Analysis I für TPH

Eduard Nigsch

18. März 2025

Inhaltsverzeichnis

Vorwort

Dieses Skriptum entstand begleitend zur Vorlesung *Analysis I für TPH* im Wintersemester 2024/2025.

Obwohl es den behandelten Stoff möglichst sinnvoll abzubilden versucht, ersetzt es weder den Besuch der Vorlesung noch die intensive Beschäftigung mit einem oder mehreren geeigneten Lehrbüchern. Auf vieles, das in der Vorlesung nur verbal erklärt wird, wird hier schriftlich nicht weiter eingegangen. Außerdem lebt die Mathematik vom individuellen Nachvollziehen von Argumenten und Sachverhalten.

Der behandelte Stoff ist kanonisch; wer dieser Vorlesung folgt, sollte bald in der Lage sein, ein nahezu beliebiges Lehrbuch der Analysis studieren zu können. Nach Abschluss der Mathematik-Vorlesungen für Physiker sollte man in der Lage sein, sich auch in kompliziertere mathematische Sachverhalte selbst einzuarbeiten. Um so wichtiger ist es, sich möglichst früh mit der richtigen Denkweise und dem Schritt zum Abstrakten anzufreunden.

Was man an Voraussetzungen mitbringen sollte: elementare Algebra, reelle Funktionen, Differential- und Integralrechnung auf Maturaniveau, sowie die Inhalte der üblichen Vorkurse (101.748 Angleichungskurs Mathematik).

Großer Dank gebührt an dieser Stelle Philipp Huber, der einen wesentlichen Beitrag zur Enstehung dieses Skriptums geleistet hat.

Wien, Jänner 2025

Kapitel 1

Grundlagen

Lernziele

Aussagenlogik

- ► Wahrheitsgehalt von zusammengesetzten Aussagen über eine Wahrheitstafel bestimmen
- ► Aussagen (mit Quantoren, Konjunktion, Disjunktion und Implikation) negieren
- ► Aufzählende und beschreibende Schreibweise für Mengen ineinander überführen

Mengen

- ► Elementare Operationen mit Mengen durchführen
- ► Unterschied zwischen n-Tupel und Mengen erfassen

Abbildungen

- ► Bestandteile einer Abbildung benennen
- ► Eine Funktion auf Injektivität, Surjektivität und Bijektivität untersuchen
- ► Zusammengesetzte Funktionen auswerten
- ► Bild und Urbild einer Funktion bestimmen

Vollständige Induktion

- ► Aussagen über natürliche Zahlen durch Induktion beweisen
- ► Summen- und Produktschreibweise verwenden sowie Indexverschiebung und Zerlegung in Teilsummen/-Produkte durchführen
- ► Fakultät und Binomialkoeffizienten als Wahrscheinlichkeiten/Möglichkeiten interpretieren
- ► Binomischen Lehrsatz anwenden
- ► Geometrische Reihe berechnen

> Das Instrument, welches die Vermittlung bewirkt zwischen Theorie und Praxis, zwischen Denken und Beobachten, ist die Mathematik; sie baut die verbindende Brücke und gestaltet sie immer tragfähiger. Daher kommt es, daß unsere ganze gegenwärtige Kultur, soweit sie auf der geistigen Durchdringung und Dienstbarmachung der Natur beruht, ihre Grundlage in der Mathematik findet.
>
> D. Hilbert[1], 1930

Physikalische Gesetze werden mit Mitteln der Mathematik formuliert, Probleme mit mathematischen Methoden gelöst und damit z.B. Vorhersagen über die physikalische Welt getroffen.

Beispiel. Die Gleichung einer gleichförmig beschleunigten Bewegung (siehe Abbildung 1.1) ist

$$\ddot{r}(t) = \text{const.}$$

Mit Startpunkt $(x, y, z) = (0, 0, h)$ und Anfangsgeschwindigkeit $(v_{0x}, 0, v_{0z})$ sowie Beschleunigung $a = (0, 0, -g)$ erhält man die Wurfparabel

$$z(x) = -\frac{1}{2}\frac{g}{v_{0x}^2}x^2 + \frac{v_{0z}}{v_{0x}}x + h.$$

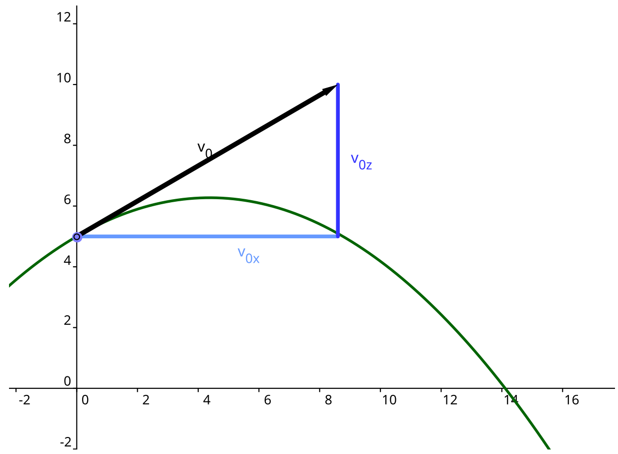


Abbildung 1.1: Wurfparabel einer gleichförmig beschleunigten Bewegung

Während dieses Beispiel noch mit Schulmathematik beherrschbar ist, verlangen anspruchsvollere Theorien der Physik eine stabilere mathematische Basis, die wir Schritt für Schritt entwickeln werden. Um die notwendige Sicherheit im Umgang damit zu erlangen, werden wir die meisten Beweise im Detail ausführen; außerdem lassen sich nur so Verständnis und Intuition hinreichend entwickeln.

[1] David Hilbert, 1862 – 1943; einer der bedeutendsten Mathematiker der Neuzeit, mit grundlegen Einflüssen auf die mathematische Physik.

1.1 Aussagenlogik

Gegenstand der Mathematik sind Aussagen über mathematische Sachverhalte; diese sind entweder wahr oder falsch, aber nie beides. Üblicherweise leitet man den Wahrheitswert einer Aussage aus bereits gesichertem Wissen oder angenommenen Voraussetzungen ab (Deduktion). Weiters beziehen sich mathematische Aussagen und Sätze üblicherweise auf einen bestimmten Gegenstandsbereich; dieser ist entweder aus dem Kontext klar oder er wird dezidiert angegeben.

Beispiel. Folgendes sind Aussagen, also wahr oder falsch:

(i) 5 ist eine natürliche Zahl.
(ii) Es regnet.
(iii) Jede Zahl ist eine Primzahl.
(iv) Die Gleichung $x^2 + 1 = 0$ hat keine Lösung.

Folgendes sind keine Aussagen:

(v) $5 + 2$.
(vi) n ist gerade.

(v) ist ein „Ausdruck“ oder „Term“ und (vi) ist eine sogenannte *Aussageform*, deren Wahrheitswert vom Wert von n abhängt.

Verknüfung von Aussagen

Zwei gegebene Aussagen, hier mit $\mathcal{A}$ bzw. $\mathcal{B}$ bezeichnet, kann man mit „und“ oder „oder“ verknüpfen sowie negieren. Die Aussage „$\mathcal{A}$ und $\mathcal{B}$“ (genannt *Konjunktion*), kurz geschrieben „$\mathcal{A} \wedge \mathcal{B}$“, ist wahr, wenn $\mathcal{A}$ und $\mathcal{B}$ beide wahr sind, und sonst falsch. Die Aussage „$\mathcal{A}$ oder $\mathcal{B}$“ (genannt *Disjunktion*), kurz „$\mathcal{A} \vee \mathcal{B}$“, ist wahr wenn mindestens eine der Aussagen $\mathcal{A}$ und $\mathcal{B}$ wahr ist, d.h. aber auch wenn beide wahr sind (also *kein entweder–oder*)! Schließlich ist die Aussage „nicht $\mathcal{A}$“, kurz $\neg\mathcal{A}$, genau dann wahr, wenn $\mathcal{A}$ falsch ist (*Negation*).

Für zwei Aussagen $\mathcal{A}$ und $\mathcal{B}$ erhält man eine neue Aussage, die *Implikation* $\mathcal{A} \implies \mathcal{B}$, *definiert* als $(\neg\mathcal{A}) \vee \mathcal{B}$, was bedeutet: wenn $\mathcal{A}$ wahr ist, dann ist auch $\mathcal{B}$ wahr. Man sagt in diesem Fall „aus $\mathcal{A}$ folgt $\mathcal{B}$“, „wenn $\mathcal{A}$, dann $\mathcal{B}$“ oder „$\mathcal{A}$ impliziert $\mathcal{B}$“. Eine weitere Sprechweise ist, dass $\mathcal{A}$ *hinreichend* für $\mathcal{B}$ ist oder $\mathcal{B}$ *notwendig* für $\mathcal{A}$.

Übersichtlich kann man das in einer *Wahrheitstabelle* (Tabelle 1.1) darstellen.

$\mathcal{A}$	$\mathcal{B}$	$\mathcal{A} \wedge \mathcal{B}$	$\mathcal{A} \vee \mathcal{B}$	$\neg\mathcal{A}$	$\mathcal{A} \implies \mathcal{B}$
w	w	w	w	f	w
f	w	f	w	w	w
w	f	f	w		f
f	f	f	f		w

Tabelle 1.1: Wahrheitstabelle

Die *Umkehrung* von $\mathcal{A} \implies \mathcal{B}$ ist die Aussage $\mathcal{B} \implies \mathcal{A}$. Gelten beide, ist das die *Äquivalenz* $\mathcal{A} \iff \mathcal{B}$. Dazu sagen wir auch „$\mathcal{A}$ und $\mathcal{B}$ sind äquivalent“, „$\mathcal{A}$ genau dann, wenn $\mathcal{B}$“ oder „$\mathcal{A}$ ist notwendig und hinreichend für $\mathcal{B}$“.

Beachte, dass eine Implikation keinen kausalen Zusammenhang ausdrückt; $\mathcal{A} \implies \mathcal{B}$ ist wahr wenn $\mathcal{B}$ wahr ist oder wenn $\mathcal{A}$ falsch ist. Die Aussagen

- „Sind Rosen blau, dann ist die Erde ein Würfel.“
- „Sind Rosen rot, dann ist die Erde rund.“

sind also beide trivialerweise wahr; die erste, weil es keine blauen Rosen gibt[2], und die zweite, weil die Erde nun einmal rund ist (nicht weil Rosen rot sind).

Beispiel.

(i) Für reelle Zahlen x, y gilt: $0 \leq x < y \Leftrightarrow x^2 < y^2$ (Anordnung reeller Zahlen).
(ii) Für ganze Zahlen p, q gilt: $\frac{p}{q}$, $p, q \in \mathbb{Z}$, $q \neq 0 \Rightarrow r^2 \neq 2$ (Rationale Zahlen).
(iii) $x \in A \cap B \Rightarrow x \in A$ (Mengen).

Wir führen ein paar grundlegende Umformungen an.

Satz 1.1.1.
(i) $\neg(\mathcal{A}$ *oder* $\mathcal{B}) \Leftrightarrow \neg\mathcal{A}$ *und* $\neg\mathcal{B}$ *(de Morgan'sche Regel).*
(ii) $\neg(\mathcal{A}$ *und* $\mathcal{B}) \Leftrightarrow \neg\mathcal{A}$ *oder* $\neg\mathcal{B}$.
(iii) $(\mathcal{A} \Rightarrow \mathcal{B}) \Leftrightarrow (\neg\mathcal{B} \Rightarrow \neg\mathcal{A})$ *(Kontraposition).*
(iv) $(\mathcal{A} \Rightarrow \mathcal{B}) \Leftrightarrow \neg(\mathcal{A} \wedge \neg\mathcal{B})$ *(Widerspruchsregel).*

Beweis. Der Beweis lässt sich durch Aufstellung einer Wahrheitstabelle wie in Tabelle 1.1 führen. □

Vorsicht mit dem Wort „nur“: liest man „$\mathcal{A}$ gilt nur, wenn $\mathcal{B}$ gilt“ als $\mathcal{A} \Rightarrow \mathcal{B}$ oder $\mathcal{B} \Rightarrow \mathcal{A}$? Wenn $\mathcal{B}$ nicht gilt kann $\mathcal{A}$ nicht gelten, nach Satz 1.1.1 (iii) bedeutet das also $\mathcal{A} \Rightarrow \mathcal{B}$.

1.2 Mengen

> Unter einer *Menge* verstehen wir jede Zusammenfassung M von bestimmten wohlunterschiedenen Objekten m unserer Anschauung oder unseres Denkens (welche die *Elemente* von M genannt werden) zu einem Ganzen.
>
> G. Cantor, 1895

[2] Wir lassen hier eingefärbte Blüten oder gentechnisch modifizierte Pflanzen außer Betracht.

DEDEKIND[3] beschrieb Mengen als „geschlossene Säcke, die ganz bestimmte Dinge enthalten, von denen man nichts wisse, außer, dass sie vorhanden und bestimmt seien.“ CANTOR[4] daraufhin, mit großer Geste: „Eine Menge stelle ich mir vor wie einen Abgrund.“

Ist m Element einer Menge M, so schreiben wir $m \in M$ („m Element M“), andernfalls $m \notin M$. Man kann Mengen durch Aufzählung ihrer Elemente angeben, wie in den folgenden Beispielen:

Beispiel.

(i) $\{\} = \emptyset$ (leere Menge),
(ii) $\{1, 2, 3\}$,
(iii) $\mathbb{N} = \{1, 2, 3, \dots\}$ (natürliche Zahlen),
(iv) $\mathbb{N}_0 = \{0, 1, 2, 3, \dots\}$ (natürliche Zahlen mit 0),
(v) $\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}$ (ganze Zahlen).

Ausdrücke wie z.B.

$E(x)$: „x ist kleiner gleich 3“
$F(y)$: „in diesem Hörsaal sitzen y Studenten“'

nennt man *Aussageformen*. Erst wenn man für x bzw. y passende Objekte (hier: Zahlen) einsetzt, werden daraus Aussagen, deren Wahrheitswert sich bestimmen lässt. So ist im Beispiel $E(2)$ wahr, $E(4)$ jedoch falsch. Ist $E(x)$ wahr, sagt man auch "x hat die Eigenschaft E". Damit lassen sich Mengen durch eine charakterisierende Eigenschaft aller ihrer Elemente angeben. Mit

$$\{x \in X \mid E(x)\} \text{ oder auch } \{x \in X : E(x)\}$$

bezeichnet man die Gesamtheit aller x aus einer Grundmenge X, welche die Eigenschaft E haben, z.B.

$$\{1, 2, 3\} = \{n \in \mathbb{N} \mid n \leq 3\}.$$

Ist schließlich $f(x)$ ein Funktionsausdruck, dann bezeichnet $\{f(x) \mid x \in M\}$ die Menge aller Funktionswerte $f(x)$ für beliebige $x \in M$, was sich auch mit weiteren Bedingungen kombinieren lässt, z.B.

$$\{n^2 \mid n \in \mathbb{N}_0, n \leq 4\} = \{0, 1, 4, 9, 16\}.$$

Zwei Mengen sind *gleich*, wenn sie die gleichen Elemente haben (beachte die Klammernsetzung auf der rechten Seite):

$$A = B \iff (x \in A \iff x \in B).$$

Insbesondere bedeutet das, dass die Elemente einer Menge *keine Reihenfolge* besitzen; so sind z.b.

$$\{1, 2, 3\} \text{ und } \{3, 1, 2\}$$

[3] RICHARD DEDEKIND, 1831 – 1916; gab die erste exakte Einführung der natürlichen Zahlen durch Axiome.

[4] GEORG CANTOR, 1845 – 1918; Begründer der Mengelehre.

nur zwei verschiedene Schreibwesen für die gleiche Menge.

Eine Menge A ist *Teilmenge* einer Menge B, wenn jedes Element von A auch in B liegt:

$$A \subseteq B \iff (x \in A \Rightarrow x \in B).$$

Bemerkung. Die rechte Seite ist so zu verstehen: für alle x impliziert $x \in A$, dass $x \in B$ ist. Hierbei gibt es genau betrachtet zwei Probleme: erstens, die Quantifizierung mit dem *Quantor* „für alle“; dies werden wir weiter unten behandeln. Viel tiefer ist die zweite Frage, nämlich wo denn das x sein soll wenn es nicht in A ist, d.h. aus welchem Vorrat schöpfen wir alle möglichen x, ob sie jetzt in A oder B enthalten sind oder nicht? Diese Frage könnte man nur durch ein tieferes Studium der Mengenlehre beantworten, was für unsere Zwecke viel zu weit gehen würde. Als gedankenliche Stütze können wir davon ausgehen, dass A und B in der Praxis immer *Zahlenmengen* sind, also Teilmengen (zum Beispiel) der reellen Zahlen, womit die rechte Seite also die Implikation $x \in A \Rightarrow x \in B$ für alle $x \in \mathbb{R}$ bedeutet.

Wollen wir betonen, dass A eine *echte* Teilmenge von B ist, d.h. in B enthalten aber nicht gleich B, schreiben wir $A \subsetneq B$.

Der *Durchschnitt* zweier Mengen A und B ist die Menge der gemeinsamen Elemente:

$$x \in A \cap B \iff (x \in A \text{ und } x \in B).$$

Ist $A \cap B = \emptyset$, nennt man A und B *disjunkt*.

Die *Vereinigung* von A und B ist die Menge aller x, die in mindestens einer der beiden Mengen liegen:

$$x \in A \cup B \iff (x \in A \text{ oder } x \in B).$$

Die *Mengendifferenz* von A und B ist gegeben durch

$$x \in A \setminus B \iff (x \in A \text{ und } x \notin B).$$

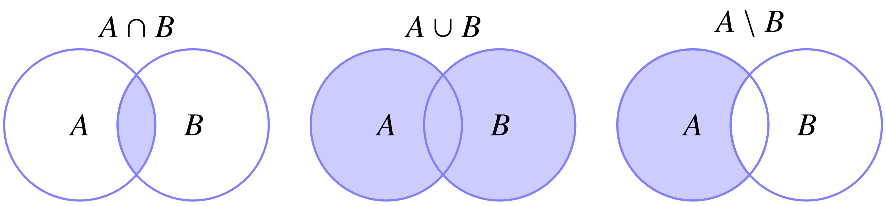


Abbildung 1.2: Mengenoperationen

Quantoren

Will man ausdrücken, dass eine Eigenschaft $E(x)$ für alle Elemente einer bestimmten Menge X gilt, verwendet man den *Allquantor* „$\forall$“:

$$\forall x \in X : E(x)$$

Will man ausdrücken, dass $E(x)$ für (mindestens) ein Element von X gilt, verwendet man den *Existenzquantor* „$\exists$“:

$$\exists x \in X : E(x)$$

Manchmal sieht man auch das Symbol „$\exists!$“; der Ausdruck

$$\exists! x \in X : E(x)$$

bedeutet, dass es *genau ein* $x \in X$ gibt, für welches $E(x)$ gilt.

Die Reihenfolge von Quantoren ist wichtig! Drückt man mit $P(m, s)$ die Tatsache aus, dass ein Schuh s einem Menschen m passt und ist M die Menge aller Menschen und S die Menge aller Schuhe, dann sieht man:

”Für jeden Menschen gibt es einen Schuh, der ihm passt.”
$\forall m \in M\ \exists s \in S : P(m, s)$

ist nicht dasselbe wie

”Es gibt einen Schuh, der jedem Menschen passt.”
$\exists s \in S\ \ \forall m \in M : P(m, s)$

Eine Aussage, welche aus Quantoren („$\forall$“, „$\exists$“), Konjunktionen („und“, „oder“) und einer Eigenschaft (oder mehreren Eigenschaften) zusammengesetzt ist, lässt sich rein mechanisch negieren, indem man die Quantoren $\forall$ und $\exists$ miteinander vertauscht, die Konjunktionen „und“ und „oder“ ebenso, und jede Eigenschaft negiert:

$$\neg(\forall m \in M\ \exists s \in S : P(m, s)) \iff \exists m \in M\ \forall s \in S : \neg P(m, s),$$

also ”Es gibt einen Menschen, dem kein Schuh passt.”

Mit Quantoren lassen sich auch Vereinigungen und Durchschnitte einer beliebigen Familie von Mengen anschreiben: ist I eine Indexmenge (d.h. eine Menge von sogenannten *Indizes*, sodass wir für jedes $i \in I$ eine Menge U_i gegeben haben) so setzen wir

$$x \in \bigcup_{i \in I} U_i \iff \exists i \in I : x \in U_i,$$
$$x \in \bigcap_{i \in I} U_i \iff \forall i \in I : x \in U_i.$$

Kartesische Produkte

Aus zwei Objekten x und y kann man das *geordnete Paar* (x, y) bilden. Im Gegensatz zu Mengen ist hier die Reihenfolge wichtig, d.h. es gilt $(1, 2) \neq (2, 1)$, obwohl $\{1, 2\} = \{2, 1\}$. Für zwei gegebene Mengen X und Y besteht das *kartesische Produkt* $X \times Y$ aus allen geordneten Paaren (x, y) mit $x \in X$ und $y \in Y$:

$$X \times Y = \{(x, y) \mid x \in X, y \in Y\}$$

Beispiel.

(i) $\{1,2\} \times \{a,b\} = \{(1,a),(1,b),(2,a),(2,b)\}$.
(ii) $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R} = \{(x,y) \mid x,y \in \mathbb{R}\}$.

Analog definiert man *n-Tupel* $(x_1,\dots,x_n)$ und das kartesische Produkt $X_1 \times \ldots \times X_n$ aus $n \in \mathbb{N}$ Faktoren.

Achtung: es gibt unterschiedliche Konventionen. Was wir als $\mathbb{N}_0$ und $\mathbb{N}$ bezeichnen, wird mancherorts als $\mathbb{N}$ und $\mathbb{N}^*$ bezeichnet; manchmal sieht man $\subset$ statt $\subsetneq$ – oder statt $\subseteq$. Kurios ist die Schreibweise von Tupeln $(a,b,c,...)$ in Demtröder, der dafür $\{a,b,c,\dots\}$ schreibt (weil dort keine Mengenklammern vorkommen).

1.3 Abbildungen

Eine *Abbildung* oder *Funktion* f von X nach Y, wobei X und Y Mengen sind, ist eine Vorschrift, die jedem Element x von X *genau ein* Element $y = f(x)$ von Y (das *Bild* von x unter f) zuordnet. Die Schreibweise dafür ist

$$f\colon X \to Y, \quad x \mapsto f(x),$$

wobei man in konkreten Fällen statt $f(x)$ natürlich den passenden Funktionsausdruck schreibt. $f(x)$ ist der *Wert* von f an der Stelle x, die Menge X heißt *Definitionsbereich* und Y ist der *Wertevorrat* oder die *Zielmenge* von f.

Beispiel.

(i) $f\colon \mathbb{N} \to \mathbb{N}$, $f(x) = x^2$ (oder $x \mapsto x^2$).
(ii) $\mathbb{R} \to \mathbb{R}$, $x \mapsto x+1$ (diese Funktion bekommt hier keinen Namen).
(iii) $\mathrm{Id}_X\colon X \to X$, $x \mapsto x$ (die *Identitätsabbildung* auf einer beliebigen Menge X).
(iv) $\mathbb{1}_A(x) = \begin{cases} 1 & \text{falls } x \in A, \\ 0 & \text{falls } x \notin A. \end{cases}$ für eine beliebige Teilmenge $A \subseteq \mathbb{R}$ (die *Indikatorfunktion* von A).

Beachte die Notation von Fallunterscheidungen in (iv).

Bemerkung. Nicht jede Zuordnungsvorschrift ergibt eine Funktion: Die Vorschrift $f\colon \mathbb{Q} \to \mathbb{Z}$, $f(\frac{p}{q}) := p - q$ für $p \in \mathbb{Z}$ und $q \in \mathbb{N}$ würde der rationalen Zahl $1/2 = 2/4$ die beiden unterschiedlichen Werte $f(\frac{1}{2}) = -1$ und $f(\frac{2}{4}) = -2$ zuordnen, was unserer Forderung nach einer eindeutigen Zuordnung widerspricht. Durch welche zusätzliche Bedingung an die Darstellung p/q einer rationalen Zahl würde man eine eindeutige Zuordnung, also eine Funktion, erhalten? [5]

Definition 1.3.1. Eine Abbildung $f\colon X \to Y$ heißt *injektiv*, falls verschiedene Elemente verschiedene Bilder haben,

$$\forall x, x' \in X : (x \neq x' \Rightarrow f(x) \neq f(x')),$$

oder, nach Satz 1.1.1 (iii) äquivalent dazu,

$$\forall x, x' \in X : (f(x) = f(x') \Rightarrow x = x').$$

[5] Man müsste voraussetzen, dass p/q immer so weit wie möglich gekürzt wird, d.h. dass p und q teilerfremd sind; damit hat nämlich jede rationale Zahl eine eindeutige Darstellung.

Beispiel.

(i) Die Funktion $f\colon \mathbb{R} \to \mathbb{R}$, $x \mapsto x^2$ ist nicht injektiv, da $f(-1) = f(1)$.
(ii) Die Matrikelnummer ist eine injektive Funktion auf der Menge aller Personen, die an einer österreichischen Universität zu einem Studium zugelassen sind.

Definition 1.3.2. Das *Bild* einer Abbildung $f\colon X \to Y$ ist die Menge

$$\mathrm{Bild}(f) := f(X) := \{f(x) \mid x \in X\}.$$

Ist $f\colon X \to Y$ eine Funktion und $A \subseteq Y$, so nennen wir $f^{-1}(A) := \{x \in X : f(x) \in A\}$ das *Urbild von A unter f*. Das ist also genau die Menge aller $x \in X$, deren Bild in B liegt.

Im Allgemeinen ist das Bild nicht gleich dem Wertevorrat.

Definition 1.3.3. Eine Abbildung $f\colon X \to Y$ heißt *surjektiv*, falls $\mathrm{Bild}(f) = Y$, d.h. wenn jedes $y \in Y$ durch f getroffen wird:

$$\forall y \in Y\ \exists x \in X\colon\ f(x) = y.$$

Definition 1.3.4. Eine Abbildung $f\colon X \to Y$ heißt *bijektiv*, falls sie sowohl injektiv als auch surjektiv ist.

Funktionen zwischen endlichen Mengen kann man wie in Abbildung 1.3 durch Diagramme veranschaulichen; so ist f ist bijektiv, g ist surjektiv, aber nicht injektiv, und h injektiv aber nicht surjektiv.

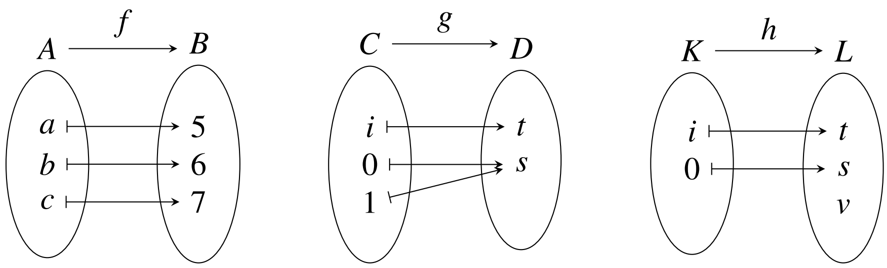


Abbildung 1.3

Wir nennen zwei Funktionen gleich und schreiben $f = g$, wenn sie den selben Definitionsbereich X und den selben Wertevorrat Y haben und $f(x) = g(x)$ für alle $x \in X$ gilt. Etwas lockerer sagt man auch „$f = g$ auf A“ falls A in den Definitionsbereichen von f und g enthalten ist und $f(x) = g(x)$ für alle $x \in A$ gilt; tatsächlich kann man den Definitionsbereich einer Funktion auch immer verkleinern:

Definition 1.3.5. Ist $f\colon X \to Y$ eine Abbildung und $A \subseteq X$, dann nennt man die Funktion $f|_A\colon A \to Y$, $x \mapsto f(x)$ die *Einschränkung von f auf A*.

Komposition

Definition 1.3.6. Seien $f\colon X \to Y$ und $g\colon Y \to Z$ Abbildungen. Die *Komposition* oder *Zusammensetzung* $g \circ f$ (sprich „g nach f“ oder „g Ring f“) ist die Abbildung

$$g \circ f\colon X \to Z, \quad (g \circ f)(x) = g(f(x)).$$

Bei der Zusammensetzung ist die Reihenfolge der Funktionen wichtig, wie das folgende Beispiel zeigt.

Beispiel. Für die Funktionen

$$f\colon \mathbb{R} \to \mathbb{R},\ f(x) := x^2, \qquad g\colon \mathbb{R} \to \mathbb{R},\ g(x) := 2x$$

gilt

$$(f \circ g)(x) = f(g(x)) = f(2x) = 4x^2$$
$$(g \circ f)(x) = g(f(x)) = g(x^2) = 2x^2,$$

und diese Funktionen sind offenbar nicht gleich, denn es gilt (außer bei $x = 0$) $4x^2 \neq 2x^2$.

Man kann auch mehr als zwei Funktionen zusammensetzen, und zwar schrittweise:

Lemma 1.3.7. Seien $f\colon X \to Y$, $g\colon Y \to Z$ und $h\colon Z \to W$ Abbildungen. Dann gilt

$$h \circ (g \circ f) = (h \circ g) \circ f.$$

Beweis. $(h \circ (g \circ f))(x) = h((g \circ f)(x)) = h(g(f(x)) = (h \circ g)(f(x)) = ((h \circ g) \circ f)(x)$. □

Deshalb ist es zulässig, $h \circ g \circ f$ zu schreiben.

Bemerkung. Das Ende eines Beweises wird üblicherweise mit „□“ oder „q.e.d.“ gekennzeichnet.

Definition 1.3.8. Eine Abbildung $f\colon X \to Y$ heißt *invertierbar*, wenn es eine Abbildung $g\colon Y \to X$ gibt mit

$$g \circ f = \mathrm{Id}_X \text{ und } f \circ g = \mathrm{Id}_Y.$$

In diesem Fall wird g mit f^{-1} bezeichnet und *Inverse* oder *Umkehrfunktion* von f genannt.

Bemerkung. Wir kennen das Symbol f^{-1} schon vom Urbild (Definition 1.3.2). Diese Notation ist konsistent mit der gerade eingeführten, denn ist eine Funktion $f\colon X \to Y$ invertierbar, so ist für eine Teilmenge $A \subseteq Y$ das Urbild $f^{-1}(A) = \{x \in X : f(x) \in A\}$ gleich dem Bild von A unter der Abbildung f^{-1}.

Beispiel. Die Abbildung

$$f\colon \mathbb{N} \to \mathbb{N},\ f(n) = \begin{cases} n-1 & \text{für } n \text{ gerade} \\ n+1 & \text{für } n \text{ ungerade} \end{cases}$$

ist invertierbar; tatsächlich ist sogar $f^{-1} = f$, wie man einfach überprüft.

Zwei Fragen drängen sich auf: wie sehen wir einer Funktion an, ob sie invertierbar ist? Und können wir wirklich von *der* Inversen sprechen, d.h. ist diese eindeutig bestimmt?

Satz 1.3.9.
(i) Ist eine Abbildung invertierbar, dann ist ihre Inverse eindeutig bestimmt.
(ii) Eine Abbildung ist genau dann invertierbar, wenn sie bijektiv ist.

Beweis. (i) Angenommen, die Funktion $f\colon X \to Y$ besitze zwei Inverse $g, h\colon Y \to X$. Dann gilt für jedes $y \in Y$

$$h(y) = h \circ \mathrm{Id}_Y(y) = h \circ (f \circ g)(y) = (h \circ f) \circ g(y) = \mathrm{Id}_X \circ g(y) = g(y),$$

also $h = g$, d.h. die zwei Inversen stimmen überein.

(ii) Angenommen, die Funktion $f\colon X \to Y$ besitze eine Inverse $f^{-1}\colon Y \to X$, möchten wir zeigen, dass f injektiv und surjektiv ist. Seien dazu $x_1, x_2 \in X$ mit $f(x_1) = f(x_2)$. Dann folgt $x_1 = f^{-1}(f(x_1)) = f^{-1}(f(x_2)) = x_2$, also ist f injektiv. Für ein gegebenes $y \in Y$ ist $f(f^{-1}(y)) = y$, für $x = f^{-1}(y)$ gilt dann $f(x) = y$ und f ist damit surjektiv, also insgesamt bijektiv.

Umgekehrt sei angenommen, f ist bijektiv. Für jedes $y \in Y$ gibt es aufgrund der Surjektivität von f mindestens ein Element $x_y \in X$ mit $f(x_y) = y$, wegen der Injektivität von f ist dieses sogar eindeutig bestimmt. Setzen wir $g(y) := x_y$ erhalten wir eine Funktion $g\colon Y \to X$, die $f(g(y)) = f(x_y) = y$ und $g(f(x_y)) = g(y) = x_y$ erfüllt, also ist f invertierbar mit der Inversen g. □

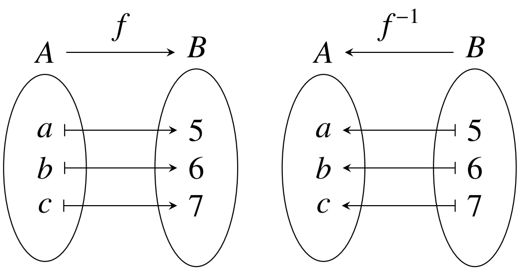


Abbildung 1.4: Funktion und Umkehrfunktion

1.4 Vollständige Induktion

Das Prinzip der *vollständigen Induktion* dient dazu, Aussagen $\mathcal{A}(n)$ (genau genommen: Aussageformen mit einer natürlichen Zahl n als Parameter) für alle natürlichen Zahlen n zu beweisen. $\mathcal{A}(n)$ gilt für alle $n \in \mathbb{N}$, wenn wir zeigen können:

(i) $\mathcal{A}(1)$ ist richtig,
(ii) für $n \geq 1$ folgt aus $\mathcal{A}(n)$ auch $\mathcal{A}(n+1)$.

Ganz analog geht man vor, wenn man eine Aussage $\mathcal{A}(n)$ nur für alle n ab einem bestimmten Startwert $n_0 \in \mathbb{N}$ zeigen will.

Im Folgenden setzen wir für $m \in \mathbb{N}_0$ und $n \in \mathbb{N}$

$$m^n := \underbrace{m \cdot m \cdot \cdots \cdot m}_{n\,\text{mal}}, \qquad m^0 := 1.$$

Beispiel 1.4.1.

(i) Für alle $n \in \mathbb{N}$ gilt $1 + 3 + 5 + \ldots + (2n - 1) = n^2$.
(ii) Für alle $n \in \mathbb{N}$ mit $n \geq 5$ gilt $2^n > n^2$.

Beweis. (i) Wir überprüfen zunächst den *Induktionsanfang* $\mathcal{A}(1)$, also $1 = 1^2$, was offensichtlich wahr ist. Für den *Induktionsschritt* $\mathcal{A}(n) \Rightarrow \mathcal{A}(n+1)$ nehmen wir dann an, dass die *Induktionsvoraussetzung* $\mathcal{A}(n)$ richtig sei. Wir müssen nun zeigen, dass $\mathcal{A}(n+1)$ gilt, also

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

gilt. Laut Voraussetzung ergibt $1+3+\ldots+(2n-1)$ genau n^2, somit kann diese Gleichung umgeformt werden zu

$$n^2 + (2n + 2 - 1) = n^2 + 2n + 1,$$

was offensichtlich richtig ist, womit wir $\mathcal{A}(n + 1)$ gezeigt haben. Nach dem Induktionsprinzip folgt daraus die Gültigkeit von $\mathcal{A}(n)$ für alle $n \in \mathbb{N}$.

(ii) Für den Induktionsanfang stellen wir fest, dass $32 = 2^5 > 5^2 = 25$ richtig ist. Gilt die Behauptung $2^n > n^2$ nun für ein $n \geq 5$, gilt sie wegen

$$2^{n+1} = 2 \cdot 2^n > 2 \cdot n^2$$

sicher auch für $n + 1$, falls $2 \cdot n^2 > (n + 1)^2$ ist, was man einfach nachprüft:

$$\begin{aligned}
& & 2 \cdot n^2 &> (n + 1)^2 \\
&\Longleftrightarrow & 2 \cdot n^2 &> n^2 + 2n + 1 \\
&\Longleftrightarrow & n^2 &> 2n + 1 \\
&\Longleftrightarrow & n^2 - 2n + 1 &> 2 \\
&\Longleftrightarrow & (n - 1)^2 &> 2 \\
&\Longleftrightarrow & n - 1 &> \sqrt{2},
\end{aligned}$$

was für $n \geq 5$ sicher richtig ist.[6] □

Folgende Ungleichung geht (u.A.) auf J. Bernoulli[7] zurück.

Satz 1.4.2 (Bernoulli'sche Ungleichung). *Für jede reelle Zahl $x \in \mathbb{R}$ mit $x \geq -1$ und jedes $n \in \mathbb{N}_0$ gilt*

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

Beweis. Wir beweisen diese Aussage mittels vollständiger Induktion.

[6] Sogar für $n \geq 3$. Warum gilt die Behauptung in (ii) trotzdem nur für $n \geq 5$?
[7] Jakob Bernoulli, 1655 – 1705

Für $n = 0$ gilt klarerweise

$$(1+x)^0 = 1 \geq 1 = 1 + 0x,$$

und daher der Induktionsanfang. Sei nun angenommen für ein $n \in \mathbb{N}_0$ gelte $(1+x)^n \geq 1 + nx$. Wegen $1 + x \geq 0$ gilt dann

$$\begin{aligned}(1+x)^{n+1} &= (1+x)^n(1+x)\\ &\overset{\text{I.V.}}{\geq} (1+nx)(1+x)\\ &= 1 + x + nx + nx^2\\ &\geq 1 + x + nx\\ &= 1 + (n+1)x,\end{aligned}$$

und somit die Aussage. □

Satz 1.4.3 (Auzinger). *Jede natürliche Zahl ist klein.*

Beweis. 1 ist klein (Induktionsanfang). Ist nun ein $n \in \mathbb{N}$ klein (Induktionsannahme), dann ist natürlich auch $n + 1$ noch klein (Induktionsschritt.) □

Summen

Für $a_1 + \ldots + a_n$ schreiben wir $\sum_{k=1}^n a_k$. Hierbei kommt es nicht auf die Wahl des *Summationsindexes* k an, d.h. $\sum_{k=1}^n a_k = \sum_{i=1}^n a_i = \sum_{j=1}^n a_j$ sind nur verschiedene Schreibweisen für dieselbe Summe. Für eine endliche Menge $I = \{n_1, \ldots, n_m\}$ setzen wir $\sum_{k\in I} a_k := a_{n_1} + \ldots + a_{n_m}$, weiters vereinbaren wir $\sum_{k\in I} = 0$ falls $I = \emptyset$ und setzen allgemein für $l, m \in \mathbb{Z}$

$$\sum_{k=l}^m a_k := \begin{cases} a_l + a_{l+1} + \ldots + a_m & \text{falls } m \geq l,\\ 0 & \text{falls } m < l.\end{cases}$$

Die *Indexverschiebung* lässt den Wert der Summe gleich, ermöglicht es aber, die Summationsgrenzen zu verschieben:

$$\sum_{k=m}^n a_k = \sum_{k=m+1}^{n+1} a_{k-1} = \ldots = \sum_{k=m+l}^{n+l} a_{k-l} = \sum_{k=m-l}^{n-l} a_{k+l}.$$

Durch Ausschreiben der Summen sieht man leicht die Rechenregeln

$$\sum_{k=m}^n (a_k + b_k) = \sum_{k=m}^n a_k + \sum_{k=m}^n b_k \text{ und } \sum_{k=m}^n (c \cdot a_k) = c \cdot \sum_{k=m}^n a_k$$

sowie

$$\sum_{k=m}^n a_k = \sum_{k=m}^{l-1} a_k + \sum_{k=l}^n a_k, \text{ wobei } m \leq l \leq n.$$

Beweis. Für Physiker: unmittelbar klar. Für Mathematiker: vollständige Induktion. □

11. 10. 2024

Beispiel 1.4.4. *Wir berechnen, für beliebige reelle Zahlen x und y,*

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

Analog zur Summe definiert man das Produkt

$$\prod_{k=l}^{m} a_k := \begin{cases} a_l \cdot a_{l+1} \cdot \ldots \cdot a_m & \text{falls } m \geq l, \\ 1 & \text{falls } m < l. \end{cases}$$

und damit die *Fakultät* oder *Faktorielle* von n,

$$n! := \prod_{k=1}^{n} k, \quad 0! := 1.$$

1.5 Binomischer Lehrsatz

Die Fakultät einer Zahl $n \in \mathbb{N}_0$ gibt die Zahl an Möglichkeiten, n verschiedene Objekte anzuordnen, wobei die Reihenfolge eine Rolle spielt. Anschaulich heißt das: für die Wahl des ersten Objektes hat man n Möglichkeiten, für die des zweiten $n-1$ und so weiter, bis man das letzte übrige Element an die letzte Stelle setzt; aufmultiplizieren der Wahlmöglichkeiten gibt genau $n \cdot (n-1) \cdots 2 \cdot 1 = n!$. So gibt es z.B. $3! = 6$ Möglichkeiten, die Zahlen $1, 2, 3$ anzuordnen:

$$(1,2,3),\ (1,3,2), (2,1,3), (2,3,1), (3,1,2), (3,2,1).$$

Binomialkoeffizient

Definition 1.5.1. Seien $n, k \in \mathbb{N}_0$ mit $n \geq k$. Wir definieren den *Binomialkoeffizienten* $\binom{n}{k}$ („n über k“) als

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

Der Binomialkoeffizient gibt die Anzahl von k-elementigen Teilmengen einer n-elementigen Menge an. Trifft man eine geordnete Auswahl von k Elementen, so hat man

$$\underbrace{n \cdot (n-1) \ldots (n-k+1)}_{k \text{ Faktoren}} = \frac{n!}{(n-k)!}$$

Möglichkeiten dafür; da $k!$ mögliche Anordnungen dieser Teilmenge auftreten und die Reihenfolge egal sein soll, muss man dann noch durch $k!$ dividieren und erhält die Anzahl von k-elementigen Teilmengen.

Beispielsweise hat die Menge $\{1, 2, 3\}$ genau $\binom{3}{2} = \frac{3!}{2!1!} = \frac{6}{2} = 3$ Teilmengen mit zwei Elementen:

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

Von besonderer Bedeutung sind die Randfälle

$$\binom{n}{0} = \binom{n}{n} = 1, \quad \binom{n}{1} = n$$

und folgendes Hilfresultat:

Lemma 1.5.2. Seien $n, k \in \mathbb{N}$ mit $k \leq n$. Dann gilt

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

Beweis. Es gilt

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

Damit lassen sich Binomialkoeffizienten für kleine n einfach per Hand berechnen:

$n = 0$							1						
$n = 1$						1		1					
$n = 2$					1		2		1				
$n = 3$				1		3		3		1			
$n = 4$			1		4		6		4		1		
$n = 5$		1		5		10		10		5		1	
$n = 6$	1		6		15		20		15		6		1

Abbildung 1.5: Pascal'sches[8] Dreieck

Satz 1.5.3 (Binomischer Lehrsatz). *Für $n \in \mathbb{N}_0$ und $x, y \in \mathbb{R}$ gilt*

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

[8] Blaise Pascal, 1623 – 1662

Beweis. Wir beweisen diesen Satz mittels vollständiger Induktion. Der Induktionsanfang ist wegen

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

richtig. Sei nun angenommen, dass die Behauptung für ein beliebiges $n \in \mathbb{N}_0$ gilt. Dann folgt mit Lemma 1.5.2

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

was die Behauptung ist. □

Beispiel. Für eine beliebige Menge M ist die *Potenzmenge* $\mathcal{P}(M)$ die Menge aller Teilmengen von M, z.B. für $M = \{1, 2, 3\}$ ist

$$\mathcal{P}(M) = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1,2\}, \{1,3\}, \{2,3\}, \{1,2,3\}\}.$$

Der Binomialkoeffizient $\binom{n}{k}$ entspricht genau der Anzahl der Teilmengen mit k Elementen; summiert man diese auf, erhält man die Zahl aller Teilmengen von M. Mit dem binomischen Lehrsatz erhalten wir

$$\text{Zahl der Teilmengen von } M = \sum_{k=0}^{n}\binom{n}{k} = \sum_{k=0}^{n}\binom{n}{k}1^k 1^{n-k} = (1+1)^n = 2^n.$$

Geometrische Reihe

Satz 1.5.4 (Geometrische Reihe). *Sei $n \in \mathbb{N}_0$ und $x \in \mathbb{R}$ mit $x \neq 1$. Dann gilt*

$$1 + x + x^2 + \ldots + x^n = \sum_{k=0}^{n} x^k = \frac{1-x^{n+1}}{1-x}.$$

Beweis. Folgt aus

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

und anschließender Division durch $1-x$. □

Alternativ könnte man das auch durch Induktion beweisen.

Beispiel. Für $a \neq \pm 1$ ist

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

17. 10. 2024

Kapitel 2

Reelle Zahlen

> Zahlen sind freie Schöpfungen des menschlichen Geistes, sie dienen als ein Mittel, um die Verschiedenheit der Dinge leichter und schärfer aufzufassen.
>
> R. Dedekind

Lernziele

Der Betrag

- Die Betragseigenschaften anwenden
- Den Betrag einer Summe/Differenz nach unten und oben abschätzen

Intervalle und beschränkte Mengen

- Intervalle ihrer Art nach kategorisieren
- Innere Punkte und Randpunkte von Intervallen unterscheiden
- Obere und untere Schranken von Mengen bestimmen
- Den Unterschied zwischen Maximum/Minimum und Supremum/Infimum erfassen
- Mit rationalen Exponenten rechnen

Komplexe Zahlen

- mit komplexen Zahlen rechnen können

2.1 Zahlbereiche

Die Menge der natürlichen Zahlen

$$\mathbb{N} = \{1, 2, 3, \dots\}$$

ist nicht abgeschlossen unter Subtraktion, d.h. die Gleichung $x + a = b$ hat nicht für alle $a, b \in \mathbb{N}$ eine Lösung $x \in \mathbb{N}$ (nämlich $x = b - a$); deshalb wird $\mathbb{N}$ zur Menge der ganzen Zahlen

$$\mathbb{Z} = \{-2, 1, 0, 1, 2, \dots\}$$

erweitert. Da diese nicht abgeschlossen unter Division ist, d.h. die Gleichung $x \cdot a = b$ nicht für alle $a, b \in \mathbb{Z}$ eine Lösung $x \in \mathbb{Z}$ (nämlich $x = b/a$) hat, wird sie zur Menge der rationalen Zahlen

$$\mathbb{Q} = \left\{ \frac{p}{q} : p, q \in \mathbb{Z},\ q > 0,\ p \text{ und } q \text{ teilerfremd} \right\}$$

erweitert.

Hierbei sagen wir, dass $r \in \mathbb{Z}$ ein *Teiler* einer Zahl $p \in \mathbb{Z}$ ist, wenn es ein $s \in \mathbb{Z}$ gibt mit $sr = p$, und $p, q \in \mathbb{Z}$ heißen teilerfremd wenn, $r = 1$ die *einzige* Zahl ist die Teiler sowohl von p als auch von q ist. So sind z.B. 4 und 8 nicht teilerfremd, da 2 ein gemeinsamer Teiler ist, aber 5 und 6 sind teilerfremd.

Auch $\mathbb{Q}$ reicht nicht für alle Zwecke:

Satz 2.1.1. *Es gibt keine rationale Zahl $r \in \mathbb{Q}$ mit $r^2 = 2$.*

Beweis. Angenommen, es gäbe $p, q \in \mathbb{Z}, q > 0$ mit $r = p/q$ und $r^2 = 2$, wobei p und q teilerfremd sind, also insbesondere nicht beide durch 2 teilbar sind. Dann folgt durch Quadrieren $2 = p^2/q^2$, also $p^2 = 2q^2$. Die Quadratzahl p^2 ist somit eine gerade Zahl, daher auch p (warum?), also ist $p = 2k$ für ein $k \in \mathbb{Z}$. Es folgt $2q^2 = p^2 = (2k)^2 = 4k^2$, also $q^2 = 2k^2$, womit auch q eine gerade Zahl ist, im Widerspruch zur Annahme, dass p und q nicht beide durch 2 teilbar sind. □

Die reellen Zahlen lösen dieses Problem: jede positive Zahl hat eine Wurzel in $\mathbb{R}$. Eine rigorose Einführung der reellen Zahlen ist für unsere Zwecke zu technisch, deshalb nehmen wir die reellen Zahlen als gegeben an und interpretieren sie im geometrischen Sinn als Punkte auf der Zahlengeraden. Wichtiger als die Frage, was reelle Zahlen genau *sind*, ist nämlich die Frage, welche *Rechenregeln* dafür gelten. Es zeigt sich, dass man alle Eigenschaften von $\mathbb{R}$ auf wenige Grundannahmen, sogannte *Axiome* zurückführen kann. Tatsächlich ist $\mathbb{R}$ der einzige *vollständig angeordnete Körper*. Nicht um das zu beweisen, sondern um das Rechnen mit rellen Zahlen zu verstehen und auf ein solides Fundament zu stellen, studieren wir im Folgenden zuerst die *Körperaxiome*, die *Anordnung* und schließlich die *Vollständigkeit*.

2.2 Körper

Wir führen nun die *Körperaxiome* an.

Für alle $a, b, c \in \mathbb{R}$ gilt:

(A1) $(a + b) + c = a + (b + c)$ (Assoziativität)
(A2) $a + b = b + a$ (Kommutativität)
(A3) $a + 0 = a$ (neutrales Element)
(A4) es gibt ein x mit $a + x = 0$; für dieses x schreiben wir $-a$ (inverses Element)

(M1) $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ (Assoziativität)

(M2) $a \cdot b = b \cdot a$ (Kommutativität)
(M3) $a \cdot 1 = a$ (neutrales Element), wobei wir $1 \neq 0$ fordern
(M4) für $a \neq 0$ gibt es ein $x \neq 0$ mit $a \cdot x = 1$, bezeichnet mit a^{-1} (inverses Element)

(D) $a \cdot (b + c) = a \cdot b + a \cdot c$ (Distributivgesetz)

Allgemein nennt man jede Menge $\mathcal{K}$, die mit einer Addition „$+ : \mathcal{K} \times \mathcal{K} \to \mathcal{K}$“ und einer Multiplikation „$\cdot : \mathcal{K} \times \mathcal{K} \to \mathcal{K}$“ versehen ist, welche diese Axiome erfüllen, einen *Körper*. $\mathbb{Q}$ und $\mathbb{R}$ sind, jeweils mit der üblichen Addition und Multiplikation, Körper. Ebenso kann man die Menge $\{0, 1\}$ mit der richtigen Definition von $+$ und $\cdot$ zu einem Körper machen. Man schreibt $a - b$ statt $a + (-b)$ und a/b statt $a \cdot b^{-1}$. Alle bekannten Rechenregeln lassen sich dann ausschließlich aus den Körperaxiomen herleiten:

Beispiel.

(i) Das Inverse Element $-a$ aus (A4) ist eindeutig bestimmt, denn aus $a + b = 0$ und $a + c = 0$ folgt

$$b \overset{(A3)}{=} b + 0 = b + (a + c) \overset{(A1)}{=} (b + a) + c \overset{(A2)}{=} (a + b) + c = 0 + c \overset{(A2)}{=} c + 0 \overset{(A3)}{=} c.$$

(ii) $(-1) \cdot a = -a$ folgt mit (i) aus

$$\begin{aligned} a + (-1) \cdot a &\overset{(M3)}{=} a \cdot 1 + (-1) \cdot a \overset{(M2)}{=} 1 \cdot a + (-1) \cdot a \overset{(D)}{=} (1 + (-1)) \cdot a \overset{(A4)}{=} 0 \cdot a \\ &\overset{(A3)}{=} 0 \cdot a + 0 \overset{(A4)}{=} 0 \cdot a + 0 \cdot a - 0 \cdot a \overset{(M2)}{=} a \cdot 0 + a \cdot 0 - 0 \cdot a \\ &\overset{(D)}{=} a \cdot (0 + 0) - 0 \cdot a \overset{(A3)}{=} a \cdot 0 - 0 \cdot a \overset{(M2)}{=} 0 \cdot a - 0 \cdot a \overset{(A4)}{=} 0. \end{aligned}$$

Wir erwähnen noch die Definition von Potenzen,

$$x^0 = 1, \quad x^1 = x, \quad x^2 = x \cdot x, \quad \ldots \quad x^{n+1} = x^n \cdot x,$$

sowie die Rechenregeln

$$x^{n+m} = x^n x^m, \quad (x^n)^m = x^{nm}, \quad x^n y^n = (xy)^n$$

für beliebige Elemente x, y eines gegebenen Körpers. Bei mehreren Exponenten wird von oben nach unten aufgelöst, d.h.

$$a^{b^c} = a^{(b^c)}.$$

Komplexe Zahlen

Auch die *komplexen Zahlen* bilden einen Körper, weswegen wir sie jetzt einführen. Wegen $a^2 \geq 0$ für alle reellen a (wie wir gleich sehen werden), gibt es keine reelle Zahl c mit der Eigenschaft $c^2 = -1$. Wir können jedoch aufbauend auf den reellen Zahlen einen Körper konstruieren, der ein Element c mit der Eigenschaft $c^2 = -1$ enthält.

Wir betrachten die Menge $\mathbb{R}^2 = \{(a, b) : a, b \in \mathbb{R}\}$ und definieren auf ihr eine Addition und eine Multiplikation wie folgt:

- $(a, b) + (c, d) := (a + c, b + d)$,

- $(a, b) \cdot (c, d) := (ac - bd, ad + bc)$.

Wie man einfach überprüft, erfüllen diese Operationen auf $\mathbb{R}^2$ die Körperaxiome. Auf der Teilmenge $\mathbb{R} \times \{0\}$ agieren diese Operationen klarerweise wie die gewöhnliche Addition und Multiplikation auf den reellen Zahlen. Wir können somit $\mathbb{R}$ mit $\mathbb{R} \times \{0\} \subseteq \mathbb{R}^2$ identifizieren.

Wir finden für jedes Tupel (a, b) ein additives Inverses $-(a, b) := (-a, -b)$ und für $(a, b) \neq (0, 0)$ ist das multiplikative Inverse gegeben durch $(a, b)^{-1} := (\frac{a}{a^2+b^2}, \frac{-b}{a^2+b^2})$.

Wir definieren nun die *imaginäre Einheit* $i := (0, 1)$ und stellen fest, dass $i^2 = (-1, 0)$. Wir definieren die *komplexen Zahlen* $\mathbb{C} := \{a + bi : a, b \in \mathbb{R}\} = \{(a, b) : a, b \in \mathbb{R}\}$ mit den obigen Operationen und schreiben ab sofort eine komplexe Zahl $z = (a, b)$ als $z = a + bi$.

Damit haben die Formeln für $-z$ und z^{-1} folgendes Aussehen:

$$-(a + ib) = -a - ib \quad \text{und} \quad (a + ib)^{-1} = \frac{a - ib}{a^2 + b^2}.$$

Wir definieren die zwei Funktionen

- $\operatorname{Re}\colon \mathbb{C} \to \mathbb{R}, \quad \operatorname{Re}(a + ib) := a,$
- $\operatorname{Im}\colon \mathbb{C} \to \mathbb{R}, \quad \operatorname{Im}(a + ib) := b,$

und nennen $\operatorname{Re} z$ den *Realteil*, sowie $\operatorname{Im} z$ den *Imaginärteil* von $z \in \mathbb{C}$. Mittels der Komplexkonjugation

$$\bar{z} := a - ib \text{ für } z = a + ib$$

hat man auch

$$\operatorname{Re} z = \frac{z + \bar{z}}{2}, \quad \operatorname{Im} z = \frac{z - \bar{z}}{2i}.$$

2.3 Anordnung

Die reellen Zahlen verfügen neben den oben behandelten Rechenregeln über Größenvergleiche, d.h. man kann Aussagen darüber treffen, ob eine Zahl größer oder kleiner als eine andere ist. Eine solche *Anordnung* gibt es nicht für komplexe Zahlen!

Die Relation $a < b$ („a ist kleiner als b") auf den reellen Zahlen hat die folgenden Eigenschaften für alle $a, b, c \in \mathbb{R}$:

(O1) Es gilt genau eine der Beziehungen $a < b$, $a = b$ oder $a > b$ (Trichotomie).
(O2) Aus $a < b$ und $b < c$ folgt $a < b$ (Transitivität).
(O3) Aus $a < b$ folgt $a + c < b + c$ (Additivität).
(O4) Aus $a < b$ und $c > 0$ folgt $a \cdot c < b \cdot c$ (Multiplikativität).

Wir schreiben $a \leq b$, wenn $a < b$ oder $a = b$ gilt. Eine Zahl a heißt *positiv*, wenn $a > 0$, *negativ*, wenn $a < 0$, *nichtnegativ*, wenn $a \geq 0$ und *nichtpositiv*, wenn $a \leq 0$. (O1)–(O4) ist alles, was wir für eine sinnvolle Anordnung brauchen; alles andere folgt daraus. Wir führen ein paar wichtige Folgerungen an:

Lemma 2.3.1. Für $a, b, x, y \in \mathbb{R}$ gilt:

(i) $x < y \Leftrightarrow 0 < y - x$,
(ii) $a < b$ und $x < y \Rightarrow a + x < b + y$,
(iii) $0 \leq a < b,\ 0 \leq x < y \Rightarrow ax < by$,
(iv) $x < y \Leftrightarrow -x > -y$,
(v) $a < b, x < 0 \Rightarrow ax > bx$,
(vi) $x \neq 0 \Rightarrow x^2 > 0$,
(vii) $x > 0 \Leftrightarrow x^{-1} > 0$,
(viii) $0 < x < y \Rightarrow x^{-1} > y^{-1}$.

Auf den komplexen Zahlen kann es keine Anordnung geben, die (O1)–(O4) erfüllt, da dann $-1 = i^2 > 0$ gelten müsste.

Der Betrag

Definition 2.3.2. Für eine reelle Zahl $x \in \mathbb{R}$ ist der *Betrag* von x definiert als

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

Es gelten die drei *Betragseigenschaften*:

Lemma 2.3.3. Für $x, y \in \mathbb{R}$ gilt

(B1) $|x| \geq 0$ für alle $x \in \mathbb{R}$, und $|x| = 0 \Leftrightarrow x = 0$ (Definitheit),
(B2) $|x \cdot y| = |x| \cdot |y|$ (Multiplikativität),
(B3) $|x + y| \leq |x| + |y|$ (Dreiecksungleichung).

Beweis. (B1): Ist $x < 0$, so ist $|x| = -x > 0$. Ist $x \geq 0$, so ist $|x| = x > 0$. Ist $|x| = 0$, so muss entweder $x = 0$ oder $-x = 0$ gelten, was in beiden Fällen $x = 0$ bedeutet.

(B2): Die Aussage ist klar für $x, y \geq 0$. Allgemein haben wir $x = \pm x_0$, $y = \pm y_0$ mit $x_0, y_0 \geq 0$ und somit $|xy| = |\pm x_0 y_0| = |x_0 y_0| = |x_0||y_0| = |x||y|$.

(B3): Aus $x \leq |x|$ und $y \leq |y|$ folgt $x + y \leq |x| + |y|$, ebenso folgt aus $-x \leq |x|$ und $-y \leq |y|$ die Ungleichung $-(x + y) \leq |x| + |y|$. Zusammen ergibt das die Behauptung. □

Es gilt

$$|x - a| < b \Longleftrightarrow -b < x - a < b \Longleftrightarrow a - b < x < a + b.$$

Für den Betrag gilt auch die folgende Ungleichung.

Lemma 2.3.4 (Umgekehrte Dreiecksungleichung)**.** Für zwei reelle Zahlen $a, b \in \mathbb{R}$ gilt

$$\big||a| - |b|\big| \leq |a - b|.$$

Beweis. Aus (B3) folgt $|a| = |b + (a - b)| \leq |b| + |a - b|$, und damit

$$|a| - |b| \leq |a - b|. \tag{2.1}$$

Analog gilt $|b| = |a + (b - a)| \leq |a| + |b - a| = |a| + |a - b|$, und daher

$$|b| - |a| \leq |b - a|. \tag{2.2}$$

Es ist $||a| - |b||$ entweder $|a| - |b|$ oder $|b| - |a|$, und daher folgt die Behauptung aus (1) oder (2). □

Lemma 2.3.5. Seien $a, b \in \mathbb{R}$. Dann gelten die folgenden Beziehungen:

(i) $|ab| \leq \frac{1}{2}(a^2 + b^2)$,
(ii) $(a + b)^2 \leq 2(a^2 + b^2)$,
(iii) Für $a < b$ ist $a < \frac{a+b}{2} < b$,
(iv) Für $a, b > 0$ ist $ab \leq \frac{(a+b)^2}{4}$, mit Gleichheit nur für $a = b$.

Der Beweis ist eine einfache Übungsaufgabe.

Wir können den Betrag auch für komplexe Zahlen definieren: für $z = a + bi$ setzen wir

$$|z| := \sqrt{a^2 + b^2}. \tag{2.3}$$

Dafür gelten die Betragseigenschaften (B1)–(B3).

2.4 Intervalle und beschränkte Mengen

Ein *Intervall* ist eine Teilmenge von $\mathbb{R}$, welche mit zwei beliebigen Punkten auch jeden Punkt dazwischen enthält.

Definition 2.4.1. Für zwei reelle Zahlen $a, b \in \mathbb{R}$ mit $a \leq b$ definieren wir die *offenen Intervalle*

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

die *abgeschlossenen Intervalle*

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

und die halboffenen Intervalle

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

Manchmal spielt es eine Rolle, ob man sich am Rand eines Intervalles befindet oder im Inneren. Daher gleich die richtigen Begriffe:

Definition 2.4.2. Ist $I \subseteq \mathbb{R}$ ein Intervall, dann heißt $x \in \mathbb{R}$

(i) *innerer Punkt* von I, falls es ein $\delta \in \mathbb{R}$, $\delta > 0$ gibt, sodass die Menge $(x - \delta, x + \delta)$ ganz in I liegt;
(ii) *Randpunkt* von I, falls für jedes $\delta \in \mathbb{R}$, $\delta > 0$ die Menge $(x - \delta, x + \delta)$ ein Element von I und eines von $\mathbb{R} \setminus I$ enthält.

Von den Intervallen in Definition 2.4.1 sind genau alle Punkte a, b Randpunkte, egal ob die Intervalle offen, halboffen oder abgeschlossen sind.

Definition 2.4.3. Eine Menge $M \subseteq \mathbb{R}$ heißt

(i) *nach oben beschränkt*, falls es eine Zahl $K \in \mathbb{R}$ gibt mit $x \leq K$ für alle $x \in M$;
(ii) *nach unten beschränkt*, falls es eine Zahl $L \in \mathbb{R}$ gibt mit $L \leq x$ für alle $x \in M$;
(iii) *beschränkt*, falls M nach oben und unten beschränkt ist, also wenn es eine Zahl $R > 0$ gibt mit $|x| \leq R$ für alle $x \in M$;
(iv) *(nach oben/unten) unbeschränkt*, falls sie nicht (nach oben/unten) beschränkt ist.

K bzw. L heißen *obere* bzw. *untere* Schranke.

Achtung: eine obere (untere) Schranke einer Menge muss nicht Element dieser Menge sein! Hat man eine obere (untere) Schranke gefunden, so ist jede größere (kleinere) Zahl auch obere (untere) Schranke.

Beispiel.

(i) Die Intervalle $[a, b]$, (a, b), $[a, b)$ und $(a, b]$ sind beschränkt.
(ii) $\mathbb{N}$ ist nach unten durch 1 beschränkt, jedoch nach oben unbeschränkt. $\mathbb{R}$ ist nach oben und unten unbeschränkt.

Maximum und Minimum

Definition 2.4.4. Das *Maximum* bzw. *Minimum* zweier Zahlen $a, b \in \mathbb{R}$ ist definiert als

$$\max\{a, b\} := \begin{cases} a & \text{falls } a \geq b \\ b & \text{falls } a < b, \end{cases} \qquad \min\{a, b\} := \begin{cases} b & \text{falls } a \geq b \\ a & \text{falls } a < b. \end{cases}$$

Zu endlich vielen Zahlen $x_1, \ldots, x_n \in \mathbb{R}$ gibt es eine größte und eine kleinste (Beweis: Induktion nach n), bezeichnet mit

$$\max\{x_1, \ldots, x_n\}, \qquad \min\{x_1, \ldots, x_n\}.$$

Ist M eine nichtleere Teilemenge von $\mathbb{R}$, so schreiben wir $m = \max M$, falls m obere Schranke von M ist und $m \in M$ gilt; m heißt dann *Maximum* von M. Analog schreiben wir $n = \min M$, falls n untere Schranke von M ist und $n \in M$ gilt; n heißt dann *Minimum* von M. Aufgrund der Antisymmetrie von „$\leq$" (d.h. $a \leq b$ und $b \leq a$ impliziert $a = b$) sind Maximum und Minimum eindeutig, falls sie existieren. Für eine beliebige Teilmenge der reellen Zahlen ist jedoch nicht garantiert, dass sie überhaupt ein Maximum bzw. Minimum besitzt! Hätte das halboffene Intervall $[0, 1)$ zum Beispiel ein Maximum m, wäre $m < (m + 1)/2 < 1$, also $(m + 1)/2$ eine obere Schranke die größer als m ist.

Vollständigkeit

Neben der Anordnung hat der Körper der reellen Zahlen eine weitere Eigenschaft, welche ihn auszeichnet und Grundvoraussetzung für die gesamte Analysis ist.

Satz 2.4.5 (Dedekind-Vollständigkeit). *Jede nach oben beschränkte, nichtleere Teilmenge $M \subseteq \mathbb{R}$ besitzt eine* kleinste obere Schranke. *Diese wird* Supremum *von M genannt und mit* $\sup M$ *bezeichnet.*

Wir werden diesen Satz nicht beweisen, sondern als *Axiom* stehen lassen. Später werden wir die Vollständigkeit noch weiter studieren.

Durch Spiegelung am Nullpunkt sieht man, dass jede nach unten beschränkte, nichtleere Teilmenge $M \subseteq \mathbb{R}$ eine größte untere Schranke besitzt; diese wird *Infimum* von M genannt und mit $\inf M$ bezeichnet.

Bemerkung. Ist eine Menge M nach oben/unten unbeschränkt, drückt man das auch durch $\sup M = \infty$ bzw. $\inf M = -\infty$ aus.

Wie beim Maximum sieht man, dass es jeweils nur *ein* Supremum bzw. Infimum einer Menge geben kann.

Lemma 2.4.6. Für $\emptyset \neq A \subseteq B \subseteq \mathbb{R}$ und B beschränkt gilt $\sup A \leq \sup B$ und $\inf A \geq \inf B$.

Beweis. Aus der Definition des Supremums wissen wir

$$\forall x \in A : x \leq \sup A \quad \text{und} \quad \forall S \in \mathbb{R} : ((\forall x \in A : x \leq S) \Rightarrow \sup A \leq S),$$

d.h. $\sup A$ ist obere Schranke von A, und es gibt keine kleinere obere Schranke. Zu zeigen ist also $x \leq \sup B$ für alle $x \in A$, was aber aus $A \subseteq B$ sofort folgt. Die Aussage für das Infimum zeigt man analog. □

Beispiel. Die Menge $\{\frac{1}{n} : n \in \mathbb{N}\}$ ist beschränkt und nichtleer. Sie besitzt das Maximum 1, jedoch kein Minimum; das Infimum ist 0, denn 0 ist untere Schranke, und für jede Zahl $\alpha > 0$ gibt es ein n mit $1/n < \alpha$ (warum eigentlich?), also ist α keine untere Schranke.

Aus der Vollständigkeit von $\mathbb{R}$ folgen einige grundlegende Tatsachen. Zuerst zeigen wir die *archimedische Eigenschaft* von $\mathbb{R}$:

Satz 2.4.7 (Satz des Archimedes).
(i) Für positive reelle Zahlen a, b gibt es $n \in \mathbb{N}$ mit $n \cdot a > b$.
(ii) Zu jeder positiven reellen Zahl r gibt es $n \in \mathbb{N}$ mit $n > r$.
(iii) Zu jedem $\varepsilon > 0$ gibt es $n \in \mathbb{N}$ mit $1/n < \varepsilon$.

Beweis. (ii): Sei $r > 0$. Angenommen es gäbe kein $n \in \mathbb{N}$ mit $n > r$ wäre $\mathbb{N}$ nach oben beschränkt mit oberer Schranke r, hätte also ein Supremum $s = \sup \mathbb{N}$. Es wäre dann auch

$n+1 \leq s$ für alle $n \in \mathbb{N}$, also $n \leq s-1$ für alle $n \in N$, damit wäre auch $s-1$ obere Schranke von $\mathbb{N}$, was ein Widerspruch zur Annahme ist dass s die kleinste obere Schranke ist.

(i) folgt aus (ii) mit $r = b/a$ und (iii) mit $r = 1/\varepsilon$. □

So grundlegend die nächsten Sätze auch sind, die Beweise im Detail durchzugehen geht für uns zu weit. Der Vollständigkeit halber (und für Interessierte) sind sie jedoch trotzdem im Skriptum angeführt.

Satz 2.4.8. *Jede nach unten beschränkte, nichtleere Teilmenge von $\mathbb{N}$ besitzt ein kleinstes Element.*

Beweis. Sei $S \subseteq \mathbb{N}$ nach unten beschränkt. Wir zeigen, dass die Annahme, S besitze kein kleinstes Element, impliziert dass S leer ist. Sei

$$A := \{n \in \mathbb{N} : 1, 2, \ldots, n \notin S\}.$$

Wäre $1 \notin A$ dann $1 \in S$, also 1 kleinestes Element, Widerspruch; also $1 \in A$. Angenommen nun $n \in A$; wäre $n+1 \notin A$, dann $n+1 \in S$, also $n+1$ kleinstes Element, Widerspruch; also $n+1 \in A$. Induktiv ist $A = \mathbb{N}$ und damit $S = \emptyset$. □

Satz 2.4.9. *Zwischen zwei reellen Zahlen $x < y$ gibt es immer eine rationale Zahl.*

Beweis. Angenommen $x > 0$; da $y - x > 0$ existiert wegen der archmiedischen Eigenschaft ein $n \in \mathbb{N}$ mit $n(y-x) > 1$, also $nx+1 < ny$. Sei $S := \{k \in \mathbb{N} : nx < k\}$. Dann ist $S \neq \emptyset$ (Archimedische Eigenschaft) und nach unten beschränkt (durch 1), besitzt also ein minimales Element m. Es gilt $m \in S$ und $s - 1 \notin S$ (ist $s - 1 \in \mathbb{N}$ kann es, da s das Minimum ist, nicht in S sein; ist $s - 1 = 0$ dann ist es sowieso nicht in S).

Damit ist $nx < s$, $nx \geq s - 1$ und insgesamt

$$nx < s \leq nx + 1 < ny \Rightarrow x < \frac{s}{n} < y.$$

Ist $x < 0$ wählt man $k \in \mathbb{N}$ mit $x + k > 0$, findet $q \in \mathbb{Q}$ mit $x + k < q < y + k$ und somit ist $x < q - k < y$ mit $q - k \in \mathbb{Q}$. □

Schließlich zeigen wir die Existenz der n-ten Wurzel jeder positiven rellen Zahl.

Satz 2.4.10. *Sei $x \in \mathbb{R}$, $x > 0$. Für jede natürliche Zahl $n \in \mathbb{N}$ gibt es eine eindeutige positive n-te Wurzel von x, also ein $y \in \mathbb{R}$, $y > 0$ mit $y^n = x$. Wir setzen $\sqrt[n]{x} = y$.*

Beweis. Wir nehmen $n \geq 2$ an, da für $n = 1$ nichts zu tun ist.

Für die Existenz betrachten wir die Menge

$$M = \{t > 0 : t^n \leq x\}.$$

Zuerst zeigen wir, dass M nicht leer ist. Dazu setzen wir $v = x/(x+1)$; aus $0 < v < 1$ folgt $0 < v^{n-1} < 1$ und damit $0 < v^n < v < x$, also $v \in M$.

Um zu zeigen, dass M nach oben beschränkt ist, setzen wir $z = x+1$. Aus $z > 1$ folgt dann $z^{n-1} > 1$ und somit $z^n > z > x$, also ist z eine obere Schranke von M: aus $t \in M$ folgt $t^n \leq x < z^n$, also muss $t \leq z$ gelten (gälte nämlich $t > z$ dann auch $t^n > z^n$).

Es gibt also ein (positives) Supremum von S. Wir setzen $y = \sup S$ und behaupten, dass $y^n = x$ ist. Dies zeigen wir, indem wir jede der Annahmen $y^n < x$ und $y^n > x$ auf einen Widerspruch führen. Dazu verwenden wir die Abschätzung

$$0 < a < b \Rightarrow b^n - a^n < (b-a)nb^{n-1} \quad \forall n \in \mathbb{N} \tag{2.4}$$

welche mittels $a^k < b^k$ aus der Formel

$$b^n - a^n = (b-a)\underbrace{(b^{n-1} + b^{n-2}a + \ldots + ba^{n-2} + a^{n-1})}_{<nb^{n-1}} \tag{2.5}$$

folgt.

Angenommen $y^n < x$. Dann wählen wir $h < 1$ mit

$$0 < h < \frac{x - y^n}{n(y+1)^{n-1}}.$$

Setzt man $a = y$ und $b = (y+h)$ in (2.4) ergibt sich

$$(y+h)^n - y^n < hn(y+h)^{n-1} < \frac{(x-y^n)n(y+h)^{n-1}}{n(y+1)^{n-1}} < \frac{(x-y^n)n(y+1)^{n-1}}{n(y+1)^{n-1}} = x - y^n,$$

also $(y+h)^n < x$ bzw. $y + h \in M$. Wegen $y < y + h$ ist das ein Widerspruch dazu, dass y eine obere Schranke für M ist.

Sei nun $y^n > x$. Dann setzen wir

$$k = \frac{y^n - x}{ny^{n-1}}$$

und bemerken

$$0 < k < \frac{y^n}{ny^{n-1}} = \frac{y}{n} < y.$$

Setzt man in (2.4) $a = y - k > 0$ und $b = y$, erhalten wir

$$y^n - (y-k)^n < kny^{n-1} = \frac{(y^n - x)ny^{n-1}}{ny^{n-1}} = y^n - x,$$

also $x < (y-k)^n$, d.h. $y - k$ ist obere Schranke für M. Wegen $y - k < y$ ist das ein Widerspruch dazu, dass y die *kleinste* obere Schranke ist.

Aus $y^n \not> x$ und $y^n \not< x$ folgt $y^n = x$.

Die Eindeutigkeit sieht man aus (2.5), indem man für a und b zwei n-te Wurzeln einsetzt. □

Über die Wurzel können wir *rationale Exponenten* einführen: für $a > 0$ und $r = m/n$ mit $m \in \mathbb{N}_0$ und $n \in \mathbb{N}$ setzen wir

$$a^r := \sqrt[n]{a^m}, \quad a^{-r} := \frac{1}{a^r}.$$

Damit gelten wieder die Rechenregeln

18. 10. 2024

$$a^{r+s} = a^r a^s, \quad a^{rs} = (a^r)^s, \quad (ab)^r = a^r b^r.$$

Kapitel 3

Folgen

Lernziele

Konvergenz

- das Konzept der Konvergenz anschaulich verstehen und formal definieren
- Arten von Divergenz benennen
- Zusammenhang von Beschränktheit und Konvergenz verstehen
- Elementare Rechenoperationen auf Folgen anwenden

Monotonie

- Arten von Monotonie benennen können
- Zusammenhang zwischen Monotonie und Konvergenz verwenden
- Maximal- und Minimalfolgen ermitteln und anwenden

Cauchyfolgen

- Folgen vs. Cauchyfolgen: Unterschied und Zusammenhang

Teilfolgen

- Divergenz und Konvergenz mit Teilfolgen in Zusammenhang bringen.
- Satz von Bolzano-Weierstraß

3.1 Konvergenz

Definition 3.1.1. Unter einer *reellen Folge* (oder Folge in $\mathbb{R}$) verstehen wir eine Funktion $x\colon \mathbb{N} \to \mathbb{R}$. Statt $x(n)$ schreiben wir x_n für die *Folgenglieder*. Die Folge wird mit $(x_n)_{n\in\mathbb{N}}$, $(x_n)_n$ oder kurz (x_n) bezeichnet. Ist $M \subseteq \mathbb{R}$ und $x_n \in M$ für alle n, so nennt man (x_n) auch eine Folge in M.

Beispiel.

(i) $(c, c, c, \ldots)$ mit $c \in \mathbb{R}$ (*konstante* Folge).
(ii) $(\frac{1}{n})_{n\in\mathbb{N}} = (\frac{1}{1}, \frac{1}{2}, \frac{1}{3}, \ldots)$.
(iii) Die Folge $x_n = (-1)^n$, also $(-1, 1, -1, 1, -1, \ldots)$.

Folgen können auch bei einem anderen Index als 1 beginnen:

(iv) $(x_n)_{n\geq 3} := (\frac{n}{n-2})_{n\geq 3} = (\frac{3}{1}, \frac{4}{2}, \frac{5}{3}, \ldots)$.

Manche Folgen sind *rekursiv* definiert,

(v) $x_0 := 0, x_1 := 1, x_n := x_{n-1} + x_{n-2}$ für $n \geq 2$, also $(x_n) = (0, 1, 1, 2, 3, 5, \ldots)$ (die sogenannte *Fibonacci-Folge*),

oder über eine geometrische Konstruktion:

(vi) r_n sei der Umfang eines dem Einheitskreis eingeschriebenen regulären Polygones mit n Ecken ($n \geq 3$).

An Folgen interessiert uns hauptsächlich das Verhalten „im Unendlichen". Dazu drei Fragen zur Motivation:

(i) Was ist die *Länge* einer Kurve?
(ii) Was ist die *momentane Änderung* einer Funktion $f\colon \mathbb{R} \to \mathbb{R}$ an einer Stelle $x \in \mathbb{R}$?
(iii) Was ist die *Fläche* unter dem Graphen einer Funktion?

Definition 3.1.2. Eine Folge $(x_n)_{n\in\mathbb{N}}$ *konvergiert* gegen eine Zahl $a \in \mathbb{R}$, falls

$$\forall \varepsilon > 0\ \exists n_0 \in \mathbb{N}\ \forall n \geq n_0 : |x_n - a| < \epsilon. \tag{3.1}$$

In diesem Fall schreibt man $x_n \to a$ und a heißt *Grenzwert* der Folge (für $n \to \infty$). Konvergiert eine Folge gegen 0, so nennen wir sie eine *Nullfolge*. Ist $M \subseteq \mathbb{R}$, (x_n) eine Folge in M und auch der Grenzwert a in M, so sagt man (x_n) konvergiert *in M*.

Ändert man von einer gegebenen Folge endlich viele Folgenglieder ab, bleibt das Konvergenzverhalten dasselbe, denn in (3.1) kann man n_0 immer so groß wählen, dass die abgeänderten Glieder keine Rolle spielen.

Satz 3.1.3 (Eindeutigkeit des Grenzwertes). *Besitzt eine konvergente Folge $(x_n)_{n\in\mathbb{N}}$ zwei Grenzwerte $a, b \in \mathbb{R}$, dann gilt $a = b$.*

Beweis. Wenn a und b Grenzwerte der Folge (x_n) sind, gibt es für jedes $\epsilon > 0$ ein $n_a, n_b \in \mathbb{N}$ sodass $\forall n \geq \max\{n_a, n_b\}$ gilt $|x_n - a| < \frac{\epsilon}{2}$ und $|x_n - b| < \frac{\epsilon}{2}$. Mit der Dreiecksungleichung folgt

$$|a - b| = |a - x_n + x_n - b| \leq |x_n - a| + |x_n - b| < \epsilon.$$

Da ε beliebig ist, muss wegen der Archmidischen Eigenschaft von $\mathbb{R}$ $|a - b| = 0$ sein, also $a = b$. Im Detail: insbesondere gilt $|a - b| \leq 1/n$ für alle $n \in \mathbb{N}$; wäre $|a - b| > 0$ gäbe es aber ein $n_0 \in \mathbb{N}$ mit $|a - b| > 1/n_0$. □

Es genügt also, von *dem* Grenzwert einer konvergenten Folge zu sprechen; konvergiert $(x_n)_n$ gegen a, so schreiben wir

$$\lim_{n\to\infty} x_n = a.$$

Beispiel.

- Die konstante Folge $x_n = c \in \mathbb{R}$ konvergiert gegen c, denn für beliebiges $\varepsilon > 0$ gilt sogar für *jedes* $n_0 \in \mathbb{N}$ die Ungleichung $|x_n - c| < \varepsilon$ für $n \geq n_0$.
- Wir untersuchen das Konvergenzverhalten der Folge $(x_n) = (\frac{1}{n})$. Sei dazu $\epsilon > 0$ gegeben. Wir finden mit Satz 2.4.7 (iii) ein $n_0 \in \mathbb{N}$ mit $\frac{1}{n_0} < \epsilon$. Für alle $n \geq n_0$ gilt $\frac{1}{n} < \frac{1}{n_0}$ und somit $|x_n - 0| = 1/n < \epsilon$, also $\lim_{n\to\infty} = 0$.
- Die Folge $(-1)^n$ nimmt abwechselnd die Werte -1 und $+1$ an und kann somit (zumindest anschaulich) nicht konvergent sein – wie beweist man das? Wir führen die Annahme der Konvergenz auf einen Widerspruch: angenommen (3.1) wäre wahr, dann gäbe es z.B. für $\varepsilon = 1/4$ ein n_0 mit $|(-1)^{n_0} - a| < \varepsilon$ und $\left|(-1)^{n_0+1} - a\right| < \varepsilon$, also
$$|1 - a| < \frac{1}{4} \quad \text{und} \quad |-1 - a| < \frac{1}{4}.$$
Dann wäre aber auch
$$2 = |1 - (-1)| = |1 - a + a - (-1)| \leq |1 - a| + |-1 - a| \leq \frac{1}{4} + \frac{1}{4} = \frac{1}{2},$$
was einen Widerspruch ergibt, also kann $(-1)^n$ nicht konvergieren.
- Für die Folge $n/(n - 2)$ könnte man erraten, dass sie gegen 1 konvergiert. Für den strengen Beweis wählen wir $\varepsilon > 0$ beliebig. Dann formen wir die gewünschte Behauptung um, um das geeignete n_0 zu erhalten:
$$\begin{aligned}\left|\frac{n}{n-2} - 1\right| = \left|\frac{n-n+2}{n-2}\right| = \left|\frac{2}{n-2}\right| &< \varepsilon \\ \Leftrightarrow 2 &< \varepsilon(n-2) \\ \Leftrightarrow n &> \frac{2}{\varepsilon} + 2,\end{aligned}$$
wobei wir $n \geq 3$ voraussetzen. Für $n_0 \in \mathbb{N}$ mit $n_0 > 2/\varepsilon + 2$ gilt also $|n/(n - 2) - 1| < \varepsilon$.
- Die Fibonacci-Folge konvergiert nicht, denn z.B. für $n \geq 3$ sieht man induktiv, dass $x_n \geq x_{n-1} + 1$. Satz 3.1.6 wird dann die Behauptung geben.

Definition 3.1.4. Ist eine Folge (x_n) nicht konvergent, so heißt sie *divergent*. Weiters heißt (x_n) *bestimmt divergent gegen* ∞, falls

$$\forall T \in \mathbb{R}\ \exists n_0 \in \mathbb{N}\ \forall n \geq n_0 : x_n > T, \tag{3.2}$$

und *bestimmt divergent gegen* $-\infty$, falls

$$\forall S \in \mathbb{R}\ \exists n_0 \in \mathbb{N}\ \forall n \geq n_0 : x_n < S. \tag{3.3}$$

Ist eine Folge divergent, aber weder gegen ∞ noch gegen $-\infty$ bestimmt divergent, so heißt sie *unbestimmt divergent*.

Manchmal sagt man für eine bestimmt gegen ∞ divergente Folge auch, dass sie „gegen ∞ konvergiert“, und analog für $-\infty$.

Bemerkung. In (3.2) und (3.3) kann man statt $<$ bzw. $>$ genauso auch $\leq$ bzw. $\geq$ fordern.

Beispiel. Das Konvergenzverhalten der Folge (x^n) ist von $|x|$ abhängig. Ist $|x| = 1$, so konvergiert die Folge im Fall $x = 1$ gegen 1, da sie die konstante Folge ist, und divergiert im Fall $x = -1$ unbestimmt, wie wir oben gesehen haben. Ist $|x| > 1$ dann ist $|x| = 1 + \delta$ für ein $\delta > 0$ und somit mit dem Satz von Bernoulli $|x|^n = (1 + \delta)^n \geq 1 + \delta n$, also ist die Folge bestimmt divergent gegen ∞: $1 + \delta n > T$ gilt nämlich sobald $n > (T - 1)/\delta$ ist. Ist schließlich $|x| < 1$, so konvergiert die Folge gegen 0, denn $1/|x| = 1 + \delta$ für ein $\delta > 0$; wieder mit dem Satz von Bernoulli folgt $1/|x|^n \geq 1 + n\delta$, also $|x^n| \leq 1/(1 + n\delta)$ und damit für beliebiges $\varepsilon > 0$ auch $|x|^n < \varepsilon$ sobald $1/(1 + n\delta) < \varepsilon$, d.h. für $n > (1/\varepsilon - 1)/\delta$.

Analog zur Beschränktheit von Teilmengen von reellen Zahlen definiert man *Beschränktheit von Folgen*:

Definition 3.1.5. Eine Folge $(x_n)_{n\in\mathbb{N}}$ ist (nach oben/unten) *beschränkt*, wenn die Menge $\{x_n : n \in \mathbb{N}\}$ (nach oben/unten) beschränkt ist. Andernfalls heißt $(x_n)_{n\in\mathbb{N}}$ (nach oben/unten) *unbeschränkt*.

Beispiel.

- $(n^2)_{n\in\mathbb{N}}$ ist nach oben unbeschränkt, nach unten durch 1 beschränkt.
- $(n(-1)^n)_{n\in\mathbb{N}}$ ist nach oben und unten unbeschränkt.

24. 10. 2024

Satz 3.1.6. *Jede konvergente Folge ist beschränkt.*

Beweis. Konvergiert eine Folge $(x_n)_n$ gegen $a \in \mathbb{R}$, so gibt es n_0 mit $|x_n - a| < 1$ für $n \geq n_0$. Ist M das Maximum der Zahlen $|x_1 - a|, |x_2 - a|, \ldots \left|x_{n_0-1} - a\right|$, dann gilt

$$|x_n| = |x_n - a + a| \leq |x_n - a| + |a| \leq M + 1 + |a|. \qquad \square$$

Umgekehrt schließen wir also, dass eine unbeschränkte Folge nicht konvergieren kann.

Satz 3.1.7 (Einschließungskriterium). *Seien $a_n \to x$ und $b_n \to x$ konvergente Folgen mit demselben Limes $x \in \mathbb{R}$. Gilt für eine Folge (c_n) und einen Index n_0 die Eigenschaft*

$$a_n \leq c_n \leq b_n \qquad \forall n \geq n_0,$$

so konvergiert auch (c_n) gegen x.

Beweis. Für jedes $\epsilon > 0$ gibt es einen Index $n_1 \geq n_0$ so, dass für alle $n \geq n_1$ die Ungleichungen $|a_n - x| < \epsilon$ und $|b_n - x| < \epsilon$ erfüllt sind. Damit folgt

$$-\epsilon < a_n - x \leq c_n - x \leq b_n - x < \epsilon,$$

also auch $|c_n - x| < \epsilon$ für $n \geq n_1$. $\square$

Wie man leicht sieht, stellt nachstehendes Lemma eine einfache Folgerung aus dem Einschließungskriterium dar:

Folgerung 3.1.8. *Sei (a_n) eine Nullfolge, $(b_n)_n$ eine Folge und es existiere $b \in \mathbb{R}$ und ein $n_0 \in \mathbb{N}$ mit $|b_n - b| \leq a_n$ für alle $n \geq n_0$. Dann gilt $b_n \to b$.*

Beispiel. Wir wollen den Grenzwert von $a_n := \frac{n^n}{2^{n^2}}$. berechnen. Wir formen a_n um zu $\frac{n^n}{2^{n^2}} = \frac{n^n}{(2^n)^n} = (\frac{n}{2^n})^n$. Wegen Beispiel 1.4.1 (ii) ist $2^n \geq n^2$ für alle $n \geq 5$ und damit $0 \leq a_n \leq (\frac{n}{n^2})^n = (\frac{1}{n})^n = \frac{1}{n^n} \leq \frac{1}{n}$, a_n konvergiert also gegen 0.

Folgenoperationen

Satz 3.1.9.
(i) Ist (a_n) eine Nullfolge und (b_n) beschränkt, dann ist $(a_n \cdot b_n)$ eine Nullfolge.
(ii) Ist $a_n \to a$ konvergent und $\lambda \in \mathbb{R}$, dann ist (λa_n) konvergent mit Grenzwert λa.
(iii) Sind $(a_n)_n$ und $(b_n)_n$ Nullfolgen, dann auch $(a_n + b_n)$.
(iv) Gilt $x_n \to x$, dann auch $|x_n| \to |x|$; gilt $|x_n| \to 0$, dann auch $x_n \to 0$.

Beweis. (i): Weil (b_n) beschränkt ist, gibt es ein $M > 0$ mit $|b_n| \leq M$ für alle $n \in \mathbb{N}$. Sei $\epsilon > 0$ beliebig. Weil (a_n) eine Nullfolge ist, gibt es ein $n_0 \in \mathbb{N}$ mit $|a_n| < \frac{\epsilon}{M}$ für $n \geq n_0$ und damit auch $|a_n \cdot b_n| < \frac{\epsilon}{M} \cdot M = \epsilon$, woraus $a_n \cdot b_n \to 0$ folgt.

(ii): Der Fall $\lambda = 0$ ist trivial, also nehmen wir $\lambda \neq 0$ an. Für ein beliebiges $\epsilon > 0$ gibt es wegen der Konvergenz von (a_n) ein $n_0 \in \mathbb{N}$ mit $|a_n - a| < \frac{\epsilon}{|\lambda|}$ für $n \geq n_0$ und daher $|\lambda a_n - \lambda a| = |\lambda||a_n - a| < |\lambda|\frac{\epsilon}{|\lambda|} = \epsilon$, also konvergiert λa_n gegen λa.

(iii): Für $\varepsilon > 0$ wählen wir n_0 so, dass $|a_n| < \varepsilon/2$ und $|b_n| < \varepsilon/2$ für $n \geq n_0$; dann ist $|a_n + b_n| \leq |a_n| + |b_n| \leq \varepsilon/2 + \varepsilon/2 = \varepsilon$ für $n \geq n_0$.

(iv): Ist $\varepsilon > 0$ und n_0 so, dass $|x_n - x| < \varepsilon$ für $n \geq n_0$, dann gilt mit der umgekehrten Dreiecksumgleichung auch

$$||x_n| - |x|| \leq |x_n - x| < \varepsilon,$$

also folgt aus $x_n \to x$ die Konvergenz $|x_n| \to |x|$. Die andere Aussage ist trivial. □

Satz 3.1.10. *Seien $a_n \to a$, $b_n \to b$ zwei konvergente Folgen. Dann gilt*
(i) $a_n + b_n \to a + b$,
(ii) $a_n b_n \to ab$,
(iii) ist $b \neq 0$, so gibt es ein $n_0 \in \mathbb{N}$ sodass $\forall n \geq n_0 : b_n \neq 0$. Außerdem gilt $\frac{a_n}{b_n} \to \frac{a}{b}$.

Beweis. (i): Sei $\epsilon > 0$ beliebig. Es gibt $n_0 \in \mathbb{N}$ mit

$$\forall n \geq n_0 : |a_n - a| < \frac{\epsilon}{2} \text{ und } |b_n - b| < \frac{\epsilon}{2}$$

und damit

$$|(a_n + b_n) - (a + b)| = |(a_n - a) + (b_n - b)| \leq |a_n - a| + |b_n - b| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

also $a_n + b_n \to a + b$.

(ii): Wir schreiben

$$|a_n b_n - ab| = |a_n b_n - a_n b + a_n b - ab| \leq |a_n(b_n - b)| + |(a_n - a)b|. \tag{3.4}$$

Nun ist (a_n) beschränkt, $b_n - b$ eine Nullfolge, also $a_n(b_n - b)$ und damit auch $|a_n(b_n - b)|$ eine Nullfolge. Ebenso ist $|(a_n - a)b|$ eine Nullfolge, also auch die Summe rechts in (3.4).

(iii): Wegen Satz 3.1.9 (iv) gilt $|b_n| \to |b|$. Für $\epsilon := \frac{|b|}{2} > 0$ gilt

$$\exists n_0 \in \mathbb{N} : \forall n \geq n_0 : ||b_n| - |b|| < \epsilon,$$

und damit

$$-(|b_n| - |b|) < \epsilon, \text{ also } |b_n| > |b| - \epsilon = \frac{|b|}{2} > 0.$$

Für alle $n \geq n_0$ ist also $b_n \neq 0$ und es gilt

$$0 \leq \left| \frac{1}{b_n} - \frac{1}{b} \right| = \left| \frac{b - b_n}{b b_n} \right| \leq \frac{2|b - b_n|}{|b|^2} \to 0.$$

Mit dem Einschließungskriterium folgt $\frac{1}{b_n} - \frac{1}{b} \to 0$, d.h. $1/b_n \to 1/b$, und mit (ii) erhalten wir $\frac{a_n}{b_n} \to \frac{a}{b}$. □

In der Praxis sind „ε-Beweise", d.h. die explizite Überprüfung von (3.1), für die Untersuchung der Konvergenz von Folgen selten notwendig; fast immer kann man die Behauptung durch geeignete Abschätzungen auf Folgen zurückführen, deren Konvergenzverhalten man kennt, und z.B. das Einschließungskriterium verwenden.

Beispiel.

$$\frac{3n^2 - 4n^5}{9n^2 - 5n^5} = \frac{\frac{3}{n^3} - 4}{\frac{9}{n^3} - 5} \to \frac{0 - 4}{0 - 5} = \frac{4}{5}.$$

3.2 Monotonie

Definition 3.2.1. Wir nennen eine Folge $(a_n)_n$ *monoton wachsend* (*monoton wachsende Folge*) falls $a_{n+1} \geq a_n$ bzw. *monoton fallend* (*monoton fallende Folge*) wenn $a_{n+1} \leq a_n$ für alle n gilt. Ersetzt man $\geq$ bzw. $\leq$ durch $>$ bzw. $<$, so nennt man sie *streng monoton wachsend* bzw. *streng monoton fallend*. Eine Folge die (streng) monoton fallend oder wachsend ist, nennt man *(streng) monoton*.

Man beachte hierbei, dass mit dieser Definition alle konstanten Folgen $(c) = (c, c, c, \ldots)$, $c \in \mathbb{R}$ sowohl monoton wachsend als auch monoton fallend sind.

Satz 3.2.2 (Monotonie des Limes). *Seien $(a_n), (b_n)$ zwei konvergente Folgen. Es gebe einen Index $n_0 \in \mathbb{N}$, sodass $a_n \leq b_n$ für alle $n \geq n_0$. Dann gilt*

$$\lim_{n\to\infty} a_n \leq \lim_{n\to\infty} b_n.$$

Beweis. Seien $a_n \to a, b_n \to b$ und sei $\epsilon > 0$. Es existiert ein n_1 mit $|a_n - a| < \frac{\epsilon}{2}, |b_n - b| < \frac{\epsilon}{2}$ für alle $n \geq n_1$. Damit gilt für alle $n \geq \max\{n_0, n_1\}$:

$$a = a_n + (a - a_n) \leq b_n + (a - a_n) = b + \underbrace{(b_n - b)}_{<\epsilon/2} + \underbrace{(a - a_n)}_{<\epsilon/2} < b + \epsilon$$

Da ϵ beliebig war, folgt $a \leq b$. □

Die Monotonie eröffnet uns die erste Möglichkeit, die Konvergenz von Folgen zu zeigen, ohne den Grenzwert zu kennen. Dabei verwenden wir die Schreibweise

$$\sup_{n \in \mathbb{N}} a_n = \sup\{a_n : n \in \mathbb{N}\}, \quad \inf_{n \in \mathbb{N}} a_n = \inf\{a_n : n \in \mathbb{N}\}$$

für das Supremum bzw. Infimum einer Folge (a_n).

Satz 3.2.3. *Ist eine monotone Folge beschränkt, so ist sie konvergent. Ist sie nicht beschränkt, so divergiert sie bestimmt gegen $+\infty$ oder $-\infty$.*

Beweis. Wir beweisen den Fall wo (a_n) monoton wachsend ist; für monoton fallende Folgen (b_n) folgt die Behauptung dann durch Betrachtung von $a_n = -b_n$.

Ist (a_n) beschränkt, zeigen wir dass $a_n \to a$ mit $a := \sup_n a_n \in \mathbb{R}$. Für beliebiges $\varepsilon > 0$ ist $a - \varepsilon$ keine obere Schranke dieser Menge, weswegen ein n_0 exisiert mit $a - \epsilon < a_{n_0} \leq a$. Aus der Monotonie folgt

$$a - \epsilon < a_{n_0} \leq a_n \leq a \quad \forall n \geq n_0$$

und damit $|a_n - a| < \epsilon$, also $a_n \to a$.

Ist (a_n) unbeschränkt, ist es – da $a_n \geq a_0$ für alle n – nach oben unbeschränkt. Sei $T > 0$; aus der Definition der (Un-)Beschränktheit nach oben gibt es dann n_0 mit $a_{n_0} > T$, wegen der Monotonie gilt dann auch $a_n > T$ für $n \geq n_0$, was genau die bestimmte Divergenz gegen ∞ bedeutet. □

Aus dem Beweis halten wir noch folgende Tatsache fest:

Folgerung 3.2.4. *Eine monoton wachsende (fallende), beschränkte Folge konvergiert gegen ihr Supremum (Infimum).*

Beispiel. Wir beweisen die Existenz des Grenzwertes einer rekursiv definierten Folge und berechnen diesen. Wir definieren (a_n) durch $a_0 := 1, a_{n+1} := \frac{5+2a_n}{10}$. Mittels vollständiger Induktion sieht man, dass (a_n) streng monoton fallend ist: tatsächlich gilt $a_{n+1} \leq a_n$, falls $a_n \geq 5/8$ ist; für n_0 ist $a_n \geq 5/8$ richtig, und induktiv folgt $a_{n+1} = (5 + 2a_n)/10 \geq 5/8$ für alle $n \geq 0$. Außerdem gilt $0 \leq a_n \leq a_0 = 1$. Nach Satz 3.2.3 ist diese Folge konvergent gegen ein $a \in \mathbb{R}$. Mit der gesicherten Existenz dieses Grenzwertes dürfen wir die folgende Rechnung anschreiben:

$$a = \lim_{n \to \infty} a_{n+1} = \frac{5 + \lim_{n \to \infty} a_n}{10} = \frac{5 + 2a}{10},$$

woraus $a = \frac{5}{8}$ folgt. Man beachte dabei, dass $\lim_{n \to \infty} a_n = \lim_{n \to \infty} a_{n+1}$, usw.

Lemma 3.2.5. Sei $\emptyset \neq M \subseteq \mathbb{R}$ nach oben beschränkt und $s := \sup M < \infty$. Dann gibt es eine Folge (a_n) in M mit $a_n \to s$ sowie eine Folge (b_n) in $\mathbb{R} \setminus M$ mit $b_n \to s$. (a_n) heißt *Maximalfolge* von M.

Beweis. Weil s das Supremum von M ist, gibt es für jedes $n \in \mathbb{N}$ ein $a_n \in M$, sodass $s - \frac{1}{n} < a_n \leq s$, da sonst $s - \frac{1}{n}$ eine obere Schranke für M wäre. Die so konstruierte Folge konvergiert offensichtlich gegen s. Ebenso sieht man, dass $b_n := s + \frac{1}{n}$ nicht in M liegt, und (b_n) ist eine Minimalfolge. □

Eine analoge Aussage gilt für das Infimum einer nach unten beschränkten Menge, d.h. man findet dafür eine Minimalfolge.

Bemerkung. Im Falle $\sup M = \infty$ bzw. $\inf M = -\infty$ gibt es eine Folge (x_n) in M mit $x_n \to \infty$ bzw. $x_n \to -\infty$.

3.3 Cauchyfolgen

Die Definition der Konvergenz beruht auf der Kenntnis des Grenzwertes. In der Praxis ist dieser jedoch oft unbekannt bzw. nur näherungsweise bestimmbar. Folgender Begriff wird uns erlauben, das Konvergenzverhalten einer Folge zu studieren, ohne den Grenzwert zu kennen. Wesentlich dafür ist die *Vollständigkeit* von $\mathbb{R}$.

Definition 3.3.1. Wir nennen eine Folge (x_n) eine *Cauchyfolge*[1], wenn

$$\forall \epsilon > 0 \; \exists n_0 \in \mathbb{N} \; \forall n, m \geq n_0 \colon |a_n - a_m| < \epsilon.$$

Satz 3.3.2. *Jede Cauchyfolge ist beschränkt.*

Beweis. Sei (x_n) eine Cauchyfolge. Für $\epsilon > 0$ gibt es ein $n_0 \in \mathbb{N}$ mit $|x_n - x_m| < \epsilon$ für $n, m \geq n_0$. Damit ist auch $|x_n - x_{n_0}| < \epsilon$ für $n \geq n_0$, also $|x_n| \leq |x_{n_0}| + \epsilon$. Insgesamt ist $|x_n| \leq \max\{|x_1|, \ldots, |x_{n_0}|, |x_{n_0}| + \epsilon\}$, also ist (x_n) beschränkt. □

Satz 3.3.3. *Eine reelle Folge ist genau dann konvergent, wenn sie eine Cauchyfolge ist.*

Beweis. Sei zunächst $(x_n)_n$ eine konvergente Folge mit Grenzwert $x \in \mathbb{R}$. Für beliebiges $\epsilon > 0$ gibt es ein $n_0 \in \mathbb{N}$ mit $|x_k - x| < \epsilon/2$ für $k \geq n_0$. Sind nun $n, m \geq n_0$, so folgt aus der Dreiecksungleichung $|x_n - x_m| \leq |x_n - x| + |x_m - x| < \epsilon$.

Sei nun $(x_n)_n$ eine Cauchyfolge in $\mathbb{R}$. Wir zeigen, dass es ein $x \in \mathbb{R}$ gibt mit $x_n \to x$. Für $N \in \mathbb{N}$ definieren wir $\tilde{x}_N := \inf\{x_N, x_{N+1}, \ldots\}$. Aus Lemma 2.4.6 folgt $\tilde{x}_N \leq \tilde{x}_{N+1}$, somit ist $(\tilde{x}_N)_N$ monoton wachsend und als Cauchyfolge nach Satz 3.3.2 auch beschränkt, insgesamt also nach Satz 3.2.3 konvergent gegen ein $x := \lim_{N \in \mathbb{N}} \tilde{x}_N \in \mathbb{R}$. Wir zeigen nun noch, dass tatsächlich $\lim_{n\to\infty} x_n = x$ gilt. Sei dazu $\epsilon > 0$ beliebig. Es existiert ein $N_0 \in \mathbb{N}$

[1] Benannt nach dem französischen Mathematiker Augustin-Louis Cauchy (1789 - 1857).

mit $|x_n - x_m| < \frac{\epsilon}{3}$ für $m, n \geq N_0$, sowie $|\tilde{x}_N - x| < \frac{\epsilon}{3}$ für alle $N \geq N_0$. Nach Lemma 3.2.5 gibt es eine Folge in $\{x_n : n \geq N_0\}$ die gegen $\tilde{x}_{N_0}$ konvergiert, also existiert insbesondere ein $n_0 \geq N_0$ mit $|x_{n_0} - \tilde{x}_{N_0}| < \frac{\epsilon}{3}$. Für $n \geq n_0$ gilt daher

$$|x_n - x| \leq |x_n - x_{n_0}| + |x_{n_0} - \tilde{x}_{N_0}| + |\tilde{x}_{N_0} - x| < \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon. \qquad \square$$

Eine wichtige Voraussetzung von Satz 3.3.3 ist die Vollständigkeit von $\mathbb{R}$. $\mathbb{Q}$ hingegen ist nicht vollständig: es gibt Folgen (x_n) in $\mathbb{Q}$, die in $\mathbb{Q}$ keinen Grenzwert haben. Für die nach Satz 2.1.1 irrationale Zahl $\sqrt{2}$ gibt es eine unendliche Dezimaldarstellung

$$\sqrt{2} = 1.41421356237\ldots$$

Wir definieren eine rationale Folge (q_n) durch das Abschneiden vor der $n-$ten Nachkommastelle, also

$$\begin{aligned} q_1 &:= 1 \\ q_2 &:= 1.4 \\ q_3 &:= 1.41 \\ q_4 &:= 1.414 \\ &\vdots \end{aligned}$$

Die so konstruierte Folge (q_n) konvergiert in $\mathbb{R}$ gegen $\sqrt{2}$, ist also auch eine Cauchyfolge; sie konvergiert aber nicht in $\mathbb{Q}$.

Andererseits kommt man mit den rationalen Zahlen den reellen beliebig nahe:

Lemma 3.3.4. Ist $x \in \mathbb{R}$ dann gibt es eine Folge (x_n) in $\mathbb{Q}$ mit $x_n \to x$.

Beweis. Das folgt direkt aus Satz 2.4.9 indem man $x < x_n < x + 1/n$ wählt. $\square$

3.4 Teilfolgen

Definition 3.4.1. Sei $(x_n)_{n\in\mathbb{N}}$ eine Folge und $(n_k)_{k\in\mathbb{N}}$ eine streng monoton wachsende Folge natürlicher Zahlen. Dann heißt $(x_{n_k})_{k\in\mathbb{N}}$ eine *Teilfolge* von (x_n).

Beispiel.

(i) Die Folge aller geraden Zahlen, $(2n)_{n\in\mathbb{N}} = (2, 4, 6, \ldots)$, ist eine Teilfolge der Folge $(n)_{n\in\mathbb{N}}$ aller natürlichen Zahlen.

(ii) Die konstanten Folgen $(1, 1, 1, \ldots)$ und $(-1, -1, -1, \ldots)$ sind (konvergente) Teilfolgen der (divergenten) Folge $((-1)^n)_{n\in\mathbb{N}}$.

Konvergenz einer Folge vererbt sich auch auf Teilfolgen:

Satz 3.4.2. *Ist die Folge (x_n) konvergent gegen $x \in \mathbb{R}$, dann auch jede Teilfolge. Divergiert (x_n) bestimmt gegen $\pm\infty$, dann auch jede Teilfolge.*

Beweis. Sei $x_n \to x$ und sei $b_k = x_{n_k}$ eine Teilfolge. Zu gegebenem $\varepsilon > 0$ gibt es ein $n_0 \in \mathbb{N}$, so dass für jedes $n \geq n_0$ die Beziehung $|x_n - x| < \varepsilon$ gilt. Es gibt ein $k_0 \in \mathbb{N}$, sodass für jedes $k \geq k_0$ der Index n_k größer als n_0 ist. Also folgt für jedes $k \geq k_0$, dass $|b_k - x| = |x_{n_k} - x| < \varepsilon$, und damit die Konvergenz der Teilfolge gegen x. Analog zeigt man die zweite Aussage. □

Satz 3.4.3 (Satz von Bolzano-Weierstraß). *Jede beschränkte reelle Folge enthält eine konvergente Teilfolge.*

Beweis. Wir zeigen, dass eine beschränkte Folge (x_n) eine monotone Teilfolge enthält. Sei $M := \{m \in \mathbb{N} \mid \forall n > m \text{ gilt } x_n < x_m\}$. Wir unterscheiden zwei Fälle:

Fall 1: M ist unendlich. Seien $m_1 < m_2 < m_3 < \ldots \in M$. Nach der Definition von M gilt dann $x_{m_1} > x_{m_2} > x_{m_3} \ldots$, also ist $(x_{m_k})_k$ eine monoton fallende Teilfolge von (x_n).

Fall 2: Ist M endlich, dann auch beschränkt und für alle $m > \sup M$ ist $m \notin M$, also

$$\forall m > \sup M \; \exists n_m > m \colon x_{n_m} \geq x_m.$$

Sei $n_1 \in \mathbb{N}$, $n_1 > \sup M$ beliebig. Für $k \in \mathbb{N}$ sei $n_{k+1} > n_k$ so gewählt, dass $x_{n_{k+1}} \geq x_{n_k}$. Damit ist (n_k) rekursiv definiert und die Folge (x_{n_k}) ist monoton wachsend.

In beiden Fällen haben wir eine monotone Teilfolge der ursprünglichen Folge gefunden, die als Teilfolge einer beschränkten Folge ebenso beschränkt ist und nach Satz 3.2.3 dann konvergent ist. □

Analog zu reellen Folgen gibt es Folgen in den komplexen Zahlen, also Abbildungen $f\colon \mathbb{N} \to \mathbb{C}$. Sinngemäß gelten auch die Definition 3.1.1, Definition 3.1.2, **??**, Definition 3.1.5, Definition 3.3.1 und Definition 3.4.1, sowie die Aussagen Satz 3.1.3, Satz 3.1.6, Folgerung 3.1.8, Satz 3.1.9, Satz 3.1.10, Satz 3.3.2, Satz 3.3.3, Satz 3.4.2 (ohne 2. Teil), sowie Satz 3.4.3. Aufgrund der in Abschnitt 2.2 behandelten Eigenschaften der komplexen Zahlen ergibt es keinen Sinn, den Begriff „Monotonie" für komplexe Folgen einzuführen.

Man beachte hierbei, dass eine komplexe Folge $(z_n)_n$ genau dann konvergiert, wenn die reellen Folgen $(\operatorname{Re} z_n)_n$ und $(\operatorname{Im} z_n)_n$ konvergieren; in diesem fall ist $\lim_{n\to\infty} z_n = \lim_{n\to\infty} \operatorname{Re} z_n + i \lim_{n\to\infty} \operatorname{Im} z_n$. Daraus folgt sofort, dass in $\mathbb{C}$ eine Folge genau dann konvergiert, wenn sie eine Cauchyfolge ist. Ebenso ist $(z_n)_n$ beschränkt in $\mathbb{C}$, d.h. $\exists R > 0$ mit $|z_n| \leq R$ für alle $n \in \mathbb{N}$, genau dann wenn $(\operatorname{Re} z_n)_n$ und $(\operatorname{Im} z_n)_n$ beschränkt sind.

Kapitel 4

Reihen

Lernziele

- Zusammenhang und Unterschied zwischen Folgen und Reihen verstehen und benennen
- Konvergenzkriterien (Cauchykriterium, Majoranten/Minorantenkriterium, Quotientenkriterium, Wurzelkriterium, Leibnizsches Kriterium) anwenden
- Unteschied zwischen Konvergenz und absoluter Konvergenz benennen
- Folgerungen der absoluten Konvergenz anführen
- Umordnungen von Reihen verstehen
- Cauchy-Produkt von Reihen berechnen
- Exponentialfunktion samt wichtigsten Eigenschaften angeben

4.1 Definition und Beispiele

Wir widmen uns in diesem Kapitel unendlichen Summen der Form

$$a_1 + a_2 + a_3 + \ldots = \sum_{k=1}^{\infty} a_k,$$

wobei $(a_n)_n$ eine Folge reeller Zahlen ist. Um diesem Ausdruck einen Sinn zu geben, bilden wir die Partialsummen

$$s_n := a_1 + a_2 + \ldots + a_n = \sum_{k=1}^{n} a_k.$$

Definition 4.1.1. Die Folge $(s_n)_n$ heißt *Reihe*, a_k ist der k-te Summand der Reihe. Existiert der Grenzwert der Folge $(s_n)_n$, so schreiben wir

$$\sum_{k=1}^{\infty} a_k = \lim_{n\to\infty} \sum_{k=1}^{n} a_k$$

und nennen die Reihe $\sum_{k=1}^{\infty} a_k$ *konvergent*, ansonsten *divergent*.

Man bezeichnet die Folge der Partialsummen selbst oft auch mit $\sum_{k=1}^{\infty} a_k$:

Beispiel 4.1.2.

(i) Die *geometrische Reihe* $\sum_{n=0}^{\infty} x^n$ konvergiert für $|x| < 1$ und hat den Grenzwert

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

Aus Satz 1.5.4 wissen wir nämlich $s_n = \sum_{k=0}^{n} x^k = \frac{1-x^{n+1}}{1-x}$; x^{n+1} konvergiert wegen $|x| < 1$ für $n \to \infty$ gegen 0 und s_n daher gegen $\frac{1}{1-x}$.

Für $x = \frac{1}{2}$ erhält man beispielweise $1 + \frac{1}{2} + \frac{1}{4} + \ldots = 2$.

(ii) Die Reihe $\sum_{k=1}^{\infty} \frac{1}{k(k+1)}$ konvergiert und hat den Grenzwert 1: Es gilt $\frac{1}{k(k+1)} = \frac{1}{k} - \frac{1}{k+1}$, also ist die n-te Partialsumme gleich

$$(1-\frac{1}{2}) + (\frac{1}{2} - \frac{1}{3}) + \cdots + (\frac{1}{n} - \frac{1}{n+1}) = 1 - \frac{1}{n+1},$$

da sich die mittleren Glieder stets paarweise aufheben. Für $n \to \infty$ geht $\frac{1}{n+1}$ gegen 0, woraus die Behauptung folgt.

(iii) Die Reihe $\sum_{k=1}^{\infty} \frac{1}{k^2}$ konvergiert, denn die Folge der Partialsummen ist monoton wachsend und nach oben beschränkt:

$$\begin{aligned} s_n &= 1 + \frac{1}{2^2} + \frac{1}{3^2} + \ldots + \frac{1}{n^2} \\ &< 1 + \frac{1}{1 \cdot 2} + \frac{1}{2 \cdot 3} + \ldots + \frac{1}{(n-1)n} = 1 + 1 - \frac{1}{n} < 2. \end{aligned}$$

(iv) Die *harmonische Reihe* $\sum_{k=1}^{\infty} \frac{1}{k}$ divergiert, denn die Folge der Partialsummen ist unbeschränkt:

$$\begin{aligned} s_{2^n} &= 1 + \frac{1}{2} + \left(\frac{1}{3} + \frac{1}{4}\right) + \left(\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}\right) + \ldots + \\ &\quad + \left(\frac{1}{2^{n-1}+1} + \ldots + \frac{1}{2^n}\right) \\ &> 1 + \frac{1}{2} + 2\frac{1}{4} + 4\frac{1}{8} + \ldots + 2^{n-1}\frac{1}{2^n} = 1 + \frac{n}{2}. \end{aligned}$$

Satz 4.1.3 (Linearkombinationen konvergenter Reihen). *Sind $\sum_{n=1}^{\infty} a_n$ und $\sum_{n=1}^{\infty} b_n$ konvergente Reihen, so ist für $\lambda, \mu \in \mathbb{R}$ auch die Reihe $\sum_{n=1}^{\infty} (\lambda a_n + \mu b_n)$ konvergent und für die Grenzwerte gilt*

$$\sum_{n=1}^{\infty} (\lambda a_n + \mu b_n) = \lambda \sum_{n=1}^{\infty} a_n + \mu \sum_{n=1}^{\infty} b_n.$$

Beweis. Folgt direkt aus Satz 3.1.10. □

4.2 Konvergenzkriterien für Reihen

Zusätzlich zu den Methoden, die uns für Folgen schon zur Verfügung stehen, gibt es für Reihen weitere Konvergenzkriterien.

Satz 4.2.1 (Cauchy-Kriterium). *Eine Reihe $\sum_{n=1}^{\infty} a_n$ konvergiert genau dann in $\mathbb{R}$, wenn*

$$\forall \varepsilon > 0\ \exists n_0 \in \mathbb{N}\ \forall m, n \geq n_0 : \left|\sum_{k=n}^{m} a_k\right| < \epsilon.$$

Beweis. Das folgt unmittelbar aus Satz 3.3.3, da $s_m - s_n = \sum_{k=n+1}^{m} a_k$ für $m > n$. □

Satz 4.2.2. *Konvergiert die Reihe $\sum_{n=1}^{\infty} a_n$, dann ist (a_n) eine Nullfolge.*

Beweis. Nach Satz 4.2.1 gibt es für $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$ sodass $\forall m, n \geq n_0$ gilt: $\left|\sum_{k=n}^{m} a_n\right| < \varepsilon$. Insbesondere gilt das für $m = n$, also $\left|\sum_{k=n}^{n} a_k\right| = |a_n| < \varepsilon$ für $n \geq n_0$, was bedeutet dass (a_n) eine Nullfolge ist. □

Die Umkehrung gilt im Allgemeinen nicht: wie wir oben gesehen haben ist $1/n$ eine Nullfolge, aber die Reihe $\sum_{n=1}^{\infty} 1/n$ konvergiert nicht!

Definition 4.2.3. Eine Reihe $\sum_{n=1}^{\infty} a_n$ heißt *absolut konvergent*, wenn die Reihe der Absolutbeträge $\sum_{n=1}^{\infty} |a_n|$ konvergiert.

Da $\sum_{k=1}^{n} |a_k|$ monoton wachsend ist, reicht es für die absolute Konvergenz, wenn diese Folge der Teilsummen nach oben beschränkt ist; kurz schreibt man dafür $\sum_{n=1}^{\infty} |a_n| < \infty$.

Satz 4.2.4. *Jede absolut konvergente Reihe ist konvergent.*

Beweis. Sei $\varepsilon > 0$. Mit dem Cauchykriterium erhalten wir ein $n_0 \in \mathbb{N}$ sodass $\forall m, n \geq n_0$ gilt: $\sum_{k=n}^{m} |a_k| < \varepsilon$. Mit der Dreiecksungleichung folgern wir

$$\left|\sum_{k=n}^{m} a_k\right| \leq \sum_{k=n}^{m} |a_k| \leq \varepsilon,$$

also ist $\sum_{k=1}^{\infty} a_k$ nach dem Cauchykriterium konvergent. □

31. 10. 2024

Nicht jede konvergente Reihe ist absolut konvergent: die *alternierende harmonische Reihe*

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

konvergiert nach dem Leibnizschen[1] Konvergenzkriterium, die Reihe der Absolutbeträge ist jedoch die harmonische Reihe $\sum_{n=1}^{\infty} 1/n$, welche divergiert.

[1] Gottfried Wilhelm Leibniz, (1646-1716); gilt als einer der Begründer der Infinitesmalrechnung

Satz 4.2.5 (Leibnizsches Konvergenzkriterium). *Ist (a_n) eine monoton fallende Nullfolge, dann konvergiert die Reihe $\sum_{n=1}^{\infty}(-1)^{n+1}a_n$.*

Beweis. Sei (s_n) die n-te Partialsumme, also $s_n = \sum_{k=1}^{n}(-1)^{n+1}a_n$. Dann ist

$$s_{2n} = \underbrace{(a_1 - a_2)}_{\geq 0} + \underbrace{(a_3 - a_4)}_{\geq 0} + \underbrace{(a_5 - a_6)}_{\geq 0} + \ldots + \underbrace{(a_{2_n-1} - a_{2n})} \quad \text{monoton wachsend,}$$
$$s_{2n+1} = a_1 - \underbrace{(a_2 - a_3)}_{\geq 0} - \underbrace{(a_4 - a_5)}_{\geq 0} - \ldots - \underbrace{(a_{2n} - a_{2n+1})}_{\geq 0} \quad \text{monoton fallend.}$$

Da $0 \leq s_{2n} \leq s_{2n+1} \leq a_1$, sind beide Folgen auch beschränkt und es existieren die Grenzwerte $S = \lim_{n\to\infty} s_{2n}$ und $S' = \lim_{n\to\infty} s_{2n+1}$. Wegen $S' - S = \lim_{n\to\infty}(s_{2n+1} - s_{2n}) = \lim_{n\to\infty} a_{2n+1} = 0$ ist $S = S'$. Damit ist auch s_n konvergent, d.h. die Reihe konvergiert. □

Die Dreiecksungleichung gilt in folgender Form auch für absolut konvergente Reihen:

Lemma 4.2.6. Ist die Reihe $\sum_{n=1}^{\infty} a_n$ absolut konvergent, so gilt

$$|\sum_{n=1}^{\infty} a_n| \leq \sum_{n=1}^{\infty} |a_n|.$$

Beweis. Für jedes $N \in \mathbb{N}$ gilt

$$\pm\sum_{n=1}^{N} a_n \leq \sum_{n=1}^{N} |a_n| \leq \sum_{n=1}^{\infty} |a_n|.$$

Durch Grenzübergang $N \to \infty$ folgt die Behauptung. □

Satz 4.2.7 (Konvergenz einer Teilsumme). *Ist die Reihe $\sum_{n=1}^{\infty} a_n$ absolut konvergent und $(a_{n_k})_{k\in\mathbb{N}}$ eine Teilfolge von $(a_n)_n$, dann konvergiert auch die Reihe $\sum_{k=1}^{\infty} a_{n_k}$ absolut.*

Beweis. Für jedes $K \in \mathbb{N}$ gilt

$$\sum_{k=1}^{K} |a_{n_k}| \leq \sum_{n=1}^{n_K} |a_n| \leq \sum_{n=1}^{\infty} |a_n|,$$

so dass nach Grenzübergang $K \to \infty$ die Ungleichung $\sum_{k=1}^{\infty} |a_{n_k}| \leq \sum_{n=1}^{\infty} |a_n| < \infty$ folgt. □

Ohne die absolute Konvergenz gilt das nicht: die alternierende harmonische Reihe

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

ist zwar konvergent, die Teilsumme mit geraden Indizes

$$\sum_{n=1}^{\infty} \frac{-1}{2n} = -\frac{1}{2} - \frac{1}{4} - \frac{1}{6} - \ldots = -\frac{1}{2}\sum_{n=1}^{\infty} \frac{1}{n}$$

jedoch divergent.

Der Ausdruck „für fast alle $n \in \mathbb{N}$" in folgendem Satz und auch später bedeutet „Für alle bis auf endlich viele $n \in \mathbb{N}$", insbesondere also auch für alle $n \geq n_0$ für ein $n_0 \in \mathbb{N}$.

Satz 4.2.8 (Majorantenkriterium/Minorantenkriterium).

(i) Ist $0 \leq a_n \leq c_n$ für fast alle $n \in \mathbb{N}$ und $\sum_{n=1}^{\infty} c_n$ konvergent, dann ist auch $\sum_{n=1}^{\infty} a_n$ konvergent.

(ii) Ist $0 \leq d_n \leq a_n$ für fast alle $n \in \mathbb{N}$ und $\sum_{n=1}^{\infty} d_n$ divergent, dann ist auch $\sum_{n=1}^{\infty} a_n$ divergent.

Die Reihe $\sum c_n$ heißt in diesem Fall konvergente Majorante *und $\sum d_n$* divergente Minorante.

Beweis. Ohne Beschränkung der Allgemeinheit können wir annehmen, dass die Ungleichungen für alle $n \in \mathbb{N}$ gelten, da wir endlich viele Reihenglieder immer abändern könnne ohne dass sich das Konvergenzverhalten der Reihe ändert.

(i) Aus der Ungleichung

$$0 \leq \sum_{n=1}^{N} a_n \leq \sum_{n=1}^{N} c_n \leq \sum_{n=1}^{\infty} c_n$$

sieht man, dass die Partialsummen $\sum_{n=1}^{N} a_n$ monoton wachsend und beschränkt sind, also konvergent.

(ii) Aus der Ungleichung

$$0 \leq \sum_{n=1}^{N} d_n \leq \sum_{n=1}^{N} a_n$$

sieht man, da $\sum_{n=1}^{N} d_n$ bestimmt gegen $+\infty$ divergiert, dass $\sum_{n=1}^{N} a_n$ für $N \to \infty$ nicht konvergieren kann. □

Beispiel.

(i) Die Reihe $\displaystyle\sum_{n=1}^{\infty} \frac{1}{3n+1}$ divergiert, denn $\frac{1}{3n+1} \geq \frac{1}{4n}$ und $\frac{1}{4}\sum_{n=1}^{\infty}\frac{1}{n}$ divergiert.

(ii) Die Reihe $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n^k}$ konvergiert für $k \geq 2$, denn $\frac{1}{n^k} \leq \frac{1}{n^2} \leq \frac{2}{n(n+1)}$, und $2\sum_{n=1}^{\infty}\frac{1}{n(n+1)}$ konvergiert nach Beispiel 4.1.2 (ii).

Satz 4.2.9 (Quotientenkriterium). *Sei $\sum_{n=1}^{\infty} a_n$ eine Reihe mit $a_n \neq 0$ für fast alle $n \in \mathbb{N}$.*

(i) Gibt es ein $q \in (0,1)$ und ein $n_0 \in \mathbb{N}$ mit $\left|\frac{a_{n+1}}{a_n}\right| \leq q$ für alle $n \geq n_0$, dann konvergiert die Reihe $\sum_{n=1}^{\infty} a_n$ absolut.

(ii) Gibt es ein $n_0 \in \mathbb{N}$ mit $\left|\frac{a_{n+1}}{a_n}\right| \geq 1$ für alle $n \geq n_0$, dann divergiert die Reihe.

Beweis. Ohne Beschränkung der Allgemeinheit sei $a_n \neq 0$ für alle n und $n_0 = 1$, denn für das Konvergenzverhalten von Reihen sind die Werte endlich vieler Glieder bis zu einem (endlichen) Index nicht relevant, man darf sie demnach derart abändern, dass $|\frac{a_{n+1}}{a_n}| \leq q$ bzw. $|\frac{a_{n+1}}{a_n}| \geq 1$ für alle $n \in \mathbb{N}$ gilt. Daraus folgt in (i)

$$|a_{n+1}| \leq q|a_n| \leq \ldots \leq q^n |a_1|,$$

also ist die Reihe $\sum_{n=1}^{\infty} q^n |a_1|$ eine konvergente Majorante. In (ii) sehen wir $|a_n| \geq |a_{n-1}| \geq \ldots \geq |a_1| \neq 0$ für alle n, also ist $(a_n)_n$ keine Nullfolge und die Reihe kann nicht konvergieren. □

Bemerkung. Gilt nur $\left|\frac{a_{n+1}}{a_n}\right| \leq 1$, kann die Reihe konvergieren oder divergieren. Für die harmonische Reihe $\sum_{n=1}^{\infty} \frac{1}{n}$ gilt

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

und für die Reihe $\sum_{n=1}^{\infty} \frac{1}{n^2}$

$$\left|\frac{a_{n+1}}{a_n}\right| = \left(\frac{n}{n+1}\right)^2 < 1.$$

In beiden Fällen konvergiert $\left|\frac{a_{n+1}}{a_n}\right|$ gegen 1, also kann es kein $q < 1$ mit $\left|\frac{a_{n+1}}{a_n}\right| \leq q$ geben. Die erste Reihe divergiert, die zweite konvergiert aber.

Satz 4.2.10 (Wurzelkriterium). *Sei $\sum_{n=1}^{\infty} a_n$ eine Reihe. Gilt für fast alle $n \in \mathbb{N}$*
(i) $\sqrt[n]{|a_n|} \leq q < 1$, dann ist $\sum_{n=1}^{\infty} a_n$ absolut konvergent;
(ii) $\sqrt[n]{|a_n|} \geq 1$, dann ist $\sum_{n=1}^{\infty} a_n$ divergent.

Beweis. (i): Aus $\sqrt[n]{|a_n|} \leq q \in (0,1)$ folgt $|a_n| \leq q^n$, also ist die geometrische Reihe eine konvergente Majorante.

(ii): Ist $\sqrt[n]{|a_n|} \geq 1$, dann ist $|a_n| \geq 1$, $(a_n)_n$ ist also keine Nullfolge und $\sum_{n=1^{\infty}} a_n$ kann nicht konvergieren. □

Gilt nur $\sqrt[n]{|a_n|} \leq 1$, ist wieder alles möglich:

Beispiel. Wir betrachten wir die Reihen

$$\sum_{n=1}^{\infty} \frac{1}{n} \quad (divergent) \quad \text{und} \quad \sum_{n=1}^{\infty} \frac{1}{n^2} \quad (konvergent).$$

Es gilt mit Beispiel 4.2.11

$$\sqrt[n]{\frac{1}{n}} = \frac{1}{\sqrt[n]{n}} \to 1, \quad \sqrt[n]{\frac{1}{n^2}} = \left(\frac{1}{\sqrt[n]{n}}\right)^2 \to 1.$$

Beide Folgen sind wegen $\sqrt[n]{n} \geq 1$ immer ≤ 1, man findet wegen der Konvergenz gegen 1 aber auch kein $q \in (0,1)$ als obere Schranke dafür.

Beispiel 4.2.11. *$\sqrt[n]{n} \to 1$ für $n \to \infty$. Tatsächlich ist $\sqrt[n]{n} \geq 1$, also können wir $\sqrt[n]{n} = 1 + \delta_n$ für ein $\delta_n \geq 0$ schreiben. Damit folgt aus der binomischen Formel für $n \geq 2$*

$$n = (1 + \delta_n)^n \geq \frac{n(n-1)}{2}\delta_n^2$$

und somit

$$\delta_n^2 \leq \frac{2}{n-1}, \quad \textit{d.h.} \quad \delta_n \leq \sqrt{\frac{2}{n-1}}.$$

Nun konvergiert $\sqrt{\frac{2}{n-1}}$ gegen 0 für $n \to \infty$, denn für $\varepsilon > 0$ ist $\sqrt{\frac{2}{n-1}} < \varepsilon$ sobald $n > \frac{2}{\varepsilon^2} + 1$ ist.

Beispiel.

- Sei $x \in (0, 1)$ und sei $k \in \mathbb{N}$. Dann konvergiert die Reihe $\sum_{n=1}^{\infty} n^k x^n$ nach dem Quotientenkriterium: für $a_n = n^k x^n$ ist $a_n \neq 0$ für jedes $n \in \mathbb{N}$ und es gilt

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

  Die Folge $\frac{n+1}{n} = 1 + \frac{1}{n}$ konvergiert gegen 1, also konvergiert $\left|\frac{a_{n+1}}{a_n}\right|$ gegen $x < 1$. Sei $x < q < 1$ beliebig, dann existiert also ein n_0 mit $\left|\frac{a_{n+1}}{a_n}\right| \leq q$ für jedes $n \geq n_0$ und damit folgt die absolute Konvergenz nach dem Quotientenkriterium.
- Die Summe $\displaystyle\sum_{n=0}^{\infty} \frac{n^{1-3n}}{4^{2n}}$ konvergiert, denn

  $$\sqrt[n]{|a_n|} = \frac{\sqrt[n]{n} \cdot (\frac{1}{n})^3}{4^2} \to \frac{1 \cdot 0}{16} = 0.$$

7. 11. 2024

4.3 Umordnungen

Sei $\sigma\colon \mathbb{N} \to \mathbb{N}$ eine bijektive Abbildung. Die Reihe $\sum_{n=1}^{\infty} a_{\sigma(n)}$ heißt dann *Umordnung* der Reihe $\sum_{n=1}^{\infty} a_n$. Man summiert also die Glieder nur in einer anderen Reihenfolge auf. Man beachte hierbei, dass die beiden Reihen im Allgemeinen nicht denselben Wert ergeben müssen, selbst wenn über dieselben Werte aufsummiert wird! Eine konvergente Reihe heißt *unbedingt konvergent*, wenn sie bei einer beliebigen Umordnung konvergent bleibt und ihr Wert sich nicht ändert; sonst heißt sie *bedingt konvergent*.

Satz 4.3.1 (Umordnungssatz). *Jede absolut konvergente Reihe ist unbedingt Konvergent, und jede Umordnung ist dann auch absolut konvergent.*

Beweis. Sei $\sigma : \mathbb{N} \to \mathbb{N}$ eine Bijektion und $A := \sum_{n=1}^{\infty} a_n$. Es ist zu zeigen, dass die Folge $s_n = \sum_{k=1}^{n} a_{\sigma(k)}$ ebenfalls gegen A konvergiert. Sei $\epsilon > 0$, so existiert wegen der absoluten Konvergenz ein $n_0 \in \mathbb{N}$ mit $\left|\sum_{n=1}^{n_0-1} |a_n| - \sum_{n=1}^{\infty} |a_n|\right| = \sum_{n=n_0}^{\infty} |a_n| < \epsilon$. Sei nun

$n_1 \geq n_0$ so groß, dass für jedes $n \geq n_1$ die Menge $S(n) := \{\sigma(1), \sigma(2), \ldots, \sigma(n)\}$ alle Zahlen $\{1, \ldots, n_0 - 1\}$ umfasst. Für jedes $n \geq n_1$ gilt dann

$$|s_n - A| = \Big| \sum_{k=1}^{n} a_{\sigma(n)} - \sum_{k=1}^{\infty} a_k \Big| = \Big| \sum_{\substack{k=1 \\ k \notin S(n)}}^{\infty} a_k \Big| \leq \sum_{\substack{k=1 \\ k \notin S(n)}}^{\infty} |a_k| \leq \sum_{k=n_0}^{\infty} |a_k| < \epsilon.$$

Das selbe Argument auf $|a_n|$ angewendet gibt die zweite Behauptung. □

Umgekehrt kann man auch zeigen, dass jede unbedingt konvergente Reihe in $\mathbb{R}$ absolut konvergent ist, d.h. diese beiden Begriffe sind (zumindest für Reihen reeller Zahlen) äquivalent.

Beispiel. Nach Satz 4.2.5 und Beispiel 4.1.2 ist

$$S := \sum_{n=1}^{\infty} \frac{(-1)^n}{n} = -1 + \frac{1}{2} - \frac{1}{3} + \frac{1}{4} - \ldots$$

konvergent, aber nicht absolut konvergent. Für die Bijektion $\sigma : \mathbb{N} \to \mathbb{N}$ definiert durch

$$\begin{aligned} &\sigma(3k-2) = 2k-1, \\ &\sigma(3k-1) = 4k-2 \quad \text{und} \\ &\sigma(3k) = 4k \end{aligned}$$

für $k \in \mathbb{N}$, erhalten wir

$$\begin{aligned} \sum_{n=1}^{\infty} \frac{(-1)^{\sigma(n)}}{\sigma(n)} &= (-1 + \frac{1}{2}) + \frac{1}{4} + (-\frac{1}{3} + \frac{1}{6}) + \frac{1}{8} + (-\frac{1}{5} + \frac{1}{10}) + \frac{1}{12} + \ldots \\ &= -\frac{1}{2} + \frac{1}{4} - \frac{1}{6} + \frac{1}{8} - \frac{1}{10} + \frac{1}{12} + \ldots \\ &= \frac{1}{2} \sum_{n=1}^{\infty} \frac{(-1)}{n} = \frac{S}{2}. \end{aligned}$$

Damit nehmen die ursprüngliche Reihe und ihre Umordnung verschiedene Werte an.

Die Umordnung von Reihen, die konvergent, aber nicht absolut konvergent sind, haben also im Allgemeinen verschiedene Grenzwerte. Es gilt sogar der folgende (erstaunliche)

Satz 4.3.2 (Riemann'scher Umordnungssatz)**.** *Sei $\sum_{n=1}^{\infty} a_n$ eine konvergente, aber nicht absolut konvergente Reihe. Dann gibt es zu jedem $a \in \mathbb{R} \cup \{\pm\infty\}$ eine Umordnung der Reihe, die gegen a konvergiert.*

Der Beweis geht für uns zu weit.

4.4 Das Produkt von Reihen

Schließlich betrachten wir noch das Produkt von Reihen.

Satz 4.4.1 (Cauchy-Produkt von Reihen). *Seien $\sum_{n=0}^{\infty} a_n$ und $\sum_{n=0}^{\infty} b_n$ absolut konvergente Reihen. Setze*

$$c_n = \sum_{k=0}^{n} a_k b_{n-k} = \sum_{k+l=n} a_k b_l.$$

Dann ist auch die Reihe $\sum_{n=0}^{\infty} c_n$ absolut konvergent und es gilt

$$\sum_{n=0}^{\infty} c_n = \Big(\sum_{n=0}^{\infty} a_n\Big)\Big(\sum_{n=0}^{\infty} b_n\Big).$$

Beweis. Sei $\sigma_N = \sum_{n=0}^{N} |c_n|$. Diese Folge ist monoton wachsend und es gilt

$$\begin{aligned}|\sigma_N| &\leq \sum_{n=0}^{N} \sum_{k+l=n} |a_k||b_l| = \sum_{k+l\leq N} |a_k||b_l| \\ &\leq \sum_{k\leq N, l\leq N} |a_k||b_l| = \sum_{k=0}^{N} |a_k| \sum_{l=0}^{N} |b_l| \leq \sum_{k=0}^{\infty} |a_k| \sum_{l=0}^{\infty} |b_l|,\end{aligned}$$

also ist $(\sigma_N)_N$ nach oben beschränkt und damit $\sum_N c_N$ absolut konvergent. Um zu zeigen, dass $\sum c_n = \sum a_n \cdot \sum b_n$ ist, setzen wir

$$I_n := \{(k,l) \mid k+l \leq 2N \text{ und } (k > N \text{ oder } l > N)\}.$$

Dann ist

$$\begin{aligned}\left|\sum_{n=0}^{2N} c_n - \sum_{k=0}^{N} a_k \cdot \sum_{l=0}^{N} b_l\right| &= \left|\sum_{(k,l)\in I_N} a_k b_l\right| \leq \sum_{(k,l)\in I_N} |a_k||b_l| \\ &\leq \sum_{k=0}^{2N} |a_k| \sum_{l=0}^{2N} |b_l| - \sum_{k=0}^{N} |a_k| \sum_{l=0}^{N} |b_l| \to 0\end{aligned}$$

für $N \to \infty$, also ist

$$\sum_{n=0}^{\infty} c_n = \lim_{N\to\infty} \sum_{n=0}^{2N} c_n = \lim_{N\to\infty} \Big(\sum_{n=0}^{N} a_k \sum_{l=0}^{N} b_l\Big) = \sum_{k=0}^{\infty} a_k \sum_{l=0}^{\infty} b_l. \qquad \square$$

4.5 Die Exponentialreihe

Ein wichtiges Beispiel einer konvergenten Reihe ist die Exponentialreihe, die das exponentielle Wachstum beschreibt.

Satz 4.5.1. *Für jedes $x \in \mathbb{R}$ ist die* Exponentialreihe

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

absolut konvergent. Für je zwei $x, y \in \mathbb{R}$ gilt die Funktionalgleichung

$$\exp(x + y) = \exp(x) \exp(y).$$

Beweis. Ist $x = 0$, so ist nichts zu zeigen. Für gegebenes $x \neq 0$ sei $a_n = \frac{x^n}{n!}$. Dann ist

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

Dies ist eine Nullfolge, so dass das Quotientenkriterium die absolute Konvergenz liefert. Zum Beweis der Funktionalgleichung wendet man Satz 4.4.1 auf die Reihen $\sum_{n=0}^{\infty} a_n = \sum_{n=0}^{\infty} \frac{x^n}{n!} = \exp(x)$ und $\sum_{n=0}^{\infty} b_n = \sum_{n=0}^{\infty} \frac{y^n}{n!} = \exp(y)$ an und erhält die Reihe $\sum_{n=0}^{\infty} c_n$, wobei die Glieder

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

sind. Es folgt

$$\exp(x) \exp(y) = \sum_{n=0}^{\infty} c_n = \exp(x + y).$$ □

Definition 4.5.2. Die Zahl

$$e := \exp(1) = \sum_{n=0}^{\infty} \frac{1}{n!}$$

wird die *Eulersche Zahl* genannt. Ihr Wert ist $e \approx 2,71828\ldots$.

Folgende Tatsachen ergeben sich aus der Definition bzw. der Funktionalgleichung:

Folgerung 4.5.3. *(i)* $\exp(0) = 1$
(ii) $\exp(x) > 0$
(iii) $\exp(-x) = \exp(x)^{-1}$.
(iv) $\forall k \in \mathbb{N}$: $\exp(k) = e^k$.

Beweis. (i) ist klar. (ii) ist klar für $x \geq 0$ und folgt mit (iii) für $x < 0$. (iii) folgt aus $\exp(x)\exp(-x) = \exp(x - x) = \exp(0) = 1$. (iv) folgt schließlich aus $\exp(k) = \exp(1 + \ldots + 1) = \exp(1) \cdots \exp(1) = e^k$. □

Kapitel 5

Grenzwerte und Stetigkeit

Lernziele

- Den Begriff des Grenzwertes exakt definieren
- Grenzwerte über Folgen charakterisieren
- Stetigkeit mit ε-δ-Kriterium definieren
- Stetigkeit über Folgen / Grenzwerte charakterisieren
- Den Zwischenwertsatz wiedergeben und anwenden
- Extremwerte stetiger Funktionen auf (kompakten) Intervallen finden
- Die Definition gleichmäßiger Stetigkeit wiedergeben
- Monotonie
- Existenz von Umkehrfunktionen

5.1 Grenzwerte von Funktionen

Häufig betrachten wir Funktionen oder Ausdrücke, die an einer bestimmten Stelle nicht definiert sind, von denen uns aber genau das Verhalten im diesem Punkt (oder in dessen Nähe) interessiert.

Beispiel.

- Die Funktion $f(x) = 1/x$ ist bei $x = 0$ nicht definiert. Wie verhält sie sich dort, d.h. wenn x beliebig klein wird?
- Wie verhalten sich die Funktionen $\sin(1/x)$, $x \cdot \sin(1/x)$ oder $\sin(x)/x$ in der Nähe von $x = 0$?
- Betrachte den Graph einer Funktion $g(x)$, die auf ganz $\mathbb{R}$ definiert ist. Wir wollen eine Tangente an diesen Graphen an der Stelle x_0 legen, d.h. eine Gerade, die ihn berührt, aber in einer Umgebung dieses Schnittpunktes nur auf *einer* Seite des Graphen liegt. Rechnerisch erhält man die Steigung der Tangente, indem man im Ausdruck

  $$\frac{g(x) - g(x_0)}{x - x_0}$$

  den Punkt x „immer näher“ an x_0 heranwandern lässt. Klarerweise kann aber man nicht $x = x_0$ setzen, da der Nenner dann verschwinden würde.

Im Folgenden werden wir die ungenauen Sprachbilder „immer näher“ und „beliebig klein“ durch exakte Definitionen ersetzen, die ein grundleges Werkzeug der Analysis darstellen und immer wieder zur Anwendung kommen werden. Zuerst ein Begriff dafür, dass man einem Punkt immer näher kommen kann (denn nur dann ergibt die Frage überhaupt Sinn):

Definition 5.1.1. Sei $D \subseteq \mathbb{R}$ eine beliebige Menge. Ein Punkt $x \in \mathbb{R}$ (der nicht in D enthalten sein muss!) heißt *Häufungspunkt* von D, wenn es für jedes $\varepsilon > 0$ ein $y \in D \setminus \{x\}$ gibt mit $|y - x| < \varepsilon$. Ein *isolierter Punkt* von D ist ein $x \in D$, das kein Häufungspunkt von D ist.

Beispiel.

(i) Jedes $x \in [a, b]$ ist Häufpungspunkt von (a, b).
(ii) 0 ist der einzige Häufungspunkt von $\{1/n : n \in \mathbb{N}\}$.

Lemma 5.1.2. Ein Punkt x ist Häufungspunkt einer Menge D genau dann, wenn es eine Folge (x_n) in $D \setminus \{x\}$ gibt, die gegen x konvergiert.

Beweis. Ist x Häufungspunkt von D, dann gibt es für jedes $n \in \mathbb{N}$ ein $x_n \in D \setminus \{x\}$ mit $|x_n - x| < 1/n$, also $x_n \to x$ nach Folgerung 3.1.8. Umgekehrt gibt es für $\varepsilon > 0$ aufgrund der Konvergenz sicher ein $x_n \neq x$ mit $|x_n - x| < \varepsilon$. □

Wir werden hauptsächlich Funktionen auf reellen Intervallen betrachten; jeder Punkt eines Intervalles, das aus mehr als einem Punkt besteht (egal ob dieses offen, halboffen oder abgeschlossen, beschränkt oder unbeschränkt ist) ist klarerweise Häufungspunkt davon, aber auch jeder Randpunkt.

Für den Rest von Abschnitt 5.1 sei $D \subseteq \mathbb{R}$ eine beliebige (nichtleere) Teilmenge und $a \in \mathbb{R}$ ein Häufungspunkt von D.

Definition 5.1.3. Sei $f\colon D \to \mathbb{R}$ eine Funktion. Eine Zahl $L \in \mathbb{R}$ heißt *Grenzwert* von f an der Stelle a, wenn

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall x \in D \setminus \{a\} : (|x - a| < \delta \implies |f(x) - L| < \varepsilon)\,. \tag{5.1}$$

Dafür schreiben wir

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

und sagen auch: *f konvergiert* gegen L, wenn x gegen a strebt.

Bevor wir Grenzwerte von Funktionen studieren, zeigen wir noch dass sie eindeutig sind; damit können wir also von *dem* Grenzwert an einer Stelle sprechen (falls er existiert).

Lemma 5.1.4. Wenn der Grenzwert $\lim_{x \to a} f(x)$ existiert, dann ist er eindeutig bestimmt.

Beweis. Angenommen $\lim_{x\to a} f(x) = L_1$ und $\lim_{x\to a} f(x) = L_2$. Sei $\varepsilon > 0$; dann gibt es für $i = 1, 2$ jeweils ein $\delta_i > 0$ sodass für $x \in D \setminus \{a\}$ mit $|x - a| < \delta_i$ die Ungleichung $|f(x) - L_i| < \varepsilon$ gilt. Für $\delta = \min(\delta_1, \delta_2)$ und $x \in D \setminus \{a\}$ mit $|x - a| < \delta$ gilt somit

$$|L_1 - L_2| \leq |L_1 - f(x)| + |f(x) - L_2| \leq 2\varepsilon.$$

Da $\varepsilon > 0$ beliebig war, folgt $L_1 = L_2$. □

Beispiel.

- Für $\lim_{x\to 1} \frac{x^2-1}{x-1}$ vermuten wir den Wert 2. Wir zeigen es:
$$\left|\frac{x^2-1}{x-1} - 2\right| < \varepsilon \Leftrightarrow |x+1-2| < \varepsilon \Leftrightarrow |x-1| < \varepsilon.$$
Tatsächlich funktionert also $\delta = \varepsilon$.
- Wir behaupten $\lim_{z\to 0} \frac{\exp(z)-1}{z} = 1$. Dazu schreiben wir
$$\frac{\exp(z)-1}{z} - 1 = \frac{z}{2}\left(1 + \frac{z}{3} + \frac{z^2}{3\cdot 4} + \frac{z^3}{3\cdot 4\cdot 5} + \dots\right)$$
und schätzen ab (für $|z| < 1$, denn wir können δ aus (5.1) immer < 1 wählen):
$$\left|\frac{\exp(z)-1}{z} - 1\right| \leq \frac{|z|}{2}\left(\underbrace{1 + |z| + \left|z^2\right| + \dots}_{\text{geom. Reihe}}\right) = \frac{|z|}{2(1-|z|)}.$$
Nun gilt
$$\frac{|z|}{2(1-|z|)} < \varepsilon \iff |z| < 2\varepsilon - 2|z|\varepsilon \iff (1+2\varepsilon)|z| < 2\varepsilon \iff |z| < \frac{2\varepsilon}{1+2\varepsilon}$$
woraus wir $\delta = (2\varepsilon)/(1+2\varepsilon)$ ablesen können.

Den Begriff des Grenzwertes kann man auf die Konvergenz von Zahlenfolgen zurückführen.

Satz 5.1.5. *Für $f\colon D \to \mathbb{R}$ gilt*
$$\lim_{x\to a} f(x) = L \tag{5.2}$$
genau dann, wenn für jede Folge (x_n) in $D \setminus \{a\}$ mit $x_n \to a$ gilt
$$\lim_{n\to\infty} f(x_n) = L. \tag{5.3}$$

Beweis. Angenommen (5.2) gilt und (x_n) ist eine Folge in $D \setminus \{a\}$ mit $x_n \to a$. Sei $\varepsilon > 0$. Aus (5.2) erhalten wir ein $\delta > 0$ sodass für $x \in D \setminus \{a\}$, $|x-a| < \delta$ die Ungleichung $|f(x) - L| < \varepsilon$ folgt. Weiters gibt es aus der Konvergenz von (x_n) ein $n_0 \in \mathbb{N}$ sodass $|x_n - a| < \delta$ für alle $n \geq n_0$ gilt. Daraus folgt
$$\forall n \geq n_0 : |f(x_n) - L| < \varepsilon.$$
Für die andere Implikation beweisen wir die Kontraposition: angenommen, (5.2) gelte nicht, d.h.
$$\exists \varepsilon > 0\ \forall \delta > 0\ \exists x \in D \setminus \{a\} : |x-a| < \delta \text{ und } |f(x) - L| \geq \varepsilon.$$
Setzen wir $\delta = 1, 1/2, 1/3, \ldots$ erhalten wir insbesondere für jedes $n \in \mathbb{N}$ ein $x_n \in D \setminus \{a\}$ mit $|x_n - a| < 1/n$ und $|f(x_n) - L| \geq \varepsilon$. Daraus sehen wir $x_n \to a$, aber $f(x_n) \not\to L$, also gilt (5.3) nicht. ∎

Bemerkung 5.1.6. Manchmal ergibt es auch Sinn, *einseitige* Grenzwerte zu betrachten. Gibt es für a eine Folge (x_n) in D mit $x_n \to a$ sodass $x_n < a$ (bzw. $x_n > a$) ist, dann erhält man den linksseitigen Grenzwert $\lim_{x\to a-} f(x)$ (bzw. den rechtsseitigen Grenzwert $\lim_{x\to a+} f(x)$), indem man in (5.1) die Ungleichung $|x-a| < \delta$ durch $a-\delta < x < a$ (bzw. $a < x < a+\delta$) ersetzt, oder äquivalent dazu, indem man in (5.3) nur Folgen mit $x_n < a$ (bzw. $x_n > a$) betrachtet.[1]

Lemma 5.1.7. $\lim_{x\to a} f(x) = L$ genau dann, wenn $\lim_{x\to a-} f(x) = \lim_{x\to a+} f(x) = L$. Das heißt, der Grenzwert existiert genau dann wenn links- und rechtsseitiger Grenzwert existieren und gleich sind, und in diesem Fall sind alle drei Grenzwerte gleich.

Beweis. Gibt es den Grenzwert, so existieren klarerweise die links- und rechtsseitigen Grenzwerte. Die Umkehrung folgt aus (5.1), denn erhält man von links ein δ_l und von rechts ein δ_r funktioniert $\min(\delta_l, \delta_r)$ für den Grenzwert. □

Beispiel 5.1.8. *Die Funktion*

$$f(x) := \begin{cases} 0 & x \le 0 \\ 1 & x > 1 \end{cases}$$

hat einen Sprung bei $x = 0$*; es ist*

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

Wie bei Folgen definiert man auch uneigentliche Grenzwerte:

Definition 5.1.9. Für $f\colon D \to \mathbb{R}$ setzen wir $\lim_{x\to a} f(x) = \infty$ bzw. $-\infty$, falls

$$\forall R \in \mathbb{R}\ \exists \delta > 0\ \forall x \in D \setminus \{a\} : |x-a| < \delta \Rightarrow f(x) > R \text{ bzw. } < R \tag{5.4}$$

oder, äquivalent dazu,

$$\forall (x_n)_n \text{ in } D \setminus \{a\} \text{ mit } x_n \to a : f(x_n) \to \infty \text{ bzw. } -\infty.$$

Die Aquivalenz ist wie folgt zu sehen: gilt (5.4) und ist eine Folge (x_n) in $D \setminus \{a\}$ sowie $R \in \mathbb{R}$ gegeben, dann gibt es ein δ wie in (5.4) und damit ein n_0 mit $|x_n - a| < \delta$ für $n \ge n_0$, damit aber auch $f(x_n) > R$ für solche n. Gilt (5.4) hingegen nicht, gibt es ein R sodass wir für alle n (mit $\delta = 1/n$) ein $x_n \in D \setminus \{a\}$ finden können sodass $|x_n - a| < 1/n$ gilt, aber $f(x_n) \le R$. Das heißt, $x_n \to a$, aber $f(x_n)$ ist nicht bestimmt divergent gegen ∞.

Definition 5.1.9 kann man genauso auch als links- und rechtsseitige Variante formulieren; die notwendigen Änderungen sind genau dieselben wie in Bemerkung 5.1.6.

Beispiel. $\lim_{x\to 0+} \frac{1}{x} = \infty$, $\lim_{x\to 0-} \frac{1}{x} = -\infty$.

8. 11. 2024

Schließlich können wir auch noch das Verhalten von f für $x \to \pm\infty$ untersuchen. Die Definitionen gehen mittlerweile schon natürlich von der Hand:

[1] Statt $lim_{x\to a-} f(x)$ liest man mitunter auch $\lim_{x\nearrow a} f(x)$ oder $f(a-0)$, ebenso statt $\lim_{x\to a+} f(x)$ mitunter $\lim_{x\searrow a} f(x)$ oder $f(a+0)$.

Definition 5.1.10. Sei $f\colon D \to \mathbb{R}$ eine Funktion. Enthält D eine Menge der Form (c, ∞), so setzt man

$$\lim_{x\to\infty} f(x) = L \;:\Longleftrightarrow\; \forall \varepsilon > 0\, \exists x_0 \in \mathbb{R}\, \forall x \geq x_0 : |f(x) - L| < \varepsilon,$$
$$\lim_{x\to\infty} f(x) = \infty \text{ bzw. } -\infty \;:\Longleftrightarrow\; \forall R \in \mathbb{R}\, \exists x_0 \in \mathbb{R}\, \forall x \geq x_0 : f(x) > R \text{ bzw. } < R.$$

Enthält D eine Menge der Form $(-\infty, d)$, so setzt man

$$\lim_{x\to-\infty} f(x) = L \;:\Longleftrightarrow\; \forall \varepsilon > 0\, \exists x_0 \in \mathbb{R}\, \forall x \leq x_0 : |f(x) - L| < \varepsilon,$$
$$\lim_{x\to-\infty} f(x) = \infty \text{ bzw. } -\infty \;:\Longleftrightarrow\; \forall R \in \mathbb{R}\, \exists x_0 \in \mathbb{R}\, \forall x \leq x_0 : f(x) > R \text{ bzw. } < R.$$

Auch hier gelten die entsprechenden Charakterisierungen über Folgen.

Lemma 5.1.11. Sei $f\colon D \to \mathbb{R}$ eine Funktion. Ist $\lim_{x\to a} f(x) = c > 0$ bzw. < 0, dann gibt es $\delta > 0$ mit $f(x) > c/2$ bzw. $< c/2$ für alle $x \in D \setminus \{a\}$ mit $|x - a| < \delta$.

Beweis. Sei $\lim_{x\to a} f(x) = c > 0$, dann gibt es für $\varepsilon = c$ ein $\delta > 0$ sodass für $x \in D \setminus \{a\}$ und $|x - a| < \delta$ die Ungleichung $|f(x) - c| < c/2$ gilt, und damit auch $f(x) > c/2$. □

Satz 5.1.12. *Existieren* $\lim_{x\to a} f(x)$ *und* $\lim_{x\to a} g(x)$*, so existieren die Grenzwerte*

(i) $\lim_{x\to a}(\alpha f(x) + \beta g(x)) = \alpha \lim_{x\to a} f(x) + \beta \lim_{x\to a} g(x)$ $(\alpha, \beta \in \mathbb{R})$,

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

(iii) $\lim_{x\to a} \frac{f(x)}{g(x)} = \frac{\lim_{x\to a} f(x)}{\lim_{x\to a} g(x)}$, *falls* $\lim_{x\to a} g(x) \neq 0$.

Hierbei kann a für eine reelle Zahl sowie for ∞ *oder* $-\infty$ *stehen. Dieselben Aussagen gelten auch für links- und rechtsseitige Grenzwerte.*

Beweis. Die Aussagen folgen direkt aus den entsprechenden Aussagen für Folgen, also Satz 3.1.9 und Satz 3.1.10. Für (iii) beachte, dass der Ausdruck $f(x)/g(x)$ für $|x - a|$ klein genug wegen Lemma 5.1.11 wohldefiniert ist. □

Achtung: für uneigentliche Grenzwerte gilt dieser Satz so nicht.

5.2 Stetigkeit

Salopp ausgedrückt bedeutet die Stetigkeit einer reellen Funktion, dass sie keine „Sprünge“ hat.

Für den Grenzwert einer Funktion an einer Stelle ist ihr Wert dort unerheblich; wichtig ist nur, dass sich die Funktion von einem festen Wert beliebig wenig unterscheidet, wenn man nur nahe genug an dieser Stelle ist. Für die Stetigkeit nimmt jedoch der Funktionswert an dieser Stelle die Rolle dieses festen Wertes ein.

Definition 5.2.1. Eine Funktion $f\colon D \to \mathbb{R}$ heißt *stetig im Punkt* $a \in D$, wenn

$$\forall \varepsilon > 0\, \exists \delta > 0\, \forall x \in D : |x - a| < \delta \implies |f(x) - f(a)| < \varepsilon \tag{5.5}$$

oder wenn

$$\forall (x_n) \text{ in } D \text{ mit } x_n \to a \text{ gilt } \lim_{n\to\infty} f(x_n) = f(a) \tag{5.6}$$

oder, falls a Häufungspunkt von D ist, wenn

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

Eine Funktion heißt (punktweise) stetig, wenn sie in jedem Punkt ihres Definitionsbereichs stetig ist.

Beachte hierbei, dass Stetigkeit in jedem Punkt des Definitionsbereiches Sinn ergibt, im Gegensatz zum Grenzwert, der nur in Häufungspunkten definiert ist. Tatsächlich ist eine Funktion in jedem isolierten Punkt ihres Definitionsbereiches stetig, wie man einfach an der Definition erkennt.

Beispiel.

(i) Die Abbildung $g\colon \mathbb{R} \to \mathbb{R}, g(x) := \begin{cases} 1, & x \in \mathbb{Q} \\ 0, & x \notin \mathbb{Q} \end{cases}$ ist in keinem Punkt stetig. Sei $x \in \mathbb{Q}$, $x_n := x + \frac{\sqrt{2}}{n} \to x$. Damit gilt $f(x_n) = 0 \nrightarrow 1 = f(x)$. Sei $x \notin \mathbb{Q}, x_n \in \mathbb{Q}$ mit $x_n \to x$. Dann gilt $f(x_n) = 1 \nrightarrow 0 = f(x)$.

(ii) $\mathrm{Id}\colon \mathbb{R} \to \mathbb{R}$ ist stetig.

(iii) $x \mapsto |x|$ ist stetig, denn $x_n \to x \implies |x_n| \to |x|$.

(iv) $x \mapsto \frac{1}{x}$ ist stetig auf $\mathbb{R} \setminus \{0\}$, denn

$$x_n \to x \neq 0 \implies \frac{1}{x_n} - \frac{1}{x} = \frac{x - x_n}{x_n x} \to 0.$$

(v) $x \mapsto \exp(x)$ ist stetig, denn

$$\begin{aligned}\left|\exp(x) - \exp(a)\right| = \left|\exp(a)(\exp(x-a) - 1)\right| &= \left|\exp(a)(x-a)\frac{\exp(x-a)-1}{x-a}\right| \\ &\leq \exp(a)\frac{|x-a|}{1-|x-a|} \to 0,\end{aligned}$$

wobei wir für $|x-a| < 1$ die Ungleichung

$$\frac{\exp(x-a)-1}{x-a} = \sum_{n=1}^{\infty} \frac{(x-a)^{n-1}}{n!} \leq \sum_{n=0}^{\infty} |x-a|^n = \frac{1}{1-|x-a|}$$

verwendet haben.

Aus den Rechenregeln für Folgen sieht man wieder sogleich:

Satz 5.2.2. *Sind $f, g\colon D \to \mathbb{R}$ stetig in a, dann auch*

(i) αf für $\alpha \in \mathbb{R}$,

(ii) $f + g$,

(iii) fg.

(iv) f/g, falls $g(a) \neq 0$.

Hierbei ist natürlich $(\alpha f)(x) = \alpha f(x)$, $(f+g)(x) = f(x) + g(x)$, $(fg)(x) = f(x)g(x)$ und $(f/g)(x) = f(x)/g(x)$ für $x \in D$.

Beispiel.

(i) Jedes Polynom $p(x) = a_n x^n + \ldots + a_1 x + a_0$ ist stetig auf $\mathbb{R}$.
(ii) Ist f stetig dann auch $|f|: x \mapsto |f(x)|$.
(iii) Mit f, g sind auch $x \mapsto f(x) \vee g(x) := \max(f(x), g(x))$ und $x \mapsto f \wedge g(x) := \min(f(x), g(x))$ stetig, denn $f \vee g = \frac{1}{2}(f + g + |f - g|)$ und $f \wedge g = \frac{1}{2}(f + g - |f - g|)$.

Lemma 5.2.3. Sie Funktionen $f: D \to \mathbb{R}$ und $g: E \to \mathbb{R}$ seien gegeben mit $f(D) \subseteq E$. Ist f stetig in a und g stetig in $f(a)$, dann ist $g \circ f: D \to \mathbb{R}$ stetig in a.

Beweis. Sei $x_n \to a$ eine konvergente Folge. Da f stetig ist, konvergiert $f(x_n)$ gegen $f(a) \in E$. Da g in $f(a)$ stetig ist, konvergiert die Folge $g(f(x_n)) = (g \circ f)(x_n)$ gegen $g(f(a))$, und damit folgt die Behauptung. □

Ganz analog wie in Definition 5.2.1 definiert man die Stetigkeit für komplexe Funktionen, wobei man nur den Betrag bzw. die Folgenkonvergenz richtig lesen muss. Damit sind z.B. auch die Komplexkonjugation sowie die Bildung von Real- und Imaginärteil stetig, d.h. die Funktionen

$$z = a + bi \mapsto \bar{z} = a - bi, \quad \mathrm{Re}: a + bi \mapsto a, \qquad \mathrm{Im}: a + bi \mapsto b.$$

Elementare Konsequenzen der Stetigkeit

Satz 5.2.4 (Zwischenwertsatz). *Sei f stetig auf einem beschränkten Intervall $I = [a, b]$ mit $a < b$. Dann nimmt f jeden Wert zwischen $f(a)$ und $f(b)$ an. Insbesondere gilt: ist $f(a) < 0$ und $f(b) > 0$, dann hat f mindestens eine Nullstelle in (a, b).*

Beweis. Wir zeigen zuerst den Spezialfall; sei dazu

$$M := \{x \in [a, b] : f(x) < 0\} \quad \text{und} \quad s := \sup M.$$

Aus Lemma Lemma 3.2.5 erhalten wir die Existenz von Folgen (a_n) in M und (b_n) in $\mathbb{R} \setminus M$ mit $a_n \to s$ und $b_n \to s$. Aus der Stetigkeit von f erhalten wir für die Randpunkte außerdem ein $\delta > 0$ mit

$$f(x) < 0 \text{ für } x \in [a, a + \delta), \qquad f(x) > 0 \text{ für } x \in (b - \delta, b].$$

Daraus folgt $s \in [a + \delta, b - \delta] \subseteq (a, b)$, wir können also auch annehmen, dass (b_n) ganz in $[a, b]$ liegt. Da $f(a_n) \leq 0$ und $f(b_n) \geq 0$ ist erhalten wir aus der Stetigkeit

$$\left.\begin{aligned} \lim_{n\to\infty} f(a_n) = f(s) \leq 0, \\ \lim_{n\to\infty} f(b_n) = f(s) \geq 0 \end{aligned}\right\} \implies f(s) = 0.$$

Nun zur allgemeinen Behauptung. Wir schließen den trivialen Fall $f(a) = f(b)$ aus und nehmen $f(a) < f(b)$ an. Ist $\eta \in (f(a), f(b))$, dann gelten für $g(x) := f(x) - \eta$ die Ungleichungen $g(a) < 0$ und $g(b) > 0$, also gibt es nach dem ersten Teil ein $c \in (a, b)$ mit

$g(c) = 0$, also $f(c) = \eta$. Im Fall $f(a) > f(b)$ wendet man das soeben gezeigte auf die Funktion $-f$ an. □

14. 11. 2024

Folgerung 5.2.5. *Ist $f\colon I \to \mathbb{R}$ eine stetige Funktion auf einem Intervall, dann ist $f(I)$ auch ein Intervall.*

Definition 5.2.6. Ein Intervall heißt *kompakt*, falls es beschränkt und abgeschlossen ist, also von der Form $[a, b]$ mit $a, b \in \mathbb{R}$.

Satz 5.2.7. *Eine stetige Funktion auf einem kompaktem Intervall $[a, b]$ nimmt ihr Maximum und Minimum an, d.h.*

$$\exists m, M \in [a, b] : f(m) = \min_{x\in[a,b]} f(x) \text{ und } f(M) = \max_{x\in[a,b]} f(x).$$

Beweis. Wir zeigen die Existenz des Maximums; das Minimum erhält man als Maximum von $-f$. Wir setzen $S := \sup\{f(x) : x \in [a, b]\} \in \mathbb{R} \cup \{\infty\}$. Da es eine Folge in $f([a, b])$ gibt, die gegen S konvergiert, gibt es eine Folge (x_n) in $[a, b]$ für welche $f(x_n)$ gegen S konvergiert. Nach Satz 3.4.3 hat (x_n) eine Konvergente Teilfolge $(x_{n_k})_k$ mit $x_{n_k} \to x \in [a, b]$. Aufgrund der Stetigkeit von f gilt $S = \lim_{n\to\infty} f(x_n) = \lim_{k\to\infty} f(x_{n_k}) = f(x)$, also nimmt f sein Maximum in x an. □

Folgerung 5.2.8. *Eine stetige Funktion bildet kompakte Intervalle auf kompakte Intervalle ab.*

Gleichmäßige Stetigkeit

In der ε-δ-Definition der punktweisen Stetigkeit gibt es für jedes (feste) $a \in D$ und für jedes $\varepsilon > 0$ ein δ mit $|f(x) - f(a)|$ für alle $x \in D$ mit $|x - a| < \delta$. Dieses δ hängt also im allgemeinen nicht nur von ε, sondern auch vom Punkt a ab; je mehr sich f dort ändert, desto kleiner muss δ sein.

Kann man das δ unabhängig vom Punkt wählen, ergibt das einen stärkeren Stetigkeitsbegriff:

Definition 5.2.9. Eine Funktion $f\colon D \to \mathbb{R}$ heißt *gleichmäßig stetig*, wenn

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall x, y \in D\colon |x - y| < \delta \implies |f(x) - f(y)| < \varepsilon. \tag{5.7}$$

Gleichmäßig stetige Funktionen sind offenbar stetig, die Umkehrung gilt jedoch nicht:

Beispiel 5.2.10. *(i) Die Funktion $f(x) = 1/x$ auf $(0, 1)$ ist nicht gleichmäßig stetig; wäre sie es, dann wäre in (5.7) mit $y = \min(\delta, 1)/2$ und beliebigem $x \in (0, y)$ sicher $|x - y| < \delta$, aber $|1/x - 1/y|$ wird für kleine x beliebig groß, also auch*

$$\frac{1}{x} - \frac{1}{y} \geq \varepsilon \iff \frac{1}{x} \geq \varepsilon + \frac{1}{y} \iff x \leq \frac{y}{\varepsilon y + 1}.$$

(ii) *Die Wurzelfunktion $x \mapsto \sqrt{x}$ ist gleichmäßig stetig auf $[0, \infty)$. Tatsächlich gilt wegen $\sqrt{|x-y|} \leq \sqrt{x} + \sqrt{y}$*

$$\left|\sqrt{x} - \sqrt{y}\right| = \frac{|x-y|}{\sqrt{x} + \sqrt{y}} \leq \frac{|x-y|}{\sqrt{|x-y|}} = \sqrt{|x-y|},$$

also kann man in (5.7) für $\varepsilon > 0$ $\delta = \varepsilon^2$ wählen. Ein analoger Beweis funktioniert für die n-te Wurzel.

Satz 5.2.11 (Satz von Heine). *Sei $I \subseteq \mathbb{R}$ ein kompaktes Intervall und $f\colon I \to \mathbb{R}$ stetig. Dann ist f gleichmäßig stetig.*

Beweis. Wäre f nicht gleichmäßig stetig gäbe es ein $\varepsilon > 0$ sodass

$$\forall n \in \mathbb{N}\ \exists x_n, y_n \in [a,b] \text{ mit } |x_n - y_n| < \frac{1}{n} \text{ aber } |f(x_n) - f(y_n)| \geq \varepsilon.$$

Nach Satz 3.4.3 hat (x_n) eine konvergente Teilfolge $(x_{n_k})_k$ mit Grenzwert $p \in [a,b]$; ebenso konvergiert wegen $\left|y_{n_k} - p\right| \leq \left|y_{n_k} - x_{n_k}\right| + \left|x_{n_k} - p\right| \to 0$ auch y_{n_k} gegen p. Aufgrund der Stetigkeit gilt

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

was einen Widerspruch zu $\left|f(x_{n_k}) - f(y_{n_k})\right| \geq \varepsilon$ für alle k ergibt. □

Definition 5.2.12. Eine Funktion $f\colon D \to \mathbb{R}$ heißt

(i) *streng monoton wachsend*, falls $x < y \Rightarrow f(x) < f(y)$,
(ii) *monoton wachsend*, falls $x < y \Rightarrow f(x) \leq f(y)$,
(iii) *monoton fallend*, falls $x < y \Rightarrow f(x) \geq f(y)$,
(iv) *streng monoton fallend*, falls $x < y \Rightarrow f(x) > f(y)$

für alle $x, y \in D$ gilt.

Beispiel. Die n-te Potenz $x \mapsto x^n$ ist (für $n \in \mathbb{N}$) nach (O4) streng monoton wachsend auf $[0, \infty)$.

Satz 5.2.13 (Stetigkeit der Umkehrfunktion). *Sei I ein Intervall und $f\colon I \to \mathbb{R}$ eine stetige, streng monoton wachsende (bzw. fallende) Funktion. Dann ist $J := f(I)$ ein Intervall, $f\colon I \to J$ bijektiv und die Umkehrfunktion $f^{-1}\colon J \to I$ ist stetig und streng monoton wachsend (bzw. fallend).*

Beweis. Wir beschränken uns auf den Fall einer streng monoton wachsenden Funktion f, da wir sonst f durch $-f$ ersetzen können um das Ergebnis zu erhalten.

J ist ein Intervall nach Folgerung 5.2.5, f injektiv aufgrund der Monotonie, also $f\colon I \to J$ bijektiv. Ist nun $x < y$ in J muss $f^{-1}(x) < f^{-1}(y)$ sein, denn sonst wäre $x = f(f^{-1}(x)) \geq f(f^{-1}(y)) = y$, d.h. f^{-1} ist streng monoton wachsend.

Wir zeigen die Stetigkeit von f^{-1} für innere Punkte $y_0 \in J$; das Argument für Randpunkte ist analog mit offensichtlichen Änderungen. Sei also $x_0 = f^{-1}(y_0)$ und $\varepsilon > 0$. Wähle $a, b \in I$ mit $x_0 - \varepsilon < a < x_0 < b < x_0 + \varepsilon$ und setze $\alpha = f(a)$, $\beta = f(b)$. Für $\delta < \min(|y_0 - f(b)|, |y_0 - f(a)|)$ und $y \in J$ gilt dann

$$\begin{aligned} |y - y_0| < \delta &\implies f(a) < y < f(b) \implies a = f^{-1}(\alpha) < f^{-1}(y) < f^{-1}(\beta) = b \\ &\implies \left| f^{-1}(y) - f^{-1}(y_0) \right| < \varepsilon, \end{aligned}$$

also ist f^{-1} stetig in y_0. □

Folgerung 5.2.14. *Sie $n \in \mathbb{N}$. Die Umkehrfunktion der Abbildung $[0, \infty) \to [0, \infty)$, $x \mapsto x^n$, also die n-te Wurzel $x \mapsto \sqrt[n]{x}$, ist stetig und streng monoton wachsend.*

21. 11. 2024

Kapitel 6

Elementare Funktionen

Lernziele

- ► Die AGM-Ungleichung kennen und anwenden können
- ► Die Exponentialfunktion als Grenzwert einer Folge darstellen
- ► Wichtigste Eigenschaften der Exponentialfunktion wiedergeben
- ► Den Logarithmus als Umkehrfunktion der Exponentialfunktion einführen
- ► Die allgemeine Potenz samt Eigenschaften erklären
- ► Trigonometrische Funktionen über die Exponentialfunktion definieren
- ► Eigenschaften der trigonometrischen Funktionen herleiten
- ► Polarkoordinaten anwenden
- ► Umkehrfunktionen der trigonometrischen Funktionen bestimmen
- ► Hyperbelfunktionen
- ► Polynomdivision durchführen
- ► Polynome in Linearfaktoren zerlegen / Linearfaktoren abspalten
- ► reelle und komplexe Partialbruchzerlegungen durchführen.

6.1 Die Exponentialfunktion

Beispiel.

(i) Wird ein Anfangskapitel K jährlich mit einem Zinssatz von p Prozent verzinst, beträgt das Kapital nach einem Jahr $K(1+\alpha)$, wobei $\alpha = p/100$. Wird unterjährig verzinst, d.h. in kürzeren als jährlichen Intervallen, beträgt das Kapital nach einem Jahr

  (i) $K(1+\frac{\alpha}{2})^2$ bei halbjähriger Verzinsung,
  (ii) $K(1+\frac{\alpha}{12})^{12}$ bei monatlicher Verzinsung,
  (iii) $K(1+\frac{\alpha}{365})^{365}$ bei täglicher Verzinsung (außer in Schaltjahren).

  Was passiert, wenn man das Verzinsungsintervall immer kürzer macht?

(ii) Bei einem radioaktiven Stoff ist die Zahl der radioaktiven Zerfälle pro Zeiteinheit proportional zur vorhandenen Anzahl N von Atomkernen:

$$\Delta N \approx -\lambda N,$$

wobei der Proportionalitätsfaktor λ die *Zerfallskonstante* genannt wird. Teilt man den Zeitraum von 0 bis T in n Teile der Länge $x = T/n$, so sind zur Zeit T noch

$$N(T) = N(0)\left(1 - \lambda\frac{T}{n}\right)^n$$

Atomkerne übrig. Was passiert, wenn n beliebig groß wird?

(iii) Ist die Wachstumsrate einer Population proportional zu ihrer Größe $N(T)$, wird das durch

$$N(T) = N_0\left(1 + \alpha\frac{T}{n}\right)^n$$

beschrieben. Auch hier wieder die Frage: was passiert für $n \to \infty$?

Wir werden nun die Folge $(1 + x/n)^n$ mit $x \in \mathbb{R}$ für $n \to \infty$ studieren. Dafür brauchen wir die *Ungleichung zwischen arithmetischem und geometrischem Mittel*:

Lemma 6.1.1. Für $x_1, \ldots, x_n \geq 0$ (mit $n \in \mathbb{N}$) gilt

$$\sqrt[n]{x_1 x_2 \cdots x_n} \leq \frac{x_1 + \ldots + x_n}{n} \tag{6.1}$$

mit Gleichheit genau für $x_1 = \ldots = x_n$.

Ein Produkt nichtnegativer Zahlen mit vorgegebener Summe ist also dann am größten, wenn alle Faktoren gleich sind.

Beweis. Wir zeigen die Behauptung durch vollständige Induktion nach $n \in \mathbb{N}$. Bezeichne G die linke und A die rechte Seite von (6.1).

Für $n = 1$ ist die Aussage trivial. Für den Schritt von $n - 1$ auf n für $n \geq 2$ unterscheiden wir 3 Fälle:

1. Fall: ist ein $x_i = 0$, dann ist $G = 0$ und damit $G \leq A$; $G = A$ impliziert $A = 0$ und $x_j = 0$ für alle j.

2. Fall: sind alle $x_j > 0$ und alle gleich, dann ist $x_j = G = A$ für alle j.

3. Fall: sind alle $x_j > 0$ und $x_i \neq A$ für ein i, dann gilt

$$\begin{aligned} x_i < A &\implies \exists j : x_j > A \\ x_j > A &\implies \exists j : x_j < A, \end{aligned}$$

da nicht alle x_j gleichzeitig größer bzw. kleiner als ihr arithmetisches Mittel sein können. In jedem Fall gibt es also i, j mit $(A - x_i)(A - x_j) < 0$, also auch $x_i x_j < A(x_i + x_j - A)$. Ersetzen wir in $x_1 \cdots x_n$ die Faktoren x_i bzw. x_j durch A bzw. $(x_i + x_j - A)$ und wenden die Induktionsvoraussetzung dann auf alle Faktoren außer A an, erhalten wir

$$\begin{aligned} x_1 \cdots x_i \cdots x_j \cdots x_n &< x_1 \cdots A \cdots (x_i + x_j - A) \cdots x_n \\ &\leq A\left(\frac{x_1 + \ldots + x_n - A}{n-1}\right)^{n-1} = A\left(\frac{nA}{n-1} - \frac{A}{n-1}\right)^{n-1} = A^n. \end{aligned}$$

□

Satz 6.1.2. *Für $x \in \mathbb{R}$ gilt*

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

Beweis. Sei zuerst $x \geq 0$ und

$$t_n = (1+\frac{x}{n})^n, \quad s_n = \sum_{k=0}^{n}\frac{x^k}{k!}.$$

Nach dem Binomischen Lehrsatz ist

$$\begin{aligned} t_n &= \sum_{k=0}^{n}\binom{n}{k}\frac{x^k}{n^k} = \sum_{k=0}^{n}\frac{x^k}{k!}\frac{n(n-1)\cdots(n-k+1)}{n^k} \\ &= \sum_{k=0}^{n}\frac{x^k}{k!}1\cdot(1-\frac{1}{n})\cdots(1-\frac{k-1}{n}) \leq s_n \leq \exp(x), \end{aligned}$$

also ist (t_n) beschränkt. Die strikte Monotonie von (t_n), also

$$\left(1+\frac{x}{n}\right)^n < \left(1+\frac{x}{n+1}\right)^{n+1} \qquad x \geq -n,\ x \neq 0 \tag{6.2}$$

folgt mit Lemma 6.1.1 indem man von

$$\sqrt[n+1]{\left(1+\frac{x}{n}\right)^n\cdot 1} < \frac{n(1+\frac{x}{n})+1}{n+1} = 1+\frac{x}{n+1}$$

die $(n+1)$-te Potenz nimmt. Daraus erhalten wir die Konvergenz von (t_n) und $\lim_n t_n \leq \exp(x)$.

Andererseits gilt für $2 \leq m \leq n$

$$\sum_{k=0}^{m}\frac{x^k}{k!}1\cdot(1-\frac{1}{n})\cdots(1-\frac{k-1}{n}) \leq t_n.$$

Für m fix und $n \to \infty$ erhalten wir

$$s_m \leq \lim_n t_n \leq \exp(x)$$

und daraus für $m \to \infty$ $\lim t_n = \exp(x)$. Um die Behauptung für $x < 0$ zu zeigen, verwenden wir $y = -x > 0$ und

$$\left(1+\frac{x}{n}\right)^n = (1-\frac{y}{n})^n = \frac{(1-\frac{y^2}{n^2})^n}{(1+\frac{y}{n})^n} \to \frac{1}{\exp(x)} = \exp(-x) = \exp(y).$$

Hierbei folgt die Konvergenz des Zählers gegen 1 aus der Abschätzung

$$1 \geq (1-\frac{x^2}{n^2})^n \geq 1-\frac{x^2}{n^2}\cdot n = 1-\frac{x^2}{n} \to 1$$

für n groß genug, die wir aus der Bernoulli'schen Ungleichung (Satz 1.4.2) erhalten. □

Lemma 6.1.3.

(i) Die Exponentialfunktion ist stetig und streng monoton wachsend.
(ii) Es gilt $\exp(x) \geq 1 + x$ für alle $x \in \mathbb{R}$, mit Gleichheit genau für $x = 0$.
(iii) Es gilt $\lim_{x\to-\infty} \exp(x) = 0$ und $\lim_{x\to+\infty} \exp(x) = +\infty$, exp bildet also $(-\infty, \infty)$ bijektiv auf $(0, \infty)$ ab.

Beweis. (i) Die Stetigkeit haben wir weiter oben schon gesehen. Für die Monotonie betrachten wir drei Fälle:

Fall $0 \leq x < y$: aus $\frac{x^n}{n!} < \frac{y^n}{n!}$ für $n \geq 1$ folgt $\exp(x) = \sum_{n=1}^{\infty} \frac{x^n}{n!} < \sum_{n=0}^{\infty} \frac{y^n}{n!} = \exp(y)$.

Fall $x < 0 \leq y$: $\exp(x) = (\underbrace{\exp(-x)}_{>\exp(0)=1})^{-1} < 1 \leq \exp(y)$.

Fall $x < y < 0$: $0 < -y < -x$ ergibt $\exp(-y) < \exp(-x)$, also $\exp(x) < \exp(y)$.

(ii) folgt mit der Ungleichung von Bernoulli: $(1 + x/n)^n \geq 1 + n \cdot x/n = 1 + x$ sobald n groß genug ist.

(iii) ist klar aus (ii), denn für eine Folge (x_n) in $\mathbb{R}$ gilt

$$\begin{aligned} x_n \to \infty &\implies \exp(x_n) \geq 1 + x_n \to \infty \\ x_n \to -\infty &\implies -x_n \to \infty \implies \exp(x_n) = \frac{1}{\exp(-x_n)} \to 0. \end{aligned}$$ □

Folgende Tatsache motiviert für exp die Bezeichnung *Exponentialfunktion*:

Satz 6.1.4. *Für $q \in \mathbb{Q}$ gilt* $\exp(q) = e^q$.

Beweis. Für $m \in \mathbb{N}_0$ haben wir

$$\exp(m) = e^m, \qquad \exp(-m) = \frac{1}{\exp(m)} = \frac{1}{e^m} = e^{-m}.$$

Für $n \in \mathbb{N}$ haben wir

$$e = \exp(1) = \exp\left(n \cdot \frac{1}{n}\right) = \exp\left(\underbrace{\frac{1}{n} + \ldots + \frac{1}{n}}_{n \text{ Summanden}}\right) = \exp\left(\frac{1}{n}\right)^n$$

woraus $e^{1/n} = \exp(1/n)$ folgt. Damit gilt

$$\begin{aligned} \exp\left(\frac{m}{n}\right) &= \exp\left(\underbrace{\frac{1}{n} + \ldots + \frac{1}{n}}_{m \text{ Summanden}}\right) = \exp\left(\frac{1}{n}\right)^m = (e^{1/n})^m = e^{m/n}, \\ \exp\left(\frac{-m}{n}\right) &= \exp\left(\frac{m}{n}\right)^{-1} = (e^{m/n})^{-1} = e^{-m/n}. \end{aligned}$$ □

Definition 6.1.5. Wir setzen

$$e^x := \exp(x) \quad \forall x \in \mathbb{R}.$$

Gewissermaßen ist das die einzige „sinnvolle“ Definition von e^x für $x \in \mathbb{R}$, denn wir kennen e^x ja schon für $x \in \mathbb{Q}$; fordert man die Stetigkeit von e^x, folgt für $x \in \mathbb{R}$ und (x_n) in $\mathbb{Q}$ mit $x_n \to x$ sofort

$$e^x = e^{\lim x_n} = \lim e^{x_n} = \lim \exp(x_n) = \exp(x).$$

Der Logarithmus

Satz 6.1.6. *Die Umkehrfunktion von* $\exp\colon (-\infty, \infty) \to (0, \infty)$*, genannt* Logarithmus*,*

$$\log\colon (0, \infty) \to (-\infty, \infty)$$

ist stetig, streng monoton wachsend und erfüllt die Funktionalgleichung

$$\log(ab) = \log(a) + \log(b)$$

für alle $a, b \in (0, \infty)$*.*

Beweis. Die Stetigkeit und Monotonie des Logarithmus folgt aus Satz 5.2.13. Die Funktionalgleichung erhält man aus

$$\log(ab) = \log(e^{\log a} e^{\log b}) = \log(e^{\log a + \log b}) = \log a + \log b. \qquad \square$$

6.2 Die allgemeine Potenz

Für $a > 0$ und $x \in \mathbb{Q}$ gilt $a^x = e^{x \log a}$, denn mit $x = m/n$, $m \in \mathbb{N}_0$ und $n \in \mathbb{N}$ ist

$$(e^{x \log a})^n = e^{m \log a} = e^{log a^m} = a^m,$$

also ist $e^{x \log a} = \sqrt[n]{a^m}$, da die n-te Wurzel eindeutig ist; ebenso ist

$$a^{-x} = \frac{1}{a^x} = \frac{1}{e^{x \log a}} = e^{-x \log a}.$$

Definition 6.2.1. Für $a > 0$ sei $a^x := e^{x \log a}$, $x \in \mathbb{R}$.

Satz 6.2.2. *Sei* $a > 0$*.*

(i) $x \mapsto a^x$ *ist stetig auf* $\mathbb{R}$, $a^{x+y} = a^x a^y$, $a^{-x} = \frac{1}{a^x}$.

(ii) $\forall x \in \mathbb{R}$: $\log(a^x) = x \log a$.

(iii) $(ab)^x = a^x b^x$ *für* $a, b > 0$ *und* $x \in \mathbb{R}$.

(iv) $(a^x)^y = a^{xy}$ *für* $x, y \in \mathbb{R}$

(v) $a^{\frac{m}{n}} = \sqrt[n]{a^m}$ *für* $m \in \mathbb{Z}$ *und* $n \in \mathbb{N}$.

Beweis. (i) ist klar aus den Eigenschaften der Exponentialfunktion, (ii) aus der Definition von a^x. (iii) sieht man aus

$$(ab)^x = e^{x\log(ab)} = e^{x(\log a+\log b)} = e^{x\log a}e^{x\log b} = a^x b^x,$$

(iv) aus

$$(a^x)^y = e^{y\log(a^x)} = e^{xy\log a} = a^{xy}$$

und (v) haben wir oben schon gezeigt. □

Satz 6.2.3.
(i) $\lim_{x\to\infty} \frac{e^x}{x^k} = \infty$ *für alle* $k \in \mathbb{N}$*, d.h.* e^x *wächst schneller als jede Potenz von* x*.*
(ii) $\lim_{x\to\infty} \frac{\log x}{\sqrt[k]{x}} = 0$ *für alle* $k \in \mathbb{N}$*, d.h.* $\log x$ *wächst langsamer als jede Wurzel von* x*.*

Beweis. (i) folgt für $x > 0$ aus

$$e^x > \frac{x^{k+1}}{(k+1)!} \implies \frac{e^x}{x^k} > \frac{x}{(k+1)!} \to \infty \text{ für } x \to \infty.$$

(ii) sieht man indem man $x = e^{ky}$ substituiert, denn wegen $x \to \infty$ genau dann wenn $y \to \infty$ ist dann

$$\lim_{x\to\infty} \frac{\log x}{\sqrt[k]{x}} = \lim_{y\to\infty} \frac{\log e^{ky}}{\sqrt[k]{e^{ky}}} = \lim_{y\to\infty} k\frac{y}{e^y} \to 0.$$ □

Bemerkung. Für $x > 0$ ist $\lim_{a\to 0+} a^x = \lim_{a\to\infty} \left(\frac{1}{a}\right)^x = \lim_{a\to\infty} \frac{1}{a^x} = 0$, also setzt man $0^x := 0$ für $x > 0$ (als stetige Erweiterung der Funktion $a \to a^x$ in den Nullpunkt).

Bemerkung. Mit $\log_a$ bezeichnet man den Logarithmus zur Basis $a > 0$, d.h. die Umkehrfunktion von $x \mapsto a^x$. Es ergibt sich sofort

$$\log_a y = \frac{\log y}{\log a}.$$

Statt log sieht man auch oft die Schreibweise ln (*Logarithmus naturalis*).

6.3 Trigonometrische Funktionen

In Kapitel 3 haben wir erwähnt, dass die meisten Begriffe Resultate auch für komplexe Folgen gelten. Ebenso überträgt sich viele auf Reihen von komplexe Zahlen: eine Reihe $\sum_{n=1}^{\infty} z_n$ mit $z_n \in \mathbb{C}$ heißt konvergent, wenn die Folge der Partialsummen $s_n = \sum_{k=1}^{n} z_k$ konvergiert, und absolut konvergent wenn $\sum_{n=1}^{\infty} |z_n| < \infty$ ist. Aus der absoluten Konvergenz folgt die Konvergenz, Majoranten-, Quotienten- und Wurzelkriterium funktionieren genauso, ebenso wie das Cauchyprodukt. Daraus sieht man:

Satz 6.3.1. *Für $z \in \mathbb{C}$ konvergiert*

$$\exp(z) := \sum_{n=0}^{\infty} \frac{z^n}{n!}$$

absolut, es gilt

$$\exp(z+w) = \exp(z) \cdot \exp(w) \qquad \forall z, w \in \mathbb{C}$$

und $\overline{\exp(z)} = \exp(\overline{z})$.

Satz 6.3.2. $\exp\colon \mathbb{C} \to \mathbb{C}$ *ist stetig.*

Damit können wir *komplexe Exponentialfunktion* definieren:

Definition 6.3.3.

$$e^z := \exp(z) \qquad \text{für } z \in \mathbb{C}.$$

Lemma 6.3.4. Für alle $z \in \mathbb{C}$ gilt $|e^z| = e^{\operatorname{Re} z}$, insbesondere ist $\left|e^{ix}\right| = 1$ für $x \in \mathbb{R}$.

Beweis. Die Behauptung folgt durch Wurzelziehen aus

$$|e^z|^2 = e^z\overline{e^z} = e^z e^{\overline{z}} = e^{z+\overline{z}} = e^{2\operatorname{Re} z} = \left(e^{\operatorname{Re} z}\right)^2,$$

der Spezialfall ist klar. □

Definition 6.3.5. Für $x \in \mathbb{R}$ setzen wir

$$\cos x := \operatorname{Re}\left(e^{ix}\right), \qquad \sin x := \operatorname{Im}\left(e^{ix}\right).$$

Es gilt die *Eulersche Formel*

$$e^{ix} = \cos x + i \sin x \qquad \forall x \in \mathbb{R}.$$

Satz 6.3.6.

(i) $\cos x = \dfrac{e^{ix} + e^{-ix}}{2}$, $\sin x = \dfrac{e^{ix} - e^{-ix}}{2i}$.

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

(iii) $\cos^2 x + \sin^2 x = 1$.

Beweis. (i) ist klar aus der Definition von Real- und Imaginärteil, (ii) und (iii) aus der Definition von Sinus bzw. Cosinus. □

Satz 6.3.7. $\cos x$ *und* $\sin x$ *sind stetig auf* $\mathbb{R}$. *Es gelten die* Additionstheoreme*:*

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

für alle $x, y \in \mathbb{R}$.

Beweis. Für die Stetigkeit betrachten wir eine Folge $x_n \to x$ in $\mathbb{R}$; damit gilt auch $e^{ix_n} \to e^{ix}$, also

$$\operatorname{Re} e^{ix_n} \to \operatorname{Re} e^{ix}, \quad \operatorname{Im} e^{ix_n} \to \operatorname{Im} e^{ix},$$

also sind Sinus und Cosinus stetig. Aus der Gleichung

$$e^{i(x+y)} = e^{ix}e^{iy}$$

folgt direkt durch Ausmultiplizieren der rechten Seite

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

was durch Vergleich von Real- und Imaginärteilen die Behauptung ergibt. □

Zuletzt bemerken wir noch *Potenzreihendarstellungen* von Sinus und Cosinus:

Satz 6.3.8. *Für Sinus und Cosinus gelten die* Potenzreihendarstellungen

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

Beweis. Dies folgt für den Cosinus aus

$$\cos x = \frac{1}{2}(e^{ix} + e^{-ix}) = \frac{1}{2}\sum_{n=}^{\infty}\left(\frac{i^n x^n}{n!} + (-1)^n\frac{i^n x^n}{n!}\right) = \sum_{n=0}^{\infty} i^{2n}\frac{x^{2n}}{(2n)!}$$

mit $i^{2n} = (-1)^n$; ganz analog sieht man die Reihendarstellung des Sinus. □

Die Zahl π

Lemma 6.3.9. Für $0 < x \leq 2$ gilt

$$1 - \frac{x^2}{2} < \cos x < 1 - \frac{x^2}{2} + \frac{x^4}{24},$$
$$x - \frac{x^3}{6} < \sin x < x.$$

Insbesondere ist $\sin x > 0$ für $x \in (0, 2]$.

Beweis. Ist $0 < x \leq 2$ dann folgt $x^2 < (2n+1)(2n+2)$ für alle $n \in \mathbb{N}_0$, also auch

$$a_{n+1} = \frac{x^{2n+2}}{(2n+2)!} < \frac{x^{2n}}{(2n)!} = a_n.$$

Damit sehen wir

$$\cos x = \sum_{n=0}^{\infty}(-1)^n a_n = 1 - \frac{x^2}{2} + \sum_{n=1}^{\infty}(a_{2n} - a_{2n+1}) > 1 - \frac{x^2}{2},$$

ebenso ist

$$\cos x = 1 - \frac{x^2}{2} + \frac{x^4}{24} - \sum_{n=1}^{\infty}(a_{2n+1} - a_{2n+2}) < 1 - \frac{x^2}{2} + \frac{x^4}{4!}.$$

Die Abschätzung des Sinus funktioniert genau gleich. □

Folgerung 6.3.10. $\cos(x)$ *fällt in* $[0,2]$ *streng monoton.*

Beweis. Es ist

$$\cos x - \cos y = -2\sin\left(\frac{x-y}{2}\right)\sin\left(\frac{x+y}{2}\right),$$

denn durch Ausmultiplizieren sieht man

$$\frac{1}{2}(e^{ix} + e^{-ix} - e^{iy} - e^{-iy}) = -2\frac{1}{(2i)^2}\left(e^{i(x-y)/2} - e^{-i(x-y)/2}\right)\left(e^{i(x+y)/2} - e^{-i(x+y)/2}\right).$$

Für $x, y \in [0,2]$ mit $x > y$ liegen $(x-y)/2$ und $(x+y)/2$ in $(0,2]$, also sind $\sin((x-y)/2)$ und $\sin((x+y)/2)$ positiv, d.h. $\cos x < \cos y$. □

Satz 6.3.11. *Der Cosinus hat in* $[0,2]$ *genau eine Nullstelle. Diese wird mit* $\pi/2$ *bezeichnet, und es gilt*

$$\cos\frac{\pi}{2} = 0, \qquad \sin\frac{\pi}{2} = 1.$$

Beweis. Wegen $\cos 0 = 1$, $\cos 2 < -1/3$ und der strengen Monotonie des Cosinus in $[0,2]$ hat er genau eine Nullstelle $\pi/2$ in $[0,2]$. Da der Sinus in $(0,2]$ positiv ist, folgt aus der Eulerschen Formel $\sin\pi/2 = 1$. □

Aus $e^{\pi i/2}$ folgt sofort:

x	$\frac{\pi}{2}$	π	$\frac{3\pi}{2}$	2π
e^{ix}	i	-1	$-i$	1
$\cos x$	0	-1	0	1
$\sin x$	1	0	-1	0

Satz 6.3.12. *Für alle* $z \in \mathbb{C}$ *gilt*

$$e^{z+\pi i/2} = ie^z, \quad e^{z+\pi i} = -e^z, \quad e^{z+3\pi/2} = -ie^z, \quad e^{z+2\pi i} = e^z.$$

Folgerung 6.3.13.

$$\cos(x + \frac{\pi}{2}) = -\sin x, \qquad \sin(x + \frac{\pi}{2}) = \cos x,$$

$$\cos(x + \pi) = -\cos x, \qquad \sin(x + \pi) = -\sin x,$$

$$\cos(x + \frac{3\pi}{2}) = \sin x, \qquad \sin(x + \frac{3\pi}{2}) = -\cos x,$$

$$\cos(x + 2\pi) = \cos x, \qquad \sin(x + 2\pi) = \sin x.$$

Folgerung 6.3.14. *Der Cosinus hat auf $\mathbb{R}$ genau die Nullstellen $\pi/2 + k\pi$ mit $k \in \mathbb{Z}$, der Sinus genau die Nullstellen $k\pi$ mit $k \in \mathbb{Z}$.*

28. 11. 2024

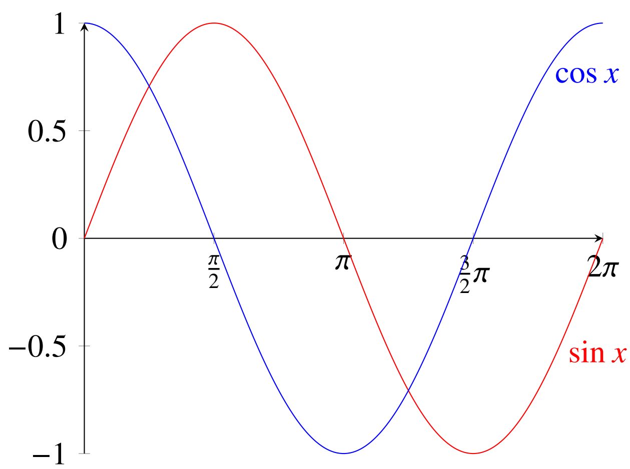
Beweis. $\cos x$ hat in $(-\pi/2, \pi/2]$ nur die Nullstelle $\pi/2$, und wegen $\cos(x+\pi) = -\cos x$ in $(-\pi/2, \pi/2+\pi]$ nur $\pi/2$ und $\pi/2+\pi$. Dieses Intervall hat die Länge 2π; da der Cosinus 2π-periodisch ist, erhält man demnach alle Nullstellen aus $\pi/2$ und $\pi/2+\pi$ durch addieren von $2k\pi$, $k \in \mathbb{Z}$. Die Nullstellen des Sinus erhält man wegen $\sin x = -\cos(x+\pi/2)$ aus den Nullstellen des Cosinus durch eine Verschiebung von $\pi/2$. □

Folgerung 6.3.15. *Für $z \in \mathbb{C}$ gilt*

$$e^z = 1 \iff z = k \cdot 2\pi i \text{ für ein } k \in \mathbb{Z}.$$

Beweis. Es gilt $e^{k2\pi i} = (e^{2\pi i})^k = 1^k = 1$. Ist umgekehrt $e^z = e^{x+iy} = 1$, folgt $|e^z| = e^x\left|e^{iy}\right| = e^x = 1$, also $x = 0$. Weiters folgt aus $e^z = e^{iy} = \cos y + i \sin y = 1$ auch $\cos y = 1$ und $\sin y = 0$. Damit ist sicher $y = m\pi$ für ein $m \in \mathbb{Z}$, und da $\cos(m\pi) = -1$ für ungerade m ist, muss m gerade sein. □

Wir können nun die wichtigsten Eigenschaften von Sinus und Cosinus, nämlich die Nullstellen, das Vorzeichen, das Monotonieverhalten und die Periodizität, in einer Grafik festhalten:



Satz 6.3.16. *Für $z \in \mathbb{C}$ gilt:*

$$|z| = 1 \iff \exists|\varphi \in [0, 2\pi) \text{ mit } e^{i\varphi} = z.$$

Beweis. Sei $|z| = 1$, $\operatorname{Im} z \geq 0$ und $z \neq -1$. Mit $z = x+iy$ ist $x \in (-1, 1] = (\cos\pi, \cos 0]$ und es gibt genau ein $\varphi \in [0, \pi)$ mit $\cos\varphi = x$. Wegen $\sin^2\varphi = \sqrt{1-\cos^2\varphi} = \sqrt{1-x^2} = y^2$ ist $\sin\varphi = \pm y$; da beide ≥ 0 sind, ist $\sin\varphi = y$ und $e^{i\varphi} = x + iy = z$.

Für $\operatorname{Im} z \leq 0$, $z \neq 1$ ist $\operatorname{Im}(-z) \geq 0$ und $-z \neq -1$, also existiert nach dem ersten Teil genau ein $\psi \in [0, \pi)$ mit $-z = e^{i\psi}$, also $z = e^{i(\pi+\psi)}$; damit gibt es aber genau ein $\varphi \in [\pi, 2\pi)$ mit $z = e^{i\varphi}$. □

Satz 6.3.17 (Polarkoordinaten). *Jede komplexe Zahl $z \in \mathbb{C}$ lässt sich in der Form*

$$z = re^{i\varphi}$$

schreiben, wobei $r = |z| \geq 0$ und $\varphi \in \mathbb{R}$ ist. Ist $z \neq 0$, dann ist φ bis auf ein ganzzahliges Vielfaches von 2π eindeutig bestimmt.

Beweis. Für $z \neq 0$ ist $z = |z| \cdot \frac{z}{|z|} = |z| \cdot e^{i\varphi}$ für ein eindeutiges $\varphi \in [0, 2\pi)$. □

Definition 6.3.18. *Tangens* und *Cotangens* werden durch

$$\tan x = \frac{\sin x}{\cos x}, \quad x \in \mathbb{R} \setminus \{\frac{\pi}{2} + k\pi : k \in \mathbb{Z}\},$$
$$\operatorname{cotan} x = \frac{\cos x}{\sin x}, \quad x \in \mathbb{R} \setminus \{k\pi : k \in \mathbb{Z}\}$$

definiert.

Zuletzt betrachten wir noch die jeweiligen Umkehrfunktionen:

Satz 6.3.19.

(i) Die Funktion $\cos x$ ist in $[0, \pi]$ streng monoton fallend und bildet dieses Intervall bijektiv auf $[-1, 1]$ ab. Die Umkehrfunktion

$$\arccos\colon [-1, 1] \to [0, \pi]$$

heißt Arcus-Cosinus.

(ii) Die Funktion $\sin x$ ist in $[-\pi/2, \pi/2]$ streng monoton wachsend und bildet dieses Intervall bijektiv auf $[-1, 1]$ ab. Die Umkehrfunktion

$$\arcsin\colon [-1, 1] \to [-\pi/2, \pi/2]$$

heißt Arcus-Sinus.

(iii) Die Funktion $\tan x$ ist im Intervall $(-\pi/2, \pi/2)$ streng monoton wachsend und bildet dieses Intervall bijektiv auf $\mathbb{R}$ ab. Die Umkehrfunktion

$$\arctan\colon \mathbb{R} \to (-\pi/2, \pi/2)$$

heißt Arcus-Tangens.

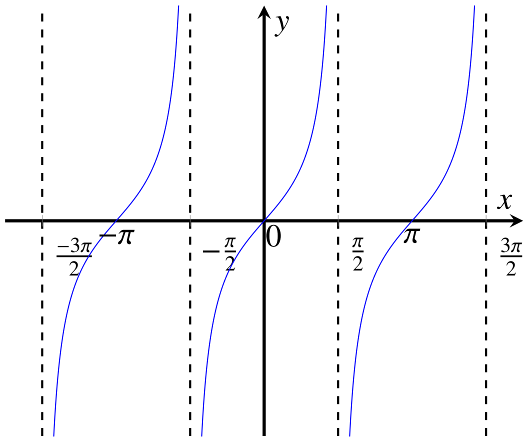
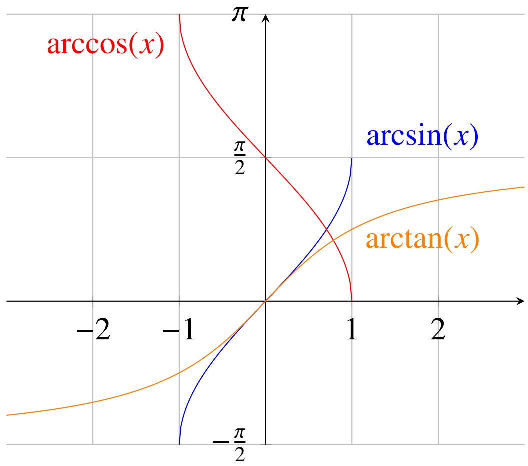


Abbildung 6.1: Graph des Tangens



Beweis. (i) und (ii) haben wir schon gesehen. Für (iii) ist für $0 \leq x < y$ $\sin x < \sin y$ und $\cos x > \cos y$, also $\tan x > \tan y$, d.h. der Tangens ist streng monoton wachsend auf $[0, \pi/2)$. Da $\tan(-x) = -\tan(x)$, ist er auch streng monoton wachsend auf $(-\pi/2, 0]$, also auf dem ganzen Intervall $(-\pi/2, \pi/2)$. Aus $\cos \pi/2 = 0$ und $\sin \pi/2 = 1$ folgt $\lim_{x \to \pi/2-} \tan x = \infty$, und aus $\tan(-x) = -\tan x$ auch $\lim_{x \to -\pi/2+} \tan x = -\infty$. □

Für die trigonometrischen Funktionen gelten eine Reihe weiterer Identitäten; diese wollen wir hier nicht alle entwickeln, sondern verweisen auf einschlägige Formelsammlungen.

6.4 Hyperbelfunktionen

Der Vollständigkeit halber erwähnen wir auch die *Hyperbelfunktionen*:

Definition 6.4.1. Für $x \in \mathbb{R}$ setzen wir

$$\begin{aligned}
\cosh x &:= \frac{e^x + e^{-x}}{2} && \text{(Cosinus Hyperbolicus)},\\
\sinh x &:= \frac{e^x - e^{-x}}{2} && \text{(Sinus Hyperbolicus)},\\
\tanh x &:= \frac{\sinh x}{\cosh x} && \text{(Tangens Hyperbolicus)},\\
\coth x &:= \frac{\cosh x}{\sinh x} && \text{(Cotangens Hyperbolicus)}.
\end{aligned}$$

Es gilt (ohne Beweis):

- $\sinh\colon \mathbb{R} \to \mathbb{R}$ ist strikt monoton wachsend und bijektiv.
- $\cosh\colon \mathbb{R} \to [1, \infty)$ bildet $[0, \infty)$ bijektiv und strikt monoton wachsend auf $[1, \infty)$ ab.
- $\tanh\colon \mathbb{R} \to (-1, 1)$ ist strikt monoton wachsend.
- $\cosh x = \cosh(-x)$, $\sinh x = -\sinh(-x)$, $\tanh x = -\tanh(-x)$,
- $\cosh^2 x - \sinh^2 x = 1$,
- $\cosh^2 x + \sinh^2 x = \cosh 2x$,
- $e^x = \cosh x + \sinh x$,
- $\sinh(x + y) = \sinh x \cosh y + \cosh x \sinh y$,
- $\cosh(x + y) = \cosh x \cosh y + \sinh x \sinh y$.

Auch von den Hyperbelfunktionen existieren Umkehrfunktionen:

- $\operatorname{arsinh} x\colon \mathbb{R} \to \mathbb{R}$ (Area-Sinus Hyperbolicus),
- $\operatorname{arcosh} x\colon [1, \infty) \to [0, \infty)$ (Area-Cosinus Hyperbolicus),
- $\operatorname{artanh} x\colon (-1, 1) \to \mathbb{R}$ (Area-Tangens Hyperbolicus).

6.5 Polynome und rationale Funktionen

Eine wichtige Klasse von Funktionen sind die *Polynome*

$$p(x) = a_n x^n + a_{n-1} x^{n-1} + \ldots + a_1 x + a_0 \tag{6.3}$$

mit $n \in \mathbb{N}_0$ und *Koeffizienten* $a_i \in \mathbb{C}$. Ist $a_n(x) \neq 0$, heißt n der *Grad* von p, $n = \operatorname{Grad}(p)$. Das *Nullpolynom* ist $p(x) = 0$; ordnet man ihm den Grad $-\infty$ zu, gelten die Regeln

$$\operatorname{Grad}(p + q) \leq \max(\operatorname{Grad} p, \operatorname{Grad} q), \qquad \operatorname{Grad}(p \cdot q) = \operatorname{Grad} p + \operatorname{Grad} q.$$

für alle Polynome, wenn man

$$-\infty < k \text{ für } k \in \mathbb{N}_0, \qquad -\infty + k = k + (-\infty) = -\infty \text{ für } k \in \mathbb{N}_0 \cup \{-\infty\}$$

postuliert, außerdem ist dann $p(x) = 0$ mitgemeint wenn man von einem Polynom vom Grad $\leq n$, $n \in \mathbb{N}$ redet. Sind alle Koeffizienten reell, heißt p *reelles Polynom*; dann ist auch $p(x) \in \mathbb{R}$ für alle $x \in \mathbb{R}$.

Ein wichtiges Hilfsmittel im Umgang mit Polynomen ist die Division mit Rest (Euklidischer Algorithmus):

Satz 6.5.1 (Division mit Rest). *Sind p und q Polynome mit $q \neq 0$, dann gibt es eindeutig bestimmte Polynome s und r mit*

$$p = sq + r \qquad \textit{mit } \operatorname{Grad} r < \operatorname{Grad} q.$$

Sind p und q reell, dann auch s und r.

Ist $r = 0$, so nennt man q einen *Teiler* von p.

Beweis. Ist $\operatorname{Grad} p < \operatorname{Grad} q$, so ist $p = 0 \cdot q + p$ die gewünschte Zerlegung, also gehen wir nun von

$$p(x) = \sum_{k=0}^{n} a_n x^n, \quad q(x) = \sum_{j=0}^{m} b_m x^m$$

mit $\operatorname{Grad} p = n \geq m = \operatorname{Grad} q$, d.h. $a_n \neq 0 \neq b_m$ aus. Subtrahiert man von p das Polynom $a_n b_m^{-1} x^{n-m} q(x)$ erhält man ein Polynom p_1 vom Grad $n_1 < n$. Ist $n_1 \geq m$ subtrahiert man ebenso von p_1 ein Vielfaches von q, so dass die Differenz ein Polynom vom Grad $n_2 < n_1$ ist. So fährt man fort, bis ein Polynom r mit $\operatorname{Grad} r < m$ übrig bleibt, und für ein geeignetes Polynom s hat man dann $p - sq = r$.

Wäre $p = s'q + r'$ eine weitere Zerlerung mit $s' \neq s$, dann folgte der Widerspruch $\operatorname{Grad}((s' - s)q) = \operatorname{Grad}(r - r') < \operatorname{Grad} q$; damit ist auch $r = p - sq$ eindeutig bestimmt.

Die letzte Behauptung folgt aus der Konstruktion. □

Wir illustrieren den Divisionsalgorithmus am Beispiel $p(x) = x^3 + 2x^2 - 2x - 1$ und $q(x) = x^2 + 1$:

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

Somit ist $p(x) = q(x) \cdot (x + 2) + (-3x - 3)$.

Lemma 6.5.2. Ist $\alpha \in \mathbb{C}$ Nullstelle eines Polynomes p mit $\operatorname{Grad} p \geq 1$, d.h. $f(\alpha = 0)$, dann ist f durch $x - \alpha$ teilbar.

Beweis. Division durch $x - \alpha$ ergibt $p(x) = s(x)(x - \alpha) + r(x)$ mit $r \in \mathbb{C}$; setzt man $x = \alpha$ folgt $r = 0$. □

Da bei jeder Division durch einen Linearfaktor $x - \alpha$ der Grad um 1 kleiner wird, folgt:

Folgerung 6.5.3. *Ein Polynom vom Grad $n \geq 1$ hat höchstens n Nullstellen.*

Folgerung 6.5.4. *Stimmen die Werte der Polynome*

$$p(x) = \sum_{k=0}^{n} a_k x^k, \quad q(x) = \sum_{k=0}^{n} b_k x^k$$

an $n+1$ verschiedenen Stellen überein, dann ist $a_k = b_k$ für $k = 0 \dots n$.

Beweis. $p - q$ hat $n+1$ verschiedene Nullstellen und Grad $\le n$, also ist $f - g = 0$ das Nullpolynom. □

Es folgt daraus, dass zwei Polynome genau dann gleich sind, wenn sie die gleichen Koeffizienten haben, d.h. die Polynomfunktion (6.3) ist durch ihre Koeffizienten eindeutig bestimmt.

Folgendes Resultat werden wir in Analysis II beweisen:

Satz 6.5.5 (Fundamentalsatz der Algebra). *Jedes Polynom vom Grad ≥ 1 besitzt eine Nullstelle in $\mathbb{C}$.*

Satz 6.5.6. *Jedes Polynom p mit* Grad $p \ge 1$ *besitzt eine Darstellung*

$$f(z) = a(z-\alpha_1)^{k_1} \cdots (z-\alpha_s)^{k_s} \tag{6.4}$$

mit $s \in \mathbb{N}$ verschiedenen Nullstellen $\alpha_1, \dots, \alpha_s \in \mathbb{C}$ und Vielfachheiten *k_i, für welche $k_1 + \ldots + k_s = n$ gilt.*

Hierbei heißt α k-fache Nullstelle von p und $k \in \mathbb{N}$ die Vielfachheit dieser Nullstelle, wenn k die größte natürliche Zahl ist, für die $(x-\alpha)^k$ Teiler von p ist.

Da nicht jedes reelle Polynom eine Nullstelle in $\mathbb{R}$ besitzt, ist eine Zerlegung der Form (6.4) mit $\alpha_i \in \mathbb{R}$ im Allgemeinen nicht möglich. Es ist aber für ein relles Polynom mit einer Nullstelle $\alpha \in \mathbb{C} \setminus \mathbb{R}$ auch $\bar{\alpha}$ eine Nullstelle, denn

$$p(\bar{\alpha}) = \sum_k a_k \bar{\alpha}^k = \overline{\sum_k a_k \alpha^k} = \overline{p(\alpha)} = 0.$$

Fasst man die konjugierten Linearfaktoren wie in

$$(x-\alpha)(x-\bar{\alpha}) = x^2 - (\alpha + \bar{\alpha}) + \alpha\bar{\alpha} = x^2 - 2\operatorname{Re}\alpha x + |\alpha|^2$$

als quadratisches Polynom zusammen, so sieht man:

Satz 6.5.7. *Jedes reelle Polynom kann als Produkt reeller Polynome mit Graden ≤ 2 dargestellt werden*

Rationale Funktionen

Unter einer rationalen Funktion verstehen wir eine Funktion der Bauart

$$R(z) = \frac{p(z)}{q(z)},$$

wobei p und q Polynome sind. R heißt *echt gebrochen* wenn $\operatorname{Grad} p < \operatorname{Grad} q$, und sonst *unecht gebrochen*. Nach Satz 6.5.1 kann man jede rationale Funktion als Summe eines Polynomes und einer echt gebrochenen Funktion darstellen, so dass wir uns auf die Untersuchung von echt gebrochenen rationalen Funktionen beschränken können.

R ist überall außer in den Nullstellen von q definiert. Ist α aber eine k-fache Nullstelle des Nenners und auch der Zähler durch $(x-\alpha)^k$ teilbar, können wir durch diesen Faktor kürzen und R ist auch im Punkt α erklärt. $\alpha \in \mathbb{R}$ heißt n-facher Pol von R, wenn es eine Darstellung $R = p/q$ gibt, in der $p(\alpha) \neq 0$ ist und q in α eine n-fache Nullstelle hat; zu jeder rationalen Funktion R gibt es also eine *gekürzte* Darstellung r. Wir werden diese nun weiter zerlegen:

Satz 6.5.8 (Partialbruchzerlegung rationaler Funktionen). *Sei $r = p/q$ eine echt gebrochene rationale Funktion und q habe die Form*

$$q(z) = a_n(z-z_1)^{\nu_1}(z-z_2)^{\nu_2}\cdots(z-z_m)^{\nu_m}$$

mit $a_n \neq 0$, $z_j \neq z_k$ für $j \neq k$ und $\nu_1 + \ldots \nu_m = n$. Dann besitzt r eine Summendarstellung der Form

$$\begin{aligned} r(z) = {} & \frac{a_{11}}{z-z_1} + \frac{a_{12}}{(z-z_1)^2} + \ldots + \frac{a_{1\nu_1}}{(z-z_1)^{\nu_1}} \\ & + \frac{a_{21}}{z-z_2} + \frac{a_{22}}{(z-z_2)^2} + \ldots + \frac{a_{2\nu_2}}{(z-z_2)^{\nu_2}} \\ & + \ldots \\ & + \frac{a_{m1}}{z-z_m} + \frac{a_{m2}}{(z-z_m)^2} + \ldots + \frac{a_{m\nu_m}}{(z-z_m)^{\nu_m}} \end{aligned} \tag{6.5}$$

Beweis. Wir führen Induktion nach dem Grad n des Nennerpolynoms durch. Für $n = 1$ ist $p \in \mathbb{C}$ und

$$r(z) = \frac{c}{a_1(z-z_1)} = \frac{a_{11}}{z-z_1} \text{ mit } a_{11} = \frac{c}{a_1}.$$

Sei nun $n > 1$ und der Satz gelte für alle P/Q mit $\operatorname{Grad} Q < n$. Dann ist

$$q(z) = (z-z_1)^{\nu_1} s(z) \text{ mit } s(z) = a_n(z-z_2)^{\nu_2}\ldots(z-z_m)^{\nu_m}$$

und für jedes $a \in \mathbb{C}$ ist

$$\frac{p(z)}{q(z)} - \frac{a}{(z-z_1)^{\nu_1}} = \frac{p(z) - as(z)}{(z-z_1)^{\nu_1} s(z)}.$$

Wählen wir $a = p(z_1)/s(z_1)$ ist $p(z_1) - as(z_1) = 0$. Ist $p(z) - as(z)$ überall 0, sind wir mit $p(z)/q(z) = a/(z-z_1)^{\nu_1}$ fertig, ansonsten können wir den Linearfaktor $z - z_1$ abspalten, d.h. es gibt ein Polynom $P(z)$ mit

$$p(z) - as(z) = (z-z_1)P(z),$$

und es folgt

$$r(z) = \frac{p(z)}{q(z)} = \frac{a}{(z-z_1)}^{\nu_1} + \frac{P(z)}{Q(z)}$$

mit

$$Q(z) = (z-z_1)^{\nu_1-1}s(z) = a_n(z-z_1)^{\nu_1-1}(z-z_2)^{\nu_2}\cdots(z-z_m)^{\nu_m}.$$

Nach Induktionsvoraussetzung können wir nun aber $P(z)/Q(z)$ in die gewünschten Partialbrüche zerlegen und sind fertig. □

Will man im Reellen bleiben, gilt ein ähnliches Resultat:

Satz 6.5.9. *Sei $r = p/q$ eine echt gebrochene reelle rationale Funktion und q habe die Form*

$$q(z) = a(x-x_1)^{\rho_1}(x-x_2)^{\rho_2}\cdots(x-x_r)^{\rho_r}(x^2+A_1x+B_1)^{\sigma_1}\cdots(x^2+A_sx+B_s)^{\sigma_s}$$

mit $\rho_1+\ldots+\rho_r+2\sigma_1+\ldots+2\sigma_s = n$, $x_j \neq x_k$ für $j \neq k$ und unter sich verschiedenen Polynomen $x^2+A_jx+B_j$ mit rellen Koeffizienten, aber ohne reelle Nullstellen. Dann besitzt r eine Summendarstellung der Form

$$\begin{aligned} r(z) = {} & \frac{a_{11}}{x-x_1} + \frac{a_{12}}{(x-x_1)^2} + \ldots + \frac{a_{1\rho_1}}{(x-x_1)^{\rho_1}} \\ & + \ldots \\ & + \frac{a_{m1}}{x-x_m} + \frac{a_{m2}}{(x-x_m)^2} + \ldots + \frac{a_{m\rho_r}}{(x-x_m)^{\rho_r}} \\ & + \frac{\alpha_{11}x+\beta_11}{x^2+A_1x+B_1} + \frac{\alpha_{12}x+\beta_{12}}{(x^2+A_1x+B_1)^2} + \ldots + \frac{\alpha_{1\sigma_1}x+\beta_{1\sigma_1}}{(x^2+A_1x+B_1)^{\sigma_1}} \\ & + \ldots \\ & + \frac{\alpha_{s1}x+\beta_{s1}}{x^2+A_sx+B_s} + \frac{\alpha_{s2}x+\beta_{s2}}{(x^2+A_sx+B_s)^2} + \ldots + \frac{\alpha_{s\sigma_1}x+\beta_{s\sigma_1}}{(x^2+A_sx+B_s)^{\sigma_s}}. \end{aligned}$$

Wie erhält man nun die Koeffizienten a_{ij} und β_{ij}? Multipliziert man (6.5) mit dem Nennerpolynom $q(z)$, so erhält man nach Kürzen durch Koeffizientenvergleich ein lineares Gleichungssystem für die Koeffizienten a_{ij}, aus welchem man diese berechnen kann. Ebenso kann man in (6.5) n verschiedene, möglichst geeignet gewählte Zahlen einsetzen, um ein anderes Gleichungsystem zu erlangen. Für die Koeffizienten der Terme, die im Nenner nur Potenzen von Linearfaktoren enthalten, ist folgende Methode oft besser: man kann in (6.5) zuerst die Koeffizienten $a_{i\nu_i}$ bestimmen, indem man mit $(z-z_i)^{\nu_i}$ multipliziert und (nach kürzen) $z = z_i$ einsetzt. Bringt man die Terme $a_{i\nu_i}/(z-z_i)^{\nu_i}$ auf die linke Seite, entsteht eine rationale Funktion in deren Nennerpolynom z_i höchstens eine (ν_i-1)-fache Nullstelle ist, und man kann so fortfahren bis alle Koeffizienten bestimmt sind.

Beispiel 6.5.10. *(i) Wir möchten eine Partialbruchzerlegung der rationalen Funktion $\frac{4x+1}{(x-2)^2}$ durchführen. Mit dem ersten Ansatz finden wir*

$$\frac{4x+1}{(x-2)^2} = \frac{A_1}{x-2} + \frac{A_2}{(x-2)^2}$$

Multiplikation auf beiden Seiten mit $(x-2)^2$ *führt zu*

$$4x+1 = A_1(x-2) + A_2 = A_1 x + A_2 - 2A_1, \tag{6.6}$$

und Koeffizientenvergleich ergibt $A_1 = 4$ *und aus* $1 = A_2 - 8$ *folgt* $A_2 = 9$. *Somit gilt*

$$\frac{4x+1}{(x-2)^2} = \frac{4}{x-2} + \frac{9}{(x-2)^2}.$$

Alternativ setzt man $x = 2$ *in* (6.6), *was* $9 = A_2$ *ergibt, und sieht aus*

$$\frac{4x+1-9}{(x-2)^2} = \frac{4x-8}{(x-2)^2} = \frac{4}{x-2} = \frac{A_1}{x-2}$$

dass $4 = A_1$ *ist.*

(ii) Für eine Partialbruchzerlegung von $\frac{x+1}{x^4-x}$ *finden wir zuerst die Nullstellen* 0 *und* 1 *von* x^4+x, *womit wir durch Division des Zählers durch* $x(x-1)$ *folgenden Ansatz erhalten:*

$$\frac{x+1}{x^4-x} = \frac{x+1}{x(x-1)(x^2+x+1)} = \frac{a}{x} + \frac{b}{x+1} + \frac{\alpha x+\beta}{x^2+x+1}.$$

Wir erhalten sofort

$$\begin{aligned} \frac{x+1}{(x-1)(x^2+x+1)} &= a + x\cdot(\dots) \implies a = -1, \\ \frac{x+1}{x(x^2+x+1)} &= b + (x-1)\cdot(\dots) \implies b = \frac{2}{3}. \end{aligned}$$

Setzen wir $x = -1$ *und* $x = 2$ *ein, erhalten wir die Gleichungen*

$$\begin{aligned} 0 &= 1 - \frac{1}{3} - \alpha + \beta \implies \alpha - \beta = \frac{2}{3} \\ \frac{3}{14} &= -\frac{1}{2} + \frac{2}{3} + \frac{2}{7}\alpha + \frac{1}{7}\beta \implies 2\alpha + \beta = \frac{1}{3} \end{aligned}$$

29. 11. 2024 *was zur Lösung* $\alpha = 1/3$, $\beta = -1/3$ *führt.*

Kapitel 7

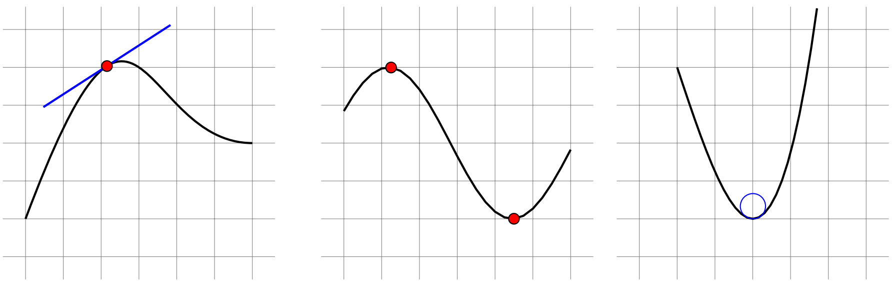
Differentialrechnung

Lernziele

- ► Definition der Ableitung wiedergeben und anschaulich verstehen
- ► Ableitung als Linearisierung verstehen
- ► Ableitungsregeln kennen und anwenden
- ► Ableitung der Umkehrfunktion bestimmen
- ► Lokale Extrema finden und berechnen
- ► Satz von Rolle und Mittelwertsatz anwenden
- ► Konvexität / Konkavität verstehen und identifizieren
- ► Grenzwert von Quotienten mit der Regel von de l'Hospital berechnen
- ► Taylorentwicklung von Funktionen bestimmen
- ► Kuvendiskussionen Durchführen

7.1 Differenzierbarkeit

Die Motivation der Differentialrechung ist, das lokale Verhalten von Funktionen zu beschreiben, z.B. über Tangenten, Extremalstellen oder Schmiegekreise.



Im Folgenden sei $D \subseteq \mathbb{R}$ eine Menge und $a \in \mathbb{R}$ Häufungspunkt von D.

Definition 7.1.1. Eine Funktion $f\colon D \to \mathbb{R}$ heißt *differenzierbar in a*, wenn der Grenzwert

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

existiert. Man nennt $f'(a) \in \mathbb{R}$ die *Ableitung* von f in a.

f heißt in D differenzierbar, wenn jeder Punkt von D Häufungspunkt ist und f in jedem Punkt $a \in D$ differenzierbar ist.

Andere Schreibweisen für die Ableitung sind $\dot{f}(a)$, $\partial f(a)$, $Df(a)$ oder $\frac{\mathrm{d}f}{\mathrm{d}x}(a)$.

Wir führen zuerst eine praktische Umformulierung der Differenzierbarkeit an.

Satz 7.1.2. *Für $f \colon D \to \mathbb{R}$ sind äquivalent:*

(i) f ist in a differenzierbar.

(ii) $\exists m_a \in \mathbb{R}$ und eine in a stetige Funktion $r \colon D \to \mathbb{R}$ mit $r(a) = 0$ und

$$f(x) = f(a) + m_a(x-a) + r(x)(x-a) \quad \forall x \in D.$$

In (ii) ist $m_a = f'(a)$.

Beweis. Für (i) $\implies$ (ii) setze

$$r(x) := \begin{cases} 0 & x = a \\ \frac{f(x)-f(a)-m_a(x-a)}{x-a} & x \neq a. \end{cases}$$

(ii) $\implies$ (i) ist klar. □

Folgerung 7.1.3. *Ist $f \colon D \to \mathbb{R}$ in a differenzierbar, so ist f in a stetig.*

Beispiel 7.1.4.

(i) Für $n \in \mathbb{N}$ ist $f(x) = x^n$ in jedem Punkt $a \in \mathbb{R}$ differenzierbar, denn

$$\begin{aligned} \frac{f(x+h)-f(x)}{h} &= \frac{1}{h}((x+h)^n - x^n) = \frac{1}{h}\sum_{k=1}^{n}\binom{n}{k}h^k x^{n-k} \\ &= \sum_{k=1}^{n}\binom{n}{k}h^{k-1}x^{n-k} \to \binom{n}{1}x^{n-1} = nx^{n-1}. \end{aligned}$$

(ii) $g(x) = e^x$ ist auf ganz $\mathbb{R}$ differenzierbar und es gilt $g'(x) = e^x$, denn

$$\begin{aligned} \frac{g(x+h)-g(x)}{h} &= \frac{1}{h}\left(e^{x+h} - e^x\right) = \frac{1}{h}e^x(e^h - 1) \\ &= e^x\frac{1}{h}\sum_{n=1}^{\infty}\frac{h^n}{n!} = e^x\sum_{n=1}^{\infty}\frac{h^{n-1}n!}{=}e^x(1 + h\sum_{n=2}^{\infty}\frac{h^{n-2}}{n!} \to e^x, \end{aligned}$$

denn für $|h| < 1$ gilt

$$\left|\sum_{n=2}^{\infty}\frac{h^{n-2}}{n!}\right| \leq \sum_{n=2}^{\infty}\frac{|h|^{n-2}}{n!} \leq \sum_{n=1}^{\infty}\frac{1}{n!} < \infty.$$

(iii) Die Betragsfunktion $x \mapsto |x|$ ist in $x = 0$ nicht differenzierbar: Für jede Folge (x_n) mit $x_n > 0$ (y_n) mit $x_n > 0$ gilt $\lim_{n\to\infty}\frac{x_n}{x_n} = 1 \neq -1 = \lim_{n\to\infty}\frac{|y_n|}{y_n}$.

Bemerkung. Definition 7.1.1 ergibt auch für $f\colon D \to \mathbb{C}$ Sinn (immer noch mit $D \subseteq \mathbb{R}$!); f ist dann differenzierbar genau dann wenn Real- und Imaginärteil differenzierbar sind, und $f'(a) = (\operatorname{Re} f)'(a) + (\operatorname{Im} f)'(a)$.

Satz 7.1.5. *Sind $f, g\colon D \to \mathbb{R}$ in $a \in D$ differenzierbar und $\lambda \in \mathbb{R}$, dann sind auch die Funktionen $f + g$, fg und λf in a differenzierbar und es gilt*

(i) Linearität: $(f + g)'(a) = f'(a) + g'(a)$, $(\lambda f)'(a) = \lambda f'(a)$;

(ii) Produktregel: $(fg)'(a) = f'(a)g(a) + f(a)g'(a)$;

(iii) Quotientenregel: ist $g(a) \neq 0$, dann ist auch f/g in a differenzierbar und es gilt

$$\left(\frac{f}{g}\right)'(x) = \frac{f'(a)g(a) - f(a)g'(a)}{g(a)^2}. \tag{7.1}$$

Beweis. (i) folgt sofort aus den Rechenregeln für konvergente Folgen bzw. Grenzwerte. (ii) sieht man aus

$$\frac{f(x)g(x) - f(a)g(a)}{x - a} = \frac{f(x) - f(a)}{x - a} g(x) + f(a)\frac{g(x) - g(a)}{x - a}, \quad x \neq a.$$

(iii): wenn $g(a) \neq 0$ ist ist auch $g(x) \neq 0$ in der Nähe von a, d.h. für $x \in D$ mit $|x| < \delta$ für ein $\delta > 0$. Für solche $x \neq a$ gibt

$$\left(\frac{f(x)}{g(x)} - \frac{f(a)}{g(a)}\right)\frac{1}{x - a} = \frac{1}{g(x)g(a)}\left(\frac{f(x) - f(a)}{x - a} g(a) - f(a)\frac{g(x) - g(a)}{x - a}\right)$$

für $x \to a$ die Behauptung. □

Bemerkung. Die Umkehrung von 7.1.5 gilt i.A. nicht. D.h. ist die Summe zweier Funktionen differenzierbar, so müssen die Summanden selbst nicht differenzierbar sein: Es ist z.B. $\mathbb{1}_{\mathbb{R}_0^+} + \mathbb{1}_{\mathbb{R}^-} = 1$, wobei $\mathbb{R}_0^+ = [0, \infty)$ und $\mathbb{R}^- = (-\infty, 0)$ ist.

Satz 7.1.6 (Kettenregel). *Sei $f\colon D \to \mathbb{R}$ in a differenzierbar, $f(D) \subseteq E \subseteq \mathbb{R}$ und $g\colon E \to \mathbb{R}$ in $f(a)$ differenzierbar. Dann ist $g \circ f$ in a differenzierbar und es gilt*

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

Beweis. Die Funktion

$$g^*(y) := \begin{cases} \frac{g(y) - g(f(a))}{y - f(a)} & y \neq f(a) \\ g'(f(a)) & y = f(a) \end{cases}$$

ist stetig in a und

$$\frac{g(f(x)) - g(f(a))}{x - a} = g^*(f(x))\frac{f(x) - f(a)}{x - a} \to g'(f(a))f'(a)$$

für $x \to a$ gibt die Behauptung. □

5. 12. 2024

Beispiel. Mit derselben Rechnung wie oben ist e^{ix} auf $\mathbb{R}$ differenzierbar und $(e^{ix})' = ie^{ix}$. Schreibt man diese Identität mit der Eulerschen Formel aus, d.h.

$$(\cos x)' + i(\sin x)' = i\cos x - \sin x$$

und vergleicht Real- und Imaginärteil auf beiden Seiten, erhält man

$$(\cos x)' = -\sin x, \qquad (\sin x)' = \cos x.$$

Alternativ kann man auch die Kettenregel verwenden:

$$(\cos x)' = \left(\frac{e^{ix} + e^{-ix}}{2}\right)' = \frac{ie^{ix} - ie^{-ix}}{2} = -\frac{e^{ix} - e^{-ix}}{2i} = -\sin(x),$$
$$(\sin x)' = \left(\frac{e^{ix} - e^{-ix}}{2i}\right)' = \frac{ie^{ix} + ie^{-ix}}{2i} = \frac{e^{ix} + e^{-ix}}{2} = \cos x.$$

Satz 7.1.7. *Sei $f\colon D \to \mathbb{R}$ injektiv und in a differenzierbar sowie $f^{-1}\colon f(D) \to D$ in $b = f(a)$ stetig. Dann ist f^{-1} in b differenzierbar genau dann wenn $f'(a) \neq 0$. In diesem Fall gilt*

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

Beweis. Ist f^{-1} in a differenzierbar, folgt aus $f^{-1} \circ f = \mathrm{Id}_D$

$$1 = (\mathrm{Id}_D)'(a) = (f^{-1})'(f(a)) \cdot f'(a),$$

also $f'(a) \neq 0$.

Umgekehrt gibt es eine Folge (x_n) in $D \setminus \{a\}$ die gegen a konvergiert (da Häufungspunkt), also gilt $f(x_n) \to f(a)$ mit $f(x_n) \neq f(a)$, d.h. $f(a)$ ist Häufungspunkt von $Y = f(D)$. Sei nun (y_n) eine Folge in $Y \setminus \{b\}$ mit $y_n \to b$, dann ist $a \neq x_n := f^{-1}(y_n) \to a$. Wegen $f'(a) \neq 0$ ist für große n auch

$$\frac{f(x_n) - f(a)}{x_n - a} \neq 0$$

und damit

$$\frac{f^{-1}(y_n) - f^{-1}(b)}{y_n - b} = \left(\frac{f(x_n) - f(a)}{x_n - a}\right)^{-1} \to \frac{1}{f'(a)}. \qquad \square$$

Beispiel 7.1.8.

(i) Die Funktion $f\colon (0,\infty) \to (0,\infty)$, $f(x) := x^2$ ist injektiv, differenzierbar und besitzt die Umkehrfunktion $g(y) = \sqrt{y}$, $f'(x) = 2x$ ist ungleich Null für $x \neq 0$. Die Ableitung von g in $y = f(x) = x^2 \neq 0$ lautet $g'(y) = \frac{1}{f'(x)} = \frac{1}{2x} = \frac{1}{2\sqrt{y}}$.

(ii) Der Logarithmus $\log(x)$ ist in jedem Punkt $x > 0$ differenzierbar und es gilt $(\log(x))' = \frac{1}{x}$ (Übung!).

(iii) Für jedes $\alpha \in \mathbb{R}$ und $x > 0$ ist x^α differenzierbar und es gilt $(x^\alpha)' = \alpha x^{\alpha-1}$. Dies folgt aus Satz 7.1.7 angewandt auf die Funktion $f(x) = \exp(\alpha \log(x))$ und der Kettenregel.

(iv) Für $a > 0$ und $x \in \mathbb{R}$ ist $(a^x)' = a^x \log a$ (Übung!).

(v) Jede Funktion $x \mapsto x^n$ für $n \in \mathbb{N}$ ist differenzierbar mit Ableitung nx^{n-1}, damit sind auch alle Polynomfunktionen $p(x) = a_n x^n + \ldots a_1 x + a_0$ differenzierbar. Wegen der Quotientenregel (7.1) sind auch alle rationalen Funktionen auf ihrem Definitionsbereich differenzierbar.

(vi) Die Funktion $\tan(x) = \frac{\sin(x)}{\cos(x)}$ ist auf $(-\pi/2, \pi/2)$ differenzierbar. Mittels der Quotientenregel folgt

$$(\tan x)' = \frac{\cos^2 x + \sin^2 y}{\cos^2 x} = 1 + \tan^2(x) = \frac{1}{\cos^2 x}.$$

Für die Umkehrfunktion arctan gilt

$$\frac{\mathrm{d}}{\mathrm{d}x} \arctan(x) = \frac{1}{1 + x^2}, \qquad x \in \mathbb{R}.$$

(vii) Für die Umkehrfunktionen arcsin und arccos gilt

$$\frac{\mathrm{d}}{\mathrm{d}x} \arcsin(x) = \frac{1}{\sqrt{1 - x^2}}, \quad \frac{\mathrm{d}}{\mathrm{d}x} \arccos(x) = -\frac{1}{\sqrt{1 - x^2}} \quad \text{für } |x| < 1.$$

7.2 Lokale Extrema, Mittelwertsatz

Definition 7.2.1. Ist $f: D \to \mathbb{R}$ differenzierbar und $f': D \to \mathbb{R}$ auch, so sagen wir, f ist *zweimal differenzierbar*, und $f'' = (f')'$ ist die zweite Ableitung. Analog definiert man n-fache Differenzierbarkeit und die n-te Ableitung $f^{(n)}$. f heißt n-mal *stetig differenzierbar*, wenn die n-te Ableitung auch stetig ist.

Definition 7.2.2. Eine Funktion $f: (a, b) \to \mathbb{R}$ hat in $x \in (a, b)$ ein *lokales Maximum* bzw. *lokales Minimum* wenn es ein $\varepsilon > 0$ gibt mit

$$\forall y \in (a, b) : |x - y| < \varepsilon \implies f(y) \leq f(x) \quad \text{bzw. } f(y) \geq f(x).$$

Ein *lokales Extremum* ist ein lokales Minimum oder Maximum. Gilt $f(y) \leq f(x)$ bzw. $f(y) \geq f(x)$ für alle y im Definitionsbereich von f, so heißt x *globales Maximum* bzw. *globales Minimum*.

Ist die Ungleichung *strikt*, d.h. mit „<“ bzw. „>“, spricht man von einem *strikten* Maximum bzw. Minimum.

Eine notwendige Bedingung für ein lokales Extremum lautet wie folgt:

Satz 7.2.3. *Besitzt $f: (a, b) \to \mathbb{R}$ in $x \in (a, b)$ ein lokales Extremum und ist f in x differenzierbar, dann ist $f'(x) = 0$.*

Beweis. Besitzt f in x ein lokales Maximum, dann ist $f(x + h) \leq f(x)$ für kleine h und damit $(f(x + h) - f(x))/h \leq 0$ für $h > 0$ und $(f(x + h) - f(x))/h \geq 0$ für $h < 0$. Damit ist $f'(x) \leq 0$ und $f'(x) \geq 0$, also $f'(x) = 0$. □

Satz 7.2.4 (Satz von Rolle)**.** *Sei $f\colon [a,b] \to \mathbb{R}$ stetig und in (a,b) differenzierbar mit $f(a) = f(b)$. Dann gibt es ein $\theta \in (a,b)$ mit $f'(\theta) = 0$.*

Beweis. f nimmt nach Satz 5.2.7 ein Maximum oder Minimum in einem Punkt $\theta \in [a,b]$ an; ist f konstant kann man $\theta \in (a,b)$ wählen, andernfalls muss wegen $f(a) = f(b)$ auch $\theta \in (a,b)$ gelten. Dami ist θ ein lokales Extremum, also $f'(\theta) = 0$. □

Satz 7.2.5 (Mittelwertsatz der Differentialrechnung)**.** *Sei $a < b$ und $f\colon [a,b] \to \mathbb{R}$ stetig und in (a,b) differenzierbar.*

(i) Es gibt ein $\theta \in (a,b)$ mit

$$f'(\theta) = \frac{f(b) - f(a)}{b - a}.$$

(ii) Ist $g\colon [a,b] \to \mathbb{R}$ auch stetig und in (a,b) differenzierbar, so gibt es $\theta \in (a,b)$ mit

$$f'(\theta)(g(b) - g(a)) = g'(\theta)(f(b) - f(a)).$$

Falls $g'(x) \neq 0$ für alle x schreibt man das auch als

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

Während man mit dem Mittelwertsatz üblicherweise Teil (i) meint, nennt man Teil (ii) oft den *verallgemeinerten Mittelwertsatz.*

Beweis. (i) folgt aus (ii) für $g(x) = x$. Für (ii) ist

$$h(x) = (f(x) - f(a))(g(b) - g(a)) - (g(x) - g(a))(f(b) - f(a)).$$

stetig auf $[a,b]$, differenzierbar in (a,b) und $h(a) = h(b) = 0$, also gibt es $\theta \in (a,b)$ mit

$$h'(\theta) = f'(\theta)(g(b) - g(a)) - g'(\theta)(f(b) - f(a)) = 0.$$ □

Folgerung 7.2.6. *Ist $f\colon [a,b] \to \mathbb{R}$ stetig und in (a,b) differenzierbar mit $f'(x) = 0$ für alle $x \in (a,b)$, dann ist f konstant.*

Beweis. Ist $x \in (a,b]$ dann gibt es $\theta \in (a,x)$ mit

$$0 = f'(\theta) = \frac{f(x) - f(a)}{x - a} \implies f(x) = f(a).$$ □

Satz 7.2.7. *Sei $a < b$ und $f\colon [a,b] \to \mathbb{R}$ stetig und in (a,b) differenzierbar. Gilt für alle $x \in (a,b)$, dass*

$$\begin{cases} f'(x) > 0, \\ f'(x) \geq 0, \\ f'(x) \leq 0, \\ f'(x) < 0 \end{cases} \quad \textit{dann ist} \quad \begin{cases} f \textit{ streng monoton wachsend}, \\ f \textit{ monoton wachsend}, \\ f \textit{ monoton fallend}, \\ f \textit{ streng monoton fallend}. \end{cases}$$

Beweis. Sei $f'(x) > 0$ für alle $x \in (a, b)$. Wäre f nicht streng monoton wachsend, gäbe es $\alpha < \beta$ in $[a, b]$ mit $f(\alpha) \geq f(\beta)$, also $f(\beta) - f(\alpha) \leq 0$ und nach dem Mittelwertsatz ein $x \in (a, b)$ mit

$$f'(x) = \frac{f(\beta) - f(\alpha)}{\beta - \alpha} \leq 0,$$

was ein Widerspruch ist. Die anderen Fälle werden analog gezeigt. □

Satz 7.2.8. *Sei $f: (a, b) \to \mathbb{R}$ differenzierbar und f' in $x \in (a, b)$ differenzierbar. Es gilt:*

$$f'(x) = 0,\ f''(x) > 0 \implies f \text{ hat strenges lokales Minimum in } x,$$
$$f'(x) = 0,\ f''(x) < 0 \implies f \text{ hat strenges lokales Maximum in } x.$$

Beweis. Ist $f''(x) > 0$ dann gibt es $\varepsilon > 0$ sodass $\forall \xi$ mit $|\xi - a| < \varepsilon$ gilt

$$\frac{f'(\xi) - f'(x)}{\xi - x} > 0.$$

Damit ist $f'(\xi) < 0$ für $x - \varepsilon < \xi < x$ und $f'(\xi) > 0$ für $x < \xi < x + \varepsilon$, also ist f streng monoton fallend in $[x - \varepsilon, x]$ und streng monoton wachsend in $[x, x + \varepsilon]$, also ist x ein strenges lokales Minumum; der Beweis im Fall $f''(x) < 0$ ist analog. □

Konvexität

Definition 7.2.9. Sei $I \subseteq \mathbb{R}$ ein Intervall. Eine Funktion $f: I \to \mathbb{R}$ heißt *konvex*, wenn

$$\forall x, y \in I\ \forall t \in [0, 1] : f(tx + (1 - t)y) \leq tf(x) + (1 - t)f(y)$$

und *konkav* wenn $-f$ konvex ist, d.h.

$$\forall x, y \in I\ \forall t \in [0, 1] : f(tx + (1 - t)y) \geq tf(x) + (1 - t)f(y)$$

Anschaulich gesprochen ist eine Funktion konvex (bzw. konkav), wenn die Verbindungsstrecke zwischen zwei Punkten des Graphens immer über (bzw. unter) dem Graphen liegt.

Satz 7.2.10. *Sei $I \subseteq \mathbb{R}$ offenes Intervall und $f: I \to \mathbb{R}$ zweimal differenzierbar. Dann ist f genau dann konvex, wenn $f''(x) \geq 0$ für alle $x \in I$ ist.*

Beweis. Angenommen f sei konvex aber $f''(x_0) < 0$ für ein $x_0 \in I$. Dann ist $\varphi(x) = f(x) - f'(x_0)(x - x_0)$ konvex, zweimal differenzierbar und $\varphi'(x_0) = 0$, $\varphi''(x_0) = f''(x_0) < 0$, also hätte φ ein strenges lokales Maximum, was im Widerspruch zur Konvexität steht.

Ist umgekehrt $f''(x) \geq 0$ für alle $x \in I$ dann ist f' monoton wachsend. Sei $x < y$ in I, $0 < t < 1$ und $w = tx + (1 - t)y$. Dann gibt es nach dem Mittelwertsatz $\theta_1 \in (x, w)$ und $\theta_2 \in (w, y)$ mit

$$\frac{f(w) - f(x)}{(1 - t)(y - x)} = \frac{f(w) - f(x)}{w - x} = f'(\theta_1) \leq f'(\theta_2) = \frac{f(y) - f(w)}{y - w} = \frac{f(y) - f(w)}{t(y - x)}$$

und damit ist $f(w) \leq tf(x) + (1 - t)f(y)$. □

6. 12. 2024

Die Regeln von de l'Hospital

Satz 7.2.11 (Regeln von de l'Hospital)**.** *Seien $f, g\colon (a,b) \to \mathbb{R}$ mit $-\infty \leq a < b \leq \infty$ differenzierbare Funktionen. Es gelte $g'(x) \neq 0$ für alle $a < x < b$ und es existiere*

$$\lim_{x \nearrow b} \frac{f'(x)}{g'(x)} = c \in \mathbb{R}.$$

Dann gilt:

(i) Falls $\lim_{x\nearrow b} g(x) = \lim_{x \nearrow b} f(x) = 0$, ist $\lim\limits_{x\nearrow b} \dfrac{f(x)}{g(x)} = c$.

(ii) Falls $\lim_{x\nearrow b} g(x) = \pm\infty$, ist $\lim\limits_{x\nearrow b} \dfrac{f(x)}{g(x)} = c$.

Analoge Aussagen gelten für $x \searrow a$.

Beweis. Zuerst bemerken wir, dass g injektiv ist, denn wäre $g(x) = g(y)$ für $x, y \in (a,b)$ mit $x \neq y$, dann hätte g' nach Satz 7.2.4 eine Nullstelle zwischen x und y, was wir ausgeschlossen haben.

Ebenso muss g streng monoton sein, denn da g' keine Nullstelle hat und stetig ist, hat es nach dem Zwischenwertsatz (Satz 5.2.4) überall das gleiche Vorzeichen.

Ohne Beschränkung der Allgemeinheit sei g nun streng mononton wachsend (sonst ersetzen wir g durch $-g$). Im Fall (i) heißt das $g(x) \neq 0$ für alle $x \in (a,b)$, im Fall (ii) $g(x) \neq 0$ für alle $x \geq x_0$ und irgendein $x_0 \in (a,b)$. Ersetzt man a durch x_0 kann man in beiden Fällen $g(x) \neq 0$ für alle x annehmen.

Ist $\varepsilon > 0$ beliebig, dann gibt es nach Voraussetzung ein $x_0 \in (a,b)$ sodass $\left|\frac{f'(\theta)}{g'(\theta)} - c\right| < \varepsilon$ für alle $\theta \in (x_0, b)$. Für $x_0 < x < y < b$ gibt es nach dem verallgemeinerten Mittelwertsatz ein $\theta \in (x,y)$ mit

$$\frac{f(y) - f(x)}{g(y) - g(x)} = \frac{f'(\theta)}{g'(\theta)}, \text{ also auch } \left|\frac{f(y) - f(x)}{g(y) - g(x)} - c\right| < \varepsilon.$$

Für (i) lässt man nun $y \to b$ gehen, dann folgt $|f(x)/g(x) - c| < \varepsilon$ für $x > x_0$ und da ε beliebig war auch die Behauptung.

Für (ii) nehmen wir an, dass $f(x)/g(x)$ nicht gegen c konvergiert. Dann gibt es eine streng monoton gegen b wachsende Folge (x_n) in (a,b) so dass $f(x_n)/g(x_n)$ nicht gegen c konvergiert, also kann man nach Übergang zu einer Teilfolge annehmen, dass für ein $\varepsilon > 0$

$$\left|\frac{f(x_n)}{g(x_n)} - c\right| \geq \varepsilon. \tag{7.2}$$

Wegen $g(x_n) \to \infty$ können wir durch Übergang zu einer weiteren Teilfolge annehmen, dass $g(x_{n+1}) \geq n g(x_n)$ und $g(x_{n+1}) \geq n|f(x_n)|$ für alle n gilt. Damit folgt

$$\frac{\frac{f(x_{n+1})}{g(x_{n+1})} - \frac{f(x_n)}{g(x_{n+1})}}{1 - \frac{g(x_n)}{g(x_{n+1})}} = \frac{f(x_{n+1}) - f(x_n)}{g(x_{n+1}) - g(x_n)} \to c.$$

Wegen $f(x_n)/g(x_{n+1}) \to 0$ und $g(x_n)/g(x_{n+1}) \to 0$ hieße das aber $f(x_{n+1})/g(x_{n+1}) \to c$, was (7.2) widerspricht. $\square$

Beispiel.

(i) Für $\alpha > 0$ gilt

$$\lim_{x\to\infty} \frac{\log x}{x^\alpha} = \lim_{x\to\infty} \frac{\frac{1}{x}}{\alpha x^{\alpha-1}} = \lim_{x\to\infty} \frac{1}{\alpha x^\alpha} = 0.$$

(ii) Es gilt

$$\lim_{x\to 0}\left(\frac{1}{\sin x} - \frac{1}{x}\right) = \lim_{x\to 0} \frac{x - \sin x}{x \sin x} = \lim_{x\searrow 0} \frac{x - \sin x}{x \sin x} = \lim_{x\searrow 0} \frac{1 - \cos x}{\sin x + x\cos x}$$
$$= \lim_{x\searrow 0} \frac{\sin x}{\cos x + \cos x - x \sin x} = 0.$$

7.3 Die Taylorentwicklung

Satz 7.3.1 (Taylorsche Formel). *Sei I ein Intervall, $x_0 \in I$ und $f\colon I \to \mathbb{R}$ eine $(n+1)$-mal stetig differenzierbare Funktion, wobei $n \in \mathbb{N}_0$ ist. Dann gilt für jedes $x \in I$*

$$f(x) = f(x_0) + \frac{f'(x_0)}{1!}(x - x_0) + \ldots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n + R_n(x)$$

mit Lagrange-Restglied

$$R_n(x) = R_n(f, x_0)(x) = f^{(n+1)}(\Theta)\frac{(x - x_0)^{n+1}}{(n+1)!}$$

wobei $\Theta = x_0 + \theta(x - x_0)$ für ein $0 < \theta < 1$ ist (d.h. Θ liegt zwischen x_0 und x).

Der Ausdruck

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

heißt Taylorpolynom n-ten Grades von f mit Entwicklungspunkt x_0*, und $R_n(f, x_0) = f - T_n(f, x_0)$ das* Restglied n-ter Ordnung von f in x_0.

Beweis. Für $x = x_0$ ist der Satz offenbar richtig; für $x \in I$ mit $x \neq x_0$ sei

$$F(t) = f(x) - f(t) - f'(t)(x - t) - \ldots - f^{(n)}(t)\frac{(x - t)^n}{n!},$$
$$G(t) = \frac{(x - t)^{n+1}}{(n+1)!} \qquad \text{für } t \in I.$$

Man überprüft einfach, dass

$$F'(t) = -f^{(n+1)}(t)\frac{(x - t)^n}{n!}, \quad G'(t) = -\frac{(x - t)^n}{n!} \quad \forall t \in I.$$

Mit dem verallgemeinerten Mittelwertsatz erhält man

$$\frac{F(x) - F(x_0)}{G(x) - G(x_0)} = \frac{F'(\Theta)}{G'(\Theta)} = f^{(n+1)}(\Theta) \tag{7.3}$$

für ein geeignetes Θ zwischen x und x_0. Für $t = x$ ist $F(x) = G(x) = 0$ und für $t = x_0$ ist

$$F(x_0) = f(x) - f(x_0) - f'(x_0)(x - x_0) - \ldots - f^{(n)}(x_0)\frac{(x - x_0)^n}{n!}$$
$$G(x_0) = \frac{(x - x_0)^{n+1}}{(n + 1)!}.$$

Aus (7.3) folgt $F(x_0) = G(x_0)f^{(n+1)}(\Theta)$ und damit die Behauptung. □

Beispiel 7.3.2. *Die Taylorentwicklung von e^x um $x_0 = 0$ ist*

$$e^x = 1 + x + \frac{x^2}{2} + \ldots + \frac{x^n}{n!} + \frac{x^{n+1}}{(n + 1)!}e^{\theta x}$$

mit $0 < \theta < 1$. Will man die Exponentialfunktion auf $[-1, 1]$ durch ein Polynom auf 2 Nachkommastellen genau approximieren, erhält man mit

$$p(x) = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \frac{x^4}{24} + \frac{x^5}{120}$$

für das Restglied die Abschätzung

$$|e^x - p(x)| = \frac{x^6}{720}e^{\theta x} < \frac{e}{720} < 0,0038.$$

7.4 Kurvendiskussionen

Um die Eigenschaften und das Verhalten einer gegebenen Funktion zu verstehen ist die Betrachtung ihres Graphen oft sehr hilfreich. Der Graph alleine reicht jedoch nicht: entweder sieht man ihm wesentliche Eigenschaften nicht an, oder sie lassen sich nur ungefähr ablesen. Deshalb führt man eine *Kurvendiskussion* durch, indem man das Wesen der Funktion analytisch bestimmt und dann anhand ihres Graphen übersichtlich darstellt. Dabei stellt man sich folgende Fragen:

- Wie lautet der *Definitionsbereich* der Funktion?
- Ist die Funktion *gerade*, *ungerade*, oder *periodisch*?
- Ist die Funktion überall *stetig* oder gibt es *Unstetigkeitsstellen*?
- Hat die Funktion *Pole*, d.h. Stellen an denen der Funktionwert gegen $\pm\infty$ divergiert?
- Falls die Funktion in *Randpunkten* nicht definiert ist: wie ist der Grenzwert dort?
- Was sind die *Nullstellen*? Wo ist die Funktion *positiv* bzw. *negativ?*
- Was sind die *Nullstellen der 1. Ableitung*? Wo ist die Funktion *monoton wachsend* bzw. *monoton fallend*?
- Was sind die *Nullstellen der 2. Ableitung*? Wo ist die Funktion *konvex* bzw. *konkav*? Wo sind die *Wendepunkte*, d.h. die Punkte an denen die 2. Ableitung ihr Vorzeichen wechselt?
- Was sind die *lokalen Extremstellen*, und welcher *Art* sind sie? Sind es *globale Extrema*?

Beispiel 7.4.1. *Wir führen für die Funktion $f(x) := (1+x)\sqrt{1-x^2}$ eine Kurvendiskussion durch.*

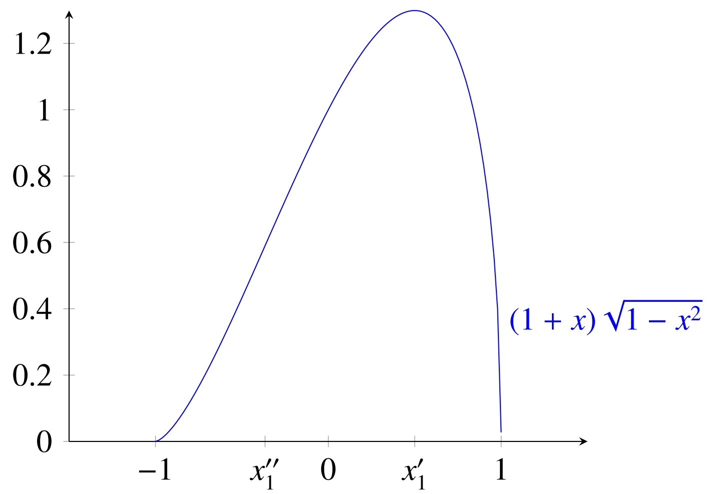
- *Die Funktion ist für $|x| \leq 1$ definiert und weder gerade noch ungerade.*
- *f ist überall stetig und besitzt keine Pole.*
- *Die Nullstellen von f sind $x_1 = -1$ und $x_2 = 1$, zwischen ihnen ist f positiv.*
- *Für $|x| < 1$ ist*

$$\begin{aligned} f'(x) &= \sqrt{1-x^2} + (1-x)\frac{-2x}{2\sqrt{1-x^2}} = \frac{1-x^2-x-x^2}{\sqrt{1-x^2}} \\ &= -\frac{2x^2+x-1}{\sqrt{1-x^2}} = 0 \iff x = -\frac{1}{4} \pm \sqrt{\frac{1}{16}+\frac{8}{10}} = -\frac{1}{4} \pm \frac{3}{4}. \end{aligned}$$

  Die einzige Nullstelle von f' ist $x_1' = 1/2$. Auf $(-1, 1/2)$ ist f' positiv, also f streng monoton wachsend, auf $(1/2, 1)$ ist f' negativ, also f streng monoton fallend.
- *Die Zweite Ableitung ist*

$$f''(x) = \frac{2x^3-3x-1}{(1-x^2)\sqrt{1-x^2}}.$$

- *Die Nullstellen -1 von f'' errät man, nach Polynomdivision durch $(x+1)$ findet man $(1+\sqrt{3})/2$ und $(1-\sqrt{3})/2$. Von diesen ist nur $x_1'' = (1-\sqrt{3})/2$ in $(-1, 1)$. Auf $(-1, x_1'')$ ist $f'' > 0$, also f konvex, und auf $(x_1'', 1)$ ist $f'' < 0$, also f konkav.*
- *In x_1' ist ein lokales Maximum von f mit Wert $f(x_1') = 3\sqrt{3}/4$ und es gibt keine anderen lokalen Extrema im Inneren des Definitionsbereiches. Am Rand befinden sich globale Minima. x_1' ist sogar globales Maximum.*
- *x_1'' ist der einzige Wendepunkt von f.*
- *Es gilt $\lim_{x\to 1-} f(x) = -\infty$ und (mit de l'Hospital) $\lim_{x\to -1+} f'(x) = 0$.*
- *Skizze:*



Kapitel 8

Integralrechnung

Lernziele

- Treppenfunktionen integrieren
- Elementare Eigenschaften des Integrals
- Gleichmäßiger Abstand
- Stetige oder monotone Funktionen durch Treppenfunktionen annähern
- Definitions des Integrales wiedergeben
- Hauptsatz der Differential- und Integralrechnung angeben
- Mittelwertsatz der Integralrechnung
- Wichtigste Stammfunktionen kennen
- Partielle Integration und Substitution anwenden
- Konkrete Integrale berechnen
- Rationale Funktionen integrieren

8.1 Treppenfunktionen und Integrierbarkeit

Definition 8.1.1. Eine *Treppenfunktion* auf einem Intervall $[a, b]$ ist eine Funktion $\varphi\colon [a, b] \to \mathbb{R}$, welche nur endlich viele Werte annimmt, d.h. für die es Zahlen $a = t_0 < t_1 < \ldots < t_n = b$ gibt, sodass φ auf jedem Teilintervall (t_j, t_{j+1}) $(j = 0, \ldots, n-1)$ konstant ist.

Die Menge $Z = \{t_0, t_1, \ldots, t_n\}$ mit $t_0 = a$, $t_n = b$ und $t_j < t_{j+1}$ heißt *Zerlegung* des Intervalls $[a, b]$. Eine zweite Zerlegung V von $[a, b]$ heißt *Verfeinerung* von Z, wenn $V \supseteq Z$.

Beispiel 8.1.2.

(i) Ist I ein beschränktes Intervall, dann ist $\mathbb{1}_I$ (die Indikatorfunktion von I) eine Treppenfunktion.
(ii) Für $\lambda \in \mathbb{R}$ ist $\lambda\mathbb{1}_I$ eine Treppenfunktion.
(iii) Für $I = [0, 2]$ und $J = [1, 3)$ ist $3\mathbb{1}_I - 2\mathbb{1}_J$ eine Treppenfunktion.

Definition 8.1.3. Zwei Treppenfunktionen φ, ψ auf $[a, b]$ heißen *fast überall gleich*, wenn $\varphi(x) = \psi(x)$ für alle bis auf endliche viele x gilt.

Es folgt, dass jede Treppenfunktion wie in Definition 8.1.1 eine Darstellung

$$\varphi = \sum_{j=1}^{n} c_j \mathbb{1}_{I_j} \quad \text{fast überall}$$

mit $n \in \mathbb{N}_0$, $I_j = (t_{j-1}, t_j)$ und $c_j \in \mathbb{R}$ für $j = 1 \ldots n$ hat.

Lemma 8.1.4. Sind φ, ψ Treppenfunktionen auf $[a, b]$, dann sind für $\lambda, \mu \in \mathbb{R}$ auch $|\varphi|$, $\lambda\varphi + \mu\psi$ und $\varphi \cdot \psi$ Treppenfunktionen auf $[a, b]$.

Der Beweis ist klar, da man die selbe Zerlegung wählen kann. Treppenfunktionen auf $[a, b]$ bilden also einen Vektorraum, den man mit $T[a, b]$ bezeichnet.

Definition 8.1.5. Ist φ eine Treppenfunktion auf $[a, b]$ zur Zerlegung $a = t_0 < \ldots < t_k = b$ und $\varphi(x) = c_j \in \mathbb{R}$ auf (t_{j-1}, t_j), d.h.

$$\varphi = \sum_{j=1}^{n} c_j \mathbb{1}_{I_j}$$

mit $I_j = (t_{j-1}, t_j)$, dann definiert man das *Integral* von φ auf $[a, b]$ durch

$$\int_a^b \varphi(x)\,\mathrm{d}x = \sum_{j=1}^{n} c_j(t_j - t_{j-1}).$$

Damit diese Definition Sinn macht, muss man sicherstellen, dass das Integral einer Treppenfunktion nicht von der Wahl der Zerlegung abhängt. Sind zwei Zerlegungen Z_1 und Z_2 gegeben, ist $V = Z_1 \cup Z_2$ eine gemeinsame Verfeinerung, die man aus Z_1 bzw. Z_2 durch schrittweises Hinzufügen jeweils eines Punktes erreichen kann. Wir zeigen also, dass sich das Integral bzgl. einer Zerlegung

$$Z = \{t_0, t_1, \ldots, t_n\},$$

welches wir mit $\int_Z \varphi(x)\,\mathrm{d}x$ bezeichnen, nicht ändert, wenn wir einen Punkt t mit $t_{\nu-1} < t < t_\nu$ für ein $1 \le \nu \le n$ hinzufügen:

$$\begin{aligned}
\int_Z \varphi(x)\,\mathrm{d}x &= \sum_{j=1}^{n} c_j(t_j - t_{j-1}) \\
&= \sum_{j=1}^{\nu-1} c_j(t_j - t_{j-1}) + \underbrace{c_\nu(t_\nu - t_{\nu-1})}_{=c_\nu(t-t_{\nu-1})+c_\nu(t_\nu-t)} + \sum_{j=\nu+1}^{n} c_j(t_j - t_{j-1}) \\
&= \int_{Z\cup\{t\}} \varphi(x)\,\mathrm{d}x.
\end{aligned}$$

Bemerkung. Mitunter schreibt man auch kurz $\int_a^b \varphi$ statt $\int_a^b \varphi(x)\,\mathrm{d}x$.

Im Folgenden schreiben wir $\varphi \le \psi$ für zwei Funktionen $\varphi \le \psi$, wenn sie denselben Definitionsbereich D haben und die Ungleichung punktweise gilt, d.h. $\varphi(x) \le \psi(x)$ für alle $x \in D$.

Satz 8.1.6. *Seien φ, ψ Treppenfunktionen auf $[a, b]$ und $\lambda \in \mathbb{R}$. Dann gilt:*
(i) $\int_a^b (\varphi + \psi) = \int_a^b \varphi + \int_a^b \psi$.
(ii) $\int_a^b \lambda\varphi = \lambda \int_a^b$
(iii) Aus $\varphi \leq \psi$ *folgt* $\int_a^b \varphi \leq \int_a^b \psi$.
(iv) $\left|\int_a^b \varphi\right| \leq \int_a^b |\varphi| \leq (b - a) \max_{x\in[a,b]} |\varphi(x)|$.

Beweis. Aussagen (i)–(iii) sieht man sofort, wenn man für φ und ψ die selbe Zerlegung wählt. Für (iv) betrachtet man

$$\left|\sum_{j=1}^{n} c_j(t_j - t_{j-1})\right| \leq \sum_{j=1}^{n} \left|c_j\right|\left|t_j - t_{j-1}\right| \leq \max_{x\in[a,b]} |\varphi(x)| \cdot (b - a).$$ □

8.2 Der gleichmäßige Abstand

Definition 8.2.1.

(i) Eine Funktion $f : D \to \mathbb{R}$ heißt *beschränkt*, wenn $f(D)$ beschränkt ist.
(ii) Für beschränkte Funktionen $f : [a, b] \to \mathbb{R}$ definieren wir die *Supremumsnorm* durch

$$\|f\| := \sup_{x\in[a,b]} |f(x)|.$$

(iii) Sind f, g zwei beschränkte Funktionen auf $[a, b]$, nennen wir $\|f - g\|$ den *gleichmäßigen Abstand* von f und g.

Satz 8.2.2. *Für beschränkte Funktionen* $f, g : [a, b] \to \mathbb{R}$ *gilt*
(i) $\|f\| \geq 0$, $\|f\| = 0 \iff f = 0$,
(ii) $\|\lambda f\| = |\lambda| \cdot \|f\|$ *für alle* $\lambda \in \mathbb{R}$,
(iii) $\|f + g\| \leq \|f\| + \|g\|$ *(Dreiecksungleichung).*

Beweis. (i) und (ii) sind klar, (iii) folgt aus

$$f(x) + g(x) \leq |f(x)| + |g(x)| \leq \|f\| + \|g\|$$

für alle $x \in [a, b]$. □

Bemerkung 8.2.3. Für Die Supremumsnorm gilt auch die *umgekehrte Dreiecksungleichung* $\big|\|f\| - \|g\|\big| \leq \|f - g\|$.

Satz 8.2.4. *Ist f stetig auf $[a, b]$, so gibt es zu jedem $\varepsilon > 0$ ein $\varphi \in T[a, b]$ mit $\|f - g\| < \varepsilon$. Insbesondere gibt es eine Folge von Treppenfunktionen φ_n auf $[a, b]$ mit $\|f - \varphi_n\| \to 0$ für $n \to \infty$.*

Beweis. Sei $\varepsilon > 0$. Da f gleichmäßig stetig ist (Satz 5.2.11) gibt es $\delta > 0$ mit

$$x, y \in [a, b],\ |x - y| < \delta \implies |f(x) - f(y)| < \varepsilon.$$

Sei $n \in \mathbb{N}$ mit $(b - a)/n < \delta$ und $t_j = a + j(b - a)/n$ für $j = 0, \ldots, n$. Dann ist $\{t_0, \ldots, t_n\}$ eine Zerlegung von $[a, b]$ und

$$\varphi(x) = \begin{cases} f(t_k) & t_{k+1} \leq x < t_k \text{ mit } k = 1, \ldots, n \\ f(b) & x = b. \end{cases}$$

eine Treppenfunktion auf $[a, b]$. Für $x \in (a, b)$ gibt es $k \in \{1, \ldots, n\}$ mit $x \in [t_{k-1}, t_k)$, d.h. $|x - t_k| < (b - a)/n < \delta$ und $|f(x) - \varphi(x)| = |f(x) - f(t_k)| < \varepsilon$, also ist $\|f - \varphi\| < \varepsilon$. □

8.3 Integrierbare Funktionen

Definition 8.3.1. Eine beschränkte Funktion $f\colon [a, b] \to \mathbb{R}$ heißt *integrierbar*, wenn es eine Folge von Treppenfunktionen $\varphi_n \in T[a, b]$ mit $\|f - \varphi_n\| \to 0$ gibt.

Satz 8.3.2. *Für integrierbare Funktionen f existiert der Grenzwert*

$$\int_a^b f := \lim_{n\to\infty} \int_a^b \varphi_n.$$

und ist unabhängig von der approximierenden Folge (φ_n). Die Zahl $\int_a^b f$ heißt (Cauchy-Riemannsches) Integral von f über $[a, b]$, *f heißt der* Integrand.

Bemerkung 8.3.3. Dieses Integral, welches auch *Cauchy-Integral* oder *Regelintegral* genannt wird, ist etwas einfacher als das häufig verwendete *Riemann-Integral*, welches die Integration einer etwas größeren Klasse von Funktionen erlauben würde.

Beweis. Für $\varepsilon > 0$ wählen wir $n_0 \in \mathbb{N}$ mit $\|f - \varphi_n\| < \varepsilon$ für alle $n \geq n_0$. Dann gilt für $n, m \geq n_0$

$$\|\varphi_n - \varphi_m\| = \|\varphi_n - f + f - \varphi_m\| \leq \|\varphi_n - f\| + \|f - \varphi_m\| < 2\varepsilon,$$

also

$$\left|\int_a^b \varphi_n - \int_a^b \varphi_m\right| = \left|\int_a^b (\varphi_n - \varphi_m)\right| \leq (b - a)\|\varphi_n - \varphi_m\| < 2(b - a)\varepsilon.$$

Die Folge $(\int \varphi_n)$ ist also eine Cauchy-Folge und besitzt einen Grenzwert.

Ist ψ_n eine zweite Folge mit $\|f - \psi_n\| \to 0$, dann gilt

$$\|\varphi_n - \psi_n\| \leq \|\varphi_n - f\| + \|f - \psi_n\| \to 0,$$

$$\left|\int_a^b \varphi_n - \int_a^b \psi_n\right| = \left|\int_a^b (\varphi_n - \psi_n)\right| \leq (b - a)\|\varphi_n - \psi_n\| \to 0,$$

$$\lim_{n\to\infty} \int_a^b \varphi_n = \lim_{n\to\infty} \int_a^b \psi_n.$$

□

Bemerkung 8.3.4.

- Für integrierbare Funktionen ist die Schreibweise
$$\int_a^b f(x)\,\mathrm{d}x$$
im Allgemeinen zu empfehlen, vor allem da nicht jede Funktion einen Namen (hier: f) hat, d.h. die Bezeichnung
$$\int_a^b y\,x^2 \sin(x)\,\mathrm{d}x$$
macht Sinn; daran sieht man auch, dass x die Integrationsvariable ist und y eine Konstante.
- Für $a > b$ vereinbaren wir
$$\int_a^b f(x)\,\mathrm{d}x = -\int_b^a f(x)\,\mathrm{d}x.$$
- Ist f über $[a, b]$ integrierbar und $[c, d]$ ein Teilintervall von $[a, b]$, dann ist auch die Einschränkung von f auf $[c, d]$ integrierbar.
- Es gilt
$$\int_a^c f = \int_a^b f + \int_b^c f.$$

Satz 8.3.5. *Stetige Funktionen sind integrierbar.*

Dies folgt direkt aus Satz 8.2.4. 13. 12. 2024

Satz 8.3.6.

(i) Mit f und g ist auch $\lambda f + \mu g$ für $\lambda, \mu \in \mathbb{R}$ über $[a, b]$ integrierbar und es gilt
$$\int_a^b (\lambda f + \mu g) = \lambda \int_a^b f + \mu \int_a^b g. \qquad \textit{(Linearität)}$$

(ii) Sind f, g auf $[a, b]$ integrierbar und $f(x) \leq g(x)$ für alle $x \in [a, b]$, dann gilt
$$\int_a^b f \leq \int_a^b g. \qquad \textit{(Monotonie)}$$

(iii) Mit f ist auch $|f|$ auf $[a, b]$ integrierbar. Aus $|f(x)| \leq M$ für alle $x \in [a, b]$ folgt
$$\left|\int_a^b f\right| \leq \int_a^b |f| \leq (b - a)M.$$

(iv) Mit f und g ist auch $f \cdot g$ über $[a, b]$ integrierbar.

Beweis. (i) Sind φ_n, ψ_n Folgen in $T[a,b]$ mit $\|f-\varphi_n\| \to 0$ und $\|g-\psi_n\| \to 0$, dann ist

$$\begin{aligned}\|(\lambda f+\mu g)-(\lambda\varphi_n+\mu\psi_n)\| &= \|\lambda(f-\varphi_n)+\mu(g-\psi_n)\|\\ &\le |\lambda|\cdot\|f-\varphi_n\|+|\mu|\cdot\|g-\psi_n\| \to 0,\end{aligned}$$

also ist $\lambda f+\mu g$ integrierbar und

$$\int_a^b(\lambda f+\mu g)=\lim_{n\to\infty}\int_a^b(\lambda\varphi_n+\mu\psi_n)=\lim_{n\to\infty}\left(\lambda\int_a^b\varphi_n+\mu\int_a^b\psi_n\right)=\lambda\int_a^b f+\mu\int_a^b g.$$

(ii) Nach (i) ist $h=g-f$ integrierbar und $\int_a^b h=\int_a^b g-\int_a^b f$. Wir zeigen, dass aus $h\ge 0$ die Ungleichung $\int_a^b h\ge 0$ folgt. Sei (φ_n) eine Folge in $T[a,b]$ mit $\|\varphi_n-h\|\to 0$, dann sind die $|\varphi_n|$ Treppenfunktionen, für welche nach der umgekehrten Dreiecksungleichung

$$\big||\varphi_n(x)|-h(x)\big|\le|\varphi_n(x)-h(x)|\le\|\varphi_n-h\|\quad\forall x\in[a,b]$$

gilt, also $\big\|h-|\varphi_n|\big\|\to 0$ und

$$\int_a^b h=\lim_{n\to\infty}\int_a^b|\varphi_n|\ge 0.$$

(iii) Sei (φ_n) eine Folge in $T[a,b]$ mit $\|f-\varphi_n\|\to 0$. Für jedes $x\in[a,b]$ hat man

$$\big||f(x)|-|\varphi_n(x)|\big|\le|f(x)-\varphi_n(x)|\le\|f-\varphi_n\|$$

und durch Übergang zum Supremum auch

$$\big\||f|-|\varphi_n|\big\|\le\|f-\varphi_n\|,$$

also ist $|f|$ auch integrierbar. Wegen $\pm f\le|f(x)|\le M$ folgt die Behauptung mit (ii) aus

$$\pm\int_a^b f\le\int_a^b|f|\le\int_a^b M\mathbb{1}_{[a,b]}=M(b-a).$$

(iv): Für $\|f-\varphi_n\|\to 0$ und $\|g-\psi_n\|\to 0$ hat man

$$\begin{aligned}\|fg-\varphi_n\psi_n\| &\le\|(f-\varphi_n)g+\varphi_n(g-\psi_n)\|\\ &\le\|f-\varphi_n\|\cdot\|g\|+\|\varphi_n\|\cdot\|g-\psi_n\|\to 0,\end{aligned}$$

da $\big|\|f\|-\|\varphi_n\|\big|\le\|f-\varphi_n\|$ ist, also $\|\varphi_n\|$ eine beschränkte reelle Folge. □

Definition 8.3.7. Eine Funktion $f\colon[a,b]\to\mathbb{R}$ heißt *stückweise stetig*, wenn es für $[a,b]$ eine Zerlegung $a=t_0<t_1<\ldots<t_n=b$ gibt, sodass f auf (t_{j-1},t_j) für alle $j=1\ldots n$ stetig ist und die einseitigen Grenzwerte $f(t_{j-1}+)$ und $f(t_j-)$ existieren.

Für solche f hat die Einschränkung von f auf (t_{j-1},t_j) eine eindeutige stetige Fortsetzung f_j auf $[t_{j-1},t_j]$ $(j=1\ldots n)$. Diese ist integrierbar, und setzt man alles Stücke zusammen erhält man:

Satz 8.3.8. *Stückweise stetige Funktionen sind integrierbar und es gilt*

$$\int_a^b f = \sum_{j=1}^{n} \int_{t_{j-1}}^{t_j} f_j.$$

Dies sieht man, indem man aus den approximierenden Treppenfunktionen für die f_j solche für f konstruiert.

Satz 8.3.9. *Monotone Funktionen $f\colon [a,b] \to \mathbb{R}$ sind integrierbar.*

Beweis. Ohne Beschränkung der Allgemeinheit sei f monoton wachsend, denn ist f monoton fallend ist $-f$ monoton wachsend und integrierbar und damit auch f.

Wir unterteilen $[f(a), f(b)]$ in n Teilintervalle der Länge $h = (f(b)-f(a))/n$ mit Teilungspunkten $y_j = f(a)+jh$, $j = 0\dots n$ und setzen $a_0 = a$ sowie $a_j = \sup\{x \in [a,b] \mid f(x) \leq y_j\}$ für $j = 1\dots n$. Es gilt $a = a_0 \leq a_1 \leq a_2 \leq \dots \leq a_n = b$ und $f(x) \in (y_{j-1}, y_j]$ für $x \in (a_{j-1}, a_j)$. Für die Treppenfunktion

$$\varphi_n = f(a_0)\mathbb{1}_{[a_0,a_1]} + \sum_{j=1}^{n} y_j \mathbb{1}_{(a_{j-1},a_j]}$$

gilt $\|f - \varphi_n\| \leq (f(b) - f(a))/n$, also $\|f - \varphi_n\| \to 0$ für $n \to \infty$. □

Beispiel 8.3.10. *Die Funktion $f(x) = \sin(1/x)$ für $x \in (1,0]$ und $f(0)$ beliebig ist nicht integrierbar.*

Der Vollständigkeit halber erwähnen wir, ohne Beweis, folgende Charakterisierung von integrierbaren Funktionen:

Satz 8.3.11. *Eine Funktion $f\colon [a,b] \to \mathbb{R}$ ist integrierbar genau dann, wenn für jedes $x \in (a,b)$ die einseitigen Grenzwerte $f(x+)$ und $f(x-)$ sowie $f(a+)$ und $f(b-)$ existieren.*

Funktionen mit dieser Eigenschaft nennt man auch *sprungstetig*. 19. 2. 2024

8.4 Der Hauptsatz der Differential- und Integralrechnung

Definition 8.4.1. Ist $f\colon [a,b] \to \mathbb{R}$ integrierbar und $x_0 \in [a,b]$, dann heißt

$$x \mapsto U(x) = \int_{x_0}^{x} f(t)\,\mathrm{d}t, \qquad x \in [a,b]$$

ein *unbestimmtes Integral* von f.

Satz 8.4.2. *Jedes unbestimmte Integral einer über $[a, b]$ integrierbaren Funktion ist dort stetig.*

Beweis. Für $x, y \in [a, b]$ ist

$$U(x) - U(y) = \int_{x_0}^{x} f - \int_{x_0}^{y} f = \int_{y}^{x} f$$

und daraus ergibt sich

$$|U(x) - U(y)| \leq \|f\| \cdot |x - y|. \qquad \square$$

Definition 8.4.3. Eine Funktion $F\colon [a, b] \to \mathbb{R}$ heißt *Stammfunktion* von $f\colon [a, b] \to \mathbb{R}$, wenn F differenzierbar ist und

$$F'(x) = f(x) \quad \forall x \in [a, b].$$

Bemerkung 8.4.4. Ist F eine Stammfunktion von f, dann erhält man alle Stammfunktionen von f aus $G = F + c$ für alle $c \in \mathbb{R}$: Klarerweise ist $G'(x) = F'(x) = f$, d.h. G ist eine Stammfunktion von f; ist umgekehrt G eine Stammfunktion von f dann gilt $(G - F)' = f - f = 0$, also ist $G - F$ nach Folgerung 7.2.6 konstant.

Satz 8.4.5. *Sei $f\colon [a, b] \to \mathbb{R}$ integrierbar und f stetig in $x_0 \in [a, b]$. Dann ist das unbestimmte Integral*

$$F(x) = \int_{a}^{x} f(t)\,\mathrm{d}t$$

in x_0 differenzierbar und es gilt $F'(x_0) = f(x_0)$.

Beweis. Für $\varepsilon > 0$ gibt es, da f stetig in x_0 ist, ein $\delta > 0$ mit $|f(t) - f(x_0)| < \varepsilon$ für alle $t \in I$ mit $|t - x_0| < \delta$. Damit folgt für $x \in I$ mit $0 < |x - x_0| < \delta$

$$\frac{F(x) - F(x_0)}{x - x_0} = \frac{1}{x - x_0} \int_{x_0}^{x} f(t)\,\mathrm{d}t = f(x_0) + \frac{1}{x - x_0} \int_{x_0}^{x} (f(t) - f(x_0))\,\mathrm{d}t$$

und es ist

$$\left| \frac{F(x) - F(x_0)}{x - x_0} - f(x_0) \right| = \left| \frac{1}{x - x_0} \int_{x_0}^{x} \underbrace{(f(t) - f(x_0))}_{\leq \varepsilon}\,\mathrm{d}t \right| \leq \varepsilon,$$

was die Behauptung ergibt. $\square$

Satz 8.4.6 (Hauptsatz der Differential- und Integralrechnung)**.**

(i) Für $f \in C([a,b])$ ist

$$F(x) = \int_a^x f(t)\,\mathrm{d}t, \qquad x \in [a,b]$$

eine Stammfunktion von f.

(ii) Für jede Stammfunktion F von $f \in C([a,b])$ gilt

$$F(x) = F(a) + \int_a^x f(t)\,\mathrm{d}t. \tag{8.1}$$

Beweis. (i) erhalten wir direkt aus Satz 8.4.5.

(ii) Aus Bemerkung 8.4.4 wissen wir, dass $F(x) = c + \int_a^x f(t)\,\mathrm{d}t$ für ein $c \in \mathbb{R}$ ist; setzt man $x = a$ ein, folgt $F(a) = c$. □

Setzt man in (ii) f' statt f, folgt:

Folgerung 8.4.7. *Für $f \in C^1([a,b])$ und $a, x \in [a,b]$ gilt*

$$f(x) = f(a) + \int_a^x f'(t)\,\mathrm{d}t.$$

Aus (8.1) sehen wir: kennen wir eine Stammfunktion F von f, dann können wir das bestimmte Integral ausrechnen:

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

Einen Grundstock von Stammfunktionen erhalten wir durch das Ableiten von stetig differenzierbaren Funktion. Ist F eine Stammfunktion von $f \in C([a,b])$, so nennen wir F auch *unbestimmtes Integral* von f auf $[a,b]$ und schreiben dafür[1]

$$\int f(x)\,\mathrm{d}x \quad \text{oder} \quad \int f\,\mathrm{d}x.$$

Das unbestimmte Integral zu berechnen bedeutet also, *irgendeine* Stammfunktion zu berechnen.

Satz 8.4.8. *Sei $f\colon [a,b] \to \mathbb{R}$ integrierbar mit $f \geq 0$. Ist f stetig in $p \in [a,b]$ und $f(p) > 0$, dann gilt*

$$\int_a^b f > 0.$$

[1] Die Notation $\int f\,\mathrm{d}x + c$ ist auch üblich, um zu betonen dass man sich die Konstante c frei aussuchen kann.

f	$\int f$	Bedingungen
x^a	$\dfrac{x^{a+1}}{a+1}$	$x \in \mathbb{R}, a \neq -1$
$\dfrac{1}{x}$	$\log \lvert x \rvert$	$x \neq 0$
a^x	$\dfrac{a^x}{\log a}$	$a > 0, a \neq 1$
e^{ax}	$\dfrac{e^{ax}}{a}$	$a \in \mathbb{C} \setminus \{0\}$
$\cos x$	$\sin x$	$x \in \mathbb{R}$
$\sin x$	$-\cos x$	$x \in \mathbb{R}$
$\cosh x$	$\sinh x$	$x \in \mathbb{R}$
$\sinh x$	$\cosh x$	$x \in \mathbb{R}$
$\dfrac{1}{1+x^2}$	$\arctan x$	$x \in \mathbb{R}$
$\dfrac{1}{1-x^2}$	$\dfrac{1}{2} \log \left\lvert \dfrac{1+x}{1-x} \right\rvert$	$x \neq \pm 1$
$\dfrac{1}{\sqrt{1+x^2}}$	$\operatorname{arsinh} x$	$x \in \mathbb{R}$
$\dfrac{1}{\sqrt{1-x^2}}$	$\arcsin x$	$x \in (-1, 1)$
$\dfrac{1}{\sqrt{x^2-1}}$	$\log \left\lvert x + \sqrt{x^2-1} \right\rvert$	$x \neq 1.$

Tabelle 8.1: Stammfunktionen

Beweis. Ist p innerer Punkt von I, dann gibt es ein $\delta > 0$ mit $[p-\delta, p+\delta] \subseteq I$ und

$$|f(x) - f(p)| \leq \frac{f(p)}{2} \quad \text{für } x \in [p-\delta, p+\delta]$$

und damit $f(x) \geq f(p)/2$ für $x \in [p-\delta, p+\delta]$. Damit erhalten wir

$$\int_a^b f = \int_a^{p-\delta} f + \int_{p-\delta}^{p+\delta} f + \int_{p+\delta}^b \geq \frac{f(p)}{2} \int_{p-\delta}^{p+\delta} 1 = \delta f(p) > 0.$$

Ist p Randpunkt, wird analog argumentiert. □

Satz 8.4.9 (Mittelwertsatz der Integralrechnung). *Seien $f, \varphi \in C([a,b])$ mit $\varphi \geq 0$. Dann gibt es ein $\xi \in [a,b]$ mit*

$$\int_a^b f(x)\varphi(x)\,\mathrm{d}x = f(\xi)\int_a^b \varphi(x)\,\mathrm{d}x.$$

Beweis. Für $\varphi = 0$ ist die Aussage trivial, also können wir annehmen dass es ein $x \in [a,b]$ mit $\varphi(x) > 0$ gibt; damit ist auch $\int_a^b \varphi(x)\,\mathrm{d}x > 0$. Setzen wir $m = \min_{x\in[a,b]} f(x)$ und $M = \max_{x\in[a,b]} f(x)$, dann gilt $m\varphi \leq f\varphi \leq M\varphi$ und damit auch

$$m\int_a^b \varphi\,\mathrm{d}x \leq \int_a^b f\varphi\,\mathrm{d}x \leq M\int_a^b \varphi(x)\,\mathrm{d}x,$$

also

$$m \leq \frac{\int_a^b f\varphi}{\int_a^b \varphi} \leq M.$$

Der Zwischenwertsatz (Satz 5.2.4) gibt nun sofort die Behauptung. □

Mit $\varphi = 1$ erhält man:

Folgerung 8.4.10. *Ist $f \in C([a,b])$, dann gibt es $\xi \in [a,b]$ mit $\int_a^b f\,\mathrm{d}x = f(\xi)(b-a)$.*

8.5 Die Technik des Integrierens

Satz 8.5.1 (Partielle Integration). *Für $u, v \in C^1([a,b])$ gilt*

$$\int_a^b uv'\,\mathrm{d}x = uv\Big|_a^b - \int_a^b u'v\,\mathrm{d}x.$$

Für das unbestimmte Integral gilt

$$\int uv'\,\mathrm{d}x = uv - \int u'v\,\mathrm{d}x.$$

Beweis. Das folgt unmittelbar aus der Produktregel $(uv)' = u'v + v'u$. □

Beispiel.

(i) $\int_a^b \cos^2 x \,\mathrm{d}x$:

$$
\begin{aligned}
\int_a^b \cos^2 x \,\mathrm{d}x &= \cos x \sin x\Big|_a^b + \int_a^b \sin x\, \sin x \,\mathrm{d}x \\
&= \sin x \cos x\Big|_a^b + \int_a^b (1 - \cos^2 x)\,\mathrm{d}x \\
\Longrightarrow\ 2\int_a^b \cos^2 x \,\mathrm{d}x &= \sin x \cos x\Big|_a^b + \int_a^b 1 \,\mathrm{d}x \\
\Longrightarrow\ \int_a^b \cos^2 x \,\mathrm{d}x &= \frac{1}{2}(x + \sin x \cos x)\Big|_a^b = \frac{b-a}{2} + (\sin b \cos b - \sin a \cos a)/2.
\end{aligned}
$$

Aus $\int_a^b \cos^2 x \,\mathrm{d}x + \int_a^b \sin^2 x \,\mathrm{d}x = \int_a^b 1 \,\mathrm{d}x = b - a$ folgt

$$\int_a^b \sin^2 x \,\mathrm{d}x = \frac{b-a}{2} - (\sin b \cos b - \sin a \cos a)/2.$$

(ii) $\int_a^b x^2 e^x \,\mathrm{d}x$:

$$
\begin{aligned}
\int_a^b x^2 e^x \,\mathrm{d}x &= x^2 e^x\Big|_a^b - \int_a^b 2x e^x \,\mathrm{d}x \\
&= x^2 e^x\Big|_a^b - 2x e^x\Big|_a^b + \int_a^b 2e^x \,\mathrm{d}x \\
&= (x^2 - 2x + 2)e^x\Big|_a^b.
\end{aligned}
$$

Satz 8.5.2 (Substitutionsregel). *Ist $f \in C([\alpha, \beta])$ und $\varphi \in C^1([a, b])$ mit $\varphi([a, b]) \subseteq [\alpha, \beta]$, dann gilt*

$$\int_a^b f(\varphi(x))\varphi'(x)\,\mathrm{d}x = \int_{\varphi(a)}^{\varphi(b)} f(y)\,\mathrm{d}y.$$

Für das unbestimmte Integral gilt

$$\int f(\varphi(x))\varphi'(x)\,\mathrm{d}x = F(\varphi(x)),$$

wobei F eine Stammfunktion von f ist.

In der Praxis liest man diese Gleichung oft rückwärts, d.h. findet man irgendein φ und a', b' mit $\varphi(a') = \alpha$ und $\varphi(b') = \beta$ (das geht z.B., aber nicht nur, wenn φ bijektiv ist), dann ist

$$\int_\alpha^\beta f(y)\,\mathrm{d}y = \int_{a'}^{b'} f(\varphi(x))\varphi'(x)\,\mathrm{d}x.$$

Schematisch schreibt man $y = \varphi(x)$, $\mathrm{d}y = \varphi'(x)\,\mathrm{d}x$, was man sinngemäß aus $\frac{\mathrm{d}y}{\mathrm{d}x} = \varphi'(x)$ durch „Multiplikation" mit $\mathrm{d}x$ erhält.

Beweis. Aus dem Fundamentalsatz erhalten wir eine Stammfunktion $F \in C^1([\alpha,\beta])$ von f, mit der Kettenregel (Satz 7.1.6) ist $F \circ \varphi \in C^1([a,b])$ und

$$(F \circ \varphi)'(x) = F'(\varphi(x))\varphi'(x) = f(\varphi(x))\varphi'(x), \quad x \in [a,b].$$

Somit erhalten wir die Behauptung aus

$$\int_a^b f(\varphi(x))\varphi'(x)\,\mathrm{d}x = F \circ \varphi\Big|_a^b = F(\varphi(b)) - F(\varphi(a)) = F\Big|_{\varphi(a)}^{\varphi(b)} = \int_{\varphi(a)}^{\varphi(b)} f(y)\,\mathrm{d}y. \qquad \square$$

Beispiel.

(i) Wir schreiben $xe^{-x^2/2} = f(\varphi(x))\varphi'(x)$ mit $f(y) = e^{-y}$ und $\varphi(x) = x^2/2$. Damit gilt

$$\int_a^b xe^{-x^2/2}\,\mathrm{d}x = \int_{\varphi(a)}^{\varphi(b)} e^{-y}\,\mathrm{d}y = -e^{-y}\Big|_{u(a)}^{u(b)} = e^{-a^2/2} - e^{-b^2/2}.$$

Üblicherweise schreibt man das Symbol φ nicht explizit an sondern nur

$$\frac{x^2}{2} = y, \quad x\,\mathrm{d}x = \mathrm{d}y.$$

Hier erkennt man also, dass der Integrad von der Form $f(\varphi(x))\varphi'(x)$ ist und muss nur $y = \varphi(x)$ setzen.

(ii) Wir berechnen die Fläche des oberen Halbkreises mit Radius r durch die Substitution $x = r\sin t$, $\mathrm{d}x = r\cos t\,\mathrm{d}t$:

$$\int_{-r}^{r} \sqrt{r^2 - x^2}\,\mathrm{d}x = \int_{-\pi/2}^{\pi/2} \sqrt{r^2 - r^2\sin^2 t}\, r\cos t\,\mathrm{d}t = r^2 \int_{-\pi/2}^{\pi/2} \cos^2 t\,\mathrm{d}t = r^2\pi/2.$$

Analong könnte man $x = r\cos t$, $\mathrm{d}x = -r\sin t\mathrm{d}t$ setzen und erhielte

$$\int_{-r}^{r} \sqrt{r^2 - x^2}\,\mathrm{d}x = -\int_{\pi}^{0} \sqrt{r^2 - r^2\cos^2 t}\, r\sin t\,\mathrm{d}t = r^2 \int_0^{\pi} \sin^2 t\,\mathrm{d}t = r^2\pi/2.$$

Hier sucht man sich also ein geeignetes $\varphi \in C^1([a,b])$ mit $\varphi([a,b]) \subseteq [\alpha,\beta]$ aus, für das es a' und b' gibt mit $\varphi(a') = \alpha$ und $\varphi(b') = \beta$.

Andere Integralbegriffe

Neben dem von uns eingeführten *Regel-* oder *Cauchy-Riemannschen Integral* gibt es auch andere Integralbegriffe; diese unterscheiden sich einerseits in der Klasse der Funktionen, welche integriert werden können, und andererseits in bestimmten Eigenschaften des Integrals. Während wir also stückweise stetige bzw. monotone Funktionen integrieren können – und im Prinzip auch alle sprungstetigen Funktionen (siehe Satz 8.3.11), obwohl wir das nicht bewiesen haben – stehen wir bei der Funktion $f(x) = \sin x$ auf $[0,1]$ an, denn diese lässt sich nicht bezüglich der Supremumnorm durch Treppenfunktionen approximieren: die erste „Treppe" wird von der form $c \cdot \mathbb{1}_{(0,a)}$ für irgendwelche $c \in \mathbb{R}$ und $a > 0$ sein, und in diesem Intervall nimmt der Sinus alle Werte zwischen -1 und 1 an, also immer auch einen der von a mindestens den Abstand 0,5 hat.

Da man das Riemann-Integral unweigerlich in vielen Lehrbüchern antreffen wird, sei seine Konstruktion der Vollständigkeit halber erwähnt: für eine Funktion $f\colon [a,b] \to \mathbb{R}$ setzt man

$$O_f := \inf\left\{\int_a^b \varphi : \varphi \in T[a,b],\ \varphi \geq f\right\},$$
$$U_f := \sup\left\{\int_a^b \varphi : \varphi \in T[a,b],\ \varphi \leq f\right\},$$

d.h. man versucht, f von oben und von unten mit Treppenfunktionen anzunähern – für deren Integral gibt es ja nur eine sinnvolle Definition, egal welche Integrationstheorie man dann damit aufbaut.

Ist $O_f = U_f$, dann heißt f *Riemann-integrierbar* und man setzt $\int_a^b f := O_f = U_f$. Sprungstetige Funktionen fallen in diese Klasse:

Satz 8.5.3. *Ist f integrierbar (in unserem Sinne), dann auch Riemann-integrierbar.*

Beweis. Nach Definition der Integrierbarkeit gibt es für alle $n \in N$ ein $\varphi_n \in T[a,b]$ mit $\varphi_n(x) - 1/n \le f(x) \le \varphi(x) + 1/n$ für alle $x \in [a,b]$. Durch Integration von $\varphi_n \pm 1/n \in T[a,b]$ erhält man

$$\int_a^b (\varphi_n - \frac{1}{n}) \le U_f \le O_f \le \int_a^b (\varphi_n + \frac{1}{n})$$

oder

$$\int_a^b \varphi_n - \frac{b-a}{n} \le U_f \le O_f \le \int_a^b \varphi_n + \frac{b-a}{n}.$$

Für $n \to \infty$ folgt daraus $U_f = O_f$, da die linke und die rechte Seite den selben Grenzwert haben. □

Um zu sehen, dass $f(x) = \sin(1/x)$ Riemann-integrierbar ist, konstruieren wir Treppenfunktionen $\varphi \ge f$ und $\psi \le f$, sodass $\int \varphi - \int \psi$ beliebig klein ist, woraus wieder $O_f = U_f$ folgt. Sei $n \in \mathbb{N}$. Wegen $f'(x) = -(1/x^2)\cos(1/x)$ ist nach dem Mittelwertsatz für x, y mit $1/n \le x < y \le 1$ $|f(x) - f(y)| \le |x - y| \cdot |f'(\xi)|$ für ein $\xi \in (x,y)$, und damit auch $|f(x) - f(y)| \le n^2|x - y|$. Wir zerlegen nun das Intervall $[0,1]$ mittels den Teilungspunkten $t_{-1} = 0$, $t_0 = 1/n$, $t_k = 1/n + \frac{1-1/n}{m} \cdot k$ für $k = \ldots m$, wobei wir $m \ge n^3$ gewählt haben. Wir setzen nun $\varphi(x) = \max_{y \in (t_i, t_{i+1})} f(y)$ bzw. $\psi(x) = \min_{y \in (t_i, t_{i+1})} f(y)$ für $x \in (t_i, t_{i+1})$ und $i = -1, \ldots, m$. Dann ist $\psi \le f \le \varphi$ und

$$\int_0^1 \varphi - \int_0^1 \psi \le \frac{2}{n} + \sum_{k=1}^m n^2 \left(\frac{1-1/n}{m}\right)^2 \le \frac{2}{n} + m\frac{n^2}{m^2} \le \frac{2}{n} + \frac{1}{n} = \frac{3}{n}.$$

Es gibt natürlich auch Funktionen, die nicht Riemann-integrierbar sind, z.B. die *Dirichlet-Funktion*

$$D(x) := \begin{cases} 1 & x \text{ irrational} \\ 0 & x \text{ rational.} \end{cases}$$

ist auf $[0,1]$ nicht Riemann-Integrierbar: eine Obersumme (d.h. das Integral einer Treppenfunktion $\ge D$) ist immer ≥ 1, eine Untersumme (d.h. das Integral einer Treppenfunktion $\le D$) immer ≤ 0. Diese Funktion ist jedoch *Lebesgue-integrierbar*.

Das Lebesgue-Integral ähnelt im Aufbau dem Regelintegral, nur ersetzt man in den Treppenfunktionen

$$\varphi = \sum c_j \mathbb{1}_{(t_{j-1}, t_j)}, \qquad \int \varphi = \sum c_j \big| t_j - t_{j-1} \big| \tag{8.2}$$

die Intervalle (t_{j-1}, t_j) durch allgemeinere Mengen $A_j \subseteq \mathbb{R}$:

$$\varphi = \sum c_j \mathbb{1}_{A_j}, \qquad \int \varphi = \sum c_j \big| A_j \big|. \tag{8.3}$$

Die Schwierigkeit hierbei liegt darin, den Mengen A_j einen *Inhalt* $\left|A_j\right|$ zuzuordnen – für Intervalle ist das klarerweise deren Länge, nur für allgemeine Mengen treten hier Schwierigkeiten auf die man mit den Methoden der *Maßtheorie* bändigen kann. Kurz gesagt, will man einen sinnvollen und gleichzeitig möglichst weit gefassten Inhaltsbegriff definieren, dann geht das nur für bestimmte, sogenannte *messbare* Teilmengen von $\mathbb{R}$. An die Stelle der sprungstetigen Funktionen, die man mit Treppenfunktionen der Form (8.2) beliebig genau annähner kann, treten dann die *messbaren* Funktionen, die man mit Funktionen der Form (8.3) beliebig genau annähnern kann.

TLDR: das Regelintegral reicht für unsere Zwecke vollkommen aus; will man einen leistungsfähigeren Integralbegriff, muss man sowieso zum Lebesgue-Integral greifen, dem das Regelintegral aber wieder konzeptuell viel näher ist als das Riemann-Integral. 20. 12. 2024

8.6 Integration rationaler Funktionen

Seien p, q reelle Polynome mit $q(x) \neq 0$ auf einem Intervall $[a, b]$. Wir machen uns daran, eine Stammfunktion der rationalen Funktion p/q auf $[a, b]$ zu bestimmen Nach Kürzen und Polynomdivision erhält man

$$\frac{p}{q} = f + \frac{r}{q}$$

wobei f ein Polynom ist und $\operatorname{Grad} r < \operatorname{Grad} q$. Die Integration von f stellt kein Problem dar. Für r/q führen wir eine Partialbruchzerlegung durch und erhalten daraus Terme der Form

$$\frac{a_1}{x-c} + \frac{a_2}{(x-c)^2} + \ldots + \frac{a_m}{(x-c)^m}$$

wenn c eine reelle Nullstelle der Vielfachheit m von q ist, bzw.

$$\frac{A_1 x + B_1}{x^2 + \alpha x + \beta} + \ldots + \frac{A_n x + B_n}{(x^2 + \alpha x + \beta)^n}$$

wenn $c, \bar{c}$ ein Paar nichtreeller Nullstellen der Vielfachheit n von q ist, wobei $\alpha = -2 \operatorname{Re} c$ und $\beta = |c|^2$ und damit $\alpha^2 < 4\beta$ ist (der Nenner entsteht damit aus $x^2 + \alpha x + \beta = (x - c)(x - \bar{c})$). Im ersten Fall haben wir direkt

$$\int_a^b \frac{\mathrm{d}x}{(x-c)^k} = \begin{cases} \log|x-c| \Big|_{x=a}^{b} & k = 1, \\ \dfrac{1}{(1-k)(x-c)^{k-1}} \Bigg|_{x=a}^{b} & k > 1. \end{cases}$$

Im zweiten Fall schreiben wir

$$\int_a^b \frac{Ax + B}{(x^2 + \alpha x + \beta)^k} \,\mathrm{d}x = \frac{A}{2} \int_a^b \frac{2x + \alpha}{(x^2 + \alpha x + \beta)^k} \,\mathrm{d}x + (B - \frac{A\alpha}{2}) \int_a^b \frac{\mathrm{d}x}{(x^2 + \alpha x + \beta)^k}.$$

Für den ersten Summanden verwenden wir

$$\int_a^b \frac{2x + \alpha}{(x^2 + \alpha x + \beta)^k} \,\mathrm{d}x = \begin{cases} \log\left|x^2 + \alpha x + \beta\right| \Big|_{x=a}^{b} & k = 1, \\ \dfrac{1}{(1-k)(x^2 + \alpha x + \beta)^{k-1}} \Bigg|_{x=a}^{b} & k > 1. \end{cases}$$

Um das verbleibende Integral zu berechnen, leiten wir für $k \geq 2$ eine Reduktionsformel her:

$$\begin{aligned} \left(\frac{2x + \alpha}{(x^2 + \alpha x + \beta)^{k-1}}\right)' &= \frac{2}{(x^2 + \alpha x + \beta)^{k-1}} - \frac{(2x + \alpha)^2 (x^2 + \alpha x + \beta)^{k-2} (k-1)}{(x^2 + \alpha x + \beta)^{2k-2}} \\ &= \frac{-2(2k-3)}{(x^2 + \alpha x + \beta)^{k-1}} + \frac{(k-1)(4\beta - \alpha^2)}{(x^2 + \alpha x + \beta)^k}, \end{aligned}$$

also mit $D = (k-1)(4\beta - \alpha^2)$

$$\frac{1}{(x^2 + \alpha x + \beta)^k} = \left(\frac{2x + \alpha}{D(x^2 + \alpha x + \beta)^{k+1}}\right)' + \frac{4k-6}{D(x^2 + \alpha x + \beta)^{k-1}}.$$

Damit ist

$$\int_a^b \frac{\mathrm{d}x}{(x^2+\alpha x+\beta)^k} = \frac{2x+\alpha}{D(x^2+\alpha x+\beta)^{k-1}}\bigg|_a^b + \frac{4k-6}{D}\int_a^b \frac{\mathrm{d}x}{(x^2+\alpha x+\beta)^{k-1}}.$$

Der Grad des Nennerpolynoms ist somit um eins weniger; dies wiederholt man bis nur noch

$$\int_a^b \frac{\mathrm{d}x}{x^2+\alpha x+\beta}$$

übrig bleibt. Wegen $\alpha^2 < 4\beta$ ist

$$x^2+\alpha x+\beta = \left(x+\frac{\alpha}{2}\right)^2+\beta-\frac{\alpha^2}{4} = \frac{1}{4}\left((2x+\alpha)^2+4\beta-\alpha^2\right) = \frac{4\beta-\alpha^2}{4}\left(\left(\frac{2x+\alpha}{\sqrt{4\beta-\alpha^2}}\right)^2+1\right) = \frac{A^2}{4}(t^2+1)$$

mit $A = \sqrt{4\beta-\alpha^2}$, $t = \frac{2x+\alpha}{A} = \varphi(x)$, $\mathrm{d}x = \frac{A}{2}\,\mathrm{d}t$. Damit ist

$$\int_a^b \frac{\mathrm{d}x}{x^2+\alpha x+\beta} = \frac{2}{A}\int_{\varphi(a)}^{\varphi(b)} \frac{\mathrm{d}t}{t^2+1} = \frac{2}{A}\arctan\left(\frac{2x+\alpha}{A}\right)\bigg|_{x=a}^b.$$

Weitere Integrale

Zum Abschluss führen wir noch die wichtigsten Substitutionen an, um einige Klassen von Integralen zu berechnen. Dazu sei $R(\xi)$ eine rationale Funktion bzw. $R(\xi,\eta)$ eine rationale Funktion in zwei Variablen, d.h. mit gemischten Potenzen von ξ und η in Zähler und Nenner.

(a) $\int R(x, \sqrt[k]{\frac{\alpha x+\beta}{\gamma x+\delta}})\,\mathrm{d}x$, $k = 2,3,\dots$ mit $\alpha\delta-\beta\gamma \neq 0$: Substitution $u = \sqrt[k]{\frac{\alpha x+\beta}{\gamma x+\delta}}$ führt auf eine rationale Funktion in u.

(b) $\int R(e^{\alpha x})$: Substitution $u = e^{\alpha x}$ führt auf eine rationale Funktion in u.

(c) $\int R(\sin x, \cos x)\,\mathrm{d}x$: Substitution $t = \tan\frac{x}{2}$ führt auf eine rationale Funktion in t.

(d) $\int R(x, \sqrt{1+x^2})\,\mathrm{d}x$: Substitution $x = \sinh u$ führt auf Typ (b).

(e) $\int R(x, \sqrt{x^2-1})\,\mathrm{d}x$: Substitution $x = \cosh u$ führt auf Typ (b).

(f) $\int R(x, \sqrt{1-x^2})\,\mathrm{d}x$: Substitution $x = \cos u$ oder $x = \sin u$ führt auf (c).

(g) $\int R(x, \sqrt{\alpha x^2+\beta x+\gamma})\,\mathrm{d}x$: quadratische Erweiterung führt auf Typ (d), (e) oder (f).

(h) $\int P(x)\cdot\begin{Bmatrix} e^{\alpha x} \\ \sin(\alpha x) \\ \cos(\alpha x)\end{Bmatrix}\mathrm{d}x$, wobei $P(x)$ ein Polynom ist: partielle Integration.

Im Fall (c) verwendet man dabei die Identitäten

$$\sin x = 2\sin\frac{x}{2}\cos\frac{x}{2} = 2\tan\frac{x}{2}\cos^2\frac{x}{2} = \frac{2\tan\frac{x}{2}}{1+\tan^2\frac{x}{2}} = \frac{2t}{1+t^2},$$

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

8.7 Uneigentliche Integrale

Definition 8.7.1. Sei $J \subseteq \mathbb{R}$ ein Intervall mit $a = \inf J$ und $b = \sup J$, und $f\colon J \to \mathbb{R}$ eine Funktion, die auf jedem kompakten Intervall $[c, d] \subseteq J$ integrierbar ist. Die Funktion f heißt *uneigentlich integrierbar*, falls es ein $c \in J$ gibt sodass

$$\lim_{\alpha \to a+} \int_\alpha^c f, \quad \lim_{\beta \to b-} \int_c^\beta f \tag{8.4}$$

existieren; dann heißt

$$\int_a^b f(x)\,\mathrm{d}x := \lim_{\alpha \to a+} \int_\alpha^c f + \lim_{\beta \to b-} \int_c^\beta f$$

uneigentliches Integral von f über J. Man sagt auch: das Integral $\int_a^b f$ *existiert* oder *konvergiert*.

Bemerkung 8.7.2.

(i) Existieren die Grenzwerte in (8.4) für ein $c \in J$, dann für *alle* $c \in J$, und der Wert von $\int_a^b f$ ist unabhängig von der Wahl von c. Dies folgt aus

$$\int_\alpha^{c_1} f = \int_\alpha^{c_2} f + \int_{c_2}^{c_1} f, \quad \int_{c_1}^\beta f = \int_{c_2}^\beta f + \int_{c_1}^{c_2} f.$$

(ii) Für $-\infty < a < b < \infty$ und $f\colon [a, b] \to \mathbb{R}$ stimmen das uneigentliche und das Cauchy-Riemannsche Integral überein. Dies sieht man aus

$$\left|\int_a^c f - \int_\alpha^c f\right| = \left|\int_a^\alpha f\right| \leq |a - \alpha| \cdot \|f\| \to 0 \quad \text{für } \alpha \to a+$$

bzw. analog für $\beta \to b-$.

Damit können wir sowohl Funktionen mit unbeschränktem Definitionsbereich als auch Funktionen mit unbeschränktem Integranden behandeln:

Beispiel 8.7.3.

(i) Sei $a > 0$ und $s \in \mathbb{R}$. Dann gilt:

$$\int_a^\infty \frac{\mathrm{d}x}{x^s} \text{ existiert} \iff s > 1, \text{ und } \int_a^\infty \frac{\mathrm{d}x}{x^s} = \frac{a^{1-s}}{s-1} \text{ für } s > 1.$$

Für $s \neq 1$ ist nämlich

$$\int_a^\infty \frac{\mathrm{d}x}{x^s} = \lim_{b \to \infty} \int_a^b \frac{\mathrm{d}x}{x^s} = \lim_{b \to \infty} \left. \frac{x^{-s+1}}{-s+1} \right|_a^b = \frac{a^{1-s}}{s-1} - \lim_{b \to \infty} \frac{b^{1-s}}{s-1}$$

konvergent genau für $s > 1$. Für $s = 1$ ist

$$\int_a^\infty \frac{\mathrm{d}x}{x} = \lim_{b \to \infty} \log x \Big|_a^b = \infty.$$

(ii) Sei $b > 0$ und $s \in \mathbb{R}$. Dann gilt:

$$\int_0^b \frac{\mathrm{d}x}{x^s} \text{ existiert} \iff s < 1, \text{ und } \int_0^b \frac{\mathrm{d}x}{x^s} = \frac{b^{1-s}}{1-s} \text{ für } s < 1.$$

(iii) $\displaystyle\int_0^\infty x^s \,\mathrm{d}x$ existiert für *kein* $s \in \mathbb{R}$.

(iv) $\displaystyle\int_{-\infty}^\infty x \,\mathrm{d}x = \lim_{\substack{a\to-\infty \\ b\to\infty}} \frac{x^2}{2}\Big|_a^b$ existiert nicht.

Kapitel 9

Funktionenfolgen und Funktionenreihen

9.1 Punktweise und gleichmäßige Konvergenz

Ist $(f_n)_n$ eine Folge von rellen Funktionen auf einem Intervall, so lassen sich zwei Arten von Konvergenz dieser Folge erklären:

Definition 9.1.1. Sei $D \subseteq \mathbb{R}$ eine beliebige Menge und $(f_n)_n$ eine Folge von Funktionen $D \to \mathbb{R}$ sowie $f\colon D \to \mathbb{R}$ eine Funktion.

(i) Die Folge $(f_n)_n$ heißt *punktweise konvergent* gegen f auf D, wenn für jedes $x \in D$ die Folge $(f_n(x))_n$ gegen $f(x)$ konvergiert, d.h.

$$\forall x \in D: \ f_n(x) \to f(x).$$

(ii) Die Folge $(f_n)_n$ heißt *gleichmäßig konvergent* gegen f auf D, wenn

$$\|f_n - f\|_D \to 0.$$

Wir schreiben $\|\cdot\|_D$, um zu verdeutlichen auf welcher Menge die Supremumsnorm genommen wird. Klarerweise ist eine gleichmäßig konvergente Folge auch punktweise konvergent mit der selben Grenzfunktion (denn $|f_n(x) - f(x)| \le \|f_n - f\|$). Die Umkehrung gilt nicht:

Beispiel. Die Funktionenfolge $f_n(x) = x^n$ auf $[0, 1)$ konvergiert punktweise gegen $f = 0$, jedoch ist für jedes n

$$\|f_n - f\| = \sup\{x^n \mid 0 \le x < 1\} = 1.$$

Lemma 9.1.2.

(i) Konvergiert eine Folge (f_n) gleichmäßig auf D, so ist sie gleichmäßig beschränkt, d.h. es gibt eine Konstante C mit

$$\|f_n\|_D \le C \text{ für alle } n.$$

(ii) Aus der gleichmäßigen Konvergenz $f_n \to f$ und $g_n \to g$ folgt die gleichmäßige Konvergenz

$$\begin{aligned} &af_n + bg_n \to af + bg \quad \text{für} \quad a, b \in \mathbb{R}, \\ &f_n \cdot g_n \to f \cdot g, \\ &|f_n| \to |f|. \end{aligned}$$

Dies sieht man wie im Beweis von Satz 8.3.6.

Bemerkung 9.1.3. Die punktweise Konvergenz einer Folge stetiger Funktionen garantiert nicht, dass die Grenzfunktion stetig ist. Nimmt man z.B. die stetigen Funktionen $f_n(x) = x^n$ auf $[0, 1]$, so ist die Grenzfunktion die unstetige Funktion

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

Ganz anders schaut es im Fall der gleichmäßigen Konvergenz aus:

Satz 9.1.4. *Sei $(f_n)_n$ eine Folge von Funktionen $D \to \mathbb{R}$, welche gleichmäßig gegen $f\colon D \to \mathbb{R}$ konvergiert. Sind die f_n alle stetig, so auch die Grenzfunktion f.*

Beweis. Für beliebige $x, y \in D$ und n gilt nach der Dreiecksungleichung

$$|f(y) - f(x)| \le |f(y) - f_n(y)| + |f_n(y) - f_n(x)| + |f_n(x) - f(x)|.$$

Wählt man zu gegebenem $\varepsilon > 0$ zuerst n so groß, dass $\|f - f_n\| < \varepsilon$ ist, und dann $\delta > 0$ so klein, dass $|f_n(y) - f_n(x)| < \varepsilon$ ist für alle $y \in D$ mit $|y - x| < \delta$, dann folgt für alle solchen y dass $|f(y) - f(x)| < 3\varepsilon$, was die Behauptung ergibt. □

Anders ausgedrück besagt dieser Satz, dass für eine gleichmäßig konvergente Folge (f_n) stetiger Funktionen und eine konvergente Folge (x_n) im Definitionsbereich dieser Funktion folgende *Vertauschbarkeit von Grenzübergängen* gilt:

$$\lim_{n\to\infty} \lim_{k\to\infty} f_n(x_k) = \lim_{k\to\infty} \lim_{n\to\infty} f_n(x_k).$$

Da auch die Ableitung und die Integration über Grenzprozesse definiert sind, studieren wir im Folgenden, unter welchen Voraussetzungen sich diese mit der Konvergenz von Funktionen vertauschen lassen, d.h. ob man den Grenzwert $f_n \to f$ in die Ableitung bzw. das Integral „hineinziehen“ kann.

Satz 9.1.5. *Sei $(f_n)_n$ eine Folge integrierbarer Funktionen auf $[a, b]$. Konvergiert diese Folge gleichmäßig gegen eine Funktion f, so ist auch f integrierbar und es gilt*

$$\int_a^b f(x)\,\mathrm{d}x = \lim_{n\to\infty} \int_a^b f_n(x)\,\mathrm{d}x.$$

Beweis. Wählt man für jedes f_n eine Treppenfunktion φ_n mit $\|f_n - \varphi_n\| < \frac{1}{n}$, dann gilt

$$\|f - \varphi_n\| \le \|f - f_n\| + \|f_n - \varphi_n\| \to 0,$$

also ist f gleichmäßiger Limes der φ_n, d.h. integrierbar. Außerdem gilt

$$\left| \int_a^b f - \int_a^b f_n \right| \le (b-a)\|f - f_n\| \to 0. \qquad \square$$

Satz 9.1.6. *Sei $(f_n)_n$ eine Folge stetig differenzierbarer Funktionen auf einem Intervall $I \subseteq \mathbb{R}$ sodass*

(i) $f_n \to f$ punktweise in I,

(ii) $f_n' \to g$ gleichmäßig auf jedem kompakten Teilintervall von I.

Dann ist f in I stetig differenzierbar und $f' = g$, d.h.

$$\frac{\mathrm{d}}{\mathrm{d}x} \lim_{n\to\infty} f_n(x) = \lim_{n\to\infty} \frac{\mathrm{d}}{\mathrm{d}x} f_n(x).$$

Beweis. Sei $x_0 \in I$ beliebig. Für $x \in I$ gilt dann (da $f \in C^1([x_0, x])$, nach Folgerung 8.4.7)

$$f_n(x) = f_n(x_0) + \int_{x_0}^x f_n'(t)\,\mathrm{d}t.$$

Für $n \to \infty$ folgt hieraus

$$f(x) = f(x_0) + \int_{x_0}^x g(t)\,\mathrm{d}t.$$

Nach Satz 9.1.4 ist g stetig auf I, also ist nach dem Hauptsatz f stetig differenzierbar und $f' = g$. $\square$

Diese Ergebnisse übertragen sich direkt auf Funktionenreihen, d.h. auf Reihen der Form

$$\sum_{n=1}^{\infty} f_n(x).$$

Eine Funktionenreihe heißt *absolut konvergent*, wenn $\sum_{n=1}^{\infty} |f_n(x)|$ für alle x konvergiert, und *gleichmäßig konvergent*, wenn die Folge ihrer Partialsummen gleichmäßig konvergiert. Ein einfaches aber wichtiges Kriterium dafür lautet wie folgt:

Satz 9.1.7 (Majorantenkriterium). *Gilt $|f_n(x)| \le a_n$ für alle $n \in \mathbb{N}$ und $x \in I$ und ist $\sum_{n=1}^{\infty} a_n$ konvergent, dann konvergiert $\sum_{n=1}^{\infty} f_n$ absolut und gleichmäßig auf I.*

Beweis. Aus

$$\left| \sum_{n=N}^{\infty} f_n(x) \right| \le \sum_{n=N}^{\infty} |f_n(x)| \le \sum_{n=N}^{\infty} a_n \to 0$$

sieht man die absolute Konvergenz, und nimmt man das Supremum über $x \in I$ auch die gleichmäßige. $\square$

Beispiel 9.1.8. *Wir berechnen (mit partieller Integration)*

$$\int_0^1 x^2 e^x \,\mathrm{d}x = e - 2,$$

es gilt aber auch die Reihendarstellung

$$\int_0^1 x^2 e^x \,\mathrm{d}x = \int_0^1 \sum_{n=0}^{\infty} \frac{x^{n+2}}{n!} \,\mathrm{d}x = \sum_{n=0}^{\infty} \int_0^1 \frac{x^{n+2}}{n!} \,\mathrm{d}x = \sum_{n=0}^{\infty} \frac{1}{n!(n+3)}.$$

Damit sehen wir, dass der Wert dieser Reihe $e - 2$ ist.

10. 1. 2025

9.2 Potenzreihen

Definition 9.2.1. Ist $(a_n)_n$ eine Folge reeller oder komplexer Zahlen, dann heißt die Reihe

$$\sum_{n=0}^{\infty} a_n(z-z_0)^n, \quad z \in \mathbb{C}$$

Potenzreihe mit *Entwicklungspunkt* $z_0 \in \mathbb{C}$.

Die wesentliche Frage ist, wann eine Potenzreihe konvergiert. Führt man als neue Variable $w = z - z_0$ ein, so kann man sich für alles Weitere auf Potenzreihen um den Nullpunkt beschränken.

Beispiel 9.2.2. *(i)* $\sum_{n=0}^{\infty} \frac{1}{n!} z^n$ *konvergiert für alle* $z \in \mathbb{C}$ *nach dem Quotientenkriterium, denn*

$$\left|\frac{a_{n+1}}{a_n}\right| = \frac{|z|}{n+1} \to 0 \quad \forall z \in \mathbb{C}.$$

(ii) $\sum_{n=0}^{\infty} z^n$ *konvergiert genau für* $|z| < 1$*, denn für* $|z| < 1$ *ist das die geometrische Reihe, und für* $|z| \geq 1$ *divergiert die Reihe weil die Summanden keine Nullfolge bilden.*

(iii) $\sum_{n=0}^{\infty} n! z^n$ *konvergiert nur für* $z = 0$*, denn für* $z \neq 0$ *ergibt*

$$\left|\frac{a_{n+1}}{a_n}\right| = (n+1)|z| \to \infty$$

die Divergenz nach dem Quotientenkriterium.

Der Bereich, in dem eine Potenzreihe konvergiert bzw. divergiert, hat eine besondere Form:

Satz 9.2.3.
(i) Konvergiert $\sum_{n=0}^{\infty} a_n z^n$ *für* $z = z_1$*, dann auch absolut für alle* z *mit* $|z| < |z_1|$*.*
(ii) Divergiert $\sum_{n=0}^{\infty} a_n z^n$ *für* $z = z_2$*, dann auch für alle* z *mit* $|z| > |z_2|$*.*

Beweis. (i) Aus der Konvergenz der Reihe folgt, dass $a_n z_1^n$ eine Nullfolge ist, insbesondere also beschränkt, d.h. $\left|a_n z_1^n\right| \leq M$ für ein M und alle n. Mit $q := |z/z_1| < 1$ ist $|a_n z^n| = \left|a_n z_1^n q^n\right| \leq M q^n$, also konvergiert $\sum_{n=0}^{\infty} |a_n z^n|$ nach dem Majorantenkriterium.

(ii) Wäre die Potenzreihe für ein z mit $|z| > |z_2|$ konvergent, dann nach (i) auch für $z = z_2$, was einen Widerspruch ergibt. □

Definition 9.2.4. Der *Konvergenzradius* einer Potenzreihe $\sum_{n=}^{\infty} a_n z_n$ ist die Zahl

$$R := \sup\left\{ |z| \;\middle|\; z \in \mathbb{C} \text{ und } \sum_{n=0}^{\infty} a_n z^n \text{ ist konvergent} \right\}$$

bzw. $R = \infty$, falls die Potenzreihe für alle $w \in \mathbb{C}$ konvergiert.

Bemerkung 9.2.5. Aus dem Beweis von Satz 9.2.3 (i) sieht man: Für $|z| < R$ konvergiert die Reihe sogar absolut; außerdem konvergiert sie sogar gleichmäßig für $|z| \leq p$, wobei p beliebig ist falls $R = \infty$ und $p < R$ falls $R < \infty$.

Der Konvergenzradius lässt sich wie folgt berechnen:

Satz 9.2.6. *Sei R der Konvergenzradius der Reihe*

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

(i) Es gilt

$$R = \frac{1}{\lim_{n\to\infty} \sqrt[n]{|a_n|}} \quad \textit{falls} \lim_{n\to\infty} \sqrt[n]{|a_n|} \in (0,\infty),$$
$$R = \infty \quad \textit{falls} \lim_{n\to\infty} \sqrt[n]{|a_n|} = 0,$$
$$R = 0 \quad \textit{falls} \quad \{\sqrt[n]{|a_n|}\} \textit{ unbeschränkt ist.}$$

(ii) Es gilt

$$R = \frac{1}{\lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right|} \quad \textit{falls} \lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| \in (0,\infty),$$
$$R = \infty \quad \textit{falls} \lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| = 0,$$
$$R = 0 \quad \textit{falls} \quad \{\sqrt[n]{|a_n|}\} \textit{ unbeschränkt ist.}$$

Beweis. Angenommen es existiere $r := \lim_{n\to\infty} \sqrt[n]{|a_n|} \in (0,\infty)$; dann betrachten wir $\sqrt[n]{|a_n z^n|} = \sqrt[n]{|a_n|}\,|z| \to r|z|$. Für $|z| < 1/r$ ist die Reihe $\sum a_n z^n$ dann nach dem Wurzelkriterium (Satz 4.2.10) konvergent, also $R \geq 1/r$. Für $|z| > 1/r$ ist sie divergent, also $R \leq 1/r$ und insgesamt $R = 1/r$.

Gilt $\lim_{n\to\infty} \sqrt[n]{|a_n|} = 0$, konvergiert die Potenzreihe ebenso nach dem Wurzelkriterium für jedes z.

Ist schließlich $\sqrt[n]{|a_n|}$ unbeschränkt, dann auch $\sqrt[n]{|a_n z^n|} = \sqrt[n]{|a_n|}|z|$ für $z \neq 0$, woraus die Divergenz folgt, da damit auch $|a_n z^n|$ unbeschränkt ist.

Die Argumentation über das Quotientenkriterium erfolgt analog. □

Über die Konvergenz für $|z| = R$ lässt sich im Allgemeinen nichts aussagen: nicht nur Konvergenz oder Divergenz, sondern auch beides ist möglich.

Beispiel.

(i) Die geometrische Reihe $\sum_{n=0}^{\infty} z^n$ hat den Konvergenzradius 1 und divergiert für $|z| = 1$.

(ii) Die Logarithmusreihe $\sum_{n=1}^{\infty}(-1)^{n-1}\frac{z^n}{n}$ hat nach dem Quotientenkriterium den Konvergenzradius 1 , denn

$$\left|\frac{a_{n+1}}{a_n}\right| = |z|\frac{n}{n+1} \to |z|.$$

Sie konvergiert für $z = 1$ (alternierende harmonische Reihe), aber divergiert für $z = -1$ (harmonische Reihe)

(iii) Die Reihe $\sum_{n=1}^{\infty}\frac{1}{n^2}z^n$ konvergiert für $|z| < 1$ und divergiert für $|z| > 1$ nach dem Quotientenkriterium, denn

$$\left|\frac{a_{n+1}}{a_n}\right| = |z|\frac{n^2}{(n+1)^2} \to |z|.$$

Für $|z| = 1$ ist die Reihe konvergent, denn

$$\left|\sum_{n=1}^{\infty}\frac{z^n}{n^2}\right| \le \sum_{n=1}^{\infty}\frac{1}{n^2} < \infty.$$

Satz 9.2.7. *Hat eine Potenzreihe mit reellen Koeffizienten*

$$f(x) = \sum_{n=0}^{\infty} a_n(x-x_0)^n$$

den Konvergenzradius $R > 0$, so ist f in $(x_0 - R, x_0 + R)$ differenzierbar und es gilt dort

$$f'(x) = \sum_{n=1}^{\infty} na_n(x-x_0)^{n-1}, \tag{9.1}$$

wobei letztere Reihe auch Konvergenzradius R hat.

Damit ist f auch beliebig oft differenzierbar und es gilt für jedes $k \in \mathbb{N}$

$$f^{(k)}(x) = \sum_{n=k}^{\infty} n(n-1)\cdots(n-k+1)a_n(x-x_0)^{n-k} \tag{9.2}$$

für $|x - x_0| < R$.

Beweis. Nach Satz 9.1.6 müssen wir nur überprüfen, dass für $0 < r < R$ die abgeleitete Reihe

$$\sum_{n=1}^{\infty} na_n(x-x_0)^{n-1}$$

auf $[x_0 - r, x_0 + r]$ gleichmäßig konvergiert. Sei $r < \rho < R$, dann ist $|a_n|\rho^n$ eine Nullfolge, also durch eine Zahl M beschränkt, und für $|x - x_0| \le r$ gilt mit $q = r/\rho$

$$\left|na_n(x-x_0)^{n-1}\right| \le n|a_n|r^{n-1} = \frac{n}{r}|a_n|\rho^n\left(\frac{r}{\rho}\right)^n \le \frac{M}{r}nq^n$$

und $\sum_{n=1}^{\infty} nq^n$ ist nach Lemma 9.2.8 konvergent. Es ist also (9.1) auf $[x_0 - r, x_0 + r]$ für alle $0 < r < R$ gleichmäßig konvergent, was die Behauptung ergibt. □

Lemma 9.2.8. Für $q < 1$ ist die Reihe

$$\sum_{n=1}^{\infty} nq^n$$

konvergent.

Beweis. Setzt man $s = \sqrt{q} \in (0,1)$ so gilt $ns^n \to 0$, denn es ist $1/s = 1 + t$ mit $t > 0$ und damit nach dem binomischen Lehrsatz

$$\frac{1}{s^n} = (1+t)^n = \sum_{k=0}^{n} \binom{n}{k} t^k \geq \frac{n(n-1)}{2} t^2 \implies ns^n \leq \frac{2}{(n-1)t^2} \to 0 \text{ für } n \to \infty.$$

Damit ist $ns^n \leq M$ für eine Zahl M und

$$\sum_{n=1}^{\infty} nq^n = \sum_{n=1}^{\infty} ns^n s^n \leq \sum_{n=1}^{\infty} Ms^n$$

konvergent. □

16. 1. 2025

Wir können nun den Zusammenhang zur Taylorentwicklung erschließen.

Satz 9.2.9 (Identitätssatz für Potenzreihen)**.**

(i) Wird f in $(x_0 - r, x_0 + r)$ durch eine reelle Potenzreihe

$$f(x) = \sum_{n=0}^{\infty} a_n (x - x_0)^n$$

dargestellt, dann gilt

$$a_n = \frac{f^{(n)}(x_0)}{n!}, \qquad n \in \mathbb{N}_0.$$

(ii) Aus

$$\sum_{n=0}^{\infty} a_n (x - x_0)^n = \sum_{n=0}^{\infty} b_n (x - x_0)^n \quad \text{für} \quad |x - x_0| < r$$

folgt $a_n = b_n$ für alle $n \in \mathbb{N}_0$.

Beweis. (i) folgt direkt aus (9.2), denn

$$f^{(k)}(x_0) = k(k-1)\cdots 1 \cdot a_k.$$

(ii) folgt, indem man die Differenz der Reihen betrachtet. □

Anders gesagt: wird eine Funktion durch eine Potenzreihe dargestellt, dann ist diese ihre *Taylorreihe* (Definition weiter unten).

Sind

$$f(x) = \sum_{n=0}^{\infty} a_n z^n, \quad g(z) = \sum_{n=0}^{\infty} b_n z^n$$

zwei Potenzreihen mit Konvergenzradien R_1 bzw. R_2, dann gilt für $|z| < r := \min\{R_1, R_2\}$

$$\alpha f(z) + \beta g(z) = \sum_{n=0}^{\infty} (\alpha a_n + \beta b_n) z^n, \qquad |z| < r,$$

$$f(z) \cdot g(z) = \sum_{n=0}^{\infty} \sum_{k=0}^{n} a_k b_{n-k} z^n = \sum_{n=0}^{\infty} c_n z^n$$

wobei für beide Reihen der Konvergenzradius $R \geq r$ ist.

9.3 Taylorreihen

In Satz 7.3.1 haben wir die Taylorsche Formel mit dem Lagrange-Restglied hergeleitet. Mit Hilfe der Integralrechnung lässt sich eine weitere Form des Restgliedes angeben:

Satz 9.3.1 (Taylorsche Formel mit Integralrestglied). *Sei I ein Intervall, $x_0 \in I$ und $f\colon I \to \mathbb{R}$ eine $(n+1)$-mal stetig differenzierbare Funktion, wobei $n \in \mathbb{N}_0$ ist. Dann gilt für jedes $x \in I$*

$$f(x) = T_n(f, x_0)(x) + R_n(f, x_0)(x)$$

mit dem n-ten Taylorpolynom

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

und dem Integralrestglied

$$R_n(x) = R_n(f, x_0)(x) = \int_{x_0}^{x} \frac{(x-t)^n}{n!} f^{(n+1)}(t)\,\mathrm{d}t.$$

Beweis. Für $n = 0$ ist die Behauptung

$$f(x) = f(x_0) + \int_{x_0}^{x} f'(t)\,\mathrm{d}t$$

gerade der Hauptsatz der Differential- und Integralrechnung. Für den Schluss von n auf $n+1$ erhält man durch partielle Integration

$$\begin{aligned} R_n(x) &= \int_{x_0}^{x} \frac{(x-t)^n}{n!} f^{(n+1)}(t)\,\mathrm{d}t \\ &= -\frac{(x-t)^{n+1}}{(n+1)!} f^{(n+1)}(t)\Big|_{t=x_0}^{x} + \int_{x_0}^{x} \frac{(x-t)^{n+1}}{(n+1)!} f^{(n+2)}(t)\,\mathrm{d}t \\ &= \frac{(x-x_0)^{n+1}}{(n+1)!} + R_{n+1}(x), \end{aligned}$$

also

$$\begin{aligned} f(x) &= T_n(x) + R_n(x) = T_n(x) + \frac{(x-x_0)^{n+1}}{(n+1)!} + R_{n+1}(x) \\ &= T_{n+1}(x) + R_{n+1}(x). \end{aligned}$$

□

Bemerkung. Aus dem Integralrestglied gewinnt man mit dem Mittelwertsatz der Integralrechnung auch das Lagrange-Restglied, und damit einen etwas kürzeren Beweis von Satz 7.3.1

Definition 9.3.2. Sei $f\colon I \to \mathbb{R}$ eine unendlich oft differenzierbare Funktion und $x_0 \in I$. Dann heißt die Potenzreihe

$$T(f, x_0)(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!}(x-a)^n$$

die *Taylorreihe* vom f im Punkt x_0.

Es stellt sich die Frage, ob diese Reihe (wie z.B. im Fall von e^x, $\cos x$ und $\sin x$) die Funktion f darstellt, d.h. ob

$$f(x) = T(f, x_0)(x)$$

gilt. Aus dem Satz von Taylor sieht man: dies ist genau dann der Fall, wenn $R_n(f, x_0)(x) \to 0$ für $n \to \infty$ gilt.

Beispiel. Nicht jede Funktion wird durch ihre Taylorreihe dargestellt! So ist z.B. die Funktion

$$F(x) = \begin{cases} e^{-\frac{1}{x^2}} & x \neq 0 \\ 0 & x = 0 \end{cases}$$

in jedem Punkt beliebig oft differenzierbar und $F^{(n)}(0) = 0$ für alle $n \in \mathbb{N}_0$. Dies sieht man wie folgt: für $x \neq 0$ ist

$$F'(x) = \frac{2}{x^3} e^{-\frac{1}{x^2}}$$

und allgemeiner ist

$$F^{(n)}(x) = \text{ eine Linearkombination von Termen der Form } \frac{1}{x^k} e^{-\frac{1}{x^2}}.$$

Nun ist

$$\lim_{x\to 0} \frac{e^{-\frac{1}{x^2}}}{x^k} = \lim_{x\to 0} \frac{\frac{1}{x^k}}{e^{\frac{1}{x^2}}} = 0,$$

denn

$$\lim_{x\to 0+} \frac{\frac{1}{x^k}}{e^{\frac{1}{x^2}}} = \lim_{y\to\infty} \frac{y^k}{e^{y^2}} = 0,$$
$$\lim_{x\to 0-} \frac{\frac{1}{x^k}}{e^{\frac{1}{x^2}}} = \lim_{y\to -\infty} \frac{y^k}{e^{y^2}} = (-1)^k \lim_{y\to\infty} \frac{y^k}{e^{y^2}} = 0,$$

denn nach Satz 6.2.3 (i) gilt

$$\lim_{y\to\infty} \frac{e^y}{y^k} = \infty \implies \lim_{y\to\infty} \frac{e^{y^2}}{y^k} = \infty.$$

Die Taylorreihe hat also Konvergenzradius ∞, sie stellt aber nicht die Funktion F dar, da letztere ja nicht überall verschwindet. Trotzdem ist das n-te Taylorpolynom, also die Funktion 0, in der Nähe von $x_0 = 0$ eine gute Approximation an f: für $x \in [-r, r]$, $r > 0$ beliebig, ist nach dem Satz von Taylor (Satz 7.3.1)

$$|f(x) - T_n(x)| \leq |R_n(x)| = \left| f^{(n+1)}(\theta) \frac{x^{n+1}}{(n+1)!} \right| \leq C x^{n+1}$$

für eine Konstante C, die jedoch von n abhängt. $n \to \infty$ genügt also nicht, um den Rest beliebig klein zu machen, denn dazu muss auch $|x|$ beliebig klein werden.

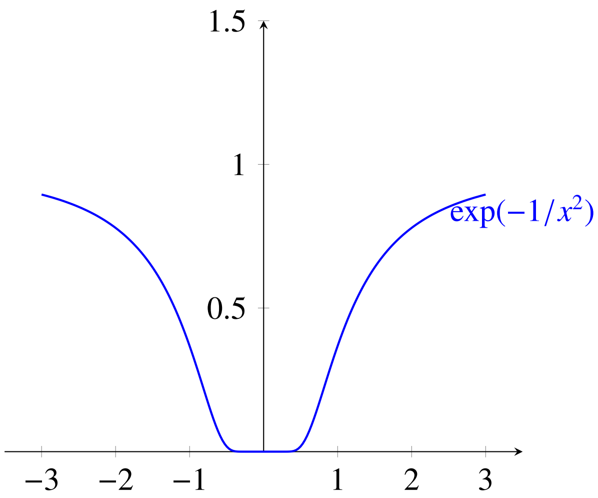


Abbildung 9.1: Die Funktion $\exp(-1/x^2)$

Satz 9.3.3. *Für $|x| < 1$ gilt*

$$\log(1+x) = \sum_{n=1}^{\infty} (-1)^{n-1} \frac{x^n}{n}.$$

Beweis. Sei f diese Reihe. Wir haben schon gesehen, dass sie den Konvergenzradius 1 hat, also gilt für ihre Ableitung

$$f'(x) = \sum_{n=1}^{\infty} (-1)^{n-1} x^{n-1} = \frac{1}{1+x}, \quad |x| < 1.$$

Da dies genau die Ableitung von $\log(1+x)$ ist, gibt es ein $c \in \mathbb{R}$ mit $\log(1+x) = f(x) + c$. Setzt man $x = 0$ ein, sieht man $c = 0$. □

Beispiel. Wir berechnen die Taylorreihe des Arcustangens. Für $f(x) = \arctan x$ gilt

$$\begin{aligned}
f'(x) &= \frac{1}{1+x^2}, \\
f''(x) &= -\frac{2x}{(1+x^2)^2}, \\
f'''(x) &= -\frac{2}{(1+x^2)^2} + \frac{8x^2}{(1+x^2)^3}, \\
f^{(4)}(x) &= \frac{24x}{(1+x^2)^3} - \frac{48x^3}{(1+x^2)^4}, \\
f^{(5)}(x) &= \frac{24}{(1+x^2)^3} + x \cdot \ldots .
\end{aligned}$$

Daraus erraten wir

$$f^{(2k+1)}(0) = (-1)^k (2k)!, \quad f^{(2k)}(0) = 0$$

und behaupten, dass

$$\arctan x = \sum_{k=0}^{n-1} (-1)^k \frac{x^{2k+1}}{2k+1} + R_n(x) \text{ mit } R_n(x) \to 0 \text{ für } n \to \infty.$$

Setzen wir

$$R_n(x) := \arctan(x) - \left(x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \ldots + (-1)^{n-1}\frac{x^{2n-1}}{2n-1}\right),$$

so sehen wir dass

$$\begin{aligned} R_n'(x) &= \frac{1}{1+x^2} - \left(1 - x^2 + x^4 - \ldots + (-1)^{n-1}x^{2n-2}\right) \\ &= \frac{1}{1+x^2} - \frac{1-((-x)^2)^n}{1-(-x^2)} = (-1)^n \frac{x^{2n}}{1+x^2}. \end{aligned}$$

Aus dem verallgemeinerten Mittelwertsatz ergibt sich mit $g(x) = \frac{x^{2n+1}}{2n+1}$ für $x \neq 0$

$$\frac{R_n(x)}{g(x)} = \frac{R_n(x) - R_n(0)}{g(x) - g(0)} = \frac{R_n'(\theta x)}{g'(\theta x)} = \frac{(-1)^n}{1+(\theta x)^2}$$

mit einem $0 < \theta < 1$, also

$$R_n(x) = \frac{(-1)^n}{1+(\theta x)^2}\frac{x^{2n+1}}{2n+1}.$$

Für $|x| \leq 1$ ergibt sich $|R_n(x)| \leq \frac{1}{2n+1}$, also

$$\arctan x = \sum_{k=0}^{\infty}(-1)^k \frac{x^{2k+1}}{2k+1}, \qquad |x| \leq 1.$$

Damit gilt zum Beispiel

$$\frac{\pi}{4} = \arctan 1 = \sum_{n=0}^{\infty}(-1)^n \frac{1}{2n+1} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \ldots .$$

Wir halten noch die uns bekannten Taylorreihen fest:

$$\begin{aligned} e^x &= \sum_{n=0}^{\infty} \frac{x^n}{n!}, && x \in \mathbb{C} \\ \cos x &= \sum_{n=0}^{\infty}(-1)^n \frac{x^{2k}}{(2k)!}, && x \in \mathbb{C} \\ \sin x &= \sum_{n=0}^{\infty}(-1)^n \frac{x^{2k+1}}{(2k+1)!}, && x \in \mathbb{C} \\ \log(1+x) &= \sum_{n=1}^{\infty}(-1)^{n-1}\frac{x^n}{n}, && x \in (-1,1) \end{aligned}$$

Kapitel 10

Iterationsverfahren

10.1 Ein Fixpunktsatz

Sei f eine reelle Funktion, von der wir wissen, dass sie eine Nullstelle ξ hat, d.h. $f(\xi) = 0$. Wie bestimmt man eine Nullstelle numerisch, d.h. näherungsweise? Die Grundidee ist folgende: Startet man in einem beliebigen Punkt x_0, bildet man zuerst die Tangente an den Graphen in der Stelle $(x_0, f(x_0))$. Deren Gleichung ist

$$x \mapsto f(x_0) + f'(x_0)(x - x_0).$$

Der Schnittpunkt dieser Tangente mit der x-Achse ist dann der neue Näherungswert x_1; berechnet man x_1 aus $f(x_0) + f'(x_0)(x_1 - x_0) = 0$, ergibt sich $x_1 = x_0 - f(x_0)/f'(x_0)$. Die Iteration dieses Vorganges nennt man *Newtonverfahren*:

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}.$$

Wählt man z.B. $f(x) = x^2 - a$ mit $a > 0$ fest und $x \geq 0$, dann gibt es nach dem Zwischenwertsatz eine eindeutige Zahl c mit $f(c) = 0$, d.h. $c = \sqrt{a}$. Das Newtonverfahren ergibt die Rekursion[1]

$$x_{n+1} = x_n - \frac{x_n^2 - a}{2x_n} = \frac{1}{2}\frac{x_n^2 + a}{x_n} = \frac{1}{2}\left(x_n + \frac{a}{x_n}\right).$$

Die Frage ist nun, ob das Verfahren konvergiert. Um das zu untersuchen, betrachten wir eine andere, geeignetere Formulierung des Problems. Setzt man — unter der Annahme dass $f'(x) \neq 0$ für alle x gilt —

$$g(x) := x - \frac{f(x)}{f'(x)}$$

dann ist

$$f(\xi) = 0 \iff g(\xi) = \xi.$$

Definition 10.1.1. Ist $g\colon D \to D$ für $D \subseteq \mathbb{R}$ eine Abbildung und $g(\xi) = \xi$, dann heißt ξ *Fixpunkt* von g.

[1] Man könnte auch $f(x) = 1 - \frac{x^2}{a}$ wählen, um eine andere Rekursionsformel für die Quadratwurzel zu erhalten.

Die Frage ist, wann eine Abbildung überhaupt einen Fixpunkt hat und ob man diesen ausgehend von einer ersten Schätzung x_0 durch *sukzessive Approximation*

$$x_{n+1} = g(x_n)$$

finden kann. Wie man in einfachen Beispiele sieht, scheint es einen Zusammenhang zwischen der Konvergenz der sukzessiven Approximation und der lokalen Änderungsrate von g zu geben. Der passende Begriff dafür ist folgender:

Definition 10.1.2. Eine Funktion $f\colon D \to \mathbb{R}$ heißt *Lipschitz*[2]*-stetig*, wenn es eine Konstante $L \geq 0$ gibt mit

$$|f(x_1) - f(x_2)| \leq L|x_1 - x_2| \quad \forall x_1, x_2 \in D.$$

17.1.2025 L heißt *Lipschitzkonstante* von f.

Ist $L < 1$, so heißt f *Kontraktion* und L *Kontraktionskonstante*.

Satz 10.1.3.
(i) Eine Lipschitz-stetige Funktion ist gleichmäßig stetig.
(ii) Ist f stetig differenzierbar auf $[a, b]$, dann ist f Lipschitz-stetig auf $[a, b]$ mit $L = \max_{x\in[a,b]} |f'(x)|$.

Man kann zeigen, dass dieses L optimal ist, d.h. es gibt keine kleinere Lipschitzkonstante für f.

Beweis. (i) Für $\varepsilon > 0$ sei $\delta = \varepsilon/L$; dann gilt

$$|x_1 - x_2| < \delta \implies |f(x_1) - f(x_2)| \leq L|x_1 - x_2| < \varepsilon.$$

(ii) Es ist, mit dem Mittelwertsatz,

$$\left|f(x_1) - f(x_2)\right| = |f'(\xi)(x_1 - x_2)| \leq \max_{x\in[a,b]} |f'(x)| \cdot |x_1 - x_2|,$$

für ein $\xi \in (x_1, x_2)$, d.h. $\max_{x\in[a,b]} |f'(x)|$ ist eine Lipschitzkonstante für f. □

Beispiel. $f(x) = \sqrt{x}$ ist gleichmäßig stetig auf $[0, 1]$ (Beispiel 5.2.10 (ii)), aber nicht Lipschitzstetig, denn wäre L eine Lipschitzkonstante für f, dann wäre

$$\sqrt{2x} - \sqrt{x} \leq L(2x - x) = Lx \text{ bzw. } \sqrt{2} - 1 \leq L\sqrt{x} \quad \forall x \in [0, 1],$$

was für x klein genug nicht stimmen kann.

Satz 10.1.4 (Fixpunktsatz). *Ist $f\colon [a, b] \to [a, b]$ eine Kontraktion, dann gilt:*
(i) f hat genau einen Fixpunkt $\xi \in [a, b]$.
(ii) Für jeden Anfangspunkt $x_0 \in [a, b]$ konvergiert die Folge $x_{n+1} := f(x_n)$, $n \in \mathbb{N}_0$, gegen ξ.

[2] Rudolf Lipschitz, 1832–1903

Hierbei ist wichtig, dass f das Intervall $[a, b]$ in sich selbst abbildet!

Beweis. (i) Ist $h(x) = f(x) - x$, dann ist

$$h(a) = f(a) - a \geq a - a = 0,$$
$$h(b) = f(b) - b \leq b - b = 0,$$

also gibt es ein $\xi \in [a, b]$ mit $h(\xi) = 0$, d.h. $f(\xi) = \xi$.

Ist η ein weiterer Fixpunkt, dann ist

$$|\xi - \eta| = |f(\xi) - f(\eta)| \leq L|\xi - \eta| \implies \xi = \eta$$

da f eine Kontraktion ist, wobei L die Kontraktionskonstante von f ist; dies ist wegen $L < 1$ nur möglich wenn $|xi - \eta| = 0$, also ist der Fixpunkt eindeutig. (ii) folgt aus

$$|x_n - \xi| = |f(x_{n-1}) - f(\xi)| \leq L|x_{n-1} - \xi| \leq L^2|x_{n-2} - \xi| \leq \ldots \leq L^n|x_0 - \xi| \to 0. \qquad \square$$

10.2 Newtonverfahren

Schließlich formulieren wir unser Resultat für die Newton-Iteration:

Satz 10.2.1. *Sei $-\infty < a < b < \infty$, $f \in C^2([a, b])$ und $f'(x) \neq 0$ für $x \in [a, b]$. Wir nehmen an, f habe eine Nullstelle $\xi \in (a, b)$, d.h. $f(\xi) = 0$. Dann gibt es ein $\delta > 0$, sodass für alle Startpunkte $x_0 \in [\xi - \delta, \xi + \delta]$ das Newtonverfahren*

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

gegen die Nullstelle ξ konvergiert.

Beweis. Aufgrund der Voraussetzungen finden wir Konstanten m, M_1, M_2 mit

$$0 < m \leq |f'(x)| \leq M_1, \qquad |f''(x)| \leq M_2.$$

Mit

$$g(x) := x - \frac{f(x)}{f'(x)}, \qquad g'(x) = \frac{f(x)f''(x)}{(f'(x))^2}$$

bekommen wir unter Verwendung des Mittelwertsatzes

$$|g'(x)| \leq \frac{M_2}{m^2}\Big|f(x) - \underbrace{f(\xi)}_{=0}\Big| \leq \frac{M_1 M_2}{m^2}|x - \xi|.$$

Sei $q \in (0, 1)$ beliebig und $\delta > 0$ mit $I := [\xi - \delta, \xi + \delta] \subseteq [a, b]$ und

$$\frac{M_1 M_2}{m^2}\delta \leq q < 1.$$

Dann ist $|g'(x)| \leq q < 1$ für $x \in I$. Außerdem gilt $g(I) \subseteq I$, denn für $x \in I$ ist

$$|g(x) - \xi| = |g(x) - g(\xi)| \leq q|x - \xi| \leq \delta,$$

also $g(x) \in I$. Nach dem Fixpunktsatz (Satz 10.1.4) konvergiert die Folge $x_{n+1} = g(x_n)$ gegen den eindeutigen Fixpunkt ξ von g. $\square$

Wir betrachten zum Abschluss noch die Konvergenzgeschwindigeiten: für die Fixpunktiteration von Satz 10.1.4 haben wir generell

$$|x_{n+1} - \xi| = |g(x_n) - g(\xi)| \leq q|x_n - \xi|,$$

wobei q die Kontraktionskonstante ist. Man sagt, $(x_n)_n$ konvergiert *linear* gegen ξ. Bei jedem Schritt wird der Fehler mit $L < 1$ multipliziert. Für die Newton-Iteration betrachten wir das Taylorpolynom 2. Grades von f mit Entwicklungspunkt x, ausgewertet an der Nullstellt ξ:

$$0 = f(\xi) = f(x) + f'(x)(\xi - x) + \frac{f''(\theta)}{2}(\xi - x)^2.$$

Daraus folgt

$$\begin{aligned} x_{n-1} - \xi &= x_n - \frac{f(x_n)}{f'(x_n)} - \xi = \frac{f''(\theta)}{2f'(x_n)}(\xi - x)^2 \\ \implies |x_{n+1} - \xi| &\leq \frac{M_2}{2m}|\xi - x|^2. \end{aligned}$$

Hier sagt man, $(x_n)_n$ konvergiert *quadratisch* gegen ξ. Bei jedem Schritt wird der Fehler quadriert (ergo die Anzahl der richtigen Nachkommastellen verdoppelt) und dann mit einer Konstanten multipliziert, geht also viel schneller gegen 0 als es der Fixpunktsatz alleine hergibt.

Symbolverzeichnis

Stichwortverzeichnis