Kapitel 4

Funktionen

Neben Folgen und Reihen sind Funktion für die Analysis von zentraler Bedeutung. In diesem und im nächsten Kapitel werden wir uns mit Funktionen und deren Eigenschaften beschäftigen. Dazu werden verschiedenen Begrifflichkeiten, die wir bereits für Folgen und Reihen kennen gelernt haben, auf Funktionen übertragen und die für die Analysis zentrale Klasse der stetigen Funktionen wird definiert.

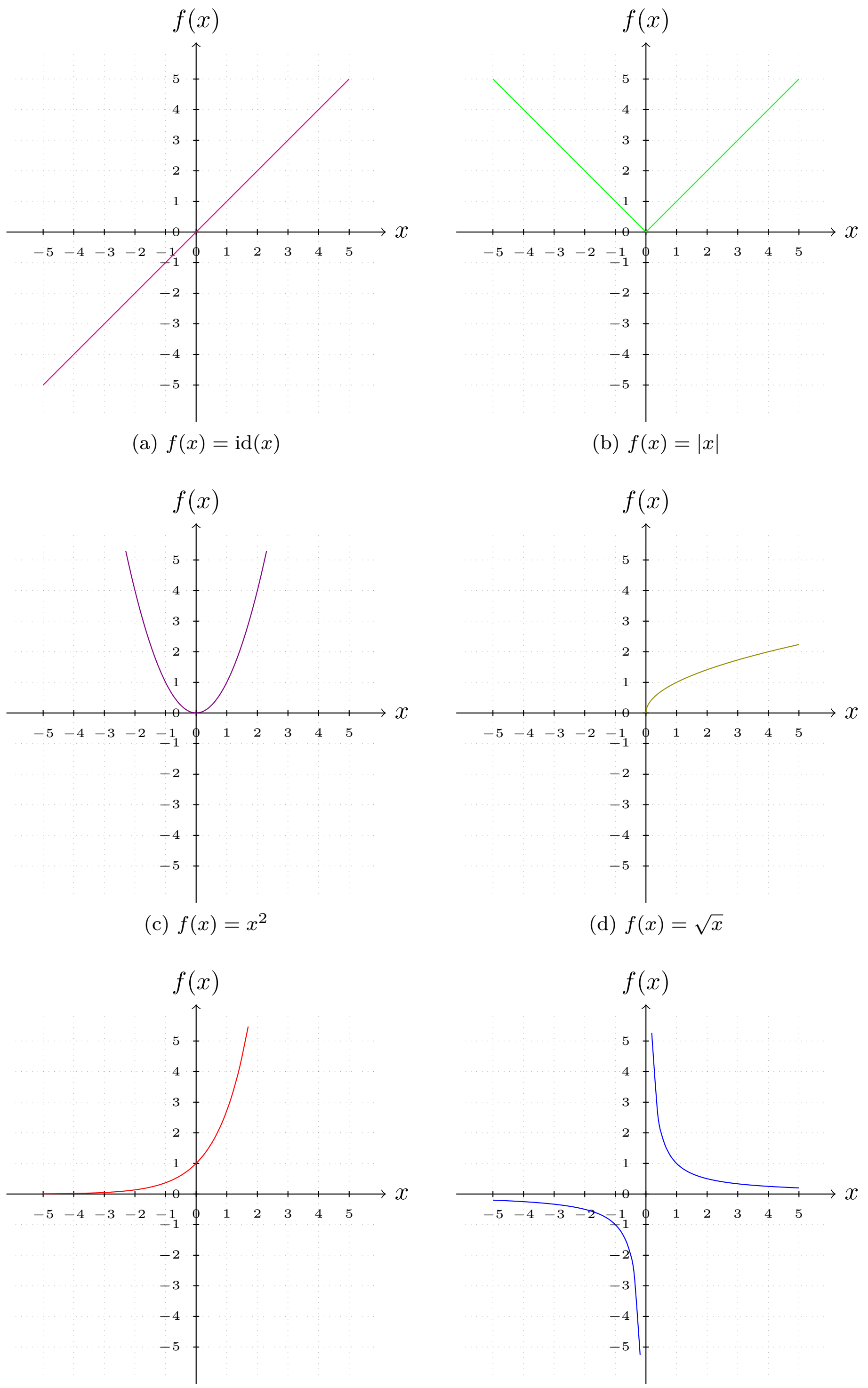
4.1 Grundlegende Definitionen

Definition 4.1 (Funktion)**.** *Seien A und B zwei nichtleere Mengen. Eine* Funktion *f mit* Definitionsbereich *A und* Zielbereich *(oder* Bildbereich*) B ist eine Vorschrift, die jedem Element aus A ein eindeutiges Element aus B zuordnet.*

Wir schreiben $f : A \to B$ und für $a \in A$ ist $f(a) \in B$ das Element, welches Element a zugeordnet wird.
Falls $B = \mathbb{R}$ ist, so sprechen wir von einer reellwertigen Funktion. Wir werden uns im Wesentlichen mit reellwertigen Funktionen beschäftigen. Man kann aber natürlich auch Funktionen über andere Defintionsbereiche bilden. So ordnet zum Beispiel der ASCII-Code den Zahlen 0 bis 127 Zeichen zu.
Falls $A \subseteq \mathbb{R}$ und $B = \mathbb{R}$, so kann man eine Funktion $f : A \to B$ als Graph interpretieren. Dazu definieren wir

$$\Gamma_f := \{(x, y) \mid x \in A \wedge y = f(x)\}$$

als eine Menge von Punkten in der Ebene, die jeweils einem Punkt im Definitionsbereich den durch die Abbildung definierten Punkt im Bildbereich zuordnen. In Abbildung 4.1 werden einige Funktionen und die zugehörigen Graphen gezeigt.



(a) $f(x) = \mathrm{id}(x)$

(b) $f(x) = |x|$

(c) $f(x) = x^2$

(d) $f(x) = \sqrt{x}$

(e) $f(x) = \exp(x)$

(f) $f(x) = \frac{1}{x}$

Abbildung 4.1: Einige Beispielfunktionen und deren Graphen.

Es gibt Funktionen, deren Graph nicht sinnvoll gezeichnet werden kann. Ein Beispiel dafür ist die Dirichlet-Funktion $f : \mathbb{R} \to \mathbb{R}$ mit

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

Die folgenden Eigenschaften von Funktionen klassifizieren diese bzgl. der Struktur der Abbildung.

Definition 4.2.

1. *Eine Funktion $f : A \to B$ heißt* injektiv, *wenn zu jedem $y \in B$ höchstens ein $x \in A$ mit $f(x) = y$ gehört (d. h. $x_1 \neq x_2 \Rightarrow f(x_1) \neq f(x_2)$)*

2. *Eine Funktion $f : A \to B$ heißt* surjektiv, *wenn jedes $y \in B$ als Abbild eines $x \in A$ auftaucht (d. h. $\forall y \in B\, \exists x \in A. f(x) = y$)*

3. *Eine Funktion $f : A \to B$ heißt* bijektiv, *wenn sie injektiv und surjektiv ist.*

Die Eigenschaften haben unter anderem eine Bedeutung bei der Lösung von Gleichungen de Form $f(x) = y$. Falls f injektiv ist, gibt es höchstens eine Lösung, falls f surjektiv ist, mindestens eine Lösung.

Definition 4.3. *Für eine bijektive Funktion $f : A \to B$ definieren wir die* Umkehrfunktion *$f^{-1} : B \to A$ als $f^{-1}(y) = x$ genau dann wenn $f(x) = y$.*

Beispiel 4.1. *$f(x) = x^2$ ist bijektiv auf $A = \mathbb{R}_{\geq 0}$, $B = \mathbb{R}_{\geq 0}$ (d. h. $x, y \geq 0$). Es gilt dann $y = x^2 \Leftrightarrow x = \sqrt{y}$, d. h. $\sqrt{y}$ ist die Umkehrfunktion von x^2. Diese Aussage gilt natürlich* nicht *für $A = \mathbb{R}$!*

Falls f injektiv, nicht aber bijektiv ist, so können wir die Umkehrfunktion ebenfalls bilden. Diese ist dann aber nicht für den Wertebereich B, sondern für den Wertebereich $f(A)$ definiert mit $f(A) = \{y \mid f(x) = y \wedge x \in A\}$. $f(A)$ bezeichnet man auch als das Bild von A unter f. Entsprechend ist A das Urbild von B, welches dem Bild der Umkehrfunktion entspricht. Die Begriffe Bild und Urbild lassen sich auch auf einzelnen Elemente anwenden. So ist $y = f(x)$ das Bild von x (unter f) und x das Urbild von y.
Die beiden folgenden Definitionen beschreiben die Kombination von Funktionen zur Erzeugung neuer, zusammengesetzter Funktionen. Wir werden anschließend untersuchen, welche Eigenschaften der Funktionen sich auf die zusammengesetzten Funktionen übertragen.

Definition 4.4 (Rationale Operationen auf Funktionen)**.** *Seien $f, g : A \to \mathbb{R}$ Funktionen und $c \in \mathbb{R}$. Dann sind die Funktionen*
$f + g : A \to \mathbb{R}$, $cf : A \to \mathbb{R}$, $fg : A \to \mathbb{R}$ definiert durch

$$\begin{array}{lcl} (f+g)(x) & = & f(x) + g(x), \\ (cf)(x) & = & cf(x), \\ (fg)(x) & = & f(x)g(x). \end{array}$$

Sei $A' = \{x \in A \mid g(x) \neq 0\}$, dann ist die Funktion $\frac{f}{g} : A' \to \mathbb{R}$ definiert durch

$$\left(\frac{f}{g}\right)(x) = \frac{f(x)}{g(x)}.$$

Definition 4.5 (Konkatenation von Funktionen)**.** *Seien $f : A \to \mathbb{R}$ und $g : B \to \mathbb{R}$ Funktionen und $f(x) \in B$ für alle $x \in A$. Dann ist die Funktion $g \circ f : A \to \mathbb{R}$ definiert durch $(g \circ f)(x) = g(f(x))$.*

Beispiel 4.2. *Sei $q : \mathbb{R} \to \mathbb{R}$ durch $q(x) = x^2$ definiert und $sqrt : \mathbb{R}_{\geq 0} \to \mathbb{R}$ durch $sqrt(x) = \sqrt{x}$. Dann lässt sich die Funktion $abs : \mathbb{R} \to \mathbb{R}$ schreiben als $(sqrt \circ q)(x) = sqrt(q(x)) = \sqrt{x^2} = |x| = abs(x)$.*

Funktionen sind ein zentrales Hilfsmittel, um den Zusammengang zwischen abhängigen Größen formal zu beschreiben. Sie werden damit in praktisch allen Wissenschaftsbereichen eingesetzt. In der Informatik existiert das klassische Konzept der Funktion, die ein Programmstück beschreibt, das aus den Werten der Eingabeparameter einen Resultatwert berechnet. Die theoretische Informatik beschäftigt sich unter anderem damit, welche Funktionen überhaupt auf einem Rechner, wie er heute genutzt wird, berechenbar sind und wie hoch der Berechnungsaufwand ist. Zur Darstellung des Berechnungsaufwandes wird dieser oft in Abhängigkeit von einer Eingabegröße mit einigen typischen Funktionen verglichen. Dies führt zur sogenannten $O(.)$-Notation, die in den Vorlesungen zur theoretischen Informatik behandelt wird. Wenn der Wert einer Variablen n die Größe der Eingabe beschreibt, so bedeutet ein Aufwand von $(O(n)$, dass der Aufwand höchstens linear wächst und damit bei einer Verdopplung der Eingabegröße maximal ein doppelt so hoher Aufwand entsteht. Bei einem Aufwand in $O(n^2)$ wächst der Aufwand quadratisch, so das bei einer Verdopplung der Eingabegröße der Berechnungsaufwand sich vervierfacht. Neben dem Berechnungsaufwand spielt natürlich auch die Genauigkeit der Berechnungen eine Rolle. Wie wir bereits gesehen haben, werden in einem Rechner Elemente aus $\mathbb{R}$ durch rationale Zahlen approximiert, so dass fast alle Berechnungen streng genommen nur Approximationen sind. Der Approximationsgüte widmet sich die Numerik, die an der Grenze zwischen angewandter Mathematik und Informatik angesiedelt ist.
Wir werden uns in dieser Vorlesung im Wesentlichen mit den mathematischen Aspekten der Funktionen beschäftigen. Da zur Beschreibung realer Vorgänge in fast allen Fällen bestimmte Basisfunktionen eingesetzt werden, widmen sich die folgenden Abschnitte der Vorstellung der wichtigsten Basisfunktionen und der Analyse grundlegender Eigenschaften von Funktionen.

4.2 Polynome und rationale Funktionen

In diesem Abschnitt betrachten wir zwei wichtige Klassen von Funktionen, die durch ihre einfache Analysierbarkeit und die Möglichkeiten mit ihnen andere, komplexere Funktionen zu approximieren, in der Analysis eine große Bedeutung haben.
Wir beginnen mit den Polynomfunktionen. Dazu seien $n \in \mathbb{N}_0$, $a_0, a_1, \ldots, a_n \in \mathbb{R}$, $p : \mathbb{R} \to \mathbb{R}$ mit

$$p(x) = a_n x^n + \cdots + a_1 x + a_0.$$

Das größte $n \in \mathbb{N}_0$ mit $a_n \neq 0$ heißt der Grad des Polynoms.
Polynomfunktionen können multipliziert werden. Seien

$$p(x) = a_n x^n + \cdots + a_1 x + a_0 \text{ und } q(x) = b_m x^m + \cdots + b_1 x + b_0$$

Polynomfunktionen vom Grade n und m. Dann ist

$$h(x) = (p \cdot q)(x) = c_{m+n}x^{m+n} + \cdots + c_1 x + c_0$$

eine Polynomfunktion vom Grad $n + m$ und $c_k = \sum_{\substack{0 \le r \le n \\ 0 \le s \le m \\ r+s=k}} a_r b_s$.

Beispiel 4.3. *Ein Anwendungsgebiet von Polynomen ist die Approximation von Funktionen. Ziel ist es, mithilfe eines Polynoms eines vorgegebenen Grades eine Funktion möglichst gut anzugleichen. Dazu werden oft* Stützstellen *verwendet. Es gibt eine Menge von Punkten (x_i, y_i), die von der zu approximierenden Funktion angenommen werden oder aus einer Messung stammen und nun durch eine Funktion beschrieben werden sollen. Es soll nun ein Polynom $p(x)$ gefunden werden, sodass $p(x_i) = y_i$ für alle vorgegebenen Wertepaare (x_i, y_i).*
Wir betrachten im Folgenden einen einfachen Algorithmus, der zu $K+1$ Stützstellen (x_i, y_i) $(0 \le i \le K)$ ein Polynom $p(x)$ vom Grad $\le K$ erzeugt, sodass $p(x_i) = y_i$. Das Polynom wird schrittweise generiert, indem das folgende Vorgehen gewählt wird. Wir nehmen dazu an, dass $x_i \neq x_j$ für alle $i \neq j, i, j \in \{0, \ldots, K\}$.

$$\begin{aligned} p_0(x) &= y_0 \\ p_{k+1}(x) &= p_k(x) + (y_{k+1} - p_k(x_{k+1})) \prod_{j=0}^{k} \tfrac{x - x_j}{x_{k+1} - x_j} \quad \textit{für } 0 \le k < K. \end{aligned}$$

$p_K(x)$ ist dann die gesuchte Polynomfunktion vom Grad $(\le K)$ und es gilt $p_K(x_i) = y_i$ für alle $i \in \{0, \ldots, K\}$. Man kann sogar zeigen, dass $p_K(x)$ eindeutig ist, d. h. es gibt keine Polynomfunktion vom gleichen oder niedrigeren Grad, die alle Punkte exakt rekonstruiert.
Wir betrachten als einfaches Beispiel die Punktemenge $\{(0,1), (2,3), (3,0)\}$, die zur Konstruktion der folgenden Polynomfunktion führt.

$$\begin{aligned} p_0(x) &= 1 \\ p_1(x) &= p_0(x) + (y_1 - p_0(x_1)) \tfrac{x - x_0}{x_1 - x_0} \\ &= 1 + (3-1)\tfrac{x-0}{2-0} \;=\; x + 1 \\ p_2(x) &= p_1(x) + (y_2 - p_1(x_2)) \tfrac{x - x_0}{x_2 - x_0} \tfrac{x - x_1}{x_2 - x_1} \\ &= 1 + x + (0-4)\tfrac{x-0}{3-0}\tfrac{x-2}{3-2} \\ &= -\tfrac{4}{3}x^2 + \tfrac{11}{3}x + 1 \end{aligned}$$

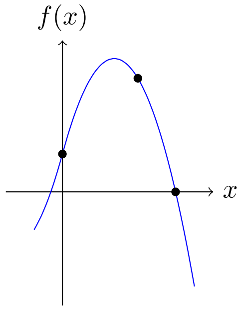


Abbildung 4.2: Darstellung der Polynomfunktion und der vorgegebenen Stützstellen

Die Beispielfunktion gibt die vorgegebenen Punkte exakt wieder und hat einen glatten Verlauf. Es sollte allerdings erwähnt werden, dass dies nicht für alle Funktionen der Fall ist und in vielen Fällen die Hinzunahme neuer Punkte zu einer Polynomfunktion führt, die sehr abrupte Änderungen aufweist und die vorgegeben Funktion nur unzureichend approximiert. Es werden deshalb in der Regel komplexere Verfahren zur Funktionsapproximation mit Hilfe von Polynomfunktionen verwendet, auf die wir nicht weiter eingehen.

Beispiel 4.4. *Das vorherige Beispiel Polynomapproximation kann auch in ganz anderen Bereichen eingesetzt werden. Zum Beispiel kann man den Ansatz nutzen, um verteilte Geheimnisse zu verwalten. Ein verteiltes Geheimnis ist eine Information, die nur gewonnen werden kann, wenn Personen aus einer Gruppe das Geheimnis gemeinsam lösen. Ein Anwendungsbeispiel ist die Sicherung geheimer Informationen, sodass diese nur von mehreren verantwortlichen Personen gemeinsam eingesehen oder geändert werden kann, um zu verhindern, dass Einzelne Missbrauch mit den Informationen treiben.*
Wir nehmen an, dass unser Geheimnis aus einer Zahl r *besteht, die zum Beispiel einen Schlüssel zur Entschlüsselung darstellt. Weiterhin soll das Geheimnis von* n *Personen verwaltet werden, sodass nur die* n *Personen zusammen* r *rekonstruieren können, nicht aber eine Teilmenge der Personen.*
Wir konstruieren ein Polynom

$$p(x) = a_{n-1}x^{n-1} + \ldots + a_1 x + a_0$$

mit $a_0 = r$ *und zufällig gewählten Koeffizienten* a_i. *Damit gilt natürlich* $p(0) = r$. *Nun verteilen wir* n *Wertepaare* $(x_i, p(x_i))$ *mit* $x_i \neq x_j$ *und* $x_i \neq 0$. *Aus den* n *Paaren kann die Polynomfunktion eindeutig rekonstruiert werden. Eine Teilmenge reicht nicht aus, um Informationen über* r *abzuleiten.*
Man kann das Verfahren so erweitern, dass eine beliebige Teilmenge von $m < n$ *Personen ausreicht, um das Geheimnis zu lösen (überlegen wie!).*

Neben der Multiplikation von Polynomen, kann man auch deren Division betrachten. Dazu betrachten wir den folgenden Satz.

Satz 4.6. *Seien* $p(x)$ *und* $q(x)$ *Polynome vom Grad* n *und* m *mit* $m \leq n$. *Dann gibt es Polynome* $s(x)$ *und* $r(x)$, *so dass*

$$p(x) = s(x)q(x) + r(x).$$

Der Grad von $s(x)$ entspricht der Differenz Grad p - Grad q und Grad $r <$ Grad q.

Beweis: Sei $p(x) = a_n x^n + \ldots + a_1 x + a_0$ und $q(x) = b_m x^m + \ldots + b_1 x + b_0$. Subtrahiert man von $p(x)$ das Polynom

$$\frac{a_n}{b_m} x^{n-m} \cdot q(x)$$

so erhält man ein Polynom $p_1(x)$ mit $n_1 =$ Grad $p_1(x) <$ Grad $p(x) = n$.
Falls $n_1 < m$, dann sei $r(x) = p_1(x)$ und $s(x) = \frac{a_n}{b_m} x^{m \cdot n}$ und die obige Darstellung wurde erreicht. Ansonsten wird mit $p_1(x)$ und $q(x)$ wie für $p(x)$ und $q(x)$ beschrieben fortgefahren. Dies Schritte werden so lange iteriert bis ein Polynom $p_i(x)$ mit Grad $p_i(x) <$ Grad $q(x)$ erzeugt wurde. Ein solches Polynom entsteht, da sich der Grad beim Übergang von $p_{i+1}(x)$ auf $p_i(x)$ jeweils um 1 reduziert.
Nun muss noch die Eindeutigkeit der Polynome $s(x)$ und $r(x)$ gezeigt werden. Sei $p(x) = s'(x)q(x) + r'(x)$ mit $s'(x) \neq s(x)$ eine weitere Darstellung, dann folgt $(s'(x) - s(x))q(x) = r(x) - r'(x)$ und damit auch Grad $(s(x) - s'(x))q(x) =$ Grad $(r(x) - r'(x)) <$ Grad $g(x) <$ Grad $p(x)$. Dies ist ein Widerspruch, da Grad $(p(x) + q(x)) \leq \max($ Grad $p(x), q(x))$. $\square$

Falls $r(x)$ im vorherigen Satz gleich 0 ist, so heißt $q(x)$ Teiler von $p(x)$. Falls es kein Polynom vom Grad ≥ 1 gibt, das $p(x)$ und $q(x)$ teilt, so heißen $p(x)$ und $q(x)$ teilerfremd.
Polynome können im Prinzip wie andere Zahlen dividiert werden, wie das folgende Beispiel zeigt. Diesen Vorgang nennt man Polynomdivision.

Beispiel 4.5. *Wir berechnen $(3x^4 + x^3 - 2x) : (x^2 + 1)$ nach folgendem Schema:*

$$\begin{array}{ccccccccc}
(& 3x^4 & +x^3 & & -2x & &) : (x^2+1) & = & 3x^2 + x - 3 \\
 & -(3x^4 & & +3x^2) & & & & & \\
\hline
 & & x^3 & -3x^2 & -2x & & & & \\
 & & -(x^3 & & +x) & & & & \\
\hline
 & & & -3x^2 & -3x & & & & \\
 & & & -(-3x^2 & & -3) & & & \\
\hline
 & & & & -3x & +3 & & &
\end{array}$$

Die Division wird abgebrochen, da Grad $(-3x + 3) <$ Grad $(x^2 + 1)$. Damit ist $s(x) = 3x^2 + x - 3$ und $r(x) = -3x + 3$.

Unter einer Nullstelle eines Polynoms $p(x)$ versteht man einen Wert $x_1 \in \mathbb{R}$, so dass $p(x_1) = 0$. Aus Satz 4.6 folgt folgendes Korollar.

Korollar 4.7. *Ein Polynom $p(x)$ lässt sich genau dann ohne Rest durch $q(x) = x - x_1$ teilen ($x_1 \in \mathbb{R}$), wenn x_1 eine Nullstelle von $p(x)$ ist.*

$q(x) = x - x_1$ bezeichnet man auch als einen Linearfaktor. Da im Korollar $p(x)$ als $(x - x_1)q(x)$ dargestellt wird und Grad $q(x) <$ Grad $p(x)$, kann ein Polynom vom Grad n höchstens n Nullstellen haben.

Beispiel 4.6. *Wir betrachten das Polynom* $p(x) = 2x^3 + 3x^2 - 5x - 6$. *Durch* Ausprobieren *kann man feststellen, dass* $p(-1) = 0$, *so dass*

$$\begin{array}{lllll}
(\quad 2x^3 & +3x^2 & -5x & -6 &) : (x+1) \quad = \quad 2x^2 + x - 6 \\
-(2x^3 & +2x^2) & & & \\
\hline
& x^2 & -5x & & \\
& -(x^2 & +x) & & \\
\hline
& & -6x & -6 & \\
& & -(-6x & -6) & \\
\hline
& & & 0 &
\end{array}$$

Die quadratische Funktion $2x^2 + x - 6$ *hat die Nullstellen* $x_1 = \frac{3}{2}$ *und* $x_2 = -2$.

Generall kann man durch die Division eines Polynoms durch seine Nullstellen den Grad des Polynom reduzieren. Dazu muss man allerdings die Nullstellen kennen. Wir werden uns in Kapitel 6 kurz mit Methoden zur Berechnung oder besser Approximation von Nullstellen für allgemeine Funktionen beschäftigen.
Aus Polynomfunktionen können die so genannten rationalen Funktionen generiert werden.
Seien $p(x) = a_n x^n + \cdots + a_1 x + a_0$ und $q(x) = b_m x^m + \cdots + b_1 x + b_0$ Polynome und $A = \{x \mid q(x) \neq 0\}$. Dann ist die rationale Funktion $r : A \to \mathbb{R}$ definiert durch

$$r(x) = \left(\frac{p}{q}\right)(x) = \frac{p(x)}{q(x)}.$$

4.3 Beschränkte und monotone Funktionen

Die folgende Definition definiert Beschränktheit, die wir bereits für Mengen, Folgen und Reihen kennen gelernt haben, für Funktionen.

Definition 4.8. *Eine Funktion* $f : A \to \mathbb{R}$ *heißt* beschränkt, *wenn* $|f(x)| \leq K$ *für* $K \in \mathbb{R}$ *und alle* $x \in A$.

Die Beschränktheit einer Funktion $f : A \to \mathbb{R}$ ist äquivalent dazu, dass die Menge $f(A)$ beschränkt ist.

Definition 4.9. *Unter einem* kompakten Intervall *versteht man ein abgeschlossenes und beschränktes Intervall* $[a, b] \subset \mathbb{R}$.

Funktionen haben spezielle Eigenschaften auf kompakten Intervallen, wie wir im Laufe der nächsten Abschnitte sehen werden.

Definition 4.10. *Sei* $A \subseteq \mathbb{R}$ *und* $f : A \to \mathbb{R}$ *eine Funktion,*

$$f \text{ heißt} \left\{ \begin{array}{c} \textsf{monoton wachsend} \\ \textsf{streng monoton wachsend} \\ \textsf{monoton fallend} \\ \textsf{streng monoton fallend} \end{array} \right\} \text{ falls } \left\{ \begin{array}{c} f(x) \leq f(x') \\ f(x) < f(x') \\ f(x) \geq f(x') \\ f(x) > f(x') \end{array} \right\}$$

für $x, x' \in A$ *mit* $x < x'$.

Beispiel 4.7.

1. *Die konstante Funktion $f(x) = c$ ist monoton wachsend und fallend, da $c \leq c$ gilt.*

2. *Die Identitätsfunktion ist (streng) monoton wachsend, da aus $x < x'$ offensichtlich $f(x) = x < f(x') = x'$ folgt.*

3. *Die Funktion $f(x) = x^2$ ist streng monoton wachsend auf dem Intervall $[0, \infty)$ und streng monoton fallend auf dem Intervall $(-\infty, 0]$.*
   Beweis: *Sei $0 \leq x < x'$ dann gilt*

$$x^2 = x \cdot x \overset{S.2.7.iii)}{<} x \cdot x' < x' \cdot x' = (x')^2$$

Sei $x < x' \leq 0$, dann gilt $x^2 = x \cdot x \overset{S.\ 2.7.v)}{>} x \cdot x' > x' \cdot x' = (x')^2$. □

Satz 4.11. *Sei $A \subseteq \mathbb{R}$ und $f : A \to \mathbb{R}$ eine Funktion. Ist f streng monoton, so ist f injektiv und die Umkehrfunktion $f^{-1} : f(A) \to A$ ist ebenfalls streng monoton (im gleichen Sinne).*

Beweis: Aus der strengen Monotonie folgt $x \neq x' \Rightarrow f(x) \neq f(x')$ und damit ist f injektiv und die Umkehrfunktion $g(x) = f^{-1}(x)$ existiert.
Wir betrachten den Fall, dass f streng monoton wachsend ist. Seien $y_1, y_2 \in f(A)$ mit $y_1 < y_2$ gegeben. Zu zeigen ist $g(y_1) < g(y_2)$.
Wäre $g(y_1) = g(y_2)$, dann wäre $y_1 = f(g(y_1)) = f(g(y_2)) = y_2$, da $f \circ g = f \circ f^{-1} = id$. Analog gilt, falls $g(y_1) > g(y_2)$, dann wäre $y_1 = f(g(y_1)) > f(g(y_2)) = y_2$ was im Widerspruch zu $y_1 < y_2$ steht. Damit bleibt nur $g(y_1) < g(y_2)$. □

Beispiel 4.8. *$f : \mathbb{R}_{>0} \to \mathbb{R}$, $f(x) = x^k$, $k \in \mathbb{N}$ und $k \geq 2$. f ist streng monoton wachsend und bildet $\mathbb{R}_{>0}$ bijektiv auf $\mathbb{R}_{>0}$ ab. Die Umkehrfunktion $f^{-1} : \mathbb{R}_{>0} \to \mathbb{R}$, $f(x) = \sqrt[k]{x}$ ist streng monoton wachsend.*

4.4 Grenzwerte von Funktionen

Im Kapitel 3 (Definition 3.18) hatten wir den Begriff des Häufungspunkts für Folgen definiert, den wir nun auf Mengen und Funktionen ausweiten.

Definition 4.12. *Sei $A \subseteq \mathbb{R}$ und $a \in \mathbb{R}$*

1. a heißt Berührungspunkt *von A, falls in jeder ε-Umgebung von a, d. h. $U_\varepsilon(a) = (a - \varepsilon, a + \varepsilon), \varepsilon > 0$, mindestens ein Punkt von A liegt.*

2. a heißt Häufungspunkt*, falls in jeder ε-Umgebung von a unendlich viele Punkte von A liegen.*

a ist ein Berührungspunkt von A, falls $a \in A$ oder eine Folge $x_n \in A$ mit $\lim\limits_{n\to\infty} x_n = a$ existiert. Im zweiten Fall ist a Häufungspunkt, wenn für diese Folge zusätzlich $x_n \in A \setminus \{a\}$ gilt.

Definition 4.13. *Sei $f : A \to \mathbb{R}$, $A \subseteq \mathbb{R}$ und $a \in \mathbb{R}$ ein Berührungspunkt von A. Man definiert dann $\lim\limits_{x\to a} f(x) = c$, falls für jede Folge $(x_n)_{n\in\mathbb{N}}$, $x_n \in A$ mit $\lim\limits_{n\to\infty} x_n = a$, $\lim\limits_{n\to\infty} f(x_n) = c$ gilt.*

Einige weitere Bezeichnungen:

1. $\lim\limits_{x \searrow a} f(x) = c$ bedeutet: a ist Berührungspunkt von $A \cap (a, \infty)$ und für jede Folge $(x_n)_{n\in\mathbb{N}}$ mit $x_n \in A$, $x_n > a$ und $\lim\limits_{n\to\infty} x_n = a$ gilt $\lim\limits_{n\to\infty} f(x_n) = c$
2. $\lim\limits_{x \nearrow a} f(x) = c$ bedeutet: a ist Berührungspunkt von $A \cap (-\infty, a)$ und für jede Folge $(x_n)_{n\in\mathbb{N}}$ mit $x_n \in A$, $x_n < a$ und $\lim\limits_{n\to\infty} x_n = a$ gilt $\lim\limits_{n\to\infty} f(x_n) = c$
3. $\lim\limits_{x\to\infty} f(x) = c$ bedeutet: A ist nach oben unbeschränkt und für jede Folge $(x_n)_{n\in\mathbb{N}}$ mit $\lim\limits_{n\to\infty} x_n = \infty$ gilt $\lim\limits_{n\to\infty} f(x_n) = c$.
4. $\lim\limits_{x\to-\infty} f(x) = c$ bedeutet: A ist nach unten unbeschränkt und für jede Folge $(x_n)_{n\in\mathbb{N}}$ mit $\lim\limits_{n\to\infty} x_n = -\infty$ gilt $\lim\limits_{n\to\infty} f(x_n) = c$.

Es gilt übrigens der folgende Zusammenhang.

$$\lim_{x\to a} f(x) = f(a) \;\Leftrightarrow \lim_{x\searrow a} f(x) = \lim_{x\nearrow a} f(x) = f(a)$$

Die Äquivalenz gilt natürlich nur, wenn die Grenzwerte $\lim\limits_{x\searrow a} f(x)$ und $\lim\limits_{x\nearrow a} f(x)$ definiert sind. Falls a ein Berührungspunkt von A ist und für alle $x \in A$ $x \geq a$ gilt, so ist existiert $\lim\limits_{x\nearrow a} f(x)$ nicht, sehr wohl aber $\lim\limits_{x\searrow a} f(x)$ und $\lim\limits_{x\to a} f(x)$. Falls a ein Berührungspunkt von A ist und für alle $x \in A$ $x \leq a$ gilt, so ist existiert $\lim\limits_{x\searrow a} f(x)$ nicht, $\lim\limits_{x\nearrow a} f(x)$ und $\lim\limits_{x\to a} f(x)$ existieren und sind dann natürlich identisch.

Beispiel 4.9.

1. *Es gilt* $\lim\limits_{x\searrow 1} \lfloor x \rfloor = 1$ *und* $\lim\limits_{x\nearrow 1} \lfloor x \rfloor = 0$. $\lim\limits_{x\to 1} \lfloor x \rfloor$ *existiert nicht!*
2. *Es gilt* $\lim\limits_{x\nearrow 0} |x| = 0$ *und* $\lim\limits_{x\searrow 0} |x| = 0$ *und damit auch* $\lim\limits_{x\to 0} |x| = 0$.
3. *Sei* $f : \mathbb{R} \setminus \{a\} \to \mathbb{R}$ *definiert als* $f(x) = \frac{x^n - a^n}{x - a}$. *Dann ist* $\lim\limits_{x\to a} f(x) = n \cdot a^{n-1}$.

Beweis: *a ist ein Häufungspunkt von $\mathbb{R}\setminus\{a\}$. Für $a \neq x$ gilt $f(x) = x^{n-1} + x^{n-2}a + \cdots + xa^{n-2} + a^{n-1}$, da $(x^{n-1} + x^{n-2}a + \cdots + xa^{n-2} + a^{n-1}) \cdot (x - a) = x^n - a^n$. Definieren wir $g : \mathbb{R} \to \mathbb{R}$ als $g(x) = x^{n-1} + x^{n-2}a + \cdots + xa^{n-2} + a^{n-1}$ dann ist $f(x) = g(x)$ für alle $x \in \mathbb{R} \setminus \{a\}$ und $\lim\limits_{x\to a} g(x) = g(a) = na^{n-1}$* □

4.5 Stetige Funktionen

Ein weiterer zentraler Begriff der Analysis ist die Stetigkeit von Funktionen, die im Prinzip aussagte, dass Funktionen ein „gutmütiges“ Verhalten in einem Intervall oder in der Umgebung eines Punktes zeigen.

Definition 4.14 (Stetigkeit)**.** *Sei $f : A \to \mathbb{R}$ eine Funktion und $a \in A$. Die Funktion f heißt* stetig im Punkt *a, falls $\lim\limits_{x\to a} f(x) = f(a)$.*

f heißt stetig *(in A), falls f in jedem Punkt aus A stetig ist.*

Um Stetigkeit in einem Punkt nachzuprüfen wird $b = f(a)$ untersucht. Es wird eine Umgebung V von b vorgegeben und geprüft, ob eine Umgebung U von a existiert, sodass $f(y) \in V$ für $y \in U$ (siehe Abbildung 4.3). Wir werden diese Darstellung der Stetigkeit später noch formalisieren.
Es ist manchmal hilfreich zum Nachweis der Stetigkeit einer Funktion $f(x)$ an einer Stelle $a \in A$, die einseitigen Grenzwerte $\lim_{x \searrow a} f(x)$ und $\lim_{x \nearrow a} f(x)$ zu untersuchen. Falls $\lim_{x \searrow a} f(x) = \lim_{x \nearrow a} f(x) = f(a)$, dann ist f in a stetig. Falls $\lim_{x \searrow a} f(x) \neq \lim_{x \nearrow a} f(x)$, dann ist f in a nicht stetig.

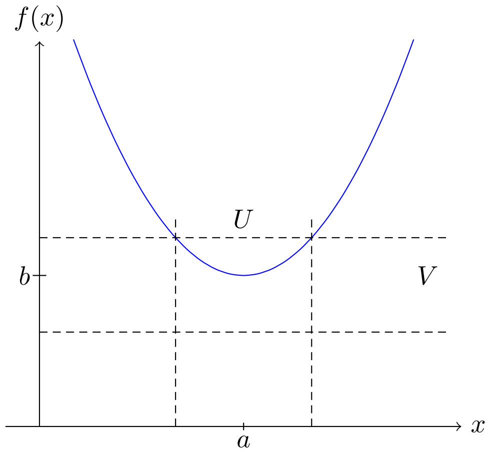


Abbildung 4.3: Skizze der Darstellung von Stetigkeit über die Umgebungen V und U.

Beispiel 4.10.

1. *Die konstante Funktion ist überall stetig.*

2. *Die Exponentialfunktion* $\exp(x)$ *ist in jedem Punkt stetig.*
   Beweis: *Sei* $a \in \mathbb{R}$.
   Zu zeigen: $\lim\limits_{x \to a} \exp(x) = \exp(a)$.
   Sei $(x_n)_{n \in \mathbb{N}}$ *eine Folge mit* $\lim\limits_{n \to \infty} x_n = a$. *Dann gilt* $\lim\limits_{n \to \infty} (x_n - a) = 0$, *also* $\lim\limits_{n \to \infty} \exp(x_n - a) = 1$.
   Damit folgt $\lim\limits_{n \to \infty} \exp(x_n) \overset{\text{Satz } 3.37.i)}{=} \lim\limits_{n \to \infty} (\exp(a) \cdot \exp(x_n - a)) =$
   $\exp(a) \cdot \lim\limits_{n \to \infty} \exp(x_n - a) = \exp(a)$ □

3. *Die Funktion* $f : \mathbb{R} \to \mathbb{R}$

$$f(x) = \begin{cases} 1 & \textit{für} \quad x \geq 0 \\ 0 & \textit{für} \quad x < 0 \end{cases}$$

   ist in jedem $a \in \mathbb{R} \setminus \{0\}$ *stetig und ist nicht stetig in* 0, *da* $\lim_{x \searrow 0} f(x) = 1 \neq \lim_{x \nearrow 0} f(x) = 0$.

4. Die Dirichlet-Funktion

$$f(x) = \begin{cases} 1 & \text{falls } x \in \mathbb{Q} \\ 0 & \text{falls } x \in \mathbb{R} \setminus \mathbb{Q} \end{cases}$$

ist in keinem Punkt stetig, da in jeder ε-Umgebung ($\varepsilon > 0$) von $a \in \mathbb{R}$ unendliche viele $x \in \mathbb{R} \setminus \mathbb{Q}$ und unendlich viele $y \in \mathbb{Q}$ liegen.

Im Prinzip kann man stetige Funktionen durchgängig zeichnen. Diese Eigenschaft bleibt erhalten, wenn man stetige Funktionen kombiniert, wie die folgenden Sätze zeigen.

Satz 4.15 (Operationen auf stetigen Funktionen)**.** *Seien $f, g : A \to \mathbb{R}$ Funktionen, die in $a \in A$ stetig sind und sei $c \in \mathbb{R}$. Dann sind auch die Funktionen*

i) $f + g : A \to \mathbb{R}$

ii) $cf : A \to \mathbb{R}$

iii) $f \cdot g : A \to \mathbb{R}$

im Punkt a stetig.
Ist $g(a) \neq 0$, so ist auch die Funktion

iv) $\frac{f}{g} : A' \to \mathbb{R}$

in a stetig. Dabei ist $A' = \{x \in A \mid g(x) \neq 0\}$.

Beweis: Sei $(x_n)_{n\in\mathbb{N}}$ eine Folge in A (bzw. A') und $\lim\limits_{n\to\infty} x_n = a$. Es ist zu zeigen:

$$\begin{aligned} \lim_{n\to\infty} (f+g)(x_n) &= (f+g)(a) \\ \lim_{n\to\infty} (cf)(x_n) &= (cf)(a) \\ \lim_{n\to\infty} (fg)(x_n) &= (fg)(a) \\ \lim_{n\to\infty} \left(\frac{f}{g}\right)(x_n) &= \left(\frac{f}{g}\right)(a) \end{aligned}$$

Nach Voraussetzung ist $\lim\limits_{n\to\infty} f(x_n) = f(a)$ und $\lim\limits_{n\to\infty} g(x_n) = g(a)$. Die Behauptung folgt aus den Rechenregeln für Folgen (siehe Satz 3.8). □

Korollar 4.16. *Jede rationale Funktion ist in ihrem Definitionsbereich stetig.*

Beweis: Wiederholte Anwendung von Satz 4.15 auf die Identitätsfunktion $id_{\mathbb{R}} : \mathbb{R} \to \mathbb{R}$ mit $id(x) = x$. □

Satz 4.17 (Komposition stetiger Funktionen)**.** *Sei $f : A \to \mathbb{R}$ und $g : B \to \mathbb{R}$ Funktionen mit $f(A) \subseteq B$. Die Funktion f sei in $a \in A$ und g in $b = f(a) \in B$ stetig. Dann ist die Funktion $g \circ f : A \to \mathbb{R}$ in a stetig.*

Beweis: Sei $(x_n)_{n\in\mathbb{N}}$ eine Folge in A mit $\lim\limits_{n\to\infty} x_n = a$. Da f stetig in a ist, gilt $\lim\limits_{n\to\infty} f(x_n) = f(a)$. Nach Voraussetzung ist $y_n = f(x_n) \in B$ und $\lim\limits_{n\to\infty} y_n = b$. Da g in b stetig ist, gilt auch $\lim\limits_{n\to\infty} g(y_n) = g(b)$. Deshalb folgt $\lim\limits_{n\to\infty} (g \circ f)(x_n) = \lim\limits_{n\to\infty} g(f(x_n)) = \lim\limits_{n\to\infty} g(y_n) = g(b) = g(f(a)) = (g \circ f)(a)$ □

Beispiel 4.11. *Sei $f(x) = \exp(x)$, $g(x) = x^2$. f, g sind stetig in $\mathbb{R}$, nach Satz 4.17 ist damit auch $h(x) = \exp(x^2)$ stetig in $\mathbb{R}$.*

Funktionen, die in ihrem Definitionsbereich stetig sind, erlauben einige interessante Aussagen, die wir in den folgenden Sätzen herleiten.

Satz 4.18 (Zwischenwertsatz). *Sei $f : [a, b] \to \mathbb{R}$ eine stetige Funktion mit $f(a) < 0$ und $f(b) > 0$ (bzw. $f(a) > 0$ und $f(b) < 0$). Dann existiert ein $c \in (a, b)$ mit $f(c) = 0$.*

Beweis: Der Beweis erfolgt über die Intervallhalbierungsmethode. Wir betrachten $f(a) < 0$, $f(b) > 0$. Der andere Fall $f(a) > 0$ und $f(b) < 0$ ist analog zu beweisen. Wir definieren zunächst eine Intervallschachtelung mit

1. $I_{n+1} = [a_{n+1}, b_{n+1}] \subset [a_n, b_n] = I_n$
2. $|I_{n+1}| = \frac{1}{2}|I_n|$
3. $f(a_n) < 0 \wedge f(b_n) > 0$

induktiv unter Nutzung der folgenden Schritte.
Induktionsanfang $[a_1, b_1] = [a, b]$.
Induktionsschritt

$$I_{n+1} = [a_{n+1}, b_{n+1}] = \begin{cases} [a_n, m] & \text{falls} \quad f(m) > 0 \\ [m, b_n] & \text{falls} \quad f(m) < 0 \\ \star & \text{falls} \quad f(m) = 0 \end{cases}$$

mit $m = \frac{a_n + b_n}{2}$, dem Mittelpunkt des Intervalls $[a_n, b_n]$. Es ist leicht nachzuprüfen, dass die obigen Bedingungen 1.-3. für die Intervalle gelten, solange wir nicht den Fall $\star$ erreicht haben. Falls $\star$ erreicht wird, so bricht die Intervallschachtelung ab, da $f(m) = 0$ und wir damit die Nullstelle bestimmt haben.
Falls das Verfahren nicht abbricht, gibt es ein eindeutiges $x \in I_n$ für alle $n \in \mathbb{N}$, da eine Intervallschachtelung vorliegt.
Es gilt ferner:

$$\lim_{n\to\infty} f(a_n) = \lim_{n\to\infty} f(b_n) = x$$

da f stetig in $[a, b]$ ist und $f(a_n) < 0, f(b_n) > 0$ nach Konstruktion. Damit folgt auch $x = 0$. □

Der Satz gilt nicht für $f : [a, b] \subset \mathbb{Q} \longrightarrow \mathbb{R}$. Wie am Beispiel $f(x) = x^2 - 2$ mit der Nullstelle $\sqrt{2} \notin \mathbb{Q}$ leicht sehen kann.
Das folgende Korollar lässt sich leicht aus dem Zwischenwertsatz herleiten.

Korollar 4.19. *Sei $f : [a, b] \to \mathbb{R}$ eine stetige Funktion und $y \in \mathbb{R}$ mit $f(a) < y < f(b)$ (bzw. $f(a) > y > f(b)$). Dann existiert ein $c \in (a, b)$ mit $f(c) = y$.*

Beweis: Sei $f(a) < y < f(b)$. Definiere $g : [a, b] \to \mathbb{R}$ mit $g(x) = f(x) - y$. g ist stetig und $g(a) = f(a) - y < 0$, $g(b) = f(b) - y > 0$, sodass nach Satz 4.18 ein c mit $g(c) = 0 \Rightarrow g(c) = f(c) - y = 0 \Rightarrow f(c) = y$ existiert.
Der Beweis für den Fall $f(a) > y > f(b)$ ist analog zu führen. □

Korollar 4.20. *Sei $I \subseteq \mathbb{R}$ ein Intervall und $f : I \to \mathbb{R}$ eine stetige Funktion. Dann ist auch $J = f(I) \subseteq \mathbb{R}$ ein Intervall.*

Beweis: Wir setzten $q = \sup(f(I)) \in \mathbb{R} \cup \{+\infty\}$ und $p = \inf(f(I)) \in \mathbb{R} \cup \{-\infty\}$. Zunächst zeigen wir $(p, q) \subseteq f(I)$.
Sei dazu $y \in \mathbb{R}$ beliebig gewählt, sodass $p < y < q$. Nach Definition von p und q gibt es dann $a, b \in I$ mit $f(a) < y < f(b)$. Nach Korollar 4.19 existiert $x \in I$ mit $f(x) = y$. Also ist $y \in f(I)$. Damit ist $(p, q) \subseteq f(I)$ bewiesen und $f(I)$ muss eines der folgenden vier Intervalle sein: $[p, q], (p, q], [p, q), (p, q)$. □

Der folgende Satz betrachtet stetige Funktionen auf kompakten Intervallen (d. h. heißt abgeschlossenen und beschränkten Intervallen; siehe Definition 4.9).

Satz 4.21 (Stetige Funktionen auf kompakten Intervallen)**.** *Jede in einem kompakten Intervall stetige Funktion $f : [a, b] \to \mathbb{R}$ ist beschränkt und nimmt ihr Minimum und Maximum an.*
D. h. es existiert ein $c \in [a, b]$, sodass $f(c) = \sup\{f(x) \mid x \in [a, b]\}$ und ein $d \in [a, b]$, sodass $f(d) = \inf\{f(x) \mid x \in [a, b]\}$.

Beweis: Wir führen den Beweis für das Maximum: Sei $y = \sup\{f(x) \mid x \in [a, b]\} \in \mathbb{R} \cup \{\infty\}$. Dann existiert eine Folge $x_n \in [a, b]$ $(n \in \mathbb{N})$, sodass $\lim\limits_{n\to\infty} f(x_n) = y$, da die Funktion f stetig ist. Da die Folge (x_n) beschränkt ist (das Intervall aus dem die Folge gewählt wird ist beschränkt), besitzt sie nach den Satz von Bolzano-Weierstraß (Satz 3.17) eine konvergente Teilfolge $(x_{n_k})_{k\in\mathbb{N}}$ mit $\lim\limits_{k\to\infty} x_{n_k} = c \in [a, b]$.
Aus der Stetigkeit folgt nun $f(c) = \lim\limits_{k\to\infty} f(x_{n_k}) = y$ und damit insbesondere auch $y = f(c) \in \mathbb{R}$. Damit ist f nach oben beschränkt und nimmt in c das Maximum an. Der Beweis für das Minimum ist analog zu führen. □

Der vorherige Satz gilt nicht für offene oder halboffene Intervalle. So ist zum Beispiel die Funktion $f(x) = \frac{1}{x}$ im Intervall $(0, 1)$ stetig, aber unbeschränkt. Die Funktion $f(x) = x$ ist im Intervall $(0, 1)$ stetig und beschränkt, nimmt aber weder das Supremum 1 noch das Infimum 0 an.
Im folgenden Satz beschreibt die Stetigkeit einer Funktion mit Hilfe von zwei Umgebungen, wie wir es zu Beginn des Abschnitts schon intuitiv in Abbildung 4.3 getan haben.

Satz 4.22 ($\varepsilon - \delta$ - Definition von Stetigkeit)**.** *Sei $A \subseteq \mathbb{R}$ und $f : A \to \mathbb{R}$ eine Funktion. f ist genau dann im Punkt $a \in A$ stetig, wenn gilt:*
Zu jedem $\varepsilon > 0$ existiert ein $\delta > 0$, sodass $|f(x) - f(a)| < \varepsilon$ für alle $x \in A$ mit $|x - a| < \delta$.

Beweis: Der Beweis wird in zwei Schritten geführt, indem in jedem Schritt eine Richtung der „genau dann“ Beziehung bewiesen wird.

ε, δ **existieren** $\Rightarrow$ **Stetigkeit:** Es gebe zu jedem $\varepsilon > 0$ ein $\delta > 0$, sodass $|f(x) - f(a)| < \varepsilon$ für alle $x \in A$ mit $|x - a| < \delta$.

Es ist zu zeigen, dass für jede Folge $(x_n)_{n\in\mathbb{N}}$ mit $x_n \in A$ und $\lim\limits_{n\to\infty} x_n = a$ gilt $\lim\limits_{n\to\infty} f(x_n) = f(a)$.

Sei $\varepsilon > 0$ vorgegeben und $\delta > 0$ gemäß der Voraussetzung gewählt. Da $\lim\limits_{n\to\infty} x_n = a$ existiert ein $n_0 \in \mathbb{N}$, sodass $|x_n - a| < \delta$ für alle $n \geq n_0$. Nach Voraussetzung ist damit $|f(x_n) - f(a)| < \varepsilon$ für alle $n \geq n_0$. Also gilt $\lim\limits_{n\to\infty} f(x_n) = f(a)$ und f ist stetig. Man kann dies auch in Abbildung

4.3 interpretieren. Zu einer Umgebung U (definiert durch ε) lässt sich ein passendes V (definiert durch δ) finden.

Stetigkeit $\Rightarrow \varepsilon, \delta$ existieren: Für jede Folge $x_n \in A$ mit $\lim_{n\to\infty} x_n = a$ gelte $\lim_{n\to\infty} f(x_n) = f(a)$.

Es ist zu zeigen, dass zu jedem $\varepsilon > 0$ ein $\delta > 0$ existiert, sodass $|f(x)-f(a)| < \varepsilon$ für alle $x \in A$ mit $|x-a| < \delta$.

Angenommen es gibt ein $\varepsilon > 0$ zu dem kein $\delta > 0$ existiert. Dann existiert zu jedem $\delta > 0$ mindestens ein $x \in A$ mit $|x-a| < \delta$, aber $|f(x)-f(a)| \geq \varepsilon$. Wähle $\delta = \frac{1}{n}$, dann gibt es für jede natürliche Zahl $n \geq 1$ ein $x_n \in A$ mit $|x_n - a| < \frac{1}{n}$ und $|f(x_n) - f(a)| \geq \varepsilon$. Da $\lim_{n\to\infty} x_n = a$ und nach Voraussetzung (f stetig) gilt folglich $\lim_{n\to\infty} f(x_n) = f(a)$. Dies steht im Widerspruch zu $|f(x_n) - f(a)| \geq \varepsilon$ für alle $n \geq 1$.

□

Korollar 4.23. *Sei $f : A \to \mathbb{R}$ stetig im Punkt $a \in A$ und $f(a) \neq 0$. Dann ist $f(x) \neq 0$ für alle x in einer Umgebung von a. D. h. es existiert ein $\delta > 0$, sodass $f(x) \neq 0$ für alle $x \in A$ mit $|x-a| < \delta$.*

Beweis: Zu $\varepsilon = |f(a)| > 0$ existiert nach Satz 4.22 ein $\delta > 0$, sodass $|f(x)-f(a)| < \varepsilon$ für alle $x \in A$ mit $|x-a| < \delta$. Also gilt nach Einsetzen

$$|f(x) - f(a)| < |f(a)| \Leftrightarrow |f(a)| - |f(x) - f(a)| > 0. \qquad (4.5.1)$$

Mit Hilfe der Dreiecksungleichung erhalten wir

$$|f(a)| = |f(a) - f(x) + f(x)| \leq |f(a) - f(x)| + |f(x)|$$

und nach Umstellung der Ungleichung

$$|f(x)| \geq |f(a)| - |f(x) - f(a)|. \qquad (4.5.2)$$

Mit Hilfe von Ungleichung (4.5.1) können wir die rechte Seite von Ungleichung (4.5.2) weiter abschätzen und erhalten schließlich $|f(x)| \geq |f(a)| - |f(x) - f(a)| > 0$ für alle $x \in A$ mit $|x-a| < \delta$. □

Korollar 4.24. *Sei $I \subseteq \mathbb{R}$ ein Intervall und $f : I \to \mathbb{R}$ stetig und streng monoton (wachsend oder fallend). Sei $J = f(I)$, dann bildet f das Intervall I bijektiv auf J ab und die Umkehrfunktion $f^{-1} : J \to \mathbb{R}$ ist stetig.*

Beweis: Die Monotonie von f^{-1} wurde bereits in Satz 4.11 gezeigt. f ist als streng monotone Funktion injektiv. Da $J = f(I)$, ist die Abbildung auch surjektiv und damit bijektiv. Damit bleibt nur noch die Stetigkeit von f^{-1} zu beweisen.
Zum Beweis der Stetigkeit benutzen wir das $\varepsilon-\delta$-Kriterium (Satz 4.22) und nehmen an, dass f streng monoton wachsend ist. Für ein $b \in J$ existiert ein $a = f^{-1}(b) \in I$, d. h. $b = f(a)$. Sei $b \in J$ kein Randpunkt von J, dann existiert $\varepsilon' > 0$, sodass $[b - \varepsilon', b + \varepsilon'] \subseteq J$. Wir zeigen nun die Stetigkeit von f^{-1} in b. Da f bijektiv abbildet, ist auch a kein Randpunkt von I. Damit können wir ein $\varepsilon > 0$ wählen,

sodass $[a-\varepsilon, a+\varepsilon] \subseteq I$. Sei $b_1 = f(a-\varepsilon)$ und $b_2 = f(a+\varepsilon)$. Wegen der strengen Monotonie von f gilt $b_1 < b < b_2$. Ferner bildet f das Intervall $[a-\varepsilon, a+\varepsilon]$ bijektiv auf das Intervall $[b_1, b_2]$ ab. Wir wählen $\delta = \min(b-b_1, b_2-b)$. Dann gilt $f^{-1}((b-\delta, b+\delta)) \subseteq (a-\varepsilon, a+\varepsilon)$ und f^{-1} ist stetig nach Satz 4.22. Damit haben wir die Stetigkeit von f^{-1} in allen Punkten gezeigt, die keine Randpunkte von J sind.
Für Randpunkte $b \in J$ erfolgt der Beweis durch Betrachtung von $[a, a-\varepsilon]$ bzw. $[a, a+\varepsilon]$ und ein ansonsten analoges Vorgehen. □

Den Begriff der Stetigkeit kann man noch verschärfen, wie die folgende Definition zeigt.

Definition 4.25 (Gleichmäßige Stetigkeit)**.** *Eine Funktion $f : A \to \mathbb{R}$ heißt in A* gleichmäßig stetig*, wenn gilt: Zu jedem $\varepsilon > 0$ existiert ein $\delta > 0$ so dass $|f(x) - f(x')| < \varepsilon$ für alle $x, x' \in A$ mit $|x - x'| < \delta$.*

Im Unterschied zu Stetigkeit hängt bei der gleichmäßigen Stetigkeit δ nur von ε nicht aber von $a \in A$ ab.

Satz 4.26 (Stetigkeit auf kompakten Intervallen)**.** *Jede auf einem kompakten Intervall stetige Funktion $f : [a, b] \to \mathbb{R}$ ist dort gleichmäßig stetig.*

Beweis: Angenommen f sei nicht gleichmäßig stetig. Dann gibt es ein $\varepsilon > 0$, sodass für alle $n \in \mathbb{N}$ Punkte $x_n, y_n \in [a, b]$ existieren mit $|x_n - y_n| < \frac{1}{n}$ $(\delta = \frac{1}{n})$ und $|f(x_n) - f(y_n)| \geq \varepsilon$. Die beschränkte Folge x_n besitzt nach dem Satz von Bolzano-Weierstraß (Satz 3.17) eine konvergente Teilfolge x_{n_k} mit $\lim\limits_{k\to\infty} x_{n_k} = c \in [a, b]$. Aus $\lim\limits_{k\to\infty} |x_{n_k} - y_{n_k}| = 0$ folgt $\lim\limits_{k\to\infty} y_{n_k} = c$. Da f stetig ist gilt auch

$$\lim_{k\to\infty} (f(x_{n_k}) - f(y_{n_k})) = f(c) - f(c) = 0$$

was im Widerspruch zur Annahme $|f(x_n) - f(y_n)| \geq \varepsilon$ für alle $n \in \mathbb{N}$ steht, sodass f gleichmäßig stetig ist. □

Der vorherige Satz gilt nicht für offene oder halboffene Intervalle, wie man sich leicht überlegen kann.

4.6 Logarithmen und allgemeine Potenzen

Die Exponentialfunktion haben wir bereits in Form der Exponentialreihe in Kapitel 3.7 kennen gelernt. Wir werden nun hier die Umkehrfunktion der Exponentialfunktion einführen und allgemeine Potenzen, bei denen die Basis aus $\mathbb{R}_{>0}$ und der Exponent aus $\mathbb{R}$ gewählt werden, definieren.

Die Exponentialfunktion ist definiert als $\exp : \mathbb{R} \to \mathbb{R}$ mit $\exp(x) = e^x = \sum\limits_{k=0}^{\infty} \frac{x^k}{k!}$ (siehe Satz 3.35). Es gilt $\exp(0) = 1$ und $\exp(1) = e$. Der Wertebereich ist $0 < \exp(x) < 1$ für $x \in (-\infty, 0)$ und $1 < \exp(x) < \infty$ für $x \in (0, \infty) \Rightarrow \exp(x) \in (0, \infty)$. Die Funktion exp bildet $\mathbb{R}$ bijektiv auf $\mathbb{R}_{>0}$ ab und ist ist streng monoton wachsend. Damit existiert die Umkehrfunktion von $\exp(x)$, die als der natürliche Logarithmus $\ln(x) : \mathbb{R}_{>0} \to \mathbb{R}$ bezeichnet wird. Es gilt

- $\ln(\exp(x)) = \exp(\ln(x)) = x$

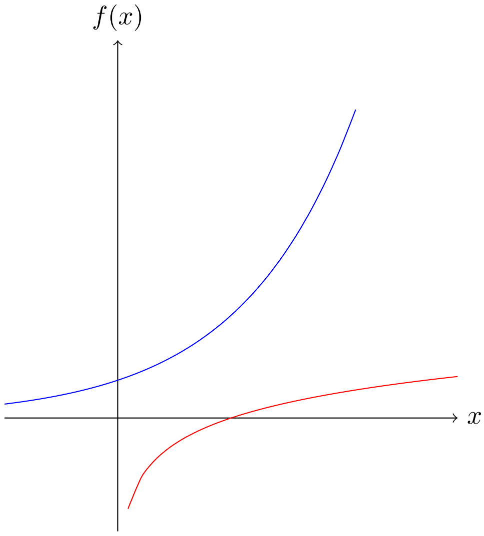


Abbildung 4.4: Verlauf der Exponential- und Logarithmusfunktion.

- $\ln(1) = 0$ und $\ln(e) = 1$
- $\ln(x) = \begin{cases} < 0 & \text{für} \quad x \in (0,1) \\ 0 & \text{für} \quad x = 1 \\ > 0 & \text{für} \quad x \in (1,\infty) \end{cases}$

Ferner gelten die beiden folgenden Rechenregeln für den Logarithmus.

i) $\ln(xy) = \ln(x) + \ln(y)$

ii) $\ln(\frac{1}{x}) = -\ln(x)$

Beweis:

i) Sei $x' = \ln(x)$, $y' = \ln(y)$ für $x, y \in (0,\infty)$. Es gilt $\exp(x' + y') = \exp(x') \cdot \exp(y') = x \cdot y$ (siehe Satz 3.37) also gilt $x' + y' = \ln(\exp(x' + y')) = \ln(x \cdot y)$.

ii) Es gilt nach *i)* $\ln(\frac{1}{x}) + \ln(x) = \ln(1) = 0 \Rightarrow \ln(\frac{1}{x}) = -\ln(x)$.

□

Bisher haben Potenzen die folgenden Fälle unterschieden, wobei $n \in \mathbb{N}$ jeweils gilt:

- $a^n = \underbrace{a \cdot a \cdot a \cdot \ldots \cdot a}_{n\text{-mal}}$ $(a \in \mathbb{R})$
- $a^0 = 1$ $(a \in \mathbb{R} \setminus \{0\})$
- $a^{-n} = \frac{1}{a^n}$ $(a \in \mathbb{R} \setminus \{0\})$
- $a^{\frac{1}{n}} = \sqrt[n]{a}$ $(a \in \mathbb{R}_{>0})$

Durch Kombination der Rechenregeln können wir damit für $a \in \mathbb{R}_{>0}$ und $n, m \in \mathbb{N}$ festlegen
$$a^{\frac{n}{m}} = \left(\sqrt[m]{a}\right)^n \ .$$
Dies Darstellung definiert Potenzen mit rationalen Exponenten. Um zu reellwertigen Exponenten zu gelangen nutzen wir die folgende Definition der Exponentialfunktion.

Definition 4.27 (Exponentialfunktion für allgemeine Basen)**.** *Für* $a \in \mathbb{R}_{>0}$ *sei die Funktion* $\exp_a : \mathbb{R} \to \mathbb{R}$ *definiert als*
$$\exp_a(x) = \exp(x \ln a).$$

Satz 4.28. *Die Funktion* $\exp_a(x)$ *ist stetig und es gilt*

i) $\exp_a(x+y) = \exp_a(x)\exp_a(y)$ *für alle* $x, y \in \mathbb{R}$

ii) $\exp_a(n) = a^n$ *für alle* $n \in \mathbb{Z}$

iii) $\exp_a\left(\frac{p}{q}\right) = \sqrt[q]{a^p}$ *für alle* $p \in \mathbb{Z}$ *und* $q \in \mathbb{N} \setminus \{1\}$

Beweis: Die Stetigkeit folgt, da $\exp_a$ die Komposition der stetigen Funktionen exp und ln ist, die nach Satz 4.17 stetig ist.
Der Beweis zu *i)* folgt aus den Rechenregeln für die Exponentialfunktion. Insbesondere gilt damit auch
$$\exp_a(-x) = \frac{1}{\exp_a(x)}.$$
$\exp_a(nx) = (\exp_a(x))^n$ für $n \in \mathbb{N}$ and $x \in \mathbb{R}$ zeigt man per vollständiger Induktion. Wenn wir $x = 1$ setzen, so gilt $\exp_a(1) = \exp(\ln a) = a$ und somit $\exp_a(n) = a^n$. Analog gilt für $x = -1$, $\exp_a(-1) = 1/a$ und damit $\exp_a(-n) = a^{-n}$. Damit ist auch *ii)* bewiesen.
Weiterhin gilt
$$a^p = \exp_a(p) = \exp_a\left(q \cdot \frac{p}{q}\right) = \left(\exp_a\left(\frac{p}{q}\right)\right)^q.$$
Da $\sqrt[q]{a^{\frac{p}{q}}} = a^p$ folgt *iii)*. □

Der vorherige Satz zeigt, dass die Bezeichnung von $\exp_a(x)$ als a^x berechtigt ist und wir damit eine allgemeine Definition von Potenzen haben. Man kann mit allgemeinen Potenzen genauso wie mit ganzzahligen Potenzen rechnen.

Satz 4.29. *Für* $a, b \in \mathbb{R}_{>0}$ *und* $x, y \in \mathbb{R}$ *gelten folgende Rechenregeln:*

i) $a^x a^y = a^{x+y}$

ii) $(a^x)^y = a^{xy}$

iii) $a^x b^x = (ab)^x$

iv) $\left(\frac{1}{a^x}\right) = a^{-x}$

Beweis: Zur Übung. □

Ähnlich zur Exponentialfunktion kann auch der Logarithmus für allgemeine Basen $a \in \mathbb{R}_{>0} \setminus \{1\}$ definiert werden.

Definition 4.30. *Sei $a \in \mathbb{R}_{>0} \setminus \{1\}$, dann ist der* Logarithmus zur Basis *a definiert als*

$$\log_a(x) = \frac{\ln(x)}{\ln(a)}.$$

Die Schreibweise zu Logarithmen ist in der Literatur nicht einheitlich. Manchmal bezeichnet $\log(x)$ den natürlichen Logarithmus, den wir mit $\ln(x)$ bezeichnen.

Beispiel 4.12. *Die Exponentialfunktion kann genutzt werden, um Wachstums- oder Zerfallsprozesse zu beschreiben. So kann der radioaktive Zerfall durch die Formel*

$$n(t) = n(0) \cdot e^{-\lambda t}$$

beschrieben werden. $t \geq 0$ ist der Zeitparameter, $n(t)$ ist die Anzahl der Kerne, die noch nicht zerfallen sind. Entsprechend ist $n(0)$ die Anzahl der Kerne zu Beginn des Experiments. λ ist eine materialspezifische Zerfallskonstante, die zum Beispiel für Jod 131 $10^{-6} sec^{-1}$ beträgt. Wenn wir diesen Wert auf Tage umrechnen erhalten wir $\lambda = 0.086$.

Die Halbwertszeit eines Stoffes ist die Zeit τ, die nötig ist, damit die Hälfte der Kerne zerfallen ist. Es muss also gelten

$$n(\tau) = n(0) \cdot e^{-\lambda\tau} = \frac{n(0)}{2}$$

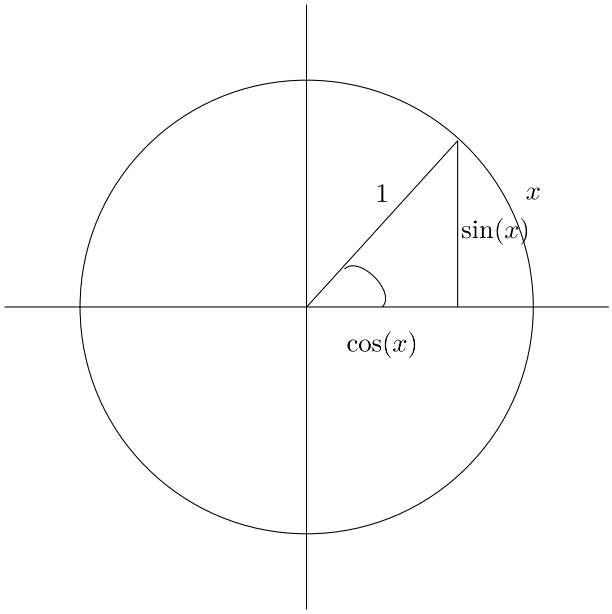
und somit

$$\ln\left(e^{-\lambda\tau}\right) = \ln\left(\frac{1}{2}\right) \Leftrightarrow -\lambda\tau = \ln\left(\frac{1}{2}\right) \Leftrightarrow \tau = \frac{\ln(2)}{\lambda},$$

da $\ln\left(\frac{1}{2}\right) = -\ln(2)$. Für Jod 131 beträgt die Halbwertszeit gut 8 Tage.

4.7 Trigonometrische Funktionen

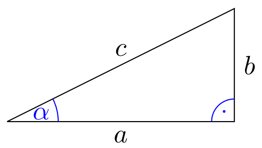
Als weitere Klasse von Funktionen betrachten wir die trigonometrischen Funktionen, die in der Geometrie und bei der Beschreibung periodischer Vorgänge eine große Bedeutung haben. In der Informatik werden trigonometrische Funktionen unter anderem in der Computergraphik breit eingesetzt. Trigonometrische Funktionen können im Einheitskreis definieren können, wie die folgende Graphik zeigt. Wir beschränken uns auf die Funktionen $\sin(x)$ und $\cos(x)$.



Der Winkel wird über die Länge des Kreisbogens angegeben, man spricht deshalb auch vom Bogenmaß. Es gilt dann

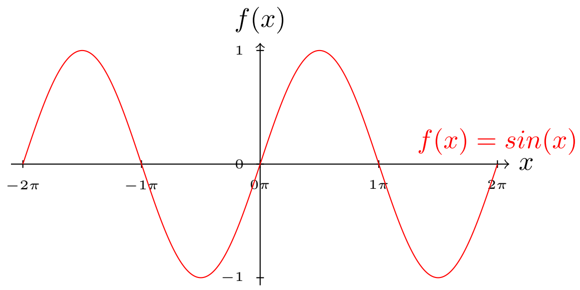
$$\text{Winkel in Bogenmaß } = \frac{2\pi}{360}\text{Winkel in Grad.}$$

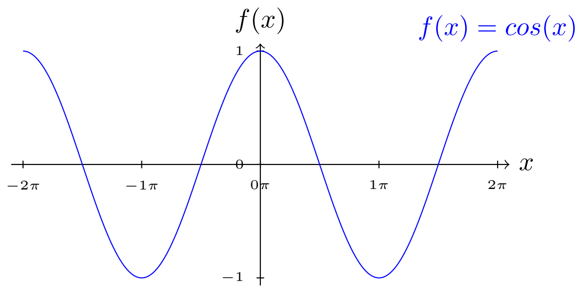
In einem rechtwinkligem Dreieck kann man sin und cos über die Seitenverhältnisse definieren, wie aus der Schule bekannt. Dies geschieht, indem die in der folgenden folgenden Graphik gezeigten Darstellungen genutzt werden, die sich aus der Darstellung im Einheitskreis ergeben.



$\sin(\alpha) = \frac{b}{c}$ $\qquad$ $\cos(\alpha) = \frac{a}{c}$

Die obige Darstellung von sin und cos im Einheitskreis kann man einfach verallgemeinern, indem man beliebige Drehungen in positive oder negative Richtungen im Einheitskreis erlaubt. Damit sind alle Kreislängen $x \in \mathbb{R}$ definierbar. Die Funktionswerte nach einer ganzen Umdrehung sind identisch und treten nach einer halben Umdrehung mit jeweils geändertem Vorzeichen wieder auf. Dies wird auch am Verlauf der beiden Funktionen deutlich, der in den folgenden Graphiken gezeigt wird.





Die Sinus- (sin) und Cosinus- (cos) Funktion kann man auch als Summe unendlicher Reihen definieren. Die Korrespondenz zwischen der Definition über die Länge des Bogens im Einheitskreis und den Reihensummen lässt sich über die Darstellung der Exponentialfunktion mit komplexen Argumenten herleiten. Da wir die komplexen Zahlen in der Vorlesung nicht behandeln, soll auf diese Herleitung verzichtet werden.

Definition 4.31. *Für alle $x \in \mathbb{R}$ sind die Funktionen $\cos : \mathbb{R} \to \mathbb{R}$ und $\sin : \mathbb{R} \to \mathbb{R}$ definiert als*

$$\cos(x) = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!} \quad \textit{und} \quad \sin(x) = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!}.$$

Satz 4.32. *Die Reihen für $\cos(x)$ und $\sin(x)$ konvergieren absolut für alle $x \in \mathbb{R}$.*

Beweis: Die Beweise nutzen das Quotientenkriterium (Satz 3.33). Zu zeigen ist

$$\left| \frac{a_{k+1}}{a_k} \right| \leq q < 1$$

für alle $k \geq n_0$. Wir beweisen die Konvergenz der Cosinus–Reihe, der Beweis für die Sinus-Reihe ist völlig analog.

$$\left| \frac{a_{k+1}}{a_k} \right| = \left| -1 \cdot \frac{x^{2k+2}(2k)!}{(2k+2)!x^{2k}} \right| = \left| \frac{x^2}{(2k+1)(2k+2)} \right|.$$

Wähle $n_0 = \lceil |x| \rceil$, dann gilt für $k \geq n_0$:

$$\left| \frac{a_{k+1}}{a_k} \right| = \left| \frac{x^2}{(2k+1)(2k+2)} \right| \leq \left| \frac{x^2}{(2|x|+1)(2|x|+2)} \right| \leq \left| \frac{x^2}{4x^2} \right| = \frac{1}{4} < 1\,. \qquad \square$$

Interessant ist die Frage, inwieweit man nach endlich vielen Summationen eine gute Approximation des Wertes der Cosinus- oder Sinus-Funktion erreicht hat. Dazu betrachten wir das so genannte Restglied, welches uns später in Kapitel 7 noch einmal in allgemeinerer Form begegnen wird. Sei

$$\cos(x) = \sum_{k=0}^{n}(-1)^k \frac{x^{2k}}{(2k)!} + r_{2n+2}(x) \text{ und } \sin(x) = \sum_{k=0}^{n}(-1)^k \frac{x^{2k+1}}{(2k+1)!} + r_{2n+3}(x).$$

Die Restglieder $r_{2n+2}(x)$ und $r_{2n+3}(x)$ umfassen den Teil der unendlichen Summe, der bei einer endlichen Summation weggelassen wird.

Satz 4.33. *Für die Restglieder der Cosinus- und Sinus-Reihe gelten die folgenden Abschätzungen*

$$\begin{aligned} |r_{2n+2}(x)| &\leq \tfrac{|x|^{2n+2}}{(2n+2)!} \quad \text{für } |x| \leq 2n+3 \\ |r_{2n+3}(x)| &\leq \tfrac{|x|^{2n+3}}{(2n+3)!} \quad \text{für } |x| \leq 2n+4 \end{aligned}$$

Beweis: Wir beschränken uns auf den Beweis für das Restglied der Cosinus-Reihe, der Beweis für das Restglied der Sinus-Reihe ist analog zu führen. Es gilt

$$r_{2n+2}(x) = \sum_{k=n+1}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!} = \pm\frac{x^{2n+2}}{(2n+2)!}\left(1 - \frac{x^2}{(2n+3)(2n+4)} + \ldots\right).$$

Sei nun $a_0 = 1$ und

$$a_k = a_{k-1}\frac{x^2}{(2n+2k+1)(2n+2k+2)} \quad \text{für } k \geq 1,$$

dann ist

$$r_{2n+2}(x) = \pm\frac{x^{2n+2}}{(2n+2)!}\left(1 - a_1 + a_2 - a_3 + a_4 - \ldots\right).$$

Für $x \leq 2n+3$ gilt

$$1 > a_1 > a_2 > a_3 > \ldots > 0.$$

Nach dem Beweis des Leibniz-Kriteriums (Satz 3.28) gilt damit

$$1 \geq 1 - a_1 + a_2 - a_3 + \ldots \geq 0$$

und damit auch $|r_{2n+2}(x)| \leq \frac{|x|^{2n+2}}{(2n+2)!}$. □

Der vorherige Satz zeigt, dass man für jedes vorgegebene $x \in \mathbb{R}$ n so wählen kann, dass der Fehler bei Abbruch der Summation nach n Schritten unterhalb einer vorgegeben Schranke $\varepsilon > 0$ liegt.

Definition 4.34 (Periodische Funktion)**.** *Eine Funktion $f : \mathbb{R} \to \mathbb{R}$ heißt* periodische Funktion*, wenn es ein $p > 0$ gibt, so dass*

$$f(x) = f(x+p)$$

für alle $x \in \mathbb{R}$.
Das kleinste $p \in \mathbb{R}_{>0}$ mit der obigen Eigenschaft heißt Periode der Funktion f.

$\sin(x)$ und $\cos(x)$ haben offensichtlich die Periode 2π, also $\sin(x) = \sin(x+2\pi) = \sin(x+4\pi) = \ldots$. Der folgende Satz fasst einige Eigenschaften von sin und cos zusammen und zeigt auch, dass sin immer durch cos ausgedrückt werden kann und umgekehrt.

Satz 4.35 (Eigenschaften von sin und cos)**.** *Es gilt:*

i) $\sin^2 x + \cos^2 x = 1$ *(Satz des Pythagoras)*

ii) $\cos(-x) = \cos(x)$
$\sin(-x) = -\sin(x)$

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

iv) $\cos(x + 2\pi) = \cos(x)$
$\sin(x + 2\pi) = \sin(x)$

v) $\cos(x) = 0 \Leftrightarrow x \in \{\frac{\pi}{2} + k\pi \mid k \in \mathbb{Z}\}$
$\sin(x) = 0 \Leftrightarrow x \in \{k\pi \mid k \in \mathbb{Z}\}$

vi) $|\sin(x)| \leq 1$
$|\cos(x)| \leq 1$

Beweis: Zur Übung unter Nutzung der Reihendarstellungen aus Definition 4.31. □

Kapitel 5

Differenzierbare Funktionen

Die Differentialrechnung, die auf Leibniz und Newton zurückgeht, bildet den eigentlichen Kern der Analysis. Die Differentialrechnung ist die Basis für Differentialgleichungen mit denen sich viele Vorgänge in der Natur und in technischen Systemen modellieren lassen.
Wir betrachten in diesem Kapitel die Differentialrechnung mit einer Veränderlichen, wie sie auch in der Schule behandelt wird. Nach einer generellen Einführung, leiten wir die Ableitungsregeln her, führen höhere Ableitungen ein und nutzen schließlich die Ableitungen um Eigenschaften von Funktionen im Rahmen der Kurvendiskussion zu analysieren.

5.1 Differenzierbarkeit einer Funktion

Eine Funktion f lässt sich, falls $f(x) \to f(a)$ für $x \to a$ (a Häufungspunkt, bzw. f ist stetig in a), in der Umgebung von a durch eine konstante Funktion annähern. Dies ist gerade die Idee der Stetigkeit. Differenzierbarkeit nutzt eine bessere Art der Approximation durch ein Polynom statt einer konstante Funktion.

Definition 5.1 (Differenzierbarkeit)**.** *Sei $a \in A \subseteq \mathbb{R}$ und sei $f : A \to \mathbb{R}$ eine Funktion. f heißt* in a differenzierbar*, falls der Grenzwert*

$$f'(a) = \lim_{\substack{x \to a \\ x \in A \setminus \{a\}}} \frac{f(x) - f(a)}{x - a}$$

existiert.

Es wird in der Definition vorausgesetzt, dass eine Folge $a_k \in A \setminus \{a\}$ mit $\lim_{k \to \infty} (a_k) = a$ existiert (d. h. a ist Häufungspunkt von A).
$f'(a)$ heißt die Ableitung von f in a. Man kann die Ableitung auch darstellen als

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

wobei wir nun Folgen h_k mit $h_k \neq 0$ und $x + h_k \in A$ zulassen. Eine Funktion f ist in einer Menge $A \subseteq \mathbb{R}$ differenzierbar, wenn f in jedem $a \in A$ differenzierbar ist.

Satz 5.2. *Eine Funktion $f : A \to \mathbb{R}$ $(A \subseteq \mathbb{R})$ ist in einem Häufungspunkt $a \in A$ genau dann differenzierbar, wenn es eine Konstante $c \in \mathbb{R}$ gibt, sodass*

$$f(x) = f(a) + c \cdot (x - a) + r(x)$$

für $x \in A$. Wobei $r(x)$ eine Funktion mit der Eigenschaft $\lim\limits_{x \to a,\ x \neq a} \frac{r(x)}{x-a} = 0$ ist. Es gilt in diesem Fall $c = f'(a)$.

Beweis:

1. f sei in a differenzierbar.
Zu zeigen: c und $r(x)$ existieren.
Wir definieren $c = f'(a)$ und $r(x) = f(x) - f(a) - f'(a)(x - a)$. Es gilt dann

$$\frac{r(x)}{x-a} + f'(a) = \frac{f(x) - f(a)}{x-a}.$$

Da $\lim\limits_{x \to a} \frac{f(x)-f(a)}{x-a}$ existiert (da f differenzierbar), existiert auch $\lim\limits_{x \to a} \frac{r(x)}{x-a}$. Damit gilt

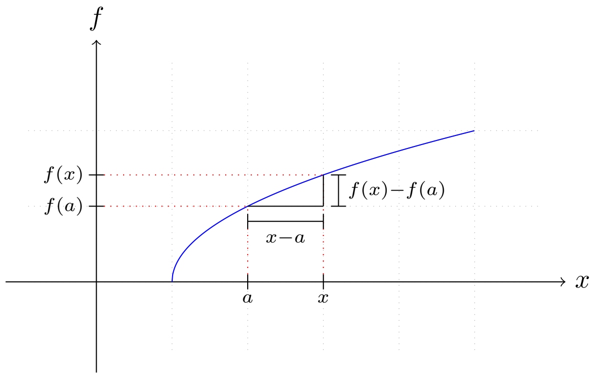
$$\lim_{\substack{x \to a \\ x \neq a}} \frac{r(x)}{x-a} = \lim_{\substack{x \to a \\ x \neq a}} \frac{f(x) - f(a)}{x-a} - f'(a) = 0.$$

2. c und $r(x)$ existieren.
Zu zeigen: f ist in a differenzierbar.
Es gilt

$$\lim_{x \to a} \left(\frac{f(x) - f(a)}{x-a} - c \right) = \lim_{x \to a} \frac{r(x)}{x-a} = 0 \Rightarrow \lim_{x \to a} \frac{f(x) - f(a)}{x-a} = c.$$

Dies ist genau die Definition von Differenzierbarkeit und $c = f'(a)$. □

Man kann die Ableitung geometrisch interpretieren, da $\frac{f(x)-f(a)}{x-a}$ der Sekante des Graphen von f im Punkt $(a, f(a))$ entspricht. Beim Grenzübergang geht die Sekante in die Tangente im Punkt a, mit der Tangentengleichung $f(a) + f'(a)(x - a)$ über. Die folgende Graphik veranschaulicht diese Interpretation. Die Ableitung einer Funktion beschreibt damit eine lineare Approximation des Wachstums der Funktion in einem Punkt. Damit liefert die Ableitung einen Zusammenhang zwischen der Änderung der Variablen x und dem Funktionswert $f(x)$. Wenn x zum Beispiel die Zeit beschreibt und $f(x)$ die Bewegung über die Zeit, so ist die Ableitung gerade die Beschleunigung.



Es gibt alternative Schreibweisen für die Ableitung, die gerade im Bereich der Differentialgleichungen oft verwendet werden und insbesondere dann benutzt werden, wenn Funktionen mit mehr als einer Veränderlichen untersucht werden. Die Ableitung wird dann Bruch $\frac{df(x)}{dx}$ dargestellt. Diese Darstellung folgt aus der Grenzwertbetrachtung von $\frac{\Delta f(x)}{\Delta x}$ mit $\frac{df(x)}{dx} = \lim\limits_{\Delta x \to 0} \frac{\Delta f(x)}{\Delta x}$. Streng genommen ist die Darstellung nicht unproblematisch, da man für $x = 0$ eigentlich nicht $\frac{df(0)}{d0}$ schreiben kann, wohl aber $f'(0)$. Alternativ werden deshalb teilweise die Formen $\frac{df}{dx}(0)$ oder $\frac{df(x)}{dx}\,|_{x=0}$ verwendet.

Beispiel 5.1. *In den folgenden Beispielen betrachten wir Folgen $(x_n)_{n\in\mathbb{N}}$ mit $\lim\limits_{n\to\infty} x_n = a$.*

1. $f : \mathbb{R} \to \mathbb{R}$, $f(x) = c$ $(c \in \mathbb{R})$

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

2. $f : \mathbb{R} \to \mathbb{R}$ $f(x) = cx$ $(c \in \mathbb{R})$

$$f'(x) = \lim_{h\to 0} \frac{f(x+h) - f(x)}{h} = \lim_{h\to 0} \frac{c(x+h) - cx}{h} = c$$

3. $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^2$

$$\begin{aligned} f'(x) &= \lim_{h\to 0} \tfrac{f(x+h)-f(x)}{h} && = \lim_{h\to 0} \tfrac{(x+h)^2-x^2}{h} \\ &= \lim_{h\to 0} \tfrac{2xh+h^2}{h} && = \lim_{h\to 0}(2x+h) && = 2x \end{aligned}$$

4. $f : \mathbb{R}\setminus\{0\} \to \mathbb{R}$ $f(x) = \frac{1}{x}$

$$\begin{aligned} f'(x) &= \lim_{h\to 0} \tfrac{f(x+h)-f(x)}{h} && = \lim_{h\to 0} \left(\tfrac{1}{h} \cdot \left(\tfrac{1}{x+h} - \tfrac{1}{x}\right)\right) \\ &= \lim_{h\to 0} \tfrac{x-(x+h)}{hx(x+h)} && = \lim_{h\to 0} \tfrac{-1}{x(x+h)} && = -\tfrac{1}{x^2} \end{aligned}$$

5. $f(x) = \exp(\lambda x),\ \lambda \in \mathbb{R}$

$$\begin{aligned}
\exp'(\lambda x) &= \lim_{h\to 0} \tfrac{\exp(\lambda(x+h))-\exp(\lambda x)}{h} &&= \lim_{h\to 0} \tfrac{\exp(\lambda x)\cdot(\exp(\lambda h)-1)}{h} \\
&= \exp(\lambda x)\cdot \lim_{h\to 0} \tfrac{\exp(\lambda h)-1}{h} &&= \exp(\lambda x) \lim_{h\to 0}(\tfrac{1}{h}\sum_{k=0}^{\infty} \tfrac{(\lambda h)^k}{k!} - 1) \\
&= \exp(\lambda x)\cdot \lim_{h\to 0}(\sum_{k=1}^{\infty} \tfrac{\lambda^k h^{k-1}}{k!}) &&= \exp(\lambda x)\cdot\lambda
\end{aligned}$$

Die letzte Umformung kann durchgeführt werden, da für $h \to 0$ alle Summanden mit Index $k \geq 2$ gegen 0 konvergieren und für $k = 1$ der Wert λ angenommen wird.

Der folgende Satz stellt eine Beziehung zwischen Stetigkeit und Differenzierbarkeit her.

Satz 5.3 (Stetigkeit - Differenzierbarkeit)**.** *Ist die Funktion $f : A \to \mathbb{R}$ ($A \subseteq \mathbb{R}$) in $a \in A$ differenzierbar, so ist sie in a auch stetig.*

Beweis: Nach Satz 5.2 gilt dann $\lim\limits_{x\to a} f(x) = f(a) + \lim\limits_{x\to a} c(x-a) + \lim\limits_{x\to a} r(x) = f(a)$ □

Die Umkehrung von Satz 5.3 gilt nicht, wie das folgende Beispiel zeigt!

Beispiel 5.2. *Sei $f : \mathbb{R} \to \mathbb{R}$, $f(x) = |x|$*
$f(x)$ ist stetig in jedem $a \in \mathbb{R}$, da $\lim\limits_{x\to a}(|x|) = |a|$ aber

$$\lim_{x\to 0} \frac{|x| - |0|}{x} = \begin{cases} 1 & falls \quad x > 0 \\ -1 & falls \quad x < 0 \end{cases}$$

Damit ist $|x|$ im Punkt 0 nicht differenzierbar.

5.2 Differentiations-Regeln

Zur Berechnung der Ableitung einer Funktion wird man in den meisten Fällen nicht direkt auf die Grenzwertbildung zurückgreifen, sondern die Ableitung auf schon bekannte elementare Fälle zurückführen. Dazu werden die in diesem Abschnitt vorgestellten Ableitungsregeln verwendet.

Satz 5.4 (Algebraische Operationen)**.** *Seien $f, g : A \to \mathbb{R}$ in $a \in A$ differenzierbar und $c \in \mathbb{R}$. Dann sind auch die Funktionen $f+g, c\cdot f, f\cdot g : A \to \mathbb{R}$ in a differenzierbar und es gelten folgende Rechenregeln:*

i) Linearität: $(f+g)'(a) = f'(a) + g'(a)$
$(c\cdot f)'(a) = c\cdot f'(a)$

ii) Produktregel: $(f\cdot g)'(a) = f'(a)g(a) + f(a)g'(a)$

Ist ferner $g(x) \neq 0$ für alle $x \in A$, so ist auch die Funktion $\frac{f}{g} : A \to \mathbb{R}$ differenzierbar und es gilt:

iii) Quotientenregel: $(\frac{f}{g})'(a) = \frac{f'(a)g(a) - f(a)g'(a)}{g^2(a)}$

Beweis:

i) Folgt direkt aus der Definition des Differenzenquotienten und den Regeln für Ableitungen.

ii) Es gilt

$$\begin{aligned}(fg)'(a) &= \lim_{h\to 0} \tfrac{f(a+h)g(a+h)-f(a)g(a)}{h} \\ &= \lim_{h\to 0}(\tfrac{1}{h}(f(a+h)(g(a+h)-g(a)) + (f(a+h)-f(a))g(a))) \\ &= \lim_{h\to 0}(f(a+h)\tfrac{g(a+h)-g(a)}{h}) + \lim_{h\to 0}(\tfrac{f(a+h)-f(a)}{h}g(a)) \\ &= f(a)g'(a) + f'(a)g(a)\end{aligned}$$

iii) Wir beweisen zuerst den Spezialfall $f = 1$.

$$\begin{aligned}\left(\tfrac{1}{g}\right)'(a) &= \lim_{h\to 0}\left(\tfrac{1}{h}\left(\tfrac{1}{g(a+h)} - \tfrac{1}{g(a)}\right)\right) \\ &= \lim_{h\to 0}\left(\tfrac{1}{g(a+h)g(a)}\left(\tfrac{g(a)-g(a+h)}{h}\right)\right) \\ &= -\tfrac{g'(a)}{g(a)^2}\ .\end{aligned}$$

Allgemein gilt dann

$$\begin{aligned}\left(\tfrac{f}{g}\right)'(a) &= \left(f\tfrac{1}{g}\right)'(a) \overset{\text{nach } ii)}{=} f'(a)\tfrac{1}{g(a)} + f(a)\tfrac{-g'(a)}{g(a)^2} \\ &= \tfrac{f'(a)g(a)-f(a)g'(a)}{g(a)^2}\end{aligned}$$

□

Beispiel 5.3. *1.* $f_n(x) = x^n$ *(*$n \in \mathbb{N}$*)*
$f_n'(x) = nx^{n-1}$
Beweis: *Vollständige Induktion über* n.
Für $n = 0$ *und* $n = 1$ *wurde die Behauptung in den Beispielen 5.1.i) und 5.1.ii) in Abschnitt 5.1 gezeigt. Damit ist der Induktionsanfang bewiesen. Ferner sei* $f_{n-1}'(x) = (n-1)x^{n-2}$.
Sei $n \geq 2$*, dann lautet der Induktionsschritt:*

$$\begin{aligned}f_n(x) &= f_1(x)f_{n-1}(x) \\ f_n'(x) &= f_1'(x)f_{n-1}(x) + f_1(x)f_{n-1}'(x) = x^{n-1} + x(n-1)x^{n-2} \\ &= nx^{n-1}\end{aligned}$$

□

2. Polynomfunktionen
Sei $n \in \mathbb{N}_0$ *und* $a_0, \ldots, a_n \in \mathbb{R}$, $f(x) = \sum\limits_{k=0}^{n} a_k x^k$
$f(x)$ *ist differenzierbar in* $x \in \mathbb{R}$ *und es gilt:*

$$f'(x) = \begin{cases} \sum\limits_{k=1}^{n} a_k k x^{k-1} & falls \quad n \geq 1 \\ 0 & falls \quad n = 0 \end{cases}$$

Der Beweis folgt aus Beispiel i) und der Summenregel.

3. $f(x) = \frac{1}{x^n}$, $x \in \mathbb{R}_{\neq 0}$, $n \in \mathbb{N}$
   Die Quotientenregel liefert:

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

4. $f : \mathbb{R} \to \mathbb{R}$, $f(x) = \sin(x)$
   $f'(x) = \sin'(x) = \lim\limits_{h\to 0} \frac{\sin(x+h)-\sin(x)}{h} = \cos(x)$
   Zum Beweis berechnen wir die Ableitung der Sinus-Reihe (Definition 4.31). Nach Satz 5.4.i) entspricht die Ableitung der Summe, der Summe der Ableitungen der einzelnen Summanden. Für die Ableitung der Summanden können wir die Ableitung von x^n (siehe 1.) verwenden.

$$\begin{aligned}
\sin'(x) &= \left(\sum_{k=0}^{\infty}(-1)^k \frac{x^{2k+1}}{(2k+1)!}\right)' &&= \left(\sum_{k=0}^{\infty}(-1)^k \frac{(2k+1)x^{2k}}{(2k+1)!}\right) \\
&= \left(\sum_{k=0}^{\infty}(-1)^k \frac{x^{2k}}{(2k)!}\right) &&= \cos(x)
\end{aligned}$$

5. $f : \mathbb{R} \to \mathbb{R}$, $f(x) = \cos(x)$
   $f'(x) = \cos'(x) = \lim\limits_{h\to 0} \frac{cos(x+h)-cos(x)}{h} = -\sin(x)$.
   Zum Beweis bestimmen wir die Ableitung der Cosinus-Reihe(Definition 4.31). Nach Satz 5.4.i) entspricht die Ableitung der Summe, der Summe der Ableitungen der einzelnen Summanden. Für die Ableitung der Summanden können wir die Ableitung von x^n (siehe 1.) verwenden.

$$\begin{aligned}
\cos'(x) &= \left(\sum_{k=0}^{\infty}(-1)^k \frac{x^{2k}}{(2k)!}\right)' &&= \left(\sum_{k=1}^{\infty}(-1)^k \frac{2k\cdot x^{2k-1}}{(2k)!}\right) \\
&= \left(\sum_{k=1}^{\infty}(-1)^k \frac{x^{2k-1}}{(2k-1)!}\right) &&= \left(\sum_{k=0}^{\infty}(-1)^{k+1} \frac{x^{2k+1}}{(2k+1)!}\right) \\
&= -\left(\sum_{k=0}^{\infty}(-1)^k \frac{x^{2k+1}}{(2k+1)!}\right) &&= -\sin(x)
\end{aligned}$$

Die folgenden beiden Sätze erweitern die Ableitungsregeln um Regeln zur Bestimmung der Ableitung der Umkehrfunktion und komponierter Funktionen.

Satz 5.5 (Ableitung der Umkehrfunktion)**.** *Sei $I \subset \mathbb{R}$ ein Intervall, das aus mehr als einem Punkt besteht und sei $f : I \to \mathbb{R}$ eine stetige, streng monotone Funktion und $g = f^{-1} : J \to \mathbb{R}$ mit $J = f(I)$ deren Umkehrfunktion.*
Ist f in $a \in I$ differenzierbar und gilt $f'(a) \neq 0$, so ist g in $b = f(a)$ differenzierbar und es gilt:

$$g'(b) = \frac{1}{f'(a)} = \frac{1}{f'(g(b))} \ .$$

Beweis: Sei $b_n \in J \setminus \{b\}$ eine Folge mit $\lim\limits_{n\to\infty} b_n = b$. Wir setzen $a_n = g(b_n)$ (a_n existiert, da $b_n \in f(I)$).
Da g stetig ist (siehe Korollar 4.24), ist $\lim\limits_{n\to\infty} a_n = a$ und da $g : J \to I$ bijektiv ist, ist $a_n \neq a$ für $b_n \neq b$. Es gilt

$$\lim_{n\to\infty} \frac{g(b_n) - g(b)}{b_n - b} = \lim_{n\to\infty} \frac{a_n - a}{f(a_n) - f(a)} = \frac{1}{f'(a)}.$$

Der Grenzwert existiert, da die Grenzwerte im Zähler und Nenner existieren. □

Satz 5.6 (Kettenregel)**.** *Seien $f : A \to \mathbb{R}$ und $g : B \to \mathbb{R}$ ($A, B \subseteq \mathbb{R}$) Funktionen. Sei f in $a \in A$ differenzierbar und g sei in $b = f(a)$ differenzierbar. Dann ist die zusammengesetzte Funktion $g \circ f : A \to \mathbb{R}$ in Punkt a differenzierbar und es gilt*

$$(g \circ f)'(a) = g'(f(a)) \cdot f'(a) \ .$$

Beweis: Definiere Funktion $g^* : B \to \mathbb{R}$ durch

$$g^*(y) := \begin{cases} \frac{g(y)-g(b)}{y-b} & , \text{ falls } y \neq b \\ g'(b) & , \text{ falls } y = b \ . \end{cases}$$

Da g in b differenzierbar ist, gilt

$$\lim_{y \to b} g^*(y) = g^*(b) = g'(b)$$

sowie

$$g(y) - g(b) = g^*(y) \cdot (y - b) \ .$$

Daraus folgt

$$\begin{aligned} (g \circ f)'(a) &= \lim_{x \to a} \frac{g(f(x)) - g(f(a))}{x - a} \\ &= \lim_{x \to a} \frac{g^*(f(x)) \cdot (f(x) - f(a))}{x - a} \\ &= \lim_{x \to a} g^*(f(x)) \cdot \lim_{x \to a} \frac{f(x) - f(a)}{x - a} \\ &= g'(f(a)) \cdot f'(a) \end{aligned}$$

□

Beispiel 5.4.

1. *Sei $F(x) = f(ax + b)$ ($a, b \in \mathbb{R}$) und seien F und f in $\mathbb{R}$ differenzierbar, dann gilt*
$$F'(x) = a \cdot f'(ax + b) \ .$$

2. *Sei $F(x) = \exp(x^2)$. Dann gilt*
$$\begin{aligned} f(x) &= x^2 \\ g(y) &= \exp(y) \\ f'(x) &= 2x \\ g'(y) &= \exp(y) \\ F'(y) &= \exp(y^2) \cdot 2y \ . \end{aligned}$$

3. *Sei* $F(x) = \sqrt{x^2+2}$. *Dann gilt*

$$\begin{aligned} f(x) &= x^2+2 \\ f'(x) &= 2x \\ g(y) &= \sqrt{y} \\ g'(y) &= \frac{1}{2\sqrt{y}}\ . \end{aligned}$$

Da $\lim_{y\to x} \frac{\sqrt{y}-\sqrt{x}}{y-x} = \lim_{y\to x} \frac{\sqrt{y}-\sqrt{x}}{(\sqrt{y}-\sqrt{x})\cdot(\sqrt{y}+\sqrt{x})} = \lim_{y\to x} \frac{1}{\sqrt{y}+\sqrt{x}} = \frac{1}{2\sqrt{x}}$ *und damit* $F'(x) = \frac{1}{2\sqrt{x^2+2}} \cdot 2x$.

4. $f(x) = \ln(x)$, $f : \mathbb{R}_{>0} \to \mathbb{R}$.
$\ln(x)$ *ist die Umkehrfunktion von* $\exp(x)$. *Damit gilt nach Satz 5.5 mit* $f(x) = \exp(x)$ *und* $g(x) = \ln(x)$

$$g'(x) = \frac{1}{f'(g(x))} = \frac{1}{\exp(\ln(x))} = \frac{1}{x}\ .$$

5.3 Ableitungen höherer Ordnung

Die Idee der Ableitung einer Funktion kann rekursiv angewendet werden, indem man die Ableitung der Ableitung bildet, sofern diese existiert. Dies führt zu folgender Definition höherer Ableitungen.

Definition 5.7 (Ableitung höherer Ordnung)**.** *Sei* $f : A \to \mathbb{R}$ *mit* $A \subseteq \mathbb{R}$. *Dann ist die* k-te Ableitung *(oder* Ableitung k-ter Ordnung*) von* f *in* $a \in A$ *definiert als* $f^{(k)}(a)$ *(*f *oben* k, $k \in \mathbb{N}_0$*) mit*

1. $f^{(0)}(a) = f(a)$
2. $f^{(k+1)}(a) = (f^{(k)}(a))' : A \to \mathbb{R}$, *falls die Ableitung von* $f^{(k)}(a)$ *in* $a \in A$ *existiert.*

Für $f : A \to \mathbb{R}$ benutzt man die Schreibweise $f^{(k)}$, wenn die k-te Ableitung für alle $a \in A$ existiert. Eine Funktion heißt k-mal differenzierbar, wenn $f^{(k)}$ existiert. Die Funktion f heißt k-mal stetig differenzierbar, wenn f k-mal differenzierbar ist und $f^{(k)}$ stetig ist. Man schreibt auch

$$\frac{d^k f(x)}{dx^k}$$

für die k-te Ableitung im Punkt x. Es wird ferner manchmal auch die Schreibweise $(f(x))^{(k)}$ verwendet.
Wie schon bei der ersten Ableitung lassen sich auch für höhere Ableitungen Regeln für zusammengesetzte Funktionen herleiten, die im folgenden Satz zusammengefasst werden.

Satz 5.8 (Operationen auf Ableitungen)**.** *Sei* $k \in \mathbb{N}$, $c \in \mathbb{R}$, $\emptyset \neq A \subseteq \mathbb{R}$ *und seien* $f : A \to \mathbb{R}$ *und* $g : A \to \mathbb{R}$ k*-mal differenzierbar.*

1. *Dann sind* $f+g$, $f-g$, $c \cdot f$, $f \cdot g$ *und falls* $g(x) \neq 0$ *für alle* $x \in A$ $\frac{f}{g}$ k*-mal differenzierbar und es gilt:*

i) $(f+g)^{(k)}(a) = f^{(k)}(a) + g^{(k)}(a)$

ii) $(f-g)^{(k)}(a) = f^{(k)}(a) - g^{(k)}(a)$

iii) $(cf)^{(k)}(a) = cf^{(k)}(a)$

iv) $(fg)^{(k)}(a) = \sum\limits_{i=0}^{k} \binom{k}{i} f^{(i)}(a) \cdot g^{(k-i)}(a)$ *(Leibnizsche Formel)*

v) $\left(\frac{f}{g}\right)^{(k)}(a) = \frac{f^{(k)}(a) - \sum\limits_{i=0}^{k-1} \binom{k}{i} \left(\frac{f}{g}\right)^{(i)}(a) g^{(k-i)}(a)}{g(a)}$

2. *Ist ferner $f(A) \subseteq B \subseteq \mathbb{R}$ und $g : B \to \mathbb{R}$ k-mal differenzierbar, so ist auch $(g \circ f)$ k-mal differenzierbar.*

Beweis: Der Beweis erfolgt durch rekursive Anwendung der Differentiationsregeln für die erste Ableitung. □

Für die Kettenregel haben wir keine allgemeine Form für Ableitungen höherer Ordnung angegeben, da diese sehr schnell unübersichtlich wird und damit nicht hilfreich ist.

Beispiel 5.5.

$$f(x) = x^n \qquad \textit{mit } n \in \mathbb{N}_0 \textit{ und } x \in \mathbb{R}$$

$$f^{(k)}(x) = \begin{cases} k!\binom{n}{k} x^{(n-k)} , & \textit{falls } 0 \leq k \leq n \\ 0 , & \textit{falls } k > n \end{cases}$$

Beweis: Per Induktion über k:
$k = 0$:

$$f(x) = x^n = 0!\binom{n}{0} x^{n-0} = x^n$$

Sei $f^{(k)}(x) = k!\binom{n}{k} x^{n-k}$. $k \to (k+1)$:

$$\begin{aligned} f^{(k+1)}(x) = (f^{(k)})'(x) &= \begin{cases} (k!\binom{n}{k} x^{n-k})' , & \text{falls } 0 \leq k \leq n \\ 0 , & \text{falls } k > n \end{cases} \\ &= \begin{cases} k!\binom{n}{k}(n-k) x^{n-k-1} , & \text{falls } 0 \leq k < n \\ k!\binom{n}{k} 0 , & \text{falls } k = n \\ 0 , & \text{falls } k > n \end{cases} \\ &= \begin{cases} (k+1)!\binom{n}{k+1} x^{n-(k+1)} , & \text{falls } 0 \leq k+1 \leq n \\ 0 , & \text{falls } k+1 > n \end{cases} \end{aligned}$$

Der Übergang vom vorletzten zum letzten Schritt gilt, da

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

□

Bevor wir uns anschauen, wozu man (höhere) Ableitungen bei der Analyse von Funktionen verwenden kann, wird im nächsten Abschnitt kurz auf die Berechnung der Ableitung einer beliebigen Funktion eingegangen.

5.4 Numerisches Differenzieren

Mit Hilfe numerischer Software sind Ableitungen numerisch berechnenbar. Die Berechnung erfolgt für allgemeine Funktionen in der Regel ohne eine symbolische Bestimmung der Stammfunktion. Da die Ableitung als Grenzwert des Differenzenquotienten definiert ist, liegt folgende Approximation nahe:

$$f'(x) \approx \frac{f(x+h)-f(x)}{h} \tag{5.4.1}$$

oder

$$f'(x) \approx \frac{f(x+h)-f(x-h)}{2h} \tag{5.4.2}$$

In beiden Fällen ist eine festes $h \neq 0$ zu wählen, um die Approximation zu berechnen. Dabei wird, da nicht der Grenzwert für $h \to 0$ berechnet wird, ein Fehler gemacht. Es stellt sich nun die Frage, wie groß der auftretende Fehler ist und wie h zu wählen ist, damit de Fehler möglichst gering ist. Es gibt zwei Arten von Fehlern die sich überlagern.

1. **Approximationsfehler:** Für ein festes $h > 0$ gilt

$$\left| \frac{f(x+h)-f(x)}{h} - \lim_{\Delta \to 0} \frac{f(x+\Delta)-f(x)}{\Delta} \right|$$

Der Fehler ist abhängig von h und f, und wird in der Regel kleiner, wenn h reduziert wird.

2. **Rundungsfehler:** Siehe auch Abschnitt 2.4.

Wir nutzen nicht x sondern $\mathrm{rd}(x)$, wobei $\mathrm{rd}(x)$ die nächste darstellbare Zahl ist. Es gilt $\mathrm{rd}(x) := x \cdot (1+\varepsilon) = x + \varepsilon x$ mit $|\varepsilon| \leq \mathsf{eps}$. Die Konstante eps ist die relative Maschinengenauigkeit der gewählten Zahlendarstellung.

Wir untersuchen kurz, was bei der Berechnung von (5.4.1) oder (5.4.2) passiert. Da $\lim_{h\to 0}(f(x+h)) - f(x) = 0)$ gilt, wird mit fallenden h der Wert im Zähler immer kleiner und damit, da eps vorgegeben ist, der relative Fehler immer größer. Zusätzlich wird durch die Division durch den kleinen Nenner auch der absolute Fehler immer größer und der Rundungsfehler wächst für kleiner werdende h.
Damit kann man zwei gegenläufige Tendenzen beobachten, einen mit fallendem h fallenden Approximationsfehler und einen mit fallendem h wachsenden Rundungsfehler.

Beispiel 5.6. *Wir betrachten die Funktion* $f(x) = \exp(x)$ *und berechnen* $\exp'(1)$. *Es ist bekannt, dass* $\exp(1) = \exp'(1) = 2.7182818\ldots$ *gilt. Die folgende Tabelle zeigt das Resultat der Berechnung*

$$\exp'(1) \approx \frac{\exp(1+h)-\exp(1)}{h} ,$$

die mit der Software octave *durchgeführt wurde, für verschiedene Werte von* h.

h	Berechneter Wert	Fehler
$1{,}0 \cdot 10^{+0}$	4,670 774 270 471 605 8	$1{,}952\,492\,442\,012\,560\,7 \cdot 10^{+0}$
$1{,}0 \cdot 10^{-2}$	2,731 918 655 787 124 5	$1{,}363\,682\,732\,807\,935\,9 \cdot 10^{-2}$
$1{,}0 \cdot 10^{-4}$	2,718 417 747 082 924 1	$1{,}359\,186\,238\,789\,611\,4 \cdot 10^{-4}$
$1{,}0 \cdot 10^{-6}$	2,718 283 187 430 614 1	$1{,}358\,971\,569\,054\,290\,3 \cdot 10^{-6}$
$1{,}0 \cdot 10^{-8}$	2,718 281 821 856 293 9	$-6{,}602\,751\,234\,652\,259\,9 \cdot 10^{-9}$
$1{,}0 \cdot 10^{-10}$	2,718 283 376 168 527 9	$1{,}547\,709\,482\,796\,477\,7 \cdot 10^{-6}$
$1{,}0 \cdot 10^{-12}$	2,718 714 142 702 082 9	$4{,}323\,142\,430\,378\,013\,0 \cdot 10^{-4}$
$1{,}0 \cdot 10^{-14}$	2,708 944 180 085 381 5	$-9{,}337\,648\,373\,663\,576\,3 \cdot 10^{-3}$
$1{,}0 \cdot 10^{-16}$	0,000 000 000 000 000 0	$-2{,}718\,281\,828\,459\,045\,1 \cdot 10^{+0}$

Ähnliche Ergebnisse erhält man für die Berechnung

$$\exp'(1) \approx \frac{\exp(1+h) - \exp(1-h)}{2h} ,$$

wie die folgende Tabelle zeigt.

h	Berechneter Wert	Fehler
$5{,}0 \cdot 10^{-1}$	2,832 967 799 637 936 3	$1{,}146\,859\,711\,788\,912\,3 \cdot 10^{-1}$
$5{,}0 \cdot 10^{-3}$	2,718 293 154 647 444 3	$1{,}132\,618\,839\,916\,332\,8 \cdot 10^{-5}$
$5{,}0 \cdot 10^{-5}$	2,718 281 829 592 328 3	$1{,}133\,283\,245\,025\,040\,7 \cdot 10^{-9}$
$5{,}0 \cdot 10^{-7}$	2,718 281 828 517 632 0	$5{,}858\,691\,309\,867\,936\,1 \cdot 10^{-11}$
$5{,}0 \cdot 10^{-9}$	2,718 281 821 856 293 9	$-6{,}602\,751\,234\,652\,259\,9 \cdot 10^{-9}$
$5{,}0 \cdot 10^{-11}$	2,718 283 376 168 527 9	$1{,}547\,709\,482\,796\,477\,7 \cdot 10^{-6}$
$5{,}0 \cdot 10^{-13}$	2,718 270 053 492 232 8	$-1{,}177\,496\,681\,226\,131\,2 \cdot 10^{-5}$
$5{,}0 \cdot 10^{-15}$	2,753 353 101 070 387 8	$3{,}507\,127\,261\,134\,268\,5 \cdot 10^{-2}$
$5{,}0 \cdot 10^{-17}$	0,000 000 000 000 000 0	$-2{,}718\,281\,828\,459\,045\,1 \cdot 10^{+0}$

Die Resultate bestätigen unsere Erwartungen. Für kleine und große Werte ist die Approximation der Ableitung sehr ungenau. Die beste Approximation erreichen wir im Bereich von 10^{-8} für die erste Approximation und $5 \cdot 10^{-7}$ für die zweite Approximation.

Aus dem bisherigen Beobachtungen kann man folgende generellen Aussagen ableiten:

1. Für eine feste Zahlendarstellung kann nur eine bestimmte funktionsabhängige Genauigkeit erreicht werden.

2. Das optimale h hängt von der Zahlendarstellung, der Funktion und der Approximation der Ableitung ab. In unserem Fall sind $h \approx \sqrt{\mathsf{eps}} \approx 10^{-8}$ beziehungsweise $h \approx \sqrt[3]{\mathsf{eps}} \approx 10^{-5}$ gute Werte.

Die ausführlichere Analyse von Approximationfehlern ist Thema der Numerik und geht über den Inhalt dieser Vorlesung weit hinaus.

5.5 Lokale Extrema und Mittelwertsätze

Wir wollen nun mit Hilfe der Ableitungen von Funktionen deren Eigenschaften analysieren. Dazu wird eine Reihe von Sätzen hergeleitet, die das Verhalten von Funktionen auf Intervallen beschreiben. Zuerst definieren wir den Begriffe des lokalen Extremums.

Definition 5.9 (Lokale Extrema)**.** *Sei $f : (a, b) \to \mathbb{R}$ eine Funktion, dann hat f in $x \in (a, b)$ ein* lokales Maximum *(bzw.* lokales Minimum*), wenn ein $\varepsilon > 0$ existiert, sodass $f(x) \geq f(y)$ (bzw. $f(x) \leq f(y)$) für alle y mit $|x - y| < \varepsilon$ gilt.*
Gilt sogar $f(x) > f(y)$ (bzw. $f(x) < f(y)$) für alle $y \neq x$ mit $|x - y| < \varepsilon$ so spricht man von einem strikten lokalen Maximum *(bzw.* striktem lokalen Minimum*).*

Extremum ist der Oberbegriff für Minimum und Maximum. Während ein striktes lokales Extremum beschreibt, dass ein Funktionswert größer oder kleiner als alle Punkte in der Umgebung ist, wird der Begriff striktes absolutes Extremum (oder striktes globales Extremum) für Punkte benutzt, deren Funktionswerte größer oder kleiner als alle anderen Funktionswerte sind. Für $f : (a, b) \to \mathbb{R}$ ist $x \in (a, b)$ ein striktes absolutes Maximum falls $f(x) > f(y)$ für alle $y \in (a, b)$ und $y \neq x$. Analog kann man strikte absolute Minima, sowie absolute Minima und Maxima definieren. Jedes (strikte) absolute Extremum ist gleichzeitig auch (striktes) lokales Extremum. Wir werden Bedingungen für lokale Extrema herleiten, da nur diese auf Basis der Ableitung, die das Verhalten einer Funktion in der Umgebung eines Punktes beschreiben, bestimmt werden können. Die Bestimmung globaler Minima ist nur in bestimmten Funktionsklassen möglich und geht über den Inhalt dieser Vorlesung hinaus.

Satz 5.10 (Notwendige Bedingung für lokale Extrema)**.** *Die Funktion $f : (a, b) \to \mathbb{R}$ besitze im Punkt $x \in (a, b)$ ein lokales Extremum und sei in x differenzierbar. Dann ist $f'(x) = 0$.*

Beweis: Die Funktion f in x ein lokales Maximum, dann existiert $\varepsilon > 0$, sodass $f(y) \leq f(x)$ für alle $y \in (x - \varepsilon, x + \varepsilon)$ gilt. Damit gilt auch

$$f'(x) = \lim_{y \to x} \frac{f(y) - f(x)}{y - x} = \underbrace{\lim_{y \nearrow x} \frac{f(y) - f(x)}{y - x}}_{\geq 0} = \underbrace{\lim_{y \searrow x} \frac{f(y) - f(x)}{y - x}}_{\leq 0}$$

Woraus $f'(x) = 0$ folgt, da beide Grenzwerte für differenzierbare Funktionen übereinstimmen.
Der Beweis für lokale Minima ist analog zu führen. □

Die Bedingung $f'(x) = 0$ ist nicht hinreichend, wie die Funktion $f(x) = x^3$ zeigt, für die in $x = 0$ zwar $f'(0) = 0$ gilt, die aber im Punkt 0 kein Extremum hat.
Ferner ist zu beachten, dass wir die Funktion f über dem offenen Intervall (a, b) untersucht haben. Wenn wir $f : [a, b] \Rightarrow \mathbb{R}$ untersuchen, so folgt aus Satz 4.21 zwar, dass die Funktion, falls sie stetig ist, ihre Minimum und Maximum im Intervall annimmt. Wenn das Minimum oder Maximum aber in einem der beiden Randpunkte angenommen wird, so ist dort die erste Ableitung nicht notwendigerweise gleich 0, wie man sich leicht überlegen kann.
Die nun folgenden Sätze beschäftigen sich mit dem Verhalten von Funktionen in Intervallen, wenn Informationen über die Randpunkte vorliegen. Wir betrachten dabei oft Funktionen die auf dem abgeschlossenen Intervall $[a, b]$ definiert sind und im offenen Intervall (a, b) differenzierbar sind. Damit $f'(x)$ nur für $x \in (a, b)$ definiert. Um die Intervallgrenzen einzubeziehen, müssten wir den Ableitungsbegriff erweitern, was wir aber im Rahmen dieser Vorlesung nicht tun werden.

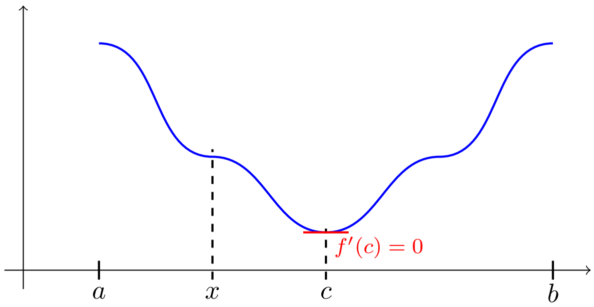


Abbildung 5.1: Veranschaulichung von Satz 5.11

Satz 5.11 (Satz von Rolle)**.** *Sei $a < b$ und $f : [a, b] \to \mathbb{R}$ eine stetige Funktion mit $f(a) = f(b)$. Die Funktion f sei in (a, b) differenzierbar. Dann existiert eine $c \in (a, b)$ mit $f'(c) = 0$.*

Beweis: Falls f konstant, so gilt $f'(c) = 0$ für alle $c \in (a, b)$.
Ist f nicht konstant, so gibt es ein $x \in (a, b)$ mit $f(x) < f(a)$ oder $f(x) > f(a)$. Damit wird ein (absolutes) Extremum in einem Punkt $c \in (a, b)$ angenommen und $f'(c) = 0$ (nach Satz 5.10). □

Satz 5.12 (Mittelwertsatz)**.** *Sei $a < b$ und $f : [a, b] \to \mathbb{R}$ eine stetige Funktion, die in (a, b) differenzierbar ist. Dann existiert ein $c \in (a, b)$, sodass $\frac{f(b)-f(a)}{b-a} = f'(c)$.*

Beweis: Definiere eine Funktion $g : [a, b] \to \mathbb{R}$, mit $g(x) = f(x) - \frac{f(b)-f(a)}{b-a}(x - a)$. g ist stetig in $[a, b]$ und differenzierbar in (a, b) und $g(b) = f(a) = g(a)$. Dann existiert nach Satz 5.11 ein $c \in (a, b)$ mit $g'(c) = 0$. Aus $g'(c) = f'(c) - \frac{f(b)-f(a)}{b-a}$ folgt die Behauptung. □

Den Mittelwertsatz kann man geometrisch interpretieren. Die Steigung der Sekante durch die Punkte $(a, f(a))$ und $(b, f(b))$ entspricht der Steigung der Tangente im Punkt $(c, f(c))$, wie die folgende Graphik zeigt.

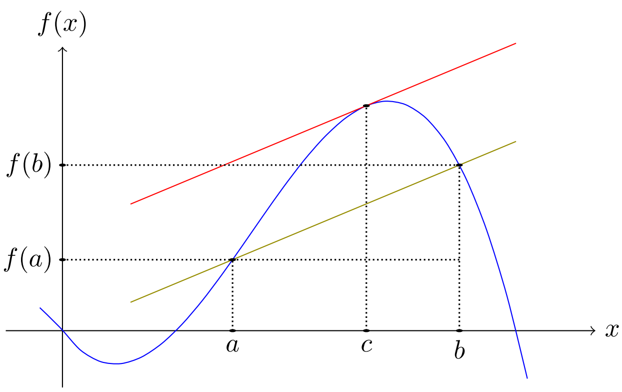


Abbildung 5.2

Aus dem Mittelwertsatz folgt unmittelbar das folgende Korollar, mit dem man Schranken für das Wachstum einer Funktion herleiten kann.

Korollar 5.13. *Sei $f : [a, b] \to \mathbb{R}$ eine stetige und in (a, b) differenzierbare Funktion. Für die Ableitung gelte $K^- \leq f'(x) \leq K^+$ für alle $x \in (a, b)$ mit $K^-, K^+ \in \mathbb{R}$. Für alle $c, d \in [a, b]$ mit $c \leq d$ gilt $K^-(d - c) \leq f(d) - f(c) \leq K^+(d - c)$*

Mit Hilfe der Ableitungen lassen sich auch Aussagen über die Monotonie von Funktionen gewinnen.

Satz 5.14 (Monotonie von Funktionen)**.**
Sei $f : [a, b] \to \mathbb{R}$ stetig und in (a, b) differenzierbar. Dann gilt:

i) $\forall x \in (a, b) : f'(x) \geq 0 \Leftrightarrow f$ *monoton wachsend in* $[a, b]$

ii) $\forall x \in (a, b) : f'(x) > 0 \Rightarrow f$ *<u>streng</u> monoton wachsend in* $[a, b]$

iii) $\forall x \in (a, b) : f'(x) \leq 0 \Leftrightarrow f$ *monoton fallend in* $[a, b]$

iv) $\forall x \in (a, b) : f'(x) < 0 \Rightarrow f$ *<u>streng</u> monoton fallend in* $[a, b]$

Beweis:

i) Sei $c, d \in [a, b]$ mit $c < d$. Der Mittelwertsatz (Satz 5.12) angewendet auf das Intervall $[c, d]$ liefert ein $y \in (c, d)$ mit $f'(y) = \frac{f(d)-f(c)}{d-c}$.
Sei nun $f'(x) \geq 0$ für alle $x \in (a, b)$, dann folgt $\frac{f(d)-f(c)}{d-c} \geq 0$ für alle $d > c$, $c, d \in [a, b]$ und damit auch $f(d) - f(c) \geq 0$ und die Funktion ist monoton wachsend. Damit ist eine Richtung bewiesen.

Nehmen wir nun an, dass f monoton wachsend in $[a, b]$ ist. Dann gilt für jedes $y \in (a, b)$ $f'(y) = \lim\limits_{x \to y} \frac{f(x)-f(y)}{x-y} \geq 0$. Daraus folgt durch den Grenzübergang $f'(y) \geq 0$ und damit der Beweis für die andere Richtung.

Die Beweise für *ii)*, *iii)* und *iv)* sind analog zu führen, wobei zu beachten ist, dass in den Fällen *ii)* und *iv)* nur eine Richtung bewiesen werden muss und kann. □

Es sollte beachtet werden, dass in dem vorherigen Satz nur die beiden Aussagen *i)* und *iii)* genau-dann-wenn-Beziehungen sind, d. h. dass die Folgerungen in beide Richtungen gelten. Für die beiden Punkte *ii)* und *iv)* gilt nur die Folgerung von links nach rechts. Dies kann man sich leicht anhand der Funktion $f(x) = x^3$ überlegen. Zwar ist mit $f'(0) = 0$ für diese Funktion, trotzdem ist die Funktion streng monoton wachsend.
Aus den vorherigen Sätzen können wir nun eine hinreichende Bedingung für strenge lokale Extrema herleiten.

Satz 5.15 (Strenges lokales Maximum/Minimum)**.** *Sei $f : (a, b) \to \mathbb{R}$ eine differenzierbare Funktion, die im Punkt $x \in (a, b)$ zweimal differenzierbar ist. Falls $f'(x) = 0$ und $f''(x) > 0$ (bzw. $f''(x) < 0$), dann besitzt f in x ein strenges lokales Minimum (bzw. Maximum).*

Beweis: Wir beweisen den Satz für strenge lokale Minima. Sei $f''(x) > 0$. Da $f''(x) = \lim\limits_{y \to x} \frac{f'(y)-f'(x)}{y-x} > 0$ ist, existiert ein $\varepsilon > 0$, sodass $\frac{f'(y)-f'(x)}{y-x} > 0$ für alle $y \in (x - \varepsilon, x + \varepsilon)$. Da ferner $f'(x) = 0$ folgt $f'(y) < 0$ für $y \in (x - \varepsilon, x)$ und $f'(y) > 0$

für $y \in (x, x+\varepsilon)$. Nach Satz 5.14 *iii*), *iv*) ist f damit streng monoton fallend in $(x-\varepsilon, x)$ und streng monoton wachsend in $(x, x+\varepsilon)$. Damit muss f in x ein strenges lokales Minimum besitzen.
Der Beweis für strenge lokale Maxima ist analog zu führen. □

Die Bedingungen sind nur hinreichend, aber nicht notwendig. Wie man an der Funktion $f(x) = x^4$ sieht. Diese besitzt in $x = 0$ ein strenges Minimum, trotzdem ist $f''(0) = 0$. Eine hinreichendes Kriterium für lokale Extrema liefert der folgende Satz, den wir ohne Beweis einführen Der Beweis basiert auf der Taylor-Reihenentwicklung, die wir in Kapitel 7 kurz einführen werden.

Satz 5.16. *Sei $f : (a,b) \to \mathbb{R}$ eine differenzierbare Funktion, die im Punkt $x \in (a,b)$ $n+1$mal differenzierbar ist. Falls $f'(x) = f^{(2)}(x) = \ldots = f^{(n)}(x) = 0$ und $f^{(n+1)}(x) \neq 0$, dann besitzt f in x*

i) ein strenges lokales Minimum, falls n ungerade ist und $f^{(n+1)}(x) > 0$,

ii) ein strenges lokales Maximum, falls n ungerade ist und $f^{(n+1)}(x) < 0$,

iii) kein Extremum, falls n gerade ist.

Ein Punkt dem sich das Vorzeichen der zweiten Ableitung ändert bezeichnet man auch als Wendepunkt. Formal bedeutet dies, dass Funktion f in $x \in A$ einen Wendepunkt besitzt, falls $f''(x) = 0$ und $f''(y) < 0$ sowie $f''(z) > 0$ für entweder $y \in (x-\varepsilon, x)$, $z \in (x, x+\varepsilon)$ oder $z \in (x-\varepsilon, x)$, $y \in (x, x+\varepsilon)$, sowie $\varepsilon > 0$.
Eine wichtige Eigenschaft von Funktionen ist die Konvexität, die wie folgt definiert wird.

Definition 5.17. *Sei $(a,b) \subset \mathbb{R}$. Eine Funktion $f : (a,b) \to \mathbb{R}$ heißt* konvex, *wenn für alle $x_1, x_2 \in (a,b)$ und alle $\lambda \in (0,1)$ gilt*

$$f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda) f(x_2).$$

Die Funktion heißt konkav, *falls $-f$ konvex ist.*

Abbildung 5.3 zeigt die Skizze einer konvexen Funktion. Die Funktion „hängt nach unten durch“, d.h. die Funktionswerte liegen unterhalb der Verbindungsgeraden zwischen den Funktionswerten an den Intervallgrenzen. Konvexe Funktionen haben einige sehr angenehme Eigenschaften, so fallen zum Beispiel lokale und globale Minima bei diesen Funktionen zusammen. Dies kann bei der Optimierung, d.h. dem Finden von globalen Extrema, genutzt werden.

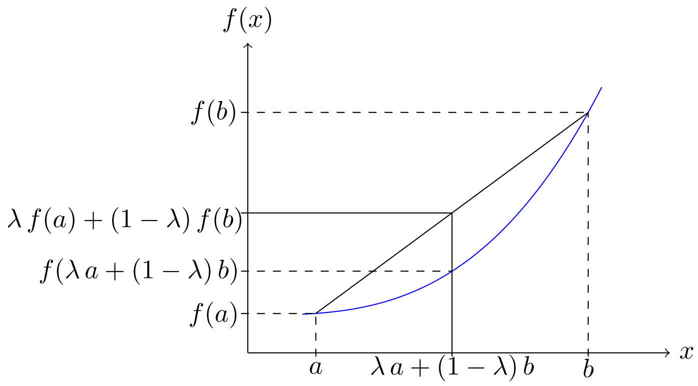


Abbildung 5.3: Skizze einer konvexen Funktion

Der folgende Satz stellt eine Zusammenhang zwischen der Konvexität und der zweiten Ableitung einer Funktion her.

Satz 5.18. *Sei $f : (a, b) \to \mathbb{R}$ eine zweimal differenzierbare Funktion. f ist genau dann konvex, wenn $f''(x) \geq 0$ für alle $x \in (a, b)$.*

Beweis: Sei $f''(x) \geq 0$ für all $x \in (a, b)$. Dann ist f' nach Satz 5.14 monoton wachsend. Sei $x = \lambda x_1 + (1-\lambda)x_2$ für $x_1, x_2 \in (a, b)$ mit $x_1 < x_2$ und $0 < \lambda < 1$, dann gilt offensichtlich $x_1 < x < x_2$. Nach Satz 5.12 existieren dann $y_1 \in (x_1, x)$ und $y_2 \in (x, x_2)$ mit

$$\frac{f(x) - f(x_1)}{x - x_1} = f'(y_1) \leq f'(y_2) = \frac{f(x_2) - f(x)}{x_2 - x}.$$

Da $x - x_1 = (1-\lambda)(x_2 - x_1)$ und $x_2 - x = \lambda(x_2 - x_1)$ folgt

$$\frac{f(x) - f(x_1)}{1 - \lambda} \leq \frac{f(x_2) - f(x)}{\lambda} \quad \Rightarrow f(x) \leq \lambda f(x_1) + (1-\lambda) f(x_2).$$

Damit ist die Funktion konvex.
Sei nun f konvex. Angenommen es gäbe ein $x_0 \in (a, b)$ mit $f''(x_0) < 0$. Sei $c = f'(x_0)$ und $g(x) = f(x) - c(x - x_0)$. g ist zweimal differenzierbar und $g'(x_0) = 0$ sowie $g''(x_0) = f''(x_0) < 0$. Damit besitzt g nach Satz 5.15 an der Stelle x_0 ein strenges lokales Maximum. Es gibt damit ein $\varepsilon > 0$, so dass $(x_0 - \varepsilon, x_0 + \varepsilon) \subset (a, b)$ und $g(x_0 - \varepsilon) < g(x_0)$, $g(x_0 + \varepsilon) < g(x_0)$. Aufgrund der Definition von g folgt damit auch

$$f(x_0) = g(x_0) > \frac{1}{2}\left(g(x_0 - \varepsilon) + g(x_0 + \varepsilon)\right) = \frac{1}{2}\left(f(x_0 - \varepsilon) + f(x_0 + \varepsilon)\right).$$

Wähle nun $x_1 = x_0 - \varepsilon$, $x_2 = x_0 + \varepsilon$ und $\lambda = \frac{1}{2}$. Dann $x_0 = \lambda x_1 + (1-\lambda)x_2$ und

$$f(\lambda x_1 + (1-\lambda)x_2) > \lambda f(x_1) + (1-\lambda) f(x_2).$$

Dies steht im Widerspruch zur Konvexität von f. □

Wir leiten nun noch einige Ergebnisse her, die es manchmal erlauben Grenzwerte von Funktionen einfacher zu bestimmen. Dazu beginnen wir mit einer Erweiterung des Mittelwertsatzes.

Satz 5.19 (Zweiter Mittelwertsatz). *Seien $f : [a,b] \to \mathbb{R}$ und $g : [a,b] \to \mathbb{R}$ zwei Funktionen, die auf $[a,b]$ stetig und auf (a,b) differenzierbar sind. Sei ferner $g'(x) \neq 0$ für alle $x \in (a,b)$, dann ist $g(a) \neq g(b)$ und es existiert ein $c \in (a,b)$ mit $\frac{f(b)-f(a)}{g(b)-g(a)} = \frac{f'(c)}{g'(c)}$.*

Beweis:

i) Wir beweisen zuerst $g(a) \neq g(b)$.
Wäre $g(a) = g(b)$, so gäbe es nach Satz 5.11 (Satz von Rolle) ein $c \in (a,b)$ mit $g'(c) = 0$, was aber laut Voraussetzung ausgeschlossen ist.

ii) Für die Hilfsfunktion $h(x) = f(x) - \frac{f(b)-f(a)}{g(b)-g(a)}(g(x)-g(a))$ gilt $h(a) = h(b) = f(a)$. Damit existiert nach Satz 5.11 (Satz von Rolle) $c \in (a,b)$ mit $h'(c) = 0$ und somit gilt $h'(c) = f'(c) - \frac{f(b)-f(a)}{g(b)-g(a)}g'(c) = 0$.

Da $g'(c) \neq 0$ folgt die Behauptung. □

Eine Anwendung des zweiten Mittelwertsatzes ist die Regel von L'Hospital.

Satz 5.20 (Regel von L'Hospital $\frac{0}{0}$). *Seien $f : [a,b] \to \mathbb{R}$ und $g : [a,b] \to \mathbb{R}$ zwei auf $[a,b]$ stetige und auf (a,b) differenzierbare Funktionen. Sei $c \in [a,b]$ und $g'(x) \neq 0$ für $x \in (a,b) \setminus \{c\}$.*
Gilt $\lim\limits_{x\to c} f(x) = 0$ und $\lim\limits_{x\to c} g(x) = 0$ und existiert $\lim\limits_{x\to c} \frac{f'(x)}{g'(x)} \in \mathbb{R}$, so existiert auch $\lim\limits_{x\to c} \frac{f(x)}{g(x)}$ und es gilt $\lim\limits_{x\to c} \frac{f(x)}{g(x)} = \lim\limits_{x\to c} \frac{f'(x)}{g'(x)}$.

Beweis: Wir definieren zwei Funktionen

$$F(x) = \begin{cases} f(x) & \text{falls } x \in [a,b] \setminus \{c\} \\ 0 & \text{falls } x = c \end{cases} \quad \text{und } G(x) = \begin{cases} g(x) & \text{falls } x \in [a,b] \setminus \{c\} \\ 0 & \text{falls } x = c \end{cases}$$

F und G sind stetig auf $[a,b]$ und differenzierbar in $(a,b) \setminus \{c\}$, da f und g diese Eigenschaften aufweisen. Ferner ist $F'(x) = f'(x)$ und $G'(x) = g'(x) \neq 0$ für alle $x \in (a,b) \setminus \{c\}$.
Wir zeigen zuerst, dass $G(x) \neq 0$ für alle $x \in (a,b) \setminus \{c\}$ gilt. Gäbe es ein $x \in (a,b) \setminus \{c\}$ mit $G(x) = G(c) = 0$, so würde nach Satz 5.11 angewendet auf das Intervall $[x,c]$ falls $x < c$ (und $[c,x]$ falls $x > c$) folgen, dass ein $y \in [x,c]$ mit $G'(y) = 0$ und damit $g'(y) = 0$ existiert, was aber der Annahme $g'(y) \neq 0$ für alle $y \in (a,b)$ widerspricht.
Sei nun $x_0 \in (a,b) \setminus \{c\}$ vorgegeben. Wir nehmen an, dass $x_0 < c$, und betrachten das Intervall $[x_0, c]$. Der Fall $x_0 > c$ ist analog zu behandeln. Es existiert dann ein $x_1 \in (x_0, c)$, sodass

$$\frac{f(x_0)}{g(x_0)} = \underbrace{\frac{F(x_0) - F(c)}{G(x_0) - G(c)}}_{\text{da } F(c)=G(c)=0} \underset{(\text{Satz } 5.19)}{=} \frac{F'(x_1)}{G'(x_1)} = \frac{f'(x_1)}{g'(x_1)}$$

Ausgehend von x_1 kann man nun ein $x_2 \in (x_1, c)$ finden, so dass $\frac{f(x_1)}{g(x_1)} = \frac{F(x_1)-F(c)}{G(x_1)-G(c)} = \frac{F'(x_2)}{G'(x_2)} = \frac{f'(x_2)}{g'(x_2)}$. Diesen Prozess kann man beliebig weiter fortführen. Sei $\xi(x_i)$ eine Funktion, die zu x_i den Wert von $x_{i+1} \in (x_i, c)$ berechnet.

Die Funktion ξ kann man auf beliebige $x \in (a,c)$ anwenden und es gilt $\frac{f(x)}{g(x)} = \frac{f'(\xi(x))}{g'(\xi(x))}$ und $|c - \xi(x)| < |c - x|$. Damit gilt auch $\lim\limits_{x\to c} \xi(x) = c$ und somit auch

$$\lim_{x\to c} \frac{f(x)}{g(x)} = \lim_{x\to c} \frac{F(x) - F(c)}{G(x) - G(c)} = \lim_{x\to c} \frac{f'(\xi(x))}{g'(\xi(x))} = \frac{f'(c)}{g'(c)} \,. \qquad \square$$

Bevor wir eine zweite Version der Regel von L'Hospital betrachten, muss der Begriff des Grenzwertes erweitert werden. Wir definieren dazu uneigentliche Grenzwerte.

Definition 5.21 (Uneigentlicher Grenzwert)**.** *Sei $f : A \to \mathbb{R}$ und a ein Häufungspunkt von A. Falls für alle $K \in \mathbb{R}$ ein $\delta > 0$ existiert, sodass $f(x) > K$ für $|x - a| < \delta$, so schreibt man $\lim\limits_{x\to a} f(x) = \infty$.*
Falls $\lim\limits_{x\to a} -f(x) = \infty \Leftrightarrow \lim\limits_{x\to a} f(x) = -\infty$.

Wie der Name schon ausdrückt, ist ein uneigentlicher Grenzwert kein richtiger Grenzwert, da keine Konvergenz vorliegt. Uneigentliche Grenzwerte sind aber die Grundlage für die folgende zweite Regel von L'Hospital.

Satz 5.22 (Regel von L'Hospital $\frac{\infty}{\infty}$)**.** *Seien $f : [a,b] \to \mathbb{R}$ und $g : [a,b] \to \mathbb{R}$ zwei auf $[a,b]$ stetig und auf (a,b) differenzierbare Funktionen. Sei $c \in [a,b]$ und $g'(x) \neq 0$ für $x \in (a,b) \setminus \{c\}$.*
Gilt $\lim\limits_{x\to c} f(x) = \infty$ und $\lim\limits_{x\to c} g(x) = \infty$ und existiert $\lim\limits_{x\to c} \frac{f'(x)}{g'(x)}$, so existiert auch $\lim\limits_{x\to c} \frac{f(x)}{g(x)}$ und es gilt $\lim\limits_{x\to c} \frac{f(x)}{g(x)} = \lim\limits_{x\to c} \frac{f'(x)}{g'(x)}$.

Beweis: Wir beschränken uns im Beweis auf den Fall $f(x) \neq 0$ für alle $x \in [a,b]\setminus\{c\}$. Definiere $F(x) = \frac{1}{g(x)}$ und $G(x) = \frac{1}{f(x)}$, dann ist

$$\frac{f(x)}{g(x)} = \frac{\frac{1}{g(x)}}{\frac{1}{f(x)}} = \frac{F(x)}{G(x)}$$

und $\lim\limits_{x\to c} F(x) = \lim\limits_{x\to c} G(x) = 0$, so dass Satz 5.20 anwendbar ist.
Also gilt

$$\lim_{x\to c} \frac{f(x)}{g(x)} = \lim_{x\to c} \frac{F(x)}{G(x)} = \lim_{x\to c} \frac{F'(x)}{G'(x)} = \lim_{x\to c} \frac{\frac{-g'(x)}{(g(x))^2}}{\frac{-f'(x)}{(f(x))^2}} = \lim_{x\to c} \left(\frac{f(x)}{g(x)}\right)^2 \cdot \frac{g'(x)}{f'(x)}$$

da $\left(\frac{1}{f(x)}\right)' = \frac{-f'(x)}{(f(x))^2}$. Damit gilt auch

$$1 = \lim_{x\to c} \left(\frac{f(x)}{g(x)}\right) \cdot \frac{g'(x)}{f'(x)} \Rightarrow \lim_{x\to c} \frac{f(x)}{g(x)} = \lim_{x\to c} \frac{f'(x)}{g'(x)} \,. \qquad \square$$

Es gibt einige weitere Varianten der Regel von L'Hospital, die wir kurz ohne Beweis vorstellen wollen.
Seien $f : [a,\infty] \to \mathbb{R}$ und $g : [a,\infty] \to \mathbb{R}$ zwei auf (a,∞) differenzierbare Funktionen, und sei $g(x) \neq 0$ für $x \in (a,\infty)$.

i) Sei $\lim\limits_{x\to\infty} f(x) = \lim\limits_{x\to\infty} g(x) = 0$ und $\lim\limits_{x\to\infty} \frac{f'(x)}{g'(x)} \in \mathbb{R}$ existiert, dann existiert auch $\lim\limits_{x\to\infty} \frac{f(x)}{g(x)} \in \mathbb{R}$ und es gilt $\lim\limits_{x\to\infty} \frac{f(x)}{g(x)} = \lim\limits_{x\to\infty} \frac{f'(x)}{g'(x)}$.

ii) Sei $\lim\limits_{x\to\infty} f(x) = \lim\limits_{x\to\infty} g(x) = \infty$ und $\lim\limits_{x\to\infty} \frac{f'(x)}{g'(x)} \in \mathbb{R}$ existiert, dann gilt existiert auch und $\lim\limits_{x\to\infty} \frac{f(x)}{g(x)} \in \mathbb{R}$ es gilt $\lim\limits_{x\to\infty} \frac{f(x)}{g(x)} = \lim\limits_{x\to\infty} \frac{f'(x)}{g'(x)}$

Beispiel 5.7.

i) $f(x) = x^2, \quad g(x) = x$
$\lim\limits_{x\to 0} \frac{x^2}{x} = \lim\limits_{x\to 0} \frac{2x}{1} = 0$
$\lim\limits_{x\to\infty} \frac{x^2}{x} = \lim\limits_{x\to\infty} \frac{2x}{1} = \infty$

ii) $f(x) = e^x - 1, \quad g(x) = \sin x$
$\lim\limits_{x\to 0} \frac{e^x-1}{\sin x} = \lim\limits_{x\to 0} \frac{e^x}{\cos x} = \frac{1}{1} = 1$.

iii) $\lim\limits_{x\searrow 0} x \ln x = \lim\limits_{x\searrow 0} \frac{\ln x}{\frac{1}{x}} = \lim\limits_{x\searrow 0} \frac{\frac{1}{x}}{-\frac{1}{x^2}} = -\lim\limits_{x\searrow 0} \frac{x^2}{x} \overset{wie\ in\ i)}{=} 0$

Das dritte Beispiel zeigt, dass man die Regel auch für einseitige Grenzwerte anwenden kann. Ferner kann man die Regel auch rekursiv anwenden, wenn die Grenzwerte der ersten Ableitungen die Bedingungen der Regel von L'Hospital erfüllen. Es ist allerdings immer darauf zu achten, dass die Voraussetzungen erfüllt sind. Wenn die Grenzwerte der Funktionen $\neq 0$ (bzw. $\neq \infty$) sind, so gelten die Regeln nicht!

5.6 Kurvendiskussion

Zum Abschluss des Kapitels fassen wir einige der bisherigen Resultate noch einmal zusammen und nutzen sie zur Analyse von Funktionen in Form einer Kurvendiskussion. Sei $f : A \to \mathbb{R}$ eine in A differenzierbare Funktion, für die wir folgende Punkte untersuchen.

1. Symmetrie
2. Verhalten am Rand des Definitionsbereichs
3. Nullstellen
4. Extrempunkte
5. Wendepunkte
6. Funktionsgraph

Im Einzelnen werden die folgenden Aspekte dabei untersucht.
1) Symmetrie
Man unterscheidet Achsensymmetrie zur y-Achse, die vorliegt wenn $f(x) = f(-x)$ für alle $x \in A$ gilt.
Weiterhin wird die Punktsymmetrie zum Nullpunkt (bzw. Ursprung) untersucht, die vorliegt, wenn $f(-x) = -f(x)$ für alle $x \in A$ gilt.
Es ist zu beachten, dass für beide Arten von Symmetrie $x \in A \Leftrightarrow -x \in A$ eine notwendige Bedingung ist.
2) Verhalten am Rand des Definitionsbereichs

Interessant sind Häufungspunkte a von A, die nicht zu A gehören, sowie $-\infty$ und ∞, falls A nach unten oder oben unbeschränkt ist. Die folgenden Grenzwerte werden für Häufungspunkte untersucht:

$$\lim_{x \searrow a} f(x) \text{ oder } \lim_{x \nearrow a} f(x)$$

Dabei ist festzustellen, ob die Funktionen konvergieren oder der uneigentliche Grenzwert gegen $\pm\infty$ existiert. Für unbeschränkte Definitionsbereiche A werden die Grenzwerte

$$\lim_{x \to \infty} f(x) \text{ oder } \lim_{x \to -\infty} f(x)$$

untersucht. Auch hier wird wieder die Frage nach der Konvergenz oder der Existenz der uneigentlichen Grenzwerte gestellt.
Für die Fälle $\lim\limits_{x \to \infty} f(x)$ und $\lim\limits_{x \to -\infty} f(x)$ kann das asymptotische Verhalten analysiert werden. Dazu wird eine Gerade $g(x) = \alpha x + \beta$ $(\alpha, \beta \in \mathbb{R})$ gesucht, sodass

$$\lim_{x \to \infty} (f(x) - g(x)) = 0 \text{ bzw. } \lim_{x \to -\infty} (f(x) - g(x)) = 0$$

Die Gerade heißt Asymptote von f für $x \to \infty$ (bzw. $x \to -\infty$). Die folgenden Bedingungen charakterisieren eine Asymptote und erlauben eine direkte Bestimmung der Werte von α und β.

$$\begin{array}{ll} \lim\limits_{x \to \infty} \frac{f(x)}{x} = \alpha & \left(\text{bzw. } \lim\limits_{x \to -\infty} \frac{f(x)}{x} = \alpha\right) \\ \lim\limits_{x \to \infty} (f(x) - \alpha x) = \beta & \left(\text{bzw. } \lim\limits_{x \to -\infty} (f(x) - \alpha x) = \beta\right) \end{array}$$

Die zweite Bedingung folgt direkt aus der Bedingung an die Asymptote, die erste Bedingung folgt aus

$$\lim_{x \to \infty} \left(\frac{f(x) - g(x)}{x} \right) = \lim_{x \to \infty} \left(\frac{f(x) - \alpha x - \beta}{x} \right) = \lim_{x \to \infty} \left(\frac{f(x)}{x} - \alpha \right)$$

oder aus dem entsprechenden Grenzwert gegen $-\infty$.
Unter Umständen wird auch die senkrechte Gerade als Asymptote angesehen, wenn der Funktionswert gegen $-\infty$ oder ∞ konvergiert.

3) Bestimmung von Nullstellen

Eine Nullstelle liegt vor, wenn $f(x) = 0$. Methoden zur Berechnung von $f(x) = 0$ werden wir im nächsten Kapitel kennen lernen.

4) Berechnung von Extrempunkten

Eine notwendige Bedingung für einen lokalen Extrempunkte $a \in A$ lautet $f'(a) = 0$ falls eine Umgebung $(a - \varepsilon, a + \varepsilon) \subseteq A$ für ein $\varepsilon > 0$ existiert. Randpunkte von A, falls diese existieren, sind separat zu untersuchen, da dort die notwendige Bedingung nicht greift.
Eine hinreichende Bedingung für die Existenz eines lokalen Extremums im Punkt $a \in A$ lautet $f'(a) = 0$ sowie $f''(a) < 0$ für ein lokales Maximum oder $f''(a) > 0$ für ein lokales Minimum. Eine weitergehende hinreichende Bedingung haben wir in Satz 5.16 kennen gelernt.

5) Wendepunkt

Ein Wendepunkt liegt dann vor, wenn sich das Vorzeichen der zweiten Ableitung ändert. Ein Punkt $a \in A$ mit $(a-\varepsilon, a+\varepsilon) \subseteq A$ für $\varepsilon > 0$ ist ein Wendepunkt, falls

$$f''(x) \begin{cases} < 0 & \text{für } x \in (a-\varepsilon, a) \\ = 0 & \text{für } x = a \\ > 0 & \text{für } x \in (a, a+\varepsilon) \end{cases} \quad \text{oder} \quad f''(x) \begin{cases} > 0 & \text{für } x \in (a-\varepsilon, a) \\ = 0 & \text{für } x = a \\ < 0 & \text{für } x \in (a, a+\varepsilon) \end{cases}$$

Eine notwendige Bedingung für einen Wendepunkt in $x \in A$, wobei x kein Randpunkt von A sein darf, lautet $f''(x) = 0$.
Eine hinreichende Bedingung lautet $f''(x) = 0$ und $f^{(3)}(x) \neq 0$. Diese Bedingung ist nicht notwendig. Eine besondere Form des Wendepunkts ist der Sattelpunkt, bei dem zusätzlich noch die erste Ableitung gleich 0 sein muss.

6) Funktionsgraph

Aus den bisherigen Analysen lassen sich erste Aussagen über den Verlauf der Funktion machen. Diese können durch die Bestimmung des Verlauf der Funktion, durch punktweises Abtasten, und die graphische Darstellung noch ergänzt werden.

Beispiel 5.8. *Als Beispiel untersuchen wir die Funktion*

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

mit Definitionsbereich $A = \mathbb{R} \setminus \{-1\}$.
zu 1) *Da $1 \in A$ aber $-1 \notin A$, kann keine Symmetrie vorliegen.*
zu 2) *Interessant ist das Verhalten der Funktion für x gegen einen der Werte $-\infty, \infty, -1$.*
Im Punkt -1 gilt:
Der Nenner $(x+1)^2$ ist größer 0 für alle $x \in \mathbb{R} \setminus \{-1\}$. Der Zähler $x^2(x+2)$ ist größer 0 für $x > -2$. Damit ist $f(x)$ positiv für $x > -2$. Es gilt

$$\lim_{x \nearrow -1} f(x) = \lim_{x \searrow -1} f(x) = \infty$$

da $\lim\limits_{x \nearrow -1} x^2(x+2) = \lim\limits_{x \searrow -1} x^2(x+2) = 1$ und $\lim\limits_{x \nearrow -1} (x+1)^2 = \lim\limits_{x \searrow -1} (x+1)^2 = 0$.
Für die Grenzwerte x gegen ∞ oder $-\infty$ gilt

$$\begin{array}{rclclcl} \lim\limits_{x\to\infty} f(x) & = & \lim\limits_{x\to\infty} \frac{x^3+2x^2}{x^2+2x+1} & = & \lim\limits_{x\to\infty} \frac{x+2}{1+\frac{2}{x}\frac{1}{x^2}} & = & \infty \\ \lim\limits_{x\to-\infty} f(x) & = & \lim\limits_{x\to-\infty} \frac{x+2}{1+\frac{2}{x}\frac{1}{x^2}} & = & -\infty \end{array}$$

Die Asymptote für $x \to \infty$ wird wie folgt berechnet:

$$\begin{array}{l} \alpha = \lim\limits_{x\to\infty} \frac{x^2(x+2)}{x(x+1)^2} = \lim\limits_{x\to\infty} \frac{x^2+2x}{x^2+2x+1} = 1 \\ \beta = \lim\limits_{x\to\infty} \left(\frac{x^2(x+2)}{(x+1)^2} - x \right) = \lim\limits_{x\to\infty} \left(\frac{x^3+2^2}{x^2+2x+1} - \frac{x^3+2x^2+x}{x^2+2x+1} \right) = \lim\limits_{x\to\infty} \left(\frac{-x}{x^2+2x-1} \right) = 0 \end{array}$$

zu 3) *Es gilt $f(x) = \frac{x^2(x+2)}{(x+1)^2} = 0$ für $x = 0$ oder $x = -2$.*
Der Schnittpunkt mit der y-Achse liegt an der Stelle 0, da $f(0) = 0$.

zu 4) *Zur Bestimmung der Extremstellen berechnen wir die erste Ableitung mithilfe der Quotientenregel.*

$$\begin{aligned} f'(x) &= \tfrac{(3x^2+4x)(x+1)^2-2(x+1)(x^3+2x^2)}{(x+1)^4} \\ &= \tfrac{3x^3+4x^2+3x^2+4x-2x^3-4x^2}{(x+1)^3} \\ &= \tfrac{x^3+3x^2+4x}{(x+1)^3} \\ &= \tfrac{x(x^2+3x+4)}{(x+1)^3} \quad \text{für } x \in \mathbb{R} \setminus \{-1\} \end{aligned}$$

$f'(x)$ besitzt eine Nullstelle im Punkt 0. Für den Term in der Klammer des Zählers gilt $x^2+3x+4 = (x+\frac{3}{2})^2 + \frac{7}{4} \geq \frac{7}{4} > 0$ für alle $x \in \mathbb{R}$. Damit ist der Zähler > 0 für $x > 0$, $= 0$ für $x = 0$ und < 0 für $x < 0$. Der Nenner ist > 0 für $x > -1$ und < 0 für $x < -1$. sodass

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

Mithilfe von Satz 5.14 können wir die folgenden Aussagen über den Verlauf von f machen.

- *f ist streng monoton wachsend auf $(-\infty,-1)$*
- *f streng monoton fallend auf $(-1,0)$*
- *f ist streng monoton wachsend auf $(0,\infty)$*
- *f hat in $x = 0$ ein lokales Minimum*

Wie wir gleich sehen werden gilt $f''(0) = 2 > 0$, sodass auch die hinreichende Bedingung für das lokale Extremum erfüllt ist. Ein globales Extremum existiert nicht, da die Funktionswerte gegen $-\infty$ und ∞ konvergieren.

zu 5) *Zur Bestimmung der Wendepunkte bestimmen wir die zweite Ableitung.*

$$\begin{aligned} f''(x) &= \left(\tfrac{x(x^2+3x+4)}{(x+1)^3}\right)' \\ &= \tfrac{(3x^2+6x+4)(x+1)^3-(x^3+3x^2+4x)3(x+1)^2}{(x+1)^6} \\ &= \tfrac{3x^3+6x^2+4x+3x^2+6x+4-3x^3-9x^2-12x}{(x+1)^4} \\ &= \tfrac{-2x+4}{(x+1)^4} = \tfrac{2(2-x)}{(x+1)^4} \quad \text{für } x \in \mathbb{R} \setminus \{-1\} \end{aligned}$$

der Nenner ist im Definitionsbereich der Funktion positiv, während der Zähler $x < 2$ positiv und für $x > 2$ negativ ist. Im Punkt $x = 2$ hat $f''(x)$ eine Nullstelle. Damit gilt

$$f''(x) = \begin{cases} > 0 & \text{für } (-\infty,-1) \cup (-1,2) \\ = 0 & \text{für } x = 2 \\ < 0 & \text{für } (2,\infty) \end{cases}$$

Die Funktion hat an der Stelle $x = 2$ einen Wendepunkt. Für die dritte Ableitung gilt

$$\begin{aligned} f^{(3)}(x) &= \left(\tfrac{4-2x}{(x+1)^4}\right)' \\ &= \tfrac{-2(x+1)^4-(4-2x)\cdot 4\cdot(x+1)^3}{(x+1)^8} \\ &= \tfrac{-2(x+1)-(4-2x)4}{(x+1)^5} \\ &= \tfrac{6x-18}{(x+1)^5} = \tfrac{6(x-3)}{(x+1)^5} \quad \text{für } x \in \mathbb{R} \setminus \{-1\} \end{aligned}$$

Da $f''(2) = 0$ *und* $f^{(3)}(2) = \frac{-6}{3^5} = -\frac{1}{27} \neq 0$ *ist die hinreichende Bedingung für einen Wendepunkt erfüllt.*

zu 6) *Die folgende Graphik zeigt den Verlauf der Funktion im Intervall* $[-5, 5]$.