Analysis1

Prof. M. Griebel

Wintersemester 2000/2001

Inhaltsverzeichnis

1 Zahlen

1.1 Die natürlichen Zahlen $\mathbb{N}$

1.1.1 Zahldarstellungen

Es gibt viele verschiedene Darstellungen natürlicher Zahlen. Jedoch sind nicht alle auch gleich gut zum Rechnen geeignet. Historisch sind vor allem folgende Zahldarstellungen von Bedeutung:

- Die Strichzahlen, bei denen die Anzahl der Striche der natürlichen Zahl entspricht, sind wenig nützlich zur Darstellung größerer Zahlen und zum Rechnen. Bsp: $|||$ entspricht der Zahl drei.

- Die antiken Römer verwendeten die Symbole I, V, X, L, C, D und M, um auch große Zahlen darstellen zu zu können. So bezeichnet z.B. MDCLXVII die gleiche Zahl wie 1667 im Dezimalsystem. Allerdings ist auch diese Zahldarstellung kaum zum Rechnen geeignet

- In Positionszahldarstellungen repräsentiert eine k- stellige Zahl, geschrieben als $s_{k-1}s_{k-2}\dots s_0$, die Summe $\sum_{i=0}^{k-1} b^i$, wobei b die Basis des Positionszahlsystems ist und die Faktoren $s_{k-1}, s_{k-2}, \dots s_0$ aus einem Vorrat von b Ziffern sind.
  Von Mayas und Kelten wurde die Basis 20 verwendet; die Sumerer rechneten mit der Basis 60. Relikte dieser Konvention finden sich noch heute in der Zeit- und der Winkelmessung.
  Im alltäglichen Gebrauch ist heute die Basis 10 (Dezimalsystem) üblich; in der Informationstechnik werden vor allem die Basen 2 (Binärdarstellung) und 16 (Hexadezimaldarstellung) verwendet.

1.1.2 Das Axiomensystem von Peano

Definition 1.1 (Axiome der natürlichen Zahlen nach Peano[1]).
Sei $\mathbb{N}$ eine Menge und $n \mapsto n'$ eine Abbildung von $\mathbb{N}$ auf sich mit den Eigenschaften

(1) Die Abbildung $n \mapsto n'$ ist injektiv, d.h. $\quad n' = m' \quad \Rightarrow \quad n = m.$

(2) Es gibt ein ausgezeichnetes Element $1 \in \mathbb{N}$, für das gilt:

(2.1) $n' \neq 1 \quad \forall\ n \in \mathbb{N}$

[1] *Giuseppe Peano* (1858-1932)

(2.2) (Axiom der vollständigen Induktion):
Ist τ eine Teilmenge von $\mathbb{N}$ mit

$$1 \in \tau \qquad \text{und}$$
$$n \in \tau \quad \Rightarrow \quad n' \in \tau,$$

so ist $\tau = \mathbb{N}$.

Die Elemente einer solchen Menge $\mathbb{N}$ heißen **natürliche Zahlen**. Die Zahl n' ist der **Nachfolger** der Zahl n. Diese wiederum heißt **Vorgänger** von n'.

Bemerkung. Die Menge $\mathbb{N}$ der natürlichen Zahlen ist durch obiges Axiomensystem eindeutig bestimmt (bis auf Isomorphie).

1.1.3 Operationen mit natürlichen Zahlen

Die Addition natürlicher Zahlen ist eine Abbildung $\mathbb{N} \times \mathbb{N} \to \mathbb{N}$, definert durch

$$m + 1 = m'$$

$$m + (n') = (m + n)'$$

Auch die Multiplikation natürlicher Zahlen ist eine Abbildung $\mathbb{N} \times \mathbb{N} \to \mathbb{N}$.
Sie wird definiert durch

$$m \cdot 1 = m$$

$$m \cdot (n') = m \cdot n + m$$

Bemerkung. Aus diesen Definitionen lassen sich Rechengesetze ableiten *(Übungsaufgabe 7)*:

- Das Assoziativgesetz der Addition und Multiplikation:
  $(k + m) + n = k + (m + n) \quad \text{und} \quad (km)n = k(mn) \quad \forall\, k, m, n \in \mathbb{N}$
- Das Kommutativgesetz der Addition und Multiplikation:
  $m + n = n + m \quad \text{und} \quad mn = nm \quad \forall\, k, m, n \in \mathbb{N}$
- Das Distributivgesetz: $(k + m)n = kn + mn \quad \forall\, k, m, n \in \mathbb{N}$

Definition 1.2. Sei $n \in \mathbb{N}$. Die Menge $\{m \in \mathbb{N} \mid m \leq n\}$ heißt
n-ter Abschnitt von $\mathbb{N}$.
Eine Menge heißt **endlich**, wenn sie leer ist oder sie sich bijektiv (umkehrbar eindeutig) auf einen Abschnitt von $\mathbb{N}$ abbilden läßt.
Jede nicht endliche Menge ist **unendlich**.
Eine Menge, welche sich bijektiv auf die Menge $\mathbb{N}$ der natürlichen Zahlen abbilden läßt, heißt **abzählbar unendlich**.

CANTOR[2] bemerkte, daß die These, ein Teil könne nicht so groß sein, wie das Ganze, für unendliche Mengen aufgegeben werden muß.

Dies illustriert HILBERTS[3] Hotel: In einem Hotel mit abzählbar unendlich vielen Zimmern seien alle Zimmer belegt. Kommt nun ein neuer Gast, so erhält er dennoch ein Zimmer; jeder der Gäste zieht ein Zimmer weiter und der neue Gast bekommt das erste.
Bei k neuen Gästen zieht jeder Gast k Zimmer weiter und die neuen Gäste beziehen die ersten k Zimmer.
Wollen abzählbar unendlich viele Gäste ein Zimmer in dem voll belegten Hotel, so zieht der Gast mit der Zimmernummer n in das $2n$-te Zimmer und die neuen Gäste erhalten die Zimmer mit ungeraden Nummern.

CANTOR stellte weiterhin fest, daß es Mengen gibt, die noch größer als $\mathbb{N}$ sind. So ist zum Beispiel die Menge aller Teilmengen von $\mathbb{N}$ überabzählbar.

1.1.4 Ordnungsrelation

Definition 1.3. Seien $m, n \in \mathbb{N}$. Hat die Gleichung $m + x = n$ eine Lösung $x \in \mathbb{N}$, so heißt m **kleiner** als n: $m < n$.
Es ist $m \leq n$ genau dann, wenn $m < n$ oder $m = n$.

Eigenschaften dieser Ordnungsrelation sind

1. Für je zwei Zahlen $m, n \in \mathbb{N}$ gilt genau eine der drei Relationen
   $m < n, \qquad m = n \quad \text{oder} \qquad m > n$
2. Aus $m < n$ und $n < k$ folgt $m < k$ (Transitivität)
3. $m < n \quad \Rightarrow \quad m + k < n + k$
4. $m < n \quad \Rightarrow \quad mk < nk$

Satz 1.1 (Wohlordnung der natürlichen Zahlen). *Jede nichtleere Menge natürlicher Zahlen enthält ein kleinstes Element.*

1.1.5 Vollständige Induktion

Das Beweisprinzip der vollständigen Induktion folgt aus den Axiomen der natürlichen Zahlen.

[2] *Georg Cantor* (1845-1918)

[3] nach *David Hilbert* (1862-1943)

Satz 1.2 (Induktionsbeweis). *Zu jedem $n \in \mathbb{N}$ sei die Aussage $A(n)$ gegeben. Alle Aussagen $A(n)$ sind wahr, wenn gilt*

(1) $A(1)$ ist wahr. (Induktionsanfang)

(2) $A(n+1)$ ist wahr, falls $A(n)$ wahr ist. (Induktionsschritt)

Beispiel.

$$1 + 2 + 3 + \ldots + n = \sum_{i=1}^{n} i = \frac{n(n+1)}{2}$$

Beweis. Induktionsanfang: Für $n = 1$ ist die Behauptung offenbar wahr.
Induktionsschritt: Die Behauptung sei wahr für n, d.h. $\sum_{i=1}^{n} i = \frac{n(n+1)}{2}$.
Dann ist $\sum_{i=1}^{n+1} i = \sum_{i=1}^{n} i + (n+1) = \frac{n(n+1)}{2} + (n+1) = \frac{(n+1)(n+2)}{2}$. □

Bemerkung. GAUß[4] hatte als Kind die Idee, die Zahlen von 1 bis 100 folgendermaßen aufzusummieren:
$1 + 2 + \ldots + 100 = (1 + 100) + (2 + 99) + \ldots + (50 + 51) = 50 \cdot 101 = 5050.$

Beispiel. Für $x \neq 1$ gilt:

$$1 + x + x^2 + x^3 + \ldots + x^n = \sum_{i=0}^{n} x^i = \frac{1 - x^{n+1}}{1 - x}$$

Beweis. Induktionsanfang: Für $n = 1$ ist die Behauptung offenbar wahr.
Induktionsschritt: Die Behauptung sei wahr für n. Dann gilt
$\sum_{i=0}^{n+1} x^i = \sum_{i=0}^{n} x^i + x^{n+1} = \frac{1-x^{n+1}}{1-x} + x^{n+1} = \frac{1-x^{n+1}+x^{n+1}-x^{n+2}}{1-x} = \frac{1-x^{n+2}}{1-x}$. □

Bemerkung. Das Prinzip der vollständigen Induktion läßt sich noch verallgemeinern. Die Aussage $A(n)$ ist wahr für alle $n \geq n_1$, falls $A(n_1)$ wahr ist und $A(n+1)$ aus $A(n)$ für alle $n \geq n_1$ folgt.

Definition 1.4 (Rekursion). Eine Abbildung $f : \mathbb{N} \to X$ ist **rekursiv definiert** durch

(1) Angabe von $f(1)$

(2) Eine Vorschrift, die es erlaubt, $f(n+1)$ aus $f(1), f(2), \ldots f(n)$ zu gewinnen.

Beispiel. Die n-te Potenz m^n einer Zahl m wird definiert durch

(1) $m^1 = m$

(2) $m^{n+1} = m \cdot m^n$

Analog zu diesem Beispiel kann die Multiplikation auf die Addition und diese wiederum auf die Nachfolgeoperation $n \mapsto n'$ zurückgeführt werden.

[4] *Carl Friedrich Gauß* (1777-1855)

1.1.6 Fakultät und Binomialkoeffizient

Definition 1.5. Die **Fakultät** einer natürlichen Zahl n ist $n! := 1 \cdot 2 \cdot 3 \ldots n$.
Die rekursive Definition lautet
$1! := 1, \quad (n+1)! = (n+1)n!$

Bemerkung. Die Definition der Fakultät wird erweitert auf 0 durch $0! := 1$.

Definition 1.6. Seien $k, n \in \mathbb{N}$. Der Binomialkoeffizient $\binom{n}{k}$ ist

$$\binom{n}{k} := \frac{n!}{k!(n-k)!} = \frac{(n-k+1)\cdot(n-k+2)\cdot\ldots\cdot n}{k!}$$

Satz 1.3. *Die Anzahl aller möglichen Anordnungen von n paarweise verschiedenen Elementen ist $n!$.*

Beispiel. Die Zahlen $1, 2$ lassen sich auf $2! = 2$ Arten anordnen:
$(1,2)$ und $(2,1)$.
Die Zahlen $1, 2, 3$ lassen sich auf $3! = 6$ Arten anordnen:
$(1,2,3), (1,3,2), (2,1,3), (2,3,1), (3,1,2), (3,2,1)$.

Satz 1.4. *Die Anzahl der k- elementigen Teilmengen einer nichtleeren Menge mit n Elementen ist $\binom{n}{k}$.*

Beweis: Übungsaufgabe 8.

Beispiel. Im Lotto „6 aus 49“ sind $\binom{49}{6} = 13983816$ verschiedene Tipps möglich.

Satz 1.5 (Binomialentwicklung). *Für alle $n \in \mathbb{N}$ gilt*

$$(1+x)^n = \binom{n}{0} + \binom{n}{1}x + \binom{n}{2}x^2 + \ldots + \binom{n}{n}x^n = \sum_{i=0}^{n} \binom{n}{i} x^i$$

Bemerkung. Es gilt

$$\binom{n}{k} = \binom{n}{n-k}$$

sowie die Rekursionsformel

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

Beweis der Rekursionsformel. $\binom{n}{k} + \binom{n}{k+1} = \frac{n(n-1)\cdots(n-k+1)}{k!} + \frac{n(n-1)\cdots(n-k)}{k!(k+1)}$
$= \frac{(k+1+n-k)n(n-1)\cdots(n-k+1)}{(k+1)!} = \frac{(n+1)n\cdots(n-k+1)}{(k+1)!} = \binom{n+1}{k+1}$ $\square$

Die Rekursionsformel für Binomialkoeffizienten veranschaulicht das PASCALsche Dreieck[5]. Dabei ist die $(k+1)$-te Zahl der $(n+1)$-ten Zeile der Wert des Binomialkoeffizienten $\binom{n}{k}$. Gemäß der Rekursionsformel ergibt sich jede Zahl im Inneren als Summe der beiden jeweils darüberstehenden.

$$\begin{array}{ccccccccc} & & & & 1 & & & & \\ & & & 1 & & 1 & & & \\ & & 1 & & 2 & & 1 & & \\ & 1 & & 3 & & 3 & & 1 & \\ 1 & & 4 & & 6 & & 4 & & 1 \\ & & & & \vdots & & & & \end{array}$$

1.2 Die ganzen Zahlen $\mathbb{Z}$ und die rationalen Zahlen $\mathbb{Q}$

1.2.1 Die ganzen Zahlen $\mathbb{Z}$

Man betrachte die Menge $\mathbb{N} \times \mathbb{N}$ aller Paare natürlicher Zahlen und definiere auf dieser eine Äquivalenzrelation $\sim$ derart, daß
$(a,b)\sim(c,d) \qquad \Leftrightarrow \qquad a+d=b+c.$

Definition 1.7. Die Äquivalenzklassen dieser Relation heißen **ganze Zahlen**.
Die Menge der ganzen Zahlen wird mit $\mathbb{Z}$ bezeichnet.
Für die ganze Zahl $[(a,b)]$ schreibt man $a-b$. Die Zahl $[a,a]$ wird mit 0 bezeichnet. Anstelle von $0-a$ schreibt man auch $-a$.

Die Operationen mit natürlichen Zahlen lassen sich auf die ganze Zahlen erweitern. Auch die Ordnungsrelation läßt sich übertragen.
Zusätzlich ist die Subtraktion, Umkehroperation der Addition, stets erklärt:
Für $m,n \in \mathbb{Z}$ hat die Gleichung $m+x=n$ stets eine Lösung $x \in \mathbb{Z}$.

Satz 1.6. *Sind $a,b \in \mathbb{Z}$, so auch $a+b$ und $a \cdot b$.*

Satz 1.7. *Die Menge der ganzen Zahlen ist abzählbar.*

Beweisidee. Die Abbildung $f: \mathbb{N} \to \mathbb{Z}$ mit $f(n) = \begin{cases} \frac{n}{2} & \text{für } n \text{ gerade} \\ \frac{1-n}{2} & \text{für } n \text{ ungerade} \end{cases}$
ist eine Bijektion von $\mathbb{N}$ auf $\mathbb{Z}$, womit beide Mengen gleichmächtig sind. □

Bemerkung. Obgleich $\mathbb{N}$ und $\mathbb{Z}$ gleichmächtig sind, ist die Menge der natürlichen Zahlen eine echte Teilmenge der ganzen Zahlen.

[5]untersucht von *Blaise Pascal* (1623-1662)

1.2.2 Die rationalen Zahlen $\mathbb{Q}$

Man betrachte nun die Äquivalenzrelation $\sim$ auf $\mathbb{Z} \times \mathbb{Z}\backslash\{0\}$ mit
$(a, b) \sim (c, d) \qquad \Leftrightarrow \qquad ad = bc.$
Die Äquivalenzklassen dieser Relation sind die rationalen Zahlen. Das Paar $(a, b) \in \mathbb{Z} \times \mathbb{Z}\backslash\{0\}$ wird geschrieben als Bruch $\frac{a}{b}$. Statt $\frac{a}{1}$ schreibt man a.
Um eine rationale Zahl eindeutig als Bruch darstellen zu können, wird der Repräsentant der Klasse, bei dem Zähler und Nenner teilerfremd sind und der Nenner positiv ist, für die Darstellung verwendet.
Die rationalen Zahlen sind abgeschlossen unter den 4 Grundoperationen, die in bekannter Weise erklärt werden.

Satz 1.8. *Die Menge der rationalen Zahlen ist abzählbar.*

Beweis. Die Menge der natürlichen Zahlen ist eine echte Teilmenge der rationalen Zahlen, daher ist $\mathbb{Q}$ mindestens so mächtig, wie $\mathbb{N}$.
Für die umgekehrte Relation konstruiere man ein Abzählverfahren, welches jeder rationalen Zahl eine natürliche zuordnet. Dazu betrachte man die folgende Tabelle aller positiven Brüche:

$$\begin{array}{ccccccccc}
\frac{1}{1} & & \frac{2}{1} & \rightarrow & \frac{3}{1} & & \frac{4}{1} & \cdots \\
\downarrow & \nearrow & & \swarrow & & \nearrow & & \\
\frac{1}{2} & & \frac{2}{2} & & \frac{3}{2} & & & \\
 & \swarrow & & \nearrow & & & & \\
\frac{1}{3} & & \frac{2}{3} & & & & & \\
\downarrow & \nearrow & & & & & & \\
\frac{1}{4} & & & & & & & \\
\vdots & & & & & & &
\end{array}$$

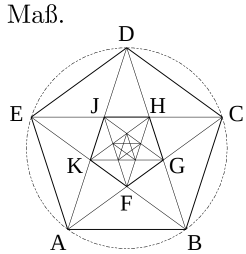
Alle Brüche einer Zeile haben den gleichen Nenner; Brüche dersselben Spalte haben den gleichen Zähler.
Die Tabelle wird in skizzierter Reihenfolge durchlaufen, wobei die Brüche, deren Zähler und Nenner teilerfremd sind, in dieser Reihenfolge aufgeschrieben werden: $\left(1, \frac{1}{2}, 2, 3, \frac{1}{3}, \frac{1}{4}, \frac{2}{3}, \frac{3}{2}, 4, \ldots\right)$.
Damit erhält man eine Bijektion $h : \mathbb{N} \rightarrow \mathbb{Q}^+$.
Setzt man $g(n) = h(n)$, sowie $g(-n) = -h(n)$ für $n \in \mathbb{N}$ und $g(0) = 0$, dann ist $g : \mathbb{Z} \rightarrow \mathbb{Q}$ auch eine bijektive Abbildung. $\mathbb{Z}$ ist abzählbar, d.h. es gibt eine Bijektion $f : \mathbb{N} \rightarrow \mathbb{Z}$. Somit ist $g \circ f : \mathbb{N} \rightarrow \mathbb{Q}$ eine Abzählung von $\mathbb{Q}$. $\square$

1.3 Die reellen Zahlen $\mathbb{R}$

1.3.1 Die Entdeckung der Existenz irrationaler Zahlen

Die Pythagoräer waren der Ansicht, alle Größen ließen sich in ganzen Zahlen ausdrücken; bis einer ihrer Schüler, Hippasus von Metapont, im 5. Jhd. v. Chr. inkommensurable Strecken, d.h. solche ohne gemeinsames Maß, im Ordenssymbol der Pytagoräer, dem Pentagramm, entdeckte.

Eine Strecke wird gemessen, indem man ein Maß der Länge e n-mal auf die Strecke legt, die dann $n \cdot e$ lang ist.
Zwei Strecken a und b haben dann ein gemeinsames Maß e, wenn es natürliche m und n gibt, sodaß $a = me$ und $b = ne$, also $\frac{m}{n}$ rational ist.
Ein gemeinsames Maß zweier kommensurabler Strecken a und b läßt sich mit dem Euklidischen Algorithmus[6] finden:
Die kleinere Strecke b wird sooft wie möglich auf a abgetragen; dabei bleibt ein Rest r, d.h. $a = b \cdot q + r$ mit $r < b$ für ein $q \in \mathbb{N}$.
Nun wird das Verfahren mit b und r fortgesetzt, d.h. r wird auf b abgetragen, wobei wieder ein Rest $\bar{r}$ bleiben kann.
Sind a und b kommensurabel, so muß das Verfahren in endlich vielen Schritten zum Ende kommen. Der letzte Rest ungleich 0 ist dann das größte gemeinsame Maß.



Im regulären Fünfeck (siehe Figur) beträgt der Zentriwinkel über einer Seite $\frac{360°}{5} = 72°$. Dementsprechend sind die Innenwinkel $180° - 72° = 108°$. Der obere Peripheriewinkel über einer Sehne ist halb so groß wie der Zentriwinkel über der gleichen Sehne; daher ist z.B. $\angle ADB = \frac{1}{2}72° = 36°$, $\angle BDC = 36°$ und $\angle ECD = 36°$. Somit gilt $\angle DJC = 180° - 3 \cdot 36° = 72° = \angle ADC$, d.h. $\triangle DJC$ ist gleichschenklig, also $|\overline{CJ}| = a_1$, wenn a_1 die Länge einer Seite ist. Analoges gilt für entsprechende Strecken, die von anderen Ecken des regulären Fünfecks ausgehen.
Sei a_0 die Länge der Diagonalen. Damit wird $|\overline{EJ}| = a_0 - a_1 := a_2$.
Es ist $\angle EJD = 180° - \angle DJC = 108°$, d.h. $\triangle DEJ$ und $\triangle ECD$ sind beide gleichschenklig, wobei die beiden gleichlangen Seiten einen Winkel von 108°

[6] nach *Euklid von Alexandria* (ca. 325 v.Chr. - ca. 265 v.Chr.)

einschließen. Damit sind die beiden Dreiecke ähnlich, woraus $\frac{|\overline{EC}|}{|\overline{ED}|} = \frac{|\overline{ED}|}{|\overline{EJ}|}$, also $\frac{a_0}{a_1} = \frac{a_1}{a_1 - a_0} = \frac{a_1}{a_2}$ folgt.
Die Schnittpunkte der Diagonalen bilden nun wieder ein reguläres Fünfeck $(FGHJK)$, welches die Seitenlänge $|\overline{KJ}| = a_0 - 2a_2 = a_1 - a_2 := a_3$ hat.
Wegen $\frac{|\overline{EH}|}{|\overline{EC}|} = \frac{|\overline{EF}|}{|\overline{EB}|} = \frac{a_1}{a_0}$ ist $\overline{HF}$ parallel zu $\overline{BC}$, woraus mit dem Strahlensatz $\frac{a_0}{a_1} = \frac{|\overline{EC}|}{|\overline{CB}|} = \frac{|\overline{EH}|}{|\overline{HF}|} = \frac{a1}{|\overline{HF}|}$ folgt. Die Länge $|\overline{HF}|$ der Diagonalen des kleinen Fünfeck ist damit $\frac{a_1}{a_0} a_1 = \frac{a_2}{a_1} a_1 = a_2$.
Alle regulären Fünfecke sind ähnlich zueinander; insbesondere ist das Verhältnis von Diagonalen- und Seitenlänge stets gleich. Für die betrachteten Fünfecke $ABCDE$ und $FGHJK$ bedeutet dies $\frac{a_0}{a_1} = \frac{a_1}{a_2} = \frac{a_2}{a_3}$.
Auf $FGHJK$ und das durch seine Diagonalenschnittpunkte entstehende regelmäßige Fünfeck lassen sich obige Überlegungen ebenfalls anwenden.
Allgemein gilt also: Sind a_0, a_1 die Diagonalen- bzw. Seitenlänge eines regelmäßigen Fünfecks und $a_{i+2} = a_i - a_{i+1}$ für $i = 0, 1, 2, \ldots$, so sind alle Verhältnisse $\frac{a_{i+1}}{a_i}$ gleich: $\frac{a_0}{a_1} = \frac{a_1}{a_2} = \frac{a_2}{a_3} = \ldots$
Wendet man den Euklidischen Algorithmus auf a_0 und a_1 and, so erhält man mit $a_0 > a_1$ und $a_0 < 2a_1$ (Dreiecksungleichung für $\triangle ECD$):
$a_0 = 1 \cdot a_1 + (a_0 - a_1) = a_1 + a_2$
$a_1 = 1 \cdot a_2 + (a_1 - a_2) = a_2 + a_3$
$\ldots$
Das Verfahren endet nicht, da alle $\frac{a_{i+1}}{a_i}$ gleich und demzufolge alle Reste $a_i \neq 0$ sind.

Diagonale und Seite eines regelmäßigen Fünfecks sind somit inkommensurabel. Ihr Verhältnis $\frac{a_0}{a_1} = \frac{a_1}{a_1 - a_0}$ ist als „Goldener Schnitt“ bekannt.

1.3.2 Die reellen Zahlen als Dedekindsche Schnitte

Eine weitere nicht rationale Zahl ist $\sqrt{2}$. Sie existiert, da sie die Länge der Diagonalen eines gleichschenkligen, rechtwinkligen Dreiecks mit der Kathetenlänge 1 darstellt.

Beweis, daß $\sqrt{2}$ irrational. Man nehme an, $\sqrt{2}$ sei rational. Dann hat hat sie eine Darstellung als $\frac{p}{q}$ mit $p, q \in \mathbb{N}$, wobei o.B.d.A. p und q als teilerfremd angenommen werden können.
$\Rightarrow \quad 2q^2 = p^2$. Somit ist p^2 und folglich auch p durch 2 teilbar, d.h. $p = 2r$ für ein $r \in \mathbb{N}$.
$\Rightarrow \quad q^2 = 2r^2$, womit auch q^2 durch 2 teilbar ist, was jedoch der Teilerfremdheit von p und q widerspricht. □

Dieses Beispiel zeigt nochmals, daß es Zahlen gibt, die nicht in $\mathbb{Q}$ liegen; die rationalen Zahlen sind nicht vollständig. Dennoch läßt sich bemerken:

Bemerkung. Die rationalen Zahlen liegen dicht, d.h. zu je zwei Zahlen $a, b \in \mathbb{Q}$ mit $a < b$ gibt es stets eine Zahl $c \in \mathbb{Q}$, sodaß $a < c < b$. Zum Beispiel ist $c = \frac{a+b}{2}$ eine solche.

Durch Vervollständigung mittels DEDEKINDscher Schnitte soll nun eine erweiterte Menge von Zahlen konstruiert werden.

Definition 1.8. Sei M eine Menge mit Ordnungsrelation. Ein Paar (α, β) von Teilmengen heißt **Dedekindscher Schnitt**[7] in M, wenn gilt:

(1) Jedes Element von M gehört einer der beiden Teilmengen an.

(2) Keine der beiden Teilmengen ist leer.

(3) Wenn $a \in \alpha, b \in \beta$, so ist $a < b$.

(4) β besitzt kein kleinstes Element.

α heißt Untermenge; β heißt Obermenge.

Definition 1.9. Ein DEDEKINDscher Schnitt in $\mathbb{Q}$ heißt **reelle Zahl**. Die Menge aller DEDEKINDschen Schnitte in $\mathbb{Q}$ ist $\mathbb{R}$.

Bemerkung. Jeder Schnitt (α, β) ist durch eine der beiden Teilmengen eindeutig bestimmt. Im folgenden kann er daher ohne Einschränkung mit der Obermenge β identifiziert werden.

Bemerkung. Die Menge $\mathbb{Q}$ der rationalen Zahlen ist in $\mathbb{R}$ enthalten. Der rationalen Zahl s entspricht der Schnitt mit der Obermenge $\{r \in \mathbb{Q} : \ r > s\}$ und der Untermenge $\{r \in \mathbb{Q} : \ r \leq s\}$.
Ein Schnitt ist genau dann rational, wenn die Untermenge ein größtes Element besitzt.
Nicht alle Schnitte sind rational.

Beispiel. $\sqrt{2}$ entspricht dem Schnitt mit der Obermenge $\beta = \{r \in \mathbb{Q} : \ r^2 > 2 \ \wedge \ r > 0\}$.
Dies ist ein DEDEKINDscher Schnitt in $\mathbb{Q}$, denn die Eigenschaften $(1) - (3)$ aus Definition 1.8 sind offensichtlich erfüllt und für den Nachweis von (4) läßt sich folgende Überlegung anstellen: Sei $r \in \beta$. Dann ist $s := \frac{2r+2}{r+2} > 0$. Außerdem gilt $s^2 - 2 = 2\frac{2(r+1)^2}{(r+2)^2} - 2 = 2\frac{r^2-2}{(r+2)^2} > 0$, da $r^2 > 2$ ist; damit folgt $s^2 > 2$. Schließlich ist $r - s = \frac{r^2+2r-2r-2}{r+2} = \frac{r^2-2}{r+2} > 0$ und folglich $r > s$.

[7] nach *Richard Dedekind* (1831-1916)

Somit gibt es zu jedem $r \in \beta$ noch ein $s \in \beta$, das kleiner ist als r; β hat kein kleinstes Element.

Bemerkung. Die in $\mathbb{Q}$ definierten Operationen $+, \cdot$ und deren Umkehrungen, sowie die Ordnungsrelation lassen sich auf $\mathbb{R}$ übertragen. Es gelten die gleichen Rechengesetze.

Bemerkung. $\mathbb{R}$ ist vollständig im Gegensatz zu $\mathbb{Q}$. Die Untermenge eines DEDEKINDschen Schnittes in $\mathbb{R}$ besitzt stets ein größtes Element. Wird also der Vervollständigungsprozeß auf $\mathbb{R}$ angewant, so liefert er wieder $\mathbb{R}$.

Bemerkung. Die konstruktive Definition von $\mathbb{R}$ ist auch möglich mittels

- Intervallschachtelung
- Fundamentalfolgen (CAUCHY- Folgen)
- Prinzip der oberen Grenze
- Dezimalbruchentwicklung

Satz 1.9. *Die Menge der reellen Zahlen ist nicht abzählbar.*

Beweis. (CANTORsches Diagonalverfahren). Man nehme an, $\mathbb{R}$ sei abzählbar. Dann ist auch das Intervall $(0,1)$ abzählbar, d.h. es existiert eine Folge $x_1, x_2, x_3, \ldots$ reeller Zahlen, sodaß $(0,1) = \{x_n \mid n \in \mathbb{N}\}$. Zu jeder Zahl x_n betrachte man eine Darstellung dieser als Dezimalbruch $0, a_{n1}a_{n2}a_{n3}a_{n4}\ldots$

$x_1 = 0, \boxed{a_{11}}\, a_{12}a_{13}a_{14}\ldots$

$x_2 = 0, a_{21}\boxed{a_{22}}\, a_{23}a_{24}\ldots$

$x_3 = 0, a_{31}a_{32}\boxed{a_{33}}\, a_{34}\ldots$

$x_4 = 0, a_{41}a_{42}a_{43}\boxed{a_{44}}\ldots$

$\vdots$

Nun definiere man eine Zahl $c \in (0,1)$ mit der Dezimalbruchdarstellung

$0, c_1c_2c_3\ldots$ so, daß $\quad c_k = \begin{cases} 4 & \text{falls } a_{kk} = 5 \\ 5 & \text{falls } a_{kk} \neq 5 \end{cases}$

Dann ist $c \neq x_n \ \forall n \in \mathbb{N}$, denn: c und x_n unterscheiden sich in mindestens einer Dezimalstelle, d.h. $\exists k \in \mathbb{N}: \ c_k \neq a_{nk}$. Sei k der kleinste Index mit dieser Eigenschaft. Mit der verallgemeinerten Dreiecksungleichung für unendliche Reihen, die in Abschnitt 2.8 hergeleitet werden soll, und dem Grenzwert der geometrischen Reihe aus 2.7 ist

$|c - x_n| = \left|\sum_{j=k}^{\infty} 10^{-j}(c_j - a_{nj})\right| \geq 10^{-k}|c_k - a_{nk}| - \left|\sum_{j=k+1}^{\infty} 10^{-j}(c_j - a_{nj})\right|$

$\geq 10^{-k} - \sum_{j=k+1}^{\infty} 10^{-j}|c_j - a_{nj}| \geq 10^{-k} - 5\sum_{j=k+1}^{\infty} 10^{-j} = \frac{4}{9}10^{-k} > 0$

Somit kommt c in obiger Aufzählung nicht vor. Folglich ist das Intervall $(0,1)$ und mithin ganz $\mathbb{R}$ nicht abzählbar. $\square$

Bemerkung. Die Darstellung einer Zahl als Dezimalbruch ist im allgemeinen nicht eindeutig. So ist z.B. $0.7 = 0.6999\ldots$

Bemerkung. Jedes nichtleere Intervall (a, b) von $\mathbb{R}$ ist überabzählbar, denn $x \mapsto \frac{x-a}{b-a}$ bildet (a, b) bijektiv auf $(0, 1)$ ab, womit beide Mengen gleichmächtig sind.

Lemma 1.10. $\mathbb{R}\backslash\mathbb{Q}$ *ist nicht abzählbar.*

Beweis. $\mathbb{R} = \mathbb{R}\backslash\mathbb{Q} \cup \mathbb{Q}$. Wäre also $\mathbb{R}\backslash\mathbb{Q}$ abzählbar, dann wäre $\mathbb{R}$ die Vereinigung zweier abzählbarer Mengen und somit selber abzählbar, was nicht der Fall ist. □

1.3.3 Axiomatik der reellen Zahlen

Die reellen Zahlen werden charakterisiert durch

- die Körperaxiome,
- die Anordnungsaxiome
- und das Vollständigkeitsaxiom

Definition 1.10. Eine Menge M bildet zusammen mit zwei Operationen „+“ und „·“ einen **Körper**, wenn für alle $a, b, c \in M$ gilt:

(A1)	$(a + b) + c = a + (b + c)$	(Assoziativgesetz der Addition)
(A2)	$a + b = b + a$	(Kommutativgesetz der Addition)
(A3)	$\exists\ 0 \in M:\ \ a + 0 = a$	(Existenz der Null)
(A4)	$\forall\ a \in M\ \ \exists\ (-a) \in M:\ \ a + (-a) = 0$	(Existenz des additiv Inversen)
(M1)	$(a \cdot b) \cdot c = a \cdot (b \cdot c)$	(Assoziativgesetz der Multiplikation)
(M2)	$a \cdot b = b \cdot a$	(Kommutativgesetz der Multiplikation)
(M3)	$\exists\ 1 \in M:\ \ a \cdot 1 = a$	(Existenz der Eins)
(M4)	$\forall\ a \in M\ \ \exists\ (a^{-1}) \in M:\ \ a \cdot (a^{-1}) = 1$	(Existenz des multiplikativ Inversen)
(D)	$(a + b) \cdot c = a \cdot c + b \cdot c$	(Distributivgesetz)

Bemerkung. Sowohl $\mathbb{Q}$ als auch $\mathbb{R}$ sind Körper.

Definition 1.11. Sei M ein Körper. Es gebe eine Teilmenge $\tau \subset M$, deren Elemente als positiv ausgezeichnet sind: $x > 0\ \forall x \in \tau$. Dann ist M ein **angeordneter Körper**, wenn gilt:

(O1) Für alle $a \in M$ gilt genau eine der drei Beziehungen
$a > 0$, $a = 0$, $-a > 0$.

(O2) $a > 0 \wedge b > 0 \quad \Rightarrow \quad a + b > 0$

(O3) $a > 0 \wedge b > 0 \quad \Rightarrow \quad ab > 0$

Definition 1.12. Ein Körper M heißt **archimedisch angeordnet**, falls er angeordnet ist und zudem $\forall\, a, b \in M$ mit $b > 0 \ \exists\ n \in \mathbb{N}: \quad nb - a > 0$.

Vollständigkeitsaxiom (Supremumseigenschaft).

(V) Jede nichtleere, nach oben beschränkte Teilmenge τ von M besitzt eine kleinste obere Schranke, das Supremum von τ.

Definition 1.13. Eine nichtleere Teilmenge τ von M heißt **nach oben beschränkt**, falls es ein $k \in M$ gibt, sodaß $a \leq k \ \ \forall\, a \in \tau$.
τ heißt **nach unten beschränkt**, falls es ein $k \in M$ gibt, sodaß $a \geq k \ \ \forall\, a \in \tau$.

Definition 1.14. Ein $k \in M$ heißt **kleinste obere Schranke (Supremum)** einer nichtleeren Menge $\tau \subset M$, wenn k obere Schranke von τ ist und es keine kleinere obere Schranke von τ gibt.

Die größte untere Schranke, das Infimum, wird analog definiert.

Bemerkung. Die Menge M, welche die Körper-, die Anordnungs- und das Vollständigkeitsaxiom erfüllt, ist eindeutig bestimmt (bis auf Isomorphie) und heißt Menge $\mathbb{R}$ der reellen Zahlen. $\mathbb{R}$ ist ein archimedisch angeordneter Körper.

Bemerkung. Die Axiome der Dedekindschen Schnitte sind gleichwertig mit dem archimedischen und dem Vollständigkeitsaxiom.

Im Folgenden werden Regeln aufgelistet, die sich aus den Anordnungsaxiomen ableiten lassen. Dabei wird $a > b$ anstelle von $a - b > 0$ geschrieben.

1. $a > b,\ b > c \quad \Rightarrow \quad a > c \qquad$ (Transitivität)
2. $a > b > 0 \quad \Rightarrow \quad \frac{1}{a} < \frac{1}{b}$
3. $a > b \quad \Rightarrow \quad a + c > b + c$
4. $a > b,\ c > 0 \quad \Rightarrow \quad ac > bc$
5. $a > b,\ c < 0 \quad \Rightarrow \quad ac < bc$
6. $a > b,\ \alpha > \beta \quad \Rightarrow \quad a + \alpha > b + \beta$
7. $a > b > 0,\ \alpha > \beta > 0 \quad \Rightarrow \quad a\alpha > b\beta$
8. $a \neq 0 \quad \Rightarrow \quad a^2 > 0$
9. Bernoulli- Ungleichung:[8]

$$x \in \mathbb{R},\ x \geq -1,\ n \in \mathbb{N} \quad \Rightarrow \quad (1 + x)^n \geq 1 + nx$$

(Beweis mittels vollständiger Induktion, vgl. *Übungsaufgabe 21a*)

[8]nach *Jakob Bernoulli* (1654-1705)

Definition 1.15. Sei $a \in \mathbb{R}$. Der **Absolutbetrag** von a ist
$|a| := \begin{cases} a & \text{falls } a \geq 0 \\ -a & \text{falls } a < 0 \end{cases}$

Es gilt:

$$|ab| = |a| \cdot |b|$$

$$|a+b| \leq |a| + |b| \qquad \text{(Dreiecksungleichung)}$$

$$||a| - |b|| \leq |a-b|$$

Satz 1.11. *Sei $b \in \mathbb{R}$ und $b > 1$. Dann existiert zu jedem $k \in \mathbb{R}$ ein $n \in \mathbb{N}$, sodaß $b^n > k$ ist.*

Beweis. $\mathbb{R}$ ist ein archimedisch angeordneter Körper, d.h. zu $b-1$ und $k-1$ existiert ein $n \in \mathbb{N}$ mit $n(b-1) > k-1$.
Aus der Bernoulli- Ungleichung folgt für dieses n:
$b^n = (1 + (b-1))^n \geq 1 + n(b-1) > k$. □

Lemma 1.12. *Ist $0 < b < 1$, dann gilt: $\forall \ \varepsilon > 0 \ \exists n \in \mathbb{N} : \ b^n < \varepsilon$.*

Beweis. $b < 1 \quad \Rightarrow \quad \frac{1}{b} > 1$. Nach Satz 1.11 existiert ein $n \in \mathbb{N}$ mit $\left(\frac{1}{b}\right)^n > \frac{1}{\varepsilon}$.
Für dieses n folgt dann $b^n < \varepsilon$. □

2 Folgen und Reihen

In diesem Abschnitt sollen Grenzprozesse und Grenzwerte untersucht werden. Ein Beispiel dafür ist die Berechnung des Umfanges π eines Kreises mit dem Durchmesser 1. Der Kreisumfang wird dabei durch den Umfang ihm um- und einbeschriebener regelmäßiger Polygone approximiert. Je mehr Ecken diese Polygone haben, um so genauer kann der Wert von π angegeben werden. So erhält man z.B. für 96- Ecke die Abschätzung $3\frac{10}{71} < \pi < 3\frac{1}{7}$.
Ein weiteres Beispiel für Grenzprozesse ist das Babylonische Wurzelziehen: Sei $a > 0$, $x_1 = 1$ und x_{n+1} errechnet sich aus x_n durch $x_{n+1} = \frac{1}{2}\left(x_n + \frac{a}{x_n}\right)$. Die Werte von x_n kommen mit zunehmenden n der $\sqrt{a}$ immer näher. Für $a = 2$ beispielsweise ist $x_2 = 1.5\,, \quad x_3 = 1.41\bar{6}\,, \quad x_4 = 1.414215686...$
Siehe dazu: *Übungsaufgabe 12, Programmieraufgabe 3.*

2.1 Folgen und Grenzwerte

Definition 2.1. Eine Abbildung $\mathbb{N} \to \mathbb{R}$, die jedem $n \in \mathbb{N}$ ein $a_n \in \mathbb{R}$ zuordnet, heißt **Folge** reeller Zahlen (Folge in $\mathbb{R}$). Sie wird bezeichnet mit $(a_n)_{n \in \mathbb{N}} \quad \text{oder} \quad (a_1, a_2, a_3, \ldots)$.

Allgemeiner ist, wenn $n_0 \in \mathbb{Z}$, auch $(a_n)_{n \geq n_0}$ eine Folge, da es eine Bijektion von $\{n \in \mathbb{Z} : \ n \geq n_0\}$ auf $\mathbb{N}$ gibt.
Eine Folge wird häufig durch ein Bildungsgesetz beschrieben.

Beispiel. (1) $a_n = a$, $a \in \mathbb{R}$ liefert die konstante Folge $(a, a, a, \ldots)$.

(2) $a_n = (-1)^n$ stellt eine alternierende Folge $(-1, 1, -1, \ldots)$ dar.

(3) $a_n = \frac{1}{n}$, $(1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots)$

(4) $a_n = \frac{n}{n+1}$, $(1, \frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \ldots)$

(5) $a_n = \frac{n}{2^n}$, $(\frac{1}{2}, \frac{1}{2}, \frac{3}{8}, \frac{1}{4}, \ldots)$

(6) Die Folge $(a_n)_{n \in \mathbb{N}}$, rekursiv definiert durch $a_1 = 1$, $a_2 = 1$, $a_{n+2} = a_{n+1} + a_n$ besteht aus den FIBONACCI[9]- Zahlen: $(1, 1, 2, 3, 5, 8, 13, 21, \ldots)$

(7) $b \in \mathbb{R}$, $a_n = b^{n-1}$: $(1, b, b^2, b^3, \ldots)$

Definition 2.2. Eine Folge $(a_n)_{n \in \mathbb{N}}$ heißt **konvergent** gegen $a \in \mathbb{R}$, falls

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

Die Zahl a ist der **Grenzwert (Limes)** der Folge $(a_n)_{n \in \mathbb{N}}$: $a = \lim (a_n)_{n \in \mathbb{N}}$.
$U_\varepsilon(a) = \{x \in \mathbb{R} : \ |x - a| < \varepsilon\}$ heißt ε**- Umgebung** von a.
In dieser liegen **fast alle**, d.h. alle bis auf endlich viele, Glieder der Folge.

Definition 2.3. Die Folge $(a_n)_{n \in \mathbb{N}}$ ist **divergent**, wenn sie gegen keine reelle Zahl konvergiert.

Satz 2.1. *Der Grenzwert einer konvergenten Folge ist eindeutig bestimmt.*

Beweis. Man nehme an, es gäbe eine Folge $(a_n)_{n \in \mathbb{N}}$ mit zwei Grenzwerten $a \neq a'$. Dann existieren zu $\varepsilon = \frac{|a-a'|}{3}$ Indizes N und N' mit
$|a_n - a| < \varepsilon \quad \forall n \geq N$ und $|a_n - a'| < \varepsilon \quad \forall n \geq N'$.
Für alle $n \geq \max(N, N')$ gilt dann:
$|a - a'| = |a - a_n + a_n - a'| \leq |a_n - a| + |a_n - a'| < \frac{2}{3}|a - a'|$,
was nicht sein kann. □

Nun sollen die Beispiele für Folgen auf Konvergenz untersucht werden.

(1) Die konstante Folge $(a_n)_{n \in \mathbb{N}}$ mit $a_n = a$ hat den Grenzwert a, denn für alle $\varepsilon > 0$ ist $|a_n - a| = 0 < \varepsilon \ \forall n \geq 1$.

[9] *Leonardo Pisano (Fibonacci)* (1170-1250)

(2) $(a_n)_{n\in\mathbb{N}}$ mit $a_n = (-1)^n$ ist divergent.

Beweis. Angenommen, (a_n) sei konvergent gegen ein $a \in \mathbb{R}$. Dann gibt es zu $\varepsilon = 1$ ein $N \in \mathbb{N}$, sodaß $|a_n - a| < 1 \ \forall n \geq N$.
$\Rightarrow \quad 2 = |a_{n+1} - a_n| \leq |a_{n+1} - a| + |a_n - a| < 1 + 1 = 2,$
was offenbar falsch ist. □

(3) Für $a_n = \frac{1}{n}$ ist $\lim (a_n) = 0$, denn $|a_n - 0| = \frac{1}{n} < \varepsilon$ für alle $n > \frac{1}{\varepsilon}$.

Definition 2.4. Eine gegen 0 konvergente Folge heißt **Nullfolge**.

(4) Sei $a_n = \frac{n}{n+1}$. Dann ist $\lim (a_n) = 1$, denn $|\frac{n}{n+1} - 1| = \frac{1}{n+1} < \varepsilon \ \forall n \geq \frac{1}{\varepsilon}$.

(5) Mittels vollständiger Induktion läßt sich $n^2 \leq 2^n$ für $n \geq 4$ zeigen.
Daher gilt $\forall \ \varepsilon > 0$ und $\forall n > \max\left(\frac{1}{\varepsilon}, 4\right)$:
$\left|\frac{n}{2^n}\right| = \frac{n}{2^n} \leq \frac{1}{n} < \varepsilon$, d.h. $a_n = \frac{n}{2^n}$ ist eine Nullfolge.

Definition 2.5. Die Folge $(a_n)_{n\in\mathbb{N}}$ heißt **nach oben (unten) beschränkt**, wenn es ein $k \in \mathbb{R}$ gibt, sodaß $a_n \leq k$ (bzw. $a_n \geq k$) $\forall n \in \mathbb{N}$.
$(a_n)_{n\in\mathbb{N}}$ heißt **beschränkt**, wenn sie nach oben und nach unten beschränkt ist, d.h. $\exists M \geq 0 : \ |a_n| \leq M \ \forall n \in \mathbb{N}$.

Satz 2.2. *Jede konvergente Folge ist auch beschränkt.*

Beweis. Sei $(a_n)_{n\in\mathbb{N}}$ eine Folge mit $\lim (a_n) = a$. Für ein $N \in \mathbb{N}$ gilt dann
$|a_n - a| < 1 \ \forall n \geq N \quad \Rightarrow \quad |a_n| < |a| + 1$ für $n \geq N$ und somit
$|a_n| \leq \max (|a_1|, |a_2|, \ldots, |a_{N-1}|, |a| + 1)$ für alle $n \in \mathbb{N}$. □

Bemerkung. Die Umkehrung von Satz 2.2 gilt nicht, z.B. ist $a_n = (-1)^n$ beschränkt, aber divergent.

(6) Die Folge der Fibonacci- Zahlen, definiert durch $a_{n+2} = a_{n+1} + a_n$ und $a_1 = 1 \ \ a_2 = 1$, ist divergent, denn $a_n \geq (n-1) \ \forall n$.

(7) Man betrachte die Folge $(b^n)_{n\in\mathbb{N}}$.
1. Fall: $|b| < 1. \quad \Rightarrow \quad \lim (b^n) = 0$, denn nach Lemma 1.12 existiert ein $N \in \mathbb{N}$ mit $|b|^N < \varepsilon$ und somit ist $|b|^n \leq |b|^N < \varepsilon \ \forall n \geq N$.
2. Fall: $b = 1 \quad \Rightarrow \quad b^n = 1 \quad \Rightarrow \quad \lim (b^n) = 1$
3. Fall: $b = -1 \quad \Rightarrow \quad (b^n)$ divergiert; siehe Beispiel (2).
4. Fall: $|b| > 1 \quad \Rightarrow \quad$ Die Folge (b^n) divergiert, da sie nach Satz 1.11 unbeschränkt ist.

2.2 Rechenregeln für Grenzwerte

Satz 2.3. *Seien* $(a_n)_{n\in\mathbb{N}}$ *und* $(b_n)_{n\in\mathbb{N}}$ *zwei konvergente Folgen mit* $\lim(a_n)_{n\in\mathbb{N}} = a$ *und* $\lim(b_n)_{n\in\mathbb{N}} = b$*. Dann ist auch* $(c_n)_{n\in\mathbb{N}}$ *konvergent mit*

a) $c_n = a_n + b_n$ *und* $\lim(c_n) = a + b$

b) $c_n = a_n - b_n$ *und* $\lim(c_n) = a - b$

c) $c_n = a_n \cdot b_n$ *und* $\lim(c_n) = a \cdot b$

d) $c_n = \frac{a_n}{b_n}$ *und* $\lim(c_n) = \frac{a}{b}$ *für* $b, b_n \neq 0$

e) $c_n = \lambda a_n$ *und* $\lim(c_n) = \lambda a$ *für alle* $\lambda \in \mathbb{R}$.

Beweis. zu a) Sei $\varepsilon > 0$. Dann existieren $N_1, N_2 \in \mathbb{N}$, sodaß $|a_n - a| < \frac{\varepsilon}{2}\ \forall n \geq N_1$ und $|b_n - b| < \frac{\varepsilon}{2}\ \forall n \geq N_2$, da (a_n) und (b_n) konvergent sind.
$\forall n \geq \max(N_1, N2)$ gilt dann
$|(a_n + b_n) - (a + b)| \leq |a_n - a| + |b_n - b| < \varepsilon.$

zu c) (a_n) und (b_n) sind beschränkt, d.h.
$\exists\ K', K'' :\ |a_n| < K',\ |b_n| < K''\ \forall n.$ Sei $K = \max(K', K'')$ und $\varepsilon > 0$ beliebig. Aufgrund der Konvergenz von (a_n) und (b_n) gibt es N_1, N_2 mit $|a_n - a| < \frac{\varepsilon}{2K}\ \forall n \geq N_1$ und $|b_n - b| < \frac{\varepsilon}{2K}\ \forall n \geq N_2$.
Für sämtliche $n \geq \max(N_1, N_2)$ folgt daraus
$|a_n \cdot b_n - ab| = |a_n(b_n - b) + b(a_n - a)| \leq$
$|a_n| \cdot |(b_n - b)| + |b| \cdot |(a_n - a)| < K\frac{\varepsilon}{2K} + K\frac{\varepsilon}{2K} = \varepsilon.$
Dabei gilt die Abschätzung $|b| \leq K$ wegen Satz 2.6.

zu e) Sei (b_n) die konstante Folge mit $b_n = \lambda$. Dann folgt aus b) die Behauptung.

zu b) Nach e) konvergiert $(-b_n)_{n\in\mathbb{N}}$ gegen $-b$ und aus a) folgt damit, daß $(a_n + (-b_n))_{n\in\mathbb{N}}$ gegen $a - b$ konvergiert.

zu d) Zunächst soll die Folge $\left(\frac{1}{b_n}\right)_{n\in\mathbb{N}}$ betrachtet werden.
(b_n) hat den Grenzwert b, daher gibt es ein $N_1 \in \mathbb{N}$ mit
$|b_n - b| < \frac{|b|}{2}\ \forall n \geq N_1 \quad \Rightarrow \quad |b_n| > \frac{|b|}{2}$ für diese n.
Weiterhin existiert ein $N_2 \in \mathbb{N}$ mit $|b_n - b| < \frac{\varepsilon|b|^2}{2}\ \forall n \geq N_2$.
Somit ist $\left|\frac{1}{b_n} - \frac{1}{b}\right| = \left|\frac{b - b_n}{b \cdot b_n}\right| = \left|\frac{1}{b \cdot b_n}\right| |b_n - b| < \frac{2}{|b|^2} \cdot \frac{\varepsilon|b|^2}{2} = \varepsilon$
$\forall n \geq N := max(N_1, N_2)$, d.h. $\lim\left(\frac{1}{b_n}\right) = \frac{1}{b}$.
Mit c) folgt daraus die Konvergenz von $\left(\frac{a_n}{b_n}\right)_{n\in\mathbb{N}}$ gegen $\frac{a}{b}$.

□

Beispiel. $c_n = \frac{3n^2+13n}{n^2-2} = \frac{3+\frac{13}{n}}{1-\frac{2}{n^2}}$.
Es ist $\lim\left(\frac{1}{n}\right) = 0$, woraus $\lim\left(\frac{1}{n^2}\right) = \lim\left(\frac{1}{n}\right) \cdot \lim\left(\frac{1}{n}\right) = 0$ und
$\lim\left(\frac{13}{n}\right) = 13 \cdot \lim\left(\frac{1}{n}\right) = 0$ folgt.
$\Rightarrow \quad \lim\left(3+\frac{13}{n}\right) = 3, \quad \lim\left(1-\frac{1}{n^2}\right) = 1.$
$\Rightarrow \quad \lim(c_n) = 3.$

Bemerkung. Sind λ, $\mu \in \mathbb{R}$ und $a_n \to a$, $b_n \to b$, dann folgt aus Satz 2.3 (a) und e)): $\lambda a_n + \mu b_n \to \lambda a + \mu b$, d.h. die konvergenten Folgen in $\mathbb{R}$ bilden einen Vektorraum über $\mathbb{R}$ und $\lim: \ (a_n)_{n\in\mathbb{N}} \mapsto \lim(a_n)_{n\in\mathbb{N}}$ ist eine lineare Abbildung.

Bemerkung. Die Summe zweier divergenter Folgen kann konvergent sein.

Bemerkung. Das Produkt einer Nullfolge und einer beschränkten Folge ist eine Nullfolge *(Übungsaufgabe 19)*.

Bemerkung. Arithmetische, geometrische und harmonische Mittelbildungen von konvergenten Folgen sind wieder konvergent.

Satz 2.4. *Sind* $(a_n)_{n\in\mathbb{N}}$ *und* $(b_n)_{n\in\mathbb{N}}$ *zwei Folgen mit* $a_n \leq b_n \ \forall n$, *so gilt auch* $\lim(a_n) \leq \lim(b_n)$.

Beweis. Sei $a = \lim(a_n)$ und $b = \lim(b_n)$. Annahme: $a > b$.
Dann gibt es zu $\varepsilon := \frac{a-b}{2}$ Indizes N_1, N_2 mit $|a_n - a| < \varepsilon \ \forall n \geq N_1$ und $|a_n - a| < \varepsilon \ \forall n \geq N_1$.
Somit ist $a - \varepsilon < a_n \leq b_n < b + \varepsilon$ für alle $n \geq \max(N_1, N_2)$, was jedoch $a - \varepsilon = a - \frac{a-b}{2} = \frac{a+b}{2} = b + \frac{a-b}{2} = b + \varepsilon$, also $a - \varepsilon = b + \varepsilon$ widerspricht. $\square$

Bemerkung. Aus $a_n < b_n \ \forall n$ folgt nicht $\lim(a_n) < \lim(b_n)$, wie das Beispiel $a_n = 0$, $b_n = \frac{1}{n}$ zeigt.

Satz 2.5. *Sind* $(a_n)_{n\in\mathbb{N}}$ *und* $(b_n)_{n\in\mathbb{N}}$ *zwei konvergente Folgen mit dem Grenzwert* a *und es gilt* $a_n \leq c_n \leq b_n$ *für fast alle* n, *so konvergiert auch die Folge* $(c_n)_{n\in\mathbb{N}}$ *gegen* a.

Satz 2.6. *Ist* $(a_n)_{n\in\mathbb{N}}$ *konvergent und* $A \leq a_n \leq B$ *für fast alle* n, *so ist auch* $A \leq \lim(a_n) \leq B$.

Beweis von Satz 2.5 und 2.6: Übungsaufgabe 20.

2.3 Bestimmte Divergenz

Definition 2.6. Eine Folge $(a_n)_{n\in\mathbb{N}}$ heißt **bestimmt divergent** (oder **uneigentlich konvergent**) gegen ∞ $(-\infty)$, wenn

$$\forall K \in \mathbb{R} \ \exists N \in \mathbb{N}: \ a_n > K \ (\text{bzw.}\, a_n < K) \ \forall n \geq N$$

Man schreibt: $\lim(a_n) = \infty$ (bzw. $\lim(a_n) = -\infty$)

Beispiel. Die Folge $a_n = n$ divergiert bestimmt gegen ∞.
$(-2^n)_{n\in\mathbb{N}}$ hat den uneigentlichen Grenzwert $-\infty$.
Die Folge $a_n = (-1)^n$ divergiert, jedoch nicht bestimmt.

Bemerkung. $+\infty$ und $-\infty$ sind Symbole, keine reellen Zahlen.
So sind z.B. die Operationen $\infty - \infty$, $0 \cdot \infty$, $\frac{\infty}{\infty}$ nicht definiert. Dies ist schnell einzusehen, wenn man die Folgen $a_n = n$, $b_n = 1$, $c_n = a_n + b_n = n + 1$ betrachtet. Es ist $\lim(a_n) = \infty$, $\lim(b_n) = 1$, $\lim(c_n) = \infty$, woraus, wäre ∞ eine reelle Zahl, $\infty = \infty + 1$ und $0 = 1$ folgen würde.
Jedoch läßt sich $\infty \pm x = \infty$, $\frac{x}{\infty} = 0$, $\forall x \in \mathbb{R}$ und $\infty \cdot x = \infty$ für $x > 0$ formulieren.
In der Nichtstandard-Analysis wird $\mathbb{R}$ kompaktifiziert und $\overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty, -\infty\}$ betrachtet.

Satz 2.7. *Ist $(a_n)_{n\in\mathbb{N}}$ bestimmt divergent gegen $\pm\infty$, so*
$\exists\, n_0 \in \mathbb{N}:\ a_n \neq 0\ \forall\, n \geq n_0$ *und* $\left(\frac{1}{a_n}\right)_{n\geq n_0}$ *ist eine Nullfolge.*

Beweis. Sei $\lim(a_n) = \infty$. Dann existiert ein $n_0 \in \mathbb{N}$, sodaß $a_n > 0\ \forall\, n \geq n_0$.
Überdies $\exists\, N \in \mathbb{N}:\ a_n > \frac{1}{\varepsilon}\ \forall\, n \geq N$. Damit ist $\left|\frac{1}{a_n}\right| < \varepsilon$ für diese n. □

Satz 2.8. *Ist $(a_n)_{n\in\mathbb{N}}$ eine Nullfolge und $a_n > 0\ \forall\, n$, so divergiert $\frac{1}{a_n}$ gegen ∞.*

2.4 Monotone Folgen

Definition 2.7. Eine Folge $(a_n)_{n\in\mathbb{N}}$ heißt **monoton wachsend** (**fallend**), wenn $a_n \leq a_{n+1}$ (bzw. $a_n \geq a_{n+1}$) für alle n gilt.
$(a_n)_{n\in\mathbb{N}}$ ist **streng monoton wachsend** (**fallend**),
falls $a_n < a_{n+1}$ (bzw. $a_n > a_{n+1}$) gilt.

Satz 2.9. *Jede monotone und beschränkte Folge ist konvergent.*

Beweis. Sei $(a_n)_{n\in\mathbb{N}}$ monoton wachsend und beschränkt.
Dann gibt es ein $\gamma = \sup(a_n)$. Dieses hat die Eigenschaft, daß
$\forall\ \varepsilon > 0\ \exists\, n_0:\ a_{n_0} > \gamma - \varepsilon$.
Aus der Monotonie folgt damit $a_n > \gamma - \varepsilon\ \forall\, n \geq n_0$, also
$|a_n - \gamma| < \varepsilon$ für fast alle n.
Der Beweis für monoton fallende Folgen verläuft analog. □

Beispiel. (1) Für $0 \leq x < 1$ ist $a_n = x^n$ monoton fallend und durch 0 nach unten beschränkt, also konvergent.

(2) Ist $a_n = \sum_{k=n}^{2n} \frac{1}{k}$, so gilt $a_{n+1} - a_n = \frac{1}{2n+1} + \frac{1}{2n+2} - \frac{1}{n} < 0$, d.h. auch diese Folge ist monoton fallend und beschränkt und folglich konvergent.

Definition 2.8. Ein Paar $((a_n),(b_n))$ von Folgen heißt **Intervallschachtelung**, wenn (a_n) monoton wächst, (b_n) monoton fällt und $(a_n - b_n)$ eine Nullfolge ist.

Beispiel. Bei der fortgesetzten Halbierung ergibt sich das Intervall $[a_{n+1}, b_{n+1}]$ durch $[a_{n+1}, b_{n+1}] = [a_n, \frac{a_n+b_n}{2}]$ oder $[a_{n+1}, b_{n+1}] = [\frac{a_n+b_n}{2}, b_n]$.
Hier gilt $|a_n - b_n| = \frac{1}{2^{n-1}}|a_1 - b_1| \to 0$.

Für eine Intervallschachtelung, bezeichnet mit $(a_n|b_n)$, ist offensichtlich $a_n \leq b_m \;\; \forall\, n, m$. Es gilt das

Intervallschachtelungsprinzip.

(ISP) Zu jeder Intervallschachtelung $(a_n|b_n)$ in $\mathbb{R}$ gibt es eine Zahl, die in allen Intervallen $[a_n, b_n]$ liegt.

Das Intervallschachtelungsprinzip ist äquivalent zum Vollständigkeitsaxiom.

Beweis. (V) $\Rightarrow$ (ISP): Die Menge $A = \{a_1, a_2, \ldots\}$ ist durch alle b_n nach oben beschränkt. Für $s = \sup A$ gilt also $a_n \leq s \leq b_n \;\; \forall\, n$, d.h. $s \in [a_n, b_n] \;\; \forall\, n$.
(ISP) $\Rightarrow$ (V): Zu einer nichtleeren, nach oben beschränkten Menge M konstruiere man rekursiv eine Intervallschachtelung (a_n, b_n) in der folgenden Weise. Sei b_1 irgend eine obere Schranke von M und a_1 eine reelle Zahl, die keine obere Schranke von M ist. Das Intervall $[a_{n+1}, b_{n+1}]$ ergibt sich aus $[a_n, b_n]$ durch die Vorschrift

$$[a_{n+1}, b_{n+1}] = \begin{cases} [a_n, \frac{a_n+b_n}{2}] & \text{falls } \frac{a_n+b_n}{2} \text{ obere Schranke von } M \text{ ist} \\ [\frac{a_n+b_n}{2}, b_n] & \text{sonst} \end{cases}$$

Dann ist die Zahl s, die allen Intervallen $[a_n, b_n]$ angehört, eine obere Schranke von M, denn andernfalls gäbe es ein $m \in M$, sodaß $m > s$ ist, womit jedes Intervall von einer Länge $< m - s$ die Zahl m nicht enthielte, was der Eigenschaft der b_n, obere Schranken von M zu sein, widerspräche.
s ist auch die kleinste obere Schranke, denn für jede Zahl r mit $r < s$ existiert ein $n \in \mathbb{N}$, sodaß $r \notin [a_n, b_n]$, d.h. $r < a_n$ ist. Nach Konstruktion der Intervallschachtelung jedoch ist kein a_n obere Schranke von M. □

Bemerkung. Die Zahl g, die in allen Intervallen $[a_n, b_n]$ einer Intervallschachtelung $(a_n|b_n)$ liegt, ist eindeutig bestimmt, denn für alle $h \neq g$ liegt h nicht in den Intervallen von einer Länge $< |g - h|$.
g ist der gemeinsame Grenzwert der Folgen $(a_n)_{n\in\mathbb{N}}$ und $(b_n)_{n\in\mathbb{N}}$:

$$g = \lim\,(a_n) = \lim\,(b_n) = \bigcap_{n=1}^{\infty} [a_n, b_n]$$

2.5 Cauchy- Folgen

Satz 2.10 (Cauchy- Kriterium[10]). *Eine Folge* $(a_n)_{n\in\mathbb{N}}$ *ist konvergent* $\Leftrightarrow$

$$\forall\ \varepsilon > 0 \quad \exists N \in \mathbb{N}: \quad |a_m - a_n| < \varepsilon \qquad \forall m, n \geq N \tag{1}$$

Definition 2.9. $(a_n)_{n\in\mathbb{N}}$ mit der Eigenschaft (1) heißt **Cauchy-Folge.**

Bemerkung. Die Richtung „$\Rightarrow$" von von 2.10 bedeutet, daß jede konvergente Folge eine CAUCHY- Folge ist. Die Umkehrung („In $\mathbb{R}$ konvergiert jede CAUCHY- Folge") läßt sich als alternatives Axiom zum Vollständigkeitsaxiom verwenden. Zu einem späteren Zeitpunkt soll gezeigt werden, daß sie auch aus dem Satz von BOLZANO-WEIERSTRAß folgt.

Beweis „$\Rightarrow$" von Satz 2.10. Sei a der Grenzwert der Folge $(a_n)_{n\in\mathbb{N}}$, d.h.
$\forall\ \varepsilon > 0\ \exists\ N \in \mathbb{N}: \quad |a_n - a| < \frac{\varepsilon}{2} \qquad \forall n \geq\ N$
$\Rightarrow \quad |a_n - a_m| = |(a_n - a) + (a - a_m)| \leq |a_n - a| + |a - a_m| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$ ☐

Beispiel. Die Folge $(a_n)_{n\in\mathbb{N}\cup\{0\}}$ mit $a_0 = 1$ und $a_{n+1} = \frac{1}{1+a_n}$ ist konvergent, denn: Offensichtlich gilt $\frac{1}{2} \leq a_n \leq 1 \quad \forall\ n.$ Damit ist für beliebige $n, k \geq 0:$
$|a_{n+k+1} - a_{n+1}| = \left|\frac{1}{1+a_{n+k}} - \frac{1}{1+a_n}\right| = \left|\frac{a_n - a_{n+k}}{(1+a_n)(1+a_{n+k})}\right| = \frac{|a_{n+k}-a_n|}{(1+a_n)(1+a_{n+k})}$
$\leq \frac{1}{\left(1+\frac{1}{2}\right)^2}\ |a_{n+k} - a_n| = \frac{4}{9}\ |a_{n+k} - a_n|$
Nimmt man diese Ungleichung im Fall $n = 0$ als Induktionsanfang und
$|a_{n+k} - a_n| \leq \left(\frac{4}{9}\right)^n |a_k - a_0|$ als Induktionsvoraussetzung, so folgt:
$|a_{n+k+1} - a_{n+1}| \leq \frac{4}{9}|a_{n+k} - a_n| \leq \left(\frac{4}{9}\right)^{n+1} |a_k - a_0|$ und damit
$|a_{n+k} - a_n| \leq \left(\frac{4}{9}\right)^n |a_k - a_0| \leq 2\left(\frac{4}{9}\right)^n \quad \forall\ n.$
Weil nun $\left(2\left(\frac{4}{9}\right)^n\right)_{n\in\mathbb{N}}$ eine Nullfolge ist, wird damit
$|a_{n+k} - a_n| = |a_m - a_n| < \varepsilon$ für beliebig kleine ε und entsprechend große m, n.
Somit ist (a_n) eine CAUCHY- Folge, also konvergent.
Es sei bemerkt, daß der Grenzwert a der rekursiv definierten Folge (a_n) die Fixpunktgleichung erfüllt: $a = \frac{1}{1+a}$.

2.6 Der Satz von Bolzano-Weierstraß

Definition 2.10. Eine Zahl h heißt **Häufungspunkt** (oder: Häufungswert) der Folge $(a_n)_{n\in\mathbb{N}}$, wenn $\forall\ \varepsilon > 0: \ |a_n - h| < \varepsilon$ für unendlich viele n.

Beispiel. • $(a_n) = (-1)^n$ hat die 2 Häufungswerte 1 und -1.

[10] nach *Augustin Louis Cauchy* (1789-1857)

- Ist ein Folge konvergent, so ist der Grenwert wegen dessen Eindeutigkeit der einzige Häufungswert.
- $(a_n) = n$ hat keinen Häufungswert.

Definition 2.11. Sei $(a_n)_{n\in\mathbb{N}}$ eine Folge und $(n_k)_{k\in\mathbb{N}}$ mit $n_i < n_j$ für $i < j$ eine Folge aufsteigender natürlicher Zahlen. Dann heißt
$(a_{n_k})_{k\in\mathbb{N}} = (a_{n_1}, a_{n_2}, \ldots)$ eine **Teilfolge** von $(a_n)_{n\in\mathbb{N}}$.

Folgerung. *Ist $(a_n)_{n\in\mathbb{N}}$ eine konvergente Folge mit $\lim a_n = a$, dann konvergiert auch jede Teilfolge von $(a_n)_{n\in\mathbb{N}}$ gegen a.*

Satz 2.11. *Ist h ein Häufungswert einer Folge $(a_n)_{n\in\mathbb{N}}$, dann existiert eine Teilfolge (a_{n_k}), die gegen h konvergiert.*

Beweis. Es soll gezeigt werden, daß es möglich ist, eine Teilfolge (a_{n_k}) mit der Eigenschaft

$$|a_{n_k} - h| < \frac{1}{k} \qquad \forall\ k \in \mathbb{N} \tag{2}$$

zu konstruieren. Dann ist $|a_{n_k} - h| < \frac{1}{k} \leq \varepsilon \quad \forall\ k \geq \frac{1}{\varepsilon}$ und somit (a_{n_k}) eine Teilfolge, die gegen h konvergiert.
Für den Beweis der Existenz einer solchen Konstruktion wird auf das Prinzip der vollständigen Induktion zurückgegriffen.
Induktionsanfang: Weil h Häufungswert sein sollte, gibt es mindestens ein n_1, sodaß (2) erfüllt ist.
Induktionsschritt: Es seien k Folgenglieder mit der Eigenschaft (2) gefunden. Nun ist h Häufungswert, d.h. $|a_n - h| < \varepsilon := \frac{1}{k+1}$ für unendlich viele n. Diese können nicht alle $< n_k$ sein, da es nur endlich viele natürliche Zahlen $< n_k$ gibt. Also $\exists\ n_{k+1} > n_k$ mit $|a_{k+1} - h| < \frac{1}{k+1}$. □

Satz 2.12 (Bolzano-Weierstraß[11]). *Jede beschränkte Folge besitzt mindestens einen Häufungswert.*

Beweis. Sei $(a_n)_{n\in\mathbb{N}}$ eine beschränkte Folge, d.h. $\exists M_0, N_0$ mit $M_0 < N_0$ und $M_0 \leq a_n \leq N_0 \quad \forall\ n$.
Man nehme das Intervall $[M_0, N_0]$ als Startwert und konstruiere daraus durch fortgesetzte Halbierung eine Intervallschachtelung $(M_k|N_k)$ mit der Eigenschaft, daß $\forall k: \quad M_k \leq a_n \leq N_k$ für unendlich viele a_n. Dies wird möglich, indem der Prozess der Halbierung stets mit einem Intervall, welches unendlich viele a_n enthält, fortgesezt wird.
Es sei h die durch eine solche Intervallschachtelung bestimmte Zahl. Sie ist der

[11] nach *Bernard Bolzano* (1781-1848) und *Karl Weierstraß* (1815-1897)

gemeinsame Grenzwert der Folgen (M_k) und (N_k).
$\Rightarrow \quad \forall\ \varepsilon > 0\ \exists\ K: \quad h - \varepsilon < M_K < N_K < h + \varepsilon$,
d.h. $\quad h - \varepsilon < a_n < h + \varepsilon \quad$ für unendlich viele a_n.
$\Rightarrow$ h ist der gesuchte Häufungswert. $\square$

Folgerung 2.12. *Eine Zahl ist Häufungswert einer Folge dann und nur dann, wenn sie auch Grenzwert einer konvergenten Teilfolge ist.*

Folgerung 2.12 (Alternative Formulierung von Satz 2.12). *Jede beschränkte Folge enthält eine konvergente Teilfolge.*

Nun soll wie angekündigt die Richtung „$\Leftarrow$" von Satz 2.10 (Cauchy-Kriterium), d.h. „Jede Cauchy- Folge ist konvergent" mit Hilfe von Satz 2.12 bewiesen werden.

Beweis. Sei $\varepsilon > 0$ und $|a_n - a_m| < \varepsilon \quad \forall\ n, m \geq N$.
Insbesondere gilt $|a_n - a_N| < \varepsilon \quad \forall\ n \geq N \quad \Rightarrow \quad |a_n| < |a_N| + \varepsilon$
Folglich ist die größere der beiden Zahlen $|a_N| + \varepsilon$ und $\max_{n<N} |a_n|$ eine Schranke für (a_n). Nach dem Satz von Bolzano-Weierstraß hat (a_n) damit mindestens einen Häufungswert h, d.h. $|a_k - h| < \varepsilon$ für unendlich viele k.
Es sei ein solches $k \geq N$ fest gewählt.
$\Rightarrow \quad |a_n - h| \leq |a_n - a_k| + |a_k - h| < 2\varepsilon\ \ \forall n \geq N$
Also ist h auch Grenzwert. $\square$

Zusammenfassung. Bisher wurden folgende Implikationen bewiesen:

(1) Vollständigkeistaxiom
$\Updownarrow$
(2) Intervallschachtelungsprinzip
$\Downarrow$
(3) Satz von Bolzano-Weierstraß
$\Downarrow$
(4) Cauchy- Kriterium

Außerdem läßt sich das Intervallschachtelungsprinzip (2) aus dem Cauchy-Kriterium (4) folgern. Es ergibt sich der Ringschluß (2) $\Rightarrow$ (3) $\Rightarrow$ (4) $\Rightarrow$ (2); damit sind alle vier Aussagen äquivalent. Sie formulieren jeweils die Vollständigkeit der reellen Zahlen.
Die Anordnungsaxiome und die Vollständigkeit von $\mathbb{R}$ zusammen sind wiederum gleichwertig mit der Konstruktion der reellen Zahlen über Dedekindsche Schnitte.
Zur Begründung der letzten Aussage betrachte man zunächst zu einer gegebenen Intervallschachtelung $(a_n|b_n)$ den Dedekindschen Schitt in $\mathbb{Q}$ mit

der Obermenge $\beta = \bigcup_{n\in\mathbb{N}} ((b_n,\infty)\cap\mathbb{Q})$. Er repräsentiert die reelle Zahl, welche allen $[a_n, b_n]$ gemeinsam ist.
Sei andererseits zu einem DEDEKINDschen Schnitt $S = (\alpha, \beta)$ in $\mathbb{Q}$ die Intervallschachtelung $(a_n|b_n)$ gegeben durch $a_1 \in \alpha$, $b_1 \in \beta$ beliebig und
$[a_{n+1}, b_{n+1}] = \begin{cases} \left[a_n, \frac{a_n+b_n}{2}\right] & \text{falls } \frac{a_n+b_n}{2} \in \beta \\ \left[\frac{a_n+b_n}{2}, b_n\right] & \text{falls } \frac{a_n+b_n}{2} \in \alpha \end{cases}$
Dann ist die durch S repräsentierte Zahl in allen $[a_n, b_n]$ enthalten.

Im Folgenden soll noch die oben erwähnte Implikation: CAUCHY-*Kriterium* $\Rightarrow$ *Intervallschachtelungsprinzip* bewiesen werden:

Beweis. Sei $(a_n|b_n)$ eine Intervallschachtelung.
$\Rightarrow \quad \forall\ \varepsilon > 0\ \exists\ N: \quad |b_N - a_N| < \varepsilon$
Wegen $a_m, a_n \in [a_N, b_N]\ \forall\ n, m > N$ ist auch $|a_n - a_m| < \varepsilon$.
$\Rightarrow \quad (a_n)$ ist CAUCHY- Folge. Somit existiert $g := \lim(a_n)$.
(a_n) ist nach Definition der Intervallschachtelung monoton wachsend.
$\Rightarrow \quad a_n \le g\ \ \forall n$. Überdies ist $a_n \le b_n\ \ \forall n$.
$\Rightarrow \quad b_n \ge g \ge a_n\ \ \forall n$, d.h. g liegt in allen Intervallen. □

Satz 2.13 (Cauchyscher Grenzwertsatz). *Ist $(a_n)_{n\in\mathbb{N}}$ eine konvergente Folge mit dem Grenzwert a, so konvergiert auch die Folge*

$$(b_n)_{n\in\mathbb{N}} \quad \textit{mit} \quad b_n := \frac{1}{n}\sum_{i=1}^{n} a_i \qquad \textit{gegen} \quad a.$$

Beweis: Übungsaufgabe 17.

Beispiel. $(a_n) = \left(\frac{1}{n}\right)$ ist eine Nullfolge, daher ist auch $\lim\left(\frac{1}{n}\sum_{i=1}^{n}\frac{1}{i}\right) = 0$; wohingegen $\left(\sum_{i=1}^{n}\frac{1}{i}\right)_{n\in\mathbb{N}}$ nicht beschränkt ist, was im nächsten Abschnitt gezeigt werden soll.

2.7 Reihen und Konvergenz

Definition 2.12. Sei $(a_n)_{n\in\mathbb{N}}$ eine Folge reeller Zahlen. Dann heißt die Folge (s_n) mit $s_n = \sum_{i=1}^{n} a_i$ die **Summenfolge** von (a_n).
Die zugeordnete **unendliche Reihe** ist Ausdruck $\sum_{i=1}^{\infty} a_i$.
s_n sind die **Partialsummen** der Reihe.

Definition 2.13. Die Reihe $\sum_{i=1}^{\infty} a_i$ heißt **konvergent** mit der Summe s, genau dann wenn (s_n) gegen s konvergiert.
Im Falle der Konvergenz schreibt man

$$s = \sum_{i=1}^{\infty} a_i = \lim_{n\to\infty}\sum_{i=1}^{n} a_i$$

Analog zu Folgen können die Indizes auch bei einer natürlichen Zahl $k \neq 1$ beginnen. $\sum_{i=k}^{\infty} a_i$ ist auch eine unendliche Reihe.

Bemerkung. Aus einer gegebenen Summenfolge (s_n) ist die Folge $(a_n)_{n\in\mathbb{N}}$ ihrer Summanden in der folgenden Weise rekonstruierbar:

$a_1 = s_1$

$a_n = s_n - s_{n-1}$

In diesem Sinne sind Folgen und Reihen äquivalent.

Beispiel.

- Die geometrische Reihe $\sum_{i=1}^{\infty} x^i = \frac{1}{1-x}$ konvergiert für $|x| < 1$, denn für die zugeordnete Summenfolge (s_n) gilt:
  $s_n = 1 + x + \ldots + x^n = \frac{1-x^{n+1}}{1-x}$

- Die harmonische Reihe $\sum_{i=1}^{\infty} \frac{1}{i}$ hingegen ist bestimmt divergent gegen ∞, denn die Summenfolge (s_n) ist monoton wachsend und $\forall\ n = 2^k$ ist
  $s_n = 1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{n}$
  $= 1 + \frac{1}{2} + \left(\frac{1}{3} + \frac{1}{4}\right) + \left(\frac{1}{5} + \ldots + \frac{1}{8}\right) + \ldots + \left(\frac{1}{2^{k-1}} + \ldots \frac{1}{2^k}\right)$
  $\geq 1 + \frac{1}{2} + 2\frac{1}{4} + \ldots + 2^{k-1}\frac{1}{2^k} = 1 + \frac{k}{2}$,
  d.h. (s_n) wächst über alle Schranken.

2.8 Konvergenzkriterien für Reihen

Dieser Abschnitt beschäftigt sich mit klassischen Techniken, Folgen auf Konvergenz zu untersuchen. Alle Konvergenzkriterien für Folgen können selbstverständlich auf die Partialsummen von Reihen angewendet werden. Unter Benutzung von Satz 2.10 erhält man

Satz 2.14 (Cauchy- Kriterium für Reihen). *Die Reihe* $\sum_{i=1}^{\infty} a_i$ *konvergiert dann und nur dann, wenn*

$$\forall\ \varepsilon > 0\, \exists\ N \in \mathbb{N}: \ |s_n - s_m| = \left|\sum_{i=m+1}^{n} a_i\right| < \varepsilon\ \forall\ n > m > N$$

Folgerung.

- *Eine Reihe konvergiert nur (höchstens) dann, wenn die Folge ihrer Glieder eine Nullfolge ist.*

- *Die Abänderung endlich vieler Summanden einer Reihe verändert die Konvergenz bzw. Divergenz nicht.*

Satz 2.15 (Leibniz- Kriterium[12]). *Ist* $(a_n)_n$ *eine monoton fallende Nullfolge, so konvergiert die alternierende Reihe* $\sum_{n=0}^{\infty} (-1)^n a_n$.

[12] nach *Gottfried Wilhelm Leibniz* (1646-1716)

Beweis. Seien $s_k = \sum_{n=0}^{k} (-1)^n a_n$ die Partialsummen der Reihe $\sum_{n=0}^{\infty} (-1)^n a_n$. Die Summenfolge wird zunächst in zwei Teilfolgen $(s_{2k})_k$ und $(s_{2k+1})_k$ zerlegt.
Aufgrund der Monotonie von (a_n) ist
$s_{2k+2} - s_{2k} = -a_{2k+1} + a_{2k+2} \leq 0 \quad \forall k$,
d.h. die Teilfolge $(s_{2k})_k$ ist monoton fallend.
Analog dazu ist $s_{2k+3} - s_{2k+1} = a_{2k+2} - a_{2k+3} \geq 0 \quad \forall k$,
womit die Teilfolge $(s_{2k+1})_k$ monoton wachsend ist.
Zudem gilt $s_{2k+1} - s_{2k} = -a_{2k+1} \leq 0 \quad \Rightarrow \quad s_{2k} \geq s_{2k+1} \geq s_1 \ \forall k$
und $s_0 \geq s_{2k} \geq s_{2k+1} \ \forall k$.
Also ist $(s_{2k})_k$ nach unten und $(s_{2k+1})_k$ nach oben beschränkt.
$\Rightarrow \quad s := \lim s_{2k}$ und $s' := \lim s_{2k+1}$ existieren.
Es bleibt zu zeigen, daß die Grenzwerte der beiden Teilfolgen gleich sein müssen. $s - s' = \lim s_{2k} - \lim s_{2k+1} = \lim s_{2k} - s_{2k+1} = \lim a_{2k+1}$
Letzteres ist nach Voraussetzung 0 und somit gilt $s = s'$.
Damit konvergiert auch die gesamte Summenfolge:
$\forall \ \varepsilon > 0 \ \exists \ N_1, N_2 : \quad |s_{2k} - s| < \varepsilon \ \forall k \geq N_1, \quad |s_{2k+1} - s| < \varepsilon \ \forall k \geq N_2$.
$\Rightarrow \quad |s_n - s| < \varepsilon \quad \forall \ n \geq N := \max(2N_1, 2N_2 + 1)$ □

Beispiel.
- Die Leibniz- Reihe $\sum_{n=0}^{\infty} (-1)^n \frac{1}{2n+1} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} \dots$ konvergiert. Die Summe ist $\frac{\pi}{4}$, was in einem späteren Kapitel noch gezeigt werden soll.
- Auch die alternierende harmonische Reihe $\sum_{n=1}^{\infty} (-1)^{n+1} \frac{1}{n} = 1 - \frac{1}{2} + \frac{1}{3} - \dots$ ist nach Satz 2.15 konvergent.

Satz 2.16 (Majoranten- Kriterium). *Sind $(a_n)_{n \in \mathbb{N}}$ und $(c_n)_{n \in \mathbb{N}}$ zwei Folgen mit $|a_n| \leq |c_n| \ \forall n \geq p$ und $\sum_{n=p}^{\infty} |c_n|$ konvergiert, dann gilt*

1. $\sum_{n=p}^{\infty} a_n$ *und* $\sum_{n=p}^{\infty} |a_n|$ *sind konvergent.*
2. $\left| \sum_{n=p}^{\infty} a_n \right| \leq \sum_{n=p}^{\infty} |c_n|$.

Beweis. Da die Reihe $\sum_{n=p}^{\infty} |c_n|$ gemäß Voraussetzung konvergent ist, erfüllt sie das Cauchy- Kriterium und es gilt
$\forall \ \varepsilon > 0 \ \exists \ N : \quad |\sum_{k=m+1}^{n} a_k| \leq \sum_{k=m+1}^{n} |a_k| \leq \sum_{k=m+1}^{n} |c_k| < \varepsilon$
$\forall \ n > m \geq N$, d.h. $\sum_{n=p}^{\infty} a_n$ und $\sum_{n=p}^{\infty} |a_n|$ erfüllen ebenfalls das Cauchy-Kriterium und sind somit konvergent.
Ferner ist
$\left| \sum_{k=p}^{\infty} a_k \right| = \lim_{n \to \infty} \left| \sum_{k=p}^{n} a_k \right| \leq \lim_{n \to \infty} \sum_{k=p}^{n} |c_k| = \sum_{k=p}^{\infty} |c_k|$ □

Definition 2.14. Eine solche Reihe $\sum_{n=p}^{\infty} c_n$ heißt **Majorante** zu $\sum_{n=p}^{\infty} a_n$.

Satz 2.17. *Eine Reihe* $\sum_{n=1}^{\infty} a_n$ *mit* $a_n \geq 0\ \forall n$ *konvergiert genau dann, wenn die Folge* (s_n) *ihrer Partialsummen beschränkt ist.*

Beweis. Die Folge (s_n) wächst monoton, da nach Voraussetzung $a_n \geq 0$ gilt. Ist sie zudem beschränkt, so muß auch Konvergenz vorliegen. Ist in der anderen Richtung (s_n) konvergent, so folgt die Beschränktheit direkt. $\square$

Definition 2.15 (Absolute Konvergenz). Die Reihe $\sum_{n=1}^{\infty} a_n$ heißt **absolut konvergent**, wenn $\sum_{n=1}^{\infty} |a_n|$ konvergiert.

Die Konvergenz von Reihen dieser Gestalt ist also nicht von möglichen Aufhebungen von Summengliedern abhängig.

Folgerung.

- *Eine Reihe* $\sum_{n=1}^{\infty} a_n$ *konvergiert auf jeden Fall, wenn sie absolut konvergiert, da sie dann* $\sum_{n=1}^{\infty} |a_n|$ *als Majorante hat. Die Umkehrung gilt nicht, wie das Beispiel der alternierenden harmonischen Reihe zeigt.*

- *Für absolut konvergente Reihen läßt sich mit dem 2. Teil von Satz 2.16 (Majoranten- Kriterium) eine verallgemeinerte Dreiecksungleichung formulieren:*

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

Satz 2.18 (Wurzelkriterium). *Sei* $L := \limsup \sqrt[n]{|a_n|}$*. Dann ist* $\sum_{n=1}^{\infty} a_n$

- *absolut konvergent für* $L < 1$*;*

- *divergent für* $L > 1$*.*

Satz 2.19 (Quotientenkriterium). *Ist* $\sum_{n=1}^{\infty} a_n$ *eine Reihe, für die* $a_n \neq 0\ \ \forall\, n$ *und*

$$\exists\ N \in \mathbb{N},\ \ \vartheta \in \mathbb{R}\ \ mit\ \ 0 < \vartheta < 1 :\ \ \left|\frac{a_{n+1}}{a_n}\right| \leq \vartheta\ \ \forall\ \ n \geq N$$

gilt, so ist sie absolut konvergent.

Beweis. Die Abänderung endlich vieler Summanden verändert die Konvergenz bzw. Divergenz nicht. Sei also o.B.d.A. $\left|\frac{a_{n+1}}{a_n}\right| \leq \vartheta < 1\ \ \forall\ \ n \in \mathbb{N}$.
Durch vollständige Induktion über n wird zunächst $|a_n| \leq |a_1|\vartheta^{n-1}\ \ \forall n$ gezeigt. Der Induktionsanfang ($n = 1$) ist trivial; ist nun $|a_n| \leq |a_1|\vartheta^{n-1}$, so folgt $|a_{n+1}| = \left|\frac{a_{n+1}}{a_n} a_n\right| \leq \vartheta|a_n| \leq |a_1|\vartheta^{n-1}$.
Daher ist $|a_n| \leq |a_1|\vartheta^{n-1}\ \ \forall n$. Mit der absoluten Konvergenz der geometrischen Reihe $\sum_{n=1}^{\infty} |a_1|\vartheta^{n-1}$ (für $0 < \vartheta < 1$) folgt, daß $\sum_{n=1}^{\infty} |a_1|\vartheta^{n-1}$ Majorante zu $\sum_{n=1}^{\infty} a_n$ ist, womit $\sum_{n=1}^{\infty} a_n$ konvergiert. $\square$

Beispiel. $\sum_{n=1}^{\infty} \frac{n^2}{2^n}$ konvergiert, da für $n \geq 3$ gilt:
$\frac{a_{n+1}}{a_n} = \frac{(n+1)^2 2^n}{n^2 2^{n+1}} = \frac{(n+1)^2}{2n^2} = \frac{1}{2}\left(1+\frac{1}{n}\right)^2 \leq \frac{1}{2}\left(1+\frac{1}{3}\right)^2 = \frac{8}{9} < 1$

Bemerkung. Das Quotientenkriterium ist einfach anzuwenden, jedoch schwächer als das Wurzelkriterium.

Beispiel. Sei $\sum_{n=1}^{\infty} a_n = 1 + a^3 + a^2 + a^5 + a^4 + a^7 + \ldots$ mit
$a_n = \begin{cases} a^{n-1} & \text{für } n \text{ ungerade} \\ a^{n+1} & \text{für } n \text{ gerade} \end{cases}$ und $a < 1$.
Das Wurzelkriterium liefert den Nachweis der Konvergenz, denn:
$\limsup \sqrt[n]{|a_n|} = a < 1$
Das Quotientenkriterium hingegen läßt sich nicht anwenden, da $\left|\frac{a_{n+1}}{a_n}\right|$ alternierend a^3 und a^{-1}, d.h. nicht stets $\leq \vartheta < 1$ ist.

2.9 Rechnen mit Reihen

Es stellt sich die Frage nach der Konvergenz von Reihen, die aus bekannten durch Anwenden der Grundoperationen konstruiert wurden.

Satz 2.20 (Linearität). *Seien $\alpha, \beta \in \mathbb{R}$. Sind $\sum_{n=1}^{\infty} a_n$ und $\sum_{n=1}^{\infty} b_n$ konvergent, so auch $\sum_{n=1}^{\infty} (\alpha a_n + \beta b_n)$ und es gilt*

$$\sum_{n=1}^{\infty} (\alpha\, a_n + \beta\, b_n) = \alpha \sum_{n=1}^{\infty} a_n + \beta \sum_{n=1}^{\infty} b_n$$

Das Assoziativ- und das Kommutativgesetz lassen sich nicht ohne Weiteres auf unendliche Reihen übertragen.

Beispiel. Würde das Assoziativgesetz gelten, so ließe sich
$\sum_{n=1}^{\infty} (-1)^{n-1} = 1 - 1 + 1 - 1 + 1 - \ldots$ schreiben als
$\sum_{n=1}^{\infty} (1-1) = (1-1) + (1-1) + \ldots = 0$ und
$1 + \sum_{n=1}^{\infty} (-1+1) = 1 + (-1+1) + (-1+1) + \ldots = 1$, was nicht sein kann.
Man nehme jetzt an, das Kommutativgesetz ließe sich auf alle Reihen anwenden und betrachte die alternierende harmonische Reihe
$s = \sum_{n=1}^{\infty} (-1)^{n-1} \frac{1}{n} = 1 - \frac{1}{2} + \ldots + \frac{1}{2k-1} - \frac{1}{2k} + \ldots$
Durch Kommutieren der Summanden ergibt sich
$t = 1 + \left(-\frac{1}{2} - \frac{1}{4}\right) + \frac{1}{3} + \left(-\frac{1}{6} - \frac{1}{8}\right) + \ldots + \frac{1}{2k-1} + \left(-\frac{1}{4k-2} - \frac{1}{4k}\right) + \ldots$
Wegen $\left(\frac{1}{2k-1} - \frac{1}{4k-2} - \frac{1}{4k}\right) = \frac{1}{2}\left(\frac{1}{2k-1} - \frac{1}{2k}\right)$ gilt für die Partialsummenfolgen (t_n) und (s_n) die Beziehung $t_{3m} = \frac{1}{2} s_{2m}$.
Da nun (s_n) und mithin die Teilfolge (s_{2m}) gegen s konvergiert und überdies $|t_{3m+1} - t_{3m}| = \frac{1}{2m+1}$ und $|t_{3m+2} - t_{3m}| = \frac{1}{2m+1} - \frac{1}{4m+2}$ beliebig klein für genügend große m sind, gilt $\forall\, m \geq N$ hinreichend groß:
$|t_{3m} - \frac{1}{2}s| \leq |t_{3m} - \frac{1}{2}s_{2m}| + \frac{1}{2}|s_{2m} - s| = \frac{1}{2}|s_{2m} - s| < \varepsilon$

$|t_{3m+1} - \frac{1}{2}s| \leq |t_{3m} - \frac{1}{2}s| + |t_{3m+1} - t_{3m}| < \varepsilon$
$|t_{3m+2} - \frac{1}{2}s| \leq |t_{3m} - \frac{1}{2}s| + |t_{3m+2} - t_{3m}| < \varepsilon$
Demnach konvergiert die gesamte Folge (t_n) gegen $\frac{1}{2}s$ und mit $s > 0$ folgt daraus $s \neq t$; das Kommutativgesetz gilt in diesem Beispiel nicht.

Bemerkung. Für absolut konvergente Reihen bleibt das Kommutativgesetz jedoch gültig.

Satz 2.21 (Großer Umordnungssatz). *Ist eine Reihe $\sum_{n=1}^{\infty} a_n$ absolut konvergent und eine Zerlegung der Reihe in endlich oder unendlich viele Teilreihen hat die Eigenschaft, daß jedes Glied der Folge (a_n) in genau einer Teilreihe als Summand auftritt, so gilt:*

1. *Jede dieser Teilreihen konvergiert absolut*

2. *Sind $s_1, s_2, \dots$ die Summen der Teilreihen, dann konvergiert auch $\sum s_i$ absolut und $\sum s_i = \sum_{n=1}^{\infty} a_n$.*

Beweis der 1. Behauptung. Sei a_{ki} das i-te Glied der k-ten Teilreihe. Nun ist jede endliche Summe von absoluten Gliedern $|a_n|$ durch $s_a = \sum_{n=1}^{\infty} |a_n|$ nach oben beschränkt. Somit ist auch jede endliche Summe von absoluten Gliedern $|a_{ki}|$ der k-ten Teilreihe durch s_a beschränkt. Mit der Monotonie der Partialsummenfolgen $\left(\sum_{i=1}^{n} |a_{ki}|\right)_n$ ergibt sich daraus die absolute Konvergenz jeder Teilreihe. □

Damit kann das Assoziativgesetz auf absolut konvergente Reihen ausgeweitet werden:

Satz 2.22 (Umordnungssatz). *Jede aus Umordnung einer absolut konvergenten Reihe mit der Summe s hervorgegangene Reihe ist ebenfalls absolut konvergent mit der Summe s.*

Eine Anwendung findet der große Umordnungssatz u.a. bei Doppelreihen.
Ist $\sum_{i,k=0}^{\infty} a_{ik}$ eine solche, dann heißt

- $z_i = \sum_{k=0}^{\infty} a_{ik}$ die i-te Zeilensumme,

- $s_k = \sum_{i=0}^{\infty} a_{ik}$ die k-te Spaltensumme,

- $D_n = \sum_{i+k=n} a_{ik} = \sum_{i=0}^{n} a_{i(n-i)}$ die n-te Diagonalensumme .

Satz 2.23 (Doppelreihensatz). *Für eine Reihe $\sum_{i,k=0}^{\infty} a_{ik}$ sei die Menge aller endlicher Summen von absoluten Gliedern $|a_{ik}|$ beschränkt. Sei ferner $z_i = \sum_{k=0}^{\infty} a_{ik}$, $s_k = \sum_{i=0}^{\infty} a_{ik}$ und $D_n = \sum_{i+k=n} a_{ik}$.*
Dann gilt

1. $\sum_{i=0}^{\infty} z_i$, $\sum_{k=0}^{\infty} s_k$ *und* $\sum_{n=0}^{\infty} D_n$ *konvergieren absolut.*

2. $\sum_{i=0}^{\infty} z_i = \sum_{k=0}^{\infty} s_k = \sum_{n=0}^{\infty} D_n$

Beweis. Die Mengen der Indizes $\{i\} = \mathbb{N} \cup \{0\}$ und $\{k\} = \mathbb{N} \cup \{0\}$ sind abzählbar, also auch die Menge der Doppelindizes $(i,k) \in \mathbb{N}_0 \times \mathbb{N}_0$, d.h. es gibt eine bijektive Abbildung $(i,k) \mapsto (m),\ \ m \in \mathbb{N}$.
Somit läßt sich $\sum_{i,k=0}^{\infty} a_{ik}$ auch als einfache Summe $\sum_{m=1}^{\infty} a_m$ schreiben.
Diese Reihe ist absolut konvergent aufgrund von Monotonie und Beschränktheit.
Nach dem großen Umordnungssatz konvergieren damit auch alle Teilreihen absolut, also insbesondere alle z_i, s_k und D_n und es ist
$\sum_{m=1}^{\infty} a_m = \sum_{i=0}^{\infty} z_i = \sum_{k=0}^{\infty} s_k = \sum_{n=0}^{\infty} D_n$. □

Beispiel. Es ist $\sum_{i,j=2}^{\infty} \left(\frac{1}{j}\right)^i = 1$. Dazu sei zunächst bemerkt, daß alle Summanden positiv sind. Desweiteren ist
$\sum_{i,j=2}^{\infty} \left(\frac{1}{j}\right)^i = \sum_{j=2}^{\infty} \left(\sum_{i=2}^{\infty} \left(\frac{1}{j}\right)^i\right) = \sum_{j=2}^{\infty} \left(\frac{1}{j-1} - \frac{1}{j}\right) = 1$

Doppelsummen treten auch bei der Multiplikation von Reihen auf.

Definition 2.16. Es heißt $\sum_{i,k=0}^{\infty} (a_i b_k)$ das **Produkt** (die Produktreihe) von $\sum_{i=0}^{\infty} a_i$ und $\sum_{k=0}^{\infty} b_k$.

Satz 2.24. *Das Produkt zweier absolut konvergenter Reihen mit den Summen a und b ist ebenfalls absolut konvergent und hat den Wert ab.*

Definition 2.17. Sind $\sum_{i=0}^{\infty} a_i,\ \ \sum_{k=0}^{\infty} b_k$ zwei Reihen und $D_n = \sum_{i+k=n} a_i b_k$, dann heißt $\sum_{n=0}^{\infty} D_n$ das **Cauchy- Produkt** der beiden Reihen.

Satz 2.25. *Konvergieren $\sum_{i=0}^{\infty} a_i$ und $\sum_{k=0}^{\infty} b_k$ absolut, so auch ihr* Cauchy- *Produkt. Sein Wert ist $(\sum_{i=0}^{\infty} a_i)(\sum_{k=0}^{\infty} b_k)$.*

Beweis. Aus der absoluten Konvergenz von $\sum_{i=0}^{\infty} a_i$ und $\sum_{k=0}^{\infty} b_k$ folgt, daß jede endliche Summe von $|a_i b_k|$ durch $\sum_{i=0}^{\infty} |a_i| \sum_{k=0}^{\infty} |b_k|$ beschränkt ist.
Es ist $D_n = \sum_{i+k=n} a_i b_k$ die n-te Diagonalensumme der Doppelreihe $\sum_{i,k=0}^{\infty} a_i b_k$.
Nach dem Doppelreihensatz ist also $\sum_{n=0}^{\infty} D_n$ absolut konvergent und $\sum_{n=0}^{\infty} D_n = \sum_{i,k=0}^{\infty} a_i b_k$.
Mit Satz 2.24 folgt daraus der zweite Teil der Behauptung. □

Beispiel. Die geometrische Reihe konvergiert für $|x| < 1$. In dem Fall gilt $\left(\frac{1}{1-x}\right)^2 = (1 + x + x^2 + \ldots)^2$. Das Cauchy- Produkt der beiden Faktoren ist

$1 + 2x + 3x^2 + 4x^3 + \ldots = \sum_{n=1}^{\infty} nx^{n-1}$.
Nach Satz 2.25 ist also $\sum_{n=1}^{\infty} nx^{n-1} = \left(\frac{1}{1-x}\right)^2$

Satz 2.26 (Satz von Mertens[13]). *Ist eine Reihe absolut konvergent und eine weitere (nicht notwendigerweise absolut) konvergent, so konvergiert auch ihr* CAUCHY- *Produkt.*

Satz 2.27 (Abel[14]). *Konvergiert von zwei konvergenten Reihen mit den Summen a und b auch das Cauchy- Produkt, so ist dessen Wert gleich ab.*

2.10 Potenzreihen

Definition 2.18. Eine Reihe der Form

$$P(x) = \sum_{n=0}^{\infty} a_n x^n = a_0 + a_1 x + a_2 x^2 + \ldots$$

heißt **Potenzreihe**.

Es stellt sich die Frage, für welche Argumente x eine solche Potenzreihe konvergiert.

Lemma 2.28. *Konvergiert eine Potenzreihe $P(x) = \sum_{n=0}^{\infty} a_n x^n$ in $x_0 \in \mathbb{R}\backslash\{0\}$, dann konvergiert $P(x)$ auch absolut in jedem Punkt $x \in \mathbb{R}$ mit $|x| < |x_0|$.*

Beweis. $\exists S: \ |a_n {x_0}^n| \leq S$
Damit ist $|a_n x^n| = \left|a_n {x_0}^n \left(\frac{x}{x_0}\right)^n\right| = |a_n {x_0}^n| \left|\left(\frac{x}{x_0}\right)\right|^n \leq S\,q^n$ mit $q = \left|\left(\frac{x}{x_0}\right)\right| < 1$. Daher hat $P(x)$ die geometrische Reihe $S\sum_{n=0}^{\infty} q^n$ als Majorante und ist folglich absolut konvergent. □

Satz 2.29. *Ist $P(x)$ eine Potenzreihe und $R(x) = \sup\{r \in \mathbb{R} \ | \ P(r)\,\text{konvergiert}\}$, so konvergiert $P(x)$ absolut $\forall\, x$ mit $|x| < R$; für $|x| > R$ divergiert $P(x)$.*

Dies führt zur folgenden Definition:

Definition 2.19. Eine reelle Zahl R heißt **Konvergenzradius** einer Potenzreihe $P(x)$ genau dann, wenn $P(x)$ konvergiert für $|x| < P(x)$ und divergiert für $|x| > P(x)$.

---

[13] *Franz Mertens* (1840-1927)

[14] *Niels Henrik Abel* (1802-1829)

Beispiel. Bekanntermaßen hat die geometrische Reihe $\sum_{n=0}^{\infty} x^n$ den Konvergenzradius 1.

Bemerkung. Aus dem Wurzel- und dem Quotientenkriterium folgen Möglichkeiten, den Konvergenzradius für Potenzreihen zu berechnen (vgl. *Übungsaufgabe 31*):

- $R = \frac{1}{\limsup \sqrt[n]{|a_n|}}$ (CAUCHY - HADAMARD[15])
- $R = \frac{1}{\lim \left|\frac{a_{n+1}}{a_n}\right|}$, sofern $\lim \left|\frac{a_{n+1}}{a_n}\right|$ existiert. (EULER[16])

Im Folgenden sollen drei wichtige Potenzreihen aufgeführt werden.

- Die Exponentialreihe ist definiert als

  $$\exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \dots$$

  Ihr Konvergenzradius ist ∞, da $\lim \frac{(n+1)!}{n!} = \lim\,(n+1) = \infty$.

- Die Logarithmusreihe

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

  hat den Konvergenzradius 1.

- Für den Exponenten s ist die Binomialreihe

  $$\mathrm{B}_s(x) = \sum_{n=0}^{\infty} \binom{s}{n} x^n$$

  Dabei ist der Binomialkoeffizient $\binom{s}{n}$ für $n \in \mathbb{Z}$, $s \in \mathbb{R}$:

  $$\binom{s}{n} = \begin{cases} \frac{s(s-1)\dots(s-n+1)}{n!} & \text{für } n > 0 \\ 1 & \text{für } n = 0 \cdot \\ 0 & \text{für } n < 0 \end{cases}$$

  Ist $s \in \mathbb{N}_0$, so ist $\binom{s}{n} = 0$, falls $n > s$ und $B_s(x)$ wird zu der endlichen Reihe $\sum_{n=0}^{s} \binom{s}{n} x^n$ mit der Summe $(1+x)^s$.
  Für $s \notin \mathbb{N}_0$ hat $B_s(x)$ den Konvergenzradius 1, denn
  $\lim_{n\to\infty} \left|\frac{\binom{s}{n}}{\binom{s}{n+1}}\right| = \lim_{n\to\infty} \left|\frac{n+1}{s-n}\right| = 1.$

[15] *Jacques Hadamard* (1865-1963)

[16] *Leonhard Euler* (1707-1783)

Bemerkung. Aus dem Multiplikationssatz folgen direkt Additionstheoreme für bestimmte Funktionen.

Beispiel. $\exp(x)$ ist absolut konvergent $\forall x \in \mathbb{R}$. Daher ist
$\exp(x)\exp(y) = \left(\sum_{i=0}^{\infty} \frac{x^i}{i!}\right)\left(\sum_{k=0}^{\infty} \frac{y^k}{k!}\right)$
$= \sum_{j=0}^{\infty}\left(\sum_{i=0}^{j} \frac{x^i y^{j-i}}{i!(j-i)!}\right)$ (Cauchy- Produkt)
$= \sum_{j=0}^{\infty} \frac{1}{j!}\left(\sum_{i=0}^{j} \binom{j}{i} x^i y^{y-i}\right) = \sum_{j=0}^{\infty} \frac{1}{j!}(x+y)^j$
$= \exp(x+y)$

In der Praxis werden Funktionen häufig dadurch approximiert, daß von Potenzreihen, die gegen den gesuchten Funktionswert konvergieren, nur endliche viele Glieder aufsummiert werden:

$$\sum_{k=0}^{\infty} a_k x^k = \sum_{k=0}^{n-1} a_k x^k + R_n(x) \quad \approx \quad \sum_{k=0}^{n-1} a_k x^k.$$

Die restliche Reihe, das Restglied $R_n(x) = \sum_{k=n}^{\infty} a_k x^k$, welches nicht berechnet wird, stellt einen Fehler dar, für den es eine Abschätzung zu finden gilt.

Satz 2.30. *$\sum_{k=0}^{\infty} a_k x^k$ habe den Konvergenzradius $R > 0$. Dann folgt*

$$\forall\, r < R,\, r > 0 \quad \exists\, C : \; |R_n(x)| < C_n |x|^n \qquad \textit{für} \quad |x| \leq r.$$

Fällt überdies $(a_k)_k$ monoton, so ist $R > 1$ und

$$|R_n(x)| \leq |a_n| \frac{|x|^n}{1-|x|} \qquad \textit{für} \quad |x| < 1.$$

Beweis. Für $|x| \leq r < R$ ist $\sum_{k=0}^{\infty} a_k x^k$ und mithin $R_n(x)$ absolut konvergent.
Die verallgemeinerte Dreiecksungleichung ergibt
$|R_n(x)| = \left|\sum_{k=n}^{\infty} a_k x^k\right| \leq \sum_{k=n}^{\infty} |a_k||x|^k$.
Mit dem Majorantenkriterium ist außerdem
$\sum_{k=n}^{\infty} |a_k||x|^k \leq |x|^n \sum_{k=n}^{\infty} |a_k||r|^{k-n}$
und folglich $|R_n(x)| \leq C_n |x|^n$ mit $C_n = \sum_{j=0}^{\infty} |a_{n+j}||r|^j$.

Ist nun $(a_k)_k$ auch monoton fallend, so ist die geometrische Reihe
$|a_0| \sum_{k=0}^{\infty} |x|^k$ Majorante für $\sum_{k=0}^{\infty} a_k x^k$.
Damit folgt $R \geq 1$ und
$|R_n(x)| \leq |a_n| \sum_{k=n}^{\infty} |x|^k = |a_n||x|^n \sum_{k=0}^{\infty} |x|^k = |a_n| \frac{|x|^n}{1-|x|}$. □

Beispiel. Sei $|x| < 1$ und $|s| < 1$. Man betrachte die lineare Näherung $1 + sx$ für $(1+x)^s = \sum_{k=0}^{\infty} \binom{s}{k} x^n = 1 + sx + R_2(x)$.
Wegen $|s| < 1$ ist $\left|\frac{s-k}{k+1}\right| < 1$ und mit $\binom{s}{k+1} = \frac{s-k}{k+1}\binom{s}{k}$ folgt, daß $\left(\left|\binom{s}{k}\right|\right)_k$

monoton fallend ist.
Nach Satz 2.30 läßt sich das Restglied abschätzen durch
$|R_2(x)| \leq \left|\frac{s(s-1)}{2}\right| \frac{|x|^2}{1-|x|}$

Beispiel. $\exp(x) = \sum_{k=0}^{n-1} \frac{x^k}{k!} + R_n(x)$. Das Restglied soll mit einer geometrischen Reihe abgeschätzt werden:
$|R_n(x)| \leq \sum_{k=n}^{\infty} \left|\frac{x^k}{k!}\right| = \frac{|x|^n}{n!}\left(1 + \frac{|x|}{n+1} + \ldots + \frac{|x|^k}{(n+1)(n+2)\ldots(n+k)} + \ldots\right)$
$\leq \frac{|x|^n}{n!}\left(1 + \frac{|x|}{n+1} + \ldots + \frac{|x|^k}{(n+1)^k} + \ldots\right) = \frac{|x|^n}{n!} \frac{1}{1-\frac{|x|}{n+1}} = \frac{|x|^n}{n!} \frac{n+1}{(n+1)-|x|}$.
Für den Fall $|x| \leq 1 + \frac{n-1}{2}$ vereinfacht sich dies zu $|R_n(x)| \leq \frac{2|x|^n}{n!}$.

3 Reelle Funktionen

3.1 Grundlagen

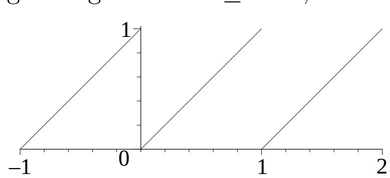
Im vorangegangenen Abschnitt wurden Folgen, d.h. Abbildungen der Menge $\mathbb{N}$ auf $\mathbb{R}$ betrachtet. Nun sollen allgemeiner Abbildungen einer beliebigen Menge auf die Menge der reellen Zahlen $\mathbb{R}$ betrachtet werden.

Definition 3.1. Eine **reelle Funktion** auf der Menge X ist eine Abbildung $f : X \to \mathbb{R}$, die jedem $x \in X$ eindeutig eine reelle Zahl $f(x)$ zuordnet.
Es heißt X der **Definitionsbereich** von f.
Die Menge $f(X) = \{f(x) \in \mathbb{R} \mid x \in X\}$ ist der **Wertebereich** (das **Bild**) von f.

Definition 3.2. Die Menge $\mathcal{G}(f) = \{(x, f(x)) \mid x \in X\} \subset X \times \mathbb{R}$ heißt der **Graph** von f.

Durch Visualisierung des Graphen können einige Funktionen veranschaulicht werden.

Beispiel. Der Graph der Sägezahnfunktion $f(x) = x - \lfloor x \rfloor$, wobei $\lfloor x \rfloor$ die größte ganze Zahl $\leq x$ ist, hat z.B. die Darstellung



Bemerkung. Der Graph muß nicht zusammenhängend sein.

Beispiel. $f(x) = \begin{cases} 1 & \text{für } x \in \mathbb{Q} \\ 0 & \text{für } x \in \mathbb{R}\setminus\mathbb{Q} \end{cases}$ läßt sich nicht zeichnen.

Bemerkung. Ist der Definitionsbereich einer Funktion die Menge der natürlichen Zahlen, so ist diese Funktion auch eine Folge.

Dementsprechend können einige Definitionen, die sich auf Folgen beziehen, auf reelle Funktionen übertragen werden.

Definition 3.3. Ist $X \subset \mathbb{R}$, so heißt $f : X \to \mathbb{R}$ **monoton wachsend**, wenn $\forall\, x_1, x_2 \in X$ mit $x_1 < x_2$ gilt: $f(x_1) \leq f(x_2)$.
f heißt **monoton fallend**, wenn $\forall\, x_1, x_2 \in X$ mit $x_1 < x_2$ gilt: $f(x_1) \geq f(x_2)$

Analog können die Definitionen der strengen Monotonie, der Schranke, des Infimums und des Supremums von Folgen übertragen werden.

Definition 3.4. $x_0 \in X$ heißt **Nullstelle** einer Funktion $f : X \to \mathbb{R}$, wenn $f(x_0) = 0$ ist.
f **verschwindet identisch** auf dem Intervall I, wenn $f(x) = 0 \ \forall x \in I$.
Die Funktion f heißt **konstante Funktion**, wenn $f(x) = const$.
Ist $X \subset \mathbb{R}$, so heißt $f : X \to \mathbb{R}$ mit $f(x) = x \ \ \forall\, x \in X$ die **identische Funktion**, in Zeichen: id_X.

Definition 3.5. Haben zwei Funktionen f und g den gleichen Definitionsbereich, so ist
$(f \pm g)(x) = f(x) + g(x)$,
$(f \cdot g)(x) = f(x) \cdot g(x)$,
$\left(\frac{f}{g}\right)(x) = \frac{f(x)}{g(x)}$ für $g(x) \neq 0$.

Definition 3.6. Seien $f : X \to \mathbb{R}$ und $g : Y \to \mathbb{R}$ zwei Funktionen mit $f(X) \subset Y$. Dann ist die zusammengesetzte Funktion $g \circ f : X \to \mathbb{R}$ mit $(g \circ f)(x) = g(f(x))$.

Definition 3.7. Ist $f : X \to \mathbb{R}$ injektiv und $X \subset \mathbb{R}$, so heißt $g : f(X) \to \mathbb{R}$ mit $(g \circ f)(x) = x \ \ \forall\, x \in X$ die **inverse Funktion** (die **Umkehrfunktion**) von f.

Bemerkung. Streng monotone Funktionen sind stets injektiv und haben somit auch eine Inverse.

Bemerkung. Ist g die inverse Funktion von f, so ist der Graph von g: $\mathcal{G}(f) = \{(f(x), x) \mid x \in X\}$ der an der Geraden $\{(x, x) \mid x \in \mathbb{R}\}$ gespiegelte Graph von f.

3.2 Polynome und rationale Funktionen

Definition 3.8. Eine Funktion $p : \mathbb{R} \to \mathbb{R}$ mit $p(x) = \sum_{i=0}^{n} a_i x^i$ und $a_i \in \mathbb{R}$ heißt **Polynom** über $\mathbb{R}$ in der Unbekannten x.
Die Zahlen a_i sind die **Koeffizienten** des Polynoms p.

Sind alle Koeffizienten 0, dann ist p das **Nullpolynom**: $p = 0$.
Die größte ganze Zahl $n \geq 0$, für die der Koeffizient $a_n \neq 0$ ist, heißt **Grad** des Polynoms $p \,:\, n = grad(p)$. Der Grad des Nullpolynoms wird als $-\infty$ definiert.
Mit LT(p) wird der **Leitterm** $a_{grad(p)}x^{grad(p)}$ des Polynoms p bezeichnet.
Bei der Schreibweise eines Polynoms als $\sum_{i=0}^{n} a_i x^i$ soll o.B.d.A. $a_n \neq 0$ sei.

Bemerkung.

- Zwei Polynome sind gleich, wenn sie koeffizientenweise übereinstimmen (algebraische Gleichheit).
- Zwei Polynome p und $\tilde{p}$ sind auch gleich, wenn $p(x) = \tilde{p}(x) \;\; \forall\, x \in \mathbb{R}$ (analytische Gleichheit).
- Der Vektorraum der Polynome über $\mathbb{R}$ vom Grad n ist isomorph zum $\mathbb{R}$-Vektorraum $\mathbb{R}^{n+1}$.

Bemerkung.

- Die Menge der Polynome ist abgeschlossen unter Addition, Subtraktion und Multiplikation.
- Die Summe zweier Polynome vom Grad m und n hat einen Grad $\leq \max{(m, n)}$.
- Das Produkt zweier Polynome vom Grad m und n hat den Grad $m + n$.

Satz 3.1. *Jedes Polynom $p \neq 0$ vom Grad n hat höchstens n Nullstellen.*

Beweis. Induktionsanfang: Das konstante Polynom $p(x) = a, \;\; a \neq 0$ hat den Grad 0 und keine Nullstelle.
Induktionsschritt: Die Behauptung sei wahr für den Grad $n - 1$. Wegen der Identität

$$x^m - y^m = (x - y) \sum_{i=0}^{m-1} x^i y^{m-1-i}$$

ist $p(x) - p(x_0) = \sum_{i=0}^{n} a_i(x^i - x_0^i) = (x - x_0)p_0(x)$, wobei $p_0(x)$ ein Polynom vom Grad (n-1) ist.
Sei nun x_0 eine Nullstelle von p. Damit ist $p(x) = (x - x_0)p_0(x)$. Dieses Produkt kann wegen der Körpereigenschaften von $\mathbb{R}$ nur verschwinden, wenn einer der Faktoren verschwindet. $(x - x_0)$ hat die eine Nullstelle x_0 und p_0 hat nach Induktionsvoraussetzung maximal $n - 1$ Nullstellen. Somit verschwindet p über höchstens n Argumenten. $\square$

Bemerkung. Es gibt nichtkonstante Polynome, die keine Nullstelle in $\mathbb{R}$ haben.

Beispiel. $x^2 + 1$

Definition 3.9. Sind $p(x)$ und $q(x)$ mit $q \neq 0$ Polynome, so ist $R(x) = \frac{p(x)}{q(x)}$ eine **rationale Funktion**. An den Nullstellen von q ist sie nicht definiert.

Bemerkung. Analog zur Definition der rationalen Zahlen lassen sich rationale Funktionen durch Paare von Polynomen charakterisieren, die in Äquivalenzklassen zerlegt werden können.

Bemerkung. Die Menge der rationalen Funktionen ist abgeschlossen unter den Operationen aus Definition 3.5.

Satz 3.2 (Polynomdivision). *Für jede rationale Funktion* $\frac{p(x)}{q(x)}$*, wobei* $p(x) = \sum_{i=0}^{m} a_i x^i$ *und* $q(x) = \sum_{i=0}^{n} b_i x^i$ *Polynome in der Unbekannten* x *sind, gibt es Polynome* $g(x)$ *und* $r(x)$*, sodaß* $\quad grad(r) < grad(q)$ *oder* $r = 0 \quad$ *und*

$$\frac{p}{q} = g + \frac{r}{q} \tag{3}$$

Beweis. Der Beweis gelingt durch Angabe eines Divisionsalgorithmus, welcher die Darstellung (3) als Ergebnis liefert.
Ist $grad(p) < grad(q)$, dann sind $g = 0$ und $r = p$ die gesuchten Polynome.
Ist $grad(p) \geq grad(q)$, so wird im ersten Schritt
$r_1 = p - \frac{a_m}{b_n} x^{m-n} q$ gebildet, womit

$$p = g_1 q + r_1 \quad mit \quad g_1 = \frac{a_m}{b_n} x^{m-n}$$

ist. Für $r_1 = 0$ oder $grad(r_1) < grad(q)$ liefern $g = g_1$ und $r = r_1$ die gewünschte Darstellung.
Im anderen Falle wird der Algorithmus mit der Bildung von $r_2 = r_1 - \frac{\mathrm{LT}(r_1)}{b_n x^n} q$ fortgeführt. Damit ist $p = g_1 q + r_1 = \left(g_1 + \frac{\mathrm{LT}(r_1)}{b_n x^n}\right) q + r_2 = g_2 + r_2$.
Sei allgemein $p = g_i q + r_i$ die Darstellung von p nach dem i-ten Divisionsschritt. Für den Fall $r_i \neq 0$ und $grad(r_i) \geq grad(q)$ besteht der $(i+1)$-te Schritt dann aus der Bildung von $r_{i+1} = r_i - \frac{\mathrm{LT}(r_i)}{b_n x^n} q$, was eine Darstellung von p als

$$p = \left(g_i + \frac{\mathrm{LT}(r_i)}{b_n x^n}\right) q + r_{i+1} = g_{i+1} q + r_{i+1}$$

liefert. Dabei ist stets entweder $r_{i+1} = 0$ oder $grad(r_{i+1}) < grad(r_i)$, da sich bei der Bildung von r_{i+1} der $\mathrm{LT}(r_i)$ und $\frac{\mathrm{LT}(r_i)}{b_n x^n} \mathrm{LT}(q)$ aufheben.
Der Grad der Reste r_i verringert sich also in jedem Schritt, d.h. nach endlich vielen Schritten erhält man eine Darstellung
$p = gq + r \qquad mit \qquad r = 0 \quad oder \quad grad(r) < grad(q)$ □

Bemerkung. Die Polynome g und r aus Gleichung (3) sind eindeutig bestimmt.

Definition 3.10. Ist in $p = gq$ für ein Polynom g, so heißt p **teilbar** durch q. q heißt **trivialer Teiler** von p, wenn $grad(q) \in \{grad(p), 0, -\infty\}$; andernfalls ist q ein **echter Teiler** von p.

Definition 3.11. Ein Polynom ohne echten Teiler heißt **irreduzibel** (oder: **prim**).

Bemerkung. Jedes reduzible Polynom läßt sich eindeutig bis auf konstante Vielfache in irreduzible Faktoren zerlegen.

Mit dem EUKLIDischen Algorithmus läßt sich der größte gemeinsame Teiler zweier Polynome $p_1(x)$ und $p_2(x)$ bestimmen:
Man starte mit p_1 und p_2 und dividiere nach dem Algorithmus aus dem Beweis von Satz 3.2 in jedem Schritt p_j durch p_{j+1}, um als Rest das Polynom p_{j+2} zu erhalten: $p_j = q_j p_{j+1} + p_{j+2}$.
Dieses ist entweder 0 oder hat einen kleineren Grad als p_{j+1}.
Somit ist nach endlich vielen Schritten ein $p_{k+2} = 0$, d.h. $p_k = q_k p_{k+1}$ für zwei Polynome p_k und p_{k+1} aus obiger Konstruktion.
p_{k+1} ist nun der gesuchte größte gemeinsame Teiler von p_1 und p_2, da der ggT zum einen jedes der p_j und damit auch p_{k+1} teilt und zum anderen wegen $p_k = q_k p_{k+1}$ und $p_j = q_j p_{j+1} + p_{j+2}$ alle Polynome p_j, also insbesondere p_1 und p_2, durch p_{k+1} teilbar sind.

Satz 3.3. *Ist $g(x)$ der größte gemeinsame Teiler zweier Polynome $p_1(x)$ und $p_2(x)$, dann gibt es Polynome $A(x)$ und $B(x)$, sodaß*

$$g = Ap_1 + Bp_2$$

Sind speziell $p_1(x)$ und $p_2(x)$ teilerfremd, so gibt es $A(x)$ und $B(x)$ mit

$$1 = Ap_1 + Bp_2$$

Definition 3.12. Sind p und q Polynome mit $grad(p) < grad(q)$, wobei q irreduzibel und nichtkonstant ist, dann heißt $\frac{p}{q^k}$ mit $k \in \mathbb{N}$ **Partialbruch**.

Satz 3.4. *Jede rationale Funktion ist darstellbar als Summe von Partialbrüchen und eines Polynoms.*

3.3 Stetigkeit von Funktionen

Definition 3.13. Die Funktion $f: X \to \mathbb{R}$ heißt **stetig** an der Stelle $x_0 \in X$ genau dann, wenn

$$\forall\ \varepsilon > 0\ \ \exists\ \ \delta >\ 0:\ |f(x) - f(x_0)| < \varepsilon \qquad \textit{für} \quad |x - x_0| < \delta, \quad x \in X$$

Weiterhin heißt f stetig in X, wenn f in allen Punkten $x \in X$ stetig ist.

Definition 3.14. Eine Funktion $f : X \to \mathbb{R}$ heißt **stetig** in $x_0 \in X$, wenn es zu jeder Umgebung V von $f(x_0)$ eine Umgebung U von x_0 gibt, sodaß $\forall x \in U \cap X$ gilt: $f(x) \in V$, d.h. $f(U \cap X) \subset V$.

Bemerkung. Definition 3.13 und 3.14 sind äquivalent. 3.13 folgt aus 3.14, denn: Ist $f : X \to \mathbb{R}$ stetig in x_0, so gibt es für jede Umgebung U_ε von $f(x_0)$ eine Umgebung U_δ von x_0 so, daß $f(U_\delta \cap X) \subset U_\varepsilon$, d.h. $|f(x) - f(x_0)| < \varepsilon$ für $|x - x_0| < \delta$.

Beispiel.
- Polynome und rationale Funktionen sind stetig in jedem Punkt ihres Definitionsbereiches.
- $\operatorname{sgn}(x)$ ist unstetig für $x = 0$ und stetig für alle $x \neq 0$.
- $|x|$ ist stetig.
- $\lfloor x \rfloor$ ist unstetig für $x \in \mathbb{Z}$ und stetig für alle übrigen x.
- $f(x) = \begin{cases} 1 & \text{für } x \in \mathbb{Q} \\ 0 & \text{für } x \in \mathbb{R}\backslash\mathbb{Q} \end{cases}$ ist für kein $x \in \mathbb{R}$ stetig, denn:
  Sei $x_0 \in \mathbb{R}$ beliebig. Jede Umgebung $U_\delta(x_0)$ enthält sowohl rationale als auch irrationale Punkte und somit auch ein $x \in \mathbb{R}$, sodaß $|f(x) - f(x_0)| = 1$ ist. Für $\varepsilon < 1$ gibt es damit aber kein δ, sodaß $f(x) \in U_\varepsilon(f(x_0)) \quad \forall x \in U_\delta(x_0)$.

Satz 3.5 (Folgenkriterium). *Eine Funktion $f : X \to \mathbb{R}$ ist stetig in $x_0 \in X$ dann und nur dann, wenn für jede Folge $(x_n)_n$ in X, die gegen x_0 konvergiert, die Folge $(f(x_n))_n$ gegen $f(x_0)$ konvergiert.*

Beweis. Sei $\lim(f(x_n)) = f(x_0)$ für jede Folge (x_n) in X mit $\lim(x_n) = x_0$.
Man nehme an, es gäbe ein $\varepsilon > 0$, sodaß kein δ mit der Eigenschaft $|f(x) - f(x_0)| < \varepsilon \quad \forall x$ mit $|x - x_0| < \delta$ existiert.
Dann müßte es eine Folge $(x_n)_n$ so geben, daß $|x_n - x_0| < \frac{1}{n}$ und $|f(x_n) - f(x_0)| \geq \varepsilon$ ist. Diese Folge hat den Grenzwert x_0, aber $(f(x_n))_n$ konvergiert nicht gegen $f(x_0)$.
Sei nun f stetig in x_0 im Sinne von Definition 3.14, d.h.
$\forall\ \varepsilon > 0\, \exists\, U(x_0)$, sodaß $|f(x) - f(x_0)| < \varepsilon$ für $x \in U$. Hat die Folge $(x_n)_n$ in X den Grenzwert x_0, so liegen fast alle x_n in U. Und mit $|f(x_n) - f(x_0)| < \varepsilon$ für diese x_n ist die Folge $(f(x_n))_n$ konvergent gegen $f(x_0)$. □

Definition 3.15. Ist $f : X \to \mathbb{R}$ eine Funktion und $\tilde{X} \subset X$, dann heißt $f|_{\tilde{X}} : \tilde{X} \to \mathbb{R}$ **Einschränkung** von f auf $\tilde{X}$.

Bemerkung. Ist f stetig, so auch $f|_{\tilde{X}}$. Die Umkehrung gilt nicht, wie das Beispiel $f(x) = \operatorname{sgn}(x)$, $X = \mathbb{R}$, $\tilde{X} = \mathbb{R}^+$ zeigt.

Satz 3.6. *Sind $f, g : X \to \mathbb{R}$ in $x_0 \in X$ stetig, dann auch $f \pm g$ und $f \cdot g$. Ist ferner $g(x_0) \neq 0$, so ist auch $\frac{f}{g}$ stetig in x_0.*

Beweis. Sei (x_n) eine Folge in X, die gegen x_0 konvergiert.
Mit der Stetigkeit von f und g und Satz 3.5 folgt
$\lim\,(f(x_n)) = f(x_0)$ und $\lim\,(g(x_n)) = g(x_0)$.
Damit ist $\lim\,((f \pm g)(x_n)) = \lim\,(f(x_n) \pm g(x_n)) =$
$\lim\,(f(x_n)) \pm \lim\,(g(x_n)) = f(x_0) \pm g(x_0) = (f \pm g)(x_0)$ und analog
$\lim\,((fg)(x_n)) = fg(x_0)$, d.h. $f \pm g$ und fg sind stetig nach dem Folgenkriterium.
Für $g(x_0) \neq 0$ gibt es eine Umgebung von x_0, in der g keine Nullstelle hat. In dieser liegen fast alle Glieder der Folge (x_n), da diese gegen x_0 konvergieren sollte. Somit existiert $\lim\left(\frac{f}{g}(x_n)\right)$ und $\lim\left(\frac{f}{g}(x_n)\right) = \frac{\lim\,(f(x_n))}{\lim\,(g(x_n))} = \frac{f}{g}(x_0)$ □

Bemerkung. Aus dem vorangegangenen Beweis bleibt noch zu zeigen, daß es für eine in x_0 stetige Funktion g mit $g(x_0) \neq 0$ eine Umgebung U von x_0 gibt, sodaß g in U keine Nullstelle hat.
Dazu betrachte man eine Umgebung U von x_0, für die
$|g(x) - g(x_0)| < \varepsilon := \frac{1}{2}|g(x_0)|\ \ \forall x \in U$ ist. Diese muß existieren, weil g stetig ist. Mit der Dreicksungleichung folgt
$|g(x)| \geq |g(x_0)| - |g(x) - g(x_0)| > |g(x_0)| - \varepsilon = \frac{1}{2}|g(x_0)| > 0 \quad \forall\, x \in U$, d.h.
$g(x) \neq 0\ \ \forall\, x \in U$ und U ist die gesuchte Umgebung.

Satz 3.7. *Sind $f : X \to \mathbb{R}$ und $g : Y \to \mathbb{R}$, wobei $f(X) \subset Y$, zwei in $x_0 \in X$ bzw. $f(x_0) \in Y$ stetige Funktionen, so ist auch $g \circ f$ stetig in x_0.*

Beweis. Sei $(x_n)_n$ eine gegen x_0 konvergente Folge in X. Dann ist $(f(x_n))_n$ wegen der Stetigkeit von f eine Folge in Y mit dem Grenzwert $f(x_0)$. Und da auch g stetig ist, konvergiert $(g(f(x_n)))_n$ gegen $g(f(x_0))$. □

Bemerkung. Aus Satz 3.7 folgt direkt, daß $|f|$ an einer Stelle stetig ist, wenn dies für f gilt.

Satz 3.8. *Sei $f : [a, b] \to \mathbb{R}$ streng monoton und stetig. Dann existiert die Umkehrfunktion und sie ist stetig.*

Beweis: Übungsaufgabe 36.

Satz 3.9 (Zwischenwertsatz). *Sei $f : [a, b] \to \mathbb{R}$ eine stetige Funktion. Dann existiert für jedes $\gamma \in [f(a), f(b)]$ ein $c \in [a, b]$, sodaß $f(c) = \gamma$ ist.*

Beweis. Sei o.B.d.A. $f(a) < \gamma < f(b)$. Zu der nichtleeren, nach oben beschränkten Menge $M = \{t \in [a, b] \mid f(t) < \gamma\}$ gibt es nach dem

Vollständigkeitsaxiom ein $c = \sup(M) \in [a,b]$.
Für eine Folge (t_n) in M, die gegen c konvergiert, muß mit dem Folgenkriterium $\lim(t_n) = f(c) \leq \gamma$ gelten.
Nun ist $f(x) \geq \gamma \ \forall x \in [c,b]$ nach Definition von c. Somit gilt für eine Folge (x_n) in $[c,b]$ mit $\lim(x_n) = c$: $\lim(f(x_n)) = f(c) \geq \gamma$.
Also ist $f(c) = \gamma$ für ein $c \in [a,b]$. □

Bemerkung. Der Zwischenwertsatz ist Grundlage zahlreicher Existenzbeweise der Analysis.

Folgerung.

- *Das Polynom $p(x) = x^n - a$ hat für alle $n \in \mathbb{N}$ und $a > 0$ eine positive Nullstelle, da zum einen $p(0) = -a < 0$ und mit der* Bernoulli- *Ungleichung zum anderen $p(1+a) > 0$ ist.*

- *Allgemein läßt sich die Existenz von Nullstellen in einigen Fällen mit Hilfe des Zwischenwertsatzes zeigen.*

- *Jedes reelle Polynom $p(x)$ mit ungeradem Grad hat mindestens eine Nullstelle.*

  Beweis. Sei o.B.d.A. $p(x) = x^n + q(x)$ mit $q(x) = a_1x^{n-1} + \ldots + a_n$.
  Sei weiterhin $b = 1 + |a_1| + \ldots + |a_n|$. Damit ist
  $|q(\pm b)| \leq |a_1|b^{n-1} + \ldots + |a_n| \leq (|a_1| + \ldots + |a_n|)\, b^{n-1} = (b-1)b^{n-1} < b^n$
  Mit der Voraussetzung n ist ungerade erhält man schließlich
  $p(b) \geq b^n - |q(b)| > 0$ und $p(-b) \leq -b^n + |g(-b)| < 0$ und die Existenz einer Nullstelle im Intervall $[-b,b]$ nach dem Zwischenwertsatz. □

Satz 3.10 (Minimum und Maximum). *Jede in einem abgeschlossenen und beschränkten Intervall $[a,b]$ stetige Funktion $f : [a,b] \to \mathbb{R}$ ist beschränkt und nimmt ihr Maximum und Minimum an, d.h. $\exists\ \xi_1, \xi_2 \in [a,b]$ mit $f(\xi_1) = \sup\{f(x), x \in [a,b]\}$ und $f(\xi_2) = \inf\{f(x), x \in [a,b]\}$.*

Beweis. Der Beweis soll für das Maximum geführt werden. Die Betrachtung von $-f$ liefert den Beweis für das Minimum.
Sei $A := \sup\{f(x), x \in [a,b]\}$. Dabei soll vorläufig auch $A = \infty$ zugelassen werden.
Dann existiert eine Folge $(x_n)_n$ in $[a,b]$, sodaß $\lim_{n\to\infty} f(x_n) = A$ gilt.
Diese hat aufgrund ihrer Beschränktheit durch a und b nach dem Satz von Bolzano-Weierstraß mindestens einen Häufungswert und daher gibt es eine gegen $\xi_1 \in [a,b]$ konvergente Teilfolge $(x_{n_k})_k$.
Wegen der Stetigkeit von f muß damit $A = \lim(f(x_{n_k})) = f(\xi_1)$ sein, d.h. f ist nach oben beschränkt und nimmt ihr Maximum an. □

Bemerkung. Satz 3.10 gilt nicht für offene oder halboffene Intervalle.

Beispiel.
- Die reelle Funktion $f(x) = \frac{1}{x}$ ist im Intervall $(0, 1]$ zwar stetig, jedoch nicht beschränkt.
- Die Funktion $f : (0, 1) \to \mathbb{R}$ mit $f(x) = x$ ist beschränkt, nimmt aber weder ihr Minimum noch ihr Maximum an.

Definition 3.16. Eine Teilmenge K von $\mathbb{R}$ heißt **kompakt**, wenn jede Folge von Punkten in K eine konvergente Teilfolge mit Grenzwert in K enthält.

Beispiel.
- Jedes abgeschlossene Intervall $[a, b] \subset \mathbb{R}$ ist kompakt.
- Offene, nichtleere Intervalle sind nicht kompakt.

Satz. *Die Vereinigung endlich vieler und der Durchschnitt beliebig vieler kompakter Mengen sind wieder kompakt.*

Siehe dazu auch *Übungsaufgabe 38.*

Beispiel. Die Konstruktion des CANTORschen Diskontinuums beginnt mit dem Intervall $C_0 := [0, 1] \in \mathbb{R}$. Dieses wird gedrittelt und die beiden äußeren Teilintervalle werden einschließlich der Ränder zu der Menge $C_1 := C_0 \setminus \left(\frac{1}{3}, \frac{2}{3}\right)$ vereinigt.
Die Menge C_2 entsteht durch Drittelung der beiden Intervalle von C_1 in gleicher Weise: $C_2 := C_1 \setminus \left(\left(\frac{1}{9}, \frac{2}{9}\right) \cup \left(\frac{7}{9}, \frac{8}{9}\right)\right)$. Allgemein entsteht C_i durch Drittelung der Intervalle von C_{j-1}.
Jede der Mengen C_i ist dann die Vereinigung endlich vieler kompakter Mengen. Daher ist auch $C := \bigcap C_i$ kompakt.

Definition 3.17. $f : X \to \mathbb{R}$ heißt **gleichmäßig stetig** auf X, wenn

$$\forall\ \varepsilon > 0\ \exists\ \ \delta >\ 0 :\ |f(x) - f(\tilde{x})| < \varepsilon \quad \forall\, x, \tilde{x} \in X \quad mit \quad |x - \tilde{x}| < \delta$$

Bemerkung. Gleichmäßige Stetigkeit impliziert normale Stetigkeit nach Definition 3.13 im gesamten Definitionsbereich.
Die Umkehrung gilt nicht.

Beispiel. Man betrachte die Funktion $f : (0, 1) \to \mathbb{R}$, wobei $f(x) = \frac{1}{x}$ ist.
Sei nun $\tilde{x} \in (0, 1)$ und $\ \varepsilon > 0$. Mit $\delta = \min\left(\frac{\tilde{x}}{2}, \frac{\tilde{x}^2\varepsilon}{2}\right)$ gilt
$\forall\, x \in (0, 1)$ mit $|x - \tilde{x}| < \delta :\quad \tilde{x} - x \leq |x - \tilde{x}| < \frac{\tilde{x}}{2} \quad \Rightarrow \quad x > \frac{\tilde{x}}{2}$
Und somit ist $|f(x) - f(\tilde{x})| = \left|\frac{1}{x} - \frac{1}{\tilde{x}}\right| = \left|\frac{\tilde{x}-x}{x\tilde{x}}\right| < \frac{2|x-\tilde{x}|}{\tilde{x}^2} < \frac{2\delta}{\tilde{x}^2} \leq \varepsilon$,
d.h. f ist stetig in allen $\tilde{x} \in (0, 1)$.
Jedoch liegt keine gleichmäßige Stetigkeit vor, da beispielsweise für $\varepsilon = 1$ zu jedem $\ \delta >\ 0$ ein $n \in \mathbb{N}$ existiert, sodaß einerseits $\left|\frac{1}{n} - \frac{1}{2n}\right| < \delta$ und andererseits $\left|f\left(\frac{1}{n}\right) - f\left(\frac{1}{2n}\right)\right| = n \geq \varepsilon$ ist.

Die Lage kann sich ändern, wenn der Definitionsbereich ein anderer ist:

Beispiel. Die reelle Funktion $f(x) = \frac{1}{x}$ ist gleichmäßig stetig auf dem abgeschlossenen Intervall $[a, b] \quad (a > 0)$, denn für alle $\varepsilon > 0$ gilt $\forall\, x, \tilde{x} \in [a, b]$ mit $|x - \tilde{x}| < \delta := a^2\varepsilon$: $\quad |f(x) - f(\tilde{x})| = \left|\frac{1}{x} - \frac{1}{\tilde{x}}\right| = \left|\frac{\tilde{x}-x}{x\tilde{x}}\right| < \frac{\delta}{a^2} = \varepsilon$.

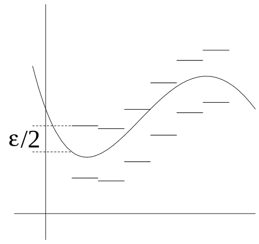
Satz 3.11. *Jede stetige Funktion $f : K \to \mathbb{R}$ mit kompaktem Definitionsbereich K ist auch gleichmäßig stetig.*

Beweis. Wäre f nicht gleichmäßig stetig, dann gäbe es ein $\varepsilon > 0$, sodaß stets zwei $x_n, \tilde{x}_n \in K$ existieren, für die $|x_n - \tilde{x}_n| < \frac{1}{n}$ und $|f(x_n) - f(\tilde{x}_n)| \geq \varepsilon$ gilt. Die Folge $(x_n)_n$ enthält nach Definition der Kompaktheit eine gegen $\tilde{x} \in K$ konvergente Teilfolge $(x_{n_k})_k$. Und weil $|x_{n_k} - \tilde{x}_{n_k}| < \frac{1}{n_k}$ ist, konvergiert auch $(\tilde{x}_{n_k})_k$ gegen $\tilde{x}$.
f ist stetig nach Voraussetzung; daher gilt aufgrund des Folgenkriteriums $\lim_{k\to\infty} (f(x_{n_k}) - f(\tilde{x}_{n_k})) = \lim_{k\to\infty} (f(x_{n_k})) - \lim_{k\to\infty} (f(\tilde{x}_{n_k})) = f(\tilde{x}) - f(\tilde{x}) = 0$, was jedoch einen Widerspruch zu der Annahme $|f(x_{n_k}) - f(\tilde{x}_{n_k})| \geq \varepsilon \ \ \forall\, x_{n_k}, \tilde{x}_{n_k}$ darstellt. □

Eine **Folgerung** aus der gleichmäßigen Stetigkeit einer Funktion ist die Approximierbarkeit dieser mittels stückweise konstanter Funktionen:

Satz 3.12. *Ist die Funktion $f : [a, b] \to \mathbb{R}$ stetig, so gibt es zu jedem $\varepsilon > 0$ Treppenfunktionen $\varphi, \psi : [a, b] \to \mathbb{R}$ mit $\varphi(x) \leq f(x)$ und $\psi(x) \geq f(x)$, sodaß $|\varphi(x) - \psi(x)| \leq \varepsilon \quad \forall\, x \in [a, b]$.*

Beweis. Zunächst sei bemerkt, daß f wegen der Kompaktheit von [a,b] auch gleichmäßig stetig ist, d.h. zu dem gegebenen $\varepsilon > 0$
$\exists\ \ \delta > \ 0 : \ |f(x) - f(\tilde{x})| < \frac{\varepsilon}{2} \quad \forall\, x, \tilde{x} \in [a, b]$ mit $|x - \tilde{x}| < \delta$.
Nun wird $[a, b]$ in n äquidistante Teilintervalle der Länge $\frac{b-a}{n}$ zerlegt, wobei n so groß gewählt wird, daß die Gitterweite $\frac{b-a}{n} < \delta$ ist.
Die Zahlen $t_k = a + k\frac{b-a}{n}, \quad k = 0 \ldots n$ sind dann die Ränder der Teilintervalle. Sei für $t_{k-1} < x \leq t_k$: $\ \varphi(x) = f(t_k) + \frac{\varepsilon}{2}$ und $\psi(x) = f(t_k) - \frac{\varepsilon}{2}$.
Die so definierten Treppenfunktionen erfüllen offensichtlich $|\varphi(x) - \psi(x)| \leq \varepsilon$.
Überdies ist $\varphi(x) = \psi(x) = f(x)$, wenn $x = a = t_0$ und für die übrigen x aus einem Intervall $(t_{k-1}, t_k]$ gilt $|x - t_k| < \delta$ und damit
$-\frac{\varepsilon}{2} < f(x) - f(t_k) < \frac{\varepsilon}{2} \quad \Rightarrow \quad \varphi(x) = f(t_k) - \frac{\varepsilon}{2} < f(x) < f(t_k) + \frac{\varepsilon}{2} = \psi(x)$. □



Definition 3.18. Eine Menge $X \subset \mathbb{R}$ heißt **offen**, falls

$$\forall\, x \in X \;\; \exists \;\varepsilon > 0 : \quad U_\varepsilon(x) \subset X$$

Eine Menge X heißt **abgeschlossen**, wenn jede CAUCHY-Folge in X einen Grenzwert in X besitzt.

Beispiel. Sind $a, b \in \mathbb{R}$ und ist $a < b$, so ist

- das offene Intervall (a, b) offen,
- das abgeschlossene Intervall $[a, b]$ abgeschlossen
- und das halboffene Intervall $(a, b]$ weder offen noch abgeschlossen.

$\emptyset$ und $\mathbb{R}$ sind sowohl offen als auch abgeschlossen.

Satz 3.13. *Die Menge $X \subset \mathbb{R}$ ist genau dann abgeschlossen, wenn ihr Komplement $\mathbb{R} \backslash X$ offen ist.*

Beweis: Übungsaufgabe 38a.

Satz 3.14. *Eine Menge $K \subset \mathbb{R}$ ist genau dann kompakt, wenn sie abgeschlossen und beschränkt ist.*

Beweis. Sei K abgeschlossen und beschränkt. Dann ist auch jede Folge $(x_n)_n$ aus K beschränkt. Diese muß nach dem Satz von BOLZANO-WEIERSTRAß und Satz 2.11 eine konvergente Teilfolge $(x_{n_k})_k$ enthalten, deren Grenzwert nach Definition der Abgeschlossenheit in K liegt, d.h. K ist kompakt.
In der umgekehrten Richtung soll die Kontraposition der zu zeigenden Aussage bewiesen werden, d.h. ist K nicht abgeschlossen oder nicht beschränkt, dann auch nicht kompakt.
Sei zunächst K nicht beschränkt. Dann gibt es eine Folge (x_n) in K, sodaß $|x_n| > n$ ist. Für jede Teilfolge $(x_{n_k})_k$ ist dann auch $|x_{n_k}| > n_k$, d.h. jede Teilfolge divergiert. Also besitzt nicht jede Folge in K eine konvergente Teilfolge; K ist nicht kompakt.
Sei schließlich K nicht abgeschlossen. Dann existiert eine CAUCHY- Folge (x_n) in K, deren Grenzwert g nicht in K liegt. Jede Teilfolge von (x_n) hat auch den Grenzwert $g \notin K$. Somit ist K nicht kompakt. □

Bemerkung. Neben der Folgenkompaktheit von Definition 3.16 gibt es noch die Überdeckungskompaktheit. Der Überdeckungssatz von Heine-Borel[17] besagt, daß in $\mathbb{R}$ beide äquivalent zueinander sind.

Satz 3.15 (Verallgemeinerung von 3.10). *Jede über einer nichtleeren, kompakten Menge stetige reelle Funktion ist beschränkt und nimmt ihr Minimum und Maximum an.*

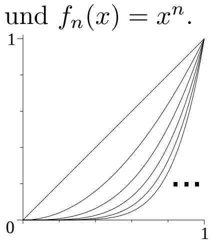
Beweis. Wie jener von Satz 3.10, wobei die Existenz der konvergenten Teilfolge durch die Kompaktheit gesichert wird. □

3.4 Funktionenfolgen und -reihen

Man betrachte die Folge $(f_n)_{n\in\mathbb{N}}$ von Funktionen $f_n : X \to \mathbb{R}$.

Definition 3.19. Die Funktionenfolge (f_n) **konvergiert punktweise** gegen die Funktion $f : X \to \mathbb{R}$, wenn $\forall\, x \in X$ gilt: $\lim_{n\to\infty} f_n(x) = f(x)$.
In diesem Falle schreibt man $f_n \to f$.

Beispiel. Sei (f_n) eine Folge von Funktionen mit dem Definitionsbereich $[0, 1]$ und $f_n(x) = x^n$.



Sie konvergiert gegen $f(x) = \begin{cases} 0 & \text{für } x < 1 \\ 1 & \text{für } x = 1 \end{cases}$
Die einzelnen f_n sind stetig, nicht jedoch f. Die punktweise Konvergenz der Folge impliziert also nicht die Stetigkeit des Grenzwertes.

Definition 3.20. Die Funktionenfolge $(f_n)_{n\in\mathbb{N}}$ **konvergiert gleichmäßig** in X, falls $\forall\ \varepsilon > 0\ \exists\ N \in \mathbb{N},$ sodaß $\forall\, n \geq N$ und $\forall\, x \in X$ gilt: $|f(x) - f_n(x)| < \varepsilon$.
Dann schreibt man $f_n \rightrightarrows f$.

Im Unterschied zur punktweisen Konvergenz ist bei der gleichmäßigen Konvergenz das N, für das $|f_n(x) - f(x)| < \varepsilon\ \ \forall\, n \geq N$, nicht vom Argument x abhängig.
Analog dazu wurde definiert, daß eine Funktion f stetig in X ist, wenn
$\forall\, x \in X,\ \forall\ \varepsilon > 0\ \exists\ \ \delta >\ 0,$ sodaß
$\forall\, y \in X :\ |x - y| < \delta \quad \Rightarrow \quad |f(x) - f(y)| < \varepsilon$

[17] *Eduard Heine* (1821-1881), *Emile Borel* (1871-1956)

und f ist gleichmäßig stetig, wenn
$\forall \;\; \varepsilon > 0 \; \exists \;\; \delta > \; 0$, sodaß $\forall\, x, y \in X: \;\; |x-y| < \delta \quad \Rightarrow \quad |f(x) - f(y)| < \varepsilon$.
Bei gleichmäßiger Stetigkeit ist also das δ aus obiger Definition nicht vom Argument x abhängig.

Satz 3.16. *Sind die Funktionen $f_n : X \to \mathbb{R}$ stetig und die Folge $(f_n)_n$ konvergiert gleichmäßig gegen f, so ist auch f stetig.*

Beweis. Sei $x \in X$ und $\varepsilon > 0$.
Sei überdies N so groß gewählt, daß $\forall\, y \in X: \;\; |f_N(y) - f(y)| < \frac{\varepsilon}{3}$.
Dies ist möglich wegen der gleichmäßigen Konvergenz von $(f_n)_n$.
Weiterhin läßt sich aufgrund der Stetigkeit von f_N ein $\delta > 0$ finden, sodaß $|f_N(y) - f_N(x)| < \frac{\varepsilon}{3} \quad \forall\, y \in X$ mit $|x-y| < \delta$.
Damit folgt für alle $y \in X$ mit $|x-y| < \delta$:
$|f(x) - f(y)| \leq |f(x) - f_N(x)| + |f_N(x) - f_N(y)| + |f_N(y) - f(y)| < \varepsilon$. □

Satz 3.17. *Ist (f_n) eine gleichmäßig konvergente Folge gleichmäßig stetiger Funktionen, dann ist auch ihr Grenzwert f gleichmäßig stetig.*

Beweis. Wie Satz 3.16. □

Analoge Sätze lassen sich für Funktionenreihen formulieren.

Definition 3.21. Sei V ein reeller Vektorraum. Eine **Norm** auf V ist eine Abbildung $||\cdot|| : V \to \mathbb{R}$, für die gilt:

a) $||v|| \geq 0 \;\; \forall\, v \in V$. Gleichheit gilt nur für $v = 0$.

b) $\forall\, v \in V$, $c \in \mathbb{R}$ ist $||cv|| = |c| \, ||v||$.

c) $\forall\, x, y \in V$ ist $||v + w|| \leq ||v|| + ||w||$ (Dreiecksungleichung)

Beispiel.

- Der Betrag $|\cdot|$ einer reellen Zahl ist eine Norm auf dem reellen Vektorraum $\mathbb{R}$.
- Die euklidische Norm auf $\mathbb{R}^n$ ist definiert als $|(x_1, \ldots, x_n)| = \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2}$.

Definition 3.22. Sei $X \subset \mathbb{R}$. Der $\mathbb{R}$-Vektorraum der stetigen Funktionen auf X wird bezeichnet mit $\mathcal{C}^0(X) = \{f : X \to \mathbb{R} \mid f \text{ ist stetig auf } X\}$.
Sei $||\cdot||_X : \mathcal{C}^0(X) \to \mathbb{R} \cup \{\infty, -\infty\}$ eine Abbildung mit

$$||f||_X = \sup\{|f(x)|, \, x \in X\}$$

Lemma 3.18. *Ist K eine nichtleere, kompakte Teilmenge von $\mathbb{R}$, so ist $||\cdot||_K$ wie in Definition 3.22 eine Norm auf $\mathcal{C}^0(X)$. Insbesondere ist* $||f||_K \neq \pm\infty \;\; \forall f \in \mathcal{C}^0(X)$.

Beweis. Ist f stetig, so auch $|f|$ nach Satz 3.7. Mit Satz 3.15 nimmt $|f|$ ihr Minimum und Maximum auf K an, ist also beschränkt und
$||f||_K \neq \pm\infty \;\; \forall f \in \mathcal{C}^0(X)$.
Die übrigen Eigenschaften einer Norm können durch Nachrechnen verifiziert werden. □

Bemerkung. Ist X nicht kompakt, dann ist $||f||_X = \infty$ für einige $f \in \mathcal{C}^0$.

Damit ist eine Umformulierung der gleichmäßigen Konvergenz von Funktionenfolgen möglich: (f_n) heißt gleichmäßig konvergent, wenn $||f_n - f||_X \to 0$ für $n \to \infty$.

Satz 3.19 (Weierstraßsches Konvergenzkriterium). *Sei $X \subset \mathbb{R}$ und $f_n : X \to \mathbb{R}$. Falls $\sum_{n=1}^{\infty} ||f_n||_X$ existiert, dann konvergiert $\sum_{n=1}^{\infty} f_n$ absolut und gleichmäßig.*

Beweis. Es ist $\sum_{n=1}^{\infty} ||f_n||_X$ eine Majorante zu jeder Reihe $\sum_{n=1}^{\infty} f_n(x)$, denn $|f_n(x)| \leq ||f_n||_X$ per definitionem, d.h. $\sum_{n=1}^{\infty} f_n$ konvergiert punktweise absolut. Sei f der Grenzwert: $f(x) = \sum_{n=1}^{\infty} f_n(x)$.
Weil $\sum_{n=1}^{\infty} ||f_n||_X$ konvergiert, gibt es $\forall \;\; \varepsilon > 0$ ein $N \in \mathbb{N}$, sodaß für $M \geq N$ gilt: $\sum_{n=M+1}^{\infty} ||f_n||_X < \varepsilon$. Damit und mit der verallgemeinerten Dreiecksungleichung für absolut konvergente Reihen ergibt sich
$|f(x) - \sum_{n=1}^{M} f_n(x)| = |\sum_{n=M+1}^{\infty} f_n(x)| \leq \sum_{n=M+1}^{\infty} |f_n(x)|$
$\leq \sum_{n=M+1}^{\infty} ||f_n||_X < \varepsilon$. Dies gilt $\forall x \in X$; daher ist $\sum_{n=1}^{\infty} f_n$ gleichmäßig konvergent. □

Satz 3.20. *Sei $P(x) = \sum_{n=0}^{\infty} a_n x^n$ eine Potenzreihe mit dem Konvergenzradius R. Dann ist $P|_{(-R,R)}$ stetig.*

Beweis. Sei $r \in (0, R)$, $K = [-r, r]$ und $f_n : K \to \mathbb{R}$ mit $f_n(x) = a_n x^n$.
Für jedes $x \in X$ gilt $|f_n(x)| = |a_n x^n| \leq ||f_n||_K = |a_n r^n|$.
Nach Satz 2.29 konvergiert P in r absolut, also auch $\sum_{n=0}^{\infty} ||f_n||_K$.
Mit dem Weierstraßschen Konvergenzkriterium folgt, daß $\sum_{n=0}^{\infty} f_n$ gleichmäßig gegen eine Funktion f konvergiert. Diese ist stetig nach Satz 3.16, da alle f_n und somit auch alle endlichen Summen von f_n stetig sind.
Damit ist $P|_{[-r,r]}$ stetig für alle $r \in (0, R)$, d.h. $P|_{(-R,R)}$ ist stetig. □

Folgerung. $\exp(x)$ *ist stetig auf* $\mathbb{R}$.

3.5 Grenzwerte von Funktionen

Definition 3.23. Eine Zahl $x_0 \in \mathbb{R}$ heißt Häufungspunkt einer Menge X, wenn jede Umgebung von x_0 unendlich viele Punkte von X enthält.

Beispiel.

- Die Menge der Häufungspunkte von (a, b) ist $[a, b]$.
- Die Menge der Häufungspunkte von $\mathbb{Q}$ ist $\mathbb{R}$.
- Der einzige Häufungspunkt von $\{\frac{1}{n}\}$ ist 0.

Definition 3.24. Sei $f : X \to \mathbb{R}$ und $x_0 \in \mathbb{R}$. Eine in x_0 stetige Funktion, welche auf $X \backslash \{x_0\}$ mit f übereinstimmt, heißt **stetige Fortsetzung** von f in x_0.

Bemerkung. Der Begriff der stetigen Fortsetzung ist nur dann interessant, wenn x_0 ein Häufungspunkt von X ist, da sonst jede Funktion, die auf $X \backslash \{x_0\}$ mit f übereinstimmt und in x_0 irgendeinen Wert hat, eine stetige Fortsetzung von f ist.

Satz 3.21. *Ist x_0 Häufungspunkt von X, so hat jede Funktion f auf $X \backslash \{x_0\}$ höchstens eine stetige Fortsetzung in x_0.*

Definition 3.25. Die Funktion $f : X \to \mathbb{R}$ hat im Häufungspunkt x_0 von X den **Grenzwert** a, wenn es zu jedem $\varepsilon > 0$ ein $\delta > 0$ gibt, sodaß
$|f(x) - a| < \varepsilon$ für alle $x \in X \backslash \{x_0\}$ mit $|x - x_0| < \delta$.
In diesem Fall schreibt man $\lim_{x \to x_0} f(x) = a$ oder $f(x) \to a$ für $x \to x_0$.

Bemerkung.

- Für Grenzwerte von Funktionen gelten die üblichen Rechenregeln: wenn $f(x) \to a$ und $g(x) \to b$, so gilt
  $f(x) \pm g(x) \to a \pm b$,
  $f(x) \cdot g(x) \to ab$ und
  $\frac{f(x)}{g(x)} \to \frac{a}{b}$, falls $b \neq 0$.
- Wenn $f(x) \to a$ für $x \to x_0$ und die Funktion g in a stetig ist, so geht $(g \circ f)(x) \to g(a)$ für $x \to x_0$.
- Aus $f \leq g$ folgt $\lim_{x \to x_0} f(x) \leq \lim_{x \to x_0} g(x)$, falls die Grenzwerte existieren.

Satz 3.22 (Folgenkriterium). *Die Funktion $f : X \to \mathbb{R}$ hat in x_0 den Grenzwert a, wenn für jede Folge (x_n) in $X \backslash \{x_0\}$ mit $(x_n) \to x_0$ gilt:*
$\lim_{x \to x_0} f(x_n) = a$.

Satz 3.23 (Cauchy-Kriterium für Funktionen). *$f : X \to \mathbb{R}$ ist konvergent in x_0 genau dann, wenn $\forall\ \varepsilon > 0\ \exists\ U^*(x_0)$, sodaß $|f(x) - f(x')| < \varepsilon\ \ \forall\, x, x' \in U^*(x_0) \cap X$, wobei $U^*(x_0)$ eine Umgebung $U(x_0)$ ohne den Punkt x_0 ist.*

Definition 3.26. Die Funktion $f : X \to \mathbb{R}$ hat in x_0 den **linksseitigen (rechtsseitigen) Grenzwert** a, wenn $\forall\ \ \varepsilon > 0\ \exists\ \ \delta >\ 0 :\ \ |f(x) - a| < \varepsilon$ für alle $x \in X \cap (x_0 - \delta, x_0)$ (bzw. $x \in X \cap (x_0, x_0 + \delta)$).
Für den linksseitigen Grenzwert schreibt man $a = \lim_{x \nearrow x_0} f(x)$ und für den rechtsseitigen $a = \lim_{x \searrow x_0} f(x)$.

Beispiel. Die Funktion $f(x) = \lfloor x \rfloor$ hat in $g \in \mathbb{Z}$ den linksseitigen Grenzwert $g - 1$ und den rechtsseitigen Grenzwert g. Daraus folgt insbesondere, daß f unstetig in allen $g \in \mathbb{Z}$ ist.

Bemerkung. Die Rechenregeln und Existenzkriterien für beidseitige Grenzwerte gelten analog auch für einseitige.

Definition 3.27. Sei $f : X \to \mathbb{R}$ und X nicht nach oben beschränkt. Dann heißt a **Grenzwert** von f **bei** ∞, wenn $\forall\ \ \varepsilon > 0\ \exists\ N \in \mathbb{R}$, sodaß $|f(x) - a| < \varepsilon\ \ \forall\, x \in X$ mit $x > N$. In Zeichen: $a = \lim_{x\to\infty} f(x)$ oder $f(x) \to a$ für $x \to \infty$.
Analog läßt sich der Grenzwert von f bei $-\infty$ definieren.

Bemerkung. Der Grenzwert einer Funktion bei ∞ verallgemeinert den Begriff des Grenzwertes einer Folge, denn die Folge $(a_n)_{n\in\mathbb{N}}$ ist eine Funktion $\mathbb{N} \to \mathbb{R}$.

Beispiel. $\lim_{x\to\infty} \left(\sqrt{x+1} - \sqrt{x}\right) = 0$, denn für $x > N = \frac{1}{4\varepsilon^2} > 0$ gilt $\sqrt{x+1} - \sqrt{x} = \frac{1}{\sqrt{x+1}+\sqrt{x}} < \frac{1}{2\sqrt{x}} < \varepsilon$.

Bemerkung. Auch für die Grenzwerte von Funktionen bei $\pm\infty$ bleiben die Rechenregeln und Konvergenzkriterien erhalten.

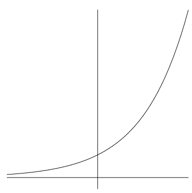
Definition 3.28. Die Funktion $f : X \to \mathbb{R}$ hat in $x_0 \in \bar{\mathbb{R}} = \mathbb{R} \cup \{\infty, -\infty\}$ den **uneigentlichen Grenzwert** ∞ $(-\infty)$, falls $\forall\, K \in \mathbb{R}\ \ \exists\ U^*(x_0)$, sodaß $f(x) > K$ (bzw. $f(x) < K$) für alle $x \in U^*(x_0) \cap X$.
Die Schreibweise ist $\lim_{x\to x_0} f(x) = \infty$.

3.6 Spezielle Funktionen

Satz 3.24. *Die Exponentialfunktion* $\exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!}$ *hat den Wertebereich $\mathbb{R}^+$.* $\exp : \mathbb{R} \to \mathbb{R}^+$ *ist bijektiv und streng monoton wachsend.*

Beweis. Für $x \geq 0$ ist $\exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!} > 0$. Mit $1 = \exp(0) = \exp(x - x) = \exp(x)\exp(-x)$ folgt auch $\exp(-x) > 0$, also

besteht der Wertebereich von $\exp(x)$ nur aus positiven reellen Zahlen.
Sei $\tilde{x} > x$. Dann läßt sich $\tilde{x}$ schreiben als $x + h$, wobei $h > 0$. Wegen letzterem ist $\exp(h) > 1$ und somit $\exp(\tilde{x}) = \exp(x + h) = \exp(x)\exp(h) > \exp(x)$, d.h. $\exp(x)$ ist streng monoton wachsend und mithin injektiv.
Weiterhin gilt $\exp(x) \geq 1 + x$ für $x \geq 0$ und damit $\lim_{x\to\infty} \exp(x) = \infty$, und mit $\exp(-x) = \frac{\exp(0)}{\exp(x)}$ folgt $\lim_{x\to-\infty} \exp(x) = \lim_{x\to\infty} \frac{1}{\exp(x)} = 0$.
Zusammen mit dem Zwischenwertsatz liefert dies die Surjektivität von $\exp(x) : \mathbb{R} \to \mathbb{R}^+$. □



Definition 3.29. Die Umkehrfunktion von $\exp(x)$ heißt der **natürliche Logarithmus** $\ln : \mathbb{R}^+ \to \mathbb{R}$.

Satz 3.25 (Funktionalgleichung des Logarithmus). *Für $xy > 0$ gilt* $\ln(xy) = \ln(x) + \ln(y)$.

Beweis. Sei $\xi = \ln(x)$ und $\eta = \ln(y)$, d.h. $\exp(\xi) = x$ und $\exp(\eta) = y$.
$\Rightarrow \quad \ln(xy) = \ln(\exp(\xi)\exp(\eta)) = \ln(\exp(\xi + \eta)) = \xi + \eta = \ln(x) + \ln(y)$ □

Bisher wurden Potenzen a^b erklärt für $a \in \mathbb{R}$ und
natürliche Exponenten $b \in \mathbb{N}$ durch $a^b = a \cdot a \ldots a$,
für ganzzahlige Exponenten $b \in \mathbb{Z}$ unter Zuhilfenahme von $a^{-b} = \left(a^b\right)^{-1} = \frac{1}{a^b}$
und für rationale Exponenten $b = \frac{p}{q} \in \mathbb{Q}$ durch $a^b = \sqrt[q]{a^p}$.
Aufgrund des Additionstheorems für die Exponentialfunktion ist außerdem $e^{\frac{p}{q}} = exp(\frac{p}{q})$ mit $e = exp(1)$.

Definition 3.30. Für $x \in \mathbb{R}$ und $a \in \mathbb{R}^+$ ist $a^x = \exp(x\ln(a))$

Man betrachte die Exponentialfunktion $\exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!}$. Es ist

$$\lim_{x\to 0} \frac{\exp(x) - 1}{x} = 1\,,$$

denn $\exp(x) = 1 + x + R_2(x)$ mit $|R_2(x)| \leq \frac{2|x|^2}{2!} = |x|^2$.
$\Rightarrow \quad |\exp(x) - x - 1| \leq |x|^2 \quad \Rightarrow \quad \left|\frac{\exp(x)-1}{x} - 1\right| \leq |x|$, woraus die Behauptung folgt.

Weiterhin gilt

$$\lim_{x\to 0} \frac{\ln(1+x)}{x} = 1\,.$$

Satz 3.26. *$\forall\, n \in \mathbb{N}$ ist $\lim_{x\to\infty} \frac{\exp(x)}{x^n} = \infty$ und $\lim_{x\to-\infty} \frac{\exp(x)}{x^{-n}} = 0$, d.h. für $x \to \infty$ wächst die Exponentialfunktion schneller als jede positive Potenzfunktion und für $x \to -\infty$ fällt sie schneller als jede negative Potenzfunktion.*

Beweis. Wegen $\exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!}$ gilt für $x > 0$:
$\exp(x) > \frac{x^{n+1}}{(n+1)!} \quad \Rightarrow \quad \frac{\exp(x)}{x^n} > \frac{x}{(n+1)!}$ und somit $\lim_{x\to\infty} \frac{\exp(x)}{x^n} = \infty$.
Die Substitution $y = -x$ liefert zudem
$\lim_{x\to-\infty} \frac{\exp(x)}{x^{-n}} = \lim_{y\to\infty} \frac{y^n}{exp(y)} = 0$. □

Satz 3.27.

$$\lim_{x\to\infty} \frac{\ln(x)}{\sqrt[n]{x}} = 0\,,$$

d.h. der Logarithmus wächst schwächer als jede Wurzelfunktion.

Definition 3.31. Die EULERsche Zahl ist

$$e := \exp(1) = \sum_{n=0}^{\infty} \frac{1}{n!} = \lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^n$$

Für die schnelle Berechnung von e^x wird das Argument getrennt:
$x = \lfloor x \rfloor + (x - \lfloor x \rfloor) = g + \vartheta$, wobei g eine ganze und ϑ eine positive Zahl kleiner 1 ist.
Damit wird $e^x = e^g e^\vartheta$. Potenzen mit ganzzahligen Exponenten lassen sich schnell berechnen, wenn die Basis e gegeben ist, und für
$\exp(\vartheta) = \sum_{k=0}^{n} \frac{x^k}{k!} + R_{n+1}(\vartheta)$ gilt die Restgliedabschätzung
$|R_{n+1}(\vartheta)| \le 2\frac{\vartheta^{n+1}}{(n+1)!} < \frac{2}{(n+1)!}$, d.h. der Fehler wird (bis auf den Faktor 2) durch den ersten fehlenden Summanden bestimmt und fällt überdies sehr schnell.

Beispiel. Bei der Berechnung von e mittels der Exponentialreihe ist das Restglied $|R_{n+1}| < \frac{2}{(n+1)!}$. Somit ergibt sich z.B. für $n = 10$
$R_{11} < \frac{2}{11!} \approx 6 \cdot 10^{-8}$, d.h. eine Genauigkeit von etwa 8 Dezimalstellen.

Bemerkung. Die Berechnung von e mit Hilfe der Exponentialreihe hat gegenüber der mittels der Folge $\left(1 + \frac{1}{n}\right)^n$ den Vorteil einer sehr schnellen Konvergenz. Die Reihe konvergiert in der Größenordnung $\frac{1}{n!}$, die Folge mit $\frac{1}{n}$. Siehe auch *Programmieraufgabe 2.*

Eine weitere Anwendung der Restgliedabschätzung ist der folgende Satz.

Satz 3.28. *e ist irrational.*

Beweis. Wäre e rational, so gäbe es teilerfremde $p \in \mathbb{Z}$ und $q \in \mathbb{N}$ mit $e = \frac{p}{q}$.
Außerdem ist dann $q \ge 2$, da e offensichtlich nicht ganzzahlig ist.
$\Rightarrow \quad e \cdot q! \in \mathbb{Z}$ und somit auch $q!\left(e - 1 - \frac{1}{1!} - \frac{1}{2!} - \ldots - \frac{1}{q!}\right) = q! R_{q+1} \in \mathbb{Z}$.

Dabei ist R_{q+1} das entsprechende Restglied der Reihe, für welches jedoch $R_{q+1} < \frac{2}{(q+1)!}$ gilt, also $q!R_{q+1} < \frac{2}{q} \leq 1$, was $q!R_{q+1} \in \mathbb{Z}$ widerspricht. $\square$

Im Folgenden sollen noch einige Grenzwerte von speziellen Funktionen angegeben werden.

(1) $\lim_{a\to\infty} a^x = \begin{cases} \infty & \text{für } x > 0 \\ 0 & \text{für } x < 0 \end{cases}$,
denn $a^x = \exp(x\ln(a))$ wächst für $x > 0$ monoton und der Wertebereich ist nicht nach oben beschränkt.
Für $x < 0$ ist $a^x = \frac{1}{a^y}$ mit $-x = y > 0$, also $\lim_{a\to\infty} a^x = 0$

(2) $\lim_{a\to 0} a^x = \begin{cases} 0 & \text{für } x > 0 \\ \infty & \text{für } x < 0 \end{cases}$

(3) $\lim_{a\to\infty} \frac{\ln(a)}{a^x} = 0$ für $x > 0$, denn $\exists\, n \in \mathbb{N}: \ \frac{1}{n} \leq x$, womit ab $a \geq 1$ gilt: $0 \leq \frac{\ln(a)}{a^x} \leq \frac{\ln(a)}{\sqrt[n]{a}}$ und nach Satz 3.27 geht $\frac{\ln(a)}{\sqrt[n]{a}} \to 0$.

(4) $\lim_{a\to 0} a^x \ln(a) = 0$ für $x > 0$.

(5) $\lim_{x\to\infty} \left(1 + \frac{c}{x}\right)^x = e^c$ mit $c \in \mathbb{R}$.
Zur Begründung wird der Folgengrenzwert $\lim_{n\to\infty} \left(1 + \frac{c}{n}\right)^n = e^c$ verallgemeinert:
$\lim_{x\to\infty} \left(1 + \frac{c}{x}\right)^x = \lim_{\zeta\to 0} (1 + c\zeta)^{\frac{1}{\zeta}} = \lim_{\zeta\to 0} \exp\left(\frac{\ln(1+c\zeta)}{\zeta}\right)$
Und weil die Exponentialfunktion stetig in $\mathbb{R}$ ist, ist dies gleich
$\exp\left(\lim_{\zeta\to 0} \frac{1+c\zeta}{\zeta}\right) = \exp\left(c \cdot \lim_{y\to 0} \frac{1+y}{y}\right) = \exp(c)$.
Dabei ist $\zeta = \frac{1}{x}$ und $y = c\zeta$.

Satz 3.29. $\forall\, s \in \mathbb{R}$, $x \in (-1, 1)$ *gilt*

$$(1+x)^s = \sum_{n=0}^{\infty} \binom{s}{n} x^n = 1 + sx + \binom{s}{2} x^2 + \dots$$

und

$$\ln(1+x) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} x^n = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \dots$$

Bemerkung. Die Logarithmusreihe divergiert für $x > 1$, obgleich die Logarithmusfunktion dort definiert ist.
Im Falle $x = 1$ konvergiert sie und demnach ist
$\ln(2) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} = 1 - \frac{1}{2} + \frac{1}{3} - \dots$.
Gregory[18] fand 1668 eine Möglichkeit, Logarithmen für beliebige $y > 0$ zu berechnen: Die Reihen für $\ln(1-x)$ und $\ln(1+x)$ konvergieren absolut für $x \in (-1, 1)$. Somit ist
$\ln\left(\frac{1+x}{1-x}\right) = \ln(1+x) - \ln(1-x) = 2\sum_{n=1}^{\infty} \frac{x^{2n-1}}{2n-1} = 2\left(x + \frac{x^3}{3} + \frac{x^5}{5} + \dots\right)$.

[18] *James Gregory* (1638-1675)

Beispiel. $2 = \frac{1+\frac{1}{3}}{1-\frac{1}{3}} \quad \Rightarrow \quad \ln(2) = 2\left(\frac{1}{3} + \frac{1}{3}\left(\frac{1}{3}\right)^3 + \dots\right)$.
Dabei liefern 6 Summanden bereits ca. 6 Dezimalstellen.

Bisher wurden Folgen und Reihen in $\mathbb{R}$ betrachtet. Die Überlegungen dazu lassen sich weitgehend auf die komplexen Zahlen $\mathbb{C}$ übertragen.

Definition 3.32. Die Menge $\mathbb{R} \times \mathbb{R}$ zusammen mit den Verknüpfungen $+$ und $\cdot$, definiert durch

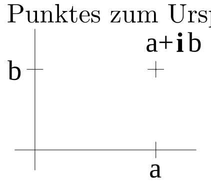
$$(a, b) + (c, d) = (a + c, b + d)$$

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

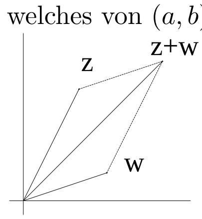
heißt Menge der **komplexen Zahlen** $\mathbb{C}$. Anstelle von (a, b) schreibt man $a + \mathbf{i}b$, wobei $a, b \in \mathbb{R}$ sind.
$\mathrm{Re}\,(a + \mathbf{i}b) = a$ ist der **Realteil**, $\mathrm{Im}\,(a + \mathbf{i}b) = b$ der **Imaginärteil** von $a + \mathbf{i}b$.
$\overline{z} = a - \mathbf{i}b$ ist die zu $z = a + \mathbf{i}b$ **konjugiert komplexe Zahl**.
Der **Betrag** einer komplexen Zahl $z = a + \mathbf{i}b$ wird definiert als $|z| := \sqrt{z\overline{z}} = \sqrt{a^2 + b^2}$.

Bemerkung. $\mathbb{C}$ bildet einen nicht anordnenbaren Körper (vgl. *Übungsaufgabe 15*).
Der Betrag einer komplexen Zahl ist eine Norm auf $\mathbb{C}$.

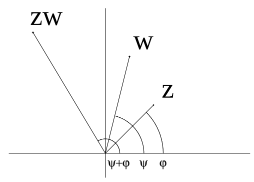
Komplexe Zahlen lassen sich mit der Gaußschen Zahlenebene veranschaulichen. Dabei repräsentiert ein Punkt (a, b) der Ebene $\mathbb{R}^2$ die komplexe Zahl $a + \mathbf{i}b$. Der Betrag der Zahl entspricht dem Abstand des Punktes zum Ursprung.



Die Summe zweier Zahlen $z = a + \mathbf{i}b$ und $w = c + \mathbf{i}d$ ist dann der vom Ursprung verschiedene Endpunkt einer Diagonalen des Parallelogramms, welches von $(a, b), (c, d) \in \mathbb{R}^2$ aufgespannt wird:



Zur Interpretation der Multiplikation zweier komplexer Zahlen z und w betrachte man die Polarkoordinaten (r_1, φ) und (r_2, ψ) von z bzw. w.
Das Produkt zw hat dann die Koordinaten $(r_1 r_2, \varphi + \psi)$.



Definition 3.33. Sei $x \in \mathbb{R}$. Dann ist

$$\cos(x) := \operatorname{Re}(\exp(\mathbf{i}x)) = \frac{\exp(\mathbf{i}x) + \exp(-\mathbf{i}x)}{2}$$

$$\sin(x) := \operatorname{Im}(\exp(\mathbf{i}x)) = \frac{\exp(\mathbf{i}x) - \exp(-\mathbf{i}x)}{2\mathbf{i}}$$

Aus der Definition folgt

$$\exp(\mathbf{i}x) = \cos(x) + \mathbf{i}\sin(x)$$

Sodann ergeben sich Additionstheoreme für den Sinus und Cosinus:
$\cos(x+y) + \mathbf{i}\sin(x+y) = \exp(\mathbf{i}(x+y)) = \exp(\mathbf{i}x)\exp(\mathbf{i}y) =$
$(\cos(x) + \mathbf{i}\sin(x))(\cos(y) + \mathbf{i}\sin(y)) =$
$(\cos(x)\cos(y) - \sin(x)\sin(y)) + \mathbf{i}(\sin(x)\cos(y) + \sin(y)\cos(x))$

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

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

Überdies sind Sinus- und Cosinusfunktion periodisch (*Übungsaufgabe 45)*:
$\cos(x + 2\pi) = \cos(x)$ und $\sin(x + 2\pi) = \sin(x)$
Weiterhin gilt $\cos(-x) = \cos(x)$ und $\sin(-x) = -\sin(x)$
und damit

$$1 = \cos(0) = \cos(x - x) = \cos(x)\cos(-x) - \sin(x)\sin(-x) = \cos^2(x) + \sin^2(x)$$

Die Reihendarstellung von Sinus und Cosinus erhält man aus der Exponentialreihe. Nach dem großen Umordnungssatz ist
$\exp(\mathbf{i}x) = \sum_{n=0}^{\infty} \frac{(\mathbf{i}x)^n}{n!} = \sum_{n=0}^{\infty} \frac{(\mathbf{i}x)^{2n}}{(2n)!} + \sum_{n=0}^{\infty} \frac{(\mathbf{i}x)^{2n+1}}{(2n+1)!} =$
$\sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!} + \mathbf{i}\sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!}$.

$$\Rightarrow \qquad \cos(x) = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \dots$$

$$\sin(x) = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \dots$$

Bemerkung. Die Sinus- und die Cosinusreihe konvergieren absolut für $\forall\, x \in \mathbb{R}$.

Mit Hilfe der Exponentialreihe erhält man eine Abschätzung für die Restglieder:
$\cos(x) = \sum_{k=0}^{n} (-1)^k \frac{x^{2k}}{(2k)!} + R_{2n+2}(x)$ mit $|R_{2n+2}(x)| \leq \frac{|x|^{2n+2}}{(2n+2)!}$ für $|x| \leq 2n+3$ und
$\sin(x) = \sum_{k=0}^{n} (-1)^k \frac{x^{2k+1}}{(2k+1)!} + R_{2n+3}(x)$ mit $|R_{2n+3}(x)| \leq \frac{|x|^{2n+3}}{(2n+3)!}$ für $|x| \leq 2n+4$

Bemerkung. Diese Restgliedabschätzung gilt sogar für alle $x \in \mathbb{R}$, was später gezeigt werden soll.

Folgerung 3.30.
$$\lim_{x\to 0} \frac{\sin(x)}{x} = 1$$

Beweis. $\sin(x) = x + R_3(x)$ mit $|R_3(x)| \leq \frac{|x|^3}{3!}$
$\Rightarrow \quad |\sin(x) - x| \leq \frac{|x|^3}{6}$
$\Rightarrow \quad |\frac{\sin(x)}{x} - 1| \leq \frac{|x|^2}{6}$
und mit $\lim_{x\to 0} \frac{|x|^2}{6} = 0$ folgt die Behauptung. □

Definition 3.34. Für $x \in \mathbb{R}\backslash\{\frac{\pi}{2} + k\pi\,,\, k \in \mathbb{Z}\}$ ist $\tan(x) = \frac{\sin(x)}{\cos(x)}$.
Für $x \in \mathbb{R}\backslash\{k\pi\,,\, k \in \mathbb{Z}\}$ ist $\cot(x) = \frac{\cos(x)}{\sin(x)}$.

Die Cosinusfunktion ist im Intervall $[0, \pi]$ streng monoton fallend, bildet also dieses Intervall bijektiv auf $[-1, 1]$ ab.
Die Sinusfunktion ist in $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ monoton wachsend und bildet das Intervall bijektiv auf $[-1, 1]$ ab.
$\tan(x)$ wächst streng monoton in $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ und $\lim_{x\to\pm\frac{\pi}{2}} \tan(x) = \pm\infty$.
Schließlich ist der Cotangens monoton fallend in $\left(0, \frac{\pi}{2}\right)$ und $\lim_{x\to 0} \cot(x) = \infty$, sowie $\lim_{x\to\pi} \cot(x) = -\infty$.
Somit lassen sich die Umkehrfunktionen definieren.

Definition 3.35. Die Umkehrfunktion des

- cos ist $\arccos : [-1, 1] \to \mathbb{R}$ (Arcuscosinus)
- sin ist $\arcsin : [-1, 1] \to \mathbb{R}$ (Arcussinus)
- tan ist $\arctan : \mathbb{R} \to \mathbb{R}$ (Arcustangens)
- cot ist $\operatorname{arccot} : \mathbb{R} \to \mathbb{R}$ (Arcuscotangens)

Definition 3.36 (Hyperbelfunktionen). Sei $x \in \mathbb{R}$. Dann ist

- $\cosh(x) = \frac{1}{2}(e^x + e^{-x})$ (Cosinus hyperbolicus)
- $\sinh(x) = \frac{1}{2}(e^x - e^{-x})$ (Sinus hyperbolicus)

- $\tanh(x) = \frac{\sinh(x)}{\cosh(x)} = \frac{e^x - e^{-x}}{e^x + e^{-x}}$ (Tangens hyperbolicus)
- $\coth(x) = \frac{\sinh(x)}{\cosh(x)} = \frac{e^x - e^{-x}}{e^x + e^{-x}}$, falls $x \neq 0$ (Cotangens hyperbolicus)

Die Inversen der Hyperbelfunktionen sind arsinh (Areasinus hyperbolicus), arcosh, artanh und arcoth.

4 Differentialrechnung

Eine klassische Fragestellung der Differentialrechnung ist das Tangentenproblem. Dabei wird zu einer gegebenen Funktion die Tangente an deren Graphen in einem Punkt gesucht.
Optimierungsprobleme können häufig dadurch gelöst werden, daß mit Hilfe der Differentialrechnung Extremwerte von Funktionen gefunden werden.
Ferner ist die Differentialrechnung Grundlage für Differentialgleichungen, welche zahlreiche Sachverhalte der Naturwissenschaft und Technik modellieren.

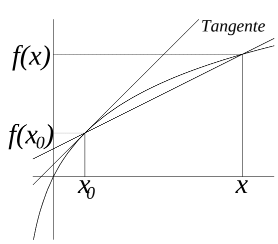
4.1 Differentiation

Definition 4.1. Sei $X \subset \mathbb{R}$ und $f : X \to \mathbb{R}$ eine reelle Funktion. f heißt **differenzierbar** in $x_0 \in X$, wenn der Grenzwert

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

existiert. Er heißt **Differentialquotient** (oder **Ableitung**) von f an der Stelle x_0. Statt $f'(x_0)$ schreibt man auch $\frac{\mathrm{d}f}{\mathrm{d}x}(x_0)$ oder $\mathrm{D}f(x_0)$.
f heißt **differenzierbar in** X, wenn f differenzierbar in allen $x \in X$ ist.

Alternativ dazu läßt sich auch $f'(x_0) = \lim_{h \to 0} \frac{f(x_0+h) - f(x_0)}{h}$ formulieren.
Die geometrische Interpretation der Ableitung erschließt sich durch Betrachtung der Funktion $s(x) = f(x_0) + \frac{f(x_0+h) - f(x_0)}{h}(x - x_0)$. Ihr Graph ist die Sekante des Graphen von f durch die Punkte $(x_0, f(x_0))$ und $(x_0 + h, f(x_0 + h))$. Ist f differenzierbar in x_0, so geht mit $h \to 0$ die Steigung $\frac{f(x_0+h) - f(x_0)}{h}$ von $s(x)$ gegen die der Tangenten an den Graphen von f im Punkt $(x_0, f(x_0))$.



Beispiel.

- $f(x) = c \quad \Rightarrow \quad f'(x_0) = \lim_{x \to x_0} \frac{c-c}{x-x_0} = 0$

- $f(x) = cx \quad \Rightarrow \quad f'(x_0) = \lim_{x \to x_0} \frac{cx - cx_0}{x-x_0} = c$

- $f(x) = x^n \quad \Rightarrow \quad f'(x_0) = \lim_{h \to 0} \frac{(x_0+h)^n - {x_0}^n}{h} =$
$\lim_{h \to 0} \sum_{k=0}^{n-1} \binom{n}{k} {x_0}^k h^{n-k-1} = \binom{n}{n-1} {x_0}^{n-1} = n{x_0}^{n-1}$

- $f(x) = \frac{1}{x} \quad \Rightarrow \quad f'(x) = \lim_{h \to 0} \frac{1}{h} \left(\frac{1}{x_0+h} - \frac{1}{x_0} \right) = \lim_{h \to 0} \frac{x_0 - (x_0+h)}{hx_0(x_0+h)} =$
$\lim_{h \to 0} \frac{-1}{x_0(x_0+h)} = -\frac{1}{{x_0}^2}$

- $f(x) = \frac{1}{x^n} \quad \Rightarrow \quad f'(x) = \frac{-n}{x^{n+1}}$

- $f(x) = \ln(x) \quad \Rightarrow \quad f'(x) = \lim_{h \to 0} \frac{\ln(x+h) - \ln(x)}{h} = \lim_{h \to 0} \frac{\ln\left(1 + \frac{h}{x}\right)}{x \frac{h}{x}} =$
$\frac{1}{x} \lim_{h \to 0} \frac{\ln\left(1 + \frac{h}{x}\right)}{\frac{h}{x}} = \frac{1}{x}$

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

  Bemerkung. Die Exponentialfunktion ist invariant unter Differentiation.

- $f(x) = \sin(x) \quad \Rightarrow \quad f'(x) = \lim_{h \to 0} \frac{\sin(x+h) - \sin(x)}{h} =$
$\lim_{h \to 0} \frac{2\cos\left(\frac{2x+h}{2}\right)\sin\left(\frac{h}{2}\right)}{h} = \lim_{h \to 0} \cos\left(x + \frac{h}{2}\right) \lim_{h \to 0} \frac{\sin\left(\frac{h}{2}\right)}{\frac{h}{2}} = \cos(x),$
denn $\quad \sin(x) - \sin(y) = \sin\left(\frac{x+y}{2} + \frac{x-y}{2}\right) - \sin\left(\frac{x+y}{2} - \frac{x-y}{2}\right) =$
$2\cos\left(\frac{x+y}{2}\right)\sin\left(\frac{x-y}{2}\right)$

- $f(x) = \cos(x) \quad \Rightarrow \quad f'(x) = -\sin(x)$

- $\operatorname{abs}(x) = |x|$ ist in 0 nicht differenzierbar, denn:
Sei $h_n = (-1)^n \frac{1}{n}$. Dann ist $\lim (h_n) = 0$. Die Folge
$q_n = \frac{|h_n + 0| - |0|}{h_n} = (-1)^n$ besitzt jedoch keinen Grenzwert, d.h. *abs* ist in 0 nicht differenzierbar.

Definition 4.2. Die Funktion $f : X \to \mathbb{R}$ heißt in x_0 **von links differenzierbar**, wenn der Grenzwert $f'_-(x_0) = \lim_{x \nearrow x_0} \frac{f(x) - f(x_0)}{x - x_0}$ existiert. Analog dazu heißt f **von rechts differenzierbar**, falls $f'_+(x_0) = \lim_{x \searrow x_0} \frac{f(x) - f(x_0)}{x - x_0}$ existiert.

Bemerkung. Die Absolutfunktion ist in 0 von rechts und von links differenzierbar. So ist $\operatorname{abs}'_+(0) = 1$ und $\operatorname{abs}'_-(0) = -1$.

Satz 4.1. *Sei $a \in X$ und es existiere eine Folge $(x_n)_{n \in \mathbb{N}}$ in $X \backslash \{a\}$ mit $\lim (x_n) = a$. Dann ist $f : X \to \mathbb{R}$ genau dann differenzierbar in a, wenn $\exists\, c \in \mathbb{R} : \quad f(x) = f(a) + c(x-a) + \varphi(x)$ mit $\lim_{x \to a} \frac{\varphi(x)}{x-a} = 0$.
Es ist dann $c = f'(a)$.*

Bemerkung. Die Differenzierbarkeit einer Funktion f an der Stelle x_0 ist gleichbedeutend mit der Approximierbarkeit durch eine lineare Funktion, welche die Tangente an den Graphen von f im Punkt $(a, f(a))$ beschreibt.

Satz 4.2. *Ist die Funktion $f: X \to \mathbb{R}$ differenzierbar in $a \in X$, so ist sie dort auch stetig.*

Beweis. Nach Satz 4.1 $\exists\, c \in \mathbb{R}: \quad f(x) = f(a) + c(x-a) + \varphi(x)$ mit $\lim_{x\to a} \frac{\varphi(x)}{x-a} = 0$.
$\Rightarrow \quad \lim_{x\to a} f(x) = f(a) + \lim_{x\to a} c(x-a) + \lim_{x\to a} (x-a)\frac{\varphi(x)}{x-a} =$
$f(a) + \lim_{x\to a}(x-a) \cdot \lim \frac{\varphi(x)}{x-a} = f(a)$, d.h. f ist stetig in a. □

Satz 4.3. *Sind $f, g: X \to \mathbb{R}$ differenzierbar in $x \in X$ und ist $\lambda \in \mathbb{R}$, dann sind auch $f \pm g, \lambda f, fg: X \to \mathbb{R}$, sowie $\frac{f}{g}: X \to \mathbb{R}$ mit $g(x) \neq 0$ differenzierbar an der Stelle x und es gilt:*

(1) $(f \pm g)'(x) = f'(x) \pm g'(x)$

(2) $(\lambda f)'(x) = \lambda f'(x)$

(3) $(fg)'(x) = f'(x)g(x) + f(x)g'(x)$ *(Produktregel)*

(4) $\left(\frac{f}{g}\right)'(x) = \frac{f'(x)g(x) - f(x)g'(x)}{[g(x)]^2}$ *(Quotientenregel)*

Beweis. Die Behauptungen (1) und (2) folgen direkt aus den Rechenregeln für Grenzwerte.

zu (3) $(fg)'(x) = \lim_{h\to 0} \frac{f(x+h)g(x+h) - f(x)g(x)}{h} =$
$\lim_{h\to 0} \frac{f(x+h)(g(x+h)-g(x))}{h} + \lim_{h\to 0} \frac{g(x)(f(x+h)-f(x))}{h} =$
$f(x) \lim_{h\to 0} \frac{g(x+h)-g(x)}{h} + g(x) \lim_{h\to 0} \frac{(f(x+h)-f(x))}{h} =$
$g'(x)f(x) + g(x)f'(x)$

zu (4) g ist stetig in x und $g(x) \neq 0$; somit gibt es eine Umgebung $U_\varepsilon(x)$ mit $g(y) \neq 0 \ \ \forall\, y \in U_\varepsilon$. Für $|h| < \varepsilon$ folgt
$\left(\frac{1}{g}\right)'(x) = \lim_{h\to 0} \frac{1}{h}\left(\frac{1}{g(x+h)} - \frac{1}{g(x)}\right) =$
$\lim_{h\to 0} \frac{g(x)-g(x+h)}{h} \lim_{h\to 0} \frac{1}{g(x)g(x+h)} = \frac{-g'(x)}{[g(x)]^2}$
Mit (3) folgt daraus $\left(\frac{f}{g}\right)'(x) = \frac{f'(x)g(x) - f(x)g'(x)}{[g(x)]^2}$.

□

Beispiel.

- $f_n(x) = x^n \quad \Rightarrow \quad f_n'(x) = nx^{n-1}$, denn für $n = 0$ ist dies offenbar wahr und aus $f_{n-1}'(x) = (n-1)x^{n-2}$ folgt $f_n'(x) = (f_{n-1}f_1)'(x) = (n-1)x^{n-2} \cdot x + x^{n-1} = nx^{n-1}$.
- Für $f(x) = \frac{1}{x^n}$ folgt aus der Quotientenregel $f'(x) = \frac{-nx^{n-1}}{x^{2n}} = -\frac{n}{x^{n+1}}$

- $f(x) = \tan(x) = \frac{\sin(x)}{\cos(x)} \quad \Rightarrow \quad f'(x) = \frac{\cos^2(x)+\sin^2(x)}{\cos^2(x)} = \frac{1}{\cos^2(x)}$
oder $f'(x) = \frac{\cos^2(x)+\sin^2(x)}{\cos^2(x)} = 1 + \tan^2(x)$

Satz 4.4 (Ableitung der Umkehrfunktion). *Sei $X \subset \mathbb{R}$ abgeschlossen und $f : X \to \mathbb{R}$ stetig und streng monoton.*
Weiterhin sei $\varphi : f(X) \to \mathbb{R}$ die Umkehrfunktion von f. Ist dann f differenzierbar in x und $f'(x) \neq 0$, so auch φ und es gilt $\varphi'(y) = \frac{1}{f'(\varphi(y))}$.

Beispiel.
- ln ist die Umkehrfunktion von exp. Mit Satz 4.4 folgt daraus $\ln'(x) = \frac{1}{\exp'(\ln(x))} = \frac{1}{x}$.
- $\arcsin'(x) = \frac{1}{\cos(\arcsin(x))} = \frac{1}{\sqrt{1-\sin^2(\arcsin(x))}} = \frac{1}{\sqrt{1-x^2}}$
- $\arctan'(x) = \frac{1}{1+\tan^2(\arctan(x))} = \frac{1}{1+x^2}$

Satz 4.5 (Kettenregel). *Ist $f : X \to \mathbb{R}$ differenzierbar in $x \in X$ und $g : Y \to \mathbb{R}$ mit $f(X) \subset Y$ differenzierbar in $f(x)$, so ist $g \circ f$ differenzierbar in x und es gilt $(g \circ f)'(x) = g'(f(x)) \cdot f'(x)$.*

Beispiel.
- Wenn $f : \mathbb{R} \to \mathbb{R}$ differenzierbar und $F(x) = f(ax + b)$, dann ist $F'(x) = a \cdot f(ax + b)$
- Sei $\alpha \in \mathbb{R}$ und $x \in \mathbb{R}^+$. Mit $x^\alpha = \exp(\alpha \ln(x))$ folgt
$\frac{\mathrm{d}x^\alpha}{\mathrm{d}x} = \exp(\alpha \ln(x)) \left(\alpha \frac{1}{x}\right) = x^\alpha \frac{\alpha}{x} = \alpha x^{\alpha-1}$ für beliebige $\alpha \in \mathbb{R}$.
- Mit Hilfe der Kettenregel läßt sich die Quotientenregel herleiten:
Ist $g : X \to \mathbb{R}$ differenzierbar in x, dann gilt $\left(\frac{1}{g}\right)'(x) = -\frac{1}{[g(x)]^2} \cdot g'(x)$
und somit $\left(\frac{f}{g}\right)'(x) = \frac{f'(x)g(x)-f(x)g'(x)}{[g(x)]^2}$ aufgrund der Produktregel.
- (logarithmische Ableitung) Ist $f : [a, b] \to \mathbb{R}$ in $[a, b]$ differenzierbar und hat dort keine Nullstelle, so ist $\ln|f|$ differenzierbar in (a, b) mit $(\ln|f|)' = \frac{f'}{f}$.

  Beispiel. $(\ln\cosh)' = \frac{\sinh}{\cosh} = \tanh$

  Seien $f_1, f_2, \ldots, f_n$ in (a, b) differenzierbar und dort ohne Nullstelle.
  Für $F := |f_1|^{a_1}|f_2|^{a_2} \cdot \ldots \cdot |f_n|^{a_n}$ mit $a_1, a_2, \ldots, a_n \in \mathbb{R}$ gilt dann

  $$\frac{F'}{F} = a_1 \frac{f_1'}{f_1} + a_2 \frac{f_2'}{f_2} + \ldots + a_n \frac{f_n'}{f_n},$$

  denn $\frac{F'}{F} = (\ln|F|)' = (a_1 \ln|f_1| + a_2 \ln|f_2| + \ldots + a_n \ln|f_n|)'$.
  Wegen $F' = \frac{F'}{F} F$ lassen sich mit dieser Regel Produkte ableiten.

Beispiel. $F(x) = \sqrt{\frac{1+x^2}{1-x^2}}$, wobei $|x| < 1$ gelten soll.
$F(x) = (1+x^2)^{\frac{1}{2}}(1-x^2)^{-\frac{1}{2}}$
$\Rightarrow \quad \frac{F'}{F}(x) = \frac{1}{2}\frac{2x}{1+x^2} + \frac{1}{2}\frac{2x}{1-x^2} = \frac{2x}{(1+x^2)(1-x^2)}$
$\Rightarrow \quad F'(x) = \frac{2x}{(1+x^2)(1-x^2)}\sqrt{\frac{1+x^2}{1-x^2}}$

Definition 4.3. Sei f differenzierbar. Ist auch f' differenzierbar, dann heißt $\frac{\mathrm{d}^2 f(x)}{\mathrm{d}x^2} = f''(x) := (f'(x))'$ die **zweite Ableitung** von f nach x.
Analog dazu ist die **k-te Ableitung** ($k \in \mathbb{N}$) von f, sofern sie existiert, rekursiv definiert durch $f^k(x) = \frac{\mathrm{d}^k f(x)}{\mathrm{d}x^k} := \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\mathrm{d}^{k-1} f(x)}{\mathrm{d}x^{k-1}}\right)$ und $f^0 = f$.
f heißt **k mal stetig differenzierbar**, falls die k-te Ableitung von f existiert und stetig ist.

4.2 Extrema und Mittelwertsätze

Definition 4.4. Die Funktion $f : X \to \mathbb{R}$ hat in $x_0 \in X$ ein **globales Maximum**, wenn $f(x) \leq f(x_0) \;\; \forall x \in X$.
f hat in x_0 ein **lokales Maximum**, wenn es eine Umgebung $U(x_0)$ gibt, sodaß $f(x) \leq f(x_0) \;\; \forall x \in U \cap X$

Lokales und globales Minimum werden analog definiert.

Bemerkung. Jede auf einer nichtleeren, kompakten Menge stetige Funktion hat dort nach Satz 3.15 stets ein globales Minimum und Maximum.

Satz 4.6. *Ist $f : (a,b) \to \mathbb{R}$ differenzierbar in $x \in (a,b)$ und hat dort ein lokales Extremum, so folgt $f'(x) = 0$.*

Beweis. Sei x eine lokale Minimumsstelle. Dann $\exists \;\; \varepsilon > 0$:
$(x-\varepsilon, x+\varepsilon) \subset (a,b) \quad$ und $\quad f(\xi) \geq f(x) \;\; \forall \xi \in (x-\varepsilon, x+\varepsilon)$.
$\Rightarrow \quad f'_+(x) = \lim_{\xi \searrow x} \frac{f(\xi)-f(x)}{\xi - x} \leq 0$ und $f'_-(x) = \lim_{\xi \nearrow x} \frac{f(\xi)-f(x)}{\xi - x} \geq 0$.
Da nun f differenzierbar in x ist, folgt $f'_+(x) = f'_-(x) = f'(x)$ und damit $f'(x) = 0$. □

Bemerkung.

- Mit vorangegangenem Satz ist die notwendige Bedingung $f'(x) = 0$ für ein lokales Extremum in x gegeben. Sie ist nicht hinreichend, wie das Beispiel $f(x) = x^3$ zeigt; es ist $f'(0) = 0$, jedoch hat f kein lokales Extremum.

- Ist der Definitionsbereich der Funktion f ein kompaktes Intervall, so besitzt f immer ein Extremum. Liegt dieses am Rand, so muß die Ableitung dort nicht 0 sein.

  Beispiel.
  $f : [0,1] \to \mathbb{R}, \;\; f(x) = x \quad \Rightarrow \quad f(x) \leq f(1) \;\; \forall x \in [0,1], \;\; f'(1) = 1$

Satz 4.7 (Satz von Rolle[19]). *Ist* $f\,[a,b] \to \mathbb{R}$ *stetig in* $[a,b]$*, differenzierbar in* (a,b) *und es gilt* $f(a) = f(b)$*, dann*

$$\exists\, \xi \in (a,b): \quad f'(\xi) = 0$$

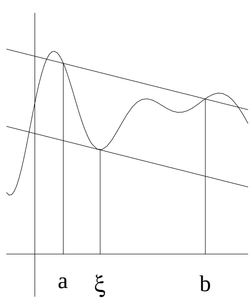
Beweis. Ist f konstant, so ist $f'(\xi) = 0 \;\; \forall\, \xi \in (a,b)$.
Falls f nicht konstant ist, so nimmt sie in einem $\xi \in (a,b)$ ihr Maximum oder Minimum an. Nach Satz 4.6 ist $f'(\xi) = 0$ für dieses ξ. □

Bemerkung. Insbesondere folgt aus dem Satz von Rolle, daß zwischen zwei Nullstellen einer differenzierbaren Funktion stets eine Nullstelle der Ableitung liegt.

Satz 4.8 (Mittelwertsatz). *Ist* $f\,[a,b] \to \mathbb{R}$ *stetig in* $[a,b]$ *und differenzierbar in* (a,b)*, dann*

$$\exists\, \xi \in (a,b): \quad \frac{f(a) - f(b)}{a - b} = f'(\xi)$$

Beweis. Die Hilfsfunktion $F: [a,b] \to \mathbb{R}$ sei definiert durch
$F(x) = f(x) - \frac{f(a)-f(b)}{a-b}(x-a)$. Sie ist ebenso wie f stetig in $[a,b]$ und differenzierbar in (a,b).
Außerdem gilt $F(a) = F(b) = f(a)$. Damit folgt aus dem Satz von Rolle die Existenz eines $\xi \in (a,b)$, für das $F'(\xi) = f'(\xi) - \frac{f(a)-f(b)}{a-b} = 0$, also $\frac{f(a)-f(b)}{a-b} = f'(\xi)$ ist. □



Als Folgerung ergibt sich

Satz 4.9 (Monotoniekriterium). *Sei* $f: (a,b) \to \mathbb{R}$ *differenzierbar in* (a,b)*. Wenn* $f'(x) \geq 0 \;\; \forall\, x \in (a,b)$*, so ist* f *monoton wachsend in* (a,b)*.*
f *wächst streng monoton, falls* $f'(x) > 0 \;\; \forall\, x$ *gilt.*
Wenn $f'(x) \leq 0$ $(f'(x) < 0)$ $\forall\, x \in (a,b)$ *gilt, dann fällt* f *in* (a,b) *(streng) monoton.*

[19]nach *Michel Rolle (1652-1719)*

Bemerkung. Für eine in (a, b) differenzierbare Funktion, die außerdem in a und b stetig ist, gilt Satz 4.9 auch im abgeschlossenen Intervall $[a, b]$.

Satz 4.10. *Sei $f : (a, b) \to \mathbb{R}$ differenzierbar in (a, b) und $f'(x_0) = 0$ für ein $x_0 \in (a, b)$.*
Dann hat f in x_0 ein lokales Minimum, wenn es eine Umgebung (α, β) von x_0 gibt, sodaß $f'(x) \leq 0 \;\; \forall\, x \in (\alpha, x_0)$ und $f'(x) \geq 0 \;\; \forall\, x \in (x_0, \beta)$ ist.
Analog hat f in x_0 ein lokales Maximum, falls es eine Umgebung (α, β) von x_0 gibt, sodaß $f'(x) \geq 0 \;\; \forall\, x \in (\alpha, x_0)$ und $f'(x) \leq 0 \;\; \forall\, x \in (x_0, \beta)$ ist.

Satz 4.11. *Eine Funktion $f : (a, b) \to \mathbb{R}$ sei zweimal stetig differenzierbar und für $x_0 \in (a, b)$ sei $f'(x_0) = 0$.*
Dann besitzt f in x_0 ein lokales Minimum (Maximum), wenn $f''(x_0) > 0$ ($f''(x_0) < 0$) ist.

Beweis. Da f'' stetig ist, existiert für den Fall $f''(x_0) > 0$ eine Umgebung (α, β) von x_0, in der $f''(x) > 0$ ist. Also ist f' nach Satz 4.9 streng wonoton wachsend in (α, β). Mit der Voraussetzung $f'(x_0) = 0$ und Satz 4.10 folgt, daß f in x_0 ein lokales Minimum hat.
Wenn $f''(x_0) < 0$ ist, dann existiert eine Umgebung (α, β) von x_0, in der $f''(x) < 0$ ist, womit f' streng monoton fällt und f ein lokales Maximum in x_0 hat. □

Das Kriterium $f'(x) = 0$, $f''(x) \neq 0$ für ein lokales Extremum von f in x ist hinreichend, jedoch nicht notwendig.

Beispiel. $f(x) = x^4$ hat in 0 ein lokales Minimum und es ist $f''(x) = 0$.

Eine weitere Folgerung aus dem Mittelwertsatz ist

Satz 4.12. *Sei $f : [a, b] \to \mathbb{R}$ stetig in $[a, b]$ und differenzierbar in (a, b). Ist die Ableitung von f beschränkt, d.h. $\exists\, m, M \in \mathbb{R}$ mit $m \leq f'(\xi) \leq M \;\; \forall\, \xi \in (a, b)$, dann gilt $\forall\, x, y \in [a, b]$ mit $x \leq y :\;\; m(y - x) \leq f(y) - f(x) \leq M(y - x)$.*

Für $M = m = 0$ ergibt sich

Folgerung 4.13. *Ist $f : [a, b] \to \mathbb{R}$ stetig in $[a, b]$, differenzierbar in (a, b) und $f'(x) = 0 \;\; \forall\, x \in (a, b)$, so ist f eine konstante Funktion.*

Damit läßt sich die Exponentialfunktion durch eine Differentialgleichung charakterisieren:

Satz 4.14. *Ist $f : \mathbb{R} \to \mathbb{R}$ überall differenzierbar, $c \in \mathbb{R}$ und es gilt $f'(x) = c \cdot f(x) \;\; \forall\, x$, dann folgt $f(x) = Ae^{cx}$ mit $A = f(0)$.*

Beweis. Sei $F(x) = f(x)e^{-cx}$.
$\Rightarrow \quad F'(x) = f'(x)e^{-cx} - cf(x)e^{-cx} = e^{-cx}(f'(x) - cf(x)) = 0$
Also ist F wegen Folgerung 4.13 eine konstante Funktion.
$\Rightarrow \quad F(x) = F(0) = f(0) = A$, d.h. $f(x)$ kann höchstens Ae^{cx} sein.
Einsetzen von $f(x) = Ae^{cx}$ in die Differentialgleichung ergibt, daß dieses f auch wirklich eine Lösung ist. $\square$

Beispiel (Methode der kleinsten Quadrate). Zu gegebenen $a_1, a_2, \ldots a_n \in \mathbb{R}$ wird ein $a \in \mathbb{R}$ gesucht, sodaß $\sum_{i=1}^{n} (a - a_i)^2$ minimal ist, d.h. es ist dasjenige Argument von $f(x) = (x - a_1)^2 + (x - a_2)^2 + \ldots + (x - a_n)^2$ zu finden, in dem f ein Minimum hat.
$f'(x) = 2((x - a_1)^2 + \ldots + (x - a_2)^2) = 0 \quad \Rightarrow \quad x = \frac{a_1 + a_2 + \ldots + a_n}{n}$, d.h. das arithmetische Mittel $a = \frac{a_1 + a_2 + \ldots + a_n}{n}$ der a_i ist die einzige potentielle (lokale) Extremstelle. $f''(a) = 2n > 0$ zeigt, daß f in a tatsächlich ein Minimum hat.

Satz 4.15 (Erweiterter Mittelwertsatz). *Seien $f, g : [a, b] \to \mathbb{R}$ stetig im Definitionsbereich und differenzierbar in (a, b). Sei überdies $g(x) \neq 0 \;\forall x \in (a, b)$. Dann ist $g(a) \neq g(b)$ und $\exists \xi \in (a, b)$ mit*

$$\frac{f(a) - f(b)}{g(a) - g(b)} = \frac{f'(\xi)}{g'(\xi)}$$

Beweis. Es ist $g(a) \neq g(b)$, da es andernfalls ein $\xi \in (a, b)$ mit $f(\xi) = 0$ gäbe, was laut Voraussetzung nicht der Fall ist.
Sei nun $F(x) = f(x) - \frac{f(a) - f(b)}{g(a) - g(b)}(g(x) - g(a))$.
F ist stetig in $[a, b]$ und differenzierbar in (a, b). Es gilt $F(a) = F(b) = f(a)$, daher $\exists \xi \in (a, b)$ mit $F'(\xi) = f'(\xi) - \frac{f(a) - f(b)}{g(a) - g(b)} g'(\xi) = 0$. $\square$

Der folgende Satz ist eine Anwendung des erweiterten Mittelwertsatzes.

Satz 4.16 (Regel von de l'Hôspital[20]). *Sind $f, g : (a, b) \to \mathbb{R}$ differenzierbar, $g'(x) \neq 0 \;\forall x \in (a, b)$ und es gilt $\lim_{x \searrow a} f(x) = \lim_{x \searrow a} g(x) = 0$ oder $\lim_{x \searrow a} f(x) = \lim_{x \searrow a} g(x) = \infty$, so folgt: Wenn $\lim_{x \searrow a} \frac{f'(x)}{g'(x)}$ existiert, dann auch $\lim_{x \searrow a} \frac{f(x)}{g(x)}$ und*

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

Beispiel. $\lim_{x \searrow 0} x \ln(x) = \lim_{x \searrow 0} \frac{\ln(x)}{\frac{1}{x}} = \lim_{x \searrow 0} \frac{\frac{1}{x}}{-\frac{1}{x^2}} = 0$

Bemerkung. In einigen Fällen läßt sich der Grenzwert einer Funktion erst durch mehrmaliges Anwenden der Regel von de l'Hôspital bestimmen.

---

[20] *Guillaume de l'Hôspital* (1661-1704)

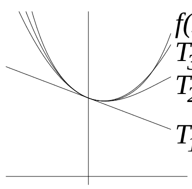
4.3 Taylorpolynome

Wegen Satz 4.1 läßt sich eine in $a \in X$ differenzierbare Funktion $f: X \to \mathbb{R}$ durch die Funktion $L(x) = f(a) + f'(a)(x-a)$ linear approximieren, wobei für den Fehler $R(x) = f(x) - L(x)$ gilt: $\lim_{x \to a} \frac{R(x)}{x-a} = 0$. Es ist außerdem $L(a) = f(a)$ und $L'(a) = f'(a)$.
Sei nun f eine in a n mal differenzierbare Funktion. Gesucht wird ein Polynom T mit $T(a) = f(a)$, $T'(a) = f'(a), \ldots T^{(n)}(a) = f^{(n)}(a)$.
Der Ansatz $T(x) = \sum_{k=0}^{n} a_k (x-a)^k$ liefert mit $T^{(k)}(a) = a_k k! = f^{(k)}(a)$ die Koeffizienten $a_k = \frac{f^{(k)}(a)}{k!}$

Definition 4.5. Sei $f: X \to \mathbb{R}$ eine in $a \in X$ n-mal differenzierbare Funktion. Dann heißt

$$T_n f(x,a) := f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \ldots + \frac{f^{(n)}(a)}{n!}(x-a)^n$$

das **n-te Taylorpolynom**[21] von f in a.



Für eine Funktion f mit der Darstellung $f(x) = \sum_{k=0}^{\infty} a_k (x-a)^k$ ist das n-te Taylorpolynom die n-te Partialsumme der Reihe.

Satz 4.17 (Lagrange-Form des Restgliedes[22]). *Sei $f: [a,b] \to \mathbb{R}$ eine in (a,b) n+1 mal differenzierbare und in $[a,b]$ n mal stetig differenzierbare Funktion. Dann gibt es zu jedem $x \in (a,b]$ ein $\xi \in (a,x)$, sodaß gilt*

$$R_{n+1}(x) := f(x) - T_n f(x,a) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-a)^{n+1}$$

Beweis. Sei $g: [a,b] \to \mathbb{R}$ definiert durch
$g(y) := f(x) - f(y) - \frac{f'(y)}{1!}(x-y) - \frac{f''(y)}{2!}(x-y)^2 - \ldots - \frac{f^{(n)}(y)}{n!}(x-y)^n - m\frac{(x-y)^{n+1}}{(n+1)!}$
mit der Konstanten $m = R_{n+1}(x)\frac{(n+1)!}{(x-a)^{n+1}}$.
Diese Funktion ist stetig in $[a,b]$ und differenzierbar in (a,b), da f in diesen Intervallen n mal stetig differenzierbar bzw. n+1 mal differenzierbar ist.
$g'(y) = -f'(y) - \left(\frac{f''(y)}{1!}(x-y) - \frac{f'(y)}{1!}\right) - \left(-2\frac{f''(y)}{2!}(x-y) + \frac{f'''(y)}{2!}(x-y)^2\right)$
$- \left(\frac{f^{(4)}(y)}{3!}(x-y)^3 - 3\frac{f'''(y)}{3!}(x-y)^2\right) - \left(-4\frac{f^{(4)}(y)}{4!}(x-y)^3 + \frac{f^{(5)}(y)}{4!}(x-y)^4\right)$

[21]nach *Brook Taylor* (1685-1731)

[22]nach *Joseph Louis Lagrange* (1736-1813)

$- \ldots\ldots - \left(\frac{f^{(n+1)}(y)}{n!}(x-y)^n - n\frac{f^{(n)}(y)}{n!}(x-y)^{n-1}\right) + m\frac{(x-y)^n}{n!}$
$= -\frac{f^{(n+1)}(y)}{n!}(x-y)^n + m\frac{(x-y)^n}{n!}$
Es gilt $g(x) = 0$ und $g(a) = R_{n+1}(x) - m\frac{(x-a)^{n+1}}{(n+1)!} = 0$.
Nach dem Satz von ROLLE existiert daher ein ξ zwischen a und x, sodaß $g'(\xi) = -\frac{f^{(n+1)}(\xi)}{n!}(x-\xi)^n + m\frac{(x-\xi)^n}{n!} = 0$, also $m = f^{(n+1)}(\xi)$ ist. Mit der Definition von m folgt daraus die Behauptung. □

Beispiel. Der Cosinus hat eine Darstellung als
$\cos x = \sum_{k=0}^{\infty}(-1)^k\frac{x^{2k}}{(2k)!} = \sum_{k=0}^{n}(-1)^k\frac{x^{2k}}{(2k)!} + R_{2n+2}$
Aus dem vorangegangenen Satz folgt für das Restglied
$|R_{2n+2}| = \frac{\left|\cos^{(2n+2)}(\xi)\right|}{(2n+2)!}|x|^{2n+2} \leq \frac{|x|^{2n+2}}{(2n+2)!}$

Als weitere Anwendung von Satz 4.17 läßt sich Satz 4.11 verallgemeinern:

Beispiel. Sei f n mal stetig differenzierbar und es seien
$f'(a), f''(a), \ldots f^{(n-1)}(a) = 0$, sowie $f^{(n)}(a) \neq 0$.
Dann hat f in a

(i) ein lokales Minimum, falls n gerade und $f^{(n)}(a) > 0$ ist, denn ist U eine Umgebung von a, in der $f^{(n)}(x) > 0$ ist, so gilt
$f(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \ldots + \frac{f^{(n-1)}(a)}{n!}(x-a)^{n-1} + \frac{f^{(n)}(\xi)}{n!}(x-a)^n = f(a) + \frac{f^{(n)}(\xi)}{n!}(x-a)^n > f(a) \ \ \forall x \in U$

(ii) ein lokales Maximum, falls n gerade und $f^{(n)}(a) < 0$ ist.

(iii) kein Extremum, falls n ungerade ist, denn:
Sei $f^{(n)}(x) > 0$ in einer Umgebung von a. Dort ist dann
$R_n(x) = \frac{f^{(n)}(\xi)}{n!}(x-a)^n > 0$ für $x > a$ und $R_n(x) < 0$ für $x < a$, d.h.
$f(x) = f(a) + R_n(x) > f(a)$, falls $x > a$ und $f(x) < f(a)$ für $x < a$.

Folgerung 4.18 (Qualitative Taylorformel). *Ist $f : I \to \mathbb{R}$ n mal stetig differenzierbar auf dem Intervall I, so gibt es eine auf I stetige Funktion r mit $r(a) = 0$ und*

$$f(x) = T_n f(x,a) + (x-a)^n r(x)$$

Beweis. Nach Satz 4.17 existiert ein ξ zwischen a und x, sodaß
$f(x) - T_{n-1}f(x,a) = \frac{f^{(n)}(\xi)}{n!}(x-a)^n$ ist. Mit
$r(x) = \frac{1}{n!}(f^{(n)}(\xi) - f^{(n)}(x)) = \frac{1}{(x-a)^n}\left(f(x) - T_{n-1}f(x,a) - \frac{f^{(n)}(x)}{n!}(x-a)^n\right)$
$= \frac{1}{(x-a)^n}(f(x) - T_n f(x,a))$ ergibt sich $f(x) = T_n f(x,a) + (x-a)^n r(x)$ und da $f^{(n)}$ stetig ist und ξ zwischen a und x liegt, gilt zudem $\lim_{x\to a} r(x) = 0$. □

Definition 4.6. Ist $f: X \to \mathbb{R}$ eine in $a \in X$ beliebig oft differenzierbare Funktion, dann heißt die Potenzreihe

$$Tf(x,a) := \sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!}(x-a)^k$$

Taylorreihe von f in a.
Konvergiert $Tf(x,a)$ gegen $f(x)$ für alle x aus einer Umgebung U von a, so **besitzt** f in U eine **Taylorentwicklung** mit Entwicklungspunkt a.

Bemerkung.

- Hat die Funktion f eine Darstellung als Potenzreihe $f(x) = \sum_{k=0}^{\infty} a_k(x-a)^k$, so ist dies die Taylorreihe von f in a.
- Die Taylorreihe $Tf(x,a)$ kann für $x \neq a$ divergieren. Auch muß im Konvergenzfall der Grenzwert nicht mit $f(x)$ übereinstimmen.

  Beispiel. $f(x) = \begin{cases} e^{-\frac{1}{x}} & \text{für } x > 0 \\ 0 & \text{für } x \leq 0 \end{cases}$
  ist in 0 unendlich oft differenzierbar, $Tf(x,0) = 0 \ \forall\, x$ und $f(x) \neq 0 \ \forall\, x > 0$.

Definition 4.7. Eine Funktion $f: [a,b] \to \mathbb{R}$ heißt **analytisch** in $x_0 \in [a,b]$, wenn es ein $\delta > 0$ und eine Potenzreihe mit Konvergenzradius $\geq \delta$ gibt, sodaß $f(x) = \sum_{k=0}^{\infty} a_k(x-a)^k \ \forall\, x \in U_\delta(x_0)$.
f ist analytisch in $[a,b]$, wenn f analytisch in allen $x \in [a,b]$ ist.

Bemerkung. Ist f eine analytische Funktion, so ist $f \in \mathcal{C}^\infty$. Die Umkehrung gilt nicht; es gibt unendlich oft differenzierbare, nichtanalytische Funktionen.

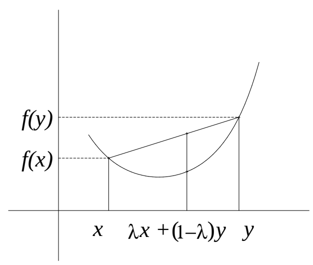
4.4 Konvexität

Definition 4.8. Eine Funktion $f: [a,b] \to \mathbb{R}$ heißt **konvex**, falls $\forall\, x_1, x_2 \in [a,b]$, $\lambda \in (0,1)$ gilt:

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

f heißt **konkav**, wenn $-f$ konvex ist.

Geometrisch interpretiert bedeutet die Konvexität einer Funktion, daß ihr Graph in jedem Intervall $[x_1, x_2]$ unterhalb der Sekante durch $(x_1, f(x_1))$ und $(x_2, f(x_2))$ liegt.



Bemerkung. Wird in Definition 4.8 die Gleichheit ausgeschlossen, dann heißt die Funktion streng konvex (bzw. streng konkav).

Bemerkung. Konvexität einer Funktion impliziert nicht die Differenzierbarkeit dieser.

Beispiel. $f(x) = |x|$ ist konvex, in 0 jedoch nicht differenzierbar.

Satz 4.19. *Ist die Funktion $f : (a, b) \to \mathbb{R}$ zweimal differenzierbar in (a, b), so ist sie genau dann konvex, wenn $f''(x) \geq 0 \;\; \forall\, x \in (a, b)$ ist.*

Beweis. „$\Leftarrow$": Sei $f''(x) \geq 0 \;\; \forall\, x \in (a, b)$.
Dann folgt aus Satz 4.9, daß f' auf (a, b) monoton wachsend ist.
Seien nun $x_1, x_2 \in (a, b)$, wobei o.B.d.A. $x_1 < x_2$ ist, sowie $\lambda \in (0, 1)$ und
$x = \lambda x_1 + (1 - \lambda)x_2$.
Nach dem Mittelwertsatz $\exists\, \xi_1 \in (x_1, x),\; \xi_2 \in (x, x_2)$, sodaß
$\frac{f(x)-f(x_1)}{x-x_1} = f'(\xi_1) \leq f'(\xi_2) = \frac{f(x_2)-f(x)}{x_2-x}$.
Mit $x - x_1 = (1 - \lambda)(x_2 - x_1)$ und $x_2 - x = \lambda(x_2 - x_1)$ ergibt sich
$\frac{f(x)-f(x_1)}{1-\lambda} \leq \frac{f(x_2)-f(x)}{\lambda} \;\Rightarrow\; \quad f(x) \leq \lambda f(x_1) + (1 - \lambda)f(x_2)$, d.h. f ist konvex.
„$\Rightarrow$": Sei nun f konvex. Man nehme an, es gäbe ein $x_0 \in (a, b)$ mit
$f''(x_0) < 0$. Die Funktion $\varphi(x) := f(x) - f'(x_0)(x - x_0)$ ist zweimal differenzierbar und $\varphi'(x_0) = 0$, $\varphi''(x_0) = f''(x_0) < 0$.
Ähnlich wie im Beweis von Satz 4.11 läßt sich zeigen, daß φ in x_0 ein isoliertes Maximum hat, d.h. $\exists\, U_\varepsilon(x_0) : \;\; \varphi|_{U_\varepsilon \setminus \{x_0\}} < \varphi(x_0)$.
Sei nun $h > 0$ so gewählt, daß $x_0 + h, x_0 - h \in U_\varepsilon$ also $\varphi(x_0 - h) < \varphi(x_0)$ und $\varphi(x_0 + h) < \varphi(x_0)$ sind (z.B. $h = \frac{\varepsilon}{2}$).
Damit folgt:
$f(x_0) = \varphi(x_0) > \frac{1}{2}\left(\varphi(x_0 - h) + \varphi(x_0 + h)\right) = \frac{1}{2}\left(f(x_0 - h) + f(x_0 + h)\right)$
Setzt man schließlich $x_1 = x_0 - \frac{h}{2}$, $x_2 = x_0 + \frac{h}{2}$ und $\lambda = \frac{1}{2}$, so erhält man
$f(\lambda(x_1) + (1 - \lambda)(x_2)) > \lambda f(x_1) + (1 - \lambda)f(x_2)$ im Widerspruch zur Konvexität von f. $\square$

Definition 4.9. Sei $f : (a, b) \to \mathbb{R}$ stetig. $(x_0, f(x_0))$ heißt **Wendepunkt** von f, wenn es Intervalle (α, x_0) und (x_0, β) gibt, sodaß f

(1) in (α, x_0) konvex und in (x_0, β) konkav
oder
(2) konvex in (x_0, β) und konkav in (α, x_0) ist.

Bemerkung. Für eine zweimal differenzierbare Funktion f ist Forderung (1) aus Definition 4.9 äquivalent zu
$f''(x) \geq 0 \ \forall x \in (\alpha, x_0) \quad \wedge \quad f''(x) \leq 0 \ \forall x \in (x_0, \beta)$
Analog ist dazu ist Forderung (2) aus Definition 4.9 für zweimal differenzierbare f äquivalent zu
$f''(x) \leq 0 \ \forall x \in (\alpha, x_0) \quad \wedge \quad f''(x) \geq 0 \ \forall x \in (x_0, \beta)$, d.h.
$(x_0, f(x_0))$ ist Wendepunkt von f, wenn f' ein lokales Extremum in x_0 hat.
Eine notwendige Bedingung dafür ist $f''(x_0) = 0$.
Ist f in einer Umgebung von x_0 dreimal stetig differenzierbar, dann ist $f''(x_0) = 0$, $f'''(x_0) \neq 0$ eine hinreichende Bedingung für einen Wendepunkt (vgl Satz 4.11).

Mit dem Begriff der Konvexität lassen sich einige wichtige Ungleichungen formulieren.

Satz 4.20 (Jensensche Ungleichung[23]). *Sei f eine im Intervall I konvexe Funktion und für $\lambda_1, \lambda_2, \dots \lambda_n \in \mathbb{R}^+$ gelte $\sum_i \lambda_i = 1$. Dann gilt $\forall\, x_1, x_2, \dots x_n \in I$:*

$$f(\lambda_1 x_1 + \lambda_2 x_2 + \ldots + \lambda_n x_n) \leq \lambda_1 f(x_1) + \lambda_2 f(x_2) + \ldots + \lambda_n f(x_n)$$

Ist f auch streng konvex, dann tritt Gleichheit nur für $x_1 = x_2 = \ldots = x_n$ ein.

Der Beweis kann mittels vollständiger Induktion über n geführt werden.

Beispiel. Die Funktion $f(x) = \ln(x)$ ist konkav, denn $f''(x) = -\frac{1}{x^2} < 0 \ \forall x$.
Somit gilt wegen der Jensenschen Ungleichung für $x_1, x_2, \dots x_n \in \mathbb{R}^+$ und $\lambda_1, \lambda_2, \dots \lambda_n \in \mathbb{R}^+$ mit $\sum_i \lambda_i = 1$:
$\ln(\lambda_1 x_1 + \lambda_2 x_2 + \ldots + \lambda_n x_n) \geq \lambda_1 \ln(x_1) + \lambda_2 \ln(x_2) + \ldots + \lambda_n \ln(x_n)$
Die Exponentialfunktion ist überall streng monoton wachsend, daher folgt
$\lambda_1 x_1 + \lambda_2 x_2 + \ldots + \lambda_n x_n \geq \exp(\lambda_1 \ln(x_1)) \exp(\lambda_2 \ln(x_2)) \cdot \ldots \cdot \exp(\lambda_n \ln(x_n)) = {x_1}^{\lambda_1} {x_2}^{\lambda_2} \ldots {x_n}^{\lambda_n}$, also die Ungleichung zwischen dem gewichteten geometrischen und artithetischen Mittel:

$${x_1}^{\lambda_1} {x_2}^{\lambda_2} \ldots {x_n}^{\lambda_n} \leq \lambda_1 x_1 + \lambda_2 x_2 + \ldots + \lambda_n x_n$$

Setzt man alle $\lambda_i = \frac{1}{n}$, so erhält man den Spezialfall

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

[23] nach *Johan Jensen* (1859-1925)

Bemerkung. Gleichheit tritt nur für $x_1 = x_2 = \ldots = x_n$ ein, da der $\ln$ streng konkav ist.

Folgerung 4.21 (Youngsche Ungleichung [24]). *Seien $a, b \geq 0,\ p, q > 1$ und $\frac{1}{p} + \frac{1}{q} = 1$. Dann gilt*

$$ab \leq \frac{1}{p}a^p + \frac{1}{q}b^q$$

Definition 4.10. Sei $p \in \mathbb{R}^+$ und $(x_1, x_2, \ldots x_n) \in \mathbb{R}^n$. Dann heißt

$$||x||_p = \sqrt[p]{\sum_{i=1}^{n} |x_i|^p}$$

die **p-Norm** von x.

$$||x||_\infty = \max |x_i|$$

heißt **Maximumsnorm** (oder **Supremumsnorm**) von x.

Die Euklidische Norm ist nach dieser Bezeichnung die 2-Norm.

Bemerkung. Es ist $\lim_{x\to\infty} ||x||_p = ||x||_\infty$.
Beweis: Übungsaufgabe 57.

$||x||_p$ ist eine Norm auf $\mathbb{R}^n$ gemäß Definition 3.21, denn

- $||x||_p \geq 0$ und $\quad ||x||_p = 0 \quad \Leftrightarrow \quad x = 0$
- $||cx||_p = |c| \cdot ||x||_p$ mit $c \in \mathbb{R}$.
- Die Dreiecksungleichung gilt (Satz 4.24).

Satz 4.22 (Höldersche Ungleichung[25]). *Seien $p, q > 1$ mit $\frac{1}{p} + \frac{1}{q} = 1$ oder $p = 1, q = \infty$. Seien weiter $x, y \in \mathbb{R}^n$. Dann gilt*

$$||x \bullet y||_1 \leq ||x||_p \cdot ||y||_q\,,$$

wobei $x \bullet y$ das Skalarprodukt von x und y bezeichnet.

Beweis. Für $p = 1, q = \infty$ ist die Behauptung trivial, ebenso für $x = 0$ oder $y = 0$. Sei nun also $||x||_p||y||_q \neq 0$.
$\frac{||x\bullet y||_1}{||x||_p||y||_q} \leq \sum_{i=1}^n \frac{|x_i y_i|}{||x||_p||y||_q}$
$\leq \sum_{i=1}^n \frac{1}{p}\frac{|x_i|^p}{||x||_p^p} + \frac{1}{q}\frac{|x_i|^q}{||x||_q^q}$ (Youngsche Ungleichung)
$= \frac{1}{p}\frac{\sum |x_i|^p}{||x||_p{}^p} + \frac{1}{q}\frac{\sum |x_i|^q}{||x||_q^q} = \frac{1}{p} + \frac{1}{q} = 1$ □

[24] nach *William Young* (1863-1942)

[25] nach *Otto Hölder* (1859-1937)

Setzt man $p = q = 2$, so erhält man

Folgerung 4.23 (Cauchy-Schwarz[26]sche Ungleichung).

$$||x \bullet y||_1 \leq ||x||_2 \cdot ||y||_2$$

Satz 4.24 (Minkowski-Ungleichung[27]). *Sei $p \geq 1$ oder $p = \infty$ und $x, y \in \mathbb{R}^n$. Dann folgt*

$$||x + y||_p \leq ||x||_p + ||y||_p$$

Beweis. Für $p = 1$ und $p = \infty$ ergibt sich die Dreiecksungleichung in $\mathbb{R}$.
Sei nun $p > 1$ und $q = \frac{p}{p-1}$. Dann gilt $\frac{1}{p} + \frac{1}{q} = 1$. Damit ist
$||x + y||_p^p = \sum |x_i + y_i|^p = \sum |x_i + y_i| \cdot |x_i + y_i|^{p-1}$
$\leq \sum |x_i| \cdot |x_i + y_i|^{p-1} + \sum |y_i| \cdot |x_i + y_i|^{p-1}$
Aufgrung der HÖLDERsche Ungleichung ist dies
$\leq \left(\sum |x_i|^p\right)^{\frac{1}{p}} \left(\sum |x_i + y_i|^{(p-1)q}\right)^{\frac{1}{q}} + \left(\sum |y_i|^p\right)^{\frac{1}{p}} \left(\sum |x_i + y_i|^{(p-1)q}\right)^{\frac{1}{q}}$
$= (||x||_p + ||y||_p) \, ||x + y||_p^{\frac{p}{q}}$
Letzteres folgt aus $(p - 1)q = p$.
Für $||x + y||_p = 0$ ist die Behauptung trivial; für die übrigen Fälle liefert die Division durch $||x + y||_p^{\frac{p}{q}}$:
$||x + y||_p = ||x + y||_p^{p-\frac{p}{q}} \leq ||x||_p + ||y||_p$ □

Bemerkung. Die Ungleichungen können auf Reihen und Funktionen verallgemeinert werden. (Integralnormen)

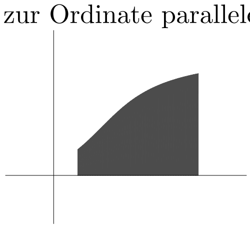
5 Integralrechnung

Im Folgenden soll das Problem der Bestimmung von Flächeninhalten behandelt werden.
Die Fläche eines Dreiecks ($A = \frac{g \cdot h}{2}$, wenn g die Grundseite und h die Höhe ist) oder eines Rechtecks ($A = a \cdot b$, wobei a und b die Seitenlängen sind) läßt sich relativ einfach einfach berechnen.
Anders verhält es sich beispielsweise mit der Fläche eines Kreises. Diese kann durch den Flächeninhalt vieler einfacher Figuren approximiert werden; in der Hoffnung, daß die Abweichung vom tatsächlichen Wert kleiner wird, je feiner die Approximation ist.

[26] *Herman Schwarz* (1843-1921)

[27] nach *Hermann Minkowski* (1864-1909)

Um diese Idee zu verallgemeinern und analytisch umzusetzen, wird nun der Inhalt der Fläche, die von dem Graphen einer Funktion, der Abszisse und zwei zur Ordinate parallelen Geraden eingeschlossen wird, betrachtet.



5.1 Treppenfunktionen

Definition 5.1. Sei $a < b$. Eine Funktion $t : [a, b] \to \mathbb{R}$ heißt **Treppenfunktion**, falls es eine Unterteilung von $[a, b]$ in endlich viele Punkte $x_0 = a < x_1 < \ldots < x_{N-1} < x_N = b$ gibt, sodaß t auf jedem Intervall (x_{i-1}, x_i) mit $i = 1, \ldots N$ konstant ist.
Die Menge der Treppenfunktionen auf $[a, b]$ ist
$T(a, b) := \{t : [a, b] \to \mathbb{R} : t \text{ istTreppenfunktion}\}$

Satz 5.1. *$T(a, b)$ ist ein $\mathbb{R}$-Vektorraum.*

Beweis. Seien $t_1, t_2 \in T(a, b)$ zu den Unterteilungen $x_1^1, x_2^1, \ldots x_n^1$ und $x_1^2, x_2^2, \ldots x_n^2$. Dann ist $t_1 + t_2$ eine Treppenfunktion zu der Unterteilung mit den Punkten $\{x_i^1\} \cup \{x_i^2\}$.
Die übrigen Eigenschaften eines Vektorraums lassen sich ebenso leicht überprüfen. □

Definition 5.2. $\mathcal{I} : T(a, b) \to \mathbb{R}$ ist die Abbildung mit

$$\mathcal{I}(t) = \sum_{i=1}^{n} (x_i - x_{i-1}) \cdot t\left(\frac{x_{i-1} + x_i}{2}\right),$$

falls $x_0, x_1, \ldots x_n$ eine Unterteilung zu t ist.
$\mathcal{I}(t)$ heißt **Integral der Treppenfunktion** t über $[a, b]$.

$\mathcal{I}(t)$ ist der Flächeninhalt unter der Treppenfunktion t.

Im Folgenden soll gezeigt werden, daß $\mathcal{I}(t)$ wohldefiniert, d.h. unabhängig von der Unterteilung zu t ist.
Dazu seien $x_1, x_2, \ldots x_M$ und $y_1, y_2, \ldots y_N$ Unterteilungen zu t.
Dann ist auch $\{z_i\} = \{x_i\} \cup \{y_i\}$ eine Unterteilung zu t.
Seien nun $0 = j_0 < j_1 < \ldots < j_M = K$ so gewählt, daß $x_i = z_{j_i}$ für alle i ist.
Dann gilt: $\sum_{i=1}^{M} (x_i - x_{i-1})\ t\left(\frac{x_{i-1}+x_i}{2}\right)$
$= \sum_{i=1}^{M} \left(\sum_{j=j_{i-1}+1}^{j_i} (z_j - z_{j-1})\right)\ t\left(\frac{x_{i-1}+x_i}{2}\right) = \sum_{j=1}^{K} (z_{j-1} - z_j)\ t\left(\frac{z_{j-1}+z_j}{2}\right)$

Das letzte Gleichheitszeichen gilt, weil t in jedem Intervall (x_{i-1}, x_i) konstant ist. Mit analoger Argumentation folgt
$\sum_{i=1}^N (y_i - y_{i-1})\ t\left(\frac{y_{i-1}+y_i}{2}\right) = \sum_{j=1}^K (z_{j-1} - z_j)\ t\left(\frac{z_{j-1}+z_j}{2}\right)$, also
$\sum_{i=1}^N (y_i - y_{i-1})\ t\left(\frac{y_{i-1}+y_i}{2}\right) = \sum_{i=1}^M (x_i - x_{i-1})\ t\left(\frac{x_{i-1}+x_i}{2}\right)$.

Bemerkung. Die Funktionswerte einer Treppenfunktion t an den Unterteilungsstellen sind unerheblich für $\mathcal{I}(t)$.

Satz 5.2. *Die Abbildung $\mathcal{I} : T(a,b) \to \mathbb{R}$ ist linear und monoton, d.h.* $t_1 \leq t_2 \quad \Rightarrow \quad \mathcal{I}(t_1) \leq \mathcal{I}(t_2)$.

5.2 Das Riemann-Integral

Definition 5.3. Sei $f : [a,b] \to \mathbb{R}$ eine beschränkte Funktion. Dann heißt

$$\int_a^{b\ *} f(x)\mathrm{d}x := \inf\{\mathcal{I}(t):\ \ t \in T(a,b),\, t \geq f\}$$

das **Oberintegral** und

$$\int_{a\ *}^{b} f(x)\mathrm{d}x := \sup\{\mathcal{I}(t):\ \ t \in T(a,b),\, t \leq f\}$$

das **Unterintegral** von f über $[a,b]$.

Abkürzend wird im Folgenden auch $\int^* f$ anstelle von $\int_a^{b\ *} f(x)\mathrm{d}x$ und $\int_* f$ statt $\int_{a\ *}^b f(x)\mathrm{d}x$ geschrieben

Satz 5.3. *Seien $f, g : [a,b] \to \mathbb{R}$ beschränkt und $\lambda \in \mathbb{R}$. Dann gilt*

(a) $\int^*(f+g) \leq \int^* f + \int^* g$

(b) $\int_*(f+g) \geq \int_* f + \int_* g$

(c) $\int^* \lambda f = \lambda \int^* f$, *falls* $\lambda \geq 0$

(d) $\int_* \lambda f = \lambda \int_* f$, *falls* $\lambda \geq 0$

(e) $\int^* \lambda f = \lambda \int_* f$, *falls* $\lambda \leq 0$

(f) $\int_* f \leq \int^* f$

Beweis. (a) Zu $\varepsilon > 0$ gibt es $\varphi, \psi \in T(a,b)$ mit $\varphi \geq f$, $\psi \geq g$,
$\mathcal{I}(\varphi) \leq \int_* f + \varepsilon$ und $\mathcal{I}(\psi) \leq \int^* g + \varepsilon$.
$\Rightarrow \quad \varphi + \psi \geq f + g$
$\Rightarrow \quad \int^*(f+g) \leq \mathcal{I}(\varphi+\psi) = \mathcal{I}(\varphi) + \mathcal{I}(\psi) \leq \int^* f + \int^* g\ + 2\varepsilon$
Diese Relation bleibt für $\varepsilon \to 0$ erhalten.

$\Rightarrow \quad \int^* (f+g) \leq \int^* f + \int^* g$

(c) Zu $\varepsilon > 0$ existiert $\varphi \in T(a,b),$ sodaß $\varphi \geq f$ und $\mathcal{I}(\varphi) \leq \int^* f + \varepsilon$ ist.

Damit ist auch $\lambda\varphi \geq \lambda f$, denn $\lambda \geq 0$ nach Voraussetzung.

$\Rightarrow \quad \int^* \lambda f \leq \mathcal{I}(\lambda\varphi) = \lambda\mathcal{I}(\varphi) \leq \lambda \int^* f + \lambda\varepsilon$

Für $\varepsilon \to 0$ ergibt sich daraus $\int^* \lambda f \leq \lambda \int^* f$.

Die Behauptung ist trivial für $\lambda = 0$; andernfalls gibt es ein $\overline{\psi} \in T(a,b)$ mit $\overline{\psi} \geq \lambda f$ und $\mathcal{I}(\overline{\psi}) \leq \int^* \lambda f + \varepsilon. \quad \Rightarrow \quad \psi := \frac{\overline{\psi}}{\lambda} \geq f$

$\Rightarrow \quad \lambda \int^* f \leq \lambda\mathcal{I}(\psi) = \mathcal{I}(\lambda\psi) = \mathcal{I}(\overline{\psi}) \leq \int^* \lambda f + \varepsilon.$

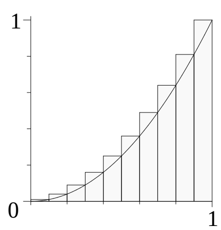
Für $\varepsilon \to 0$ folgt $\int^* \lambda f \geq \lambda \int^* f$ und schließlich $\int^* \lambda f = \lambda \int^* f$.

Mit $\int_* f = -\int^* (-f)$ folgen die Behauptungen (b), (d) und (e).

(f) ist eine direkte Folgerung aus der Monotonie von $\mathcal{I}$. □

Beispiel. Sei $f : [0,1] \to \mathbb{R}$ definiert durch $f(x) = x^2$.

Zu $N \in \mathbb{N}$ betrachte man die äquidistante Unterteilung des Intervalls $[0,1]$ mit den Punkten $x_i = \frac{i}{N}$, $i = 0 \ldots N$ und die Treppenfunktion $t_N \in T(0,1)$ mit $t_N|_{(x_i-1,x_i]} = f(x_i)$.



f wächst streng monoton, daher ist $t_N \geq f$.

$\mathcal{I}(t_N) = \sum_{i=1}^N (x_i - x_{i-1})\ t_N \left(\frac{x_i + x_{i-1}}{2}\right) = \sum_{i=1}^N \frac{1}{N} \left(\frac{i}{N}\right)^2$

$= \frac{1}{N^3} \sum_{i=1}^N i^2 = \frac{N(N+1)(2N+1)}{6N^3}$

Somit ist $\lim_{N\to\infty} \mathcal{I}(t_N) = \frac{1}{3}$ und damit $\int_0^{1\,*} x^2 \mathrm{d}x \leq \frac{1}{3}$.

Analog läßt sich $\int_{0\,*}^1 x^2 \mathrm{d}x \geq \frac{1}{3}$ zeigen.

Mit Satz 5.3 (f) folgt $\int_{0\,*}^1 x^2 \mathrm{d}x = \int_0^{1\,*} x^2 \mathrm{d}x = \frac{1}{3}$.

Beispiel. Sei $f : [0,1] \to \mathbb{R}$ mit $f(x) = \begin{cases} 1 & \text{für } x \in \mathbb{Q} \\ 0 & \text{für } x \in \mathbb{R}\backslash\mathbb{Q} \end{cases}$

$\mathbb{Q}$ liegt dicht in $\mathbb{R}$, daher gilt für jedes $t \in T(0,1)$ mit $t \geq f$:

$t(x) \geq 1$ für alle x, bis auf die Unterteilungsstellen zu t.

$\Rightarrow \quad \int_0^{1\,*} f(x) \mathrm{d}x \geq 1$

Zudem ist $t_0 = 1$ eine Treppenfunktion mit der Eigenschaft $t_0 \geq f$ und $\mathcal{I}(t_0) = 1 \quad \Rightarrow \quad \int_0^{1\,*} f(x)\mathrm{d}x = 1$.

In ähnlicher Weise folgt $\int_{0\,*}^1 f(x)\mathrm{d}x = 0$.

Dieses Beispiel zeigt, daß Ober- und Unterintegral einer Funktion im allgemeinen nicht übereinstimmen müssen.

Definition 5.4. Sei $f : [a, b] \to \mathbb{R}$ beschränkt. f heißt **Riemann-integrierbar,**[28] falls $\int_{a\,*}^{b} f(x)\mathrm{d}x = \int_{a}^{b\,*} f(x)\mathrm{d}x$ ist. Der Wert

$$\int_a^b f(x)\mathrm{d}x := \int_{a\,*}^{b} f(x)\mathrm{d}x = \int_a^{b\,*} f(x)\mathrm{d}x$$

heißt **Riemann-Integral** von f über $[a, b]$.

Satz 5.4. *Die Menge* $V = \{f : [a, b] \to \mathbb{R}\,,\ \ f \,\text{ist}\, \textsc{Riemann} - \text{integrierbar}\}$ *bildet einen reellen Vektorraum. Das* Riemann*-Integral ist eine Linearform auf* V *(eine lineare Abbildung* $V \to \mathbb{R}$*) und monoton, d.h.*

(a) $f, g \in V \quad \Rightarrow \quad f + g \in V$ *und*
$\int_a^b f(x)\mathrm{d}x + \int_a^b g(x)\mathrm{d}x = \int_a^b (f(x) + g(x))\mathrm{d}x$

(b) $f \in V,\ \lambda \in \mathbb{R} \quad \Rightarrow \quad \lambda f \in V$ *und* $\int_a^b \lambda f(x)\mathrm{d}x = \lambda \int_a^b f(x)\mathrm{d}x$

(c) $f, g \in V,\ f \leq g \quad \Rightarrow \quad \int_a^b f(x)\mathrm{d}x \leq \int_a^b g(x)\mathrm{d}x$

Der Beweis erfolgt mit Satz 5.3.

Im Folgenden wird auch „integrierbar" anstelle von „Riemann-integrierbar" geschrieben.

Satz 5.5. *Ist die Funktion* $f : [a, b] \to \mathbb{R}$ *stetig, so ist sie auch integrierbar.*

Beweis. Nach Satz 3.12 gibt zu jedem $\varepsilon > 0$ Treppenfunktionen $\varphi, \psi \in T(a, b)$ mit $\varphi \leq f \leq \psi$ und $\psi - \varphi \leq \varepsilon$.
Damit wird $\int_a^{b\,*} f(x)\mathrm{d}x - \int_{a\,*}^{b} f(x)\mathrm{d}x \leq \mathcal{I}(\psi) - \mathcal{I}(\phi) = \mathcal{I}(\psi - \phi) \leq (b - a)\varepsilon$.
Für $\varepsilon \to 0$ hat dies den Grenzwert 0, woraus $\int_a^{b\,*} f(x)\mathrm{d}x = \int_{a\,*}^{b} f(x)\mathrm{d}x$, d.h. die Integrierbarkeit von f folgt. □

Satz 5.6. *Ist* $f : [a, b] \to \mathbb{R}$ *monoton, dann ist* f *auch integrierbar.*

Beweis. O.B.d.A. sei f monoton wachsend.
Die Treppenfunktionen $t_*, t^* \in T(a, b)$ zu der Unterteilung mit den Punkten $x_i = \frac{i(b-a)}{N}$, $N \in \mathbb{N}$ seien durch $t_*|_{(x_{i-1},x_i)} = f(x_{i-1})$ und $t^*|_{(x_{i-1},x_i)} = f(x_i)$ definiert. Dann gilt $t_* \leq f \leq t^*$ und somit
$\int_a^{b\,*} f(x)\mathrm{d}x - \int_{a\,*}^{b} f(x)\mathrm{d}x \leq \mathcal{I}(t^*) - \mathcal{I}(t_*) = \mathcal{I}(t^* - t_*)$
$= \sum_{i=1}^{N} \frac{1}{N}(f(x_i) - f(x_{i-1})) = \frac{1}{N}(f(b) - f(a))$
Für $N \to \infty$ bleibt die Relation erhalten, d.h.
$\int_a^{b\,*} f(x)\mathrm{d}x - \int_{a\,*}^{b} f(x)\mathrm{d}x \leq \lim_{N\to\infty} \frac{1}{N}(f(b) - f(a)) = 0$
$\Rightarrow \quad \int_a^{b\,*} f(x)\mathrm{d}x = \int_{a\,*}^{b} f(x)\mathrm{d}x$ □

[28] *Bernhard Riemann* (1826-1866)

Definition 5.5. Sei $f : [a, b] \to \mathbb{R}$ eine reelle Funktion,
$a = x_0 < x_1 < \ldots < x_N = b$ eine Unterteilung von $[a, b]$ und
$\xi_i \in [a, b]\,, \;\; i = 1 \ldots N$ Stützstellen. Dann heißt

$$\sum_{i=1}^{N} (x_i - x_{i-1}) f(\xi_i)$$

die **Riemannsche Summe** von f bezüglich der Unterteilung
$a = x_0 < x_1 < \ldots < x_N = b$ und der Stützstellen $\xi_1, \ldots \xi_N$.
Die Zahl $\eta := \max\{|x_i - x_{i-1}|, \;\; i = 1 \ldots N\}$ heißt **Feinheit** der Unterteilung.

Satz 5.7. *$f : [a, b] \to \mathbb{R}$ sei* RIEMANN-*integrierbar.*
Dann existiert zu jedem $\varepsilon > 0$ ein $\eta > 0$, sodaß für jede Unterteilung $a = x_0 < \ldots x_N = b$ der Feinheit $\leq \eta$ und jede Wahl von Stützstellen $\xi_i \in [x_{i-1}, x_i]$ gilt:

$$\left| \int_a^b f(x)\mathrm{d}x - \sum_{i=1}^{N} (x_i - x_{i-1}) f(\xi_i) \right| \leq \varepsilon$$

Beweis. Sei $\varepsilon > 0$. Nach Voraussetzung ist f integrierbar, d.h.
$\exists\, \varphi, \psi \in T(a, b)$ mit $\varphi \leq f \leq \psi$ und $\mathcal{I}(\psi - \varphi) \leq \frac{\varepsilon}{2}$.
Sei $a = y_0 < y_1 < \ldots < y_M = b$ die gemeinsame Unterteilung zu φ und ψ.
Es ist $-B \leq \psi$, sowie $\varphi \leq B$, wobei $B = ||f||_{[a,b]}$.
Nun wähle man $\eta := \frac{\varepsilon}{8BM}$
Zu einer Unterteilung $a = x_0 < \ldots < x_N = b$ von $[a, b]$ der Feinheit $\leq \eta$ und Stützstellen $\xi_i \in [x_{i-1}, x_i]$ sei $t \in T(a, b)$ die Treppenfunktion der dazugehörigen RIEMANN- Summe, d.h. $t|_{(x_{i-1}, x_i)} = f(\xi_i)$.
Dann ist $\varphi - 2B \leq -B \leq t \leq B \leq \psi + 2B$
Falls $[x_{i-1}, x_i] \subset (y_{j-1}, y_j)$ ist, dann gilt sogar
$\varphi(x) \leq t(x) \leq \psi(x) \;\; \forall\, x \in [x_{i-1}, x_i]$
Sei eine weitere Treppenfunktion $s \in T(a, b)$ definiert durch

$$s(x) = \begin{cases} 0 & \text{falls } \exists\, i, j : \;\; x \in (x_{i-1}, x_i) \text{ und } [x_{i-1}, x_i] \subset (y_{j-1}, y_j) \\ 2B & \text{sonst} \end{cases}$$

Es folgt $\varphi - s \leq t \leq \psi + s$.
Überdies gibt es maximal $2M$ Intervalle $[x_{i-1}, x_i]$, auf denen $s = 2B$ ist.
Damit wird
$\int_a^b f(x)\mathrm{d}x \leq \mathcal{I}(\varphi) + \frac{\varepsilon}{2} \leq \mathcal{I}(\varphi - s) + \frac{\varepsilon}{2} + \eta \cdot (2M) \cdot (2B) = \mathcal{I}(\varphi - s) + \varepsilon \leq \mathcal{I}(t) + \varepsilon$
und $\int_a^b f(x)\mathrm{d}x \geq \mathcal{I}(\psi) - \frac{\varepsilon}{2} \geq \mathcal{I}(\psi + s) - \varepsilon \geq \mathcal{I}(t) - \varepsilon$,
also $\left| \int_a^b f(x) - \mathcal{I}(t)\mathrm{d}x \right| \leq \varepsilon$. ☐

Bemerkung. Es gilt auch die Umkehrung von Satz 5.7 (vgl. *Übungsaufgabe 60*).

Beispiel. Für $a > 0$ soll $\int_0^a \cos(x)\mathrm{d}x$ bestimmt werden.
Die Riemann-Summe bzgl. der äquidistanten Unterteilung von $[0, a]$ mit den Punkten $x_k = \frac{ka}{n}$ und der Stützstellen $\xi_k = x_k$ ist $s_n = \sum_{k=1}^n \frac{a}{n}\cos\left(\frac{ka}{n}\right)$.

Lemma. Ist $t \in \mathbb{R}$ kein ganzzahliges Vielfaches von 2π, dann gilt $\forall\, n \in \mathbb{N}$:

$$\frac{1}{2} + \sum_{k=1}^{n} \cos(kt) = \frac{\sin\left(t\left(n+\frac{1}{2}\right)\right)}{2\sin\left(\frac{t}{2}\right)}$$

Beweis. $\cos kt = \frac{1}{2}\left(e^{\mathbf{i}kt} + e^{-\mathbf{i}kt}\right)$.
$\Rightarrow \quad \frac{1}{2} + \sum_{k=1}^n \cos(kt) = \frac{1}{2}\sum_{k=-n}^n e^{\mathbf{i}kt} = \frac{1}{2}e^{-\mathbf{i}nt}\sum_{k=0}^{2n} e^{\mathbf{i}kt} = \frac{1}{2}e^{-\mathbf{i}nt}\frac{1-e^{(2n+1)\mathbf{i}t}}{1-e^{\mathbf{i}t}}$

$$= \frac{1}{2}\frac{-e^{-\mathbf{i}\frac{t}{2}}\left(e^{-\mathbf{i}nt} - e^{(n+1)\mathbf{i}t}\right)}{-e^{-\mathbf{i}\frac{t}{2}}\left(1-e^{\mathbf{i}t}\right)} = \frac{1}{2}\frac{e^{\mathbf{i}\left(n+\frac{1}{2}\right)t} - e^{-\mathbf{i}\left(n+\frac{1}{2}\right)t}}{e^{\mathbf{i}\frac{t}{2}} - e^{-\mathbf{i}\frac{t}{2}}} = \frac{\sin\left(t\left(n+\frac{1}{2}\right)\right)}{2\sin\left(\frac{t}{2}\right)}$$

□

Damit ist

$$s_n = \frac{a}{n}\left(\frac{\sin\left(\frac{a}{n}\left(n+\frac{1}{2}\right)\right)}{2\sin\left(\frac{a}{2n}\right)} - \frac{1}{2}\right) = \frac{\frac{a}{2n}}{\sin(\frac{a}{2n})}\sin\left(a + \frac{a}{2n}\right) - \frac{a}{2n}$$

Zusammen mit $\lim_{n\to\infty} \frac{\frac{a}{2n}}{\sin(\frac{a}{2n})} = 1$ liefert dies

$$\int_0^a \cos(x)\mathrm{d}x = \lim_{n\to\infty} s_n = \sin(a)$$

Definition 5.6. Sei $f : X \to \mathbb{R}$ eine reelle Funktion. Dann ist

$$f_+ := \begin{cases} f(x) & \text{falls } f(x) > 0 \\ 0 & \text{sonst} \end{cases}$$

$$f_- := \begin{cases} -f(x) & \text{falls } f(x) < 0 \\ 0 & \text{sonst} \end{cases}$$

Es gilt $f = f_+ - f_-$, sowie $|f| = f_+ + f_-$.

Satz 5.8. *Sind $f, g : [a, b] \to \mathbb{R}$ integrierbar, dann folgt*

- *f_+ und f_- sind integrierbar*
- *$\forall\, p \in [1, \infty)$ ist $|f|^p$ integrierbar.*
- *$fg : [a, b] \to \mathbb{R}$ ist integrierbar.*

Beweis. Nach Voraussetzung gibt es $\forall\ \varepsilon > 0$ Treppenfunktionen $\varphi, \psi \in T(a, b)$ mit $\varphi \leq f \leq \psi$ und $\int_a^b (\varphi - \psi)(x)\mathrm{d}x \leq \varepsilon$.
Also sind φ_+, ψ_+ Treppenfunktionen mit $\varphi_+ \leq f_+ \leq \psi_+$ und
$\int_a^b (\varphi_+ - \psi_+)(x)\mathrm{d}x \leq \int_a^b (\varphi - \psi)(x)\mathrm{d}x \leq \varepsilon$, d.h. f_+ ist integrierbar.

Entsprechend folgt die Integrierbarkeit von f_-.
Wegen Satz 5.4 und $|f| = f_+ + f_-$ ist damit auch $|f|$ integrierbar.
Insbesondere ist $|f|$ beschränkt auf $[a,b]$. Nun soll gezeigt werden, daß $|f|^p$ integrierbar ist. Sei dazu zunächst $0 \leq ||f||_{[a,b]} \leq 1$.
$|f|$ ist integrierbar, d.h. $\forall\ \varepsilon > 0\ \exists\, \varphi, \psi \in T(a,b)$ mit $0 \leq \varphi \leq |f| < \psi$ und $\int_a^b (\psi - \varphi)(x)\ \mathrm{d}x \leq \frac{\varepsilon}{p}$.
Dann sind φ^p, ψ^p Treppenfunktionen mit $\varphi^p \leq |f|^p \leq \psi^p$.
Nach dem Mittelwertsatz der Differentialrechnung $\exists\, \xi \in [\varphi(x), \psi(x)]$, sodaß $\frac{(\psi(x))^p - (\varphi(x))^p}{\psi(x) - \varphi(x)} = \left(\frac{\mathrm{d}}{\mathrm{d}\alpha}\alpha^p\right)(\xi) = p\ \xi^{p-1}$ ist.
$\Rightarrow \quad \psi^p - \varphi^p \leq p(\psi - \varphi)$
$\Rightarrow \quad \int_a^b (\psi^p - \varphi^p)(x)\ \mathrm{d}x \leq p \int_a^b (\psi - \varphi)(x)\ \mathrm{d}x \leq \varepsilon$, d.h. $|f|^p$ ist integrierbar.
Für den Fall $||f||_{[a,b]} > 1$ sei $\tilde{f} := \frac{1}{||f||_{[a,b]}}\ f$. Dann ist $||\tilde{f}||_{[a,b]} \leq 1$.
Somit ist $|\tilde{f}|^p$ und nach Satz 5.4 auch $|f|^p = ||f||_{[a,b]}^p \cdot |\tilde{f}|^p$ integrierbar.
Schließlich folgt aus obigem, daß auch

$$fg = \frac{1}{4}\left[(f+g)^2 - (f-g)^2\right]$$

integrierbar ist. $\square$

Man beachte, daß im allgemeinen $\int_a^b f(x)g(x)\mathrm{d}x \neq \int_a^b f(x)\mathrm{d}x \int_a^b g(x)\mathrm{d}x$ ist.

Satz 5.9 (Mittelwertsatz der Integralrechnung). *Sind $f, \varphi : [a,b] \to \mathbb{R}$ stetig und ist $\varphi \geq 0$, so existiert ein $\xi \in [a,b]$ mit*

$$\int_a^b f(x)\varphi(x)\mathrm{d}x = f(\xi) \int_a^b \varphi(x)\mathrm{d}x$$

Für den Spezialfall $\varphi = 1$ ergibt sich

$$\int_a^b f(x)\mathrm{d}x = f(\xi)(b-a)$$

Beweis. Sei $m = \inf\{f(x) : x \in [a,b]\}$ und $M = \sup\{f(x) : x \in [a,b]\}$.
Dann ist $m\varphi \leq f\varphi \leq M\varphi$.
Mit Satz 5.4 folgt daraus $\quad m\int_a^b \varphi(x)\mathrm{d}x \leq \int_a^b f(x)\varphi(x)\mathrm{d}x \leq M\int_a^b \varphi(x)\mathrm{d}x$.
Daher existiert ein $\quad \mu \in [m, M]$ mit $\int_a^b f(x)\varphi(x)\mathrm{d}x = \mu \int_a^b \varphi(x)\mathrm{d}x$.
Und nach dem Zwischenwertsatz ist $\quad \mu = f(\xi)$ für ein $\xi \in [a,b]$. $\square$

Satz 5.10. *Sei $a < b < c$. $f : [a,c] \to \mathbb{R}$ ist dann und nur dann integrierbar, wenn $f|_{[a,b]}$ und $f_{[b,c]}$ dies sind und $\int_a^c f(x)\mathrm{d}x = \int_a^b f(x)\mathrm{d}x + \int_b^c f(x)$ gilt.*

Der Beweis kann unter Verwendung von Treppenfunktionen und der Definition des Integrals erfolgen.

5.3 Der Hauptsatz der Differential- und Integralrechnung

An dieser Stelle sei nocheinmal herausgestellt, daß es sich bei der Differenzierbarkeit einer Funktion um eine lokale Eigenschaft handelt; daher ist die Ableitung einer Funktion in natürlicher Weise wieder eine Funktion. Das Integral hingegen ist über einem Intervall des Definitionsbereiches, also global definiert. Es ist ein reeller Wert, welcher der Funktion und dem Intervall zugeordnet ist, läßt sich also als Funktion von der oberen Integrationsgrenze auffassen: $F(x) = \int_a^x f(x)\mathrm{d}x$

Satz 5.11. *Ist $f : I \to \mathbb{R}$ stetig auf dem Intervall I und $a \in I$, dann ist $F : I \to \mathbb{R}$ mit $F(x) = \int_a^x f(t)\mathrm{d}t$ differenzierbar auf I und es gilt $F' = f$.*

Beweis. Für $h \neq 0$ ist $\frac{F(x+h)-F(x)}{h} = \frac{1}{h}\left(\int_a^{x+h} f(t)\mathrm{d}t - \int_a^x f(t)\mathrm{d}t\right)$
$= \frac{1}{h}\left(\int_a^x f(t)\mathrm{d}t + \int_x^{x+h} f(t)\mathrm{d}t - \int_a^x f(t)\mathrm{d}t\right) = \frac{1}{h}\int_x^{x+h} f(t)\mathrm{d}t.$
Nach dem Mittelwertsatz der Integralrechnung existiert ein $\xi_h \in [x, x+h]$ (bzw. $\xi_h \in [x+h, x]$, falls $h < 0$), sodaß $f(\xi_h) \cdot h = \int_x^{x+h} f(t)\mathrm{d}t$ ist.
Mit der Stetigkeit von f folgt daraus
$F'(x) = \lim_{h\to 0} \frac{F(x+h)-F(x)}{h} = \lim_{h\to 0} f(\xi_h) = f(x)$ □

Definition 5.7. Eine Funktion $F : I \to \mathbb{R}$ heißt **Stammfunktion** von $f : I \to \mathbb{R}$, falls sie differenzierbar und $F' = f$ ist.

Satz 5.12. *Ist $F : I \to \mathbb{R}$ eine Stammfunktion von $f : I \to \mathbb{R}$, so ist $G : I \to \mathbb{R}$ genau dann Stammfunktion von f, wenn $F - G$ konstant ist.*

Beweis. Zunächst sei $F - G = c$ für ein $c \in \mathbb{R}$.
$\Rightarrow \quad G' = (F-c)' = F' = f,$ d.h. G ist eine Stammfunktion von f.
Die Umkehrung folgt aus Satz 4.13:
$G' = f = F' \quad \Rightarrow \quad G' - F' = 0 \quad \Rightarrow \quad G - F = const.$ □

Satz 5.13 (Hauptsatz der Differential- und Integralrechnung). *Sei $f : I \to \mathbb{R}$ stetig und F eine Stammfunktion von f. Dann gilt $\forall\, a, b \in I$:*

$$\int_a^b f(x)\mathrm{d}x = F(b) - F(a)$$

Statt $F(b) - F(a)$ wird auch $F(x)|_a^b$ geschrieben.

Beweis. Nach Satz 5.11 ist $F_0 : I \to \mathbb{R},\ F_0(x) = \int_a^x f(t)\mathrm{d}t$ eine Stammfunktion von f. Es ist $F_0(a) = 0$ und $F_0(b) = \int_a^b f(t)\mathrm{d}t$.
Für jede weitere Stammfunktion F von f gibt es ein $c \in \mathbb{R}$ mit $F - F_0 = c$.
$\Rightarrow \quad F(b) - F(a) = F_0(b) - F_0(a) = \int_a^b f(t)\mathrm{d}t$ □

Anstelle von $\int_a^b f(x)\mathrm{d}x = F(x)|_a^b$ schreibt man auch $\int f(x)\mathrm{d}x = F(x)$.

Beispiel.

- Für $s \in \mathbb{R}$, $s \neq -1$ ist $\int_a^b x^s \ \mathrm{d}x = \frac{x^{s+1}}{s+1}|_a^b$.
  Dabei sind die Integationsgrenzen für $s \in \mathbb{N}$ beliebig, für $s \in \mathbb{Z}$, $s \leq -2$ müssen entweder $a, b > 0$ oder $a, b < 0$ sein und für $s \notin \mathbb{Z}$ müssen $a, b > 0$ sein.
- $\int \sin(x) \ \mathrm{d}x = -\cos(x), \quad \int \cos(x) \ \mathrm{d}x = \sin(x)$
- $\int \exp(x) \ \mathrm{d}x = \exp(x)$
- $\int \frac{1}{1+x^2} \ \mathrm{d}x = \arctan(x)$

5.4 Integrationsregeln

Zusammen mit dem Hauptsatz der Differential- und Integralrechnung ergeben sich aus den Differentiationsregeln Möglichkeiten, Integrale umzuformen bzw. zu lösen.

Satz 5.14 (Partielle Integration). *Sind $f, g : [a, b] \to \mathbb{R}$ stetig differenzierbar, so ist*

$$\int_a^b f(x)g'(x) \ \mathrm{d}x = f(x)g(x)|_a^b - \int_a^b f'(x)g(x) \ \mathrm{d}x$$

Beweis. Für $F = f \cdot g$ gilt aufgrund der Produktregel
$F'(x) = f'(x)g(x) + f(x)g'(x)$
Und mit dem Hauptsatz folgt daraus
$\int_a^b f(x)g'(x) \ \mathrm{d}x + \int_a^b f'(x)g(x) \ \mathrm{d}x = F(x)|_a^b = f(x)g(x)|_a^b$ □

Satz 5.15 (Substitutionsregel). *Ist $f : I \to \mathbb{R}$ stetig, $\varphi : [a, b] \to \mathbb{R}$ stetig differenzierbar und $\varphi([a, b]) \subset I$, dann ist*

$$\int_a^b f(\varphi(t))\varphi'(t) \ \mathrm{d}t = \int_{\varphi(a)}^{\varphi(b)} f(x) \ \mathrm{d}x$$

Beweis. Sei $F : I \to \mathbb{R}$ eine Stammfunktion von f.
$\Rightarrow \quad (F \circ \varphi)'(t) = F'(\varphi(t))\varphi'(t) = f(\varphi(t))\varphi'(t)$
$\Rightarrow \quad \int_a^b f(\varphi(t))\varphi'(t) \ \mathrm{d}t = F(\varphi(b)) - F(\varphi(a)) = \int_{\varphi(a)}^{\varphi(b)} f(x) \ \mathrm{d}x$ □

Mit $\mathrm{d}\varphi(t) = \varphi'(t)\mathrm{d}t$ schreibt sich obige Substitutionsregel als
$\int_a^b f(\varphi(t)) \ \mathrm{d}\varphi = \int_{\varphi(a)}^{\varphi(b)} f(x) \ \mathrm{d}x$.

Beispiel. Zur Substitutionsregel:

- $\int_a^b f(ct + d) \ \mathrm{d}t = \frac{1}{c} \int_{ac+d}^{bc+d} f(x) \ \mathrm{d}x$
  Hier wurde $ct + d = \varphi(t)$ substituiert.

- $\int_a^b 2tf(t^2)\ \mathrm{d}t = \int_{a^2}^{b^2} f(x)\ \mathrm{d}x$ mit der Substitution $t^2 = \varphi(t)$
- $\int_a^b \frac{t}{\sqrt{1-t^2}}\ \mathrm{d}t = \int_{1-a^2}^{1-b^2} \frac{-1}{2\sqrt{x}}\ \mathrm{d}x = -\sqrt{x}\ |_{1-a^2}^{1-b^2} = -\sqrt{1-t^2}\ |_a^b$
  mit $x = 1 - t^2$.

Zur partiellen Integration:

- $\int_a^b \ln(x)\ \mathrm{d}x = \int_a^b 1 \cdot \ln(x)\ \mathrm{d}x = x\ln(x)|_a^b - \int_a^b 1\ \mathrm{d}x = x(\ln(x) - 1)|_a^b$
- $\int_a^b \arcsin(x)\ \mathrm{d}x = x\arcsin(x)|_a^b - \int_a^b \frac{x}{\sqrt{1-x^2}}\ \mathrm{d}x$
  $= \left(x\arcsin(x) + \sqrt{1-x^2}\right)|_a^b$

Beispiel. Sei $I_m = \int \sin^m x\ \mathrm{d}x$. Durch partielle Integration ergibt sich für
$m \geq 2$: $\quad I_m = -\cos x \sin^{m-1} x + (m-1)\int \cos^2 x \sin^{m-2} x\ \mathrm{d}x$
$= -\cos x \sin^{m-1} x + (m-1)\int (1 - \sin^2 x)\sin^{m-2} x\ \mathrm{d}x$
$= -\cos x \sin^{m-1} x + (m-1)(I_{m-2} - I_m)$
Es folgt die Rekursionsformel
$I_m = -\frac{1}{m}\cos x \sin^{m-1} x + \frac{m-1}{m} I_{m-2}$ mit $I_0 = x$ und $I_1 = -\cos x$.
Setzt man die Integrationsgrenzen 0 und $\frac{\pi}{2}$ ein, so ergibt sich für
$A_m = \int_0^{\frac{\pi}{2}} \sin^m x\ \mathrm{d}x$ die Rekursion $A_m = \frac{m-1}{m} A_{m-2}$, $A_0 = \frac{\pi}{2}$, $A_1 = 1$;
somit ist $A_{2n} = \frac{(2n-1)(2n-3)\ldots 3\cdot 1}{2n(2n-2)\ldots 4\cdot 2}\frac{\pi}{2}$, $\quad A_{2n+1} = \frac{2n(2n-2)\ldots 4\cdot 2}{(2n+1)(2n-1)\ldots 3}$.
Für $x \in [0, \frac{\pi}{2}]$ ist $\sin^{2n+2} x \leq \sin^{2n+1} x \leq \sin^{2n} x$ und folglich
$A_{2n+2} \leq A_{2n+1} \leq A_{2n}$. $\quad \Rightarrow \quad \frac{A_{2n+2}}{A_{2n}} \leq \frac{A_{2n+1}}{A_{2n}} \leq 1$
Mit $\lim_{n\to\infty} \frac{A_{2n+2}}{A_{2n}} = \lim_{n\to\infty} \frac{2n+1}{2n+2} = 1$ folgt daraus
$\lim_{n\to\infty} \frac{A_{2n+1}}{A_{2n}} = \lim_{n\to\infty} \frac{2n\cdot 2n \ldots 4\cdot 4\cdot 2\cdot 2}{(2n+1)(2n-1)\ldots 5\cdot 3\cdot 3\cdot 1}\frac{2}{\pi} = 1.$

$$\Rightarrow \quad \frac{\pi}{2} = \prod_{n=1}^{\infty} \frac{4n^2}{4n^2 - 1}$$

Dieser als WALLISsches[29] Produkt bekannte Ausdruck ist nach dem VIETAschen Produkt[30] die zweite bekannt gewordene analytische Darstellung für die Zahl π.

Satz 5.16. *Sei $f : [a,b] \to \mathbb{R}$ stetig differenzierbar und $F(k) = \int_a^b f(x)\sin kx\ \mathrm{d}x$, $k \in \mathbb{R}$. Dann ist* $\lim_{|k|\to\infty} F(k) = 0$.

Beweis. Für $k \neq 0$ liefert partielle Integration
$F(k) = \int_a^b f(x)\sin kx\ \mathrm{d}x = -f(x)\frac{\cos kx}{k}\Big|_a^b + \frac{1}{k}\int_a^b f'(x)\cos kx\ \mathrm{d}x$
Nach Voraussetzung sind f und f' stetig auf $[a,b]$.

[29] nach *John Wallis* (1616-1703)

[30] nach *François Viéte* (1540-1603)

$\Rightarrow \quad \exists\, M \in \mathbb{R}: \quad |f(x)| \leq M, \; |f'(x)| \leq M \quad \forall\, x \in [a,b]$
Überdies ist $|\cos kx| \leq 1. \quad \Rightarrow \quad |F(k)| \leq \frac{2M}{|k|} + \frac{M}{|k|}(b-a)$
$\Rightarrow \quad \lim_{|k|\to\infty} F(k) = 0.$ $\square$

Beispiel. Für $0 < x < 2\pi$ ist

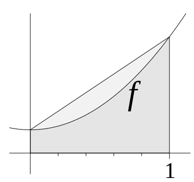
$$\sum_{k=1}^{\infty} \frac{\sin kx}{k} = \frac{\pi - x}{2}\,,$$

denn: Mit $\frac{\sin kx}{k} = \int_\pi^x \cos kt \;\mathrm{d}t$ und der in 5.2 bewiesenen Beziehung
$\sum_{k=1}^n \cos kt = \frac{\sin\left(\left(n+\frac{1}{2}\right)t\right)}{2\sin\frac{t}{2}} - \frac{1}{2}$ folgt
$S_n = \sum_{k=1}^n \frac{\sin kx}{k} = \int_\pi^x \frac{1}{2\sin\frac{t}{2}} \sin\left(\left(n+\frac{1}{2}\right)t\right) \;\mathrm{d}t - \int_\pi^x \frac{1}{2} \;\mathrm{d}t$
Und nach vorangegangenem Satz ist $\lim_{n\to\infty} S_n = -\int_\pi^x \frac{1}{2} \;\mathrm{d}t = \frac{\pi - x}{2}$.

Eine weitere Anwendung der partiellen Integration ist die Trapezregel.

Satz 5.17 (Trapezregel). *Ist $f:[0,1] \to \mathbb{R}$ zweimal stetig differenzierbar, so* $\exists\, \xi \in [0,1]:$

$$\int_0^1 f(x) \;\mathrm{d}x = \frac{1}{2}\left(f(0) + f(1)\right) - \frac{1}{12} f''(\xi)$$



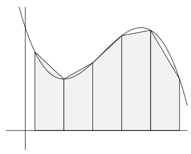
Beweis. Für die Hilfsfunktion $\varphi(x) := \frac{1}{2}x(1-x)$ gilt $\varphi \geq 0 \;\; \forall\, x \in [0,1]$,
$\varphi'(x) = \frac{1}{2} - x$ und $\varphi'' = -1$.
Zweimalige partielle Integration von f führt zu
$\int_0^1 f(x) \;\mathrm{d}x = -\int_0^1 \varphi''(x) f(x) \;\mathrm{d}x = -\varphi'(x)f(x)|_0^1 + \int_0^1 \varphi'(x) f'(x) \;\mathrm{d}x$
$= \frac{1}{2}\left(f(0) + f(1)\right) + \varphi(x)f'(x)|_0^1 - \int_0^1 \varphi(x) f''(x) \;\mathrm{d}x$
$= \frac{1}{2}\left(f(0) + f(1)\right) - \int_0^1 \varphi(x) f''(x) \;\mathrm{d}x$
Und nach dem Mittelwertsatz der Integralrechnung $\exists\, \xi \in [0,1]:$
$\int_0^1 \varphi(x) f''(x) \;\mathrm{d}x = f''(\xi) \int_0^1 \varphi(x) \;\mathrm{d}x = \frac{1}{12} f''(\xi).$ $\square$

Damit gelangt man zu einer einfachen numerischen Quadraturformel mit Fehlerabschätzung:

Satz 5.18. *Sei $f:[a,b] \to \mathbb{R}$ zweimal stetig differenzierbar und $k := ||f''||_{[a,b]}$. Weiter sei $h = \frac{b-a}{n}$, $n \in \mathbb{N}$ die Länge eines Teilintervalls der äqidistanten Unterteilung von $[a,b]$ mit den Punkten $a + ih$, $i = 0, \ldots, n$. Dann ist*

$$\int_a^b f(x) \;\mathrm{d}x = h \cdot \left(\frac{1}{2} f(a) + \sum_{i=1}^{n-1} f(a+ih) + \frac{1}{2} f(b) \right) + R$$

mit $|R| \le \frac{k}{12}(b-a)h^2$, *d.h. der Fehler* $|R|$ *nimmt quadratisch mit der Anzahl der Gitterpunkte ab.*



Beweis. Für jedes $i = 0, \ldots, n-1$ sei $g_i : [0,1] \to \mathbb{R}$ definiert durch
$g_i(x) = f(a + ih + hx)$.
g_i ist zweimal stetig differenzierbar; $g_i''(x) = h^2 f''(a + ih + hx)$.
Außerdem gilt $g_i(0) = f(a + ih)$ und $g_i(1) = f(a + (i+1)h)$.
Nach Satz 5.17 $\exists \xi_i \in [0,1]$ mit $\int_0^1 g_i(x)\, \mathrm{d}x = \frac{1}{2}\left(g_i(0) + g_i(1)\right) - \frac{1}{12} g_i''(\xi_i)$
$= \frac{1}{2}\left(f(a+ih) + f(a+(i+1)h)\right) - \frac{h^2}{12} f''(\tilde{\xi_i})$, wobei $\tilde{\xi_i} = a + ih + h\xi_i$, also
$\tilde{\xi_i} \in [a+ih, a+(i+1)h]$ ist.

$$\Rightarrow \quad I_i f = \int_{a+ih}^{a+(i+1)h} f(x) = h \int_0^1 g_i(x)\, \mathrm{d}x$$
$$= \frac{h}{2}\left(f(a+ih) + f(a+(i+1)h)\right) - \frac{h^3}{12} f''(\tilde{\xi_i})$$
$$\Rightarrow \quad \int_a^b f(x)\, \mathrm{d}x = \sum_{i=0}^{n-1} I_i f = h \cdot \left(\frac{1}{2} f(a) + \sum_{i=1}^{n-1} f(a+ih) + \frac{1}{2} f(b)\right) + R$$

mit $R = -\frac{h^3}{12} \sum_{i=0}^{n-1} f''(\tilde{\xi_i})$, für das gilt:
$|R| \le \frac{h^3}{12} \sum_{i=0}^{n-1} k = \frac{h^3}{12} nk = \frac{k}{12}(b-a)h^2$ □

Dies ist ein Beispiel für eine Quadraturformel des Typs $I_n f = \sum_{i=0}^{n} \gamma_i f(x_i)$.

5.5 Uneigentliche Integrale

In den vorangegangenen Abschnitten wurden stets Integrale von beschränkten Funktionen über endlichen Intervallen betrachtet.
Nun werden Fälle betrachtet, in denen die Funktion im Integrationsgebiet teilweise unbeschränkt ist oder in denen über einem unendlichen Intervall integriert werden soll.

Definition 5.8. Ist $f : [a, \infty) \to \mathbb{R}$ über jedem Intervall $[a, R]$ integrierbar und es existiert $\lim_{R\to\infty} \int_a^R f(x)\, \mathrm{d}x$, dann heißt $\int_a^\infty f(x)\, \mathrm{d}x$ **konvergent** und

$$\int_a^\infty f(x)\, \mathrm{d}x := \lim_{R\to\infty} \int_a^R f(x)\, \mathrm{d}x$$

Das uneigentliche Integral $\int_{-\infty}^a f(x)\, \mathrm{d}x$ wird analog definiert.

Beispiel. $\int_1^\infty \frac{1}{x^s}\, \mathrm{d}x$ konvergiert für $s > 1$ und divergiert, falls $s \le 1$ ist, denn:
Für $s \ne 1$ ist $\int_1^R \frac{1}{x^s}\, \mathrm{d}x = \frac{1}{(1-s)(x^{s-1})}\Big|_1^R = \frac{1}{s-1}\left(1 - \frac{1}{R^{s-1}}\right)$. Im Falle $s > 1$ ist $\lim_{R\to\infty} \frac{1}{R^{s-1}} = 0$, also $\int_1^\infty \frac{1}{x^s}\, \mathrm{d}x = \frac{1}{s-1}$. Für $s < 1$ ist $\lim_{R\to\infty} \frac{1}{R^{s-1}} = \infty$.
Das Integral divergiert auch, wenn $s = 1$: $\int_1^R \frac{\mathrm{d}x}{x} = \ln(R)$.

Definition 5.9. Ist $f: (a,b] \to \mathbb{R}$ über jedem Intervall $[a+\varepsilon, b]$ mit $0 < \varepsilon < b-a$ integrierbar und es existiert $\lim_{\varepsilon \searrow 0} \int_{a+\varepsilon}^b f(x)\ \mathrm{d}x$, dann heißt $\int_a^b f(x)\ \mathrm{d}x$ **konvergent** und

$$\int_a^b f(x)\ \mathrm{d}x := \lim_{\varepsilon \searrow 0} \int_{a+\varepsilon}^b f(x)\ \mathrm{d}x$$

Beispiel. $\int_0^1 \frac{1}{x^s}\ \mathrm{d}x$ konvergiert für $s < 1$ und divergiert, wenn $s \geq 1$ ist, denn: $\int_\varepsilon^1 \frac{1}{x^s}\ \mathrm{d}x = \frac{1}{(1-s)(x^{s-1})}\Big|_\varepsilon^1 = \frac{1}{1-s}\left(1-\varepsilon^{1-s}\right)$, falls $s \neq 1$. Für $s < 1$ ist $\lim_{\varepsilon \searrow 0} \varepsilon^{1-s} = 0$, also $\int_0^1 \frac{1}{x^s}\ \mathrm{d}x = \frac{1}{1-s}$. Im Fall $s > 1$ ergibt sich $\lim_{\varepsilon \searrow 0} \varepsilon^{1-s} = \infty$. $s = 1$ liefert $\int_\varepsilon^1 \frac{\mathrm{d}x}{x} = -\ln(\varepsilon)$ und $\lim_{\varepsilon \to 0} -\ln(\varepsilon) = \infty$.

Definition 5.10. $f: (a,b) \to \mathbb{R}$ mit $a \in \mathbb{R} \cup \{-\infty\}$, $b \in \mathbb{R} \cup \{\infty\}$ sei über jedem Intervall $[\alpha, \beta] \subset (a,b)$ integrierbar. Falls $\lim_{\alpha \searrow a} \int_\alpha^c f(x)\ \mathrm{d}x$ und $\lim_{\beta \nearrow b} \int_c^\beta f(x)\ \mathrm{d}x$ existieren, wobei $c \in (a,b)$ beliebig ist, so ist $\int_a^b f(x)\ \mathrm{d}x$ **konvergent** und hat den Wert

$$\int_a^b f(x)\ \mathrm{d}x := \lim_{\alpha \searrow a} \int_\alpha^c f(x)\ \mathrm{d}x + \lim_{\beta \nearrow b} \int_c^\beta f(x)\ \mathrm{d}x$$

Beispiel.

- $\int_{-1}^1 \frac{\mathrm{d}x}{\sqrt{1-x^2}} = \lim_{\varepsilon \searrow 0} \int_{\varepsilon-1}^0 \frac{\mathrm{d}x}{\sqrt{1-x^2}} + \lim_{\varepsilon \searrow 0} \int_0^{1-\varepsilon} \frac{\mathrm{d}x}{\sqrt{1-x^2}}$
  $= -\lim_{\varepsilon \searrow 0} \arcsin(\varepsilon - 1) + \lim_{\varepsilon \searrow 0} \arcsin(1-\varepsilon) = \pi$

- $\int_{-\infty}^\infty \frac{\mathrm{d}x}{1+x^2} = \lim_{R \to \infty} \int_{-R}^0 \frac{\mathrm{d}x}{1+x^2} + \lim_{R \to \infty} \int_0^R \frac{\mathrm{d}x}{1+x^2}$
  $= \lim_{R \to \infty} \arctan(-R) + \lim_{R \to \infty} \arctan(R) = \pi$

- $\int_0^\infty \frac{\mathrm{d}x}{x^s}$ divergiert $\forall\, s \in \mathbb{R}$, denn: für $s \neq 1$ ist
  $\int_\varepsilon^R \frac{\mathrm{d}x}{x^s} = \int_\varepsilon^1 \frac{\mathrm{d}x}{x^s} + \int_1^R \frac{\mathrm{d}x}{x^s} = \frac{1}{1-s}\left(1-\varepsilon^{1-s}\right) + \frac{1}{s-1}\left(1 - \frac{1}{R^{s-1}}\right)$
  Der erste Summand konvergiert nur für $s < 1$, in diesem Fall aber divergiert der zweite Summand. Letzterer konvergiert, falls $s > 1$ ist; dann jedoch divergiert der erste Summand.
  Für $s = 1$ divergiert sowohl $\int_\varepsilon^1 \frac{\mathrm{d}x}{x}$, als auch $\int_1^R \frac{\mathrm{d}x}{x}$.

Definition 5.11. Ist $f: [a,b]\backslash\{c\} \to \mathbb{R}$ mit $c \in (a,b)$ über jedem Intervall $[a, c-\varepsilon]$, $0 < \varepsilon < c-a$ und $[c+\varepsilon, b]$, $0 < \varepsilon < b-c$ integrierbar und es existiert $\lim_{\varepsilon \searrow 0} \int_a^{c-\varepsilon} f(x)\ \mathrm{d}x$, sowie $\lim_{\varepsilon \searrow 0} \int_{c+\varepsilon}^b f(x)\ \mathrm{d}x$, dann ist

$$\int_a^b f(x)\ \mathrm{d}x := \lim_{\varepsilon \searrow 0} \int_a^{c-\varepsilon} f(x)\ \mathrm{d}x + \lim_{\varepsilon \searrow 0} \int_{c+\varepsilon}^b f(x)\ \mathrm{d}x$$

In einigen Fällen kann durch Betrachtung uneigentlicher Integrale eine Aussage über die Konvergenz einer unendlichen Reihe gemacht werden:

Satz 5.19 (Integralkriterium). *Ist $f: [1,\infty) \to \mathbb{R}$ monoton fallend, so konvergiert $\sum_{n=1}^\infty f(n)$ dann und nur dann, wenn $\int_1^\infty f(x)\ \mathrm{d}x$ konvergiert.*

Beweis. Nach Voraussetzung fällt f monoton; damit folgt $f \geq 0$, wenn $\int_1^\infty f(x)\ \mathrm{d}x$ oder $\sum_{n=1}^\infty f(n)$ konvergiert.
Für $n-1 \leq x \leq n$ ist $f(n) \leq f(x) \leq f(n-1)$.
$\Rightarrow \quad f(n) \leq \int_{n-1}^n f(x)\ \mathrm{d}x \leq f(n-1)$
$\Rightarrow \quad \sum_{n=2}^N f(n) \leq \int_1^N f(x)\ \mathrm{d}x \leq \sum_{n=1}^{N-1} f(n)$, d.h.
falls $\int_1^\infty f(x)\ \mathrm{d}x$ konvergiert, dann ist $\sum_{n=1}^N f(n)$ beschränkt und wegen $f \geq 0$ monoton wachsend, also konvergent.
Ist andererseits $\sum_{n=1}^N f(n)$ konvergent, so ist $\int_1^N f(x)\ \mathrm{d}x$ beschränkt und monoton wachsend. □

Beispiel. $\sum_{n=1}^\infty \frac{1}{n^s}$ konvergiert für $s > 1$ und divergiert, falls $s < 1$ ist.

5.6 Weitere Anwendungen der Integralrechnung

In Abschnitt 4.3 wurde eine n mal differenzierbare Funktionen $f : I \to \mathbb{R}$ als Summe eines Taylorpolynoms
$T_n f(x,a) = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \ldots + \frac{f^{(n)}(a)}{n!}(x-a)^n\,, \quad a \in I$
und eines Restgliedes $R_{n+1}(x) = f(x) - T_n f(x,a)$ dargestellt.
Für n+1 mal stetig differenzierbare Funktionen ergab sich die Lagrangesche Form des Restgliedes: $R_{n+1}(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-a)^{n+1}$ mit ξ zwischen a und x.
Nun soll eine weitere Restgliedform vorgestellt werden.

Satz 5.20 (Integralform des Restgliedes). *Sei $f : I \to \mathbb{R}$ n+1 mal stetig differenzierbar auf I und $a \in I$. Dann gilt für alle $x \in I$:*

$$R_{n+1}(x) = f(x) - T_n f(x,a) = \frac{1}{n!}\int_a^x (x-t)^n f^{(n+1)}(t)\ \mathrm{d}t$$

Beweis. Der Beweis erfolgt durch vollständige Induktion über n.
Für $n = 0$ ist nach Satz 5.13 $R_1 = f(x) - f(a) = \int_a^x f(t)\ \mathrm{d}t$.
Die Behauptung sei wahr für $n-1$, d.h.
$f(x) - T_{n-1}f(x,a) = \frac{1}{(n-1)!}\int_a^x (x-t)^{n-1} f^{(n)}(t)\ \mathrm{d}t$.
Durch partielle Integation ergibt sich daraus
$f(x) - T_{n-1}f(x,a) = -\frac{(x-t)^n}{n!} f^{(n)}(t)\Big|_a^x + \frac{1}{n!}\int_a^x (x-t)^n f^{(n+1)}(t)\ \mathrm{d}t$
$= \frac{f^{(n)}(a)}{n!}(x-a)^n + \frac{1}{n!}\int_a^x (x-t)^n f^{(n+1)}(t)\ \mathrm{d}t$
$\Rightarrow \quad f(x) - T_n f(x,a) = \frac{1}{n!}\int_a^x (x-t)^n f^{(n+1)}(t)\ \mathrm{d}t$ □

Die Ungleichungen für Normen aus 4.4 lassen sich auf Integralnormen erweitern. Im Folgenden wird dies am Beispiel der Hölderschen Ungleichung dargestellt.

Definition 5.12. Sei $f : [a,b] \to \mathbb{R}$ Riemann-integrierbar und $p \geq 1$. Die **p-Norm** von f bezüglich $[a,b]$ ist

$$||f||_p = \left(\int_a^b |f(x)|^p \, \mathrm{d}x \right)^{\frac{1}{p}}$$

Die so definierte p-Norm erfüllt die Bedingungen von Definition 3.21.

Satz 5.21 (Höldersche Ungleichung). *Sind $f, g : [a,b] \to \mathbb{R}$ integrierbar und $p, q > 0$ mit $\frac{1}{p} + \frac{1}{q}$, so gilt*

$$||fg||_1 = \int_a^b |f(x)g(x)| \, \mathrm{d}x \leq ||f||_p \cdot ||g||_q$$

Beweis. Man betrachte eine äquidistante Zerlegung von $[a,b]$ in Teilintervalle der Länge $h_n = \frac{b-a}{n}$, $n \in \mathbb{N}$ und beliebige Stützstellen
$\xi_i \in [a + ih_n, a + (i+1)h_n], \;\; i = 0, \ldots, n-1$.
Für die dazugehörigen Riemann-Summen wurde die Höldersche Ungleichung schon gezeigt (Satz 4.22).

$$\begin{aligned}
\sum_{k=1}^{n} |h_n f(\xi_k) g(\xi_k)| &= \sum_{k=1}^{n} \left| \left(h_n^{\frac{1}{p}} f(\xi_k) \right) \left(h_n^{\frac{1}{q}} g(\xi_k) \right) \right| \\
&\leq \left(\sum_{k=1}^{n} \left| h_n^{\frac{1}{p}} f(\xi_k) \right|^p \right)^{\frac{1}{p}} \left(\sum_{k=1}^{n} \left| h_n^{\frac{1}{q}} g(\xi_k) \right|^q \right)^{\frac{1}{q}} \\
&= \left(\sum_{k=1}^{n} h_n \, |f(\xi_k)|^p \right)^{\frac{1}{p}} \left(\sum_{k=1}^{n} h_n \, |g(\xi_k)|^q \right)^{\frac{1}{q}}
\end{aligned}$$

Nach Satz 5.7 ist $\lim_{n\to\infty} \left(\sum_{k=1}^{n} h_n \, |f(\xi_k)|^p \right) = \int_a^b |f(x)|^p \, \mathrm{d}x$ und $\lim_{n\to\infty} \left(\sum_{k=1}^{n} h_n \, |g(\xi_k)|^q \right) = \int_a^b |g(x)|^q \, \mathrm{d}x$. Damit folgt die Behauptung. $\square$

5.7 Integation und Grenzübergang

Satz 5.22. *Sei I ein endliches Intervall und $(f_n)_{n\in\mathbb{N}}$ eine Folge stetiger Funktionen $f_n : I \to \mathbb{R}$, die gleichmäßig gegen eine Funktion f konvergiert. Weiterhin sei F_n eine Stammfunktion von f_n und für ein $x_0 \in I$ existiere* $\lim_{n\to\infty} F_n(x_0)$. *Dann konvergiert $(F_n)_{n\in\mathbb{N}}$ gleichmäßig gegen eine Stammfunktion F von f. Interpretation: $\forall \, a, b \in I$:*
$\lim_{n\to\infty} \int_a^b f_n(x) \, \mathrm{d}x = F|_a^b = \int_a^b f(x)\mathrm{d}x = \int_a^b \lim_{n\to\infty} f_n(x) \, \mathrm{d}x$

Beweis. L möge die Länge des Intervalls I bezeichnen.
Man setze $F(x) := \lim_{n\to\infty} F_n(x_0) + \int_{x_0}^{x} f(t) \, \mathrm{d}t$. Dann ist F eine Stammfunktion von f und es bleibt noch $F_n \rightrightarrows F$ zu zeigen.

Mit dem Hauptsatz der Differential- und Integralrechnung ist
$|F_n(x) - F(x)| = |F_n(x_0) + \int_{x_0}^{x} f_n(t)\ \mathrm{d}t - F(x_0) - \int_{x_0}^{x} f(t)\ \mathrm{d}t|$
$\leq |F_n(x_0) - F(x_0)| + |\int_{x_0}^{x} (f_n(t) - f(t))\ \mathrm{d}t| \leq |F_n(x_0) - F(x_0)| + L||f_n - f||_I$
Da $(f_n)_n$ gleichmäßig gegen f konvergiert, existiert zu jedem $\varepsilon > 0$ ein
$N \in \mathbb{N}: \ ||f_n - f||_I < \frac{\varepsilon}{2L} \ \forall n \geq N$.
Außerdem gibt es nach Voraussetzung ein
$N' \in \mathbb{N}: \ |F_n(x_0) - F(x_0)| < \frac{\varepsilon}{2} \ \forall n \geq N'$.
$\Rightarrow \quad \forall n \geq \max(N, N'): \ \forall x \in I: \ |F_n(x) - F(x)| < \varepsilon.$ □

Bemerkung. Die Konvergenz von (F_n) in einem Punkt x_0 läßt sich immer erreichen:
Ist G_n eine Stammfunktion von f_n, dann auch $F_n(x) = G_n(x) - G_n(x_0)$ und für F_n gilt $F_n(x_0) = 0 \ \forall n$, d.h. $\lim_{n\to\infty} F_n$ existiert.

Bemerkung. Bei Reihen vertauschen Integration und Grenzwertbildung in analoger Weise.

Eine mögliche Anwendung von Satz 5.22 ist die gliedweisen Integration gleichmäßig konvergenter Reihen.

Beispiel. $e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$. Im Beweis von Satz 3.20 wurde gezeigt, daß Potenzreihen innerhalb des Konvergenzradius gleichmäßig konvergieren.
$\Rightarrow \quad \int e^x\ \mathrm{d}x = \sum_{n=0}^{\infty} \int \frac{x^n}{n!}\ \mathrm{d}x = \sum_{n=1}^{\infty} \frac{x^n}{n!} = e^x - 1$

Satz 5.23. *Ist $(f_n)_{n\in\mathbb{N}}$ eine Folge* RIEMANN*-integrierbarer Funktionen, welche gleichmäßig gegen eine Funktion f konvergiert, so ist die Folge $\left(\int_a^b f_n(x)\mathrm{d}x\right)_n$ konvergent und*
$\lim_{n\to\infty} \int_a^b f_n(x)\mathrm{d}x = \int_a^b \lim_{n\to\infty} f_n(x)\mathrm{d}x = \int_a^b f(x)\mathrm{d}x.$

Beispiel. Für $|t| < 1$ betrachte man $(1 - t^2)^{-\frac{1}{2}} = \sum_{n=0}^{\infty} \binom{-\frac{1}{2}}{n}(-1)^n t^{2n}$.
Diese Reihe ist in jedem kompakten Teilintervall von $[-1, 1]$ gleichmäßig konvergent. Für $|x| < 1$ gilt somit: $\int_0^x \frac{\mathrm{d}t}{\sqrt{1-t^2}} = \sum_{n=0}^{\infty} \binom{-\frac{1}{2}}{n}(-1)^n \int_0^x t^{2n}$
Mit $\binom{-\frac{1}{2}}{n} = (-1)^n \frac{1\cdot 3\cdot 5 \ldots (2n-1)}{2^n n!}$ ergibt sich daraus eine Darstellung von $\arcsin x$ als Reihe: $\int_0^x \frac{\mathrm{d}t}{\sqrt{1-t^2}} = \arcsin x = x + \sum_{n=1}^{\infty} \frac{1\cdot 3\cdot 5 \ldots (2n-1)}{2^n n!(2n+1)} x^{2n+1}$

Bemerkung (Satz von ARZELA-OSGOOD[31]*).* Konvergiert $(f_n)_{n\in\mathbb{N}}$ punktweise gegen eine Funktion f, so ist $\left(\int_a^b f_n(x)\mathrm{d}x\right)_n$ konvergent, wenn $(||f_n||)_n$ beschränkt ist. In diesem Fall gilt $\lim_{n\to\infty} \int_a^b f_n(x)\mathrm{d}x = \int_a^b f(x)\mathrm{d}x$.

---

[31] *Cesare Arzela* (1847-1912), *William Osgood* (1864-1943)

6 Fourier-Reihen

Analytische Funktionen sind dadurch charakterisiert, daß sie sich mit Hilfe von Potenzreihen darstellen lassen. Eine weitere Möglichkeit der Darstellung von Funktionen als Reihe bieten die trigonometrischen Reihen.

6.1 Trigonometrische Polynome und Fourier-Reihen

Definition 6.1.

$$T(x) = \sum_{k=-n}^{n} c_k e^{\mathbf{i}kx}$$

mit $c_k \in \mathbb{C}$ und $x \in \mathbb{R}$ heißt **trigonometrisches Polynom** vom Grad n.

Bemerkung. Summen und Produkte trigonometrischer Polynome sind wieder trigonometrische Polynome.

Aus der Beziehung $e^{\mathbf{i}kx} = \cos kx + \mathbf{i} \sin kx$ ergibt sich eine alternative Darstellung trigonometrischer Polynome:

$$T(x) = \frac{a_0}{2} + \sum_{k=1}^{n} a_k \cos kx + b_k \sin kx$$

mit $a_0 = 2c_0\,,\ a_k = c_k + c_{-k}\,,\ b_k = \mathbf{i}(c_k - c_{-k})$
bzw. $c_k = \frac{1}{2}(a_k - \mathbf{i}b_k)\,,\ c_{-k} = \frac{1}{2}(a_k + \mathbf{i}b_k)$.

Es besteht die Orthogonalitätsrelation

$$\frac{1}{2\pi} \int_0^{2\pi} e^{\mathbf{i}\mu x} e^{-\mathbf{i}\nu x} \mathrm{d}x = \delta_{\mu\nu}$$

Damit ist $c_k = \frac{1}{2\pi} \int_0^{2\pi} T(x) e^{-\mathbf{i}kx} \mathrm{d}x$ und die Koeffizienten c_k sind für gegebenes $T(x)$ eindeutig bestimmt.
$T(x)$ ist reell, wenn $\overline{c_k} = c_{-k}$ ist, d.h. alle a_k und b_k reell sind.

Definition 6.2.

$$\sum_{k=-\infty}^{\infty} c_k e^{\mathbf{i}kx}$$

mit $c_k \in \mathbb{C}$ und $x \in \mathbb{R}$ heißt **trigonometrische Reihe**. Sie konvergiert, wenn $\lim_{n\to\infty} \sum_{k=-n}^{n} c_k e^{\mathbf{i}kx}$ existiert.

Hat die Funktion f eine Darstellung als trigonometrische Reihe $\sum_{k=-\infty}^{\infty} c_k e^{\mathbf{i}kx}$ und kann diese gliedweise integriert werden, was nach Satz 5.23 z.B. bei gleichmäßiger Konvergenz der Fall ist, so gilt

$$c_k = \frac{1}{2\pi} \int_0^{2\pi} f(x) e^{-\mathbf{i}kx} \mathrm{d}x\,,$$

denn wegen der Orthogonalitätsrelation ist
$\int_0^{2\pi} f(x)e^{-\mathbf{i}kx}\mathrm{d}x = \sum_{n=-\infty}^{\infty} \int_0^{2\pi} c_n e^{\mathbf{i}n} e^{-\mathbf{i}k} = 2\pi c_k$
Es folgt $\quad a_k = \frac{1}{\pi}\int_0^{2\pi} f(x)\cos kx \ \mathrm{d}x \quad$ und $\quad b_k = \frac{1}{\pi}\int_0^{2\pi} f(x)\sin kx \ \mathrm{d}x$.

Bemerkung. Bei der Bestimmung der Koeffizienten kann über ein beliebiges Intervall der Länge 2π integriert werden, da die Integranden 2π-periodisch sind. Ist f eine gerade Funktion, so ist $b_k = 0 \ \forall k$; für ungerade Funktionen sind alle $a_k = 0$.

Somit sind die Koeffizienten einer trigonometrischen Reihe eindeutig bestimmt, wenn diese punktweise konvergiert und gliedweise integriert werden darf.

Satz 6.1. *Konvergiert* $\sum_{n=-\infty}^{\infty} c_n$ *absolut, dann ist die Funktion* $f(x) = \sum_{-\infty}^{\infty} c_n e^{\mathbf{i}nx}$ *stetig und* 2π*-periodisch. Die Darstellung von* f *als* $\sum_{n=-\infty}^{\infty} c_n e^{\mathbf{i}nx}$ *ist eindeutig.*

Beweis. Für jedes nichtleere, abgeschlossene Intervall $I \subset \mathbb{R}$ ist $||c_n e^{\mathbf{i}nx}||_I = |c_n|$. Mit dem Weierstraßschen Konvergenzkriterium folgt, daß $\sum_{n=-\infty}^{\infty} c_n e^{\mathbf{i}nx}$ gleichmäßig konvergiert. Wegen Satz 3.16 ist die Grenzfunktion f damit stetig.
Überdies ist die Reihe gliedweise integrierbar, d.h. die Koeffizienten c_n sind nach obigen Betrachtungen eindeutig bestimmt.
f ist 2π-periodisch, weil jedes $e^{\mathbf{i}nx}$ dies ist. $\square$

Beispiel. Für $s > 1$ sind $f(x) = \sum_{n=1}^{\infty} \frac{\cos nx}{n^s}$ und $g(x) = \sum_{n=1}^{\infty} \frac{\cos nx}{n^s}$ stetige Funktionen mit der Periode 2π.

Es stellt sich nun nicht mehr allein die Frage, ob die Reihe $\sum_{k=-\infty}^{\infty} c_k e^{\mathbf{i}kx}$ konvergiert, sondern ob sie auch gegen die gegebene Funktion f konvergiert, wenn sich die Koeffizienten c_k aus $c_k = \frac{1}{2\pi}\int_0^{2\pi} f(x)e^{-\mathbf{i}kx}\mathrm{d}x$ ergeben, d.h. inwiefern f durch eine trigonometrische Reihe dargestellt werden kann.

Definition 6.3. Ist die Funktion f 2π-periodisch und es existiert $\int_0^{2\pi} f(x)e^{-\mathbf{i}kx}\mathrm{d}x$, so sind die Zahlen $c_k = \frac{1}{2\pi}\int_0^{2\pi} f(x)e^{-\mathbf{i}kx}\mathrm{d}x$ die **Fourier-Koeffizienten**[32] von f.

Definition 6.4. Sind c_k die Fourier-Koeffizienten der 2π-periodischen Funktion f, dann heißt $\sum_{k=-\infty}^{\infty} c_k e^{\mathbf{i}kx}$ die zu f gehörige **Fourier-Reihe**.

Beispiel. Sei $f: [0, 2\pi] \to \mathbb{R}$ mit $f(x) = \begin{cases} -x + \frac{\pi}{2} & \text{für } x \leq \pi \\ x - \frac{3}{2}\pi & \text{für } x > \pi \end{cases}$
und $g(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos nx + \mathbf{i}\sin nx$ die zu f gehörige Fourier-Reihe.
f ist gerade, daher sind die Koeffizienten $b_n = 0$.

[32] nach *Joseph Fourier* (1768-1830)

$a_n = \frac{1}{\pi}\int_0^{\pi}\left(-x+\frac{\pi}{2}\right)\cos nx\ \mathrm{d}x + \frac{1}{\pi}\int_{\pi}^{2\pi}\left(x-\frac{3}{2}\pi\right)\cos nx\ \mathrm{d}x$

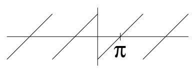
$\Rightarrow \quad a_0 = 0$ und für $n \geq 1$ ergibt sich mittels partieller Integration

$a_n = \frac{2}{\pi n^2}(1-(-1)^n)$, d.h. $\quad g(x) = \frac{4}{\pi}\left(\cos x + \frac{\cos 3x}{3^2} + \frac{\cos 5x}{5^2} + \ldots\right)$.

g ist stetig, da $\sum_{n=1}^{\infty}\frac{1-(-1)^n}{n^2}$ absolut konvergiert.

Beispiel. Man betrachte die an den Stellen $(2k+1)\cdot\pi,\ k \in \mathbb{Z}$ unstetige

Funktion $f(x) = \begin{cases} x - 2\pi\lfloor\frac{x}{2\pi}\rfloor - \pi & \text{für } x \neq (2k+1)\pi,\ k \in \mathbb{Z} \\ 0 & \text{für } x = (2k+1)\pi \end{cases}$



Diese Funktion ist ungerade, daher hat die zugehörige FOURIER-Reihe die Form $\sum_{n=1}^{\infty} b_n \sin nx$. $b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} x \sin nx\ \mathrm{d}x = \frac{2(-1)^{n+1}}{n}$;

damit ist $2\left(\sin x - \frac{\sin 2x}{2} + \frac{\sin 3x}{3} - \ldots\right)$ die FOURIER-Reihe von f.

Sie konvergiert punktweise gegen die gegebene Funktion.

Beispiel. Sei $f(x) = \ln\left(2\cos\left(\frac{x}{2}\right)\right)$. f ist gerade und die uneigentlichen Integrale $a_n = \frac{1}{\pi}\int_{-\pi}^{\pi}\ln\left(2\cos\left(\frac{x}{2}\right)\right)\cos nx\ \mathrm{d}x$ sind absolut konvergent:

$a_n = \frac{(-1)^{n+1}}{n}$. Die zu f gehörige FOURIER-Reihe ist somit

$\cos x - \frac{cos2x}{2} + \frac{cos3x}{3} - \ldots.$ Sie divergiert für $x = (2k+1)\pi\ k \in \mathbb{Z}$.

Die Frage nach der Darstellbarkeit einer Funktion durch ihre FOURIER-Reihe ist auch Gegenstand des folgenden Abschnitts.

6.2 Der Satz von Fejér

Definition 6.5. Eine Funktion $f : [a,b] \to \mathbb{R}$ heißt **Regelfunktion**, wenn es eine Folge von Treppenfunktionen $g_n \in T(a,b)$ gibt, die gleichmäßig gegen f konvergiert.

Bemerkung. $f : [a,b] \to \mathbb{R}$ ist genau dann eine Regelfunktion, wenn sie in a einen rechtsseitigen, in b einen linksseitigen und in jedem $x \in (a,b)$ einen rechts- und linksseitigen Grenzwert hat.

Satz 6.2 (Satz von Fejér). [33] *Sei f eine 2π-periodische Regelfunktion und für alle x_0 gelte* $f(x_0) = \frac{1}{2}(\lim_{x\nearrow x_0} f(x) + \lim_{x\searrow x_0} f(x))$. *Weiter seien* $c_k = \frac{1}{2\pi}\int_0^{2\pi} f(x)e^{-\mathbf{i}kx}\mathrm{d}x$, $s_m(x) = \sum_{k=-m}^{m} c_k e^{\mathbf{i}kx}$ *und*

$\sigma_n(x) = \frac{1}{n}(s_0(x) + s_1(x) + \ldots + s_{n-1}(x))$.

Dann konvergiert σ_n punktweise gegen f und $(||\sigma_n||)_{n\in\mathbb{N}}$ ist beschränkt.

Beweis. Zunächst soll eine Integraldarstellung für $\sigma_n(x)$ hergeleitet werden.

$s_m(x) = \sum_{k=-m}^{m}\left(e^{\mathbf{i}kx}\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-\mathbf{i}kt}f(t)\ \mathrm{d}t\right) = \frac{1}{2\pi}\int_{-\pi}^{\pi}\left(\sum_{k=-m}^{m} e^{\mathbf{i}k(x-t)}f(t)\right)\ \mathrm{d}t$

[33] *Lipót Fejér* (1880-1959)

Die Substitution $x - t = y$ führt zu $s_m(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \left(\sum_{k=-m}^{m} e^{\mathbf{i}ky} f(x+y)\right)\, \mathrm{d}y$

Die Summe im Integranden läßt sich zusammenfassen:

$\sum_{k=-m}^{m} e^{\mathbf{i}ky} = \sum_{k=0}^{m} e^{\mathbf{i}ky} + \sum_{k=0}^{m} e^{\mathbf{i}k(-y)} - 1 = \frac{1-e^{\mathbf{i}y(m+1)}}{1-e^{\mathbf{i}y}} + \frac{1-e^{-\mathbf{i}y(m+1)}}{1-e^{-\mathbf{i}y}} - 1$
$= \frac{\left(e^{\mathbf{i}my}+e^{-\mathbf{i}my}\right)-\left(e^{\mathbf{i}(m+1)y}+e^{-\mathbf{i}(m+1)y}\right)-\left(e^{\mathbf{i}y}+e^{-\mathbf{i}y}\right)+2}{2-\left(e^{\mathbf{i}y}+e^{-\mathbf{i}y}\right)} - 1 = \frac{\cos my - \cos{(m+1)y}}{1-\cos y}$

Damit wird $s_m(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{\cos my - \cos{(m+1)y}}{1-\cos y} f(x+y)\, \mathrm{d}y$

Der Integrand ist in 0 nicht definiert; das Integral jedoch konvergiert.

Es folgt $\sigma_n(x) = \frac{1}{2\pi n}\int_{-\pi}^{\pi} \frac{1-\cos ny}{1-\cos y} f(x+y)\, \mathrm{d}y$

und mit $\cos 2\alpha = 1 - 2\sin^2\alpha$ ist

$$\sigma_n(x) = \frac{1}{2\pi n}\int_{-\pi}^{\pi} \left(\frac{\sin\frac{ny}{2}}{\sin\frac{y}{2}}\right)^2 f(x+y)\, \mathrm{d}y$$

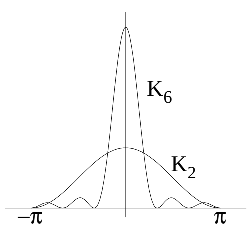
Definition 6.6.

$$K_n(y) = \frac{1}{2\pi n}\left(\frac{\sin\frac{ny}{2}}{\sin\frac{y}{2}}\right)^2$$

heißt n-ter **Fejérscher Kern**.

Die Fejérschen Kerne haben einige nützliche Eigenschaften:

- $K_n \geq 0$
- $K_n(y) = K_n(-y)$, d.h. die Kerne sind gerade Funktionen.
- $\int_{-\pi}^{\pi} K_n(y)\, \mathrm{d}y = 1$
- Für alle $\delta > 0$ konvergiert $(K_n|_{[-\pi,-\delta]\cup[\delta,\pi]})_{n\in\mathbb{N}}$ gleichmäßig gegen 0.



Mit dieser Definition schreibt sich $\sigma_n(x)$ als $\int_{-\pi}^{\pi} K_n(y) f(x-y)\, \mathrm{d}y$,

d.h. σ_n ist das Faltungsprodukt von f und $\widetilde{K_n}(x) := \begin{cases} K_n(x) & \text{für } x \in [-\pi,\pi] \\ 0 & \text{sonst} \end{cases}$

Mit der Voraussetzung $f(x) = \frac{1}{2}\left(f(x_+) + f(x_-)\right)$, wobei $f(x_+) = \lim_{\tilde{x}\searrow x} f(x)$ und $f(x_-) = \lim_{\tilde{x}\nearrow x} f(x)$ ist, folgt

$f(x) = \frac{1}{2}\int_{-\pi}^{\pi} K_n(y)(f(x_+) + f(x_-))\, \mathrm{d}y$

Außerdem gilt wegen $K_n(y) = K_n(-y)$:

$\sigma_n(x) = \frac{1}{2}\int_{-\pi}^{\pi} K_n(y)(f(x+y) + f(x-y))\, \mathrm{d}y$.

Hiernach ist $\sigma_n(x) - f(x) = \int_{-\pi}^{\pi} K_n g(y)\ \mathrm{d}y$ mit

$g(y) = \frac{1}{2}(f(x+y) + f(x-y) - f(x_+) - f(x_-))$.

Nun wird die punktweise Konvergenz von σ_n gezeigt.

Die Funktion g ist stetig in 0 und verschwindet dort, daher existiert $\forall\ \varepsilon > 0$ ein (von x abhängiges) $\delta > 0$, sodaß $\forall\, y : \ |y| < \delta \quad \Rightarrow \quad |g(y)| < \varepsilon$.

Zusammen mit $K_n \geq 0$ liefert dies die Abschätzung

$\left|\int_{-\delta}^{\delta} K_n(y) g(y)\ \mathrm{d}y\right| \leq \int_{-\delta}^{\delta} K_n(y)|g(y)|\ \mathrm{d}y$

$\leq \varepsilon \int_{-\delta}^{\delta} K_n(y)\ \mathrm{d}y < \varepsilon \int_{-\pi}^{\pi} K_n(y)\ \mathrm{d}y = \varepsilon$

Für die beiden anderen Teile des Integrals für σ_n nutze man

$K_n(y) \leq \frac{1}{2\pi n} \cdot \frac{1}{sin^2 \frac{y}{2}}$, was direkt aus der Definition der Fejérschen Kerne folgt und $|g(y)| \leq 2||f||$: $\ |\int_{-\pi}^{\delta} K_n g(y)\ \mathrm{d}y| \leq \frac{1}{n} \frac{||f||}{\sin^2 \frac{\delta}{2}}$, $|\int_{\delta}^{\pi} K_n g(y)\ \mathrm{d}y| \leq \frac{1}{n} \frac{||f||}{\sin^2 \frac{\delta}{2}}$.

$\Rightarrow \quad |\sigma_n(x) - f(x)| \leq \varepsilon + \frac{1}{n} \cdot \frac{2||f||}{\sin^2 \frac{\delta}{2}}$

Somit läßt sich für alle x und $\widetilde{\varepsilon} > 0$ ein $N \in \mathbb{N}$ mit

$|\sigma_n(x) - f(x)| < \widetilde{\varepsilon} \quad \forall\, n \geq N$ finden, d.h. (σ_n) konvergiert punktweise gegen f.

Mithin ist $(||\sigma_n||)_{n \in \mathbb{N}}$ beschränkt:

$||\sigma_n|| = ||\int_{-\pi}^{\pi} K_n(y) f(x-y)\ \mathrm{d}y|| \leq ||f|| \int_{-\pi}^{\pi} K_n(y)\ \mathrm{d}y = ||f||$ □

Bemerkung. Ist die Funktion f stetig und 2π-periodisch, so ist sie auch gleichmäßig stetig. Damit läßt sich δ in vorangegangenem Beweis unabhängig von x wählen, womit σ_n gleichmäßig gegen f konvergiert. Demzufolge kann jede stetige, periodische Funktion gleichmäßig durch trigonometrische Polynome approximiert werden.

6.3 Konvergenz in der Hilbert-Norm

Bemerkung. Die komplexwertigen Regelfunktionen der Periode 2π bilden einen Vektorraum über $\mathbb{C}$.

Definition 6.7. Seien $f, g : \mathbb{R} \to \mathbb{C}$ Regelfunktionen. Dann heißt

$$\langle f, g \rangle = \int_0^{2\pi} f(x)\overline{g(x)}\ \mathrm{d}x$$

Skalarprodukt von f und g.

Bemerkung. Für $\langle \cdot, \cdot \rangle$ gilt

- $\langle f, g \rangle = \overline{\langle g, f \rangle}$. Sind f und g reell, so ist $\langle f, g \rangle$ reell und $\langle f, g \rangle = \langle g, f \rangle$
- $\langle \cdot, \cdot \rangle$ ist linear in der ersten Kompoente, d.h.
  $\langle f_1 + f_2, g \rangle = \langle f_1, g \rangle + \langle f_2, g \rangle \quad$ und $\quad \langle \lambda f, g \rangle = \lambda \langle f, g \rangle \quad$ mit $\lambda \in \mathbb{C}$

und konjugiert linear in der zweiten Komponente, d.h.
$\langle f, g_1 + g_2 \rangle = \langle f, g_1 \rangle + \langle f, g_2 \rangle$ und $\langle f, \lambda g \rangle = \overline{\lambda} \langle f, g \rangle$ mit $\lambda \in \mathbb{C}$

- Für stetige f ist $\langle \cdot, \cdot \rangle$ positiv definit, d.h. $\langle f, f \rangle \geq 0$ und
$\langle f, f \rangle = 0 \quad \Leftrightarrow \quad f = 0$

- $\langle f, g \rangle \overline{\langle f, g \rangle} \leq \langle f, f \rangle \langle g, g \rangle$ (CAUCHY-SCHWARZsche Ungleichung)

$\langle \cdot, \cdot \rangle$ induziert eine Norm auf dem Vektorraum der komplexwertigen 2π-periodischen Regelfunktionen.

Definition 6.8.

$$||f||_2 := \sqrt{\langle f, g \rangle} = \sqrt{\int_0^{2\pi} |f(x)|^2 \, \mathrm{d}x}$$

heißt **Hilbert-Norm** von f.
f und g heißen **orthogonal** zueinander, wenn $\langle f, g \rangle = 0$ ist.

Definition 6.9. Sei A eine beliebige Indexmenge. $(f_\alpha)_{\alpha \in A}$ heißt **Orthonormalsystem**, falls $\langle f_\alpha, f_\beta \rangle = \delta_{\alpha\beta}$ gilt.

Beispiel. $\left(\frac{1}{\sqrt{2\pi}} \, e^{\mathbf{i}nx} \right)_{n \in \mathbb{Z}}$ ist ein Orthonormalsystem.
Bemerkung. Jedem f sind bezüglich eines Orthonormalsystems $(f_\alpha)_{\alpha \in A}$ eindeutig Zahlen $c_\alpha = \langle f, f_\alpha \rangle$ zugeordnet. Sie heißen FOURIER-Koeffizienten von f bzgl. $(f_\alpha)_{\alpha \in A}$.

Satz 6.3. *Sei $(f_\alpha)_{\alpha \in A}$ ein Orthonormalsystem und E eine endliche Teilmenge von A. Dann wird die* HILBERT*-Norm $||f - \sum_{\alpha \in E} u_\alpha f_\alpha||_2$ mit $u_\alpha \in \mathbb{C}$ minimal, wenn u_α die* FOURIER*-Koeffizienten, also $u_\alpha = c_\alpha = \langle f, f_\alpha \rangle$ sind.*

Beweis. $||f - \sum_{\alpha \in E} u_\alpha f_\alpha||_2^2 = \langle f - \sum_{\alpha \in E} u_\alpha f_\alpha \ , \ f - \sum_{\alpha \in E} u_\alpha f_\alpha \rangle$
$= \langle f, f \rangle - \sum_{\alpha \in E} u_\alpha \langle f_\alpha, f \rangle - \sum_{\alpha \in E} \overline{u_\alpha} \langle f, f_\alpha \rangle + \langle \sum_{\alpha \in E} u_\alpha f_\alpha \ , \ \sum_{\alpha \in E} u_\alpha f_\alpha \rangle$
Die f_α bilden ein Orthonormalsystem; folglich ist
$\langle \sum_{\alpha \in E} u_\alpha f_\alpha \ , \ \sum_{\alpha \in E} u_\alpha f_\alpha \rangle = \sum_{\alpha \in E} u_\alpha \overline{u_\alpha}$

$$\Rightarrow \quad ||f - \sum_{\alpha \in E} u_\alpha f_\alpha||_2^2 = ||f||_2^2 - \sum_{\alpha \in E} |c_\alpha|^2 + \sum_{\alpha \in E} |c_\alpha - u_\alpha|^2 \qquad (*)$$

Dieser Ausdruck wird minimal, wenn der letzte Term verschwindet. □

Bemerkung. Aus $(*)$ ergibt sich die Ungleichung $0 \leq ||f||_2^2 - \sum_{\alpha \in E} |c_\alpha|^2$
$= ||f - \sum_{\alpha \in E} c_\alpha f_\alpha||_2^2 \leq ||f - \sum_{\alpha \in E} u_\alpha f_\alpha||_2^2$ mit c_α, u_α wie oben.

Es folgt weiter

Satz 6.4 (Besselsche Ungleichung[34]). *Für jede Teilmenge $B \subset A$ ist $\sum_{\alpha\in B} |c_\alpha|^2$ konvergent und $\sum_{\alpha\in B} |c_\alpha|^2 \leq ||f||_2^2$.*

Definition 6.10. Eine Funktionenfolge $(g_n)_{n\in\mathbb{N}}$ **konvergiert in der Hilbert-Norm** gegen f, wenn $\lim_{n\to\infty} ||f - g_n||_2 = 0$.

Es stellt sich nun die Frage, ob für ein gegebenes Orthonormalsystem $(f_\alpha)_{\alpha\in A}$ die Summe $\sum_{\alpha\in A} \langle f, f_\alpha\rangle f_\alpha$ gegen f konvergiert.
Wegen $||f - \sum_{\alpha\in E} c_\alpha f_\alpha||_2^2 = ||f||_2^2 - \sum_{\alpha\in E} |c_\alpha|^2$ aus $(*)$ ist dies äquivalent zu $||f||_2^2 = \sum_{\alpha\in A} |c_\alpha|^2$.

Definition 6.11. Ein Orthonormalsystem $(f_\alpha)_{\alpha\in A}$ heißt **vollständig**, falls für alle f gilt: $||f||_2^2 = \sum_{\alpha\in A} |\langle f, f_\alpha\rangle|^2$.

Bemerkung. Ist $(f_\alpha)_{\alpha\in A}$ ein vollständiges Orthonormalsystem, so gilt für alle f, g die PARSEVALsche Gleichung[35]

$$\langle f, g\rangle = \sum_{\alpha\in A} \langle f, f_\alpha\rangle \overline{\langle g, f_\alpha\rangle}$$

Beispiel. $\left(\frac{1}{\sqrt{2\pi}}\, e^{\mathbf{i}nx}\right)_{n\in\mathbb{Z}}$ ist ein vollständiges Orthonormalsystem.
Dies soll als Satz formuliert werden:

Satz 6.5. *Ist f eine Regelfunktion der Periode 2π, dann konvergiert die* FOURIER-*Reihe von f in der* HILBERT-*Norm gegen f, d.h.*
$||f - \sum_{k=-\infty}^{\infty} c_k e^{\mathbf{i}kx}||_2 = 0$ *mit* $c_k = \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} f(t) e^{-\mathbf{i}kt}\, \mathrm{d}t$.
Das Orthonormalsystem $\left(\frac{1}{\sqrt{2\pi}}\, e^{\mathbf{i}nx}\right)_{n\in\mathbb{Z}}$ *ist vollständig, d.h.*
$\int_0^{2\pi} |f(x)|^2\, \mathrm{d}x = \sum_{k=-\infty}^{\infty} |c_k|^2$

Beweis. Sei $s_n(x) = \sum_{k=-n}^{n} c_k e^{\mathbf{i}kx}$ und $\sigma_{n+1}(x) = \frac{1}{n+1}(s_0(x) + \ldots + s_n(x))$.
Nach Satz 6.2 konvergiert $(\sigma_n)_n$ punktweise gegen f und $(||\sigma_n||)_n$ ist beschränkt, woraus mit dem Satz von ARZELA-OSGOOD
$\lim_{n\to\infty} ||f - \sigma_{n+1}||_2 = \int_{-\pi}^{\pi} \lim_{n\to\infty} |f(x) - \sigma_{n+1}(x)|^2\, \mathrm{d}x = 0$ folgt.
Nach Satz 6.3 gilt außerdem $0 \leq ||f - s_n||_2^2 \leq ||f - \sigma_{n+1}||_2^2$; somit ist auch $\lim_{n\to\infty} ||f - \sigma_n||_2 = 0$. □

6.4 Ausblick: Punktweise Konvergenz

Satz 6.6. *Sei $(a_n)_{n\in\mathbb{N}}$ eine Folge komplexer Zahlen, $s_n = a_1 + a_2 + \ldots + a_n$ und $\sigma_n = \frac{1}{n}(s_1 + s_2 + \ldots + s_n)$.*

[34] nach *Wilhelm Bessel* (1784-1846)

[35] nach *Marc-Antoine Parseval* (1755-1836)

Konvergiert σ_n gegen eine Zahl s und $|a_n| \leq \frac{s}{n} \;\; \forall\, n$, so konvergiert auch $(s_n)_{n\in\mathbb{N}}$ gegen s.

Mit dem Satz von FEJÉR folgt daraus

Satz 6.7. *Ist f eine Regelfunktion der Periode 2π, c_k die FOURIER-Koeffizienten von f und es existiert ein $c \geq 0$, mit $|c_k| \leq \frac{c}{|k|} \;\; \forall\, k \in \mathbb{Z}\backslash\{0\}$, dann konvergiert die FOURIER-Reihe $\sum_{k=-\infty}^{\infty} c_k e^{\mathbf{i}kx}$ punktweise gegen f.*

Dies ist z.B. der Fall, wenn f eine stetig differenzierbare Funktion ist, denn dann gilt für $k \neq 0$: $\int_{-\pi}^{\pi} f(t)e^{-\mathbf{i}kt}\,\mathrm{d}t = f(t)\frac{\mathbf{i}}{k}e^{-\mathbf{i}kt}\Big|_{-\pi}^{\pi} - \frac{\mathbf{i}}{k}\int_{-\pi}^{\pi} f'(t)e^{-\mathbf{i}kt}\,\mathrm{d}t$ $= -\frac{\mathbf{i}}{k}\int_{-\pi}^{\pi} f'(t)e^{-\mathbf{i}kt}\,\mathrm{d}t$, also $|c_k| \leq \frac{||f'||}{|k|}$
Auch 2π-periodische Treppenfunktionen erfüllen die Voraussetzung von Satz 6.7. Ist $-\pi = x_0 < x_1 < \ldots < x_r = \pi$ eine Zerlegung von $[-\pi, \pi]$ und $g \in T(-\pi, \pi)$ mit $g|_{(x_{i-1},x_i)} = \gamma_i, \;\; \gamma_i \in \mathbb{C}, \;\; i = 1, 2, \ldots n$, dann sind die FOURIER- Koeffizienten von g
$c_k = \frac{1}{2\pi}\sum_{j=1}^{r}\int_{x_{j-1}}^{x_j} \gamma_j e^{-\mathbf{i}kt}\,\mathrm{d}t = \frac{1}{2\pi}\frac{\mathbf{i}}{k}\sum_{j=1}^{r}\gamma_j(e^{-\mathbf{i}kx_j} - e^{-\mathbf{i}kx_{j-1}})$
$\Rightarrow \quad |c_k| \leq \frac{1}{|k|}\left|\frac{\mathbf{i}}{2\pi}\sum_{j=1}^{r}\gamma_j(e^{-\mathbf{i}kx_j} - e^{-\mathbf{i}kx_{j-1}})\right|$

Satz 6.8. *Jede stückweise stetig differenzierbare 2π-periodische Funktion wird durch ihre FOURIER-Reihe dargestellt.*