Skriptum zur Vorlesung

Analysis I

Patrizio Neff

nach einer Vorlage von Hans-Dieter Alber, TU Darmstadt

Inhaltsverzeichnis

Literatur:

M. Barner, F. Flohr: Analysis I. Berlin: de Gruyter 1987

O. Forster: Analysis I. Vieweg Verlag 2005

1 Grundbegriffe

a.) Mengen

Grundlegend ist der von Georg Cantor (1845 – 1918) eingeführte Begriff der Menge. Nach der Definition von Cantor ist eine Menge eine

„Zusammenfassung von wohlbestimmten und wohlunterschiedenen Dingen zu einem Ganzen.“

Diese Dinge werden Elemente der Menge genannt. Mengen kann man auf zwei Arten festlegen:

1.) Man führt alle ihre Elemente in geschweiften Klammern auf: z. Bsp.

$$\begin{aligned} M_1 &= \{1, 2, 3, 4\} \\ M_2 &= \{-1, 2, 4, 6, 7.5\} \\ M_3 &= \{2, 4, 6, 8, 10\}\,. \end{aligned}$$

2.) Oder man gibt eine definierende Eigenschaft an:

$$\begin{aligned} M_4 &= \{x \mid x \text{ ist gerade Zahl zwischen 1 und 10}\} \\ M_5 &= \{a \mid a^2 = 1\} \\ M_6 &= \{x \mid x \text{ ist Primzahl}\}\,. \end{aligned}$$

Es kommt nicht auf die Reihenfolge der Elemente an. Also ist

$$M_1 = \{1, 2, 3, 4\} = \{4, 3, 1, 2\}\,.$$

Die Definition der Menge muß so gefaßt sein, daß von jedem Ding feststeht, ob es zu der Menge gehört oder nicht (jedes Element muß „wohlbestimmt“ sein). In einer Menge darf ein Element nur einmal vorkommen („wohlunterschieden“).

Man sagt, zwei Mengen A und B seien gleich,

$$A = B\,,$$

wenn sie dieselben Elemente enthalten. Also ist

$$M_1 = \{2, 4, 6, 8, 10\} = M_4 = \{x \mid x \text{ ist gerade Zahl zwischen 1 und 10}\}\,.$$

Man sagt, A sei Teilmenge von B,

$$A \subseteq B\,,$$

wenn jedes Element von A auch in B enthalten ist:

$$\begin{aligned} \{1,4\} &\subseteq \{1,2,3,4\}, \\ \{1,2,3,4\} &\subseteq \{a \mid a \text{ ist ganze Zahl}\}\,. \end{aligned}$$

B ist dann Obermenge von A. Für die Aussage, „m ist Element von M“ führt man das Symbol

$$m \in M$$

ein. Die Negation dieser Aussage ist

$$m \notin M\,.$$

Dieses Symbol bedeutet also „m ist nicht Element von M“.

Eine Menge, die kein Element enthält, wird leere Menge genannt. Das Symbol für diese Menge ist $\emptyset$. Es ist sinnvoll, die leere Menge zuzulassen, weil sonst in vielen Aussagen Fallunterscheidungen vorgenommen werden müssen. Die leere Menge $\emptyset$ wird als Teilmenge jeder Menge angesehen.

Für eine Reihe von Mengen hat man Standardbezeichnungen eingeführt:

$$\begin{aligned} \mathbb{N} &= \{1,2,3,\ldots\} \quad \text{(Menge der natürlichen Zahlen)}, \\ \mathbb{Z} &= \{0,1,-1,2,-2,\ldots\} \quad \text{(Menge der ganzen Zahlen)}, \\ \mathbb{Q} &= \{q \mid q \text{ ist rationale Zahl}\}, \\ \mathbb{R} &= \{r \mid r \text{ ist reelle Zahl}\}. \end{aligned}$$

Sei $M_1 \subseteq M$. Dann bezeichnet man die Menge

$$C = \{m \mid m \text{ gehört zu } M \text{ aber nicht zu } M_1\}$$

als Komplement von M_1 in M und schreibt dafür

$$C = M \setminus M_1\,.$$

Die Menge aller Teilmengen einer Menge M heißt die Potenzmenge $\mathcal{P}(M)$ von M. Insbesondere gilt

$$\emptyset \in \mathcal{P}(M)\,, \quad M \in \mathcal{P}(M)$$

(M ist Teilmenge von sich selbst).

Beispiel. Teilmengen von $\{1,2,3,4\}$:

$$\begin{Bmatrix} \emptyset \\ \{1\},\{2\},\{3\},\{4\} \\ \{1,2\},\{1,3\},\{1,4\},\{2,3\},\{2,4\},\{3,4\} \\ \{1,2,3\},\{1,2,4\},\{1,3,4\},\{2,3,4\} \\ \{1,2,3,4\} \end{Bmatrix} = \mathcal{P}(\{1,2,3,4,\}) .$$

Die Potenzmenge enthält $16 = 2^4$ Elemente.

Der Mengenbegriff ist nicht so scharf definiert, wie man das sonst in der Mathematik gewohnt ist. Er ist ein Grundbegriff, auf dem die Mathematik aufbaut. Jedoch muß man vorsichtig sein, da bestimmte Mengendefinitionen auf Widersprüche führen und daher nicht zugelassen werden können.

Beispiel. „Russelsche Antinomie“ (B. Russell, 1872 – 1970):
Sei M die Menge aller Mengen, die sich nicht selbst als Element enthalten. Gilt $M \in M$?

Aus $M \in M$ folgt $M \notin M$. Umgekehrt folgt aus $M \notin M$ aber $M \in M$. Also kann weder $M \in M$ noch $M \notin M$ gelten. Da jedoch eine der beiden Aussagen richtig sein muß, ergibt sich ein Widerspruch. Diese Mengendefinition ist also nicht zugelassen.

Um diese Widersprüche zu vermeiden, legt man in der „axiomatischen“ Mengenlehre die zwischen Mengen gültigen Relationen durch ein System von formalen Axiomen fest. Meistens stellt man sich aber auf den Standpunkt der „naiven Mengenlehre“ und benützt die Cantorsche Mengendefinition, vermeidet aber auf Widersprüche führende Mengendefinitionen. Auch wir gehen so vor.

Weitere Definitionen und Bezeichnungen: Seien M_1 und M_2 Mengen. Dann bezeichnet man

$$M_1 \cup M_2 = \{x \mid x \in M_1 \text{ oder } x \in M_2\}$$

als Vereinigungsmenge. Das Wort „oder“ wird hier wie üblich im mathematischen Sprachgebrauch nicht als „exklusives oder“ benutzt. Ein Element x darf also zu beiden Mengen M_1 und M_2 gehören. Für ein beliebiges endliches oder unendliches System S von Mengen schreibt man auch

$$\bigcup_{M \in S} M = \{x \mid \text{Es gibt ein } M \in S \text{ mit } x \in M\} .$$

Man benützt auch die Bezeichnungen

$$\bigcup_{i=1}^{n} M_i = M_1 \cup M_2 \cup \ldots M_n\,; \quad \bigcup_{i=1}^{\infty} M_i = M_1 \cup M_2 \cup M_3 \cup \ldots .$$

Ebenso definiert man durch

$$M_1 \cap M_2 = \{x \mid x \in M_1 \text{ und } x \in M_2\}$$

den Durchschnitt der Mengen M_1 und M_2. Für ein beliebiges Mengensystem S schreibt man

$$\bigcap_{M \in S} M = \{x \mid \text{Für alle } M \in S \text{ gilt } x \in M\}\,.$$

Man benützt auch die Bezeichnungen

$$\begin{aligned}
\bigcap_{i=1}^{n} M_i &= M_1 \cap M_2 \cap \ldots \cap M_n\,, \\
\bigcap_{i=1}^{\infty} M_i &= M_1 \cap M_2 \cap M_3 \cap \ldots .
\end{aligned}$$

Gilt $M_1 \cap M_2 = \emptyset$, dann sagt man, die Mengen M_1 und M_2 seien disjunkt.

Beispiele. 1.) Sei $M_k = \{1, 2, \ldots, k\}$. Dann gilt

$$\begin{aligned}
\bigcup_{k=1}^{n} M_k &= M_n\,, \\
\bigcup_{k=1}^{\infty} M_k &= \mathbb{N}\,, \\
\bigcap_{k=1}^{n} M_k &= \bigcap_{k=1}^{\infty} M_k = \{1\}\,.
\end{aligned}$$

2.) Sei a eine positive reelle Zahl und sei $M_a = \{x \mid x \text{ reelle Zahl mit } 0 < x < a\}$.
Für

$$S = \{M_a \mid a > 0\}$$

gilt dann

$$\begin{aligned}
\bigcap_{M \in S} M &= \emptyset\,, \\
\bigcup_{M \in S} M &= \{x \mid x \text{ ist positive reelle Zahl}\}\,.
\end{aligned}$$

3.) Sei a eine positive reelle Zahl und sei

$$M'_a = \{x \mid x \text{ ist reelle Zahl mit } 0 \le x < a\}\,.$$

Für

$$S' = \{M'_a \mid a > 0\}$$

gilt dann

$$\begin{aligned} \bigcap_{M \in S'} M &= \{0\}\,, \\ \bigcup_{M \in S'} M &= \{x \mid x \text{ ist reelle Zahl mit } x \ge 0\}\,. \end{aligned}$$

Regeln von de Morgan. (1806 – 1871) Seien alle Mengen $M \in S$ Teilmenge einer gemeinsamen Obermenge K. Das Komplement von M in K bezeichne ich als M'. Dann gilt

$$\begin{aligned} \Big(\bigcup_{M \in S} M\Big)' &= \bigcap_{M \in S} M'\,, \\ \Big(\bigcap_{M \in S} M\Big)' &= \bigcup_{M \in S} M'\,. \end{aligned}$$

Cartesisches Produkt. Seien A, B Mengen. Unter dem cartesischen Produkt $A \times B$ der beiden Mengen A, B versteht man die Menge aller geordneten Paare (x, y) mit $x \in A$ und $y \in B$:

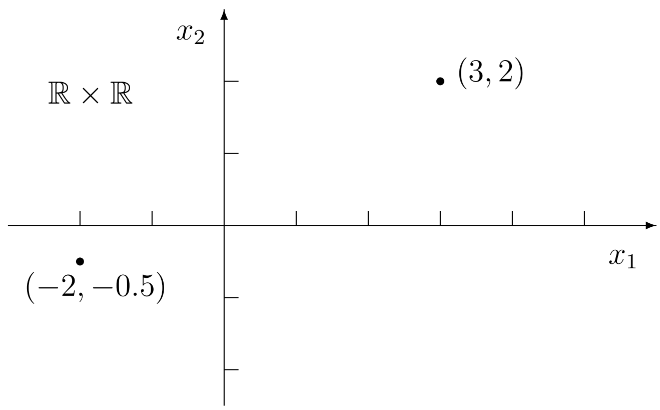
$$A \times B = \{z \mid z = (x, y),\ x \in A,\ y \in B\}\,.$$

Bei geordneten Paaren muß man also die Reihenfolge beachten. Zwei geordnete Paare (x, y) und (x', y') heißen gleich, genau dann wenn $x = x'$ und $y = y'$ gilt.

Mengentheoretisch einwandfrei kann man geordnete Paare folgendermaßen definieren:

$$(x, y) := \big\{x, \{x, y\}\big\}\,.$$

Beispiel. $A = B = \mathbb{R}$. Jeden Punkt der Ebene kann man durch ein Koordinatenpaar $(x_1, x_2) \in \mathbb{R} \times \mathbb{R} = \mathbb{R}^2$ festlegen; jeden Punkt des Raumes kann man durch ein Koordinatentripel $(x_1, x_2, x_3) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R} = \mathbb{R}^3$ darstellen.



Relationen. Auf der Menge der reellen Zahlen ist die „kleiner als“ Beziehung $<$ erklärt. Zum Beispiel gilt $1 < 2$. Man kann dies auch so ausdrücken: Die Relation $<$ trifft auf das geordnete Paar $(1, 2)$ zu. Dagegen trifft die Relation $<$ auf das geordnete Paar $(2, 1)$ nicht zu. Die Relation $<$ auf den reellen Zahlen ist also durch die Menge derjenigen geordneten Paare aus $\mathbb{R} \times \mathbb{R}$ bestimmt, auf die $<$ zutrifft. Das legt nahe, $<$ direkt als die Menge aller geordneten Paare

$$\{(x, y) \mid x \in \mathbb{R} \text{ und } y \in \mathbb{R} \text{ mit } x < y\} \subseteq \mathbb{R} \times \mathbb{R}$$

aufzufassen. Allgemein definiert man: Wenn M eine Menge ist, bezeichnet man jede Teilmenge R von $M \times M$ als zweistellige Relation auf M.

Beispiel. Sei M eine Menge. Eine endliche oder unendliche Menge P von Teilmengen von M heißt Partition von M, wenn

$$M = \bigcup_{T \in P} T$$

und $S \cap T = \emptyset$ für alle $S, T \in P$ mit $S \neq T$ gilt. Eine Menge P von Teilmengen von M ist also genau dann eine Partition von M, wenn jedes Element von M in genau einer Menge von P liegt.

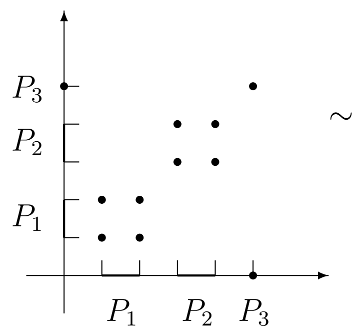
Sei P eine Partition von M. Man nennt zwei Elemente $x, y \in M$ äquivalent und schreibt $x \sim y$, genau dann wenn x und y zur selben Menge von P gehören.

Hierdurch wird eine Relation $\sim$ auf M erklärt:

$$\begin{aligned} \sim \;&= \{(x, y) \mid (x, y) \in M \times M \text{ und } x \sim y\} \\ &= \{(x, y) \mid x \text{ und } y \text{ gehören zur selben Menge aus } P\}. \end{aligned}$$

Beispiel.

$$\begin{aligned} M &= \{1,2,3,4,5\} \\ P_1 &= \{1,2\}\,, \quad P_2 = \{3,4\}\,, \quad P_3 = \{5\} \\ P &= \{P_1, P_2, P_3\} \quad \text{ist eine Partition}\,. \end{aligned}$$



$\sim$ hat folgende Eigenschaften:

(1) $x \sim x$ $\quad$ ($\sim$ ist reflexiv)

(2) Aus $x \sim y$ folgt $y \sim x$ $\quad$ ($\sim$ ist symmetrisch)

(3) Aus $x \sim y$ und $y \sim z$ folgt $x \sim z$ $\quad$ ($\sim$ ist transitiv)

Eine Relation, für die (1), (2), (3) gelten, heißt Äquivalenzrelation. Das einfachste Beispiel für eine Äquivalenzrelation ist $\quad = \quad$ (Gleichheit).

b.) **Verknüpfung von Aussagen**

Seien $A, B, C, \ldots$ Aussagen. Eine Aussage kann wahr (W) oder falsch (F) sein. Durch logische Verknüpfung von Aussagen erhält man neue Aussagen. Deren Wahrheitswert hängt allein von den Wahrheitswerten der ursprünglichen Aussagen ab. Ich will die Wahrheitstafeln für die vier Verknüpfungsoperationen $\wedge$ (‚und'), $\vee$ (‚oder'), $\Rightarrow$ (‚Implikation'), $\Leftrightarrow$ (‚Äquivalenz') angeben.

1.) ‚Und' $\wedge$

A	B	$A \wedge B$
W	W	W
F	W	F
W	F	F
F	F	F

Damit A ‚und' B wahr ist, müssen sowohl A als auch B wahr sein. Man versteht diese Festlegung besser, wenn man folgendes Beispiel betrachtet: Man kann die Definition einer

Menge

$$M = \{m \mid \text{definierende Aussage}\}$$

folgendermaßen auffassen: Es soll $m \in M$ gelten, wenn die definierende Aussage für m richtig ist (den Wahrheitswert W hat), und es soll $m \notin M$ sein, wenn die definierende Aussage für m falsch ist.

Beispiel. Seien M_1, M_2 Mengen. Den Durchschnitt $M_1 \cap M_2$ kann man folgendermaßen definieren

$$M_1 \cap M_2 := \{m \mid (m \in M_1) \wedge (m \in M_2)\}\,.$$

2.) ‚Oder' $\vee$

A	B	$A \vee B$
W	W	W
F	W	W
W	F	W
F	F	F

Damit A ‚oder' B wahr ist, genügt es, daß eine der beiden Aussagen A, B wahr ist.

Beispiel.

$$M_1 \cup M_2 := \{m \mid (m \in M_1) \vee (m \in M_2)\}\,.$$

3.) ‚Implikation' $\Rightarrow$ (aus A folgt B)

A	B	$A \Rightarrow B$
W	W	W
F	W	W
W	F	F
F	F	W

In der Umgangssprache interessiert man sich bei der Aussage „aus A folgt B" nur für den Fall, daß A wahr ist. Nur dann macht man eine Aussage über B. Was passiert, wenn A nicht wahr ist, läßt man offen:

Beispiel. „Wenn es regnet, wird die Straße naß." „Wenn es regnet, folgt daß die Straße naß wird." Was passiert, wenn es nicht regnet, bleibt offen. In diesem Fall kann die Straße trocken bleiben, oder aber aus anderen Gründen doch naß werden. Dem wird man in der formalen Festlegung des Wahrheitswertes von $A \Rightarrow B$ gerecht, indem man diese Aussage immer als wahr ansieht, wenn A falsch ist. Die Implikation ist eine „einseitige Aussage".

Beispiel.

$$\begin{aligned} N &= \{m \mid m \in M \wedge (m \in M_1 \Rightarrow m \in M_2)\} = M \cap \big((M\backslash M_1) \cup (M_1 \cap M_2)\big) \\ &= M \cap (M \setminus M_1) \cup (M \cap M_1 \cap M_2) = (M \setminus M_1) \cup (M \cap M_1 \cap M_2). \end{aligned}$$

$$\begin{aligned} L &= \{m \mid m \in M \wedge \big[(m \in M_1 \Rightarrow m \in M_2) \wedge (m \in M_2 \Rightarrow m \in M_1)\big]\} \\ &= M \cap \big[(M\backslash M_1) \cup (M_1 \cap M_2)\big] \cap \big[(M\backslash M_2) \cup (M_1 \cap M_2)\big] \\ &= (M \cap M_1 \cap M_2) \cup \big[M\backslash(M_1 \cup M_2)\big]\,. \end{aligned}$$

4.) ‚Äquivalenz‘ $\Leftrightarrow$ (genau dann, wenn)

A	B	$A \Leftrightarrow B$
W	W	W
F	W	F
W	F	F
F	F	W

Beispiel.

$$M = \{a \mid (a \in \mathbb{R}) \wedge (a^2 > 1 \Leftrightarrow a > 1)\} = \{a \mid a \geq -1\}\,.$$

Sei a eine reelle Zahl. Wenn man sagt „Ich beweise, daß $a^2 > 1$ ist genau dann wenn $|a| > 1$ ist“, dann meint man damit, daß man zeigt, daß $(a^2 > 1) \Leftrightarrow (|a| > 1)$ immer wahr ist. Hierzu muß man beweisen:

1.) Ist $a^2 > 1$ dann ist $|a| > 1\,.$

2.) Ist $|a| > 1$ dann ist $a^2 > 1\,.$

Wenn man dagegen sagt „Ich beweise, daß aus $a > 1$ folgt, daß $a^2 > 1$ ist“, dann meint man damit, daß man zeigt, daß $(a > 1) \Rightarrow (a^2 > 1)$ immer wahr ist. Hierzu muß bewiesen werden: Ist $a > 1$ dann ist $a^2 > 1\,.$

c.) **Quantoren, Negation von Aussagen**

Zur Abkürzung von Aussagen benützt man die Quantoren $\forall, \exists\,.$

All–Quantor: $\forall$ ‚für alle‘

Existenz–Quantor: $\exists$ ‚es gibt‘

Beispiele:

$$\forall_{a\in\mathbb{R}} \quad \exists_{n\in\mathbb{N}} \quad : \quad n > a.$$

„Für jede reelle Zahl a gibt es (mindestens) eine natürliche Zahl n mit $n > a\,.$“

$$\forall_{\substack{a\in\mathbb{R}\\ a>1}} \quad \forall_{n\in\mathbb{N}} \quad \exists_{m\in\mathbb{N}} \quad : \quad n \leq a^m.$$

„Für jede reelle Zahl a, die größer als 1 ist, und für jede natürliche Zahl n gibt es eine natürliche Zahl m so daß $n \leq a^m$ gilt.“

Die Negation einer mit Quantoren geschriebenen Aussage kann leicht bestimmt werden. Die Negation der Aussage aus dem ersten Beispiel ist

$$\exists_{a\in\mathbb{R}} \quad \forall_{n\in N} \quad : \quad n \leq a.$$

„Es gibt (mindestens) ein $a \in \mathbb{R}$ so daß für alle $n \in \mathbb{N}$ gilt $n \leq a$.“
Die Negation der zweiten Aussage ist:

$$\exists_{\substack{a\in\mathbb{R}\\ a>1}} \quad \exists_{n\in\mathbb{N}} \quad \forall_{m\in\mathbb{N}} \quad : n > a^m.$$

„Es gibt eine reelle Zahl a, die größer als 1 ist, und es gibt eine natürliche Zahl n, so daß für alle natürlichen Zahlen m gilt $n > a^m$.“

2 Reelle Zahlen

$$\begin{aligned} \frac{10}{7} &= 1,428571\,428571\ \ldots \\ \sqrt{2} &= 1,4142135\ \ldots \\ \pi &= 3,1415926\ \ldots \end{aligned}$$

sind reelle Zahlen. Was diese unendlichen Dezimalbrüche eigentlich sein sollen und wie man damit rechnet, soll nun geklärt werden. Dazu muß man genau angeben, was man unter „den reellen Zahlen“ verstehen will, und welche Eigenschaften sie haben sollen. Ich will zunächst das Vorgehen skizzieren:

Zunächst beachte man, daß es gar nicht darauf ankommt, was die reellen Zahlen eigentlich sind, sondern nur wie man damit rechnet. Man stellt daher zunächst eine Liste der Rechenregeln und Eigenschaften auf, die die reellen Zahlen haben sollen. Dann versucht man alle diese Eigenschaften auf möglichst wenige, möglichst klare „Grundsätze“, sogenannte „Axiome“, zurückzuführen. Nachdem man diese Axiome aufgestellt hat, vergißt man die konkrete Vorstellung einer reellen Zahl, die man aus Erfahrung hat, und nimmt an, man hätte irgendeine Menge und Rechenoperation auf dieser Menge, für die diese Axiome gelten. Dann zeigt man, daß wenn man eine zweite solche Menge hat, diese auf die erste in geeignetem Sinn „abgebildet“ werden kann, also in einem gewissen Sinn „äquivalent“ zur ersten Menge ist. Man zeigt also, daß es in diesem Sinn nur eine einzige solche Menge gibt, und diese Menge (zusammen mit den Rechenoperationen, für die die Axiome gelten), nennt man die „reellen Zahlen“. Schließlich muß man sich überlegen, daß die übliche Dezimalbruchdarstellung der reellen Zahlen mit der üblichen Erklärung der Rechenoperationen $+$, $\cdot$, $<$ usw. gerade eine solche Menge ist, ein sogenanntes „Modell“ für die reellen Zahlen darstellt.

Man nennt dies den axiomatischen Aufbau der reellen Zahlen. Man führt also die Eigenschaften der reellen Zahlen auf einige einfache Axiome, d.h. Grundannahmen zurück. Diese Axiome führt man nicht weiter zurück, leitet sie also nicht aus noch einfacheren Annahmen her. Denn irgendwo muß auch die Mathematik beginnen. Natürlich gibt es verschiedene Möglichkeiten, das Axiomensystem zu wählen. Auf jeden Fall aber darf das Axiomensystem keinen Widerspruch enthalten, und muß man aus diesen Axiomen alle Eigenschaften herleiten können, die die reellen Zahlen haben sollen. Ich werde gleich ein solches Axiomensystem angeben. Ich muß aber erwähnen, daß man auch andere Axiome wählen kann, und zwar Axiome, die nur die natürlichen Zahlen $\{1, 2, 3, \ldots\}$ betreffen. Dadurch hat man die natürlichen Zahlen gegeben. Aus diesen natürlichen Zahlen konstruiert man dann die rationalen, und aus den rationalen Zahlen die reellen Zahlen. Man nennt dies den konstruktiven Aufbau der reellen Zahlen. Das Ergebnis ist genau dasselbe. Jedoch spielt der konstruktive Aufbau eine große Rolle, weil es viel einfacher, natürlicher und überzeugender scheint, nur Annahmen über die natürlichen Zahlen zu machen als gleich über die reellen Zahlen. Dieser konstruktive Aufbau erfordert jedoch viel Arbeit. Daher führt man diesen Aufbau in einer Vorlesung über Infinitesimalrechnung nicht vor, sondern beginnt mit Axiomen über die reellen Zahlen.

a.) Körperaxiome

Wir nehmen an, je zwei reellen Zahlen a, b sei eindeutig eine reelle Zahl $a+b$, ihre Summe, und eine andere reelle Zahl $a \cdot b$, ihr Produkt zugeordnet. Für diese Zuordnung, die man Addition und Multiplikation nennt, soll gelten

Addition	A1.)	$a+b=b+a$	Kommutativgesetz
	A2.)	$a+(b+c)=(a+b)+c$	Assoziativgesetz
	A3.)	Es gibt genau eine Zahl 0 so daß für alle a gilt $a+0=a$	Existenz- und Eindeutigkeit der 0
	A4.)	Zu jedem a gibt es genau ein x so daß gilt $a+x=0$	Eindeutige Lösbarkeit der Gleichung $a+x=0$.
Multiplikation	A5.)	$a \cdot b = b \cdot a$	Kommutativgesetz
	A6.)	$a \cdot (b \cdot c) = (a \cdot b) \cdot c$	Assoziativgesetz
	A7.)	Es gibt genau eine von 0 verschiedene Zahl 1, so daß für alle a gilt $a \cdot 1 = a$	Existenz- und Eindeutigkeit der 1
	A8.)	Zu jedem $a \neq 0$ gibt es genau ein x, so daß gilt $a \cdot x = 1$	Eindeutige Lösbarkeit der Gleichung $a \cdot x = 1$.
	A9.)	$a \cdot (b+c) = a \cdot b + a \cdot c$	Distributivgesetz

Wegen der Assoziativgesetze 2.) und 6.) kann man Klammern weglassen.

$$\begin{aligned} a+(b+c) &= (a+b)+c =: a+b+c \\ a(bc) &= (ab)c =: abc \end{aligned}$$

Folgerungen aus den Körperaxiomen:

1.) Es gilt

$$(a+b)^2 \qquad = a^2 + 2ab + b^2$$

(binomische Formel) .

Hierbei sei $c^2 := cc\,;\quad 2 := 1+1$

Beweis:

$$\begin{aligned}
(a+b)(a+b) &= (a+b)a + (a+b)b \\
&= a(a+b) + b(a+b) \\
&= (aa+ab) + (ba+bb) \\
&= \big((aa+ab)+ba\big) + bb \\
&= \big(aa+(ab+ba)\big) + bb \\
&= \big(aa+(ab+ab)\big) + bb \\
&= \big(aa+ab(1+1)\big) + bb \\
&= a^2 + 2ab + b^2
\end{aligned}$$

Nach A4) gibt es genau eine Lösung x zur Gleichung $a + x = 0$. Diese Lösung bezeichnet man mit $(-a)$.

2.) Für reelle Zahlen a, b besitzt die Gleichung $a + x = b$ genau eine Lösung.

Beweis: Addiere zur Gleichung $a + x = b$ die Zahl $-a$. Es folgt

$$x = (-a) + a + x = (-a) + b = b + (-a)\,.$$

Umgekehrt erfüllt diese Zahl auch die Gleichung. Denn es gilt

$$a + \big(b + (-a)\big) = a + (-a) + b = b\,.$$

Man definiert

$$b - a = b + (-a) \qquad \text{Differenz}\,.$$

Also ist die Subtraktion definiert.

3.) Dasselbe gilt für die Multiplikaton: Nach A8.) gibt es zu jedem $a \neq 0$ eine Zahl a^{-1} mit $aa^{-1} = 1$; man schreibt auch $a^{-1} = \frac{1}{a}$.
Für $a \neq 0$ besitzt die Gleichung $ax = b$ genau eine Lösung $x = ba^{-1} =: \frac{b}{a}$ (Quotient).

Hiermit ist die Division erklärt.

Daß man in Axiom A8.) $a \neq 0$ voraussetzen muß, folgt aus

4.) Für jedes a gilt $a \cdot 0 = 0$.

Beweis:

$$a \cdot 1 = a \cdot (1+0) = a \cdot 1 + a \cdot 0 .$$

Subtraktion von $a = a \cdot 1$ auf beiden Seiten dieser Gleichung ergibt

$$0 = a \cdot 0 .$$

Ist in einem Produkt ein Faktor 0, so ist das Produkt 0.

Umgekehrt gilt auch:
Ist ein Produkt 0, so ist mindestens einer der Faktoren Null. Denn sei $ab = 0$ und $a \neq 0$. Dann folgt

$$0 = \frac{1}{a}(ab) = (\frac{1}{a}a)b = 1 \cdot b = b .$$

Also gilt: Ein Produkt ist Null, genau dann wenn wenigstens einer der Faktoren Null ist.

5.) Für alle a, b gilt

$$(-a)b = -(ab) .$$

Hierbei ist $-(ab)$ nach Definition die eindeutig bestimmte Lösung der Gleichung $ab + x = 0$. Andererseits gilt auch

$$ab + (-a)b = \big(a + (-a)\big)b = 0 \cdot b = 0 .$$

Also ist auch $(-a)b$ Lösung dieser Gleichung. Nach Voraussetzung A4.) ist diese Lösung eindeutig bestimmt, also ist $(-a)b = -(ab)$.

b.) Anordnungsaxiome

Auf der Menge der reellen Zahlen sei eine Relation $<$ definiert, die folgende Eigenschaften habe:

A10.)	Aus $a \neq b$ folgt, daß entweder $a < b$ oder $b < a$ richtig ist	Trichotomiegesetz
A11.)	Aus $a < b$ und $b < c$ folgt $a < c$	Transitivität
A12.)	Aus $a < b$ folgt $a + c < b + c$	Monotoniegesetz der Addition
A13.)	Aus $a < b$ und $0 < c$ folgt $ac < bc$	Monotoniegessetz der Multiplikation.

Gilt $0 < a$, so heißt a positiv. Gilt $a < 0$, so heißt a negativ. Eine Beziehung $a < b$ heißt Ungleichung oder Abschätzung.

Folgerungen aus den Ordnungsaxiomen.

1.) Aus $a < b$ und $c < d$ folgt $a + c < b + d$.

Beweis: Aus $a < b$ folgt

$$a + c < b + c\,.$$

Aus $c < d$ folgt

$$b + c < b + d\,.$$

Also folgt aus A11.)

$$a + c < b + d\,.$$

■

2.) Aus $a < b$ und $c < 0$ folgt $bc < ac$.

Beweis: Zunächst zeige ich

$$c < 0 \quad \Leftrightarrow \quad 0 < (-c)\,.$$

Denn durch Addition von $-c$ zur Ungleichung $c < 0$ folgt

$$(-c) + c < -c\,, \quad \text{also} \quad 0 < (-c)\,.$$

Umgekehrt folgt durch Addition von c zur Ungleichung $0 < (-c)$

$$c < (-c) + c\,, \quad \text{d.h.} \quad c < 0\,.$$

Also gilt

$$a < b \quad \text{und} \quad 0 < (-c) \quad \Rightarrow \quad (-c)a < (-c)b$$

also

$$-(ca) < -(cb)\,.$$

Durch Addition von $ac + bc$ folgt

$$bc < ac\,.$$

3.) Für jedes $a \neq 0$ gilt $0 < a^2$.

Beweis: Sei $0 < a$. Dann folgt

$$0 = 0 \cdot a < a \cdot a = a^2\,.$$

Sei $a < 0$. Dann folgt $0 < (-a)$, also

$$0 = 0 \cdot (-a) < (-a)(-a) = a \cdot a = a^2$$

Denn $\quad (-a)(-a) = -(a(-a)) = -(-aa) = -(-(aa)) = aa\,.$

4.) Aus $0 < a$ folgt $0 < \frac{1}{a}$.

Beweis: Es gilt $\frac{1}{a} \neq 0$. Denn sonst müßte gelten $1 = a \cdot \frac{1}{a} = 0$. Dies ist ein Widerspruch zu Axiom A7.).

Es kann auch nicht sein, daß $\frac{1}{a} < 0$ gilt. Denn hieraus und aus $0 < a$ könnte man folgern

$$1 = \frac{1}{a} \cdot a < 0 \cdot a = 0\,.$$

Das kann nicht sein wegen $1 = 1 \cdot 1 > 0$, nach 3.).

5.) Aus $a < b$ und $0 < ab$ folgt $\frac{1}{b} < \frac{1}{a}$.
Aus $a < b$ und $ab < 0$ folgt $\frac{1}{a} < \frac{1}{b}$.

Beweis: Sei $0 < ab$. Dann folgt $0 < \frac{1}{ab}$. Also erhalten wir aus $a < b$

$$\frac{1}{b} = \frac{1}{b} \cdot \frac{a}{a} = a \cdot \frac{1}{ab} \ < \ b \cdot \frac{1}{ab} = \frac{1}{a} \cdot \frac{b}{b} = \frac{1}{a}\,.$$

Ebenso folgt aus $ab < 0$

$$\frac{1}{a} = \frac{1}{a} \cdot \frac{b}{b} = b \cdot \frac{1}{ab} \ < \ a \cdot \frac{1}{ab} = \frac{1}{b} \cdot \frac{a}{a} = \frac{1}{b}\,.$$

Man schreibt $a > b$, genau dann wenn $b < a$ gilt. Die Relation $a \leq b$ gilt genau dann, wenn entweder $a < b$ oder $a = b$ ist.

6.) Für alle a, b mit $a \leq b$ gilt

$$a \leq \frac{a+b}{2} \leq b$$

Beweis: Aus $0 < 1$ folgt $1 < 2$, also $0 < 2$, also $0 < \frac{1}{2}$, also $\frac{a}{2} \leq \frac{b}{2}$, also

$$\begin{aligned} a &= \left(\frac{2}{2}\right)a &= \left(\frac{1}{2} + \frac{1}{2}\right)a = \frac{1}{2}a + \frac{1}{2}a \leq \frac{b}{2} + \frac{a}{2} \\ &\leq \frac{b}{2} + \frac{b}{2} &= b\,. \end{aligned}$$

Man kann nun Intervalle definieren: Sei $a \leq b$.

$$\begin{aligned} [a,b] &:= \{x \mid a \leq x \leq b\} && \text{abgeschlossenes Intervall} \\ [a,b) &:= \{x \mid a \leq x < b\} && \\ (a,b] &:= \{x \mid a < x \leq b\} && \text{halboffene Intervalle} \\ (a,b) &:= \{x \mid a < x < b\} && \text{offenes Intervall}\,. \end{aligned}$$

Man bezeichnet auch die folgenden Mengen als Intervalle

$$\begin{aligned}
[a,\infty) &:= \{x \mid a \le x\} \\
(a,\infty) &:= \{x \mid a < x\} \\
(-\infty,a] &:= \{x \mid x \le a\} \\
(-\infty,a) &:= \{x \mid x < a\} \\
(-\infty,\infty) &:= \mathbb{R}
\end{aligned}$$

$[a,a]$ ist ein einelementiges Intervall und (a,a) ist die leere Menge.

Definition

$$|a| := \begin{cases} a\,, & a \ge 0 \\ -a\,, & a < 0 \end{cases} \quad \text{„Betrag von } a\text{“}$$

Also gilt $|a| \ge 0$ und $-|a| \le a \le |a|$.

7.) Es gilt:

a.) $|a| = 0$ genau dann, wenn $a = 0$

b.) $|ab| = |a|\,|b|$

c.) $|a+b| \le |a| + |b|$ Dreiecksungleichung.

Beweis: a.) „$\Longleftrightarrow$“ klar.

b.) Unterscheide vier Fälle:

$$\begin{aligned}
a \ge 0\,, &\quad b \ge 0 \\
a \ge 0\,, &\quad b < 0 \\
a < 0\,, &\quad b \ge 0 \\
a < 0\,, &\quad b < 0\,.
\end{aligned}$$

Vierter Fall:

$$\begin{aligned}
a < 0\,, b < 0 \Rightarrow -a > 0\,, &\;\Rightarrow\; -ab < 0 \Rightarrow ab > 0 \Rightarrow |ab| = ab \\
a < 0\,, b < 0 \Rightarrow |a| &= -a\,, |b| = -b \Rightarrow \\
|a|\,|b| = (-a)\cdot(-b) &= -\big((-a)\cdot b\big) = -\big(-(ab)\big) = ab
\end{aligned}$$

Zusammen folgt $|ab| = |a|\,|b|$.

c.) Aus $-|a| \le a \le |a|$ und $-|b| \le b \le |b|$ folgt

$$-(|a|+|b|) = (-|a|) + (-|b|) \le a + b \le |a| + |b|\,.$$

Ist $a + b \geq 0$, dann folgt die Behauptung aus der rechten Seite dieser Ungleichung. Ist $a + b < 0$, so folgt $|a + b| = -(a + b)$. Multiplikation der linken Seite der obenstehenden Ungleichung mit -1 ergibt also

$$|a + b| = -(a + b) \leq |a| + |b|$$

8.) Sei $b \neq 0$. Dann gilt $|\frac{a}{b}| = \frac{|a|}{|b|}$.
Beweis: Wegen Behauptung 7b) gilt $|\frac{a}{b}| \cdot |b| = |\frac{a}{b} \cdot b| = |a|$. Division durch $|b|$ ergibt die Behauptung.

9.) $|\,|a| - |b|\,| \leq |a + b|$
Beweis: Sei $c := a + b$. Dann gilt $b = c - a$. Die Dreiecksungleichung ergibt

$$|b| = |c - a| = |c + (-a)| \leq |c| + |-a| = |c| + |a| = |a + b| + |a|,$$

also

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

Durch Vertauschen von a, b ergibt sich

$$-(|b| - |a|) = |a| - |b| \leq |a + b|,$$

also

$$|\,|a| - |b|\,| \leq |a + b|.$$

c.) Natürliche Zahlen, vollständige Induktion
Es ist naheliegend zu sagen, man erhalte die natürlichen Zahlen indem man von 1 ausgehend fortgesetzt 1 addiert. Besser definiert man die natürlichen Zahlen folgendermaßen:
Definition: Eine Menge M von reellen Zahlen heißt induktiv, wenn gilt
(a) $1 \in M$
(b) wenn $x \in M$ gilt, dann ist auch $x + 1 \in M$.
Induktive Mengen existieren. Beispiele sind die Menge aller reellen Zahlen und die Menge der positiven reellen Zahlen.

Hat man ein beliebiges System von induktiven Mengen, so ist auch der Durchschnitt aller dieser induktiven Mengen eine induktive Menge. Denn 1 gehört zu jeder induktiven Menge, also auch zum Durchschnitt. Gehört x zum Durchschnitt, dann gehört x auch zu jeder induktiven Menge des Systems, also gehört nach (b) auch $x + 1$ zu jeder dieser Mengen, also auch zum Durchschnitt.

Definition: Der Durchschnitt aller induktiven Mengen reeller Zahlen heißt Menge der natürlichen Zahlen und wird mit $\mathbb{N}$ bezeichnet.

1 ist die kleinste natürliche Zahl, weil die Menge aller reellen Zahlen, die größer oder gleich 1 sind, eine induktive Menge ist, und folglich die natürlichen Zahlen Teilmenge diese Menge sein müssen. Dagegen gibt es keine größte natürliche Zahl. Denn zu jedem $n \in \mathbb{N}$ gibt es noch $n+1 \in \mathbb{N}$, und wegen $1 > 0$ ist $n+1 > n$.

Es gilt folgender Satz:

Induktionssatz: Sei $W \subseteq \mathbb{N}$, und es gelte

(a) $1 \in W$

(b) gilt $n \in W$, dann gilt auch $n+1 \in W$.

Dann ist $W = \mathbb{N}$.

Beweis: Es gilt nach Voraussetzung $W \subseteq \mathbb{N}$. Andererseits ist W eine induktive Menge. Nach Definition ist $\mathbb{N}$ Teilmenge jeder induktiven Menge, also gilt $\mathbb{N} \subseteq W$, folglich $W = \mathbb{N}$.

Dieser Satz ist die Grundlage zur Beweismethode der vollständigen Induktion. Diese Beweismethode verläuft folgendermaßen:

Es sei $A(n)$ eine Aussage über die natürliche Zahl n. Man zeige nun:

(a) Induktionsanfang: $A(1)$ ist richtig

(b) Induktionsschritt: Aus der Annahme, $A(n)$ sei richtig (Induktionsannahme), folgt, daß auch $A(n+1)$ richtig ist.

Dann ist $A(n)$ für jede natürliche Zahl n richtig. Denn ist W die Menge aller natürlichen Zahlen, für die $A(n)$ richtig ist, so erfüllt W die Bedingungen des Induktionssatzes, also gilt $W = \mathbb{N}$, also ist $A(n)$ für alle natürlichen Zahlen richtig.

Beispiele für Beweise durch vollständige Induktion. In den folgenden Beispielen treten Summen der Form

$$a_1 + a_2 + a_3 + \ldots + a_{n-1} + a_n$$

und Produkte der Form

$$a_1 \cdot a_2 \cdot a_3 \ldots a_{n-1} \cdot a_n$$

auf. Diese Ausdrücke ergeben zunächst keinen Sinn, weil nicht klar ist, in welcher Reihenfolge die Summation oder die Produktbildung durchzuführen ist. Auf Grund der

Assoziativ– und Kommutativgesetze ist der Wert dieser Summen und Produkte jedoch von der Reihenfolge unabhängig. Dies ist jedoch nicht selbstverständlich, sondern muß erst durch vollständige Induktion nach n gezeigt werden. Ich überspringe diesen Beweis.

Weil diese Summen und Produkte von der Reihenfolge unabhängig sind, können zur Abkürzung folgende Symbole eingeführt werden:

$$\begin{aligned} \sum_{i=1}^{n} a_i &:= a_1 + \ldots + a_n \,. \\ \prod_{i=1}^{n} a_i &:= a_1 \cdot \ldots \cdot a_n \,. \end{aligned}$$

Es ist klar, daß der Summationsindex beliebig umbenannt werden darf. Außerdem definiert man

$$\begin{aligned} a^0 &:= 1 \qquad (0^0 := 1) \\ a^1 &:= a \\ a^2 &:= a \cdot a \\ &\vdots \\ a^n &:= \underbrace{aa\ldots a}_{\text{n Faktoren}} \end{aligned}$$

1.) Es sei $a > -1$. Dann gilt für alle natürlichen Zahlen n

$$(1+a)^n \geq 1 + an \,.$$

Beweis:
1.) Induktionsanfang: Für $n = 1$ gilt

$$(1+a)^n = (1+a)^1 = 1 + a \geq 1 + na \,.$$

2.) Induktionsschritt: Angenommen, die Aussage sei für eine natürliche Zahl n bewiesen. Also gelte

$$(1+a)^n \geq 1 + na \,.$$

Zu zeigen ist, daß sie dann auch für $n+1$ gilt:

$$\begin{aligned} (1+a)^{n+1} &= (1+a)^n(1+a) \geq (1+na)(1+a) \\ &= 1 + na + a + na^2 = 1 + (n+1)a + na^2 \\ &\geq 1 + (n+1)a \,, \end{aligned}$$

weil $na^2 > 0$ ist. ■

2.) Für alle natürlichen Zahlen n gilt

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

Dies kann auch in der Form

$$\sum_{j=1}^{n} (2j-1) = n^2$$

geschrieben werden.

1.) Induktionsanfang: Für $n = 1$ ist die Behauptung klar.

2.) Induktionsschritt: Angenommen, die Aussage sei für eine natürliche Zahl n richtig. Also gelte

$$\sum_{j=1}^{n} (2j-1) = n^2 \,.$$

Zu zeigen ist $\sum_{j=1}^{n+1} (2j-1) = (n+1)^2$.

Es gilt

$$\begin{aligned} \sum_{j=1}^{n+1} (2j-1) &= \sum_{j=1}^{n} (2j-1) + \big(2(n+1)-1\big) \\ = n^2 + \big(2(n+1)-1\big) &= n^2 + 2n + 1 = (n+1)^2 \,. \end{aligned}$$

■

Sei n eine natürliche Zahl und $0 \le k \le n$, $\quad k \in \{0\} \cup \mathbb{N}$. Dann sei

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

Hierbei bedeutet

$$n! := 1 \cdot 2 \cdot 3 \cdot \ldots \cdot n$$

(Bessere Definition: $1! := 1$; $(n+1)! := (n+1)(n!)$) und $0! := 1$. Man nennt $\binom{n}{k}$ Binomialkoeffizient.

3.) Für alle natürlichen Zahlen n und für alle natürlichen Zahlen $1 \le k \le n$ gilt

$$\binom{n}{k-1} + \binom{n}{k} = \binom{n+1}{k} \qquad (k \in \mathbb{N} \wedge k \ge 2 \Rightarrow k-1 \in \mathbb{N}, \quad \text{siehe später}) \,.$$

Beweis: Dies ergibt sich einfach durch nachrechnen. Die Beweismethode der vollständigen Induktion braucht dazu nicht verwendet zu werden:

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

Dies kann nun benützt werden, um folgende Formel zu beweisen:

4.) a, b seien reelle Zahlen. Dann gilt für alle natürlichen Zahlen $n \geq 1$

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

(Binomische Formel). Für $n=2$ ergibt sich $(a+b)^2=\sum_{k=0}^{2}\binom{n}{k}a^{n-k}b^k=a^2+2ab+b^2$.

Beweis: 1.) Induktionsanfang: Sei $n=1$:

$$a+b=\sum_{k=0}^{1}\binom{n}{k}a^{n-k}b^k=\binom{1}{0}a+\binom{1}{1}b=a+b\,.$$

2.) Induktionsschritt: Sei für ein $n \in \mathbb{N}$ die Formel richtig:

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

Dann gilt

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

Setze $j=k+1$. Dann gilt $k=j-1$, und es folgt

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

Nun werden einige Eigenschaften der natürlichen Zahlen hergeleitet.

Satz: Ist $n \geq 2$ eine natürliche Zahl, dann auch $n-1$.
Denn wenn das nicht richtig wäre, gäbe es ein $n_0 \in \mathbb{N}, n_0 \geq 2$, mit: $n_0 - 1$ ist keine natürliche Zahl. Dann wäre aber $\mathbb{N} \backslash \{n_0\}$ eine induktive Menge, die echte Teilmenge von $\mathbb{N}$ wäre. Dies ist nicht möglich, weil $\mathbb{N}$ die kleinste induktive Menge ist.

Satz: Für jede natürliche Zahl n gilt: Es gibt keine natürliche Zahl m mit der Eigenschaft $n < m < n+1$.

Beweis: 1.) Induktionsanfang: Die Behauptung ist richtig für $n = 1$. Denn $\{1\} \cup \{x \mid x \geq 2\}$ ist eine induktive Menge, also ist $\mathbb{N}$ eine Teilmenge hiervon.
2.) Induktionsschritt: Es gebe keine natürliche Zahl m mit $n < m < n+1$. Dann kann es auch keine natürliche Zahl m' mit $n+1 < m' < n+2$ geben, denn sonst wäre nach dem vorangehenden Satz $m' - 1$ eine natürliche Zahl mit $n < m' - 1 < n+1$.

Satz: In jeder nichtleeren Menge A von natürlichen Zahlen gibt es ein kleinstes Element. (Die natürlichen Zahlen sind wohlgeordnet.)
Beweis: Betrachte die Menge W von natürlichen Zahlen

$$W = \{n \mid n \leq a \text{ für alle } a \in A\}.$$

Es gilt $1 \in W$. Würde mit jeder natürlichen Zahl n auch $n+1$ zu W gehören, dann wäre $W = \mathbb{N}$. Dies kann nicht sein, weil A nichtleer ist. Denn wenn es in A ein Element a gibt, dann ist $a+1$ eine natürliche Zahl, die nach Definition von W nicht zu W gehören kann.

Also gibt es eine Zahl $k \in W$ mit $k+1 \notin W$. Dieses k ist kleinstes Element von A. Hierzu ist zu zeigen: $k \in A$ und

$$\forall_{a \in A} : k \leq a.$$

Daß $k \leq a$ gilt für alle $a \in A$ ist klar, da $k \in W$ ist. Also muß noch bewiesen werden, daß $k \in A$ ist.

Wegen $k+1 \notin W$ gibt es $a_0 \in A$ mit $k+1 > a_0$, also ist $a_0 \leq k$, weil es keine natürliche Zahl zwischen k und $k+1$ gibt. Also ist

$$\begin{aligned} a_0 &\leq k \\ \text{und} \quad k &\leq a_0, \end{aligned}$$

also $k = a_0$.

Oft will man bei Induktionsbeweisen nicht mit der Zahl 1 sondern mit beliebigem n_0 als Induktionsanfang beginnen. Der folgende Satz zeigt, daß dies richtig ist:

Satz: Es sei $W \subseteq \mathbb{N}$ und es gelte

1.) $n_0 \in W$

2.) wenn $n \geq n_0$ und $n \in W$, so $n+1 \in W$.

Dann gilt $\{n \mid n \in \mathbb{N} \wedge n \geq n_0\} \subseteq W$.

Beweis: Die Menge

$$W' = \{k \mid k \in \mathbb{N} \wedge k \leq n_0 - 1\} \cup W$$

ist eine induktive Menge. Denn es gilt $1 \in W'$. Außerdem gilt: wenn $n \in W'$, dann auch $n+1 \in W'$. Im Beweis muß man drei Fälle unterscheiden:

a.) $n < n_0 - 1$

b.) $n = n_0 - 1$

c.) $n > n_0 - 1$.

Im Fall a.) ist $n+1 < n_0$. Da zwischen $n_0 - 1$ und n_0 keine natürliche Zahl liegt, bedeutet dies $n+1 \leq n_0 - 1$, also $n+1 \in W'$.

Im Fall b.) ist $n+1 = n_0$, also $n+1 \in W \subseteq W'$.

Im Fall c.) ist $n \geq n_0$, da zwischen $n_0 - 1$ und n_0 keine natürliche Zahl liegt. Also ist $n \in W$, und dann gehört $n+1$ zu $W \subseteq W'$ nach Voraussetzung. Somit ist W' eine induktive Menge natürlicher Zahlen, also $W' = \mathbb{N}$. Hieraus folgt die Behauptung.

Als Übung beweise man:

Satz: Es seien m und n natürliche Zahlen. Dann sind $m+n$ und $m \cdot n$ natürliche Zahlen.

Satz: Es seien m, n natürliche Zahlen mit $m > n$. Dann ist auch $m - n$ eine natürliche Zahl.

Man kann nun auch die Menge der ganzen Zahlen

$$\mathbb{Z} := \{m \mid m = 0 \text{ oder } m \in \mathbb{N} \text{ oder } -m \in \mathbb{N}\}$$

und die Menge der rationalen Zahlen

$$\mathbb{Q} := \{q \mid q = \frac{m}{n}\,,\; m \in \mathbb{Z}\,,\; n \in \mathbb{N}\}$$

definieren. Man beachte, daß die Darstellung einer rationalen Zahl $q = \frac{m}{n}$ nicht eindeutig ist. Denn es gilt zum Beispiel

$$q = \frac{m}{n} = \frac{2m}{2n} = \frac{3m}{3n} = \ldots .$$

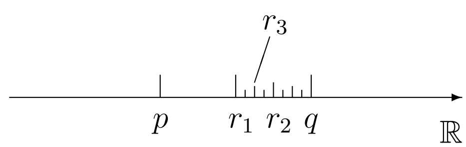
Satz: Zu je zwei rationalen Zahlen p, q mit $p < q$ gibt es noch eine dritte rationale Zahl r mit

$$p < r < q.$$

Beweis: Setze $r = \frac{p+q}{2}$.

d.) Reelle Zahlen, Vollständigkeit, Dedekindscher Schnitt

Die rationalen Zahlen erfüllen die Körperaxiome und die Anordnungsaxiome. Man kann also mit den natürlichen Zahlen wie üblich rechnen. Außerdem sieht man aus dem letzten Satz, daß die rationalen Zahlen dicht liegen auf der Zahlengerade:



Es ist jedoch bekannt, daß die rationalen Zahlen die Zahlengerade nicht ausfüllen, und daß dazwischen noch die irrationalen Zahlen liegen. Um dies einzusehen benötigt man einige weitere Grundbegriffe.

Definition: Eine Menge M von reellen Zahlen heißt nach oben beschränkt, wenn gilt

$$\exists_{s \in \mathbb{R}} \quad \forall_{x \in M} : x \leq s\,.$$

M heißt nach unten beschränkt, wenn gilt

$$\exists_{s \in \mathbb{R}} \quad \forall_{x \in M} : s \leq x\,.$$

s heißt obere bzw. untere Schranke.

Eine nach oben und unten beschränkte Menge heißt beschränkt. (Die leere Menge ist beschränkt.)

Definition: Eine obere bzw. untere Schranke von M, die zu M gehört, heißt Maximum beziehungsweise Minimum von M (größte beziehungsweise kleinste Zahl von M).

Jede endliche Menge reeller Zahlen ist beschränkt und besitzt ein Maximum und Minimum. Man beweist dies durch vollständige Induktion nach der Anzahl n der Elemente. Es wurde auch schon gezeigt, daß jede beliebige Teilmenge natürlicher Zahlen nach unten beschränkt ist und ein Minimum besitzt. Dagegen wird später noch gezeigt werden, daß die natürlichen Zahlen nicht nach oben beschränkt sind. Beschränkt sind auch die Intervalle

$$[a,b]\,,\quad (a,b)\,,\quad [a,b)\,,\quad (a,b]\,,$$

unbeschränkt sind

$$[a,\infty)\,,\quad (a,\infty)\,,\quad (-\infty,a)\,,\quad (-\infty,a],\quad (-\infty,\infty)=\mathbb{R}\,.$$

Man beachte, daß das Intervall $(0,1)$ nach oben und unten beschränkt ist, jedoch weder Maximum noch Minimum besitzt. Denn es gilt

$$(0,1)=\{x\mid 0<x<1\}\,.$$

Wäre $m\in(0,1)$ Maximum, dann müßte gelten

$$0<m<1$$

und

$$\forall_{x\in(0,1)}\quad : x\leq m\,.$$

Das kann aber nicht sein, denn es gilt

$$m<\frac{m+1}{2}<1\,,$$

also $\frac{m+1}{2}\in(0,1)$.

Definition: Sei M eine nach oben beschränkte Menge reeller Zahlen. Die obere Schranke s_0 heißt kleinste obere Schranke oder Supremum von M , wenn für jede obere Schranke s von M die Ungleichung

$$s_0\leq s$$

gilt. Bezeichnung: $\quad s_0=\sup M$.

Wenn M nach unten beschränkt ist, definiert man entsprechend die größte untere Schranke: Die untere Schranke s_0 heißt größte untere Schranke oder Infimum von M , wenn für jede untere Schranke s die Ungleichung

$$s\leq s_0$$

gilt. Bezeichnung: $\quad s_0:=\inf M$.

Die Menge $(0,1)$ besitzt sowohl Infimum als auch Supremum:

$$\begin{aligned}\inf(0,1) &= 0\\ \sup(0,1) &= 1\,.\end{aligned}$$

Denn 1 ist obere Schranke, und gäbe es eine echt kleinere obere Schranke, dann könnte wie oben ein Widerspruch hergeleitet werden.

Man hätte nun gerne, daß jede nach oben beschränkte Menge ein Supremum, (und entsprechend jede nach unten beschränkte Menge ein Infimum) hat. Später wird sich zeigen, daß dies eine in der Analysis überaus wichtige Eigenschaft ist. Aus den Körper– und Anordnungsaxiomen kann man dies aber nicht folgern. Um dies einzusehen beachte man, daß die Menge der rationalen Zahlen die Körper– und Anordnungsaxiome erfüllt. Ich zeige aber jetzt, daß in dieser Menge nicht jede nach oben beschränkte Menge ein Supremum besitzt.

Die Menge $M = \{x \mid x^2 < 2\}$ ist nach oben beschränkt. Denn etwa $s = \frac{3}{2}$ ist obere Schranke. Würde es nämlich $x \in M$ geben mit $x > s$, so müßte nach den Anordnungsaxiomen $x^2 > s^2 = \frac{9}{4} > 2$ sein, und dies ist ein Widerspruch zur Annahme $x \in M$.

Diese Menge kann jedoch keine rationale Zahl als Supremum besitzen. Um dies einzusehen beachte man, daß das Supremum s_0, falls es existiert, positive sein muss wegen $1 \in M$. Ausserdem muss s_0 die Gleichung $s_0^2 = 2$ erfüllen. Zum Beweis schließt man die Fälle
a.) $s_0^2 < 2$
b.) $s_0^2 > 2$
aus.

a.) Wäre $s_0^2 < 2$, könnte man $h > 0$ so bestimmen, daß auch $(s_0 + h)^2 < 2$ gelten würde. Denn

$$(s_0 + h)^2 < 2 \Longleftrightarrow s_0^2 + 2s_0 h + h^2 < 2 \Longleftrightarrow 2s_0 h + h^2 < 2 - s_0^2 .$$

Also genügt es, $0 < h < 1$ zu finden mit $2s_0 h + h < 2 - s_0^2$, und dies gilt genau dann wenn

$$h < \frac{2 - s_0^2}{2s_0 + 1}$$

ist. Die letzte Ungleichung wird zum Beispiel erfüllt durch $h = \frac{1}{2}\,\frac{2-s_0^2}{2s_0+1} > 0$. Im Widerspruch zur Annahme könnte also s_0 nicht Supremum sein.

b.) Wäre $s_0^2 > 2$, dann wäre auch $s = \frac{s_0}{2} + \frac{1}{s_0} = s_0 - \frac{s_0^2-2}{2s_0} < s_0$ noch obere Schranke von M, wegen

$$s^2 = \Big(s_0 - \frac{s_0^2 - 2}{2s_0}\Big)^2 = s_0^2 - (s_0^2 - 2) + \Big(\frac{s_0^2 - 2}{2s_0}\Big)^2 = 2 + \Big(\frac{s_0^2 - 2}{2s_0}\Big)^2 > 2 .$$

Also könnte s_0 nicht kleinste obere Schranke sein, im Widerspruch zur Annahme. Folglich muß $s_0^2 = 2$ gelten.

Als nächstes zeige ich, daß es keine rationale Zahl s_0 gibt mit $s_0^2 = 2$.
Denn wäre $s_0 = \frac{m}{n}$, dann müßte gelten

$$\frac{m^2}{n^2} = s_0^2 = 2\,,$$

also

$$m^2 = 2n^2\,.$$

Folglich muß m eine gerade Zahl sein, also $m = 2m'$ für eine geeignete natürliche Zahl m'. Folglich würde gelten $4m'^2 = 2n^2$, also $2m'^2 = n^2$, und hieraus könnte man folgern, daß auch n gerade sein muß. Man kann aber davon ausgehen, daß von vornherein m und n nicht beide gerade sind, weil man sonst kürzen könnte. Also hat die Annahme, s_0 sei rational, zu einem Widerspruch geführt. Diese Überlegungen zeigen, daß die Menge $M = \{x \mid x^2 < 2\}$ in der Menge der rationalen Zahlen kein Supremum besitzt. Also bleibt einem nichts weiter übrig, als ein neues Axiom hinzuzunehmen.

A 14.) **Vollständigkeitsaxiom:** Jede nicht leere, nach oben beschränkte Menge reeller Zahlen besitzt eine kleinste obere Schranke.

Natürlich fragt man sich, ob dieses Axiom sinnvoll ist, d.h. ob es überhaupt eine Menge gibt die neben den Körper– und Anordnungsaxiomen auch das Vollständigkeitsaxiom erfüllt. Daher will ich an dieser Stelle kurz auf den konstruktiven Aufbau der reellen Zahlen eingehen und zeigen, daß man aus den rationalen Zahlen eine größere Menge, nämlich die Menge der reellen Zahlen konstruieren kann, die dieses Vorständigkeitsaxiom erfüllt. Ich will zeigen, wie man diese Menge mit dem Dedekinschen Schnitt konstruieren kann. Daneben gibt es andere Methoden; beispielsweise kann man Cauchyfolgen oder Intervallschachtelung benützen.

Um aus den rationalen Zahlen die reellen zu konstruieren, muß man zu den rationalen Zahlen neue Zahlen hinzunehmen. Denn die reellen Zahlen sollen ja alle Axiome A 1 – A 14 erfüllen, und hieraus folgt, daß die reellen Zahlen eine Zahl s_0 enthalten müssen, deren Quadrat 2 ist. Diese neuen Zahlen konstruiert man prinzipiell so, daß man zur Menge der rationalen Zahlen neue „Dinge“ hinzunimmt, und auf diesen neuen Dingen die Rechenregeln so erklärt, daß alle Axiome erfüllt sind.

Definition: Ein Paar $(\underline{U}, \overline{U})$ von Teilmengen von $\mathbb{Q}$ heißt ein Dedekindscher Schnitt in $\mathbb{Q}$, wenn folgendes gilt:

1. $\mathbb{Q} = \underline{U} \cup \overline{U}$,

2. $\underline{U} \neq \emptyset$ und $\overline{U} \neq \emptyset$,

3. Aus $\underline{q} \in \underline{U}$ und $\overline{q} \in \overline{U}$ folgt $\underline{q} < \overline{q}$,

4. $\overline{U}$ besitzt kein kleinstes Element.

$\underline{U}$ heißt die Unterklasse, $\overline{U}$ die Oberklasse des Schnittes. (Richard Dedekind, 1831 – 1916.)

Jeder Dedekindsche Schnitt heißt eine reelle Zahl. Eine reelle Zahl $(\underline{U}\,,\overline{U})$, deren Unterklasse $\underline{U}$ ein größtes Element p besitzt, wird mit der rationalen Zahl p identifiziert. Dadurch erscheint $\mathbb{Q}$ als Teil der Menge $\mathbb{R}$ aller reellen Zahlen (aller Schnitte). Auf der so konstruierten Menge der reellen Zahlen werden die Addition, die Multiplikation und die Ordnungsrelation folgendermaßen erklärt:

Es seien $r_1 = (\underline{U}_1, \overline{U}_1)$ und $r_2 = (\underline{U}_2, \overline{U}_2)$. Die Summe $r_1 + r_2$ sei die reelle Zahl $(\underline{V}, \overline{V})$ mit

$$\begin{aligned}
\underline{V} &= \{q_1 + q_2 \mid q_1 \in \underline{U}_1\,,\ q_2 \in \underline{U}_2\} &&=: \underline{U}_1 + \underline{U}_2 \\
\overline{V} &= \{q_1 + q_2 \mid q_1 \in \overline{U}_1\,,\ q_2 \in \overline{U}_2\} &&=: \overline{U}_1 + \overline{U}_2\,.
\end{aligned}$$

Es gelte $r_1 < r_2$, genau dann wenn $\underline{U}_1 \subsetneq \underline{U}_2$ und $\overline{U}_1 \supsetneq \overline{U}_2$. Für $r_1 \geq 0\,, r_2 \geq 0$ sei das Produkt $r_1 \cdot r_2$ gleich der reellen Zahl $(\underline{V}, \overline{V})$ mit

$$\underline{V} = \{q_1 \cdot q_2 \mid q_1 \in \underline{U}_1\,,\ q_2 \in \underline{U}_2\,,\ q_1 \geq 0\,,\ q_2 \geq 0\} \cup \{q \mid q < 0\}\,.$$

Hiermit definiert man das Produkt für beliebige Zahlen r_1, r_2 durch

$$r_1 \cdot r_2 = \begin{cases} -(|r_1|\,|r_2|)\,, & r_1 < 0\,, \quad r_2 \geq 0, \\ -(|r_1| \cdot |r_2|)\,, & r_1 \geq 0\,, \quad r_2 < 0, \\ \ \ |r_1|\,|r_2|\,, & r_1 < 0\,, \quad r_2 < 0\,. \end{cases}$$

Wenn man Addition, Multiplikation und Ordnungsrelation so erklärt, dann sind die Körper– und Anordnungsaxiome erfüllt. Außerdem ist das Vollständigkeitsaxiom erfüllt. Denn sei $M \neq \emptyset$ eine nach oben beschränkte Menge reeller Zahlen. Diese Menge hat ein Supremum $s_0 = (\underline{V}, \overline{V})$ mit

$$\begin{aligned}
\underline{V} &= \bigcap_{(\underline{U},\overline{U}) \in M'} \underline{U}\,, \\
\overline{V} &= \bigcup_{(\underline{U},\overline{U}) \in M'} \overline{U}\,,
\end{aligned}$$

wobei $M' = \{r \mid r \text{ ist obere Schranke von } M\}$ sei.

Ich diskutiere nun noch einige Eigenschaften von Supremum und Infimum.

Satz: Die Zahl s ist genau dann das Supremum der Menge M, wenn folgende Bedingungen erfüllt sind:

(1) Für jedes $x \in M$ gilt $x \leq s$

(2) Zu jedem $\varepsilon > 0$ gibt es ein $x \in M$ mit $s - \varepsilon \leq x \leq s$.

Beweis: Falls s Supremum von M ist, sind beide Bedingungen erfüllt. Denn (1) ist klar. Wäre (2) nicht erfüllt, so würde es ein $\varepsilon > 0$ geben mit $x < s - \varepsilon$ für alle $x \in M$, und somit wäre $s - \varepsilon$ eine obere Schranke von M, die kleiner wäre als s, im Widerspruch zu Annahme. Sind umgekehrt beide Bedingungen erfüllt, dann muß s obere Schranke von M sein. Wegen (2) kann es keine kleinere obere Schranke geben. ■

Satz: Jede nicht leere, nach unten beschränkte Menge reeller Zahlen hat eine größte untere Schranke.

Beweis: Die Menge $M' = \{x \mid (-x) \in M\}$ ist nach oben beschränkt. Also hat diese Menge ein Supremum s'_0, und $u_0 = -s'_0$ ist das Infimum.

Denn sei $x \in M$. Dann ist $-x \in M'$, also $s'_0 \geq (-x)$, also $u_0 = -s'_0 \leq x$. Also ist u_0 untere Schranke von M. Gäbe es eine größere untere Schranke u_1, so wäre $-u_1$ kleinere obere Schranke von M'. ■

e.) Archimedische Anordnung der reellen Zahlen

Satz: Die Menge der natürlichen Zahlen ist nicht beschränkt.

Beweis: Wäre $\mathbb{N}$ nach oben beschränkt, dann würde für $\mathbb{N}$ eine kleinste obere Schranke s_0 existieren. Also würde gelten

$$\forall_{n \in \mathbb{N}} : \quad n \leq s_0 .$$

Nach dem im letzten Abschnitt bewiesenen Satz müßte es also ein $n_0 \in \mathbb{N}$ geben mit $n_0 > s_0 - 1$, und hieraus würde folgen $n_0 + 1 > s_0$. Da aber $n_0 + 1 \in \mathbb{N}$ ist, könnte s_0 keine obere Schranke sein. ■

Satz: Zu jedem $a > 0$ und jedem $b \in \mathbb{R}$ gibt es eine natürliche Zahl n derart, daß gilt

$$na > b .$$

($\mathbb{R}$ ist archimedisch angeordnet.)

Beweis: Gäbe es keine solche Zahl n, dann würde für alle $n \in N$ gelten $na \leq b$, also

$$n \leq \frac{b}{a}$$

also wäre $\mathbb{N}$ beschränkt. ■

Satz: Ist $a \geq 0$ und gilt für jede natürliche Zahl n die Ungleichung $a \leq \frac{1}{n}$, dann ist $a = 0$.

Beweis: Wäre $a > 0$ würde folgen

$$\forall_{n\in\mathbb{N}} : \quad n \leq \frac{1}{a}$$

d.h. $\mathbb{N}$ wäre nach oben beschränkt. ■

Satz: a.) Sind a, b reelle Zahlen mit $a < b$, so gibt es mindestens eine rationale Zahl r mit $a < r < b$.

b.) Sind a, b rationale Zahlen mit $a < b$, dann gibt es mindestens eine irrationale Zahl r mit $a < r < b$.

Beweis:

a.) Sei $a \geq 1$. Wegen $b > a$ existiert $n \in \mathbb{N}$ mit $n \cdot (b-a) > 1$, also $0 < \frac{1}{n} < b - a$. Halte n fest. Betrachte nun die Menge

$$\{m \mid m \in \mathbb{N} \wedge \frac{m}{n} > a\} .$$

Diese Menge ist nicht leer, weil es mindestnes ein $m \in \mathbb{N}$ gibt mit $m \cdot \frac{1}{n} > a$, und wegen $a \geq 1$ gilt $m \geq 2$ für alle Elemente der Menge. Außerdem ist sie eine Teilmenge der natürlichen Zahlen, also enthält sie ein kleinstes Element k . Für k gilt

$$a < \frac{k}{n} .$$

Also genügt es zu beweisen, daß $\frac{k}{n} < b$ gilt. Wäre $\frac{k}{n} \geq b$, würde gelten

$$\frac{k-1}{n} = \frac{k}{n} - \frac{1}{n} \geq b - \frac{1}{n} > b - (b-a) = a ,$$

also wäre k nicht Minimum. Widerspruch, also gilt

$$a < \frac{k}{n} < b .$$

Für $a < 1$ erhält man die Behauptung durch Verschieben von $[a, b]$.

b.) Angenommen, die Behauptung sei falsch. Dann gäbe es rationale Zahlen a, b mit $a < b$ so daß das Intervall $[a, b]$ ganz aus rationalen Zahlen bestehen würde. Man beachte nun, daß jede reelle Zahl, die sich als Summe zweier rationaler Zahlen darstellen läßt, rational ist, weil $\mathbb{Q}$ ein Körper ist. Also ist $\frac{a+b}{2}$ rational, und folglich besteht das Intervall $[-\frac{b-a}{2} , \frac{b-a}{2}]$ nur aus rationalen Zahlen. Denn für $x \in [-\frac{b-a}{2} , \frac{b-a}{2}]$ gilt

$$a = -\frac{b-a}{2} + \frac{a+b}{2} \leq x + \frac{a+b}{2} \leq \frac{b-a}{2} + \frac{a+b}{2} = b ,$$

also ist $q = x + \frac{a+b}{2} \in [a, b]$, also ist q rational, also ist auch $x = q - \frac{a+b}{2}$ rational.

Man beachte weiter, daß jede reelle Zahl, die sich als Produkt zweier rationaler Zahlen darstellen läßt, rational ist. Insbesondere ist das Produkt einer rationalen Zahl und einer natürlichen Zahl rational. Also besteht jedes der Intervalle

$$[-n\,\frac{b-a}{2}\,,\; n\,\frac{b-a}{2}]$$

aus rationalen Zahlen. Sei nun r eine beliebige reelle Zahl. Da $\frac{b-a}{2} > 0$ ist, gibt es ein $n \in \mathbb{N}$ mit

$$r \in [-n\,\frac{b-a}{2}\,,\; n\,\frac{b-a}{2}]\,,$$

also ist r rational, also wäre jede reelle Zahl rational. Es wurde aber schon bewiesen, daß $\sqrt{2}$ irrational ist, also hat die Annahme zu einem Widerspruch geführt, und die Behauptung muß richtig sein.

Folgerung: Seien a, b reelle Zahlen mit $a < b$. Dann gibt es mindestens eine irrationale Zahl r mit

$$a < r < b\,.$$

$\blacksquare$

3 Funktionen

3a.) Elementare Begriffe

Seien A, B Mengen. Jedem $x \in A$ sei ein eindeutiges Element aus B zugeordnet, das mit $f(x)$ bezeichnet sei. Man sagt, daß durch diese Zuordnung eine Abbildung f von A in B gegeben ist. Zur Bezeichnung von Abbildungen verwendet man die Schreibweisen

$$f : A \to B$$

und, wenn man auf die Elemente abhebt,

$$x \mapsto f(x)\,.$$

Anstelle des Begriffes Abbildung verwendet man auch den Begriff Funktion. Man beachte, daß der Name der Funktion f ist; nicht $f(x)$. Dies ist das eindeutige Element $y = f(x) \in B$, das dem Element $x \in A$ zugeordnet wird.

Jedem $x \in A$ darf nur ein eindeutiges Element $y \in B$ zugeordnet werden. Dagegen darf zwei verschiedenen Elementen $x_1, x_2 \in A$ mit $x_1 \neq x_2$ ein und dasselbe $y \in B$ zugeordnet werden: $\quad f(x_1) = f(x_2)\,.$

Beispiele:

1.) Sei $f : \mathbb{R} \to \mathbb{R} \quad x \mapsto f(x) := 1 \quad$ (Konstante Funktion).

2.) Sei $A = B = \mathbb{R}$. Es werde eine Funktion $f : \mathbb{R} \to \mathbb{R}$ definiert durch

$$\forall x \in \mathbb{R} : \;\; x \mapsto f(x) := x\,.$$

Diese Abbildung nennt man auch die Identität.

3.) Sei $g : \mathbb{R} \to \mathbb{R}$ definiert durch

$$x \mapsto g(x) := x^2\,.$$

4.) $h_1 : [0, \infty] \to \mathbb{R}\,, \quad x \mapsto \sqrt{x}\,; \quad h_2 : [0, \infty) \to \mathbb{R}\,, \quad x \mapsto -\sqrt{x}\,.$

5.) Der Weg s, den ein Körper beim freien Fall zurücklegt, hängt von der Fallzeit t ab. Man sagt, „s ist eine Funktion von t“. Der Zusammenhang ist

$$s = \frac{1}{2} g t^2\,. \quad (g = \text{ Erdbeschleunigung}\,.)$$

Dieser Zusammenhang definiert die Funktion:

$$\begin{aligned} &S : [0, \infty) \to [0, \infty) \\ &t \mapsto S(t) := \frac{1}{2} g t^2\,. \end{aligned}$$

6.) Sei

$$Q : [0, \infty) \to \mathcal{P}(\mathbb{R}) \quad \text{(Potenzmenge von } \mathbb{R}\text{)}$$

definiert durch

$$x \mapsto Q(x) := \left\{\sqrt{x}, -\sqrt{x}\right\} .$$

7.) Sei $G : \mathbb{R} \to \mathbb{Z}, \quad x \mapsto G(x) := [x] \quad := \quad$ größte ganze Zahl $\ \leq x$.

Ist $f : A \to B$ eine Abbildung, dann nennt man A den Definitionsbereich von f. B heißt Zielmenge, die Menge

$$W(f) = f(A) := \{y \mid \exists_{x \in A} : y = f(x)\} \subseteq B$$

heißt Wertebereich. Man sagt, „die Funktion f hängt von $x \in A$ ab“. Jedes $x \in A$ heißt auch „Argument“ von f. Für $x \in A$ nennt man $y = f(x) \in B$ das „Bild von x unter f“.

Seien $C \subseteq A$ und $D \subseteq B$. Dann heißt

$$f(C) := \{y \in B \mid \exists_{x \in C} : y = f(x)\} \subseteq B$$

das „Bild von C unter f“, und

$$f^{-1}(D) := \{x \in A \mid \exists_{y \in D} : y = f(x)\} \subseteq A$$

„das Urbild von D unter f“. Es gilt

$$\begin{aligned} C &\subseteq f^{-1}\big(f(C)\big) , && \text{für alle} \quad C \subseteq A \\ f\big(f^{-1}(D)\big) &\subseteq D , && \text{für alle} \quad D \subseteq B . \end{aligned}$$

Definition: Zwei Abbildungen f, g heißen gleich, genau dann wenn die Definitionsbereiche $D(f)$ und $D(g)$ der Abbildungen f und g übereinstimmen: $D(f) = D(g) = D$, und wenn für alle $x \in D$ gilt $f(x) = g(x)$.

Seien $A = D(f)$, $B = D(g)$ mit $A \subseteq B$, und für alle $x \in A$ gelte

$$f(x) = g(x) .$$

dann heißt g Forsetzung von f und umgekehrt f Einschränkung von g. Man schreibt auch

$$f = g\big|_A \quad \text{(Einschränkung von } g \text{ auf } A\text{)} .$$

Seien $f : A \to B_1, \quad g : B_2 \to C$ Abbildungen mit $B_1 \subseteq B_2$. Dann definiert man die Abbildung $g \circ f : A \to C$ durch

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

$g \circ f$ heißt Hintereinanderausführung von f und g.

Damit $g \circ f$ erklärt werden kann, muß natürlich der Wertebereich von f im Definitionsbereich von g enthalten sein.

Satz: f, g, h seien Abbildungen, so daß $g \circ f$ und $h \circ g$ erklärt sind. Dann sind auch die Abbildungen $h \circ (g \circ f)$ und $(h \circ g) \circ f$ erklärt, und es gilt

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

(Assoziativität der Hintereinanderausführung).

Beweis: $h \circ (g \circ f)$ kann definiert werden, wenn $W(g \circ f) \subseteq D(h)$ gilt. Dies ist richtig, denn

$$W(g \circ f) \subseteq W(g) \subseteq D(h)\,,$$

weil nach Voraussetzung $h \circ g$ definiert ist. Ebenso ist $(h \circ g) \circ f$ erklärt, wenn $W(f) \subseteq D(h \circ g)$ gilt. Dies ist richtig, wegen $W(f) \subseteq D(g) = D(h \circ g)$.

Sei nun $x \in D(h \circ (g \circ f))$. Dann gilt

$$\begin{aligned} [h \circ (g \circ f)](x) &= h\big(g \circ f(x)\big) = h\Big(g\big(f(x)\big)\Big) \\ &= h \circ g\big(f(x)\big) = [(h \circ g) \circ f](x)\,. \end{aligned}$$

■

Seien A, B Mengen, und id_A bzw. id_B die Identitäten auf A, B. Sei $f : A \to B$. Dann gilt

$$f \circ \mathrm{id}_A = f = \mathrm{id}_B \circ f\,.$$

id_A bzw. id_B sind also rechtsneutrale bzw. linksneutrale Elemente bezüglich der Hintereinanderausführung.

Definition: Seien A, B Mengen. $f : A \to B$ heißt

1.) injektiv: $:\iff \quad (f(x_1) = f(x_2) \Rightarrow x_1 = x_2)$

2.) surjektiv: $:\iff \quad W(f) = B$

3.) bijektiv: $:\iff \quad f$ ist injektiv und surjektiv.

Sei $f : A \to B$ bijektiv. Dann existiert die Umkehrabbildung $f^{-1} : B \to A$ zu f, die folgendermaßen definiert ist: Für alle $y \in B$ gilt $y \in f(A)$, weil f surjektiv ist, also existiert mindestens ein $x \in A$ mit $f(x) = y$. Da f injektiv ist, existiert genau ein solches x. Sei nun

$$f^{-1}(y) := x\,.$$

Man kann dies auch so schreiben: Sei $y \in W$

$$f^{-1}(y) = x \quad \Longleftrightarrow \quad y = f(x)\,.$$

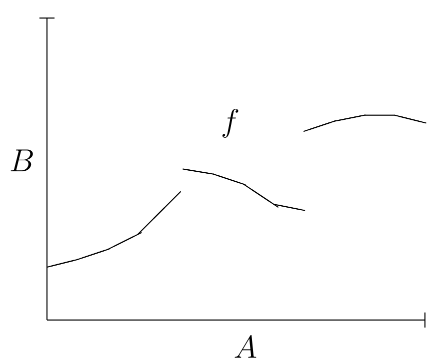
Natürlich sind die Abbildungen $f^{-1} \circ f : A \to A$ und $f \circ f^{-1} : B \to B$ erklärt, und es gilt

$$f^{-1} \circ f = \mathrm{id}_A\,, \quad f \circ f^{-1} = \mathrm{id}_B\,.$$

Ist $f : A \to B$ surjektiv, dann sagt man auch, f sei eine Abbildung von A auf B. Ist f injektiv, so sagt man auch, f sei eine eineindeutige Abbildung.

Man kann den nicht präzise definierten Begriff der Funktion auf den nicht präzise definierten Begriff der Menge zurückführen. Dies geht folgendermaßen:

Definition: Seien A, B Mengen. Eine Abbildung f von A in B ist eine Teilmenge von $A \times B$ mit der Eigenschaft: Zu jedem $x \in A$ gibt es genau ein $y \in B$ mit $(x, y) \in f$



Diese Teilmenge von $A \times B$ nennt man auch „Graph der Funktion f".

3b.) Reelle Funktionen

Eine Funktion, deren Definitionsbereich und Wertebereich Teilmengen der reellen Zahlen sind, heißt reelle Funktion. Ist nur der Wertebereich der Funktion Teilmenge der reellen Zahlen, dann heißt die Funktion reellwertig.

Beispiele:

1.) Ganzrationale Funktionen oder Polynome.

Seien $a_0, \dots, a_n \in \mathbb{R}$ mit $a_n \neq 0$, und sei die Funktion $p : \mathbb{R} \to \mathbb{R}$ definiert durch

$$x \mapsto p(x) = \sum_{m=0}^{n} a_m x^m\,.$$

p heißt Polynom oder ganzrationale Funktion vom Grad n (Beachte: $a_n \neq 0$).
In der Einführung wurde folgender Satz bewiesen:
Satz: Seien p_1, p_2 Polynome vom Grade n, und seien $x_0, \ldots, x_n \in \mathbb{R}$ paarweise verschiedene Zahlen. Gilt $p_1(x_i) = p_2(x_i)$ für $i = 0, \ldots, n$, dann stimmen die Koeffizienten des ersten Polynoms mit den Koeffizienten des zweiten Polynoms überein.

Korollar: Gilt $p_1 = p_2$ im Funktionssinn $(\forall_{x \in \mathbb{R}} : p_1(x) = p_2(x))$, dann stimmen die Koeffizienten der Polynome überein.

Die Summe $p_1 + p_2$ und das Produkt $p_1 p_2$ zweier Polynome sind wieder Polynome, und es gilt Grad $(p_1 \cdot p_2) =$ Grad p_1+ Grad p_2.

Ganzrationale Funktionen, deren Grad höchstens gleich 1 ist, heißen affine Funktionen:

$$x \mapsto a_1 x + a_0 .$$

Auch die konstanten Funktionen $x \mapsto a_0$ gehören dazu.

2.) Gebrochen rationale Funktionen
Sind g, h ganzrationale Funktionen, und sei M die Menge der Nullstellen von h. Sei $r : \mathbb{R} \backslash M \to \mathbb{R}$ definiert durch

$$x \to r(x) = \frac{g(x)}{h(x)} .$$

Dann heißt r „gebrochen rationale Funktion“ oder einfach „rationale Funktion“. Jede rationale Funktion kann man in der Form

$$r = \frac{g}{h} = p + \frac{s}{h}$$

schreiben, wobei p und s Polynome sind mit Grad $s <$ Grad h und mit

$$\text{Grad } p = \ (\text{Grad } g) \ - \ (\text{Grad } h) ,$$

falls Grad $g \geq$ Grad h. Ansonsten ist $p = 0$.

Diese Darstellung gewinnt man mit dem folgenden Divisionsalgorithmus. (Mit diesem Algorithmus kann man auch beweisen, daß eine solche Darstellung immer möglich ist.)

Beispiel: $g(x) = x^4 + x^3 - 2, \quad h(x) = x^2 - 1.$

$$\begin{array}{l}
(x^4 + x^3 - 2) : (x^2 - 1) = x^2 + x + 1 + \frac{x-1}{x^2-1} \\
\underline{-(x^4 - x^2)} \\
\qquad x^3 + x^2 - 2 \\
\quad \underline{-(x^3 - x)} \\
\qquad\quad x^2 + x - 2 \\
\qquad \underline{-(x^2 - 1)} \\
\qquad\qquad x - 1
\end{array}$$

also

$$\frac{x^4 + x^3 - 2}{x^2 - 1} = (x^2 + x + 1) + \frac{x - 1}{x^2 - 1} = (x^2 + x + 1) + \frac{1}{x + 1}\,,$$

wegen $x^2 - 1 = (x - 1)(x + 1)$.

Man sieht an diesem Beispiel, daß anders als bei den ganzrationalen Funktionen, eine rationale Funktion sich auf verschiedene Weisen darstellen läßt. Denn für die rationale Funktion $\frac{s}{h}$ aus diesem Beispiel gilt

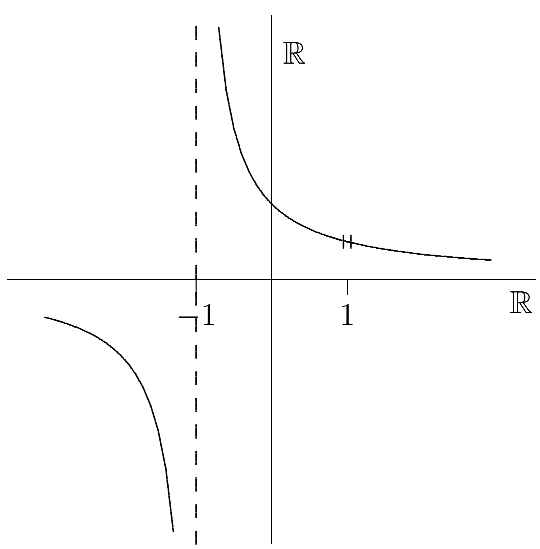
$$\frac{s(x)}{h(x)} = \frac{x - 1}{x^2 - 1} = \frac{x - 1}{(x - 1)(x + 1)} = \frac{1}{x + 1}\,.$$

Man beachte aber, daß $x^2 - 1$ die Nullstellen $x = \pm 1$ hat. Nach unserer Definition ist also $f_1(x) := \frac{x-1}{x^2-1}$ auf $R\backslash\{-1, 1\}$ definiert:

$$f_1 : \mathbb{R}\backslash\{-1, 1\} \to \mathbb{R}\,.$$

Dagegen hat $x + 1$ nur die Nullstelle $x = -1$. Also ist $f_2(x) := \frac{1}{x+1}$ auf $\mathbb{R}\backslash\{-1\}$ definiert:

$$f_2 : \mathbb{R}\backslash\{-1\} \to \mathbb{R}\,.$$

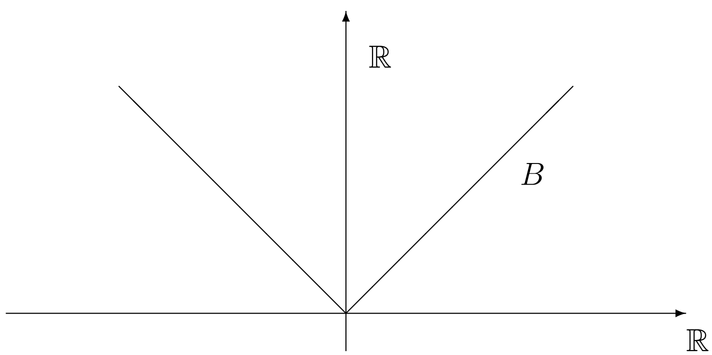


Wegen $D(f_1) \subseteq D(f_2)$ und wegen $f_1(x) = f_2(x)$ für alle $x \in D(f_1)$ ist also f_2 eine Fortsetzung von f_1. Weil der Zähler $x-1$ von f_1 an der Stelle $x=1$ eine Nullstelle hat, wird die Nullstelle $x=1$ des Nenners x^2-1 „kompensiert“. f_1 kann „stetig“ zu f_2 fortgesetzt werden.

3.) Betragsfunktion

Sei $B : \mathbb{R} \to \mathbb{R}$ definiert durch

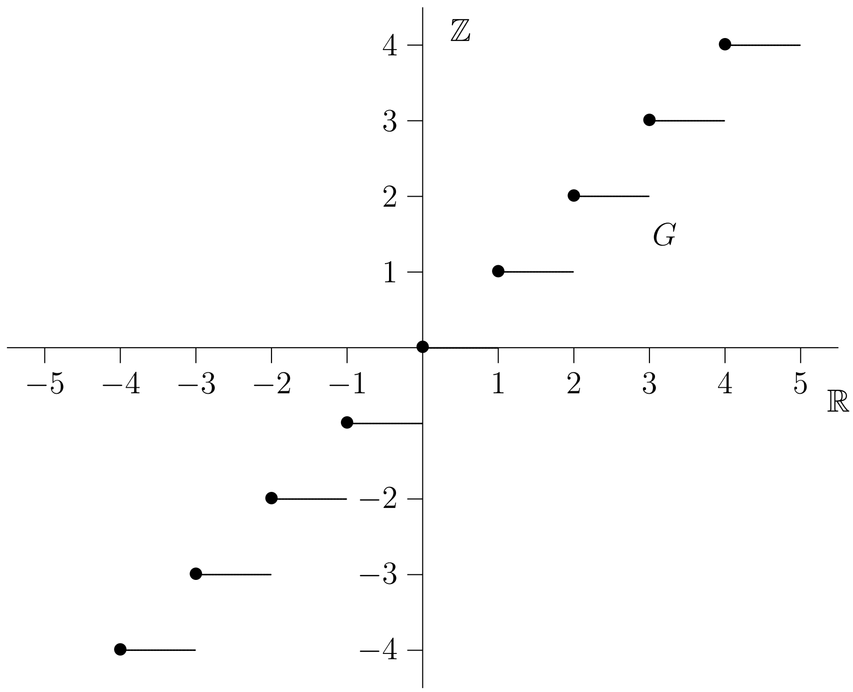
$$x \mapsto B(x) := |x| = \begin{cases} x\,, & \text{für} \quad x \geq 0 \\ -x\,, & \text{für} \quad x \leq 0 \end{cases}$$



B kann auch als Funktion $\hat{B} : \mathbb{R} \to [0, \infty)$ aufgefaßt werden. $\hat{B}$ ist surjektiv, aber nicht injektiv.

4.) „Größte–Ganze–Funktion“

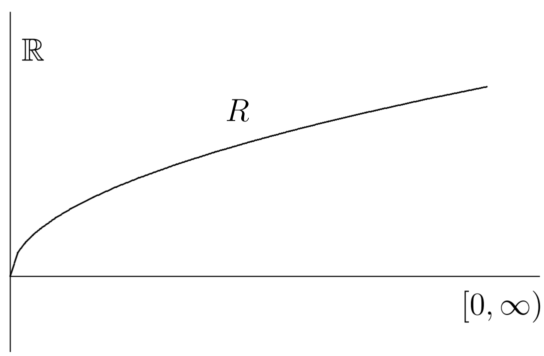
Als Beispiel am Anfang dieses Abschnittes habe ich schon die „Größte–Ganze–Funktion“ eingeführt: Sei $G : \mathbb{R} \to \mathbb{R}$ definiert durch $x \mapsto G(x) := [x] :=$ größte ganze Zahl, die kleiner oder gleich als x ist.



Es gilt $W(G) = \mathbb{Z}$. $G : \mathbb{R} \to \mathbb{R}$ ist weder surjektiv noch injektiv. Es gilt beispielsweise $G^{-1}(\{0\}) = [0, 1)$.

5.) Wurzelfunktion

Auch die Wurzelfunktion $W : [0, \infty) \to \mathbb{R}$, $\quad x \mapsto w(x) := \sqrt{x}$ wurde schon betrachet.



Es gilt $W(W) = [0, \infty)$. Denn nach Definition gilt $W(W) \subseteq [0, \infty)$. Zu zeigen ist also:

$$[0, \infty) \subseteq W(W).$$

Hier sei $a \in [0, \infty)$. Dann ist auch $a^2 \in [0, \infty) = D(W)$, und $W(a^2) = \sqrt{a^2} = a$, also ist $a \in W(W)$, also $[0, \infty) \subseteq W(W)$, also

$$[0, \infty) = W(W).$$

$\blacksquare$

Die Funktion W ist also injektiv, aber nicht surjektiv. Dagegen ist $\hat{W} : [0,\infty) \to [0,\infty)$, $x \mapsto \hat{W}(x) := \sqrt{x}$, nach dem eben Bewiesenen surjektiv und injektiv, also existiert $\hat{W}^{-1}$. Natürlich gilt $\hat{W}^{-1} : [0,\infty) \to [0,\infty)$, $\quad x \mapsto \hat{W}^{-1}(x) := x^2$.

6.) Dirichlet Funktion

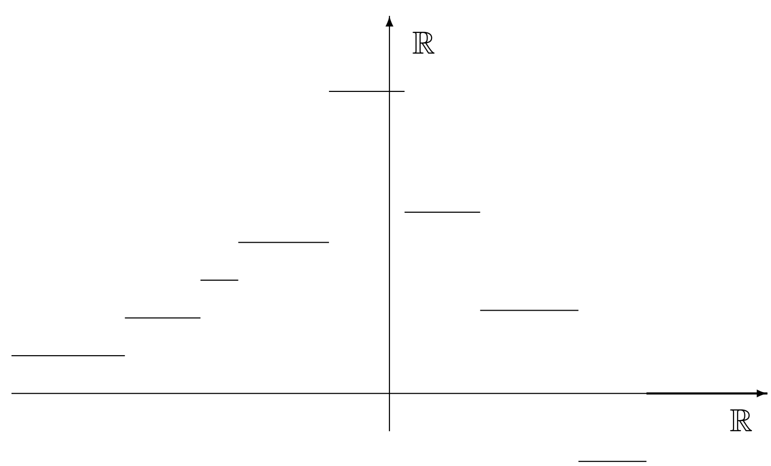
Sei $f : \mathbb{R} \to \mathbb{R}$ definiert durch

$$f(x) := \begin{cases} 1\,, & x \in \mathbb{Q} \\ 0\,, & x \notin \mathbb{Q} \end{cases} .$$

Also ist $W(f) = \{0,1\}$. Diese Dirichlet Funktion kann nicht gezeichnet werden, weil es zwischen zwei rationalen Zahlen immer noch eine irrationale, und zwischen zwei irrationalen Zahlen immer noch eine rationale gibt.

7.) Treppenfunktion

Eine Funktion $t : \mathbb{R} \to \mathbb{R}$ heißt Treppenfunktion, wenn ihr Wertebereich $W(t)$ eine endliche Menge ist, und das Urbild $t^{-1}(\{y\})$ eine Vereinigung von endlich vielen Intervallen ist für jedes $y \in \mathbb{R}$.



3c.) Folgen, Abzählbarkeit

Definition: Abbildungen mit dem Definitionsbereich $\mathbb{N}$ heißen Folgen.

Die Bildmenge von Folgen kann irgendeine Menge sein. Ist die Bildmenge die Menge der reellen Zahlen, dann spricht man von Zahlenfolgen. Folgen bezeichnet man folgendermaßen:

$$\{a_1, a_2, a_3, \ldots\} = \{a_n\}_{n\in\mathbb{N}} .$$

Hierbei ist a_n das Bild der Zahl $n \in \mathbb{N}$. Man bezeichnet a_n auch als n–tes Element der Folge $\{a_n\}_{n\in\mathbb{N}}$. Diese Bezeichnung weicht von den üblichen Bezeichnungen bei Funktionen ab. Man verwechsele nicht die Folge $\{a_n\}_{n\in\mathbb{N}}$ mit der Menge $\{a_n \mid n \in \mathbb{N}\}$, dem Wertebereich der Folge.

Einfache Beispiele für Zahlenfolgen:

$$\{0,0,0,\ldots\}, \quad \{0,1,0,1,0,1,\ldots\}, \quad \{n\}_{n\in\mathbb{N}}, \quad \{\frac{1}{n}\}_{n\in\mathbb{N}}, \quad \{\frac{-1}{n^2+1}\}_{n\in\mathbb{N}} .$$

Folgen werden oft rekursiv definiert:

Beispiel. Sei die Folge $\{x_n\}_{n\in\mathbb{N}}$ definiert durch

1.) $x_1 = 1$

2.) $x_{n+1} = \frac{1}{2} \cdot (x_n + \frac{2}{x_n}) \quad$ für $\quad n \geq 1$.

Intuitiv ist klar, daß durch diese Definition eine eindeutig bestimmte Folge definiert wird. Tatsächlich muß aber bewiesen werden, daß es genau eine Folge gibt, die dieser Definition genügt. Wir verzichten auf diesen Beweis.

Mit Folgen definiert man den Begriff der Abzählbarkeit: Sei $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$.

Definition: Für $n \in \mathbb{N}_0$ sei $A_n = \{k \mid k < n,\ k \in \mathbb{N}_0\}$ der zu n gehörende Abschnitt von $\mathbb{N}_0$. $(A_0 = \emptyset)$.

Definition: Eine Menge M heißt endlich, wenn es eine Zahl $n \in \mathbb{N}_0$ gibt derart, daß eine bijektive Abbildung von A_n auf M existiert. Die Zahl n heißt „die Anzahl der Elemente von M“.

Eine Menge M heißt abzählbar unendlich, wenn eine bijektive Abbildung von $\mathbb{N}_0$ auf M existiert.

Man sagt, zwei Mengen M_1 und M_2 „seien von gleicher Mächtigkeit“, wenn es eine bijektive Abbildung von M_1 auf M_2 gibt.

Also kann man auch sagen, eine Menge M sei abzählbar unendlich, wenn sie von gleicher Mächtigkeit wie die Menge der natürlichen Zahlen ist (da $\mathbb{N}$ und $\mathbb{N}_0$ gleiche Mächtigkeit haben).

Definition: Eine Menge heißt abzählbar, wenn sie endlich ist oder abzählbar unendlich ist. Im anderen Fall heißt sie überabzählbar.

Die erste Definition ist sinnvoll wegen des folgenden Satzes:

Satz: Seien $n, m \in \mathbb{N}_0$ mit $n > m$. Dann gibt es keine injektive Abbildung $f : A_n \to A_m$.

Hieraus folgt dann auch, daß eine Menge M nicht gleichzeitig endlich und abzählbar unendlich sein kann. Denn sonst gäbe es $m \in \mathbb{N}$ und bijektive Abbildungen $f : A_m \to M$, $g : \mathbb{N}_0 \to M$, also wäre $f^{-1} \circ g : \mathbb{N}_0 \to A_m$ eine bijektive Abbildung. Für $n > m$ wäre dann

$$(f^{-1} \circ g)\big|_{A_n} : A_n \to A_m$$

eine injektive Abbildung, im Widerspruch zum Satz.

Beweis des Satzes:

Angenommen, die Behauptung sei falsch. Dann gibt es mindestens einen Abschnitt A_n, der injektiv in einen Abschnitt A_m mit $m < n$ abgebildet werden kann. Dann kann A_n auch injektiv in A_{n-1} abgebildet werden. Weil $n - 1 \in \mathbb{N}_0$ gilt, ist $n \in \mathbb{N}$. Die Menge aller $n \in \mathbb{N}$ so daß A_n injektiv in A_{n-1} abgebildet werden kann, ist also nicht leer. In dieser Menge gibt es also ein kleinstes Element k. Wir zeigen, daß hieraus folgt, daß A_{k-1} injektiv in A_{k-2} abgebildet werden kann, und erhalten einen Widerspruch. Sei $f : A_k \to A_{k-1}$ die injektive Abbildung.

Es gibt drei Fälle:

(i) Die Zahl $k - 2$ tritt nicht als Bild auf bei der Abbildung f.

(ii) Die Zahl $k - 2$ ist Bild von $k - 1$ bei der Abbildung f.

(iii) Die Zahl $k - 2$ ist Bild einer Zahl j mit $j \leq k - 2$ bei der Abbildung f.

In den beiden ersten Fällen ist $f\big|_{A_{k-1}} : A_{k-1} \to A_{k-2}$ eine injektive Abbildung. Im letzten Fall wird eine injektive Abbildung $g : A_{k-1} \to A_{k-2}$ definiert durch

$$g(\ell) = \begin{cases} f(\ell)\,, & \ell \in A_{k-1}, \ell \neq j \\ f(k-1)\,, & \ell = j\,. \end{cases}$$

Damit erhalten wir den behaupteten Widerspruch.

Satz: Die Menge $\mathbb{Q}$ der rationalen Zahlen ist abzählbar.

Beweis: Man ordnet die positiven rationalen Zahlen in ein doppelt unendliches Schema:

$$\begin{array}{cccccccc}
\frac{1}{1} & \rightarrow & \frac{2}{1} & & \frac{3}{1} & & \frac{4}{1} & \dots \\
 & \swarrow & & \swarrow & & \swarrow & & \swarrow \\
\frac{1}{2} & & \frac{2}{2} & & \frac{3}{2} & & \frac{4}{2} & \dots \\
 & \swarrow & & \swarrow & & \swarrow & & \swarrow \\
\frac{1}{3} & & \frac{2}{3} & & \frac{3}{3} & & \frac{4}{3} & \dots \\
 & \swarrow & & \swarrow & & \swarrow & & \swarrow \\
\frac{1}{4} & & \frac{2}{4} & & \frac{3}{4} & & \frac{4}{4} & \dots \\
 & \swarrow & & \swarrow & & \swarrow & & \\
\vdots & & \vdots & & \vdots & & \vdots &
\end{array}$$

Auf diese Weise erhält jede positive rationale Zahl eine Nummer. Sei $\{q_n\}_{n\in\mathbb{N}}$ diese Folge. Alle rationalen Zahlen enthält dann die Folge $\{0, q_1, -q_1, q_2, -q_2, q_3, -q_3, \dots\}$.

Satz: Die Menge aller reellen Zahlen ist überabzählbar.

Beweis: Zum Beweis nimmt man an, die Menge $\mathbb{R}$ sei abzählbar. Dann können ihre Elemente in einer Folge geordnet werden (durchnumeriert werden):

$$\{x_1, x_2, x_3, x_4, \dots\}.$$

Man konstruiert nun eine Zahl, die sicher nicht zu dieser Folge gehört, so daß diese Folge doch nicht alle Elemente umfassen kann. Also hat man einen Widerspruch zur Annahme erhalten, und $\mathbb{R}$ muß überabzählbar sein.

Hierzu wähle ein Intervall $I_1 = [a_1, b_1]$ mit $a_1 < b_1$, das x_1 nicht enthält, und zerlege I_1 in drei gleiche abgeschlossene Teilintervalle. x_2 kann nicht in allen drei Teilintervallen zugleich liegen. Wähle aus diesen drei Teilintervallen ein Intervall $I_2 = [a_2, b_2]$ aus, das x_2 nicht enthält. Fahre so fort. Dies ergibt eine Folge von Intervallen $\{I_n\}_{n\in\mathbb{N}}$ mit

$$I_1 \supseteq I_2 \supseteq I_3 \supseteq \dots$$

und mit $x_n \notin I_n$, für alle $n \in \mathbb{N}$. Die Folge $\{a_n\}_{n\in\mathbb{N}}$ ist nach oben beschränkt; jedes b_n ist obere Schranke. Also existiert das Supremum s der Menge $\{a_n \mid n \in \mathbb{N}\}$. s gehört zu jedem der Intervalle I_n. Wäre das nicht so, müßte es $n_0 \in \mathbb{N}$ geben mit $s \notin I_{n_0}$, also $s \notin I_n$, für alle $n \geq n_0$, also $s > b_n$, für alle $n \geq n_0$, weil nach Konstruktion für alle n die Ungleichung $s \geq a_n$ gilt. Da aber jedes b_n obere Schranke ist, könnte s nicht kleinste obere Schranke sein. Also gilt $s \in I_n$ für alle $n \in \mathbb{N}$. Andererseits gilt $x_n \notin I_n$, für alle $n \in \mathbb{N}$, also muß $s \neq x_n$ gelten für alle $n \in \mathbb{N}$. $\blacksquare$

3d.) Vektorräume reeller Funktionen

Sei D eine Menge und sei $F(D) = F(D, \mathbb{R})$ die Menge aller reellwertigen Funktionen $f : D \to \mathbb{R}$.

Auf der Menge $F(D)$ erklärt man eine Verknüpfung von Elementen aus $F(D)$, die „Addition" von Funktionen durch

$$(f+g)(x) := f(x) + g(x)\,, \quad \forall x \in D\,.$$

Mit dieser Verknüpfung ist $F(D)$ eine kommutative Gruppe, d.h. die Gruppenaxiome (A1) – (A4) sind erfüllt: Es gilt

V1.) $f + g = g + f$

V2.) $f + (g + h) = (f + g) + h$

V3.) Es gibt genau ein neutrales Element, nämlich die Funktion $f : D \to \mathbb{R}$, $x \mapsto f(x) := 0$. Man bezeichnet diese Funktion mit 0.

V4.) Zu jedem f gibt es genau ein g, so daß $f + g = 0$ gilt, nämlich die Funktion

$$x \mapsto g(x) := -f(x)\,.$$

Man bezeichnet diese Funktion mit $-f$.

Man definiert nun noch eine Verknüpfung zwischen Elementen von $F(D)$ und reellen Zahlen. Ist $a \in \mathbb{R}$ und $f \in F(D)$ so definiert man das Produkt $af \in F(D)$ durch

$$x \mapsto (af)(x) := a \cdot f(x)\,.$$

Für diese Multiplikation gelten die Regeln:

V5.) $(a + b)f = af + bf\,, \quad a, b \in \mathbb{R}\,, \quad f \in F(D)$

V6.) $a(f + g) = af + ag \qquad a \in \mathbb{R}\,, \quad f, g \in F(D)$

V7.) $a(bf) = (ab)f$

V8.) $1 \cdot f = f$

Eine Menge, auf der die Addition erklärt ist so daß (V1) – (V4) gelten, und auf der die Multiplikation mit reellen Zahlen erklärt ist so daß (V5) – (V8) gelten, bezeichnet man als (reellen) *Vektorraum*, oder Vektorraum über $\mathbb{R}$. Zum Vektorraum gehört also eine Menge und zwei Verknüpfungen. Somit ist ein Vektorraum ein Trippel $(V, +, \cdot)$. (Beachte, daß $+, \cdot$ Funktionen sind und wie beschrieben durch Mengen erklärt werden können.)

$(F(D), +, \cdot)$ ist also ein reeller Vektorraum. Man bezeichnet $(U, +, \cdot)$ als Untervektorraum eines Vektorraumes $(V, +, \cdot)$, wenn $U \subseteq V$ gilt, und wenn U mit den durch $(V, +, \cdot)$ auf

U induzierten Verknüpfungen ein Vektorraum ist.

In der Analysis spielen eine Reihe von Untervektorräumen des Raumes $F(D)$ eine wichtige Rolle. Ich betrachte einige einfache Beispiele. Weitere Untervektorräume werden später eingeführt.

1.) Die Menge der beschränkten Funktionen $f : D \to \mathbb{R}$ ist ein Untervektorraum von $F(D)$. Hierbei heißt eine Funktion beschränkt, genau dann wenn

$$\exists_{C>0} \forall_{x \in D} : |f(x)| \leq C .$$

gilt.

Um zu zeigen, daß dies ein Untervektorraum ist, genügt es zu zeigen, daß mit f, g auch $f+g$ sowie mit $a \in \mathbb{R}$ auch af beschränkte Funktionen sind.

Sei

$$\forall_{x \in D} : |f(x)| \leq C_1 , \quad |g(x)| \leq C_2 .$$

Dann folgt:

$$\forall_{x \in D} : |f(x) + g(x)| \leq |f(x)| + |g(x)| \leq C_1 + C_2 .$$

Also ist $f+g$ beschränkt. Außerdem folgt

$$\forall_{x \in D} : |af(x)| \leq |a| \, |f(x)| \leq |a| \cdot C_1 .$$

Also ist auch af beschränkt.

Wählt man $D = \mathbb{N}$, dann erhält man den Vektorraum der beschränkten Zahlenfolgen.

2.) Die Menge der ganzrationalen Funktionen ist ein Untervektorraum von $F(\mathbb{R})$. Denn seien

$$x \mapsto p_1(x) = \sum_{m=0}^{n} a_m x^m , \quad x \mapsto p_2(x) = \sum_{m=0}^{k} b_m x^m$$

Polynome. Dann ist auch

$$(p_1 + p_2)(x) = p_1(x) + p_2(x) = \sum_{m=0}^{\max(n,k)} (a_m + b_m) x^m$$

ein Polynom. Hierbei setzt man

$$\begin{aligned} a_m &= 0 , \quad \text{für} \quad n < m \leq \max(n,k) \\ b_m &= 0 , \quad \text{für} \quad k < m \leq \max(n,k) . \end{aligned}$$

Auch das Produkt

$$(ap_1)(x) = ap_1(x) = a\sum_{m=0}^{n} a_m x^m = \sum_{m=0}^{n}(aa_m)x^m$$

ist ein Polynom.
Die Menge der Polynome, deren Grad höchstens gleich n ist, ist ebenfalls ein Untervektorraum von $F(\mathbb{R})$. Die Dimension dieses Raumes ist $n+1$. (Die Dimension eines Vektorraumes wird in der Vorlesung „Lineare Algebra" definiert.)

3.) Die Menge aller Treppenfunktionen ist ein Untervektorraum von $F(\mathbb{R})$. Denn seien t_1, t_2 Treppenfunktionen. Sei $c \in W(t_1+t_2)$. Dann gibt es $x \in \mathbb{R}$ mit $c = (t_1+t_2)(x) = t_1(x) + t_2(x)$. Also gilt

$$c \in \left\{ b \mid \exists_{c_1 \in W(t_1)} \exists_{c_2 \in W(t_2)} : b = c_1 + c_2 \right\}$$

Dies ist eine endliche Menge. (Zur Übung konstruiere man eine bijektive Abbildung von einem geeigneten Abschnitt von $\mathbb{N}$ auf diese Menge.)
Außerdem ist $(t_1+t_2)^{-1}(\{c\})$ eine endliche Vereinigung von Intervallen für jedes $c \in W(t_1+t_2)$. Denn seien $c_1 \in W(t_1), c_2 = W(t_2)$ mit $c = c_1 + c_2$ und seien $I_1^{(1)}, \ldots, I_n^{(1)}; I_1^{(2)}, \ldots, I_m^{(2)}$ Intervalle mit

$$t_1^{-1}(\{c_1\}) = \bigcup_{i=1}^{n} I_i^{(1)}, \quad t_2^{-1}(\{c_2\}) = \bigcup_{i=1}^{m} I_i^{(2)}.$$

Für jedes x aus der endlichen Vereinigung von Intervallen

$$M_{(c_1,c_2)} = \left(\bigcup_{i=1}^{n} I_i^{(1)}\right) \cap \left(\bigcup_{j=1}^{m} I_j^{(2)}\right) = \bigcup_{\substack{i=1,\ldots,n \\ j=1,\ldots,m}} (I_i^{(1)} \cap I_j^{(2)})$$

gilt dann

$$x \in t_1^{-1}(\{c_1\}) \cap t_2^{-1}(\{c_2\}),$$

also $(t_1+t_2)(x) = t_1(x) + t_2(x) = c_1 + c_2 = c$.

Diese endliche Vereinigung von Intervallen, die zur Wertekombination (c_1, c_2) gehört, ist also eine Teilmenge des Urbildes von $\{c\}$). Das ganze Urbild von $\{c\}$ ist

$$M = \bigcup_{c_1+c_2=c} M_{(c_1,c_2)}.$$

Da es nur endlich viele Wertepaare (c_1, c_2) mit $c_1 + c_2 = c$ gibt, ist also auch M endliche Vereinigung von Intervallen.

Ebenso ist der Wertebereich von $at_1 \quad (a \in \mathbb{R})$ endlich: $W(at_1) = \{b \mid \exists_{c \in W(t_1)} : b = ac\}$.

Die Dimension des Vektorraumes aller Treppenfunktionen ist unendlich.

Auf dem Vektorraum $F(D)$ kann man auch die Multiplikation zweier Funktionen $f, g : D \to \mathbb{R}$ erklären:

$$\forall_{x \in D} : (fg)(x) := f(x) \cdot g(x) .$$

Diese Multiplikation ist assoziativ und kommutativ. Die konstante Funktion $x \mapsto 1$ ist das neutrale Element. Man beachte aber, daß zu einer Funktion $f \in F(D)$ nicht unbedingt ein inverses Element in $F(D)$ existiert, nämlich dann nicht, wenn f Nullstellen hat, d.h. wenn es $x \in D$ gibt mit $f(x) = 0$. Für die Multiplikation gilt auch das Distributivgesetz.

Also ist $F(D)$ eine kommutative, assoziative Algebra mit Einselement.

Diese Algebra ist nicht nullteilerfrei, denn man kann leicht Funktionen $f, g \in F(D)$ mit $f \neq 0, \quad g \neq 0$ konstruieren, für die aber

$$fg = 0$$

ist.

Alle oben angegebenen Beispiele sind auch Unteralgebren von $F(D)$, bis auf den Vektorraum der Polynome höchstens n–ten Grades. (Die Begriffe „Algebra“, „Unteralgebra“, „nullteilerfrei“ werden in der Vorlesung „Lineare Algebra“ definiert.)

3e.) Einige einfache Eigenschaften reeller Funktionen

Definition: Eine reelle Funktion f heißt nach oben beschränkt, genau dann wenn die Wertemenge von f nach oben beschränkt ist.
Sie heißt nach unten beschränkt, genau dann wenn die Wertemenge nach unten beschränkt ist.
Also: f ist beschränkt, genau dann wenn f nach oben und unten beschränkt ist.

Definition: Eine reelle Funktion $f : D \to \mathbb{R}$ heißt monton wachsend, genau dann wenn gilt:

$$\forall_{x_1, x_2 \in D} : x_1 < x_2 \Rightarrow f(x_1) \leq f(x_2) .$$

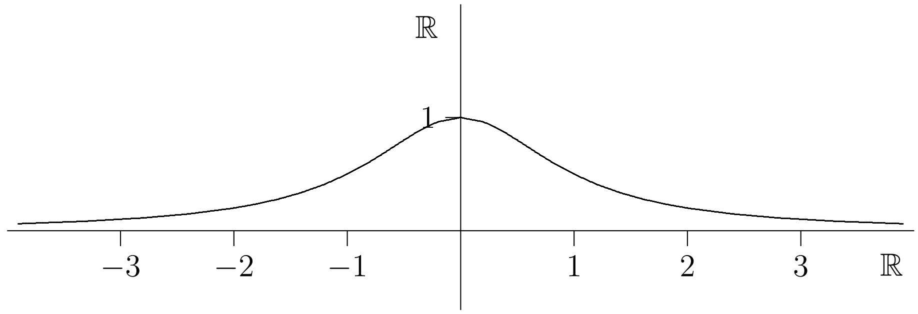
f heißt streng monoton wachsend, genau dann wenn gilt:

$$\forall_{x_1, x_2 \in D} : x_1 < x_2 \Rightarrow f(x_1) < f(x_2) .$$

Ebenso definiert man monoton fallend und streng monoton fallend. Eine Funktion heißt (streng) monoton, wenn sie (streng) monoton wachsend oder fallend ist.

Beispiele:

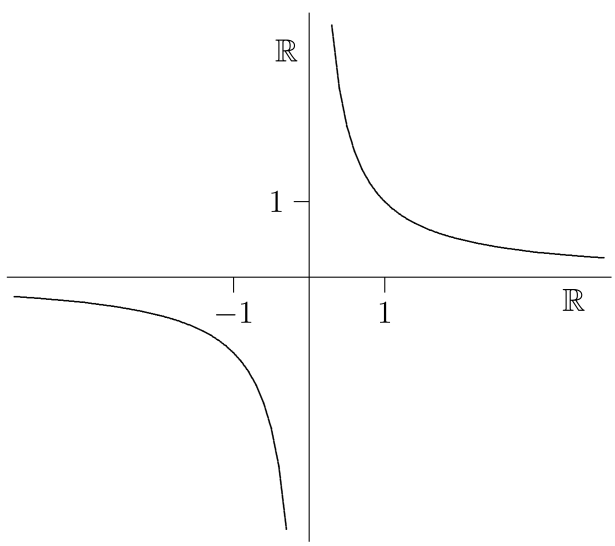
1.) $x \mapsto \dfrac{1}{1+x^2}\,, \quad x \in \mathbb{R}\,.$



Diese Funktion ist beschränkt, aber weder monoton fallend noch monoton wachsend. Jedoch ist die Einschränkung dieser Funktion auf $[0,\infty)$ streng monoton fallend. Denn sei $0 \leq x_1 < x_2$. Es gilt

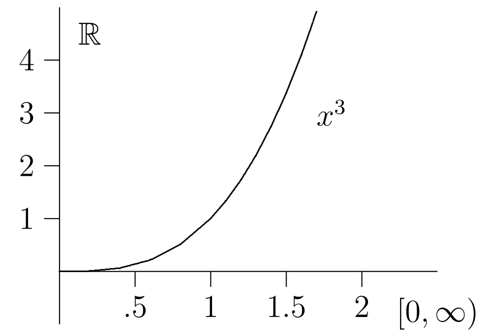
$$x_1^2 < x_2^2 \Rightarrow 1 + x_1^2 < 1 + x_2^2 \Rightarrow \frac{1}{1+x_1^2} > \frac{1}{1+x_2^2}$$

2.) $x \mapsto \dfrac{1}{x}\,, \quad x \in \mathbb{R}\backslash\{0\}\,.$



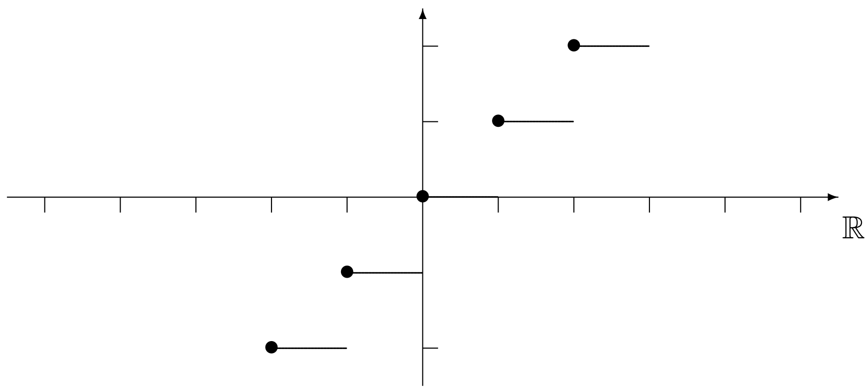
Diese Funktion ist nicht beschränkt. Die Einschränkung dieser Funktion auf $(0,\infty)$ ist jedoch streng monoton fallend, und auch die Einschränkung auf $(-\infty,0)$.

3.) $x \to x^n : [0,\infty) \to \mathbb{R}\,, \quad n \in \mathbb{N}\,.$



Diese Funktion ist nicht beschränkt, aber streng monoton wachsend. (Zur Übung beweise man dies durch Induktion.) Der Wertebereich ist ganz $[0,\infty)$.

4.) $x \mapsto [x] : \mathbb{R} \to \mathbb{Z}$.



Diese Funktion ist unbeschränkt, monoton wachsend aber nicht streng monoton wachsend. Sie ist auch nicht injektiv. Es gilt aber:

Satz: Sei f eine reelle, streng monotone Funktion. Dann ist f injektiv, also existiert auf $W(f)$ die Umkehrfunktion $f^{-1} : W(f) \to D(f)$. f^{-1} ist ebenfalls streng monoton.

Beweis: O.B.d.A. sei f streng monoton wachsend. Ist $x_1 \neq x_2$, dann muß entweder $x_1 < x_2$ oder $x_1 > x_2$ sein, also entweder $f(x_1) < f(x_2)$ oder $f(x_1) > f(x_2)$, also $f(x_1) \neq f(x_2)$. Somit ist f injektiv und es existiert $f^{-1} : W(f) \to D(f)$. Auch f^{-1} ist streng monoton wachsend. Denn sonst würde $y_1, y_2 \in W(f)$ existieren mit $y_1 < y_2$ und $f^{-1}(y_1) \geq f^{-1}(y_2)$. Hieraus würde wegen der strengen Monotonie von f folgen

$$y_1 = f\big(f^{-1}(y_1)\big) \geq f(\big(f^{-1}(y_2)\big) = y_2\,,$$

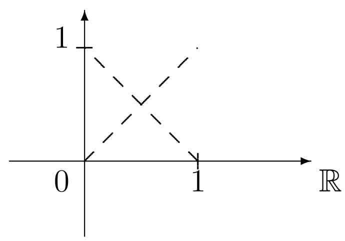
im Widerspruch zu $y_1 < y_2$. ■

Die strenge Monotonie ist eine hinreichende, aber keine notwendige Bedingung für die Umkehrbarkeit einer reellen Funktion. Dies sieht man an folgendem

Beispiel: Sei $D = [0,1]$. $f : D \to \mathbb{R}$ sei definiert durch

$$x \to f(x) := \begin{cases} x\,, & x \in \mathbb{Q} \\ 1-x\,, & x \notin \mathbb{Q} \end{cases} .$$

Diese Funktion bildet $[0,1]$ bijektiv auf $[0,1]$ ab, ist jedoch in keinem Teilintervall monoton.



Anwendung des Satzes: Da $x \mapsto x^n : [0,\infty) \to [0,\infty)$ streng monoton und surjektiv ist, existiert die inverse Funktion

$$x \mapsto x^{1/n} : [0,\infty) \to [0,\infty)\,, \quad n \in \mathbb{N}\,,$$

die auch streng monoton ist.

Die allgemeine Potenz mit rationalen Exponenten $q = \frac{m}{n} > 0$ kann man auf zwei Weisen erklären:

$$x^{\frac{m}{n}} = (x^{\frac{1}{n}})^m \quad \text{oder} \quad x^{\frac{m}{n}} = (x^m)^{\frac{1}{n}}\,.$$

Beide Definitionen stimmen überein. Denn es gilt

$$x^m = \left[\left(x^{\frac{1}{n}}\right)^n\right]^m = \left(x^{\frac{1}{n}}\right)^{n\cdot m} = \left[\left(x^{\frac{1}{n}}\right)^m\right]^n,$$

also

$$\left(x^m\right)^{\frac{1}{n}} = \left(x^{\frac{1}{n}}\right)^m.$$

Also ist x^q für positives rationales q und $x \geq 0$ erklärt. Für negatives rationales q setzt man

$$x^q := \frac{1}{x^{-q}}\,, \quad x > 0\,.$$

Außerdem definiert man

$$x^0 := 1 \quad \text{für} \quad x \geq 0\,.$$

Mit diesen Definitionen gilt

$$\begin{aligned} x^r \cdot x^s &= x^{r+s} \\ x^r y^r &= (x \cdot y)^r \\ (x^r)^s &= x^{rs}\,. \end{aligned}$$

4 Konvergente Folgen

4a.) Grundlegende Definitionen und Eigenschaften

Eine Abbildung

$$f : \mathbb{N} \to \mathbb{R}$$

heißt Zahlenfolge. Als Beispiel betrachte man die Folge

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

Alle Glieder dieser Folge sind größer als Null, aber intuitiv meint man, daß die Folgenglieder sich immer mehr der Null annähern mit wachsendem n. Anders ist dies bei der Folge

$$\left\{(-1)^n \frac{n}{n+1}\right\}_{n\in\mathbb{N}} = \left\{-\frac{1}{2}, \frac{2}{3}, -\frac{3}{4}, \frac{4}{5}, \dots\right\}.$$

Diese intuitiven Vorstellungen sollen präzisiert werden:

Definition: Eine Zahlenfolge $\{x_n\}_{n\in\mathbb{N}}$ heißt konvergent gegen die Zahl $a \in \mathbb{R}$, wenn es zu jeder Zahl $\varepsilon > 0$ eine Zahl $n_0 \in \mathbb{N}$ gibt so daß für alle $n \in \mathbb{N}$ mit $n \geq n_0$ gilt

$$|a - x_n| < \varepsilon .$$

a heißt Grenzwert der Folge $\{x_n\}_{n\in\mathbb{N}}$.

Mit Quantoren läßt sich dies auch folgendermaßen schreiben: $\{x_n\}_{n\in\mathbb{N}}$ heißt konvergent gegen $a \in \mathbb{R}$, genau dann wenn

$$\forall_{\varepsilon>0} \quad \exists_{n_0\in\mathbb{N}} \quad \forall_{\substack{n\in\mathbb{N} \\ n\geq n_0}} \quad : \quad |a - x_n| < \varepsilon .$$

Führt man den Begriff der „Umgebung“ ein, so läßt sich diese Definition auch etwas anschaulicher fassen:

Definition: Sei $\varepsilon > 0$ und $a \in \mathbb{R}$. Als ε–Umgebung von a bezeichnet man die Menge

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

Hiermit läßt sich die Definition der Konvergenz folgendermaßen formulieren:

Eine Zahlenfolge $\{x_n\}_{n\in\mathbb{N}}$ heißt konvergent gegen den Grenzwert $a \in \mathbb{R}$, wenn es zu jeder Umgebung U von a eine Zahl $n_0 \in \mathbb{N}$ gibt, so daß für alle $n \in \mathbb{N}$ mit $n \geq n_0$ gilt

$$x_n \in U .$$

Mit Quantoren:

$$\forall_{\substack{\text{Umgebung} \\ U \text{ von } a}} \quad \exists_{n_0\in\mathbb{N}} \quad \forall_{\substack{n\in\mathbb{N} \\ n\geq n_0}} : \quad x_n \in U .$$

Wenn eine Zahlenfolge $\{x_n\}_{n\in\mathbb{N}}$ gegen den Grenzwert a konvergiert, dann schreibt man

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

Gilt für eine Zahlenfolge $\{x_n\}_{n\in\mathbb{N}}$

$$\lim_{n\to\infty} x_n = 0\,,$$

dann heißt $\{x_n\}_{n\in\mathbb{N}}$ Nullfolge. Aus der Konvergenzdefinition folgt sofort:
Eine Zahlenfolge $\{x_n\}_{n\in\mathbb{N}}$ konvergiert gegen $a\in\mathbb{R}$, genau dann wenn $\{a-x_n\}_{n\in\mathbb{N}}$ Nullfolge ist.

Satz: Jede Folge besitzt höchstens einen Grenzwert.

Beweis: Seien a, b Grenzwerte von $\{x_n\}_{n\in\mathbb{N}}$. Sei $\varepsilon > 0$ beliebig. Nach Voraussetzung gibt es Zahlen $n_0, n_1 \in \mathbb{N}$ mit $|a-x_n| < \varepsilon$ für $n \geq n_0$ und $|b-x_n| < \varepsilon$ für $n \geq n_1$, also gilt $|a-x_n| < \varepsilon\,, \quad |b-x_n| < \varepsilon$ für $n \geq \max(n_0, n_1)$. Für solche n gilt daher

$$|a-b| = |(a-x_n)-(b-x_n)| \leq |a-x_n| + |b-x_n| < 2\varepsilon\,,$$

also folgt $a-b=0$, weil ε beliebig klein gewählt werden kann, also $a=b$. ■

Eine Zahlenfolge heißt konvergent, wenn sie einen Grenzwert besitzt, andernfalls divergent. Abändern von endlich vielen Gliedern einer Zahlenfolge ändert das Konvergenzverhalten dieser Folge nicht. (D.h., hat die Folge den Grenzwert a, so konvergiert auch die abgeänderte Folge gegen a, ist die Folge divergent, so ist auch die abgeänderte Folge divergent.)

Beispiele: 1.) $\{\frac{1}{n}\}_{n\in\mathbb{N}}$ konvergiert gegen Null und ist daher Nullfolge.

Beweis: Man muß zu jedem $\varepsilon > 0$ eine Zahl $n_0 \in \mathbb{N}$ finden so daß $|0-\frac{1}{n}| = \frac{1}{n} < \varepsilon$ ist für alle $n \geq n_0$. Wegen der archimedischen Anordnung der reellen Zahlen gilt

$$\Big\{n \mid n \in \mathbb{N} \wedge n > \frac{1}{\varepsilon}\Big\} \neq \emptyset$$

also hat diese Menge ein kleinstes Element, das ich mit n_0 bezeichne. Somit gilt für alle $n \geq n_0$, daß $n > \frac{1}{\varepsilon}$, also $\frac{1}{n} < \varepsilon$. Hiermit ist n_0 gefunden. ■

Bevor ich weitere Beispiele betrachte, beweise ich noch einige Sätze über das Rechnen mit Zahlenfolgen, und führe noch weitere Definitionen ein.

Satz: Jede konvergente Folge ist beschränkt.

Beweis: Sei $\{x_n\}_{n\in\mathbb{N}}$ eine Zahlenfolge mit

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

Dann existiert $n_0 \in \mathbb{N}$ mit $|a - x_n| < 1$ für alle $n \geq n_0$, also

$$|x_n| = |x_n - a + a| \leq |a - x_n| + |a| < |a| + 1$$

für alle $n \geq n_0$. Also gilt

$$\forall_{n\in\mathbb{N}}: \quad |x_n| \leq \max\Big(\Big\{x_n \mid n < n_0\Big\} \cup \Big\{|a| + 1\Big\}\Big)\,.$$

■

Satz: Seien $\{x_n\}_{n\in\mathbb{N}}, \{y_n\}_{n\in\mathbb{N}}$ konvergente Zahlenfolgen mit $\lim_{n\to\infty} x_n = a$, $\lim_{n\to\infty} y_n = b$. Dann sind auch die Folgen $\{x_n + y_n\}_{n\in\mathbb{N}}, \{x_n - y_n\}_{n\in\mathbb{N}}$ und $\{x_n y_n\}_{n\in\mathbb{N}}$ konvergent, und es gilt

$$\lim_{n\to\infty}(x_n \pm y_n) = a \pm b\,, \quad \lim_{n\to\infty} x_n y_n = ab\,.$$

Wenn alle y_n und der Grenzwert b von 0 verschieden sind, so ist auch die Quotientenfolge $\{\frac{x_n}{y_n}\}_{n\in\mathbb{N}}$ konvergent, und es gilt

$$\lim_{n\to\infty} \frac{x_n}{y_n} = \frac{a}{b}\,.$$

Man beachte, daß aus diesem Satz folgt, daß die Folge $\{cx_n\}_{n\in\mathbb{N}}$ gegen ca konvergiert wenn c eine reelle Zahl ist und $\lim_{n\to\infty} x_n = a$ gilt, weil die konstante Folge $\{c\}_{n\in\mathbb{N}}$ natürlich gegen c konvergiert.

Beweis des Satzes: Sei $\varepsilon > 0$ gegeben. Es existieren $n_0, n_1 \in \mathbb{N}$ mit $|a - x_n| < \frac{\varepsilon}{2}$ für alle $n \geq n_0$, $|b - y_n| < \frac{\varepsilon}{2}$ für alle $n \geq n_1$. Also gilt für $n \geq n_2 = \max(n_0, n_1)$:

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

Damit ist die Aussage über Summe und Differenz bewiesen. Um die Aussage über das Produkt zu beweisen beachte man, daß $\{x_n\}_{n\in\mathbb{N}}$ eine konvergente Folge ist also insbesondere beschränkt ist. Sei $c > 0$ eine obere Schranke für $\{|x_n|\}_{n\in\mathbb{N}}$.

Sei nun $\varepsilon > 0$ gegeben. Falls $b \neq 0$ ist, existiert $n_0 \in \mathbb{N}$ mit $|a - x_n| < \frac{\varepsilon}{2|b|}$, für $n \geq n_0$. Für $b = 0$ sei $n_0 = 1$. Außerdem existiert $n_1 \in \mathbb{N}$ mit $|b - y_n| < \frac{\varepsilon}{2c}$ für $n \geq n_1$. Also gilt für $n \geq n_2 = \max(n_0, n_1)$

$$\begin{aligned} |ab - x_n y_n| &= |ab - bx_n + bx_n - x_n y_n| \\ &= |b(a - x_n) + x_n(b - y_n)| \leq |b|\,|a - x_n| + |x_n|\,|b - y_n| \\ &< |b|\frac{\varepsilon}{2|b|} + c\,\frac{\varepsilon}{2c} = \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon\,. \end{aligned}$$

Hiermit ist auch die Aussage für das Produkt bewiesen. Um die Aussage für den Quotienten zu beweisen, genügt es zu zeigen, daß

$$\lim_{n\to\infty} \frac{1}{y_n} = \frac{1}{b}$$

gilt, weil Produkte bereits behandelt wurden.
Es gilt

$$\frac{1}{b} - \frac{1}{y_n} = \frac{y_n - b}{y_n \cdot b}\,.$$

Die Folge $\{\frac{1}{y_n}\}_{n\in\mathbb{N}}$ ist beschränkt. Denn weil $\lim_{n\to\infty} y_n = b$ gilt, existiert $n_0 \in \mathbb{N}$ so daß für $n \geq n_0$

$$|b - y_n| < \frac{|b|}{2}\,, \quad \text{also} \quad |y_n| = |y_n - b + b| \geq |b| - |y_n - b| > \frac{|b|}{2}\,,$$

also $|\frac{1}{y_n}| \leq \frac{2}{|b|}$ gilt. Sei m das Maximum der endlichen Menge $\{|\frac{1}{y_n}| \;\; | \; n < n_0\}$. Dann ist $c = \max(m, \frac{2}{|b|})$ eine obere Schranke für $\{|\frac{1}{y_n}|\}_{n\in\mathbb{N}}$. Also folgt für alle $n \in \mathbb{N}$

$$|\frac{1}{b} - \frac{1}{y_n}| = |\frac{y_n - b}{by_n}| \leq \frac{c}{|b|}\,|y_n - b|\,.$$

Sei nun $\varepsilon > 0$ gegeben. Dann existiert $n_0 \in \mathbb{N}$ so daß für alle $n \geq n_0$ gilt $|y_n - b| < \frac{|b|}{c}\varepsilon$, also

$$|\frac{1}{b} - \frac{1}{y_n}| \leq \frac{c}{|b|}\,|y_n - b| < \varepsilon\,.$$

Hiermit ist der Satz bewiesen. ■

Man beachte, daß Zahlenfolgen reelle Funktionen sind. Also sind Summe und Produkt zweier Zahlenfolgen $\{x_n\}_{n\in\mathbb{N}}, \{y_n\}_{n\in\mathbb{N}}$ erklärt als $\{x_n + y_n\}_{n\in\mathbb{N}}, \{x_n y_n\}_{n\in\mathbb{N}}$.

Folgerung: Die konvergenten Zahlenfolgen bilden einen Vektorraum über $\mathbb{R}$, ebenso die Nullfolgen.
Dies ergibt sich aus dem eben bewiesenen Satz. Als weitere Folgerung ergibt sich der Satz:

Satz: Sei F eine rationale Funktion mit Definitionsbereich D, und sei $\{x_n\}_{n\in\mathbb{N}}$ eine

Folge mit $x_n \in D$ für alle n, die gegen $a \in D$ konvergiert. Dann ist die Folge $\{F(x_n)\}_{n\in\mathbb{N}}$ konvergent, es es gilt

$$\lim_{n\to\infty} F(x_n) = F(a)\,.$$

Beweis: Es gilt $F(x) = \frac{a_0+\ldots+a_r x^r}{b_0+\ldots+b_s x^s}$, also

$$F(x_n) = \frac{a_0 + \ldots + a_r x_n^r}{b_0 + \ldots + b_s x_n^s}\,.$$

Nach dem eben bewiesenen Satz sind die Folgen im Zähler und im Nenner konvergent mit Grenzwerten $a_0 + \ldots + a_r a^r$ und $b_0 + \ldots + b_s a^s$. Nach Voraussetzung ist $b_0 + \ldots + b_s x_n^s \neq 0$ für alle n und $b_0 + \ldots + b_s a^s \neq 0$, also kann wieder der eben bewiesene Satz angewendet werden, womit die Behauptung folgt. ■

Satz: Sei $\{x_n\}_{n\in\mathbb{N}}$ eine Nullfolge und $\{y_n\}_{n\in\mathbb{N}}$ eine beschränkte Folge. Dann ist $\{x_n y_n\}_{n\in\mathbb{N}}$ eine Nullfolge.

Beweis: Sei $c > 0$ eine obere Schranke für $\{|y_n|\}_{n\in\mathbb{N}}$. Sei $\varepsilon > 0$ gegeben. Dann existiert $n_0 \in \mathbb{N}$ mit $|0 - x_n| = |x_n| < \frac{\varepsilon}{c}$ für alle $n \geq n_0$, und somit auch

$$|0 - x_n y_n| = |x_n y_n| = |y_n|\,|x_n| \leq c|x_n| < c \cdot \frac{\varepsilon}{c} = \varepsilon\,.$$

■

Satz: $\lim_{n\to\infty} x_n = a \quad \Longrightarrow \quad \lim_{n\to\infty} |x_n| = |a|\,.$

Beweis: Sei $\varepsilon > 0$ gegeben. Dann existiert $n_0 \in \mathbb{N}$ mit $|x_n - a| < \varepsilon$ für alle $n \geq n_0$, also auch

$$\Big|\,|a| - |x_n|\,\Big| \leq |a - x_n| < \varepsilon\,.$$

■

Die Umkehrung ist falsch. Man überlege sich dies anhand eines Gegenbeispiels.

Satz: (i) Sei $\lim_{n\to\infty} x_n = a$, $\lim_{n\to\infty} y_n = b$ und $x_n \leq y_n$ für alle n. Dann gilt $a \leq b$.

(ii) Sei $\lim_{n\to\infty} x_n = \lim_{n\to\infty} y_n = a$, und für die Folge z_n gelte $x_n \leq z_n \leq y_n$, für alle $n \in \mathbb{N}$. Dann gilt

$$\lim_{n\to\infty} z_n = a\,.$$

Beweis: (i) Sei $\varepsilon > 0$ beliebig. Dann existieren $n_0, n_1 \in \mathbb{N}$ mit $|a - x_n| < \varepsilon$ für alle $n \geq n_0$ und $|b - y_n| < \varepsilon$ für alle $n \geq n_1$. Also gilt für $n \geq n_2 = \max(n_0, n_1)$:

$$\begin{aligned} a = a - x_n + x_n &\leq x_n + \varepsilon \leq y_n + \varepsilon \\ = y_n - b + b + \varepsilon &\leq b + \varepsilon + \varepsilon = b + 2\varepsilon\,, \end{aligned}$$

also

$$a \leq b + 2\varepsilon \,.$$

Da $\varepsilon > 0$ beliebig gwählt war, folgt $a \leq b$.

(ii) Sei $\varepsilon > 0$ beliebig. Es existieren $n_0, n_1 \in \mathbb{N}$ mit $|a - x_n| < \varepsilon$ für alle $n \geq n_0$, $|a - y_n| < \varepsilon$ für alle $n \geq n_1$, also gilt für alle $n \geq n_2 = \max(n_0, n_1)$:

$$-\varepsilon < -|a - x_n| \leq -(a - x_n) = x_n - a \leq z_n - a \leq y_n - a \leq |y_n - a| < \varepsilon \,,$$

also

$$|z_n - a| \leq \varepsilon \,.$$

■

Definition: Sei $\{x_n\}_{n\in\mathbb{N}}$ eine Zahlenfolge und $\varphi : \mathbb{N} \to \mathbb{N}$ eine streng monoton wachsende Abbildung. Dann heißt die Folge $(n \mapsto x_{\varphi(n)}) : \mathbb{N} \to \mathbb{R}$ eine Teilfolge von $\{x_n\}_{n\in\mathbb{N}}$. Ich bezeichne diese Teilfolge auch mit $\{x_{\varphi(n)}\}_{n\in\mathbb{N}}$. Da φ selber eine Folge ist, benützt man meistens die Bezeichnung $\{x_{j_n}\}_{n\in\mathbb{N}}$, wobei $j_n := \varphi(n)$ gesetzt ist.

Satz: Sei $\{x_n\}_{n\in\mathbb{N}}$ eine Zahlenfolge mit $\lim_{n\to\infty} x_n = a$, und sei $\{x_{\varphi(n)}\}_{n\in\mathbb{N}}$ eine Teilfolge. Dann gilt auch

$$\lim_{n\to\infty} x_{\varphi(n)} = a \,.$$

Beweis: Zunächst beachte man:
Weil $\varphi : \mathbb{N} \to \mathbb{N}$ streng monoton ist, gilt

$$\varphi(n) \geq n \quad \text{für alle} \quad n \in \mathbb{N} \,.$$

Man beweist dies durch vollständige Induktion:
a.) Induktionsanfang $n = 1$:

$$\varphi(1) \in \mathbb{N} \Longrightarrow \varphi(1) \geq 1 \,.$$

b.) Induktionsschritt: Sei die Behauptung richtig für ein $n \in \mathbb{N}$. Dann gilt, weil φ streng monoton ist,

$$\varphi(n+1) > \varphi(n) \geq n \,,$$

also $\varphi(n+1) \geq n+1$.
Nun kann der Satz bewiesen werden: Sei $\varepsilon > 0$ beliebig gegeben. Dann existiert $n_0 \in \mathbb{N}$, so daß $|a - x_n| < \varepsilon$ gilt für alle $n \geq n_0$, also auch $|a - x_{\varphi(n)}| < \varepsilon$, wegen $\varphi(n) \geq n \geq n_0$.
Damit ist der Satz bewiesen. ■

Beispiele: 1.) Sei $p \in \mathbb{N}$ (also $p \geq 1$). Dann gilt

$$\lim_{n\to\infty} \frac{1}{n^{1/p}} = \lim_{n\to\infty} n^{-\frac{1}{p}} = 0$$

Beweis: Sei $\varepsilon > 0$. Es gilt

$$|0 - n^{-\frac{1}{p}}| = n^{-\frac{1}{p}} < \varepsilon \Longleftrightarrow n > \varepsilon^{-p}.$$

Die Menge $\{n \mid n \in \mathbb{N} \wedge n > \varepsilon^{-p}\}$ ist nicht leer, und hat somit ein kleinestes Element n_0. Für alle $n \geq n_0$ gilt also $n^{-\frac{1}{p}} < \varepsilon$. ■

2.) **Folgerung:** Sei $q > 0$ rational. Dann gilt

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

Beweis: Sei $q = \frac{r}{s}$, $r, s \in \mathbb{N}$, und sei die rationale Funktion F erklärt durch

$$x \mapsto F(x) := x^r : \mathbb{R} \to \mathbb{R}.$$

Nach 1.) und dem oben bewiesenen Satz folgt

$$\lim_{n\to\infty} n^{-\frac{r}{s}} = \lim_{n\to\infty} F(n^{-\frac{1}{s}}) = F(0) = 0.$$

■

3.) Sei $q \in \mathbb{R}$. Dann gilt für

(i) $|q| < 1 :\ \lim_{n\to\infty} q^n = 0$

(ii) $q = 1 :\ \lim_{n\to\infty} q^n = 1$

(iii) $q = -1 :\ \{q^n\}$ ist divergent

(iv) $|q| > 1 :\ \{q^n\}$ ist divergent.

Beweis: (i) Für $q = 0$ ist die Behauptung richtig. Sei $0 < |q| < 1$. Zu zeigen ist: Es existiert $n_0 \in \mathbb{N}$ so daß für alle $n \geq n_0$

$$|q^n| = |q|^n < \varepsilon$$

gilt.

Sei $h := \frac{1}{|q|} - 1$. Dann (Bernoullische Ungleichung)

$$\frac{1}{|q|^n} = (1+h)^n \geq 1 + nh > nh = n\Big(\frac{1}{|q|} - 1\Big),$$

also

$$|q|^n \leq \frac{1}{n}\,\frac{|q|}{1-|q|}.$$

Setze

$$n_0 := \min\Big\{n \mid n \in \mathbb{N} \wedge \Big(\frac{1}{n}\,\frac{|q|}{1-|q|} < \varepsilon\Big)\Big\}$$

(ii) $q = 1 \Longrightarrow q^n = 1 \Longrightarrow \lim_{n\to\infty} q^n = 1$

(iii) $q = -1 \Longrightarrow \quad q^n = (-1)^n \Longrightarrow$

$$\{q^n\}_{n\in\mathbb{N}} = \{-1, 1, -1, 1, \ldots\}\,.$$

Diese Folge enthält als Teilfolgen die beiden konstanten Folgen $\{-1, -1, -1, \ldots\}$ und $\{1, 1, 1, \ldots\}$, die gegen -1 und gegen 1 konvergieren. Wäre $\{q^n\}_{n\in\mathbb{N}}$ konvergent, müßten beide Teilfolgen gegen denselben Wert, den eindeutig bestimmten Grenzwert von $\{q^n\}_{n\in\mathbb{N}}$ konvergieren. Also ist $\{q^n\}_{n\in\mathbb{N}}$ divergent.

(iv) $|q| > 1$: Sei

$$\begin{aligned} h &:= |q| - 1 \\ \Longrightarrow \qquad |q|^n &= (1+h)^n \geq 1 + nh > h = n(|q|-1)\,. \end{aligned}$$

Wegen $|q| - 1 > 0$ und weil die reellen Zahlen archimedisch angeordnet sind, folgt hieraus, daß $\{q^n\}_{n\in\mathbb{N}}$ unbeschränkt ist, also muß $\{q^n\}_{n\in\mathbb{N}}$ divergieren.

4.) Die Folge

$$\Big\{(-1)^n \frac{n}{n+1}\Big\}_{n\in\mathbb{N}}$$

diverigert.

Beweis: Diese Folge enthält die Teilfolgen

$$\Big\{\frac{2n}{2n+1}\Big\}_{n\in\mathbb{N}} \quad (n \to 2n)$$

und

$$\Big\{-\frac{2n+1}{2n+2}\Big\}_{n\in\mathbb{N}} \quad (n \to 2n+1)\,.$$

Wir zeigen, daß $\lim_{n\to\infty} \frac{2n}{2n+1} = 1$ und $\lim_{n\to\infty} -\frac{2n+1}{2n+2} = -1$ gilt. Dann muß die ursprüngliche Folge divergieren. Es gilt

$$\frac{2n}{2n+1} = \frac{2n+1}{2n+1} - \frac{1}{2n+1} = 1 - \frac{1}{2n+1}\,.$$

Da $\{1, 1, 1, \ldots\}$ gegen 1 konvergiert, und $\lim_{n\to\infty} \frac{1}{2n+1} = 0$ gilt, folgt für die Folge aus den Differenzen $\lim_{n\to\infty}(1 - \frac{1}{2n+1}) = 1 - 0 = 1$. Ebenso beweist man die Behauptung für die zweite Folge.

5.) Man betrachte die rekursiv definierte Folge

$$\begin{aligned} x_1 &= b > 0 \\ x_{n+1} &:= \frac{1}{2}\left(x_n + \frac{a}{x_n}\right), \quad a > 0\,. \end{aligned}$$

Unter der Annahme, daß diese Folge konvergiert, kann der Grenzwert c bestimmt werden. c muß > 0 sein, weil $(x_n + \frac{a}{x_n}) \geq 2\sqrt{a}$ gelten muß. Denn

$$\begin{aligned} 0 \leq \left(\sqrt{x_n} - \sqrt{\frac{a}{x_n}}\right)^2 &= x_n - 2\sqrt{a} + \frac{a}{x_n}, \quad \text{also} \\ x_{n+1} &\geq \sqrt{a} > 0\,. \end{aligned}$$

Hieraus folgt $c = \lim_{n\to\infty} x_n = \lim_{n\to\infty} x_{n+1} \geq \sqrt{a} > 0$. Da $F(x) = \frac{1}{2}(x + \frac{a}{x})$ eine rationale Funktion ist, folgt aus dem früher bewiesenen Satz und aus der Definition der Folge, daß

$$\begin{aligned} c = \lim_{n\to\infty} x_n = \lim_{n\to\infty} x_{n+1} = \lim_{n\to\infty} F(x_n) &= \frac{1}{2}\left(c + \frac{a}{c}\right), \quad \text{also} \\ c &= \sqrt{a}\,. \end{aligned}$$

Falls diese rekursiv definierte Folge konvergiert, kann sie zur näherungsweisen Berechnung von $\sqrt{a}$ benützt werden. Konvergenz ergibt sich aus den folgenden Konvergenzkriterien.

4b.) Konvergenzkriterien, Limes superior, Limes inferior

Satz: Jede monotone beschränkte Folge ist konvergent.

Beachte: Eine Folge ist monoton, wenn sie als Funktion monoton ist. Also ist eine Folge monoton wachsend, wenn

$$x_n \leq x_{n+1}$$

gilt für alle n, und monoton fallend, wenn

$$x_n \geq x_{n+1}$$

gilt für alle n.

Beweis: Sei $\{x_n\}_{n\in\mathbb{N}}$ eine monoton wachsende, beschränkte Folge. Sei c das Supremum dieser Folge. Dann existiert zu jedem $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$ mit

$$c - \varepsilon \leq x_{n_0} \leq c\,.$$

Da die Folge monoton wachsend ist, folgt also für alle $n \geq n_0$

$$c - \varepsilon \leq x_n \leq c\,,$$

also $|c - x_n| \leq \varepsilon$. ■

Folgerung: Eine monton wachsende Folge konvergiert gegen ihr Supremum. Ebenso erhält man: Eine monoton fallende Folge konvergiert gegen ihr Infimum.

Beispiel: Nun kann gezeigt werden, daß die Folge

$$\begin{aligned} x_1 &= b > 0 \\ x_{n+1} &= \frac{1}{2}\left(x_n + \frac{a}{x_n}\right) \end{aligned}$$

konvergiert.
Hierzu beachte man, daß diese Folge ab $n = 2$ monoton fallend ist. Denn es wurde gezeigt, daß $x_n \geq \sqrt{a}$ für $n \geq 2$ gilt, also $\frac{a}{x_n} \leq x_n$, folglich

$$x_{n+1} = \frac{1}{2}\left(x_n + \frac{a}{x_n}\right) \leq \frac{1}{2}\,x_n + \frac{1}{2}\,x_n = x_n\,.$$

Es wurde schon gezeigt, daß $\min(\sqrt{a}, b)$ untere Schranke ist, also konvergiert diese Folge gegen $\sqrt{a}$.

Intervallschachtelung:

Satz: Sei $\{J_n\}_{n\in\mathbb{N}}$ eine Folge von Intervallen $J_n = [a_n, b_n]$ mit

$$J_1 \supseteq J_2 \supseteq \ldots \supseteq J_n \supseteq \ldots\,.$$

Ist $b_n - a_n$ eine Nullfolge, dann enthält der Durchschnitt $\bigcap_{n\in\mathbb{N}} J_n$ genau einen Punkt.

Beweis: Die Folge $\{a_n\}_{n\in\mathbb{N}}$ ist monoton wachsend und beschränkt, denn b_1 ist obere Schranke. Ebenso ist $\{b_n\}_{n\in\mathbb{N}}$ monoton fallend und beschränkt, also konvergieren beide Folgen, und es gilt nach Voraussetzung

$$\lim_{n\to\infty} a_n - \lim_{n\to\infty} b_n = \lim_{n\to\infty} (a_n - b_n) = 0\,,$$

also $\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n = a$.
Da $\{a_n\}_{n\in\mathbb{N}}$ monoton wachsend und $\{b_n\}_{n\in\mathbb{N}}$ monoton fallend ist, ist a das Supremum von $\{a_n\}_{n\in\mathbb{N}}$ und das Infimum von $\{b_n\}_{n\in\mathbb{N}}$, also gilt

$$a_n \leq a \leq b_n\,, \quad \text{für alle} \quad n \in \mathbb{N}\,,$$

also

$$a \in \bigcap_{n\in\mathbb{N}} J_n\,.$$

Natürlich ist a der einzige Punkt, der in allen Intervallen liegt. ■

Diese Behauptung gilt nicht, wenn man anstelle abgeschlossener Intervalle offene Intervalle nimmt.

Definition: Eine Zahl a heißt Häufungspunkt der Folge $\{x_n\}_{n\in\mathbb{N}}$, wenn zu jeder Umgebung U von a unendlich viele $n \in \mathbb{N}$ existieren mit $x_n \in U$.

Äquivalent dazu ist: a ist Häufungspunkt, genau dann wenn gilt:

$$\forall_{\varepsilon>0} \quad \forall_{n\in\mathbb{N}} \quad \exists_{\substack{m\in\mathbb{N}\\ m\geq n}} : \quad |a - x_m| < \varepsilon\,.$$

Satz: Wenn a Häufungspunkt der Folge $\{x_n\}_{n\in\mathbb{N}}$ ist, dann gibt es eine Teilfolge $\{x_{j_n}\}_{n\in\mathbb{N}}$, die gegen a konvergiert.

Beweis: Es genügt, die folgende Behauptung zu beweisen:
Für jedes $n \in \mathbb{N}$ kann eine Zahl $j_n \in \mathbb{N}$ gefunden werden mit $j_n > j_{n-1}$ (falls $n \geq 2$), und

$$|a - x_{j_n}| < \frac{1}{n}\,.$$

Hieraus folgt sofort, daß die so konstruierte Teilfolge $\{x_{j_n}\}_{n\in\mathbb{N}}$ gegen a konvergiert. (Beachte, daß $\{x_{j_n}\}_{n\in\mathbb{N}}$ eine Teilfolge von $\{x_n\}_{n\in\mathbb{N}}$ ist, weil $n \mapsto j_n$ nach Konstruktion eine streng monoton wachsende Folge ist.)
Es soll nun diese Behauptung durch vollständige Induktion bewiesen werden.

1.) Induktionsanfang: $n = 1$. Da a Häufungspunkt von $\{x_n\}_{n\in\mathbb{N}}$ ist, gibt es n_0 mit

$$|a - x_{n_0}| < 1\,.$$

Setze $j_1 := n_0$.

2. Induktionsschritt: Sei j_n gefunden. Weil a Häufungspunkt von $\{x_n\}_{n\in\mathbb{N}}$ ist, gibt es unendlich viele $m \in \mathbb{N}$ mit

$$|a - x_m| < \frac{1}{n+1}\,,$$

also insbesondere auch ein solches $m_0 \in \mathbb{N}$ mit $m_0 > j_n$. Setze $j_{n+1} := m_0$.

Satz von Bolzano (1781 – 1848) **Weierstrass** (1815 – 1897)**:** Jede beschänkte Folge besitzt wenigstens einen Häufungspunkt.

Beweis: Sei a_1 eine untere Schranke und b_1 eine obere Schranke für $\{x_n\}_{n\in\mathbb{N}}$. Halbiert man das Intervall $J_1 = [a_1, b_1]$, dann liegen wenigstens in einem der beiden abgeschlossenen

Teilintervalle unendlich viele Glieder der Folge. Man wähle ein solches Teilintervall aus und nenne es J_2. Setzt man dieses Verfahren fort, dann erhält man eine Intervallschachtelung $J_1 \supseteq J_2 \supseteq \ldots, \quad J_n = [a_n, b_n]$, mit der Eigenschaft daß in jedem Intervall unendlich viele Glieder von $\{x_n\}_{n\in\mathbb{N}}$ liegen. Da $b_n - a_n = \frac{b_1-a_1}{2^{n-1}} \to 0$ gilt für $n \to \infty$, enthält der Durchschnitt $\bigcap_{n\in\mathbb{N}} J_n$ einen einzigen Punkt c.

c ist Häufungspunkt. Denn sei $\varepsilon > 0$. Wähle $n \in \mathbb{N}$ so groß, daß $b_n - a_n = \frac{b_1-a_1}{2^{n-1}} < \varepsilon$ ist. Dann gilt wegen $c \in J_n$

$$a_n \geq c - \frac{b_1 - a_1}{2^{n-1}} > c - \varepsilon$$

und

$$b_n \leq c + \frac{b_1 - a_1}{2^{n-a}} < c + \varepsilon\,,$$

also

$$J_n = [a_n, b_n] \subseteq (c - \varepsilon, x + \varepsilon)\,,$$

also liegen unendlich viele Glieder von $\{x_n\}_{n\in\mathbb{N}}$ in $(c - \varepsilon, c + \varepsilon)$.

Folgerung: Jede beschränkte Zahlenfolge besitzt eine konvergente Teilfolge.

Beweis: Jede beschränkte Folge besitzt mindestens einen Häufungspunkt. Also kann man eine Teilfolge auswählen, die gegen den Häufungspunkt konvergiert.

Im folgenden werde ich ein Konvergenzkriterium angeben, das für alle Folgen anwendbar ist, und mit dem man Konvergenz untersuchen kann, ohne den Grenzwert der Folge zu kennen. Hierzu benötigt man folgende Definition:

Definition: Eine Zahlenfolge $\{x_n\}_{n\in\mathbb{N}}$ heißt Cauchy–Folge wenn gilt:
Zu jedem $\varepsilon > 0$ existiert ein $n_0 \in \mathbb{N}$ so daß für alle $n, m \geq n_0$ gilt

$$|x_n - x_m| < \varepsilon\,.$$

Satz: Eine Folge ist genau dann konvergent, wenn sie Cauchyfolge ist.

Beweis: „$\Longrightarrow$“ Angenommen, die Folge $\{x_n\}_{n\in\mathbb{N}}$ sei konvergent. Zu zeigen ist, daß diese Folge Cauchyfolge ist. Sei $\varepsilon > 0$ gegeben, und sei a der Grenzwert. Dann existiert $n_0 \in \mathbb{N}$ so daß für alle $n \geq n_0$ gilt

$$|x_n - a| < \varepsilon/2\,,$$

also gilt für $n, m \geq n_0$

$$|x_n - x_m| = |x_n - a + a - x_m| \leq |x_n - a| + |x_m - a| < \varepsilon/2 + \varepsilon/2 = \varepsilon\,.$$

„$\Longleftarrow$“ Sei $\{x_n\}_{n\in\mathbb{N}}$ Cauchyfolge. Zu zeigen ist daß $\{x_n\}_{n\in\mathbb{N}}$ konvergiert. Zunächst zeigt man, daß jede Cauchyfolge beschränkt ist. Sei $\varepsilon > 0$. Dann existiert $n_0 \in \mathbb{N}$ so daß für alle $n \geq n_0$

$$|x_n - x_{n_0}| < \varepsilon$$

gilt, also

$$\begin{aligned} |x_n| &= |x_n - x_{n_0} + x_{n_0}| \leq |x_n - x_{n_0}| + |x_{n_0}| \\ &\leq \varepsilon + |x_{n_0}| . \end{aligned}$$

Also gilt für alle $n \in \mathbb{N}$

$$|x_n| \leq \Big(\max\{|x_1|, \ldots, |x_{n_0}|\}\Big) + \varepsilon ,$$

also ist $\{x_n\}_{n\in\mathbb{N}}$ beschränkt. Somit hat die Folge $\{x_n\}_{n\in\mathbb{N}}$ einen Häufungspunkt a.
Als nächstes wird gezeigt, daß a Grenzwert der Folge ist. Sei $\varepsilon > 0$. Da $\{x_n\}_{n\in\mathbb{N}}$ Cauchyfolge ist, existiert $n_0 \in \mathbb{N}$ mit $|x_n - x_m| < \varepsilon/2$ für alle $n, m, \geq n_0$. Nach Definition des Häufungspunktes existieren unendlich viele $n \in \mathbb{N}$ mit

$$|x_n - a| < \varepsilon/2 ,$$

also existiert auch ein $n_1 \geq n_0$ mit $|x_{n_1} - a| < \varepsilon/2$. Für $n \geq n_0$ gilt also

$$|x_n - a| = |x_n - x_{n_1} + x_{n_1} - a| \leq |x_n - x_{n_1}| + |x_{n_1} - a| < \varepsilon .$$

$\varepsilon > 0$ ist beliebig gewählt, also ist a Grenzwert von $\{x_n\}_{n\in\mathbb{N}}$.

Beispiel: Sei die Folge $\{x_n\}_{n\in\mathbb{N}}$ rekursiv definiert durch

$$\begin{aligned} x_1 &= 1 \\ x_{n+1} &= \frac{1}{1+x_n} , \quad n \geq 1 . \end{aligned}$$

Diese Folge konvergiert. Denn es gilt für alle n

$$\frac{1}{2} \leq x_n \leq 1 .$$

Dies folgt durch vollständige Induktion. $n = 1$ ist klar. Sei die Annahme für n richtig, dann folgt

$$\begin{aligned} && \frac{1}{2} \leq \frac{1}{1+x_n} \leq \frac{2}{3} , \\ \text{also} && \frac{1}{2} \leq x_{n+1} \leq \frac{2}{3} \leq 1 . \end{aligned}$$

Hieraus folgt

$$\begin{aligned} x_{n+1+k} - x_{n+1} &= \frac{1}{1+x_{n+k}} - \frac{1}{1+x_n} &&= -\frac{x_{n+k}-x_n}{(1+x_{n+k})(1+x_n)}\,, \\ \text{d.h.} \qquad |x_{n+1+k} - x_{n+1}| &= \frac{|x_{n+k}-x_n|}{(1+x_{n+k})(1+x_n)} &&\le \frac{1}{(1+\frac{1}{2})^2}\,|x_{n+k}-x_n| \\ & &&= \frac{4}{9}\,|x_{n+k}-x_n|\,. \end{aligned}$$

Durch vollständige Induktion folgt für alle n und k

$$|x_{n+k}-x_n| \le \Big(\frac{4}{9}\Big)^{n-1}|x_{k+1}-x_1| \le 2\cdot\Big(\frac{4}{9}\Big)^{n-1}.$$

Wegen $(\frac{4}{9})^{n-1} \to 0$ für $n\to\infty$ kann das Cauchy–Kriterium angewendet werden: Sei $\varepsilon > 0$. Wähle $n_0 \in \mathbb{N}$ so, daß $(\frac{4}{9})^{n-1} < \varepsilon/2$ ist für $n \ge n_0$. Dann folgt für alle $m > n \ge n_0$

$$|x_m - x_n| = |x_{n+(m-n)} - x_n| \le 2\Big(\frac{4}{9}\Big)^{n-1} \le 2\cdot\frac{\varepsilon}{2} = \varepsilon\,,$$

also ist $\{x_n\}_{n\in\mathbb{N}}$ Cauchyfolge, also konvergent. Sei a der Grenzwert. Es gilt $\frac{1}{2} \le a \le 1$, also konvergiert $\frac{1}{1+x_n}$ gegen $\frac{1}{1+a}$. Wegen $x_{n+1} = \frac{1}{1+x_n}$ muß also gelten

$$a = \lim_{n\to\infty} x_{n+1} = \lim_{n\to\infty}\frac{1}{1+x_n} = \frac{1}{1+a}\,,$$

also $\quad (1+a)a = a + a^2 = 1\,, \quad$ somit $\quad a = -\frac{1}{2} + \frac{1}{2}\sqrt{5}\,.$

Satz: Die Menge M der Häufungspunkte einer nach oben (bzw. nach unten) beschränkten Folge $\{x_n\}_{n\in\mathbb{N}}$ sei nicht leer. Dann besitzt M ein Maximum (bzw. ein Minimum).

Beweis: Sei $\{x_n\}_{n\in\mathbb{N}}$ nach oben beschränkt. Dann ist auch M nach oben beschränkt, also existiert $s_0 = \sup M$. Dann gibt es zu jedem $\varepsilon > 0$ ein $a \in M$ mit $a > s_0 - \varepsilon/2$. Da a Häufungspunkt von $\{x_n\}_{n\in\mathbb{N}}$ ist, gibt es unendlich viele $n \in \mathbb{N}$ mit

$$|x_n - a| < \varepsilon/2\,,$$

also gilt für diese unendlich vielen n

$$\begin{aligned} |x_n - s_0| &= |x_n - a + a - s_0| \\ &\le |x_n - a| + |a - s_0| \le \varepsilon/2 + \varepsilon/2 = \varepsilon\,, \end{aligned}$$

also ist s_0 selbst Häufungspunkt von $\{x_n\}_{n\in\mathbb{N}}$, und somit gilt $s_0 \in M$, dies bedeutet

$$s_0 = \max M\,.$$

Für nach unten beschränkte Folgen verläuft der Beweis genauso. ■

Definition: Das Maximum der Menge M aus dem vorangehenden Satz heißt „oberer Limes“ oder „Limes superior“ von $\{x_n\}_{n\in\mathbb{N}}$, das Minimum von M heißt „unterer Limes“ oder “Limes inferior“ von $\{x_n\}_{n\in\mathbb{N}}$. Den Limes superior bezeichnet man mit

$$\overline{\lim}_{n\to\infty} x_n\,,$$

den Limes inferior mit

$$\underline{\lim}_{n\to\infty} x_n\,.$$

Satz: (i) Es gilt $a = \overline{\lim}_{n\to\infty} x_n$, genau dann wenn für jedes $\varepsilon > 0$ gilt

$$(*) \qquad \begin{array}{ll} x_n > a-\varepsilon & \text{für unendlich viele } n \\ x_n < a+\varepsilon & \text{für fast alle } n\,. \end{array}$$

(ii) $b = \underline{\lim}_{n\to\infty} x_n$, genau dann wenn für jedes $\varepsilon > 0$ gilt

$$\begin{array}{ll} x_n < b+\varepsilon & \text{für unendlich viele } n \\ x_n > b-\varepsilon & \text{für fast alle } n\,. \end{array}$$

(„Fast alle n“ heißt: „Für alle bis auf endlich viele n“.)

Beweis: (i) Sei $a = \overline{\lim}_{n\to\infty} x_n$, und sei $\varepsilon > 0$.
Im vorangehenden Beweis wurde schon gezeigt, daß dann für unendlich viele $n \in \mathbb{N}$ gilt $x_n > a-\varepsilon$. Die Ungleichung $x_n \geq a+\varepsilon$ kann höchstens für endlich viele $n \in \mathbb{N}$ gelten. Denn sonst könnte man eine Teilfolge $\{x_{n_k}\}_{k\in\mathbb{N}}$ konstruieren mit

$$x_{n_k} \geq a+\varepsilon$$

für alle $k \in \mathbb{N}$. Diese Folge wäre auch beschränkt, hätte also einen Häufungspunkt c, und für diesen Häufungspunkt müßte gelten $c \geq a+\varepsilon$, also könnte a nicht das Maximum der Häufungspunkte von $\{x_n\}_{n\in\mathbb{N}}$ sein, weil auch c Häufungspunkt von $\{x_n\}_{n\in\mathbb{N}}$ wäre.

Angenommen, die beiden Aussagen $(*)$ seien richtig für eine Zahl $a \in \mathbb{R}$. Dann ist a Häufungspunkt von $\{x_n\}_{n\in\mathbb{N}}$, wegen der ersten und zweiten Aussage zusammen. Wegen der zweiten Aussage kann $\{x_n\}_{n\in\mathbb{N}}$ keinen Häufungspunkt haben der größer als a ist, also gilt $a = \overline{\lim}_{n\to\infty} x_n$.
Teil (ii) des Satzes wird ganz entsprechend bewiesen. ■

Für eine beschränkte Folge $\{x_n\}_{n\in\mathbb{N}}$ existieren $\overline{\lim}_{n\to\infty} x_n$ und $\underline{\lim}_{n\to\infty} x_n$ immer, und es gilt

$$\underline{\lim}_{n\to\infty} x_n \leq \overline{\lim}_{n\to\infty} x_n\,.$$

Satz: Es gilt $\underline{\lim}_{n\to\infty} x_n = \overline{\lim}_{n\to\infty} x_n$ genau dann, wenn die Folge $\{x_n\}_{n\in\mathbb{N}}$ konvergiert. In diesem Fall gilt

$$\underline{\lim}_{n\to\infty} x_n = \lim_{n\to\infty} x_n = \overline{\lim}_{n\to\infty} x_n\,.$$

Beweis: Gilt $a = \underline{\lim}_{n\to\infty} x_n = \overline{\lim}_{n\to\infty} x_n$, dann folgt aus dem vorangehenden Satz für alle $\varepsilon > 0$

$$a - \varepsilon < x_n < a + \varepsilon$$

für fast alle n, also existiert $n_0 \in \mathbb{N}$ so daß für alle $n \geq n_0$

$$|a - x_n| < \varepsilon$$

gilt, folglich gilt $\lim_{n\to\infty} x_n = a$.
Gilt $a = \lim_{n\to\infty} x_n$, dann folgt

$$a - \varepsilon < x_n < a + \varepsilon$$

für fast alle $n \in \mathbb{N}$, also ist $a = \overline{\lim}_{n\to\infty} x_n = \underline{\lim}_{n\to\infty} x_n$, nach dem vorangehenden Satz.

5 Reihen

Ich habe schon endliche Summen der Form

$$\sum_{m=1}^{n} a_m$$

betrachtet. Aus den Assoziativ– und Kommutativgesetzen folgt, daß der Wert dieser Summe nicht von der Reihenfolge der Summanden abhängt. In vielen Fällen stößt man aber auf unendliche Summen

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

Man kann solchen Summen einen präzisen Sinn geben mit Hilfe des Grenzwertbegriffes.

Definition: Sei $\{a_k\}_{k\in\mathbb{N}}$ eine Zahlenfolge, und sei $s_n = \sum_{k=1}^{n} a_k$. Die Zahlenfolge $\{s_n\}_{n=1}^{\infty} = \{\sum_{k=1}^{n} a_k\}_{n=1}^{\infty}$ wird als die zu $\{a_k\}_{k=1}^{\infty}$ gehörende unendliche Reihe bezeichnet.

Ist $\{s_n\}_{n=1}^{\infty}$ konvergent mit dem Grenzwert s, so heißt s die Summe der unendlichen Reihe, und man schreibt

$$s = \sum_{k=1}^{\infty} a_k \,.$$

Die Reihe selbst bezeichnet man auch mit $\sum_{k=1}^{\infty} a_k$.

Man muß also untersuchen, ob die Reihe $\{s_n\}_{n=1}^{\infty}$ konvergiert, und wenn ja, gegen welchen Grenzwert. Es zeigt sich, daß es hierbei im allgemeinen auf die Reihenfolge der Glieder a_n der Reihe ankommt. Daneben gibt es Reihen, die absolut konvergenten Reihen, die unabhängig von der Reihenfolge der Glieder immer gegen denselben Grenzwert konvergieren.

Beispiel: Sei $|q| < 1$ und sei $a_k = q^k$. Sei $s_n = \sum_{k=0}^{n} q^k$. Mann nennt diese Reihe „geometrische Reihe“. Ich möchte die Konvergenz dieser Reihe untersuchen. Es gilt

$$(1-q)s_n = \sum_{k=0}^{n} q^k - \sum_{k=1}^{n+1} q^k = 1 - q^{n+1} \,,$$

also

$$s_n = \frac{1-q^{n+1}}{1-q} = \frac{1}{1-q} - \frac{q^{n+1}}{1-q} \,.$$

Wegen $q^{n+1} \to 0$ für $n \to \infty$ folgt

$$\sum_{k=0}^{\infty} q^k = \lim_{n=\infty} s_n = \frac{1}{1-q} - \frac{1}{1-q} \lim_{n\to\infty} q^{n+1} = \frac{1}{1-q} \,.$$

Aus der Theorie der Folgen ergibt sich folgender Satz:

Satz: Sei $\sum_{n=1}^{\infty} a_n = a$, $\sum_{n=1}^{\infty} b_n = b$, $c \in \mathbb{R}$. Dann folgt

$$\begin{aligned} \sum_{n=1}^{\infty} (a_n + b_n) &= a + b \\ \sum_{n=1}^{\infty} c a_n &= c a \, . \end{aligned}$$

Gilt $a_n \leq b_n$, dann ist $a \leq b$.

5a.) Konvergenzkriterien

Aus dem Cauchyschen Konvergenzkriterium für Folgen erhält man unmittelbar ein Konvergenzkriterium für Reihen:

Satz: Die Reihe $\sum_{k=1}^{\infty} a_k$ konvergiert genau dann wenn zu jedem $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$ existiert so daß für alle $n \geq n_0$ und alle $m \geq n$ gilt

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

Folgerung: Ist die Reihe $\sum_{k=1}^{\infty} a_k$ konvergent, dann ist $\{a_k\}_{k=1}^{\infty}$ eine Nullfolge.

Daß $\{a_k\}_{k=1}^{\infty}$ eine Nullfolge ist, ist also notwendig für die Konvergenz der Reihe $\sum_{k=1}^{\infty} a_k$, aber nicht hinreichend. Dies sieht man aus folgendem Beispiel:

Beispiel: Die Reihe $\sum_{n=1}^{\infty} \frac{1}{n}$ ist divergent. Denn betrachte

$$\begin{aligned} s_{2^m} &= \sum_{k=1}^{2^m} \frac{1}{k} \\ &= 1 + \frac{1}{2} + \Big(\frac{1}{3} + \frac{1}{4}\Big) + \Big(\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}\Big) + \ldots + \Big(\frac{1}{2^{i-1}+1} + \ldots + \Big(\frac{1}{2^i}\Big) \\ &\quad + \underbrace{\Big(\frac{1}{2^{m-1}+1} + \frac{1}{2^{m-1}+2} + \ldots + \frac{1}{2^m}\Big)}_{2^{m-1} \text{ Summanden}} \, . \end{aligned}$$

Jede der Klammern enthält 2^{i-1} Summanden, und jeder Summand in der Klammer ist größer als $\frac{1}{2^i}$, also ist der Wert jeder der Klammern größer oder gleich $2^{-i} \cdot 2^{i-1} = \frac{1}{2}$, also

$$s_{2^m} \geq 1 + m \cdot \frac{1}{2} \, .$$

Also ist die Teilfolge $\{s_{2^m}\}_{m=1}^{\infty}$ nicht beschränkt, also auch nicht $\{s_m\}_{m=1}^{\infty}$, also divergiert die Reihe.

Die Reihe $\sum_{n=1}^{\infty} \frac{1}{n}$ heißt harmonische Reihe.

Eine Reihe heißt alternierend, wenn die Glieder abwechselndes Vorzeichen haben.

Satz (Kriterium von Leibniz)**:** Sei $\sum_{n=1}^{\infty} a_n$ eine alternierende Reihe, und sei $\{|a_n|\}_{n=1}^{\infty}$ eine monoton fallende Nullfolge. Dann ist

$$\sum_{n=1}^{\infty} a_n$$

konvergent.

Beweis: o.B.d.A. nehme ich an, daß $a_1 \geq 0$ ist. Sei $s_n = \sum_{m=1}^{n} a_n$. Die Teilfolge $\{s_{2n}\}_{n=1}^{\infty}$ von $\{s_n\}_{n=1}^{\infty}$ ist monoton wachsend, wegen

$$s_{2(n+1)} = s_{2n} + (a_{2n+1} + a_{2n+2}) \geq s_{2n} ,$$

da nach Voraussetzung

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

gilt. Außerdem ist diese Folge durch $s_1 = a_1 \geq 0$ nach oben beschränkt, wegen

$$\begin{aligned} s_{2n} &= a_1 + (-|a_2| + |a_3|) + (-|a_4| + |a_5|) + \dots \\ &\quad + (-|a_{2n-2}| + |a_{2n-1}|) - |a_{2n}| \\ &\leq a_1 . \end{aligned}$$

Ebenso ist die Folge $\{s_{2n-1}\}_{n=1}^{\infty}$ monoton fallend, und durch $s_2 = a_1 + a_2$ nach unten beschränkt, also sind beide Teilfolgen konvergent. Es gilt

$$s_{2n} - s_{2n-1} = a_{2n} .$$

Da $\{a_{2n}\}_{n=1}^{\infty}$ eine Nullfolge ist, stimmen die Grenzwerte der beiden Teilfolgen überein. Natürlich ist dann auch $\{s_n\}_{n=1}^{\infty}$ selbst konvergent gegen denselben Grenzwert s . ■

Aus diesem Beweis folgt

$$s_{2n} \leq s \leq s_{2n-1} , \quad s_{2n} \leq s \leq s_{2n+1}$$

also, wegen $s_{2n} - s_{2n-1} = a_{2n}$, $\quad s_{2n+1} - s_{2n} = a_{2n+1}$,

$$|s - s_n| \leq |a_{n+1}| .$$

Dies ist eine „Fehlerabschätzung“ .
Aus diesem Satz folgt, daß die Reihen

$$\sum_{n=0}^{\infty}(-1)^n \frac{1}{2n+1} \quad \text{(Leibnizsche Reihe)}$$
$$\sum_{n=1}^{\infty}(-1)^n \frac{1}{n}$$

konvergieren.

5b.) Umordnung von Reihen, absolut konvergente Reihen

Ich möchte ein Beispiel angeben das zeigt, daß das Konvergenzverhalten und der Grenzwert einer Reihe von der Reihenfolge der Glieder abhängt. Die Reihe $\sum_{n=1}^{\infty}(-1)^{n-1}\frac{1}{n}$ ist konvergent. Ich bezeichne ihren Grenzwert mit s. Es gilt $\frac{1}{2} \leq s \leq 1$. Man addiere nun die beiden Reihen

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \frac{1}{7} - \frac{1}{8} + \frac{1}{9} - \frac{1}{10} + \frac{1}{11} - \frac{1}{12} + - \ldots = s$$
$$0 + \frac{1}{2} + 0 - \frac{1}{4} + 0 + \frac{1}{6} + 0 - \frac{1}{8} + 0 + \frac{1}{10} + 0 - \frac{1}{12} + - \ldots = \frac{1}{2}s$$

und erhält

$$1 \quad + \frac{1}{3} - \frac{1}{2} + \frac{1}{5} \quad + \frac{1}{7} - \frac{1}{4} + \frac{1}{9} \quad + \frac{1}{11} - \frac{1}{6} + - \ldots = \frac{3}{2}s\,.$$

Dies ist eine Umordnung der urspünglichen Reihe, und man sieht, daß auch die umgeordnete Reihe konvergiert, allerdings gegen den Grenzwert $\frac{3}{2}s \neq s$, weil $s \neq 0$ ist.

Definition: Sei $\sigma : \mathbb{N} \to \mathbb{N}$ eine bijektive Abbildung. Dann heißt die Reihe $\sum_{k=1}^{\infty} a_{\sigma(k)}$ eine Umordnung der Reihe $\sum_{k=1}^{\infty} a_k$.

Satz: Es sei $a_k \geq 0$ für alle $k \in \mathbb{N}$, und die Reihe $\sum_{k=1}^{\infty} a_k$ konvergiere gegen s. Dann konvergiert auch jede Umordnung $\sum_{k=1}^{\infty} a_{\sigma(k)}$ gegen s.

Beweis: Seien $s_n = \sum_{k=1}^{n} a_k\,, \quad s'_n = \sum_{k=1}^{n} a_{\sigma(k)}$.
Die Folgen $\{s_n\}_{n=1}^{\infty}$ und $\{s'_n\}_{n=1}^{\infty}$ sind beide monoton wachsend. Hieraus folgt, daß der Grenzwert s der Folge $\{s_n\}$ gleich dem Supremum der Menge $\{s_n \mid n \in \mathbb{N}\}$ ist. Ich zeige nun, daß $s'_n \leq s$ gilt für alle $n \in \mathbb{N}$.

Hierzu wähle zu $n \in \mathbb{N}$ ein $m \in \mathbb{N}$ mit

$$m = \max\{\sigma(1), \ldots, \sigma(n)\}\,.$$

Dann folgt $\{\sigma(1),\dots,\sigma(n)\} \subseteq \{1,\dots,m\}$, also

$$s'_n = \sum_{k=1}^{n} a_{\sigma(k)} \le \sum_{k=1}^{m} a_k = s_m \le s\,,$$

also ist die Folge $\{s'_n\}_{n\in\mathbb{N}}$ nach oben beschränkt also konvergent gegen s' mit $s' \le s$. Da die Reihe $\sum_{k=1}^{\infty} a_k$ durch Umordnung aus $\sum_{k=1}^{\infty} a_{\sigma(k)}$ hervorgeht, können in diesen Überlegungen jetzt die Rollen von $\sum_{k=1}^{\infty} a_k$ und $\sum_{k=1}^{\infty} a_{\sigma(k)}$ vertauscht werden, und es folgt $s \le s'$, also $s = s'$. ■

Ich definiere nun absolut konvergente Reihen. Es wird sich zeigen, daß dies genau die Reihen sind, für die jede Umordnung gegen denselben Grenzwert konvergiert.

Definition: Eine Reihe $\sum_{k=1}^{\infty} a_k$ heißt absolut konvergent, wenn die Reihe $\sum_{k=1}^{\infty} |a_k|$ konvergiert.

Satz: Jede absolut konvergente Reihe ist konvergent.

Beweis: Um zu zeigen, daß $\sum_{k=1}^{\infty} a_k$ konvergiert, verwendet man das Cauchy–Kriterium. Sei $\varepsilon > 0$. Da die Reihe absolut konvergiert, gibt es $n_0 \in \mathbb{N}$ so daß für alle $m \ge n \ge n_0$ gilt

$$\sum_{k=n}^{m} |a_k| \le \varepsilon\,,$$

also

$$|\sum_{k=n}^{m} a_k| \le \sum_{k=n}^{m} |a_k| \le \varepsilon\,.$$

Hierbei habe ich die Dreiecksungleichung angewendet. Also konvergiert $\sum_{k=1}^{\infty} a_k$. ■

Satz: Eine Reihe $\sum_{k=1}^{\infty} a_k$ ist genau dann absolut konvergent, wenn die Folge der Partialsummen $\{\sum_{k=1}^{n} |a_k|\}_{n=1}^{\infty}$ zur Reihe $\sum_{k=1}^{\infty} |a_k|$ beschränkt ist.

Beweis: klar.

Satz: Sei $\sum_{k=1}^{\infty} a_k$ absolut konvergent. Dann ist auch jede Umordnung $\sum_{k=1}^{\infty} a_{\sigma(k)}$ absolut konvergent und es gilt

$$\sum_{k=1}^{\infty} a_k = \sum_{k=1}^{\infty} a_{\sigma(k)}\,.$$

Beweis: Für $n \in \mathbb{N}$ sei k_n eine natürliche Zahl mit den Eigenschaften $k_n > k_{n-1}$ falls $n \ge 2$, und

$$\{0,1,\dots,n\} \subseteq \{\sigma(1),\dots,\sigma(k_n)\}\,.$$

Sei $s'_{k_n} = \sum_{m=1}^{k_n} a_{\sigma(m)}$, $\quad s_n = \sum_{m=1}^{n} a_m$. Die Differenz $s'_{k_n} - s_n$ enthält nur Reihenglieder a_m mit $m > n$. Da $\sum_{k=1}^{\infty}$ absolut konvergiert, kann zu jedem $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$ gefunden werden so daß für alle $p \geq \ell \geq n_0$ gilt

$$\sum_{m=\ell}^{p} |a_m| < \varepsilon\,,$$

also folgt für alle $n \geq n_0$

$$|s'_{k_n} - s_n| < \varepsilon\,,$$

also ist $\{s'_{k_n} - s_n\}_{n\in\mathbb{N}}$ eine Nullfolge. Man beachte nun, daß $\{s'_n\}_{n=1}^{\infty}$ und $\{s_n\}_{n=1}^{\infty}$ konvergieren, weil $\sum_{k=1}^{\infty} |a_{\sigma(k)}|$ eine Umordnung von $\sum_{k=1}^{\infty} |a_k|$ ist, also konvergiert. Hieraus folgt, daß auch $\sum_{k=1}^{\infty} a_{\sigma(k)}$ konvergiert. Sei $s' = \sum_{k=1}^{\infty} a_{\sigma(k)}$, $\quad s = \sum_{k=1}^{\infty} a_k$. Wegen $s'_{k_n} \to s'$, $s_n \to s$ für $n \to \infty$, und da $s'_{k_n} - s_n$ Nullfolge ist, daß $s' = s$ gelten. ■

Darüberhinaus gilt sogar folgender Satz.

Satz: Eine Reihe $\sum_{k=1}^{\infty} a_k$ ist genau dann absolut konvergent, wenn jede ihrer Umordnungen konvergiert.

Beweis: Es wurde schon bewiesen, daß jede Umordnung konvergiert, falls die Reihe absolut konvergent ist. Zum Beweis des Satzes genügt es also zu zeigen, daß man jede konvergente Reihe, die nicht absolut konvergiert, so umordnen kann, daß sie divergiert. Sei also $\sum_{k=1}^{\infty} a_k$ eine Reihe, die nicht absolut konvergiert. Sei

$$\begin{aligned} b_k &= \frac{a_k + |a_k|}{2} = \begin{cases} a_k\,, & \text{falls} \quad a_k \geq 0 \\ 0\,, & \text{falls} \quad a_k < 0 \end{cases} \\ c_k &= \frac{a_k - |a_k|}{2} = \begin{cases} 0\,, & \text{falls} \quad a_k > 0 \\ a_k\,, & \text{falls} \quad a_k \leq 0\,. \end{cases} \end{aligned}$$

Es gilt $a_k = b_k + c_k$. Bis auf zusätzliche Nullen sind die Reihen $\sum_{k=1}^{\infty} b_k$ und $\sum_{k=1}^{\infty} c_k$ die „Teilreihen“ aller positiven Glieder beziehungsweise aller negativen Glieder von $\sum_{k=1}^{\infty} a_k$.

Da $\sum_{k=1}^{\infty} a_k = \sum_{k=1}^{\infty}(b_k + c_k)$ gilt, und da $\sum_{k=1}^{\infty} a_k$ konvergiert, können die Reihen $\sum_{k=1}^{\infty} b_k$ und $\sum_{k=1}^{\infty} c_k$ nur beide konvergieren oder beide divergieren. Wegen $|a_k| = b_k - c_k$ müssen beide divergieren, weil sonst $\sum_{k=1}^{\infty} |a_k| = \sum_{k=1}^{\infty} b_k - \sum_{k=1}^{\infty} c_k$ konvergieren würde, im Widerspruch zur Annahme.

Da $\sum_{k=1}^{\infty} b_k$ nur aus nichtnegativen Gliedern besteht und divergiert, kann man zu jedem

$C > 0$ und zu jedem $n_0 \in \mathbb{N}$ ein $m \geq n_0$ finden so daß

$$\sum_{k=n_0}^{m} b_k \geq C$$

gilt, weil sonst $\{s_n\}_{n=1}^{\infty} = \{\sum_{k=1}^{n} b_k\}_{n=1}^{\infty}$ nach oben beschränkt wäre, und foglich als monoton wachsende Folge doch konvergieren würde. Ebenso kann zu jedem $n_0 \in \mathbb{N}$ ein $m \geq n_0$ gefunden werden mit

$$\sum_{k=n_0}^{m} c_k \leq -C \, .$$

Nun konstruiere ich die umgeordnete Reihe, die divergiert. Sei k_0 die kleinste natürliche Zahl mit

$$\sum_{k=1}^{k_0} b_k \geq 1 \, ,$$

sei k_1 die kleinste natürliche Zahl mit

$$\sum_{k=1}^{k_0} b_k + \sum_{k=1}^{k_1} c_k \leq -1 \, ,$$

und sei k_2 die kleinste natürliche Zahl mit

$$\sum_{k=1}^{k_0} b_k + \sum_{k=1}^{k_1} c_k + \sum_{k=k_0+1}^{k_2} b_k \geq 1 \, .$$

Durch Fortsetzung des Verfahrens erhält man eine Reihe

$$\sum_{k=1}^{\infty} d_k$$

mit

$$d_k = \begin{cases} b_k & , \quad 1 \leq k \leq k_0 \\ c_{k-k_0} & , \quad k_0 + 1 \leq k \leq k_1 + k_0 \\ b_{k-k_1} & , \quad k_1 + k_0 + 1 \leq k \leq k_2 + k_1 + k_0 \\ \vdots & \end{cases}$$

die eine Umordnung von $\sum_{k=1}^{\infty} a_k$ ist, weil man zusätzliche Nullen weglassen darf, und für die $\overline{\lim}_{n\to\infty} \sum_{k=1}^{n} d_k \geq 1$, $\underline{\lim}_{n\to\infty} \sum_{k=1}^{n} d_k \leq -1$, also $\overline{\lim}_{n\to\infty} \sum_{k=1}^{n} d_k \neq \underline{\lim}_{n\to\infty} \sum_{k=1}^{n} d_k$ gilt. Folglich konvergiert diese Reihe nicht.
Damit ist der Satz bewiesen. ∎

Satz (Großer Umordnungssatz)**:** Sei eine Doppelreihe $\sum_{k,\ell=1}^{\infty} a_{k\ell}$ gegeben. Es existiere eine Zahl M mit

$$\sum_{k,\ell=1}^{m} |a_{k\ell}| \leq M\,, \quad \text{für alle} \quad m \in \mathbb{N}\,.$$

Dann gilt:

1.) Jede Anordnung der Doppelreihe in eine Einfachreihe ist absolut konvergent mit stets gleicher Summe s.

2.) Die Reihen $\sum_{\ell=1}^{\infty} a_{k\ell}$ und $\sum_{k=1}^{\infty} a_{k\ell}$ sind absolut konvergent.

3.) Seien $s_k = \sum_{\ell=1}^{\infty} a_{k\ell}\,, \quad s'_\ell = \sum_{k=1}^{\infty} a_{k\ell}$. Die beiden Reihen $\sum_{k=1}^{\infty} s_k$ und $\sum_{\ell=1}^{\infty} s'_\ell$ sind absolut konvergent mit Grenzwert s.

Bemerkungen:

Beachte, daß 3.) bedeutet:

$$\begin{aligned}
\sum_{k=1}^{\infty}\Big(\sum_{\ell=1}^{\infty} a_{k\ell}\Big) &= \lim_{m\to\infty}\Big(\sum_{k=1}^{m}(\lim_{j\to\infty}\sum_{\ell=1}^{j} a_{k\ell})\Big) \\
&= \lim_{m\to\infty}\Big(\lim_{j\to\infty}(\sum_{k=1}^{m}\sum_{\ell=1}^{j} a_{k\ell})\Big) = \lim_{j\to\infty}\Big(\lim_{m\to\infty}(\sum_{\ell=1}^{j}\sum_{k=1}^{m} a_{k\ell})\Big) \\
&= \lim_{j\to\infty}\Big(\sum_{\ell=1}^{j}(\lim_{m\to\infty}\sum_{k=1}^{m} a_{k\ell})\Big) = \sum_{\ell=1}^{\infty}\sum_{k=1}^{\infty} a_{k\ell}\,.
\end{aligned}$$

Also ist 3.) eine Aussage über Vertauschung von Grenzübergängen.

Die Menge $\mathbb{N}\times\mathbb{N}$ ist abzählbar, daß heißt es existiert mindestens eine bijektive Abbildung $\sigma : \mathbb{N} \to \mathbb{N}\times\mathbb{N}$ (eine sogenannte Abzählung). Dies sieht man, wenn man die Elemente $(k,\ell) \in \mathbb{N}\times\mathbb{N}$ in der Form eines doppelt unendlichen Schemas anordnet:

$$\begin{array}{cccc}
(1,1) & (1,2) & (1,3) & \dots \\
(2,1) & (2,2) & (2,3) & \dots \\
(3,1) & (3,2) & (3,3) & \dots \\
\vdots & \vdots & \vdots &
\end{array}\,.$$

Natürlich ist σ nicht eindeutig bestimmt. Mit 1.) ist genauer gemeint: Für jede Abzählung σ von $\mathbb{N}\times\mathbb{N}$ ist die Reihe $\sum_{k=1}^{\infty} a_{\sigma(k)}$ absolut konvergent mit Grenzwert s.

Beweis: 1.) Sei σ eine Abzählung von $\mathbb{N}\times\mathbb{N}$. Sei $n \in \mathbb{N}$, und sei m die größte natürliche Zahl, die in einem der Paare $\sigma(1), \sigma(2), \dots, \sigma(n)$ auftritt. Dann ist

$$\sum_{j=1}^{n} |a_{\sigma(j)}| \leq \sum_{k,\ell=1}^{m} |a_{k\ell}| \leq M\,.$$

also ist $\sum_{j=1}^{\infty} a_{\sigma(j)}$ absolut konvergent. Wenn σ' eine zweite Abzählung von $\mathbb{N} \times \mathbb{N}$ ist, dann ist $\sum_{j=1}^{\infty} a_{\sigma'(j)}$ eine Umordnung von $\sum_{j=1}^{\infty} a_{\sigma(j)}$ und ist damit absolut konvergent mit derselben Summe.

2.) Für $k, n \in \mathbb{N}$ sei $m = \max(k, n)$. Dann folgt

$$\sum_{\ell=1}^{n} |a_{k\ell}| \le \sum_{k,\ell=1}^{m} |a_{k\ell}| \le M ,$$

also ist $\sum_{\ell=1}^{\infty} a_{k\ell}$ absolut konvergent. Ebenso folgt, daß $\sum_{k=1}^{\infty} a_{k\ell}$ absolut konvergiert.

3.)

$$\begin{aligned}
\sum_{k=1}^{m} |s_k| &= \sum_{k=1}^{m} \Big|\sum_{\ell=1}^{\infty} a_{k\ell}\Big| = \sum_{k=1}^{m} \Big|\lim_{n\to\infty} \sum_{\ell=1}^{n} a_{k\ell}\Big| \\
&\le \sum_{k=1}^{m} \lim_{n\to\infty} \Big|\sum_{\ell=1}^{n} a_{k\ell}\Big| \le \sum_{k=1}^{m} \lim_{n\to\infty} \sum_{\ell=1}^{n} |a_{k\ell}| \\
&= \lim_{n\to\infty} \sum_{k=1}^{m} \sum_{\ell=1}^{n} |a_{k\ell}| \le \lim_{n\to\infty} \sum_{k,\ell=1}^{n} |a_{k\ell}| \le M ,
\end{aligned}$$

also ist $\sum_{k=1}^{\infty} s_k$ absolut konvergent. Ebenso folgt, daß $\sum_{k=1}^{\infty} s'_\ell$ absolut konvergiert. Also bleibt zu zeigen, daß die Grenzwerte beider Reihen übereinstimmen. Sei $\sigma : \mathbb{N} \to \mathbb{N} \times \mathbb{N}$ eine Abzählung und sei $\varepsilon > 0$. Da $\sum_{k=1}^{\infty} a_{\sigma(k)} = s$ gilt und absolut konvergiert, gibt es $\nu \in \mathbb{N}$ mit $\sum_{k=n+1}^{\infty} |a_{\sigma(k)}| < \varepsilon$ für $n \ge \nu$.

Sei μ die größte natürliche Zahl, die in einem der Paare $\sigma(1), \ldots, \sigma(n)$ vorkommt. Sei $m, j \ge \mu$. Die Summe $\sum_{k=1}^{m} \sum_{\ell=1}^{j} a_{k\ell} - \sum_{k=1}^{n} a_{\sigma(k)}$ enthält dann nur Glieder $a_{k\ell}$ mit $(k, \ell) \ne \sigma(j)$ für alle $j = 1, \ldots, n$, also folgt

$$\Big|\sum_{k=1}^{m} \sum_{\ell=1}^{j} a_{k\ell} - \sum_{k=1}^{n} a_{\sigma(k)}\Big| \le \sum_{k=n+1}^{\infty} |a_{\sigma(k)}| < \varepsilon ,$$

folglich

$$
\begin{aligned}
|\sum_{k=1}^{\infty}(\sum_{\ell=1}^{\infty} a_{k\ell}) - \sum_{k=1}^{n} a_{\sigma(k)}| &= |\lim_{m\to\infty}\Big(\lim_{j\to\infty}\sum_{k=1}^{m}\sum_{\ell=1}^{j} a_{k\ell}\Big) - \sum_{k=1}^{n} a_{\sigma(k)}| \\
&= |\lim_{m\to\infty}\Big(\lim_{j\to\infty}\Big[\sum_{k=1}^{m}\sum_{\ell=1}^{j} a_{k\ell} - \sum_{k=1}^{n} a_{\sigma(k)}\Big]\Big)| \\
&\leq \lim_{m\to\infty}|\lim_{j\to\infty}\Big[\sum_{k=1}^{m}\sum_{\ell=1}^{j} a_{k\ell} - \sum_{k=1}^{n} a_{\sigma(k)}\Big]| \\
&\leq \lim_{m\to\infty}\lim_{j\to\infty}|\sum_{k=1}^{m}\sum_{\ell=1}^{j} a_{k\ell} - \sum_{k=1}^{n} a_{\sigma(k)}| \\
&\leq \lim_{m\to\infty}\lim_{j\to\infty}\varepsilon = \varepsilon\,.
\end{aligned}
$$

Hieraus folgt $|\sum_{k=1}^{\infty}(\sum_{\ell=1}^{\infty} a_{k\ell}) - \sum_{k=1}^{\infty} a_{\sigma(k)}| \leq \varepsilon$. Da $\varepsilon > 0$ beliebig gewählt war, resultiert also

$$\sum_{k=1}^{\infty}\Big(\sum_{\ell=1}^{\infty} a_{k\ell}\Big) = \sum_{k=1}^{\infty} a_{\sigma(k)} = s\,.$$

Ebenso ergibt sich

$$\sum_{\ell=1}^{\infty}\Big(\sum_{k=1}^{\infty} a_{k\ell}\Big) = s\,.$$

Damit ist der Satz bewiesen. ■

Dieser Satz kann auf Produkte von Reihen angewendet werden. Multipliziert man zwei endliche Summen, dann gilt nach dem Distributivgesetz

$$\Big(\sum_{k=1}^{n} a_k\Big)\cdot\Big(\sum_{\ell=1}^{n} b_\ell\Big) = \sum_{k,\ell=1}^{n}(a_k b_\ell)\,.$$

Die Reihenfolge der Summation ist beliebig. Für absolut konvergente Reihen gilt:

Satz: Jedes Produkt zweier absolut konvergenter Reihen $\sum_{k=1}^{\infty} b_k, \sum_{\ell=1}^{\infty} c_\ell$ ist absolut konvergent mit stets gleichem Wert:

$$\sum_{k,\ell=1}^{\infty} b_k c_\ell = \Big(\sum_{k=1}^{\infty} b_k\Big)\Big(\sum_{\ell=1}^{\infty} c_\ell\Big)\,.$$

Präziser: Sei $\sigma : \mathbb{N} \to \mathbb{N}\times\mathbb{N}$ eine Abzählung, $\sigma(k) = (\sigma_1(k), \sigma_2(k)) \in \mathbb{N}\times\mathbb{N}$. Dann gilt

$$(*) \qquad \sum_{k=1}^{\infty} b_{\sigma_1(k)}\, c_{\sigma_2(k)} = \Big(\sum_{k=1}^{\infty} b_k\Big)\cdot\Big(\sum_{\ell=1}^{\infty} c_\ell\Big)\,.$$

Beweis: Die Voraussetzungen des großen Umordnungssatzes sind erfüllt mit $a_{k\ell} = b_k c_\ell$, wegen

$$\begin{aligned} \sum_{k,\ell=1}^{m} |a_{k\ell}| &= \sum_{k,\ell=1}^{m} |b_k c_\ell| = \sum_{k,\ell=1}^{m} |b_k|\,|c_\ell| \\ &= \sum_{k=1}^{m}\Big(\sum_{\ell=1}^{m} |b_k|\,|c_\ell|\Big) = \Big(\sum_{k=1}^{m}|b_k|\Big)\Big(\sum_{\ell=1}^{m}|c_\ell|\Big) \\ &\le \Big(\sum_{k=1}^{\infty}|b_k|\Big)\Big(\sum_{\ell=1}^{\infty}|c_\ell|\Big) =: M \end{aligned}$$

$(*)$ folgt aus Aussage 3.) des großen Umordnungssatzes wegen

$$\begin{aligned} s_\ell &= \sum_{\ell=1}^{\infty} b_k c_\ell = b_k\Big(\sum_{\ell=1}^{\infty} c_\ell\Big), \\ s'_\ell &= \sum_{k=1}^{\infty} b_k c_\ell = \Big(\sum_{k=1}^{\infty} b_k\Big) c_\ell . \end{aligned}$$

■

Sei $\sigma : \mathbb{N} \to \mathbb{N}_0 \times \mathbb{N}_0$ die Abzählung nach den Diagonalen

$$\begin{array}{ccccccc} (0,0) & & (0,1) & & (0,2) & & \dots \\ & \swarrow & & \swarrow & & \swarrow & \\ (1,0) & & (1,1) & & (1,2) & & \dots \\ & \swarrow & & \swarrow & & & \\ (2,0) & & (2,1) & & (2,2) & & \dots \\ \vdots & \swarrow & \vdots & & \vdots & & \end{array} .$$

Benützt man im vorangehenden Satz diese Abzählung, dann erhält man das Cauchy–Produkt:

$$\sum_{k=0}^{\infty}\sum_{\ell=0}^{k} b_\ell c_{k-\ell} .$$

Also folgt aus diesem Satz:

Satz: Das Cauchysche Produkt $\sum_{k=0}^{\infty}(\sum_{\ell=0}^{k} b_\ell c_{k-\ell})$ zweier absolut konvergenter Reihen $\sum_{k=0}^{\infty} b_k$, $\sum_{k=0}^{\infty} c_k$ ist absolut konvergent, und es gilt

$$\sum_{k=0}^{\infty}\Big(\sum_{\ell=0}^{k} b_\ell c_{k-\ell}\Big) = \Big(\sum_{k=0}^{\infty} b_k\Big)\Big(\sum_{k=0}^{\infty} c_k\Big) .$$

Beispiel: Für jedes $x \in \mathbb{R}$ konvergiert die Reihe $\sum_{k=0}^{\infty} \frac{x^k}{k!}$ absolut. Denn sei k_0 die kleinste natürliche Zahl k mit $k_0 \geq 2|x|$. Dann folgt für $m > k_0$

$$\begin{aligned}
\sum_{k=0}^{m} \frac{|x|^k}{k!} &\leq \sum_{k=0}^{k_0-1} \frac{|x|^k}{k!} + \sum_{k=k_0}^{m} \frac{|x|^k}{k!} \\
&\leq \sum_{k=0}^{k_0-1} \frac{|x|^k}{k!} + \sum_{k=k_0}^{m} \frac{|x|^k}{k_0^{k-k_0+1}} \\
&\leq \sum_{k=0}^{k_0-1} \frac{|x|^k}{k!} + \sum_{k=k_0}^{m} \Big(\frac{1}{2}\Big)^k \frac{k_0^k}{k_0^{k-k_0+1}} \\
&\leq \sum_{k=0}^{k_0-1} \frac{|x|^k}{k!} + k_0^{k_0-1} \Big(\frac{1}{2}\Big)^{k_0} \sum_{k=0}^{m-k_0} \Big(\frac{1}{2}\Big)^k \\
&= \sum_{k=0}^{k_0-1} \frac{|x|^k}{k!} + k_0^{k_0-1} \Big(\frac{1}{2}\Big)^{k_0-1} \Big(1 - \Big(\frac{1}{2}\Big)^{m-k_0+1}\Big) \\
&\leq \sum_{k=0}^{k_0-1} \frac{|x|^k}{k!} + \Big(\frac{k_0}{2}\Big)^{k_0-1} =: M
\end{aligned}$$

also ist $\sum_{k=0}^{\infty} \frac{x^k}{k!}$ absolut konvergent. Aus dem vorangehenden Satz und aus der binomischen Formel folgt somit

$$\begin{aligned}
\sum_{k=0}^{\infty} \frac{x^k}{k!} \cdot \sum_{\ell=0}^{\infty} \frac{y^\ell}{\ell!} &= \sum_{k=0}^{\infty} \Big[\sum_{\ell=0}^{k} \Big(\frac{x^\ell}{\ell!} \frac{y^{k-\ell}}{(k-\ell)!} \Big) \Big] \\
&= \sum_{k=0}^{\infty} \Big[\frac{1}{k!} \sum_{\ell=0}^{k} \binom{k}{\ell} x^\ell y^{k-\ell} \Big] = \sum_{k=0}^{\infty} \frac{(x+y)^k}{k!} .
\end{aligned}$$

Definiert man eine Funktion $\exp : \mathbb{R} \to \mathbb{R}$ durch $x \mapsto \exp(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!}$, dann folgt also

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

Wegen $\exp(0) = 1$ folgt insbesondere für $y = -x$

$$\exp(x) \cdot \exp(-x) = 1 ,$$

also

$\binom{*}{*}$ $$\exp(-x) = \frac{1}{\exp(x)} .$$

Seit Leonhard Euler (1707 – 1783) bezeichnet man die Zahl $\exp(1)$ mit e:

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

Aus $(*)$ folgt durch vollständige Induktion für jedes $x \in \mathbb{R}\,,\;\; x \geq 0$

$$\exp(nx) = [\exp(x)]^n\,.$$

Für $x = \frac{1}{m}\,,\;\; m \in \mathbb{N}$, folgt hieraus insbesondere, daß $\exp\left(\frac{1}{m}\right)$ die eindeutig bestimmte positive Lösung der Gleichung

$$y^m = e$$

ist, also ist $\exp\left(\frac{1}{m}\right) = e^{1/m}$, also $\exp\left(\frac{n}{m}\right) = e^{n/m}$. Hieraus und aus $\binom{*}{*}$ folgt für jedes rationale q

$$\exp(q) = e^q\,.$$

Also ist es sinnvoll, für jedes reelle $x \in \mathbb{R}$ zu definieren

$$e^x := \sum_{k=0}^{\infty} \frac{x^k}{k!}$$

$(*)$ und $\binom{*}{*}$ lauten dann

$$e^{x+y} = e^x e^y\,, \quad e^{-x} = \frac{1}{e^x}\,.$$

5c.) Kriterien für absolute Konvergenz

Definition: Sei $\sum_{k=1}^{\infty} c_k$ eine Reihe mit lauter nichtnegativen Gliedern. Sei heißt Majorante für die Reihe $\sum_{k=1}^{\infty} a_k$, wenn sie konvergiert und wenn außerdem

$$|a_k| \leq c_k$$

gilt für alle k. Sei heißt Minorante für $\sum_{k=1}^{\infty} a_k$, wenn sie divergiert und wenn

$$|a_k| \geq c_k$$

gilt für alle k.

Satz: Die Reihe $\sum_{k=1}^{\infty} a_k$ ist absolut konvergent, genau dann wenn sie eine Majorante besitzt, und nicht absolut konvergent, genau dann wenn sie eine Minorante besitzt.

Beweis: Wenn eine Majorante existiert, gilt

$$\sum_{k=1}^{n} |a_k| \leq \sum_{k=1}^{n} c_k \leq \sum_{k=1}^{\infty} c_k\,,$$

also ist $\sum_{k=1}^{\infty} a_k$ absolut konvergent. Ist $\sum_{k=1}^{\infty} a_k$ absolut konvergent, dann ist $\sum_{k=1}^{\infty} |a_k|$ eine Majorante. Also ist die erste Aussage bewiesen. Sei $\sum_{k=1}^{\infty} c_k$ eine Minorante. Dann existiert zu jedem M ein $n \in \mathbb{N}$ mit

$$\sum_{k=1}^{n} |a_k| \geq \sum_{k=1}^{n} c_k \geq M\,,$$

also ist $\sum_{k=1}^{n} a_k$ nicht absolut konvergent. Ist $\sum_{k=1}^{\infty} a_k$ nicht absolut konvergent, dann ist $\sum_{k=1}^{\infty} |a_k|$ eine Minorante. ■

Satz (Wurzelkriterium)**:** Sei $\overline{\lim}_{k\to\infty} \sqrt[k]{|a_k|} = a\,.$
Die Reihe $\sum_{k=1}^{\infty} a_k$ ist

$$\begin{cases} \text{absolut konvergent}\,, & \text{falls} \quad a < 1 \\ \text{divergent}\,, & \text{falls} \quad a > 1\,. \end{cases}$$

(Falls $\sqrt[k]{|a_k|}$ nicht nach oben beschränkt ist, der Limes superior also nicht existiert, ist $\sum_{k=1}^{\infty} a_k$ divergent.)

Beweis: Falls $a < 1$ ist gibt es ein $k_0 \in \mathbb{N}$, so daß für alle $k \geq k_0$

$$\sqrt[k]{|a_k|} \leq \theta = \frac{a+1}{2} < 1\,,$$

also

$$|a_k| \leq \theta^k$$

gilt. Also ist $\sum_{k=1}^{\infty} \theta^k$ Majorante.

Wenn $a > 1$ ist, dann gibt es unendlich viele k derart, daß $\sqrt[k]{|a_k|} \geq 1$, also $|a_k| \geq 1$ ist. Also ist $\{a_k\}_{k=1}^{\infty}$ keine Nullfolge, also kann $\sum_{k=1}^{\infty} a_k$ nicht konvergent sein. ■

Satz (Quotientenkriterium)**:**
a.) Für fast alle $k \in \mathbb{N}$ gelte $a_k \neq 0$ und es sei

$$\overline{\lim}_{k\to\infty} \left| \frac{a_{k+1}}{a_k} \right| = b\,.$$

Falls $b < 1$ ist, ist die Reihe $\sum_{k=1}^{\infty} a_k$ absolut konvergent.
b.) Für fast alle $k \in \mathbb{N}$ gelte $a_k \neq 0$ und

$$\left| \frac{a_{k+1}}{a_k} \right| \geq 1\,.$$

Dann ist die Reihe $\sum_{k=1}^{\infty} a_k$ divergent.

Beweis: a.) Sei $b < 1$ und sei $\theta = \frac{b+1}{2} < 1$. Dann existiert ein $k_0 \in \mathbb{N}$ so daß für alle $k \geq k_0$

$$\frac{a_{k+1}}{a_k} \leq \theta\,,$$

also $a_{k+1} \leq \theta a_k$ gilt. Durch vollständige Induktion folgt hieraus $a_k \leq \theta^{(k-k_0)} a_{k_0}$, also ist $\frac{a_{k_0}}{\theta^{k_0}} \sum_{k=1}^{\infty} \theta^k$ eine Majorante (eventuell nach Vergrößerung der ersten $k_0 - 1$ Glieder), also

konvergiert $\sum_{k=1}^{\infty} a_k$ absolut, somit ist a.) bewiesen.
b.) Da $|a_{k+1}| \geq |a_k|$ gilt für fast alle k, und da $|a_k| \neq 0$ ist für fast alle k, kann $\{a_k\}_{k=1}^{\infty}$ keine Nullfolge sein, also muß $\sum_{k=1}^{\infty} a_k$ divergieren. ■

Beispiel: $\displaystyle\sum_{\nu=1}^{\infty} \left(\frac{\nu}{\nu+1}\right)^{\nu^2}$ ist konvergent. Der Beweis kann mit dem Wurzelkriterium geführt werden:

$$\begin{aligned} \overline{\lim}_{\nu\to\infty} \sqrt[\nu]{(\frac{\nu}{\nu+1})^{\nu^2}} &= \overline{\lim}_{\nu\to\infty} \left(\frac{\nu}{\nu+1}\right)^{\nu} \\ = \quad \overline{\lim}_{\nu\to\infty} \frac{1}{(1+\frac{1}{\nu})^{\nu}} &\leq \overline{\lim}_{\nu\to\infty} \frac{1}{1+\nu\frac{1}{\nu}} = \frac{1}{2}\,, \end{aligned}$$

nach der Bernoullischen Ungleichung.

Dagegen ist $\sum_{\nu=1}^{\infty}(\frac{\nu}{\nu+1})^{\nu}$ divergent.

5d.) *g*–adische Entwicklungen

Sei $g \geq 2$ eine gegebene natürliche Zahl. Die ganzen Zahlen a mit $0 \leq a < g$ nennt man g–adische Ziffern. Es sei

$$\sum_{k=\ell}^{\infty} a_k g^{-k}$$

eine Reihe mit g–adischen Ziffern a_k. Hierbei sei ℓ eine beliebige gegebene ganze Zahl. Diese Reihe konvergiert, denn die geometrische Reihe

$$\sum_{k=\ell}^{\infty} g^{-k+1}$$

ist Majorante. Sei x der Grenzwert:

$$x = \sum_{k=\ell}^{\infty} a_k g^{-k}\,.$$

Falls nicht fast alle $a_k = g-1$ sind, nennt man diese Reihe die g–adische Entwicklung von x. Man sagt auch, x sei als Systembruch im g–adischen System dargestellt.
($g = 10$: Dezimalbruchentwicklung oder dekadische Entwicklung, $g = 2$: Dualbruchentwicklung oder dyadische Entwicklung.)

Satz: Jede positive reelle Zahl hat eine eindeutig bestimmte g–adische Entwicklung

$$x = \sum_{k=\ell}^{\infty} a_k g^{-k}\,,$$

mit $a_\ell \neq 0$, und so daß für unendlich viele k gilt $a_k \neq g-1$.

Beweis: Zur Vereinfachung der Schreibweise sei $q = g^{-1}$. Im Beweis wird die Formel

$$(g-1)\sum_{k=n}^{\infty} q^k = (g-1)q^n \sum_{k=0}^{\infty} q^k = \big(\frac{1}{q}-1\big)q^n \frac{1}{1-q} = \frac{q^n}{q} = gq^n. \qquad (*)$$

gebraucht, die für alle $n \in \mathbb{N}$ gilt.

Zunächst wird gezeigt, dass es für jede reelle Zahl $x > 0$ höchstens eine g-adische Entwicklung geben kann. Sei

$$x = \sum_{k=\ell}^{\infty} a_k g^{-k} = \sum_{k=\ell}^{\infty} a_k q^k$$

eine solche Entwicklung. Mit

$$x_n = \sum_{k=n}^{\infty} a_k q^k$$

folgt dann aus $(*)$, dass

$$\begin{aligned} a_n \leq g^n x_n &= a_n + g^n \sum_{k=n+1}^{\infty} a_k q^k < a_n + g^n \sum_{k=n+1}^{\infty} (g-1)q^k \\ &= a_n + g^{n+1} q^{n+1} = a_n + 1. \end{aligned}$$

Dies bedeutet, dass

$$a_n = [g^n x_n]$$

sein muss. Also erfüllen die Koeffizienten a_k der g-adischen Entwicklung von x das folgende rekursive Schema:

Es sei ℓ die größte ganze Zahl mit $g^{\ell-1}x < 1$, und es sei

$$\begin{aligned} x_\ell &= x, \\ a_\ell &= [g^\ell x_\ell]. \end{aligned}$$

Ist $k \in \mathbb{N}$ mit $k \geq \ell$ und sind x_k und a_k bekannt, dann gilt

$$\begin{aligned} x_{k+1} &= x_k - a_k q^k, \\ a_{k+1} &= [g^{k+1} x_{k+1}]. \end{aligned} \qquad (**)$$

Dieses Schema bestimmt die Zahlen a_k und ℓ eindeutig, also gibt es zu x höchstens eine g-adische Entwicklung.

Als nächstes wird gezeigt, dass durch dieses Schema eine g-adische Entwicklung bestimmt wird, die gegen x konvergiert. Wir beweisen erst durch vollständige Induktion, dass

$$g^{k-1}\, x_k < 1, \quad a_k < g \tag{$***$}$$

gilt. Denn nach Konstruktion von ℓ ist $g^{\ell-1}\, x_\ell < 1$, also

$$a_\ell = [g^\ell x_\ell] \le g(g^{\ell-1} x_\ell) < g.$$

Sei nun $k \ge \ell$. Ist $g^{k-1} x_k < 1$ und $a_k < g$, dann folgt

$$g^k x_{k+1} = g^k x_k - a_k = g^k x_k - [g^k x_k] < 1,$$

und somit auch

$$a_{k+1} = [g^{k+1} x_{k+1}] < g.$$

Damit ist $(***)$ nachgewiesen. Es gilt auch $a_k \neq g-1$ für unendliche viele k. Denn wäre $a_k = g-1$ für fast alle k, dann würde n_0 existieren mit $a_k = g-1$ für alle $k \ge n_0$. Aus $(*)$ folgte

$$g^{n_0-1} x_{n_0} = g^{n_0-1} \sum_{k=n_0}^{\infty} (g-1) q^k = g^{n_0}\, q^{n_0} = 1,$$

im Widerspruch zu $(***)$. Damit ist die Behauptung bewiesen. Zusammen ergibt sich, dass die aus dem Schema konstruierte Entwicklung

$$\sum_{k=\ell}^{\infty} a_k q^k$$

g-adisch ist.

Es bleibt zu zeigen, dass $x = \sum_{k=\ell}^{\infty} a_k q^k$ gilt. Durch vollständige Induktion ergibt sich aus $(**)$, dass

$$x_n = x_\ell - \sum_{k=\ell}^{n-1} a_k q^k = x - \sum_{k=\ell}^{n-1} a_k q^k, \quad n \ge \ell.$$

Zusammen mit $(***)$ folgt hieraus

$$x - \sum_{k=\ell}^{n-1} a_k q^k = q^{n-1}(g^{n-1} x_n) < q^{n-1},$$

also

$$x - \sum_{k=\ell}^{\infty} a_k q^k = \lim_{n\to\infty} \left(x - \sum_{k=\ell}^{n-1} a_k q^k \right) \le \lim_{n\to\infty} q^{n-1} = 0,$$

wegen $0 < q < 1$.

6 Stetige Funktionen

6a.) Topologie von $\mathbb{R}$

Unter den Intervallen werden die offenen, abgeschlossenen und die halboffenen Intervalle betrachtet. Außerdem sind bisher schon Begriffe wie Umgebung, Häufungspunkt, usw. vorgekommen. Diese Dinge sollen nun in einem allgemeineren Rahmen untersucht werden.

Definition: Sei $M \subseteq \mathbb{R}$. $x \in \mathbb{R}$ heißt innerer Punkt von M, wenn es eine Umgebung U von x gibt, die in M enthalten ist.

Definition: Sei $M \subseteq \mathbb{R}$. $x \in \mathbb{R}$ heißt Häufungspunkt von M, wenn es in jeder Umgebung von x ein $y \in M$ gibt mit $x \neq y$.

Beachte, daß ein innerer Punkt von M stets zu M gehört. Ein Häufungspunkt von M dagegen braucht nicht zu M zu gehören. Es wurde schon der Begriff Häufungspunkt einer Folge eingeführt. Man beachte den Unterschied. Ein Häufungspunkt einer Folge braucht nicht Häufungspunkt des Wertebereiches der Folge zu sein. Ein innerer Punkt ist stets Häufungspunkt von M.

Beispiel: Betrachte das Intervall (a, b) mit $a < b$. Dann sind a und b Häufungspunkte, aber keine inneren Punkte. Das Intervall besteht nur aus inneren Punkten. Auch für das Intervall $[a, b]$ sind a, b Häufungspunkte, die diesmal zum Intervall gehören, jedoch sind auch hier a, b keine inneren Punkte.

Jede reelle Zahl x ist Häufungspunkt der Menge $\mathbb{Q}$. Jedoch hat $\mathbb{Q}$ keine inneren Punkte.

Definition: $M \subseteq \mathbb{R}$ heißt offen, genau dann wenn jeder Punkt von M innerer Punkt ist.

Definition: $M \subseteq \mathbb{R}$ heißt abgeschlossen, genau dann wenn jeder Häufungspunkt von M zu M gehört.

Beispiele: Jedes offene Intervall (a, b) ist eine offene Menge. Offene Mengen sind auch $\mathbb{R}$, $\emptyset$, $(a, b) \cup (c, d)$. Nicht offen sind zum Beispiel $\mathbb{N}, \mathbb{Z}, \mathbb{Q}$.
Jedes abgeschlossene Intervall $[a, b]$ ist abgeschlossen (auch die einpunktigen Mengen $[a, a] = \{a\}$ sind abgeschlossen). Abgeschlossen sind auch die Intervalle

$$(-\infty, a], \quad [b, \infty), \quad (-\infty, \infty) = \mathbb{R}, \quad \emptyset,$$

und die Mengen $\mathbb{N}, \mathbb{Z}$, dagegen nicht $\mathbb{Q}$, weil jede reelle Zahl Häufungspunkt von $\mathbb{Q}$ ist.

Satz: Eine Menge M ist genau dann abgeschlossen, wenn ihre Komplementärmenge offen ist.

Beweis: Sei M abgeschlossen, und sei $x \in \mathbb{R} \backslash M$. Dann ist x nicht Häufungspunkt von M, also existiert eine Umgebung U von x mit $U \cap M = \emptyset$, also $U \subseteq \mathbb{R} \backslash M$, also ist x innerer Punkt von $\mathbb{R} \backslash M$, also ist $\mathbb{R} \backslash M$ offen. Sei umgekehrt $\mathbb{R} \backslash M$ offen, und sei x Häufungspunkt von M. Dann gibt es in jeder Umgebung von x ein Element aus M, also ist x kein innerer Punkt von $\mathbb{R} \backslash M$, also gilt $x \notin \mathbb{R} \backslash M$, da $\mathbb{R} \backslash M$ nur aus inneren Punkten besteht, also $x \in M$, also ist M abgeschlossen.

Satz: Die Vereinigung eines beliebigen Systems offener Mengen ist eine offene Menge. Der Durchschnitt eines endlichen Systems offener Mengen ist eine offene Menge.
Der Durchschnitt eines beliebigen Systems abgeschlossener Mengen ist eine abgeschlossene Menge. Die Vereinigung eines endlichen Systems abgeschlossener Mengen ist eine abgeschlossene Menge.

Beweis: Sei S ein System offener Mengen, und sei $x \in \bigcup_{M \in S} M$. Dann gibt es $N \in S$ mit $x \in N$. Da N offen ist, gibt es eine Umgebung U von x, die ganz zu N gehört, also auch zu $\bigcup_{M \in S} M$, also ist diese Menge offen.

Seien $M_1, \dots, M_n$ offene Mengen, und sei $x \in \bigcap_{i=1}^n M_i$. Dann gilt $x \in M_i$ für alle $i = 1, \dots, n$, und da M_i offen ist, gibt es $\varepsilon_i > 0$, $i = 1, \dots, n$, mit $U_{\varepsilon_i}(x) = \{y \in \mathbb{R} \mid |y - x| < \varepsilon_i\} \subseteq M_i$. Sei $\varepsilon := \min\{\varepsilon_1, \dots, \varepsilon_n\}$. Es gilt $\varepsilon > 0$, und $U_\varepsilon(x) \subseteq \bigcap_{i=1}^n U_{\varepsilon_i} \subseteq \bigcap_{i=1}^n M_i$, also ist $\bigcap_{i=1}^n M_i$ offen.

Die Aussagen über die abgeschlossenen Mengen folgen aus den de Morganschen Regeln

$$\mathbb{R} \backslash \Big(\bigcap_{M \in S} M \Big) = \bigcup_{M \in S} (\mathbb{R} \backslash M), \quad \mathbb{R} \backslash \Big(\bigcup_{M \in S} M \Big) = \bigcap_{M \in S} (\mathbb{R} \backslash M),$$

und aus dem vorangehenden Satz.

Beachte, daß der Durchschnitt eines unendlichen Systems offener Mengen nicht offen zu sein braucht, und die Vereinigung eines unendlichen Systems abgeschlossener Mengen nicht abgeschlossen zu sein braucht.

Beispiele:

$$\begin{aligned}\bigcap_{a>0}(-a,a) &= \{0\}\\ \bigcup_{0<a<1}[0,a] &= [0,1)\,.\end{aligned}$$

Sei $M \subseteq \mathbb{R}$. Betrachte die Menge S aller abgeschlossener Mengen $A \subseteq \mathbb{R}$ mit $M \subseteq A$. Es gibt mindestens eine solche Menge, nämlich $\mathbb{R}$. Also ist $S \neq \emptyset$, also ist $\overline{M} = \bigcap_{A\in S} A \supseteq M$, und $\overline{M}$ ist eine abgeschlossene Menge.

Definition: $\overline{M}$ heißt die abgeschlossene Hülle von M .

Satz: $\overline{M}$ besteht aus den Punkten von M und aus allen Häufungspunkten von M .

Beweis: Da $M \subseteq \overline{M}$ gilt, muß $\overline{M}$ mindestens aus den im Satz angegebenen Punkten bestehen, weil $\overline{M}$ sonst nicht abgeschlossen wäre. (Ein Häufungspunkt von M muß auch Häufungspunkt von $\overline{M}$ sein wegen $M \subseteq \overline{M}$.)

Sei umgekehrt $x \notin M$, und sei x kein Häufungspunkt von M . Dann ist x innerer Punkt von $\mathbb{R}\backslash M$, also existiert $\varepsilon > 0$ mit $U_\varepsilon(x) = (x-\varepsilon, x+\varepsilon) \subseteq \mathbb{R}\backslash M$, also $U_\varepsilon(x) \cap M = \emptyset$. Da $U_\varepsilon(x)$ offen ist, ist $\mathbb{R}\backslash U_\varepsilon(x)$ abgeschlossen, also folgt aus $M \subseteq \mathbb{R}\backslash U_\varepsilon(x)$ daß auch $\overline{M} \subseteq \mathbb{R}\backslash U_\varepsilon(x)$ gelten muß, folglich ist $x \notin \overline{M}$.

Definition: x heißt Berührungspunkt von $M \subseteq \mathbb{R}$, wenn in jeder Umgebung U von x ein $y \in M$ liegt.

x heißt isolierter Punkt von M , wenn $x \in M$ und wenn eine Umgebung U von x existiert, die außer x keine weiteren Punkte von M enthält.

x heißt Randpunkt von M , wenn in jeder Umgebung U von x ein $y \in M$ und ein $z \in \mathbb{R}\backslash M$ liegt.

Ähnlich wie für Folgen gilt:

Satz: a ist Häufungspunkt einer Menge $M \subseteq \mathbb{R}$ genau dann, wenn es eine Folge $\{x_n\}_{n=1}^\infty$ von lauter verschiedenen Zahlen $x_n \in M$ gibt, die gegen a konvergiert.

Beweis: Übung

Satz: (Satz von Bolzano–Weierstrass für Mengen)
Jede beschränkte unendliche Zahlenmenge M besitzt wenigstens einen Häufungspunkt.

Beweis: Da M unendlich ist, gibt es eine Folge $\{x_n\}_{n=1}^{\infty}$ aus lauter verschiedenen Zahlen $x_n \in M$. $\{x_n\}_{n=1}^{\infty}$ ist beschränkt, und besitzt folglich einen Häufungspunkt a (im Sinne von Folgen). Außerdem besitzt $\{x_n\}_{n=1}^{\infty}$ eine Teilfolge, die gegen a konvergiert. Nach dem vorangehenden Satz ist dann a auch Häufungswert von M (im Sinne von Zahlenmengen).

Definition: Eine Menge $M \subseteq \mathbb{R}$ heißt kompakt, genau dann wenn sie beschränkt und abgeschlossen ist.

Beispiele: Abgeschlossene Intervalle $[a,b]$ $(a, b \neq \infty)$ und alle endlichen Mengen sind kompakt. Eine endliche Vereinigung kompakter Mengen ist wieder kompakt. Nicht kompakt ist $\{\frac{1}{n} \mid n \in \mathbb{N}\}$, aber kompakt ist $\{\frac{1}{n} \mid n \in \mathbb{N}\} \cup \{0\}$.

Satz: Eine nichtleere kompakte Menge besitzt ein Minimum und ein Maximum.

Beweis: Da M beschränkt ist, existieren $s = \sup M$ und $u = \inf M$. Es muß gezeigt werden, daß $s, u \in M$ gilt. Für das Supremum s gilt entweder $s \in M$, oder s ist Häufungspunkt von M . Für kompakte Mengen gehört also s immer zu M . Genauso folgt dies für das Infimum u .

Der Satz von Bolzano–Weierstrass ergibt:

Satz: Jede unendliche Teilmenge einer kompakten Menge M gesitzt einen Häufungspunkt in M .

Die folgenden beiden Sätze geben Charakterisierungen kompakter Mengen:

Satz: Eine Menge M ist genau dann kompakt, wenn jede Folge mit Werten in M eine konvergente Teilfolge mit Grenzwert in M besitzt.

Beweis: Aus dem vorangehenden folgt, daß eine kompakte Menge diese Eigenschaft besitzt. Also muß bewiesen werden, daß M kompakt ist, wenn jede Folge eine in M konvergente Teilfolge hat. Zunächst ist klar, daß dann M beschränkt sein muß. Denn sonst gäbe es zu jedem $n \in \mathbb{N}$ ein $x_n \in M$ mit $|x_n| > n$. Für die Folge $\{x_n\}_{n \in \mathbb{N}}$ könnte keine Teilfolge konvergieren. Denn für die Teilfolge $\{x_{n_m}\}_{m \in \mathbb{N}}$ muß gelten $n_m \geq m$, also $|x_{n_m}| > m$, also ist jede Teilfolge unbeschränkt.
Außerdem muß M abgeschlossen sein. Denn sei x Häufungspunkt von M . Nach einem der vorangehenden Sätzen gibt es dann eine Folge $\{x_n\}_{n \in \mathbb{N}}$ mit Werten in M , die gegen x konvergiert. Diese Folge besitzt eine Teilfolge, die gegen ein $y \in M$ konvergiert. Da aber $\{x_n\}_{n \in \mathbb{N}}$ schon gegen x konvergiert, konvergiert also auch jede Teilfolge gegen x , also

$x = y \in M$, somit ist M abgeschlossen. ■

Ich komme nun zur zweiten Charakterisierung. Sei $M \subseteq \mathbb{R}$. Wie immer bezeiche ich für $x \in \mathbb{R}$ und $\varepsilon > 0$ mit $U_\varepsilon(x)$ die offene Umgebung

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

von x . Sei zu jedem $x \in \mathbb{R}$ eine solche Umgebung $U(x) = U_{\varepsilon(x)}(x)$ vorgegeben. Dann gilt natürlich

$$M \subseteq \bigcup_{x \in \mathbb{R}} U(x)$$

und

$$M \subseteq \bigcup_{x \in M} U(x)\,.$$

Man nennt ein System $\mathcal{U}$ derartiger offener Umgebungen mit

$$M \subseteq \bigcup_{U(x) \in \mathcal{U}} U(x)$$

eine offene Überdeckung von M .

Satz: (Heine (1821 – 1881) – Borel (1871 – 1956)) $M \subseteq \mathbb{R}$ ist genau dann kompakt, wenn man aus jeder offenen Überdeckung $\mathcal{U} = \{U(x)\}$ von M endlich viele $U(x_1), \ldots, U(x_n)$ auswählen kann, so daß schon

$$M \subseteq \bigcup_{i=1}^{n} U(x_i)$$

gilt. (Heine – Borelsche Überdeckungseigenschaft.)

Beweis: Sei M kompakt. Angenommen, M hätte nicht die Heine – Borelsche Überdeckungseigenschaft. Sei $\mathcal{U}$ eine offene Überdeckung, aus der nicht endlich viele Umgebungen zur Überdeckung von M ausgewählt werden können. Da M beschränkt ist, existieren $a_1, b_1 \in \mathbb{R}$ mit $M \subseteq [a_1, b_1] =: J_1$. Sei $m = \frac{a_1+b_1}{2}$. Nach Annahme sind wenigstens zur Überdeckung einer der beiden kompakten Mengen $[a_1, m] \cap M$ und $[m, b_1] \cap M$ unendlich viele Umgebungen aus $\mathcal{U}$ notwendig, weil sonst M doch durch endlich viele überdeckt werden könnte. Sei dies etwa die Menge $[a_1, m] \cap M$. Nun halbiere das Intervall $[a_1, m]$ und fahre mit der Konstruktion fort. Dies ergibt eine Folge $\{J_n\}_{n \in \mathbb{N}}$ von Intervallen mit

$$J_1 \supseteq J_2 \supseteq J_3 \supseteq \ldots\,,$$

so daß zur Überdeckung jeder Menge $M \cap J_n$ unendlich viele Umgebungen aus $\mathcal{U}$ notwendig sind. Oben wurde bewiesen, daß es für diese Intervallschachtelung genau ein $s \in \bigcap_{n=1}^{\infty} J_n$

gibt. Es gilt $s \in M$, weil jede der Mengen $J_n \cap M$ unendlich viele Elemente der Menge M enthält, also s Häufungspunkt von M ist, folglich zu M gehört. Also gibt es $U(z) = U_{\varepsilon(z)}(z) \in \mathcal{U}$ mit $s \in U_{\varepsilon(z)}(z)$, also $\delta = \varepsilon(z) - |s - z| > 0$. Da die Intervallfolge $\{J_n\}_{n=1}^{\infty}$ sich auf s zusammenzieht, gibt es $n_0 \in \mathbb{N}$ so daß für $J_{n_0} = [a_{n_0}, b_{n_0}]$ gilt $|b_{n_0} - a_{n_0}| < \delta$, also gilt für alle $x \in J_{n_0}$ wegen $s \in J_{n_0}$

$$\begin{aligned} |x - z| &= |x - s + s - z| \le |x - s| + |s - z| \\ &\le |b_{n_0} - a_{n_0}| + |s - z| < \delta + |s - z| = \varepsilon(z), \end{aligned}$$

also

$$x \in U(z), \quad \text{d.h.} \quad J_{n_0} \subseteq U_{\varepsilon(z)}(z),$$

also genügt zur Überdeckung von J_{n_0} im Widerspruch zur Annahme nur eine einzige Umgebung aus $\mathcal{U}$, also muß die kompakte Menge M doch die Heine – Borelsche Überdeckungseigenschaft haben.

Umgekehrt habe M die Heine – Borelsche Überdeckungseigenschaft. Zu zeigen ist, daß M kompakt ist. M ist beschränkt, weil es von endlich vielen beschränkten Umgebungen überdeckt werden kann. Wäre M nicht abgeschlossen, dann würde es einen Häufungspunkt a von M geben mit $a \notin M$. Dann würde die folgende offene Überdeckung von M nicht die Heine – Borelsche Überdeckungseigenschaft haben:

Für $x \in M$ setze $\varepsilon(x) := \frac{1}{2}|x - a| > 0$. Die Menge $\mathcal{U} = \{U_{\varepsilon(x)}(x) \mid x \in M\}$ ist eine offene Überdeckung von M, aber zur Überdeckung von M reichen nicht endlich viele Umgebungen $U_{\varepsilon(x_1)}(x_1), \ldots, U_{\varepsilon(x_n)}(x_n)$ aus. Denn setze

$$\varepsilon := \min\{\varepsilon(x_1), \ldots, \varepsilon(x_n)\} > 0.$$

Dann gilt

$$U_\varepsilon(a) \cap U_{\varepsilon(x_i)}(x_i) = \emptyset,$$

für alle $i = 1, \ldots, n$, aber es gibt $x \in M$ mit $x \in U_\varepsilon(a)$, weil a Häufungspunkt von M ist, also ist M nicht Teilmenge von $\bigcup_{i=1}^{n} U_{\varepsilon(x_i)}(x_i)$. Damit also M die Heine – Borel Eigenschaft haben kann, muß M abgeschlossen sein. ∎

6b.) Stetigkeit von Funktionen

Bei sehr vielen Funktionen, die in der Mathematik und ihren Anwendungen vorkommen, ändert sich der Funktionswert nur wenig, wenn man das Argument der Funktion nur wenig ändert. Diese Funktionen nennt man stetige Funktionen. Genau werden stetige

Funktionen folgendermaßen definiert:

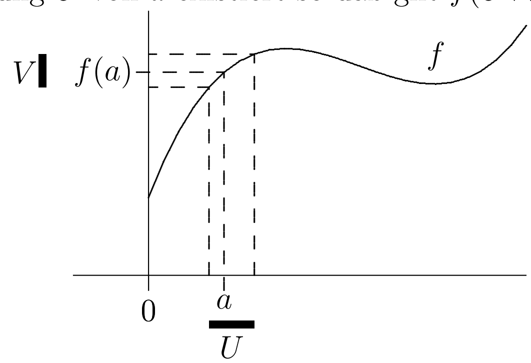
Definition: Sei $D \subseteq \mathbb{R}$. Die Funktion $f : D \to \mathbb{R}$ heißt stetig in $a \in D$, wenn zu jedem $\varepsilon > 0$ ein $\delta > 0$ existiert so daß für alle $x \in D$ mit $|x - a| < \delta$ gilt $|f(x) - f(a)| < \varepsilon$.

Mit Quantoren:

$$\forall_{\varepsilon>0} \quad \exists_{\delta>0} \quad \forall_{\substack{x \in D \\ |x-a|<\delta}} \quad : \quad |f(x) - f(a)| < \varepsilon\,.$$

Äquivalent dazu ist folgende Definition:

Definition: Die Funktion $f : D \to \mathbb{R}$ heißt stetig in $a \in D$, wenn zu jeder Umgebung V von $f(a)$ eine Umgebung U von a existiert so daß gilt $f(U \cap D) \subseteq V$.



Beispiele: 1.) Die identische Abbildung $(x \mapsto \mathrm{id}\,(x) = x) : \mathbb{R} \to \mathbb{R}$ ist an jeder Stelle $a \in \mathbb{R}$ stetig.

Beweis: Sei $\varepsilon > 0$ gegeben. Wähle $\delta = \varepsilon$. Dann folgt trivialerweise für $x \in \mathbb{R}$ mit $|x - a| < \delta$

$$|\,\mathrm{id}\,(x) - \mathrm{id}\,(a)| = |x - a| < \delta = \varepsilon\,.$$

2.) Die Quadratwurzelfunktion $Q : \mathbb{R}_0^+ \to \mathbb{R}$,

$$x \mapsto Q(x) = \sqrt{x}\,,$$

ist in jedem $a \in \mathbb{R}_0^+$ stetig.

Beweis: Zunächst sei $a = 0$. Sei $\varepsilon > 0$ gegeben. Wähle $\delta = \varepsilon^2 > 0$. Dann folgt für alle $x \in \mathbb{R}_0^+$ mit $|x - a| = x < \delta$ daß gilt

$$|Q(x) - Q(a)| = |\sqrt{x} - \sqrt{a}| = \sqrt{x} < \sqrt{\delta} = \varepsilon\,,$$

wegen der strengen Monotonie der Quadratwurzel.

Sei $a > 0$ und sei $\varepsilon > 0$ gegeben. Um $\delta > 0$ zu finden, betrachte man folgende Rechnung

$$\sqrt{x} - \sqrt{a} = \frac{(\sqrt{x} - \sqrt{a})(\sqrt{x} + \sqrt{a})}{\sqrt{x} + \sqrt{a}} = \frac{x - a}{\sqrt{x} + \sqrt{a}}\,.$$

Hieraus folgt, daß $\delta = \sqrt{a} \cdot \varepsilon$ gewählt werden kann. Denn für $x \in \mathbb{R}_0^+$ und $|x - a| < \delta$ folgt dann

$$|\sqrt{x} - \sqrt{a}| = \frac{|x - a|}{\sqrt{x} + \sqrt{a}} \leq \frac{|x - a|}{\sqrt{a}} < \frac{\delta}{\sqrt{a}} = \frac{\sqrt{a}}{\sqrt{a}}\varepsilon = \varepsilon\,.$$

3.) Sei $f : \mathbb{R}^+ \to \mathbb{R}$ definiert durch

$$x \mapsto f(x) = \begin{cases} 0\,, & \text{für irrationales} \quad x \\ \dfrac{1}{p+q}\,, & \text{für} \quad x = \dfrac{p}{q} \quad \text{mit teilerfremden Zahlen} \quad p, q \in \mathbb{N}\,. \end{cases}$$

Dann ist f unstetig in jedem rationalen $x > 0$, aber stetig in jedem irrationalen $x > 0$.

Denn sei $x = \frac{p}{q}$ rational. Um zu zeigen, daß f in x unstetig ist muß man zeigen, daß die Aussage in der Stetigkeitsdefinition falsch ist für rationales x, also daß die Negation dieser Aussage wahr ist. Diese Negation lautet

$$(*) \qquad \exists_{\varepsilon>0} \quad \forall_{\delta>0} \quad \exists_{\substack{y\in\mathbb{R}^+ \\ |y-x|<\delta}} \quad : \quad |f(y) - f(x)| \geq \varepsilon\,.$$

Es gilt $f(x) = \frac{1}{p+q} > 0$. Wähle $\varepsilon = \frac{1}{2}\frac{1}{p+q} > 0$. Dann gibt es zu jedem $\delta > 0$ noch eine irrationale Zahl y mit $x < y < x + \delta$, also mit $|x - y| < \delta$, weil zwischen zwei reellen Zahlen immer noch eine irrationale liegt. Für dieses y gilt

$$|f(x) - f(y)| = |\frac{1}{p+q} - 0| = \frac{1}{p+q} = 2\varepsilon > \varepsilon\,,$$

also ist $(*)$ wahr, und die Funktion f kann in rationalem x nicht stetig sein.
Sei nun x irrational. Dann ist $f(x) = 0$. Sei $\varepsilon > 0$ gegeben. Gesucht ist $\delta > 0$. Es gibt höchstens endlich viele Paare (p, q) natürlicher Zahlen mit $p + q \leq \frac{1}{\varepsilon}$, also $\frac{1}{p+q} \geq \varepsilon$.

Sei

$$\delta = \frac{1}{2}\min\left\{|\frac{p}{q} - x| \;\Big|\; \frac{1}{p+q} \geq \varepsilon\right\} > 0\,,$$

oder $\delta = 1$, falls es kein solches Paar gibt. Für alle rationalen $y = \frac{p}{q}$ mit $|y - x| < \delta$ gilt also

$$|f(x) - f(y)| = f(y) = \frac{1}{p+q} < \varepsilon\,.$$

Für alle irrationalen y mit $|x - y| < \delta$ gilt sowieso $|f(x) - f(y)| = 0 < \varepsilon$, also ist f stetig in irrationalem x.

4.) Die Dirchlet–Funktion

$$f(x) = \begin{cases} 1\,, & x \in \mathbb{Q} \\ 0\,, & x \in \mathbb{R}\backslash\mathbb{Q}\,, \end{cases}$$

ist natürlich in keinem Punkt $x \in \mathbb{R}$ stetig.

Definition: Sei $D \subseteq \mathbb{R}$. Die Funktion $f : D \to \mathbb{R}$ heißt stetig, wenn sie in jedem Punkt von D stetig ist.

Satz: Eine Funktion $f : D \to \mathbb{R}$ ist in $a \in D$ stetig genau dann, wenn für jede Folge $\{x_n\}_{n=1}^{\infty}$ mit $x_n \in D$ und $\lim_{n\to\infty} x_n = a$ gilt

$$\lim_{n\to\infty} f(x_n) = f(\lim_{n\to\infty} x_n) = f(a)\,.$$

(Insbesondere bedeutet dies, daß für stetige Funktionen das Funktionssymbol mit dem Grenzwertsymbol vertauscht werden darf.)

Beweis: Sei f in $a \in D$ stetig, und sei $\{x_n\}_{n=1}^{\infty}$ mit $x_n \in D$ und $\lim_{n\to\infty} x_n = a$. Sei $\varepsilon > 0$ gegeben. Zu zeigen ist, daß $n_0 \in \mathbb{N}$ existiert mit $|f(x_n) - f(a)| < \varepsilon$ für alle $n \geq n_0$. Da f in a stetig ist, existiert $\delta > 0$ mit $|f(x) - f(a)| < \varepsilon$ für alle $|x - a| < \delta$. Wegen $\lim_{n\to\infty} x_n = a$ gibt es auch $n_0 \in \mathbb{N}$ mit $|x_n - a| < \delta$ für alle $n \geq n_0$, also

$$|f(x_n) - f(a)| < \varepsilon\,, \quad \text{für alle} \quad n \geq n_0\,.$$

Sei f in $a \in D$ nicht stetig. Zu zeigen ist, daß es dann eine Folge $\{x_n\}_{n=1}^{\infty}$ gibt mit $x_n \in D$, $\lim_{n\to\infty} x_n = a$, für die aber $\{f(x_n)\}_{n=1}^{\infty}$ nicht gegen $f(a)$ konvergiert.

Da f in a nicht stetig ist, gilt

$$\exists_{\varepsilon>0} \quad \forall_{\delta>0} \quad \exists_{\substack{x\in D \\ |x-a|<\delta}} \quad : \quad |f(x) - f(a)| \geq \varepsilon\,.$$

Also gibt es $\varepsilon > 0$, so daß man insbesondere für jedes $n \in \mathbb{N}$ zu $\delta = \frac{1}{n}$ ein x_n mit $|x_n - a| < \frac{1}{n}$ finden kann mit $|f(x_n) - f(a)| \geq \varepsilon > 0$. Die so konstruierte Folge $\{x_n\}_{n=1}^{\infty}$ konvergiert gegen a wegen

$$|x_n - a| < \frac{1}{n}\,,$$

aber wegen $|f(x_n) - f(a)| \geq \varepsilon > 0$ für alle $n \in \mathbb{N}$ konvergiert $\{f(x_n)\}_{n=1}^{\infty}$ nicht gegen $f(a)$. ∎

Folgerung: Alle rationalen Funktionen sind stetig.

Beweis: Sei $F : D \to \mathbb{R}$ rationale Funktion. Es wurde gezeigt daß $\lim_{n\to\infty} F(x_n) = F(\lim_{n\to\infty} x_n)$ gilt für alle $\{x_n\}_{n=1}^{\infty}$ mit $x_n \in D$ und $\lim_{n\to\infty} x_n \in D$. Also ist F in ganz D stetig nach dem vorangehenden Satz. ∎

Auf diese Weise können weitere Resultate erhalten werden:

Satz: Seien $f : D \to \mathbb{R}$ und $g : D \to \mathbb{R}$ in $a \in D$ stetig. Dann sind auch die in D erklärten Funktionen

$$f+g\,, \quad cf \quad (c \in \mathbb{R})\,, \quad f \cdot g, \quad |f|$$

an der Stelle $a \in D$ stetig.

Gilt $f(a) \neq 0$, dann ist auch die Funktion $\frac{1}{f}$ an der Stelle a stetig.

Beachte, daß $\frac{1}{f}$ erklärt ist auf der Menge $D' = \{x \mid x \in D \text{ und } f(x) \neq 0\}$. Wenn $f(a) \neq 0$ ist, dann gibt es wegen der Stetigkeit von f in a noch eine ganze Umgebung U von a mit $f(x) \neq 0$ für alle $x \in U$, also ist $\frac{1}{f}$ auch noch in einer ganzen Umgebung U von a erklärt. Denn für $\varepsilon = \frac{|f(a)|}{2} > 0$ existiert $\delta > 0$ so daß $|f(x) - f(a)| < \varepsilon$ ist für $|x-a| < \delta$, also

$$\begin{aligned} |f(x)| &= |f(x) - f(a) + f(a)| \geq \Big| |f(a)| - |f(x) - f(a)| \Big| \\ &\geq |f(a)| - |f(x) - f(a)| > \frac{1}{2}|f(a)| = \varepsilon > 0 \end{aligned}$$

für $|x-a| < \delta$.

Beweis des Satzes: Für jede Folge $\{x_n\}_{n=1}^{\infty}$ mit $\lim_{n\to\infty} x_n = a \in D$ gilt $\lim_{n\to\infty} f(x_n) = f(a)$, $\quad \lim_{n\to\infty} g(x_n) = g(a)$, also gilt auch

$$\begin{aligned} \lim_{n\to\infty} (f+g)(x_n) &= \lim_{n\to\infty} \Big(f(x_n) + g(x_n) \Big) = f(a) + g(a) = (f+g)(a)\,, \\ \lim_{n\to\infty} (cf)(x_n) &= c \lim_{n\to\infty} f(x_n) = (cf)(a)\,, \end{aligned}$$

also sind $f+g$ und cf stetig. Ebenso folgt dies für $f \cdot g$ und $|f|$. Wenn $f(a) \neq 0$ ist, dann gibt es ein $n_0 \in \mathbb{N}$ mit $f(x_n) \neq 0$ für $n \geq n_0$. Die Folge $\{\frac{1}{f(x_n)}\}_{n=1}^{\infty}$ konvergiert dann gegen $\frac{1}{f(a)}$, also ist auch $\frac{1}{f}$ stetig. ∎

Folgerung: Die Menge aller Funktionen $f : D \to \mathbb{R}$, die an der Stelle $a \in D$ stetig sind, ist ein Vektorraum über $\mathbb{R}$. Ebenso ist die Menge $C(D)$ aller in D stetigen Funktionen ein Vektorraum über $\mathbb{R}$.

Satz: Seien $f : D_1 \to \mathbb{R}$, $g : D_2 \to \mathbb{R}$ $(D_1, D_2 \subseteq \mathbb{R})$ und sei $g \circ f$ definiert. Sei f an der Stelle $a \in D_1$ stetig und g an der Stelle $b = f(a)$. Dann ist $g \circ f$ an der Stelle a stetig.

Beweis: Sei $\{x_n\}_{n=1}^{\infty}$ eine Folge mit $x_n \in D_1$ und $\lim_{n\to\infty} x_n = a$. Dann gilt $\lim_{n\to\infty} f(x_n) = f(a) \in D_2$ und $f(x_n) \in D_2$, also auch $\lim_{n\to\infty}(g \circ f)(x_n) = \lim_{n\to\infty} g(f(x_n)) = g(f(a)) = (g \circ f)(a)$, also ist $g \circ f$ an der Stelle a stetig. ∎

6c.) Abbildungseigenschaften stetiger Funktionen

Sei $D \subseteq \mathbb{R}$ offen und $f : D \to \mathbb{R}$ stetig. Dann braucht $f(D)$ nicht offen zu sein.
Beispiel: $x \mapsto x^2$ bildet das offene Intervall $(-1, 1)$ auf das halboffene Intervall $[0, 1)$ ab.

Wenn D abgeschlossen ist, muß $f(D)$ nicht abgeschlossen sein.
Beispiel: $x \mapsto \frac{1}{1+x^2}$ bildet $\mathbb{R}$ auf $(0, 1]$ ab.
Es gilt jedoch:

Satz: Ist $D \subseteq \mathbb{R}$ kompakt und $f : D \to \mathbb{R}$ stetig, dann ist $f(D)$ kompakt.

Beweis: Ich zeige, daß mit D auch $f(D)$ die Heine – Borel–Überdeckungseigenschaft hat. Sei $\mathcal{U}$ eine offene Überdeckung von $f(D)$. Dann bildet das Mengensystem $\{f^{-1}(U)\}_{U \in \mathcal{U}}$ eine Überdeckung von D. Ich zeige nun, daß bereits endlich viele Mengen $U_1, \ldots, U_n$ genügen um $f(D)$ zu überdecken. Wenn dies gezeigt ist folgt, daß $f(D)$ kompakt ist.

Sei $U \in \mathcal{U}$. Man ordne jedem $x \in f^{-1}(U)$ eine Umgebung $V(x)$ zu mit $f(D \cap V(x)) \subseteq U$. Eine solche Umgebung von x existiert, weil U offen ist, also noch eine ganze Umgebung $W(f(x))$ von $f(x)$ enthält. Wegen der Stetigkeit von f kann nun $V(x)$ gefunden werden mit $f(D \cap V(x)) \subseteq W(f(x)) \subseteq U$.

Führt man diese Konstruktion für jedes $U \in \mathcal{U}$ durch, erhält man ein Mengensystem $\{V(x)\}$, das zu jedem $x \in D$ mindestens eine Umgebung $V(x)$ enthält. Also ist $\{V(x)\}$ eine offene Überdeckung von D, und es genügen bereits endlich viele $V(x_1), \ldots, V(x_n)$ zur Überdeckung von D. Dann gilt auch $D = (V(x_1) \cap D) \cup \ldots \cup (V(x_n) \cap D)$. Zu jedem $x_i, i = 1, \ldots, n$ wähle $U_i \in \mathcal{U}$ mit $(f(V(x_i) \cap D) \subseteq U_i$, also

$$\begin{aligned} f(D) &= f\Big(\Big(V(x_1) \cap D\Big) \cup \ldots \cup \Big(V(x_n) \cap D\Big)\Big) \\ &= f\Big(V(x_1) \cap D\Big) \cup \ldots \cup f\Big(V(x_n) \cap D\Big) \\ &\subseteq U_1 \cup \ldots \cup U_n . \end{aligned}$$

■

Folgerung: Ist $D \neq \emptyset$ kompakt und $f : D \to \mathbb{R}$ stetig, dann nimmt f das Minimum und das Maximum in D an, d.h. es gibt $x_1, x_2 \in D$ mit $f(x_1) = \min f(D)$, $\quad f(x_2) = \max f(D)$.

Beweis: klar, weil $f(D) \neq \emptyset$ Minimum und Maximum besitzt als kompakte Menge. ■

Die Umkehrfunktion einer stetigen Funktion muß nicht stetig sein. Es gilt aber

Satz: Sei $D \subseteq \mathbb{R}$ kompakt, und $f : D \to \mathbb{R}$ injektiv und stetig. Dann ist auch

$$f^{-1} : f(D) \to \mathbb{R}$$

stetig.

Beweis: Ich nehme an, f^{-1} sei in $b \in f(D)$ nicht stetig. Dann gibt es eine Folge $\{y_n\}_{n=1}^{\infty}$ mit $y_n \in f(D)$ und $\lim_{n\to\infty} y_n = b$, so daß $\{x_n = f^{-1}(y_n)\}_{n=1}^{\infty}$ nicht gegen $a := f^{-1}(b)$ konvergiert.

Die Folge $\{x_n\}_{n=1}^{\infty}$ ist beschränkt und besitzt daher Häufungspunkte. Es gibt mindestens einen Häufungspunkt $a' \in D$ mit $a' \neq a$, denn sonst wäre

$$\overline{\lim_{n\to\infty}} x_n = \underline{\lim}_{n\to\infty} x_n = a\,,$$

also $\lim_{n\to\infty} x_n = a$, im Widerspruch zur Annahme. Sei $\{x_{n_j}\}_{j=1}^{\infty}$ eine Teilfolge von $\{x_n\}_{n=1}^{\infty}$, die gegen a' konvergiert. Wegen der Stetigkeit von f folgt dann

$$\begin{aligned} f(a') &= f(\lim_{j\to\infty} x_{n_j}) = \lim_{j\to\infty} f(x_{n_j}) \\ &= \lim_{j\to\infty} y_{n_j} = b\,, \end{aligned}$$

also $a' = f^{-1}(b) = a$, im Widerspruch zur Annahme $a \neq a'$. Also muß f^{-1} in jedem Punkt $b \in f(D)$ stetig sein. ∎

Satz (Zwischenwertsatz): Sei f im Intervall $[a, b]$ stetig, und es gelte $f(a) < f(b)$. Dann gibt es zu jeder Zahl y zwischen $f(a)$ und $f(b)$ ein z, $a < z < b$, mit $f(z) = y$.

Anders formuliert: Das Intervall $[f(a), f(b)]$ ist im Wertebereich von f enthalten.

Beweis: Sei $f(a) < y < f(b)$, und sei

$$M_y = \{x \mid f(x) \leq y\,,\ a \leq x \leq b\}\,.$$

Es gilt $M_y \neq \emptyset$, wegen $a \in M_y$. Außerdem ist M_y durch b nach oben beschränkt, also existiert

$$z = \sup M_y\,.$$

Ich zeige, daß

$$f(z) = y$$

gilt. Hierzu beweist man, daß weder

1.) $f(z) < y$

noch

2.) $f(z) > y$

gelten kann.

Angenommen, 1.) sei richtig. Zunächst beachte man, daß $z \neq b$ ist, denn sonst müßte wegen 1.) $f(b) < y < f(b)$ sein, also erhielte man einen Widerspruch.

Sei $\varepsilon = y - f(z)$. Dann ist $\varepsilon > 0$, und es existiert $\delta > 0$, so daß für alle $x \in \mathbb{R}$ mit $|x - z| < \delta$ gilt $|f(x) - f(z)| < \varepsilon$, also

$$\begin{aligned} f(x) &= f(x) - f(z) + f(z) \\ &\leq f(z) + |f(x) - f(z)| < f(z) + \varepsilon \\ &= f(z) + y - f(z) = y \end{aligned}$$

also gibt es $x \in [a, b]$ mit $z < x < z + \delta$ und $f(x) < y$, also $x \in M_y$. Dies ist ein Widerspruch zur Annahme $z = \sup M_y$.

Angenommen, 2.) sei richtig. Sei $\varepsilon = f(z) - y$. Dann ist $\varepsilon > 0$, und es kann $\delta > 0$ so gewählt werden, daß für alle $x \in [a, b]$ mit $|x - z| < \delta$ gilt $|f(x) - f(z)| < \varepsilon$, also

$$\begin{aligned} f(x) &= f(x) - f(z) + f(z) \\ &\geq f(z) - |f(x) - f(z)| > f(z) - \varepsilon = y\,, \end{aligned}$$

also kann in der Umgebung $U_\delta(z)$ kein Element von M_y enthalten sein, also kann z nicht Supremum von M_y sein, im Widerspruch zur Annahme. ■

Folgerung: Ist f in $[a, b]$ stetig und gilt $f(a) \cdot f(b) < 0$, dann besitzt f eine Nullstelle. Insbesondere besitzen alle ganzrationalen Funktionen mit ungeradem Grad mindestens eine Nullstelle.

Beweis: Sei zum Beispiel $f(a) < 0$, $f(b) > 0$. Dann gilt $0 \in [f(a), f(b)]$, also folgt die erste Behauptung aus dem Zwischenwertsatz.

Es sei

$$p(x) = a_{2n+1}\, x^{2n+1} + \ldots + a_1 x + a_0$$

ein Polynom ungeraden Grades mit $a_{2n+1} > 0$. Dann gilt

$$p(x) = x^{2n+1}\left[a_{2n+1} + \frac{a_{2n}}{x} + \ldots + \frac{a_1}{x^{2n}} + \frac{a_0}{x^{2n+1}}\right].$$

Für genügend große $|x|$ ist

$$a_{2n+1} + \frac{a_{2n}}{x} + \ldots + \frac{a_0}{x^{2n+1}} > 0\,,$$

also ist $p(x) > 0$ für genügend große $x > 0$, und $p(x) < 0$ für genügend kleine $x < 0$, also hat p eine Nullstelle. ∎

Satz: Das stetige Bild eines kompakten Intervalls ist ein kompaktes Intervall.

Beweis: Sei $f[a,b] \to \mathbb{R}$ stetig. $f([a,b])$ ist kompakt, also existieren $x_1, x_2 \in [a,b]$ mit $f(x_1) = c = \min f([a,b])$, $f(x_2) = d = \max f([a,b])$. Natürlich gilt

$$f\Big([a,b]\Big) \subseteq [c,d] .$$

Aus dem Zwischenwertsatz folgt

$$[c,d] \subseteq f\Big([a,b]\Big) ,$$

also

$$[c,d] = f\Big([a,b]\Big) .$$

∎

Satz: Sei $f : [a,b] \to \mathbb{R}$ stetig und injektiv. Dann ist f streng monoton in $[a,b]$.

Eine in einem Intervall definierte stetige Funktion ist also genau dann injektiv, wenn sie streng monoton ist.

Beweis: Es gilt entweder $f(a) < f(b)$ oder $f(a) > f(b)$. O.B.d.A. sei $f(a) < f(b)$. Das Minimum von f wird an der Stelle a angenommen. Denn würde das Minimum an der Stelle $x > a$ angenommen, dann müßte wegen der Injektivität von f für das Minimum $f(x) < f(a)$ gelten. Nach dem Zwischenwert gäbe es somit $y \in (x,b)$ mit $f(y) = f(a)$, also könnte f nicht injektiv sein. Ebenso folgt, daß das Maximum von f an der Stelle b angenommen wird.

Wäre f nicht streng monoton wachsend, gäbe es $x_1, x_2 \in [a,b]$ mit $x_1 < x_2$ und $f(x_1) > f(x_2)$. Wegen

$$f(a) < f(x_2) < f(x_1)$$

müßte es dann im Intervall (a, x_1) ein y geben mit $f(y) = f(x_2)$, also wäre f nicht injektiv, also muß f doch streng monoton wachsend sein. ∎

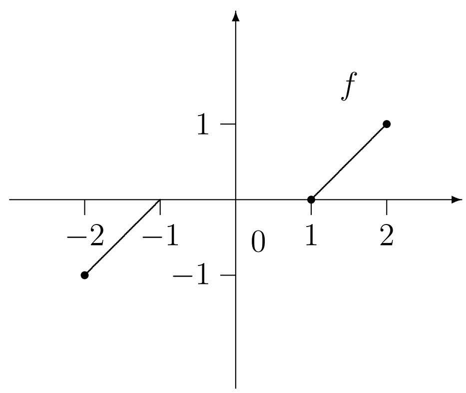
Satz: Sei D offen und $f : D \to \mathbb{R}$ stetig und injektiv. Dann ist $f^{-1} : f(D) \to \mathbb{R}$ stetig.

Beweis: Sei $y \in f(D)$, und sei U eine Umgebung von $x = f^{-1}(y)$. Da D offen ist, kann man $U \subseteq D$ voraussetzen. Sei $[a,b] \subseteq U$ ein abgeschlossenes Intervall, das x als inneren Punkt enthält. Da f stetig und injektiv, also streng monoton auf $[a,b]$ ist, wird $[a,b]$

durch f auf ein Intervall abgebildet, das $y = f(x)$ als inneren Punkt enthält. Sei V eine Umgebung von y, die in diesem Intervall enthalten ist. Dann gilt $f^{-1}(V) \subseteq U$, also ist f^{-1} stetig in jedem beliebigen $y \in f(D)$. ■

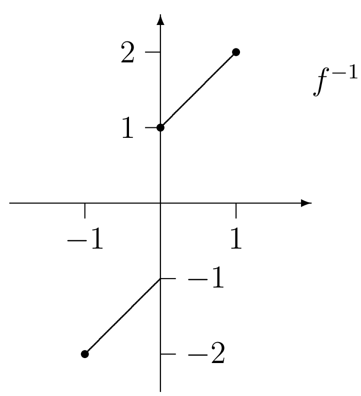
Beispiel für eine stetige, streng monoton wachsende Funktion, deren Inverse nicht stetig ist: Sei $D = [-2, -1) \cup [1, 2]$ und

$$f(x) = \begin{cases} x + 1\,, & \text{falls} \quad x \in [-2, -1) \\ x - 1\,, & \text{falls} \quad x \in [1, 2]\,. \end{cases}$$



$f : D \to [-1, 1]$ ist stetig und streng monoton wachsend, und

$$f^{-1}(x) = \begin{cases} x - 1\,, & \text{falls} \quad x \in [-1, 0) \\ x + 1\,, & \text{falls} \quad x \in [0, 1]\,. \end{cases}$$



Die Funktion f^{-1} ist nicht stetig.

6d.) Grenzwerte von Funktionen, gleichmäßige Stetigkeit

Sei $f : D \to \mathbb{R}$ stetig, und sei a ein Häufungspunkt von D, der nicht zu D gehört. Wann kann f zu einer stetigen Funktion auf $D \cup \{a\}$ fortgesetzt werden? Zu dieser Frage betrachte man folgende drei Beispiele: Es seien $a = 0$, $D = [-1, 0) \cup (0, 1]$ und

1.) $x \mapsto f(x) = \dfrac{1}{|x|}$,

2.) $x \mapsto f(x) = \dfrac{x}{|x|}$,

3.) $x \mapsto f(x) = \dfrac{x^2}{|x|}$.

Alle drei Funktionen sind auf D stetig, aber in den Fällen 1.) und 2.) existiert keine stetige Funktion $g : D \cup \{a\} = [-1, 1] \to \mathbb{R}$ mit $g\big|_D = f$. Im ersten Fall ist dies sofort klar, weil $[-1, 1]$ kompakt ist, also g beschränkt sein müßte, also auch f. Im zweiten Fall sieht man dies folgendermaßen: Wenn g existierte, müßte für alle Folgen $\{x_n\}_{n=1}^{\infty}$ mit $x_n \in D$ und $\lim_{n\to\infty} x_n = a$ wegen $g(x_n) = f(x_n)$ auch

$$\lim_{n\to\infty} f(x_n) = \lim_{n\to\infty} g(x_n) = g(a)$$

gelten. Das zweite Beispiel erfüllt diese Bedingung nicht, denn für $\{(-1)^n \frac{1}{n}\}_{n=1}^{\infty}$ gilt $\lim_{n\to\infty}(-1)^n \frac{1}{n} = 0$, aber $\{f((-1)^n \frac{1}{n})\}_{n=1}^{\infty} = \{(-1)^n\}_{n=1}^{\infty}$ konvergiert nicht.

Es gilt: Falls für alle Folgen $\{x_n\}_{n=1}^{\infty}$ mit $x_n \in D$ und $\lim_{n\to\infty} x_n = a$ die Folge $\{f(x_n)\}_{n=1}^{\infty}$ konvergiert, dann konvergieren alle diese Folgen gegen denselben Grenzwert. Denn seien $\{x'_n\}_{n=1}^{\infty}$ und $\{x''_n\}_{n=1}^{\infty}$ solche Folgen, dann ist auch $\{x_n\}_{n=1}^{\infty}$ mit

$$x_n = \begin{cases} x'_m\,, & n = 2m-1 \\ x''_m\,, & n = 2m \end{cases}$$

eine solche Folge, für die $\{f(x_n)\}_{n=1}^{\infty}$ konvergiert, und da $\{f(x'_n)\}_{n=1}^{\infty}$ und $\{f(x''_n)\}_{n=1}^{\infty}$ Teilfolgen sind, muß gelten

$$\lim_{n\to\infty} f(x_n) = \lim_{n\to\infty} f(x'_n) = \lim_{n\to\infty} f(x''_n)\,.$$

Definition: Seien $f : D \to \mathbb{R}$, $D \subseteq \mathbb{R}$, und a Häufungspunkt von D. Wenn für jede Folge $\{x_n\}_{n=1}^{\infty}$ mit $x_n \in D$ und $\lim_{n\to\infty} x_n = a$ gilt

$$\lim_{n\to\infty} f(x_n) = b\,,$$

dann sagt man, f habe an der Stelle a den Grenzwert b. Man schreibt in diesem Fall

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

Beachte, daß f im Punkt a nicht definiert zu sein braucht.

Mit dieser Definition folgt: Sei a ein Häufungspunkt von D, der zu D gehört. Die Funktion $f : D \to \mathbb{R}$ ist genau dann stetig in a, wenn $\lim_{x \to a} f(x) = f(a)$ gilt.

Satz: Seien $D \subseteq \mathbb{R}$, a Häufungspunkt von D und $f : D \to \mathbb{R}$. Genau dann existiert eine Funktion $g : D \cup \{a\} \to \mathbb{R}$ mit $g|_D = f$, die in a stetig ist, wenn f einen Grenzwert besitzt an der Stelle a. $g(a)$ ist dann eindeutig bestimmt, und es gilt

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

Beweis: Es wurde schon gezeigt, daß $g(a) = \lim_{x \to a} f(x)$ gelten muß, falls g existiert. Umgekehrt gilt: Hat f an der Stelle a einen Grenzwert, dann ist die Funktion

$$g(x) = \begin{cases} f(x) , & x \in D \backslash \{a\} \\ \lim\limits_{x \to a} f(x) , & x = a \end{cases}$$

eine Fortsetzung von f auf $D \cup \{a\}$. Nach Definition des Grenzwertes $\lim_{x \to a} f(x)$ gilt, daß für jede Folge $\{x_n\}_{n=1}^{\infty}$ mit $x_n \in D \cup \{a\}$ und $\lim_{n \to \infty} x_n = a$ gilt $\lim_{n \to \infty} g(x_n) = g(a)$, also ist g stetig in a. ∎

Ebenso wie im Fall der Stetigkeit kann man bei der Definition des Grenzwertes von Funktionen auch mit ε und δ arbeiten. Denn es gilt:

Satz: Seien $D \subseteq \mathbb{R}$, a Häufungspunkt von D, $b \in \mathbb{R}$ und $f : D \to \mathbb{R}$. Dann sind die folgenden Aussagen äquivalent:

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

(ii) Zu jedem $\varepsilon > 0$ existiert ein $\delta > 0$ mit

$$|f(x) - b| < \varepsilon$$

für alle $x \in D$ mit $|x - a| < \delta$.

Die zweite Aussage kann mit Quantoren folgendermaßen formuliert werden:

$$\forall_{\varepsilon > 0} \quad \exists_{\delta > 0} \quad \forall_{\substack{x \in D \\ |x-a| < \delta}} : \quad |f(x) - b| < \varepsilon .$$

Mit dem Umgebungsbegriff kann die Aussage des Satzes auch folgendermaßen formuliert werden: f hat in a den Grenzwert b, genau dann wenn zu jeder Umgebung U von b eine Umgebung V von a existiert mit $f(V \cap D) \subseteq U$.

Beweis des Satzes: Sei (i) erfüllt. Nach dem vorangehenden Satz ist dann

$$g(x) = \begin{cases} f(x) , & x \in D \backslash \{a\} \\ b , & x = a \end{cases}$$

stetig in a, also gilt:

$$\forall_{\varepsilon>0} \quad \exists_{\delta>0} \quad \forall_{\substack{x\in D\\|x-a|<\delta}} : \quad |f(x)-b| = |g(x)-g(a)| < \varepsilon\,.$$

Sei (ii) erfüllt und sei

$$g(x) = \begin{cases} f(x)\,, & x \in D\backslash\{a\} \\ b\,, & x = a\,. \end{cases}$$

Dann gilt nach Voraussetzung

$$\forall_{\varepsilon>0} \quad \exists_{\delta>0} \quad \forall_{\substack{x\in D\\|x-a|<\delta}} : \quad |g(x)-g(a)| < \varepsilon\,,$$

also ist g stetig in a. Somit gilt für jede Folge $\{x_n\}_{n=1}^\infty$ mit $x_n \in D$ und $\lim_{n\to\infty} x_n = a$, daß $\lim_{n\to\infty} f(x_n) = \lim_{n\to\infty} g(x_n) = g(a)$ ist, also hat f den Grenzwert $g(a) = b$ in a. ■

Satz: Seien $D \subseteq \mathbb{R}$, a Häufungspunkt von D und $f : D \to \mathbb{R}$. Genau dann hat f in a einen Grenzwert, wenn zu jedem $\varepsilon > 0$ ein $\delta > 0$ existiert so daß $|f(x)-f(y)| < \varepsilon$ gilt für alle $x, y \in D$ mit $|x-a| < \delta$, $|y-a| < \delta$.

Dies ist das Cauchykriterium für die Existenz eines Grenzwertes von f in a. Mit Quantoren kann das Cauchykriterium auch folgendermaßen geschrieben werden:

$$\forall_{\varepsilon>0} \quad \exists_{\delta>0} \quad \forall_{\substack{x,y\in D\\|x-a|<\delta\\|y-a|<\delta}} : \quad |f(x)-f(y)| < \varepsilon\,.$$

Beweis: Gilt $\lim_{x\to a} f(x) = b$, dann existiert zu $\varepsilon > 0$ ein $\delta > 0$ mit $|f(x)-b| < \frac{\varepsilon}{2}$ für alle $x \in D$ mit $|x-a| < \delta$. Also folgt

$$|f(x)-f(y)| = |f(x)-b+b-f(y)| \le |f(x)-b| + |f(y)-b| < \varepsilon$$

für alle $x, y \in D$ mit $|x-a| < \delta$, $|y-a| < \delta$.

Umgekehrt sei das Cauchykriterium erfüllt. Sei $\{x_n\}_{n=1}^\infty$ eine Folge mit $x_n \in D$ und $\lim_{n\to\infty} x_n = a$. Die Folge $\{f(x_n)\}_{n=1}^\infty$ ist dann eine Cauchy–Folge. Denn zu $\varepsilon > 0$ wähle $\delta > 0$ wie in der Bedingung. Es existiert $n_0 \in \mathbb{N}$ mit $|x_n - a| < \delta$ für alle $n \ge n_0$. Also folgt für alle $n, m \ge n_0$

$$|f(x_n) - f(x_m)| < \varepsilon\,,$$

also ist $\{f(x_n)\}_{n=1}^\infty$ Cauchy–Folge, also konvergiert jede derartige Folge $\{f(x_n)\}_{n=1}^\infty$, also hat f einen Grenzwert in a. ■

Sei $f : D \to \mathbb{R}$ stetig. Wenn f in jedem Häufungspunkt von D einen Grenzwert besitzt, dann gibt es eine eindeutig bestimmte stetige Fortsetzung g von f auf die abgeschlossene Hülle $\overline{D}$. Hinreichend dafür, daß f diese Eigenschaft hat, ist die gleichmäßige Stetigkeit von f.

Definition: Die Funktion $f : D \to \mathbb{R}$ heißt gleichmäßig stetig, wenn es zu jedem $\varepsilon > 0$ ein $\delta > 0$ gibt so daß

$$|f(x) - f(y)| < \varepsilon$$

ist für alle $x, y \in D$ mit $|x - y| < \delta$.

Mit Quantoren lautet diese Bedingung

$$\forall_{\varepsilon>0} \quad \exists_{\delta>0} \quad \forall_{\substack{x,y\in D \\ |x-y|<\delta}} : \quad |f(x) - f(y)| < \varepsilon .$$

Jede gleichmäßige stetige Funktion f ist stetig. Denn hält man $y = a \in D$ fest, dann folgt aus der Definition, daß f stetig ist in a. Bevor ich Beispiele betrachte, beweise ich folgenden Satz:

Satz: Sei $D \subseteq \mathbb{R}$ und $f : D \to \mathbb{R}$ gleichmäßig stetig. Dann besitzt f eine eindeutig bestimmte stetige Fortsetzung auf $\overline{D}$.

Beweis: Es genügt zu zeigen, daß f in jedem Punkt $a \in \overline{D}$ einen Grenzwert besitzt. Da f gleichmäßig stetig ist existiert zu $\varepsilon > 0$ ein $\delta > 0$ mit $|f(x) - f(y)| < \varepsilon$ für $x, y \in D$ mit $|x - y| < \delta$. Für $x, y \in D$ mit $|x - a| < \delta/2$, $|y - a| < \delta/2$ gilt $|x - y| = |x - a + a - y| \le |x - a| + |y - a| < \delta$, also $|f(x) - f(y)| < \varepsilon$, also besitzt f nach dem vorangehenden Satz einen Grenzwert in a. ∎

Beispiel: Sei $f : D = [-1, 0) \cup (0, 1] \to \mathbb{R}$, $x \mapsto f(x) = \frac{x^2}{|x|}$. Diese Funktion ist gleichmäßig stetig, und hat also eine stetige Fortsetzung auf $\overline{D} = [-1, 1]$. Diese Fortsetzung ist

$$x \mapsto g(x) = |x| .$$

Um zu sehen, daß f gleichmäßig stetig ist, sei $\varepsilon > 0$. Setze $\delta = \varepsilon$. Dann gilt für alle $x, y \in D$ mit $|x - y| < \delta$

$$|f(x) - f(y)| = \Big| |x| - |y| \Big| \le |x - y| < \varepsilon ,$$

nach der umgekehrten Dreiecksungleichung.

Nicht gleichmäßig stetig ist die Funktion $x \mapsto f(x) = \frac{1}{|x|}$ mit Definitionsbereich $D =$

$[-1,0)\cup(0,1]$. Ich zeige, daß für diese Funktion die Negation zur Aussage in der Definition der gleichmäßigen Stetigkeit richtig ist:

$$\exists_{\varepsilon>0} \quad \forall_{\delta>0} \quad \exists_{\substack{x,y\in D\\ |x-y|<\delta}} : \quad |f(x)-f(y)| \geq \varepsilon .$$

Für $\varepsilon = 1$ ist diese Aussage richtig. Denn sei $x > 0$ und $y = \frac{1}{2}x$. Dann folgt

$$|x-y| = \frac{1}{2}x \quad \text{und} \quad |f(x)-f(y)| = \Big| \frac{1}{|x|} - \frac{1}{|y|} \Big| = \frac{1}{x} ,$$

also ergibt sich: Sei $\delta > 0$. Wähle $x = \min(\delta, 1)$ und $y = \frac{1}{2}x = \min(\frac{1}{2}, \frac{1}{2}\delta)$. Dann folgt

$$|x-y| \leq \frac{1}{2}\delta < \delta , \quad |f(x)-f(y)| = \frac{1}{x} \geq 1 = \varepsilon ,$$

also sind zu beliebigem δ Zahlen x, y gefunden.

Satz: Seien $D \subseteq \mathbb{R}$ kompakt, $f : D \to \mathbb{R}$ stetig. Dann ist f gleichmäßig stetig.

Beweis: Angenommen, f sei nicht gleichmäßig stetig. Dann gibt es eine Zahl $\varepsilon > 0$, so daß für alle $n \in \mathbb{N}$ Zahlen $x_n, y_n \in D$ existieren mit $|x_n - y_n| < \frac{1}{n}$ und $|f(x_n) - f(y_n)| \geq \varepsilon$. Da D kompakt ist, besitzt $\{x_n\}_{n=1}^{\infty}$ eine konvergente Teilfolge $\{x_{n_j}\}_{j=1}^{\infty}$ mit Grenzwert $x_0 \in D$. Auch die Folge $\{y_{n_j}\}_{n=1}^{\infty}$ konvergiert gegen x_0 , denn zu $\eta > 0$ existiert $j_0 \in \mathbb{N}$ mit

$$|x_{n_j} - x_0| < \eta/2$$

für $j \geq j_0$, also

$$|x_0 - y_{n_j}| \leq |x_0 - x_{n_j}| + |x_{n_j} - y_{n_j}| < \eta/2 + \frac{1}{n_j} < \eta$$

für $j \geq \max(j_0, j_1)$, wobei j_1 die kleinste natürliche Zahl ist mit $j_1 \geq \frac{2}{\eta}$. Also folgt

$$\begin{aligned} 0 &= f(x_0) - f(x_0) \\ &= \lim_{j\to\infty} f(x_{n_j}) - \lim_{j\to\infty} f(y_{n_j}) \\ &= \lim_{j\to\infty} \Big(f(x_{n_j}) - f(y_{n_j}) \Big) \\ &= \lim_{j\to\infty} |f(x_{n_j}) - f(y_{n_j})| \geq \lim_{j\to\infty} \varepsilon = \varepsilon > 0 . \end{aligned}$$

Dies ist ein Widerspruch, also muß die Annahme falsch sein. ∎

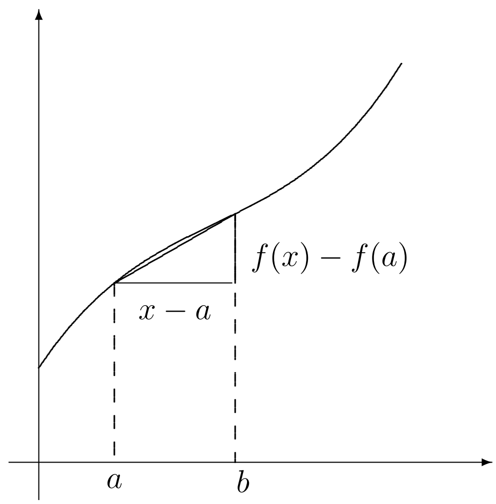
Beispiel: Die Funktion $Q : [0,1] \to \mathbb{R}$, $x \to Q(x) := \sqrt{x}$ ist gleichmäßig stetig.

Zur Übung zeige man, daß auch $Q : [0,\infty) \to \mathbb{R}$, $x \mapsto Q(x) := \sqrt{x}$ gleichmäßig stetig ist.

7 Differenzierbare Funktionen

7a.) Definition und Rechenregeln

Sei $J \subseteq \mathbb{R}$ ein Intervall, sei $a \in J$ und sei $f : J \to \mathbb{R}$. Die Funktion $\frac{f(x)-f(a)}{x-a}$ ist in ganz J mit Ausnahme des Punktes $x = a$ definiert. Man nennt diese Funktion den Differenzenquotienten. Diese Funktion gibt die Steigung der Sehne durch die Punkte $(a, f(a))$ und $(x, f(x))$ an.



Falls x gegen a strebt, geht die Sehne in die Tangente an den Graphen der Funktion f im Punkt $(a, f(a))$ über, falls diese Tangente existiert. Man kann die Tangente als diejenige affine Funktion auffassen, die die Funktion f in der Umgebung des Punktes a am besten approximiert. In der Differentialrechnung geht es also darum, gegebene komplizierte Funktionen durch möglichst einfache Funktionen zu approximieren.

Definition: Sei $D \subseteq \mathbb{R}$ und sei $a \in D$ ein Häufungspunkt von D. Sei $f : D \to \mathbb{R}$. Die Funktion f heißt differenzierbar im Punkt $a \in D$, wenn der Grenzwert

$$m = \lim_{\substack{x \to a \\ x \neq a}} \frac{f(x) - f(a)}{x - a}$$

existiert.

Der folgende Satz gibt eine hierzu äquivalente Bedingung:

Satz: f ist differenzierbar in a, genau dann wenn eine Zahl m und eine an der Stelle a stetige Funktion $r : D \to \mathbb{R}$ mit $r(a) = 0$ existiert so daß

$$f(x) = f(a) + m(x - a) + r(x)(x - a)$$

gilt für $x \in D$.

Beweis: Angenommen,

$$m = \lim_{\substack{x \to a \\ x \neq a}} \frac{f(x) - f(a)}{x - a}$$

existiere. Definiere die Funktion r durch

$$r(x) = \begin{cases} \dfrac{f(x) - f(a)}{x - a} - m\,, & x \neq a \\ 0\,, & x = a\,. \end{cases}$$

Dann gilt $f(x) = f(a) + m(x - a) + r(x)(x - a)$, und wegen

$$\lim_{\substack{x \to a \\ x \neq a}} r(x) = \lim_{\substack{x \to a \\ x \neq a}} \left[\frac{f(x) - f(a)}{x - a} - m\right] = 0 = r(a)$$

ist r stetig im Punkt a.

Falls umgekehrt m und r die im Satz angegebenen Eigenschaften haben, folgt aus der Stetigkeit von r, daß der Grenzwert des Differenzenquotienten existiert wegen

$$\lim_{\substack{x \to a \\ x \neq a}} \frac{f(x) - f(a)}{x - a} = m + \lim_{x \to a} r(x) = m\,.$$

■

Definition: Sei $D \subseteq \mathbb{R}$, $a \in D$ Häufungspunkt von D, und sei die Funktion $f : D \to \mathbb{R}$ differenzierbar in a. Dann heißt die eindeutig bestimmte Zahl

$$m = \lim_{\substack{x \to a \\ x \neq a}} \frac{f(x) - f(a)}{x - a}$$

die Ableitung von f im Punkt a. Man schreibt auch $f'(a) = m$.

Satz: Sei $f : D \to \mathbb{R}$ in $a \in D$ differenzierbar. Dann ist f in a stetig.

Beweis: Es gilt für alle $x \in D$

$$f(x) = f'(a)(x - a) + r(x)(x - a) + f(a)\,,$$

wobei r in a stetig ist, also ist f stetig in a. ■

Bevor ich Beispiele betrachte, beweise ich noch einige Rechenregeln für differenzierbare Funktionen:

Satz: Sei $a \in D \subseteq \mathbb{R}$ und sei $f : D \to \mathbb{R}$ in a differenzierbar mit $f'(a) > 0$. Dann gibt es eine Umgebung U von a, so daß $f(x) < f(a)$ gilt für alle $x \in U$ mit $x < a$ und $f(x) > f(a)$ für alle $x \in U$ mit $x > a$.

Beweis: Aus $f'(a) = \lim_{\substack{x \to a \\ x \neq a}} \frac{f(x)-f(a)}{x-a} > 0$ folgt für alle genügend nahe bei a liegenden $x \neq a$, daß $\frac{f(x)-f(a)}{x-a} > 0$, also $f(x) > f(a)$ für $x > a$ und $f(x) < f(a)$ für $x < a$. ■

Man beachte, daß aus $f'(a) \neq 0$ nicht gefolgert werden kann, daß f in einer Umgebung von a injektiv ist. Dies sieht man an Beispielen.

Satz: Sei $a \in D \subseteq \mathbb{R}$ und $f, g : D \to \mathbb{R}$ seien in a differenzierbar. Dann sind auch $f+g, f-g, cf$ und $f \cdot g$ in a differenzierbar, und es gilt

$$\begin{aligned}
(f \pm g)'(a) &= f'(a) \pm g'(a) \\
(fg)'(a) &= f'(a)g(a) + f(a)g'(a) \qquad \text{(Produktregel)} \\
(cf)'(a) &= cf'(a)\,.
\end{aligned}$$

Falls $g \neq 0$ ist in a, ist auch $\frac{f}{g}$ in a differenzierbar, und es gilt

$$\left(\frac{f}{g}\right)'(a) = \frac{g(a)f'(a) - f(a)g'(a)}{[g(a)]^2}\,, \qquad \text{(Quotientenregel)}$$

insbesondere also

$$\left(\frac{1}{g}\right)'(a) = -\frac{g'(a)}{[g(a)]^2}\,.$$

Beweis:

$$\begin{aligned}
&\lim_{\substack{x \to a \\ x \neq a}} \frac{f(x) \pm g(x) - f(a) \mp g(a)}{x-a} \\
&= \lim_{\substack{x \to a \\ x \neq a}} \frac{f(x)-f(a)}{x-a} \pm \lim_{\substack{x \to a \\ x \neq a}} \frac{g(x)-g(a)}{x-a} = f'(a) \pm g'a)\,.
\end{aligned}$$

Der Beweis für cf verläuft ebenso. Für das Produkt fg gilt

$$\begin{aligned}
&\lim_{\substack{x \to a \\ x \neq a}} \frac{f(x)g(x) - f(a)g(a)}{x-a} \\
&= \lim_{\substack{x \to a \\ x \neq a}} \left[\frac{f(x)-f(a)}{x-a}\, g(x) + \frac{g(x)-g(a)}{x-a}\, f(a)\right] \\
&= \lim_{\substack{x \to a \\ x \neq a}} \frac{f(x)-f(a)}{x-a} \lim_{x \to a} g(x) + \lim_{\substack{x \to a \\ x \neq a}} \frac{g(x)-g(a)}{x-a} f(a) \\
&= f'(a)g(a) + g'(a)f(a)\,,
\end{aligned}$$

weil g stetig ist im Punkt a.

Falls $g(a) \neq 0$ ist, ist auch $g(x) \neq 0$ für alle x aus einer Umgebung von a, da g in a stetig ist. Also folgt

$$\lim_{\substack{x\to a\\x\neq a}} \frac{\frac{1}{g(x)}-\frac{1}{g(a)}}{x-a} = \lim_{\substack{x\to a\\x\neq a}} \Big[-\frac{1}{g(x)g(a)}\,\frac{g(x)-g(a)}{x-a}\Big]$$
$$= \quad -\lim_{\substack{x\to a\\x\neq a}} \frac{1}{g(x)g(a)} \lim_{\substack{x\to a\\x\neq a}} \frac{g(x)-g(a)}{x-a} = -\frac{g'(a)}{g^2(a)}\,.$$

Die Formel für den allgemeinen Quotienten $(\frac{f}{g})'$ ergibt sich hieraus und aus der Produktregel. ■

Folgerung: Die Menge der Funktionen $f : D \to \mathbb{R}$, die in $a \in D$ differenzierbar sind, ist ein Vektorraum.

Satz: Seien $D_1, D_2 \subseteq \mathbb{R}$, $f : D_1 \to \mathbb{R}$, $g : D_2 \to \mathbb{R}$, und $g \circ f$ sei definiert. f sei in $a \in D_1$, g in $b = f(a) \in D_2$ differenzierbar. Dann ist $g \circ f$ in a differenzierbar, und es gilt

$$(g\circ f)'(a) = g'\Big(f(a)\Big)\cdot f'(a)\,. \qquad \text{(Kettenregel)}$$

Beweis: Nach Voraussetzung existiert eine in b stetige Funktion $r : D_2 \to \mathbb{R}$ mit $r(b) = 0$, so daß für $x \in D_2$ gilt

$$g(x) - g(b) = g'(b)(x-b) + r(x)(x-b)\,.$$

Also ergibt sich

$$\lim_{\substack{x\to a\\x\neq a}} \frac{g(f(x))-g(f(a))}{x-a} = \lim_{\substack{x\to a\\x\neq a}} \Big[g'\Big(f(a)\Big)\frac{f(x)-f(a)}{x-a} + r\Big(f(x)\Big)\frac{f(x)-f(a)}{x-a}\Big]$$
$$= \quad \Big[g'\Big(f(a)\Big) + \lim_{x\to a} r\Big(f(x)\Big)\Big] \lim_{\substack{x\to a\\x\neq a}} \frac{f(x)-f(a)}{x-a} = g'\Big(f(a)\Big)f'(a)\,,$$

weil $r \circ f$ im Punkt a stetig ist, und weil $r(f(a)) = r(b) = 0$ gilt. ■

Satz: Sei $f : D \to \mathbb{R}$ injektiv und differenzierbar in $a \in D$ mit $f'(a) \neq 0$. Wenn die Umkehrfunktion g von f an der Stelle $b = f(a)$ stetig ist, dann ist g sogar differenzierbar in b mit

$$g'(b) = \frac{1}{f'(a)} = \frac{1}{f'(g(b))}\,.$$

Beweis: Es ist

$$\lim_{\substack{y\to b\\y\neq b}} \frac{g(y)-g(b)}{y-b} = \lim_{\substack{y\to b\\y\neq b}} \frac{g(y)-a}{f(g(y))-f(a)} = \frac{1}{\lim_{\substack{y\to b\\y\neq b}} \frac{f(g(y))-f(a)}{g(y)-a}} = \frac{1}{f'(a)}\,,$$

da nach Voraussetzung die Funktion

$$h(x) = \begin{cases} \dfrac{f(x)-f(a)}{x-a}\,, & x \neq a \\[2ex] f'(a)\,, & x = a \end{cases}$$

stetig ist in a und somit $\lim_{x\to a} h(x) = h(a) = f'(a)$ gilt. Weil nach Voraussetzung auch g stetig ist in b, folgt daß $h \circ g$ stetig ist in $b = f(a)$ mit

$$\lim_{y\to b}(h\circ g)(y) = h\Big(\lim_{y\to b} g(y)\Big) = h\Big(g(b)\Big) = h(a) = f'(a)\,.$$

Wegen $f'(a) \neq 0$ ergibt sich hieraus, daß auch $\frac{1}{h\circ g}$ stetig ist in b. ■

Definition: Sei $f : D \to \mathbb{R}$ in jedem Punkt des Definitionsbereiches differenzierbar. Dann heißt die Funktion f differenzierbar. Für differenzierbare Funktionen $f : D \to \mathbb{R}$ wird durch

$$\Big(x \mapsto f'(x)\Big) : D \to \mathbb{R}$$

eine neue Funktion f' definiert. f' heißt Ableitung von f. Ist f' stetig, dann heißt f stetig differenzierbar.

Häufig wird auch die von Leibniz stammende Schreibweise

$$\frac{d}{dx} f(x) := f'(x)$$

benutzt. Dies ist insbesondere dann zweckmäßig, wenn f von mehreren Variablen abhängt, weil dann klargestellt ist, nach welcher Variablen abgeleitet wird.

Aus den bewiesenen Sätzen folgt:

Satz: Die Menge aller in D differenzierbaren Funktionen ist ein Vektorraum über $\mathbb{R}$. Auch die Menge aller in D stetig differenzierbaren Funktionen ist ein Vektorraum über $\mathbb{R}$, der mit $C_1(D,\mathbb{R}) = C_1(D)$ bezeichnet wird.

Auch das Produkt zweier differenzierbarer (bzw. stetig differenzierbarer) Funktionen ist (stetig) differenzierbar. Wenn f' differenzierbar ist, bezeichnet man die Ableitung von f' mit f'' oder mit $f^{(2)}$, und nennt diese Ableitung die zweite Ableitung von f. Allgemein definiert man die k–te Ableitung $f^{(k)}$ von f rekursiv durch

$$f^{(k+1)} := (f^{(k)})'\,.$$

Falls f k–mal differenzierbar ist, dann sind die Ableitungen $f^{(0)} = f, f^{(1)}, \ldots, f^{(k-1)}$ stetig. Falls auch noch $f^{(k)}$ stetig ist, sagt man, f sei k–mal stetig differenzierbar. Die Menge der k–mal stetig differenzierbaren Funktionen $f : D \to \mathbb{R}$ bildet einen Vektorraum über $\mathbb{R}$, den man mit $C_k(D, \mathbb{R})$ oder $C_k(D)$ bezeichnet. $C_0(D) = C(D)$ sind dann die stetigen Funktionen. Existiert $f^{(k)}$ für alle k, dann sagt man, f sei unendlich oft differenzierbar, oder f sei beliebig oft differenzierbar. Den Vektorraum dieser Funktionen bezeichnet man mit $C_\infty(D, \mathbb{R})$ oder mit $C_\infty(D)$. Man benutzt die Bezeichnung:

$$\frac{d^k}{dx^k} f(x) = f^{(k)}(x) .$$

7b.) Beispiele und Anwendungen

1. Die konstanten Funktionen $f : \mathbb{R} \to \mathbb{R}, \; x \mapsto f(x) := c$ sind in allen Punkten $a \in \mathbb{R}$ differenzierbar mit Ableitung $f'(a) = 0$.

2. Die Identität $x \mapsto f(x) := x$ ist in allen $a \in \mathbb{R}$ differenzierbar mit Ableitung

$$f'(a) = \lim_{\substack{x \to a \\ x \neq a}} \frac{x-a}{x-a} = 1 .$$

3. Hieraus folgt, daß alle Polynome $p(x) = \sum_{n=0}^m a_n x^n$ differenzierbar sind, und daß alle rationalen Funktionen in ihrem Definitionsbereich differenzierbar sind. Um die Ableitung des Polynoms p im Punkt $x \in \mathbb{R}$ zu berechnen, genügt es, die Ableitung der Potenzen x^n für ganze Zahlen $n \geq 0$ zu kennen, weil

$$\Big(\sum_{n=0}^m a_n x^n \Big)' = \sum_{n=0}^m a_n (x^n)'$$

gilt.
Durch vollständige Induktion ergibt sich

$$(x^n)' = \begin{cases} 0 , & n = 0 \\ n x^{n-1} , & n \geq 1 . \end{cases}$$

Denn für $n = 0$ und $n = 1$ ist diese Gleichung richtig. Angenommen, die Gleichung sei für eine Zahl $n \in \mathbb{N}$ richtig. Dann folgt nach der Produktregel

$$\begin{aligned} (x^{n+1})' &= (x\, x^n)' = x' x^n + x(x^n)' \\ &= x^n + x(n\, x^{n-1}) = (n+1) x^n . \end{aligned}$$

Also ergibt sich: Die Ableitung eines Polynoms ist wieder ein Polynom. Bei der Differentiation erniedrigt sich der Grad eines Polynoms um 1. Nach der Quotientenregel ist also auch jede rationale Funktion differenzierbar in jedem Punkt ihres Definitionsbereiches. Insbesondere ergibt sich für $n \in \mathbb{N}$

$$(x^{-n})' = \left(\frac{1}{x^n}\right)' = -\frac{1}{x^{2n}}\, n\, x^{n-1} = -n\, x^{-n-1}\,.$$

4. Sei $f : \mathbb{R} \to \mathbb{R}\,,$

$$x \mapsto f(x) := \exp(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!} \quad (=: e^x)$$

die Exponentialfunktion. Es gilt

$$\begin{aligned}
\lim_{\substack{x\to 0\\ x\neq 0}} \frac{e^x - e^0}{x-0} &= \lim_{\substack{x\to 0\\ x\neq 0}} \frac{e^x-1}{x} = \lim_{\substack{x\to 0\\ x\neq 0}} \left(1 + \sum_{k=2}^{\infty} \frac{x^{k-1}}{k!}\right)\\
&= 1 + \lim_{\substack{x\to 0\\ x\neq 0}} \sum_{k=2}^{\infty} \frac{x^{k-1}}{k!} = 1\,,
\end{aligned}$$

wegen

$$\begin{aligned}
&\Big|\lim_{x\to 0} \sum_{k=2}^{\infty} \frac{x^{k-1}}{k!}\Big| \leq \lim_{x\to 0} \sum_{k=2}^{\infty} \frac{|x|^{k-1}}{k!}\\
&= \lim_{x\to 0} |x| \sum_{k=2}^{\infty} \frac{|x|^{k-2}}{k!} \leq \lim_{x\to 0} |x| \sum_{k=2}^{\infty} \frac{|x|^{k-2}}{(k-2)!}\\
&= \lim_{x\to 0} |x| e^{|x|} \leq \lim_{x\to 0} |x| e^1 = 0\,,
\end{aligned}$$

also

$$\exp'(0) = 1 \quad \Big(= (e^x)'\big|_{x=0}\Big)\,.$$

Aus dem Additionstheorem $e^{x+y} = e^x e^y$ folgt

$$\lim_{\substack{h\to 0\\ h\neq 0}} \frac{e^{x+h} - e^x}{h} = \lim_{\substack{h\to 0\\ h\neq 0}} e^x \frac{e^h - 1}{h} = e^x\,,$$

also

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

oder

$$(e^x)' = e^x\,.$$

Also ist $\exp : \mathbb{R} \to \mathbb{R}$ differenzierbar in allen Punkten $x \in \mathbb{R}$, folglich ist die Exponentialfunktion eine stetige Funktion.

5. **Definition des natürlichen Logarithmus:** Aus der Beziehung $e^x e^{-x} = 1$ folgt $e^x \neq 0$ für alle $x \in \mathbb{R}$. Aus dem Zwischenwertsatz folgt damit, daß entweder $e^x > 0$ gilt

für alle x oder $e^x < 0$ für alle x. Wegen $e^0 = 1$ ergibt sich somit $e^x > 0$ für alle x.

Für $x > 0$ folgt

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!} > 1 + x > 1\,,$$

also ergibt sich für $x_2 > x_1$

$$e^{x_2} = e^{x_2 - x_1 + x_1} = e^{x_2 - x_1} e^{x_1} > e_1^x\,,$$

also ist e^x auf ganz $\mathbb{R}$ streng monoton wachsend, folglich hat e^x eine Umkehrfunktion, die mit log bezeichnet wird.

Der Wertebereich von exp ist ganz $\mathbb{R}^+ = (0, \infty)$. Denn wegen $e^x > x+1$ für $x > 0$ nimmt e^x beliebig große Werte an, und somit folgt wegen $e^0 = 1$ aus dem Zwischenwertsatz, daß $[1, \infty) \subseteq W(\exp)$ ist. Mit $e^{-x} = \frac{1}{e^x}$ resultiert hieraus auch $(0, 1) \subseteq W(\exp)$, also $W(\exp) = \mathbb{R}^+$. Folglich ist $\log : \mathbb{R}^+ \to \mathbb{R}$ bijektiv.

Da $\exp : \mathbb{R} \to \mathbb{R}^+$ und da $\mathbb{R}$ offen ist, ist die Umkehrfunktion log von exp stetig. Also ist log sogar differenzierbar wegen $\exp'(x) = \exp(x) \neq 0$ mit

$$(\log x)' := \log' x = \frac{1}{\exp'(\log x)} = \frac{1}{\exp(\log x)} = \frac{1}{x}\,.$$

6. Die Funktion $(x \mapsto \log(-x)) : \mathbb{R}^- \to \mathbb{R}$ hat nach der Kettenregel die Ableitung

$$\Big(\log(-x)\Big)' = \log'(-x)(-x)' = \frac{-1}{-x} = \frac{1}{x}\,.$$

7. **Definition der allgemeinen Potenz:** Mit Hilfe von exp und log kann die allgemeine Potenz definiert werden: Bei der Definition der Exponentialfunktion wurde gezeigt, daß für alle $x \in \mathbb{Q}$

$$e^x = \exp(x)$$

gilt mit $\exp(x) := \sum_{k=0}^{\infty} \frac{x^k}{k!}$ und mit $e := \exp(1)$. Mit denselben Schlüssen zeigt man, daß für rationales x und beliebiges $y \in \mathbb{R}$

$$\Big(\exp(y)\Big)^x = \exp(xy)$$

gilt. Also definiert man für $x, y \in \mathbb{R}$

$$\Big(\exp(y)\Big)^x := \exp(xy)\,.$$

Für $y = 1$ erhält man hieraus wieder die bereits früher angegebene Definition $e^x := \exp(x)$, $x \in \mathbb{R}$. Sei nun $a > 0$. Dann gilt $a = \exp(\log a) = e^{\log a}$, also folgt aus dieser Definition für beliebiges $x \in \mathbb{R}$

$$a^x = \Big[\exp(\log a)\Big]^x = \exp(x \log a) = e^{x \log a}.$$

Also ist a^x erklärt für alle $x \in \mathbb{R}$ und alle $a > 0$.
Aus diesen Definitionen folgt für die Funktionen $\log x$ und a^x

(i) $\quad \log(xy) = \log(e^{\log x} e^{\log y}) = \log e^{\log x + \log y} = \log x + \log y$

(ii) $\quad \log a^x = \log(e^{x \log a}) = x \log a$

(iii) $\quad a^x a^y = e^{x \log a} e^{y \log a} = a^{x+y}$

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

(v) $\quad a^x b^x = e^{x \log a} e^{x \log b} = e^{x \log(ab)} = (ab)^x$.

Die Funktion $(x \mapsto a^x) : \mathbb{R} \to \mathbb{R}^+$ ist differenzierbar, und es gilt wegen der Kettenregel

$$\begin{aligned} \frac{d}{dx}(a^x) &= (e^{x \log a})' = \exp'(x \log a)(\log a) \\ &= (\log a) e^{x \log a} = (\log a) a^x . \end{aligned}$$

Für $c \in \mathbb{Z}$ wurde gezeigt, daß $a \mapsto a^c$ differenzierbar ist, mit $(a^c)' = c\, a^{c-1}$. Die Funktion $(a \mapsto a^c) : \mathbb{R}^+ \to \mathbb{R}^+$ ist auch für beliebiges $c \in \mathbb{R}$ differenzierbar, und es gilt

$$\begin{aligned} \frac{d}{da}(a^c) &= \frac{d}{da}(e^{c \log a}) = \frac{d}{da}\Big(\exp(c \log a)\Big) \\ &= \exp'(c \log a) \frac{d}{da}(c \log a) = a^c \frac{c}{a} = c a^{c-1} . \end{aligned}$$

8. Hieraus folgt: Die Quatratwurzel

$$\Big(x \mapsto Q(x) := \sqrt{x}\Big) : \mathbb{R}_0^+ \to \mathbb{R}_0^+$$

ist für alle $x \in \mathbb{R}^+$ differenzierbar mit

$$Q'(x) = (x^{1/2})' = \frac{1}{2} x^{-1/2} = \frac{1}{2\sqrt{x}} .$$

Im Punkt $x = 0$ ist Q dagegen nicht differenzierbar, denn es gilt für $x > 0$

$$\frac{Q(x) - Q(0)}{x - 0} = \frac{1}{\sqrt{x}} ,$$

und diese Funktion ist unbeschränkt in jeder Umgebung des Nullpunktes, also kann sie keinen Grenzwert an der Stelle 0 haben.

9. Es gilt

$$\lim_{\substack{x\to 0\\ x\neq 0}} \frac{\log(1+x)}{x} = \frac{d}{dx} \log(1+x)\Big|_{x=0} = 1\,.$$

Seien $a \in \mathbb{R}$ mit $a \neq 0$ und $n \in \mathbb{N}$. Wegen

$$\frac{\log(1+\frac{a}{n})}{\frac{a}{n}} = \frac{1}{a} \log\left[\left(1+\frac{a}{n}\right)^n\right]$$

folgt somit

$$\lim_{n\to\infty} \log\left[\left(1+\frac{a}{n}\right)^n\right] = a\,,$$

also

$$\lim_{n\to\infty}\left(1+\frac{a}{n}\right)^n = \lim_{n\to\infty} e^{\log[(1+\frac{a}{n})^n]} = e^a\,,$$

wegen der Stetigkeit der Exponentialfunktion.

10. Die Funktion $(x \mapsto |x|) : \mathbb{R} \to \mathbb{R}_0^+$ ist überall differenzierbar mit Ausnahme des Nullpunktes. Es gilt

$$\frac{d}{dx}|x| = \begin{cases} -1\,, & x<0 \\ \;\;\,1\,, & x>0\,. \end{cases}$$

Im Nullpunkt existieren jedoch der „linksseitige“ und der „rechtsseitige“ Grenzwert

$$\lim_{\substack{x\to 0\\ x<0}} \frac{|x|}{x} = -1 \quad \text{und} \quad \lim_{\substack{x\to 0\\ x>0}} \frac{|x|}{x} = 1\,.$$

Man sagt, $|x|$ sei im Nullpunkt „linksseitig differenzierbar“ und „rechtsseitig differenzierbar“.

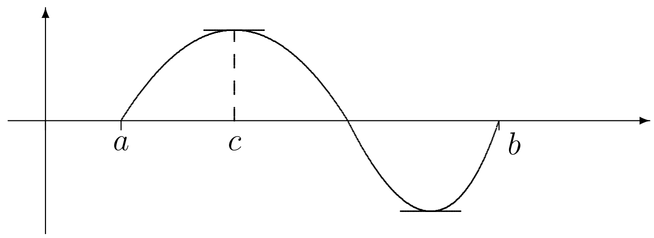
Bemerkung: Jede differenzierbare Funktion ist stetig. Es gibt aber stetige Funktionen, die nirgends differenzierbar sind. Ein Beispiel findet man im Buch von Barner–Flohr, S.261 ff.

7c.) Mittelwertsatz

Satz (von Rolle): Sei $-\infty < a < b < \infty$. Ist $f : [a,b] \to \mathbb{R}$ stetig mit $f(a) = f(b) = 0$, und ist f in (a,b) differenzierbar, dann gibt es ein $c \in (a,b)$ mit

$$f'(c) = 0\,.$$

Beweis: Falls $f \equiv 0$ ist in $[a,b]$, ist jedes $c \in (a,b)$ geeignet. Sei also $f \not\equiv 0$. O.B.d.A. existiere $x \in (a,b)$ mit $f(x) > 0$. (Sonst gehe man zu $-f$ über.)



Damit nimmt f das Maximum in einem Punkt $c \in (a,b)$ an. Es gilt $f'(c) = 0$. Denn für alle $x \in [a,b]$ muß $f(x) \leq f(c)$ sein, folglich muß

$$\frac{f(x)-f(c)}{x-c} \begin{cases} \leq 0\,, & \text{für} \quad x > c \\ \geq 0\,, & \text{für} \quad x < c \end{cases}$$

sein, also

$$f'(c) = \lim_{\substack{x \to c \\ x \neq 0}} \frac{f(x)-f(c)}{x-c} = 0\,.$$

■

Aus diesem Satz erhält man den „ersten Mittelwertsatz der Differentialrechnung“:

Erster Mittelwertsatz der Differentialrechnung: Sei $f : [a,b] \to \mathbb{R}$ stetig und sei f in (a,b) differenzierbar. Dann gibt es ein $c \in (a,b)$ mit

$$f(b) - f(a) = f'(c)(b-a)\,.$$

Beweis: Sei $g : [a,b] \to \mathbb{R}$ definiert durch

$$g(x) = f(x) - \frac{f(b)-f(a)}{b-a}\,(x-a) - f(a)\,.$$

Diese Funktion erfüllt die Voraussetzungen des Satzes von Rolle. Es gilt

$$g(a) = g(b) = 0\,.$$

Also gibt es ein $c \in (a,b)$ mit $g'(c) = 0$. Aus

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

folgt die Behauptung des Satzes. ■

Es ist unmittelbar klar, daß man den Mittelwertsatz auch auf folgende Weise formulieren kann: Seien $x, x+h \in [a,b]$. Dann gibt es $\theta, 0 < \theta < 1$, mit

$$f(x+h) = f(x) + f'(x+\theta h)h\,.$$

(Hierbei darf h auch kleiner als Null sein.)

Einfache Anwendungen des Mittelwertsatzes

Falls $f : D \to \mathbb{R}$ konstant ist, gilt $f'(x) = 0$ für alle $x \in D$. Die Umkehrung ist für beliebiges D nicht richtig, aber es gilt:

Satz: Sei J ein Intervall und sei $f : J \to \mathbb{R}$ mit $f'(x) = 0$ für alle $x \in J$. Dann ist f konstant.

Beweis: Sei $a \in J$. Für jedes $x \in J$ gilt $[a, x] \subseteq J$, also gibt es nach dem Mittelwertsatz ein $c \in (a, x) \subseteq J$ mit

$$f(x) = f(a) + f'(c)(x - a) = f(a)\,,$$

also ist $f \equiv f(a)$. ■

Definition: Seien $f : D \to \mathbb{R}$ und $F : D \to \mathbb{R}$ mit $F' = f$. Dann heißt F Stammfunktion zu f.

Satz: Sei $f : J \to \mathbb{R}$, wobei J ein Intervall sei, und seien F, G Stammfunktionen zu f. Dann gilt $F - G = \text{const}$.

Beweis: $(F - G)' = F' - G' = f - f = 0$, also ist $F - G = \text{const}$ im Intervall. ■

Anwendung auf Differentialgleichungen: Sei $f : \mathbb{R} \to \mathbb{R}$ eine differenzierbare Funktion, die die „Differentialgleichung“ $f' = f$ erfüllt. Dann ist $f(x) = c\,e^x$ mit beliebiger Konstanten $c \in \mathbb{R}$.

Beweis: Die Funktion $g(x) = f(x)e^{-x}$ ist differenzierbar in $\mathbb{R}$, und es gilt

$$\begin{aligned} \Big(f(x)e^{-x}\Big)' &= f'(x)e^{-x} - f(x)e^{-x} \\ &= f(x)e^{-x} - f(x)e^{-x} = 0\,, \end{aligned}$$

also ist $f(x)e^{-x} = c$, also $f(x) = c\,e^x$.
(Dies ist ein „Eindeutigkeitssatz“). ■

Der zweite Mittelwertsatz der Differentialrechnung enthält den ersten als Spezialfall. Er wird später zur Ermittlung von Grenzwerten benutzt:

Zweiter Mittelwertsatz der Differentialrechnung: Seien $f, g : [a, b] \to \mathbb{R}$ stetige Funktionen, und seien diese Funktionen in (a, b) differenzierbar. Dann gibt es $c \in (a, b)$ mit

$$g'(c)\Big(f(b) - f(a)\Big) = f'(c)\Big(g(b) - g(a)\Big)\,.$$

Beweis: Sei

$$h(x) = \Big(f(b) - f(a)\Big) \cdot \Big(g(x) - g(a)\Big) - \Big(g(b) - g(a)\Big)\Big(f(x) - f(a)\Big).$$

Dann gilt $h(a) = h(b) = 0$. Also existiert nach dem Satz von Rolle ein $c \in (a, b)$ mit $h'(c) = 0$, also

$$0 = h'(c) = \Big(f(b) - f(a)\Big)g'(c) - \Big(g(b) - g(a)\Big)f'(c).$$

Hieraus folgt die Behauptung. ■

Der erste Mittelwertsatz folgt hieraus mit $g(x) = x$. Falls $g'(x) \neq 0$ ist für alle $x \in (a, b)$, dann folgt aus dem ersten Mittelwertsatz $g(b) \neq g(a)$, also kann die Formel im zweiten Mittelwertsatz auch in der Form

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

geschrieben werden.

7d.) Taylorsche Formel

Sei

$$f(x) = \sum_{k=0}^{n} c_k x^k$$

ein Polynom. Weil $|x|^k$ in einer Umgebung des Nullpunktes $x = 0$ klein ist für große Werte von k, kann man das Verhalten von f in einer Umgebung des Nullpunktes gut an den Koeffizienten c_k ablesen. Diese Koeffizienten können auf einfache Weise aus den Ableitungen von f im Nullpunkt berechnet werden. Denn es gilt $f'(x) = \sum_{k=1}^{n} k c_k x^{k-1}$, also

$$\begin{aligned}
f(0) &= c_0 \\
f'(0) &= c_1 \\
f^{(2)}(0) &= 2c_2 \\
&\vdots \\
f^{(n)}(0) &= n!\ c_n, \qquad n! = 1 \cdot 2 \cdot \ldots \cdot n \\
f^{(k)}(0) &= 0 \quad \text{für} \quad k > n.
\end{aligned}$$

Um das Verhalten von f in der Umgebung eines beliebigen Punktes $a \in \mathbb{R}$ zu untersuchen, ist es zweckmäßig, f in der Form

4

$$f(x) = \sum_{k=0}^{n} b_k (x-a)^k \tag{$*$}$$

darzustellen, mit geeignet zu bestimmenden Koeffizienten b_k. Ich werde nachher zeigen, daß es solche Koeffizienten b_k gibt. Zunächst nehme ich an, dies sei möglich. Es ist dann einfach, die Koeffizienten b_k aus den höheren Ableitungen von f zu bestimmen, denn wie oben folgt aus

$$f'(x) = \sum_{k=1}^{n} k\, b_k (x-a)^{k-1}$$

daß

$$b_k = \frac{f^{(k)}(a)}{k!}\,, \quad k = 0, 1, \ldots, n\,,$$

sein muß. Falls für ein Polynom f höchstens n–ten Grades die Darstellung $(*)$ möglich ist, muß also gelten

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

Falls f kein solches Polynom ist, ist eine Darstellung der Form $(*)$ nicht möglich. Dann ist diese Gleichung nicht mehr richtig. Man kann aber hoffen, daß der rechtsstehende Ausdruck noch eine gute Approximation für f liefert in einer Umgebung des Punktes $x = a$, falls f genügend oft differenzierbar ist. Den Fehler, den man macht, wenn man f durch den rechtsstehenden Ausdruck ersetzt, kann man mit der Formel von Taylor abschätzen:

Satz (Taylorformel) (Brook Taylor, 1685 – 1731): Sei $f : [a,b] \to \mathbb{R}$ n–mal stetig differenzierbar, und sei f im offenen Intervall (a,b) $(n+1)$–mal differenzierbar. Dann existiert $c \in (a,b)$ mit

$$f(b) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(b-a)^k + \frac{f^{(n+1)}(c)}{(n+1)!}(b-a)^{n+1}\,.$$

Beweis: Sei

$$g(x) = f(b) - \sum_{k=0}^{n} \frac{f^{(k)}(x)}{k!}(b-x)^k - \frac{m}{(n+1)!}(b-x)^{n+1}$$

wobei die Zahl $m \in \mathbb{R}$ so bestimmt wird, daß $g(a) = 0$ gilt, also

$$m = \frac{(n+1)!}{(b-a)^{n+1}} \Big[f(b) - \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(b-a)^k \Big]. \tag{$*$}$$

Es gilt auch $g(b) = 0$ und

$$\begin{aligned} g'(x) &= -\sum_{k=0}^{n} \frac{f^{(k+1)}(x)}{k!}\,(b-x)^k \\ &\quad + \sum_{k=1}^{n} \frac{f^{(k)}(x)}{k!}\,k(b-x)^{k-1} + \frac{m}{n!}\,(b-x)^n \\ &= -\frac{f^{(n+1)}(x)}{n!}\,(b-x)^n + \frac{m}{n!}\,(b-x)^n\,. \end{aligned}$$

Aus dem Satz von Rolle folgt nun, daß ein $c \in (a,b)$ existiert mit $g'(c) = 0$, also

$$0 = [m - f^{(n+1)}(c)]\,\frac{(b-c)^n}{n!}\,,$$

folglich ergibt sich $m = f^{(n+1)}(c)$. Einsetzen in $(*)$ und Auflösen nach $f(b)$ liefert die Taylorformel. ■

Folgerung: Sei f ein Polynom höchstens n–ten Grades. Dann gilt für jedes $a \in \mathbb{R}$

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

Beweis: Es ist $f^{(n+1)}(x) \equiv 0$, also folgt die Behauptung aus der Taylorformel. ■

Folgerung: Sei $f : [a,b] \to \mathbb{R}$ eine $(n+1)$–mal differenzierbare Funktion mit $f^{(n+1)}(x) = 0$ für alle $x \in [a,b]$. Dann ist f ein Polynom höchstens n–ten Grades.

Beweis: Aus der Taylorformel folgt $f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}\,(x-a)^k$ für alle $x \in [a,b]$, und dies ist ein Polynom höchstens n–ten Grades. ■

Man kann die Taylorformel auch in folgender Form schreiben: Seien $x, x+h \in [a,b]$. Dann existiert $\theta, 0 < \theta < 1$, mit

$$f(x+h) = \sum_{k=0}^{n} \frac{f^{(k)}(x)}{k!}\,h^k + \frac{f^{(n+1)}(x+\theta h)}{(n+1)!}\,h^{n+1}\,.$$

Für $n = 0$ geht die Taylorformel in den Mittelwertsatz über. Die Funktion

$$x \mapsto \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}\,(x-a)^k$$

heißt Taylorpolynom n–ter Ordnung zur Funktion f an der Stelle a. Aus der Größe des Restgliedes

$$R_n = f(b) - \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}\,(b-a)^k = \frac{f^{(n+1)}(c)}{(n+1)!}\,(b-a)^{n+1}$$

ergibt sich, wie gut das Taylorpolynom die Funktion f approximiert.
$R_n = \frac{f^{(n+1)}(c)}{(n+1)!}\,(b-a)^{n+1}$ heißt Lagrangesche Darstellung des Restgliedes. Es gibt andere Möglichkeiten, das Restglied in der Taylorformel darzustellen.

Man kann daher die Taylorformel zur näherungsweisen Berechnung von Funktionen benutzen.

Beispiele:

1.) Berechnung von $\log x$ in einer Umgebung des Punktes $x = 1$:

Es gilt
$$\frac{d}{dx}\,\log x = \frac{1}{x}\,,$$
und
$$\frac{d^n}{dx^n}\,\log x = (n-1)!(-1)^{n-1}\,\frac{1}{x^n}\,,$$
also folgt aus der Taylorformel
$$\begin{aligned} \log(1+h) &= \sum_{k=0}^{n} \frac{\log^{(k)}(1)}{k!}\,h^k + \frac{\log^{(n+1)}(1+\theta h)}{(n+1)!}\,h^{n+1} \\ &= \sum_{k=1}^{n} \frac{(-1)^{k-1}}{k}\,h^k + \frac{(-1)^n}{n+1}\,\frac{h^{n+1}}{(1+\theta h)^{n+1}}\,. \end{aligned} \tag{$*$}$$
Das Restglied
$$R_n = (-1)^n\,\frac{1}{n+1}\,\frac{h^{n+1}}{(1+\theta h)^{n+1}}$$
kann für $0 \le h \le 1$ abgeschätzt werden durch
$$|R_n| = \frac{1}{n+1}\left|\frac{h}{1+\theta h}\right|^{n+1} \le \frac{h^{n+1}}{n+1}\,,$$
und für $-1 < h \le 0$ durch
$$|R_n| = \frac{1}{n+1}\,\frac{|h|^{n+1}}{(1-|h|)^{n+1}}\,.$$
Für $-\frac{1}{2} \le h \le 1$ folgt somit
$$|R_n| \le \frac{1}{n+1}\,,$$
also $\lim_{n\to\infty} R_n = 0$. Aus $(*)$ ergibt sich
$$\begin{aligned} \log(1+h) &= \lim_{n\to\infty}\Big(\log(1+h) - R_n\Big) \\ &= \lim_{n\to\infty}\sum_{k=1}^{n}(-1)^{k-1}\,\frac{h^k}{k} = \sum_{k=1}^{\infty}(-1)^{k-1}\,\frac{h^k}{k}\,. \end{aligned}$$

Zum Beispiel ergibt sich für $h = 1$

$$\log 2 = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \dots .$$

Diese Reihe konvergiert jedoch sehr schlecht und ist daher nicht für numerische Berechnungen geeignet. (Es gilt ja nur $|R_n| < \frac{1}{n+1}$). Bessere Ergebnisse erhält man für Werte von h, die näher bei Null liegen. Zum Beispiel ergibt sich für $h = \frac{1}{2}$ und $n = 5$

$$\sum_{k=1}^{5} (-1)^{k-1} \frac{1}{k2^k} = \frac{1}{2} - \frac{1}{8} + \frac{1}{24} - \frac{1}{64} + \frac{1}{160} \approx 0,40729$$

und

$$|R_5| \leq \frac{1}{6 \cdot 2^6} = \frac{1}{384} < 0,0027$$

(Genauer gilt $\quad \log 1,5 \approx 0,405465108$)

2.) Für $s \in \mathbb{R}$ betrachte die Funktion

$$(x \mapsto x^s) : \mathbb{R}^+ \to \mathbb{R} .$$

Sei $a > 0$ und sei $h \in \mathbb{R}$ mit $|h| < a$. Dann folgt aus der Taylorformel mit

$$\frac{d^k}{dx^k}(x^s) = s(s-1)(s-2)\dots(s-k+1)x^{s-k} ,$$

$$(a+h)^s = \sum_{k=0}^{n} \binom{s}{k} a^{s-k} h^k + \binom{s}{n+1} (a+\theta h)^{s-n-1} h^{n+1}$$

mit geeignetem $0 < \theta < 1$. Hierbei sei

$$\binom{s}{k} = \frac{s(s-1)\dots(s-k+1)}{k!} .$$

Als Übungsaufgabe zeige man: Für $0 \leq h < a$ gilt

$$\lim_{n\to\infty} R_n = \lim_{n\to\infty} \binom{s}{n+1} \frac{h^{n+1}}{(a+\theta h)^{n+1-s}} = 0 .$$

Hieraus ergibt sich für $0 \leq h < a$

$$(a+h)^s = \sum_{k=0}^{\infty} \binom{s}{k} a^{s-k} h^k .$$

Man vergleiche dieses Ergebnis mit der Binomischen Formel.

In beiden betrachteten Beispielen galt $\lim_{n\to\infty} R_n = 0$. Hieraus folgte, daß die beiden beliebig oft differenzierbaren Funktionen $\log x$ und x^s in gewissen Bereichen in Reihen entwickelt werden konnten (Taylorreihe). Es ist jedoch im allgemeinen nicht richtig, daß das Restglied in der Taylorformel einer beliebig oft differenzierbaren Funktion gegen Null konvergiert. In jedem Spezialfall muß untersucht werden, ob das Restglied gegen Null konvergiert oder nicht. Funktionen, die in eine derartige Potenzreihe entwickelt werden können, nennt man analytische Funktionen. Die „Funktionentheorie“ befaßt sich ausschließlich mit derartigen Funktionen. Ich werde später darauf zurückkommen.

7e.) Monotonie, Extrema, Grenzwerte

Satz: Sei J ein Intervall und $f : J \to \mathbb{R}$ differenzierbar. f ist genau dann monoton wachsend, wenn

$$f'(x) \geq 0$$

ist für alle $x \in J$.

Beweis: Ist f monoton wachsend, dann gilt für $x \neq y$

$$\frac{f(y) - f(x)}{y - x} \geq 0\,,$$

also

$$f'(x) = \lim_{\substack{y \to x \\ y \neq x}} \frac{f(y) - f(x)}{y - x} \geq 0\,.$$

Ist umgekehrt $f'(x) \geq 0$ für alle x, dann folgt aus dem Mittelwertsatz für $y > x$

$$f(y) - f(x) = f'(z)(y - x) \geq 0\,.$$

■

Bemerkung: Ist $f'(x) > 0$ für alle $x \in J$, dann ist f streng monoton wachsend. Dies ist jedoch nur eine hinreichende Bedingung. Zum Beispiel ist $x \mapsto x^3$ streng monoton wachsend, aber $(x^3)'\big|_{x=0} = 3x^2\big|_{x=0} = 0$.

Definition: Sei $D \subseteq \mathbb{R}$, $a \in D$, und sei $f : D \to \mathbb{R}$. Wenn es eine Umgebung U von a gibt, so daß für alle $x \in D \cap U$ gilt

$$f(x) \geq f(a)\,,$$

dann heißt a relative Minimalstelle von f. Gilt für alle $x \in D \cap U$

$$f(x) \leq f(a)\,,$$

dann heißt a relative Maximalstelle von f. Der Punkt a heißt relative Extremstelle von f, wenn a relative Minimal– oder Maximalstelle ist.

Satz: Sei a innerer Punkt von D, sei $f : D \to \mathbb{R}$ differenzierbar, und sei a relative Extremalstelle von f. Dann gilt

$$f'(a) = 0 \,.$$

Beweis: Siehe Beweis zum Satz von Rolle. ■

$f'(a) = 0$ ist nur eine notwendige Bedingung für eine Extremstelle. Zum Beispiel gilt für $f(x) = x^3$

$$f'(0) = 0 \,,$$

aber $x = 0$ ist keine relative Extremalstelle. Eine hinreichende Bedingung gibt folgender Satz:

Satz: Sei $D \subseteq \mathbb{R}$, und sei a innerer Punkt von D. $f : D \to \mathbb{R}$ sei n–mal stetig differenzierbar, und es sei

$$f'(a) = f''(a) = \ldots = f^{(n-1)}(a) = 0 \,,$$

aber $f^{(n)}(a) \neq 0$.
Ist n ungerade, dann ist a keine Extremalstelle. Ist n gerade, dann ist a Minimalstelle falls $f^{(n)}(a) > 0$ ist, und Maximalstelle falls $f^{(n)}(a) < 0$ ist.

Beweis: Wegen $f'(a) = \ldots = f^{(n-1)}(a) = 0$ liefert die Taylorformel für alle x aus einer Umgebung von a, daß

$$f(x) = f(a) + \frac{f^{(n)}(y)}{n!}\,(x-a)^n \,,$$

wobei y zwischen a und x liegt. Da $f^{(n)}(a) \neq 0$ ist, folgt aus der Stetigkeit von $f^{(n)}$ daß $f^{(n)}(x) \neq 0$ ist für alle x aus einer Umgebung von a und dasselbe Vorzeichen hat wie $f^{(n)}(a)$. Hieraus schließt man folgendes: Ist n ungerade, dann hat

$$f(x) - f(a) = \frac{f^{(n)}(y)}{n!}\,(x-a)^n$$

für $x > a$ und für $x < a$ verschiedenes Vorzeichen, also kann kein f in a kein Extremum haben.

Ist n gerade, dann ergibt sich für alle x aus einer Umgebung von a, daß

$$\begin{aligned} f(x) - f(a) &\geq 0 \,, &\quad \text{falls} \quad f^{(n)}(y) > 0 \,,\\ f(x) - f(a) &\leq 0 \,, &\quad \text{falls} \quad f^{(n)}(y) < 0 \,, \end{aligned}$$

und hieraus resultiert die Behauptung des Satzes. ■

Grenzwertbestimmung durch Differentiation

Es gilt

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

falls die Grenzwerte $\lim_{n\to a} f(x), \lim_{x\to a} g(x) \neq 0$ existieren. Ist $\lim_{x\to a} f(x) \neq 0$, aber $\lim_{x\to a} g(x) = 0$, dann existiert der Grenzwert $\lim_{x\to a}\frac{f(x)}{g(x)}$ nicht. Es ist aber möglich, daß dieser Grenzwert existiert, falls $\lim_{x\to a} f(x) = \lim_{x\to a} g(x) = 0$ sind. Hierüber machen die sogenannten „Regeln von de l' Hospital“ eine Aussage:

Satz: Die Funktionen f, g seien für $x > a$ definiert und differenzierbar. Es sei $g(x) \neq 0$ und $g'(x) \neq 0$ für $x > a$, und $\lim_{x\to a} f(x) = \lim_{x\to a} g(x) = 0$.

Wenn $\lim_{x\to a}\frac{f'(x)}{g'(x)}$ existiert, dann existiert auch $\lim_{x\to a}\frac{f(x)}{g(x)}$. Außerdem gilt

$$\lim_{x\to a}\frac{f(x)}{g(x)} = \lim_{x\to a}\frac{f'(x)}{g'(x)}\,.$$

Beweis: Man definiere $f(a) = g(a) = 0$. Dann sind die Funktionen f, g stetig fortgesetzt. Sei $x > a$. Aus dem zweiten Mittelwertsatz folgt, daß $y \in (a, x)$ existiert mit

$$\Big(f(x) - f(a)\Big)g'(y) = \Big(g(x) - g(a)\Big)f'(y)\,,$$

also

$$\frac{f(x)}{g(x)} = \frac{f'(y)}{g'(y)}\,.$$

Um zu zeigen, daß $\lim_{x\to a}\frac{f(x)}{g(x)}$ existiert, zeige man, daß für jede Folge $\{x_n\}_{n=1}^{\infty}$ mit $x_n > a$ und $\lim_{n\to\infty} x_n = a$ auch

$$\lim_{n\to\infty}\frac{f(x_n)}{g(x_n)}$$

existiert. Hierzu wähle man zu x_n ein $y_n \in (a, x_n)$ mit

$$\frac{f(x_n)}{g(x_n)} = \frac{f'(y_n)}{g'(y_n)}\,.$$

Dann gilt $\lim_{n\to\infty} y_n = a$, also konvergiert die Folge

$$\left\{\frac{f'(y_n)}{g'(y_n)}\right\}_{n=1}^{\infty} = \left\{\frac{f(x_n)}{g(x_n)}\right\}_{n=1}^{\infty}\,.$$

■

Durch wiederholte Anwendung dieses Satzes ergibt sich

Satz: Die Funktionen f, g seien in $x > a$ definiert und dort n–mal differenzierbar. Für $x > a$ sei $g(x) \neq 0\,, \ldots\,, g^{(n)}(x) \neq 0$ und

$$\lim_{x \to a} g^{(k)}(x) = \lim_{x \to a} f^{(k)}(x) = 0\,, \quad k = 0, 1, \ldots, n-1\,.$$

Wenn dann $\lim_{x \to a} \frac{f^{(n)}(x)}{g^{(n)}(x)}$ existiert, dann auch $\lim_{x \to a} \frac{f(x)}{g(x)}$, und es ist $\lim_{x \to a} \frac{f(x)}{g(x)} = \lim_{x \to a} \frac{f^{(n)}(x)}{g^{(n)}(x)}$.

Sei $f : D \to \mathbb{R}$ und D nach oben unbeschränkt. Um den „Grenzwert von f im Unendlichen“ zu untersuchen, definiere man

$$\lim_{x \to \infty} f(x) = \lim_{y \to 0} f\left(\frac{1}{y}\right).$$

Mit dieser Definition ist die Untersuchung von $\lim_{x \to \infty} \frac{f(x)}{g(x)}$ auf die Untersuchung von $\lim_{y \to 0} \frac{f(1/y)}{g(1/y)}$ zurückgeführt. Die vorangehenden Sätze sind dann auch gültig, wenn man a durch ∞ und die Voraussetzung $x > a$ durch $x < \infty$ ersetzt. Also gilt insbesondere

$$\lim_{x \to \infty} \frac{f(x)}{g(x)} = \lim_{x \to \infty} \frac{f'(x)}{g'(x)}\,.$$

Ähnlich kann auch der Fall behandelt werden, daß $g(x)$ und $f(x)$ unbeschränkt sind in der Umgebung von a :

Satz: Die Funktionen f, g seien in $x > a$ definiert und differenzierbar. Seien $\lim_{x \to a} \frac{1}{g(x)} = 0$ und $\lim_{x \to a} \frac{1}{f(x)} = 0$. Wenn dann $\lim_{x \to a} \frac{f'(x)}{g'(x)}$ existiert, dann auch $\lim_{x \to a} \frac{f(x)}{g(x)}$, und es gilt $\lim_{x \to a} \frac{f(x)}{g(x)} = \lim_{x \to a} \frac{f'(x)}{g'(x)}$.

Beweis: Der Beweis ist etwas komplizierter, weil man f und g nicht stetig nach a fortsetzen kann:
Sei $c = \lim_{x \to a} \frac{f'(x)}{g'(x)}$. Dann gibt es zu $\varepsilon > 0$ eine Zahl $b > a$ mit $\left|\frac{f'(y)}{g'(y)} - c\right| < \varepsilon$ für alle $a < y \leq b$. Es gilt

$$\frac{f(x)}{g(x)} = \frac{f(x) - f(b)}{g(x) - g(b)} \cdot \frac{f(x)}{f(x) - f(b)} \cdot \frac{g(x) - g(b)}{g(x)}\,.$$

Falls x genügend nahe bei a liegt, sind alle Nenner von Null verschieden, da $f(x)$ und $g(x)$ für $x \to a$ „gegen Unendlich konvergieren“. Also folgt für diese x

$$\begin{aligned}
\left|\frac{f(x)}{g(x)} - c\right| \;&\leq\; \left|\frac{f(x) - f(b)}{g(x) - g(b)} - c\right| \, \left|\frac{f(x)}{f(x) - f(b)}\right| \, \left|\frac{g(x) - g(b)}{g(x)}\right| \\
&\quad + |c| \, \left|\frac{f(x)}{f(x) - f(b)} \cdot \frac{g(x) - g(b)}{g(x)} - 1\right|,
\end{aligned}$$

somit, nach dem zweiten Mittelwertsatz,

$$\Big|\frac{f(x)}{g(x)} - c\Big| \leq \varepsilon \Big|\frac{1}{1-\frac{f(b)}{f(x)}}\Big| \cdot \Big|1 - \frac{g(b)}{g(x)}\Big| + |c| \; \Big|\frac{1-\frac{g(b)}{g(x)}}{1-\frac{f(b)}{f(x)}} - 1\Big| .$$

Wegen $\lim_{x\to a} \frac{1}{f(x)} = 0$, $\lim_{x\to a} \frac{1}{g(x)} = 0$, ist $\lim_{x\to a}(1 - \frac{f(b)}{f(x)}) = 1$, $\lim_{x\to a}(1 - \frac{g(b)}{g(x)}) = 1$, und somit auch

$$\lim_{x\to a} \Big(\frac{1-\frac{g(b)}{g(x)}}{1-\frac{f(b)}{f(x)}} - 1\Big) = 0 .$$

Folglich existiert $a < \delta \leq b$, so daß für alle $x \in (a, \delta)$ gilt

$$\Big|\frac{f(x)}{g(x)} - c\,\Big| \leq 2\varepsilon + |c|\varepsilon = (2 + |c|)\varepsilon .$$

Hieraus folgt

$$\lim_{x\to a} \frac{f(x)}{g(x)} = c = \lim_{x\to a} \frac{f'(x)}{g'(x)} .$$

■

Durch wiederholte Anwendung dieses Satzes folgt

Satz: Die Funktionen f, g seien in $x > a$ definiert und n–mal differenzierbar. Sei

$$\lim_{x\to a} \frac{1}{g(x)} = \ldots = \lim_{x\to a} \frac{1}{g^{(n-1)}(x)} = \lim_{x\to a} \frac{1}{f(x)} = \ldots = \lim_{x\to a} \frac{1}{f^{(n-1)}(x)} = 0 .$$

Wenn dann $\lim_{x\to a} \frac{f^{(n)}(x)}{g^{(n)}(x)}$ existiert, dann auch $\lim_{x\to a} \frac{f(x)}{g(x)}$, und es gilt $\lim_{x\to a} \frac{f(x)}{g(x)} = \lim_{x\to a} \frac{f^{(n)}(x)}{g^{(n)}(x)}$.

Außerdem darf in diesen Sätzen auch $a = \infty$ gesetzt werden.

Beispiele: 1.) Sei $s > 0$ und $f(x) = x^s$, $g(x) = e^x$. Es soll untersucht werden, ob

$$\lim_{x\to\infty} \frac{x^s}{e^x}$$

einen Grenzwert hat. Es gilt $\lim_{x\to\infty} x^{-s} = \lim_{x\to\infty} e^{-x} = 0$ und

$$\lim_{x\to\infty} \frac{(x^s)^{(n)}}{(e^x)^{(n)}} = \lim_{x\to\infty} \frac{s(s-1)\ldots(s-n+1)}{x^{n-s}e^x} = 0 ,$$

falls $n \geq s$ gewählt wird. Also folgt aus dem voranstehenden Satz mit $a = \infty$, daß $\lim_{x\to\infty} \frac{x^s}{e^x} = 0$, d.h. die Exponentialfunktion wächst stärker als jede Potenz.

2.) Sei $s > 0$. Wegen $\lim_{x\to\infty} \frac{1}{\log x} = \lim_{x\to\infty} \frac{1}{x^s} = 0$ folgt aus dem voranstehenden Satz mit $a = \infty$ und $n = 1$, daß

$$\lim_{x\to\infty} \frac{\log x}{x^s} = \lim_{x\to\infty} \frac{1}{sx^s} = 0 .$$

Wegen $\lim_{x\to 0} \frac{1}{\log x} = \lim_{x\to 0} \frac{1}{x^{-s}} = 0$ folgt aus dem vorangehenden Satz mit $a = 0$ und $n = 1$, daß

$$\lim_{x\to 0} x^s \log x = \lim_{x\to 0} \frac{\log x}{x^{-s}} = \lim_{x\to 0} \Big(-\frac{1}{sx^{-s}} \Big) = \lim_{x\to 0} \Big(-\frac{1}{s} x^s \Big) = 0\,.$$

Der Logarithmus wächst für $x \to \infty$ und für $x \to 0$ langsamer als jede Potenz.

7f.) Konvexe Funktionen

Definition: Sei J ein Intervall. $f < J \to \mathbb{R}$ heißt konvex, wenn für alle $x_1, x_2 \in J$ und alle $t \in [0,1]$ gilt

$$f\Big((1-t)x_1 + tx_2\Big) \le (1-t)f(x_1) + tf(x_2)\,.$$

$(1-t)f(x_1) + tf(x_2)$

$f((1-t)x_1 + tx_2)$

0 x_1 x_2 $\mathbb{R}$

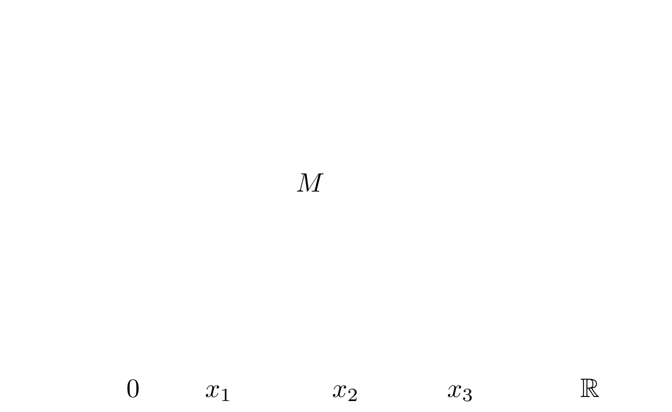
$(1-t)x_1 + tx_2$

f ist also konvex, wenn der Graph von f unterhalb jeder Strecke liegt, die zwei Punkte des Graphen verbindet. Wenn der Graph von f oberhalb jeder Strecke liegt, die zwei Punkte des Graphen verbindet, dann heißt f konkav. Ist f konkav, dann ist $-f$ konvex.

Lemma: Sei f eine konvexe Funktion und seien $x_1, \ldots, x_n \in J$, $t_1 \ldots, t_n \in [0,1]$ mit $t_1 + t_2 + \ldots + t_n = 1$. Dann ist

$$f(t_1 x_1 + \ldots + t_n x_n) \le t_1 f(x_1) + \ldots + t_n f(x_n)\,.$$

Für $n = 3$ ist die geometrische Bedeutung dieser Ungleichung in der folgenden Skizze dargestellt:



Die Punktemenge

$$M = \left\{ \Big(\sum_{i=1}^{3} t_i x_i,\ \sum_{i=1}^{3} t_i f(x_i) \Big) \,\Big|\, t_1 + t_2 + t_3 = 1\,,\ \ t_i \geq 0 \right\}$$

liegt ganz oberhalb des Graphen von f.

Beweis: (Durch Induktion). Nach Definition konvexer Funktionen ist die Aussage richtig für $n = 2$. Angenommen, die Aussage sei richtig für n.
Sei $t_1 + \ldots + t_{n+1} = 1\,,\ \ t_i \geq 0$.
Es gilt

$$\begin{aligned}
&f(t_1 x_1 + t_2 x_2 + \ldots + t_{n+1} x_{n+1}) \\
&\quad = \ f\Big(\Big[\frac{t_1}{1 - t_{n+1}}\, x_1 + \ldots + \frac{t_n}{1 - t_{n+1}}\, x_n \Big] (1 - t_{n+1}) + t_{n+1} x_{n+1} \Big) \\
&\quad \leq \ (1 - t_{n+1}) f\Big(\frac{t_1}{1 - t_{n+1}}\, x_1 + \ldots + \frac{t_n}{1 - t_{n+1}}\, x_n \Big) + t_{n+1} f(x_{n+1}) \\
&\quad \leq \ (1 - t_{n+1}) \Big[\frac{t_1}{1 - t_{n+1}}\, f(x_1) + \ldots + \frac{t_n}{1 - t_{n+1}}\, f(x_n) \Big] + t_{n+1} f(x_{n+1}) \\
&\quad = \ t_1 f(x_1) + \ldots + t_{n+1} f(x_{n+1})\,,
\end{aligned}$$

da

$$\frac{t_1}{1 - t_{n+1}} + \ldots + \frac{t_n}{1 - t_{n+1}} = \frac{1 - t_{n+1}}{1 - t_{n+1}} = 1\,.$$

■

Satz: $f : J \to \mathbb{R}$ sei differenzierbar. Die Funktion f ist genau dann konvex, wenn ihre Ableitung monoton wachsend ist.
Also ist eine zweimal differenzierbare Funktion genau dann konvex, wenn $f''(x) \geq 0$ ist für alle $x \in J$.

Beweis: Sei $x_1 < x < x_2$. Dann folgt

$$\frac{x - x_1}{x_2 - x_1} + \frac{x_2 - x}{x_2 - x_1} = 1 \quad \text{und} \quad \frac{x_2 - x}{x_2 - x_1}\, x_1 + \frac{x - x_1}{x_2 - x_1}\, x_2 = x$$

also

$$\begin{array}{rrcl} & f(x) & \leq & \dfrac{x_2 - x}{x_2 - x_1}\, f(x_1) + \dfrac{x - x_1}{x_2 - x_1}\, f(x_2) \\ \Longleftrightarrow & (x_2 - x_1) f(x) & \leq & (x_2 - x) f(x_1) + (x - x_1) f(x_2) \qquad (*) \\ \Longleftrightarrow & (x_2 - x) f(x) + (x - x_1) f(x) & \leq & (x_2 - x) f(x_1) + (x - x_1) f(x_2) \\ \Longleftrightarrow & \dfrac{f(x) - f(x_1)}{x - x_1} & \leq & \dfrac{f(x_2) - f(x)}{x_2 - x} \end{array}$$

Also ist diese letzte Ungleichung äquivalent zu der Konvexitätsbedingung. Sei nun f' monoton wachsend. Dann folgt aus dem ersten Mittelwertsatz

$$\frac{f(x) - f(x_1)}{x - x_1} = f'(z_1) \leq f'(z_2) = \frac{f(x_2) - f(x)}{x_2 - x}$$

für geeignete $z_1 \in (x, x_1)\,,\ \ z_2 \in (x, x_2)\,$, also ist f konvex. Sei umgekehrt f konvex. Aus $(*)$ folgt

$$(x_2 - x_1) f(x) \leq (x_2 - x_1) f(x_1) + (x_1 - x) f(x_1) + (x - x_1) f(x_2)$$

und

$$(x_2 - x_1) f(x) \leq (x_2 - x) f(x_1) + (x - x_2) f(x_2) + (x_2 - x_1) f(x_2)\,,$$

also

$$\frac{f(x) - f(x_1)}{x - x_1} \leq \frac{f(x_2) - f(x_1)}{x_2 - x_1}$$

und

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} \leq \frac{f(x_2) - f(x)}{x_2 - x}\,,$$

also

$$\begin{array}{rcl} f'(x_1) & = & \displaystyle\lim_{\substack{x \to x_1 \\ x > x_1}} \frac{f(x) - f(x_1)}{x - x_1} \leq \frac{f(x_2) - f(x_1)}{x_2 - x_1} \\ & \leq & \displaystyle\lim_{\substack{x \to x_2 \\ x < x_2}} \frac{f(x_2) - f(x)}{x_2 - x} = f'(x_2)\,, \end{array}$$

also ist f' monoton wachsend. ∎

Anwendung: Sei $f(x) = e^x$. Dann ist $f''(x) = e^x > 0$, also ist e^x konvex. Nach dem obenstehenden Lemma gilt somit für $x_1, x_2, \ldots, x_n \in \mathbb{R}$ und $t_1 + \ldots + t_n = 1\,,\ \ t_i \geq 0$:

$$e^{t_1 x_1 + \ldots + t_n x_n} \leq t_1 e^{x_1} + \ldots + t_n e^{x_n}\,.$$

Sei $y_1, \ldots, y_n \in \mathbb{R}^+$. Setze $x_i = \log y_i$. Dann folgt hieraus

$$y_1^{t_1} y_2^{t_2} \ldots y_n^{t_n} \leq t_1 y_1 + \ldots + t_n y_n \,.$$

Mit $t_i = \frac{1}{n}$ für $i = 1, \ldots, n$ ergibt sich also

$$\sqrt[n]{y_1 \cdot y_2 \cdot \ldots \cdot y_n} \leq \frac{y_1 + \ldots + y_n}{n} \,.$$

Dies ist die Ungleichung zwischen geometrischem und arithmetischem Mittel.

8 Funktionenfolgen, gleichmäßige Konvergenz, Potenzreihen

8a.) Punktweise Konvergenz

Die Exponentialfunktion wurde definiert durch

$$x \mapsto \exp(x) := \sum_{k=0}^{\infty} \frac{x^k}{k!} .$$

Für $n \in \mathbb{N}$ betrachte man die Funktion $f_n : \mathbb{R} \to \mathbb{R}$,

$$f_n(x) := \sum_{k=0}^{n} \frac{x^k}{k!} .$$

Hierdurch wird eine Folge von Funktionen, eine Funktionenfolge $\{f_n\}_{n=1}^{\infty}$, definiert. Wegen

$$\exp(x) = \lim_{n\to\infty} f_n(x) , \quad x \in \mathbb{R} ,$$

konvergiert diese Funktionenfolge „punktweise“ gegen die Exponentialfunktion.

Definition: Sei D eine Menge (nicht notwendig Teilmenge der reellen Zahlen) und $\{f_n\}_{n=1}^{\infty}$ eine Folge von Funktionen $f_n : D \to \mathbb{R}$. Diese Funktionenfolge heißt punktweise konvergent, wenn eine Funktion $f : D \to \mathbb{R}$ existiert mit

$$f(x) = \lim_{n\to\infty} f_n(x)$$

für alle $x \in D$.

Selbstverständlich ist die Grenzfunktion einer punktweise konvergenten Funktionenfolge eindeutig bestimmt, denn:
Notwendig und hinreichend dafür, daß die Funktionenfolge $\{f_n\}_{n=1}^{\infty}$ punktweise konvergiert ist, daß für jedes $x \in D$ die Zahlenfolge $\{f_n(x)\}_{n=1}^{\infty}$ konvergiert. Denn falls jede dieser Zahlenfolgen konvergiert, kann durch

$$f(x) := \lim_{n\to\infty} f_n(x)$$

eine Funktion $f : D \to \mathbb{R}$ definiert werden, die natürlich Grenzfunktion für die Funktionenfolge $\{f_n\}_{n=1}^{\infty}$ ist. Somit ist auch folgender Satz unmittelbar klar:

Satz: Die Funktionenfolge $\{f_n\}_{n=1}^{\infty}$, $f_n : D \to \mathbb{R}$, ist genau dann punktweise konvergent, wenn zu jedem $\varepsilon > 0$ und zu jedem $x \in D$ ein n_0 existiert mit

$$|f_n(x) - f_m(x)| < \varepsilon ,$$

für alle $n, m \geq n_0$ (Cauchykriterium).

Man beachte, daß die Zahl n_0 sowohl von ε als auch von $x \in D$ abhängen darf.

Beispiele:

1.) $D = [0,1]$, $x \mapsto f_n(x) := x^n$. Für $x \in [0,1)$ gilt $\lim_{n\to\infty} f_n(x) = \lim_{n\to\infty} x^n = 0$, und für $x = 1$ gilt $\lim_{n\to\infty} f_n(x) = \lim_{n\to\infty} 1^n = 1$, also konvergiert die Funktionenfolge $\{f_n\}_{n=1}^{\infty}$ punktweise gegen die Grenzfunktion

$$f(x) = \begin{cases} 0\,, & 0 \leq x < 1 \\ 1\,, & x = 1\,. \end{cases}$$

2.) Sei $f_n : \mathbb{R} \to \mathbb{R}$ erklärt durch

$$x \mapsto f_n(x) = \sum_{k=0}^{n} \frac{x^k}{k!}\,.$$

Dann konvergiert $\{f_n\}_{n=1}^{\infty}$ punktweise gegen die Grenzfunktion exp . Man kann die Folge $\{f_n\}_{n=1}^{\infty} = \{\sum_{k=0}^{n} \frac{x^k}{k!}\}_{n=1}^{\infty}$ auch als „Funktionenreihe“ bezeichnen.

3.) $D = [0,2]$,

$$f_n(x) = \begin{cases} n \cdot x\,, & 0 \leq x \leq \frac{1}{n} \\ 2 - n \cdot x\,, & \frac{1}{n} < x < \frac{2}{n} \\ 0\,, & \frac{2}{n} \leq x \leq 2 \end{cases}$$

Diese Funktionenfolge $\{f_n\}_{n=1}^{\infty}$ konvergiert punktweise gegen die Nullfunktion in $[0,2]$.

Beweis: Zu zeigen ist: Für alle $x \in D$ gilt

$$\lim_{n\to\infty} f_n(x) = 0\,.$$

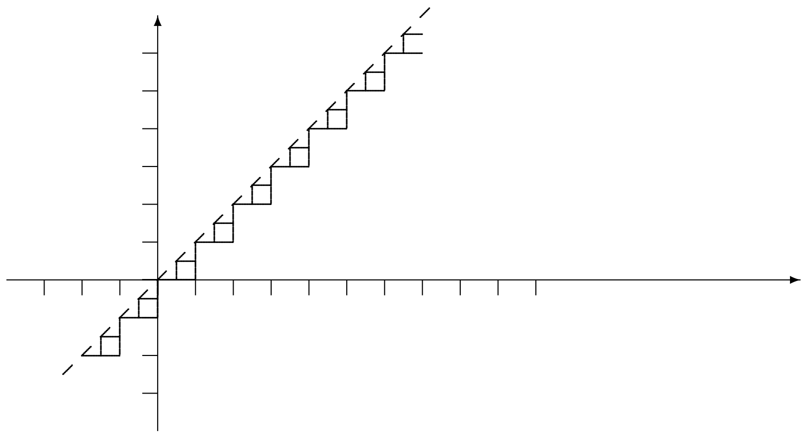
Für $x = 0$ gilt $\lim_{n\to\infty} f_n(x) = \lim_{n\to\infty} 0 = 0$.

Für $x > 0$ gibt es ein $n_0 \in \mathbb{N}$, so daß $\frac{2}{n_0} \leq x$ ist. Für $n \geq n_0$ gilt dann $\frac{2}{n} \leq \frac{2}{n_0} \leq x$, also folgt aus der Definition von f_n für $n \geq n_0$, daß $f_n(x) = 0$ ist, also

$$\lim_{n\to\infty} f_n(x) = 0\,.$$

■

4.) $D = \mathbb{R}$, $x \mapsto f_n(x) = \frac{1}{n}[nx]$. Hierbei bedeutet $[nx]$ die größte ganze Zahl, die kleiner oder gleich n ist. (Gaußklammer).



$\{f_n\}_{n=1}^{\infty}$ konvergiert punktweise gegen die identische Abbildung $x \mapsto f(x) := x$.

Beweis: Seien $x \in \mathbb{R}$ und $n \in \mathbb{N}$. Dann existiert $k \in \mathbb{Z}$ mit $x \in [\frac{k}{n}, \frac{k+1}{n})$, also $nx \in [k, k+1)$, somit

$$f_n(x) = \frac{1}{n}[nx] = \frac{k}{n}\,.$$

Wegen $\frac{k}{n} \leq x < \frac{k+1}{n}$ folgt

$$0 \leq x - \frac{k}{n} < \frac{1}{n}\,,$$

also $|x - f_n(x)| = |x - \frac{k}{n}| < \frac{1}{n}$. Hieraus folgt

$$\lim_{n\to\infty} f_n(x) = x\,.$$

■

8b.) Gleichmäßige Konvergenz, Stetigkeit der Grenzfunktion

Sei $\{f_n\}_{n=1}^{\infty}$ eine Funktionenfolge aus lauter stetigen Funktionen $f_n : D \to \mathbb{R}$. Diese Funktionenfolge konvergiere punktweise gegen die Grenzfunktion $f : D \to \mathbb{R}$. Ist dann auch die Grenzfunktion stetig? Das obenstehende Beispiel 1.) zeigt, daß dies nicht der Fall zu sein braucht, denn

$$(x \mapsto f_n(x) = x^n) : [0,1] \to \mathbb{R}$$

ist stetig, aber die Grenzfunktion

$$f(x) = \begin{cases} 0\,, & x \in [0,1) \\ 1\,, & x = 1 \end{cases}\,,$$

ist nicht stetig. Um auf die Stetigkeit der Grenzfunktion schließen zu können, muß man einen stärkeren Konvergenzbegriff für Funktionenfolgen einführen:

Definition: Sei D eine Menge (nicht notwendig Teilmenge der reellen Zahlen), und sei für jedes $n \in \mathbb{N}$ eine Funktion $f_n : D \to \mathbb{R}$ erklärt. Die Funktionenfolge $\{f_n\}_{n=1}^{\infty}$ heißt gleichmäßig konvergent, wenn eine Funktion $f : D \to \mathbb{R}$ existiert, so daß zu jedem $\varepsilon > 0$ eine Zahl $n_0 \in \mathbb{N}$ existiert mit

$$|f_n(x) - f(x)| < \varepsilon$$

für alle $x \in D$ und alle $n \geq n_0$. Die Funktion f heißt Grenzfunktion.

Mit Quantoren:

$$\forall_{\varepsilon>0}\ \exists_{n_0\in\mathbb{N}}\ \forall_{x\in D}\ \forall_{n\geq n_0}\ :\ |f_n(x) - f(x)| < \varepsilon\,.$$

Der Unterschied zur punktweisen Konvergenz besteht darin, daß zu jedem $\varepsilon > 0$ eine Zahl $n_0 \in \mathbb{N}$ gefunden werden muß, die nicht von $x \in D$ abhängen darf. Konvergiert $\{f_n\}_{n=1}^{\infty}$ gleichmäßig, dann auch punktweise, und die Grenzfunktionen bei gleichmäßiger Konvergenz und bei punktweiser Konvergenz stimmen überein.

Beispiele: 1.) Sei $D = [0,1]$ und $f_n(x) := x^n$, für alle $x \in D$. Die Folge $\{f_n\}_{n=1}^{\infty}$ ist nicht gleichmäßig konvergent.

Beweis: Wäre diese Folge gleichmäßig konvergent, dann müßte sie gegen

$$f(x) = \begin{cases} 0\,, & x \in [0,1) \\ \\ 1\,, & x = 1 \end{cases}$$

konvergieren, weil $\{f_n\}_{n=1}^{\infty}$ punktweise dagegen konvergiert. Ich zeige, daß für dieses f die Negation der Aussage in der Definition der gleichmäßigen Konvergenz gilt:

$$\exists_{\varepsilon>0}\ \forall_{n_0\in\mathbb{N}}\ \exists_{x\in D}\ \exists_{n\geq n_0}\ :\ |f_n(x) - f(x)| \geq \varepsilon\,.$$

Diese Aussage ist für $\varepsilon = \frac{1}{2}$ richtig. Denn sei $n_0 \in \mathbb{N}$. Es genügt $x \in (0,1)$ zu finden mit

$$|f_{n_0}(x) - f(x)| = |f_{n_0}(x)| = x^{n_0} = \frac{1}{2}\,.$$

Dies ist äquivalent zu

$$x = \left(\frac{1}{2}\right)^{1/n_0} = 2^{-1/n_0} = e^{-\frac{\log 2}{n_0}}\,.$$

Wegen $\frac{\log 2}{n_0} > 0$ ist $0 < e^{-\frac{\log 2}{n_0}} < e^0 = 1$, weil die Exponentialfunktion streng monoton wachsend ist, also gilt $0 < (\frac{1}{2})^{\frac{1}{n_0}} < 1$, und somit hat $x = (\frac{1}{2})^{\frac{1}{n_0}}$ die gewünschten Eigenschaften. ∎

2.) Sei $\{f_n\}_{n=1}^{\infty}$ die Funktionenfolge aus dem dritten Beispiel von Abschnitt 9a.). Auch diese Funktionenfolge konvergiert nicht gleichmäßig. Denn sonst müßte sie gegen $f \equiv 0$ konvergieren, weil $\{f_n\}_{n=1}^{\infty}$ punktweise dagegen konvergiert. Für alle $n \in \mathbb{N}$ gilt aber

$$|f_n\Big(\frac{1}{n}\Big) - f\Big(\frac{1}{n}\Big)| = |f_n\Big(\frac{1}{n}\Big)| = 1\,,$$

also ist die Negation der Aussage in der Definition der gleichmäßigen Konvergenz richtig mit $\varepsilon = 1$.

3.) Sei $D = \mathbb{R}$, $x \mapsto f_n(x) = \frac{1}{n}[nx]$. Die Funktionenfolge $\{f_n\}_{n=1}^{\infty}$ konvergiert gleichmäßig gegen $x \mapsto f(x) := x$, denn im vierten Beispiel von Abschnitt 9a.) wurde gezeigt, daß

$$|x - f_n(x)| < \frac{1}{n}$$

gilt für alle $x \in \mathbb{R}$ und $n \in \mathbb{N}$. Sei $\varepsilon > 0$. Wähle $n_0 \in \mathbb{N}$, so daß $\frac{1}{n_0} < \varepsilon$ ist. Dann gilt für alle $x \in \mathbb{R}$ und $n \geq n_0$

$$|f(x) - f_n(x)| = |x - f_n(x)| < \frac{1}{n} \leq \frac{1}{n_0} < \varepsilon\,.$$

Es gilt nun folgender Satz:

Satz: Seien $D \subseteq \mathbb{R}$, $a \in D$, und die Funktionen $f_n : D \to \mathbb{R}$ seien in $a \in D$ stetig. Die Funktionenfolge $\{f_n\}_{n=1}^{\infty}$ konvergiere gleichmäßig gegen $f : D \to \mathbb{R}$. Dann ist f in a stetig.

Beweis: Sei $\varepsilon > 0$. Gesucht ist $\delta > 0$ mit $|f(x) - f(a)| < \varepsilon$ für alle $x \in D$ mit $|x - a| < \delta$. Es gilt für alle $x \in D$ und $n \in \mathbb{N}$

$$\begin{aligned} |f(x) - f(a)| &= |f(x) - f_n(x) + f_n(x) - f_n(a) + f_n(a) - f(a)| \\ &\leq \ |f(x) - f_n(x)| + |f_n(x) - f_n(a)| + |f_n(a) - f(a)|\,. \end{aligned}$$

Da $\{f_n\}_{n=1}^{\infty}$ gleichmäßig gegen f konvergiert, existiert $n_0 \in \mathbb{N}$ mit $|f_n(y) - f(y)| < \frac{\varepsilon}{3}$ für alle $n \geq n_0$ und alle $y \in D$, also folgt

$$|f(x) - f(a)| \leq \frac{2}{3}\varepsilon + |f_{n_0}(x) - f_{n_0}(a)|\,.$$

Da f_{n_0} stetig ist, existert $\delta > 0$ mit $|f_{n_0}(x) - f_{n_0}(a)| < \frac{\varepsilon}{3}$ für $|x - a| < \delta$, also folgt

$$|f(x) - f(a)| < \frac{2}{3}\varepsilon + \frac{1}{3}\varepsilon = \varepsilon$$

für $|x - a| < \delta$. ■

Dieser Satz zeigt, daß

$$\lim_{x\to a}\lim_{n\to\infty} f_n(x) = \lim_{x\to a} f(x) = f(a) = \lim_{n\to\infty} f_n(a) = \lim_{n\to\infty}\lim_{x\to a} f_n(x)$$

gilt. Bei einer gleichmäßig konvergenten Funktionenfolge dürfen also die beiden Grenzwerte $\lim_{x\to a}$ und $\lim_{n\to\infty}$ vertauscht werden.

Beispiel 2.) zeigt, daß die Grenzfunktion f stetig sein kann, ohne daß die Folge $\{f_n\}_{n=1}^{\infty}$ gleichmäßig konvergiert. Der folgende Satz von Dini zeigt, dass man allerdings unter sehr starken Zusatzvoraussetzungen aus der Stetigkeit der Grenzfunktion auf gleichmäßige Konvergenz schließen kann.

Man sagt, eine Funktionenfolge $\{f_n\}_{n=1}^{\infty}$ konveriert punktweise monoton gegen f, wenn $\{|f_n(x) - f(x)|\}_{n=1}^{\infty}$ für alle x aus dem Definitionsbereich eine monoton fallende Nullfolge ist.

Satz (Ulisse Dini, 1845 – 1918): Es sei $D \subseteq \mathbb{R}$ eine kompakte Menge, $f_n : D \to \mathbb{R}$ und $f : D \to \mathbb{R}$ seien stetig, und die Funktionenfolge $\{f_n\}_{n=1}^{\infty}$ konvergiere punktweise monoton gegen f. Dann konvergiert $\{f_n\}_{n=1}^{\infty}$ sogar gleichmäßig gegen f.

Beweis: Sei $\varepsilon > 0$. Jedem $x \in D$ ordne ich in folgender Weise eine Umgebung $U(x)$ zu: Wegen $\lim_{n\to\infty} f_n(x) = f(x)$ existiert eine Zahl $n_0 = n_0(x, \varepsilon)$ mit $|f_{n_0}(x) - f(x)| < \varepsilon$. Da f und f_{n_0} stetig sind, ist auch $|f_{n_0} - f|$ stetig, also existiert eine Umgebung $U(x)$ von x mit $|f_{n_0}(y) - f(y)| < \varepsilon$ für alle $y \in U(x) \cap D$. Das System $\mathcal{U} = \{U(x) \mid x \in D\}$ dieser Umgebungen ist eine offene Überdeckung von D, also genügen bereits endlich viele dieser Umgebungen $U(x_1), \dots, U(x_m)$ zur Überdeckung von D, weil D kompakt ist. Sei nun

$$\tilde{n} = \max\Big\{n_0(x_i, \varepsilon) \mid i = 1, \dots, m\Big\}.$$

Zu jedem $x \in D$ existiert $i \in \{1, \dots, m\}$ mit $x \in U(x_i)$. Nach Konstruktion von $U(x_i)$ gilt dann

$$|f_{n_0(x_i,\varepsilon)}(x) - f(x)| < \varepsilon,$$

und somit, weil $\{f_n(x)\}_{n=1}^{\infty}$ monoton gegen $f(x)$ konvergiert,

$$|f_n(x) - f(x)| < \varepsilon$$

für alle $n \geq n_0(x_i, \varepsilon)$, also insbesondere für alle $n \geq \tilde{n}$. Da $\tilde{n}$ unabhängig von x ist, ist gleichmäßige Konvergenz gezeigt. ■

8c.) Supremumsnorm

Bei der Definition von Konvergenz und Grenzwert von Folgen reeller Zahlen wurde entscheidend benützt, daß man auf den rellen Zahlen einen Abstandsbegriff hat, den Absolutbetrag der Differenz zweier Zahlen. In Abschnitt 9 a.) wurde gezeigt, daß man mit diesem Konvergenzbegriff von $\mathbb{R}$ auch Konvergenz auf der Algebra $F(D,\mathbb{R})$ der reellwertigen Funktionen erklären kann, die punktweise Konvergenz, ohne einen vergleichbaren einfachen Abstandsbegriff auf dieser Algebra zu haben. Auf der Algebra der *beschränkten* reellwertigen Funktionen kann man aber einen einfachen Abstandsbegriff definieren. Hierzu führt man die Norm einer Funktion ein, die ähnliche Eigenschaften hat wie der Absolutbetrag einer reellen Zahl. Definiert man mit dieser Norm Konvergenz von Funktionen ähnlich wie man mit dem Absolutbetrag Konvergenz von Zahlenfolgen definiert, dann ergibt sich die gleichmäßige Konvergenz.

Definition: Sei D eine Menge, die nicht notwendig eine Teilmenge der reellen Zahlen zu sein braucht. Sei $f : D \to \mathbb{R}$ eine beschränkte Funktion.
Die nichtnegative reelle Zahl

$$\|f\| := \sup_{x\in D} |f(x)|$$

heißt Norm (speziell: Supremumsnorm) von f.

Satz: Seien $f, g : D \to \mathbb{R}$ beschränkte Funktionen. Dann gilt

(i) $\|f\| = 0 \iff f(x) = 0$ für alle $x \in D$

(ii) $\|cf\| = |c|\,\|f\|$, für alle $c \in \mathbb{R}$

(iii) $\|f + g\| \le \|f\| + \|g\|$

(iv) $\|f \cdot g\| \le \|f\| \cdot \|g\|$.

Beweis: (i), (ii) sind klar.

(iii) Für alle $x \in D$ gilt

$$\begin{aligned} |(f+g)(x)| = |f(x) + g(x)| &\le |f(x)| + |g(x)| \\ &\le \|f\| + \|g\|. \end{aligned}$$

Also ist $\|f\| + \|g\|$ eine obere Schranke für die Menge $\{|(f+g)(x)| \mid x \in D\}$, also ist

$$\|f+g\| = \sup_{x\in D} |(f+g)(x)| \le \|f\| + \|g\|.$$

Ebenso gilt für alle $x \in D$

$$|(f\cdot g)(x)| = |f(x)\cdot g(x)| = |f(x)|\cdot|g(x)| \le \|f\|\,\|g\|,$$

also

$$\|f \cdot g\| = \sup_{x \in D} |(fg)(x)| \leq \|f\| \cdot \|g\| .$$

■

Definition: Sei V ein Vektorraum. Eine Abbildung $\|\cdot\| : V \to \mathbb{R}_0^+$, die die Eigenschaften (i), (ii), (iii) hat, heißt Norm auf V . Ist A eine Algebra, dann heißt $\|\cdot\| : A \to \mathbb{R}_0^+$ eine Norm, wenn die Eigenschaften (i) – (iv) erfüllt sind.

Satz: Sei D eine Menge, $f_n, f : D \to \mathbb{R}$ seien beschränkte Funktionen. Die Funktionenfolge $\{f_n\}_{n=1}^{\infty}$ konvergiert genau dann gleichmäßig gegen f , wenn

$$\lim_{n\to\infty} \|f_n - f\| = 0$$

gilt.

Beweis: Es konvergiere $\{f_n\}_{n=1}^{\infty}$ gleichmäßig gegen f . Sei $\varepsilon > 0$. Dann existiert $n_0 \in \mathbb{N}$ so daß für alle $n \geq n_0$ und alle $x \in D$

$$|f_n(x) - f(x)| < \varepsilon$$

gilt. Somit ist

$$\|f_n - f\| = \sup_{x \in D} |f_n(x) - f(x)| \leq \varepsilon$$

für alle $n \geq n_0$, also gilt $\lim_{n\to\infty} \|f_n - f\| = 0$.

Umgekehrt gelte $\lim_{n\to\infty} \|f_n - f\| = 0$. Sei $\varepsilon > 0$. Dann existiert $n_0 \in \mathbb{N}$ so daß

$$\|f_n - f\| = \sup_{x \in D} |f_n(x) - f(x)| < \varepsilon$$

ist für alle $n \geq n_0$, also konvergiert $\{f_n\}_{n=1}^{\infty}$ gleichmäßig gegen f . ■

Satz: Sei $D \subseteq \mathbb{R}$ eine Menge, $f_n : D \to \mathbb{R}$ seien beschränkte Funktionen. Dann gilt: Die Funktionenfolge konvergiert genau dann gleichmäßig, wenn zu jedem $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$ existiert, so daß für alle $n, m \geq n_0$ gilt

$$\|f_n - f_m\| < \varepsilon .$$

($\{f_n\}_{n=1}^{\infty}$ ist Cauchyfolge bezüglich der Norm.)

Beweis: Falls $\{f_n\}_{n=1}^{\infty}$ gleichmäßig konvergiert, existiert eine beschränkte Funktion $f : D \to \mathbb{R}$, so daß $\{\|f_n - f\|\}_{n=1}^{\infty}$ eine Nullfolge ist. Somit existiert zu $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$ mit

$$\|f_n - f_m\| = \|f_n - f + f - f_m\| \leq \|f_n - f\| + \|f - f_m\| < 2\varepsilon$$

für alle $n, m \geq n_0$.

Sei umgekehrt $\{f_n\}_{n=1}^\infty$ eine Cauchyfolge bezüglich der Norm. Um zu zeigen, daß $\{f_n\}_{n=1}^\infty$ gleichmäßig konvergiert, muß zunächst die Grenzfunktion f bestimmt werden.

Für jedes $x \in D$ ist $\{f_n(x)\}_{n=1}^\infty$ eine Cauchyfolge. Denn sei $\varepsilon > 0$. Dann existiert nach Voraussetzung $n_0 \in \mathbb{N}$ mit

$$|f_n(x) - f_m(x)| \leq \|f_n - f_m\| < \varepsilon\,,$$

für alle $n, m, \geq n_0$. Somit ist $\{f_n(x)\}_{n=1}^\infty$ Cauchyfolge in $\mathbb{R}$ für alle $x \in D$, also ist $\{f_n\}_{n=1}^\infty$ punktweise konvergent. Sei $f : D \to \mathbb{R}$ die Grenzfunktion im Sinn der punktweisen Konvergenz. f ist auch Grenzfunktion im Sinn der gleichmäßigen Konvergenz. Denn sei $\varepsilon > 0$. Es existiert $n_0 \in \mathbb{N}$ mit $\|f_n - f_m\| < \varepsilon$ für $n, m \geq n_0$. Sei $x \in D$ und $n \geq n_0$. Dann folgt

$$|f_n(x) - f(x)| = |f_n(x) - \lim_{m\to\infty} f_m(x)| = \lim_{m\to\infty} |f_n(x) - f_m(x)| \leq \varepsilon\,,$$

also

$$\|f_n - f\| = \sup_{x\in D} |f_n(x) - f(x)| \leq \varepsilon\,,$$

für $n \geq n_0$, weil n_0 nicht von $x \in D$ abhängt. ■

8d.) Gleichmäßige Konvergenz von Funktionenreihen

Sei D eine Menge, $f_n : D \to \mathbb{R}$ seien Funktionen. Die Funktionenreihe $\sum_{n=1}^\infty f_n$ heißt gleichmäßig konvergent, wenn die Folge $\{\sum_{n=1}^m f_n\}_{m=1}^\infty$ der Partialsummen gleichmäßig konvergiert.

Satz (Majorantenkriterium von Weierstraß): Sei D eine Menge, seien die Funktionen $f_n : D \to \mathbb{R}$ beschränkt mit $\|f_n\| \leq c_n$, und sei $\sum_{n=1}^\infty c_n$ konvergent. Dann konvergiert die Funktionenreihe $\sum_{n=1}^\infty f_n$ gleichmäßig.

Beweis: Nach dem vorangehenden Satz genügt es zu zeigen, daß $\sum_{n=1}^\infty f_n$ das Cauchysche Konvergenzkriterium bezüglich der Norm erfüllt, d. h. daß zu jedem $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$ existiert mit

$$\|\sum_{k=n}^m f_k\| < \varepsilon\,,$$

für alle $m \geq n \geq n_0$. Da $\sum_{k=1}^\infty c_k$ konvergiert, existiert $n_0 \in \mathbb{N}$ mit $|\sum_{k=n}^m c_k| = \sum_{k=n}^m c_k < \varepsilon$, für $m \geq n \geq n_0$, also ergibt sich

$$\|\sum_{k=n}^m f_k\| \leq \sum_{k=n}^m \|f_k\| \leq \sum_{k=n}^m c_k < \varepsilon\,,$$

für $m \geq n \geq n_0$. ■

8e.) Differenzierbarkeit der Grenzfunktion

Sei $D \subseteq \mathbb{R}$ und die Folge $\{f_n\}_{n=1}^{\infty}$, $f_n : D \to \mathbb{R}$, konvergiere gleichmäßig gegen $f : D \to \mathbb{R}$. Wenn alle f_n differenzierbar sind, ist dann f differenzierbar? An einfachen Beispielen sieht man, daß dies im allgemeinen nicht richtig ist. Man kann auch fragen, ob die Folge $\{f_n'\}_{n=1}^{\infty}$ der Ableitungen gegen f' konvergiert, falls f' existiert. Das folgende Beispiel zeigt, daß auch dies nicht der Fall zu sein braucht. Betrachte $D = [0,1]$, $f_n = \frac{1}{n}x^n$. Die Folge $\{f_n\}_{n=1}^{\infty}$ konvergiert gleichmäßig gegen $f \equiv 0$, aber $f_n' = x^{n-1}$ konvergiert nicht gleichmäßig auf $[0,1]$. Punktweise konvergiert $\{f_n'\}_{n=1}^{\infty}$ gegen

$$g(x) = \begin{cases} 0\,, & 0 \leq x < 1 \\ 1\,, & x = 1 \end{cases}\,,$$

aber $g \neq f' \equiv 0$.

Es gilt jedoch folgender Satz:

Satz: Sei $-\infty < a < b < \infty$, und seien die Funktionen $f_n : [a,b] \to \mathbb{R}$ differenzierbar. Die Folge $\{f_n'\}_{n=1}^{\infty}$ sei gleichmäßig konvergent, und die Folge $\{f_n\}_{n=1}^{\infty}$ konvergiere wenigstens in einem Punkt $x_0 \in [a,b]$. Dann kovergiert $\{f_n\}_{n=1}^{\infty}$ gleichmäßig gegen eine differenzierbare Funktion $f : [a,b] \to \mathbb{R}$, und es gilt

$$f'(x) = \lim_{n\to\infty} f_n'(x)$$

für alle $x \in [a,b]$.

Dies bedeutet, daß man die Bildung der Ableitung (Grenzübergang!) mit dem Grenzübergang bezüglich n vertauschen darf:

$$(\lim_{n\to\infty} f_n)' = \lim_{n\to\infty} f_n'\,.$$

Beweis: Zunächst zeigt man, daß $\{f_n\}_{n=1}^{\infty}$ gleichmäßig konvergiert. Sei $\varepsilon > 0$. Es gilt für $x \in [a,b]$

$$\begin{aligned} |f_m(x) - f_n(x)| \;\leq\; & \Big|\Big(f_m(x) - f_n(x)\Big) - \Big(f_m(x_0) - f_n(x_0)\Big)\Big| \\ & + |f_m(x_0) - f_n(x_0)|\,. \end{aligned} \tag{$*$}$$

Da die Funktion $f_m - f_n$ differenzierbar ist, folgt aus dem ersten Mittelwertsatz

$$\Big|\Big(f_m(x) - f_n(x)\Big) - \Big(f_m(x_0) - f_n(x_0)\Big)\Big| = |f_m'(z) - f_n'(z)|\,|x - x_0|\,,$$

mit geeignetem z zwischen x_0 und x.

Da die Folge der Ableitungen gleichmäßig konvergiert, existiert ein $n_0 \in \mathbb{N}$ mit

$$|f_m'(z) - f_n'(z)| < \frac{\varepsilon}{2(b-a)}$$

für alle $m, n \geq n_0$, also folgt

$$\Big|\Big(f_m(x) - f_n(x)\Big) - \Big(f_m(x_0) - f_n(x_0)\Big)\Big| \leq \frac{\varepsilon}{2},$$

für alle $m, n \geq n_0$ und alle $x \in [a, b]$. Da die Folge $\{f_n(x_0)\}_{n=1}^{\infty}$ konvergiert, existiert $n_1 \in \mathbb{N}$ mit

$$|f_m(x_0) - f_n(x_0)| \leq \frac{\varepsilon}{2}$$

für alle $m, n \geq n_1$. Somit folgt aus $(*)$

$$\|f_m - f_n\| \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

für $m, n \geq n_2 = \max(n_0, n_1)$, also ist $\{f_n\}_{n=1}^{\infty}$ gleichmäßig konvergent. Die Grenzfunktion sei f.

Sei $c \in [a, b]$, und sei für $x \in [a, b]$

$$F(x) = \begin{cases} \dfrac{f(x) - f(c)}{x - c} - m\,, & x \neq c \\[2ex] 0\,, & x = c \end{cases}$$

mit $m = \lim_{n\to\infty} f_n'(c)$. Die Behauptung des Satzes folgt, wenn F stetig ist im Punkt $x = c$, weil dann f differenzierbar ist mit Ableitung $f'(c) = m = \lim_{n\to\infty} f_n'(c)$. Zum Beweis, daß F stetig ist in c, setze

$$F_n(x) = \begin{cases} \dfrac{f_n(x) - f_n(c)}{x - c} - f_n'(c) & x \neq c \\[2ex] 0\,, & x = c\,. \end{cases}$$

Es gilt $F(x) = \lim_{n\to\infty} F_n(x)$, und da F_n stetig ist, weil f_n differenzierbar ist, genügt es zu zeigen, daß $\{F_n\}_{n=1}^{\infty}$ gleichmäßig konvergiert. Dies ergibt sich durch Anwendung des Mittelwertsatzes auf die differenzierbare Funktion $f_m - f_n$:

$$\begin{aligned} F_m(x) - F_n(x) \quad &= \quad \begin{cases} \dfrac{(f_m(x) - f_n(x)) - (f_m(c) - f_n(c))}{x - c} - \Big(f_m'(c) - f_n'(c)\Big)\,, & x \neq c \\ 0\,, & x = c\,, \end{cases} \\ &= \quad \Big(f_m'(z) - f_n'(z)\Big) - \Big(f_m'(c) - f_n'(c)\Big) \end{aligned}$$

für geeignetes z zwischen x und c falls $x \neq c$, und für $z = c$ falls $x = c$. Da $\{f_n'\}_{n=1}^{\infty}$ gleichmäßig konvergiert, existiert $n_0 \in \mathbb{N}$ mit

$$\begin{aligned} |F_m(x) - F_n(x)| &\leq |f_m'(z) - f_n'(z)| + |f_m'(c) - f_n'(c)| \\ &< \varepsilon + \varepsilon = 2\varepsilon \end{aligned}$$

für $m, n \geq n_0$. Da n_0 unabhängig von x gewählt werden kann, ist also $\{F_n\}_{n=1}^{\infty}$ gleichmäßig konvergent. ■

8f.) Potenzreihen

Sei eine Folge $\{a_n\}_{n=0}^{\infty}$ und eine reelle Zahl x_0 gegeben. Für beliebiges $x \in \mathbb{R}$ betrachte man die Reihe

$$\sum_{n=0}^{\infty} a_n (x - x_0)^n .$$

Man nennt diese Reihe eine Potenzreihe. a_m heißen Koeffizienten, x_0 Entwicklungspunkt der Potenzreihe. Beispiele sind Taylorreihen und die Reihe für die Exponentialfunktion. An diesen Beispielen sieht man, daß Potenzreihen hauptsächlich interessant sind, wenn man sie als Funktionenfolgen $\{f_m\}_{m=1}^{\infty}$

$$x \mapsto f_m(x) := \sum_{n=0}^{m} a_n (x - x_0)^n$$

auffaßt. Zunächst muß die Konvergenz von Potenzreihe untersucht werden.

Satz: Sei $\overline{\lim}_{n\to\infty} \sqrt[n]{|a_n|} = a < \infty$. Die Potenzreihe

$$\sum_{n=0}^{\infty} a_n (x - x_0)^n$$

ist im Fall

$$\begin{aligned} a &= 0: \quad \text{absolut konvergent für alle } x \in \mathbb{R} \\ a &> 0: \quad \begin{cases} \text{absolut konvergent für} \quad |x - x_0| < \dfrac{1}{a} \\ \text{divergent für} \quad |x - x_0| > \dfrac{1}{a} \\ \text{konvergent oder divergent für} \quad |x - x_0| = \dfrac{1}{a} . \end{cases} \end{aligned}$$

Falls $\{\sqrt[n]{|a_n|}\}_{n=0}^{\infty}$ nicht beschränkt ist, konvergiert die Potenzreihe nur für $x = x_0$.

Beweis: Nach dem Wurzelkriterium gilt: $\sum_{n=0}^{\infty} a_n (x - x_0)^n$ ist absolut konvergent, falls

$$\overline{\lim}_{n\to\infty} \sqrt[n]{|a_n|\, |x - x_0|^n} = |x - x_0|\, \overline{\lim}_{n\to\infty} \sqrt[n]{|a_n|} < 1$$

ist, also falls $|x-x_0| < \frac{1}{a}$ gilt. Falls $\overline{\lim}_{n\to\infty} \sqrt[n]{|a_n|\,|x-x_0|^n} = |x-x_0|\,\overline{\lim}_{n\to\infty} \sqrt[n]{|a_n|} > 1$ ist, also falls $|x-x_0| > \frac{1}{a}$ ist, divergiert die Reihe. Falls $\{\sqrt[n]{|a_n|}\}_{n=0}^\infty$ unbeschränkt ist, ist $\{|x-x_0|\sqrt[n]{|a_n|}\}_{n=0}^\infty$ für alle $x \neq x_0$ unbeschränkt, also ist $\{a_n(x-x_0)^n\}_{n=0}^\infty$ keine Nullfolge, also kann $\sum_{n=0}^\infty a_n(x-x_0)^n$ nicht konvergieren für $x \neq x_0$. ■

Definition: Sei $a = \overline{\lim}_{n\to\infty} \sqrt[n]{|a_n|}$. Die Zahl

$$r = \begin{cases} \dfrac{1}{a}\,, & \text{für} \quad a \neq 0 \\ \infty\,, & \text{für} \quad a = 0\,, \end{cases}$$

heißt der Konvergenzradius der Potenzreihe

$$\sum_{n=0}^\infty a_n(x-x_0)^n\,.$$

Beispiele: 1.) Der Konvergenzradius der Reihen

$$\sum_{n=0}^\infty x^n\,, \qquad \sum_{n=1}^\infty \frac{1}{n}\,x^n$$

ist 1, wegen

$$\lim_{n\to\infty} \sqrt[n]{n} = \lim_{n\to\infty} e^{1/n\,\log n} = e^{\lim_{n\to\infty} 1/n \log n} = e^0 = 1\,,$$

da $\lim_{x\to\infty} \frac{\log x}{x} = 0$, nach der Regel von de l' Hospital, also

$$r = \frac{1}{\overline{\lim}_{n\to\infty} \sqrt[n]{\frac{1}{n}}} = \frac{1}{\frac{1}{\lim_{n\to\infty} \sqrt[n]{n}}} = \lim_{n\to\infty} \sqrt[n]{n} = 1\,.$$

Für $x = 1$ divergieren beide Reihen, für $x = -1$ divergiert die erste, aber die zweite konvergiert.

2.) Der Konvergenzradius von

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

ist unendlich, wegen $\overline{\lim}_{n\to\infty} \sqrt[n]{\frac{1}{n!}} = 0$.

Bemerkung: In Abschnitt 8 d.) wurde gezeigt, daß

$$\log(1+x) = \sum_{n=1}^\infty (-1)^{n-1} \frac{x^n}{n} \quad \text{für} \quad -\frac{1}{2} < x \leq 1$$

also

$$\log x = \sum_{n=1}^\infty \frac{(-1)^{n-1}}{n}\,(x-1)^n\,, \quad \text{für} \quad \frac{1}{2} < x \leq 2$$

gilt. Tatsächlich gilt die letzte Gleichung für alle $0 < x \leq 2$, log wird durch diese Reihe also für alle x im „Konvergenzkreis“ dargestellt.

Beweis: Siehe später

Satz: Seien $\sum_{n=0}^{\infty} a_n(x-x_0)^n$, $\sum_{n=0}^{\infty} b_n(x-x_0)^n$ Potenzreihen mit Konvergenzradius r_1 bzw. r_2. Dann gilt für alle x mit $|x-x_0| < r$, $r = \min(r_1, r_2)$:

$$\begin{aligned} \sum_{n=0}^{\infty} a_n(x-x_0)^n + \sum_{n=0}^{\infty} b_n(x-x_0)^n &= \sum_{n=0}^{\infty}(a_n+b_n)(x-x_0)^n \\ \Big[\sum_{n=0}^{\infty} a_n(x-x_0)^n\Big]\Big[\sum_{n=0}^{\infty} b_n(x-x_0)^n\Big] &= \sum_{n=0}^{\infty}\Big(\sum_{k=0}^{n} a_k b_{n-k}\Big)(x-x_0)^n . \end{aligned}$$

Beweis: Die Aussagen folgen unmittelbar aus den Sätzen über das Rechnen mit Reihen und aus dem Satz über das Cauchyprodukt zweier Reihen (Beachte, daß der Konvergenzradius der rechtsstehenden Reihen mindestens gleich r ist, aber größer sein kann.) ∎

Satz: Sei $\sum_{n=0}^{\infty} a_n(x-x_0)^n$ eine Potenzreihe mit Konvergenzradius r. Dann ist die Potenzreihe in jedem kompakten Intervall $x \in [x_0 - r_1, x_0 + r_1]$ mit $0 \leq r_1 < r$ gleichmäßig konvergent.

Beweis: Dies folgt aus dem Weierstraßschen Majorantenkriterium: Es gilt

$$\overline{\lim}_{n\to\infty} \sqrt[n]{|a_n| r_1^n} = r_1 \frac{1}{r} < 1 ,$$

also liefert das Wurzelkriterium die Konvergenz der Reihe

$$\sum_{n=0}^{\infty} |a_n| r_1^n .$$

Wegen $|a_n(x-x_0)^n| \leq |a_n| r_1^n$ für $|x-x_0| \leq r_1$ folgt nun die Behauptung aus dem Majorantenkriterium. ∎

Folgerung: Sei $\sum_{n=0}^{\infty} a_n(x-x_0)^n$ eine Potenzreihe mit Konvergenzradius $r > 0$. Dann ist die durch diese Potenzreihe dargestellte Funktion $f : (x_0 - r, x_0 + r) \to \mathbb{R}$, $f(x) = \sum_{n=0}^{\infty} a_n(x-x_0)^n$, stetig.

Beweis: Da $\sum_{n=0}^{m} a_n(x-x_0)^n$ stetig ist, und da die Potenzreihe in jedem kompakten Intervall $[x_0 - r_1, x_0 + r_1]$ gleichmäßig konvergiert, ist die Grenzfunktion f in jedem derartigen Intervall stetig. Wegen $(x_0 - r, x_0 + r) = \bigcup_{0 \leq r_1 < r} [x_0 - r_1, x_0 + r_1]$ ist f im ganzen Konvergenzkreis stetig. ∎

Jede der Funktionen $f_m(x) = \sum_{n=0}^{m} a_n(x-x_0)^n$ ist differenzierbar mit Ableitung

$$f_m'(x) = \sum_{n=1}^{m} na_n(x-x_0)^{n-1} .$$

Die Potenzreihe $\sum_{n=1}^{\infty} na_n(x-x_0)^{n-1}$ hat denselben Konvergenzradius wie die Potenzreihe $\sum_{n=0}^{\infty} a_n(x-x_0)^n$. Zum Beweis beachte man, daß wegen

$$\sum_{n=1}^{\infty} na_n(x-x_0)^{n-1} = \frac{1}{x-x_0} \sum_{n=1}^{\infty} na_n(x-x_0)^n$$

und wegen

$$\begin{aligned} \overline{\lim}_{n\to\infty} \sqrt[n]{n|a_n|} &= \lim_{n\to\infty} \sqrt[n]{n}\ \overline{\lim}_{n\to\infty} \sqrt[n]{|a_n|} \\ &= \overline{\lim}_{n\to\infty} \sqrt[n]{|a_n|} \end{aligned}$$

die Potenzreihe $\sum_{n=1}^{\infty} na_n(x-x_0)^{n-1}$ absolut konvergiert für alle x mit $0 < |x-x_0| < \frac{1}{\overline{\lim}_{n\to\infty} \sqrt[n]{|a_n|}}$, also stimmt diese Zahl mit dem Konvergenzradius überein. Also ist die Folge der Ableitungen $\{f_m'\}_{m=1}^{\infty}$ in jedem kompakten Teilintervall des Konvergenzkreises gleichmäßig konvergent. Aus dem Satz über die Differenzierbarkeit der Grenzfunktion ergibt sich also, daß die Funktion $f(x) = \sum_{n=0}^{\infty} a_n(x-x_0)^n$ in jedem kompakten Teilintervall des Konvergenzkreises differenzierbar ist, also sogar im ganzen Konvergenzkreis, und daß die Ableitung gliedweise berechnet werden darf:

$$f'(x) = \sum_{n=1}^{\infty} na_n(x-x_0)^{n-1} .$$

Durch Wiederholung dieses Schlusses ergibt sich:

Satz: Sei $f(x) = \sum_{n=0}^{\infty} a_n(x-x_0)^n$ eine Potenzreihe mit Konvergenzradius $r > 0$. Dann ist f beliebig oft differenzierbar im Konvergenzkreis, und es gilt

$$f^{(k)}(x) = \sum_{n=k}^{\infty} n(n-1)\dots(n-k+1)a_n(x-x_0)^{n-k}$$

Anwendung: Es gilt

$$\log(1+x) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n}\, x^n$$

für alle $-1 < x < 1$. Zum Beweis beachte man, daß der Konvergenzradius der rechtsstehenden Reihe 1 ist, also konvergiert diese Reihe für $|x| < 1$ und stellt dort eine beliebig oft differenzierbare Funktion dar. Für die Ableitung dieser Funktion ist

$$\sum_{n=1}^{\infty} (-1)^{n-1}x^{n-1} = \sum_{n=0}^{\infty} (-x)^n = \frac{1}{1+x} = \Big[\log(1+x)\Big]' ,$$

also gilt

$$\log(1+x) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n}\, x^n + C$$

mit geeigneter Konstanten C. Um C zu bestimmen, setze man $x = 0$. Wegen $\log(1) = 0$ ergibt sich $C = 0$.

Identitätssatz für Potenzreihen: Seien die Konvergenzradien r_1, r_2 der Potenzreihen $\sum_{n=0}^{\infty} a_n(x-x_0)^n$ und $\sum_{n=0}^{\infty} b_n(x-x_0)^n$ größer als Null, und sei $0 < r \leq \min(r_1, r_2)$. In der Umgebung $U = U_r(x_0) = \{x \in \mathbb{R} \,\big|\, |x - x_0| < r\}$ gelte

$$\sum_{n=0}^{\infty} a_n(x-x_0)^n = \sum_{n=0}^{\infty} b_n(x-x_0)^n\,,$$

für alle $x \in U$. Dann gilt $a_n = b_n$, für alle $n = 0, 1, 2, \ldots$.

Beweis: Zunächst wähle man $x = x_0$. Dann folgt

$$a_0 = b_0\,.$$

Angenommen, es sei für ein $n \in \mathbb{N} \cup \{0\}$ gezeigt, daß $a_k = b_k$ gilt für alle $0 \leq k \leq n$. Zu zeigen ist, daß dann auch $a_{n+1} = b_{n+1}$ gilt. Aus der Induktionsannahme und aus der Voraussetzung des Satzes folgt

$$\sum_{k=n+1}^{\infty} a_k(x-x_0)^k = \sum_{k=n+1}^{\infty} b_k(x-x_0)^k\,,$$

also

$$\sum_{k=n+1}^{\infty} (a_k - b_k)(x-x_0)^k = 0 \quad \text{für alle } x \in U\,.$$

Für $x \in U$, $x \neq x_0$, ergibt sich hieraus

$$\sum_{k=n+1}^{\infty} (a_k - b_k)(x-x_0)^{k-n-1} = 0\,,$$

also

$$\begin{aligned} 0 &= \lim_{\substack{x \to x_0 \\ x \neq x_0}} \sum_{k=n+1}^{\infty} (a_k - b_k)(x-x_0)^{k-n-1} \\ &= \sum_{k=n+1}^{\infty} (a_k - b_k)(x_0 - x_0)^{k-n-1} = a_{n+1} - b_{n+1}\,, \end{aligned}$$

also $a_{n+1} = b_{n+1}$, weil

$$\sum_{k=n+1}^{\infty} (a_k - b_k)(x - x_0)^{k-n-1} = \sum_{k=0}^{\infty} (a_{k+n+1} - b_{k+n+1})(x - x_0)^k$$

eine in U konvergente Potenzreihe ist, also dort stetig ist. ∎

Abelscher Grenzwertsatz: Jede Potenzreihe definiert eine im Innern des Konvergenzkreises stetige Funktion. Eine Aussage über die Stetigkeit der Potenzreihe auf dem Rand des Konvergenzkreises erhält man aus folgendem Satz:

Satz: Sei $\sum_{n=0}^{\infty} a_n(x - x_0)^n$ eine Potenzreihe und sei $z \in \mathbb{R}$ ein Punkt auf dem Rand des Konvergenzkreises, so daß $\sum_{n=0}^{\infty} a_n(z - x_0)^n$ konvergiert. Dann ist die Potenzreihe im Intervall $[z, x_0]$ (falls $z < x_0$) bzw. im Intervall $[x_0, z]$ gleichmäßig konvergent.

Folgerung (Abelscher Grenzwertsatz): (Niels Hendrik Abel, 1802 – 1829) Sei die Potenzreihe $\sum_{n=0}^{\infty} a_n(x - x_0)^n$ in einem Punkt z auf dem Rand des Konvergenzkreises noch konvergent. Dann ist die Potenzreihe im Punkt z stetig.

Beweis des Satzes: Sei o.B.d.A. $z > x_0$ und sei

$$g_k = \sum_{j=k+1}^{\infty} a_j(z - x_0)^j .$$

Dann gilt

$$\begin{aligned}
\sum_{k=1}^{n} a_k(x - x_0)^k &= \sum_{k=1}^{n} a_k(z - x_0)^k \Big(\frac{x - x_0}{z - x_0}\Big)^k \\
&= -\sum_{k=1}^{n} (g_k - g_{k-1}) \Big(\frac{x - x_0}{z - x_0}\Big)^k \\
&= \sum_{k=1}^{n-1} g_k \Big[\Big(\frac{x - x_0}{z - x_0}\Big)^{k+1} - \Big(\frac{x - x_0}{z - x_0}\Big)^k\Big] - g_n \Big(\frac{x - x_0}{z - x_0}\Big)^n + g_0 \Big(\frac{x - x_0}{z - x_0}\Big) .
\end{aligned}$$

Aus dieser Gleichung folgt: Um zu zeigen, daß die Funktionenreihe $\sum_{k=0}^{\infty} a_k(x - x_0)^k$ in $[x_0, z]$ gleichmäßig konvergiert, genügt es zu zeigen, daß die Funktionenfolge

$$\Big\{ g_n \Big(\frac{x - x_0}{z - x_0}\Big)^n \Big\}_{n=1}^{\infty}$$

und die Funktionenreihe

$$\sum_{k=1}^{\infty} g_k \Big[\Big(\frac{x - x_0}{z - x_0}\Big)^{k+1} - \Big(\frac{x - x_0}{z - x_0}\Big)^k\Big]$$

gleichmäßig konvergieren. Es gilt

$$\begin{aligned} \sup_{x_0 \le x \le z} \Big| g_n \Big(\frac{x-x_0}{z-x_0} \Big)^n \Big| &= \Big| \sum_{j=n+1}^{\infty} a_j (z-x_0)^j \Big| \sup_{x_0 \le x \le z} \Big| \frac{x-x_0}{z-x_0} \Big|^n \\ &\le \Big| \sum_{j=n+1}^{\infty} a_j (z-x_0)^j \Big| \to 0 \quad \text{für} \quad n \to \infty \,, \end{aligned}$$

also ist die Folge $\{g_n(\frac{x-x_0}{z-x_0})^n\}_{n=1}^{\infty}$ gleichmäßig konvergent gegen 0. Weiter gilt: Sei $\varepsilon > 0$. Wähle n_0 so groß, daß $|g_k| < \varepsilon$ gilt für alle $k \ge n_0$. Wegen

$$\Big(\frac{x-x_0}{z-x_0} \Big)^{k+1} - \Big(\frac{x-x_0}{z-x_0} \Big)^{k} \le 0$$

für alle $x \in [x_0, z]$ folgt für alle solche x und alle $m > n \ge n_0$

$$\begin{aligned} &\Big| \sum_{k=n+1}^{m} g_k \Big[\Big(\frac{x-x_0}{z-x_0} \Big)^{k+1} - \Big(\frac{x-x_0}{z-x_0} \Big)^{k} \Big] \Big| \\ &\le \sum_{k=n+1}^{m} |g_k| \Big[\Big(\frac{x-x_0}{z-x_0} \Big)^{k} - \Big(\frac{x-x_0}{z-x_0} \Big)^{k+1} \Big] \\ &\le \varepsilon \sum_{k=n+1}^{m} \Big[\Big(\frac{x-x_0}{z-x_0} \Big)^{k} - \Big(\frac{x-x_0}{z-x_0} \Big)^{k+1} \Big] \\ &= \varepsilon \Big[\Big(\frac{x-x_0}{z-x_0} \Big)^{n+1} - \Big(\frac{x-x_0}{z-x_0} \Big)^{m+1} \Big] \le \varepsilon \Big(\frac{x-x_0}{z-x_0} \Big)^{n+1} \le \varepsilon \,, \end{aligned}$$

folglich ist nach dem Cauchyschen Konvergenzkriterium auch die Reihe

$$\sum_{k=1}^{\infty} g_k \Big[\Big(\frac{x-x_0}{z-x_0} \Big)^{k+1} - \Big(\frac{x-x_0}{z-x_0} \Big)^{k} \Big]$$

gleichmäßig konvergent für x aus dem Intervall $[x_0, z]$. ■

Beispiel: $\sum_{n=1}^{\infty} \frac{1}{n} x^n$ hat Konvergenzradius $r = 1$. Da $\sum_{1=0}^{\infty} \frac{1}{n} (-1)^n$ konvergiert, ist die Funktion $x \mapsto \sum_{n=1}^{\infty} \frac{1}{n} x^n$ stetig in $[-1, 1)$.

8g.) Trigonometrische Funktionen

Sinus und Cosinus: Die Potenzreihen $\sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!}$ und $\sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!}$ haben Konvergenzradius $r = \infty$, konvergieren also auf der ganzen reellen Achse. Denn es ist

$$\sum_{n=0}^{\infty} \Big| \frac{x^{2n+1}}{(2n+1)!} \Big| \le \sum_{n=0}^{\infty} \frac{|x|^n}{n!} = \exp(|x|) \,,$$

also ist die Reihe zur Exponentialfunktion eine konvergente Majorante. Somit sind die Funktionen

$$\sin x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!}\,, \quad \cos x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!}$$

auf der ganzen reellen Achse definierte beliebig oft differenzierbare Funktionen. Es gilt

$$\sin' x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!} = \cos x$$

und

$$\cos' x = \sum_{n=1}^{\infty} (-1)^n \frac{x^{2n-1}}{(2n-1)!} = -\sin x\,,$$

und

$$\sin 0 = \cos' 0 = 0\,, \quad \sin' 0 = \cos 0 = 1\,.$$

Hieraus resultiert

$$\begin{aligned} \sin'' x + \sin x &= 0\,; \\ \sin 0 = 0\,, \quad \sin' 0 &= 1 \\ \cos'' x + \cos x &= 0 \\ \cos 0 = 1\,, \quad \cos' 0 &= 0\,. \end{aligned}$$

Satz: Seien $a, b \in \mathbb{R}$ und sei $f : \mathbb{R} \to \mathbb{R}$ eine zweimal differenzierbare Funktion mit $f''(x) + f(x) = 0$ für alle $x \in \mathbb{R}$, und mit $f(0) = a$, $f'(0) = b$. Dann gilt

$$f(x) = a \cos x + b \sin x\,.$$

Dies bedeutet, daß die „Differentialgleichung“ $f'' + f = 0$ eine eindeutig bestimmte Lösung hat, die auch noch die Anfangsbedingungen $f(0) = a$, $f'(0) = b$ erfüllt.

Beweis: Es ist klar, daß $g(x) = a \cos x + b \sin x$ die gewünschten Eigenschaften hat. Sei nun $h(x) = f(x) - g(x)$. Im Folgenden wird gezeigt, daß $h \equiv 0$ ist. Hierzu beachte man, daß für alle $x \in \mathbb{R}$

$$h''(x) + h(x) = f'' + f - g'' - g = 0\,,$$

und $h(0) = h'(0) = 0$ ist. Durch Multiplikation der Differentialgleichung mit $h'(x)$ folgt

$$h''(x)h'(x) + h(x)h'(x) = 0\,.$$

Wegen

$$\begin{aligned} \frac{d}{dx}\Big[h'(x)\Big]^2 &= 2h''(x)\,h'(x)\,, \\ \frac{d}{dx}\Big[h(x)\Big]^2 &= 2h(x)\,h'(x)\,, \end{aligned}$$

kann diese Gleichung in der Form

$$\frac{d}{dx}\Big\{\Big[h'(x)\Big]^2 + \Big[h(x)\Big]^2\Big\} = 0$$

geschrieben werden. Hieraus folgt, daß $[h'(x)]^2+[h(x)]^2$ konstant ist. Um diesen konstanten Wert zu berechnen setze man $x = 0$. Wegen $[h'(0)]^2 + [h(0)]^2 = 0$ folgt dann $[h'(x)]^2 + [h(x)]^2 = 0$, also $h'(x) = h(x) = 0$ für alle $x \in \mathbb{R}$. Dies ergibt

$$f(x) = g(x) = a\cos x + b\sin x\,.$$

Damit ist der Satz bewiesen. ∎

Mit diesem Satz kann das Additionstheorem für sin und cos bewiesen werden.

Satz: (Additionstheoreme) Für alle $x, y \in \mathbb{R}$ gilt

$$\begin{aligned} \sin(x+y) &= \sin x \cos y + \sin y \cos x \\ \cos(x+y) &= \cos x \cos y - \sin x \sin y\,. \end{aligned}$$

Beweis: Für die Funktion $x \mapsto f(x) := \sin(x+y)$ folgt

$$f''(x) + f(x) = 0\,, \quad \text{für alle } x \in \mathbb{R}\,,$$

$$f(0) = \sin y\,, \quad f'(0) = \cos y\,.$$

Also folgt aus dem vorangehenden Satz

$$\sin(x+y) = f(x) = \sin y \cos x + \cos y \sin x\,.$$

Ebenso zeigt man dies für $\cos(x+y)$. ∎

Aus diesen Additionstheoremen kann man einige weitere Formeln herleiten. Man beachte zunächst, daß sin eine ungerade Funktion ist, d.h. es gilt

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

Ebenso folgt, daß cos eine gerade Funktion ist:

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

Aus den Additionstheoremen erhält man also

$$\begin{aligned}\sin(x-y) &= \sin\Big(x+(-y)\Big) = \sin x\cos(-y) + \sin(-y)\cos x \\ &= \sin x\cos y - \sin y\cos x\end{aligned}$$

und

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

Für $x = y$ erhält man aus der letzten Formel wegen $\cos 0 = 1$

$$\cos^2 x + \sin^2 x = 1\,.$$

Hieraus folgt $|\cos x| \leq 1\,,\ \ |\sin x| \leq 1\,.$

Definition von π: Aus der Taylorformel folgt mit $\cos 0 = -\cos'' 0 = 1\,,$ $\cos' 0 = \cos^{(3)}(0) = 0\,,\ \ \cos^{(4)}(x) = \cos x\,,$ daß

$$\cos x = 1 - \frac{x^2}{2} + \frac{\cos(\vartheta x)}{24}\, x^4\,,$$

für geeignetes $\vartheta\,,\ \ 0 < \vartheta < 1\,.$ Für $x = 2$ ergibt sich hieraus wegen $|\cos(\vartheta x)| \leq 1\,,$ daß

$$\cos 2 = 1 - 2 + 16\,\frac{\cos(2\vartheta)}{24} = -1 + \frac{2}{3}\,\cos(2\vartheta) \leq -\frac{1}{3}\,.$$

Wegen $\cos 0 = 1$ muß also die Funktion cos nach dem Zwischenwertsatz eine Nullstelle zwischen 0 und 2 haben. Natürlich kann cos mehrere Nullstellen zwischen 0 und 2 haben. Das Infimum der Menge der Nullstellen von cos zwischen 0 und 2 werde $\frac{\pi}{2}$ genannt. Da cos stetig ist, ist $\frac{\pi}{2}$ selbst eine Nullstelle von cos . Zum Beweis wähle man eine Folge $\{x_n\}_{n=1}^{\infty}$ mit $\lim_{n\to\infty} x_n = \frac{\pi}{2}$ und mit $\cos x_n = 0\,.$ Es folgt dann

$$\cos\frac{\pi}{2} = \cos(\lim_{n\to\infty} x_n) = \lim_{n\to\infty}\cos(x_n) = 0\,.$$

Also ist π definiert als das Doppelte der kleinsten positiven Nullstelle von cos .

Aus $\cos^2 x + \sin^2 x = 1$ folgt

$$\left|\sin\frac{\pi}{2}\right| = 1\,.$$

Da der Cosinus zwischen 0 und $\frac{\pi}{2}$ positiv ist, folgt aus

$$\sin' x = \cos x\,,$$

daß sin streng monoton wachsend ist im Intervall $(0, \frac{\pi}{2})$. Wegen $\sin 0 = 0$ ergibt sich hieraus, daß $\sin \frac{\pi}{2} > 0$ sein muß, also

$$\sin \frac{\pi}{2} = 1 .$$

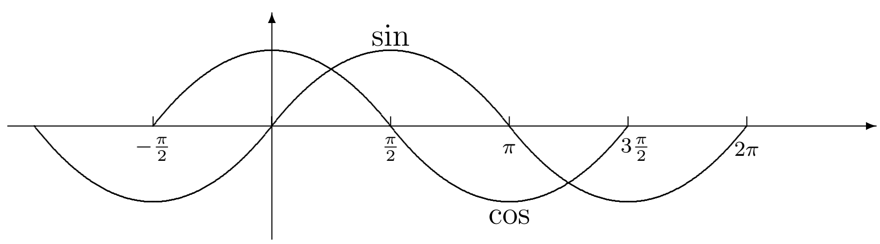
Aus dem Additionstheorem folgt nun

$$\sin\left(\frac{\pi}{2} + x\right) = \sin\frac{\pi}{2}\cos x + \cos\frac{\pi}{2}\sin x = \cos x ,$$

also

$$\sin\left(\frac{\pi}{2} - x\right) = \cos(-x) = \cos x = \sin\left(\frac{\pi}{2} + x\right). \qquad (*)$$

Diese Formel liefert die Werte von sin im Intervall $[\frac{\pi}{2}, \pi]$ aus den Werten im Intervall $[0, \frac{\pi}{2}]$.



Aus $\sin(-x) = -\sin x$ ergeben sich dann die Werte im Intervall $[-\pi, \pi]$. Die weitere Fortsetzung erhält man aus der Periodizität:

$$\sin(x + 2\pi) = \sin x .$$

Diese Formel erhält man folgendermaßen: Aus $(*)$ ergibt sich

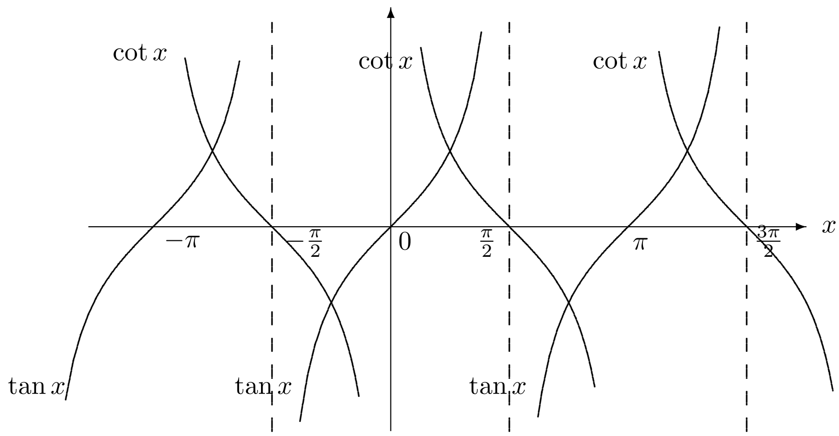
$$\sin(x + \pi) = \sin\left(\frac{\pi}{2} + \frac{\pi}{2} + x\right) = \sin\left(\frac{\pi}{2} - (\frac{\pi}{2} + x)\right) = \sin(-x) = -\sin x .$$

Also

$$\sin(x + 2\pi) = \sin(x + \pi + \pi) = -\sin(x + \pi) = \sin x .$$

Tangens und Cotangens: Man definiert

$$\tan x = \frac{\sin x}{\cos x}, \quad \cot x = \frac{\cos x}{\sin x} = \frac{1}{\tan x} .$$



Aus den Additionstheoremen für Sinus und Cosinus ergeben sich Additionstheoreme für Tangens und Cotanges:

$$\begin{aligned}\tan(x+y) &= \frac{\tan x + \tan y}{1 - \tan x \tan y}\\ \cot(x+y) &= \frac{\cot x \cot y - 1}{\cot x + \cot y}\end{aligned}$$

Als Ableitung erhält man:

$$\begin{aligned}\tan' x = \left(\frac{\sin x}{\cos x}\right)' &= \frac{\cos^2 x + \sin^2 x}{\cos^2 x} = \frac{1}{\cos^2 x}\\ \cot' x = \left(\frac{\cos x}{\sin x}\right)' &= \frac{-\sin^2 x - \cos^2 x}{\sin^2 x} = \frac{-1}{\sin^2 x}\,.\end{aligned}$$

Arcus–Funktionen: sin und cos sind periodisch, also nicht injektiv, folglich existieren zu sin und cos keine Umkehrfunktionen. Wenn man aber sin und cos auf geeignete Intervalle einschränkt, existieren Umkehrfunktionen.

Es ist $\cos x > 0$ für $x \in (-\frac{\pi}{2}, \frac{\pi}{2})$, nach Definition von π. Wegen $\sin' x = \cos x$ ist also sin streng monoton wachsend im Intervall $[-\frac{\pi}{2}, \frac{\pi}{2}]$, also injektiv, also existiert zu $\sin : [-\frac{\pi}{2}, \frac{\pi}{2}] \to [-1, 1]$ die Umkehrfunktion, ebenso existieren die Umkehrfunktionen in den Intervallen

$$\sin \left[\frac{\pi}{2} + n\pi, \frac{\pi}{2} + (n+1)\pi\right] \to [-1, 1]\,, \quad n \in \mathbb{Z}\,.$$

Man muß daher immer genau angeben, welche dieser unendlich vielen Umkehrfunktionen man meint. Ist nichts weiter gesagt, meint man die Umkehrfunktion

$$\arcsin : [-1, 1] \to \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$$

zu $\sin : [-\frac{\pi}{2}, \frac{\pi}{2}] \to [-1, 1], .$ (Arcus Sinus).

Aus Gründen, die in der Funktionentheorie ihren Ursprung haben, bezeichnet man die unendlich vielen verschiedenen Umkehrfunktionen

$$x \mapsto (\arcsin x) + 2n\pi\,, \quad n \in \mathbb{Z}\,,$$

und

$$x \mapsto -(\arcsin x) + (2n+1)\pi\,, \quad n \in \mathbb{Z}\,,$$

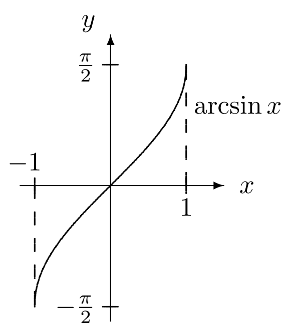
auch als die Zweige der Umkehrfunktion von Sinus. Die Funktion $\arcsin : [-1,1] \to [-\frac{\pi}{2}, \frac{\pi}{2}]$ heißt dann Hauptwert der Umkehrfunktion.

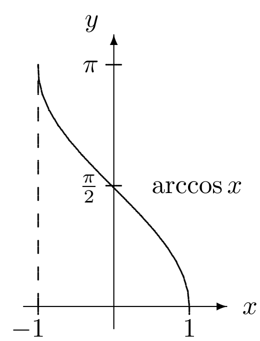
Entsprechend bezeichnet man die Umkehrfunktion

$$\arccos : [-1,1] \to [0,\pi]\,, \qquad \text{(Arcus Cosinus)}$$

zur Funktion $\cos : [0,\pi] \to [-1,1]$ als Hauptwert der Umkehrfunktion zu Cosinus. Es existieren aber die unendlich vielen weiteren Umkehrfunktionen

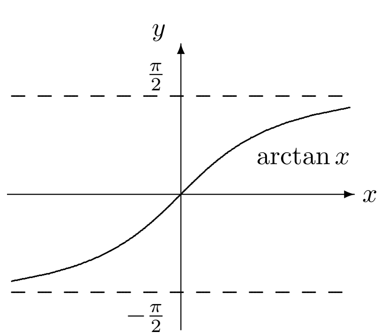
$$x \mapsto \pm(\arccos x) + 2n\pi\,, \quad n \in \mathbb{Z}\,.$$





Ähnlich liegen die Verhältnisse bei tan und cot. Als Hauptwert zur Umkehrfunktion von tan bezeichnet man die Funktion

$$\arctan : [-\infty,\infty] \to \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]. \qquad \text{(Arcus Tangens)}$$



Ich betrachte im Folgenden die Hauptwerte der Arcusfunktionen. Für die Ableitungen erhält man:

$$\begin{aligned}
(\arcsin x)' &= \frac{1}{\sin'(\arcsin x)} = \frac{1}{\cos(\arcsin x)} \\
&= \frac{1}{\sqrt{1-[\sin(\arcsin x)]^2}} = \frac{1}{\sqrt{1-x^2}} \\
(\arccos x)' &= \frac{1}{\cos'(\arccos x)} = \frac{-1}{\sin(\arccos x)} \\
&= \frac{-1}{\sqrt{1-[\cos(\arccos x)]^2}} = \frac{-1}{\sqrt{1-x^2}} \\
\arctan' x &= \frac{1}{\tan'(\arctan x)} = \Big[\cos(\arctan x)\Big]^2 \\
&= \frac{1}{1+[\tan(\arctan x)]^2} = \frac{1}{1+x^2}\,.
\end{aligned}$$

Man kann die Funktionen arcsin, arccos und arctan in Potenzreihen entwickeln. Zum Beispiel gilt

$$\frac{d}{dt}(\arctan x) = \frac{1}{1+x^2} = \sum_{n=0}^{\infty}(-1)^n x^{2n}\,,$$

für $|x| < 1$. Auch die Potenzreihe

$$\sum_{n=0}^{\infty}\frac{(-1)^n}{2n+1}\,x^{2n+1}$$

hat den Konvergenzradius 1, und ist Stammfunktion zu $\sum_{n=0}^{\infty}(-1)^n x^{2n}$, also gilt

$$\arctan x = \sum_{n=0}^{\infty}\frac{(-1)^n}{2n+1}\,x^{2n+1} + C$$

in $|x| < 1$, mit einer geeigneten Konstanten C. Wegen $\arctan 0 = 0$ folgt $C = 0$, also folgt

$$\arctan x = \sum_{n=0}^{\infty}\frac{(-1)^n}{2n+1}\,x^{2n+1}$$

für $x \in \mathbb{R}$ mit $|x| < 1$. Nach dem Leibnizkriterium ist die rechtsstehende Reihe für $x = 1$ noch konvergent, also definiert die Potenzreihe eine stetige Funktion im Punkt $x = 1$. Da auch arctan stetig ist, sind durch die Potenzreihe und durch die Funktion arctan zwei stetige Fortsetzungen der Funktion arctan vom Intervall $(-1, 1)$ auf $(-1, 1]$ gegeben. Da die stetige Fortsetzung eindeutig ist, muß also gelten

$$\arctan 1 = \sum_{n=0}^{\infty}\frac{(-1)^n}{2n+1}\,.$$

Wegen

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

folgt

$$0 = 2\Big(\cos\frac{\pi}{4}\Big)^2 - 1\,,$$

also

$$\cos\frac{\pi}{4} = \sqrt{\frac{1}{2}}\,,$$

und

$$\sin\frac{\pi}{4} = \sqrt{1 - \Big(\cos\frac{\pi}{4}\Big)^2} = \sqrt{\frac{1}{2}}\,,$$

somit

$$\tan\frac{\pi}{4} = \frac{\sin\frac{\pi}{4}}{\cos\frac{\pi}{4}} = 1\,.$$

Dies ergibt $\arctan 1 = \frac{\pi}{4}$, d.h.

$$\frac{\pi}{4} = \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1}\,.$$

Theoretisch kann aus dieser Reihe π berechnet werden, die Konvergenz ist aber sehr langsam.

9 Komplexe Zahlen (Ausblick)

Es gibt keine relle Zahl x, die die Gleichung

$$x^2 + 1 = 0$$

löst. Denn sonst müßte $x^2 = -1 < 0$ sein, was unmöglich ist, weil aus den Anordnungsaxiomen $x^2 \geq 0$ folgt für jede reelle Zahl x.

Wenn man erreichen will, daß diese Gleichung lösbar ist, muß man zur Menge $\mathbb{R}$ der reellen Zahlen also noch weitere Elemente hinzunehmen, also muß man zu einer Obermenge $\mathbb{C}$ der reellen Zahlen $\mathbb{R}$ übergehen. Man kann nicht erreichen, daß in $\mathbb{C}$ alle Axiome (A 1) – (A 14) für die reellen Zahlen erfüllt sind, weil durch diese Axiome $\mathbb{R}$ bis auf Isomorphie eindeutig bestimmt ist. Man möchte aber erreichen, daß möglichst viele dieser Axiome erfüllt sind. Insbesondere möchte man natürlich, daß $\mathbb{C}$ die Körperaxiome erfüllt, daß man in $\mathbb{C}$ also addieren, subtrahieren, multiplizieren und dividieren kann wie in $\mathbb{R}$. Da $\mathbb{C}$ eine Obermenge von $\mathbb{R}$ sein soll möchte man auch, daß die in $\mathbb{C}$ neu definierte Addition und Multiplikation bei Einschränkung auf die Teilmenge $\mathbb{R}$ mit der üblichen Addition und Multiplikation übereinstimmt. Außerdem muß es in $\mathbb{C}$ ein Element i geben mit

$$i^2 = -1\,.$$

Diese Forderungen werden erfüllt, wenn man $\mathbb{C}$ folgendermaßen konstruiert:

Es sei $\mathbb{C}$ die Menge der geordneten Paare (a, b) mit $a, b \in \mathbb{R}$, also $\mathbb{C} = \mathbb{R} \times \mathbb{R}$. Auf $\mathbb{C}$ erkläre man Addition und Multiplikation folgendermaßen:

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

Satz: $(\mathbb{C}, +, \cdot)$ ist ein kommutativer Körper, der einen zu $\mathbb{R}$ isomorphen Unterkörper enthält. In $\mathbb{C}$ ist die Gleichung $z^2 + 1 = 0$ lösbar.

Beweis: $(\mathbb{C}, +)$ ist eine kommutative Gruppe: Denn es gilt

$$\begin{aligned} (a,b) + (a',b') &= (a',b') + (a,b)\,, \\ (a,b) + \Big((c,d) + (e,f)\Big) &= \Big((a,b) + (c,d)\Big) + (e,f)\,. \end{aligned}$$

Das eindeutig bestimmte neutrale Element ist $(0,0)$. Die Gleichung $(a,b) + (x,y) = 0$ ist eindeutig lösbar in $\mathbb{C}$. Die Lösung ist $(x,y) = (-a,-b)$.

Durch nachrechnen sieht man, daß auch für die Multiplikation das Kommutativgesetz und das Assoziativgesetz erfüllt sind, und daß das Distributivgesetz gilt.

Für die Multiplikation ist $(1,0)$ das eindeutig bestimmte neutrale Element. Denn es gilt

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

Außerdem hat für jedes $(a,b) \neq 0$ die Gleichung

$$(a,b)\cdot(x,y) = (1,0)$$

eine eindeutig bestimmte Lösung. Denn (x,y) ist Lösung dieser Gleichung genau dann wenn

$$(ax - by\,,\ \ bx + ay) = (1,0)\,,$$

gilt, d. h. genau dann wenn x, y Lösung sind des linearen Gleichungssystems

$$\begin{aligned} ax - by &= 1 \\ bx + ay &= 0\,, \end{aligned}$$

oder

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Wegen

$$\det \begin{pmatrix} a & -b \\ b & a \end{pmatrix} = a^2 + b^2 > 0$$

ist dieses Gleichungssystem eindeutig lösbar, die Lösung ist

$$x = \frac{a}{a^2+b^2}\,, \quad y = \frac{-b}{a^2+b^2}\,.$$

Also ist $(\mathbb{C}, +, \cdot)$ ein kommutativer Körper.
Die Abbildung

$$x \mapsto I(x) = (x,0) : \mathbb{R} \to \mathbb{C}$$

ist injektiv und ein Homomorphismus, d. h. es gilt

$$\begin{aligned} I(x+y) &= (x+y,0) &&= (x,0)+(y,0) &&= I(x)+I(y) \\ I(x\cdot y) &= (x\cdot y,0) &&= (x,0)\cdot(y,0) &&= I(x)\cdot I(y)\,, \end{aligned}$$

also ist $\mathbb{R}$ isomorph zu dem Unterkörper $\{(x,0) \mid x \in \mathbb{R}\}$ von $\mathbb{C}$. Das Element $z = (0,1)$ ist Lösung der Gleichung

$$z^2 + (1,0) = 0\,.$$

Damit ist der Satz bewiesen. ■

Man bezeichnet $\mathbb{C}$ als den Körper der komplexen Zahlen.

Zur Abkürzung setzt man

$$1 := (1,0)\,, \quad i := (0,1)\,.$$

Man kann dann jede komplexe Zahl $z = (a,b)$ in der Form

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

schreiben. Daher bezeichnet man a als Realteil und b als Imaginärteil von z, und man schreibt auch

$$\operatorname{Re} z = a\,, \quad \operatorname{Im} z = b\,.$$

Es ist nicht möglich auf dem Körper $\mathbb{C}$ eine Anordnung $<$ zu erklären, die die im Kapitel 2 angegebenen Anordnungsaxiome (A 10) – (A 13) erfüllt. Denn es wurde gezeigt, daß aus den Körperaxiomen und Anordnungsaxiomen für das neutrale Element der Multiplikation 1 und für jedes Element des Körpers z immer folgt

$$-1 < 0\,, \quad z^2 \geq 0\,.$$

Dies steht im Widerspruch zur Gleichung

$$i^2 = -1\,,$$

die in $\mathbb{C}$ erfüllt ist. Man kann jedoch als Ersatz den Absolutbetrag von $\mathbb{R}$ auf $\mathbb{C}$ fortsetzen. Zunächst definiert man: Sei $z = a + ib \in \mathbb{C}$. Dann heißt

$$\overline{z} = a - ib$$

die zu z konjugiert komplexe Zahl. Natürlich gilt $\overline{\overline{z}} = z$. Es gilt auch

$$\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}\,, \quad \overline{z_1 z_2} = \overline{z_1}\,\overline{z_2}\,.$$

Definition: Die Zahl $|x| := \sqrt{z \cdot \overline{z}}$ heißt absoluter Betrag von z.

Beachte, daß $|z| = \sqrt{a^2 + b^2} \geq 0$ ist für alle $z = a + ib$.

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

1.) $|z| = 0 \iff z = 0$

2.) $|zw| = |z|\,|w|$

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

Beweis:

1.) $|z| = 0 \iff \sqrt{a^2 + b^2} = 0 \iff a = 0 \quad \text{und} \quad b = 0\,.$

2.) $|zw| = \sqrt{zw \cdot \overline{zw}} = \sqrt{z\overline{z} \cdot w\overline{w}} = \sqrt{z\overline{z}}\,\sqrt{w\overline{w}} = |z| \cdot |w|$

3.) Es gilt

$$|z + w|^2 = (z + w)(\overline{z} + \overline{w}) = z\overline{z} + z\overline{w} + w\overline{z} + w\overline{w} = |z|^2 + z\overline{w} + w\overline{z} + |w|^2\,.$$

Für die beiden mittleren Terme folgt

$$z\overline{w} + \overline{z}w = z\overline{w} + \overline{z\overline{w}} = 2\mathrm{Re}\,(z\overline{w}) \leq 2|z\overline{w}| = 2|z|\,|\overline{w}| = 2|z|\,|w|\,,$$

also

$$|z + w|^2 \leq |z|^2 + 2|z|\,|w| + |w|^2 = (|z| + |w|)^2\,,$$

somit

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

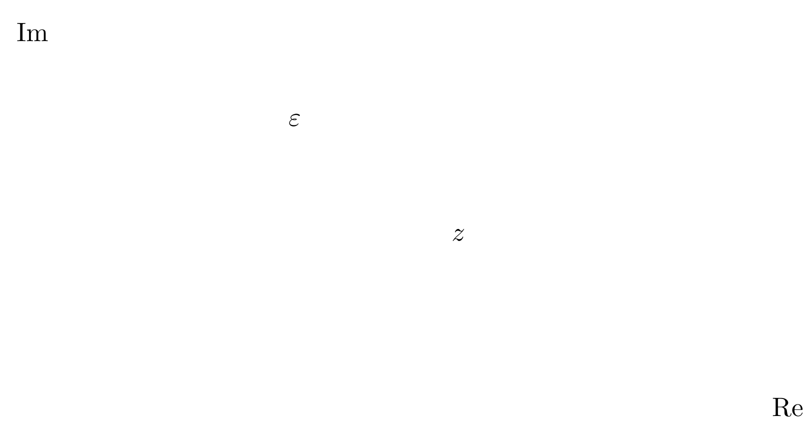
■

Definition: Eine Folge $\{z_n\}_{n=1}^{\infty}$ komplexer Zahlen heißt konvergent, wenn $z \in \mathbb{C}$ existiert, so daß zu jedem $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$ existiert mit

$$|z_n - z| < \varepsilon$$

für alle $n \geq n_0$. Die Zahl z heißt Grenzwert der Folge.

Dies bedeutet, daß für $n \geq n_0$ alle Glieder z_n innerhalb eines Kreises um z mit Radius ε liegen müssen.



Satz: (i) Eine Folge $\{z_n\}_{n=1}^{\infty} = \{x_n + iy_n\}_{n=1}^{\infty}$ ist genau dann konvergent mit dem Grenzwert $z = x + iy$, wenn in $\mathbb{R}$

$$\lim_{n\to\infty} x_n = x\,, \quad \lim_{n\to\infty} y_n = y$$

gilt.
(ii) Eine Folge $\{z_n\}_{n=1}^{\infty}$ komplexer Zahlen $z_n = x_n + iy_n$ ist genau dann Cauchyfolge, wenn $\{x_n\}_{n=1}^{\infty}$ und $\{y_n\}_{n=1}^{\infty}$ Cauchyfolgen in $\mathbb{R}$ sind.

Der Beweis bleibt dem Leser überlassen.

Folgerung: Eine Folge $\{z_n\}_{n=1}^{\infty}$ komplexer Zahlen ist genau dann konvergent, wenn sie Cauchyfolge ist.

Definition: Eine Reihe $\sum_{n=0}^{\infty} z_n$ mit $z_n \in \mathbb{C}$ heißt absolut konvergent, wenn $\sum_{n=0}^{\infty} |z_n|$ in $\mathbb{R}$ konvergiert.

Satz: Jede absolut konvergente Reihe ist konvergent.

Beweis: Da $\sum_{n=0}^{\infty} |z_n|$ konvergiert, existiert zu $\varepsilon > 0$ eine Zahl n_0, so daß für alle $m \geq n \geq n_0$ gilt

$$|\sum_{\ell=n}^{m} z_n| \leq \sum_{\ell=n}^{m} |z_n| < \varepsilon\,,$$

also ist $\sum_{n=0}^{\infty} z_n$ konvergent nach dem Cauchyschen Konvergenzkriterium. ■

Folgerung: Seien $z_0, a_n \in \mathbb{C}$. Die Potenzreihe

$$\sum_{n=0}^{\infty} a_n(z - z_0)^n$$

konvergiert absolut für $z \in \mathbb{C}$ mit

$$|z - z_0| < r = \frac{1}{\overline{\lim}_{n\to\infty} \sqrt[n]{|a_n|}}\,.$$

r heißt Konvergenzradius und $\{z \in \mathbb{C} \,\big|\, |z - z_0| < r\}$ Konvergenzkreis.

Beweis: Für $|z - z_0| < r$ ist

$$\overline{\lim}_{n\to\infty} \sqrt[n]{|a_n|\,|z - z_0|^n} = |z - z_0|\,\overline{\lim}_{n\to\infty} \sqrt[n]{|a_n|} < 1\,,$$

also konvergiert $\sum_{n=0}^{\infty} a_n(z - z_0)^n$ absolut. ■

Folgerung: Die Reihen

$$\sum_{n=0}^{\infty} \frac{z^n}{n!} \;, \quad \sum_{n=0}^{\infty} (-1)^n \frac{z^{2n+1}}{(2n+1)!} \;, \quad \sum_{n=0}^{\infty} (-1)^n \frac{z^{2n}}{(2n)!}$$

konvergieren für alle $z \in \mathbb{C}$. Also können die Exponentialfunktion

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

der Sinus

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

und der Cosinus

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

auf ganz $\mathbb{C}$ erklärt werden. Die Reihe $\sum_{n=1}^{\infty} (-1)^{n-1} \frac{(z-1)^n}{n}$ konvergiert für alle $z \in \mathbb{C}$ mit $|z| < 1$, also ist

$$\log z := \sum_{n=1}^{\infty} (-1)^{n-1} \frac{(z-1)^n}{n}$$

im ganzen Kreis $|z-1| < 1$ erklärt.

Im

$|z-1| < 1$

1 Re

Die Sätze über das Cauchyprodukt absolut konvergenter Reihen gelten natürlich auch für komplexe Reihen. Also gilt

$$\exp(z_1 + z_2) = \exp(z_1) \exp(z_2)$$

für alle $z_1, z_2 \in \mathbb{C}$.

Sei nun $x \in \mathbb{R}$. Dann gilt

$$\begin{aligned}
e^{ix} = \sum_{n=0}^{\infty} \frac{(ix)^n}{n!} &= \sum_{n=0}^{\infty} i^n \frac{x^n}{n!} \\
&= \sum_{n=0}^{\infty} i^{2n} \frac{x^{2n}}{(2n)!} + \sum_{n=0}^{\infty} i^{2n+1} \frac{x^{2n+1}}{(2n+1)!} \\
&= \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!} + i \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!} \\
&= \cos x + i \sin x \,.
\end{aligned}$$

Also folgt für $z = x + iy\,,\ \ x, y \in \mathbb{R}$:

$$e^z = e^{x+iy} = e^x\, e^{iy} = e^x (\cos y + i \sin y)\,,$$

folglich

$$\begin{aligned}
|e^z| &= |e^x(\cos y + i \sin y)| = |e^x|\, |\cos y + i \sin y| \\
&= e^x \sqrt{\cos^2 y + \sin^2 y} = e^x \,.
\end{aligned}$$

Also liegt e^z auf dem Kreis mit Radius e^x um den Nullpunkt:

Im

$e^{x+\frac{\pi}{2}i}$

e^x

$e^{x+\pi i}$ Re

$e^{x+\frac{3}{2}\pi i}$

Insbesondere folgt:

$$\begin{aligned}
e^{\frac{\pi}{2}i} &= \cos\frac{\pi}{2} + i \sin\frac{\pi}{2} = i\,, \\
e^{\pi i} &= \cos\pi + i \sin\pi = -1\,, \\
e^{\frac{3}{2}\pi i} &= \cos\frac{3}{2}\pi + i \sin\frac{3}{2}\pi = -i\,, \\
e^{2\pi i} &= \cos 2\pi + i \sin 2\pi = 1\,.
\end{aligned}$$

Schließlich gilt für $x, y \in \mathbb{R}$:

$$\begin{aligned} e^{i(x+y)} &= e^{ix}e^{iy} \\ &= (\cos x + i \sin x)(\cos y + i \sin y) \\ &= [\cos x \cos y - \sin x \sin y] + i[\sin x \cos y + \cos x \sin y] , \end{aligned}$$

somit

$$\begin{aligned} \sin(x+y) &= \operatorname{Im} e^{i(x+y)} = \sin x \cos y + \cos x \sin y \\ \cos(x+y) &= \operatorname{Re} e^{i(x+y)} = \cos x \cos y - \sin x \sin y , \end{aligned}$$

d. h. die Additionstheoreme für sin und cos sind eine Folgerung aus dem Additionstheorem für die Exponentialfunktion.