Analysis I

Vorlesung im Wintersemester2002/03.

Manfred Lehn

Mainz, 11.4.2003

Inhaltsverzeichnis

Die mit Buchstaben statt Zahlen indizierten Kapitel sind Vorlesungsstoff, den auszulassen die Zeitnot mich zwang. Da er nun einmal vorbereitet war, habe ich ihn für interessierte Hörer im Skript gelassen. Diese Teile sind nicht klausur- oder prüfungsrelevant, aber trotzdem sehr interessant.

§1 Die natürlichen Zahlen

Ein systematischer Aufbau der Mathematik würde mit einer Diskussion der Mengenlehre beginnen und aus der leeren Menge $\emptyset$ der Reihe nach die natürlichen Zahlen $\mathbb{N}$, die ganzen Zahlen $\mathbb{Z}$, die rationalen Zahlen $\mathbb{Q}$, die reellen Zahlen $\mathbb{R}$ und schließlich die komplexen Zahlen $\mathbb{C}$ konstruieren. Zwischen diesen bestehen die Inklusionsbeziehungen

$$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}.$$

Wir gehen nicht ganz systematisch vor, sondern beginnen mit den natürlichen Zahlen.

1.1 Die Peano-Axiome

Die natürlichen Zahlen bilden die Menge $\mathbb{N} = \{1, 2, 3, \ldots\}$. Häufig nimmt man noch die Null hinzu und schreibt $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$. Die Menge $\mathbb{N}$ ist folgendermaßen charakterisiert:

Es gibt ein ausgezeichnetes Element $1 \in \mathbb{N}$ und eine Abbildung $f : \mathbb{N} \longrightarrow \mathbb{N}$, $n \mapsto f(n)$, die jedem Element seinen *Nachfolger* zuordnet. Diese haben die folgenden Eigenschaften:

P1 1 ist kein Nachfolger, d.h. es gibt kein $n \in \mathbb{N}$ mit $f(n) = 1$, und jedes Element ist durch seinen Nachfolger eindeutig bestimmt, d.h. aus $f(n) = f(m)$ folgt $n = m$.

P2 Ist $M \subset \mathbb{N}$ eine Teilmenge mit den Eigenschaften

i) $1 \in M$ und

ii) Für alle $n \in M$ gilt $f(n) \in M$,

so gilt: $M = \mathbb{N}$.

Man nennt P1 und P2 die Peanoaxiome[1]. Aus dem zweiten Peanoaxiom folgen zwei wichtige Prinzipien:

- Die Definition durch Rekursion.
- Der Beweis durch vollständige Induktion.

[1] Giuseppe Peano, $*$ 27.8.1858 † 20.4.1932.

1.2 Rekursive Definitionen

Satz 1.1 — *Es sei X eine Menge, $x_1 \in X$ und $g : X \longrightarrow X$ eine beliebige Abbildung. Dann gibt es genau eine Abbildung $\varphi : \mathbb{N} \longrightarrow X$ mit den Eigenschaften*

- $\varphi(1) = x_1$
- *Für alle $n \in \mathbb{N}$ gilt $\varphi(f(n)) = g(\varphi(n))$.*

Beweis. Der Beweis ist eher aufwendig als schwierig oder interessant. Ich verweise auf einschlägige Lehrbücher der Mengenlehre. □

Im folgenden bezeichne $f : \mathbb{N} \longrightarrow \mathbb{N}$ stets die Nachfolgerabbildung. Wir verwenden die folgende Sprechweise: „Wir definieren die Abbildung $\varphi : \mathbb{N} \longrightarrow X$ durch $\varphi(1) := x_1$ und $\varphi(f(n)) := g(\varphi(n))$ für alle $n \in \mathbb{N}$".

Definition 1.2 — Es sei $m \in \mathbb{N}$ gegeben. Wir definieren eine Additionsabbildung $\mathbb{N} \longrightarrow \mathbb{N}$, $n \mapsto m + n$, durch $m + 1 := f(m)$ und $m + f(n) := f(m + n)$.

Beispiel 1.3 — So ergibt sich formal korrekt, aber ziemlich schwerfällig

$$\begin{aligned} 3+3 &= 3 + f(2) = f(3+2) = f(3+f(1)) = f(f(3+1)) \\ &= f(f(f(3))) = f(f(4)) = f(5) = 6. \end{aligned}$$

Definition 1.4 — Es sei $m \in \mathbb{N}$. Wir definieren eine Multiplikationsabbildung $\mathbb{N} \longrightarrow \mathbb{N}$, $n \mapsto m \cdot n$ durch $m \cdot 1 := m$ und $m \cdot f(n) := m \cdot n + m$.

Beispiel 1.5 — Es sei $x \in \mathbb{R}$ [2]. Wir definieren den Binomialkoeffizienten $\binom{x}{n}$ für $n \in \mathbb{N}$ durch:

$$\binom{x}{1} := x, \quad \binom{x}{n+1} := \frac{x}{n+1} \cdot \binom{x-1}{n}.$$

Ausgeschrieben mit 'Pünktchen' heißt das

$$\binom{x}{n} = \frac{x \cdot (x-1) \cdot \ldots \cdot (x-n+1)}{n \cdot (n-1) \cdot \ldots \cdot 2 \cdot 1}.$$

Schließlich setzen wir noch $\binom{x}{0} := 1$.

[2] Im folgenden verwenden wir die reellen Zahlen in den Beispielen und greifen dabei auf Schulwissen zurück, obwohl wir uns vom logischen Standpunkt her erst in einem Stadium befinden, wo wir reelle Zahlen und ihre Eigenschaften noch gar nicht kennen.

Sind $a_1, \ldots, a_n \in \mathbb{R}$, so definieren wir das Summensymbol $\sum_{i=1}^{n} a_i$ durch $\sum_{i=1}^{1} a_i := a_1$ und $\sum_{i=1}^{n} a_i := a_n + \sum_{i=1}^{n-1} a_i$. Etwas schlampiger schreibt man $\sum_{i=1}^{n} a_i = a_1 + \ldots + a_n$. Analog definiert man die Produkte $\prod_{i=1}^{n} a_i = a_1 \cdot \ldots \cdot a_n$. Statt von 1 bis n kann der Index zwischen beliebigen ganzen Zahlen laufen. Sind oberer und unterer Index gleich, so besteht die Summe bzw. das Produkt nur aus einem Summanden bzw. einem Faktor. Ist der untere Index größer als der obere, so spricht man von einer leeren Summe bzw. einem leeren Produkt. Nach Konvention ist die leere Summe gleich 0, das leere Produkt gleich 1.

1.3 Vollständige Induktion

Satz 1.6 *(Prinzip der vollständigen Induktion) — Es sei für jedes $n \in \mathbb{N}$ eine Ausssage $A(n)$ gegeben und es gelte:*

- *$A(1)$ ist wahr (Induktionsverankerung), und*
- *$A(n) \Rightarrow A(n+1)$ für alle $n \in \mathbb{N}$ (Induktionsschritt).*

Dann ist die Aussage $A(n)$ für alle $n \in \mathbb{N}$ richtig.

Beweis. Von der Menge $M := \{n \in \mathbb{N} \mid A(n) \text{ ist richtig } \} \subset \mathbb{N}$ wissen wir: $1 \in M$ wegen der Induktionsverankerung, und für alle $n \in M$ gilt $n+1 \in M$ wegen des Induktionsschritts. Aus dem zweiten Peanoaxiom folgt nun $M = \mathbb{N}$, d.h. die Aussage $A(n)$ gilt für alle $n \in \mathbb{N}$. □

Satz 1.7 *(Assoziativität der Addition) — Für alle $n, m, p \in \mathbb{N}$ gilt:*

$$m + (n + p) = (m + n) + p.$$

Beweis. Es bezeichne $A(p)$ die Aussage

$$m + (n + p) = (m + n) + p, \quad \forall m, n \in \mathbb{N}.$$

Wir beweisen $A(p)$ für alle $p \in \mathbb{N}$ durch Induktion nach p.

Induktionsverankerung: Aussage $A(1)$ sagt: $m + (n + 1) = (m + n) + 1$ für alle $m, n \in \mathbb{N}$. Aber diese Aussage ist richtig nach Definition der Addition.

Induktionsschritt: $A(n) \Rightarrow A(n+1)$: Angenommen, die Aussage $A(p)$ ist wahr (Induktionshypothese). Es ist zu zeigen, dass dann auch $A(p+1)$ wahr ist.

In der Tat:

$$\begin{aligned}(m+n)+(p+1) &= ((m+n)+p)+1 \\ &\quad \text{nach Definition der Addition} \\ &= (m+(n+p))+1 \\ &\quad \text{weil nach Annahme } A(p) \text{ wahr ist} \\ &= m+((n+p)+1) \\ &\quad \text{nach Definition der Addition} \\ &= m+(n+(p+1)) \\ &\quad \text{nach Definition der Addition.}\end{aligned}$$

Das war zu zeigen. Durch Induktion folgt nun die Behauptung: $A(p)$ ist wahr für alle $p \in \mathbb{N}$. □

Satz 1.8 *(Kommutativität der Addition) — Für alle $n, m \in \mathbb{N}$ gilt:*

$$n + m = m + n.$$

Beweis. Übung! □

Satz 1.9 *— Für alle $n \in \mathbb{N}$ gilt: $\sum_{i=1}^{n} i = \binom{n+1}{2}$.*

Beweis.— durch vollständige Induktion. Wir bezeichnen mit $B(n)$ die Aussage $\sum_{i=1}^{n} i = \binom{n+1}{2}$.

Induktionsverankerung: $B(1)$ sagt: $1 = \sum_{i=1}^{1} i$ ist gleich $\binom{1+1}{2}$. Das ist sicher richtig.

Induktionsschritt: Es sei $n \geq 1$, und die Aussage $B(n)$ sei wahr. Es ist zu zeigen, dass $B(n+1)$ gilt.

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

Das war zu zeigen. Nach Induktion ist also $B(n)$ wahr für alle $n \in \mathbb{N}$. □

Beispiel 1.10 — $2^n \geq n^2 - 1$ für alle $n \in \mathbb{N}$. Wir beweisen die Behauptung durch Induktion. Angenommen, die Behauptung sei schon gezeigt für ein n. Gilt

es auch für $n+1$? Wir rechnen:

$$2^{n+1} = 2^n \cdot 2 \geq 2(n^2-1) = 2n^2 - 2 = [(n+1)^2 - 1] + [n^2 - 2n - 2].$$

Zu zeigen wäre: $n^2 - 2n - 2 \geq 0$. Wenn wir dies hätten, so könnten wir weiter schließen:

$$2^{n+1} \geq \ldots \geq (n+1)^2 - 1$$

und wären fertig. Aber gilt wirklich $0 \leq n^2 - 2n - 2 = (n-1)^2 - 3$? Dazu brauchen wir $n - 1 \geq \sqrt{3}$, d.h. $n \geq 3$. Damit also die hier vorgeschlagene Rechnung für den Induktionsschritt funktioniert, muss $n \geq 3$ gelten. Es genügt deshalb nicht, die Verankerung nur für $n = 1$ zu machen. Also zeigen wir direkt:

$$\begin{aligned} n = 1: &\quad 2^1 \geq 1^2 - 1 = 0 \\ n = 2: &\quad 2^2 \geq 2^2 - 1 = 3 \\ n = 3: &\quad 2^3 \geq 3^2 - 1 = 8. \end{aligned}$$

Ab jetzt greift unser Induktionsargument.

Bemerkung 1.11 — Induktion muss nicht bei 1 anfangen:

Es sei $n_0 \in \mathbb{Z}$. Gegeben seien Aussagen $A(n)$ für alle $n \in \mathbb{Z}$, $n \geq n_0$. Gilt nun:

1) $A(n_0)$ ist wahr.

2) Für alle $n \geq n_0$ gilt $A(n) \Rightarrow A(n+1)$,

dann gilt $A(n)$ für alle $n \geq n_0$.

Diese Form der vollständigen Induktion können wir formal auf die oben diskutierte wie folgt zurückführen: Wir definieren für alle $n \in \mathbb{N}$ die Aussage $B(n) := A(n_0 + n - 1)$, d.h. $B(1) = A(n_0)$, $B(2) = A(n_0 + 1)$ usw. Die Annahmen über A sagen dann, dass B die Bedingungen für die vollständige Induktion im Sinne des Satzes 1.6 erfüllt. Folglich gilt $B(n)$ für alle $n \in \mathbb{N}$ und somit $A(n)$ für alle $n \geq n_0$.

In der Praxis ist $n_0 = 0$ der häufigste Fall.

Satz 1.12 *(Binomischer Lehrsatz) — Es sei $n \in \mathbb{N}_0$ und $a, b \in \mathbb{R}$. Dann gilt:*

$$(a+b)^n = \sum_{i=0}^{n} \binom{n}{i} a^i b^{n-i}.$$

Beweis.— durch Induktion nach n. *Induktionsverankerung:* Für $n = 0$ steht links $(a+b)^0 = 1$ und rechts $\binom{0}{0} a^0 b^0 = 1$. *Induktionsschritt* Es sei $n \geq 0$ und

die Aussage für n bereits bewiesen. Dann gilt die Behauptung auch für $n+1$, denn:

$$\begin{aligned}
(a+b)^{n+1} &= (a+b)^n \cdot (a+b) = (a+b)^n \cdot a + (a+b)^n \cdot b \\
&= \sum_{k=0}^{n} \binom{n}{k} a^{k+1} b^{n-k} + \sum_{k=0}^{n} \binom{n}{k} a^k \cdot b^{n-k+1} \\
&= a^{n+1} + \sum_{\ell=1}^{n} \binom{n}{\ell-1} a^\ell b^{n+1-\ell} + \sum_{\ell=1}^{n} \binom{n}{\ell} a^\ell b^{n+1-\ell} + b^{n+1} \\
&= a^{n+1} + \sum_{\ell=1}^{n} \left\{ \binom{n}{\ell-1} + \binom{n}{\ell} \right\} a^\ell b^{n+1-\ell} + b^{n+1} \\
&= \sum_{\ell=0}^{n+1} \binom{n+1}{\ell} a^\ell b^{n+1-\ell}.
\end{aligned}$$

□

Satz 1.13 *(Variante zur Induktion) — Für jedes $n \in \mathbb{N}$ sei eine Aussage $A(n)$ gegeben. Es gelte:*

1. *$A(1)$ ist wahr.*
2. *Für jedes $k > 1$ folge aus der Richtigkeit aller $A(p)$, $p < k$, die Richtigkeit von $A(k)$.*

Dann gilt $A(n)$ für alle $n \in \mathbb{N}$.

Beweis. Wir betrachten die Aussage

$$B(n) := \quad A(k) \text{ ist richtig für alle } k \leq n.$$

Dann übersetzen sich die Voraussetzungen des Satzes in die folgenden Aussagen:
– $B(1)$ ist wahr.
– Ist $n \geq 1$, so folgt aus $B(n)$ die Aussage $A(n+1)$. Da nun nach Definition $B(n+1) = B(n) \wedge A(n+1)$, folgt aus $B(n)$ auch $B(n+1)$. Nach Induktion ist dann $B(n)$ richtig für alle $n \in \mathbb{N}$, also erst recht $A(n)$ für alle n. □

Beispiel 1.14 — Auch hierzu ein Beispiel. Eine natürliche Zahl n heißt bekanntlich Primzahl, falls $n \neq 1$ und falls aus $m|n$ folgt[3] $m = 1$ oder $m = n$.

Satz 1.15 — *Jede natürliche Zahl lässt sich als Produkt $n = p_1 \cdot \ldots \cdot p_s$ von Primzahlen $p_1, \ldots, p_s$ schreiben.*

[3]Das Symbol $m|n$ bedeutet: m teilt n, d.h. es gibt eine natürliche Zahl p mit $n = mp$.

Hier ist das leere Produkt, d.h. $s = 0$ und es gibt gar keine Faktoren, ausdrücklich zugelassen. Das leere Produkt hat definitionsgemäß den Wert 1. Anders könnten wir die 1 natürlich nicht als Produkt von Primzahlen darstellen.

Beweis. Für $n = 1$ ist der Satz nach der Vorbemerkung wahr. Es sei also ohne Einschränkung $n > 1$, und die Aussage sei schon für alle $k \in \mathbb{N}$, $k < n$ bewiesen (Induktionshypothese). Entweder ist nun n selbst eine Primzahl, dann sind wir mit der Wahl $s = 1$ und $p_1 = n$ fertig; oder es existiert eine natürliche Zahl m, $m \neq 1$ und $m \neq n$, die n teilt. Es sei $\ell := n/m \in \mathbb{N}$. Da m weder 1 noch n ist, gilt dasselbe für ℓ. Außerdem sind m und ℓ kleiner als n. Nach Induktionsvoraussetzung gibt es Primzahlen $p_1, \ldots, p_s$ und $q_1, \ldots, q_t$ mit $m = p_1 \cdot \ldots \cdot p_s$ und $\ell = q_1 \cdot \ldots \cdot q_t$. Es folgt nun

$$n = m\ell = (p_1 \cdot \ldots \cdot p_s) \cdot (q_1 \cdot \ldots \cdot q_t).$$

□

Schwieriger als die Existenz der Primfaktorzerlegung ist der Beweis der Eindeutigkeit. Das gehört aber eher in die Algebra als in die Analysis.

1.4 Mächtigkeit

Definition 1.16 — Eine Abbildung $f : X \longrightarrow Y$ heißt
– *surjektiv*, falls es zu jedem $y \in Y$ ein $x \in X$ mit $f(x) = y$ gibt,
– *injektiv*, falls für alle $x, x' \in X$, $x \neq x'$ folgt $f(x) \neq f(x')$, und
– *bijektiv*, falls f injektiv und surjektiv ist.
Statt von einer surjektiven, injektiven oder bijektiven Abbildung spricht man auch von einer Surjektion, einer Injektion bzw. einer Bijektion.

Es sei ausdrücklich darauf hingewiesen, dass die Frage, ob eine Abbildung injektiv ist oder nicht, völlig unabhängig davon ist, ob sie surjektiv ist oder nicht. Alle vier denkbaren Kombinantion sind tatsächlich möglich. Machen Sie sich das an ganz einfachen Beispielen unbedingt klar.

Definition 1.17 — Es seien X und Y Mengen.

1. X und Y heißen *gleichmächtig*, wenn es eine Bijektion $f : X \to Y$ gibt. In Zeichen: $|X| = |Y|$.

2. X ist *höchstens so mächtig* wie Y, wenn es eine Injektion $f : X \to Y$ gibt. In Zeichen: $|X| \leq |Y|$.

Die identische Abbildung $\mathrm{id}_X : X \longrightarrow X$ ist eine Bijektion von X in sich. Trivialerweise gilt daher $|X| = |X|$. Gibt es eine Bijektion $f : X \longrightarrow Y$, so gibt es zu jedem $y \in Y$ genau ein $x \in X$ mit $f(x) = y$. Für eine bijektive Abbildung f ist die Umkehrabbildung $f^{-1} : Y \longrightarrow X$ definiert, die y auf eben dieses x schickt. Die Umkehrabbildung ist selbst wieder bijektiv. So sieht man, dass aus $|X| = |Y|$ auch $|Y| = |X|$ folgt. Schließlich ist die Zusammensetzung $g \circ f$ zweier Bijektionen $f : X \longrightarrow Y$ und $g : Y \longrightarrow Z$ wieder bijektiv, und das impliziert: Aus $|X| = |Y|$ und $|Y| = |Z|$ folgt $|X| = |Z|$.

Ähnlich schnell überlegt man sich die folgenden Aussagen:

Satz 1.18 — *X, Y, Z seien Mengen. Dann gilt:*

1. *Aus $|X| = |Y|$ folgt $|X| \leq |Y|$.*
2. *Aus $|X| \leq |Y|$ und $|Y| \leq |Z|$ folgt $|X| \leq |Z|$.*

Satz 1.19 — *Es seien X und Y Mengen, $X \neq \emptyset$. Die folgenden Aussagen sind äquivalent:*

1. *Es gibt eine injektive Abbildung $f : X \longrightarrow Y$.*
2. *Es gibt eine surjektive Abbildung $g : Y \longrightarrow X$.*

Beweis.— Es sind zwei Richtungen zu zeigen:

1. $\Rightarrow$ 2: Es sei eine injektive Abbildung $f : X \longrightarrow Y$ gegeben und $A := f(X) \subset Y$ das Bild von f. Dann ist $f : X \longrightarrow A$ bijektiv und besitzt eine Umkehrabbildung $h : A \longrightarrow X$. Wähle $x_0 \in X$ beliebig und setze

$$g(y) := \begin{cases} h(y), & \text{falls } y \in A, \\ x_0, & \text{falls } y \in Y \setminus A. \end{cases}$$

Die so definierte Abbildung $g : Y \longrightarrow X$ ist surjektiv, denn ist $x \in X$ gegeben, so liegt $z := f(x)$ in A und es gilt $g(z) = h(z) = h(f(x)) = x$.

2. $\Rightarrow$ 1. Es sei eine surjektive Abbildung $g : Y \longrightarrow X$ gegeben. Das bedeutet, dass für jedes $x \in X$ die Urbildmenge $A_x := \{y \in Y | g(y) = x\}$ nicht leer ist. Wir wählen aus jeder Teilmenge A_x ein Element $a_x \in A_x$ aus und definieren $f : X \longrightarrow Y$ durch $f(x) := a_x$. Diese Abbildung ist injektiv: Für jedes $x \in X$ gilt nämlich: $f(x) = a_x \in A_x = g^{-1}(x)$ und $g(f(x)) = x$. Hätten wir also $f(x) = f(x')$, so folgte $x = g(f(x)) = g(f(x')) = x'$. □

Bemerkung 1.20 — In einer logisch fundierten Mengenlehre benötigt man ein besonderes Axiom, das sogenannte *Auswahlaxiom*, um den obigen Wahlvorgang zu rechtfertigen. (Literatur: Halmos, Naive Mengenlehre; Ebbinghaus et al., Zahlen)

Satz 1.21 *(Äquivalenzsatz von Schröder[4]-Bernstein[5]) — Es seien A und B Mengen mit $|A| \leq |B|$ und $|B| \leq |A|$. Dann gilt: $|A| = |B|$.*

Beweis. Nach Voraussetzung existieren injektive Abbildungen $f : A \longrightarrow B$ und $g : B \longrightarrow A$.

Wir konstruieren nun zu jedem $b \in B$ eine Folge von Elementen, die abwechselnd aus A und B kommen: Entweder ist $b \notin f(A)$ oder es gibt genau ein $a \in A$ mit $f(a) = b$. Entweder ist $a \notin g(B)$, oder es gibt genau ein $b_2 \in B$ mit $g(b_2) = a$. Wir fahren auf diese Weise fort, solange die Konstruktion nicht abbricht, d.h. solange $b_i \in f(A)$ bzw. $a_i \in g(B)$. Wir erhalten auf diese Weise eine *eindeutig* bestimmte Folge von Elementen $a_n \in A$, $b_n \in B$, mit

$$b = b_1 \overset{f}{\longleftarrow} a = a_1 \overset{g}{\longleftarrow} b_2 \overset{f}{\longleftarrow} a_2 \longleftarrow \ldots .$$

Diese Folge ist entweder unendlich oder bricht ab, unter Umständen schon bei b_1. Wir zerlegen B in drei disjunke Teilmengen $B = B_A \dot\cup B_B \dot\cup B_\infty$, wobei die Zugehörigkeit von b zu einer der Mengen davon abhängt, ob die bei b beginnende Folge bei einem Element aus A abbricht, bei einem Element in B abbricht, oder gar nicht abbricht.

Genauso zerlegen wir A in drei disjunkte Teilmengen $A = A_A \dot\cup A_B \dot\cup A_\infty$.

Man überlegt sich nun, dass für jedes der Symbole $i \in \{A, B, \infty\}$ gilt:

$$f(A_i) \subset B_i, \quad g(B_i) \subset A_i.$$

Außerdem müssen die Abbildungen

$$f : A_A \longrightarrow B_A, \quad f : A_\infty \longrightarrow B_\infty$$

und

$$g : B_B \longrightarrow A_B, \quad g : B_\infty \longrightarrow A_\infty$$

surjektiv sein, denn bei einem Element $b \in B$, das nicht im Bild von f liegt, bricht das Verfahren sofort ab, b liegt also in B_B, und für g schließt man analog. Da aber f und g injektiv sind, sind diese Abbildungen sogar bijektiv!

Daher können wir eine Bijektion $h : A \longrightarrow B$ durch Zusammenkleben definieren:

$$h(a) := \begin{cases} f(a) & \text{, falls} \quad a \in A_A \cup A_\infty \\ (g|_{B_B})^{-1}(a) & \text{, falls} \quad a \in A_B. \end{cases}$$

□

[4]Friedrich Wilhelm Karl Ernst Schröder, ∗ 25.11.1841 † 16.6.1902.

[5]Felix Bernstein, ∗ 24.2.1878 † 3.12.1956.

Satz 1.22 — *Es sei X eine unendliche Menge. Dann ist $|\mathbb{N}| \leq |X|$, d.h. es gibt eine injektive Abbildung $f : \mathbb{N} \longrightarrow X$.*

Beweis. Es sei $A \subset X$ eine beliebige endliche Menge. Da X unendlich ist, ist $X \setminus A$ nicht leer, und wir können ein Element $x_A \in X \setminus A$ auswählen. Wir definieren $f : \mathbb{N} \longrightarrow X$ rekursiv wie folgt: $f(1) := x_\emptyset$; sind $f(1), \ldots, f(n)$ schon definiert so sei $f(n+1) := x_{\{f(1),\ldots,f(n)\}}$.

Nach Konstruktion gilt also

$$f(n+1) \notin \{f(1), \ldots, f(n)\}$$

für alle $N \in \mathbb{N}$. Also ist f injektiv. Sind nämlich $k, n \in \mathbb{N}$ gegeben und ist $k \neq n$, so können wir ohne Einschränkung annehmen, dass $k < n$. Nach Konstruktion gilt aber $f(n) \notin \{f(1), \ldots, f(n-1)\} \ni f(k)$ und daraus folgt $f(n) \neq f(k)$. □

Definition 1.23 — Eine Menge X heißt abzählbar unendlich, wenn $|X| = |\mathbb{N}|$. Eine Menge heißt abzählbar, wenn sie endlich oder abzählbar unendlich ist.

Satz 1.24 — *Die folgenden Aussagen über eine Menge $X \neq \emptyset$ sind äquivalent:*

1. *X ist abzählbar.*
2. *Es gibt eine Injektion $f : X \longrightarrow \mathbb{N}$.*
3. *Es gibt eine Surjektion $g : \mathbb{N} \longrightarrow X$.*

Beweis.— Die Äquivalenz von 2. und 3. folgt aus einem früheren Satz.

1. $\Rightarrow$ 2. Ist X abzählbar, so ist X entweder abzählbar unendlich und somit sogar bijektiv zu $\mathbb{N}$, oder endlich und damit bijektiv zu einer Teilmenge $\{1, \ldots, n\}$ von $\mathbb{N}$.

2. $\Rightarrow$ 1. Es sei $f : X \longrightarrow \mathbb{N}$ eine injektive Abbildung. Falls X endlich ist, so ist X nach Definition abzählbar. Falls X unendlich ist, so gilt nach dem vorigen Satz $|\mathbb{N}| \leq |X|$. Nach dem Äquivalenzsatz von Schröder-Bernstein sind $\mathbb{N}$ und X gleichmächtig. □

Definition 1.25 — Eine Menge, die nicht abzählbar ist, heißt überabzählbar.

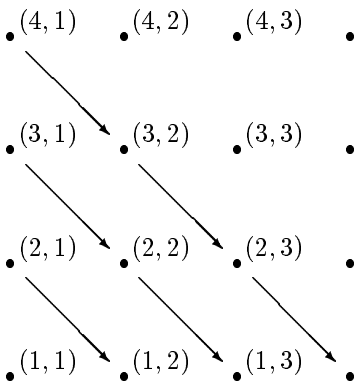
Beispiele 1.26 —

i) Die Menge $M = \{2n \mid n \in \mathbb{N}\} = \{2, 4, 6, \ldots\}$ der geraden Zahlen ist abzählbar: $f : \mathbb{N} \longrightarrow M, n \mapsto 2n$ ist eine Bijektion.

ii) Die Menge der ganzen Zahlen $\mathbb{Z} = \{\ldots, -1, 0, 1, 2, \ldots\}$ ist abzählbar, wie man aus der Darstellung $\mathbb{Z} = \{0, -1, 1, -2, 2, \ldots\}$ sofort sieht. Für die, die auf einer expliziten Bijektion bestehen: $f : \mathbb{N} \longrightarrow \mathbb{Z}, n \mapsto (-1)^n \left\lfloor \frac{n}{2} \right\rfloor$. (Hier steht $\lfloor u \rfloor$ für die größte ganze Zahl, die kleiner oder gleich der reellen Zahl u ist, d.h. die Abrundung von u.)

Satz 1.27 — *Das Produkt $\mathbb{N} \times \mathbb{N}$ ist abzählbar.*

1. Beweis: Erstes Cantorsches Diagonalverfahren. Wir zählen die Zahlenpaare (m, n) aus $\mathbb{N} \times \mathbb{N}$ in der folgenden Weise ab:



Übung: Man gebe eine Bijektion $f : \mathbb{N} \longrightarrow \mathbb{N} \times \mathbb{N}$, $n \mapsto (f_1(n), f_2(n))$, explizit an.

2. Beweis: Wir definieren $g : \mathbb{N} \times \mathbb{N} \mapsto \mathbb{N}, (n, m) \mapsto 2^n \cdot 3^m$, und zeigen: g ist injektiv. Angenommen, es wäre $2^n \cdot 3^m = 2^{n'} \cdot 3^{m'}$. Ohne Einschränkung sei $n \leq n'$. Dann gilt: $3^m = 2^{n'-n} \cdot 3^{m'}$. Die linke Seite ist ungerade, also auch die rechte: Daraus folgt $n = n'$ und somit $3^m = 3^{m'}$ und $m = m'$. Aber das bedeutet $(n, m) = (n', m')$. □

Lemma 1.28 — *Ist X abzählbar und $f : X \longrightarrow Y$ surjektiv, so ist Y ebenfalls abzählbar.*

Beweis.— Da X abzählbar ist, existiert eine Surjektion $\mathbb{N} \longrightarrow X$. Dann ist auch die zusammengesetzte Abbildung $\mathbb{N} \longrightarrow X \longrightarrow Y$ surjektiv und somit Y abzählbar. □

Folgerung 1.29 — *$\mathbb{Q}$ ist abzählbar.*

Beweis.— Wir wählen Bijektionen $f : \mathbb{N} \longrightarrow \mathbb{Z}$ und $\varphi : \mathbb{N} \longrightarrow \mathbb{N} \times \mathbb{N}$. Die Abbildung $\psi : \mathbb{Z} \times \mathbb{N} \longrightarrow \mathbb{Q}$, $(z, n) \mapsto \frac{z}{n}$, ist surjektiv. Die zusammengesetzte

Abbildung

$$\mathbb{N} \xrightarrow{\varphi} \mathbb{N} \times \mathbb{N} \xrightarrow{f \times \mathrm{id}_{\mathbb{N}}} \mathbb{Z} \times \mathbb{N} \xrightarrow{\psi} \mathbb{Q}$$

ist dann ebenfalls surjektiv. □

Satz 1.30 *(Cantor[6])— Die Menge $\mathbb{R}$ ist überabzählbar.*

Beweis.(2. Cantorsches Diagonalverfahren) Es genügt zu zeigen, dass das halboffene Intervall $[0,1) := \{x \in \mathbb{R} \mid 0 \leq x < 1\}$ überabzählbar ist.

Jedes $x \in [0,1)$ besitzt eine Dezimalentwicklung

$$x = 0, x_1 x_2 x_3 \ldots$$

mit $x_i \in \{0, 1, \ldots, 9\}$. Diese Entwicklung ist eindeutig, wenn wir verabreden, dass es kein $n \in \mathbb{N}$ gibt mit $x_\ell = 9$ für alle $\ell \geq n$. Wenn wir umgekehrt eine beliebige Folge $x_i \in \{0, 1, \ldots, 9\}$ wählen, die nicht fast alle gleich 9 sind, so ist $x = 0, x_1 x_2 x_3 \ldots$ eine eindeutig bestimmte reelle Zahl aus dem Intervall $[0,1)$. Cantor argumentiert wie folgt:

Angenommen, es gibt eine Bijektion $f : \mathbb{N} \longrightarrow [0,1)$. Wir können dann alle Zahlen $f(n)$ in ihrer Dezimalbruchentwicklung hinschreiben:

$$\begin{aligned} f(1) &= 0, x_{11}\ x_{12}\ x_{13}\ \ldots \\ f(2) &= 0, x_{21}\ x_{22}\ x_{23}\ \ldots \\ f(3) &= 0, x_{31}\ x_{32}\ x_{33}\ \ldots \end{aligned}$$

Für jedes $n \in \mathbb{N}$ wählen wir

$$y_n \in \{0,1\} \setminus \{x_{nn}\}.$$

(Eigentlich kommt es hier nur darauf an, dass wir irgendein y_n aus der Ziffernmenge $\{0, 1, 2, \ldots, 8\}$ auswählen, wobei wir die 9 vermeiden, damit wir nicht das Problem mit den periodischen Dezimalbrüchen bekommen, und das Element x_{nn} vermeiden, damit das folgende Argument funktioniert.) Wir setzen die Ziffern z_n zu einer Dezimalzahl zusammen:

$$y := 0, y_1 y_2 y_3 \ldots \in [0,1)$$

Es muss nun ein $N \in \mathbb{N}$ geben, so dass $f(N) = y$. Damit haben wir zwei Darstellungen von y:

$$\begin{aligned} y &= 0, y_1 y_2 y_3 \ldots \\ &= 0, x_{N1} x_{N2} x_{N3} \ldots, \end{aligned}$$

[6]Georg Ferdinand Ludwig Philipp Cantor, ∗ 3.3.1845 † 6.1.1918.

die wegen der Eindeutigkeit der Dezimalbruchentwicklung gleich sein müssen, d.h. für alle $i \in \mathbb{N}$ muss gelten $y_i = x_{N_i}$, also insbesondere auch $y_n = x_{NN}$. Aber das steht im Widerspruch zur Wahl von y_N!

Also war die Annahme, es gäbe eine Bijektion $f : \mathbb{N} \longrightarrow [0, 1)$, falsch. $\square$

Satz 1.31 *(Cantor) — Es sei X eine Menge, $\mathfrak{P}(X)$ ihre Potenzmenge. Dann gilt $|X| < |\mathfrak{P}(X)|$.*

Beweis.— Die Abbildung $X \longrightarrow \mathfrak{P}(X)$, $x \mapsto \{x\}$, ist injektiv, d.h. es gilt: $|X| \leq |\mathfrak{P}(X)|$. Deshalb genügt es, die Möglichkeit $|X| = |\mathfrak{P}(X)|$ auszuschließen. Angenommen, es gäbe eine Bijektion $\varphi : X \longrightarrow \mathfrak{P}(X)$. Für die Menge $A := \{x \in X \mid x \notin \varphi(x)\} \in \mathfrak{P}(X)$ muss es dann ein $a \in X$ mit $\varphi(a) = A$ geben. Nun ist $a \in A$ nach Definition von A äquivalent mit $a \notin \varphi(a) = A$. Ein feiner Widerspruch. Also gibt es keine Bijektion $\varphi : X \to \mathfrak{P}(X)$. $\square$

Satz 1.32 — *Sind X und Y beliebige Mengen, so gilt stets $|X| \leq |Y|$ oder $|Y| \leq |X|$.*

Der Beweis dieser harmlos aussehenden Aussage ist etwas aufwendiger. Ich verweise auf das Buch von Halmos zur Mengenlehre.

§2 Die reellen Zahlen

Wir charakterisieren die reellen Zahlen axiomatisch, d.h. wir beschreiben sie als eine Menge $\mathbb{R}$ mit zwei Verknüpfungen

$$\begin{aligned} + &: \mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{R} \quad \text{(Addition)} \\ \cdot &: \mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{R} \quad \text{(Multiplikation)} \end{aligned}$$

und einer Relation $\leq$ (Kleinerrelation), die die unten diskutierten Axiome erfüllen. Wir arbeiten danach ausschließlich mit diesen Axiomen und den daraus ableitbaren Sätzen. Dagegen lassen wir die Fragen, ob es überhaupt eine Realisierung dieser Axiome gibt, und ob es vielleicht verschiedene Realisierungen gibt, also die Fragen nach Existenz und Eindeutigkeit, vorläufig außer Betracht.

2.1 Axiome der Verknüpfungen und der Anordnung

1. Axiom $(\mathbb{R}, +, \cdot)$ ist ein Körper.

Definition 2.1 — Eine Menge K mit zwei Verknüpfungen $+$ und $\cdot$ ist ein Körper, wenn die folgenden Bedingungen erfüllt sind:

1. $(K, +)$ ist eine abelsche Gruppe, d.h. die Addition $+$ ist assoziativ und kommutativ. Es gibt ein Neutralelement $0 \in K$, die Null, mit der Eigenschaft $a + 0 = a$ für alle $a \in K$. Zu jedem $a \in K$ existiert ein additives Inverses $b \in K$ mit der Eigenschaft $a + b = 0$.

2. Die Multiplikation $\cdot$ ist assoziativ und kommutativ. Es gibt ein Neutralelement $1 \in K \setminus \{0\}$, die Eins, mit der Eigenschaft $a \cdot 1 = a$ für alle $a \in K$. Zu jedem $a \in K \setminus \{0\}$ existiert ein multiplikatives Inverses $c \in K$ mit der Eigenschaft $a \cdot c = 1$.

3. Addition und Multiplikation verhalten sich distributiv, d.h. für alle $a, b, c \in K$ gilt:
$$a \cdot (b + c) = ab + ac.$$

Bemerkungen 2.2 1. Die Elemente $0 \in K$ und $1 \in K$ sind eindeutig. Angenommen, 0 und $0'$ wären beides Nullelemente. Weil 0 Nullelement ist, folgt $0 + 0' = 0' + 0 = 0'$, aber weil auch $0'$ ein Nullelement ist, folgt $0 + 0' = 0$. Das zeigt $0 = 0'$. Genauso schließt man für das Einselement.

2. Das additive Inverse zu jedem $a \in K$ ist eindeutig. Es wird mit $-a$ bezeichnet. Es gilt $-0 = 0$ und $-(-a) = a$. Denn angenommen, b und b'

wären zwei Inverse von a. Dann gilt:

$$\begin{aligned} b &= b + 0 = b + (a + b') \\ &= (b + a) + b' = 0 + b' = b'. \end{aligned}$$

Liest man die Identität $a + (-a) = 0$ von $-a$ aus, so folgt $-(-a) = a$. In gleicher Weise folgert man: Das multipikative Inverse einer Zahl $a \in K \setminus \{0\}$ ist eindeutig. Es wird mit a^{-1} bezeichnet.

3. Die Assoziativität der Addition und der Multiplikation erlauben es, beliebige Summen bzw. Produkte beliebig umzuklammern. Wir können Klammern deshalb auch ganz weglassen. Das beweist man durch Induktion über die Länge n eines Summenausdrucks $((a_1+a_2)+a_3)+(\ldots a_n)$ etc. (Übung).

4. Für alle $a, b \in K$ gilt: $0 \cdot a = 0$ und $a \cdot (-b) = -(ab), (-a) \cdot (-b) = ab$. Es gilt nämlich $0 \cdot a = (0 + 0) \cdot a = 0 \cdot a + 0 \cdot a$. Nach Addition von $-(0 \cdot a)$ bleibt $0 = 0 \cdot a$. Nun folgt $0 = a \cdot 0 = a \cdot (b + (-b)) = ab + a \cdot (-b)$ und hieraus $a \cdot (-b) = -(ab)$. Nimmt man jetzt noch $-a$ statt a, so erhält man $(-a) \cdot (-b) = -((-a) \cdot b) = -(-ab) = ab$.

Beispiele 2.3 — 1. Die rationalen Zahlen $\mathbb{Q}$ mit der gewöhnlichen Addition und Multiplikation bilden einen Körper.
2. Die Menge $\mathbb{F}_2 = \{0, 1\}$ mit den Verknüpfungen, die durch die folgenden Tabellen gegeben sind, ist ein Körper:

$+$	0	1
0	0	1
1	1	0

$\cdot$	0	1
0	0	0
1	0	1

2. Axiom $(\mathbb{R}, \leq)$ ist eine geordnete Menge.

Definition 2.4 — Es sei X eine Menge und $\leq$ eine Relation auf X. Man nennt $\leq$ eine Halbordnung und das Paar $(X, \leq)$ eine halbgeordnete Menge, falls für alle $a, b, c \in X$ die folgenden Bedingungen erfüllt sind.

1. Reflexivität: $a \leq a$.

2. Antisymmetrie: $a \leq b, b \leq a \Rightarrow a = b$.

3. Transitivität: $a \leq b, b \leq c \Rightarrow a \leq c$.

Die Relation $\leq$ heißt Ordnung auf X und $(X, \leq)$ eine geordnete Menge, falls zusätzlich die folgende Vergleichbarkeitsbedingung erfüllt ist:

4. Für alle $a, b \in X$ gilt $a \leq b$ oder $b \leq a$.

Beispiel 2.5 — Es sei $M = \{a, b, \ldots\}$ eine Menge mit mindestens zwei Elementen a und b. Wir betrachten auf der Potenzmenge $X = \mathfrak{P}(M)$ die durch die Inklusion zweier Teilmengen gegebene Relation $\subset$. Dann ist $(\mathfrak{P}(M), \subset)$ eine halbgeordnete Menge. Aber sie ist nicht geordnet, weil weder $\{a\} \subset \{b\}$ noch $\{b\} \subset \{a\}$.

Bezeichnungen 2.6 — Wir benutzen für Elemente a und b in einer halbgeordneten Menge die folgenden Bezeichnungen:

$$\begin{aligned} a < b &:\Longleftrightarrow a \leq b \quad \text{und} \quad a \neq b, \\ a \geq b &:\Longleftrightarrow b \leq a, \\ a > b &:\Longleftrightarrow b < a. \end{aligned}$$

Axiom 1 beschreibt die arithmetischen Eigenschaften von $\mathbb{R}$, Axiom 2 die Anordnung von $\mathbb{R}$ entlang der Zahlengerade. Bisher sind die beiden Strukturen aber völlig unverbunden. Die Verbindung stellt das folgende Axiom her.

3. Axiom $(\mathbb{R}, +, \cdot, \leq)$ ist ein angeordneter Körper.

Definition 2.7 — Es sei $(K, +, \cdot)$ ein Körper mit einer Ordnungsrelation $\leq$. Dann heißt K angeordnet, falls für alle $a, b, c \in K$ die folgenden Bedingungen erfüllt sind:

1. $a > b \Longleftrightarrow a - b > 0$.
2. $a > 0, b > 0 \Longrightarrow a \cdot b > 0$.

Folgerung 2.8 — *Es sei K ein angeordneter Körper. Für alle $a, b, c \in K$ gilt:*

1. $a > b \implies a + c > b + c$.
2. $a > 0 \iff -a < 0$.
3. $a > b, c > 0 \implies ac > bc$.
4. $a > b, c < 0 \implies ac < bc$.

Beweis. Zu 1.: $a > b$ ist äquivalent zu der Aussage $0 < a - b = (a + c) - (b + c)$ und damit zu $a + c > b + c$.
Zu 2.: $a = (0 - (-a)) > 0 \Longleftrightarrow 0 > -a$.
Zu 3.: $a > b \Longleftrightarrow (a - b) > 0 \Longrightarrow (a - b) \cdot c > 0 \Longleftrightarrow ac - bc > 0 \Longleftrightarrow ac > bc$.
Zu 4.: Genauso. □

Folgerung 2.9 — *Es sei K ein angeordneter Körper. Für alle $a \in K$ gilt $a^2 \geq 0$, und Gleichheit besteht nur für $a = 0$.*

Beweis. Für jedes $a \in K$ gilt $a > 0$, $a = 0$ oder $a < 0$. Falls $a > 0$, so ist auch $a \cdot a > 0$ nach dem Anordnungsaxiom. Falls $a < 0$, so ist $(-a) > 0$ und $a^2 = (-a)^2 > 0$. □

2.2 Die natürlichen Zahlen

Im angeordneten Körper $\mathbb{R}$ gilt $1 = 1^2 > 0$ und induktiv $0 < 1 < 1 + 1 < 1 + 1 + 1 < \ldots$. Da die Zahlen immer größer werden, tritt keine mehrfach auf. Wir nutzen dies aus, um die natürlichen Zahlen in die reellen Zahlen einzubetten. Dazu müssen wir formal zunächst zwischen der 1 in $\mathbb{N}$ und der 1 in $\mathbb{R}$ unterscheiden. Dafür schreiben wir $1_{\mathbb{N}}$ und $1_{\mathbb{R}}$.

Satz 2.10 — *Es gibt eine eindeutig bestimmte Abbildung $f : \mathbb{N} \longrightarrow \mathbb{R}$ mit den Eigenschaften $f(1_{\mathbb{N}}) = 1_{\mathbb{R}}$ und $f(n + 1_{\mathbb{N}}) = f(n) + 1_{\mathbb{R}}$ für alle $n \in \mathbb{N}$. Diese Abbildung ist injektiv.*

Beweis. Die Existenz und Eindeutigkeit von f folgt aus dem Prinzip der rekursiven Definition. Die Injektivität ist eine unmittelbare Folgerung aus der Behauptung: $k < n \Rightarrow f(k) < f(n)$. Wir zeigen diese Aussage durch Induktion nach n. Für $n = 1_{\mathbb{N}}$ ist nichts zu zeigen.

Induktionsschritt: $n \longrightarrow n + 1$. Nach Voraussetzung gilt für alle $k < n$, dass $f(k) < f(n) < f(n) + 1_{\mathbb{R}} = f(n + 1_{\mathbb{N}})$. Damit ist auch der Fall $k = n$ und damit alle Fälle $k < n + 1$ abgedeckt. □

Der Satz erlaubt uns, $\mathbb{N}$ mit dem Bild unter f in $\mathbb{R}$ zu identifizieren. Wir können dann auch die vorübergehende Unterscheidung zwischen $1_{\mathbb{N}}$ und $1_{\mathbb{R}}$ aufheben. (Dasselbe Argument gilt natürlich für beliebige angeordnete Körper und erlaubt eine eindeutige Einbettung der natürliche Zahlen). Insbesondere ist jeder angeordnete Körper unendlich. Umgekehrt heißt das, dass endliche Körper keine Anordnung zulassen.

Definition 2.11 — $\mathbb{Z} := \mathbb{N} \cup \{0\} \cup \{-n \mid n \in \mathbb{N}\}$ heißt die Menge der ganzen Zahlen, $\mathbb{Q} := \{\frac{z}{n} \mid z \in \mathbb{Z}, n \in \mathbb{N}\}$ Menge der rationalen Zahlen.

Man rechnet leicht nach, dass $\mathbb{Z}$ und $\mathbb{Q}$ abgeschlossen unter Addition und Multiplikation sind, und dass $\mathbb{Q}$ ein angeordneter Körper ist.

Da wir jetzt wissen, dass in einem angeordneten Körper gilt $2 := 1 + 1 \neq 0$, können wir durch 2 dividieren und den folgenden Satz formulieren.

Satz 2.12 *(Ungleichung vom harmonischen und arithmetischen Mittel) — Es seien $x, y \in \mathbb{R}$, $x, y > 0$. Das arithmetische Mittel a und das harmonische Mittel h von x und y werden durch*

$$a = \frac{1}{2}(x+y) \quad \text{und} \quad \frac{1}{h} = \frac{1}{2}\left(\frac{1}{x} + \frac{1}{y}\right)$$

definiert. Es gilt $h \leq a$.

Beweis. Für alle $x, y > 0$ gilt $0 \leq (x-y)^2 = x^2 - 2xy + y^2$. Durch Addition von $4xy$ folgt $4xy \leq x^2 + 2xy + y^2 = (x+y)^2$ und somit

$$h = \frac{2xy}{x+y} \leq \frac{1}{2}(x+y) = a.$$

□

2.3 Das Axiom vom Dedekindschen Schnitt

Alle bisherigen Axiome sind auch für den Körper $\mathbb{Q}$ der rationalen Zahlen richtig. Das folgende Axiom begründet den wesentlichen Unterschied zwischen $\mathbb{Q}$ und $\mathbb{R}$.

Definition 2.13 — Es sei $(K, +, \cdot, \leq)$ ein angeordneter Körper. Ein Dedekindscher Schnitt ist ein Paar (A, B) von Teilmengen von K mit den folgenden Eigenschaften:

1. $A \cup B = K$ und $A \cap B = \emptyset$.
2. Für alle $a \in A$ und $b \in B$ gilt $a \leq b$.
3. $A \neq \emptyset$, $B \neq \emptyset$.

Ein $s \in K$ heißt Schnittzahl zum Dedekindschen Schnitt (A, B), falls $a \leq s$ für alle $a \in A$ und $b \geq s$ für alle $b \in B$.

Im folgenden schreiben wir kürzer $s \leq B$, falls $s \leq b$ für alle $b \in B$ etc.

Lemma 2.14 — *Zu jedem Dedekindschen Schnitt (A, B) gibt es höchstens eine Schnittzahl.*

Beweis. Angenommen, s und s' wären zwei Schnittzahlen und $s \neq s'$. Ohne Einschränkung sei $s < s'$. Wir betrachten $t := \frac{1}{2}(s + s')$. Dann gilt $s < t < s'$. Wäre $t \in A$, so hätte man $t \leq s$, wäre dagegeben $t \in B$, so hätte man $t \geq s'$. Beides ist falsch. Da $K = A \cup B$, hat man einen Widerspruch. □

4. Axiom vom Dedekindschen Schnitt.
Zu jedem Dedekindschen Schnitt in $\mathbb{R}$ gibt es eine Schnittzahl.

Wir werden später sehen, dass das Axiom vom Dedekindschen Schnitt in $\mathbb{Q}$ nicht gilt.

Definition 2.15 — Es sei $A \subset \mathbb{R}$ eine nicht leere Teilmenge, $a \in \mathbb{R}$.

1. a heißt eine obere Schranke von A, falls $A \leq a$.
2. A heißt nach oben beschränkt, falls es eine obere Schranke gibt.
3. a heißt kleinste obere Schranke (oder kurz: Supremum) von A, wenn a eine obere Schranke von A ist und wenn für jede andere obere Schranke s gilt: $a \leq s$.
4. a heißt Maximum von A, falls a eine kleinste obere Schranke ist und $a \in A$ gilt.

Analog definiert man die Begriffe: nach unten beschränkt, untere Schranke, größte untere Schranke (= Infimum), Minimum.

Wir schreiben $\sup(A)$, $\max(A)$, $\inf(A)$ und $\min(A)$ für das Supremum, das Maximum, das Infimum bzw. das Minimum einer Menge A, falls diese existieren.

Beispiel 2.16 — 3 ist eine obere Schranke von $A = \{x \in \mathbb{R} \mid x < 2\}$, und 2 ist das Supremum von A: Es ist klar, dass 2 eine obere Schranke ist. Wir müssen zeigen, dass es keine kleinere obere Schranke gibt. Angenommen, x wäre eine obere Schranke mit $x < 2$. Mit $t := \frac{x+2}{2}$ folgt $x < t < 2$, d.h. $t \in A$ und somit $t \leq x$, Widerspruch.

Satz 2.17 — *Jede nichtleere nach oben beschränkte Teilmenge von $\mathbb{R}$ hat ein eindeutig bestimmtes Supremum, jede nach unten beschränkte Teilmenge ein eindeutig bestimmtes Infimum.*

Beweis. Es sei M eine nicht leere, nach oben beschränkte Teilmenge von $\mathbb{R}$ und s eine obere Schranke.

Eindeutigkeit: Sind a, b Suprema von M, so ist b obere Schranke. Deshalb gilt $b \geq a$, weil a eine *kleinste* obere Schranke ist. Das Argument ist symmetrisch: es gilt also auch $a \geq b$. Zusammengenommen: $a = b$.

Die Menge $B := \{x \in \mathbb{R} \mid x > M\}$ enthält s und ist deshalb nicht leer, und $A := \mathbb{R} \setminus B$ ist auch nicht leer, denn $M \subset A$. Weiterhin gilt $A < B$. Denn andernfalls gäbe es ein $a \in A$ und ein $b \in B$ mit $a \geq b$. Wegen $b > M$ hätte

man auch $a > M$, also $a \in B$, im Widerspruch zu $A \cap B = \emptyset$. Zusammengefasst: (A, B) ist ein Dedekindscher Schnitt.

Es sei t die zugehörige Schnittzahl. Für diese gilt also $A \leq t \leq B$. Da $M \subset A$, ist t eine obere Schranke. Angenommen, t' wäre eine obere Schranke von M mit $t' < t$. Für $u := \frac{1}{2}(t + t')$ gilt: $t' < u < t$, also insbesondere $u > M$. Dann läge u in B, und $u < t$ widerspräche $t \leq B$.

Für das Infimum nach unten beschränkter Mengen schließt man analog. □

Satz 2.18 *(Archimedisches Axiom) — Für jedes $x \in \mathbb{R}$ existiert ein $n \in \mathbb{N}$ mit $n > x$.*

Beweis. Andernfalls wäre die Menge $\mathbb{N}$ nach oben beschränkt und hätte deshalb ein Supremum s. Dann kann $s - 1$ keine obere Schranke von $\mathbb{N}$ sein und es gibt ein $n \in \mathbb{N}$ mit $s - 1 < n$. Aber jetzt folgt $s < n + 1 \in \mathbb{N}$ im Widerspruch dazu, dass s ein Supremum sein soll. □

Bemerkung 2.19 — Es sei $(K, +, \cdot, \leq)$ ein angeordneter Körper. Man sagt, K sei archimedisch angeordnet, wenn für jedes $a \in K$ ein $n \in \mathbb{N}$ mit $a < n$ existiert. Mit dieser Bezeichnung sagt der Satz, dass jeder angeordnete Körper, in dem das Axiom vom Dedekindschen Schnitt gilt, archimedisch angeordnet ist. Satz 2.20 und Satz 2.21 gelten allgemeiner in archimedisch angeordneten Körpern.

Satz 2.20 *(Bernoullische Ungleichung) — Es sei $x \in \mathbb{R}$, $x \geq -1$, und $n \in \mathbb{N}$. Dann gilt*

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

und Gleichheit tritt nur für $n = 1$ oder $x = 0$ ein.

Beweis. Für $x = -1$ wird behauptet, dass $0^n \geq 1 - n$, mit Gleichheit genau für $n = 0$. Das ist sicher wahr. Wir nehmen also an, dass $(x + 1) > 0$, und beweisen die Behauptung durch Induktion nach n. Für $n = 1$ ist nichts zu zeigen.

Induktionsschritt $n \longrightarrow n + 1$: Wir nehmen an, die Behauptung sei wahr für n. Es folgt:

$$\begin{aligned} (1 + x)^{n+1} &= (1 + x)^n \cdot (1 + x) \geq (1 + nx) \cdot (1 + x) \\ &= 1 + (n + 1) \cdot x + nx^2 \\ &\geq 1 + (n + 1) \cdot x \end{aligned}$$

Gleichheit kann hier nur eintreten, wenn $nx^2 = 0$, also $x = 0$. Und dann gilt tatsächlich auch Gleichheit. □

Satz 2.21 —

1. *Es sei $g \in \mathbb{R}$, $g > 0$. Für alle $a \in \mathbb{R}$ existiert ein $n \in \mathbb{N}$ derart, dass $n \cdot g > a$.*
2. *Es sei $g \in \mathbb{R}$, $g > 1$. Für alle $a \in \mathbb{R}$ existiert ein $n \in \mathbb{N}$ derart, dass $g^n > a$.*
3. *Für alle $\varepsilon \in \mathbb{R}$, $\varepsilon > 0$, existiert ein $n \in \mathbb{N}$ derart, dass $0 < \frac{1}{n} < \varepsilon$.*
4. *Für alle $\varepsilon \in \mathbb{R}$, $\varepsilon > 0$, und alle $g \in \mathbb{R}$, $0 < g < 1$, existiert ein $n \in \mathbb{N}$ derart, dass $0 < g^n < \varepsilon$.*

Beweis. Zu 1.: Es gibt ein $n \in \mathbb{N}$ mit $n > \frac{a}{g}$. Da $g > 0$, folgt $n \cdot g > a$.
Zu 2.: Da $g > 1$, ist $(g-1) > 0$. Es existiert also ein $n \in \mathbb{N}$ mit $n(g-1) > a$. Mit der Bernoullischen Ungleichung folgt:

$$g^n = (1 + (g-1))^n \geq 1 + n(g-1) > 1 + a > a.$$

Zu 3.: Es gibt ein $n \in \mathbb{N}$ mit $n > \frac{1}{\varepsilon}$. Da $\varepsilon > 0$ und $\frac{1}{n} > 0$, folgt: $\varepsilon > \frac{1}{n}$.
Zu 4.: Aus $0 < g < 1$ folgt: $\frac{1}{g} > 1$. Daher existiert ein $n \in \mathbb{N}$ mit $(\frac{1}{g})^n = \frac{1}{g^n} > \frac{1}{\varepsilon}$.
Es folgt $0 < g^n < \varepsilon$. □

Bezeichnungen 2.22 — Wir schreiben $\mathbb{R}_{>0} := \{x \in \mathbb{R} \,|\, x > 0\}$ und analog $\mathbb{R}_{\geq 0}$, $\mathbb{R}_{<0}$ und $\mathbb{R}_{\leq 0}$.

Satz 2.23 — *Es sei $x \in \mathbb{R}_{\geq 0}$, und $n \in \mathbb{N}$. Dann existiert eine eindeutig bestimmte Zahl $y \in \mathbb{R}_{\geq 0}$ mit $y^n = x$. Die Zahl y heißt die n-te Wurzel aus x und wird mit $\sqrt[n]{x} := y$ bezeichnet.*

Ich weise ausdrücklich darauf hin, dass das Symbol $\sqrt[n]{x}$ nur für $x \geq 0$ definiert ist. Für den Beweis brauchen wir das folgende Lemma:

Lemma 2.24 — *Es seien $y, z \in \mathbb{R}_{\geq 0}$ und $n \in \mathbb{N}$. Dann gilt:*

$$y < z \quad \Leftrightarrow y^n < z^n.$$

Beweis. Es sei ohne Einschränkung $n \geq 2$. Dann gilt

$$y^n - z^n = (y - z)(y^{n-1} + y^{n-2}z + \ldots + z^{n-1}).$$

Falls $y = z = 0$, ist nichts zu zeigen, und andernfalls ist der Ausdruck $a := y^{n-1} + y^{n-2}z + \ldots + z^{n-1}$ positiv. Daraus folgt: Ist $y - z > 0$, so gilt auch $y^n - z^n = a(y - z) > 0$, und ist umgekehrt $y^n - z^n > 0$, so gilt auch $y - z = (y^n - z^n)/a > 0$. □

Zum Beweis des Satzes. Aus dem Lemma folgt sofort die Eindeutigkeit der Wurzel. Wir müssen die Existenz zeigen. Dazu betrachten wir die Menge $A := \{z \in \mathbb{R}_{\geq 0} \mid z^n \leq x\}$. Da $0 \in A$, ist A nicht leer. Falls $x \geq 1$, so gilt $x^n \geq x$. Für $u := \max\{1, x\}$ folgt in jedem Falle $u^n \geq x$, und aus dem Lemma erhält man $A \leq u$. Demnach ist A nach oben beschränkt und besitzt eine Supremum y.

Wir wollen zeigen, dass $y^n = x$. Dazu führen wir die Annahmen $y^n > x$ und $y^n < x$ zum Widerspruch.

1. Fall: $y^n > x$. Es sei $C := ny^{n-1} + \binom{n}{2}y^{n-2} + \ldots + 1$. Dann gilt für $\varepsilon = \min\{1, (y^n - x)/C\}$ die folgende Abschätzung:

$$\begin{aligned}(y-\varepsilon)^n &= y^n - ny^{n-1}\varepsilon + \binom{n}{2}y^{n-2}\varepsilon^2 - \ldots + (-1)^n\varepsilon^n \\ &\geq y^n - \varepsilon\left(ny^{n-1} + \binom{n}{2}y^{n-2}\varepsilon + \ldots + \varepsilon^{n-1}\right) \\ &\geq y^n - \varepsilon C, \text{ weil } \varepsilon \leq 1, \\ &\geq x, \text{ weil } \varepsilon C \leq y^n - x.\end{aligned}$$

Also wäre auch noch $y - \varepsilon$ eine obere Schranke von A, im Widerspruch zur Definition von y.

2. Fall: $y^n < x$. Es sei ε so gewählt, dass $0 < \varepsilon < \min\{1, (x - y^n)/C\}$, und C wie im ersten Fall. Dann zeigt man ganz wie im ersten Fall, dass $(y + \varepsilon)^n \leq y^n + \varepsilon C < x$. Es folgt, dass $y + \varepsilon \in A$. Aber die Ungleichung $y < y + \varepsilon$ widerspricht der Annahme, dass y eine Supremum von A sein soll.

Es bleibt also insgesamt nur die Möglichkeit $y^n = x$, wie gewünscht. □

2.4 Der Absolutbetrag

Definition 2.25 — Für jedes $x \in \mathbb{R}$ heißt $|x| := \max\{x, -x\}$ der Absolutbetrag (oder Betrag) von x.

Die folgenden Rechenregeln ergeben sich unmittelbar aus der Definition.

Satz 2.26 *(Eigenschaften des Betrags) — Für alle reellen Zahlen x, y, z gilt:*

1) $|x| \geq 0$, *und* $|x| = 0$ *genau dann, wenn* $x = 0$.

2) $|x + y| \leq |x| + |y|$ *(Dreiecksungleichung).*

3) $|x \cdot y| = |x| \cdot |y|$

4) $-|x| \leq x \leq |x|$.

Folgerung 2.27 — *Für alle reellen Zahlen $a, b, c, a_i \in \mathbb{R}$ gilt:*

1. $|\sum_{i=1}^{n} a_i| \leq \sum_{i=1}^{n} |a_i|$.

2. $|a-c| \leq |a-b| + |b-c|$ *(Dreiecksungleichung)*.

3. $|a-b| \geq ||a| - |b||$.

Beweis. 1. folgt durch Induktion aus der Dreiecksungleichung.
2. folgt aus der Dreiecksungleichung, angewendet auf $(a-b)+(b-c) = (a-c)$.
Zu 3.: Wendet man die Dreiecksungleichung auf $(a-b)+b = a$ an, so erhält man $|a| \leq |a-b| + |b|$, also $|a| - |b| \leq |a-b|$. Die Situation ist symmetrisch in a und b, also gilt auch $|b| - |a| \leq |b-a| = |a-b|$. Daraus folgt $||a| - |b|| \leq |a-b|$. □

2.5 Die erweiterte Zahlengerade

Wir fügen der Menge $\mathbb{R}$ zwei neue Elemente hinzu, die mit den Symbolen $-\infty$ und ∞ bezeichnet werden.

Definition 2.28 — Die Menge $\overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, \infty\}$ heißt erweiterte Zahlengerade. Die Relation $<$ wird auf $\overline{\mathbb{R}}$ durch $-\infty < a < \infty$ für alle $a \in \mathbb{R}$ erweitert. Die Rechenoperationen sind auf $\overline{\mathbb{R}}$ nur teilweise definiert:

1. Für alle $a \in \mathbb{R} \cup \{\infty\}$ sei $a + \infty := a - (-\infty) := \infty$, und für alle $a \in \mathbb{R} \cup \{-\infty\}$ sei $a - \infty := a + (-\infty) := -\infty$.

2. Für alle $a > 0$ setzen wir $a \cdot \infty := \infty$, $a \cdot (-\infty) := -\infty$, und für alle $a < 0$ entsprechend $a \cdot \infty := -\infty$, $a \cdot (-\infty) := \infty$.

3. Für alle $a \in \mathbb{R}$ sei $\frac{a}{\infty} := \frac{a}{-\infty} := 0$.

Man beachte, dass alle Ausdrücke, die nicht in der obigen Liste vorkommen, nicht definiert sind und keine Bedeutung haben, also etwa $\infty - \infty$, $\frac{a}{0}$, $\frac{\infty}{\infty}$ usw.

2.A Beispiel eines nichtarchimedisch angeordneten Körpers

Man kann angeordnete Körper auch in einer anderen Weise charakterisieren, die zur Definition 2.7 äquivalent ist.

Definition 2.29 — Es sei K ein Körper. Eine Teilmenge $P \subset K$ heißt Positivbereich, wenn die folgenden Bedingungen erfüllt sind:

1. Für alle $a, b \in P$ gilt $a + b \in P$ und $ab \in P$.

2. Für jede Zahl $a \in K$ gilt genau eine der drei Möglichkeiten $a \in P$, $-a \in P$ oder $a = 0$.

Es ist nicht schwer sich das folgende klarzumachen:

1. Ist $\leq$ eine Halbordnung auf K, die K zu einem angeordneten Körper macht, so ist die Menge $P := \{x \in K \mid x > 0\}$ ein Positivbereich.

2. Ist dagegen $P \subset K$ ein Positivbereich, so wird durch

$$a \leq b \quad :\Leftrightarrow \quad a = b \text{ oder } b - a \in P$$

   eine Anordnung auf K definiert.

3. Die beiden Konstruktionen sind invers zueinander und geben eine Bijektion zwischen der Menge aller Anordnungen auf K und der Menge aller Positivbereiche auf K.

Wir machen uns diese Beschreibung einer Anordnung zunutze, um ein Beispiel für einen angeordneten Körper anzugeben, der *nicht* archimedisch angeordnet ist.

Wir betrachten den Körper $\mathbb{Q}(X) = \{f/g \mid f, \in \mathbb{Q}[X], g \neq 0\}$ aller rationalen Funktionen in einer Unbestimmten. Die rationalen Zahlen $\mathbb{Q}$ bilden einen Unterkörper von K: Wir identifizieren eine Zahl mit der entsprechenden konstanten Abbildung. Ein beliebiges Element $a \in \mathbb{Q}(X)^*$ lässt sich immer auf die folgende Weise beschreiben:

$$a = \frac{f_n X^n + \ldots + f_1 X + f_0}{g_m X^m + \ldots + g_0},$$

wobei f_i, g_i rationale Zahlen und $F_n, g_m \neq 0$ sind. Der Quotient f_n/g_m hängt nicht von der Wahl der Darstellung $a = f/g$ mit Polynomen f und g ab. Deshalb ist die Menge

$$P := \{a \in \mathbb{Q}(X) \mid f_n/g_m > 0\}$$

wohldefiniert. Außerdem ist nach den Regeln für das Rechnen mit Brüchen klar, dass P ein Positivbereich ist. Insbesondere liefert uns diese Konstruktion eine Anordnung auf $\mathbb{Q}(X)$.

Diese Anordnung ist nicht archimedisch: Es sei $n \in \mathbb{N}$ eine beliebige natürliche Zahl. Das Polynome $X - n \in \mathbb{Q}(X)$ hat den Leitkoeffizienten 1 und liegt deshalb in P. Anders gesagt: $X - n > 0$. Durch Addition von n folgt $X > n$. Aber n war beliebig, d.h. $X > n$ für alle $n \in \mathbb{N}$.

§3 Folgen

3.1 Folgen und Grenzwerte

Definition 3.1 — Unter einer Folge reeller Zahlen versteht man eine Abbildung $f : \mathbb{N} \longrightarrow \mathbb{R}$. Ist $a_n := f(n)$, so schreibt man (a_n) anstelle von f. Häufig betrachtet man auch Folgen, deren Index von einem anderen Wert als 1 an läuft, so bedeutet $(a_n)_{n \geq n_0}$ formal eine Abbildung $f : \{n \in \mathbb{Z} \mid n \geq n_0\} \longrightarrow \mathbb{R}$.

Definition 3.2 — Es sei $a \in \mathbb{R}$ und $\varepsilon > 0$. Die Menge

$$U_\varepsilon(a) := \{x \in \mathbb{R} \mid |x - a| < \varepsilon\}$$

heißt ε-Umgebung von a.

Definition 3.3 — Eine Folge (a_n) reeller Zahlen heißt konvergent mit Grenzwert $a \in \mathbb{R}$, wenn es zu jedem $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$ mit der Eigenschaft gibt, dass $a_n \in U_\varepsilon(a)$ für alle $n \geq n_0$. In diesem Falle schreibt man

$$a = \lim_{n \to \infty} (a_n)$$

oder $a_n \longrightarrow a$ für $n \longrightarrow \infty$. Eine Folge, die gegen 0 konvergiert, heißt Nullfolge. Eine nicht konvergente Folge heißt divergent.

Satz 3.4 — *Ist (a_n) eine konvergente Folge, so ist der Grenzwert eindeutig bestimmt.*

Beweis. Angenommen, (a_n) konvergiere sowohl gegen a als auch gegen a' und es gelte $a \neq a'$. Es sei $\delta := |a - a'|$. Dann gibt es ein n_1 so, dass für alle $n \geq n_1$ gilt: $|a_n - a| < \frac{\delta}{2}$, und es gibt weiter ein n_2 so, dass für alle $n \geq n_2$ gilt $|a_n - a'| < \frac{\delta}{2}$. Für alle $n \geq n_3 := \max\{n_1, n_2\}$ folgt:

$$\delta = |a - a'| \ \leq \ |a - a_n| \ + \ |a_n - a'| < \frac{\delta}{2} + \frac{\delta}{2} = \delta.$$

Widerspruch! □

Beispiele 3.5 —

i) Die konstante Folge $(a_n) = (a)$ konvergiert gegen a.

ii) Die harmonische Folge $(\frac{1}{n})$ ist eine Nullfolge. Es sei nämlich $\varepsilon > 0$ vorgegeben. Dann gibt es nach dem archimedischen Axiom ein $n_0 \in \mathbb{N}$ mit $\frac{1}{n_0} < \varepsilon$. Für alle $n \geq n_0$ folgt: $\frac{1}{n} \leq \frac{1}{n_0} < \varepsilon$, und das heißt: $\frac{1}{n} \in U_\varepsilon(0)$.

iii) Das Konvergenzverhalten der geometrischen Folge (a^n) hängt vom Absolutbetrag von a ab:

1.Fall $0 \leq |a| < 1$. Zu jedem $\varepsilon > 0$ existiert dann ein $n_0 \in \mathbb{N}$ derart, dass $|a^{n_0}| < \varepsilon$. Für alle $n \geq n_0$ folgt $|a^n| \leq |a^{n_0}| < \varepsilon$. Das bedeutet: $\lim_{n\to\infty} a^n = 0$.

2. Fall $|a| = 1$. Das heißt entweder: $a = 1$, und dann ist (a^n) konstant und $\lim 1^n = 1$, oder: $a = -1$, und dann alterniert das Vorzeichen $((-1)^n) = (-1, 1, -1, \ldots)$: Die Folge ist divergent.

3. Fall $|a| > 1$. Zu jedem $N \in \mathbb{N}$ existiert dann ein n mit $|a^n| > N$. Die Folge (a^n) ist unbeschränkt und muss gemäß dem folgenden Lemma divergieren.

Definition 3.6 — Eine Folge (a_n) heißt beschränkt (bzw. nach oben oder nach unten beschränkt), wenn dies für die Menge $\{a_n \mid n \in \mathbb{N}\}$ gilt.

Lemma 3.7 — *Jede konvergente Folge ist beschränkt.*

Beweis. Es sei (a_n) eine konvergente Folge und $a := \lim(a_n)$. Es gibt ein $n_0 \in \mathbb{N}$ derart, dass für alle $n \geq n_0$ gilt: $|a - a_n| \leq 1$. Es folgt für alle $n \geq n_0$:

$$\begin{aligned} |a_n| &= |a + (a_n - a)| \leq |a| + |a_n - a| \\ &\leq |a| + 1. \end{aligned}$$

Dann ist $s := \max\{|a| + 1, |a_1|, \ldots, |a_{n_0-1}|\}$ eine Schranke für (a_n). □

Satz 3.8 — *Sind (a_n) und (b_n) konvergente Folgen, so ist auch $(a_n + b_n)$ konvergent, und es gilt*

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

Beweis. Es sei $a := \lim_{n\to\infty} a_n$, $b := \lim_{n\to\infty} b_n$. Ist $\varepsilon > 0$ vorgegeben, so existieren $n_1, n_2 \in \mathbb{N}$ derart, dass

$$\begin{aligned} |a - a_n| &< \frac{\varepsilon}{2} \quad \forall\, n \geq n_1 \\ |b - b_n| &< \frac{\varepsilon}{2} \quad \forall\, n \geq n_2. \end{aligned}$$

Jetzt gilt für alle $n \geq n_0 := \max\{n_1, n_2\}$

$$\begin{aligned} |(a + b) - (a_n + b_n)| &\leq |a - a_n| + |b - b_n| \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

□

Satz 3.9 — *Sind (a_n) und (b_n) konvergente Folgen, so ist auch $(a_n \cdot b_n)$ konvergent, und es gilt*

$$\lim_{n\to\infty}(a_n \cdot b_n) = \lim_{n\to\infty}(a_n) \cdot \lim_{n\to\infty}(b_n).$$

Beweis. Es sei $a := \lim_{n\to\infty} a_n$ und $b := \lim_{n\to\infty} b_n$. Für $n \in \mathbb{N}$ gilt:

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

Die Folge (b_n) ist konvergent, also auch beschränkt. Wir wählen $s > 0$ so, dass $|b_n| < s \; \forall \, n \in \mathbb{N}$. Es sei nun $\varepsilon > 0$ vorgegeben. Es existieren $n_1, n_2 \in \mathbb{N}$ derart, dass

$$\begin{aligned} |b - b_n| &< \tfrac{\varepsilon}{2(1+|a|)} \quad \forall \, n \geq n_1 \\ |a - a_n| &< \tfrac{\varepsilon}{2 \cdot s} \quad \forall \, n \geq n_2. \end{aligned}$$

Für alle $n \geq n_0 := \max\{n_1, n_2\}$ gilt somit

$$\begin{aligned} |ab - a_n b_n| &\leq |a| \cdot |b - b_n| + |b_n| \cdot |a - a_n| \\ &< |a| \cdot \frac{\varepsilon}{2(1+|a|)} + s \cdot \frac{\varepsilon}{2s} \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

Das bedeutet: $\lim_{n\to\infty} a_n b_n = ab$. □

Folgerung 3.10 — *Ist $\lambda \in \mathbb{R}$ und ist (a_n) konvergent, so konvergiert auch (λa_n) und*

$$\lim_{n\to\infty}(\lambda a_n) = \lambda \cdot \lim_{n\to\infty}(a_n).$$

Beweis. Wähle $b_n := \lambda \; \forall \, n \in \mathbb{N}$. □

Folgerung 3.11 — *Sind (a_n), (b_n) konvergent, so auch $(a_n - b_n)$, und es gilt* $\lim_{n\to\infty}(a_n - b_n) = \lim_{n\to\infty} a_n - \lim_{n\to\infty} b_n$.

Satz 3.12 — *Es sei (b_n) konvergent, $b_n \neq 0 \; \forall \, n \in \mathbb{N}$, und $b := \lim_{n\to\infty} b_n \neq 0$. Dann ist auch $(\frac{1}{b_n})$ konvergent, und es gilt* $\lim_{n\to\infty} 1/b_n = 1/b$.

Beweis. Es sei $\varepsilon > 0$ vorgegeben. Es gibt ein $n_1 \in \mathbb{N}$ so, dass für alle $n \geq n_1$ gilt: $|b_n - b| < \frac{|b|}{2}$. Insbesondere folgt:

$$\begin{aligned} |b_n| &= |b - (b - b_n)| \\ &\geq |b| - |b - b_n| \\ &\geq |b| - \frac{1}{2}|b| = \frac{1}{2}|b|. \end{aligned}$$

Es gibt weiter ein n_2 so, dass für alle $n \geq n_2$ gilt:

$$|b_n - b| \ < \varepsilon \cdot \frac{1}{2}|b|^2.$$

Damit folgt für alle $n \geq n_0 := \max\{n_1, n_2\}$:

$$|\frac{1}{b_n} - \frac{1}{b}| = \frac{|b_n - b|}{|b_n| \cdot |b|} < \frac{\varepsilon \cdot \frac{1}{2}|b|^2}{\frac{1}{2}|b| \cdot |b|} = \varepsilon.$$

□

Folgerung 3.13 — *Sind (a_n), (b_n) konvergente Folgen, $a := \lim a_n$, $b := \lim b_n$, und gilt $b_n \neq 0 \ \forall \ n \in \mathbb{N}$ und $b \neq 0$, so ist auch $(\frac{a_n}{b_n})$ konvergent, und es gilt*

$$\lim_{n\to\infty} \frac{a_n}{b_n} = \frac{a}{b}.$$

Beweis. Betrachte $(\frac{a_n}{b_n}) = (a_n \cdot \frac{1}{b_n})$. □

Beispiel 3.14 — Es gilt

$$\lim_{n\to\infty} \frac{n^2 + 7n + 2}{3n^2 + 8n + 7} = \frac{1}{3},$$

denn

$$\frac{n^2 + 7n + 2}{3n^2 + 8n + 7} = \frac{1 + 7\frac{1}{n} + 2\frac{1}{n^2}}{3 + 8\frac{1}{n} + 7\frac{1}{n^2}}$$

und $\lim_{n\to\infty} 1 + 7\frac{1}{n} + 2\frac{1}{n^2} = 1$ sowie $\lim_{n\to\infty} 3 + 8\frac{1}{n} + 7\frac{1}{n^2} = 3$.

Beispiel 3.15 (Die Fibonaccizahlen) — Leonardo di Pisa[7] (=Fibonacci) hat die folgende Zahlenfolge in die Mathematik eingeführt, und zwar um das Wachstum einer Kaninchenpopulation zu beschreiben:

$$f_0 = 1, \quad f_1 = 1, \quad f_2 = 2, \quad f_3 = 3, \quad f_4 = 5, \quad f_5 = 8, \ldots$$

Das Bildungsgesetz der Folge lautet

$$f_{n+1} = f_n + f_{n-1} \quad \text{für alle } n \geq 1.$$

Man sieht induktiv, dass die Abschätzungen $f_n \leq f_{n+1} \leq 2f_n$ bestehen. Wir betrachten die Folge der Quotienten auf einander folgender Fibonaccizahlen: $x_n := f_{n+1}/f_n$. Wir bekommen:

$$(x_n)_{n\geq 0} = (1, 2, \frac{3}{2}, \frac{5}{3}, \frac{8}{5}, \ldots)$$

[7] Leonardo di Pisa, $*$ 1170 † 1250.

Die Rekursionsgleichung der f_n übersetzt sich in die Gleichung

$$x_n = 1 + \frac{1}{x_{n-1}}. \tag{3.1}$$

Wir zeigen, dass die Folge (x_n) konvergiert.

1. Schritt. Wir nehmen zunächst an, wir wüssten bereits, dass die Folge konvergiert, und berechnen den Grenzwert α. Indem wir in (3.1) auf beiden Seiten zum Grenzwert übergehen, folgt

$$\alpha = 1 + \frac{1}{\alpha}. \tag{3.2}$$

Außerdem wissen wir, dass für alle n gilt $1 \leq x_n \leq 2$. Von den beiden Wurzeln der quadratischen Gleichung $X^2 - X - 1 = 0$ muss α die positive sein, also

$$\alpha = \frac{\sqrt{5}+1}{2} \approx 1,618.$$

2. Schritt. Wir zeigen $\lim x_n = \alpha$. Dabei nutzen wir aus, dass $\frac{\sqrt{5}-1}{2} = \frac{2}{\sqrt{5}+1}$. Es folgt nun:

$$\begin{aligned} \left| x_{n+1} - \frac{\sqrt{5}+1}{2} \right| &= \left| (1 + \frac{1}{x_n}) - \frac{\sqrt{5}+1}{2} \right| \\ &= \left| \frac{1}{x_n} - \frac{1}{\frac{\sqrt{5}+1}{2}} \right| = \frac{|x_n - \frac{\sqrt{5}+1}{2}|}{|x_n| \cdot |\frac{\sqrt{5}+1}{2}|} < \frac{2}{3} \left| x_n - \frac{\sqrt{5}+1}{2} \right|, \end{aligned}$$

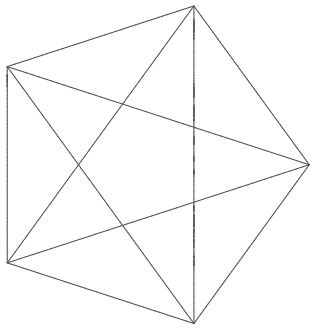
weil $x_n \geq 1$ und $\frac{\sqrt{5}+1}{2} > \frac{3}{2}$. Induktiv bekommt man

$$\left| x_{n+1} - \frac{\sqrt{5}+1}{2} \right| < \left(\frac{2}{3} \right)^n \left| x_1 - \frac{\sqrt{5}+1}{2} \right|. \tag{3.3}$$

Für $n \to \infty$ geht die rechte Seite gegen 0.

Die Zahl α, oder manchmal auch $1/\alpha$, heißt der *Goldene Schnitt*.

Bemerkung 3.16 — Der goldene Schnitt taucht geometrisch unter anderem am regulären Fünfeck auf, nämlich als Verhältnis einer Diagonalen zur Kantenlänge. Beweis?



Satz 3.17 — *Es seien (a_n), (b_n) konvergente Folgen, und es gelte $a_n \geq b_n$ für fast alle $n \in \mathbb{N}$. Dann gilt auch* $\lim a_n \geq \lim b_n$.

Beweis. Angenommen, $a := \lim a_n < b := \lim b_n$. Wir setzen $\varepsilon := b - a$. Dann gibt es $n_1, n_2 \in \mathbb{N}$ derart, dass $\forall\, n \geq n_1 : |a - a_n| < \frac{\varepsilon}{2}$ und $\forall\, n \geq n_2 : |b - b_n| < \frac{\varepsilon}{2}$. Es folgt für $n \geq \max\{n_1, n_2\}$:

$$a_n < a + \frac{\varepsilon}{2} = b - \frac{\varepsilon}{2} < b_n \leq a_n \qquad \text{Widerspruch!}$$

□

Definition 3.18 — Eine Folge (a_n) in $\mathbb{R}$ heißt uneigentlich konvergent mit Grenzwert ∞ , falls für alle $c \in \mathbb{R}$ ein $n_0 \in \mathbb{N}$ existiert derart, dass $a_n > c$ für alle $n \geq n_0$. Man schreibt $\lim_{n\to\infty} a_n = \infty$. Analog definiert man $\lim_{n\to\infty} a_n = -\infty$.

Satz 3.19 — *Es seien $(a_n), (b_n)$ (eigentlich oder uneigentlich) konvergente Folgen. Falls in den folgenden Gleichungen die rechte Seite definiert ist, so ist die Folge auf der linken Seite (uneigentlich) konvergent, und es besteht die behauptete Gleichheit:*

1) $\lim(a_n + b_n) = \lim a_n + \lim b_n$

2) $\lim(a_n - b_n) = \lim a_n - \lim b_n$

3) $\lim(a_n \cdot b_n) = \lim a_n \cdot \lim b_n$

4) $\lim(\frac{a_n}{b_n}) = \frac{\lim a_n}{\lim b_n}$, *falls* $b_n \neq 0$ *und* $\lim b_n \neq 0$.

Beweis. Zur Übung. Geht ähnlich wie bei den Sätzen 3.8 und 3.9. □

Definition 3.20 — Eine Folge (x_n) reeller Zahlen heißt monoton wachsend (streng monoton wachsend, monoton fallend, streng monoton fallend), falls für alle $n \in \mathbb{N}$ gilt: $a_n \leq a_{n+1}$ ($a_n < a_{n+1}$, $a_n \geq a_{n+1}$, bzw. $a_n > a_{n+1}$), und sie heißt monoton, wenn sie monoton wachsend oder fallend ist.

Satz 3.21 — *Jede monotone Folge konvergiert (uneigentlich) in $\overline{\mathbb{R}}$.*

Beweis. Wir nehmen an, dass (a_n) monoton wächst. Den Fall einer fallenden Folge behandelt man analog. Es sei $a := \sup\{a_n \mid n \in \mathbb{N}\} \in \overline{\mathbb{R}}$. Es sind zwei Fälle möglich:

1. Fall: $a = \infty$. Es sei dann $M > 0$ vorgegeben. Nach Definition des Supremums gibt es ein $n \in \mathbb{N}$ mit $a_n > M$. Wegen der Monotonie gilt $a_m \geq a_n > M$ für alle $m \geq n$. Das bedeutet: $\lim a_m = \infty$.

2. Fall: $a \in \mathbb{R}$. Es sei $\varepsilon > 0$ vorgegeben. Dann ist $a - \varepsilon$ keine obere Schranke mehr. Es gibt also ein $n \in \mathbb{N}$ derart, dass $a - \varepsilon < a_n \leq a$. Wegen Monotonie folgt für alle $m \geq n$, dass $a - \varepsilon < a_n \leq a_m \leq a$. Das bedeutet: $\lim a_m = a$. $\square$

Folgerung 3.22 — *Beschränkte monotone Folgen konvergieren in* $\mathbb{R}$.

Beispiel 3.23 — Kann man das Größenwachstum der Folge der mittleren Binomialkoeffizienen

$$\binom{2n}{n}: \quad 2, 6, 20, 70 \ldots$$

genauer beschreiben?

Wir betrachten die beiden Folgen

$$\begin{aligned} a_n &:= \frac{2}{1} \cdot \frac{4}{3} \cdot \ldots \cdot \frac{2n}{2n-1} \cdot \frac{1}{\sqrt{n}} \\ b_n &:= \frac{2}{1} \cdot \frac{4}{3} \cdot \ldots \cdot \frac{2n}{2n-1} \cdot \frac{1}{\sqrt{n+1}} \end{aligned}$$

und stellen fest: Für alle n gilt $a_n > b_n$, und

$$\lim \frac{a_n}{b_n} = \lim \sqrt{\frac{n+1}{n}} = 1.$$

Falls also eine der beiden Folgen konvergiert, so konvergiert auch die andere, und zwar gegen denselben Grenzwert.

Aus der Ungleichung

$$\left(\frac{a_{n+1}}{a_n}\right)^2 = \left(\frac{2n+2}{2n+1}\sqrt{\frac{n}{n+1}}\right)^2 = \frac{4n^2+4n}{4n^2+4n+1} < 1.$$

folgt, dass (a_n) monoton fällt. Analog sieht man, dass (b_n) monoton wächst. Da stets $b_n \leq a_n$, sind beide Folgen beschränkt und damit konvergent. Es sei $p := \lim a_n = \lim b_n$. Wir formen a_n etwas um:

$$a_n = \frac{2 \cdot 4 \cdots 2n}{1 \cdot 3 \cdots (2n-1)\sqrt{n}} = \frac{2^2 \cdot 4^2 \cdots (2n)^2}{(2n)!\sqrt{n}} = \frac{2^{2n}}{\binom{2n}{n}\sqrt{n}}. \tag{3.4}$$

Die Antwort auf die Frage nach dem Wachstum der Binomialkoeffizienten lautet:

$$\binom{2n}{n} \sim \frac{4^n}{p\sqrt{n}} \tag{3.5}$$

in dem Sinne, dass das Verhältnis der beiden Ausdrücke für $n \to \infty$ gegen 1 geht. Wir werden später im Beweis von Satz 9.32 sehen, dass $p = \sqrt{\pi}$.

3.2 Der Satz von Bolzano-Weierstraß

Definition 3.24 — Es sei (a_n) eine Folge und $(n_k)_k$ eine streng monoton wachsende Folge natürlicher Zahlen. Dann heißt die Folge $(a_{n_k})_k$ eine Teilfolge von (a_n).

Definition 3.25 — Es sei (a_n) ein Folge. Eine Zahl $a \in \mathbb{R}$ heißt Häufungspunkt von (a_n), wenn es zu jedem $\varepsilon > 0$ und jedem $n \in \mathbb{N}$ ein $m > n$ mit $a_m \in U_\varepsilon(a)$ gibt.

Mit anderen Worten: a ist genau dann ein Häufungspunkt von (a_n), wenn in jeder ε-Umgebung von a unendlich viele Folgenglieder liegen.

Lemma 3.26 — *Die folgenden Aussagen über eine Folge (a_n) und eine Zahl $a \in \mathbb{R}$ sind äquivalent:*

1. *a ist ein Häufungspunkt.*
2. *Es gibt eine Teilfolge (a_{n_k}) mit $\lim_{k\to\infty} a_{n_k} = a$.*

Beweis. 1. $\Rightarrow$ 2.: Da a Häufungspunkt ist, gibt es zu je zwei natürlichen Zahlen k und n eine Zahl $N(k,n) \in \mathbb{N}$ derart, dass $a_{N(k,n)} \in U_{1/k}(a)$ und $N(k,n) > n$. Wir definieren jetzt rekursiv $n_1 := N(1,1)$ und $n_{k+1} := N(k+1, n_k)$. Dann gilt stets $n_{k+1} > n_k$, d.h. die Folge $(n_k)_k$ ist streng monoton wachsend, und $a_{n_k} \in U_{1/k}(a)$, d.h. $\lim_{k\to\infty} a_{n_k} = a$.

2. $\Rightarrow$ 1.: Ist a Grenzwert einer Folge, so liegen in jeder ϵ-Umgebung alle Folgenglieder bis auf endlich viele Ausnahmen. Also ist a sicher ein Häufungspunkt. Aber jeder Häufungspunkt einer Teilfolge ist auch Häufungspunkt der Folge selbst. □

Beispiele 3.27 — 1. Falls (a_n) konvergiert, so ist $\lim a_n$ der einzige Häufungspunkt.
2. Die Folge $((-1)^n)$ hat zwei Häufungspunkte, nämlich 1 und -1.
3. Es sei $f : \mathbb{N} \to \mathbb{Q}$ eine Bijektion. Jede (!) reelle Zahl ist Häufungspunkt dieser Folge. □

Es sei nun (a_n) eine nach oben beschränkte Folge von reellen Zahlen und

$$b_n := \sup\{a_k \mid k \geq n\}, \quad \text{für alle } n \in \mathbb{N}.$$

Aus der Definition ist sofort klar, dass $b_n \geq b_{n+1}$, d.h. (b_n) fällt monoton. Daher konvergiert die Folge (b_n) eigentlich oder uneigentlich in $\overline{\mathbb{R}}$.

Definition 3.28 — Unter dem Limes superior einer Folge (a_n) reeller Zahlen verstehen wir den folgenden Punkt in $\overline{\mathbb{R}}$:

$$\overline{\lim}\, a_n = \inf\{\sup\{a_k \mid k \geq n\} \mid n \in \mathbb{N}\},$$

falls (a_n) nach oben beschränkt ist, und

$$\overline{\lim}\, a_n := \infty$$

andernfalls. Analog ist der Limes inferior gegeben durch

$$\underline{\lim}\, a_n = \sup\{\inf\{a_k \mid k \geq n\} \mid n \in \mathbb{N}\},$$

falls (a_n) nach unten beschränkt ist, und

$$\underline{\lim}\, a_n := -\infty$$

andernfalls.

Eine andere Schreibweise für den Limes superior bzw. inferior ist $\overline{\lim}\,(a_n) = \limsup(a_n)$ bzw. $\underline{\lim}\,(a_n) = \liminf(a_n)$.

Satz 3.29 — *Es sei (a_n) eine Folge reeller Zahlen. Dann gilt:*

1. $\underline{\lim}\,(a_n) = -\overline{\lim}\,(-a_n)$.
2. $\overline{\lim}\,(a_n) = \infty \Leftrightarrow (a_n)$ *ist nicht nach oben beschränkt.*
3. $\overline{\lim}\,(a_n) = -\infty \Leftrightarrow (a_n)$ *konvergiert gegen* $-\infty$.
4. *Ist* $\overline{\lim}\,(a_n) = a \in \mathbb{R}$, *so ist a ein Häufungspunkt von (a_n), und ist a' ein weiterer Häufungspunkt, so gilt* $a' \leq a$.

Analoge Aussagen gelten für den Limes inferior.

Beweis. Die Aussage 1. folgt aus der Bemerkung, dass für jede Menge $X \subset \mathbb{R}$ und $-X := \{-x | x \in \mathbb{R}\}$ gilt: $\sup(-X) = -\inf(X)$.
Die Aussage 2. ist klar nach Definition.
Zu 3. und 4.: Da wir nicht in der Situation der Aussage 2. sind, ist (a_n) nach oben beschränkt und $\overline{\lim}\,(a_n) = \lim b_n$ mit $b_n := \sup\{a_k | k \geq n\}$.

Zu 3.: Nach Annahme gilt $\lim b_n = -\infty$. Für jedes $M > 0$ existiert daher ein n mit $b_n < -M$. Nach Definition von b_n folgt $a_k \leq b_n < -M$ für alle $k \geq n$. Das zeigt $\lim a_n = -\infty$.

Zu 4.: Es sei $\varepsilon > 0$ vorgegeben. Dann existiert ein n_0 so, dass für alle $n \geq n_0$ gilt $b_n \in U_\varepsilon(a)$. Da die Folge b_n monoton fällt, gilt genauer $a \leq b_n < a + \varepsilon$ für

alle $n \geq n_0$. Nach Definition von b_n existiert ein $k_n \geq n$ mit $b_n - \varepsilon < a_{k_n} \leq b_n$. Das ergibt insgesamt:

$$a - \varepsilon \leq b_n - \varepsilon < a_{k_n} \leq b_n < a + \varepsilon,$$

also $a_{k_n} \in U_\varepsilon(a)$. Es gibt also in jeder ε-Umgebung von a unendlich viele Folgenglieder a_{k_n}, und das heißt: a ist ein Häufungspunkt.

Es sei nun $a' > a$. Wir setzen $\varepsilon := a' - a$. Nach Definition von a gibt es ein n_0 so, dass für alle $n \geq n_0$ gilt $a \leq b_n < a + \frac{\varepsilon}{2}$. Nach Definition von b_{n_0} gilt für alle $k \geq n_0$, dass $a_k \leq b_{n_0}$. Insbesondere gilt daher $a_k < a + \frac{\varepsilon}{2} = a' - \frac{\varepsilon}{2}$ für alle $k \geq n_0$. Damit kann a' kein Häufungspunkt von (a_n) sein. □

Folgerung 3.30 — *Es sei (a_n) eine beschränkte Folge. Dann sind $\overline{\lim}\,(a_n)$ und $\underline{\lim}\,(a_n)$ Häufungspunkte, und für jeden anderen Häufungspunkt a' gilt*

$$\underline{\lim}\,(a_n) \leq a' \leq \overline{\lim}\,(a_n).$$

Die Folge (a_n) konvergiert genau dann, wenn $\underline{\lim}\,(a_n) = \overline{\lim}\,(a_n)$.

Der folgende Satz fasst die bisherigen Ergebnisse zusammen:

Satz 3.31 *(Bolzano[8]-Weierstraß[9]) — Jede beschränkte Folge reeller Zahlen besitzt einen Häufungspunkt und daher eine in $\mathbb{R}$ konvergente Teilfolge.*

Definition 3.32 — Eine Folge (a_n) reeller Zahlen heißt Cauchyfolge[10] wenn es zu jedem $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$ gibt derart, dass für alle $n, m \geq n_0$ gilt $|a_n - a_m| < \varepsilon$.

Satz 3.33 — *Eine Folge reller Zahlen konvergiert genau dann, wenn sie eine Cauchyfolge ist.*

Beweis. Es sei zunächst (a_n) eine konvergente Folge mit Grenzwert $a = \lim a_n$. Zu jedem $\varepsilon > 0$ existiert daher ein n_0 so, dass für alle $n \geq n_0$ gilt: $|a - a_n| < \varepsilon/2$. Für alle $n, m \geq n_0$ folgt mit der Dreiecksungleichung:

$$|a_n - a_m| \leq |a_n - a| + |a - a_m| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

D.h. (a_n) ist ein Cauchyfolge.

Es sei nun umgekehrt (a_n) eine Cauchyfolge. Es gibt dann ein n_0 derart, dass $|a_n - a_{n_0}| < 1$ für alle $n \geq n_0$. Insbesondere gilt $|a_n| \leq |a_{n_0}| + 1$ für alle $n \geq n_0$. Die Folge (a_n) ist durch die Zahl $\max\{|a_1|, \ldots, |a_{n_0-1}|, |a_{n_0}| + 1\}$ beschränkt.

[8] Bernard Placidus Johann Nepomuk Bolzano, * 5.10.1781 † 18.12.1848.
[9] Karl Wilhelm Theodor Weierstraß, * 31.10.1815 † 19.2.1897.
[10] Augustin Louis Cauchy, * 21.8.1789 † 23.5.1857.

Nach dem Satz von Bolzano-Weierstraß besitzt (a_n) einen Häufungspunkt a. Es bleibt noch zu zeigen, dass $\lim a_n = a$. Dazu sei ein $\varepsilon > 0$ vorgegeben. Es gibt ein n_0 derart, dass $|a_n - a_m| < \varepsilon/2$ für alle $n, m \geq n_0$. Nun ist a ein Häufungspunkt. Es gibt folglich ein $m \geq n_0$ mit $|a - a_m| < \varepsilon/2$. Für alle $n \geq n_0$ weiß man jetzt, dass

$$|a_n - a| \leq |a_n - a_m| + |a - a_m| < \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

Also konvergiert (a_n) gegen a. □

§4 Die komplexen Zahlen

Wir versehen die Menge

$$\mathbb{C} := \mathbb{R} \times \mathbb{R} = \{(a, b) \mid a, b \in \mathbb{R}\}$$

mit zwei Verknüpfungen:

$$\begin{aligned} (a,b) + (a',b') &:= (a + a', b + b') \\ (a,b) \cdot (a',b') &:= (aa' - bb', ab' + a'b). \end{aligned}$$

Satz 4.1 — $(\mathbb{C}, +, \cdot)$ *ist ein Körper.*

Beweis. Kommutativität, Assoziativität und Distributivität von Addition und Multiplikation verifiziert man leicht durch Rechnung. Ebenso sieht man schnell, dass $0_{\mathbb{C}} := (0,0)$ und $1_{\mathbb{C}} := (1,0)$ Neutralelemente der Addition bzw. der Multiplikation sind und dass $(-a, -b)$ ein additives Inverses zu (a, b) ist. Es bleibt zu zeigen, dass jedes Element $(a, b) \neq (0, 0)$ ein multiplikatives Inverses besitzt. Aus $(a, b) \neq (0, 0)$ folgt $a^2 + b^2 > 0$. Also ist das Element

$$\left(\frac{a}{a^2+b^2}, \frac{-b}{a^2+b^2}\right)$$

definiert und es gilt

$$\begin{aligned} &(\frac{a}{a^2+b^2}, \frac{-b}{a^2+b^2}) \cdot (a,b) \\ &\quad = (a \cdot \frac{a}{a^2+b^2} - b \cdot \frac{-b}{a^2+b^2}, a \cdot \frac{-b}{a^2+b^2} + b\frac{a}{a^2+b^2}) \\ &\quad = (1,0). \end{aligned}$$

□

4.1 Eigenschaften der komplexen Zahlen

Durch die Zuordnung $a \mapsto (a, 0)$ wird der Körper $\mathbb{R}$ injektiv nach $\mathbb{C}$ abgebildet. Diese Abbildung ist ein Körperhomomorphismus, d.h. sie schickt die Eins und die Null aus $\mathbb{R}$ auf die entsprechenden Elemente in $\mathbb{C}$ und ist mit allen Rechenoperationen verträglich. Wir identifizieren $\mathbb{R}$ mit dem Bild dieser Abbildung, fassen also $\mathbb{R}$ immer als Unterkörper von $\mathbb{C}$ auf, und schreiben $(a, 0) = a$ für alle $a \in \mathbb{R}$. Das Element $i := (0, 1)$ heißt imaginäre Einheit. Mit seiner Hilfe können wir nun beliebige komplexe Zahlen wie folgt beschreiben:

$$(a,b) = (a,0) + (0,b) = (a,0) + (b,0) \cdot (0,1) = a + bi.$$

Außerdem gilt:

$$i^2 = (0,1)^2 = (-1 \cdot 1, 0) = -1.$$

Dies ist der entscheidende Vorteil, den die komplexen Zahlen gegenüber den reellen Zahlen haben: In $\mathbb{C}$ besitzen auch negative Zahlen Quadratwurzeln. Das bezahlen wir teuer:

Satz 4.2 — *Der Körper $\mathbb{C}$ lässt sich nicht anordnen.*

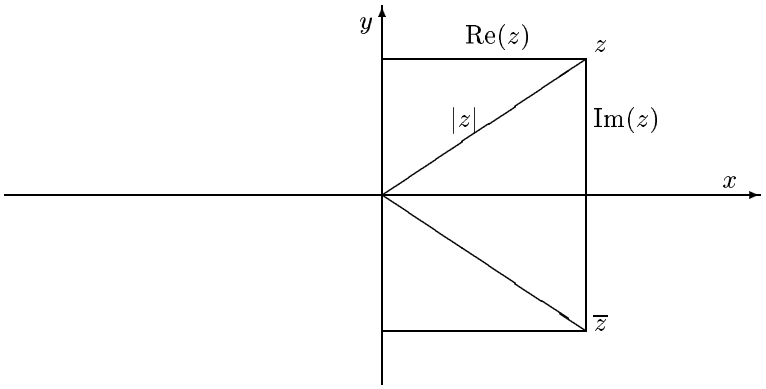
Beweis. In einem angeordneten Körper K gilt für alle $a \in K$ stets $a^2 \geq 0$, und insbesondere $-1 < 0$. In $\mathbb{C}$ haben wir dagegen $i^2 = -1$. □

Ist $z = (a, b) \in \mathbb{C}$, so heißt $\mathrm{Re}(z) := a$ der Realteil und $\mathrm{Im}(z) := b$ der Imaginärteil von z. Ferner ist $\bar{z} := a - bi$ die zu z konjugierte komplexe Zahl und

$$|z| := \sqrt{\mathrm{Re}(z)^2 + \mathrm{Im}(z)^2}$$

der Absolutbetrag von z.

Wir stellen komplexe Zahlen geometrisch durch Punkte in der sogenannten Gaußschen[11] Zahlenebene dar:



Der Zahl z entspricht ein Punkt mit den Koordinaten $(\mathrm{Re}(z), \mathrm{Im}(z))$. Der Absolutbetrag von $|z|$ ist der Abstand zum Koordinatenursprung. Die zu z konjugierte Zahl erhält man durch Spiegelung an der x-Achse. In diesem Bild entspricht die Addition zweier komplexer Zahlen der Vektoraddition.

Lemma 4.3 — *Für alle $z, w \in \mathbb{C}$ gilt:*

1. *$\overline{\overline{z}} = z$. Es gilt $\overline{z} = z$ genau dann, wenn $z \in \mathbb{R}$, und $\overline{z} = -z$ genau dann, wenn $z \in i\mathbb{R} = \{ia \mid a \in \mathbb{R}\}$.*

2. *$\overline{z + w} = \overline{z} + \overline{w}$, $\overline{z \cdot w} = \overline{z} \cdot \overline{w}$.*

[11] Carl Friedrich Gauß, ∗ 30.4.1777 † 23.2.1855.

3. $z \cdot \overline{z} = |z|^2$. *Insbesondere gilt für* $z \neq 0$, *dass* $z^{-1} = \frac{\overline{z}}{|z|^2}$.

4. $z = \mathrm{Re}(z) + \mathrm{Im}(z)i$ *und* $\overline{z} = \mathrm{Re}(z) - \mathrm{Im}(z)i$. *Umgekehrt gilt* $\mathrm{Re}(z) = \frac{1}{2}(z + \overline{z})$ *und* $\mathrm{Im}(z) = \frac{1}{2i}(z - \overline{z})$.

Beweis. Übung. □

Lemma 4.4 — *Für alle* $z, w \in \mathbb{C}$ *gilt:*

1. $|z| \geq 0$, *und* $|z| = 0$ *gilt genau dann, wenn* $z = 0$.

2. $|z \cdot w| = |z| \cdot |w|$.

3. $|z + w| \leq |z| + |w|$.

4. $\max\{|\mathrm{Re}(z)|, |\mathrm{Im}(z)|\} \leq |z| \leq |\mathrm{Re}(z)| + |\mathrm{Im}(z)|$.

Beweis. Zu 3.: Wir schreiben $z = a + bi$ und $w = x + yi$. Aus $0 \leq (ay - bx)^2$ erhält man

$$2abxy \leq a^2y^2 + b^2x^2$$

und

$$(ax+by)^2 = a^2x^2+2abxy+b^2y^2 \leq a^2x^2+a^2y^2+b^2x^2+b^2y^2 = (a^2+b^2)(x^2+y^2).$$

Durch Wurzelziehen folgt hieraus

$$ax + by \leq \sqrt{a^2 + b^2}\sqrt{x^2 + y^2}$$

und

$$\begin{aligned}(a + x)^2 + (b + y)^2 &= (a^2 + x^2 + 2ax + 2by + b^2 + y^2) \\ &\leq (a^2 + b^2) + 2\sqrt{a^2 + b^2}\sqrt{x^2 + y^2} + (x^2 + y^2) \\ &= (\sqrt{a^2 + b^2} + \sqrt{x^2 + y^2})^2.\end{aligned}$$

Das besagt aber gerade, dass $|z + w|^2 \leq (|z| + |w|)^2$. □

4.2 Komplexe Folgen

Die Gültigkeit der Dreiecksungleichung auch für den komplexen Absolutbetrag hat zur Folge, dass wir alle Aussagen und Beweise, die mit dem reellen Betrag arbeiten, aus $\mathbb{R}$ übernehmen können. Vorsicht ist allerdings geboten mit Aussagen und Beweisen, die die Kleinerrelation benutzen, da man $\mathbb{C}$, wie wir gesehen haben, nicht anordnen kann. Aus Bequemlichkeit vereinbaren wir, dass eine Formulierung wie $\varepsilon > 0$ implizit immer mit einschließt, dass ε eine reelle Zahl ist.

Definition 4.5 — 1. Für jedes $\varepsilon > 0$ und jedes $z \in \mathbb{C}$ heißt die Menge

$$U_\varepsilon(z) := \{w \in \mathbb{C} \,|\, |z - w| < \varepsilon\}$$

die ε-Umgebung von z.
2. Eine Folge (z_n) komplexer Zahlen heißt konvergent mit Grenzwert $z \in \mathbb{C}$, wenn es zu jedem $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$ gibt derart, dass für alle $n \geq n_0$ gilt: $|z - z_n| < \varepsilon$.
3. Eine komplexe Folge (z_n) heißt beschränkt, wenn es ein $M > 0$ gibt derart, dass $|z_n| < M$ für alle $n \in \mathbb{N}$.
4. Eine komplexe Folge (z_n) heißt Cauchyfolge, wenn es für jedes $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$ gibt derart, dass für alle $n, m \geq n_0$ gilt $|z_n - z_m| < \varepsilon$.

Entsprechend verallgemeinern wir die Begriffe Teilfolge und Häufungspunkt, aber nicht die Begriffe monoton wachsend, limes superior usw., die zu ihrer Definition die Anordnung von $\mathbb{R}$ benötigen.

Lemma 4.6 — *Eine komplexe Folge (z_n) konvergiert genau dann, wenn die Folgen der Real- und Imaginärteile konvergieren, und in diesem Falle gilt:*

$$\mathrm{Re}(\lim z_n) = \lim \mathrm{Re}(z_n), \quad \mathrm{Im}(\lim z_n) = \lim(\mathrm{Im} z_n).$$

Beweis. Für jedes $z \in \mathbb{C}$ gilt:

$$\max\{|\mathrm{Re}(z - z_n)|, |\mathrm{Im}(z - z_n)|\} \leq |z - z_n| \leq |\mathrm{Re}(z - z_n)| + |\mathrm{Im}(z - z_n)|.$$

Falls nun (z_n) gegen z konvergiert, falls also $|z - z_n| \longrightarrow 0$, so folgt aus der ersten Abschätzung, dass auch $\mathrm{Re}(z_n) \longrightarrow \mathrm{Re}(z)$ und $\mathrm{Im}(z_n) \longrightarrow \mathrm{Im}(z)$. Hat man umgekehrt die Konvergenzen $\lim \mathrm{Re}(z_n) = a$ und $\lim \mathrm{Im}(z_n) = b$ und setzt $z = a + bi$, so folgt aus der zweiten Abschätzung, dass $z_n \longrightarrow z$. □

Satz 4.7 — *(Bolzano-Weierstraß, komplexe Fassung)*
Jede beschränkte Folge komplexer Zahlen besitzt einen Häufungspunkt und damit eine konvergente Teilfolge.

Beweis. Ist (z_n) eine beschränkte Folge komplexer Zahlen, so sind auch die Folgen $(\mathrm{Re}(z_n))$ und $(\mathrm{Im}(z_n))$ beschränkt. Aus der reellen Version des Satzes von Bolzano-Weierstraß folgt zunächst, dass es eine Teilfolge (z_{n_k}) gibt derart, dass $\mathrm{Re}(z_{n_k})$ konvergiert. Erneute Anwendung des Satzes auf die Teilfolge liefert uns eine Teilfolge $(z_{n_{k_\ell}})$ mit der Eigenschaft, dass $(\mathrm{Im}(z_{n_{k_\ell}}))$ konvergiert. Natürlich konvergiert die Folge der Realteile immer noch. Nach dem vorigen Lemma konvergiert die Folge $(z_{n_{k_\ell}})$. □

Ebenso zeigt man:

Satz 4.8 — *Jede Cauchyfolge komplexer Zahlen konvergiert.*

4.3 Kubische Gleichungen

Bekanntlich lassen sich quadratische Gleichungen

$$x^2 + ax + b = 0$$

durch quadratische Ergänzung lösen, und man stößt auf die Lösungsformeln

$$x_1 = -\frac{a}{2} + \sqrt{\frac{a^2}{4} - b}, \quad x_2 = -\frac{a}{2} - \sqrt{\frac{a^2}{4} - b}$$

für die Lösungen. Bei positiver Diskriminante $\frac{a^2}{4} - b$ erhält man zwei verschiedene Lösungen, die bei verschwindender Diskriminante zusammenfallen, bei negativer Diskriminante findet man überhaupt keine reelle Lösung, weil negative reelle Zahlen keine Quadratwurzel besitzen. Erweitern wir die reellen Zahlen zu den komplexen, so finden wir stets zwei (manchmal zusammenfallende) Lösungen.

Historisch traten die komplexen Zahlen allerdings im Zusammenhang mit den kubischen Gleichungen ins Rampenlicht. Die Verwendung komplexer Zahlen bei quadratischen Gleichungen hat nämlich etwas rein formales. Wir können zwar Lösungen hinschreiben, aber diese sind eben nicht reell. Bei kubischen Gleichungen tritt das Phänomen auf, dass man zwischendurch komplex rechnen muss, um am Ende wieder zu einer reellen Lösungen zu gelangen. Die Verwendung komplexer Zahlen führt hier also nicht auf formale, sondern sozusagen echte Lösungen. Der erste, der kubische Gleichungen systematisch lösen konnte, scheint del Ferro[12] gewesen zu sein, der seine Lösungsmethode aber geheimhielt.

Wir beginnen mit einer allgemeinen kubischen Gleichung

$$x^3 + ax^2 + bx + c = 0.$$

Durch kubische Ergänzung können wir den quadratischen Term eliminieren:

$$\left(x + \frac{a}{3}\right)^3 = \left(\frac{a^2}{3} - b\right) x + \left(\frac{a^3}{27} - c\right).$$

Damit haben wir die ursprüngliche Gleichung auf die Form

$$y^3 = \alpha y + \beta$$

[12] Scipione del Ferro, $*$ 6.2.1465 † 29.10.1526.

mit den Abkürzungen $y = x + \frac{a}{3}$, $\alpha = \frac{a^2}{3} - b$ und $\beta = \frac{ab}{3} - \frac{2a^3}{27} - c$ reduziert. Man löst diese Gleichung mit dem Ansatz $y = u + v$. Daraus gewinnt man

$$y^3 = (u+v)^3 = u^3 + 3uv(u+v) + v^3 = 3uvy + (u^3 + v^3).$$

Durch Vergleich mit der Ausgangsgleichung erhalten wir die Gleichungen

$$3uv = \alpha, \quad u^3 + v^3 = \beta$$

zur Bestimmung von u und v. Multiplikation mit u^3 liefert:

$$\beta u^3 = u^6 + u^3 v^3 = u^6 + \left(\frac{\alpha}{3}\right)^3.$$

Dies ist eine quadratische Gleichung für u^3. Auf die übliche Weise finden wir also

$$u = \sqrt[3]{\frac{\beta}{2} + \sqrt{\left(\frac{\beta}{2}\right)^2 - \left(\frac{\alpha}{3}\right)^3}}$$

und am Ende

$$y = u + \frac{\alpha}{3u}.$$

Dies sind die nach Cardano[13] benannten Cardanischen Formeln, der sie zwar nicht selbst gefunden, aber im Jahre 1540 in seinem Buch *Ars Magna* – unter Nennung des Entdeckers Fontana[14]– veröffentlicht hat.

Beispiele 4.9 1. Die Gleichung $x^3 = 6x + 9$ führt auf

$$u = \sqrt[3]{\frac{9}{2} + \sqrt{\frac{81}{4} - 8}} = \sqrt[3]{\frac{9}{2} + \frac{7}{2}} = 2$$

und $x = 2 + \frac{6}{3 \cdot 2} = 3$.

2. Wie man sofort einsieht, hat die Gleichung $x^3 = 15x + 4$ die reelle Lösung $x = 4$. Die Cardanischen Formeln führen auf das Problem, eine dritte Wurzel aus einer komplexen Zahl zu ziehen:

$$u = \sqrt[3]{2 + \sqrt{4 - 125}} = \sqrt[3]{2 + 11i}.$$

Tatsächlich zeigt eine kurze Rechnung, dass $(2+i)^3 = 2 + 11i$, und damit findet man $x = (2+i) + \frac{5}{2+i} = (2+i) + (2-i) = 4$. In diesem Beispiel finden wir also eine reelle Lösung nur über den Umweg über die komplexen Zahlen.

Satz 4.10 — *Jede quadratische Gleichung $z^2 + az + b = 0$ mit $a, b \in \mathbb{C}$ hat eine Lösung in $\mathbb{C}$.*

[13] Girolamo Cardano, $*$ 24.9.1501 † 21.9.1576.

[14] Nicolo Fontana (Tartaglia = Der Stotterer), $*$ 1500 † 13.12.1557.

Beweis. Mit quadratischer Ergänzung reduziert man das Problem darauf, eine Lösung der reinen Gleichung $z^2 = w$ für vorgegebenes $w \in \mathbb{C}$ zu finden. Mit dem Ansatz $w = x + iy$ und $z = \xi + \eta i$, $x, y, \xi, \eta \in \mathbb{R}$, führt das auf die Gleichungen:

$$\xi^2 - \eta^2 = x, \quad 2\xi\eta = y.$$

Der Vergleich der Absolutbeträge liefert außerdem noch $\xi^2 + \eta^2 = \sqrt{x^2 + y^2}$. Damit findet man

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

und

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

Durch diese Gleichungen sind ξ und η bis auf die Vorzeichen eindeutig bestimmt. Letztere wählt man so, dass $2\xi\eta = y$ gilt. Man verifiziert nun, dass die so gewonnenen Zahlen ξ und η sich zu einer Lösung $z = \xi + i\eta$ zusammensetzen. □

§5 Reihen

5.1 Reihen

Es sei $(a_n)_{n\in\mathbb{N}_0}$ eine Folge komplexer Zahlen. Wir bezeichnen

$$s_n := \sum_{\nu=0}^{n} a_\nu$$

als die n-te Partialsumme. Die Folge (s_n) heißt unendliche Reihe und wird mit

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

oder mit

$$a_0 + a_1 + a_2 + \cdots$$

bezeichnet. Falls die Folge (s_n) gegen s konvergiert, so heißt die Reihe konvergent, und wir schreiben

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

Reihen können natürlich nicht nur bei 0 sondern auch bei jedem anderen ganzzahligen Index anfangen. Reihen, die nicht konvergieren, heißen divergent.

Beispiel 5.1 — Die Reihe $\frac{1}{1\cdot 2} + \frac{1}{2\cdot 3} + \frac{1}{3\cdot 4} + \ldots$ konvergiert und es gilt:

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

Beweis. Für alle $n \in \mathbb{N}$ gilt $\frac{1}{n(n+1)} = \frac{1}{n} - \frac{1}{n+1}$. Damit erhält man für die n-te Partialsumme:

$$s_n = \sum_{k=1}^{n} \left(\frac{1}{k} - \frac{1}{k+1}\right) = \left(1 - \frac{1}{2}\right) + \left(\frac{1}{2} - \frac{1}{3}\right) \ldots \left(\frac{1}{k} - \frac{1}{k+1}\right) = 1 - \frac{1}{k+1}.$$

Im Grenzübergang $k \longrightarrow \infty$ folgt $s_k \longrightarrow 1$. □

Beispiel 5.2 — Es sei $q \in \mathbb{C}$. Die Reihe $\sum_{n=0}^{\infty} q^n$ heißt geometrische Reihe. Die Konvergenz hängt davon ab, ob $|q| < 1$ oder$|q| \geq 1$ ist. Man findet für die n-te Partialsumme den Ausdruck

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

Ist nun $|q| < 1$, so gilt $s_n \longrightarrow \frac{1}{1-q}$, d.h.

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

Ist dagegen $|q| \geq 1$, so muss die Reihe nach dem Kriterium des folgenden Satzes divergieren.

Satz 5.3 — *Konvergiert die Reihe $\sum_{n=0}^{\infty} a_n$, so ist (a_n) eine Nullfolge.*

Beweis. Es sei $s_n = \sum_{k=0}^{n} a_k$, und es gelte $\lim(s_n) = s$. Dann gilt in gleicher Weise $\lim(s_{n-1}) = s$ und

$$\lim(a_n) = \lim(s_n - s_{n-1}) = \lim(s_n) - \lim(s_{n-1}) = s - s = 0.$$

□

Dass (a_n) eine Nullfolge ist, ist aber nur ein notwendiges Kriterium für die Konvergenz der Reihe $\sum a_n$ und keinesfalls hinreichend, wie das folgende Beispiel zeigt:

Beispiel 5.4 — Die harmonische Reihe $\sum_{n=1}^{\infty} \frac{1}{n}$ divergiert.

Beweis. Für jedes k und n mit $2^{k-1} < n \leq 2^k$ können wir den Bruch $\frac{1}{n}$ durch $\frac{1}{n} \geq \frac{1}{2^k}$ abschätzen. Das ergibt die folgende Abschätzung für die 2^k-te Partialsumme:

$$s_{2^k} = 1 + \sum_{\ell=1}^{k} \sum_{n=2^{\ell-1}-1}^{2^\ell} \frac{1}{n} \geq 1 + \sum_{\ell=1}^{k} \sum_{n=2^{\ell-1}-1}^{2^\ell} \frac{1}{2^\ell} = 1 + \sum_{\ell=1}^{k} \frac{1}{2} = 1 + \frac{k}{2}.$$

Da die Folge (s_n) monoton wächst, zeigt diese Rechnung, dass $\lim(s_n) = \infty$. □

Bemerkung 5.5 — Euler[15] hat die Divergenz der harmonischen Reihe benutzt, um einen anderen Beweis für den Satz von Euklid[16] zu geben, dass es unendlich viele Primzahlen gibt. Er argumentiert wie folgt: Es seien $p_1, \ldots, p_s$ verschiedene Primzahlen und $\mathbb{N}' \subset \mathbb{N}$ die Menge aller natürlichen Zahlen, in deren Primfaktorzerlegung nur diese Primzahlen vorkommen. Dann gilt:

$$\prod_{i=1}^{s} \frac{1}{1 - \frac{1}{p_i}} = (1 + \frac{1}{p_1} + \frac{1}{p_1^2} + \ldots) \cdot \ldots \cdot (1 + \frac{1}{p_s} + \frac{1}{p_s^2} + \ldots) = \sum_{n \in \mathbb{N}'} \frac{1}{n}.$$

Wären nun $p_1, \ldots, p_s$ *alle* Primzahlen, so wäre auch $\mathbb{N}' = \mathbb{N}$. Das Produkt $\prod_{i=1}^{s} \frac{1}{1-\frac{1}{p_i}}$ ist nun eine endliche Zahl, während die Reihe $\sum_{n \in \mathbb{N}} \frac{1}{n}$ divergiert. Aus dem Widerspruch folgt, dass es unendlich viele Primzahlen geben muss. Man kann das Argument verfeinern, um genauere Aussagen über die Primzahlverteilung zu gewinnen. Euler hat bereits die später nach Riemann[17] benannte

[15] Leonhard Euler, ∗ 15.4.1707 † 18.9.1783.
[16] Euklid, ∗ ca. 325 v. Chr. † ca. 265 v.Chr.
[17] Bernhard Riemann, ∗17.9.1826 † 20.7.1866.

Zetafunktion

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

studiert. Die Riemannsche Vermutung macht eine Aussage über die Lage der Nullstellen dieser Funktion.

Satz 5.6 — *Sind $\sum a_n$ und $\sum b_n$ konvergente Reihen und $\lambda \in \mathbb{C}$, so konvergieren auch die Reihen $\sum(a_n + b_n)$ und $\sum(\lambda a_n)$, und es gilt $\sum(a_n + b_n) = \sum a_n + \sum b_n$ und $\sum(\lambda a_n) = \lambda \sum a_n$.*

Beweis. Dies folgt sofort aus dem analogen Satz für Folgen. □

5.2 Konvergenzkriterien

Wir stellen eine Reihe von hinreichenden Kriterien für die Konvergenz einer Reihe zusammen. Vorher halten wir die einfache Beobachtung fest, dass wir bei Konvergenzbetrachtungen immer eine beliebige endliche Anzahl von Anfangstermen vernachlässigen können. Diese beeinflussen natürlich das Ergebnis der Summation, nicht aber die Frage, ob die Reihe überhaupt konvergiert.

Satz 5.7 *(Cauchykriterium) — Eine Reihe $\sum a_n$ konvergiert genau dann, wenn es zu jedem $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$ gibt derart, dass für alle $m \geq n \geq n_0$ gilt:*

$$\left| \sum_{k=n}^{m} a_k \right| < \varepsilon.$$

Beweis. Es sei $s_N = \sum_{k=0}^{N} a_k$ die N-te Partialsumme. Dann sind die folgenden Aussagen äquivalent:

1. Die Reihe $\sum a_n$ konvergiert.
2. Die Folge (s_n) konvergiert.
3. (s_n) ist eine Cauchyfolge.
4. Für jedes $\varepsilon > 0$ gibt es ein n_0 derart, dass für alle $m, n \geq n_0$ gilt: $|s_m - s_n| < \varepsilon$.

Da $s_m - s_n = \sum_{k=n+1}^{m} a_k$, folgt hieraus die Behauptung. □

Satz 5.8 *(Majorantenkriterium) — Es sei $\sum b_n$ eine konvergente Reihe mit reellen Reihengliedern und (a_n) eine Folge komplexer Zahlen. Gilt $|a_n| \leq b_n$ für alle $n \geq n_0$ für irgendein $n_0 \in \mathbb{N}$, so ist die Reihe $\sum a_n$ konvergent.*

Beweis. Es sei $\varepsilon > 0$ vorgegeben. Da $\sum b_n$ konvergiert, existiert ein n_1 derart, dass für alle $m \geq n \geq n_1$ die Abschätzung $\sum_{k=n}^{m} b_k < \varepsilon$ besteht. Für alle $m \geq n \geq \max\{n_0, n_1\}$ folgt:

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

Nach dem Cauchykriterium konvergiert die Reihe $\sum a_n$. □

In der Situation des Satzes heißt die Reihe $\sum b_n$ eine Majorante der Reihe $\sum a_n$.

Beispiel 5.9 — Für jedes $s \in \mathbb{R}$, $s \geq 2$, konvergiert die Reihe $\sum_{n=1}^{\infty} \frac{1}{n^s}$. Für jedes $n \geq 2$ gilt nämlich die Abschätzung $\frac{1}{n^s} \leq \frac{1}{n^2} \leq \frac{1}{n(n-1)}$. Da die Reihe $\sum_{n=2}^{\infty} \frac{1}{n(n-1)}$ konvergiert, folgt aus dem Majorantenkriterium, dass auch die Reihe $\sum_{n=1}^{\infty} \frac{1}{n^s}$ konvergiert.

Definition 5.10 — Eine Reihe $\sum a_n$ heißt absolut konvergent, wenn die Reihe $\sum |a_n|$ konvergiert.

Satz 5.11 — *Jede absolut konvergente Reihe konvergiert.*

Beweis.. $\sum |a_n|$ ist eine Majorante von $\sum a_n$. □

Die Umkehrung des Satzes ist falsch: Es gibt sehr wohl Reihen, die konvergieren, ohne absolut konvergent zu sein. Ein Beispiel ist die Reihe $1 - \frac{1}{2} + \frac{1}{3} - \ldots$, von der aus dem nächsten Satz folgt, dass sie konvergiert, während wir bereits gesehen haben, dass die harmonische Reihe divergiert.

Satz 5.12 *(Leibnizkriterium[18])* — *Ist (a_n) eine monoton fallende Nullfolge reeller Zahlen, so ist die Reihe $\sum_{n=1}^{\infty} (-1)^{n-1} a_n$ konvergent.*

Beweis. Wir setzen $s_n := \sum_{k=1}^{n} (-1)^{k-1} a_k$ und betrachten die Teilfolgen (s_{2n}) und (s_{2n+1}). Dann wächst die Folge (s_{2n}) monoton, denn

$$s_{2n+2} - s_{2n} = a_{2n+1} - a_{2n+2} \geq 0.$$

Ähnlich sieht man, dass (s_{2n+1}) monoton fällt. Da andererseits aber auch $0 \leq s_{2n+1} - s_{2n} = a_{2n+1} \longrightarrow 0$ für $n \longrightarrow \infty$, müssen die Folgen (s_{2n}) und (s_{2n+1}) konvergieren, und zwar gegen denselben Grenzwert. Das war zu zeigen. □

[18] Gottfried Wilhelm Leibniz, $*$ 1.7.1646 † 14.11.1716.

Satz 5.13 *(Quotientenkriterium) — Es sei (a_n) eine Folge komplexer Zahlen. Es gebe ein $q \in \mathbb{R}$, $0 < q < 1$, und ein $n_0 \in \mathbb{N}$ derart, dass für alle $n \geq n_0$ gilt: $a_n \neq 0$ und $\left|\frac{a_{n+1}}{a_n}\right| \leq q$. Dann ist die Reihe $\sum a_n$ absolut konvergent.*

Beweis. Für alle $n \geq n_0$ gilt:

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

Ebenso

$$|a_{n+2}| \leq q|a_{n+1}| \leq q^2|a_n|$$

und induktiv

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

für alle $n \geq n_0$. Da die geometrische Reihe konvergiert, konvergiert die Reihe $\sum |a_n|$ nach dem Majorantenkriterium. □

Beispiel 5.14 — Die Reihe $\sum_{n=0}^{\infty} \frac{z^n}{n!}$ heißt Exponentialreihe. Sie konvergiert sicherlich für $z = 0$. Ist $z \neq 0$, so ist $a_n = \frac{z^n}{n!} \neq 0$ und

$$\left|\frac{a_{n+1}}{a_n}\right| = \frac{|z|}{n+1} < \frac{1}{2},$$

sobald $n > 2|z|$. Nach dem Quotientenkriterium konvergiert die Exponentialreihe nunmehr für alle $z \in \mathbb{C}$.

Beispiel 5.15 — Es seien $\alpha, z \in \mathbb{C}$. Die Reihe

$$B_\alpha(z) := \sum_{n=0}^{\infty} \binom{\alpha}{n} z^n$$

heißt Binomialreihe. Zur Erinnerung: $\binom{\alpha}{0} = 1$ und

$$\binom{\alpha}{n} = \frac{\alpha \cdot (\alpha - 1) \cdot \ldots \cdot (\alpha + 1 - n)}{n \cdot (n-1) \cdot \ldots \cdot 1}.$$

Falls $z = 0$, so konvergiert die Reihe trivialerweise. Falls $\alpha \in \mathbb{N}$, so sind nur endlich viele der Binomialkoeffizienten von Null verschieden. In diesem Falle ist die Summe endlich, und wir wissen, dass

$$B_\alpha(z) = (1+z)^\alpha.$$

Wir nehmen also an, dass $z \neq 0$ und dass $\alpha \notin \mathbb{N}_0$. In diesem Falle ist $a_n = \binom{\alpha}{n} z^n \neq 0$ für alle $n \in \mathbb{N}_0$, und es gilt

$$\left|\frac{a_{n+1}}{a_n}\right| = \frac{|\alpha - n|}{n+1} \cdot |z|.$$

Für $n \longrightarrow \infty$ geht dieser Ausdruck gegen $|z|$. Falls $|z| < 1$, ist das Quotientenkriterium anwendbar, und es folgt, dass die Binomialreihe $B_\alpha(z)$ für beliebiges $\alpha \in \mathbb{C}$ konvergiert, solange nur $|z| < 1$.

Satz 5.16 *(Wurzelkriterium) — Ist (a_n) eine Folge komplexer Zahlen und gilt $\overline{\lim} \sqrt[n]{|a_n|} < 1$, so konvergiert die Reihe $\sum a_n$ absolut.*

Beweis. Es sei $p = \overline{\lim} \sqrt[n]{|a_n|}$ und q eine Zahl mit $p < q < 1$. Dann existiert ein n_0 derart, dass für alle $n \geq n_0$ gilt: $\sqrt[n]{|a_n|} < q$. Das ist gleichbedeutend mit $|a_n| < q^n$. Die Behauptung folgt also aus der Konvergenz der geometrischen Reihe und dem Majorantenkriterium. □

Das Wurzelkriterium lässt sich genauso gut umdrehen: Gilt $\underline{\lim} \sqrt[n]{|a_n|} > 1$, so muss die Reihe $\sum a_n$ divergieren, weil die Folge (a_n) nicht einmal eine Nullfolge ist.

Satz 5.17 *(Verdichtungskriterium) — Es sei (a_n) eine monoton fallende Folge reeller Zahlen. Dann konvergiert die Reihe $\sum a_n$ genau dann, wenn die Reihe $\sum 2^n a_{2^n}$ konvergiert.*

Beweis. Wir nehmen ohne Einschränkung an, dass (a_n) eine Nullfolge ist, denn andernfalls divergieren beide Reihen. Insbesondere sind also alle $a_n \geq 0$.

Für $n = 2^k$ folgt:

$$\sum_{i=1}^{n} a_i = a_1 + \sum_{\ell=1}^{k} \sum_{i=2^{\ell-1}+1}^{2^\ell} a_i \geq a_1 + \sum_{\ell=1}^{k} \sum_{i=2^{\ell-1}+1}^{2^\ell} a_{2^\ell} = a_1 + \frac{1}{2} \sum_{\ell=1}^{k} 2^\ell a_{2^\ell}.$$

Divergiert also die Reihe $\sum 2^n a_{2^n}$, so auch die Reihe $\sum a_n$. Andererseits gilt für $n = 2^k - 1$:

$$\sum_{i=1}^{n} a_i = \sum_{\ell=1}^{k} \sum_{i=2^{\ell-1}}^{2^\ell - 1} a_i \leq \sum_{\ell=1}^{k} \sum_{i=2^{\ell-1}}^{2^\ell - 1} a_{2^{\ell-1}} = \sum_{\ell=1}^{k} 2^{\ell-1} a_{2^{\ell-1}}.$$

Konvergiert also die Reihe $\sum 2^n a_{2^n}$, so auch die Reihe $\sum a_n$. □

5.3 g-adische Entwicklung

Es sei $g \in \mathbb{N}$, $g \geq 2$.

Satz 5.18 *— (Existenz einer g-adischen Entwicklung)*
Zu jeder reellen Zahl $x \geq 0$ gibt es ein $N \in \mathbb{N}_0$ und $a_n \in \{0, 1, \ldots, g-1\}$ für alle $n \geq -N$ derart, dass

$$x = \sum_{n=-N}^{\infty} \frac{a_n}{g^n}.$$

Beweis. Es sei $x \in \mathbb{R}_{\geq 0}$ vorgegeben. Da $g > 1$, existiert ein $N \in \mathbb{N}_0$ mit $g^{N+1} > x$. Wir setzen $x_{-N-1} := x$ und konstruieren induktiv für alle $n \geq -N$ reelle Zahlen x_n mit $0 \leq x_n < g^{-n}$ und *Ziffern* $a_n \in \{0, 1, \ldots, g-1\}$ wie folgt:

Es sei $n \geq -N$ und x_{n-1} bereits konstruiert, so dass also $0 \leq x_{n-1} g^n < g$. Mit $a_n := \lfloor x_{n-1} g^n \rfloor$ gilt dann

$$0 \leq g^n x_{n-1} - a_n < 1.$$

Und für $x_n := x_{n-1} - \frac{a_n}{g^n}$ besteht die Beziehung $0 \leq x_n < g^{-n}$.

Diese Konstruktion liefert also eine beschränkte Folge (a_n) und eine Nullfolge x_n. Folglich konvergiert die Reihe $\sum_{n=-N}^{\infty} \frac{a_n}{g^n}$, und zwar gegen x, denn

$$x - \sum_{\nu=-N}^{n} \frac{a_\nu}{g^\nu} = x_n \longrightarrow 0 \quad \text{für } n \longrightarrow \infty.$$

□

Diese Reihendarstellung heißt g-adische Entwicklung von x. Wir schreiben statt $x = \sum_{n=-N}^{\infty} \frac{a_n}{g^n}$ in der Regel

$$x = a_{-N} a_{-N+1} \cdots a_{-1} a_0, a_1 a_2 a_3 \cdots,$$

wobei die Zahl g aus dem Kontext klar sein muss. Außerdem werden führende Nullen weggelassen. Diese Darstellung ist eindeutig nur bis auf die aus dem elementaren Rechnen vertraute Tatsache, dass

$$4,376999999\cdots = 4,377000000\cdots$$

Satz 5.19 — *(Eindeutigkeit der g-adischen Entwicklung)*
Sind $x = \sum_{n=-N}^{\infty} \frac{a_n}{g^n} = \sum_{n=-M}^{\infty} \frac{b_n}{g^n}$ zwei g-adische Entwicklungen von x, so gilt entweder $a_\nu = b_\nu$ für alle $\nu \in \mathbb{Z}$, oder es gibt ein $n \in \mathbb{Z}$ mit $a_\nu = b_\nu$ für alle $\nu < n$ und $a_n \neq b_n$. Ist dann ohne Einschränkung $a_n < b_n$, so gilt weiter $b_n = a_n + 1$ und $a_\nu = g - 1$, $b_\nu = 0$ für alle $\nu > n$.

Beweis. Indem wir mit führenden Nullen anfüllen, können wir annehmen, dass $N = M$. Falls $a_\nu = b_\nu$ für alle $\nu \geq -N$, so ist nichts zu zeigen. Es sei also n die kleinste ganze Zahl mit $a_n \neq b_n$. Wie im Satz können wir wegen der Symmetrie zwischen den beiden Reihen annehmen, dass $b_n > a_n$. Wir betrachten die Differenz der beiden Reihen und multiplizieren mit g^n. Mit der Bezeichnung $c_\nu := b_\nu - a_\nu$ folgt:

$$0 = c_n + \sum_{i=1}^{\infty} \frac{c_{n+i}}{g^i}.$$

Nach Konstruktion gilt $c_{n+i} \in \{1-g, \ldots, -1, 0, 1, \ldots, g-1\}$ für alle $i \in \mathbb{N}$ und $c_n \in \{1, \ldots, g-1\}$. Damit haben wir die folgenden Abschätzungen:

$$c_n = -\sum_{i=1}^{\infty} \frac{c_{n+i}}{g^i} \leq \sum_{i=1}^{\infty} \frac{g-1}{g^i} = 1 \leq c_n.$$

Da die linke und die rechte Seite gleich sind, muss überall Gleichheit gelten, und das ist gleichbedeutend damit, dass $c_n = 1$ und $c_{n+i} = 1 - g$ für alle $i \in \mathbb{N}$. Für die a_ν und b_ν bedeutet dies: $b_n = a_n + 1$, und $a_{n+i} = g - 1$, $b_{n+i} = 0$ für alle $i \in \mathbb{N}$. □

Bemerkung 5.20 — Im Beweis des Existenzsatzes ist genau genommen ein präziser Algorithmus angegeben, wie man ausgehend von einer nichtnegativen reellen Zahl x die Ziffern a_n der g-adischen Entwicklung berechnet. Daraus kann man noch die genauere Aussage herleiten, dass rationale Zahlen eine periodische Entwicklung besitzen:

Ist x eine rationale Zahl, so bleiben auch die durch den Algorithmus konstruierten Zahlen x_n alle rational, genauer gilt für alle $n \geq -N$:

$$g^n x_n = g(g^{n-1} x_{n-1}) - a_n.$$

Lässt sich nun $g^{n-1}x_{n-1}$ als Bruch mit Nenner q schreiben, so gilt das für $g^n x_n$ auch, denn g und a_n sind ganze Zahlen. Nun gilt einerseits $0 \leq x_n g^n < 1$; andererseits liegen zwischen 0 und 1 nur endlich viele Brüche mit festem Nenner q. Deshalb muss nach spätestens q Schritten eine der Zahlen $x_n g^n$ ein zweites Mal auftreten. Von jetzt an wiederholen sich alle Rechnungen, und die Entwicklung wird periodisch.

Umgekehrt ist es eine leichte Übung zu zeigen, dass $x = \sum_{n=-N}^{\infty} \frac{a_n}{g^n}$ eine rationale Zahl ist, wenn es ein n und ein $p \geq 1$ gibt mit $a_\nu = a_{\nu+p}$ für alle $\nu \geq n$.

5.4 Umordnungssätze

Die zentrale Frage dieses Abschnitts lautet: *Darf man Reihen umordnen, oder hat die Reihenfolge der Reihenglieder Einfluss auf die Konvergenz oder den Wert der Summe?* Das folgende Beispiel von Dirichlet[19] zeigt, dass man im allgemeinen sehr vorsichtig sein muss:

Beispiel 5.21 — Nach dem Leibnizkriterium konvergiert die Reihe

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \ldots =: \ell.$$

[19] Johann Peter Gustav Lejeune Dirichlet, ∗ 13.2.1805 † 5.5.1859.

Dann gilt auch

$$\begin{aligned}\frac{\ell}{2} &= \frac{1}{2}-\frac{1}{4}+\frac{1}{6}-\frac{1}{8}+\dots \\ &= 0+\frac{1}{2}+0-\frac{1}{4}+0+\frac{1}{6}+0-\frac{1}{8}+0+\dots\end{aligned}$$

Addieren wir die beiden Reihen von ℓ und $\ell/2$, so heben sich die Hälfte der Terme mit geradem Nenner weg und wir erhalten:

$$\frac{3}{2}\ell = 1+\frac{1}{3}-\frac{1}{2}+\frac{1}{5}+\frac{1}{7}-\frac{1}{4}+\dots$$

Aber rechts stehen jetzt dieselben Reihenglieder wie in der ursprünglichen Reihe von ℓ, nur in anderer Reihenfolge!

Das Beispiel zeigt, dass man im Allgemeinen Reihen eben nicht ohne schwerwiegende Folgen umordnen darf. Um zu untersuchen, unter welchen Bedingungen man umordnen darf, führen wir den Begriff der summierbaren Familie ein.

Eine Abbildung $a : M \to N$ nennt man auch eine von M indizierte Familie von Elementen in N und schreibt a_m statt $a(m)$ und $(a_m)_{m\in M}$ anstelle von $a : M \to N$. M heißt die Indexmenge der Familie. Eine von $\mathbb{N}$ indizierte Familie ist also dasselbe wie eine Folge.

Im Folgenden betrachten wir Familien $(a_i)_{i\in I}$ von komplexen Zahlen. Für jede endliche Teilmenge $J \subset I$ heißt

$$s_J := \sum_{i\in J} a_i$$

eine Partialsumme der Familie $(a_i)_{i\in I}$. Eigentlich müssten wir in der Definition der Partialsumme genauer festlegen, in welcher Reihenfolge wir die einzelnen Summanden addieren wollen. Aber da die Indexmenge J endlich ist und da in $\mathbb{C}$ das Kommutativ- und das Assoziativgesetz gelten, kommt es darauf nicht an.

Definition 5.22 — Eine Familie $(a_i)_{i\in I}$ komplexer Zahlen heißt summierbar mit Summe s, falls für jedes $\varepsilon > 0$ eine endliche Teilmenge $I_0 \subset I$ existiert derart, dass für jede endliche Menge J, $I_0 \subset J \subset I$, gilt: $|s - s_J| < \varepsilon$. Wir schreiben dann

$$s = \sum_{i\in I} a_i.$$

Der folgende Satz gibt ein schlagkräftiges Kriterium für Summierbarkeit. Um ihn bequem formulieren zu können, setzen wir noch $\tilde{s}_J := \sum_{i\in J} |a_i|$ für jede endliche Teilmenge $J \subset I$.

Satz 5.23 — *Eine Familie $(a_i)_{i\in I}$ ist genau dann summierbar, wenn die Menge $\{\tilde{s}_J \mid J \subset I \text{ endlich}\}$ beschränkt ist.*

Beweis. 1. Wir beginnen mit einer Vorüberlegung: Ist die Familie $(a_i)_{i\in I}$ summierbar, so gilt dies auch für die Familien $(\mathrm{Re}(a_i))_{i\in I}$ und $(\mathrm{Im}(a_i))_{i\in I}$ und umgekehrt. Da ferner $|z| \leq |\mathrm{Re}(z)| + |\mathrm{Im}(z)|$, können wir im folgenden ohne Einschränkung annehmen, dass die Familie $(a_i)_{i\in I}$ reell ist.

2. Es sei zunächst $(a_i)_{i\in I}$ eine summierbare Familie reeller Zahlen und $s := \sum_{i\in I} a_i$. Es existiert eine endliche Teilmenge $I_0 \subset I$ derart, dass für jede endliche Menge J, $I_0 \subset J \subset I$, gilt: $|s - s_J| < 1$. Es sei nun $K \subset I$ eine beliebige endliche Teilmenge. Dann gilt $K = (K \cup I_0) \setminus (I_0 \setminus K)$ und somit $s_K = s_{K\cup I_0} - s_{I_0\setminus K}$. Damit erhalten wir die Abschätzung

$$|s_K| = |(s_{K\cup I_0} - s) + (s - s_{I_0}) + (s_{I_0} - s_{I_0\setminus K})| < 1 + 1 + |s_{I_0\cap K}| \leq 2 + \tilde{s}_{I_0}.$$

Wendet man diese Abschätzung speziell auf die Teilmengen $K_+ := \{i \in K \mid a_i \geq 0\}$ und $K_- := \{i \in K \mid a_i < 0\}$ von K an, so folgt:

$$\tilde{s}_K = s_{K_+} - s_{K_-} \leq 4 + 2\tilde{s}_{I_0},$$

und das ist die gesuchte Beschränktheitsaussage.

3. Es sei nun umgekehrt die Menge $\{\tilde{s}_K \mid K \subset I \text{ endlich}\}$ beschränkt und $\tilde{s}$ ihr Supremum. Zu jedem $n \in \mathbb{N}$ existiert dann eine endliche Teilmenge $I_n \subset I$ mit $\tilde{s} - \frac{1}{n} < \tilde{s}_{I_n}$. Für jede endliche Teilmenge $J \subset I$ mit $J \cap I_n = \emptyset$ folgt nun:

$$\tilde{s}_{I_n} + \tilde{s}_J = \tilde{s}_{I_n+J} \leq \tilde{s} < \tilde{s}_{I_n} + \frac{1}{n}$$

und somit

$$|s_J| \leq \tilde{s}_J < \frac{1}{n}.$$

Für alle $n.m \in \mathbb{N}$ folgt daher:

$$|s_{I_n} - s_{I_m}| = |s_{I_n\setminus I_m} - s_{I_m\setminus I_n}| < \frac{1}{m} + \frac{1}{n} \leq \frac{2}{\min\{m,n\}}.$$

Das zeigt, dass die Folge $(s_{I_n})_n \in \mathbb{N}$ eine Cauchyfolge ist. Es sei $s = \lim(s_{I_n})$. Wir zeigen jetzt: $s = \sum_{i\in I} a_i$.

Dazu sei $\varepsilon > 0$ vorgegeben und $n_0 \in \mathbb{N}$ so gewählt, dass für alle $n \geq n_0$ gilt: $|s - s_{I_n}| < \frac{\varepsilon}{2}$. Wir wählen $m_0 \in \mathbb{N}$ mit $m_0 > \max\{n_0, \frac{2}{\varepsilon}\}$. Jetzt gilt für jede endliche Menge J, $I_{m_0} \subset J \subset I$:

$$|s - s_J| = |(s - s_{I_{m_0}}) - s_{J\setminus I_{m_0}}| < \frac{\varepsilon}{2} + \frac{1}{m_0} \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

□

Satz 5.24 — *Eine Familie $(a_n)_{n\in\mathbb{N}}$ ist genau dann summierbar, wenn die Reihe $\sum_{n=1}^{\infty} a_n$ absolut konvergiert, und in diesem Falle gilt $\sum_{n=1}^{\infty} a_n = \sum_{n\in\mathbb{N}} a_n$.*

Beweis.— Ist die Reihe $\sum a_n$ absolut konvergent, so folgt für jede endliche Teilmenge $J \subset \mathbb{N}$ mit $\max J =: N$, dass

$$\tilde{s}_J = \sum_{i \in J} |a_i| \leq \sum_{n=1}^{N} |a_n| \leq \sum_{n=1}^{\infty} |a_n|.$$

Somit ist die Familie summierbar.

Es sei nun umgekehrt die Familie summierbar und $s = \sum_{n \in \mathbb{N}} a_n$. Dann bleibt zunächst einmal $\sum_{\nu=1}^{n} |a_\nu| = \sum_{\nu \in \{1,2,\ldots,n\}} |a_\nu|$ für alle n beschränkt, d.h. die Reihe konvergiert absolut.

Es bleibt zu zeigen, dass $\sum_{n=1}^{\infty} a_n = s$. Ist dazu $\varepsilon > 0$ vorgegeben, so existiert eine endliche Teilmenge $I_0 \subset \mathbb{N}$ derart, dass für jede endliche Menge J, $I_0 \subset J \subset \mathbb{N}$, gilt: $|s - s_J| < \varepsilon$. Das wenden wir speziell auf die Menge $J := \{1, 2, \ldots, N\}$ an, wobei N eine beliebige natürliche Zahl mit $N \geq \max\{I_0\}$ ist. Es folgt:

$$|s - \sum_{n=1}^{N} a_n| = |s - s_J| < \varepsilon.$$

□

Satz 5.25 *(Umordnungssatz) — Ist $\sum_{n=1}^{\infty} a_n$ eine absolut konvergente Reihe und ist $\sigma : \mathbb{N} \to \mathbb{N}$ eine beliebige Bijektion, dann ist die Reihe $\sum_{n=1}^{\infty} a_{\sigma(n)}$ ebenfalls absolut konvergent, und es gilt $\sum_{n=1}^{\infty} a_{\sigma(n)} = \sum_{n=1}^{\infty} a_n$.*

Beweis. Die beiden Reihen sind nach Satz 5.24 genau dann absolut konvergent, wenn die entsprechenden Familien summierbar sind. Aber für eine beliebig indizierte Familie $(a_i)_{i \in I}$ ist es völlig klar, dass eine Bijektion der Indexmenge in sich keinen Einfluss auf die Summierbarkeit und die Summe hat. □

Satz 5.26 *— Ist $(a_i)_{i \in I}$ summierbar, dann ist die Menge $I' = \{i \in I \,|\, a_i \neq 0\}$ abzählbar.*

Beweis. Es gibt ein $M \in \mathbb{R}$ mit der Eigenschaft, dass für jede endliche Teilmenge $J \subset I$ gilt: $\sum_{j \in J} |a_j| < M$. Wir definieren nun $I_0 := \{i \in I \,|\, |a_i| \geq 1\}$ und $I_k := \{i \in I \,|\, \frac{1}{k} > |a_i| \geq \frac{1}{k+1}\}$ für jedes $k \in \mathbb{N}$. Dann ist offensichtlich I' die disjunkte Vereinigung der Mengen I_k, $k \in \mathbb{N}_0$. Ist $J \subset I_k$ endlich, so folgt nun

$$\frac{|J|}{k+1} \leq \sum_{i \in J} |a_i| < M$$

und

$$|J| \leq M(k+1).$$

Da aber J eine beliebige endliche Teilmenge von I_k war, muss jede der Mengen I_k, $k \in \mathbb{N}_0$, endlich sein. Damit ist I' eine abzählbare Vereinigung von endlichen Mengen und somit selbst abzählbar. □

Geht man die letzten Sätze noch einmal durch, so könnte der Eindruck entstehen, der Begriff der Summierbarkeit sei völlig überflüssig: Der letzte Satz besagt, dass jede summierbare Familie bis auf irrelevante Glieder, die Null sind, abzählbar sein muss, sich also genauso gut durch $\mathbb{N}$ indizieren ließe und damit eine Folge wäre, und nach Satz 5.24 ist eine durch $\mathbb{N}$ indizierte Familie genau dann summierbar, wenn die Reihe absolut konvergiert. Dem ist zunächst nicht zu widersprechen. Allerdings gibt uns der Begriff der summierbaren Familie eine große Freiheit, die Indizierung der Familie nach Belieben und nach Bedarf zu ändern, und dass macht gewisse Schlussweisen durchsichtiger, als sie es bei sturem Festhalten an der Reihennotation wären. Das sah man schon beim Umordnungssatz und wird beim folgenden Satz und den daraus abgeleiteten Aussagen von Neuem deutlich.

Satz 5.27 *(Großer Umordnungssatz) — Es sei $(a_i)_{i\in I}$ eine summierbare Familie komplexer Zahlen und $(I_\mu)_{\mu\in M}$ eine Familie von Teilmengen von I mit $I = \coprod_{\mu\in M} I_\mu$. Dann ist für jedes $\mu \in M$ die Familie $(a_i)_{i\in I_\mu}$ summierbar. Ist $s_\mu = \sum_{i\in I_\mu} a_i$, so ist auch die Familie $(s_\mu)_{\mu\in M}$ summmierbar, und es gilt*

$$\sum_{i\in I} a_i = \sum_{\mu\in M} s_\mu = \sum_{\mu\in M} \left(\sum_{i\in I} a_i \right).$$

Beweis. 1. Schritt: Summmierbarkeit. Es ist klar, dass für jedes $\mu \in M$ die Familie $(a_i)_{i\in I_\mu}$ summierbar ist; es sei s_μ die Summe. Wir setzen noch $\tilde{s} := \sum_{i\in I} |a_i|$.

Es ist zu zeigen, dass es eine obere Schranke für die Menge

$$\{\sum_{\mu\in L} |s_\mu| \,|\, L \subset M \text{ endlich}\}$$

gibt. Es sei dazu $\tilde{s}_\mu := \sum_{i\in I_\mu} |a_i|$ für jedes $\mu \in M$. Ist $L \subset M$ eine beliebige nicht leere endliche Menge, so existiert zu jedem $\mu \in L$ eine endliche Teilmenge $I'_\mu \subset I_\mu$ derart, dass

$$\tilde{s}_\mu - \frac{1}{|L|} < \sum_{i\in I'_\mu} |a_i| \leq \tilde{s}_\mu.$$

Daraus folgt:

$$\sum_{\mu\in L} |s_\mu| \leq \sum_{\mu\in L} \tilde{s}_\mu \leq \sum_{\mu\in L} \left(\sum_{i\in I'_\mu} |a_i| + \frac{1}{|L|} \right) \leq 1 + \sum_{i\in \bigcup_{\mu\in L} I'_\mu} |a_i| \leq 1 + \tilde{s}.$$

2. Schritt: Es ist zu zeigen, dass die beiden Summen $s := \sum_{i\in I} a_i$ und $s' := \sum_{\mu\in M} s_\mu$ gleich sind.

Es sei dazu ein $\varepsilon > 0$ vorgegeben. Falls I leer ist, ist die Behauptung trivial. Wir nehmen also an, dass $I \neq \emptyset$.

Die Summierbarkeit der Familie $(a_i)_{i\in I}$ bedeutet, dass es eine endliche Menge $I_0 \subset I$ gibt derart, dass für jede endliche Menge I' mit $I_0 \subset I' \subset I$ gilt: $|s - \sum_{i\in I'} a_i| < \varepsilon$. Für jedes $\mu \in M$ sei $I_{0\mu} := I_0 \cap I_\mu$, ferner $M_0 := \{\mu \in M \mid I_{0\mu} \neq \emptyset\}$. Die Mengen $I_{0\mu}$, $\mu \in M$, und M_0 sind endlich.

Weil die Familie $(s_\mu)_{\mu\in M}$ summierbar ist, gibt es eine endliche Teilmenge $M_1 \subset M$ derart, dass für jede endliche Menge M' mit $M_1 \subset M' \subset M$ gilt: $|s' - \sum_{\mu\in M'} s_\mu| < \varepsilon$. Wir wählen speziell $M' = M_0 \cup M_1$. Ohne Einschränkung können wir annehmen, dass $M' \neq \emptyset$.

Weiter ist nach Voraussetzung für jedes $\mu \in M'$ die Familie $(a_i)_{i\in I_\mu}$ summierbar. Es gibt daher eine endliche Teilmenge $I_{\mu 0} \subset I_\mu$ derart, dass für jede endliche Menge I'_μ mit $I_{\mu 0} \subset I'_\mu \subset I_\mu$ gilt:

$$|s_\mu - \sum_{i\in I'_\mu} a_i| < \frac{\varepsilon}{|M'|}.$$

Wir wählen nun speziell $I'_\mu := I_{\mu 0} \cup I_{0\mu}$ und $I' := \cup_{\mu\in M'} I'_\mu$ und sammeln alle bisherigen Abschätzungen zusammen:

Wir haben einerseits, weil $I_0 \subset I'$:

$$|s - \sum_{i\in I'} a_i| < \varepsilon.$$

Andererseits gilt mit der Dreiecksungleichung:

$$\begin{aligned} |s' - \sum_{i\in I'} a_i| &= |s' - \sum_{\mu\in M'} \sum_{i\in I'_\mu} a_i| \\ &\leq |s' - \sum_{\mu\in M'} s_\mu| + \sum_{\mu\in M'} |s_\mu - \sum_{i\in I'_\mu} a_i| \\ &\leq \varepsilon + \sum_{\mu\in M'} \frac{\varepsilon}{|M'|} = 2\varepsilon. \end{aligned}$$

Zusammengenommen folgt hieraus, dass $|s - s'| < 3\varepsilon$. Da ε aber beliebig war, folgt $s = s'$, was zu zeigen war. □

Satz 5.28 *(Doppelreihensatz)— Es sei $(a_{mn})_{(m,n)\in\mathbb{N}_0^2}$ eine summierbare Familie. Dann ist für jedes $m \in \mathbb{N}_0$ auch die Familie $(a_{mn})_{n\in\mathbb{N}_0}$ summierbar. Ferner ist die Familie $(\sum_{n\in\mathbb{N}_0} a_{mn})_{m\in\mathbb{N}_0}$ summierbar und es gilt:*

$$\sum_{(m,n)\in\mathbb{N}_0^2} a_{mn} = \sum_{m\in\mathbb{N}_0} \left(\sum_{n\in\mathbb{N}_0} a_{mn} \right).$$

Beweis. Das folgt sofort aus dem großen Umordnungssatz, wenn man die Menge $\mathbb{N}_0 \times \mathbb{N}_0$ in die Teilmengen $\{m\} \times \mathbb{N}_0$ zerlegt. □

Definition 5.29 — Es seien $\sum_{n=0}^{\infty} a_n$ und $\sum_{n=0}^{\infty} b_n$ zwei Reihen. Ist $c_n = \sum_{k=0}^{n} a_k b_{n-k}$ so heißt die Reihe $\sum_{n=0}^{\infty} c_n$ das Cauchyprodukt der beiden Reihen $\sum_{n=0}^{\infty} a_n$ und $\sum_{n=0}^{\infty} b_n$.

Im allgemeinen ist nicht klar, ob das Cauchyprodukt zweier konvergenter Reihen wieder konvergiert und ob in diesem Falle der Wert der Reihe gleich dem Produkt der Werte der Ausgangsreihen ist (vgl. Folgerung 8.29). Der folgende Satz gibt ein hinreichendes Kriterium.

Satz 5.30 — *Sind $\sum_{n=0}^{\infty} a_n$ und $\sum_{n=0}^{\infty} b_n$ absolut konvergente Reihen, so ist auch ihr Cauchyprodukt $\sum_{n=0}^{\infty} c_n$ absolut konvergent und es gilt*

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

Beweis. Wir betrachten die Familie $(a_m b_n)_{(m,n)\in\mathbb{N}_0}$.

1. Schritt: Die Familie $(a_m b_n)_{(m,n)\in\mathbb{N}_0}$ ist summierbar.

Ist nämlich $J \subset \mathbb{N}_0 \times \mathbb{N}_0$ eine beliebige endliche Teilmenge, so gibt es $m_0, n_0 \subset \mathbb{N}_0$ so, dass $J \subset \{0, 1, \ldots, m_0\} \times \{0, 1, \ldots, n_0\}$. Nun folgt:

$$\sum_{(m,n)\in J} |a_m b_n| \leq \sum_{m=0}^{m_0} |a_m| \cdot \sum_{n=0}^{n_0} |b_n| \leq \sum_{n=0}^{\infty} |a_m| \cdot \sum_{n=0}^{\infty} |b_n|.$$

Das bedeutet, dass wir für alle J eine einheitliche Schranke für die Summe $\sum_{(m,n)\in J} |a_m b_n|$ gefunden haben. Daraus folgt die Summierbarkeit.

2. Schritt. Nach dem Doppelreihensatz gilt:

$$\begin{aligned}
\sum_{(m,n)\in\mathbb{N}_0} a_m b_n &= \sum_{m\in\mathbb{N}_0} \left(\sum_{n\in\mathbb{N}_0} a_m b_n \right) \\
&= \sum_{m\in\mathbb{N}_0} a_m \left(\sum_{n\in\mathbb{N}_0} b_n \right) \\
&= \left(\sum_{n\in\mathbb{N}_0} b_n \right) \cdot \left(\sum_{m\in\mathbb{N}_0} a_m \right)
\end{aligned}$$

3. Schritt. Wir zerlegen die Menge $\mathbb{N}_0 \times \mathbb{N}_0$ in die Teilmengen

$$I_\mu := \{(m,n) \,|\, m + n = \mu\}, \quad \mu \in \mathbb{N}_0 \,.$$

Dann ist

$$\sum_{(m,n)\in I_\mu} a_m b_n = \sum_{k=0}^{\mu} a_k b_{\mu-k} =: c_\mu.$$

Nach dem großen Umordnungssatz ist $\sum_{\mu\in\mathbb{N}_0} c_\mu$ summierbar, also das Cauchyprodukt $\sum_{\mu=0}^{\infty} c_\mu$ absolut konvergent, und es gilt:

$$\left(\sum_{m=0}^{\infty} a_m\right)\cdot\left(\sum_{n=0}^{\infty} b_n\right) = \sum_{(m,n)\in\mathbb{N}_0^2} a_m b_n = \sum_{\mu=0}^{\infty} c_\mu.$$

□

5.5 Potenzreihen

Es sei (a_n) eine Folge komplexer Zahlen. Eine Reihe der Form

$$\sum_{n=0}^{\infty} a_n z^n$$

mit $z\in\mathbb{C}$ heißt Potenzreihe. Die Konvergenz einer solchen Reihe ist dann nicht allein durch die Koeffizienten a_n bestimmt, sondern hängt von z ab.

Satz 5.31 — *Konvergiert die Reihe $\sum a_n z^n$ für ein $z\in\mathbb{C}$, so konvergiert die Reihe $\sum a_n w^n$ absolut für alle $w\in\mathbb{C}$ mit $|w|<|z|$.*

Beweis. Ist $\sum_{n=0}^{\infty} a_n z^n$ konvergent, so ist die Folge $(a_n z^n)$ eine Nullfolge, also insbesondere beschränkt. Es sei $M=\sup\{|a_n z^n|\,|\,n\in\mathbb{N}_0\}$. Für jede komplexe Zahl w mit $|w|<|z|$ und jedes $n\in\mathbb{N}_0$ folgt dann:

$$|a_n w^n| = |a_n z^n|\cdot\left|\frac{w}{z}\right|^n \le M\cdot\left|\frac{w}{z}\right|^n.$$

Nach dem Majorantenkriterium konvergiert $\sum a_n w^n$ absolut. □

Definition 5.32 — $\sup\{|z|\,|\,\sum a_n z^n \text{ konvergiert}\}\in\overline{\mathbb{R}}$ heißt Konvergenzradius der Potenzreihe $\sum a_n z^n$.

Nach Definition kann der Konvergenzradius R einer Potenzreihe $\sum a_n z^n$ den Wert $R=0$, $R\in\mathbb{R}_{>0}$ oder $R=\infty$ annehmen. Falls $R=0$, so konvergiert die Reihe $\sum a_n z^n$ nur für $z=0$. Falls $0<R<\infty$, so ist die Reihe $\sum a_n z^n$ für jedes z in der offenen Kreisscheibe $U_R(0)$ absolut konvergent und divergiert für jedes z mit $|z|>R$. Über die Konvergenz für solche z mit $|z|=R$ kann man keine allgemeingültigen Aussagen treffen, sondern muss diese Frage für jede Potenzreihe und jedes fragliche z getrennt untersuchen. Falls $R=\infty$, so konvergiert die Reihe $\sum a_n z^n$ absolut für alle $z\in\mathbb{C}$.

Beispiel 5.33 — Wir haben schon früher gesehen, dass die Exponentialreihe $\sum_{n=0}^{\infty}\frac{z^n}{n!}$ für jede komplexe Zahl z absolut konvergiert. Der Konvergenzradius der Exponentialreihe ist also ∞.

Beispiel 5.34 — Die geometrische Reihe $\sum_{n=0}^{\infty} z^n$ hat den Konvergenzradius 1. Denn einerseits ist die Reihe für alle $z \in \mathbb{C}$ mit $|z| < 1$ absolut konvergent, andererseits für $z = 1$ sicher divergent.

Beispiel 5.35 — Die Arkustangensreihe

$$z - \frac{z^3}{3} + \frac{z^5}{5} - \frac{z^7}{7} + \ldots = \sum_{n=0}^{\infty} (-1)^n \frac{z^{2n+1}}{2n+1}$$

hat Konvergenzradius 1, denn sie konvergiert nach dem Leibnizkriterium für $z = 1$, divergiert aber für $z = i$, weil für jedes $n \in \mathbb{N}$ gilt:

$$|\sum_{\nu=0}^{n} (-1)^\nu \frac{i^{2\nu+1}}{2\nu+1}| = |i \sum_{\nu=0}^{n} \frac{1}{2\nu+1}| \geq \frac{1}{2} \sum_{\nu=0}^{n} \frac{1}{\nu+1},$$

und die harmonische Reihe divergiert!

Satz 5.36 *(Konvergenzradius nach Cauchy-Hadamard[20])* — *Es sei (a_n) eine Folge komplexer Zahlen und $L := \overline{\lim} \sqrt[n]{|a_n|}$. Dann ist der Konvergenzradius der Potenzreihe $\sum a_n z^n$ gegeben durch L^{-1}, falls $L \neq 0$, und ist ∞, falls $L = 0$.*

Beweis. Es sei $R := L^{-1}$, falls $L \neq 0$, und $R := \infty$, falls $L = 0$. Gilt $|z| < R$ für eine komplexe Zahl z, so ist $\overline{\lim} \sqrt[n]{|a_n z^n|} < 1$, und die Reihe $\sum a_n z^n$ konvergiert nach dem Wurzelkriterium. Ist dagegen $|z| > R$, und ist a_{n_k} eine Teilfolge mit $\lim \sqrt[n_k]{|a_{n_k}|} = L$, so ist die Folge $a_{n_k} z^{n_k}$ keine Nullfolge, die Reihe $\sum a_n z^n$ muss also divergieren. □

Satz 5.37 — *Es sei (a_n) eine Folge komplexer Zahlen und $n_0 \in \mathbb{N}$ so, dass $a_n \neq 0$ für alle $n \geq n_0$. Existiert der uneigentliche Grenzwert $R = \lim \left|\frac{a_n}{a_{n+1}}\right|$, so ist R der Konvergenzradius der Reihe $\sum a_n z^n$.*

Beweis. Der Grenzwert $R = \lim \left|\frac{a_n}{a_{n+1}}\right|$ existiere. Für z mit $0 < |z| < R$ folgt: $\lim \left|\frac{a_{n+1}}{a_n} z\right| < 1$. Damit konvergiert die Reihe $\sum a_n z^n$ absolut nach dem Quotientenkriterium 5.13. Umgekehrt ist für z mit $|z| > R$ die Folge $|a_n z^n|$ unbeschränkt und die Reihe $\sum a_n z^n$ muss divergieren. □

Sind $\sum_{n=0}^{\infty} a_n z^n$ und $\sum_{n=0}^{\infty} b_n z^n$ zwei Potenzreihen mit Konvergenzradien R_1 und R_2 so haben die Reihen $\sum_{n=0}^{\infty} (a_n + b_n) z^n$ und $\sum_{n=0}^{\infty} \left(\sum_{k=0}^{n} a_k b_{n-k}\right) z^n$ jeweils einen Konvergenzradius, der größer oder gleich $R := \min\{R_1, R_2\}$ ist, und für alle z mit $|z| < R$ gilt:

$$\sum_{n=0}^{\infty} a_n z^n + \sum_{n=0}^{\infty} b_n z^n = \sum_{n=0}^{\infty} (a_n + b_n) z^n$$

[20] Jacques Salomon Hadamard, ∗ 8.12.1865 † 17.10.1963.

und

$$\sum_{n=0}^{\infty} a_n z^n \cdot \sum_{n=0}^{\infty} b_n z^n = \sum_{n=0}^{\infty} \left(\sum_{k=0}^{n} a_k b_{n-k} \right) z^n.$$

Ersteres ist sofort klar, und letzteres folgt aus Satz 5.30.

Lemma 5.38 — *Es sei $\sum_{n=0}^{\infty} a_n z^n$ eine Potenzreihe mit Konvergenzradius $R > 0$. Dann existiert zu jedem $N \in \mathbb{N}_0$ und jedem $r < R$ eine Konstante C derart, dass für alle $z \in \mathbb{C}$, $|z| \leq r$, gilt:*

$$\left| \sum_{n=N+1}^{\infty} a_n z^n \right| \leq C \cdot |z|^{N+1}.$$

Beweis. Da $r < R$, konvergiert die Reihe $\sum a_n r^n$ absolut. Es sei

$$C := \sum_{n=N+1}^{\infty} |a_n| r^{n-N-1}.$$

Dann gilt für alle z, $|z| \leq r$:

$$\begin{aligned} \left| \sum_{n=N+1}^{\infty} a_n z^n \right| &= |z|^{N+1} \left| \sum_{n=N+1}^{\infty} a_n r^{n-N-1} \left(\frac{z}{r} \right)^{n-N-1} \right| \\ &\leq |z|^{N+1} \sum_{n=N+1}^{\infty} |a_n| r^{n-N-1} \\ &= C|z|^{N+1}. \end{aligned}$$

□

Satz 5.39 — *Es sei $a(z) = \sum_{n=0}^{\infty} a_n z^n$ eine Potenzreihe mit positivem Konvergenzradius R. Ferner seien nicht alle $a_n = 0$. Dann existiert ein ε, $0 < \varepsilon < R$, derart, dass $a(z)$ in der offenen Kreisscheibe $U_\varepsilon(0)$ höchstens eine Nullstelle im Ursprung hat.*

Beweis. Wir zeigen dass in einer hinreichend kleinen Umgebung der 0 die einzige mögliche Nullstelle $z = 0$ ist. Wir nehmen das Gegenteil an. Dann gibt es zu jedem $k \in \mathbb{N}$ ein z_k mit $0 < |z_k| < \frac{R}{2k}$ und $a(z_k) = 0$. Nach Voraussetzung verschwinden nicht alle a_n. Es sei N minimal mit $a_N \neq 0$. Da z_k eine Nullstelle ist, gilt:

$$a_N z_k^N = - \sum_{n=N+1}^{\infty} a_n z_k^n.$$

Nach dem vorigen Lemma exisitiert aber eine Konstante C so, dass für alle z mit $|z| < R/2$ gilt:

$$\left| \sum_{n=N+1}^{\infty} a_n z^n \right| < C|z|^{N+1}.$$

Wendet man dies auf die z_k an, so erhält man:

$$0 < |a_N| < C|z_k|$$

für alle k, im Widerspruch dazu, dass (z_k) eine Nullfolge ist. □

Folgerung 5.40 — *Sind $a(z) = \sum_{n=0}^{\infty} a_n z^n$ und $b(z) = \sum_{n=0}^{\infty} b_n z^n$ Potenzreihen mit positiven Konvergenzradien und ist (z_k) eine Nullfolge in $\mathbb{C} \setminus \{0\}$ mit $a(z_k) = b(z_k)$ für alle k, so folgt $a_n = b_n$ für alle $n \in \mathbb{N}_0$.*

Beweis. Man wende den Satz auf die Potenzreihe $\sum(a_n - b_n)z^n$ an. □

5.5.1 Die Exponentialfunktion

Die Exponentialreihe hat den Konvergenzradius ∞. Wir können also für $z \in \mathbb{C}$ die Exponentialfunktion

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

definieren.

Satz 5.41 — *Für alle komplexen Zahlen z und w gilt*

$$\exp(z) \cdot \exp(w) = \exp(z + w). \tag{5.1}$$

Insbesondere ist $\exp(z) \neq 0$ für alle $z \in \mathbb{C}$.

Beweis. Es seien z und w beliebige komplexe Zahlen. Dann gilt

$$\begin{aligned}
\exp(z) \cdot \exp(w) &= \sum_{n=0}^{\infty} \frac{z^n}{n!} \cdot \sum_{n=0}^{\infty} \frac{w^n}{n!} \\
&= \sum_{n=0}^{\infty} \left(\sum_{k=0}^{n} \frac{z^k}{k!} \frac{w^{n-k}}{(n-k)!} \right) \\
&= \sum_{n=0}^{\infty} \frac{1}{n!} \left(\sum_{k=0}^{n} \binom{n}{k} z^k w^{n-k} \right) \\
&= \sum_{n=0}^{\infty} \frac{(z+w)^n}{n!} \\
&= \exp(z + w)
\end{aligned}$$

Die zweite Aussage folgt als Spezialfall aus der allgemeinen Multiplikativität mit $w = -z$. □

Satz 5.42 — *Für alle $z \in \mathbb{C}$ gilt*

$$\sum_{n=0}^{\infty} \frac{z^n}{n!} = \lim_{n \longrightarrow \infty} \left(1 + \frac{z}{n}\right)^n.$$

Beweis. Es sei $z \in \mathbb{C}$. Die Reihe $\sum_{n=0}^{\infty} \frac{z^n}{n!}$ konvergiert absolut. Zu vorgegebenem $\varepsilon > 0$ finden wir also ein N so, dass $\sum_{n=N}^{\infty} \frac{|z|^n}{n!} < \varepsilon$. Wir müssen die beiden folgenden Reihen vergleichen:

$$\begin{aligned} \exp(z) &= 1 + z + \frac{z^2}{2!} + \frac{z^3}{3!} + \frac{z^4}{4!} + \dots \\ \left(1 + \frac{z}{n}\right)^n &= 1 + z + \frac{z^2}{2!}\frac{n(n-1)}{n^2} + \frac{z^3}{3!}\frac{n(n-1)(n-2)}{n^3} + \dots . \end{aligned}$$

Daraus ergibt sich:

$$\begin{aligned} \left|\exp(z) - \left(1 + \frac{z}{n}\right)^n\right| &= \left|\sum_{k=2}^{\infty} \frac{z^k}{k!}\left(1 - \left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right)\cdots\left(1 - \frac{k-1}{n}\right)\right)\right| \\ &\leq \left|\sum_{k=2}^{N-1} \frac{z^k}{k!}\left(1 - \left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right)\cdots\left(1 - \frac{k-1}{n}\right)\right)\right| \\ &\quad + \sum_{k=N}^{\infty} \frac{|z|^k}{k!}\left(1 + \left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right)\cdots\left(1 - \frac{k-1}{n}\right)\right). \end{aligned}$$

Im zweiten Summanden auf der rechten Seite ist der geklammerte Faktor kleiner als 2, der gesamte Summand nach Wahl von N also $< 2\varepsilon$. Der erste Summand ist eine endliche Summe, die für $n \longrightarrow \infty$ gegen 0 geht. Da ε beliebig war, folgt $\exp(z) = \lim \left(1 + \frac{z}{n}\right)^n$. □

Bemerkung 5.43 — Der Grenzwert $\lim \left(1 + \frac{z}{n}\right)^n$ taucht bei der Diskussion stetiger Verzinsung auf: Ein Kapital K erhöht sich bei einem Zins z nach Ablauf eines Jahres auf $K(1 + z)$. Wird halbjährlich verzinst und der Zins wieder mitangelegt, so erhält man stattdessen $K(1 + z/2)^2$, bei monatlicher Verzinsung $K(1 + z/12)^{12}$, bei stetiger Verzinsung also $K \lim \left(1 + \frac{z}{n}\right)^n$. Die Existenz dieses Grenzwertes besagt also insbesondere, dass das Kapital bei ständiger Verzinsung nicht ins Unermessliche steigt, sondern durch die Zahl $\exp(z)$ beschränkt bleibt.

Derselbe mathematische Zusammenhang besteht bei allen Wachstums- (Bakterienkulturen) oder Zerfallsprozessen (radioaktive Stoffe) bei denen das Wachstum oder der Zerfall proportional zur vorhandenen Stoffmenge ist. Wir kommen später darauf zurück.

Definition 5.44 — Die Zahl

$$e := \sum_{n=0}^{\infty} \frac{1}{n!} = \lim \left(1 + \frac{1}{n}\right)^n \tag{5.2}$$

heißt Eulersche Zahl.

Satz 5.45 (Euler) — *e ist eine irrationale Zahl.*

Beweis. Angenommen, es gälte $e = \frac{p}{q}$ mit $p, q \in \mathbb{N}$. Durch Multiplikation von

$$0 < e - \sum_{n=0}^{q} \frac{1}{n!} = \sum_{n=q+1}^{\infty} \frac{1}{n!}$$

mit $q!$ erhält man:

$$\begin{aligned} 0 &< N := p(q-1)! - \sum_{n=0}^{q} \frac{q!}{n!} \\ &= \frac{1}{q+1} + \frac{1}{(q+1)(q+2)} + \frac{1}{(q+1)(q+2)(q+3)} + \dots \\ &< \frac{1}{q+1} + \frac{1}{(q+1)^2} + \frac{1}{(q+1)^3} + \dots \\ &= \frac{1}{q} \leq 1. \end{aligned}$$

Man sieht nun leicht, dass N einerseits eine ganze Zahl ist, andererseits zwischen 0 und 1 liegen soll. Das ist unmöglich. Die Annahme, e wäre rational, war falsch. □

5.5.2 Die trigonometrischen Funktionen

Wir definieren für beliebiges $z \in \mathbb{C}$ die Sinusfunktion

$$\sin(z) := \sum_{n=0}^{\infty} (-1)^n \frac{z^{2n+1}}{(2n+1)!} \tag{5.3}$$

und die Kosinusfunktion

$$\cos(z) := \sum_{n=0}^{\infty} (-1)^n \frac{z^{2n}}{(2n)!} \tag{5.4}$$

Beide Reihen haben den Konvergenzradius ∞. Offensichtlich ist $\sin(z)$ eine ungerade und $\cos(z)$ eine gerade Funktion, d.h.

$$\sin(-z) = -\sin(z) \quad \text{und} \quad \cos(-z) = \cos(z).$$

Durch Vergleich mit der Exponentialreihe erhält man die Eulersche Formel:

Satz 5.46 — *Für jedes $z \in \mathbb{C}$ gilt* $\exp(iz) = \cos(z) + \sin(z)i$.

Da andererseits $\exp(-iz) = \cos(z) - \sin(z)i$, gelten für jedes $z \in \mathbb{C}$ die Umkehrformeln

$$\cos(z) = \frac{1}{2}(\exp(iz) + \exp(-iz)), \quad \sin(z) = \frac{1}{2i}(\exp(iz) - \exp(-iz)).$$

Benutzt man nun noch die Multiplikativität der Exponentialfunktion (5.1) so erhält man die Additionstheoreme

Satz 5.47 — *Für alle $z, w \in \mathbb{C}$ gilt:*

$$\sin(z+w) = \sin(z)\cos(w) + \cos(z)\sin(w), \tag{5.5}$$

$$\cos(z+w) = \cos(z)\cos(w) - \sin(z)\sin(w). \tag{5.6}$$

Insbesondere gilt für alle $z \in \mathbb{C}$:

$$\cos(z)^2 + \sin(z)^2 = 1. \tag{5.7}$$

5.5.3 Die Binomialreihe

Die Binomialreihe

$$B_\alpha(z) = \sum_{n=0}^{\infty} \binom{\alpha}{n} z^n$$

hat den Konvergenzradius ∞, falls $\alpha \in \mathbb{N}_0$, denn dann sind alle bis auf die ersten α Terme identisch 0. Falls $\alpha \notin \mathbb{N}_0$, so zeigt die Überlegung in Beispiel 5.15 zusammen mit dem Kriterium 5.37, dass der Konvergenzradius genau 1 ist.

Satz 5.48 — *Für beliebige $\alpha, \beta \in \mathbb{C}$ und für alle $z \in U_1(0)$ gilt*

$$B_\alpha(z) \cdot B_\beta(z) = B_{\alpha+\beta}(z).$$

Insbesondere gilt für alle $n \in \mathbb{N}$:

$$B_\alpha(z)^n = B_{n\alpha}(z).$$

Beweis. Mit dem Satz über Cauchyprodukte 5.30 folgt für alle $z \in U_1(0)$:

$$\begin{aligned} B_\alpha(z)B_\beta(z) &= \sum_{n=0}^{\infty} \binom{\alpha}{n} z^n \sum_{m=0}^{\infty} \binom{\beta}{m} z^m \\ &= \sum_{n=0}^{\infty} z^n \left(\sum_{k=0}^{n} \binom{\alpha}{k} \binom{\beta}{n-k} \right). \end{aligned}$$

Zum Beweis der Behauptung genügt es also, die folgende Identität von Binomialkoeffizienten zu beweisen:

$$\sum_{k=0}^{n} \binom{\alpha}{k}\binom{\beta}{n-k} = \binom{\alpha+\beta}{n}. \tag{5.8}$$

Wir tun dies durch Induktion über n. Für $n = 0$ oder $n = 1$ ist die Behauptung offensichtlich richtig. Wir nehmen an, die Behauptung sei schon bewiesen für ein $n \geq 1$ und zeigen, dass sie dann auch für $n+1$ richtig ist. Nun gilt in der Tat:

$$\begin{aligned}
\binom{\alpha+\beta}{n+1} &= \frac{\alpha+\beta-n}{n+1}\binom{\alpha+\beta}{n} \\
&= \frac{\alpha+\beta-n}{n+1}\sum_{k=0}^{n}\binom{\alpha}{k}\binom{\beta}{n-k} \\
&= \sum_{k=0}^{n}\left[\frac{\beta-n+k}{n+1}\binom{\alpha}{k}\binom{\beta}{n-k} + \frac{\alpha-k}{n+1}\binom{\alpha}{k}\binom{\beta}{n-k}\right] \\
&= \sum_{k=0}^{n}\left[\frac{n+1-k}{n+1}\binom{\alpha}{k}\binom{\beta}{n+1-k} + \frac{k+1}{n+1}\binom{\alpha}{k+1}\binom{\beta}{n-k}\right] \\
&= \sum_{k=0}^{n+1}\left[\frac{n+1-k}{n+1}\binom{\alpha}{k}\binom{\beta}{n+1-k} + \frac{k}{n+1}\binom{\alpha}{k}\binom{\beta}{n+1-k}\right] \\
&= \sum_{k=0}^{n+1}\binom{\alpha}{k}\binom{\beta}{n+1-k}.
\end{aligned}$$

Das beweist die erste Aussage des Satzes, die zweite Aussage des Satzes folgt durch Induktion aus der ersten. □

Satz 5.49 — *Es sei $z \in \mathbb{C}$ und $n \in \mathbb{N}$. Dann gibt es ein $w \in \mathbb{C}$ mit $w^n = z$.*

Beweis. 0. Schritt: Die Behauptung ist klar für $z = 0$.

1. Schritt: Es gelte $\mathrm{Re}(z) > 0$. Dann ist $|z-1| < |z+1|$, und folglich liegt $u := \frac{z-1}{z+1}$ in $U_1(0)$. Umgekehrt ist $z = \frac{1+u}{1-u}$. Daher sind die Zahlen $B_{1/n}(u)$ und $B_{1/n}(-u)$ definiert und haben die Eigenschaft

$$B_{1/n}(u)^n = 1+u, \quad B_{1/n}(-u)^n = 1-u \neq 0.$$

Schließlich ist $w := B_{1/n}(u)/B_{1/n}(-u)$ definiert, und es gilt

$$w^n = (1+u)/(1-u) = z,$$

wie verlangt.

2. Schritt: Ist z eine beliebige komplexe Zahl $\neq 0$, so wissen wir aufgrund von Satz 4.10, dass es ein $v \in \mathbb{C}$ mit $v^4 = z$ gibt. Aber jede der Zahlen $v_\nu := vi^\nu$

hat die Eigenschaft $v_\nu^4 = z$. Wir wählen ν so, dass $z' := vi^\nu$ einen positiven Realteil hat. Nach Schritt 1 gibt es dann ein $w' \in \mathbb{C}$ mit $(w')^n = z'$. Mit der Wahl $w := (w')^4$ folgt $w^n = (z')^4 = z$. □

5.6 Vom richtigen Zählen

Potenzreihen können in vielen Fällen erfolgreich in der Kombinatorik eingesetzt werden. Das Vorgehen ist wie folgt: Angenommen, a_n ist für jedes $n \in \mathbb{N}_0$ die Lösung eines von n abhängigen Zählproblems. Wir bilden die Potenzreihe $f(z) = \sum_{n=0}^{\infty} a_n z^n$. Diese Potenzreihe heißt in diesem Zusammenhang die erzeugende Funktion zu der Folge (a_n). Man leite für die a_n rekursive Beziehungen her und übersetze diese in Aussagen über die Funktion $f(z)$. Die Analyse der Eigenschaften von $f(z)$ hilft oft, das ursprüngliche kombinatorische Problem zu lösen. Wir besprechen diese Idee exemplarisch an zwei Beispielen.

Beispiel 5.50 — Die Fibonaccizahlen
Wir haben die Fibonaccizahlen $(f_n)_{n\in\mathbb{N}_0}$ rekursiv wie folgt definiert:

$$f_0 := 1, \quad f_1 := 1, \quad f_{n+1} := f_n + f_{n-1} \quad \text{für alle } n \geq 1.$$

Das Problem lautet nun: Kann man für die f_n eine explizite Formel angeben, die also erlauben würde, direkt f_n auszurechnen, ohne vorher alle $f_0, f_1, \ldots, f_{n-1}$ zu berechnen? Wir bilden die Potenzreihe

$$f(z) := \sum_{n=0}^{\infty} f_n z^n$$

Wir wissen außerdem, dass $\lim \frac{f_{n+1}}{f_n} = \frac{\sqrt{5}+1}{2}$. Also ist der Konvergenzradius der Reihe durch

$$R = \frac{2}{\sqrt{5}+1} = \frac{\sqrt{5}-1}{2} \approx 0,618$$

gegeben. Die Beziehung $f_{n+1} = f_n + f_{n-1}$ ist zusammen mit den Anfangsbedingungen $f_0 = f_1 = 1$ äquivalent zu der Aussage:

$$f(z) - zf(z) - z^2 f(z) = 1.$$

Demnach gilt für alle z, $|z| < R$:

$$f(z) = \frac{1}{1 - z - z^2}.$$

Die Nullstellen des Polynoms $z^2 + z - 1$ sind $\beta = -\frac{\sqrt{5}+1}{2}$ und $\gamma = \frac{\sqrt{5}-1}{2}$. Wir machen den Ansatz

$$\frac{1}{1 - z - z^2} = \frac{B}{\beta - z} + \frac{C}{\gamma - z}$$

Das führt auf die Bedingungen

$$B\gamma + C\beta = -1 \quad \text{und} \quad B + C = 0$$

an B und C mit der eindeutigen Lösung

$$B = -C = -\frac{1}{\gamma - \beta} = -\frac{1}{\sqrt{5}}.$$

Damit haben wir:

$$f(z) = -\frac{1}{\sqrt{5}} \left(\frac{1}{\beta - z} + \frac{1}{\gamma - z} \right).$$

Die rechte Seite besteht aus zwei geometrischen Reihen. Lösen wir nach diesen auf, so folgt:

$$\sum_{n=0}^{\infty} f_n z^n = -\frac{1}{\sqrt{5}} \sum_{n=0} \left(\frac{1}{\beta^{n+1}} - \frac{1}{\gamma^{n+1}} \right) z^n.$$

Folgerung 5.40 erlaubt den Koeffizientenvergleich der beiden Reihen. Damit finden wir schließlich die explizite Formel

$$f_n = \frac{1}{\sqrt{5}} \left(\left(\frac{\sqrt{5}+1}{2} \right)^{n+1} + (-1)^n \left(\frac{\sqrt{5}-1}{2} \right)^{n+1} \right). \tag{5.9}$$

Man beachte, dass der zweite Term in der Summe stets kleiner als 1 ist und überdies für $n \longrightarrow \infty$ gegen 0 geht. Man erhält also f_n vereinfacht durch Rundung der irrationalen Zahl

$$\frac{1}{\sqrt{5}} \left(\frac{\sqrt{5}+1}{2} \right)^{n+1}. \tag{5.10}$$

Beispiel 5.51 — Die Catalanschen[21] Zahlen

Es bezeichne c_n die Anzahl der Möglichkeiten, ein (nicht assoziatives) Produkt aus n Faktoren (in vorgegebener Reihenfolge) zu klammern. Man findet sofort, dass $c_1 = 1$, $c_2 = 1$, $c_3 = 2$, $c_4 = 5$. Kann man eine allgemeine Formel für c_n angeben?

Wir leiten zunächst die folgende Rekursionsformel für die c_n her: Für alle $n \geq 2$ gilt

$$c_n = \sum_{i=1}^{n-1} c_i c_{n-i}. \tag{5.11}$$

Das Bestehen dieser Beziehung sieht man so: Bei einem Produkt aus n Faktoren werden $n-1$ Verknüpfungen ausgeführt. Die zuletzt vorgenommene Multiplikation verknüpft ein irgendwie geklammertes Produkt der ersten i Faktoren mit einem irgendwie geklammerten Produkt der letzten $n-i$ Faktoren für ein gewisses $i \in \{1, 2, \ldots, n-1\}$. Für festes i gibt es c_i Möglichkeiten den ersten Faktor

[21] Eugène Charles Catalan, $*$ 30.5.1814 † 14.2.1894.

und c_{n-i} Möglichkeiten, den zweiten Faktor zu klammern. Da jede Kombination von Möglichkeiten zulässig ist und zu verschiedenen Produkten führt, hat man $c_i c_{n-i}$ Möglichkeiten. Lässt man jetzt noch i laufen, so erhält man die Formel (5.11).

Wir betrachten die erzeugende Funktion

$$c(z) := \sum_{n=0}^{\infty} c_n z^n. \tag{5.12}$$

Es ist zweckmäßig, die noch unbestimmte Zahl c_0 als 0 zu wählen, denn dann kann man die Beziehung (5.11) in die Form

$$c_n = \sum_{i=0}^{n} c_i c_{n-i} \tag{5.13}$$

bringen, in der die rechte Seite, abgesehen von einem Faktor z^n, der n-te Koeffizient des Cauchyproduktes von $c(z)$ mit sich ist. Wir erhalten formal die Gleichung

$$c(z) = c(z)^2 + z. \tag{5.14}$$

Der Summand z kommt hinzu, weil die Gleichung (5.13) nur für $n \geq 2$ gilt. Gleichung (5.14) ist zunächst nur eine formale Gleichung, da wir das Wachstumsverhalten der Koeffizienten c_n nicht kennen und daher auch keine Aussagen über den Konvergenzradius der Reihe $c(z)$ machen können. Insbesondere wissen wir nicht, ob die Gleichung (5.14) überhaupt für irgendeine Zahl $z \neq 0$ gilt. Um weiterarbeiten zu können, nehmen wir an, der Konvergenzradius sei positiv, müssen aber daran denken, diese Annahme am Ende zu rechtfertigen.

Aus (5.14) erhält man:

$$1 - 4z = 1 - 4c(z) + 4c(z)^2 = (1 - 2c(z))^2.$$

Wir kennen bereits eine Potenzreihe, deren Quadrat $1 - 4z$ ist, nämlich die Binomialreihe $B_{1/2}(-4z)$ mit dem Konvergenzradius $\frac{1}{4}$. Das führt auf den Ansatz

$$\tilde{c}(z) := \frac{1}{2}(1 - B_{1/2}(-4z)) \tag{5.15}$$

Diese Potenzreihe $\tilde{c}(z)$ hat nun die folgenden Eigenschaften: Erst einmal ist der konstante Term $\tilde{c}_0 = 0$. Zweitens hat sie denselben Konvergenzradius wie die Binomialreihe, also $\frac{1}{4}$. Deshalb gilt für alle z mit $|z| < \frac{1}{4}$ die Beziehung

$$1 - 4z = (1 - 2\tilde{c}(z))^2 = 1 - 4\tilde{c}(z) + 4\tilde{c}(z)^2,$$

und dies übersetzt sich wegen $\tilde{c}_0 = 0$ in die Rekursionsgleichungen

$$\tilde{c}_n = \sum_{i=1}^{n-1} \tilde{c}_i \tilde{c}_{n-i} \quad \text{und} \quad \tilde{c}_1 = 1.$$

Durch diese Gleichung sind die Koeffizienten $\tilde{c}_n$ aber bereits vollständig bestimmt, müssen also mit den c_n, die derselben Beziehung genügen, übereinstimmen. Damit haben wir die Hypothese über den Konvergenzradius gerechtfertigt und können daran gehen, die $c_n = \tilde{c}_n$ auszurechnen. Wir finden:

$$c_n = -\frac{1}{2}\binom{\frac{1}{2}}{n}(-4)^n = \frac{(2n-2)!}{(n-1)!n!} = \frac{1}{n}\binom{2(n-1)}{n-1}.$$

§6 Stetigkeit

6.1 Stetige Funktionen

Es sei D eine Menge. Abbildungen $f : D \longrightarrow \mathbb{R}$ oder $f : D \longrightarrow \mathbb{C}$ heißen reelle bzw. komplexe Funktionen.

Definition 6.1 — Es sei $D \subset \mathbb{C}$. Eine Funktion $f : D \longrightarrow \mathbb{C}$ heißt stetig im Punkt $a \in D$, wenn es zu jedem $\varepsilon > 0$ ein $\delta > 0$ gibt derart, dass für alle $z \in D$ mit $|z - a| < \delta$ gilt $|f(z) - f(a)| < \varepsilon$. Ist f in allen Punkten $a \in D$ stetig, so nennt man f kurz eine auf D stetige Funktion.

Hier ist die Reihenfolge, in der ε und δ erscheinen von entscheidender Wichtigkeit: Erst wird ein $\varepsilon > 0$ vorgegeben. Dazu muss man ein passendes $\delta > 0$ mit bestimmten Eigenschaften finden. Dieses δ hängt in der Regel sowohl von dem ϵ wie von der Stelle a ab, für die Stetigkeit behauptet wird. Häufig versucht man dies dadurch deutlich auszudrücken, dass man schreibt: „... gibt es ein $\delta = \delta(a, \varepsilon)$ ...“. Die logische Abhängigkeit zwischen ϵ und δ muss in Fleisch und Blut übergehen. Man stelle sich den Sachverhalt wie einen Dialog vor: A behauptet die Stetigkeit. B widerspricht. Daraufhin darf B ein beliebig kleines ε vorgeben, und A muss in der Lage sein, mit einem geeigneten δ zu antworten.

Beispiel 6.2 — Die Funktion $f : \mathbb{C} \longrightarrow \mathbb{C}, z \mapsto z^2$, ist in allen Punkten stetig.
Beweis. Es sei $a \in \mathbb{C}$ und $\varepsilon > 0$ vorgegeben. Es sei $\delta := \min\{1, \varepsilon/(2|a| + 1)\}$. Dann gilt für alle z mit $|z - a| < \delta$:

$$|f(z) - f(a)| = |z^2 - a^2| = |z - a| \cdot |2a + (z - a)| < \delta(2|a| + \delta) \leq \delta(2|a| + 1) \leq \varepsilon.$$

□

In diesem Beispiel sieht man deutlich die Abhängigkeit von δ sowohl von $a \in \mathbb{C}$ wie von ε. Gelegentlich gilt aber etwas mehr:

Definition 6.3 — Es sei $D \subset \mathbb{C}$. Eine Funktion $f : D \longrightarrow \mathbb{C}$ heißt gleichmäßig stetig, wenn es zu jedem $\varepsilon > 0$ ein $\delta > 0$ gibt so, dass für alle $z, a \in D$ gilt: Aus $|z - a| < \delta$ folgt $|f(z) - f(a)| < \varepsilon$.

Es ist klar, dass aus der gleichmäßigen Stetigkeit die Stetigkeit in allen Punkten des Definitionsbereichs folgt.

Beispiel 6.4 — Für jedes $n \in \mathbb{N}$ ist die Funktion $\sqrt[n]{-} : \mathbb{R}_{\geq 0} \longrightarrow \mathbb{R}$ gleichmäßig stetig.

Beweis. Sind zunächst $s, t \geq 0$ beliebige reelle Zahlen, so folgt

$$(s+t)^n = \sum_{i=0}^{n} \binom{n}{i} s^i t^{n-i} \geq s^n + t^n.$$

Speziell mit $s = \sqrt[n]{a}$ und $t = \sqrt[n]{b}$ folgt hieraus

$$\sqrt[n]{a} + \sqrt[n]{b} \geq \sqrt[n]{a+b},$$

und daraus für beliebige $x, y \geq 0$:

$$|\sqrt[n]{x} - \sqrt[n]{y}| \leq \sqrt[n]{|x-y|}.$$

Damit bekommt man schnell die gleichmäßige Stetigkeit: Ist $\varepsilon > 0$ vorgegeben, so sei $\delta = \varepsilon^n$. Für beliebige $z, a \geq 0$ mit $|z-a| < \delta$ folgt nun:

$$|\sqrt[n]{z} - \sqrt[n]{a}| \leq \sqrt[n]{|z-a|} < \sqrt[n]{\delta} = \varepsilon.$$

Das war zu zeigen. □

Beispiele 6.5 — Andere Beispiele für gleichmäßig stetige Funktionen sind die identische Funktion $\mathrm{id} : \mathbb{C} \to \mathbb{C}$, die komplexe Konjugation $\overline{} : \mathbb{C} \to \mathbb{C}$, und die Projektionen auf Real- und Imaginärteil $\mathrm{Re} : \mathbb{C} \longrightarrow \mathbb{R}$, $\mathrm{Im} : \mathbb{C} \longrightarrow \mathbb{R}$, sowie der Absolutbetrag $|-| : \mathbb{C} \longrightarrow \mathbb{R}$.

Satz 6.6 *(Folgenkriterium für Stetigkeit) — Es sei $f : D \longrightarrow \mathbb{C}$ eine Funktion und $a \in D \subset \mathbb{C}$. Die folgenden Aussagen sind äquivalent:*

1. *f ist in $a \in D$ stetig.*
2. *Für jede Folge (z_n) in D mit $\lim z_n = a$ gilt $\lim f(z_n) = f(a)$.*

Beweis. 1.⇒2.: Es sei also f stetig in a und (z_n) eine Folge in D mit $\lim z_n = a$. Wir müssen zeigen, dass $f(z_n)$ gegen $f(a)$ konvergiert. Ist $\varepsilon > 0$ vorgegeben, so existiert wegen Stetigkeit ein $\delta > 0$ derart, dass für alle $z \in D$ mit $|z-a| < \delta$ gilt $|f(z) - f(a)| < \varepsilon$. Da die Folge (z_n) gegen a konvergiert, gibt es ein $n_0 \in \mathbb{N}$ derart, dass für alle $n \geq n_0$ gilt: $|z_n - a| < \delta$, nach Konstruktion also auch $|f(z_n) - f(a)| < \varepsilon$. Also konvergiert $f(z_n)$ gegen $f(a)$.

2.⇒1.: Es gelte also die Bedingung 2. Wäre f nicht stetig in a, so gäbe es ein $\epsilon > 0$, zu dem für jedes $\delta > 0$ ein $z \in D$ existierte mit $|z-a| < \delta$, aber $|f(z) - f(a)| \geq \epsilon$. Insbesondere gibt es dann zu jedem $n \in \mathbb{N}$ ein $z_n \in D$ mit $|z_n - a| < \frac{1}{n}$ und $|f(z_n) - f(a)| \geq \varepsilon$. Das bedeutet aber, dass $\lim z_n = a$, ohne dass $\lim f(z_n) = f(a)$, im Widerspruch zur Annahme. □

Sind $f, g : D \to \mathbb{C}$ Funktionen, so bezeichnet man mit $f+g, fg : D \longrightarrow \mathbb{C}$ die Funktionen

$$(f+g)(z) := f(z) + g(z) \quad \text{und} \quad (fg)(z) := f(z)g(z).$$

Analog definiert man zu $f : D \longrightarrow \mathbb{C}$ die Funktionen $\mathrm{Re}(f), \mathrm{Im}(f), |f| : D \longrightarrow \mathbb{R}$ und $\overline{f} : D \longrightarrow \mathbb{C}$. Gilt $g(z) \neq 0$ für alle $z \in D$, so ist $\frac{f}{g} : D \longrightarrow \mathbb{C}$ durch $\frac{f}{g}(z) = \frac{f(z)}{g(z)}$ definiert.

Satz 6.7 — *Sind $f, g : D \longrightarrow \mathbb{C}$ in $a \in D$ stetig, so auch $f+g$ und fg. Ist $g(a) \neq 0$, so gibt es ein $\eta > 0$ derart, dass $g(z) \neq 0$ für alle $z \in D \cap U_\eta(a)$, und $\frac{f}{g}$ ist auf $D \cap U_\eta(a)$ definiert und in a stetig.*

Beweis. Die erste Aussage folgt direkt aus dem Folgenkriterium und den Rechenregeln für Grenzwerte von Folgen: Ist nämlich (z_n) eine Folge in D mit $\lim(a_n) = a$, so gilt:

$$\begin{aligned} \lim(f+g)(z_n) = \lim(f(z_n) + g(z_n)) &= \lim f(z_n) + \lim g(z_n) \\ &= f(a) + g(a) = (f+g)(a). \end{aligned}$$

Analog beweist man die Aussage für das Produkt fg.

Ist nun $g(a) \neq 0$, so gibt es wegen der Stetigkeit von g in a ein $\eta > 0$ derart, dass für alle $z \in D \cap U_\eta(a)$ gilt: $|g(z) - g(a)| < \frac{1}{2}|g(a)|$. Mit der Dreiecksungleichung folgt für solche z: $|g(z)| > \frac{1}{2}|g(a)|$, also insbesondere $g(z) \neq 0$. Der Rest der Aussage folgt wieder mit dem Folgenkriterium. □

Satz 6.8 — *Es sei $f : D \to D' \subset \mathbb{C}$ stetig in $a \in D$ und $g : D' \longrightarrow \mathbb{C}$ stetig in $f(a) \in D'$. Dann ist auch die zusammengesetzte Funktion $g \circ f : D \to \mathbb{C}$ stetig in a.*

Beweis. Folgt sofort aus dem Folgenkriterium. □

Folgerung 6.9 — *Ist $f : D \to \mathbb{C}$ stetig in $a \in D$, so auch die Funktionen $\mathrm{Re}(f)$, $\mathrm{Im}(f)$, $|f|$, und $\overline{f}$.*

Beispiel 6.10 — Jedes Polynom $f = f_0 + f_1 X + \ldots + f_n X^n \in \mathbb{C}[X]$ definiert eine stetige Funktion $f : \mathbb{C} \longrightarrow \mathbb{C}$, $z \mapsto f(z) = f_0 + f_1 z + \ldots + f_n z^n$. Denn trivialerweise sind konstante Funktionen und die identische Funktion $X : \mathbb{C} \longrightarrow \mathbb{C}$, $z \mapsto z$, stetig, und nach den obigen Sätzen alle daraus durch Addition und Multiplikation gebildeten Funktionen.

Beispiel 6.11 — Als zusammengesetzte Funktion ist auch $\sqrt{3 + X^2} : \mathbb{R} \to \mathbb{R}$ stetig.

6.2 Grenzwerte

Definition 6.12 — Es sei $D \subset \mathbb{C}$ und $a \in \mathbb{C}$. Eine Funktion $f : D \to \mathbb{C}$ hat in a den Grenzwert $c \in \mathbb{C}$, wenn die beiden folgenden Bedingungen erfüllt sind:

1. Es gibt eine Folge (z_n) in $D \setminus \{a\}$ mit $\lim z_n = a$.

2. Für jede solche Folge gilt $\lim f(z_n) = c$.

In diesem Falle schreibt man $\lim_{z \to a} f(z) = c$.

Man beachte, dass in dieser Definition a kein Punkt von D zu sein braucht. Insbesondere ist f also in a nicht notwendigerweise definiert. Andererseits hat der Wert $f(a)$, falls a doch in D liegt, keinen Einfluss auf den Grenzwert, da die Folgen ja nur in $D \setminus \{a\}$ gebildet werden.

Definition 6.13 — 1. Es sei $D \subset \mathbb{C}$. Ein Punkt $a \in \mathbb{C}$ heißt Häufungspunkt von D, wenn es eine Folge (z_n) in $D \setminus \{a\}$ mit $\lim z_n = a$ gibt.
2. Eine Menge $D \subset \mathbb{C}$ heißt abgeschlossen, wenn D alle ihre Häufungspunkte enthält.

Lemma 6.14 — *1. Es sei $D \subset \mathbb{C}$. Ein Punkt $a \in \mathbb{C}$ ist genau dann Häufungspunkt von D, wenn für jedes $\varepsilon > 0$ gilt $(D \setminus \{a\}) \cap U_\varepsilon(a) \neq \emptyset$.*
2. Eine Menge $D \subset \mathbb{C}$ ist genau dann abgeschlossen, wenn es zu jedem $a \in \mathbb{C} \setminus D$ ein $\varepsilon > 0$ mit $U_\varepsilon(a) \subset \mathbb{C} \setminus D$ gibt.

Bemerkung 6.15 — Eine Funktion $f : D \to \mathbb{C}$ ist stetig in $a \in D$ genau dann, wenn entweder a kein Häufungspunkt von D ist oder $\lim_{z \to a} f(z) = f(a)$ gilt.

Für Funktionen, die auf Teilmengen von $\mathbb{R}$ definiert sind, sind die Begriffe des rechtsseitigen und linksseitigen Grenzwertes zweckmäßig:

Definition 6.16 — Es sei $D \subset \mathbb{R}$ und $a \in D$. Eine Funktion $f : D \to \mathbb{C}$ besitzt in a einen rechtsseitigen Grenzwert c, wenn die beiden folgenden Bedingungen erfüllt sind:

1. Es gibt eine Folge (z_n) in D mit $z_n > a$ für alle n und $\lim z_n = a$.

2. Für jede solche Folge gilt $\lim f(z_n) = c$.

In diesem Falle schreibt man $\lim_{z \downarrow a} f(z) = c$. Analog definiert man den linksseitigen Grenzwert $\lim_{z \uparrow a} f(z)$.

Schließlich treten häufig noch die folgenden uneigentlichen Grenzwerte auf. Es sei $D \subset \mathbb{R}$ und $f : D \longrightarrow \mathbb{R}$.

1. $\lim_{z\to+\infty} f(z) = c$ bedeutet: Es gibt eine Folge (z_n) in D mit $\lim z_n = \infty$, und für jede solche Folge gilt $\lim f(z_n) = c$. Analog: $\lim_{z\to-\infty} f(z) = c$.

2. $\lim_{z\to a} f(z) = \infty$ bedeutet: Es gibt eine Folge (z_n) in $D \setminus \{a\}$ mit $\lim z_n = a$, und für jede solche Folge gilt $\lim f(z_n) = \infty$. Analog: $\lim_{z\to a} f(z) = -\infty$.

3. Schließlich gibt es weitere Grenzwerte, deren Definition sich aber unmittelbar aus der Notation erschließt, wie etwa $\lim_{z\to\infty} f(z) = \infty$ oder $\lim_{z\downarrow a} f(z) = -\infty$.

Satz 6.17 — *(Rechenregeln für Grenzwerte)*
Es sei $D \subset \mathbb{R}$ und $f, g : D \to \mathbb{R}$ Funktionen. Existieren die Grenzwerte $\lim_{z\to a} f(z)$ und $\lim_{z\to a} g(z)$ im eigentlichen oder uneigentlichen Sinne, und sind die rechten Seiten der folgenden Ausdrücke definiert, so existieren auch die Grenzwerte auf der linken Seite und es gilt:

1. $\lim_{z\to a}(f(z) + g(z)) = \lim_{z\to a} f(z) + \lim_{z\to a} g(z)$.

2. $\lim_{z\to a}(f(z)g(z)) = (\lim_{z\to a} f(z)) \cdot (\lim_{z\to a} g(z))$.

3. *Ist $g(z) \neq 0$ für alle $z \in D$, so gilt auch* $\lim_{z\to a} \frac{f(z)}{g(z)} = \frac{\lim_{z\to a} f(z)}{\lim_{z\to a} g(z)}$.

Die entsprechenden Aussagen gelten für einseitige Grenzwerte $\lim_{z\downarrow a}$ etc. und uneigentliche Grenzwerte $\lim_{z\to\infty}$ etc.

Beweis. Der Satz ergibt sich aus den analogen Aussagen für Grenzwerte von Folgen. □

6.3 Zwischenwertsatz und Anwendungen

Satz 6.18 *(Zwischenwertsatz) — Es seien $a, b \in \mathbb{R}$, $a \leq b$, und es sei $f : [a, b] \to \mathbb{R}$ eine stetige Funktion. Gilt $f(a) \leq y \leq f(b)$ oder $f(a) \geq y \geq f(b)$, so existiert ein $x \in [a, b]$ mit $f(x) = y$.*

Beweis. Falls $y = f(a)$ oder $y = f(b)$, so sind wir sofort fertig. Wir nehmen also an, dass $f(a)$, $f(b)$ und y verschieden sind, und können sogar noch ohne Einschränkung annehmen, dass $f(a) < y < f(b)$, indem wir andernfalls f durch $-f$ ersetzen.

Dann ist die Menge $M := \{z \in [a, b] \mid f(z) \geq y\}$ nicht leer, denn $b \in M$. Es sei $x := \inf M$. Nach Konstruktion gibt es dann eine Folge (z_n) in M mit $\lim z_n = x$, und wegen Stetigkeit folgt: $f(x) = \lim f(z_n) \geq y > f(a)$. Insbesondere ist $x > a$,

und es gibt eine Folge (z_n) in $[a,b]$ mit $z_n \uparrow x$. Für diese gilt $f(z_n) < y$ für alle n, also auch $f(x) = \lim f(z_n) \leq y$. Zusammengenommen folgt $f(x) = y$. □

Beispiel 6.19 — Ist $f := f_0 + f_1 X + \ldots + f_n X^n \in \mathbb{R}[X]$ ein reelles Polynom von ungeradem Grad, so besitzt f eine reelle Nullstelle, d.h. es gibt ein $a \in \mathbb{R}$ mit $f(a) = 0$.

Beweis. Wir können ohne Einschränkung annehmen, dass $f_n = 1$. Wir schreiben

$$f(z) = z^n \left(1 + \frac{f_{n-1}}{z} + \ldots + \frac{f_0}{z^n}\right)$$

für alle $z \in \mathbb{R}^*$. Da $\lim_{z\to\infty} \frac{1}{z} = 0$ und $\lim_{z\to-\infty} \frac{1}{z} = 0$, folgt:

$$\lim_{z \longrightarrow \pm\infty} \left(1 + \frac{f_{n-1}}{z} + \ldots + \frac{f_0}{z^n}\right) = 1$$

und somit

$$\lim_{z\to\infty} f(z) = \infty, \quad \lim_{z\to-\infty} f(z) = -\infty.$$

Also gibt es ein $N \in \mathbb{N}$ mit $f(N) > 0$ und $f(-N) < 0$. Nach dem Zwischenwertsatz existiert ein $a \in [-N, N]$ mit $f(a) = 0$. □

Wir kennen den Begriff eines abgeschlossenen Intervalls $[a,b]$. Daneben betrachtet man noch halboffene Intervalle

$$[a,b) = \{x \in \mathbb{R} \,|\, a \leq x < b\},$$

(und entsprechend $(a,b]$), offene Intervalle

$$(a,b) = \{x \in \mathbb{R} \,|\, a < x < b\},$$

sowie uneigentliche Intervalle

$$(a,\infty) = \{x \in \mathbb{R} \,|\, a < x\},$$

(und entsprechend $[a,\infty)$, $(-\infty,b)$, $(-\infty,b]$, $(-\infty,\infty)$). Man überlegt sich leicht, dass eine Teilmenge $I \subset \mathbb{R}$ genau dann ein Intervall ist, wenn für alle $a,b \in I$ mit $a < b$ gilt: $[a,b] \subset I$.

Es sei $D \subset \mathbb{R}$. Eine Funktion $f : D \to \mathbb{R}$ heißt monoton wachsend, wenn für alle $a,b \in D$, $a < b$, gilt $f(a) \leq f(b)$, und streng monoton wachsend, wenn schärfer $f(a) < f(b)$ gilt. Analog: (streng) monoton fallend.

Satz 6.20 — *Es sei $I \subset \mathbb{R}$ ein Intervall und $f : I \to \mathbb{R}$ eine injektive stetige Funktion. Dann gilt:*

1. *$J := f(I)$ ist ebenfalls ein Intervall.*

2. *$f : I \to J$ und die Umkehrfunktion $f^{-1} : J \to I$ sind entweder beide streng monoton wachsend oder beide streng monoton fallend.*

3. *Die Umkehrfunktion f^{-1} ist stetig.*

Beweis. zu 1.: Es sei $J = f(I)$. Sind $a, b \in I$ und $a < b$, so liegt nach dem Zwischenwertsatz das ganze abgeschlossene Intervall mit den Grenzen $f(a)$ und $f(b)$ in J. Somit ist J selbst ein Intervall.

zu 2.: Falls I nur aus einem Punkt besteht, ist nichts zu zeigen. Es seien also $a, b \in I$ und $a < b$. Wir nehmen weiterhin an, dass $f(a) < f(b)$. Ist nun $x \in (a, b)$ und wäre $f(x) < f(a)$, so gäbe es nach dem Zwischenwertsatz ein $z \in (x, b)$ mit $f(z) = f(a)$. Aber das widerspräche der Injektivität von f. Ebenso schließt man den Fall $f(x) > f(b)$ aus. Demnach bleibt nur die Möglichkeit $f(a) < f(x) < f(b)$. Sind nun $x, y \in (a, b)$ mit $a < x < y < b$, so folgt durch erneute Anwendung dieser Schlussweise, dass $f(a) < f(x) < f(y) < f(b)$. Demnach folgt aus der Annahmen $f(a) < f(b)$, dass f auf dem Intervall $[a, b]$ streng monoton wächst.

In gleicher Weise zeigt man, dass aus der Annahme $f(a) > f(b)$ folgt, dass f auf dem Intervall $[a, b]$ streng monoton fällt.

Damit legen aber a und b bereits das Monotonieverhalten von f auf ganz I fest, denn jede endliche Teilmenge von I liegt in einem geeigneten abgeschlossenen Teilintervall, das a und b enthält.

Es ist klar, dass f^{-1} genau dann streng monoton wachsend bzw. fallend ist, wenn dies für f gilt.

zu 3.: Wir müssen zeigen, dass für jede Folge (z_n) in J mit $\lim z_n = z \in J$ gilt $\lim f^{-1}(z_n) = f^{-1}(z)$. Wir nehmen wieder ohne Einschränkung an, dass f streng monoton wächst.

Es sei $M := \{z_n \mid n \in \mathbb{N}\}$ und $a = \inf M$ und $b = \sup M$. Dann liegen a und b selbst in J. Für b sieht man das so: Entweder ist b das Maximum von M, d.h. liegt selbst in M, oder b ist der Limes einer Teilfolge der z_n. Aber eine solche konvergiert zwangsläufig gegen z, so dass also $b = z \in J$. Für a schließt man genauso.

Jedenfalls ist f^{-1} in a und b definiert, und wegen der Monotonie folgt aus $a \leq z_n \leq b$ auch $f^{-1}(a) \leq f^{-1}(z_n) \leq f^{-1}(b)$ für alle n. Die Folge $f^{-1}(z_n)$ ist daher beschränkt. Es sei $u := \underline{\lim} f^{-1}(z_n)$ und $v := \overline{\lim} f^{-1}(z_n)$. Ist z_{n_k} eine Teilfolge mit $\lim f^{-1}(z_{n_k}) = u$, so folgt aus der Stetigkeit von f, dass $f(u) = f(\lim f^{-1}(z_{n_k})) = \lim f(f^{-1}(z_{n_k})) = \lim z_{n_k} = z$, und ebenso $f(v) = z$.

Nun ist f injektiv, und man bekommt $u = v$. Aber das bedeutet gerade, dass die Folge $f^{-1}(z_n)$ konvergiert, und zwar gegen $u = v = f^{-1}(z)$. □

Beispiel 6.21 — Für jedes $n \in \mathbb{N}$ ist die Abbildung $X^n : \mathbb{R}_{\geq 0} \longrightarrow \mathbb{R}_{\geq 0}$ streng monoton wachsend und stetig. Nach dem vorigen Satz ist also auch die Umkehrfunktion $\sqrt[n]{-} : \mathbb{R}_{\geq 0} \longrightarrow \mathbb{R}_{\geq 0}$ streng monoton wachsend und stetig. Das ist ein neuer Beweis einer früheren Aussage.

6.4 Die Exponentialfunktion

Lemma 6.22 — *Für jede komplexe Zahl z mit $|z| < 1$ gilt:*

$$|\exp(z) - 1| < |z|(e-1).$$

Beweis. Für z mit $|z| < 1$ gilt:

$$|\exp(z) - 1| = |\sum_{n=1}^{\infty} \frac{z^n}{n!}| \leq |z| \sum_{n=1}^{\infty} \frac{|z|^{n-1}}{n!} < |z| \sum_{n=1}^{\infty} \frac{1}{n!} = |z|(e-1).$$

□

Satz 6.23 — *Die Exponentialfunktion* $\exp : \mathbb{C} \to \mathbb{C}$ *ist stetig. Insbesondere sind auch die Sinus- und die Kosinusfunktion stetig.*

Beweis. Es sei $a \in \mathbb{C}$ und $\varepsilon > 0$ vorgegeben. Es sei $\delta := \min\{1, \frac{\varepsilon}{(e-1)|\exp(a)|}\}$. Dann gilt für alle $z \in U_\delta(a)$ nach dem vorigen Lemma :

$$\begin{aligned} |\exp(z) - \exp(a)| &= |\exp(a)| \cdot |\exp(z-a) - 1| \\ &\leq |\exp(a)| \cdot |z-a| \cdot (e-1) \\ &< |\exp(a)|\delta(e-1) \leq \varepsilon. \end{aligned}$$

Der Rest ist klar. □

Wir schränken uns jetzt auf die reelle Exponentialfunktion ein. Es ist klar, dass für $x \in \mathbb{R}$ auch $\exp(x)$ eine reelle Zahl ist.

Satz 6.24 — *Für alle $x \in \mathbb{R}$ ist $\exp(x) > 0$. Die Funktion $\exp : \mathbb{R} \to \mathbb{R}_{>0}$ ist stetig, bijektiv und streng monoton wachsend. Es gilt $\lim_{x\to\infty} \exp(x) = \infty$ und $\lim_{x\to-\infty} \exp(x) = 0$.*

Beweis. Für positive x gilt $\exp(x) > 1 + x > 1$. Daraus erhält man für $x < 0$, dass $0 < \exp(x) < \frac{1}{1-x} < 1$. Aus diesen beiden Abschätzungen folgt erstens,

dass $\exp(x) > 0$ für alle $x \in \mathbb{R}$, und zweitens ergeben sich die Grenzwerte $\lim_{x\to\infty} \exp(x) = \infty$ und $\lim_{x\to-\infty} \exp(x) = 0$. Mit dem Zwischenwertsatz folgt, dass $\exp(\mathbb{R}) = \mathbb{R}_{>0}$. Außerdem gilt für beliebige reelle Zahlen mit $x < y$, dass $\exp(y) - \exp(x) = \exp(x)(\exp(y-x) - 1) > 0$. Demnach ist exp streng monoton wachsend. □

Definition 6.25 — Die Umkehrfunktion

$$\log : \mathbb{R}_{>0} \longrightarrow \mathbb{R}$$

der Exponentialfunktion heißt Logarithmusfunktion.

Häufig spricht man auch vom natürlichen Logarithmus (*logarithmus naturalis*) und schreibt $\ln(x)$ statt $\log(x)$.

Satz 6.26 — *Der Logarithmus* $\log : \mathbb{R}_{>0} \longrightarrow \mathbb{R}$ *ist stetig, bijektiv und streng monoton wachsend. Für alle* $a, b > 0$ *gilt:*

$$\log(ab) = \log(a) + \log(b), \text{ und } \log(a^{-1}) = -\log(a).$$

Beweis. Die ersten Aussagen folgen direkt aus Satz 6.20. Die beiden letzten Aussagen ergeben sich aus der Funktionalgleichung $\exp(u+v) = \exp(u)\exp(v)$ der Exponentialfunktion. □

Da $\lim_{x\to\infty} \exp(x) = \infty$, gilt auch $\lim_{x\to\infty} \log(x) = \infty$. Der nächste Satz zeigt aber einen großen Unterschied: Die Exponentialfunktion wächst schneller als jede Potenz, der Logarithmus dagegen langsamer als jede noch so hohe Wurzel:

Satz 6.27 — *Für alle* $n \in \mathbb{N}$ *gilt:*

1. $\lim_{x\to\infty} \frac{\exp(x)}{x^n} = \infty$.
2. $\lim_{x\to\infty} \frac{\log(x)}{\sqrt[n]{x}} = 0$.

Beweis. Es sei $n \in \mathbb{N}$. Für alle $x > 0$ gilt dann

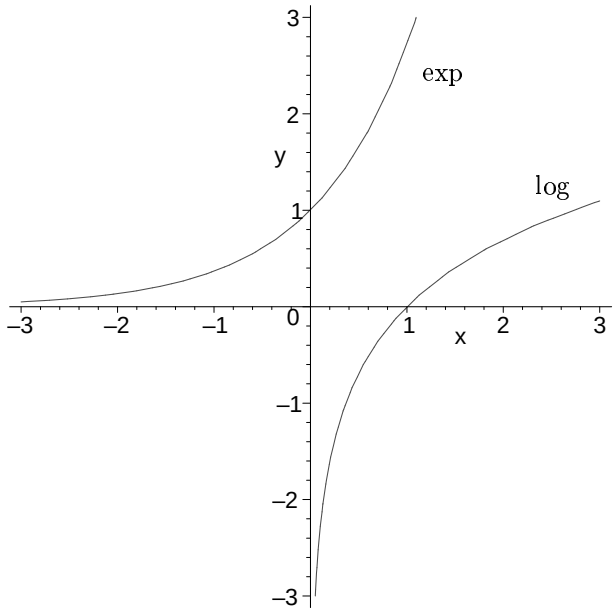
$$\frac{\exp(x)}{x^n} \geq \frac{x}{(n+1)!}.$$

Im Grenzübergang folgt hieraus die erste Behauptung. Wählt man $x = \log(y)$, so erhält man stattdessen $\frac{y}{\log(y)^n} \geq \frac{\log(y)}{(n+1)!}$. Zieht man die n-te Wurzel und geht zum Kehrwert über, so bleibt:

$$\frac{\log(y)}{\sqrt[n]{y}} \leq \sqrt[n]{\frac{(n+1)!}{\log(y)}}.$$

Im Grenzübergang $y \longrightarrow \infty$ folgt die zweite Behauptung. $\square$

Ein Bild der Exponential- und der Logarithmusfunktion:



Dass exp und log Umkehrfunktionen von einander sind, bedeutet algebraisch, dass

$$\exp(x) = y \qquad \Leftrightarrow \qquad \log(y) = x,$$

und geometrisch, dass die Graphen der beiden Funktionen bezüglich der Diagonalen $\{(x, y) \in \mathbb{R}^2 \mid x = y\}$ spiegelsymmetrisch zueinander liegen.

6.5 Sinus- und Kosinusfunktion. Die Zahl π

Aus den Eulerschen Formeln $\sin(z) = \frac{1}{2i}(\exp(zi) - \exp(-zi))$ und $\cos(z) = \frac{1}{2}(\exp(iz) + \exp(-iz))$ und der Stetigkeit der Exponentialfunktion folgt, dass die Sinus- und die Kosinusfunktion auf ganz $\mathbb{C}$ stetig sind. Da andererseits alle Koeffizienten in der Reihendarstellung der beiden Funktionen reell sind, ist klar, dass $\sin(x)$ und $\cos(x)$ reelle Zahlen sind, wenn das Argument x reell ist. In diesem Abschnitt wollen wir die stetigen, reellen Funktionen $\sin : \mathbb{R} \to \mathbb{R}$ und $\cos : \mathbb{R} \to \mathbb{R}$ näher betrachten. Wir wissen bereits, dass für alle $x \in \mathbb{R}$ die Beziehung $\sin(x)^2 + \cos(x)^2 = 1$ besteht. Sinus und Kosinus können also auf ganz $\mathbb{R}$ nur Werte im Intervall $[-1, 1]$ annehmen.

Lemma 6.28 — *Für alle $x \in [0, 2]$ gilt:*

$$x - \frac{x^3}{6} < \sin(x) < x$$

und

$$1 - \frac{x^2}{2} < \cos(x) < 1 - \frac{x^2}{2} + \frac{x^4}{24}.$$

Beweis. Für alle $x \in [0,2]$ und alle $n \in \mathbb{N}$ gilt die Beziehung

$$\frac{x^n}{n!} \geq \frac{x^{n+2}}{(n+2)!}.$$

Aus den Darstellungen

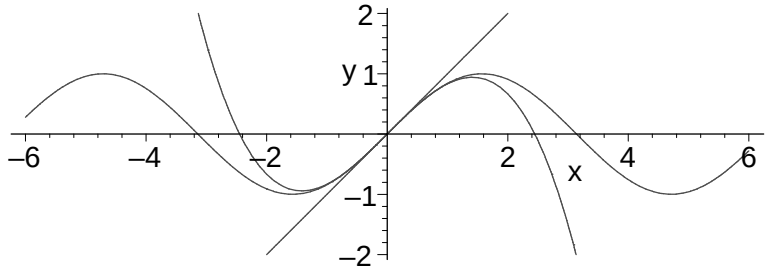
$$\begin{aligned} \sin(x) &= x - \left(\frac{x^3}{3!} - \frac{x^5}{5!}\right) - \left(\frac{x^7}{7!} - \frac{x^9}{9!}\right) - \dots \\ &= \left(x - \frac{x^3}{3!}\right) + \left(\frac{x^5}{5!} - \frac{x^7}{7!}\right) + \dots \end{aligned}$$

und

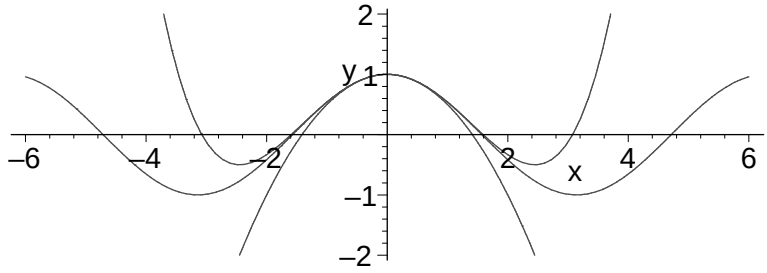
$$\begin{aligned} \cos(x) &= \left(1 - \frac{x^2}{2!}\right) + \left(\frac{x^4}{4!} - \frac{x^6}{6!}\right) + \dots \\ &= 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \left(\frac{x^6}{6!} - \frac{x^8}{8!}\right) - \dots \end{aligned}$$

folgt die Behauptung. □

Das Lemma in Bildern für die Sinusfunktion



und für die Kosinusfunktion



Folgerung 6.29 — *Es gilt:*

1. $\cos(x) > 0$ *für alle* $x \in [0, \sqrt{2}]$.
2. $\cos(\sqrt{6 - \sqrt{12}}) < 0$.

3. *Es gibt ein* $x \in (\sqrt{2}, \sqrt{6-\sqrt{12}})$ *mit* $\cos(x) = 0$.

Beweis. Offenbar gilt $1 - x^2/2 \geq 0$ für alle $x \in [0, \sqrt{2}]$, und aus dem vorigen Lemma folgt die erste Abschätzung. Das Polynom $1 - \frac{x^2}{2!} + \frac{x^4}{4!}$ hat bei $\sqrt{6-\sqrt{12}} \approx 1,592$ eine Nullstelle, und aus dem vorigen Lemma folgt die zweite Behauptung. Schließlich folgt aus $\cos(\sqrt{2}) > 0$ und $\cos(\sqrt{6-\sqrt{12}}) < 0$ die letzte Aussage nach dem Zwischenwertsatz. □

Definition 6.30 — Die Zahl

$$\pi := 2 \inf\{x \in [0,2] \mid \cos(x) = 0\}$$

heißt Kreiszahl oder Ludolphsche Zahl[22].

Lemma 6.31 — *Es gilt*

$$\cos\frac{\pi}{2} = 0 \quad \textit{und} \quad \sin\frac{\pi}{2} = 1.$$

Insbesondere gilt

$$\exp(\pi i/2) = i,\ \exp(\pi i) = -1,\ \exp(3\pi i/2) = -i,\ \exp(2\pi i) = 1.$$

Beweis. Nach Definition des Infimums und wegen Stetigkeit gilt $\cos(\pi/2) = 0$. Wegen $\sin(\alpha)^2 + \cos(\alpha)^2 = 1$ für alle $\alpha \in \mathbb{C}$, kann der Sinus in einer Nullstelle des Kosinus nur die Werte 1 oder -1 annehmen. Wir wissen aber nach Lemma 6.28, dass der Sinus auf dem Intervall $(0, 2]$ strikt positiv ist. Es folgt, dass $\sin(\pi/2) = 1$. Die restlichen Aussagen folgen aus den Eulerschen Formeln und $\exp(x+y) = \exp(x)\exp(y)$ für $x, y \in \mathbb{C}$. □

Lemma 6.32 — *Für alle* $z \in \mathbb{C}$ *gilt:*

1. $\sin(z + \frac{\pi}{2}) = \cos(z)$ *und* $\cos(z + \frac{\pi}{2}) = -\sin(z)$.
2. $\sin(z + \pi) = -\sin(z)$ *und* $\cos(z + \pi) = -\cos(z)$.
3. $\sin(z + 2\pi) = \sin(z)$ *und* $\cos(z + 2\pi) = \cos(z)$.

Sinus- und Kosinusfunktion sind also 2π*-periodisch.*

Beweis. Die erste Aussage folgt aus den Additionstheoremen für die trigonometrischen Funktionen und den Beziehungen $\cos(\pi/2) = 0$ und $\sin(\pi/2) = 1$. Die zweite und dritte Aussage folgen durch wiederholte Anwendung. □

[22] Ludolph van Ceulen, $*$ 28.1.1540 † 31.12.1610.

Folgerung 6.33 — *Für alle $x \in \mathbb{R}$ gilt:*

1. $\sin(x) = 0$ *genau dann, wenn* $x = k\pi$ *für ein* $k \in \mathbb{Z}$.

2. $\cos(x) = 0$ *genau dann, wenn* $x = \pi/2 + k\pi$ *für ein* $k \in \mathbb{Z}$.

Beweis. Der Sinus ist eine ungerade Funktion, aus $\sin(x) > 0$ für alle $x \in (0, \pi/2]$ folgt daher $\sin(x) < 0$ für alle $x \in [-\pi/2, 0)$. Die Periodizität des vorigen Lemmas liefert $\sin(x) \neq 0$ für alle $x \in \mathbb{R} \setminus \mathbb{Z}\pi$ und $\sin(x) = 0$ für $x \in \mathbb{Z}\pi$. Für den Kosinus schließt man ähnlich. □

Satz 6.34 — *Die Abbildung*

$$p : [0, 2\pi) \longrightarrow S^1 := \{z \in \mathbb{C} \,|\, |z| = 1\}, \quad \alpha \mapsto \exp(i\alpha),$$

ist stetig und bijektiv.

Beweis. Es ist klar, dass die Abbildung p stetig ist. Wir zeigen die Surjektivität. Es sei zunächst ein $z \in S^1$ mit $x = \mathrm{Re}(z) \geq 0$ und $y = \mathrm{Im}(z) \geq 0$ vorgegeben. Wegen $x^2 + y^2 = 1$ hat man $y \in [0, 1]$. Daher existiert nach dem Zwischenwertsatz ein $\alpha \in [0, \pi/2]$ mit $\sin(\alpha) = y$. Da $x^2 = 1 - y^2 = 1 - \sin(\alpha)^2 = \cos(\alpha)^2$, und da sowohl $x \geq 0$ wie $\cos(\alpha) \geq 0$, muss $\cos(\alpha) = x$ gelten. Das bedeutet aber gerade, dass $\exp(i\alpha) = \cos(\alpha) + i\sin(\alpha) = z$. Ist $z \in S^1$ beliebig, so gibt es eine Potenz i^m, $m \in \{0, 1, 2, 3\}$, derart, dass $\mathrm{Re}(z/i^m) \geq 0$ und $\mathrm{Im}(z/i^m) \geq 0$. Folglich existiert ein $\alpha \in [0, \pi/2]$ mit $z/i^m = \exp(i\alpha)$. Da $\exp(i\pi/2) = i$, folgt: $z = \exp(i(\alpha + m\pi/2))$.

Es bleibt die Injektivität zu zeigen. Angenommen, wir haben $\alpha, \alpha' \in [0, 2\pi)$ mit $\alpha < \alpha'$ und $\exp(i\alpha) = \exp(i\alpha')$. Für die Differenz $\delta := \alpha' - \alpha \in (0, 2\pi)$ gilt dann $\exp(i\delta) = 1$. Folglich ist $\sin(\delta) = 0$. Das lässt im angegebenen Intervall nur die Möglichkeit $\delta = \pi$. Aber $\exp(i\pi) = -1$, Widerspruch. □

Satz 6.35 — *Die Abbildung* $\exp : \mathbb{C} \to \mathbb{C}^*$ *ist ein surjektiver Gruppenhomomorphismus von der additiven Gruppe* $(\mathbb{C}, +)$ *in die multiplikative Gruppe* $(\mathbb{C}^*, \cdot)$. *Der Kern des Homomorphismus ist die Untergruppe* $2\pi i\mathbb{Z} \subset \mathbb{C}$.

Beweis. Wir wissen bereits, dass $\exp(z + w) = \exp(z) \cdot \exp(w)$ für alle $z, w \in \mathbb{C}$, und das bedeutet gerade, dass exp ein Gruppenhomomorphismus ist.

Es sei $z \in \mathbb{C}^*$ vorgegeben. Dann gibt es nach dem vorigen Satz ein $\alpha \in \mathbb{R}$ mit $\exp(i\alpha) = z/|z|$. Wir wählen $r = \log|z|$. Dann gilt $\exp(r + i\alpha) = \exp(r) \cdot \exp(i\alpha) = z$.

Angenommen, $\exp(z) = 1$. Da $|\exp(z)| = \exp(\mathrm{Re}(z))$, ist $z = iy$ für ein $y \in \mathbb{R}$. Es sei $k := \lfloor \frac{y}{2\pi} \rfloor \in \mathbb{Z}$. Dann ist $\alpha := y - 2\pi k \in [0, 2\pi)$, und es gilt

$\exp(\alpha i) = \exp(yi)/\exp(2\pi ik) = 1/1 = 1$. Aus dem vorigen Satz folgt, dass $\alpha = 0$, also $x = 2\pi ik$. □

Wir haben also gezeigt, dass sich jede komplexe Zahl z in der Form

$$z = r\exp(i\varphi)$$

mit reellen Zahlen $r \geq 0$ und φ darstellen lässt. Dabei ist r der Absolutbetrag von z. Für $r = 0$ ist φ völlig irrelevant, für $r \neq 0$ ist φ eindeutig bis auf die Addition ganzzahliger Vielfache von 2π. Die Restklasse

$$[\varphi] \in \mathbb{R}/2\pi\mathbb{Z}$$

heißt das Argument von z. Jede Klasse $[\varphi]$ hat einen eindeutigen Repräsentanten aus dem Intervall $[0, 2\pi)$. Diese Darstellung erlaubt eine geometrische Deutung der Multiplikation zweier komplexer Zahlen: Gilt $z = r\exp(i\varphi)$ und $z' = r'\exp(i\varphi')$, so folgt:

$$zz' = (rr')\exp(i(\varphi + \varphi')).$$

Man multipliziert also zwei komplexe Zahlen, indem man ihre Radien multipliziert und ihre Argumente addiert.

6.6 Kompakte Teilmengen von $\mathbb{C}$

Definition 6.36 — Eine Teilmenge $K \subset \mathbb{C}$ heißt kompakt, wenn jede Folge (z_n) in K eine Teilfolge (z_{n_k}) besitzt, die gegen ein $z \in K$ konvergiert.

Satz 6.37 — *Es sei $K \subset \mathbb{C}$. Dann sind die folgenden Eigenschaften äquivalent:*

1. *K ist abgeschlossen und beschränkt.*
2. *K ist kompakt.*

Beweis. 1. $\Rightarrow$ 2.: Es sei (z_n) eine Folge in K. Da K beschränkt ist, gibt es nach dem Satz von Bolzano-Weierstraß eine konvergente Teilfolge (z_{n_k}). Es sei $z = \lim z_{n_k}$. Es ist zu zeigen, dass $z \in K$. Nun gilt entweder $z = z_n$ für ein n, so dass $z \in K$, oder $z_n \neq z$ für alle n, dann ist z ein Häufungspunkt von K. Nach Voraussetzung ist K abgeschlossen. In jedem Falle gilt also $z \in K$.

2. $\Rightarrow$ 1.: Wäre K unbeschränkt, so gäbe es zu jedem $n \in \mathbb{N}$ ein $z_n \in K$ mit $|z_n| \geq n$. Dann gilt für jede Teilfolge (z_{n_k}) von (z_n), dass $|z_{n_k}| \to \infty$ für $k \to \infty$. Keine Teilfolge kann demnach konvergieren, im Widerspruch zur Annahme, dass K kompakt ist. Somit ist K beschränkt. Es sei ferner a ein Häufungspunkt von K. Nach Annahme existiert eine Folge (z_n) in K mit $\lim z_n = a$. Da K kompakt

ist, gibt es eine Teilfolge (z_{n_k}), die gegen ein $z \in K$ konvergiert. Es gilt aber auch $\lim z_{n_k} = z$. Das zeigt $a = z \in K$. Da a ein beliebiger Häufungspunkt war, ist K abgeschlossen. □

Lemma 6.38 — *Es sei $K \subset \mathbb{R}$ kompakt und nicht leer. Dann besitzt K ein Maximum und ein Minimum.*

Beweis. K ist beschränkt und nicht leer, also besitzt K Supremum und Infimum in $\mathbb{R}$. Es sei $M := \sup K$. Es gibt eine Folge $x_n \in K$ mit $\lim x_n = M$. Da K abgeschlossen ist, liegt M in K und ist damit nicht nur ein Supremum von K, sondern ein Maximum. Für das Minimum schließt man analog. □

Lemma 6.39 — *Es sei $K \subset \mathbb{C}$ kompakt und nicht leer und $f : K \to \mathbb{C}$ eine stetige Funktion. Dann ist auch $f(K)$ kompakt.*

Beweis. Es sei (z_n) eine Folge in $f(K)$. Für jedes n wählen wir ein $x_n \in K$ mit $f(x_n) = z_n$. Da K kompakt ist, hat die Folge (x_n) eine Teilfolge (x_{n_k}) mit $\lim x_{n_k} = x \in K$. Die Funktion f ist stetig. Es folgt daher: $\lim z_{n_k} = \lim f(x_{n_k}) = f(x) \in f(K)$. □

Satz 6.40 — *Es sei $K \subset \mathbb{C}$ kompakt und nicht leer. Jede stetige Funktion $f : K \to \mathbb{R}$ nimmt auf K ihr Maximum und Minimum an, d.h. es gibt Punkte $a \in K$ mit $f(a) = \inf\{f(z) \mid z \in K\}$ und $b \in K$ mit $f(b) = \sup\{f(z) \mid z \in K\}$.*

Beweis. Nach den beiden vorstehenden Lemmata ist $f(K) \subset \mathbb{R}$ kompakt und besitzt daher ein Maximum und ein Minimum. □

Ein wichtiger Spezialfall des Satzes sind stetige Funktionen $f : [a, b] \to \mathbb{R}$ auf abgeschlossenen Intervallen. Es gibt dann stets einen Punkt $x_0 \in [a, b]$ mit $f(x_0) \geq f(x)$ für alle $x \in [a, b]$. Die Aussage ist falsch, wenn das Intervall nicht abgeschlossen oder wenn f nicht stetig ist. Beispiele?

6.7 Fundamentalsatz der Algebra

Satz 6.41 *(Fundamentalsatz der Algebra) — Es sei $f \in \mathbb{C}[X]$ ein nichtkonstantes Polynom. Dann besitzt f in $\mathbb{C}$ eine Nullstelle.*

Beweis. Es sei $f = f_0 + f_1 X + \ldots + f_n X^n$ ein Polynom vom Grad $n \in \mathbb{N}$ mit komplexen Koeffizienten $f_i \in \mathbb{C}$ und $M := \inf\{|f(z)| \mid z \in \mathbb{C}\}$.
1. Schritt: Es gibt ein $a \in \mathbb{C}$ mit $|f(a)| = M$.

Indem wir gegebenfalls f durch f/f_n ersetzen, können wir erreichen, dass f normiert ist, also einen Leitkoeffizienten $f_n = 1$ hat. Wir setzen

$$R_1 := \max\{\sqrt[i]{2n|f_{n-i}|} \mid i = 1, \ldots, n\}$$

und

$$R_2 = \sqrt[n]{2(M+1)}.$$

Für alle z mit $|z| > R := \max\{R_1, R_2\}$ folgt

$$\begin{aligned} |f(z)| &= |z^n| \cdot \left|1 + f_{n-1}\frac{1}{z} + \ldots + f_0\frac{1}{z^n}\right| \\ &\geq |z^n| \cdot \left(1 - \sum_{i=1}^{n} \frac{|f_{n-i}|}{R^i}\right) \\ &\geq |z^n| \cdot \left(1 - \sum_{i=1}^{n} \frac{1}{2n}\right) \\ &= |z^n| \cdot \frac{1}{2} > (M+1). \end{aligned}$$

Wir betrachten jetzt die abgeschlossene Kreisscheibe

$$K := \{z \in \mathbb{C} \mid |z| \leq R\}$$

von Radius R. Nach dem gerade Bewiesenen gilt $|f(z)| > M + 1$ für alle $z \notin K$. Das bedeutet, dass $M = \inf\{|f(z)| \mid z \in K\}$. Aber die stetige Funktion $|f|$ muss ihr Minimum auf der kompakten Menge K annehmen, etwa im Punkt $a \in K$. Das zeigt $|f(a)| = M$.

2. Schritt: $M = 0$. Insbesondere ist a eine Nullstelle von f.

Wir führen die Annahme $M \neq 0$ wie folgt zum Widerspruch: Zunächst vereinfachen wir die Rechnung, indem wir zu dem Polynom

$$g(z) := f(z + a), \quad g(z) = g_0 + g_1 z + \ldots + g_n z^n,$$

übergehen. Dann hat $|g(z)|$ bei $z = 0$ ein globales Minimum $M = |g_0|$. Indem wir g durch g/g_0 ersetzen, können wir außerdem erreichen, dass $1 = g_0 = g(0) = M$. Es sei $m \in \mathbb{N}$ minimal mit $g_m \neq 0$. Wir haben dann:

$$g(z) = 1 + g_m z^m + g_{m+1} z^{m+1} + \ldots + g_n z^n.$$

Wir wählen $w \in \mathbb{C}$ so, dass $w^m = -1/g_m$, und setzen $N := \sum_{i=m+1}^{n} |g_i w^i|$. Für $\varepsilon \in \mathbb{R}$ mit $0 < \varepsilon < \min\{1, 1/N\}$ folgt:

$$g(\varepsilon w) = 1 - \varepsilon^m + \varepsilon^{m+1} \sum_{i=m+1}^{n} g_i w^i \varepsilon^{i-m-1},$$

also

$$|g(\varepsilon w)| \leq (1-\varepsilon^m) + \varepsilon^{m+1}N \leq 1 - \varepsilon^m(1-\varepsilon N) < 1.$$

Das ist ein Widerspruch zur Annahme, dass $|g|$ das Minimum 1 hat. □

Folgerung 6.42 — *Jedes komplexe Polynom f vom Grad $n \in \mathbb{N}$ zerfällt in ein Produkt*

$$f(X) = a(X - z_1)(X - z_2)\cdots(X - z_n)$$

mit $a \in \mathbb{C}^*$ und $z_i \in \mathbb{C}$ für $i = 1, \ldots, n$.

Beweis. Nach dem Fundamentalsatz besitzt f eine Nullstelle z. Mit Polynomdivision sieht man, dass es ein Polynom g vom Grad $n-1$ gibt so, dass $f(X) = (X - z)g(X)$. Durch Induktion über den Grad findet an eine Zerlegung der behaupteten Form für f. □

§7 Differenzierbare Funktionen

7.1 Differentialquotienten und Rechenregeln

Definition 7.1 — Es sei $M \subset \mathbb{R}$ und $a \in M$. Eine Funktion $f : M \to \mathbb{R}$ heißt in a differenzierbar, wenn

1. a ein Häufungspunkt von f ist, und
2. der Grenzwert
$$\lim_{x \to a} \frac{f(x) - f(a)}{x - a}$$
existiert.

In diesem Falle wird der Grenzwert mit $f'(a)$ bezeichnet und heißt die Ableitung oder der Differentialquotient von f in a. Die Funktion f heißt differenzierbar, wenn f in allen Punkten $a \in M$ differenzierbar ist. Dann versteht man unter der Ableitung von f die Funktion $f' : M \to \mathbb{R}$, $a \mapsto f'(a)$.

Bemerkung 7.2 — Es sei $f : M \to \mathbb{R}$ in a differenzierbar. Die Funktion

$$\Delta_a f : M \to \mathbb{R}, x \mapsto \begin{cases} \frac{f(x)-f(a)}{x-a} & \text{für } x \neq a, \text{ und} \\ f'(a) & \text{für } x = a, \end{cases}$$

heißt der Differenzenquotient von f im Punkt a. Nach Definition der Differenzierbarkeit ist die Funktion $\Delta_a f$ im Punkt a stetig, und es gilt für alle $x \in M$:

$$f(x) = f(a) + \Delta_a f(x) \cdot (x - a). \tag{7.1}$$

Umgekehrt könnte man so auch die Differenzierbarkeit definieren: Eine Funktion $f : M \to \mathbb{R}$ heißt in $a \in M$ differenzierbar, wenn a ein Häufungspunkt von M ist und wenn es eine in a stetige Funktion $\Delta : M \to \mathbb{R}$ mit der Eigenschaft $f(x) = f(a) + \Delta(x)(x - a)$ für alle $x \in M$ gibt. In diesem Falle gilt für $x \neq a$

$$\frac{f(x) - f(a)}{x - a} = \Delta(x). \tag{7.2}$$

Da die rechte Seite in a stetig ist, existiert der Grenzwert der linken Seite für $x \to a$, und es gilt $f'(a) = \Delta(a)$ und $\Delta = \Delta_a f$.

Lemma 7.3 — *Ist $f : M \to \mathbb{R}$ in a differenzierbar, so ist f in a stetig.*

Beweis. In der Gleichung (7.1) ist die rechte Seite stetig in a, also auch die linke. □

Beispiele 7.4 — 1. Konstante Funktionen $c : \mathbb{R} \to \mathbb{R}$, $x \mapsto c$, sind differenzierbar mit der Ableitung $c' = 0$.
2. Die identische Abbildung $X : \mathbb{R} \to \mathbb{R}$, $x \mapsto x$, ist differenzierbar mit der Ableitung $X' = 1$.
3. Der Absolutbetrag $|-| : \mathbb{R} \to \mathbb{R}$ ist in 0 nicht differenzierbar, denn

$$\lim_{x \searrow 0} \frac{|x| - |0|}{x - 0} = 1, \text{ aber } \lim_{x \nearrow 0} \frac{|x| - |0|}{x - 0} = -1.$$

Satz 7.5 *(Summenregel) — Sind $f, g : M \to \mathbb{R}$ in $a \in M$ differenzierbar, so ist auch $f + g$ in a differenzierbar, und es gilt*

$$(f + g)'(a) = f'(a) + g'(a).$$

Beweis. Wenn die Voraussetzungen des Satzes erfüllt sind, ist a sicherlich ein Häufungspunkt. Für alle $x \in M$, $x \neq a$ gilt:

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

Nach den Rechenregeln für Grenzwerte existiert der Grenzwert auf der linken Seite, wenn die Grenzwerte auf der rechten Seite existieren, und ist gleich der Summe dieser Grenzwerte. Daraus folgt die Behauptung. □

Satz 7.6 *(Produktregel) — Sind $f, g : M \to \mathbb{R}$ in $a \in M$ differenzierbar, so ist auch fg in a differenzierbar, und es gilt*

$$(fg)'(a) = f'(a)g(a) + f(a)g'(a).$$

Beweis. Für alle $x \in M$, $x \neq a$, gilt:

$$(fg)(x) - (fg)(a) = (f(x) - f(a))g(x) + f(a)(g(x) - g(a)),$$

also

$$\frac{(fg)(x) - (fg)(x)}{x - a} = \frac{f(x) - f(a)}{x - a} g(x) + f(a) \frac{g(x) - g(a)}{x - a}.$$

Der Grenzwert der rechten Seite für $x \to a$ existiert, weil f und g differenzierbar sind und weil g nach Lemma 7.3 in a stetig ist. Für $x \to a$ erhält man $(fg)'(a) = f'(a)g(a) + f(a)g'(a)$. □

Satz 7.7 *(Quotientenregel) — Sind $f, g : M \to \mathbb{R}$ in $a \in M$ differenzierbar und ist $g(a) \neq 0$, so existiert ein $\varepsilon > 0$ so, dass $g(x) \neq 0$ für alle $x \in M \cap U_\varepsilon(a)$. Weiterhin ist die Funktion*

$$\frac{f}{g} : M \cap U_\varepsilon(a) \longrightarrow \mathbb{R}$$

in a differenzierbar und hat die Ableitung

$$\left(\frac{f}{g}\right)'(a) = \frac{f'(a)g(a) - f(a)g'(a)}{g(a)^2}.$$

Beweis. Wir betrachten zunächst den Fall $f = 1$. Da g stetig ist, gibt es nach Satz 6.7 ein $\varepsilon > 0$ derart, dass $g(x) \neq 0$ für alle $x \in M \cap U_\varepsilon(a)$. Für alle $x \in M \cap U_\varepsilon(a)$, $x \neq a$, gilt:

$$\frac{\frac{1}{g(x)} - \frac{1}{g(a)}}{x - a} = -\frac{g(x) - g(a)}{x - a} \cdot \frac{1}{g(x)g(a)}.$$

Der Grenzwert auf der rechten Seite existiert, weil g in a differenzierbar und daher insbesondere in a stetig ist. Demnach ist die Funktion $1/g$ in a differenzierbar, und es gilt

$$\left(\frac{1}{g}\right)'(a) = -\frac{g'(a)}{g(a)^2}.$$

Die Aussage für allgemeines f folgt hieraus und aus der Produktregel. □

Folgerung 7.8 — *1. Ist $f = f_0 + f_1X + f_2X^2 + \ldots + f_nX^n$ ein Polynom, so ist $f : \mathbb{R} \to \mathbb{R}$ differenzierbar und hat die Ableitung*

$$f' = f_1 + 2f_2X + \ldots + nf_nX^{n-1}.$$

2. Rationale Funktionen sind auf ihrem Definitionsbereich differenzierbar.

Beweis. Zu 1.: Konstante Polynome sind differenzierbar mit Ableitung 0, und die identische Funktion X ist differenzierbar mit Ableitung 1. Durch vollständige Induktion und mit Hilfe der Produktregel folgt, dass X^n für jedes $n \in \mathbb{N}$ differenzierbar ist und die Ableitung $(X^n)' = nX^{n-1}$ hat. Die Behauptung für beliebige Polynome folgt hieraus und aus der Summenregel.
Zu 2.: Die Aussage folgt aus 1. und der Quotientenregel. □

Satz 7.9 *(Kettenregel) — Es sei $f : M \to M'$ in a und $g : M' \to \mathbb{R}$ in $f(a)$ differenzierbar. Dann ist auch die zusammengesetzte Funktion $g \circ f : M \to \mathbb{R}$ in a differenzierbar, und es gilt*

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

Man spricht in diesem Zusammenhang häufig von der äußeren Ableitung $g'(f(a))$ und der inneren Ableitung $f'(a)$.

Beweis. Der folgende Beweis ist *falsch*!
„Für alle $x \in M$, $x \neq a$ gilt:

$$\frac{g(f(x)) - g(f(a))}{x - a} = \frac{g(f(x)) - g(f(a))}{f(x) - f(a)} \cdot \frac{f(x) - f(a)}{x - a}.$$

Für $x \to a$ geht $f(x) \to f(a)$. Daher existieren die Grenzwerte auf der rechten Seite, und es folgt:

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

was zu zeigen war."

Das Problem ist, dass wir bei dieser Vorgehensweise unter Umständen durch 0 teilen. Wir haben zwar sorgfältig den Fall $x = a$ ausgeschlossen, wir wissen aber nicht, ob nicht $f(x) = f(a)$. Tatsächlich bricht der Beweis schon zusammen, wenn wir für f die konstante Funktion $f = 1$ wählen.

Korrekt kann man den Satz wie folgt beweisen:

Mit dem in $f(a)$ stetigen Differenzenquotienten $\Delta_{f(a)}g : M' \longrightarrow \mathbb{R}$ schließt man wie folgt: Für alle $y \in M'$ gilt

$$g(y) = g(f(a)) + (y - f(a))\ \Delta_{f(a)}g(y),$$

also speziell für alle $x \in M \setminus \{a\}$

$$g(f(x)) = g(f(a)) + (f(x) - f(a))\ \Delta_{f(a)}g(f(x)),$$

und somit

$$\frac{g(f(x)) - g(f(a))}{x - a} = \frac{f(x) - f(a)}{x - a}(\Delta_{f(a)}g \circ f)(x).$$

Der Grenzwert der rechten Seite für $x \to a$ existiert, weil f in a differenzierbar und $\Delta_{f(a)}g \circ f$ in a stetig ist, und man erhält

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

□

Es sei nun I ein Intervall und $f : I \to J$ stetig und bijektiv. Angenommen, sowohl f also auch die Umkehrfunktion $f^{-1} : J \to I$ sind differenzierbar. Dann hängen die Ableitungen von f und f^{-1} eng miteinander zusammen. Für alle $x \in I$ gilt $f^{-1}(f(x)) = x$, und aus der Kettenregel folgt $(f^{-1})'(f(x)) \cdot f'(x) = 1$, also $(f^{-1})'(y) = 1/f'(f^{-1}(y))$ für alle $y \in J$. Wir haben allerdings die Differenzierbarkeit von f^{-1} in dieser Herleitung vorausgesetzt, und müssen sie erst noch beweisen.

Satz 7.10 *(Ableitung inverser Funktionen) — Es sei $I \subset \mathbb{R}$ ein Intervall und $f : I \to J$ eine bijektive Funktion. Weiter sei f in $a \in I$ differenzierbar und f^{-1} in $b := f(a) \in J$ stetig. Gilt $f'(a) \neq 0$, so ist die inverse Funktion $f^{-1} : J \to I$ in b differenzierbar und es gilt*

$$(f^{-1})'(b) = \frac{1}{f'(f^{-1}(b))}.$$

Beweis. Für alle $x \in I \setminus \{a\}$ gilt

$$\frac{f(x) - f(a)}{x - a} = \Delta_a f(x). \tag{7.3}$$

Da f bijektiv ist, ist die linke Seite stets $\neq 0$, und wir können zum Kehrwert übergehen. Für alle $y \in J \setminus \{b\}$ und $x := f^{-1}(y)$ gilt folglich

$$\frac{f^{-1}(y) - f^{-1}(b)}{y - b} = \frac{x - a}{f(x) - f(a)} = \frac{1}{\Delta_a f(f^{-1}(y))}. \tag{7.4}$$

Nun sind f^{-1} in b und $\Delta_a f$ in $f^{-1}(b)$ stetig, und $\Delta_a f(f^{-1}(b)) = f'(a) \neq 0$. Folglich ist die rechte Seite von (7.4) stetig in a, und der Grenzwert der linken Seite für $y \to b$ existiert und hat den Wert $1/\Delta_a f(f^{-1}(b))$. Das bedeutet: f^{-1} ist in b differenzierbar mit Ableitung $1/f'(a)$. □

Bezeichnungen 7.11 (Höhere Ableitungen) — Eine Funktion $f : M \to \mathbb{R}$ heißt einmal differenzierbar, wenn sie differenzierbar ist. Wir schreiben $f^{(0)} := f$, $f^{(1)} := f'$. Weiter ist f k-mal differenzierbar, $k \in \mathbb{N}$, wenn f differenzierbar ist, und wenn $f' : M \to \mathbb{R}$ selbst $(k-1)$-mal differenzierbar ist. Die k-te Ableitung ist $f^{(k)} := (f')^{(k-1)}$. Ferner heißt f k-mal stetig differenzierbar, wenn f k-mal differenzierbar und $f^{(k)}$ stetig ist. Man schreibt

$$\mathcal{C}^0(M) := \{f : M \to \mathbb{R} \mid f \text{ ist stetig}\}$$

und

$$\mathcal{C}^k(M) := \{f : M \to \mathbb{R} \mid f \text{ ist } k\text{-mal stetig differenzierbar}\},$$

sowie

$$\mathcal{C}^\infty(M) := \bigcap_{k \geq 0} \mathcal{C}^k(M).$$

Bemerkung 7.12 — Aus der Summenregel folgt, dass $\mathcal{C}^\alpha(M)$ für alle $\alpha \in \mathbb{N}_0 \cup \{\infty\}$ ein reeller Vektorraum ist. Der folgende Satz zeigt außerdem, dass $\mathcal{C}^\alpha(M)$ sogar eine $\mathbb{R}$-Algebra ist.

Satz 7.13 *(Leibnizregel) — Es seien $f, g : M \to \mathbb{R}$ k-mal differenzierbare Funktionen. Dann ist auch das Produkt fg k-mal differenzierbar, und es gilt*

$$(fg)^{(k)} = \sum_{i=0}^{k} \binom{k}{i} f^{(i)} g^{(k-i)}.$$

Beweis. Durch Induktion nach k mit Hilfe der Produktregel. □

7.2 Die Ableitungen der Standardfunktionen

Beispiel 7.14 (Die Exponentialfunktion und der Logarithmus) — Für alle reellen Zahlen x, $|x| < 1$, gilt

$$|\exp(x) - 1 - x| = \left| \sum_{n \geq 2} \frac{x^n}{n!} \right| < |x|^2 \sum_{n \geq 2} \frac{1}{n!} = |x|^2 (e-2).$$

Es folgt für alle x mit $x \neq 0$ und $|x| < 1$:

$$\left| \frac{\exp(x) - 1}{x} - 1 \right| < |x|(e-2).$$

Für $x \to 0$ geht die rechte Seite gegen 0. Das zeigt:

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

Für beliebige $a \in \mathbb{R}$ folgt:

$$\begin{aligned} \lim_{x \to a} \frac{\exp(x) - \exp(a)}{x - a} &= \exp(a) \lim_{x \to a} \frac{\exp(x-a) - 1}{x - a} \\ &= \exp(a) \lim_{h \to 0} \frac{\exp(h) - 1}{h} = \exp(a). \end{aligned}$$

Die Exponentialfunktion ist also differenzierbar mit der Ableitung

$$\exp'(x) = \exp(x) \quad \text{für alle } x \in \mathbb{R}.$$

Mit Satz 7.10 über die Ableitung der inversen Funktion können wir auch den Logarithmus behandeln. Da $\exp(x) \neq 0$ für alle $x \in \mathbb{R}$, folgt für alle $y \in \mathbb{R}_{>0}$ und $x = \log(y)$: Die Logarithmusfunktion ist in y differenzierbar und es gilt:

$$\log'(y) = \frac{1}{\exp(x)} = \frac{1}{y}.$$

Beispiel 7.15 (Potenz- und Wurzelfunktionen) — Es sei $q \in \mathbb{N}$, $p \in \mathbb{Z}$, und $r := p/q$. Nach den Rechenregeln für Logarithmus und Exponentialfunktion gilt für alle $x > 0$

$$(\exp(r \log(x)))^q = \exp(qr \log(x)) = \exp(p \log(x)) = \exp(\log(x^p)) = x^p$$

und somit

$$\exp(r \log(x)) = \sqrt[q]{x^p} = x^r.$$

Für beliebiges $r \in \mathbb{R}$ ist x^r als Grenzwert $\lim x^{r_n}$ für eine beliebige Folge (r_n) rationaler Zahlen mit $r_n \to r$ definiert. Da $\exp(\alpha \log(x))$ bei festem x als Funktion von α stetig ist, erhält man die Identität

$$x^\alpha = \exp(\alpha \log(x)) \tag{7.5}$$

für jedes $\alpha \in \mathbb{R}$. Die Kettenregel liefert nun die folgenden Aussagen:

1. Für $\alpha \in \mathbb{R}$ ist die Funktion $X^\alpha : \mathbb{R}_{>0} \to \mathbb{R}$, $x \mapsto x^\alpha$, differenzierbar und $(X^\alpha)' = \alpha X^{\alpha-1}$.

2. Für $b > 0$ ist die Funktion $b^X : \mathbb{R} \to \mathbb{R}$, $x \mapsto b^x$, differenzierbar und $(b^X)' = \log(b) b^X$.

3. Die Funktion $X^X : \mathbb{R}_{>0} \to \mathbb{R}$, $x \mapsto x^x$, ist differenzierbar und $(X^X)' = (\log(X) + 1) X^X$.

Beispiel 7.16 (Die trigonometrischen Funktionen) — Für alle $x \in (0, 2)$ gelten nach Lemma (6.28) die Ungleichungen

$$-\frac{x}{2} \leq \frac{\cos(x) - 1}{x} \leq 0$$

und

$$1 - \frac{x^2}{6} \leq \frac{\sin(x)}{x} \leq 1.$$

Im Grenzübergang folgt:

$$\lim_{x \searrow 0} \frac{\cos(x) - 1}{x} = 0 \quad \text{und} \quad \lim_{x \searrow 0} \frac{\sin(x)}{x} = 1.$$

Da der Kosinus eine gerade und der Sinus eine ungerade Funktion ist, gelten dieselben Aussagen auch für den linksseitigen Grenzwert.

Es sei nun $a \in \mathbb{R}$ beliebig. Mit dem Additionstheorem für Sinus und Kosinus, angewendet auf das Argument $x = a + (x - a)$, folgt:

$$\begin{aligned} \lim_{x \to a} \frac{\sin(x) - \sin(a)}{x - a} &= \sin(a) \lim_{x \to a} \frac{\cos(x - a) - 1}{x - a} + \cos(a) \lim_{x \to a} \frac{\sin(x - a)}{x - a} \\ &= \cos(a) \end{aligned}$$

und

$$\begin{aligned} \lim_{x \to a} \frac{\cos(x) - \cos(a)}{x - a} &= \cos(a) \lim_{x \to a} \frac{\cos(x - a) - 1}{x - a} - \sin(a) \lim_{x \to a} \frac{\sin(x - a)}{x - a} \\ &= -\sin(a). \end{aligned}$$

Das zeigt: Die Funktionen sin und cos sind auf ganz $\mathbb{R}$ differenzierbar, und es gilt

$$\sin'(x) = \cos(x) \quad \text{und} \quad \cos'(x) = -\sin(x).$$

Die Ableitung der Tangensfunktion ergibt sich jetzt leicht mit Hilfe der Quotientenregel: tan ist auf dem ganzen Definitionsgebiet differenzierbar und es gilt

$$\tan'(x) = \frac{1}{\cos(x)^2}.$$

Beispiel 7.17 (Die Bogenfunktionen) —
1. Da $\sin'(x) = \cos(x) = \sqrt{1-\sin(x)^2} \neq 0$ für alle $x \in (-\pi/2, \pi/2)$, ist die Arkussinusfunktion in allen Punkten $y \in (-1,1)$ differenzierbar, aber nicht in $y = -1$ oder $y = 1$. Und mit der Bezeichnung $x = \arcsin(y)$ gilt:

$$\arcsin'(y) = \frac{1}{\sin'(x)} = \frac{1}{\sqrt{1-\sin(x)^2}} = \frac{1}{\sqrt{1-y^2}}.$$

2. Da $\cos'(x) = -\sin(x) = -\sqrt{1-\cos(x)^2} \neq 0$ für alle $x \in (0,\pi)$, ist die Arkuskosinusfunktion in allen Punkten $y \in (-1,1)$ differenzierbar, aber nicht in $y = -1$ oder $y = 1$. Und mit der Bezeichnung $x = \arccos(y)$ gilt:

$$\arccos'(y) = \frac{1}{-\sin(x)} = -\frac{1}{\sqrt{1-\cos(x)^2}} = -\frac{1}{\sqrt{1-y^2}}.$$

3. Da $\tan'(x) = \frac{1}{\cos(x)^2} = 1 + \tan(x)^2 \neq 0$ für alle $x \in (-\pi/2, \pi/2)$, ist die Arkustangensfunktion in allen Punkten $y \in \mathbb{R}$ differenzierbar, und mit der Bezeichnung $x = \arctan(y)$ folgt:

$$\arctan'(y) = \frac{1}{\tan'(x)} = \frac{1}{1+\tan(x)^2} = \frac{1}{1+y^2}.$$

7.3 Der Mittelwertsatz

Definition 7.18 — Es sei $I \subset \mathbb{R}$ ein Intervall und $a \in I$, aber keine Intervallgrenze. Eine Funktion $f : I \to \mathbb{R}$ hat bei a ein lokales Maximum (bzw. Minimum), wenn es ein $\varepsilon > 0$ gibt derart, dass für alle $x \in I \cap U_\varepsilon(a)$ gilt $f(x) \leq f(a)$ (bzw. $f(x) \geq f(a)$). Wenn einer der beiden Fälle eintritt, sagt man, f habe bei a ein lokales Extremum.

Satz 7.19 — *Es sei $f : I \to \mathbb{R}$ differenzierbar und $a \in I$ ein lokales Extremum von f. Dann gilt $f'(a) = 0$.*

Beweis. Es sei a ein lokales Maximum. (Den Fall eines Minimums behandelt man analog.) Dann gilt:

$$f'(a) = \lim_{x \searrow a} \frac{f(x)-f(a)}{x-a} \leq 0,$$

weil für jedes $x > a$ in einer kleinen Umgebung von a der Differenzenquotient negativ ist. Ebenso sieht man

$$f'(a) = \lim_{x \nearrow a} \frac{f(x)-f(a)}{x-a} \geq 0.$$

Zusammengenommen folgt die Behauptung. □

Vorsicht: Die Umkehrung des Satzes ist im allgemeinen falsch!

Satz 7.20 *(Satz von Rolle) — Es sei $f : [a,b] \to \mathbb{R}$ stetig und $f|_{(a,b)}$ differenzierbar. Gilt $f(a) = f(b)$, so gibt es ein $c \in (a,b)$ mit $f'(c) = 0$.*

Beweis. Da f stetig und $[a,b]$ kompakt ist, gibt es Punkte $c_1, c_2 \in [a,b]$ mit $f(c_1) = \max\{f(x) \mid x \in [a,b]\}$ und $f(c_2) = \min\{f(x) \mid x \in [a,b]\}$.

Falls $f(c_1) > f(a) = f(b)$, so ist c_1 ein lokales Maximum, und es gilt $f'(c) = 0$ nach dem obigen Satz.

Falls $f(c_2) < f(a) = f(b)$, so ist c_2 ein lokales Minimum, und es gilt $f'(c) = 0$ nach dem obigen Satz.

Falls dagegen $f(c_1) = f(a) = f(b) = f(c_2)$, so ist f auf dem ganzen Intervall konstant, und für jeden Punkt $c \in (a,b)$ gilt $f'(c) = 0$. □

Satz 7.21 *(Verallgemeinerter Mittelwertsatz) — Es seien $f, g : [a,b] \to \mathbb{R}$ stetige Funktionen, die auf (a,b) differenzierbar sind. Weiterhin gelte $g(a) \neq g(b)$ und $g'(x) \neq 0$ für alle $x \in (a,b)$. Dann gibt es ein $c \in (a,b)$ mit*

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

Beweis. Wir definieren für $x \in [a,b]$ die Funktion

$$F(x) = f(x) - \frac{f(b) - f(a)}{g(b) - g(a)}(g(x) - g(a)).$$

Dann gilt $F(a) = f(a) = F(b)$ und

$$F'(x) = f'(x) - \frac{f(b) - f(a)}{g(b) - g(a)} g'(x).$$

Nach dem Satz von Rolle existiert ein $c \in (a,b)$ mit $F'(c) = 0$, und dies ist äquivalent zur Behauptung. □

Speziell für $g(x) = x$ erhält man den gewöhnlichen Mittelwertsatz:

Satz 7.22 *(Mittelwertsatz) — Es sei $f : [a,b] \to \mathbb{R}$ stetig und auf (a,b) differenzierbar. Dann gibt es ein $c \in (a,b)$ mit*

$$\frac{f(b) - f(a)}{b - a} = f'(c).$$

Folgerung 7.23 *Es sei I ein Intervall und $f : I \to \mathbb{R}$ differenzierbar.*

1. *Ist $f'(x) = 0$ für alle $x \in I$, so ist f konstant.*
2. *Ist $f'(x) > 0$ für alle $x \in I$, so ist f streng monoton wachsend.*
3. *Ist $f'(x) < 0$ für alle $x \in I$, so ist f streng monoton fallend.*

Beweis. Es seien $a, b \in I$, $a < b$. Dann gibt es ein $c \in (a, b)$ mit

$$f(b) - f(a) = (b - a) \cdot f'(c).$$

Da $(b - a) > 0$, ist die linke Seite $= 0$, > 0 bzw. < 0 je nachdem, ob $f'(c) =$, > 0 oder < 0. □

Folgerung 7.24 — *Es sei I ein Intervall und $f : I \to \mathbb{R}$ differenzierbar. Gibt es $m, M \in \mathbb{R}$ mit $m \leq f'(x) \leq M$ für alle $x \in I$, so gilt für alle $a, b \in I$, $a < b$:*

$$m(b - a) \leq f(b) - f(a) \leq M(b - a).$$

Insbesondere gilt $|f(b) - f(a)| \leq (b - a) \sup\{|f'(x)| \,|\, x \in I\}$.

Satz 7.25 — *Es sei $f \in \mathcal{C}^2(I)$ und $a \in I$ keine Intervallgrenze von I. Gilt $f'(a) = 0$ und $f''(a) < 0$ (bzw. $f''(a) > 0$), dann hat f in a ein lokales Maximum (bzw. Minimum).*

Beweis. Wir betrachten den Fall $f''(a) > 0$. (Den Fall $f''(a) < 0$ behandelt man analog.) Weil f'' stetig ist, gibt es ein solches $\varepsilon > 0$, dass $f''(x) > 0$ für jedes $x \in I$ mit $|a - x| < \varepsilon$. Für ein solches x existiert nach dem Mittelwertsatz ein c zwischen a und x mit $f(x) - f(a) = f'(c)(x - a)$. Wieder nach dem Mittelwertsatz existiert ein d zwischen a und c mit $f'(c) = f'(c) - f'(a) = (c - a)f''(d)$. Zusammengenommen zeigt dies:

$$f(x) - f(a) = (x - a)(c - a)f''(d).$$

Die rechte Seite ist nicht negativ, weil x und c auf derselben Seite von a liegen und weil $f''(d) > 0$ ist. Folglich ist a ein lokales Minimum. □

Satz 7.26 *(Regel von de l'Hospital[23])* — *Es sei I ein Intervall, $a \in I$, und $f, g : I \to \mathbb{R}$ stetige Funktionen, die auf $I \setminus \{a\}$ differenzierbar sind. Es sei $g(a) = 0 = f(a)$, aber $g(x) \neq 0$, $g'(x) \neq 0$ für alle $x \in I \setminus \{a\}$. Existiert nun der Limes*

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

im eigentlichen oder uneigentlichen Sinne, so existiert auch der Limes

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

und die beiden Grenzwerte sind gleich.

[23] Guillaume François Antoine Marquis de l'Hospital, $*$ 1661 † 2.2.1704.

Beweis. Für alle $x \in I$, $x \neq a$, existiert nach dem verallgemeinerten Mittelwertsatz ein c_x strikt zwischen x und a mit

$$\frac{f(x)}{g(x)} = \frac{f(x) - f(a)}{g(x) - g(a)} = \frac{f'(c_x)}{g'(c_x)}.$$

Für $x \to a$ geht auch $c_x \to a$. Jetzt folgt die Behauptung. □

Satz 7.27 *(Regel von de l'Hospital II) — Es sei $b > 0$, und $f, g : [b, \infty) \to \mathbb{R}$ seien differenzierbare Funktionen, mit $g(x) \neq 0$, $g'(x) \neq 0$ für alle $x \in [b, \infty)$. Ferner gelte $\lim_{x\to\infty} f(x) = 0$ und $\lim_{x\to\infty} g(x) = 0$. Existiert nun der Limes*

$$\lim_{x\to\infty} \frac{f'(x)}{g'(x)}$$

im eigentlichen oder uneigentlichen Sinne, so existiert auch der Limes

$$\lim_{x\to\infty} \frac{f(x)}{g(x)},$$

und die beiden Grenzwerte sind gleich.

Beweis. Man führt diese Regel auf den obigen Satz zurück, indem man auf $[0, 1/b]$ die Funktion $\tilde{f}(x) = f(1/x)$ für $x \neq 0$, und $\tilde{f}(0) = 0$, und entsprechend $\tilde{g}$ einführt und die Regel von de l'Hospital auf $\tilde{f}$ und $\tilde{g}$ anwendet. □

Satz 7.28 *(Regel von de l'Hospital III) — Es seien $f, g : [b, \infty) \to \mathbb{R}$ differenzierbare Funktionen, mit $g(x) \neq 0$, $g'(x) \neq 0$ für alle $x \in [b, \infty)$. Ferner gelte $\lim_{x\to\infty} f(x) = \infty$ und $\lim_{x\to\infty} g(x) = \infty$. Existiert nun der Limes*

$$\lim_{x\to\infty} \frac{f'(x)}{g'(x)}$$

in $\mathbb{R}$, so existiert auch der Limes

$$\lim_{x\to\infty} \frac{f(x)}{g(x)},$$

und die beiden Grenzwerte sind gleich.

Beweis. Es sei $c := \lim_{x\to\infty} f'(x)/g'(x)$. Zu vorgegebenem $\varepsilon > 0$ existiert demnach ein x_0 so, dass für alle $\xi \geq x_0$ gilt $|f'(\xi)/g'(\xi) - c| < \varepsilon$.

Weil nach Voraussetzung $\lim_{x\to\infty} f(x) = \infty \lim_{x\to\infty} g(x)$, gibt es ein $x_1 \geq b$ derart, dass für alle $x \geq x_1$ gilt $f(x) \neq 0$ und $|f(x_0)/f(x)| < 1$, sowie $|g(x_0)/g(x)| < 1$. Nach dem verallgemeinerten Mittelwertsatz gibt es zu jedem $x > \max\{x_0, x_1\}$ ein $\xi \in (x_0, x)$ mit

$$\frac{f'(\xi)}{g'(\xi)} = \frac{f(x) - f(x_0)}{g(x) - g(x_0)} = \frac{f(x)}{g(x)} \cdot \frac{1 - \frac{f(x_0)}{f(x)}}{1 - \frac{g(x_0}{g(x)}},$$

oder etwas umgeformt,

$$\frac{f(x)}{g(x)} = \frac{f'(\xi)}{g'(\xi)} \cdot \frac{1 - \frac{g(x_0)}{g(x)}}{1 - \frac{f(x_0)}{f(x)}} = \frac{f'(\xi)}{g'(\xi)} \cdot \frac{\frac{f(x_0)}{f(x)} - \frac{g(x_0)}{g(x)}}{1 - \frac{f(x_0)}{f(x)}} + \frac{f'(\xi)}{g'(\xi)}.$$

Hieraus folgt

$$\left| \frac{f(x)}{g(x)} - c \right| \leq \left| \frac{f'(\xi)}{g'(\xi)} \right| \cdot \left| \frac{\frac{f(x_0)}{f(x)} - \frac{g(x_0)}{g(x)}}{1 - \frac{f(x_0)}{f(x)}} \right| + \left| \frac{f'(\xi)}{g'(\xi)} - c \right|$$

für alle $x > \max\{x_0, x_1\}$. Geht nun x gegen ∞, so bleibt $|f'(\xi)/g'(\xi) - c|$ beschränkt durch ε und $|f'(\xi)/g'(\xi)|$ durch $|c| + \varepsilon$, während der Faktor

$$\left| \frac{\frac{f(x_0)}{f(x)} - \frac{g(x_0)}{g(x)}}{1 - \frac{f(x_0)}{f(x)}} \right|$$

gegen Null geht. Folglich gibt es ein $x_2 \geq \max\{x_0, x_1\}$ derart, dass

$$\left| \frac{f(x)}{g(x)} - c \right| \leq 2\varepsilon$$

für alle $x \geq x_2$. Das zeigt: Der Grenzwert $\lim_{x \to \infty} f(x)/g(x)$ existiert und hat den Wert c. □

§8 Funktionenfolgen

8.1 Gleichmäßige Konvergenz

Definition 8.1 — Es sei M eine Menge und (f_n) eine Folge von Funktionen $f_n : M \to \mathbb{C}$.

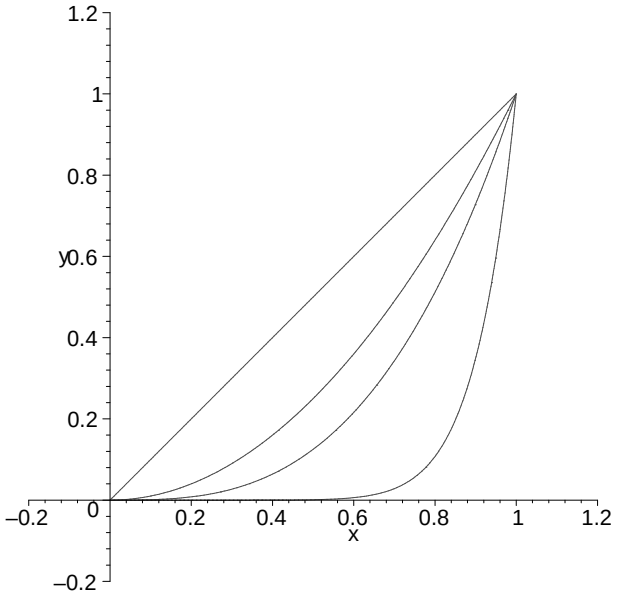
1. Die Folge (f_n) konvergiert *punktweise* gegen die Grenzfunktion $f : M \to \mathbb{C}$, wenn für jedes $x \in M$ gilt $\lim_{n\to\infty} f_n(x) = f(x)$.

2. Die Folge (f_n) konvergiert *gleichmäßig* gegen f, wenn für jedes $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$ existiert so, dass für alle $n \geq n_0$ gilt: $|f_n(x) - f(x)| < \varepsilon$ für alle $x \in M$.

Wir schreiben $f_n \xrightarrow{p.k.} f$ und $f_n \xrightarrow{g.k.} f$ für punktweise bzw. gleichmäßige Konvergenz.

Jede gleichmäßig konvergente Funktionenfolge konvergiert auch punktweise. Die Umkehrung gilt nicht, wie das folgende Beispiel zeigt:

Beispiel 8.2 — Wir betrachten auf dem Intervall $[0, 1]$ die reellen Funktionen $f_n = X^n|_{[0,1]}$. Für jedes $x \in [0, 1)$ ist die Folge (x^n) eine Nullfolge. Und für $x = 1$ ist die Folge (x^n) sowieso konstant. Man sieht, dass die Folge f_n punktweise gegen eine Grenzfunktion f konvergiert, die durch $f(x) = 0$ für $x \in [0, 1)$ und $f(1) = 1$ gegeben ist. Diese Konvergenz ist nicht gleichmäßig, wie man der Skizze für die Funktionen X^n, $n = 1, 2, 3, 10$, entnimmt,



und wie folgt beweist: Es sei $\varepsilon \in (0, 1/2)$ und $n \in \mathbb{N}$ beliebig. Wegen der Stetigkeit von X^n an der Stelle 1 gibt es ein $\delta > 0$ so, dass für alle $x \in (1-\delta, 1]$ gilt $|1 - f_n(x)| < 1/2$. Für $x \in (1-\delta, 1)$ folgt $|f(x) - f_n(x)| = f_n(x) \geq 1/2 > \varepsilon$.

Definition 8.3 — Es sei M eine Menge. Für jede Funktion $f : M \to \mathbb{C}$ definieren wir die Supremumsnorm von f durch

$$\|f\|_M := \sup\{|f(x)| \mid x \in M\} \in \mathbb{R} \cup \{\infty\}.$$

Die Funktion f heißt beschränkt, wenn $\|f\|_M < \infty$.

Lemma 8.4 — *Die Supremumsnorm hat die folgenden Eigenschaften:*

1. $\|f\|_M \geq 0$, *und es gilt* $\|f\|_M = 0$ *genau dann, wenn* $f = 0$.
2. *Für jedes* $f : M \to \mathbb{C}$ *und* $\alpha \in \mathbb{C}$ *gilt* $\|\alpha f\|_M = |\alpha| \, \|f\|_M$.
3. *Für alle* $f, g : M \to \mathbb{C}$ *gilt* $\|f + g\|_M \leq \|f\|_M + \|g\|_M$.
4. *Für alle* $f, g : M \to \mathbb{C}$ *gilt* $\|fg\| \leq \|f\|_M \cdot \|g\|_M$, *falls die rechte Seite definiert ist.*

Beweis. Die beiden ersten Aussagen sind klar. Für jedes $x \in M$ gilt außerdem $|f(x) + g(x)| \leq |f(x)| + |g(x)| \leq \|f\|_M + \|g\|_M$. Da dies für alle $x \in M$ gilt, hat man auch $\|f+g\|_M \leq |f\|_M + \|g\|_M$. Bei der letzten Aussage muss man nur den Fall ausschließen, dass $f = 0$ und g unbeschränkt ist oder umgekehrt. □

Bemerkung 8.5 — Mit dem Begriff der Supremumsnorm kann man gleichmäßige Konvergenz auch so formulieren: Für eine Folge (f_n) von Funktionen auf M gilt

$$f_n \xrightarrow{g.k.} f \quad \Leftrightarrow \quad \|f - f_n\|_M \to 0.$$

Lemma 8.6 *(Eigenschaften der gleichmäßigen Konvergenz)*

1. $f_n \xrightarrow{g.k.} f \quad \Rightarrow \quad f_n \xrightarrow{p.k.} f$.
2. $f_n \xrightarrow{g.k.} f \quad \Rightarrow \alpha f_n \xrightarrow{g.k.} \alpha f$ *für jedes* $\alpha \in \mathbb{C}$.
3. *Aus* $f_n \xrightarrow{g.k.} f$ *und* $g_n \xrightarrow{g.k.} g$ *folgt* $f_n + g_n \xrightarrow{g.k.} f + g$.
4. *Sind* f *und* g *beschränkt, so folgt aus* $f_n \xrightarrow{g.k.} f$ *und* $g_n \xrightarrow{g.k.} g$, *dass* $f_n g_n \xrightarrow{g.k.} fg$.

Beweis. Zur letzten Aussage: Nach Voraussetzung ist f beschränkt. Da f_n gleichmäßig gegen f konvergiert, gibt es ein n_0 so, dass für alle $n \geq n_0$ auch f_n beschränkt ist. Für solche n gilt

$$\begin{aligned} \|f_n g_n - fg\| &\leq \|f_n(g_n - g)\| + \|(f_n - f)g\| \\ &\leq \|f_n\| \, \|g_n - g\| + \|f_n - f\| \, \|g\| \\ &\xrightarrow{n\to\infty} \|f\| \cdot 0 + 0 \cdot \|g\| = 0. \end{aligned}$$

□

Definition 8.7 — Es sei $M \subset \mathbb{C}$. Eine Folge (f_n) von Funktionen $f_n : M \to \mathbb{C}$ konvergiert *lokal gleichmäßig* gegen die Grenzfunktion $f : M \to \mathbb{C}$, wenn es zu jedem $a \in M$ ein $\eta > 0$ gibt derart, dass $f_n|_{U_\eta(a)\cap M} \xrightarrow{g.k.} F|_{U_\eta(a)\cap M}$. Wir schreiben in diesem Falle $f_n \xrightarrow{l.g.k.} f$.

Satz 8.8 — *Es sei $M \subset \mathbb{C}$ und (f_n) eine Folge stetiger Funktionen $f_n : M \to \mathbb{C}$. Gilt $f_n \xrightarrow{l.g.k.} f$ für eine Funktion $f : M \to \mathbb{C}$, so ist die Grenzfunktion f stetig.*

Beweis. Es sei $a \in M$. Nach Voraussetzung gibt es ein $\eta > 0$ derart, dass die Folge $f_n|_{U_\eta(a)\cap M}$ gleichmäßig konvergiert. Da die Stetigkeit von f in a äquivalent zur Stetigkeit von $f|_{U_\eta(a)\cap M}$ ist, können wir M durch $M \cap U_\eta(a)$ ersetzen und somit annehmen, dass die Konvergenz schon auf M gleichmäßig ist.

Es sei nun ein $\varepsilon > 0$ vorgegeben. Wegen der Gleichmäßigkeit der Konvergenz gibt es ein n mit $\|f - f_n\|_M < \varepsilon/3$. Die Funktion f_n ist stetig in a. Daher gibt es in $\delta > 0$ derart, dass für alle $x \in M$, $|x - a| < \delta$, gilt $|f_n(x) - f_n(a)| < \varepsilon/3$. Es folgt

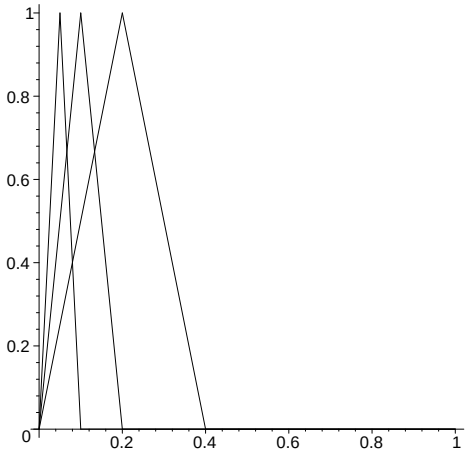
$$\begin{aligned} |f(a) - f(x)| &< |f(a) - f_n(a)| + |f_n(a) - f_n(x)| + |f_n(x) - f(x)| \\ &\leq \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon. \end{aligned}$$

□

Beispiel 8.9 — Wir haben oben ein Beispiel gesehen, wo die Grenzfunktion einer punktweise konvergierenden Folge von stetigen Funktionen nicht mehr stetig war. Damit keine Mißverständnisse aufkommen, gebe ich hier noch ein Beispiel an, wo die Grenzfunktion stetig ist, obwohl die Konvergenz nur punktweise, aber nicht gleichmäßig ist. Dazu betrachten wir auf dem Intervall $[0, 1]$ die Funktionen f_n, $n \geq 2$, mit

$$f_n(x) = \begin{cases} nx & x \in [0, 1/n] \\ 2 - nx \quad \text{falls} & x \in [1/n, 2/n] \\ 0 & x \in [2/n, 1]. \end{cases}$$

Diese Folge konvergiert punktweise gegen die Funktion $f = 0$.



Denn für $x = 0$ ist $f_n(x) = 0$ für alle n, und für jedes $x > 0$ gilt $f_n(x) = 0$ für alle $n \geq 2/x$.

Bemerkung 8.10 — Die analoge Aussage für differenzierbare Funktionen, d.h. „die Grenzfunktion einer gleichmäßig konvergenten Folge von differenzierbaren Funktionen ist differenzierbar“, ist falsch. Hier müssen wir etwas mehr fordern. Das geschieht im nächsten Satz. Bevor Sie den Beweis des Satzes studieren, überlegen Sie sich an einfachen Gegenbeispielen, was passiert, wenn man auf die eine oder die andere der beiden Bedingungen im Satz verzichtet.

Satz 8.11 — *Es sei I ein abgeschlossenes Intervall und $f_n : I \to \mathbb{R}$ eine Folge differenzierbarer Funktionen. Gilt*

1. *die Folge (f_n') konvergiert gleichmäßig, und*
2. *es gibt ein $x_0 \in I$ derart, dass $f_n(x_0)$ konvergiert,*

so konvergiert (f_n) gleichmäßig gegen eine differenzierbare Funktion f, und (f_n') konvergiert gleichmäßig gegen f'.

Beweis. Es sei g die Grenzfunktion der Folge f_n'. Nach Satz 8.8 ist g stetig. Ferner sei $c := \lim f_n(x_0)$.

Wir betrachten für ein $a \in I$ die Funktionen $\Delta_a f_n$ und zeigen zunächst, dass die Folge $(\Delta_a f_n)$ auf I gleichmäßig konvergiert. Es sei $\varepsilon > 0$ vorgegeben. Dann existiert ein n_0 derart, dass für alle $n \geq n_0$ gilt $\|f_n' - g\| < \varepsilon/2$. Für $n, m \geq n_0$ existiert nach dem Mittelwertsatz für jedes $x \in I$ ein z_x in I mit

$$|(\Delta_a f_n - \Delta_a f_m)(x)| = |(f_n' - f_m')(z_x)| \leq \|f_n' - f_m'\| \leq \|f_n' - g\| + \|g - f_m'\| < \varepsilon. \tag{8.1}$$

Das zeigt, dass für jedes x die Folge $(\Delta_a f_n(x))$ eine Cauchyfolge ist und deshalb konvergieren muss. Es sei $\Delta_a(x) := \lim \Delta_a f_n(x)$. Wir haben gesehen, dass $\Delta_a f_n$ punktweise gegen Δ_a konvergiert. Diese Konvergenz ist aber sogar gleichmäßig, denn aus (8.1) folgt durch den Grenzübergang $m \to \infty$, dass

$$\|\Delta_a f_n - \Delta_a\| \leq \varepsilon,$$

und das bedeutet gerade, dass $\Delta_a f_n$ gleichmäßig gegen Δ_a konvergiert.

Wir wenden dies speziell auf die Wahl $a = x_0$ an. Es gilt für alle $x \in I$:

$$f_n(x) = f_n(x_0) + (x - x_0)\Delta_{x_0} f_n(x). \tag{8.2}$$

Da das Intervall I nach Voraussetzung abgeschlossen ist, ist die lineare Funktion $X - x_0$ ebenso wie die stetige Funktion Δ_{x_0} beschränkt. Daher konvergiert die rechte Seite von Gleichung (8.2) gleichmäßig gegen eine Funktion f.

Jetzt, wo wir schon wissen, dass f_n gleichmäßig gegen f konvergiert, können wir denselben Schritt mit beliebigem $a \in I$ wiederholen und sehen, dass

$$f(x) = f(a) + (x - a)\Delta_a(x), \quad \forall x \in I,$$

mit der stetigen Funktion Δ_a. Damit ist f in a differenzierbar und hat die Ableitung $f'(a) = \Delta_a(a) = \lim \Delta_a f_n(a) = \lim f_n'(a)$. □

8.2 Funktionenreihen

Definition 8.12 — Es sei (f_n) eine Folge von Funktionen $f_n : M \to \mathbb{C}$. Unter einer Reihe $\sum_{n=0}^{\infty} f_n$ verstehen wir die Folge (s_n) der Partialsummen $s_n := \sum_{i=0}^{n} f_i$. Entsprechend sprechen wir von punktweiser, gleichmäßiger oder lokal gleichmäßiger Konvergenz der Reihe $\sum_{n=0}^{\infty} f_n$, wenn dies für die Folge (s_n) gilt, und schreiben $\sum_{n=0}^{\infty} f_n = f$, wenn die Folge (s_n) punktweise gegen die Grenzfunktion f konvergiert.

Satz 8.13 — *Es sei (f_n) eine Folge von Funktionen $f_n : M \to \mathbb{C}$. Konvergiert die Reihe $\sum_n \|f_n\|_M$, so konvergiert die Reihe $\sum_n f_n$ gleichmäßig und in jedem Punkt $x \in M$ absolut. Für die Grenzfunktion f gilt $\|f\|_M \leq \sum_n \|f_n\|_M$.*

Beweis. Für jeden Punkt $x \in M$ ist nach Voraussetzung $\|f_n\|_M$ eine Majorante der Folge $f_n(x)$. Daher konvergiert die Reihe $\sum_{n=0}^{\infty} f_n(x)$ absolut. Wir

bezeichnen den Grenzwert mit $f(x)$. Außerdem gilt für jeden Punkt $x \in M$:

$$\begin{aligned} |f(x) - \sum_{i=0}^{n} f_i(x)| &\leq \sum_{i=n+1}^{\infty} |f_i(x)| \\ &\leq \sum_{i=n+1}^{\infty} \|f_i\|_M \\ &\longrightarrow 0 \text{ für } n \to \infty. \end{aligned}$$

Das beweist die gleichmäßige Konvergenz. Aus $\| \sum_{i=0}^{n} f_i \| \leq \sum_{i=0}^{n} \|f_i\|$ folgt im Grenzübergang $n \to \infty$ die letzte Behauptung. □

Lemma 8.14 — *Die Potenzreihen $\sum_{n=0}^{\infty} a_n z^n$ und $\sum_{n=0}^{\infty} (n+1) a_{n+1} z^n$ haben denselben Konvergenzradius.*

Beweis. Wir betrachten die Reihen $A = \sum_{n=0}^{\infty} (n+1) a_{n+1} z^n = \sum_{n=0}^{\infty} n a_n z^{n-1}$, $B = \sum_{n=0}^{\infty} n a_n z^n$ und $C = \sum_{n=0}^{\infty} a_n z^n$. Die Reihe A konvergiert genau dann, wenn die Reihe B konvergiert, denn sie unterscheiden sich nur um den Faktor z. Um zu sehen, dass die Reihen B und C denselben Konvergenzradius haben, genügt es nach Satz 5.36, sich das Folgende zu überlegen: Es gilt

$$\overline{\lim}\, \sqrt[n]{|n a_n|} = \lim \sqrt[n]{n}\, \overline{\lim}\, \sqrt[n]{|a_n|} = \overline{\lim}\, \sqrt[n]{|a_n|},$$

weil $\lim \sqrt[n]{n} = 1$. □

Satz 8.15 — *Es sei $f(z) = \sum_{n=0}^{\infty} a_n z^n$ eine komplexe Potenzreihe mit Konvergenzradius R. Dann konvergiert für jedes r, $0 < r < R$, die Potenzreihe auf der abgeschlossenen Kreisscheibe $K_r := \{z \in \mathbb{C} \mid |z| \leq r\}$ gleichmäßig. Insbesondere ist $f : U_R(0) \to \mathbb{C}$ stetig.*

Beweis. Da $r < R$, konvergiert die Reihe $\sum_n a_n r^n$ absolut. Außerdem gilt $\|a_n X^n\|_{K_r} \leq |a_n| r^n$. Nach Satz 8.13 konvergiert die Reihe $\sum_n a_n X^n|_{K_r}$ gleichmäßig gegen $f|_{K_r}$. Für jeden Punkt $z \in U_R(0)$, $r := (|z| + R)/2$ und $\varepsilon := r - |z|$ liegt $U_\varepsilon(z)$ ganz in K_r. Demnach konvergiert die Reihe $\sum a_n X^n$ auf $U_R(0)$ gleichmäßig in einer Umgebung eines jeden Punktes von $U_R(0)$. Die einzelnen Glieder der Reihe sind stetige Funktionen, nach Satz 8.8 ist auch f stetig. □

Satz 8.16 — *Es sei $f(x) = \sum_n a_n x^n$ eine Potenzreihe mit reellen Koeffizienten a_n und Konvergenzradius $R > 0$. Dann ist die Funktion $f : (-R, R) \to \mathbb{R}$ differenzierbar, und für alle $x \in (-R, R)$ gilt*

$$f'(x) = \sum_{n \geq 1} n a_n x^{n-1}.$$

Insbesondere ist $f \in \mathcal{C}^\infty((-R,R))$.

Beweis. Nach Lemma 8.14 haben die Reihen $\sum_{n=0}^\infty a_n X^n$ und $\sum_{n=0}^\infty n a_n X^{n-1}$ denselben Konvergenzradius und konvergieren nach Satz 8.15 auf jedem abgeschlossenen Intervall $[-r,r]$, $r < R$ gleichmäßig. Nach Satz 8.11 ist f auf jedem dieser Intervalle (und somit auch auf $(-R,R)$) differenzierbar, und die Ableitung von f ist gegeben durch die Reihe $\sum_{n\geq 1} n a_n X^{n-1}$. Die letzte Behauptung des Satzes folgt durch Induktion. □

Beispiel 8.17 — Die auf $(-1,\infty)$ definierte Funktion $\log(1+X)$ hat die Ableitung $1/(1+X)$. Diesen Bruch können wir als geometrische Reihe entwickeln: Für alle $x \in (-1,1)$ gilt

$$\frac{1}{1+x} = 1 - x + x^2 - x^3 + x^4 - \ldots.$$

Wir definieren eine Funktion F, indem wir die Reihenglieder formal integrieren:

$$F(x) := x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \ldots = \sum_{n\geq 1} \frac{(-1)^{n-1}}{n} x^n.$$

Nach den obigen Sätzen hat die Reihe F denselben Konvergenzradius wie die geometrische Reihe, also 1, und ist auf $(-1,1)$ eine $\mathcal{C}^\infty$-Funktion mit der Ableitung

$$F' = \frac{1}{1+X} = \log(1+X)'.$$

Insbesondere ist die Funktion $F - \log(1+X)$ auf $(-1,1)$ konstant. Da außerdem $F(0) = 0 = \log(1+0)$, gilt $F = \log(1+X)$ auf dem Intervall $(-1,1)$. Mit anderen Worten:

Für alle x, $-1 < x < 1$, *gilt*

$$\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \ldots = \sum_{n\geq 1} \frac{(-1)^{n-1}}{n} x^n. \tag{8.3}$$

Die Reihe gestattet die Berechnung von $\log(y)$ zunächst nur für $y \in (0,2)$. Für $y \geq 2$ wählt man ein $m \in \mathbb{N}$ mit $y/e^m < 2$ und bekommt

$$\log(y) = m + \log(y/e^m).$$

Alternativ kann man so vorgehen: Wir betrachten neben der Reihe (8.3) noch die folgende, die durch Auswertung in $-x$ entsteht:

$$\log(1-x) = -x - \frac{x^2}{2} - \frac{x^3}{3} - \frac{x^4}{4} - \ldots \tag{8.4}$$

Indem man die Reihen voneinander subtrahiert, gewinnt man die Darstellung

$$\log \frac{1+x}{1-x} = 2 \sum_{n=0}^{\infty} \frac{x^{2n+1}}{2n+1} \tag{8.5}$$

Die Abbildung $x \mapsto y = (1+x)/(1-x)$ liefert eine Bijektion von $(-1,1)$ auf $(0,\infty)$. Deshalb kann man die Formel (8.5) zur Berechnung von $\log(y)$ für beliebig große y verwenden.

Beispiel 8.18 — Die Arkustangensfunktion $\arctan : \mathbb{R} \longrightarrow (-1,1)$ hat die Ableitung $\arctan'(x) = \frac{1}{1+x^2}$. Im Intervall $(-1,1)$ können wir die Ableitung als geometrische Reihe mit Konvergenzradius 1 entwickeln:

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

Die Reihe

$$g(x) := \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{2n+1},$$

die durch gliedweises Integrieren entsteht, hat darum ebenfalls den Konvergenzradius 1, ist auf $(-1,1)$ differenzierbar und hat dort dieselbe Ableitung wie der Arkustangens. Folglich ist $\arctan - g$ eine konstante Funktion. Andererseits gilt $\arctan(0) = 0 = g(0)$, woraus $\arctan = g$ folgt. Wir halten fest:

Für alle $x \in (-1,1)$ gilt

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

Könnten wir hier $x = 1$ einsetzen, so würden wir die folgende Entwicklung erhalten (Leibnizsche Reihe)

$$\frac{\pi}{4} = \arctan(1) = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \ldots$$

Diese Entwicklung ist tatsächlich gültig. Zu ihrer Begründung reichen die bisherigen Sätze aber nicht aus. Wir kommen später darauf zurück. Unabhängig von der Rechtfertigung dieser Reihe sieht man nach einigem Ausprobieren, dass die Reihe sehr langsam konvergiert. Eine sehr viel schnellere Konvergenz erhält man folgendermaßen:

Für den Tangens gilt die Additionsformel

$$\tan(\alpha \pm \beta) = \frac{\tan(\alpha) \pm \tan(\beta)}{1 \mp \tan(\alpha)\tan(\beta)}$$

für alle α, β mit $\alpha, \beta, \alpha \pm \beta \in \mathbb{R} \setminus (\pi/2 + \pi\mathbb{Z})$. Das folgt aus den Additionsformeln für Sinus und Kosinus. Speziell für $\alpha = \arctan 1/5$ und $\beta = \arctan 1/239$ erhält man

$$\tan(2\alpha) = \frac{5}{12}, \quad \tan(4\alpha) = \frac{120}{119}, \quad \tan(4\alpha - \beta) = 1.$$

Es folgt

$$\frac{\pi}{4} = 4\alpha - \beta = 4 \arctan \frac{1}{5} - \arctan \frac{1}{239}.$$

Das liefert die schnell konvergente Darstellung

$$\pi = 16 \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)5^{2n+1}} - 4 \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)239^{2n+1}}.$$

8.3 Taylorreihen

Es sei $f := \sum_{n \geq 0} a_n X^n$ eine Potenzreihe mit Konvergenzradius $R > 0$. Wir haben gesehen, dass f auf $(-R, R)$ unendlich oft differenzierbar ist, und dass die k-te Ableitung durch

$$f^{(k)}(x) = \sum_{n \geq k} n(n-1) \cdots (n-k+1) a_n X^{n-k}$$

gegeben ist. Insbesondere findet man für den Wert in 0:

$$f^{(k)}(0) = k! a_k.$$

Will man also eine gegebene Funktion f in einer Umgebung der 0 durch eine Potenzreihe darstellen, so muss, wenn es überhaupt möglich ist, diese Funktion unendlich oft differenzierbar sein, und die Koeffizienten der Reihe müssen durch

$$a_n = \frac{f^{(n)}(0)}{n!}$$

gegeben sein. Wir werden sehen, dass diese Bedingungen aber nicht hinreichend sind.

Definition 8.19 — Es sei I ein Intervall und $f : I \to \mathbb{R}$ n-mal differenzierbar. Das Polynom

$$T_a^n f := \sum_{\nu=0}^{n} \frac{f^{(\nu)}(a)}{\nu!} (X-a)^\nu$$

heißt das n-te Taylorpolynom von f in a. Falls $f \in \mathcal{C}^\infty(I)$, heißt

$$T_a f := \sum_{\nu=0}^{\infty} \frac{f^{(\nu)}(a)}{\nu!} (X-a)^\nu$$

die Taylorreihe von f im Punkte a.

Wir stehen den folgenden Problemen gegenüber:

1. Wie genau approximiert das n-te Taylorpolynom die Funktion f?
2. Konvergiert die Taylorreihe $T_a f$, d.h. hat die Reihe einen positiven Konvergenzradius?

3. Falls die Taylorreihe konvergiert, gilt $T_a f(x) = f(x)$ und für welchen Argumentbereich?

Lemma 8.20 — *Es sei I ein Intervall, $a \in I$, und f auf I n-mal differenzierbar. Für alle $k < n$ gelte $f^{(k)}(a) = 0$. Dann existiert für jedes $x \in I$ ein $t \in (0,1)$ mit*

$$f(x) = \frac{f^{(n)}(a + t(x-a))}{n!}(x-a)^n.$$

Beweis. Für $x = a$ ist die Aussage klar. Wir nehmen also an, dass $x \neq a$.

Wir wenden den verallgemeinerten Zwischenwertsatz auf die Funktionen f und $(X-a)^n$ an und schließen: Es gibt eine Zahl x_1 strikt zwischen a und x mit

$$\frac{f(x)}{(x-a)^n} = \frac{f(x)-f(a)}{(x-a)^n - (a-a)^n} = \frac{f'(x_1)}{n(x_1-a)^{n-1}}.$$

Für den Quotienten auf der rechten Seite können wir das Argument wiederholen. Induktiv finden wir also Elemente $x_1, \ldots, x_n$, jeweils x_{i+1} strikt zwischen x_i und a, mit der Eigenschaft:

$$\frac{f(x)}{(x-a)^n} = \frac{f'(x_1)}{n(x_1-a)^{n-1}} = \frac{f''(x_2)}{n(n-1)(x_2-a)^{n-2}} = \cdots = \frac{f^{(n)}(x_n)}{n!}.$$

Setze $t = (x_n - a)/(x-a)$. □

Satz 8.21 *(Abschätzung mit Lagrangeschem Restglied) — Es sei I ein Intervall, $a \in I$, und $f : I \to \mathbb{R}$ sei $(n+1)$-mal differenzierbar. Dann existiert für jedes $x \in I$ ein $t \in (0,1)$ so, dass*

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \ldots + \frac{f^{(n)}(a)}{n!}(x-a)^n + R_{n+1} \quad (8.6)$$

mit

$$R_{n+1} = \frac{f^{(n+1)}(a + t(x-a))}{(n+1)!}(x-a)^{n+1}. \quad (8.7)$$

Natürlich kann man eine Gleichung wie (8.6) immer hinschreiben, wenn man das Restglied R_{n+1} passend wählt. Die eigentliche Aussage des Satzes ist also, dass das Restglied R_{n+1} durch die Formel (8.7) beschrieben werden kann. Diese Beschreibung nennt man das Lagrangesche Restglied. Es gibt auch andere Darstellungen.

Beweis. Die Funktion $g := f - T_a^n f$ ist $n+1$-mal differenzierbar, und es gilt

$$g(a) = \ldots = g^{(n)}(a) = 0, \quad g^{(n+1)} = f^{(n+1)}.$$

Nach Lemma 8.20 existiert ein $t \in (0,1)$ mit

$$R_{n+1} = g(x) = \frac{f^{(n+1)}(a+t)x - a))}{(n+1)!}(x-a)^{n+1}.$$

□

Satz 8.22 — *Es sei I ein Intervall, $a \in I$, und $f \in \mathcal{C}^\infty(I)$. Gibt es Konstanten $M, r > 0$ mit $|f^{(k)}(x)|/k! \leq Mr^{-k}$ für alle $k \in \mathbb{N}_0$ und $x \in I \cap U_r(a)$, so konvergiert die Taylorreihe $T_a f$ auf $I \cap U_r(a)$ lokal gleichmäßig gegen f. Insbesondere gilt*

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n$$

für jedes $x \in I \cap U_r(a)$.

Beweis. Es sei $n \in \mathbb{N}$. Nach Satz 8.21 gibt es für jedes $x \in I$ ein $t \in (0,1)$ mit

$$|f(x) - T_a^n f(x)| = \left| \frac{f^{(n+1)}(a + t(x-a))}{(n+1)!}(x-a)^{n+1} \right| \leq M \left(\frac{|x-a|}{r} \right)^{n+1}.$$

Für jedes s, $0 < s < r$, folgt:

$$\|f - T_a^n f\|_{I \cap U_s(a)} \leq M(s/r)^{n+1}.$$

Dies bedeutet gerade, dass die Folge $T_a^n f$ auf $I \cap U_s(a)$ gleichmäßig und somit auf $I \cap U_r(a)$ lokal gleichmäßig gegen f konvergiert. □

Beispiel 8.23 — Wir betrachten die Funktion $X^\alpha : \mathbb{R} > 0 \to \mathbb{R}$ mit $\alpha \in \mathbb{R}$. Diese Funktion ist beliebig oft differenzierbar, die k-te Ableitung ist gegeben durch

$$\alpha \cdot (\alpha - 1) \cdots (\alpha - k + 1) X^{\alpha - k},$$

und indem wir an der Stelle 1 auswerten, finden wir, dass Taylorreihe an der Stelle 1 im wesentlichen, d.h. bis auf eine Verschiebung im Argument, die Binomialreihe zu α ist:

$$T_1 X^\alpha = \sum_{n=0}^{\infty} \binom{n}{\alpha} (X-1)^n = B_\alpha(X-1).$$

Auf einer r-Umgebung von 1 ist die n-te Ableitung beschränkt durch

$$\binom{\alpha}{n}(1-r)^{\alpha-n}$$

für $n \geq \alpha$. Wir wissen, dass die Folge der Binomialkoeffizienten $\binom{\alpha}{n}$ beschränkt ist, etwa durch C. Für $0 < r < 1/2$ folgt $r < 1 - r$ und somit

$$\binom{\alpha}{n}(1-r)^{\alpha-n} \leq C(1-r)^\alpha \cdot r^{-n}$$

für alle $k \geq \alpha$. Der obige Satz sagt nun, dass die Taylorreihe auf $(-1/2, 3/2)$ lokal gleichmäßig gegen X^α konvergiert.

8.3.1 Eine nicht analytische Funktion

$\mathcal{C}^\infty$-Funktionen $f : I \to \mathbb{C}$, die in einer Umgebung eines jeden Punktes ihres Definitionsgebietes in eine Potenzreihe entwickelt werden können, heißen analytisch. Wir werden uns damit genauer in der Vorlesung Funktionentheorie beschäftigen. Sehr viele, wenn nicht die meisten in der Praxis auftretenden Funktionen sind analytisch, aber in einem gewissen Sinne gibt es viel mehr Funktionen, die nicht analytisch sind. Bei vielen Konstruktionen, mit denen man sich für bestimmte Zwecke Funktionen mit vorgegebenen Eigenschaften nach Maß zusammenschneidert, spielen aber gerade nicht analytische Funktionen eine große Rolle.

Das folgende Beispiel beschreibt eine sehr nützliche $\mathcal{C}^\infty$-Funktion, die nicht analytisch ist.

Lemma 8.24 — *Für alle $n \in \mathbb{Z}$ gilt* $\lim_{x\to 0} x^n e^{-\frac{1}{x^2}} = 0$.

Beweis. Für $n > 0$ ist weniger zu zeigen als für $n \leq 0$. Es genügt also, das Problem

$$\lim_{x\to 0} \frac{e^{-\frac{1}{x^2}}}{x^k} \stackrel{?}{\to} 0$$

für $k \in \mathbb{N}_0$ zu betrachten. Aus Symmetriegründen genügt es ferner, sich den einseitigen Grenzwert $\lim_{x\searrow 0}$ anzuschauen. Schließlich reduziert die Substitution $x = 1/y$ das Problem darauf,

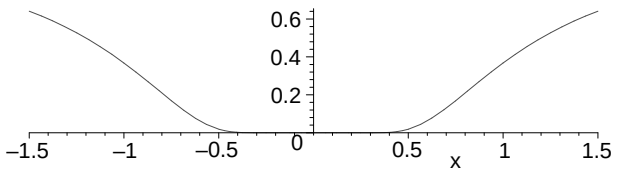
$$\lim_{y\to\infty} \frac{y^{k/2}}{e^y} = 0$$

zu beweisen. Aber wir wissen schon, dass die Funktion exp schneller wächst als jede Potenz (s. Satz 6.27). □

Es sei $\Phi : \mathbb{R} \to \mathbb{R}$ die Funktion, die durch

$$\Phi(x) = \begin{cases} \exp(-\frac{1}{x^2}) & \text{falls } x \neq 0, \\ 0 & \text{sonst.} \end{cases}$$

gegeben ist. Hier ist ein Bild von Φ. Man beachte, wie sehr sich der Graph der Funktion an die x-Achse anschmiegt.



Es ist klar, dass die Funktion Φ außerhalb des Koordinatenursprungs eine $\mathcal{C}^\infty$-Funktion ist und dass ihre Ableitungen durch

$$\Phi'(x) = \frac{2}{x^3} e^{-\frac{1}{x^2}}, \quad \Phi''(x) = \frac{4 - 6x^2}{x^6} e^{-\frac{1}{x^2}}$$

und allgemeiner

$$\Phi^{(k)}(x) = \frac{p_k(x)}{x^{3k}} e^{-\frac{1}{x^2}}$$

gegeben sind, wobei $p_k(x)$ ein hier nicht weiter interessierendes Polynom vom Grad $2k$ ist. Jetzt folgt induktiv, dass die Funktion Φ auch im Nullpunkt unendlich oft differenzierbar ist und dass alle Ableitungen verschwinden: Für $k = 0$ folgt dies aus dem vorigen Lemma; und ist bereits bekannt, dass $\Phi^{(k)}(0) = 0$, so ergibt sich mit erneuter Anwendung des Lemmas, dass der Grenzwert

$$\lim_{x \to 0} \frac{\Phi^{(k)}(x)}{x}$$

existiert, d.h. dass die $(k+1)$-te Ableitung in 0 existiert, und den Wert 0 hat.

Die Konsequenz hieraus ist: Die Taylorreihe $T_0\Phi$ verschwindet identisch und stellt – außer im Nullpunkt – nirgends die Funktion Φ dar.

8.A Der Abelsche Grenzwertsatz

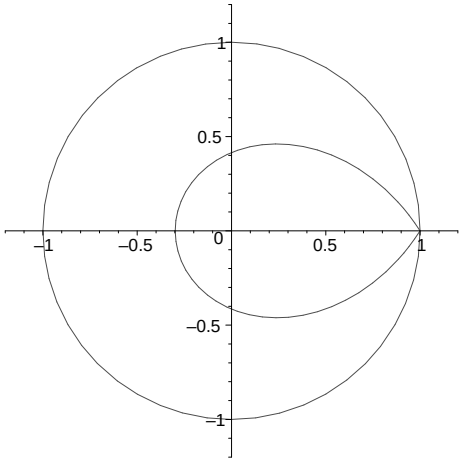
Theorem 8.25 *(Abelscher Grenzwertsatz) — Konvergiert die komplexe Potenzreihe $\sum_n a_n z^n$ in $w \in \mathbb{C}$, $w \neq 0$, so konvergiert die Reihe für jedes $c \geq 1$ auf der Menge $M_c = \{z \in \mathbb{C} \,|\, |1 - z/w| \leq c(1 - |z/w|)\}$ gleichmäßig gegen eine stetige Grenzfunktion.*

Folgerung 8.26 — *Unter den Voraussetzungen des Satzes gilt*

$$\lim_{t \nearrow 1} \sum_{n=0}^{\infty} a_n (tw)^n = \sum_{n=0}^{\infty} a_n w^n.$$

Beweis. Wie man anhand der Definition von M_c sofort sieht, liegt die Strecke $\{tw \,|\, t \in [0,1]\}$ in jeder der Mengen M_c, $c \geq 1$. Daraus folgt die Behauptung. □

Im übrigen haben die Mengen M_c etwa die folgende Form. Das Bild zeigt den Rand der Menge M_c für $w = 1$ und $c = 1.85$. Man beachte, dass die Randkurve bei $z = 1$ nicht glatt ist, sondern eine Spitze hat.



Beweis des Satzes. Nach Voraussetzung konvergiert für jedes $n \geq 0$ die Reihe

$$s_n = \sum_{i \geq n} a_i w^i,$$

und (s_n) ist eine Nullfolge. Weiter gilt $a_k w^k = s_k - s_{k+1}$.

Für $|z| < |w|$ konvergieren die Reihen $\sum_n s_n (z/w)^n$ absolut. Deshalb dürfen wir für solche z die Umordnungen in der folgenden Rechnung vornehmen:

$$\begin{aligned}
\sum_{k=n}^{\infty} a_k z^k &= \sum_{k=n}^{\infty} a_k w^k (z/w)^k \\
&= \sum_{k=n}^{\infty} (s_k - s_{k+1})(z/w)^k \\
&= \sum_{k=n}^{\infty} s_k (z/w)^k - \sum_{k=n}^{\infty} s_{k+1} (z/w)^k \\
&= \sum_{k=n}^{\infty} s_k (z/w)^k - \sum_{k=n+1}^{\infty} s_k (z/w)^{k-1} \\
&= s_n (z/w)^n + (z/w - 1)(z/w)^n \sum_{k=0}^{\infty} s_{n+1+k} (z/w)^k.
\end{aligned}$$

Mit der Bezeichnung $C_n := \max\{|s_k| \mid k \geq n\}$ können wir die Summe jetzt wie folgt abschätzen:

$$\left| \sum_{k=n}^{\infty} a_k z^k \right| \leq C_n \left(1 + \frac{|1 - z/w|}{1 - |z/w|} \right).$$

Für $z = w$ gilt sowieso

$$\left| \sum_{k=n}^{\infty} a_k z^k \right| \leq C_n.$$

Da für jedes $c \geq 1$ und $z \in M_c$ entweder $z = w$ oder $|z| < |w|$ gilt, folgt:

$$\left| \sum_{k=n}^{\infty} a_k z^k \right| \leq C_n (1 + c).$$

Wegen der Konvergenz der Reihe $\sum_n a_n w^n$ ist (C_n) eine Nullfolge. Zu jedem $\varepsilon > 0$ gibt es daher ein n_0 derart, dass $C_n < \frac{\varepsilon}{1+c}$. Für alle $n \geq n_0$ und $z \in M_c$ folgt

$$\left| \sum_{k=n}^{\infty} a_k z^k \right| < \varepsilon.$$

Das beweist die gleichmäßige Stetigkeit auf M_c. □

Folgerung 8.27 — $\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \ldots$.

Beweis. Nach dem Leibnizkriterium konvergiert die Reihe $\sum_{n\geq 0}(-1)^n \frac{x^{2n+1}}{2n+1}$ auch noch im Punkt $x = 1$. Auf dem offenen Intervall $(-1, 1)$ wird durch die Reihe die auf ganz $\mathbb{R}$ stetige Funktion arctan dargestellt. Nach dem Abelschen Grenzwertsatz stellt die Reihe auch noch auf $(-1, 1]$ eine stetige Funktion dar. Es folgt

$$1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \ldots = \arctan(1) = \frac{\pi}{4}.$$

□

Folgerung 8.28 — $\log(2) = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \ldots$.

Beweis. Die Reihe $\sum_{n\geq 1}(-1)^{n-1} \frac{x^n}{n}$ stellt auf $(-1, 1)$ die auf $(-1, \infty)$ stetige Funktion $\log(1 + x)$ dar und konvergiert noch im Punkt $x = 1$. Nach dem Abelschen Grenzwertsatz stellt die Reihe auf $(-1, 1]$ eine stetige Funktion dar. Es folgt:

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \ldots = \log(1 + 1).$$

□

Folgerung 8.29 — *Es seien $\sum_{n\geq 0} a_n$ und $\sum_{n\geq 0} b_n$ konvergente Reihen und $\sum_{n\geq 0} c_n$ ihr Cauchyprodukt, d.h. $c_n = \sum_{k=0}^{n} a_k b_{n-k}$ für alle $n \geq 0$. Ist die Reihe $\sum_{n\geq 0} c_n$ konvergent, so gilt*

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

Beweis. Für alle $z \in \mathbb{C}$ mit $|z| < 1$ konvergieren die Potenzreihen

$$\sum_{n\geq 0} a_n z^n, \quad \sum_{n\geq 0} b_n z^n, \quad \sum_{n\geq 0} c_n z^n$$

nach Satz 5.30 absolut, und es gilt

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

Für $z = 1$ konvergieren alle Potenzreihen immer noch und stellen auf den im Abelschen Grenzwertsatz eingeführten Mengen M_c stetige Funktionen dar. Aus der Stetigkeit für $z \to 1$, $z \in (-1, 1]$, folgt die Behauptung. $\square$

§9 Integration

9.1 Das Riemannintegral

Definition 9.1 — Eine Zerlegung des abgeschlossenen Intervalls $I = [a, b]$ ist eine endliche Folge $\mathcal{Z} = (t_0, \ldots, t_n)$, $n \in \mathbb{N}_0$, mit $a = t_0 < \cdots < t_n = b$. Eine weitere Zerlegung $\mathcal{Z}' = (t'_0, \ldots, t'_m)$ heißt feiner als $\mathcal{Z}$ oder eine Verfeinerung von $\mathcal{Z}$, in Zeichen: $\mathcal{Z}' \geq \mathcal{Z}$, wenn $\{t_0, \ldots, t_n\} \subset \{t'_0, \ldots, t'_m\}$.

Bemerkung 9.2 — Sind $\mathcal{Z}' = (t'_0, \ldots, t'_n)$ und $\mathcal{Z}'' = \{t''_0, \ldots, t''_m\}$ zwei Zerlegungen von I, so gibt es stets eine gemeinsame Verfeinerung $\mathcal{Z}$, d.h. eine Zerlegung $\mathcal{Z}$ mit $\mathcal{Z} \geq \mathcal{Z}'$ und $\mathcal{Z} \geq \mathcal{Z}''$. Dazu muss man nur die Menge $\{t'_0, \ldots, t'_n\} \cup \{t''_0, \ldots, t''_m\}$ ordnen und neu numerieren.

Definition 9.3 — Es sei I ein abgeschlossenes Intervall, $f : I \to \mathbb{R}$ eine beschränkte Funktion und $\mathcal{Z} = (t_0, \ldots, t_n)$ eine Zerlegung von I. Dann heißen

$$O(f, \mathcal{Z}) = \sum_{i=1}^{n} (t_i - t_{i-1}) \sup f([t_{i-1}, t_i])$$

und

$$U(f, \mathcal{Z}) = \sum_{i=1}^{n} (t_i - t_{i-1}) \inf f([t_{i-1}, t_i])$$

die Ober- bzw. Untersumme von f bezüglich der Zerlegung $\mathcal{Z}$.

Lemma 9.4 — *Ober- und Untersumme haben die folgenden Eigenschaften:*

1. *Aus $\mathcal{Z}' \geq \mathcal{Z}$ folgt $O(f, \mathcal{Z}') \leq O(f, \mathcal{Z})$ und $U(f, \mathcal{Z}') \geq U(f, \mathcal{Z})$.*
2. *Für beliebige Zerlegungen $\mathcal{Z}$ und $\mathcal{Z}'$ von I gilt $U(f, \mathcal{Z}) \leq O(f, \mathcal{Z}')$.*
3. *$U(-f, \mathcal{Z}) = -O(f, \mathcal{Z})$.*
4. *Für alle $\alpha > 0$ gilt $U(\alpha f, \mathcal{Z}) = \alpha U(f, \mathcal{Z})$, $O(\alpha f, \mathcal{Z}) = \alpha O(f, \mathcal{Z})$.*
5. *$U(f, \mathcal{Z}) + U(g, \mathcal{Z}) \leq U(f + g, \mathcal{Z}) \leq O(f + g, \mathcal{Z}) \leq O(f, \mathcal{Z}) + O(g, \mathcal{Z})$.*

Beweis. Zu 1.: Es sei $\mathcal{Z}' = (t'_0, \ldots t'_m)$ eine Verfeinerung von $\mathcal{Z} = (t_0, \ldots, t_n)$. Dann gibt es Indizes $0 = \mu_0 < \cdots < \mu_n = m$ mit $t'_{\mu_i} = t_i$ für alle $i = 0, \ldots, n$. Mit diesen Bezeichnungen folgt

$$t_i - t_{i-1} = \sum_{j=\mu_{i-1}+1}^{\mu_i} (t'_j - t'_{j-1})$$

und

$$\sup f([t'_{j-1}, t'_j]) \leq \sup f([t_{i-1}, t_i]), \quad \inf f([t'_{j-1}, t'_j]) \geq \inf f([t_{i-1}, t_i])$$

für alle j mit $\mu_{i-1} < j \leq \mu_i$. Nun folgt

$$\begin{aligned} O(f, \mathcal{Z}') &= \sum_{j=1}^{m} (t'_j - t'_{j-1}) \sup f([t'_{j-1}, t'_j]) \\ &= \sum_{i=1}^{n} \sum_{j=\mu_{i-1}+1}^{\mu_i} (t'_j - t'_{j-1}) \sup f([t'_{j-1}, t'_j]) \\ &\leq \sum_{i=1}^{n} \sum_{j=\mu_{i-1}+1}^{\mu_i} (t'_j - t'_{j-1}) \sup f([t_{i-1}, t_i]) \\ &= \sum_{i=1}^{n} (t_i - t_{i-1}) \sup f([t_{i-1}, t_i]) \\ &= O(f, \mathcal{Z}). \end{aligned}$$

Für die Untersumme schließt man analog.
Zu 2.: Es sei $\tilde{\mathcal{Z}}$ eine gemeinsame Verfeinerung von $\mathcal{Z}$ und $\mathcal{Z}'$. Mit dem Ergebnis aus 1. folgt

$$U(f, \mathcal{Z}) \leq U(f, \tilde{\mathcal{Z}}) \leq O(f, \tilde{\mathcal{Z}}) \leq O(f, \mathcal{Z}').$$

Zu 3. und 4.: Für jede beschränkte Teilmenge $A \subset \mathbb{R}$ und jedes $\alpha > 0$ gilt $\inf(-A) = -\sup(A)$, $\inf(\alpha A) = \alpha \inf(A)$ und $\sup(\alpha A) = \alpha \sup(A)$. Wendet man dies auf alle Summanden in der Ober- bzw. Untersumme an, so folgt die Behauptung.
Zu 5.: Für jedes Teilintervall $J \subset I$ gilt: Für $x \in J$ hat man $f(x) + g(x) \leq \sup f(J) + \sup g(J)$, also auch $\sup(f+g)(J) \leq \sup f(J) + \sup g(J)$. Daraus folgt die Behauptung.

□

Definition 9.5 — Es sei I ein abgeschlossenes Intervall und $f : I \to \mathbb{R}$ beschränkt. Dann heißen

$$O(f, I) := \inf\{O(f, \mathcal{Z}) \mid \mathcal{Z} \text{ Zerlegung von } I\}$$

und

$$U(f, I) := \sup\{U(f, \mathcal{Z}) \mid \mathcal{Z} \text{ Zerlegung von } I\}$$

Ober- bzw. Unterintegral von f auf I.

Lemma 9.6 — *Es sei I ein abgeschlossenes Intervall und $f : I \to \mathbb{R}$ beschränkt. Dann gilt*

1. $U(-f, I) = -O(f, I)$.
2. $U(\alpha f, I) = \alpha U(f, I)$ *und* $O(\alpha f, I) = \alpha O(f, I)$ *für alle* $\alpha > 0$.

3. $U(f,I)+U(g,I) \leq U(f+g,I) \leq O(f+g,I) \leq O(f,I)+O(g,I)$.

Beweis. Alle Aussagen folgen direkt aus Lemma 9.4. □

Definition 9.7 — Es sei $I=[a,b]$. Eine beschränkte Funktion $f: I \to \mathbb{R}$ heißt im Riemannschen Sinne integrierbar, wenn $O(f,I)=U(f,I)$. In diesem Falle nennt man

$$\int_a^b f(x)dx := \int_I f(x)dx := O(f,I) = U(f,I)$$

das Riemannintegral von f zwischen den Grenzen a und b.

Die Funktion f nennt man den Integranden.

Satz 9.8 — *Es sei $I=[a,b]$. Das Riemannintegral hat die folgenden Eigenschaften:*

1. *Ist $f: I \to \mathbb{R}$ integrierbar und $\alpha \in \mathbb{R}$, so ist auch αf integrierbar, und es gilt*
$$\int_a^b \alpha f(x)dx = \alpha \int_a^b f(x)dx.$$
2. *Sind $f,g: I \to \mathbb{R}$ integrierbar, so ist auch $f+g$ integrierbar, und es gilt*
$$\int_a^b (f+g)(x)dx = \int_a^b f(x)dx + \int_a^b g(x)dx.$$
3. *Sind $f,g: I \to \mathbb{R}$ integrierbare Funktionen und gilt $f(x) \geq g(x)$ für alle $x \in I$, so gilt auch*
$$\int_a^b f(x)dx \geq \int_a^b g(x)dx.$$

Beweis. Zu 1.: Die Aussage folgt aus Lemma 9.6, und zwar zunächst für $\alpha > 0$ und $\alpha = -1$, und durch Verbindung der beiden Aussagen für beliebige α.
Zu 2.: Sind f und g integrierbar, so gilt $U(f,I)=O(f,I)$ und $U(g,I)=O(g,I)$. Aus Lemma 9.6 Aussage 3. folgt $O(f+g,I)=U(f+g,I)=U(f,I)+U(g,I)$. Das war zu zeigen.
Zu 3.: Ist $f(x) \geq g(x)$ für jedes $x \in I$, so gilt $\inf f(J) \geq \inf g(J)$ für jede Teilmenge $J \subset I$. Für jede Zerlegung $\mathcal{Z}$ folgt $U(f,\mathcal{Z}) \geq U(g,\mathcal{Z})$. Durch Übergang zum Supremum über alle Zerlegungen folgt $U(f,I) \geq U(g,I)$, und da f und g als integrierbar vorausgesetzt sind, folgt die Behauptung. □

Bemerkung 9.9 — Der Satz sagt, dass die integrierbaren Funktionen einen Untervektorraum $\mathcal{I}([a,b])$ im Raum $B([a,b])$ der auf $[a,b]$ beschränkten Funktionen bilden und dass das Integral

$$\int_a^b dx : \mathcal{I}([a,b]) \to \mathbb{R}, \;\; f \mapsto \int_a^b f(x)dx,$$

eine lineare Abbildung ist, die überdies monoton ist:

$$f(x) \geq g(x) \;\; \forall x \in [a,b] \quad \Rightarrow \quad \int_a^b f(x)dx \geq \int_a^b g(x)dx.$$

Satz 9.10 — *Es sei I ein abgeschlossenes Intervall und $f : I \to \mathbb{R}$ eine beschränkte Funktion. Die folgenden Aussagen sind äquivalent:*

1. *f ist integrierbar.*

2. *Zu jedem $\varepsilon > 0$ gibt es eine Zerlegung $\mathcal{Z}$ von I mit der Eigenschaft*

$$O(f,\mathcal{Z}) - U(f,\mathcal{Z}) < \varepsilon \tag{9.1}$$

Beweis. 1. $\Rightarrow$ 2.: Zu jedem vorgegebenen $\varepsilon > 0$ gibt es eine Zerlegung $\mathcal{Z}'$ mit $O(f,\mathcal{Z}') < O(f,I) + \varepsilon/2$. Ebenso gibt es eine Zerlegung $\mathcal{Z}''$ mit $U(f,\mathcal{Z}'') > U(f,I) - \varepsilon/2$. Es sei $\mathcal{Z}$ eine gemeinsame Verfeinerung von $\mathcal{Z}'$ und $\mathcal{Z}''$. Dann gilt

$$O(f,\mathcal{Z}) - U(f,\mathcal{Z}) \leq O(f,\mathcal{Z}') - U(f,\mathcal{Z}'') < (O(f,I) + \varepsilon/2) - (U(f,I) - \varepsilon/2) = \varepsilon.$$

2. $\Rightarrow$ 1.: Es sei $\varepsilon > 0$ und $\mathcal{Z}$ eine Zerlegung mit $O(f,\mathcal{Z}) - U(f,\mathcal{Z}) < \varepsilon$. Es folgt

$$0 \leq O(f,I) - U(f,I) \leq O(f,\mathcal{Z}) - U(f,\mathcal{Z}) < \varepsilon$$

Da ε beliebig war, folgt $O(f,I) - U(f,I) = 0$, d.h. f ist integrierbar. □

Beispiel 9.11 — Es sei $\alpha \neq -1$ und $0 < a < b$. Wir wollen zeigen, dass die Funktion $X^\alpha : [a,b] \to \mathbb{R}$ integrierbar ist, und dass

$$\int_a^b x^\alpha dx = \frac{b^{\alpha+1} - a^{\alpha+1}}{\alpha+1}.$$

Dazu betrachten wir die folgende Zerlegung von $[a,b]$. Für $N \in \mathbb{N}$ sei $q := \sqrt[N]{b/a}$ und $\mathcal{Z}_n := (t_0, \ldots, t_N)$ mit $t_i = aq^i$. Je nachdem, ob $\alpha \geq 0$ oder $\alpha \leq 0$, ist die Funktion X^α auf $[a,b]$ monoton wachsend oder fallend. Wir rechnen hier nur den Fall $\alpha \geq 0$ weiter, der andere Fall geht ganz analog. Dann wird das Supremum bzw. das Infimum von X^α auf jedem Teilintervall $[t_{i-1}, t_i]$ an der rechten bzw.

an der linken Intervallgrenze angenommen und hat dementsprechend den Wert $t_i^\alpha = a^\alpha q^{i\alpha}$ bzw. $t_{i-1}^\alpha = a^\alpha q^{(i-1)\alpha}$. Damit ergibt sich für die Obersumme:

$$\begin{aligned} O(X^\alpha, \mathcal{Z}_N) &= \sum_{i=1}^{N} (t_i - t_{i-1}) \sup X^\alpha([t_{i-1}, t_i]) \\ &= \sum_{i=1}^{N} (aq^i - aq^{i-1}) a^\alpha q^{i\alpha} \\ &= a^{\alpha+1}(q-1)q^\alpha \sum_{i=1}^{N} (q^{\alpha+1})^{i-1} \\ &= a^{\alpha+1} \frac{q-1}{q^{\alpha+1}-1} q^\alpha (q^{N(\alpha+1)} - 1) \\ &= a^{\alpha+1} \frac{q-1}{q^{\alpha+1}-1} ((b/a)^{\alpha+1} - 1) q^\alpha. \end{aligned}$$

(An dieser Stelle geht die Annahme $\alpha \neq -1$ ein, denn für $\alpha = -1$ wird $q^{1+\alpha} - 1$ Null.) Geht nun N gegen Unendlich, so geht q gegen 1, und mit der Regel von de l'Hospital folgt:

$$\lim_{q \to 1} \frac{q-1}{q^{\alpha+1}-1} = \lim_{q \to 1} \frac{1}{(\alpha+1)q^\alpha} = \frac{1}{\alpha+1}.$$

Es folgt

$$\lim_{N \to \infty} O(X^\alpha, \mathcal{Z}_N) = \frac{b^{\alpha+1} - a^{\alpha+1}}{\alpha+1}.$$

Für die Untersummen erhält man denselben Grenzwert. Nach dem Integrationskriterium Satz 9.10 ist die Funktion X^α integrierbar, und das Integral hat den angegebenen Wert.

Satz 9.12 — *Es seien $f, g : I = [a, b] \to \mathbb{R}$ integrierbare Funktionen. Dann sind auch die folgenden Funktionen integrierbar:*

1. $|f|$, $\max\{f, g\}$, $\min\{f, g\}$,
2. *für jedes $p \in [1, \infty)$ die Potenzen $|f|^p$,*
3. *das Produkt fg.*

Beweis. Zu 1.: Es gilt

$$\max\{f, g\} = \frac{f+g}{2} + \frac{|f-g|}{2} \quad \text{und} \quad \min\{f, g\} = \frac{f+g}{2} - \frac{|f-g|}{2}.$$

Es genügt daher zu zeigen, dass $|f|$ integrierbar ist.

Für beliebige reelle Zahlen x, y in einem Teilintervall J von I gilt

$$\big|\, |f(x)| - |f(y)| \,\big| \leq |f(x) - f(y)| \leq \sup f(J) - \inf f(J)$$

und somit

$$\sup|f(J)| - \inf|f(J)| \leq \sup f(J) - \inf f(J).$$

Für jede Zerlegung $\mathcal{Z}$ von I folgt

$$0 \leq O(|f|, \mathcal{Z}) - U(|f|, \mathcal{Z}) \leq O(f, \mathcal{Z}) - U(f, \mathcal{Z})$$

und im Grenzübergang

$$0 \leq O(|f|, I) - U(|f|, I) \leq O(f, I) - U(f, I).$$

Ist f integrierbar, so verschwindet die rechte Differenz, also auch die linke. Damit erweist sich $|f|$ als integrierbar.
Zu 2.: Der Mittelwertsatz sagt für die Funktion $y \mapsto y^p$ und Zahlen $y_1, y_2 \geq 0$, dass $y_1^p - y_2^p = p \cdot (y_1 - y_2)\xi^{p-1}$ für eine Zahl ξ zwischen y_1 und y_2. Insbesondere gilt für beliebige Punkte x_1, x_2 in einem Teilintervall J von I, dass

$$|f(x_1)|^p - |f(x_2)|^p = (|f(x_1)| - |f(x_2)|) \cdot p\xi^{p-1}$$

für einen Wert ξ zwischen $|f(x_1)|$ und $|f(x_2)|$. Da f beschränkt ist, gilt in jedem Falle $\xi \leq \|f\|_I$. Es folgt:

$$\sup|f|^p(J) - \inf|f|^p(J) \leq p(\sup|f|(J) - \inf|f|(J))\|f\|_I^{p-1}. \tag{9.2}$$

Insbesondere gilt

$$O(|f|^p, \mathcal{Z}) - U(|f|^p, \mathcal{Z}) \leq p\left(O(|f|, \mathcal{Z}) - U(|f|, \mathcal{Z})\right)\|f\|_I^{p-1} \tag{9.3}$$

für jede Zerlegung $\mathcal{Z}$, und im Grenzübergang zu feineren Zerlegungen:

$$0 \leq O(|f|^p, I) - U(|f|^p, I) \leq p\left(O(|f|, I) - U(|f|, I)\right)\|f\|_I^{p-1}.$$

Nach 1. ist $|f|$ integrierbar. Damit verschwindet die rechte Seite, und wir schließen, dass auch $|f|^p$ integrierbar ist.
Zu 3.: Wegen

$$fg = \frac{(f+g)^2 - (f-g)^2}{4} \tag{9.4}$$

folgt die Behauptung aus Teilaussage 2. mit der Wahl $p = 2$, angewendet auf $f + g$ und $f - g$. □

Satz 9.13 — *Es seien $a < b < c$ reelle Zahlen und $f : [a, c] \to \mathbb{R}$ eine reelle Funktion. f ist genau dann integrierbar, wenn $f|_{[a,b]}$ und $f|_{[b,c]}$ integrierbar sind, und in diesem Falle gilt*

$$\int_a^b f(x)dx + \int_b^c f(x)dx = \int_a^c f(x)dx. \tag{9.5}$$

Beweis. Es sei $\mathcal{Z}$ eine Zerlegung von $[a,c]$. Durch Hinzunahme von b als Zerlegungspunkt erhält man eine Verfeinerung $\tilde{\mathcal{Z}} = (t_0, \ldots, t_N) \geq \mathcal{Z}$ mit $b = t_n$ für ein $n < N$. Dann sind $\mathcal{Z}' = (t_0, \ldots, t_n)$ und $\mathcal{Z}'' = (t_n, \ldots, t_N)$ Zerlegungen von $[a,b]$ bzw. $[b,c]$, und es gilt

$$U(f, \mathcal{Z}) \leq U(f, \tilde{\mathcal{Z}}) = U(f|_{[a,b]}, \mathcal{Z}') + U(f|_{[b,c]}, \mathcal{Z}'').$$

Umgekehrt kann man beliebige Zerlegungen $\mathcal{Z}'$ und $\mathcal{Z}''$ von $[a,b]$ bzw. $[b,c]$ auf diese Weise erhalten. Durch Übergang zu den Suprema folgt:

$$U(f,[a,c]) = U(f,[a,b]) + U(f,[b,c]). \tag{9.6}$$

Ebenso schließt man für die Oberintegrale und erhält

$$O(f,[a,c]) = O(f,[a,b]) + O(f,[b,c]).$$

Wir subtrahieren die beiden Gleichungen voneinander und erhalten

$$0 \leq O(f,[a,c]) - U(f,[a,c]) = (O(f,[a,b]) - U(f,[a,b])) + (O(f,[b,c]) - U(f,[b,c])). \tag{9.7}$$

Wir schließen wie folgt: Sind $f|_{[a,b]}$ und $f|_{[b,c]}$ integrierbar, so verschwindet die rechte Seite von (9.7). Folglich ist f auf $[a,c]$ integrierbar, und es gilt wegen (9.6) die im Satz behauptete Additivität des Integrals.

Ist umgekehrt f auf $[a,c]$ integrierbar, so verschwindet die rechte Seite von (9.7). Da die einzelnen Summanden aber in jedem Falle ≥ 0 sind, müssen sie jeweils für sich verschwinden. Folglich ist f auf $[a,b]$ wie auf $[b,c]$ integrierbar. □

Definition 9.14 — Ist $f : [a,b] \to \mathbb{R}$ integrierbar, so setzen wir

$$\int_b^a f(x)dx := -\int_a^b f(x)dx.$$

Folgerung 9.15 — *Ist I ein abgeschlossenes Intervall und $f : I \to \mathbb{R}$ eine integrierbare Funktion, so gilt für beliebige $a, b, c \in I$*

$$\int_a^b f(x)dx + \int_b^c f(x)dx = \int_a^c f(x)dx.$$

9.2 Der Hauptsatz der Differential- und Integralrechnung

Lemma 9.16 — *Es sei $K \subset \mathbb{C}$ eine kompakte Menge und $f : K \to \mathbb{C}$ stetig. Dann ist f sogar gleichmäßig stetig.*

Beweis. Andernfalls gibt es ein $\varepsilon > 0$ und zu jedem $\delta > 0$ Punkte $x, y \in K$ mit $|x - y| < \delta$ und $|f(x) - f(y)| \geq \varepsilon$. Insbesondere finden wir für jedes $n \in \mathbb{N}$ Punkte $x_n, y_n \in K$ mit $|x_n - y_n| < 1/n$ und $|f(x_n) - f(y_n)| > \varepsilon$. Da K kompakt ist, gibt es eine konvergente Teilfolge (x_{n_k}) mit Grenzwert x. Weiter gibt es eine konvergente Teilfolge $y_{n_{k_\ell}}$ mit Grenzwert y. Wegen der Stetigkeit von f gilt auch $f(x_{n_{k_\ell}}) \to f(x)$ und $f(y_{n_{k_\ell}}) \to f(y)$. Das führt auf den folgenden Widerspruch:

1. Da $|x_{n_{k_\ell}} - y_{n_{k_\ell}}| < 1/n_{k_\ell}$ für alle ℓ, folgt im Grenzübergang für $\ell \to \infty$: $x = y$.

2. Aus $|f(x_{n_{k_\ell}}) - f(y_{n_{k_\ell}})| > \varepsilon$ folgt im Grenzübergang $|f(x) - f(y)| \geq \varepsilon$!

□

Satz 9.17 — *Es sei $f : I = [a, b] \to \mathbb{R}$ stetig. Dann ist f integrierbar.*

Beweis. Nach dem vorigen Lemma ist f gleichmäßig stetig. Ist daher $\varepsilon > 0$ vorgegeben, so existiert ein $\delta > 0$ derart, dass $|f(x) - f(y)| < \frac{\varepsilon}{b-a}$ für alle $x, y \in [a, b]$ mit $|x - y| < \delta$. Wir wählen $N \in \mathbb{N}$ so groß, dass $(b - a)/N < \delta$, und betrachten die Zerlegung $\mathcal{Z} = (t_0, \ldots, t_N)$ mit $t_i = a + i(b - a)/N$. Für jedes Teilintervall $J = [t_{i-1}, t_i]$ folgt

$$\sup f(J) - \inf f(J) \leq \frac{\varepsilon}{b - a}.$$

Für die Differenz von Ober- und Untersumme ergibt sich

$$\begin{aligned} O(f, \mathcal{Z}) - U(f, \mathcal{Z}) &= \sum_{i=1}^{N} \frac{b-a}{N} (\sup f([t_{i-1}, t_i]) - \inf f([t_{i-1}, t_i])) \\ &\leq \sum_{i=1}^{N} \frac{b-a}{N} \frac{\varepsilon}{b-a} = \varepsilon \end{aligned}$$

Nach dem Integrationskriterium Satz 9.10 ist f integrierbar. □

Satz 9.18 *(Mittelwertsatz der Integralrechnung) — Es sei I ein Intervall und $f : I \to \mathbb{R}$ eine stetige Funktion. Dann gibt es zu beliebigen $a, b \in I$ ein ξ zwischen a und b mit*

$$\int_a^b f(x)dx = (b - a)f(\xi).$$

Beweis. Ohne Einschränkung ist $a < b$ und $I = [a, b]$. Da f stetig und I kompakt ist, gibt es Punkte x' und x'' in I mit

$$f(x') = \inf f(I) \quad \text{und} \quad f(x'') = \sup f(I).$$

Es folgt

$$f(x')\,(b-a) \le \int_a^b f(x)dx \le f(x'')\,(b-a)$$

wegen der Monotonie des Integrals. Nach dem Zwischenwertsatz 6.18 für stetige Funktionen gibt es ein ξ zwischen x' und x'' mit

$$\frac{1}{b-a}\int_a^b f(x)dx = f(\xi).$$

Das war zu zeigen. □

Satz 9.19 *(Hauptsatz der Differential- und Integralrechnung) — Es sei I ein Intervall und $f : I \to \mathbb{R}$ eine stetige Funktion. Ferner sei $a \in I$ und $F : I \to \mathbb{R}$ die durch*

$$F(x) := \int_a^x f(t)dt$$

definierte Funktion. Dann ist F differenzierbar und hat die Ableitung $F' = f$.

Beweis. Es seien b, x aus I. Nach dem Mittelwertsatz gibt es ein γ_x zwischen b und x mit

$$(x-b)f(\gamma_x) = \int_b^x f(t)dt = \int_a^x f(t)dt - \int_a^b f(t)dt = F(x) - F(b).$$

Für $x \to b$ geht auch $\gamma_x \to b$, weil γ_x zwischen x und b liegt. Außerdem ist die Funktion f stetig. Daher existiert der Grenzwert

$$\lim_{x\to b} \frac{F(x)-F(b)}{x-b} = \lim_{x\to b} f(\gamma_x) = f(b).$$

□

Definition 9.20 — Es sei $I \subset \mathbb{R}$ ein Intervall und $f : I \to \mathbb{R}$ eine Funktion.

1. Eine Funktion $F : I \to \mathbb{R}$ heißt Stammfunktion von f, wenn F differenzierbar ist und die Ableitung $F' = f$ hat.

2. Eine Funktion $F : I \to \mathbb{R}$ heißt unbestimmtes Integral von f, wenn f auf jedem abgeschlossenen Teilintervall $[a, b] \subset I$ integrierbar ist, und wenn das Integral durch $\int_a^b f(x)dx = F(b) - F(a)$ gegeben ist. Häufig werden unbestimmte Integrale durch $\int f(x)dx$ bezeichnet.

Lemma 9.21 — *1. Falls f eine Stammfunktion besitzt, ist diese bis auf eine additive Konstante eindeutig bestimmt.*
2. Falls f ein unbestimmtes Integral besitzt, ist dieses bis auf eine additive Konstante eindeutig bestimmt.

Beweis. Zu 1.: Sind F und G zwei Stammfunktionen von f, so folgt $(F-G)' = f - f = 0$. Die Funktion $F - G$ ist also auf I konstant.
Zu 2.: Sind F und G unbestimmte Integrale von f, so ist $F-G$ ein unbestimmtes Integral von $f - f = 0$. Es folgt $(F-G)(b) - (F-G)(a) = \int_a^b 0dx = 0$. Also ist $(F-G)(b) = (F-G)(a) = const.$ □

Bemerkung 9.22 — Mit diesen Bezeichnungen sagt der Hauptsatz der Differential- und Integralrechnung, dass jede stetige Funktion eine Stammfunktion besitzt und dass für stetige Funktionen Stammfunktionen und unbestimmte Integrale dasselbe sind.

Beispiel 9.23 — Die Funktion

$$\operatorname{sgn} : \mathbb{R} \to \mathbb{R}, \quad f(x) = \begin{cases} 1, & x > 0, \\ 0, & \text{falls } x = 0, \\ -1, & x > 0, \end{cases}$$

besitzt ein unbestimmtes Integral, nämlich die Betragsfunktion $|-|$, aber keine Stammfunktion, da die Betragsfunktion bekanntlich im Nullpunkt nicht differenzierbar ist.

Bemerkung 9.24 — Ist $f : I \to \mathbb{R}$ auf jedem Teilintervall integrierbar, so ist jede Stammfunktion F von f ein unbestimmtes Integral von f. Denn angenommen, $[a,b] \subset I$ ist ein abgeschlossenes Teilintervall von I und $\mathcal{Z} = (t_0, \ldots, t_n)$ eine Zerlegung von $[a,b]$. Nach dem Mittelwertsatz der Differentialrechnung, angewendet auf F, gibt es zu jedem i ein ξ_i zwischen t_{i-1} und t_i mit $f(\xi_i)(t_i - t_{i-1}) = F(t_i) - F(t_{i-1})$. Damit folgt:

$$F(b) - F(a) = \sum_{i=1}^{n} (F(t_i) - F(t_{i-1})) = \sum_{i=1}^{n} (t_i - t_{i-1}) f(\xi_i).$$

Wegen $\inf f([t_{i-1}, t_i]) \leq f(\xi_i) \leq \sup f([t_{i-1}, t_i])$ gilt

$$U(f, \mathcal{Z}) \leq F(b) - F(a) \leq O(f, \mathcal{Z}).$$

Diese Abschätzung gilt für jede Zerlegung. Außerdem ist f nach Voraussetzung auf $[a,b]$ integrierbar, und somit folgt $F(b) - F(a) = \int_a^b f(x)dx$.

Bemerkung 9.25 — Mit dem Hauptsatz der Differential- und Integralrechnung findet man mit einem Schlage unbestimmte Integrale für eine große Klasse

von Funktionen:

$f(x)$	$\int f(x)dx$
X^α, für $\alpha \in \mathbb{R} \setminus \{-1\}$	$\frac{1}{1+\alpha}X^{1+\alpha}$
X^{-1} auf $\mathbb{R} \setminus \{0\}$	$\log\lvert X\rvert$
$\frac{1}{1+X^2}$	$\arctan(X)$
$\sin(X)$	$-\cos(X)$
$\cos(X)$	$\sin(X)$
$\tan(X)$ auf $(-\pi/2, \pi/2)$	$\log(\cos(X))$
$\exp(X)$	$\exp(X)$

Satz 9.26 *(verallgemeinerter Mittelwertsatz der Integralrechnung) —*
Es seien $f, g : [a, b] \to \mathbb{R}$ stetige Funktionen, und es gelte $g(x) \neq 0$ für alle $x \in (a, b)$. Dann gibt es ein $\xi \in (a, b)$ mit

$$\int_a^b f(x)g(x)dx = f(\xi)\int_a^b g(x)dx.$$

Beweis. Unter den Voraussetzungen des Satzes hat g auf (a, b) stets dasselbe Vorzeichen. Für die Funktion $G(x) := \int_a^x g(t)dt$, die nach dem Hauptsatz differenzierbar ist, gilt $G(x) \neq 0$ für alle $x > a$. Damit sind für die Funktionen G und $F(x) := \int_a^x f(t)g(t)dt$ die Voraussetzungen des verallgemeinerten Mittelwertsatzes 7.21 erfüllt, und es gibt ein $\xi \in (a, b)$ mit

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

Ersetzt man noch $F(b) - F(a) = \int_a^b f(x)g(x)dx$ und $G(b) - G(a) = \int_a^b g(x)dx$, so erhält man die Behauptung des Satzes. □

9.3 Partielle Integration

Satz 9.27 *(Partielle Integration) — Für $f, g \in \mathcal{C}^1([a, b])$ gilt*

$$\int_a^b f(x)g'(x)dx = f(x)g(x)\Big|_a^b - \int_a^b f'(x)g(x)dx.$$

Dabei bedeutet $f(x)g(x)\Big|_a^b := f(b)g(b) - f(a)g(a)$.

Schreibt man die Regel mit unbestimmten Integralen, so bekommt sie die folgende Form

$$\int fg'dx = fg - \int f'gdx.$$

Beweis des Satzes. Nach der Produktregel für das Differenzieren gilt

$$(fg)' = fg' + f'g.$$

Das bedeutet nach dem Hauptsatz

$$fg\Big|_a^b = \int_a^b (fg)'dx = \int_a^b fg'dx + \int_a^b f'gdx.$$

□

Beispiel 9.28 — Ein häufig brauchbarer Trick besteht darin, $f = x$ zu wählen:

$$\int g(x)dx = xg(x) - \int xg'(x).$$

Das funtioniert zum Beispiel sehr gut, wenn man den Logarithmus integrieren will: Es bezeichne F_n eine Stammfunktion für $\log^n$. Dann gilt:

$$F_n = \int \log(x)^n dx = x\log(x)^n - \int xn\log(x)^{n-1}\frac{1}{x}dx = x\log(x)^n - nF_{n-1}.$$

Beginnt man mit $F_0 = x$, so erhält man rekursiv

$$\begin{aligned}
\int \log(x)dx &= x(\log(x) - 1),\\
\int \log(x)^2dx &= x(\log(x)^2 - 2\log(x) + 2),\\
\int \log(x)^3dx &= x(\log(x)^3 - 3\log(x)^2 + 6\log(x) - 6),
\end{aligned}$$

etc.

Beispiel 9.29 — Partielle Integration liefert

$$\int \cos(x)^2dx = \sin(x)\cos(x) + \int \sin(x)^2dx = \sin(x)\cos(x) + \int (1 - \cos(x)^2)dx,$$

also

$$2\int \cos(x)^2dx = x + \sin(x)\cos(x).$$

Beispiel 9.30 — Wir betrachten für $n \geq 0$ das Integral

$$I_n := \int_0^{\pi/2} \sin(x)^n dx.$$

Man verifiziert sofort

$$I_0 = \pi/2 \quad \text{und} \quad I_1 = 1.$$

Weil $0 \leq \sin(x) \leq 1$ für alle $x \in [0, \pi/2]$, besteht die Abschätzung

$$I_{n+1} \leq I_n \quad \text{für } n \geq 0. \tag{9.8}$$

Weiter sieht man mit partieller Integration für $n \geq 2$:

$$\begin{aligned}\int_0^{\pi/2} \sin(x)^n(x)dx &= -\cos(x)\sin(x)^{n-1}(x)\Big|_0^{\pi/2} \\ &\quad +(n-1)\int_0^{\pi/2} \cos(x)^2 \sin(x)^{n-2}dx. \\ &= (n-1)\int_0^{\pi/2} \sin(x)^{n-2}dx \\ &\quad -(n-1)\int_0^{\pi/2} \sin(x)^n dx.\end{aligned}$$

Das ergibt die Rekursionsgleichung

$$I_n = \frac{n-1}{n} I_0.$$

Für $n = 2k$ heißt das

$$I_{2k} = \frac{2k-1}{2k} \cdot \frac{2k-3}{2k-2} \cdot \ldots \cdot \frac{3}{4} \cdot \frac{1}{2} \cdot \frac{\pi}{2}$$

und für $n = 2k+1$

$$I_{2k+1} = \frac{2k}{2k+1} \cdot \frac{2k-2}{2k-1} \cdot \ldots \cdot \frac{2}{3}.$$

Nun erinnern wir uns, dass wir im Abschnitt §3 die folgende Folge betrachtet haben:

$$a_n = \frac{2}{1} \cdot \frac{4}{3} \cdot \ldots \frac{2n}{2n-1} \cdot \frac{1}{\sqrt{n}} = \frac{2}{3} \cdot \frac{4}{5} \cdot \ldots \cdot \frac{2n-2}{2n-1} \cdot 2\sqrt{n}.$$

Mit dieser Bezeichnung gilt also

$$\frac{a_n}{\sqrt{n}} = 2\, I_{2n-1} \quad \text{und} \quad \frac{\pi}{\sqrt{n}\, a_n} = 2\, I_{2n}.$$

Wegen der Ungleichung (9.8) erhält man

$$\frac{a_{n+1}}{\sqrt{n+1}} \leq \frac{\pi}{\sqrt{n}\, a_n} \leq \frac{a_n}{\sqrt{n}}$$

und umgeformt

$$\sqrt{\frac{n+1}{n}} a_{n+1} a_n \leq \pi \leq a_n^2$$

Wir haben gesehen, dass die Folge a_n konvergiert. Es folgt

$$\lim_{n\to\infty} a_n = \sqrt{\pi}.$$

Satz 9.31 *(Formel von Wallis)*

$$\frac{\pi}{2} = \frac{2}{1} \cdot \frac{2}{3} \cdot \frac{4}{3} \cdot \frac{4}{5} \cdot \ldots$$

Beweis. Behauptet wird, dass die Folge

$$b_1 = \frac{2}{1}, \quad b_2 = \frac{2}{1} \cdot \frac{2}{3}, \quad b_3 = \frac{2}{1} \cdot \frac{2}{3} \cdot \frac{4}{3}, \ldots$$

gegen $\pi/2$ konvergiert. Da die Faktoren $\frac{2n}{2n-1}$ und $\frac{2n}{2n+1}$ gegen 1 konvergieren, genügt es zu zeigen, dass die Folge

$$c_n = \frac{2}{1} \cdot \frac{2}{3} \cdot \ldots \cdot \frac{2n}{2n-1} \cdot \frac{2n}{2n+1}$$

gegen $\pi/2$ konvergiert. Nun gilt

$$c_n = \frac{1}{2} a_n a_{n+1} \sqrt{\frac{n}{n+1}} \xrightarrow{n\to\infty} \frac{\pi}{2}.$$

□

Satz 9.32 — *Es gilt*

$$\binom{2n}{n} \sim \frac{4^n}{\sqrt{\pi n}} \tag{9.9}$$

in dem Sinne, dass der Quotient $\binom{2n}{n} \Big/ \frac{4^n}{\sqrt{\pi n}}$ *für* $n \to \infty$ *gegen* 1 *konvergiert.*

Beweis. Dies ist einfach Formel (3.5), wo der Grenzwert $p = \lim a_n$ durch den gerade berechneten Wert $p = \sqrt{\pi}$ ersetzt ist. □.

9.4 Substitutionsregel

Satz 9.33 *(Substitutionsregel) — Es sei* $f : I \to \mathbb{R}$ *eine stetige Funktion und* $\varphi : [a, b] \to I$ *eine stetig differenzierbare Funktion. Dann gilt*

$$\int_a^b f(\varphi(t))\varphi'(t)dt = \int_{\varphi(a)}^{\varphi(b)} f(x)dx.$$

Beweis. Es sei $F : I \to \mathbb{R}$ eine Stammfunktion von f. Nach der Kettenregel gilt

$$(F \circ \varphi)'(t) = F'(\varphi(t))\varphi'(t) = f(\varphi(t))\varphi'(t).$$

Demnach ist $F \circ \varphi$ eine Stammfunktion für $(f \circ \varphi) \cdot \varphi'$. Es folgt:

$$\int_a^b f(\varphi(t))\varphi'(t)dt = F(\varphi(b)) - F(\varphi(a)) = \int_{\varphi(a)}^{\varphi(b)} f(x)dx.$$

□

Beispiel 9.34 — Mit der Substitution $x = \varphi(t) := t^4 + 1$ bekommt man

$$\begin{aligned}\int t^3 \log(1+t^4)dt &= \frac{1}{4}\int \log(x)dx \\ &= \frac{1}{4}\left(x\log(x) - x\right) \\ &= \frac{1}{4}((1+t^4)\log(1+t^4) - (t^4+1)).\end{aligned}$$

Beispiel 9.35 — Integrale der Form

$$\int R(\sin(\varphi), \cos(\varphi))d\varphi,$$

wo R irgendeine rationale Funktion in zwei Variablen bezeichnet, können durch die folgende Substitution in Integrale über rationale Funktionen in einer Variablen überführt werden:

$$\varphi := 2\arctan(t), \qquad t = \tan(\varphi/2).$$

Man findet:

$$\sin(\varphi) = \frac{2t}{1+t^2},\ \cos(\varphi) = \frac{1-t^2}{1+t^2},\ \tan(\varphi) = \frac{2t}{1-t^2},\ \varphi'(t) = \frac{2}{1+t^2},$$

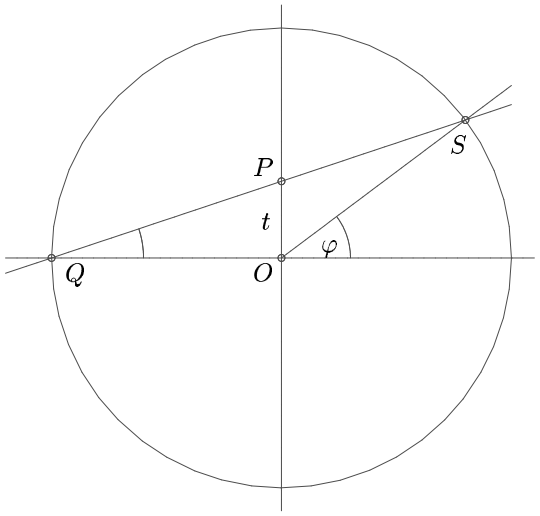
und das bedeutet:

$$\int R(\sin(\varphi), \cos(\varphi)d\varphi = 2\int R\left(\frac{2t}{1+t^2}, \frac{1-t^2}{1+t^2}\right)\frac{dt}{1+t^2}.$$

Die rechte Seite ist jetzt das Integral über eine rationale Funktion, das man nach den üblichen Regeln löst, gegebenenfalls mit Partialbruchzerlegung. Zum Beispiel:

$$\int \frac{d\varphi}{\sin(\varphi)} = \int \frac{1+t^2}{2t}\cdot\frac{2dt}{1+t^2} = \int \frac{dt}{t} = \log(t) = \log\tan\frac{\varphi}{2}.$$

Der geometrische Hintergrund für diese Regel ist der folgende:



Es sei $S = (\cos(\varphi), \sin(\varphi))$ ein Punkt auf dem Einheitskreis und $P = (0, t)$ der Schnittpunkt der Geraden durch S und $Q = (-1, 0)$ mit der y-Achse. Durchläuft φ den Bereich $(-\pi, \pi)$, also S den gesamten Kreis ohne den Punkt Q, so läuft t durch alle relle Zahlen und umgekehrt. Aus dem Diagramm ergeben sich leicht die Umrechnungsformeln

$$t = \tan(\varphi/2) = \frac{\sin(\varphi)}{1 + \cos(\varphi)}. \tag{9.10}$$

Die Abbildung $t \mapsto (\frac{1-t^2}{1+t^2}, \frac{2t}{1+t^2})$ ist eine Parametrisierung des Kreises durch rationale Funktionen. Man spricht deshalb von einer rationalen Parametrisierung des Kreises, im Gegensatz etwa zu der transzendenten Parametrisierung $\varphi \mapsto (\cos(\varphi), \sin(\varphi))$.

Bemerkung 9.36 — Ein pythagoräisches Zahlentripel ist ein Tripel (a, b, c) von natürlichen Zahlen mit der Eigenschaft $a^2 + b^2 = c^2$. Nach dem Satz des Pythagoras ist ein Dreieck mit den Seiten a, b und c rechtwinklig. Bekannt sind seit Jahrtausenden die Lösungen

$$3^2 + 4^2 = 5^2, \quad 5^2 + 12^2 = 13^2, \quad 7^2 + 24^2 = 25^2.$$

Wenn man ein Tripel (a, b, c) hat, kann durch Erweitern mit einer natürlichen Zahl d trivialerweise ein neues bekommen: (da, db, dc). Besitzen umgekehrt zwei der drei Zahlen a, b, c einen gemeinsamen Teiler d, so teilt d auch die dritte Zahl, und man kann d herauskürzen. Interessant sind also speziell solche Tripel, die paarweise teilerfremd sind.

Die oben besprochene rationale Parametrisierung des Kreises erlaubt es, einen Überblick über *alle* teilerfremden Lösungen zu gewinnen. Denn ist (a, b, c) ein pythagoräisches Zahlentripel, so ist $(a/c, b/c)$ ein Punkt auf dem Einheitskreis mit rationalen Koordinaten. Ist umgekehrt (x, y) ein Punkt auf dem Einheitskreis mit positiven rationalen Koordinaten und ist c der Hauptnenner von x und y, so ist (xc, yc, c) ein teilerfremdes pythagoräisches Zahlentripel. Wir müssen also alle Punkte auf dem Einheitskreis mit rationalen Koordinaten bestimmen.

Mit den Bezeichnungen aus Beispiel 9.35 gilt nun: Sind x und y rational, so ist auch $t = \frac{y}{1+x}$ rational, und ist umgkehrt t rational, so auch $y = 2t/(1 + t^2)$ und $x = (1-t^2)/(1+t^2)$. Dabei sind x und y genau dann positiv, wenn $t \in (0, 1)$. Mit dem Ansatz $t = \frac{m}{n}$, $m, n \in \mathbb{N}$, erhält man

$$a := n^2 - m^2, \quad b := 2mn, \quad c := n^2 + m^2.$$

Das Tripel (a, b, c) ist genau dann teilerfremd, wenn m und n teilerfremd und nicht beide ungerade sind. Zulässige t mit kleinem Nenner sind also

$$t = \frac{1}{2}, \frac{2}{3}, \frac{1}{4}, \frac{3}{4}, \frac{2}{5}, \frac{4}{5},$$

was auf die Tripel

$$(a, b, c) = (3, 4, 5), (5, 12, 13), (15, 8, 17), (7, 24, 25), (23, 20, 29), (9, 40, 41),$$

führt.

Beispiel 9.37 — Integrale der Form $\int R(x, \sqrt{1-x^2})dx$, wo R irgendeine rationale Funktion in zwei Variablen ist, lassen sich wie folgt behandeln: Durch die Substitution $x = \sin(\psi)$ erhält man

$$\int R(x, \sqrt{1-x^2})dx = \int R(\sin(\psi), \cos(\psi))cos(\psi)d\psi.$$

Falls sich dieses Integral nicht mit anderen Mitteln integrieren lässt, kann man die im vorigen Beispiel besprochene Substitution $\tan(\psi/2) = t$ anwenden. Die beiden Substitutionen kann man natürlich auch in einem Schritt durchführen und bekommt mit $x = \frac{2t}{1+t^2}$, $t = (1 - \sqrt{1+x^2})/x$:

$$\int R(x, \sqrt{1-x^2})dx = \int 2R(\frac{2t}{1+t^2}, \frac{1-t^2}{1+t^2})\frac{1-t^2}{(1+t^2)^2}dt.$$

Beispiel 9.38 — Integrale der Form $\int R(x, \sqrt{1+x^2})dx$, wo R irgendeine rationale Funktion in zwei Variablen ist, lassen sich mithilfe der Substitution $x = \tan(\psi)$ lösen. Man bekommt

$$\int R(x, \sqrt{1+x^2})dx = \int R\left(\tan(\psi), \frac{1}{\cos(\psi)}\right)\frac{d\psi}{cos(\psi)^2}.$$

Auch hier muss man gegebenenfalls erneut substituieren, etwa mit $t = \tan(\psi/2)$. Und in einem Schritt bedeutet dies $x = \frac{2t}{1-t^2}$, $t = (\sqrt{1+x^2} - 1)/x$ und

$$\int R(x, \sqrt{1+x^2})dx = \int 2R\left(\frac{2t}{1-t^2}, \frac{1+t^2}{1-t^2}\right)\frac{1+t^2}{(1-t^2)^2}dt.$$

9.5 Hyperbelfunktionen

Wir definieren für beliebige komplexe Argumente den sinus hyperbolicus und den cosinus hyperbolicus durch die Formeln

$$\sinh(z) := \sum_{k=0}^{\infty} \frac{z^{2k+1}}{(2k+1)!} \quad \text{und} \quad \cosh(z) := \sum_{k=0}^{\infty} \frac{z^{2k}}{(2k)!}.$$

Beide Reihen haben unendlichen Konvergenzradius.

Lemma 9.39 *(Rechenregeln für* sinh *und* cosh*) — Für alle $z \in \mathbb{C}$ gilt:*

1. $\sinh(z) := (e^z - e^{-z})/2$, $\cosh(z) := (e^z + e^{-z})/2$.
2. $\cosh(z)^2 - \sinh(z)^2 = 1$.
3. $\sinh(z+w) = \sinh(z)\cosh(w) + \cosh(z)\sinh(w)$.
4. $\cosh(z+w) = \cosh(z)\cosh(w) + \sinh(z)\sinh(w)$.
5. $\sinh(iz) = i\sin(z)$, $\cosh(iz) = \cos(z)$.

Beweis. Aussage 1. ergibt sich sofort aus der Definition, und alle anderen Aussagen folgen hieraus und den Eigenschaften der e-Funktion. □

Analog definiert man den tangens hyperbolicus und den cotangens hyperbolicus:

$$\tanh(z) := \frac{\sinh(z)}{\cosh(z)} = \frac{e^{2z}-1}{e^{2z}+1}$$

für $z \neq (2k-1)\pi i$, $k \in \mathbb{Z}$, und

$$\operatorname{ctgh}(z) := \frac{\cosh(z)}{\sinh(z)} = \frac{e^{2z}+1}{e^{2z}-1}$$

für $z \neq 2k\pi i$, $k \in \mathbb{Z}$.

Für reelle Argumente x sind $\sinh(x)$, $\cosh(x)$, $\tanh(x)$ und $\operatorname{ctgh}(x)$ ebenfalls reell. Diese reellen Funktionen sind auf ganz $\mathbb{R}$ bzw. im Falle des ctgh auf $\mathbb{R}\setminus\{0\}$ definiert und dort uendlich oft differenzierbar mit den folgenden Ableitungen:

1. $\sinh'(x) = \cosh(x)$, $\cosh'(x) = \sinh(x)$.
2. $\tanh'(x) = 1 - \tanh^2(x) = \frac{1}{\cosh^2(x)}$.
3. $\operatorname{ctgh}'(x) = 1 - \operatorname{ctgh}^2(x) = -\frac{1}{\sinh^2(x)}$.

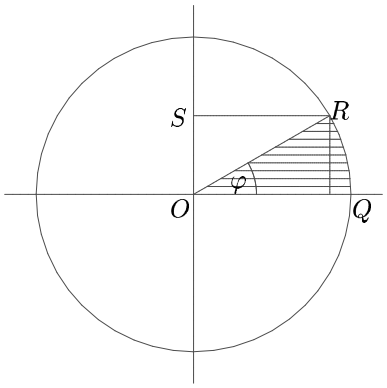
Die Funktionen $\sinh : \mathbb{R} \to \mathbb{R}$ und $\tanh : \mathbb{R} \to (-1,1)$ sind bijektiv und besitzen Umkehrfunktionen $\operatorname{arsinh} : \mathbb{R} \to \mathbb{R}$, den *area sinus hyperbolicus*, und $\operatorname{artanh} : (-1,1) \to \mathbb{R}$, den *area tangens hyperbolicus*. Diese lassen sich auch wie folgt ausdrücken:

$$\operatorname{arsinh}(x) = \log(x + \sqrt{1+x^2}) \quad \text{und} \quad \operatorname{artanh}(x) = \frac{1}{2}\log\frac{1+y}{1-y}.$$

Nach dem Satz über die Ableitung inverser Funktionen sind die Areafunktionen differenzierbar. Die Ableitungen sind

$$\operatorname{arsinh}'(x) = \frac{1}{\sqrt{1+x^2}} \quad \text{und} \quad \operatorname{artanh}'(x) = \frac{1}{1-x^2}$$

Beispiel 9.40 — Wir berechnen den Flächeninhalt des Kreissektors, der im Einheitskreis durch den Punkt $P = (\cos(\varphi), \sin(\varphi))$ bestimmt wird. Diese Fläche F ist gleich dem Flächenstück $OQPS$, vermindert um das Dreieck OPS.



Daraus folgt:

$$F + \frac{1}{2}\sin\varphi\cos\varphi = \int_0^{\sin(\varphi)} \sqrt{1-y^2}dy.$$

Mit der Substitution $y = \sin(t)$, $dy = \cos(t)dt$, erhält man

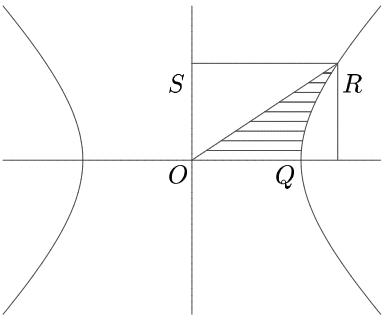
$$\begin{aligned} F + \frac{1}{4}\sin(2\varphi) &= \int_0^{\varphi} \cos(t)^2 dt \\ &= \left(\frac{t}{2} + \frac{1}{4}\sin(2t)\right)|_0^{\varphi} \\ &= \frac{\varphi}{2} + \frac{1}{4}\sin(2\varphi). \end{aligned}$$

Es folgt $F = \varphi/2$. Wir erhalten als geometrische Deutung des Parameters φ, dass φ gleich dem doppelten Flächeninhalt des durch φ bestimmten Kreissektors ist. Insbesondere ergibt sich für $\varphi = 2\pi$, dass der Flächeninhalt des Einheitskreises π ist. Im nächsten Semester werden wir sehen, dass sich φ auch als Länge des zugehörigen Kreisbogens auffassen lässt. Daher leitet sich der Name Bogenfunktion für die Umkehrfunktionen der Kreisfunktionen her.

Beispiel 9.41 — Die Hyperbelfunktionen lassen eine ähnliche Deutung an der Hyperbel mit der Gleichung

$$x^2 - y^2 = 1$$

zu. Für jedes $\varphi \in \mathbb{R}$ ist $P = (\cosh(\varphi), \sinh(\varphi))$ ein Punkt auf dieser Hyperbel.



Wir berechnen den Flächeninhalt F des Hyperbelsegments OQR auf dieselbe Weise wie oben:

$$F + \frac{1}{2}\sinh(\varphi)\cosh(\varphi) = \int_0^{\sinh(\varphi)} \sqrt{1+y^2}dy.$$

Mit der Substitution $y = \sinh(t)$, $dy = \cosh(t)dt$, erhält man hieraus

$$\begin{aligned} F + \frac{1}{4}\sinh(2\varphi) &= \int_0^{\varphi} \cosh(t)^2 dt \\ &= \int_0^{\varphi} \frac{1}{2}(\cosh(2t)+1)dt \\ &= \left(\frac{t}{2} + \frac{1}{4}\sinh(2t)\right)\Big|_0^{\varphi} = \frac{\varphi}{2} + \frac{1}{4}\sinh(2\varphi). \end{aligned}$$

Ähnlich wie im vorigen Beispiel hat φ die geometrische Deutung als das Doppelte des Flächeninhalts des Hyperbelsegments. Dieses Ergebnis rechtfertigt auch die Benennung der *Flächen*funktionen *area* sinus hyperbolicus etc.

9.6 Partialbruchzerlegung

Definition 9.42 — Es sei K ein Körper. Ein Polynom

$$f = f_n X^n + f_{n-1}X^{n-1} + \ldots + f_0 \in K[X]$$

heißt normiert, wenn der Leitkoeffizient $f_n = 1$ ist. Ein normiertes, nicht konstantes Polynom f heißt irreduzibel, wenn es keine normierten Polynome kleineren Grades $g, h \in K[X]$ mit der Eigenschaft $f = gh$ gibt.

Nach dem Fundamentalsatz der Algebra besitzt jedes normierte, nicht konstante Polynom $f \in \mathbb{C}[X]$ eine Zerlegung

$$f = (X - z_1)^{n_1} \cdot \ldots \cdot (X - z_s)^{n_s}$$

mit komplexen Nullstellen z_i und Vielfachheiten $n_i \in \mathbb{N}$. Die z_i und die n_i sind eindeutig bestimmt. Ist nämlich $z \in \mathbb{C}$ beliebig, so ist die Nullstellenordnung von f in z durch

$$n := \inf\{k \mid f^{(k)}(z) \neq 0\}$$

gegeben. Hier ist der Fall $n = 0$ durchaus zugelassen und bedeutet eben nichts anderes, als dass z keine Nullstelle von f ist.

Die einzigen komplexen irreduziblen Polynome sind also die Polynome vom Grad 1. Im Reellen ist dagegen das quadratische Polynom $X^2 + 1$ irreduzibel, denn wäre es zerlegbar, $X^2 + 1 = gh$, so müssten g und h lineare Polynome sein, d.h. $X^2 + 1$ müsste reelle Nullstellen haben, aber die einzigen Nullstellen sind bekanntlich i und $-i$.

Ist $f = f_n X^n + \ldots + f_0$ ein komplexes Polynome, so sei

$$\overline{f} = \overline{f}_n X^n + \ldots + \overline{f}_0$$

das zu f konjugierte komplexe Polynom. Die folgenden Rechenregeln sind offensichtlich.

Lemma 9.43 — *Für alle $f, g \in \mathbb{C}[X]$ und $z \in \mathbb{C}$ gilt:*

1. $\overline{f+g} = \overline{f} + \overline{g}$, $\overline{fg} = \overline{f}\,\overline{g}$.
2. $\overline{f(z)} = \overline{f}(\overline{z})$.
3. $\overline{f} = f$ *genau dann, wenn* $f \in \mathbb{R}[X]$.

Jedes reelle normierte Polynom $g \in \mathbb{R}[X]$ ist natürlich auch ein komplexes Polynom und besitzt als solches eine Zerlegung

$$g(X) = (X - z_1)^{n_1} \cdot \ldots \cdot (X - z_s)^{n_s}$$

mit komplexen Nullstellen. Konjugieren wir das Polynom, so erhalten wir

$$\overline{g}(X) = (X - \overline{z}_1)^{n_1} \cdot \ldots \cdot (X - \overline{z}_s)^{n_s}.$$

Aus der Eindeutigkeit der Faktorzerlegung folgt für jede Nullstelle z_i:

1. Entweder ist z_i reell,
2. oder es gibt einen eindeutig bestimmten Index $j \neq i$ mit $z_j = \overline{z}_i$ und $n_j = n_i$.

Die Nullstellen von g sind also entweder reell, oder sie kommen in Paaren $(z_i, \bar{z}_i)$ komplex konjugierter Zahlen vor. Wir fassen die Faktoren, die zu nicht reellen Nullstellen gehören, in diesem Sinne paarweise zusammen:

$$(X - z_i)(X - \overline{z}_i) = X^2 - 2\mathrm{Re}(z_i)X + |z_i|^2 = X^2 + u_i X + v_i.$$

Damit folgt:

Lemma 9.44 — *Es sei $g \in \mathbb{R}[X]$ ein normiertes Polynom. Dann existiert eine Zerlegung*

$$g(X) = \prod_{i=1}^{s} (X - x_i)^{n_i} \prod_{k=1}^{t} (X + u_k X + v_k)^{m_k}$$

mit paarweise verschiedenen irreduziblen Faktoren aus $\mathbb{R}[X]$. Die Zerlegung ist eindeutig bis auf die Anordnung der Faktoren.

Beispiel 9.45 — Es sei $\epsilon = \exp(2\pi i/8) = (1+i)/\sqrt{2}$. Das Polynom $X^4 + 1$ hat die Nullstellen ϵ^i mit $i = 1, 3, 5, 7$. Damit finden wir die Produktdarstellungen

$$\begin{aligned} X^4 + 1 &= (X + \epsilon)(X + \epsilon^3)(X + \epsilon^5)(X + \epsilon^7) \\ &= [(X + \epsilon)(X + \epsilon^7)] \cdot [(X + \epsilon^3)(X + \epsilon^5)] \\ &= (X^2 + \sqrt{2}X + 1)(X^2 - \sqrt{2}X + 1). \end{aligned}$$

Die erste Zeile ist die Zerlegung von $X^4 + 1$ in irreduzible Faktoren in $\mathbb{C}[X]$, die letzte Zeile die Zerlegung in irreduzible Faktoren in $\mathbb{R}[X]$. Letztere kann man auch mit dem folgenden Trick finden:

$$\begin{aligned} X^4 + 1 &= (X^4 + 2X^2 + 1) - 2X^2 \\ &= (X^2 + 1)^2 - (\sqrt{2}X)^2 = (X^2 + \sqrt{2}X + 1)(X^2 - \sqrt{2}X + 1). \end{aligned}$$

Lemma 9.46 — *Es sei $K = \mathbb{R}$ oder $\mathbb{C}$. Es sei $g \in K[X]$ ein nicht konstantes Polynom, $z \in K$ eine Nullstelle von g der Vielfachheit n und $h \in K[X]$ so, dass $g(X) = (X - z)^n h(X)$. Zu jedem $f \in K[X]$ gibt es eindeutig bestimmte $b \in K$ und $\tilde{f} \in K[X]$ derart, dass*

$$\frac{f(X)}{g(X)} = \frac{b}{(X - z)^n} + \frac{\tilde{f}(X)}{(X - z)^{n-1} h(X)}.$$

Beweis. Wenn es überhaupt b und $\tilde{f}$ mit den verlangten Eigenschaften gibt, so muss gelten

$$f(X) = bh(X) + (X - z)\tilde{f}(X).$$

Dadurch ist b eindeutig festgelegt: $b = f(z)/h(z)$. Definiert man umgekehrt b auf diese Weise, so hat das Polynom $f(X) - bh(X)$ bei z eine Nullstelle.

Durch Polynomdivision findet man das eindeutig bestimmte Polynom $\tilde{f}(X) = (f(X) - bh(X))/(X - z)$. □

Satz 9.47 *(Partialbruchzerlegung, komplexe Fassung) — Es sei $F = f/g$ eine rationale Funktion mit $f, g \in \mathbb{C}[X]$, g normiert vom Grad n. Es sei*

$$g(X) = (X - z_1)^{n_1} \cdot \ldots \cdot (X - z_s)^{n_s}, \quad s \in \mathbb{N}_0, z_i \in \mathbb{C}$$

die Zerlegung von g in Linearfaktoren. Dann gibt es ein eindeutig bestimmtes Polynom $r \in \mathbb{C}[X]$ und eindeutig bestimmte Zahlen a_{ij}, $i = 1, \ldots, s$, $j = 1, \ldots, n_i$, so, dass

$$F = r + \sum_{i=1}^{s} \sum_{j=1}^{n_i} \frac{a_{ij}}{(X - z_i)^j}. \tag{9.11}$$

Diese Zerlegung heißt die Partialbruchzerlegung von F.

Beweis. Wir beweisen Existenz und Eindeutigkeit durch Induktion nach dem Grad von g.

Falls $\operatorname{grad}(g) = 0$, also $g = 1$, ist F ein Polynom, und es ist nichts weiter zu zeigen. Die Behauptung sei also wahr für alle rationalen Funktionen mit einem Nenner vom Grad $< \operatorname{grad}(g)$.

Nach dem Lemma gibt es ein eindeutig bestimmtes $b =: a_{1n_1}$ derart, dass

$$F(X) = \frac{b}{(X - z_1)^{n_1}} + \frac{\tilde{f}(X)}{(X - z_1)^{n-1} \prod_{i \neq 1} (X - z_i)^{n_i}}.$$

Subtrahieren wir $\frac{b}{(X-z_1)^{n_1}}$ von $F(X)$, so hat der Rest einen Nenner kleineren Grades als g. Durch Induktion folgt die Behauptung. □

Lemma 9.48 *— Es sei $g \in \mathbb{R}[X]$ ein nicht konstantes Polynom, $z \in \mathbb{C} \setminus \mathbb{R}$ eine komplexe, nicht relle Nullstelle von g der Vielfachheit n und $(X - z)(X - \overline{z}) = X^2 + Xu + v$). Ferner sei $h \in \mathbb{R}[X]$ so, dass $g(X) = (X^2 + Xu + v)^n h(X)$. Zu jedem $f \in \mathbb{R}[X]$ gibt es eindeutig bestimmte $b, c \in \mathbb{R}$ und $\tilde{f} \in \mathbb{R}[X]$ derart, dass*

$$\frac{f(X)}{g(X)} = \frac{b + cx}{(X^2 + Xu + v)^n} + \frac{\tilde{f}(X)}{(X^2 + Xu + v)^{n-1} h(X)}.$$

Beweis. Wenn es überhaupt b, c und $\tilde{f}$ mit den verlangten Eigenschaften gibt, so muss gelten

$$f(X) = (bX + c)h(X) + (X^2 + Xu + v)\tilde{f}(X).$$

Dadurch sind b und c als Lösungen des Gleichungssystems

$$\begin{aligned} f(z)/h(z) &= bz + c \\ f(\overline{z})/h(\overline{z}) &= b\overline{z} + c \end{aligned}$$

eindeutig festgelegt:

$$b = \frac{f(z)/h(z) - f(\overline{z})/h(\overline{z})}{z - \overline{z}}, \qquad c = \frac{\overline{z}f(z)/h(z) - zf(\overline{z})/h(\overline{z})}{\overline{z} - z}.$$

Definiert man umgekehrt b und c auf diese Weise, so sind b und c reelle Zahlen, und das Polynom $f(X) - (bX + c)h(X)$ hat Nullstellen bei z und bei $\overline{z}$, spaltet also den reellen Faktor $X^2 + Xu + v$ ab:

$$f(X) - (bX + c)h(X) = (X^2 + Xu + v)\tilde{f}(X).$$

□

Satz 9.49 *(Partialbruchzerlegung, reelle Fassung) — Es sei $F = f/g$ eine rationale Funktion mit $f, g \in \mathbb{R}[X]$, g normiert vom Grad 1. Es sei*

$$g(X) = \prod_{i=1}^{s} (X - x_i)^{n_i} \prod_{k=1}^{t} (X^2 + u_k X + v_k)^{m_k}$$

mit $n_i, m_j \in \mathbb{N}$, und $x_i, u_j, v_j \in \mathbb{R}$, $u_j^2 - 4v_j < 0$, die Zerlegung von g in irreduzible Faktoren in $\mathbb{R}[X]$. Dann gibt es eindeutig bestimmte reelle Zahlen a_{ij}, $i = 1, \ldots, s$, $j = 1, \ldots, n_i$ und $b_{k\ell}, c_{k\ell}$, $k = 1, \ldots, t$, $\ell = 1, \ldots, m_k$, und ein eindeutig bestimmtes Polynom $r \in \mathbb{R}[X]$ mit

$$F = r + \sum_{i=1}^{s} \sum_{j=1}^{n_i} \frac{a_{ij}}{(X - x_i)^j} + \sum_{k=1}^{t} \sum_{\ell=1}^{m_k} \frac{b_{k\ell} X + c_{k\ell}}{(X^2 + u_k X + v_k)^\ell}. \tag{9.12}$$

Beweis. Der Beweis verläuft wie im Beweis zu Satz 9.47: Für reelle Nullstellen verwendet man wieder Lemma 9.46, und komplexe Nullstellen behandelt man mit Lemma 9.48. □

Der Beweis liefert zugleich einen Algorithmus, wie man die Partialbruchzerlegung einer rationalen Funktion f/g wirklich durchführt. In jedem Falle braucht man dazu zunächst die Faktorzerlegung des Nenners. Ist diese gegeben, so kann man entweder dem Algorithmus des Beweises folgen, oder man macht den Ansatz nach Gleichung (9.12) mit Unbestimmten a_{ij}, $b_{k\ell}$, $c_{k\ell}$. Bringt man nun alles auf den Hauptnenner, so führt Koeffizientenvergleich auf ein lineares Gleichungssystem für die Unbestimmten, das man nach den üblichen Regeln der linearen Algebra löst.

Satz 9.50 — *Reelle rationale Funktionen F lassen sich geschlossen integrieren.*

Beweis. Das soll heißen, dass rationale Funktionen Stammfunktionen besitzen, die sich durch elementare Funktionen beschreiben lassen. Genauer geschieht das wie folgt: Wegen der Existenz der Partialbruchzerlegung genügt es, Integranden von folgendem Typ zu betrachten:

1. Polynome.

2. rationale Funktionen der Form

$$\frac{1}{(X-a)^n}, \quad a \in \mathbb{R}, n \in \mathbb{N}.$$

3. rationale Funktionen der Form

$$\frac{bX+c}{(X^2+Xu+v)^n}, \quad b,c,u,v \in \mathbb{R}, n \in \mathbb{N},$$

   mit irreduziblem Nenner.

Es ist klar, wie man Polynome integriert. Ferner gilt

$$\int \frac{dx}{(x-a)^n} = -\frac{1}{(n-1)(x-a)^{n-1}}, \quad \text{falls } n > 1$$

und

$$\int \frac{dx}{x-a} = \log|x-a|.$$

Interessant ist darum nur noch der Fall 3. Mit der linearen Substitution

$$x = \sqrt{v-u^2/4}\; t - u/2$$

bekommt man

$$x^2 + ux + v = (v - u^2/4) \cdot (t^2+1).$$

Diese Substitution führt und also auf die Betrachtung der folgenden Integrale zurück:

$$\int \frac{t}{(1+t^2)^n} dt \quad \text{und} \quad \int \frac{1}{1+t^2} dt$$

Für den ersten Term findet man leicht die Stammfunktion

$$\int \frac{tdt}{(1+t^2)^n} = -\frac{1}{2(n-1)(1+t^2)^{n-1}}, \quad \text{falls } n > 1$$

und

$$\int \frac{tdt}{1+t^2} = \frac{1}{2}\log(1+t^2) = \log\sqrt{1+t^2},$$

wie man durch Differenzieren verifiziert. Durch partielle Integration erhält man weiter

$$\begin{aligned} \int \frac{dt}{(1+t^2)^n} &= \frac{t}{(1+t^2)^n} + 2n\int \frac{t^2 dt}{(1+t^2)^{n+1}} \\ &= \frac{t}{(1+t^2)^n} + 2n\int \frac{dt}{(1+t^2)^n} - 2n\int \frac{dt}{(1+t^2)^{n+1}}. \end{aligned}$$

Daraus ergibt sich die Rekursionsgleichung

$$\int \frac{dt}{(1+t^2)^{n+1}} = \frac{1}{2n}\frac{t}{(1+t^2)^n} + \frac{2n-1}{2n}\int \frac{dt}{(1+t^2)^n}.$$

Die Rekursion endet bei dem Integral

$$\int \frac{dt}{1+t^2} = \arctan(t).$$

□

§10 Die Bernoullizahlen

10.1 Definition der Bernoullizahlen

Die Reihe

$$\frac{e^z-1}{z} = 1 + \frac{z}{2!} + \frac{z^2}{3!} + \dots \tag{10.1}$$

konvergiert überall in $\mathbb{C}$. Der Kehrwert dieser Funktion hat eine Reihenentwicklung

$$\frac{z}{e^z-1} = 1 - \frac{z}{2} + \frac{z^2}{6} + \dots = \sum_{k=0}^{\infty} B_k \frac{z^k}{k!} \tag{10.2}$$

mit einem positiven Konvergenzradius.

Definition 10.1 — Die Koeffizienten B_k heißen Bernoullizahlen[24]

Nach Definition der Bernoullizahlen gilt

$$\sum_{k=0}^{\infty} B_k \frac{z^k}{k!} \cdot \sum_{\ell} \frac{z^\ell}{(\ell+1)!} = 1. \tag{10.3}$$

Koeffizientenvergleich für das Monom z^m, $m \geq 1$, gibt uns die Bedingung

$$\sum_{k=0}^{m} \frac{1}{k!(m+1-k)!} B_k = 0. \tag{10.4}$$

Daraus lassen sich die Bernoullizahlen rekursiv gewinnen. Die Gleichung wird etwas schöner, wenn man sie mit $(m+1)!$ multipliziert. Man erhält so die erste Aussage des folgenden Lemmas:

Lemma 10.2 — *1. Es gilt $B_0 = 1$ und*

$$B_m = -\frac{1}{m+1} \sum_{k=0}^{m-1} \binom{m+1}{k} B_k \tag{10.5}$$

für alle $m \geq 1$. Insbesondere sind alle B_k rationale Zahlen.
2. Für alle ungeraden $k > 1$ ist $B_k = 0$.

Beweis. Die erste Aussage folgt aus der Gleichung (10.4). Die zweite Aussage ist eine Konsequenz der Tatsache, dass die Funktion

$$\sum_{k\geq 0, k\neq 1} B_k \frac{z^k}{k!} = \frac{z}{e^z-1} + \frac{z}{2} = \frac{z}{2} \cdot \frac{e^z+1}{e^z-1} = \frac{z}{2}\mathrm{ctgh}\frac{z}{2} \tag{10.6}$$

gerade ist, so dass also alle ungeraden Taylorkoeffizienten verschwinden. □

[24] Jacob Bernoulli, ∗ 27.12.1654 † 16.8.1705.

Bemerkung 10.3 — Für kleine k findet man die folgenden Werte von B_k:

k	0	1	2	3	4	5	6	7	8	9	10	11	12
B_k	1	$-\frac{1}{2}$	$\frac{1}{6}$	0	$-\frac{1}{30}$	0	$\frac{1}{42}$	0	$-\frac{1}{30}$	0	$\frac{5}{66}$	0	$-\frac{691}{2730}$

Satz 10.4 — *In einer Umgebung von $0 \in \mathbb{C}$ gelten die Entwicklungen*

$$\frac{z}{2}\mathrm{ctgh}\frac{z}{2} = \sum_{k=0}^{\infty} B_{2k}\frac{z^{2k}}{(2k)!}, \tag{10.7}$$

$$\frac{z}{2}\mathrm{ctg}\frac{z}{2} = \sum_{k=0}^{\infty} (-1)^k B_{2k}\frac{z^{2k}}{(2k)!}, \tag{10.8}$$

$$\tan(z) = \sum_{k=1}^{\infty} (-1)^{k-1} 4^k (4^k - 1) B_{2k}\frac{z^{2k-1}}{(2k)!} \tag{10.9}$$

Beweis. Die Aussage des ersten Satzes haben wir bereits im Beweis des Lemmas gesehen. Durch Einsetzen von iz für z geht (10.7) in (10.8) über. Weiter gilt die Verdopplungsformel

$$\tan(z) = \mathrm{ctg}(z) - 2\mathrm{ctg}(2z). \tag{10.10}$$

Setzt man jetzt die Potenzreihen aus (10.8) für den Kotangens ein, so folgt die Behauptung. □

Folgerung 10.5 — *Für alle $k \geq 1$ gilt*

$$(-1)^{k-1} 4^k (4^k - 1)\frac{B_{2k}}{2k} \in \mathbb{N}. \tag{10.11}$$

Beweis. Vergleicht man die rechte Seite von (10.9) mit der Taylorformel für den Tangens, so sieht man, dass der Ausdruck in Gleichung (10.11) genau $\tan^{(2k-1)}(0)$ ist. Wir müssen also zeigen, dass die ungeraden Ableitungen des Tangens, ausgewertet in 0, stets positive ganze Zahlen sind. Differenziert man den Tangens, so findet man nach der Kettenregel die folgenden Ausdrücke

$$\begin{aligned}
\tan'(x) &= 1 + \tan(x)^2 \\
\tan''(x) &= 2\tan(x)(1 + \tan(x)^2) = 2\tan(x) + 2\tan(x)^3 \\
\tan'''(x) &= (2 + 6\tan(x)^2)(1 + \tan(x)^2) = 2 + 8\tan(x)^2 + 6\tan(x)^4
\end{aligned}$$

usw. Der Mechanismus ist klar: Definiert man ganzzahlige Polynome mit positiven Koeffzienten rekursiv durch $P_0 := X$ und $P_{n+1} := P_n' \cdot (1 + X^2)$, so gilt:

1. $\tan^{(m)} = P_m(\tan)$ für alle $m \in \mathbb{N}_0$,
2. $\mathrm{grad}(P_m) = m + 1$, und schreibt man $P_m(x) = \sum_{i=0}^{m+1} a_{mi} x^i$, so folgt induktiv: $a_{mi} = 0$, falls $m + i$ gerade, und $a_{mi} \in \mathbb{N}$, falls $m + i$ ungerade.

Insbesondere ist also $\tan^{(2k-1)}(0) = P_{2k-1}(\tan(0) = a_{2k-1,0} \in \mathbb{N}$. $\square$

Satz 10.6 — *Für alle $z \in \mathbb{C} \setminus \pi\mathbb{Z}$ gilt*

$$\operatorname{ctg}(z) = \frac{1}{z} + \sum_{k=1}^{\infty} \frac{2z}{z^2 - k^2\pi^2}. \tag{10.12}$$

Wir zerlegen den Beweis des Satzes in mehrere Lemmata:

Lemma 10.7 — *Der Kotangens hat die folgenden Eigenschaften:*

1. $\operatorname{ctg} : \mathbb{C} \setminus \pi\mathbb{Z} \longrightarrow \mathbb{C}$ *ist eine ungerade, stetige Funktion,*
2. $\lim_{z\to 0}(\operatorname{ctg}(z) - \frac{1}{z}) = 0$.

Für alle $z \in \mathbb{C} \setminus \pi\mathbb{Z}$ gilt außerdem

3. $\operatorname{ctg}(z + \pi) = \operatorname{ctg}(z)$,
4. $\operatorname{ctg}(z) = \frac{1}{2}\operatorname{ctg}\frac{z}{2} + \frac{1}{2}\operatorname{ctg}\frac{z+\pi}{2}$.

Beweis. Als Quotient zweier stetiger Funktionen ist $\operatorname{ctg}(z) = \cos(z)/\sin(z)$ auf $\mathbb{C} \setminus \pi\mathbb{Z}$ stetig. Nach Gleichung (10.8) folgt

$$\operatorname{ctg}(z) - \frac{1}{z} = \sum_{k=1}^{\infty} \frac{B_{2k}}{(2k)!}(2z)^{2k-1}. \tag{10.13}$$

Die rechte Seite ist eine stetige Funktion. Setzt man $z = 0$, so erhält man die zweite Behauptung.

Kosinus und Sinus wechseln bei Verschiebung des Arguments um π das Vorzeichen, der Kotangens ist daher π-periodisch. Weiterhin ist cos eine gerade und sin eine ungerade Funktion, und somit ctg ungerade. Für $z \notin \pi\mathbb{Z}$ gilt schließlich

$$\begin{aligned} \operatorname{ctg}(z/2) + \operatorname{ctg}((z+\pi)/2) &= \operatorname{ctg}(z/2) - \tan(z/2) = \frac{\cos(z/2)}{\sin(z/2)} - \frac{\sin(z/2)}{\cos(z/2)} \\ &= \frac{\cos(z/2)^2 - \sin(z/2)^2}{\sin(z/2)\cos(z/2)} = 2\frac{\cos(z)}{\sin(z)} = 2\operatorname{ctg}(z). \end{aligned}$$

$\square$

Man sieht leicht, dass die Reihe $F(z) := \frac{1}{z} + \sum_{k=1}^{\infty} \frac{2z}{z^2 - k^2\pi^2}$ für jedes $z \notin \mathbb{C}\setminus\pi\mathbb{Z}$ absolut konvergiert, denn bis auf Vorfaktoren ist $\sum_{k=1}^{\infty} \frac{1}{k^2}$ eine Majorante.

Lemma 10.8 — *Die Funktion F hat die folgenden Eigenschaften:*

1. $F : \mathbb{C} \setminus \pi\mathbb{Z} \longrightarrow \mathbb{C}$ *ist eine ungerade, stetige Funktion,*

2. $\lim_{z\to 0}(F(z)-\frac{1}{z})=0$.

Für alle $z\in\mathbb{C}\setminus\pi\mathbb{Z}$ gilt außerdem

3. $F(z+\pi)=F(z)$,

4. $F(z)=\frac{1}{2}F(z/2)+\frac{1}{2}F((z+\pi)/2)$.

Beweis. Für $n\in\mathbb{N}_0$ sei

$$S_n(z):=\frac{1}{z}+\sum_{k=1}^{n}\frac{2z}{z^2-k^2\pi^2}=\sum_{k=-n}^{n}\frac{1}{z-k\pi}. \tag{10.14}$$

Es sei nun $R\in\mathbb{R}_{>0}$. Dann gelten für $z\in\mathbb{C}$ und $k,n\in\mathbb{N}$ mit $|z|<R\le n\le k$ die Abschätzungen $|z|^2<k^2$ und $|z^2-k^2\pi^2|\ge k^2(\pi^2-1)$. Daraus folgt:

$$|F(z)-S_{n-1}(z)|=\left|\sum_{k=n}^{\infty}\frac{2z}{z^2-k^2\pi^2}\right|\le\frac{2|z|}{\pi^2-1}\sum_{k=n}^{\infty}\frac{1}{k^2}.$$

Das zeigt, dass die auf $\mathbb{C}\setminus\pi\mathbb{Z}$ stetigen Funktionen $S_n(z)$ auf jeder Teilmenge $U_R(z)\setminus\pi\mathbb{Z}$ gleichmäßig gegen F konvergieren. Folglich ist F auf $\mathbb{C}\setminus\pi\mathbb{Z}$ stetig. Man sieht sofort aus der Definition, dass F ungerade ist.

Ferner gilt

$$|F(z)-\frac{1}{z}|=|F(z)-S_0(z)|\le\frac{2|z|}{\pi^2-1}\sum_{k=1}^{\infty}\frac{1}{k^2}.$$

Für $z\to 0$ geht die rechte Seite gegen 0.

Für $n\to\infty$ geht $S_n(z)-S_n(z+\pi)$ bei festem z gegen $F(z)-F(z+\pi)$. Andererseits gilt

$$\begin{aligned}S_n(z)-S_n(z+\pi) &= \sum_{k=-n}^{n}\frac{1}{z-k\pi}-\sum_{k=-n}^{n}\frac{1}{z+\pi-k\pi}\\ &= \frac{1}{z-n\pi}-\frac{1}{z+(n+1)\pi}\\ &= \frac{(2n+1)\pi}{(z^2+\pi z)-(n^2+n)\pi^2)}\xrightarrow{n\to\infty}0.\end{aligned}$$

Daraus folgt $F(z+\pi)-F(z)=0$.

Ähnlich schließen wir für die letzte Aussage: Die Differenz

$$S_{2n}(z)-\frac{1}{2}S_n(z/2)-\frac{1}{2}S_n((z+\pi)/2) \tag{10.15}$$

geht bei festem z für $n\to\infty$ gegen $F(z)-F(z/2)/2-F((z+\pi)/2)/2$. Andererseits ist der Ausdruck (10.15) gleich

$$\sum_{k=-2n}^{2n}\frac{1}{z-k\pi}-\sum_{k=-n}^{n}\frac{1}{z-2k\pi}-\sum_{k=-n}^{n}\frac{1}{z-(2k-1)\pi}=-\frac{1}{z+(2n+1)\pi},$$

und geht für $n \to \infty$ gegen 0. $\square$

Wir definieren nun $\Phi(z) : \mathbb{C} \to \mathbb{C}$ durch

$$\Phi(z) := \begin{cases} \operatorname{ctg}(z) - F(z) & \text{falls } z \in \mathbb{C} \setminus \pi\mathbb{Z}. \\ 0 & \text{sonst.} \end{cases}$$

Lemma 10.9 — *Φ hat die folgenden Eigenschaften:*

1. *Φ ist ungerade und stetig, und $\Phi(0) = 0$.*
2. *$\Phi(z) = \Phi(z + \pi)$ für alle $z \in \mathbb{C}$.*
3. *$\Phi(z) = \frac{1}{2}\Phi(z/2) + \frac{1}{2}\Phi((z + \pi)/2)$ für alle $z \in \mathbb{C}$.*

Die einzige Funktion $\Phi : \mathbb{C} \to \mathbb{C}$ mit den Eigenschaften 1.-3. ist die Nullfunktion.

Mit diesem Lemma ist offenbar auch Satz 10.6 bewiesen.

Beweis. Die Funktion Φ ist π-periodisch, d.h. $\Phi(z + \pi) = \Phi(z)$: Für $z \notin \pi\mathbb{Z}$ folgt das aus der analogen Aussage für ctg und F, und für $z \in \pi\mathbb{Z}$ ist es nach Definition klar.

Φ ist stetig auf ganz $\mathbb{C}$: Für $z \in \mathbb{C} \setminus \pi\mathbb{Z}$ folgt dies aus der Stetigkeit für ctg und F. Die Behauptung muss also noch für alle $z \in \pi\mathbb{Z}$ gezeigt werden. Wegen der π-Periodizität kann man sich ferner auf den Fall $z = 0$ zurückziehen. Da nach Definition $\Phi(z) = 0$, muss gezeigt werden, dass $\lim_{z\to 0} \Phi(z) = 0$. Nun gilt nach den obigen Lemmata

$$\lim_{z\to 0} \Phi(z) = \lim_{z\to 0} \left(\operatorname{ctg}(z) - \frac{1}{z} \right) - \lim_{z\to 0} \left(F(z) - \frac{1}{z} \right) = 0.$$

Die Verdopplungsformel $\Phi(z) = \frac{1}{2}\Phi(z/2) + \frac{1}{2}\Phi((z + \pi)/2)$ ist für $z \in \mathbb{C} \setminus \pi\mathbb{Z}$ wahr, weil sie für ctg und F gilt. Für $z \in \pi\mathbb{Z}$ kann man sich wegen π-Periodizität auf den Fall $z = 0$ beschränken. Zu zeigen ist dann: $0 = \Phi(0) = \frac{1}{2}\Phi(0) + \frac{1}{2}\Phi(\pi/2)$, also $\Phi(\pi/2) = 0$. Nun ist Φ ungerade und π-periodisch, woraus

$$\Phi(\pi/2) = -\Phi(-\pi/2) = -\Phi(\pi/2)$$

und somit $\Phi(\pi/2) = 0$ folgt.

Es bleibt zu zeigen, dass Φ identisch verschwindet. Angenommen, Φ ist nicht die Nullfunktion, etwa $\Phi(x) \neq 0$ für ein $x \in \mathbb{C}$. Wir wählen $R \geq \max\{|x|, \pi\}$ und setzen $K_R := \{z \in \mathbb{C} \,|\, |z| \leq R\}$. Dann ist $|\Phi|$ auf der kompakten Kreisscheibe K_R nicht Null. Wegen der Kompaktheit von K_R gibt es ein $a \in K_R$ mit $|\Phi(a)| = \sup |\Phi(K_R)| =: M \neq 0$. Die Kreisscheibe ist aber auch konvex. Mit a liegen also auch die Punkte $a/2$ und $(a + \pi)/2$ in K_R. Aus der Verdopplungsformel folgt:

$$M = |\Phi(a)| \leq \frac{1}{2}|\Phi(a/2)| + \frac{1}{2}|\Phi((a + \pi)/2)| \leq \frac{1}{2}M + \frac{1}{2}M = M.$$

Daher muss überall Gleichheit gelten, insbesondere haben wir $|\Phi(a)| = |\Phi(a/2)|$. Das Argument lässt sich mit $a/2$ statt a wiederholen. Induktiv folgt $|\Phi(a/2^n)| = M$. Im Grenzübergang erhalten wir wegen der Stetigkeit von Φ den Widerspruch

$$M = \lim_{n\to\infty} |\Phi(a/2^n)| = |\Phi(\lim_{n\to\infty} a/2^n)| = |\Phi(0)| = 0.$$

Das zeigt $\Phi(z) = 0$ für alle $z \in \mathbb{C}$. □

Jetzt können wir ernten. Ich erinnere daran, dass die Riemannsche Zetafunktion für $s > 1$ wie folgt definiert ist:

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}. \tag{10.16}$$

Satz 10.10 — *Für alle $\ell \in \mathbb{N}$ gilt*

$$\zeta(2\ell) = \sum_{k=1}^{\infty} \frac{1}{k^{2\ell}} = (-1)^{\ell-1} \frac{2^{2\ell-1}\pi^{2\ell}}{(2\ell)!} B_{2\ell}. \tag{10.17}$$

Speziell ergeben sich die Formeln

$$\sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{6}, \quad \sum_{k=1}^{\infty} \frac{1}{k^4} = \frac{\pi^4}{90}, \quad \sum_{k=1}^{\infty} \frac{1}{k^6} = \frac{\pi^6}{945}. \tag{10.18}$$

Beweis. Kombiniert man die Aussage von Satz 10.6 mit der Potenzreihenentwicklung des Kotangens, so folgt:

$$\begin{aligned} \sum_{\ell=0}^{\infty} (-1)^\ell B_{2\ell} \frac{(2z)^{2\ell}}{(2\ell)!} &= z\,\mathrm{ctg}(z) = 1 + \sum_{k=1}^{\infty} \frac{2z^2}{z^2 - k^2\pi^2} \\ &= 1 - 2\sum_{k=1}^{\infty}\sum_{n=1}^{\infty} \Big(\frac{z}{k\pi}\Big)^{2n} \end{aligned}$$

Die Doppelsumme konvergiert absolut für alle z mit $|z| < \pi$ und darf umgeordnet werden:

$$\ldots = 1 - 2\sum_{n=1}^{\infty} \frac{z^{2n}}{\pi^{2n}} \sum_{k=1}^{\infty} \frac{1}{k^{2n}}.$$

Koeffizientenvergleich liefert also die Beziehung

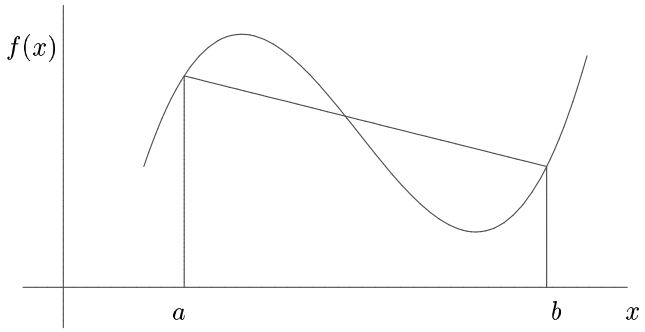
$$-\frac{2}{\pi^{2n}} \sum_{k=1}^{\infty} \frac{1}{k^{2n}} = (-1)^n B_{2n} \frac{2^{2n}}{(2n)!}.$$

Das war zu zeigen. □

10.A Die Trapezregel

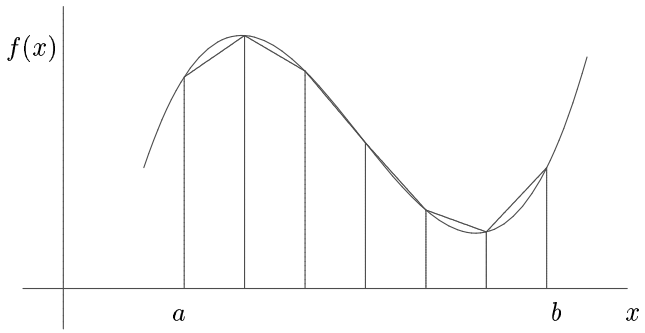
Es sei $f : [a, b] \to \mathbb{R}$ eine stetige Funktion. Eine sehr grobe Näherung an das Integral $\int_a^b f(x)dx$ ist durch die Fläche des Trapezes gegeben, das durch die Punkte $(a, 0)$, $(a, f(a))$, $(b, f(b))$ und $(b, 0)$ abgesteckt wird. Der Flächeninhalt ist

$$F = (b - a)\frac{(f(a) + f(b)}{2}.$$



Die Näherung wird etwas besser, wenn man das Intervall weiter unterteilt. Es sei $N \in \mathbb{N}$ und $h := (b-a)/N$. Betrachtet man die Trapeze über allen Teilintervallen $[t_{i-1}, t_i]$ mit $t_i = a + ih$, $i = 0, \ldots, N$, so erhält man als Gesamtfläche

$$F = h(\frac{f(a)}{2} + \sum_{i=1}^{N-1} f(t_i) + \frac{f(b)}{2}). \tag{10.19}$$



Die Frage, wie gut die Trapezsumme F das Integral $\int_a^b f(x)dx$ approximiert, lässt sich quantitativ beantworten, wenn f zweimal stetig differenzierbar ist.

Satz 10.11 — *Es sei $f \in \mathcal{C}^2([a, b])$. Ferner sei $N \in \mathbb{N}$, $h = (b - a)/N$ und $t_i = a + ih$ für $i = 0, \ldots, N$. Dann gibt es ein $\xi \in (a, b)$ mit*

$$\int_a^b f(x)dx = h\left[\frac{f(a)}{2} + \sum_{i=1}^{N-1} f(t_i) + \frac{f(b)}{2}\right] - (b - a)\frac{h^2}{12}f''(\xi).$$

Beweis. 1. Schritt: Es sei $a = 0$, $b = 1$, $N = 1$. Durch zweimalige partielle Integration folgt:

$$\begin{aligned}\int_0^1 f(x)dx &= (x-\frac{1}{2})f(x)\Big|_0^1 - \int_0^1 (x-\frac{1}{2})f'(x)dx \\ &= \frac{f(1)+f(0)}{2} - \left[\frac{x^2-x}{2}f'(x)\right]\Big|_0^1 + \int_0^1 \frac{x^2-x}{2}f''(x)dx\end{aligned}$$

Der Term $\left[\frac{x^2-x}{2}f'(x)\right]\Big|_0^1$ verschwindet. Die Funktion x^2-x hat auf $(0,1)$ keine Nullstellen. Nach dem verallgemeinerten Mittelwertsatz der Integralrechnung 9.26 gibt es ein $\xi \in (0,1)$ mit

$$\int_0^1 \frac{x^2-x}{2}f''(x)dx = f''(\xi)\int_0^1 \frac{x^2-x}{2}dx = -\frac{f''(\xi)}{12}.$$

Das bedeutet:

$$\int_0^1 f(x)dx = \frac{f(1)+f(0)}{2} - \frac{f''(\xi)}{12}.$$

2. Schritt: $a < b$ beliebig, aber $N = 1$. Mit der Substitutionsregel und der Substitution $x = a + (b-a)t$, $t \in [0,1]$, folgt aus dem Ergebnis des 1. Schritts:

$$\int_a^b f(x)dx = (b-a)\frac{f(a)+f(b)}{2} - \frac{(b-a)^3}{12}f''(\xi). \tag{10.20}$$

3. Schritt: $a < b$ beliebig, $N \in \mathbb{N}$. Wir wenden (10.20) auf jedes Teilintervall $[t_{i-1}, t_i]$ an. Durch Aufsummieren folgt:

$$\begin{aligned}\int_a^b f(x)dx &= h\left[\frac{f(a)}{2} + \sum_{i=1}^{N-1} f(t_i) + \frac{f(b)}{2}\right] - \frac{h^3}{12}\sum_{i=1}^{N} f''(\xi_i) \\ &= h\left[\frac{f(a)}{2} + \sum_{i=1}^{N-1} f(t_i) + \frac{f(b)}{2}\right] - \frac{h^3}{12}Nf''(\xi),\end{aligned}$$

wobei die $\xi_i \in (t_{i-1}, t_i)$ geeignete Zwischenpunkte für jedes $i = 1, \ldots, N$ sind und ξ mit Hilfe des Zwischenwertsatzes so gewählt ist, dass $f(\xi_1)+\ldots+f(\xi_N) = N\,f(\xi)$. □

10.B Die Summenformel von Euler-Maclaurin

Die Methode, wie wir die Trapezregel mit partieller Integration hergeleitet haben, lässt sich noch weiter ausbeuten. Dies führt auf die Summenformel von Euler und Maclaurin.

Satz 10.12 — *Es sei $m \geq 1$ und $g \in \mathcal{C}^{2m}([0,1])$. Dann gibt es ein $\xi \in (0,1)$ so, dass*

$$\begin{aligned}\int_0^1 g(x)dx &= \frac{g(0)+g(1)}{2} - \sum_{k=1}^{m-1} \frac{B_{2k}}{(2k)!}\Big(g^{(2k-1)}(1) - g^{(2k-1)}(0)\Big) \\ &\quad - \frac{B_{2m}}{(2m)!} g^{(2m)}(\xi).\end{aligned}$$

Wir brauchen ein paar Vorbereitungen:

Definition 10.13 — Die Polynome

$$B_m(x) := \sum_{k=0}^{m} \binom{m}{k} B_k x^{m-k}, \quad m \in \mathbb{N}_0 \tag{10.21}$$

heißen Bernoullipolynome.

Die ersten Bernoullipolynome lauten also

$$B_0(x) = 1, \quad B_1(x) = x - \frac{1}{2}, \quad B_2(x) = x^2 - x + \frac{1}{6}$$

$$B_3(x) = x^3 - \frac{3}{2}x^2 + \frac{1}{2}x, \quad B_4(x) = x^4 - 2x^3 + x^2 - \frac{1}{30}.$$

Lemma 10.14 — *Die Bernoullipolynome haben die folgenden Eigenschaften:*

1. $B_0 = 1$,
2. $\frac{1}{m} B_m(x)$ *ist eine Stammfunktion von* B_{m-1}.
3. $\int_0^1 B_m(x)dx = 0$ *für alle* $m \geq 1$.

Die Bernoullipolynome sind durch diese Eigenschaften eindeutig gekennzeichnet.

Beweis. Unmittelbar aus der Definition folgt:

$$\begin{aligned}B_m(x)' &= \sum_{k=0}^{m-1} B_k \binom{m}{k} (m-k) x^{m-1-k} \\ &= m \sum_{k=0}^{m-1} B_k \binom{m-1}{k} x^{(m-1)-k} \\ &= m B_{m-1}(x).\end{aligned}$$

D. h. $B_m(x)/m$ ist eine Stammfunktion von B_{m-1}. Deshalb ist die Aussage 3. äquivalent zu der Aussage

$$B_m(0) = B_m(1)$$

für $m \geq 2$. Und das sieht man so:

$$B_m(1) = \sum_{k=0}^{m} B_k \binom{m}{k} = B_m + \sum_{k=0}^{m-1} B_k \binom{m}{k} = B_m = B_m(0).$$

Beim vorletzten Schritt wurde die Rekursionsgleichung (10.4) für $m-1$ statt m benutzt. Die Eigenschaften 1.-3. charakterisieren die Bernoullipolynome in der Tat eindeutig, denn sind $B_0, \ldots, B_{m-1}$ schon gegeben, so ist B_m als Stammfunktion von B_{m-1} eindeutig bestimmt bis auf eine Integrationskonstante, und diese wird durch die Normierungsbedingung 3. festgelegt. Alle B_m sind also durch B_0 und die Forderungen 1. und 2. bestimmt. □

Lemma 10.15 — *Die Bernoullipolynome haben die folgenden Eigenschaften:*

1. *Es gilt $B_m(1-x) = (-1)^m B_m(x)$ für alle $x \in \mathbb{R}$ und alle $m \geq 0$. Insbesondere gilt $B_m(1) = B_m(0) = B_m$ für alle $m \geq 2$.*

2. *Für alle ungeraden m verschwindet $B_m(x)$ auf dem Intervall $(0,1)$ nur an der Stelle $x = 1/2$.*

3. *Für alle geraden $m \geq 2$ gilt $B_m(x) \neq B_m$ für alle $x \in (0,1)$.*

Beweis. Zu 1.: Die Polynome $(-1)^m B_m(1-x)$, $m \in \mathbb{N}_0$, erfüllen die Bedingungen von Lemma 10.14 ebenso wie die $B_m(x)$ selbst. Da die Bernoullipolynome durch diese Bedingungen eindeutig charakterisiert sind, folgt $B_m(x) = (-1)^m B_m(1-x)$. Die Identität $B_m(1) = B_m(0) = B_m$ haben wir schon im Beweis des vorigen Lemmas eingesehen.
Zu 2.: Die Behauptung ist offensichtlich wahr für B_1. Wir wissen weiter, dass für alle ungeraden $m > 1$ gilt $B_m(0) = 0 = B_m(1)$. Wegen der Symmetrie $B_m(1-x) = -B_m(x)$ muss auch $B_m(1/2) = 0$ sein. Wir zeigen durch Induktion über alle ungeraden m, dass B_m in $(0,1)$ keine weiteren Nullstellen haben kann. Angenommen, dies ist für $m-2$ bereits gezeigt. Wäre nun $\xi \in (0,1/2)$ eine weitere Nullstelle — den Fall $\xi \in (1/2,1)$ behandelt man analog —, so gäbe es nach dem Satz von Rolle ein ξ' strikt zwischen 0 und ξ und ein ξ'' strikt zwischen ξ und $1/2$ derart, dass $B'_m(\xi') = B'_m(\xi'') = 0$. Erneute Anwendung des Satzes von Rolle gibt uns ein η strikt zwischen ξ' und ξ'' mit $0 = B''_m(\eta) = m(m-1)B_{m-2}(\eta)$. Aber dies widerspricht der Annahme, dass B_{m-2} keine Nullstellen zwischen 0 und $1/2$ hat.
Zu 3.: Angenommen, es gäbe ein $x \in (0,1)$ mit $B_m(x) = B_m$. Nach dem Satz von Rolle gäbe es dann zwei verschiedene Nullstellen ζ' und ζ'' von $B'_m(x) =$

$mB_{m-1}(x)$ im Intervall $(0,1)$. Mindestens eine davon ist ungleich $1/2$ und widerspricht der Teilaussage 2. des Lemmas, wonach B_{m-1} auf $(0,1)$ nur in $1/2$ verschwindet. □

Es sei nun $g \in \mathcal{C}^{2m}([0,1])$ für ein $m \geq 1$. Durch wiederholte partielle Integration erhält man die folgenden Identitäten:

$$\begin{aligned}
\int_0^1 g(x)dx &= B_1(x)g(x)\Big|_0^1 - \int_0^1 B_1(x)g'(x)dx \\
&= (B_1(x)g(x) - \frac{1}{2}B_2(x)g'(x))\Big|_0^1 + \frac{1}{2}\int_0^1 B_2(x)g''(x)dx \\
&\vdots \\
&= \sum_{k=1}^{2m-1}(-1)^{k-1}\frac{B_k(x)}{k!}g^{(k-1)}(x)\Big|_0^1 - \int_0^1 \frac{B_{2m-1}(x)}{(2m-1)!}g^{(2m-1)}(x)dx \\
&= \frac{g(0)+g(1)}{2} - \sum_{\ell=1}^{m-1}\frac{B_{2\ell}}{(2\ell)!}(g^{(2\ell)}(1) - g^{(2\ell)}(0)) \\
&\quad - \int_0^1 \frac{B_{2m-1}(x)}{(2m-1)!}g^{(2m-1)}(x)dx
\end{aligned}$$

Für den Beweis von Satz 10.12 genügt es deshalb zu zeigen:

Lemma 10.16 — *Es sei $m \geq 1$ und $g \in \mathcal{C}^{2m}([0,1])$. Dann gibt es ein $\xi \in (0,1)$ mit*

$$\int_0^1 \frac{B_{2m-1}(x)}{(2m-1)!}g^{(2m-1)}(x)dx = \frac{B_{2m}}{(2m)!}g^{(2m)}(\xi).$$

Beweis. Erneute partielle Integration liefert

$$\begin{aligned}
\int_0^1 \frac{B_{2m-1}(x)}{(2m-1)!}g^{(2m-1)}(x)dx &= \frac{B_{2m}(x) - B_{2m}}{(2m)!}g^{(2m-1)}(x)\Big|_0^1 \\
&\quad - \int_0^1 \frac{B_{2m}(x) - B_{2m}}{(2m)!}g^{(2m)}(x)dx.
\end{aligned}$$

Der erste Summand auf der rechten Seite verschwindet, weil $B_{2m}(0) = B_{2m}(1) = B_{2m}$. Auf den zweiten Summanden wenden wir den verallgemeinerten Mittelwertsatz 9.26 der Integralrechnung an. Dies ist zulässig, weil nach Lemma 10.15 $B_m(x) \neq B_m$ für alle $x \in (0,1)$. Folglich existiert ein $\xi \in (0,1)$ mit

$$\int_0^1 \frac{B_{2m}(x) - B_{2m}}{(2m)!}g^{(2m)}(x)dx = g^{(2m)}(\xi)\int_0^1 \frac{B_{2m}(x) - B_{2m}}{(2m)!}dx,$$

und weil

$$\int_0^1 B_{2m}(x)dx = \frac{1}{2m+1}B_{2m+1}(x)\Big|_0^1 = 0,$$

folgt schließlich

$$\int_0^1 \frac{B_{2m}(x) - B_{2m}}{(2m)!} g^{(2m)}(x)dx = -g^{(2m)}(\xi)\frac{B_{2m}}{(2m)!}.$$

□

Satz 10.17 *(Euler Maclaurinsche[25] Summenformel) — Es sei $N \in \mathbb{N}$ und $g \in \mathcal{C}^{2m}([0,N])$. Dann gibt es für jedes $i \in \{1,\ldots,N\}$ ein $\xi_i \in (i-1,i)$ so, dass*

$$\begin{aligned}\frac{g(0)}{2} + \sum_{i=0}^{N-1} g(i) + \frac{g(N)}{2} &= \int_0^N g(x)dx + \frac{B_{2m}}{(2m)!}\sum_{i=1}^{N} g^{(2m)}(\xi_i) \\ &\quad + \sum_{k=1}^{m-1} \frac{B_{2k}}{(2k)!}\Big(g^{(2k-1)}(N) - g^{(2k-1)}(0)\Big).\end{aligned}$$

Beweis. Wir wenden Satz 10.12 nacheinander auf die Teilintervalle $[i-1,i]$, $i = 1,\ldots,N$, an und summieren auf. □

Folgerung 10.18 *— Für alle $n, N \in \mathbb{N}$ gilt:*

$$\sum_{i=1}^{N-1} i^n = \frac{1}{n+1}(B_{n+1}(N) - B_{n+1}). \tag{10.22}$$

Das verallgemeinert unsere früher durch Induktion bewiesenen Formeln

$$\sum_{i=1}^{N-1} i = \frac{N^2}{2} - \frac{N}{2}, \quad \sum_{i=1}^{N-1} i^2 = \frac{N^3}{3} - \frac{N}{2} + \frac{N}{6}, \quad \sum_{i=1}^{N-1} i^3 = \frac{N^4}{4} - \frac{N^3}{2} + \frac{N^2}{4}.$$

Vermutlich hat Jacob Bernoulli die nach ihm benannten Zahlen gefunden, indem er die Formeln (10.22) für viele kleine n bestimmt und die dabei auftretenden Koeffizienten untersucht hat.

Beweis. Wir wenden die Euler-Maclaurinsche Summenformel auf die Funktion $g(x) = x^n$ an. Wählen wir $m > n$, so ist $g^{(2m)} = 0$, und die ξ_i enthaltenden Terme verschwinden. Außerdem lässt sich das Integral leicht auswerten:

$$\int_0^N g(x)dx = \int_0^N x^n dx = \frac{N^{n+1}}{n+1}.$$

Man erhält

$$\sum_{i=1}^{N-1} i^n + \frac{N^n}{2} = \frac{N^{n+1}}{n+1} + \sum_{k=1}^{n+1} \frac{B_{2k}}{n+1}\binom{n+1}{2k} N^{n-(2k-1)}$$

[25] Colin Maclaurin, ∗ Februar 1698 † 14.6.1746.

Daraus folgt nach Definition der Bernoullipolynome, dass

$$\sum_{i=1}^{N-1} i^n = \frac{1}{n+1}(B_{n+1}(N) - B_{n+1}).$$

□

Satz 10.19 *(Stirlingformel[26]) — Es gilt*

$$\lim_{n\to\infty} \frac{n!}{\sqrt{2\pi n}\left(\frac{n}{e}\right)^n} = 1. \tag{10.23}$$

Genauer existiert für jedes $m \geq 1$ eine Zahl $t \in (0,1)$ so, dass

$$n! = \sqrt{2\pi n}\left(\frac{n}{e}\right)^n \exp\left(\sum_{k=1}^{m-1} \frac{B_{2k}}{2k(2k-1)n^{2k-1}} + \frac{tB_{2m}}{(2m)^2(n-1)^{2m-1}}\right) \tag{10.24}$$

Beweis. Zunächst sieht man, dass die Formel (10.23) aus der Formel (10.24) folgt. Bei festem m geht das Argument der Exponentialfunktion in (10.24) gegen 0, der ganze Faktor also gegen 1.

Wir zeigen jetzt (10.24). Dazu seien $m, n \in \mathbb{N}$ und $n \geq 2$. Wir wollen die Euler-Maclaurinsche Summenformel Satz 10.17 auf die Funktion $g(x) = \log(x+1)$ und $N = n-1$ anwenden. Das Integral lässt sich leicht ausrechnen:

$$\begin{aligned}\int_0^{n-1} \log(x+1)dx &= [(x+1)\log(x+1) - x]\Big|_0^{n-1} \\ &= n\log(n) - n + 1 = n\log\frac{n}{e} + 1.\end{aligned}$$

Satz 10.17 sagt: Es gibt Punkte $\xi_i \in (i-1, i)$ für $i = 1, \ldots, n-1$ so, dass

$$\begin{aligned}\sum_{i=2}^{n-1} \log(i) + \frac{1}{2}\log(n) &= n\log\frac{n}{e} + 1 \\ &\quad - \frac{B_{2m}}{(2m)!}\sum_{i=1}^{n-1} \frac{(2m-1)!}{(1+\xi_i)^{2m}} \\ &\quad + \sum_{k=1}^{m-1} \frac{B_{2k}}{(2k)!}\left[\frac{(2k-2)!}{n^{2k-1}} - \frac{(2k-2)!}{1^{2k-1}}\right]\end{aligned}$$

[26] James Stirling, ∗ Mai 1692 † 5.12.1770.

Dabei beachte man, dass die Zwischenpunkte ξ_i nur von i und m, nicht aber von n abhängen. Diese Gleichung können wir wie folgt umschreiben:

$$\begin{aligned}\log\frac{n!}{\sqrt{n}\left(\frac{n}{e}\right)^n} &= \left[1-\sum_{k=1}^{m-1}\frac{B_{2k}}{2k(2k-1)}-\frac{B_{2m}}{2m}\sum_{i=1}^{\infty}\frac{1}{(1+\xi_i)^{2m}}\right]\\ &\quad+\sum_{k=1}^{m-1}\frac{B_{2k}}{2k(2k-1)}\frac{1}{n^{2k-1}}\\ &\quad+\frac{B_{2m}}{2m}\sum_{i=n}^{\infty}\frac{1}{(1+\xi_i)^{2m}}\end{aligned}\tag{10.25}$$

Wir bezeichnen die drei Summanden auf der rechten Seite mit A, B und C. Nun gilt:

1. A ist unabhängig von n.
2. Bei festem m ist B eine endliche Summe und geht für $n\to\infty$ gegen 0.
3. C lässt sich wie folgt abschätzen:

$$\begin{aligned}0\le\frac{C}{B_{2m}} &= \frac{1}{2m}\sum_{i=n}^{\infty}\frac{1}{(1+\xi_i)^{2m}}\le\frac{1}{2m}\sum_{i=n}^{\infty}\frac{1}{i^{2m}}\\ &< \frac{1}{2m}\int_{n-1}^{\infty}\frac{du}{u^{2m}}\\ &= \frac{1}{(2m)^2(n-1)^{2m-1}}\end{aligned}\tag{10.26}$$

   Insbesondere gilt $C\to 0$ für $n\to\infty$.

Es folgt: Der Grenzwert

$$\lim_{n\to\infty}\log\frac{n!}{\sqrt{n}\left(\frac{n}{e}\right)^n}$$

existiert und ist gleich A. Aber das zeigt, dass A auch von m nicht abhängt. Zur Bestimmung von A greifen wir auf die Formel (9.9)

$$\lim_{n\to\infty}\frac{4^n}{\sqrt{\pi n}\binom{2n}{n}}=1$$

zurück. Wegen

$$\binom{2n}{n}=\frac{(2n)!}{n!\cdot n!}$$

folgt nun

$$\begin{aligned}\exp(A) &= \frac{\exp(A)\exp(A)}{\exp(A)}\\ &= \lim_{n\to\infty}\left(\frac{n!}{\sqrt{n}(n/e)^n}\right)^2\left(\frac{(2n)!}{\sqrt{2n}(2n/e)^{2n}}\right)^{-1}\\ &= \lim_{n\to\infty}\frac{4^n\sqrt{\frac{2}{n}}}{\binom{2n}{n}}=\sqrt{2\pi}\end{aligned}$$

Setzen wir $A = \log(\sqrt{2\pi})$ in der obigen Rechnung (10.25) ein, so bleibt:

$$\log \frac{n!}{\sqrt{2\pi n}(n/e)^n} - \sum_{k=1}^{m-1} \frac{B_{2k}}{2k(2k-1)} \frac{1}{n^{2k-1}} = C,$$

und wir haben bereits oben gesehen (10.26), dass

$$0 < \frac{C}{B_{2m}} < \frac{1}{(2)^2 m(n-1)^{2m-1}}.$$

Definiert man jetzt t als den Quotienten der beiden Ausdrücke dieser Ungleichung, so folgt die Behauptung. □

Bemerkung 10.20 — Die Reihe

$$\sum_{k=1}^{\infty} \frac{B_{2k}}{2k(2k-1)} \frac{1}{n^{2k-1}}$$

konvergiert für kein n! Man kann also nicht einfach in der Stirlingformel (10.24) m gegen unendlich gehen lassen.