Analysis I

Franz Merkl[1]
Universität München
(sehr vorläufige Version [2], 20. März 2017)

Inhaltsverzeichnis

[1]Der Autor dankt Herrn V. Braungardt, Herrn P. Eichinger, Herrn M. Hamilton, Herrn M. Mair und Frau E. Roth für die Hilfe beim Korrekturlesen sowie Frau G. Bach, Herrn P. Eichinger, Frau D. Mader und Herrn M. Mair für die Hilfe bei der technischen Herstellung des Skripts.

[2]Dies ist nur ein Entwurf eines Analysis I Skripts. Ohne jede Garantie. Für Hinweise auf Fehler aller Art ist der Autor dankbar.

0 Zeittafel

Die folgende Tabelle gibt einen unvollständigen historischen Überblick über die Entwicklung der Analysis:

17. Jahrhundert	*Newton, Leibniz*: "Infinitesimalrechnung". Rechnung mit unendlich kleinen, "infinitesimalen" Größen wie "dx" in Termen wie "$\frac{dy}{dx}$, $\int f(x)\,dx$".
18. Jahrhundert	*Euler*: Weiterentwicklung des Infinitesimalkalküls.
19. Jahrhundert	*Cauchy, Weierstraß*: Formale Präzisierung der Analysis *Riemann*: Präzisierung des Integralbegriffs *Cantor*: Mengenlehre
20. Jahrhundert	*Borel, Lebesgue*: maßtheoretische Version des Integrals *Robinson*: "Nonstandardanalysis": Formale Begründung des Infinitesimalkalküls mit modelltheoretischen Methoden.

1 Grundlagen

1.1 Logik

Mephisto:
Mein teurer Freund, ich rat' Euch drum
Zuerst Collegium Logicum.
Da wird der Geist Euch wohl dressiert,
In spanische Stiefeln eingeschnürt,
Daß er bedächtiger so fortan
Hinschleiche die Gedankenbahn,
Und nicht etwa, die Kreuz und Quer,
Irrlichteliere hin und her.

Goethe, Faust I

Mit der formalen Präzisierung der Differential- und Integralrechnung im 19. Jahrhundert, insbesondere durch Cauchy und Weierstraß, verschwanden "unendlich kleine", "infinitesimale" Größen aus der (standard) Analysis. Sie wurden durch "beliebig kleine" Größen ersetzt. Der subtile logische Unterschied zwischen "infinitesimal" und "beliebig klein" soll in diesem Abschnitt thematisiert werden.
Dabei streben wir keine systematische Abhandlung der logischen Grundlagen der Mathematik an; dies ist Spezialvorlesungen vorbehalten, z. B. über Logik oder Mengenlehre. Vielmehr wollen wir nur die logische Standardsprache der Mathematik soweit umreißen, wie wir sie zur Arbeit benötigen, ähnlich wie man eine Fremdsprache lernen kann, ohne alle grammatischen Regeln genau zu kennen.

1.1.1 Aussagenlogik

Aussagen können durch Verknüpfungen "und", oder "nicht", "impliziert", "ist äquivalent zu" verbunden werden. Diese Operationen werden auch "Junktoren" genannt. Der Wahrheitswert der Verknüpfung hängt nur vom Wahrheitswert "wahr" oder "falsch" der Argumente ab. Er wird durch folgende Tabellen definiert.

Einstelliger Junktor:

a	$\neg a$
w	f
f	w

Zweistellige Junktoren:

a	b	$a \wedge b$ a und b Konjunktion	$a \vee b$ a oder b Disjunktion	$a \Rightarrow b$ a impliziert b wenn a, dann b	$a \Leftrightarrow b$ a ist äquivalent zu b
w	w	w	w	w	w
w	f	f	w	f	f
f	w	f	w	w	f
f	f	f	f	w	w

Beispiel: "$1 + 1 = 3 \Rightarrow 2 > 3$" ist wahr, denn sowohl "$1 + 1 = 3$" als auch "$2 > 3$" sind falsch.
Die Implikation beschreibt nicht inhaltliche "Kausalität", etwa "a ist die Ursache für b", sondern nur formale Konstellationen von Wahrheitwerten: "Wenn a wahr ist, dann ist auch b wahr".

Nullstellige Junktoren:

Manchmal nimmt man auch die Aussagenkonstanten "$\top$" (die stets wahre Aussage) und "$\perp$" (die stets falsche Aussage, den Widerspruch) hinzu. Sie besitzen die Wahrheitswerte w bzw. f.

Konventionen zur Klammerersparnis:

- "$\neg$" bindet stärker als "$\wedge$"
- "$\wedge$" bindet stärker als "$\vee$"
- "$\vee$" bindet stärker als "$\Rightarrow$" und "$\Leftrightarrow$"
- "$\Rightarrow$" und "$\Leftrightarrow$" binden gleich stark.
- "bindet stärker" ist transitiv.

Beispiel: Die Formel

$$\neg a \vee \neg b \wedge c \Rightarrow d$$

bedeutet:

$$((\neg a) \vee ((\neg b) \wedge c)) \Rightarrow d.$$

Einige aussagenlogische Regeln:

1.

de Morgan'sche Regeln der Aussagenlogik
$\neg(\neg a \wedge \neg b)$ ist gleichwertig mit $a \vee b$.
$\neg(\neg a \vee \neg b)$ ist gleichwertig mit $a \wedge b$.

Wir begründen das mit je einer Wahrheitstabelle:

a	b	$\neg a$	$\neg b$	$\neg a \wedge \neg b$	$\neg(\neg a \wedge \neg b)$	$a \vee b$
w	w	f	f	f	w	w
w	f	f	w	f	w	w
f	w	w	f	f	w	w
f	f	w	w	w	f	f

a	b	$\neg a$	$\neg b$	$\neg a \vee \neg b$	$\neg(\neg a \vee \neg b)$	$a \wedge b$
w	w	f	f	f	w	w
w	f	f	w	w	f	f
f	w	w	f	w	f	f
f	f	w	w	w	f	f

Merkregel: "Erst Negation, dann Konjunktion" ist gleichwertig zu "erst Disjunktion, dann Negation", und umgekehrt.

2.

Aussagenlogische Distributivgesetze
$a \wedge (b \vee c)$ ist gleichwertig mit $(a \wedge b) \vee (a \wedge c)$.
$a \vee (b \wedge c)$ ist gleichwertig mit $(a \vee b) \wedge (a \vee c)$.

Wir begründen das durch je eine Wahrheitstabelle, die alle möglichen Kombinationen der Wahrheitswerte von a, b und c umfasst:

a	b	c	$b \vee c$	$a \wedge (b \vee c)$	$a \wedge b$	$a \wedge c$	$(a \wedge b) \vee (a \wedge c)$
w	w	w	w	w	w	w	w
w	w	f	w	w	w	f	w
w	f	w	w	w	f	w	w
w	f	f	f	f	f	f	f
f	w	w	w	f	f	f	f
f	w	f	w	f	f	f	f
f	f	w	w	f	f	f	f
f	f	f	f	f	f	f	f

a	b	c	$b \wedge c$	$a \vee (b \wedge c)$	$a \vee b$	$a \vee c$	$(a \vee b) \wedge (a \vee c)$
w	w	w	w	w	w	w	w
w	w	f	f	w	w	w	w
w	f	w	f	w	w	w	w
w	f	f	f	w	w	w	w
f	w	w	w	w	w	w	w
f	w	f	f	f	w	f	f
f	f	w	f	f	f	w	f
f	f	f	f	f	f	f	f

Man beachte, dass die beiden Distributivgesetze über die de Morgan'sche Regeln miteinander zusammenhängen (Vertauschen von "und" und "oder").

3.

Kontraposition
$a \Rightarrow b$ ist gleichwertig mit $\neg b \Rightarrow \neg a$

Begründung durch eine Wahrheitstabelle:

a	b	$\neg a$	$\neg b$	$\neg b \Rightarrow \neg a$	$a \Rightarrow b$
w	w	f	f	w	w
w	f	f	w	f	f
f	w	w	f	w	w
f	f	w	w	w	w

Die Kontraposition wird oft wie folgt in Beweisen verwendet: Um eine Behauptung b aus einer Prämisse a zu beweisen, nimmt man an, b sei falsch, und folgert daraus, dass dann auch a falsch sein muss.

Aussagenlogische Herleitungsregeln. Wir besprechen nun einige aussagenlogische Herleitungsregeln, wie sie in Beweisen öfter vorkommen:

1. **Beweis einer Konjunktion.** Um $A \wedge B$ zu zeigen, zeigt man einerseits A und andererseits B.

2. **Vorwärtsschliessen:** Sind A und $A \Rightarrow B$ gegeben, so können wir daraus B schließen. Das Vorwärtsschließen beruht auf der Tautologie

$$A \wedge (A \Rightarrow B) \Rightarrow B.$$

(Tautologie = aussagenlogisch allgemeingültige Formel)

3. **Beweis einer Implikation $A \Rightarrow B$:**

Um dies zu zeigen, nehmen wir an, dass A gilt. Unter (möglicher) Verwendung von A wird dann B gezeigt. Damit ist $A \Rightarrow B$ gezeigt.

4. **Beweis einer Negation $\neg A$:**

   Dies ist ein Spezialfall der vorhergehenden Strategie, wenn wir die Gleichwertigkeit von $\neg A$ zu $A \Rightarrow \perp$ verwenden, wobei $\perp$ für die Aussagenkonstante "stets falsche Aussage", synonym "Widerspruch", steht:

   Um $\neg A$ zu zeigen, nehmen wir an, dass A gilt. Unter (möglicher) Verwendung von A wird dann ein Widerspruch gezeigt. Damit ist $\neg A$ gezeigt.

5. **Indirekter Beweis:**

   Um A zu zeigen, nehmen wir an, dass $\neg A$ gilt. Unter (möglicher) Verwendung von $\neg A$ wird dann ein Widerspruch gezeigt. Damit ist A gezeigt. Dem indirekten Beweis liegt die Tautologie $(\neg A \Rightarrow \perp) \Rightarrow A$ zugrunde.

6. **Beweis durch Kontraposition:** Um $A \Rightarrow B$ zu zeigen, nehmen wir an, dass $\neg B$ gilt. Unter (möglicher) Verwendung davon wird dann $\neg A$ gezeigt. Damit ist $A \Rightarrow B$ gezeigt.

7. **Beweis durch Fallunterscheidung:** *Genaueres dazu und Varianten davon lernen Sie in den Übungen.* Um eine Aussage B zu zeigen, zeigt man zunächst für geeignete Aussagen $A_1, A_2, \ldots, A_n$ die Disjunktion $A_1 \vee A_2 \vee \ldots \vee A_n$. Dann:
   1. Fall: Wir nehmen A_1 an und zeigen damit B.
   2. Fall: Wir nehmen A_2 an und zeigen damit B.
   $\vdots$
   n. Fall: Wir nehmen A_n an und zeigen damit B.
   Damit ist B gezeigt.

   Dieser Herleitungsregel liegt folgendes Schema von Tautologien zugrunde:

   $$(A_1 \vee A_2 \vee \ldots \vee A_n) \wedge (A_1 \Rightarrow B) \wedge (A_2 \Rightarrow B) \wedge \ldots \wedge (A_n \Rightarrow B) \Rightarrow B$$

8. **Beweis einer Disjunktion – Strategie 1: Zurückführen auf eine Implikation.** Um $A \vee B$ zu zeigen, nehmen wir zunächst $\neg A$ an. Unter dieser Annahme zeigen wir B. Dann ist $A \vee B$ gezeigt. Dieser Herleitungsregel liegt zugrunde, dass $A \vee B$ und $\neg A \Rightarrow B$ den gleichen Wahrheitswert besitzen.

9. **Beweis einer Disjunktion – Strategie 2: Fallunterscheidung.** Um $A \vee B$ zu zeigen, machen wir eine Fallunterscheidung mit einer geeigneten Aussage C:
   1. Fall: Wir nehmen C an. Unter dieser Annahme zeigen wir A.
   2. Fall: Wir nehmen $\neg C$ an. Unter dieser Annahme zeigen wir B.
   Dann ist $A \vee B$ gezeigt.

   Dieser aussagenlogischen Herleitungsregel liegt die Tautologie

   $$(C \Rightarrow A) \wedge (\neg C \Rightarrow B) \Rightarrow A \vee B$$

   zugrunde.

10. **Beweis einer Äquivalenz:** Der Beweis einer Äquivalenz $A \Leftrightarrow B$ kann durch folgende beiden Beweisteile erfolgen:
    Beweisteil "$\Rightarrow$": Man nimmt die Aussage A an und zeigt unter dieser Annahme die Aussage B.
    Beweisteil "$\Leftarrow$": Man nimmt die Aussage B an und zeigt unter dieser Annahme die Aussage A.

    Dieser aussagenlogischen Herleitungsregel liegt zugrunde, dass $A \Leftrightarrow B$ den gleichen Wahrheitswert wie $(A \Rightarrow B) \wedge (B \Rightarrow A)$ besitzt.

1.1.2 Prädikatenlogik

Aussagen mit freien Variablen haben a priori keinen Wahrheitswert.

Beispiel: "$x > 2$" hat keinen Wahrheitswert, solange wir x nicht spezifizieren. Erst durch *Belegung* von x mit einem Wert oder durch *Binden* von x mit einem *Quantor* "$\forall, \exists$" wird die Formel wahr oder falsch.

Bedeutung der Quantoren:

- "$\forall x : \varphi(x)$" bedeutet: "Für alle x gilt $\varphi(x)$".
- "$\exists x : \varphi(x)$" bedeutet: "Es existiert ein x, für das $\varphi(x)$ gilt".

Üblicherweise, wenn sich der Bereich, den die Variable x durchlaufen darf, nicht von selbst versteht, spezifiziert man noch diesen Bereich.

Beispiel:

$$\forall x \in \mathbb{N}_0 \colon x > 2$$

bedeutet: "*Alle natürlichen Zahlen sind größer als 2*". (Dies ist natürlich falsch, denn 1 ist eine natürliche Zahl, die nicht größer als 2 ist.)

$$\exists x \in \mathbb{N}_0 \colon x > 2$$

bedeutet: "*Es gibt eine natürliche Zahl, die größer als 2 ist*".
(Diese Aussage ist wahr, denn 3 ist eine natürliche Zahl größer als 2.)

Einige Regeln für Quantoren:

- $\neg \forall x \colon \varphi(x)$ ist gleichwertig mit $\exists x \colon \neg \varphi(x)$.

  Wir geben zwei Begründungen hierfür an, eine "kurze", die ungefähr den Detaillierungsgrad in Lehrbuchbeweisen hat, und eine "lange", in der die zugrundeliegenden aussagenlogischen und prädikatenlogischen Schlüsse im Detail ausgearbeitet sind:

Begründung (Kurzform):
"$\Rightarrow$": Wenn $\forall x\colon \varphi(x)$ *nicht* gilt, dann muß es mindestens ein x geben, für das $\varphi(x)$ nicht gilt. Für dieses x gilt dann $\neg\varphi(x)$. Wir folgern daraus: $\exists x\colon \neg\varphi(x)$.

"$\Leftarrow$": Es gelte $\exists x\colon \neg\varphi(x)$. Wir können also ein x wählen, für das $\varphi(x)$ nicht gilt. Demnach kann $\varphi(x)$ nicht für alle x gelten, d.h. $\neg\forall x\colon \varphi(x)$ gilt.

□

Diesem kurzen Beweis stellen wir nun eine ausführliche Variante zur Seite, wie man sie nur selten in Lehrbüchern findet. In Ihren Beweisen brauchen Sie (normalerweise) keinen so hohen Detaillierungsgrad anwenden.

Begründung (Langform):
Das Beweisziel ist die Äquivalenz

$$[\neg\forall x\colon \varphi(x)] \;\Leftrightarrow\; \exists x\colon \neg\varphi(x). \tag{1}$$

Dies beweisen wir in zwei Beweisteilen:

Beweisteil "$\Rightarrow$": Das Beweisziel lautet hier

$$[\neg\forall x\colon \varphi(x)] \;\Rightarrow\; \exists x\colon \neg\varphi(x).$$

Mit Kontraposition können wir das auch gleichwertig in der Form

$$[\neg\exists x\colon \neg\varphi(x)] \;\Rightarrow\; \neg\neg\forall x\colon \varphi(x)$$

schreiben, also auch in der Form

$$[\neg\exists x\colon \neg\varphi(x)] \;\Rightarrow\; \forall x\colon \varphi(x).$$

Zum Beweis dieser Formel nehmen wir $\neg\exists x\colon \neg\varphi(x)$ an, anders gesagt

$$[\exists x\colon \neg\varphi(x)] \Rightarrow \bot. \tag{2}$$

Zu zeigen ist nun $\forall x\colon \varphi(x)$. Hierzu sei x gegeben; für dieses x ist nun $\varphi(x)$ zu zeigen. Das beweisen wir indirekt und nehmen hierzu $\neg\varphi(x)$ an. Aus dieser Annahme schließen wir $\exists x :\ \neg\varphi(x)$, was zusammen mit der Annahme (2) einen Widerspruch $\bot$ ergibt. Damit ist $\varphi(x)$, also die verbleibende Behauptung, indirekt gezeigt.

Beweisteil "$\Leftarrow$": Das Beweisziel lautet nun

$$[\exists x\colon \neg\varphi(x)] \;\Rightarrow\; \neg\forall x\colon \varphi(x).$$

Hierzu nehmen wir

$$\exists x\colon \neg\varphi(x) \tag{3}$$

an; nun ist $\neg\forall x\colon \varphi(x)$ zu zeigen, anders gesagt $[\forall x\colon \varphi(x)] \Rightarrow \perp$. Führen wir hierzu die Annahme

$$\forall x\colon \varphi(x) \tag{4}$$

wie folgt zu einem Widerspruch $\perp$: Aufgrund der Annahme (3) können wir ein x mit $\neg\varphi(x)$ nehmen, also anders geschrieben $\varphi(x) \Rightarrow \perp$. Andererseits liefert uns die Annahme (4) auch $\varphi(x)$ für dieses x, was zusammen den gewünschten Widerspruch $\perp$ liefert.

□

- $\neg\exists x\colon \psi(x)$ ist gleichwertig mit $\forall x\colon \neg\psi(x)$.

  Begründung: Statt dies direkt zu begründen, was analog zum vorhergehenden Beweis auch möglich wäre, wenden wir die eben bewiesene Äquivalenz (1) an, wobei $\varphi(x)$ darin durch $\neg\psi(x)$ ersetzt wird:

  $$[\neg\forall x\colon \neg\psi(x)] \;\Leftrightarrow\; \exists x\colon \neg\neg\psi(x).$$

  Nun ist $\neg\neg\psi(x)$ gleichwertig mit $\psi(x)$; wir können also die doppelte Negation weglassen:

  $$[\neg\forall x\colon \neg\psi(x)] \;\Leftrightarrow\; \exists x\colon \psi(x). \tag{5}$$

  Weil für beliebige Aussagen A und B die Äquivalenz $\neg A \Leftrightarrow B$ den gleichen Wahrheitswert wie $A \Leftrightarrow \neg B$ hat (sozusagen "doppelte Kontraposition"), können wir die Formel (5) auch in der Form

  $$[\forall x\colon \neg\psi(x)] \;\Leftrightarrow\; \neg\exists x\colon \psi(x)$$

  Die letzte Formel war zu zeigen.

  □

Merkregel: Eine Aussage mit einer Quantorenfolge zu Beginn, z.B. **Wichtig!**

$$\forall x_1\, \exists x_2\, \forall x_3\, \exists x_4 : \ldots, \tag{6}$$

wird negiert, indem man die Allquantoren mit Existenzquantoren vertauscht und den aussagenlogischen "Kern" der Formel negiert. Zum Beispiel lautet das Gegenteil des Formelfragments (6):

$$\exists x_1\, \forall x_2\, \exists x_3\, \forall x_4 : \neg \ldots.$$

Man darin ein prädikatenlogisches Analogon zu den aussagenlogischen de-Morgan-Regeln sehen.

□

Typisch für die Analysis sind komplexe Kombinationen alternierender Quantoren, z.B.

$$\forall x \in \mathbb{R}\, \forall \varepsilon > 0\, \exists \delta > 0\, \forall y \in \mathbb{R}\colon (|x - y| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon).$$

Hierbei muß man sehr genau auf die Reihenfolge der Quantoren achten.

Beispiel 1:

$$\text{“}\forall n \in \mathbb{N}_0 \, \exists m \in \mathbb{N}_0 \colon m > n\text{”}$$

versus

$$\text{“}\exists m \in \mathbb{N}_0 \, \forall n \in \mathbb{N}_0 \colon m > n\text{”}.$$

Die erste Formel besagt, daß es zu jeder natürlichen Zahl n eine größere natürliche Zahl m gibt. Das ist wahr; man nehme z.B. $m = n + 1$.
Die zweite Formel besagt, daß es eine natürliche Zahl m gibt, die größer als alle natürlichen Zahlen n ist. Das ist falsch: Gegeben $m \in \mathbb{N}_0$, können wir $n = m$ wählen; dann gilt $m > n$ jedoch nicht.
Die erste Formel behauptet also nur die Existenz “beliebig großer” natürlicher Zahlen; aber die zweite Formel behauptet die Existenz “unendlich großer” natürlicher Zahlen.
Ähnlich verhält es sich mit dem Unterschied zwischen “infinitesimalen Zahlen” (die es im Rahmen der Standard-Analysis nicht gibt) und “beliebig kleinen” Zahlen (mit denen wir in dieser Vorlesung viel arbeiten werden).

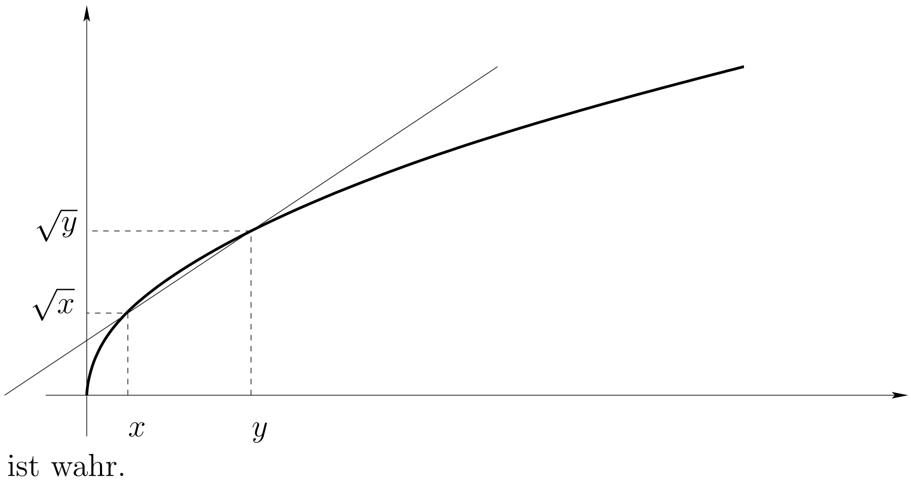
Beispiel 2:

$$\text{“}\forall x > 0 \, \exists M > 0 \, \forall y > 0 \colon |\sqrt{x} - \sqrt{y}| \leq M|x - y|\text{”} \tag{7}$$

versus

$$\text{“}\exists M > 0 \, \forall x > 0 \, \forall y > 0 \colon |\sqrt{x} - \sqrt{y}| \leq M|x - y|\text{”} \tag{8}$$

Die Formel (7) behauptet, daß die Steigung der Sekante durch die Punkte $(x, \sqrt{x})$ und $(y, \sqrt{y})$ bei *festgehaltenem* x nicht beliebig groß werden kann, also durch ein $M > 0$ beschränkt werden kann.



Formel (7) ist wahr.
Beweis: Es sei $x > 0$ gegeben. Wir wählen $M = 1/\sqrt{x}$. Es sei nun $y > 0$ gegeben. Es gilt $\sqrt{y} \geq 0$, also $\sqrt{x} + \sqrt{y} \geq \sqrt{x}$; wegen $|x - y| \geq 0$ folglich

$$\frac{|x - y|}{\sqrt{x} + \sqrt{y}} \leq \frac{|x - y|}{\sqrt{x}} = M|x - y|.$$

Mit Hilfe von $(\sqrt{x}-\sqrt{y})(\sqrt{x}+\sqrt{y}) = x-y$ folgt hieraus die Behauptung:

$$|\sqrt{x}-\sqrt{y}| = \frac{|x-y|}{\sqrt{x}+\sqrt{y}} \le M|x-y|.$$

□

Die Formel (8) behauptet jedoch, daß die Steigung der Sekante *gleichmäßig* für alle $x > 0$ und $y > 0$ durch ein $M > 0$ beschränkt werden kann. "*Gleichmäßig*" bedeutet hier, daß M weder von x noch von y abhängen darf.
Formel (8) ist falsch. Anschaulich ist das plausibel: Wählen wir x und y beide "beliebig nahe" bei 0, so wird die Sekantensteigung "beliebig groß".
Wir beweisen nun das Gegenteil[3] von (8). Wir zeigen also:

$$\forall M > 0 \quad \exists x > 0 \quad \exists y > 0\colon |\sqrt{x}-\sqrt{y}| > M|x-y| \tag{9}$$

Beweis: Es sei $M > 0$ gegeben. Wir wählen

$$x = \frac{1}{4M^2} > 0 \qquad \text{und} \qquad y = \frac{x}{4} > 0.$$

Dann gilt:

$$\sqrt{x}-\sqrt{y} = \frac{1}{2M} - \frac{1}{4M} = \frac{1}{4M}$$

und

$$x - y = \frac{1}{4M^2} - \frac{1}{16M^2} = \frac{3}{16M^2},$$

folglich

$$|\sqrt{x}-\sqrt{y}| = \frac{1}{4M} > \frac{3}{16M} = M \cdot \frac{3}{16M^2} = M|x-y|,$$

also die Behauptung

$$|\sqrt{x}-\sqrt{y}| > M|x-y|.$$

□

Zur Quantorenbehandlung in den Beweisen. Die Behandlung der Quantoren in den beiden Beweisen erfolgt genau nach der Reihenfolge der Quantoren in der zu beweisenden Formel. Um das zu sehen, analysieren wir für jeden Beweisschritt, der einen Quantor behandelt, die aktuell gegebenen Voraussetzungen und verbleibende Behauptung:

Analyse des Beweises der Formel (7):

`Zu zeigen:` $\forall x > 0\, \exists M > 0\, \forall y > 0\colon |\sqrt{x}-\sqrt{y}| \le M|x-y|$

Es sei $x > 0$ gegeben.

[3] Man beachte, dass das Gegenteil hier durch "Umdrehen" der Quantoren und Negieren der Ungleichung gebildet wird.

$\boxed{\texttt{Noch zu zeigen: } \exists M > 0\, \forall y > 0 : |\sqrt{x} - \sqrt{y}| \leq M|x-y| \texttt{ gegeben } x > 0.}$

Wir wählen $M = 1/\sqrt{x}$.

$\boxed{\texttt{Noch zu zeigen: } \forall y > 0 : |\sqrt{x} - \sqrt{y}| \leq M|x-y| \texttt{ gegeben } x > 0,\ M = 1/\sqrt{x}.}$

Es sei nun $y > 0$ gegeben.

$\boxed{\texttt{Noch zu zeigen: } |\sqrt{x} - \sqrt{y}| \leq M|x-y| \texttt{ gegeben } x > 0,\ M = 1/\sqrt{x},\ y > 0.}$

$(\ldots)$ Es folgt die Behauptung:

$$|\sqrt{x} - \sqrt{y}| = \ldots \leq M|x-y|.$$

$\boxed{\texttt{Nichts mehr zu zeigen.}}$

Analyse des Beweises des Gegenteils (9) der Formel (8):

$\boxed{\texttt{Zu zeigen: } \forall M > 0\ \exists x > 0\ \exists y > 0 : |\sqrt{x} - \sqrt{y}| > M|x-y|}$

Es sei $M > 0$ gegeben.

$\boxed{\texttt{Noch zu zeigen: } \exists x > 0\ \exists y > 0 : |\sqrt{x} - \sqrt{y}| > M|x-y| \texttt{ gegeben } M > 0.}$

Wir wählen $x = \frac{1}{4M^2} > 0$.

$\boxed{\texttt{Noch zu zeigen: } \exists y > 0 : |\sqrt{x} - \sqrt{y}| > M|x-y| \texttt{ gegeben } M > 0,\ x = \frac{1}{4M^2} > 0.}$

Wir wählen $y = \frac{x}{4} > 0$.

$\boxed{\texttt{Noch zu zeigen: } |\sqrt{x} - \sqrt{y}| > M|x-y| \texttt{ gegeben } M > 0,\ x = \frac{1}{4M^2} > 0,\ y = \frac{x}{4} > 0.}$

$(\ldots)$ Folglich gilt die Behauptung

$$|\sqrt{x} - \sqrt{y}| > M|x-y|.$$

$\boxed{\texttt{Nichts mehr zu zeigen.}}$

Abstrahieren wir aus den Beispielen zwei prädikatenlogische Herleitungsregeln zur Behandlung der Quantoren:

Entfernung des $\forall$-Quantors aus der Behauptung:

$\boxed{\texttt{Zu zeigen: } \quad \forall x \text{ vom Typ } T : B(x) \quad \texttt{gegeben } A}$

Es sei x vom Typ T gegeben.

`Noch zu zeigen:` $B(x)$ `gegeben` A, x vom Typ T.

Wenn über x im gegebenen A Annahmen gemacht wurden, muss man x mit einer neuen, noch unverbrauchten Variable umbenennen.

Entfernung des $\exists$-Quantors aus der Behauptung:

`Zu zeigen:` $\exists x$ vom Typ $T : B(x)$ `gegeben` A

Wir wählen $x =$ Term.

Hierbei muss Term *den Typ T besitzen.*

`Noch zu zeigen:` $B(x)$ `gegeben` A, $x =$ Term.

Auch hier darf x in der Annahme A nicht frei vorkommen; notfalls umbenennen! Alternativ, ohne Verwendung von x, kann man das verbleibende Beweisziel auch so formulieren:

`Noch zu zeigen:` $B(\text{Term})$ `gegeben` A.

Der "Term" muss beim Finden des Beweises geschickt gewählt werden. Man überlegt ihn sich am besten vorab mit Hilfe anschaulicher Überlegungen oder mit Hilfe einer Nebenrechnung, die im eigentlichen Beweis nicht mehr vorkommt.

Verwendung von gegebenen All- und Existenzaussagen. Wir besprechen nun noch zwei weitere prädikatenlogische Herleitungsregeln, mit deren Hilfe gegebene All- und Existenzaussagen angewandt werden können.
Hier ein klassisches Beispiel zur Herleitung aus einer gegebenen Allaussage:

Alle Menschen sind sterblich. Sokrates ist ein Mensch. Also ist Sokrates sterblich.

Abstrahieren wir daraus eine weitere prädikatenlogische Herleitungsregel:

Entfernung des $\forall$-Quantors aus einer gegebenen Aussage:

`Zu zeigen:` Φ. `Gegeben:` A, $\forall x$ vom Typ $T :\ B(x)$.

Wir wenden die gegebene Allaussage auf $x =$ Term an.

Hierbei muss Term *den Typ T besitzen.*

`Weiterhin zu zeigen:` Φ. `Gegeben:` A, $B(\text{Term})$, $\forall x$ vom Typ $T :\ B(x)$.

Bei diesem Schluss kommt $B(\text{Term})$ als neu gegebene Aussage hinzu. Die zu zeigende Behauptung Φ bleibt bei diesem Schluss unverändert. Auch die Allaussage $\forall x$ vom Typ $T :\ B(x)$

bleibt weiterhin gegeben; sie kann zum Beispiel später noch auf andere Terme angewandt werden.

Die folgende Herleitungsregel beschreibt, wie man *gegebene* Existenzaussagen verwenden kann:

Entfernung des $\exists$-Quantors aus einer gegebenen Aussage:

`Zu zeigen:` Φ. `Gegeben:` $A, \exists x$ vom Typ $T: B(x)$.

Wir nehmen so ein x vom Typ T und nennen es y.

Hierbei muss y eine neue, noch nicht frei vorkommende Variable sein. [4]

`Weiterhin zu zeigen:` Φ. `Gegeben:` A, y vom Typ T, $B(y)$.

Die zu zeigende Behauptung Φ ändert sich bei diesem Schluss nicht. Natürlich bleibt die Existenzaussage $\exists x$ vom Typ $T: \;B(x)$ auch weiterhin gegeben, doch ist es nun nicht mehr nützlich, sie noch bei den gegebenen Aussagen aufzulisten, weil sie nicht mehr Information enthält, als schon in Aussage $B(y)$ mit der neuen freien Variablen y vom Typ T steht.

Zur Illustration der Verwendung der beiden Herleitungsregeln zur Entfernung von Quantoren aus *gegebenen* Aussagen wird hier ein typisches Beweisfragment in der Analysis dargestellt:

Gegeben seien $\epsilon > 0$ und eine Folge $(a_n)_{n\in\mathbb{N}}$ reeller Zahlen mit[5]

$$\forall \tilde{\epsilon} > 0\; \exists m \in \mathbb{N}\; \forall n > m : \; |a_n| < \tilde{\epsilon}.$$

Wir wenden diese Aussage auf $\tilde{\epsilon} = \frac{\epsilon}{2}$ an und erhalten

$$\exists m \in \mathbb{N}\; \forall n > m : \; |a_n| < \frac{\epsilon}{2}.$$

Wir nehmen so ein m und nennen es k. Damit wissen wir $k \in \mathbb{N}$ und

$$\forall n > k : \; |a_n| < \frac{\epsilon}{2}.$$

[4] Die Umbenennung der *gebundenen* Variable x in die *freie* Variable y kann zweckmäßig sein, wenn man die Bezeichnung x später noch für eine andere *freie* Variable mit einer anderen Bedeutung zur Verfügung haben will. Dieser Fall tritt recht häufig auf, doch in einfachen Fällen kann man auch auf die Umbenennung verzichten.

[5] Um Verwechslungen der *freien* Variable ϵ mit der durch den Allquantor *gebundenen* Variable $\tilde{\epsilon}$ zu vermeiden, wurden hier verschiedene Variablennamen gewählt. Die Verwendung der gleichen Bezeichnung für freie und gebundene Variablen kann manchmal zu Missverständnissen führen.

Ausblick. Die vier hier besprochenen prädikatenlogischen Herleitungsregeln "$\forall$-Entfernung und $\exists$-Entfernung aus der Behauptung und aus gegebenen Aussagen" reichen zusammen mit den aussagenlogischen Herleitungsregeln aus, um alle Aussagen über Objekte, die in allen "Modellen" inhaltlich gültig sind, auch formal herzuleiten. Eine Präzisierung dieser Aussage ist der berühmte *Gödelsche Vollständigkeitssatz*, den Sie in der mathematischen Logik kennenlernen können.

1.2 Mengen

Mengenlehre ist die Sprache der modernen Mathematik. Man kann sogar die gesamte Mathematik auf Axiome der Mengenlehre und logisches Schließen aufbauen. Die Behandlung dieser "axiomatischen Mengenlehre" ist Spezialvorlesungen vorbehalten; wir besprechen hier nur die "naive Mengenlehre" als unsere mathematische Standardsprache, ohne Axiome der Mengenlehre vollständig aufzuzählen.

Eine *Menge* ist eine Zusammenfassung von mathematischen Objekten zu einem neuen mathematischen Objekt. Die zusammengefassten Objekte heißen *Elemente* der Menge. Ist x ein Objekt und M eine Menge, so bedeutet $x \in M$, dass x ein Element von M ist. $x \notin M$ bezeichnet das Gegenteil von $x \in M$. Dabei werden Mengen eindeutig durch die Gesamtheit ihrer Elemente charakterisiert:

Extensionalitätsaxiom

Zwei Mengen M und N sind genau dann gleich, wenn sie die gleichen Elemente besitzen.
In Formeln:

$$\forall M \text{ Menge } \forall N \text{ Menge } ((\forall x(x \in M \Leftrightarrow x \in N)) \Leftrightarrow M = N)$$

Teilmengen. Eine Menge M heißt eine *Teilmenge* einer Menge N, in Zeichen $M \subseteq N$ oder auch $N \supseteq M$, wenn jedes Element von M auch ein Element von N ist. In Formeln:

$$\forall M \text{Menge } \forall N \text{Menge } ((\forall x(x \in M \Rightarrow x \in N)) \Leftrightarrow M \subseteq N)$$

Das Extensionalitätsaxiom kann man damit auch in der folgenden Form schreiben:

$$\forall M \text{Menge } \forall N \text{Menge } (M \subseteq N \land M \supseteq N \Leftrightarrow M = N)$$

Dies ist ein wichtiges Prinzip, um die Gleichheit zweier Mengen zu beweisen.
Das Symbol $M \subset N$ soll $M \subseteq N \land M \neq N$ bedeuten; wir verwenden es nur selten.

Widersprüchlichkeit des unbeschränkten Komprehensionsprinzips. Eine Intention der Cantorschen Mengenlehre war es, *jeder* Aussage $\phi(x)$ über Objekte x eine Menge $\{x|\ \phi(x)\}$ zuzuordnen, die genau die Objekte x mit der Eigenschaft $\phi(x)$ als Elemente besitzt. Bertrand Russell entdeckte ein einfaches Argument, dass dieses "unbeschränkte Komprehensionsprinzip" widersprüchlich ist: Setzt man $R = \{x|\ x \notin x\}$, so erhält man den Widerspruch $R \in R \Leftrightarrow R \notin R$. Dieser Widerspruch kann zum Beispiel dadurch aufgelöst werden, dass man nicht unbeschränkte Komprehension zulässt, sondern nur beschränkte Mengenbildungen erlaubt, z.B. das folgende *Aussonderungsprinzip* (synonym: beschränkte Komprehensionsprinzip) und weitere ausgewählte Mengenbildungen, von denen manche unten aufgeführt sind. Das unbeschränkte Komprehensionsprinzip wird also verworfen. [6]

Aussonderungsprinzip. Zu jeder Aussage $\phi(x)$ über Objekte x und zu jeder Menge M gibt es eine Menge, deren Elemente genau die Elemente x von M mit der Eigenschaft $\phi(x)$ sind; wir schreiben $\{x \in M |\ \phi(x)\}$ für diese Menge.

Durchschnitt, Vereinigung und Mengendifferenz zweier Mengen. Zu je zwei Mengen M und N gibt es die *Schnittmenge* $M \cap N$, deren Elemente genau diejenigen Objekte sind, die Elemente von M *und* von N sind. Weiter gibt es die *Vereinigungsmenge* $M \cup N$, deren Elemente genau diejenigen Objekte sind, die Elemente von M *oder* von N sind. In Formeln:

$$\begin{aligned} &\forall M \text{ Menge } \forall N \text{ Menge } \forall x (x \in M \cap N \Leftrightarrow x \in M \wedge x \in N) \\ &\forall M \text{ Menge } \forall N \text{ Menge } \forall x (x \in M \cup N \Leftrightarrow x \in M \vee x \in N) \end{aligned}$$

oder auch

$$\begin{aligned} M \cap N &= \{x |\ x \in M \wedge x \in N\} \\ M \cup N &= \{x |\ x \in M \vee x \in N\}. \end{aligned}$$

Die *Mengendifferenz* $M \setminus N$ zweier Mengen M und N wird durch

$$M \setminus N = \{x \in M |\ x \notin N\}$$

definiert; die *symmetrische Differenz* $M \triangle N$ durch

$$M \triangle N = \{x |\ \neg(x \in M \Leftrightarrow x \in N)\} = (M \setminus N) \cup (N \setminus M).$$

Die aussagenlogischen Gesetze haben Entsprechungen in der Mengenlehre; zum Beispiel gelten die Distributivgesetze für Mengen A, B, C:

$$\begin{aligned} A \cap (B \cup C) &= (A \cap B) \cup (A \cap C), \\ A \cup (B \cap C) &= (A \cup B) \cap (A \cup C). \end{aligned}$$

Ebenso haben die de-Morgan'schen Regeln der Aussagenlogik Entsprechungen in der Mengenlehre. Um sie formulieren zu können, brauchen wir folgendes mengentheoretisches

[6] Alternativ kann man allgemeiner als "Mengen" auch "*Klassen*" einführen und "Klassenbildungen" $\{x |\ \phi(x)\}$ auf *Mengen* x beschränken. "Zu große Klassen" sind dann keine Mengen mehr. Diesen alternativen Aufbau der Mengenlehre mit "Mengen" und "Klassen" verwenden wir in dieser Vorlesung nicht.

Analogon zur Negation: Wir fixieren eine Menge Ω, "das Universum", und definieren die "Komplementbildung im Universum Ω" durch $M^c = \Omega \setminus M$ für $M \subseteq \Omega$. Die de-Morgan'schen Regeln der Mengenlehre für Mengen A, B mit $A \subseteq \Omega$ und $B \subseteq \Omega$ lauten damit:

$$(A^c \cap B^c)^c = A \cup B,$$
$$(A^c \cup B^c)^c = A \cap B.$$

Man beachte im Vergleich mit den aussagenlogischen Regeln, dass einfach nur Aussagen durch Mengen und die Junktoren $\wedge$, $\vee$ und $\neg$ durch die Mengenoperationen $\cap$, $\cup$ und ^{c} ersetzt wurden.

Mengen in aufzählender Notation. Die einfachste Menge ist die *leere Menge* $\emptyset$, die kein Element enthält. Zu jedem Objekt x gibt es auch eine Menge $\{x\}$, die genau das Element x enthält. Die Menge $\{x\}$ heißt *Einermenge* oder auch *Singleton* zu x und darf nicht mit x verwechselt werden. Ebenso gibt es zu zwei Objekten x, y die Menge $\{x, y\}$, die genau x und y als Elemente enthält. Allgemeiner gibt es zu jeder Liste $x_1, \ldots, x_n$ von Objekten die Menge $\{x_1, \ldots, x_n\}$, die genau die Elemente $x_1, \ldots, x_n$ enthält. Eine Menge heißt *endlich*, wenn sie leer ist oder die Gestalt $\{x_1, \ldots, x_n\}$ mit einer natürlichen Zahl n hat; andernfalls nennt man sie *unendlich*.

Paare und Tupel. Je zwei Objekten x, y kann man auch ein *Paar* (x, y) zuordnen. Es besitzt die fundamentale Eigenschaft

$$\forall x \, \forall y \, \forall u \, \forall v \, ((x, y) = (u, v) \Leftrightarrow x = u \wedge y = v)$$

Im Gegensatz zur Menge $\{x, y\} = \{y, x\}$ kommt es also beim Paar auf die Reihenfolge der Einträge an: $(x, y) = (y, x)$ gilt nur für $x = y$. Allgemeiner ordnet man jeder Liste $x_1, \ldots, x_n$ von Objekten ein *n-Tupel* $(x_1, \ldots, x_n)$ zu, das wieder die die fundamentale Eigenschaft

$$\forall x_1, \ldots, x_n \, \forall y_1, \ldots, y_n \, ((x_1, \ldots, x_n) = (y_1, \ldots y_n) \Leftrightarrow x_1 = y_1 \wedge \ldots \wedge x_n = y_n)$$

besitzt. Im Gegensatz zur Menge $\{x_1, \ldots, x_n\}$ ist beim Tupel $(x_1, \ldots, x_n)$ die Reihenfolge der Einträge bedeutsam. 3-Tupel (x_1, x_2, x_3) werden auch *Tripel* genannt; 4-Tupel (x_1, x_2, x_3, x_4) auch *Quadrupel*.

Kartesisches Produkt. Zu je zwei Mengen M und N gibt es eine Menge $M \times N$, deren Elemente genau die Paare (x, y) mit $x \in M$ und $y \in N$ sind:

$$M \times N = \{z | \, \exists x \in M \, \exists y \in N : \; z = (x, y)\}$$

oder dasselbe in Kurznotation:

$$M \times N = \{(x, y) | \; x \in M, \; y \in N\}$$

Die Menge $M \times N$ wird *kartesisches Produkt* von M und N genannt.
Allgemeiner gibt es zu n Mengen $M_1, \ldots, M_n$ eine Menge $M_1 \times \ldots \times M_n$, deren Elemente genau die Paare $(x_1, \ldots, x_n)$ mit $x_1 \in M_1, \ldots, x_n \in M_n$ sind:

$$M_1 \times \ldots \times M_n = \{(x_1, \ldots, x_n) |\ x_1 \in M_1, \ldots, x_n \in M_n\}$$

Auch diese Menge wird (n-faches) kartesisches Produkt genannt. Im Spezialfall, dass alle $M_1, \ldots, M_n$ mit einer Menge M übereinstimmen, $M_1 = \ldots = M_n = M$, schreiben wir auch $M^n = M_1 \times \ldots \times M_n$.

Relationen. Es seien M und N Mengen. Eine Teilmenge $R \subseteq M \times N$ heißt auch *(zweistellige) Relation* auf M und N. Man schreibt auch xRy statt $(x, y) \in R$. Allgemeiner werden Teilmengen R eines n-fachen kartesischen Produkts $M_1 \times \ldots \times M_n$ (n-stellige) Relationen auf $M_1, \ldots, M_n$ genannt.
Die zu einer zweistelligen Relation $R \subseteq M \times N$ gebildete Relation

$$R^{-1} := \{(y, x) \in N \times M |\ (x, y) \in R\}$$

heißt *Umkehrrelation* von R. Für drei Mengen L, M und N und zwei Relationen $R \subseteq L \times M$ und $S \subseteq M \times N$ wird die *Komposition* $S \circ R \subseteq L \times N$ durch

$$S \circ R = \{(x, z) \in L \times N |\ \exists y \in M : \ (x, y) \in R \wedge \ (y, z) \in S\}$$

definiert. Man beachte, dass die Notation $S \circ R$ am anschaulichsten von rechts nach links gelesen wird.

Eine zweistellige Relation $R \subseteq M \times M$ auf einer Menge M heißt *reflexiv*, falls xRx für alle $x \in M$ gilt. Sie heißt *symmetrisch*, falls gilt:

$$\forall x \in M \ \forall y \in M (xRy \Rightarrow yRx).$$

Die Relation R nennt man *transitiv*, falls gilt

$$\forall x \in M \ \forall y \in M \ \forall z \in M (xRy \wedge yRz \Rightarrow xRz).$$

Ist R reflexiv, symmetrisch und transitiv, so wird R eine *Äquivalenzrelation* genannt; in diesem Fall heißt $[x]_R := \{y \in M |\ xRy\}$ die *Äquivalenzklasse* von $x \in M$.

Funktionen und Abbildungen. Eine zweistellige Relation $f \subseteq M \times N$ auf zwei Mengen M und N heißt *Funktion*, wenn gilt:

$$\forall x \in M \ \forall y \in N \ \forall z \in N : \ ((x, y) \in f \wedge (x, z) \in f \Rightarrow y = z)$$

In diesem Fall heißt

$$D(f) := \{x \in M |\ \exists y \in N : \ (x, y) \in f\}$$

der *Definitionsbereich* (engl. "domain") von f und

$$R(f) := \{y \in N |\ \exists x \in M: \ (x,y) \in f\}$$

der *Wertebereich* (engl. "range") von f. Für $x \in D(f)$ bezeichnet $f(x)$ das eindeutig bestimmte $y \in N$ mit $(x,y) \in f$. Der Term $f(x)$ wird der Wert von f an der Stelle x genannt.
Ein Tripel (f, M, N), bestehend aus einer Funktion $f \subseteq M \times N$ zusammen mit zwei Mengen M und N heißt eine *Abbildung* von M nach N, in Zeichen $f : M \to N$, wenn $D(f) = M$ gilt. In diesem Fall heißt M der Definitionsbereich und N der *Zielbereich* (engl. "codomain") der Abbildung $f : M \to N$. Soll eine Abbildung $f : M \to N$ durch einen von $x \in M$ abhängigen Term term(x) definiert werden, so schreibt man oft

$$f : M \to N,\ f(x) := \mathrm{term}(x)$$

oder synonym

$$f : M \to N,\ x \mapsto \mathrm{term}(x).$$

Man beachte hier die Verwendung zweier verschiedener Pfeilsymbole.
Stimmen Wertebereich $R(f)$ und Zielbereich N einer Abbildung $f : M \to N$ überein, d.h. wenn es für alle $y \in N$ ein $x \in M$ mit $f(x) = y$ gibt, so heißt die Abbildung $f : M \to N$ *surjektiv* (synonym: eine *Surjektion*). Man beachte, dass dieser Begriff nur für Abbildungen $f : M \to N$ sinnvoll ist, jedoch nicht für die zugrundeliegende Funktion f alleine. Falls für alle $x_1 \in M$ und $x_2 \in M$ mit $f(x_1) = f(x_2)$ gilt: $x_1 = x_2$, so heißt die Abbildung $f : M \to N$ *injektiv* (synonym: eine *Injektion*). Eine Abbildung $f : M \to N$ heißt *bijektiv* (synonym: eine *Bijektion*), wenn sie injektiv und surjektiv ist. Ist $f : M \to N$ eine Bijektion, so liefert die Umkehrrelation $f^{-1} \subseteq N \times M$ eine Bijektion $f^{-1} : N \to M$ in umgekehrter Richtung; sie wird *Umkehrabbildung* (synonym: *inverse Abbildung* oder *Inverse*) von $f : M \to N$ genannt. Sie besitzt die Eigenschaften

$$\begin{aligned} \forall x \in M: \ f^{-1}(f(x)) &= x, \\ \forall y \in N: \ f(f^{-1}(y)) &= y. \end{aligned}$$

Für Mengen L, M und N und Abbildungen $f : L \to M$ und $g : M \to N$ ist auch die Komposition $g \circ f$ wieder eine Abbildung: $g \circ f : L \to N$; für sie gilt $(g \circ f)(x) = g(f(x))$ für alle $x \in L$. Diese Gleichung motiviert die Konvention, die erste Abbildung in einer Komposition rechts, die zweite links zu schreiben, da Funktionen f in Funktionsauswertungen $f(x)$ links von ihren Argumenten x geschrieben werden.

Bilder und Urbilder von Mengen. Es sei $f : M \to N$ eine Abbildung. Für eine Teilmenge $U \subseteq M$ heißt

$$f[U] = \{y \in N | \exists x \in U: \ y = f(x)\},$$

kurz $f[U] = \{f(x) |\ x \in U\}$, das *Bild* der Menge U unter f. Für eine Teilmenge $V \subseteq N$ heißt

$$f^{-1}[V] = \{x \in M |\ \exists y \in V: \ y = f(x)\}$$

das *Urbild* der Menge V unter f. Man beachte, dass dieses Urbild für beliebige Abbildungen $f : M \to N$ existiert, nicht nur für Bijektionen $f : M \to N$. Man verwechsle die Urbildbildung also nicht mit der Umkehrfunktion, trotz der Ähnlichkeit der Schreibweisen! Falls $f : M \to N$ eine Bijektion ist, kann man die Notation $f^{-1}[V]$ auf zwei Weisen lesen: Als Urbild von V unter f oder als Bild von V unter der Umkehrabbildung f^{-1}. Da diese beiden Interpretationen aber übereinstimmen, führt diese Zweideutigkeit nicht zu Missverständnissen.

Mengensysteme. Da Mengen selbst wieder mathematische Objekte sind, können sie wieder als Elemente von Mengen auftreten. Mengen, deren Elemente selbst Mengen sind, nennt man *Mengensysteme*. Mengensysteme kann man sich in einem hierarchischen Aufbau des Mengenuniversums als "Mengen höherer Stufe" vorstellen; naturgemäß sind sie abstrakter als die bisher betrachteten Mengen und bereiten Studierenden beim ersten Kennenlernen daher oft Schwierigkeiten. Zum Beispiel besitzt jede Menge M eine *Potenzmenge* $\mathcal{P}(M)$, deren Elemente genau die Teilmengen von M sind:

$$\mathcal{P}(M) := \{N|\ N \subseteq M\}.$$

Ein weiteres Beispiel für Mengensysteme ist der *Quotientenraum* (synonym: *Faktorraum*) einer Menge M nach einer Äquivalenzrelation $\sim \subseteq M \times M$ auf M, der als die Menge aller Äquivalenzklassen bezüglich $\sim$ definiert ist:

$$M/\sim := \{[x]_\sim|\ x \in M\}$$

Die Abbildung $f : M \to M/\sim$, $x \mapsto [x]_\sim$, wird *kanonische Abbildung* in den Quotientenraum genannt.

Familien und allgemeine kartesische Produkte. Ist f eine Funktion mit Definitionsbereich $D(f) = I$, so verwendet man statt der "Klammernotation" $f(i)$ für den Wert von f an der Stelle $i \in I$ manchmal auch die "Indexnotation" f_i. In diesem Zusammenhang nennt man die Funktion f auch eine "*Familie*" mit der "*Indexmenge*" I und schreibt $f = (f_i)_{i\in I}$ statt $f : I \to R(f), i \mapsto f(i)$. Zum Beispiel kann man ein n-Tupel $(x_1, \dots, x_n)$ auch mit der Familie $(x_i)_{i\in\{1,\dots,n\}}$ identifizieren. Eine *Folge* ist eine Familie $(x_n)_{n\in\mathbb{N}_0}$ mit der Indexmenge $\mathbb{N}_0$, der Menge der natürlichen Zahlen inklusive 0; oft lässt man auch den Index 0 weg. Veranschaulichend schreibt man manchmal auch $x_0, x_1, x_2, \dots$ für eine Folge $(x_n)_{n\in\mathbb{N}_0}$. Ist $(M_i)_{i\in I}$ eine Familie von Mengen, so nennt man

$$\prod_{i\in I} M_i = \{(x_i)_{i\in I}|\ \forall i \in I : x_i \in M_i\}$$

das (allgemeine) kartesische Produkt der Mengen M_i, $i \in I$. Das kartesische Produkt $\prod_{i\in I} M_i$ besteht also genau aus allen Funktionen x mit Definitionsbereich I, so dass für alle $i \in I$ gilt: $x(i) \in M_i$. Im Spezialfall, dass alle M_i gleich einer Menge M sind, schreibt man auch

$$M^I := \prod_{i\in I} M$$

und nennt dies die *kartesische Potenz* von M mit I. Die kartesische Potenz M^I enthält also genau alle Funktionen f mit $f : I \to M$ als Elemente. Zum Beispiel ist $M^{\mathbb{N}}$ die Menge aller Folgen[7] $(a_n)_{n\in\mathbb{N}}$, veranschaulichend notiert $a_1, a_2, a_3, \ldots$, mit Werten in M. Identifiziert man wie oben n-Tupel mit Familien zur Indexmenge $\{1, \ldots, n\}$, so identifiziert man die kartesischen Produkte

$$\prod_{i=1}^{n} M_i := \prod_{i\in\{1,\ldots,n\}} M_i = M_1 \times \ldots \times M_n.$$

Damit wird für eine Menge M die Menge M^n von n-Tupeln mit der kartesischen Potenz $M^{\{1,\ldots,n\}}$ identifiziert. Zum Beispiel besteht M^3 aus allen Tripeln (m_1, m_2, m_3) mit Einträgen $m_1, m_2, m_3 \in M$, während $M^{\{1,2,3\}}$ aus allen Familien $(m_i)_{i=1,2,3}$, ebenfalls mit Einträgen $m_1, m_2, m_3 \in M$, besteht. Das Tripel (m_1, m_2, m_3) wird also mit der Funktion mit dem Definitionsbereich $\{1, 2, 3\}$ und den Zuordnungen $1 \mapsto m_1$, $2 \mapsto m_2$ und $3 \mapsto m_3$ identifiziert.

Allgemeine Vereinigungen und Durchschnitte. Zu jeder Familie von Mengen $(M_i)_{i\in I}$ gibt es auch die (allgemeine) Vereinigungsmenge

$$\bigcup_{i\in I} M_i = \{x |\ \exists i \in I : \ x \in M_i\}$$

und, falls I nichtleer[8] ist, den (allgemeinen) Durchschnitt

$$\bigcap_{i\in I} M_i = \{x |\ \forall i \in I : \ x \in M_i\}.$$

Auswahlaxiom. In der axiomatischen Mengenlehre fordert man als Axiom, dass das kartesische Produkt $\prod_{i\in I} M_i$ nichtleerer Mengen M_i nichtleer ist. Das bedeutet, dass es zu jeder Familie $(M_i)_{i\in I}$ von nichtleeren Mengen $M_i \neq \emptyset$ eine "*Auswahlfunktion*" $(x_i)_{i\in I}$ gibt, so dass $\forall i \in I : x_i \in M_i$ gilt. Dieses *Auswahlaxiom* ist zwar auf dem ersten Blick anschaulich plausibel. Dennoch ist es sehr nichtkonstruktiv. Wer sich davon überzeugen will, versuche, ein Element von $\prod_{i\in I} M_i$ für $I = \mathcal{P}(\mathcal{P}(\mathbb{N})) \setminus \{\emptyset\}$ und $M_i = i$ explizit anzugeben. Eine alternative äquivalente Formulierung des Auswahlaxioms lautet: *Zu jeder surjektiven Abbildung $f : A \to B$ gibt es eine Rechtsinverse $g : B \to A$ bezüglich der Komposition $\circ$, also eine Abbildung $g : B \to A$, so dass für alle $b \in B$ gilt: $f \circ g(b) = b$.* Um diese Variante aus dem Auswahlaxiom zu folgern, beachte man, dass solch ein g genau das Gleiche wie eine Auswahlfunktion des kartesischen Produkts $\prod_{b\in B} f^{-1}[\{b\}]$ ist, und dass die Surjektivität von f besagt, dass das Urbild $f^{-1}[\{b\}]$ für jedes $b \in B$ nichtleer ist.

[7] hier mit dem Index 1 beginnend geschrieben

[8] Der Durchschnitt über die leere Indexmenge ergibt keinen Sinn, da er jedes Objekt als Element besitzen müsste.

Ausblick. In der Zermelo-Fraenkelschen axiomatischen Mengenlehre besteht das gesamte mathematische Universum aus Mengen, die letztlich durch verschachtelte Elementbeziehungen aus der leeren Menge aufgebaut werden. Die Unterscheidung zwischen Mengen und Mengensystemen entfällt dann. Das Paar (x, y) kann man hier als Menge $\{\{x\}, \{x, y\}\}$ realisieren, die Null als leere Menge $0 = \emptyset$, und die Nachfolgerfunktion N auf den natürlichen Zahlen durch $N(m) = m \cup \{m\}$. Bei diesem Aufbau des mathematischen Universums wird dann $1 = \{0\}$, $2 = \{0, 1\}$, $3 = \{0, 1, 2\}$, etc.; die Kleinerrelation auf den natürlichen Zahlen ist dann die Elementrelation.

1.3 Zahlen

In diesem Abschnitt besprechen wir aus der Hierarchie der Zahlenmengen

$$\mathbb{N}_0 \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$$

also der natürlichen, ganzen, rationalen, reellen und komplexen Zahlen, die Bereiche $\mathbb{N}_0, \mathbb{R}$ und $\mathbb{C}$ etwas genauer.
Eine weitergehende Beschreibung des Aufbaus des Zahlensystems findet sich im lesenswerten Buch "Zahlen" von Ebbinghaus et.al.[EHH$^+$83]

1.3.1 Natürliche Zahlen

Wir beginnen mit einer axiomatischen Charakterisierung der natürlichen Zahlen.

Peano-Axiome

1. 0 *ist eine natürliche Zahl.*

   In Formeln: $0 \in \mathbb{N}_0$.

2. *Jede natürliche Zahl x hat eine natürliche Zahl $N(x)$ als Nachfolger.*

   Mit anderen Worten: Wir haben eine Abbildung $N\colon \mathbb{N}_0 \to \mathbb{N}_0$.

3. *Wenn zwei natürliche Zahlen den gleichen Nachfolger haben, sind sie gleich.*

   In Formeln:
   $$\forall n, m \in \mathbb{N}_0\colon (N(n) = N(m) \Rightarrow n = m).$$
   Anders gesagt: Die Abbildung $N\colon \mathbb{N}_0 \to \mathbb{N}_0$ ist injektiv.

4. *Keine natürliche Zahl hat den Nachfolger* 0.

   In Formeln:
   $$\forall n \in \mathbb{N}_0\colon N(n) \neq 0.$$
   Anders gesagt: $0 \notin N[\mathbb{N}_0]$.

5. **Induktionsschema:** Für jede Aussage $\varphi(n)$ über natürliche Zahlen n gilt:

   Wenn $\varphi(0)$ gilt, und wenn für alle $n \in \mathbb{N}_0$ aus $\varphi(n)$ die Aussage $\varphi(N(n))$ folgt, dann gilt für alle $n \in \mathbb{N}_0$ die Aussage $\varphi(n)$.

   In Formeln:
   $$[\varphi(0) \wedge \forall n \in \mathbb{N}_0\colon (\varphi(n) \Rightarrow \varphi(N(n)))] \Rightarrow \forall n \in \mathbb{N}_0\colon \varphi(n).$$

Wir kürzen ab:

$$1 = N(0), \quad 2 = N(1), \quad 3 = N(2), \quad 4 = N(3), \quad \text{etc.}$$

Das Induktionsschema ist anschaulich plausibel:
Es gelte $\varphi(0)$ und für alle $n\colon \varphi(n) \Rightarrow \varphi(N(n))$.
Damit erhalten wir

$$\varphi(1) \text{ wegen } \varphi(0) \text{ und } \varphi(0) \Rightarrow \varphi(1), \tag{10}$$
$$\varphi(2) \text{ wegen } \varphi(1) \text{ und } \varphi(1) \Rightarrow \varphi(2), \tag{11}$$
$$\varphi(3) \text{ wegen } \varphi(2) \text{ und } \varphi(2) \Rightarrow \varphi(3), \tag{12}$$
$$\vdots \tag{13}$$

und schließlich – gewissermaßen nach unendlich vielen Schritten – $\forall n \in \mathbb{N}_0\colon \varphi(n)$.
Um eine Aussage $\forall n \in \mathbb{N}_0\colon \varphi(n)$ zu beweisen, kann man also so vorgehen:

Prinzip der vollständigen Induktion:

Induktionsanfang: *Man zeige* $\varphi(0)$.
Induktionsvoraussetzung: *Es sei* $n \in \mathbb{N}_0$. *Annahme:* $\varphi(n)$ *gilt.*
Induktionsschritt: *Unter Verwendung der Induktionsvoraussetzung zeige man* $\varphi(N(n))$.

Wichtiges Beweis-prinzip!

Beispiel: Wir zeigen die Bernoullische Ungleichung

$$\forall x \geq -1 \forall n \in \mathbb{N}_0 \colon (1+x)^n \geq 1 + nx.$$

Beweis: Sei $x \geq -1$. Wir beweisen

$$\forall n \in \mathbb{N}_0 : \ (1+x)^n \geq 1 + nx$$

durch vollständige Induktion.

Induktionsanfang: $(1+x)^0 = 1 \geq 1 + 0 \cdot x$
Induktionsvoraussetzung: Es sei $n \in \mathbb{N}_0$, und es gelte

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

Induktionsschritt $n \rightsquigarrow n+1$*:*

$$\begin{aligned}
(1+x)^{n+1} &= (1+x)(1+x)^n \\
&\geq (1+x)(1+nx) \qquad \text{wegen der Induktionsvoraussetzung und } 1+x \geq 0 \\
&= 1 + (n+1)x + nx^2 \\
&\geq 1 + (n+1)x \qquad \text{wegen } nx^2 \geq 0.
\end{aligned}$$

$\square$

Hier ist eine Variante des Induktionsschemas:
Es gelte $\varphi(0)$ *und*

$$\forall n \geq 1 \colon [\varphi(0) \wedge \ldots \wedge \varphi(n-1) \Rightarrow \varphi(n)].$$

Dann gilt

$$\forall n \in \mathbb{N}_0 \colon \varphi(n).$$

Etwas formaler ausgedrückt:

Induktionsschema – Variante:

$$[\forall n : ((\forall m < n \colon \varphi(m)) \Rightarrow \varphi(n))] \Rightarrow \forall n \colon \varphi(n)$$

Hier laufen alle Quantoren über natürliche Zahlen.
Anschaulich ist das Schema plausibel: Es gelte die Prämisse des Schemas.

- Zunächst gilt $\varphi(0)$, weil es kein $m < 0, m \in \mathbb{N}_0$ gibt.
- Dann gilt $\varphi(1)$ wegen $\varphi(0) \Rightarrow \varphi(1)$.
- Dann gilt $\varphi(2)$ wegen $\varphi(0) \wedge \varphi(1) \Rightarrow \varphi(2)$.
- Dann gilt $\varphi(3)$ wegen $\varphi(0) \wedge \varphi(1) \wedge \varphi(2) \Rightarrow \varphi(3)$.
  $\vdots$

Wir verzichten auf eine präzise Herleitung des zweiten Induktionsschemas.

Rekursion Ähnlich wie *Induktion* den *Beweis* von Aussagen durch Rückgriff auf frühere Instanzen erlaubt, dient *Rekursion* zur *Definition* von Objekten durch Rückgriff auf frühere Instanzen.

Beispiel 1: Die Fakultätsfunktion wird rekursiv wie folgt definiert:

Rekursionsanfang:

$$0! := 1 \tag{14}$$

Rekursionsschritt:

$$(n+1)! := (n+1) \cdot n! \qquad \text{für } n \in \mathbb{N}_0 \tag{15}$$

Also: $n! = n \cdot (n-1) \cdot (n-2) \cdot \ldots \cdot 3 \cdot 2 \cdot 1$.

Abstrahieren wir aus dem Beispiel folgenden

Rekursionssatz über $\mathbb{N}_0$ (einfachste Form):
Gegeben seien eine Menge M, ein Element $a \in M$ und eine Abbildung $g : \mathbb{N}_0 \times M \to M$. Dann gibt es genau eine Abbildung $f : \mathbb{N}_0 \to M$ mit

$$\begin{aligned} &f(0) = a, \\ &\forall n \in \mathbb{N}_0 : \ f(n+1) = g(n, f(n)). \end{aligned}$$

Der Beweis davon gehört zur Mengenlehre; wir verzichten hier darauf. Im obigen Beispiel ist $M = \mathbb{N}_0$, $g(n,m) = (n+1)m$, und $f(n) = n!$. Oft verwendet man Varianten und Verallgemeinerungen des hier zitierten einfachsten Rekursionssatzes; wir formulieren diese nicht explizit.

Beispiel 2: Es sei $f\colon \mathbb{N}_0 \to \mathbb{R}$ eine Funktion.
Für $n, m \in \mathbb{N}_0$ mit $m + 1 \geq n$ definieren wir die Summe $\sum_{k=n}^{m} f(k)$ rekursiv.

$$\sum_{k=n}^{n-1} f(k) := 0 \qquad \text{(die “leere Summe” ist 0),}$$

$$\sum_{k=n}^{m} f(k) := f(m) + \sum_{k=n}^{m-1} f(k) \qquad \text{für } m \geq n.$$

Das bedeutet:

$$\sum_{k=n}^{m} f(k) = f(n) + f(n+1) + \ldots + f(m).$$

Beispiel 3: Ebenso wird das Produkt definiert:

$$\prod_{k=n}^{n-1} f(k) \quad := \quad 1, \tag{16}$$

$$\prod_{k=n}^{m} f(k) \quad := \quad f(m) \cdot \prod_{k=n}^{m-1} f(k) \text{ für } m \geq n. \tag{17}$$

Zum Beispiel ist die Fakultätsfunktion gegeben durch

$$n! = \prod_{k=1}^{n} k = 1 \cdot 2 \cdot \ldots \cdot n.$$

Beispiel 4: Die Kleiner-Relation $<$ auf $\mathbb{N}_0 \times \mathbb{N}_0$ kann man rekursiv wie folgt definieren:

$$\forall n \in \mathbb{N}_0 \colon \neg(n < 0), \tag{18}$$

$$\forall n, m \in \mathbb{N}_0 \colon (n < N(m) :\Leftrightarrow n = m \vee n < m). \tag{19}$$

Hierbei soll $A :\Leftrightarrow B$ bedeuten, dass A als äquivalent zu B definiert wird. Die Rekursion läuft dabei über m.
Alternativ (und äquivalent) kann man "$<$" auch so rekursiv definieren:

$$\forall n \in \mathbb{N}_0 : \neg\, n < 0,$$
$$\forall m \in \mathbb{N}_0 : 0 < N(m),$$
$$\forall n \in \mathbb{N}_0 \; \forall m \in \mathbb{N}_0 : (N(n) < N(m) \quad :\Leftrightarrow \; n < m).$$

Beispiel 5: Für $x \in \mathbb{R}, n \in \mathbb{N}_0$ wird der *Binomialkoeffizient* $\binom{x}{n}$ rekursiv wie folgt definiert:

$$\binom{x}{0} := 1, \qquad \binom{x}{n+1} := \frac{x}{n+1} \cdot \binom{x-1}{n}$$

Anders ausgedrückt:

$$\boxed{\binom{x}{n} = \frac{x \cdot (x-1) \cdot (x-2) \cdot \ldots \cdot (x-n+1)}{n \cdot (n-1) \cdot (n-2) \cdot \ldots \cdot 1} = \frac{\prod_{k=0}^{n-1}(x-k)}{n!}}$$

Für $m, n \in \mathbb{N}_0$ ist $\binom{m}{n}$ gleich der Anzahl $N(m,n)$ der n-elementigen Teilmengen einer m-elementigen Menge.
In der Tat erfüllt diese Anzahl die gleiche Rekursionsformel und stimmt daher wegen der Eindeutigkeitsaussage im Rekursionssatz mit dem Binomialkoeffizienten überein:

- Jede Menge hat nur eine 0-elementige Teilmenge, nämlich $\emptyset$.

- Um $(n+1)N(m, n+1) = mN(m-1, n)$ zu beweisen, hier eine Illustration:

  Ein Parlament von m Abgeordneten wählt einen Ausschuß von $n + 1$ Mitgliedern [$N(m, n + 1)$ Möglichkeiten] und dann daraus den Ausschußvorsitzenden [$n + 1$ Möglichkeiten]. Es gibt also $(n + 1)N(m, n + 1)$ mögliche Zusammensetzungen des Ausschusses mit Vorsitzenden.

  Anders gezählt: Erst wählt das Parlament den Vorsitzenden [m Möglichkeiten], dann die restlichen Ausschußmitglieder [$N(m - 1, n)$ Möglichkeiten]. So gezählt gibt es $mN(m - 1, n)$ mögliche Zusammensetzungen.

Beispiel: Die 4-elementige Menge $\{1, 2, 3, 4\}$ hat $\binom{4}{2} = \frac{4 \cdot 3}{2 \cdot 1} = 6$ zweielementige Teilmengen, nämlich $\{1, 2\}$, $\{1, 3\}$, $\{1, 4\}$, $\{2, 3\}$, $\{2, 4\}$ und $\{3, 4\}$.

Hier ist eine weitere Rekursionsformel für $\binom{m}{n}$ für $m, n \in \mathbb{N}_0$:

$$\boxed{\binom{m+1}{n+1} = \binom{m}{n+1} + \binom{m}{n}}$$

In der Tat: $\binom{m}{n+1}$ ist die Anzahl der $n + 1$-elementigen Teilmengen von $\{1, \ldots, m + 1\}$, die $m + 1$ *nicht* enthalten, und $\binom{m}{n}$ ist die Anzahl der $n + 1$-elementigen Teilmengen von $\{1, \ldots, m + 1\}$, die $m + 1$ enthalten. Hat man nämlich schon das Element $m + 1$ aus $\{1, \ldots, m + 1\}$ gewählt, so bleiben noch $\binom{m}{n}$ Möglichkeiten, die übrigen n Elemente aus der verbleibenden Menge $\{1, \ldots, m\} = \{1, \ldots, m + 1\} \setminus \{m + 1\}$ zu auszuwählen.

$\square$

Im obigen Beispiel enthalten drei der sechs zweielementigen Teilmengen von $\{1, 2, 3, 4\}$ die Zahl 4 nicht, die übrigen drei Mengen enthalten die Zahl 4:

$$\binom{4}{2} = \binom{3}{2} + \binom{3}{1} = 3 + 3.$$

Die folgende Tabelle zeigt einige Werte der Binomialkoeffizienten:

Pascal-Dreieck:

$\binom{m}{n}$	$n=0$	$n=1$	$n=2$	$n=3$	$n=4$	$n=5$
$m=0$	1					
$m=1$	1	1				
$m=2$	1	2	1			
$m=3$	1	3	3	1		
$m=4$	1	4	6	4	1	
$m=5$	1	5	10	10	5	1

Es ist nützlich, diese Werte zu kennen, weil sie in den folgenden häufig verwendeten Spezialfällen

$$\begin{aligned}
(x+y)^2 &= 1x^2 + 2xy + 1y^2 \\
(x+y)^3 &= 1x^3 + 3x^2y + 3xy^2 + 1y^3 \\
(x+y)^4 &= 1x^4 + 4x^3y + 6x^2y^2 + 4xy^3 + 1y^4 \\
(x+y)^5 &= 1x^5 + 5x^4y + 10x^3y^2 + 10x^2y^3 + 5xy^4 + 1y^5
\end{aligned}$$

der binomischen Formel (21), siehe unten, vorkommen.

Weitere Beispiele zur vollständigen Induktion:

a) **geometrische Summe:** Es gilt für alle $x \in \mathbb{R} \setminus \{1\}, n \in \mathbb{N}_0$:[9]

wichtigste Summenformel der gesamten Mathematik!

$$\boxed{\sum_{k=0}^{n-1} x^k = \frac{x^n - 1}{x - 1}} \tag{20}$$

Beweis: (Induktion über n):

$\boxed{n = 0}$

$$\sum_{k=0}^{-1} x^k \quad = \quad 0 = \frac{x^0 - 1}{x - 1}$$

[9]Die geometrische Summenformel gilt mit identischem Beweis ebenso für alle *komplexen* Zahlen $x \neq 1$, die wir allerdings erst im übernächsten Abschnitt einführen. Analoges gilt für die binomische Formel, die unten behandelt wird.

$\boxed{n \rightsquigarrow n+1}$ Es gelte die Behauptung (20) für ein $n \in \mathbb{N}_0$.

$$\begin{aligned}
\sum_{k=0}^{n} x^k &= x^n + \sum_{k=0}^{n-1} x^k \\
&= x^n + \frac{x^n - 1}{x - 1} \text{ nach der Induktionsvoraussetzung} \\
&= \frac{(x^{n+1} - x^n) + (x^n - 1)}{x - 1} \\
&= \frac{x^{n+1} - 1}{x - 1}
\end{aligned}$$

$\square$

b) **binomische Formel** Es gilt für $x, y \in \mathbb{R}, m \in \mathbb{N}_0$: *sehr wichtig!*

$$\boxed{(x+y)^m = \sum_{k=0}^{m} \binom{m}{k} x^k y^{m-k}.} \tag{21}$$

Beweis: (Induktion über m):

$\boxed{m = 0}$

$$(x+y)^0 = 1 = \binom{m}{0} x^0 y^0.$$

$\boxed{m \rightsquigarrow m+1}$ Es gelte die Aussage (21) für ein $m \in \mathbb{N}_0$.

$$\begin{aligned}
(x+y)^{m+1} &= (x+y)(x+y)^m \\
&= (x+y) \sum_{k=0}^{m} \binom{m}{k} x^k y^{m-k} \text{ nach der Induktionsvoraussetzung} \\
&= x \sum_{k=0}^{m} \binom{m}{k} x^k y^{m-k} + y \sum_{k=0}^{m} \binom{m}{k} x^k y^{m-k} \\
&= \sum_{k=0}^{m} \binom{m}{k} x^{k+1} y^{m-k} + \sum_{k=0}^{m} \binom{m}{k} x^k y^{(m+1)-k}. \qquad (22)
\end{aligned}$$

Wir führen nun in der linken Summe in der letzten Zeile einen neuen Summationsindex $l := k+1$ ein. Dieser läuft von 1 bis $m+1$, da $k = l-1$ von 0 bis m läuft. Anschließend benennen wir den neuen Summationsindex l in k um und spalten dann den Summanden mit Index $k = m + 1$ ab:

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

wobei im letzten Schritt $\binom{m}{m} = 1 = \binom{m+1}{m+1}$ verwendet wurde. In der rechten Summe in der letzten Zeile in (22) spalten wir den Summanden mit Index $k = 0$ ab:

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

Wir verwendeten hier $\binom{m}{0} = 1 = \binom{m+1}{0}$. Eingesetzt in (22) folgt die Induktionsbehauptung so:

$$\begin{aligned}
(x+y)^{m+1} &= \sum_{k=0}^{m} \binom{m}{k} x^{k+1} y^{m-k} + \sum_{k=0}^{m} \binom{m}{k} x^k y^{(m+1)-k} \\
&= \binom{m+1}{m+1} x^{m+1} y^0 + \sum_{k=1}^{m} \binom{m}{k-1} x^k y^{(m+1)-k} \\
&\quad + \sum_{k=1}^{m} \binom{m}{k} x^k y^{(m+1)-k} + \binom{m+1}{0} x^0 y^{m+1} \\
&= \binom{m+1}{m+1} x^{m+1} y^0 + \sum_{k=1}^{m} \left[\binom{m}{k-1} + \binom{m}{k} \right] x^k y^{(m+1)-k} + \binom{m+1}{0} x^0 y^{m+1} \\
&= \binom{m+1}{m+1} x^{m+1} y^0 + \sum_{k=1}^{m} \binom{m+1}{k} x^k y^{(m+1)-k} + \binom{m+1}{0} x^0 y^{m+1} \\
&= \sum_{k=0}^{m+1} \binom{m+1}{k} x^k y^{(m+1)-k},
\end{aligned}$$

wobei wir $\binom{m}{k-1} + \binom{m}{k} = \binom{m+1}{k}$ für $k = 1, \ldots, m$ verwendet haben.

$\square$

Potenz-Exponentialsummen. Als eine etwas komplexere Anwendung der Variante des Induktionsschemas und der binomischen Formel analysieren wir nun eine Rekursionsformel für Potenz-Exponentialsummen:
Wir definieren rekursiv die Funktion $f : \mathbb{N}_0 \times \mathbb{N}_0 \times (\mathbb{R} \setminus \{1\}) \to \mathbb{R}$ durch

$$f(m, 0, x) := \frac{x^{m+1} - 1}{x - 1}, \tag{23}$$

$$f(m, n, x) := \frac{1}{x - 1} \left[(m+1)^n x^{m+1} - x \sum_{l=0}^{n-1} \binom{n}{l} f(m, l, x) \right] \quad \text{für } n \in \mathbb{N}. \tag{24}$$

Dann gilt für alle $m, n \in \mathbb{N}_0$ und $x \in \mathbb{R} \setminus \{1\}$:

$$\sum_{k=0}^{m} k^n x^k = f(m, n, x). \tag{25}$$

Beweis: Es seien $m \in \mathbb{N}_0$ und $x \in \mathbb{R} \setminus \{1\}$ gegeben. Wir zeigen die folgende Aussage mit Hilfe der Variante des Induktionsschemas:

$$\forall n \in \mathbb{N}_0 : \sum_{k=0}^{m} k^n x^k = f(m, n, x). \tag{26}$$

Hierzu ist zu zeigen:

$$\forall n \in \mathbb{N}_0 : \left[\left(\forall l \in \mathbb{N}_0 : \left(l < n \Rightarrow \sum_{k=0}^{m} k^l x^k = f(m, l, x) \right) \right) \Rightarrow \sum_{k=0}^{m} k^n x^k = f(m, n, x) \right].$$

Zum Beweis davon nehmen wir folgende Induktionsvoraussetzung an:
Es sei $n \in \mathbb{N}_0$ gegeben, und es gelte

$$\forall l \in \mathbb{N}_0 : \left(l < n \Rightarrow \sum_{k=0}^{m} k^l x^k = f(m, l, x) \right). \tag{27}$$

Zu zeigen ist $\sum_{k=0}^{m} k^n x^k = f(m, n, x)$, also die Formel (25). Wir unterscheiden zwei Fälle:

1. Fall, "Induktionsanfang", $n = 0$. Hier ist

$$\begin{aligned}
\sum_{k=0}^{m} k^n x^k &= \sum_{k=0}^{m} x^k \quad \text{(wegen } n = 0\text{)} \\
&= \frac{x^{m+1} - 1}{x - 1} \quad \text{(geometrische Summe)} \\
&= f(m, 0, x) \quad \text{(wegen (23))} \\
&= f(m, n, x) \quad \text{(wegen } n = 0\text{).}
\end{aligned}$$

2. Fall, “Induktionsschritt”, $n \neq 0$. Es gilt also in diesem Fall $n \in \mathbb{N}$. Hier erhalten wir:

$$
\begin{aligned}
f(m,n,x) &= \frac{1}{x-1}\left[(m+1)^n x^{m+1} - x\sum_{l=0}^{n-1}\binom{n}{l} f(m,l,x)\right] \quad \text{(mit dem Rekursionsschritt (24))}\\
&= \frac{1}{x-1}\left[(m+1)^n x^{m+1} - x\sum_{l=0}^{n-1}\binom{n}{l}\sum_{k=0}^{m} k^l x^k\right] \quad \text{(wegen der Induktionsvoraussetzung (27))}\\
&= \frac{1}{x-1}\left[(m+1)^n x^{m+1} - \sum_{k=0}^{m}\left(\sum_{l=0}^{n-1}\binom{n}{l}k^l\right) x^{k+1}\right] \quad \begin{array}{l}\text{(mit Vertauschung}\\ \text{der Summationsreihenfolge}\\ \text{und Ausklammern)}\end{array}
\end{aligned} \tag{28}
$$

Nun gilt für $k \in \{0,\dots,m\}$ mit Hilfe von $\binom{n}{n} = 1$ und der binomischen Formel:

$$
\sum_{l=0}^{n-1}\binom{n}{l}k^l = \sum_{l=0}^{n}\binom{n}{l}k^l - \binom{n}{n}k^n = \sum_{l=0}^{n}\binom{n}{l}k^l - k^n = \sum_{l=0}^{n}\binom{n}{l}k^l 1^{n-l} - k^n = (k+1)^n - k^n.
$$

In Formel (28) eingesetzt:

$$
\begin{aligned}
f(m,n,x) &= \frac{1}{x-1}\left[(m+1)^n x^{m+1} - \sum_{k=0}^{m}((k+1)^n - k^n)x^{k+1}\right]\\
&= \frac{1}{x-1}\left[(m+1)^n x^{m+1} - \sum_{k=0}^{m}(k+1)^n x^{k+1} + \sum_{k=0}^{m} k^n x^{k+1}\right]\\
&= \frac{1}{x-1}\left[-\sum_{k=0}^{m-1}(k+1)^n x^{k+1} + \sum_{k=0}^{m} k^n x^{k+1}\right] \quad \begin{array}{l}\text{(Summand zu } k=m\\ \text{in linker Summe weggeho-}\\ \text{ben)}\end{array}\\
&= \frac{1}{x-1}\left[x\sum_{k=0}^{m} k^n x^k - \sum_{k=0}^{m-1}(k+1)^n x^{k+1}\right]
\end{aligned} \tag{29}
$$

In der rechten Summe der letzten Zeile in (29) führen wir einen neuen Summationsindex $j := k+1$ ein. Dieser läuft über $j = 1,\dots m$, wenn $k = j-1$ über $0,\dots,m-1$ läuft. Anschließend benennen wir den Summationsindex j wieder in k um:

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

Zum letzten Gleichheitszeichen bemerken wir, dass der Summand zu $k = 0$ in der letzten Summe nichts beiträgt, da $0^n = 0$ wegen $n \neq 0$ gilt.
Eingesetzt in (29) folgt die Behauptung (25) so:

$$
f(m,n,x) = \frac{1}{x-1}\left[x\sum_{k=0}^{m} k^n x^k - \sum_{k=0}^{m} k^n x^k\right] = \frac{1}{x-1}[x-1]\sum_{k=0}^{m} k^n x^k = \sum_{k=0}^{m} k^n x^k.
$$

Damit ist die Behauptung (25) in beiden Fällen gezeigt.

1.3.2 Reelle Zahlen

Wir führen die reellen Zahlen hier axiomatisch ein. Erst später besprechen wir kurz ihre Konstruktion in allgemeinerem Rahmen.
Die Axiome für die reellen Zahlen bestehen aus vier Teilen: den Körperaxiomen, den Anordnungsaxiomen, dem Archimedischen Axiom und dem Vollständigkeitsaxiom.

Körperaxiome

Es sind zwei Operationen gegeben, die Addition $+ : \mathbb{R}\times\mathbb{R} \to \mathbb{R}$ *und die Multiplikation* $\cdot : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$*, mit denen* $\mathbb{R}$ *ein Körper ist.*

Das heißt:

a) $(\mathbb{R}, +)$ ist eine abelsche Gruppe mit neutralem Element 0,

b) $(\mathbb{R} \setminus \{0\}, \cdot)$ ist eine abelsche Gruppe mit neutralem Element 1,

c) Es gilt das Distributivgesetz:

$$\forall a, b, c \in \mathbb{R} : \ a \cdot (b + c) = a \cdot b + a \cdot c, \ (b + c) \cdot a = b \cdot a + c \cdot a$$

Zur Erinnerung: In der Linearen Algebra wird der Begriff der abelschen Gruppe behandelt: Ein Paar $(G, *)$, bestehend aus einer Menge G und einer Abbildung $* : G \times G \to G$, $(a, b) \mapsto a * b$, zusammen mit einem Element $I \in G$ heißt *abelsche Gruppe mit neutralem Element* I, wenn gilt:

1. Die Operation $*$ ist assoziativ, d.h. $\forall a, b, c \in G : \ (a * b) * c = a * (b * c)$.
2. Die Operation $*$ ist kommutativ, d.h. $\forall a, b \in G : \ a * b = b * a$.
3. I ist (links-)neutral bezüglich $*$, d.h. $\forall a \in G : I * a = a$.
4. Jedes Element von G besitzt ein (Links-)Inverses bezüglich $*$ und I, d.h. $\forall a \in G \ \exists b \in G : \ b*a = I$.

Das neutrale Element und die Inverse sind dann eindeutig bestimmt. Für $x \in \mathbb{R}$ bezeichnet $-x$ die Inverse von x in $(\mathbb{R}, +)$. Weiter steht $x - y$ für $x + (-y)$, wobei $x, y \in \mathbb{R}$. Ebenso bezeichnet x^{-1} die multiplikative Inverse von $x \in \mathbb{R} \setminus \{0\}$, und $\frac{x}{y} = x/y$ steht für $x \cdot y^{-1}$, wobei $x \in \mathbb{R}$ und $y \in \mathbb{R} \setminus \{0\}$.

Insbesondere gilt $0 \cdot a = 0$ für alle $a \in \mathbb{R}$ wegen $0 \cdot a + 0 = 0 \cdot a = (0 + 0) \cdot a = 0 \cdot a + 0 \cdot a$, sowie ebenso $a \cdot 0 = 0$.

Anordnungsaxiome

Auf $\mathbb{R}$ ist eine einstellige Relation "positiv" definiert, also eine Teilmenge $\mathbb{R}^+$ von $\mathbb{R}$, so daß gilt:

a) Für jedes $x \in \mathbb{R}$ gilt *genau eine* der Relationen $x \in \mathbb{R}^+, x = 0$ oder $-x \in \mathbb{R}^+$.

b) $\forall x, y \in \mathbb{R}$: $(x \in \mathbb{R}^+ \wedge y \in \mathbb{R}^+ \Rightarrow x + y \in \mathbb{R}^+ \wedge x \cdot y \in \mathbb{R}^+)$

Definition: Es seien $x, y \in \mathbb{R}$. Die Aussage $x > y$ (synonym $y < x$) bedeutet: $x - y \in \mathbb{R}^+$. Die Aussage $x \geq y$ (synonym $y \leq x$) bedeutet: $x > y$ oder $x = y$. Wir setzen $\mathbb{R}^- := \{-x |\ x \in \mathbb{R}^+\}$. Die Elemente von $\mathbb{R}^+$ heißen *positive Zahlen*, die Elemente von $\mathbb{R}^-$ *negative Zahlen*.

Mit dieser Definition wird $x \in \mathbb{R}^+$ gleichbedeutend mit $x > 0$, und $x \in \mathbb{R}^-$ gleichbedeutend mit $x < 0$.

Wir können $\mathbb{N}_0 \subset \mathbb{R}$ wiederfinden als die kleinste Menge, die 0 enthält und mit jedem x auch $N(x) = x + 1$ enthält. Wir setzen $\mathbb{N} = \mathbb{N}_0 \setminus \{0\}$.

Archimedisches Axiom

$$\forall x \in \mathbb{R}\ \exists n \in \mathbb{N}_0 :\ n > x.$$

In Worten:

Jede reelle Zahl wird von mindestens einer natürlichen Zahl übertroffen.

Wir verwenden auch häufig die folgende äquivalente Formulierung des Archimedischen Axioms:

Archimedisches Axiom – Variante

$$\forall \varepsilon > 0\ \exists n \in \mathbb{N} :\ \frac{1}{n} < \varepsilon$$

In Worten:

Zu jeder positiven reellen Zahl $\varepsilon > 0$ *gibt es einen kleineren Stammbruch* $1/n$, $n \in \mathbb{N}$.

Die Variante des Archimedischen Axioms besagt also, dass es keine infinitesimalen reellen Zahlen gibt.

Bevor wir das letzte Axiom der reellen Zahlen, das Vollständigkeitsaxiom, besprechen, hier einige Vorbereitungen:

Definition 1.1 Sei $M \subseteq \mathbb{R}$ und $x \in \mathbb{R}$. x heißt eine *obere Schranke* von M, wenn

$$\forall y \in M :\ y \leq x$$

gilt. M heißt *nach oben beschränkt*, wenn es eine obere Schranke besitzt, d.h. wenn

$$\exists x \in \mathbb{R}\ \forall y \in M :\ y \leq x$$

gilt. Andernfalls heisst M *nach oben unbeschränkt.*
Analog heißt x eine *untere Schranke* von M, wenn

$$\forall y \in M: \ y \geq x$$

gilt. M heißt *nach unten beschränkt*, wenn es eine untere Schranke besitzt, d.h. wenn

$$\exists x \in \mathbb{R}\ \forall y \in M: \ y \geq x$$

gilt. Andernfalls heisst M *nach unten unbeschränkt.*
M heißt *beschränkt*, wenn es nach oben und nach unten beschränkt ist.

Das Archimedische Axiom besagt also: $\mathbb{N}_0 \subset \mathbb{R}$ *ist nach oben unbeschränkt.*

Supremum und Infimum

Definition 1.2 Es seien $M \subseteq \mathbb{R}$ und $x \in \mathbb{R}$.
x heißt *Supremum* von M, wenn es die kleinste obere Schranke von M ist.

Mit anderen Worten: x ist Supremum von M, wenn es eine obere Schranke von M ist, und wenn für jede obere Schranke x' von M gilt: $x \leq x'$.
Analog:
x heißt *Infimum* von M, wenn es eine größte untere Schranke von M ist.

sehr wichtig!

Das Supremum ist eindeutig bestimmt, falls es existiert. Wir schreiben: $\sup M$ für das Supremum von M und $\inf M$ für das Infimum von M.

In Formeln können wir die Definitionen des Supremums und des Infimums auch so ausdrücken:

$$\begin{array}{rl} s = \sup M & \Leftrightarrow (\forall x \in M : x \leq s) \wedge \forall x \in \mathbb{R}: \ (x < s \Rightarrow \exists y \in M : y > x) \\ i = \inf M & \Leftrightarrow (\forall x \in M : x \geq i) \wedge \forall x \in \mathbb{R}: \ (x > i \Rightarrow \exists y \in M : y < x) \end{array}$$

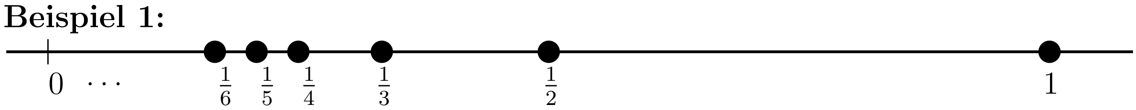
Vollständigkeitsaxiom

Jede nichtleere, nach oben beschränkte Teilmenge von $\mathbb{R}$ besitzt ein Supremum.

fundamental!

Äquivalente Formulierung: *Jede nichtleere, nach unten beschränkte Teilmenge von $\mathbb{R}$ besitzt ein Infimum.*

Beispiel 1:



Es sei M die Menge der Stammbrüche

$$M = \{\frac{1}{n} \mid n \in \mathbb{N}\} = \left\{1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \dots\right\}.$$

Die Menge M ist nichtleer. Weiter ist 0 eine untere Schranke von M, weil $\frac{1}{n} \geq 0$ für alle $n \in \mathbb{N}$ gilt.
0 ist sogar das Infimum von M.
Beweis: Es sei x eine untere Schranke von M. Wir müssen zeigen: $x \leq 0$.
Nehmen wir also an: $x > 0$. Nach dem Archimedischen Axiom gibt es $n \in \mathbb{N}$ mit $\frac{1}{n} < x$, was nicht gelten kann, da x eine untere Schranke von M ist. Die Annahme $x > 0$ ist also falsch, so daß $x \leq 0$ folgt.

Beispiel 2: Wenn M ein *Maximum* besitzt, d.h. ein größtes Element besitzt, also ein Element $x \in M$ mit $\forall y \in M: \ y \leq x$, dann ist x das Supremum von M.
Beweis: x ist obere Schranke von M, und für jede obere Schranke x' von M gilt:

$$\forall y \in M: \ y \leq x' \text{ also } x \leq x' \text{ wegen } x \in M.$$

x ist also die kleinste obere Schranke von M.

$\square$

Analog gilt: Wenn ein M *Minimum* besitzt, d.h. ein kleinstes Element, so ist das Minimum von M gleich dem Infimum von M.
Wir schreiben $\max M$ für das Maximum von M und $\min M$ für das Minimum von M, falls sie existieren.

Achtung: Es gibt Mengen reeller Zahlen, die zwar ein Supremum, aber kein Maximum besitzen, zum Beispiel die Menge

$$[0, 1[= \{x \in \mathbb{R} \mid 0 \leq x < 1\}.$$

Diese Menge hat das Supremum 1, aber $1 \notin [0, 1[$. Die Zahl 1 ist also *kein* Maximum von $[0, 1[$.

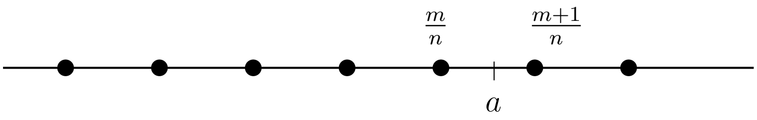
Jede nichtleere Menge $M \subseteq \mathbb{N}_0$ besitzt ein kleinstes Element. Jede nichtleere, nach oben (bzw. nach unten) beschränkte Menge $M \subseteq \mathbb{Z}$ besitzt ein größtes (bzw. kleinstes) Element. Der Beweis hiervon wird in den Übungen behandelt.

Beispiel 3: Es sei $a \in \mathbb{R}$ und $M = \{x \in \mathbb{Q} \mid x < a\}$. Dann gilt $\sup M = a$.
Beweis:

1. a ist eine obere Schranke von M, weil für alle $x \in M$ gilt: $x \leq a$,

2. Es sei b eine obere Schranke von M. Wir müssen zeigen: $a \leq b$. Nehmen wir das Gegenteil an: $b < a$. Dann ist $\varepsilon := a - b > 0$. Nach dem archimedischen Axiom gibt es $n \in \mathbb{N}$ mit $\frac{1}{n} < \varepsilon$. Wir wählen solch ein n. Die Menge $A := \{x \in \mathbb{Z} : \ x < na\}$ ist nach oben beschränkt. Aus dem archimedischen Axiom folgt, daß A nichtleer ist, denn es gibt $k \in \mathbb{N}_0$ mit $k > -na$, d.h. $-k \in A$. Also hat A ein größtes Element $m \in \mathbb{Z}$. Insbesondere gilt $m \in A$, aber $m + 1 \notin A$; das bedeutet $m < na \leq m + 1$, also $\frac{m}{n} < a$ und $a - \frac{1}{n} \leq \frac{m}{n}$. Es folgt $\frac{m}{n} \in M$, und $b = a - \varepsilon < a - \frac{1}{n} \leq \frac{m}{n}$, somit $b < \frac{m}{n} \in M$, was ein Widerspruch dazu ist, daß b eine obere Schranke von M ist. Die Annahme $b < a$ ist also falsch, und wir schließen $b \geq a$.

$\square$



Folgerung: Zwischen je zwei verschiedenen reellen Zahlen liegt mindestens eine rationale Zahl.

$$\forall a, b \in \mathbb{R} : \ (a \neq b \Rightarrow \quad \exists q \in \mathbb{Q} : \ (a < q < b \vee b < q < a)).$$

Beweis: Wir können annehmen: $a > b$; der Fall $b > a$ wird analog behandelt.
Nach dem eben Gezeigten ist b keine obere Schranke von $\{q \in \mathbb{Q} \mid q < a\}$, so daß es $q \in \mathbb{Q}$ mit $q < a$ und $b < q$ gibt.

$\square$

1.3.3 Komplexe Zahlen

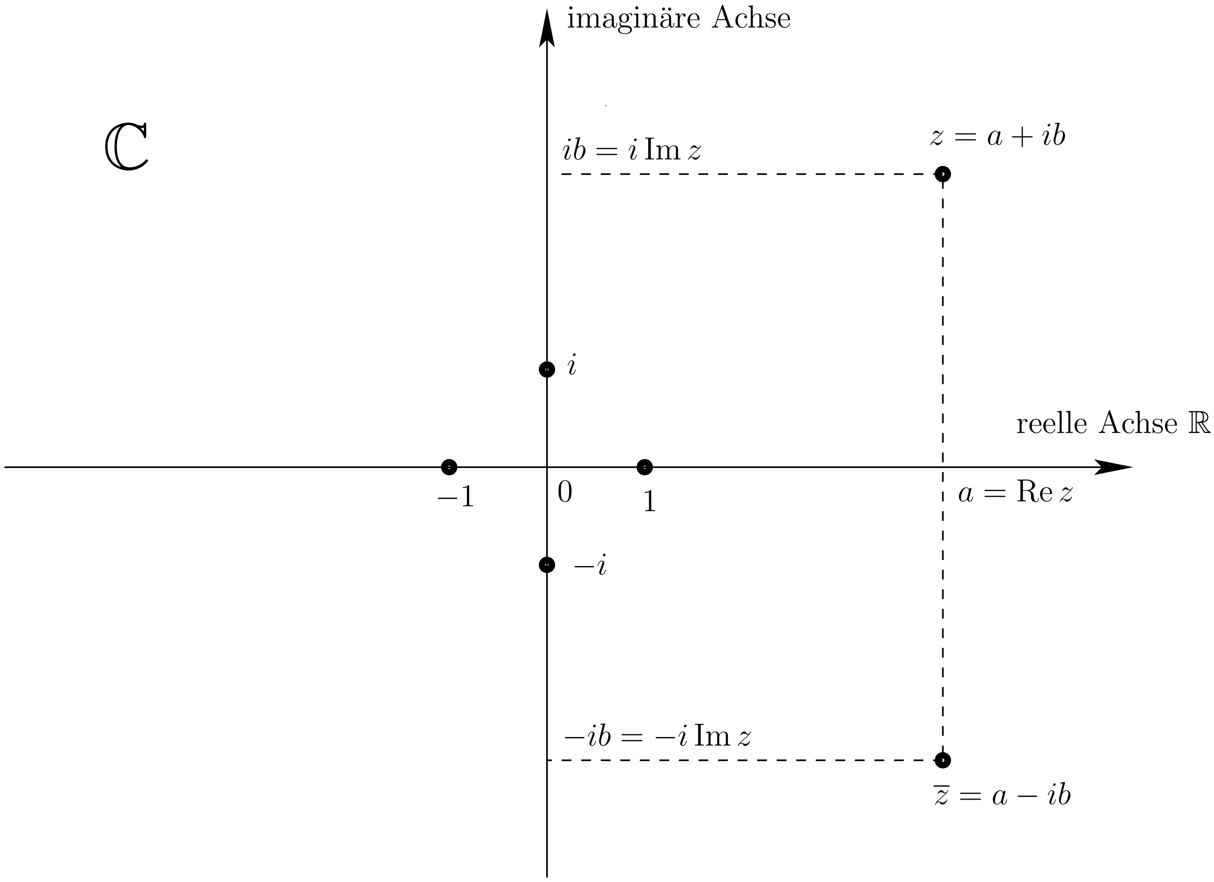
Historisch wurden komplexe Zahlen aus dem Wunsch eingeführt, auch Quadratwurzeln von negativen Zahlen ziehen zu können.

Definition 1.3 Wir definieren die Menge der komplexen Zahlen

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

Wir identifizieren reelle Zahlen $a \in \mathbb{R}$ mit den komplexen Zahlen $(a, 0) \in \mathbb{C}$ und schreiben deshalb: $\mathbb{R} \subset \mathbb{C}$. Die komplexe Zahl $i := (0, 1)$ heißt *imaginäre Einheit*. Komplexe Zahlen der Form $(0, b)$ heißen *imaginär*. Für eine komplexe Zahl $z = (a, b) \in \mathbb{C}$ heißt a der *Realteil* von z, $a = \operatorname{Re} z$, und b der *Imaginärteil* von z, $b = \operatorname{Im} z$. Die Zahl $\overline{z} := (a, -b)$ heißt *komplex Konjugierte* von z. Wir schreiben normalerweise $z = a + ib$ statt $z = (a, b)$.

Veranschaulichung: “Gaußsche Zahlenebene”



Addition und Multiplikation komplexer Zahlen werden wie folgt definiert:

$$(a,b) + (c,d) = (a+c, b+d), \tag{30}$$

$$(a,b) \cdot (c,d) = (ac - bd, ad + bc). \tag{31}$$

Insbesondere gilt:

$$\boxed{i^2 = -1}$$

Begründung:

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

wobei wir im letzten Schritt die Identifikation von reellen Zahlen mit komplexen Zahlen mit Imaginärteil 0 verwendet haben.
Mit diesen Operationen $+, \cdot$ wird $\mathbb{C}$ ein *Körper*.
Mit der Schreibweise $a + ib$ statt (a,b) wird die Multiplikationsregel motiviert:

$$(a+ib) \cdot (c+id) = ac + i^2 bd + iad + ibc = (ac - bd) + i(ad + bc).$$

Die multiplikative Inverse $\frac{1}{z}$ einer komplexen Zahl $z = a + ib \neq 0$ ist

$$\frac{1}{z} = \frac{a}{a^2+b^2} - i\frac{b}{a^2+b^2},$$

denn

$$\begin{aligned}
&\left(\frac{a}{a^2+b^2} - i\frac{b}{a^2+b^2}\right) \cdot (a+ib) \\
&= \frac{a}{a^2+b^2} \cdot a - i^2 \frac{b}{a^2+b^2} \cdot b - i\frac{b}{a^2+b^2} \cdot a + i\frac{a}{a^2+b^2} b \\
&= 1.
\end{aligned}$$

Der Quotient zweier komplexer Zahlen $a+ib$ und $c+id \neq 0$ lautet

$$\frac{a+ib}{c+id} = \frac{(a+ib)(c-id)}{(c+id)(c-id)} = \frac{ac+bd}{c^2+d^2} + i\frac{bc-ad}{c^2+d^2}.$$

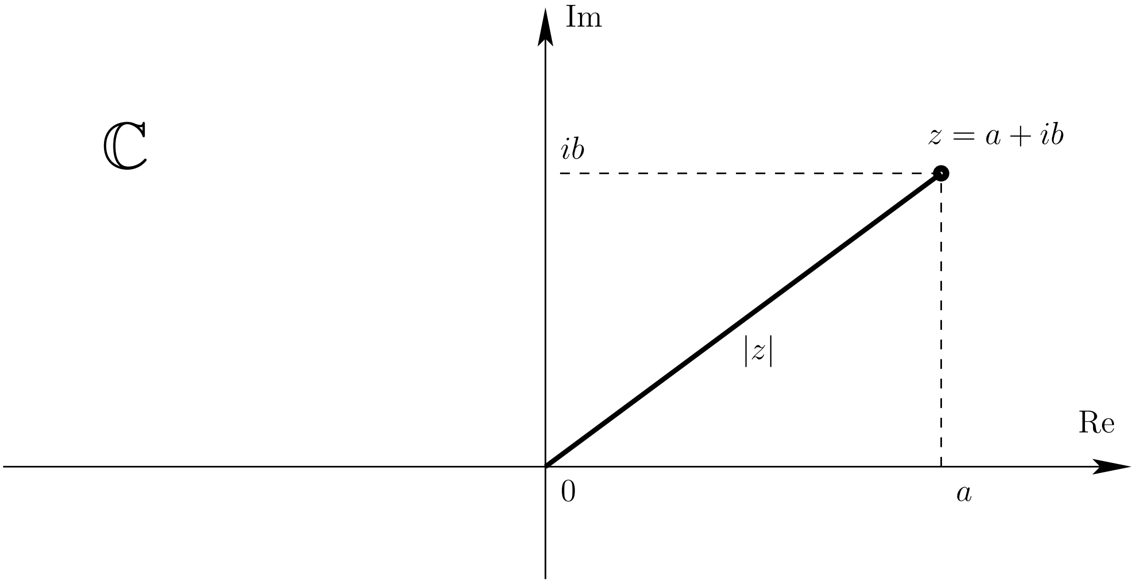
Der Quotient wird also durch Erweitern mit dem konjugiert Komplexen des Nenners berechnet.

Rechenregeln für komplexe Zahlen: Es seien $z, w \in \mathbb{C}$. Dann gilt:

1. $\overline{zw} = \overline{z}\,\overline{w}$
2. $\overline{\left(\frac{z}{w}\right)} = \frac{\overline{z}}{\overline{w}}$, falls $w \neq 0$
3. $\overline{z \pm w} = \overline{z} \pm \overline{w}$
4. $\overline{\overline{z}} = z$
5. $z + \overline{z} = 2\,\mathrm{Re}\,z$
6. $z - \overline{z} = 2i\,\mathrm{Im}\,z$

Der Absolutbetrag $|z| \in \mathbb{R}$ einer komplexen Zahl $z = a+ib \in \mathbb{C}$, $(a, b \in \mathbb{R})$ ist wie folgt definiert:

$$\boxed{|z| = \sqrt{a^2+b^2} = \sqrt{z\overline{z}}}$$



Damit kann man die Regel zur Quotientenbildung auch anders schreiben:

$$\frac{z}{w} = \frac{z\overline{w}}{w\overline{w}} = \frac{1}{|w|^2} z\overline{w}$$

für $z \in \mathbb{C}$ und $w \in \mathbb{C} \setminus \{0\}$.

Rechenregeln für den Absolutbetrag: Für alle $z, w \in \mathbb{C}$ gilt:

$$\boxed{\begin{aligned} &|zw| = |z||w|, \\ &\left|\frac{z}{w}\right| = \frac{|z|}{|w|} \quad \text{falls } w \neq 0 \end{aligned}}$$

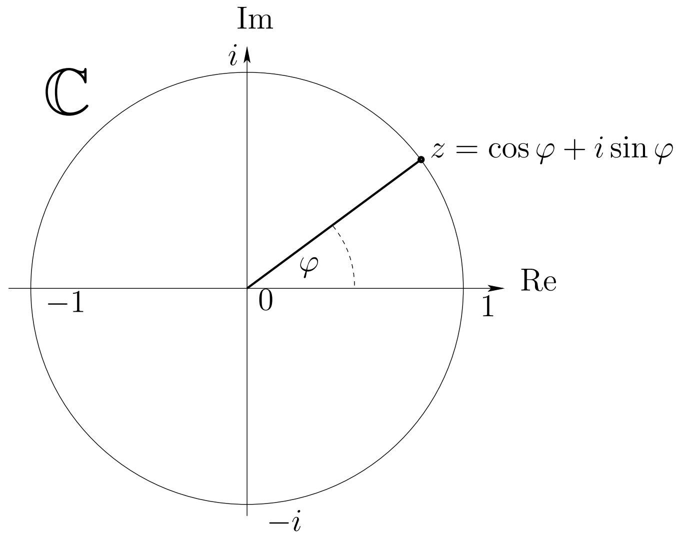
denn

$$|zw| = \sqrt{zw\overline{zw}} = \sqrt{z\overline{z}}\sqrt{w\overline{w}} = |z||w|.$$

Polarkoordinaten komplexer Zahlen. Jede komplexe Zahl z vom Betrag $|z| = 1$ läßt sich in der Form

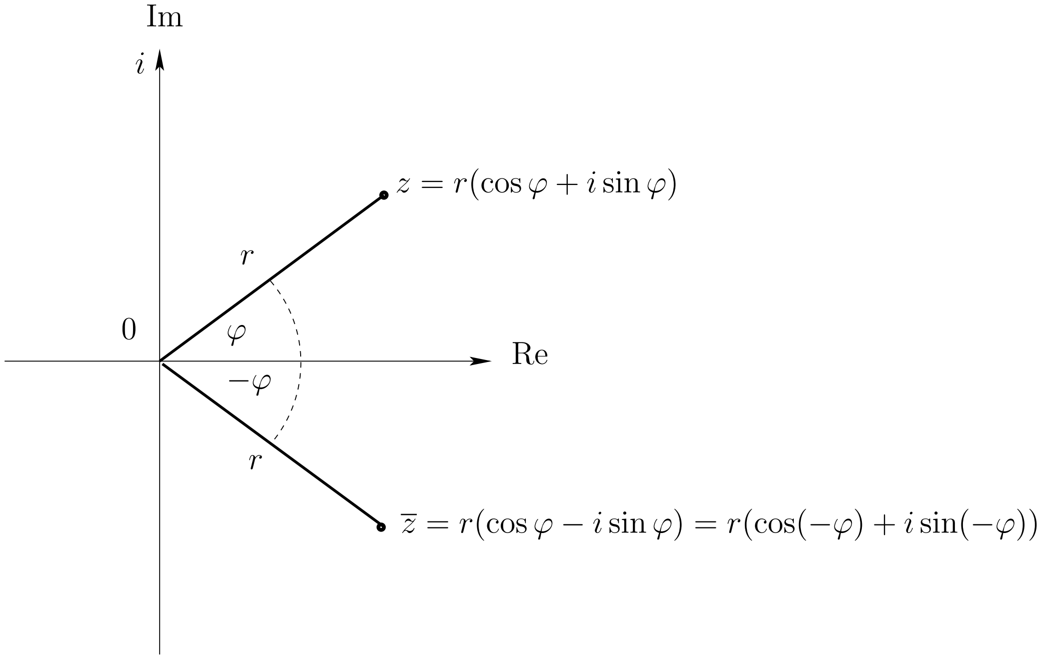
$$z = \cos\varphi + i \sin\varphi$$

schreiben, wobei $\varphi \in \mathbb{R}$ nur bis auf ganzzahlige Vielfache von 2π bestimmt ist.



Jede komplexe Zahl hat eine Polardarstellung

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



In dieser Darstellung wird die Multiplikation und Division komplexer Zahlen besonders einfach: Für komplexe Zahlen in Polarkoordinatendarstellung

$$\begin{aligned} z &= r(\cos\varphi + i\sin\varphi) \\ w &= s(\cos\psi + i\sin\psi) \end{aligned}$$

erhalten wir mit Hilfe der Additionstheoreme von Sinus und Kosinus:

$$\begin{aligned}
z \cdot w &= rs(\cos\varphi\cos\psi - \sin\varphi\sin\psi + i(\cos\varphi\sin\psi + \sin\varphi\cos\psi)) \\
&= rs(\cos(\varphi+\psi) + i\sin(\varphi+\psi)) \\
\frac{z}{w} &= \frac{r}{s}(\cos\varphi\cos\psi + \sin\varphi\sin\psi + i(\sin\varphi\cos\psi - \cos\varphi\sin\psi)) \\
&= \frac{r}{s}(\cos(\varphi-\psi) + i\sin(\varphi-\psi))
\end{aligned}$$

wobei wir bei der Division $w \neq 0$ annehmen. Die Multiplikation mit w ist also eine *Drehstreckung* um 0 mit Streckungsfaktor $|w| = s$ und Drehwinkel ψ.

Die folgende Ungleichung ist ein Spezialfall einer gleichnamigen Ungleichung, die in der linearen Algebra behandelt wird:

Lemma 1.4 (Cauchy-Schwarz Ungleichung für komplexe Zahlen) *Für alle $z, w \in \mathbb{C}$ gilt:*

$$\boxed{|z||w| \geq |\operatorname{Re}(z\overline{w})|}$$

Beweis:

$$\begin{aligned}
(|z||w|)^2 &= (|z||\overline{w}|)^2 \\
&= |z\overline{w}|^2 \\
&= |\operatorname{Re}(z\overline{w})|^2 + |\operatorname{Im}(z\overline{w})|^2 \\
&\geq |\operatorname{Re}(z\overline{w})|^2.
\end{aligned}$$

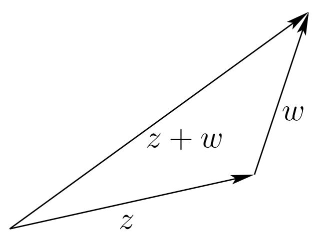
Hieraus folgt die Behauptung, da $|z||w| \geq 0$.

□

Folgerung: (Dreiecksungleichung für komplexe Zahlen) *extrem wichtig!*
Für alle $z, w \in \mathbb{C}$ gilt:

$$\boxed{|z+w| \leq |z| + |w|}$$

Illustration:



Beweis:

$$\begin{aligned}
|z+w|^2 &= (z+w)\overline{(z+w)} \\
&= z\overline{z} + z\overline{w} + \overline{z}w + w\overline{w} \\
&= |z|^2 + 2\operatorname{Re}(z\overline{w}) + |w|^2 \\
&\leq |z|^2 + 2|z||w| + |w|^2 \\
&= (|z|+|w|)^2,
\end{aligned}$$

wobei wir die Cauchy-Schwarz Ungleichung verwendeten. Hieraus folgt die Behauptung wegen $|z| + |w| \geq 0$.

$\square$

Aus der Dreiecksungleichung folgt einfach die folgende Variante:

Dreiecksungleichung – Variante: Für $z, w \in \mathbb{C}$ gilt:

$$\big||z| - |w|\big| \leq |z - w|$$

Warnung: Die Anordnung $\leq$ auf den reellen Zahlen wird *nicht* auf die komplexen Zahlen fortgesetzt. Die komplexen Zahlen bilden *keinen* angeordneten Körper.

1.3.4 Unendlich ferne Punkte

Erweitert reelle Zahlen. Wir fügen zwei "unendlich ferne" Punkte $+\infty$ und $-\infty$ zu $\mathbb{R}$ hinzu und erhalten die erweitert reellen Zahlen $\mathbb{R} \cup \{\pm\infty\}$
Die Anordnung $\leq$ wird wie folgt fortgesetzt: Für alle $x \in \mathbb{R} \cup \{\pm\infty\}$ definiert man

$$-\infty \leq x \leq +\infty.$$

Mit dieser Erweiterung können wir das Vollständigkeitsaxiom einfacher formulieren.

Vollständigkeitsaxiom – alternative Fassung

Jede Teilmenge von $\mathbb{R} \cup \{\pm\infty\}$ besitzt ein Supremum und ein Infimum in $\mathbb{R} \cup \{\pm\infty\}$.

Zum Beispiel ist $\sup \emptyset = -\infty$, $\inf \emptyset = +\infty$ und falls $M \subseteq \mathbb{R}$ nach oben unbeschränkt ist gilt $\sup M = +\infty$, insbesondere $\sup \mathbb{R} = \sup \mathbb{N} = +\infty$.

Warnung: Manchmal – wenn keine Missverständnisse zu befürchten sind – schreibt man einfach ∞ statt $+\infty$. Dies darf nicht mit dem *komplex* unendlich fernen Punkt verwechselt werden, den wir gleich einführen. In Zweifelsfällen sollte man besser $+\infty$ statt ∞ schreiben.

Der komplex unendlich ferne Punkt ∞. Üblicherweise wird $\mathbb{C}$ nur um einen einzigen "unendlichen" Punkt ∞ erweitert, den man sich "weit weg" (aber nicht in einer bestimmten Richtung) vorstellen soll.

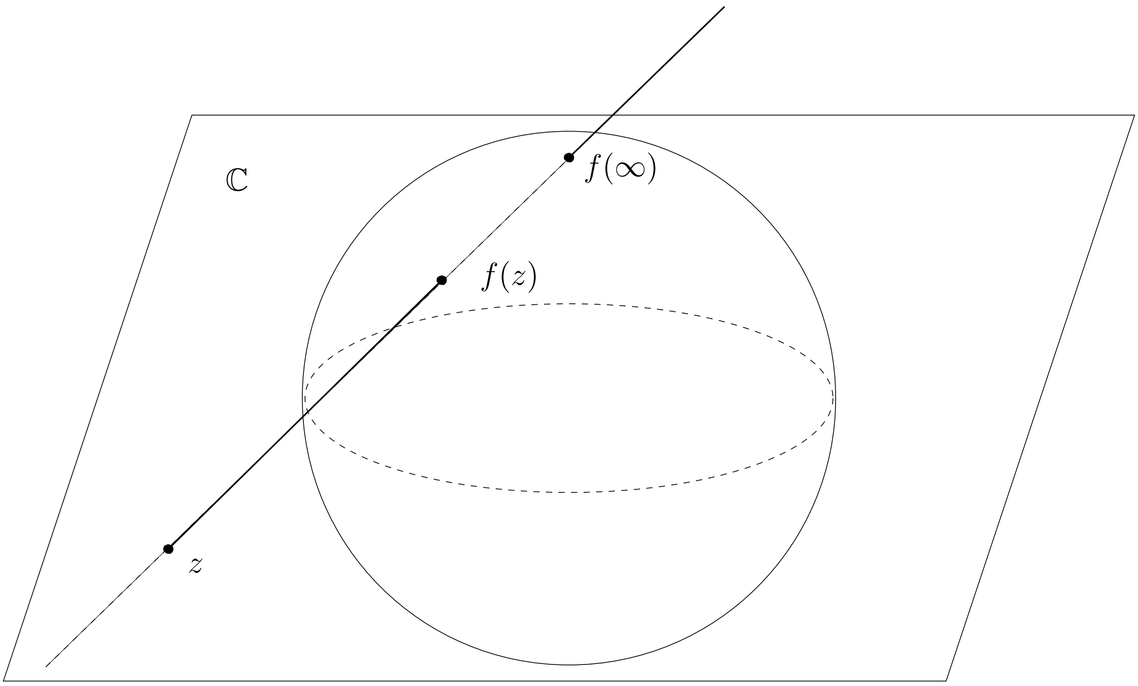
Einfacher kann man sich den unendlich fernen Punkt ∞ vorstellen, indem man die Zahlenebene zur "Riemannschen Zahlenkugel" zusammenbiegt. Dies leistet die "stereographische Projektion", die wie folgt definiert wird:

$$\begin{array}{rcl} f : \mathbb{C} \cup \{\infty\} & \longrightarrow & S^2 = \{(x,y,z) \in \mathbb{R}^3 \mid x^2+y^2+z^2=1\}, \\ a+ib & \mapsto & \text{von } (0,0,1) \text{ verschiedener Schnittpunkt} \\ & & \text{der Geraden durch } (a,b,0) \text{ und } (0,0,1) \text{ mit } S^2 \\ \infty & \mapsto & (0,0,1) \end{array}$$

In Formeln:

$$\begin{array}{rcl} f(z) & = & \left(\dfrac{2\,\mathrm{Re}\, z}{|z|^2+1}, \dfrac{2\,\mathrm{Im}\, z}{|z|^2+1}, \dfrac{|z|^2-1}{|z|^2+1} \right) \\ f(\infty) & = & (0,0,1). \end{array}$$

Veranschaulichung in einem 3-dimensionalen Bild



2 Topologische Grundbegriffe

Topologie ist ein Nachbargebiet der Geometrie, in dem unter anderem der anschauliche Begriff "nahe benachbart sein" durch ein genaues mathematische Konzept präzisiert wird.

2.1 Topologie von $\mathbb{R}$ und $\mathbb{C}$

Wir quantifizieren den intuitiven Begriff "nahe benachbart zu x sein" durch "näher als eine Toleranz ε bei x liegen". Das führt uns auf den Begriff der ε-Umgebung:

Definition 2.1 *Sei $\varepsilon > 0$ und $x \in \mathbb{R}$. Wir definieren die ε-Umgebung von x durch*

$$U_\varepsilon(x) := \{y \in \mathbb{R} \mid |y - x| < \varepsilon\} =]x - \varepsilon, x + \varepsilon[$$

zentraler Grund-begriff!

Mit Hilfe des Grundbegriffs der ε-Umgebung klassifizieren wir die Lage von Punkten bezüglich einer Menge:

Definition 2.2 (topologische Klassifikation von Punkten) Es sei $M \subseteq \mathbb{R}$. Ein Punkt $x \in \mathbb{R}$ heißt

- *innerer Punkt* von M, wenn gilt: $\exists\, \varepsilon > 0 : U_\varepsilon(x) \subseteq M$.
  Anders gesagt: $\exists \varepsilon > 0\ \forall y \in \mathbb{R} : (|x - y| < \varepsilon \ \Rightarrow\ y \in M)$.
  In Worten: x ist ein innerer Punkt von M, wenn alle Punkte, die genügend nahe bei x liegen, zu M gehören.

- *äußerer Punkt* von M, wenn gilt: $\exists\, \varepsilon > 0 : U_\varepsilon(x) \cap M = \emptyset$
  Anders gesagt: $\exists \varepsilon > 0\ \forall\, y \in M :\ |x - y| \geq \varepsilon$
  In Worten: x ist ein äußerer Punkt von M, wenn genügend nahe bei x keine Punkte mehr in M liegen.

- *Berührpunkt* von M, wenn gilt: $\forall\, \varepsilon > 0 : U_\varepsilon(x) \cap M \neq \emptyset$
  Anders gesagt: $\forall \varepsilon > 0\ \exists y \in M :\ |x - y| < \varepsilon$
  In Worten: x ist ein Berührpunkt von M, wenn beliebig nahe bei x Punkte in M liegen.

- *Randpunkt* von M, wenn gilt: $\forall \varepsilon > 0 :\ (U_\varepsilon(x) \cap M \neq \emptyset\ \wedge\ U_\varepsilon(x) \setminus M \neq \emptyset)$, d.h. wenn beliebig nahe bei x sowohl Punkte in M als auch Punkte im Komplement von M liegen.

Jeder Punkt x ist entweder ein innerer Punkt von M oder ein äußerer Punkt von M oder ein Randpunkt von M. Ein Punkt x ist genau dann ein Berührpunkt von M, wenn er ein innerer Punkt oder ein Randpunkt von M ist.

Beispiel: Es seien $a, b \in \mathbb{R}$, $a < b$. Wir betrachten das Intervall $I =]a, b]$. Die Menge der inneren Punkte von I ist $]a, b[$, die Menge der Berührpunkte von I ist $[a, b]$, und die Menge der Randpunkte von I ist $\{a, b\}$.

Definition 2.3 (topologische Grundbegriffe) *Es sei* $M \subseteq \mathbb{R}$.

- Die Menge der inneren Punkte von M heißt *Inneres* M° von M.
- Die Menge der Berührpunkte von M heißt *Abschluss* $\overline{M}$ von M.
- Die Menge der Randpunkte von M heißt *Rand* ∂M von M.

Eine Teilmenge $M \subseteq U$ einer Menge $U \subseteq \mathbb{R}$ heißt *dicht* in U, wenn $\overline{M} \supseteq U$, d.h. wenn jeder Punkt in U ein Berührpunkt von M ist.

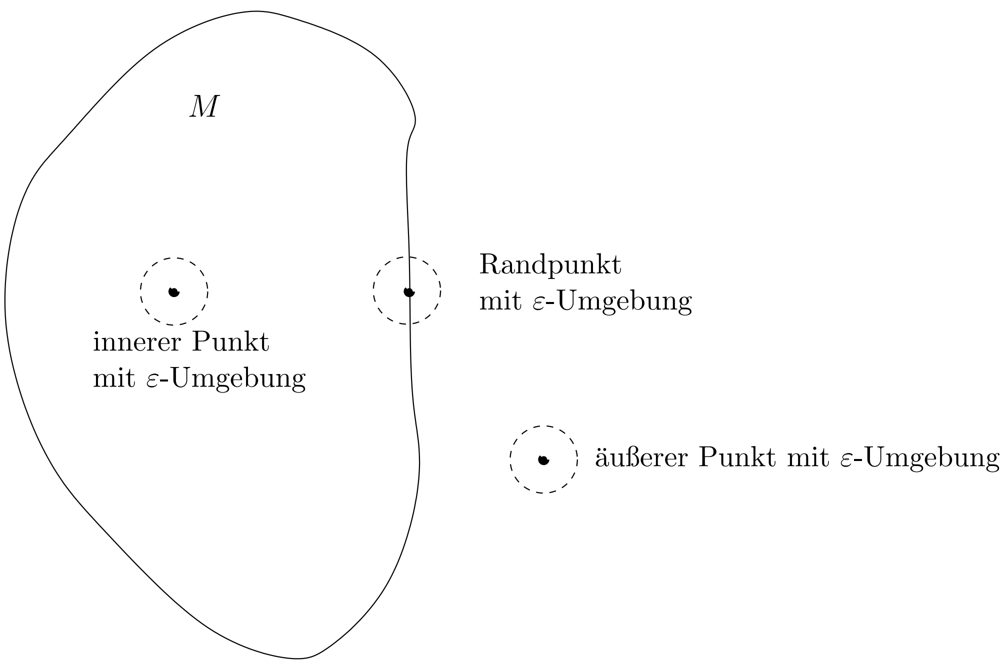
Beispiel: Das Innere der Menge der rationalen Zahlen ist leer: $\mathbb{Q}^\circ = \emptyset$, der Abschluss hingegen ist der ganze reelle Zahlenraum $\overline{\mathbb{Q}} = \mathbb{R}$, ebenso der Rand: $\partial\mathbb{Q} = \mathbb{R}$. Insbesondere ist die Menge der rationalen Zahlen $\mathbb{Q}$ dicht in der Menge $\mathbb{R}$ der reellen Zahlen.
In der Tat: Zu jedem $x \in \mathbb{R}$ gibt es keine ε-Umgebung, die nur rationale Zahlen enthält, aber jede ε-Umgebung $U_\varepsilon(x)$ enthält auch rationale Zahlen.

Nun übertragen wir den Begriff der ε-Umgebung auf komplexe Zahlen. Die Definition sieht fast genauso wie im reellen Fall aus:

Definition 2.4 *Für* $\varepsilon > 0$ *und* $x \in \mathbb{C}$ *definieren wir die* ε*-Umgebung von* x *in* $\mathbb{C}$*:*

$$U_\varepsilon^{\mathbb{C}}(x) := \{y \in \mathbb{C} \mid |y - x| < \varepsilon\}$$

Innere Punkte, Berührpunkte, etc. werden über $\mathbb{C}$ analog wie über $\mathbb{R}$ definiert; man verwendet $U_\varepsilon^{\mathbb{C}}(x)$ statt $U_\varepsilon(x)$.



Falls Missverständnisse zu befürchten sind, über welchem Raum die topologischen Begriffe “Inneres”, “Rand”, etc. gemeint sind, muss man “über $\mathbb{R}$” oder “über $\mathbb{C}$” spezifizieren. Falls jedoch keine Missverständnisse zu befürchten sind, schreiben wir $U_\varepsilon(x)$ statt $U_\varepsilon^{\mathbb{C}}(x)$.

Definition 2.5 (offene und abgeschlossene Mengen) Eine Menge $U \subseteq \mathbb{R}$ (oder $\subseteq \mathbb{C}$) heißt *offen* (über $\mathbb{R}$ bzw. über $\mathbb{C}$), wenn $U^\circ = U$. Das bedeutet: in offenen Mengen ist jeder Punkt ein innerer Punkt.

Die Menge U heißt *abgeschlossen*, wenn $\overline{U} = U$, das heißt wenn alle Berührpunkte von U Elemente von U sind.

fundamentale Begriffe!

Eine Menge $V \subseteq \mathbb{R}$ (oder $V \subseteq \mathbb{C}$) heißt eine *Umgebung* von x, wenn x ein innerer Punkt von V ist. Anders gesagt: V ist eine Umgebung von x, wenn V eine ε-Umgebung von x enthält.

Mit anderen Worten:

Eine Menge U ist offen genau dann, wenn zu jedem $x \in U$ es eine ε-Umgebung von x gibt, die in U enthalten ist:

$$\begin{aligned} U \text{ ist offen} \quad &\Leftrightarrow \quad \forall x \in U\ \exists \varepsilon > 0 :\ U_\varepsilon(x) \subseteq U. \\ &\Leftrightarrow \quad \forall x \in U\ \exists \varepsilon > 0\ \forall z \in \mathbb{R} :\ (|z - x| < \epsilon\ \Rightarrow\ z \in U) \end{aligned} \tag{32}$$

(wobei im komplexen Fall $\mathbb{R}$ durch $\mathbb{C}$ ersetzt wird). Die Menge U ist abgeschlossen genau dann, wenn jeder Punkt x, für den jede ε-Umgebung $U_\varepsilon(x)$ die Menge U trifft, zu U gehört:

$$U \text{ ist abgeschlossen} \quad \Leftrightarrow \quad \forall x\ [(\forall \varepsilon > 0 :\ U_\varepsilon(x) \cap U \neq \emptyset) \Rightarrow x \in U].$$

Eine Menge $U \subseteq \mathbb{R}$ (oder $U \subseteq \mathbb{C}$) ist eine *offene Umgebung* von einem Punkt $x \in \mathbb{R}$ (bzw. $x \in \mathbb{C}$), wenn U offen ist und x in U liegt.

Beispiel: Für zwei reelle Zahlen $a < b$ ist $]a, b[\ = \{x \in \mathbb{R} \mid a < x < b\}$ offen (in $\mathbb{R}$) und $[a, b] = \big\{x \in \mathbb{R} \big|\ a \leq x \leq b\big\}$ abgeschlossen (in $\mathbb{R}$). Die Menge $]a, b[$ ist also eine offene Umgebung all ihrer Elemente. Die Menge $[a, b[$ ist weder offen noch abgeschlossen.

Gelegentlich ist es nützlich, von offenen bzw. abgeschlossenen Mengen relativ zu einem “Teiluniversum” M zu sprechen:

Definition 2.6 (relativ offene bzw. abgeschlossene Mengen) Es sei $M \subseteq \mathbb{C}$. Eine Menge $U \subseteq M$ heißt *offen in* M, wenn es eine offene Menge $V \subseteq \mathbb{C}$ mit $U = V \cap M$gibt.
Analog heißt $U \subseteq M$ *abgeschlossen in* M, wenn es eine abgeschlossene Menge $V \subseteq \mathbb{C}$ mit $U = V \cap M$ gibt.

Zum Beispiel ist eine Menge $M \subseteq \mathbb{R}$ offen in $\mathbb{R}$ im Sinn von Definition 2.6 genau dann, wenn sie offen über $\mathbb{R}$ im Sinn von Definition 2.5 ist.

Satz 2.7 *Es sei $x \in \mathbb{R}$, $M \subseteq \mathbb{R}$. Dann gilt: Der Punkt x ist genau dann ein innerer Punkt von M, wenn er ein äußerer Punkt des Komplements von M ist.*
In Formeln:

$$x \in M^\circ \iff x \notin \overline{\mathbb{R}\backslash M}$$

Beweis. Es gilt:

$$\begin{aligned} x \in M^\circ &\Leftrightarrow \exists\, \varepsilon > 0 : U_\varepsilon(x) \subseteq M \\ &\Leftrightarrow \exists\, \varepsilon > 0 : U_\varepsilon(x) \cap (\mathbb{R}\backslash M) = \emptyset \\ &\Leftrightarrow \neg\, \forall\, \varepsilon > 0 : U_\varepsilon(x) \cap (\mathbb{R}\backslash M) \neq \emptyset \\ &\Leftrightarrow x \notin \overline{\mathbb{R}\backslash M}. \end{aligned}$$

□

Folgerung: *Eine Menge $M \subseteq \mathbb{R}$ ist genau dann offen, wenn ihr Komplement $\mathbb{R} \setminus M$ abgeschlossen ist. Analoges gilt über $\mathbb{C}$.*
Beweis: Es gelten die folgenden Äquivalenzen:

$$\begin{aligned} M \text{ ist offen} &\Leftrightarrow M = M^\circ \\ &\Leftrightarrow M = \mathbb{R} \setminus \overline{\mathbb{R} \setminus M} \\ &\Leftrightarrow \mathbb{R} \setminus M = \overline{\mathbb{R} \setminus M} \\ &\Leftrightarrow \mathbb{R} \setminus M \text{ ist abgeschlossen} \end{aligned}$$

□

Übungsaufgabe: Es sei $U \subseteq \mathbb{R}$ (bzw. $\mathbb{C}$). Dann gilt: $\overline{\overline{U}} = \overline{U}$. Das bedeutet: Der Abschluss von U ist abgeschlossen. Weiter gilt: $U^{\circ\circ} = U^\circ$. Das bedeutet: Das Innere von U offen.

Satz 2.8 (ε-Umgebungen sind offen.) *Es seien $x \in \mathbb{R}$ und $\varepsilon > 0$. Dann ist $U_\varepsilon(x)$ offen. Eine analoge Aussage gilt über $\mathbb{C}$.*

Beweis: Wir verwenden das Kriterium (32) für Offenheit. Für unser gegebenes $x \in \mathbb{R}$ und $\varepsilon > 0$ ist also zu zeigen:

$$\forall y \in U_\varepsilon(x)\ \exists \delta > 0 :\ U_\delta(y) \subseteq U_\epsilon(x)$$

Hierzu sei $y \in U_\varepsilon(x)$ gegeben. Es gilt also $|y - x| < \varepsilon$. Wir setzen $\delta = \varepsilon - |y - x| > 0$. Zu zeigen ist nun $U_\delta(y) \subseteq U_\varepsilon(x)$. Das sieht man so ein: Es sei $z \in U_\delta(y)$ gegeben; zu zeigen ist nun $z \in U_\varepsilon(x)$. Wir schätzen ab:

$$\begin{aligned} |z - x| &= |(z - y) + (y - x)| \\ &\leq |z - y| + |y - x| && \text{wegen der Dreiecksungleichung} \\ &< \delta + |y - x| && \text{weil } |z - y| < \delta \text{ wegen } z \in U_\delta(y) \\ &= \varepsilon, \end{aligned}$$

also $|z - x| < \varepsilon$ und damit $z \in U_\varepsilon(x)$. Das war zu zeigen.

$\square$

Satz 2.9 (fundamentale Eigenschaften offener Mengen in $\mathbb{R}$)

1. *Die leere Menge $\emptyset$ und der ganze Raum $\mathbb{R}$ sind offen (in $\mathbb{R}$).*
2. *Es seien $A, B \subseteq \mathbb{R}$ offen. Dann ist $A \cap B$ offen.*
3. *Es seien $(A_i)_{i \in I}$ eine beliebige Familie offener Mengen $A_i \subseteq \mathbb{R}$. Dann ist die Menge*

$$\bigcup_{i \in I} A_i = \{x \mid \exists i \in I : x \in A_i\}$$

   ebenfalls offen.

Analoges gilt für offene Mengen in $\mathbb{C}$ und für relativ zu einem Universum M offene Mengen.

Beweis:

1. Die leere Menge $\emptyset$ ist trivialerweise offen, denn sie enthält keine Punkte. $\mathbb{R}$ ist offen, denn $\forall x \in \mathbb{R} : U_1(x) \subseteq \mathbb{R}$.
2. Es sei $x \in A \cap B$. Dann gibt es ein $\varepsilon > 0$ mit $U_\varepsilon(x) \subseteq A$, weil A offen ist. Ebenso gibt es ein $\delta > 0$ mit $U_\delta(x) \subseteq B$, weil B offen ist. Dann ist $U_{\min\{\varepsilon,\delta\}}(x) = U_\varepsilon(x) \cap U_\delta(x) \subseteq A \cap B$, wobei $\min\{\varepsilon, \delta\} > 0$ die kleinere der beiden Zahlen ε und δ bezeichnet. Also ist $A \cap B$ offen.
3. Es sei $x \in \bigcup_{i \in I} A_i$. Es gibt einen Index $j \in I$ mit $x \in A_j$. Für solch ein j gibt es ein $\varepsilon > 0$ mit $U_\varepsilon(x) \subseteq A_j$, weil A_j offen ist. Es folgt: $U_\varepsilon(x) \subseteq A_j \subseteq \bigcup_{i \in I} A_i$, also $U_\varepsilon(x) \subseteq \bigcup_{i \in I} A_i$. Folglich gibt es zu jedem Punkt $x \in \bigcup_{i \in I} A_i$ eine ε-Umgebung, die in $\bigcup_{i \in I} A_i$ enthalten ist. Das bedeutet: $\bigcup_{i \in I} A_i$ ist offen.

$\square$

Insbesondere sind Durchschnitte *endlich vieler* offener Mengen wieder offen. Unendliche Durchschnitte offener Mengen brauchen dagegen im Allgemeinen nicht offen zu sein. Zum Beispiel ist $\bigcap_{\epsilon > 0} U_\epsilon(0) = \{0\}$ nicht offen.

Folgerung: (fundamentale Eigenschaften abgeschlossener Mengen in $\mathbb{R}$)

1. Der ganze Raum $\mathbb{R}$ und die leere Menge $\emptyset$ sind abgeschlossen (in $\mathbb{R}$).
2. Es seien $A, B \subseteq \mathbb{R}$ abgeschlossen. Dann ist $A \cup B$ abgeschlossen.

3. Es seien $(A_i)_{i\in I}$ eine beliebige Familie abgeschlossener Mengen $A_i \subseteq \mathbb{R}$ mit nichtleerer Indexmenge I. Dann ist die Menge

$$\bigcap_{i\in I} A_i = \{x \mid \forall i \in I : x \in A_i\}$$

ebenfalls abgeschlossen.

Analoges gilt über $\mathbb{C}$.
Die Folgerung wird durch Komplementbildung unter Verwendung der de-Morgan Regeln der Mengenlehre bewiesen; abgeschlossene Mengen sind ja genau die Komplemente von offenen Mengen. Wir verzichten hier auf die Ausführung der Details. Die Beschränkung auf nichtleere Indexmengen I ist nötig, da $\bigcap_{i\in\emptyset} A_i$ nicht definiert ist.

Ausblick: topologische Räume. Die topologischen Begriffe "offen", "abgeschlossen", "Berührpunkt", "Rand" etc. sind nicht auf $\mathbb{R}$ und $\mathbb{C}$ beschränkt, sondern können auf allgemeinere Räume übertragen werden. Im Rahmen dieser Vorlesung spielen nur zwei weitere Räume eine Rolle, nämlich die erweiterten Zahlenräume $\mathbb{R}\cup\{\pm\infty\}$ und $\mathbb{C}\cup\{\infty\}$; später werden jedoch noch viele weitere Beispiele dazukommen. Um eine gemeinsame Sprechweise für all diese Beispiele zur Verfügung zu stellen, geben wir hier einen Ausblick auf eine weitgehende Abstraktion: den Begriff des *topologischen Raums*. In der Vorlesung Analysis I spielt dieser Begriff nur eine Nebenrolle als eine praktische Sprechweise.

Definition 2.10 (Abstraktion des Satzes 2.9) Es sei M eine beliebige Menge und $\mathcal{T}$ eine Menge von Teilmengen von M. $\mathcal{T}$ heißt eine *Topologie* auf M, wenn gilt:

1. $\emptyset \in \mathcal{T}$ und $M \in \mathcal{T}$
2. $\forall A, B \in \mathcal{T}: \ A \cap B \in \mathcal{T}$
3. Für alle Familien $(A_i)_{i\in I}$ von Mengen in $\mathcal{T}$ gilt: $\bigcup_{i\in I} A_i \in \mathcal{T}$.

Ein Paar $(M, \mathcal{T})$ heißt *topologischer Raum*, wenn $\mathcal{T}$ eine Topologie auf M ist. Die Elemente von $\mathcal{T}$ werden auch *offene Mengen* in M (bezüglich der Topologie $\mathcal{T}$) genannt.

Beispiel: Die Menge der offenen Mengen von $\mathbb{R}$ bilden eine Topologie.

Alle topologischen Begriffe (z. B. abgeschlossen, Abschluss, innerer Punkt, Inneres, Berührpunkt, Rand, (offene) Umgebung etc.) lassen sich auf den Begriff der offenen Mengen zurückführen und damit auf topologische Räume erweitern. Wir verzichten hier auf Details.

2.2 Topologie von $\mathbb{R}\cup\{\pm\infty\}$ und $\mathbb{C}\cup\{\infty\}$

Definition 2.11 Eine Teilmenge $M \subseteq \mathbb{R}\cup\{\pm\infty\}$ heißt *offen in* $\mathbb{R}\cup\{\pm\infty\}$, wenn gilt:

1. $M \cap \mathbb{R}$ ist offen in $\mathbb{R}$.

2. Falls $+\infty \in M$, so gilt:

$$\exists\, r \in \mathbb{R}\, \forall x \in \mathbb{R} :\; (x > r \;\Rightarrow\; x \in M).$$

Anders gesagt: Wenn $+\infty \in M$, so liegen alle genügend großen x in M.

3. Falls $-\infty \in M$, so gilt:

$$\exists\, r \in \mathbb{R}\, \forall\, x \in \mathbb{R} :\; (x < r \;\Rightarrow\; x \in M),$$

d. h. wenn $-\infty \in M$, so liegen alle genügend kleinen x in M.

Die letzten beiden Bedingungen kann man auch so ausdrücken:

2'. Falls $+\infty \in M$, so ist $\mathbb{R}\backslash M$ nach oben beschränkt.

3'. Falls $-\infty \in M$, so ist $\mathbb{R}\backslash M$ nach unten beschränkt.

Eine Menge $M \subseteq \mathbb{C} \cup \{\infty\}$ heißt *offen in* $\mathbb{C} \cup \{\infty\}$, wenn gilt:

1. $M \cap \mathbb{C}$ ist offen in $\mathbb{C}$.
2. Falls $\infty \in M$, so gilt: $\exists r \in \mathbb{R}\, \forall x \in \mathbb{C} :\; (|x| > r \;\Rightarrow\; x \in M)$.

Wir nennen eine Menge $M \subseteq \mathbb{C}$ beschränkt, wenn $\exists N > 0\, \forall z \in M :\; |z| \leq N$ gilt, d.h. wenn M ganz in einer Kreisscheibe um 0 enthalten ist. Die letzte Bedingung kann man damit so formulieren:

2'. falls $\infty \in M$, so ist das Komplement von M beschränkt.

Bemerkung: Die Menge der offenen Mengen in $\mathbb{R} \cup \{\pm\infty\}$ (bzw. in $\mathbb{C} \cup \{\infty\}$) bilden eine Topologie über $\mathbb{R} \cup \{\pm\infty\}$ (bzw. über $\mathbb{C} \cup \{\infty\}$).

2.3 Häufungspunkte

Definition 2.12 (Häufungspunkt) *Es sei $a = (a_n)_{n\in\mathbb{N}_0}$ eine Folge*[10] *von reellen Zahlen. Anders gesagt: $a : \mathbb{N}_0 \to \mathbb{R}$.*
Eine Zahl $x \in \mathbb{R}$ heißt Häufungspunkt von a, wenn gilt:

$$\forall\, \varepsilon > 0\, \forall\, n \in \mathbb{N}_0\, \exists m > n :\; a_m \in U_\varepsilon(x)$$

Anders gesagt lautet die definierende Bedingung für einen Häufungspunkt:
Für alle $\varepsilon > 0$ gibt es beliebig große $m \in \mathbb{N}$, so dass $|a_m - x| < \varepsilon$ gilt.

[10] Analog kann man auch nur $(a_n)_{n\in\mathbb{N}}$ betrachten. Es ist Geschmacksache, ob man Folgen mit dem Index 0 oder dem Index 1 beginnen lassen will.

Nochmal anders gesagt: *Jede ε-Umgebung von x enthält unendlich viele Folgenglieder.* Notation mit einer anderen Formel: *Die Menge der Häufungspunkte von a ist*

$$\bigcap_{n\in\mathbb{N}} \overline{\{a_m | m > n\}}.$$

Beispiel 1: Die Folge[11] $(1/n)_{n\in\mathbb{N}}$ hat den Häufungspunkt 0, denn jede ε-Umgebung $U_\varepsilon(0) =]-\varepsilon,\ \varepsilon\,[$ von 0 enthält unendlich viele Folgenglieder, nämlich sogar alle Folgenglieder von einer bestimmten Stelle an:

$$\forall\, \varepsilon > 0\ \forall\, n \in \mathbb{N}\ \exists\, m > n : \left|\frac{1}{m}\right| < \varepsilon.$$

Dies folgt aus dem archimedischen Axiom.

Beispiel 2: Wir zählen $\mathbb{Q}^+ = \left\{\frac{n}{m} | n, m \in \mathbb{N}\right\}$ wie folgt ab:

$$\begin{array}{ccccccccccccc}
a_0 = \frac{1}{1} & & a_2 = \frac{1}{2} & & a_5 = \frac{1}{3} & & a_9 = \frac{1}{4} & & a_{14} = \frac{1}{5} & & a_{20} = \frac{1}{6} & & \cdots \\
 & \nearrow & & \nearrow & & \nearrow & & \nearrow & & \nearrow & & & \\
a_1 = \frac{2}{1} & & a_4 = \frac{2}{2} & & a_8 = \frac{2}{3} & & a_{13} = \frac{2}{4} & & a_{19} = \frac{2}{5} & & \cdots & & \\
 & \nearrow & & \nearrow & & \nearrow & & \nearrow & & & & & \\
a_3 = \frac{3}{1} & & a_7 = \frac{3}{2} & & a_{12} = \frac{3}{3} & & a_{18} = \frac{3}{4} & & \cdots & & & & \\
 & \nearrow & & \nearrow & & \nearrow & & & & & & & \\
a_6 = \frac{4}{1} & & a_{11} = \frac{4}{2} & & a_{17} = \frac{4}{3} & & \cdots & & & & & & \\
 & \nearrow & & \nearrow & & & & & & & & & \\
a_{10} = \frac{5}{1} & & a_{16} = \frac{5}{2} & & \cdots & & & & & & & & \\
 & \nearrow & & & & & & & & & & & \\
a_{15} = \frac{6}{1} & & \cdots & & & & & & & & & & \\
\vdots & & & & & & & & & & & &
\end{array}$$

Die so konstruierte Folge $a : \mathbb{N} \to \mathbb{Q}^+$ ist surjektiv. Die Folge $(a_n)_{n\in\mathbb{N}}$ hat alle positiven reellen Zahlen und 0 als Häufungspunkte, denn in jeder ε-Umgebung $]x - \varepsilon, x + \varepsilon[$ einer reellen Zahl $x \geq 0$ liegen unendlich viele rationale Zahlen.

Bemerkung: Eine Menge Q heißt *abzählbar*, wenn $Q = \emptyset$ oder es eine surjektive Abbildung $a : \mathbb{N} \to Q$ gibt. Das Abzählverfahren zeigt: $\mathbb{Q}^+$ ist abzählbar. Ebenso: $\mathbb{N} \times \mathbb{N} = \{(n, m) \mid n, m \in \mathbb{N}\}$ ist abzählbar, und allgemeiner: $Q_1 \times Q_2$ ist abzählbar, wenn Q_1 und Q_2 abzählbar sind.

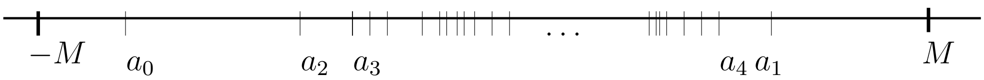
[11] Man störe sich nicht daran, dass die Folge im Beispiel mit $\mathbb{N}$ indiziert ist, in der Definition 2.12 jedoch mit $\mathbb{N}_0$.

Beispiel 3: Die Folge $(2^n)_{n\in\mathbb{N}}$ hat keine Häufungspunkte in $\mathbb{R}$. Anschaulich gesehen häufen sich die Folgenglieder bei $+\infty$. Wir erweitern die Definition von "Häufungspunkten" auf unendlich ferne Punkte.

Definition 2.13 (Häufungspunkte – abstrahiert) *Es sei* $X = \mathbb{R}$, $\mathbb{C}$,$\mathbb{R} \cup \{\pm\infty\}$, $\mathbb{C} \cup \{\infty\}$, *(oder allgemeiner* X *ein topologischer Raum). Ein Punkt* $x \in X$ *heißt Häufungspunkt einer Folge* $(a_n)_{n\in\mathbb{N}}$ *mit Werten in* X*, wenn jede offene Umgebung von* x *in* X *unendlich viele Folgenglieder enthält.*

Beispiel 4: Die Folge $((-2)^n)_{n\in\mathbb{N}}$ hat die Häufungspunkte $+\infty$ und $-\infty$, denn die Menge $\{(-2)^n | n \in \mathbb{N}\}$ ist sowohl nach oben als auch nach unten unbeschränkt.
Betrachten wir eine beschränkte Folge $(a_n)_{n\in\mathbb{N}}$ in $\mathbb{R}$, d. h. $\exists\, M \in \mathbb{R}\, \forall n \in \mathbb{N}: \ |a_n| < M$.

Illustration:



Anschaulich ist plausibel: Die Punkte häufen sich irgendwo. Dies zu beweisen, wird unsere erste Anwendung des Vollständigkeitsaxioms:

Satz 2.14 (Satz von Bolzano-Weierstraß)
Jede beschränkte Folge $(a_n)_{n\in\mathbb{N}}$ *in* $\mathbb{R}$ *besitzt mindestens einen Häufungspunkt.* *wichtig!*

Beweis: Mit Hilfe der Beschränktheit der gegebenen Folge $(a_n)_{n\in\mathbb{N}}$ wählen wir ein $M \geq 0$, so dass für alle $n \in \mathbb{N}$ gilt: $|a_n| \leq M$. Betrachten wir die Menge

$$\begin{aligned} K &= \{x \in \mathbb{R} \mid \text{für alle genügend großen } n \text{ gilt: } a_n \geq x\} \\ &= \{x \in \mathbb{R} \mid \exists\, m \in \mathbb{N}\ \forall\, n > m: \ a_n \geq x\}. \end{aligned}$$

Es gilt $-M \in K$, weil für alle n (und damit erst recht für alle genügend großen $n \in \mathbb{N}$) gilt: $a_n \geq -M$. Also ist $K \neq \emptyset$.
Weiter ist M eine obere Schranke von K. (*Beweis hierzu:* Es sei $x \in K$ gegeben; zu zeigen ist nun $x \leq M$. Weil für alle genügend großen $n \in \mathbb{N}$ gilt: $a_n \geq x$, finden wir ein $n \in \mathbb{N}$ mit $a_n \geq x$, so dass $x \leq a_n \leq |a_n| \leq M$ und damit die Behauptung $x \leq M$ folgt.)
Nach dem Vollständigkeitsaxiom existiert das Supremum $s := \sup K \in \mathbb{R}$ der nichtleeren, nach oben beschränkten Menge K.
Wir zeigen nun: s ist ein Häufungspunkt von $(a_n)_{n\in\mathbb{N}}$. Hierzu sei $\varepsilon > 0$ gegeben. Zu zeigen ist nun, dass unendlich viele Folgenglieder von $(a_n)_{n\in\mathbb{N}}$ in $U_\epsilon(s)$ liegen. Weil s eine obere Schranke von K ist und $s+\varepsilon > s$ gilt, folgt $s+\varepsilon \notin K$. Nach der Definition von K bedeutet das:

$$\neg\exists\, m \in \mathbb{N}\ \forall\, n > m: \ a_n \geq s + \varepsilon,$$

anders gesagt:

$$\forall\, m \in \mathbb{N}\ \exists\, n > m: \ a_n < s + \varepsilon,$$

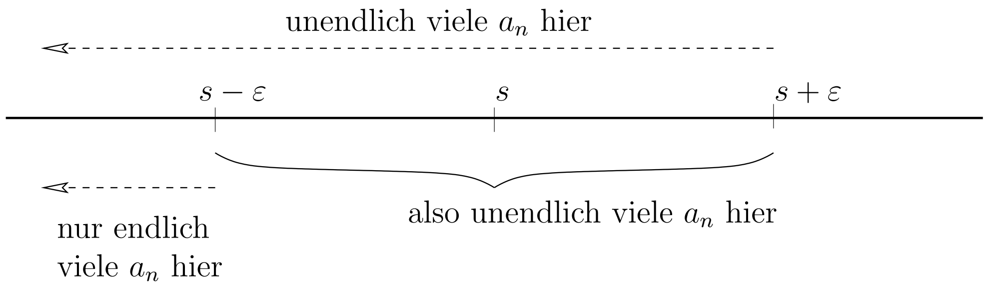
d.h. es gibt unendlich viele $n \in \mathbb{N}$ mit $a_n < s + \varepsilon$. Die Zahl s ist obere Schranke von K, die Zahl $s - \varepsilon$ jedoch nicht, da $s - \epsilon$ kleiner als die *kleinste* obere Schranke $s = \sup K$ von K ist. Wir finden also ein $x \in K$ mit $s - \varepsilon < x \leq s$, d.h. für alle genügend großen $n \in \mathbb{N}$ gilt: $a_n \geq x > s - \varepsilon$, anders gesagt

$$\exists\, m \in \mathbb{N}\ \forall\, n > m :\ a_n \geq x > s - \varepsilon.$$

Wir schließen: Unendlich viele der Folgenglieder müssen zwischen $s - \varepsilon$ und $s + \varepsilon$ liegen, d. h. die ε-Umgebung $U_\varepsilon(s) =]s - \varepsilon, s + \varepsilon[$ von s enthält unendlich viele Folgenglieder. Das war zu zeigen.

$\square$

Illustration zum Beweis des Satzes von Bolzano-Weierstraß:



Bemerkung: Der im Beweis des Satzes von Bolzano-Weierstraß gefundene Häufungspunkt s ist der *kleinste* Häufungpunkt der Folge $(a_n)_{n\in\mathbb{N}}$; später nennen wir ihn den *Limes inferior* der Folge, in Zeichen $s = \liminf_{n\to\infty} a_n$.

Ausblick: Varianten des Satzes von Bolzano-Weierstraß:

- Jede Folge in $\mathbb{R} \cup \{\pm\infty\}$ besitzt mindestens einen Häufungspunkt in $\mathbb{R} \cup \{\pm\infty\}$.
- Jede beschränkte Folge in $\mathbb{C}$ besitzt mindestens einen Häufungspunkt.
  [Das beweisen wir erst später.]
- Jede Folge in $\mathbb{C} \cup \{\infty\}$ besitzt mindestens einen Häufungspunkt in $\mathbb{C} \cup \{\infty\}$.
  [Das folgt aus der vorhergehenden Aussage.]

2.4 Kompaktheit

Wir stellen dem Satz von Bolzano-Weierstraß in diesem Abschnitt ein abstraktes Gegenstück zur Seite, den Satz von Heine-Borel. Beide Sätze beruhen letztlich auf dem Vollständigkeitsaxiom. Kompaktheit spielt in der Analysis eine ähnliche Rolle wie Endlichkeit; sie erlaubt es in verschiedensten Anwendungen, Probleme auf *endliche* Situationen zu reduzieren. Dieser Abschnitt ist vielleicht der abstrakteste der gesamten Vorlesung. Wir arbeiten nämlich hier höher als bisher in der Hierarchie der Mengen, nicht mit *einer* offenen Menge, sondern mit Systemen offener Mengen, also mit möglicherweise unendlichen *Familien* offener Mengen.

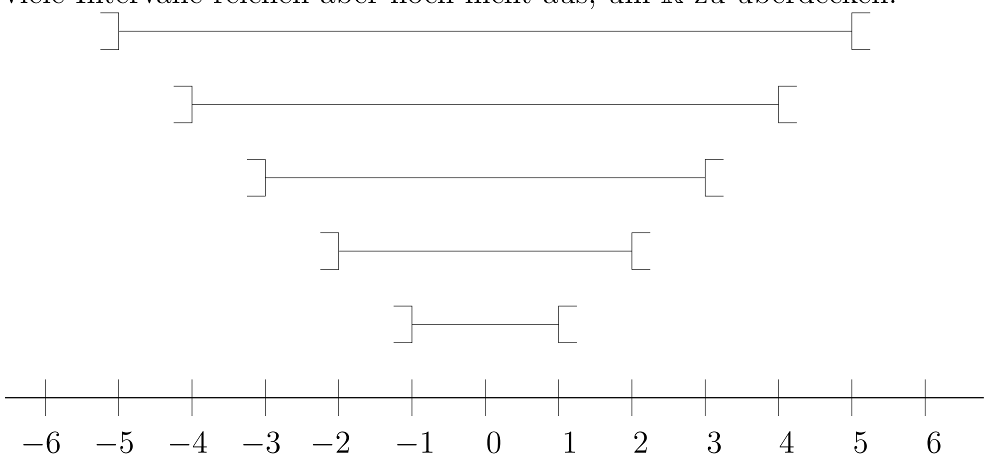
Definition 2.15 *Es sei $M \subseteq \mathbb{R}$. Eine Familie $(U_i)_{i \in I}$ von Mengen $U_i \subseteq \mathbb{R}$ heißt offene Überdeckung von M, wenn gilt:*

1. *Für jedes $i \in I$ ist U_i offen.*
2. $\bigcup\limits_{i \in I} U_i \supseteq M$.

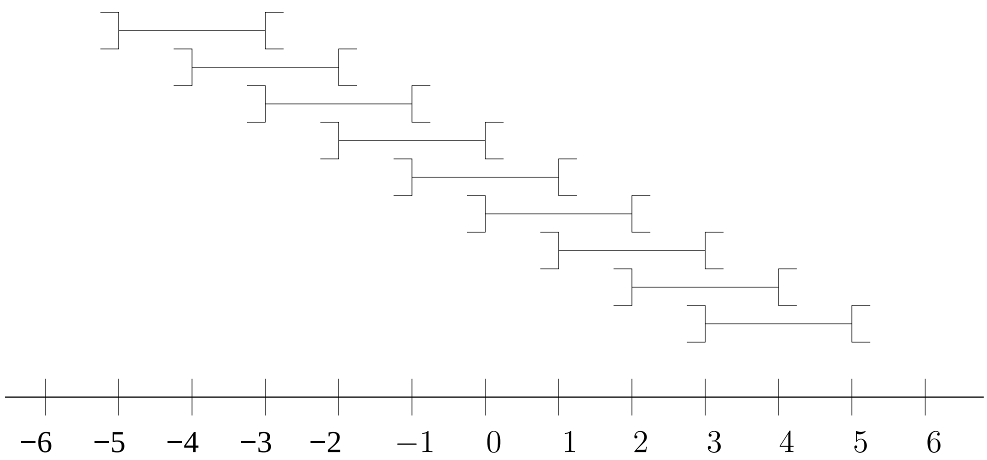
Die zweite Bedingung in dieser Definition bedeutet: Jedes Element von M ist in mindestens einem U_i, $i \in I$, enthalten.

Bemerkung: Analog kann man offene Überdeckungen für beliebige topologische Räume definieren.

Beispiel 1: Die Familie $\big(\,]-M, M[\, \big)_{M \in \mathbb{N}}$ ist eine offene Überdeckung von $\mathbb{R}$. Endlich viele Intervalle reichen aber noch nicht aus, um $\mathbb{R}$ zu überdecken.



Beispiel 2: Die Familie $\big(\,]k-1, k+1[\, \big)_{k \in \mathbb{Z}}$ ist auch eine offene Überdeckung von $\mathbb{R}$. Sobald man aber auch nur mindestens eine der Mengen $]k-1, k+1[$ weglässt, wird $\mathbb{R}$ nicht mehr überdeckt, denn k wird dann nicht mehr überdeckt.



Beispiel 3: Für jedes $\varepsilon > 0$ ist $\big(U_\varepsilon(x)\big)_{x\in[0,1]}$ eine offene Überdeckung des Einheitsintervalls $[0,1]$. Diesmal reichen endlich viele dieser offenen Mengen aus, um $[0,1]$ zu überdecken. Wählen wir $n \in \mathbb{N}$ mit $\frac{1}{n} < \varepsilon$, so ist schon $\big(U_\varepsilon(\frac{k}{n})\big)_{k=0,1,\ldots,n}$ eine endliche offene Teilüberdeckung dieser Überdeckung von $[0,1]$.

Beispiel 4: *(typisches Beispiel für das Auftreten offener Überdeckungen)* Es sei $a = (a_n)_{n\in\mathbb{N}}$ eine Folge mit Werten in einer Menge $B \subseteq \mathbb{R}$. Wir setzen

$$C = \{x \in B \mid x \text{ ist kein Häufungspunkt von } a\}.$$

Wir wählen für jedes $x \in C$ eine offene Umgebung U_x von x aus, die nur endlich viele Folgenglieder von a enthält, d. h. $\{n \in \mathbb{N} \mid a_n \in U_x\}$ soll endlich sein. Dann ist $(U_x)_{x\in C}$ eine offene Überdeckung von C.

Definition 2.16 (Kompaktheit) *Eine Menge $M \subseteq \mathbb{R}$ heißt kompakt, wenn jede offene Überdeckung $(U_i)_{i\in I}$ von M eine endliche Teilüberdeckung besitzt.*

Das bedeutet: Es muss zu jeder offenen Überdeckung $(U_i)_{i\in I}$ von M eine endliche Menge $E = \{i_1, \ldots, i_n\} \subseteq I$ geben, so dass $\bigcup\limits_{i\in E} U_i \supseteq M$ gilt.

abstrakt, aber wichtig!

Kompaktheit für beliebige topologische Räume $(M, \mathcal{T})$ statt $M \subseteq \mathbb{R}$ wird wörtlich genauso definiert.

Die obigen Beispiele 1 und 2 zeigen: $\mathbb{R}$ ist nicht kompakt. $[0,1[$ ist auch nicht kompakt, denn $(]-\infty, x[)_{x\in [0,1[}$ ist eine offene Überdeckung von $[0,1[$, die keine endliche Teilüberdeckung besitzt. Der folgende Satz gibt jedoch Beispiele für kompakte Mengen:

Satz 2.17 (Satz von Heine-Borel) *Jedes abgeschlossene, beschränkte Intervall $[a,b]$ ist kompakt.*

wichtig!

Hier sind $a \leq b$ zwei reelle Zahlen.

Beweis:[12] Es sei $(U_i)_{i\in I}$ eine offene Überdeckung von $[a,b]$. Wir setzen

$$M = \left\{ x \in [a,b] \;\middle|\; \text{Es gibt eine endliche Menge } E \subseteq I \text{ mit } \bigcup_{i\in E} U_i \supseteq [a,x] \right\}.$$

Wir müssen also $b \in M$ zeigen. Es gilt $a \in M$, denn $\{a\} = [a,a]$ wird schon von einer einzigen Menge U_{i_0} für ein $i_0 \in I$ überdeckt. Also ist M nichtleer. Weiter gilt: $M \subseteq [a,b]$, also ist M nach oben beschränkt. Nach dem Vollständigkeitsaxiom existiert demnach $s = \sup M$. Wegen $a \in M$ ist $s \geq a$, und wegen $M \subseteq [a,b]$ ist $s \leq b$. Es folgt: $s \in [a,b]$. Weil $(U_i)_{i\in I}$ eine offene Überdeckung von $[a,b]$ ist, können wir $j \in I$ wählen, so dass $s \in U_j$ gilt. Weil U_j eine offene Umgebung von s ist, gibt es ein $\varepsilon > 0$, so dass gilt:

$$U_\varepsilon(s) =]s-\varepsilon, s+\varepsilon[\subseteq U_j. \tag{33}$$

Wir fixieren solch ein $\varepsilon > 0$. Wegen $s-\varepsilon < s = \sup M$ gibt es ein $x \in M$ mit $s-\varepsilon < x \leq s$; nehmen wir solch ein x. Nach der Definition von M und wegen $x \in M$ gibt es eine endliche Menge $E \subseteq I$ mit

$$\bigcup_{i\in E} U_i \supseteq [a,x], \tag{34}$$

d.h. $[a,x]$ wird schon von endlich vielen der U_i überdeckt. Fixieren wir solch eine endliche Menge E. Zusammen mit der Formel (33) und $x \in]s-\varepsilon, s+\varepsilon[$ schließen wir

$$[a, s+\varepsilon[\subseteq [a,x] \cup]s-\varepsilon, s+\varepsilon[\subseteq \bigcup_{i\in E} U_i \cup U_j = \bigcup_{i\in E\cup\{j\}} U_i, \tag{35}$$

d.h. das Intervall $[a, s+\varepsilon[$ wird von den endlich vielen Mengen U_i, $i \in E \cup \{j\}$, überdeckt. Wir unterscheiden nun zwei Fälle:[13]

- *1. Fall:* Es gelte $b < s+\varepsilon$. Das impliziert $[a,b] \subseteq [a, s+\varepsilon[$, so dass auch $[a,b]$ von den endlich vielen Mengen U_i, $i \in E \cup \{j\}$, überdeckt wird. Anders gesagt gilt $b \in M$, was zu zeigen war.

- *2. Fall:* Es gelte $b \geq s+\varepsilon$. In diesem Fall nehmen wir ein y mit $s < y < s+\varepsilon$. Insbesondere gilt $y \in [a,b]$ wegen $a \leq s < y < s+\varepsilon \leq b$. Einerseits wird das Intervall $[a,y] \subseteq [a, s+\varepsilon[$ ebenfalls von den endlich vielen Mengen U_i, $i \in E \cup \{j\}$, überdeckt. Das bedeutet $y \in M$. Andererseits folgt $y \notin M$ wegen $y > s = \sup M$. Aus dem somit gezeigten Widerspruch $y \notin M \wedge y \in M$ schließen wir, dass der 2. Fall nicht auftreten kann.

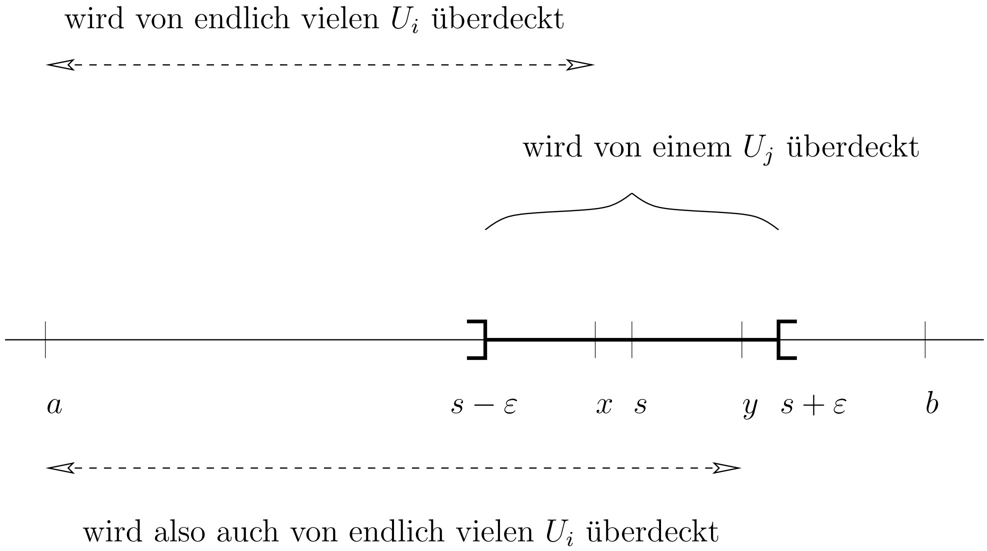
Im somit einzig möglichen 1. Fall war aber die Behauptung $b \in M$ schon gezeigt.

$\square$

[12] Obwohl dieser Beweis keine vollständige Induktion verwendet, denn es ist ja gar nicht von natürlichen Zahlen die Rede, hat er doch etwas Ähnlichkeit mit einem Induktionsbeweis: Wir arbeiten uns "von links nach rechts", ausgehend vom "Beweisanfang" $x = a$, also vom Intervall $[a,a] = \{a\}$, zum gewünschten Fall $x = b$, also dem ganzen Intervall $[a,b]$ vorwärts. Die Aussage (34) kann man als Analogie zu einer Induktionsvoraussetzung ansehen; das Hinzufügen von $\{j\}$ zur endlichen Menge E, siehe Formel (35), entspricht dann einem Induktionsschritt.

[13] Dem nun folgenden Schluss liegt aussagenlogisch gesehen die Tautologie $[(A \Rightarrow B) \wedge (\neg A \Rightarrow \bot)] \Rightarrow B$ zugrunde, wobei A für $b < s+\varepsilon$ und B für $b \in M$ steht.

Illustration zum Beweis des Satzes von Heine-Borel: Die Graphik illustriert den 2. Fall, der nicht auftreten kann, und der im Verlauf des Beweises widerlegt wird.



Korollar 2.18 (Charakterisierung der Kompaktheit in $\mathbb{R}$) *Es sei $M \subseteq \mathbb{R}$. Folgende Aussagen sind äquivalent:*

wichtig!

a) M ist abgeschlossen und beschränkt.

b) M ist kompakt.

c) Jede Folge $(a_n)_{n\in\mathbb{N}}$ mit Werten in M besitzt mindestens einen Häufungspunkt in M.

Sprechweise für die Aussage c): "M ist folgenkompakt."

Beweis:

$\boxed{\text{a)} \Rightarrow \text{b)}}$ Es sei M abgeschlossen und beschränkt. Wegen der Beschränktheit von M gibt es zwei reelle Zahlen $a < b$ mit $M \subseteq [a,b]$; wir fixieren solche Zahlen a und b. Es sei $(U_i)_{i\in I}$ eine offene Überdeckung von M. Nehmen wir noch die offene[14] Menge $\mathbb{R}\backslash M$ zu dieser offenen Überdeckung dazu, erhalten wir eine offene Überdeckung von $[a,b]$. Nach dem Satz von Heine-Borel reichen schon endliche viele Mengen $(U_i)_{i=i_1\dots i_n}$ zusammen mit $\mathbb{R}\backslash M$ aus, um das Intervall $[a,b]$ zu überdecken. Wenn wir $\mathbb{R}\backslash M$ in dieser Überdeckung weglassen, erhalten wir immerhin noch eine endliche Teilüberdeckung $(U_i)_{i=i_1\dots i_n}$ von M, denn $(\mathbb{R}\backslash M)\cap M = \emptyset$. Also ist M kompakt.

[14]Die Menge $\mathbb{R}\backslash M$ ist offen, weil sie das Komplement der abgeschlossenen Menge M ist.

$\boxed{\text{b)} \Rightarrow \text{c)}}$ Es sei M kompakt, und es sei eine Folge $(a_n)_{n\in\mathbb{N}}$ mit Werten in M gegeben. Wir zeigen indirekt, dass $(a_n)_{n\in\mathbb{N}}$ einen Häufungspunkt in M besitzt. Hierzu nehmen wir das Gegenteil an, also dass $(a_n)_{n\in\mathbb{N}}$ *keinen* Häufungspunkt in M besitze.[15] Mit Hilfe dieser Annahme wählen wir[16] wie in Beispiel 4 zu jedem $x \in M$ eine offene Umgebung U_x aus, die nur endlich viele Folgenglieder trifft, d.h. für die die Menge $\{n \in \mathbb{N} \mid a_n \in U_x\}$ endlich ist. Weil M kompakt ist, besitzt die offene Überdeckung $(U_x)_{x\in M}$ von M eine endliche Teilüberdeckung $(U_x)_{x\in E}$. Hier ist E eine endliche Teilmenge von M. Dann ist auch die Menge

$$\bigcup_{x\in E}\{n \in \mathbb{N} \mid a_n \in U_x\} = \left\{ n \in \mathbb{N} \;\middle|\; a_n \in \bigcup_{x\in E} U_x \right\}$$

endlich, denn sie ist eine Vereinigung endlich vieler endlicher Mengen. Andererseits gilt für alle $n \in \mathbb{N}$, also für *unendlich viele* n, die Aussage

$$a_n \in M \subseteq \bigcup_{x\in E} U_x.$$

Das liefert den gewünschten Widerspruch. Wir haben also indirekt gezeigt, dass $(a_n)_{n\in\mathbb{N}}$ einen Häufungspunkt in M besitzt.

$\boxed{\text{c)} \Rightarrow \text{a)}}$ Es sei M folgenkompakt. Wir müssen zeigen, dass M abgeschlossen und beschränkt ist.

Wir zeigen zunächst: Die Menge M ist abgeschlossen. Hierzu sei ein Berührpunkt x von M gegeben. Zu zeigen ist nun $x \in M$.[17] Für jede natürliche Zahl $n \in \mathbb{N}$ wählen wir ein $a_n \in M$ mit $|a_n - x| < \frac{1}{n}$ aus.[18] Wegen der vorausgesetzten Folgenkompaktheit von M besitzt die Folge $(a_n)_{n\in\mathbb{N}}$ einen Häufungspunkt y mit $y \in M$. Wir zeigen nun indirekt $x = y$. Hierzu nehmen wir $x \neq y$ an. Dann gibt es disjunkte offene Umgebungen $U_\varepsilon(x)$ bzw. $U_\varepsilon(y)$ von x bzw. von y, d.h. $U_\varepsilon(x) \cap U_\varepsilon(y) = \emptyset$; man wähle z.B. $\varepsilon = |x - y|/2$. Die offene Umgebung $U_\varepsilon(x)$ von x enthält für alle genügend großen n das Folgenglied a_n, also kann $U_\varepsilon(y)$ höchstens endlich viele Folgenglieder enthalten. Das ergibt einen Widerspruch, da y ein Häufungspunkt der Folge $(a_n)_{n\in\mathbb{N}}$ ist. Damit ist $x = y$ indirekt gezeigt. Wegen $y \in M$ folgt hieraus $x \in M$, was zu zeigen war.

Wir zeigen nun indirekt: Die Menge M ist beschränkt. Nehmen wir also an, dass M unbeschränkt sei. Dann können wir zu jedem $n \in \mathbb{N}$ ein $b_n \in M$ mit $|b_n| > n$ wählen.[19] Die Folge $(b_n)_{n\in\mathbb{N}}$ besitzt dann keinen Häufungspunkt in $\mathbb{R}$; höchstens die unendlich fernen Punkte $+\infty$ und $-\infty$ können Häufungspunkte der Folge sein. Insbesondere gibt es keinen

[15] Unser Beweisziel wird damit ein Widerspruch.

[16] mit dem Auswahlaxiom

[17] Hier verwenden wir die Definition der Abgeschlossenheit: M ist abgeschlossen, wenn für jeden Berührpunkt x von M gilt: $x \in M$.

[18] Wegen $x \in \overline{M}$ ist das nach dem Auswahlaxiom möglich; hier wird die Definition von Berührpunkten verwendet: $x \in \overline{M} \Rightarrow \forall n \in \mathbb{N}\ \exists a \in M : |a - x| < \frac{1}{n}$.

[19] Hier verwenden wir wieder das Auswahlaxiom.

Häufungspunkt in M der Folge $(b_n)_{n\in\mathbb{N}}$. Das widerspricht der Voraussetzung c), also der Folgenkompaktheit von M. Damit ist das Beweisziel indirekt gezeigt: Die Menge M ist beschränkt. $\square$

Korollar 2.19 *Es sei $A_0 \supseteq A_1 \supseteq A_2 \ldots$ eine absteigende Folge von abgeschlossenen, beschränkten, nichtleeren Teilmengen von $\mathbb{R}$. Dann ist der Durchschnitt aller Folgenglieder nichtleer:*

$$\bigcap_{n\in\mathbb{N}_0} A_n \neq \emptyset.$$

Beweis: Wäre $\bigcap_{n\in\mathbb{N}_0} A_n = \emptyset$, so wäre $\bigcup_{n\in\mathbb{N}_0}(\mathbb{R}\backslash A_n) = \mathbb{R}\backslash \bigcap_{n\in\mathbb{N}_0} A_n = \mathbb{R}$. Demnach wäre die Folge $(\mathbb{R}\backslash A_n)_{n\in\mathbb{N}_0}$ eine offene Überdeckung von A_0. Man beachte hier, dass die Komplemente der $\mathbb{R}\backslash A_n$ der abgeschlossenen Mengen A_n offen sind. Nun ist A_0 abgeschlossen und beschränkt, also kompakt. Dann reichen schon endlich viele der Mengen $\mathbb{R}\backslash A_n$, sagen wir $(\mathbb{R}\backslash A_n)_{n=0,1,\ldots,N}$, zur Überdeckung von A_0 aus. Andererseits trifft keine der Mengen $\mathbb{R}\backslash A_n$ die Menge A_N, wenn n die Zahlen $0, 1, \ldots, N$ durchläuft:

$$\forall n \in \{0, 1, \ldots, N\} : \; A_N \cap (\mathbb{R}\backslash A_n) = \emptyset,$$

also

$$\emptyset \neq A_N = A_N \cap A_0 \subseteq A_N \cap \bigcup_{n=0}^{N}(\mathbb{R}\backslash A_n) = \emptyset.$$

Das ist ein Widerspruch. Die Annahme $\bigcap_{n\in\mathbb{N}_0} A_n = \emptyset$ ist also falsch.

$\square$

Beispiel: Intervallschachtelungen. Es seien $a_0 \leq a_1 \leq a_2 \leq \ldots$ eine aufsteigende Folge und $b_0 \geq b_1 \geq b_2 \geq \ldots$ eine absteigende Folge reeller Zahlen, so dass $a_n \leq b_n$ für alle $n \in \mathbb{N}_0$ gilt. Dann gibt es mindestens eine Zahl $x \in \mathbb{R}$, die in allen Intevallen $[a_n, b_n]$ liegt, $n \in \mathbb{N}_0$. Anders gesagt:

$$\bigcap_{n\in\mathbb{N}_0} [a_n, b_n] \neq \emptyset.$$

Falls $\sup_{n\in\mathbb{N}_0} a_n = \inf_{n\in\mathbb{N}_0} b_n$ gilt, so ist $x = \sup_{n\in\mathbb{N}_0} a_n = \inf_{n\in\mathbb{N}_0} b_n$ der einzige Punkt im Durchschnitt der Intervalle $[a_n, b_n]$.

Folgerung: Nichtabzählbarkeit von $\mathbb{R}$

Satz 2.20 ($\mathbb{R}$ ist nicht abzählbar.) *Das bedeutet: Jede Folge $(x_n)_{n\in\mathbb{N}_0}$ reeller Zahlen trifft mindestens eine reelle Zahl nicht. Anders gesagt: Jede Abbildung $x : \mathbb{N}_0 \to \mathbb{R}$ ist nicht surjektiv.*

Beweis: Wir zeigen, dass das Intervall $[0, 1]$ nicht im Bild von x enthalten sein kann.

Dazu definieren wir rekursiv[20] eine absteigende Folge $[a_n, b_n]$, $n \in \mathbb{N}_0$, von Intervallen, mit nichtleerem Inneren, d.h. $a_n \le a_{n+1} < b_{n+1} \le b_n$, so dass $[a_n, b_n]$ keinen der Punkte x_k, $0 \le k < n$, trifft.

- Als Rekursionsanfang setzen wir $[a_0, b_0] := [0, 1]$.
- Als Rekursionsvoraussetzung fixieren wie ein $n \in \mathbb{N}_0$ und nehmen an, dass

  $$a_0 \le \ldots \le a_n < b_n \le \ldots \le b_0$$

  mit

  $$\{x_0, \ldots, x_{n-1}\} \cap [a_n, b_n] = \emptyset \tag{36}$$

  schon gefunden seien.
- Im Rekursionsschritt $n \rightsquigarrow n+1$ wählen[21] wir ein Intervall $[a_{n+1}, b_{n+1}] \subseteq [a_n, b_n] \setminus \{x_n\}$ mit $a_{n+1} < b_{n+1}$. Wegen $a_n < b_n$ ist diese Wahl möglich. Nach Konstruktion gilt dann die Rekursionsvoraussetzung auch in der nächsten Rekursionsstufe:

  $$a_0 \le \ldots \le a_n \le a_{n+1} < b_{n+1} \le b_n \le \ldots \le b_0$$

  und

  $$\{x_0, \ldots, x_{n-1}, x_n\} \cap [a_{n+1}, b_{n+1}] = \emptyset. \tag{37}$$

Nun enthält der Durchschnitt $\bigcap_{n\in\mathbb{N}_0} [a_n, b_n]$ keinen der Punkte $(x_n)_{n\in\mathbb{N}_0}$.[22] Dieser Durchschnitt ist aber nach dem Korollar 2.19 nichtleer, so dass $(x_n)_{n\in\mathbb{N}_0}$ mindestens einen Punkt in $[0, 1]$ nicht trifft.

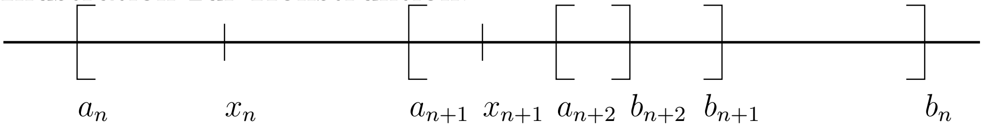
$\square$

[20] Formal-logisch gesehen wird in der nun folgenden Konstruktion *simultan* eine Rekursion und eine Induktion ausgeführt: *Rekursiv* wird die Folge $([a_n, b_n])_{n\in\mathbb{N}_0}$ *definiert* und gleichzeitig *induktiv* ihre wesentlichen Eigenschaften, formuliert in der Rekursionsvoraussetzung, *bewiesen*. Die Verwendung des Worts "Rekursionsvoraussetzung" statt "Induktionsvoraussetzung" soll diese simultane Behandlung andeuten. Es wäre auch möglich gewesen, diese beiden Teile zu trennen, also zuerst nur die Folge $([a_n, b_n])_{n\in\mathbb{N}_0}$ durch Angabe der Rekursionsvorschrift zu definieren und erst anschliessend die Eigenschaften der Folge induktiv zu beweisen. Nachteil davon: Beim Rekursionsschritt hätte man dann noch nicht die Rekursionsvoraussetzung zur Verfügung, weil die Induktion erst später erfolgen würde. Dann müsste man die Rekursionsvorschrift auch auf den Fall erweitern, dass die Rekursionsvoraussetzung nicht gilt, nur um gleich anschließend im Induktionsbeweis zu zeigen, dass dieser Fall doch nicht auftritt.

[21] Obwohl man es sehr leicht übersehen kann, wird hier das Auswahlaxiom verwendet, denn es geschieht ja nicht nur die Auswahl eines einzigen Intervalls $[a_n, b_n]$, sondern die Auswahl der Rekursionsvorschrift mit Hilfe einer Auswahlfunktion.

[22] Formal sieht man das so: Es ist $\forall n \in \mathbb{N}_0 : \ x_n \notin \bigcap_{k\in\mathbb{N}_0} [a_k, b_k]$ zu zeigen (Umbenennung der gebundenen Variablen $n \rightsquigarrow k$ rechts, um Verwechslungen zu vermeiden). Hierzu sei $n \in \mathbb{N}_0$ gegeben. Nun wissen wir $x_n \notin [a_{n+1}, b_{n+1}]$ wegen Formel (37). Wir schließen $\neg\forall k \in \mathbb{N}_0 : \ x_n \in [a_k, b_k]$, was anders geschrieben die Behauptung $x_n \notin \bigcap_{k\in\mathbb{N}_0} [a_k, b_k]$ ist.

Illustration zur Konstruktion:



Satz 2.21 *Jede nichtleere, kompakte Menge* $K \subseteq \mathbb{R}$ *besitzt ein Minimum* $\min K$ *und ein Maximum* $\max K$.

Beweis: Wir setzen $s = \sup K$. Es gilt $s \in \mathbb{R}$, denn K ist nichtleer und nach oben beschränkt. Die Folge $\left([s - \frac{1}{n}, s] \cap K\right)_{n \in \mathbb{N}_0}$ steigt ab. Jedes Folgenglied $[s - \frac{1}{n}, s] \cap K$ ist nichtleer, weil $s = \sup K$. Als Durchschnitt zweier abgeschlossener, beschränkter Mengen ist $[s - \frac{1}{n}, s] \cap K$ abgeschlossen und beschränkt, also kompakt. Nach dem Korollar 2.19 ist der Durchschnitt $\bigcap_{n \in \mathbb{N}} [s - \frac{1}{n}, s] \cap K$ nichtleer. Der Durchschnitt kann aber nur ein einziges Element enthalten, nämlich s. Es folgt $s \in K$. Das Supremum von K ist also auch das Maximum von K.
Der Beweis für das Minimum verläuft analog; statt dem Supremum von K verwendet man hier das Infimum von K.

$\square$

Satz 2.22 *Jede beschränkte Folge* $(a_n)_{n \in \mathbb{N}}$ *in* $\mathbb{R}$ *besitzt einen kleinsten und einen größten Häufungspunkt in* $\mathbb{R}$.

Beweis:[23] Die Menge H der Häufungspunkte der Folge $(a_n)_{n \in \mathbb{N}}$ kann wie folgt als eine Durchschnittsmenge geschrieben werden:

$$H = \bigcap_{m \in \mathbb{N}} \overline{\{a_n | n \geq m\}}.$$

Das ist ein Durchschnitt einer absteigenden Folge von nichtleeren, abgeschlossenen und beschränkten Mengen. Nach dem Korollar 2.19 ist H nichtleer. Als Durchschnitt von abgeschlossenen Mengen ist H abgeschlossen. Weil H auch beschränkt ist, folgt: H ist kompakt. Also besitzt H sowohl ein Minimum als auch ein Maximum.

$\square$

Variante: Jede Folge $(a_n)_{n \in \mathbb{N}}$ mit Werten in $\mathbb{R} \cup \{\pm\infty\}$ besitzt einen größten und einen kleinsten Häufungspunkt in $\mathbb{R} \cup \{\pm\infty\}$.

[23] Dieser Beweis gibt uns als Nebenprodukt auch einen alternativen Beweis des Satzes von Bolzano-Weierstraß. Man beachte, dass er *nicht* den Satz von Bolzano-Weierstraß verwendet, sondern letztlich nur den Satz von Heine-Borel. Man kann auch das Vollständigkeitsaxiom (Existenz von Suprema) aus dem Satz von Bolzano-Weierstraß herleiten. Damit sieht man, dass das Vollständigkeitsaxiom, der Satz von Bolzano-Weierstraß und der Satz von Heine-Borel auf der Basis der übrigen Axiome der reellen Zahlen äquivalent sind.

Definition 2.23 ($\liminf$ **und** $\limsup$)

- Der größte Häufungspunkt einer Folge $(a_n)_{n\in\mathbb{N}}$ in $\mathbb{R}\cup\{\pm\infty\}$ wird *Limes superior* dieser Folge genannt. In Zeichen: $\limsup_{n\to\infty}\ a_n$.
- Der kleinste Häufungspunkt einer Folge $(a_n)_{n\in\mathbb{N}}$ in $\mathbb{R}\cup\{\pm\infty\}$ wird *Limes inferior* dieser Folge genannt. In Zeichen: $\liminf_{n\to\infty}\ a_n$.

Übungsaufgabe: Es gilt:

$$\limsup_{n\to\infty}\ a_n = \inf_{m\in\mathbb{N}_0}\ \sup_{n>m} a_n,$$
$$\liminf_{n\to\infty}\ a_n = \sup_{m\in\mathbb{N}_0}\ \inf_{n>m} a_n.$$

3 Konvergenz und Stetigkeit

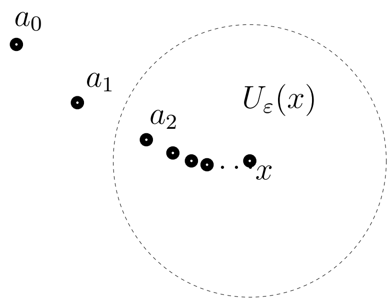
3.1 Konvergenz von Folgen in $\mathbb{R}$ oder $\mathbb{C}$

Zentraler Begriff!

Definition 3.1 *Eine Folge $(a_n)_{n\in\mathbb{N}_0}$ in $\mathbb{R}$ oder $\mathbb{C}$ heißt konvergent gegen $x\in\mathbb{R}$ bzw. $\mathbb{C}$, wenn gilt:*

$$\forall\varepsilon>0\ \exists m\in\mathbb{N}_0\ \forall n>m:\ |a_n-x|<\varepsilon.$$

Anders gesagt: Konvergenz gegen x bedeutet:
Für alle $\varepsilon>0$ liegen höchstens endlich viele Folgenglieder außerhalb von $U_\varepsilon(x)$.
Nochmal anders gesagt: *Für alle $\varepsilon>0$ gilt für alle genügend großen $n\in\mathbb{N}_0$: $a_n\in U_\varepsilon(x)$.*
Die Sprechweise "Für alle genügend großen n gilt $\phi(n)$" bedeutet also $\exists m\in\mathbb{N}_0\ \forall n>m:\ \phi(n)$.
Ist die Folge $(a_n)_{n\in\mathbb{N}_0,n\geq k}$ nur auf einem Endstück $\{n\in\mathbb{N}_0|\ n\geq k\}$ der natürlichen Zahlen definiert, wird Konvergenz ganz analog definiert, wobei "$m\in\mathbb{N}_0$" durch "$m\in\mathbb{N}_0, m\geq k$" ersetzt wird.



Bemerkung: Eine Folge kann höchstens gegen *eine* Zahl konvergieren. In der Tat: Wäre $(a_n)_{n\in\mathbb{N}_0}$ sowohl gegen x als auch gegen $y\neq x$ konvergent, so gälte für alle genügend großen $n\in\mathbb{N}_0$ sowohl $|a_n-x|<|x-y|/2$ als auch $|a_n-y|<|x-y|/2$, was zu folgendem

Widerspruch führt:

$$|x-y| \leq |a_n - x| + |a_n - y| < \frac{|x-y|}{2} + \frac{|x-y|}{2} = |x-y|.$$

Schreibweisen: $a_n \overset{n\to\infty}{\longrightarrow} x$, "$a_n$ hat den Grenzwert x", "a_n hat den Limes x", $\lim_{n\to\infty} a_n = x$. Eine Folge heißt *divergent*, falls sie nicht konvergiert.

Beispiel 1: Es gilt

$$\frac{1}{n} \overset{n\to\infty}{\longrightarrow} 0,$$

denn

$$\forall \varepsilon > 0\ \exists m \in \mathbb{N}\ \forall\ n > m : \left|\frac{1}{n}\right| < \varepsilon.$$

In der Tat: Gegeben $\varepsilon > 0$, wählen wir $m \in \mathbb{N}$ mit $m \geq \frac{1}{\varepsilon}$. Dann gilt für alle $n > m$:

$$\left|\frac{1}{n} - 0\right| = \frac{1}{n} < \frac{1}{m} \leq \varepsilon,$$

also

$$\left|\frac{1}{n} - 0\right| < \varepsilon.$$

Beispiel 2: Für alle $x \in \mathbb{C}$ mit $|x| < 1$ gilt:

$$\lim_{n\to\infty} x^n = 0.$$

Beweis: Wir führen den Beweis zunächst im Fall $0 < |x| < 1$. Es sei $\varepsilon > 0$. Dann ist $x^{-1} = 1/x$ definiert, und

$$|x^{-1}| = \frac{1}{|x|} > 1.$$

Nach dem archimedischen Axiom wählen wir $m \in \mathbb{N}$ so groß, dass

$$m \geq \frac{1}{(|x^{-1}| - 1)\varepsilon} \tag{38}$$

gilt. Anders gesagt bedeutet das:

$$m(|x^{-1}| - 1) \geq \frac{1}{\varepsilon}.$$

Für alle $n \in\ \mathbb{N}$ mit $n > m$ gilt:

$$\begin{aligned}
\frac{1}{|x^n|} &= (1 + (|x^{-1}| - 1))^n \\
&\geq 1 + n(|x^{-1}| - 1) && \text{wegen der Bernoullischen Ungleichung} \\
&> m(|x^{-1}| - 1) && \text{wegen } |x^{-1}| - 1 > 0 \text{ und } n > m \\
&\geq \frac{1}{\varepsilon},
\end{aligned}$$

also $1/|x^n| > 1/\varepsilon$. Es folgt: $|x^n - 0| < \varepsilon$.
Im Fall $x = 0$ ist $x^n = 0$ für alle $n \geq 1$, also $|x^n - 0| < \varepsilon$.

$\square$

Veranschaulichung der Konvergenz durch ein Spiel. Wir können uns die Frage nach der Konvergenz einer Folge mit Hilfe eines Spiels zwischen zwei Personen, dem *Proponenten* und dem *Opponenten* vorstellen. Der Proponent verteidigt die Konvergenz, während der Opponent sie zu widerlegen sucht. Dazu wählt der Opponent $\varepsilon > 0$, worauf der Proponent mit einem $m \in \mathbb{N}_0$ antworten muss. Verfolgen wir so ein Spiel für $\lim_{n\to\infty} 2^{-n} = 0$:

PROPONENT: Ich behaupte: $2^{-n} \overset{n\to\infty}{\longrightarrow} 0$.

OPPONENT: Das glaube ich nicht. Nimm $\varepsilon = \frac{1}{3}$.

PROPONENT: (*rechnet mit $x = 1/2$ in der Formel (38): $\frac{1}{(2-1)\varepsilon} = 3$.*) Dann nimm $m = 3$.

OPPONENT: In der Tat, ich kann kein $n > 3$ finden, so dass $|2^{-n}| \geq 1/3$.
Dann nehmen wir $\varepsilon = \frac{1}{10}$.

PROPONENT: (*rechnet wieder: $\frac{1}{(2-1)\cdot\frac{1}{10}} = 10$.*)

OPPONENT (*etwas verlegen*): Hm, ich finde auch kein $n > 10$, so dass $|2^{-n}| \geq \frac{1}{10}$ gilt.

$\vdots$

Hier verliert der Proponent nicht, d.h. er wird niemals widerlegt.

In dem folgenden Beispiel vertritt der Proponent die (falsche) These "$\sum_{k=1}^{n} \frac{1}{k} \overset{n\to\infty}{\longrightarrow} 2$":

PROPONENT: Ich behaupte $\lim_{n\to\infty} \sum_{k=1}^{n} \frac{1}{k} \overset{n\to\infty}{\longrightarrow} 2$.

OPPONENT (*voller Überzeugung*): Das ist falsch. Nimm doch $\varepsilon = \frac{1}{12}$.

PROPONENT (*verlegen, probiert es trotzdem*): Dazu nehmen wir $m = 3$.

OPPONENT: Nein, das geht nicht: Für $n = 4 > 3$ gilt:

$$\left| \left(\frac{1}{1} + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} \right) - 2 \right| = \frac{1}{12} = \varepsilon, \qquad \text{nicht "} < \varepsilon \text{"}.$$

PROPONENT: Ich habe verloren. Ich finde kein geeignetes $m \in \mathbb{N}$.

Hier gewinnt der Opponent: $\sum_{k=1}^{n} \frac{1}{k}$ konvergiert für $n \to \infty$ nicht gegen 2.

Andere Formulierung der Konvergenzdefinition:

$a_n \stackrel{n\to\infty}{\longrightarrow} x$ bedeutet: *Es gibt eine Funktion* $N_0 :]0,\infty[\to \mathbb{N}$, *so dass für alle* $\varepsilon > 0$ *und alle natürlichen Zahlen* $n > N_0(\varepsilon)$ *gilt:* $|a_n - x| < \varepsilon$.

Beispiele: Für den Nachweis von $\lim_{n\to\infty} 1/n = 0$ können wir

$$N_0(\varepsilon) = \left\lceil \frac{1}{\varepsilon} \right\rceil$$

wählen, wobei wir die Aufrundung einer Zahl zur nächsten ganzen Zahl wie folgt definieren:

$$\lceil y \rceil := \min\{m \in \mathbb{Z} \mid m \geq y\}.$$

Für den Beweis von $x^n \stackrel{n\to\infty}{\longrightarrow} 0$ für $0 < |x| < 1$ setzen wir

$$N_0(\varepsilon) = \left\lceil \frac{1}{(|x^{-1}| - 1)\varepsilon} \right\rceil.$$

Satz 3.2 (geometrische Reihe) *Für* $x \in \mathbb{C}$ *mit* $|x| < 1$ *gilt:*

wichtigste Reihe der Mathematik!

$$\boxed{\sum_{k=0}^{n} x^k \stackrel{n\to\infty}{\longrightarrow} \frac{1}{1-x}}$$

Beweis: Es sei $\varepsilon > 0$. Weil $\lim_{n\to\infty} x^n = 0$, können wir m so groß wählen, dass $|x^{m+1}| < \varepsilon|1 - x|$ gilt; man beachte hierbei $|1 - x| > 0$. Dann gilt für alle $n > m$:

$$\begin{aligned}
\left| \sum_{k=0}^{n} x^k - \frac{1}{1-x} \right| &= \left| \frac{1 - x^{n+1}}{1-x} - \frac{1}{1-x} \right| \qquad \text{mit der geometrischen Summe} \\
&= \left| \frac{x^{n+1}}{1-x} \right| = \frac{|x^{n+1}|}{|1-x|} \\
&= \underbrace{\frac{|x^{m+1}|}{|1-x|}}_{<\varepsilon} \underbrace{|x^{n-m}|}_{<1} \\
&< \varepsilon
\end{aligned}$$

□

Definition 3.3 Es sei $(a_n)_{n\in\mathbb{N}_0}$ eine Folge in $\mathbb{R}$ oder $\mathbb{C}$. Die Folge $\left(\sum_{k=0}^n a_k\right)_{n\in\mathbb{N}_0}$ heißt die *Reihe* zu $(a_n)_{n\in\mathbb{N}_0}$. Ebenso wird auch manchmal der Grenzwert der Folge $\left(\sum_{k=0}^n a_k\right)_{n\in\mathbb{N}_0}$, falls er existiert, Reihe der $(a_n)_{n\in\mathbb{N}_0}$ genannt. Die Folgenglieder $\sum_{k=0}^n a_k$ heißen *Partialsummen* der Reihe. Im Fall der Existenz des Grenzwerts schreiben wir:

$$\boxed{\sum_{k=0}^{\infty} a_k := \lim_{n\to\infty} \sum_{k=0}^{n} a_k}$$

Wir verwenden dann die Sprechweise "Die Reihe $\sum_{k=0}^{\infty} a_k$ konvergiert". Analog werden Reihen zu Folgen $(a_k)_{k\geq l}$ definiert, die nur auf einem Endstück $\{k \in \mathbb{N}_0 |\ k \geq l\}$ der natürlichen Zahlen definiert sein brauchen; man schreibt $\sum_{k=l}^{\infty} a_k := \lim_{n\to\infty} \sum_{k=l}^{n} a_k$, falls der Limes existiert.

Mit dieser Definition können wir die geometrische Reihe auch so schreiben:

Für $x \in \mathbb{C}$ mit $|x| < 1$ gilt:

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

Der nächste Satz zeigt, dass man Grenzwerte und arithmetische Operationen vertauschen darf:

Satz 3.4 Es seien $(a_n)_{n\in\mathbb{N}_0}$ und $(b_n)_{n\in\mathbb{N}_0}$ zwei konvergente Folgen reeller oder auch komplexer Zahlen. Es gelte $\lim_{n\to\infty} a_n = a$ und $\lim_{n\to\infty} b_n = b$. Dann folgt:

a) $a_n + b_n \overset{n\to\infty}{\longrightarrow} a + b$

b) $a_n - b_n \overset{n\to\infty}{\longrightarrow} a - b$

c) $a_n \cdot b_n \overset{n\to\infty}{\longrightarrow} a \cdot b$

d) Falls $b \neq 0$, gilt $\dfrac{a_n}{b_n} \overset{n\to\infty}{\longrightarrow} \dfrac{a}{b}$

Beweis:

a,b) Es sei $\varepsilon > 0$. Wir wählen $m_1 \in \mathbb{N}_0$ so groß, dass

$$\forall n > m_1 : \ |a_n - a| < \frac{\varepsilon}{2} \tag{39}$$

gilt.[24] Diese Wahl ist möglich, da $a_n \overset{n\to\infty}{\longrightarrow} a$. Ebenso wählen wir $m_2 \in \mathbb{N}_0$ so groß, dass auch

$$\forall n > m_2 : \ |b_n - b| < \frac{\varepsilon}{2}$$

[24] Prädikatenlogisch gesehen liegt diesem Schritt die folgende Argumentation zugrunde: Gegeben ist $a_n \overset{n\to\infty}{\longrightarrow} a$, also

$$\forall \epsilon > 0\ \exists m \in \mathbb{N}_0\ \forall n > m : \ |a_n - a| < \frac{\varepsilon}{2}, \tag{40}$$

wobei ε hier durch den Allquantor gebunden ist, also nicht mit dem gegebenen $\varepsilon > 0$ verwechselt werden darf. Wir wenden die Formel (40) auf $\varepsilon/2$ an und erhalten

$$\exists m \in \mathbb{N}_0\ \forall n > m : \ |a_n - a| < \frac{\varepsilon}{2}.$$

Wir nehmen so ein m und benennen es in m_1 um, damit wir später im Beweis den Variablennamen m wieder frei zur Verfügung haben. Damit erhalten wir Formel (39).

gilt. Dann folgt für alle $n > \max\{m_1, m_2\}$:

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

Das heißt: $a_n \pm b_n \stackrel{n\to\infty}{\longrightarrow} a \pm b$.

c) Wir setzen $M := \max\{|a|, |b| + 1\}$. Dann gilt $|b_n| < M$ für alle genügend großen[25] $n \in \mathbb{N}_0$. In der Tat: Weil $\lim_{n\to\infty} b_n = b$, folgt

$$|b_n| - |b| \leq |b_n - b| < 1$$

für alle genügend großen $n \in \mathbb{N}_0$.
Es sei nun $\varepsilon > 0$. Für alle genügend großen n gilt:

$$|a_n - a| < \frac{\varepsilon}{2M} \quad \text{und} \quad |b_n - b| < \frac{\varepsilon}{2M}.$$

Es folgt für alle genügend großen n:

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

d) Wir beweisen zunächst:

$$\frac{1}{b_n} \stackrel{n\to\infty}{\longrightarrow} \frac{1}{b}.$$

Für alle genügend großen n gilt:

$$|b_n - b| \leq \frac{|b|}{2},$$

also[26]

$$|b_n| \geq |b| - |b_n - b| \geq |b| - \frac{|b|}{2} = \frac{|b|}{2} > 0.$$

[25] "Für alle genügend großen n gilt $\Phi(n)$" ist eine Sprechweise für "$\exists m \forall n > m : \Phi(n)$". Hier laufen die Quantoren über natürliche Zahlen, doch die gleiche Sprechweise verwendt man auch bei reellen Zahlen. Gleichbedeutend sagt man auch manchmal "$\Phi(n)$ gilt für $n \to \infty$ schließlich". Wir wenden öfter den folgenden Schluss an: "Wenn $\Phi(n)$ für alle genügend großen n gilt und $\Psi(n)$ für alle genügend großen n gilt, dann gilt auch $\Phi(n) \wedge \Psi(n)$ für alle genügend großen n". Das sieht man so ein: Unter Verwendung von $\exists m \forall n > m : \Phi(n)$ nehmen wir ein m_1, so dass $\forall n > m_1 : \Phi(n)$ gilt. Ebenso nehmen wir unter Verwendung von $\exists m \forall n > m : \Phi(n)$ ein m_2, so dass $\forall n > m_2 : \Phi(n)$ gilt. Wir wählen nun $m = \max\{m_1, m_2\}$. Gegeben ein $n > m$, schließen wir $\Phi(n) \wedge \Psi(n)$.

[26] Wir sagen: Für genügend große n ist b_n *von der* 0 *weg beschränkt*.

Es sei $\varepsilon > 0$. Wählen wir m so groß, daß für $n \geq m$ gilt:

$$|b_n - b| < \min\left\{\frac{1}{2}|b|^2\varepsilon, \frac{|b|}{2}\right\}.$$

Hierbei haben wir $\lim_{n\to\infty} b_n = b$ verwendet. Es folgt:

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

Das bedeutet:

$$\frac{1}{b_n} \stackrel{n\to\infty}{\longrightarrow} \frac{1}{b}.$$

Hieraus und aus $a_n \stackrel{n\to\infty}{\longrightarrow} a$ folgt $\frac{a_n}{b_n} \stackrel{n\to\infty}{\longrightarrow} \frac{a}{b}$ wegen der bereits bewiesenen Aussage c).

□

Gelegentlich ist folgende Monotonieaussage für Grenzwerte nützlich:

Lemma 3.5 *Es seien $(a_n)_{n\in\mathbb{N}_0}$ und $(b_n)_{n\in\mathbb{N}_0}$ zwei konvergente reelle Folgen mit $a_n \leq b_n$ für alle $n \in \mathbb{N}_0$. Dann gilt* $\lim_{n\to\infty} a_n \leq \lim_{n\to\infty} b_n$.

Beweis: Wir setzen $\lim_{n\to\infty} a_n = a$ und $\lim_{n\to\infty} b_n = b$. Es sei $\varepsilon > 0$. Dann folgt für alle genügend großen $n \in \mathbb{N}_0$:

$$a \leq a_n + \frac{\varepsilon}{2} \leq b_n + \frac{\varepsilon}{2} \leq b + \varepsilon.$$

Weil dies für alle $\varepsilon > 0$ gilt, schließen wir: $a \leq b$.

□

3.2 Cauchyfolgen

Ein Problem beim Nachweis der Konvergenz nach der Definition ist, daß man den Grenzwert x kennen muß. Die folgende, eng verwandte Definition vermeidet dies:

Definition 3.6 Eine Folge $(a_n)_{n\in\mathbb{N}_0}$ mit Werten in $\mathbb{R}$ oder $\mathbb{C}$ heißt *Cauchyfolge*, wenn gilt:

$$\forall\varepsilon > 0\ \exists m \in \mathbb{N}_0\ \forall k > m\ \forall l > m :\ |a_k - a_l| < \varepsilon.$$

Es wird also nur verlangt, daß die Folgenglieder untereinander beliebig nahe kommen, wenn nur die Indices genügend groß sind.

Satz 3.7 (Cauchyfolgen in $\mathbb{R}$ sind konvergent) *Eine reelle Folge $(a_n)_{n\in\mathbb{N}_0}$ ist genau dann konvergent in $\mathbb{R}$, wenn sie eine Cauchyfolge ist.*

Beweis:

"$\Rightarrow$": Es gelte $a_n \stackrel{n\to\infty}{\longrightarrow} x$. Es sei $\varepsilon > 0$. Wir wählen $m \in \mathbb{N}_0$ so groß, daß für alle $n > m$ gilt: $|a_n - x| < \frac{\varepsilon}{2}$. Dann folgt für alle $k, l > m$:

$$|a_k - a_l| \leq |a_k - x| + |a_l - x| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Also ist $(a_n)_{n\in\mathbb{N}_0}$ eine Cauchyfolge.

"$\Leftarrow$": Es sei $(a_n)_{n\in\mathbb{N}_0}$ eine Cauchyfolge in $\mathbb{R}$. Es gibt $N \in \mathbb{N}_0$, so daß für alle $n > N$ gilt:

$$|a_n - a_{N+1}| < 1,$$

also

$$|a_n| < |a_{N+1}| + 1,$$

folglich ist $(a_n)_n$ beschränkt durch

$$M := \max\{|a_0|, |a_1|, \ldots, |a_N|, |a_{N+1}| + 1\}.$$

Nach dem Satz von Bolzano-Weierstraß gibt es einen Häufungspunkt $x \in \mathbb{R}$ von $(a_n)_{n\in\mathbb{N}_0}$. Wir zeigen nun: $a_n \stackrel{n\to\infty}{\longrightarrow} x$. Hierzu sei $\varepsilon > 0$. Wir wählen $m \in \mathbb{N}_0$, so daß für alle $k, l > m$ gilt:

$$|a_k - a_l| < \frac{\varepsilon}{2}.$$

Weil x ein Häufungspunkt von $(a_n)_{n\in\mathbb{N}_0}$ ist, gibt es $l > m$ mit $|a_l - x| < \frac{\varepsilon}{2}$. Es folgt für alle $k > m$:

$$|a_k - x| \leq |a_k - a_l| + |a_l - x| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Wir schließen $a_k \stackrel{n\to\infty}{\longrightarrow} x$.

□

Eine analoge Aussage gilt auch über den komplexen Zahlen. Um dies zu zeigen, braucht man aber den Satz von Bolzano-Weierstraß in $\mathbb{C}$, den wir noch nicht bewiesen haben. Das holen wir nun nach. Wir benötigen dazu den Begriff der *Teilfolge* als Hilfsmittel:

Definition 3.8 *Es sei $(n_k)_{k\in\mathbb{N}_0}$ eine streng monoton steigende Folge von natürlichen Zahlen und $(a_n)_{n\in\mathbb{N}_0}$ eine beliebige Folge. Dann heißt $(a_{n_k})_{k\in\mathbb{N}_0}$ eine Teilfolge von $(a_n)_{n\in\mathbb{N}_0}$.*

Beispiel: $(a_{2n})_{n\in\mathbb{N}_0}$ und $(a_{n^2})_{n\in\mathbb{N}_0}$ sind Teilfolgen von $(a_n)_{n\in\mathbb{N}_0}$.

Satz 3.9 *Es seien $(a_n)_{n\in\mathbb{N}_0}$ eine Folge in $\mathbb{C}$ und $x \in \mathbb{C}$. Dann sind folgende Aussagen äquivalent:*

a) x ist Häufungspunkt von $(a_n)_{n\in\mathbb{N}_0}$.

b) Es gibt eine Teilfolge $(a_{n_k})_{k\in\mathbb{N}_0}$ mit $a_{n_k} \stackrel{k\to\infty}{\longrightarrow} x$.

Beweis:
$\boxed{a) \Rightarrow b)}$ Wir wählen rekursiv eine streng monoton steigende Folge $(n_k)_{k\in\mathbb{N}_0}$: Rekursionsanfang: $n_0 = 0$. Im Rekursionsschritt $k \rightsquigarrow k+1$ wählen wir $n_{k+1} > n_k$ so, daß

$$|a_{n_{k+1}} - x| < \frac{1}{k+1}.$$

Dann gilt $a_{n_k} \stackrel{k\to\infty}{\longrightarrow} x$.

$\boxed{b) \Rightarrow a)}$ Es sei $\varepsilon > 0$. Unendlich viele Folgenglieder von $(a_n)_{n\in\mathbb{N}_0}$, nämlich a_{n_k} für alle genügend großen k, liegen in der ε-Umgebung $U_\varepsilon(x)$ von x. Also ist x ein Häufungspunkt der Folge $(a_n)_{n\in\mathbb{N}_0}$.

□

Als Folgerung erhalten wir:

Satz 3.10 (Satz von Bolzano-Weierstraß in $\mathbb{C}$) Jede beschränkte Folge $(a_n)_{n\in\mathbb{N}_0}$ in $\mathbb{C}$ besitzt eine konvergente Teilfolge und somit einen Häufungspunkt.

Beweis: Die beschränkte reelle Folge $(\operatorname{Re} a_n)_{n\in\mathbb{N}_0}$ besitzt einen Häufungspunkt, also eine konvergente Teilfolge $(\operatorname{Re} a_{n_k})_{k\in\mathbb{N}_0}$. Die Folge $(\operatorname{Im} a_{n_k})_{k\in\mathbb{N}_0}$ ist beschränkt, besitzt also eine konvergente Teilfolge $(\operatorname{Im} a_{n_{k_j}})_{j\in\mathbb{N}_0}$. Wir setzen $b_j := a_{n_{k_j}}$ für $j \in \mathbb{N}_0$ und erhalten

$$\operatorname{Re} b_j \stackrel{j\to\infty}{\longrightarrow} \operatorname{Re} z \qquad \text{und} \qquad \operatorname{Im} b_j \stackrel{j\to\infty}{\longrightarrow} \operatorname{Im} z$$

für ein $z \in \mathbb{C}$. Hieraus folgt $\lim_{j\to\infty} b_j = z$. In der Tat: Sei $\varepsilon > 0$. Dann gilt $|\operatorname{Re} b_j - \operatorname{Re} z| < \varepsilon/2$ und $|\operatorname{Im} b_j - \operatorname{Im} z| < \varepsilon/2$ für alle genügend großen $j \in \mathbb{N}_0$. Für diese j folgt:

$$|b_j - z| \leq |\operatorname{Re} b_j - \operatorname{Re} z| + |\operatorname{Im} b_j - \operatorname{Im} z| < \varepsilon.$$

Die Teilfolge $(b_j)_{j\in\mathbb{N}_0}$ von $(a_n)_{n\in\mathbb{N}_0}$ konvergiert also gegen z. Insbesondere ist z ein Häufungspunkt von $(a_n)_{n\in\mathbb{N}_0}$.

□

Wir erhalten nun:

Satz 3.11 *Eine Folge in $\mathbb{C}$ ist genau dann konvergent, wenn sie eine Cauchyfolge ist.*

3.3 Vergleichskriterien für die Konvergenz von Reihen

Satz 3.12 (Majorantenkriterium) *Es seien $a_n \geq 0, n \in \mathbb{N}_0$, so daß die Reihe $\sum_{n=0}^\infty a_n$ in $\mathbb{R}$ konvergiert. Weiter sei $(b_n)_{n\in\mathbb{N}_0}$ eine Folge in $\mathbb{C}$, so daß für alle $n \in \mathbb{N}_0$ gilt: $|b_n| \leq a_n$. Dann konvergiert $\sum_{n=0}^\infty b_n$ in $\mathbb{C}$, und es gilt*

$$\left|\sum_{n=0}^\infty b_n\right| \leq \sum_{n=0}^\infty |b_n| \leq \sum_{n=0}^\infty a_n.$$

Die letzte Aussage kann man als eine unendliche Version der Dreiecksungleichung auffassen.

Beweis des Majorantenkriteriums: $\left(\sum_{n=0}^m b_n\right)_{m\in\mathbb{N}_0}$ ist eine Cauchyfolge in $\mathbb{C}$, also konvergent. In der Tat: Es sei $\varepsilon > 0$. Dann gilt für alle genügend großen $k, l \in \mathbb{N}_0$ mit $k \geq l$:

$$\begin{aligned}\left|\sum_{n=0}^k b_n - \sum_{n=0}^l b_n\right| &= \left|\sum_{n=l+1}^k b_n\right| \leq \sum_{n=l+1}^k |b_n| \\ &\leq \sum_{n=l+1}^k a_n = \left|\sum_{n=0}^k a_n - \sum_{n=0}^l a_n\right| < \varepsilon,\end{aligned}$$

da $\left(\sum_{n=0}^m a_n\right)_{m\in\mathbb{N}_0}$ eine Cauchyfolge ist. Die letzte Aussage des Satzes folgt im Limes $k \to \infty$ aus

$$\left|\sum_{n=0}^k b_n\right| \leq \sum_{n=0}^k |b_n| \leq \sum_{n=0}^k a_n$$

mit Hilfe von Lemma 3.5. Man beachte, dass auch die Reihe $\sum_{n=0}^\infty |b_n|$ konvergiert, denn auch sie wird durch die Reihe $\sum_{n=0}^\infty a_n$ majorisiert.

□

Sprechweise: Wir sagen, eine Reihe $\left(\sum_{n=0}^m b_n\right)_{m\in\mathbb{N}_0}$ *konvergiert absolut*, wenn $\left(\sum_{n=0}^m |b_n|\right)_{m\in\mathbb{N}_0}$ in $\mathbb{R}$ konvergiert.

Aus dem Satz folgt also:

Korollar 3.13 *Absolute Konvergenz von Reihen impliziert Konvergenz:*
$\sum_{n=0}^\infty |b_n|$ *konvergent in* $\mathbb{R}$ $\quad\Rightarrow\quad$ $\sum_{n=0}^\infty b_n$ *konvergent in* $\mathbb{C}$ $\quad\Rightarrow\quad$ $b_n \overset{n\to\infty}{\longrightarrow} 0$

Beweis: Nur die letzte Implikation ist noch zu zeigen. Es sei also $\sum_{n=0}^\infty b_n$ konvergent in $\mathbb{C}$. Dann ist die Folge der Partialsummen $\left(\sum_{n=0}^m b_n\right)_{m\in\mathbb{N}_0}$ eine Cauchyfolge. Gegeben $\varepsilon > 0$, gilt also

$$|b_m| = \left|\sum_{n=0}^m b_n - \sum_{n=0}^{m-1} b_n\right| < \varepsilon$$

für alle genügend großen $m \in \mathbb{N}_0$.

□

Bemerkung: Die Umkehrung des Korollars gilt nicht: Zum Beispiel ist die Reihe $\sum_{n=1}^\infty \frac{(-1)^n}{n}$ konvergent in $\mathbb{R}$ (Übungsaufgabe), $\sum_{n=1}^\infty \frac{1}{n}$ jedoch nicht:

Satz 3.14 (Die harmonische Reihe divergiert.) $\left(\sum_{n=1}^m \frac{1}{n}\right)_{m\in\mathbb{N}_0}$ *ist nach oben unbeschränkt, also nicht konvergent in* $\mathbb{R}$.

Beweis: Wir zeigen durch Induktion für alle $k \in \mathbb{N}_0$:

$$\sum_{n=1}^{2^k} \frac{1}{n} \geq \frac{k}{2}. \tag{41}$$

Das ist sicher richtig für $k = 0$. Für den Induktionsschritt $k \rightsquigarrow k+1$ nehmen wir an, daß die Ungleichung (41) für ein $k \in \mathbb{N}_0$ gilt, und schließen:

$$\begin{aligned}
\sum_{n=1}^{2^{k+1}} \frac{1}{n} &= \sum_{n=1}^{2^k} \frac{1}{n} + \sum_{n=2^k+1}^{2^{k+1}} \frac{1}{n} \\
&\geq \frac{k}{2} + \sum_{n=2^k+1}^{2^{k+1}} \frac{1}{n} && \text{(nach der Induktionsvoraussetzung)} \\
&\geq \frac{k}{2} + \sum_{n=2^k+1}^{2^{k+1}} \frac{1}{2^{k+1}} && \left(\text{weil } \tfrac{1}{n} \geq \tfrac{1}{2^{k+1}} \text{ für } n \leq 2^{k+1}\right) \\
&= \frac{k}{2} + \frac{2^k}{2^{k+1}} = \frac{k+1}{2} && \text{(weil die rechte Summe } 2^k \text{ Summanden umfasst.)}
\end{aligned}$$

Weil die Folge $\left(\sum_{n=1}^{k} \frac{1}{n}\right)_{n\in\mathbb{N}_0}$ monoton steigt, folgt hieraus:

$$\liminf_{k\to\infty} \sum_{n=1}^{k} \frac{1}{n} = +\infty.$$

□

Veranschaulichung zur Divergenz der harmonischen Reihe:

$$1 + \frac{1}{2} + \underbrace{\frac{1}{3} + \frac{1}{4}}_{\geq 2\cdot\frac{1}{4}=\frac{1}{2}} + \underbrace{\frac{1}{5} + \ldots + \frac{1}{8}}_{\geq 4\cdot\frac{1}{8}=\frac{1}{2}} + \underbrace{\frac{1}{9} + \ldots + \frac{1}{16}}_{\geq 8\cdot\frac{1}{16}=\frac{1}{2}} + \underbrace{\frac{1}{17} + \ldots + \frac{1}{32}}_{\geq 16\cdot\frac{1}{32}=\frac{1}{2}} + \ldots$$

Beispiel zum Majorantenkriterium: Für alle $x \in \mathbb{C}$ mit $|x| < 1$ konvergiert $\sum_{n=1}^{\infty} \frac{x^n}{n}$ nach dem Majorantenkriterium. In der Tat: Es gilt $|\frac{x^n}{n}| \leq |x|^n$ für $n \geq 1$, und die geometrische Reihe $\sum_{n=1}^{\infty} |x|^n = \frac{|x|}{1-|x|}$ konvergiert.

3.3.1 Konvergenz und Divergenz von Potenzreihen

Die folgende Klasse von Reihen spielt eine wichtige Rolle in der reellen Analysis und darüber hinaus eine fundamentale Rolle in der komplexen Analysis:

Definition 3.15 (Potenzreihen) *Es sei $(a_n)_{n\in\mathbb{N}_0}$ eine Folge in $\mathbb{C}$ und $x \in \mathbb{C}$. Eine Reihe der Gestalt $\sum_{n=0}^{\infty} a_n x^n$ heißt Potenzreihe in x.*

sehr wichtig!

Der folgende Satz liefert ein Konvergenzkriterium für Potenzreihen:

Satz 3.16 *Es seien* $x, y \in \mathbb{C}$ *mit* $|x| < |y|$ *und* $(a_n)_{n\in\mathbb{N}_0}$ *eine Folge in* $\mathbb{C}$. *Wenn* $(|a_n y^n|)_{n\in\mathbb{N}_0}$ *beschränkt ist, dann konvergiert die Potenzreihe* $\sum_{n=0}^{\infty} a_n x^n$ *in* $\mathbb{C}$ *absolut.*

Beweis: Es sei $M \in \mathbb{R}$ eine obere Schranke von $|a_n y^n|$, $n \in \mathbb{N}_0$. Dann folgt:

$$|a_n x^n| = |a_n y^n| \left|\frac{x}{y}\right|^n \leq M \left|\frac{x}{y}\right|^n.$$

Nun ist die geometrische Reihe

$$\sum_{n=0}^{\infty} M \left|\frac{x}{y}\right|^n = \frac{M}{1 - |\frac{x}{y}|}$$

konvergent in $\mathbb{C}$ wegen $|\frac{x}{y}| < 1$, also folgt die Behauptung nach dem Majorantenkriterium.

□

Hieraus erhalten wir:

Satz 3.17 (Konvergenzkreisscheibe von Potenzreihen) Es sei $(a_n)_{n\in\mathbb{N}_0}$ eine Folge in $\mathbb{C}$. Dann gibt es ein $r \in [0, +\infty]$, so daß für alle $x \in \mathbb{C}$ gilt:

- Falls $|x| < r$, so konvergiert die Potenzreihe $\sum_{n=0}^{\infty} a_n x^n$ in $\mathbb{C}$.
- Falls $|x| > r$, so divergiert diese Potenzreihe in $\mathbb{C}$.

Die Zahl r heißt *Konvergenzradius* der Potenzreihe $\sum_{n=0}^{\infty} a_n x^n$.

Beweis: Wir setzen:

$$r = \sup\{y \geq 0 \mid (|a_n y^n|)_{n\in\mathbb{N}_0} \text{ ist beschränkt}\}.$$

- Falls $x \in \mathbb{C}$ mit $|x| < r$, so gibt es $y \geq 0$ mit $|x| < y \leq r$, so daß $|a_n y^n|, n \in \mathbb{N}_0$ beschränkt ist, also ist $\sum_{n=0}^{\infty} a_n x^n$ konvergent.
- Ist umgekehrt $x \in \mathbb{C}$ mit $|x| > r$, so ist $|a_n x^n|, n \in \mathbb{N}_0$, unbeschränkt, also kann $\sum_{n=0}^{\infty} a_n x^n$ nicht konvergieren.

□

Bemerkung: Für $|x| = r$ sind beide Fälle möglich: Konvergenz oder Divergenz.

Beispiele:

1. Für $r > 0$ hat die geometrische Reihe $\sum_{n=0}^{\infty} r^{-n} x^n$ als Potenzreihe in x den Konvergenzradius r. Für $|x| < r$ gilt

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

2. Die Reihe
$$\left(\sum_{n=0}^{m} n!\, x^n\right)_{m\in\mathbb{N}_0}$$
hat den Konvergenzradius 0. In der Tat ist $(|n!\, y^n|)_{n\in\mathbb{N}_0}$ für alle $y>0$ unbeschränkt.
3. Die Reihe
$$\sum_{n=0}^{\infty} \frac{x^n}{n!}$$
konvergiert für alle $x\in\mathbb{C}$, sie hat also den Konvergenzradius $+\infty$.

Beweis: Für alle $y>0$ ist $(y^n/n!)_{n\in\mathbb{N}_0}$ beschränkt. In der Tat: Für alle $n\in\mathbb{N}_0$ mit $n>y$ gilt
$$\frac{y^n}{n!} = \frac{y}{n}\frac{y^{n-1}}{(n-1)!} \le \frac{y^{n-1}}{(n-1)!},$$
also fällt die Folge $(y^n/n!)_{n\ge\lceil y\rceil}$ monoton. Es folgt also
$$0 \le \frac{y^n}{n!} \le \frac{y^{\lceil y\rceil}}{\lceil y\rceil!}$$
für alle $n\ge\lceil y\rceil$. Also ist das Endstück $(y^n/n!)_{n\ge\lceil y\rceil}$ der Folge $(y^n/n!)_{n\in\mathbb{N}_0}$ beschränkt, also auch die Folge $(y^n/n!)_{n\in\mathbb{N}_0}$ selbst, weil nur endlich viele weitere Folgenglieder hinzukommen.
Mit dem Satz 3.16 schließen wir hieraus: Die Potenzreihe $\sum_{n=0}^{\infty}\frac{x^n}{n!}$ konvergiert für alle $x\in\mathbb{C}$ mit $|x|<y$. Weil $y>0$ beliebig genommen werden kann, hat sie also den Konvergenzradius $+\infty$.

$\square$

Definition 3.18 Die Abbildung $\exp\colon \mathbb{C}\to\mathbb{C}$
$$\exp(x) = \sum_{n=0}^{\infty}\frac{x^n}{n!}$$
heißt *Exponentialfunktion*. Die Reihe $\sum_{n=0}^{\infty}\frac{x^n}{n!}$ heißt *Exponentialreihe.*

extrem wichtig!

Wie wir erst später sehen werden, stimmt die Exponentialfunktion auf den rationalen Zahlen mit der gleichnamigen Funktion, die Sie aus der Schule kennen, überein. *Die geometrische Reihe und die Exponentialreihe sind die beiden wichtigsten Reihen der Mathematik.* Wir studieren die Exponentialfunktion später noch intensiv.

3.3.2 Vergleichskriterien mit der geometrischen Reihe

Das folgende beiden Kriterien zur Konvergenz von Reihen beruhen beide auf einem Vergleich mit einer geometrischen Reihe und damit auf dem Majorantenkriterium. Sie sind in manchen Fällen sehr praktisch, aber sie versagen beide, wenn eine absolut konvergente Reihe langsamer als alle geometrische Reihen konvergiert. Ein Beispiel für solch eine langsamer als geometrisch konvergente Reihe ist $\sum_{n=1}^{\infty} n^{-2}$.

Satz 3.19 (Quotientenkriterium) *Es sei $(a_n)_{n\in\mathbb{N}_0}$ eine Folge in $\mathbb{C}$. Wenn es ein $M \in]0,1[$ gibt, so daß für alle genügend großen $n \in \mathbb{N}_0$ gilt: $|a_{n+1}| \le M|a_n|$, so konvergiert die Reihe $\sum_{n=0}^{\infty} a_n$ absolut.*

Der Name "Quotientenkriterium" beruht darauf, dass man für $a_n \neq 0$ die obige Bedingung auch in der Form $|a_{n+1}/a_n| \le M$ schreiben kann.

Beweis des Quotientenkriteriums: Wir wählen $m \in \mathbb{N}_0$, so daß für alle $n \ge m$ gilt: $|a_{n+1}| \le M|a_n|$. Wir folgern daraus

$$\forall n \ge m\colon |a_n| \le M^{n-m}|a_m|$$

durch vollständige Induktion über n:

Induktionsanfang $n = m$: Es gilt $|a_n| = M^0|a_m|$.
Induktionsschritt $n \rightsquigarrow n+1$: Es sei $n \ge m$, und es gelte die Induktionsvoraussetzung $|a_n| \le M^{n-m}|a_m|$. Dann folgt

$$|a_{n+1}| \le M|a_n| \le M \cdot M^{n-m}|a_m| = M^{(n+1)-m}|a_m|.$$

Nun konvergiert

$$\sum_{n=m}^{\infty} M^{n-m}|a_m| = |a_m|\sum_{n=0}^{\infty} M^n = \frac{|a_m|}{1-M}$$

wegen $M \in]0,1[$, also konvergiert nach dem Majorantenkriterium $\sum_{n=m}^{\infty} a_n$ absolut und damit auch $\sum_{n=0}^{\infty} a_n$ absolut.

□

Beispiel: Es sei $x \in \mathbb{C} \setminus \{0\}$. Für alle genügend großen n gilt:

$$\frac{\left|\frac{x^{n+1}}{(n+1)!}\right|}{\left|\frac{x^n}{n!}\right|} = \frac{|x|}{n+1} \le \frac{1}{2},$$

also ist die Exponentialreihe $\sum_{n=0}^{\infty} \frac{x^n}{n!}$ absolut konvergent.

Satz 3.20 (Wurzelkriterium) *Es sei $(a_n)_{n\in\mathbb{N}_0}$ eine Folge in $\mathbb{C}$ und $M \in]0,1[$. Es gelte $\sqrt[n]{|a_n|} \le M$ für alle genügend großen n.*
Dann ist $\sum_{n=0}^{\infty} a_n$ absolut konvergent.

Die Voraussetzung des Wurzelkriteriums kann man äquivalent auch in der folgenden Form schreiben:

$$\limsup_{n\to\infty} \sqrt[n]{|a_n|} < 1.$$

Beweis des Wurzelkriteriums: Aus der Voraussetzung folgt $|a_n| \le M^n$ für alle genügend großen n. Die Reihe $\sum_{n=m}^{\infty} a_n$ wird also durch die konvergente geometrische Reihe $\sum_{n=m}^{\infty} M^n$ majorisiert. Hieraus folgt die Behauptung.

□

3.4 Konvergenz in $\mathbb{R} \cup \{\pm\infty\}$ und $\mathbb{C} \cup \{\infty\}$

Wir erweitern nun den Konvergenzbegriff auf unendlich ferne Punkte und allgemeiner auf topologische Räume.

Definition 3.21 *Es sei $M = \mathbb{R} \cup \{\pm\infty\}, M = \mathbb{C} \cup \{\infty\}$, oder allgemeiner M ein topologischer Raum. Wir sagen, eine Folge $(a_n)_{n\in\mathbb{N}_0}$ konvergiert gegen $x \in M$, wenn jede offene Umgebung U von x höchstens endlich viele Folgenglieder nicht enthält, d.h. wenn folgendes gilt:*

$$\exists m \in \mathbb{N}_0 \, \forall n > m \colon a_n \in U$$

Bemerkung: Der Grenzwert in $\mathbb{R}\cup\{\pm\infty\}$ oder $\mathbb{C}\cup\{\infty\}$ ist ebenfalls eindeutig bestimmt, falls er existiert.[27] Wir schreiben also wieder $x = \lim_{n\to\infty} a_n$ für den Grenzwert x der Folge $(a_n)_n$, oder auch $a_n \stackrel{n\to\infty}{\longrightarrow} x$. Insbesondere bedeutet das für reelle $a_n, n \in \mathbb{N}_0$:

$$a_n \stackrel{n\to\infty}{\longrightarrow} +\infty \iff \forall M \in \mathbb{R} \, \exists m \in \mathbb{N}_0 \, \forall n > m \colon \; a_n > M, \tag{42}$$

$$a_n \stackrel{n\to\infty}{\longrightarrow} -\infty \iff \forall M \in \mathbb{R} \, \exists m \in \mathbb{N}_0 \, \forall n > m \colon \; a_n < M, \tag{43}$$

und für komplexe a_n:

$$a_n \stackrel{n\to\infty}{\longrightarrow} \infty \text{ in } \mathbb{C} \cup \{\infty\} \iff \forall M > 0 \, \exists m \in \mathbb{N}_0 \, \forall n > m \colon \; |a_n| > M.$$

[27] In allgemeinen topologischen Räumen können dagegen Grenzwerte mehrdeutig sein.

Beispiele:

1. **harmonische Reihe:** Es gilt $\sum_{n=1}^{m} \frac{1}{n} \stackrel{m\to\infty}{\longrightarrow} +\infty$. Anders gesagt: $\sum_{n=1}^{\infty} \frac{1}{n} = +\infty$.

2. Für $x \in \mathbb{C}$ mit $|x| > 1$ gilt $x^n \stackrel{n\to\infty}{\longrightarrow} \infty$ in $\mathbb{C} \cup \{\infty\}$.

Sprechweise: Statt "Konvergenz gegen $\pm\infty$ oder ∞" sagt man auch "bestimmte Divergenz gegen $\pm\infty$ oder ∞". Wenn Mißverständnisse zu befürchten sind, ob Konvergenz in $\mathbb{R}$, in $\mathbb{C}$, in $\mathbb{R} \cup \{\pm\infty\}$ oder in $\mathbb{C} \cup \{\infty\}$ gemeint ist, muß man das explizit spezifizieren; bei fehlender Spezifikation ist meist Konvergenz in $\mathbb{R}$ oder in $\mathbb{C}$ gemeint.

Lemma 3.22 *Jede monoton steigende Folge $(a_n)_{n\in\mathbb{N}_0}$ in $\mathbb{R} \cup \{\pm\infty\}$ konvergiert in $\mathbb{R} \cup \{\pm\infty\}$, und es gilt*

$$\lim_{n\to\infty} a_n = \sup\{a_n |\ n \in \mathbb{N}_0\}.$$

Beweis: Übung.

3.5 Operationen mit Reihen

3.5.1 Vertauschung von Limes und unendlicher Summe

Wir wissen: Sind $(a_n^{(1)})_{n\in\mathbb{N}_0}, (a_n^{(2)})_{n\in\mathbb{N}_0}, \ldots, (a_n^{(j)})_{n\in\mathbb{N}_0}$ konvergente Folgen, so gilt

$$\lim_{n\to\infty} \sum_{i=1}^{j} a_n^{(i)} = \sum_{i=1}^{j} \lim_{n\to\infty} a_n^{(i)}.$$

Unendliche Summen und Limes können jedoch manchmal *nicht* vertauscht werden.

Gegenbeispiel: Wir kürzen ab:

$$\delta_{in} = \begin{cases} 1, \text{ falls} & i = n \\ 0, \text{ falls} & i \neq n \end{cases}$$

δ_{in} wird *Kronecker-Delta* genannt. Es gilt $\lim_{n\to\infty} \delta_{in} = 0$ für alle $i \in \mathbb{N}_0$, also

$$\sum_{i=0}^{\infty} \lim_{n\to\infty} \delta_{in} = 0.$$

Andererseits gilt $\sum_{i=0}^{\infty} \delta_{in} = 1$ für alle $n \in \mathbb{N}_0$, also

$$\lim_{n\to\infty} \sum_{i=0}^{\infty} \delta_{in} = 1 \neq 0 = \sum_{i=0}^{\infty} \lim_{n\to\infty} \delta_{in}.$$

Unter Zusatzvoraussetzungen kann man Limes und unendliche Summe dennoch vertauschen. Wir beweisen zwei Sätze hierzu, den *Satz von der dominierten Konvergenz* für Reihen und den *Satz von der monotonen Konvergenz* für Reihen.

Der Satz von der dominierten Konvergenz.

Satz 3.23 (Satz von der dominierten Konvergenz für Reihen)
Es sei $(a_n^{(i)})_{n,i\in\mathbb{N}_0}$ eine Doppelfolge, d.h. eine Folge von Folgen in $\mathbb{C}$. Wir nehmen an:

1. *Für alle $i \in \mathbb{N}_0$ existiere $\lim_{n\to\infty} a_n^{(i)}$ in $\mathbb{C}$.*

2. *Es existieren $b^{(i)} \geq 0$, $i \in \mathbb{N}_0$, mit $\sum_{i=0}^{\infty} b^{(i)} < +\infty$, so daß für alle $n, i \in \mathbb{N}_0$ gilt: $|a_n^{(i)}| \leq b^{(i)}$.*

Dann existiert $\lim_{n\to\infty} \sum_{i=0}^{\infty} a_n^{(i)}$ in $\mathbb{C}$, und es gilt

$$\lim_{n\to\infty} \sum_{i=0}^{\infty} a_n^{(i)} = \sum_{i=0}^{\infty} \lim_{n\to\infty} a_n^{(i)}.$$

wichtigster Satz zur Vertauschung von Limes und unendlicher Summe!

Bemerkungen:

1. Sprechweise für die Voraussetzung 2.:
"$(b^{(i)})_{i\in\mathbb{N}_0}$ ist eine summierbare Majorante der $a_n^{(i)}$, $i, n \in \mathbb{N}_0$".

2. Man beachte, dass die Majorante $(b^{(i)})_{i\in\mathbb{N}_0}$ nur *einen* Index i besitzt. Die Majorante darf also *nicht* von n abhängen!

3. Der Satz ist ein Spezialfall eines gleichnamigen, aber viel allgemeineren Satzes von Lebesgue aus der Integrationstheorie, den wir erst in der Maßtheorie behandeln werden.

4. Manchmal wird der Satz auch "Satz von der *majorisierten* Konvergenz" genannt.

Beweis des Satzes von der dominierten Konvergenz: Nach dem Majorantenkriterium existiert $\sum_{i=0}^{\infty} a_n^{(i)} \in \mathbb{C}$ für alle $n \in \mathbb{N}_0$. Weiter gilt für alle $i \in \mathbb{N}_0$:

$$\left|\lim_{n\to\infty} a_n^{(i)}\right| \leq b^{(i)},$$

da[28] für alle $i, n \in \mathbb{N}_0$ gilt: $|a_n^{(i)}| \leq b^{(i)}$. Also ist nach dem Majorantenkriterium

$$\sum_{i=0}^{\infty} \lim_{n\to\infty} a_n^{(i)}$$

absolut konvergent, da

$$\sum_{i=0}^{\infty} b^{(i)} < +\infty.$$

[28] Bitte überlegen Sie sich, warum das gilt.

Wir zeigen nun die Behauptung

$$\sum_{i=0}^{\infty} a_n^{(i)} \overset{n\to\infty}{\longrightarrow} \sum_{i=0}^{\infty} \lim_{k\to\infty} a_k^{(i)},$$

also

$$\forall \epsilon > 0 \; \exists m \in \mathbb{N}_0 \; \forall n \geq m : \left| \sum_{i=0}^{\infty} a_n^{(i)} - \sum_{i=0}^{\infty} \lim_{k\to\infty} a_k^{(i)} \right| < \epsilon.$$

Hierzu sei $\varepsilon > 0$ gegeben. Wir wählen $N \in \mathbb{N}_0$ so groß, dass gilt:

$$\frac{\varepsilon}{3} > \left| \sum_{i=0}^{N} b^{(i)} - \sum_{i=0}^{\infty} b^{(i)} \right| = \sum_{i=N+1}^{\infty} b^{(i)}.$$

Nun gilt

$$\lim_{n\to\infty} \sum_{i=0}^{N} a_n^{(i)} = \sum_{i=0}^{N} \lim_{n\to\infty} a_n^{(i)}.$$

Hier vertauschen wir nämlich nur eine *endliche* Summe mit dem Limes. Wir können also ein $m \in \mathbb{N}_0$ wählen, so daß für alle $n \geq m$ gilt:

$$\left| \sum_{i=0}^{N} a_n^{(i)} - \sum_{i=0}^{N} \lim_{k\to\infty} a_k^{(i)} \right| < \frac{\varepsilon}{3}.$$

Es folgt für diese n:

$$\begin{aligned}
&\left| \sum_{i=0}^{\infty} a_n^{(i)} - \sum_{i=0}^{\infty} \lim_{k\to\infty} a_k^{(i)} \right| \\
&\leq \left| \sum_{i=0}^{\infty} a_n^{(i)} - \sum_{i=0}^{N} a_n^{(i)} \right| + \left| \sum_{i=0}^{N} a_n^{(i)} - \sum_{i=0}^{N} \lim_{k\to\infty} a_k^{(i)} \right| + \left| \sum_{i=0}^{N} \lim_{k\to\infty} a_k^{(i)} - \sum_{i=0}^{\infty} \lim_{k\to\infty} a_k^{(i)} \right| \\
&\leq \left| \sum_{i=N+1}^{\infty} a_n^{(i)} \right| + \frac{\varepsilon}{3} + \left| \sum_{i=N+1}^{\infty} \lim_{k\to\infty} a_k^{(i)} \right| \\
&\leq \sum_{i=N+1}^{\infty} b^{(i)} + \frac{\varepsilon}{3} + \sum_{i=N+1}^{\infty} b^{(i)} \qquad \text{(wegen } |a_n^{(i)}| \leq b^{(i)} \text{ und } \lim_{k\to\infty} |a_k^{(i)}| \leq b^{(i)}\text{)} \\
&< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon,
\end{aligned}$$

also die Behauptung.

□

Beispiel: Für alle $x \in \mathbb{C}$ gilt:

$$\boxed{\left(1+\frac{x}{n}\right)^n \stackrel{n\to\infty}{\longrightarrow} \exp(x)}$$

Beweis: Nach der binomischen Formel gilt:

$$\begin{aligned}
\left(1+\frac{x}{n}\right)^n &= \sum_{k=0}^{n}\binom{n}{k}\left(\frac{x}{n}\right)^k 1^{n-k} \\
&= \sum_{k=0}^{\infty}\binom{n}{k}\left(\frac{x}{n}\right)^k \qquad \left(\text{weil } \tbinom{n}{k}=0 \text{ für } k>n\right) \\
&= \sum_{k=0}^{\infty}\frac{x^k}{k!}\prod_{l=0}^{k-1}\frac{n-l}{n}.
\end{aligned}$$

Die verwendete Formel für die Summanden

$$\binom{n}{k}\left(\frac{x}{n}\right)^k = \frac{x^k}{k!}\prod_{l=0}^{k-1}\frac{n-l}{n}$$

folgt unmittelbar aus der Definition des Binomialkoeffizienten.
Wir wollen nun zeigen, dass

$$\lim_{n\to\infty}\sum_{k=0}^{\infty}\frac{x^k}{k!}\prod_{l=0}^{k-1}\frac{n-l}{n} = \sum_{k=0}^{\infty}\lim_{n\to\infty}\frac{x^k}{k!}\prod_{l=0}^{k-1}\frac{n-l}{n}$$

gilt. Wir berechnen dazu zuerst den Grenzwert der Summanden für $n\to\infty$. Es gilt für alle $l\in\mathbb{N}_0$:

$$\frac{n-l}{n} = 1-\frac{l}{n} \stackrel{n\to\infty}{\longrightarrow} 1,$$

also für alle $k\in\mathbb{N}_0$:

$$\prod_{l=0}^{k-1}\frac{n-l}{n} \stackrel{n\to\infty}{\longrightarrow} 1.$$

(Dies gilt auch für $k=0$, denn das leere Produkt ist gleich 1.)
Es folgt für alle $k\in\mathbb{N}_0$:

$$\frac{x^k}{k!}\prod_{l=0}^{k-1}\frac{n-l}{n} \stackrel{n\to\infty}{\longrightarrow} \frac{x^k}{k!}.$$

Um den Satz von der majorisierten Konvergenz anwenden zu können, zeigen wir nun die Existenz einer summierbaren Majorante. Es gilt für alle $n,k\in\mathbb{N}_0$ mit $n\geq 1$:

- Falls $k \leq n$:
$$\left| \frac{x^k}{k!} \prod_{l=0}^{k-1} \frac{n-l}{n} \right| = \frac{|x|^k}{k!} \prod_{l=0}^{k-1} \left| 1 - \frac{l}{n} \right| \leq \frac{|x|^k}{k!}.$$
Man beachte, dass die Faktoren $\left|1 - \frac{l}{n}\right|$ hier zwischen 0 und 1 liegen.

- Falls $k > n$:
$$\left| \frac{x^k}{k!} \prod_{l=0}^{k-1} \frac{n-l}{n} \right| = 0 \leq \frac{|x|^k}{k!},$$
denn das Produkt enthält hier einen Faktor 0, nämlich für $l = n$.

Die Reihen
$$\sum_{k=0}^{\infty} \frac{x^k}{k!} \prod_{l=0}^{k-1} \frac{n-l}{n}, \qquad n \in \mathbb{N},$$
werden also durch die Exponentialreihe
$$\sum_{k=0}^{\infty} \frac{|x|^k}{k!}$$
majorisiert, und alle Summanden konvergieren für $n \to \infty$ gegen die Summanden der Exponentialreihe $\sum_{k=0}^{\infty} x^k/k!$. Aus dem Satz von der dominierten Konvergenz folgt
$$\begin{aligned}
\lim_{n\to\infty} \left(1 + \frac{x}{n}\right)^n &= \lim_{n\to\infty} \sum_{k=0}^{\infty} \frac{x^k}{k!} \prod_{l=0}^{k-1} \frac{n-l}{n} \\
&= \sum_{k=0}^{\infty} \lim_{n\to\infty} \frac{x^k}{k!} \prod_{l=0}^{k-1} \frac{n-l}{n} \\
&= \sum_{k=0}^{\infty} \frac{x^k}{k!} = \exp(x).
\end{aligned}$$
□

Interpretation: Eine Bank bietet einen Zinssatz $x = x \cdot 100\%$ pro Jahr an. Bei halbjährlicher Zinszahlung wächst das Kapital (inklusive Zinseszins) um den Faktor $(1 + \frac{x}{2})^2$ statt $1 + x$, bei dritteljährlicher Zinszahlung um den Faktor $(1 + \frac{x}{3})^3$, etc. Im Limes kontinuierlicher Zinszahlung erhält man den Faktor $\lim_{n\to\infty}(1 + \frac{x}{n})^n = \exp(x)$, Zinseszinsen eingerechnet.

Zum Beispiel: Für $x = 100\%$ vervielfacht sich bei kontinuierlicher Zinszahlung inklusive Zinseszins das Kapital um den Faktor
$$\boxed{e := \exp(1) = 2,71828\ldots}$$
Eulersche Zahl

(statt des Faktors 2 bei einmaliger Zinszahlung). Die Zahl e heißt *Eulersche Zahl*.

Der Satz von der monotonen Konvergenz. Nun besprechen wir einen weiteren Satz zur Vertauschung von Limes und unendlicher Summe, der die Dominiertheitsbedingung durch eine Monotoniebedingung ersetzt. Für die Formulierung des Satzes ist es zweckmäßig, die Addition $+ : \mathbb{R}\times\mathbb{R}\to\mathbb{R}$ auf $+ : [(\mathbb{R}\cup\{\pm\infty\})\times(\mathbb{R}\cup\{\pm\infty\})]\setminus\{(+\infty,-\infty),(-\infty,+\infty)\}$ durch die Festsetzungen $a+(+\infty) = +\infty + a = +\infty$ für $a\in\mathbb{R}\cup\{+\infty\}$ und $b+(-\infty) = -\infty + b = -\infty$ für $b\in\mathbb{R}\cup\{-\infty\}$ zu erweitern.

Satz 3.24 (Satz von der monotonen Konvergenz für Reihen)
Es sei $(a_n^{(i)})_{n,i\in\mathbb{N}_0}$ eine Doppelfolge in $[0,+\infty]$, so daß für alle $i\in\mathbb{N}_0$ die Folge $(a_n^{(i)})_{n\in\mathbb{N}_0}$ monoton steigt. Dann folgt

$$\lim_{n\to\infty}\sum_{i=0}^{\infty} a_n^{(i)} = \sum_{i=0}^{\infty}\lim_{n\to\infty} a_n^{(i)}, \tag{44}$$

wobei die Grenzwerte in $[0,+\infty]$ zu verstehen sind.

Auch dieser Satz hat eine weitgehende Verallgemeinerung in der Lebesgueschen Integrationstheorie, die wir erst in der Maßtheorie besprechen.
Beweis des Satzes: Man beachte, daß alle Grenzwerte im folgenden Beweis in $[0,+\infty]$ wegen Lemma 3.22 existieren. Weil $a_n^{(i)}$ in n monoton steigt, gilt

$$a_n^{(i)} \le \lim_{k\to\infty} a_k^{(i)}$$

für alle $n,i\in\mathbb{N}_0$, also

$$\sum_{i=0}^{m} a_n^{(i)} \le \sum_{i=0}^{m}\lim_{k\to\infty} a_k^{(i)}$$

für alle $n,m\in\mathbb{N}_0$ und folglich

$$\sum_{i=0}^{\infty} a_n^{(i)} \le \sum_{i=0}^{\infty}\lim_{k\to\infty} a_k^{(i)}$$

für alle $n\in\mathbb{N}_0$ und schließlich

$$\lim_{n\to\infty}\sum_{i=0}^{\infty} a_n^{(i)} \le \sum_{i=0}^{\infty}\lim_{n\to\infty} a_n^{(i)}. \tag{45}$$

Umgekehrt: Es sei

$$x < \sum_{i=0}^{\infty}\lim_{n\to\infty} a_n^{(i)}.$$

Dann gibt es $m\in\mathbb{N}_0$ mit

$$
\begin{aligned}
x &< \sum_{i=0}^{m} \lim_{n\to\infty} a_n^{(i)} \\
&= \lim_{n\to\infty} \sum_{i=0}^{m} a_n^{(i)} && \text{(wegen der Vertauschbarkeit von lim und \textit{endlicher} Summe)} \\
&\le \lim_{n\to\infty} \sum_{i=0}^{\infty} a_n^{(i)} && \text{(weil } \textstyle\sum_{i=0}^{m} a_n^{(i)} \le \sum_{i=0}^{\infty} a_n^{(i)} \text{ für alle } m \in \mathbb{N}_0\text{),}
\end{aligned}
$$

wobei wir $a_n^{(i)} \ge 0$ verwendeten. Man beachte, dass hier Limes und *endliche* Summe auch dann vertauschen, wenn Werte $+\infty$ auftreten. Wir schließen

$$
\lim_{n\to\infty} \sum_{i=0}^{\infty} a_n^{(i)} \ge \sum_{i=0}^{\infty} \lim_{n\to\infty} a_n^{(i)}. \tag{46}
$$

Die Aussagen (45) und (46) zusammen implizieren die Behauptung.

$\square$

Ausblick. Wie schon oben angedeutet, besitzen der Satz von der monotonen Konvergenz für Reihen und der Satz von der dominierten Konvergenz für Reihen weitgehende Verallgemeinerungen in der Integrationstheorie. Diese Verallgemeinerungen geben hinreichende Bedingungen dafür an, dass man Limes und Integral vertauschen kann, dass also

$$
\lim_{n\to\infty} \int_{\Omega} f_n(x)\,dx = \int_{\Omega} \lim_{n\to\infty} f_n(x)\,dx
$$

gilt. Zwei hinreichende Bedingungen hierfür sind (neben Voraussetzungen, die die Existenz der Integrale garantieren):

- **monotone Konvergenz:** Es gilt $f_n \ge 0$ für alle $n \in \mathbb{N}$, und $(f_n(x))_{n\in\mathbb{N}}$ steigt für alle $x \in \Omega$ monoton an;
- **dominierte Konvergenz:** Für alle $x \in \Omega$ konvergiert $(f_n(x))_{n\in\mathbb{N}}$, und es existiert eine Funktion $g : \Omega \to \mathbb{R}_0^+$ mit endlichem Integral

  $$
  \int_{\Omega} g(x)\,dx < \infty
  $$

  so dass $|f_n| \le g$ für alle $n \in \mathbb{N}$ gilt.

Allerdings ist das hier verwendete Integral *nicht* das Riemann-Integral, das wir in diesem Semester besprechen, sondern das Lebesgue-Integral oder allgemeiner das Integral bezüglich eines Maßes, das wir erst im dritten Semester besprechen. Auch absolut konvergente Reihen oder auch Reihen mit beliebigen nichtnegativen Summanden sind Spezialfälle dieses allgemeinen Integrals.

3.5.2 Umordnung von Reihen

Satz 3.25 (Umordnung absolut konvergenter oder nichtnegativer Reihen) *Es sei* $(a_n)_{n\in\mathbb{N}_0}$ *eine Folge in* $\mathbb{C}$, *so daß* $\sum_{n=0}^{\infty}|a_n| < +\infty$, *oder* $(a_n)_{n\in\mathbb{N}_0}$ *eine Folge in* $[0,+\infty]$. *Dann gilt für jede Bijektion* $J\colon \mathbb{N}_0 \to \mathbb{N}_0$*:*

$$\sum_{i=0}^{\infty} a_{J(i)} = \sum_{n=0}^{\infty} a_n.$$

Zum Beweis ist die folgende Notation nützlich: Für eine Aussage φ, in der Variablen $x_1,\ldots,x_l$ frei vorkommen dürfen, definieren wir die *Indikatorfunktion* von φ:

$$1_{\{\varphi\}} = \begin{cases} 1, \text{ falls } \varphi \text{ gilt} \\ 0, \text{ falls } \varphi \text{ nicht gilt.} \end{cases}$$

Zum Beispiel können wir das Kronecker-Delta damit so schreiben: $\delta_{ij} = 1_{\{i=j\}}$.

Beweis des Satzes: Wir beweisen den Satz unter der Voraussetzung $\sum_{n=0}^{\infty}|a_n| < \infty$. Es gilt:

$$a_n 1_{\{J^{-1}(n)\le j\}} \overset{j\to\infty}{\longrightarrow} a_n$$

für alle $n \in \mathbb{N}_0$, denn für alle genügend großen j gilt sogar

$$a_n 1_{\{J^{-1}(n)\le j\}} = a_n.$$

Weiter haben die Folgen $\left(a_n 1_{\{J^{-1}(n)\le j\}}\right)_{n\in\mathbb{N}_0}$, $j \in \mathbb{N}_0$, die summierbare Majorante $(|a_n|)_{n\in\mathbb{N}_0}$, denn es gilt

$$\left|a_n 1_{\{J^{-1}(n)\le j\}}\right| \le |a_n|.$$

Es folgt nach dem Satz von der dominierten Konvergenz:

$$\sum_{i=0}^{j} a_{J(i)} = \sum_{n=0}^{\infty} a_n 1_{\{J^{-1}(n)\le j\}} \overset{j\to\infty}{\longrightarrow} \sum_{n=0}^{\infty} \lim_{j\to\infty} \left(a_n 1_{\{J^{-1}(n)\le j\}}\right) = \sum_{n=0}^{\infty} a_n.$$

Um das erste Gleichheitszeichen in dieser Formel einzusehen, beachte man: In der unendlichen Summe rechts vom Gleichheitszeichen tauchen gerade die $a_{J(i)}$ mit $i = 0,\ldots,j$ als Summanden auf, für die $1_{\{J^{-1}(n)\le j\}}$ nicht gleich 0 ist.

Für $a_n \ge 0$ beweist man den Satz analog; statt dominierter Konvergenz verwendet man montone Konvergenz.

$\square$

Im allgemeinen kann man unendliche Reihen jedoch *nicht* beliebig umordnen. Das zeigt der folgende Satz:

Satz 3.26 *Die Reihe* $\sum_{n=0}^{\infty} a_n$ *in* $\mathbb{R}$ *sei konvergent, aber nicht absolut konvergent. Dann gibt es zu jedem* $x \in \mathbb{R}$ *eine Bijektion* $J : \mathbb{N}_0 \to \mathbb{N}_0$*, so daß gilt:*

$$\sum_{i=0}^{\infty} a_{J(i)} = x.$$

Beispiel: Der konvergenten, aber nicht absolut konvergenten Reihe $\sum_{n=1}^{\infty} \frac{(-1)^n}{n}$ kann man also durch Umordnen der Summanden jeden beliebigen Wert geben, z.B. $0, \pi, -37, \ldots$.

Beweisskizze zum Satz: Wir unterteilen die Summanden in zwei Hälften, die negativen und die nichtnegativen Summanden:

$$\begin{aligned} &b_0, b_1, b_2, \ldots \text{ seien die negativen der } a_n,\ n \in \mathbb{N}_0; \\ &c_0, c_1, c_2, \ldots \text{ seien die positiven der } a_n,\ n \in \mathbb{N}_0, \text{ inklusive } 0. \end{aligned}$$

Weil $\sum_{n=0}^{\infty} a_n$ konvergiert, aber nicht absolut konvergiert, gilt sowohl $\sum_{k=0}^{\infty} b_k = -\infty$ als auch $\sum_{l=0}^{\infty} c_l = +\infty$. Gegeben ein beliebiger Wert $x \in \mathbb{R}$, definieren wir rekursiv eine umgeordnete Version $\sum_{n=0}^{\infty} d_n$ von $\sum_{n=0}^{\infty} a_n$ wie folgt:

Rekursionsanfang: Wir setzen

$$d_0 = \begin{cases} b_0, & \text{falls } x < 0 \\ c_0, & \text{falls } x \geq 0 \end{cases}$$

Rekursionsschritt $n - 1 \rightsquigarrow n$*:* Wenn wir $d_0, \ldots, d_{n-1}$ schon gewählt haben, so setzen wir

- $d_n =$ das erste noch nicht verwendete b_k, falls $\sum_{k=0}^{n-1} d_k \geq x$,
- $d_n =$ das erste noch nicht verwendete c_l, falls $\sum_{k=0}^{n-1} d_k < x$,

Das bedeutet: Wir verwenden negative Summanden, wenn die Summe bisher oberhalb x liegt, und positive Summanden, wenn die Summe bisher unterhalb x liegt. Man erhält: $\sum_{n=0}^{\infty} d_n = x$. Wir führen die Details hier nicht näher aus.

$\square$

Notation: Es sei $(a_i)_{i \in I}$ eine Familie reeller oder komplexer Zahlen. Die Menge I sei abzählbar. Falls $I = \{i_1, \ldots, i_k\}$ endlich ist, setzen wir

$$\sum_{i \in I} a_i = \sum_{n=1}^{k} a_{i_n}.$$

Dieser Wert hängt nicht von der Aufzählung von I ab. Falls I abzählbar unendlich ist, wählen wir eine beliebige Bijektion $J : \mathbb{N}_0 \to I$ und setzen

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

Dies ist unabhängig von der Wahl von J.
Falls

$$\sum_{i\in I} |a_i| < \infty,$$

so setzen wir

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

Nach dem Umordnungssatz ist dies ebenfalls unabhängig von der Wahl von J.

Der folgende Satz zeigt, dass man bei absolut konvergenten Reihen die Summanden in beliebige Teile aufteilen darf, die einzelnen Teile aufsummieren, und dann die Summen nochmals aufsummieren, ohne die Gesamtsumme dabei zu ändern:

Satz 3.27 (Großer Umordnungssatz) *Es sei I eine abzählbare Indexmenge und $I = \bigcup_{k\in K} I_k$ eine Zerlegung von I in paarweise durchschnittsfremde Mengen, indiziert mit einer abzählbaren Indexmenge K. Es sei $(a_i)_{i\in I}$ eine Familie, für die (mindestens) eine der folgenden beiden Voraussetzungen gilt:*

1. *Für alle $i \in I$ gelte $0 \le a_i \le +\infty$.*
2. *Oder: Für alle $i \in I$ gelte $a_i \in \mathbb{C}$, und $\sum_{i\in I} |a_i| < \infty$.*

Dann folgt:

$$\sum_{i\in I} a_i = \sum_{k\in K} \sum_{i\in I_k} a_i.$$

Beweis: Der Satz ist nur für unendliche I und K nichttrivial; wir setzen also voraus, daß I und K abzählbar unendlich sind. Durch Umbenennung der Indices können wir annehmen: $I = \mathbb{N}_0$ und $K = \mathbb{N}_0$. Es sei

$$a_i^{(n)} = a_i 1_{\{i \in I_k \text{ für ein } k \le n\}} = \sum_{k=0}^{n} a_i 1_{\{i\in I_k\}}.$$

Man beachte, dass in der letzten Summe höchstens ein Summand von 0 verschieden ist.
Wir nehmen nun $0 \le a_i \le +\infty$ für alle $i \in I$ an. Dann folgt:

$$\sum_{i\in I} a_i^{(n)} = \sum_{i\in I} \sum_{k=0}^{n} a_i 1_{\{i\in I_k\}} = \sum_{k=0}^{n} \sum_{i\in I} a_i 1_{\{i\in I_k\}} = \sum_{k=0}^{n} \sum_{i\in I_k} a_i.$$

Hier haben wir nur eine unendliche Summation mit einer *endlichen* Summe vertauscht.
Weiter gilt für alle $i \in I$:

$$a_i^{(n)} \nearrow a_i \quad \text{für } n \to \infty$$

d.h.: $a_i^{(n)}$ konvergiert *monoton steigend* für $n \to \infty$ gegen a_i. Es folgt nach dem Satz von der monotonen Konvergenz:

$$\sum_{i\in I} a_i = \sum_{i\in I} \lim_{n\to\infty} a_i^{(n)} = \lim_{n\to\infty} \sum_{i\in I} a_i^{(n)} = \lim_{n\to\infty} \sum_{k=0}^{n} \sum_{i\in I_k} a_i = \sum_{k\in K} \sum_{i\in I_k} a_i.$$

Falls $a_i \in \mathbb{C}$ für alle $i \in I$ und $\sum_{i\in I} |a_i| < \infty$ gilt, so folgt:

$$\begin{aligned}\sum_{i\in I} a_i &= \sum_{i\in I} \lim_{n\to\infty} a_i^{(n)} = \lim_{n\to\infty} \sum_{i\in I} a_i^{(n)} = \lim_{n\to\infty} \sum_{i\in I} \sum_{k=0}^{n} a_i 1_{\{i\in I_k\}} \\ &= \lim_{n\to\infty} \sum_{k=0}^{n} \sum_{i\in I} a_i 1_{\{i\in I_k\}} = \lim_{n\to\infty} \sum_{k=0}^{n} \sum_{i\in I_k} a_i = \sum_{k\in K} \sum_{i\in I_k} a_i,\end{aligned}$$

wobei wir den Satz von der dominierten Konvergenz zur Vertauschung von unendlicher Summe und Limes verwendet haben. Die Voraussetzungen des Satzes von der dominierten Konvergenz sind wegen der vorausgesetzten Summierbarkeit der $(|a_i|)_{i\in I}$ erfüllt, denn einerseits gilt $|a_i^{(n)}| \le |a_i|$ für alle $i \in I$, $n \in \mathbb{N}_0$, und andererseits $a_i^{(n)} \overset{n\to\infty}{\longrightarrow} a_i$ für alle $i \in I$.

□

Als Beispiel beweisen wir die Funktionalgleichung der Exponentialfunktion:

Satz 3.28 *Für alle $x, y \in \mathbb{C}$ gilt:*

$$\boxed{\exp(x+y) = \exp(x) \cdot \exp(y)}$$

Wichtigste Rechenregel für exp!

Beweis. Wir bemerken zunächst für $n \in \mathbb{N}_0$, $k = 0, \ldots, n$:

$$\binom{n}{k} = \frac{\prod_{j=n-k+1}^{n} j}{k!} = \frac{n!}{k!(n-k)!}.$$

In der folgenden Rechnung ist die Umordnung aufgrund der absoluten Konvergenz der Exponentialreihen erlaubt:

$$\begin{aligned}\exp(x+y) &= \sum_{n=0}^{\infty} \frac{(x+y)^n}{n!} = \sum_{n=0}^{\infty} \sum_{k=0}^{n} \frac{1}{n!}\binom{n}{k} x^k y^{n-k} \\ &= \sum_{n=0}^{\infty} \sum_{k=0}^{n} \frac{x^k}{k!} \frac{y^{n-k}}{(n-k)!} = \sum_{n=0}^{\infty} \sum_{\substack{k,l\in\mathbb{N}_0 \\ k+l=n}} \frac{x^k}{k!} \frac{y^l}{l!} \\ &= \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \frac{x^k}{k!} \frac{y^l}{l!} = \sum_{k=0}^{\infty} \frac{x^k}{k!} \sum_{l=0}^{\infty} \frac{y^l}{l!} \\ &= \exp(x)\exp(y).\end{aligned}$$

Aufgrund der Funktionalgleichung der Exponentialfunktion schreibt man auch e^x statt $\exp(x)$, $x \in \mathbb{C}$.

Korollar 3.29 *Für alle $y \in \mathbb{R}$ ist $e^{-iy} = \overline{e^{iy}}$ und $|e^{iy}| = 1$.*

Beweis: Es gilt

$$e^{-iy} = \sum_{n=0}^{\infty} \frac{(-iy)^n}{n!} = \sum_{n=0}^{\infty} \frac{\overline{iy}^n}{n!} = \sum_{n=0}^{\infty} \frac{\overline{(iy)^n}}{n!} = \overline{\sum_{n=0}^{\infty} \frac{(iy)^n}{n!}} = \overline{e^{iy}},$$

wobei wir verwendet haben, dass für komplexe Folgen $(a_n)_n$ aus $a_n \overset{n\to\infty}{\longrightarrow} a$ folgt: $\overline{a_n} \overset{n\to\infty}{\longrightarrow} \overline{a}$.
Wir erhalten:

$$|e^{iy}|^2 = e^{iy}\overline{e^{iy}} = e^{iy}e^{-iy} = e^{iy-iy} = e^0 = 1,$$

also $|e^{iy}| = 1$ wegen $|e^{iy}| > 0$.

3.6 Stetigkeit

3.6.1 Definition und Charakterisierung der Stetigkeit

„Stetigkeit" ist ein Grundbegriff der Analysis und Topologie. Er wird ähnlich wie Konvergenz definiert:

Definition 3.30 Es sei $M \subseteq \mathbb{R}$ (oder $M \subseteq \mathbb{C}$) und $x \in M$. Eine Funktion $f\colon M \to \mathbb{R}$ (bzw. $\mathbb{C}$) heißt *stetig* in x, wenn gilt:

$$\forall \varepsilon > 0 \quad \exists \delta > 0 \quad \forall y \in M : \ (|y - x| < \delta \Rightarrow |f(y) - f(x)| < \varepsilon)$$

zentraler Grundbegriff!

Anders gesagt: "f ist stetig in x" bedeutet: Für alle $\varepsilon > 0$ gibt es $\delta > 0$, so daß

$$f[U_\delta(x) \cap M] \subseteq U_\varepsilon(f(x))$$

Nochmal anders gesagt: Für jede ε-Umgebung von $f(x)$ gilt: Wenn nur $y \in M$ genügend nahe bei x liegt, gilt $f(y) \in U_\varepsilon(f(x))$.
Nochmal anders gesagt: Es gibt eine Funktion $\Delta\colon]0,\infty[\to]0,\infty[$, so daß für alle $\varepsilon > 0$ gilt: $f[U_{\Delta(\varepsilon)}(x) \cap M] \subseteq U_\varepsilon(f(x))$.

Beispiel:

1. Die Funktion $f\colon \mathbb{R} \to \mathbb{R}, f(x) = 2x$, ist stetig in jedem Punkt $x \in \mathbb{R}$. Wir müssen also zeigen:

   $$\forall \varepsilon > 0 \, \exists \delta > 0 \, \forall y \in \mathbb{R} : (|y - x| < \delta \ \Rightarrow |f(y) - f(x)| < \varepsilon).$$

   In der Tat:

   Es sei $\varepsilon > 0$. Wir wählen $\delta = \varepsilon/2 > 0$. Dann gilt für alle $y \in \mathbb{R}$ mit $|y - x| < \delta$:

   $$|f(y) - f(x)| = |2y - 2x| = 2|y - x| < 2\delta = \varepsilon.$$

Metabemerkung zum Beweisaufbau:

Beachten Sie, dass der Beweis parallel zur zu beweisenden Formel aufgebaut ist: Wir arbeiten die Quantorenkette "$\forall\varepsilon\ \exists\delta\ \forall y$" von links nach rechts ab:

(a) Für den Allquantor "$\forall\varepsilon$" geben wir uns ε vor: "*Es sei* $\varepsilon > 0$".

(b) Für den Existenzquantor "$\exists\delta$" müssen wir dann δ (von ε abhängig) angeben: "*Wir wählen* $\delta = \ldots$". Welche Wahl von δ zweckmäßig ist, sieht man erst im Nachhinein, wenn man die Abschätzungen schon ausgeführt hat.

(c) Den Allquantor "$\forall y$" und die Prämisse der Implikation behandeln wir mit "*Dann gilt für alle* $y \in \mathbb{R}$ *mit* $|y - x| < \delta \ldots$". Die Prämisse $|y - x| < \delta$ wird also als eine Annahme über y behandelt.

(d) Erst jetzt beginnen wir mit den Abschätzungen unter Verwendung der Annahme "$|y - x| < \delta$" über y: "$|f(y) - f(x)| = \ldots < 2\delta$". Jetzt – also erst am Schluss des Beweises – sieht man, welche Wahl von δ zweckmäßig ist: Wir wollen $|f(y)-f(x)| < \varepsilon$ erreichen; somit bietet sich an, δ so zu wählen, dass $2\delta = \varepsilon$ gilt. Somit ist $\delta = \varepsilon/2$ im Schritt (b) eine gute Wahl.

2. Die Funktion

$$g\colon \mathbb{R} \to \mathbb{R}, \qquad g(x) = \begin{cases} x, & \text{falls } x \in \mathbb{Q} \\ -x, & \text{falls } x \in \mathbb{R} \setminus \mathbb{Q} \end{cases}$$

ist stetig in 0, aber unstetig in allen anderen $x \in \mathbb{R} \setminus \{0\}$.

Beweis hierzu: Gegeben $\varepsilon > 0$, setzen wir $\delta = \varepsilon$.

Dann gilt für alle $y \in U_\delta(0)\colon |g(y) - g(0)| = |g(y)| = |y| < \delta = \varepsilon$. Also ist g stetig in 0.

Für alle $x \neq 0$ zeigen wir nun:

$$\exists \varepsilon > 0 \quad \forall \delta > 0 \quad \exists y \in \mathbb{R} : (|y - x| < \delta \wedge |g(y) - g(x)| \geq \varepsilon)$$

also Unstetigkeit von g in x:

Wir setzen $\varepsilon = |x| > 0$. [29] Es sei $\delta > 0$.[30] Wir setzen $\delta' = \min\{\varepsilon, \delta\} > 0$. Falls $x \in \mathbb{Q}$, wählen wir ein beliebiges $y \in U_{\delta'}(x) \cap (\mathbb{R} \setminus \mathbb{Q})$, und falls $x \in \mathbb{R} \setminus \mathbb{Q}$, wählen wir ein beliebiges $y \in U_{\delta'}(x) \cap \mathbb{Q}$. [31]

Insbesondere haben x und y das gleiche Vorzeichen, aber $g(x)$ und $g(y)$ haben verschiedene Vorzeichen.

Es folgt: $y \in U_\delta(x)$, weil $\delta \geq \delta'$, und

$$|g(y) - g(x)| \geq |g(x)| = |x| = \varepsilon.$$

Also ist g unstetig in x.

[29] *Metabemerkung:* Hier wird der Existenzquantor $\exists\varepsilon$ behandelt.

[30] *Metabemerkung:* Das ist die Behandlung des Allquantors $\forall\delta$.

[31] *Metabemerkung:* Das ist die Behandlung des Existenzquantors $\exists y$. Die Variable δ' ist nur eine Hilfsgröße zur Wahl von y.

Das $\varepsilon - \delta$-Spiel Veranschaulichen wir die Stetigkeitsdefinition wieder mit einem Spiel zwischen Proponenten und Opponenten. Der Proponent verteidigt die Stetigkeit, der Opponent versucht sie zu widerlegen.

Proponent: $f\colon x \mapsto 2x$, $x \in \mathbb{R}$ ist stetig in 1.

Opponent: Das glaube ich nicht: Nimm $\varepsilon = 0.1$!

Proponent: Ich wähle $\delta = 0.05$.

Opponent: Hm, ich finde kein $y \in]0.95; 1.05[$ mit $|2y - 2 \cdot 1| \geq 0.1$. Doch nimm nun $\varepsilon = 0.01$!

Proponent: Ich wähle $\delta = 0.005$.

Opponent: Wieder kann ich kein $y \in]0.995; 1.005[$ mit $|2y - 2 \cdot 1| \geq 0.01$ finden.

$\vdots$

Hier wird der Proponent niemals widerlegt.

Satz 3.31 (Charakterisierung der Stetigkeit in einem Punkt) Es sei $M \subseteq \mathbb{R}$ oder $M \subseteq \mathbb{C}$.
Folgende Aussagen für eine Funktion $f\colon M \to \mathbb{R}$ (bzw. $\mathbb{C}$) und $x \in M$ sind äquivalent:

a) f ist stetig in x.

b) Für jede offene Umgebung U von $f(x)$ enthält $f^{-1}[U]$ eine offene Umgebung von x in M. (D.h.: Es gibt eine offene Umgebung V von x, so daß $V \cap M \subseteq f^{-1}[U]$.)

c) "f ist *folgenstetig* in x": Für jede Folge $(a_n)_{n\in\mathbb{N}_0}$ in M mit $a_n \overset{n\to\infty}{\longrightarrow} x$ gilt $f(a_n) \overset{n\to\infty}{\longrightarrow} f(x)$.

d) *„f reißt den Berührpunkt x nicht ab“:* Für jede Menge $N \subseteq M$ gilt: Wenn x Berührpunkt von N ist, dann ist $f(x)$ Berührpunkt von $f[N]$.

Beweis:
a) $\Rightarrow$ b):
Es sei f stetig in x, und es sei U eine offene Umgebung von $f(x)$. Dann gibt es $\varepsilon > 0$ mit $U_\varepsilon(f(x)) \subseteq U$. Zu solch einem $\varepsilon > 0$ wählen wir $\delta > 0$ mit $f[U_\delta(x) \cap M] \subseteq U_\varepsilon(f(x))$ und setzen $V = U_\delta(x)$. Es folgt $f[V \cap M] \subseteq U$, also $V \cap M \subseteq f^{-1}[U]$, da $V \cap M$ im Definitionsbereich von f enthalten ist.

b) $\Rightarrow$ c)
Es gelte b), und es sei $(a_n)_{n\in\mathbb{N}_0}$ eine gegen x konvergente Folge in M. Wir zeigen $f(a_n) \overset{n\to\infty}{\longrightarrow} f(x)$. Sei hierzu $\varepsilon > 0$. Wegen b) enthält $f^{-1}[U_\varepsilon(f(x))]$ eine offene Umgebung von x in M. Es gibt also $\delta > 0$ mit

$$U_\delta(x) \cap M \subseteq f^{-1}[U_\varepsilon(f(x))], \quad \text{also}$$

$$\forall y \in M\colon (|x-y| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon).$$

Wegen $a_n \overset{n\to\infty}{\longrightarrow} x$ gibt es $m \in \mathbb{N}_0$, so daß $\forall n > m\colon |x - a_n| < \delta$, also auch $|f(x) - f(a_n)| < \varepsilon$ für diese n. Das bedeutet: $f(a_n) \overset{n\to\infty}{\longrightarrow} f(x)$.

c) $\Rightarrow$ d):
Es gelte c), und es sei x ein Berührpunkt von N. Dann können wir zu jedem $n \in \mathbb{N}$ ein $a_n \in N$ mit $|x - a_n| < \frac{1}{n}$ wählen. Die Folge $(a_n)_{n\in\mathbb{N}_0}$ konvergiert gegen x, also konvergiert die Folge $(f(a_n))_{n\in\mathbb{N}_0}$ gegen $f(x)$ wegen c). Nun ist $(f(a_n))_{n\in\mathbb{N}_0}$ eine Folge in $f[N]$. Zu jedem $\varepsilon > 0$ gibt es also ein $n \in \mathbb{N}_0$ mit $f(a_n) \in f[N] \cap U_\varepsilon(f(x))$. Folglich ist $f(x)$ Berührpunkt von $f[N]$.

d) $\Rightarrow$ a):
Es gelte d), und es sei $\varepsilon > 0$. Wir setzen $N = \{y \in M \mid |f(x) - f(y)| \geq \varepsilon\}$. Jedes Element von $f[N]$ hat also mindestens Abstand ε von $f(x)$, d.h. $f(x)$ ist kein Berührpunkt von $f[N]$. Wegen d) ist dann x kein Berührpunkt von N, d.h. es gibt $\delta > 0$ mit $U_\delta(x) \cap N = \emptyset$. Es folgt $|f(x) - f(y)| < \varepsilon$ für alle $y \in U_\delta(x) \cap M$. Wir schließen: f ist stetig in x.

$\square$

Beispiel:
$f\colon \mathbb{R} \to \mathbb{R}$, $f(x) = 1_{\{x>0\}}$ ist unstetig in 0. In der Tat ist 0 ein Berührpunkt von $]0,\infty[$, aber $0 = f(0)$ ist kein Berührpunkt von $f\big[\,]0,\infty[\,\big] = \{1\}$.
Ebenso ist f nicht folgenstetig in 0. Es gilt nämlich $\frac{1}{n} \overset{n\to\infty}{\longrightarrow} 0$, aber $1 = f(\frac{1}{n}) \overset{n\to\infty}{\longrightarrow} 1 \neq 0 = f(0)$.

Satz 3.32 *Es seien $f, g\colon M \to \mathbb{C}$ stetig in $x \in M$. Dann sind auch*

$$f \pm g\colon y \mapsto f(y) \pm g(y),$$
$$f \cdot g\colon y \mapsto f(y) \cdot g(y)$$

und für $g(x) \neq 0$ auch

$$\frac{f}{g}\colon y \mapsto \frac{f(y)}{g(y)}$$

stetig in x.

Beweis: Es sei $(a_n)_{n\in\mathbb{N}_0}$ eine Folge in M mit $a_n \overset{n\to\infty}{\longrightarrow} x$. Dann gilt $f(a_n) \overset{n\to\infty}{\longrightarrow} f(x)$ und $g(a_n) \overset{n\to\infty}{\longrightarrow} g(x)$, da f und g stetig in x sind. Es folgt

$$f(a_n) \pm g(a_n) \overset{n\to\infty}{\longrightarrow} f(x) + g(x),$$
$$f(a_n)g(a_n) \overset{n\to\infty}{\longrightarrow} f(x)g(x)$$

und für $g(x) \neq 0$ auch

$$\frac{f(a_n)}{g(a_n)} \overset{n\to\infty}{\longrightarrow} \frac{f(x)}{g(x)},$$

also folgt die Behauptung.

$\square$

Definition 3.33 Eine Funktion $f\colon M \to \mathbb{R}$ bzw. $\mathbb{C}$ heißt *stetig*, wenn sie in jedem $x \in M$ stetig ist.

Satz 3.34 (Charakterisierung der Stetigkeit) Es sei $M \subseteq \mathbb{R}$ (oder $\mathbb{C}$). Folgende Aussagen für eine Funktion $f\colon M \to \mathbb{R}$ bzw. $\mathbb{C}$ sind äquivalent:

a) f ist stetig.

b) *„Urbilder offener Mengen sind offen“:* Für jedes offene U ist $f^{-1}[U]$ offen in M, d.h. es gibt eine offene Menge W in $\mathbb{R}$ (oder $\mathbb{C}$), so daß $f^{-1}[U] = W \cap M$.

c) *„Urbilder abgeschlossener Menge sind abgeschlossen“:* Für jedes abgeschlossene V ist $f^{-1}[V]$ abgeschlossen in M, d.h. es gibt eine abgeschlossene Menge W in $\mathbb{R}$ (oder $\mathbb{C}$) mit $f^{-1}[V] = W \cap M$.

Beweis:

a) $\Rightarrow$ b):
Es sei U offen und $x \in f^{-1}[U]$. Wir müssen zeigen: x ist innerer Punkt von $f^{-1}[U]$ in M. In der Tat: U ist offene Umgebung von $f(x)$, also enthält $f^{-1}[U]$ eine offene Umgebung von x in M. Der Punkt x ist also ein innerer Punkt von $f^{-1}[U]$ in M.

b) $\Rightarrow$ c):
Es sei V abgeschlossen. Dann ist $\mathbb{R} \setminus V$ (bzw. $\mathbb{C} \setminus V$) offen. Also ist $f^{-1}[\mathbb{R}\setminus V]$ (bzw. $f^{-1}[\mathbb{C}\setminus V]$) offen in M wegen a), also $f^{-1}[V] = M \setminus f^{-1}[\mathbb{R}\setminus V]$ abgeschlossen in M.

c) $\Rightarrow$ b):
Ebenso wie b) $\Rightarrow$ c), mit Vertauschung von „offen“ und „abgeschlossen“

b) $\Rightarrow$ a):
Sei $x \in M$ und U eine offene Umgebung von $f(x)$. Dann ist $f^{-1}[U]$ offen in M wegen b). Ferner gilt $x \in f^{-1}[U]$, also ist $f^{-1}[U]$ eine offene Umgebung von x. Also ist f stetig in x wegen der Charakterisierung b) der Stetigkeit in einem Punkt.

Anwendung der topologischen Charakterisierung der Stetigkeit. Diese Charakterisierung der Stetigkeit ist oft sehr praktisch, um die Offenheit oder die Abgeschlossenheit einer konkret gegebenen Menge A zu zeigen:
Man muss nur A als Urbild einer “einfachen” offenen bzw. abgeschlossenen Menge unter einer stetigen Abbildung schreiben. *wichtig!*

Beispiele:

1. Die Einheitskreislinie $S^1 = \{z \in \mathbb{C}|\ |z| = 1\}$ ist abgeschlossen, denn sie ist das Urbild der abgeschlossenen Menge $\{1\} \subset \mathbb{R}$ unter der stetigen Abbildung $f : \mathbb{C} \to \mathbb{R}$, $f(z) = |z|^2 = z\overline{z}$.
2. Die obere Halbebene $H = \{z \in \mathbb{C}|\ \operatorname{Im} z > 0\}$ ist offen, denn sie ist das Urbild der offenen Menge $\mathbb{R}^+ \subset \mathbb{R}$ unter der Imaginärteilabbildung $\operatorname{Im} : \mathbb{C} \to \mathbb{R}$, die stetig ist.

3.6.2 Ausblick: Die allgemeine Stetigkeitsdefinition

Nachdem wir nun Stetigkeit auf den Begriff der offenen Mengen zurückgespielt haben, können wir den Begriff auf beliebige topologische Räume, insbesondere auf $\mathbb{R} \cup \{\pm\infty\}$ und $\mathbb{C} \cup \{\infty\}$ erweitern:

Definition 3.35 (allgemeine Stetigkeitsdefinition) *Es seien $(X, \mathcal{T})$ bzw. $(Y, \mathcal{S})$ topologische Räume. Wir nennen die Elemente von $\mathcal{T}$ bzw. $\mathcal{S}$ "offen" in X bzw. Y.*
Eine Abbildung $f\colon X \to Y$ heißt stetig, wenn die Urbilder $f^{-1}[U]$ aller offenen Mengen $U \in \mathcal{S}$ offen in X sind.
Sie heißt stetig in $x \in X$, wenn für jede offene Umgebung $U \in \mathcal{S}$ von $f(x)$ das Urbild $f^{-1}[U]$ eine offene Umgebung $V \in \mathcal{T}$ von x enthält.

Beispiel: Die stetige Abbildung $q\colon \mathbb{C} \setminus \{0\} \to \mathbb{C}$, $q(x) = \frac{1}{x}$ hat eine stetige Fortsetzung

$$Q\colon \mathbb{C} \cup \{\infty\} \to \mathbb{C} \cup \{\infty\},\ Q(x) = \begin{cases} \frac{1}{x} & \text{für } x \in \mathbb{C} \setminus \{0\}, \\ \infty & \text{für } x = 0, \\ 0 & \text{für } x = \infty. \end{cases}$$

In der Tat ist Q stetig in 0:
Ist nämlich U eine offene Umgebung von ∞, so enthält U alle genügend betragsgroßen $z \in \mathbb{C}$; also enthält $Q^{-1}[U]$ die Zahlen $\frac{1}{z}$ für alle genügend betragsgroßen z. Zudem gilt $0 \in Q^{-1}[U]$. Also enthält $Q^{-1}[U]$ alle genügend betragskleinen $z \in \mathbb{C}$, also eine offene Umgebung von 0.
Ebenso sieht man, daß Q stetig in ∞ ist.

Bemerkung: Die Charakterisierung der Stetigkeit in einem Punkt von Satz 3.31 gilt auch für Funktionen mit Argumenten und Werten in $\mathbb{R} \cup \{\pm\infty\}$ oder $\mathbb{C} \cup \{\infty\}$. Wir verzichten hier auf Details.

3.6.3 Grundlegende Eigenschaften stetiger Funktionen

Satz 3.36 (Bilder kompakter Mengen sind kompakt) *Es seien $M, N \subseteq \mathbb{R}, \mathbb{C}$ oder beliebige topologische Räume. Weiter sei $f\colon M \to N$ stetig. Dann gilt: Wenn M kompakt ist, so ist auch das Bild $f[M]$ kompakt.*

Beweis: Es sei $(U_i)_{i \in I}$ eine offene Überdeckung von $f[M]$. Weil f stetig ist, ist für alle $i \in I$ das Urbild $f^{-1}[U_i] = \{x \in M \mid f(x) \in U_i\}$ offen. Jedes $x \in M$ ist in einem $f^{-1}[U_i]$ enthalten, so daß $(f^{-1}[U_i])_{i \in I}$ eine offene Überdeckung von M ist. Weil M kompakt ist, hat diese offene Überdeckung von M eine endliche Teilüberdeckung $(f^{-1}[U_i])_{i \in E}$, $E \subseteq$

I endlich. Dann ist $(U_i)_{i\in E}$ eine endliche Teilüberdeckung von $f[M]$. Folglich ist $f[M]$ kompakt.

□

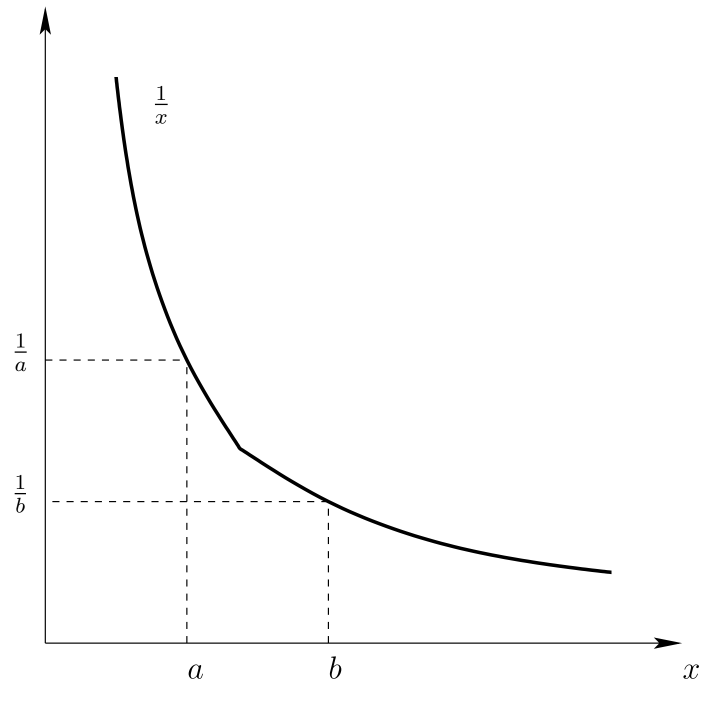
Korollar 3.37 (Satz vom Maximum) *Es sei $f\colon M \to \mathbb{R}$ stetig, $M \neq \emptyset$ kompakt. Dann nimmt f sowohl ein Maximum* $\max f$ *als auch ein Minimum* $\min f$ *als Werte an, d.h. es gilt:*

$$\exists x \in M \quad \forall y \in M\colon \; f(x) \geq f(y) \quad [\textit{bzw. } f(x) \leq f(y)]$$

Beweis: $f[M] \subseteq \mathbb{R}$ ist kompakt und nichtleer, besitzt also nach dem Satz 2.21 sowohl ein Maximum als auch ein Minimum.

□

Beispiel: $f(x) = \frac{1}{x}$, $x > 0$, besitzt weder ein Maximum noch ein Minimum. Schränken wir jedoch diese stetige Funktion auf ein kompaktes Intervall $[a, b]$ ein, wobei $0 < a \leq b$, so nimmt die Einschränkung das Maximum $\frac{1}{a}$ und das Minimum $\frac{1}{b}$ an.



Wir betrachten nun eine Folge von Funktionen $f_n\colon M \to \mathbb{C}$, $n \in \mathbb{N}_0$:

Definition 3.38 f_n heißt *punktweise* konvergent gegen $f\colon M \to \mathbb{C}$, wenn gilt:

$$\forall x \in M\colon f_n(x) \overset{n\to\infty}{\longrightarrow} f(x),$$

anders gesagt:

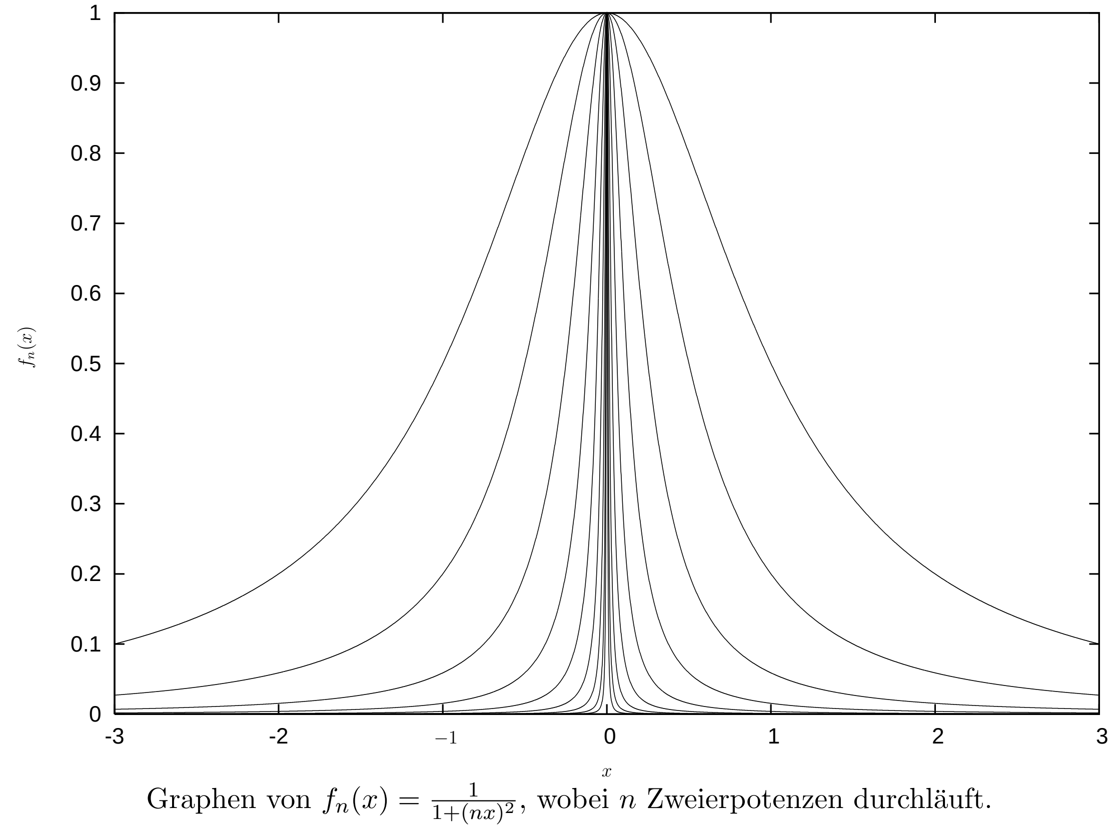
$$\forall x \in M \quad \forall \varepsilon > 0 \quad \exists m \in \mathbb{N}_0 \quad \forall n > m\colon |f_n(x) - f(x)| < \varepsilon.$$

$(f_n)_{n\in\mathbb{N}_0}$ heißt *gleichmäßig* konvergent gegen $f\colon M \to \mathbb{C}$, wenn gilt:

$$\forall \varepsilon > 0 \quad \exists m \in \mathbb{N}_0 \quad \forall n > m \quad \forall x \in M\colon |f_n(x) - f(x)| < \varepsilon.$$

Man beachte die unterschiedliche Quantorenstellung! Bei *punktweiser* Konvergenz darf $m \in \mathbb{N}_0$ von $\varepsilon > 0$ und von $x \in M$ abhängen, bei *gleichmäßiger* Konvergenz jedoch nur von $\varepsilon > 0$, nicht von x.

Beispiel: $f_n\colon \mathbb{R} \to \mathbb{R}$, $f_n(x) = \frac{1}{1+(nx)^2}$ konvergiert punktweise, aber *nicht* gleichmäßig gegen $f(x) = 1_{\{x=0\}}$.



Graphen von $f_n(x) = \frac{1}{1+(nx)^2}$, wobei n Zweierpotenzen durchläuft.

In der Tat gilt zwar

$$\forall x \in \mathbb{R} \setminus \{0\}\colon \frac{1}{1+(nx)^2} \overset{n\to\infty}{\longrightarrow} 0 \qquad \text{und} \qquad \frac{1}{1+(n\cdot 0)^2} \overset{n\to\infty}{\longrightarrow} 1,$$

aber zu $\varepsilon = \frac{1}{2}$ läßt sich kein $m \in \mathbb{N}_0$ finden, so daß für alle $n > m$ und für alle $x \in \mathbb{R}$ gilt:

$$\left| \frac{1}{1+(nx)^2} - 1_{\{x=0\}} \right| < \frac{1}{2},$$

denn z.B. für $x_n = \frac{1}{n}$ gilt:

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

also

$$|f_n(x_n) - f(x_n)| = \frac{1}{2},$$

gleichgültig, wie groß n ist.

Man beachte auch, daß die Grenzfunktion $x \mapsto 1_{\{x=0\}}$ unstetig ist, obwohl alle f_n stetig sind.

Bei gleichmäßiger Konvergenz gibt es dieses Phänomen nicht:

Satz 3.39 *Die Folge stetiger Funktionen $f_n\colon M \to \mathbb{C}$ konvergiere gleichmäßig gegen $f\colon M \to \mathbb{C}$. Dann ist f stetig.*

Beweis: Es sei $\varepsilon > 0$ und $x \in M$. Weil $(f_n)_{n\in\mathbb{N}_0}$ gleichmäßig gegen f konvergiert, können wir ein $n \in \mathbb{N}_0$ mit $\forall z \in M\colon |f_n(z) - f(z)| < \frac{\varepsilon}{3}$ wählen. Weil f_n in x stetig ist, gibt es ein $\delta > 0$, so daß gilt:

$$\forall y \in U_\delta(x) \cap M\colon |f_n(y) - f_n(x)| < \frac{\varepsilon}{3}.$$

Es folgt für alle $y \in U_\delta(x) \cap M$:

$$|f(y) - f(x)| \leq |f(y) - f_n(y)| + |f_n(y) - f_n(x)| + |f_n(x) - f(x)| < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.$$

Wir haben also gezeigt:

$$\forall x \in M \quad \forall \varepsilon > 0 \quad \exists \delta > 0 \quad \forall y \in U_\delta(x) \cap M\colon\ f(y) \in U_\varepsilon(f(x)),$$

d.h. f ist stetig.

□

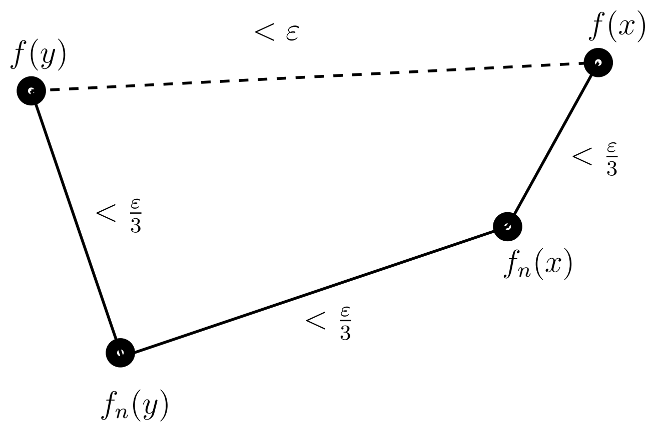


Illustration zur Abschätzung im letzten Beweis

Als Anwendung beweisen wir, daß Potenzreihen in ihrem (offenen) Konvergenzkreis stetige Funktionen beschreiben.

Satz 3.40 *Es sei $f(x) = \sum\limits_{k=0}^{\infty} a_k x^k$ eine Potenzreihe mit Konvergenzradius R. Dann ist f in $U_R(0) = \{x \in \mathbb{C} \mid |x| < R\}$ stetig.*

Im allgemeinen konvergiert $\sum\limits_{k=0}^{\infty} a_k x^k$ *nicht* gleichmäßig in $U_R(0)$, so daß wir den vorhergehenden Satz nicht direkt anwenden können.
Gegenbeispiel: Die geometrische Reihe $\sum\limits_{k=0}^{\infty} x^k$ konvergiert in $U_1(0)$ *nicht* gleichmäßig. Es gilt nämlich

$$\left| \sum_{k=0}^{n} x^k \right| \leq n+1$$

für alle $n \in \mathbb{N}_0$ und alle $|x| < 1$, aber es gibt nahe bei 1 Zahlen $x \in U_1(0)$ mit

$$\left| \sum_{k=0}^{\infty} x^k \right| = \left| \frac{1}{1-x} \right| \geq n+2.$$

Wir zeigen folgendes Lemma, das den Satz 3.40 impliziert.

Lemma 3.41 *Für alle $r < R$ ist die Potenzreihe $\sum\limits_{k=0}^{\infty} a_k x^k$ in $U_r(0)$ gleichmäßig konvergent und daher stetig.*

Beweis Für alle $x \in U_r(0)$ gilt:

$$\left| \sum_{k=0}^{n} a_k x^k - \sum_{k=0}^{\infty} a_k x^k \right| = \left| \sum_{k=n+1}^{\infty} a_k x^k \right| \leq \sum_{k=n+1}^{\infty} \left| a_k x^k \right| \leq \underbrace{\sum_{k=n+1}^{\infty} |a_k| r^k}_{\text{unabhängig von } x!} \stackrel{n\to\infty}{\longrightarrow} 0.$$

Es sei nun $\varepsilon > 0$. Wählen wir $m \in \mathbb{N}_0$ so groß, daß für alle $n > m$ gilt:

$$\sum_{k=n+1}^{\infty} |a_k| r^k < \varepsilon,$$

so folgt für diese n und alle $x \in U_r(0)$:

$$\left| \sum_{k=0}^{n} a_k x^k - \sum_{k=0}^{\infty} a_k x^k \right| < \varepsilon.$$

Die Potenzreihe $\sum\limits_{k=0}^{\infty} a_k x^k$ ist also gleichmäßig konvergent in $U_r(0)$. Als gleichmäßiger Limes stetiger Funktionen ist sie dort stetig.

□

Beispiel: Die Exponentialfunktion $\exp\colon \mathbb{C} \to \mathbb{C}$ ist stetig, denn die Exponentialreihe $\sum\limits_{k=0}^{\infty} \frac{x^k}{k!}$ ist in jedem beschränkten Kreis $U_r(0)$, $r < +\infty$, gleichmäßig konvergent.

Der folgende Satz zeigt, daß die Komposition stetiger Funktionen stetig ist.

Satz 3.42 *Es seien $M, N, L \subseteq \mathbb{C}$ (oder auch M, N, L mit einer Topologie versehen). Sind $f\colon M \to N$ und $g\colon N \to L$ stetig, so ist auch $g \circ f : M \to L$ stetig, wobei $g \circ f(x) = g(f(x))$.*

Beweis: Es sei $U \subseteq L$ offen. Dann ist $g^{-1}[U] \subseteq N$ offen, da g stetig ist, also $f^{-1}[g^{-1}[U]] = (g \circ f)^{-1}[U] \subseteq M$ offen, da f stetig ist. Also ist $g \circ f$ stetig.

□

Hier ist eine "lokale Version" des Satzes:

Satz 3.43 *Es seien $M, N, L \subseteq \mathbb{C}$ (oder auch M, N, L mit einer Topologie versehen) und $x \in M$. Sind $f\colon M \to N$ stetig in x und $g\colon N \to L$ stetig in $f(x)$, so ist auch $g \circ f : M \to L$ stetig in x.*

Wir verzichten hier auf die Darstellung des Beweises. Es wird als Übung empfohlen, den Satz auf vier verschiedene Weisen zu zeigen, und zwar mit den vier verschiedenen Charakterisierungen der Stetigkeit in einem Punkt aus Satz 3.31.

Satz 3.44 (Zwischenwertsatz) *Es sei $f\colon [a,b] \to \mathbb{R}$ stetig, $a \le b$. Dann nimmt f alle Zahlen zwischen $f(a)$ und $f(b)$ als Werte an.*

Beweis: Wir unterscheiden 3 Fälle:

1. $f(a) = f(b)$,
2. $f(a) < f(b)$,
3. $f(a) > f(b)$.

Der Fall 1. ist trivial, und der Fall 3. folgt aus dem Fall 2., indem wir $-f$ statt f betrachten. Wir beschränken uns daher auf den 2. Fall und müssen

$$\forall y \in [f(a), f(b)]\ \exists x \in [a,b]\colon\ y = f(x)$$

zeigen. Im Fall $y = f(b)$ ist nichts zu zeigen; wir nehmen also $f(a) \le y < f(b)$ an.
Wir setzen

$$K = \{z \in [a,b] \mid f(z) \le y\} = f^{-1}\Big[\,]-\infty, y]\Big].$$

Als Urbild einer abgeschlossenen Menge ist K abgeschlossen, da f stetig ist, und wegen $K \subseteq [a,b]$ ist K beschränkt. Ferner ist $K \neq \emptyset$ wegen $a \in K$. Also existiert $x = \max K \in K$, denn K ist kompakt. Insbesondere ist $f(x) \leq y$.

Um auch $f(x) \geq y$ zu zeigen, gehen wir so vor: Wegen $b \notin K$ ist $x < b$. Es folgt $]x,b] \neq \emptyset$, und x ist ein Berührpunkt von $]x,b]$. Wegen der Stetigkeit von f folgt hieraus, dass $f(x)$ ein Berührpunkt von $f(]x,b]) \subseteq]y,+\infty[$ ist. (Zum Nachweis der letzten Inklusion beachte man, daß für alle z mit der Eigenschaft $x < z \leq b$ folgt: $z \notin K$, also $f(z) > y$.) Es gilt also $f(x) \in \overline{]y,+\infty[} = [y,+\infty[$, d.h. $f(x) \geq y$.

Zusammen folgt $f(x) = y$.

□

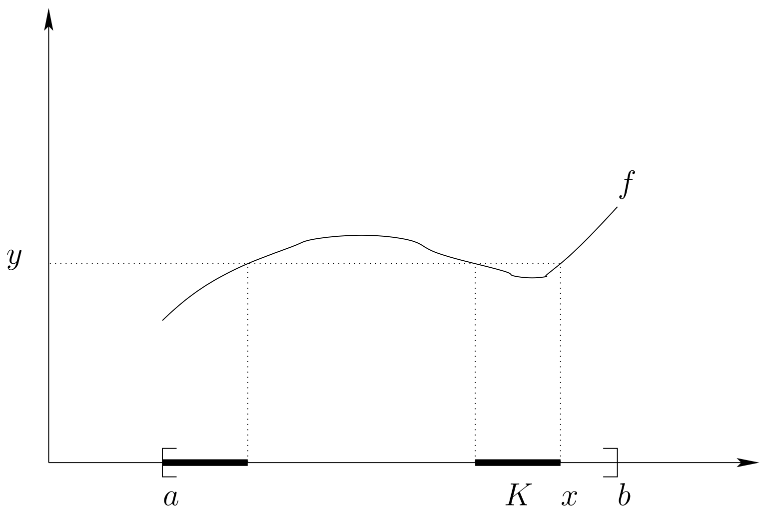


Illustration zum Beweis des Zwischenwertsatzes

Anwendung: Die reelle Exponentialfunktion $\exp\colon \mathbb{R} \to \mathbb{R}$ nimmt alle positiven Zahlen als Werte an.

Beweis: Sei $y > 0$. Einerseits gilt

$$\exp(y) = \sum_{n=0}^{\infty} \frac{y^n}{n!} \geq \frac{y^1}{1!} = y,$$

andererseits

$$\exp\left(-\frac{1}{y}\right) = \frac{1}{\exp\left(\frac{1}{y}\right)} \leq \frac{1}{\frac{1}{y}} = y.$$

Nach dem Zwischenwertsatz gibt es also ein $x \in [-\frac{1}{y}, y]$ mit $e^x = y$, weil $\exp\colon \mathbb{R} \to \mathbb{R}$ stetig ist.

□

Lemma 3.45 *1. Für alle $x \in \mathbb{R}$ ist $\exp(x) > 0$. Für $x > 0$ ist sogar $\exp(x) > 1$.*

2. $\exp : \mathbb{R} \to \mathbb{R}$ *ist streng monoton wachsend.*

Beweis:

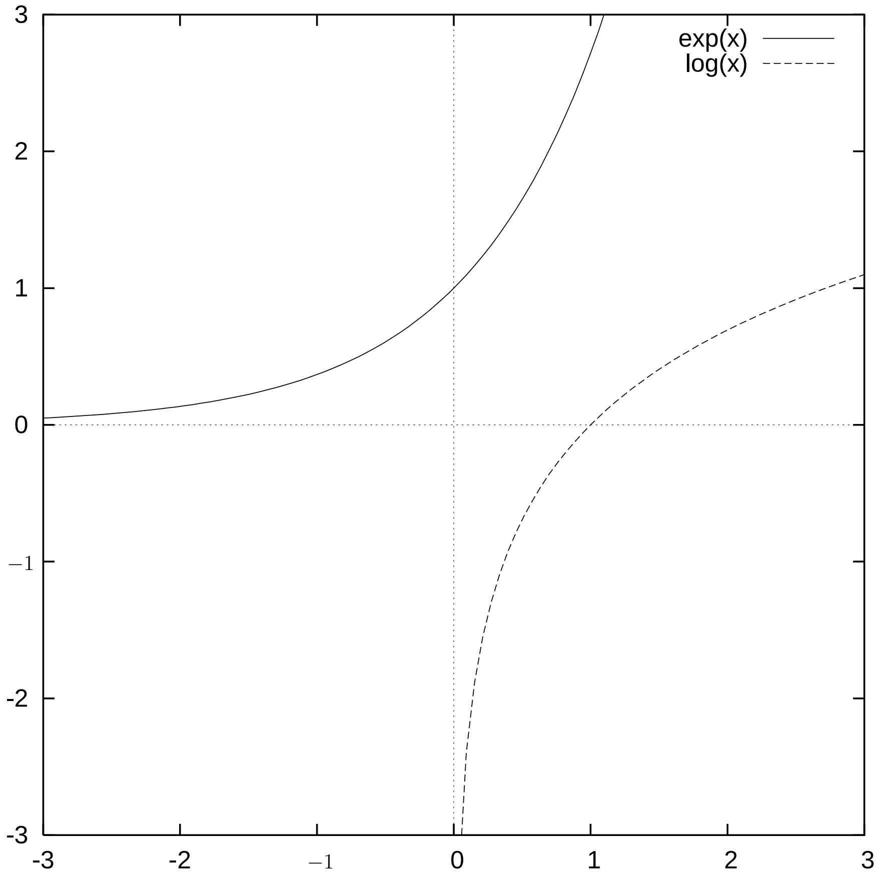
1. Für $x > 0$ ist
$$\exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!} > 1 > 0,$$
denn der 0-te Summand in der Reihe ist 1, und alle anderen Summanden sind positiv. Für $x = 0$ ist $\exp(0) = 1$. Für $x < 0$ folgt $-x > 0$, also $\exp(-x) > 0$ nach dem eben Gezeigten, also $\exp(x) = 1/\exp(-x) > 0$.

2. Für $y, z \in \mathbb{R}$ mit $y > z$ folgt $y - z > 0$, also $\exp(y)/\exp(z) = \exp(y - z) > 1$ wegen Teil 1. Mit $\exp(z) > 0$ folgt hieraus $\exp(y) > \exp(z)$.

Fassen wir zusammen:

Satz 3.46 $\exp\colon \mathbb{R} \to]0, \infty[$ ist eine streng monoton wachsende, stetige Bijektion. Die Umkehrabbildung $\log\colon]0, \infty[\to \mathbb{R}$ wird *(natürlicher) Logarithmus* genannt.

Die Symbole "$\log x$" und "$\ln x$" sind Synonyme.



Graphen der Exponentialfunktion und des Logarithmus

Ausblick: Die komplexe Exponentialfunktion nimmt alle komplexen Zahlen außer 0 als Werte an. Sie ist aber nicht injektiv, wie wir später sehen werden. Deshalb ist der Logarithmus im Komplexen mehrdeutig; wir stellen seine Untersuchung zurück.

Die Logarithmusfunktion $\log\colon]0,\infty[\to \mathbb{R}$ ist stetig.
Es gilt nämlich allgemeiner:

Satz 3.47 *Sei $I =]a,b[\subseteq \mathbb{R}$ ein offenes Intervall und $f\colon I \to \mathbb{R}$ streng monoton steigend oder auch streng monoton fallend. Dann ist $f^{-1}\colon f[I] \to I$ stetig.*

Beweis: Wir behandeln nur den Fall, dass f streng monoton steigt; der andere Fall wird analog behandelt. Zu zeigen ist:[32]

$$\forall y = f(x) \in f[I]\ \forall \varepsilon > 0\ \exists \delta > 0 :\ f^{-1}[U_\delta(y) \cap f[I]] \subseteq U_\varepsilon(x)$$

[32] Man beachte, dass die beiden Lesarten von $f^{-1}[\ldots]$, nämlich 1) Urbild unter f oder 2) Bild unter der Umkehrabbildung f^{-1}, genau die gleiche Menge bezeichnen, so dass die Schreibweise nicht missverständlich ist.

Hierzu seien $y = f(x) \in f[I]$ mit $x \in I$ und $\varepsilon > 0$ gegeben. Wir dürfen zusätzlich annehmen, dass $\varepsilon > 0$ so klein ist, daß $x \pm \varepsilon \in I$. (Durch Verkleinern von ε kann man das stets erreichen, weil I offen ist.) Dann ist $U =]f(x-\varepsilon), f(x+\varepsilon)[$ wegen der strengen Monotonie von f eine offene Umgebung von $f(x)$. Wir finden also ein $\delta > 0$ mit $U_\delta(y) \subseteq U$. Wir müssen nun $f^{-1}[U_\delta(y) \cap f[I]] \subseteq U_\varepsilon(x)$ zeigen. Hierzu sei $z = f(w) \in U_\delta(y) \cap f[I]$ gegeben; zu zeigen ist nun $w \in U_\varepsilon(x)$. Nun gilt

$$f(x-\varepsilon) < f(w) < f(x+\varepsilon),$$

also wegen der strengen Monotonie von f:

$$x - \varepsilon < w < x + \varepsilon,$$

anders gesagt $w \in U_\varepsilon(x)$, wie behauptet.

□

Bemerkung: Man beachte, daß wir nicht voraussetzen brauchen, daß f stetig ist. Zum Beispiel ist

$$f(x) = \begin{cases} x+1, & x \geq 0 \\ x, & x < 0 \end{cases}$$

unstetig in 0, aber $f^{-1}\colon]-\infty, 0[\cup [1, \infty[\to \mathbb{R}$ ist stetig, selbst in 1.

3.6.4 Varianten des Stetigkeitsbegriffs

Definition 3.48 Eine Funktion $f\colon M \to \mathbb{C}$, $M \subseteq \mathbb{C}$ heißt *gleichmäßig stetig*, wenn gilt

$$\forall \varepsilon > 0 \quad \exists \delta > 0 \quad \forall x \in M \quad \forall y \in M\colon (|x-y| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon). \tag{47}$$

Man beachte die andere Quantorenreihenfolge als bei der Stetigkeitsdefinition:

$$f: M \to \mathbb{C} \text{ stetig}$$
$$\Longleftrightarrow$$
$$\forall x \in M \quad \forall \varepsilon > 0 \quad \exists \delta > 0 \quad \forall y \in M\colon (|x-y| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon).$$

Bei gleichmäßiger Stetigkeit darf also δ – im Gegensatz zur Stetigkeit – nicht von $x \in M$ abhängen.

Jede gleichmäßig stetige Funktion ist stetig, aber nicht umgekehrt:

Beispiel: $f\colon \mathbb{R} \to \mathbb{R}, f(x) = x^2$ ist stetig, aber *nicht* gleichmäßig stetig.
Beweis: $x \mapsto x$ ist stetig, also ist auch $f: x \mapsto x^2$ als Produkt stetiger Funktionen stetig. Es gilt aber das Gegenteil von (47):

$$\exists \varepsilon > 0 \quad \forall \delta > 0 \quad \exists x \in M \quad \exists y \in M\colon (|x-y| < \delta \wedge |f(x) - f(y)| \geq \varepsilon).$$

In der Tat: Wir wählen $\varepsilon = 1$. Es sei $\delta > 0$. Wir wählen $x = \frac{1}{\delta}$ und $y = x + \frac{\delta}{2}$. Dann gilt $|x - y| = \frac{\delta}{2}$ und

$$|f(x) - f(y)| = \left(x + \frac{\delta}{2}\right)^2 - x^2 = 2x\frac{\delta}{2} + \frac{\delta^2}{4} \geq x\delta = 1 = \varepsilon.$$

□

Unter einer Zusatzvoraussetzung fallen gleichmäßige Stetigkeit und Stetigkeit zusammen:

Satz 3.49 *Es sei $f\colon M \to \mathbb{C}$ stetig und $M \subseteq \mathbb{C}$ kompakt. Dann ist f gleichmäßig stetig. Insbesondere gilt dies für abgeschlossene und beschränkte Intervalle $M = [a, b]$.*

Beweis: Es sei $\varepsilon > 0$. Weil f stetig ist, können wir zu jedem $z \in M$ ein $\Delta(z, \varepsilon) > 0$ wählen, so daß

$$f\left[U_{\Delta(z,\varepsilon)}(z) \cap M\right] \subseteq U_{\varepsilon/2}(f(z)).$$

Nun ist $(U_{\Delta(z,\varepsilon)/2}(z))_{z\in M}$ eine offene Überdeckung von M. Weil M kompakt ist, hat sie eine endliche Teilüberdeckung $(U_{\Delta(z,\varepsilon)/2}(z))_{z\in E}$. Wir setzen

$$\delta = \frac{1}{2} \min_{z\in E} \Delta(z, \varepsilon) > 0.$$

Nun seien $x, y \in M$ mit $|x - y| < \delta$. Dann gibt es $z \in E$ mit $x \in U_{\Delta(z,\varepsilon)/2}(z)$. Es folgt einerseits $x \in U_{\Delta(z,\varepsilon)}(z)$, also

$$|f(x) - f(z)| < \frac{\varepsilon}{2},$$

und andererseits

$$|y - z| \leq |y - x| + |x - z| < \delta + \frac{\Delta(z, \varepsilon)}{2} \leq \frac{\Delta(z, \varepsilon)}{2} + \frac{\Delta(z, \varepsilon)}{2} = \Delta(z, \varepsilon)$$

also $|f(y) - f(z)| < \frac{\varepsilon}{2}$ wegen der Wahl von $\Delta(z, \varepsilon)$.
Insgesamt:

$$|f(x) - f(y)| \leq |f(x) - f(z)| + |f(y) - f(z)| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

□

Beispiel: $f\colon [-1, 1] \to \mathbb{R}$, $x \mapsto x^2$ ist gleichmäßig stetig.

Hier ist noch eine einfachere Variante des Stetigkeitsbegriffs, bei der δ *linear* von ε abhängen soll:

Definition 3.50 Eine Funktion $f\colon M \to \mathbb{C}$ $(M \subseteq \mathbb{C})$ heißt *lokal Lipschitz-stetig* in $x \in M$, wenn gilt:

$$\exists L > 0 \quad \exists \varepsilon > 0 \quad \forall y \in U_\varepsilon(x) \cap M\colon |f(y) - f(x)| \leq L|x - y|.$$

Sie heißt *gleichmäßig Lipschitz-stetig*, wenn gilt:

$$\exists L > 0 \quad \forall x, y \in M\colon |f(y) - f(x)| \leq L|x - y|.$$

Das Beispiel 2 im Logik-Abschnitt zeigt, daß die Wurzelfunktion in $]0, \infty[$ überall lokal Lipschitz-stetig, aber nicht gleichmäßig Lipschitz-stetig ist.

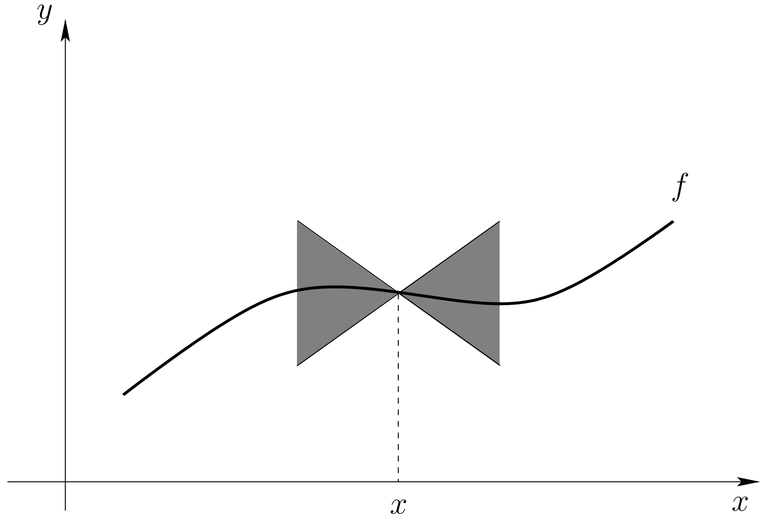


Illustration zur Lipschitzstetigkeit

3.6.5 Konvergenz für $x \to x_0$

Es seien $M, N \subseteq \mathbb{R} \cup \{\pm\infty\}$ oder $\mathbb{C} \cup \{\infty\}$ und $x_0 \in M$. x_0 sei kein *isolierter Punkt*, d.h. x_0 sei ein Berührpunkt von $M \setminus \{x_0\}$. Es sei $f : M \setminus \{x_0\} \to N$ eine Funktion und $y \in N$.

Definition 3.51 *Wir sagen, $f(x)$ konvergiert für $x \to x_0$ gegen den Grenzwert y, wenn die Abbildung $g : M \to N$,*

$$g(x) = \begin{cases} f(x) & \text{für } x \in M \setminus \{x_0\} \\ y & \text{für } x = x_0 \end{cases}$$

stetig in x_0 ist. In Zeichen:

$$f(x) \stackrel{x \to x_0}{\longrightarrow} y$$

oder auch

$$\lim_{x \to x_0} f(x) = y.$$

Der Grenzwert y ist eindeutig bestimmt, falls er existiert. Im Fall $M = \mathbb{N}_0 \cup \{+\infty\}$, $x_0 = +\infty$ stimmt diese neue Definition von $f(n) \stackrel{n \to +\infty}{\longrightarrow} y$ mit der früheren überein.

Für $M, N \subseteq \mathbb{C}$ und $y \in \mathbb{C}$ kann man Konvergenz auch so formulieren:

$\lim_{x \to x_0} f(x) = y \in \mathbb{C}$ ist äquivalent zu

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall x \in M \setminus \{x_0\} : \ (|x - x_0| < \delta \ \Rightarrow\ |f(x) - y| < \varepsilon).$$

Beispiele:

1.
$$\frac{\sqrt{x}-1}{x-1} \xrightarrow{x\to 1} \frac{1}{2}.$$

Begründung: Die Funktion

$$x \mapsto \frac{\sqrt{x}-1}{x-1} = \frac{1}{\sqrt{x}+1}, \quad x \geq 0,\ x \neq 1$$

ist stetig nach $x = 1$ fortsetzbar mit dem Wert $1/2$. Dies sieht man so: $x \mapsto \sqrt{x}$, $x \geq 0$, ist die Umkehrfunktion der streng monoton steigenden Funktion $y \mapsto y^2$, $y \geq 0$, also stetig. Folglich ist auch $x \mapsto \sqrt{x}+1$, $x \geq 0$, und damit auch $x \mapsto 1/(\sqrt{x}+1)$, $x \geq 0$, stetig.

2. Für $a > 0$, $x \in \mathbb{C}$ definieren wir[33]

$$\boxed{a^x := e^{x \log a}}$$

Diese Definition ist für $x \in \mathbb{Q}$ konsistent mit der aus der Schule bekannten.
Es gilt:

$$\frac{a^x - 1}{x} \xrightarrow{x\to 0} \log a.$$

Beweis: Es gilt für $x \in \mathbb{C} \setminus 0$:

$$\begin{aligned}
\frac{a^x - 1}{x} &= \frac{1}{x}\left(\sum_{n=0}^{\infty} \frac{(\log a)^n}{n!} x^n - 1\right) \\
&= \frac{1}{x}\sum_{n=1}^{\infty} \frac{(\log a)^n}{n!} x^n \\
&= \sum_{n=0}^{\infty} \frac{(\log a)^{n+1}}{(n+1)!} x^n.
\end{aligned}$$

Die rechte Seite ist eine Potenzreihe in x mit Konvergenzradius $+\infty$, also stetig fortsetzbar in $x = 0$ mit dem Wert $\log a$.

□

[33] Der Leser mag sich fragen, warum wir als Basis a der Potenz a^x hier nur positive Zahlen zulassen, während der Exponent x eine beliebige komplexe Zahl sein darf. Der Grund ist, dass die Exponentialfunktion im Komplexen nicht injektiv ist, so dass der Logarithmus im Komplexen mehrdeutig wird. Zum Beispiel ist $e^{i\pi(2k+1)} = -1$ für alle $k \in \mathbb{Z}$, so dass man mit gleichem Recht alle komplexen Zahlen der Gestalt $i\pi(2k+1)$, $k \in \mathbb{Z}$, als Logarithmen von -1 auffassen kann. Diese Mehrdeutigkeit vererbt sich dann auf Potenzen mit nichtganzzahligen Exponenten: $(-1)^{1/2} = \sqrt{-1}$ kann man mit gleichem Recht als $i = e^{\frac{1}{2}i\pi(2k+1)}$ für gerade $k \in \mathbb{Z}$ oder als $-i = e^{\frac{1}{2}i\pi(2k+1)}$ für ungerade $k \in \mathbb{Z}$ auffassen. Um diese Mehrdeutigkeit zu vermeiden, beschränken wir die Definition auf positive a und verwenden nur den *reellen* Zweig des Logarithmus.

Für $f : \mathbb{R} \to \mathbb{R}$ formulieren wir die Aussage

$$\lim_{x \to +\infty} f(x) = +\infty$$

noch einmal anders mit einer Formel:

$$\forall r \in \mathbb{R} \, \exists s \in \mathbb{R} \, \forall x > s : \; f(x) > r.$$

Beispiel: Für alle $n \in \mathbb{N}_0$ gilt

$$\lim_{x \to +\infty} \frac{e^x}{x^n} = +\infty.$$

Beweis: Gegeben $r \in \mathbb{R}$, wählen wir $s = (n+1)! \max\{r, 0\} \geq 0$. Dann gilt für alle $x > s$:

$$\frac{e^x}{x^n} = \frac{1}{x^n} \sum_{k=0}^{\infty} \frac{x^k}{k!} \geq \frac{1}{x^n} \frac{x^{n+1}}{(n+1)!} = \frac{x}{(n+1)!} > \frac{s}{(n+1)!} \geq r.$$

$\square$

3.6.6 Der Abelsche Grenzwertsatz

Am Rand der Konvergenzkreisscheibe einer Potenzreihe ist beides möglich: Konvergenz oder Divergenz. Ein Beispiel liefert die "umgedrehte Logarithmusreihe"

$$\sum_{n=1}^{\infty} \frac{(-z)^n}{n},$$

die den Konvergenzradius 1 besitzt. Für $z = -1$ ist sie die harmonische Reihe, die in $\mathbb{C}$ divergiert, aber für $z = 1$ ist sie die alternierende harmonische Reihe, die konvergiert. Im Fall der Konvergenz stellt sich die Frage nach Stetigkeit im betroffenen Randpunkt. Der Abelsche Grenzwertsatz, den wir in diesem Abschnitt behandeln, gibt eine Teilantwort auf diese Frage. Bevor wir ihn formulieren, starten wir mit einer geometrischen Vorüberlegung: Wir können uns auf die Einheitskreisscheibe und den Punkt $z = 1$ beschränken, denn mit einer Drehstreckung kann man auch den allgemeinen Fall darauf zurückführen. Für $M \geq 0$ und $c > 0$ definieren wir die Menge

$$\Delta_{M,c} := \{z \in \mathbb{C} |\; 1 - c \leq \operatorname{Re} z \leq 1, \; |\operatorname{Im} z| \leq M(1 - \operatorname{Re} z)\}.$$

Für $M > 0$ ist das die abgeschlossene Dreiecksscheibe in der komplexen Ebene mit den Eckpunkten 1, $1 - c + iMc$ und $1 - c - iMc$; für $M = 0$ ist es nur das Intervall $[1 - c, 1]$. Als Vorüberlegung beweisen wir:

Lemma 3.52 *Für alle $M \geq 0$ gibt es ein $c > 0$ und ein $L > 0$, so dass für alle $z \in \Delta_{M,c}$ gilt:*

$$|z - 1| \leq L(1 - |z|). \tag{48}$$

Insbesondere folgt für solche $c > 0$:

$$\Delta_{M,c} \subseteq \{z \in \mathbb{C} |\; |z| < 1 \; \vee \; z = 1\}. \tag{49}$$

Beweis: Gegeben $M \geq 0$, setzen wir

$$c := \frac{1}{1+M^2} > 0 \tag{50}$$

und

$$L := 2(M+1) > 0. \tag{51}$$

Nun sei $z \in \Delta_{M,c}$ gegeben. Wir kürzen ab: $x := 1 - \operatorname{Re} z$ und $y := \operatorname{Im} z$. Insbesondere gilt $z = 1 + iy - x$ sowie $0 \leq x \leq c$ und $|y| \leq Mx$. Dann folgt:

$$|z-1| = |iy - x| \leq |y| + |x| \leq Mx + x = (1+M)x \tag{52}$$

Wir schätzen andererseits ab:

$$\begin{aligned}
|z| &= \sqrt{(1-x)^2 + y^2} \leq \sqrt{(1-x)^2 + (Mx)^2} = \sqrt{1 - 2x + (1+M^2)x^2} \\
&\leq \sqrt{1 - 2x + (1+M^2)cx} \quad \text{(wegen } 0 \leq x \leq c\text{)} \\
&= \sqrt{1 - 2x + x} \quad \text{(wegen (50))} \\
&= \sqrt{1-x} \leq \sqrt{1 - x + \frac{x^2}{4}} \\
&= \sqrt{\left(1 - \frac{x}{2}\right)^2} = 1 - \frac{x}{2} \quad \text{(wegen } x < 1\text{)}
\end{aligned}$$

also

$$L(1 - |z|) \geq L\frac{x}{2} \geq L\frac{|z-1|}{2(M+1)} = |z-1| \quad \text{(wegen (52) und (51))}$$

Damit ist die Behauptung (48) gezeigt. Zum Beweis der verbleibenden Behauptung (49) sei $z \in \Delta_{M,c}$ mit $z \neq 1$ gegeben. Dann folgt

$$1 - |z| \geq \frac{|z-1|}{L} > 0$$

und daher $|z| < 1$, was zu zeigen war. □

Satz 3.53 (Abelscher Grenzwertsatz) *Gegeben sei eine konvergente Reihe*

$$A = \sum_{n=0}^{\infty} a_n \in \mathbb{C}$$

zu einer Folge $(a_n)_{n \in \mathbb{N}}$ mit Werten in $\mathbb{C}$. Weiter sei $M \geq 0$ gegeben. Dann existiert $c > 0$ mit

$$\Delta_{M,c} \subseteq \{z \in \mathbb{C} \mid |z| < 1 \ \vee \ z = 1\}, \tag{53}$$

so dass die auf $\Delta_{M,c}$ eingeschränkte Potenzreihe

$$f : \Delta_{M,c} \to \mathbb{C}, \quad f(z) := \sum_{n=0}^{\infty} a_n z^n$$

stetig ist.

Beweis: Wir wählen $c > 0$ und $L > 0$ nach dem vorhergehenden Lemma; insbesondere gilt dann die Behauptung (53). Die Potenzreihe $\sum_{n=0}^{\infty} a_n z^n$ besitzt mindestens den Konvergenzradius 1, weil sie für $z = 1$ nach Voraussetzung konvergiert. Insbesondere ist sie in allen Punkten $z \in \mathbb{C}$ mit $|z| < 1$ stetig. Wegen Formel (53) ist f in allen Punkten $z \in \Delta_{M,c} \setminus \{1\}$ stetig, da Potenzreihen im *Inneren* ihrer Konvergenzkreisscheibe stetig sind. Es bleibt nur noch der interessanteste Fall $z = 1$ zu behandeln. Zu zeigen ist also:

$$\forall \epsilon > 0\, \exists \delta > 0\, \forall z \in \Delta_{M,c} : \left(|z-1| < \delta \;\Rightarrow\; \left| \sum_{n=0}^{\infty} a_n z^n - A \right| < \epsilon \right) \tag{54}$$

Hierzu sei $\epsilon > 0$ gegeben. Weil die Reihe $\sum_{n=0}^{\infty} a_n$ in $\mathbb{C}$ konvergiert, bildet die Folge der Partialsummen

$$b_m := \sum_{n=0}^{m} a_n, \quad m \in \mathbb{N}_0, \tag{55}$$

eine Cauchyfolge. Wir finden also ein $n_0 \in \mathbb{N}_0$, so dass für alle $l, m \in \mathbb{N}_0$ mit $n_0 \le l \le m$ gilt:

$$|b_m - b_l| < \frac{\epsilon}{2L} \tag{56}$$

Wegen der Stetigkeit von Polynomfunktionen gilt

$$\sum_{n=0}^{n_0} a_n z^n \overset{z\to 1}{\longrightarrow} \sum_{n=0}^{n_0} a_n.$$

Wir nehmen also ein $\delta > 0$, so dass für alle $z \in \mathbb{C}$ mit $|z-1| < \delta$ gilt:

$$\left| \sum_{n=0}^{n_0} a_n z^n - \sum_{n=0}^{n_0} a_n \right| < \frac{\epsilon}{2}. \tag{57}$$

Mit dieser Wahl von δ ist vom ursprünglichen Beweisziel (54) nur mehr das folgende Beweisziel übrig geblieben:

$$\forall z \in \Delta_{M,c} : \left(|z-1| < \delta \;\Rightarrow\; \left| \sum_{n=0}^{\infty} a_n z^n - A \right| < \epsilon \right) \tag{58}$$

Zum Beweis hiervon sei $z \in \Delta_{M,c}$ mit $|z-1| < \delta$ gegeben. Wir schätzen ab:

$$\begin{aligned}
&\left|\sum_{n=0}^{\infty} a_n z^n - A\right| = \left|\sum_{n=0}^{n_0} a_n z^n + \sum_{n=n_0+1}^{\infty} a_n z^n - \sum_{n=0}^{n_0} a_n - \sum_{n=n_0+1}^{\infty} a_n\right| \\
\leq &\left|\sum_{n=0}^{n_0} a_n z^n - \sum_{n=0}^{n_0} a_n\right| + \left|\sum_{n=n_0+1}^{\infty} a_n z^n - \sum_{n=n_0+1}^{\infty} a_n\right| \\
< &\frac{\epsilon}{2} + \left|\sum_{n=n_0+1}^{\infty} a_n z^n - \sum_{n=n_0+1}^{\infty} a_n\right| = \frac{\epsilon}{2} + \left|\sum_{n=n_0+1}^{\infty} a_n (z^n - 1)\right| \\
= &\frac{\epsilon}{2} + \lim_{m\to\infty} \left|\sum_{n=n_0+1}^{m} a_n (z^n - 1)\right|
\end{aligned}$$

wobei wir die Stetigkeit des Absolutbetrags verwendet haben. Es genügt nun, noch zu zeigen:

$$\forall m > n_0 : \left|\sum_{n=n_0+1}^{m} a_n (z^n - 1)\right| \leq \frac{\epsilon}{2}, \tag{59}$$

denn damit folgt die zu zeigende Behauptung so:

$$\left|\sum_{n=0}^{\infty} a_n z^n - A\right| < \frac{\epsilon}{2} + \lim_{m\to\infty} \left|\sum_{n=n_0+1}^{m} a_n (z^n - 1)\right| \leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

Zum Beweis der Behauptung (59) sei $m \in \mathbb{N}_0$ mit $m > n_0$ gegeben. Die Hauptidee des Beweises besteht nun darin, die Differenz $z^n - 1$ in der folgenden Rechnung mit der geometrischen Summe auszudrücken und dann die Summationsreihenfolge zu vertauschen:

$$\begin{aligned}
&\sum_{n=n_0+1}^{m} a_n (z^n - 1) = \sum_{n=n_0+1}^{m} a_n (z-1) \sum_{k=0}^{n-1} z^k \\
= &(z-1) \sum_{\substack{(n,k)\in\mathbb{N}_0^2: \\ n_0 < n \leq m, \\ k < n}} a_n z^k = (z-1) \sum_{k=0}^{m-1} z^k \sum_{n=\max\{n_0,k\}+1}^{m} a_n \\
= &(z-1) \sum_{k=0}^{m-1} z^k (b_m - b_{\max\{n_0,k\}}) \quad \text{(siehe (55))}
\end{aligned}$$

Schätzen wir den Betrag davon mit Dreiecksungleichung ab und verwenden nochmal die geometrische Summe, diesmal für die Absolutbeträge:

$$
\begin{aligned}
&\left|\sum_{n=n_0+1}^{m} a_n(z^n-1)\right| = |z-1|\left|\sum_{k=0}^{m-1} z^k(b_m - b_{\max\{n_0,k\}})\right| \\
\leq&|z-1|\sum_{k=0}^{m-1}|z|^k|b_m - b_{\max\{n_0,k\}}| \leq |z-1|\sum_{k=0}^{m-1}|z|^k\frac{\epsilon}{2L} \quad \text{(wegen (56))} \\
=&|z-1|\frac{1-|z|^m}{1-|z|}\frac{\epsilon}{2L} \leq \frac{|z-1|}{1-|z|}\frac{\epsilon}{2L} \quad \text{(wegen } |z|^m \leq 1\text{)} \\
\leq&L\frac{\epsilon}{2L} = \frac{\epsilon}{2} \quad \text{(wegen (48) in Lemma 3.52)}
\end{aligned}
$$

Damit ist die Behauptung (59) gezeigt. □

Schon der einfachste Spezialfall $M = 0$ des Abelschen Grenzwertsatzes ist bemerkenswert:

Korollar 3.54 *Konvergiert eine Reihe $\sum_{n=0}^{\infty} a_n$ in $\mathbb{C}$, so ist die zugehörige Potenzreihe*

$$f :]-1,1] \to \mathbb{C}, \quad f(z) := \sum_{n=0}^{\infty} a_n z^n,$$

auf dem Intervall $]-1,1]$ stetig.

Beispiel: Weil die alternierende harmonische Reihe $\sum_{n=1}^{\infty} \frac{(-1)^n}{n}$ in $\mathbb{R}$ konvergiert, liefert die Logarithmusreihe

$$f(z) = \sum_{n=1}^{\infty} \frac{(-z)^n}{n}$$

eine stetige Funktion $f :]-1,1] \to \mathbb{R}$. Bemerkenswert ist hier die Stetigkeit in $z = 1$! Später werden wir zeigen: $f(z) = -\log(1+z)$ für $z \in]-1,1[$, so dass mit der Stetigkeit des Logarithmus dann die folgende bemerkenswerte Formel folgt:

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

Für praktische Zwecke, also zur Berechnung einiger Dezimalstellen von $\log 2$, ist sie allerdings ziemlich unbrauchbar, weil sie viel zu langsam konvergiert. Das liegt daran, dass wir an den Rand des Konvergenzkreises gegangen sind.

3.6.7 Konvergenzgeschwindigkeit

Die Funktionen f und g seien für alle genügend großen $x \in \mathbb{R}$ definiert, und es gelte $f \geq 0$ und $g > 0$.

Definition 3.55 *Wir sagen, $f(x)$ ist asymptotisch klein relativ zu $g(x)$ für $x \to +\infty$, wenn gilt:*

$$\frac{f(x)}{g(x)} \stackrel{x\to+\infty}{\longrightarrow} 0.$$

In Zeichen:

$$f(x) \ll g(x) \text{ für } x \to +\infty.$$

Andere Sprechweise, wenn f und g beide wachsen: "f wächst asymptotisch langsamer als g".

Die Aussage "$f(x) \ll g(x)$ für $x \to x_0$" wird analog definiert; statt Umgebungen von $+\infty$ verwendet man Umgebungen von x_0.

Warnung: Das Symbol "$\ll$" wird in der Mathematik nicht einheitlich benutzt. Daher kann seine Benutzung zu Missverständnissen führen.

Beispiele:

1. Für alle $n \in \mathbb{N}_0$ gilt:
$$x^n \ll e^x \text{ für } x \to +\infty,$$
d.h. die Exponentialfunktion wächst schneller als Potenzfunktionen.

2. Für alle $\alpha > 0$ gilt:
$$\log x \ll x^\alpha \text{ für } x \to +\infty.$$
Beweis: Zunächst gilt
$$\frac{y}{e^y} \stackrel{y\to+\infty}{\longrightarrow} 0,$$
also
$$\frac{1}{\alpha}\frac{y}{e^y} \stackrel{y\to+\infty}{\longrightarrow} 0.$$
Setzen wir $y = \alpha \log x$ ein: Wegen $\alpha > 0$ und
$$\log x \stackrel{x\to+\infty}{\longrightarrow} +\infty$$
folgt
$$\alpha \log x \stackrel{x\to+\infty}{\longrightarrow} +\infty,$$
also zusammen
$$\frac{\log x}{x^\alpha} = \frac{1}{\alpha}\frac{\alpha \log x}{e^{\alpha \log x}} \stackrel{x\to+\infty}{\longrightarrow} 0.$$
Wir verwenden hier, dass die Komposition der stetigen Funktionen
$$y \mapsto \begin{cases} \frac{1}{\alpha}\frac{y}{e^y}, & y \in \mathbb{R} \\ 0, & y = +\infty \end{cases}$$

und

$$x \mapsto \begin{cases} \alpha \log x, & 0 < x < +\infty \\ +\infty, & x = +\infty \end{cases}$$

stetig in $+\infty$ ist.

$\square$

Ordnen wir einige wichtige monoton steigende Funktionen nach ihrer Wachstumsgeschwindigkeit an:

Für $0 < \alpha < \beta$ gilt für $x \to +\infty$:

$$1 \ll \log\log x \ll \log x \ll x^\alpha \ll x^\beta \ll e^x \ll e^{x^2} \ll e^{e^x}$$

langsam wachsend — moderat wachsend — schnell wachsend

wird in Anwendungen sehr oft benutzt!

Durch Kehrwertbildung erhalten wir folgende Hierarchie fallender Funktionen:

Für $0 > -\alpha > -\beta$ gilt für $x \to +\infty$:

$$1 \gg \frac{1}{\log\log x} \gg \frac{1}{\log x} \gg x^{-\alpha} \gg x^{-\beta} \gg e^{-x} \gg e^{-x^2} \gg e^{-e^x} \gg 0$$

langsam fallend — moderat fallend — schnell fallend

Landau-Symbole. Wir führen nun einige sehr gebräuchliche Schreibweisen ein, die es oft erlauben, Grenzwertaussagen recht kompakt zu schreiben:

- **"groß O"**: "$f(x) = O(g(x))$ für $x \to x_0$" bedeutet: Es gibt $C > 0$ und eine Umgebung U von x_0, so dass für alle $x \in U$ gilt: $|f(x)| \leq C|g(x)|$.
- **"klein o"**: "$f(x) = o(g(x))$ für $x \to x_0$" bedeutet: $g(x) \neq 0$ nahe bei x_0 und

$$\frac{f(x)}{g(x)} \xrightarrow{x \to x_0} 0.$$

Bei der Verwendung von O und o muss man stets spezifizieren, auf welchen Grenzübergang $x \to x_0$ sie sich beziehen.

Warnung: Die Symbole $O(g(x))$ und $o(g(x))$ werden also für *verschiedene* Funktionen verwendet, möglicherweise von einer Verwendung zur nächsten innerhalb einer Formel verschieden. *Ihre unbedachte Verwendung ist deshalb sehr gefährlich.* Die Notationen O und o erlauben es, komplexe Grenzwertaussagen "stenographisch" zu schreiben.

Vorsicht, Fehlerquelle!

Beispiele:

1.

$$\log x = o(x^\alpha) \quad \text{für } x \to +\infty,\ \alpha > 0.$$

2.

$$e^x = 1 + x + O(x^2) \quad \text{für } x \to 0 \tag{60}$$

ist eine andere Schreibweise für[34]

$$\exists \varepsilon > 0\ \exists C > 0\ \forall x \in U_\varepsilon(0) \setminus \{0\}:\ |e^x - 1 - x| \le C|x^2|. \tag{61}$$

3.

$$e^x = 1 + x + o(x) \quad \text{für } x \to 0 \tag{62}$$

ist eine andere Schreibweise für[35]

$$\frac{e^x - 1 - x}{x} \xrightarrow{x \to 0} 0.$$

4. "$\sqrt{x+1} = 1 + \frac{x}{2} + o(x)$ für $x \to 0$" ist eine andere Notation für

$$\lim_{x \to 0} \frac{\sqrt{x+1} - 1 - \frac{x}{2}}{x} = 0.$$

5.

$$\begin{aligned} \frac{\sqrt{1+x} - 1}{e^x - 1} &= \frac{(1 + \frac{x}{2} + o(x)) - 1}{(1 + x + o(x)) - 1} \\ &= \frac{\frac{1}{2} + o(1)}{1 + o(1)} \\ &= \left(\frac{1}{2} + o(1)\right)(1 + o(1)) \\ &= \frac{1}{2} + o(1) \quad \text{für } x \to 0 \end{aligned}$$

[34]Die Formel (61) wird so bewiesen: Aus der Exponentialreihe erhalten wir

$$e^x - 1 - x = x^2 r(x)$$

mit der Potenzreihe

$$r(x) = \sum_{k=0}^{\infty} \frac{x^k}{(k+2)!}.$$

Diese Potenzreihe hat den Konvergenzradius $+\infty$, denn sie wird durch die Exponentialreihe majorisiert. Insbesondere ist $r : \mathbb{C} \to \mathbb{C}$ stetig und daher auf kompakten Mengen, z.B. der Einheitskreisscheibe, beschränkt. Wählen wir also $\varepsilon = 1$ und $C = \sup\{|r(x)||\ x \in \mathbb{C},\ |x| \le 1\} \in \mathbb{R}^+$, so folgt für alle $x \in U_\varepsilon(0) \setminus \{0\}$ die Abschätzung $|x^2 r(x)| \le C|x^2|$.

[35]Die Aussage (62) ist eine Abschwächung der Aussage (60), denn aus der Beschränktheit des Restterms r nahe bei 0 folgt $xr(x) \to 0$ für $x \to 0$.

ist eine stenographisch aufgeschriebene Rechnung, die

$$\lim_{x\to 0} \frac{\sqrt{x+1}-1}{e^x-1} = \frac{1}{2}$$

zeigt. Man beachte, dass die verschiedenen Auftreten des Symbols $o(\ldots)$ in dieser Rechnung völlig verschiedene Funktionen bezeichnen.

4 Differentialrechnung

4.1 Definition und grundlegende Eigenschaften

Es sei $U \subseteq \mathbb{R}$ (oder auch $U \subseteq \mathbb{C}$) offen, $f\colon U \to \mathbb{C}$ und $x \in U$.

Definition 4.1 f heißt *differenzierbar* in x, wenn $\frac{1}{h}(f(x+h)-f(x))$ für $h \to 0$ konvergiert. In diesem Fall schreiben wir

$$f'(x) := \frac{df(x)}{dx} := \lim_{h\to 0} \frac{1}{h}(f(x+h)-f(x)) \tag{63}$$

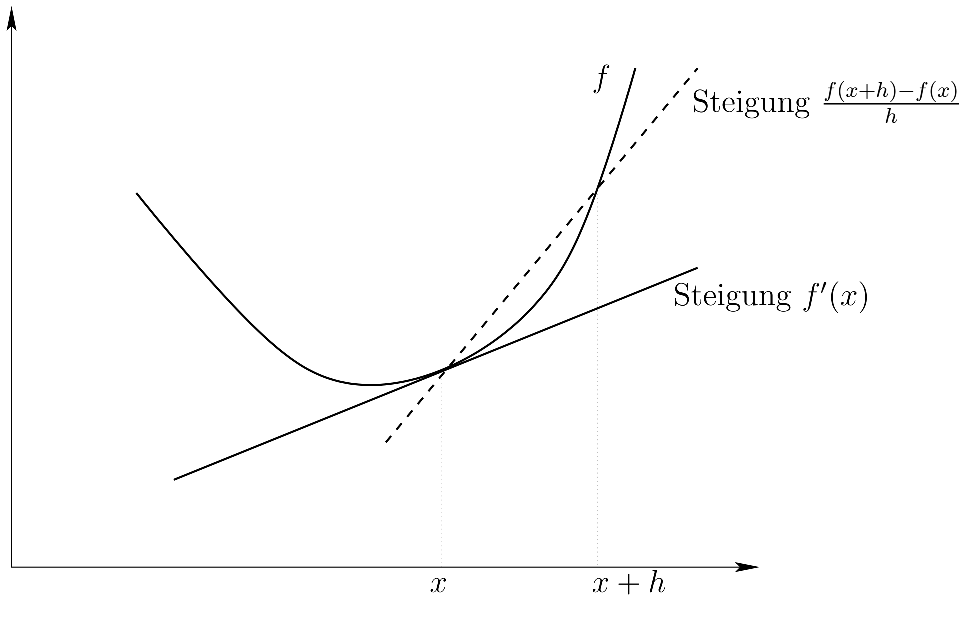
fundamentale Definition!

f' heißt *Ableitung* oder *Differentialquotient*.

Für $y = f(x)$ schreiben wir auch $\dfrac{dy}{dx}$ statt $\dfrac{df}{dx}$.

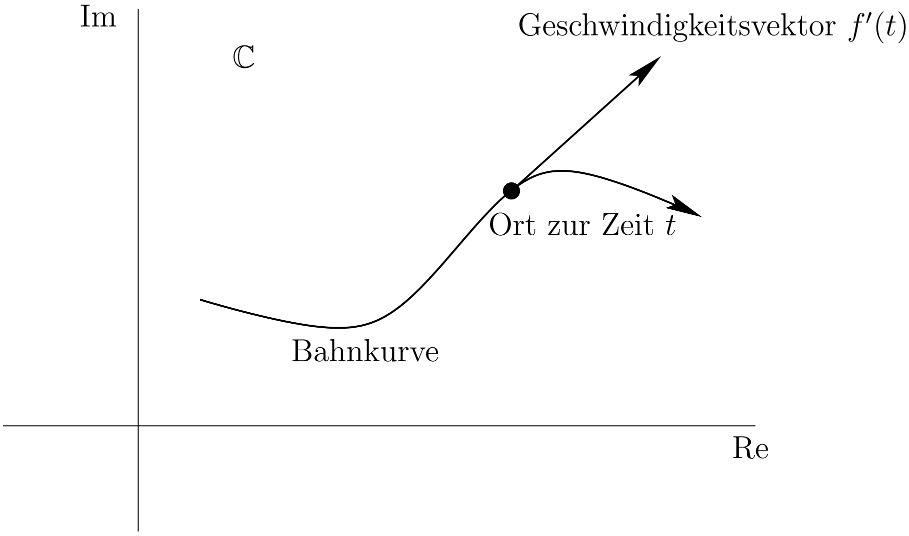
Interpretation der Ableitung:

a) **Tangentensteigung als Grenzwert von Sekantensteigungen:**



b) **Momentangeschwindigkeit**

Ein Teilchen bewege sich auf der reellen Achse (oder in der komplexen Ebene). Ist $f(t)$ der Ort des Teilchens zur Zeit t, so ist $\frac{df}{dt}(t)$ die momentane Geschwindigkeit des Teilchens zur Zeit t. Bei Bewegungen in der komplexen Ebene ist es der Geschwindigkeitsvektor zur Zeit t.



c) **Wachstumsrate**

In einem Reaktorgefäß entstehe die chemische Substanz A bei einer Reaktion. Ist $N(t)$ die Stoffmenge von A zur Zeit t, so beschreibt $\frac{dN}{dt}(t)$ die Reaktionsgeschwindigkeit zur Zeit t.

(Dabei behandeln wir N als Funktion mit kontinuierlichen Werten, lassen also die atomare Struktur der Materie im Modell unberücksichtigt.)

Anders geschrieben lautet die Definition (63) der Ableitung: *Sehr wichtig!*

$$\boxed{f(x+h) = f(x) + f'(x)h + o(h) \quad \text{für } h \to 0}$$

Nochmal anders gesagt:

f ist differenzierbar in x mit $f'(x) = a$.

$\Leftrightarrow$

Es gibt eine in x stetige Funktion $F\colon U \to \mathbb{C}$ mit $F(x) = a$ und

$$f(x+h) = f(x) + F(x+h) \cdot h$$

für $x + h \in U$.

Korollar 4.2 *Ist eine Funktion $f : U \to \mathbb{C}$ differenzierbar in $x \in U$ ($U \subseteq \mathbb{R}$ offen), so ist f dort auch stetig.*

Dies folgt unmittelbar aus der vorherigen Charakterisierung der Differenzierbarkeit.

$\square$

Beispiel 1: Für $f(x) = x^2$, $x \in \mathbb{C}$, gilt $f'(x) = 2x$.
Beweis:

$$\frac{(x+h)^2 - x^2}{h} = \frac{2hx + h^2}{h} = 2x + h \quad \stackrel{h \to 0}{\longrightarrow} \; 2x.$$

$\square$

Beispiel 2: Für $f(x) = e^x$, $x \in \mathbb{C}$, gilt $f'(x) = e^x$.
Beweis:

$$\frac{e^{x+h} - e^x}{h} = e^x \cdot \frac{e^h - 1}{h} \stackrel{h \to 0}{\longrightarrow} e^x \text{ , da } \quad \frac{e^h - 1}{h} \stackrel{h \to 0}{\longrightarrow} 1$$

$\square$

Wir können das auch so formulieren:

$\boxed{\text{Die Exponentialfunktion exp ist eine Lösung der Differentialgleichung } f' = f.}$

Fassen wir zusammen:

Differenzierbarkeit = Approximierbarkeit durch eine *lineare* Funktion.

$$\begin{array}{c} y = f(x) \\ \text{(nichtlinear)} \end{array} \xrightarrow{\text{Linearisierung bei } x_0} \begin{array}{c} y = f(x_0) + (x - x_0) f'(x_0) \\ \text{Tangentengleichung:} \\ \text{lineare Approximation bei } x_0 \end{array}$$

Wichtigste Bedeutung der Ableitung!

Zur Illustration siehe auch die Folien in
`http://www.mathematik.uni-muenchen.de/~merkl/ws12/ana1/diffbarkeit.pdf`
oder
`http://www.mathematik.uni-muenchen.de/~merkl/ws12/ana1/diffbarkeit.ps`

Es seien $f, g \colon U \to \mathbb{C}$, $x \in U$, U offen in $\mathbb{R}$ oder $\mathbb{C}$.

Satz 4.3 (Rechenregeln für die Ableitung) *Sind f und g differenzierbar in x, so sind auch $f+g, f-g, f\cdot g$ und für $g(x)\neq 0$ auch $\frac{f}{g}$ differenzierbar in x, und es gilt im Punkt x:*

a) $(f\pm g)' = f'\pm g'$

b) $(fg)' = f'g + fg'$ *„Produktregel"*

c) $\left(\dfrac{f}{g}\right)' = \dfrac{f'g - fg'}{g^2}$ *„Quotientenregel"*

Beweis: Es seien

$$\begin{aligned} f(x+h) &= f(x) + F(x+h)h \\ g(x+h) &= g(x) + G(x+h)h \end{aligned}$$

mit in x stetigen Funktionen $F, G\colon U\to\mathbb{C}$ mit $F(x)=f'(x)$, $G(x)=g'(x)$.
Dann gilt:

a)
$$f(x+h)\pm g(x+h) = f(x)\pm g(x) + [F(x+h)\pm G(x+h)]h$$
wobei $F(x+h)\pm G(x+h)$ stetig in $h=0$ ist und
$$F(x)\pm G(x) = f'(x)\pm g'(x).$$

b)
$$\begin{aligned} f(x+h)g(x+h) &= [f(x)+F(x+h)h][g(x)+G(x+h)h] \\ &= f(x)g(x) + [F(x+h)g(x) + f(x)G(x+h) + F(x+h)G(x+h)h]h \end{aligned}$$

Der Term in eckigen Klammern auf der rechten Seite ist stetig in $h=0$ mit dem Wert $f'(x)g(x)+f(x)g'(x)$.

c) Wir zeigen zunächst, daß an der Stelle x gilt:

$$\boxed{\left(\frac{1}{g}\right)' = -\frac{g'}{g^2}}$$

In der Tat:
$$\frac{1}{h}\left(\frac{1}{g(x+h)} - \frac{1}{g(x)}\right) = -\underbrace{\frac{1}{g(x+h)g(x)}}_{\xrightarrow{h\to 0}\frac{1}{g(x)^2}}\underbrace{\frac{g(x+h)-g(x)}{h}}_{\xrightarrow{h\to 0} g'(x)} \xrightarrow{h\to 0} -\frac{g'(x)}{g(x)^2}.$$

Es folgt mit Hilfe der Produktregel an der Stelle x:

$$\left(\frac{f}{g}\right)' = \left(f \cdot \frac{1}{g}\right)' = f' \cdot \frac{1}{g} + f \cdot \left(\frac{1}{g}\right)' = \frac{f'}{g} - \frac{fg'}{g^2} = \frac{f'g - fg'}{g^2}.$$

$\square$

Folgerung: Es gilt für alle $n \in \mathbb{N}_0$:

$$\boxed{\frac{d(x^n)}{dx} = nx^{n-1}}$$

Beweis durch Induktion über n:

$\boxed{n = 0}$

$$\frac{d(x^0)}{dx} = \frac{d}{dx}1 = 0.$$

$\boxed{n = 1}$

$$\frac{dx}{dx} = \lim_{h \to 0} \frac{(x+h) - x}{h} = 1.$$

$\boxed{n \rightsquigarrow n+1}$ Es gelte die Behauptung für n. Dann folgt:

$$\begin{aligned} \frac{d}{dx} x^{n+1} &= \frac{d}{dx}(x \cdot x^n) \\ &= 1 \cdot x^n + x \cdot \frac{d(x^n)}{dx} && \text{nach der Produktregel} \\ &= 1 \cdot x^n + x \cdot nx^{n-1} && \text{nach der Induktionsvoraussetzung} \\ &= (n+1)x^n. \end{aligned}$$

Satz 4.4 (Kettenregel) *Es seien U und V in $\mathbb{R}$ oder in $\mathbb{C}$ offen, $f\colon U \to V$, $g\colon V \to \mathbb{C}$. Die Funktion f sei differenzierbar in $x \in U$, und g sei differenzierbar in $f(x) \in V$. Dann ist $g \circ f$ differenzierbar in x, und es gilt*

Wichtigste Ableitungsregel!

$$\boxed{(g \circ f)'(x) = g'(f(x)) \cdot f'(x)}$$

Bemerkung: Schreiben wir $y = f(x)$ und $z = g(y)$, so kann man die Kettenregel in der folgenden intuitiven Notation schreiben:

$$\boxed{\frac{dz}{dx} = \frac{dz}{dy}\frac{dy}{dx}}$$

In dieser Notation bleibt die Stelle, an der die Ableitungen $\frac{dz}{dy}$ bzw. $\frac{dy}{dx}$ auszuwerten sind, implizit.

Beweis der Kettenregel: Wir kürzen ab: $y = f(x)$.
Es sei

$$\begin{aligned} f(x+h) &= f(x) + F(x+h)\cdot h, \qquad x+h \in U,\\ g(y+k) &= g(y) + G(y+k)\cdot k, \qquad y+k \in V, \end{aligned}$$

wobei F stetig in x mit $F(x) = f'(x)$ und G stetig in y mit $G(y) = g'(y)$ sein soll.
Durch Einsetzen erhalten wir

$$\begin{aligned} g(f(x+h)) &= g(f(x) + F(x+h)\cdot h)\\ &= g(y) + G(y + F(x+h)\cdot h)\cdot F(x+h)\cdot h \end{aligned}$$

Nun gilt $F(x+h) \overset{h\to 0}{\longrightarrow} f'(x)$ und $y + F(x+h)\cdot h \overset{h\to 0}{\longrightarrow} y$, also

$$G(y + F(x+h)\cdot h) \overset{h\to 0}{\longrightarrow} G(y) = g'(f(x)),$$

weil G stetig in y ist. Es folgt:

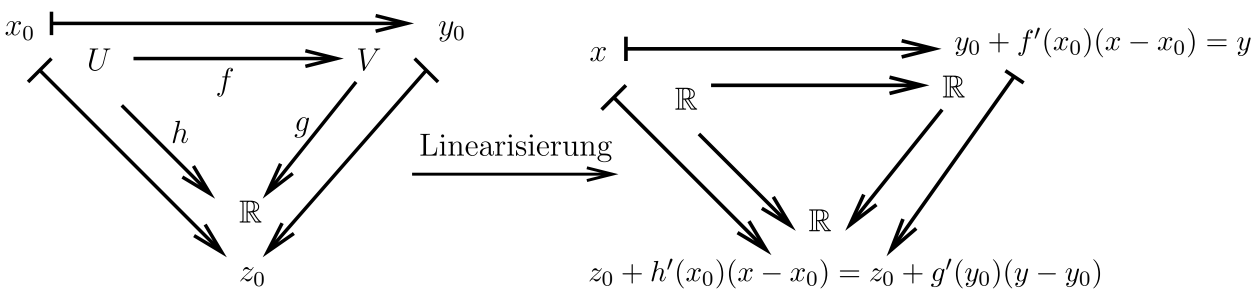
$$G(y + F(x+h)\cdot h)F(x+h) \overset{h\to 0}{\longrightarrow} g'(f(x))\cdot f'(x).$$

Interpretation der Kettenregel

Verträglichkeit von Linearisierung mit Komposition *wichtig!*

nichtlinear **lineare Approximation**

$h = g\circ f$ $h'(x_0) = g'(y_0)\cdot f'(x_0)$



Linearisierung der Komposition = Komposition der Linearisierungen

Diese Interpretation ist auch im Hinblick auf eine höherdimensionale Verallgemeinerung im kommenden Semester essentiell für ein gutes Verständnis der Differentialrechung!

Beispiele:

1. Es sei $\lambda \in \mathbb{C}$. Man berechne $\frac{d}{dx}e^{\lambda x}$.

   Lösung: Wir setzen $y = \lambda x, z = e^y$, und erhalten

   $$\frac{dz}{dx} = \frac{dz}{dy}\frac{dy}{dx} = e^y \cdot \lambda = \lambda e^{\lambda x}.$$

   Das bedeutet:

   $$\boxed{\text{Die Differentialgleichung } f' = \lambda f \text{ hat eine Lösung } f(x) = e^{\lambda x}.}$$

2. Man berechne $\frac{d}{dx}\log x$ für $x > 0$ unter der Annahme, daß die Logarithmusfunktion differenzierbar ist.

Lösung:

$$1 = \frac{dx}{dx} = \frac{d}{dx}e^{\log x} = e^{\log x} \cdot \frac{d}{dx}\log x = x \cdot \frac{d}{dx}\log x,$$

also

$$\boxed{\frac{d}{dx}\log x = \frac{1}{x}}$$

In der Tat ist der natürliche Logarithmus differenzierbar. Es gilt nämlich allgemeiner:

Satz 4.5 (Ableitung der Umkehrfunktion) *Es sei $U \subseteq \mathbb{R}$ ein offenes Intervall. $f\colon U \to \mathbb{R}$ stetig und streng monoton, und $x \in U$. Wenn f an der Stelle x differenzierbar ist mit Ableitung $f'(x) \neq 0$, so ist die Umkehrfunktion f^{-1} von f differenzierbar in $f(x)$, und es gilt:*

$$\boxed{(f^{-1})'(f(x)) = \frac{1}{f'(x)}}$$

Beweis: Nach dem Zwischenwertsatz ist $f(x)$ ein innerer Punkt von $f[U]$, denn für genügend kleine $\varepsilon > 0$ liegen alle Zahlen zwischen $f(x-\varepsilon)$ und $f(x+\varepsilon)$ in $f[U]$. Es sei $(y_n)_{n\in\mathbb{N}_0}$ eine Folge in $f[U] \setminus \{f(x)\}$ mit $y_n \overset{n\to\infty}{\longrightarrow} y := f(x)$. Die Funktion f^{-1} ist stetig, weil $f\colon U \to f[U]$ streng monoton ist. Wir setzen $x_n := f^{-1}(y_n)$. Es folgt:

$$x_n = f^{-1}(y_n) \overset{n\to\infty}{\longrightarrow} f^{-1}(y) = x,$$

also

$$\frac{f(x_n) - f(x)}{x_n - x} \overset{n\to\infty}{\longrightarrow} f'(x)$$

und folglich

$$\frac{f^{-1}(y_n) - f^{-1}(y)}{y_n - y} = \frac{x_n - x}{f(x_n) - f(x)} \overset{n\to\infty}{\longrightarrow} \frac{1}{f'(x)}.$$

Das bedeutet: f^{-1} ist differenzierbar in y mit $(f^{-1})'(y) = \frac{1}{f'(x)}$.

□

In der Kurznotation läßt sich die Ableitung der Umkehrfunktion so schreiben: Für $y = f(x)$, $x = f^{-1}(y)$:

$$\boxed{\frac{dx}{dy} = \frac{1}{\frac{dy}{dx}}}$$

Wieder bleibt implizit, aus welcher Stelle die Funktionen ausgewertet werden.

Beispiel: Für $x = \sqrt{y}$, $y > 0$ erhalten wir $y = x^2$, also

$$\frac{d}{dy}\sqrt{y} = \frac{dx}{dy} = \frac{1}{\frac{dy}{dx}} = \frac{1}{2x} = \frac{1}{2\sqrt{y}}.$$

Weitere Beispiele zur Kettenregel:
Für $x > 0$ und feste $s \in \mathbb{C},\ a > 0$ berechne man

1. $\frac{d}{dx}x^s$,
2. $\frac{d}{dx}a^x$,
3. $\frac{d}{dx}x^x$.

Lösung:

1. Es gilt: $x^s = e^{s \log x}$.

   Es folgt:

   $$\begin{aligned}\frac{d}{dx}x^s &= e^{s\log x}\frac{d}{dx}(s\log x) = e^{s\log x} \cdot s \cdot \frac{1}{x} \\ &= \frac{sx^s}{x} = sx^{s-1}.\end{aligned}$$

   Die Formel

   $$\boxed{\frac{d}{dx}x^s = sx^{s-1}}$$

   gilt also für alle $s \in \mathbb{C}$.

   Beispiele dazu:

   $$\frac{d}{dx}\sqrt{x} = \frac{d}{dx}x^{\frac{1}{2}} = \frac{1}{2}x^{\frac{1}{2}-1} = \frac{1}{2\sqrt{x}}$$

   $$\frac{d}{dx}\frac{1}{x} = \frac{d}{dx}x^{-1} = -1 \cdot x^{-2} = -\frac{1}{x^2}.$$

2. Es gilt:

   $$\frac{d}{dx}a^x = \frac{d}{dx}e^{x\log a} = e^{x\log a}\frac{d}{dx}(x\log a) = a^x \log a.$$

3.

$$\frac{d}{dx}x^x = \frac{d}{dx}e^{x\log x} = e^{x\log x}\frac{d}{dx}(x\log x)$$
$$= x^x\left(\frac{dx}{dx}\log x + x\frac{d}{dx}(\log x)\right) = x^x\left(\log x + x\cdot\frac{1}{x}\right)$$
$$= x^x(\log x + 1)$$

4.2 Exkurs: Trigonometrische Funktionen und Hyperbelfunktionen

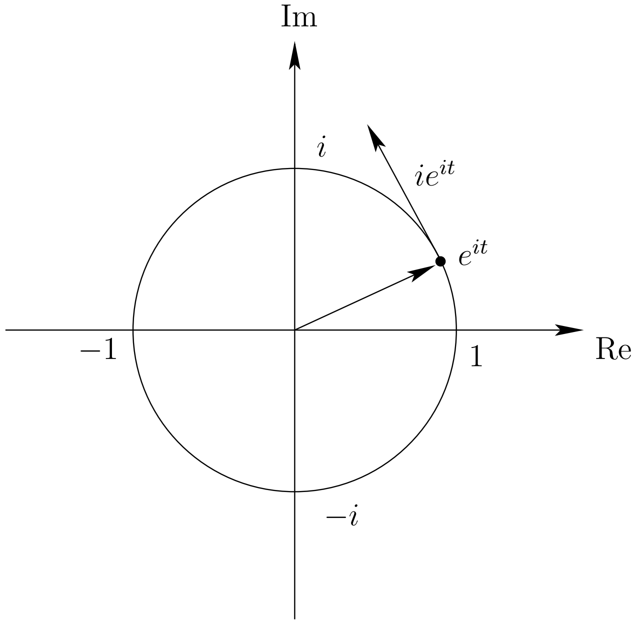
In diesem Abschnitt untersuchen wir die komplexe Exponentialfunktion etwas genauer. Wir berechnen nun die Geschwindigkeit, mit der e^{it} auf dem Einheitskreis läuft:

$$\frac{d}{dt}\,e^{it} = ie^{it}$$

also:

$$\left|\frac{d}{dt}\,e^{it}\right| = 1$$

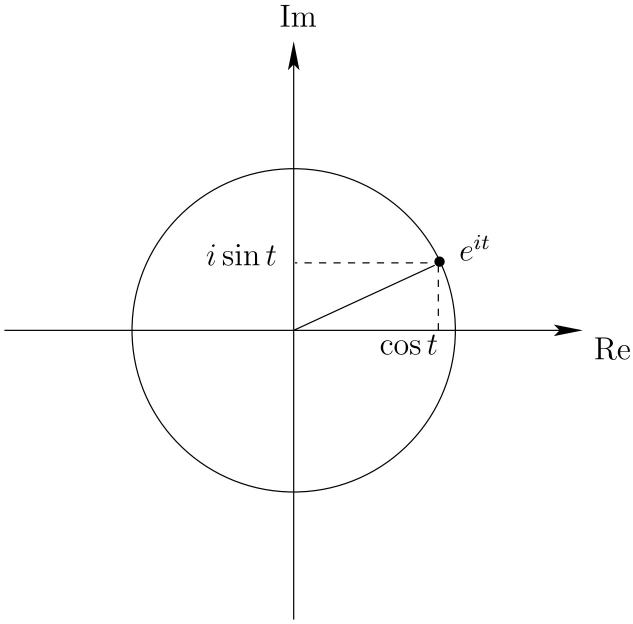
e^{it} läuft also mit Geschwindigkeit 1 in positiver Richtung um den Einheitskreis und es gilt: $e^{i0} = 1$.



Ein Vergleich mit der elementargeometrischen Definition von sin und cos zeigt:
Für $t \in \mathbb{R}$ gilt die *Eulersche Formel*:

$$\boxed{e^{it} = \cos t + i\sin t}$$

(Zum Beispiel gilt: $e^{i\frac{\pi}{2}} = i,\ e^{i\pi} = -1,\ e^{2\pi i} = 1$)



Anders gesagt: Für $t \in \mathbb{R}$ gilt:

$$\boxed{\begin{aligned} \cos t &= \operatorname{Re} e^{it} = \frac{e^{it} + e^{-it}}{2} \\ \sin t &= \operatorname{Im} e^{it} = \frac{e^{it} - e^{-it}}{2i} \end{aligned}}$$

Die Additionstheoreme von sin und cos ergeben sich nun als einfache Folgerungen aus der Funktionalgleichung der Exponentialfunktion:
Für $\alpha, \beta \in \mathbb{R}$ gilt:

$$\begin{aligned} \cos(\alpha+\beta) + i\sin(\alpha+\beta) &= e^{i(\alpha+\beta)} \\ &= e^{i\alpha} e^{i\beta} \\ &= (\cos\alpha + i\sin\alpha)(\cos\beta + i\sin\beta) \\ &= (\cos\alpha\cos\beta - \sin\alpha\sin\beta) + i(\sin\alpha\cos\beta + \cos\alpha\sin\beta) \end{aligned}$$

also:

$$\boxed{\begin{aligned} \cos(\alpha+\beta) &= \cos\alpha\cos\beta - \sin\alpha\sin\beta \\ \sin(\alpha+\beta) &= \sin\alpha\cos\beta + \cos\alpha\sin\beta \end{aligned}}$$

Im Spezialfall $\alpha + \beta = 0$ erhalten wir:

$$\boxed{\cos^2\alpha + \sin^2\alpha = 1}$$

Interpretation der Additionstheoreme:

Die um $-\beta$ verschobene Cosinus- bzw. Sinusfunktion

$$x \mapsto \cos(x+\beta) \quad \text{bzw.} \quad x \mapsto \sin(x+\beta)$$

ist eine Linearkombination der Cosinus- und Sinusfunktion.

Die Polardarstellung einer komplexen Zahl $z = r(\cos\varphi + i\sin\varphi)$ lässt sich nun einfacher so schreiben:

$$\boxed{z = re^{i\varphi}} \qquad r = |z|$$

Für $z = re^{i\varphi} \in \mathbb{C}\setminus\{0\};\ r, \varphi \in \mathbb{R}$ erhalten wir:

$$z = e^{\log r + i\varphi} = e^{\log r + i(2\pi k + \varphi)}$$

für alle $k \in \mathbb{Z}$. Das bedeutet:

Der Logarithmus im Komplexen ist mehrdeutig. Er ist nur bis auf ganzzahlige Vielfache von $2\pi i$ bestimmt.

Potenzreihe von sin und cos: Es gilt für $x \in \mathbb{R}$:

$$\begin{aligned} e^{ix} &= 1 + ix + \frac{i^2x^2}{2} + \frac{i^3x^3}{3!} + \frac{i^4x^4}{4!} + \frac{i^5x^5}{5!} + \dots \\ &= 1 + ix - \frac{x^2}{2} - i\frac{x^3}{3!} + \frac{x^4}{4!} + i\frac{x^5}{5!} \pm \dots \end{aligned}$$

Fassen wir Real- und Imaginärteil zusammen:

$$\boxed{\begin{aligned} \cos x &= 1 - \frac{x^2}{2} + \frac{x^4}{4!} - \frac{x^6}{6!} \pm \dots = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} x^{2k} \\ \sin x &= x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} \pm \dots = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} x^{2k+1} \end{aligned}}$$

Zur Illustration siehe auch die Folien in
`http://www.mathematik.uni-muenchen.de/~merkl/ws12/ana1/reihe.pdf`
oder
`http://www.mathematik.uni-muenchen.de/~merkl/ws12/ana1/reihe.ps`

Ableitung von Sinus und Cosinus:
Aus

$$\frac{d}{dx}(\cos x + i\sin x) = \frac{d}{dx}e^{ix} = ie^{ix} = i(\cos x + i\sin x) = -\sin x + i\cos x, \qquad x \in \mathbb{R},$$

erhalten wir durch Real- und Imaginärteilbildung:

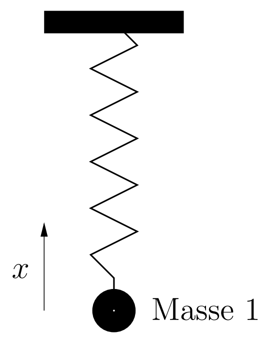
$$\boxed{\begin{aligned} \frac{d}{dx}\cos x &= -\sin x \\ \frac{d}{dx}\sin x &= \cos x \end{aligned}}$$

Anders gesagt:

> Das Differentialgleichungssystem
>
> $$y_1' = -y_2$$
> $$y_2' = y_1$$
>
> hat die Lösungen
>
> $$y_1^{(1)} = \cos x,\ y_2^{(1)} = \sin x$$
>
> zusammen mit
>
> $$y_1^{(2)} = -\sin x,\ y_2^{(2)} = \cos x.$$

Weitere Lösungen ergeben sich durch Linearkombination, denn mit $y_1^{(1)}, y_2^{(1)}$ und $y_1^{(2)}, y_2^{(2)}$ sind auch $y_1 = \alpha y_1^{(1)} + \beta y_1^{(2)}$ zusammen mit $y_2 = \alpha y_2^{(1)} + \beta y_2^{(2)}$ Lösungen des Systems, wobei α und β beliebige reelle (oder auch komplexe) Zahlen sind.

Interpretation: Schwingungsvorgänge:
Eine Einheitsmasse hängt an einer Feder (Federkonstante 1, Ort $x(t)$ zur Zeit t, Geschwindigkeit $v(t)$ zur Zeit t, Ruhelage 0). Bei Auslenkung um x aus der Ruhelage wirkt effektiv die Kraft $-x$ auf die Masse (entgegengesetzt zu x).



Für die Geschwindigkeit der Masse gilt also:

$$\begin{aligned} \frac{dv}{dt} &= -x \qquad &\text{(Newtonsches Gesetz)} \\ \frac{dx}{dt} &= v \qquad &\text{(Definition der Geschwindigkeit)} \end{aligned}$$

Wir haben also die Schwingungslösungen:

$$\boxed{\begin{aligned} x(t) &= \alpha \cos t + \beta \sin t \\ v(t) &= -\alpha \sin t + \beta \cos t \end{aligned}}$$

wobei $\alpha, \beta \in \mathbb{R}$.
Nach den Additionstheoremen sind diese Lösungen Vielfache von verschobenen Sinus- und Cosinusfunktionen.
Schreiben wir für die zweite Ableitung, also die Ableitung der Ableitung:

$$f''(x) = \frac{d}{dx}\frac{d}{dx} f(x) = \frac{d^2}{dx^2} f(x)$$

so erhalten wir:

> *Die Schwingungsgleichung*
>
> $$f'' + f = 0$$
>
> hat mit $\alpha, \beta, x \in \mathbb{R}$ die folgenden Funktionen als Lösungen:
>
> $$f(x) = \alpha \cos x + \beta \sin x$$

Wir werden später sehen, daß dies die einzigen Lösungen auf $\mathbb{R}$ sind.

Die Zahl π. Bisher haben wir die Kreiszahl π aus der Elementargeometrie als bekannt vorausgesetzt, aber sie noch nicht im strengen Sinn ohne Rückgriff auf die Elementargeometrie eingeführt. Das holen wir nun nach, zusammen mit einer Einführung der Polarkoordinaten, die ebenfalls nicht auf die geometrische Anschauung zurückgreift. Wir starten mit einer oberen Schranke der Cosinusfunktion im Intervall $]0, 2]$:

Lemma 4.6 *Für $x \in]0, 2]$ gilt*

$$\cos x < 1 + \frac{x^2}{2}\left[-1 + \frac{x^2}{12}\right] < 1. \tag{64}$$

Insbesondere ist

$$\cos 2 < -\frac{1}{3}.$$

Beweis. Gegeben $x \in]0,2]$, rechnen wir mit der Cosinusreihe:

$$\begin{aligned}
\cos x &= 1 + \sum_{n=1}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n} \\
&= 1 + \sum_{m=0}^{\infty} \left[\frac{(-1)^{2m+1}}{(4m+2)!} x^{4m+2} + \frac{(-1)^{2m+2}}{(4m+4)!} x^{4m+4} \right] \\
&= 1 + \sum_{m=0}^{\infty} \frac{x^{4m+2}}{(4m+2)!} \left[-1 + \frac{x^2}{(4m+3)(4m+4)} \right]
\end{aligned}$$

wobei die Umordnung der Cosinusreihe wegen ihrer absoluten Konvergenz erlaubt ist. Alle Summanden in der letzten Summe sind negativ, denn für alle $m \in \mathbb{N}_0$ gilt:

$$-1 + \frac{x^2}{(4m+3)(4m+4)} \leq -1 + \frac{1}{(4m+3)(m+1)} \leq -1 + \frac{1}{3} = -\frac{2}{3} < 0$$

Wir schätzen damit alle Summanden nach oben durch 0 ab, schreiben den Summanden zu $m = 0$ aber nochmal gesondert hin:

$$\cos x = 1 + \sum_{m=0}^{\infty} \frac{x^{4m+2}}{(4m+2)!} \left[-1 + \frac{x^2}{(4m+3)(4m+4)} \right] < 1 + \frac{x^2}{2!} \left[-1 + \frac{x^2}{3 \cdot 4} \right] = 1 - \frac{x^2}{2} + \frac{x^4}{24} < 1$$

Speziell für $x = 2$ erhalten wir:

$$\cos 2 < 1 - \frac{2^2}{2} + \frac{2^4}{24} = -\frac{1}{3} < 0.$$

□

Lemma 4.7 *Die Cosinusfunktion* $\cos : \mathbb{R} \to \mathbb{R}$ *besitzt eine kleinste positive Nullstelle. Sie liegt im Intervall* $]0,2[$.

Beweis. Es gilt $\cos 0 = 1 > 0$ und nach dem eben Gezeigten $\cos 2 < 0$. Da die Cosinusfunktion stetig ist, besitzt sie nach dem Zwischenwertsatz also eine Nullstelle im Intervall $]0,2[$. Als Urbild der 0 unter der Cosinusfunktion, die stetig ist, ist das Nullstellengebilde der Cosinusfunktion abgeschlossen. Also ist $M = \{x \in [0,2] |\ \cos x = 0\}$ abgeschlossen und beschränkt, also kompakt, und zusätzlich nach dem eben Gezeigten nichtleer. Die Menge M besitzt also ein Minimum. Dieses Minimum kann nicht gleich 0 sein, da $\cos 0 = 1$, und nicht gleich 2, da $\cos 2 < 0$. Es liegt also im offenen Intervall $]0,2[$ und ist damit die kleinste positive Nullstelle der Cosinusfunktion.

□

Definition 4.8 (Die Kreiszahl π) *Es sei π das Doppelte der kleinsten positiven Nullstelle der Cosinusfunktion.*

Nach dem eben Gezeigten wissen wir $\pi \in]0,4[$ und $\cos x \neq 0$ für alle $x \in [0, \frac{\pi}{2}[$. Es gilt sogar $\cos x > 0$ für diese x, weil sonst nach dem Zwischenwertsatz eine weitere Nullstelle der Cosinusfunktion im Intervall $[0, x]$ existieren müsste, was nicht der Fall ist.

Lemma 4.9 *Es gilt* $\sin x > 0$ *für alle* $x \in]0, \frac{\pi}{2}]$. *Weiter gilt* $\sin \frac{\pi}{2} = 1$,

$$\boxed{e^{i\frac{\pi}{2}} = i, \quad e^{i\pi} = -1, \quad e^{2\pi i} = 1}$$

und damit auch

$$\boxed{\forall k \in \mathbb{Z}: \; e^{2\pi ik} = 1.}$$

Beweis. Da

$$\lim_{x\to 0} \frac{\sin x}{x} = \sin'(0) = \cos 0 = 1 > 0$$

folgt $\sin x > 0$ für alle genügend kleinen $x > 0$. Nun kann die Sinusfunktion keine Nullstelle im Intervall $]0, \frac{\pi}{2}]$ besitzen, denn sonst müsste wegen $\cos^2 + \sin^2 = 1$ der Kosinus dort den Wert $+1$ oder -1 annehmen, was nicht der Fall ist, da die Kosinusfunktion im Intervall $]0, \frac{\pi}{2}]$ Werte im Intervall $[0, 1[$ annimmt. Da die Sinusfunktion stetig ist, folgt nach dem Zwischenwertsatz, dass $\sin x > 0$ sogar für alle $x \in]0, \frac{\pi}{2}]$ gilt. Wegen $\sin^2 \frac{\pi}{2} = \cos^2 \frac{\pi}{2} + \sin^2 \frac{\pi}{2} = 1$ folgt hieraus $\sin \frac{\pi}{2} = 1$. Damit ist gezeigt:

$$e^{i\frac{\pi}{2}} = \cos \frac{\pi}{2} + i \sin \frac{\pi}{2} = i.$$

Es folgt:

$$e^{i\pi} = e^{2i\frac{\pi}{2}} = (e^{i\frac{\pi}{2}})^2 = i^2 = -1$$

und

$$e^{2\pi i} = (e^{i\pi})^2 = (-1)^2 = 1.$$

□

Lemma 4.10 (Der Kern der komplexen Exponentialabbildung) *Es gilt:*

$$\{z \in \mathbb{C} |\; e^z = 1\} = \{2\pi ik |\; k \in \mathbb{Z}\}$$

In der Sprache der Algebra können wir das auch so sagen: Der Kern des Gruppenhomomorphismus

$$\exp : (\mathbb{C}, +) \to (\mathbb{C} \setminus \{0\}, \cdot)$$

besteht genau aus den ganzzahligen Vielfachen[36] *von* $2\pi i$:

$$\boxed{\ker(\exp) = 2\pi i\mathbb{Z}}$$

[36] verwendete Abkürzung: $wM = \{wz |\; z \in M\}$ für $w \in \mathbb{C}$, $M \subseteq \mathbb{C}$.

Beweis: "$\supseteq$:" Für $k \in \mathbb{N}_0$ gilt:

$$e^{2\pi ik} = (e^{2i\pi})^k = 1^k = 1,$$

und hieraus auch

$$e^{2\pi i(-k)} = \frac{1}{e^{2\pi ik}} = 1,$$

also $e^{2\pi ik} = 1$ für alle $k \in \mathbb{Z}$.

"$\subseteq$:" Es sei $z \in \mathbb{C}$ mit $e^z = 1$ gegeben. Wir zerlegen z in Real- und Imaginärteil: $z = a + ib$ mit $a, b \in \mathbb{R}$. Dann gilt $e^a > 0$ und $|e^{ib}| = 1$, also

$$e^0 = 1 = |e^z| = e^a|e^{ib}| = e^a$$

und daher $a = 0$, da die *reelle* Exponentialfunktion $\exp : \mathbb{R} \to \mathbb{R}$ injektiv ist. Die Zahl $z = ib$ ist also imaginär. Wir setzen

$$k = \left\lfloor \frac{b}{2\pi} \right\rfloor \in \mathbb{Z}$$

und

$$x = b - 2\pi k \in [0, 2\pi[,$$

also $b = 2\pi k + x$. Zu zeigen ist nun $x = 0$, denn hieraus folgt $z = ib = 2\pi ik \in 2\pi i\mathbb{Z}$. Wir zeigen das indirekt und nehmen hierzu $x \in]0, 2\pi[$ an. Einerseits gilt nun

$$e^{ix} = e^{2\pi ik}e^{ix} = e^{ib} = 1$$

Setzen wir nun $w = e^{i\frac{x}{4}}$, so folgt:

$$(w + 1)(w - 1)(w + i)(w - i) = w^4 - 1 = e^{4i\frac{x}{4}} - 1 = e^{ix} - 1 = 0,$$

also $w \in \{\pm 1, \pm i\}$. Andererseits folgt $\frac{x}{4} \in]0, \frac{\pi}{2}[$ und daher $\sin\frac{x}{4} > 0$ und $\cos\frac{x}{4} > 0$. Es folgt $w = \cos\frac{x}{4} + i\sin\frac{x}{4} \notin \{\pm 1, \pm i\}$, ein Widerspruch.

□

Lemma 4.11 (Bildbereiche der Exponentialfunktion) *Der Wertebereich der reellen Exponentialfunktion* $\exp : \mathbb{R} \to \mathbb{R}$ *ist die Menge* $\mathbb{R}^+$ *der positiven reellen Zahlen. Das Bild der imaginären Achse* $i\mathbb{R} = \{ib|\ b \in \mathbb{R}\}$ *unter der Exponentialfunktion ist die Einheitskreislinie*

$$S^1 := \{z \in \mathbb{C}|\ |z| = 1\}.$$

Der Wertebereich der komplexen Exponentialfunktion ist die Menge $\mathbb{C} \setminus \{0\}$*. In Formeln:*

$$\boxed{\exp[\mathbb{R}] = \mathbb{R}^+, \quad \exp[i\mathbb{R}] = S^1, \quad \exp[\mathbb{C}] = \mathbb{C} \setminus \{0\}}$$

Beweis: Wir wissen schon, dass $\exp : \mathbb{R} \to \mathbb{R}^+$ eine Bijektion ist. Insbesondere ist $\exp[\mathbb{R}] = \mathbb{R}^+$. Weiter gilt $\exp[i\mathbb{R}] \subseteq S^1$, da für imaginäre $ib \in i\mathbb{R}$ gilt: $|e^{ib}| = 1$. Um auch die umgekehrte Inklusion $\exp[i\mathbb{R}] \supseteq S^1$ zu zeigen, sei $z = a + ib \in S^1$ gegeben, $a, b \in \mathbb{R}$. Insbesondere gilt $a^2 + b^2 = 1$. Wir unterscheiden mehrere Fälle:

- **1. Fall:** $a \geq 0$.
  - **1.1. Fall:** $b \geq 0$. Wegen $a^2 + b^2 = 1$ ist $a, b \in [0, 1]$. Wegen $\cos 0 = 1$ und $\cos \frac{\pi}{2} = 0$ gibt es nach dem Zwischenwertsatz ein $x \in [0, \frac{\pi}{2}]$ mit $\cos x = a$. Es folgt $b^2 = 1 - a^2 = 1 - \cos^2 x = \sin^2 x$ und daher $b = \sin x$ wegen $b \geq 0$ und $\sin x \geq 0$. Das bedeutet: $e^{ix} = a + ib = z$, also $z \in \exp[i\mathbb{R}]$.
  - **1.2. Fall:** $b < 0$. Nach dem 1.1. Fall gibt es ein $x \in \mathbb{R}$ mit $e^{ix} = a - ib = \overline{z}$. Es folgt $e^{-ix} = \overline{\overline{z}} = z$, also $z \in \exp[i\mathbb{R}]$.
- **2. Fall:** $a < 0$. Dann ist $-z = -a - ib$ mit $-a > 0$. Nach dem 1. Fall gibt es ein $x \in \mathbb{R}$ mit $-z = e^{ix}$. Es folgt $z = -ze^{i\pi} = e^{ix}e^{i\pi} = e^{i(x+\pi)}$, also ebenfalls $z \in \exp[i\mathbb{R}]$.

Damit ist $\exp[i\mathbb{R}] = S^1$ gezeigt.
Es gilt $\exp[\mathbb{C}] \subseteq \mathbb{C} \setminus \{0\}$, denn für $z \in \mathbb{C}$ gilt $e^z e^{-z} = e^{z-z} = e^0 = 1$, also $e^z \neq 0$. Um auch die umgekehrte Inklusion $\exp[\mathbb{C}] \supseteq \mathbb{C} \setminus \{0\}$ zu zeigen, sei $z \in \mathbb{C} \setminus \{0\}$ gegeben. Setzen wir $r = |z|$, so ist $r > 0$ und $z/r \in S^1$ wegen $|z/r| = |z|/r = 1$; also existieren $a = \log r \in \mathbb{R}$ und ein $\varphi \in \mathbb{R}$ mit $e^{i\varphi} = z/r$. Es folgt $z = re^{i\varphi} = e^{a+i\varphi} \in \exp[\mathbb{C}]$.

□

Wir können mit der nun entwickelten Theorie auch Polarkoordinaten rigoros, ohne Rückgriff auf die geometrische Anschauung, einführen:

Korollar 4.12 (Polarkoordinaten) *Jede von 0 verschiedene komplexe Zahl $z \in \mathbb{C} \setminus \{0\}$ besitzt eine Darstellung der Gestalt*

$$z = re^{i\varphi} = r(\cos \varphi + i \sin \varphi)$$

mit $r > 0$, $\varphi \in \mathbb{R}$. Dabei ist $r = |z|$ eindeutig bestimmt, während φ nur bis auf ganzzahlige Vielfache von 2π eindeutig bestimmt ist.

Beweis: *Existenz der Polardarstellung:* Setzen wir wie oben $r = |z| > 0$, so finden wir wie eben ein $\varphi \in \mathbb{R}$ mit $e^{i\varphi} = z/r$. Es folgt $z = re^{i\varphi}$.
Eindeutigkeit der Polardarstellung modulo 2π: Es seien

$$z = r_1 e^{i\varphi_1} = r_2 e^{i\varphi_2}$$

zwei Polardarstellungen von z, wobei $r_1, r_2 > 0$ und $\varphi_1, \varphi_2 \in \mathbb{R}$. Dann folgt die Eindeutigkeit des "Radius"

$$r_1 = r_1 |e^{i\varphi_1}| = |r_1 e^{i\varphi_1}| = |z| = |r_2 e^{i\varphi_2}| = r_2 |e^{i\varphi_2}| = r_2$$

und hieraus

$$1 = \frac{r_1 e^{i\varphi_1}}{r_2 e^{i\varphi_2}} = \frac{e^{i\varphi_1}}{e^{i\varphi_2}} = e^{i(\varphi_1 - \varphi_2)}.$$

Wir schließen $i(\varphi_1 - \varphi_2) \in 2\pi i\mathbb{Z}$, also $\varphi_1 - \varphi_2 \in 2\pi\mathbb{Z}$. Damit ist auch die Eindeutigkeit des "Winkels" in der Polardarstellung modulo 2π rigoros gezeigt, ohne Rückgriff auf die geometrische Anschauung.

$\square$

Obwohl die eben vorgestellte Theorie die geometrische Anschauung nicht verwendet, ist es trotzdem wichtig, ein möglichst anschauliches Bild von der komplexen Exponentialfunktion zu bekommen. Zum Beispiel gilt: Variiert man den Realteil einer komplexen Zahl z bei festgehaltenem Imaginärteil, so läuft e^z auf einem von 0 ausgehenden Strahl. Variiert man dagegen den Imaginärteil bei festgehaltenem Realteil, so läuft e^z auf einer Kreislinie mit Mittelpunkt 0. Variiert man dagegen z auf einer allgemeinen Geraden, weder parallel zur reellen Achse noch zur imaginären Achse, so läuft e^z auf einer Spiralkurve um den Nullpunkt.

Weitere Ableitungen trigonometrischer Funktionen und Arcusfunktionen *Siehe auch die Folien unter folgenden Adressen:*
`http://www.mathematik.uni-muenchen.de/~merkl/ws12/ana1/trigon.ps` oder
`http://www.mathematik.uni-muenchen.de/~merkl/ws12/ana1/trigon.pdf`

a) **Ableitung des Tangens**

$$\frac{d}{dx}\tan x = \frac{d}{dx}\frac{\sin x}{\cos x} = \frac{(\frac{d}{dx}\sin x)\cos x - \sin x(\frac{d}{dx}\cos x)}{\cos^2 x} = \frac{\cos^2 x + \sin^2 x}{\cos^2 x} = \frac{1}{\cos^2 x}$$

Hier haben wir die Quotientenregel der Differentialrechnung verwendet.

b) **Ableitung des Cotangens**

$$\frac{d}{dx}\cot x = \frac{d}{dx}\frac{\cos x}{\sin x} = \frac{(\frac{d}{dx}\cos x)\sin x - \cos x(\frac{d}{dx}\sin x)}{\sin^2 x} = -\frac{1}{\sin^2 x}$$

c) **Ableitung des Arcussinus**
Setzen wir $y = \sin x$, $x = \arcsin y$, $x \in \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$, dann gilt:

$$\frac{dy}{dx} = \cos x = \sqrt{1 - \sin^2 x} = \sqrt{1 - y^2},$$

da $\cos x$ in unserem Intervall positiv ist. Damit folgt:

$$\frac{d}{dy}\arcsin y = \frac{dx}{dy} = \frac{1}{\sqrt{1 - y^2}}$$

Hier geht unser Wissen über die Ableitung der Umkehrfunktion ein. Man beachte, dass der Arcussinus am Rand des Intervalls zwar definiert, aber nicht differenzierbar ist.

d) **Ableitung des Arcuscosinus**
Setzen wir $y = \cos x, x = \arccos y, x \in]0, \pi[$, dann gilt:

$$\frac{dy}{dx} = -\sin x = -\sqrt{1 - \cos^2 x} = -\sqrt{1 - y^2},$$

da $\sin x$ in unserem Intervall positiv ist. Damit folgt:

$$\frac{d}{dy} \arccos y = \frac{dx}{dy} = -\frac{1}{\sqrt{1 - y^2}}$$

e) **Ableitung des Arcustangens**
Setzen wir $y = \tan x$, $x = \arctan y$, $|x| < \frac{\pi}{2}$, dann gilt:

$$\frac{dy}{dx} = \frac{1}{\cos^2 x} = \frac{\sin^2 x + \cos^2 x}{\cos^2 x} = \tan^2 x + 1 = y^2 + 1$$

Also:

$$\frac{d}{dy} \arctan y = \frac{dx}{dy} = \frac{1}{y^2 + 1}$$

f) **Ableitung des Arcuscotangens**
Setzen wir $y = \cot x$, $x = \operatorname{arccot} y$, $x \in]0, \pi[$, dann gilt:

$$\frac{dy}{dx} = -\frac{1}{\sin^2 x} = -\frac{\sin^2 x + \cos^2 x}{\sin^2 x} = -(1 + \cot^2 x) = -(1 + y^2)$$

Also:

$$\frac{d}{dy} \operatorname{arccot} y = \frac{dx}{dy} = -\frac{1}{1 + y^2}$$

Sinus und Cosinus im Komplexen: Hyperbelfunktionen

Die meisten behandelten Aussagen über *Sinus* und *Cosinus* lassen sich auf komplexe Argumente erweitern.

Definition 4.13 (Sinus und Cosinus in $\mathbb{C}$) *Für $z \in \mathbb{C}$ definieren wir*

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

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

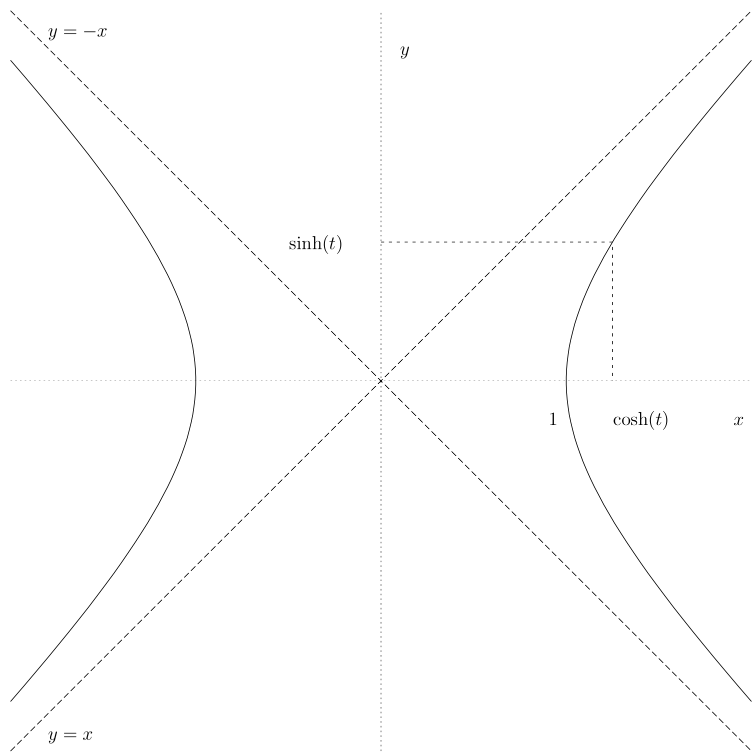
In diesem Abschnitt studieren wir diese Funktionen auf der imaginären Achse.
Wir setzen für $t \in \mathbb{C}$:

Definition 4.14 (Sinus hyperbolicus und Cosinus hyperbolicus)

$$\cosh t := \cos(it) = \frac{e^{i^2 t} + e^{-i^2 t}}{2} = \frac{e^t + e^{-t}}{2}$$

$$\sinh t := \frac{1}{i}\sin(it) = \frac{e^{i^2 t} - e^{-i^2 t}}{2i^2} = \frac{e^t - e^{-t}}{2}$$

Diese Funktionen werden *hyperbolischer Sinus* und *hyperbolischer Cosinus* genannt. Der Grund dafür liegt darin, daß $(\cosh, \sinh)$ einen Zweig der Hyperbel $\{(x,y) \in \mathbb{R}^2 | x^2 - y^2 = 1\}$ parametrisieren:



In der Tat gilt:

$$\cosh^2 t - \sinh^2 t = (\cosh t + \sinh t)(\cosh t - \sinh t) = e^t e^{-t} = 1$$

Analogie:

$$\boxed{\begin{gathered}\cosh^2 t - \sinh^2 t = 1 \\ \text{versus} \\ \cos^2 t + \sin^2 t = 1\end{gathered}}$$

Die Hyperbelfunktionen spielen also eine ähnliche Rolle zur Parametrisierung eines Hyperbelzweiges wie (cos,sin) zur Parametrisierung des Einheitskreises.

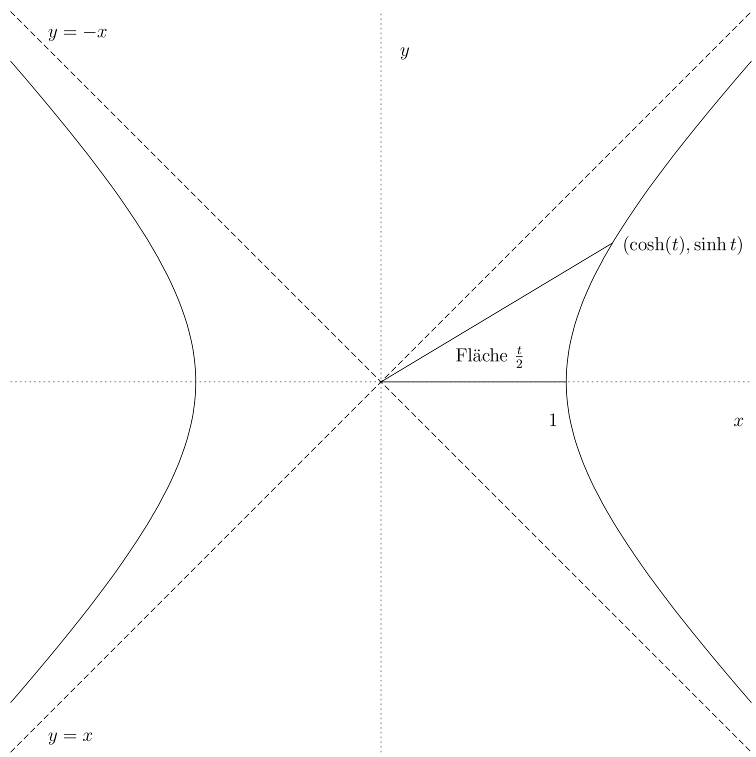
Reihe und Ableitung von sinh, cosh:
Aus den Reihen $e^t = \sum_{n=0}^{\infty} \frac{x^n}{n!}$ und $e^{-t} = \sum_{n=0}^{\infty} \frac{(-1)^n x^n}{n!}$ folgt:

$$\boxed{\begin{aligned} \cosh t &= \sum_{n=0}^{\infty} \frac{1+(-1)^n}{2} \frac{x^n}{n!} = \sum_{k=0}^{\infty} \frac{x^{2k}}{(2k)!} \\ \sinh t &= \sum_{n=0}^{\infty} \frac{1-(-1)^n}{2} \frac{x^n}{n!} = \sum_{k=0}^{\infty} \frac{x^{2k+1}}{(2k+1)!} \end{aligned}}$$

und aus $\frac{d}{dt}e^t = e^t$, $\frac{d}{dt}e^{-t} = -e^{-t}$ folgt:

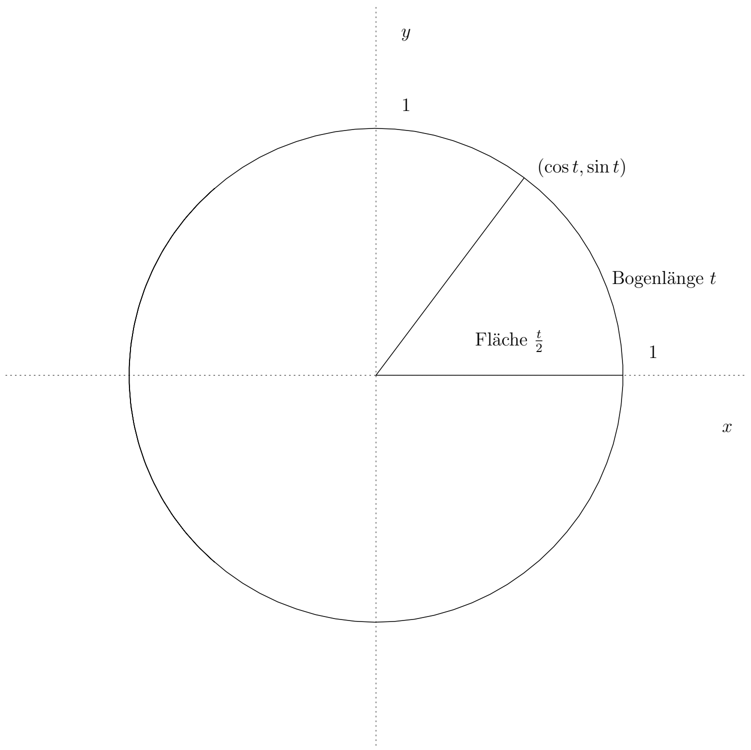
$$\frac{d}{dt}\cosh t = \sinh t$$
$$\frac{d}{dt}\sinh t = \cosh t$$

Interpretation von tanh, sinh, cosh an der Hyperbel:



Die Fläche des Hyperbelsegments mit den Ecken $(0,0)$, $(1,0)$ und $(\cosh(t), \sinh t)$ beträgt $t/2$. Die Gerade durch den Ursprung und den Punkt $(\cosh(t), \sinh t)$ hat die Steigung $\tanh t = 1/\coth t$. (Die Begründung der Flächenformel können wir erst später mit Hilfe der Integralrechnung geben.) Das Argument t der Hyperbelfunktionen hat *keine* Interpretation als Bogenlänge, sondern *nur als Fläche* eines Hyperbelsegments.

Analogie zum Kreis:



Die Fläche des Kreissegments mit den Ecken $(0,0)$, $(1,0)$ und $(\cos(t), \sin t)$ beträgt $t/2$. Das ist die Hälfte der Bogenlänge t des Kreisbogens zwischen $(1,0)$ und $(\cos(t), \sin t)$. Die Gerade durch den Ursprung und den Punkt $(\cos(t), \sin t)$ hat die Steigung $\tan t = 1/\cot t$.

Areafunktionen (Umkehrungen der Hyperbelfunktionen)

Aufgrund der fehlenden Interpretation des Arguments t von $(\cosh t, \sinh t)$ als Bogenlänge heißen die Umkehrfunktionen der Hyperbelfunktionen nicht Arcusfunktionen (arcus = Bogen), sondern Areafunktionen (area = Fläche).
Sie können mit Hilfe des Logarithmus ausgedrückt werden und erfüllen ähnliche Ableitungsregeln wie die Arcusfunktionen:

Wir setzen $y = \sinh x$, $x = \text{arsinh}\ y$ mit $x \in \mathbb{R}$

$$\frac{dy}{dx} = \cosh x = \sqrt{1 + \sinh^2 x} = \sqrt{1 + y^2}$$

Man beachte, daß $\cosh x > 0$, und $\cosh^2 x - \sinh^2 x = 1$.
Also:

$$\frac{d}{dy}\text{arsinh} y = \frac{dx}{dy} = \frac{1}{\sqrt{1+y^2}} \quad \text{(in Analogie zu } \frac{d}{dy} \arcsin y = \frac{1}{\sqrt{1-y^2}}\text{)}$$

Darstellung der Umkehrfunktion mit Hilfe des Logarithmus:

$$\begin{aligned} & y = \sinh x \\ \Rightarrow \quad & y + \sqrt{1+y^2} = \sinh x + \cosh x = e^x \\ \Rightarrow \quad & x = \log(y + \sqrt{1+y^2}) \end{aligned}$$

Das bedeutet:

$$\text{arsinh}\ y = \log(y + \sqrt{1+y^2})$$

Die Rechnung für $\text{arcosh}\ y$ geht ganz analog.

4.3 Varianten von Stetigkeit und Differenzierbarkeit: Einseitig stetige und differenzierbare Funktionen

Definition 4.15 Sei $f : U \to \mathbb{C}$, $U \subseteq \mathbb{R}$, und $x \in U$.
Es sei x ein Berührpunkt von $U \cap \,]x, +\infty[$. f heißt *rechtsseitig stetig* in x, wenn die Einschränkung von f auf $U \cap [x, +\infty[$ stetig in x ist.
In diesem Fall schreiben wir

$$\lim_{y \downarrow x} f(y) = f(x)$$

"$y \downarrow x$" steht also für den Grenzübergang $y \to x$ mit der Einschränkung $y > x$.

f heißt *rechtsseitig differenzierbar* in x, wenn $\lim\limits_{y \downarrow x} \dfrac{f(y) - f(x)}{y - x}$ existiert.
Der Grenzwert heißt dann *rechtsseitige Ableitung* von f in x.

Analog werden in x *linksseitig stetige* Funktionen und *linksseitig differenzierbare* Funktionen mit der Einschränkung auf $U \cap \,]-\infty, x]$ definiert;
ebenso der linksseitige Grenzwert $\lim\limits_{y \uparrow x} f(y) = \lim\limits_{\substack{y \to x \\ y < x}} f(y)$.

Beispiel 4.16 $x \mapsto |x|$, $x \in \mathbb{R}$, ist für $x = 0$ links- und rechtsseitig differenzierbar mit links- und rechtsseitiger Ableitung -1 bzw. $+1$.

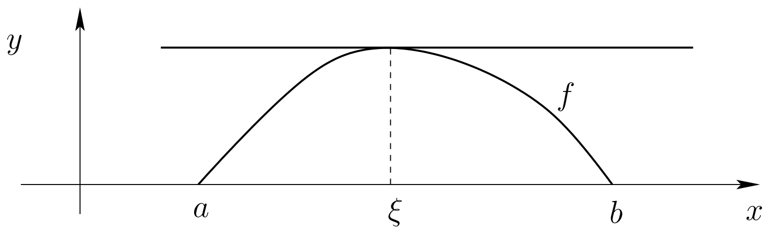
4.4 Der Mittelwertsatz der Differentialrechnung

4.4.1 Der Satz von Rolle und der einfache Mittelwertsatz

Im Folgenden seien $a, b \in \mathbb{R}$ mit $a < b$.

Satz 4.17 (Satz von Rolle) Es sei $f : [a, b] \to \mathbb{R}$ stetig und in $]a, b[$ differenzierbar. Es gelte $f(a) = f(b) = 0$. Dann gibt es $\xi \in]a, b[$ mit $f'(\xi) = 0$.

Illustration:



Beweis: Weil f stetig und $[a, b]$ kompakt ist, nimmt f ein Maximum und ein Minimum an. Wir unterscheiden 3 Fälle:

1.Fall: $\max f > 0$. Dann gibt es $\xi \in]a, b[$ mit $f(\xi) = \max f$.
Es folgt für $a \leq x < \xi : f(x) \leq f(\xi)$, also

$$\frac{f(x) - f(\xi)}{x - \xi} \geq 0.$$

Ebenso folgt für $\xi < x \leq b$:

$$\frac{f(x) - f(\xi)}{x - \xi} \leq 0.$$

Wir schließen:

$$0 \leq \lim_{x \uparrow \xi} \frac{f(x) - f(\xi)}{x - \xi} = f'(\xi) = \lim_{x \downarrow \xi} \frac{f(x) - f(\xi)}{x - \xi} \leq 0.$$

also $f'(\xi) = 0$.

2.Fall: $\min f < 0$. Dann gibt es $\xi \in]a, b[$ mit $f(\xi) = \min f$,
und analog zum 1.Fall erhalten wir $f'(\xi) = 0$.

3.Fall: $\min f = \max f = 0$. Dann ist $f = 0$, und wir erhalten
$f'(\xi) = 0$ für alle $\xi \in]a, b[$.

$\square$

Folgerung:

Satz 4.18 (Mittelwertsatz der Differentialrechnung) Es sei $f : [a, b] \to \mathbb{R}$ stetig und in $]a, b[$ differenzierbar. Dann gibt es $\xi \in]a, b[$ mit

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

Beweis: Wir betrachten die Funktion $g : [a, b] \to \mathbb{R}$,

$$g(x) = f(x) - \frac{f(b) - f(a)}{b - a}(x - a) - f(a).$$

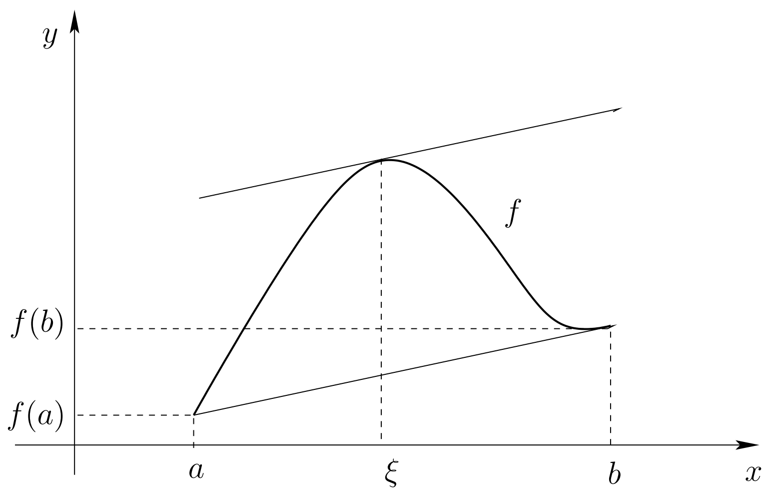
Dann gilt $g(a) = 0 = g(b)$, und nach dem Satz von Rolle gibt es ein $\xi \in]a, b[$ mit $g'(\xi) = 0$ Das bedeutet:

$$f'(\xi) - \frac{f(b) - f(a)}{b - a} = 0,$$

also die Behauptung.

□

Illustration:



Der Mittelwertsatz dient dazu, aus dem Verhalten der Ableitung auf das Verhalten der Funktion zurückzuschließen. Hierzu einige Beispiele.

Beispiel 4.19 Wenn f' nur positive bzw. negative Werte auf einem Intervall I annimmt, so ist f dort streng monoton steigend bzw. streng monoton fallend.
In der Tat gilt für $a, b \in I,\ a < b$:
Es gibt $\xi \in I$ mit $f(b) - f(a) = f'(\xi)(b - a)$, also hat $f(b) - f(a)$ das gleiche Vorzeichen wie $f'(\xi)$.

Beispiel 4.20 *Ist $f : \mathbb{R} \to \mathbb{R}$ differenzierbar mit beschränkter Ableitung, so ist f global Lipschitz-stetig.*
Beweis: Es gelte $|f'| \leq L$, $L \in \mathbb{R}$. Für gegebene $x, y \in \mathbb{R}$ finden wir mit dem Mittelwertsatz ein ξ zwischen x und y mit $f(y) - f(x) = f'(\xi)(y - x)$. Es folgt:

$$|f(y) - f(x)| = |f'(\xi)||y - x| \leq L|y - x|,$$

also ist f global Lipschitz-stetig.

4.4.2 Anwendung auf Differentialgleichungen

Die Differentialgleichung

$$f'(x) = 0$$

für eine unbekannte Funktion $f : I \to \mathbb{C}$
auf einem Intervall $I \subseteq \mathbb{R}$ hat nur die Lösungen:

$$f = \text{konstant}.$$

Beweis: Indem wir Real- und Imaginärteil von f einzeln betrachten, erhalten wir: $(\mathrm{Re}\, f)' = 0$ und $(\mathrm{Im}\, f)' = 0$. Es genügt also, reellwertige f zu untersuchen.
Hier erhalten wir für alle $a < b$ in I: Es gibt $\xi \in]a, b[$ mit $f(b) - f(a) = f'(\xi)(b - a) = 0$, also $f(b) = f(a)$.
Also ist f konstant.

$\square$

Folgerung 1: Es sei $\lambda \in \mathbb{C}$.

Die Differentialgleichung
$y' = \lambda y$, mit $y : I \to \mathbb{C}$
auf I hat nur die Lösungen $y(x) = ce^{\lambda x}$, $c \in \mathbb{C}$ konstant

Beweis: Es gelte $y' = \lambda y$. Wir setzen $f(x) = e^{-\lambda x} y(x)$. Dann folgt:

$$f'(x) = -\lambda e^{-\lambda x} y(x) + e^{-\lambda x} y'(x) = 0,$$

also $f = c =$ konstant für ein $c \in \mathbb{C}$. Das bedeutet $y(x) = ce^{\lambda x}$.

$\square$

Folgerung 2:

Die Schwingungsgleichung
$y'' + y = 0$, mit $y : I \to \mathbb{R}$
hat den 2-dimensionalen $\mathbb{R}$-Vektorraum mit Basis $\{\sin, \cos\}$ als Lösungsraum.

Anders gesagt: Jede reelle Lösung der Schwingungsgleichung auf I läßt sich eindeutig als $y(x) = \alpha\cos x + \beta\sin x$ schreiben.
Beweis: Es gelte $y'' + y = 0$. Wir setzen $z = y'$, $f = y + iz$ und erhalten

$$f' = y' + iz' = z + iy'' = z - iy = -i(y + iz) = -if.$$

Es folgt $f(x) = ce^{-ix}$ für ein $c \in \mathbb{C},\ c = \alpha + i\beta;\ \alpha, \beta \in \mathbb{R}$ also

$$y(x) = \mathrm{Re} f(x) = \mathrm{Re}[(\alpha + i\beta)(\cos x - i\sin x)] = \alpha\cos x + \beta\sin x$$

α und β sind eindeutig bestimmt, weil sin und cos linear unabhängig sind.

□

Folgerung 3:

Die Differentialgleichung

$$y'' = 0$$

auf einem Intervall I hat die allgemeine Lösung
$y(x) = \alpha x + \beta$, mit $\alpha,\ \beta \in \mathbb{C}$

Beweis: Zunächst folgt $y' = \alpha =$ konstant für ein $\alpha \in \mathbb{C}$.
Wir schließen $\frac{d}{dx}(y(x) - \alpha x) = y'(x) - \alpha = 0$, also $y(x) - \alpha x = \beta$ für ein $\beta \in \mathbb{C}$.

□

4.4.3 Der verallgemeinerte Mittelwertsatz

Wieder seien $a, b \in \mathbb{R}$ mit $a < b$.

Satz 4.21 (verallgemeinerter Mittelwertsatz der Differentialrechnung)
Es seien $f, g : [a, b] \to \mathbb{R}$ zwei stetige und auf $]a, b[$ differenzierbare Funktionen. Dann gibt es $\xi \in\ ]a, b[$, so daß gilt:

$$[f(b) - f(a)]g'(\xi) = [g(b) - g(a)]f'(\xi).$$

Beweis: Wir setzen für $x \in [a, b]$:

$$h(x) = [f(b) - f(a)][g(x) - g(a)] - [g(b) - g(a)][f(x) - f(a)]$$

Dann gilt $h(a) = 0 = h(b)$, und h ist differenzierbar in $]a, b[$ und stetig in $[a, b]$. Aus dem Satz von Rolle folgt für ein $\xi \in\ ]a, b[$ die Behauptung:

$$0 = h'(\xi) = [f(b) - f(a)]g'(\xi) - [g(b) - g(a)]f'(\xi)$$

□

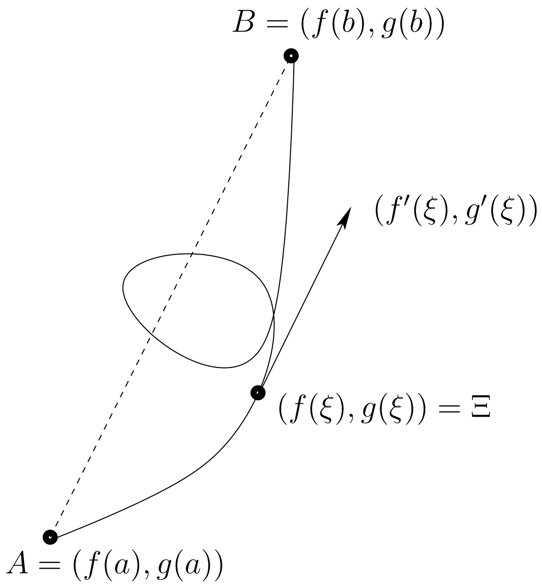
Mit Hilfe einer 2×2-Determinante kann man die Formel im verallgemeinerten Mittelwertsatz auch so schreiben:

$$\begin{vmatrix} f(b)-f(a) & f'(\xi) \\ g(b)-g(a) & g'(\xi) \end{vmatrix} = 0$$

Das Gleiche in der Sprache der Linearen Algebra gesagt: Für die Zwischenstelle ξ sind die Vektoren $\begin{pmatrix} f(b) \\ g(b) \end{pmatrix} - \begin{pmatrix} f(a) \\ g(a) \end{pmatrix}$ und $\begin{pmatrix} f'(\xi) \\ g'(\xi) \end{pmatrix}$ linear abhängig. Für $g(b) \neq g(a)$ und $g'(\xi) \neq 0$ kann man den verallgemeinerten Mittelwertsatz auch in folgender Form schreiben:

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

Interpretation: Betrachten wir die durch $x \mapsto (f(x), g(x))$ parametrisierte Kurve in der Ebene:



Dann gibt es einen Punkt Ξ auf der Kurve, an dem der Geschwindigkeitsvektor $(f'(\xi), g'(\xi))$ parallel zur Strecke $\overline{AB}$ ist.
Wir können Ξ als einen Punkt wählen, an dem das Dreieck $A,\ B,\ \Xi$ maximalen Flächeninhalt hat.
In der Tat ist dieser Flächeninhalt (bis auf das Vorzeichen) gleich:

$$\frac{1}{2}\begin{vmatrix} f(b)-f(a) & f(\xi)-f(a) \\ g(b)-g(a) & g(\xi)-g(a) \end{vmatrix} = \frac{1}{2}h(\xi)$$

Der verallgemeinerte Mittelwertsatz ist in Anwendungssituationen nützlich, in denen $f'(\xi)$ für $\xi \in [a,b]$ über viele Größenordnungen variiert. Der einfache Mittelwertsatz ist dort jedoch fast nutzlos. Man kann dann Zuwächse $f(b)-f(a)$ einer Funktion mit den Zuwächsen

$g(b) - g(a)$ einer "einfacheren" Funktion g vergleichen, deren Ableitung g' ähnlich stark wie f' variiert.

Folgerung: (Regel von l'Hôpital) Es sei U eine offene Umgebung von $x \in \mathbb{R}$, $f, g : U \to \mathbb{R}$ stetig und in $U \setminus \{x\}$ differenzierbar mit $f(x) = g(x) = 0$, aber $g(y) \neq 0$ und $g'(y) \neq 0$ für $y \in U \setminus \{x\}$.
Wenn $\lim\limits_{y \to x} \frac{f'(y)}{g'(y)}$ existiert, so auch $\lim\limits_{y \to x} \frac{f(x)}{g(x)}$, und es gilt:

$$\lim_{y \to x} \frac{f'(y)}{g'(y)} = \lim_{y \to x} \frac{f(y)}{g(y)}$$

Bemerkung: Ähnliche Varianten gibt es für $f(y) \overset{y \to x}{\longrightarrow} \pm\infty$ und $g(y) \overset{y \to x}{\longrightarrow} \pm\infty$ oder auch für $y \to \pm\infty$ statt $y \to x$. Diese Varianten beweisen wir hier nicht.

Beweis der Regel von l'Hôpital: Wir dürfen annehmen, daß U ein Intervall ist.
Nach dem verallgemeinerten Mittelwertsatz können wir ein $\xi(y)$ zwischen y und x wählen, so daß

$$\frac{f(y)}{g(y)} = \frac{f(y) - f(x)}{g(y) - g(x)} = \frac{f'(\xi(y))}{g'(\xi(y))} \overset{y \to x}{\longrightarrow} \lim_{\xi \to x} \frac{f'(\xi)}{g'(\xi)}$$

denn $\xi(y) \overset{y \to x}{\longrightarrow} x$.

$\square$

4.4.4 Konvexe Funktionen

Definition 4.22 Sei $I \subseteq \mathbb{R}$ ein Intervall und $f : I \to \mathbb{R}$.
f heißt *konvex* wenn für alle $x,\ y \in I$ und $\alpha,\ \beta \in [0, 1]$ mit $\alpha + \beta = 1$ gilt:

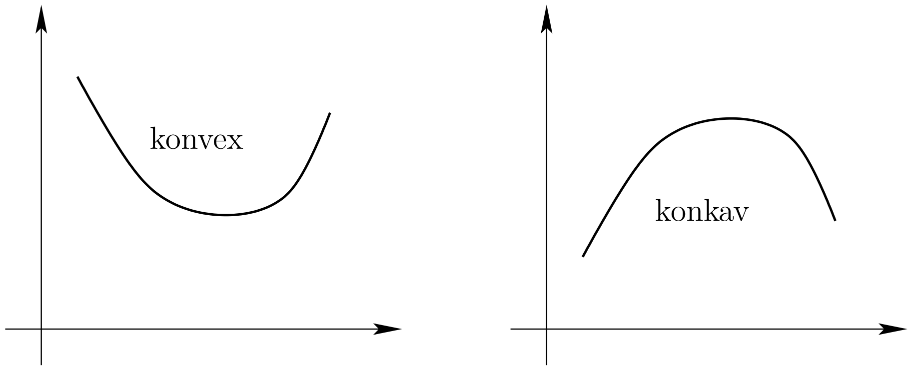
$$f(\alpha x + \beta y) \leq \alpha f(x) + \beta f(y).$$

Gilt hingegen:

$$f(\alpha x + \beta y) \geq \alpha f(x) + \beta f(y),$$

so heißt f *konkav.*

Bemerkung: Insbesondere in der Schulmathematik werden manchmal die Begriffe "konvex" und "konkav" mit vertauschten Bedeutungen gebraucht. Die hier gewählte Konvention hat sich allerdings weitgehend durchgesetzt.



Nun nehmen wir an, daß $f : I \to \mathbb{R}$ zweimal differenzierbar sei.

Satz 4.23 Gilt $f'' \geq 0$ auf I, so ist f konvex.
Gilt hingegen $f'' \leq 0$ auf I, so ist f konkav.

Beweis: Wir beweisen den Fall $f'' \geq 0$, der andere Fall folgt analog.
Es seien $x \leq y$ in I und $0 \leq \alpha \leq 1$, $\alpha + \beta = 1$.
Wir setzen $z = \alpha x + \beta y,\ z \in [x, y]$.
Im Fall $z = x$ oder $z = y$ ist nichts zu zeigen.
Andernfalls gilt $x < z < y$. Nach dem Mittelwertsatz gibt es dann $\xi_1,\ \xi_2$ mit

$$x < \xi_1 < z < \xi_2 < y \text{ und } f(z) - f(x) = f'(\xi_1)(z - x) \text{ und } f(y) - f(z) = f'(\xi_2)(y - z).$$

Weiter gilt $f'(\xi_1) \leq f'(\xi_2)$, denn wegen $f'' \geq 0$ ist f' monoton steigend.
Es folgt die Behauptung:

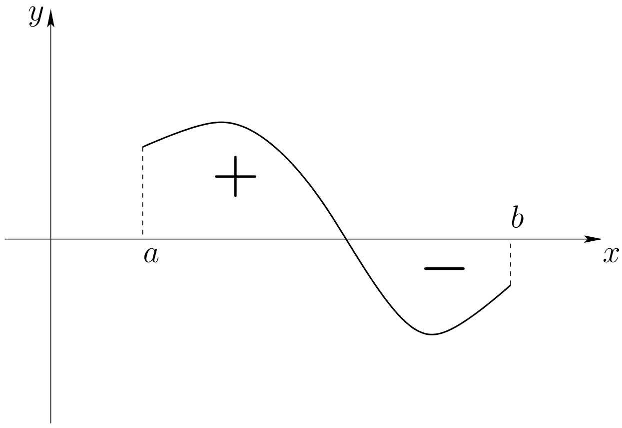
$$\begin{aligned}
\alpha f(x) + \beta f(y) - f(z) &= \alpha f(x) + \beta f(y) - (\alpha + \beta) f(z) \\
&= \beta[f(y) - f(z)] - \alpha[f(z) - f(x)] \\
&= \beta f'(\xi_2)(y - z) - \alpha f'(\xi_1)(z - x) \\
&\geq \beta f'(\xi_1)(y - z) - \alpha f'(\xi_1)(z - x) \\
&= f'(\xi_1)(\alpha x + \beta y - (\alpha + \beta) z) \\
&= 0.
\end{aligned}$$

□

5 Integralrechnung

5.1 Das Riemann-Integral

Das Riemann-Integral misst die Fläche zwischen einem Funktionsgraphen und der x-Achse; Teile unter der x-Achse zählen negativ.



Wir definieren das Integral zunächst für Treppenfunktionen, also stückweise konstante Funktionen.

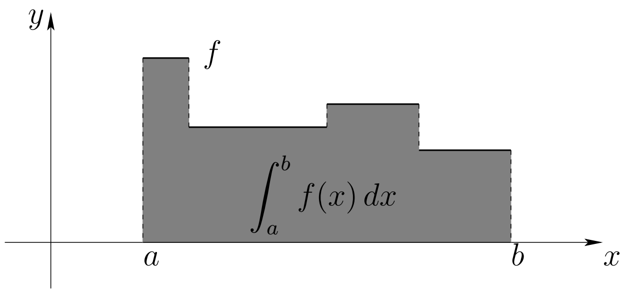
Definition 5.1 Eine Funktion $f : [a, b] \to \mathbb{R}$ $(a \leq b)$ heißt *Treppenfunktion*, wenn es eine Zerlegung $a = \xi_0 \leq \xi_1 \leq \cdots \leq \xi_n = b$ von $[a, b]$ gibt, so daß f auf allen Intervallen $[\xi_j, \xi_{j+1}[\,,\ j = 0, \ldots, n-1$, konstant ist. $\mathcal{T}[a, b]$ sei die Menge der Treppenfunktionen auf $[a, b]$. Wir definieren das Integral einer Treppenfunktion f:

$$\int_a^b f(x)\,dx := \sum_{j=0}^{n-1} f(\xi_j)(\xi_{j+1} - \xi_j).$$

Bemerkungen:

1) Das Integral hängt nicht von der Wahl der Zerlegung ab, wenn f auf allen $[\xi_j, \xi_{j+1}[$ konstant ist.

2) Interpretation als Fläche:



3) Der Raum $\mathcal{T}[a, b]$ bildet einen Untervektorraum des $\mathbb{R}$-Vektorraums $\mathbb{R}^{[a,b]}$ aller Funktionen $f : [a, b] \to \mathbb{R}$.

Lemma 5.2 (Linearität des Integrals für Treppenfunktionen) *Das Integral für Treppenfunktionen ist eine lineare Funktion $\mathcal{T}[a,b] \to \mathbb{R}$. Das bedeutet:*

a) Aus $f, g \in \mathcal{T}[a,b]$ folgt $f + g \in \mathcal{T}[a,b]$ und

$$\int_a^b [f(x) + g(x)]\, dx = \int_a^b f(x)\, dx + \int_a^b g(x)\, dx.$$

b) Aus $f \in \mathcal{T}[a,b]$ und $\alpha \in \mathbb{R}$ folgt $\alpha f \in \mathcal{T}[a,b]$ und

$$\int_a^b \alpha f(x)\, dx = \alpha \int_a^b f(x)\, dx.$$

Beweis:

a) Es sei $a = \xi_0 \leq \cdots \leq \xi_n = b$ eine so feine Zerlegung von $[a,b]$, daß sowohl f als auch g auf allen Teilintervallen $[\xi_j, \xi_{j+1}[$ konstant ist. Dann ist auch $f + g$ auf allen $[\xi_j, \xi_{j+1}[$ konstant, und es gilt:

$$\begin{aligned}\int_a^b [f(x) + g(x)]\, dx &= \sum_{j=0}^{n-1} [f(\xi_j) + g(\xi_j)]\,(\xi_{j+1} - \xi_j) \\ &= \sum_{j=0}^{n-1} f(\xi_j)(\xi_{j+1} - \xi_j) + \sum_{j=0}^{n-1} g(\xi_j)(\xi_{j+1} - \xi_j) \\ &= \int_a^b f(x)\, dx + \int_a^b g(x)\, dx.\end{aligned}$$

b) αf ist wieder stückweise konstant, also eine Treppenfunktion, und es gilt:

$$\int_a^b \alpha f(x)\, dx = \alpha \sum_{j=0}^{n-1} f(\xi_j)(\xi_{j+1} - \xi_j) = \alpha \int_a^b f(x)\, dx$$

□

Lemma 5.3 (Monotonie des Integrals für Treppenfunktionen) *Es seien $f, g \in \mathcal{T}[a,b]$ mit $f \leq g$. Dann gilt:*

$$\int_a^b f(x)\, dx \leq \int_a^b g(x)\, dx.$$

Beweis: Wir wählen eine so feine Zerlegung $a = \xi_0 \leq \xi_1 \leq \cdots \leq \xi_n = b$ von $[a,b]$, daß sowohl f als auch g auf allen Intervallen $[\xi_j, \xi_{j+1}[$ konstant ist. Dann folgt:

$$\begin{aligned}\int_a^b f(x)\, dx &= \sum_{j=0}^{n-1} f(\xi_j)(\xi_{j+1} - \xi_j) \\ &\leq \sum_{j=0}^{n-1} g(\xi_j)(\xi_{j+1} - \xi_j) = \int_a^b g(x)\, dx.\end{aligned}$$

$\square$

Wir erweitern nun den Integralbegriff auf eine größere Funktionenklasse:

fundamentale Definition!

Definition 5.4 (Riemann-Integral) Eine Funktion $f : [a,b] \to \mathbb{R}$ heißt *Riemann-integrierbar* (in Zeichen: $f \in \mathcal{R}[a,b]$), wenn es zu jedem $\varepsilon > 0$ Treppenfunktionen $g, h \in \mathcal{T}[a,b]$ mit $g \leq f \leq h$ und

$$\int_a^b h(x)\,dx - \int_a^b g(x)\,dx \leq \varepsilon$$

gibt. In diesem Fall definieren wir das *Riemann-Integral* von f:

$$\int_a^b f(x)\,dx := \sup_{\substack{g \in \mathcal{T}[a,b] \\ g \leq f}} \int_a^b g(x)\,dx = \inf_{\substack{h \in \mathcal{T}[a,b] \\ h \geq f}} \int_a^b h(x)\,dx. \tag{65}$$

Bemerkungen:

1. Zum Beweis der letzten Gleicheit (65):
   Wir setzen

$$A = \left\{ \int_a^b g(x)\,dx \,\middle|\, g \in \mathcal{T}[a,b],\, g \leq f \right\},$$
$$B = \left\{ \int_a^b h(x)\,dx \,\middle|\, h \in \mathcal{T}[a,b],\, h \geq f \right\}.$$

   Zu zeigen ist

$$\sup A = \inf B \in \mathbb{R}.$$

   Wegen der Riemann-Integrierbarkeit von f sind A und B nichtleer. Weiter gilt

$$\forall \alpha \in A\ \forall \beta \in B :\ \alpha \leq \beta. \tag{66}$$

   Ist nämlich $\alpha = \int_a^b g(x)\,dx \in A$ mit $g \in \mathcal{T}[a,b]$, $g \leq f$ und $\beta = \int_a^b h(x)\,dx \in B$ mit $h \in \mathcal{T}[a,b]$, $h \geq f$, so folgt $g \leq h$ und damit $\alpha \leq \beta$ wegen Lemma 5.3. Damit ist Aussage (66) bewiesen. Sie besagt mit anderen Worten: Jedes $\alpha \in A$ ist eine untere Schranke von B. Da B nichtleer und nach unten beschränkt ist, folgt $\inf B \in \mathbb{R}$. Da $\inf B$ die größte untere Schranke von B ist, folgt aus (66) weiter

$$\forall \alpha \in A :\ \alpha \leq \inf B.$$

   Das bedeutet: $\inf B$ ist eine obere Schranke von A. Da $\sup A$ die kleinste obere Schranke von A ist, erhalten wir

$$\sup A \leq \inf B.$$

Umgekehrt: Ist $\varepsilon > 0$ gegeben, so gibt es nach Definition der Riemann-Integrierbarkeit $\beta \in B$ und $\alpha \in A$ mit $\beta - \alpha < \varepsilon$. Wir schließen:

$$\sup A \geq \alpha > \beta - \varepsilon \geq \inf B - \varepsilon.$$

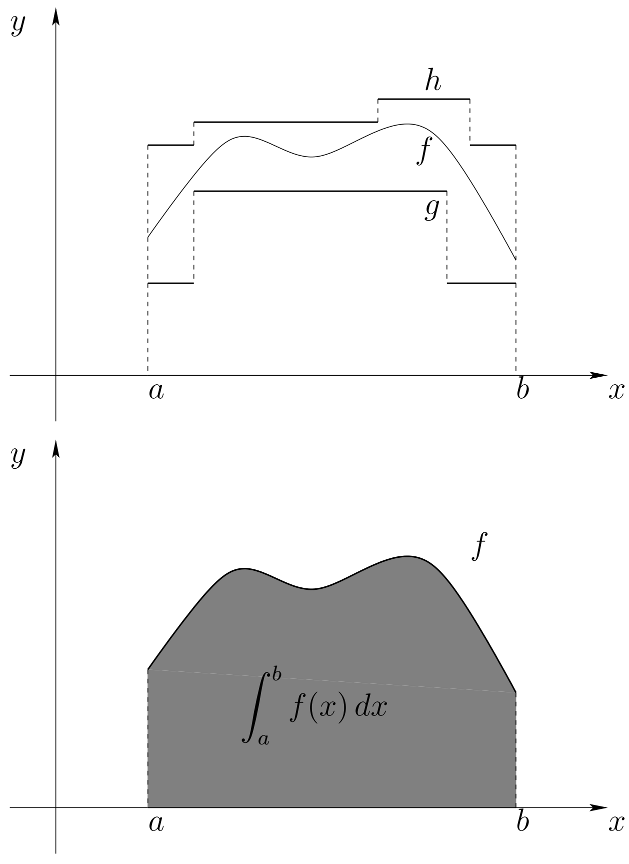
Weil dies für alle $\varepsilon > 0$ gilt, folgt

$$\sup A \geq \inf B,$$

also zusammen die Behauptung $\sup A = \inf B$.

2. Für Treppenfunktionen $f \in \mathcal{T}[a, b]$ stimmt das zuerst definierte Integral mit dem Riemann-Integral überein: Das Supremum (bzw. Infimum) in der Definition (65) des Riemann-Integrals ist in diesem Fall nämlich ein Maximum (bzw. Minimum) und wird für $g = f$ (bzw. $h = f$) angenommen. Wir dürfen also beide Integrale mit dem gleichen Symbol bezeichnen.

3. Interpretation des Integrals als Fläche:



4. Für $f \in \mathcal{R}[a, b]$ und eine Zerlegung

$$a = \xi_0 \leq x_0 \leq \xi_1 \leq x_1 \leq \xi_2 \leq \ldots \leq \xi_{n-1} \leq x_{n-1} \leq \xi_n = b$$

von $[a,b]$ mit zusätzlichen Zwischenpunkten $x_0, \dots, x_{n-1}$ wird

$$R = \sum_{j=1}^{n-0} f(x_j)(\xi_{j+1} - \xi_j)$$

eine *Riemannsumme* zu f genannt. Sind g, h Treppenfunktionen zur gleichen Zerlegung $\xi_0, \xi_1, \dots, \xi_n$ mit $g \le f \le h$, so gilt

$$\int_a^b g(x)\,dx \le R \le \int_a^b h(x)\,dx.$$

Die Riemannsche Integrationstheorie lässt sich alternativ auch mit Hilfe von Grenzwerten von Riemannsummen aufbauen.

Satz 5.5 Jede stetige Funktion $f : [a,b] \to \mathbb{R},\ a < b$, ist Riemann-integrierbar.

Beweis: Weil f stetig und $[a,b]$ kompakt ist, ist f gleichmäßig stetig. Es sei $\varepsilon > 0$. Wegen der gleichmäßigen Stetigkeit gibt es $\delta > 0$, so daß für alle $x, y \in [a,b]$ mit $|x - y| < \delta$ gilt:

$$|f(x) - f(y)| < \frac{\varepsilon}{2(b-a)}.$$

Wir wählen eine Zerlegung $a = \xi_0 \le \xi_1 \le \dots \le \xi_n = b$ von $[a,b]$, so daß $\xi_{j+1} - \xi_j < \delta$ für alle $j = 0, \dots, n-1$ gilt. Wir definieren die Treppenfunktion

$$\tilde{f} : [a,b] \to \mathbb{R}, \qquad \tilde{f}(x) = \begin{cases} f(\xi_j) & \text{für } \xi_j \le x < \xi_{j+1}, \\ f(b) & \text{für } x = b. \end{cases}$$

Nach Konstruktion gilt für alle $x \in [a,b]$:

$$\left| f(x) - \tilde{f}(x) \right| < \frac{\varepsilon}{2(b-a)}$$

Wir definieren Treppenfunktionen

$$g = \tilde{f} - \frac{\varepsilon}{2(b-a)}, \qquad h = \tilde{f} + \frac{\varepsilon}{2(b-a)}.$$

Dann gilt $g \le f \le h$ und

$$\int_a^b g(x)\,dx = \sum_{j=0}^{n-1} \left[\tilde{f}(\xi_j) - \frac{\varepsilon}{2(b-a)} \right] (\xi_{j+1} - \xi_j) = \int_a^b \tilde{f}(x)\,dx - \frac{\varepsilon}{2}$$

$$\int_a^b h(x)\,dx = \sum_{j=0}^{n-1} \left[\tilde{f}(\xi_j) + \frac{\varepsilon}{2(b-a)} \right] (\xi_{j+1} - \xi_j) = \int_a^b \tilde{f}(x)\,dx + \frac{\varepsilon}{2}$$

also

$$\int_a^b h(x)\,dx - \int_a^b g(x)\,dx = \varepsilon.$$

Es folgt: f ist Riemann-integrierbar.

$\square$

Satz 5.6 (Monotonie des Integrals) Es seien $a \leq b$. Sind $f, g : [a,b] \to \mathbb{R}$ Riemann-integrierbar mit $f \leq g$, so gilt

$$\int_a^b f(x)\,dx \leq \int_a^b g(x)\,dx.$$

Dies folgt unmittelbar daraus, daß für jede Treppenfunktion h mit $h \leq f$ auch $h \leq g$ gilt:

$$\int_a^b f(x)\,dx = \sup_{\substack{h \in \mathcal{T}[a,b] \\ h \leq f}} \int_a^b h(x)\,dx \leq \sup_{\substack{h \in \mathcal{T}[a,b] \\ h \leq g}} \int_a^b h(x)\,dx = \int_a^b g(x)\,dx$$

$\square$

Satz 5.7 Die Menge $\mathcal{R}[a,b]$ der Riemann-integrierbaren Funktionen $f : [a,b] \to \mathbb{R}$ bilden einen $\mathbb{R}$-Vektorraum. Das Riemann-Integral

$$\mathcal{R}[a,b] \to \mathbb{R}, \qquad f \mapsto \int_a^b f(x)\,dx$$

ist eine Linearform, d.h. eine lineare Abbildung mit Werten in $\mathbb{R}$.

Mit anderen Worten:

a) Es gilt $0 \in \mathcal{R}[a,b]$. Hierbei bezeichnet 0 die Nullfunktion auf $[a,b]$.

b) Aus $f, g \in \mathcal{R}[a,b]$ folgt $f + g \in \mathcal{R}[a,b]$ und

$$\int_a^b [f(x) + g(x)]\,dx = \int_a^b f(x)\,dx + \int_a^b g(x)\,dx.$$

c) Aus $f \in \mathcal{R}[a,b]$ und $\alpha \in \mathbb{R}$ folgt $\alpha f \in \mathcal{R}[a,b]$ und

$$\int_a^b \alpha f(x)\,dx = \alpha \int_a^b f(x)\,dx.$$

Beweis:

a) $0 \in \mathcal{T}[a,b] \subseteq \mathcal{R}[a,b]$

b) Es sei $\varepsilon > 0$. Wir wählen $\underline{f},\ \overline{f},\ \underline{g},\ \overline{g} \in \mathcal{T}[a,b]$ mit $\underline{f} \leq f \leq \overline{f},\ \underline{g} \leq g \leq \overline{g}$ und

$$\int_a^b \overline{f}(x)\,dx - \int_a^b \underline{f}(x)\,dx \leq \frac{\varepsilon}{2} \quad \text{und} \quad \int_a^b \overline{g}(x)\,dx - \int_a^b \underline{g}(x)\,dx \leq \frac{\varepsilon}{2}.$$

Dann gilt:

$$\begin{aligned}
&\int_a^b \left[\overline{f}(x)+\overline{g}(x)\right]\,dx-\int_a^b \left[\underline{f}(x)+\underline{g}(x)\right]\,dx\\
&=\left[\int_a^b \overline{f}(x)\,dx-\int_a^b \underline{f}(x)\,dx\right]+\left[\int_a^b \overline{g}(x)\,dx-\int_a^b \underline{g}(x)\,dx\right]\\
&\le \frac{\varepsilon}{2}+\frac{\varepsilon}{2}=\varepsilon
\end{aligned} \tag{67}$$

Also ist $f+g\in\mathcal{R}[a,b]$, und es gilt

$$\begin{array}{ccccc}
\displaystyle\int_a^b \left[\underline{f}(x)+\underline{g}(x)\right]\,dx & \le & \displaystyle\int_a^b \left[f(x)+g(x)\right]\,dx & \le & \displaystyle\int_a^b \left[\overline{f}(x)+\overline{g}(x)\right]\,dx\\
\| & & & & \|\\
\displaystyle\int_a^b \underline{f}(x)\,dx+\int_a^b \underline{g}(x)\,dx & \le & \displaystyle\int_a^b f(x)\,dx+\int_a^b g(x)\,dx & \le & \displaystyle\int_a^b \overline{f}(x)\,dx+\int_a^b \overline{g}(x)\,dx
\end{array}$$

Hierbei stimmen die linken Seiten in der ersten und zweiten Zeile stimmen überein; ebenso die beiden rechten Seiten. Zusammen mit obiger Feststellung (67) folgt:

$$\left|\int_a^b \left[f(x)+g(x)\right]\,dx-\left[\int_a^b f(x)\,dx+\int_a^b g(x)\,dx\right]\right|\le\varepsilon,$$

also die Behauptung, da $\varepsilon>0$ beliebig war.

c) Es sei $\varepsilon>0$. Wir wählen $\delta>0$ so klein, daß $|\alpha|\,\delta<\varepsilon$, und $\underline{f},\ \overline{f}\in\mathcal{T}[a,b]$ mit $\underline{f}\le f\le\overline{f}$ und $\int_a^b \overline{f}\,dx-\int_a^b \underline{f}\,dx<\delta$. Dann folgt

$$\begin{cases}\alpha\underline{f}\le\alpha f\le\alpha\overline{f} & \text{für } \alpha\ge 0,\\ \alpha\overline{f}\le\alpha f\le\alpha\underline{f} & \text{für } \alpha<0,\end{cases}$$

sowie $\alpha\overline{f},\ \alpha\underline{f}\in\mathcal{T}[a,b]$ und

$$\left|\int_a^b \alpha\overline{f}(x)\,dx-\int_a^b \alpha\underline{f}(x)\,dx\right|=|\alpha|\left|\int_a^b \overline{f}(x)\,dx-\int_a^b \underline{f}(x)\,dx\right|\le|\alpha|\,\delta<\varepsilon,$$

also ist $f\in\mathcal{R}[a,b]$.

Es folgt im Fall $\alpha\ge 0$:

$$\begin{array}{ccccccc}
\displaystyle\int_a^b \alpha\underline{f}(x)\,dx & \le & \displaystyle\int_a^b \alpha f(x)\,dx & \le & \displaystyle\int_a^b \alpha\overline{f}(x)\,dx & \le & \displaystyle\int_a^b \alpha\underline{f}(x)\,dx+\varepsilon\\
\| & & & & \| & &\\
\displaystyle\alpha\int_a^b \underline{f}(x)\,dx & \le & \displaystyle\alpha\int_a^b f(x)\,dx & \le & \displaystyle\alpha\int_a^b \overline{f}(x)\,dx & &
\end{array}$$

Es gilt also
$$\left|\int_a^b \alpha f(x)\,dx - \alpha\int_a^b f(x)\,dx\right| < \varepsilon.$$
Hieraus folgt die Behauptung, da $\varepsilon > 0$ beliebig ist.
Der Fall $\alpha < 0$ wird analog behandelt.

□

Satz 5.8 Es sei $a < b < c$ und $f : [a,c] \to \mathbb{R}$. Die Funktion f ist genau dann Riemann-integrierbar, wenn die Einschränkungen auf $[a,b]$ und auf $[b,c]$ Riemann-integrierbar sind. Es gilt dann:
$$\int_a^c f(x)\,dx = \int_a^b f(x)\,dx + \int_b^c f(x)\,dx \tag{68}$$

Beweis: Übung.

Definition 5.9 (Erweiterung der Integralnotation)

a) Es seien $a < b$ und $f \in \mathcal{R}[a,b]$. Wir definieren
$$\int_b^a f(x)\,dx := -\int_a^b f(x)\,dx.$$
Für das so erweiterte Riemann-Integral gilt die Formel (68) unabhängig von der Anordnung von a, b und c.

b) Eine komplexwertige Funktion $f : [a,b] \to \mathbb{C}$ heißt Riemann-integrierbar, wenn sowohl $\operatorname{Re} f$ als auch $\operatorname{Im} f$ Riemann-integrierbar sind. In diesem Fall setzen wir:
$$\int_a^b f(x)\,dx := \int_a^b \operatorname{Re} f(x)\,dx + i\int_a^b \operatorname{Im} f(x)\,dx.$$
Das so komplexifizierte Integral ist eine $\mathbb{C}$-Linearform.

Satz 5.10 (Mittelwertsatz der Integralrechnung - allgemeine Version) *Es seien $f, g : [a,b] \to \mathbb{R}$ stetig, $g \geq 0,\ a \leq b$. Dann gibt es $\xi \in [a,b]$, so daß gilt:*
$$\int_a^b f(x)g(x)\,dx = f(\xi)\int_a^b g(x)\,dx.$$

Beweis: Wir dürfen $a < b$ annehmen; für $a = b$ sind beide Seiten gleich 0. Die Funktion f nimmt ein Maximum M und eine Minimum m an, da sie stetig ist, und da $[a,b]$ kompakt ist. Aus $m \leq f \leq M$ und $g \geq 0$ folgt $mg \leq fg \leq Mg$, also wegen der Monotonie des Integrals:
$$m\int_a^b g(x)\,dx \leq \int_a^b f(x)g(x)\,dx \leq M\int_a^b g(x)\,dx$$

Im Fall $\displaystyle\int_a^b g(x)\,dx = 0$ sind wir fertig. Im anderen Fall gilt

$$\frac{\int_a^b f(x)g(x)}{\int_a^b g(x)\,dx} \in [m, M]\,.$$

Nach dem Zwischenwertsatz nimmt f diesen Wert an einer Stelle $\xi \in [a,b]$ an. Hieraus folgt die Behauptung. $\square$

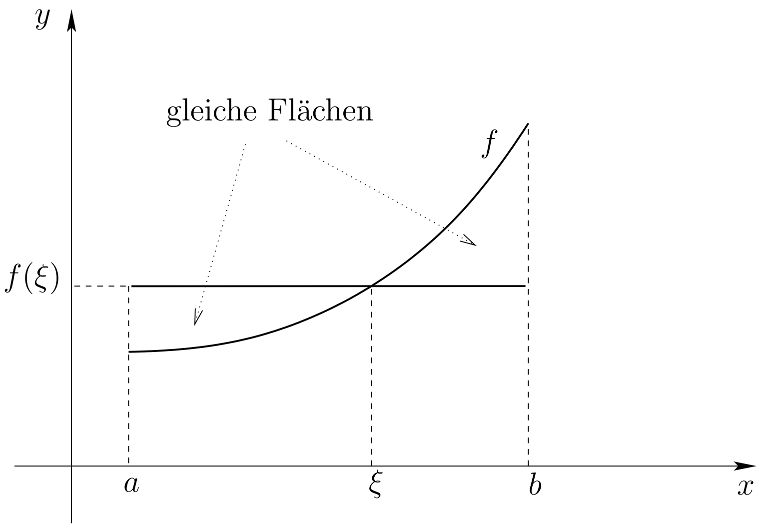
Bemerkung (ohne Beweis): Stetigkeit von g ist hier unwichtig, Riemann-Integrierbarkeit reicht aus.

Der Spezialfall $g = 1$ ist besonders wichtig:

Satz 5.11 (Mittelwertsatz der Integralrechnung - einfache Version) *Es sei $f : [a,b] \to \mathbb{R}$ stetig, $a \leq b$. Dann gibt es $\xi \in [a,b]$, so daß gilt:*

$$\int_a^b f(x)\,dx = f(\xi)(b-a).$$

Illustration:



5.2 Der Hauptsatz der Differential- und Integralrechnung

Integral und Ableitung sind inverse Operationen zueinander:

wichtigster Satz der Vorlesung!

Satz 5.12 (Hauptsatz - erste Version) Es sei $f : [a,b] \to \mathbb{R}$ stetig. Dann ist

$$F : [a,b] \to \mathbb{R}, \qquad F(x) = \int_a^x f(t)\,dt$$

differenzierbar (einseitig differenzierbar am Rand) und es gilt $F' = f$.

Mit anderen Worten:

$$\boxed{\frac{d}{dx}\int_a^x f(t)\,dt = f(x)}$$

Beweis: Es seien $x, y \in [a, b]$ mit $x \neq y$. Dann gibt es nach dem Mittelwertsatz der Integralrechnung ein $\xi_{x,y}$ zwischen x und y, so daß gilt:

$$F(y) - F(x) = \int_x^y f(t)\,dt = f(\xi_{x,y})(y - x).$$

Für $y \to x$ folgt $\xi_{x,y} \to x$, also $f(\xi_{x,y}) \to f(x)$, weil f stetig ist. Es folgt:

$$\frac{F(y) - F(x)}{y - x} = f(\xi_{x,y}) \overset{y \to x}{\longrightarrow} f(x)$$

und demnach

$$F'(x) = f(x).$$

$\square$

Satz 5.13 (Hauptsatz - zweite Version) Es sei $F : [a, b] \to \mathbb{R}$ stetig differenzierbar, d.h differenzierbar mit stetiger Ableitung (einseitig differenzierbar am Rand). Dann gilt:

$$\int_a^b F'(x)\,dx = F(b) - F(a).$$

Beweis: Wir betrachten die Funktion $g : [a, b] \to \mathbb{R}$, $g(x) = \displaystyle\int_a^x F'(t)\,dt - F(x)$.
Nach der ersten Version des Hauptsatzes gilt:

$$g'(x) = F'(x) - F'(x) = 0,$$

also ist g konstant (Mittelwertsatz der Differentialrechnung). Es folgt:

$$\int_a^b F'(x)\,dx - F(b) = g(b) = g(a) = -F(a).$$

$\square$

Bemerkung: Durch Real- und Imaginärteilbildung kann man den Hauptsatz sofort auf komplexwertige Funktionen verallgemeinern.

Notation:

- Wir schreiben $[F(x)]_a^b$ oder auch $[F(x)]_{x=a}^b$ für $F(b) - F(a)$.
- Eine Funktion F heißt *Stammfunktion* von f, wenn $F' = f$ gilt. Wir schreiben $\displaystyle\int f(x)\,dx$ für eine beliebige Stammfunktion von f. Sie ist auf einem Intervall nur bis auf eine additive Konstante bestimmt.

5.3 Integrationsregeln

5.3.1 Einige wichtige Integrale

Grundwissen!

$$\int x^s\,dx = \frac{x^{s+1}}{s+1} + C$$

($s \in \mathbb{C} \setminus \{-1\}$, $x > 0$ oder $s \in \mathbb{N}_0$, $x \in \mathbb{R}$)

$$\int \frac{dx}{x} = \log|x| + C = \log(\alpha x)$$

($x > 0$, $\alpha > 0$ oder $x < 0$, $\alpha < 0$)

$$\int e^x\,dx = e^x + C$$

eng verwandte Varianten davon:

$\int a^x\,dx = \frac{a^x}{\log a} + C \qquad (a > 0)$

$\int e^{\lambda x}\,dx = \frac{e^{\lambda x}}{\lambda} + C \qquad (\lambda \in \mathbb{C} \setminus \{0\})$

$$\int \sin x\,dx = -\cos x + C$$

$$\int \cos x\,dx = \sin x + C$$

$$\int \tan x\,dx = -\log\cos x + C \qquad (-\tfrac{\pi}{2} < x < \tfrac{\pi}{2})$$

$$\int \cot x\,dx = \log\sin x + C \qquad (0 < x < \pi)$$

$$\int \frac{dx}{1+x^2} = \arctan x + C$$

$$\int \frac{dx}{\sqrt{1-x^2}} = \arcsin x + C \qquad (-1 < x < 1)$$

5.3.2 Partielle Integration und Substitutionsregel

Satz 5.14 (partielle Integration) Es seien $f, g : [a, b] \to \mathbb{C}$ stetig differenzierbar, d.h. differenzierbar mit stetiger Ableitung. Dann gilt:

$$\boxed{\int_a^b f'(x)g(x)\,dx = [f(x)g(x)]_a^b - \int_a^b f(x)g'(x)\,dx}$$

Beweis: Nach der Produktregel gilt:

$$f'g = (fg)' - fg'$$

Mit dem Hauptsatz folgt die Behauptung durch Integration über $[a, b]$. $\square$

Beispiel 5.15 Für $b > 0$ gilt:

$$\begin{aligned}
\int_1^b \log x \, dx = \int_1^b 1 \log x \, dx &= \int_1^b \left(\frac{d}{dx} x\right) \log x \, dx \\
&= [x \log x]_1^b - \int_1^b x \left(\frac{d}{dx} \log x\right) \, dx \\
&= b \log b - \int_1^b x \cdot \frac{1}{x} \, dx \\
&= b \log b - [x]_1^b = b \log b - b + 1.
\end{aligned}$$

Satz 5.16 (Substitutionsregel) Es seien $U \subseteq \mathbb{R}$ offen, $g : [a, b] \to U$ stetig differenzierbar, und $f : U \to \mathbb{C}$ stetig. Dann gilt:

$$\boxed{\int_{g(a)}^{g(b)} f(y) \, dy = \int_a^b f(g(x)) g'(x) \, dx}$$

Bemerkung: Schreiben wir $y = y(x) = g(x)$, so können wir die Substitutionsregel auch in der intuitiven Notation

$$\int f(y) \, dy = \int f(y(x)) \frac{dy}{dx} \, dx$$

schreiben. Implizit bleibt hier, an welchen Grenzen die Integrale ausgewertet werden.

Beweis der Substitutionsregel: Mit der Kettenregel und dem Hauptsatz (1.Version) erhalten wir für $x \in [a, b]$:

$$\frac{d}{dx} \int_{g(a)}^{g(x)} f(y) \, dy = f(g(x)) g'(x)$$

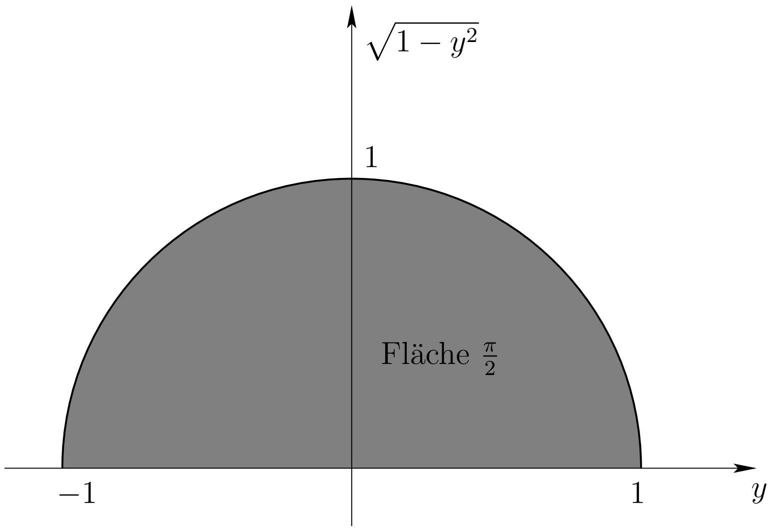
Mit dem Hauptsatz (2.Version) folgt die Behauptung durch Integration über $[a, b]$. $\square$

Beispiel 5.17 Wir berechnen $\int_{-1}^{1} \sqrt{1-y^2}\, dy$ mit der Substitution $y = \sin x$:

$$\begin{aligned}
\int_{-1}^{1} \sqrt{1-y^2}\, dy &= \int_{\sin(-\frac{\pi}{2})}^{\sin(\frac{\pi}{2})} \sqrt{1-y^2}\, dy = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \sqrt{1-\sin^2 x}\, \frac{d\sin x}{dx}\, dx \\
&= \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos^2 x\ dx = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \left(\frac{e^{ix}+e^{-ix}}{2}\right)^2 dx = \frac{1}{4}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \left[e^{2ix}+2+e^{-2ix}\right]\, dx \\
&= \frac{1}{4}\left[\frac{e^{2ix}}{2i}+2x+\frac{e^{-2ix}}{-2i}\right]_{-\frac{\pi}{2}}^{\frac{\pi}{2}} = \left[\frac{x}{2}+\frac{1}{4}\sin(2x)\right]_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \\
&= \frac{\pi}{2}
\end{aligned}$$

Wir hätten natürlich auch das Additionstheorem $\cos^2 x = \frac{1}{2}(1+\cos(2x))$ direkt verwenden können.

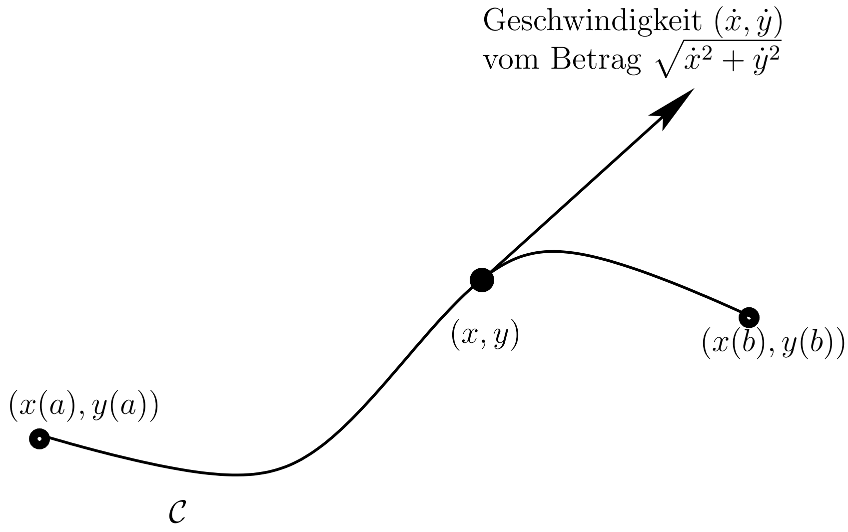
Geometrisch ist das Ergebnis offensichtlich, denn es handelt sich um die Fläche der oberen Hälfte des Einheitskreises:



5.4 Anwendungen

Neben der Interpretation als Fläche unter Funktionsgraphen hat das Integral unzählige weitere Anwendungen, von denen wir ein paar ausgewählte hier besprechen. Zum Teil werden die Anwendungen hier heuristisch-intuitiv präsentiert.

a) **Bogenlänge:**
   Betrachten wir eine glatte Kurve $\mathcal{C}$ in der Ebene oder im Raum, parametrisiert durch $(x(t), y(t))$ bzw. $(x(t), y(t), z(t))$, $t \in [a, b]$.



Wird die Kurve mit Geschwindigkeit 1 durchlaufen, d.h. falls

$$\sqrt{\dot{x}^2 + \dot{y}^2} = 1 \qquad \text{bzw.} \qquad \sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2} = 1$$

(wobei der Punkt $\dot{}$ für $\frac{d}{dt}$ steht), so gilt

$$\text{Länge von } \mathcal{C} = \text{ benötigte Zeit zum Durchlaufen } = b - a.$$

Im allgemeinen erhält man die Bogenlänge L als das Integral über die Geschwindigkeit:

$$L = \int_a^b \sqrt{\dot{x}(t)^2 + \dot{y}(t)^2}\, dt \qquad \text{bzw.} \qquad L = \int_a^b \sqrt{\dot{x}(t)^2 + \dot{y}(t)^2 + \dot{z}(t)^2}\, dt$$

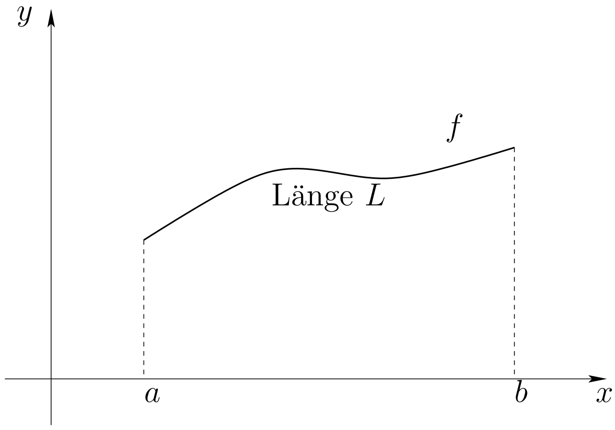
In der Tat hängt dieses Integral nicht von der Parametrisierung ab, d.h. nicht von der Geschwindigkeit, mit der die Kurve durchlaufen wird.
Ist nämlich $[c, d] \to [a, b]$, $s \mapsto t(s)$ ein Wechsel der Parametrisierung (bijektiv mit positiver stetiger Ableitung $\dfrac{dt}{ds}$) so folgt mit der Substitutionsregel:

$$\begin{aligned}
\int_c^d \sqrt{\left[\frac{d}{ds} x(t(s))\right]^2 + \left[\frac{d}{ds} y(t(s))\right]^2} ds &= \int_c^d \sqrt{\left[\dot{x}(t(s)) \frac{dt}{ds}\right]^2 + \left[\dot{y}(t(s)) \frac{dt}{ds}\right]^2} ds \\
&= \int_c^d \sqrt{\dot{x}(t(s))^2 + \dot{y}(t(s))^2}\, \frac{dt}{ds}\, ds \\
&= \int_a^b \sqrt{\dot{x}(t)^2 + \dot{y}(t)^2}\, dt.
\end{aligned}$$

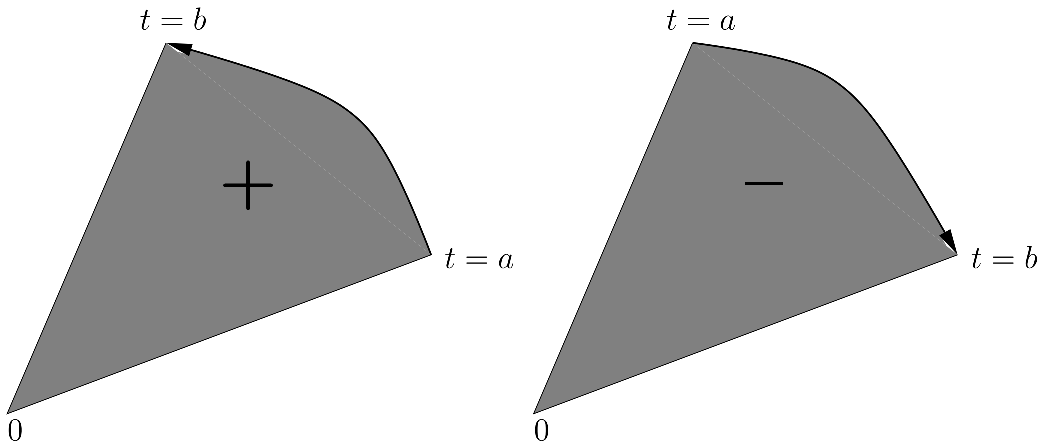
Parametrisiert man insbesondere mit der Bogenlänge $t(s) = \displaystyle\int_a^s \sqrt{\dot{x}(u)^2 + \dot{y}(u)^2}\, du$, so wird die Durchlaufgeschwindigkeit gleich 1.

Beispiel 5.18 *Die Länge eines Funktionsgraphen zu $f : [a, b] \to \mathbb{R}$ ergibt sich mit der Parametrisierung $(x(t), y(t)) = (t, f(t))$ als*

$$L = \int_a^b \sqrt{\dot{x}(t)^2 + \dot{y}(t)^2}\, dt = \int_a^b \sqrt{1 + f'(x)^2}\, dx.$$

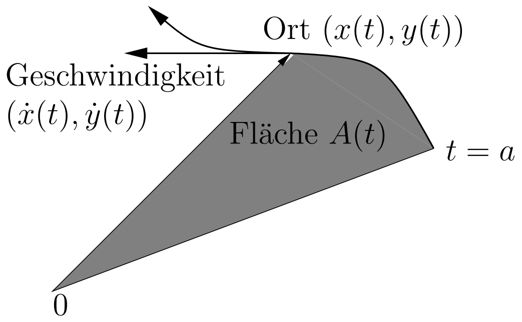


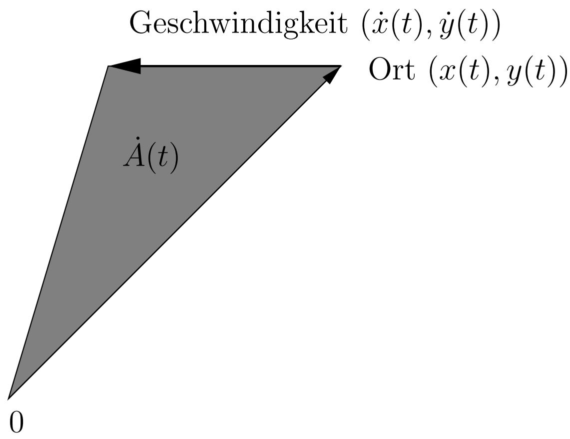
b) **Vom Ortsvektor überstrichene Fläche:**
Betrachten wir eine Kurve $(x(t), y(t))$ in der Ebene, $t \in [a, b]$, und die vom Ortsvektor überstrichene Fläche:



$$\text{In } \left.\begin{matrix}\text{positiver}\\ \text{negativer}\end{matrix}\right\} \text{Umlaufrichtung zählt die Fläche } \left\{\begin{matrix}\text{positiv}\\ \text{negativ}\end{matrix}\right. .$$

Flächenzuwachsrate $\dot{A}$:

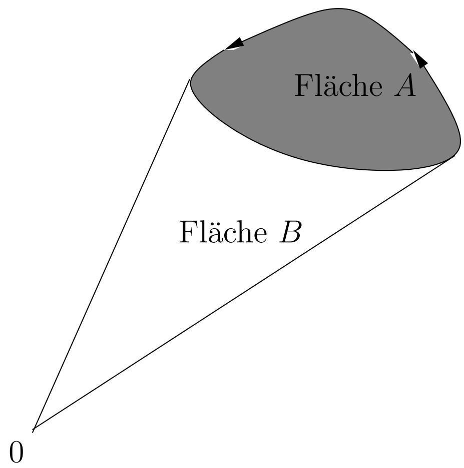




$$\dot{A} = \frac{1}{2} \begin{vmatrix} x & \dot{x} \\ y & \dot{y} \end{vmatrix} = \frac{1}{2} (x\dot{y} - y\dot{x})$$

$$\boxed{\text{überstrichene Fläche} = \frac{1}{2} \int_a^b [x(t)\dot{y}(t) - y(t)\dot{x}(t)] \ dt}$$

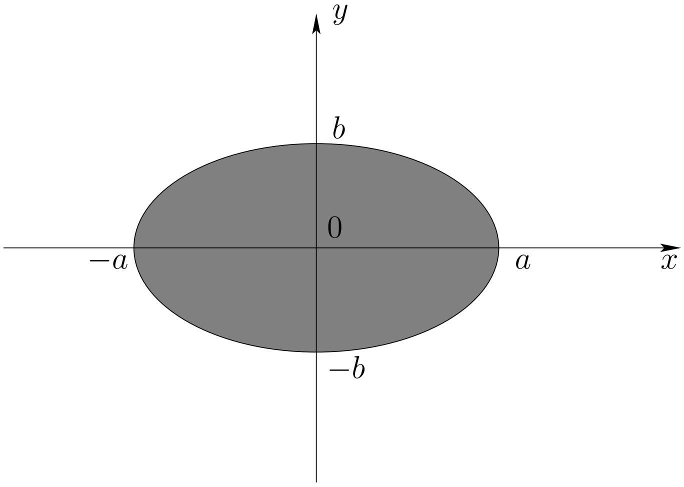
Flächeninhalt, der von geschlossenen Kurven begrenzt wird:



Fläche A wird einmal in positive Richtung überstrichen, zählt also einfach positiv. Fläche B wird zweimal überstrichen: einmal in positive Richtung und einmal in negative Richtung, zählt also nicht.

$$A = \frac{1}{2} \int\limits_{1 \text{ Umlauf}} \begin{vmatrix} x & \dot{x} \\ y & \dot{y} \end{vmatrix} dt = \frac{1}{2} \int\limits_{1 \text{ Umlauf}} [x(t)\dot{y}(t) - y(t)\dot{x}(t)] \ dt$$

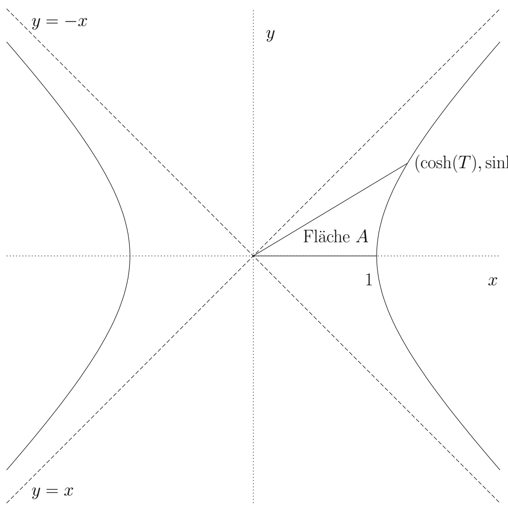
Beispiel Ellipsenfläche:



$\left(\frac{x}{a}\right)^2 + \left(\frac{y}{b}\right)^2 = 1,$
parametrisiert durch
$x = a\cos t$
$y = b\sin t,\ \ t \in [0, 2\pi].$

$$\begin{aligned}\text{Ellipsenfläche} &= \frac{1}{2}\int_0^{2\pi}\left[a\cos t\frac{d}{dt}(b\sin t) - b\sin t\frac{d}{dt}(a\cos t)\right] dt\\ &= \frac{1}{2}\int_0^{2\pi} ab\left(\cos^2 t + \sin^2 t\right)\, dt = \pi ab.\end{aligned}$$

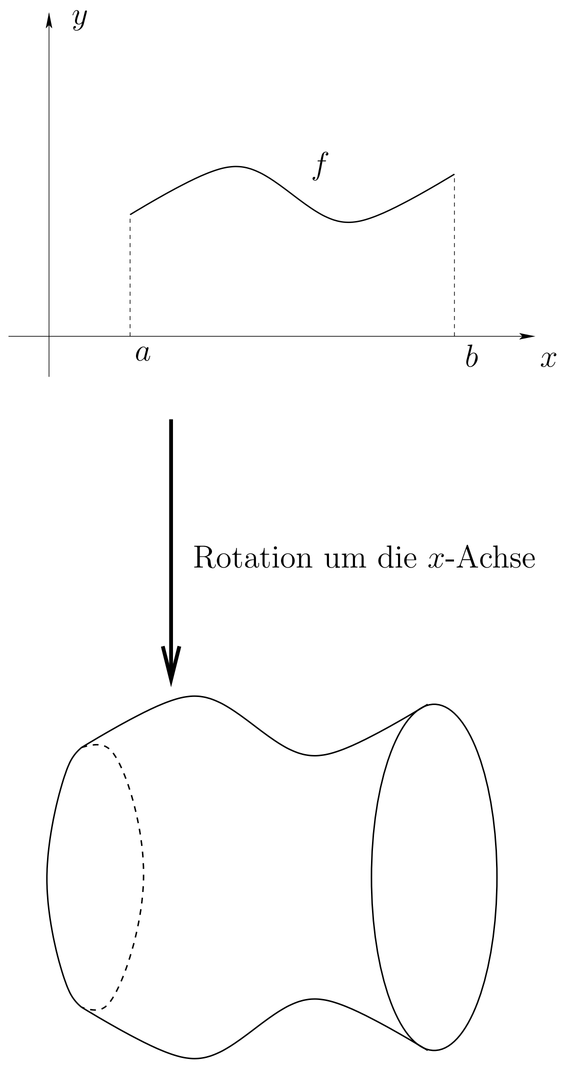
Beispiel Hyperbelsektor:



$$A = \frac{1}{2}\int_0^T \left[\cosh t \frac{d}{dt}(\sinh t) - \sinh t \frac{d}{dt}(\cosh t)\right] dt$$
$$= \frac{1}{2}\int_0^T \left[\cosh^2 t - \sinh^2 t\right] dt = \frac{1}{2}\int_0^T dt = \frac{T}{2}.$$

Diese Rechnung rechtfertigt nachträglich die Bezeichnung “Areafunktionen” für die Umkehrfunktionen der Hyperbelfunktionen.

c) **Volumen von Rotationskörpern**



Für eine Treppenfunktionen $g(x)$ erhält man Zylinderscheiben mit einem gesamten Volumen von $\pi \int_a^b g(x)^2\, dx$.

Durch Approximation einer Funktion $f : [a,b] \to [0,\infty[$ mit Treppenfunktionen erhält man für das Volumen des Rotationskörpers:

$$\boxed{V = \pi \int_a^b f(x)^2 \, dx}$$

Beispiel 5.19 (Kugelvolumen) Durch Rotation um die x-Achse des Halbkreises

$$\{(x, \sqrt{r^2 - x^2}) \mid -r \leq x \leq r\}$$

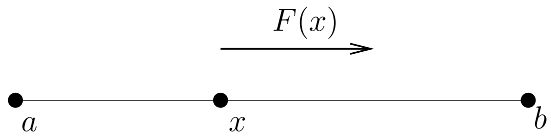
mit Radius r erhält man eine Kugel mit dem Radius r. Wir berechnen das Volumen V dieser Kugel:

$$\begin{aligned} V &= \pi \int_{-r}^{r} \left(\sqrt{r^2 - x^2}\right)^2 dx = \pi \int_{-r}^{r} (r^2 - x^2) \, dx \\ &= \pi \left[r^2 x - \frac{1}{3}x^3\right]_{x=-r}^{r} = \pi \left[2r^3 - \frac{2}{3}r^3\right] \\ &= \frac{4}{3}r^3\pi \end{aligned}$$

d) **physikalische Arbeit:**
Ein Körper bewege sich in einem Kraftfeld (in einer Dimension). Am Ort x wirke die Kraft $F(x)$. Wenn der Körper von a nach b bewegt wird, wird an ihm die Arbeit

$$W = \int_a^b F(x) \, dx$$

verrichtet.



Wenn F stückweise konstant ist, etwa auf den Teilstücken $[x_j, x_{j+1}[$ einer Zerlegung $a = x_0 \leq x_1 \leq \cdots \leq x_n = b$, folgt das unmittelbar aus der physikalischen Regel „Arbeit = Kraft · Weg“:

$$W = \sum_{j=0}^{n-1} F(x_j)(x_{j+1} - x_j) = \int_a^b F(x) \, dx$$

Wenn hingegen F stetig (oder allgemeiner: Riemann-integrierbar) ist, folgt die Formel durch Approximation von oben und unten durch Treppenfunktionen.

Beispiel 5.20 Auf zwei positive elektrische Ladungen im Abstand x wirkt eine Kraft $F(x) = c/x^2$ mit einer Konstanten c. Bei der Bewegung von r_1 nach r_2 ($r_1 < r_2$) wird die Arbeit

$$W = \int_{r_1}^{r_2} \frac{c}{x^2}\,dx = \left[-\frac{c}{x}\right]_{r_1}^{r_2} = \frac{c}{r_1} - \frac{c}{r_2}$$

freigesetzt.

5.5 Uneigentliche Riemann-Integrale

Es seien $a, b \in \mathbb{R} \cup \{\pm\infty\},\ a < b$ und $f\colon]a,b[\ \to \mathbb{C}$.
Die Funktion f heißt uneigentlich Riemann-integrierbar, wenn sie auf allen kompakten Teilintervallen $[a_1, b_1] \subset]a,b[$ Riemann-integrierbar ist und wenn

$$\lim_{a_1 \downarrow a} \lim_{b_1 \uparrow b} \int_{a_1}^{b_1} f(x)\,dx$$

existiert und endlich ist.
Der Grenzwert heißt *uneigentliches Riemann-Integral* und wird auch mit

$$\int\limits_a^b f(x)\,dx$$

bezeichnet.

Beispiel 5.21 Betrachten wir die Funktion $f\colon]0,1] \to \mathbb{C},\ f(x) = x^s$ mit $s \in \mathbb{C}$.

- Für $\operatorname{Re} s > 0$ ist f nach 0 stetig fortsetzbar mit dem Wert 0, denn

  $$|x^s| = |e^{s \log x}| = e^{(\operatorname{Re} s)\log x} = x^{\operatorname{Re} s} \ \to 0 \ \text{ für } x \downarrow 0.$$

  Wir erhalten

  $$\int_0^1 x^s\,dx = \left[\frac{x^{s+1}}{s+1}\right]_0^1 = \frac{1}{s+1}$$

  als gewöhnliches Riemann-Integral.

- Für $s = 0$ ist x^0 konstant gleich 1, also $\int_0^1 x^0\,dx = 1$.

- Für $-1 < \operatorname{Re} s < 0$ gilt $x^s \to \infty$ für $x \downarrow 0$ in $\mathbb{C} \cup \{\infty\}$. Dennoch existiert hier das uneigentliche Riemann-Integral:

$$\begin{aligned}\int_0^1 x^s\,dx = \lim_{\varepsilon\downarrow 0}\int_\varepsilon^1 x^s\,dx &= \lim_{\varepsilon\downarrow 0}\left[\frac{x^{s+1}}{s+1}\right]_\varepsilon^1 \\ &= \lim_{\varepsilon\downarrow 0}\left(\frac{1}{s+1} - \frac{\varepsilon^{s+1}}{s+1}\right) \\ &= \frac{1}{s+1}\end{aligned}$$

  Man beachte $\operatorname{Re}(s+1) > 0$.

- Für $s = -1$ gilt:

$$\int_\varepsilon^1 \frac{dx}{x} = -\log\varepsilon \overset{\varepsilon\downarrow 0}{\to} +\infty$$

  Also ist das uneigentliche Riemann-Integral $\int\limits_0^1 \frac{dx}{x}$ nicht endlich.

  Auch für alle anderen $s \in \mathbb{C}$ mit $\operatorname{Re} s \leq -1$ existiert $\int\limits_0^1 x^s\,dx$ in $\mathbb{C}$ nicht.

Beispiel 5.22 Die Funktion $f : [1, +\infty[\to \mathbb{C},\ f(x) = x^s$ ist uneigentlich Riemann-integrierbar für $\operatorname{Re} s < -1$ mit

$$\int_1^{+\infty} x^s\,dx = \left[\frac{x^{s+1}}{s+1}\right]_0^{+\infty} = -\frac{1}{s+1}$$

Für $\operatorname{Re} \geq -1$ ist sie nicht Riemann-integrierbar.

5.5.1 Die Gammafunktion

Für $s \in \mathbb{C}$ mit $\operatorname{Re} s > 0$ definieren wir die Gammafunktion

$$\Gamma(s) = \int_0^{+\infty} e^{-x}x^{s-1}\,dx$$

Für $\operatorname{Re} s \leq 0$ ist dieses Integral divergent bei 0 (ohne Beweis).
Dieses Integral ist ein uneigentliches Riemann-Integral bei $+\infty$. Für $0 < \operatorname{Re} s \leq 1$ fassen wir es als uneigentliches Riemann-Integral bei 0 auf.
Man beachte, dass $\displaystyle\int_0^1 e^{-x}x^{s-1}\,dx$ wegen $|e^{-x}x^{s-1}| \leq x^{\operatorname{Re} s-1}$ durch

$$\frac{1}{\operatorname{Re} s} = \int_0^1 x^{\operatorname{Re} s-1}\,dx \geq \int_0^1 |e^{-x}x^{s-1}|\,dx$$

dominiert wird.

Satz 5.23 (Funktionalgleichung der Γ-Funktion) *Für $s \in \mathbb{C}$, $\operatorname{Re} s > 0$ gilt:*

$$\boxed{\Gamma(s+1) = s\Gamma(s)}$$

Beweis: Durch partielle Integration folgt:

$$\begin{aligned}
\Gamma(s+1) &= \int_0^\infty e^{-x} x^s \, dx \\
&= -\int_0^\infty \left(\frac{d}{dx} e^{-x}\right) x^s \, dx \\
&= -\left[e^{-x} x^s\right]_0^\infty + \int_0^\infty e^{-x} \left(\frac{d}{dx} x^s\right) dx \\
&= s\Gamma(s)
\end{aligned}$$

Dabei haben wir $e^{-x} \ll x^{-s}$ für $x \to \infty$ und $e^{-x}x^s \to 0$ für $x \downarrow 0$ verwendet.

□

Folgerung: Für $n \in \mathbb{N}_0$ gilt:

$$\boxed{n! = \Gamma(n+1)}$$

Beweis: (Induktion über n)

- $\boxed{n = 0}$

$$\Gamma(1) = \int_0^\infty e^{-x} \, dx = -\left[e^{-x}\right]_0^\infty = 1 = 0!$$

- $\boxed{n - 1 \rightsquigarrow n}$ Es gelte die Induktionsvoraussetzung $(n-1)! = \Gamma(n)$. Dann folgt mit der Funktionalgleichung:

$$\Gamma(n+1) = n\Gamma(n) = n(n-1)! = n!$$

□

5.6 Symbolische Integrationsverfahren für einige Funktionenklassen

In diesem Abschnitt besprechen wir Verfahren, mit denen man in manchen Fällen die Stammfunktion von Funktionen explizit berechnen kann. Die Bezeichnung "symbolisch" steht hier für die Berechnungen auf der Ebene von Termen mit Variablen, im Gegensatz zu numerischen Verfahren.

Vergleich von Ableitung und Integration:

- *Symbolische Ableitung*:
  Die symbolische Ableitung von Termen, die aus elementaren Termen wie x, e^x, $\log x$ etc. durch Zusammensetzen mit Komposition und arithmetischen Operationen gebildet sind, erfolgt nach einem sehr einfachen algorithmischen Schema, das immer zum Ziel führt.

- *Symbolische Integration*:
  Im Gegensatz dazu ist die symbolische Integration, also das Auffinden einer expliziten Darstellung einer Stammfunktion mit elementaren Funktionen, nicht immer elementar lösbar und in manchen Fällen sehr schwierig. Wichtige Beispiele nicht elementar darstellbarer Stammfunktionen sind:

  - das in der Stochastik wichtige *“Fehlerintegral”*
  $$\int e^{-x^2}\,dx,$$
  - der *“Integrallogarithmus”*
  $$\int \frac{dx}{\log x},$$
  - die *“Integralexponentialfunktion”*
  $$\int \frac{e^x}{x}\,dx.$$

  Effiziente Algorithmen zur symbolischen Integration spielen in Computeralgebrasystemen eine wichtige Rolle. Die genaue Beschreibung von Algorithmen zur symbolischen Integration führt weit über den Rahmen dieser Vorlesung hinaus. Tieferes für fortgeschrittene Leser findet sich in der Spezialliteratur, z.B. [Bro05]. Hier haben wir nur ein bescheideneres Ziel: Die Integration einiger Funktionsklassen.

5.6.1 Rationale Funktionen

Wir betrachten nun Integrale vom Typ
$$\int \frac{p(x)}{q(x)}\,dx, \qquad p, q \in \mathbb{C}[x],\ q \neq 0.$$
Durch Polynomdivision kann man den Integranden stets in die Form
$$\frac{p(x)}{q(x)} = s(x) + \frac{r(x)}{q(x)}, \quad s, r \in \mathbb{C}[x],\ \deg r < \deg q$$
bringen. Hierbei bezeichnet $\mathbb{C}[x]$ den Raum der Polynome in der Unbestimmten x mit komplexen Koeffizienten und $\deg : \mathbb{C}[x] \to \mathbb{N}_0 \cup \{-\infty\}$ die Gradabbildung für Polynome.

Beispiel 5.24

$$\frac{x^3+2x^2+2x-1}{x^2+1} = x+2+\frac{x-3}{x^2+1}$$

Der ganzrationale Anteil lässt sich sofort integrieren:

$$\int \frac{x^3+2x^2+2x-1}{x^2+1}\,dx = \frac{x^2}{2}+2x+\int \frac{x-3}{x^2+1}\,dx$$

Wir setzen nun vorraus, dass eine Faktorisierung des Nennerpolynoms q in Linearfaktoren bekannt ist.

Effiziente Algorithmen für Computeralgebrasysteme verwenden verfeinerte Methoden, die keine solche Faktorisierung benötigen.

Es gilt:

Fundamentalsatz der Algebra:

Jedes Polynom $q \in \mathbb{C}[x]$ vom Grad $n = \deg q > 0$ besitzt eine Zerlegung in Linearfaktoren:

$$q(x) = c \cdot \prod_{k=1}^{m} (x-\alpha_k)^{n_k}$$

mit $c \in \mathbb{C} \setminus \{0\},\ \alpha_1, \ldots, \alpha_m \in \mathbb{C},\ n_k \in \mathbb{N},\ \sum\limits_{k=1}^{m} n_k = n.$

Hierbei sind $\alpha_1, \ldots, \alpha_m$ die verschiedenen Nullstellen von q und n_k die Vielfachheit von α_k. Die Linearfaktorzerlegung ist bis auf Vertauschung der Faktoren eindeutig.

Für einen eleganten Beweis des Satzes fehlen uns noch die Hilfsmittel. Wir verzichten hier darauf und verweisen auf eine Algebra- oder Funktionentheorie-Vorlesung.

Beispiel 5.25 (Fortsetzung)

$$x^2+1 = (x+i)(x-i)$$

Wir dürfen annehmen, dass das Nennerpolynom q normiert ist, d.h. dass $c = 1$ gilt, sonst kürzen wir $\frac{p}{q}$ mit c.

Satz 5.26 (Partialbruchzerlegung über $\mathbb{C}$) *Für jedes $r \in \mathbb{C}[x]$ mit $\deg r < n = \deg q$ besteht eine eindeutig bestimmte Zerlegung:*

$$\frac{r(x)}{q(x)} = \sum_{k=1}^{m}\sum_{j=1}^{n_k} \frac{c_{kj}}{(x-\alpha_k)^j},\ c_{kj} \in \mathbb{C}$$

Anders geschrieben (durch Multiplikation mit q):

$$r(x) = \sum_{k=1}^{m} \sum_{j=1}^{n_k} c_{kj} b_{kj}(x)$$

wobei

$$b_{kj}(x) = (x-\alpha_k)^{n_k-j} \prod_{\substack{l=1\\l\neq k}}^{m} (x-\alpha_l)^{n_l},\ k=1,\dots,m,\ j=1,\dots,n_k$$

In der Sprache der linearen Algebra heißt das:

Lemma 5.27 *Die Polynome* $(b_{kj})_{\substack{k=1\dots m,\\ j=1\dots n_k}}$ *bilden eine Basis des* $\mathbb{C}$*-Vektorraums* $R := \{r \in \mathbb{C}[x] \mid \deg r < n\}$

Obwohl das eine rein algebraische Aussage ist, ziemt sich für eine Analysis-Vorlesung ein analytischer Beweis:

Beweis: Weil $\dim R = n = \sum\limits_{k=1}^{m} n_k$ gleich der Anzahl der b_{kj} ist, genügt es zu zeigen, dass die b_{kj} linear unabhängig sind.

Betrachten wir also eine beliebige Linearkombination

$$r(x) = \sum_{k=1}^{m} \sum_{j=1}^{n_k} c_{kj} b_{kj}(x)$$

so dass nicht alle c_{kj} gleich 0 sind.

Wählen wir ein k_0 und ein möglichst großes $j_0 \geq 1$ mit $c_{k_0 j_0} \neq 0$. Dann gilt:

$$\frac{r(x)}{q(x)} = \frac{c_{k_0 j_0}}{(x-\alpha_{k_0})^{j_0}} + \sum_{j=1}^{j_0-1} \frac{c_{k_0 j}}{(x-\alpha_{k_0})^{j_0}} + \sum_{\substack{k=1\\k\neq k_0}}^{m} \sum_{j=1}^{n_k} \frac{c_{kj}}{(x-\alpha_k)^j} \tag{69}$$

Betrachten wir die Asymptotik $x \to \alpha_{k_0}$:

Die Summe in der Mitte ist von der Ordnung

$$O(|x-\alpha_{k_0}|^{1-j_0}) = o(|x-\alpha_{k_0}|^{-j_0})$$

und die Summe rechts in (69) hat in dieser Asymptotik folgende Schranke:

$$O(1) = o(|x-\alpha_{k_0}|^{-j_0}),$$

weil die α_k alle verschieden sind.

Nun ist aber der erste Summand $\dfrac{c_{k_0 j_0}}{(x-\alpha_{k_0})^{j_0}}$ in (69) nicht von der Ordnung $o(|x-\alpha_{k_0}|^{-j_0})$.

Also ist $r \neq 0$ und damit die lineare Unabhängigkeit der b_{kj} gezeigt.

$\square$

Zur effektiven Berechnung der c_{kj} kann man die Gleichung

$$r(x) = \sum_{k=1}^{m} \sum_{j=1}^{n_k} c_{kj} b_{kj}(x)$$

an den Stellen $x = \alpha_k$ auswerten, ebenso die Ableitungen bis zur Ordnung $n_k - 1$. Es gilt nämlich $b_{lj}(\alpha_k) = 0$ für $k \neq l$ (ebenso für die Ableitungen bis zur Ordnung $n_k - 1$).

Beispiel 5.28 (Fortsetzung)

$$\frac{x-3}{x^2+1} = \frac{c_{11}}{x+i} + \frac{c_{21}}{x-i},$$

anders geschrieben:

$$x - 3 = c_{11}(x-i) + c_{21}(x+i)$$

$x = -i$ eingesetzt:

$$-3 - i = -2ic_{11}, \qquad \text{also} \qquad c_{11} = \frac{1}{2} - \frac{3}{2}i$$

$x = +i$ eingesetzt:

$$-3 + i = 2ic_{21}, \qquad \text{also} \qquad c_{21} = \frac{1}{2} + \frac{3}{2}i$$

Das bedeutet:

$$\frac{x-3}{x^2+1} = \frac{1}{2}(1-3i)\frac{1}{x+i} + \frac{1}{2}(1+3i)\frac{1}{x-i}.$$

Die Partialbrüche $\dfrac{c_{kj}}{(x-\alpha_k)^j}$ lassen sich sofort integrieren:

$$\int \frac{c_{kj}}{(x-\alpha_k)^j}\,dx = \frac{c_{kj}}{1-j} \cdot \frac{1}{(x-\alpha_k)^{j-1}} + C \quad \text{für } j \neq 1$$

$$\int \frac{c_{kj}}{x-\alpha_k}\,dx = c_{kj} \log\left((x-\alpha_k)\beta\right) \quad \text{für } j = 1,\ \alpha_k \in \mathbb{R}$$

mit einer multiplikativen Integrationskonstanten β. Für komplexe α_k bedarf die letzte Formel einer besonderen Interpretation. Man erinnere sich an das Problem, dass der Logarithmus im Komplexen mehrdeutig ist wegen $e^{z+2\pi ik} = e^z$ für $k \in \mathbb{Z}$, $z \in \mathbb{C}$.
Wir schreiben für $\operatorname{Im}\alpha_k \neq 0, x \in \mathbb{R}$:

$$\frac{1}{x-\alpha_k} = \frac{x-\overline{\alpha}_k}{(x-\alpha_k)(x-\overline{\alpha}_k)} = \frac{x - \operatorname{Re}\alpha_k + i\operatorname{Im}\alpha_k}{(x-\operatorname{Re}\alpha_k)^2 + (\operatorname{Im}\alpha_k)^2} =$$

$$= \frac{d}{dx}\left[\frac{1}{2}\log\left[(x-\operatorname{Re}\alpha_k)^2 + (\operatorname{Im}\alpha_k)^2\right] + i\arctan\frac{x-\operatorname{Re}\alpha_k}{\operatorname{Im}\alpha_k}\right]$$

kurz geschrieben:

$$\int \frac{dx}{x-\alpha_k} = \log|x-\alpha_k| + i \arctan \frac{x - \operatorname{Re}\alpha_k}{\operatorname{Im}\alpha_k} + C \text{ für } \alpha_k \in \mathbb{C}\setminus\mathbb{R}, \tag{70}$$

wobei wir

$$\log\left[(x-\operatorname{Re}\alpha_k)^2 + (\operatorname{Im}\alpha_k)^2\right] = \log(|x-\alpha_k|^2) = 2\log|x-\alpha_k|$$

verwendet haben. Man vergleiche die Formel (70) mit $z = e^{\log|z| + i\arg z}$, wobei $\arg z$ den (bis auf Vielfache von 2π bestimmten) Winkel in der Polardarstellung von $z \in \mathbb{C}\setminus\{0\}$ bezeichnet.

im Bsp:

$$\int \frac{dx}{x \pm i} = \frac{1}{2}\log(x^2+1) \mp i \arctan x,$$

also

$$\int \frac{x^3+2x^2+2x-1}{x^2+1}\,dx = \frac{x^2}{2} + 2x + \frac{1}{2}\log(x^2+1) - 3\arctan x + C$$

Reelle Variante: Für reelle Polynome gibt es eine Variante des Verfahrens, die ohne komplexe Zahlen auskommt. Wir skizzieren diese Variante hier nur, ohne Beweise.

Reelle Variante des Fundamentalsatzes der Algebra:
Jedes $q \in \mathbb{R}[x]$, $q \neq 0$, besitzt eine Zerlegung

$$q(x) = c\prod_{k=1}^{m_1}(x-\alpha_k)^{n_k}\prod_{l=1}^{m_2}((x-\beta_l)^2+\gamma_l^2)^{\widetilde{n}_l}$$

mit $\gamma_l > 0$, $\beta_l, \alpha_k \in \mathbb{R}$ und Vielfachheiten $n_k, \widetilde{n}_l \in \mathbb{N}$ der Faktoren. Die komplexen Nullstellen des Polynoms q sind also α_k (Vielfachheit n_k) und $\beta_l \pm \gamma_l$ (beide mit Vielfachheit $\widetilde{n}_l$).
Damit erhält man eine Partialbruchzerlegung im Reellen:

$$\frac{r(x)}{q(x)} = \sum_{k=1}^{m_1}\sum_{j=1}^{n_k}\frac{c_{kj}}{(x-\alpha_k)^j} + \sum_{l=1}^{m_2}\sum_{j=1}^{\widetilde{n}_l}\frac{\widetilde{c}_{lj} + (x-\beta_l)\widetilde{d}_{lj}}{((x-\beta_l)^2+\gamma_l^2)^j}$$

Die quadratischen Partialbrüche integriert man so: Zerlegen wir einen typischen Summanden:

$$\int \frac{\widetilde{c} + (x-\beta)\widetilde{d}}{[(x-\beta)^2+\gamma^2]^j}\,dx = \widetilde{d}\int \frac{x-\beta}{[(x-\beta)^2+\gamma^2]^j}\,dx + \widetilde{c}\int \frac{dx}{[(x-\beta)^2+\gamma^2]^j}. \tag{71}$$

Mit der Substitution $(x-\beta)^2+\gamma^2 = t$ erhalten wir für das erste Integral auf der rechten Seite in (71):

$$\int \frac{x-\beta}{[(x-\beta)^2+\gamma^2]^j}\,dx = \frac{1}{2}\int \frac{dt}{t^j} = \begin{cases} \dfrac{[(x-\beta)^2+\gamma^2]^{1-j}}{2(1-j)} + C & \text{für } j>1 \\ \dfrac{1}{2}\log[(x-\beta)^2+\gamma^2] + C & \text{für } j=1 \end{cases}$$

Das zweite Integral auf der rechten Seite in (71) berechnen wir mit der Substitution $u := \frac{x-\beta}{\gamma}$:

$$\int \frac{dx}{[(x-\beta)^2+\gamma^2]^j} = \gamma^{1-2j}\int \frac{du}{(u^2+1)^j}$$

Zur Berechnung von $\displaystyle\int \frac{du}{(u^2+1)^j}$ kann man rekursiv vorgehen:

- $\boxed{j=1}$:

$$\int \frac{du}{u^2+1} = \arctan u + C \tag{72}$$

- $\boxed{j \rightsquigarrow j+1}$:

$$\int \frac{du}{(u^2+1)^{j+1}} = \int \frac{du}{(u^2+1)^j} - \int \frac{u^2}{(u^2+1)^{j+1}}du \tag{73}$$

  Für den zweiten Term auf der rechten Seite in (73) erhält man durch partielle Integration:

$$\begin{aligned} -\int \frac{u^2}{(u^2+1)^{j+1}}du &= \frac{1}{2j}\int u\frac{d}{du}[(u^2+1)^{-j}]\,du \\ &= \frac{1}{2j}\frac{u}{(u^2+1)^j} - \frac{1}{2j}\int \frac{du}{(u^2+1)^j} \end{aligned}$$

  also die gewünschte Rekursionsformel:

$$\boxed{\int \frac{du}{(u^2+1)^{j+1}} = \left(1-\frac{1}{2j}\right)\int \frac{du}{(u^2+1)^j} + \frac{1}{2j}\frac{u}{(u^2+1)^j}}$$

5.6.2 Integration einiger anderer Funktionsklassen

1. **Integrale vom Typ $\int R(e^x)\,dx$ mit einer rationalen Funktion R**

   Mit der Substitution $u = e^x$ wird das Integral auf ein Integral mit rationalem Integranden zurückgeführt:

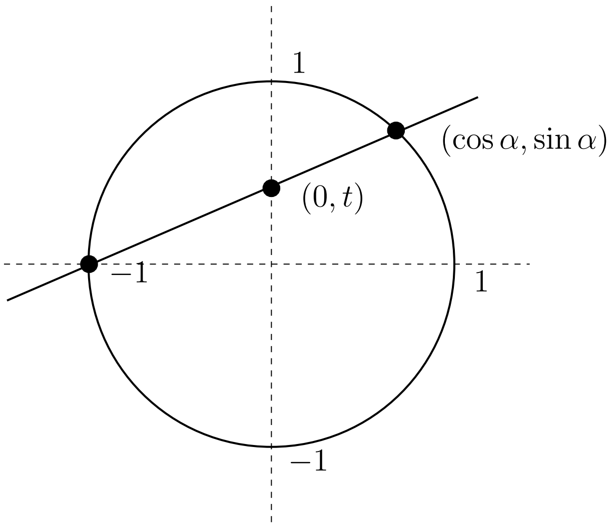
$$\int R(e^x)\,dx = \int R(u)\frac{du}{u}, \quad \text{da } \frac{du}{dx} = e^x = u.$$

2. **Integrale vom Typ $\int R(\cos\alpha, \sin\alpha)\, d\alpha$, ($R$ rational)**

Folgende Substitution macht das Integral rational:[37]

$$]-\pi,\pi[\to \mathbb{R}, \quad \alpha \mapsto t = \frac{\sin\alpha}{1+\cos\alpha} = \tan\frac{\alpha}{2} \tag{74}$$

Geometrisch gesehen ist das die stereographische Projektion, die Sie schon von der Riemannschen Zahlenkugel kennen, nun in zwei Dimensionen:



Die Umkehrtransformation lautet:

$$\cos\alpha = \frac{1-t^2}{1+t^2}, \quad \sin\alpha = \frac{2t}{1+t^2}, \quad \frac{d\alpha}{dt} = \frac{2}{1+t^2}$$

In der Tat: Der Punkt $\left(\frac{1-t^2}{1+t^2}, \frac{2t}{1+t^2}\right) \neq (-1, 0)$, $t \in \mathbb{R}$, liegt auf dem Einheitskreis wegen

$$\left(\frac{1-t^2}{1+t^2}\right)^2 + \left(\frac{2t}{1+t^2}\right)^2 = \frac{(1-2t^2+t^4)+4t^2}{(1+t^2)^2} = \frac{1+2t^2+t^4}{(1+t^2)^2} = 1,$$

[37] Die letzte Gleichung in Formel (74) kann man z.B. so einsehen:

$$\begin{aligned}\frac{\sin\alpha}{1+\cos\alpha} &= \frac{\frac{1}{2i}(e^{i\alpha}-e^{-i\alpha})}{1+\frac{1}{2}(e^{i\alpha}+e^{-i\alpha})} = \frac{1}{i}\frac{e^{i\alpha}-e^{-i\alpha}}{2+e^{i\alpha}+e^{-i\alpha}} \\ &= \frac{1}{i}\frac{(e^{i\alpha/2}+e^{-i\alpha/2})(e^{i\alpha/2}-e^{-i\alpha/2})}{(e^{i\alpha/2}+e^{-i\alpha/2})^2} = \frac{1}{i}\frac{e^{i\alpha/2}-e^{-i\alpha/2}}{e^{i\alpha/2}+e^{-i\alpha/2}} = \tan\frac{\alpha}{2}\end{aligned}$$

ist also von der Gestalt $(\cos\alpha, \sin\alpha)$ für ein eindeutiges $\alpha \in]-\pi,\pi[$, und es gilt

$$\frac{\sin\alpha}{1+\cos\alpha} = \frac{\frac{2t}{1+t^2}}{1+\frac{1-t^2}{1+t^2}} = \frac{2t}{(1+t^2)+(1-t^2)} = t.$$

Die Formel für die Ableitung $d\alpha/dt$ erhält man zum Beispiel mit

$$\begin{aligned}\frac{dt}{d\alpha} = \frac{d}{d\alpha}\frac{\sin\alpha}{1+\cos\alpha} &= \frac{\cos\alpha(1+\cos\alpha)+\sin^2\alpha}{(1+\cos\alpha)^2} \\ &= \frac{1+\cos\alpha}{(1+\cos\alpha)^2} = \frac{1}{1+\cos\alpha} = \frac{1}{1+\frac{1-t^2}{1+t^2}} = \frac{1+t^2}{2}\end{aligned}$$

unter Verwendung der Ableitungsformel für Umkehrfunktionen.
Fassen wir zusammen:

$$\int R(\cos\alpha, \sin\alpha)\, d\alpha = \int R\left(\frac{1-t^2}{1+t^2}, \frac{2t}{1+t^2}\right)\frac{2}{1+t^2}\, dt$$

Auf der rechten Seite steht hier das Integral einer rationalen Funktion.

Bei Integration über die Singularität bei $(\cos\alpha,\sin\alpha) = (-1,0)$ muß man das Integral eventuell in mehrere uneigentliche Riemann-Integrale zerlegen.

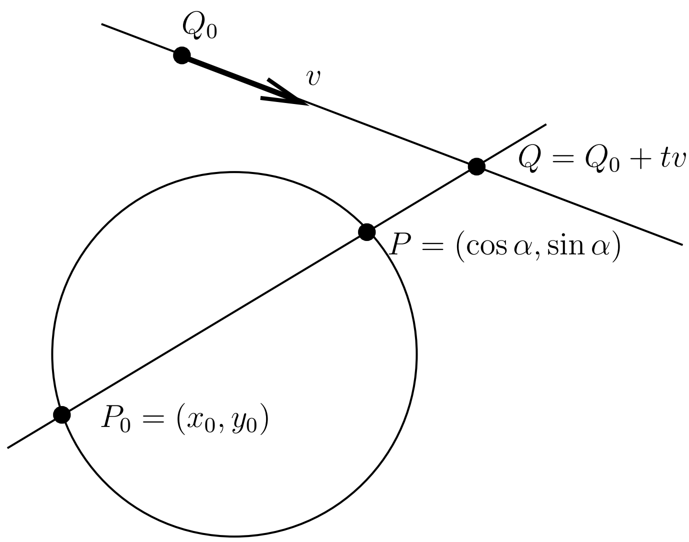
Beispiel 5.29

$$\begin{aligned}\int_{-\pi}^{\pi}\frac{d\alpha}{2+\sin\alpha} &= \int_{-\infty}^{\infty}\frac{1}{2+\frac{2t}{1+t^2}}\frac{2}{1+t^2}\,dt \\ &= \int_{-\infty}^{\infty}\frac{dt}{1+t+t^2} = \int_{-\infty}^{\infty}\frac{dt}{(t+\frac{1}{2})^2+\frac{3}{4}} \\ &= \left[\frac{2}{\sqrt{3}}\arctan\left[\frac{2}{\sqrt{3}}\left(t+\frac{1}{2}\right)\right]\right]_{-\infty}^{\infty} = \frac{2\pi}{\sqrt{3}}\end{aligned}$$

Bemerkungen:

(a) Der Punkt $(-1,0)$ und die Gerade $\{(0,t)|t\in\mathbb{R}\}$ spielen keine ausgezeichnete Rolle. Stattdessen hätten wir auch einen beliebigen Punkt $P_0 = (x_0, y_0)$ auf dem Einheitskreis sowie eine beliebige Gerade, die nicht durch (x_0, y_0) geht, wählen können:

Euler-Substitution $P \mapsto Q$



Die Euler-Substitution $P \mapsto Q$ und ihre Umkehrung $Q \mapsto P$ sind beide rational.

Grund: Die Gleichungen zur Bestimmung von Q bei gegebenem P beschreiben den Schnitt zweier Geraden, also eine rationale Operation. Umgekehrt ist zur Bestimmung von P bei gegebenem Q eine quadratische Gleichung zu lösen (Schnitt von Gerade mit Kreis), von der eine Lösung (nämlich P_0) bekannt ist. Auch das führt auf eine rationale Operation.

(b) Oft sind andere Verfahren einfacher, z.B. das Einsetzen von

$$\cos\alpha = \frac{e^{i\alpha} + e^{-i\alpha}}{2}, \qquad \sin\alpha = \frac{e^{i\alpha} - e^{-i\alpha}}{2i}.$$

3. **Integrale vom Typ $\int R\left(x, \sqrt{ax^2 + bx + c}\right)\, dx$, ($R$ rational)**

Hier setzen wir voraus, dass der Radikand keine doppelte Nullstelle besitzt, d.h. die Diskriminante $b^2 - 4ac$ soll nicht gleich 0 sein. Andernfalls hebt sich nämlich die Wurzel weg:

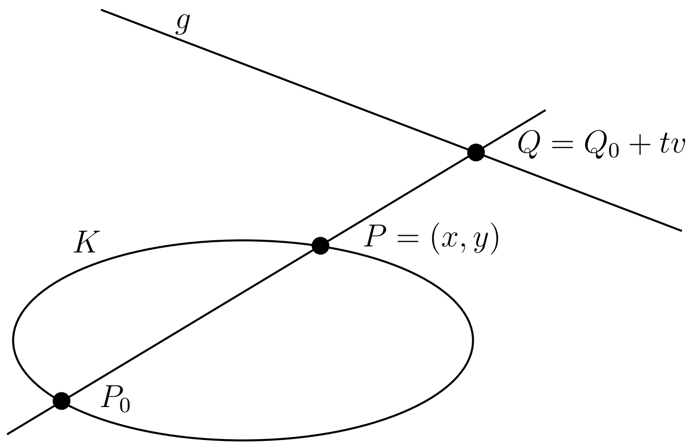
$$\sqrt{ax^2 + bx + \frac{b^2}{4a}} = \sqrt{a}\left|x + \frac{b}{2a}\right|.$$

Weiter soll der Radikand mindestens in einem Intervall positiv sein.

Auch diese Integrale lassen sich mit Euler-Substitutionen rational machen.
Man betrachte den Kegelschnitt

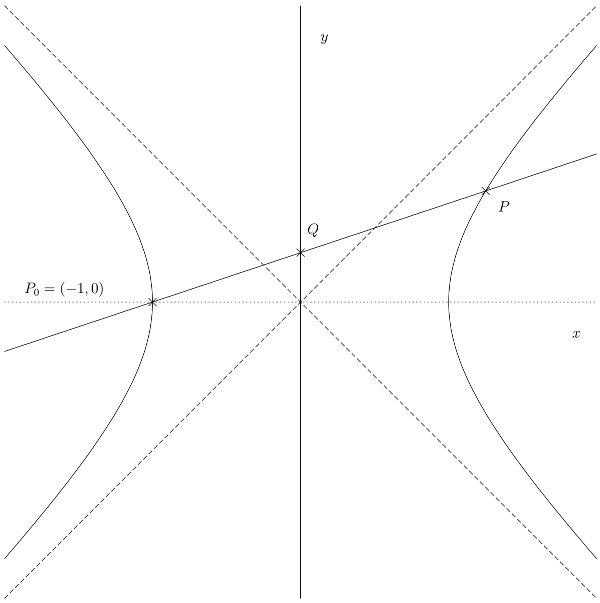
$$K = \left\{(x, y) \in \mathbb{R}^2 \mid y^2 = ax^2 + bx + c\right\}$$

und einen beliebigen Punkt $P_0 = (x_0, y_0) \in K$, sowie eine beliebige Gerade $g = \{Q_0 + tv | t \in \mathbb{R}\}$, die nicht durch P_0 läuft.
Der Kegelschnitt K ist eine Ellipse oder Hyperbel für $a \neq 0$ und eine Parabel im entarteten Fall $a = 0$.



Die Euler-Substitution $P \mapsto Q$ bzw. $x \mapsto t$ macht auch hier das Integral rational.

Beispiel 5.30 $\int R\left(x, \sqrt{x^2 - 1}\right)\, dx.$
Hier geht es um die durch die Gleichung $x^2 - y^2 = 1$ beschriebene Hyperbel:



Die Euler-Substitution $P \mapsto Q$ bzw. $]1, \infty[\ni x \to t \in]0, 1[$ mit

$$\begin{aligned} &P = (x, \sqrt{x^2 - 1}) = (x, y) \\ &\mapsto Q = (0, t) = \left(0, \frac{\sqrt{x^2 - 1}}{x + 1}\right) = \left(0, \sqrt{\frac{x - 1}{x + 1}}\right), \quad x > 1, \end{aligned}$$

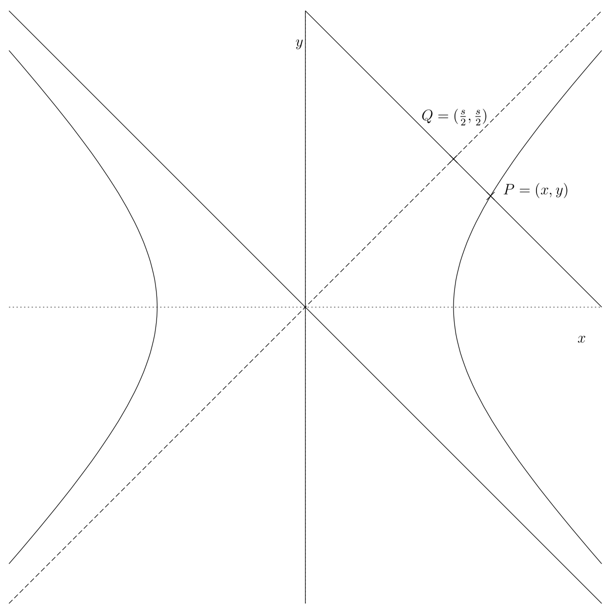
bildet den Hyperbelzweig im 1. Quadranten auf ein Stück der y-Achse ab.
Sie besitzt die folgende *rationale* Umkehrung:

$$y = \frac{2t}{1-t^2}, \quad x = \frac{1+t^2}{1-t^2}, \quad \frac{dx}{dt} = \frac{4t}{(1-t^2)^2}.$$

Damit erhalten wir folgende rationale Form des Integrals:

$$\int R\left(x, \sqrt{x^2-1}\right)\, dx = \int R\left(\frac{1+t^2}{1-t^2}, \frac{2t}{1-t^2}\right) \frac{4t}{(1-t^2)^2}\, dt$$

Variante: $P_0 =$ „unendlich ferner Punkt auf der Asymptoten $y = x$“:



Die Euler-Substitution

$$s = x + y = x + \sqrt{x^2-1} \qquad (x > 1, s > 1)$$

besitzt die Umkehrung

$$x = \frac{1}{2}\left(s + \frac{1}{s}\right), \quad y = \frac{1}{2}\left(s - \frac{1}{s}\right), \quad \frac{dx}{ds} = \frac{1}{2} - \frac{1}{2s^2}.$$

Damit erhalten wir folgende weitere rationale Form des Integrals:

$$\int R\left(x, \sqrt{x^2-1}\right)\, dx = \int R\left(\frac{1}{2}\left(s + \frac{1}{s}\right), \frac{1}{2}\left(s - \frac{1}{s}\right)\right)\left(\frac{1}{2} - \frac{1}{2s^2}\right)\, ds$$

Ausblick: Integrale vom Typ

$$\int R\left(x, \sqrt{p(x)}\right)\,dx$$

mit Polynomen p vom Grad 3,4 („*elliptische Integrale*“) oder höher („*hyperelliptische Integrale*“) lassen sich nur in Ausnahmefällen elementar ausdrücken. Sie spielen in der algebraischen Geometrie eine wichtige Rolle und treten z.B. bei der Berechnung der Bogenlänge einer Ellipse auf.

5.7 Vertauschung von Integral und Grenzwert

Es seien $a, b \in \mathbb{R},\ a < b$, und $f_n : [a,b] \to \mathbb{R},\ n \in \mathbb{N}_0$ eine Folge von Funktionen, sowie $f : [a,b] \to \mathbb{R}$.

Erinnerung an die unterschiedliche Quantorenstellung bei punktweiser und gleichmäßiger Konvergenz:

$f_n \overset{n\to\infty}{\longrightarrow} f$ punktweise
$\Leftrightarrow\ \forall x \in [a,b]\ \forall \varepsilon > 0\ \exists m \in \mathbb{N}_0\ \forall n > m : |f_n(x) - f(x)| < \varepsilon.$

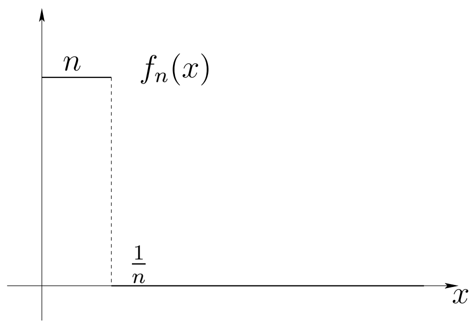
$f_n \overset{n\to\infty}{\longrightarrow}$ gleichmäßig
$\Leftrightarrow\ \forall \varepsilon > 0\ \exists m \in \mathbb{N}_0\ \forall n > m\ \forall x \in [a,b] : |f_n(x) - f(x)| < \varepsilon.$

Warnung: Aus $f_n \overset{n\to\infty}{\longrightarrow} f$ punktweise folgt i.a. *nicht* $\displaystyle\int_a^b f_n(x)\,dx \overset{n\to\infty}{\longrightarrow} \int_a^b f(x)\,dx$, selbst wenn alle Integrale existieren.

Beispiel 5.31 Es sei

$$f_n(x) = \begin{cases} n & \text{für } 0 < x < \frac{1}{n} \\ 0 & \text{sonst.} \end{cases}$$

Dann gilt $f_n \overset{n\to\infty}{\longrightarrow} 0$ punktweise, aber $\displaystyle\int_0^1 f_n(x)\,dx = 1 \neq 0 = \int_0^1 0\,dx.$



Es gilt aber:

Satz 5.32 (Vertauschbarkeit von Integral und gleichmäßigem Limes) Es seien $a, b \in \mathbb{R}$. Sind alle f_n Riemann-integrierbar und gilt $f_n \overset{n\to\infty}{\longrightarrow} f$ *gleichmäßig*, so ist auch f Riemann-integrierbar, und es gilt

$$\int_a^b f_n(x)\,dx \overset{n\to\infty}{\longrightarrow} \int_a^b f(x)\,dx.$$

Erinnerung: Sind die f_n sogar stetig, so ist auch die Grenzfunktion f stetig.

Beweis: Sei $\varepsilon > 0$. Wir wählen $m \in \mathbb{N}_0$ so groß, daß gilt:

$$\forall n > m\ \forall x \in [a, b] : |f_n(x) - f(x)| < \frac{\varepsilon}{4(b-a)}$$

Es sei $n > m$. Wegen der Riemann-Integrierbarkeit von f_n gibt es Treppenfunktionen $g_n, h_n \in \mathcal{T}[a, b]$ mit $g_n \leq f_n \leq h_n$ und

$$\int_a^b h_n(x)\,dx - \int_a^b g_n(x)\,dx < \frac{\varepsilon}{2}.$$

Wir setzen

$$h = h_n + \frac{\varepsilon}{4(b-a)}, \qquad g = g_n - \frac{\varepsilon}{4(b-a)}.$$

Es folgt:

$$g \leq f_n - \frac{\varepsilon}{4(b-a)} \leq f \leq f_n + \frac{\varepsilon}{4(b-a)} \leq h$$

und

$$\int_a^b h(x)\,dx - \int_a^b g(x)\,dx = \int_a^b h_n(x)\,dx - \int_a^b g_n(x)\,dx + 2\frac{\varepsilon}{4(b-a)}(b-a) \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Also ist f Riemann-integrierbar, und es gilt:

$$\left|\int_a^b f_n(x)\,dx - \int_a^b f(x)\,dx\right| \leq \frac{\varepsilon}{4(b-a)}(b-a) = \frac{\varepsilon}{4} \leq \varepsilon.$$

Es folgt

$$\int_a^b f_n(x)\,dx \overset{n\to\infty}{\longrightarrow} \int_a^b f(x)\,dx.$$

Die Stetigkeit der Grenzfunktion f bei stetigen f_n wurde früher gezeigt.

$\square$

Korollar 5.33 (Vertauschbarkeit von Grenzwert und Ableitung bei lokal gleichmäßiger Konvergenz der Ableitung) *Es sei $U \subseteq \mathbb{R}$ ein offenes Intervall und $F_n : U \to \mathbb{C}$ eine Folge stetig differenzierbarer Funktionen, so dass die F_n für $n \to \infty$ gleichmäßig gegen eine Funktion $F : U \to \mathbb{C}$ konvergieren und die Ableitungen F_n' gleichmäßig für $n \to \infty$ gleichmäßig gegen eine Funktion $f : U \to \mathbb{C}$ konvergieren Dann ist auch F stetig differenzierbar mit der Ableitung $F' = f$.*

Beweis: Wegen der vorausgesetzten gleichmäßigen Konvergenz und der Stetigkeit der Funktionen F_n und F_n' sind auch die Grenzfunktionen F und f stetig. Nach dem Hauptsatz gilt für alle $a, x \in U$:

$$\int_a^x F_n'(t)\,dt = F_n(x) - F_n(a)$$

Bilden wir den Limes $n \to \infty$, können wir wegen der vorausgesetzten gleichmäßigen Konvergenz von F_n' gegen f nach dem vohergehenden Satz Limes und Integral vertauschen:

$$\int_a^x f(t)\,dt = \int_a^x \lim_{n\to\infty} F_n'(t)\,dt = \lim_{n\to\infty} \int_a^x F_n'(t)\,dt = \lim_{n\to\infty} (F_n(x) - F_n(a)) = F(x) - F(a).$$

Nochmals mit dem Hauptsatz folgt durch Differenzieren nach x die Differenzierbarkeit von F mit der Ableitung $f(x) = F'(x)$, wie behauptet.

$\square$

6 Taylorapproximationen und Potenzreihen

6.1 Die Taylorformel

Im folgenden seien $x \in \mathbb{R}$, $U \subseteq \mathbb{R}$ eine offene Umgebung von x, $f : U \to \mathbb{C}$ eine Abbildung und $n \in \mathbb{N}_0$.

Definition 6.1 (höhere Ableitungen) Die n-te Ableitung $f^{(n)}(x) = \frac{d^n}{dx^n} f(x)$ von f an der Stelle $x \in U$ wird rekursiv definiert: $f^{(0)}(x) := f(x)$, $f^{(n+1)}(x) := \left(f^{(n)}\right)'(x)$, falls die Ableitungen existieren. f heißt n-mal differenzierbar in x, wenn $f^{(n)}(x)$ existiert. f heißt n-mal differenzierbar (in U), wenn f n-mal differenzierbar in allen $x \in U$ ist. Ist f n-mal differenzierbar und $f^{(n)}$ stetig, so heißt f n-mal stetig differenzierbar.

Wir wollen nun f nahe bei $x \in U$ durch ein Polynom approximieren.

Definition 6.2 (Taylorpolynom) Ist f n-mal differenzierbar in x, so heißt

$$T_{x,n}f : \mathbb{R} \to \mathbb{C}, \quad T_{x,n}f(y) = \sum_{k=0}^{n} \frac{f^{(k)}(x)}{k!}(y-x)^k$$

das n-te *Taylorpolynom* von f an der Stelle x. Für $y \in U$ heißt

$$R_{x,n}f(y) := f(y) - T_{x,n}f(y)$$

das *Restglied* der n-ten Taylorapproximation von f bei y an der (Entwicklungs-)Stelle x. Damit gilt die *Taylorformel*:

$$f(y) = \sum_{k=0}^{n} \frac{f^{(k)}(x)}{k!}(y-x)^k + R_{x,n}f(y)$$

Wir setzen auch noch formal $T_{x,-1}f = 0$, also $R_{x,-1}f = f$.
Natürlich ist die Taylorformel nur nützlich, wenn man etwas über das Restglied aussagen kann. Es gilt:

Lemma 6.3 *Ist f n-mal differenzierbar in x, so besitzen f und $T_{x,n}f$ die gleichen Ableitungen an der Stelle x bis zur n-ten Stufe:*

$$(T_{x,n}f)^{(j)}(x) = f^{(j)}(x) \text{ für } j = 0, \dots, n.$$

Durch Differenzbildung folgt:

$$(R_{x,n}f)^{(j)}(x) = 0 \text{ für } j = 0, \dots, n.$$

Beweis: Durch Induktion über $j = 0, \dots, n$ zeigen wir

$$(T_{x,n}f)^{(j)}(y) = \sum_{k=j}^{n} \frac{\prod_{l=k-j+1}^{k} l}{k!} f^{(k)}(x)(y-x)^{k-j}. \tag{75}$$

Für $j = 0$ folgt dies unmittelbar aus der Definition des Taylorpolynoms; man beachte, dass das leere Produkt den Wert 1 besitzt. Als Induktionsvoraussetzung nehmen wir die Behauptung (75) für ein gegebenes $j = 0, \dots, n-1$ und alle $y \in U$ an. Durch Ableiten folgt (der Summand zu $k = j$ fällt beim Ableiten weg):

$$\begin{aligned}(T_{x,n}f)^{(j+1)}(y) &= \sum_{k=j+1}^{n} \frac{\prod_{l=k-j+1}^{k} l}{k!} f^{(k)}(x)(k-j)(y-x)^{k-j-1} \\ &= \sum_{k=j+1}^{n} \frac{\prod_{l=k-j}^{k} l}{k!} f^{(k)}(x)(y-x)^{k-j-1},\end{aligned}$$

also die Behauptung (75) für $j+1$ statt j.
Speziell $y = x$ fallen in (75) alle Summanden ausser dem zu $k = j$ weg. Es folgt die Behauptung $(T_{x,n}f)^{(j)}(x) = f^{(j)}(x)$.

$\square$

äußerst wichtig!!!

Satz 6.4 (Restglieddarstellungen in der Taylorformel) Es seien U ein offenes Intervall, $x \in U$ und $f : U \to \mathbb{R}$. Dann gilt:

1. Ist f n-mal differenzierbar in x, wobei $n \in \mathbb{N}$, so folgt
$$R_{x,n}f(y) = o((y-x)^n) \text{ für } y \to x, \tag{76}$$
also
$$\frac{R_{x,n}f(y)}{(y-x)^n} \xrightarrow{y\to x} 0.$$

2. Ist f sogar $(n+1)$-mal differenzierbar in U, $n \in \mathbb{N}_0$, so gibt es für $y \in U \setminus \{x\}$ ein ξ echt zwischen x und y mit
$$R_{x,n}f(y) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(y-x)^{n+1}. \tag{77}$$
Ist zudem $f^{(n+1)}$ nahe bei x beschränkt, so folgt hieraus
$$R_{x,n}f(y) = O((y-x)^{n+1}) \text{ für } y \to x. \tag{78}$$

3. Ist f sogar $(n+1)$-mal *stetig* differenzierbar in U, $n \in \mathbb{N}_0$, so gilt für $y \in U$:
$$R_{x,n}f(y) = \int_x^y \frac{f^{(n+1)}(t)}{n!}(y-t)^n \, dt. \tag{79}$$

Bemerkung: Die Formel (79) wird *Lagrange-Darstellung* des Restglieds genannt. Die Restglieddarstellungen (76), (78) und (79) gelten ebenso, wenn f komplexe Werte annimmt. Dies folgt durch Zerlegung in Real- und Imaginärteil. Bei der Darstellung (77) kann jedoch die Zwischenstelle ξ für Real- und Imaginärteil verschieden sein.

Beweis des Satzes: Wir beweisen zuerst 2. Wir kürzen ab: $g = R_{x,n}f$. Insbesondere gilt $g^{(j)}(x) = 0$ für $j = 0, \ldots, n$ und $g^{(n+1)} = f^{(n+1)}$. Gegeben $y \in U$ mit $y \neq x$ zeigen wir nun durch Induktion über j: Für alle $j = 0, \ldots, n+1$ gibt es ein ξ_j zwischen x und y mit $\xi_j \neq x$ (sogar echt zwischen x und y für $j > 0$), so dass gilt:
$$\frac{g(y)}{(y-x)^{n+1}} = \frac{g^{(j)}(\xi_j)}{\left[\prod_{k=n+2-j}^{n+1} k\right](\xi_j - x)^{n+1-j}}. \tag{80}$$

Im Fall $j = n+1$ folgt hieraus die Behauptung 2.
Induktionsanfang: Für $j = 0$ wird die Behauptung (80) mit der Wahl $\xi_0 = y$ trivial.
Induktionsvoraussetzung: Gegeben sei $j \in \{1, \ldots, n\}$, und es gelte (80) für dieses j. Dann folgt mit dem verallgemeinerten Mittelwertsatz der Differentialrechnung: Es gibt ein ξ_{j+1}

echt zwischen x und ξ_j (und damit echt zwischen x und y) mit

$$\begin{aligned}
&g^{(j+1)}(\xi_{j+1})(\xi_j - x)^{n+1-j} \\
&= g^{(j+1)}(\xi_{j+1})[(\xi_j - x)^{n+1-j} - (x - x)^{n+1-j}] \\
&= [g^{(j)}(\xi_j) - g^{(j)}(x)] \left. \frac{d}{dz}(z - x)^{n+1-j} \right|_{z=\xi_{j+1}} \\
&= g^{(j)}(\xi_j)(n + 1 - j)(\xi_{j+1} - x)^{n-j}.
\end{aligned}$$

Es folgt unter Verwendung der Induktionsvoraussetzung:

$$\begin{aligned}
\frac{g(y)}{(y - x)^{n+1}} &= \frac{g^{(j)}(\xi_j)}{\left[\prod_{k=n+2-j}^{n+1} k\right] (\xi_j - x)^{n+1-j}} \\
&= \frac{g^{(j+1)}(\xi_j)}{\left[\prod_{k=n+2-j}^{n+1} k\right] (n + 1 - j)(\xi_{j+1} - x)^{n-j}} \\
&= \frac{g^{(j+1)}(\xi_j)}{\left[\prod_{k=n+1-j}^{n+1} k\right] (\xi_{j+1} - x)^{n-j}}.
\end{aligned}$$

Damit ist die Behauptung (80) auch für $j + 1$ statt j gezeigt.

Nun beweisen wir die Behauptung 1. Es sei f n-mal differenzierbar in x. Insbesondere ist f $(n - 1)$-mal differenzierbar in einer Umgebung von x. Wir setzen $g := R_{x,n}f$. Dann ist auch g $(n - 1)$-mal differenzierbar in einer Umgebung von x und n-mal differenzierbar in x mit $g^{(j)}(x) = 0$ für $j = 0, \dots, n$. Wegen der schon gezeigten Aussage 2., angewandt auf $n - 1$ statt n, folgt: Für $y \neq x$ genügend nahe bei x gibt es ein $\xi(y)$ echt zwischen x und y mit

$$g(y) = R_{x,n-1}g(y) = \frac{g^{(n-1)}(\xi(y))}{(n - 1)!}(y - x)^{n-1}.$$

Weil $g^{(n-1)}$ differenzierbar bei x mit Ableitung $g^{(n)}(x) = 0$ und Wert $g^{(n-1)}(x) = 0$ ist, folgt:

$$g^{(n-1)}(\xi(y)) = g^{(n-1)}(x)+g^{(n)}(x)(\xi(y)-x)+o(\xi(y)-x) = o(\xi(y)-x) = o(y-x) \text{ für } y \to x.$$

Oben eingesetzt folgt die Behauptung 1.:

$$g(y) = \frac{o(y - x)}{(n - 1)!}(y - x)^{n-1} = o((y - x)^n) \text{ für } y \to x.$$

Wir beweisen nun die Lagrange-Darstellung 3. des Restglieds durch vollständige Induktion über n:

$\boxed{n = 0}$ $\quad f(y) = f(x) + \int_x^y f'(x)\, dt$ $\quad$ gilt nach dem Hauptsatz.

$\boxed{n-1 \rightsquigarrow n}$ Es gelte die Induktionsvoraussetzung:

$$f(y) = \sum_{k=0}^{n-1} \frac{f^{(k)}(x)}{k!}(y-x)^k + \int_x^y \frac{f^{(n)}(k)}{(n-1)!}(y-t)^{n-1}\, dt.$$

Partielle Integration liefert:

$$\begin{aligned} \int_x^y \frac{f^{(n)}(t)}{(n-1)!}(y-t)^{n-1}\, dt &= \left[-\frac{f^{(n)}(t)}{n!}(y-t)^n \right]_{t=x}^{t=y} + \int_x^y \frac{f^{(n+1)}(t)}{n!}(y-t)^n\, dt \\ &= \frac{f^{(n)}(x)}{n!}(y-x)^n + \int_x^y \frac{f^{(n+1)}(t)}{n!}(y-t)^n\, dt \end{aligned}$$

Eingesetzt folgt die Behauptung 3. für n.

$\square$

Bemerkungen:

1. Aus der Lagrange-Darstellung (79) des Restglieds folgt die Darstellung (77) auch alternativ mit dem allgemeinen Mittelwertsatz der Integralrechnung:
   Für ein ξ zwischen x und y gilt:

   $$\int_x^y \frac{f^{(n+1)}(t)}{n!}(y-t)^n\, dt = \frac{f^{(n+1)}(\xi)}{n!}\int_x^y (y-t)^n\, dt = \frac{f^{(n+1)}(\xi)}{(n+1)!}(y-x)^{n+1}.$$

   Man beachte zur Anwendbarkeit des Mittelwertsatzes, dass der Integrand $(y-t)^n$ für t zwischen x und y ein einheitliches Vorzeichen besitzt: entweder er ist nur nichtnegativ oder nur nichtpositiv. Allerdings setzt dieser alternative Beweis der Darstellung (77) die Stetigkeit der $(n+1)$-ten Ableitung voraus.

2. Die Restglieddarstellung 1. im Satz 6.4 verallgemeinert die Definition der Differenzierbarkeit, denn sie besagt für $n = 1$:

   $$f(y) = f(x) + f'(x)(y-x) + o(y-x), \quad y \to x.$$

   Ebenso verallgemeinert die Restglieddarstellung 2. im Satz 6.4 den Mittelwertsatz der Differentialrechnung, denn sie besagt für $n = 0$:

   $$f(y) - f(x) = f'(\xi)(y-x)$$

   für ein ξ echt zwischen x und y.
   Schließlich verallgemeinert die Restglieddarstellung 3. im Satz 6.4 den Hauptsatz, denn sie besagt für $n = 0$:

   $$f(y) - f(x) = \int_x^y f'(t)\, dt.$$

   Die Restglieddarstellungen umfassen also drei der wichtigsten Formeln der Differential- und Integralrechnung.

Beispiel 6.5 *Die geometrische Reihe und die Exponentialreihe sind Beispiele zur Taylorformel:*

$$\frac{1}{1-y} = \sum_{k=0}^{n} y^k + O(y^{n+1}), \quad y \to 0$$

und

$$e^y = \sum_{k=0}^{n} \frac{y^k}{k!} + O(y^{n+1}), \quad y \to 0$$

Beispiel 6.6 Für $f(x) = \arctan x$ erhalten wir

$$f'(x) = \frac{1}{1+x^2} \tag{81}$$

$$f''(x) = -\frac{2x}{(1+x^2)^2} \tag{82}$$

$$f'''(x) = \frac{-2}{(1+x^2)^2} + \frac{8x^2}{(1+x^2)^3} \tag{83}$$

also $f(0) = 0$, $f'(0) = 1$, $f''(0) = 0$ und $f'''(0) = -2$.
Es folgt:

$$\arctan x = 0 + \frac{1}{1!}x + \frac{0}{2!}x^2 - \frac{2}{3!}x^3 + o(x^3) = x - \frac{x^3}{3} + o(x^3) \text{ für } x \to 0.$$

Manchmal ist es einfacher, die Taylorapproximation durch Einsetzen der Taylorapproximationen der Bausteine als durch direktes Ableiten zu bestimmen:

Beispiel 6.7 Man bestimme die Taylorapproximation von $f(x) = \dfrac{1}{1+\frac{1}{2}\sinh x}$ für $x \to 0$ mit einem Fehlerterm $o(x^2)$.

1. Lösung: Es gilt:

$$\frac{d}{dx}\frac{1}{1+\frac{1}{2}\sinh x} = -\frac{\frac{1}{2}\cosh x}{(1+\frac{1}{2}\sinh x)^2}, \tag{84}$$

$$\frac{d^2}{dx^2}\frac{1}{1+\frac{1}{2}\sinh x} = -\frac{\frac{1}{2}\sinh x}{(1+\frac{1}{2}\sinh x)^2} + \frac{\frac{1}{2}\cosh^2 x}{(1+\frac{1}{2}\sinh x)^3}, \tag{85}$$

also $f(0) = 1$, $f'(0) = -\frac{1}{2}$, $f''(0) = \frac{1}{2}$ und damit

$$f(x) = 1 - \frac{1}{2}x + \frac{1}{4}x^2 + o(x^2) \quad \text{für } x \to 0.$$

2. Lösung: Wir setzen

$$y = \frac{1}{2}\sinh x = \frac{1}{2}(x + o(x^2)) \quad \text{für } x \to 0.$$

Insbesondere gilt $y \to 0$ für $x \to 0$. Mit der geometrischen Reihe erhalten wir:

$$\frac{1}{1+y} = 1 - y + y^2 + o(y^2) \quad \text{für } y \to 0.$$

Eingesetzt:

$$\begin{aligned} f(x) &= 1 - \frac{1}{2}(x + o(x^2)) + \left(\frac{1}{2}(x + o(x^2))\right)^2 + o((x + o(x^2))^2) & (86) \\ &= 1 - \frac{1}{2}x + \frac{1}{4}x^2 + o(x^2) \quad \text{für } x \to 0. & (87) \end{aligned}$$

6.2 Ableitung von Potenzreihen

Im Hinblick auf Anwendungen mit Taylorreihen $\sum_{n=0}^{\infty} \frac{f^{(n)}(x)}{n!} h^n$ studieren wir in diesem Abschnitt die Ableitung von Potenzreihen. Es gilt:

Satz 6.8 (Gliedweise Differentiation von Potenzreihen) *Es seien $(a_n)_{n \in \mathbb{N}_0}$ eine Folge in $\mathbb{C}$,*

$$f(x) = \sum_{n=0}^{\infty} a_n x^n$$

eine Potenzreihe und

$$g(x) = \sum_{n=1}^{\infty} n a_n x^{n-1}$$

die gliedweise abgeleitete Reihe.
Dann haben f und g den gleichen Konvergenzradius R, und es gilt für $|x| < R$:

$$\frac{d}{dx} f(x) = g(x).$$

Beweis: Wenn die Reihe für g absolut konvergiert, so auch die Reihe für f, denn

$$\sum_{n=1}^{\infty} |a_n n x^{n-1}| \cdot |x|$$

majorisiert

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

Also ist der Konvergenzradius von f mindestens so groß wie derjenige von g.
Umgekehrt: Wenn $a_n x^n = O(e^{-\varepsilon n})$ für $n \to \infty$ mit einem $\varepsilon > 0$, d.h. wenn die Reihe für f durch eine geometrische Reihe majorisiert wird, so folgt $n a_n x^{n-1} = O(e^{-\delta n})$ für $n \to \infty$ für alle $\delta \in]0, \varepsilon[$ wegen $n e^{-\varepsilon n} = O(e^{-\delta n})$.
Also ist der Konvergenzradius von g mindestens so groß wie der von f.

Es sei nun $|x| < R$. Wir wählen $r \in]x, R[$. Es genügt für jede gegen x konvergente Folge $(x_k)_{k\in\mathbb{N}_0}$ in $U_r(0)$ zu zeigen:

$$\frac{f(x_k) - f(x)}{x_k - x} \stackrel{k\to\infty}{\longrightarrow} g(x).$$

Es gilt

$$\begin{aligned}
\frac{f(x_k) - f(x)}{x_k - x} &= \sum_{n=1}^{\infty} a_n \frac{x_k^n - x^n}{x_k - x} && (88)\\
&= \sum_{n=1}^{\infty} a_n \sum_{l=0}^{n-1} x_k^l x^{n-1-l} && (89)
\end{aligned}$$

nach der geometrischen Summe.
Wir überprüfen nun die Voraussetzungen des Satzes von der dominierten Konvergenz:

$$\lim_{k\to\infty} a_n \sum_{l=0}^{n-1} x_k^l x^{n-1-l} = a_n n x^{n-1},$$

und

$$\left| a_n \sum_{l=0}^{n-1} x_k^l x^{n-1-l} \right| \leq |a_n| n r^{n-1}.$$

Die Reihe $\sum_{n=1}^{\infty} |a_n| n r^{n-1}$ konvergiert wegen $r < R$ absolut.
Mit dem Satz von der dominierten Konvergenz folgt:

$$\begin{aligned}
\lim_{k\to\infty} \frac{f(x_k) - f(x)}{x_k - x} &= \sum_{n=1}^{\infty} \lim_{k\to\infty} a_n \sum_{l=0}^{n-1} x_k^l x^{n-1-l} && (90)\\
&= \sum_{n=1}^{\infty} a_n n x^{n-1}, && (91)
\end{aligned}$$

also die Behauptung.

$\square$

Beispiel 6.9 *: Für die Exponentialreihe gilt:*

$$\frac{d}{dx}\exp(x) = \frac{d}{dx}\sum_{n=0}^{\infty}\frac{x^n}{n!} = \sum_{n=1}^{\infty}\frac{d}{dx}\frac{x^n}{n!} = \sum_{n=1}^{\infty}\frac{x^{n-1}}{(n-1)!} = \exp(x).$$

Das liefert einen neuen Beweis von $\exp' = \exp$.

6.3 Beispiele für Potenzreihen

6.3.1 Die Logarithmusreihen und die Arcustangensreihe

Beispiel 6.10 (Logarithmusreihen) *Es gilt*

$$\boxed{\sum_{n=1}^{\infty} \frac{x^n}{n} = -\log(1-x) \quad \textit{für } |x| < 1} \tag{92}$$

Beweis: Beide Seiten haben die gleiche Ableitung:

$$\frac{d}{dx}\sum_{n=1}^{\infty} \frac{x^n}{n} = \sum_{n=1}^{\infty} \frac{d}{dx}\frac{x^n}{n} = \sum_{n=1}^{\infty} x^{n-1} \tag{93}$$

$$= \frac{1}{1-x} = \frac{d}{dx}(-\log(1-x)). \tag{94}$$

Also ist

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

konstant, denn

$$\frac{d}{dx}\left(\sum_{n=1}^{\infty} \frac{x^n}{n} + \log(1-x)\right) = 0.$$

Für $x = 0$ erhalten wir den Wert 0; also folgt die Behauptung.

□

Durch die Substitution $x \rightsquigarrow -x$ erhalten wir

$$\boxed{\sum_{n=1}^{\infty} \frac{(-1)^n}{n} x^n = -\log(1+x) \quad \text{für } |x| < 1} \tag{95}$$

Durch Subtraktion der Formeln (92) und (95) schließen wir:

$$\boxed{2\sum_{k=0}^{\infty} \frac{x^{2k+1}}{2k+1} = \log(1+x) - \log(1-x) = \log\frac{1+x}{1-x} \quad \text{für } |x| < 1}$$

Für x nahe bei 0 eignen sich diese Formeln auch für numerische Berechnungen des Logarithmus.

Beispiel 6.11 (Alternierende harmonische Reihe) Für $0 < x \leq 1$ ist $\frac{x^{2k-1}}{2k-1} - \frac{x^{2k}}{2k} > 0$ monoton steigend in x, da

$$\frac{d}{dx}\left(\frac{x^{2k-1}}{2k-1} - \frac{x^{2k}}{2k}\right) = x^{2k-2} - x^{2k-1} = x^{2k-2}(1-x) \geq 0.$$

Es folgt mit dem Satz von der monotonen Konvergenz und der Stetigkeit des Logarithmus:

$$\begin{aligned}
\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} &= \sum_{k=1}^{\infty} \left(\frac{1}{2k-1} - \frac{1}{2k} \right) \\
&= \sum_{k=1}^{\infty} \lim_{x \uparrow 1} \left(\frac{x^{2k-1}}{2k-1} - \frac{x^{2k}}{2k} \right) \\
&= \lim_{x \uparrow 1} \sum_{k=1}^{\infty} \left(\frac{x^{2k-1}}{2k-1} - \frac{x^{2k}}{2k} \right) \\
&= -\lim_{x \uparrow 1} \sum_{n=1}^{\infty} \frac{(-x)^n}{n} \\
&= \lim_{x \uparrow 1} \log(1+x) \\
&= \log 2.
\end{aligned}$$

Fassen wir zusammen:

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

Zur effizienten numerischen Berechnung von $\log 2$ ist diese Reihe ungeeignet, da sie zu langsam konvergiert.

Beispiel 6.12 (Arcustangensreihe) Es gilt:

$$\boxed{\arctan x = \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} x^{2n+1} \quad \text{für } |x| < 1}$$

Beweis: Für $x = 0$ haben beide Seiten den Wert 0, und es gilt für $|x| < 1$:

$$\frac{d}{dx} \arctan x = \frac{1}{1+x^2} = \sum_{n=0}^{\infty} (-1)^n x^{2n} = \frac{d}{dx} \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{2n+1}$$

aufgrund der geometrischen Reihe.

□

Im Limes $x \uparrow 1$ erhalten wir mit der Stetigkeit des Arcustangens und dem Satz von der monotonen Konvergenz:

$$\begin{aligned}
\frac{\pi}{4} &= \arctan 1 \\
&= \lim_{x \uparrow 1} \arctan x \\
&= \lim_{x \uparrow 1} \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} x^{2n+1} \\
&= \lim_{x \uparrow 1} \sum_{k=0}^{\infty} \left(\frac{1}{4k+1} x^{4k+1} - \frac{1}{4k+3} x^{4k+3} \right) \\
&= \sum_{k=0}^{\infty} \lim_{x \uparrow 1} \left(\frac{1}{4k+1} x^{4k+1} - \frac{1}{4k+3} x^{4k+3} \right) \\
&= \sum_{k=0}^{\infty} \left(\frac{1}{4k+1} - \frac{1}{4k+3} \right) \\
&= \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1}.
\end{aligned}$$

Man beachte bei der Vertauschung von Limes und Reihe, dass $\frac{1}{4k+1} x^{4k+1} - \frac{1}{4k+3} x^{4k+3} \geq 0$ für $0 \leq x \leq 1$, $k \in \mathbb{N}_0$, in x monoton steigt, denn

$$\frac{d}{dx} \left(\frac{1}{4k+1} x^{4k+1} - \frac{1}{4k+3} x^{4k+3} \right) = x^{4k} - x^{4k+2} \geq 0$$

Fassen wir zusammen:

Leibniz-Reihe für $\frac{\pi}{4}$:

$$\boxed{\frac{1}{1} - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \frac{1}{9} \pm \ldots = \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} = \frac{\pi}{4}}$$

Zur effizienten numerischen Berechnung von π ist diese Reihe ungeeignet, da auch sie zu langsam konvergiert.

6.3.2 Die binomische Reihe und die Arcussinusreihe

Satz 6.13 (binomische Reihe) *Für $s \in \mathbb{C}$ und $x \in \mathbb{R}$ gilt:*

neben geometrischer Reihe und Exponentialreihe die drittwichtigste Reihe überhaupt!

$$\boxed{(1+x)^s = \sum_{n=0}^{\infty} \binom{s}{n} x^n, \quad \textit{falls } |x| < 1}$$

Hierbei wird der Binomialkoeffizient durch

$$\binom{s}{n} = \frac{1}{n!} \prod_{k=0}^{n-1} (s-k)$$

gegeben.

Bemerkungen:

1. Die binomische Reihe ergibt die binomische Formel im Spezialfall $s \in \mathbb{N}_0$.

   Im Spezialfall $s = -1$ erhält man die geometrische Reihe $(1+x)^{-1} = \sum_{n=0}^{\infty} (-x)^n$ für $|x| < 1$, denn $\binom{-1}{n} = (-1)^n$.

2. Für $z \in \mathbb{C} \setminus \mathbb{R}_0^-$ sei $\operatorname{Log} z$ die eindeutig bestimmte Zahl $w \in \mathbb{R} + i]-\pi, \pi[$, für die $e^w = z$ gilt. Die Funktion $\operatorname{Log} : \mathbb{C} \setminus \mathbb{R}_0^- \to \mathbb{R} + i]-\pi, \pi[$ wird *Hauptzweig* des Logarithmus genannt. Andere Zweige des Logarithmus im Komplexen erhält man durch Addition ganzzahliger Vielfacher von $2\pi i$. Für $z \in \mathbb{C} \setminus \mathbb{R}_0^-$ und $s \in \mathbb{C}$ wird

   $$z^s := e^{s \operatorname{Log} z}$$

   der Hauptzweig der Potenz genannt. Bei dieser Definition von $\operatorname{Log}(1+x)$ und von $(1+x)^s$ als Hauptzweig für $x \in \mathbb{C}$ mit $|x| < 1$ gelten die Logarithmusreihen und die binomische Reihe auch für komplexe x.

3. Die binomische Reihe ist die Taylorreihe von $f(x) = (1+x)^s$ an der Stelle 0, denn es gilt

   $$f^{(n)}(x) = s(s-1) \cdot \ldots \cdot (s-n+1)(1+x)^{s-n} = n! \binom{s}{n} (1+x)^{s-n}$$

   und daher $\frac{f^{(n)}(0)}{n!} = \binom{s}{n}$.

Ein möglicher Beweis der binomischen Reihe verwendet die Lagrange-Darstellung des Restglieds in der Taylorformel. Wir verwenden hier einen alternativen Beweis mit einer Differentialgleichung:

Beweis der binomischen Reihe:
Wir zeigen zunächst mit dem Quotientenkriterium, daß die Reihe für $|x| < 1$ konvergiert. Wir dürfen $x \neq 0$ annehmen:

$$\left| \frac{\binom{s}{n} x^n}{\binom{s}{n-1} x^{n-1}} \right| = \left| \frac{s-n+1}{n} x \right| \overset{n\to\infty}{\longrightarrow} |x| < 1.$$

Als nächstes zeigen wir, daß

$$y(x) = \sum_{k=0}^{\infty} \binom{s}{k} x^k \quad \text{für } |x| < 1 \tag{96}$$

die Differentialgleichung

$$(1+x)y' = sy \tag{97}$$

erfüllt:

$$(1+x)y'(x) = (1+x)\frac{d}{dx}\sum_{k=0}^{\infty}\binom{s}{k}x^k = (1+x)\sum_{k=1}^{\infty}\binom{s}{k}kx^{k-1}$$ (gliedweise Ableitung einer Potenzreihe)

$$= \sum_{k=1}^{\infty}\binom{s}{k}kx^{k-1} + \sum_{k=1}^{\infty}\binom{s}{k}kx^k$$

$$= \sum_{k=0}^{\infty}\binom{s}{k+1}(k+1)x^k + \sum_{k=0}^{\infty}\binom{s}{k}kx^k$$ (Indexshift um 1 in linker Summe, Summand zu $k=0$ in rechter Summe verschwindet)

$$= \sum_{k=0}^{\infty}\binom{s}{k}(s-k)x^k + \sum_{k=0}^{\infty}\binom{s}{k}kx^k$$ (weil $\binom{s}{k+1}(k+1) = \frac{1}{k!}\prod_{j=0}^{k}(s-j) = (s-k)\binom{s}{k}$ gilt)

$$= \sum_{k=0}^{\infty}\binom{s}{k}((s-k)+k)x^k = \sum_{k=0}^{\infty}\binom{s}{k}sx^k$$

$$= s\sum_{k=0}^{\infty}\binom{s}{k}x^k = sy(x)$$

Wir zeigen nun, dass der Quotient von $y(x)$ mit $(1+x)^s$ konstant ist, indem wir die Ableitung des Quotienten bilden: Es folgt

$$\frac{d}{dx}\frac{y(x)}{(1+x)^{-s}} = \frac{d}{dx}[y(x)(1+x)^{-s}] \; = \; y'(x)(1+x)^{-s} - y(x)s(1+x)^{-s-1} = 0$$

aufgrund der Differentialgleichung (97), also ist $y(x)(1+x)^{-s}$ konstant. Für $x=0$ erhalten wir den Wert $y(0) = \binom{s}{0} = 1$. Es folgt die Behauptung: $y(x)(1+x)^{-s} = 1$, also $y(x) = (1+x)^s$.

□

Als Folgerung erhalten wir

Satz 6.14 (Arcussinusreihe) *Es gilt*

$$\arcsin x = \sum_{n=0}^{\infty}\frac{(-1)^n}{2n+1}\binom{-\frac{1}{2}}{n}x^{2n+1} \quad \textit{für } |x|<1$$

oder anders geschrieben:

$$\boxed{\arcsin x = \sum_{n=0}^{\infty}\binom{n-\frac{1}{2}}{n}\frac{x^{2n+1}}{2n+1} \quad \textit{für } |x|<1}$$

Beweis: Beide Seiten haben den Wert 0 für $x = 0$ sowie die gleiche Ableitung:

$$\begin{aligned}\frac{d}{dx}\arcsin x &= (1-x^2)^{-\frac{1}{2}} \\ &= \sum_{n=0}^{\infty}\binom{-\frac{1}{2}}{n}(-x^2)^n \\ &= \frac{d}{dx}\sum_{n=0}^{\infty}(-1)^n\binom{-\frac{1}{2}}{n}\frac{x^{2n+1}}{2n+1}\end{aligned}$$

Für die zweite Darstellung des Arcussinus beachte man

$$(-1)^n\binom{-\frac{1}{2}}{n} = \frac{1}{n!}\prod_{k=0}^{n-1}\left(\frac{1}{2}+k\right) = \frac{1}{n!}\prod_{l=0}^{n-1}\left(n-\frac{1}{2}-l\right) = \binom{n-\frac{1}{2}}{n}.$$

□

Zum Beispiel gilt wegen $\arcsin\frac{1}{2} = \frac{\pi}{6}$:

$$\boxed{\pi = 3\sum_{n=0}^{\infty}\binom{n-\frac{1}{2}}{n}\frac{1}{4^n(2n+1)}}$$

Mit dieser Formel kann man mit kleinem Aufwand einige Nachkommastellen von π berechnen.

6.4 Approximation des Absolutbetrags durch Polynome

Als eine Anwendung zeigen wir, dass der Absolutbetrag auf $[-1, 1]$ gleichmäßig durch Polynome approximiert werden kann.
Hierzu beobachten wir: Die folgende Binomialreihe für die Quadratwurzel

$$1-\sqrt{1-x} = \sum_{n=1}^{\infty}\binom{\frac{1}{2}}{n}(-1)^{n+1}x^n$$

hat positive Koeffizienten:

$$\binom{\frac{1}{2}}{n}(-1)^{n+1} = \frac{1}{2\cdot n!}\prod_{k=1}^{n-1}\left(k-\frac{1}{2}\right) > 0, \quad n\in\mathbb{N}.$$

Insbesondere steigt $\binom{\frac{1}{2}}{n}(-1)^{n+1}x^n$ für $0 \le x \le 1$ monoton in x. Im Limes $x \uparrow 1$ erhalten wir mit dem Satz von der monotonen Konvergenz:

$$\begin{aligned}
1 &= \lim_{x\uparrow 1}(1-\sqrt{1-x}) \\
&= \lim_{x\uparrow 1}\sum_{n=1}^{\infty}\binom{\frac{1}{2}}{n}(-1)^{n+1}x^n \\
&= \sum_{n=1}^{\infty}\binom{\frac{1}{2}}{n}(-1)^{n+1}\lim_{x\uparrow 1}x^n \\
&= \sum_{n=1}^{\infty}\binom{\frac{1}{2}}{n}(-1)^{n+1}
\end{aligned}$$

und daher für $k \in \mathbb{N}_0$ und $0 \le x \le 1$:

$$0 \le \sum_{n=0}^{k}\binom{\frac{1}{2}}{n}(-x)^n - \sqrt{1-x} = \sum_{n=k+1}^{\infty}\binom{\frac{1}{2}}{n}(-1)^{n+1}x^n \le \sum_{n=k+1}^{\infty}\binom{\frac{1}{2}}{n}(-1)^{n+1} \overset{k\to\infty}{\longrightarrow} 0.$$

Damit ist gezeigt:

$$\sup_{0\le x\le 1}\left|\sqrt{1-x}-\sum_{n=0}^{k}\binom{\frac{1}{2}}{n}(-x)^n\right| = \sup_{0\le x\le 1}\sum_{n=k+1}^{\infty}\binom{\frac{1}{2}}{n}(-1)^{n+1}x^n = \sum_{n=k+1}^{\infty}\binom{\frac{1}{2}}{n}(-1)^{n+1} \overset{k\to\infty}{\longrightarrow} 0.$$

Setzen wir hier $x = 1 - y^2$ mit $-1 \le y \le 1$ ein, also $|y| = \sqrt{y^2} = \sqrt{1-x}$, so erhalten wir folgende gleichmäßige Approximation des Absolutbetrags auf $[-1, 1]$ durch Polynome:

$$\boxed{\sup_{-1\le y\le 1}\left||y|-\sum_{n=0}^{k}\binom{\frac{1}{2}}{n}(y^2-1)^n\right| \overset{k\to\infty}{\longrightarrow} 0.}$$

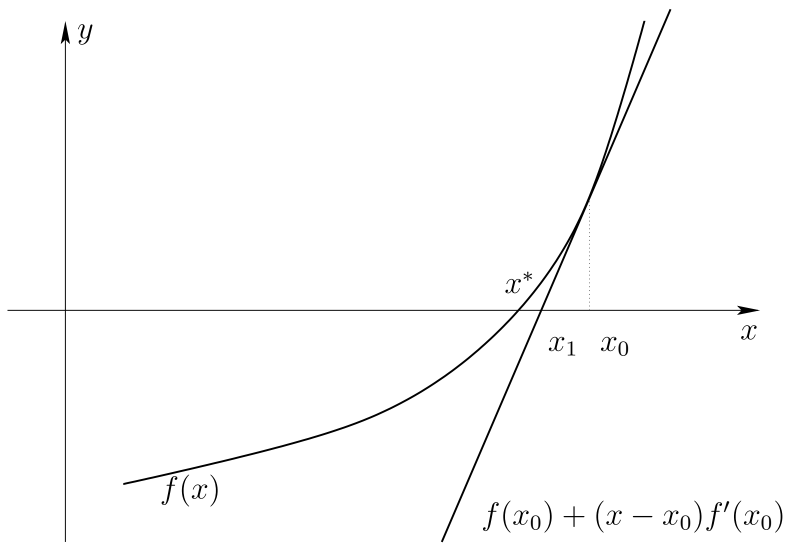
Mit Hilfe dieser Formel werden wir in der Analysis 2 beliebige stetige Funktionen auf einem kompakten Intervall gleichmäßig durch Polynome approximieren (Weierstraßscher Approximationssatz).

6.5 Das Newtonverfahren

Das Newtonverfahren ist ein numerisches Verfahren zur näherungsweisen Lösung einer Gleichung $f(x) = 0$. Ausgehend von einer Startnäherung x_0 (genügend nahe bei der gesuchten Lösung x^*) findet man eine bessere Näherung x_1 an x^*. Das Prinzip ist einfach: Statt der (nichtlinearen) Gleichung $f(x^*) = 0$ löst man die lineare Approximation $f(x_0) + (x_1 - x_0)f'(x_0) = 0$, also

$$\boxed{x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}}.$$

Das ist ein Schritt des Newtonverfahrens.



Lemma 6.15 *Es sei I ein Intervall, $f\colon I \to \mathbb{R}$ zweimal differenzierbar, f' von 0 weg beschränkt: $c_1 := \inf_{x\in I} |f'(x)| > 0$ und f'' beschränkt: $c_2 := \sup_{x\in I} |f''(x)| < +\infty$. Es sei $x^* \in I$ eine Nullstelle von f, also $f(x^*) = 0$. Weiter sei $x_0 \in I$ und $x_1 = x_0 - f(x_0)/f'(x_0)$. Dann gilt:*

$$|x_1 - x^*| \le \frac{c_2}{2c_1}|x_0 - x^*|^2.$$

Beweis: Nach der Taylorformel gilt für ein ξ zwischen x_0 und x^*:

$$0 = f(x^*) = f(x_0) + f'(x_0)(x^* - x_0) + \frac{1}{2}f''(\xi)(x^* - x_0)^2. \tag{98}$$

Nun ist x_1 wegen $x_1 = x_0 - f(x_0)/f'(x_0)$ eine Nullstelle der Tangente an den Graphen von f im Punkt $(x_0, f(x_0))$. In Formeln:

$$f(x_0) + f'(x_0)(x_1 - x_0) = 0,$$

anders geschrieben

$$f(x_0) + f'(x_0)(x^* - x_0) = f'(x_0)(x^* - x_1).$$

In (98) eingesetzt erhalten wir:

$$0 = f'(x_0)(x^* - x_1) + \frac{1}{2}f''(\xi)(x^* - x_0)^2.$$

Das bedeutet:

$$x_1 - x^* = \frac{f''(\xi)}{2f'(x_0)}(x_0 - x^*)^2.$$

Wegen $|f'(x_0)| \ge c_1$ und $|f''(\xi)| \le c_2$ folgt die Behauptung.

□

Betrachten wir nun das iterierte Verfahren:

Rekursionsschritt des Newtonverfahrens

$$\boxed{x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, n \in \mathbb{N}_0.}$$

Mit der Abkürzung $c := \frac{c_2}{2c_1}$ erhalten wir:

Satz 6.16 *Es sei f mindestens in $U_\varepsilon(x^*)$ definiert und zweimal differenzierbar. Es gelte $f(x^*) = 0$ und $c\varepsilon < 1$. Dann konvergiert das Newtonverfahren für alle Startwerte $x_0 \in U_\varepsilon(x^*)$ gegen x^*, und es gilt die Fehlerabschätzung*

$$\forall n \in \mathbb{N}\colon c|x_n - x^*| \leq (c|x_0 - x^*|)^{2^n}.$$

Beweis: Dies folgt unmittelbar durch Iteration aus dem Lemma. Man beachte, daß die Voraussetzung $c|x_0 - x^*| \leq c\varepsilon < 1$ garantiert, daß $(c|x_0 - x^*|)^{2^n}$ superexponentiell schnell für $n \to \infty$ gegen 0 konvergiert.

"Daumenregel" zur Fehlerabschätzung im Newton-Verfahren.
Liegt der Startwert x_0 genügend nahe bei x^*, so wird in jedem Schritt des Newtonverfahrens die Anzahl der korrekten Dezimalstellen ungefähr verdoppelt.

Beispiel 6.17 *Für $f(x) = x^2 - a$, also die Gleichung $x^2 - a = 0$, erhalten wir*

$$x_{n+1} = x_n - \frac{x_n^2 - a}{2x_n} = \frac{1}{2}\left(x_n + \frac{a}{x_n}\right),$$

Dies ist das Heronverfahren zur Bestimmung von Quadratwurzeln.

Literatur

[Bro05] Manuel Bronstein. *Symbolic integration. I*, volume 1 of *Algorithms and Computation in Mathematics*. Springer-Verlag, Berlin, second edition, 2005. Transcendental functions, With a foreword by B. F. Caviness.

[EHH+83] H.-D. Ebbinghaus, H. Hermes, F. Hirzebruch, M. Koecher, K. Mainzer, A. Prestel, and R. Remmert. *Zahlen*, volume 1 of *Grundwissen Mathematik [Basic Knowledge in Mathematics]*. Springer-Verlag, Berlin, 1983. Edited and with an introduction by K. Lamotke.

Index