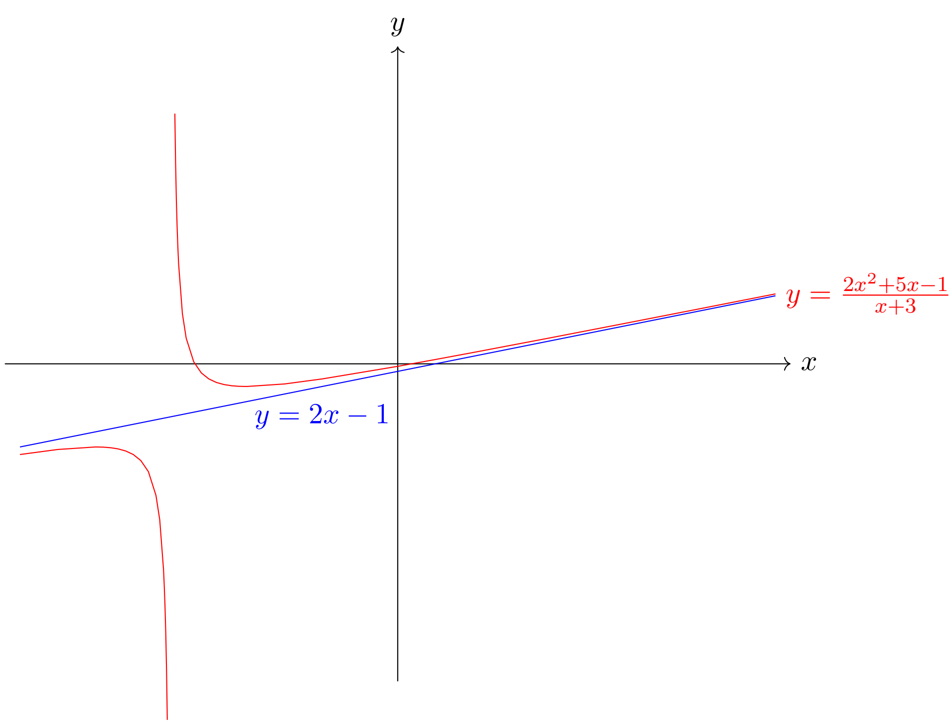

Abbildung 3.6: Zu Beispiel 3.18.

Es wurde bereits erwähnt, dass der Wert einer Funktion an einem Punkt nicht unbedingt gleich dem Grenzwert sein muss. Wenn die beiden übereinstimmen, so spricht man davon, dass die Funktion dort *stetig* ist.

Definition 3.19. Eine Funktion f auf einem Intervall I heißt stetig im Punkt $a \in I$, wenn dort

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

gilt.

Äquivalent kann man Stetigkeit in einem Punkt auch wie folgt definieren:

- Eine Funktion f auf einem Intervall I ist stetig in a, wenn für jede Folge von Zahlen $x_n \in I \setminus \{a\}$ mit $\lim_{n\to\infty} x_n = a$ gilt, dass $\lim_{n\to\infty} f(x_n) = f(a)$.
- Eine Funktion f auf einem Intervall I ist stetig in a, wenn zu jedem $\epsilon > 0$ ein $\delta > 0$ existiert (das von ϵ abhängen darf), sodass für alle $x \in I$ mit $|x - a| < \delta$ stets $|f(x) - f(a)| < \epsilon$ gilt.

So sind etwa konstante Funktionen, die Funktion $f(x) = x$ oder die Funktion $f(x) = |x|$ in allen Punkten stetig. Die Funktion $f(x) = \{x\}$ ist in allen nichtganzzahligen Punkten stetig, an den ganzen Zahlen dagegen *unstetig*. Die Funktion in (3.1) ist in allen Punkten stetig, bis auf den Nullpunkt.

Man definiert Stetigkeit nicht nur in einzelnen Punkten, sondern auch auf größeren Bereichen:

Definition 3.20. Eine Funktion f heißt stetig auf einem Intervall I, wenn sie in jedem Punkt von I stetig ist.

Anschaulich bedeutet die Stetigkeit einer Funktion auf einem Intervall, dass man den Funktionsgraphen „ohne Absetzen“ zeichnen kann.

Aus den Rechenregeln für Grenzwerte ergibt sich der folgende wichtige Satz:

Satz 3.21. *Sind Funktionen f und g in einem Punkt a (bzw. auf einem Intervall I) stetig, so sind es auch die Funktionen $f+g$, $f-g$, $f \cdot g$ und $\frac{f}{g}$ (letzteres solange der Nenner nicht* 0 *ist).*

Daraus folgt unter anderem, dass Polynome und rationale Funktionen auf ihrem gesamten Definitionsbereich stetig sind. Auch Wurzeln, Exponentialfunktionen und die trigonometrischen Funktionen $\sin, \cos, \tan, \cot$ sind Beispiele von Funktionen, die auf dem gesamten Definitionsbereich stetig sind. Wichtig ist weiters die folgende Regel über zusammengesetzte Funktionen.

Satz 3.22. *Ist die Funktion f stetig in b, und gilt* $\lim_{x\to a} g(x) = b$, *dann folgt*

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

Ist nämlich x_n eine Folge, die sich a annähert, dann muss sich die Folge $g(x_n)$ dem Punkt b annähern, und es folgt aus der Stetigkeit von f, dass der Grenzwert von $f(g(x_n))$ gleich $f(b)$ ist.

Beispiel 3.23. Wir interessieren uns für den Grenzwert

$$\lim_{x \to 0} \frac{\cos x - 1}{x^2}.$$

Nun ist

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

Da die Funktion $f(x) = x^2$ stetig ist, gilt

$$\lim_{x\to 0}\left(\frac{\sin x}{x}\right)^2 = \left(\lim_{x\to 0}\frac{\sin x}{x}\right)^2 = 1,$$

und damit letztlich

$$\lim_{x\to 0}\frac{\cos x - 1}{x^2} = -\frac{1}{2}\cdot 1 = -\frac{1}{2}.$$

Für unstetiges f muss Satz 3.22 nicht gelten: zum Beispiel gilt für die Zusammensetzung $f(x) = \{\cos x\}$ der Cosinusfunktion und des gebrochenen Anteils $\lim_{x\to 0} f(x) = 1$, aber $f(0) = \{1\} = 0$. Wenn nämlich x nach 0 strebt, dann geht $\cos x$ gegen 1, und zwar von links, denn $\cos x \le 1$ für alle x. Damit strebt aber $\{\cos x\}$ auch gegen 1.

Aus Satz 3.22 ergibt sich weiters

Satz 3.24. *Ist g stetig in a und f stetig in $b = g(a)$, dann ist auch $f \circ g$ stetig in a. Ist g stetig auf einem Intervall I und f stetig auf einem Intervall J, das $g(I)$ enthält, dann ist auch $f \circ g$ stetig auf I.*

Damit ergibt sich unmittelbar, dass auch kompliziertere zusammengesetzte Funktionen wie

$$f(x) = \frac{e^{1/x}(x^3 - 4\sin(2^{x-4}))}{\cos(2^x - x^2) + \tan^2(5x)}$$

auf ihrem gesamten Definitionsbereich stetig sind. Auch auf Umkehrfunktionen überträgt sich die Stetigkeit:

Satz 3.25. *Ist f im Punkt a stetig, und existiert die Umkehrfunktion auf einem Intervall um a, dann ist auch f^{-1} im Punkt $f(a)$ stetig. Ist f auf einem Intervall I stetig, auf dem die Umkehrfunktion existiert, dann ist auch f^{-1} auf $f(I)$ stetig.*

Insbesondere sind also Logarithmen und die Umkehrfunktionen der trigonometrischen Funktionen wie arcsin auf ihrem Definitionsbereich stetig.

Durch Stetigkeit lassen sich Eigenschaften in einem Punkt auf eine ganze Umgebung des Punktes übertragen, wie im folgenden Satz:

Satz 3.26. *Ist f stetig in a, und gilt $f(a) > 0$, dann gibt es ein Intervall um a (eine „Umgebung“), auf dem die Funktion f positiv ist.*

Es muss nämlich für $\epsilon = \frac{f(a)}{2}$ ein $\delta > 0$ geben, sodass $|f(x) - f(a)| < \epsilon$ ist, sofern $|x - a| < \delta$. Damit ist aber auf dem Intervall $(a - \delta, a + \delta)$

$$f(x) > f(a) - \epsilon = f(a) - \frac{f(a)}{2} = \frac{f(a)}{2} > 0.$$

So wie es einseitige Grenzwerte gibt, spricht man auch von *einseitiger Stetigkeit*:

Definition 3.27. Eine Funktion f auf einem Intervall I heißt *linksseitig stetig* oder *linksstetig* im Punkt $a \in I$, wenn dort

$$\lim_{x\to a^-} f(x) = f(a)$$

gilt. Sie heißt *rechtsseitig stetig* oder *rechtsstetig* im Punkt $a \in I$, wenn dort

$$\lim_{x\to a^+} f(x) = f(a)$$

gilt.

So ist beispielsweise die Funktion $f(x) = \{x\}$ in den ganzen Zahlen zwar rechts- aber nicht linksstetig.

Man klassifiziert verschiedene Typen von Unstetigkeitsstellen:

- Eine *hebbare* Unstetigkeit liegt vor, wenn der Grenzwert existiert, aber nicht gleich dem Funktionswert ist, wie im Beispiel (3.1).
- Eine *Sprungstelle* liegt vor, wenn links- und rechtsseitiger Grenzwert verschieden sind, wie etwa im Fall der Vorzeichenfunktion $\operatorname{sgn}(x)$ oder der Funktion $\{x\}$.
- Ein *Pol* liegt vor, wenn der links- oder rechtsseitige Grenzwert unendlich wird (oder beide), wie beispielsweise für Potenzfunktionen mit negativem Exponenten.
- Eine *wesentliche Unstetigkeit* liegt vor, wenn mindestens einer der einseitigen Grenzwerte nicht existiert. Ein Beispiel ist die Funktion $f(x) = \sin \frac{1}{x}$, die um 0 zwischen -1 und 1 hin- und heroszilliert und keinen Grenzwert annimmt, siehe Abbildung 3.7.

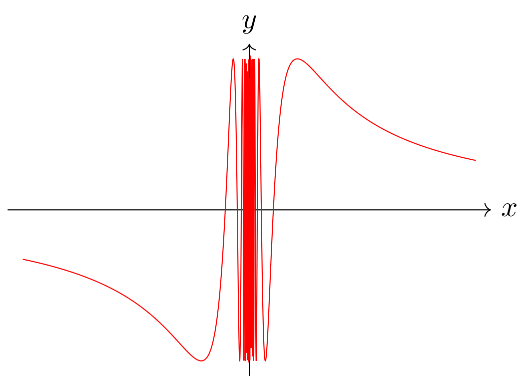

Abbildung 3.7: Graph der Funktion $f(x) = \sin \frac{1}{x}$.

Es kann sogar passieren, dass eine Funktion gar nirgends stetig ist: ein bekanntes Beispiel ist die sogenannte *Dirichlet-Funktion*

$$f(x) = \begin{cases} 1 & x \in \mathbb{Q}, \\ 0 & x \notin \mathbb{Q}. \end{cases}$$

Da es in jedem noch so kleinen Intervall sowohl rationale als auch irrationale Zahlen gibt, werden in jedem Intervall sowohl der Wert 1 als auch der Wert 0 angenommen. Der Grenzwert existiert damit in keinem Punkt.

Wir beenden diesen Abschnitt mit einigen theoretischen Resultaten über stetige Funktionen. Hier ist zunächst der *Zwischenwertsatz* zu nennen.

Satz 3.28 (Zwischenwertsatz)**.** *Ist eine Funktion f stetig auf dem Intervall $[a, b]$, und liegt y zwischen $f(a)$ und $f(b)$ (wobei $f(a) \neq f(b)$), dann gibt es ein $c \in (a, b)$, sodass $f(c) = y$.*

Beweis. Es gelte $f(a) < y < f(b)$ (der umgekehrte Fall wird ähnlich abgehandelt). Wir betrachten die Menge

$$M = \{x \in [a, b] : f(x) < y\}.$$

Diese Menge enthält a und ist daher nicht leer und nach oben durch b beschränkt. Also hat die Menge ein Supremum: $s = \sup M$.

Wir wollen zeigen, dass $f(s) = y$. Für jede positive ganze Zahl n gibt es ein $x_n \in M$ mit $x_n > s - \frac{1}{n}$, weil sonst $s - \frac{1}{n}$ eine obere Schranke für M wäre, im Widerspruch zur Definition des Supremums. Damit haben wir eine Folge von Punkten x_n, die $s - \frac{1}{n} < x_n \leq s$ erfüllt und somit den Grenzwert s hat. Aufgrund der Stetigkeit ist $f(s) = \lim_{n\to\infty} f(x_n) \leq y$, da $f(x_n) \leq y$ für alle n nach Definition von M gilt. Insbesondere ist $s \neq b$.

Andererseits können wir die Folge $w_n = s + \frac{1}{n}$ betrachten: für genügend großes n liegt w_n im Intervall $[a, b]$, und $\lim_{n\to\infty} w_n = s$. Da $w_n > s = \sup M$ gilt, liegt w_n nicht in M, also $f(w_n) \geq y$. Damit folgt wiederum mit Hilfe der Stetigkeit $f(s) = \lim_{n\to\infty} f(w_n) \geq y$.

Aus den Ungleichungen $f(s) \leq y$ und $f(s) \geq y$ folgt $f(s) = y$, womit der Beweis abgeschlossen ist.

Das sogenannte *Bisektionsverfahren* zur numerischen Bestimmung von Nullstellen beruht auf dem Zwischenwertsatz. Es sei dazu f eine stetige Funktion, von der wir eine Nullstelle bestimmen wollen. Weiters seien x_0 und x_1 Punkte ($x_0 < x_1$), für die $f(x_0)$ und $f(x_1)$ verschiedenes Vorzeichen haben. Nach dem Zwischenwertsatz muss f eine Nullstelle zwischen x_0 und x_1 haben.

Betrachte nun den Halbierungspunkt $x_2 = \frac{x_0+x_1}{2}$. Wenn das Vorzeichen von $f(x_2)$ mit jenem von $f(x_0)$ übereinstimmt, suchen wir im Intervall $[x_2, x_1]$ weiter: da $f(x_2)$ und $f(x_1)$ in diesem Fall verschiedenes Vorzeichen haben, muss es dort eine Nullstelle geben. Andernfalls betrachten wir das Intervall $[x_0, x_2]$. Dann bestimmen wir wieder den Wert am Halbierungspunkt und wiederholen das Argument. Auf diese Art erhalten wir eine Folge von immer kleineren Intervallen und eine Folge von Punkten $x_0, x_1, x_2, \ldots$, die sich einer Nullstelle von f annähert.

Beispiel 3.29. Wir wollen eine Nullstelle der Funktion $f(x) = \cos x - x$ bestimmen. Dazu können wir mit $x_0 = 0$ und $x_1 = 1$ beginnen, da $f(x_0) = 1 > 0$ und $f(x_1) \approx -0.45970 < 0$ verschiedenes Vorzeichen haben. Das Verfahren setzt sich wie folgt fort:

$$\begin{aligned}
x_2 &= 0.5,\ \ f(x_2) \approx 0.37758 > 0, \\
x_3 &= 0.75,\ \ f(x_3) \approx -0.01831 < 0, \\
x_4 &= 0.625,\ \ f(x_4) \approx 0.18596 > 0, \\
x_5 &= 0.6875,\ \ f(x_5) \approx 0.08533 > 0, \\
x_6 &= 0.71875,\ \ f(x_6) \approx 0.03388 > 0,
\end{aligned}$$

und so weiter. Man sieht, wie die Intervalle schrittweise kleiner werden und sich die Intervallgrenzen dem tatsächlichen Wert der Nullstelle (der ungefähr 0.73909 ist) annähern.

In späterer Folge werden wir uns für Maxima und Minima von Funktionen interessieren. Dabei ist wichtig festzuhalten, dass solche Extremwerte für stetige Funktionen auf einem abgeschlossenen Intervall immer existieren.

Satz 3.30. *Es sei f eine stetige Funktion auf dem Intervall $[a, b]$. Dann gibt es einen Punkt $x^* \in [a, b]$, wo die Funktion ein Maximum annimmt: $f(x^*) \geq f(x)$ für alle $x \in [a, b]$. Eine analoge Aussage gilt für das Minimum.*

Beweis. Es sei $s = \sup\{f(x) : x \in [a, b]\}$ das Supremum der Funktionswerte. Dann gilt definitionsgemäß $f(x) \leq s$ für alle $x \in [a, b]$.

Ist s endlich, dann muss es für jede positive ganze Zahl n ein $x_n \in [a, b]$ geben, für das $f(x_n) > s - \frac{1}{n}$, da sonst $s - \frac{1}{n}$ eine kleinere obere Schranke für die Funktionswerte wäre, im Widerspruch zur Supremumsdefinition. Es gilt also $s - \frac{1}{n} \leq f(x_n) \leq s$ und damit $\lim_{n\to\infty} f(x_n) = s$. Ist s unendlich, dann können wir mit der gleichen Argumentation eine Folge von Zahlen $x_n \in [a, b]$ wählen, für die $f(x_n) > n$ und damit $\lim_{n\to\infty} f(x_n) = \infty = s$ gilt.
Die Folge x_n ist auf das Intervall $[a, b]$ beschränkt und hat daher einen Häufungswert $x^* \in [a, b]$. Also gibt es eine Teilfolge x_{n_k} mit $\lim_{k\to\infty} x_{n_k} = x^*$. Aufgrund der Stetigkeit von f gilt $f(x^*) = \lim_{k\to\infty} f(x_{n_k}) = s$. Insbesondere ist s also endlich, und es gilt $f(x^*) \geq f(x)$ für alle $x \in [a, b]$ nach Definition von s.

Zum Abschluss sei noch das Konzept der *gleichmäßigen Stetigkeit* erwähnt.

Definition 3.31. Eine Funktion f heißt auf dem Intervall I gleichmäßig stetig, wenn zu jedem $\epsilon > 0$ ein $\delta > 0$ existiert (das von ϵ abhängen darf), sodass für alle $x, a \in I$ mit $|x - a| < \delta$ stets $|f(x) - f(a)| < \epsilon$ gilt.

Der Unterschied zu „gewöhnlicher Stetigkeit“ liegt darin, dass das δ zwar von ϵ, nicht aber von a und x abhängen darf (für eine stetige Funktion auf einem Intervall ist es erlaubt, dass δ auch vom jeweiligen Punkt abhängt).

Eine gleichmäßig stetige Funktion ist immer stetig, aber nicht notwendigerweise umgekehrt: wenn wir etwa die Funktion $f(x) = \frac{1}{x}$ auf dem Intervall $I = (0, \infty)$ betrachten, dann sehen wir etwa mit $\epsilon = 1$: damit für ein gegebenes $a \in I$

$$|f(x) - f(a)| < 1$$

ist, muss

$$\frac{1-a}{a} = \frac{1}{a} - 1 < \frac{1}{x} < \frac{1}{a} + 1 = \frac{1+a}{a}$$

gelten, oder äquivalent dazu

$$\frac{a}{1+a} < x < \frac{a}{1-a}.$$

Je kleiner a, desto kleiner muss δ gewählt werden, damit dies für alle x mit $|x - a| < \delta$ gilt. Dazu ist nämlich $\frac{a}{1+a} \leq a - \delta < a + \delta \leq \frac{a}{1-a}$ erforderlich, was wiederum auf $\delta \leq \frac{a^2}{1+a}$ führt. Kein positives δ erfüllt diese Bedingung jedoch für alle $a \in I$ gleichzeitig. Die Funktion f ist stetig, aber nicht gleichmäßig stetig.

Auf beschränkten abgeschlossenen Intervallen sind solche Gegenbeispiele jedoch nicht möglich. Es gilt der folgende Satz:

Satz 3.32. *Eine stetige Funktion f auf einem beschränkten abgeschlossenen Intervall $[a, b]$ ist gleichmäßig stetig.*

3.6 Polynome

Wie schon erwähnt sind Polynome Ausdrücke der Form

$$P(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0,$$

wobei die a_i konstante reelle oder komplexe Zahlen sind. Hier wird n der *Grad* des Polynoms genannt, und $a_0, a_1, \ldots$ die *Koeffizienten*. Eine wesentliche Eigenschaft ist der sogenannte *Fundamentalsatz der Algebra*, der bereits im Kapitel über komplexe Zahlen erwähnt wurde:

Satz 3.33 (Fundamentalsatz der Algebra). *Eine polynomielle Gleichung n-ten Grades ($n > 0$), also eine Gleichung der Form*

$$a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 = 0,$$

wobei die Koeffizienten a_i komplexe Zahlen mit $a_n \neq 0$ sind, hat immer eine Lösung in komplexen Zahlen.

Wenn x_1 eine solche Lösung (also eine *Nullstelle* des Polynoms $P(x)$ ist, dann ist $x - x_1$ ein Faktor des Polynoms, und es ist

$$P(x) = (x - x_1)Q(x)$$

für ein neues Polynom $Q(x)$. Falls dieses nicht konstant ist, hat es wiederum eine komplexe Nullstelle x_2 und damit einen Faktor $x - x_2$, und so weiter. Damit erhält man immer eine Faktorisierung des Polynoms:

$$P(x) = a_n(x - x_1)(x - x_2) \cdots (x - x_n).$$

Beispiel 3.34. Das Polynom $P(x) = x^3 + 3x^2 + x - 5$ hat eine Nullstelle bei $x = 1$ (die man durch Ausprobieren finden kann). Also hat es eine Faktorisierung:

$$P(x) = (x - 1)(x^2 + 4x + 5).$$

Die Nullstellen des quadratischen Polynoms $x^2 + 4x + 5$ bestimmt man mit der Lösungsformel für quadratische Gleichungen. Sie sind in diesem Fall $-2 - i$ und $-2 + i$. Die vollständige Faktorisierung wäre damit also

$$P(x) = (x - 1)(x - (-2 - i))(x - (-2 + i)) = (x - 1)(x + 2 + i)(x + 2 - i).$$

Ein lineares Polynom der Form $P(x) = ax + b$ hat als Funktionsgraph eine Gerade. Eine solche ist durch zwei Punkte eindeutig bestimmt. Allgemein bestimmen $n + 1$ Punkte (mit verschiedenen x-Koordinaten) (x_0, y_0), (x_1, y_1), ..., (x_n, y_n) ein Polynom n-ten Grades, das genau durch diese Punkte geht, ein sogenanntes *Interpolationspolynom*. Dieses kann allgemein als

$$P(x) = \sum_{k=0}^{n} y_k \prod_{\substack{i=0 \\ i \neq k}}^{n} \frac{x - x_i}{x_k - x_i}$$

ausgedrückt werden. In der Tat gilt hier $P(x_m) = y_m$ für jedes m: wenn man nämlich $x = x_m$ einsetzt, ergibt sich

$$P(x_m) = \sum_{k=0}^{n} y_k \prod_{\substack{i=0 \\ i \neq k}}^{n} \frac{x_m - x_i}{x_k - x_i}.$$

Falls $k \neq m$, dann enthält das Produkt einen Faktor $x_m - x_m$ und ist damit 0. Es bleibt also nur der Term $k = m$ übrig, und der ist

$$y_m \prod_{\substack{i=0 \\ i \neq m}}^{n} \frac{x_m - x_i}{x_m - x_i} = y_m \prod_{\substack{i=0 \\ i \neq m}}^{n} 1 = y_m,$$

also $P(x_m) = y_m$.

Beispiel 3.35. Wir wollen ein Polynom bestimmen, dessen Funktionsgraph durch die Punkte $(-2,1)$, $(0,5)$ und $(3,4)$ geht. Die Formel liefert

$$\begin{aligned} P(x) &= 1 \cdot \frac{(x-0)(x-3)}{(-2-0)(-2-3)} + 5 \cdot \frac{(x-(-2))(x-3)}{(0-(-2))(0-3)} + 4 \cdot \frac{(x-(-2))(x-0)}{(3-(-2))(3-0)} \\ &= \frac{x(x-3)}{10} + 5 \cdot \frac{(x+2)(x-3)}{-6} + 4 \cdot \frac{(x+2)x}{15} = -\frac{7x^2}{15} + \frac{16x}{15} + 5. \end{aligned}$$

Man prüft leicht nach, dass $P(-2) = 1$, $P(0) = 5$ und $P(3) = 4$.

So wie sich beliebige Punkte durch ein Polynom interpolieren lassen, können auch Funktionen durch Polynome approximiert werden. In späterer Folge werden wir unter anderem eine Form der Approximation durch sogenannte Taylorpolynome kennenlernen. Es gilt sogar der folgende, recht starke Satz:

Satz 3.36 (Approximationssatz von Weierstraß)**.** *Für jede stetige Funktion f auf einem endlichen abgeschlossenen Intervall $[a,b]$ und jedes $\epsilon > 0$ gibt es ein Polynom P, das sich von f nirgends mehr als ϵ unterscheidet:*

$$|f(x) - P(x)| < \epsilon \text{ für alle } x \in [a,b].$$

3.7 Reelle und komplexe Exponentialfunktion

Die Exponentialfunktion wurde schon zuvor als eine wichtige elementare Funktion genannt. Will man formal definieren, was mit einem Ausdruck wie 2^π eigentlich gemeint ist, so ist das gar nicht so offensichtlich wie bei $2^3 = 2 \cdot 2 \cdot 2$. Man kann die (natürliche) Exponentialfunktion $\exp(x) = e^x$ auf verschiedene Arten einführen. Zwei Möglichkeiten seien hier genannt:

- als Grenzwert einer Folge: $\exp(x) = \lim_{n\to\infty}(1+\frac{x}{n})^n$. Es lässt sich zeigen, dass dieser Grenzwert tatsächlich für jedes reelle x existiert. Als Spezialfall sei hier genannt:

  $$e = e^1 = \exp(1) = \lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n \approx 2.71828.$$

- als unendliche Reihe:

  $$\exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \cdots .$$

  Diese Reihe konvergiert (absolut) für alle x, wie man mit Hilfe des Quotientenkriteriums zeigen kann. Insbesondere gilt

  $$e = e^1 = \exp(1) = \sum_{n=0}^{\infty} \frac{1}{n!} = \frac{1}{0!} + \frac{1}{1!} + \frac{1}{2!} + \cdots \approx 2.71828.$$

Den Zusammenhang zwischen den beiden Varianten sieht man unter Zuhilfenahme des binomischen Lehrsatzes:

$$\left(1+\frac{x}{n}\right)^n = \sum_{k=0}^{n} \binom{n}{k} 1^{n-k} \left(\frac{x}{n}\right)^k = \sum_{k=0}^{n} \binom{n}{k} \frac{x^k}{n^k}.$$

Für jedes feste k gilt:

$$\binom{n}{k}\frac{x^k}{n^k} = \frac{n!x^k}{k!(n-k)!n^k} = \frac{n(n-1)\cdots(n-k+1)x^k}{k!n^k} = \frac{n}{n}\cdot\frac{n-1}{n}\cdot\frac{n-2}{n}\cdots\frac{n-k+1}{n}\cdot\frac{x^k}{k!}$$
$$= 1\cdot\Big(1-\frac{1}{n}\Big)\cdot\Big(1-\frac{2}{n}\Big)\cdots\Big(1-\frac{k-1}{n}\Big)\cdot\frac{1}{k!}.$$

Jeder der Faktoren geht für $n\to\infty$ gegen 1. Der k-te Term der Summe ist also im Grenzwert $\frac{x^k}{k!}$, was die Identität

$$\lim_{n\to\infty}\Big(1+\frac{x}{n}\Big)^n = \sum_{k=0}^{\infty}\binom{n}{k}\frac{x^k}{n^k}$$

erklärt[2].

Aus der Darstellung als unendliche Reihe sieht man, dass die erwarteten Eigenschaften der Exponentialfunktion tatsächlich gelten: es ist $\exp(0)=1$, und es gilt die Identität $\exp(x)\exp(y) = e^xe^y = e^{x+y} = \exp(x+y)$. Dies erhält man indem man das Produkt der Reihen ausmultipliziert, wie im Kapitel über Reihen beschrieben:

$$\exp(x)\exp(y) = \Big(\sum_{n=0}^{\infty}\frac{x^n}{n!}\Big)\Big(\sum_{n=0}^{\infty}\frac{y^n}{n!}\Big) = \sum_{n=0}^{\infty}c_n,$$

wobei

$$c_n = \sum_{k=0}^{n}\frac{x^k}{k!}\frac{y^{n-k}}{(n-k)!} = \sum_{k=0}^{n}\frac{1}{n!}\cdot\frac{n!}{k!(n-k)!}\cdot x^ky^{n-k} = \frac{1}{n!}\sum_{k=0}^{n}\binom{n}{k}x^ky^{n-k} = \frac{1}{n!}(x+y)^n.$$

Hier wurde im letzten Schritt der binomische Lehrsatz verwendet. Damit gilt also

$$\exp(x)\exp(y) = \sum_{n=0}^{\infty}\frac{(x+y)^n}{n!} = \exp(x+y). \tag{3.2}$$

Es folgt aus der Definition, dass $\exp(x)$ für positives x streng monoton wachsend ist (da dies für jeden Term der Reihe gilt). Da wegen (3.2) $\exp(-x)=\frac{1}{\exp(x)}$ gilt, muss $\exp(x)$ auch für negatives x streng monoton wachsend sein. Daher ist exp injektiv, und zudem stetig[3]. Wegen $\exp(x)\geq 1+x$ muss $\lim_{x\to\infty}\exp(x)=\infty$ gelten, und wegen $\exp(-x)=\frac{1}{\exp(x)}$ zusätzlich $\lim_{x\to-\infty}\exp(x)=0$.

Damit wissen wir, dass der Wertebereich $(0,\infty)$ ist, womit sich für die Umkehrfunktion $\ln x$ wie schon erwähnt der Definitionsbereich $(0,\infty)$ ergibt. Unter Zuhilfenahme der Exponentialfunktion und des Logarithmus kann man nun a^b für allgemeine Werte von a und b definieren: ist a eine positive Zahl, so bildet man zunächst den Logarithmus $\ln a$ und *definiert*

$$a^b = \exp\big((\ln a)\cdot b\big).$$

[2]Man muss hier streng genommen argumentieren, warum man den Grenzwert dadurch bilden kann, dass man für jeden Term der Summe einzeln den Grenzwert bestimmt. Die Details seien hier ausgespart.

[3]Dass die Summe *unendlich vieler* stetiger Funktionen noch immer stetig ist, ist im allgemeinen falsch; im konkreten Fall kann man aber zeigen, dass es der Fall ist. Siehe Beispiel 5.43 in Kapitel 5.4.3.

Damit hat man auch dann eine formale Definition, wenn b keine ganze oder rationale Zahl ist. Alle gewohnten Rechenregeln sind mit dieser Definition kompatibel: zum Beispiel gilt $\ln a^b = (\ln a) \cdot b$ (gemäß obiger Definition), und damit

$$(a^b)^c = e^{(\ln a) \cdot b \cdot c} = a^{b \cdot c}.$$

Allgemeine Logarithmen lassen sich durch ln ausdrücken: $\log_a b = c$ bedeutet $b = a^c$. Damit ist

$$\ln b = \ln a^c = (\ln a) \cdot c,$$

und damit

$$\log_a b = c = \frac{\ln b}{\ln a}.$$

All das lässt sich auch auf die komplexen Zahlen erweitern: beide Definitionen von $\exp x$ bleiben sinnvoll und gültig, wenn x eine komplexe Zahl ist. Betrachten wir den Fall, dass x rein imaginär ist, also $x = iy$ für ein reelles y. Die Potenzen von i sind der Reihe nach $1, i, -1, -i, 1, i, -1, -i, \ldots$. Es ergibt sich daher

$$\exp(iy) = 1 + iy - \frac{y^2}{2!} - i\frac{y^3}{3!} + \frac{y^4}{4!} + i\frac{y^5}{5!} - \frac{y^6}{6!} - i\frac{y^7}{7!} + \cdots$$

Wie wir später (Kapitel 5.4.2) sehen werden, sind der reelle und imaginäre Teil genau die Reihendarstellungen von Cosinus und Sinus:

$$\cos y = 1 - \frac{y^2}{2!} + \frac{y^4}{4!} - \frac{y^6}{6!} + \cdots$$

und

$$\sin y = y - \frac{y^3}{3!} + \frac{y^5}{5!} - \frac{y^7}{7!} + \cdots$$

Darin begründet sich die sogenannte *Eulersche Formel*:

$$e^{iy} = \cos y + i \sin y.$$

Daraus folgt auch für jede reelle Zahl y, dass $|e^{iy}| = \sqrt{\cos^2 y + \sin^2 y} = \sqrt{1} = 1$. Allgemeiner gilt:

Satz 3.37. *Für eine komplexe Zahl $z = x + iy$ ist*

$$e^{x+iy} = e^x \big(\cos y + i \sin y \big).$$

Beispiel 3.38. Es ist

$$e^{2+i\pi/3} = e^2 \left(\cos \frac{\pi}{3} + i \sin \frac{\pi}{3} \right) = e^2 \left(\frac{1}{2} + i \frac{\sqrt{3}}{2} \right) \approx 3.69453 + 6.39911i.$$

Man beachte, dass sich die Polarform einer komplexen Zahl, also

$$a + bi = r(\cos \varphi + i \sin \varphi),$$

damit als

$$a + bi = re^{i\varphi}$$

ausdrücken lässt. Damit wird etwa aus der vom Potenzieren von komplexen Zahlen schon bekannten Formel

$$(\cos\varphi + i\sin\varphi)^n = \cos(n\varphi) + i\sin(n\varphi),$$

der sogenannten *Formel von de Moivre*, der kompakte Ausdruck

$$\left(e^{i\varphi}\right)^n = e^{in\varphi}.$$

Mit Hilfe der Polarform kann man auch einen komplexen Logarithmus definieren[4]. Für eine komplexe Zahl in Polardarstellung $z = re^{i\varphi}$, wobei $r = |z|$ und φ das Argument von z im Intervall $(-\pi, \pi]$ ist, setzen wir $\ln z = \ln r + i\varphi$. Dann gilt tatsächlich $e^{\ln z} = e^{\ln r} \cdot e^{i\varphi} = re^{i\varphi} = z$.

Auf diese Art kann man auch Ausdrücken wie i^i einen Sinn geben[5], indem man die Definition $a^b = e^{(\ln a)\cdot b}$ verwendet: in Polarform ist $i = e^{i\pi/2}$, und damit

$$i^i = e^{i\pi/2 \cdot i} = e^{-\pi/2}.$$

Auch trigonometrische und hyperbolische Funktionen lassen sich für komplexe Zahlen erklären. Aus

$$e^{iy} = \cos y + i\sin y$$

und

$$e^{-iy} = \cos(-y) + i\sin(-y) = \cos y - i\sin y$$

ergibt sich durch Addition $2\cos y = e^{iy} + e^{-iy}$ und durch Subtraktion $2i\sin y = e^{iy} - e^{-iy}$. Daher gilt

$$\cos y = \frac{e^{iy} + e^{-iy}}{2} \text{ und } \sin y = \frac{e^{iy} - e^{-iy}}{2i}.$$

Durch diese Gleichungen kann man Sinus und Cosinus auch für beliebige komplexe Werte definieren.

Beispiel 3.39. Es ist

$$\begin{aligned}\cos(\pi + i) &= \frac{e^{i\pi-1} + e^{-i\pi+1}}{2} = \frac{e^{-1}(\cos\pi + i\sin\pi) + e(\cos(-\pi) + i\sin(-\pi))}{2} \\ &= \frac{-e^{-1} - e}{2} \approx -1.54308\end{aligned}$$

und

$$\begin{aligned}\sin(\pi + i) &= \frac{e^{i\pi-1} - e^{-i\pi+1}}{2i} = \frac{e^{-1}(\cos\pi + i\sin\pi) - e(\cos(-\pi) + i\sin(-\pi))}{2i} \\ &= \frac{-e^{-1} + e}{2i} = \frac{e^{-1} - e}{2} i \approx -1.17520i.\end{aligned}$$

Man beachte den folgenden Zusammenhang zwischen den trigonometrischen Funktionen sin und cos sowie den hyperbolischen Funktionen sinh und cosh: es gilt

$$\sin(iz) = i\sinh z \text{ und } \cos(iz) = \cosh z$$

für alle komplexen Zahlen z.

[4] Streng genommen den sogenannten *Hauptwert*, denn der komplexe Logarithmus ist nicht eindeutig.
[5] Wieder handelt es sich im Folgenden um den *Hauptwert*.

Beispiel 3.40. Man bestimme den Wert der Reihe

$$\sum_{n=0}^{\infty} \frac{\cos n}{2^n}.$$

Lösung. Wir schreiben $\cos n$ als $\frac{e^{in}+e^{-in}}{2}$. Damit ergibt sich

$$\sum_{n=0}^{\infty} \frac{\cos n}{2^n} = \sum_{n=0}^{\infty} \frac{e^{in}+e^{-in}}{2\cdot 2^n} = \frac{1}{2}\Big(\sum_{n=0}^{\infty} \frac{e^{in}}{2^n} + \sum_{n=0}^{\infty} \frac{e^{-in}}{2^n}\Big) = \frac{1}{2}\Big(\sum_{n=0}^{\infty} \Big(\frac{e^{i}}{2}\Big)^n + \sum_{n=0}^{\infty} \Big(\frac{e^{-i}}{2}\Big)^n\Big).$$

Die Reihe ist also die Summe zweier (komplexer) geometrischer Reihen, wobei $|\frac{e^i}{2}| = \frac{|e^i|}{2} = \frac{1}{2}$ und $|\frac{e^{-i}}{2}| = \frac{|e^{-i}|}{2} = \frac{1}{2}$. Die Reihen sind damit konvergent, und es gilt

$$\begin{aligned}\sum_{n=0}^{\infty} \frac{\cos n}{2^n} &= \frac{1}{2}\Big(\frac{1}{1-e^i/2} + \frac{1}{1-e^{-i}/2}\Big) = \frac{1}{2}\cdot\frac{1-e^{-i}/2+1-e^{i}/2}{(1-e^i/2)(1-e^{-i}/2)} \\ &= \frac{1}{2}\cdot\frac{2-e^i/2-e^{-i}/2}{1-e^i/2-e^{-i}/2+1/4} = \frac{1}{2}\cdot\frac{2-\cos 1}{5/4-\cos 1} \\ &= \frac{4-2\cos 1}{5-4\cos 1} \approx 1.02839.\end{aligned}$$

Kapitel 4

Lineare Algebra

4.1 Geometrie in zwei und drei Dimensionen

In Arbeit.

4.2 Vektorräume

In Arbeit.

4.3 Matrizen und lineare Abbildungen

4.3.1 Matrizen

Definition 4.1. Eine *Matrix* ist ein rechteckiges Zahlenschema. Eine $m \times n$-Matrix hat m *Zeilen* und n *Spalten*. Wir bezeichnen die Einträge einer Matrix A üblicherweise mit a_{ij} $(1 \leq i \leq m,$ $1 \leq j \leq n)$, wobei der erste Index i die Zeile und der zweite Index j die Spalte angibt:

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

Man beachte, dass Zeilenvektoren als $1 \times n$-Matrizen und Spaltenvektoren als $m \times 1$-Matrizen betrachtet werden können.

Beispiel 4.2. Die folgenden Matrizen haben die Dimensionen 3×4, 2×5 und 3×1:

$$\begin{pmatrix} 3 & 0 & -\pi & \sqrt{2} \\ e^5 & -1 & \frac{1}{2} & 4 \\ 2 & \ln 3 & -7 & 0 \end{pmatrix}, \qquad \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 4 & 3 & 2 & 1 \end{pmatrix}, \qquad \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}.$$

Zwei Matrizen gelten als gleich, wenn sie die gleichen Dimensionen haben und alle Einträge gleich sind: $A = B$, wenn für die entsprechenden Einträge a_{ij} und b_{ij} die Gleichheit $a_{ij} = b_{ij}$ für alle i und j gilt.

Wir definieren zunächst einige Operationen auf Matrizen:

- *Addition* und *Subtraktion* von Matrizen A und B ist genau dann definiert, wenn A und B die gleichen Dimensionen haben. Entsprechende Einträge werden dabei jeweils miteinander addiert bzw. subtrahiert:
$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 4 & 3 & 2 & 1 \end{pmatrix} + \begin{pmatrix} -1 & 2 & -3 & 4 & -5 \\ 0 & 1 & 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 4 & 0 & 8 & 0 \\ 5 & 5 & 3 & 3 & 1 \end{pmatrix}$$
und
$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 4 & 3 & 2 & 1 \end{pmatrix} - \begin{pmatrix} -1 & 2 & -3 & 4 & -5 \\ 0 & 1 & 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 0 & 6 & 0 & 10 \\ 5 & 3 & 3 & 1 & 1 \end{pmatrix}.$$
- Das *skalare Vielfache* λA einer Matrix A mit einem Skalar λ wird gebildet, indem man jeden Eintrag von A mit λ multipliziert:
$$3 \cdot \begin{pmatrix} 1 & 2 & 3 \\ -4 & -5 & -6 \end{pmatrix} = \begin{pmatrix} 3 & 6 & 9 \\ -12 & -15 & -18 \end{pmatrix}.$$
- *Multiplikation* von Matrizen A und B ist genau dann möglich, wenn A genauso viele Spalten hat, wie B Zeilen hat: es sei also A eine $m \times n$-Matrix und B eine $n \times \ell$-Matrix. Das Produkt C ist dann eine $m \times \ell$-Matrix.

  Es sei $\vec{a}_i$ die i-te Zeile von A (mit Einträgen $a_{i1}, a_{i2}, \ldots$) und $\vec{b}_j$ die j-te Spalte von B (mit Einträgen $b_{1j}, b_{2j}, \ldots$). Man erhält den Eintrag c_{ij} der Matrix $C = A \cdot B$, indem man das Skalarprodukt von $\vec{a}_i$ und $\vec{b}_j$ bildet, also
$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj} = \sum_{k=1}^{n} a_{ik}b_{kj}.$$
  So ist zum Beispiel
$$\begin{aligned} &\begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \cdot \begin{pmatrix} 0 & -2 & 4 & -6 \\ -1 & 3 & -5 & 7 \end{pmatrix} \\ &\qquad = \begin{pmatrix} 1 \cdot 0 + 2 \cdot (-1) & 1 \cdot (-2) + 2 \cdot 3 & 1 \cdot 4 + 2 \cdot (-5) & 1 \cdot (-6) + 2 \cdot 7 \\ 3 \cdot 0 + 4 \cdot (-1) & 3 \cdot (-2) + 4 \cdot 3 & 3 \cdot 4 + 4 \cdot (-5) & 3 \cdot (-6) + 4 \cdot 7 \\ 5 \cdot 0 + 6 \cdot (-1) & 5 \cdot (-2) + 6 \cdot 3 & 5 \cdot 4 + 6 \cdot (-5) & 5 \cdot (-6) + 6 \cdot 7 \end{pmatrix} \\ &\qquad = \begin{pmatrix} -2 & 4 & -6 & 8 \\ -4 & 6 & -8 & 10 \\ -6 & 8 & -10 & 12 \end{pmatrix}. \end{aligned}$$

Beispiel 4.3. Die Drehung eines zweidimensionalen Vektors $\begin{pmatrix} a \\ b \end{pmatrix}$ um 90° gegen den Uhrzeigersinn liefert den Vektor $\begin{pmatrix} -b \\ a \end{pmatrix}$. Man kann dies durch eine Matrixmultiplikation darstellen:
$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} -b \\ a \end{pmatrix}.$$
Allgemeiner lassen sich Drehungen um einen Winkel α durch Multiplikation mit der Matrix
$$\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$$
ausdrücken.

Die *Transponierte* A^t einer $m \times n$-Matrix A ist die $n \times m$-Matrix, die durch das Vertauschen der Rollen von Zeilen und Spalten ergibt: beispielsweise ist

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}^t = \begin{pmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{pmatrix}.$$

Die $m \times n$-*Nullmatrix* $\mathbf{0}$ ist jene Matrix, deren Einträge sämtlich 0 sind. Eine *Einheitsmatrix* I ist eine quadratische Matrix, für die die Einträge in der Diagonale von links oben nach rechts unten sämtlich gleich 1 sind, alle anderen dagegen 0. So ist die 3×3-Einheitsmatrix

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

Sie ist ein Spezialfall einer sogenannten *Diagonalmatrix*: alle Einträge einer Diagonalmatrix, die nicht in der Diagonalen von links oben nach rechts unten stehen, sind gleich 0.

Die Rechenoperationen erfüllen einige wichtige Eigenschaften, die im Folgenden aufgelistet sind.

- Addition ist kommutativ: $A + B = B + A$. Multiplikation dagegen ist nicht kommutativ: es ist möglich, dass $A \cdot B$ definiert ist, aber nicht $B \cdot A$; selbst wenn $A \cdot B$ und $B \cdot A$ wohldefiniert sind und die gleiche Dimension haben, müssen sie nicht gleich sein.

- Addition und Multiplikation erfüllen das Assoziativgesetz:

  $$A + (B + C) = (A + B) + C, \qquad A \cdot (B \cdot C) = (A \cdot B) \cdot C.$$

  Zudem gilt

  $$\lambda(A \cdot B) = (\lambda A) \cdot B = A \cdot (\lambda B)$$

  für Skalare λ.

- Addition und Multiplikation erfüllen das Distributivgesetz:

  $$\lambda(A + B) = \lambda A + \lambda B$$

  für die Multiplikation mit Skalaren und

  $$A \cdot (B + C) = A \cdot B + A \cdot C, \qquad (A + B) \cdot C = A \cdot C + B \cdot C$$

  für die Matrixmultiplikation.

- Nullmatrizen und Einheitsmatrizen spielen die Rollen von 0 und 1:

  $$A + \mathbf{0} = \mathbf{0} + A = A$$

  sowie

  $$A \cdot \mathbf{0} = \mathbf{0}, \qquad \mathbf{0} \cdot A = \mathbf{0}$$

  und

  $$A \cdot I = A, \qquad I \cdot A = A$$

  sofern die entsprechenden Summen und Produkte wohldefiniert sind (also die Dimensionen geeignet sind).

- Es gilt $(A+B)^t = A^t + B^t$ und $(A \cdot B)^t = B^t \cdot A^t$. Beim Transponieren eines Produkts wird also die Reihenfolge vertauscht!

Aus den genannten Eigenschaften folgt außerdem, dass die $m \times n$-Matrizen einen Vektorraum bilden. Die Dimension ist mn, und eine Basis besteht aus allen Matrizen, für die ein Eintrag gleich 1 und alle anderen gleich 0 sind.

4.3.2 Lineare Abbildungen

Die Rotation von Vektoren in Beispiel 4.3 ist ein Beispiel einer *linearen Abbildung*. Eine lineare Abbildung ist eine Funktion $f : V \to W$ zwischen zwei Vektorräumen V und W, die die folgenden Eigenschaften für alle Vektoren $\vec{v}, \vec{w} \in V$ und alle Skalare λ hat:

$$f(\vec{v} + \vec{w}) = f(\vec{v}) + f(\vec{w}), \qquad f(\lambda \vec{v}) = \lambda f(\vec{v}).$$

Einige weitere Beispiele linearer Abbildungen:

- Die Abbildung $\pi_i : \mathbb{R}^n \to \mathbb{R}$, die jeden Vektor $\vec{a} = (a_1, a_2, \ldots, a_n)$ auf seine i-te Koordinate $\pi_i(\vec{a}) = a_i$ abbildet, ist linear (sogenannte *Projektion*).

- Die Abbildung $\alpha_{m,n} : \mathbb{R}^m \to \mathbb{R}^n$ $(m \leq n)$, die jeden Vektor $\vec{a} = (a_1, a_2, \ldots, a_m)$ mit $n-m$ Nullen auffüllt – also $\alpha_{m,n}(\vec{a}) = (a_1, a_2, \ldots, a_m, 0, 0, \ldots, 0)$ – ist linear (sogenannte *Einbettung*).

- Die Multiplikation mit einer Konstanten ($f(\vec{x}) = \lambda \vec{x}$) ist eine lineare Abbildung.

- Das Skalarprodukt mit einem festen Vektor $\vec{a}$ (also $f(\vec{x}) = \langle \vec{a}, \vec{x} \rangle = \vec{a} \cdot \vec{x}$) ist eine lineare Abbildung.

- Die Abbildung $f : V \to W$, die jedem Vektor $\vec{v} \in V$ den Nullvektor $\vec{0}$ in W zuordnet (also $f(\vec{v}) = \vec{0}$) ist für beliebige Vektorräume V, W eine lineare Abbildung.

Aus der Definition einer linearen Abbildung $f : V \to W$ folgt, dass der Nullvektor von V immer auf den Nullvektor von W abgebildet werden muss (dies ist der Spezialfall $\lambda = 0$ der zweiten Eigenschaft). Durch wiederholtes Anwenden der beiden Eigenschaften erhält man zudem die allgemeine Identität

$$f(\lambda_1 \vec{v}_1 + \lambda_2 \vec{v}_2 + \cdots + \lambda_k \vec{v}_k) = \lambda_1 f(\vec{v}_1) + \lambda_2 f(\vec{v}_2) + \cdots + \lambda_k f(\vec{v}_k). \tag{4.1}$$

Die Multiplikation mit einer Matrix A ist ebenfalls eine lineare Abbildung. Dies folgt aus den genannten Rechenregeln:

$$A \cdot (\vec{v} + \vec{w}) = A \cdot \vec{v} + A \cdot \vec{w}$$

und

$$A \cdot (\lambda \vec{v}) = \lambda (A \cdot \vec{v}).$$

Tatsächlich lassen sich alle linearen Abbildungen in gewissem Sinne durch eine Matrixmultiplikation beschreiben. Es sei $f : V \to W$ linear, und es seien $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ und $\vec{w}_1, \vec{w}_2, \ldots, \vec{w}_m$

Basen der Vektorräume V und W. Die Bilder $f(\vec{v}_1), f(\vec{v}_2), \ldots, f(\vec{v}_n)$ der Basisvektoren von V können als Linearkombinationen der Basisvektoren $\vec{w}_1, \vec{w}_2, \ldots, \vec{w}_m$ ausgedrückt werden:

$$\begin{aligned}
f(\vec{v}_1) &= a_{11}\vec{w}_1 + a_{21}\vec{w}_2 + \cdots + a_{m1}\vec{w}_m, \\
f(\vec{v}_2) &= a_{12}\vec{w}_1 + a_{22}\vec{w}_2 + \cdots + a_{m2}\vec{w}_m, \\
\vdots \quad & \quad \vdots \qquad\qquad\quad \vdots \\
f(\vec{v}_n) &= a_{1n}\vec{w}_1 + a_{2n}\vec{w}_2 + \cdots + a_{mn}\vec{w}_m.
\end{aligned}$$

Jeder Vektor $\vec{v} \in V$ lässt sich als Linearkombination der Basisvektoren schreiben:

$$\vec{v} = b_1\vec{v}_1 + b_2\vec{v}_2 + \cdots + b_n\vec{v}_n.$$

Aus (4.1) folgt nun, dass

$$f(\vec{v}) = \sum_{k=1}^{n} b_k f(\vec{v}_k) = \sum_{k=1}^{n} b_k \sum_{j=1}^{m} a_{jk}\vec{w}_j = \sum_{j=1}^{m} \Big(\sum_{k=1}^{n} a_{jk} b_k \Big) \vec{w}_j.$$

Die Koordinaten von $f(\vec{v})$ bezüglich der Basis $\vec{w}_1, \vec{w}_2, \ldots, \vec{w}_m$ sind damit genau die Einträge des Matrixprodukts

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}.$$

Für den Spezialfall $V = \mathbb{R}^n$, $W = \mathbb{R}^m$ (die wir jeweils mit ihren Standardbasen betrachten) gilt insbesondere folgender Satz:

Satz 4.4. *Jede lineare Abbildung $f : \mathbb{R}^n \to \mathbb{R}^m$ lässt sich als Multiplikation mit einer $m \times n$-Matrix ausdrücken:*

$$f(\vec{x}) = A \cdot \vec{x}.$$

Die Hintereinanderausführung zweier linearer Abbildungen entspricht genau der Multiplikation der Matrizen: es seien $f : \mathbb{R}^n \to \mathbb{R}^m$ und $g : \mathbb{R}^\ell \to \mathbb{R}^n$ zwei lineare Abbildungen, und es seien A und B die entsprechenden Matrizen, sodass

$$f(\vec{x}) = A \cdot \vec{x} \quad \text{und} \quad g(\vec{y}) = B \cdot \vec{y}$$

für alle Vektoren $\vec{x}, \vec{y}$ gilt. Dann ist

$$f(g(\vec{y})) = f(B \cdot \vec{y}) = A \cdot (B \cdot \vec{y}) = (A \cdot B) \cdot \vec{y}.$$

Die Matrix, die der Abbildung $f \circ g$ entspricht, ist also genau $A \cdot B$.

Beispiel 4.5. Die lineare Abbildung $f : \mathbb{R}^2 \to \mathbb{R}^3$ erfüllt[1] $f(2,7) = (5,-1,0)$ und $f(1,5) = (4,1,-3)$. Man bestimme die zugehörige Matrix.

[1]Der Einfachheit halber schreiben wir $f(2,3)$ anstelle von $f((2,3))$.

Lösung. Für die Matrix benötigen wir die Bilder der Basisvektoren $(1,0)$ und $(0,1)$. Diese können wir aus $(2,7)$ und $(1,5)$ durch Elimination jeweils einer Koordinaten gewinnen: $(3,0) = 5\cdot(2,7) - 7\cdot(1,5)$, also $(1,0) = \frac{5}{3}\cdot(2,7) - \frac{7}{3}\cdot(1,5)$. Damit folgt aus der Linearität

$$f(1,0) = \frac{5}{3}f(2,7) - \frac{7}{3}f(1,5) = (-1,-4,7).$$

Ebenso ist $(0,3) = 2\cdot(1,5) - (2,7)$, also $(0,1) = \frac{2}{3}\cdot(1,5) - \frac{1}{3}\cdot(2,7)$. Es folgt

$$f(0,1) = \frac{2}{3}f(1,5) - \frac{1}{3}f(2,7) = (1,1,-2).$$

Die Matrix, die f zugeordnet ist, ist somit

$$A = \begin{pmatrix} -1 & 1 \\ -4 & 1 \\ 7 & -2 \end{pmatrix}.$$

Bemerkung 4.6. Lineare Abbildungen $f, g : V \to W$ können addiert und mit Skalaren multipliziert werden:

$$(f+g)(\vec{x}) = f(\vec{x}) + g(\vec{x}), \qquad (\lambda f)(\vec{x}) = \lambda f(\vec{x}).$$

Man kann nachprüfen, dass $f+g$ und λf wieder lineare Abbildungen sind. Die Addition bzw. Multiplikation mit Skalaren entspricht auch der Addition bzw. Multiplikation mit Skalaren der zugehörigen Matrizen. Die linearen Abbildungen bilden also ihrerseits einen Vektorraum. Wenn die Dimensionen von V und W dabei m und n sind, dann hat der Vektorraum der linearen Abbildungen Dimension mn. Insbesondere bildet für einen reellen Vektorraum V die Menge der linearen Abbildungen $f : V \to \mathbb{R}$ einen Vektorraum, der dieselbe Dimension wie V hat, den sogenannten *Dualraum*.

Definition 4.7. Der *Kern* $\operatorname{Ker} f$ einer linearen Abbildung $f : V \to W$ ist die Menge aller Vektoren $\vec{v} \in V$, die auf den Nullvektor abbilden, für die also $f(\vec{v}) = \vec{0}$ gilt. Das *Bild* $\operatorname{Im} f$ einer linearen Abbildung $f : V \to W$ ist die Menge aller Vektoren der Form $\vec{w} = f(\vec{v})$.

Man beachte:

- $\vec{0} \in \operatorname{Ker} f$, und für zwei Vektoren $\vec{u}, \vec{v} \in \operatorname{Ker} f$ und einen Skalar λ gilt aufgrund der Linearität von f

  $$f(\vec{u} + \vec{v}) = f(\vec{u}) + f(\vec{v}) = \vec{0} + \vec{0} = \vec{0}$$

  sowie

  $$f(\lambda\vec{v}) = \lambda f(\vec{v}) = \lambda\vec{0} = \vec{0}.$$

  Also liegen $\vec{u} + \vec{v}$ und $\lambda\vec{v}$ wiederum in $\operatorname{Ker} f$. Dies bedeutet, dass $\operatorname{Ker} f$ ein Untervektorraum von V ist.

- $\vec{0} = f(\vec{0}) \in \operatorname{Im} f$, und für zwei Vektoren $\vec{w} = f(\vec{u})$ und $\vec{x} = f(\vec{v})$ und einen Skalar λ gilt

  $$\vec{w} + \vec{x} = f(\vec{u} + \vec{v})$$

  sowie

  $$\lambda\vec{w} = f(\lambda\vec{u}).$$

  Also liegen $\vec{w} + \vec{x}$ und $\lambda\vec{w}$ wiederum in $\operatorname{Im} f$. Dies bedeutet, dass $\operatorname{Im} f$ ein Untervektorraum von W ist.

Beispiel 4.8. Die lineare Abbildung $f : \mathbb{R}^2 \to \mathbb{R}^3$ sei durch $f(x, y) = (2x+2y, x+y, 0)$ gegeben. Der Kern besteht dann aus allen Vektoren (x, y) mit $y = -x$, also allen Vektoren der Form $(x, -x)$. Diese formen einen eindimensionalen Vektorraum. Das Bild dagegen besteht aus allen Vielfachen des Vektors $(2, 1, 0)$ und ist ebenfalls eindimensional.

Dieses Beispiel illustriert den folgenden allgemeinen Satz.

Satz 4.9 (Dimensionssatz)**.** *Die Summe aus der Dimension des Kerns und der Dimension des Bildes einer linearen Abbildung $f : V \to W$ ist gleich der Dimension von V:*

$$\dim \operatorname{Ker} f + \dim \operatorname{Im} f = \dim V.$$

Beweis. Es sei $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_r$ eine Basis des Kerns, und $\vec{w}_1, \vec{w}_2, \ldots, \vec{w}_s$ eine Basis des Bildes. Nach Definition des Bildes gibt es Vektoren $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_s$, sodass $f(\vec{u}_i) = \vec{w}_i$ für $1 \leq i \leq s$ gilt. Wir beweisen nun, dass $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_s, \vec{v}_1, \vec{v}_2, \ldots, \vec{v}_r$ eine Basis von V ist.

- Zunächst zeigen wir, dass diese Vektoren V aufspannen: es sei $\vec{x}$ ein beliebiger Vektor in V. Nun hat $f(\vec{x})$ (was im Bild liegt) eine Darstellung durch die Basisvektoren des Bildes:

  $$f(\vec{x}) = a_1\vec{w}_1 + a_2\vec{w}_2 + \cdots a_s\vec{w}_s = a_1 f(\vec{u}_1) + a_2 f(\vec{u}_2) + \cdots + a_s f(\vec{u}_s).$$

  Aus der Linearität von f folgt nun

  $$f(\vec{x} - a_1\vec{u}_1 - a_2\vec{u}_2 - \cdots - a_s\vec{u}_s) = \vec{0}.$$

  Damit liegt $\vec{x} - a_1\vec{u}_1 - a_2\vec{u}_2 - \cdots - a_s\vec{u}_s$ im Kern von f und kann als $b_1\vec{v}_1 + b_2\vec{v}_2 + \cdots + b_r\vec{v}_r$ geschrieben werden. Das liefert eine Darstellung von $\vec{x}$:

  $$\vec{x} = a_1\vec{u}_1 + a_2\vec{u}_2 + \cdots + a_s\vec{u}_s + b_1\vec{v}_1 + b_2\vec{v}_2 + \cdots + b_r\vec{v}_r.$$

- Nun zeigen wir noch lineare Unabhängigkeit: es gelte

  $$\vec{0} = a_1\vec{u}_1 + a_2\vec{u}_2 + \cdots + a_s\vec{u}_s + b_1\vec{v}_1 + b_2\vec{v}_2 + \cdots + b_r\vec{v}_r.$$

  Wir wenden f auf diese Gleichung an und verwenden die Linearität sowie die Tatsache, dass $f(\vec{u}_i) = \vec{w}_i$ und $f(\vec{v}_j) = \vec{0}$ für alle i und j gilt:

  $$\begin{aligned}\vec{0} = f(\vec{0}) &= a_1 f(\vec{u}_1) + a_2 f(\vec{u}_2) + \cdots + a_s f(\vec{u}_s) + b_1 f(\vec{v}_1) + b_2 f(\vec{v}_2) + \cdots + b_r f(\vec{v}_r) \\ &= a_1\vec{w}_1 + a_2\vec{w}_2 + \cdots + a_s\vec{w}_s + b_1\vec{0} + b_2\vec{0} + \cdots + b_r\vec{0} \\ &= a_1\vec{w}_1 + a_2\vec{w}_2 + \cdots + a_s\vec{w}_s.\end{aligned}$$

  Da die Vektoren $\vec{w}_1, \vec{w}_2, \ldots, \vec{w}_s$ linear unabhängig sind, folgt damit $a_1 = a_2 = \cdots = a_s = 0$, also

  $$\vec{0} = b_1\vec{v}_1 + b_2\vec{v}_2 + \cdots + b_r\vec{v}_r.$$

  Aus der linearen Unabhängigkeit von $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_r$ folgt nun aber $b_1 = b_2 = \cdots = b_r = 0$.

Wir haben also gezeigt, dass $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_s, \vec{v}_1, \vec{v}_2, \ldots, \vec{v}_r$ den Vektorraum V aufspannen und linear unabhängig sind. Also bilden sie eine Basis, und damit ist die Dimension von V gleich $s + r = \dim \operatorname{Im} f + \dim \operatorname{Ker} f$.

Es sei $f : \mathbb{R}^n \to \mathbb{R}^m$ eine lineare Abbildung, und A die zugeordnete $m \times n$-Matrix. Weiters seien $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n$ die Spaltenvektoren von A. Man kann $f(x_1, x_2, \ldots, x_n)$ als

$$A \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = x_1\vec{a}_1 + x_2\vec{a}_2 + \cdots + x_n\vec{a}_n$$

ausdrücken. Das Bild $\operatorname{Im} f$ besteht also genau aus allen Linearkombinationen der Spaltenvektoren $\vec{a}_1, \ldots, \vec{a}_n$. Man bezeichnet dies als den *Spaltenraum* von A. Der *Zeilenraum* wird analog definiert.

Definition 4.10. Der Spaltenraum von A ist der von den Spalten von A aufgespannte Raum. Seine Dimension wird *Spaltenrang* genannt. Der Zeilenraum von A ist der von den Zeilen von A aufgespannte Raum. Seine Dimension wird *Zeilenrang* genannt.

Zeilen- bzw. Spaltenrang sind gleich der maximalen Anzahl linear unabhängiger Zeilen bzw. Spalten. Offensichtlich muss der Zeilenrang einer $m \times n$-Matrix $\leq m$ und der Spaltenrang $\leq n$ sein.

Unter *elementaren Zeilen- und Spaltenumformungen* versteht man:

- Vertauschen zweier Zeilen oder zweier Spalten,
- Multiplikation einer Zeile oder Spalte mit einer Zahl $\lambda \neq 0$,
- Addition des λ-fachen einer Zeile (Spalte) zu einer anderen Zeile (Spalte), wobei λ ein Skalar ist.

Es gelten die folgenden wichtigen Sätze:

Satz 4.11. *Elementare Zeilen- und Spaltenumformungen verändern weder den Zeilenrang noch den Spaltenrang.*

Satz 4.12. *Jede Matrix lässt sich durch elementare Zeilen- und Spaltenumformungen auf die Form*

$$\begin{pmatrix} I_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \tag{4.2}$$

bringen, wobei I_r eine r-dimensionale Einheitsmatrix ist und $\mathbf{0}$ Nullmatrizen geeigneter Dimensionen bezeichnet.

Man beachte, dass für eine Matrix der Form (4.2) Spalten- und Zeilenrang gleich r sind. Wir können daraus folgern:

Satz 4.13. *Für jede Matrix A sind Zeilenrang und Spaltenrang gleich.*

Wir sprechen vom *Rang* der Matrix:

$$\text{Zeilenrang} = \text{Spaltenrang} = \text{Rang}.$$

Der Rang einer Matrix A wird als $\operatorname{rang} A$ geschrieben[2]. Nur Nullmatrizen haben den Rang 0, alle anderen Matrizen haben positiven Rang. Das Konzept des Ranges wird in weiterer Folge für die Behandlung linearer Gleichungssysteme von entscheidender Bedeutung sein.

Beschränkt man sich auf elementare Zeilenumformungen, dann ändert sich nicht nur die Dimension des Zeilenraumes nicht, sondern sogar der Zeilenraum selbst bleibt der gleiche. Um eine Basis dieses Raumes zu bestimmen, kann man eine Matrix durch elementare Zeilenumformungen auf sogenannte *Zeilenstufenform* bringen. Diese hat folgende Eigenschaften:

- Alle Nullzeilen stehen ganz unten.
- Der erste Eintrag $\neq 0$ in einer Zeile, die keine Nullzeile ist, wird *Pivotelement* genannt. Sind die i-te und j-te Zeile keine Nullzeilen, und ist $i < j$, dann muss das Pivotelement der j-ten Zeile rechts vom Pivotelement der i-ten Zeile stehen.

Man spricht von *reduzierter Zeilenstufenform*, wenn zudem gilt:

- Alle Pivotelemente sind gleich 1.
- Alle Elemente, die in der gleichen Spalte wie ein Pivotelement sind, sind gleich 0.

So ist die Matrix A hier in Zeilenstufenform, aber nicht in reduzierter Zeilenstufenform. Die Matrix B dagegen ist in reduzierter Zeilenstufenform:

$$A = \begin{pmatrix} 3 & -2 & 7 & 0 \\ 0 & 5 & 2 & -4 \\ 0 & 0 & 0 & -6 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 0 & 7 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Um eine beliebige Matrix in die Zeilenstufenform überzuführen, geht man folgendermaßen vor (*Gauß-Elimination*):

1. Suche die erste Spalte, die nicht nur aus Nullen besteht (wenn alle Spalten nur aus Nullen bestehen, ist die Matrix die Nullmatrix und wir sind fertig). Es sei dies die j-te Spalte.
2. Wähle eine Zeile, die in dieser Spalte einen Eintrag $\neq 0$ hat, und tausche sie gegen die erste Zeile aus (wenn die erste Zeile bereits die gewünschte Eigenschaft hat, ist nichts zu tun).
3. Das erste Pivotelement ist nun in der ersten Zeile und j-ten Spalte. Wir nennen es b_{1j}. Für jede Zeile in der der Eintrag in der j-ten Spalte nicht 0 ist, eliminieren wir ihn durch Abziehen eines geeigneten Vielfachen der ersten Zeile: ist der Eintrag in der i-ten Zeile/j-ten Spalte $b_{ij} \neq 0$, so wird das $\frac{b_{ij}}{b_{1j}}$-fache der ersten Zeile von der i-ten Zeile abgezogen. Dies wird so lange wiederholt, bis die j-te Spalte außer dem Pivoteintrag in der ersten Zeile nur Nullen enthält.

[2] Auch Varianten wie $\operatorname{rg} A$ kommen vor.

4. Ab diesem Punkt kann die erste Zeile ignoriert werden, und dieselben Schritte werden mit den restlichen Zeilen wiederholt. Man beachte, dass diese in den ersten j Spalten nur Nullen enthalten.

Will man zudem die reduzierte Zeilenstufenform erreichen, so führt man zusätzlich noch die folgenden Schritte durch:

5. Dividiere jede Zeile, die keine Nullzeile ist, durch ihr Pivotelement, das damit zu 1 wird (wenn es bereits 1 ist, ist dies natürlich nicht nötig).

6. In der Reihenfolge von unten nach oben werden geeignete Vielfache jeder Zeile mit einem Pivotelement von den darüberliegenden Zeilen subtrahiert, um die Elemente über dem Pivotelement zu eliminieren (d.h., in 0 zu verwandeln), auf dieselbe Art wie in Schritt 3.

In seiner Gesamtheit spricht man vom *Gauß–Jordan-Algorithmus.* Die Zeilen in der (reduzierten) Zeilenstufenform, die nicht nur Nullen enthalten, sind linear unabhängig und bilden eine Basis des Zeilenraumes. Damit ist es auch möglich, den Rang einer Matrix zu bestimmen.

4.4 Lineare Gleichungssysteme

4.4.1 Grundbegriffe und Lösungsverfahren

Eine *lineare Gleichung* in den Variablen $x_1, x_2, \ldots, x_n$ ist eine Gleichung der Form

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b.$$

Ein System aus mehreren solchen Gleichungen, also

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

wird *lineares Gleichungssystem* genannt. Wir interessieren uns für die Lösungsmenge solcher Gleichungssysteme, also die Menge aller Vektoren $(x_1, x_2, \ldots, x_n)$, die alle Gleichungen gleichzeitig erfüllen. In Matrixschreibweise kann man dies auch als

$$A \cdot \vec{x} = \vec{b}$$

ausdrücken, wobei

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \qquad \vec{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix},$$

und

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

Die Matrix A wird *Koeffizientenmatrix* genannt. Fügt man den Vektor $\vec{b}$ als zusätzliche Spalte am Ende hinzu, spricht man von der *erweiterten Koeffizientenmatrix* $(A, \vec{b})$. Ein lineares Gleichungssystem heißt

- *überbestimmt*, wenn $m > n$ (mehr Gleichungen als Variablen),
- *unterbestimmt*, wenn $m < n$ (mehr Variablen als Gleichungen),
- *quadratisch*, wenn $m = n$,
- *homogen*, wenn $\vec{b} = \vec{0}$,
- *inhomogen*, wenn $\vec{b} \neq \vec{0}$.

Manipulationen der Gleichungen können als elementare Zeilenoperationen der erweiterten Koeffizientenmatrix interpretiert werden:

- Das Vertauschen von Zeilen entspricht dem Vertauschen von Gleichungen.
- Das Multiplizieren einer Zeile mit einer Zahl $\lambda \neq 0$ entspricht der Multiplikation der entsprechenden Gleichung mit λ.
- Addition (eines Vielfachen) einer Zeile zu einer anderen Zeile entspricht der Addition (eines Vielfachen) einer Gleichung zu einer anderen Gleichung.

Keiner dieser Schritte ändert die Lösungsmenge. Man kann daher die Lösungsmenge dadurch bestimmen, dass man die erweiterte Koeffizientenmatrix durch Zeilenoperationen auf (reduzierte) Zeilenstufenform bringt und dann die Lösung durch Rückwärtseinsetzen bestimmt.

Bemerkung 4.14. Vorsicht: elementare Spaltenoperationen verändern die Lösungsmenge (z.B. entspricht das Vertauschen von Spalten dem Vertauschen von Variablen) und sollen daher in dieser Methode nicht verwendet werden!

Am Ende können folgende Szenarien auftreten:

1. Eine Zeile der Zeilenstufenform besteht nur aus Nullen, mit Ausnahme der letzten Spalte. Dies entspricht einer Gleichung der Form
$$0x_1 + 0x_2 + \cdots + 0x_n = b,$$
wobei $b \neq 0$. Dies ist klarerweise unmöglich, und es gibt keine Lösungen. Dieser Fall tritt dann ein, wenn $\operatorname{rang}(A, \vec{b}) > \operatorname{rang} A$.

2. Es gibt keine Zeile wie im ersten Fall, und in der Zeilenstufenform bleiben genauso viele Nicht-Nullzeilen übrig wie es Variablen gibt. Dann bestimmt man die letzte Variable aus der letzten Gleichung, damit die vorletzte Variable aus der vorletzten Gleichung, und so weiter. Die Lösung des Gleichungssystems ist eindeutig. Dieser Fall tritt dann ein, wenn $\operatorname{rang}(A, \vec{b}) = \operatorname{rang} A = n$.

3. Es gibt keine Zeile wie im ersten Fall, und in der Zeilenstufenform bleiben weniger Nicht-Nullzeilen übrig als es Variablen gibt. Dann kann man alle Variablen, die Spalten ohne Pivotelement entsprechen, als freie Parameter wählen und wie im zweiten Fall die restlichen Variablen von unten nach oben bestimmen. Man erhält unendlich viele Lösungen. Die Anzahl der freien Variablen wird auch *Freiheitsgrade* genannt. Dieser Fall tritt dann ein, wenn $\operatorname{rang}(A, \vec{b}) = \operatorname{rang} A = n - k$, wobei k die Anzahl der Freiheitsgrade ist. Für $k = 1$ bilden die Lösungen eine Gerade, für $k = 2$ eine Ebene, und so weiter.

Es ergeben sich einige unmittelbare Schlussfolgerungen:

- Ein lineares Gleichungssystem hat entweder keine Lösung, genau eine Lösung, oder unendlich viele Lösungen.
- Nur quadratische oder überbestimmte Gleichungssysteme können eindeutige Lösungen haben.
- Ist ein unterbestimmtes Gleichungssystem lösbar, so ist die Anzahl der Freiheitsgrade mindestens $n - m$.

4.4.2 Homogene und inhomogene Gleichungssysteme

Homogene Gleichungssysteme der Form

$$A \cdot \vec{x} = \vec{0}$$

haben immer mindestens eine Lösung, nämlich $\vec{x} = \vec{0}$. Daher können nur der zweite oder der dritte Fall eintreten (genau eine oder unendlich viele Lösungen). Die Lösungsmenge bildet einen Vektorraum, der genau der Kern der zu A gehörigen linearen Abbildung ist. Da das Bild dieser linearen Abbildung wie schon erwähnt der Spaltenraum ist, folgt aus dem Dimensionssatz, dass die Lösungsmenge Dimension $n - \operatorname{rang} A$ hat.

Betrachtet man zwei Lösungen $\vec{x}_1$ und $\vec{x}_2$ eines inhomogenen Gleichungssystems

$$A \cdot \vec{x} = \vec{b},$$

so gilt

$$A \cdot (\vec{x}_1 - \vec{x}_2) = A \cdot \vec{x}_1 - A \cdot \vec{x}_2 = \vec{b} - \vec{b} = \vec{0}.$$

Die Differenz zweier Lösungen eines inhomogenen Gleichungssystems ist also eine Lösung des homogenen Gleichungssystems. Addiert man umgekehrt zu einer Lösung des inhomogenen Systems eine Lösung des homogenen Systems, so ergibt sich eine weitere Lösung des inhomogenen Systems (wie man durch eine analoge Rechnung begründet). Diesen Zusammenhang kann man so ausdrücken: ist $\vec{x}_0$ eine spezielle Lösung des inhomogenen Systems, so haben alle Lösungen die Form

$$\vec{x}_0 + \vec{y},$$

wobei $\vec{y}$ eine Lösung des homogenen Systems ist. In Worten:

allgemeine Lösung = spezielle Lösung+allgemeine Lösung des homogenen Gleichungssystems

Die Lösung eines inhomogenen linearen Gleichungssystems (falls eine existiert) ist also genau dann eindeutig, wenn das zugehörige homogene Gleichungssystem nur die Lösung $\vec{0}$ hat.
Für quadratische Systeme ergibt sich insbesondere der folgende Satz.

Satz 4.15. *Es sei A eine $n \times n$-Matrix. Es tritt genau einer der folgenden Fälle ein:*

- *Das Gleichungssystem $A \cdot \vec{x} = \vec{0}$ hat nichttriviale Lösungen $\vec{x} \neq \vec{0}$. In diesem Fall hat ein Gleichungssystem der Form $A \cdot \vec{x} = \vec{b}$ entweder keine oder unendlich viele Lösungen.*
- *Jedes Gleichungssystem der Form $A \cdot \vec{x} = \vec{b}$ hat genau eine Lösung.*

Beweis. Der erste Fall tritt genau dann auf, wenn die Lösungsmenge Dimension > 0 hat. Wir wissen bereits, dass die Dimension $n - \operatorname{rang} A$ ist, also gilt dies genau für $\operatorname{rang} A < n$. Ist andererseits $\operatorname{rang} A = n$, dann auch $\operatorname{rang}(A, \vec{b}) = \operatorname{rang} A$ für alle $\vec{b}$. Der Rang von $(A, \vec{b})$ kann nämlich nicht kleiner als der von A sein, aber auch nicht größer als die Anzahl n der Zeilen; da beide gleich sind, folgt $\operatorname{rang}(A, \vec{b}) = \operatorname{rang} A = n$. Damit sind wir aber für jedes $\vec{b}$ in dem Fall, in dem es eine eindeutige Lösung gibt.

4.5 Inverse und Determinante

In Arbeit.

4.6 Orthogonalsysteme

In Arbeit.

4.7 Eigenwerte

In Arbeit.

Kapitel 5

Differentialrechnung

5.1 Ableitungen

Der Begriff der Ableitung ergibt sich aus dem Problem, in einem gewissen Punkt a die Tangente an den Graphen einer Funktion $f(x)$ zu bestimmen. Deren Steigung entspricht der Änderungsrate des Funktionswerts in diesem Punkt. Wenn beispielsweise $f(x)$ eine Bewegung repräsentiert, dann handelt es sich um die aktuelle Geschwindigkeit.

Wenn man einen Wert x in der Nähe von a betrachtet, dann ist die relative Veränderung des Funktionswerts zwischen a und x der sogenannte *Differenzenquotient*

$$\frac{\Delta y}{\Delta x} = \frac{f(x) - f(a)}{x - a},$$

siehe Abbildung 5.1. Wenn sich nun x an a annähert, dann nähert sich die Verbindungsgerade zwischen den Punkten $(a, f(a))$ und $(x, f(x))$ an die Tangente bei a an, und man erhält den Grenzwert, der die *Ableitung* von f in a genannt wird.

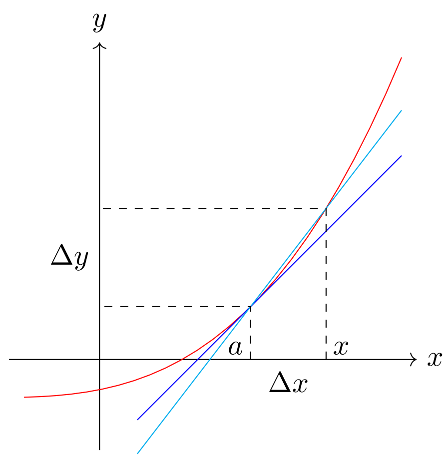

Abbildung 5.1: Zur Definition der Ableitung.

Definition 5.1. Es sei f eine reelle Funktion auf einem Intervall I, und a ein innerer Punkt von I. Die Ableitung von f in a ist der Grenzwert

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

Der Grenzwert in der Definition von $f'(a)$ kann auch in einer anderen Form ausgedrückt werden, die für die Berechnung oft günstig ist. Wenn wir nämlich für den Nenner $x-a$ eine neue Variable h substituieren, dann ist $x \to a$ äquivalent zu $h \to 0$, und wir erhalten

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

Die Tangente im Punkt $(a, f(a))$ ist jene Gerade, die Steigung $f'(a)$ hat und durch den Punkt $(a, f(a))$ geht. Diese hat die Gleichung

$$y - f(a) = f'(a)(x-a).$$

Beispiel 5.2. Wir betrachten die Funktion $f(x) = x^2 + x$ im Punkt $a = 2$. Dort ist $f(2) = 6$, und nach Definition

$$f'(2) = \lim_{x\to 2} \frac{f(x)-f(2)}{x-2} = \lim_{x\to 2} \frac{x^2+x-6}{x-2} = \lim_{x\to 2} \frac{(x-2)(x+3)}{x-2} = \lim_{x\to 2} x+3 = 5.$$

Die Steigung der Tangente in diesem Punkt ist also 5. Die Gleichung der Tangente ist damit

$$y = 5(x-2)+6 = 5x-4.$$

Wenn der Grenzwert in Definition 5.1 existiert, dann wird f *differenzierbar* in a genannt. Ist f in allen Punkten eines Intervalls I differenzierbar, so heißt f differenzierbar auf I. Die Ableitung f' kann dann auch als neue Funktion aufgefasst werden. Ist f' als Funktion wiederum stetig, dann heißt f *stetig differenzierbar*.

Es gibt verschiedene Notationen, die für Ableitungen verwendet werden: ist $y = f(x)$, dann schreibt man für die Ableitung neben f' auch y' oder auch $\frac{dy}{dx}$, $\frac{df}{dx}$ oder $\frac{d}{dx}f(x)$. Die Notation $\frac{dy}{dx}$ ist dabei an den Differenzenquotienten $\frac{\Delta y}{\Delta x} = \frac{f(x)-f(a)}{x-a}$ angelehnt. Man spricht in diesem Zusammenhang vom *Differentialquotienten* $\frac{dy}{dx}$, und vom *Differential* dy.

Nicht alle Funktionen sind überall differenzierbar. Als Beispiel betrachten wir zunächst den Absolutbetrag $f(x) = |x|$ im Punkt $a = 0$. Dann ist

$$\frac{f(x)-f(a)}{x-a} = \frac{|x|}{x} = \begin{cases} 1 & x > 0, \\ -1 & x < 0. \end{cases}$$

Damit existiert der Grenzwert in der Definition von $f'(0)$ nicht: links- und rechtsseitige Grenzwerte existieren, aber sie sind nicht gleich (-1 bzw. 1). Dies stimmt mit der geometrischen Intuition überein: an der Stelle $a = 0$ hat der Graph der Funktion eine Spitze und damit keine Tangente.

Ein anderer Grund, weswegen die Ableitung in einem Punkt nicht existiert, ist eine Unstetigkeit: die Ableitung kann nur in Punkten existieren, wo die Funktion stetig ist.

Satz 5.3. *Wenn die Ableitung $f'(a)$ einer Funktion f im Punkt a existiert, dann ist f in a stetig.*

Beweis. Wir schreiben $f(x)$ als

$$f(x) = \frac{f(x)-f(a)}{x-a} \cdot (x-a) + f(a).$$

Wenn wir nun den Grenzwert bilden, dann folgt

$$\begin{aligned}\lim_{x\to a} f(x) &= \lim_{x\to a}\Big(\frac{f(x)-f(a)}{x-a}\cdot(x-a)+f(a)\Big) = \Big(\lim_{x\to a}\frac{f(x)-f(a)}{x-a}\Big)\cdot\Big(\lim_{x\to a} x-a\Big)+f(a)\\ &= f'(a)\cdot 0 + f(a) = f(a).\end{aligned}$$

Damit ist f im Punkt a stetig.

Wir bestimmen nun zunächst die Ableitungen einiger wichtiger Funktionen.

5.1.1 Ableitungen elementarer Funktionen

Konstante Funktionen

Für eine konstante Funktion $f(x) = c$ ($c \in \mathbb{R}$) ist die Ableitung in jedem Punkt 0:

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

Potenzen

Es sei n eine natürliche Zahl. Wir bestimmen die Ableitung von $f(x) = x^n$. Aus dem binomischen Lehrsatz ergibt sich (da $\binom{n}{0} = 1$ und $\binom{n}{1} = n$)

$$\begin{aligned}\frac{f(x+h)-f(x)}{h} &= \frac{\binom{n}{0}x^n + \binom{n}{1}x^{n-1}h + \binom{n}{2}x^{n-2}h^2 + \cdots + \binom{n}{n}h^n - x^n}{h}\\ &= \frac{\binom{n}{1}x^{n-1}h + \binom{n}{2}x^{n-2}h^2 + \cdots + \binom{n}{n}h^n}{h}\\ &= nx^{n-1} + \binom{n}{2}x^{n-2}h + \cdots + \binom{n}{n}h^{n-1}.\end{aligned}$$

Für $h \to 0$ erhalten wir nun

$$f'(x) = \lim_{h\to 0}\frac{f(x+h)-f(x)}{h} = nx^{n-1}.$$

Später werden wir sehen, dass diese Formel nicht nur für natürliche Zahlen n gilt, sondern für beliebige reelle Zahlen.

Exponentialfunktion

Für die natürliche Exponentialfunktion $f(x) = e^x$ erhalten wir

$$f'(x) = \lim_{h\to 0}\frac{e^{x+h}-e^x}{h} = \lim_{h\to 0} e^x\frac{e^h-1}{h}.$$

Wir benötigen also den Grenzwert von $\frac{e^h-1}{h}$ für $h \to 0$. Dazu erinnern wir uns an die Reihendefinition der Exponentialfunktion:

$$e^h = \sum_{k=0}^{\infty}\frac{h^k}{k!} = 1 + \sum_{k=1}^{\infty}\frac{h^k}{k!} = 1 + h\sum_{k=1}^{\infty}\frac{h^{k-1}}{k!},$$

also

$$\frac{e^h-1}{h}=\sum_{k=1}^{\infty}\frac{h^{k-1}}{k!}=1+\sum_{k=2}^{\infty}\frac{h^{k-1}}{k!}=1+h\sum_{k=2}^{\infty}\frac{h^{k-2}}{k!}.$$

Nun gilt

$$\Big|\sum_{k=2}^{\infty}\frac{h^{k-2}}{k!}\Big|\le\sum_{k=2}^{\infty}\frac{|h|^{k-2}}{k!}\le\sum_{k=2}^{\infty}\frac{|h|^{k-2}}{(k-2)!}=\sum_{\ell=0}^{\infty}\frac{|h|^{\ell}}{\ell!}=e^{|h|}$$

und daher

$$-|h|e^{|h|}\le h\sum_{k=2}^{\infty}\frac{h^{k-2}}{k!}\le|h|e^{|h|}.$$

Aus Satz 3.15 folgt damit

$$\lim_{h\to0}h\sum_{k=2}^{\infty}\frac{h^{k-2}}{k!}=0$$

und schließlich

$$\lim_{h\to0}\frac{e^h-1}{h}=1.$$

Daraus können wir schließen, dass die Ableitung von $f(x)=e^x$ durch $f'(x)=e^x$ gegeben ist. Die natürliche Exponentialfunktion ist also gleich ihrer eigenen Ableitung.

Sinus und Cosinus

Zur Bestimmung der Ableitungen der Winkelfunktionen Sinus und Cosinus verwenden wir zunächst die Additionstheoreme (1.2):

$$\frac{\sin(x+h)-\sin x}{h}=\frac{\sin x\cos h+\cos x\sin h-\sin x}{h}=\sin x\frac{\cos h-1}{h}+\cos x\frac{\sin h}{h}$$

und

$$\frac{\cos(x+h)-\cos x}{h}=\frac{\cos x\cos h-\sin x\sin h-\cos x}{h}=\cos x\frac{\cos h-1}{h}-\sin x\frac{\sin h}{h}.$$

Der Grenzwert

$$\lim_{h\to0}\frac{\sin h}{h}=1$$

ist uns bereits in Beispiel 3.16 begegnet. Weiters können wir das Resultat aus Beispiel 3.23 verwenden, demzufolge $\lim_{h\to0}\frac{\cos h-1}{h^2}=-\frac{1}{2}$ ist. Damit folgt nämlich

$$\lim_{h\to0}\frac{\cos h-1}{h}=\lim_{h\to0}\frac{\cos h-1}{h^2}\cdot\lim_{h\to0}h=-\frac{1}{2}\cdot0=0.$$

Damit folgt

$$\lim_{h\to0}\frac{\sin(x+h)-\sin x}{h}=\cos x$$

sowie

$$\lim_{h\to0}\frac{\cos(x+h)-\cos x}{h}=-\sin x,$$

und wir können schließen dass die Ableitungen der Funktionen $\sin x$ und $\cos x$ gleich $\cos x$ bzw. $-\sin x$ sind.

5.1.2 Differentiationsregeln

Als Nächstes untersuchen wir, wie sich verschiedene Operationen auf die Ableitung übertragen. Die einfachsten Regeln in diesem Zusammenhang betreffen Summen, Differenzen und konstante Vielfache.

Addition und Subtraktion

Die Ableitung einer Summe oder Differenz zweier differenzierbarer Funktionen ist die Summe bzw. Differenz der Ableitungen. Dies sieht man direkt aus der Definition, indem man den Grenzwert aufteilt:

$$\begin{aligned}(f \pm g)'(x) = \lim_{h\to 0} \frac{(f \pm g)(x+h) - (f \pm g)(x)}{h} &= \lim_{h\to 0} \left(\frac{f(x+h) - f(x)}{h} \pm \frac{g(x+h) - g(x)}{h} \right) \\ &= \lim_{h\to 0} \frac{f(x+h) - f(x)}{h} \pm \lim_{h\to 0} \frac{g(x+h) - g(x)}{h} \\ &= f'(x) \pm g'(x).\end{aligned}$$

Konstante Vielfache

Ebenso übertragen sich konstante Faktoren direkt auf die Ableitung: für jede reelle Konstante c gilt

$$\begin{aligned}(cf)'(x) &= \lim_{h\to 0} \frac{(cf)(x+h) - (cf)(x)}{h} = \lim_{h\to 0} \frac{c(f(x+h) - f(x))}{h} \\ &= c \lim_{h\to 0} \frac{f(x+h) - f(x)}{h} = cf'(x).\end{aligned}$$

Beispiel 5.4. Linearkombinationen wie $f(x) = 4e^x + 2x^5 - 3\sin x$ lassen sich nun leicht differenzieren:

$$f'(x) = 4e^x + 10x^4 - 3\cos x.$$

Produkte

Im Allgemeinen sind Produkte etwas komplizierter: insbesondere ist die Ableitung eines Produkts *nicht* gleich dem Produkt der Ableitungen. Stattdessen gilt die sogenannte *Produktregel*: sind $f(x)$ und $g(x)$ in einem Punkt bzw. auf einem Intervall differenzierbar, dann auch $f(x)g(x)$, und es gilt

$$\frac{d}{dx} f(x)g(x) = f'(x)g(x) + f(x)g'(x).$$

Dies ergibt sich durch folgende Rechnung:

$$\begin{aligned}&\lim_{h\to 0} \frac{f(x+h)g(x+h) - f(x)g(x)}{h} \\ &\quad= \lim_{h\to 0} \frac{f(x+h)g(x+h) - f(x)g(x+h) + f(x)g(x+h) - f(x)g(x)}{h} \\ &\quad= \lim_{h\to 0} \frac{f(x+h) - f(x)}{h} g(x+h) + \lim_{h\to 0} f(x) \frac{g(x+h) - g(x)}{h} \\ &\quad= f'(x)g(x) + f(x)g'(x).\end{aligned}$$

Hier ist zu beachten, dass f und g stetig sein müssen (da sie differenzierbar sind), weswegen $\lim_{h\to 0} g(x+h) = g(x)$ gilt.

Beispiel 5.5. Die Ableitung von $f(x) = e^x \cos x + 3x \sin x$ ergibt sich mit der Produktregel wie folgt:

$$f'(x) = e^x \cos x + e^x(-\sin x) + 3\sin x + 3x\cos x = e^x(\cos x - \sin x) + 3(\sin x + x\cos x).$$

Quotienten

Die sogenannte *Quotientenregel* für die Ableitung von Quotienten ähnelt der Produktregel: sind $f(x)$ und $g(x)$ in einem Punkt bzw. auf einem Intervall differenzierbar, dann auch $\frac{f(x)}{g(x)}$ überall wo $g(x) \neq 0$, und es gilt

$$\frac{d}{dx}\frac{f(x)}{g(x)} = \frac{f'(x)g(x) - f(x)g'(x)}{g(x)^2}.$$

Dies wird auf ähnliche Weise wie die Produktregel bewiesen:

$$\begin{aligned}
\lim_{h\to 0} \frac{\frac{f(x+h)}{g(x+h)} - \frac{f(x)}{g(x)}}{h} &= \lim_{h\to 0} \frac{f(x+h)g(x) - f(x)g(x+h)}{hg(x)g(x+h)} \\
&= \lim_{h\to 0} \frac{f(x+h)g(x) - f(x)g(x) + f(x)g(x) - f(x)g(x+h)}{hg(x)g(x+h)} \\
&= \lim_{h\to 0} \frac{\frac{(f(x+h)-f(x))g(x)}{h} - \frac{f(x)(g(x+h)-g(x))}{h}}{g(x)g(x+h)} \\
&= \frac{f'(x)g(x) - f(x)g'(x)}{g(x)^2}.
\end{aligned}$$

Beispiel 5.6. Die Ableitung von $f(x) = \frac{\cos x}{e^x+3}$ ergibt sich mit der Quotientenregel wie folgt:

$$f'(x) = \frac{-\sin x \cdot (e^x+3) - \cos x \cdot e^x}{(e^x+3)^2} = -\frac{e^x(\sin x + \cos x) + 3\sin x}{(e^x+3)^2}.$$

Mit Hilfe der Quotientenregel können wir nunmehr auch Tangens und Cotangens differenzieren: es gilt

$$\frac{d}{dx}\tan x = \frac{d}{dx}\frac{\sin x}{\cos x} = \frac{\cos x \cdot \cos x - \sin x \cdot (-\sin x)}{\cos^2 x} = \frac{\cos^2 x + \sin^2 x}{\cos^2 x} = \frac{1}{\cos^2 x}$$

und

$$\frac{d}{dx}\cot x = \frac{d}{dx}\frac{\cos x}{\sin x} = \frac{-\sin x \cdot \sin x - \cos x \cdot \cos x}{\sin^2 x} = -\frac{\sin^2 x + \cos^2 x}{\sin^2 x} = -\frac{1}{\sin^2 x}.$$

Zusammengesetzte Funktionen: die Kettenregel

Die *Kettenregel* ermöglicht es, zusammengesetzte Funktionen der Form $f(g(x))$ zu differenzieren. Es seien f und g Funktionen und a ein Punkt, für den g in a und f in $g(a)$ differenzierbar ist. Dann gilt[1]

$$\begin{aligned}
\lim_{x\to a}\frac{f(g(x))-f(g(a))}{x-a} &= \lim_{x\to a}\frac{f(g(x))-f(g(a))}{g(x)-g(a)}\cdot\frac{g(x)-g(a)}{x-a}\\
&= \lim_{x\to a}\frac{f(g(x))-f(g(a))}{g(x)-g(a)}\cdot\lim_{x\to a}\frac{g(x)-g(a)}{x-a}\\
&= f'(g(a))\cdot g'(a).
\end{aligned}$$

Ist g auf einem ganzen Intervall I differenzierbar, und f auf $g(I)$ differenzierbar, dann ist $f\circ g$ auf I differenzierbar, und es gilt

$$\frac{d}{dx}f(g(x)) = f'(g(x))g'(x).$$

Beispiel 5.7. Für die Funktion $h(x)=\sin(x^3)$, die aus $f(x)=\sin x$ und $g(x)=x^3$ zusammengesetzt ist, gilt

$$h'(x)=\cos(x^3)\cdot\frac{d}{dx}x^3=\cos(x^3)\cdot 3x^2.$$

Mitunter kann es nötig sein, die Kettenregel mehrfach anzuwenden, weil mehrere Funktionen ineinander verschachtelt sind.

Beispiel 5.8. Betrachte die Funktion $g(x)=e^{\cos((3x+1)^2)}$. Diese kann als Verknüpfung $g=f_1\circ f_2\circ f_3\circ f_4$ von vier Funktionen betrachtet werden: $f_1(x)=e^x$, $f_2(x)=\cos x$, $f_3(x)=x^2$ und $f_4(x)=3x+1$. Damit ergibt sich durch wiederholtes Anwenden der Kettenregel

$$\begin{aligned}
\frac{d}{dx}e^{\cos((3x+1)^2)} &= e^{\cos((3x+1)^2)}\cdot\frac{d}{dx}\cos((3x+1)^2)\\
&= e^{\cos((3x+1)^2)}\cdot\big(-\sin((3x+1)^2)\big)\cdot\frac{d}{dx}(3x+1)^2\\
&= e^{\cos((3x+1)^2)}\cdot\big(-\sin((3x+1)^2)\big)\cdot 2(3x+1)\cdot\frac{d}{dx}(3x+1)\\
&= e^{\cos((3x+1)^2)}\cdot\big(-\sin((3x+1)^2)\big)\cdot 2(3x+1)\cdot 3\\
&= -6(3x+1)e^{\cos((3x+1)^2)}\sin((3x+1)^2).
\end{aligned}$$

Mit Hilfe der Kettenregel können wir nun auch weitere elementare Funktionen differenzieren. So können wir etwa a^x als $e^{x\ln a}$ schreiben, woraufhin die Kettenregel

$$\frac{d}{dx}a^x = e^{x\ln a}\cdot\ln a = a^x\ln a$$

liefert. Aus der Kettenregel folgt außerdem

$$\frac{d}{dx}e^{-x}=-e^{-x},$$

womit sich in weiterer Folge die Ableitungen der hyperbolischen Funktionen ergeben:

$$\frac{d}{dx}\cosh x=\frac{d}{dx}\frac{e^x+e^{-x}}{2}=\frac{e^x-e^{-x}}{2}=\sinh x,$$

und auf analoge Weise $\frac{d}{dx}\sinh x=\cosh x$.

[1]Etwas Vorsicht ist geboten, weil der Nenner $g(x)-g(a)$ auch für $x\neq a$ gleich 0 sein kann. Wenn es eine Folge x_n von Punkten mit $x_n\neq a$, $g(x_n)=g(a)$ und $x_n\to a$ gibt, dann muss $g'(a)$ gleich 0 sein, ebenso wie die Ableitung von $f(g(x))$ an der Stelle a. Die Details seien hier ausgespart.

Ableitung inverser Funktionen

Um Funktionen wie den Logarithmus differenzieren zu können, betrachten wir als Nächstes inverse Funktionen. Angenommen, f ist auf einem Intervall I differenzierbar und bijektiv mit Umkehrfunktion f^{-1} auf $J = f(I)$. Wir wollen nun für einen Punkt $b = f(a)$ im Inneren von J die Ableitung von f^{-1} bestimmen.

Aus der Definition erhalten wir

$$(f^{-1})'(b) = \lim_{y \to b} \frac{f^{-1}(y) - f^{-1}(b)}{y - b}.$$

Wenn wir nun $y = f(x)$ und $b = f(a)$ substituieren, dann gilt mit $y \to b$ gleichzeitig auch $x = f^{-1}(y) \to a = f^{-1}(b)$ (f und f^{-1} sind Umkehrfunktionen voneinander; wenn eine von ihnen differenzierbar und damit stetig ist, dann auch die andere). Damit erhalten wir

$$(f^{-1})'(b) = \lim_{x \to a} \frac{x - a}{f(x) - f(a)} = \frac{1}{f'(a)} = \frac{1}{f'(f^{-1})(b)}.$$

Zusammenfassend können wir also sagen:

Satz 5.9. *Ist f auf einem Intervall I bijektiv und differenzierbar, dann ist f^{-1} in jedem Punkt $b = f(a)$ differenzierbar, für den $f'(a) \neq 0$. In diesem Fall gilt*

$$(f^{-1})'(b) = \frac{1}{f'(a)} = \frac{1}{f'(f^{-1}(b))}.$$

Als wichtige Spezialfälle erhalten wir nun die Ableitungen des Logarithmus sowie der inversen trigonometrischen und hyperbolischen Funktionen.

- Für $f(x) = e^x$ ist $f'(x) = e^x \neq 0$ für alle $x \in \mathbb{R}$. Daher ist die Umkehrfunktion, nämlich der natürliche Logarithmus ln, auf dem gesamten Definitionsbereich differenzierbar, und es gilt
$$\frac{d}{dx} \ln x = \frac{1}{e^{\ln x}} = \frac{1}{x}.$$
Für den allgemeinen Logarithmus $\log_a$ schreiben wir
$$\log_a x = \frac{\ln x}{\ln a} = \frac{1}{\ln a} \cdot \ln x$$
(siehe Kapitel 3.7), und damit
$$\frac{d}{dx} \log_a x = \frac{1}{\ln a} \cdot \frac{1}{x} = \frac{1}{x \ln a}.$$

- Für $f(x) = \sin x$ ist $f'(x) = \cos x \neq 0$ für alle $x \in (-\frac{\pi}{2}, \frac{\pi}{2})$. Damit ist die Umkehrfunktion arcsin im Inneren ihres Definitionsbereichs[2] differenzierbar: es gilt
$$\frac{d}{dx} \arcsin x = \frac{1}{\cos(\arcsin x)}.$$

[2] Aber nicht an den Rändern: dort ist der Nenner 0, wie sich zeigt.

Aus dem Satz von Pythagoras erhalten wir

$$\cos^2(\arcsin x) + \sin^2(\arcsin x) = \cos^2(\arcsin x) + x^2 = 1.$$

Daher ist $\cos(\arcsin x) = \sqrt{1-x^2}$ (das Vorzeichen muss positiv sein, denn $\arcsin x$ liegt nach Definition im Intervall $[-\frac{\pi}{2}, \frac{\pi}{2}]$, wo der Cosinus ≥ 0 ist) und damit

$$\frac{d}{dx} \arcsin x = \frac{1}{\sqrt{1-x^2}}, \qquad x \in (-1, 1).$$

Für $f(x) = \cos x$ ergibt sich in gleicher Weise, dass

$$\frac{d}{dx} \arccos x = -\frac{1}{\sqrt{1-x^2}}, \qquad x \in (-1, 1).$$

- Für $f(x) = \tan x$ ist $f'(x) = \frac{1}{\cos^2 x} \neq 0$ für alle $x \in (-\frac{\pi}{2}, \frac{\pi}{2})$. Daher ist die Umkehrfunktion arctan überall differenzierbar. Aus der Identität $\sin^2 x + \cos^2 x = 1$ folgt zudem

  $$\frac{1}{\cos^2 x} = \frac{\sin^2 x + \cos^2 x}{\cos^2 x} = \frac{\sin^2 x}{\cos^2 x} + 1 = \tan^2 x + 1,$$

  daher

  $$\frac{d}{dx} \arctan x = \frac{1}{\cos^2(\arctan x)} = \frac{1}{1 + \tan^2(\arctan x)} = \frac{1}{1+x^2}.$$

  Analog dazu erhält man auch

  $$\frac{d}{dx} \operatorname{arccot} x = -\frac{1}{1+x^2}.$$

Die Ableitung des natürlichen Logarithmus kann verwendet werden, um die Ableitungsregel für Potenzfunktionen x^a auf beliebigen reellen Exponenten a auszudehnen. So wie bei der Exponentialfunktion a^x schreiben wir zunächst um: $x^a = e^{a \ln x}$. Damit folgt nach der Kettenregel

$$\frac{d}{dx} x^a = e^{a \ln x} \cdot \frac{a}{x} = x^a \cdot \frac{a}{x} = a x^{a-1},$$

nunmehr für beliebiges reelles a, und nicht nur für natürliche Zahlen. Auf diese Art kann man auch beliebige Funktionen der Form $f(x)^{g(x)}$ differenzieren, bei denen sowohl Basis als auch Exponent von x abhängig sind. So ist zum Beispiel

$$\frac{d}{dx} x^x = \frac{d}{dx} e^{x \ln x} = e^{x \ln x}\Big(x \cdot \frac{1}{x} + 1 \cdot \ln x\Big) = x^x(1 + \ln x).$$

Als Zusammenfassung dieses Abschnitts sind in Tabelle 5.1 die Ableitungen wichtiger Funktionen aufgelistet.

5.1.3 Höhere Ableitungen

Wenn die Ableitung $f'(x)$ einer Funktion $f(x)$ wieder differenzierbar ist, dann ist deren Ableitung die *zweite Ableitung* $f''(x)$ von $f(x)$. Die Ableitung davon ist die *dritte Ableitung* $f'''(x)$, und so fort. Allgemein erhält man die n-te Ableitung einer Funktion durch n-faches Differenzieren. Die übliche Schreibweise für die n-te Ableitung einer Funktion $y = f(x)$ ist[3]

$$\left(\frac{d}{dx}\right)^n f(x) = \frac{d^n f}{dx^n} = f^{(n)}(x) = y^{(n)}.$$

[3] Bis zur dritten Ableitung sind f' („f Strich"), f'' („f zwei Strich"), f''' („f drei Strich") üblich. Danach setzt man mit $f^{(4)}$, $f^{(5)}$, … oder auch mit f^{IV}, f^{V}, … fort.

Funktion $f(x)$	Ableitung $f'(x)$	Funktion $f(x)$	Ableitung $f'(x)$
C (konstant)	0	$\sinh x$	$\cosh x$
x^a	ax^{a-1}	$\cosh x$	$\sinh x$
e^x	e^x	$\tanh x$	$\frac{1}{\cosh^2 x}$
a^x	$a^x \ln a$	$\coth x$	$-\frac{1}{\sinh^2 x}$
$\ln x$	$\frac{1}{x}$	$\arcsin x$	$\frac{1}{\sqrt{1-x^2}}$
$\log_a x$	$\frac{1}{x \ln a}$	$\arccos x$	$-\frac{1}{\sqrt{1-x^2}}$
$\sin x$	$\cos x$	$\arctan x$	$\frac{1}{1+x^2}$
$\cos x$	$-\sin x$	$\operatorname{arsinh} x$	$\frac{1}{\sqrt{x^2+1}}$
$\tan x$	$\frac{1}{\cos^2 x}$	$\operatorname{arcosh} x$	$\frac{1}{\sqrt{x^2-1}}$
$\cot x$	$-\frac{1}{\sin^2 x}$	$\operatorname{artanh} x$	$\frac{1}{1-x^2}$

Tabelle 5.1: Ableitungen wichtiger Funktionen.

Beispiel 5.10. Für die Funktion $f(x) = x^2 \cos x$ sind die erste, zweite, dritte und vierte Ableitung, die man durch wiederholtes Anwenden der Produktregel erhält,

$$\begin{aligned}
f'(x) &= 2x\cos x - x^2 \sin x,\\
f''(x) &= 2\cos x - 2x\sin x - 2x\sin x - x^2\cos x = (2-x^2)\cos x - 4x\sin x,\\
f'''(x) &= -2x\cos x - (2-x^2)\sin x - 4\sin x - 4x\cos x = (x^2-6)\sin x - 6x\cos x,\\
f^{(4)}(x) &= 2x\sin x + (x^2-6)\cos x - 6\cos x + 6x\sin x = (x^2-12)\cos x + 8x\sin x.
\end{aligned}$$

5.2 Ableitungen und Eigenschaften von Funktionen

In diesem Abschnitt untersuchen wir, wie sich Eigenschaften einer Funktion in den Ableitungen widerspiegeln.

5.2.1 Stationäre Punkte

Ein Punkt a wird *stationärer Punkt* einer Funktion genannt, wenn dort die Ableitung 0 ist: $f'(a) = 0$. In diesem Fall ist die Tangente in dem Punkt waagrecht. Stationäre Punkte spielen eine wichtige Rolle bei der Bestimmung der Extremwerte einer Funktion.

Definition 5.11. Es sei $f : D \to \mathbb{R}$ eine reelle Funktion, und es sei $a \in \mathbb{R}$ derart, dass für ein $\delta > 0$ das Intervall $(a-\delta, a+\delta)$ in D enthalten ist und für alle $x \in (a-\delta, a+\delta)$ gilt, dass $f(x) \leq f(a)$. Dann sagen wir, dass f in a ein lokales Maximum hat. Gilt umgekehrt $f(x) \geq f(a)$ für alle $x \in (a-\delta, a+\delta)$, dann sagen wir, dass f in a ein lokales Minimum hat.

Satz 5.12. *Hat f im Punkt a ein lokales Maximum oder Minimum, und ist f dort differenzierbar, dann muss a ein stationärer Punkt sein.*

Beweis. Wir führen den Beweis für das Maximum; für das Minimum funktioniert er analog. Aus der Bedingung folgt, dass für $x \in (a-\delta, a+\delta)$ der Zähler des Quotienten

$$\frac{f(x)-f(a)}{x-a}$$

immer ≤ 0 ist. Der Nenner dagegen ist positiv für $x > a$ und negativ für $x < a$. Damit folgt

$$\lim_{x \to a^+} \frac{f(x)-f(a)}{x-a} \leq 0$$

und gleichzeitig

$$\lim_{x \to a^-} \frac{f(x)-f(a)}{x-a} \geq 0$$

(sofern die Grenzwerte existieren). Wenn also die Ableitung

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

existiert, dann muss sie 0 sein.

Zu beachten ist, dass die Umkehrung nicht gilt: nicht jeder stationäre Punkt ist zwangsläufig ein lokales Minimum oder Maximum. Beispielsweise hat die Funktion $f(x) = x^3$ einen stationären Punkt bei 0, der aber weder ein Maximum noch ein Minimum ist. Etwas später werden wir sehen, wie man feststellt ob in einem gegebenen stationären Punkt ein Minimum, ein Maximum oder keines von beiden vorliegt.

5.2.2 Der Mittelwertsatz

Wenn wir eine Funktion f über ein Intervall $[a, b]$ betrachten, dann stellt der Quotient

$$\frac{f(b)-f(a)}{b-a} \tag{5.1}$$

die relative Änderung über das Intervall hinweg dar. Der Mittelwertsatz sagt nun: ist f differenzierbar, dann gibt es einen Punkt wo die lokale Änderung, also die Ableitung, der Funktion gleich diesem Quotienten ist.

Satz 5.13 (Mittelwertsatz)**.** *Ist die Funktion f auf dem Intervall $[a, b]$ stetig und auf (a, b) differenzierbar, dann gibt es ein $c \in (a, b)$ sodass*

$$f'(c) = \frac{f(b)-f(a)}{b-a}.$$

Wenn wir f etwa als eine Bewegung betrachten, dann stellt $f'(c)$ die Geschwindigkeit zum Zeitpunkt c dar, und der Quotient (5.1) ist die durchschnittliche Geschwindigkeit über das gesamte Zeitintervall $[a, b]$. Die Aussage des Mittelwertsatzes ist dann also: es gibt einen Zeitpunkt zu dem die Geschwindigkeit exakt gleich der Durchschnittsgeschwindigkeit ist.

Zum Beweis brauchen wir zunächst einen Spezialfall:

Satz 5.14 (Satz von Rolle). *Ist die Funktion f auf dem Intervall $[a, b]$ stetig und auf (a, b) differenzierbar, und ist zudem $f(b) = f(a)$, dann gibt es mindestens einen stationären Punkt, also ein $c \in (a, b)$ sodass*

$$f'(c) = 0.$$

Beweis. Wir wissen (Satz 3.30), dass f aufgrund der Stetigkeit ein Maximum und ein Minimum haben muss. Sind beide gleich $f(a) = f(b)$, dann ist die Funktion auf dem ganzen Intervall konstant, und die Ableitung ist überall 0. In diesem Fall gilt der Satz offensichtlich. Andernfalls wird entweder das Maximum oder das Minimum in einem inneren Punkt $c \in (a, b)$ angenommen, und wir wissen (nach Satz 5.12), dass dann c ein stationärer Punkt ist. Damit ist der Satz von Rolle bewiesen.

Der allgemeinere Mittelwertsatz kann nun auf den Satz von Rolle zurückgeführt werden: erfüllt die Funktion f die Bedingungen des Mittelwertsatzes, dann gilt für die Funktion

$$g(x) = f(x) - \frac{f(b) - f(a)}{b - a} \cdot x,$$

dass sie alle Bedingungen des Satzes von Rolle erfüllt: sie ist stetig auf $[a, b]$, differenzierbar auf (a, b), und es gilt

$$\begin{aligned} g(b) - g(a) &= \left(f(b) - \frac{f(b) - f(a)}{b - a} \cdot b\right) - \left(f(a) - \frac{f(b) - f(a)}{b - a} \cdot a\right) \\ &= f(b) - f(a) - \frac{f(b) - f(a)}{b - a} \cdot (b - a) = (f(b) - f(a)) - (f(b) - f(a)) = 0. \end{aligned}$$

Damit gibt es ein $c \in (a, b)$ sodass

$$g'(c) = f'(c) - \frac{f(b) - f(a)}{b - a} = 0,$$

also

$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$

Damit ist auch der Mittelwertsatz bewiesen.

5.2.3 Monotonie

Aus dem Mittelwertsatz lässt sich schließen, wie das Monotonieverhalten einer Funktion mit der Ableitung zusammenhängt. Es gilt nämlich der folgende Satz.

Satz 5.15. *Die reelle Funktion f sei auf dem Intervall (a, b) differenzierbar. Es gilt:*

- *Ist $f'(x) > 0$ für alle $x \in (a, b)$, dann ist f auf (a, b) streng monoton steigend.*
- *Ist $f'(x) \geq 0$ für alle $x \in (a, b)$, dann ist f auf (a, b) monoton steigend.*
- *Ist $f'(x) < 0$ für alle $x \in (a, b)$, dann ist f auf (a, b) streng monoton fallend.*
- *Ist $f'(x) \leq 0$ für alle $x \in (a, b)$, dann ist f auf (a, b) monoton fallend.*

Ist f zudem stetig in a und b (also auf ganz $[a, b]$), dann gilt die Monotonie auch auf $[a, b]$.

Beweis. Wir behandeln nur den ersten der vier Fälle, die weiteren drei ergeben sich in gleicher Weise. Es sei also die Ableitung von f auf ganz (a, b) positiv. Wir betrachten nun zwei beliebige Punkte x_1 und x_2 im Intervall (a, b), $x_1 < x_2$. Nach dem Mittelwertsatz gilt

$$f(x_2) - f(x_1) = (x_2 - x_1)f'(c)$$

für ein $c \in (a, b)$. Nach Voraussetzung ist $f'(c) > 0$, also $f(x_2) - f(x_1) > 0$ bzw. $f(x_2) > f(x_1)$. Damit ist f streng monoton steigend auf (a, b). Ist f zudem stetig auf $[a, b]$, dann gilt derselbe Schluss auch für $x_1 = a$ bzw. $x_2 = b$, und f ist damit streng monoton steigend auf $[a, b]$.

Beispiel 5.16. Man bestimme die Intervalle, auf denen die Funktion $f(x) = \frac{10x}{x^2+4}$ monoton steigend bzw. fallend ist.

Lösung. Die Ableitung ist nach der Quotientenregel

$$f'(x) = \frac{10(x^2+4) - 10x \cdot 2x}{(x^2+4)^2} = \frac{40 - 10x^2}{(x^2+4)^2} = \frac{10(4-x^2)}{(x^2+4)^2}.$$

Dies ist positiv, falls $x^2 < 4$, also für $x \in (-2, 2)$. Für $x > 2$ oder $x < -2$ ist die Ableitung dagegen negativ. Da f in den Punkten 2 und -2 zudem stetig ist, folgt: f ist auf $[-2, 2]$ streng monoton steigend, auf $(-\infty, -2]$ und $[2, \infty)$ dagegen streng monoton fallend, siehe Abbildung 5.2.

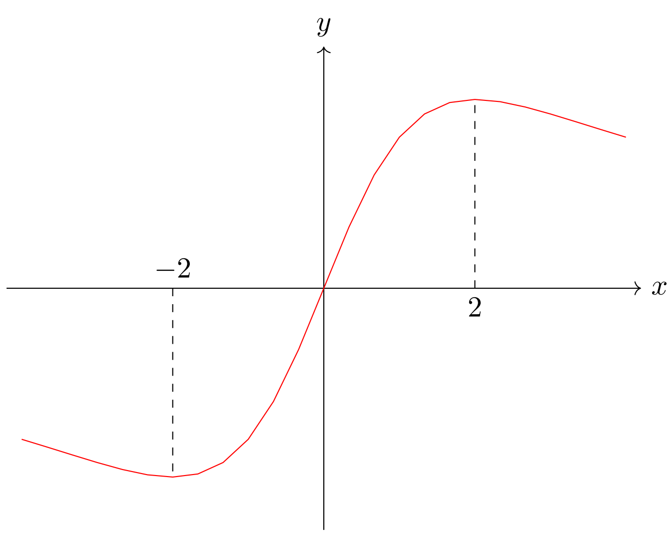

Abbildung 5.2: Die Funktion in Beispiel 5.16.

5.2.4 Konvexität und Konkavität

So wie die erste Ableitung einer Funktion das Monotonieverhalten bestimmt, hängt die zweite Ableitung mit dem Krümmungsverhalten zusammen. Wir definieren zunächst *Konvexität* und *Konkavität* von Funktionen formal.

Definition 5.17. Eine Funktion f heißt konvex auf einem Intervall I, wenn für je zwei Zahlen $x_1, x_2 \in I$ gilt, dass die Strecke zwischen den Punkten $(x_1, f(x_1))$ und $(x_2, f(x_2))$ über dem Funktionsgraphen liegt, wenn also für jedes $\lambda \in (0, 1)$ gilt, dass

$$(1-\lambda)f(x_1) + \lambda f(x_2) \geq f\big((1-\lambda)x_1 + \lambda x_2\big).$$

Analog dazu heißt f konkav auf I, wenn für je zwei Zahlen $x_1, x_2 \in I$ die umgekehrte Ungleichung gilt:

$$(1-\lambda)f(x_1) + \lambda f(x_2) \le f\big((1-\lambda)x_1 + \lambda x_2\big).$$

Gilt sogar strikte Ungleichung (also $>$ bzw. $<$), so spricht man von einer streng konvexen bzw. konkaven Funktion.

Abbildung 5.3 illustriert diese Definition. Eine (streng) konvexe Funktion ist also eine „nach oben gekrümmte“ Funktion, eine konkave Funktion eine „nach unten gekrümmte“. Man beachte, dass Funktionen der Form $f(x) = ax+b$, deren Funktionsgraph eine Gerade ist, gleichzeitig konvex und konkav sind.

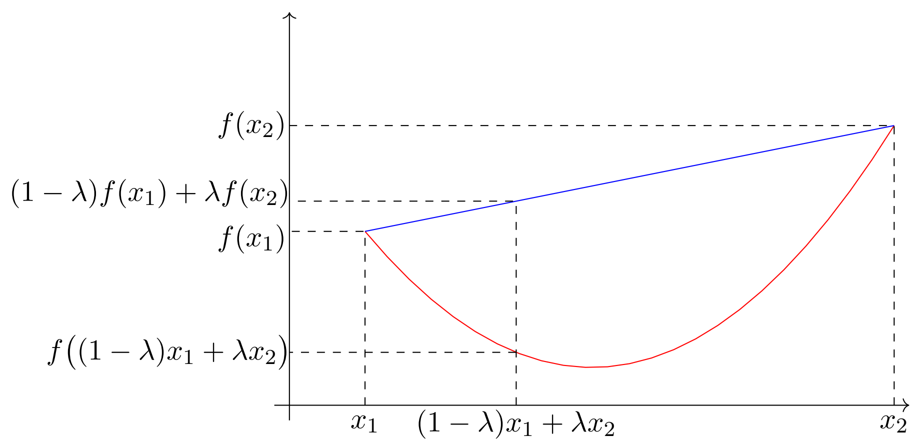

Abbildung 5.3: Zur Definition von Konvexität.

Es gilt nun eine ähnliche Aussage wie Satz 5.15.

Satz 5.18. *Die reelle Funktion f sei auf dem Intervall (a,b) zweimal differenzierbar. Es gilt:*

- *Ist $f''(x) > 0$ für alle $x \in (a,b)$, dann ist f auf (a,b) streng konvex.*
- *Ist $f''(x) \ge 0$ für alle $x \in (a,b)$, dann ist f auf (a,b) konvex.*
- *Ist $f''(x) < 0$ für alle $x \in (a,b)$, dann ist f auf (a,b) streng konkav.*
- *Ist $f''(x) \le 0$ für alle $x \in (a,b)$, dann ist f auf (a,b) konkav.*

Ist f zudem stetig in a und b (also auf ganz $[a,b]$), dann gilt die Aussage auch auf $[a,b]$.

Beweis. Wir führen den Beweis wieder nur im ersten Fall. Dazu betrachten wir zwei Punkte x_1, x_2 $(x_1 < x_2)$ im Intervall (a,b) sowie einen Punkt $x^* = (1-\lambda)x_1 + \lambda x_2$ zwischen den beiden. Nach dem Mittelwertsatz gibt es $c_1 \in (x_1, x^*)$ und $c_2 \in (x^*, x_2)$ derart, dass

$$f(x^*) - f(x_1) = (x^* - x_1)f'(c_1) \qquad \text{und} \qquad f(x_2) - f(x^*) = (x_2 - x^*)f'(c_2).$$

Da nach Voraussetzung $f''(x) > 0$ auf dem ganzen Intervall gilt, ist f' nach Satz 5.15 streng monoton steigend. Wegen $c_1 < x^* < c_2$ gilt daher $f'(c_1) < f'(c_2)$, also

$$\frac{f(x^*) - f(x_1)}{x^* - x_1} < \frac{f(x_2) - f(x^*)}{x_2 - x^*}.$$

Durch Multiplikation mit dem gemeinsamen Nenner $(x^* - x_1)(x_2 - x^*)$ und Umformen wird daraus

$$f(x_1)(x_2 - x^*) + f(x_2)(x^* - x_1) > f(x^*)(x_2 - x_1).$$

Nun ist $x_2 - x^* = (1-\lambda)(x_2 - x_1)$ und $x^* - x_1 = \lambda(x_2 - x_1)$. Wir können also $x_2 - x_1$ kürzen und erhalten

$$(1-\lambda)f(x_1) + \lambda f(x_2) > f(x^*),$$

was genau die gewünschte Ungleichung ist. So wie im Beweis von Satz 5.15 kann man die Aussage auf das abgeschlossene Intervall $[a, b]$ ausdehnen, wenn f in a und b stetig ist.

Bemerkung 5.19. So wie eine Funktion monoton sein kann ohne dass sie differenzierbar (oder auch nur stetig) ist, kann eine Funktion auch konvex sein ohne dass die Ableitung überall existiert. Beispielsweise ist der Absolutbetrag $f(x) = |x|$ eine konvexe Funktion.

Beispiel 5.20. Man bestimme die Intervalle, auf denen die Funktion $f(x) = \sqrt{x} \cdot \frac{x^2+15}{10}$ konvex bzw. konkav ist.

Lösung. Die Funktion ist ausmultipliziert $f(x) = \frac{1}{10}(x^2\sqrt{x} + 15\sqrt{x}) = \frac{1}{10}(x^{5/2} + 15x^{1/2})$. Damit ergibt sich

$$f'(x) = \frac{1}{10}\Big(\frac{5}{2}x^{3/2} + \frac{15}{2}x^{-1/2}\Big) = \frac{1}{4}x^{3/2} + \frac{3}{4}x^{-1/2}$$

und

$$f''(x) = \frac{3}{8}x^{1/2} - \frac{3}{8}x^{-3/2} = \frac{3}{8}x^{-3/2}(x^2 - 1).$$

Man beachte, dass die Funktion nur auf $[0, \infty)$ definiert ist. Wir sehen anhand der Faktorisierung von $f''(x)$, dass $f''(x) > 0$ für $x > 1$ und $f''(x) < 0$ für $x < 1$ gilt. Also ist die Funktion auf $[0, 1]$ (streng) konkav, auf $[1, \infty)$ dagegen (streng) konvex. Siehe Abbildung 5.4.

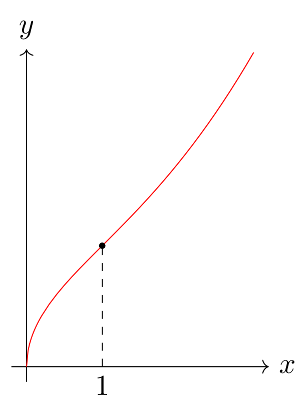

Abbildung 5.4: Die Funktion in Beispiel 5.20.

Ein Punkt auf dem Funktionsgraphen einer Funktion f wird *Wendepunkt* genannt, wenn sich dort das Krümmungsverhalten ändert: von konvex auf konkav oder umgekehrt. So ändert sich in Beispiel 5.20 etwa das Krümmungsverhalten bei $x = 1$, und $f(1) = \frac{16}{10} = \frac{8}{5}$. Also ist der Punkt $(1, \frac{8}{5})$ ein Wendepunkt.

5.2.5 Minima und Maxima

Satz 5.12 sagt aus, dass lokale Extrema (also Minima und Maxima) differenzierbarer Funktionen nur in stationären Punkten vorkommen können. Wir wollen uns nun der Frage widmen, wie man für einen stationären Punkt feststellt, welcher Fall vorliegt. Dazu gibt es zwei Methoden:

Methode 1: erste Ableitung

Diese Methode beruht darauf, die erste Ableitung in der Nähe eines stationären Punktes x_0 näher zu untersuchen. Hier sind im Wesentlichen drei Fälle[4] zu unterscheiden:

- Wenn die Ableitung links von x_0 positiv ist, rechts davon dagegen negativ, dann wechselt die Funktion von steigend auf fallend. Es liegt damit ein Maximum in x_0 vor.
- Wenn die Ableitung links von x_0 negativ ist, rechts davon dagegen positiv, dann wechselt die Funktion von fallend auf steigend. Es liegt damit ein Minimum in x_0 vor.
- Wenn die Ableitung links und rechts von x_0 dasselbe Vorzeichen hat, dann hat die Funktion bei x_0 weder ein Minimum noch ein Maximum.

Methode 2: zweite Ableitung

Wenn die zweite Ableitung in einem stationären Punkt x_0 existiert, dann kann man diese heranziehen um den Typ von x_0 zu bestimmen.

- Ist $f''(x_0) > 0$, dann liegt ein Minimum in x_0 vor.
- Ist $f''(x_0) < 0$, dann liegt ein Maximum in x_0 vor.
- Ist $f''(x_0) = 0$, dann liefert dieser Test kein Ergebnis.

Eine Verfeinerung, die auf dem Taylorpolynom beruht, wird etwas später besprochen.

Beispiel 5.21. Wir betrachten die Funktion $f(x) = x^2 \ln x$ auf $(0, \infty)$. Diese hat die erste Ableitung

$$f'(x) = 2x \ln x + x^2 \cdot \frac{1}{x} = x(2 \ln x + 1).$$

Da die Funktion nur für positives x definiert ist, muss ein stationärer Punkt $2 \ln x = -1$ bzw. $\ln x = -\frac{1}{2}$ erfüllen. Dies ist für $x_0 = e^{-1/2}$ der Fall. Wir können beide Methoden verwenden um festzustellen, ob dort ein Minimum oder Maximum (oder keines von beiden) vorliegt.

Einerseits sehen wir, dass der Faktor $2 \ln x + 1$ und damit auch $f'(x)$ bei $x_0 = e^{-1/2}$ das Vorzeichen von negativ auf positiv wechselt. Also liegt bei x_0 ein lokales Minimum mit dem Wert $f(e^{-1/2}) = -\frac{1}{2e}$. Siehe Abbildung 5.5.

Andererseits können wir auch die zweite Ableitung

$$f''(x) = 2 \ln x + 1 + x \cdot \frac{2}{x} = 2 \ln x + 3$$

verwenden: es ist $f''(x_0) = 2 \ln e^{-1/2} + 3 = 2 \cdot (-\frac{1}{2}) + 3 = 2 > 0$. Wiederum folgt, dass bei x_0 ein lokales Minimum vorliegt.

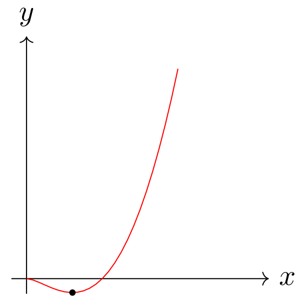

Abbildung 5.5: Die Funktion in Beispiel 5.21.

Ein *globales* Maximum (Minimum) einer Funktion $f : D \to \mathbb{R}$ liegt in einem Punkt x_0 vor, wenn $f(x) \leq f(x_0)$ (bzw. $f(x) \geq f(x_0)$) für *alle* $x \in D$ gilt, nicht nur in einer Umgebung von x_0. Ein globales Extremum, das nicht am Rand liegt, muss nach Definition automatisch ein lokales Extremum sein. Damit ergibt sich die folgende Methode zum Auffinden globaler Extrema einer Funktion f:

- Bestimme alle stationären Punkte sowie gegebenenfalls Punkte, wo f nicht differenzierbar ist.
- Bestimme den Wert von f an allen diesen Punkten sowie Randpunkten von f.
- Der größte der so gefundenen Werte von f liefert das Maximum, der kleinste das Minimum.

Beispiel 5.22. Man bestimme das globale Maximum und das globale Minimum der Funktion $f(x) = 3x + 4\sqrt{5x - x^2}$ auf ihrem (maximalen) Definitionsbereich.

Lösung. Der Ausdruck unter der Wurzel muss ≥ 0 sein, also $5x - x^2 = x(5 - x) \geq 0$. Dies gilt für $0 \leq x \leq 5$, der maximale Definitionsbereich ist also das Intervall $[0, 5]$. Wir bestimmen nun die stationären Punkte:

$$f'(x) = 3 + 4 \cdot \frac{1}{2\sqrt{5x - x^2}} \cdot (5 - 2x) = 3 + \frac{10 - 4x}{\sqrt{5x - x^2}}.$$

Aus $f'(x) = 0$ folgt damit

$$3\sqrt{5x - x^2} = 4x - 10$$

beziehungsweise durch Quadrieren

$$9(5x - x^2) = 16x^2 - 80x + 100.$$

Weitere Vereinfachung führt auf die quadratische Gleichung

$$25x^2 - 125x + 100 = 25(x^2 - 5x + 4) = 0$$

[4] Ohne ins Detail zu gehen sei angemerkt, dass grundsätzlich auch andere Fälle auftreten können, wie etwa die Funktion $f(x) = x^2 \sin(1/x)$ (wobei zusätzlich $f(0) = 0$ definiert wird) illustriert: diese hat bei 0 einen stationären Punkt. Die Ableitung $f'(x) = 2x \sin(1/x) - \cos(1/x)$ wechselt aber in der Nähe von 0 unendlich oft das Vorzeichen.

mit den Lösungen $x = 1$ und $x = 4$. Für $x = 1$ ist allerdings $f'(1) = 3 + \frac{6}{\sqrt{4}} = 6$, nur $x = 4$ ist tatsächlich ein stationärer Punkt: $f'(4) = 3 - \frac{6}{\sqrt{4}} = 0$.
Wir vergleichen daher die Werte von f an drei Punkten: dem stationären Punkt $x = 4$ und den Randpunkten $x = 0$ und $x = 5$. Es ist

$$f(0) = 0 + 4\sqrt{0} = 0, \qquad f(4) = 12 + 4\sqrt{4} = 20, \qquad f(5) = 15 + 4\sqrt{0} = 15.$$

Also hat die Funktion ihr globales Minimum bei 0 mit $f(0) = 0$ und das globale Maximum bei 4 mit $f(4) = 20$.

5.2.6 Kurvendiskussion

In der Kurvendiskussion werden die verschiedenen Punkte, die in diesem Kapitel bisher besprochen wurden, zu einem Ganzen zusammengefasst. Gegeben sei eine reelle Funktion, deren Gestalt und Verhalten wir analysieren wollen. Die folgenden Komponenten gehen in die Kurvendiskussion ein:

- Definitionsbereich
- Nullstellen
- Stationäre Punkte, Minima und Maxima
- Monotonieverhalten (Intervalle, auf denen die Funktion steigt/fällt)
- Krümmungsverhalten (Intervalle, auf denen die Funktion konvex/konkav ist, Wendepunkte)
- Verhalten am Rand, Asymptoten

Mit deren Hilfe kann man zumeist schon eine recht genaue Skizze des Funktionsgraphen machen, die dessen wesentliche Eigenschaften widerspiegelt. Wir illustrieren dies anhand von zwei Beispielen.
Beispiel 5.23. Wir betrachten die Funktion $f(x) = xe^{-x^2/2}$.

- Der Definitionsbereich ist hier ganz $\mathbb{R}$.
- Da die Exponentialfunktion niemals den Wert 0 annimmt, ist $x = 0$ die einzige Nullstelle.
- Wir bestimmen die erste Ableitung mit Hilfe der Produktregel und der Kettenregel:

  $$f'(x) = e^{-x^2/2} + x \cdot e^{-x^2/2} \cdot \Big(-\frac{2x}{2}\Big) = (1 - x^2)e^{-x^2/2}.$$

  Diese hat die Nullstellen $x = 1$ und $x = -1$, die damit stationäre Punkte sind. Der Faktor $e^{-x^2/2}$ ist immer positiv, es kommt also nur auf den Faktor $1 - x^2$ an. Es gilt:

  - Die erste Ableitung ist für $x < -1$ und $x > 1$ negativ, die Funktion f daher streng monoton fallend auf $(-\infty, -1]$ und $[1, \infty)$.
  - Die erste Ableitung ist für $-1 < x < 1$ positiv, die Funktion f daher streng monoton steigend auf $[-1, 1]$.

- Die Funktion hat somit bei $x = -1$ ein lokales Minimum mit dem Funktionswert $f(-1) = -e^{-1/2}$ und bei $x = 1$ ein lokales Maximum mit dem Funktionswert $f(1) = e^{-1/2}$.

- Die zweite Ableitung ist

$$\begin{aligned} f''(x) &= -2xe^{-x^2/2} + (1-x^2) \cdot e^{-x^2/2} \cdot \Big(-\frac{2x}{2}\Big) = -xe^{-x^2/2}(2+1-x^2) \\ &= x(x^2-3)e^{-x^2/2} = x(x-\sqrt{3})(x+\sqrt{3})e^{-x^2/2}. \end{aligned}$$

  Sie ändert an drei Stellen ihr Vorzeichen: bei $x = 0$ und bei $x = \pm\sqrt{3}$. Es gilt:

  - Die zweite Ableitung ist für $x < -\sqrt{3}$ und $0 < x < \sqrt{3}$ negativ, die Funktion f daher streng konkav auf $(-\infty, -\sqrt{3}]$ und $[0, \sqrt{3}]$.
  - Die zweite Ableitung ist für $x > \sqrt{3}$ und $-\sqrt{3} < x < 0$ positiv, die Funktion f daher streng konvex auf $[\sqrt{3}, \infty)$ und $[-\sqrt{3}, 0]$.
  - Es gibt drei Wendepunkte: $(0, 0)$, $(-\sqrt{3}, -\sqrt{3}e^{-3/2})$ und $(\sqrt{3}, \sqrt{3}e^{-3/2})$.

- Für $x \to \pm\infty$ strebt die Funktion $f(x) = xe^{-x^2/2} = \frac{x}{e^{x^2/2}}$ jeweils gegen 0. Dies kann man mit der Regel von de l'Hospital, die im folgenden Abschnitt behandelt wird, gezeigt werden. Die x-Achse ist damit eine Asymptote.

Die verschiedenen Eigenschaften des Funktionsgraphen sind in Abbildung 5.6 illustriert.

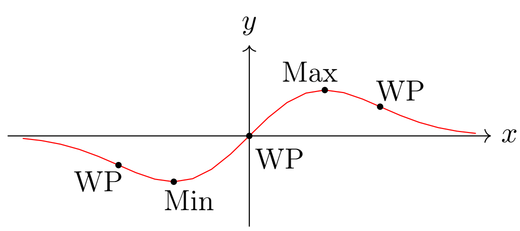

Abbildung 5.6: Die Funktion in Beispiel 5.23.

Beispiel 5.24. Wir betrachten die Funktion $f(x) = \frac{x^2+x-2}{4x-8}$.

- Die Funktion ist für alle reellen x definiert, mit Ausnahme der Nullstelle des Nenners, die bei $x = 2$ liegt. Der Definitionsbereich ist also $\mathbb{R} \setminus \{2\}$.

- Der Zähler hat die Faktorisierung $x^2 + x - 2 = (x+2)(x-1)$. Also sind $x = -2$ und $x = 1$ Nullstellen.

- Wir bestimmen die erste Ableitung mit Hilfe der Quotientenregel:

$$\begin{aligned} f'(x) &= \frac{(2x+1)(4x-8) - (x^2+x-2) \cdot 4}{(4x-8)^2} = \frac{8x^2 - 12x - 8 - 4x^2 - 4x + 8}{(4x-8)^2} \\ &= \frac{4x^2 - 16x}{(4x-8)^2} = \frac{4(x^2-4x)}{4^2(x-2)^2} = \frac{x^2-4x}{4(x-2)^2}. \end{aligned}$$

  Diese hat die Nullstellen $x = 0$ und $x = 4$, die damit stationäre Punkte sind. Der Nenner ist immer positiv, der Zähler wechselt bei den Nullstellen das Vorzeichen. Es gilt:

- Die erste Ableitung ist für $x < 0$ und $x > 4$ positiv, die Funktion f daher streng monoton steigend auf $(-\infty, 0]$ und $[4, \infty)$.
- Die erste Ableitung ist für $0 < x < 2$ und $2 < x < 4$ negativ, die Funktion f daher streng monoton fallend auf $[0, 2)$ und $(2, 4]$ (man beachte, dass $x = 2$ nicht im Definitionsbereich liegt und daher ausgenommen werden muss!).
- Die Funktion hat somit bei $x = 0$ ein lokales Maximum mit dem Funktionswert $f(0) = \frac{1}{4}$ und bei $x = 4$ ein lokales Minimum mit dem Funktionswert $f(4) = \frac{9}{4}$.

- Die zweite Ableitung ist

$$\begin{aligned} f''(x) &= \frac{(2x-4)(x-2)^2 - (x^2-4x)\cdot 2(x-2)}{4(x-2)^4} = \frac{(x-2)((2x-4)(x-2) - 2(x^2-4x))}{4(x-2)^4} \\ &= \frac{(x-2)(2x^2-8x+8-2x^2+8x)}{4(x-2)^4} = \frac{(x-2)\cdot 8}{4(x-2)^4} = \frac{2}{(x-2)^3}. \end{aligned}$$

  Sie ändert nur bei $x = 2$ ihr Vorzeichen. Es gilt:

  - Die zweite Ableitung ist für $x < 2$ negativ, die Funktion f daher streng konkav auf $(-\infty, -2)$.
  - Die zweite Ableitung ist für $x > 2$ positiv, die Funktion f daher streng konvex auf $(2, \infty)$.
  - Es gibt keinen Wendepunkt, denn $x = 2$ liegt nicht im Definitionsbereich.

- Bei $x = 2$ hat die Funktion eine vertikale Asymptote. Der Zähler strebt für $x \to 2$ gegen 4 und ist damit positiv. Der Nenner $4x - 8$ ist links von $x = 2$ negativ, rechts davon positiv. Also gilt:

$$\lim_{x\to 2^-} f(x) = -\infty, \qquad \lim_{x\to 2^+} f(x) = \infty.$$

  Für $x \to \pm\infty$ gibt es eine schiefe Asymptote der Form $y = ax + b$, die wir wie in Beispiel 3.18 bestimmen können:

$$\lim_{x\to\infty} \frac{f(x)}{x} = \lim_{x\to\infty} \frac{x^2+x-2}{4x^2-8x} = \lim_{x\to\infty} \frac{1+x^{-1}-2x^{-2}}{4-8x^{-1}} = \frac{1}{4}.$$

  Nun ist

$$f(x) - \frac{x}{4} = \frac{x^2+x-2}{4x-8} - \frac{x}{4} = \frac{x^2+x-2}{4x-8} - \frac{x(x-2)}{4(x-2)} = \frac{x^2+x-2}{4x-8} - \frac{x^2-2x}{4x-8} = \frac{3x-2}{4x-8}$$

  und damit

$$\lim_{x\to\infty}\left(f(x) - \frac{x}{4}\right) = \lim_{x\to\infty} \frac{3-2x^{-1}}{4-8x^{-1}} = \frac{3}{4}.$$

  Dieselben Grenzwerte gelten auch für $x \to -\infty$. Also ist $y = \frac{x}{4} + \frac{3}{4}$ eine Asymptote.

Die verschiedenen Eigenschaften des Funktionsgraphen sind in Abbildung 5.7 illustriert.

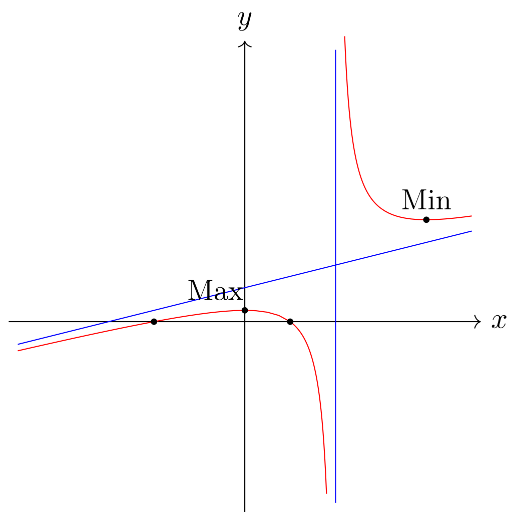

Abbildung 5.7: Die Funktion in Beispiel 5.24.

5.3 Die Regel von de l'Hospital

Die Regel von de l'Hospital ermöglicht es, Grenzwerte zu berechnen, die einem *unbestimmten Ausdruck* wie $\frac{0}{0}$, $\frac{\infty}{\infty}$ oder $0 \cdot \infty$ entsprechen. Wir werden diese Regel zunächst in einem Fall formulieren und beweisen und dann anhand verschiedener Beispiele illustrieren. Zum Beweis brauchen wir zunächst eine Erweiterung des Mittelwertsatzes, den *zweiten* oder *Cauchy'schen* Mittelwertsatz:

Satz 5.25 (2. Mittelwertsatz)**.** *Es seien f und g stetig auf dem Intervall $[a, b]$ und differenzierbar auf (a, b). Dann gibt es ein $c \in (a, b)$, sodass*

$$(f(b) - f(a))g'(c) = (g(b) - g(a))f'(c).$$

Man beachte, dass der Spezialfall $g(x) = x$ den (ersten) Mittelwertsatz ergibt.

Beweis. Wenn $g(b) = g(a)$, dann folgt die Aussage sofort aus dem Satz von Rolle (angewandt auf g): es gibt dann nämlich ein $c \in (a, b)$ mit $g'(c) = 0$, und beide Seiten der Gleichung sind 0. Wir nehmen nun an, dass $g(b) \neq g(a)$, und betrachten eine neue Funktion $h(x) = f(x) - kg(x)$, wobei die Konstante k so gewählt wird, dass $h(b) = h(a)$: dazu muss

$$f(b) - kg(b) = f(a) - kg(a)$$

sein, woraus sich $k = \frac{f(b)-f(a)}{g(b)-g(a)}$ ergibt. Die Funktion h erfüllt nun alle Bedingungen des Satzes von Rolle. Es gibt daher ein c derart, dass

$$0 = h'(c) = f'(c) - kg'(c) = f'(c) - \frac{f(b) - f(a)}{g(b) - g(a)} g'(c).$$

Durch Umschreiben ergibt sich die gewünschte Gleichung.

Nun sind wir bereit für die Regel von de l'Hospital, die wir für Grenzwerte der Form $\frac{0}{0}$ formulieren.

Satz 5.26 (Regel von de l'Hospital)**.** *Es seien f und g differenzierbar in einem Intervall I um a, und es sei $f(a) = g(a) = 0$ sowie $g'(x) \neq 0$ für $x \in I \setminus \{a\}$. Wenn der Grenzwert*

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

existiert, dann auch der Grenzwert

$$\lim_{x \to a} \frac{f(x)}{g(x)},$$

und die beiden sind gleich:

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

Beweis. Es muss $g(x) \neq 0$ für $x \in I \setminus \{a\}$ gelten, da sonst wegen $g(a) = 0$ aus dem Satz von Rolle die Existenz eines $c \in I \setminus \{a\}$ zwischen a und x mit $g'(c) = 0$ folgen würde, im Widerspruch zur Annahme. Aus der Annahme, dass $f(a) = g(a) = 0$, folgt mit Hilfe des zweiten Mittelwertsatzes

$$\frac{f(x)}{g(x)} = \frac{f(x) - f(a)}{g(x) - g(a)} = \frac{f'(c)}{g'(c)}$$

für ein c zwischen a und x. Wenn nun $x \to a$, dann folgt auch $c \to a$, und damit

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

Beispiel 5.27. Man bestimme den Grenzwert

$$\lim_{x \to 1} \frac{\ln x}{2^x - x - 1}.$$

Lösung. Der gegebene Grenzwert hat die Form $\frac{0}{0}$, denn sowohl Zähler als auch Nenner sind an der Stelle $x = 1$ gleich 0. Wir können daher die Regel von de l'Hospital anwenden und erhalten

$$\lim_{x \to 1} \frac{\ln x}{2^x - x - 1} = \lim_{x \to 1} \frac{\frac{1}{x}}{2^x \ln 2 - 1} = \frac{1}{2 \ln 2 - 1}.$$

Mitunter ist es nötig, die Regel wiederholt anzuwenden, wenn der Quotient $\frac{f'(x)}{g'(x)}$ im Grenzwert wieder einen unbestimmten Ausdruck ergibt.

Beispiel 5.28. Man bestimme den Grenzwert

$$\lim_{x \to 0} \frac{\sinh x - \sin x}{x^2 \ln(1 + 2x)}.$$

Lösung. Ein erstes Anwenden der Regel von de l'Hospital liefert

$$\lim_{x \to 0} \frac{\sinh x - \sin x}{x^2 \ln(1 + 2x)} = \lim_{x \to 0} \frac{\cosh x - \cos x}{2x \ln(1 + 2x) + \frac{2x^2}{1+2x}} = \lim_{x \to 0} \frac{(\cosh x - \cos x)(1 + 2x)}{(2x + 4x^2) \ln(1 + 2x) + 2x^2}.$$

Der neue Quotient hat für $x = 0$ wieder die Form $\frac{0}{0}$. Wir können die Regel daher ein zweites Mal anwenden und erhalten

$$\begin{aligned}
\lim_{x\to 0} \frac{(\cosh x - \cos x)(1+2x)}{(2x+4x^2)\ln(1+2x)+2x^2} &= \lim_{x\to 0}(1+2x) \cdot \lim_{x\to 0} \frac{\cosh x - \cos x}{(2x+4x^2)\ln(1+2x)+2x^2} \\
&= \lim_{x\to 0} \frac{\cosh x - \cos x}{(2x+4x^2)\ln(1+2x)+2x^2} \\
&= \lim_{x\to 0} \frac{\sinh x + \sin x}{(2+8x)\ln(1+2x) + \frac{2(2x+4x^2)}{1+2x} + 4x} \\
&= \lim_{x\to 0} \frac{\sinh x + \sin x}{(2+8x)\ln(1+2x) + 4x + 4x} \\
&= \lim_{x\to 0} \frac{\sinh x + \sin x}{(2+8x)\ln(1+2x) + 8x}.
\end{aligned}$$

Wiederum hat der Grenzwert die Form $\frac{0}{0}$. Damit ist es möglich, die Regel von de l'Hospital ein drittes Mal anzuwenden. Wir erhalten

$$\begin{aligned}
\lim_{x\to 0} \frac{\sinh x + \sin x}{(2+8x)\ln(1+2x) + 8x} &= \lim_{x\to 0} \frac{\cosh x + \cos x}{8\ln(1+2x) + \frac{2(2+8x)}{1+2x} + 8} \\
&= \frac{1+1}{8\cdot 0 + \frac{4}{1} + 8} = \frac{2}{12} = \frac{1}{6}.
\end{aligned}$$

Der gegebene Grenzwert hat also den Wert $\frac{1}{6}$.

Bemerkung 5.29.

- Die Regel von de l'Hospital gilt auch, wenn der Grenzwert die Form $\frac{\infty}{\infty}$ anstelle von $\frac{0}{0}$ hat, und auch für links- oder rechtsseitige Grenzwerte sowie Grenzwerte bei $\pm\infty$.

- Die Regel kann nicht verwendet werden, wenn der Grenzwert nicht von der Form $\frac{0}{0}$ oder $\frac{\infty}{\infty}$ ist. Zum Beispiel ist

  $$\lim_{x\to 1} \frac{\ln x}{2^x - x} = \frac{\ln 1}{2^1 - 1} = \frac{0}{2-1} = 0,$$

  und nicht etwa

  $$\lim_{x\to 1} \frac{\frac{1}{x}}{2^x \ln 2 - 1} = \frac{1}{2\ln 2 - 1},$$

  da der Nenner bei $x = 1$ nicht den Wert 0 hat.

Das folgende Beispiel illustriert den Fall $\frac{\infty}{\infty}$:

Beispiel 5.30. Man bestimme den Grenzwert

$$\lim_{x\to\infty} \frac{\ln x}{\sqrt{x}}.$$

Lösung. Es ist nach der Regel von de l'Hospital

$$\lim_{x\to\infty} \frac{\ln x}{\sqrt{x}} = \lim_{x\to\infty} \frac{\ln x}{x^{1/2}} = \lim_{x\to\infty} \frac{\frac{1}{x}}{\frac{1}{2}x^{-1/2}} = \lim_{x\to\infty} \frac{2}{\sqrt{x}} = 0.$$

Bemerkung 5.31. In gleicher Weise lässt sich zeigen, dass $\lim_{x\to\infty} \frac{\ln x}{x^\alpha} = 0$ für alle positiven reellen Zahlen α gilt. Der Logarithmus wächst also langsamer als jede beliebige Potenz von x.

Es lassen sich auch andere unbestimmte Ausdrücke mit Hilfe der Regel von de l'Hospital behandeln. Dazu gehören neben $\frac{0}{0}$ und $\frac{\infty}{\infty}$ auch $0\cdot\infty$, $\infty - \infty$, 0^0, ∞^0 und 1^∞.

Der Fall $0 \cdot \infty$

Durch Umschreiben eines Produkts $f(x)g(x)$ in einen Quotienten $\frac{f(x)}{\frac{1}{g(x)}}$ oder $\frac{g(x)}{\frac{1}{f(x)}}$ lässt sich dieser Fall auf $\frac{0}{0}$ oder $\frac{\infty}{\infty}$ zurückführen.

Beispiel 5.32. Man bestimme den Grenzwert

$$\lim_{x\to\infty} x\Big(\frac{\pi}{2} - \arctan x\Big).$$

Lösung. Man beachte, dass der Faktor x gegen ∞ strebt, der Faktor $\frac{\pi}{2} - \arctan x$ dagegen gegen 0. Wenn wir also umschreiben auf

$$x\Big(\frac{\pi}{2} - \arctan x\Big) = \frac{\frac{\pi}{2} - \arctan x}{\frac{1}{x}},$$

dann sind wir wieder im Fall $\frac{0}{0}$ und können die Regel von de l'Hospital anwenden. Diese liefert

$$\lim_{x\to\infty} x\Big(\frac{\pi}{2} - \arctan x\Big) = \lim_{x\to\infty} \frac{\frac{\pi}{2} - \arctan x}{\frac{1}{x}} = \lim_{x\to\infty} \frac{-\frac{1}{1+x^2}}{-\frac{1}{x^2}} = \lim_{x\to\infty} \frac{x^2}{1+x^2}.$$

Diesen Grenzwert können wir wiederum bestimmen, indem wir Zähler und Nenner durch x^2 dividieren. Damit ergibt sich

$$\lim_{x\to\infty} x\Big(\frac{\pi}{2} - \arctan x\Big) = \lim_{x\to\infty} \frac{1}{x^{-2}+1} = \frac{1}{0+1} = 1.$$

Der Fall $\infty - \infty$

Der Grenzwert einer Differenz $f(x) - g(x)$, in der sowohl $f(x)$ als auch $g(x)$ gegen ∞ strebt, lässt sich ebenso mit der Regel von de l'Hospital behandeln. Zunächst schreibt man

$$f(x) - g(x) = g(x)\Big(\frac{f(x)}{g(x)} - 1\Big).$$

Nun betrachtet man den Grenzwert von $\frac{f(x)}{g(x)} - 1$:

- Hat $\frac{f(x)}{g(x)} - 1$ einen positiven Grenzwert, dann ist der Grenzwert von $f(x) - g(x) = g(x)\Big(\frac{f(x)}{g(x)} - 1\Big)$ gleich $+\infty$.
- Hat $\frac{f(x)}{g(x)} - 1$ einen negativen Grenzwert, dann ist der Grenzwert von $f(x) - g(x) = g(x)\Big(\frac{f(x)}{g(x)} - 1\Big)$ gleich $-\infty$.
- Hat $\frac{f(x)}{g(x)} - 1$ den Grenzwert 0, dann ist man im Fall $0 \cdot \infty$.

Beispiel 5.33. Man bestimme den Grenzwert

$$\lim_{x\to 0} \cot x - \frac{1}{e^x - 1}.$$

Lösung. Man kann nachprüfen, dass der Grenzwert des Quotienten 1 ist:

$$\lim_{x\to 0} \frac{\cot x}{\frac{1}{e^x-1}} = \lim_{x\to 0} \cos x \cdot \frac{e^x-1}{\sin x} = \cos 0 \cdot \lim_{x\to 0} \frac{e^x-1}{\sin x} = \lim_{x\to 0} \frac{e^x}{\cos x} = \frac{e^0}{\cos 0} = 1.$$

Um den gegebenen Grenzwert zu bestimmen, formen wir zunächst um:

$$\cot x - \frac{1}{e^x-1} = \frac{\cos x}{\sin x} - \frac{1}{e^x-1} = \frac{(e^x-1)\cos x - \sin x}{(e^x-1)\sin x}.$$

Daher gilt nach der Regel von de l'Hospital (die wir zweimal anwenden)

$$\begin{aligned}
\lim_{x\to 0} \cot x - \frac{1}{e^x-1} &= \lim_{x\to 0} \frac{(e^x-1)\cos x - \sin x}{(e^x-1)\sin x} \\
&= \lim_{x\to 0} \frac{e^x\cos x - (e^x-1)\sin x - \cos x}{e^x \sin x + (e^x-1)\cos x} \\
&= \lim_{x\to 0} \frac{e^x(\cos x - \sin x) + \sin x - \cos x}{e^x(\sin x + \cos x) - \cos x} \\
&= \lim_{x\to 0} \frac{e^x(\cos x - \sin x) + e^x(-\sin x - \cos x) + \cos x + \sin x}{e^x(\sin x + \cos x) + e^x(\cos x - \sin x) + \sin x} \\
&= \lim_{x\to 0} \frac{-2e^x \sin x + \cos x + \sin x}{2e^x\cos x + \sin x} \\
&= \frac{0+1+0}{2+0} = \frac{1}{2}.
\end{aligned}$$

Die Fälle 0^0, ∞^0 und 1^∞

Durch die Umformung $f(x)^{g(x)} = e^{g(x)\ln f(x)}$ werden diese unbestimmten Ausdrücke auf die Form $0 \cdot \infty$ (im Exponenten) zurückgeführt.

Beispiel 5.34. Man bestimme den Grenzwert

$$\lim_{x\to 1} x^{\tan(\pi x/2)}.$$

Lösung. Wir schreiben zunächst um:

$$x^{\tan(\pi x/2)} = \exp\Big(\ln x \tan(\pi x/2)\Big) = \exp\Big(\frac{\ln x}{\cot(\pi x/2)}\Big).$$

Nun ist der Grenzwert im Exponenten

$$\lim_{x\to 1} \frac{\ln x}{\cot(\pi x/2)} = \lim_{x\to 1} \frac{\frac{1}{x}}{-\sin^2(\pi x/2)\cdot \pi/2} = -\frac{1}{\sin^2(\pi/2)\cdot\pi/2} = -\frac{2}{\pi}.$$

Damit ist der ursprüngliche Grenzwert

$$\lim_{x\to 1} x^{\tan(\pi x/2)} = e^{-2/\pi}.$$

5.4 Taylorpolynome und Taylorreihen

5.4.1 Die Taylor-Formel

Polynome sind besonders einfache Funktionen und haben als solche eine zentrale Stellung. Damit erklärt sich auch der Wunsch kompliziertere Funktionen durch Polynome zu approximieren. Wie der Approximationssatz von Weierstraß (Satz 3.36) zeigt, ist es grundsätzlich möglich, beliebige stetige Funktionen auf einem gegebenen Intervall durch Polynome in dem Sinn zu approximieren, dass die Differenz überall klein ist.

Eine andere Art der Approximation besteht darin, eine gegebene Funktion in der Umgebung eines Punktes a durch ein Polynom zu approximieren. Eine natürliche Wahl besteht darin ein Polynom zu bestimmen, dessen Wert und Ableitungen bis zu einem gewissen Grad übereinstimmen.

Wir betrachten zunächst den Fall $a = 0$. Es sei eine Funktion f gegeben, die im Punkt 0 mindestens n Mal differenzierbar ist. Wir suchen ein Polynom P mit

$$P(0) = f(0),\ P'(0) = f'(0),\ P''(0) = f''(0), \dots, P^{(n)}(0) = f^{(n)}(0). \tag{5.2}$$

Wie sich herausstellt, gibt es genau ein Polynom vom Grad $\leq n$, das diese Eigenschaft hat: es sei

$$P(x) = c_n x^n + c_{n-1}x^{n-1} + \cdots + c_2x^2 + c_1x + c_0.$$

Dann ist

$$\begin{aligned}
P'(x) &= nc_nx^{n-1} + (n-1)c_{n-1}x^{n-2} + \cdots + 2c_2x + c_1,\\
P''(x) &= n(n-1)c_nx^{n-2} + (n-1)(n-2)c_{n-1}x^{n-3} + \cdots + 2c_2,\\
P'''(x) &= n(n-1)(n-2)c_nx^{n-3} + (n-1)(n-2)(n-3)c_{n-1}x^{n-4} + \cdots + 6c_3,\\
&\ \vdots\\
P^{(n)}(x) &= n!c_n.
\end{aligned}$$

Wenn wir $x = 0$ einsetzen, dann fallen jeweils alle Terme bis auf den letzten weg: es gilt also $P(0) = c_0$, $P'(0) = c_1$, $P''(0) = 2c_2$, $P'''(0) = 6c_3$, $\dots$, $P^{(n)}(0) = n!c_n$. Wenn wir also (5.2) erfüllen wollen, dann müssen wir

$$c_k = \frac{1}{k!}f^{(k)}(0)$$

für alle k setzen. Das resultierende Polynom

$$P(x) = f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \cdots + \frac{f^{(n)}}{n!}x^n = \sum_{k=0}^{n} \frac{f^{(k)}(0)}{k!}x^k$$

wird n-tes *Taylorpolynom* von f an der Stelle 0 genannt. Ein solches Polynom kann man an jeder Stelle a bestimmen, an der f genügend oft differenzierbar ist. Setzt man

$$P(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(n)}}{n!}(x-a)^n = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k,$$

dann hat dieses Polynom die Eigenschaft, dass

$$P(0) = f(a),\ P'(a) = f'(a),\ P''(a) = f''(a), \dots,\ P^{(n)}(a) = f^{(n)}(a).$$

Man spricht vom n-ten Taylorpolynom an der Stelle a. Der folgende Satz ist nützlich um die Differenz zwischen einer Funktion und ihrem Taylorpolynom zu beschränken.

Satz 5.35 (Taylor-Formel)**.** *Ist die Funktion f auf einem Intervall I, das den Punkt a enthält, $n+1$ Mal differenzierbar, dann gibt es für jedes $x \in I$ ein ξ zwischen a und x, für das*

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n + \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-a)^{n+1}.$$

Beweis. Man betrachte die Funktion

$$\begin{aligned} F(t) &= f(t) + f'(t)(x-t) + \frac{f''(t)}{2!}(x-t)^2 + \cdots + \frac{f^{(n)}(t)}{n!}(x-t)^n + C(x-t)^{n+1} \\ &= \sum_{k=0}^{n} \frac{f^{(k)}(t)}{k!}(x-t)^k + C(x-t)^{n+1}, \end{aligned}$$

wobei die Konstante C so gewählt wird, dass $F(a) = F(x) = f(x)$. Wenn wir diese Funktion nach t differenzieren, so erhalten wir unter Zuhilfenahme der Produktregel

$$\begin{aligned} F'(t) &= \sum_{k=0}^{n} \Big(\frac{f^{(k+1)}(t)}{k!}(x-t)^k - \frac{f^{(k)}(t)}{k!} k(x-t)^{k-1} \Big) - C(n+1)(x-t)^n \\ &= \sum_{k=0}^{n} \Big(\frac{f^{(k+1)}(t)}{k!}(x-t)^k \Big) - \sum_{k=1}^{n} \Big(\frac{f^{(k)}(t)}{(k-1)!}(x-t)^{k-1} \Big) - C(n+1)(x-t)^n \\ &= f'(t) + \frac{f''(t)}{1!}(x-t) + \frac{f'''(t)}{2!}(x-t)^2 + \cdots + \frac{f^{(n+1)}(t)}{n!}(x-t)^n \\ &\quad - f'(t) - \frac{f''(t)}{1!}(x-t) - \frac{f'''(t)}{2!}(x-t)^2 - \cdots - \frac{f^{(n)}(t)}{(n-1)!}(x-t)^{n-1} - C(n+1)(x-t)^n. \end{aligned}$$

Es kürzen sich fast alle Terme paarweise weg, und übrig bleibt nur

$$F'(t) = \frac{f^{(n+1)}(t)}{n!}(x-t)^n - C(n+1)(x-t)^n = \Big(\frac{f^{(n+1)}(t)}{n!} - (n+1)C \Big)(x-t)^n.$$

Da C derart gewählt wurde, dass $F(a) = F(x)$, können wir den Satz von Rolle verwenden: es gibt ein ξ zwischen a und x derart, dass

$$F'(\xi) = \Big(\frac{f^{(n+1)}(\xi)}{n!} - (n+1)C \Big)(x-\xi)^n = 0.$$

Daraus folgt (weil $x \neq \xi$), dass $C = \frac{f^{(n+1)}(\xi)}{(n+1)!}$. Es gilt also

$$f(x) = F(x) = F(a) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n + \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-a)^{n+1},$$

womit der Satz bewiesen ist.

Beispiel 5.36. Wir bestimmen das Taylorpolynom sechsten Grades der Funktion $f(x) = \sin x$ im Punkt 0. Die Ableitungen sind

$$f'(x) = \cos x,\ f''(x) = -\sin x,\ f'''(x) = -\cos x,\ f^{(4)}(x) = \sin x,\ f^{(5)}(x) = \cos x,\ f^{(6)}(x) = -\sin x,$$

und damit

$$f'(0) = 1,\ f''(0) = 0,\ f'''(0) = -1,\ f^{(4)}(0) = 0,\ f^{(5)}(x) = 1,\ f^{(6)}(x) = 0.$$

Also ist das Taylorpolynom für $n = 6$

$$T(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} = x - \frac{x^3}{6} + \frac{x^5}{120}.$$

Wir können dieses Polynom verwenden um numerische Werte der Sinusfunktion zu bestimmen, zum Beispiel

$$\sin 0.6 \approx T(0.6) = 0.6 - \frac{0.6^3}{6} + \frac{0.6^5}{120} = 0.564648.$$

Der Fehler kann mit Hilfe von Satz 5.35 abgeschätzt werden: die siebte Ableitung ist $f^{(7)}(x) = -\cos x$, also gilt

$$\sin 0.6 - T(0.6) = -\frac{\cos\xi}{7!} 0.6^7$$

für ein $\xi \in (0, 0.6)$. Da $|\cos\xi| \leq 1$ jedenfalls gilt, folgt

$$|\sin 0.6 - T(0.6)| \leq \frac{0.6^7}{7!} \leq 6 \cdot 10^{-6}.$$

Also ist $0.564642 = 0.564648 - 6 \cdot 10^{-6} \leq \sin 0.6 \leq 0.564648 + 6 \cdot 10^{-6} = 0.564654$. Man kann also mit einem Taylorpolynom relativ niedrigen Grades schon eine sehr hohe Genauigkeit erreichen.

Stationäre Punkte

Es sei a ein stationärer Punkt von f. Wir können das Taylorpolynom verwenden um zu bestimmen, ob ein Maximum, ein Minimum, oder keines von beiden vorliegt. Es sei n die kleinste natürliche Zahl, für die die n-te Ableitung von f an der Stelle a ungleich 0 ist, also

$$f'(a) = f''(a) = \cdots = f^{(n-1)}(a) = 0, \quad f^{(n)}(a) \neq 0.$$

Das Taylorpolynom n-ten Grades ist dann

$$f(a) + \frac{f^{(n)}(a)}{n!}(x-a)^n,$$

da alle anderen Terme 0 sind. In der Nähe von a wird f durch das Taylorpolynom approximiert:

$$f(x) \approx f(a) + \frac{f^{(n)}(a)}{n!}(x-a)^n.$$

Nun können folgende Fälle auftreten:

- n ist gerade, und $f^{(n)}(a) > 0$: in diesem Fall ist $\frac{f^{(n)}(a)}{n!}(x-a)^n \geq 0$, womit f bei a ein lokales Minimum hat.
- n ist gerade, und $f^{(n)}(a) < 0$: in diesem Fall ist $\frac{f^{(n)}(a)}{n!}(x-a)^n \leq 0$, womit f bei a ein lokales Maximum hat.
- n ist ungerade: in diesem Fall wechselt $\frac{f^{(n)}(a)}{n!}(x-a)^n$ das Vorzeichen bei a, und es liegt kein lokales Extremum vor.

5.4.2 Taylorreihen

Wenn in der Taylor-Formel das Restglied $\frac{f^{(n+1)}(\xi)}{(n+1)!}(x-a)^{n+1}$ für $n \to \infty$ gegen 0 strebt, dann ist $f(a)$ gleich dem Wert der *Taylorreihe*

$$\sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots$$

Dies ist nicht notwendigerweise immer der Fall, aber gilt für viele wichtige Funktionen.

Beispiel 5.37. Wir setzen Beispiel 5.36 fort: die Ableitungen der Sinusfunktion setzen sich periodisch fort: die geradzahligen Ableitungen sind $f^{(2n)}(x) = (-1)^n \sin x$, die ungeradzahligen $f^{(2n+1)}(x) = (-1)^n \cos x$. Damit gilt $f^{(2n)}(0) = 0$ und $f^{(2n+1)}(x) = (-1)^n$. Also ist die Taylorreihe an der Stelle 0

$$\sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots$$

Wir zeigen, dass diese Reihe tatsächlich für jedes x gleich $\sin x$ ist, indem wir das Restglied abschätzen: jede Ableitung ist in diesem Beispiel entweder $\pm \sin x$ oder $\pm \cos x$. Damit gilt jedenfalls $|f^{(n+1)}(\xi)| \leq 1$ und daher

$$\left| \frac{f^{(n+1)}(\xi)}{(n+1)!} x^{n+1} \right| \leq \frac{|x|^{n+1}}{(n+1)!}.$$

Wir zeigen, dass dieser Quotient für $n \to \infty$ immer gegen 0 strebt. Es sei K so groß, dass $K \geq |x|$. Es gilt

$$\begin{aligned}
\frac{|x|^{n+1}}{(n+1)!} &= \frac{|x|^{n+1}}{1 \cdot 2 \cdots (n+1)} \\
&= \frac{|x|^K}{K!} \cdot \frac{|x|^{n+1-K}}{(K+1)(K+2)\cdots(n+1)} \\
&\leq \frac{|x|^K}{K!} \cdot \frac{|x|^{n+1-K}}{(K+1)(K+1)\cdots(K+1)} \\
&= \frac{|x|^K}{K!} \cdot \left(\frac{|x|}{K+1}\right)^{n+1-K} \\
&\leq \frac{K^K}{K!} \cdot \left(\frac{K}{K+1}\right)^{n+1-K}.
\end{aligned}$$

Weil $\frac{K}{K+1} < 1$ gilt, hat die geometrische Folge am Ende der Ungleichungskette den Grenzwert 0. Also erhalten wir (unter Zuhilfenahme von Satz 2.9), dass das Restglied tatsächlich gegen 0 geht. Die Funktion $\sin x$ ist also überall gleich ihrer Taylorreihe.

Auf ähnliche Art erhält man auch die Taylorreihe der Cosinusfunktion und zeigt, dass die Funktion überall durch die Taylorreihe dargestellt wird:

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

Da alle Ableitungen der Exponentialfunktion e^x wieder gleich e^x sind, ist die Taylorreihe im Punkt 0 genau

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

also genau die Reihe, durch die wir die Exponentialfunktion in Kapitel 3.7 definiert haben.

Nicht alle Funktionen werden jedoch durch ihre Taylorreihe dargestellt. So kann man etwa zeigen, dass alle Ableitungen der Funktion

$$f(x) = \begin{cases} e^{-1/x^2} & x \neq 0, \\ 0 & x = 0, \end{cases}$$

im Punkt 0 gleich 0 sind. Die Taylorreihe wäre damit auch überall 0 und damit nicht gleich $f(x)$. Es gibt auch, wie wir im nächsten Abschnitt sehen werden, Taylorreihen, die nur auf einem begrenzten Bereich überhaupt konvergent sind.

5.4.3 Funktionenfolgen und Funktionenreihen

Wenn eine Folge von Funktionen $f_1(x), f_2(x), \ldots$ mit demselben Definitionsbereich D die Eigenschaft hat, dass

$$\lim_{n\to\infty} f_n(x) = f(x)$$

für alle $x \in D$, dann sagen wir, dass die Funktionenfolge f_n *punktweise* gegen f konvergiert. Zum Beispiel gilt für $f_n(x) = 1 + x + \frac{x^2}{n}$, dass $\lim_{n\to\infty} f_n(x) = 1 + x$ für alle $x \in \mathbb{R}$. Wie sich zeigt, bleiben Eigenschaften der Funktionen f_n im Grenzwert nicht notwendigerweise erhalten. Etwa muss der Grenzwert einer Folge stetiger Funktionen nicht stetig sein, wie das folgende Beispiel illustriert.

Beispiel 5.38. Die stetigen Funktionen $f_n(x) = x^n$ auf dem Intervall $[0, 1]$ streben punktweise gegen den Grenzwert

$$f(x) = \begin{cases} 1 & x = 1, \\ 0 & x < 1. \end{cases}$$

Diese Funktion ist offensichtlich unstetig bei 1.

Eine stärkere Bedingung als punktweise Konvergenz, die solche Beispiele ausschließt, ist die *gleichmäßige Konvergenz.*

Definition 5.39. Eine Folge von Funktionen f_n mit gemeinsamem Definitionsbereich D heißt gleichmäßig konvergent gegen f, wenn es zu jedem $\epsilon > 0$ ein N_ϵ derart gibt, dass

$$|f_n(x) - f(x)| < \epsilon$$

für alle $n \geq N_\epsilon$ und alle $x \in D$ gilt.

Die Definition bedeutet, dass sich der Graph von f_n für jedes ϵ ab einem gewissen Punkt in einem Streifen der Breite 2ϵ (ϵ in jede Richtung) um den Graphen von f befindet. Der Unterschied zur punktweisen Konvergenz ist etwas subtil: der wesentliche Punkt ist, dass N_ϵ nicht von x abhängt. Aus der punktweisen Konvergenz folgt nämlich die Existenz eines solchen N_ϵ für jedes feste x (nach der Definition der Konvergenz von Folgen), aber es muss nicht notwendigerweise ein einziges N_ϵ für alle x geben.

Beispiel 5.40. Wir setzen Beispiel 5.38 fort. Die Folge ist nicht gleichmäßig konvergent, denn zu jedem $\epsilon > 0$ und jeder natürlichen Zahl n gibt es Werte von x in $[0,1]$ für die $|f_n(x) - f(x)| > \epsilon$, nämlich alle $x \in (\epsilon^{1/n}, 1)$.

Wenn wir uns auf ein geeignetes kleineres Intervall beschränken, dann wird die Folge gleichmäßig konvergent: betrachten wir nur das Intervall $[0, \frac{1}{2}]$, so ist die Folge $f_n(x) = x^n$ gleichmäßig konvergent gegen $f(x) = 0$. Wenn wir nämlich $N_\epsilon > -\log_2 \epsilon$ wählen, dann gilt für alle $x \in [0, \frac{1}{2}]$ und alle $n \geq N_\epsilon$, dass

$$|f_n(x) - f(x)| = x^n \leq \left(\tfrac{1}{2}\right)^n < \left(\tfrac{1}{2}\right)^{-\log_2 \epsilon} = 2^{\log_2 \epsilon} = \epsilon.$$

Satz 5.41. *Ist f_n eine gleichmäßig konvergente Folge stetiger Funktionen, dann ist die Grenzfunktion f auch stetig.*

Ein wichtiger Spezialfall ist jener, bei dem f_n die Folge der Partialsummen einer Reihe ist:

$$f_n(x) = \sum_{k=0}^{n} g_k(x).$$

Hier gilt das folgende Kriterium für die gleichmäßige Konvergenz zur Reihe $\sum_{k=0}^{\infty} g_k(x)$:

Satz 5.42 (Weierstraß-Kriterium)**.** *Es sei g_k eine Folge von Funktionen mit gemeinsamem Definitionsbereich D, und es sei M_k eine Folge reeller Zahlen, für die gilt, dass*

$$|g_k(x)| \leq M_k$$

für alle k und alle $x \in D$, und $\sum_{k \geq 0} M_k < \infty$. Dann ist die Funktionenreihe

$$\sum_{k \geq 0} g_k(x)$$

gleichmäßig konvergent, d.h., die Folge der Partialsummen $\sum_{k=0}^{n} g_k(x)$ konvergiert gleichmäßig.

Beispiel 5.43. Als ein Beispiel betrachten wir die Reihe für die Exponentialfunktion, die in Kapitel 3.7 eingeführt wurde:

$$\exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!}.$$

Wenn wir uns auf ein Intervall $[-A, A]$ beschränken, dann gilt für die Summanden

$$\left|\frac{x^n}{n!}\right| \leq \frac{A^n}{n!}.$$

Wir können daher $M_n = \frac{A^n}{n!}$ setzen. Die Reihe

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

ist nach dem Quotientenkriterium konvergent. Daher können wir das Weierstraß-Kriterium verwenden: auf jedem Intervall der Form $[-A, A]$ ist die Reihe gleichmäßig konvergent. Da die einzelnen Summanden stetige Funktionen sind, ist es damit auch die Reihe.

Die Reihe in Beispiel 5.43 ist eine sogenannte *Potenzreihe*. Als solche bezeichnet man eine Reihe der Form

$$\sum_{n=0}^{\infty} c_n (x-a)^n.$$

Die *Koeffizienten* c_n der Reihe können dabei reelle oder komplexe Zahlen sein. Die Zahl a heißt *Entwicklungspunkt* der Potenzreihe. Die Taylorreihe einer Funktion um den Punkt a ist eine solche Potenzreihe. In dem sehr wichtigen Spezialfall $a = 0$ heißt die Taylorreihe auch *Maclaurin-Reihe*.

Eine Potenzreihe ist nicht immer konvergent. Als ein einfaches Beispiel ist uns bereits die geometrische Reihe

$$\sum_{n=0}^{\infty} x^n$$

begegnet, die nur für $|x| < 1$ konvergiert und dann den Wert $\frac{1}{1-x}$ hat. Allgemein hat jede Potenzreihe einen sogenannten *Konvergenzradius*. Es gilt der folgende Satz:

Satz 5.44. *Jede (reelle oder komplexe) Potenzreihe $\sum_{n=0}^{\infty} c_n(x-a)^n$ hat einen Konvergenzradius R, für den folgendes gilt:*

- *Die Reihe konvergiert für $|x-a| < R$.*
- *Die Reihe divergiert für $|x-a| > R$.*

Die Potenzreihe konvergiert also immer auf dem Intervall $(a-R, a+R)$. An den Rändern $a \pm R$ kann die Reihe konvergent oder divergent sein (es ist dabei möglich, dass sie an einem Rand konvergiert, am anderen divergiert). Es ist auch möglich, dass $R = 0$: in diesem Fall ist die Reihe nur im Punkt $x = a$ konvergent.

Um den Konvergenzradius zu bestimmen, kann man das Quotienten- oder Wurzelkriterium heranziehen: es gilt nämlich

$$\lim_{n\to\infty} \left| \frac{c_{n+1}(x-a)^{n+1}}{c_n(x-a)^n} \right| = |x-a| \lim_{n\to\infty} \left| \frac{c_{n+1}}{c_n} \right|,$$

sofern der Grenzwert existiert. Daher konvergiert die Reihe, falls

$$|x-a| < \frac{1}{\lim_{n\to\infty} \left| \frac{c_{n+1}}{c_n} \right|} = \lim_{n\to\infty} \left| \frac{c_n}{c_{n+1}} \right|,$$

und divergiert, falls

$$|x-a| > \frac{1}{\lim_{n\to\infty} \left| \frac{c_{n+1}}{c_n} \right|} = \lim_{n\to\infty} \left| \frac{c_n}{c_{n+1}} \right|.$$

Ebenso gilt

$$\lim_{n\to\infty} \left| c_n(x-a)^n \right|^{1/n} = |x-a| \lim_{n\to\infty} |c_n|^{1/n},$$

sofern der Grenzwert existiert. Daher konvergiert die Reihe, falls

$$|x-a| < \frac{1}{\lim_{n\to\infty} |c_n|^{1/n}} = \lim_{n\to\infty} |c_n|^{-1/n},$$

und divergiert, falls

$$|x-a| > \frac{1}{\lim_{n\to\infty}|c_n|^{1/n}} = \lim_{n\to\infty}|c_n|^{-1/n}.$$

Zusammenfassend haben wir also folgenden Satz.

Satz 5.45. *Wenn der Grenzwert* $\lim_{n\to\infty}|\frac{c_n}{c_{n+1}}|$ *existiert, dann ist dieser gleich dem Konvergenzradius* R *der Potenzreihe* $\sum_{n=0}^{\infty} c_n(x-a)^n$.
Wenn der Grenzwert $\lim_{n\to\infty}|c_n|^{-1/n}$ *existiert, dann ist dieser gleich dem Konvergenzradius* R *der Potenzreihe* $\sum_{n=0}^{\infty} c_n(x-a)^n$.

Beispiel 5.46. Man bestimme den Konvergenzradius der Potenzreihe

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

Lösung. Wir verwenden das Quotientenkriterium: für $c_n = \frac{n!^2}{(2n+1)!}$ gilt

$$\begin{aligned}\frac{c_n}{c_{n+1}} &= \frac{\frac{n!^2}{(2n+1)!}}{\frac{(n+1)!^2}{(2n+3)!}} = \frac{\frac{(2n+3)!}{(2n+1)!}}{(\frac{(n+1)!}{n!})^2} = \frac{\frac{(2n+1)!(2n+2)(2n+3)}{(2n+1)!}}{(\frac{n!(n+1)}{n!})^2} \\ &= \frac{(2n+2)(2n+3)}{(n+1)^2} = \frac{2(n+1)(2n+3)}{(n+1)^2} = \frac{4n+6}{n+1},\end{aligned}$$

und damit

$$R = \lim_{n\to\infty}\frac{c_n}{c_{n+1}} = \lim_{n\to\infty}\frac{4n+6}{n+1} = \lim_{n\to\infty}\frac{4+\frac{6}{n}}{1+\frac{1}{n}} = 4.$$

Der Konvergenzradius ist also 4, und die Reihe ist im Intervall $(1-4, 1+4) = (-3, 5)$ konvergent.

Ähnlich zur Exponentialfunktion in Beispiel 5.43 kann man argumentieren, dass eine Potenzreihe mit Entwicklungspunkt a und Konvergenzradius R auf jedem Teilintervall der Form $[a-r, a+r]$ mit $r < R$ gleichmäßig konvergiert und daher stetig ist. Zudem kann man Potenzreihen differenzieren, indem man die einzelnen Summanden differenziert.

Satz 5.47. *Es sei* R *der Konvergenzradius der Potenzreihe* $f(x) = \sum_{n=0}^{\infty} c_n(x-a)^n$. *Dann ist* $f(x)$ *auf dem Intervall* $(a-R, a+R)$ *stetig und differenzierbar. Die Ableitung ist*

$$f'(x) = \sum_{n=1}^{\infty} nc_n(x-a)^{n-1}.$$

Wenn die Potenzreihe in einem Randpunkt ($a-R$ oder $a+R$) konvergiert, dann ist sie dort auch weiterhin stetig.

Beispiel 5.48. Die Ableitung der Taylorreihe der Sinusfunktion, also

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots$$

ist erwartungsgemäß die Taylorreihe der Cosinusfunktion:

$$\frac{d}{dx}\sin x = 1 - \frac{3x^2}{3!} + \frac{5x^4}{5!} - \frac{7x^6}{7!} + \cdots = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots$$

Wir beenden das Kapitel mit einigen weiteren Beispielen von Reihendarstellungen wichtiger Funktionen.

Beispiel 5.49. Für die Funktion $f(x) = \ln(1+x)$ gilt

$$f'(x) = \frac{1}{1+x},\ f''(x) = -\frac{1}{(1+x)^2},\ f'''(x) = \frac{2}{(1+x)^3},$$

und allgemein (wie man mittels vollständiger Induktion beweisen kann)

$$f^{(n)}(x) = \frac{(-1)^{n-1}(n-1)!}{(1+x)^n}.$$

Daher ist $f^{(n)}(0) = (-1)^{n-1}(n-1)!$ für $n \geq 1$ (sowie $f(0) = 0$), und die Taylorreihe an der Stelle 0 ist

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

Der Konvergenzradius ist in diesem Fall 1. Die Taylorreihe ist im Intervall $(-1, 1]$ gleich der Funktion f:

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \cdots, \qquad -1 < x \leq 1.$$

Die alternierende harmonische Reihe

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots$$

ist also etwa gleich $\ln 2 \approx 0.69315$.

Beispiel 5.50. Es sei α eine beliebige reelle Zahl. Wir betrachten die Funktion $f(x) = (1+x)^\alpha$. Deren Ableitungen sind

$$f'(x) = \alpha(1+x)^{\alpha-1},\ f''(x) = \alpha(\alpha-1)(1+x)^{\alpha-2},\ f'''(x) = \alpha(\alpha-1)(\alpha-2)(1+x)^{\alpha-3},$$

und allgemein

$$f^{(n)}(x) = \alpha(\alpha-1)(\alpha-2)\cdots(\alpha-n+1)(1+x)^{\alpha-n}.$$

Damit ergibt sich die Taylorreihe

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

die für $|x| < 1$ konvergiert und gleich $f(x)$ ist. Die Identität

$$(1+x)^\alpha = \sum_{n=0}^{\infty} \binom{\alpha}{n} x^n$$

verallgemeinert den binomischen Lehrsatz

$$(1+x)^k = \sum_{n=0}^{k} \binom{k}{n} x^n$$

auf nicht-ganzzahlige Exponenten (der Binomialkoeffizient $\binom{k}{n}$ ist für $n > k$ gleich 0, falls k eine natürliche Zahl ist; die Reihe wird damit endlich).

Tabelle 5.2 fasst die Taylorreihen einiger wichtiger Funktionen zusammen.

Taylorreihe	Konvergenzbereich
$(1+x)^\alpha = 1 + \binom{\alpha}{1}x + \binom{\alpha}{2}x^2 + \binom{\alpha}{3}x^3 + \cdots$	$\|x\| < 1$
$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \cdots$	$-1 < x \leq 1$
$\ln(1-x) = x + \frac{x^2}{2} + \frac{x^3}{3} + \frac{x^4}{4} + \cdots$	$-1 \leq x < 1$
$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots$	$x \in \mathbb{R}$
$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots$	$x \in \mathbb{R}$
$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots$	$x \in \mathbb{R}$
$\sinh x = x + \frac{x^3}{3!} + \frac{x^5}{5!} + \frac{x^7}{7!} + \cdots$	$x \in \mathbb{R}$
$\cosh x = 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \frac{x^6}{6!} + \cdots$	$x \in \mathbb{R}$
$\arcsin x = x + \frac{1 \cdot x^3}{2 \cdot 3} + \frac{1 \cdot 3 \cdot x^5}{2 \cdot 4 \cdot 5} + \frac{1 \cdot 3 \cdot 5 \cdot x^7}{2 \cdot 4 \cdot 6 \cdot 7} + \cdots$	$\|x\| \leq 1$
$\arctan x = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \cdots$	$\|x\| < 1$

Tabelle 5.2: Einige wichtige Taylorreihen.

5.5 Das Newton-Verfahren

Das Newton-Verfahren ist eine Methode um Nullstellen von Funktionen zu finden. Es sei $f(x)$ eine differenzierbare Funktion, von der wir eine Nullstelle x^* suchen, also eine reelle Zahl mit $f(x^*) = 0$. Wir nehmen an, dass eine erste Näherung x_0 bekannt sei. Die Grundidee des Newton-Verfahrens besteht darin, die Funktion f durch die Tangente an f im Punkt x_0 zu approximieren. Diese hat die Gleichung

$$y - f(x_0) = f'(x_0)(x - x_0).$$

Der Schnittpunkt mit der x-Achse ($y = 0$) ist durch

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

gegeben. Dieser Wert wird nun als neuer Näherungswert x_1 verwendet (siehe Abbildung 5.8). Das Verfahren wird nun wiederholt: wir verwenden die Tangente im Punkt x_1 als nächste Annäherung und erhalten einen neuen Schnittpunkt x_2.

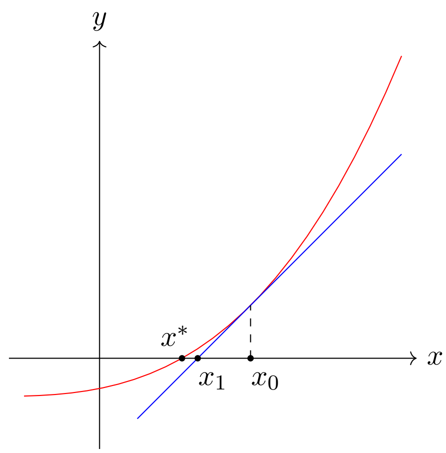

Abbildung 5.8: Newton-Verfahren.

Das Newton-Verfahren kann also folgendermaßen zusammengefasst werden:

- Wähle einen Startwert x_0.
- Definiere die Folge x_n rekursiv durch

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

  für $n \geq 0$.

Es handelt sich also um ein *iteratives* Verfahren, bei dem man versucht sich schrittweise der Lösung anzunähern. Unter geeigneten Voraussetzungen (siehe Satz 5.53) konvergiert die so definierte Folge gegen eine Nullstelle x^*: $\lim_{n\to\infty} x_n = x^*$. Die Wahl des Startwerts x_0 ist bedeutsam und mitunter schwierig: man kann etwa offensichtlich nicht in einem stationären Punkt x_0 beginnen, da dann der Nenner $f'(x_0)$ gleich 0 ist. Es kann aber auch in anderen Fällen vorkommen, dass das Verfahren nicht konvergiert.

Beispiel 5.51. Wir greifen Beispiel 3.29 wieder auf und verwenden das Newton-Verfahren um eine Nullstelle der Funktion $f(x) = \cos x - x$ zu bestimmen. Als Startwert verwenden wir den Punkt $x_0 = 0.5$. Die Rekursion lautet nun

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} = x_n - \frac{\cos x_n - x_n}{-\sin x_n - 1} = x_n + \frac{\cos x_n - x_n}{\sin x_n + 1}.$$

Es ergeben sich die Werte

$$\begin{aligned} x_0 &= 0.5, \\ x_1 &\approx 0.75522, \\ x_2 &\approx 0.73914, \\ x_3 &\approx 0.73909, \\ x_4 &\approx 0.73909. \end{aligned}$$

Die Werte von x_3 und x_4 stimmen also bereits auf fünf Dezimalstellen überein.

Das Newton-Verfahren konvergiert hier also bedeutend schneller als das Bisektionsverfahren, das in Kapitel 3.5 besprochen wurde. Zur Analyse des Verfahrens verwenden wir den sogenannten *Banach'schen Fixpunktsatz.*

Satz 5.52 (Banach'scher Fixpunktsatz)**.** *Es sei $f : [a, b] \to [a, b]$ eine Funktion. Wir nehmen an, dass es ein $K < 1$ derart gibt, dass*

$$|f(x_1) - f(x_2)| \leq K|x_1 - x_2| \tag{5.3}$$

für alle $x_1, x_2 \in [a, b]$. Dann gelten die folgenden Aussagen:

1. *Es gibt genau einen Fixpunkt x^* der Funktion f auf dem Intervall $[a, b]$, also einen Punkt mit $f(x^*) = x^*$.*
2. *Eine rekursive Folge x_n mit beliebigem Startwert $x_0 \in [a, b]$ und $x_{n+1} = f(x_n)$ konvergiert gegen x^*: $\lim_{n\to\infty} x_n = x^*$.*

Eine Funktion wie in Satz 5.52 wird eine *Kontraktion* genannt. Die Ungleichung (5.3) ist unter anderem dann erfüllt, wenn f differenzierbar ist und $|f'(x)| \leq K$ für alle $x \in (a, b)$ gilt. Dies folgt direkt aus dem Mittelwertsatz.

Beweis. Wenn entweder $f(a) = a$ oder $f(b) = b$ gilt, dann hat f einen Fixpunkt. Andernfalls ist $f(a) > a$ und $f(b) < b$, die Funktion $g(x) = f(x) - x$ damit an der Stelle a positiv und an der Stelle b negativ. Also hat g nach dem Zwischenwertsatz eine Nullstelle und damit f einen Fixpunkt.

Wenn es zwei verschiedene Fixpunkte x_1^* und x_2^* gibt, dann gilt

$$\frac{f(x_1^*) - f(x_2^*)}{x_1^* - x_2^*} = \frac{x_1^* - x_2^*}{x_1^* - x_2^*} = 1.$$

Andererseits muss dieser Quotient nach dem Mittelwertsatz gleich $f'(c)$ für ein c zwischen x_1^* und x_2^* sein, im Widerspruch zur Annahme $|f'(c)| \leq K < 1$. Wir schließen also, dass es genau einen Fixpunkt x^* gibt.

Für die durch $x_{n+1} = f(x_n)$ rekursiv definierte Funktion gilt nun

$$|x_{n+1} - x^*| = |f(x_n) - f(x^*)| \leq K|x_n - x^*|,$$

wodurch mit Hilfe von vollständiger Induktion die Ungleichung

$$|x_n - x^*| \leq K^n|x_0 - x^*|$$

folgt. Da K nach Annahme kleiner als 1 ist, gilt $\lim_{n\to\infty} K^n = 0$, also $\lim_{n\to\infty} |x_n - x^*| = 0$ und damit $\lim_{n\to\infty} x_n = x^*$.

Wir können nun die Konvergenz des Newton-Verfahrens analysieren. Um Satz 5.52 verwenden zu können, brauchen wir, dass die Funktion

$$F(x) = x - \frac{f(x)}{f'(x)},$$

die im Newton-Verfahren iteriert wird, eine Kontraktion ist. Die Ableitung dieser Funktion ist

$$F'(x) = 1 - \frac{f'(x)f'(x) - f(x)f''(x)}{f'(x)^2} = \frac{f(x)f''(x)}{f'(x)^2}. \tag{5.4}$$

Damit können wir nun folgenden Satz beweisen:

Satz 5.53. *Wenn die Funktion f in einem Intervall um die Nullstelle x^* zweimal stetig differenzierbar ist und $f'(x^*) \neq 0$ gilt, dann konvergiert das Newton-Verfahren, wenn der Startwert x_0 nahe genug an x^* gewählt wird.*

Beweis. Wenn f zweimal stetig differenzierbar ist, dann ist die Ableitung $F'(x)$ in (5.4) stetig. Zudem folgt aus $f(x^*) = 0$ und $f'(x^*) \neq 0$, dass $F(x^*) = x^*$ und $F'(x^*) = 0$. Aufgrund der Stetigkeit muss es daher ein Intervall $[x^* - \delta, x^* + \delta]$ um x^* geben, auf dem $|F'(x)| \leq \frac{1}{2}$ gilt. Für jedes $x \in [x^* - \delta, x^* + \delta]$ gilt nach dem Mittelwertsatz zudem

$$|F(x) - x^*| = |F(x) - F(x^*)| = F'(c)|x - x^*| \leq \frac{1}{2} \cdot \delta < \delta$$

für ein c zwischen x und x^*. Damit gilt $F(x) \in [x^* - \delta, x^* + \delta]$. Es sind also alle Voraussetzungen erfüllt, um Satz 5.52 auf F (mit $K = \frac{1}{2}$) anzuwenden: wenn der Startwert x_0 im Intervall $[x^* - \delta, x^* + \delta]$ gewählt wird, dann konvergiert das Newton-Verfahren gegen x^*.