Skript zur Vorlesung Analysis 1

Wintersemester 2012/2013

Prof. Dr. Benjamin Schlein

Diese Notizen sind eine Zusammenfassung der Vorlesung "Analysis 1", die ich im Wintersemester 2011-2012 an der Universität Bonn gehalten habe, und sie sind nur als Hilfsmittel für die Studenten dieser Vorlesung gemeint. Die Notizen wurden aus verschiedenen Quellen zusammengestellt, z.B. aus den folgenden Büchern:

- O. Forster. Analysis 1, Vieweg Verlag.
- S. Hildebrandt. Analysis 1, Springer Verlag.
- K. Knigsberger. Analysis 1, Springer Verlag.
- C. Blatter. Analysis 1, Springer Verlag.

Inhaltsverzeichnis

1 Grundbegriffe

1.1 Logik

In der Mathematik versucht man die Wahrheit von Aussagen zu überprüfen.
Axiom vom ausgeschlossenen Dritten: Jede Aussage ist entweder wahr oder falsch.
Triviales Beispiel: die Aussage $1 > 0$ ist wahr, die Aussage $1 > 2$ ist falsch.
Ein weniger triviales Beispiel: Die Aussage

$$a > b \;\Rightarrow\; a + 1 > b + 1$$

ist wahr, unabhängig von der Wahl der Zahlen a, b. Diese Aussage besagt nicht, ob $a > b$ wahr ist; sie besagt nur, dass die Implikation wahr ist.

Für zwei Aussagen A, B definieren wir die weitere Aussagen:

$\neg A$ ist die Verneigung von A
$A \wedge B$ ist die Aussage: A und B
$A \vee B$ ist die Aussage: A oder B (oder beide)
$A \Rightarrow B$ ist die Aussage: A impliziert B
$A \iff B$ ist die Aussage: A genau dann wenn B

Diese Aussagen werden durch die folgende Wahrheitstafel definiert (w steht für wahr, f für falsch)

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

Bemerke, noch einmal, dass die Wahrheit von $A \Rightarrow B$ nichts über die Wahrheit von A besagt. Falls A falsch ist, so ist die Aussage $A \Rightarrow B$ immer wahr.

Satz 1.1. *Es gelten die Aussagen*

$$\begin{aligned} \neg(A \wedge B) &\iff \neg A \vee \neg B \\ \neg(A \vee B) &\iff \neg A \wedge \neg B \\ A \Rightarrow B &\iff \neg B \Rightarrow \neg A\,. \end{aligned}$$

Beweis. Betrachte verschiedene Fälle. □

Es werden oft Aussagen untersucht, die von einer Variablen abhängen. Z.B. die Wahrheit der Aussage

$$n + 5 \geq n^2$$

für $n \in \mathbb{N}$, hängt von dem Wert von n ab. Die Aussage ist wahr für $n = 0, 1, 2$, sie ist falsch für $n \geq 3$. Aussagen der Form $A(n)$, die von einem Parameter n abhängen (im

Beispiel $n \in \mathbb{N}$, im Allgemeinen kann der Parameter in irgendeiner Menge sein) heissen *Aussagenformen.*

Zur Formulierung von mathematischen Aussagen werden oft die Quantoren

$$\begin{aligned}
\exists \equiv\ & \text{es existiert} \\
\exists! \equiv\ & \text{es existiert genau ein} \\
\forall \equiv\ & \text{für alle}
\end{aligned}$$

benutzt. Z.B. die Aussage

$$A(x) \quad \forall x \in \mathbb{R}$$

trifft zu, falls die Aussageform $A(x)$ wahr ist, für alle $x \in \mathbb{R}$. Die Aussage

$$\exists\, x \in \mathbb{R} : A(x)$$

trifft dagegen zu, falls ein $x \in \mathbb{R}$ existiert, so dass $A(x)$ stimmt.

Bemerkung: Es gilt

$$\neg\,(A(x)\;\forall x \in X) \iff \exists x \in X : \neg A(x)$$

1.2 Mengen

Eine Menge ist eine ungeordnete Zusammenfassung von Elementen, Objekten. Zwei Mengen sind gleich, wenn sie die selben Elemente enthalten. Man kann Mengen durch Angaben von allen Elementen oder durch eine Aussage definieren. Z.B.:

$$M = \{0, 1, 2\} \quad \text{oder} \quad M = \{n \in \mathbb{N} : n + 5 > n^2\}$$

Besonderes Beispiel: $\emptyset$ bezeichnet die leere Menge, die kein Element enthält.

Die Notation $x \in A$ bedeutet, dass x Element der Menge A ist. Die Notation $A \subset B$ bedeutet, dass A eine Teilmenge von B ist. Diese Aussage trifft zu, falls

$$x \in A \Rightarrow x \in B$$

Beachte, dass $A \subset B$ wahr ist, falls $A = B$. Bemerke auch, dass $\emptyset \subset A$, für alle Mengen A. Ferner gilt

$$A = B \iff A \subset B \;\wedge\; B \subset A$$

Konstruktion von Mengen: Seien A und B zwei Mengen. Dann definieren wir

$$\begin{aligned}
A \cup B &:= \{x : x \in A \;\vee x \in B\} = \text{ Vereinigung von } A \text{ und } B \\
A \cap B &:= \{x : x \in A \wedge x \in B\} = \text{ Durchschnitt von } A \text{ und } B \\
A \backslash B &:= \{x : x \in A \wedge x \notin B\} = \text{Differenz} \\
A \Delta B &:= (A \backslash B) \cup (B \backslash A) = \text{Symmetrische Differenz}
\end{aligned}$$

Sind alle Menge unter Betrachtung Teilmenge einer gemeinsame Menge X, so bezeichnen wir mit $A^c := X \backslash A$ das Komplement von A in X. Bemerke, dass

$$(A \cup B)^c = A^c \cap B^c \quad \text{und} \quad (A \cap B)^c = A^c \cup B^c$$

Diese Beziehungen werden als Morgan'sche Regeln bezeichnet.

Definition 1.2. *Sei M eine Menge. Dann bezeichnet*

$$P(M) = \{A : A \subset M\}$$

die Potenzmenge von M; das ist die Menge aller Teilmengen von M. $P(M)$ ist nie leer, weil sie immer die leere Menge enthält. Manchmal wird $P(M)$ auch mit 2^M bezeichnet, weil, falls $m < \infty$ die Anzahl Elementen in M bezeichnet, so enthält $P(M)$ 2^m Elemente.

Beispiel: Sei $M = \{0, 1\}$. Dann gilt $P(M) = \{\emptyset, \{0\}, \{1\}, \{0, 1\}\}$.

Für zwei Mengen A, B, definieren wir das Produkt (oder das Kreuzprodukt) der Mengen A und B als die neue Menge

$$A \times B := \{(x, y) : x \in A \ \wedge \ y \in B\}$$

Elemente der Produktmenge $A \times B$ sind geordnete Paare (x, y) wobei $x \in A$ und $y \in B$. Zwei Paare (x_1, y_1) und (x_2, y_2) sind genau dann gleich, wenn $x_1 = x_2$ und $y_1 = y_2$. Bemerke, dass $(1, 2) \neq (2, 1)$ (dagegen stimmen die zwei Mengen $\{1, 2\}$ und $\{2, 1\}$ überein). Also $A \times B \neq B \times A$. Bespiel:

$$\mathbb{R} \times \mathbb{R} = \{(x, y) : x \in \mathbb{R}, y \in \mathbb{R}\}$$

ist die euklidische Ebene.

1.3 Funktionen

Definition 1.3. *Seien M und N zwei Mengen. Eine Funktion $f : M \to N$ ist eine Abbildung, eine Vorschrift, die jedem $m \in M$ genau einen Wert $f(m) \in N$ zuordnet. M heisst der Definitionsbereich von f, N der Zielbereich. Die Bildmenge oder der Wertebereich von f ist die Menge*

$$\mathit{Ran}\ f \equiv f(M) = \{f(m) : m \in M\}$$

und besteht aus allen Werten in N, die tatsächlich angenommen werden.

Ein einfaches Beispiel ist die Identitätsfunktion. Sei M eine Menge. Dann bezeichnet $\mathrm{id}_M : M \to M$ die Funktion $\mathrm{id}_M(x) = x$ für alle $x \in M$.

Ein anderes Beispiel von Funktionen sind Folgen. Sei M eine Menge. Eine Folge auf M ist eine Abbildung $f : \mathbb{N} \to M$ definiert auf der Menge der natürlichen Zahlen. Man kann eine Folge als eine strukturierte Liste von Elementen aus M interpretieren. Man benutzt manchmal die Notation $(f_n)_{n \in \mathbb{N}}$.

Sei nun $f : M \to N$ eine Funktion. Dann definieren wir für $X \subset M$, $Y \subset N$

$$f(X) = \{f(x) \in N : x \in X\} \equiv \text{ Bild von } X$$
$$f^{-1}(Y) = \{x \in M : f(x) \in Y\} \equiv \text{ Urbild von } Y$$

Das definiert zwei Abbildungen $\phi_1 : P(M) \to P(N)$ und $\phi_2 : P(N) \to P(M)$ (ϕ_1 bildet jede Teilmenge $X \subset M$ in $f(X) \subset N$ ab, ϕ_2 bildet dagegen jede Teilmenge $Y \subset Y$ in $f^{-1}(Y)$ ab).

Beispiel: Seien $M = \{1,2,3,4,5\}$ und $N = \{1,2,3\}$. Wir definieren die Funktion $f : M \to N$ durch die Angaben $f(1) = f(2) = f(4) = 1$ und $f(3) = f(5) = 2$. f ist wohldefiniert, weil jedem $m \in M$ genau ein Wert zuordnet wird. Die Bildmenge ist $f(M) = \{1,2\}$. Weiter

$$f^{-1}(\{1\}) = \{1,2,4\}, \quad f^{-1}(\{2\}) = \{3,5\}, \quad f^{-1}(\{3\}) = \emptyset$$

Gegeben Funktionen $f : M \to N$ und $g : N \to P$ so kann man die Verknüpfung $g \circ f : M \to P$ durch die Angabe

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

definieren.

Satz 1.4. *Die Verknüpfung von Funktionen ist assoziativ, d.h.*

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

Bemerkung: Die Verknüpfung von Funktionen ist dagegen im Allgemeinen nicht kommutativ, d.h. $f \circ g \neq g \circ f$.

Definition 1.5. *Eine Funktion $f : M \to N$ heisst surjektiv, falls $f(M) = N$, d.h. falls Zielbereich und Bildmenge übereinstimmen und jeder Wert in N angenommen wird. f heisst injektiv, falls*

$$f(x) = f(y) \;\Rightarrow x = y$$

f heisst bijektiv, falls f surjektiv und injektiv ist.

Ist $f : M \to N$ bijektiv, so

$$\forall\, n \in N \quad \exists!\, m \in M \quad \text{mit } f(m) = n$$

Man kann also in diesem Fall eine neue Funktion durch $f^{-1} : N \to M$ durch die Angabe

$$f^{-1}(n) = m \quad \text{falls} \quad f(m) = n$$

definieren. Die Funktion f^{-1} heisst die Inverse von f. Sie erfüllt $f \circ f^{-1} = \mathrm{id}_M$ und $f^{-1} \circ f = \mathrm{id}_N$ (diese zwei Eigenschaften charakterisieren f^{-1} eindeutig). Beachte: Das Urbild $f^{-1}(Y)$ einer Teilmenge $Y \subset N$ ist immer wohldefiniert. Die inverse Funktion f^{-1} dagegen ist nur wohldefiniert, falls f bijektiv ist.

Sei $f : M \to N$ eine Funktion. Dann definiert

$$\mathcal{G}(f) := \{(x, f(x)) \in M \times N : x \in M\}$$

den Graphen der Funktion f. Per Definition $\mathcal{G}(f) \subset M \times N$.

1.4 Relationen

Definition 1.6. *Sei M eine Menge. Eine Relation auf M ist eine Teilmenge R von $M \times M$. Man schreibt $x \sim y$ oder $x \sim_R y$ und man sagt x sei in Relation mit y bezüglich R, falls $(x, y) \in R$.*

Beispiel: Der Graph jeder Funktion $f : M \to M$ ist eine Relation. Graphen von Funktionen sind spezielle Relationen mit der Eigenschaft, dass für alle $x \in M$, $(x, y) \in \mathcal{G}(f)$ genau für ein $y \in M$ stimmt. Im Allgemeinen ist das für Relationen nicht der Fall.

Eine wichtige Rolle spielen sogenannten Äquivalenzrelationen.

Definition 1.7. *Sei M eine Menge. Eine Teilmenge $R \subset M \times M$ heisst eine Äquivalenzrelation, falls:*

i) $x \sim_R x$ für alle $x \in M$ (R heisst in diesem Fall reflexiv).

ii) $x \sim_R y \Rightarrow y \sim_R x$, für alle $x, y \in M$ (R heisst dann symmetrisch).

iii) $x \sim_R y$ und $y \sim_R z \Rightarrow x \sim_R z$, für alle $x, y, z \in M$ (R heisst dann transitiv).

Beispiel: Sei M eine Menge. Dann ist $R = \{(x, x) : x \in M\}$ eine Äquivalenzrelation.

Definition 1.8. *Sei R eine Äquivalenzrelation auf einer Menge M. Für $m \in M$ definiert dann*

$$[m] := \{x \in M : x \sim_R m\}$$

die Äquivalenzklasse von m. Ein beliebiges $x \in [m]$ heisst ein Repräsentant der Äquivalenzklasse $[m]$. Wir definieren die Quotientenmenge von M bezüglich R als die Menge aller Äquivalenzklassen, d.h.

$$M/_\sim \; = \; \{[m] : m \in M\}$$

Satz 1.9. *Sei R eine Äquivalenzrelation auf einer Menge M. Dann gilt*

$$[x] \cap [y] \neq \emptyset \quad \Rightarrow \quad [x] = [y]$$

Es folgt, dass $M = \bigcup_{[m] \in M/_\sim} [m]$ eine disjunkte Zerlegung von M definiert.

Eine andere Art von Relationen die in der Analysis eine sehr wichtige Rolle spielen sind sogenannte Ordnungsrelationen.

Definition 1.10. *Eine Relation R auf M heisst eine Ordnungsrelation, falls*

i) $x \sim_R x$ für alle $x \in M$ (reflexiv).

ii) $x \sim_R y$ und $y \sim_R z \Rightarrow x \sim_R z$ für alle $x, y, z \in M$ (transitiv).

iii) $x \sim_R y$ und $y \sim_R x \Rightarrow x = y$ für alle $x, y \in M$ (identiv).

Ist R eine Ordnungsrelation, so schreiben wir $x \preceq y$ statt $x \sim_R y$.

Eine Menge M, versehen mit einer Ordnungsrelation $\preceq$, heisst eine geordnete Menge. Eine Ordnung auf M heisst total, falls, für alle $x, y \in M$ entweder $x \preceq y$ oder $y \preceq x$ gilt.

Beispiele: $(\mathbb{R}, \leq)$ ist eine total geordnete Menge. Für eine beliebige Menge M ist $(P(M), \subset)$ eine geordnete, aber im Allgemeinen nicht total geordnete Menge.

Definition 1.11. *Sei $(M, \preceq)$ eine geordnete Menge, $A \subset M$, $A \neq \emptyset$. $m \in A$ heisst maximal in A, falls*

$$(x \in A) \wedge (m \preceq x) \Rightarrow x = m$$

$m \in A$ heisst minimal in A, falls

$$(x \in A) \wedge (x \preceq m) \Rightarrow x = m$$

Bemerke, dass, im Allgemeinen, maximale und minimale Elemente weder existieren noch eindeutig sein müssen. Ist aber die Ordnung auf M total, so sind maximale und minimale Elemente, falls sie existieren, eindeutig. In der Tat, sind $x_1, x_2 \in A$ zwei maximale Elemente von $A \subset M$ für eine total geordnete Menge M, so muss entweder $x_1 \preceq x_2$ oder $x_2 \preceq x_1$ gelten. Gilt $x_1 \preceq x_2$, so finden wir durch Anwendung der Tatsache, dass x_1 maximal ist, $x_2 = x_1$. Gilt dagegen $x_2 \preceq x_1$, so folgt aus der Maximalität von x_2, dass $x_1 = x_2$. Also $x_1 = x_2$ muss auf jedem Fall gelten. Analog kann man die Eindeutigkeit von minimalen Elementen zeigen.

Definition 1.12. *Sei $(M, \preceq)$ eine geordnete Menge, $A \subset M$, $A \neq \emptyset$. $m \in M$ heisst eine obere Schranke für A, falls $x \preceq m$ für alle $x \in A$. $m \in M$ heisst eine untere Schranke für A, falls $m \preceq x$ für alle $x \in A$. Die Menge A heisst nach oben (bzw. unten) beschränkt, falls eine obere (bzw. untere) Schranke existiert.*

Bemerke, obere und untere Schranken für eine Menge A, im Gegensatz zu maximalen und minimalen Elementen, müssen nicht in der Menge A enthalten sein.

Sei nun $(M, \preceq)$ eine total geordnete Menge und $A \subset M$ eine nach oben beschränkte Menge. Sei

$$B = \{x \in M : x \text{ ist obere Schranke für } A\}$$

Nach Voraussetzung $B \neq \emptyset$. Enthält B ein minimales Element $x \in B$, so heisst x das Supremum von A und wird mit $x = \sup A$ bezeichnet. Aus der Eindeutigkeit von maximalen und minimalen Elementen (in total geordneten Mengen), erhalten wir, dass das $\sup A$, falls es existiert, eindeutig ist. Im Allgemeinen muss das Supremum $x = \sup A$ nicht zur Menge A gehören (per Definition gehört $\sup A$ zur Menge der obere Schranken von A). Gehört $x = \sup A$ zu der Menge A, dann heisst x das Maximum von A und wird mit $\max A$ bezeichnet.

Analog kann man das Infimum von A, bezeichnet mit $\inf A$, als die grösste untere Schranke von A definieren. Ist $\inf A \in A$, so heisst $\inf A$ das Minimum von A, bezeichnet mit $\min A$.

2 Zahlen

2.1 Die natürlichen Zahlen

Wir kennen die natürlichen Zahlen vom Zählen. Alle Eigenschaften der natürlichen Zahlen lassen sich aus der Nachfolge-Abbildung $\nu(n) = n + 1$ herleiten. Das ist die Idee der Peano-Axiome, die die Einführung der natürlichen Zahlen formalisieren.

Definition 2.1 (Peano Axiome)**.** *Die natürlichen Zahlen bilden eine Menge* $\mathbb{N}$ *mit einem ausgezeichneten Element* 0 *und mit einer Abbildung* $\nu : \mathbb{N} \to \mathbb{N}$ *mit den folgenden Eigenschaften:*

P1) ν *ist injektiv und* $0 \notin$ *Ran* ν*.*

P2) Prinzip der vollständigen Induktion: Ist $A \subset \mathbb{N}$*, s.d.* $0 \in A$ *und, s.d.*

$$n \in A \;\Rightarrow\; \nu(n) \in A$$

so ist $A = \mathbb{N}$*.*

Die Zahlen $\nu(0), \nu(\nu(0)), \ldots$ *werden als* $1, 2, \ldots$ *bezeichnet.*

Alle Eigenschaften von $\mathbb{N}$ lassen sich aus den Peano-Axiomen herleiten. Z.B. können wir auf $\mathbb{N}$ eine Ordnung definieren: Wir sagen $k \leq m$, falls $m = k$ oder falls m durch wiederholte Anwendung von ν auf k erreicht werden kann. Es ist einfach, sich zu überzeugen, dass dann $\leq$ eine (totale) Ordnungsrelation definiert. Aus den Axiomen folgt, dass $0 \leq n$, $\forall\, n \in \mathbb{N}$. Auch die Summe auf $\mathbb{N}$ lässt sich durch wiederholte Anwendung von der Abbildung ν definieren.

Satz 2.2 (Wohlordnungsprinzip)**.** *Jede nicht leere Teilmenge* $M \subset \mathbb{N}$ *hat ein Minimum.*

Beweis. Wir betrachten die Menge

$$A = \{k \in \mathbb{N} : k \leq m, \quad \forall\, m \in M\}$$

der unteren Schranken von M. Wir wissen, dass $0 \in A$, weil $0 \leq m$, $\forall m \in \mathbb{N}$. Aus Annahme wissen wir, dass M nicht leer ist; d.h. $\exists\, m_0 \in M$. Dann gilt $m_0 + 1 \notin A$. Also $A \neq \mathbb{N}$. Das Prinzip der vollständigen Induktion besagt, dass

$$A \neq \mathbb{N} \;\Rightarrow\; (0 \notin A) \;\vee\; \neg((n \in A) \;\Rightarrow\; (n+1) \in A)$$

Da aber $0 \in A$, erhalten wir

$$\neg((n \in A) \;\Rightarrow\; (n+1) \in A) \quad \Longleftrightarrow \quad \exists n_0 \in A \,:\, (n_0 + 1) \notin A$$

Wir behaupten nun, n_0 ist das gesuchte Minimum von M. Um diese Behauptung zu zeigen, bemerken wir, dass

$$(n_0 + 1) \notin A \;\Rightarrow\; \exists p \in M : p < n_0 + 1 \quad \Longleftrightarrow \quad \exists p \in M : p \leq n_0$$

Da aber $n_0 \in A$, muss $n_0 \leq p$. Also $n_0 = p \in M$. Das impliziert, dass $k \leq n_0$ für alle $k \in A$. D.h. n_0 ist ein maximales Element von A. D.h. $n_0 = \inf M$. Da $n_0 \in M$ ist also $n_0 = \min M$. □

Die Peano-Axiome liefern eine nützliche Methode, um die Wahrheit von Aussagenformen über $\mathbb{N}$ zu beweisen.

Proposition 2.3 (Induktiver Beweis)**.** *Es sei $A(n)$ eine Aussageform über $\mathbb{N}$. Trifft $A(0)$ zu und gilt die Aussage*

$$A(n) \;\Rightarrow\; A(n+1) \quad \textit{für alle } n \in \mathbb{N}$$

so trifft $A(n)$ für alle $n \in \mathbb{N}$ zu.

Beweis. Sei $B = \{n \in \mathbb{N} : A(n) \text{ trifft zu}\}$. Wegen (P2) folgt aus

$$(0 \in B) \;\wedge\; (n \in B \;\Rightarrow\; (n+1) \in B)\,,$$

dass $B = \mathbb{N}$. □

Der Beweis von $A(0)$ heisst die Induktionsverankerung. Der Beweis der Implikation $A(n) \Rightarrow A(n+1)$ ist der Induktionschritt.

Verallgemeinerungen: Induktive Beweise können auch mit $n = n_0 > 0$ beginnen. D.h. es gilt:

$$A(n_0) \;\wedge\; (A(n) \;\Rightarrow\; A(n+1) \text{ für alle } n \geq n_0) \;\Rightarrow\; A(n) \text{ für alle } n \geq n_0$$

Bei dem Induktionsschritt kann man auch die Gültigkeit aller Aussagen $A(k)$ für $k \leq n$ annehmen. D.h.

$$A(0) \;\wedge\; (A(0) \wedge \cdots \wedge A(n) \;\Rightarrow\; A(n+1) \text{ für alle } n \in \mathbb{N}) \;\Rightarrow\; A(n) \text{ für alle } n \in \mathbb{N}$$

Beispiele:

- Für $n \geq 1$ sei $A(n)$ die Aussage, dass

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

  Um $A(n)$ für alle $n \in \mathbb{N}$ zu zeigen, genügt es zu bemerken, dass $A(1)$ zutrifft und dass $A(n) \Rightarrow A(n+1)$ für alle $n \in \mathbb{N}$ zutrifft. Die Verankerung ist klar. Um den Schritt zu zeigen, nehmen wir an, dass $A(n)$ zutrifft und wir zeigen $A(n+1)$. Dazu bemerken wir, dass

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

  wobei wir die Annahme $A(n)$ in der zweite Gleichheit benutzt haben.

- Wir behaupten, jedes $n \in \mathbb{N}$ kann als Produkt von Primzahlen geschrieben werden. Um diese Behauptung zu beweisen, definieren wir für $n \in \mathbb{N}\backslash\{0\}$, $A(n)$ als die Aussage, dass n als Produkt von Primzahlen geschrieben werden kann. $A(1)$ trifft offenbar zu (weil 1 eine Primzahl ist). Wir nehmen nun an, dass $A(1) \wedge \cdots \wedge A(n)$ zutrifft (d.h., dass $A(k)$ für alle $k \leq n$ zutrifft) und wir zeigen, dass $A(n+1)$ zutrifft. Dazu unterscheiden wir zwei Fälle: Ist $(n+1)$ eine Primzahl, so sind wir fertig. Ist dagegen $(n+1)$ keine Primzahl, so existieren $p, q \in \mathbb{N}$ mit $n+1 = p \cdot q$. Da aber $p, q < n$ sind, folgt aus der Induktionsannahme, dass p und q als Produkt von Primzahlen faktorisiert werden können. Aus $n+1 = p \cdot q$ können wir dann auch n als Produkt von Primzahlen schreiben.

Das Prinzip der vollständigen Induktion kann man benutzen, um induktive (oder rekursive) Definitionen zu formulieren. $n!$, gennant n Fakultät, wird als die Multiplikation aller ganzen Zahlen zwischen 1 und n definiert, d.h.

$$n! = 1 \cdot 2 \cdot 3 \cdot \dots \cdot n$$

mit der Konvention, dass $0! = 1$. Dieselbe Grösse kann man auch induktiv wie folgt definieren: Man setzt $0! = 1$ und $(n+1)! = (n+1) \cdot n!$ für alle $n \in \mathbb{N}$. D.h., $(n+1)!$ ist durch Verwendung von $n!$ definiert, $n!$ durch Verwendung von $(n-1)!$, usw. Die Tatsache, dass in dieser Weise $n!$ wirklich für jede $n \in \mathbb{N}$ definiert wird, ist eine Folgerung aus dem Prinzip der vollständigen Induktion.

Proposition 2.4 (Rekursive Definitionen). *Sei S eine Menge (im Beispiel oben $S = \mathbb{N}$) und $s_0 \in S$ (im Beispiel $s_0 = 1$). Für jedes $n \in \mathbb{N}$ sei $F_n : S \to S$ eine Funktion (oben $F_n(m) = (n+1) \cdot m$). Dann existiert genau eine Funktion $f : \mathbb{N} \to S$ mit den Eigenschaften $f(0) = s_0$ und $f(n+1) = F_n(f(n))$ (im Beispiel $f(n) = n!$).*

Z.B. für eine Folge $(a_n)_{n\in\mathbb{N}}$ mit Werten in irgendeiner Menge, wo eine Addition definiert ist, definieren wir die Summe $\sum_{j=0}^{n} a_j$ für alle $n \in \mathbb{N}$ durch die rekursive Definition

$$\sum_{j=0}^{0} a_j := a_0 \quad \text{und} \quad \sum_{j=0}^{n+1} a_j := \sum_{j=0}^{n} a_j + a_{n+1}$$

Ähnlich für eine Folge $(a_j)_{j\in\mathbb{N}}$ mit Werten in irgendeiner Menge, wo eine Multiplikation definiert ist, definieren wir $\prod_{j=0}^{n} a_j$ durch die rekursive Definition

$$\prod_{j=0}^{0} a_j := a_0 \quad \text{und} \quad \prod_{j=0}^{n+1} a_j := \prod_{j=0}^{n} a_j \cdot a_{n+1}$$

2.2 Die rationalen Zahlen

Problem bei den natürlichen Zahlen: Summe und Multiplikation können im Allgemeinen nicht invertiert werden. Gegeben $n, m \in \mathbb{N}$ ist es i.A. nicht möglich, $x, y \in \mathbb{N}$ mit $n = m + x$ und $n = m \cdot y$ zu finden. Aus diesem Grund erweitert man die Menge $\mathbb{N}$ und betrachtet grössere Zahlensysteme.

Man führt zunächst die Menge der ganzen Zahlen $\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}$ ein. Auf $\mathbb{Z}$ ist die Summe invertierbar. Die Multiplikation ist aber immer noch nicht invertierbar. So geht man weiter und führt die Menge der rationalen Zahlen $\mathbb{Q}$ ein.

Damit man auf $\mathbb{Q}$ die Multiplikation invertieren kann, muss $\mathbb{Q}$ Zahlen der Form p/q für alle $p \in \mathbb{Z}$ und $q \in \mathbb{Z}\backslash\{0\}$ enthalten ($x = p/q$ bezeichnet die Lösung der Gleichung $p = x \cdot q$). Man könnte also überlegen, $\mathbb{Q}$ als den Kreuzprodukt $\mathbb{Z} \times (\mathbb{Z}\backslash\{0\})$ einzuführen. Das Problem damit ist aber, dass die Darstellung p/q nicht eindeutig ist (weil z.B. die Gleichungen $p = xq$ und $2p = x2q$ die selbe Lösung haben). Das bedeutet, wir müssen einige Elemente in $\mathbb{Z} \times (\mathbb{Z}\backslash\{0\})$ miteinander identifizieren. Auf der Menge $\mathbb{Z} \times (\mathbb{Z}\backslash\{0\})$ definieren wir dazu die Relation

$$(p_1, q_1) \sim (p_2, q_2) \ :\Longleftrightarrow \ p_1 \cdot q_2 = p_2 \cdot q_1$$

Es ist einfach zu überprüfen, dass $\sim$ eine Äquivalenzrelation definiert. Die Äquivalenzklasse

$$[(p,q)] = \{(r,s) \in \mathbb{Z} \times (\mathbb{Z}\backslash\{0\}) : rq = sp\}$$

enthält genau alle Paare (r,s) mit der Eigenschaft $p/q = r/s$, also alle Paare, die dieselbe rationale Zahl darstellen sollen. Deswegen definieren wir

$$\mathbb{Q} = (\mathbb{Z} \times (\mathbb{Z}\backslash\{0\}))/_\sim \; = \{[p/q] : p \in \mathbb{Z}, q \in \mathbb{Z}\backslash\{0\}\}$$

Auf $\mathbb{Q}$ kann man dann eine Ordnung sowie eine Addition, eine Multiplikation, eine Subtraktion (die Inverse der Addition) und eine Division (die Inverse der Multiplikation) definieren. Alle bekannten Rechenregeln in $\mathbb{Q}$ werden dann aus dem folgenden Theorem zusammengefasst.

Theorem 2.5. $\mathbb{Q}$ *ist ein geordneter Körper.*

Das Theorem werden wir nicht beweisen, aber wir wollen zumindest seine Bedeutung erklären.

Definition 2.6. *Eine Menge K, versehen mit einer Addition $+ : K \times K \to K$ und einer Multiplikation $\cdot : K \times K \to K$ heisst ein Körper, falls:*

A1) $x + y = y + x \;\forall\, x, y \in K$ *(Kommutativität).*

A2) $x + (y + z) = (x + y) + z \;\forall\, x, y, z \in K$ *(Associativität).*

A3) $\exists\, 0 \in K$ *(Null) mit* $x + 0 = x \;\forall\, x \in K$ *(Existenz eines neutralen Elements).*

A4) $\forall\, x \in K \;\exists\, y \in K$ *mit* $x + y = 0$.

M1) $x \cdot y = y \cdot x$, $\forall\, x, y \in K$ *(Kommutativität).*

M2) $x \cdot (y \cdot z) = (x \cdot y) \cdot z \;\forall\, x, y, z \in K$ *(Associativität).*

M3) $\exists$ *ein Element* $1 \neq 0$ *so dass* $x \cdot 1 = x \;\forall\, x \in K$ *(Existenz eines neutrales Elements).*

M4) $\forall x \neq 0 \;\exists y \in K$ *mit* $x \cdot y = 1$ *(Existenz der Inverse).*

D) $x \cdot (y + z) = x \cdot y + x \cdot z$, $\forall x, y, z \in K$ *(Distributivität).*

Bemerkungen:

- Die neutralen Elemente sind eindeutig. Gibt es in der Tat $0, 0' \in K$ mit $a + 0 = a$ und $a + 0' = a$ für alle $a \in K$, so gilt insbesondere $0' + 0 = 0'$, $0 + 0' = 0$. Aus der Kommutativität folgt, dass $0 = 0'$ (ähnlich ist 1 eindeutig).

- Für beliebige $x \in K$ ist die additive Inverse eindeutig. Existieren in der Tat $y_1, y_2 \in K$ mit $x + y_1 = x + y_2 = 0$, so muss

$$y_1 = y_1 + 0 = y_1 + (x + y_2) = (y_1 + x) + y_2 = (x + y_1) + y_2 = 0 + y_2 = y_2$$

  Die eindeutige $y \in K$ mit $x + y = 0$ bezeichnet man mit $-x$. Für beliebige $a, b \in K$ definieren wir dann die Subtraktion $a - b := a + (-b)$.

- Ähnlich zeigt man die Eindeutigkeit der multiplikativen Inverse zu jeder $x \in K\backslash\{0\}$. Die eindeutige multiplikative Inverse zu $x \in K\backslash\{0\}$ bezeichnet man mit $1/x$. Für beliebige $a, b \in K$, mit $b \neq 0$ definieren wir dann die Division $a/b = a \cdot (1/b)$.

- Die Rechenregeln $-(-a) = a$, $1/(1/a) = a$ (für $a \neq 0$), $0 \cdot a = 0$, $(-1) \cdot a = -a$, $(-1) \cdot (-1) = 1$, $a \cdot (b - c) = ab - ac$, $a \cdot b = 0 \Rightarrow a = 0 \vee b = 0$ folgen alle aus den Axiomen von Körpern.

Definition 2.7. *Ein geordneter Körper ist ein Körper K, versehen mit einer totalen Ordnung, bezeichnet mit $\leq$, mit folgenden Kongruenzeigenschaften*

O1) $a \leq b \Rightarrow a + c \leq b + c$, *für alle* $a, b, c \in K$.

O2) $(a \leq b) \wedge (0 \leq c) \;\Rightarrow\; ac \leq bc$, *für alle* $a, b, c \in K$.

Die Notation $a \geq b$ bedeutet $b \leq a$, $a < b$ bedeutet $a \leq b$ und $a \neq b$, $a > b$ bedeutet $b < a$.

Die folgenden bekannten Rechenregeln für Ungleichungen folgen aus den Axiomen von geordneten Körpern (Beweis: Übung). Seien $a, b, c \in K$. Dann

- $a \leq b \iff a + c \leq b + c$.

- $a < b \iff a + c < b + c$.

- $a < b \iff 0 < b - a$. Insbesondere $a < 0 \iff 0 < -a$.

- $a > 0 \,\wedge\, b > 0 \;\Rightarrow\; ab > 0$.

- $a \neq 0 \;\Rightarrow\; a^2 > 0$.

- $1 > 0$.

- $a > 0 \;\Rightarrow\; 1/a > 0$.

Auf geordneten Körpern kann man den absoluten Betrag

$$|a| := \begin{cases} a & \text{falls } a \geq 0 \\ -a & \text{falls } a < 0 \end{cases}$$

definieren. Aus der Rechenregel folgt, dass $|a| \geq 0$ für alle $a \in K$.

Aus Theorem 2.5 folgt, dass alle beschriebenen Rechenregeln auf $\mathbb{Q}$ gelten.

2.3 Die reellen Zahlen

Auf $\mathbb{Q}$ sind Addition und Multiplikation invertierbar. Das bedeutet, es ist immer möglich auf $\mathbb{Q}$ lineare Gleichungen zu lösen. Mit anderen Worten, für beliebige $a \in \mathbb{Q}\backslash\{0\}$ und $b \in \mathbb{Q}$ gibt es genau eine Lösung $x \in \mathbb{Q}$ der Gleichung $ax + b = 0$. Schon bei quadratischen Gleichungen ist aber die Lage mit rationalen Zahlen nicht so befriedigend, wie bei linearen Gleichungen.

Satz 2.8. $\sqrt{2} \notin \mathbb{Q}$, *d.h. es existiert keine* $x \in \mathbb{Q}$ *mit* $x^2 = 2$.

Beweis. Der Satz wird durch Widerspruch bewiesen. Wir nehmen an, es existiert $x \in \mathbb{Q}$ mit $x^2 = 2$. Dann existiert $q \in \mathbb{N}\backslash\{0\}$ mit $q \cdot x \in \mathbb{Z}$. Setze

$$M = \{q \in \mathbb{N}\backslash\{0\} : q \cdot x \in \mathbb{Z}\}$$

Da $M \neq \emptyset$, impliziert das Wohlordnungprinzip (Satz 2.2), dass M ein Minimum hat. Sei $g \in M$ das minimale Element und $p = g \cdot x \in \mathbb{Z}$. Dann gilt $x = p/g$. Die Minimalität von g impliziert, dass p und g Teilerfremde sind. Es gilt

$$2 = p^2/g^2 \;\Rightarrow\; p^2 = 2g^2$$

Das impliziert, dass p^2 gerade ist. Deswegen muss auch p gerade sein (weil $(2n+1)^2 = 4n^2 + 4n + 1$ immer ungerade ist). Das bedeutet, dass ein $m \in \mathbb{Z}$ mit $p = 2m$ existiert. Dann

$$4m^2 = (2m)^2 = p^2 = 2g^2 \quad \Rightarrow \quad g^2 = 2m^2$$

Also ist auch g gerade. Das ergibt einen Widerspruch zu der Tatsache, dass g, p teilerfremd sind (oder, äquivalent, ein Widerspruch zur Minimalität von g in M). □

Die rationalen Zahlen haben also bei $\sqrt{2}$ eine "Lücke". Es ist einfach, sich zu überzeugen, dass die rationalen Zahlen viele Lücken haben. Wegen dieser Lücken gibt es Schwierigkeiten mit $\mathbb{Q}$. Eine erste Schwierigkeit ist, dass einfache quadratische Gleichungen, wie $x^2 = 2$ (und viele Gleichungen höherer Ordnung) keine Lösung haben. Ausser dieser algebraischen Schwierigkeit gibt es auch analytische Schwierigkeiten mit den rationalen Zahlen; wir werden sie diskutieren, wenn wir den Begriff von Konvergenz von Folgen betrachten werden. Es gibt also mehrere Gründe, um die rationalen Zahlen zu erweitern und die "Lücken" zu füllen. Dazu führen wir die reellen Zahlen ein.

Definition 2.9. *Eine totalgeordnete Menge $(M, \preceq)$ heisst ordnungsvollständig, falls die folgende Bedingung erfüllt ist: Sind $A, B \subset M$ mit $a \preceq b$ für alle $a \in A$ und $b \in B$, dann existiert $c \in M$ mit $a \preceq c$ für alle $a \in A$ und $c \preceq b$ für alle $b \in B$.*

Bemerkung: $\mathbb{Q}$ ist nicht ordnungsvollständig. Mit $A = \{x \in \mathbb{Q} : x > 0 \text{ und } x^2 < 2\}$ und $B = \{x \in \mathbb{Q} : x > 0 \text{ und } x^2 > 2\}$, existiert kein $c \in \mathbb{Q}$ mit $a \leq c \leq b$ für alle $a \in A$ und $b \in B$. Würde so ein c existieren, dann wäre $c = \sqrt{2}$.

Satz 2.10. *Es existiert ein ordnungsvollständig geordneter Körper und dieses Objekt ist eindeutig bis auf den Isomorphismus.*

In Satz 2.10 bedeutet der Ausdruck "eindeutig bis auf den Isomorphismus", dass falls $(K_1, +_1, \cdot_1, \leq_1)$ und $(K_2, +_2, \cdot_2, \leq_2)$ zwei ordnungsvollständig geordnete Körper sind, so müssen sie isomorph sein. Zwei geordnete Körper $(K_1, +_1, \cdot_1, \leq_1), (K_2, +_2, \cdot_2, \leq_2)$ heissen isomorph, falls eine Bijektion $\phi : K_1 \to K_2$ existiert, mit

$$\phi(a +_1 b) = \phi(a) +_2 \phi(b), \quad \phi(a \cdot_1 b) = \phi(a) \cdot_2 \phi(b), \quad a \leq_1 b \;\Rightarrow\; \phi(a) \leq_2 \phi(b)$$

D.h. falls eine Bijektion ϕ existiert, die alle relevanten Strukturen von K_1 auf den entsprechenden Strukturen auf K_2 abbildet.

Wir werden in der Vorlesung Satz 2.10 nicht beweisen. Ein Beweis kann zB. in dem Buch von Blatter gefunden werden.

Definition 2.11. *Wir bezeichnen mit $\mathbb{R}$ den eindeutig ordnungsvollständig geordneten Körper.*

Bemerkung: Man kann $\mathbb{N}$, $\mathbb{Z}$ und $\mathbb{Q}$ mit Teilmengen von $\mathbb{R}$ identifizieren. Aus den Körperaxiomen folgt in der Tat die Existenz von $0, 1 \in \mathbb{R}$. Durch wiederholte Anwendung der Addition mit 1 bekommt man dann $\mathbb{N} \subset \mathbb{R}$. Aus der Existenz der additiven Inversen folgt die Existenz von $\mathbb{Z} \subset \mathbb{R}$. Aus der Existenz der multiplikativen Inverse folgt dann die Existenz von $\mathbb{Q} \subset \mathbb{R}$. Die Ordnung auf $\mathbb{R}$ ist aus den Axiomen der geordneten Körper (siehe Def. 2.7) mit der Ordnung auf $\mathbb{Q}$ verträglich (und damit auch mit der Ordnung auf $\mathbb{N}$ und $\mathbb{Z}$).

Seien $a, b \in \mathbb{R}$, mit $a < b$. Wir benutzen die Notationen

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

Proposition 2.12 (Supremumprinzip)**.** *Jede nicht leere nach oben beschränkte Teilmenge $A \subset \mathbb{R}$ hat ein Supremum.*

Beweis. Sei $A \subset \mathbb{R}$ nicht leer und nach oben beschränkt. Sei

$$B = \{b \in \mathbb{R} : a \le b \text{ für alle } a \in A\} = \{\text{Menge aller oberen Schranken von } A\}$$

Dann gilt $a \le b$ für alle $a \in A$ und alle $b \in B$. Die Ordnungsvollständigkeit von $\mathbb{R}$ impliziert, dass $c \in \mathbb{R}$ mit $a \le c \le b$ für alle $a \in A$, $b \in B$ existiert. $c \ge a$ für alle $a \in A$ bedeutet, dass c eine obere Schranke von A ist, d.h. $c \in B$. $c \le b$ für alle $b \in B$ impliziert, dass c ein minimales Element von B ist. Das bedeutet, dass $c = \sup\ A$. □

Ähnlich hat jede nach unten beschränkte Teilmenge von $\mathbb{R}$ ein Infimum. Um Suprema zu berechnen, ist das folgende Kriterium oft wichtig.

Proposition 2.13. *Sei $A \subset \mathbb{R}$ nicht leer und nach oben beschränkt. Dann ist $s = \sup\ A$ genau dann, wenn $a \le s$ für alle $a \in A$ und $\forall\, \varepsilon > 0$ existiert ein $a \in A$ mit $s - \varepsilon < a \le s$.*

Beweis. Sei $s = \sup\ A$. Dann ist s eine obere Schranke von A, und also $a \le s$ für alle $a \in A$. Wir behaupten nun, dass $\forall\, \varepsilon > 0$ existiert $a \in A$ mit $s - \varepsilon < a \le s$. Wenn nicht, so würde $\varepsilon > 0$ existieren, s.d. $a \le s - \varepsilon$ für alle $a \in A$. Dann wäre $s - \varepsilon$ eine obere Schranke für A, in Widerspruch zur Tatsache, dass s die kleinste obere Schranke ist.

Nehmen wir nun an, dass für alle $a \in A$ $a \le s$ und dass $\forall \varepsilon > 0$ ein $a \in A$ mit $s - \varepsilon < a \le s$ existiert. Wir behaupten, dass $s = \sup A$. Die Bedingung $a \le s$ für alle $a \in A$ impliziert, dass s eine obere Schranke ist. Wir müssen noch zeigen, dass s die minimale obere Schranke ist. Nehmen wir an, es existiert eine obere Schranke $t < s$. Dann gilt $a \le t$ und deswegen auch $a \le s - (s - t)$ für alle $a \in A$. Setzen wir $\varepsilon = s - t > 0$, so existiert kein $a \in A$ mit $s - \varepsilon < a$, in Widerspruch zur Annahme. □

Ein ähnliches Kriterium gilt natürlich auch für das Infimum. Wir benutzen die Notationen $\sup\ A = +\infty$, falls A nach oben unbeschränkt ist und $\inf\ A = -\infty$, falls A nach unten unbeschränkt ist. Weiter schreiben wir $\sup \emptyset = -\infty$ und $\inf \emptyset = +\infty$.

Satz 2.14 (Archimedes Prinzip)**.** *$\mathbb{N} \subset \mathbb{R}$ ist nach oben unbeschränkt.*

Beweis. Nehmen wir an, $\mathbb{N} \subset \mathbb{R}$ ist nach oben beschränkt. Sei $s = \sup \mathbb{N}$. Dann existiert für jeden $\varepsilon > 0$ $n \in \mathbb{N}$ mit $s - \varepsilon < n \leq s$. Insbesondere, für $\varepsilon = 1$ finden wir $s - 1 < n$, d.h. $s < n + 1$, in Widerspruch zu der Tatsache, dass $s = \sup \mathbb{N}$. □

Korollar 2.15. *Für alle $\varepsilon > 0$ existiert $n \in \mathbb{N}\backslash\{0\}$ mit*

$$0 < 1/n < \varepsilon$$

Proposition 2.16. *Sei $x \in \mathbb{R}$. Dann $\exists!\, n \in \mathbb{Z}$ mit $n \leq x < n + 1$.*

Beweis der Existenz. Nehmen wir zunächst an $x \geq 0$. Sei $M = \{j \in \mathbb{N} : j > x\}$. Dann ist $M \neq \emptyset$. Wir setzen $m = \min \{j \in \mathbb{N} : j > x\}$. Es gilt $x < m$, und deswegen $m \geq 1$. Also $m - 1 \in \mathbb{N}$ und $(m-1) \notin M$ (sonst wäre m nicht das Minimum). Es folgt, dass $x \geq (m - 1)$. Den Fall $x < 0$ kann man ähnlich behandeln. □

Proposition 2.17. *Seien $x, y \in \mathbb{R}$, mit $x < y$. Dann existiert $r \in \mathbb{Q}$ mit $x < r < y$.*

Beweis. Sei $q \in \mathbb{N}\backslash\{0\}$ mit $0 < 1/q < (y - x)$. Dann ist $qy > qx + 1$. Sei nun $p \in \mathbb{Z}$ mit $(p - 1) \leq qx < p$. Dann ist $qy > (p - 1) + 1 = p$. Das bedeutet, dass $qx < p < qy$. Division mit q gibt $x < (p/q) < y$. □

Bemerkung: Es folgt aus Prop. 2.17, dass für jedes $x \in \mathbb{R}$ und $\varepsilon > 0$ ein $r \in \mathbb{Q}$ mit $|x - r| < \varepsilon$ existiert. Man sagt deswegen, dass $\mathbb{Q}$ dicht in $\mathbb{R}$ liegt.

Bemerkung: Gilt $x < y$, so existieren unendlich viele $r \in \mathbb{Q}$ zwischen $x, y \in \mathbb{R}$. Man kann nämlich $x < \cdots < r_3 < r_2 < r_1 < y$ rekursiv definieren.

Nun kommen wir zurück zur Frage, die die Einführung von $\mathbb{R}$ motivierte.

Proposition 2.18. *Es existiert $c \in \mathbb{R}$, $c > 0$, mit $c^2 = 2$.*

Beweis. Sei

$$A = \{x \in \mathbb{R} : (x \geq 0) \wedge (x^2 \leq 2)\}$$

A ist offenbar nach oben beschränkt. Sei $c = \sup A$. Wir behaupten, dass $c^2 = 2$. Nehmen wir an, $c^2 < 2$. Wir wählen nun $0 < \varepsilon < 1$ mit $\varepsilon < (2 - c^2)/(2c + 1)$. Dann gilt

$$(c + \varepsilon)^2 = c^2 + 2c\varepsilon + \varepsilon^2 \leq c^2 + (2c + 1)\varepsilon \leq 2\,.$$

Deswegen $(c + \varepsilon) \in A$, im Widerspruch zur Tatsache, dass c eine obere Schranke für A ist. Nehmen wir nun an, dass $c^2 > 2$. Für $0 < \varepsilon < (c^2 - 2)/2c$ gilt

$$(c - \varepsilon)^2 = c^2 - 2\varepsilon c + \varepsilon^2 > c^2 - 2\varepsilon c > 2$$

Das impliziert, dass $(c - \varepsilon)$ eine obere Schranke für A ist, in Widerspruch zur Tatsache, dass c die kleinste obere Schranke von A ist. Es folgt, dass $c^2 = 2$. □

Bemerkung (siehe Übungen): Sei $a \in \mathbb{R}$, mit $a > 0$, und $p \in \mathbb{N}\backslash\{0\}$. Dann existiert genau ein reelles $x > 0$ mit $x^p = a$.

Zu gegebenem $x \in \mathbb{R}$, existiert $n \in \mathbb{Z}$ mit $n \leq x < n+1$. n ist der ganzzahlige Teil von x. Sei nun $x_0 = x - n$. Per Definition $x_0 \in [0; 1)$. Wir zeigen nun, wie die Zahl x_0 durch eine Dezimalbruchdarstellung beschrieben werden kann. Die Bezeichnung

$$x_0 = 0.a_1 a_2 a_3 \ldots$$

für eine Folge $a : \mathbb{N} \to \{0, 1, \ldots, 9\}$ bedeutet, dass

$$x_0 = a_1 \cdot 10^{-1} + a_2 \cdot 10^{-2} + \ldots$$

(Wir werden später in der Vorlesung unendliche Summen genauer betrachten). Es gibt ein kleines Problem mit der Beschreibung von reellen Zahlen in $[0, 1)$ durch Folgen mit Werten in $\{0, 1, \ldots, 9\}$. Z.B. die Folge $(4, 9, 9, 9, \ldots)$ und die Folge $(5, 0, 0, \ldots)$ entsprechen der selben Zahl, weil

$$90 * 0.499 \cdots = 100 * 0.4999 \cdots - 10 * 0.4999 \cdots = 45$$

und deswegen $0.499 \cdots = 45/90 = 0.5$. Wir sagen eine Folge $a : \mathbb{N} \to \{0, 1, \ldots, 9\}$ ist zulässig, falls es unendlich viele $j \in \mathbb{N}$ mit $a_j \neq 9$ gibt. Dann ist die Folge $(4, 9, 9, \ldots)$ nicht zulässig, während die Folge $(5, 0, 0, \ldots)$ zulässig ist. Wir bezeichnen die Menge der zulässigen Folgen mit $\mathcal{A}$. Dann ist die Behauptung, dass jedes $x_0 \in [0, 1)$ genau einer zulässigen Folge in $\mathcal{A}$ zugeordnet werden kann.

Definition 2.19 (Dezimalbruchdarstellung)**.** *Sei $x_0 \in [0, 1)$. Wir definieren $\widetilde{x}_1 = 10x_0$ und wir bezeichnen mit a_1 den ganzzahligen Teil von $\widetilde{x}_1$. Offenbar ist $a_1 \in \{0, 1, \ldots, 9\}$. Wir setzen dann $x_1 = \widetilde{x}_1 - a_1 \in [0, 1)$. Rekursiv, für beliebige $n \in \mathbb{N}$, definieren wir $\widetilde{x}_{n+1} = 10x_n$, $a_{n+1} =$ ganzzahligen Teil von x_{n+1} und $x_{n+1} = \widetilde{x}_{n+1} - a_{n+1}$. Das definiert die Folge $A(x) = (a_1, a_2, \ldots)$. Es ist einfach, sich zu überzeugen, dass die konstruierte Folge $A(x)$ immer zulässig ist. Das heisst, wir haben eine Abbildung $A : [0, 1) \to \mathcal{A}$ definiert.*

Satz 2.20. *Die Abbildung $A : [0, 1) \to \mathcal{A}$, die die Dezimalbruchdarstellung von reellen Zahlen in $[0, 1)$ definiert, ist bijektiv.*

2.4 Abzählbare und überabzählbare Mengen

Wie kann man die Kardinalität von unendlichen Mengen vergleichen?

Definition 2.21. *Wir sagen, zwei Mengen A und B haben die gleiche Kardinalität (sind gleichmächtig), falls eine Bijektion $\phi : A \to B$ existiert.*

Ist A eine nicht leere endliche Menge, so existiert eine $n \in \mathbb{N}$ und eine Bijektion $\phi : A \to \{1, 2, \ldots, n\}$ (das kann als Definition für eine endliche Menge genommen werden). Die Zahl n ist dann eindeutig bestimmt und wird als die Kardinalität von A bezeichnet.

Definition 2.22. *Eine Menge A heisst unendlich abzählbar, falls sie mit $\mathbb{N}$ gleichmächtig ist. A heisst abzählbar, falls sie endlich oder unendlich abzählbar ist.*

Eine Bijektion $\mathbb{N} \to A$ kann als eine Aufzählung der Elemente von A interpretiert werden. Abzählbare Mengen sind Mengen, deren Elemente in einer strukturierten Liste organisiert werden können.

Beispiele:

- Die Abbildung $\phi : \mathbb{N} \to \{n^2 : n \in \mathbb{N}\}$ bildet $\mathbb{N}$ auf der Menge der Quadratzahlen bijektiv ab. Deswegen ist die Menge der Quadratzahlen mit $\mathbb{N}$ gleichmächtig.
- $\mathbb{Z}$ ist abzählbar. $0, -1, 1, -2, 2, -3, 3, \ldots$ ist eine Aufzählung der Elemente von $\mathbb{Z}$. Mit anderen Worten, die Funktion $f : \mathbb{N} \to \mathbb{Z}$, die durch $f(2n) = n$ und $f(2n-1) = -n$ definiert wird, ist eine Bijektion.

Wir diskutieren nun einige wichtige Rechenregeln für Abzählbarkeit.

Proposition 2.23. *Alle Teilmengen von* $\mathbb{N}$ *sind abzählbar.*

Beweis. Sei $A \subset \mathbb{N}$. Ist A nach oben beschränkt, so ist A endlich. Ist A nach oben unbeschränkt, so definieren wir die Funktion $\phi : \mathbb{N} \to A$ rekursiv durch $\phi(0) = \min A$ und

$$\phi(n+1) = \min \{a \in A : a > \phi(n)\}.$$

Die Funktion ist wohldefiniert, weil jede nichtleere Teilmenge von $\mathbb{N}$ ein Minimum hat (die Teilmenge $\{a \in A : a > \phi(n)\}$ ist nicht leer, weil A unbeschränkt ist). Wir behaupten nun ϕ ist bijektiv. ϕ ist offenbar injektiv, weil $\phi(n+1) > \phi(n)$ für alle $n \in \mathbb{N}$. Wir zeigen nun die Surjektivität. Sei $a \in A$ und setze

$$B = \{n \in \mathbb{N} : \phi(n) \geq a\}$$

B ist nicht leer. Deswegen existiert $m = \min B$. Es gilt $\phi(m) \geq a$ (weil $m \in B$). Anderseits, falls $m > 0$, ist $(m-1) \in \mathbb{N}$ (ist dagegen $m = 0$, so muss $\phi(0) - \min A \leq a$; das gibt $\phi(0) = a$ und zeigt, dass $a \in \operatorname{Ran} \phi$). Da $(m-1) \notin B$ (das würde der Minimalität von m widersprechen), muss $\phi(m-1) < a$. Nach Definition von ϕ gilt aber

$$\phi(m) = \min \{b \in A : b > \phi(m-1)\} \;\Rightarrow\; \phi(m) \leq a$$

Hier benutzen wir, dass $a \in A$. Also $a = \phi(m)$ und ϕ ist injektiv. □

Verallgemeinerung: Jede Teilmenge einer abzählbaren Menge ist abzählbar.

Proposition 2.24. *Sei B abzählbar und $\psi : B \to A$ surjektiv. Dann ist A abzählbar.*

Beweis. O.B.d.A. dürfen wir annehmen, dass $B \subset \mathbb{N}$ (weil B sich bijektiv auf einer Teilmenge von $\mathbb{N}$ abbilden lässt). Für $a \in A$ definieren wir

$$\phi(a) = \min \{b \in B : \psi(b) = a\}$$

Dann ist $\phi : A \to B \subset \mathbb{N}$. ϕ offenbar injektiv. ϕ bildet also A auf der Teilmenge $B' = \operatorname{Ran} \phi \subset B$ bijektiv ab. D.h. A ist mit B' gleichmächtig. Da $B' \subset B \subset \mathbb{N}$, ist B sicher abzählbar. Also ist A auch abzählbar. □

Satz 2.25. *Das Produkt* $\mathbb{N} \times \mathbb{N}$ *ist abzählbar.*

Beweis. Es ist einfach, die Elemente von $\mathbb{N} \times \mathbb{N}$ in einer Liste zu organisieren. $\square$

Verallgemeinerung: Seien $A_1, \ldots, A_n$ abzählbare Mengen. Dann ist $A_1 \times \cdots \times A_n$ auch abzählbar.

Beweis. Wir können O.B.d.A. annehmen, dass $A_1, \ldots, A_n \subset \mathbb{N}$. Wir benutzen nun die Induktion über n. Für $n = 2$, impliziert $A_1, A_2 \subset \mathbb{N}$, dass $A_1 \times A_2 \subset \mathbb{N} \times \mathbb{N}$ sicher abzählbar ist. Nehmen wir nun an, dass $A_1 \times \cdots \times A_n$ abzählbar ist. Dann ist aber

$$A_1 \times \cdots \times A_{n+1} = (A_1 \times \cdots \times A_n) \times A_{n+1}$$

als Produkt zweier abzählbarer Mengen, nach der Überlegung für $n = 2$, abzählbar. $\square$

Eine Folgerung von Satz 2.25 ist die Abzählbarkeit von $\mathbb{Q}$.

Satz 2.26. $\mathbb{Q}$ *ist abzählbar.*

Beweis. $\mathbb{Z}, \mathbb{Z}\backslash\{0\}$ sind beide abzählbar. $\mathbb{Z} \times \mathbb{Z}\backslash\{0\}$ ist auch abzählbar. Die Abbildung $\phi : \mathbb{Z} \times \mathbb{Z}\backslash\{0\} \to \mathbb{Q}$, die durch $\phi(p, q) = p/q$ definiert ist, ist surjektiv. Also ist $\mathbb{Q}$ auch abzählbar. $\square$

Eine andere interessante Bemerkung ist, dass abzählbare Vereinigungen abzählbarer Mengen wieder abzählbar sind.

Satz 2.27. *Für alle $n \in \mathbb{N}$, sei A_n eine abzählbare Menge. Dann ist auch*

$$\bigcup_{n=1}^{\infty} A_n$$

abzählbar.

Beweis. Für jede $n \in \mathbb{N}$ existiert $B_n \subset \mathbb{N}$ und eine Bijektion $\phi_n : B_n \to A_n$. Nun setzen wir

$$B := \{(n, m) \in \mathbb{N} \times \mathbb{N} : m \in B_n\}$$

Es gilt $B \subset \mathbb{N} \times \mathbb{N}$, also ist B abzählbar. Wir definieren nun die Abbildung $\phi : B \to \bigcup_{n=1}^{\infty} A_n$ durch die Angabe $\phi(n, m) = \phi_n(m)$. Die Abbildung ist surjektiv, deswegen ist $\bigcup_{n=1}^{\infty} A_n$ abzählbar. $\square$

Anwendung: Sei

$$S_0 = \{(a_i)_{i \in \mathbb{N}} : a_i \in \mathbb{Z} \text{ für alle } i \in \mathbb{N} \text{ und, s.d. ein } n \in \mathbb{N} \text{ existiert, mit } a_i = 0 \; \forall i \geq n\}$$

die Menge der Folgen mit Werten in $\mathbb{Z}$, die nur endlich viele nicht verschwindende Elemente haben. Wir behaupten, S_0 ist abzählbar. Sei, in der Tat, $S^{(n)} = \{(a_i)_{i \in \mathbb{N}} \in S_0 \text{ mit } a_i = 0 \; \forall \; i > n\}$. Dann ist $S^{(n)}$ mit $\mathbb{Z} \times \cdots \times \mathbb{Z}$ (n Kopien) gleichmächtig. $S^{(n)}$ ist deswegen für alle $n \in \mathbb{N}$ abzählbar. Da

$$S_0 = \bigcup_{n=1}^{\infty} S^{(n)}$$

ist auch S_0 als Vereinigung abzählbar vieler abzählbarer Mengen abzählbar.

Welche Mengen sind nicht abzählbar?

Satz 2.28 (Cantor)**.** *Sei M eine Menge. Dann sind M und $P(M)$ nicht gleichmächtig.*

Beweis. Wir beweisen den Satz indirekt durch Widerspruch. Nehmen wir an $\phi : M \to P(M)$ ist eine Bijektion. Setze

$$B = \{x \in M : x \notin \phi(x)\}$$

Dann gilt $\phi(y) \neq B$ für alle $y \in M$. Ist in der Tat $y \in \phi(y)$, so ist per Definition $y \notin B$, und deswegen ist $\phi(y) \neq B$. Ist anderseits $y \notin \phi(y)$, so ist per Definition $y \in B$. Auch in diesem Fall ist $\phi(y) \neq B$. Das impliziert, dass im Widerspruch zur Annahme ϕ nicht surjektiv ist. □

Definition 2.29. *Eine nicht abzählbare Menge heisst überabzählbar.*

Z.B.: $P(\mathbb{N})$ ist eine überabzählbare Menge.

Proposition 2.30. *$\mathbb{R}$ ist überabzählbar.*

Beweis. $P(\mathbb{N})$ ist überabzählbar. Sei $\{0,1\}^{\mathbb{N}}$ die Menge der Folgen mit Werten 0 und 1. Wir definieren nun eine Abbildung $\phi : P(\mathbb{N}) \to \{0,1\}^{\mathbb{N}}$ wie folgt:

$$(\phi(X))_i = 1 \text{ falls } i \in X, \quad (\phi(X))_i = 0 \text{ sonst}$$

Es ist einfach zu sehen, dass ϕ eine Bijektion ist. Deswegen ist $\{0,1\}^{\mathbb{N}}$ überabzählbar. Sei nun $X_0 \subset [0,1)$ die Teilmenge aller reellen Zahlen in $[0,1)$ deren Dezimalbruchdarstellung nur die Zahlen 0 und 1 enthält (erinnere, die Dezimalbruchdarstellung einer reellen Zahl $x \in [0,1)$ ist eine Folge $A(x) = (a_1, a_2, \dots)$, wobei $a_j \in \{0,1,\dots,9\}$ für alle $j \in \mathbb{N}$ (und $a_j \neq 9$ für unendlich viele n)). Dann gibt es eine Bijektion $\{0,1\}^{\mathbb{N}} \to X_0$. Es folgt, dass X_0 überabzählbar ist. Da $\mathbb{R} \supset X_0$, ist auch $\mathbb{R}$ überabzählbar. □

Tatsache: $\mathbb{R}$ ist mit X_0 gleichmächtig.

Frage: Ist die Kardinalität von $\mathbb{R}$ die kleinste überabzählbare Kardinalität? Oder gibt es Teilmengen von $\mathbb{R}$, die nicht abzählbar sind, aber die nicht mit $\mathbb{R}$ gleichmächtig sind?

Kontinuumhypothese (Cantor): jede Teilmenge von $\mathbb{R}$ ist entweder abzählbar oder gleichmächtig mit $\mathbb{R}$ selbst.

P. Cohen (1960): Weder die Kontinuumhypothese noch ihre Verneigung lassen sich auf Grund der üblichen Axiome der Mengelehre beweisen.

2.5 Die komplexen Zahlen

Die Ordnungsvollständigkeit von $\mathbb{R}$ ist für die Analysis sehr nützlich. Wir werden in dieser Vorlesung meistens mit reellen Zahlen arbeiten. Manchmal gibt es aber mit $\mathbb{R}$ algebraische Schwierigkeiten, die mit der Tatsache zu tun haben, dass in $\mathbb{R}$ viele Polynome keine Nullstellen haben. Aus diesem Grund ist es manchmal nützlich, komplexe Zahlen zu betrachten.

Einer der Gründe, die reellen Zahlen einzuführen war, dass man in $\mathbb{R}$ die Gleichung $x^2 = 2$ (und allgemeiner, die Gleichung $x^p = a$, für beliebige $a > 0$ und $p \in \mathbb{N}\backslash\{0\}$) lösen kann (was nicht in $\mathbb{Q}$ möglich war). Es bleiben aber viele quadratische Gleichungen (und natürlich auch Gleichungen höherer Ordnung), die in $\mathbb{R}$ keine Lösung haben. Z.B. hat die Gleichung $x^2 = -1$ keine Lösung auf $\mathbb{R}$. Um dieses Problem zu lösen, führen wir die komplexen Zahlen ein.

Die komplexen Zahlen bilden einen Körper, den wir mit $\mathbb{C}$ bezeichnen, welcher $\mathbb{R}$ enthält, und der auch ein Element i, mit der Eigenschaft $i^2 = -1$, enthält. Es ist einfach zu sehen, dass der kleinste Körper mit diesen Eigenschaften aus allen Zahlen der Form $x + iy$, mit $x, y \in \mathbb{R}$, besteht:

$$\mathbb{C} = \{x + iy : x, y \in \mathbb{R}\}$$

versehen mit der Summe

$$(x_1 + iy_1) + (x_2 + iy_2) = (x_1 + x_2) + i(y_1 + y_2)$$

und der Multiplikation

$$(x_1 + iy_1) \cdot (x_2 + iy_2) = (x_1 x_2 - y_1 y_2) + i(x_1 y_2 + x_2 y_1)\,.$$

Formell, wird $\mathbb{C}$ als das Produkt

$$\mathbb{C} = \mathbb{R} \times \mathbb{R} = \{(x, y) : x, y \in \mathbb{R}\}$$

definiert. Summe und Multiplikation werden dann durch

$$(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2)$$

und

$$(x_1, y_1) \cdot (x_2, y_2) = (x_1 x_2 - y_1 y_2, x_1 y_2 + x_2 y_1)$$

gegeben.

Proposition 2.31. *$\mathbb{C}$, versehen mit den Operationen $+, \cdot$ ist ein Körper.*

Beweis. Die Null ist aus $(0, 0)$ gegeben. Die additive Inverse ist aus $-(x, y) = (-x, -y)$ gegeben, die Eins aus $(1, 0)$. Die einzige Eigenschaft, die nicht ganz trivial ist, ist die Existenz der multiplikativen Inverse. Aus

$$\frac{1}{x + iy} = \frac{(x - iy)}{(x + iy)(x - iy)} = \frac{x}{x^2 + y^2} - i\frac{y}{x^2 + y^2}$$

können wir raten, dass

$$(x, y)^{-1} = \left(\frac{x}{x^2 + y^2}, \frac{-y}{x^2 + y^2}\right)$$

falls $x^2 + y^2 \neq 0$ (d.h. falls $(x, y) \neq (0, 0)$). Man kann dann in der Tat zeigen, dass $(x, y)^{-1}$ die Inverse von (x, y) ist. □

Bemerkung: Es gibt auf $\mathbb{C}$ keine Ordnung (unmöglich, weil $i^2 = -1 < 0$).

Die Abbildung $\mathbb{R} \ni x \to (x, 0) \in \mathbb{C}$ identifiziert $\mathbb{R}$ mit einer Teilmenge (sogar einem Unterkörper) von $\mathbb{C}$. Wir definieren dann $i = (0, 1)$. Per Definition gilt $i^2 = (0, 1) \cdot (0, 1) =$

$(-1, 0)$. Statt (x, y) werden wir einfach $x + iy$ schreiben; diese Schreibweise hilft bei der Berechnung von Produkten von komplexen Zahlen.

Gegeben $z = x + iy$, sagt man x ist der Realteil von z, y der Imaginärteil. Sie werden mit $x = \text{Re } z$ und $y = \text{Im } z$ bezeichnet. Die komplexe Konjugierte von z ist dann aus $\bar{z} = x - iy$ gegeben. Es gelten dann die Regeln:

- $\overline{z_1 + z_2} = \bar{z}_1 + \bar{z}_2$.
- $\overline{z_1 \cdot z_2} = \bar{z}_1 \cdot \bar{z}_2$.
- $z\bar{z} = \bar{z}z = x^2 + y^2 \geq 0$ (insbesondere ist $z\bar{z} \in \mathbb{R}$).

Definition 2.32. *Der Absolutbetrag von $z = x + iy \in \mathbb{C}$ wird durch*

$$|z| = \sqrt{\bar{z}z} = \sqrt{x^2 + y^2}$$

definiert.

Es gelten die Eigenschaften:

- $|z_1 \cdot z_2| = |z_1||z_2|$.
- $|z| = 0 \iff z = 0$.
- Für $x \in \mathbb{R}$ ist $|x|_{\mathbb{C}} = |x|_{\mathbb{R}}$.
- $|\text{Re } z| \leq |z|$.
- $|\text{Im } z| \leq |z|$.
- Dreieckungleichung: $|z_1 + z_2| \leq |z_1| + |z_2|$.

Da $\mathbb{C} = \mathbb{R} \times \mathbb{R}$, lassen sich komplexe Zahlen auf der euklidischen Ebene darstellen. Die x-Achse ist die reelle Achse, die y-Achse die imaginäre Achse. Die geometrische Bedeutung der Summe ist dann einfach zu beschreiben: Die Summe $z_1 + z_2$ ist aus der vektoriellen Summe der Vektoren in $\mathbb{R}^2$, die z_1, z_2 zugeordnet sind. Die komplexe Konjugation entspricht geometrisch der Spiegelung um die x-Achse. $|z|$ ist einfach der (euklidische) Abstand zwischen z und dem Ursprung 0. Allgemeiner: $|z_1 - z_2|$ gibt den euklidischen Abstand zwischen z_1 und z_2.

Um die Multiplikation geometrisch zu interpretieren, führen wir Polarkoordinaten ein. Für $z \in \mathbb{C}$ sei $r = |z|$ der Abstand zu 0 und $\varphi \in [0, 2\pi]$ der Winkel zwischen der reellen Achse und z. Dann

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

Die Multiplikation von $z_1 = r_1(\cos\varphi_1 + i\sin\varphi_1)$ und $z_2 = r_2(\cos\varphi_2 + i\sin\varphi_2)$ ist aus

$$\begin{aligned} z_1 \cdot z_2 &= r_1 r_2 \left[(\cos\varphi_1\cos\varphi_2 - \sin\varphi_1\sin\varphi_2) + i\,(\cos\varphi_1\sin\varphi_2 + \sin\varphi_1\,\cos\varphi_2)\right] \\ &= r_1 r_2 \left[\cos(\varphi_1 + \varphi_2) + i\sin(\varphi_1 + \varphi_2)\right] \end{aligned}$$

gegeben. Also ist der Abstand von $z_1 \cdot z_2$ zum Ursprung aus der Mulitplikation der Abstände $|z_1||z_2|$ gegeben. Der Winkel zwischen $z_1 \cdot z_2$ und der x-Achse ist dagegen aus der Summe der Winkel φ_1 und φ_2 gegeben.

Wir werden nach Einführung der Exponentialfunktion sehen, dass $e^{i\varphi} = \cos\varphi + i\sin\varphi$. D.h., ist $r = |z|$ und φ der Winkel zwischen z und der x-Achse, so gilt $z = re^{i\varphi}$. Eine wichtige Eigenschaft der Exponentialfunktion ist $e^x e^y = e^{x+y}$. Das gibt eine andere, einfachere Erklärung der Tatsache, dass bei der Multiplikation von zwei komplexen Zahlen die Winkel addiert werden sollen. Die Darstellung $z = re^{i\varphi}$ erlaubt uns auch die Gleichung $w^p = z$, für gegebene $z \in \mathbb{C}$ zu lösen. Ist $z = re^{i\varphi}$, für $r > 0$ und $\varphi \in [0, 2\pi)$, so ist $w = r^{1/p}e^{i\varphi/p}$ offenbar eine Lösung (weil $(e^x)^n = e^{nx}$ für jede $n \in \mathbb{N}$, $x \in \mathbb{R}$). Da aber $(\varphi + 2\pi j)$ den selben Winkel wie φ darstellt, ist

$$w_j = r^{1/p}e^{i(\varphi+2\pi j)/p}$$

für jede $j = 0, 1, \ldots, (p-1)$ eine Lösung von der Gleichung $w^p = z$ (die Lösung mit $j = p$ ist äquivalent zur Lösung mit $j = 0$, und so weiter). D.h. für beliebige $z \in \mathbb{C}$ haben wir genau p Lösungen der Gleichung $w^p = z$ (in $\mathbb{R}$ gibt es manchmal gar keine Lösung, zB. falls $p = 2$ und $z = -1$). Allgemeiner kann man zeigen, dass jede polynomische Gleichung auf $\mathbb{C}$ mindestens eine Lösung hat (das impliziert auch, dass jede polynomische Gleichung der Ordnung n genau n Lösungen hat, falls man jede Lösung mit der korrekten Vielfachkeit zählt). Das erklärt die Nützlichkeit der komplexen Zahlen.

Satz 2.33 (Fundamentalsatz der Algebra)**.** *Jedes Polynom*

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

mit komplexen Koeffizienten, von Grad $n \geq 1$ ($a_n \neq 0$) hat mindestens eine Nullstelle in $\mathbb{C}$.

2.6 Die Vektorräume $\mathbb{R}^m$

Der Raum

$$\mathbb{R}^m = \{x = (x_1, \ldots, x_m) : x_j \in \mathbb{R} \text{ für alle } j = 1, \ldots, m\}$$

besteht aus allen m-Tupeln von reellen Zahlen. Diese Räume spielen eine extrem wichtige Rolle in der Mathematik und in Anwendungen in natürlichen Wissenschaften (z.B. die Position eines Teilchens wird durch einen Punkt $x = (x_1, x_2, x_3) \in \mathbb{R}^3$ beschrieben, ihr Zustand wird durch die Position und die Geschwindigkeit, also durch einen Punkt $(x_1, x_2, x_3, v_1, v_2, v_3) \in \mathbb{R}^6$ beschrieben).

Auf $\mathbb{R}^m$ kann man eine Addition definieren. Für $x = (x_1, \ldots, x_m), y = (y_1, \ldots, y_m) \in \mathbb{R}^m$ setzen wir

$$x + y = (x_1 + y_1, x_2 + y_2, \ldots, x_m + y_m)$$

Man kann auf $\mathbb{R}^m$ auch eine skalare Multiplikation einführen. Ist $\alpha \in \mathbb{R}$ und $x = (x_1, \ldots, x_m) \in \mathbb{R}^m$, so setzen wir

$$\alpha x = (\alpha x_1, \alpha x_2, \ldots, \alpha x_m) \in \mathbb{R}^m$$

Definition 2.34. *Eine Menge V, versehen mit einer Addition $+ : V \times V \to V$ und einer skalaren Multiplikation $\cdot : \mathbb{R} \times V \to V$ heisst ein Vektorraum über $\mathbb{R}$, falls*

- $x + y = y + x$ *für alle* $x, y \in V$..
- $(x + y) + z = x + (y + z)$ *für alle* $x, y, z \in V$.
- *Es existiert* $0 \in V$ *mit* $x + 0 = x$ *für alle* $x \in V$.
- *Für alle* $x \in V$ *existiert* $y \in V$, *bezeichnet mit* $-x$, *mit* $x + y = 0$.
- $\alpha \cdot (\beta \cdot x) = (\alpha\beta) \cdot x$, *für alle* $\alpha, \beta \in \mathbb{R}$ *und alle* $x \in V$.
- $\alpha \cdot (x + y) = \alpha \cdot x + \alpha \cdot y$ *für alle* $\alpha \in \mathbb{R}$, $x, y \in V$.
- $(\alpha + \beta) \cdot x = \alpha \cdot x + \beta \cdot x$ *für alle* $\alpha, \beta \in \mathbb{R}$, $x \in V$.
- $1 \cdot x = x$ *für alle* $x \in V$.

Man kann ähnlich Vektorräume über einen allgemeinen Körper K definieren.

$\mathbb{R}^m$ ist dann ein Vektorraum über $\mathbb{R}$, für alle $m \in \mathbb{N}\backslash\{0\}$ (und ähnlich ist $\mathbb{C}^m$ ein Vektorraum über $\mathbb{C}$). Bemerke, dass, für $m \geq 2$, $\mathbb{R}^m$ kein Körper ist, weil keine Multiplikation von zwei Vektoren in $\mathbb{R}^m$ definiert ist ($\mathbb{R}^2 \simeq \mathbb{C}$ kann aber zu einem Körper gemacht werden). Auf $\mathbb{R}^m$, $m \geq 2$, ist keine Ordnung definiert.

Es ist sehr nützlich, auf $\mathbb{R}^m$ eine Funktion zu definieren, die die Länge von Vektoren misst. Für $x = (x_1, \dots, x_m) \in \mathbb{R}^m$ setzen wir

$$\|x\| = \left(\sum_{j=1}^{m} x_j^2 \right)^{1/2} \tag{1}$$

Dann hat die Funktion $\|.\| : \mathbb{R}^m \to \mathbb{R}$ die Eigenschaften

i) $\|x\| \geq 0$ für alle $x \in \mathbb{R}^n$, mit $\|x\| = 0$ genau dann, wenn $x = 0$.

ii) $\|\alpha x\| = |\alpha| \|x\|$ für alle $\alpha \in \mathbb{R}$ und $x \in \mathbb{R}^m$.

iii) $\|x + y\| \leq \|x\| + \|y\|$ für alle $x, y \in \mathbb{R}^m$.

Eine Funktion $\|.\|$ mit den Eigenschaften i)-ii)-iii) heisst eine Norm. Ein Vektorraum, in welchem eine Norm definiert ist, heisst ein normierter Raum. $\mathbb{R}^m$, versehen mit der Norm (1), ist ein Beispiel eines normierten Raumes. Bemerkung: Die Definition der Norm ist nicht eindeutig. Z.B. auf $\mathbb{R}^m$ definiert

$$\|x\|_p = \left[\sum_{j=1}^{m} |x_j|^p \right]^{1/p}$$

für alle $p > 1$ eine Norm.

3 Konvergenz von Folgen

Nachdem wir jetzt die nötigen Werkzeuge eingeführt haben, beginnen wir die wichtigen Begriffe der Analysis zu betrachten. In diesem Kapitel untersuchen wir Folgen und ihre Konvergenz. Erinnere, dass eine Folge auf einer Menge M eine Funktion $a : \mathbb{N} \to M$ ist. Wir schreiben üblicherweise a_n oder $(a_n)_{n\in\mathbb{N}}$, um die Elemente der Folge zu bezeichnen. In diesem Kapitel werden wir Folgen auf $\mathbb{R}$ betrachten, d.h. Folgen $a : \mathbb{N} \to \mathbb{R}$. Wir werden später Folgen auf allgemeineren Räumen (sogenannten metrischen Räume) untersuchen.

Beispiele: Die einfachsten Folgen sind die konstanten Folgen, definiert durch $a_n = c$, für alle $n \in \mathbb{N}$, und für ein $c \in \mathbb{R}$. Andere Beispiele sind $a_n = n$, oder $a_n = 1/n$. Manchmal werden Folgen durch Summen definiert, z.B.

$$a_n = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n}$$

oder

$$a_n = 1 + \frac{1}{2^2} + \frac{1}{3^2} + \cdots + \frac{1}{n^2}$$

Diese Folgen heissen Reihen. Wir werden Reihen detaillierter im nächsten Kapitel betrachten.

3.1 Konvergenz: Definition und elementare Eigenschaften

Einer der wichtigsten Begriffe der Analysis ist der Begriff von Konvergenz von Folgen. Betrachte z.B. die Folge $a_n = 1/n$. Falls wir n sehr gross wählen, so kommt a_n immer näher zu 0 (obwohl a_n immer verschieden von 0 bleibt). In diesem Fall sagen wir, dass a_n gegen 0 konvergiert, also n gegen Unendlich strebt. Man braucht hier eine genauere mathematische Definition. Die Definition von Konvergenz übersetzt, in der Sprache der Mathematik, die Idee, dass a_n gegen a konvergiert ($a = 0$ im Beispiel oben), falls a_n beliebig nah zu a kommt, für n genügend gross.

Definition 3.1. *Sei a_n eine Folge auf $\mathbb{R}$. Wir sagen a_n konvergiert gegen $a \in \mathbb{R}$, dass a der Limes der Folge a_n ist, oder dass a der Grenzwert von a_n ist, und wir schreiben $a_n \to a$, oder $\lim_{n\to\infty} a_n = a$, falls*

$$\text{für alle } \varepsilon > 0 \text{ existiert } n_0 = n_0(\varepsilon), \text{ s.d. } n > n_0 \ \Rightarrow\ |a_n - a| < \varepsilon\,.$$

Bemerkungen:

- Sei a_n eine Folge auf $\mathbb{R}$. Dann gilt $a_n \to a$ genau dann, wenn $a_n - a \to 0$ genau dann, wenn $|a_n - a| \to 0$. Das folgt direkt aus der Definition.

- Wir haben

  $$|a_n - a| < \varepsilon \iff a - \varepsilon < a_n < a + \varepsilon \iff a_n \in (a - \varepsilon, a + \varepsilon)$$

  Der Absolutbetrag $|a_n - a|$ misst den Abstand zwischen a_n und a. Die Bedingung $|a_n - a| < \varepsilon$ besagt also, dass der Abstand zwischen a_n und a kleiner als ε ist. Da

$\varepsilon > 0$ beliebig klein sein kann, bedeutet $a_n \to a$, dass der Abstand zwischen a_n und a kleiner als eine beliebige Fehlergrenze ist, falls wir n gross genug wählen. Wie gross n sein muss, wird aus n_0 bestimmt. Bemerke, dass n_0 von ε abhängt. Um so kleiner $\varepsilon > 0$ gewählt wird, um so grösser wird (typischerweise) n_0 sein.

Beispiele:

- Sei $a_n = c$ für alle n (a_n ist eine konstante Folge). Dann gilt $a_n \to c$. In diesem Fall gilt $|a_n - c| < \varepsilon$ für alle $n \in \mathbb{N}$ und alle $\varepsilon > 0$ (d.h. man kann immer $n_0 = 1$ wählen).

- Sei $a_n = 1/n$. Dann $a_n \to 0$.

  Beweis. Sei $\varepsilon > 0$ fest gewählt. Nach Satz 2.14 existiert $n_0 \in \mathbb{N}\backslash\{0\}$ mit $n_0 > (1/\varepsilon)$. Für $n > n_0$ gilt
  $$0 \leq a_n = \frac{1}{n} < \frac{1}{n_0} < \varepsilon$$
  Das impliziert $|a_n| < \varepsilon$, für alle $n > n_0$. □

- Sei $a_n = n^{1/n}$. Wir behaupten, dass $a_n \to 1$.

  Beweis. Sei $\varepsilon > 0$ fest. Dann gilt
  $$(1+\varepsilon)^n = \sum_{j=0}^{n} \binom{n}{j} \varepsilon^j = 1 + n\varepsilon + \frac{n(n-1)}{2}\varepsilon^2 + \cdots + \varepsilon^n \geq \frac{n(n-1)}{2}\varepsilon^2 = n\left[\frac{(n-1)\varepsilon^2}{2}\right]$$
  Ist $(n-1)\varepsilon^2/2 > 1$, so gilt
  $$(1+\varepsilon)^n > n \;\Rightarrow\; n^{1/n} < (1+\varepsilon)$$
  Da für alle $n \in \mathbb{N}$ ist $n^{1/n} > 1$ folgt, dass
  $$1 < n^{1/n} < 1 + \varepsilon$$
  falls $(n-1)\varepsilon^2/2 > 1$. Also
  $$|n^{1/n} - 1| < \varepsilon$$
  falls $(n-1)\varepsilon^2/2 > 1$. Um diese Bedingung zu erfüllen, wählen wir einfach $n > 1 + (2/\varepsilon^2)$. Zusammenfassend setzen wir $n_0(\varepsilon) = 1 + (2/\varepsilon^2)$. Dann impliziert $n > n_0$, dass $|n^{1/n} - 1| < \varepsilon$. Da $\varepsilon > 0$ beliebig ist, folgt die Behauptung. □

- Sei $a_n = (-1)^n$. Die Folge oszilliert zwischen 1 und -1. Sie konvergiert nicht.

- Die Folge $a_n = n$ wird immer grösser, da n gegen Unendlich strebt. Also konvergiert sie nicht.

Die erste wichtige Eigenschaft von Konvergenz ist die Eindeutigkeit des Grenzwertes.

Proposition 3.2. *Jede Folge hat höchstens einen Grenzwert.*

Beweis. Nehmen wir an, $a_n \to a$ und $a_n \to a'$, mit $a, a' \in \mathbb{R}$ und $a \neq a'$. Sei $\varepsilon = |a-a'|/2$. Für $a_n \to a$, existiert n_0 mit $|a_n - a| < \varepsilon$ für alle $n > n_0$. Für $a_n \to a'$, existiert auch n_1 mit $|a_n - a'| < \varepsilon$ für alle $n > n_1$. Also, für $n > \max\{n_0, n_1\}$, gilt $|a_n - a| < \varepsilon$ und $|a_n - a'| < \varepsilon$. Dann muss aber

$$|a - a'| = |a - a_n + a_n - a'| \leq |a - a_n| + |a_n - a'| < 2\varepsilon = |a - a'|$$

was unmöglich ist. $\square$

Eine andere wichtige Eigenschaft von konvergenten Folgen auf $\mathbb{R}$ ist ihre Beschränktheit.

Definition 3.3. *Eine Folge a_n in $\mathbb{R}$ heisst nach unten beschränkt, falls $b \in \mathbb{R}$ existiert, mit $b \leq a_n$ für alle $n \in \mathbb{N}$. Sie heisst nach oben beschränkt, falls $b \in \mathbb{R}$ existiert, mit $a_n \leq b$ für alle $n \in \mathbb{N}$. Sie heisst beschränkt, falls sie nach unten und nach oben beschränkt ist, d.h. falls $b > 0$ existiert, mit $|a_n| < b$ für alle $n \in \mathbb{N}$.*

Proposition 3.4. *Jede konvergente Folge ist beschränkt.*

Beweis. Sei a_n eine Folge, mit $a_n \to a$. Per Definition existiert mit $\varepsilon = 1$ $n_0 \in \mathbb{N}$ mit $|a_n - a| < 1$ für alle $n > n_0$. Das impliziert, dass

$$|a_n| = |a + (a_n - a)| \leq |a| + |a_n - a| \leq |a| + 1$$

für alle $n > n_0$. Also

$$|a_n| \leq \max\{|a_1|, |a_2|, \dots, |a_{n_0}|, |a| + 1\}$$

für alle $n \in \mathbb{N}$. $\square$

Ordnungsrelationen (aber nicht strikte Ordnungsrelationen) werden durch den Limes erhalten.

Lemma 3.5. *Seien a_n und b_n zwei Folgen, mit $a_n \to a$ und $b_n \to b$ und mit $a_n \leq b_n$ für alle $n \in \mathbb{N}$ (eigentlich genügt es, dass ein $n_0 \in \mathbb{N}$ existiert, mit $a_n \leq b_n$ für alle $n > n_0$). Dann gilt $a \leq b$.*

Beweis. Nehmen wir an $a > b$. Dann setzen wir $\varepsilon = (a-b)/2 > 0$. Für $a_n \to a$, existiert $n_1 \in \mathbb{N}$ mit $|a_n - a| < \varepsilon$ für alle $n > n_1$. Für $b_n \to b$, existiert $n_2 \in \mathbb{N}$ mit $|b_n - b| < \varepsilon$ für alle $n > n_2$. Für $n \geq \max\{n_1, n_2\}$ gilt also

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

in Widerspruch zur Annahme $a_n \leq b_n$ für alle $n \in \mathbb{N}$. $\square$

Bemerkung: Gilt $a_n \to a$, $b_n \to b$ und $a_n < b_n$ für alle $n \in \mathbb{N}$, so muss $a < b$ nicht gelten. Im Allgemeinen kann man aus $a_n < b_n$ für alle $n \in \mathbb{N}$ nur folgen, dass $a \leq b$ (z.B. $a_n = 1/n > 0$ für alle $n \in \mathbb{N}$, aber $\lim_{n\to\infty} a_n = 0$).

Wir untersuchen nun die Beziehung zwischen Limes und den Körperoperationen, die auf $\mathbb{R}$ definiert sind.

Proposition 3.6. *Nehmen wir an, a_n und b_n sind zwei Folgen in $\mathbb{R}$, mit $a_n \to a$ und $b_n \to b$. Dann gilt:*

i) $a_n + b_n \to a + b$.

ii) $a_n - b_n \to a - b$.

iii) $a_n b_n \to ab$.

iv) Ist $b_n \neq 0$ für alle $n \in \mathbb{N}$, und $b \neq 0$, so gilt auch

$$a_n/b_n \to a/b$$

v) Ist $a_n \geq 0$ für alle $n \in \mathbb{N}$, $p \in \mathbb{Z}$ und $q \in \mathbb{N}\backslash\{0\}$, dann gilt

$$a_n^{p/q} \to a^{p/q}$$

Beweis. i) Sei $\varepsilon > 0$ fest. Dann existiert $n_1 \in \mathbb{N}$ s.d. $|a_n - a| < \varepsilon/2$ für alle $n > n_1$, und $n_2 \in \mathbb{N}$ s.d. $|b_n - b| < \varepsilon/2$ für alle $n > n_2$. Für $n > \max\{n_1, n_2\}$ gilt also

$$|(a_n + b_n) - (a + b)| \leq |(a_n - a) + (b_n - b)| \leq |a_n - a| + |b_n - b| < \varepsilon\,.$$

ii) Ähnlich.

iii) $a_n \to a$ impliziert, dass a_n beschränkt ist. D.h. es existiert $c > 0$ mit $|a_n| \leq c$ für alle $n \in \mathbb{N}$. Dann

$$\begin{aligned} |a_n b_n - ab| &= |a_n b_n - a_n b + a_n b - ab| \leq |a_n(b_n - b)| + |b(a_n - a)| \\ &\leq |a_n||b_n - b| + |b||a_n - a| \leq c|b_n - b| + |b||a_n - a| \end{aligned} \tag{2}$$

Sei nun $\varepsilon > 0$ fest. Dann existiert $n_1 \in \mathbb{N}$ mit $|a_n - a| < \varepsilon/(2|b|)$ für alle $n > n_1$. Weiter existiert $n_2 \in \mathbb{N}$ mit $|b_n - b| < \varepsilon/(2c)$ für alle $n > n_2$. Für $n > \max\{n_1, n_2\}$ folgt aus (2), dass $|a_n b_n - ab| < \varepsilon$.

iv) Es genügt zu zeigen, dass $1/b_n \to 1/b$ (dann folgt iv) aus iii)). Aus $b_n \to b$ folgt, dass ein $n_1 \in \mathbb{N}$ existiert, mit $|b_n - b| \leq |b|/2$ für alle $n > n_1$. Das impliziert, dass

$$|b_n| = |b + (b_n - b)| \geq |b| - |b_n - b| \geq |b| - |b|/2 = |b|/2$$

für alle $n > n_1$. Also haben wir, wieder für $n > n_1$,

$$\left|\frac{1}{b_n} - \frac{1}{b}\right| = \left|\frac{b - b_n}{b b_n}\right| = \frac{1}{|b||b_n|}\,|b_n - b| \leq \frac{2}{|b|^2}\,|b_n - b| \tag{3}$$

Sei nun $\varepsilon > 0$ fest. Dann finden wir $n_2 \in \mathbb{N}$ mit

$$|b_n - b| < \frac{|b|^2}{2}\,\varepsilon$$

für alle $n > n_2$. Wir setzen nun $n_0 = \max\{n_1, n_2\}$. Für $n > n_0$ gilt, aus (3),

$$\left|\frac{1}{b_n} - \frac{1}{b}\right| < \varepsilon$$

v) Übung.

$\square$

Beispiele: Sei

$$a_n = \frac{n^2 + 3n + 5}{3n^2 + 8n + 1}$$

Dann, nach Division mit n^2,

$$a_n = \frac{1 + (3/n) + (5/n^2)}{3 + (8/n) + (1/n^2)}$$

Für $3/n \to 0$, $5/n^2 \to 0$, $8/n \to 0$ und $(1/n^2) \to 0$, folgt aus Proposition 3.6, dass $a_n \to 1/3$. Sei nun

$$b_n = n\left(\sqrt{1 + \frac{1}{n}} - 1\right)$$

Rechenregeln können nicht direkt angewandt werden, weil n nicht konvergiert. Wir schreiben b_n um:

$$b_n = n\frac{\left(\sqrt{1 + \frac{1}{n}} - 1\right) \cdot \left(\sqrt{1 + \frac{1}{n}} + 1\right)}{\left(\sqrt{1 + \frac{1}{n}} + 1\right)} = n\frac{1 + \frac{1}{n} - 1}{\sqrt{1 + \frac{1}{n}} + 1} = \frac{1}{\sqrt{1 + \frac{1}{n}} + 1}$$

Aus den Rechenregeln folgt nun, dass $b_n \to 1/2$.

Lemma 3.7. *Seien a_n und b_n zwei Folgen, so dass $a_n \to 0$ und b_n beschränkt ist (braucht nicht zu konvergieren). Dann gilt $a_n b_n \to 0$.*

Beweis. Übung. $\square$

Proposition 3.8. *Seien a_n, b_n, c_n Folgen auf $\mathbb{R}$, mit $a_n \to a$, $b_n \to a$, und, s.d. $n_0 \in \mathbb{N}$ existiert, mit*

$$a_n \le c_n \le b_n$$

für alle $n > n_0$. Dann gilt $c_n \to a$.

Beweis. Sei $\varepsilon > 0$ fest. Dann existiert $n_1 \in \mathbb{N}$ mit $|a_n - a| < \varepsilon$ für alle $n > n_1$ und $n_2 \in \mathbb{N}$ mit $|b_n - a| < \varepsilon$ für alle $n > n_2$. Für alle $n > \max\{n_0, n_1, n_2\}$ gilt also

$$a - \varepsilon < a_n \le c_n \le b_n < a + \varepsilon$$

Das bedeutet, dass $|c_n - a| < \varepsilon$ für alle $n > n_0$. $\square$

Ein wichtiges Beispiel für konvergierende Folgen ist q^n, für $|q| < 1$. Es gilt dann $q^n \to 0$. Eigentlich konvergiert q^n gegen Null, schneller als jede Potenz von n. Mit anderen Worten: $n^p q^n \to 0$, für beliebige $p \in \mathbb{N}$.

Lemma 3.9. *Sei $|q| < 1$ und $p \in \mathbb{N}$. Dann gilt $n^p q^n \to 0$.*

Beweis. O.B.d.A $q > 0$ (sonst schreibe $q = -r$, für ein $r > 0$; aus $n^p r^n \to 0$ folgt auch $n^p(-r)^n = (-1)^n n^p r^n \to 0$ aus Lemma 3.7). Wir betrachten zunächst den Fall $p = 1$. Aus $q < 1$ folgt, dass $1/q > 1$. D.h. wir finden $r > 0$ mit $1/q = 1 + r$. Dann ist $q = 1/(1 + r)$ und

$$0 \leq nq^n = n\frac{1}{(1+r)^n} = \frac{n}{1 + nr + \frac{n(n-1)}{2}r^2 + \cdots + r^n} \leq \frac{2n}{n(n-1)r^2} = \frac{2}{r^2}\frac{1}{n-1} \to 0$$

Also $nq^n \to 0$. Für $p > 1$ setzen wir $r = q^{1/p}$. Dann gilt $0 < r < 1$, $q = r^p$ und $q^n = r^{np}$. Das impliziert, dass

$$n^p q^n = n^p r^{np} = (nr^n)^p$$

Für $nr^n \to 0$, folgt aus Anwendung von Prop. 3.6, dass $n^p q^n \to 0$. Für $p = 0$ schreiben wir

$$q^n = \frac{1}{n} nq^n$$

Für $nq^n \to 0$ und $1/n \to 0$, es folgt, dass $q^n \to 0$. □

3.2 Monotone Folgen, Limessuperior und Limesinferior

Definition 3.10. *Eine Folge a_n heisst monoton wachsend, falls $m > n$ impliziert, dass $a_m \geq a_n$. Sie heisst streng monoton wachsend, falls $m > n$ impliziert, dass $a_m > a_n$. Sie heisst monoton fallend, falls $n > m$ impliziert, dass $a_n \leq a_m$ und streng monoton fallend, falls $n > m$ impliziert, dass $a_n < a_m$.*

Proposition 3.11. *Eine monoton wachsende Folge ist genau dann konvergent, wenn sie nach oben beschränkt ist. Eine monoton fallende Folge ist genau dann konvergent, wenn sie nach unten beschränkt ist.*

Beweis. Wir wissen schon, dass jede konvergente Folge beschränkt ist (nach oben und nach unten beschränkt). Zu zeigen bleibt, dass jede nach oben beschränkte monoton wachsende Folge und jede nach unten beschränkte monoton fallende Folge konvergiert. Sei z.B. a_n monoton wachsend mit $a_n \leq b$ für alle $n \in \mathbb{N}$ und für ein $b \in \mathbb{R}$. Die Menge

$$A = \{a_n : n \in \mathbb{N}\} \subset \mathbb{R}$$

ist nach oben beschränkt. Also existiert $a = \sup A$. Dann gilt $a_n \leq a$ für alle $n \in \mathbb{N}$. Für gegebenes $\varepsilon > 0$ existiert weiter (aus der Charakterisierung des Supremum in Proposition 2.13) ein $n_0 \in \mathbb{N}$ mit

$$a - \varepsilon < a_{n_0} \leq a$$

Da aber a_n monoton wachsend ist, muss

$$a - \varepsilon < a_{n_0} \leq a_n \leq a$$

für alle $n \geq n_0$. Das impliziert, dass $|a_n - a| \leq \varepsilon$ für alle $n \geq n_0$. Da $\varepsilon > 0$ beliebig ist, folgt, dass $a_n \to a$. Ähnlich zeigt man, dass jede nach unten beschränkte monoton fallende Folge konvergiert. □

Wir definieren nun den Begriff von Limessuperior und Limesinferior. Sei dazu a_n eine beschränkte Folge mit $|a_n| \leq b$ für alle $n \in \mathbb{N}$. Dann setzen wir

$$\begin{aligned} \overline{a}_n &:= \sup \ \{a_j : j \geq n\} \quad \text{ und} \\ \underline{a}_n &:= \inf \ \{a_j : j \geq n\} \end{aligned}$$

Es gilt $-b \leq \underline{a}_n \leq \overline{a}_n \leq b$ für alle $n \in \mathbb{N}$. Also sind $\overline{a}_n$ und $\underline{a}_n$ beschränkte Folgen. Weiter ist $\overline{a}_n$ monoton fallend und $\underline{a}_n$ monoton wachsend. Es folgt, dass $\overline{a}_n$ und $\underline{a}_n$ konvergieren.

Definition 3.12. *Der Limessuperior einer beschränkten Folge a_n wird durch*

$$\limsup_{n\to\infty} a_n := \lim_{n\to\infty} \overline{a}_n$$

definiert. Der Limesinferior einer beschränkten Folge a_n wird durch

$$\liminf_{n\to\infty} a_n := \lim_{n\to\infty} \underline{a}_n$$

definiert. Limessuperior und Limesinferior existieren immer für beschänkte Folgen.

Besipiel: Sei $a_n = (-1)^n(1+1/n)$. Dann gilt $\overline{a}_n = (1+1/n)$, falls n gerade und $\overline{a}_n = (1+1/(n+1))$, falls n ungerade. Es folgt $\limsup_{n\to\infty} a_n = 1$. Analog $\underline{a}_n = -(1+1/n)$, falls n ungerade und $\underline{a}_n = -(1+1/(n+1))$, falls n gerade. Also $\liminf_{n\to\infty} a_n = -1$.

Satz 3.13. *Sei a_n eine beschränkte Folge. Dann gilt*

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

Die Folge a_n ist genau dann konvergent, falls

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

In diesem Fall gilt auch

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

Beweis. Sei $\overline{a}_n = \sup\{a_j : j \geq n\}$, $\underline{a}_n = \inf\{a_j : j \geq n\}$. Da $\overline{a}_n \geq \underline{a}_n$ für alle $n \in \mathbb{N}$, folgt aus Lemma 3.5, dass

$$\liminf_{n\to\infty} a_n = \lim_{n\to\infty} \underline{a}_n \leq \lim_{n\to\infty} \overline{a}_n = \limsup_{n\to\infty} a_n\,.$$

Zweiter Teil: Übung. □

Definition 3.14. *Sei $A(n)$ eine Aussageform über $\mathbb{N}$. Wir sagen $A(n)$ gilt für fast alle $n \in \mathbb{N}$ falls sie falsch ist, für höchstens endlich viele $n \in \mathbb{N}$, d.h. falls $\{n : \neg A(n)\}$ eine endliche Menge ist.*

Bemerkung: Mit dem Begriff von Aussagen $A(n)$ die für fast alle $n \in \mathbb{N}$ gelten, kann man die Definition von Konvergenz von Folgen umformulieren. Sei a_n eine Folge auf $\mathbb{R}$. Dann gilt $a_n \to a$ genau dann, wenn für alle $\varepsilon > 0$, $|a_n - a| < \varepsilon$ für fast alle $n \in \mathbb{N}$.

Die folgende Charakterisierung vom Limessuperior ist nützlich:

Satz 3.15. *Sei a_n eine beschränkte Folge. Dann gilt $\overline{a} = \limsup_{n\to\infty} a_n$ genau dann, wenn für alle $\varepsilon > 0$, $a_n \leq \overline{a} + \varepsilon$ für fast alle $n \in \mathbb{N}$ und es gibt unendlich viele $n \in \mathbb{N}$ mit $a_n > \overline{a} - \varepsilon$.*

Beweis. Nehmen wir zunächst an, $\overline{a} = \limsup_{n\to\infty} a_n$. Dann gilt $\overline{a}_n = \sup\{a_j : j > n\} \to \overline{a}$. Daher existiert, für alle $\varepsilon > 0$, ein $n_0 \in \mathbb{N}$ mit $|\overline{a}_n - \overline{a}| < \varepsilon$ für alle $n > n_0$. Insbesondere

$$a_n \leq \sup\{a_j : j \geq n\} < \overline{a} + \varepsilon$$

für alle $n > n_0$. Das bedeutet, dass $a_n < \overline{a} + \varepsilon$ für fast alle $n \in \mathbb{N}$. Wir möchten noch zeigen, dass $a_n > \overline{a} - \varepsilon$ für unendlich viele $n \in \mathbb{N}$. Wäre das nicht der Fall, so müsste $n_1 \in \mathbb{N}$ existieren, mit $a_n \leq \overline{a} - \varepsilon$ für alle $n > n_1$. Das würde implizieren, dass

$$\overline{a}_n = \sup\{a_j : j > n\} \leq \overline{a} - \varepsilon$$

für alle $n > n_1$ in Widerspruch zur Tatsache, dass $\overline{a}_n \to \overline{a}$.

Anderseits nehmen wir an, dass für alle $\varepsilon > 0$, $a_n \leq \overline{a}+\varepsilon$ für fast alle $n \in \mathbb{N}$ und es gibt unendlich viele $n \in \mathbb{N}$ mit $a_n > \overline{a}-\varepsilon$. Wir behaupten dann, dass $\overline{a} = \limsup_{n\to\infty} a_n$. Um die Behauptung zu zeigen, bemerken wir, dass $a_n \leq \overline{a} + \varepsilon$ für fast alle $n \in \mathbb{N}$ impliziert, dass ein $n_1 \in \mathbb{N}$ mit $a_n \leq \overline{a} + \varepsilon$ für alle $n > n_1$ existiert. Dann ist aber

$$\overline{a}_n = \sup\{a_j : j > n\} \leq \overline{a} + \varepsilon$$

für alle $n > n_1$. Andererseits impliziert die Annahme, dass $a_n > \overline{a}-\varepsilon$ für unendlich viele n, dass $\overline{a}_n = \sup\{a_j : j \geq n\} > \overline{a}-\varepsilon$ für alle $n \in \mathbb{N}$. Wir haben gezeigt, dass $|\overline{a}_n - \overline{a}| < \varepsilon$ für alle $n > n_1$; d.h. $\overline{a}_n \to \overline{a}$. □

Bemerkung: Eine analoge Charakterisierung gilt natürlich auch für den Limesinferior.

3.3 Häufungspunkte und Teilfolgen

Definition 3.16. *Sei a_n eine Folge in $\mathbb{R}$, nicht notwendigerweise beschränkt. Wir sagen $b \in \mathbb{R}$ ist einen Häufungspunkt von a_n falls, für alle $\varepsilon > 0$, $|a_n - b| < \varepsilon$ für unendlich viele $n \in \mathbb{N}$.*

Bemerkung: Die Definition von Häufungspunkt sollte mit der Definition von Limes verglichen werden. b ist Limes von a_n falls $|a_n - b| < \varepsilon$ für fast alle $n \in \mathbb{N}$, b ist dagegen Häufungspunkt, falls $|a_n - b| < \varepsilon$ für unendlich viele $n \in \mathbb{N}$.

Beispiel: Die Folge $a_n = (-1)^n$ konvergiert nicht, sie hat aber die Häufungspunkte ± 1.

Es folgt aus Satz 3.15, dass für jede beschränkte Folge $\limsup_{n\to\infty} a_n$ ein Häufungspunkt ist. Analog ist $\liminf_{n\to\infty} a_n$ auch ein Häufungspunkt (die zwei könnten natürlich übereinstimmen, falls die Folge konvergiert).

Korollar 3.17. *Jede beschränkte Folge auf $\mathbb{R}$ besitzt mindestens einen Häufungspunkt. Weiter, eine beschränkte Folge konvergiert genau dann, wenn sie nur einen Häufungspunkt besitzt.*

Häufungspunkte, also insbesondere Limessuperior und Limesinferior, sind Grenzwerte von Teilfolgen.

Definition 3.18. *Sei a_n eine Folge. Eine Teilfolge von a_n ist eine Folge der Form a_{n_j}, wobei $1 \leq n_1 < n_2 < \dots$. Mit anderen Worten: eine Teilfolge der Folge $a : \mathbb{N} \to \mathbb{R}$ ist eine Folge der Form $a \circ \phi : \mathbb{N} \to \mathbb{R}$, wobei $\phi : \mathbb{N} \to \mathbb{N}$ eine strikt monotone Abbildung ist (d.h. $\phi(n_1) < \phi(n_2)$ falls $n_1 < n_2$).*

Beispiel: Ist a_n eine Folge, so ist a_{2n} die Teilfolge, die aus den geraden Elementen von a_n besteht.

Proposition 3.19. *Sei a_n eine Folge mit $a_n \to a$. Dann konvergiert jede Teilfolge von a_n auch gegen a.*

Beweis. Sei $\varepsilon > 0$ fest gewählt. Dann existiert n_0 mit $|a_n - a| < \varepsilon$ für alle $n > n_0$. Also gilt für $i > n_0$, $n_i \geq i > n_0$ und deswegen $|a_{n_i} - a| < \varepsilon$. Das impliziert, dass $a_{n_i} \to a$. □

Bemerkung: Ist a_n nicht konvergent, so könnten trotzdem konvergente Teilfolgen existieren. Z.B. $a_n = (-1)^n$ konvergiert nicht, aber $a_{2n} \to 1$ und $a_{2n+1} \to -1$.

Proposition 3.20. *Sei a_n eine Folge auf $\mathbb{R}$ (nicht unbedingt beschränkt). Dann ist b ein Häufungspunkt von a_n genau dann, wenn eine Teilfolge $n_j : \mathbb{N} \to \mathbb{N}$ mit $a_{n_j} \to b$ existiert.*

Beweis. Sei b ein Häufungspunkt. Wir wählen $n_1 \in \mathbb{N}$ mit $|a_{n_1} - b| < 1$ und dann rekursiv $n_j > n_{j-1}$ mit $|a_{n_j} - b| < 1/j$ (das ist sicher möglich, weil es unendlich viele n mit $|a_n - b| < 1/j$ existieren). Dann ist $j \to a_{n_j}$ eine Teilfolge von a_n und offenbar $a_{n_j} \to b$. Anderseits, nehmen wir an, dass eine Teilfolge a_{n_j} mit $a_{n_j} \to b$ existiert. Wir behaupten dann, dass b ein Häufungspunkt von a_n ist. In der Tat, für $a_{n_j} \to b$, gibt es, für ein beliebiges $\varepsilon > 0$, ein $j_0 \in \mathbb{N}$, mit $|a_{n_j} - b| < \varepsilon$ für alle $j > j_0$. Das bedeutet, dass $|a_n - b| < \varepsilon$ für alle $n \in \{n_{j_0+1}, n_{j_0+2}, \dots\}$. D.h. $|a_n - b| < \varepsilon$ für unendlich viele $n \in \mathbb{N}$. □

Da für jede beschränkte Folge, $\limsup_{n\to\infty} a_n$ ein Häufungspunkt ist, bekommen wir das folgende Korollar:

Korollar 3.21 (Satz von Bolzano-Weierstrasse)**.** *Jede beschränkte Folge auf $\mathbb{R}$ besitzt eine konvergente Teilfolge.*

3.4 Cauchy-Folgen

Sei a_n eine Folge auf $\mathbb{R}$, mit $a_n \to a$, und $\varepsilon > 0$. Dann existiert ein $n_0 \in \mathbb{N}$ mit $|a_n - a| < \varepsilon/2$ für alle $n > n_0$. Für $n, m \in \mathbb{N}$ mit $n, m > n_0$ gilt also

$$|a_n - a_m| \leq |a_n - a + a - a_m| \leq |a_n - a| + |a_m - a| < \varepsilon$$

Definition 3.22. *Eine Folge a_n heisst eine Cauchy-Folge, falls für alle $\varepsilon > 0$ $n_0 \in \mathbb{N}$ existiert mit*

$$n, m > n_0 \quad \Rightarrow \quad |a_n - a_m| < \varepsilon$$

Die Idee ist, dass bei Cauchy-Folgen Abstände zwischen Elementen der Folge kleiner werden, als eine beliebige Fehlergrenz $\varepsilon > 0$. Wir haben schon bewiesen, dass jede konvergente Folge eine Cauchy-Folge ist. Wir zeigen nun, dass auf $\mathbb{R}$ auch die umgekehrte Aussage gilt.

Satz 3.23 (Cauchy-Kriterium). *Eine Folge a_n auf $\mathbb{R}$ ist genau dann konvergent, wenn sie eine Cauchy-Folge ist.*

Beweis. Der Beweis ist in zwei Schritte unterteilt.

Schritt 1: a_n Cauchy impliziert, dass a_n beschränkt ist.
Beweis: Sei $\varepsilon = 1$. Dann existiert $n_0 \in \mathbb{N}$ mit $|a_n - a_m| < 1$ für alle $n, m > n_0$. Das bedeutet, dass $|a_n - a_{n_0+1}| < 1$ für alle $n > n_0$. Also

$$|a_n| = |a_{n_0+1} + (a_n - a_{n_0+1})| \leq |a_{n_0+1}| + 1$$

für alle $n > n_0$. Das impliziert, dass

$$|a_n| \leq \max\{|a_1|, \ldots, |a_{n_0}|, |a_{n_0+1}| + 1\}$$

für alle $n \in \mathbb{N}$.

Schritt 2: Sei a_n eine Cauchy-Folge, und b ein Häufungspunkt von a_n. Dann konvergiert $a_n \to b$.
Beweis: Sei $\varepsilon > 0$ fest. Da a_n eine Cauchy-Folge ist, existiert $n_0 \in \mathbb{N}$ mit $|a_n - a_m| < \varepsilon/2$ für alle $n, m > n_0$. Da b ein Häufungspunkt ist, folgt, dass $|a_n - b| < \varepsilon/2$ für unendlich viele $n \in \mathbb{N}$. Insbesondere existiert sicher $n_1 > n_0$ mit $|a_{n_1} - b| < \varepsilon/2$. Das impliziert, dass, für beliebige $n > n_0$,

$$|a_n - b| \leq |a_n - a_{n_1}| + |a_{n_1} - b| < \varepsilon$$

Also $a_n \to b$, wie behauptet.

Schritt 1 impliziert, dass a_n mindestens einen Häufungspunkt hat. Schritt 2 impliziert also, dass a_n konvergiert. □

3.5 Uneigentliche Grenzwerte

Viele Folgen, wie zum Beispiel $a_n = n$, konvergieren nicht, weil sie ins Unendliche divergieren. Es ist manchmal nützlich, die reellen Zahlen mit plus und minus unendlich zu erweitern, und Folgen wie $a_n = n$ als konvergente Folgen in den erweiterten reellen Zahlen zu betrachten.

Definition 3.24. *Wir definieren die abgeschlossene reelle Achse*

$$\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$$

Wir erweitern die Ordnung auf $\mathbb{R}$ durch

$$\begin{aligned} &a < \infty, \quad \textit{für alle } a \in \mathbb{R} \\ &a > -\infty \quad \textit{für alle } a \in \mathbb{R} \\ &-\infty < \infty \end{aligned}$$

Bemerke, dass $\overline{\mathbb{R}}$ kein Körper ist.

Definition 3.25 (Uneigentliche Grenzwerte)**.** *Die Folge a_n in $\overline{\mathbb{R}}$ konvergiert gegen $+\infty$ (man sagt auch, sie divergiert gegen $+\infty$), falls für alle $M \in \mathbb{R}$ $n_0 \in \mathbb{N}$ existiert, so dass $a_n > M$ für alle $n > n_0$. Ähnlich definiert man die Bedeutung von $a_n \to -\infty$.*

Eigenschaften von uneigentlichen Grenzwerten.

- Sei $a_n \to +\infty$, $b_n \to b \in \mathbb{R} \cup \{+\infty\}$. Dann gilt $a_n + b_n \to +\infty$. Analog, falls $a_n \to -\infty$ und $b_n \to b \in \mathbb{R} \cup \{-\infty\}$, so gilt $a_n + b_n \to -\infty$. Gilt $a_n \to +\infty$ und $b_n \to -\infty$, so gibt es keine allgemeine Rechenregel, um den Limes von $a_n + b_n$ zu bestimmen.

- Sei $a_n \to +\infty$, $b_n \to b$. Ist $b > 0$, so gilt $a_n b_n \to +\infty$. Ist $b < 0$, so gilt $a_n b_n \to -\infty$. Ist $b = 0$, so gibt es keine allgemein gültige Rechenregel, um den Limes von $a_n b_n$ zu finden.

- Sei $a_n \to \pm\infty$ (d.h. $a_n \to +\infty$ oder $a_n \to -\infty$). Dann gilt $1/a_n \to 0$.

- Eine monotone Folge in $\overline{\mathbb{R}}$ hat immer einen (eigentlichen oder uneigentlichen) Grenzwert.

- Für eine beliebige Folge a_n in $\overline{\mathbb{R}}$ definieren wir $\overline{a}_n = \sup\{a_j : j > n\} \in \overline{\mathbb{R}}$ und $\limsup_{n\to\infty} a_n = \lim_{n\to\infty} \overline{a}_n$. Ähnlich kann man auch den Limesinferior einer beliebigen Folge in $\overline{\mathbb{R}}$ definieren. Limessuperior und Limesinferior existieren immer. Eine Folge a_n konvergiert, im eigentlich oder uneigentlichen Sinne, genau dann, wenn $\limsup_{n\to\infty} a_n = \liminf_{n\to\infty} a_n$.

- Wir sagen $+\infty$ ist Häufungspunkt von a_n, falls für alle $M \in \mathbb{R}$ gilt $a_n > M$ für unendlich viele $n \in \mathbb{N}$. Jede Folge in $\overline{\mathbb{R}}$ hat dann mindestens einen (eigentlichen oder uneigentlichen) Häufungspunkt.

4 Reihen

Reihen sind ein besonderes Beispiel von Folgen, die durch Summen definiert werden.

4.1 Definition und elementare Eigenschaften

Sei $(a_n)_{n\in\mathbb{N}}$ eine Folge auf $\mathbb{R}$. Der Ausdruck $\sum a_j$ oder $\sum_j a_j$ oder $\sum_{j=0}^{\infty} a_j$ heisst eine Reihe, mit Gliedern a_n. Die Reihe ist als Grenzwert von endlichen Summen zu verstehen.

Definition 4.1. *Sei a_n eine Folge auf $\mathbb{R}$. Wir setzen, für jede $n \in \mathbb{N}$,*

$$s_n = a_0 + a_1 + \cdots + a_n = \sum_{j=0}^{n} a_j$$

s_n heisst die Folge der Partialsummen. Ist s_n konvergent, so sagen wir, dass die Reihe $\sum_{j=0}^{\infty} a_j$ konvergiert, und wir setzen

$$\sum_{j=0}^{\infty} a_j := \lim_{n\to\infty} s_n = \lim_{n\to\infty} \sum_{j=0}^{n} a_j$$

Falls $s_n \to \pm\infty$, so sagen wir dass die Reihe $\sum_{j=0}^{\infty} a_j$ divergiert gegen $\pm\infty$. Ist $\sum_{j=0}^{\infty} a_j$ weder konvergent noch nach $\pm\infty$ divergent, so sagen wir $\sum_{j=0}^{\infty} a_j$ existiert nicht. Ähnlich, können wir auch Reihen $\sum_{j=1}^{\infty} a_j$, oder allgemeiner Reihen der Form $\sum_{j=n_0}^{\infty} a_j$ definieren.

Bemerke, dass die Frage, ob eine Folge konvergiert oder nicht, per Difinition nur von dem asymptotischen Verhalten von a_n abhängt. Mit anderen Worten:

$$\sum_{j=0}^{\infty} a_j \quad \text{konvergiert, g.d.w.} \quad \sum_{j=m}^{\infty} a_j \quad \text{konvergiert}$$

für ein beliebiges $m \in \mathbb{N}$.

Beispiele:

- Die geometrische Reihe: Sei $q \in \mathbb{R}$. Dann gilt

  $$s_n = \sum_{j=0}^{n} q^j = 1 + q + q^2 + \cdots + q^n = \frac{1 - q^{n+1}}{1 - q}$$

  Ist $|q| < 1$, so konvergiert $q^{n+1} \to 0$, und deswegen $s_n \to 1/(1-q)$. Also konvergiert für $|q| < 1$ die geometrische Reihe

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

- Die harmonische Reihe: Wir untersuchen hier die Reihe

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

  Für $n \in \mathbb{N}$ setzen wir

  $$s_n = 1 + \frac{1}{2} + \cdots + \frac{1}{n} = \sum_{j=1}^{n} \frac{1}{j}$$

  Wir bemerken

  $$s_{2n} - s_n = \frac{1}{n+1} + \frac{1}{n+2} + \cdots + \frac{1}{2n} \geq n \cdot \frac{1}{2n} = \frac{1}{2}$$

  Das bedeutet, die Folge s_n kann keine Cauchy-Folge sein. Deswegen ist s_n divergent (man kann zeigen, dass s_n logarithmisch wächst). Also divergiert die harmonische Reihe.

Eine erste wichtige Bemerkung ist, dass $a_n \to 0$ eine notwendige Bedingung ist, damit die Reihe $\sum a_n$ konvergiert.

Proposition 4.2. *Ist die Reihe $\sum_{j=0}^{\infty} a_j$ konvergent, so gilt $a_n \to 0$.*

Beweis. Sei

$$s_n = \sum_{j=0}^{n} a_j$$

die Folge der Partialsummen. Die Konvergenz der Reihe bedeutet, dass s_n konvergiert. Sei $s = \lim_{n\to\infty} s_n$. Dann gilt

$$a_{n+1} = s_{n+1} - s_n \to s - s = 0$$

(weil die Folge $s_{n+1} \to s$ und weil der Grenzwert der Differenz zweier konvergenten Folgen gleich zur Differenz der Grenzwerte ist). $a_{n+1} \to 0$ impliziert auch, dass $a_n \to 0$. □

Folgerung: Für $|q| \geq 1$ konvergiert die Folge q^n nicht gegen Null. Deswegen ist in diesem Fall die Reihe $\sum_n q^n$ nicht konvergent (falls $q > 0$ divergiert die Reihe nach $+\infty$, falls $q < 0$ existiert $\sum_{j=0}^{\infty} q^j$ nicht).

Bemerkung: $a_n \to 0$ ist notwendig, aber nicht hinreichend für die Konvergenz der Reihe $\sum a_n$. Z.B. die harmonische Reihe divergiert, obwohl $1/n \to 0$.

Die folgende Proposition folgt aus Proposition 3.6 (angewandt auf der Folge der Partialsummen).

Proposition 4.3. *Seine $\sum a_n$, $\sum b_n$ konvergenten Reihen und $\alpha \in \mathbb{R}$ beliebig.*

- *Die Reihe $\sum(a_n + b_n)$ konvergiert, und*

$$\sum_{j=0}^{\infty}(a_j + b_j) = \sum_{j=0}^{\infty} a_j + \sum_{j=0}^{\infty} b_j$$

- *Die Reihe $\sum \alpha a_n$ konvergiert, und*

$$\sum_{j=0}^{\infty} \alpha a_j = \alpha \sum_{j=0}^{\infty} a_j$$

4.2 Konvergenzkriterien und Anwendungen

Satz 4.4. *Sei $a_n \geq 0$ für alle $n \in \mathbb{N}$. Dann konvergiert die Reihe $\sum_n a_n$ genau dann, wenn die Folge $s_n = a_0 + a_1 + \cdots + a_n$ der Partialsummen beschränkt ist.*

Beweis. Da $a_n \geq 0$ für alle n, ist die Folge s_n monoton steigend. Aus Proposition 3.11 konvergiert dieFolge s_n genau dann wenn sie nach oben beschränkt ist. □

Eine analoge Aussage gilt offenbar für Reihen mit $a_n \leq 0$ für alle $n \in \mathbb{N}$ (in diesem Fall ist die Folge der Partialsummen monoton fallend).

Das Cauchy-Kriterium für die Folge der Partialsummen impliziert das folgende Cauchy-Kriterium für Reihen.

Satz 4.5 (Cauchy-Kriterium für Reihen)**.** *Die Reihe* $\sum a_n$ *konvergiert genau dann, wenn*

$$\forall\, \varepsilon > 0\ \exists\, n_0 \in \mathbb{N},\ s.d.\ \left| \sum_{j=n+1}^{m} a_j \right| < \varepsilon \text{ für alle } n_0 < n < m$$

Beweis. Nach der Definition konvergiert die Reihe $\sum a_n$ g.d.w. die Folge $s_n = \sum_{j=0}^{n} a_j$ konvergiert. Aus dem Cauchy-Kriterium für Folgen, ist die Konvergenz von s_n aber äquivalent zur Bedingung, dass s_n eine Cauchy-Folge ist: Für alle $\varepsilon > 0$ existiert $n_0 \in \mathbb{N}$ mit $|s_m - s_n| < \varepsilon$. Da $|s_n - s_m| = |s_m - s_n|$, genügt es $|s_m - s_n| < \varepsilon$ für alle $n_0 < n < m$ zu überprüfen. Dann ist aber

$$|s_m - s_n| = \left| \sum_{j=n+1}^{m} a_j \right|$$

□

Definition 4.6. *Eine Reihe* $\sum a_n$ *heisst absolut konvergent, falls* $\sum |a_n|$ *konvergiert.*

Proposition 4.7. *Jede absolut konvergente Reihe ist auch konvergent. Weiter gilt (verallgemeinerte Dreiecksungleichung):*

$$\left| \sum_{j=0}^{\infty} a_j \right| \leq \sum_{j=0}^{\infty} |a_j|$$

Beweis. Sei $\sum |a_n|$ konvergent. Aus dem Cauchy-Kriterium für Reihen, existiert für alle $\varepsilon > 0$ $n_0 \in \mathbb{N}$ mit $\sum_{j=n+1}^{m} |a_j| < \varepsilon$ für alle $n_0 < n < m$. Nach der Dreiecksungleichung ist

$$\left| \sum_{j=n+1}^{m} a_j \right| \leq \sum_{j=n+1}^{m} |a_j| < \varepsilon$$

für alle $n_0 < n < m$. Also erfüllt $\sum a_n$ die Cauchy-Bedingung und konvergiert deswegen. Um die verallgemeinerte Dreiecksungleichung zu zeigen, setzen wir $s_n = \sum_{j=0}^{n} a_j$. Da $\sum a_n$ konvergiert, gilt $s_n \to s := \sum_{j=0}^{\infty} a_j$. Wir haben

$$\left| \sum_{j=0}^{\infty} a_j \right| = |s| \leq |s_n| + |s - s_n| \leq \sum_{j=0}^{n} |a_j| + |s - s_n| \leq \sum_{j=0}^{\infty} |a_j| + |s - s_n|$$

Da die linke Seite unabhängig von n ist, können wir n beliebig gross wählen. Für $n \to \infty$ konvergiert aber $|s - s_n| \to 0$. Das ergibt die gewünschte Ungleichung. □

Bemerkung: Absolute Konvergenz ist hinreichend für Konvergenz, aber nicht notwendig. Wir werden bald sehen, dass konvergente Reihen existieren, die aber nicht absolut konvergieren.

Proposition 4.8 (Majorantenkriterium)**.** *Sei* $\sum c_n$ *eine konvergente Reihe positiver Glieder, und sei* $|a_n| < c_n$ *für fast alle* $n \in \mathbb{N}$. *Dann ist* $\sum a_n$ *absolut konvergent.*

Beweis. O.B.d.A können wir annehmen, dass $|a_n| < c_n$ für alle $n \in \mathbb{N}$ (weil Konvergenznur von asymptotischen Verhalten abhängt). Die Folge $\sum |a_n|$ konvergiert genau dann wenn sie beschränkt ist. Es gilt

$$s_n = \sum_{j=0}^{n} |a_j| \leq \sum_{j=0}^{n} c_j \leq \sum_{j=0}^{\infty} c_j$$

für alle $n \in \mathbb{N}$. Da die rechte Seite unabhängig von n ist, ist die Folge s_n beschränkt. $\square$

Anwendung: Die Folge

$$\sum_{j=0}^{\infty} \frac{1}{j!}$$

konvergiert. In der Tat, für $j \geq 1$, haben wir

$$\frac{1}{(j+1)!} = \frac{1}{j+1} \frac{1}{j!} \leq \frac{1}{2} \frac{1}{j!} \leq \frac{1}{2^2} \frac{1}{(j-1)!} \leq \cdots \leq \frac{1}{2^j}$$

Also

$$\frac{1}{j!} \leq \frac{1}{2^{j-1}} = 2 \frac{1}{2^j}$$

für alle $j \geq 1$ (der Fall $j = 1$ wird separat behandelt). Da die Reihe $\sum 2(1/2)^n$ konvergiert, konvergiert wegen dem Majorantenkriterium auch $\sum 1/n!$.

Definition 4.9. *Wir setzen*

$$e := \sum_{j=0}^{\infty} \frac{1}{j!}$$

e wird als Euler'sche Zahl bezeichnet.

Bemekrung: es gilt

$$\left(1 + \frac{1}{n}\right)^n \to e$$

Beweis: Übung.

Ähnlich zum Majorantenkriterium haben wir das folgende Vergleichskriterium.

Proposition 4.10 (Vergleichskriterium)**.** *Seien $a_n, b_n > 0$ für fast alle $n \in \mathbb{N}$.*

i) Ist

$$\limsup_{n\to\infty} \frac{b_n}{a_n} < \infty$$

und $\sum a_n$ konvergent, so ist auch $\sum b_n$ konvergent.

ii) Ist

$$\liminf_{n\to\infty} \frac{b_n}{a_n} > 0$$

und $\sum a_n$ divergent, so ist auch $\sum b_n$ divergent.

Beweis. i) Für $\limsup_{n\to\infty}(b_n/a_n) < \infty$ existieren $r > 0$ und $n_0 \in \mathbb{N}$ mit $b_n < ra_n$ für alle $n > n_0$ (man kann z.B. $r = \limsup_{n\to\infty}(b_n/a_n)+1$ wählen). Da $\sum c_n$, mit $c_n = ra_n$, konvergiert, folgt aus dem Majorantenkriterium, dass $\sum b_n$ konvergiert. Der Beweis von ii) geht analog (man braucht hier ein Minorantenkriterium: Sei c_n eine Folge positiver Zahlen, s.d. $\sum c_n$ nach ∞ divergiert. Sei $|a_n| > c_n$ für fast alle $n \in \mathbb{N}$. Dann ist $\sum a_n$ nicht absolut konvergent). $\square$

Beispiel: Wir betrachten die Reihe

$$\sum_{j=0}^{\infty} \frac{1}{\sqrt{j^2+1}}$$

Wir möchten $b_n = 1/\sqrt{n^2+1}$ mit $a_n = 1/n$ vergleichen. Da

$$\frac{b_n}{a_n} = \frac{n}{\sqrt{n^2+1}} = \frac{1}{\sqrt{1+(1/n^2)}} \to 1$$

gilt

$$\liminf_{n\to\infty} \frac{b_n}{a_n} = 1 > 0$$

Da $\sum a_n$ divergiert, muss auch $\sum b_n$ divergieren.

Proposition 4.11 (Wurzelkriterium)**.** *Sei a_n eine Folge auf $\mathbb{R}$. Gilt*

$$\limsup_{n\to\infty} |a_n|^{1/n} < 1$$

so konvergiert die Reihe $\sum a_n$ absolut. Ist dagegen

$$\liminf_{n\to\infty} |a_n|^{1/n} > 1$$

so ist die Reihe $\sum a_n$ nicht absolut konvergent.

Beweis. Aus $\limsup |a_n|^{1/n} < 1$ folgt, dass $0 < q < 1$ existiert, mit $|a_n|^{1/n} \leq q$ für alle $n > n_0$. Das impliziert, dass $|a_n| \leq q^n$ für fast alle n. Die Konvergenz von $\sum q^n$ impliziert nach Majorantenkriterium die Konvergenz von $\sum |a_n|$. Die zweite Aussage kann man ähnlich zeigen. $\square$

Satz 4.12 (Quotientenkriterium)**.** *Ist $a_n \neq 0$ für fast alle $n \in \mathbb{N}$, und*

$$\limsup_{n\to\infty} \frac{|a_{n+1}|}{|a_n|} < 1$$

so konvergiert $\sum a_n$ absolut.

Beweis. Es existieren $0 < q < 1$ und $n_0 \in \mathbb{N}$ mit

$$\frac{|a_{n+1}|}{|a_n|} \leq q$$

für alle $n > n_0$. Also

$$\frac{|a_{n_0+1}|}{|a_{n_0}|} \leq q \quad \Rightarrow \quad |a_{n_0+1}| \leq q\,|a_{n_0}|$$

Ein Schritt weiter

$$\frac{|a_{n_0+2}|}{|a_{n_0+1}|} \leq q \quad \Rightarrow \quad |a_{n_0+2}| \leq q|a_{n_0+1}| \leq q^2|a_{n_0}|$$

Nach j Schritten

$$|a_{n_0+j}| \leq q^j|a_{n_0}|$$

Das gibt

$$|a_n| \leq q^{n-n_0}|a_{n_0}|$$

für alle $n > n_0$. Da $\sum q^{n-n_0}|a_{n_0}|$ konvergiert, folgt aus dem Majorantenkriterium, dass $\sum |a_n|$ konvergiert. $\square$

Beispiel: Wir betrachten die Reihe

$$\sum_{j=0}^{\infty} \frac{(j!)^2}{(2j)!}$$

Es gilt

$$\frac{a_{n+1}}{a_n} = \frac{((n+1)!)^2\,(2n)!}{(n!)^2\,(2n+2)!} = \frac{(n+1)^2}{(2n+2)(2n+1)} \to \frac{1}{4} < 1$$

Aus dem Quotientenkriterium ist also die Reihe konvergent.

In vielen Situationen kann das Quotientenkriterium nicht weiter helfen. Z.B. falls $a_n = 1/n^\alpha$, $\alpha \in \mathbb{Q}$ und positiv. Dann gilt

$$\frac{a_{n+1}}{a_n} = \frac{n^\alpha}{(n+1)^\alpha} = \left(1 - \frac{1}{n+1}\right)^\alpha \to 1$$

Wir kommen später auf dieses Beispiel zurück.

Proposition 4.13 (Teleskopsumme)**.** *Sei b_n eine Folge auf $\mathbb{R}$, und $a_n = b_n - b_{n-1}$, für alle $n \geq 1$. Dann konvergiert die Reihe $\sum a_n$ g.d.w. die Folge b_n konvergiert. In diesem Fall*

$$\sum_{j=1}^{\infty} a_j = \lim_{n\to\infty} b_n - b_0$$

Beweis. Es gilt

$$\sum_{j=1}^{n} a_j = \sum_{j=1}^{n}(b_j - b_{j-1}) = (b_1 - b_0) + (b_2 - b_1) + \cdots + (b_n - b_{n-1}) = b_n - b_0$$

$\square$

Beispiel: Sei $b_n = -1/(n+1)$ und

$$a_n = b_n - b_{n-1} = -\frac{1}{n+1} + \frac{1}{n} = \frac{1}{n(n+1)}$$

Für $b_n \to 0$, folgt, dass $\sum a_n$ konvergiert, und

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

Mit dem Vergleichskriterium folgt, dass auch $\sum_{n=1}^{\infty}(1/n^2)$ konvergiert.

Proposition 4.14 (Cauchy's Verdichtungssatz)**.** *Sei a_n eine monoton fallende Folge positiver Zahlen. Dann konvergiert $\sum_n a_n$ genau dann, wenn $\sum 2^k a_{2^k}$ konvergiert.*

Beweis. Da $a_n > 0$ für alle $n \in \mathbb{N}$ ist die Reihe $\sum a_n$ konvergent g.d.w. die Folge $s_n = \sum_{j=1}^{n} a_j$ beschränkt ist. Das ist aber äquivalent (da s_n monoton wachsend ist) zur Beschränktheit der Teilfolge

$$s_{2^k-1} = \sum_{j=1}^{2^k-1} a_j$$

Wir setzen

$$b_\ell = \sum_{j=2^{\ell-1}}^{2^\ell-1} a_j$$

Dann gilt

$$s_{2^k-1} = \sum_{j=1}^{2^k-1} a_j = \sum_{\ell=1}^{k} b_\ell$$

Da a_j monoton fallend ist, und weil b_ℓ die Summe von $2^{\ell-1}$ Gliedern ist, haben wir auch

$$2^{\ell-1} a_{2^\ell} \le b_\ell \le 2^{\ell-1} a_{2^{\ell-1}}$$

Wir erhalten, dass

$$\sum_{\ell=1}^{k} 2^{\ell-1} a_{2^\ell} \le \sum_{\ell=1}^{k} b_\ell \le \sum_{\ell=1}^{k} 2^{\ell-1} a_{2^{\ell-1}}$$

und deswegen

$$\frac{1}{2}\sum_{\ell=1}^{k} 2^\ell a_{2^\ell} \le s_{2^k-1} \le a_1 + \sum_{\ell=1}^{k-1} 2^\ell a_{2^\ell}$$

Es folgt, dass die Folge s_{2^k-1} genau dann beschränkt ist, wenn die Reihe $\sum_{\ell=1}^{\infty} 2^\ell a_{2^\ell}$ konvergiert. □

Anwendung: Sei $\alpha = p/q$ eine rationale Zahl mit $p, q \in \mathbb{N}\backslash\{0\}$. Dann betrachten wir die Folge $\sum 1/n^\alpha$. Wir wissen schon, dass die Folge divergiert für $\alpha = 1$, und dass sie konvergiert für $\alpha = 2$. Mit Hilfe von Cauchy's Verdichtungsatz, konvergiert $\sum 1/n^\alpha$ genau dann, wenn

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

konvergiert. Das ist aber genau dann der Fall, falls

$$(1/2^{\alpha-1}) < 1 \quad \Longleftrightarrow \quad 2^{\alpha-1} > 1 \quad \Longleftrightarrow \quad \alpha > 1$$

Wir haben somit gezeigt, dass die Reihe $\sum 1/n^\alpha$ genau dann konvergiert, falls $\alpha > 1$.

Alternierende Reihen sind Reihen, deren Glieder alternierende Vorzeichen haben.

Proposition 4.15 (Leibnitz)**.** *Sei a_n eine monoton fallende Folge positiver Zahlen, mit $a_n \to 0$. Dann konvergiert die alternierende Reihe*

$$\sum_{j=1}^{\infty} (-1)^{j-1} a_j = a_1 - a_2 + a_3 - a_4 + \dots .$$

Beweis. Wir zeigen zunächst, dass, für alle $n \in \mathbb{N}$,

$$0 \leq a_1 - a_2 + \dots \pm a_n \leq a_1 \tag{4}$$

In der Tat, ist $n = 2m$ gerade, so gilt

$$a_1 - a_2 + \dots - a_{2m} = (a_1 - a_2) + \dots + (a_{2m-1} - a_{2m}) \geq 0$$

und

$$a_1 - a_2 + \dots - a_{2m} = a_1 - (a_2 - a_3) - \dots - (a_{2m-2} - a_{2m-1}) - a_{2m} \leq a_1$$

weil die Folge a_n monoton fallend ist. Ist anderseits $n = 2m+1$ ungerade, so haben wir

$$a_1 - a_2 + \dots + a_{2m+1} = (a_1 - a_2) + \dots + (a_{2m-1} - a_{2m}) + a_{2m+1} \geq 0$$

und

$$a_1 - a_2 + \dots + a_{2m+1} = a_1 - (a_2 - a_3) - \dots - (a_{2m} - a_{2m+1}) \leq a_1$$

Das zeigt (4) und impliziert auch, dass

$$0 \leq a_{n+1} - a_{n+2} + a_{n+3} - \dots \pm a_m \leq a_{n+1}$$

für alle $n < m$. Wir erhalten also

$$\left| \sum_{j=n+1}^{m} (-1)^j a_j \right| \leq a_{n+1}$$

für alle $n < m$. Die Konvergenz $a_n \to 0$ impliziert also, dass die Reihe $\sum (-1)^j a_j$ die Cauchy-Bedingung erfüllt, und also, dass sie konvergiert. □

Beispiel: Die alternierende harmonische Reihe

$$\sum_{j=1}^{\infty}(-1)^{j-1}\frac{1}{j} = 1 - \frac{1}{2} + \frac{1}{3} - \dots$$

konvergiert. Wir werden später sehen, dass der Grenzwert $\log 2$ ist ($\log$ bezeichnet den natürlichen Logarithmus mit Basis e). Da $\sum(1/n)$ divergiert, ist die alternierende harmonische Reihe ein Beispiel einer konvergierenden, aber nicht absolut konvergierenden, Reihe.

Satz 4.16 (Cauchy-Schwarz Ungleichung)**.** *Seien a_n, b_n Folgen auf $\mathbb{R}$, mit $\sum_n a_n^2$ und $\sum_n b_n^2$ konvergent. Dann ist $\sum_n a_n b_n$ absolut konvergent, und*

$$\sum_{n=0}^{\infty}|a_n|\,|b_n| \leq \left(\sum_{n=0}^{\infty} a_n^2\right)^{1/2} \left(\sum_{n=0}^{\infty} b_n^2\right)^{1/2}$$

Beweis. Für jede $\alpha > 0$, haben wir

$$2|a_n||b_n| \leq \alpha a_n^2 + \alpha^{-1} b_n^2$$

Wir summieren über $n \in \mathbb{N}$ und bekommen

$$2\sum_n |a_n||b_n| \leq \alpha \sum_n a_n^2 + \alpha^{-1}\sum_n b_n^2$$

Die Wahl

$$\alpha = \frac{\left(\sum_n b_n^2\right)^{1/2}}{\left(\sum_n a_n^2\right)^{1/2}}$$

zeigt den Satz. □

4.3 Umordnungen von Reihen

Wir untersuchen in diesem Abschnitt, ob die Konvergenz und der Wert einer Reihe von der Ordnung der Glieder abhängt. Ändert sich der Wert der Glieder, falls wir sie in eine andere Ordnung summieren. Für endliche Summen wissen wir natürlich, dass die Ordnung der Summanden keine Rolle spielt. Wir werden sehen, dass das bei unendlichen Reihen nicht immer der Fall ist.

Satz 4.17 (Umordnungsatz)**.** *Sei $\sum a_n$ eine absolut konvergente Reihe und $\phi : \mathbb{N} \to \mathbb{N}$ eine Bijektion. Dann ist $\sum a_{\phi(j)}$ absolut konvergent, und*

$$\sum_{j=0}^{\infty} a_{\phi(j)} = \sum_{j=0}^{\infty} a_j$$

Notation: Sei $\sum_n a_n$ eine Reihe und $X \subset \mathbb{N}$. Dann definieren wir $\sum_{n\in X} a_n$ als die Reihe $\sum_n a_n^X$, wobei $a_n^X = a_n$, falls $n \in X$ und sonst $a_n^X = 0$.

Zum Beweis von Satz 4.17 benutzen wir das Resultat vom folgenden Lemma.

Lemma 4.18. *Sei $\sum a_n$ absolut konvergent. Dann ist auch $\sum_{n\in X} a_n$ absolut konvergent, für alle $X \subset \mathbb{N}$. Weiter, falls $X_1, X_2 \subset \mathbb{N}$ mit $X_1 \cap X_2 = \emptyset$, so gilt*

$$\sum_{n\in(X_1\cup X_2)} a_n = \sum_{n\in X_1} a_n + \sum_{n\in X_2} a_n \tag{5}$$

Beweis. Die absolute Konvergenz von $\sum_{n\in X} a_n$ folgt aus dem Majorantenkriterium, weil $|a_n^X| \leq |a_n|$ für alle $n \in \mathbb{N}$. Die Gleichung (5) folgt, weil $a_n^{X_1\cup X_2} = a_n^{X_1} + a_n^{X_2}$ für alle $n \in \mathbb{N}$ (das folgt indem man die Fälle $n \in X_1$, $n \in X_2$, $n \in (X_1 \cup X_2)^c$ separat untersucht). □

Wir kommen nun zum Beweis des Umordnungsatzes.

Beweis von Satz 4.17. Sei $\varepsilon > 0$. Da $\sum a_n$ absolut konvergent ist, finden wir $m_0 \in \mathbb{N}$ mit

$$\sum_{j=m_0+1}^{\infty} |a_j| < \varepsilon$$

Das ist möglich, weil die Folge der Partialsummen $s_n = \sum_{j=0}^{n} |a_j|$ gegen $s = \sum_{j=0}^{\infty} |a_j|$ konvergiert; das bedeutet, dass $s - s_n = \sum_{j=n+1}^{\infty} |a_j| \to 0$. Wir setzen nun

$$n_0 = \max\{\phi^{-1}(0), \phi^{-1}(1), \dots, \phi^{-1}(m_0)\}$$

Es folgt, dass

$$\{0, 1, \dots, m_0\} \subset \{\phi(0), \phi(1), \dots, \phi(n_0)\}$$

Wir behaupten nun, dass

$$\left| \sum_{j=0}^{n} a_{\phi(j)} - \sum_{j=0}^{\infty} a_j \right| < \varepsilon \tag{6}$$

für alle $n > n_0$. In der Tat, für $n > n_0$, setze

$$X_n = \{\phi(0), \phi(1), \dots, \phi(n)\}$$

Dann gilt

$$\sum_{j\in X_n} a_j = \sum_{j=1}^{n} a_{\phi(j)}$$

Aus Lemma 4.18 haben wir

$$\sum_{j=0}^{\infty} a_j = \sum_{j\in X_n} a_j + \sum_{j\in X_n^c} a_j$$

D.h.

$$\sum_{j=1}^{\infty} a_j - \sum_{j=1}^{n} a_{\phi(j)} = \sum_{j\in X_n^c} a_j$$

Das gibt, aus der verallgemeinerten Dreiecksungleichung

$$\left| \sum_{j=1}^{\infty} a_j - \sum_{j=1}^{n} a_{\phi(j)} \right| \leq \sum_{j\in X_n^c} |a_j| \leq \sum_{j=m_0+1}^{\infty} |a_j| < \varepsilon$$

Hier haben wir benutzt, dass $\{0,1,\dots,m_0\} \subset \{\phi(0),\phi(1),\dots,\phi(n_0)\} \subset X_n$, und deswegen $X_n^c \subset \{0,1,\dots,m_0\}^c = \{m_0+1, m_0+2,\dots\}$. Das zeigt (6) und impliziert, dass $\sum a_{\phi(j)}$ gegen den Grenzwert $\sum_{j=0}^\infty a_j$ konvergiert. Absolute Konvergenz folgt ähnlich durch Anwendung dieses Arguments auf der Reihe $\sum |a_{\phi(j)}|$. $\square$

Proposition 4.19 (Riemann). *Sei $\sum a_n$ eine konvergente, aber nicht absolut konvergente, Reihe und sei $\sigma \in \mathbb{R}$ beliebig. Es existiert dann eine Bijektion $\phi : \mathbb{N} \to \mathbb{N}$ mit*

$$\sum_{j=1}^\infty a_{\phi(j)} = \sigma$$

Es gibt auch Bijektionen, s.d. die Reihe $\sum_{j=0}^\infty a_{\phi(j)}$ gegen $+\infty$ oder $-\infty$ divergiert (Beweis: Übung).

Auf Grund der letzten Proposition sagt man manchmal, dass absolut konvergierende Reihen unbedingt konvergent sind, während konvergierende, aber nicht absolut konvergierende Reihe, nur bedingt konvergieren (die Konvergenz hängt in diesem Fall nämlich von der Wahl der Ordnung der Glieder ab).

Beweis. Wir definieren $a_n^+ = \max\{a_n, 0\}$ und $a_n^- = \max\{-a_n, 0\}$. Dann gilt $a_n = a_n^+ - a_n^-$ und $|a_n| = a_n^+ + a_n^-$. Wären die zwei Reihen $\sum a_n^+$ und $\sum a_n^-$ beide konvergent, so wäre auch $\sum (a_n^+ + a_n^-)$ konvergent, und $\sum a_n$ wäre absolut konvergent. Also mindestens eine der zwei Reihen $\sum a_n^+$ und $\sum a_n^-$ muss nach ∞ divergieren. Wir behaupten nun, dass beide Reihen divergieren. In der Tat, falls zB. $\sum a_n^+$ konvergiert und $\sum a_n^-$ divergiert, so würde

$$\sum_{j=0}^n a_j = \sum_{j=0}^n a_j^+ - \sum_{j=0}^n a_j^-$$

gegen $-\infty$ divergieren, im Widerspruch zur Konvergenz von $\sum a_n$.

Wir setzen

$$I_+ = \{n \in \mathbb{N} : a_n \geq 0\}$$
$$I_- = \{n \in \mathbb{N} : a_n < 0\}$$

Es gilt $I_+ \cap I_- = \emptyset$ und $I_+ \cup I_- = \mathbb{N}$. Weiter, I_+ und I_- sind beide unendliche Mengen, weil sonst $\sum_{j \in I_+} a_j = \sum a_j^+$ oder $\sum_{j\in I_-} a_j = -\sum a_j^-$ konvergieren würde.

Für gegebene $\sigma \in \mathbb{R}$ definieren wir nun rekursiv eine Funktion $\phi : \mathbb{N} \to \mathbb{N}$. Ist $\sigma \geq 0$, so setzen wir $\phi(0) = \min I_+$. Ist dagegen $\sigma < 0$, so definieren wir $\phi(0) = \min I_-$. Angenommen $\phi(0),\dots,\phi(n)$ sind schon gewählt, definieren wir $\phi(n+1)$ wie folgt. Sei

$$s_n = a_{\phi(0)} + \dots + a_{\phi(n)}$$

Gilt $s_n \leq \sigma$, so setzen wir

$$\phi(n+1) = \min\{n \in I_+ : n > \phi(j) \text{ für alle } j \text{ mit } \phi(j) \in I_+\}$$

Ist dagegen $s_n > \sigma$, so definieren wir

$$\phi(n+1) = \min\{n \in I_- : n > \phi(j) \text{ für alle } j \text{ mit } \phi(j) \in I_-\}$$

Mit anderen Worten: $\phi(n+1)$ ist das erste noch nicht benutzte Element von I_+, falls $s_n \leq \sigma$ und das erste noch nicht benutzte Element von I_- falls $s_n > \sigma$. Die Funktion ϕ ist wohldefiniert, weil I_+, I_- unendliche Mengen sind. Wir behaupten nun, dass ϕ eine Bijektion ist. ϕ ist offenbar injektiv, weil nur noch nicht benutzte Elemente von I_+, I_- benutzt werden. Um zu zeigen, dass ϕ auch surjektiv ist, müssen wir zeigen, dass für alle $n \in \mathbb{N}$ mit $\phi(n) \in I_+$ ein $m > n$ existiert mit $\phi(m) \in I_-$ und, umgekehrt, für jede $n \in \mathbb{N}$ mit $\phi(n) \in I_-$ ein $m > n$ existiert mit $\phi(m) \in I_+$. Das impliziert, dass, früher oder später, alle Elemente von I_+ und von I_- benutzt werden. Nehmen wir an $\phi(n) \in I_+$. Falls kein $m > n$ mit $\phi(m) \in I_-$ existiert, dann wäre $\phi(m) \in I_+$ für alle $m > n$. Dann wäre aber

$$s_m = a_{\phi(0)} + \cdots + a_{\phi(n-1)} + a_{\phi(n)} + \cdots + a_{\phi(m)} \leq \sigma$$

für alle $m > n$, wobei $\phi(n), \ldots, \phi(m) \in I_+$. Da aber $\sum_{j \in I_+} a_j = \infty$, und weil $a_{\phi(0)} + \cdots + a_{\phi(n-1)}$ eine von m unabhängige Konstante ist, würde dann $s_m \to \infty$ als $m \to \infty$, in Gegensatz zu $s_m \leq \sigma$ für alle $m > n$. Wir haben gezeigt, dass ϕ eine Bijektion ist.

Nun behaupten wir, dass

$$s_n = \sum_{j=1}^{n} a_{\phi(j)} \to \sigma$$

Sei $\varepsilon > 0$. Dann existiert $m_1 \in \mathbb{N}$ mit $|a_m| < \varepsilon$ für alle $m > m_1$. Wir setzen

$$n_1 = \max\{\phi^{-1}(0), \phi^{-1}(1), \ldots, \phi^{-1}(m_1)\}$$

Für $n > n_1$ ist $\phi(n) > m_1$ und deswegen $|a_{\phi(n)}| < \varepsilon$. Wir behaupten nun, dass, für alle $n > n_1$,

$$|s_n - \sigma| \leq \max\{|s_{n-1} - \sigma|, \varepsilon\} \tag{7}$$

Das folgt aus der Bemerkung, dass aus Konstruktion von ϕ nur einer der folgenden Fälle auftreten kann: 1) $s_{n-1} \leq s_n \leq \sigma$ oder $\sigma \leq \sigma_n \leq \sigma_{n-1}$, 2) $s_{n-1} \leq \sigma \leq s_n$ oder $s_n \leq \sigma \leq s_{n-1}$. Im Fall 1) gilt $|s_n - \sigma| \leq |s_{n-1} - \sigma|$. Im Fall 2) gilt dagegen $|s_n - \sigma| \leq |s_n - s_{n-1}| \leq |a_{\phi(n)}| < \varepsilon$. Sei nun $n_2 = \min\{n > n_1 : s_{n-1} \leq \sigma \leq s_n\} < \infty$. Dann $|s_{n_2} - \sigma| < \varepsilon$ und, durch wiederholte Anwendung von (7), $|s_n - \sigma| < \varepsilon$ für alle $n > n_2$. Das zeigt, dass $s_n \to \sigma$. □

4.4 Reihen mit abzählbaren Indexmengen

Sei I eine abzählbare Menge. Wir nennen eine Funktion $a : I \to \mathbb{R}$ eine Folge mit Indexmenge I (wird oft mit $(a_i)_{i \in I}$ bezeichnet). Wir untersuchen hier die Frage, ob $\sum_{i \in I} a_i$ vernünftigt definiert werden kann, und, falls ja, wie diese Summe berechnet werden kann. Ein typisches Beispiel sind Folgen $a_{n,m}$ mit Indexmengen $\mathbb{N} \times \mathbb{N}$; wann kann man $\sum_{(n,m) \in \mathbb{N} \times \mathbb{N}} a_{n,m}$ definieren, und wie kann man dann die Summe berechnen?

Definition 4.20. *Sei $(a_i)_{i \in I}$ eine Folge mit abzählbarer Indexmenge I. Existiert eine Bijektion $\psi : \mathbb{N} \to I$ mit $\sum_{n=0}^{\infty} |a_{\psi(n)}| < \infty$, dann definieren wir*

$$\sum_{i \in I} a_i := \sum_{n=0}^{\infty} a_{\psi(n)}$$

Die Definition ist in diesem Fall unabhängig von der Wahl der Bijektion ψ. Ist nämlich $\varphi : \mathbb{N} \to I$ eine andere Bijektion, so ist, aus Satz 4.17,

$$\sum_{n=1}^{\infty} a_{\varphi(n)} = \sum_{n=1}^{\infty} a_{\psi(\psi^{-1}\circ\varphi(n))} = \sum_{n=1}^{\infty} a_{\psi(n)}$$

weil $\psi^{-1} \circ \varphi : \mathbb{N} \to \mathbb{N}$ eine Bijektion ist, und weil die Folge $a_{\psi(n)}$ absolut konvergent ist. In diesem Fall heisst die Funktion $a : I \to \mathbb{R}$ (die Folge $(a_i)_{i\in I}$) summierbar.

Um Reihen mit abzählbaren Mengen I zu berechnen, ist es manchmal nützlich, I als abzählbare Vereinigung von disjunkten Mengen $I_1, I_2, \dots$ zu schreiben, separat auf den verschiedenen Indexmengen I_j zu summieren, und schlussendlich die Resultaten über j zu summieren. Dazu ist die nächste Proposition nützlich.

Proposition 4.21. *Sei I eine abzählbare Menge, $(a_i)_{i\in I}$ eine Folge mit Indexmenge I (eine Funktion auf I). Sei $I_j \subset I$, für alle $j \in \mathbb{N}$, s.d. $I = \bigcup_{j=0}^{\infty} I_j$ und $I_i \cap I_m = \emptyset$ für alle $i \neq m$. Dann gilt:*

a) Sei a_i über I summierbar. Dann a_i ist über I_n summierbar, für alle $n \in \mathbb{N}$, die Reihe

$$\sum_{n=0}^{\infty} \left(\sum_{i\in I_n} a_i \right)$$

ist absolut konvergent und

$$\sum_{n=0}^{\infty} \left(\sum_{i\in I_n} a_i \right) = \sum_{i\in I} a_i$$

b) Sei a_i summierbar über I_n, für alle $n \in \mathbb{N}$ und sei die Reihe $\sum_{n=0}^{\infty} \left(\sum_{i\in I_n} |a_i| \right)$ konvergent. Dann ist a_i summierbar über I.

Beweis. O.B.d.A. können wir annehmen, dass $I = \mathbb{N}$ (sonst finden wir eine Bijektion $\psi : \mathbb{N} \to I$ und wir ersetzen a_i durch die Folge $a_{\psi(n)}$ mit Indexmenge $\mathbb{N}$). a) Wir nehmen hier an, dass $\sum_{i=0}^{\infty} |a_i| < \infty$. Das impliziert, dass $\sum_{i\in I_n} a_i$ absolut konvergent ist, für alle $n \in \mathbb{N}$. Weiter

$$\sum_{n=0}^{m} \left| \sum_{i\in I_n} a_i \right| \leq \sum_{n=0}^{m} \sum_{i\in I_n} |a_i| = \sum_{i\in I_1\cup\dots\cup I_m} |a_i| \leq \sum_{i\in I} |a_i|$$

Die rechte Seite gibt eine obere Schranke, unabhängig aus m. Das impliziert, dass die Reihe $\sum_{n=0}^{\infty} \left(\sum_{i\in I_n} a_i \right)$ absolut konvergent ist, und, dass

$$\sum_{n=0}^{\infty} \left| \sum_{i\in I_n} a_i \right| \leq \sum_{i\in I} |a_i| .$$

Wir möchten nun zeigen, dass

$$\sum_{n=0}^{\infty} \left(\sum_{i\in I_n} a_i \right) = \sum_{i\in I} a_i$$

Dazu bemerken wir, dass

$$\left|\sum_{i=0}^{\infty} a_i - \sum_{n=0}^{m}\sum_{i\in I_n} a_i\right| = \left|\sum_{i=0}^{\infty} a_i - \sum_{i\in I_0\cup\cdots\cup I_m} a_i\right| = \left|\sum_{i\in (I_0\cup\cdots\cup I_m)^c} a_i\right| \leq \sum_{i\in (I_0\cup\cdots\cup I_m)^c} |a_i|$$

Sei nun $\varepsilon > 0$ fest gewählt. Dann existiert $n_0 \in \mathbb{N}$ mit

$$\sum_{j=n_0+1}^{\infty} |a_j| < \varepsilon$$

Wir finden dann $m_0 \in \mathbb{N}$ mit

$$\{0,1,2,\dots,n_0\} \subset I_0 \cup I_1 \cup \cdots \cup I_{m_0}$$

Dann gilt für $m > m_0$

$$\left|\sum_{i=0}^{\infty} a_i - \sum_{n=0}^{m}\sum_{i\in I_n} a_i\right| \leq \sum_{i\in (I_0\cup\cdots\cup I_m)^c} |a_i| \leq \sum_{i=n_0+1}^{\infty} |a_i| < \varepsilon$$

Das zeigt Teil a). Nun zu b). Wir müssen zeigen, dass $\sum_{i=0}^{\infty} |a_i|$ konvergiert. Für beliebige $r \in \mathbb{N}$ wählen wir $m \in \mathbb{N}$, s.d. $\{0,1,\dots,r\} \subset I_0 \cap I_1 \cap \cdots \cap I_m$. Dann gilt

$$\sum_{i=0}^{r} |a_i| \leq \sum_{i\in I_0\cup\cdots\cup I_m} |a_i| = \sum_{n=0}^{m}\sum_{i\in I_n} |a_i| \leq \sum_{n=0}^{\infty}\sum_{i\in I_n} |a_i|$$

Die rechte Seite ist eine obere Schranke, unabhängig von r. Das impliziert, dass $\sum_{i\in I} |a_i| < \infty$. □

Sei $a : \mathbb{N}\times\mathbb{N} \to \mathbb{R}$ eine Folge $(a_{n,m})_{(n,m)\in\mathbb{N}\times\mathbb{N}}$ mit Indexmenge $\mathbb{N}\times\mathbb{N}$. Um die Summierbarkeit von $a_{n,m}$ zu überprüfen, untersuchen wir, ob

$$\sum_{n=0}^{\infty}\left(\sum_{m=0}^{\infty} |a_{n,m}|\right) < \infty$$

Dann impliziert Prop. 4.21, Teil b), dass $a_{n,m}$ summierbar ist, und dass

$$\sum_{(n,m)\in\mathbb{N}\times\mathbb{N}} a_{n,m} = \sum_{n=0}^{\infty}\left(\sum_{m=0}^{\infty} a_{n,m}\right)$$

Im Fall $a_{n,m} = a_n b_m$ heisst $\sum_{n,m} a_n b_m$ das Cauchy-Produkt der Reihen $\sum_n a_n$ und $\sum_m b_m$.

Satz 4.22. *Seien $\sum_n a_n$ und $\sum_m b_m$ absolut konvergent.*

a) $a_n b_m$ ist summierbar über $\mathbb{N}\times\mathbb{N}$.

b) Es gilt

$$\sum_{(n,m)\in\mathbb{N}\times\mathbb{N}} a_n b_m = \left(\sum_{n=0}^{\infty} a_n\right)\left(\sum_{m=0}^{\infty} b_m\right) = \sum_{k=0}^{\infty}\sum_{j=0}^{k} a_j b_{k-j}$$

Beweis. Für $n \in \mathbb{N}$ sei $I_n = \{(n,0),(n,1),\dots\}$. Dann ist $a_n b_m$ summierbar über I_n, für alle $n \in \mathbb{N}$ (weil $\sum b_m$ absolut konvergent ist) und

$$\sum_{n=0}^{\infty}\left(\sum_{m=0}^{\infty}|a_n b_m|\right) = \sum_{n=0}^{\infty}|a_n|\sum_{m=0}^{\infty}|b_m| < \infty\,.$$

Proposition 4.21, Teil b), impliziert, dass $a_n b_m$ summierbar über $\mathbb{N}\times\mathbb{N}$ ist. Proposition 4.21, Teil a), impliziert auch, dass

$$\sum_{(n,m)\in\mathbb{N}\times\mathbb{N}} a_n b_m = \left(\sum_{n=0}^{\infty} a_n\right)\left(\sum_{m=0}^{\infty} b_m\right)$$

Um die zweite Gleichheit in b) zu zeigen, setzen wir $I_k = \{(n,m) : n+m = k\}$. Wir bemerken, dass

$$\mathbb{N}\times\mathbb{N} = \bigcup_{k=0}^{\infty} I_k$$

und dass $I_k \cap I_\ell = \emptyset$ für alle $k \neq \ell$. Proposition 4.21, Teil a), impliziert also,dass

$$\sum_{(n,m)\in\mathbb{N}\times\mathbb{N}} a_n b_m = \sum_{k=0}^{\infty}\sum_{(n,m)\in I_k} a_n b_m = \sum_{k=0}^{\infty}\sum_{j=0}^{k} a_j b_{k-j}$$

□

5 Metrische Räume

5.1 Konvergenz auf metrischen Räumen

Um den Begriff von Konvergenz zu definieren, braucht man nicht unbedingt, dass die Folge Werte auf $\mathbb{R}$ annimmt. Die Definition von Konvergenz kann in der Tat einfach verallgemeinert werden zu Räumen, in denen man Abstände zwischen Punkten messen kann. In der Mathematik heisst eine Funktion, die den Abstand zwischen Punkten misst, eine Metrik, und ein Raum, in welchem eine Metrik definiert ist, ein metrischer Raum. Auf $\mathbb{R}$ war bis jetzt der Abstand zwischen x und y durch $|x-y|$ gegeben; das ist ein Beispiel einer Metrik. Nun nehmen wir die wichtigsten Eigenschaften des Absolutbetrags und wir verlangen axiomatisch, dass jede Metrik solche Eigenschaften hat.

Definition 5.1. *Ein metrischer Raum ist eine Menge M, versehen mit einer Funktion $d : M \times M \to \mathbb{R}$, genannt Metrik, so dass*

- $d(x,y) \geq 0$ *für alle* $x,y \in M$ *und* $d(x,y) = 0$ *genau dann, wenn* $x = y$.
- $d(x,y) = d(y,x)$ *(d ist symmetrisch).*
- $d(x,y) \leq d(x,z) + d(z,y)$ *(d erfüllt die Dreiecksungleichung).*

Beispiele.

- $M = \mathbb{R}$, mit $d(x,y) = |x-y|$ ist ein metrischer Raum. Auch $M = \mathbb{Q}$, mit der selben Metrik, definiert einen metrischen Raum.

- $M = \mathbb{C}$, mit $d(z,w) = |z-w|_{\mathbb{C}}$ ist ein metrischer Raum. Sind $z = x_1 + iy_1$ und $w = x_2 + iy_2$, so gilt $d(z,w) = ((x_1-x_2)^2 + (y_1-y_2)^2)^{1/2}$.

- $M = \mathbb{R}^m$. Für $x = (x_1, \dots, x_m), y = (y_1, \dots, y_m) \in \mathbb{R}^m$, definieren wir

$$d(x,y) = \|x-y\| = \left(\sum_{j=1}^{m} (x_i - y_i)^2 \right)^{1/2} \tag{8}$$

  Dann ist (M,d) ein metrischer Raum (d heisst der euklidische Abstand zwischen x und y). Allgemeiner, jeder normierte Raum $(V, \|.\|)$ (erinnere, dass ein normierter Raum ein Vektorraum versehen mit einer Norm $\|.\|$ ist, der die Eigenschaften i)-ii)-iii) in Kapitel 2.6 hat) wird mit der Definition

$$d(x,y) = \|x-y\|$$

  ein metrischer Raum. Man nennt d die aus der Norm induzierte Metrik. Nicht alle Metriken sind durch eine Norm induziert; insbesondere können Metriken auf beliebige Mengen definiert werden, während Normen nur auf Vektorräumen sinnvoll sind.

- Für eine beliebige Menge M definiert

$$d(x,y) = \begin{cases} 0 & \text{falls } x = y \\ 1 & \text{sonst} \end{cases}$$

  eine Metrik auf M (man nennt diese Metrik die diskrete Metrik auf M).

- Sei (M,d) ein metrischer Raum und $N \subset M$. Dann ist N, versehen mit der Einschränkung von d auf $N \times N$, ein metrischer Raum.

Wir definieren nun den Begriff von Konvergenz von Folgen auf metrischen Räumen.

Definition 5.2. *Sei (M,d) ein metrischer Raum und $a : \mathbb{N} \to M$ eine Folge auf M. Wir sagen $a_n \to a$ (für ein $a \in M$), falls für alle $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$ existiert, mit $d(a_n, a) < \varepsilon$ für alle $n > n_0$.*

Bemerkung: Es gilt $a_n \to a$ auf M genau dann, wenn $d(a_n, a) \to 0$ auf $\mathbb{R}$.

Bemerkung: Die Konvergenz einer Folge hängt von der Wahl der Metrik ab. Z.B. für eine beliebige Menge M, versehen mit der diskreten Metrik $d(x,y) = 0$, falls $x = y$ und $d(x,y) = 1$, sonst gilt $a_n \to a$ g.d.w. $a_n = a$ für fast alle $n \in \mathbb{N}$. D.h. nur konstante Folgen sind bzgl. der diskreten Metrik konvergent. Z.B. falls $M = \mathbb{R}$, ist die Folge $a_n = 1/n$ bzgl. der diskrete Metrik, nicht konvergent.

Viele Eigenschaften von Konvergenz, die wir aus Konvergenz von Folgen auf $\mathbb{R}$ kennen, lassen sich mit Hilfe der Eigenschaften, die wir in der Definition der Metrik verlangt haben, auf beliebige metrische Räume erweitern. Sei M ein metrischer Raum.

- Jede Folge $a : \mathbb{N} \to M$ hat höchstens einen Grenzwert. In der Tat, falls $a_n \to a$ und $a_n \to b$ mit $a \neq b$, so existiert $n_1 \in \mathbb{N}$ und $n_2 \in \mathbb{N}$ mit $d(a_n, a) < \varepsilon/2$ für alle $n > n_1$ und $d(a_n, b) < \varepsilon/2$ für alle $n > n_2$. Also, für $n > \max\{n_1, n_2\}$, gilt

$$d(a,b) \leq d(a,a_n) + d(a_n, b) < \varepsilon = d(a,b)$$

was einen Widerspruch ergibt.

- Sei $a : \mathbb{N} \to M$ eine Folge auf M. $a \in M$ heisst ein Häufungspunkt von a_n, falls für jede $\varepsilon > 0$, unendlich viele $n \in \mathbb{N}$ existieren, mit $d(a_n, a) < \varepsilon$. Es gilt: a ist ein Häufungspunkt von a_n g.d.w. eine Teilfolge n_j mit $a_{n_j} \to a$ existiert. Der Beweis ist ähnlich wie in Proposition 3.20.

- Eine Folge $a : \mathbb{N} \to M$ auf M heisst eine Cauchy-Folge, falls für alle $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$ existiert, mit

$$n, m > n_0 \quad \Rightarrow \quad d(a_n, a_m) < \varepsilon\,.$$

Jede konvergente Folge auf M ist eine Cauchy-Folge.

Definition 5.3. *Ein metrischer Raum M heisst vollständig, falls jede Cauchy Folge konvergiert.*

Der Raum $\mathbb{R}$ ist vollständig. Es ist einfach, Teilmengen von $\mathbb{R}$ zu finden, die nicht vollständig sind. Z.B. auf $M = (0;1]$, versehen mit der Metrik $d(x,y) = |x-y|$, ist die Folge $a_n = 1/n$ eine Cauchy-Folge die nicht konvergiert. $\mathbb{Q}$, versehen wieder mit der Metrik $d(x,y) = |x-y|$, ist nicht vollständig. In der Tat, es ist einfach, eine Folge in $\mathbb{Q}$ zu konstruieren, die in $\mathbb{R}$ gegen $\sqrt{2}$ konvergiert. Da sie in $\mathbb{R}$ konvergiert, ist eine solche Folge offenbar eine Cauchy-Folge, die aber auf $\mathbb{Q}$ nicht konvergiert.

Ein wichtiges Beispiel vom metrischen Raum ist der Raum $\mathbb{R}^m$, für $m \in \mathbb{N}$.

Satz 5.4 (Konvergenz auf $\mathbb{R}^m$)**.** *Sei $M = \mathbb{R}^m$, versehen mit der Euklidischen Metrik*

$$d(x,y) = \|x-y\| = \left[\sum_{j=1}^m (x_j - y_j)^2\right]^{1/2}$$

für $x = (x_1, \dots, x_m), y = (y_1, \dots, y_m)$. Sei $a_n = (a_n^1, \dots, a_n^m)$ eine Folge auf $\mathbb{R}^m$, und $a = (a^1, \dots, a^m) \in \mathbb{R}^m$. Dann gilt $a_n \to a$ in $\mathbb{R}^m$ genau dann, wenn $a_n^j \to a^j$ für alle $j = 1, \dots, m$.

Beweis. Die Konvergenz $a_n \to a$ impliziert, dass

$$d(a_n, a) = \left[\sum_{j=1}^m (a_n^j - a^j)^2\right] \to 0$$

Für beliebiges $j = 1, \dots, m$ gilt

$$0 \leq |a_n^j - a^j| \leq d(a_n, a)$$

Das impliziert, dass $|a_n^j - a^j| \to 0$. Anderseits, nehmen wir an $|a_n^j - a^j| \to 0$ für $j = 1, \dots, m$. Wir möchten zeigen, dass $a_n \to a$. Sei dazu $\varepsilon > 0$ fest. Dann existiert $n_1 \in \mathbb{N}$ mit $|a_n^1 - a^1| < \varepsilon/\sqrt{m}$ für alle $n > n_1$. Analog gibt es für jede $j = 2, 3, \dots, m$ $n_j \in \mathbb{N}$ mit $|a_n^j - a^j| < \varepsilon/\sqrt{m}$ für alle $n > n_j$. Für $n > n_0 = \max(n_1, \dots, n_m)$ gilt also

$$d(a_n, a) = \left(\sum_{j=1}^{m} (a_n^j - a^j)^2 \right)^{1/2} < \left(\sum_{j=1}^{n} \varepsilon^2/m \right)^{1/2} = \varepsilon$$

□

Der metrische Raum $\mathbb{R}^2$ ist per Definition der Norm $\|.\|$ identisch mit dem metrischen Raum $\mathbb{C}$, versehen mit der Metrik $d(z, w) = |z - w|_{\mathbb{C}}$. Satz 5.4 bedeutet, dass eine Folge z_n auf $\mathbb{C}$ gegen $z \in \mathbb{C}$ genau dann konvergiert, wenn $\operatorname{Re} z_n \to \operatorname{Re} z$, und $\operatorname{Im} z_n \to \operatorname{Im} z$. Somit ist die Frage, ob eine Folge auf $\mathbb{R}^m$ oder auf $\mathbb{C}$ konvergiert, auf die Untersuchung der Konvergenz von Folgen auf $\mathbb{R}$ reduziert. Damit erhalten wir die Rechenregel für Konvergenz von Folgen auf $\mathbb{R}^m$ und $\mathbb{C}$. Seien a_n, b_n Folgen auf $\mathbb{R}^m$ mit $a_n \to a$ und $b_n \to b$, und α_n eine Folge auf $\mathbb{R}$ mit $\alpha_n \to \alpha$. Dann gilt

$$a_n + b_n \to a + b, \quad a_n - b_n \to a - b \quad \text{und} \quad \alpha_n a_n \to \alpha a\,.$$

Sind a_n und b_n Folgen auf $\mathbb{C}$, so gilt

$$a_n + b_n \to a + b, \quad a_n - b_n \to a - b \quad \text{und} \quad a_n b_n \to ab$$

Ist $b_n \neq 0$ für alle $n \in \mathbb{N}$ und $b \neq 0$, so gilt auch $a_n/b_n \to a/b$.

Eine Folge a_n auf $\mathbb{R}^m$ heisst beschränkt, falls eine Konstante $C > 0$ mit $\|a_n\| \leq C$ für alle $n \in \mathbb{N}$ existiert (eine Folge a_n auf $\mathbb{C}$ ist beschränkt falls $|a_n|_{\mathbb{C}} \leq C$ für alle $n \in \mathbb{N}$). Ähnlich wie auf $\mathbb{R}$ (siehe Korollar 3.21), ist jede konvergente Folge auf $\mathbb{R}^m$ beschränkt. Anderseits, wie schon auf $\mathbb{R}$, hat jede beschränkte Folge auf $\mathbb{R}^m$ (oder auf $\mathbb{C}$) mindestens eine konvergente Teilfolge.

Satz 5.5 (Satz von Bolzano-Weierstrass auf $\mathbb{R}^m$)**.** *Jede beschränkte Folge auf $\mathbb{R}^m$ hat eine konvergente Teilfolge.*

Beweis. Sei $a_n = (a_n^1, \dots, a_n^m)$ eine beschränkte Folge auf $\mathbb{R}^m$. Dann ist a_n^j, für jedes $j = 1, \dots, m$ eine beschränkte Folge auf $\mathbb{R}$. Da a_n^1 eine beschränkte Folge auf $\mathbb{R}$ ist, existiert aus Korollar 3.21 eine Teilfolge $n_{1,j}$ und ein a^1, mit $a_{n_{1,j}}^1 \to a^1$. Nun definiert $a_{n_{1,j}}^2$ eine beschränkte Folge auf $\mathbb{R}$; deswegen existiert $n_{2,j}$, Teilfolge von $n_{1,j}$, und $a^2 \in \mathbb{R}$ mit $a_{n_{2,j}}^2 \to a^2$. Da $n_{2,j}$ eine Teilfolge von $n_{1,j}$ ist, gilt wieder $a_{n_{2,j}}^1 \to a^1$. Rekursiv finden wir eine Teilfolge $n_{m,j}$ und $a^1, \dots, a^m \in \mathbb{R}$ mit $a_{n_{m,j}}^\ell \to a^\ell$ für alle $\ell = 1, \dots, m$. Das impliziert, dass $a_{n_{m,j}} \to (a^1, \dots, a^m)$. □

Ähnlich wie $\mathbb{R}$, ist der metrische Raum $\mathbb{R}^m$ vollständig für jedes $m \in \mathbb{N}$ (damit ist auch $\mathbb{C}$ vollständig).

Satz 5.6 (Cauchy-Kriterium auf $\mathbb{R}^m$)**.** *Sei a_n eine Folge auf $\mathbb{R}^m$. Dann ist a_n genau dann konvergent, wenn a_n eine Cauchy-Folge ist.*

Beweis. Ähnlich wie im Fall $m = 1$. Zunächst zeigt man, dass Cauchy-Folgen immer beschränkt sind. Satz 5.5 impliziert dann, dass jede Cauchy-Folge eine konvergente Teilfolge hat. Dann zeigt man, dass eine Cauchy-Folge mit einer konvergenten Teilfolge konvergent ist (wie im Schritt 2 vom Beweis von Satz 3.23). $\square$

Da auf $\mathbb{R}^m$ (und auf $\mathbb{C}$) eine Addition definiert ist (das ist auf allgemeinen metrischen Räumen nicht der Fall), kann man auf $\mathbb{R}^m$ und auf $\mathbb{C}$, wie auf $\mathbb{R}$, Reihen betrachten. Sei $a : \mathbb{N} \to \mathbb{R}^m$ eine Folge auf $\mathbb{R}^m$. Ein Ausdruck $\sum_{j=0}^{\infty} a_j$ heisst eine Reihe, mit Gliedern a_n. Für $n \in \mathbb{N}$ definieren wir die n-te Partialsumme

$$s_n = a_0 + \cdots + a_n = \sum_{j=0}^{n} a_j$$

Dann ist $s : \mathbb{N} \to \mathbb{R}^m$ eine Folge auf $\mathbb{R}^m$. Ist die Folge konvergent, so definieren wir

$$\sum_{j=0}^{\infty} a_j = \lim_{n\to\infty} s_n = \lim_{n\to\infty} \sum_{j=0}^{n} a_j$$

Eine Reihe $\sum_{j=0}^{\infty} a_j$ auf $\mathbb{R}^m$ heisst absolut konvergent, falls $\sum_{j=0}^{\infty} \|a_j\|$ (Reihe auf $\mathbb{R}$) konvergiert. Wie auf $\mathbb{R}$, gilt (hier ist die Vollständigkeit von $\mathbb{R}^m$ wichtig)

$$\sum_{n=0}^{\infty} a_n \text{ absolut konvergent } \Rightarrow \sum_{n=0}^{\infty} a_n \text{ konvergent}$$

Das bedeutet, dass die verschiedenen Kriterien, die wir für die absolute Konvergenz von Reihen auf $\mathbb{R}$ diskutiert haben, auch auf $\mathbb{R}^m$ und auf $\mathbb{C}$ gelten.

5.2 Offene und abgeschlossene Mengen

Sei M ein metrischer Raum mit Metrik d, $a \in M$ und $r > 0$. Dann ist

$$B_r(a) := \{x \in M : d(x, a) < r\}$$

die offene Kugel von Radius r um a. Für $M = \mathbb{R}$, ist $B_r(a)$ das offene Interval $B_r(a) = (a - r; a + r)$. Für $M = \mathbb{C}$, $B_r(a) =$ Kreisscheibe von Radius r um a (ohne Rand). Ist $M = \mathbb{R}^m$ und $a = (a_1, \dots, a_m) \in M$, so ist

$$B_r(a) = \left\{ x = (x_1, \dots, x_m) \in \mathbb{R}^m : \sum_{j=1}^{m} (x_j - a_j)^2 < r^2 \right\}$$

Definition 5.7. *Eine Menge $G \subset M$ heisst offen, falls für alle $x \in G$ existiert $\varepsilon > 0$, s.d. $B_\varepsilon(x) \subset G$. D.h. G ist offen, falls für jeden $x \in G$, alle Punkte in G sind, die genügend nahe bei x liegen. Eine Menge $F \subset M$ heisst abgeschlossen, falls ihr Komplement $F^c = M \backslash F$ offen ist.*

Beispiel: $M = \mathbb{R}$ mit der üblichen Metrik, die aus dem Absolutbetrag definiert wird, $a < b$. Dann ist

$$(a; b) = \{x \in \mathbb{R} : a < x < b\}$$

eine offene Menge. Dagegen ist

$$[a;b] = \{x \in \mathbb{R} : a \leq x \leq b\}$$

nicht offen (die Bedingung ist bei $x = a$ und $x = b$ nicht erfüllt). Da

$$[a;b]^c = (-\infty; a) \cup (b; \infty)$$

offen ist, ist $[a;b]$ eine abgeschlossene Menge. Insbesondere, die Menge $\{x\}$, die aus einem einzigen Punkt $x \in \mathbb{R}$ besteht, ist abgeschlossen.

Bemerkung: $M, \emptyset$ sind immer offen und abgeschlossen.

Bemerkung: Welche Mengen offen und abgeschlossen sind, hängt von der Wahl der Metrik ab. Z.B. für eine belibige Menge M, versehen mit der diskreten Metrik $d(x,y) = 0$, falls $x = y$ und $d(x,y) = 1$, falls $x \neq y$, ist $B_\varepsilon(x) = \{x\}$ für alle $\varepsilon < 1$. Das impliziert, dass alle Teilmengen von M offen sind. Deswegen sind auch alle Teilmengen von M abgeschlossen.

Proposition 5.8. *Die offene Kugel $B_r(a)$ ist für alle $a \in M$ und $r > 0$ offen. Die abgeschlossene Kugel $\{x \in M : d(x,a) \leq r\}$ ist dagegen abgeschlossen.*

Beweis. Sei $x \in B_r(a)$. Dann $d(x,a) < r$. Wir wählen $\varepsilon = r - d(x,a)$ und behaupten, dass $B_\varepsilon(x) \subset B_r(a)$. In der Tat,

$$y \in B_\varepsilon(x) \;\Rightarrow\; d(y,x) < \varepsilon \;\Rightarrow\; d(a,y) \leq d(a,x) + d(x,y) < d(a,x) + \varepsilon = r \;\Rightarrow\; y \in B_r(a)$$

Um zu zeigen, dass $\{x \in M : d(x,a) \leq r\}$ abgeschlossen ist, bemerken wir, dass

$$M \backslash \{x \in M : d(x,a) \leq r\} = \{x \in M : d(x,a) > r\}$$

Sei nun $x \in M$ s.d. $d(x,a) > r$ und $\varepsilon = d(x,a) - r$. Dann behaupten wir, dass $B_\varepsilon(x) \subset \{x \in M : d(x,a) > r\}$. In der Tat, ist $y \in B_\varepsilon(x)$, so gilt aus der Dreiecksungleichung $d(y,a) \geq d(a,x) - d(x,y) > d(a,x) - \varepsilon = r$. □

Satz 5.9. *Sei M ein metrischer Raum.*

i) $G_1, G_2 \subset M$ offen. Dann ist auch $G_1 \cap G_2$ offen.

ii) G_i offen, für alle i in eine beliebige Indexmenge I (braucht nicht abzählbar zu sein). Dann ist $\bigcup_{i \in I} G_i$ offen.

iii) $F_1, F_2 \subset M$ abgeschlossen. Dann ist auch $F_1 \cup F_2$ abgeschlossen.

iv) F_i abgeschlossen, für alle i in eine beliebige Indexmenge I (braucht nicht abzählbar zu sein). Dann ist $\bigcap_{i \in I} F_i$ abgeschlossen.

Bemerkung: durch wiederholte Anwendung von Teil i), es folgt, dass

$$G_1, \ldots, G_n \text{ offen } \Rightarrow \bigcap_{j=1}^{n} G_j \text{ ist offen}$$

Also besagen i) und ii), dass endliche Durchschnitte und beliebige Vereinigungen von offenen Mengen offen sind. Analog implizieren iii) und iv), dass endliche Vereinigungen und beliebige Durchschnitte von abgeschlossenen Mengen abgeschlossen sind. Unendliche Durchschnitte offener Mengen brauchen nicht offen zu sein. Z.B. $G_n = (-1/n; 1/n) \subset \mathbb{R}$ (versehen mit der üblichen Metrik) ist offen für alle $n \in \mathbb{N}$, aber

$$\bigcap_{n=1}^{\infty} G_n = \{0\}$$

ist abgeschlossen. Analog brauchen unendliche Vereinigungen abgeschlossener Mengen nicht abgeschlossen zu sein. Z.B. $\bigcup_{x \in (0,1)} \{x\} = (0,1)$ ist offen.

Beweis. i) Sei $x \in G_1 \cap G_2$. Dann

$$\begin{aligned} x \in G_1 &\Rightarrow \exists\, r_1 > 0 : \; B_{r_1}(x) \subset G_1 \\ x \in G_2 &\Rightarrow \exists\, r_2 > 0 : \; B_{r_2}(x) \subset G_2 \end{aligned}$$

Sei $r = \min(r_1, r_2)$. Dann gilt $B_r(x) \subset B_{r_1}(x) \subset G_1$ und $B_r(x) \subset B_{r_2}(x) \subset G_2$. Also, $B_r(x) \subset G_1 \cap G_2$. ii) Sei

$$x \in \bigcup_{i \in I} G_i$$

Dann existiert $i \in I$ mit $x \in G_i$. Also existiert $r > 0$ mit $B_r(x) \subset G_i$. Das impliziert, dass $B_r(x) \subset \bigcup_{i \in I} G_i$. iii) Wir benutzen die Morgan'sche Regel

$$(F_1 \cup F_2)^c = F_1^c \cap F_2^c \,.$$

F_1, F_2 abgeschlossen impliziert, dass F_1^c, F_2^c offen sind. Teil i) impliziert, dass $F_1^c \cap F_2^c$ offen ist. Deswegen ist $F_1 \cup F_2$ abgeschlossen. iv) Aus der Morgan'schen Regel:

$$\left(\bigcap_{i \in I} F_i \right)^c = \bigcup_{i \in I} F_i^c$$

F_i abgeschlossen impliziert, dass F_i^c offen ist, für alle $i \in I$. Teil ii) impliziert also, dass $\bigcup_{i \in I} F_i^c$ offen ist, und deswegen, dass $\bigcap_{i \in I} F_i$ abgeschlossen ist. □

Definition 5.10. *Sei M ein metrischer Raum, $A \subset M$, $a \in A$. A heisst eine Umgebung von a, falls eine offene Menge G mit $a \in G$ und $G \subset A$ existiert (wir können immer $G = B_\varepsilon(a)$ wählen, für genügend kleines $\varepsilon > 0$).*

Bemerkungen: Die offene Kugel $B_r(a)$ ist eine Umgebung von a, für alle $r > 0$. Eine Menge $B \subset M$ ist genau dann offen, wenn B eine Umgebung jedes Punktes in B ist.

Proposition 5.11. *Sei M ein metrischer Raum, $(x_n)_{n \in \mathbb{N}}$ eine Folge in M, $a \in M$. Dann $x_n \to a$ g.d.w. für jede Umgebung U von a, $x_n \in U$ für fast alle $n \in \mathbb{N}$.*

Bemerkung: Diese Charakterisierung von Konvergenz benutzt die Metrik nicht direkt (die Metrik wird nur in der Definition von offenen Mengen und von Umgebungen benutzt). Das erlaubt den Begriff von Konvergenz auf sogenannte topologische Räume zu erweitern, wo keine Metrik vorhanden ist, sondern nur eine Topologie (die Topologie definiert, welche Mengen offen sind).

Beweis. Sei $x_n \to a$ und U eine Umgebung von a. Dann existiert $\varepsilon > 0$ mit $B_\varepsilon(a) \subset U$. Weiter existiert $n_0 \in \mathbb{N}$ mit $d(x_n, a) < \varepsilon$ für alle $n > n_0$. Das bedeutet, dass $x_n \in B_\varepsilon(a) \subset U$ für alle $n > n_0$. Also $x_n \in U$ für fast alle $n \in \mathbb{N}$. Anderseits, nehmen wir an, dass für jede Umgebung U von a, $x_n \in U$ für fast alle $n \in \mathbb{N}$. Wir möchten zeigen, dass $x_n \to a$. Sei dazu $\varepsilon > 0$. Dann ist $B_\varepsilon(a)$ eine Umgebung von a, und deswegen $x_n \in B_\varepsilon(a)$ für fast alle $n \in \mathbb{N}$. Das bedeutet, es existiert $n_0 \in \mathbb{N}$ mit $d(x_n, a) < \varepsilon$ für alle $n > n_0$. $\square$

Anderseits kann man auf metrischen Räumen abgeschlossene (und deswegen auch offene) Mengen durch Untersuchung von Konvergenz von Folgen identifizieren.

Proposition 5.12. *Sei M ein metrischer Raum, $A \subset M$. Folgende Aussagen sind äquivalent.*

a) A ist abgeschlossen.

b) $(x_n)_{n\in\mathbb{N}}$ eine Folge in A, die gegen $a \in M$ konvergiert, dann ist $a \in A$.

Beweis. a) $\Rightarrow$ b): Sei A abgeschlossen, und x_n eine Folge in A mit $x_n \to a$ in M. Wir behaupten, dass $a \in A$. Wäre nämlich $a \in A^c$, dann impliziert A^c offen, dass $\varepsilon > 0$ existiert, mit $B_\varepsilon(a) \subset A^c$. Für $x_n \to a$ existiert $n_0 \in \mathbb{N}$ mit $x_n \in B_\varepsilon(a) \subset A^c$ für alle $n > n_0$, im Gegensatz zur Annahme $x_n \in A$ für alle $n \in \mathbb{N}$.

b) $\Rightarrow$ a): Wir zeigen, dass $\neg$ a) $\Rightarrow$ $\neg$ b). Sei A nicht abgeschlossen, dann ist A^c nicht offen. Also, es existiert $a \in A^c$ mit $B_\varepsilon(a) \cap A \neq \emptyset$ für jede $\varepsilon > 0$. Für jede $n \in \mathbb{N}$ finden wir also $x_n \in B_{1/n}(a) \cap A$. Dann gilt $x_n \in A$ für alle $n \in \mathbb{N}$, $x_n \to a$, aber $a \in A^c$. Das zeigt $\neg$ b). $\square$

Definition 5.13. *Sei M ein metrischer Raum, $A \subset M$. Wir definieren dann den Abschluss $\bar{A}$ von A durch*

$$\bar{A} = \bigcap \{F : F \text{ abgeschlossen und } F \supset A\}$$

Ähnlicherweise ist das Innere A° von A durch

$$A^\circ = \bigcup \{G : G \text{ offen und } G \subset A\}$$

definiert. Aus Satz 5.9 folgt, dass $\bar{A}$ abgeschlossen ist, und dass A° offen ist. $\bar{A}$ ist die kleinste abgeschlossene Menge, die A enthält. A° ist dagegen die grösste offene Menge, die in A enthalten ist.

Bemerkung: Ist A abgeschlossen, so gilt $\bar{A} = A$. Ist A offen, dann ist $A^\circ = A$. Es gilt immer $A^\circ \subset A \subset \bar{A}$.

Bemerkung: Für jede $A \subset M$ gilt $(\bar{A})^c = (A^c)^\circ$. Beweis: Übung.

Proposition 5.14. *Sei $M = \mathbb{R}$ mit der üblichen Metrik. Dann gilt $\mathbb{Q}^\circ = \emptyset$ und $\bar{\mathbb{Q}} = \mathbb{R}$.*

Beweis. Sei $G \subset \mathbb{R}$ eine nicht leere offene Menge, $a \in G$. Dann $(a - \varepsilon; a + \varepsilon) \subset G$ für $\varepsilon > 0$ klein genug. Aber $(a - \varepsilon; a + \varepsilon)$ enthält eine irrationale Zahl, für beliebige $\varepsilon > 0$. Also $G \not\subset \mathbb{Q}$. Deswegen ist die einzige offene Menge $G \subset \mathbb{Q}$ die leere Menge, und $\mathbb{Q}^\circ = \emptyset$. Aus der Bemerkung oben folgt, dass $\bar{\mathbb{Q}} = \mathbb{R}$. $\square$

Definition 5.15. *Sei M ein metrischer Raum, $A \subset M$. Der Rand ∂A von A ist die Menge aller $x \in M$ s.d. jede Umgebung von x A und A^c schneidet, d.h.*

$$\partial A = \{x \in M : B_\varepsilon(x) \cap A \neq \emptyset \text{ und } B_\varepsilon(x) \cap A^c \neq \emptyset \text{ für alle } \varepsilon > 0\}$$

Beispiel: Ist $M = \mathbb{R}^2$ und $A = B_1(0)$, dann ist $\partial A = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$ der Einheitskreis. Ist $M = \mathbb{R}$ und $A = \mathbb{Q}$, so gilt $\partial\mathbb{Q} = \mathbb{R}$.

Bemerkung: Es gilt $\partial A = \bar{A} \backslash A^\circ$. Beweis: Übung.

5.3 Kompaktheit

Als wir Folgen auf $\mathbb{R}$ betrachtet haben, haben wir gelernt, dass beschränkte Folgen immer konvergente Teilfolgen besitzen (Bolzano-Weierstrass). Diese Eigenschaft ist sehr nützlich in der Analysis. Es stellt sich die Frage, wann man auf allgemeinen metrische Räumen auf die Existenz von konvergenten Teilfolgen schliessen kann. Da Cauchy-Folgen mit konvergenten Teilfolgen konvergent sind, ist diese Frage auch mit der Vollständigkeit vom metrischen Raum verbunden.

Definition 5.16. *Sei M ein metrischer Raum. Eine Teilmenge $A \subset M$ heisst folgenkompakt, falls jede Folge in A eine in A konvergente Teilfolge besitzt (eine konvergente Teilfolge, deren Limes in A enthalten ist).*

Beispiel: Der abgeschlossene Intervall $[0; 1]$ ist folgenkompakt. In der Tat, jede Folge auf $[0; 1]$ besitzt aus Bolzano-Weierstrass eine in $\mathbb{R}$ konvergente Teilfolge. Da $[0; 1]$ abgeschlossen ist, muss die Teilfolge in $[0; 1]$ konvergieren. Ähnlich werden wir sehen, dass jede abgeschlossene und beschränkte Teilmenge von $\mathbb{R}^m$ folgenkompakt ist.

Auf metrischen Räumen ist der Begriff von Folgenkompaktheit mit dem Begriff von Kompaktheit (oder Überdeckungskompaktheit) äquivalent.

Definition 5.17. *Sei M ein metrischer Raum und $A \subset M$. Eine offene Überdeckung von A ist eine Familie $(G_i)_{i \in I}$ offener Mengen G_i, mit*

$$A \subset \bigcup_{i \in I} G_i \,.$$

Eine Teilmenge $A \subset M$ heisst kompakt, falls jede offene Überdeckung von A eine endliche Teilüberdeckung enthält. D.h:

$$A \subset \bigcup_{i \in I} G_i, \quad G_i \text{ offen für alle } i \in I \;\Rightarrow\; \exists\, i_1, \dots, i_n \in I : \; A \subset \bigcup_{j=1}^{n} G_{i_j}$$

Beispiele:

- Sei $M = \mathbb{R}$ und $A = \{1, 1/2, 1/3, \dots\}$. Für $n \in \mathbb{N}$ sei $G_n = (1/(n+1); 2)$. Dann ist G_n offen für alle $n \in \mathbb{N}$, und

$$A \subset \bigcup_{n \in \mathbb{N}} G_n$$

Also ist $(G_n)_{n\in\mathbb{N}}$ eine offene Überdeckung von A. Wir behaupten, es existiert keine endliche Teilüberdeckung. In der Tat, nehmen wir an, dass $A \subset G_{n_1} \cup \cdots \cup G_{n_m}$. Dann, mit $n_0 = \max\{n_1, \ldots, n_m\}$ wir haben $G_{n_j} \subset G_{n_0}$ für alle $j = 1, \ldots, m$. Deswegen ist $G_{n_1} \cup \cdots \cup G_{n_m} = G_{n_0}$. Aber $A \subset G_{n_0}$ kann offenbar nicht gelten. Also ist A nicht kompakt.

- Sei wieder $M = \mathbb{R}$, und $A = \{0, 1, 1/2, 1/3, \ldots\}$. Wir behaupten A ist kompakt. Sei in der Tat $(G_i)_{i\in I}$ eine offene Überdeckung von A. Da $0 \in A$ existiert $i_0 \in I$ mit $0 \in G_{i_0}$. Da G_{i_0} offen ist, existiert $n_0 \in \mathbb{N}$ mit $1/n \in G_{i_0}$ für alle $n > n_0$. Für jede $1 \leq j \leq n_0$ finden wir $i_j \in I$ mit $1/j \in G_{i_j}$. Dann ist aber

$$A \subset \bigcup_{j=0}^{n_0} G_{i_j}$$

  Wir haben also eine endliche Teilüberdeckung gefunden. Das zeigt, dass A kompakt ist.

Theorem 5.18. *Sei M ein metrischer Raum. Dann ist $A \subset M$ genau dann kompakt, wenn A folgenkompakt ist.*

Beweis. "$\Leftarrow$": Sei A folgenkompakt und $(G_i)_{i\in I}$ eine offene Überdeckung von A. Für $B \subset A$ werden wir sagen, dass B überdeckbar ist, falls $n \in \mathbb{N}$ und $i_1, \ldots, i_n \in I$ existieren, mit $B \subset G_{i_1} \cup \cdots \cup G_{i_n}$. Bemerke, dass $B \subset G_i$ impliziert, dass B überdeckbar ist. Weiter, falls $B_1, \ldots, B_m$ überdeckbar sind, so ist auch $B_1 \cup \cdots \cup B_m$ überdeckbar. Wir müssen zeigen, dass A überdeckbar ist. Wir werden dazu das folgende Hilflemma benutzen.

Lemma 5.19. *Sei A folgenkompakt. Dann existieren $m \in \mathbb{N}$ und $y_1, \ldots, y_m \in A$ mit*

$$A \subset B_\varepsilon(y_1) \cup B_\varepsilon(y_2) \cup \cdots \cup B_\varepsilon(y_m)$$

für ein beliebiges $\varepsilon > 0$.

Beweis des Lemmas. Nehmen wir an, die Behauptung sei falsch für ein $\varepsilon > 0$. Dann wählen wir $y_1 \in A$ beliebig, und rekursiv

$$y_n \in A \setminus (B_\varepsilon(y_1) \cup \cdots \cup B_\varepsilon(y_{n-1}))$$

Das ist möglich, weil die Menge nie leer wird (sonst wäre $A \subset B_\varepsilon(y_1) \cup \cdots \cup B_\varepsilon(y_m)$). Die rekursiv definierte Folge y_n ist, s.d. $d(y_n, y_m) > \varepsilon$ für alle $n \neq m$. Der Abstand zwischen zwei Elementen irgendeiner Teilfolge ist immer grösser als ε. Es existiert also keine Teilfolge, die das Cauchy-Kriterium erfüllt. Deswegen hat y_n keine konvergente Teilfolge, in Widerspruch zur Annahme. □

Zurück zum Beweis von der Implikation "$\Leftarrow$" von Theorem 5.18. Wir nehmen an, A ist nicht überdeckbar. Aus dem Hilfslemma, mit $\varepsilon = 1$, finden wir $y_1^{(1)}, \ldots, y_{m_1}^{(1)} \in A$ mit

$$A \subset B_1(y_1^{(1)}) \cup \cdots \cup B_1(y_{m_1}^{(1)})$$

A nicht überdeckbar impliziert, dass für mindestens ein $j \in \{1, \dots, m_1\}$, $A \cap B_1(y_j^{(1)})$ nicht überdeckbar ist. Wir setzen $x_1 = y_j^{(1)}$. Aus dem Hilfslemma mit $\varepsilon = 1/2$ folgt, dass

$$A \subset B_{1/2}(y_1^{(2)}) \cup \dots B_{1/2}(y_{m_2}^{(2)})$$

für geeignete $y_1^{(2)}, \dots, y_{m_2}^{(2)} \in A$. Da $A \cap B_1(x_1)$ nicht überdeckbar ist, existiert $j \in \{1, \dots, m_2\}$ mit $A \cap B_1(x_1) \cap B_{1/2}(y_j^{(2)})$ nicht überdeckbar. Wir setzen $x_2 = y_j^{(2)}$. Rekursiv, finden wir eine Folge $(x_n)_{n\in\mathbb{N}}$, s.d. $x_n \in A$ und

$$A \cap B_1(x_1) \cap \dots \cap B_{1/n}(x_n)$$

nicht überdeckbar ist, für jede $n \in \mathbb{N}$. Da A folgenkompakt ist, hat die Folge x_n einen Häufungspunkt $\bar{x}$. Es gilt

$$\bar{x} \in A \subset \bigcup_{i\in I} G_i$$

Also existiert $i_0 \in I$ mit $\bar{x} \in G_{i_0}$. Da G_{i_0} offen ist, existiert $\varepsilon > 0$ mit $B_{2\varepsilon}(\bar{x}) \subset G_{i_0}$. Da $\bar{x}$ ein Häufungspunkt ist, finden wir $n \in \mathbb{N}$ mit $1/n < \varepsilon$ und $x_n \in B_\varepsilon(\bar{x})$ (weil unendlich viele x_n in $B_\varepsilon(\bar{x})$) liegen. Dann gilt

$$B_{1/n}(x_n) \subset B_{\varepsilon+1/n}(\bar{x}) \subset B_{2\varepsilon}(\bar{x}) \subset G_{i_0}$$

Deswegen ist

$$A \cap B_1(x_0) \cap B_{1/2}(x_1) \cap \dots \cap B_{1/n}(x_n) \subset G_{i_0}$$

sicher überdeckbar, in Widerspruch zur Annahme.

"$\Rightarrow$": Sei A kompakt und (x_n) eine Folge auf A. Wir behaupten, es existiert $x^* \in A$ Häufungspunkt von x_n (d.h. für alle $\varepsilon > 0$ existieren unendlich viele $n \in \mathbb{N}$ mit $x_n \in B_\varepsilon(x^*)$). Falls nicht, so gäbe es für jede $x \in A$ ein $\varepsilon_x > 0$ mit $x_n \in B_{\varepsilon_x}(x)$ für nur endlich viele $n \in \mathbb{N}$. Da $B_{\varepsilon_x}(x)$ offen ist, gibt

$$A \subset \bigcup_{x\in A} B_{\varepsilon_x}(x)$$

eine offene Überdeckung von A. Da A kompakt ist, existiert eine endliche Teilüberdeckung. D.h. es existieren $y_1, \dots, y_m \in A$ mit

$$A \subset \bigcup_{j=1}^m B_{\varepsilon_{y_j}}(y_j)$$

Da jede $B_{\varepsilon_{y_j}}(y_j)$ nur endlich viele Elemente der Folge x_n enthält, gibt es in A nur endlich viele Elemente von dieser Folge, was ein Widerspruch ist. Die Existenz eines Häufungspunktes in A impliziert die Existenz einer in A konvergenten Teilfolge. $\square$

Die Kompaktheit von $A \subset M$ impliziert, dass A abgeschlossen ist, und dass jede abgeschlossene Teilmenge von A kompakt ist.

Satz 5.20. *Sei M ein metrischer Raum und $A \subset M$ kompakt. Dann ist A abgeschlossen. Ist weiter $B \subset A$ abgeschlossen, dann ist B kompakt.*

Beweis. Sei x_n eine Folge in A, mit $x_n \to x$ in M. Aus Kompaktheit folgt, dass x_n eine in A konvergente Teilfolge hat. Also existiert Teilfolge x_{n_j} und $y \in A$, mit $x_{n_j} \to y$. Es folgt aus Proposition 3.19, dass $x = y \in A$. Also ist A abgeschlossen. Sei nun $B \subset A$ abgeschlossen und x_n eine Folge in B. Da A kompakt ist, hat x_n eine in A konvergente Teilfolge $x_{n_j} \to x \in A$. Da B abgeschlossen ist, muss aber $x \in B$. Deswegen ist B kompakt. □

Kompakte metrische Räume sind immer vollständig.

Satz 5.21. *Sei M ein metrischer Raum. Ist M kompakt, dann ist M vollständig.*

Beweis. Sei x_n eine Cauchy-Folge auf M. Da M kompakt ist (und deswegen auch folgenkompakt) existiert eine Teilfolge n_j, s.d. x_{n_j} in M konvergiert. Sei $x = \lim_{j\to\infty} x_{n_j}$. Wir behaupten, dass $x_n \to x$. In der Tat

$$d(x_n, x) \leq d(x_n, x_{n_j}) + d(x_{n_j}, x)$$

wird kleiner als ein beliebiges $\varepsilon > 0$ falls j und n genügend gross sind (weil x_n eine Cauchy-Folge ist und weil $x_{n_j} \to x$). □

Für Teilmengen von $\mathbb{R}^m$ (und deswegen auch von $\mathbb{C}$) gibt es eine sehr einfache Charakterisierung von Kompaktheit und Folgenkompaktheit. Erinnere, dass eine Teilmenge $A \subset \mathbb{R}^m$ beschränkt ist, falls ein $C > 0$ existiert, mit $\|x\| \leq C$, für alle $x \in A$.

Satz 5.22 (Satz von Heine-Borel)**.** *$A \subset \mathbb{R}^m$ ist kompakt g.d.w. A beschränkt und abgeschlossen ist.*

Beweis. Ist A kompakt, so folgt aus Satz 5.20, dass A abgeschlossen ist. Wäre A nicht beschränkt, so könnten wir eine Folge x_n konstruieren, mit $\|x_n\| \geq n$. x_n hat dann keine konvergente Teilfolge. Anderseits, falls A beschränkt ist, so folgt aus Satz 5.5 (Bolzano-Weierstrass), dass jede Folge x_n auf A eine konvergente Teilfolge hat. Da A abgeschlossen ist, muss der Limes der konvergenten Teilfolge in A liegen. Also ist A folgenkompakt und also kompakt. □

6 Stetigkeit

6.1 Definition und elementare Eigenschaften

Wir betrachten in dieser Sektion Funktionen $f : M_1 \to M_2$ zwischen metrischen Räumen. Bemerke, dass, falls f nur auf einer Teilmenge $A \subset M_1$ definiert ist, kann man A als metrischen Raum mit der von M_1 induzierten Metrik betrachten. Deswegen werden wir annehmen, f ist auf dem vollen metrischen Raum M_1 definiert. Wir führen nun den wichtigen Begriff der Stetigkeit von Funktionen zwischen metrischen Räumen ein.

Definition 6.1. *Seien (M_1, d_1), (M_2, d_2) zwei metrische Räume und $f : M_1 \to M_2$ eine Funktion. f heisst stetig an der Stelle $a \in M_1$, falls für jede Folge $(x_n)_{n\in\mathbb{N}}$ in M_1, mit $x_n \to a$ bezüglich d_1 gilt $f(x_n) \to f(a)$ bezüglich d_2. Die Funktion f heisst stetig, falls sie steig an der Stelle a ist, für jede $a \in M_1$.*

Beispiele.

- Sei $f : \mathbb{R} \to \mathbb{R}$ durch $f(x) = x^m$, für ein $m \in \mathbb{N}$, definiert. Dann ist f stetig. In der Tat, falls $x_n \to a$ in $\mathbb{R}$, so folgt aus der Rechenregel für Konvergenz von Folgen, dass $x_n^m \to a^m$. Auch $f : \mathbb{R}_+ \to \mathbb{R}$, definiert durch $f(x) = x^{p/q}$, für $p \in \mathbb{Z}, q \in \mathbb{N}\backslash\{0\}$ ist stetig (hier ist $\mathbb{R}_+ = \{x \in \mathbb{R} : x > 0\}$).

- Die Funktion $f : \mathbb{R} \to \mathbb{R}$, die durch $f(x) = 0$ für $x \leq 0$ und $f(x) = 1$ für $x > 0$ definiert wird, ist an der Stelle $x = 0$ nicht stetig. Sei zB. $x_n = 1/n$. Es gilt $f(x_n) = 1$ für alle $n \in \mathbb{N}$ und also $f(x_n) \to 1$. Da aber $f(0) = 0$, folgt, dass $f(x_n) \not\to f(0)$.

Proposition 6.2. *Seien M_1, M_2 metrische Räume.*

- *Konstante Funktionen sind stetig.*

- *Seien $f : M_1 \to M_2$ stetig an der Stelle $a \in M_1$ und $g : M_2 \to M_3$ stetig an der Stelle $f(a)$. Dann ist $g \circ f : M_1 \to M_3$ stetig an der Stelle a.*

Beweis. i) Sei $f(x) = c \in M_2$ für alle $x \in M_1$. Für jede Folge x_n mit $x_n \to a$ in M_1, gilt $f(x_n) = f(a)$ für alle $n \in \mathbb{N}$. Insbesondere $f(x_n) \to f(a)$. ii) Gilt $x_n \to a$ in M_1, so impliziert die Stetigkeit von f an der Stelle a, dass $f(x_n) \to f(a)$ in M_2. Nun aber impliziert die Stetigkeit von g an der Stelle $f(a)$, dass $g(f(x_n)) \to g(f(a))$. □

Für Funktionen $f : M \to \mathbb{R}$ mit Werten in $\mathbb{R}$ erhalten wir auch die folgenden Rechenregeln.

Proposition 6.3. *Sei M ein metrischer Raum und seien $f, g : M \to \mathbb{R}$ stetig an der Stelle $a \in M$. Dann sind auch die Funktionen $f + g : M \to \mathbb{R}$ und $fg : M \to \mathbb{R}$, die durch $(f + g)(x) = f(x) + g(x)$ und $(fg)(x) = f(x)g(x)$ definiert werden, stetig an der Stelle $a \in M$. Ist $g(a) \neq 0$, so ist auch $f/g : M \to \mathbb{R}$ stetig an der Stelle a.*

Beweis. Folgt aus den Rechenregeln für Folgen. □

Eine besondere Rolle spielen Funktionen auf $\mathbb{R}$ oder auf $\mathbb{C}$ mit Werten in $\mathbb{R}$ oder in $\mathbb{C}$. Beispiele solchen Funktionen sind Polynomen, nämlich Funktionen der Typ

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

für $n \in \mathbb{N}$, $x \in \mathbb{R}$ oder $x \in \mathbb{C}$ und Koeffizienten $a_0, \dots, a_n \in \mathbb{R}$ oder in $\mathbb{C}$.

Proposition 6.4. *Polynomen sind überall stetig. Rationale Funktionen, d.h. Funktionen der Form $p(x)/q(x)$ für zwei Polynome p, q sind stetig auf $\{x : g(x) \neq 0\}$.*

Beweis. Folgt aus Proposition 6.3. □

Die intuitive Idee ist, dass eine Funktion f genau dann stetig ist, wenn eine kleine Änderung von $x \in M_1$ nur eine kleine Änderung von $f(x)$ verursacht. Diese Idee wird im folgenden Kriterium mathematisch genau ausgedrückt.

Satz 6.5 (Weierstrass $\varepsilon - \delta$ Kriterium)**.** *Seien M_1, M_2 metrische Räume und $a \in M_1$. Dann ist $f : M_1 \to M_2$ genau dann an der Stelle a stetig , wenn für alle $\varepsilon > 0t$ $\delta > 0$ existier, s.d.*

$$d_1(x, a) < \delta \;\Rightarrow\; d_2(f(x), f(a)) < \varepsilon$$

Beweis. Sei f stetig an der Stelle $a \in M_1$. Nehmen wir an, das $\varepsilon - \delta$ Kriterium sei nicht erfüllt. Dann gibt es $\varepsilon_0 > 0$, s.d. für alle $\delta > 0$ ein $x \in M_1$ existiert, mit $d_1(x,a) < \delta$ und $d_2(f(x), f(a)) \geq \varepsilon_0$. Insbesondere für jede $n \in \mathbb{N}$ finden wir $x_n \in M_1$ mit $d_1(x_n, a) < 1/n$ und $d_2(f(x_n), f(a)) > \varepsilon_0$. Dann ist x_n eine Folge auf M_1 mit $x_n \to a$ aber $f(x_n) \not\to f(a)$, in Widerspruch zur Stetigkeit von f an der Stelle a. Sei nun das $\varepsilon - \delta$ Kriterium an der Stelle $a \in M_1$ erfüllt. Wir behaupten, f ist stetig an der Stelle a. Sei x_n eine Folge in M_1, mit $x_n \to a$, und sei $\varepsilon > 0$ gegeben. Aus dem $\varepsilon - \delta$ Kriterium existiert $\delta > 0$ mit

$$d_1(x_n, a) < \delta \;\Rightarrow\; d_2(f(x_n), f(a)) < \varepsilon$$

Für $x_n \to a$ existiert $n_0 \in \mathbb{N}$ mit

$$n > n_0 \;\Rightarrow\; d_1(x_n, a) < \delta$$

Also gilt für alle $n \in \mathbb{N}$ mit $n > n_0$, $d_2(f(x_n), f(a)) < \varepsilon$. Das zeigt, dass $f(x_n) \to f(a)$. Also ist f stetig an der Stelle a. □

Beispiele.

- Sei $M = \mathbb{Q}$ und $f : M \to \mathbb{R}$, mit $f(x) = 0$, falls $x^2 < 2$ und $f(x) = 1$, falls $x^2 > 2$. Dann ist f überall stetig. Für beliebige $a \in \mathbb{Q}$, ist $\delta = |a - \sqrt{2}| > 0$. Dann gilt, für ein beliebiges $\varepsilon > 0$,

$$|x - a| < \delta \;\Rightarrow\; f(x) = f(a) \;\Rightarrow\; |f(x) - f(a)| = 0 < \varepsilon$$

- Die Funktion $f : \mathbb{R}\backslash\{0\} \to \mathbb{R}$, die durch $f(x) = 1/x$ definiert ist, ist überall stetig. Seien, in der Tat, $a \neq 0$ und $\varepsilon > 0$ gegeben. Für $|x| > |a|/2$ gilt

$$|f(x) - f(a)| = \left|\frac{1}{x} - \frac{1}{a}\right| = \frac{1}{|a||x|}|x - a| \leq \frac{2}{|a|^2}\,|x - a|$$

  Wir setzen also $\delta = \min(|a|/2, \varepsilon|a|^2/2)$. Dann

$$|x - a| < \delta \;\Rightarrow\; |x| > |a|/2 \;\wedge\; |x - a| < \varepsilon|a|^2/2 \;\Rightarrow\; |f(x) - f(a)| < \varepsilon$$

- Sei $f : \mathbb{R} \to \mathbb{R}$ durch

$$f(x) = \begin{cases} 0 & \text{falls } x \in \mathbb{R}\backslash\mathbb{Q} \\ 1/q & \text{falls } x = p/q \text{ und } p, q \text{ Teilerfremd sind} \end{cases}$$

  definiert. Dann gilt: f ist stetig an der Stelle a, für alle $a \in \mathbb{R}\backslash\mathbb{Q}$ und f ist unstetig an der Stelle a, für alle $a \in \mathbb{Q}$. Beweis: Sei $a \in \mathbb{R}\backslash\mathbb{Q}$. O.B.d.A. nehmen wir an, dass $a > 0$. Wir finden $n \in \mathbb{N}$, mit $a < n$. Sei nun $\varepsilon > 0$ gegeben. Wir finden $Q \in \mathbb{N}_+$ mit $1/Q < \varepsilon$. Dann setzen wir

$$\delta := \min\{|a - p/q| : p, q \in \mathbb{N} \text{ mit } 1 \leq q \leq Q \text{ und } 0 \leq p \leq 2nq\}$$

  Da $a \notin \mathbb{Q}$, gilt $\delta > 0$. Ist nun $|x - a| < \delta$, dann unterscheiden wir zwei Fälle: Ist $x \in \mathbb{R}\backslash\mathbb{Q}$, dann gilt $f(x) = 0 = f(a)$, und also $|f(x) - f(a)| < \varepsilon$. Ist dagegen $x \in \mathbb{Q}$, so gilt $x = p_1/q_1$ für geeignete $p_1, q_1 \in \mathbb{N}$ ($\delta < |a| = a$ und $|x - a| < \delta$ implizieren, dass $x > 0$ ist). Aus Definition von δ muss entweder $q_1 > Q$ sein, oder $q_1 \leq Q$ und $p_1 > 2nq_1$. Im zweiten Fall wäre aber $x = p_1/q_1 > 2n > 2a$ und deswegen $|x - a| > a$ im Widerspruch zur Annahme, dass $|x - a| < \delta < a$. Also, es muss $q_1 > Q$ gelten, und deswegen $f(x) = 1/q_1 < 1/Q < \varepsilon$. Da $f(a) = 0$, krigen wir auch in diesem Fall, dass $|f(x) - f(a)| < \varepsilon$. Das $\varepsilon - \delta$ Kriterium ist also erfüllt. Der Beweis der Unstetigkeit von f an a, für alle $a \in \mathbb{Q}$ ist als Übung gelassen.

6.2 Grenzwerte von Funktionen

Definition 6.6. *Sei M ein metrischer Raum und $A \subset M$. Ein Punkt $a \in M$ heisst ein Häufungspunkt von A, falls für jede Umgebung U von a gilt*

$$(U \backslash \{a\}) \cap A \neq \emptyset$$

Man nennt $U \backslash \{a\}$ eine punktierte Umgebung von a.

Beispiele. $M = \mathbb{R}$, $A = \mathbb{R}$. Jedes $x \in \mathbb{R}$ ist Häufungspunkt von $\mathbb{R}$. $M = \mathbb{R}$, $A = \mathbb{Q}$. Jedes $x \in \mathbb{R}$ ist Häufungspunkt von $\mathbb{Q}$. Analog $M = \mathbb{R}$, $A = \mathbb{R} \backslash \mathbb{Q}$. Jeder Punkt in $\mathbb{R}$ ist Häufungspunkt von $\mathbb{R} \backslash \mathbb{Q}$. $M = \mathbb{R}$, $A = \mathbb{Z}$. Kein Punkt in $\mathbb{R}$ ist Häufungspunkt von $\mathbb{Z}$.

Bemerkung. Ist a ein Häufungspunkt von A so, enthält jede Umgebung von a unendlich viele Elemente von A. Wäre nämlich U eine Umgebung von a, mit

$$(U \backslash \{a\}) \cap A = \{b_1, \dots, b_n\}$$

endlich, dann würde mit $\varepsilon = \min\{d(a, b_1), \dots, d(a, b_n)\} > 0$,

$$(B_\varepsilon(a) \backslash \{a\}) \cap A = \emptyset$$

gelten, in Widerspruch zur Tatsache, dass a ein Häufungspunkt von A ist.

Proposition 6.7. *Sei M ein metrischer Raum und $A \subset M$. Dann ist $a \in M$ ein Häufungspunkt von A g.d.w. eine Folge $(x_n)_{n \in \mathbb{N}}$ in $A \backslash \{a\}$ mit $x_n \to a$ existiert.*

Beweis. Übung □

Definition 6.8. *Seien $(M_1, d_1), (M_2, d_2)$ metrische Räume, $A \subset M_1$ und $a \in M_1$ ein Häufungspunkt von A. Sei $f : A \backslash \{a\} \to M_2$. Wir schreiben*

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

falls für alle $\varepsilon > 0$ ein $\delta > 0$ existiert, s.d.

$$x \in A \ \wedge \ 0 < d_1(x, a) < \delta \ \Rightarrow \ d(f(x), b) < \varepsilon$$

Gemäss dieser Definition ist es möglich, dass a ein Häufungspunkt von zwei verschiedenen Mengen $A, B \subset M$ ist, mit $\lim_{x \to a, x \in A} f(x) \neq \lim_{x \to a, x \in B} f(x)$.

Verlangen wir, dass A eine Umgebung von a ist, dann ist der Limes $\lim_{x \to a, x \in A} f(x)$ unabhängig aus der Wahl von A. Es ist in der Tat klar, dass für $A_1 \subset A_2$ gilt $\lim_{x \to a, x \in A_1} f(x) = b$ gdw. $\lim_{x \to a, x \in A_2} f(x) = b$. Anderseits, für zwei beliebige Umgebungen A_1, A_2 von a ist $A_1 \cap A_2$ wieder eine Umgebung. Da $A_1 \cap A_2 \subset A_1$ und $A_1 \cap A_2 \subset A_2$ ist

$$\lim_{x \to a, x \in A_1} f(x) = b \iff \lim_{x \to a, x \in A_1 \cap A_2} f(x) = b \iff \lim_{x \to a, x \in A_2} f(x) = b$$

Wir benutzen diese Bemerkung um $\lim_{x \to a} f(x)$ zu definieren.

Definition 6.9. *Die Schreibweise*

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

bedeutet, dass eine Umgebung A von a, mit f definiert auf $A\backslash\{a\}$ existiert, und dass

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

Bemerkung: Für $\lim_{x\to a} f(x)$ gelten die üblichen Rechenregeln für Grenzwerten. Insbesondere ist der Grenzwert, falls er existiert, eindeutig. Weiter, falls f, g reelwertige oder complexwertige Funktionen sind, dann sind Grenzwerte von Summen und Multiplikationen durch Summen und Multiplikationen der Grenzwerte gegeben (ähnlich für Subtraktion und Division, falls der Nenner nicht Null ist).

Proposition 6.10. *Seien M_1, M_2 metrische Räume, $A \subset M_1$ und $a \in A$ ein Häufungspunkt von A. Eine Funktion f, definiert auf A, ist stetig an der Stelle a gdw.* $\lim_{x\to a} f(x)$ *existiert und*

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

Beweis. Folgt aus dem $\varepsilon - \delta$-Kriterium. □

Proposition 6.11 (Fortsetzung durch Stetigkeit)**.** *Sei M ein metrischer Raum und a ein Häufungspunkt von M. Sei f mindestens auf $M\backslash\{a\}$ definiert und es existiere*

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

Wir definieren eine neue Funktion $\widetilde{f}$ auf M durch

$$\widetilde{f}(x) = \begin{cases} f(x) & \text{falls } x \neq a \\ b & \text{falls } x = a \end{cases}$$

Dann ist $\widetilde{f}$ stetig an der Stelle a.

Beweis. Sei $\varepsilon > 0$ fest gewählt. Für $\lim_{x\to a} f(x) = b$ existiert $\delta > 0$, s.d.

$$0 < d(x,a) < \delta \;\Rightarrow\; d(f(x), b) < \varepsilon \tag{9}$$

Wir behaupten nun, dass $d(x,a) < \delta$ impliziert, dass $d(\widetilde{f}(x), \widetilde{f}(a)) = d(\widetilde{f}(x), b) < \varepsilon$. Um diese Behauptung zu zeigen, unterscheiden wir zwei Fälle. Ist $d(x,a) < \delta$ und $x \neq a$, so gilt $0 < d(x,a) < \delta$ und $\widetilde{f}(x) = f(x)$. Aus (9) erhalten wir, dass $d(\widetilde{f}(x), b) < \varepsilon$. Anderseits, gilt $x = a$, dann ist $\widetilde{f}(x) = b$, und deswegen $d(\widetilde{f}(x), b) = 0 < \varepsilon$. □

Beispiel: Sei

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

für alle $x \in \mathbb{R}\backslash\{1\}$. Es gilt

$$\lim_{x\to 1} f(x) = 2$$

Die Funktion

$$\widetilde{f}(x) = \begin{cases} \frac{x^2-1}{x-1} & \text{falls } x \neq 1 \\ 2 & \text{falls } x = 1 \end{cases}$$

ist stetig an der Stelle $x = 1$. Da f als rationale Funktion stetig auf $\mathbb{R}\backslash\{1\}$, ist $\widetilde{f}$ stetig auf $\mathbb{R}$ (in diesem Fall ist es einfach zu sehen, dass $\widetilde{f}(x) = x + 1$, für alle $x \in \mathbb{R}$).

Wir kommen nun zurück zur Definition 6.8 von $\lim_{x\to a, x\in A} f(x)$ für eine beliebige Menge A mit der Eigenschaft, dass a ein Häufungspunkt von A ist. Wir verlangen hier nicht, dass A eine Umgebung von a ist. Eine Anwendung dieser Definition sind einseitige Grenzwerte für Funktionen auf $\mathbb{R}$.

Definition 6.12. *Sei $f : A \to M$, für eine Teilmenge $A \subset \mathbb{R}$ und ein beliebiger metrischer Raum M. Wir schreiben*

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

falls $r > 0$ existiert, s.d. f auf dem Intervall $(a - r; a)$ definiert ist und

$$\lim_{x\to a, x\in(a-r;a)} f(x) = b$$

Ähnlich bedeutet die Schreibweise

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

dass ein $r > 0$ existiert, s.d. f auf $(a; a + r)$ definiert ist, und dass

$$\lim_{x\to a, x\in(a;a+r)} f(x) = b$$

Proposition 6.13. *Sei f eine Funktion einer reellen Veränderlichen, die in einer Umgebung von $a \in \mathbb{R}$ definiert ist. Dann existiert $\lim_{x\to a} f(x)$ g.d.w. $\lim_{x\uparrow a} f(x)$ und $\lim_{x\downarrow a} f(x)$ existieren, und*

$$\lim_{x\uparrow a} f(x) = \lim_{x\downarrow a} f(x)$$

In diesem Fall gilt auch

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

Beweis. Übung. $\square$

Bemerkung: Es folgt aus der Proposition, dass eine Funktion f einer reellen Veränderlichen genau dann stetig an der Stelle a ist, wenn $\lim_{x\uparrow a} f(x)$ und $\lim_{x\downarrow a} f(x)$ existieren, und

$$\lim_{x\uparrow a} f(x) = \lim_{x\downarrow a} f(x) = f(a)$$

Definition 6.14. *Eine Funktion $f : [a; b] \to \mathbb{R}$ heisst monoton wachsend, falls $a \leq x_1 \leq x_2 \leq b$ impliziert, dass $f(x_1) \leq f(x_2)$. f heisst monoton fallend, falls $a \leq x_1 \leq x_2 \leq b$ impliziert, dass $f(x_1) \geq f(x_2)$.*

Für monotone Funktionen existieren immer einseitige Grenzwerte.

Proposition 6.15. *Sei $f : [a; b] \to \mathbb{R}$ monoton wachsend, $a < c < b$. Dann*

$$f(c^-) := \lim_{x \uparrow c} f(x) = \sup \{f(x) : x < c\}$$
$$f(c^+) := \lim_{x \downarrow c} f(x) = \inf \{f(x) : x > c\}$$

und $f(c^-) \leq f(c) \leq f(c^+)$.

Bemerkung: Es folgt, dass eine monotone Funktion $f : [a; b] \to \mathbb{R}$ an der Stelle $a < c < b$ genau dann stetig ist, falls $f(c^-) = f(c^+)$.

Beweis. Sei $s = \sup\{f(x) : x < c\}$ und $\varepsilon > 0$. Dann es existiert $x_\varepsilon < c$ mit $f(x_\varepsilon) > s - \varepsilon$. Für $x_\varepsilon < x < c$, haben wir

$$s - \varepsilon < f(x_\varepsilon) \leq f(x) \leq s$$

Also, mit $\delta = c - x_\varepsilon$ haben wir

$$x \in (c - \delta, c) \;\Rightarrow\; |f(x) - s| < \varepsilon$$

Da $\varepsilon > 0$ beliebig ist, folgt, dass

$$\lim_{x \uparrow c} f(x) = s$$

Analog kann man zeigen, dass

$$\lim_{x \downarrow c} f(x) = \inf\{f(x) : x > c\}$$

□

Wie für Folgen, kann man auch für Funktionen uneigentliche Grenzwerte definieren. Sei $f : A \to M$ für eine Teilmenge $A \subset \mathbb{R}$ und ein beliebiger metrischer Raum M. Dann bedeutet die Schreibweise

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

dass ein $r_0 \in \mathbb{R}$ existiert, s.d. f für alle $x > r_0$ definiert ist, und, dass für alle $\varepsilon > 0$ ein $r > r_0$ existiert, mit $d(f(x), b) < \varepsilon$ für alle $x > r$. Analog wird die Schreibweise

$$\lim_{x \to -\infty} f(x) = b$$

definiert. Sei nun M ein metrischer Raum und f eine reelwertige Funktion, definiert auf einer Teilmenge von M. Die Schreibweise

$$\lim_{x \to a, x \in A} f(x) = +\infty$$

bedeutet, dass $a \in M$ ein Häufungspunkt von $A \subset M$ ist, dass f mindestens auf $A \backslash \{a\}$ definiert ist, und dass für alle $r \in \mathbb{R}$ $\delta > 0$ existiert, s.d.

$$x \in A \text{ und } 0 < d(x, a) < \delta \;\Rightarrow\; f(x) > r$$

Die Notation
$$\lim_{x\to a} f(x) = +\infty$$
bedeutet dagegen, dass eine Umgebung A von $a \in M$ existiert, s.d. f auf $A\backslash\{a\}$ definiert ist, und dass
$$\lim_{x\to a, x\in A} f(x) = +\infty$$
Analog ist die Bedeutung von $\lim_{x\to a, x\in A} f(x) = -\infty$ und von $\lim_{x\to a} f(x) = -\infty$ definiert.
Beispiele: Es gilt
$$\lim_{x\downarrow 0} \frac{1}{x} = +\infty, \quad \lim_{x\uparrow 0} \frac{1}{x} = -\infty, \quad \lim_{x\to 0} \frac{1}{x} \text{ existiert nicht}, \quad \lim_{x\to 0} \frac{1}{x^2} = +\infty$$

6.3 Zwischenwertsatz

In dieser Sektion betrachten wir reel-wertige Funktionen einer reellen Veränderlichen.

Satz 6.16. *Sei $f : [a; b] \to \mathbb{R}$ stetig, mit $f(a) < 0 < f(b)$. Dann existiert $c \in (a; b)$ mit $f(c) = 0$. Die Behauptung gilt auch falls $f(b) < 0 < f(a)$.*

Beweis. Sei $A = \{x \in [a; b] : f(x) < 0\}$.
$$f(a) < 0, \quad f \text{ stetig} \quad \Rightarrow \quad \exists\, \delta_1 > 0 \;:\; f(x) < 0 \quad \text{für alle } x \in [a; a + \delta_1]$$
D.h. $[a; a+\delta_1] \subset A$. Ähnlicherweise existiert $\delta_2 > 0$ mit $[b-\delta_2; b] \cap A = \emptyset$. Sei $c = \sup A$. Dann gilt $a + \delta_1 \leq c \leq b - \delta_2$. Wir behaupten $f(c) = 0$. Nehmen wir an, $f(c) > 0$. Dann existiert, wegen Stetigkeit, $\eta > 0$ mit $f(x) > 0$ auf $[c - \eta, c + \eta]$. Das impliziert, dass $A \cap [c - \eta; c] = \emptyset$, in Widerspruch zu $c = \sup A$. Anderseits, falls $f(c) < 0$, existiert $\eta > 0$ mit $f(x) < 0$ auf $[c - \eta, c + \eta]$. Das impliziert, dass $c + \eta \in A$, in Widerspruch zu $c = \sup A$. Es bleibt nur die Möglichkeit, dass $f(c) = 0$. $\square$

Proposition 6.17. *Sei $f : [a; b] \to \mathbb{R}$ streng monoton wachsend und stetig. Sei $c = f(a)$ und $d = f(b)$. Dann ist $f : [a; b] \to [c; d]$ bijektiv, und $f^{-1} : [c; d] \to [a; b]$ ist stetig. Mit anderen Worten: Die Umkehrabbildung einer stetigen, streng monotonen Funktion ist stetig.*

Beweis. f streng monoton wachsend impliziert, dass f injektiv ist, und dass $c \leq f(x) \leq d$ für alle $a \leq x \leq b$. Wir behaupten nun, dass $f : [a; b] \to [c; d]$ surjektiv ist. Sei $c < y_0 < d$, und setze $g(x) = f(x) - y_0$. Dann g ist stetig auf $[a; b]$, $g(a) = f(a) - y_0 = c - y_0 < 0$ und $g(b) = f(b) - y_0 = d - y_0 > 0$. Der Zwischenwertsatz impliziert, dass ein $x_0 \in (a; b)$ existiert, mit $g(x_0) = 0$, d.h. mit $f(x_0) = y_0$. Um die Stetigkeit der Inverse zu zeigen, wenden wir das $\varepsilon - \delta$ Kriterium an. Sei $y_0 \in (c; d)$ und $\varepsilon > 0$ gegeben. Wir setzen $x_0 = f^{-1}(y_0) \in (a; b)$. O.B.d.A. können wir annehmen, dass $a < x_0 - \varepsilon < x_0 + \varepsilon < b$ (sonst ersetzen wir die Fehlergrenze $\varepsilon > 0$ durch ein $\widetilde{\varepsilon} < \varepsilon$ klein genug). Sei $x_1 = x_0 - \varepsilon$, $x_2 = x_0 + \varepsilon$,$y_1 = f(x_1)$, $y_2 = f(x_2)$. Aus der Monotonie von f folgt, dass $y_1 < y_0 < y_2$ und dass $f : [x_1; x_2] \to [y_1; y_2]$ bijektiv. Wir wählen $\delta > 0$ so klein, dass $(y_0 - \delta; y_0 + \delta) \subset (y_1; y_2)$. Dann
$$|y - y_0| < \delta \;\Rightarrow\; y \in (y_1; y_2) \;\Rightarrow\; f^{-1}(y) \in (x_0; x_1) \;\Rightarrow\; |f^{-1}(y) - f^{-1}(y_0)| < \varepsilon$$
Das zeigt die Stetigkeit von f^{-1} an der Stelle y_0, für jede $y_0 \in (c; d)$. Stetigkeit von f^{-1} an der Stelle c und an der Stelle d kann analog gezeigt werden. $\square$

6.4 Stetigkeit und Topologie

In dieser Sektion geben wir eine neue Charakterisierung der Stetigkeit von Funktionen zwischen metrischen Räumen, die nicht direkt von den Metriken abhängt, sondern nur davon, was offene Mengen sind. Diese Charakterisierung erlaubt es, den Begriff von Stetigkeit auf topologischen Räumen zu verallgemeinern, wo keine Metrik vorhanden ist, sondern nur eine Topologie, die bestimmt, welche Mengen offen sind.

Proposition 6.18. *Seien $(M_1, d_1), (M_2, d_2)$ metrische Räume, und $f : M_1 \to M_2$.*

a) f ist stetig an der Stelle $a \in M_1$ g.d.w. für jede Umgebung U von $f(a)$, $f^{-1}(U)$ eine Umgebung von a ist.

b) f ist überall auf M_1 stetig g.d.w.

$$G \subset M_2 \text{ offen } \Rightarrow f^{-1}(G) \subset M_1 \text{ offen}$$

Beweis. a) Wir nehmen zunächst an, f ist stetig an der Stelle a. Sei $U \subset M_2$ eine Umgebung von $f(a)$. Dann existiert $\varepsilon > 0$, s.d. $B_\varepsilon(f(a)) \subset U$. Nach dem $\varepsilon - \delta$-Kriterium existiert $\delta > 0$, s.d.

$$d_1(x, a) < \delta \Rightarrow f(x) \in B_\varepsilon(f(a)) \Rightarrow f(x) \in U \Rightarrow x \in f^{-1}(U)$$

Das bedeutet, dass $B_\delta(a) \subset f^{-1}(U)$. D.h. $f^{-1}(U)$ eine Umgebung von a ist. Nehmen wir nun an, dass, für jede Umgebung U von $f(a)$, $f^{-1}(U)$ eine Umgebung von a ist. Wir möchten dann zeigen, dass f stetig an der Stelle a ist. Wir zeigen dafür, dass das $\varepsilon - \delta$ Kriterium erfüllt ist. Sei $\varepsilon > 0$ fest. $B_\varepsilon(f(a))$ ist eine Umgebung von $f(a)$. Aus Annahme, ist $f^{-1}(B_\varepsilon(f(a)))$ eine Umgebung von a. D.h. es existiert $\delta > 0$, s.d. $B_\delta(a) \subset f^{-1}(B_\varepsilon(f(a)))$. Wir erhalten:

$$\begin{aligned} d_1(x, a) < \delta \Rightarrow x \in B_\delta(a) \Rightarrow x \in f^{-1}(B_\varepsilon(f(a))) \\ \Rightarrow f(x) \in B_\varepsilon(f(a)) \Rightarrow d_2(f(x), f(a)) < \varepsilon \end{aligned}$$

Also ist f stetig an der Stelle a. b) Sei f stetig und $G \subset M_2$ offen. Wir behaupten, dass $f^{-1}(G)$ offen ist. Sei dafür $a \in f^{-1}(G)$. Dann ist $f(a) \in G$. Da G offen ist, existiert eine Umgebung U von $f(a)$ mit $U \subset G$. Aus a) ist $f^{-1}(U) \subset f^{-1}(G)$ eine Umgebung von a. Da $a \in f^{-1}(G)$ beliebig ist, es folgt, dass $f^{-1}(G)$ offen ist. Nehmen wir nun an, dass $f^{-1}(G)$ offen ist, für alle $G \subset M_2$ offen. Wir behaupten, dass f überall stetig ist. Sei $a \in M_1$ und $\varepsilon > 0$ beliebig. $B_\varepsilon(f(a))$ ist eine offene Menge in M_2. Aus Annahme, $f^{-1}(B_\varepsilon(f(a)))$ ist offen. Da $a \in f^{-1}(B_\varepsilon(f(a)))$, existiert $\delta > 0$ mit $B_\delta(a) \subset f^{-1}(B_\varepsilon(f(a)))$. Das bedeutet

$$\begin{aligned} d_1(x, a) < \delta \Rightarrow x \in B_\delta(a) \Rightarrow x \in f^{-1}(B_\varepsilon(f(a))) \\ \Rightarrow f(x) \in B_\varepsilon(f(a)) \Rightarrow d_2(f(x), f(a)) < \varepsilon \end{aligned}$$

Also f ist stetig an der Stelle a, für alle $a \in M_1$. □

Bemerkung: Es folgt aus der Proposition, dass $f : M_1 \to M_2$ stetig ist g.d.w. $f^{-1}(F)$ abgeschlossen ist, für alle abgeschlossenen $F \subset M_2$.

6.5 Stetige Funktionen auf kompakten Mengen

Wir haben schon gesehen, dass der Begriff von Kompaktheit sehr nützlich ist, wann man die Konvergenz von Folgen untersucht. Auf kompakten metrischen Räumen (oder, äquivalent, auf kompakten Teilmengen von metrischen Räumen) hat nämlich jede Folge eine konvergente Teilfolge. Wir werden in dieser Sektion sehen, dass der Begriff von Kompaktheit auch bei der Untersuchung von stetigen Funktionen nützlich ist; stetige Funktionen auf kompakten Mengen haben nämlich wichtige zusätzliche Eigenschaften.

Die erste Bemerkung ist, dass stetige Funktionen kompakte Mengen auf kompakten Mengen abbilden.

Proposition 6.19. *Seien M_1, M_2 metrische Räume. $A \subset M_1$ kompakt, und $f : A \to M_2$ stetig. Dann ist $f(A) \subset M_2$ kompakt.*

Beweis. Sei (y_n) eine Folge in $f(A)$. Dann finden wir für alle $n \in \mathbb{N}$ ein $x_n \in A$ mit $f(x_n) = y_n$. Das definiert eine Folge (x_n) in A. Da A kompakt ist (und deswegen auch folgenkompakt), existiert eine Teilfolge n_j und eine $x \in A$ mit $x_{n_j} \to x$. Da f stetig ist, erhalten wir $y_{n_j} = f(x_{n_j}) \to f(x) \in f(A)$. Wir haben gezeigt, dass jede Folge in $f(A)$ eine in $f(A)$ konvergente Teilfolge hat. Also ist $f(A)$ folgenkompakt und deswegen kompakt. $\square$

Der nächste Satz vom Maximum spielt eine sehr wichtige Rolle in der Analysis, insbesondere in der Untersuchung partieller Differentialgleichungen. Er besagt, dass stetige reelwertige Funktionen auf kompakten Mengen das Supremum und das Infimum annehmen. Diese Aussage ist offenbar falsch, falls man stetige Funktionen auf nicht kompakten Mengen untersucht (z.B. die Funktion $f(x) = 1/x$ ist auf $(0;1)$ stetig, aber sie ist nicht beschränkt, und sie nimmt den Wert $\sup\{f(x) : x \in (0;1)\} = \infty$ nicht an).

Satz 6.20 (Satz vom Maximum)**.** *Sei A ein kompakter metrischer Raum und $f : A \to \mathbb{R}$ stetig. Dann ist f beschränkt auf A und f nimmt $\sup_{x\in A} f(x)$ unf $\inf_{x\in A} f(x)$ an. D.h. es existieren $x_{min}, x_{max} \in A$ mit*

$$f(x_{min}) \leq f(x) \leq f(x_{max}),$$

für alle $x \in A$.

Beweis. Wir zeigen zunächst, dass f nach oben beschränkt ist. Nehmen wir an, f ist nicht nach oben beschränkt. Dann finden wir, für alle $n \in \mathbb{N}$, $x_n \in A$, mit $f(x_n) \geq n$. (x_n) definiert eine Folge in A. Da A kompakt ist, existiert eine Teilfolge n_j und $x \in A$ mit $x_{n_j} \to x$. Da f stetig ist, muss $f(x_{n_j}) \to f(x) \in \mathbb{R}$ konvergieren, was natürlich der Annahme $f(x_{n_j}) \geq n_j \to \infty$ widerspricht. Analog kann man zeigen, f ist nach unten beschränkt. Wir zeigen nun, dass f den Wert

$$s = \sup\{f(x) : x \in A\}$$

annimmt. Aus der Definition von Supremum finden wir, für alle $n \in \mathbb{N}$, ein $x_n \in A$ mit $s - (1/n) \leq f(x_n) \leq s$. (x_n) definiert dann eine Folge auf A. Da A kompakt ist, existiert eine Teilfolge n_j und ein $x \in A$ mit $x_{n_j} \to x$. Da f stetig ist, gilt $f(x_{n_j}) \to f(x)$. Aus Wahl von x_n gilt auch $f(x_{n_j}) \to s$. Das impliziert, dass $f(x) = s$, und zeigt die Behauptung. Analog zeigt man, dass das Infimum angenommen wird. $\square$

Definition 6.21. *Seien (M_1, d_1) und (M_2, d_2) metrische Räume. Eine Funktion $f : M_1 \to M_2$ heisst gleichmässig stetig falls, für alle $\varepsilon > 0$ $\delta > 0$ existiert, s.d.*

$$d_1(x, y) < \delta \;\Rightarrow\; d_2(f(x), f(y)) < \varepsilon$$

Bemerkung: Vergleiche diese Definition mit dem $\varepsilon-\delta$ Kriterium in Satz 6.5. Die Differenz zwischen Stetigkeit und gleichmässiger Stetigkeit, ist, dass für stetige Funktionen, δ darf von x (und natürlich auch von ε) abhängen. Bei gleichmässig stetige Funktionen, dagegen, δ nur von ε abhängen darf, nicht von x. Dasselbe δ muss für alle $x \in M_1$ gelten.

Beispiele: Die Funktion $f(x) = x^2$ auf $[0; 1]$ ist gleichmässig stetig, weil

$$|x^2 - y^2| = |x - y||x + y| \leq 2|x - y| \leq \varepsilon \quad \text{falls } |x - y| \leq \varepsilon/2 =: \delta$$

und δ ist unabhängig aus x. Die Funktion $f(x) = x^2$ auf $[0; +\infty)$ ist dagegen nicht gleichmässig stetig, weil für feste $|x - y|$ kann $|x^2 - y^2| = |x - y||x + y|$ beliebig gross werden, indem wir x (und y, damit $|x - y|$ fest bleibt) genügend gross wählen.

Satz 6.22. *Eine stetige Funktion auf einer kompakten Menge ist gleichmässig stetig.*

Beweis. Sei A kompakt (mit Metrik d_1), (M_2, d_2) ein beliebiger metrischer Raum , $f : A \to M_2$ stetig und $\varepsilon > 0$. Für jedes $x \in A$ finden wir ein $\delta_x > 0$ s.d.

$$d_1(x, y) < \delta_x \;\Rightarrow\; d_2(f(x), f(y)) < \varepsilon/2$$

Wir definieren $G_x = B_{\delta_x/2}(x)$. $(G_x)_{x \in A}$ ist eine offene Überdeckung von A. Da A kompakt ist, existiert eine endliche Teilüberdeckung. D.h. es existieren $x_1, \dots, x_n \in A$ mit

$$A \subset G_{x_1} \cup \dots \cup G_{x_n}$$

Wir setzen nun $\delta = (1/2) \min\{\delta_{x_1}, \dots, \delta_{x_n}\}$. Wir behaupten nun, dass,für beliebige $x, y \in A$,

$$d_1(x, y) < \delta \;\Rightarrow\; d_2(f(x), f(y)) < \varepsilon\,.$$

In der Tat, für beliebige $x \in A$ existiert $j \in \{1, \dots, n\}$ mit $x \in G_{x_j}$. Dann gilt $d_1(x, x_j) < \delta_{x_j}/2$ und also $d_2(f(x), f(x_j)) < \varepsilon/2$. Gilt weiter $d_1(x, y) < \delta$, so muss auch $d_1(x, y) < \delta_{x_j}/2$ und

$$d_1(y, x_j) \leq d_1(y, x) + d_1(x, x_j) < \delta_{x_j}$$

Deswegen gilt auch

$$d_2(f(y), f(x_j)) < \varepsilon/2$$

Wir finden, dass

$$d_2(f(x), f(y)) \leq d_2(f(x), f(x_j)) + d_2(f(x_j), f(y)) < \varepsilon$$

Das zeigt die Behauptung. □

6.6 Funktionenfolgen und Stetigkeit

Wir betrachten in dieser Sektion Folgen von Funktionen. Sei A eine beliebige Menge und (M, d) ein metrischer Raum. Eine Funktionenfolge ist eine Folge $(f_n)_{n\in\mathbb{N}}$, wobei für jedes $n \in \mathbb{N}$, $f_n : A \to M$ eine Funktion auf A ist, mit Werten auf M. Wir sagen die Funktionenfolge f_n konvergiert punktweise gegen eine Funktion $f : A \to M$, falls, für alle $x \in A$, $f_n(x) \to f(x)$ (wir benutzen hier die metrische Struktur auf M).

Beispiel: Sei $A = M = [0; 1]$ (mit der Standardmetrik $d(x, y) = |x - y|$). Für $n \in \mathbb{N}$ sei $f_n : A \to M$ durch $f_n(x) = x^n$ definiert. Für $0 \leq x < 1$ gilt $x^n \to 0$, als $n \to \infty$. Für $x = 1$ gilt $x^n \to 1$. Es gilt also $f_n \to f$ punktweise, wobei

$$f(x) = \begin{cases} 0 & \text{falls } 0 \leq x < 1 \\ 1 & \text{falls } x = 1 \end{cases}$$

In diesem Beispiel ist $A = [0; 1]$ ein metrischer Raum. Für alle $n \in \mathbb{N}$ ist die Funktion f_n stetig. Dagegen ist f nicht stetig. Das Beispiel zeigt, dass Stetigkeit bei punktweiser Konvergenz nicht erhalten bleibt.

Definition 6.23. *Die Folge $f_n : A \to M$ konvergiert gleichmässig auf A gegen $f : A \to M$, falls*

$$\forall \varepsilon > 0 \, \exists n_0 \in \mathbb{N} : \forall n > n_0, \forall x \in A \Rightarrow d(f_n(x), f(x)) < \varepsilon.$$

Gleichmässig bedeutet, dass n_0 unabhängig von x ist. Äquivalent: f_n konvergiert gleichmässig gegen f, falls die reelle Folge

$$\sup_{x\in A} d(f_n(x), f(x)) \to 0$$

Bemerkung: $f_n \to f$ gleichmässig impliziert insbesondere, dass $f_n \to f$ punktweise.

Beispiel: Die Funktionenfolge $f_n(x) = x^n$ auf $[0; 1/2]$ konvergiert gleichmässig gegen 0. In der Tat

$$\sup_{x\in[0;1/2]} x^n = 2^{-n} \to 0$$

Dagegen konvergiert die Funktionenfolge $f_n(x) = x^n$ auf $[0; 1]$ nicht gleichmässig. Nehmen wir an, es existiert $f : [0; 1] \to [0; 1]$ mit $f_n \to f$ gleichmässig. Dann müsste $f_n \to f$ auch punktweise. D.h. es müsste $f(x) = 0$ für $0 \leq x < 1$ und $f(1) = 1$ gelten. Da aber

$$\sup_{x\in[0;1]} |x^n - f(x)| \geq \left(1 - \frac{1}{n^2}\right)^n \to 1$$

ist die Konvergenz nicht gleichmässig.

Proposition 6.24. *Seien M_1, M_2 metrische Räume, und, für alle $n \in \mathbb{N}$, $f_n : M_1 \to M_2$ stetig. Sei $f : M_1 \to M_2$, s.d. $f_n \to f$ gleichmässig. Dann ist f stetig.*

Bemerkung: Die Proposition besagt, dass Stetigkeit bei gleichmässiger Konvergenz erhalten bleibt. Hier ist natürlich wichtig, dass der Definitionsbereich auch ein metrischer Raum ist, sonst macht der Begriff Stetigkeit keinen Sinn.

Beweis. Sei $x_0 \in M_1$. Wir zeigen, dass f stetig an der Stelle x_0 ist. Sei $\varepsilon > 0$ fest. Es existiert $n_0 \in \mathbb{N}$ mit

$$\sup_{x \in M_1} d_2(f_{n_0}(x), f(x)) < \varepsilon/3$$

Da f_{n_0} stetig an der Stelle x_0 ist, existiert ein $\delta > 0$ s.d.

$$d_1(x, x_0) < \delta \;\Rightarrow\; d_2(f_n(x), f_n(x_0)) < \varepsilon/3$$

Also gilt für $x \in M_1$ mit $d_1(x, x_0) < \delta$

$$\begin{aligned} d_2(f(x_0), f_n(x)) &\leq d_2(f(x_0), f_{n_0}(x_0)) + d_2(f_{n_0}(x_0), f_{n_0}(x)) + d_2(f_{n_0}(x), f(x)) \\ &\leq 2 \sup_{x \in M_1} d_2(f_{n_0}(x), f(x)) + d_2(f_{n_0}(x_0), f_{n_0}(x)) < \varepsilon \end{aligned}$$

□

Proposition 6.25 (Cauchy-Kriterium für Funktionenfolge)**.** *Sei A eine Menge, (M, d) ein vollständiger metrischer Raum, und $f_n : A \to M$ eine Funktionenfolge. Die Folge f_n konvergiert gleichmässig g.d.w.*

$$\forall\, \varepsilon > 0\; \exists\, n_0 \in \mathbb{N} : \forall n, m > n_0, \forall x \in A \;\Rightarrow\; d(f_n(x), f_m(x)) < \varepsilon \tag{10}$$

Bemerkung: Die Bedingung (10) ist äquivalent zu:

$$\forall \varepsilon > 0,\; \exists\, n_0 \in \mathbb{N} :\; \forall n, m > n_0 \;\Rightarrow\; \sup_{x \in A} d(f_n(x), f_m(x)) < \varepsilon$$

Beweis. Die Tatsache, dass gleichmässige Konvergenz das Cauchy-Kriterium impliziert ist Standard. Wir zeigen, dass das Cauchy-Kriterium (10) die gleichmässige Konvergenz impliziert. Aus der Cauchy-Bedingung folgt, dass für alle $x \in A$, $f_n(x)$ eine Cauchy-Folge auf M ist. Da M vollständig ist, ist für jedes feste $x \in A$, die Folge $f_n(x)$ konvergent. Wir bezeichnen den Grenzwert mit $f(x)$. Wir haben gezeigt, dass $f_n(x) \to f(x)$ für alle $x \in A$ (bis jetzt haben wir nur punktweise Konvergenz bewiesen). Sei nun $\varepsilon > 0$. Dann existiert $n_0 \in \mathbb{N}$ mit

$$d(f_n(x), f_m(x)) < \varepsilon$$

für alle $n, m > n_0$, und für alle $x \in A$. Wir halten nun n und x fest und lassen $m \to \infty$ streben. Für $f_m(x) \to f(x)$ muss auch

$$d(f_n(x), f(x)) < \varepsilon$$

für alle $n > n_0$ und $x \in M$. Also, $f_n \to f$ gleichmässig. □

Sei M ein metrischer Raum. Wir sagen eine Funktion $f : M \to \mathbb{R}$ ist beschränkt, falls

$$\sup_{x \in M} |f(x)| < \infty$$

Wir bezeichnen mit $C_b^0(M)$ den Raum aller stetigen und beschränkten Funktionen $f : M \to \mathbb{R}$. Auf $C_b^0(M)$ kann man eine Addition und eine Multiplikation mit Skalaren aus $\mathbb{R}$ definieren. Für $f, g \in C_b^0(M)$ und $\alpha \in \mathbb{R}$ setzen wir nämlich

$$(f + g)(x) := f(x) + g(x) \quad \text{und} \quad (\alpha f)(x) = \alpha f(x)$$

Dann sind $f+g$ und αf wieder beschrnkt (aus der Dreiecksungleichung) und stetig, also wieder in $C_b^0(M)$. Damit hat $C_b^0(M)$ die Struktur eines Vektorraumes. Auf $C_b^0(M)$ kann man auch eine Norm durch

$$\|f\|_{C_b^0(M)} = \sup_{x\in M} |f(x)|$$

definieren. Es ist dann einfach zu überprüfen, dass die Eigenschaften unten (1) erfüllt sind (und also, dass $\|.\|_{C_b^0(M)}$ eine Norm ist). Damit wird $C_b^0(M)$ zu einem normierten Raum. Die Norm induziert wie immer eine Metrik auf $C_b^0(M)$, die wir mit

$$d_\infty(f,g) = \|f-g\|_{C_b^0(M)}$$

bezeichnen. Damit ist $C_b^0(M)$ auch ein metrischer Raum. Konvergenz im Sinne der Metrik d_∞ ist mit gleichmässiger Konvergenz äquivalent, d.h. eine Funktionenfolge f_n in $C_b^0(M)$ konvergiert gleichmässig gegen $f \in C_b^0(M)$ g.d.w. $d_\infty(f_n, f) \to 0$. Es folgt aus Proposition 6.25, dass jede Cauchy-Folge auf $C_b^0(M)$ (bzgl. d_∞) konvergiert (bzgl. d_∞). D.h. $C_b^0(M)$, versehen mit der Metrik d_∞, ist ein vollständiger metrischer Raum. Analog kann man den Raum $C_b^0(M;\mathbb{C})$ (oder allgemeiner, den Raum $C_b^0(M;\mathbb{R}^m)$) der beschränkten und stetigen Funktionen auf dem metrischen Raum M, mit Werten auf $\mathbb{C}$ (bzw. auf $\mathbb{R}^m$) definieren. Auch $C_b^0(M;\mathbb{C})$ (und $C_b^0(M;\mathbb{R}^m)$), versehen mit der Metrik d_∞ definiert mit dem komplexen Absolutbetrag (bzw. mit der Euklidischen Norm auf $\mathbb{R}^m$), ist ein vollständiger metrischer Raum.

Wir haben auf dem Raum $C_b^0(M)$ eine Struktur eingeführt, die ähnlich zu der Struktur von $\mathbb{R}^m$ ist. Es gibt aber einen wichtigen Unterschied: das Theorem von Bolzano-Weierstrass gilt auf $C_b^0(M)$ nicht. Die Folge $f_n(x) = x^n$ auf $M = [0;1]$ ist ein Beispiel einer beschränkten Folge, die in $C_b^0(M)$ keine konvergente Teilfolge hat. Folgich existieren auf $C_b^0(M)$ beschränkte und abgeschlossene Teilmengen, die nicht folgenkompakt (oder, äquivalent dazu, kompakt) sind.

7 Potenzreihen und Elementarfunktionen

Ein wichtiges Beispiel von Funktionenfolgen mit Werten auf $\mathbb{R}^m$ (insbeondere auf $\mathbb{R}$ oder auf $\mathbb{C}$) sind Funktionenreihen. Sei M ein metrischer Raum und, für jede $n \in \mathbb{N}$, sei $f_n : M \to \mathbb{R}^m$ eine Funktion auf M, mit Werten in $\mathbb{R}^m$. Wir sagen die Funktionenreihe $\sum_{j=0}^\infty f_j(x)$ konvergiert punktweise, falls die Partialsummenfolge $s_n(x) = \sum_{j=0}^n f_j(x)$ punktweise konvergiert. Wir sagen $\sum_{j=0}^\infty f_j(x)$ konvergiert gleichmässig, falls die Partialsummenfolge $s_n(x)$ gleichmässig konvergiert.

Satz 7.1 (Majorantenkriterium). *Sei $(f_n)_{n\in\mathbb{N}}$ eine Folge von $\mathbb{R}^m$-wertigen Funktionen auf einem metrischen Raum M. Es existiere eine konvergente Reihe $\sum_n m_n$ mit $|f_n(x)| \le m_n$ für alle $n \in \mathbb{N}$ und alle $x \in M$. Dann*

- *konvergiert die Reihe $\sum_n f_n$ gleichmässig.*
- *konvergiert für jedex $x \in M$, die Reihe $\sum_n f_n(x)$ absolut.*

Beweis. Wir benutzen das Cauchy-Kriterium in Proposition 6.25. Sei

$$s_n(x) = \sum_{j=0}^{n} f_j(x)$$

und $m, n \in \mathbb{N}$ mit $m > n$. Dann

$$s_m(x) - s_n(x) = \sum_{j=n+1}^{m} f_j(x)$$

Da $\sum_n m_n$ konvergiert, finden wir, für jede $\varepsilon > 0$, ein $n_0 \in \mathbb{N}$ mit

$$\sum_{j=n+1}^{m} m_j < \varepsilon$$

für alle $m > n > n_0$. Also gilt für $m > n > n_0$

$$\|s_m(x) - s_n(x)\| \leq \sum_{j=n+1}^{m} \|f_j(x)\| \leq \sum_{j=n+1}^{m} m_j < \varepsilon \tag{11}$$

für alle $x \in M$. Aus Proposition 6.25 folgt, dass s_n gleichmässig konvergiert. Gleichung (11) impliziert auch, dass $\sum_n \|f_n(x)\|$ konvergiert, d.h., dass $\sum_n f_n(x)$ für alle $x \in M$ absolut konvergiert. □

7.1 Potenzreihen

Ein wichtiges Beispiel von Funktionenreihen sind Potenzreihen. Potenzreihen sind Funktionenreihen der Form $\sum_n a_n x^n$, mit Konstanten a_n. Allgemeiner, Potenzreihen mit Mittelpunkt x_0 sind Funktionenreihen der Form $\sum_n a_n(x - x_0)^n$. Der natürliche Bereich für Potenzreihen sind die komplexen Zahlen $\mathbb{C}$; die Konstanten a_n, sowie die Veränderliche x werden aus $\mathbb{C}$ gewählt. Potenzreihen werden also Funktionenreihen auf $\mathbb{C}$ sein, mit Werten in $\mathbb{C}$. Um zu erinnern, dass die Veränderliche komplex ist, werden wir z statt x schreiben.

Der Konvergenzbereich der Potenzreihe $\sum_n a_n z^n$ ist definiert als die Menge $\{z \in \mathbb{C} : \sum_{n=0}^{\infty} a_n z^n \text{ konvergiert}\}$.

Der Konvergenzradius der Potenzreihe $\sum_n a_n z^n$ ist definiert als

$$\rho := \frac{1}{\limsup_{k\to\infty} |a_k|^{1/k}} . \tag{12}$$

mit der Konvention, dass $1/\infty = 0$ und, dass $1/0 = \infty$.

Satz 7.2. *Sei der Konvergenzradius ρ der Potenzreihe $\sum_n a_n z^n$ wie in (12) definiert.*

a) Für $|z| < \rho$ konvergiert $\sum_n a_n z^n$ absolut.

b) Für $|z| > \rho$ ist die Folge $a_k z^k$ unbeschränkt. Die Reihe $\sum_n a_n z^n$ divergiert.

c) *Für $0 < r < \rho$ ist die Potenzreihe $\sum_n a_n z^n$ auf $\{z \in \mathbb{C} : |z| < r\}$ gleichmässig konvergent.*

d) *$\sum_{n=0}^{\infty} a_n z^n$ ist stetig auf $\{z \in \mathbb{C} : |z| < \rho\}$.*

Beweis. Sei

$$r < \rho = \frac{1}{\limsup_{k\to\infty} |a_k|^{1/k}}$$

Dann gilt

$$\limsup_{k\to\infty} (r^k |a_k|)^{1/k} < 1$$

Nach dem Wurzelkriterium ist die Reihe $\sum_k |a_k| r^k$ konvergent. Für $|z| \leq r$ gilt $|a_k z^k| \leq |a_k| r^k$. Nach dem Majorantenkriterium in Satz 7.1 ist die Potenzreihe $\sum_n a_n z^n$ auf $\{z \in \mathbb{C} : |z| \leq r\}$ gleichmässig konvergent. Aus Satz 7.1 folgt auch, dass für alle z mit $|z| \leq r$, $\sum_n a_n z^n$ absolut konvergiert. Das zeigt a) und c). Um b) zu zeigen, bemerken wir, dass, für $|z| > \rho$,

$$\limsup_{k\to\infty} |a_k z^k|^{1/k} > 1$$

Das impliziert, dass die Folge $a_k z^k$ unbeschränkt ist, und also, dass die Folge $\sum_n a_n z^n$ nicht konvergiert. Schlussendlich beweisen wir d). Sei $|z| < \rho$. Wir wählen $r > 0$ mit $|z| < r < \rho$. Da $s_n(z) = \sum_{j=0}^{n} a_j z^j$ stetig auf $\{z \in \mathbb{C} : |z| \leq r\}$ für alle $n \in \mathbb{N}$ (Polynome sind stetig) ist, und da s_n gleichmässig konvergiert, ist der Limes $\sum_{j=0}^{\infty} a_j z^j$ stetig auf $\{z \in \mathbb{C} : |z| \leq r\}$ (nach Proposition 6.24). $\square$

Besondere Fälle sind

- $\rho = 0$. In diesem Fall divergiert die Potenzreihe $\sum_n a_n z^n$ für alle $z \neq 0$.
- $\rho = \infty$. In diesem Fall konvergiert die Potenzreihe $\sum_n a_n z^n$ für alle $z \in \mathbb{C}$ (und die Konvergenz ist gleichmässig auf jeder beschränkten Teilmenge von $\mathbb{C}$).

Sei ρ der Konvergenzradius der Potenzreihe $\sum_n a_n z^n$. Wir setzen $D = \{z \in \mathbb{C} : |z| < \rho\}$. Der Abschluss von D ist dann $\overline{D} = \{z \in \mathbb{C} : |z| \leq \rho\}$. Es gilt

$$D \subset \text{ Konvergenzbereich von } \sum_n a_n z^n \ \subset \overline{D}$$

Das Verhalten der Potenzreihe $\sum_n a_n z^n$ auf der Kreislinie $\partial D = \{z \in \mathbb{C} : |z| = \rho\}$ hängt von der besonderen Potenzreihe ab.
Beispiele:

- $\sum_{k=0}^{\infty} z^k$ hat Konvergenzradius $\rho = 1$. Die Reihe konvergiert für alle $|z| < 1$. Für $|z| \geq 1$, ist $z^k \not\to 0$. Also divergiert die Reihe für alle $|z| \geq 1$.
- $\sum_{k=0}^{\infty} z^k/k$ hat Konvergenzradius $\rho = 1$. Die Reihe konvergiert für alle $|z| < 1$. Für $z = 1$ ist die Reihe die harmonische Reihe und divergiert. Für $z = -1$ ist die Reihe die alternierende harmonische Reihe und konvergiert (eigentlich konvergiert die Reihe für alle $|z| = 1$, ausser $z = 1$).
- $\sum_{k=0}^{\infty} z^k/k^2$ hat Konvergenzradius $\rho = 1$. Die Reihe konvergiert für alle $|z| \leq 1$.

Der Konvergenzradius kann auch durch das Quotientenkriterium berechnet werden.

Proposition 7.3. *Sei ρ der Konvergenzradius der Potenzreihe $\sum_n a_n z^n$, definiert in (12). Es gilt*

$$\rho = \lim_{k\to\infty} \frac{|a_k|}{|a_{k+1}|}$$

falls der Grenzwert existiert (oder falls der Limes ∞ ist).

Beweis. Wir nehmen an, der Grenzwert $\lim_{k\to\infty} |a_k|/|a_{k+1}|$ existiert. Sei $0 < |z| < \lim_{k\to\infty} |a_k|/|a_{k+1}|$. Dann gilt

$$1 < \lim_{k\to\infty} \frac{|a_k||z|^k}{|a_{k+1}||z|^{k+1}}$$

und

$$1 > \lim_{k\to\infty} \frac{|a_{k+1}||z|^{k+1}}{|a_k||z|^k} = \lim_{k\to\infty} \frac{|a_{k+1}z^{k+1}|}{|a_k z^k|}$$

Das Quotientenkriterium impliziert, dass $\sum_k a_k z^k$ konvergiert. Analog impliziert das Quotientenkriterium, dass die Reihe divergiert, falls $|z| > \lim_{k\to\infty} |a_k|/|a_{k+1}|$. Das zeigt, dass $\rho = \lim_{k\to\infty} |a_k|/|a_{k+1}|$. $\square$

7.2 Die Exponentialfunktion

Wir definieren die Exponentialfunktion $\exp : \mathbb{C} \to \mathbb{C}$ durch

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

für alle $z \in \mathbb{C}$. Der Konvergenzradius der Reihe ist aus Proposition 7.3 aus

$$\rho = \lim_{k\to\infty} \frac{(k+1)!}{k!} = \lim_{k\to\infty} (k+1) = \infty$$

gegeben. Das heisst, die Reihe konvergiert absolut für alle $z \in \mathbb{C}$, und die Konvergenz ist gleichmässig auf jeder beschränkten Teilmenge von $\mathbb{C}$. Aus 7.2 folgt auch, dass die Exponentialfunktion auf $\mathbb{C}$ stetig ist. Es gilt $\exp(0) = 1$. Eine wichtige Eigenschaft der Exponentialfunktion ist, dass

$$\exp(z_1 + z_2) = \exp(z_1)\exp(z_2) \quad \text{für alle } z_1, z_2 \in \mathbb{C} \tag{13}$$

In der Tat, aus Satz 4.22 über Cauchy-Produkte von Reihen, gilt

$$\begin{aligned}
\exp(z_1)\exp(z_2) &= \left(\sum_{i=0}^{\infty} \frac{z^i}{i!}\right)\left(\sum_{j=0}^{\infty} \frac{z^j}{j!}\right) \\
&= \sum_{n=0}^{\infty} \left(\sum_{m=0}^{n} \frac{z_1^m}{m!} \frac{z_2^{n-m}}{(n-m)!}\right) \\
&= \sum_{n=0}^{\infty} \frac{1}{n!} \left(\sum_{m=0}^{n} \binom{n}{m} z_1^m z_2^{n-m}\right) \\
&= \sum_{n=0}^{\infty} \frac{(z_1+z_2)^n}{n!}
\end{aligned}$$

Aus (13) folgt, dass

$$\exp(z)\exp(-z) = \exp(0) = 1 \;\Rightarrow\; \exp(-z) = \frac{1}{\exp(z)}$$

Insbesondere $\exp(z) \neq 0$, für alle $z \in \mathbb{C}$.

Sei nun $x \in \mathbb{R}$. Dann gilt offenbar $\exp(x) \in \mathbb{R}$. Für $x \geq 0$ ist

$$\exp(x) = 1 + x + \frac{x^2}{2!} + \dots$$

positiv und streng monoton wachsend (weil jeder Summand monoton wachsend ist und alle Summanden, ausser dem ersten streng monoton wachsend sind). Da offenbar $\exp(x) \geq 1 + x$ gilt $\exp(x) \to \infty$, für $x \to \infty$. Da $\exp(x) \geq x^{p+1}/(p+1)!$ gilt sogar

$$\frac{\exp(x)}{x^p} \to \infty$$

Für $x \to \infty$, für beliebige $p \in \mathbb{N}$. D.h. die Exponentialfunktion divergiert schneller als jede Potenz von x, im Limes $x \to \infty$. Wir betrachten nun $x < 0$.

$$\exp(x) = \frac{1}{\exp(-x)}$$

ist die Exponentialfunktion auch auf $(-\infty; 0)$ streng monoton wachsend. In der Tat, für $x_1 < x_2 < 0$ gilt $0 < -x_2 < -x_1$. Also $\exp(-x_2) < \exp(-x_1)$ und deswegen

$$\exp(x_1) = \frac{1}{\exp(-x_1)} < \frac{1}{\exp(-x_2)} = \exp(x_2)$$

Als $x \to -\infty$ gilt $-x \to \infty$ und also $\exp(-x) \to \infty$. Deswegen

$$\exp(x) = \frac{1}{\exp(-x)} \to 0$$

Es folgt, dass die $\exp : \mathbb{R} \to \mathbb{R}_+ = \{x \in \mathbb{R} : x > 0\}$ bijektiv ist (Injektivität folgt aus strenger Monotonie, Surjektivität aus Stetigkeit und aus dem Zwischenwertsatz).

Sei nun $x \in \mathbb{R}$. Wir behaupten, dass

$$(\exp(x))^\alpha = \exp(\alpha x) \tag{14}$$

für alle $\alpha \in \mathbb{Q}$ (linke Seite ist wohldefiniert, weil $\exp(x) > 0$ ist). In der Tat, für $p \in \mathbb{N}$ gilt

$$(\exp(x))^p = (\exp(x)) \cdot (\exp(x)) \dots (\exp(x)) = \exp(x + x + \cdots + x) = \exp(px)$$

Für $\alpha = p = -1, -2, \dots$ gilt weiter

$$(\exp(x))^p = \frac{1}{(\exp(x))^{-p}} = \frac{1}{\exp(-px)} = \exp(px)$$

Für $p \in \mathbb{Z}$ und $q \in \mathbb{N}\backslash\{0\}$ gilt dann

$$\left(\exp\left(\frac{p}{q}x\right)\right)^q = \exp\left(q\frac{p}{q}x\right) = \exp(px)$$

und deswegen

$$\exp\left(\frac{p}{q}x\right) = \exp(px)^{1/q}.$$

Da die rechte Seite von (14) sinnvoll für alle $\alpha \in \mathbb{C}$ ist, kann man sie benutzen, um die linke Seite für $\alpha \in \mathbb{C}\backslash\mathbb{Q}$ zu definieren. Für beliebige $x \in \mathbb{R}$ und $z \in \mathbb{C}$ setzen wir also

$$(\exp(x))^z := \exp(zx) \tag{15}$$

Für $\exp : \mathbb{R} \to (0;\infty)$ bijektiv, definiert die letze Gleichung a^z für alle $a \in (0;+\infty)$ und $z \in \mathbb{C}$. Die Funktion a^z hat dann die folgenden Eigenschaften:

$$(a \cdot b)^z = a^z \cdot b^z, \quad a^{z_1+z_2} = a^{z_1} \cdot a^{z_2}, \quad a^{-z} = 1/a^z$$

Wie schon in Definition 4.9, bezeichnen wir mit

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

die Euler'sche Zahl. Gemäss der Definition (15) gilt dann $\exp(z) = e^z$, für alle $z \in \mathbb{C}$.

Proposition 7.4. *Die Euler'sche Zahl e ist irrational.*

Beweis. Nehmen wir an $e \in \mathbb{Q}$. Dann existieren $p \in \mathbb{N}$ und $q \in \mathbb{N}\backslash\{0\}$ mit $e = p/q$. Es gilt

$$q!e = q!(p/q) = p(q-1)! \in \mathbb{N}$$

Anderseits

$$q!e = q! + \frac{q!}{2!} + \dots + \frac{q!}{q!} + \frac{1}{q+1} + \frac{1}{(q+1)(q+2)} + \dots$$

Die Summe der ersten $(q+1)$ Termen in offenbar in $\mathbb{Z}$. Damit $q!e$ auch in $\mathbb{Z}$ sein kann,muss auch die Summe

$$\frac{1}{q+1} + \frac{1}{(q+1)(q+2)} + \dots \in \mathbb{Z}$$

sein. Das ist aber unmöglich, weil $q \geq 1$ und deswegen

$$\begin{aligned} 0 &< \frac{1}{q+1} + \frac{1}{(q+1)(q+2)} + \dots < \frac{1}{q+1} + \frac{1}{(q+1)^2} + \frac{1}{(q+1)^3} + \dots \\ &\leq \frac{1}{2} + \frac{1}{2^2} + \frac{1}{2^3} + \dots = \sum_{j=1}^{\infty} \frac{1}{2^j} = 1. \end{aligned}$$

$\square$

7.3 Der Logarithmus

Die Funktion $\exp : \mathbb{R} \to (0;\infty)$ ist streng monoton wachsend und bijektiv. Wir definieren die Logarithmusfunktion, also die Inverse dieser Funktion. D.h. $\log : (0;\infty) \to \mathbb{R}$ ist durch

$$\log x = y \quad \text{falls} \quad \exp(y) = x$$

definiert. Da exp streng monoton wachsend ist, folgt aus Proposition 6.17, dass log streng monoton wachsend und stetig ist. Es gilt

- $\exp(0) = 1 \Rightarrow \log(1) = 0$.
- $\exp(1) = e \Rightarrow \log(e) = 1$.
- $\exp x \to \infty$, als $x \to \infty$ impliziert, dass $\log x \to \infty$, als $x \to \infty$.
- Für $x \to \infty$ $\log x$ gegen ∞ langsamer konvergiert, als jede Potenz von x, gilt für jede $\alpha > 0$
$$\lim_{x\to\infty} \frac{\log x}{x^\alpha} = 0$$
In der Tat, sei $y = \log x$. Dann ist $x = e^y$ und $x^\alpha = (e^y)^\alpha = e^{\alpha y}$. Also
$$\lim_{x\to\infty} \frac{\log x}{x^\alpha} = \lim_{y\to\infty} \frac{y}{e^{\alpha y}} = 0$$
- $\exp x \to 0$, als $x \to -\infty$ impliziert, dass $\log x \to -\infty$, als $x \to 0$.
- Für $x \to 0$ $\log x \to -\infty$ langsamer konvergiert, als jede Potenz von x konvergiert gegen 0, gilt für jede $\alpha > 0$
$$\lim_{x\to 0} x^\alpha \log x = 0$$
In der Tat, sei $y = \log x$. Dann ist $x = e^y$ und $x^\alpha = e^{\alpha y}$. Das impliziert
$$\lim_{x\to 0} x^\alpha \log x = \lim_{y\to -\infty} y e^{\alpha y} = 0$$
- $\log(x \cdot y) = \log x + \log y$, für alle $x, y > 0$. In der Tat
$$\log(xy) = \log\left(e^{\log x} e^{\log y}\right) = \log\left(e^{\log x + \log y}\right) = \log x + \log y$$
- $(\exp(x))^z = \exp(zx)$ impliziert, dass für jedes $a > 0$,
$$a^z = (\exp(\log a))^z = \exp(z \log a)$$
Diese Formel definiert a^z, für jedes $z \in \mathbb{C}$ und $a > 0$.

7.4 Hyperbolische Funktionen

Für $z \in \mathbb{C}$ definieren wir die Funktionen
$$\cosh(z) = \frac{e^z + e^{-z}}{2}, \quad \sinh(z) = \frac{e^z - e^{-z}}{2}, \quad \tanh(z) = \frac{\sinh(z)}{\cosh(z)}$$
Aus Definition der Exponentialfunktion gilt
$$\cosh(z) = \sum_{j=0}^{\infty} \frac{z^{2j}}{(2j)!}, \quad \sinh(z) = \sum_{j=0}^{\infty} \frac{z^{2j+1}}{(2j+1)!}$$
Eigenschaften der hyperbolischen Funktionen:

- $\cosh(-z) = \cosh z$ und $\sinh(-z) = -\sinh z$ für alle $z \in \mathbb{C}$.

- Sei nun $x \in \mathbb{R}$. Für $x \gg 1$ gilt $\cosh x \simeq e^x/2$ und $\sinh x \simeq e^x/2$. Deswegen

$$\lim_{x\to\infty} \cosh x = \lim_{x\to\infty} \sinh x = \infty \qquad \text{und} \qquad \lim_{x\to\infty} \tanh x = 1$$

  Für $x \to -\infty$ gilt $\cosh x \simeq e^{-x}/2$ und $\sinh x \simeq -e^{-x}/2$. Wir erhalten

$$\lim_{x\to-\infty} \cosh x = \infty, \qquad \lim_{x\to-\infty} \sinh x = -\infty, \quad \text{und} \quad \lim_{x\to-\infty} \tanh x = -1$$

- Für jedes $z \in \mathbb{C}$ gilt

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

  In der Tat

$$\begin{aligned}\cosh^2 z - \sinh^2 z &= \left(\frac{e^z + e^{-z}}{2}\right)^2 - \left(\frac{e^z - e^{-z}}{2}\right)^2 \\ &= \frac{1}{4}\left(e^{2z} + e^{-2z} + 2 - e^{2z} - e^{-2z} + 2\right) = 1\end{aligned}$$

- Für $x \in \mathbb{R}$ sind die Funktionen e^x und $-e^{-x}$ streng monoton wachsend. Deswegen ist $\sinh(x)$ streng monoton wachsend.

- Da für $\sinh x \to -\infty$ $x \to -\infty$ und für $\sinh x \to \infty$ $x \to \infty$, ist $\sinh : \mathbb{R} \to \mathbb{R}$ bijektiv. Die Inverse Funktion wird mit $\operatorname{arcsinh} : \mathbb{R} \to \mathbb{R}$ bezeichnet. Es gilt

$$\operatorname{arcsinh}(x) = \log\left(x + \sqrt{x^2+1}\right)$$

  Aus $\cosh x = (1 + \sinh^2 x)^{1/2}$ folgt, dass $\cosh x$ monoton wachsend auf $[0;\infty)$ ist. Da $\cosh 0 = 1$ folgt, dass $\cosh : [0,\infty) \to [1;\infty)$ bijektiv ist. Die Inverse Funktion wird mit $\operatorname{arccosh} : [1;\infty) \to [0;\infty)$ bezeichnet. Es gilt

$$\operatorname{arccosh}(x) = \log\left(x + \sqrt{x^2-1}\right)$$

7.5 Trigonometrische Funktionen

Für $z \in \mathbb{C}$ definieren wir

$$\cos z := \frac{e^{iz} + e^{-iz}}{2} \quad \text{und} \quad \sin z := \frac{e^{iz} - e^{-iz}}{2i}$$

Aus Definition der Exponentialfunktion folgt, dass

$$\cos z = \sum_{j=0}^{\infty} (-1)^j \frac{z^{2j}}{(2j)!}, \quad \sin z = \sum_{j=0}^{\infty} (-1)^j \frac{z^{2j+1}}{(2j+1)!}$$

Also $\cos(-z) = \cos z$ und $\sin(-z) = -\sin z$ für alle $z \in \mathbb{C}$.

Für $t \in \mathbb{R}$ gilt $\overline{\exp(it)} = \exp(-it)$ (weil $\overline{\exp(z)} = \exp(\overline{z})$ für alle $z \in \mathbb{C}$). Deswegen

$$\cos t = \frac{e^{it} + e^{-it}}{2} = \operatorname{Re} e^{it}, \quad \sin t = \frac{e^{it} - e^{-it}}{2i} = \operatorname{Im} e^{it} \quad \Rightarrow \quad e^{it} = \cos t + i \sin t$$

Aus

$$|e^{it}|^2 = e^{it}e^{-it} = e^0 = 1$$

folgt, dass $|e^{it}| = 1$ für alle $t \in \mathbb{R}$. Es folgt, dass $\cos^2 t + \sin^2 t = 1$ für alle $t \in \mathbb{R}$.

Für alle $x, y \in \mathbb{R}$ gelten die Additionsformeln

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

Sie folgen aus der Bemerkung, dass einerseits

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

und anderseits

$$\begin{aligned} e^{i(x+y)} &= e^{ix}e^{iy} = (\cos x + i\sin x)(\cos y + i\sin y) \\ &= (\cos x \cos y - \sin x \sin y) + i(\cos x \sin y + \sin x \cos y) \end{aligned}$$

Satz 7.5. *Es existiert eine Zahl $\pi \in [2; 4]$ mit*

$$\cos(t + 2\pi) = \cos t, \quad \sin(t + 2\pi) = \sin t, \quad \textit{für alle } t \in \mathbb{R}$$

Beweis. Sei $t \in (0; 1]$. Die Termen in der Reihe für cos und sin sind alternierend und ihr Absolutbetrag ist monoton fallend. Also gilt, aus dem Beweis von Prop. 4.15

$$1 - \frac{t^2}{2} < \cos t < 1 - \frac{t^2}{2} + \frac{t^4}{4!} \tag{16}$$

und

$$t - \frac{t^3}{3!} < \sin t < t$$

für alle $t \in (0; 1]$. Insbesondere $\cos t, \sin t > 0$ für alle $t \in (0; 1]$. Aus (16) folgt, mit $t = 1/2$, dass

$$\cos(1/2) > 1 - \frac{1}{8} = \frac{7}{8}$$

und dass $\cos(1) < 1 - 1/2 + 1/24 = 13/24$. Da $\cos t$ stetig ist, folgt aus dem Zwischenwertsatz, dass $t_0 \in [1/2, 1]$ mit $\cos t_0 = 1/\sqrt{2}$ existiert (weil $13/24 < 1/\sqrt{2} < 7/8$). Aus $\sin^2(t_0) + \cos^2(t_0) = 1$ und $\sin t_0 > 0$ folgt, dass

$$\sin t_0 = \sqrt{1 - \cos^2 t_0} = 1/\sqrt{2}$$

Also $\cos t_0 = \sin t_0 = 1/\sqrt{2}$. Wir setzen $\pi = 4t_0$. Dann gilt

$$e^{i\pi/4} = \cos t_0 + i \sin t_0 = \frac{1}{\sqrt{2}}(1 + i)$$

Das gibt

$$e^{i\pi/2} = (e^{i\pi/4})^2 = \frac{1}{2}(1+i)^2 = i \;\Rightarrow\; e^{i\pi} = -1 \;\Rightarrow\; e^{2i\pi} = 1$$

Das bedeutet, dass

$$e^{i(t+2\pi)} = e^{it}e^{2i\pi} = e^{it}$$

für alle $t \in \mathbb{R}$, und deswegen

$$\cos(t + 2\pi) = \cos t, \quad \sin(t + 2\pi) = \sin t$$

für alle $t \in \mathbb{R}$. □

Aus dem Beweis folgt auch, dass $e^{i\pi} = -1$ und $e^{i\pi/2} = i$. Das gibt

$$e^{i(\pi+t)} = -e^{it} = -\cos t - i\sin t$$

und

$$e^{i(\pi/2-t)} = e^{i\pi/2}e^{-it} = ie^{-it} = i(\cos t - i\sin t) = \sin t + i\cos t$$

Deswegen gilt

$$\cos(t+\pi) = -\cos t, \quad \text{und} \quad \sin(t+\pi) = -\sin t$$

und auch

$$\cos(\pi/2 - t) = \sin t \quad \text{und} \quad \sin(\pi/2 - t) = \cos t$$

Proposition 7.6. *Die Funtkion $t \to e^{it} = \cos t + i\sin t$ bildet das Interval $[0; 2\pi)$ auf der Kreislinie $\{z = x + iy : x^2 + y^2 = 1\}$ bijektiv ab.*

Beweis. Schritt 1: $\cos t, \sin t > 0$ für $0 < t < \pi/2$.

Wir haben die Behauptung für $0 < t \leq 1$ schon bewiesen. Da wir aus Definition $\pi/4 < 1$ wissen, dass $\cos t, \sin t > 0$ für $t \in (0; \pi/4]$. Nun, für $t \in (\pi/4; \pi/2)$ ist $\pi/2 - t \in (0; \pi/4)$. Deswegen gilt, für $t \in (\pi/4; \pi/2)$,

$$\cos t = \sin(\pi/2 - t) > 0, \quad \text{und} \quad \sin t = \cos(\pi/2 - t) > 0\,.$$

Schritt 2: $\sin t$ ist streng monoton wachsend auf $[0; \pi/2]$.

Um die Behauptung zu zeigen, bemerken wir, dass wir für beliebige $s, t \in \mathbb{R}$

$$\begin{aligned}
\sin s &= \sin\left(\frac{s+t}{2} + \frac{s-t}{2}\right) = \sin\left(\frac{s+t}{2}\right)\cos\left(\frac{s-t}{2}\right) + \cos\left(\frac{s+t}{2}\right)\sin\left(\frac{s-t}{2}\right) \\
\sin t &= \sin\left(\frac{s+t}{2} - \frac{s-t}{2}\right) = \sin\left(\frac{s+t}{2}\right)\cos\left(\frac{s-t}{2}\right) - \cos\left(\frac{s+t}{2}\right)\sin\left(\frac{s-t}{2}\right)
\end{aligned}$$

haben. Durch Subtraktion folgt, dass

$$\sin s - \sin t = 2\cos\left(\frac{s+t}{2}\right)\sin\left(\frac{s-t}{2}\right)$$

Nun gilt für $0 \leq t < s \leq \pi/2$ $0 < (s+t)/2 < \pi/2$ und $0 < (s-t)/2 < \pi/2$. Aus der letzen Gleichung und aus Schritt 1 folgt, dass $\sin s - \sin t > 0$.

Schritt 3: Die Abbildung $t \to e^{it} = \cos t + i\sin t$ bildet $[0; \pi/2]$ bijektiv auf $\{x + iy : x^2 + y^2 = 1, \text{ mit } x, y \geq 0\}$ ab.

Da $\sin t$ streng monoton wachsend ist, ist die Injektivität der Abbildung klar. Um die Surjektivität zu zeigen, benutzen wir, dass $\sin(0) = 0$, $\sin(\pi/2) = 1$. Sei $u + iv \in \mathbb{C}$ mit $u^2 + v^2 = 1$ und $u, v \geq 0$ gegeben. Da $v \in [0; 1]$ und sin stetig ist, folgt aus dem Zwischenwertsatz, dass $t_0 \in [0; \pi/2]$ mit $\sin t_0 = v$ existiert. Dann ist (da $\cos t_0 > 0$ aus Schritt 1)

$$\cos t_0 = \sqrt{1 - \sin^2 t_0} = \sqrt{1 - v^2} = u\,.$$

D.h. $u + iv = \cos t_0 + i\sin t_0$.

Schritt 4: die Abbildung $t \to e^{it} = \cos t + i \sin t$ bildet $[0; 2\pi)$ bijektiv auf $\{x + iy : x^2 + y^2 = 1\}$.

Aus Schritt 3 und aus $\cos(-t) = \cos t$ und $\sin(-t) = -\sin t$ folgt, dass $t \to e^{it}$ das Interval $[-\pi/2; 0]$ bijektiv auf $\{x + iy : x^2 + y^2 = 1, x \geq 0, y \leq 0\}$ abbildet. Diese Bemerkung impliziert, dass die Abbildung $t \to e^{it}$ einerseits das Interval $[\pi/2; \pi]$ auf $\{x + iy : x^2 + y^2 = 1, x \leq 0, y \geq 0\}$ (weil $\cos(t + \pi) = -\cos t$ und $\sin(t + \pi) = -\sin t$), und anderseits das Interval $[3\pi/2; 2\pi]$ auf $\{x + iy : x^2 + y^2 = 1, x \geq 0, y \leq 0\}$ (weil $\cos(t + 2\pi) = \cos t$ und $\sin(t + 2\pi) = \sin t$) bijektiv abbildet. Aus Schritt 3 folgt auch, dass $t \to e^{it}$ $[\pi; 3\pi/2]$ bijektiv auf $\{x + iy : x^2 + y^2 = 1, \text{ mit } x, y \leq 0\}$ abbildet (weil $\cos(t + \pi) = -\cos t$ und $\sin(t + \pi) = -\sin t$). $\square$

Definition 7.7. *Eine Funktion $f : \mathbb{R} \to \mathbb{C}$ heisst periodisch, falls ein $p \neq 0$ existiert mit*

$$f(t + p) = f(t), \quad \text{für alle } t \in \mathbb{R}$$

heisst p dann eine Periode von f. Die Fundamentalperiode von f ist die kleinste positive Periode. Ist die Funktion f periodisch und stetig, dann existiert die Fundamentalperiode. Ein Fundamentalbereich ist ein Intervall der Form $[x; x + p)$, wobei p die Fundamentalperiode ist.

Die Funktionen $t \to e^{it}$, $t \to \cos t$, $t \to \sin t$ sind periodisch, weil

$$e^{i(t+2\pi)} = e^{it}, \quad \cos(t + 2\pi) = \cos t, \quad \sin(t + 2\pi) = \sin t$$

für alle $t \in \mathbb{R}$. Aus Proposition 7.6 folgt, dass 2π die Fundamentalperiode von $t \to e^{it}$ ist (aus Injektivität kann kein $0 < T < 2\pi$ mit $e^{iT} = 1$) existieren. Da

$$e^{it} = \cos t + i \sin t = \cos t + i \cos(\pi/2 - t)$$

ist jede Periode von $\cos t$ auch eine Periode von e^{it}. Das zeigt, dass 2π die Fundamentalperiode von $\cos t$ ist. Ähnlich ist 2π die Fundamentalperiode von $\sin t$.

Aus der Injektivität von $t \to e^{it}$ auf $[0; 2\pi)$ folgt auch, dass

$$\begin{aligned}
e^{it} = 1 &\iff t = 2\pi n, \text{ für ein } n \in \mathbb{Z} \\
e^{it} = -1 &\iff t = \pi(2n + 1), \text{ für ein } n \in \mathbb{Z} \\
e^{it} = i &\iff t = \pi/2 + 2\pi n, \text{ für ein } n \in \mathbb{Z} \\
e^{it} = -i &\iff t = 3\pi/2 + 2\pi n, \text{ für ein } n \in \mathbb{Z}
\end{aligned}$$

Das impliziert, dass

$$\begin{aligned}
\sin t = 0 &\iff t = \pi n \text{ für ein } n \in \mathbb{Z} \\
\cos t = 0 &\iff t = \pi \left(n + \frac{1}{2} \right) \text{ für ein } n \in \mathbb{Z}
\end{aligned}$$

Da $\sin t$ streng monoton wachsend auf $[0; \pi/2]$ ist und $\sin(-t) = -\sin t$ folgt, dass $\sin t$ streng monoton wachsend auf $[-\pi/2; \pi/2]$ ist, mit $\sin(-\pi/2) = -1$ und $\sin(\pi/2) =$

1. Es folgt, dass $\sin : [-\pi/2; \pi/2] \to [-1; 1]$ bijektiv. Wir bezeichnen die Inverse Funktion mit $\arcsin : [-1; 1] \to [-\pi/2; \pi/2]$. Da sin stetig und streng monoton wachsend ist, ist, wegen Proposition 6.17, arcsin auch stetig und monoton wachsend. Da $\cos t = \sin(\pi/2 - t) = -\sin(t - \pi/2)$ folgt, dass $\cos : [0; \pi] \to [-1; 1]$ streng monoton fallend und bijektiv ist. Die inverse Funktion bezeichnen wir mit $\arccos : [-1; 1] \to [0; \pi]$. Aus den Eigenschaften von cos folgt, dass arccos stetig und streng monoton fallend ist.

Für $t \in \mathbb{R}$, $t \neq (n + 1/2)\pi$, wir definieren

$$\tan t := \frac{\sin t}{\cos t}$$

Auf $[0; \pi/2)$ gilt $\sin t, \cos t \geq 0$, $\sin t$ ist streng monoton wachsend, $\cos t$ ist streng monoton fallend. Es folgt, dass $\tan t \geq 0$ und streng monoton wachsend auf $[0; \pi/2]$ ist. $\tan 0 = 0$ und

$$\lim_{t \uparrow \pi/2} \tan t = +\infty$$

implizieren, dass $t \to \tan t$ $[0; \pi/2)$ bijektiv auf $[0; \infty)$ abbildet. Aus $\sin(-t) = -\sin t$ und $\cos(-t) = \cos t$ folgt, dass $\tan(-t) = \tan t$. Deswegen bildet $t \to \tan t$ $(-\pi/2; 0]$ bijektiv auf $(-\infty; 0]$ ab. Es folgt, dass $t \to \tan t$ monoton wachsend auf $(-\pi/2; \pi/2)$ ist und $(-\pi/2; \pi/2)$ bijektiv auf $\mathbb{R}$ abbildet. Aus der Stetigkeit von sin und cos folgt auch, dass $t \to \tan t$ stetig auf $(-\pi/2; \pi/2)$ ist. Die inverse Abbildung $x \to \arctan(x)$ bildet $\mathbb{R}$ bijektiv und stetig auf $(-\pi/2; \pi/2)$ ab.

Für $z \in \mathbb{C}$ mit $|z| = 1$ existiert, wegen Proposition 7.6, genau ein $\varphi \in [0; 2\pi)$ mit $z = e^{i\varphi}$. Für beliebige $z \in \mathbb{C}$ mit $z \neq 0$ existiert $\varphi \in [0; 2\pi)$ mit

$$\frac{z}{|z|} = e^{i\varphi} \;\Rightarrow\; z = |z|e^{i\varphi}$$

Man nennt $z = re^{i\varphi}$ die Polardarstellung der komplexen Zahl z. $r = |z| \geq 0$ ist dabei eindeutig bestimmt. φ ist dagegen nur modulo 2π eindeutig, weil $e^{i(\varphi + 2n\pi)} = e^{i\varphi}$ für alle $n \in \mathbb{Z}$. Jedes φ mit $z = re^{i\varphi}$ heisst ein Argument von z. Für jedes $z \neq 0$, und beliebige $\alpha \in \mathbb{R}$ existiert genau ein Argument von z in $[\alpha; \alpha + 2\pi)$ und genau ein Argument von z in $(\alpha; \alpha + 2\pi]$. Insbesondere existiert genau ein Argument $\varphi \in (-\pi; \pi]$; man nennt das eindeutige Argument von z in $(-\pi; \pi]$ den Hauptwert des Arguments $\arg z$. Sei $z = x + iy$, mit $x > 0$. Ist $z = re^{i\varphi} = r(\cos\varphi + i \sin\varphi)$ so muss $r = (x^2 + y^2)^{1/2}$ und

$$x = r\cos\varphi, \quad y = r\sin\varphi \;\Rightarrow\; \tan\varphi = y/x \;\Rightarrow\; \varphi = \arctan(y/x)$$

Aus der Definition von arctan ist $\varphi \in (-\pi/2; \pi/2)$ der Hauptwert des Arguments von z. Eine kurze Überlegung zeigt, dass

$$\arg(x + iy) = \begin{cases} \arctan(y/x), & \text{falls } x > 0 \\ \arctan(y/x) + \pi, & \text{falls } x < 0, y \geq 0 \\ \arctan(y/x) - \pi, & \text{falls } x < 0, y < 0 \\ \pi/2, & \text{falls } x = 0, y > 0 \\ -\pi/2, & \text{falls } x = 0, y < 0 \end{cases}$$

Die Funktion $\arg z$ ist auf $\mathbb{C} \backslash (-\infty, 0]$ stetig.

Die Polardarstellung $z = re^{i\varphi}$ ist besonders nützlich, um Potenzen und Wurzeln von komplexen Zahlen zu berechnen. Für ein beliebiges $p \in \mathbb{N}\backslash\{0\}$ und $z = re^{i\varphi}$ ist

$$w_n = r^{1/p} e^{i(\varphi+2\pi n)/p}$$

für $n \in \mathbb{Z}$ eine p-Wurzel von z, weil offenbar $w_n^p = z$. Sind $n, m \in \mathbb{Z}$ mit $n - m = jp$, für ein $j \in \mathbb{Z}$, dann ist $e^{2i\pi n/p} = e^{2i\pi m/p}$, und also $w_n = w_m$. Das bedeutet, dass es für jedes$z \in \mathbb{C}\backslash\{0\}$ genau p verschiedene komplexe Zahlen $w_0, w_1, \dots w_{p-1}$ mit $w_j^p = z$ gibt. Die Berechnung der Wurzeln $w_0, \dots, w_p$ wird in Polarkoordinaten auf die Berechnung der p-ten Wurzel $r^{1/p}$ der positiven reellen Zahl $r > 0$ zurückgeführt.

8 Differentialrechnung

8.1 Grundbegriffe

Wir betrachten in dieser Sektion Funktionen $f : A \to \mathbb{R}^m$, definiert in einer Teilmenge $A \subset \mathbb{R}$, mit Werten in $\mathbb{R}^m$. Für diese Funktionen definieren wir den Begriff von Differenzierbarkeit und von Ableitung.

Definition 8.1. *Sei f eine $\mathbb{R}^m$-wertige Funktion, definiert in einer Umgebung von $x_0 \in \mathbb{R}$. Wir sagen, f ist an der Stelle x_0 differenzierbar, falls der Limes*

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

existiert. In diesem Fall heisst $f'(x_0)$ die Ableitung von f an der Stelle x_0. Für $f'(x_0)$ wird manchmal auch die Bezeichnung $df/dx(x_0)$ benutzt. Sei $G \subset \mathbb{R}$ eine offene Menge. Eine Funktion $f : G \to \mathbb{R}^m$ heisst differenzierbar auf G, falls f differenzierbar an der Stelle x_0 ist, für alle $x_0 \in G$.

Bemerkung: Differenzierbar ist für Funktionen $f : G \to \mathbb{R}^m$ definiert, unter der Annahme, dass G eine offene Menge ist. Diese Annahme garantiert, dass für jedes $x_0 \in G$, die Funktion f in einer Umgebung von x_0 definiert ist (damit ist der Limes sinnvoll). Um Stetigkeit von $f : A \to M$ zu definieren, braucht man dagegen nicht, dass A offen ist.

Da $(f(x_0 + h) - f(x_0))/h$ die mittlere Zuwachsrate von f im Intervall $[x_0; x_0 + h]$ (oder im Intervall $[x_0 + h; x_0]$, falls $h < 0$) ist, kann man die Ableitung $f'(x_0)$ als die momentane Zuwachsrate bezeichnen. Ist f reelwertig, dann ist $f'(x_0)$ die Steigung der Tangente zum Graph von f im Punkt $(x_0; f(x_0))$. Nimmt f Werte in $\mathbb{R}^m$, so ist $f'(x_0)$ (falls es existiert) auch ein Vektor in $\mathbb{R}^m$; geometrisch beschreibt $x \to f(x)$ eine Kurve in $\mathbb{R}^m$, und $f'(x_0)$ ist ein Tangentenvektor zu dieser Kurve, im Punkt $f(x_0)$ (die Länge von $f'(x_0)$, d.h. $\|f'(x_0)\|_{\mathbb{R}^m}$ gibt die "Geschwindigkeit" der Parametrisierung der Kurve.

Proposition 8.2. *Sei f eine $\mathbb{R}^m$-wertige Funktion, definiert in einer Umgebung von $x_0 \in \mathbb{R}$. Ist f differenzierbar an der Stelle x_0, so ist f auch stetig an der Stelle x_0.*

Beweis. Es gilt

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

Für $h \to 0$ konvergiert $(f(x_0+h) - f(x_0))/h \to f'(x_0)$. Also

$$\lim_{h\to 0} f(x_0+h) - f(x_0) = 0 \;\Rightarrow\; \lim_{h\to 0} f(x_0+h) = f(x_0)$$

Das impliziert die Stetigkeit von f an der Stelle x_0. $\square$

Die Umkehrung von Proposition 8.2 gilt nicht; Stetigkeit von f an der Stelle x_0 impliziert nicht, dass f an der Stelle x_0 differenzierbar ist. Zum Beispiel, die Funktion $f : \mathbb{R} \to \mathbb{R}$, die durch $f(x) = |x|$ definiert ist, ist stetig auf $\mathbb{R}$ aber, da

$$\lim_{h\to 0} \frac{|h|}{h}$$

nicht existiert, ist sie nicht differenzierbar an der Stelle $x = 0$. Tatsache: Es existieren überall stetige Funktionen, die nirgends differenzierbar sind.

Bemerkung: Sei $f : G \to \mathbb{R}^m$ eine $\mathbb{R}^m$-wertige Funktion, definiert auf $G \subset \mathbb{R}$. Dann existieren Funktionen $f_j : G \to \mathbb{R}$, für $j = 1,\dots,m$, mit $f(x) = (f_1(x), f_2(x), \dots, f_m(x))$. Die Funktion f ist an der Stelle $x_0 \in G$ differenzierbar g.d.w. f_j an der Stelle x_0, für alle $j = 1,\dots,m$ differenzierbar ist. Das folgt natürlich aus der Bemerkung, dass eine Folge $a_n = (a_n^{(1)}, a_n^{(2)}, \dots, a^(m)_n)$ auf $\mathbb{R}^m$ genau dann konvergiert, wenn jede Komponente $a_n^{(j)}$ konvergiert. In diesem Sinn kann die Frage, ob eine $\mathbb{R}^m$-wertige Funktion differenzierbar ist, immer auf die Differenzierbarkeit von $\mathbb{R}$-wertigen Funktionen zurückgeführt werden.

Satz 8.3. *Seien f, g reelwertige Funktionen definiert in einer Umgebung von $x_0 \in \mathbb{R}$. Seien f, g differenzierbar an der Stelle x_0. Dann*

a) ist $f + g$ differenzierbar an der Stelle x_0 und $(f+g)'(x_0) = f'(x_0) + g'(x_0)$.

b) ist $f \cdot g$ differenzierbar an der Stelle x_0 und $(fg)'(x_0) = f'(x_0)g(x_0) + f(x_0)g'(x_0)$.

c) ist $g(x_0) \neq 0$, so ist f/g differenzierbar an der Stelle x_0, und

$$\left(\frac{f}{g}\right)'(x_0) = \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{g^2(x_0)}$$

Beweis. a) folgt aus Linearität des Limes. Um b) zu zeigen, bemerken wir, dass

$$\begin{aligned}\frac{(f\cdot g)(x_0+h) - (f\cdot g)(x_0)}{h} &= \frac{f(x_0+h)g(x_0+h) - f(x_0)g(x_0+h)}{h} \\ &\quad + \frac{f(x_0)g(x_0+h) - f(x_0)g(x_0)}{h} \\ &= \frac{f(x_0+h) - f(x_0)}{h} g(x_0+h) + f(x_0)\frac{g(x_0+h) - f(x_0)}{h} \\ &\to f'(x_0)g(x_0) + f(x_0)g'(x_0)\end{aligned}$$

wo wir benutzt haben, dass f und g differenzierbar an der Stelle x_0 sind (und deswegen ist g stetig an der Stelle x_0, d.h. $g(x_0+h) \to g(x_0)$, für $h \to 0$). Um c) zu zeigen, genügt es, den Fall $f = 1$ zu betrachten. Der allgemeine Fall folgt dann aus b), indem wir f/g als $f \cdot (1/g)$ schreiben. Für $f = 1$ finden wir

$$\begin{aligned} \frac{1}{h}\left(\frac{1}{g(x_0+h)} - \frac{1}{g(x_0)}\right) &= \frac{g(x_0) - g(x_0+h)}{hg(x_0)g(x_0+h)} \\ &= -\frac{g(x_0+h) - g(x_0)}{h} \frac{1}{g(x_0)g(x_0+h)} \\ &\to -\frac{g'(x_0)}{g^2(x_0)} \end{aligned}$$

weil g differenzierbar ist, und also, insbesondere, stetig. $\square$

Wir berechnen nun die Ableitung von elementaren $\mathbb{R}$-wertigen Funktionen.

- Die Ableitung der konstanten Funktion $f(x) = c$ ist $f'(x) = 0$.
- Sei $f(x) = x$. Dann

  $$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} = \lim_{h \to 0} \frac{x + h - x}{h} = 1$$

  für alle $x \in \mathbb{R}$.
- Sei $f(x) = x^n$, für $n \in \mathbb{N}$. Wir behaupten, dass $f'(x) = nx^{n-1}$ (insbesondere ist f überall auf $\mathbb{R}$ differenzierbar). Das zeigen wir durch Induktion. Die Behauptung gilt für $n = 1$. Nehmen wir an, die Behauptung gilt für $n \leq m$, für ein $m \in \mathbb{N}$. Wir zeigen sie für $n = m + 1$:

  $$\frac{d}{dx} x^{m+1} = \frac{d}{dx} x \cdot x^m = 1 \cdot x^m + x \frac{d}{dx} x^m = x^m + mx^m = (m+1)x^m$$

  wo wir die Induktionsannahme benutzt haben. Es folgt, mit Satz 8.3, dass Polynome immer auf $\mathbb{R}$ überall differenzierbar sind, und dass rationale Funktionen überall dort differenzierbar sind, wo der Nenner nicht verschwindet.
- Sei $f(x) = x^{-n} = 1/x^n$, für ein $n \in \mathbb{N}$. Dann ist f differenzierbar an der Stelle x, für alle $x \in \mathbb{R}$, mit $x \neq 0$. Für $x \neq 0$ es gilt weiter $f'(x) = -n/x^{n+1} = -nx^{-n-1}$. Das heisst die Formel $(d/dx)x^n = nx^{n-1}$ gilt eigentlich für alle $n \in \mathbb{Z}$. Das folgt aus Satz 8.3 (Teil c)), weil

  $$\frac{d}{dx} x^{-n} = \frac{d}{dx} \frac{1}{x^n} = \frac{-nx^{n-1}}{x^{2n}} = -\frac{n}{x^{n+1}}$$

- Sei $f(x) = \exp(x)$ die Exponentialfunktion, definiert auf $\mathbb{R}$. Wir behaupten, dass f überall differenzierbar ist, und dass $f'(x) = \exp(x)$. In der Tat

  $$\lim_{h \to 0} \frac{e^{x+h} - e^x}{h} = \lim_{h \to 0} e^x \frac{e^h - 1}{h} = e^x \tag{17}$$

  weil

  $$\frac{e^h - 1}{h} = \sum_{n=1}^{\infty} \frac{h^{n-1}}{n!} = \sum_{n=0}^{\infty} \frac{h^n}{(n+1)!}$$

Die Potenzreihe $\sum h^n/(n+1)!$ konvergiert auf $|h| \leq 1$ gleichmässig und ist deswegen stetig als gleichmässige Limes stetiger Funktionen. Das bedeutet

$$\lim_{h \to 0} \frac{e^h - 1}{h} = \lim_{h \to 0} \sum_{n=0}^{\infty} \frac{h^n}{(n+1)!} = 1 \tag{18}$$

und zeigt (17).

- Sei $f(x) = \log x$, für $x > 0$. Wir behaupten, f ist auf $(0; \infty)$ differenzierbar, und $f'(x) = 1/x$. Beweis: Wir bemerken, dass

  $$\lim_{x \to 1} \frac{\log x}{x - 1} = 1 \tag{19}$$

  In der Tat, mit $z = \log x$ ist $x = e^z$. Da $\log x$ stetig an der Stelle 1 ist, finden wir

  $$\lim_{x \to 1} \frac{\log x}{x-1} = \lim_{z \to 0} \frac{z}{e^z - 1} = \frac{1}{\lim_{z \to 0} \frac{e^z - 1}{z}} = 1$$

  Aus (19) erhalten wir, für $x > 0$,

  $$\lim_{h \to 0} \frac{\log(x+h) - \log x}{h} = \lim_{h \to 0} \frac{\log(1 + h/x)}{h} = \frac{1}{x} \lim_{h \to 0} \frac{\log(1 + h/x)}{h/x} = \frac{1}{x}$$

- Die Funtkionen sin und cos sind auf $\mathbb{R}$ differenzierbar. Weiter $(\sin)'(x) = \cos(x)$ und $(\cos)'(x) = -\sin x$. Wir haben

  $$\sin x = \frac{e^{ix} - e^{-ix}}{2i}, \quad \text{und} \quad \cos x = \frac{e^{ix} + e^{-ix}}{2}$$

  Wir bemerken, dass, ähnlich wie in (18),

  $$\lim_{h \to 0} \frac{e^{ih} - 1}{ih} = 1 \quad \text{und} \quad \lim_{h \to 0} \frac{e^{-ih} - 1}{ih} = -1$$

  Das gibt

  $$\begin{aligned} \lim_{h \to 0} \frac{\sin(x+h) - \sin x}{h} &= \lim_{h \to 0} \frac{e^{i(x+h)} - e^{-i(x+h)} - e^{ix} + e^{-ix}}{2ih} \\ &= \lim_{h \to 0} \frac{e^{i(x+h)} - e^{ix}}{2ih} - \lim_{h \to 0} \frac{e^{-i(x+h)} - e^{-ix}}{2ih} \\ &= \frac{e^{ix}}{2} \lim_{h \to 0} \frac{e^{ih} - 1}{ih} - \frac{e^{-ix}}{2} \lim_{h \to 0} \frac{e^{-ih} - 1}{ih} \\ &= \frac{e^{ix} + e^{-ix}}{2} = \cos x \end{aligned}$$

  Analog kann man zeigen, dass $(\cos)'(x) = -\sin x$. Satz 8.3 impliziert dann, dass $\tan x = \sin x / \cos x$ an der Stelle x differnzierbar, für alle $x \neq (n + 1/2)\pi$, für ein $n \in \mathbb{Z}$.

Die folgende Proposition wird nützlich sein.

Proposition 8.4. *Sei f eine $\mathbb{R}^m$-wertige Funktion, definiert auf einer Umgebung U von $x_0 \in \mathbb{R}$. Dann ist f differenzierbar an der Stelle $x_0 \in \mathbb{R}$ g.d.w. eine $\mathbb{R}^m$-wertige Funktion k existiert, definiert auf U und stetig an der Stelle x_0, mit*

$$f(x) = f(x_0) + k(x)(x - x_0)$$

für alle $x \in U$. Dann gilt $k(x_0) = f'(x_0)$.

Beweis. Für $x \neq x_0$ muss

$$k(x) = \frac{f(x) - f(x_0)}{x - x_0} \tag{20}$$

Die Behauptung ist äquivalent zur folgenden Aussage: f ist differenzierbar an der Stelle x_0 g.d.w. die Funktion k, definiert auf $U \backslash \{x_0\}$ durch (20), eine stetige Fortsetzung auf $x = x_0$ besitzt. Das folgt aber aus Proposition 6.11, weil k eine stetige Fortsetzung in $x = x_0$ g.d.w. der Limes

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} \tag{21}$$

existiert, hat. In diesem Fall gibt (21) den Wert der stetigen Fortsetzung an der Stelle x_0; also $k(x_0) = f'(x_0)$. □

Die Kettenregel erlaubt uns die Ableitung von verknüpften Funktionen zu berechnen.

Proposition 8.5 (Kettenregel)**.** *Sei f eine reelwertige Funktion, definiert in einer Umgebung von $x_0 \in \mathbb{R}$. Sei g eine $\mathbb{R}^m$-wertige Funktion, definiert in einer Umgebung von $y_0 = f(x_0)$. Sei f differenzierbar an der Stelle x_0 und g differenzierbar an der Stelle y_0. Dann ist $g \circ f$ differenzierbar an der Stelle x_0 und*

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

Beweis. Aus Proposition 8.4 folgt, dass

$$g(y) - g(y_0) = k(y)(y - y_0)$$

für eine Funktion k (definiert in einer Umgebung von $y_0 = f(x_0)$) mit $k(y_0) = g'(y_0)$. Also

$$\frac{g(f(x_0 + h)) - g(f(x_0))}{h} = k(f(x_0 + h)) \, \frac{f(x_0 + h) - f(x_0)}{h} \to k(f(x_0)) \, f'(x_0)$$

für $h \to 0$ (weil k stetig an der Stelle $f(x_0)$ ist und f differenzierbar an der Stelle x_0, und also insbesondere stetig ist). □

Beispiele: Sei $f(x) = x^\alpha = \exp(\alpha \log x)$ definiert für $x \in (0; \infty)$ und für ein $\alpha \in \mathbb{R}$. Dann gilt

$$\frac{d}{dx} \, x^\alpha = \exp(\alpha \log x) \, \frac{\alpha}{x} = \alpha x^{\alpha - 1}$$

Ein anderes Beispiel: $f(x) = x^x = \exp(x \log x)$. Also

$$\frac{d}{dx} \, x^x = \exp(x \log x) \, (\log x + 1) = x^x (\log x + 1)$$

Proposition 8.6 (Differenzierbarkeit der Umkehrfunktion)**.** *Sei $I \subset \mathbb{R}$ ein Intervall. $f : I \to \mathbb{R}$ stetig und streng monoton. Sei $x_0 \in I^\circ$ (nicht ein Endpunkt des Intervals) und f an der Stelle x_0 differenzierbar, mit $f'(x_0) \neq 0$. Dann ist f^{-1} definiert in einer Umgebung von $f(x_0)$ und an der Stelle $f(x_0)$ differenzierbar mit*

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

Beweis. Sei $g = f^{-1}$. Da f differenzierbar an der Stelle x_0 ist, finden wir k, stetig an der Stelle x_0, mit $k(x_0) = f'(x_0)$ und

$$f(x) - f(x_0) = k(x)(x - x_0)$$

in einer Umgebung von x_0. Wir setzen $y_0 = f(x_0)$. Dann ist $x_0 = g(y_0)$. Wir setzen auch $x = g(y_0 + h)$. Es gilt

$$y_0 + h = f(x) = f(x_0) + k(x)(x - x_0) = y_0 + k(g(y_0 + h))(g(y_0 + h) - g(y_0))$$

Das gibt

$$\frac{g(y_0 + h) - g(y_0)}{h} = \frac{1}{k(g(y_0 + h))}$$

Für $h \to 0$, nach Stetigkeit von g and der Stelle y_0, gilt $g(y_0 + h) \to g(y_0)$. Da k stetig ist, gilt auch $k(g(y_0 + h)) \to k(g(y_0))$. Also

$$\lim_{h \to 0} \frac{g(y_0 + h) - g(y_0)}{h} = \frac{1}{k(g(y_0))} = \frac{1}{f'(x_0)}$$

Insbesondere der Limes existiert. □

Beispiel: Sei $\exp : \mathbb{R} \to (0; \infty)$ streng monoton wachsend und differenzierbar. Die Inverse Funktion ist $\log : (0; \infty) \to \mathbb{R}$. Aus Proposition 8.6 folgt, dass log auf $(0; \infty)$ differenzierbar ist, und

$$(\log)'(\exp x) = \frac{1}{\exp x} \;\Rightarrow\; (\log)'(y) = \frac{1}{y}$$

für alle $y \in \mathbb{R}$, wie wir schon bewiesen hatten.

Beispiel: Sei $f(x) = \tan x$. $f : (-\pi/2; \pi/2) \to \mathbb{R}$ streng monoton wachsend und differenzierbar, mit $f'(x) = \tan^2 x + 1$. Die Inverse Funktion ist $\arctan : \mathbb{R} \to (-\pi/2; \pi/2)$. Für beliebige $x \in \mathbb{R}$ ist arctan an der Stelle x differenzierbar, mit

$$(\arctan)'(\tan x) = \frac{1}{\tan^2 x + 1} \;\Rightarrow\; (\arctan)'(y) = \frac{1}{y^2 + 1}$$

für alle $y \in \mathbb{R}$.

8.2 Ableitungen und Optimierung

Sei f reelwertig, definiert auf $A \subset \mathbb{R}$. Wir sagen $x_0 \in A$ ist an der Stelle x_0 lokal maximal (bzw. minimal), falls eine Umgebung U von x_0 mit $f(x) \leq f(x_0)$ (bzw. mit $f(x) \geq f(x_0)$) für alle $x \in U \cap A$ existiert. Wir sagen f ist lokal extremal an der Stelle x_0 falls sie entweder lokal maximal oder lokal minimal an der Stelle x_0 ist.

Proposition 8.7. *Sei f eine reelwertige Funktion, definiert in einer Umgebung von $x_0 \in \mathbb{R}$. Sei f an der Stelle x_0 differenzierbar, mit $f'(x_0) \neq 0$. Dann ist f an der Stelle x_0 nicht extremal.*

Beweis. Nehmen wir z.B. an, dass $f'(x_0) > 0$. Dann existiert $\delta > 0$ s.d.

$$\frac{f(x_0+h) - f(x_0)}{h} > 0 \quad \text{für alle } h \in \mathbb{R} \text{ mit } 0 < |h| < \delta.$$

Dann gilt $f(x_0+h) - f(x_0) > 0$, d.h. $f(x_0+h) > f(x_0)$, falls $0 < h < \delta$, und $f(x_0 + h) - f(x_0) < 0$, d.h. $f(x_0+h) < f(x_0)$, falls $-\delta < h < 0$. Also x_0 kann weder ein lokales Maximum noch ein lokales Minimum sein. Analog kann man der Fall $f'(x_0) < 0$ betrachten. □

Folgerung: Nehmen wir an, wir suchen das (globale) Maximum einer stetigen Funktion $f : [a;b] \to \mathbb{R}$. Das Maximum existiert aus dem Satz des Maximums. Das Maximum ist sicher auch ein lokales Maximum. Alle Punkte $x \in (a;b)$ mit $f'(x) \neq 0$ können deswegen nicht das globale Maximum sein. Die einzigen bleibenden Kandidaten sind: a, b, alle Punkte in $(a;b)$, wo $f'(x) = 0$, und alle Punkte in $(a;b)$, wo f nicht differenzierbar ist.

Beispiel: Welches Rechteck mit Umfang p hat den grössten Flächeninhalt? Seien x, y die zwei Kantenlängen. Es muss gelten: $0 \leq x, y \leq p/2$ und $2(x+y) = p$. Das gibt $y = p/2 - x$. Der Flächeninhalt des Rechtecks ist dann $A = xy = x(p/2 - x)$. Um den maximalen Flächeninhalt zu finden, müssen wir das Maximum der Funktion $A(x) = x(p/2 - x)$, definiert auf dem Intervall $[0; p/2]$, finden. A ist überall in $(0; p/2)$ differenzierbar, mit $A'(x) = p/2 - 2x$. $A'(x) = 0$ genau dann, wenn $x = p/4$. Mögliche Maxima sind $A(0) = 0$, $A(p/2) = 0$ und $A(p/4) = p^2/16 > 0$. Wir schliessen, dass das Maximum $p^2/16$ ist, was einem Quadrat entspricht.

8.3 Mittelwertsatz

Satz 8.8 (Satz von Rolle)**.** *Sei $f : [a;b] \to \mathbb{R}$ stetig, mit $f(a) = f(b) = 0$. Sei f differenzierbar auf $(a;b)$. Dann existiert mindestens ein $\xi \in (a;b)$ mit $f'(\xi) = 0$.*

Beweis. Ist $f(x) = 0$ für alle $x \in [a;b]$, so ist die Behauptung offenbar erfüllt. Wir können also annehmen, es existiert mindestens ein $x \in (a;b)$ mit $f(x) > 0$ oder mit $f(x) < 0$. Nehmen wir zum Beispiel, an es gibt $x \in (a;b)$ mit $f(x) > 0$. Nach dem Satz vom Maximum existiert $\xi \in [a;b]$ mit

$$f(\xi) = \sup\{f(x) : x \in [a;b]\}$$

Da $f(a) = f(b) = 0$ ist $a < \xi < b$. Da f differenzierbar an der Stelle ξ ist, muss $f'(\xi) = 0$. □

Bemerkung: Satz 8.8 gilt allgemeiner, falls $f(a) = f(b)$ (man braucht nicht, dass $f(a) = f(b) = 0$).

Satz 8.9 (Mittelwertsatz). *Sei $f : [a; b] \to \mathbb{R}$ stetig, differenzier auf $(a; b)$. Dann existiert $\xi \in (a; b)$ mit*

$$f'(\xi) = \frac{f(b) - f(a)}{b - a}$$

Beweis. Wir definieren

$$h(x) = (b - a)(f(x) - f(a)) - (f(b) - f(a))(x - a)$$

Dann gilt: h ist stetig auf $[a; b]$ und differenzierbar auf $(a; b)$, $h(a) = h(b) = 0$. Aus Satz 8.8 folgt, dass ein $\xi \in (a; b)$ existiert mit

$$0 = h'(\xi) = (b - a)f'(\xi) - (f(b) - f(a)) \Rightarrow f'(\xi) = \frac{f(b) - f(a)}{b - a}$$

□

Bemerkung: Allgemeiner gilt $f, g : [a; b] \to \mathbb{R}$ stetig und auf $(a; b)$ differenzierbar, mit $g'(x) \neq 0$ für alle $x \in (a; b)$. Dann existiert $\xi \in (a; b)$ mit

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(\xi)}{g'(\xi)}$$

Beweis: Übung.

Bemerkung: Mittelwertsatz ist i.A. falsch für Funktionen mit Werten auf $\mathbb{R}^m$, $m > 1$, oder auf $\mathbb{C}$. Sei zum Beispiel $f(x) = \exp(ix)$. Es gilt $f(0) = f(2\pi) = 1$. Also $(f(2\pi) - f(0))/(2\pi - 0) = 0$. Andererseits gilt $f'(x) = i \exp(ix)$. Deswegen existiert kein $\xi \in [0; 2\pi]$ mit $f'(\xi) = 0$.

Proposition 8.10. *Sei $f : [a; b] \to \mathbb{R}^m$ stetig, differenzierbar auf $(a; b)$. Es gelte $\|f'(x)\| \leq M$ für alle $x \in (a; b)$. Dann gilt $\|f(b) - f(a)\| \leq M|b - a|$.*

Beweis. O.B.d.A können wir annehmen, dass $f(b) \neq f(a)$. Wir setzen dann $c = f(b) - f(a)$, und

$$h(x) = c \cdot (f(x) - f(a)) = \sum_{j=1}^{m} c_j (f_j(x) - f_j(a))$$

h ist dann eine reelwertige stetige Funktion, differenzierbar auf $(a; b)$, mit $h(a) = 0$, $h(b) = \|f(b) - f(a)\|^2$, und $h'(x) = c \cdot f'(x) = \sum_{j=1}^{m} c_j f_j'(x)$. Aus Cauchy-Schwarz, folgt

$$|h'(x)| \leq \|c\| \|f'(x)\| \leq M\|c\|$$

für alle $x \in (a; b)$. Aus dem Mittelwertsatz für h, existiert ein $\xi \in (a; b)$ mit

$$\frac{h(b) - h(a)}{b - a} = h'(\xi)$$

Also

$$\|f(b) - f(a)\|^2 = |h(b) - h(a)| \leq |h'(\xi)||b - a| \leq M|b - a|\|f(b) - f(a)\|$$

Wir dividieren links und rechts durch $\|f(b) - f(a)\|$ und bekommen $\|f(b) - f(a)\| \leq M|b - a|$. □

Korollar 8.11. *Sei $f : [a;b] \to \mathbb{R}^m$ stetig, mit $f'(x) = 0$ für alle $x \in (a;b)$. Dann ist f konstant auf $[a;b]$.*

Beweis. Sei $a < x < b$. Aus der letze Proposition, angewandt auf $[a;x]$ folgt $\|f(x) - f(a)\| = 0$, und also $f(x) = f(a)$. $\square$

Proposition 8.12. *Sei $f : [a;b] \to \mathbb{R}$ stetig, differenzierbar auf $(a;b)$. Dann ist f monoton wachsend (bzw. fallend) auf $[a;b]$ g.d.w. $f'(x) \geq 0$ (bzw. $f'(x) \leq 0$) für alle $x \in (a;b)$.*

Beweis. Ist f wachsend, dann gilt $f(x_0 + h) - f(x_0) \geq 0$ für alle $x_0 \in (a;b)$ und $h > 0$ klein genug. Das impliziert, dass $(f(x_0 + h) - f(x_0))/h \geq 0$ für alle $h > 0$ klein genug, und also (weil wir angenommen haben, dass der Limes existiert, dass

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

Anderseits, falls $f'(x) \geq 0$ für alle $x \in (a;b)$, so gilt wegen dem Mittelwertsatz auf dem Intervall $[x;y]$, mit $a \leq x \leq y \leq b$,

$$f(y) - f(x) = f'(\xi)(y - x) \geq (y - x) \geq 0$$

Das zeigt, dass f monoton wachsend ist. Analog kann man die Behauptung für monoton fallende Funktionen anwenden. $\square$

Das Beispiel $f(x) = x^3$ zeigt, dass f streng monoton wachsend sein kann, obwohl ein x existiert (in diesem Beispiel $x = 0$) mit $f'(x) = 0$. Ist f monoton wachsend aber nicht streng monoton wachsend auf $[a;b]$, dann gibt es $a \leq c < d \leq b$ mit $f(c) = f(d)$. Wegen der Monotonie muss dann $f(x) = f(c)$ für alle $x \in [c;d]$. Also $f'(x) = 0$ für alle $x \in (c;d)$. D.h. monoton wachsende Funktionen die nicht streng monoton wachsend sind, sind auf einem Intervall konstant; das impliziert, dass die Ableitung auf einem Intervall positiver Länge verschwinden muss.

Proposition 8.13. *Sei $f : [a;b] \to \mathbb{R}$ stetig, differenzierbar auf $(a;b)$. Dann ist f streng monoton wachsend auf $[a;b]$ g.d.w. $f'(x) \geq 0$ für alle $x \in (a;b)$ und es gibt kein Intervall positiver Länge, wo f' verschwindet.*

Satz 8.14 (Satz von Bernoulli-de L' Hôpital)**.** *Seien f, g differenzierbar auf $(a;b)$ ($a = -\infty$ ist hier zugelassen). Sei $g(x) \neq 0$ und $g'(x) \neq 0$ für alle $x \in (a;b)$. Es gelte*

$$\lim_{x \downarrow a} f(x) = \lim_{x \downarrow a} g(x) = 0$$

und es existiere der Limes

$$\lim_{x \downarrow a} \frac{f'(x)}{g'(x)} = \lambda \in \mathbb{R} \cup \{\pm\infty\}$$

Dann gilt

$$\lim_{x \downarrow a} \frac{f(x)}{g(x)} = \lambda$$

Bemerkungen: Es gibt eine analoge Version des Satzes von Bernoulli-L'Hôpital für den linkseitigen Limes. Die Aussage gilt auch, falls $\lim_{x\downarrow a} f(x) = \lim_{x\downarrow a} g(x) = +\infty$ oder $\lim_{x\downarrow a} f(x) = \lim_{x\downarrow a} g(x) = -\infty$ (Beweis: Übung). Der Satz gilt i.A. nicht für komplexwertige Funktionen.

Beweis. Wir behandeln nur den Fall $\lambda \in \mathbb{R}$. Sei $\varepsilon > 0$. Dann existiert $x_0 \in (a; b)$ mit

$$\left| \frac{f'(x)}{g'(x)} - \lambda \right| < \varepsilon \quad \text{für alle } a < x < x_0$$

Nach der ersten Bemerkung nach Satz 8.9 gilt, für beliebige $a < x_1 < x_0$:

$$\frac{f(x) - f(x_1)}{g(x) - g(x_1)} = \frac{f'(\xi)}{g'(\xi)}$$

für ein $\xi \in (x_1; x_0)$. Also

$$\left| \frac{f(x) - f(x_1)}{g(x) - g(x_1)} - \lambda \right| < \varepsilon$$

Für alle $a < x_1 < x < x_0$. Wir lassen nun $x_1 \to a$ streben; dann $f(x_1) \to 0$ und $g(x_1) \to 0$. Es folgt, dass

$$\left| \frac{f(x)}{g(x)} - \lambda \right| < \varepsilon$$

für alle $a < x < x_0$. $\square$

Anwendung: Wir berechnen

$$\lim_{x\to 0} \frac{1}{\sin x} - \frac{1}{x} = \lim_{x\to 0} \frac{x - \sin x}{x \sin x} = \lim_{x\to 0} \frac{1 - \cos x}{\sin x + x \cos x} = \lim_{x\to 0} \frac{\sin x}{2\cos x - x \sin x} = 0$$

wobei wir den Satz von Bernoulli-L' Hôpital zwei Mal angewandt haben.

8.4 Höhere Ableitungen

Wir bezeichnen mit f'', mit $f^{(2)}$ oder mit $d^2 f/dx^2$ die zweite Ableitung von f, d.h., die Ableitung von f'. Rekursiv definieren wir $f^{(n)} = (f^{(n-1)})'$. Statt $f^{(n)}$ benutzen wir manchmal auch die Notation $d^n f/dx^n$. Genauer: $f^{(n)}$ ist an der Stelle $x_0 \in \mathbb{R}$ definiert, wenn $f^{(n-1)}(x)$ in einer Umgebung von x_0 definiert ist, und wenn die Funktion $x \to f^{(n-1)}(x)$ an der Stelle x_0 differenzierbar ist. In diesem Fall sagen wir, dass f an der Stelle x_0, n-mal differenzierbar ist. Ist f an der Stelle x n-mal differenzierbar, für alle x in einer offenen Menge $G \subset \mathbb{R}$, dann sagen wir f ist auf G n-mal differenzierbar. Ist weiter $x \to f^{(n)}(x)$ stetig auf G, dann sagen wir f ist auf G n-mal stetig differenzierbar. Für eine offene Menge $G \subset \mathbb{R}$, definieren wir

$$C^n(G) := \{f : f \text{definiert, reelwertig, } n\text{-mal stetig differenzierbar auf } G\}$$

mit der Konvention, dass $C^0(G) = C(G)$ die Menge der stetigen Funktionen auf G ist (wir bezeichnen dagegen mit $C_b(G)$ die Menge der stetigen und beschränkten Funktionen auf G). Wir bezeichnen mit $C^n(G; \mathbb{C})$ die Menge der komplex-wertigen n-mal stetigen, differenzierbaren Funktionen auf G.

Da die Existenz von $f^{(m)}$ die Stetigkeit von $f^{(m-1)}$ impliziert, und deswegen auch die Existenz von $f^{(j)}$, für alle $0 \leq j \leq m$, haben wir die Inklusionen

$$C^0(G) \supset C^1(G) \supset \dots$$

Wir setzen

$$C^\infty(G) = \bigcap_{n=0}^{\infty} C^n(G)$$

Ist $f \in C^\infty(G)$, dann sagen wir, dass f auf G beliebig oft differenzierbar ist.

Aus Definition ist klar: $f \in C^m(G)$ g.d.w. $f \in C^1(G)$ und $f' \in C^{m-1}(G)$.

Proposition 8.15. *Seien $f, g \in C^m(G)$. Dann gilt $f + g, fg \in C^{(m-1}(G)$.*

Beweis. Die Tatsache $f + g \in C^m$ ist klar. Wir zeigen nun, dass

$$f, g \in C^m \;\Rightarrow\; fg \in C^m$$

durch Induktion über m. Für $m = 1$ folgt die Behauptung aus $(fg)' = f'g + fg'$. Wir nehmen nun an, die Behauptung gilt für $m = n-1$, und wir zeigen sie für $m = n$. Dazu bemerken wir, dass

$$f, g \in C^n \;\Rightarrow\; f, g \in C^1 \;\Rightarrow\; fg \in C^1$$

mit $(fg)' = f'g + fg'$. Da $f' \in C^{m-1}$ und $g \in C^m \subset C^{m-1}$ folgt aus der Induktionsannahme, dass $f'g \in C^{m-1}$. Analog ist $fg' \in C^{m-1}$. Wir erhalten, dass $(fg)' \in C^{m-1}$. Das zeigt, dass $fg \in C^m$. □

Proposition 8.16. *Sei $G \subset \mathbb{R}$ offen, $f \in C^m(G)$, $f(G) \subset H \subset \mathbb{R}$ offen, und $g \in C^m(H)$. Dann ist $g \circ f \in C^m(G)$.*

Beweis. Wir benutzen Induktion. Für $m = 1$ folgt die Behauptung aus der Kettenregel, mit $(g \circ f)'(x) = g'(f(x))f'(x)$. Nehmen wir nun an, die Behauptung gilt für $m = n-1$; wir zeigen sie für $m = n$. wir bemerken, dass

$$f, g \in C^m \;\Rightarrow\; f, g \in C^1 \;\Rightarrow\; g \circ f \in C^1, \text{ mit } (g \circ f)' = (g' \circ f)f'$$

Aus $g \in C^m$ folgt, dass $g' \in C^{m-1}$. Da auch $f \in C^m \subset C^{m-1}$, gilt aus Induktionsannahme $g' \circ f \in C^{m-1}$. Aus der Proposition 8.15 folgt auch, dass $(g' \circ f)f' \in C^{m-1}$. □

Proposition 8.17. *Sei I ein offenes Intervall, $f \in C^m(I)$, $f'(x) \neq 0$ für alle $x \in I$ (d.h. f ist streng monoton). Dann ist $f^{-1} \in C^m(f(I))$.*

Beweis. Wir benutzen wieder Induktion nach m. Für $m = 1$ folgt die Behauptung aus Proposition 8.6. Wir nehmen nun an, die Bahauptung gilt für $m = n-1$ und wir zeigen sie für $m = n$. Ist $f \in C^m$, dann ist insbesondere $f \in C^1$ und also $f^{-1} \in C^1$. Da $(f^{-1})' = j \circ f' \circ f^{-1}$, wobei $j(t) = 1/t$ beliebig oft differenzierbar ist, $f' \in C^{m-1}$ und $f^{-1} \in C^{m-1}$ (aus Induktionsannahme), folgt aus Prop. 8.16, dass $(f^{-1})' \in C^{m-1}$, also, dass $f^{-1} \in C^m$. □

Wir haben bis jetzt den Raum $C^m(G)$ für $G \subset \mathbb{R}$ offen definiert. Auf $C^m(G)$ kann man die Struktur eines Vektorraumes einführen (Funktionen werden punktweise summiert und mit skalaren $\lambda \in \mathbb{R}$ multipliziert). Es gibt aber auf $C^m(G)$ keine natürliche Norm (die Supremum-Norm muss nicht endlich sein; z.B. $1/x \in C^\infty((0;1))$ ist nicht beschränkt). Es ist deswegen nützlich, weitere Räume zu definieren. Für $-\infty < a \leq b < \infty$ setzen wir

$$C^m([a;b]) := \Big\{ f \in C^m((a;b)) : f^{(j)} \text{ kann auf } [a;b] \text{ stetig fortgesetzt werden,}$$
$$\text{für alle } j = 1, \dots, m \Big\}$$

Die stetige Funktion $\phi : (a;b) \to \mathbb{R}$ ist auf $[a;b]$ stetig ergänzbar g.d.w. $\lim_{x \downarrow a} \phi(x)$ und $\lim_{x \uparrow b} \phi(x)$ existieren (und endlich sind). Da $[a;b] \subset \mathbb{R}$ kompakt ist, ist $f^{(j)} \in C^{m-j}([a;b]) \subset C^0([a;b])$ sicher beschränkt. Auf $C^m([a;b])$ kann man nun, ähnlich wie wir auf $C_b(M)$ gemacht haben, Normen einführen. Für $f \in C^m([a;b])$ setzen wir

$$\|f\|_{C^m([a;b])} := \sum_{j=0}^{m} \|f^{(j)}\|_{C^0([a;b])} = \sum_{j=0}^{m} \sup_{x \in [a;b]} |f^{(j)}(x)|$$

$C^m([a;b])$, versehen mit der Norm $\|.\|_{C^m([a;b])}$, ist ein normierter Raum. Die Norm auf $C^m([a;b])$ induziert die Metrik definiert aus

$$d(f,g) = \|f-g\|_{C^m([a;b])}$$

für alle $f, g \in C^m([a;b])$. $C^m([a;b])$, versehen mit dieser Metrik ist dann ein vollständiger metrischer Raum (Beweis: Übung).

8.5 Konvexität

Definition 8.18. *Eine reelwertige Funktion f auf einem Intervall I heisst konvex, falls*

$$f(\alpha x + (1-\alpha) y) \leq \alpha f(x) + (1-\alpha) f(y)$$

für alle $x, y \in I$, und falls $0 \leq \alpha \leq 1$. f heisst streng konvex, falls

$$f(\alpha x + (1-\alpha) y) < \alpha f(x) + (1-\alpha) f(y)$$

für alle $x, y \in I$, mit $x \neq y$ und alle $0 < \alpha < 1$. Die Funktion f heisst konkav, bzw. streng konkav, falls $-f$ konvex, bzw. streng konvex ist.

Geometrisch bedeutet Konvexität, dass das Geradenstück zwischen zwei beliebigen Punkten $(x, f(x))$ und $(y, f(y))$ auf dem Graphen von f oberhalb des Graphen von f selbst liegt.

Für differenzierbare und zwei-mal differenzierbare Funktionen gibt es einfache Kriteria für Konvexität.

Lemma 8.19. *Seif konvex auf I. Dann ist*

$$m(x,y) = \frac{f(y) - f(x)}{y - x}$$

(die Steigung der Sekante) monoton wachsend in x für y fest, und monoton wachsend in y für x fest, für alle $x < y$.

Beweis. Sei $x < z < y$. Dann gilt

$$m(x,z) \leq m(x,y) \leq m(z,y)$$

Um die erste Ungleichung zu zeigen, schreiben wir $z = \alpha x + (1-\alpha)y$, für ein $\alpha \in (0;1)$. Aus Konvexität gilt

$$f(z) \leq \alpha f(x) + (1-\alpha) f(y)$$

Also

$$m(x,z) = \frac{f(z)-f(x)}{z-x} \leq \frac{(\alpha-1)f(x) + (1-\alpha)f(y)}{(\alpha-1)x + (1-\alpha)y} = \frac{f(y)-f(x)}{y-x} = m(x,y)$$

Die zweite Ungleichung folgt analog. □

Proposition 8.20. *Sei f konvex auf dem Intervall I, und $x \in I^\circ$. Dann existieren*

$$f'(x^+) := \lim_{h \downarrow 0} \frac{f(x+h)-f(x)}{h}$$
$$f'(x^-) := \lim_{h \uparrow 0} \frac{f(x+h)-f(x)}{h}$$

f ist an der Stelle x stetig und $f'(x^-) \leq f'(x^+)$. Weiter, für $x, y \in I^\circ$ mit $x < y$, gilt

$$f'(x^+) \leq \frac{f(y)-f(x)}{y-x} \leq f'(y^-)$$

Beweis. Nach Lemma 8.19 ist $(f(x+h)-f(x))/h$ monoton wachsend in h für x fest, für $h > 0$. Da

$$\frac{f(x)-f(y)}{x-y} = m(y,x) \leq m(y, x+h) \leq m(x, x+h) = \frac{f(x+h)-f(x)}{h}$$

für jedes $y < x$ ist der Bruch $(f(x+h)-f(x))/h$ nach unten beschränkt. Es folgt, dass der Limes

$$\lim_{h \downarrow 0} \frac{f(x+h)-f(x)}{h}$$

existiert. Analog existiert der linkseitige Limes. Die anderen Aussagen folgen direkt aus Lemma 8.19. □

Proposition 8.21. *Sei f differenzierbar auf dem offenen Intervall I. Dann ist f konvex auf I g.d.w. f' monoton wachsend auf I ist. Ist ferner f zweimal differenzierbar auf I, dann ist f konvex g.d.w. $f'' \geq 0$.*

Beweis. Sei f konvex. Aus Proposition 8.20 gilt $f'(x^+) \leq f'(y^-)$ für alle $x < y$. Da f differenzierbar ist, gilt $f'(x) = f'(x^+) \leq f'(y^-) = f'(y)$. Das zeigt, dass f' monoton wachsend ist. Sei nun f' monoton wachsend. Wir zeigen, dass f konvex ist. Es genügt zu zeigen, dass

$$f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha) f(y)$$

für alle $x < y$ und $\alpha \in (0;1)$. Sei also $z = \alpha x + (1-\alpha)y$. Dann gilt $\alpha = (y-z)/(y-x)$ und $1-\alpha = (z-x)/(y-x)$. Also

$$\begin{aligned}\alpha f(x) + (1-\alpha)f(y) - f(z) &= \frac{(y-z)f(x) + (z-x)f(y)}{y-x} - \frac{(y-z)f(z) + (z-x)f(z)}{y-x} \\ &= \frac{(y-z)(f(x)-f(z)) + (z-x)(f(y)-f(z))}{y-x}\end{aligned} \tag{22}$$

Aus dem Mittelwertsatz existieren $\xi_1 \in (x;z)$ und $\xi_2 \in (z;y)$ mit

$$f(x) - f(z) = f'(\xi_1)(x-z) \quad \text{und} \quad f(y) - f(z) = f'(\xi_2)(y-z)$$

Aus (22) folgt, dass

$$\alpha f(x) + (1-\alpha)f(y) - f(z) = \frac{(z-x)(y-z)}{y-x}\big(f'(\xi_2) - f'(\xi_1)\big) \geq 0$$

weil f' monoton wachsend ist. Falls f zweimal differenzierbar ist, so ist f' monoton wachsend g.d.w. $f'' \geq 0$. □

Definition 8.22. *Sei f definiert in einer Umgebung von $x_0 \in \mathbb{R}$ und konvex. Eine Stützfunktion von f an der Stelle x_0 ist eine lineare Funktion $\phi(x) = \phi(x_0) + m(x-x_0)$ mit $\phi(x_0) = f(x_0)$ und $\phi(x) \leq f(x)$ für alle x in dem Definitionsbereich von f.*

Bemerkung: Ist f zweimal differenzierbar, so ist f streng konvex auf I g.d.w. $f''(x) \geq 0$ auf I und es gibt kein Intervall positiver Länge, wo f'' verschwindet.

Proposition 8.23. *Sei f konvex auf einem Intervall I, $x_0 \in I^\circ$ und $\overline{m} \in \mathbb{R}$ mit $f'(x_0^-) \leq \overline{m} \leq f'(x_0^+)$. Dann ist*

$$\phi(x) = f(x_0) + \overline{m}(x-x_0)$$

eine Stützfunktion von f an der Stelle x_0.

Bemerkung: Ist f differenzierbar an der Stelle x_0, dann ist die Tangente mit Steigung $f'(x_0)$ die einzige Stützfunktion.

Beweis. Für $x > x_0$ gilt, nach Proposition 8.20,

$$f'(x_0^+) \leq \frac{f(x)-f(x_0)}{x-x_0} \;\Rightarrow\; f(x) \geq f(x_0) + f'(x_0^+)(x-x_0) \geq f(x_0) + \overline{m}(x-x_0)$$

Für $x < x_0$ gilt ähnlich , dass

$$f'(x_0^-) \geq \frac{f(x_0)-f(x)}{x_0-x} \;\Rightarrow\; f(x) \geq f(x_0) + f'(x_0^-)(x-x_0) \geq f(x_0) + \overline{m}(x-x_0)$$

weil $x - x_0 < 0$. □

Proposition 8.24 (Jensen'sche Ungleichung)**.** *Sei f konvex auf dem Intervall I und seien $x_1, \ldots, x_n \in I$, $\alpha_1, \ldots, \alpha_n \in [0;1]$ mit $\alpha_1 + \cdots + \alpha_n = 1$. Dann gilt*

$$f(\alpha_1 x_1 + \cdots + \alpha_n x_n) \leq \alpha_1 f(x_1) + \cdots + \alpha_n f(x_n)$$

Bemerkung: Der Fall $n = 2$ ist genau die Definition von Konvexität.

Beweis. O.B.d.A können wir annehmen, dass $\alpha_i > 0$ für alle $i = 1, \ldots, n$. Wir setzen $\overline{x} = \alpha_1 x_1 + \cdots + \alpha_n x_n$. Ist $\overline{x}$ ein Endpunkt von I, dann müssen alle $x_i = \overline{x}$, und die Behauptung ist trivial. Wir können also annehmen, dass $\overline{x} \in I^\circ$. Nach Proposition 8.23 existiert $\overline{m} \in \mathbb{R}$, s.d.

$$\phi(x) = f(\overline{x}) + m(x - \overline{x}) \leq f(x)$$

für alle $x \in I$. Also

$$\begin{aligned} f(\overline{x}) = \phi(\overline{x}) = \phi\left(\sum_{i=1}^{n} \alpha_i x_i\right) &= f(\overline{x}) + m\left(\sum_{i=1}^{n} \alpha_i x_i - \overline{x}\right) \\ &= \sum_{i=1}^{m} \alpha_i \left(f(\overline{x}) + m(x_i - \overline{x})\right) = \sum_{i=1}^{m} \alpha_i \phi(x_i) \leq \sum_{i=1}^{n} \alpha_i f(x_i) \end{aligned}$$

wo wir die Tatsache benutzt haben, dass $\sum_i \alpha_i = 1$. $\square$

Bemerkung: Ist f streng konvex auf I, dann gilt

$$f\left(\sum_{i=1}^{n} \alpha_i x_i\right) < \sum_{i=1}^{n} \alpha_i f(x_i)$$

falls es mindestens zwei Indizes $i \neq j$, mit $\alpha_i, \alpha_j \neq 0$ und $x_i \neq x_j$ gibt.

8.6 Taylor Polynome

Satz 8.25 (Taylor'sche Approximation mit Lagrange'schen Restglied)**.** *Sei $f \in C^{m-1}([a;x])$ und m-mal differenzierbar auf $(a;x)$. Dann existiert $\xi \in (a;x)$ mit*

$$f(x) = f(a) + f'(a)(x-a) + \cdots + \frac{f^{(m-1)}(a)}{(m-1)!}(x-a)^{m-1} + \frac{f^{(m)}(\xi)}{m!}(x-a)^m$$

Es gilt natürlich eine analoge Aussage für $x < a$ und $f \in C^{m-1}([x;a])$, m-mal auf $(x;a)$ differenzierbar.

Bemerkung: Es ist wichtig zu bemerken, dass ξ von x abhängt.
Bemerkung: Man nennt

$$p_{m-1}(x) = f(a) + f'(a)(x-a) + \cdots + \frac{f^{(m-1)}(a)}{(m-1)!}(x-a)^{m-1}$$

das $(m-1)$-te Taylor'sche Polynom für f um $x = a$. Die Taylor-Polynome geben eine Approximation für f in der Nähe von $x = a$. Der Fehler dieser Approximation hat gemäss Satz 8.25 die Form

$$f(x) - p_{m-1}(x) = \frac{f^{(m)}(\xi)}{m!}(x-a)^m$$

Beweis. Wir definieren $c \in \mathbb{R}$ durch die Gleichung

$$f(x) - p_{m-1}(x) = c\frac{(x-a)^m}{m!}$$

Wir müssen zeigen, dass $\xi \in (a;x)$ mit $c = f^{(m)}(\xi)$ existiert. Dazu definieren wir

$$h(t) = f(t) + f'(t)(x-t) + \dots + \frac{f^{(m-1)}(t)}{(m-1)!}(x-t)^{m-1} + c\frac{(x-t)^m}{m!} - f(x)$$

Aus Annahme ist h stetig auf $[a;x]$ und differenzierbar auf $(a;x)$. Es gilt offenbar $h(x) = 0$ und, aus Definition von c, $h(a) = 0$. Der Satz von Rolle impliziert, dass ein $\xi \in (a;x)$ mit $h'(\xi) = 0$ existiert. Da aber

$$h'(t) = \frac{(x-t)^{m-1}}{(m-1)!}\left(f^{(m)}(t) - c\right)$$

impliziert $h'(\xi) = 0$, dass $f^{(m)}(\xi) = c$. □

Es folgt aus dem Satz, dass p_{m-1} eine gute Approximation von f in der Nähe von $x = a$ gibt. Der Fehler konvergiert gegen Null, für $x \to a$, mindestens so schnell wie $(x-a)^m$.

Korollar 8.26. *Sei $f \in C^{m-1}([a;b])$, m-mal differenzierbar auf $(a;b)$, mit $|f^{(m)}(t)| \leq M$ für alle $t \in (a;b)$. Dann gilt*

$$|f(x) - p_{m-1}(x)| \leq M\frac{(x-a)^m}{m!}$$

für alle $x \in (a;b)$.

Ist weiter $f^{(m)}$ stetig, dann verschwindet die Differenz $f - p_m$ schneller als $(x-a)^m$.

Korollar 8.27. *Sei $f \in C^m([a;b])$. Dann gilt*

$$\lim_{x \downarrow a} \frac{|f(x) - p_m(x)|}{(x-a)^m} = 0$$

Beweis. Wir haben

$$f(x) - p_m(x) = f(x) - p_{m-1}(x) - \frac{f^{(m)}(a)}{m!}(x-a)^m = \frac{(x-a)^m}{m!}\left(f^{(m)}(\xi_x) - f^{(m)}(a)\right)$$

Für ein geeignete $a < \xi_x < x$. Das gibt

$$\frac{|f(x) - p_m(x)|}{(x-a)^m} = \frac{1}{m!}\left|f^{(m)}(\xi_x) - f^{(m)}(a)\right| \to 0$$

als $x \to a$, weil $f^{(m)}$ ist, aus Annahme, stetig. □

Es ist nützlich, hier die Landau'schen O, o Symbole einzuführen. Wir betrachten einen Grenzübergang $x \to a$ (oder $x \to \infty$, $x \to -\infty$). Die Schreibweise $\psi(x) = O(h(x))$ für $x \to a$ bedeutet, dass

$$|\psi(x)/h(x)| \le C < \infty$$

bleibt beschränkt im Limes $x \to a$. Die Schreibweise $\psi(x) = o(h(x))$ für $x \to a$ bedeutet, dass

$$|\psi(x)/h(x)| \to 0$$

für $x \to a$. ZB. $(4x+1)/(x^2-3x+5) = O(1/x)$ oder $(4x+1)/(x^2-3x+5) = o(1)$ für $x \to \infty$. Die zwei letzten Korollare können dann wie folgt umgeschrieben werden:

- Sei $f \in C^{(m-1)}([a;b])$, mit $f^{(m)}$ beschränkt auf $[a;b]$. Dann gilt

  $$f(x) - p_{m-1}(x) = O((x-a)^m)$$

  im Limes $x \downarrow a$.

- Sei $f \in C^m([a;b])$. Dann gilt, für $x \downarrow a$,

  $$f(x) - p_m(x) = o((x-a)^m)\,.$$

Analoge Aussagen gelten offenbar für den linksseitigen Limes.

Eine Anwendung der Approximation von f durch Polynome ist die folgende Charackterisierung von lokaler Extremalstelle.

Proposition 8.28. *Sei f m-mal stetig differenzierbar in einer Umgebung von $a \in \mathbb{R}$. Es sei $f'(a) = f''(a) = \cdots = f^{(m-1)}(a) = 0$ aber $f^{(m)}(a) \neq 0$. Dann*

a) Ist m gerade und $f^{(m)}(a) > 0$, dann ist a ein lokales Minimum von f.

b) Ist m gerade und $f^{(m)}(a) < 0$, dann ist a ein lokales Maximum von f.

c) Ist m ungerade, dann ist a keine lokale Extremalstelle.

Im Fall c) (falls $m \ge 3$) wechselt auch das Vorzeichen von f'' an der Stelle $x = a$.

Bemerkung: Man nennt einen Punkt a, wo das Vorzeichen von f'' wechselt, einen Wendepunkt. Die Funktion wechselt an einem Wendepunkt von konvex zu konkav oder von konkav zu konvex. Die Tangenten sind auf einer Seite der Wendepunkt unterhalb des Graphen und auf der andere Seite der Wendepunkt oberhalb des Graphen der Funktion. Ein Wendepunkt $a \in \mathbb{R}$ mit $f'(a) = 0$ heisst ein Sattelpunkt.

Beweis. Aus Satz 8.25 folgt, dass, für alle x in einer punktierten Umgebung von a $\xi_x \in (x;a)$ existiert, falls $x < a$ und $\xi_x \in (a;x)$, falls $x > a$, mit

$$f(x) = f(a) + \frac{f^{(m)}(\xi_x)}{m!}(x-a)^m$$

Da $f^{(m)}$ stetig an der Stelle a ist, hat $f^{(m)}(\xi_x)$ dasselbe Vorzeichen wie $f^{(m)}(a)$ für $|x-a|$ klein genug. Ist m gerade, dann ist $(x-a)^m > 0$; deswegen hat $f(x) - f(a)$ dasselbe

Vorzeichen wie $f^{(m)}(a)$, für $|x-a|$ klein genug. Das zeigt die Behauptungen a) und b). Ist dagegen m ungerade, so wechselt das Vorzeichen von $(x-a)^m$ bei $x=a$. Deswegen wird auch das Vorzeichen von $f(x)-f(a)$ bei $x=a$ wechseln, und $x=a$ ist weder ein lokales Maximum noch ein lokales Minimum. Wir behaupten, dass in diesem Fall (für $m \geq 3$) auch das Vorzeichen von $f''(x)$ in $x=a$ wechseln wird. In der Tat, $f \in C^m(I)$ (wobei I ein Intervall um a ist) impliziert, dass $f'' \in C^{m-2}(I)$. Deswegen finden wir für $x \in I$ ξ_x mit

$$f''(x) = f''(a) + \frac{f'''(a)}{1!}(x-a) + \cdots + \frac{f^{(m)}(\xi_x)}{(m-2)!}(x-a)^{m-2} = \frac{f^{(m)}(\xi_x)}{(m-2)!}(x-a)^{m-2}$$

Für $|x-a|$ klein genug, hat $f^{(m)}(\xi_x)$ dasselbe Vorzeichen wie $f^{(m)}(a)$. Da m ungerade ist, ist auch $m-2$ ungerade. Das Vorzeichen von $(x-a)^{m-2}$ wird also bei $x=a$ wechseln, und so wird auch das Vorzeichen von $f''(x)$. □

8.7 Taylorreihen und analytische Funktionen

Nehmen wir nun an, dass $f \in C^\infty(I)$ für ein offenes Interval I um einen Punkt $a \in \mathbb{R}$. Wir können dann Taylor Polynome

$$p_m(x) = \sum_{j=0}^{m} \frac{f^{(j)}(a)}{j!}(x-a)^j$$

beliebiger Ordnung definieren. Satz 8.25 und seine Korollare implizieren, p_m gibt, mit steigenden m, eine Approximation von f neben a, die immer genauer wird.

Es ist dann natürlich sich zu fragen, ob man $m \to \infty$ streben lassen kann. D.h. man kann sich fragen, ob die Taylorreihe $\sum_{m=0}^\infty f^{(m)}(a)(x-a)^m/m!$ konvergiert (in einer Umgebung von $x=a$) und, falls ja, ob sie die Funktion f darstellt. Die Antwort zu diesen Fragen hängt von der Funktion f ab. Wir betrachten einige Beispiele:

- Sei $f(x) = \exp(x)$. Dann gilt $f^{(n)}(x) = \exp(x)$ für alle $n \in \mathbb{N}$. Also $f^{(n)}(0) = 1$ für alle n. Die Taylor-Reihe um $x=0$ ist gerade $\sum_{n\geq 0} x^n/n!$. Der Konvergenzradius der Taylorreihe ist ∞ und, aus Definition von $\exp(x)$, stellt die Taylorreihe die Funktion f dar, für alle $x \in \mathbb{R}$. Man kann auch die Taylorreihe von $\exp(x)$ um $a \neq 0$ betrachten. Sie ist aus

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

  gegeben. Also hat, für alle $a \in \mathbb{R}$, die Taylorreihe von $f(x) = \exp(x)$ um $x=a$ Konvergenzradius ∞ und stellt die Funktion $f(x)$ dar, für alle $x \in \mathbb{R}$.

- Die Situation ist ähnlich für die Funktionen $\sin x$ und $\cos x$.

- Sei nun $f(x) = (1+x)^\alpha$, für ein $\alpha \neq 0, 1, 2, \dots$. $f \in C^\infty(I)$ für eine Umgebung I von 0 (zB. für $I = (-1;1)$). Wir untersuchen die Taylorreihe für f um $x=0$. Es gilt

$$f^{(j)}(x) = \alpha(\alpha-1)\dots(\alpha-j+1)(1+x)^{\alpha-j} \;\Rightarrow\; f^{(j)}(0) = \alpha(\alpha-1)\dots(\alpha-j+1)$$

Die Taylorreihe für f um $x = 0$ ist also aus

$$\sum_{j=0}^{\infty} \frac{\alpha(\alpha-1)\dots(\alpha-j+1)}{j!} x^j$$

definiert. Für beliebige $\alpha \neq 0$ und $j \in \mathbb{N}$ führen wir die Bezeichnung

$$\binom{\alpha}{j} := \frac{\alpha(\alpha-1)\dots(\alpha-j+1)}{j!}$$

ein (für $\alpha \in \mathbb{N}$ ist das der gewöhnliche binomiale Koeffizient). Der Konvergenzradius der Taylorreihe ist

$$\rho = \lim_{j\to\infty} \frac{\binom{\alpha}{j}}{\binom{\alpha}{j+1}} = \lim_{j\to\infty} \left| \frac{j+1}{\alpha - j} \right| = 1$$

Die Taylorreihe konvergiert also für alle $|x| < 1$. Stellt die Taylorreihe die Funktion f dar? Die Antwort ist ja, d.h.

$$\sum_{j=0}^{\infty} \binom{\alpha}{j} x^j = (1+x)^\alpha$$

gilt für alle $|x| < 1$. Wir zeigen die Behauptung für alle $-1/4 < x \leq 1/4$. Aus Satz 8.25 haben wir

$$(1+x)^\alpha = \sum_{j=0}^{n-1} \binom{\alpha}{j} x^j + \frac{f^{(n)}(\xi)}{n!} x^n$$

für ein $\xi \in (-1/4; 1/4)$. Für $n > \alpha$ gilt

$$\left| \frac{f^{(n)}(\xi)}{n!} \right| = \binom{\alpha}{n} |(1+\xi)^{\alpha-n}| \leq \left(\frac{4}{3}\right)^n \binom{\alpha}{n}$$

Also, für $x \in (-1/4; 1/4)$, haben wir

$$\left| \frac{f^{(n)}(\xi)}{n!} x^n \right| \leq \frac{1}{3^n} \binom{\alpha}{n} =: a_n$$

Da

$$\left| \frac{a_{n+1}}{a_n} \right| = \frac{1}{3} \frac{n-\alpha}{n+1} \to \frac{1}{3} < 1$$

konvergiert $a_n \to 0$, für $n \to \infty$. Das zeigt, dass

$$(1+x)^\alpha = \sum_{j=0}^{\infty} \binom{\alpha}{j} x^j$$

für alle $|x| \leq 1/4$ (wie oben erwähnt, gilt die Formel eigentlich für alle $|x| < 1$). Es folgt, dass die Taylorreihe für f um $x = 0$ konvergiert und stellt die Funktion f in einer Umgebung von $x = 0$ dar.

- Sei nun

$$f(x) = \begin{cases} e^{-1/x}, & \text{für } x > 0 \\ 0, & \text{für } x \leq 0 \end{cases}$$

Wir behaupten nun $f \in C^\infty(\mathbb{R})$. Es ist klar, dass $f \in C^\infty(\mathbb{R}\backslash\{0\})$. Wir zeigen, dass f an der Stelle $x = 0$ beliebig oft differenzierbar ist. Dazu müssen wir beweisen, dass $f^{(n)}(0) = 0$ für alle $j \in \mathbb{N}$ (weil $f^{(n)}(x) = 0$ für alle $x < 0$, und, falls $f^{(n)}$ stetig ist, muss auch $f^{(n)}(0) = 0$). Für $n = 1$ gilt

$$\lim_{h\downarrow 0} \frac{f(h) - f(0)}{h} = \lim_{h\downarrow 0} \frac{1}{h} e^{-1/h} = 0$$

Da

$$\lim_{h\uparrow 0} \frac{f(h) - f(0)}{h} = 0$$

es folgt, dass f an der Stelle 0 differenzierbar ist, und, dass $f'(0) = 0$. Nehmen wir nun an $f^{(m)}(0) = 0$ für alle $m \leq n$. Wir zeigen, dass $f^{(n)}(0) = 0$. Dazu bemerken wir, dass für alle $n \in \mathbb{N}$, ein Polynom p_n existiert mit

$$\frac{d^n}{dx^n} e^{-1/x} = p_n(1/x) e^{-1/x}$$

für alle $x > 0$ (diese Behauptung kann auch induktiv gezeigt werden). Also, für alle $h > 0$,

$$\frac{f^{(n)}(h) - f^{(n)}(0)}{h} = \frac{1}{h} p_n(1/h) e^{-1/h} \to 0$$

für $h \downarrow 0$. Das zeigt, dass

$$\lim_{h\downarrow 0} \frac{f^{(n)}(h) - f^{(n)}(0)}{h} = \lim_{h\uparrow 0} \frac{f^{(n)}(h) - f^{(n)}(0)}{h} = 0$$

Es folgt, dass $f^{(n)}$ differenzierbar an der Stelle $x = 0$ ist, und dass $f^{(n+1)}(0) = 0$. Wir haben bewiesen, $f \in C^\infty(\mathbb{R})$, mit $f^{(n)}(0) = 0$ für alle $n \geq 0$. Die Taylorreihe von f ist also identisch Null. Der Konvergenzradius der Reihe ist unendlich, aber die Reihe gibt in keiner Umgebung von 0 eine Darstellung von $f(x)$ (trotzdem ist die Reihe eine gute Approximation für f, weil $f(x) = o(x^n)$ für $x \to 0$, für alle $n \geq 0$).

Im letzten Beispiel lässt sich die Funktion nicht durch die Taylorreihe darstellen. Ist es möglich, dass f durch eine andere Potenzreihe in einer Umgebung um $x = 0$ darstellbar ist? Die Antwort ist nein. Lässt sich eine Funktion in einer Umgebung von $x = a$ durch eine Potenzreihe $\sum_{n=0}^\infty a_n (x-a)^n$ darstellen, dann ist insbesondere f unendlich oft differenzierbar und $a_n = f^{(n)}(a)/n!$. D.h. die Taylorreihe ist die einzige mögliche Potenzreihe, die die Funktion f darstellen kann. Formal ist es einfach, diese Tatsache zu verstehen. Gilt $f(x) = \sum_{n=0}^\infty a_n (x-a)^n$ in einer Umgebung von $x = a$, dann, angenommen man kann Ableitungen mit der unendlichen Summe vertauschen, finden wir

$$f^{(j)}(x) = \sum_{n=j}^\infty a_n n(n-1)\dots(n-j+1)(x-a)^{n-j} \;\Rightarrow\; f^{(j)}(a) = a_j j! \;\Rightarrow a_j = \frac{f^{(j)}(a)}{j!}$$

und deswegen

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n$$

Um dieses Argument präzise zu machen, möchten wir nun zeigen, dass die Ableitung einer Funktionenreihe (einer Potenzreihe) gleich zum Limes der Ableitung der Partialsummen ist. Allgemeiner kann man fragen ob die Ableitung des Limes einer differenzierbaren Funktionenfolge gleich zum Limes der Ableitung der Funktionenfolge ist. Das ist nicht immer der Fall, wie die folgenden Beispiele zeigen.

- Sei $f_n(x) = (1/n)\sin(nx)$, $x \in \mathbb{R}$. Dann gilt $f_n \to 0$ gleichmässig. Aber $f_n'(x) = \cos(nx)$ konvergiert nicht.
- Sei $f_n(x) = (x^2+1/n^2)^{1/2}$ auf $\mathbb{R}$. Dann ist f_n überall differenzierbar. Aber $f_n(x) \to |x|$ gleichmässig und $|x|$ ist nicht differenzierbar an der Stelle $x = 0$.

Proposition 8.29. *Sei $I \subset \mathbb{R}$ ein offenes Intervall. $f_n \in C^1(I)$ eine Funktionenfolge, mit $f_n \to f$ gleichmässig auf I und mit $f_n' \to g$ gleichmässig auf I. Dann ist $f \in C^1(I)$ und $f' = g$.*

Bemerkung: Die Proposition bedeutet, dass

$$\frac{d}{dx} \lim_{n\to\infty} f_n(x) = \lim_{n\to\infty} \frac{df_n}{dx}(x)$$

angenommen f_n und die Ableitung f_n' konvergieren gleichmässig.

Beweis. $f_n \in C^1$ für alle n impliziert, dass f_n' stetig für alle $n \in \mathbb{N}$. $f_n' \to g$ gleichmässig impliziert, dass g stetig ist. Sei nun $x_0 \in I$ beliebig. Wir möchten zeigen, dass

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

Sei $\varepsilon > 0$ fest. Da g stetig ist, existiert $\delta > 0$ mit $|g(x_0) - g(x)| < \varepsilon$ für alle x mit $|x - x_0| < \delta$. Sei nun $0 < |h| < \delta$. Dann, aus dem Mittelwertsatz, existiert für alle $n \in \mathbb{N}$ ein ξ_n, mit $|\xi_n - x_0| < \delta$ und

$$\frac{f_n(x_0+h) - f_n(x_0)}{h} = f_n'(\xi_n)$$

Das gibt

$$\begin{aligned}
\Big|\frac{f(x_0+h) - f(x_0)}{h} - g(x_0)\Big| & \\
&\leq \frac{1}{h}|f(x_0+h) - f_n(x_0+h)| + \frac{1}{h}|f(x_0) - f_n(x_0)| + \big|f_n'(\xi_n) - g(x_0)\big| \\
&\leq \frac{2}{h}\sup_x |f(x) - f_n(x)| + \sup_x |f_n'(x) - g(x)| + |g(\xi_n) - g(x_0)| \\
&\leq \frac{2}{h}\sup_x |f(x) - f_n(x)| + \sup_x |f_n'(x) - g(x)| + \varepsilon
\end{aligned}$$

Da die linke Seite unabängig von n ist, können wir $n \to \infty$ streben lassen. Wir kriegen

$$\left| \frac{f(x_0+h)-f(x_0)}{h} - g(x_0) \right| \leq \frac{2}{h} \lim_{n\to\infty} \sup_x |f(x)-f_n(x)| + \lim_{n\to\infty} \sup_x |f_n'(x)-g(x)| + \varepsilon = \varepsilon$$

für alle h mit $0 < |h| < \delta$. Das zeigt die Behauptung. $\square$

Wir können nun das Resultat von Proposition 8.29 auf Potenzreihen anwenden.

Satz 8.30. *Sei $\sum_n a_n z^n$ eine Potenzreihe mit Konvergenzradius $\rho > 0$. Sei $f(x) = \sum_{n=0}^{\infty} a_n x^n$, für $|x| < \rho$. Dann*

a) Die Potenzreihe $\sum_{n=0}^{\infty}(n+1)a_{n+1}x^n$ hat den Konvergenzradius ρ.

b) Die Funktion f ist differenzierbar auf $(-\rho;\rho)$, mit

$$f'(x) = \sum_{n=0}^{\infty}(n+1)a_{n+1}x^n$$

Bemerkung: Der Satz impliziert, dass konvergente Potenzreihen gliedweise differenziert werden dürfen.

Beweis. Es gilt $\limsup_{n\to\infty} |(n+1)a_{n+1}|^{1/n} = \limsup_{n\to\infty} |a_n|^{1/n}$. Das zeigt a). Um b) zu beweisen wählen wir $0 < r < \rho$. Wir bemerken, dass

$$s_n(x) = \sum_{j=0}^{n} a_j x^j \to \sum_{j=0}^{\infty} a_j x^j =: f(x)$$

gleichmässig in $[-r;r]$, und dass

$$s_n'(x) = \sum_{j=1}^{n} j a_j x^{j-1} = \sum_{j=0}^{n-1}(j+1)a_{j+1}x^j \to \sum_{j=0}^{\infty}(j+1)a_{j+1}x^j$$

gleichmässig in $[-r;r]$. Proposition 8.29 impliziert, dass f differenzierbar auf $[-r;r]$, und dass $f'(x) = \sum_{j=0}^{\infty}(j+1)a_{j+1}x^j$ ist. $\square$

Korollar 8.31. *Sei $\rho > 0$ der Konvergenzradius der Potenzreihe $\sum_{n\geq 0} a_n x^n$. Wir setzen $f(x) = \sum_{n\geq 0} a_n x^n$ für $|x| < \rho$. Dann $f \in C^{\infty}((-\rho;\rho))$,*

$$f^{(j)}(x) = \sum_{n=0}^{\infty}(n+1)(n+2)\dots(n+j)a_{n+j}x^n$$

für alle $|x| < \rho$ und alle $j \in \mathbb{N}$. Insbesondere

$$f^{(j)}(0) = j! a_j \;\Rightarrow\; a_j = \frac{f^{(j)}(0)}{j!}$$

für alle $j \in \mathbb{N}$.

Definition 8.32. *Eine Funktion f, definiert in einer Umgebung von $x_0 \in \mathbb{R}$, heisst analytisch an der Stelle x_0, falls eine Potenzreihe mit $f(x) = \sum_{n \geq 0} a_n(x - x_0)^n$ für alle x in einer Umgebung von x_0 existiert. Die Funktion f, definiert auf einem offenen Gebiet $G \subset \mathbb{R}$, heisst analytisch auf G, falls sie analytisch an der Stelle x ist, für alle $x \in G$.*

Bemerkung: Es folgt aus Korollar 8.31, dass, falls f analytisch an der Stelle x_0 ist, dann ist f unendlich oft differenzierbar an der Stelle x_0 und, dass die Potenzreihe $\sum_n a_n(x - x_0)^n$, die f in einer Umgebung von x_0 darstellt, die Koeffizienten

$$a_n = \frac{f^{(n)}(x_0)}{n!}$$

hat. D.h., f ist analytisch an der Stelle x_0 g.d.w. die Taylorreihe für f an der Stelle x_0 Konvergenzradius $\rho > 0$ hat, und, falls sie die Funktion f in einer Umgebung von x_0 darstellt.

Bemerkung: Die Funktion $f : \mathbb{R} \to \mathbb{R}$, definiert durch $f(x) = e^{-1/x}$ für $x > 0$ und $f(x) = 0$ für $x \leq 0$, ist ein Beispiel einer Funktion $f \in C^\infty(\mathbb{R})$, die aber in $x = 0$ nicht analytisch ist. Für eine beliebige offene Menge $G \subset \mathbb{R}$ haben wir die Inklusionen

$$C^0(G) \supset C^1(G) \supset C^2(G) \supset \cdots \supset C^\infty(G) \supset \text{Menge der analytischen Funktionen auf } G$$

Da per Definition die Summe von zwei analytischen Funktionen und das Produkt einer analytische Funktion mit einer skalare $\lambda \in \mathbb{R}$ offensichtlich wieder analytisch sind, hat die Menge der analytischen Funktionen die Struktur eines Vektorraumes.

Analytische Funktionen haben sehr schöne Eigenschaften. Z.B. sind die Nullstellen einer analytischen Funktion immer isoliert.

Proposition 8.33. *Sei f analytisch an der Stelle x_0, mit $f(x_0) = 0$. Dann ist entweder f identisch Null in einer Umgebung von x_0 oder es gibt eine Umgebung U von x_0 mit $f(x) \neq 0$ für alle $x \in U \backslash \{0\}$.*

Beweis. Nehmen wir an f, ist nicht identisch Null in einer Umgebung von x_0. Da f analytisch an der Stelle x_0, gilt

$$f(x) = \sum_{n=0}^{\infty} a_n(x - x_0)^n$$

in einer Umgebung von x_0. $f(x_0) = 0$ impliziert, dass $a_0 = 0$. Da f nicht identisch Null ist, existiert $m \in \mathbb{N}\backslash\{0\}$ mit $a_n = 0$ für alle $n < m$ aber $a_m \neq 0$. Dann ist

$$f(x) = (x - x_0)^m \left(a_m + a_{m+1}(x - x_0) + \dots\right) =: (x - x_0)^m g(x)$$

für die Funktion

$$g(x) = \sum_{j=0}^{\infty} a_{m+j}(x - x_0)^j$$

die analytisch an der Stelle x_0 ist (der Konvergenzradius der Potenzreihe für g ist gleich zum Konvergenzradius der Potenzreihe für f, und deswegen strikt positiv). Es gilt $g(x_0) = a_m \neq 0$ und weil g stetig ist, ist $g(x) \neq 0$ in einer Umgebung von x_0. Deswegen ist auch $f(x) = (x - x_0)^m g(x)$ nicht Null für alle $x \neq x_0$ in einer Umgebung von x_0. □

Bemerkung: Die letze Proposition gilt für C^∞ Funktionen nicht. Z.B. die Funktion

$$f(x) = \begin{cases} e^{-1/|x|}\sin(1/x), & \text{falls } x \neq 0 \\ 0, & \text{falls } x = 0 \end{cases}$$

ist in $C^\infty(\mathbb{R})$, hat aber an der Stelle $x = 0$ eine nicht isolierte Nullstelle.

Eine wichtige Folgerung aus der letzten Proposition ist das Identitätsprinzip für analytische Funktionen.

Theorem 8.34 (Identitätsprinzip)**.** *Es seien f, g analytisch an der Stelle x_0. Es existiere eine Folge x_n mit $x_n \neq x_0$, $x_n \to x_0$ und $f(x_n) = g(x_n)$ für alle $n \in \mathbb{N}$. Dann ist $f(x) = g(x)$ in einer Umgebung von x_0.*

Beweis. Die Funktion $f - g$ ist analytisch an der Stelle x_0 und $(f - g)(x_k) = 0$ für alle $k \in \mathbb{N}$. Da $f - g$ stetig ist, es folgt auch, dass $f(x_0) = g(x_0)$. D.h. x_0 ist eine nicht isolierte Nullstelle von $f - g$. Aus Proposition 8.33 muss $f - g$ in einer Umgebung von x_0 identisch verschwinden. □

Bemerkung: Analytische Funktionen können in natürlicher Weise auf komplexe Gebiete erweitert werden. Die Untersuchung von komplex analytischen Funktionen ist das Hauptthema der Vorlesung komplexe Analysis (Funktionentheorie).

Proposition 8.35. *Sei f analytisch an der Stelle x_0, und es gelte $f(x) = \sum_n a_n(x-x_0)^n$ für alle x mit $|x - x_0| < \rho$. Dann ist f auch analytisch an der Stelle x_1, für alle x_1 mit $|x_1 - x_0| < \rho$. Für $|x - x_1| < \widetilde{\rho} = \rho - |x_1 - x_0|$ gilt weiter*

$$f(x) = \sum_{j=0}^{\infty} b_j (x - x_0)^j$$

mit den Koeffizienten

$$b_j = \sum_{n \geq j} \binom{n}{j} a_n (x_1 - x_0)^{n-j} = \sum_{m=0}^{\infty} \binom{j+m}{j} a_{m+j} (x_1 - x_0)^m$$

Idee des Beweis. Aus dem Satz über das Cauchy-Produkt zwei absolut konvergenter Reihen finden wir

$$\begin{aligned} f(x) &= \sum_{n=0}^{\infty} a_n (x - x_0)^n = \sum_{n=0}^{\infty} a_n (x - x_1 + x_1 - x_0)^n \\ &= \sum_{n=0}^{\infty} a_n \sum_{j=0}^{n} \binom{n}{j} (x_1 - x_0)^{n-j} (x - x_1)^j \\ &= \sum_{j=0}^{\infty} \left(\sum_{n=j}^{\infty} \binom{n}{j} a_n (x_1 - x_0)^{n-j} \right) (x - x_1)^j = \sum_{j=0}^{\infty} b_j (x - x_1)^j \end{aligned}$$

□

Anwendung: Betrachte die Funktion $f(x) = (1+x)^\alpha$, für ein $\alpha \notin \mathbb{N}$. Wir haben schon gezeigt, dass f analytisch an der Stelle $x = 0$ ist, mit

$$f(x) = \sum_{j=0}^{\infty} \binom{\alpha}{j} x^j \tag{23}$$

für alle $|x| \leq 1/4$. Wir behaupten nun, dass (23) für alle $|x| < 1$ gilt. Um das zu zeigen, bemerken wir zunächst, dass, für alle $x_0 > -1$,

$$\begin{aligned}(1+x_0+h)^\alpha &= (1+x_0)^\alpha \left(1 + \frac{h}{1+x_0}\right)^\alpha \\ &= (1+x_0)^\alpha \sum_{j=0}^{\infty} \binom{\alpha}{j} \left(\frac{h}{1+x_0}\right)^j = \sum_{j=0}^{\infty} \binom{\alpha}{j} \frac{1}{(1+x_0)^{j-\alpha}} h^j\end{aligned}$$

für alle $|h|$ klein genug. Das bedeutet, dass

$$(1+x)^\alpha = \sum_{j=0}^{\infty} \binom{\alpha}{j} \frac{1}{(1+x_0)^{j-\alpha}} (x-x_0)^j$$

für alle x genügend nah bei x_0, d.h. $(1+x)^\alpha$ ist analytisch auf $(-1;+\infty)$. Anderseits ist die Funktion $g(x) = \sum_{j=0}^{\infty} \binom{\alpha}{j} x^j$ analytisch auf $(-1;1)$ (weil die Reihe Konvergenzradius $\rho = 1$ hat). Also sind f und g analytisch auf $(-1;1)$ und sie sind identisch auf $[-1/4;1/4]$. Aus dem Identitätsprinzip folgt dann, dass $f = g$ auf $(-1;1)$. In der Tat, nehmen wir an, dass $f \not\equiv g$ auf $(-1;-1/4)$. Wir setzen dann $x_0 = \sup\{x \in (-1;-1/4) : f(x) \neq g(x)\}$. Aus Definition gilt dann $f(x_0+1/n) = g(x_0+1/n)$ für alle $n \in \mathbb{N}$. Das Identitätsprinzip impliziert, dass $f = g$ in einer Umgebung von x_0, in Widerspruch zur Definition von x_0. Analog kann man zeigen, dass $f \equiv g$ auf $(1/2;1)$.

Weitere Tatsache über analytische Funktionen (ohne Beweis):

- Seien f, g analytisch an der Stelle x_0. Dann ist auch das Produkt fg analytisch an der Stelle x_0 (auch hier verwendet man das Cauchy-Produkt zweier absolut konvergenten Reihen).

- Ist f analytisch an der Stelle x_0 und $f(x_0) \neq 0$, dann ist $1/f$ analytisch an der Stelle x_0.

- Ist f analytisch an der Stelle x_0 und g analytisch an der Stelle $f(x_0)$, dann ist $g \circ f$ analytisch an der Stelle x_0.

- Ist f analytisch an der Stelle x_0 und invertierbar, und es gelte $f'(x_0) \neq 0$, dann ist f^{-1} analytisch an der Stelle $f(x_0)$.

Die Beweise sind zum Teil komplizierte Übungen in Umordnung von Reihen. Es gibt bessere Beweise mit den Methoden der komplexen Analysis.

9 Riemann'sches Integral

9.1 Definition und elementare Eigenschaften

Ziel: für eine Funktion $f : [a; b] \to [0; \infty)$ möchten wir den Flächeninhalt von $\{(x, y) : a \leq x \leq b, 0 \leq y \leq f(x)\}$ berechnen (und definieren).

Intuitive Konstruktion. Wir wählen Punkte $x_0, x_1, \dots, x_n \in \mathbb{R}$, mit $a = x_0 < x_1 < \dots < x_{n-1} < x_n = b$. Das zerlegt $[a; b]$ in n Intervalle $[x_0; x_1], [x_1; x_2], \dots, [x_{n-1}; x_n]$. Für jede $j = 1, \dots, n$, wählen wir einen Repräsentanten $\xi_j \in [x_{j-1}; x_j]$. Die Fläche von $\{(x, y) : x_{j-1} \leq x \leq x_j, 0 \leq y \leq f(\xi_j)\}$ ist ungefähr $f(\xi_j)(x_j - x_{j-1})$. Die gesamte Fläche von $\{(x, y) : a \leq x \leq b, 0 \leq y \leq f(x)\}$ ist ungefähr aus der Riemann'sche Summe

$$\sum_{j=1}^{n} f(\xi_j)(x_j - x_{j-1})$$

gegeben. Die Hoffnung ist dann, dass die Riemann'sche Summe konvergiert, da die Teilung unendlich fein wird. Das Integral ist dann als der Grenzwert der Riemann'sche Summe definiert.

Genaue Konstruktion. Wir betrachten ein kompaktes Intervall $I = [a; b]$ und eine beschränkte $\mathbb{R}$-wertige Funktion f auf I. Eine Teilung von $[a; b]$ ist eine endliche Teilmenge $T = \{x_0 < x_1 < \dots < x_{n-1} < x_n\}$ von $[a; b]$, mit $x_0 = a$ und $x_n = b$. Zu einer gegebenen Teilung T, definieren wir die Intervalle $I_j = [x_{j-1}, x_j]$, für $j = 1, 2, \dots, n$. Es gilt

$$I = I_1 \cup I_2 \cup \dots \cup I_n$$

Die Intervalle I_j sind fast disjunkt, mit

$$I_j \cap I_{j+1} = \{x_j\} \text{ und } I_i \cap I_j = \emptyset, \text{ falls } i \neq j, j \pm 1$$

Sei T eine Teilung von $[a; b]$. Eine zu T entsprechende Familie von Repräsentanten ist ein n-Tupel $\xi = (\xi_1, \xi_2, \dots, \xi_n\}$ mit $\xi_j \in I_j$ für alle $j = 1, \dots, n$. Für gegebene Teilung T und Familie von Representanten ξ definieren wir die Riemann'sche Summe

$$S(T, \xi) = \sum_{j=1}^{n} f(\xi_j)(x_j - x_{j-1})$$

wobei $(x_j - x_{j-1}) = |I_j|$ die Länge von I_j ist. Wir definieren nun die obere Riemann'sche Summe zur Teilung T:

$$\overline{S}(T) = \sup_{\xi} S(T, \xi) = \sum_{j=1}^{n} \sup\{f(x) : x \in I_j\} \, |I_j|$$

Die untere Riemann'sche Summe zur Teilung T ist

$$\underline{S}(T) = \inf_{\xi} S(T, \xi) = \sum_{j=1}^{n} \inf\{f(x) : x \in I_j\} |I_j|$$

Offenbar gilt $\overline{S}(T) \geq \underline{S}(T)$ für alle Teilungen T von $[a; b]$. Intuitiv $\underline{S}(T) \leq$ Flächeninhalt $\leq \overline{S}(T)$.

Wir sagen, eine Teilung T' ist eine Verfeinerung von T falls $T' \supset T$ (eine feinere Teilung ist eine grössere Menge).

Lemma 9.1. *Wir haben die folgenden Eigenschaften:*

a) Sei $T' \supset T$ eine Verfeinerung von T. Dann gilt $\overline{S}(T') \leq \overline{S}(T)$ und $\underline{S}(T') \geq \underline{S}(T)$.

b) Es gilt

$$\sup_T \underline{S}(T) \leq \inf_T \overline{S}(T)$$

Beweis. a) Es genügt, den Fall $|T'| = |T|+1$ zu betrachten. Wir haben also $T' = T \cup \{\hat{x}\}$, für ein $\hat{x} \in [a;b]$. Ist $T = \{x_0 < x_1 < \cdots < x_{n-1} < x_n\}$, dann gibt es ein $k \in \{1, 2, \ldots, n\}$ mit $x_{k-1} < \hat{x} < x_k$. Also $T' = \{x_0 < x_1 < \cdots < x_{k-1} < \hat{x} < x_k < \cdots < x_n\}$. Es gilt $I'_j = I_j$ für $1 \leq j \leq k-1$, $I'_{j+1} = I_j$ für $k+1 \leq j \leq n$, und $I_k = I'_k \cup I'_{k+1}$. Da

$$\sup_{I'_k} f \leq \sup_{I_k} f, \quad \sup_{I'_{k+1}} f \leq \sup_{I_k} f$$

finden wir

$$\sup_{I'_k} f\,|I'_k| + \sup_{I'_{k+1}} f|I'_{k+1}| \leq \sup_{I_k} f(|I'_k| + |I'_{k+1}|) = \sup_{I_k} f|I_k|$$

Also

$$\begin{aligned}\overline{S}(T) &= \sup_{I_1} f|I_1| + \cdots + \sup_{I_k} f|I_k| + \cdots + \sup_{I_n} f|I_n| \\ &\geq \sup_{I'_1} f|I'_1| + \cdots + \sup_{I'_k} f|I'_k| + \sup_{I'_{k+1}} f|I'_{k+1}| + \cdots + \sup_{I_n} f|I_n| \\ &= \overline{S}(T')\end{aligned}$$

Ähnlich kann man zeigen, dass $\underline{S}(T) \leq \underline{S}(T')$.

b) Seien T_1, T_2 zwei Teilungen von $[a;b]$. Wir setzen $T_3 = T_1 \cup T_2$. T_3 ist eine Verfeinerung von T_1 und von T_2. Aus a), und weil offenbar $\underline{S}(T) \leq \overline{S}(T)$ für jede Teilung T, bekommen wir

$$\underline{S}(T_1) \leq \underline{S}(T_3) \leq \overline{S}(T_3) \leq \overline{S}(T_2)$$

Also $\underline{S}(T_1) \leq \overline{S}(T_2)$ für jede zwei Teilungen T_1, T_2. Das impliziert, dass $\sup_T \underline{S}(T) \leq \inf_T \overline{S}(T)$. $\square$

Definition 9.2. *Die reelwertige beschränkte Funktion f auf $[a;b]$ heisst Riemann integrierbar, falls*

$$\sup_T \underline{S}(T) = \inf_T \overline{S}(T)\,.$$

In diesem Fall definieren wir das (Riemann'sche) Integral von f auf $[a;b]$ durch

$$\int_a^b f(x)dx := \sup_T \underline{S}(T) = \inf_T \overline{S}(T)$$

Bemerkungen. Wir benutzen auch die Notation $\int_a^b f dx$ oder einfach $\int_a^b f$ für das Integral von f auf $[a;b]$. Das Differntial dx erinnert an der $\Delta x = x_j - x_{j-1}$ in der Riemann'schen Summe; es hat aber keine Bedeutung im Integral. In Analysis 3 werden wir sehen, dass eine andere alternative Konstruktion des Integrals existiert, das Lebesgue-Integral. Der Hauptvorteil des Lebesgue-Integrals ist, dass es für allgemeinere Funktionen existiert (und dass es bessere Eigenschaften bezüglich Vertausch von Grenzwert und Integral hat). Wann das Riemann Integral existiert, stimmt es mit dem Lebesgue Integral überein.

Proposition 9.3. *a) f ist genau dann integrierbar, wenn* $\inf_T \left(\overline{S}(T) - \underline{S}(T)\right) = 0$.

b) Sei T_n eine Familie von Teilungen mit $\overline{S}(T_n) - \underline{S}(T_n) \to 0$, als $n \to \infty$. Dann ist f integrierbar und

$$\overline{S}(T_n) \to \int_a^b f(x)dx, \quad \textit{und} \quad \underline{S}(T_n) \to \int_a^b f(x)dx$$

Weiter, falls $\xi^{(n)}$ eine beliebige Familie von Representanten zur Teilung T_n ist, haben wir

$$S(T_n, \xi^{(n)}) \to \int_a^b f(x)dx$$

Bemerkung: Da $\overline{S}(T) \geq \underline{S}(T)$ für alle T, die Bedingung $\inf \left(\overline{S}(T) - \underline{S}(T)\right) = 0$ ist mit $\inf \left(\overline{S}(T) - \underline{S}(T)\right) \leq 0$ äquivalent.

Beweis. a) Nehmen wir an, die Bedingung $\inf \left(\overline{S}(T) - \underline{S}(T)\right) = 0$ ist erfüllt. Dann

$$0 = \inf \left(\overline{S}(T) - \underline{S}(T)\right) \geq \inf \overline{S}(T) - \sup \underline{S}(T)$$

und $\sup \underline{S}(T) \geq \inf \overline{S}(T)$. Da die umgekehrte Ungleichung automatisch ist, gilt $\sup \underline{S}(T) = \inf \overline{S}(T)$. Das bedeutet, dass f integrierbar ist.

Nehmen wir nun an, f ist integrierbar. Sei nun $\varepsilon > 0$, T_1, T_2 zwei Teilungen mit

$$\overline{S}(T_1) - \inf_T \overline{S}(T) \leq \varepsilon/2$$

und

$$\sup_T \underline{S}(T) - \underline{S}(T_2) \leq \varepsilon/2$$

Da f integriebar ist, ist $\inf \overline{S} = \sup \underline{S}$. Also

$$\overline{S}(T_1) - \underline{S}(T_2) \leq \varepsilon$$

Sei nun $T_3 = T_1 \cup T_2$. Da T_3 eine Verfeinerung von T_1, T_2 ist, gilt $\overline{S}(T_3) \leq \overline{S}(T_1)$ und $\underline{S}(T_3) \geq \underline{S}(T_2)$. Deswegen

$$\overline{S}(T_3) - \underline{S}(T_3) \leq \overline{S}(T_1) - \underline{S}(T_2) \leq \varepsilon$$

Da $\varepsilon > 0$ beliebig, gilt

$$\inf_T \left(\overline{S}(T) - \underline{S}(T)\right) \leq 0$$

b) Aus $\overline{S}(T_n) - \underline{S}(T_n) \to 0$ folgt, dass $\inf(\overline{S}(T) - \underline{S}(T)) = 0$. Aus a) folgt, dass f integrierbar ist. Da

$$\underline{S}(T_n) \leq \int_a^b f(x)dx \leq \overline{S}(T_n)$$

wir finden, dass $\underline{S}(T_n) \to \int_a^b f(x)dx$ und auch $\overline{S}(T_n) \to \int_a^b f(x)dx$. Weiter, da

$$\underline{S}(T_n) \leq S(T_n, \xi^{(n)}) \leq \overline{S}(T_n)$$

muss auch $S(T_n, \xi^{(n)}) \to \int_a^b f(x)dx$. $\square$

Es folgt aus der Proposition, dass um Integrale zu berechnen (und Integrierbarkeit zu prüfen) es genügt, eine spezielle Folge von Teilungen zu betrachten, nämlich eine Folge mit der Eigenschaft $\overline{S}(T_n) - \underline{S}(T_n) \to 0$. Ist die Funktion integrierbar, so konvergieren obere und untere Summe gegen das Integral von f, für beliebige Folgen von Teilungen T_n, falls die Länge jedes Intervalls von T_n gegen Null strebt. Für eine Teilung T von $[a;b]$ setzen wir $\|T\| = \max_{j=1,\dots,n} |x_j - x_{j-1}|$. D.h. $\|T\|$ ist die Länge des grössten Intervalls in der Teilung. Für eine beliebige Menge J und eine $\mathbb{R}$-wertige Funktion f auf J definieren wir auch die Oszillation von f auf J durch

$$\sigma(f, J) = \sup\{f(x) : x \in J\} - \inf\{f(x) : x \in J\} = \sup\{|f(x) - f(y)| : x, y \in J\}$$

Dann ist, für eine beliebige Teilung T,

$$\overline{S}(T) - \underline{S}(T) = \sum_{j=1}^{n} \sigma(f, I_j)|I_j|$$

Proposition 9.4. *Sei f auf $[a;b]$ integrierbar, und T_n eine Folge von Teilungen, mit $\|T_n\| \to 0$. Dann gilt*

$$\int_a^b f dx = \lim_{n\to\infty} \overline{S}(T_n) = \lim_{n\to\infty} \underline{S}(T_n) = \lim_{n\to\infty} S(T_n, \xi^{(n)})$$

wobei, für alle $n \in \mathbb{N}$, $\xi^{(n)}$ eine Familie von Repräsentanten zur Teilung T_n ist.

Beweis. Es genügt, die folgende Tatsache zu überprüfen: Ist T eine Teilung von $[a;b]$, mit $\delta = \min_{j=1,\dots,n} |I_j|$. Dann gilt

$$\overline{S}(T') - \underline{S}(T') \leq 3\left(\overline{S}(T') - \underline{S}(T')\right) \tag{24}$$

für alle Teilungen T' von $[a;b]$ mit $\|T'\| = \max |I_j| < \delta$. In der Tat, nehmen wir an (24) ist korrekt. Dann können wir wie folgt argumentieren. Da f integrierbar ist, finden wir, für alle $\varepsilon > 0$, eine Teilung $\widetilde{T}$ mit

$$\overline{S}(\widetilde{T}) - \underline{S}(\widetilde{T}) < \frac{\varepsilon}{3}$$

Wir setzen dann $\delta = \min |\widetilde{I}_j|$, wobei $\widetilde{I}_j$ die zur Teilung $\widetilde{T}$ entsprechenden Intervalle sind. Da $\|T_n\| \to 0$, für $n \to \infty$, finden wir N mit $\|T_n\| \leq \delta$, für alle $n > N$. Gleichung (24) impliziert dann, dass

$$\overline{S}(T_n) - \underline{S}(T_n) < \varepsilon$$

für alle $n > N$. Also $\overline{S}(T_n) - \underline{S}(T_n) \to 0$, für $n \to \infty$. Proposition 9.3 zeigt dann die Behauptung.

Es bleibt (24) zu zeigen. Wir bezeichnen mit I_j und I'_j die zu T und zu T' entsprechenden Intervalle. Es gilt $\max_j |I'_j| < \min_k |I_k|$. Deswegen schneidet jedes Intervall I'_j höchstens zwei I-Intervalle. Wir definieren $\varepsilon_{ij} = 1$ falls I_i I'_j schneidet(für jedes j gibt es höchstens zwei Indizien i, mit $\varepsilon_{ij} \neq 0$). Wir bemerken, dass

$$\sigma(f, I'_j) \leq \sum_i \varepsilon_{ij}\sigma(f, I_i)$$

Deswegen finden wir

$$\begin{aligned}\overline{S}(T') - \underline{S}(T') &= \sum_j \sigma(f, I'_j)|I'_j| \leq \sum_j \left(\sum_i \sigma(f, I_i) \right) |I'_j| \\ &= \sum_i \sigma(f, I_i) \left(\sum_j \varepsilon_{ij} |I'_j| \right) \leq 3 \sum_i \sigma(f, I_i)|I_i| = 3(\overline{S}(T) - \underline{S}(T))\end{aligned}$$

weil $\sum_j \varepsilon_{ij} |I'_j|$ die gesamte Länge aller I'-Intervalle ist, die I_i schneiden, was höchstens $3|I_i|$ sein kann (weil die I'-Intevalle alle kürzer als I_i sind). □

Um die letzte Proposition anzuwenden, und Integrale mit Hilfe beliebige Folge von Teilungen T_n, mit $\|T_n\| \to 0$ zu berechnen, müssen wir zunächst wissen, ob f auf $[a; b]$ integrierbar ist. Die folgende Proposition gibt eine erste wichtige hinreichende Bedingung für Integrierbarkeit.

Proposition 9.5. *Ist f stetig auf $[a; b]$, so ist f integrierbar auf $[a; b]$.*

Beweis. Eine stetige Funktion auf $[a; b]$ ist gleichmässig stetig. Sei $\varepsilon > 0$ fest. Dann existiert $\delta > 0$ mit

$$|x - y| \leq \delta \quad \Rightarrow \quad |f(x) - f(y)| \leq \frac{\varepsilon}{b - a}$$

Sei nun $T = \{x_0 < x_1 < \cdots < x_n\}$ eine Teilung mit $|x_{j+1} - x_j| < \delta$ für alle $1 \leq j \leq n$. Dann gilt $\sigma(f, I_j) \leq \varepsilon/(b - a)$ für alle $j = 1, \ldots, n$ und also

$$\overline{S}(T) - \underline{S}(T) = \sum_{j=1}^n \sigma(f, I_j)|I_j| \leq \frac{\varepsilon}{b - a} \sum_{j=1}^n |I_j| = \varepsilon$$

Das impliziert, dass

$$\inf \left(\overline{S}(T) - \underline{S}(T) \right) \leq 0$$

und damit, dass f integrierbar ist. □

Beispiel: Die Funktion $f(x) = 1/x$ ist stetig auf $[1; a]$. Wir möchten das Integral

$$\int_1^a \frac{1}{x} \, dx$$

Wir wählen die Teilung $T_n = \{a^{j/n} : 0 \leq j \leq n\}$. Dann $I_j = [a^{(j-1)/n}, a^{j/n}]$, $j = 1, \ldots, n$. Da $1/x$ monoton fallend ist, gilt

$$\sup\{f(x) : x \in I_j\} = a^{-(j-1)/n} \quad \text{und} \quad \inf\{f(x) : x \in I_j\} = a^{-j/n}$$

Also

$$\underline{S}(T_n) = \sum_{j=1}^n a^{-j/n}(a^{j/n} - a^{(j-1)/n} = \sum_{j=1}^n (1 - a^{-1/n}) = n(1 - a^{-1/n})$$

und

$$\overline{S}(T_n) = \sum_{j=1}^n a^{-(j-1)/n}(a^{j/n} - a^{(j-1)/n} = a^{1/n} \underline{S}(T_n)$$

Sei $f(t) = a^t$. Dann gilt

$$\underline{S}(T_n) = n(1 - a^{-1/n}) = \frac{f(0) - f(-1/n)}{1/n} \to f'(0) = \log\, a$$

für $n \to \infty$. Da $\overline{S}(T_n) = a^{1/n}\underline{S}(T_n) \to \log\, a$, folgt, dass

$$\int_1^a \frac{1}{x}\, dx = \log\, a$$

Beispiel: Sei

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

und $[a; b] \subset \mathbb{R}$ ein beliebiges Intervall. Es gilt $\sigma(f, I) = 1$ für jedes Intervall $I \subset \mathbb{R}$. Also gilt, für eine beliebige Teilung T von $[a; b]$

$$\overline{S}(T) - \underline{S}(T) = \sum_{j=1}^{n} \sigma(f, I_j)|I_j| = \sum_{j=1}^{n} |I_j| = (b - a)$$

Die Funktion f ist deswegen nicht integrierbar.

Proposition 9.6. *Sei $a < b < c$, und f eine beschränkte Funktion auf $[a; c]$. Dann ist f integrierbar auf $[a; c]$ g.d.w. f integrierbar auf $[a; b]$ und auf $[b; c]$ ist. In diesem Fall gilt*

$$\int_a^c f dx = \int_a^b f dx + \int_b^c f dx$$

Beweis. Sei f integrierbar auf $[a; b]$ und auf $[b; c]$. Wir finden Folgen $T_{1,n}$ und $T_{2,n}$, Teilungen von $[a; b]$ und, bzw. von $[b; c]$, mit

$$\overline{S}_{[a;b]}(T_{1,n}) - \underline{S}_{[a;b]}(T_{1,n}) \to 0 \quad \text{und} \quad \overline{S}_{[b;c]}(T_{2,n}) - \underline{S}_{[b;c]}(T_{2,n}) \to 0$$

für $n \to \infty$. Wir setzen $T_{3,n} = T_{1,n} \cup T_{2,n}$. T_n ist dann eine Folge von Teilungen von $[a; c]$ mit

$$\begin{aligned} \overline{S}_{[a;c]}(T_{3,n}) &= \overline{S}_{[a;b]}(T_{1,n}) + \overline{S}_{[b;c]}(T_{2,n}) \\ \underline{S}_{[a;c]}(T_{3,n}) &= \underline{S}_{[a;b]}(T_{1,n}) + \underline{S}_{[b;c]}(T_{2,n}) \end{aligned}$$

für alle n. Das impliziert, dass f auf $[a; c]$ integrierbar ist, weil

$$\overline{S}_{[a;c]}(T_{3,n}) - \underline{S}_{[a;c]}(T_{3,n}) = \left(\overline{S}_{[a;b]}(T_{1,n}) - \underline{S}_{[a;b]}(T_{1,n})\right) + \left(\overline{S}_{[b;c]}(T_{2,n}) - \underline{S}_{[b;c]}(T_{2,n})\right) \to 0$$

für $n \to \infty$, und, dass

$$\int_a^c f dx = \lim_{n\to\infty} \overline{S}_{[a;c]}(T_{3,n}) = \lim_{n\to\infty} \overline{S}_{[a;b]}(T_{1,n}) + \lim_{n\to\infty} \overline{S}_{[b;c]}(T_{2,n})$$

Übung: Zeige, dass f integrierbar auf $[a; c]$ impliziert, dass f integrierbar auf $[a; b]$ und auf $[b; c]$ ist. $\square$

Als Anwendung der letzten Proposition zeigen wir, dass jede auf $[a;b]$ beschränkte Funktion, mit endlich vielen Unstetigkeitstellen, integrierbar ist.

Proposition 9.7. *Sei f auf $[a;b]$ beschränkt, mit endlich vielen Unstetigkeitstellen. Dann ist f auf $[a;b]$ integrierbar.*

Beweis. Seien $y_1 < y_2 < \cdots < y_{m-1}$ die Unstetigkeitstellen von f in $(a;b)$. Wir setzten auch $y_0 = a$ und $y_m = b$. Ist f integrierbar auf $[y_{j-1};y_j]$ für alle $j = 1, \dots m$, so ist f integrierbar auf $[a;b]$ (aus Proposition 9.6). Es genügt also zu zeigen, dass f stetig auf $(a;b)$ und beschränkt auf $[a;b]$ impliziert, dass f integrierbar auf $[a;b]$ ist. Sei dazu $M > 0$ s.d. $|f(x)| \leq M$ auf $[a;b]$. Für ein beliebiges $\varepsilon > 0$, ist f auf $I_\varepsilon := [a+(\varepsilon/8M);b-(\varepsilon/8M)]$ integrierbar, weil f stetig auf diesem Intervall ist. Es existiert also eine Teilung T von I_ε mit

$$\overline{S}_{I_\varepsilon}(T) - \underline{S}_{I_\varepsilon}(T) \leq \varepsilon/2$$

Wir betrachten nun die Teilung $T' = T \cup \{a;b\}$ von $[a;b]$. Es gilt

$$\begin{aligned}\overline{S}_{[a;b]}(T') &- \underline{S}_{[a;b]}(T') \\ &= (\sigma(f,[a;a+(\varepsilon/8M)]) + \sigma(f,[b-(\varepsilon/8M);b]))\,\frac{\varepsilon}{8M} + \overline{S}_{I_\varepsilon}(T) - \underline{S}_{I_\varepsilon}(T) \leq \varepsilon\end{aligned}$$

weil $\sigma(J;f) \leq 2M$ für jede Menge $J \subset [a;b]$. Da $\varepsilon > 0$ beliebig ist, ist f auf $[a;b]$ integrierbar. □

Bemerkung: Die Bedingung für Integrierbarkeit in der letzten Proposition ist hinreichend, aber nicht notwendig. Eine Menge $N \subset \mathbb{R}$ heisst eine Nullmenge, falls für alle $\varepsilon > 0$ eine endlich oder abzählbare Familie $\{J_i\}$ von offenen Intervallen existiert, mit

$$N \subset \bigcup_i J_i \quad \text{und} \quad \sum_i |J_i| \leq \varepsilon$$

Jede abzählbare Menge ist offenbar eine Nullmenge, aber es existieren auch überabzählbare Nullmengen. Tatsache: Eine beschränkte Funktion f auf $[a;b]$ ist genau dann auf $[a;b]$ integrierbar, falls $\{x \in [a;b]; f \text{ unstetig an der Stelle } x\}$ eine Nullmenge ist.

Wir untersuchen nun elementare Eigenschaften vom Integral.

Proposition 9.8. *Seien f, g integrierbar über $[a;b]$.*

a) Für alle $\alpha, \beta \in \mathbb{R}$ ist $\alpha f + \beta g$ integrierbar auf $[a;b]$ und

$$\int_a^b (\alpha f + \beta g)\, dx = \alpha \int_a^b f dx + \beta \int_a^b g dx$$

(Das Integral ist linear).

b) Gilt $f(x) \geq g(x)$ für alle $x \in [a;b]$, so ist

$$\int_a^b f dx \geq \int_a^b g dx$$

(Das Integral ist monoton).

c) $|f|$ ist integrierbar auf $[a;b]$ und

$$\left| \int_a^b f dx \right| \leq \int_a^b |f| dx$$

(Dreiecksungleichung für Integrale). Es folgt, dass $\max(f,g), \min(f,g)$ *integrierbar auf $[a;b]$ sind.*

Beweis: c) Aus der gewöhnlichen Dreiecksungleichung gilt

$$||f(x)| - |f(y)|| \leq |f(x) - f(y)|$$

D.h. die Oszillationen von $|f|$ auf jedem Intervall $J \subset \mathbb{R}$ sind durch die Oszillationen von f beschränkt:

$$\sigma(|f|, J) \leq \sigma(f, J) \quad \text{for all } J \subset \mathbb{R}$$

Also

$$\overline{S}(|f|, T) - \underline{S}(|f|, T) \leq \overline{S}(f, T) - \underline{S}(f, T)$$

für alle Teilungen T. Das impliziert, dass

$$\inf_T \left(\overline{S}(|f|, T) - \underline{S}(|f|, T) \right) \leq \inf_T \left(\overline{S}(f, T) - \underline{S}(f, T) \right) = 0$$

und deswegen, dass $|f|$ integrierbar ist. Die Dreiecksungleichung für Integrale folgt aus $-|f(x)| \leq f(x) \leq |f(x)|$ und aus der Monotonie des Integrals (Teil b), Beweis unten). Die Integrierbarkeit von $\max(f,g)$ und $\min(f,g)$ folgt aus der Bemerkung, dass $\max(f,g) = (f+g)/2 + |f-g|/2$ und $\min(f,g) = (f+g)/2 - |f-g|/2$, aus der Linearität (Teil a), Beweis unten), und aus der Integrierbarkeit des Absolutbetrags.

a) Es genügt zu zeigen, dass

$$\begin{aligned}
&i) \int (f+g) dx = \int f dx + \int g dx \\
&ii) \int (\alpha f) dx = \alpha \int f dx \quad \text{für alle } \alpha > 0 \\
&iii) \int (-f) dx = - \int f dx
\end{aligned}$$

Wir zeigen zunächst i). Für eine beliebige Teilung T von $[a;b]$ gilt

$$\sup\{f(x) + g(x) : x \in I_j\} \leq \sup\{f(x) : x \in I_j\} + \sup\{g(x) : x \in I_j\}$$

Also

$$\overline{S}(f+g, T) \leq \overline{S}(f, T) + \overline{S}(g, T)$$

Ähnlicherweise

$$\underline{S}(f+g, T) \geq \underline{S}(f, T) + \underline{S}(g, T)$$

Seien nun T_n^f und T_n^g Folgen von Teilungen mit der Eigenschaften

$$\overline{S}(f, T_n^f) - \underline{S}(f, T_n^f) \to 0$$

und

$$\overline{S}(f, T_n^g) - \underline{S}(g, T_n^g) \to 0$$

für $n \to \infty$. Sei nun $T_n = T_n^f \cup T_n^g$. Dann gilt (da T_n eine Verfeinerung von T_n^f ist)

$$\overline{S}(f, T_n) - \underline{S}(f, T_n) \leq \overline{S}(f, T_n^f) - \underline{S}(f, T_n^f)$$

und also $\overline{S}(f, T_n) - \underline{S}(f, T_n) \to 0$, für $n \to \infty$. Ähnlicherweise $\overline{S}(g, T_n) - \underline{S}(g, T_n) \to 0$, für $n \to \infty$. Damit

$$\begin{aligned} \overline{S}(f+g, T_n) - \underline{S}(f+g, T_n) &\leq \overline{S}(f, T_n) + \overline{S}(g, T_n) - \underline{S}(f, T_n) - \underline{S}(g, T_n) \\ &= \big(\overline{S}(f, T_n) - \underline{S}(f, T_n)\big) + \big(\overline{S}(g, T_n) - \underline{S}(g, T_n)\big) \to 0 \end{aligned}$$

Das zeigt, dass $f+g$ integrierbar ist. Es gilt

$$\underline{S}(f, T_n) + \underline{S}(g, T_n) \leq \underline{S}(f+g, T_n) \leq \int_a^b (f+g)dx \overline{S}(f+g, T_n) \leq \overline{S}(f, T_n) + \overline{S}(g, T_n)$$

Da

$$\begin{aligned} \underline{S}(f, T_n) + \underline{S}(g, T_n) &\to \int_a^b f dx + \int_a^b g dx \quad \text{und} \\ \overline{S}(f, T_n) + \overline{S}(g, T_n) &\to \int_a^b f dx + \int_a^b g dx \end{aligned}$$

für $n \to \infty$, folgt, dass

$$\int_a^b (f+g)dx = \int_a^b f dx + \int_a^b g dx$$

Wir beweisen nun ii). Offenbar gilt

$$\overline{S}(\alpha f, T) = \alpha \overline{S}(f, T), \quad \text{und} \quad \underline{S}(\alpha f, T) = \alpha \underline{S}(f, T)$$

für jede Teilung T. Ist T_n eine Folge von Teilungen, mit $\overline{S}(f, T_n) - \underline{S}(f, T_n) \to 0$, so gilt auch

$$\overline{S}(\alpha f, T_n) - \underline{S}(\alpha f, T_n) = \alpha \left(\overline{S}(f, T_n) - \underline{S}(f, T_n)\right) \to 0$$

Damit ist αf integrierbar, und

$$\int_a^b \alpha f dx = \lim_{n \to \infty} \overline{S}(\alpha f, T_n) = \alpha \lim_{n \to \infty} \overline{S}(f, T_n) = \alpha \int_a^b f dx$$

Um iii) zu zeigen, bemerken wir, dass

$$\overline{S}(-f, T) = -\underline{S}(f, T) \quad \text{und } \underline{S}(-f, T) = -\overline{S}(f, T)$$

für jede Teilung T, weil $\sup(-f) = -\inf f$. Ist T_n eine Teilung mit $\overline{S}(f, T_n) - \underline{S}(f, T_n) \to 0$, so gilt auch

$$\overline{S}(-f, T_n) - \underline{S}(-f, T_n) = -\underline{S}(f, T_n) + \overline{S}(f, T_n) \to 0$$

für $n \to \infty$. Damit ist $-f$ integrierbar, und

$$\int_a^b (-f)dx = \lim_{n\to\infty} \overline{S}(-f, T_n) = -\lim_{n\to\infty} \underline{S}(f, T_n) = -\int_a^b f dx$$

Damit ist Teil a) bewiesen.

b) Ist $f \geq g$, so ist $f - g \geq 0$ auf $[a; b]$. Da aus a) $f - g$ integrierbar ist, muss

$$\int_a^b (f - g)dx \geq 0$$

(weil jede Riemannsche Summe positiv ist). Wieder aus a) folgt, dass

$$\int_a^b f dx - \int_a^b g dx = \int_a^b (f - g)dx \geq 0 \;\Rightarrow \int_a^b f dx \geq \int_a^b g dx$$

Damit ist auch Teil b) gezeigt. □

Die elementare Eigenschaften des Integrals aus der letzten Proposition haben einige einfache aber wichtige Folgerungen.

Satz 9.9. *a) Konstante Funktionen sind integrierbar,*

$$\int_a^b c dx = c(b - a)$$

b) Sei f integrierbar auf $[a; b]$, und

$$m = \inf\{f(x) : x \in [a; b]\}, \quad M = \sup\{f(x) : x \in [a; b]\}$$

Dann gilt

$$m(b - a) \leq \int_a^b f dx \leq M(b - a)$$

c) (Mittelwertsatz für Integrale). Ist f stetig auf $[a; b]$, dann existiert $\xi \in (a; b)$ mit

$$\int_a^b f dx = f(\xi)(b - a)$$

Beweis. a) $f(x) = c$ impliziert, dass $\underline{S}(f, T) = \overline{S}(f, T) = c(b - a)$ für alle Teilungen T. Damit ist

$$\int_a^b c dx = c(b - a)$$

b) Es gilt $m \leq f(x) \leq M$ für alle $x \in [a; b]$. Die Monotonie des Integrals impliziert, dass

$$m(b - a) = \int_a^b m dx \leq \int_a^b f dx \leq \int_a^b M dx = M(b - a)$$

c) Aus b) folgt, dass

$$m \leq \frac{1}{b - a} \int_a^b f dx \leq M$$

Aus dem Satz von Maximum folgt, dass $x_0, x_1 \in [a;b]$ mit $f(x_0) = m$ und $f(x_1) = M$ existiert. Aus dem Zwischenwertsatz, existiert ξ zwischen x_0 und x_1 mit

$$f(\xi) = \frac{1}{b-a} \int_a^b f dx$$

□

9.2 Hauptsatz der Integralrechnung

Satz 9.10 (Hauptsatz)**.** *Sei f stetig auf $[a;b]$,*

$$F(x) = \int_a^x f(t)dt$$

für $x \in (a;b]$.

a) F ist differenzierbar auf $(a;b)$ mit $F'(x) = f(x)$.

b) Sei G stetig auf $[a;b]$, differenzierbar auf $(a;b)$ mit $G'(x) = f(x)$. Dann gilt

$$F(x) = G(x) - G(a)$$

für alle $a < x \leq b$.

Beweis. a) Sei $x_0 \in (a;b)$ fest. Wir berechnen $F'(x_0)$. Sei zunächst $h > 0$. Dann

$$F(x_0 + h) = \int_a^{x_0+h} f dt = \int_a^{x_0} f dt + \int_{x_0}^{x_0+h} f dt = F(x_0) + f(\xi)h$$

für ein $\xi \in (x_0; x_0 + h)$. Also

$$\frac{F(x_0+h) - F(x_0)}{h} = f(\xi) \to f(x_0)$$

für $h \downarrow 0$, aus Stetigkeit von f. Für $h < 0$ haben wir ähnlich

$$\frac{F(x_0+h) - F(x_0)}{h} = \frac{F(x_0) - F(x_0 - |h|)}{|h|} = \frac{1}{|h|} \int_{x_0-|h|}^{x_0} f dt = f(\xi)$$

für ein $x_0 - |h| < \xi < x_0$. Die Stetigkeit von f zeigt, dass

$$\lim_{h \uparrow 0} \frac{F(x_0+h) - F(x_0)}{h} = f(x_0)$$

Damit ist F differenzierbar an der Stelle x_0, mit $F'(x_0) = f(x_0)$.

b) Es gilt $G'(x) = f(x) = F'(x)$. Also $(F - G)' = 0$, und deswegen, $G(x) = F(x) + c$ für eine Konstante $c \in \mathbb{R}$, und alle $x \in (a;b)$. Da $\lim_{x \downarrow a} F(x) = 0$, und (aus Stetigkeit von G), $\lim_{x \downarrow a} G(x) = G(a)$, gilt $c = G(a)$. Damit ist $F(x) = G(x) - G(a)$, für alle $x \in (a;b)$. □

Man nennt eine Funktion G, stetig auf $[a;b]$, differenzierbar auf $(a;b)$, mit $G'(x) = f(x)$ für alle $x \in (a;b)$, eine Stammfunktion von f auf $[a;b]$. Ist G eine Stammfunktion einer stetigen Funktion f, so gilt

$$\int_a^b f dt = G(b) - G(a)$$

Beachte: Nicht stetige integrierbare Funktionen brauchen keine Stammfunktion zu haben (z.B. hat die integrierbare Funktion f, definiert durch $f(x) = 0$ für $x < 0$ und $f(x) = 1$ für $x \geq 0$, keine Stammfunktion auf $[-1;1]$. Anderseits impliziert die Existenz einer Stammfunktion von$f$ nicht, dass f integrierbar ist.

Bemerkung: Ist G eine Stammfunktion von f, dann ist auch $G + c$, für irgendeine Konstante $c \in \mathbb{R}$, eine Stammfunktion von f. Alle Stammfunktionen von f haben diese Form. Wir definieren das *unbestimmtes Integral* von f, als die Familie aller Stammfunktionen von f:

$$\int f dx = G(x) + c$$

falls $G'(x) = f(x)$. Manchmal ist es wichtig, das Intervall zu schreiben, wo die Relation $G' = f$ gilt. Zusammenfassend: Das bestimmte Integral $\int_a^b f dx$ ist eine Zahl (Grenzwert von Riemann'schen Summen). Das unbestimmte Integral $\int f dx$ ist dagegen die Familie aller Stammfunktionen von f. Gemäss Hauptsatz ist die Berechnung des unbestimmten Integrals von grosser Hilfe in der Berechnung des bestimmten Integrals.
Notation: Für $b < a$ setzen wir

$$\int_a^b f dx := -\int_b^a f dx$$

Für $b = a$, dagegen,

$$\int_a^a f dx := 0$$

Damit gilt

$$\int_a^b f dx = G(b) - G(a)$$

falls f stetig ist, und falls G eine Stammfunktion von f ist, unabhängig von der Ordnung von a, b.

Die Berechnung von unbestimmten Integralen ist nicht immer einfach. Unbestimmte Integrale sind nicht immer durch die bekannten Elementarfunktionen darstellbar (die Ableitung jeder durch Elementarfunktionen darstellbaren Funktion ist dagegen wieder durch elementare Funktionen darstellbar). Ein wichtiges Beispiel einer Funktion, deren Stammfunktion (die "error-function") nicht elementar darstellbar ist, ist $f(x) = e^{-x^2}$. Hier sind dagegen einige einfache Beispiele, für welche die Stammfunktion elementar darstellbar ist.

- $\frac{d}{dx}x^\alpha = \alpha x^{\alpha-1} \;\Rightarrow\; \int x^\alpha dx = \frac{x^{\alpha+1}}{\alpha+1} + c,$
  für alle $\alpha \neq -1$ und, falls $\alpha < 0$, für alle $x \neq 0$
- $\frac{d}{dx}\log|x| = \frac{1}{x} \;\Rightarrow\; \int \frac{1}{x}dx = \log|x| + c, \quad$ für alle $x \neq 0$
- $\frac{d}{dx}e^{\lambda x} = \lambda e^{\lambda x} \;\Rightarrow\; \int e^{\lambda x}dx = \frac{1}{\lambda}e^{\lambda x} + c$
- $\frac{d}{dx}\sin x = \cos x \;\Rightarrow\; \int \cos x\, dx = \sin x + c$
- $\frac{d}{dx}\cos x = -\sin x \;\Rightarrow\; \int \sin x dx = -\cos x + c$
- $\frac{d}{dx}\sinh x = \cosh x \;\Rightarrow\; \int \cosh x\, dx = \sinh x + c$
- $\frac{d}{dx}\cosh x = \sinh x \;\Rightarrow\; \int \sinh x\, dx = \cosh x + c$
- $\frac{d}{dx}\arctan x = \frac{1}{1+x^2} \;\Rightarrow\; \int \frac{1}{x^2+1}\, dx = \arctan x + c$
- $\frac{d}{dx}\arcsin x = \frac{1}{\sqrt{1-x^2}} \;\Rightarrow\; \int \frac{1}{\sqrt{1-x^2}}\, dx = \arcsin x + c, \quad$ für $x \in (-1;1)$
- $\frac{d}{dx}\text{arcsinh } x = \frac{1}{\sqrt{1+x^2}} \;\Rightarrow\; \int \frac{1}{\sqrt{1+x^2}}\, dx = \text{arcsinh } x + c$

9.3 Integrationsmethoden

Es gibt zwei allgemeine Bemerkungen, die bei der Berechnung von Integralen nützlich sein können; die Substitutionsformel und partielle Integration.

Satz 9.11 (Substitutionsformel)**.** *Sei f stetig und g stetig differenzierbar auf geeigneten Intervallen. Ist*

$$\int f(x)dx = F(x) + c$$

so ist

$$\int f(g(t))g'(t)dt = F(g(t)) + c \tag{25}$$

Mit anderen Worten, falls f stetig auf $[g(a); g(b)]$ ist, und g stetig differenzierbar auf $[a; b]$, haben wir

$$\int_a^b f(g(x))g'(x)dx = \int_{g(a)}^{g(b)} f(t)dt$$

(Es ist hier nicht nötig, dass $g(b) > g(a)$).

Beweis. Da F eine Stammfunktion von f ist, gilt $F'(x) = f(x)$. Aus der Kettenregel folgt, dass

$$\frac{d}{dx}F(g(x)) = f(g(x))g'(x)$$

D.h., $F(g(x))$ ist eine Stammfunktion von $f(g(x))g'(x)$. Also

$$\int_a^b f(g(x))g'(x)dx = F(g(b)) - F(g(a)) = \int_{g(a)}^{g(b)} f(t)dt$$

$\square$

Man kann den Faktor $g'(t)$ in der Substitutionsformel (25) als die Transformation des Differentials dt betrachten. Sei F eine Stammfunktion für f. Um das Integral

$$\int f(g(t))g'(t)dt$$

zu berechnen, setzen wir $x = g(t)$. Dann ist $f(g(t)) = f(x)$, und $dx = g'(t)dt$. Damit

$$\int f(g(t))g'(t)dt = \int f(x)dx = F(x) + c = F(g(t)) + c$$

Beispiele: Mit Hilfe der Substitutionsformel berechnen wir die folgenden unbestimmten Integrale.

1) Für $a \in \mathbb{R}$ fest, gilt

$$\int \frac{1}{x^2 + a^2} dx = \frac{1}{a^2} \int \frac{1 + (x/a)^2}{d} x$$

Sei $y = x/a$. Dann ist $dx = (1/a)dx$, und

$$\int \frac{1}{x^2 + a^2} dx = \frac{1}{a} \int \frac{1}{1 + y^2} dy = \frac{1}{a} \arctan y + c = \frac{1}{a} \arctan(x/a) + c$$

2) Wir untersuchen nun

$$\int \tan x \, dx = \int \frac{\sin x}{\cos x} dx$$

Sei $y = \cos x$. Dann ist $dy = -\sin x dx$ und

$$\int \tan x \, dx = -\int \frac{1}{y} dy = -\log y + c = -\log \cos x + c$$

3) Wir berechnen

$$\int x\sqrt{1 + x^2} \, dx$$

indem wir $t = 1 + x^2$ setzen. Dann ist $dt = 2xdx$ und

$$\int x\sqrt{1 + x^2} dx = \frac{1}{2} \int t^{1/2} dt = \frac{1}{3} t^{3/2} + c = \frac{1}{3}(1 + x^2)^{3/2} + c$$

Proposition 9.12 (Partielle Integration)**.** *Seien $u, v \in C^1([a; b])$. Dann gilt auf diesem Intervall*

$$\int u(x)v'(x)dx = u(x)v(x) - \int u'(x)v(x)dx$$

Es folgt

$$\int_a^b u(x)v'(x)dx = u(b)v(b) - u(a)v(a) - \int_a^b u'(x)v(x)dx$$

Beweis. Wir bemerken, dass

$$\frac{d}{dx}(u(x)v(x)) = u'(x)v(x) + u(x)v'(x)$$

Damit

$$\int u'(x)v(x)dx + \int u(x)v'(x)dx = u(x)v(x) + c$$

□

Beispiele: Wir untersuchen das Integral von xe^x. Wir setzen $u = x$ und $v' = e^x$. Dann ist $u' = 1$ und $v = e^x$. Damit

$$\int xe^x dx = xe^x - \int e^x dx = xe^x - e^x + c = e^x(x-1) + c$$

Ein anderes Beispiel ist das Integral von $x^2 \sin x$. Auch hier setzen wir $u = x^2$ und $v' = \sin x$. Dann ist $u' = 2x$ und $v = -\cos x$, und

$$\int x^2 \sin x dx = -x^2 \cos x + 2 \int x \cos x dx$$

Wir wenden noch ein Mal die partielle Integration an. Sei nun $u = x$ und $v' = \cos x$. Dann gilt $u' = 1$ und $v = \sin x$, und

$$\int x \cos x dx = x \sin x - \int \sin x dx = x \sin x + \cos x + c$$

Also

$$\int x^2 \sin x dx = -x^2 \cos x + 2x \sin x + 2 \cos x + c = (2 - x^2) \cos x + 2x \sin x + c\,.$$

Auch mit Substitution und partieller Integration Stammfunktionen zu finden ist i.A. eine schwierige Aufgabe. Es gibt aber einige spezielle Methoden, die die Berechnung von Integralen von besonderen Klassen von Funktionen erlauben. Die wichtigste Klasse von Funktionen, für die man immer eine elementare Stammfunktion finden kann, besteht aus allen rationalen Funktionen.

9.4 Integration von rationalen Funktionen: Partialbruchzerlegung

Das Integral einer rationalen Funktion kann immer mit der Methode der Partialbruchzerlegung berechnet werden. Eine rationale Funktion hat die Form p/q, wobei p, q Polynome sind. Es bezeichne $\deg p$ und $\deg q$ den Grad der Polynome p, q. Ist $\deg p \geq \deg q$, so kann man p durch q teilen. Man findet Polynome r, s mit $\deg s < \deg q$ und mit $p = rq + s$. Damit ist $p/q = r + s/q$. Das Integral von r kann sehr einfach berechnet werden; es bleibt das Integral von s/q zu berechnen. Mit anderen Worten, es genügt, rationale Funktionen p/q zu betrachten, mit $\deg p < \deg q$.

Es lohnt sich, komplexe Zahlen zu benutzen, um die Polynome q und p zu faktorisieren. Es existieren immer paarweise unterschiedliche $z_1, \dots, z_n \in \mathbb{C}$, eine Konstante a und $\alpha_1, \dots, \alpha_n \in \mathbb{N}$, mit

$$q(x) = a \prod_{j=1}^{n} (x - z_j)^{\alpha_j}$$

Die Zahlen $z_1, \dots, z_n$ sind die Nullstellen von q; sie heissen die Pole der rationalen Funktion p/q. Der Exponent α_j ist die Vielfachkeit oder die Ordnung, der Pol z_j. Es gilt $\sum_{j=1}^n \alpha_j = \deg q$. O.B.d.A können wir annehmen, dass p und q keine gemeinsame Nullstelle haben (sonst kann man die zwei Faktoren kürzen). Unter dieser Annahme finden wir, dass

$$\lim_{x \to z_j} (x - z_j)^{\alpha_j} \frac{p(x)}{q(x)} =: A \neq 0$$

Dann hat

$$\frac{p(x)}{q(x)} - \frac{A}{(x - z_j)^{\alpha_j}}$$

höchstens einen Pol der Ordnung $(m-1)$ an der Stelle z_j. In der Tat, sei das Polynom $\hat{q}$ so definiert, dass $q(x) = (x - z_j)^{\alpha_j} \hat{q}(x)$. Es gilt $\hat{q}(z_j) \neq 0$, und $A = p(z_j)/\hat{q}(z_j)$. Damit

$$\frac{p(x)}{q(x)} - \frac{A}{(x - z_j)^{\alpha_j}} = \frac{p(x) - A\hat{q}(x)}{q(x)}$$

Da der Numerator $p(z_j) - A\hat{q}(z_j) = 0$, hat $p/q - A/(x - z_j)^{\alpha_j}$ höchstens einen Pol der Vielfachkeit $\alpha_j - 1$ in z_j. Durch Wiederholung dieses Arguments, finden wir Konstanten $A_1, \dots, A_{\alpha_j}$, s.d.

$$\frac{p(x)}{q(x)} - \sum_{\ell=1}^{\alpha_j} \frac{A_\ell}{(x - z_j)^\ell}$$

keinen Pol an der Stelle z_j hat. Wiederholen wir das Argument für alle Pole, erhalten wir: es existieren (eindeutig bestimmte) Konstanten $A_{1,1}, \dots, A_{1,\alpha_1}, \dots, A_{n,1}, A_{n,\alpha_n}$ mit

$$\frac{p(x)}{q(x)} = \sum_{j=1}^{n} \sum_{\ell=1}^{\alpha_j} \frac{A_{j,\ell}}{(x - z_j)^\ell}$$

Diese Darstellung der rationalen Funktion p/q heisst eine Partialbruchzerlegung. Integration einer beliebigen rationalen Funktion wird somit auf das Problem der Berechnung der Integralen $1/(x - z_j)^k$ reduziert. Bemerke, dass auch für reellen rationale Funktionen (d.h. rationale Funktionen mit reellen Keoffizienten), die Nullstellen z_j des Polynoms q sind i.A. komplex. Nur auf $\mathbb{C}$ kann ein Polynom mit Sicherheit so einfach zerlegt werden. Wir müssen also komplex-wertige Funktionen integrieren. Wir definieren hier das (unbestimmte) Integral einer $\mathbb{C}$-wertige Funktion f durch

$$\int f dx = \int \text{Re } f dx + i \int \text{Im } f dx$$

(mit dieser Definition ist der Realteil von $\int f dx$ gleich zu $\int \text{Re } f dx$, und analog für den Imaginärteil). Für $k > 1$ gilt, ähnlich wie im Fall $z_j \in \mathbb{R}$,

$$\int \frac{1}{(x - z_j)^k} dx = \frac{-1}{k-1} \frac{1}{(x - z_j)^{k-1}} + c \tag{26}$$

Für $k = 1$ schreiben wir $z_j = a_j + ib_j$ und wir berechnen

$$\begin{aligned}\int \frac{1}{x-z_j} dx &= \int \frac{1}{(x-a_j)-ib_j}\, dx \\ &= \int \frac{(x-a_j)+ib_j}{(x-a_j)^2+b_j^2}\, dx \\ &= \int \frac{(x-a_j)}{(x-a_j)^2+b_j^2}\, dx + ib_j \int \frac{1}{(x-a_j)^2+b_j^2}\, dx \\ &= \frac{1}{2} \log\left((x-a_j)^2+b_j^2\right) + i \arctan \frac{x-a_j}{b_j} \\ &= \log|x-z_j| + i \arctan \frac{x - \mathrm{Re}\ z_j}{\mathrm{Im}\ z_j}\end{aligned} \tag{27}$$

Damit können wir das Integral (bestimmt oder unbestimmt) jeder rationalen Funktion berechnen. Zusammenfassend ist die Strategie um das Integral einer rationale Funktion zu bestimmen die folgende: Zunächst wird durch geeignete Division und Kürzung, das Problem auf die Berechnung des Integrals von p/q reduziert, wobei $\deg p < \deg q$ und p, q keine gemeinsamen Faktoren haben. Dann findet man alle Pole $z_1, \ldots, z_n$ von p/q, mit der entsprechenden Vielfachkeit α_j. Angenommen $q(x) = a \prod_{j=1}^n (x - z_j)^{\alpha_j}$, wir berechnen dann die Partialbruchzerlegung

$$\frac{p(x)}{q(x)} = \sum_{j=1}^{n} \sum_{\ell_j=1}^{\alpha_j} \frac{A_{j,\ell_j}}{(x-z_j)^{\ell}}\,.$$

von p/q. Die Berechnung der Koeffizienten A_{j,ℓ_j} reduziert sich nach Koeffizientenvergleich zur Lösung eines linearen Systems. Schlussendlich benutzen wir (26) und (27), um das Integral jeder Term auszurechnen.

Als Beispiel betrachten wir die rationale Funktion

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

Der Nenner hat die zwei Nullstellen $x = \pm i$, beide mit Multiplizität zwei. Es existieren also Konstanten $A, B, C, D \in \mathbb{C}$ mit

$$\begin{aligned}&R(x)\\ &= \frac{A}{x+i} + \frac{B}{(x+i)^2} + \frac{C}{x-i} + \frac{D}{(x-i)^2} \\ &= \frac{A(x+i)(x-i)^2 + B(x-i)^2 + C(x+i)^2(x-i) + D(x+i)^2}{(x^2+1)^2} \\ &= \frac{A(x^3 - ix^2 + x - i) + B(x^2 - 2ix - 1) + C(x^3 + ix^2 + x + i) + D(x^2 + 2ix - 1)}{(x^2+1)^2} \\ &= \frac{x^3(A+C) + x^2(-iA + B + iC + D) + x(A - 2iB + C + 2iD) - (iA + B - iC + D)}{(x^2+1)^2}\end{aligned}$$

Wir bekommen also die 4 Gleichungen

$$A + C = 0,\ -iA + B + iC + D = 0,\ A - 2iB + C + 2iD = 0,\ -iA - B + iC - D = 1$$

für die vier Unbekannten A, B, C, D. Die erste Gleichung gibt $C = -A$, die dritte also $B = D$. Die zweite Gleichung wird $iA = B$, und die vierte $A = -1/4i = i/4$. Also:

$$\frac{1}{(x^2+1)^2} = \frac{i}{4}\frac{1}{x+i} - \frac{1}{4}\frac{1}{(x+i)^2} - \frac{i}{4}\frac{1}{x-i} - \frac{1}{4}\frac{1}{(x-i)^2}$$

und damit

$$\begin{aligned}\int \frac{1}{(x^2+1)^2} &= \frac{i}{4}\left(\log|x+i| + i\arctan x - \log|x-i| + i\arctan x\right) + \frac{1}{4}\left(\frac{1}{x+i} + \frac{1}{x-i}\right) \\ &= -\frac{1}{2}\arctan x + \frac{1}{2}\frac{x}{x^2+1}\end{aligned}$$

weil $|x-i| = |x+i|$. Wie erwartet, ist das Endresultat reell (alle imaginären Beiträge kürzen sich weg).

Eine andere Klasse von Integralen, die man immer in geschlossene Form ausrechnen kann, sind Integrale der Form

$$\int R(\cos x; \sin x)dx$$

wobei $R(s;t)$ eine rationale Funktion von den zwei Variablen s, t ist (d.h. $R(s;t) = p(s;t)/q(s;t)$ für p, q Polynome in den Variablen s, t). Der Trick in diesem Fall ist die Substitution $u = \tan(x/2)$ durchzuführen. Dann ist $x = 2\arctan(u)$, und

$$dx = \frac{2}{1+u^2}\, du$$

Weiter gilt

$$\cos^2(x/2) = \frac{1}{1+\tan^2(x/2)} = \frac{1}{1+u^2}$$

und aus

$$\cos^2(x/2) = \frac{1-\cos x}{2}$$

findet man

$$\cos x = 1 - 2\cos^2(x/2) = 1 - \frac{2}{1+u^2} = \frac{u^2-1}{u^2+1}$$

und

$$\sin x = 2\sin(x/2)\cos(x/2) = 2\tan(x/2)\cos^2(x/2) = \frac{2u}{1+u^2}$$

Also, nach Substitution, wird

$$\int R(\cos x; \sin x)dx = \int \widetilde{R}(u)du$$

für eine neue rationale Funktion $\widetilde{R}$. Das Integral von $\widetilde{R}$ kann man dann durch die Methode der Partialbruchzerlegung berechnen.

Beispiel: Wir möchten das Integral

$$\int \frac{1}{\cos x} dx$$

berechnen. Wir setzen $u = \tan(x/2)$, und finden

$$\begin{aligned}\int \frac{1}{\cos x} dx &= \int \frac{u^2+1}{u^2-1} \frac{2}{1+u^2} du = 2 \int \frac{1}{u^2-1} du \\ &= \int \left(\frac{1}{u-1} - \frac{1}{u+1} \right) du \\ &= \log|u-1| - \log|u+1| + c = \log \frac{|\tan^2(x/2)-1|}{|\tan^2(x/2)+1|} + c\end{aligned}$$

Bemerke, dass trigonometrische Funktionen oft einfacher integriert werden können, ohne den Trick mit $u = \tan(x/2)$ zu benutzen.

Integrale der Form

$$\int R(x; \sqrt{1-x^2}) dx,$$

für eine rationale Funktion R, können mit der Substitution $x = \sin t$ berechnet werden. In der Tat, mit dieser Substitution finden wir $dx = \cos t \, dt$, $\sqrt{1-x^2} = \cos t$, und damit

$$\int R(x; \sqrt{1-x^2}) dx = \int \widetilde{R}(\sin t; \cos t) dt$$

für eine neue rationale Funktion $\widetilde{R}$. Das Integral auf der rechten Seite kann dann, wie oben erklärt, mit der Substitution $u = \tan(t/2)$ berechnet werden.

Integrale der Form

$$\int R(x; \sqrt{x^2-1}) dx, \text{ oder } \int R(x; \sqrt{1+x^2}) dx\,,$$

für eine rationale Funktion R, können mit der Substitution $x = \cosh t$, bzw. $x = \sinh t$ berechnet werden Mit dieser Substitution reduziert sich das Problem auf der Berechnung von Integralen der Form

$$\int \widetilde{R}(\cosh t; \sinh t) dt$$

für eine neue rationale Funktion $\widetilde{R}$. Die Substitution $u = e^t$, reduziert dann das Problem auf die Berechnung vom Integral von rationalen Funktionen in u. Da sich jeder quadratische Ausdruck $ax^2 + bx + c$ durch quadratische Ergänzung und lineare Substitution y in der Form $1 - y^2$, $1 + y^2$ oder $y^2 - 1$ schreiben lässt, folgt, dass man jedes Integral der Form

$$\int R(x; \sqrt{ax^2 + bx + c}) dx$$

explizit berechnen kann.

9.5 Vertausch von Grenzübergang und Integral

In dieser Sektion untersuchen wir die folgende Frage: Sei f_n eine Folge von auf $[a; b]$ integrierbare Funktionen, mit $f_n \to f$. Ist dann f auf $[a; b]$ integrierbar? Falls ja, ist das

Integral von f aus dem Grenzwert der Integrale von f_n gegeben? Mit anderen Worten, unter welchen Bedingungen gilt

$$\lim_{n\to\infty} \int_a^b f_n dx = \int_a^b \left(\lim_{n\to\infty} f_n \right) dx$$

Beispiel: Sei f_n die Folge von Funktionen auf $[-1;1]$, definiert durch

$$f_n(x) = \begin{cases} n + n^2 x & \text{für } -1/n \le x < 0 \\ n - n^2 x & \text{für } 0 < x < 1/n \\ 0 & \text{sonst} \end{cases}$$

Es gilt $f_n \to 0$ punktweise. Jede f_n ist auf $[-1;1]$ integrierbar, mit

$$\int f_n dx = 1$$

für alle n (der Graph von f beschreibt ein Dreieck, mit Basis $2/n$ und Höhe n). Also, in diesem Fall

$$1 = \lim_{n\to\infty} \int_{-1}^1 f_n dx \neq \int_{-1}^1 \left(\lim_{n\to\infty} f_n \right) dx = 0$$

Das Beispiel zeigt, dass punktweise Konvergenz von f_n nach f nicht genügt, um Grenzwert mit Integral zu vertauschen. Der nächste Satz zeigt, dass gleichmässige Konvergenz hinreichend ist.

Satz 9.13. *Sei f_n eine Folge von auf $[a;b]$ integrierbaren Funktionen, mit $f_n \to f$ gleichmässig auf $[a;b]$. Dann ist f auf $[a;b]$ integrierbar und*

$$\int_a^b f dx = \lim_{n\to\infty} \int_a^b f_n dx$$

Beweis. Sei $\varepsilon > 0$ fest. Wir finden dann $n \in \mathbb{N}$ mit

$$|f_n(x) - f(x)| < \frac{\varepsilon}{4(b-a)}$$

für alle $x \in [a;b]$. Da f_n integrierbar ist, finden wir auch eine Teilung T von $[a;b]$ mit

$$\overline{S}(f_n, T) - \underline{S}(f_n, T) < \frac{\varepsilon}{2}$$

Dann gilt

$$\begin{aligned} \overline{S}(f, T) &= \sum_{j=1}^m \sup\{f(x) : x \in I_j\} \, |I_j| \\ &\le \sum_{j=1}^m \left(\frac{\varepsilon}{4(b-a)} + \sup\{f_n(x) : x \in I_j\} \right) |I_j| \\ &= \frac{\varepsilon}{4} + \overline{S}(f_n, T) \end{aligned}$$

Analog

$$\underline{S}(f,t) \geq \underline{S}(f_n,T) - \frac{\varepsilon}{4}$$

und damit

$$\overline{S}(f,T) - \underline{S}(f,T) \leq \frac{\varepsilon}{2} + \overline{S}(f_n,T) - \underline{S}(f_n,T) \leq \varepsilon$$

Also, ist f integrierbar. Weiter gilt

$$\left| \int_a^b f dx - \int_a^b f_n dx \right| \leq \int_a^b |f - f_n| dx \leq (b-a) \sup_x |f - f_n|$$

Da die rechte Seite gegen Null konvergiert, muss

$$\int_a^b f_n dx \to \int_a^b f dx$$

für $n \to \infty$. □

Bemerkung: Gleichmässige Konvergenz ist zwar hinreichend, aber nicht notwendig für Konvergenz der Integrale. Mit der alternativen (und modernen) Definition des Integrals (das Lebesgue Integral, wird in der Vorlesung Analysis III diskutiert) ist es relativ einfach Bedingungen für Konvergenz von Integrale von Funktionenfolge die viel schwächer, und damit viel nützelicher sind, als gleichmässige Konvergenz der Folge.

9.6 Uneigentliche Integrale

Bis jetzt haben wir Integrale von beschränkten Funktionen auf kompakte Intervalle untersucht. Die Definition mit Riemann'schen Summen funktioniert für Integrale der Form

$$\int_0^\infty \frac{1}{1+x^2} dx \quad \text{oder} \quad \int_0^1 \frac{1}{\sqrt{x}} dx$$

nicht. Diese Integrale, die man als uneigentliche Integrale bezeichnet, kann man trotzdem als Grenzwerte von "eigentlichen Integrale" definieren. Für beliebige $y > 0$, gilt

$$\int_0^y \frac{1}{1+x^2} dx = \arctan(y) - \arctan(0) = \arctan(y)$$

Also können wir

$$\int_0^\infty \frac{1}{1+x^2} dx := \lim_{y \to \infty} \int_0^y \frac{1}{1+x^2} dx = \lim_{y \to \infty} \arctan(y) = \frac{\pi}{2}$$

definieren. Ähnlich können wir

$$\int_0^1 \frac{1}{\sqrt{x}} dx = \lim_{y \downarrow 0} \int_y^1 \frac{1}{\sqrt{x}} dx = \lim_{y \downarrow 0} (2 - 2\sqrt{y}) = 2$$

definieren. Die allgemeinere Definition ist die folgende.

Definition 9.14. *Sei f auf $[a;b)$ definiert ($b=+\infty$ ist zugelassen), und auf $[a;y]$ integrierbar, für alle $y \in (a;b)$. Existiert der Limes*

$$\lim_{y \uparrow b} \int_a^y f dx$$

dann sagen wir, f sei auf $[a;b)$ uneigentlich integrierbar, und wir definieren das uneigentliche Integral von f auf $[a;b)$ durch

$$\int_a^b f dx := \lim_{y \uparrow b} \int_a^y f dx$$

Ähnlich, falls f auf $(a;b]$ ($a=-\infty$ zugelassen) definiert ist, auf $[y;b]$ integrierbar ist, für alle $y \in (a;b)$, und falls der Limes

$$\lim_{y \downarrow a} \int_y^b f dx$$

existiert, so definieren wir

$$\int_a^b f dx := \lim_{y \downarrow a} \int_y^b f dx$$

Ist f auf $(a;b)$ definiert ($a=-\infty$ und/oder $b=+\infty$ sind zugelassen) und falls die uneigentlichen Integrale

$$\int_a^c f dx \quad \textit{und} \quad \int_c^b f dx$$

für ein $c \in [a;b]$ existieren, so definieren wir

$$\int_a^b f dx = \int_a^c f dx + \int_c^b f dx$$

Bemerkung: Ist das Integral auf beiden Seiten uneigentlich, so müssen die zwei Grenzwerte $y \downarrow a$ und $y \uparrow b$ unabhängig voneinander genommen werden. Z.B. das Integral

$$\int_{-\infty}^{\infty} x dx$$

existiert nicht, obwohl

$$\int_{-y}^{y} x dx = 0$$

für alle $y > 0$.

Beispiel: Es gilt

$$\int_1^\infty \frac{1}{x^\alpha} dx = \lim_{y \to \infty} \int_1^y \frac{1}{x^\alpha} dx = \lim_{y \to \infty} \frac{-1}{\alpha - 1} \frac{1}{y^{\alpha-1}} + \frac{1}{\alpha-1} = \frac{1}{\alpha - 1}$$

für alle $\alpha > 1$. Das Integral

$$\int_1^\infty \frac{1}{x^\alpha} dx$$

existiert dagegen nicht, für $\alpha < 1$. Es gilt weiter

$$\int_0^1 \frac{1}{x^\alpha} dx = \lim_{y \to 0} \int_y^1 \frac{1}{x^\alpha} = \lim_{y \to 0} \frac{1}{1-\alpha} - \frac{y^{1-\alpha}}{1-\alpha} = \frac{1}{1-\alpha}$$

für alle $\alpha < 1$. Das Integral

$$\int_0^1 \frac{1}{x^\alpha} dx$$

existiert nicht, für $\alpha > 1$. Die Integrale

$$\int_1^\infty \frac{1}{x} dx = \lim_{y \to \infty} \log y = +\infty$$

und

$$\int_0^1 \frac{1}{x} dx = \lim_{y \to 0} - \log y = +\infty$$

existieren nicht.

Proposition 9.15 (Vergleichskriterium). *Seien $a \in \mathbb{R} \cup \{-\infty\}, b \in \mathbb{R} \cup \{+\infty\}$, mit $a < b$. Seien f, g integrierbar auf $(\alpha; \beta)$, für alle $a < \alpha < \beta < b$. Es gelte $g(x) \geq 0$ und $|f(x)| \leq g(x)$ für alle $x \in (a; b)$ und es existiere das (uneigentliche) Integral*

$$\int_a^b g dx$$

Dann existiert auch das (uneigentliche) Integral

$$\int_a^b f dx$$

Beweis. Wir betrachten den Fall, dass f integrierbar auf $[a; y]$, für alle $y \in (a; b)$, ist. Wir möchten zeigen, dass

$$\lim_{y \uparrow b} \int_a^y f dx$$

existiert. Sei y_n eine beliebige Folge, mit $y_n < b$ für alle $n \in \mathbb{N}$ und $y_n \to b$ als $n \to \infty$. Wir zeigen, dass

$$\int_a^{y_n} f dx$$

eine Cauchy-Folge ist. Dazu bemerken wir, dass (unter Annahme, dass, zB., $y_m < y_n$)

$$\left| \int_a^{y_n} f dx - \int_a^{y_m} f dx \right| = \left| \int_{y_m}^{y_n} f dx \right| \leq \int_{y_m}^{y_n} |f| dx \leq \int_{y_m}^{y_n} g dx \leq= \int_a^{y_n} g dx - \int_a^{y_m} g dx$$

Die Existenz des Integrals von g auf $[a; b]$ impliziert, dass die Folge $\int_a^{y_n} g dx$ eine Cauchy-Folge ist. Damit ist auch $\int_a^{y_n} f dx$ eine Cauchy-Folge. Also konvergiert die Folge $\int_a^{y_n} f dx$. Es ist weiter einfach zu sehen, dass der Limes unabhängig von der Wahl der Folge y_n ist; man nimmt dazu an, es existieren zwei Folgen $y_{1,n} \to b$ und $y_{2,n} \to b$, so dass

$$\lim_{n \to \infty} \int_a^{y_{1,n}} f dx \neq \lim_{n \to \infty} \int_a^{y_{2,n}} f dx$$

Dann man definiert die Folge $y_{3,n}$, die, alternierend, Werte aus $y_{1,n}$ und $y_{2,n}$ annimmt. Die Folge

$$\lim_{n\to\infty} \int_a^{y_{3,n}} f dx$$

sollte dann zwei Häufungspunkten haben; da aber $y_{3,n} \to b$, muss die Folge $\int_a^{y_{3,n}} f dx$ konvergieren. $\square$

Anwendung: das Integral

$$\int_1^\infty \frac{(\log x)^m}{x^\alpha} dx$$

existiert, für alle $\alpha > 1$, und alle $m > 0$. Das folgt aus der Tatsache, dass, für alle $\varepsilon > 0$ es existiert eine Konstante C_ε mit

$$(\log x) \leq C_\varepsilon x^\varepsilon$$

Für $0 < \varepsilon < \alpha - 1$ gilt also

$$\frac{(\log x)^m}{x^\alpha} \leq \frac{1}{x^{1+\varepsilon}} \sup_{x\geq 1} \frac{(\log x)^m}{x^{\alpha-1-\varepsilon}} \leq \frac{C_\varepsilon}{x^{1+\varepsilon}}$$

Die Existenz des Integrals von $1/x^{1+\varepsilon}$ auf $[1;\infty)$ impliziert also die Existenz des Integrals von $(\log x)^m/x^\alpha$.

Uneigentliche Integrale können auch benutzt werden, um die Konvergenz von Reihen zu prüfen.

Proposition 9.16 (Integralkriterium für Reihen)**.** *Sei f positiv, monoton fallend auf $[1;\infty)$. Dann ist die Reihe $\sum_{n=1}^\infty f(n)$ konvergent, genau dann wenn $\int_1^\infty f dx$ existiert.*

Beweis. Da f positiv ist die Reihe $\sum_n f(n)$ konvergent, genau dann wenn sie beschränkt ist. Aus dem selben Grund, das Integral $\int_1^\infty f dx$ existiert, genau dann wenn die Folge $\int_1^m f dx$ beschränkt ist. Aus der Monotonie von f gilt

$$\int_n^{n+1} f dx \leq f(n) \leq \int_{n-1}^n f dx$$

für alle $n \in \mathbb{N}$. Also

$$\int_2^{m+1} f dx = \sum_{n=2}^m \int_n^{n+1} f dx \leq \sum_{n=2}^m f(n) \leq \sum_{n=2}^m \int_{n-1}^n f dx = \int_1^{m-1} f dx$$

für alle $m \in \mathbb{N}$, $m > 2$. Also die Reihe ist beschränkt, genau dann wenn das Integral beschränkt ist. $\square$

Beispiel: konvergiert die Reihe

$$\sum_{n=2}^\infty \frac{1}{n \log n}\,?$$

Die Funktion $f(x) = (x \log x)^{-1}$ ist positiv und monoton fallend auf $x > 1$. D.h. die Reihe konvergiert g.d.w. das uneigentliche Integral von f auf $[2; \infty)$ existiert. Da (mit der Substitution $u = \log x$)

$$\int_2^y \frac{1}{x \log x} dx = \int_{\log 2}^{\log y} \frac{du}{u} = \log \log y - \log \log 2$$

divergiert, als $y \to \infty$, es folgt, dass

$$\sum_{n \geq 2} \frac{1}{n \log n} = \infty$$

Ähnlicherweise kann man zeigen, dass

$$\sum_{n \geq 2} \frac{1}{n (\log n)^{\alpha}}$$

konvergiert, g.d.w. $\alpha > 1$.