Ralf Gerkmann

Mathematisches Institut der
Ludwig-Maximilians-Universität München

Wintersemester 2024-25

Analysis einer Variablen

Inhaltsverzeichnis

§ 1. Definition der Zahlbereiche *3*
§ 2. Konvergenz von Folgen *30*
§ 3. Häufungspunkte und Cauchyfolgen *40*
§ 4. Konvergenzkriterien für Reihen *48*
§ 5. Funktionsgrenzwerte und Stetigkeit *67*
§ 6. Potenzreihen, Exponentialfunktion, Sinus und Kosinus *82*
§ 7. Differenzierbarkeit und Ableitungsregeln *98*
§ 8. Mittelwertsatz und Extremwertbestimmung *108*
§ 9. Das Riemann-Integral *118*

Literaturverzeichnis *142*

§ 1. *Definition der Zahlbereiche*

Inhaltsübersicht

Zunächst werden wir in diesem Kapitel die axiomatische Charakterisierung der reellen Zahlen zum Abschluss bringen. Dabei setzen wir die Körperaxiome aus § 5 der Linearen Algebra als bekannt voraus. Als weitere Eigenschaft kommt bei den reellen Zahlen die Anordnung hinzu, die auf der Menge $\mathbb{R}$ eine Totalordnung definiert. Darüber hinaus wird es für die Analysis eine wichtige Rolle spielen, dass dieser angeordnete Körper vollständig ist. Wir werden sehen, dass sich die Vollständigkeit auf unterschiedliche Weise anschaulich beschreiben lässt: durch das Konzept der Dedekindschen Schnitte, durch die aus dem Kapitel § 3 der Linearen Algebra bekannten Infima und Suprema, sowie durch die Intervallschachtelungs-Eigenschaft und das archimedische Axiom.

Im weiteren Verlauf werden wir die Zahlbereiche $\mathbb{N}$, $\mathbb{N}_0$, $\mathbb{Z}$, $\mathbb{Q}$ und $\mathbb{R}$ als Teilmengen von $\mathbb{R}$ definieren. Wir werden sehen, dass es sich bei $\mathbb{Z}$ um einen Ring und bei $\mathbb{Q}$ um einen Körper handelt, und das letzterer eine dichte Teilmenge von $\mathbb{R}$ bildet. Am Ende des Kapitels werden wir sehen, dass $\mathbb{R}$ seinerseits in einem noch größeren Körper enthalten ist, dem Körper $\mathbb{C}$ der komplexen Zahlen, bei dem es sich nicht um einen angeordneten, aber um einen sog. bewerteten Körper handelt.

Wichtige Begriffe und Sätze

- Anordnung auf einem Körper, angeordneter Körper
- Intervall in einem angeordneten Körper
- archimedische Anordnung
- Dedekindscher Schnitt
- vollständig angeordneter Körper
- Intervallschachtelungs-Eigenschaft
- Definition von $\mathbb{R}$ als vollständig angeordneter Körper
- induktive Teilmenge von $\mathbb{R}$
- Definition von $\mathbb{N}$, $\mathbb{N}_0$, $\mathbb{Z}$ und $\mathbb{Q}$ als Teilmengen von $\mathbb{R}$
- dichte Teilmenge von $\mathbb{R}$ ($\mathbb{Q}$ und $\mathbb{R} \setminus \mathbb{Q}$ als Beispiele)
- bewerteter Körper
- Definition des Körpers $\mathbb{C}$ der komplexen Zahlen

Die Begriffe ***Halb-*** und ***Totalordnung*** haben wir bereits in der Linearen Algebra kennengelernt. Auf einem Körper K sollte eine Totalordnung $\leq$ natürlich nicht völlig willkürlich definiert werden, sondern auf eine Art und Weise, die mit den arithmetischen Operationen $+$ und $\cdot$ des Körpers zuammenhängt. Dies gelingt mit Hilfe des folgenden Begriffs.

(1.1) Definition Sei K ein Körper. Eine Teilmenge $K^+ \subseteq K$ wird ***Anordnung*** von K genannt, wenn die folgenden Bedingungen erfüllt sind.

(i) Für jedes $x \in K$ gilt genau eine der drei Aussagen $x = 0_K$, $x \in K^+$, $-x \in K^+$.
(ii) Aus $x, y \in K^+$ folgt $x + y \in K^+$ und $xy \in K^+$.

Das Paar (K, K^+) bezeichnet man als ***angeordneten Körper***. Ein Element $x \in K$ wird als ***positiv*** bezeichnet, wenn $x \in K^+$, und ***negativ***, wenn $-x \in K^+$ gilt.

Für die spätere Verwendung definieren wir die Bezeichnungen

$$K_+ = K^+ \cup \{0_K\} \qquad \text{und} \qquad K^\times = K \setminus \{0_K\}.$$

Wie wir gleich sehen werden, lässt sich mit Hilfe der Teilmenge K^+ eine Totalordnung auf der Menge K definieren. Zunächst aber zeigen wir

(1.2) Lemma Sei (K, K^+) ein angeordneter Körper.

(i) Für alle $a \in K^\times$ gilt $a^2 \in K^+$.
(ii) Aus $a \in K^+$ folgt jeweils $a^{-1} \in K^+$.

Beweis: zu (i) Nach Bedingung (i) in Definition (1.1) gilt ür jedes $a \in K^\times$ jeweils $a \in K^+$ oder $-a \in K^+$. Aus Bedingung (ii) folgt im ersten Fall $a^2 = a \cdot a \in K^+$. Im zweiten Fall wenden wir dieselbe Bedingung auf $-a$ an, und erhalten mit Hilfe der bekannten Rechenregeln in Ringen $a^2 = a \cdot a = (-a)(-a) \in K^+$.

zu (ii) Das Element a^{-1} besitzt wegen $(a^{-1})^{-1} = a$ einen Kehrwert, ist also ungleich 0_K. Nach (i) folgt damit $(a^{-1})^2 \in K^+$. Mit $(a^{-1})^2$ und a liegt auch das Produkt $(a^{-1})^2 a$ in K^+, und wegen $a^{-1} = a^{-1} \cdot 1_K = a^{-1}(a^{-1}a) = (a^{-1}a^{-1})a = (a^{-1})^2 a$ gilt damit auch $a^{-1} \in K^+$. □

(1.3) Satz Sei (K, K^+) ein angeordneter Körper.

(i) Durch $x \le y \Leftrightarrow y - x \in K_+$ ist eine ***Totalordnung*** auf K definiert. Für alle $x, y \in K$ gilt $x < y \Leftrightarrow y - x \in K^+$.

Seien nun $x, y, a \in K$ vorgegeben.

(ii) Es gilt $x \le y \Leftrightarrow x + a \le y + a$.
(iii) Ist a positiv, dann gilt $x \le y \Leftrightarrow ax \le ay$.
(iv) Ist a negativ, dann gilt $x \le y \Leftrightarrow ax \ge ay$.

Die drei Aussagen (ii), (iii) und (iv) bleiben gültig, wenn man $\le$ durch $<$ und $\ge$ durch $>$ ersetzt.

Beweis: zu (i) Als erstes beweisen wir die Zusatzaussage. Nach Definition ist für $x, y \in K$ die Aussage $x < y$ äquivalent zu $(x \leq y) \wedge (x \neq y)$. Dies ist äquivalent zu $(y - x \in K_+) \wedge (y - x \neq 0_K)$, und da sich K_+ disjunkt in $K^+ \cup \{0_K\}$ zerlegen lässt, ist dies wiederum gleichbedeutend mit $y - x \in K^+$.

Beweisen wir nun, dass $\leq$ eine Halbordnung ist, und beginnen wir mit der Reflexivität. Sei $x \in K$ vorgegeben. Zu zeigen ist $x \leq x$, was nach Definition gleichbedeutend mit $x - x \in K_+$ ist. Aber Letzteres ist offensichtlich wahr, denn es gilt $x - x = 0_K$ und $0_K \in K_+$. Zum Beweis der Antisymmetrie seien $x, y \in K$ mit $x \leq y$ und $y \leq x$ vorgegeben. Zu zeigen ist $x = y$. Nehmen wir an, es gilt $x \neq y$. Dann wäre zugleich $x < y$ und $y < x$ erfüllt. Wie bereits gezeigt, hätte dies zur Folge, dass $y - x$ und $x - y$ beide in K^+ liegen, wegen $x - y = -(y - x)$ also $y - x \in K^+$ und $-(y - x) \in K^+$. Aber dies widerspricht Axiom (i) in Definition (1.1). Also ist nur $x = y$ möglich.

Wenden wir uns nun der Transitivität zu und setzen dafür $x \leq y$ und $y \leq z$ voraus. Zu zeigen ist $x \leq z$. Im Fall $x = y$ erhalten wir direkt $x = y \leq z$, und aus $y = z$ folgt genauso $x \leq y = z$. Wir können uns also auf den Fall beschränken, dass $x \neq y$ und $y \neq z$ gilt. Dann folgt aus unserer Voraussetzung $x < y$ und $y < z$, und es folgt $y - x, z - y \in K^+$. Axiom (ii) in Definition (1.1) zeigt uns, dass wegen

$$(z - y) + (y - x) \quad = \quad z + (-y) + y + (-x) \quad = \quad z + (-x) \quad = \quad z - x$$

dann auch das Element $z - x$ in K^+ liegt. Daraus folgt $x < z$, insbesondere $x \leq z$.

Um nachzuweisen, dass $\leq$ nicht nur eine Halb-, sondern eine Totalordnung ist, seien $x, y \in K$ vorgegeben. Zu zeigen ist, dass $x \leq y$ oder $y \leq x$ gilt. Ist $x = y$, dann ist $x \leq y$ nach Definition erfüllt. Im Fall $x \neq y$ gilt dagegen $y - x \neq 0_K$, und aus Axiom (i) in Definition (1.1) folgt dann, dass entweder $y - x$ oder $-(y - x) = x - y$ ein Element von K^+ ist. Im ersten Fall gilt $x < y$ und somit $x \leq y$, im zweiten $y < x$ und insbesondere $y \leq x$.

zu (ii) Wir beweisen zunächst die Äquivalenz $x < y \Leftrightarrow x + a < y + a$. Wie eingangs bemerkt, ist $x < y$ gleichbedeutend mit $y - x \in K^+$. Wegen $(y + a) - (x + a) = y + a + (-x) + (-a) = y + (-x) = y - x$ ist dies äquivalent zu dass $(y + a) - (x + a) \in K^+$, also zu $x + a < y + a$. Dieselbe Aussage für $\leq$ erhält man nun durch die Äquivalenzumformung

$$x \leq y \quad \Leftrightarrow \quad (x < y) \vee (x = y) \quad \Leftrightarrow \quad (x + a < y + a) \vee (x + a = y + a) \quad \Leftrightarrow \quad x + a \leq y + a.$$

zu (iii) Auch hier betrachten wir zunächst die Äquivalenz $x < y \Leftrightarrow ax < ay$. „$\Rightarrow$“ Aus $x < y$ folgt $y - x \in K^+$. Wegen $a \in K^+$ liegt damit auch auch $ay - ax = a(y - x)$ in K^+, also gilt $ax < ay$. „$\Leftarrow$“ Da nach Lemma (1.2) (ii) mit a auch a^{-1} positiv ist, können wir die bereits bewiesene Schlussrichtung anwenden und erhalten $ax < ay \Rightarrow a^{-1}ax < a^{-1}ay \Rightarrow x < y$. Damit ist die Äquivalenz für die Relation $<$ insgesamt bewiesen.

Um auch die Äquivalenz für die Relation $\leq$ zu zeigen, bemerken wir zunächst $x = y \Rightarrow ax = ay$ und $ax = ay \Rightarrow a^{-1}ax = a^{-1}ay \Rightarrow x = y$, insgesamt also $x = y \Leftrightarrow ax = ay$. Aus den beiden Äquivalenzen $x < y \Leftrightarrow ax < ay$ und $x = y \Leftrightarrow ax = ay$ erhält man rein logisch die Äquivalenz $x \leq y \Leftrightarrow ax \leq ay$.

zu (iv) Hier zeigen wir zunächst die Äquivalenz $x < y \Leftrightarrow -x > -y$. Diese erhält man durch die Umformungen

$$x < y \quad \Leftrightarrow \quad y - x \in K^+ \quad \Leftrightarrow \quad (-x) - (-y) \in K^+ \quad \Leftrightarrow \quad -x > -y.$$

Wegen $x = y \Leftrightarrow -x = -y$ folgt daraus auch unmittelbar $x \leq y \Leftrightarrow -x \geq -y$. Ist nun a ein beliebiges negatives Element, dann ist $-a$ positiv. Durch Anwendung von (iii) und der soeben bewiesenen Äquivalenz erhalten wir $x < y \Leftrightarrow (-a)x < (-a)y \Leftrightarrow -ax < -ay \Leftrightarrow ax > ay$, und nach demselben Schema beweist man schließlich auch die Äquivalenz $x \leq y \Leftrightarrow ax \geq ay$. □

Ist (K, K^+) ein angeordneter Körper, dann verwenden wir von nun an das Relationssymbol $\leq$ (und entsprechend auch die Symbole $\leq$, $<$ und $>$) immer für die nach Satz (1.3) zu K^+ gehörende Totalordnung auf K, ohne jeweils ausdrücklich darauf hinzweisen.

In einem Ring R kann jeder Zahl $n \in \mathbb{N}_0$ auf natürliche Weise ein Element n_R in R zugeordnet werden: Zunächst ordnet man der 0 das Nullelement 0_R und der 1 das Einselement 1_R des Rings zu. Anschließend definiert man rekursiv $(n+1)_R = n_R + 1_R$ für alle $n \in \mathbb{N}$. Es gilt also $2_R = 1_R + 1_R$, $3_R = 2_R + 1_R = 1_R + 1_R + 1_R$, und so fort. Viele Rechenregeln übertragen sich von $\mathbb{N}$ auf diese Elemente. Dies ist in erster Linie darauf zurückzuführen, dass die Addition und Multiplikation auf $\mathbb{N}_0$ mit der Addition und Multiplikation der entsprechenden Körperelemente zusammenhängt.

(1.4) Lemma Für alle $m, n \in \mathbb{N}_0$ gilt $(m+n)_R = m_R + n_R$ und $(m \cdot n)_R = m_R n_R$.

Beweis: Die erste Gleichung beweisen wir durch vollständige Induktion über $n \in \mathbb{N}_0$ jeweils für alle $m \in \mathbb{N}_0$. Für $n = 0$ ist die Gleichung wegen $(m+0)_R = m_R$ und $m_R + 0_R = m_R$ offenbar für alle $m \in \mathbb{N}_0$ erfüllt. Sei nun $n \in \mathbb{N}_0$, und setzen wir die Gleichung für n voraus. Für jedes $m \in \mathbb{N}$ gilt dann

$$(m+(n+1))_R \quad = \quad ((m+n)+1)_R \quad = \quad (m+n)_R + 1_R \quad =$$

$$m_R + n_R + 1_R \quad = \quad m_R + (n+1)_R.$$

Dabei ist zu beachten, dass die Gleichungen $(m+n+1)_R = (m+n)_R + 1_R$ und $(n+1)_R = n_R + 1_R$ auf Grund der rekursiven Definition von n_R gelten. Außerdem haben wir im dritten Schritt die Induktionsvoraussetzung angewendet. Beim Beweis der zweiten Gleichung gehen wir genauso vor. Für $n = 0$ ist diese wegen $(m \cdot 0)_R = 0_R$ und $m_R \cdot 0_R = 0_R$ für alle $m \in \mathbb{N}_0$ erfüllt. Sei nun $n \in \mathbb{N}_0$, und setzen wir die Gleichung für dieses n voraus. Wir erhalten

$$(m(n+1))_R \quad = \quad (mn+m)_R \quad = \quad (mn)_R + m_R \quad = \quad m_R n_R + m_R \quad =$$

$$m_R n_R + m_R 1_R \quad = \quad m_R(n_R + 1_R) \quad = \quad m_R(n+1)_R.$$

Dabei wurde im zweiten Schritt die bereits bewiesene erste Gleichung und im dritten Schritt die Induktionsvoraussetzung angwendet. □

(1.5) Proposition *(weitere Rechenregeln in angeordneten Körpern)*

(i) Aus $0_K < a < b$ folgt $b^{-1} < a^{-1}$.

(ii) Für alle $n \in \mathbb{N}$ gilt $n_K \in K^+$.

(iii) Für alle $a, b \in K$ mit $a < b$ gilt $a < 2_K^{-1}(a+b) < b$.

Beweis: zu (i) Aus $a, b \in K^+$ folgt $ab \in K^+$. Nach Lemma (1.2) (ii) sind a^{-1} und b^{-1} und damit auch $a^{-1}b^{-1}$ in K^+ enthalten. Wegen $a < b$ liegt auch $b - a$ in K^+. Nun gilt

$$\begin{aligned} a^{-1} - b^{-1} &= 1_K a^{-1} - 1_K b^{-1} = (bb^{-1})a^{-1} - (aa^{-1})b^{-1} = b(b^{-1}a^{-1}) - a(a^{-1}b^{-1}) \\ &= b(a^{-1}b^{-1}) - a(a^{-1}b^{-1}) = (b-a)(a^{-1}b^{-1}). \end{aligned}$$

Aus $b - a \in K^+$ und $a^{-1}b^{-1} \in K^+$ folgt also $a^{-1} - b^{-1} \in K^+$ und damit $a^{-1} > b^{-1}$ wie gewünscht.

zu (ii) Wir beweisen die Aussage durch vollständige Induktion über $n \in \mathbb{N}_0$. Wegen $0_K \neq 1_K$ und $1_K^2 = 1_K$ können wir Teil (i) von Lemma (1.2), nach dem alle Quadrate ungleich Null positiv sind. Dies liefert uns $1_K \in K^+$, also die Aussage für $n = 1$. Ist nun $n \in \mathbb{N}$ beliebig, und setzen wir die Aussage für dieses n voraus, dann folgt $(n+1)_K = n_K + 1_K \in K^+$, denn die Summe zweier positiver Zahlen ist positiv.

zu (iii) Nach Definition gilt $2_K a = (1_K + 1_K)a = 1_K a + 1_K a = a + a$, und ebenso erhält man $2_K b = b + b$. Aus $a < b$ folgt $2_K a = a + a < a + b$ und $a + b < b + b = 2_K b$. Nach Teil (ii) gilt $2_K \in K^+$, und nach Lemma (1.2) (ii) ist damit auch 2_K^{-1} positiv. Aus $2_K a < a + b$ folgt nun $a = 2_K^{-1}(2_K a) \leq 2_K^{-1}(a + b)$. Aus $a + b < 2_K b$ folgt ebenso die Ungleichung $2_K^{-1}(a + b) < 2_K^{-1}(2_K b) = b$. □

(1.6) Definition Eine Teilmenge $I \subseteq K$ wird ***Intervall*** in K genannt, wenn gilt: Sind $a, b \in I$ mit $a < b$, dann sind auch alle $x \in K$ mit $a < x < b$ in I enthalten.

Für viele Anwendungen ist es hilfreich, unseren angeordneten Körper K durch Hinzunahme zweier weiterer Elemente $-\infty$ und $+\infty$ zu einer Menge $\overline{K} = K \cup \{-\infty, +\infty\}$ zu erweitern. Durch die Festlegung

$$a \leq b \quad \Leftrightarrow \quad a = -\infty \ \vee \ (a, b \in K \wedge a \leq b) \ \vee \ b = +\infty$$

wird die Totalordnung $\leq$ von K auf $\overline{K}$ ausgedehnt. Denn auf Grund der Definition gilt $a \leq a$ nicht nur für alle $a \in K$, sondern auch für $a \in \{\pm\infty\}$. Also ist $\leq$ reflexiv. Seien nun $a, b \in \overline{K}$ mit $a \leq b$ und $b \leq a$ vorgegeben. Im Fall $a, b \in K$ folgt $a = b$ direkt aus der Tatsache, dass $\leq$ eine Totalordnung auf K ist. Gilt $a \notin K$ und $b \in K$, dann folgt aus $a \leq b$, dass $a = -\infty$ sein muss, und aus $b \leq a$ ergibt sich ebenso $b = -\infty$, also $a = b$. Analog behandelt man die beiden Kombinationen $a \in K$, $b \notin K$ und $a, b \notin K$. Also ist die Relation $\leq$ antisymmetrisch.

Zum Nachweis der Transitivität seien $a, b, c \in \overline{K}$ mit $a \leq b$ und $b \leq c$ vorgegeben. Ist $a = -\infty$ oder $c = +\infty$, dann gilt auf jeden Fall $a \leq c$. Im Fall $a = +\infty$ muss $b = +\infty$ und damit auch $c = +\infty$ gelten, also ist $a \leq c$ auch hier füllt. Genauso behandelt man den Fall $c = -\infty$. Im Fall $a, c \in K$ folgt aus $a \leq b$ zunächst $b \in K$ oder $b = +\infty$; aber $b = +\infty$ würde wegen $b \leq c$ auch $c = +\infty$ bedeuten, im Widerspruch zur Voraussetzung. Also liegt auch b in K. Die Ungleichungen $a \leq b$ und $b \leq c$ implizieren dann $a \leq c$, weil $\leq$ auf K transitiv ist.

Damit ist insgesamt gezeigt, dass $\leq$ auf $\overline{K}$ eine Halbordnung ist. Darüber hinaus ist es sogar eine Totalordnung. Sind nämlich $a, b \in \overline{K}$ vorgegeben, dann liegen entweder beide Elemente in K, oder mindestens eines der Elemente liegt in $\{\pm\infty\}$. Im ersten Fall erhalten wir $a \leq b$ oder $b \leq a$, weil $\leq$ eine Totalordnung auf K ist. Im zweiten Fall sei beispielsweise $a = -\infty$. Dann gilt auf jeden Fall $a \leq b$. Nach demselben Schema geht man die Fälle $a = +\infty$, $b = -\infty$ und $b = +\infty$ durch.

(1.7) Proposition Neben der leeren Menge $\varnothing$ sind alle Teilmengen der Form

$$\begin{aligned}
]a,b[&= \{x \in K \mid a < x < b\} && a,b \in \overline{K}, a < b \\
[a,b[&= \{x \in K \mid a \leq x < b\} && a \in K, b \in \overline{K}, a < b \\
]a,b] &= \{x \in K \mid a < x \leq b\} && a \in \overline{K}, b \in K, a < b \\
[a,b] &= \{x \in K \mid a \leq x \leq b\} && a,b \in K, a \leq b
\end{aligned}$$

Intervalle in K. Insbesondere ist auch $K =]-\infty, +\infty[$ selbst ein Intervall.

Beweis: Wir beweisen die Intervalleigenschaft nur für Teilmengen der Form $[a,b[$ mit $a \in K$ und $b \in \overline{K}$, da für alle übrigen der Beweis sehr ähnlich funktioniert. Seien x, y Elemente aus $I = [a,b[$ mit $x < y$, und sei $z \in K$ mit $x < z < y$. Zu zeigen ist, dass auch z in I liegt. Wegen $x \in I$ gilt $a \leq x$, und aus $a \leq x$, $x < z$ folgt $a \leq z$ sogar $a < z$. Aus $y \in I$ folgt $y < b$, und zusammen mit $z < y$ erhalten wir $z < b$. Insgesamt gilt also $a \leq z < b$ und somit $z \in I$. Man beachte, dass es die Argumentation keine Rolle spielt, ob $b \in \mathbb{R}$ oder $b = +\infty$ gilt. □

Intervalle vom Typ $]a,b[$ bezeichnet man als ***offen***, die vom Typ $[a,b]$ als ***abgeschlossen***, und diejenigen vom Typ $[a,b[$ oder $]a,b]$ sind ***halboffen***. In jedem Fall nennt man a und b die ***Grenzen*** des Intervalls. Sind sie beide Grenzen endlich, also Elemente aus K, so spricht man von einem ***endlichen***, sonst von einem ***unendlichen Intervall***.

Die Elemente n_R kamen durch wiederholte Addition des Elements 1_R zu Stande. Genau lassen sich mit Hilfe der Multiplikation Potenzenvon Ringelementen definieren. Für beliebiges $a \in R$ setzt man $a^0 = 1_R$ und definiert anschließend rekursiv $a^{n+1} = a^n \cdot a$ für alle $n \in \mathbb{N}_0$.

(1.8) Lemma Seien R ein Ring, $a, b \in R$ und $m, n \in \mathbb{N}_0$. Dann gilt

(i) $a^{m+n} = a^m a^n$
(ii) $(a^m)^n = a^{mn}$
(iii) $(ab)^m = a^m b^m$

Beweis: zu (i) Das Beweisschema ist dasselbe wie bei Lemma (1.4): Wir beweisen die Aussage durch vollständige Induktion über $n \in \mathbb{N}_0$ jeweils für alle $m \in \mathbb{N}_0$. Für $n = 0$ ist die Gleichung erfüllt, denn es gilt $a^{m+0} = a^m$ und $a^m a^0 = a^m \cdot 1_R = a^m$. Sei nun $n \in \mathbb{N}_0$, und setzen wir die Gleichung für dieses n voraus. Dann folgt

$$a^{m+(n+1)} = a^{(m+n)+1} = a^{m+n}a = a^m a^n a = a^m a^{n+1}.$$

Dabei wurde im zweiten Schritt die rekursive Definition der höheren Potenzen und im dritten Schritt die Induktionsvoraussetzung verwendet.

zu (ii) Wegen $(a^m)^0 = 1_R$ und $a^{m0} = a^0 = 1_R$ ist die Gleichung im Fall $n = 0$ für beliebige $m \in \mathbb{N}$ gültig. Sei nun $n \in \mathbb{N}_0$, und setzen wir die Gleichung für dieses n voraus. Dann folgt

$$(a^m)^{n+1} = (a^m)^n a^m = a^{mn} a^m = a^{mn+m} = a^{m(n+1)}.$$

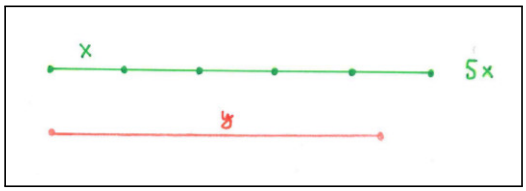
Im ersten Schritt haben wir die rekursive Definition der höheren Potenzen, im zweiten Schritt die Induktionsvoraussetzung und im dritten Schritt die Gleichung (i) verwendet.

zu (iii) Für $m = 0$ ist die Gleichung erfüllt, denn es gilt $(ab)^0 = 1_R$ und $a^0 b^0 = 1_R \cdot 1_R = 1_R$. Sei nun $m \in \mathbb{N}_0$ und die Gleichung für dieses m vorausgesetzt. Dann folgt

$$(ab)^{m+1} \quad = \quad (ab)^m (ab) \quad = \quad a^m b^m ab \quad = \quad a^m a b^m b \quad = \quad a^{m+1} b^{m+1}.$$

Also ist die Gleichung auch für $m+1$ gültig. □

(1.9) Definition *(archimedisches Axiom)*
Die Anordnung K^+ auf K wird ***archimedisch*** genannt, wenn für beliebige $x, y \in K^+$ jeweils ein $n \in \mathbb{N}$ existiert, so dass $n_K x > y$ erfüllt ist. Man bezeichnet (K, K^+) in diesem Fall als ***archimedisch angeordneten*** Körper.



Die Elemente $x, y \in K^+$ erfüllen die archimedische Bedingung: Es gilt $5_K x > y$.

Es gibt angeordnete Körper, in denen das archimedische Axiom nicht gültig ist, beispielsweise die sogenannen ***hyperreellen Zahlen***, eine Erweiterung der gewöhnlichen reellen Zahlen, in der auch „unendlich kleine“ und „unendlich große“ Elemente existieren. Wir beweisen nun noch einige Rechenregeln basierend die auf dem archimedischen Axiom.

(1.10) Satz Sei (K, K^+) ein archimedisch angeordneter Körper.

(i) Für jedes $\varepsilon \in K^+$ gibt es ein $n \in \mathbb{N}$ mit $n_K^{-1} < \varepsilon$.

(ii) Sei $x \in K$ mit $x \geq -1_K$. Dann gilt $(1_K + x)^n \geq 1_K + n_K x$ für alle $n \in \mathbb{N}_0$.
Dies ist die sogenannte ***Bernoullische Ungleichung***.

(iii) Ist $b > 1_K$, dann existiert für jedes $\kappa \in K^+$ ein $n \in \mathbb{N}$ mit $b^n > \kappa$.
Ist dagegen $0_K < b < 1_K$, so gibt es für jedes $\varepsilon \in K^+$ ein $n \in \mathbb{N}$ mit $b^n < \varepsilon$.

Beweis: zu (i) Wenden wir das archimedische Axiom auf $x = 1_K$ und $y = \varepsilon^{-1}$ an, so erhalten wir ein $n \in \mathbb{N}$ mit $n_K > \varepsilon^{-1}$. Es folgt $n_K^{-1} < \varepsilon$.

zu (ii) Wir zeigen die Ungleichung durch vollständige Induktion über $n \in \mathbb{N}_0$. Für $n = 0$ ist die Aussage offensichtlich erfüllt, denn es ist $(1_K + x)^0 = 1_K$ und $1_K + 0_K x = 1_K$. Sei nun $n \in \mathbb{N}_0$, und setzen wir die Aussage für n voraus. Es

gilt

$$\begin{aligned}(1_K+x)^{n+1} &= (1_K+x)^n(1_K+x) \geq (1_K+n_Kx)(1_K+x) = \\ 1_K+n_Kx+x+n_Kx^2 &\geq 1_K+n_Kx+x = 1_K+(n+1)_Kx \quad ,\end{aligned}$$

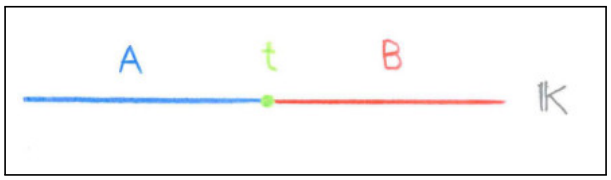
wobei wir im zweiten Schritt neben der Induktionsvoraussetzung $(1_K+x)^n \geq 1+n_Kx$ auch die Bedingung $1_K+x \geq 0$ an x verwendet haben, die sich aus der Voraussetzung $x \geq -1_K$ ergibt. Insgesamt erhalten wir die gewünschte Ungleichung für $n+1$.

zu (iii) Zunächst betrachten wir den Fall $b > 1_K$. Sei $x = b - 1_K$. Auf Grund der Bernoullischen Ungleichung gilt $b^n = (1_K+x)^n \geq 1_K + n_Kx$. Wir zeigen nun, dass ein $n \in \mathbb{N}$ mit $n_Kx > \kappa - 1_K$ existiert. Im Fall $\kappa - 1_K \leq 0_K$ ist dies bereits für $n = 1$ erfüllt; im Fall $\kappa - 1_K > 0_K$ liefert uns das archimedische Axiom wegen $x > 0_K$ ein solches Element. In beiden Fällen gilt dann $b^n \geq 1_K + n_Kx > 1_K + (\kappa - 1_K) = \kappa$. Setzen wir nun $b < 1_K$ voraus. Definieren wir $c = b^{-1}$, dann gilt $c > 1_K$. Auf Grund des bereits bewiesenen Falls, angewendet auf $\kappa = \varepsilon^{-1}$, gibt es ein $n \in \mathbb{N}$ mit $c^n > \varepsilon^{-1}$. Daraus folgt $(c^n)^{-1} < \varepsilon$. Nach Lemma (1.8) (iii) ist $c^nb^n = (cb)^n = (b^{-1}b)^n = 1_K^n = 1_K$ und somit $(c^n)^{-1} = b^n$. Auf diese Weise erhalten wir $b^n < \varepsilon$. □

(1.11) Definition Ein ***Dedekindscher Schnitt*** von K ist ein Paar (A,B) bestehend aus Teilmengen $A, B \subseteq K$ mit den folgenden Eigenschaften:

(i) Es gilt $A, B \neq \emptyset$ und $A \cup B = K$.
(ii) Für alle $a \in A$ und $b \in B$ gilt $a < b$, insbesondere ist $A \cap B = \emptyset$.

Ein Element $t \in K$ wird ***Schnittzahl*** von (A,B) genannt, wenn die Ungleichungen $a \leq t \leq b$ für alle $a \in A$ und $b \in B$ erfüllt sind.



ein Dedekindscher Schnitt (A,B) mit Schnittzahl t

(1.12) Lemma Jeder Dedekindsche Schnitt von K besitzt höchstens eine Schnittzahl.

Beweis: Sei (A,B) ein Dedekindscher Schnitt, und nehmen wir an, dass $t, t' \in K$ mit $t < t'$ beides Schnittzahlen von (A,B) sind. Nach Proposition (1.5) erfüllt das Element $u = 2_K^{-1}(t+t')$ die Ungleichungen $t < u < t'$. Wegen $t < u$ und $t \geq a$ für alle $a \in A$ gilt $u \notin A$. Andererseits gilt wegen $u < t'$ und $b \geq t'$ für alle $b \in B$ auch $u \notin B$. Das Element u liegt also weder in der Menge A noch in der Menge B. Aber wegen $A \cup B = K$ muss jedes Element von K in A oder B liegen. Dieser Widerspruch zeigt, dass keine zwei verschiedenen Schnittzahlen existieren können. □

(1.13) Definition Ein angeordneter Körper K wird ***vollständig*** genannt, wenn jeder Dedekindsche Schnitt (A,B) von K eine Schnittzahl besitzt.

Betrachten wir im Körper $\mathbb{R}$ die beiden Teilmengen

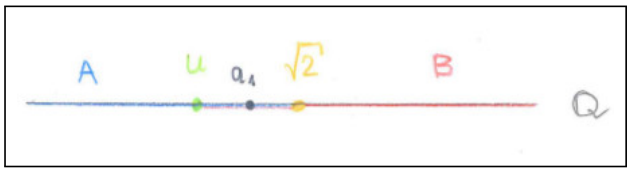
$$A = \{x \in \mathbb{R} \mid x \leq \sqrt{2}\} \qquad \text{und} \qquad B = \{x \in \mathbb{R} \mid x > \sqrt{2}\}.$$

Wir überprüfen, dass (A,B) ein Dedekindscher Schnitt ist. Die Mengen A und B sind nichtleer, denn beispielsweise gilt $1 \in A$ und $2 \in B$. Weil für jede Zahl $x \in \mathbb{R}$ entweder $x \leq \sqrt{2}$ oder $x > \sqrt{2}$ gilt, ist $A \cup B = \mathbb{R}$ erfüllt. Ist $a \in A$ und $b \in B$, dann gilt $a \leq \sqrt{2} < b$, also insbesondere $a < b$. Damit haben wir alle Bedingungen für einen Dedekindschen Schnitt verifiziert. Anhand der Ungleichungen $a \leq \sqrt{2} < b$ sieht man auch, dass $\sqrt{2}$ eine Schnittzahl von (A,B) ist.

Genauso können wir im Körper $\mathbb{Q}$ der rationalen Zahlen die beiden Teilmengen

$$A = \{x \in \mathbb{Q} \mid x \leq \sqrt{2}\} \qquad \text{und} \qquad B = \{x \in \mathbb{Q} \mid x > \sqrt{2}\} \qquad \text{definieren.}$$

Wie im vorherigen Beispiel überzeugt man sich davon, dass (A,B) ein Dedekindscher Schnitt ist. Dieser besitzt aber (im Körper $\mathbb{Q}$) keine Schnittzahl. Gehen wir nämlich davon aus, dass $u \in \mathbb{Q}$ eine Schnittzahl ist, dann gilt entweder $u < \sqrt{2}$ oder $u > \sqrt{2}$, denn die Zahl $\sqrt{2}$ ist irrational und somit nicht in $\mathbb{Q}$ enthalten. Betrachten wir zunächst den Fall $u < \sqrt{2}$.



Es gibt keine Schnittzahl $u < \sqrt{2}$.

Nach Definition der Schnittzahl muss $a \leq u$ für alle $a \in A$ gelten. Wie wir aber im nächsten Abschnitt zeigen werden, liegen die rationalen Zahlen *dicht* in $\mathbb{R}$, was bedeutet, dass in jedem noch so kleinen offenen Intervall von $\mathbb{R}$ eine rationale Zahl enthalten ist. Dies bedeutet, dass ein $a_1 \in \mathbb{Q}$ mit $u < a_1 < \sqrt{2}$ existiert. Es gilt $a_1 \in A$ nach Definition der Menge A, aber dies steht im Widerspruch zu $a \leq u$ für alle $a \in A$. Ebenso führt man den Fall $u > \sqrt{2}$ auf einen Widerspruch. Der Körper $\mathbb{Q}$ ist also nicht vollständig!

Mit Hilfe des Supremums und des Infimums können wir die Vollständigkeit angeordneter Körper auf eine andere Art und Weise charakterisieren.

(1.14) Satz Für einen angeordneten Körper (K, K^+) sind folgende Aussagen äquivalent.

(i) (K, K^+) ist vollständig.
(ii) Jede nichtleere, nach oben beschränkte Menge $M \subseteq K$ besitzt ein Supremum.
(iii) Jede nichtleere, nach unten beschränkte Menge $M \subseteq K$ besitzt ein Infimum.

Beweis: „(i) ⇒ (ii)“ Sei $M \subseteq K$ nichtleer und nach oben beschränkt. Weiter sei $B = \mathscr{S}^+(M)$ und $A = K \setminus B$. Wir zeigen, dass (A,B) ein Dedekindscher Schnitt von K ist. Weil M nach oben beschränkt ist, ist B nicht leer. Sei $x \in M$ ein beliebiges Element. Wegen $x - 1_K < x$ und $x \in M$ ist $x - 1_K$ keine obere Schranke von M. Aus $x - 1_K \notin B$ folgt $x - 1_K \in A$, also ist auch A nicht leer. Nach Definition ist auch $A \cup B = K$ erfüllt. Sei nun $a \in A$ und $b \in B$, zu zeigen ist $a < b$. Nehmen wir an, dass $a \geq b$ gilt. Weil das Element b eine obere Schranke von M ist, müsste dies erst recht für a gelten. Aber dann wäre a auch ein Element von B, was auf Grund der Definition $A = K \setminus B$ ausgeschlossen ist.

Da K nach Voraussetzung vollständig ist, existiert für (A,B) eine Schnittzahl t. Wir zeigen, dass $t = \sup M$ gilt. Nehmen wir zunächst an, t wäre keine obere Schranke. Dann gäbe es ein $x \in M$ mit $x > t$. Setzen wir $y = 2_K^{-1}(t+x)$, dann gilt $t < y < x$. Weil t die Schnittzahl von (A,B) ist, muss y in B liegen. Damit wäre y eine obere Schranke von M, was im Widerspruch zu $x \in M$ und $x > y$ steht. Gehen wir jetzt davon aus, das Element t wäre zwar eine obere Schranke, aber kein Supremum von M. Dann gäbe es ein $t' \in \mathscr{S}^+(M)$ mit $t' < t$. Weil aber t die Schnittzahl von (A,B) ist, würde daraus $t' \in A$ folgen. Somit wäre t' kein Element von B und damit keine obere Schranke von M. Insgesamt ist $t = \sup M$ damit bewiesen.

“(ii) ⇒ (iii)“ Sei $M \subseteq K$ nichtleer und nach unten beschränkt. Dann ist auch $-M = \{-x \mid x \in M\}$ eine nichtleere Teilmenge von K. Auf Grund der Äquivalenz $t \leq x \ \forall x \in M \iff -t \geq -x \ \forall \ x \in M$ ist ein Element $t \in K$ genau dann eine untere Schranke von M, wenn $-t$ eine obere Schranke von $-M$ ist. Da nach Voraussetzung die Menge M untere Schranken besitzt, gibt es für $-M$ obere Schranken. Aus unserer Voraussetzung folgt nun, dass $s = \sup(-M)$ existiert. Wir zeigen, dass $-s = \inf M$ gilt.

Weil s eine obere Schranke von $-M$ ist, handelt es sich auf Grund unserer Vorüberlegung bei $-s$ um eine untere Schranke von M. Wäre $-s$ dennoch kein Infimum von M, dann müsste es ein $s' \in \mathscr{S}^-(M)$ mit $s' > -s$ geben. Aber dann würde auch $-s' < s$ gelten, und nach unserer Vorüberlegung wäre $-s' \in \mathscr{S}^+(-M)$. Aber dies ist unmöglich, weil s nach Voraussetzung die *kleinste* obere Schranke von $-M$ ist.

“(iii) ⇒ (i)“ Sei (A,B) ein Dedekindscher Schnitt von K. Wegen $A \neq \varnothing$ gibt es ein $a \in A$, und dieses erfüllt $a \leq b$ für alle $b \in B$. Insgesamt ist B damit nichtleer und nach unten beschränkt, und auf Grund unserer Voraussetzung existiert $t = \inf B$. Wir zeigen, dass t die Schnittzahl von (A,B) ist. Weil t nach Definition eine untere Schranke von B ist, gilt $t \leq b$ für alle $b \in B$. Nehmen wir nun an, $a \leq t$ ist nicht für alle $a \in A$ erfüllt. Dann gibt es ein $a \in A$ mit $a > t$. Wegen $a \leq b$ für alle $b \in B$ ist a in $\mathscr{S}^-(B)$ enthalten. Aber dies ist unmöglich, denn nach Definition ist t das *größte* Element von $\mathscr{S}^-(B)$. Also gilt $a \leq t \leq b$ für alle $a \in A$ und $b \in B$. □

Ist $\varnothing \neq M \subseteq K$ nicht nach oben beschränkt, also $\mathscr{S}^+(M) = \varnothing$, dann definiert man $\sup M = +\infty$. Ebenso ist es üblich, $\inf M = -\infty$ zu setzen, wenn M nicht nach unten beschränkt ist. Im Gegensatz zu beliebigen angeordneten Körpern lassen sich in vollständigen Körpern die Intervalle leicht beschreiben.

(1.15) Satz Ist (K, K^+) vollständig, dann sind alle Intervalle $\neq \varnothing$ in K von der Form $]a,b[$, $]a,b]$, $[a,b[$ oder $[a,b]$ (mit a, b wie in Proposition (1.7) festgelegt).

Beweis: Sei $I \subseteq K$ ein Intervall, und gehen wir davon aus, dass I nicht leer ist. Wir beweisen den Satz, indem wir die folgenden vier Einzelfälle abarbeiten.

(i) Die Menge I ist weder nach oben noch nach unten beschränkt.
(ii) Sie ist nach oben, aber nicht nach unten beschränkt.
(iii) Sie ist nach unten, aber nicht nach oben beschränkt.
(iv) Sie ist beschränkt.

zu (i) Wir zeigen, dass in diesem Fall $I = K$ gilt. Ist $x \in K$ beliebig, dann gibt es ein $z \in I$ mit $z > x$, denn sonst wäre x eine obere Schranke von I. Ebenso gibt es ein $y \in I$ mit $y < x$, weil x ansonsten eine untere Schranke von I wäre. Aus $y, z \in I$, den Ungleichungen $y < x < z$ und der Intervall-Eigenschaft folgt $x \in I$.

zu (ii) Hier zeigen wir, dass I die Form $]-\infty, b[$ oder $]-\infty, b]$ hat. Weil I nichtleer und nach oben beschränkt ist, existiert $b = \sup I$. Setzen wir zunächt $b \in I$ voraus und zeigen wir, dass in diesem Fall $I =]-\infty, b]$ gilt. Ist $x \in I$, dann folgt $x \leq b$, weil b eine obere Schranke von I ist. Somit erhalten wir $I \subseteq]-\infty, b]$. Zum Beweis der umgekehrten Inklusion sei $x \in]-\infty, b]$ vorgegeben. Dann gilt $x \leq b$, und im Fall $x = b$ ist $x \in I$ offenbar erfüllt. Gehen wir von $x < b$ aus, dann gibt es jedenfalls ein $y \in I$ mit $y < x$, denn ansonsten wäre x eine untere Schranke von I, die nach Voraussetzung nicht existiert. Wegen $x < b$ ist x keine obere Schranke von I, somit gibt es ein $z \in I$ mit $x < z$. Aus $y, z \in I$, den Ungleichungen $y < x < z$ und der Intervall-Eigenschaft folgt $x \in I$. Insgesamt haben wir damit gezeigt, dass $I =]-\infty, b]$ gilt.

Betrachten wir nun den Fall $b \notin I$ und zeigen, dass in dieser Situation $I =]-\infty, b[$ gilt. Ist $x \in I$, dann gilt $x < b$ und somit $x \in]-\infty, b[$. Sei umgekehrt $x \in]-\infty, b[$ vorgegeben. Dann gilt $x < b$. Weil x keine untere Schranke von I ist, gibt es ein $y \in I$ mit $y < x$. Weil x keine obere Schranke von I ist, existiert ein $z \in I$ mit $x < z$. Aus $y, z \in I$, den Ungleichungen $y < x < z$ und der Intervall-Eigenschaft folgt wiederum $x \in I$, damit ist die Gleichung $I =]-\infty, b[$ insgesamt bewiesen.

zu (iii) Weil I nichtleer und nach unten beschränkt ist, existiert $a = \inf I$. Wie im vorherigen Absatz beweist man nun $I =]a, +\infty[$ im Fall $a \notin I$ und $I = [a, +\infty[$ im Fall $a \in I$. Weil der Beweis völlig analog verläuft, verzichten wir auf die Angabe der Details.

zu (iv) Weil I nichtleer und beschränkt ist, existieren $a = \inf I$ und $b = \sup I$. Wir betrachten zunächst den Fall, dass $a \notin I$ und $b \in I$ gilt und zeigen, dass dann die Gleichung $I =]a, b]$ gilt. Sei $x \in I$. Dann gilt $a \leq x \leq b$, weil a eine untere und b eine obere Schranke von I ist. Wegen $a \notin I$ gilt sogar $a < x \leq b$, und daraus folgt $x \in]a, b]$. Setzen wir nun umgekehrt $x \in]a, b]$ voraus. Ist $x = b$, dann folgt $x \in I$. Ansonsten gilt $a < x < b$. Weil x keine untere Schranke für I ist (nach Definition ist a die größte untere Schranke), gibt es ein $y \in I$ mit $y < x$. Andererseits ist x auch keine obere Schranke von I, so dass ein $z \in I$ mit $x < z$ existiert. Aus $y, z \in I$, den Ungleichungen $y < x < z$ und der Intervall-Eigenschaft folgt $x \in I$. Damit ist $I =]a, b]$ nachgewiesen. Nach demselben Schema beweist man, dass im Fall $a \in I, b \notin I$ das Intervall I durch $I = [a, b[$ gegeben ist, im Fall $a, b \notin I$ durch $I =]a, b[$ und im Fall $a, b \in I$ durch $I = [a, b]$. Insgesamt ist der Beweis damit abgeschlossen. □

Wir werden die vollständig angeordneten Körper noch auf eine weitere Art charakterisieren.

(1.16) Definition Man sagt, der angeordnete Körper (K, K^+) besitzt die ***Intervallschachtelungs-Eigenschaft***, wenn für jede Folge $(I_n)_{n\in\mathbb{N}}$ nichtleerer, abgeschlossener, endlicher Intervalle, die $I_{n+1} \subseteq I_n$ für alle $n \in \mathbb{N}$ erfüllt, die Schnittmenge $\bigcap_{n\in\mathbb{N}} I_n$ nicht leer ist.

Der Körper $\mathbb{Q}$ der rationalen Zahlen besitzt die Intervallschachtelungs-Eigenschaft nicht. Um das zu sehen, konstruieren wir uns folgendermaßen eine Folge $(I_n)_{n\in\mathbb{N}}$ abgeschlossener Intervall I_n in $\mathbb{Q}$. Weil $\mathbb{Q}$ in $\mathbb{R}$ dicht liegt, finden wir zunächst rationale Zahlen a_1, b_1 mit $\sqrt{2}-1 < a_1 < \sqrt{2}$ und $\sqrt{2} < b_1 < \sqrt{2}+1$. Sind a_n, b_n bereits konstruiert, dann wählen wir $a_{n+1}, b_{n+1} \in \mathbb{Q}$ mit

$$\max\{a_n, \sqrt{2}-\tfrac{1}{n+1}\} < a_{n+1} < \sqrt{2} \qquad \text{und} \qquad \sqrt{2} < b_{n+1} < \min\{b_n, \sqrt{2}+\tfrac{1}{n+1}\}.$$

Setzen wir nun $I_n = [a_n, b_n] \cap \mathbb{Q}$, dann sind alle I_n nichtleere, endliche und abgeschlossene Intervalle in $\mathbb{Q}$, und es gilt $I_{n+1} \subseteq I_n$ für alle $n \in \mathbb{N}$. Aber der Durchschnitt $\bigcap_{n\in\mathbb{N}} I_n$ ist leer! Denn nehmen wir an, dass $r \in \mathbb{Q}$ in allen Intervallen I_n enthalten ist. Wegen $\sqrt{2} \notin \mathbb{Q}$ muss $r < \sqrt{2}$ oder $r > \sqrt{2}$ gelten. Im ersten Fall gilt $r < \sqrt{2} - \frac{1}{n}$ für hinreichend großes n. Dann liegt r nicht in $[\sqrt{2}-\frac{1}{n}, \sqrt{2}+\frac{1}{n}]$, erst recht nicht im Intervall $[a_n, b_n]$, das eine Teilmenge davon ist. Im zweiten Fall finden wir ein $n \in \mathbb{N}$ mit $\sqrt{2}+\frac{1}{n} < r$, was erneut $r \notin [\sqrt{2}-\frac{1}{n}, \sqrt{2}+\frac{1}{n}]$ und $r \notin [a_n, b_n]$ zur Folge hat.

(1.17) Satz Ein angeordneter Körper (K, K^+) ist genau dann vollständig, wenn er die Intervallschachtelungs-Eigenschaft besitzt und die Anordnung K^+ archimedisch ist.

Beweis: „⇒“ Wir setzen voraus, dass (K, K^+) vollständig ist und beweisen zunächst die Intervallschachtelungs-Eigenschaft. Sei $I_n = [a_n, b_n]$ eine Folge nichtleerer, abgeschlossener, endlicher Intervalle mit $a_n, b_n \in K$, $a_n < b_n$ und $I_n \subseteq I_{n+1}$ für alle $n \in \mathbb{N}$. Wir definieren

$$A \quad = \quad \{a_n \mid n \in \mathbb{N}\}.$$

Offenbar ist A nichtleer, und wegen $A \subseteq I_1$ ist A beschränkt. Auf Grund der Vollständigkeit existiert somit $a = \sup A$. Zum Nachweis der Intervallschachtelungs-Eigenschaft zeigen wir, dass a im Durchschnitt $\bigcap_{n\in\mathbb{N}} I_n$ der Intervalle enthalten ist. Weil a eine obere Schranke von A ist, gilt $a \geq a_n$ für alle $n \in \mathbb{N}$. Liegt a dennoch nicht im Durchschnitt, dann gilt $a \notin I_m$ für ein $m \in \mathbb{N}$, also $a > b_m$. Wir zeigen nun, dass b_m eine obere Schranke von A ist.

Wegen $I_{n+1} \subseteq I_n$ für alle $n \in \mathbb{N}$ gilt jeweils $a_{n+1} \geq a_n$ und $b_{n+1} \leq b_n$. Ist nun $n \geq m$, dann gilt $b_m \geq b_n > a_n$. Im Fall $n \leq m$ gilt $b_m > a_m \geq a_n$. Also ist b_m tatsächlich eine obere Schranke von A. Weil a aber die *kleinste* obere Schranke von a ist, folgt daraus $a \leq b_m$, im Widerspruch zu unserer vorherigen Feststellung. Dies widerspruch zeigt, dass a im Durchschnitt sämtlicher Intervalle enthalten ist.

Nun zeigen wir noch, dass K^+ archimedisch ist. Seien $x, y \in K^+$ vorgegeben, und nehmen wir an, es gibt kein $n \in \mathbb{N}$ mit $n_K x > y$. Dann gilt $n_K x \leq y$ und somit $n_K \leq \frac{x}{y}$ für alle $n \in \mathbb{N}$, das heißt die Menge

$$N \quad = \quad \{n_K \mid n \in \mathbb{N}\}$$

ist nach oben beschränkt. Wegen $1_K \in N$ ist die Menge N außerdem nichtleer. Auf Grund der Vollständigkeit existiert somit $s = \sup N$. Weil s nach Definition die kleinste obere Schranke von N ist, kann $s - 1_K$ wegen $s - 1_K < s$ keine obere Schranke von N sein. Daraus folgt, dass ein $n \in \mathbb{N}$ mit $n_K > s - 1_K$ existiert. Wir erhalten $(n+1)_K = n_K + 1_K > s$. Aber wegen $(n+1)_K \in N$ widerspricht dies $s \in \mathscr{S}^+(N)$. Unsere Annahme, dass kein $n \in \mathbb{N}$ mit $n_K x > y$ existiert, hat zu einem Widerspruch geführt. Also gibt es ein solches n, und K^+ ist archimedisch.

„⇐“ Setzen wir nun voraus, dass (K, K^+) die Intervallschachtelungs-Eigenschaft besitzt und ein archimedisch angeordneter Körper ist. Wir müssen zeigen, dass daraus die Vollständigkeit von (K, K^+) folgt. Sei dazu $M \subseteq K$ eine nichtleere, nach oben beschränkte Menge. Zu zeigen ist, dass $s = \sup M$ existiert. Die grundlegende besteht darin, das hypothetische s in eine Folge immer kleiner werdender Intervalle „einzusperren“. Dazu definieren wir eine Folge $(I_n)_{n\in\mathbb{N}}$ von Intervallen durch folgendes Verfahren. Sei $a \in M$ und $b \in \mathscr{S}^+(M)$ mit $a < b$. Zunächst setzen wir $a_0 = a$, $b_0 = b$. Sei nun $n \in \mathbb{N}_0$, und nehmen wir an, dass a_n, b_n bereits konstruiert sind. Wir setzen dann $c_n = 2_K^{-1}(a_n + b_n)$ und führen eine Fallunterscheidung durch. Gilt $c_n \in \mathscr{S}^+(M)$, dann definieren wir $a_{n+1} = a_n$ und $b_{n+1} = c_n$. Ist c_n dagegen keine obere Schranke von M, dann setzen wir $a_{n+1} = c_n$ und $b_{n+1} = b_n$. Wir überprüfen durch vollständige Induktion für jedes $n \in \mathbb{N}_0$ die folgenden Aussagen.

(i) $b_n - a_n = (2_K^{-1})^n (b - a)$
(ii) Das Element b_n ist eine obere Schranke von M.
(iii) Für das Intervall $I_n = [a_n, b_n]$ gilt $I_n \cap M \neq \varnothing$.

Für $n = 0$ sind die Aussagen (i), (ii) und (iii) nach Definition erfüllt. Sei nun $n \in \mathbb{N}_0$, und setzen wir die drei Aussagen für n voraus. Es gilt

$$c_n - a_n \quad = \quad 2_K^{-1} a_n + 2_K^{-1} b_n - a_n \quad = \quad 2_K^{-1}(b_n - a_n) \quad = \quad (2_K^{-1})^{n+1}(b - a)$$

und ebenso $b_n - c_n = b_n - 2_K^{-1} b_n - 2_K^{-1} a_n = 2_K^{-1}(b_n - a_n) = (2_K^{-1})^{n+1}(b - a)$. Weil (a_{n+1}, b_{n+1}) entweder mit (a_n, c_n) oder mit (c_n, b_n) übereinstimmt, ist (i) für $n + 1$ auf jeden Fall erfüllt. Außerdem wird b_{n+1} so gewählt, dass auch Aussage (ii) für $n + 1$ gültig ist. Überprüfen wir nun noch die Aussage (iii). Nach Induktionsvoraussetzung gibt es ein $x \in M$ mit $a_n \leq x \leq b_n$. Ist $(a_{n+1}, b_{n+1}) = (a_n, c_n)$, dann ist $c_n = b_{n+1}$ eine obere Schranke von M. Es gilt dann $a_n \leq x \leq c_n$, also $x \in I_{n+1} \cap M$. Im Fall $(a_{n+1}, b_{n+1}) = (c_n, b_n)$ ist c_n keine obere Schranke von M, die Zahl b_n nach Induktionsvoraussetzung aber schon. Somit gibt es ein $x \in M$ mit $c_n < x \leq b_n$, und es folgt $x \in I_{n+1} \cap M$.

Damit ist der Induktionsbeweis ist abgeschlossen. Für jedes $n \in \mathbb{N}_0$ gilt wegen $a_n < c_n < b_n$ in jedem Fall $a_{n+1} \geq a_n$ und $b_{n+1} \leq b_n$, insgesamt also $I_{n+1} \subseteq I_n$. Deshalb kann die Intervallschachtelungs-Eigenschaft angewendet werden, und diese zeigt, dass ein Element $s \in K$ im Durchschnitt $\bigcap_{n\in\mathbb{N}_0} I_n$ sämtlicher Intervalle existiert. Wir werden nun zeigen, dass $s = \sup M$ gilt. Nehmen wir zunächst an, dass s keine obere Schranke von M ist. Dann gibt es ein $x \in M$ mit $x > s$. Für jedes $n \in \mathbb{N}$ gilt dann also $a_n \leq s < x \leq b_n$, Letzteres weil b_n jeweils eine obere Schranke von M ist und $x \in M$ gilt. Es folgt $x - s \leq b_n - a_n \leq 2^{-n}(b - a)$ für alle $n \in \mathbb{N}$, also $\frac{x-s}{b-a} \leq 2^{-n}$ für alle $n \in \mathbb{N}$. Aber andererseits existiert nach Satz (1.10) (iii) in $n \in \mathbb{N}$ mit $2^{-n} < \frac{x-s}{b-a}$. Der Widerspruch zeigt, dass s tatsächlich eine obere Schranke von M ist.

Nehmen wir nun an, s ist nicht die kleinste obere Schranke von M. Dann existiert eine obere Schranke $t < s$ von M. Für jedes $n \in \mathbb{N}$ existiert ein $x_n \in M \cap I_n$ auf Grund von Eigenschaft (iii) der Intervallfolge. Es gilt dann also $a_n \leq x_n \leq t < s \leq b_n$ und somit $t - s \leq b_n - a_n \leq 2^{-n}(b-a)$ für alle $n \in \mathbb{N}$. Dies zeigt, dass $\frac{t-s}{b-a} < 2^{-n}(b-a)$ für alle $n \in \mathbb{N}$ gelten müsste, was erneut zu einem Widerspruch mit Teil (iii) von Satz (1.10) führt. Also ist s tatsächlich das Supremum der Menge M. □

Mit Hilfe des Intervallschachtelungs-Kriteriums leiten wir eine weitere wichtige Eigenschaft vollständiger angeordneter Körper her.

(1.18) Satz Jeder vollständige, angeordnete Körper besitzt überabzählbar viele Elemente.

Beweis: Sei (K, K^+) vollständig, und nehmen wir an, dass K höchstens abzählbar ist. Dann gibt es eine injektive Abbildung $f : K \to \mathbb{N}$, und auf Grund der Injektivität existiert, wie wir in der Linearen Algebra gezeigt haben, eine Abbildung $g : \mathbb{N} \to K$ mit $f \circ g = \mathrm{id}_{\mathbb{N}}$. Aus dieser Gleichung wiederum folgt, dass g surjektiv ist. Wir definieren $x_n = g(n)$ für alle $n \in \mathbb{N}$ und werden die Surjektivität von g nun zu einem Widerspruch führen. Dazu definieren wir zwei Folgen $(i_\ell)_{\ell \in \mathbb{N}}$ und $(j_\ell)_{\ell \in \mathbb{N}}$ natürlicher Zahlen nach folgendem Schema. Zunächst setzen wir $i_1 = 1$ und $j_1 = \min\{n \in \mathbb{N} \mid x_n > x_{i_1}\}$. Ist nun $\ell \in \mathbb{N}$ und sind i_ℓ, j_ℓ bereits konstruiert, dann setzen wir

$$\begin{aligned} i_{\ell+1} &= \min\{ n \in \mathbb{N} \mid x_{i_\ell} < x_n < x_{j_\ell} \} \\ j_{\ell+1} &= \min\{ n \in \mathbb{N} \mid x_{i_{\ell+1}} < x_n < x_{j_\ell} \} \end{aligned}$$

Nach Konstruktion gilt für alle $\ell \in \mathbb{N}$ jeweils $x_{i_\ell} < x_{i_{\ell+1}} < x_{j_{\ell+1}} < x_{j_\ell}$, und außerdem $i_1 < j_1 < i_2 < j_2 < i_3 < \ldots$, weil die Intervalle, in denen das „nächste“ x_n liegen darf, fortwährend kleiner werden. Setzen wir $I_\ell = [x_{i_\ell}, x_{j_\ell}]$, dann gilt also jeweils $I_\ell \supsetneq I_{\ell+1}$. Auf Grund des Intervallschachtelungs-Kriteriums gibt es ein $x \in K$ im Durchschnitt $\bigcap_{\ell \in \mathbb{N}} I_\ell$ all dieser Intervalle.

Auf Grund der Surjektivität von oben gibt es ein $m \in \mathbb{N}$ mit $x = x_m$. Es gilt dann $x_{i_\ell} \leq x_m \leq x_{j_\ell}$ für alle $\ell \in \mathbb{N}_0$; wegen $x_{i_\ell} < x_{i_{\ell+1}}$ und $x_{j_{\ell+1}} < x_{j_\ell}$ muss sogar $x_{i_\ell} < x_m < x_{j_\ell}$ für alle ℓ gelten. Weil die Indizes i_ℓ eine unendliche Teilmenge von $\mathbb{N}$ durchlaufen, muss es ein minimales $k \in \mathbb{N}$ mit $i_k > m$ geben. Dabei ist $k \geq 2$, denn im Fall $k = 1$ wäre $1 = i_1 > m$, was wegen $m \in \mathbb{N}$ unmöglich ist. Nach Definition gilt

$$i_k = \min\{ n \in \mathbb{N} \mid x_{i_{k-1}} < x_n < x_{j_{k-1}} \}.$$

Aber andererseits gilt $x_{i_{k-1}} < x_m < x_{j_{k-1}}$, also widerspricht $i_k > m$ der Minimalität von i_k. Die Annahme „K höchstens abzählbar“ war also falsch, und somit ist K überabzählbar. □

Die *Existenz* vollständiger, angeordneter Körper kann aus den Axiomen der Mengenlehre abgeleitet werden. Leider können wir diese Konstruktion aus Zeitgründen hier nicht durchführen. Wir formulieren ohne Beweis

(1.19) Satz

Es gibt vollständige, angeordnete Körper. Wir wählen einen solchen Körper willkürlich aus, bezeichnen ihn mit $\mathbb{R}$ und nennen ihn den Körper der ***reellen Zahlen***.

Für Null- und Einselement $0_{\mathbb{R}}, 1_{\mathbb{R}}$ in $\mathbb{R}$ schreiben wir 0 und 1, lassen die Angabe des Körpers $\mathbb{R}$ im Index also künftig weg. An Stelle des Ausdrucks ab^{-1} schreiben wir auch $\frac{a}{b}$, d.h. wir verwenden die aus der Schule bekannte Bruchschreibweise. Wir verwenden die Bezeichnung $\mathbb{R}^+$ für die Teilmenge der *positiven* reelle Zahlen und setzen $\mathbb{R}_+ = \mathbb{R}^+ \cup \{0\}$.

Sämtliche Eigenschaften, die wir im letzten Abschnitt für vollständige Körper hergeleitet haben, sind also auch für die reellen Zahlen gültig, nämlich

(i) Jede nichtleere, nach oben beschränkte Teilmenge von $\mathbb{R}$ besitzt nach Satz (1.14) ein Supremum, und jede nichtleere, nach unten beschränkte Teilmenge besitzt ein Infimum.

(ii) Der Körper $\mathbb{R}$ besitzt nach Satz (1.15) die Intervallschachtelungs-Eigenschaft.

(iii) Die Anordnung $\mathbb{R}^+$ ist auf Grund von Satz (1.15) archimedisch. Damit sind auch die unter Satz (1.10) formulierten Aussagen für $\mathbb{R}$ gültig.

(iv) Der Körper $\mathbb{R}$ ist überabzählbar, siehe Satz (1.18).

Bei der Definition der archimedischen Eigenschaft wurden allerdings die natürlichen Zahlen als bekannt vorausgesetzt. Wir werden nun die natürlichen Zahlen als Teilmenge von $\mathbb{R}$ definieren.

(1.20) Definition Eine Teilmenge $M \subseteq \mathbb{R}$ wird ***induktiv*** genannt, wenn $1 \in M$ gilt und für jedes $x \in M$ auch $x + 1 \in M$ erfüllt ist.

Beispiele für induktive Teilmengen in $\mathbb{R}$ sind $\mathbb{R}$ selbst, $\mathbb{R}_+$ oder $\mathbb{R}^+$.

(1.21) Definition Die Teilmenge $\mathbb{N} \subseteq \mathbb{R}$ der ***natürlichen Zahlen*** ist definiert durch

$$\mathbb{N} \quad = \quad \bigcap_{M \subseteq \mathbb{R} \text{ induktiv}} M \quad = \quad \{\, n \in \mathbb{R} \mid n \in M \text{ für jede induktive Teilmenge } M \subseteq \mathbb{R} \,\}.$$

Außerdem definieren wir $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$.

Die Menge $\mathbb{N}$ ist selbst induktiv. Da nämlich 1 nach Definition in jeder induktiven Menge liegt, ist 1 auch im Durchschnitt aller induktiven Mengen, also in $\mathbb{N}$, enthalten. Ist $x \in \mathbb{N}$, dann liegt x in in jeder induktivem Menge. Nach Definition ist dann auch $x + 1$ in jeder induktiven Menge enthalten. Deshalb folgt $x + 1 \in \mathbb{N}$.

Aus der Definition von $\mathbb{N}$ folgt unmittelbar, dass $\mathbb{N} \subseteq M$ für jede induktive Teilmenge $M \subseteq \mathbb{R}$ gilt. Ist M umgekehrt eine induktive Teilmenge von $\mathbb{R}$ mit $M \subseteq \mathbb{N}$, dann muss also $M = \mathbb{N}$ gelten.

(1.22) Proposition

(i) Sind $m, n \in \mathbb{N}$, dann gilt auch $m + n \in \mathbb{N}$ und $mn \in \mathbb{N}$.
(ii) Ist $m \in \mathbb{N}$, dann liegt $m - 1$ in $\mathbb{N}_0$.
(iii) Alle natürlichen Zahlen sind positiv, d.h. es gilt $\mathbb{N} \subseteq \mathbb{R}^+$.

Beweis: zu (i) Sei $m \in \mathbb{N}$. Dann ist die Teilmenge $T = \{n \in \mathbb{N} \mid m + n \in \mathbb{N}\}$ induktiv, denn es gilt $m + 1 \in \mathbb{N}$ auf Grund der Induktivität von $\mathbb{N}$, und aus $m + n \in \mathbb{N}$ folgt aus demselben Grund $m + (n + 1) \in \mathbb{N}$. Es folgt $T = \mathbb{N}$, also $m + n \in \mathbb{N}$ für alle $n \in \mathbb{N}$. Ebenso ist $U = \{n \in \mathbb{N} \mid mn \in \mathbb{N}\}$ eine induktive Menge, denn wegen $m1 = m \in \mathbb{N}$ gilt $1 \in U$, und aus $n \in U$ folgt $mn \in \mathbb{N} \Rightarrow mn + m = m(n + 1) \in \mathbb{N} \Rightarrow n + 1 \in U$. Also ist wiederum $U = \mathbb{N}$, es gilt also $mn \in \mathbb{N}$ für alle $n \in \mathbb{N}$.

zu (ii) Die Menge $M = \{x \in \mathbb{R} \mid x - 1 \in \mathbb{N}_0\}$ ist induktiv. Denn wegen $1 - 1 = 0$ gilt $1 \in M$, und ist $n \in M$, dann gilt $n - 1 \in \mathbb{N}_0$. Sowohl für $n - 1 = 0$ als auch für $n - 1 \in \mathbb{N}$ ist $(n + 1) - 1 = n$ in $\mathbb{N}_0$ enthalten, und daraus folgt $n + 1 \in M$. Die Induktivität von M impliziert $\mathbb{N} \subseteq M$. Also gilt $n - 1 \in \mathbb{N}_0$ für alle $n \in \mathbb{N}$.

zu (iii) Dies folgt aus der Induktivität von $\mathbb{R}^+$. □

(1.23) Satz Die natürlichen Zahlen $\mathbb{N}$ erfüllen die Peano-Axiome (P1) bis (P5).

Beweis: Die Axiome (P1) (Existenz eines ausgezeichneten Elements $1 \in \mathbb{N}$) und (P2) (Existenz eines Nachfolgers $n + 1$ für jedes $n \in \mathbb{N}$) ergeben sich direkt aus der Induktivität von $\mathbb{N}$. Nehmen wir an, es gibt ein $n \in \mathbb{N}$, das 1 als Nachfolger besitzt. Dann gilt $n + 1 = 1$, und daraus folgt $n = 0$. Aber die Null ist in $\mathbb{N}$ nicht enthalten, denn $\mathbb{R}^+$ ist eine induktive Menge mit $0 \notin \mathbb{R}^+$, und $\mathbb{N}$ ist in jeder induktiven Teilmenge von $\mathbb{R}$ enthalten. Damit ist (P3) bewiesen. Sind $m, n \in \mathbb{N}$ mit $m+1 = n+1$ vorgegeben, dann können wir auf beiden Seiten der Gleichung den Wert -1 addieren und erhalten $m = n$. Dies zeigt (P4).

Nun beweisen wir noch das Induktionsprinzip (P5). Sei dazu φ ein Aussagenschema mit Parameter x, dass für jedes $n \in \mathbb{N}$ eine sinnvolle Aussage $\varphi(n)$ liefert. Außerdem setzen wir voraus, dass die Aussagen $\varphi(1)$ und $\forall x \in \mathbb{N} : \varphi(x) \Rightarrow \varphi(x + 1)$ erfüllt sind. Dann ist $A = \{n \in \mathbb{N} \mid \varphi(n)\}$ eine induktive Menge, denn wegen $\varphi(1)$ gilt $1 \in A$, und auf Grund der zweiten Voraussetzung ist für jedes $n \in A$ auch $n + 1$ in A enthalten. Wie zu Anfang des Abschnitts bemerkt, folgt aus $A \subseteq \mathbb{N}$ und der Induktivität von A die Gleichheit $A = \mathbb{N}$. Somit ist die Aussage $\forall x \in \mathbb{N} : \varphi(x)$ gültig und damit das Induktionsprinzip (P5) bewiesen. □

In den Übungen wurde gezeigt, dass aus den Peano-Axiomen folgt, dass jede nichtleere Teilmenge von $\mathbb{N}$ ein kleinstes Element besitzt. Wir bemerken noch, dass die Teilmenge $\mathbb{N} \subseteq \mathbb{R}$ nach oben ***unbeschränkt*** ist. Denn nehmen wir an, $s \in \mathbb{R}$ wäre eine obere Schranke von $\mathbb{N}$. Wegen $1 \leq s$ muss $s \in \mathbb{R}^+$ gelten. Auf Grund der archimedischen Eigenschaft gibt es aber ein $n \in \mathbb{N}$ mit $n = n \cdot 1 > s$, was zu unserer Annahme im Widerspruch steht.

(1.24) Definition Die Menge $\mathbb{Z}$ der ***ganzen Zahlen*** und die Menge $\mathbb{Q}$ der ***rationalen Zahlen*** sind als Teilmengen von $\mathbb{R}$ definiert durch

$$\begin{aligned} \mathbb{Z} &= \mathbb{N}_0 \cup \{n \in \mathbb{R} \mid -n \in \mathbb{N}\} \\ \mathbb{Q} &= \{x \in \mathbb{R} \mid \exists\, a \in \mathbb{Z}, b \in \mathbb{N} : x = \tfrac{a}{b}\}. \end{aligned}$$

(1.25) Satz

(i) Durch Einschränkung der Abbildungen $+ : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ und $\cdot : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ auf $\mathbb{Z} \times \mathbb{Z}$ erhält man zwei Verknüpfungen $+$ und $\cdot$ auf $\mathbb{Z}$. Entsprechendes gilt für die Menge $\mathbb{Q}$ der rationalen Zahlen.

(ii) Das Tripel $(\mathbb{Z}, +, \cdot)$ ist ein Ring.

(iii) Das Tripel $(\mathbb{Q}, +, \cdot)$ ist ein Körper, und durch $\mathbb{Q}^+ = \mathbb{R}^+ \cap \mathbb{Q}$ ist eine archimedische Anordnung auf $\mathbb{Q}$ definiert.

Beweis: zu (i) Durch Einschränkung der Verknüpfungen $+$ und $\cdot$ auf $\mathbb{Z} \times \mathbb{Z}$ erhält man zunächst Abbildungen $+ : \mathbb{Z} \times \mathbb{Z} \to \mathbb{R}$ und $\cdot : \mathbb{Z} \times \mathbb{Z} \to \mathbb{R}$. Um zu zeigen, dass diese Abbildungen Verknüpfungen auf $\mathbb{Z}$ sind, müssen wir nachweisen, dass Summen und Produkte ganzer Zahlen wiederum ganze Zahlen sind.

Wir beginnen mit der Multiplikation. Für vorgegebene $x, y \in \mathbb{Z}$ beweist man $xy \in \mathbb{Z}$ durch separate Betrachtung der Fälle $x \in \mathbb{N}$, $x = 0$ und $-x \in \mathbb{N}$ sowie $y \in \mathbb{N}$, $y = 0$ und $-y \in \mathbb{N}$. Ist beispielsweise $x \in \mathbb{N}$ und $-y \in \mathbb{N}$, dann folgt aus Proposition (1.22) $-xy \in \mathbb{N}$ und damit $xy \in \mathbb{Z}$. Die anderen acht Kombinationen können nach dem gleichen Prinzip abgearbeitet werden.

Bevor wir zeigen, dass auch $x + y$ in $\mathbb{Z}$ liegt, bemerken wir zunächst, dass $\mathbb{Z} \cap \mathbb{R}^+ = \mathbb{N}$ gilt. Die Inklusion „$\supseteq$“ folgt aus Proposition (1.22) (iii). Sei umgekehrt $a \in \mathbb{Z} \cap \mathbb{R}^+$. Wegen $0 \notin \mathbb{R}^+$ ist $a = 0$ ausgeschlossen. Wäre $a = -n$ für ein $n \in \mathbb{N}$, dann würden sowohl $-n$ als auch n in $\mathbb{R}^+$ liegen, was den Anordungseigenschaften von $\mathbb{R}^+$ widersprechen würde. Also bleibt $a = n$ als einzige Möglichkeit.

Nun beweisen wir durch vollständige Induktion über $n \in \mathbb{N}$: Ist $x \in \mathbb{Z}$ und $n \in \mathbb{N}_0$, dann liegt $x - n$ in $\mathbb{Z}$. Für $n = 0$ ist nichts zu zeigen. Sei nun die Aussage für n bewiesen und $x \in \mathbb{Z}$. Wir müssen zeigen, dass für $x \in \mathbb{Z}$ auch $x-(n+1) = (x-n)-1$ in $\mathbb{Z}$ liegt. Nach Induktionsvoraussetzung gilt $x - n \in \mathbb{Z}$. Im Fall $x - n > 0$ gilt $x - n \in \mathbb{N}$ auf Grund der Bemerkung von oben. Aus Proposition (1.22) (ii) folgt $x - (n+1) = (x-n) - 1 \in \mathbb{N}_0$, also insbesondere $x - (n+1) \in \mathbb{Z}$. Ist $x - n = 0$, dann gilt $x - (n+1) = -1$, und -1 ist ebenfalls in $\mathbb{Z}$ enthalten. Betrachten wir nun noch den Fall $x - n < 0$. Dann ist $n - x$ wieder in $\mathbb{Z} \cap \mathbb{R}^+ = \mathbb{N}$ enthalten. Es folgt $n - x + 1 \in \mathbb{N}$ und somit $x - (n+1) = -(n - x + 1) \in \mathbb{Z}$.

Nach diesen Vorbereitungen können wir nun zeigen, dass mit $x, y \in \mathbb{Z}$ auch $x + y$ in $\mathbb{Z}$ enthalten ist. Sind $x, y > 0$, dann folgt $x, y \in \mathbb{N}$ und somit $x + y \in \mathbb{N} \subseteq \mathbb{Z}$. Im Fall $x, y < 0$ liegen $-x$ und $-y$ beide in $\mathbb{N}$, somit gilt auch $(-x) + (-y) \in \mathbb{N}$ und $x + y = -((-x) + (-y)) \in \mathbb{Z}$. Sei nun $x > 0$ und $y < 0$, also $n = -y \in \mathbb{N}$. Wie zuvor gezeigt, gilt dann $x + y = x + (-n) = x - n \in \mathbb{Z}$. Ebenso behandelt man den Fall $x < 0$, $y > 0$.

Wir zeigen nun, dass die Einschränkungen von $+$ und $\cdot$ auf $\mathbb{Q} \times \mathbb{Q}$ Verknüpfungen auf $\mathbb{Q}$ sind, d.h. dass für $r, s \in \mathbb{Q}$ auch $r + s$ und rs in $\mathbb{Q}$ liegen. Sei dazu $r = \frac{a}{b}$ und $s = \frac{c}{d}$ mit $a, c \in \mathbb{Z}$ und $b, d \in \mathbb{N}$. Auf Grund der Gleichungen

$$r + s = \frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{bd} \qquad \text{und} \qquad rs = \frac{ac}{bd}$$

und wegen $bd \in \mathbb{N}$ sowie $ad + bc, ac \in \mathbb{Z}$ sind $r + s$ und rs Elemente von $\mathbb{Q}$.

zu (ii) Wir überprüfen die Bedingungen aus der Definition der Ringe. Da die Assoziativ- und Kommutativgesetze sowie das Distributivgesetz für alle Elemente in $\mathbb{R}$ gelten, sind sie erst recht für alle Elemente von $\mathbb{Z}$ bzw. $\mathbb{Q}$ erfüllt. Die Elemente $0_{\mathbb{Q}} = 0$ und $1_{\mathbb{Q}} = 1$ besitzen offenbar die definierende Eigenschaft von Null- und Einselement. Ist $n \in \mathbb{Z}$, dann überprüft man durch Betrachtung der Einzelfälle $n \in \mathbb{N}$, $n = 0$ und $-n \in \mathbb{N}$ leicht, dass auch $-n$ in $\mathbb{Z}$ enthalten ist. Außerdem gilt $n + (-n) = 0$. Sei nun $x \in \mathbb{Q}$, $x = \frac{a}{b}$ mit $a \in \mathbb{Z}$ und $b \in \mathbb{N}$. Dann ist auch $-x = \frac{-a}{b}$ in $\mathbb{Q}$ enthalten, und es gilt $x + (-x) = 0$. Damit haben wir gezeigt, dass $\mathbb{Z}$ und $\mathbb{Q}$ Ringe sind.

Sei $x \in \mathbb{Q}^\times$, $x = \frac{a}{b}$ mit $a \in \mathbb{Z}$ und $b \in \mathbb{N}$. Dann gilt $a \in \mathbb{N}$ oder $-a \in \mathbb{N}$. Im ersten Fall zeigt die Gleichung $x^{-1} = \frac{b}{a}$, im zweiten Fall die Gleichung $x^{-1} = \frac{(-b)}{(-a)}$, dass auch x^{-1} in $\mathbb{Q}$ enthalten ist. Wegen $x \cdot x^{-1} = 1$ besitzt x also einen Kehrwert in $\mathbb{Q}$; dies zeigt, dass $\mathbb{Q}$ ein Körper ist.

Zum Schluss zeigen wir noch, dass $\mathbb{Q}^+ = \mathbb{Q} \cap \mathbb{R}^+$ die Bedingungen für eine Anordnung aus Definition (1.1) erfüllt. Da $\mathbb{R}^+$ eine Anordnung auf $\mathbb{R}$ ist, gilt für jedes $x \in \mathbb{Q}$ genau eine der Bedingungen $x \in \mathbb{R}^+$, $x = 0$ oder $-x \in \mathbb{R}^+$. Zusammen mit $x \in \mathbb{Q}$ folgt daraus, dass genau eine der Bedingungen $x \in \mathbb{Q}^+$, $x = 0$ oder $-x \in \mathbb{Q}^+$ erfüllt ist. Sind $x, y \in \mathbb{Q}^+$, dann gilt insbesondere $x, y \in \mathbb{R}^+$. Weil $\mathbb{R}^+$ eine Anordnung auf $\mathbb{R}$ ist, folgt $x + y \in \mathbb{R}^+$ und $xy \in \mathbb{R}^+$. Wegen $x + y, xy \in \mathbb{Q}$ folgt daraus wiederum $x + y \in \mathbb{Q}^+$ und $xy \in \mathbb{Q}^+$. Die Anordnung $\mathbb{Q}^+$ ist archimedisch, denn die archimedische Bedingung aus Definition (1.9) gilt für alle $x, y \in \mathbb{R}^+$ und somit erst recht für alle $x, y \in \mathbb{Q}^+$. □

Wir bezeichnen eine Zahl $x \in \mathbb{R}$ als ***rational***, wenn sie in $\mathbb{Q}$ liegt, ansonsten als ***irrational***. Eine Teilmenge $M \subseteq \mathbb{R}$ wird als ***dichte Teilmenge*** von $\mathbb{R}$ bezeichnet, wenn jedes nicht-leere offene Intervall $I \subseteq \mathbb{R}$ ein Element aus M enthält.

(1.26) Satz Sowohl die rationalen als auch die irrationalen Zahlen bilden jeweils eine dichte Teilmenge von $\mathbb{R}$.

Beweis: Nach Definition gilt $I =]a, b[$ mit $a, b \in \bar{\mathbb{R}}$. Wir dürfen davon ausgehen, dass $a < b$ mit $a, b \in \mathbb{R}$ gilt. Denn enthält jedes solche Intervall rationale und irrationale Zahlen, dann gilt dies erst recht für Intervalle der Form $]-\infty, b[$ und $]a, +\infty[$.

Zunächst konzentrieren wir uns auf den Beweis der Aussage, dass I eine rationale Zahl enthält und führen dies auf den Fall zurück, dass $a > 0$ ist. Nehmen wir an, wir hätten den Satz in dieser Situation bereits bewiesen, und setzen wir nun $a \leq 0$ voraus. Dann können wir a und b in den positiven Bereich verschieben und dort die bereits bewiesene Aussage anwenden. Da $\mathbb{N}$ nach oben unbeschränkt ist, gibt es ein $k \in \mathbb{N}$ mit $k > -a$, und wegen $a + k, b + k > 0$ finden wir auf Grund unserer Annahme finden ein $s \in \mathbb{Q}$ mit $a + k < s < b + k$. Es folgt dann $a < r < b$ für die rationale Zahl $r = s - k$.

Setzen wir nun $0 < a < b$ voraus. Die Idee besteht darin, den Abstand zwischen x und y durch Multiplikation mit einer natürlichen Zahl so zu vergrößern, dass zwischen x und y auf jeden Fall eine natürliche Zahl liegt. Dazu definieren wir $\varepsilon = b - a$ und wählen ein auf Grund von Satz (1.10) existierendes $m \in \mathbb{N}$ mit $\frac{1}{m} < \varepsilon$. Sei nun $k \in \mathbb{N}$ minimal mit $k > ma$. Dann gilt $k - 1 \leq ma < k$ und somit

$$\frac{k-1}{m} \quad \leq \quad a \quad < \quad \frac{k}{m}.$$

Wir rechnen nach, dass auch die Ungleichung $\frac{k}{m} < b$ erfüllt ist. Tatsächlich gilt

$$b \quad = \quad a + (b-a) \quad = \quad a + \varepsilon \quad > \quad a + \frac{1}{m} \quad \geq \quad \frac{k-1}{m} + \frac{1}{m} \quad = \quad \frac{k}{m}.$$

Insgesamt gilt also $a < \frac{k}{m} < b$. Nun zeigen wir noch, dass jedes endliche, offene Intervall $I =]a, b[$ auch eine irrationale Zahl enthält. Nach eventueller Verkleinerung von I dürfen wir $0 \notin I$ annehmen. Wie bereits gezeigt, enthält $I' = \left]\frac{a}{\sqrt{2}}, \frac{b}{\sqrt{2}}\right[$ eine Zahl $r \in \mathbb{Q}$, und wegen $0 \notin I'$ ist $r \neq 0$. Die Zahl $\sqrt{2}r$ liegt dann offenbar in I. Wäre $\sqrt{2}r$ rational, dann würde auch $\sqrt{2} = (\sqrt{2}r)r^{-1} \in \mathbb{Q}$ gelten. Es ist aber bekannt, dass $\sqrt{2}$ nicht rational ist, und somit ist auch $\sqrt{2}r$ irrational. □

Es ist leicht zu sehen, dass $\mathbb{Z}$ und $\mathbb{Q}$ ***abzählbar unendlich*** sind. Zunächst sind $\mathbb{Z}$ und $\mathbb{Q}$ offenbar unendlich, da sie $\mathbb{N}$ als Teilmengen enthalten. Durch $n \mapsto -n$ ist eine Bijektion zwischen $\mathbb{N}$ und $\mathbb{N}_- = \{-n \mid n \in \mathbb{N}\}$ definiert. Also ist auch $\mathbb{N}_-$ abzählbar unendlich. Als Vereinigung der höchstens abzählbaren Mengen $\mathbb{N}$, $\{0\}$ und $\mathbb{N}_-$ ist $\mathbb{Z}$ höchstens abzählbar, insgesamt also abzählbar unendlich. Die surjektive Abbildung $\mathbb{Z} \times \mathbb{N} \to \mathbb{Q}$, $(a, b) \mapsto \frac{a}{b}$ zeigt, dass auch $\mathbb{Q}$ höchstens abzählbar und damit insgesamt abzählbar unendlich ist.

Für beliebige Körper K haben wir weiter oben Potenzen der Form a^n mit $a \in K$ und $n \in \mathbb{N}_0$ eingeführt. Dies werden wir nun verallgemeinern. Ist K ein Körper und $a \in K^\times$, so definieren wir

$$a^{-n} \quad = \quad (a^n)^{-1} \qquad \text{für} \quad n \in \mathbb{N}.$$

Insgesamt ist a^m damit für alle $m \in \mathbb{Z}$ definiert.

> **(1.27) Lemma** Die Potenzgesetze (i) bis (iii) aus Lemma (1.8) sind im Fall $a, b \neq 0_K$ für beliebige $m, n \in \mathbb{Z}$ gültig.

Beweis: Weil die Rechenregeln für $m, n \in \mathbb{N}_0$ bereits bewiesen wurden, können wir von Anfang an voraussetzen, dass mindestens einer der beiden Exponenten m, n negativ ist.

zu (i) Hier betrachten wir zunächst den Fall $m \geq 0$, $n < 0$. Darüber hinaus nehmen wir an, dass $m + n \geq 0$ gilt. Dann liegen $m + n$, $-n$ und m in $\mathbb{N}_0$. Es gilt $(m + n) + (-n) = m$, und auf Grund des bereits bewiesenen Falls gilt

deshalb $a^{m+n}a^{-n} = a^m$. Es folgt $a^{m+n} = a^m(a^{-n})^{-1} = a^m a^{-(-n)} = a^m a^n$. Ist $m+n < 0$, dann sind die Zahlen $-(m+n)$, m und $-n$ in $\mathbb{N}_0$ erhalten. Wegen $-(m+n)+m = -n$ erhalten wir $a^{-(m+n)}a^m = a^{-n}$. Daraus folgt weiter

$$a^m(a^{-n})^{-1} = (a^{-(m+n)})^{-1} \quad \Leftrightarrow \quad a^m a^{-(-n)} = a^{-(-(m+n))} \quad \Leftrightarrow \quad a^m a^n = a^{m+n}.$$

Den Fall $m < 0$, $n \geq 0$, können wir auf den vorherigen zurückführen, denn es gilt $a^{m+n} = a^{n+m} = a^n a^m = a^m a^n$. Setzen wir nun $m, n < 0$ voraus. Dann liegen $-m$, $-n$ und $-(m+n)$ in $\mathbb{N}_0$, und aus $-(m+n) = (-m)+(-n)$ folgt $a^{-(m+n)} = a^{-m}a^{-n}$. Durch die Umformungen

$$(a^{-(m+n)})^{-1} = (a^{-m}a^{-n})^{-1} \quad \Leftrightarrow \quad (a^{-(m+n)})^{-1} = (a^{-m})^{-1}(a^{-n})^{-1}$$

$$\Leftrightarrow \quad a^{-(-(m+n))} = a^{-(-m)}a^{-(-n)} \quad \Leftrightarrow \quad a^{m+n} = a^m a^n$$

sieht man, dass die Gleichung auch in dieser Situation erfüllt ist.

zu (ii) Auch hier betrachten wir zunächst den Fall $m \geq 0$, $n < 0$. Hier liegen $-n$ und $-mn$ in $\mathbb{N}_0$. Auf Grund des bereits bewiesenen Falls erhalten wir $(a^m)^n = ((a^m)^{-n})^{-1} = (a^{m(-n)})^{-1} = (a^{-mn})^{-1} = a^{-(-mn)} = a^{mn}$. Bevor wir weitere Fälle bearbeiten können, müssen wir zunächst durch vollständige Induktion die Gleichung $(a^{-1})^n a^n = 1_K$ für alle $n \in \mathbb{N}_0$ beweisen. Wegen $(a^{-1})^0 a^0 = 1_K \cdot 1_K = 1_K$ ist diese für $n = 0$ erfüllt.

Ist $n \in \mathbb{N}_0$ und die Gleichung für n vorausgesetzt, dann folgt

$$(a^{-1})^{n+1}a^{n+1} \quad = \quad (a^{-1})^n a^{-1}a^n a \quad = \quad (a^{-1})^n a^n a^{-1}a \quad = \quad 1_K \cdot 1_K \quad = \quad 1_K \quad ,$$

wobei im vorletzten Schritt die Induktionsvoraussetzung verwendet wurde. Damit ist der Induktionsbeweis abgeschlossen.

Wir beweisen nun die Gleichung $(a^m)^n = a^{mn}$ im Fall $m < 0$ und $n \in \mathbb{Z}$ und definieren dafür $c = a^{-1}$. Wie soeben gezeigt, gilt $a^{-m}c^{-m} = 1_K$, also $a^m = (a^{-m})^{-1} = c^{-m}$, und ebenso $a^{mn} = c^{-mn}$. Auf Grund des bereits bewiesenen Falls und wegen $-m \in \mathbb{N}_0$ wir

$$(a^m)^n \quad = \quad (c^{-m})^n \quad = \quad c^{(-m)n} \quad = \quad c^{-mn} \quad = \quad a^{mn}.$$

zu (iii) Hier brauchen wir nur den Fall $m < 0$ betrachten. Wegen $-m \in \mathbb{N}$ und auf Grund des bereits bewiesenen Falls erhalten wir $(ab)^m = ((ab)^{-m})^{-1} = (a^{-m})^{-1}(b^{-m})^{-1} = a^m b^m$. □

(1.28) Lemma Sei (K, K^+) ein angeordneter Körper, und sei $n \in \mathbb{N}$. Dann gilt für $a, b \in K_+$ die Äquivalenz $a < b \Leftrightarrow a^n < b^n$. Ist n ungerade, dann gilt sie sogar für alle $a, b \in K$.

Beweis: Seien $a, b \in K_+$. Im Fall $a = 0_K$ oder $b = 0_K$ ist die Äquivalenz offensichtlich, so dass wir uns auf $a, b \in K^+$ beschränken können. Es gilt

$$(b-a)\left(\sum_{k=0}^{n-1} a^{n-1-k}b^k\right) \quad = \quad \sum_{k=0}^{n-1} a^{n-1-k}b^{k+1} - \sum_{k=0}^{n-1} a^{n-k}b^k \quad = \quad \sum_{k=1}^{n} a^{n-k}b^k - \sum_{k=0}^{n-1} a^{n-k}b^k \quad =$$

$$b^n + \sum_{k=1}^{n-1} a^{n-k}b^k - a^n - \sum_{k=1}^{n-1} a^{n-k}b^k \quad = \quad b^n - a^n.$$

Mit a und b ist auch die Summe $c = \sum_{k=0}^{n-1} a^{n-1-k} b^k$ in der Klammer positiv. An der Gleichung $b^n - a^n = c(b-a)$ kann somit die Äquivalenz $b - a \in K^+ \Leftrightarrow b^n - a^n \in K^+$ abgelesen werden, und damit gilt auch $a < b \Leftrightarrow a^n < b^n$.

Sei nun n ungerade, $n = 2k+1$ mit $k \in \mathbb{N}_0$. Im Fall $a, b \geq 0_K$ wurde die Äquivalenz bereits bewiesen. Ist $a < 0_K$ und $b \geq 0_K$, dann ist wegen $a^{2k} = (a^k)^2$ auch $a^n = a^{2k}a$ negativ. Es gilt also $a < 0_K \leq b$ und $a^n < 0_K \leq b^n$. Dies zeigt, dass die Äquivalenz $a < b \Leftrightarrow a^n < b^n$ in dieser Situation immer erfüllt ist. Nun betrachten wir noch den Fall $a, b < 0_K$. Weil n ungerade ist, gilt $(-1_K)^n = (-1_K)^{2k}(-1_K) = ((-1_K)^2)^k(-1_K) = 1_K^k(-1_K) = -1_K$. Wegen $-a, -b > 0_K$ können wir die Äquivalenz auf den bereits bewiesenen Fall zurückführen, denn es gilt

$$a < b \qu\Leftrightarrow \quad -b < -a \quad \Leftrightarrow \quad (-b)^n < (-a)^n \quad \Leftrightarrow \quad (-1_K)^n b^n < (-1_K)^n a^n$$

$$\Leftrightarrow \quad (-1_K) b^n < (-1_K) a^n \quad \Leftrightarrow \quad a^n < b^n. \qquad \square$$

Wir bemerken noch, dass Lemma (1.28) richtig bleibt, wenn man das Zeichen $<$ durch eines der Symbole $>$, $\leq$, $\geq$ oder $=$ ersetzt. Bei $>$ genügt es, einfach die Rollen von a und b zu vertauschen. Für die Gleichheit argumentiert man indirekt: Die Implikation $a = b \Rightarrow a^n = b^n$ gilt offensichtlich. Andererseits folgt aus $a^n = b^n$ auch $a = b$, denn $a < b$ hätte $a^n < b^n$ zur Folge, und aus $a > b$ würde sich $a^n > b^n$ ergeben, jeweils im Widerspruch zur Voraussetzung. Damit ist auch klar, wie man die Aussagen für $\leq$ und $\geq$ herleitet.

(1.29) Satz Sei $b \in \mathbb{R}_+$ und $n \in \mathbb{N}$. Dann gibt es eine eindeutig bestimmte Zahl $\alpha \in \mathbb{R}_+$ mit $\alpha^n = b$. Sie wird mit $\sqrt[n]{b}$ bezeichnet und die ***n-te Wurzel*** von b genannt.

Beweis: Im Fall $n = 1$ oder $b = 0$ ist die Aussage offensichtlich, weil man im ersten Fall $\alpha = b$ und im zweiten $\alpha = 0$ setzen kann. Deshalb können wir $n > 1$ und $b > 0$ voraussetzen. Wir betrachten nun die Menge

$$M \quad = \quad \{x \in \mathbb{R} \mid x^n \leq b\}$$

und zeigen zunächst, dass diese ein positives Supremum besitzt. Wegen $0^n \leq b$ gilt $0 \in M$, also ist die Menge M nicht leer. Sei nun $c = \max\{1, b\}$. Wegen Lemma (1.28) und $c \geq 1$ gilt $c^n = c^{n-1}c \geq 1^{n-1}c = c$. Für jedes $x \in \mathbb{R}$ mit $x > c$ gilt folglich $x^n > c^n \geq c \geq b$ und somit $x \notin M$. Dies zeigt, dass c eine obere Schranke von M ist. Weil es sich bei M um eine nichtleere, nach oben beschränkte Menge handelt, existiert $\alpha = \sup M$. Wegen $b > 0$ enthält M auf jeden Fall positive Zahlen; wählt man beispielsweise $m \in \mathbb{N}$ so groß, dass $\frac{1}{m} < b$ ist, dann folgt

$$\left(\frac{1}{m}\right)^n \quad = \quad \frac{1}{m^n} \quad \leq \quad \frac{1}{m} \quad < \quad b$$

und somit $\frac{1}{m} \in M$. Weil α eine obere Schranke von M ist, muss also $\alpha > 0$ gelten.

Nun beweisen wir die Gleichung $\alpha^n = b$, indem wir die beiden Ungleichungen $\alpha^n > b$ und $\alpha^n < b$ einzeln widerlegen. Nehmen wir zunächst an, dass $\alpha^n > b$ gilt. In diesem Fall können wir ein $h \in \mathbb{R}^+$ so klein wählen, dass auch $(\alpha - h)^n > b$ gilt. Denn für jedes $h \in \mathbb{R}^+$ mit $h < \alpha$ ist $-\frac{h}{\alpha} > -1$. Wir können deshalb die Bernoullische Ungleichung aus Satz (1.10) anwenden und erhalten

$$(\alpha - h)^n \quad = \quad \alpha^n \left(1 - \frac{h}{\alpha}\right)^n \quad \geq \quad \alpha^n \left(1 - \frac{nh}{\alpha}\right) \quad = \quad \alpha^n - hn\alpha^{n-1}.$$

Setzen wir nun $\varepsilon = \alpha^n - b$ und wählen h so klein, dass außer $h < \alpha$ auch $hn\alpha^{n-1} < \varepsilon \Leftrightarrow h < \frac{\varepsilon}{n\alpha^{n-1}}$ erfüllt ist, dann erhalten wir

$$(\alpha - h)^n \quad \geq \quad \alpha^n - hn\alpha^{n-1} \quad > \quad \alpha^n - \varepsilon \quad = \quad \alpha^n - (\alpha^n - b) \quad = \quad b$$

wie gewünscht. Andererseits ist $\alpha - h$ keine obere Schranke von M, weil α nach Definition die kleinste obere Schranke ist. Es gibt also ein $y \in M$ mit $y > \alpha - h$. Aus $y \in M$ folgt $y^n \leq b$ und nach Lemma (1.28) somit $(\alpha - h)^n < y^n \leq b$, im Widerspruch zu $(\alpha - h)^n > b$. Dies zeigt, dass der Fall $\alpha^n > b$ ausgeschlossen ist.

Nehmen wir nun an, dass $\alpha^n < b$ gilt. In diesem Fall können wir $h \in \mathbb{R}^+$ so klein wählen, dass auch $(\alpha + h)^n < b$ noch gültig ist. Der Binomische Lehrsatz, den wir in den Übungen eingeführt haben, liefert für jedes $h \in \mathbb{R}^+$ mit $h < \alpha$ die Abschätzung

$$\begin{aligned}(\alpha + h)^n \quad &= \quad \sum_{k=0}^{n} \binom{n}{k} \alpha^{n-k} h^k \quad = \quad \alpha^n + \sum_{k=1}^{n} \binom{n}{k} \alpha^{n-k} h^k \quad = \quad \alpha^n + h \sum_{k=1}^{n} \binom{n}{k} \alpha^{n-k} h^{k-1} \\ &\leq \quad \alpha^n + h \sum_{k=1}^{n} \binom{n}{k} \alpha^{n-k} \alpha^{k-1} \quad = \quad \alpha^n + h \sum_{k=1}^{n} \binom{n}{k} \alpha^{n-1} \quad = \quad \alpha^n + h\gamma\end{aligned}$$

mit der Konstanten $\gamma = \sum_{k=1}^{n} \binom{n}{k} \alpha^{n-1} \in \mathbb{R}^+$. Wir setzen nun $\varepsilon = b - \alpha^n$ und wählen $h \in \mathbb{R}^+$ so, dass neben $h < \alpha$ auch die Abschätzung $h\gamma < \varepsilon \Leftrightarrow h < \varepsilon\gamma^{-1}$ erfüllt ist. Dann folgt

$$(\alpha + h)^n \quad \leq \quad \alpha^n + h\gamma \quad < \quad \alpha^n + \varepsilon \quad = \quad \alpha^n + (b - \alpha^n) \quad = \quad b.$$

Andererseits ist α eine obere Schranke von M und somit $\alpha + h \notin M$. Aus $\alpha + h \notin M$ folgt $(\alpha + h)^n > b$. Aber dies widerspricht unserer zuvor gefundenen Abschätzung $(\alpha + h)^n < b$. Somit scheidet auch $\alpha^n < b$ aus, und als einzige Möglichkeit bleibt $\alpha^n = b$. Es gibt kein weiteres Element $\beta \in \mathbb{R}_+$ mit $\beta^n = b$, denn wie wir im Anschluss an Lemma (1.28) festgestellt haben, folgt aus $\alpha^n = b = \beta^n$ unmittelbar $\alpha = \beta$. □

(1.30) Lemma Für $a, b \in \mathbb{R}_+$ und $m, n \in \mathbb{N}$, $p \in \mathbb{N}_0$ gilt

$$\sqrt[n]{ab} = \sqrt[n]{a} \cdot \sqrt[n]{b} \quad , \quad (\sqrt[n]{a})^p = \sqrt[n]{a^p} \qquad \text{und} \qquad \sqrt[m]{\sqrt[n]{a}} = \sqrt[mn]{a}.$$

Die zweite Gleichung gilt im Fall $a \neq 0$ für alle $p \in \mathbb{Z}$. Die beiden Zuordnungen $\mathbb{R}_+ \to \mathbb{R}_+$ gegeben durch $a \mapsto a^n$ und $a \mapsto \sqrt[n]{a}$ sind invers zueinander.

Beweis: Aus $(\sqrt[n]{a})^n = a$ und $(\sqrt[n]{b})^n = b$ folgt $(\sqrt[n]{a} \cdot \sqrt[n]{b})^n = (\sqrt[n]{a})^n (\sqrt[n]{b})^n = ab$. Also ist $\gamma = \sqrt[n]{a}\sqrt[n]{b}$ die eindeutig bestimmte Zahl in $\mathbb{R}_+$ mit $\gamma^n = ab$. Damit ist die erste Gleichung bewiesen. Ebenso zeigt die Rechnung

$$\left(\sqrt[m]{\sqrt[n]{a}}\right)^{mn} \quad = \quad \left(\left(\sqrt[m]{\sqrt[n]{a}}\right)^m\right)^n \quad = \quad \left(\sqrt[n]{a}\right)^n \quad = \quad a \quad ,$$

dass $\gamma = \sqrt[m]{\sqrt[n]{a}}$ die eindeutig bestimmte Zahl mit $\gamma^{mn} = a$ ist. Daraus folgt die dritte Gleichung. Die zweite Gleichung kann durch vollständige Induktion über p bewiesen werden. Für $p = 0$ ist sie offenbar erfüllt, denn es gilt $(\sqrt[n]{a})^0 = 1$, und wegen $1^n = 1$ gilt $\sqrt[n]{a^0} = \sqrt[n]{1} = 1$, insgesamt also $(\sqrt[n]{a})^0 = \sqrt[n]{a^0}$. Ist $p \in \mathbb{N}_0$ und setzen wir die Formel dieses p voraus, dann folgt

$$(\sqrt[n]{a})^{p+1} \quad = \quad (\sqrt[n]{a})^p \sqrt[n]{a} \quad = \quad \sqrt[n]{a^p}\sqrt[n]{a} \quad = \quad \sqrt[n]{a^p \cdot a} \quad = \quad \sqrt[n]{a^{p+1}}.$$

Beweisen wir die Gleichung nun noch im Fall $a \neq 0$ und $p < 0$. Zunächst zeigen wir für $c \in \mathbb{R}^\times$ und $n \in \mathbb{N}$ die Gleichung

$$\sqrt[n]{c^{-1}} \quad = \quad \left(\sqrt[n]{c}\right)^{-1}.$$

Wegen $1^n = 1$ gilt $\sqrt[n]{1} = 1$, und daraus folgt $\sqrt[n]{c}\sqrt[n]{c^{-1}} = \sqrt[n]{cc^{-1}} = \sqrt[n]{1} = 1$. Multipliziert man nun beide Seiten mit $(\sqrt[n]{c})^{-1}$, so erhält man das gewünschte Ergebnis. Anwendung dieser Gleichung auf $c = a^{-p}$ liefert nun

$$\sqrt[n]{a^p} \quad = \quad \sqrt[n]{(a^{-p})^{-1}} \quad = \quad \left(\sqrt[n]{a^{-p}}\right)^{-1} \quad = \quad \left(\left(\sqrt[n]{a}\right)^{-p}\right)^{-1} \quad = \quad \left(\sqrt[n]{a}\right)^p.$$

Schließlich zeigt die triviale Gleichung $a^n = a^n$, dass $\gamma = a$ das eindeutig bestimmte Element mit $\gamma^n = a^n$ sein muss. Daraus folgt $\sqrt[n]{a^n} = a$. Nach Definition der n-ten Wurzel gilt andererseits $(\sqrt[n]{a})^n = a$. Dies zeigt, dass die angegebenen Zuorndungen zueinander invers sind. □

Mit Hilfe der n-ten Wurzeln können wir Potenzen a^r für $a \in \mathbb{R}^+$ beliebige ***rationale Exponenten*** r definieren. Besitzt r die Darstellung $r = \frac{m}{n}$ mit $m \in \mathbb{Z}$ und $n \in \mathbb{N}$, dann setzen wir $a^r = \sqrt[n]{a^m}$. Die Definition ist unabhängig von der Darstellung der rationalen Zahl r als Bruch. Sind nämlich $p \in \mathbb{Z}$ und $q \in \mathbb{N}$ weitere Zahlen mit $r = \frac{p}{q}$, dann folgt $\frac{m}{n} = \frac{p}{q} \Leftrightarrow mq = np$ und

$$\sqrt[n]{a^m} \quad = \quad \sqrt[n]{\sqrt[p]{(a^m)^p}} \quad = \quad \sqrt[np]{a^{mp}} \quad = \quad \sqrt[mq]{a^{mp}} \quad = \quad \sqrt[q]{\sqrt[m]{(a^p)^m}} \quad = \quad \sqrt[q]{a^p}.$$

(1.31) Lemma Die Potenzgesetze (i) bis (iii) aus Lemma (1.8) gelten auch für $a, b \in \mathbb{R}^+$ und beliebige Exponenten $r, s \in \mathbb{Q}$.

Beweis: Wegen $r, s \in \mathbb{Q}$ gibt es $m, p \in \mathbb{Z}$ und $n, q \in \mathbb{N}$ mit $r = \frac{m}{n}$ und $s = \frac{p}{q}$. Zunächst beweisen wir die Gleichung $a^{r+s} = a^r a^s$. Aus $r + s = \frac{mq+np}{nq}$ folgt

$$\begin{aligned} a^r a^s \quad &= \quad \sqrt[n]{a^m}\sqrt[q]{a^p} \quad = \quad \sqrt[n]{\sqrt[q]{(a^m)^q}} \cdot \sqrt[q]{\sqrt[n]{(a^p)^n}} \quad = \quad \sqrt[nq]{a^{mq}}\,\sqrt[nq]{a^{np}} \\ &= \quad \sqrt[nq]{a^{mq+np}} \quad = \quad a^{r+s}. \end{aligned}$$

Die Gleichung $(a^r)^s = a^{rs}$ erhält man durch die Rechnung

$$\begin{aligned} (a^r)^s \quad &= \quad (a^{\frac{m}{n}})^{\frac{p}{q}} \quad = \quad \left(\sqrt[n]{a^m}\right)^{\frac{p}{q}} \quad = \quad \sqrt[q]{\left(\sqrt[n]{a^m}\right)^p} \quad = \quad \sqrt[q]{\sqrt[n]{a^{mp}}} \\ &= \quad \sqrt[nq]{a^{mp}} \quad = \quad a^{\frac{mp}{nq}} \quad = \quad a^{rs} \quad , \end{aligned}$$

und die Gleichung $(ab)^r = a^r b^r$ erhält man durch

$$\begin{aligned} (ab)^r \quad &= \quad (ab)^{\frac{m}{n}} \quad = \quad \sqrt[n]{(ab)^m} \quad = \quad \sqrt[n]{a^m b^m} \quad = \quad \sqrt[n]{a^m}\sqrt[n]{b^m} \\ &= \quad a^{\frac{m}{n}} b^{\frac{m}{n}} \quad = \quad a^r b^r. \end{aligned}$$

□

(1.32) Definition Eine ***Bewertung*** auf einem Körper K ist eine Abbildung $|\cdot| : K \to \mathbb{R}_+$, so dass für alle $x, y \in K$ die folgenden Bedingungen erfüllt sind.

(i) $x = 0 \Leftrightarrow |x| = 0$
(ii) $|xy| = |x||y|$
(iii) $|x + y| \leq |x| + |y|$ *(Dreiecksungleichung)*

Das Paar $(K, |\cdot|)$ bezeichnet man dann als ***bewerteten Körper***.

Wir definieren eine Abbildung $|\cdot| : \mathbb{R} \to \mathbb{R}_+$ durch

$$x \mapsto \begin{cases} x & \text{falls} \quad x \geq 0 \\ -x & \text{falls} \quad x < 0. \end{cases}$$

Man bezeichnet sie als ***Absolutbetrag*** auf den reellen Zahlen.

(1.33) Lemma Für jedes $x \in \mathbb{R}$ gilt $x \leq |x|$ und $|x| = |-x|$.

Beweis: Im Fall $x = 0$ sind die Gleichungen $|x| = 0 = |-x|$ und die Ungleichung $x \leq |x|$ offenbar gültig. Ist x positiv, dann gilt nach Definition $|x| = x$ und $|-x| = x$, und die Ungleichung ist wegen $x \leq x$ erfüllt. Ist dagegen x negativ, dann gilt $|x| = -x$ und $|-x| = -x$. Da x eine negative und $-x$ eine positive Zahl ist, gilt auch hier $x \leq -x = |x|$. □

(1.34) Satz Durch den Absolutbetrag ist eine Bewertung auf $\mathbb{R}$ definiert.

Beweis: Zu überprüfen sind die Eigenschaften (i), (ii) und (iii) einer Bewertung.

zu (i) „⇒“ Nach Definition gilt $|0| = 0$. „⇐“ Ist $|x| = 0_{\mathbb{R}}$, dann gilt $x = 0$ oder $-x = 0$, also $x = 0$ in beiden Fällen.

zu (ii) Man zeigt die Gleichung durch Betrachtung der vier Einzelfälle

(a) $x, y \geq 0$ (b) $x \geq 0, y < 0$ (c) $x < 0, y \geq 0$ (d) $x, y < 0$.

Wir beschränken uns auf Fall (b). Nach Definition gilt $|x| = x$ und $|y| = -y$. Im Fall $x = 0$ gilt $|x||y| = |0||y| = 0 \cdot |y| = 0$ und $|xy| = |0| = 0$. Sei nun $x > 0$ vorausgesetzt. Dann gilt $-xy = x(-y) > 0$, also $xy < 0$ und deshalb $|xy| = -xy$. Es folgt $|xy| = -xy = x(-y) = |x||y|$, die Gleichung ist im Fall (b) also erfüllt.

zu (iii) Wie für alle Körperelemente gilt $x \leq |x|$ und $y \leq |y|$. Aus den Anordnungsaxiomen folgt $x + y \leq x + |y| \leq |x| + |y|$. Ebenso gilt $-x \leq |-x| = |x|$ und $-y \leq |-y| = |y|$ und somit $-(x + y) \leq |x| + |y|$. Da $|x + y|$ gleich $x + y$ oder gleich $-(x + y)$ ist, ist die Ungleichung $|x + y| \leq |x| + |y|$ somit auf jeden Fall erfüllt. □

Wir werden nun noch ein weiteres Beispiel für einen bewerteten Körper kennenlernen. Der Beweis des folgenden Satzes ist Gegenstand der Algebra-Vorlesung.

(1.35) Satz Es gibt einen Körper $(\mathbb{C}, +_{\mathbb{C}}, \cdot_{\mathbb{C}})$ mit folgenden Eigenschaften.

(i) Es gilt $\mathbb{C} \supseteq \mathbb{R}$, und für alle $a, b \in \mathbb{R}$ gilt $a +_{\mathbb{C}} b = a + b$ und $a \cdot_{\mathbb{C}} b = ab$.

(ii) Es gibt ein Element $i \in \mathbb{C}$ mit $i^2 = -1$.

(iii) Jedes Element $z \in \mathbb{C}$ kann auf eindeutige Weise in der Form $z = a + ib$ mit $a, b \in \mathbb{R}$ dargestellt werden.

Man nennt $\mathbb{C}$ den Körper der ***komplexen Zahlen***, und das Element i die ***imaginäre Einheit***. Ist $z \in \mathbb{C}$, $z = a + ib$ mit $a, b \in \mathbb{R}$, so nennt man a den ***Realteil*** $\mathrm{Re}(z)$ und b den ***Imaginärteil*** $\mathrm{Im}(z)$ von z. Wieder verwenden wir zur Vereinfachung der Notation die Symbole $+$ und $\cdot$ statt $+_{\mathbb{C}}$ und $\cdot_{\mathbb{C}}$ zur Bezeichnung der Addition und Multiplikation auf $\mathbb{C}$.

Aus der Definition von $\mathbb{C}$ und der Tatsache, dass es sich um einen Körper handelt, kann direkt abgeleitet werden, nach welchen Regeln die Addition, Multiplikation, Negative- und Kehrwertbildung zu erfolgen hat. Seien $z, w \in \mathbb{C}$, $z = a + ib$ und $w = c + id$ mit $a, b, c, d \in \mathbb{R}$. Dann gilt

$$z + w \quad = \quad (a + ib) + (c + id) \quad = \quad (a + c) + i(b + d) \quad ,$$

und das Produkt erhält man durch

$$\begin{aligned} zw \quad = \quad (a + ib)(c + id) \quad = \quad ac + ibc + iad + i^2bd \quad = \quad ac + (-1)bd + ibc + iad \\ = \quad (ac - bd) + i(bc + ad). \end{aligned}$$

An den Gleichungen $0 + z = 0 + (a + ib) = (a + 0) + ib = a + ib = z$ und $1 \cdot z = 1(a + ib) = (1a) + i(1b) = a + ib = z$ sieht man, dass Null- und Einselement von $\mathbb{C}$ durch $0_{\mathbb{C}} = 0$ und $1_{\mathbb{C}} = 1$ gegeben sind. Den Kehrwert erhält man im Fall $z \neq 0$ durch die Rechnung

$$\begin{aligned} z^{-1} \quad = \quad \frac{1}{a + ib} \quad = \quad \frac{a - ib}{(a + ib)(a - ib)} \quad = \quad \frac{a - ib}{a^2 + iab - iab - i^2b^2} \quad = \\ \frac{a - ib}{a^2 + iab - iab - (-1)b^2} \quad = \quad \frac{a - ib}{a^2 + b^2} \quad = \quad \frac{a}{a^2 + b^2} + i\frac{(-b)}{a^2 + b^2}. \end{aligned}$$

(1.36) Definition Für jedes $z \in \mathbb{C}$, $z = a + ib$ mit $a, b \in \mathbb{R}$ nennt man $\bar{z} = a - ib$ das zu z ***konjugiert komplexe*** Element. Die Abbildung $\mathbb{C} \to \mathbb{C}, z \mapsto \bar{z}$ bezeichnet man dann als komplexe Konjugation.

Wir beweisen für die komplexe Konjuation einige einfache Rechenregeln.

(1.37) Proposition Für $z, w \in \mathbb{C}$ gilt

(i) $\bar{\bar{z}} = z$ (ii) $\overline{z + w} = \bar{z} + \bar{w}$ (iii) $\overline{zw} = \bar{z}\bar{w}$.

Beweis: Wir schreiben z und w in der Form $z = a + ib$ und $w = c + id$, mit $a, b, c, d \in \mathbb{R}$.

zu (i) Diese Gleichung ergibt sich durch die Rechnung

$$\bar{\bar{z}} \;=\; \overline{\overline{a+ib}} \;=\; \overline{a-ib} \;=\; \overline{a+i(-b)} \;=\; a+i(-(-b)) \;=\; a+ib \;=\; z.$$

zu (ii) Es gilt $z + w = (a + c) + i(b + d)$ und somit

$$\overline{z+w} \;=\; \overline{(a+c)+i(b+d)} \;=\; (a+c)-i(b+d) \;=\; (a-ib)+(c-id) \;=\; \bar{z}+\bar{w}.$$

zu (iii) Wegen $zw = (a + ib)(c + id) = (ac - bd) + i(ad + bc)$ gilt einerseits

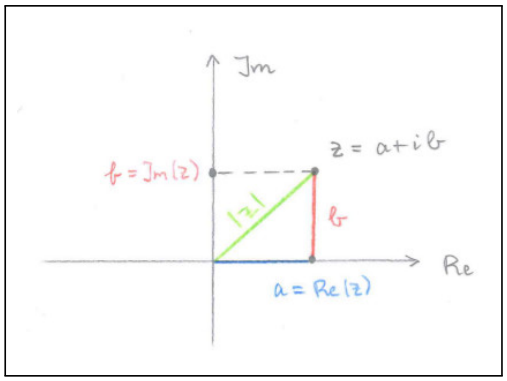
$$\overline{zw} \;=\; (ac-bd)-i(ad+bc).$$

Andererseits ist $\bar{z}\bar{w} = (a - ib)(c - id) = (ac - bd) + i(-bc - ad)$. Also ist $\overline{zw} = \bar{z}\bar{w}$ erfüllt. □

(1.38) Definition Sei $z = a + ib \in \mathbb{C}$ mit $a, b \in \mathbb{R}$. Dann nennt man

$$|z| \;=\; \sqrt{a^2+b^2} \;=\; \sqrt{(a+ib)(a-ib)} \;=\; \sqrt{z\bar{z}}$$

den komplexen ***Absolutbetrag*** von z. Durch ihn ist eine Bewertung auf $\mathbb{C}$ definiert.



Der komplexe Absolutbetrag $|z|$ ist der Abstand des Punktes z vom Nullpunkt.

Beweis: Wieder sind die Eigenschaften (i), (ii) und (iii) einer Bewertung zu überprüfen.

zu (i) „⇒“ Die Darstellung der Null durch Real- und Imaginärteil ist gegen durch $0 + i0$, und folglich gilt $|0| = \sqrt{0^2 + 0^2} = 0$. „⇐“ Sei $z = a + ib$ mit $a, b \in \mathbb{R}$. Ist $|z| = 0$, dann folgt $\sqrt{a^2 + b^2} = 0$ und somit $a^2 + b^2 = 0$. Da Quadrate von Elementen in $\mathbb{R}$ entweder positiv oder gleich Null sind, folgt $a = b = 0$ und damit $z = 0$.

zu (ii) Wegen (1.37) gilt $|zw|^2 = (zw)\overline{zw} = z\bar{z}w\bar{w} = |z|^2|w|^2$. Durch Wurzelziehen auf beiden Seiten erhalten wir $|zw| = |z||w|$.

zu (iii) Seien $z, w \in \mathbb{C}$ vorgegeben. Es gilt die Äquivalenz

$$|z+w| \le |z|+|w| \quad \Leftrightarrow \quad |z+w|^2 \le (|z|+|w|)^2 \quad \Leftrightarrow \quad |z+w|^2 \le |z|^2+2|z||w|+|w|^2 \quad \Leftrightarrow$$
$$(z+w)(\bar{z}+\bar{w}) \le z\bar{z}+w\bar{w}+2|z||w| \quad \Leftrightarrow \quad z\bar{z}+w\bar{z}+z\bar{w}+w\bar{w} \le z\bar{z}+w\bar{w}+2|z||w|$$
$$\Leftrightarrow \quad w\bar{z}+z\bar{w} \le 2|z||w|.$$

Stellen wir $w\bar{z}$ durch Real- und Imaginärteil dar, $w\bar{z} = a+ib$ mit $a, b \in \mathbb{R}$, dann ist die linke Seite der Ungleichung

$$w\bar{z}+z\bar{w} \quad = \quad w\bar{z}+\overline{w\bar{z}} \quad = \quad (a+ib)+(a-ib) \quad = \quad 2a$$

und auf der rechten Seite steht $2|z||w| = 2|\bar{z}||w| = 2|w\bar{z}| = 2\sqrt{a^2+b^2}$. Zu zeigen ist also $2a \le 2\sqrt{a^2+b^2}$, was zu $a \le \sqrt{a^2+b^2}$ äquivalent ist. Aber diese Gleichung ist offenbar erfüllt, denn es gilt $a \le |a| = \sqrt{a^2} \le \sqrt{a^2+b^2}$. □

Der Körper $\mathbb{C}$ der komplexen Zahlen unterscheidet sich von $\mathbb{R}$ in einigen wichtigen Punkten.

(1.39) Satz Auf dem Körper $\mathbb{C}$ gibt es ***keine Anordnung***. Es gibt also keine Teilmenge $\mathbb{C}^+ \subseteq \mathbb{C}$, so dass die Bedingungen (i) und (ii) aus Definition (1.1) erfüllt sind.

Beweis: Angenommen, $\mathbb{C}^+$ ist eine Teilmenge von $\mathbb{C}$ mit den Eigenschaften (i) und (ii). Nach Lemma (1.2) (i) sind Quadrate von Elementen ungleich Null stets positiv, es gilt also $-1 = i^2 \in \mathbb{C}^+$. Andererseits ist nach Proposition (1.5) (ii) das Einselement 1 positiv und -1 damit negativ. Auf Grund von Eigenschaft (i) einer Anordnung kann ein Element nicht zugleich positiv und negativ sein. Dieser Widerspruch zeigt, dass eine Menge $\mathbb{C}^+$ mit den Eigenschaften (i) und (ii) nicht existiert. □

Ein weiterer Unterschied zwischen $\mathbb{R}$ und $\mathbb{C}$ besteht darin, dass man in $\mathbb{C}$ aus beliebigen Elementen eine Quadratwurzeln ziehen kann; für jedes $z \in \mathbb{C}$ gibt es also ein $w \in \mathbb{C}$ mit $w^2 = z$. Beispielsweise sind i und $\sqrt{2}i$ die Quadratwurzeln von -1 und -2, denn es gilt $i^2 = -1$ und $(\sqrt{2}i)^2 = \sqrt{2}^2 i^2 = 2(-1) = -2$. Auch das Element i besitzt wieder eine Quadratwurzel, denn es gilt

$$\left(\frac{1}{\sqrt{2}}+\frac{i}{\sqrt{2}}\right)^2 \quad = \quad \left(\frac{1}{\sqrt{2}}\right)^2 + 2\cdot\frac{1}{\sqrt{2}}\cdot\frac{i}{\sqrt{2}} + \left(\frac{i}{\sqrt{2}}\right)^2 \quad = \quad \tfrac{1}{2}+i+(-\tfrac{1}{2}) \quad = \quad i.$$

Allgemeiner kann man beweisen

(1.40) Satz In den komplexen Zahlen besitzt jede Gleichung der Form

$$a_n z^n + a_{n-1} z^{n-1} + \ldots + a_2 z^2 + a_1 z + a_0 \quad = \quad 0$$

mit $n \in \mathbb{N}$, $a_0, a_1, \ldots, a_n \in \mathbb{C}$ und $a_n \neq 0$ mindestens eine Lösung.

Diese Eigenschaft des Körpers $\mathbb{C}$ bezeichnet man als ***algebraische Abgeschlossenheit***. Der Satz ist Thema der Algebra- und der Funktionentheorie-Vorlesung. In den reellen Zahlen ist die entsprechende Aussage falsch, denn beispielsweise besitzt die Gleichung $x^2 + 1 = 0$ keine reelle Lösung.

§ 2. *Konvergenz von Folgen*

Inhaltsübersicht

Intuitiv bedeutet die *Konvergenz* einer Folge $(a_n)_{n\in\mathbb{N}}$ reeller oder komplexer Zahlen gegen einen *Grenzwert* a, dass sich die Folgenglieder a_n mit wachsendem n dem Wert a immer weiter annähern. Präziser lässt sich dies folgendermaßen formulieren: Für jedes noch so kleine $\varepsilon \in \mathbb{R}^+$ gibt es einen Index N, ab dem alle Folgenglieder a_n von a einen Abstand kleiner als ε haben. Beispielsweise gibt es einen Index N, ab dem jedes a_n und der Grenzwert a weniger als $\varepsilon = 10^{-6}$ Längeneinheiten auseinanderliegen, und einen (eventuell größeren) Index N_1, ab dem die Abweichung sogar nur noch $\varepsilon_1 = 10^{-12}$ beträgt.

Nach der Formulierung des Konvergenzbegriffs untersuchen wir, unter welchen Operationen, die auf Folgen angwendet werden können, die Konvergenzeigenschaft erhalten bleibt. Beispielsweise werden wir zeigen, dass für zwei konvergente Folgen $(a_n)_{n\in\mathbb{N}}$ und $(b_n)_{n\in\mathbb{N}}$ mit Grenzwerten a und b auch die Folge $(a_n + b_n)_{n\in\mathbb{N}}$ konvergiert, und zwar gegen den Wert $a+b$. Außerdem werden wir die sog. *uneigentliche* Konvergenz gegen die unendlichen Werte $\pm\infty$ definieren.

Wichtige Begriffe und Sätze

- Folge in einem (bewerteten) Körper $\mathbb{K}$, zum Beispiel $\mathbb{R}$ oder $\mathbb{C}$
- Konvergenz und Grenzwert einer solchen Folge, Divergenz
- Konvergente Folgen sind beschränkt.
- Grenzwertsätze und Sandwich-Lemma
- uneigentliche Konvergenz (gegen $-\infty$ oder $+\infty$)
- Umgebung eines Punkts $a \in \bar{\mathbb{R}}$

Im gesamten Abschnitt bezeichnet $\mathbb{K}$ einen der Körper $\mathbb{R}$ oder $\mathbb{C}$. Wie wir bereits in Abschnitt § 4 definiert haben, ist eine ***Folge*** $(a_n)_{n\in\mathbb{N}}$ in $\mathbb{K}$ eine Familie reeller oder komplexer Zahlen mit $\mathbb{N}$ als Indexmenge, also eine Abbildung $\mathbb{N} \to \mathbb{K}$. Gelegentlich bietet es sich an, eine Folge nicht bei $n = 1$, sondern bei einer anderen Zahl $n_0 \in \mathbb{Z}$ beginnen zu lassen. Auch Familien mit der Indexmenge $\{n \in \mathbb{Z} \mid n \geq n_0\}$ nennen wir Folgen und verwenden in diesem Fall die Bezeichnung $(a_n)_{n\geq n_0}$. Wir betrachten einige konkrete Beispiele für Folgen reeller und komplexer Zahlen.

(i) $(a)_{n\in\mathbb{N}} = (a, a, a, a, \ldots), \quad a \in \mathbb{K}$ *(konstante Folge)*

(ii) $(\frac{1}{n})_{n\in\mathbb{N}} = (1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots)$

(iii) $(\frac{n}{n+1})_{n\in\mathbb{N}} = (\frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \frac{4}{5}, \ldots)$

(iv) $(\frac{n}{2^n})_{n\in\mathbb{N}} = (\frac{1}{2}, \frac{2}{4}, \frac{3}{8}, \frac{4}{16}, \frac{5}{32}, \frac{6}{64}, \frac{7}{128}, \ldots)$

(v) $(b^n)_{n\in\mathbb{N}_0} = (1, b, b^2, b^3, b^4, b^5, \ldots), \quad b \in \mathbb{C}$

Setzt man im letzten Beispiel $b = -1$, dann erhält man die Folge $(1, -1, 1, -1, 1, -1, 1, \ldots)$, für $b = 3$ kommt entsprechend $(1, 3, 9, 27, 81, 243, 729, \ldots)$ und für $b = i$ die Folge

$$(1, i, -1, -i, 1, i, -1, -i, 1, i, -1, -i, 1, \ldots)$$

heraus. Im letzten Beispiel wiederholen sich die Folgenglieder periodisch in Viererschritten, weil $i^4 = 1$ ist.

Wie wir anhand der Fakultätsfunktion in § 4 bereits gesehen haben, können Folgen auch ***rekursiv*** definiert werden. Beispielsweise erhält man durch die Rekursionsvorschrift $a_0 = 1$ und $a_{n+1} = (n+1)a_n$ die Folge $a_n = n!$, gegeben durch die Fakultätsfunktion. Durch die Vorschrift $f_0 = f_1 = 1$ und $f_{n+2} = f_n + f_{n+1}$ für $n \geq 0$ erhält man die bekannte ***Fibonacci-Folge***, deren erste Folgenglieder durch

$$(1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, 233, \ldots)$$

gegeben sind. Beispielsweise ist $f_2 = f_0 + f_1 = 1 + 1 = 2$, $f_3 = f_1 + f_2 = 1 + 2 = 3$, $f_4 = f_2 + f_3 = 2 + 3 = 5$, und nach dem gleichen Schema berechnet man alle weiteren Folgenglieder.

Bevor wir zur zentralen Definition dieses Abschnitts kommen, schicken wir ein paar Anmerkungen voraus. Sei dazu $(a_n)_{n\in\mathbb{N}}$ eine Folge in $\mathbb{K}$, außerdem $a \in \mathbb{K}$ und $\varepsilon \in \mathbb{R}^+$.

(i) Die Ungleichung $|a_n - a| < \varepsilon$ bedeutet, dass der ***Abstand*** zwischen dem Wert a und dem Folgenglied a_n kleiner als ε ist. Da wir uns unter ε im Allgemeinen einer „sehr kleine“ Zahl vorstellen, bedeutet die Ungleichung also, dass a_n „sehr nahe“ bei a liegt.

(ii) Im Fall $\mathbb{K} = \mathbb{C}$ ist dieser Abstand einfach der Abstand der beiden komplexen Zahlen a_n und a in der Ebene, wenn wir wie in § 8 komplexe Zahlen als Punkte in der Ebene betrachten, wobei der Realteil als x- und der Imaginärteil als y-Koordinate aufgefasst wird.

(iii) Im Fall $\mathbb{K} = \mathbb{R}$ ist die Ungleichung $|a_n - a| < \varepsilon$ gleichbedeutend mit $a_n - a < \varepsilon$ und $-(a_n - a) = a - a_n < \varepsilon$, also mit $a - \varepsilon < a_n < a + \varepsilon$. Anschaulich bedeutet die Ungleichung, dass a_n und a auf dem Zahlenstrahl einen Abstand kleiner als ε haben.

(iv) Sei nun $N \in \mathbb{N}$. Die Aussage „$|a_n - a| < \varepsilon$ für alle $n \geq N$“ bedeutet, dass die Ungleichung $|a_n - a| < \varepsilon$ für alle Folgenglieder a_n ***ab dem Index*** N erfüllt ist. Die Ungleichung gilt also für a_N, a_{N+1}, a_{N+2} usw., aber nicht unbedingt für a_{N-1}.

(2.1) Definition Sei $(a_n)_{n\in\mathbb{N}}$ eine Folge in $\mathbb{K}$ und $a \in \mathbb{K}$. Wir sagen, die Folge ***konvergiert*** gegen a und schreiben

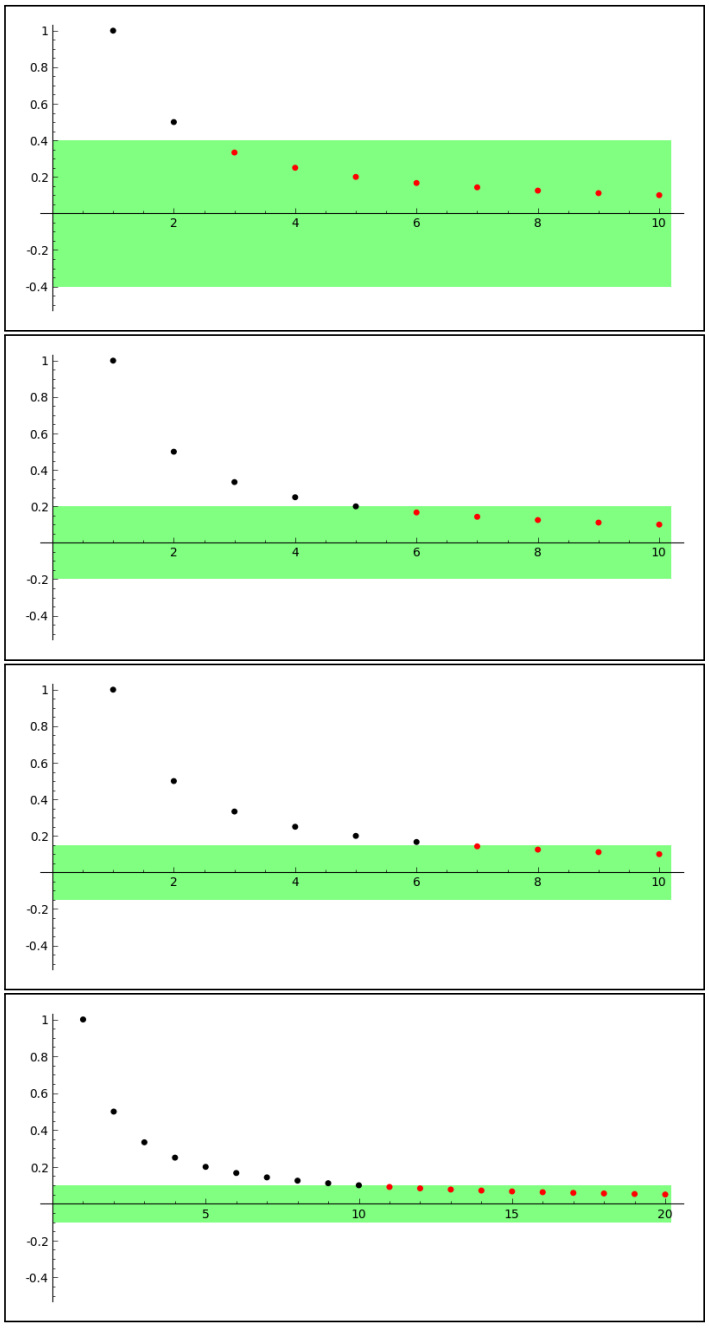
$$\lim_{n\to\infty} a_n \quad = \quad a \quad ,$$

wenn für jedes $\varepsilon \in \mathbb{R}^+$ ein $N \in \mathbb{N}$ existiert, so dass $|a_n - a| < \varepsilon$ für alle $n \geq N$ erfüllt ist. Man sagt in diesem Fall, a ist ein ***Grenzwert*** der Folge $(a_n)_{n\in\mathbb{N}}$ und bezeichnet die Folge als ***konvergent***. Folgen, die keinen Grenzwert besitzen, werden ***divergente*** Folgen genannt.

In der formalen Quantorenschreibweise lautet die Bedingung für die Konvergenz

$$\forall \varepsilon \in \mathbb{R}^+ : \exists N \in \mathbb{N} : \forall n \in \mathbb{N}, n \geq N : |a_n - a| < \varepsilon.$$

An Stelle von $\lim\limits_{n\to\infty} a_n$ verwenden wir auch die einfachere Schreibweise $\lim_n a_n$.



Veranschaulichung der Konvergenz einer Folge

Egal, wie klein der ε-Streifen $a - \varepsilon < a < a + \varepsilon$ (grün) um den Grenzwert a herum gewählt wird, es liegen immer nur endlich viele Folgenglieder (schwarz) außerhalb des Streifens, und die übrigen unendlich vielen (rot) innerhalb davon.

Wir betrachten nun einige Beispiele für konvergente und divergente Folgen.

(i) Für die Folge $(a_n)_{n\in\mathbb{N}}$ mit $a_n = \frac{1}{n}$ für alle $n \in \mathbb{N}$ gilt $\lim_n a_n = 0$. Sei nämlich $\varepsilon > 0$ beliebig vorgegeben. Nach Satz (1.10) gibt es ein $N \in \mathbb{N}$ mit $\frac{1}{N} < \varepsilon$. Für jedes $n \geq N$ folgt $|a_n - 0| = \frac{1}{n} \leq \frac{1}{N} < \varepsilon$.

(ii) Sei $a \in \mathbb{K}$. Für die konstante Folge $(a_n)_{n\in\mathbb{N}}$ gegeben durch $a_n = a$ gilt $\lim_n a_n = a$. Sei nämlich ein $\varepsilon \in \mathbb{R}^+$ vorgegeben und $N = 1$. Für alle $n \geq N$ gilt $|a_n - a| = |a - a| = 0 < \varepsilon$.

(iii) Sei $a_n = \frac{n}{n+1}$ für alle $n \in \mathbb{N}$. Dann gilt $\lim_n a_n = 1$. Für ein vorgegebenes $\varepsilon \in \mathbb{R}^+$ sei nämlich N eine natürliche Zahl mit $\frac{1}{N} < \varepsilon$. Für beliebige $n \geq N$ folgt dann

$$|a_n - 1| \quad = \quad \left|\tfrac{n}{n+1} - 1\right| \quad = \quad \left|\tfrac{n-(n+1)}{n+1}\right| \quad = \quad \tfrac{1}{n+1} \quad < \quad \tfrac{1}{N} \quad < \quad \varepsilon.$$

(iv) Für $a_n = \frac{n}{2^n}$ gilt $\lim_n a_n = 0$. Um das zu sehen, zeigt man zunächst durch vollständige Induktion, dass $2^n \geq n^2$ für alle $n \geq 4$ erfüllt ist. Für $n = 4$ ist $2^4 = 16 = 4^2$, also ist die Ungleichung offensichtlich erfüllt. Sei nun $n \in \mathbb{N}$ und $2^n \geq n^2$ bereits bewiesen. Auf Grund der Voraussetzung $n \geq 4$ gilt

$$\begin{aligned}(n-1)^2 \geq 2 \quad &\Leftrightarrow \quad n^2 - 2n + 1 \geq 2 \quad \Leftrightarrow \quad n^2 \geq 2n + 1 \\ &\Leftrightarrow \quad 2n^2 \geq n^2 + 2n + 1 \quad \Leftrightarrow \quad 2n^2 \geq (n+1)^2 \quad ,\end{aligned}$$

und die Induktionsvoraussetzung liefert $2^{n+1} = 2 \cdot 2^n \geq 2 \cdot n^2 \geq (n+1)^2$. Sei nun $\varepsilon > 0$ vorgegeben. Dann wählen wir $N_0 \in \mathbb{N}$ mit $\frac{1}{N_0} < \varepsilon$ und setzen $N = \max\{4, N_0\}$. Für alle $n \geq N$ gilt dann

$$|a_n - 0| \quad = \quad \frac{n}{2^n} \quad = \quad \frac{n^2}{2^n} \cdot \frac{1}{n} \quad \leq \quad \frac{1}{n} \quad \leq \quad \frac{1}{N_0} \quad < \quad \varepsilon.$$

(v) Die Folge $(a_n)_{n\in\mathbb{N}}$ gegeben durch $a_n = (-1)^n$ für alle $n \in \mathbb{N}$ ist divergent. Denn angenommen, $a \in \mathbb{R}$ wäre ein Grenzwert der Folge. Dann gibt es für $\varepsilon = 1$ ein $N \in \mathbb{N}$, so dass $|a_n - a| < \varepsilon$ für alle $n \geq N$ erfüllt ist. Für gerades $n \geq N$ erhalten wir $|a - 1| = |a - a_n| < 1$, insbesondere $-(a-1) < 1 \Leftrightarrow 1 - a < 1 \Leftrightarrow a > 0$. Setzen wir dagegen ein ungerades $n \geq N$ ein, so erhalten wir $|a - (-1)| = |a - a_n| < 1$, also insbesondere $a + 1 < 1 \Leftrightarrow a < 0$. Beide Ungleichungen können aber unmöglich erfüllt sein. Also hat uns die Annahme, dass ein Grenzwert a existiert, zu einem Widerspruch geführt.

(2.2) Definition Wir nennen eine Folge $(a_n)_{n\in\mathbb{N}}$ ***beschränkt***, wenn die Teilmenge von $\mathbb{R}_+$ gegeben durch $\{|a_n| \mid n \in \mathbb{N}\}$ beschränkt ist.

(2.3) Satz Jede konvergente Folge $(a_n)_{n\in\mathbb{N}}$ in $\mathbb{K}$ ist beschränkt.

Beweis: Sei a der Grenzwert der Folge. Die Konvergenzbedingung, angewendet auf $\varepsilon = 1$, liefert ein $N \in \mathbb{N}$ mit $|a_n - a| < 1$ für alle $n \geq N$. Für den Betrag von a_n folgt daraus

$$|a_n| \quad = \quad |(a_n - a) + a| \quad \leq \quad |a_n - a| + |a| \quad < \quad 1 + |a| \qquad \text{für alle} \quad n \geq N.$$

Sei nun $M = \max\{|a_1|, |a_2|, \ldots, |a_{N-1}|, 1 + |a|\}$. Dann ist $|a_n| \leq M$ für alle $n \in \mathbb{N}$ erfüllt. □

Wichtig ist auch der Umkehrschluss dieses Satzes: Jede unbeschränkte Folge ist divergent.

(2.4) Proposition Sei $b \in \mathbb{K}$ und $(a_n)_{n\in\mathbb{N}}$ gegeben durch $a_n = b^n$ für alle $n \in \mathbb{N}$.

(i) Ist $|b| < 1$, dann gilt $\lim_n a_n = 0$.
(ii) Für $b = 1$ gilt $\lim_n a_n = 1$.
(iii) In den Fällen $b = -1$ und $|b| > 1$ ist die Folge divergent.

Beweis: zu (i) Durch vollständige Induktion zeigt man zunächst, dass $|b^n| = |b|^n$ und $|b|^n < 1$ für alle $n \in \mathbb{N}$ gilt. Sei nun $\varepsilon > 0$ vorgegeben. Nach Satz (1.10) gibt es ein $N \in \mathbb{N}$ mit $|b|^N < \varepsilon$. Für alle $n \geq N$ folgt dann

$$|a_n - 0| = |b^n| = |b|^n = |b|^{n-N}|b|^N \leq |b|^N < \varepsilon.$$

zu (ii) Die Folge $(1^n)_{n\in\mathbb{N}} = (1)_{n\in\mathbb{N}}$ ist konstant und konvergiert daher, wie oben gesehen, gegen ihren konstanten Wert.

zu (iii) Im Fall $b = -1$ stimmt $(a_n)_{n\in\mathbb{N}}$ mit der Folge $((-1)^n)_{n\in\mathbb{N}}$ überein und ist somit, wie bereits gesehen, divergent. Für alle $b \in \mathbb{C}$ und $|b| > 1$ zeigen wir, dass die Folge unbeschränkt ist. Angenommen, es gibt ein $c \in \mathbb{R}^+$, $c > 0$ mit $|a_n| < c$ für alle $n \in \mathbb{N}$. Nach Satz (1.10) gibt es andererseits ein $N \in \mathbb{N}$ mit $|b^N| = |b|^N > c$. Wir erhalten somit einen Widerspruch zu unserer Annahme. □

(2.5) Proposition Ändert man bei einer Folge nur endlich viele Glieder ab, so ändert dies nichts am Konvergenzverhalten. Genauer: Seien $(a_n)_{n\in\mathbb{N}}$ und $(b_n)_{n\in\mathbb{N}}$ Folgen in $\mathbb{K}$ und $M \in \mathbb{N}$, so dass $a_n = b_n$ für alle $n \geq M$ erfüllt ist. Konvergiert nun $(a_n)_{n\in\mathbb{N}}$ gegen ein $a \in \mathbb{K}$, dann konvergiert $(b_n)_{n\in\mathbb{N}}$ gegen denselben Wert.

Beweis: Sei $\varepsilon > 0$ vorgegeben. Auf Grund der Konvergenz von $(a_n)_{n\in\mathbb{N}}$ finden wir ein $N_0 \in \mathbb{N}$ mit der Eigenschaft, dass $|a_n - a| < \varepsilon$ für alle $n \geq N_0$ gilt. Sei nun $N = \max\{M, N_0\}$. Für alle $n \geq N$ gilt dann $|b_n - a| = |a_n - a| < \varepsilon$. Also konvergiert auch $(b_n)_{n\in\mathbb{N}}$ gegen a. □

Ebenso einfach beweist man, dass sich das Konvergenzverhalten einer Folge nicht ändert, wenn man endlich viele Glieder am Anfang der Folge hinzunimmt oder entfernt.

(2.6) Satz Der Grenzwert einer Folge $(a_n)_{n\in\mathbb{N}}$ in $\mathbb{K}$, sofern er existiert, ist eindeutig bestimmt.

Beweis: Angenommen, a und a' sind zwei verschiedene Grenzwerte der Folge. Sei $\varepsilon = |a' - a|$. Wegen $\lim_n a_n = a$ gibt es ein $N_1 \in \mathbb{N}$, so dass $|a_n - a| < \frac{1}{2}\varepsilon$ für alle $n \geq N_1$ gilt. Andererseits gibt es wegen $\lim_{n\to\infty} a_n = a'$ auch ein $N_2 \in \mathbb{N}$ mit der Eigenschaft, dass $|a_n - a'| < \frac{1}{2}\varepsilon$ für alle $n \geq N_2$ erfüllt ist. Für alle $n \geq \max\{N_1, N_2\}$ gilt nun

$$\varepsilon \quad = \quad |a' - a| \quad = \quad |(a' - a_n) + (a_n - a)| \quad \leq \quad |a_n - a'| + |a_n - a| \quad < \quad \tfrac{1}{2}\varepsilon + \tfrac{1}{2}\varepsilon \quad = \quad \varepsilon.$$

Dieser Widerspruch zeigt, dass die Annahme falsch war und keine zwei verschiedenen Grenzwerte existieren. □

Wir beweisen nun eine Reihe von Aussagen, die es uns erlauben, die Berechnung der Grenzwerte von kompliziert aufgebauten Folgen auf bereits bekannte Grenzwerte von Folgen zurückzuführen. Man bezeichnet diese Aussagen zusammenfassend als ***Grenzwertsätze***.

(2.7) Satz Seien $(a_n)_{n\in\mathbb{N}}$, $(b_n)_{n\in\mathbb{N}}$ und $(c_n)_{n\in\mathbb{N}}$ Folgen in $\mathbb{K}$, wobei $c_n = a_n + b_n$ für alle $n \in \mathbb{N}$ ist. Gilt nun

$$\lim_{n\to\infty} a_n = a \qquad \text{und} \qquad \lim_{n\to\infty} b_n = b \qquad \text{mit} \quad a, b \in \mathbb{K} \quad ,$$

dann konvergiert auch $(c_n)_{n\in\mathbb{N}}$, und zwar gegen den Wert $a + b$.

Beweis: Sei $\varepsilon > 0$ vorgegeben. Auf Grund der Konvergenz von $(a_n)_{n\in\mathbb{N}}$ gibt es ein $N_1 \in \mathbb{N}$ mit $|a_n - a| < \frac{1}{2}\varepsilon$ für alle $n \geq N_1$. Die Konvergenz von $(b_n)_{n\in\mathbb{N}}$ liefert uns ein $N_2 \in \mathbb{N}$ mit $|b_n - b| < \frac{1}{2}\varepsilon$ für alle $n \geq N_2$. Sei $N = \max\{N_1, N_2\}$. Für $n \in \mathbb{N}$ mit $n \geq N$ gilt dann

$$|c_n - (a+b)| \quad = \quad |(a_n + b_n) - (a+b)| \quad = \quad |(a_n - a) + (b_n - b)| \quad \leq |a_n - a| + |b_n - b| \quad < \quad \tfrac{1}{2}\varepsilon + \tfrac{1}{2}\varepsilon \quad = \quad \varepsilon.$$

Also konvergiert $(c_n)_{n\in\mathbb{N}}$ tatsächlich gegen den Wert $a + b$. □

(2.8) Satz Seien $(a_n)_{n\in\mathbb{N}}$, $(b_n)_{n\in\mathbb{N}}$ und $(c_n)_{n\in\mathbb{N}}$ Folgen in $\mathbb{K}$ mit $c_n = a_n b_n$ für alle $n \in \mathbb{N}$. Konvergieren die Folgen $(a_n)_{n\in\mathbb{N}}$ und $(b_n)_{n\in\mathbb{N}}$ gegen a bzw. $b \in \mathbb{K}$, dann konvergiert auch die Folge $(c_n)_{n\in\mathbb{N}}$, und zwar gegen den Wert ab.

Beweis: Weil $(a_n)_{n\in\mathbb{N}}$ konvergent ist, ist sie nach Folgerung (2.3) auch beschränkt, es gibt also ein $K_1 \in \mathbb{R}$, $K_1 > 0$ mit $|a_n| < K_1$ für alle $n \in \mathbb{N}$. Sei $K = \max\{K_1, |b|\}$ und $\varepsilon > 0$ eine vorgegebene reelle Zahl. Auf Grund der Konvergenz von $(a_n)_{n\in\mathbb{N}}$ und $(b_n)_{n\in\mathbb{N}}$ finden wir natürliche Zahlen $N_1, N_2 \in \mathbb{N}$, so dass

$$|a_n - a| < \frac{\varepsilon}{2K} \quad \text{für alle} \quad n \geq N_1 \qquad \text{und} \qquad |b_n - b| < \frac{\varepsilon}{2K} \quad \text{für alle} \quad n \geq N_2$$

erfüllt ist. Sei nun $N = \max\{N_1, N_2\}$. Für alle $n \geq N$ gilt dann

$$\begin{aligned} |c_n - ab| \quad = \quad |a_n b_n - ab| \quad = \quad |a_n b_n - a_n b + a_n b - ab| \quad = \quad |a_n(b_n - b) + (a_n - a)b| \quad \leq \\ |a_n||b_n - b| + |a_n - a||b| \quad < \quad K\frac{\varepsilon}{2K} + \frac{\varepsilon}{2K}K \quad = \quad \tfrac{1}{2}\varepsilon + \tfrac{1}{2}\varepsilon \quad = \quad \varepsilon. \end{aligned}$$

Also konvergiert $(c_n)_{n\in\mathbb{N}}$ gegen den Wert ab. □

(2.9) Folgerung Seien $(a_n)_{n\in\mathbb{N}}$ und $(b_n)_{n\in\mathbb{N}}$ konvergente Folgen mit den Grenzwerten a bzw. $b \in \mathbb{K}$ und $\lambda \in \mathbb{K}$.

(i) Die Folge $(c_n)_{n\in\mathbb{N}}$ gegeben durch $c_n = \lambda a_n$ konvergiert gegen den Wert λa.
(ii) Die Folge $(d_n)_{n\in\mathbb{N}}$ mit $d_n = a_n - b_n$ konvergiert gegen den Wert $a - b$.

Beweis: Aussage (i) folgt aus Satz (2.8), wenn wir für $(b_n)_{n\in\mathbb{N}}$ die konstante Folge $(\lambda)_{n\in\mathbb{N}}$ einsetzen. Zum Beweis von (ii) wenden wir (i) zunächst auf $\lambda = -1$ an. Dies liefert die Gleichung $\lim_n -b_n = -b$. Nun erhält man die gewünschte Aussage durch Anwendung von Satz (2.7) auf die Folgen $(a_n)_{n\in\mathbb{N}}$ und $(-b_n)_{n\in\mathbb{N}}$. □

(2.10) Satz Seien $(a_n)_{n\in\mathbb{N}}$ und $(b_n)_{n\in\mathbb{N}}$ konvergente Folgen mit Grenzwerten $a, b \in \mathbb{K}$, wobei $b \neq 0$ ist. Dann gibt es ein $N_0 \in \mathbb{N}$ mit $b_n \neq 0$ für alle $n \geq N_0$, und die Folge $(c_n)_{n\geq N_0}$ mit $c_n = \frac{a_n}{b_n}$ konvergiert gegen den Wert $\frac{a}{b}$.

Beweis: Wir zeigen zunächst, dass ein $N_0 \in \mathbb{N}$ mit der angegebenen Eigenschaft existiert. Sei $\varepsilon = \frac{1}{2}|b|$. Auf Grund der Konvergenz von $(b_n)_{n\in\mathbb{N}}$ existiert ein $N_0 \in \mathbb{N}$ mit $|b_n - b| < \varepsilon$ für alle $n \in \mathbb{N}$. Für alle $n \geq N_0$ gilt dann $|b| = |b - b_n + b_n| \leq |b - b_n| + |b_n| < \frac{1}{2}|b| + |b_n|$ und somit $|b_n| > \frac{1}{2}|b|$, insbesondere ist $b_n \neq 0_{\mathbb{K}}$. Durch Übergang zum Kehrwert erhalten wir die Abschätzung $|b_n|^{-1} < 2|b|^{-1}$.

Wir beweisen nun die Konvergenz von $(c_n)_{n\geq N_0}$ zunächst für den Spezialfall, dass es sich bei $(a_n)_{n\in\mathbb{N}}$ um die konstante Folge mit Wert 1 handelt. Für vorgegebenes $\varepsilon > 0$ wählen wir uns ein $N_1 \in \mathbb{N}$ mit der Eigenschaft, dass $|b_n - b| < \frac{1}{2}\varepsilon|b|^2$ für alle $n \geq N_1$ gilt. Für alle $n \geq N$ mit $N = \max\{N_0, N_1\}$ erhalten wir dann

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

Die Aussage für eine beliebige konvergente Folge $(a_n)_{n\in\mathbb{N}}$ erhält man nun wegen $\frac{a_n}{b_n} = a_n \cdot \frac{1}{b_n}$ aus Satz (2.8). □

Als Anwendungsbeispiel für die bisherigen Sätze bestimmen wir den Grenzwert der Folge

$$a_n \quad = \quad \frac{n^2 + 7n + 3}{3n^2 - 5} \qquad \text{für alle} \quad n \in \mathbb{N}$$

Es gilt

$$\lim_{n\to\infty} a_n \quad = \quad \lim_{n\to\infty} \frac{n^2 + 7n + 3}{3n^2 - 5} \quad = \quad \lim_{n\to\infty} \frac{1 + \frac{7}{n} + \frac{3}{n^2}}{3 - \frac{5}{n^2}} \quad = \quad \frac{\lim_n \left(1 + \frac{7}{n} + \frac{3}{n^2}\right)}{\lim_n \left(3 - \frac{5}{n^2}\right)}$$

$$= \quad \frac{\lim_n 1 + \lim_n \left(\frac{7}{n}\right) + \lim_n \left(\frac{3}{n^2}\right)}{\lim_n 3 - \lim \left(\frac{5}{n^2}\right)}$$

sofern alle in der Rechnung aufgeführten Grenzwerte existieren. (Wenn auch nur einer der Grenzwerte nicht existieren würde, dann wäre die gesamte Rechnung wertlos, weil die Voraussetzungen der Grenzwertsätze nicht erfüllt und ihre Anwendung daher unzulässig wäre!) Nun gilt $\lim_n \frac{1}{n} = 0$, und dies liefert uns $\lim_n \frac{1}{n^2} = (\lim_n \frac{1}{n})(\lim_n \frac{1}{n}) = 0 \cdot 0 = 0$. Es folgt $\lim_n \frac{7}{n} = 0$, $\lim_n \frac{3}{n^2} = 0$ und $\lim_n \frac{5}{n^2} = 0$. Insgesamt erhalten wir

$$\lim_{n\to\infty} a_n \quad = \quad \frac{1 + 0 + 0}{3 - 0} \quad = \quad \frac{1}{3}.$$

(2.11) Satz Seien $(a_n)_{n\in\mathbb{N}}$, $(b_n)_{n\in\mathbb{N}}$ reelle konvergente Folgen mit Grenzwerten $a, b \in \mathbb{R}$. Gilt $a_n \leq b_n$ für alle $n \in \mathbb{N}$, dann folgt $a \leq b$.

Beweis: Nehmen wir an, dass $a > b$ ist, und sei $\varepsilon = \frac{1}{2}(a-b) > 0$. Wegen der Konvergenz von $(a_n)_{n\in\mathbb{N}}$ gibt es ein $N_1 \in \mathbb{N}$ mit $|a_n - a| < \varepsilon$, also $a_n > a - \varepsilon = \frac{1}{2}(a+b)$. Ebenso finden wir ein $N_2 \in \mathbb{N}$ mit $|b_n - b| < \varepsilon$, also $b_n < b + \varepsilon = \frac{1}{2}(a+b)$ für alle $n \geq N_2$. Sei nun $N = \max\{N_1, N_2\}$. Für alle $n \geq N$ erhalten wir dann $a_n > \frac{1}{2}(a+b) > b_n$, im Widerspruch zur Voraussetzung, dass $a_n \leq b_n$ für alle $n \in \mathbb{N}$ gilt. □

(2.12) Satz *(„Sandwich-Lemma")*
Seien $(a_n)_{n\in\mathbb{N}}$, $(b_n)_{n\in\mathbb{N}}$ und $(c_n)_{n\in\mathbb{N}}$ drei Folgen reeller Zahlen, und es gelte

$$a_n \quad \leq \quad b_n \quad \leq \quad c_n \qquad \text{für alle} \quad n \in \mathbb{N}.$$

Konvergieren die Folgen $(a_n)_{n\in\mathbb{N}}$ und $(c_n)_{n\in\mathbb{N}}$ gegen denselben Grenzwert $a \in \mathbb{R}$, dann konvergiert die Folge $(b_n)_{n\in\mathbb{N}}$, und zwar gegen den Wert a.

Beweis: Sei $\varepsilon \in \mathbb{R}^+$ vorgegeben. Wegen $\lim_n a_n = \lim_n c_n = a$ gibt es $N_1, N_2 \in \mathbb{N}$, so dass $|a_n - a| < \varepsilon$ für alle $n \geq N_1$ und $|c_n - a| < \varepsilon$ für alle $n \geq N_2$ erfüllt ist. Es folgt $a_n > a - \varepsilon$ für alle $n \geq N_1$ und $c_n < a + \varepsilon$ für alle $n \geq N_2$. Sei $N = \max\{N_1, N_2\}$ und $n \in \mathbb{N}$ mit $n \geq N$. Dann gilt einerseits

$$b_n - a \quad \geq \quad a_n - a \quad > \quad a - \varepsilon - a \quad = \quad -\varepsilon$$

und andererseits

$$b_n - a \quad \leq \quad c_n - a \quad < \quad a + \varepsilon - a \quad = \quad \varepsilon.$$

Aus $-\varepsilon < b_n - a < \varepsilon$ folgt $|b_n - a| < \varepsilon$. Damit ist $\lim_n b_n = a$ bewiesen. □

(2.13) Definition Man sagt, eine reelle Folge $(a_n)_{n\in\mathbb{N}}$ konvergiert ***uneigentlich*** gegen $+\infty$ und verwendet die Notation

$$\lim_{n\to\infty} a_n = +\infty \quad ,$$

wenn für jedes $\kappa \in \mathbb{R}^+$ ein $N \in \mathbb{N}$ existiert, so dass $a_n > \kappa$ für alle $n \geq N$ erfüllt ist. Ebenso schreiben wir $\lim_n a_n = -\infty$, wenn es für jedes $\kappa \in \mathbb{R}^+$ ein $N \in \mathbb{N}$ mit $a_n < -\kappa$ für alle $n \geq N$ gibt.

Wir betrachten einige Beispiele zur uneigentlichen Konvergenz.

(i) Die Folge $(a_n)_{n\in\mathbb{N}}$ mit $a_n = 2^n$ konvergiert uneigentlich gegen $+\infty$. Ist nämlich $\kappa > 0$ vorgegeben, dann finden wir nach Satz (1.10) ein $N \in \mathbb{N}$ mit $2^N > \kappa$. Für alle $n \geq N$ gilt dann $2^n = 2^{n-N} \cdot 2^N > \kappa$.

(ii) Die Folge $(a_n)_{n\in\mathbb{N}}$ mit $a_n = -n$ konvergiert uneigentlich gegen $-\infty$. Ist nämlich $\kappa \in \mathbb{R}^+$ vorgegeben, dann gibt es nach dem archimedischen Axiom ein $N \in \mathbb{N}$ mit $N > \kappa$. Für alle $n \geq N$ gilt dann $a_n = -n \leq -N < -\kappa$.

(iii) Die Folge $(a_n)_{n\in\mathbb{N}}$ mit $a_n = (-1)^n$ ist weder eigentlich noch uneigentlich konvergent. Dass es keinen reellen Grenzwert gibt, haben wir bereits gezeigt. Ebensowenig gilt $\lim_n a_n = +\infty$, denn es es gibt zum Beispiel für $\kappa = 1$ kein $N \in \mathbb{N}$ mit $a_n > \kappa$ für alle $n \geq N$. (Tatsächlich ist kein einziges Folgenglied größer als 1.) Genauso sieht man, dass auch $\lim_n a_n = -\infty$ nicht erfüllt ist.

Auch für uneigentliche Grenzwerte lassen sich in begrenztem Umfang „Rechenregeln“ aufstellen.

(2.14) Proposition Seien $(a_n)_{n\in\mathbb{N}}$ und $(b_n)_{n\in\mathbb{N}}$ Folgen reeller Zahlen mit

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

Dann gilt auch $\lim_n (a_n + b_n) = +\infty$.

Beweis: Sei $\kappa \in \mathbb{R}^+$ vorgegeben. Wegen $\lim_n a_n = +\infty$ gibt es ein $N_1 \in \mathbb{N}$ mit $a_n > \frac{1}{2}\kappa$ für alle $n \geq N_1$. Ebenso gibt es wegen $\lim_n b_n = +\infty$ ein $N_2 \in \mathbb{N}$ mit $b_n > \frac{1}{2}\kappa$. Sei nun $N = \max\{N_1, N_2\}$. Für alle $n \in \mathbb{N}$ mit $n \geq N$ gilt dann

$$a_n + b_n \quad > \quad \tfrac{1}{2}\kappa + \tfrac{1}{2}\kappa \quad = \quad \kappa.$$

Damit ist $\lim_n (a_n + b_n) = +\infty$ bewiesen. □

Sind andererseits $(a_n)_{n\in\mathbb{N}}$ und $(b_n)_{n\in\mathbb{N}}$ Folgen mit $\lim_n a_n = +\infty$ und $\lim_n b_n = -\infty$, dann ist über die Konvergenz der Folge $(a_n+b_n)_{n\in\mathbb{N}}$ keine allgemeine Aussage möglich. Ist beispielsweise $a_n = 2n$ und $b_n = -n$, dann gilt $a_n+b_n = n$ und somit $\lim_n (a_n + b_n) = +\infty$. Im Fall $a_n = n$ und $b_n = -2n$ konvergiert $a_n + b_n = -n$ uneigentlich gegen $-\infty$. Es kann aber auch vorkommen, dass $(a_n + b_n)_{n\in\mathbb{N}}$ gegen eine reelle Zahl konvergiert (zum Beispiel gegen 42 im Fall $a_n = n + 42$, $b_n = -n$) oder einfach divergiert (zum Beispiel bei $a_n = n + (-1)^n$ und $b_n = -n$). Folgende Tabellen geben zusammenfassend an, wie sich (uneigentliche) Grenzwerte bezüglich der Rechenoperationen $+$, $-$, $\cdot$ und bei Kehrwertbildung verhalten.

$+$	$-\infty$	$b \in \mathbb{R}$	$+\infty$
$-\infty$	$-\infty$	$-\infty$	?
$a \in \mathbb{R}$	$-\infty$	$a+b$	$+\infty$
$+\infty$	?	$+\infty$	$+\infty$

$-$	$-\infty$	$b \in \mathbb{R}$	$+\infty$
$-\infty$	?	$-\infty$	$-\infty$
$a \in \mathbb{R}$	$+\infty$	$a-b$	$-\infty$
$+\infty$	$+\infty$	$+\infty$	?

$\cdot$	$-\infty$	$b<0$	$b=0$	$b>0$	$+\infty$
$-\infty$	$+\infty$	$+\infty$	?	$-\infty$	$-\infty$
$a<0$	$+\infty$	ab	0	ab	$-\infty$
$a=0$	?	0	0	0	?
$a>0$	$-\infty$	ab	0	ab	$+\infty$
$+\infty$	$-\infty$	$-\infty$	?	$+\infty$	$+\infty$

$\lim a_n$	$-\infty$	$a<0$	$a=0$	$a>0$	$+\infty$
$\lim a_n^{-1}$	0	a^{-1}	?	a^{-1}	0

Beispielsweise bedeutet der Eintrag links oben in der „+“-Tabelle: Sind $(a_n)_{n\in\mathbb{N}_0}$ und $(b_n)_{n\in\mathbb{N}_0}$ Folgen mit $\lim_n a_n = \lim_n b_n = -\infty$, dann folgt $\lim_n (a_n+b_n) = -\infty$. Die Zeilen und Spalten mit den Einträgen $b \in \mathbb{R}$, $b < 0$ usw. beziehen

sich immer auf endliche (reelle) Werte. Der Eintrag „?“ bedeutet jeweils, dass bei der entsprechenden Kombination keine allgemeine Aussage möglich ist, so wie wir es oben bereits in einem konkreten Fall gesehen haben.

Im Hinblick auf spätere Anwendungen wird es sich als günstig herausstellen, dass wir für die Konvergenz einer reellen Folge eine einheitliche Definition zur Verfügung haben. Dazu erinnern wir an die in § 8 eingeführte Bezeichnung $\bar{\mathbb{R}}$ für die Menge $\mathbb{R} \cup \{\pm\infty\}$.

(2.15) Definition Seien $a \in \bar{\mathbb{R}}$ und $U \subseteq \bar{\mathbb{R}}$ vorgegeben. Wir bezeichnen die Menge U als ***Umgebung*** des Punktes a

(i) im Fall $a \in \mathbb{R}$, wenn ein $\varepsilon \in \mathbb{R}^+$ existiert, so dass $]a-\varepsilon, a+\varepsilon[\subseteq U$ erfüllt ist
(ii) im Fall $a = +\infty$, wenn ein $\kappa \in \mathbb{R}^+$ existiert, so dass $]\kappa, +\infty] \subseteq U$ gilt,
(iii) im Fall $a = -\infty$, wenn ein $\kappa \in \mathbb{R}^+$ existiert, so dass $[-\infty, -\kappa[\subseteq U$ gilt.

Im Fall (i) bezeichnen wir die Menge $]a-\varepsilon, a+\varepsilon[$ selbst als ε***-Umgebung*** von a.

Auf der Grundlage dieser Definition gilt nun: Eine reelle Folge $(a_n)_{n\in\mathbb{N}}$ konvergiert genau dann (eigentlich oder uneigentlich) gegen einen Punkt $a \in \bar{\mathbb{R}}$, wenn für jede Umgebung $U \subseteq \bar{\mathbb{R}}$ von a ein $N \in \mathbb{N}$ existiert, so dass $a_n \in U$ für alle $n \geq N$ erfüllt ist. Dies ergibt sich unmittelbar durch Vergleich mit den Definitionen (2.1) und (2.13).

§ 3. *Häufungspunkte und Cauchyfolgen*

Inhaltsübersicht

Häufig kommt es vor, dass eine Folge zwar im Ganzen nicht konvergiert, wohl aber bestimmte Teilfolgen. Dies ist zum Beispiel immer dann der Fall, wenn die Folge divergent und beschränkt ist, und führt auf den Begriff des ***Häufungspunkts***. Insbesondere gibt es immer einen kleinsten und einen größten Häufungspunkt, den man als ***Limes inferior*** bzw. ***Limes superior*** der Folge bezeichnet.

Desweiteren ist es in einigen Situationen notwendig, die Konvergenz einer Folge zu nachzuweisen, ohne das man einen konkreten Grenzwert angeben kann. Dies wird durch den Begriff der ***Cauchyfolge*** ermöglicht. Besonders häufig verwendet man die Cauchyfolgen im Zusammenhang mit Reihen, die das Thema des nächsten Abschnitts sein werden. Für viele Reihen lässt sich die Konvergenz nachweisen. Dass man aber auch den Grenzwert der Reihe angeben kann, ist eher die Ausnahme.

Wichtige Begriffe und Sätze

- (streng) monoton wachsende / fallende Folgen
- Häufungspunkte und Teilfolgen
- Eine Folge $(a_n)_{n\in\mathbb{N}}$ hat genau dann $a \in \mathbb{K}$ als Häufungspunkt, wenn a Grenzwert einer Teilfolge ist.
- Limes superior und Limes inferior einer Folge
- Für beschränkte Folgen ist $\liminf_n a_n$ der kleinste und $\limsup_n a_n$ der größte Häufungspunkt.
- Cauchyfolgen
- Eine Folge konvergiert genau dann in $\mathbb{R}$, wenn sie eine Cauchyfolge ist.

Wieder bezeichnet $\mathbb{K}$ im gesamten Abschnitt einen der Körper $\mathbb{R}$ oder $\mathbb{C}$.

(3.1) Definition Sei $(a_n)_{n\in\mathbb{N}}$ eine Folge in $\mathbb{K}$. Ein Punkt $a \in \mathbb{K}$ wird ***Häufungspunkt*** von $(a_n)_{n\in\mathbb{N}}$ genannt, wenn es für jedes $\varepsilon \in \mathbb{R}^+$ jeweils unendlich viele $n \in \mathbb{N}$ mit $|a_n - a| < \varepsilon$ gibt.

Für den weiteren Verlauf benötigen wir die folgenden Grundbegriffe.

(3.2) Definition Eine Folge $(a_n)_{n\in\mathbb{N}}$ reeller Zahlen ist

(i) ***monoton wachsend***, wenn $a_n \leq a_{n+1}$
(ii) ***streng monoton wachsend***, wenn $a_n < a_{n+1}$
(iii) ***monoton fallend***, wenn $a_n \geq a_{n+1}$
(iv) ***streng monoton fallend***, wenn $a_n > a_{n+1}$

für alle $n \in \mathbb{N}$ erfüllt ist.

Für den praktischen Nachweis von Häufungspunkten eignet sich die Charakterisierung durch sog. Teilfolgen.

(3.3) Definition Sei $(a_n)_{n\in\mathbb{N}}$ eine Folge in $\mathbb{K}$. Eine Folge $(b_n)_{n\in\mathbb{N}}$ wird ***Teilfolge*** von $(a_n)_{n\in\mathbb{N}}$ genannt, wenn eine streng monoton wachsende Folge $(n_k)_{n\in\mathbb{N}}$ in $\mathbb{N}$ existiert, so dass $b_k = a_{n_k}$ für alle $k \in \mathbb{N}$ erfüllt ist.

Ist zum Beispiel $a_n = \frac{1}{n}$ für alle $n \in \mathbb{N}$, dann ist die Folge $(b_n)_{n\in\mathbb{N}}$ gegeben durch $b_n = \frac{1}{2n}$ eine Teilfolge von $(a_n)_{n\in\mathbb{N}}$. Setzen wir nämlich $n_k = 2k$ für alle $k \in \mathbb{N}$, dann gilt $b_k = \frac{1}{2k} = \frac{1}{n_k} = a_{n_k}$ für alle $k \in \mathbb{N}$. Ebenso leicht überprüft man, dass die Folge $(c_n)_{n\in\mathbb{N}}$ gegeben durch $c_n = (-1)^n$ die beiden konstanten Folgen mit den Werten -1 und 1 als Teilfolgen enthält.

(3.4) Proposition Sei $(a_n)_{n\in\mathbb{N}}$ eine Folge in $\mathbb{K}$ und $a \in \mathbb{K}$. Genau dann ist a ein Häufungspunkt von $(a_n)_{n\in\mathbb{N}}$, wenn eine Teilfolge $(b_n)_{n\in\mathbb{N}}$ von $(a_n)_{n\in\mathbb{N}}$ mit $\lim_n b_n = a$ existiert.

Beweis: „⇒“ Wir konstruieren eine Teilfolge von $(a_n)_{n\in\mathbb{N}}$, indem wir eine Folge $(n_k)_{k\in\mathbb{N}}$ in $\mathbb{N}$ nach folgender Vorschrift definieren: Zunächst setzen wir $n_1 = 1$. Sei nun $k \in \mathbb{N}$ und n_k bereits definiert. Weil a ein Häufungspunkt von $(a_n)_{n\in\mathbb{N}}$ ist, existieren unendlich viele $n \in \mathbb{N}$ mit $|a_n - a| < \frac{1}{k+1}$. Insbesondere gibt es ein $n > n_k$ mit dieser Eigenschaft. Wir definieren dann $n_{k+1} = n$. Mit der Folge $(n_k)_{k\in\mathbb{N}}$ definieren wir nun eine Teilfolge $(b_k)_{k\in\mathbb{N}}$, indem wir $b_k = a_{n_k}$ für jedes $k \in \mathbb{N}$ setzen. Nach Konstruktion ist $(n_k)_{n\in\mathbb{N}}$ dann streng monoton wachsend, und es gilt $|b_k - a| = |a_{n_k} - a| < \frac{1}{k}$ für alle $k \in \mathbb{N}$ mit $k > 1$.

Nun überprüfen wir, dass $(b_k)_{k\in\mathbb{N}}$ gegen a konvergiert. Sei dazu $\varepsilon \in \mathbb{R}^+$ vorgegeben. Dann existiert ein $k_0 \in \mathbb{N}$ mit $k_0 > 1$ und $\frac{1}{k_0} < \varepsilon$. Für alle $k \geq k_0$ gilt nun $|b_k - a| < \frac{1}{k} \leq \frac{1}{k_0} < \varepsilon$. Damit ist $\lim_k b_k = a$ nachgewiesen.

„⇐“ Sei $(b_k)_{k\in\mathbb{N}}$ eine Teilfolge von A mit $\lim_k b_k = a$. Nach Definition des Teilfolgenbegriffs existiert eine streng monoton wachsende Folge $(n_k)_{k\in\mathbb{N}}$ in $\mathbb{N}$ mit $b_k = a_{n_k}$ für alle $k \in \mathbb{N}$. Zum Nachweis der Häufungspunkt-Eigenschaft sei $\varepsilon \in \mathbb{R}^+$ vorgegeben. Dann gibt es ein $K_0 \in \mathbb{N}$ mit $|a_{n_k} - a| = |b_k - a| < \varepsilon$ für alle $k \geq K_0$. Also sind durch n_k mit $k \geq K_0$ unendlich viele natürliche Zahlen n_k mit $|a_{n_k} - a| < \varepsilon$ gegeben. □

Aus der Bemerkung unmittelbar vor Prop. (3.4) folgt also, dass die Folge $(c_n)_{n\in\mathbb{N}}$ gegeben durch $c_n = (-1)^n$ die beiden Werte -1 und 1 als Häufungspunkte besitzt, denn wie wir in § 8 gesehen haben, konvergiert eine konstante Folge stets gegen ihren konstanten Wert.

(3.5) Proposition Konvergiert die Folge $(a_n)_{n\in\mathbb{N}}$ gegen ein $a \in \mathbb{K}$, dann ist a der einzige Häufungspunkt von $(a_n)_{n\in\mathbb{N}}$.

Beweis: Zunächst zeigen wir, dass a ein Häufungspunkt der Folge ist. Sei $\varepsilon \in \mathbb{R}^+$ vorgegeben. Dann gibt es ein $N \in \mathbb{N}$ mit $|a_n - a| < \varepsilon$ für alle $n \geq N$. Also ist $|a_n - a| < \varepsilon$ für unendlich viele $n \in \mathbb{N}$ erfüllt. Nehmen wir nun an, es gibt einen weiteren Häufungspunkt b. Auf Grund der Konvergenz gibt es für $\varepsilon_1 = |b - a|$ ein $N_1 \in \mathbb{N}$ mit $|a_n - a| < \frac{1}{2}\varepsilon_1$ für alle $n \geq N_1$. Sei nun $n \in \mathbb{N}$ mit $n \geq N_1$. Wäre $|a_n - b| < \frac{1}{2}\varepsilon_1$, dann würde daraus

$$\varepsilon_1 \quad = \quad |b-a| \quad \leq \quad |(b-a_n)+(a_n-a)| \quad \leq \quad |a_n-b|+|a_n-a| \quad < \quad \tfrac{1}{2}\varepsilon_1+\tfrac{1}{2}\varepsilon_1 \quad = \quad \varepsilon_1$$

folgen. Der Widerspruch $\varepsilon_1 < \varepsilon_1$ zeigt, dass $|a_n - b| < \frac{1}{2}\varepsilon_1$ nur für $n < N_1$ möglich ist. Damit kann b kein Häufungspunkt von $(a_n)_{n\in\mathbb{N}}$ sein. □

Wir werden nun die Häufungspunkte *reeller* Folgen genauer untersuchen.

(3.6) Satz Jede beschränkte, monoton wachsende oder monoton fallende Folge reeller Zahlen konvergiert.

Beweis: Sei $(a_n)_{n\in\mathbb{N}}$ monoton wachsend und beschränkt. Auf Grund der Vollständigkeit von $\mathbb{R}$ existiert $s = \sup(A)$ für die Menge $A = \{a_n \mid n \in \mathbb{N}\}$. Wir beweisen nun, dass $s = \lim_n a_n$ gilt und geben dafür ein $\varepsilon \in \mathbb{R}^+$ vor. Da $s - \varepsilon$ keine obere Schranke für die Menge A ist, gibt es ein $N \in \mathbb{N}$ mit $s - \varepsilon < a_N \leq s$. Für beliebiges $n \geq N$ gilt nun $s - \varepsilon < a_N \leq a_n \leq s$, also $a_n > s - \varepsilon$ und somit $|s - a_n| = s - a_n < \varepsilon$.

Die Argumentation für monoton fallende Folgen verläuft weitgehend analog. Anstelle des Supremums betrachtet man hier den Wert $s = \inf(A)$. Weil für beliebig vorgegebenes $\varepsilon \in \mathbb{R}^+$ die Zahl $s + \varepsilon$ keine untere Schranke von A ist, gibt es ein $N \in \mathbb{N}$ mit $s \leq a_N < s + \varepsilon$. Für alle $n \geq N$ folgt $s \leq a_n \leq a_N < s + \varepsilon$, also insbesondere $a_n < s + \varepsilon$ und somit $|s - a_n| = a_n - s < \varepsilon$. □

(3.7) Definition Sei $(a_n)_{n\in\mathbb{N}}$ eine Folge reeller Zahlen. Man nennt

$$\limsup_n a_n = \lim_{n\to\infty} \sup\{a_m \mid m \in \mathbb{N}, m \geq n\} \quad \text{bzw.}$$
$$\liminf_n a_n = \lim_{n\to\infty} \inf\{a_m \mid m \in \mathbb{N}, m \geq n\}$$

den ***Limes superior*** bzw. ***Limes inferior*** von $(a_n)_{n\in\mathbb{N}}$, sofern die Grenzwerte in $\mathbb{R}$ existieren.

(3.8) Proposition Ist $(a_n)_{n\in\mathbb{N}}$ eine beschränkte Folge in $\mathbb{R}$, dann existieren $\limsup_n a_n$ und $\liminf_n a_n$ in $\mathbb{R}$.

Beweis: Wir beweisen nur die Existenz von $\limsup_n a_n$. Dazu definieren wir für jedes $n \in \mathbb{N}$ die Menge $A_n = \{a_m \mid m \in \mathbb{N}, m \geq n\}$, außerdem sei a eine untere und b eine obere Schranke von $A_1 = \{a_m \mid m \in \mathbb{N}\}$. Für jedes $n \in \mathbb{N}$ gilt $A_n \subseteq A_1$. Also ist a auch eine untere und b eine obere Schranke von A_n. Weil A_n nichtleer und nach oben beschränkt ist, existiert $\sup A_n$, und es gilt $a \leq \sup A_n \leq b$.

Wir müssen nun zeigen, dass $\lim_n \sup A_n$ existiert. Wegen $A_n \supseteq A_{n+1}$ ist jede obere Schranke von A_n auch eine obere Schranke von A_{n+1}. Es gilt also $\mathscr{S}^+(A_n) \subseteq \mathscr{S}^+(A_{n+1})$, und daraus folgt

$$\sup A_n = \min \mathscr{S}^+(A_n) \geq \min \mathscr{S}^+(A_{n+1}) = \sup A_{n+1}.$$

Insgesamt haben wir damit nachgewiesen, dass die Folge gegeben durch die Elemente $\sup A_n$, $n \in \mathbb{N}$ monoton fallend und nach unten beschränkt ist. Aus Satz (3.6) folgt die Konvergenz. □

(3.9) Satz Sei $(a_n)_{n\in\mathbb{N}}$ eine beschränkte Folge reeller Zahlen. Dann ist $\limsup_n a_n$ der größte und $\liminf_n a_n$ der kleinste Häufungspunkt der Folge.

Beweis: Wieder beschränken wir uns beim Beweis auf den Limes superior. Zunächst zeigen wir, dass $a = \limsup_n a_n$ *ein* Häufungspunkt der Folge ist. Sei $\varepsilon \in \mathbb{R}^+$ vorgegeben. Es genügt zu zeigen, dass für jedes $N \in \mathbb{N}$ jeweils ein $n \in \mathbb{N}$ mit $n \geq N$ und $|a_n - a| < \varepsilon$ existiert, denn dann gibt es unendlich viele n mit $|a_n - a| < \varepsilon$.

Sei also $N \in \mathbb{N}$ vorgegeben. Wie im letzten Beweis definieren wir $A_n = \{a_m \mid m \geq n\}$ für alle $n \in \mathbb{N}$. Weil die Folge $(\sup A_n)_{n\in\mathbb{N}}$ gegen a konvergiert, gibt es ein $m \in \mathbb{N}$ mit $m \geq N$ und $|\sup A_m - a| < \frac{1}{2}\varepsilon$. Weil außerdem $\sup A_m$ eine obere Schranke, $\sup A_m - \frac{1}{2}\varepsilon$ aber keine obere Schranke von A_m ist, gibt es ein $n \in \mathbb{N}$ mit $n \geq m \geq N$ und

$$\sup A_m - \tfrac{1}{2}\varepsilon \quad < \quad a_n \quad \leq \quad \sup A_m.$$

Insbesondere gilt $|a_n - \sup A_m| < \frac{1}{2}\varepsilon$. Insgesamt erhalten wir $|a_n - a| = |a_n - \sup A_m + \sup A_m - a| \leq |a_n - \sup A_m| + |\sup A_m - a| < \frac{1}{2}\varepsilon + \frac{1}{2}\varepsilon = \varepsilon$.

Nun zeigen wir, dass a der *größte* Häufungspunkt der Folge ist. Nehmen wir an, dass $b \in \mathbb{R}$ mit $b > a$ ein weiterer Häufungspunkt ist. Sei $\varepsilon = b - a$. Dann gibt es für jedes $m \in \mathbb{N}$ ein $n \geq m$ mit $|a_n - b| < \frac{1}{2}\varepsilon$. Insbesondere gilt $a_n > b - \frac{1}{2}\varepsilon = a + \frac{1}{2}\varepsilon$. Weil $\sup A_m$ eine obere Schranke von $\{a_n \mid n \geq m\}$ ist, folgt daraus

$$\sup A_m \quad \geq \quad a + \tfrac{1}{2}\varepsilon.$$

Weil diese Abschätzung für jedes $m \in \mathbb{N}$ gültig ist, folgt

$$a \quad = \quad \lim_{m\to\infty} \sup A_m \quad \geq \quad a + \tfrac{1}{2}\varepsilon \quad ,$$

im Widerspruch zu $a < a + \frac{1}{2}\varepsilon$. Also muss a der größte Häufungspunkt von $(a_n)_{n\in\mathbb{N}}$ sein. □

Häufig werden in der Literatur der Limes superior und Limes inferior auch für unbeschränkte Folgen definiert. Wir verzichten auf die Angabe der allgemeinen Definition, weil wir dafür zunächst auch für Folgen, die $\pm\infty$ als Folgenglieder enthalten können, (uneigentliche) Grenzwerte definieren müssten. Zu dieser verallgemeinerten Definition erwähnen wir lediglich, dass für eine nach oben unbeschränkte Folge $(a_n)_{n\in\mathbb{N}}$ jeweils $\limsup_n a_n = +\infty$ gilt, und für eine nach unten unbeschränkte Folge entsprechend $\liminf_n a_n = -\infty$.

(3.10) Folgerung *(Satz von Bolzano-Weierstrass)*
Jede beschränkte Folge reeller Zahlen besitzt einen Häufungspunkt.

Beweis: Dies ist eine unmittelbare Konsequenz aus Satz (3.9). □

(3.11) Folgerung Eine Folge konvergiert genau dann, wenn sie beschränkt ist und genau einen Häufungspunkt besitzt.

Beweis: Die Richtung „⇒“ ergibt sich aus Proposition (3.5) und Satz (2.3). „⇐“ Sei $(a_n)_{n\in\mathbb{N}}$ eine beschränkte Folge mit genau einem Häufungspunkt. Für jedes $m \in \mathbb{N}$ sei $A_m = \{a_n \mid n \geq m\}$. Nach Satz (3.9) ist $\lim_n \sup A_n = \limsup_n a_n$ der größte und $\lim_n \inf A_n = \liminf_n a_n$ der kleinste Häufungspunkt von $(a_n)_{n\in\mathbb{N}}$. Auf Grund der Voraussetzung, dass nur ein Häufungspunkt existiert, fallen beide Werte zusammen. Wegen $a_m \in A_m$ gilt $\inf A_m \leq a_m \leq \sup A_m$ für alle $m \in \mathbb{N}$. Aus dem Sandwich-Lemmas (Satz (2.12)) folgt damit die Konvergenz der Folge $(a_n)_{n\in\mathbb{N}}$. □

Nach den Häufungspunkten wenden wir uns nun dem zweiten zentralen Begriff dieses Abschnitts zu.

(3.12) Definition Eine Folge $(a_n)_{n\in\mathbb{N}}$ in $\mathbb{K}$ heißt ***Cauchyfolge***, wenn für jedes $\varepsilon \in \mathbb{R}^+$ ein $N \in \mathbb{N}$ existiert, so dass

$$|a_n - a_m| < \varepsilon \qquad \text{für alle} \quad n, m \geq N \quad \text{gilt.}$$

Namensgeber der Cauchyfolgen ist der französische Mathematiker *Augustin-Louis Cauchy (1798-1857)*.

Wir geben ein anschauliches Kriterium für Cauchyfolgen an, dass auf der Schachtelung von Intervallen basiert. Ist $I \subseteq \mathbb{R}$ ein endliches Intervall (offen, halboffen oder abgeschlossen) mit a als links- und b als rechtsseitiger Grenze, dann nennen wir $\ell(I) = b - a$ die ***Länge*** des Intervalls. Wir bemerken außerdem: Sind $I = [a, b]$ und $J = [c, d]$ endliche abgeschlossene Intervalle und ist $I \cap J$ nichtleer, dann gilt $c \leq b$ oder $a \leq d$, und außerdem $I \cap J = [\max\{a, c\}, \min\{b, d\}]$.

(3.13) Proposition Eine Folge $(a_n)_{n\in\mathbb{N}}$ in $\mathbb{R}$ ist genau dann eine Cauchyfolge, wenn eine Folge $(J_k)_{k\in\mathbb{N}}$ endlicher abgeschlossener Intervalle $J_k \subseteq \mathbb{R}$ existiert mit $\ell(J_k) \leq 2^{-k}$ und $J_{k+1} \subseteq J_k$ für alle $k \in \mathbb{N}$, wobei jedes J_k alle bis auf endlich viele Punkte der Folge enthält.

Beweis: „⇒“ Wir konstruieren die Intervalle J_k nach folgendem Verfahren. Auf Grund der Cauchyfolgen-Eigenschaft gibt es für jedes $k \in \mathbb{N}$ ein $N_k \in \mathbb{N}$ mit $|a_m - a_n| < 2^{-(k+1)}$ für alle $m, n \geq N_k$. Dabei können die N_k so gewählt werden, dass die Folge $(N_k)_{k\in\mathbb{N}}$ streng monoton wachsend ist. Insbesondere gilt dann $|a_n - a_{N_k}| < 2^{-(k+1)}$ für alle $n \geq N_k$. Definieren wir für jedes $k \in \mathbb{N}$ jeweils

$$I_k \quad = \quad \left[a_{N_k} - 2^{-(k+1)}, a_{N_k} + 2^{-(k+1)}\right]$$

dann gilt also $a_n \in I_k$ für alle $n \geq N_k$. Sei nun $J_1 = I_1$ und $J_{k+1} = J_k \cap I_{k+1}$ für alle $k \in \mathbb{N}$. Dann gilt offenbar $J_{k+1} \subseteq J_k$ und $\ell(J_k) \leq 2^{-k}$ für alle $k \in \mathbb{N}$. Wir beweisen nun durch vollständige Induktion über $k \in \mathbb{N}$, dass J_k jeweils alle Folgenglieder $n \geq N_k$ enthält. Für $k = 1$ ist dies wegen $J_1 = I_1$ jedenfalls erfüllt. Sei nun $k \in \mathbb{N}$, und setzen wir die Aussage für k voraus. Laut Induktionsvoraussetzung gilt $a_n \in J_k$ für alle $n \geq N_k$, also insbesondere für alle $n \geq N_{k+1}$. Ebenso ist a_n für alle $n \geq N_{k+1}$ in I_{k+1} enthalten. Also gilt tatsächlich $a_n \in J_{k+1}$ für alle $n \geq N_{k+1}$.

„⇐“ Sei $(J_k)_{k\in\mathbb{N}}$ eine Intervallfolge mit den angegebenen Eigenschaften. Zum Nachweis der Cauchyfolgen-Eigenschaft sei $\varepsilon \in \mathbb{R}^+$ vorgegeben. Dann gibt es nach Satz (1.10) ein $k \in \mathbb{N}$ mit $2^{-k} < \varepsilon$. Seien r_k, s_k die Grenzen des Intervalls J_k. Weil alle bis auf endlich viele Folgenglieder in J_k liegen, gibt es ein $N \in \mathbb{N}$ mit $a_n \in J_k$ für alle $n \geq N$. Sind nun

$m, n \geq N$ vorgegeben, dann gilt $a_m, a_n \in J_k$, somit $r_k < a_m, a_n < s_k$ und somit $|a_n - a_m| < s_k - r_k = \ell(J_k) < 2^{-k} < \varepsilon$. Also ist $(a_n)_{n\in\mathbb{N}}$ tatsächlich eine Cauchyfolge. □

(3.14) Satz Die Cauchyfolgen in $\mathbb{R}$ sind genau die konvergenten Folgen.

Beweis: „⇐“ Sei $(a_n)_{n\in\mathbb{N}}$ eine konvergente Folge und $a \in \mathbb{R}$ ihr Grenzwert. Zum Nachweis der Cauchyfolgen-Eigenschaft sei $\varepsilon \in \mathbb{R}^+$ vorgegeben. Dann gibt es $N \in \mathbb{N}$ mit $|a_n - a| < \frac{1}{2}\varepsilon$ für alle $n \geq N$. Sind $m, n \geq N$ vorgegeben, so folgt

$$|a_n - a_m| \quad = \quad |(a_n - a) + (a - a_m)| \quad \leq \quad |a_n - a| + |a_m - a| \quad < \quad \tfrac{1}{2}\varepsilon + \tfrac{1}{2}\varepsilon \quad = \quad \varepsilon.$$

„⇒“ Sei $(a_n)_{n\in\mathbb{N}}$ eine reelle Cauchyfolge und $(J_k)_{k\in\mathbb{N}}$ eine Folge von Intervallen mit den in Proposition (3.13) beschriebenen Eigenschaften. Auf Grund der Vollständigkeit von $\mathbb{R}$ und der Intervallschachtelungs-Eigenschaft, die nach Satz (1.17)) vorliegt, gibt es ein $a \in \mathbb{R}$, so dass $a \in J_k$ für alle $k \in \mathbb{N}$ erfüllt ist. Wir zeigen, dass $\lim_n a_n = a$ gilt. Sei $\varepsilon \in \mathbb{R}^+$ vorgegeben und $k \in \mathbb{N}$ so groß gewählt, dass $2^{-k} < \varepsilon$ gilt. Weil das Intervall J_k alle bis auf endlich viele Folgenglieder enthält, gibt es ein $N \in \mathbb{N}$, so dass $a_n \in J_k$ für alle $n \geq N$ erfüllt ist. Aus $a, a_n \in J_k$ und $\ell(J_k) \leq 2^{-k} < \varepsilon$ folgt $|a_n - a| < \varepsilon$ für alle $n \geq N$. □

Die Richtung „⇒“ in Satz (3.14) kann auch ohne die Verwendung der Intervallschachtelungs-Eigenschaft bewiesen werden, in dem man statt dessen auf Folgerung (3.11) zurückgreift. Sei $(a_n)_{n\in\mathbb{N}}$ eine Cauchyfolge. Zunächst stellen wir fest, dass $(a_n)_{n\in\mathbb{N}}$ beschränkt ist. Denn auf Grund der Cauchyfolgen-Eigenschaft gibt es ein $N \in \mathbb{N}$, so dass $|a_n - a_m| < 1$ für alle $m, n \geq N$ erfüllt ist. Insbesondere gilt dann $|a_n - a_N| < 1$ und

$$|a_n| \quad = \quad |a_n - a_N + a_N| \quad \leq \quad |a_n - a_N| + |a_N| \quad < \quad |a_N| + 1$$

für alle $n \geq N$, so dass insgesamt $|a_n| \leq \kappa$ für alle $n \in \mathbb{N}$ mit der Konstanten $\kappa = \max\{|a_1|, ..., |a_{N-1}|, |a_N| + 1\}$ gilt. Als beschränkte Folge besitzt $(a_n)_{n\in\mathbb{N}}$ nach dem Satz von Bolzano-Weierstraß mindestens einen Häufungspunkt. Andererseits kann es bei Cauchyfolgen nicht mehr als einen Häufungspunkt geben: Sind nämlich a, b zwei verschiedene Häufungspunkte, dann sei $\varepsilon = |a - b|$ und $N_1 \in \mathbb{N}$ so groß gewählt, dass $|a_n - a_m| < \frac{1}{3}\varepsilon$ für alle $n \geq N_1$ gilt. Weil a und b beides Häufungspunkte sind, gibt es $n_1, n_2 \in \mathbb{N}$ mit $n_1, n_2 \geq N_1$ und $|a_{n_1} - a| < \frac{1}{3}\varepsilon$ und $|a_{n_2} - b| < \frac{1}{3}\varepsilon$. Es ergibt sich dann der Widerspruch

$$\varepsilon \;=\; |a - b| \;=\; |a - a_{n_1} + a_{n_1} - a_{n_2} + a_{n_2} - b| \;\leq\; |a - a_{n_1}| + |a_{n_1} - a_{n_2}| + |a_{n_2} - b| \;<\; \tfrac{1}{3}\varepsilon + \tfrac{1}{3}\varepsilon + \tfrac{1}{3}\varepsilon \;=\; \varepsilon.$$

Also hat $(a_n)_{n\in\mathbb{N}}$ genau einen Häufungspunkt. Ingesamt folgt die Konvergenz damit aus Satz (3.11).

Cauchyfolgen ermöglichen es also, die Konvergenz einer Folge auch dann nachzuweisen, wenn der Grenzwert unbekannt ist. Dies wird sich im nächsten Abschnitt bei den unendlichen Summen (den Reihen) als sehr hilfreich erweisen. Es ist nicht schwer, aus Satz (3.14) die entsprechende Aussage auch für den Körper $\mathbb{C}$ der komplexen Zahlen nachzuweisen. Aus Zeitgründen verzichten wir aber auf die Ausführung.

Wir beenden den Abschnitt mit einem Anwendungsbeispiel, indem wir die Konvergenz der Folge

$$a_n \quad = \quad \left(1+\frac{1}{n}\right)^n$$

untersuchen. Von praktischer Bedeutung ist diese Formel für die Zinsrechnung. Stellen wir uns vor, ein gewisser Geldbetrag k (in Euro) wird mit einem jährlichen Zinssatz von 100 Prozent verzinst. (Diesen unrealistisch hohen Zinssatz wählen wir nur, um das Beispiel so einfach wie möglich zu halten.) Dann hängt der Kontostand am Ende des Jahres davon ab, wie häufig der Betrag verzinst wurde. Fallen die Zinsen nur einmal im Jahr an, hätte man am Jahresende den Betrag $2a$. Bei monatlicher Verzinsung werden natürlich nicht jedesmal 100 Prozent aufgeschlagen, sondern der Zinssatz wird durch 12 geteilt. Am Jahresende hätte sich dann der Betrag $k(1+\frac{1}{12})^{12}$ angesammelt. Bei n gleichmäßig über das Jahr verteilten Verzinsungen hat man am Jahresende den Betrag

$$ka_n \quad = \quad k\left(1+\frac{1}{n}\right)^n$$

angespart. Lassen wir n gegen unendlich laufen, dann erhält man den Betrag, der bei einer ***kontinuierlichen*** Verzinsung zu Stande kommt. Die entscheidende Frage ist nun, ob dies noch ein endlicher Wert ist, und wenn ja, in welcher Größenordnung sich dieser bewegt.

(3.15) Proposition Die Folge $(a_n)_{n\in\mathbb{N}}$ gegeben durch $a_n = (1+\frac{1}{n})^n$ für alle $n \in \mathbb{N}$ ist konvergent.
Wir bezeichnen den Grenzwert als die ***Eulersche Zahl*** e.

Beweis: Neben $(a_n)_{n\in\mathbb{N}}$ betrachten wir auch die Folge $(b_n)_{n\in\mathbb{N}}$ gegeben durch $b_n = (1+\frac{1}{n})^{n+1} = a_n(1+\frac{1}{n})$. Wir zeigen, dass $(a_n)_{n\in\mathbb{N}}$ monoton wachsend und $(b_n)_{n\in\mathbb{N}}$ monoton fallend ist. Für jedes $n \in \mathbb{N}$ gilt die Äquivalenz

$$a_n \le a_{n+1} \quad \Leftrightarrow \quad (1+\tfrac{1}{n})^n \le (1+\tfrac{1}{n+1})^{n+1} \quad \Leftrightarrow \quad (1+\tfrac{1}{n})^{-1} \le \left(\frac{1+\frac{1}{n+1}}{1+\frac{1}{n}}\right)^{n+1}$$

Der Term in der Klammer auf der rechten Seite kann umgeformt werden zu

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

$$= \quad (1-\tfrac{1}{n+1})(1+\tfrac{1}{n+1}) \quad = \quad 1-\tfrac{1}{(n+1)^2}.$$

Mit der Bernoullischen Ungleichung in Satz (1.10) erhalten wir

$$\left(\frac{1+\frac{1}{n+1}}{1+\frac{1}{n}}\right)^{n+1} \quad = \quad \left(1-\frac{1}{(n+1)^2}\right)^{n+1} \quad \ge \quad 1-(n+1)\cdot\tfrac{1}{(n+1)^2} \quad = \quad 1-\tfrac{1}{n+1} \quad =$$

$$\left(\frac{1}{1-\frac{1}{n+1}}\right)^{-1} \quad = \quad \left(\frac{n+1}{n+1-1}\right)^{-1} \quad = \quad \left(\frac{n+1}{n}\right)^{-1} \quad = \quad (1+\tfrac{1}{n})^{-1} \quad ,$$

woraus auf Grund unserer Vorüberlegung $a_n \le a_{n+1}$ folgt. Also ist $(a_n)_{n\in\mathbb{N}}$ tatsächlich monoton wachsend. Nach demselben Schema kann gezeigt werden, dass $(b_n)_{n\in\mathbb{N}}$ monoton fallend ist: Es gilt die Äquivalenz

$$b_n \ge b_{n+1} \quad \Leftrightarrow \quad (1+\tfrac{1}{n})^{n+1} \ge (1+\tfrac{1}{n+1})^{n+2} \quad \Leftrightarrow \quad \left(\frac{1+\frac{1}{n}}{1+\frac{1}{n+1}}\right)^{n+1} \ge 1+\tfrac{1}{n+1}.$$

Der Ausdruck in der Klammer auf der linken Seite kann umgeformt werden zu

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

$$= \quad (1+\tfrac{1}{n})(1-\tfrac{1}{n+2}) \quad = \quad 1+\tfrac{1}{n}-\tfrac{1}{n+2}-\tfrac{1}{n(n+2)} \quad = \quad 1+\tfrac{(n+2)-n-1}{n(n+2)} \quad = \quad 1+\tfrac{1}{n(n+2)}.$$

Durch erneute Anwendung der Bernoullischen Ungleichung erhalten wir

$$\left(\frac{1+\frac{1}{n}}{1+\frac{1}{n+1}}\right)^{n+1} \quad = \quad \left(1+\tfrac{1}{n(n+2)}\right)^{n+1} \quad \geq \quad 1+\tfrac{n+1}{n(n+2)} \quad \geq \quad 1+\tfrac{1}{n+1}$$

wobei im letzten Schritt $\frac{n+1}{n(n+2)} \geq \frac{1}{n+1} \Leftrightarrow (n+1)^2 \geq n(n+2) \Leftrightarrow n^2+2n+1 \geq n^2+2n$ verwendet wurde. Also ist $(b_n)_{n\in\mathbb{N}}$ tatsächlich monoton fallend. Nun gilt für alle $n \in \mathbb{N}$ außerdem die Abschätzung

$$1 \quad < \quad 1+\tfrac{1}{n} \quad \leq \quad \left(1+\tfrac{1}{n}\right)^n \quad = \quad a_n \quad \leq \quad b_n \quad = \quad \left(1+\tfrac{1}{n}\right)^{n+1} \quad \leq \quad (1+1)^2 \quad = \quad 4.$$

Dies bedeutet, dass $(a_n)_{n\in\mathbb{N}}$ nach oben und $(b_n)_{n\in\mathbb{N}}$ nach unten beschränkt ist. Nach Satz (3.6) sind damit beide Folgen konvergent. □

Aus dem Beweis von Satz (3.6) geht hervor, dass der Grenzwert der Folge $(a_n)_{n\in\mathbb{N}}$ das Supremum der Teilmenge $\{a_n \mid n \in \mathbb{N}\}$ von $\mathbb{R}$ ist. Dies bedeutet, dass die Folge $(a_n)_{n\in\mathbb{N}}$ von unten gegen den Grenzwert e konvergiert. Ebenso konvergiert $(b_n)_{n\in\mathbb{N}}$ von oben gegen ihren Grenzwert. Es gilt also $a_n \leq e \leq b_n$ für alle $n \in \mathbb{N}$. Man erhält auf diese Weise für $n = 2$ die Abschätzung $2,25 = \frac{9}{4} < e < \frac{27}{8} = 3,375$. Für $n = 10^5$ erhält man die genauere Abschätzung $2,71827 < e < 2,71830$. Gerundet auf die ersten 60 Dezimalstellen ist die Eulersche Zahl e gegeben durch

$$e \quad \approx \quad 2,718281828459045235360287471352662497757247093699959574966 97.$$

§ 4. Konvergenzkriterien für Reihen

Inhaltsübersicht

Jeder Folge $(a_n)_{n\in\mathbb{N}}$ kann eine neue Folge $(a_n)_{n\in\mathbb{N}}$ zugeordnet werden, deren n-tes Glied durch $s_n = a_1 + a_2 + \ldots + a_n$ gegeben ist. Man nennt eine solche Folge eine ***Reihe*** und bezeichnet sie mit dem Ausdruck $\sum_{n=1}^{\infty} a_n$. Beispielsweise wird die Folge mit den Gliedern $1, 1+\frac{1}{2}, 1+\frac{1}{2}+\frac{1}{3}, \ldots$ die ***harmonische Reihe*** genannt. Wir werden in diesem Kapitel eine Reihe von Kriterien zusammenstellen, mit denen sich die Konvergenz von Reihen untersuchen lässt. Mit Hilfe der Reihen können wir dann unter anderem die aus der Schulmathematik bekannte ***Dezimalbruchentwicklung*** reeller Zahlen präzise definieren.

Wichtige Begriffe und Sätze

- Summen- und Produktzeichen
- Reihen und ihre Partialsummen
- Konvergenz, Divergenz und absolute Konvergenz von Reihen
- Konvergenzkriterien: Cauchy-, Monotonie-, Leibniz-, Majoranten-, Minoranten-, Quotienten-, Wurzel- und Verdichtungskriterium
- Umordnungssatz für absolut konvergente Reihen
- Multiplikation von Reihen, Cauchyprodukte
- Entwicklung reeller Zahlen als b-adische Brüche

Zunächst soll in diesem Kapitel das ***Summenzeichen***, das wir bereits an der einen oder anderen Stelle verwendet haben, in der notwendigen Allgemeinheit definiert werden. Zur Vorbereitung führen wir die folgende Notation ein: Ist I eine Menge, dann bezeichnen wir mit $\mathscr{P}_{\text{fin}}(I)$ die Menge der ***endlichen Teilmengen*** von I. Es handelt sich um die Vereinigung sämtlicher Mengen $\mathscr{P}_k(I)$ mit $k \in \mathbb{N}_0$.

(4.1) Satz Sei I eine Menge und $(a_i)_{i\in I}$ eine Familie natürlicher Zahlen mit Indexmenge I. Dann gibt es eine eindeutig bestimmte Zuordnung

$$\mathscr{P}_{\text{fin}}(I) \to \mathbb{N}_0 \quad , \quad J \mapsto \sum_{i\in J} a_i$$

mit den Eigenschaften $\sum_{i\in\emptyset} a_i = 0$, $\sum_{i\in\{k\}} a_i = a_k$ für alle $k \in I$ und

$$\sum_{i\in J\cup K} a_i = \sum_{i\in J} a_j + \sum_{i\in K} a_k \qquad \text{für } J, K \in \mathscr{P}_{\text{fin}}(I) \text{ mit } J \cap K = \emptyset.$$

Die Existenzaussage im diesem Satz beweist man durch eine rekursive Definition des Summenzeichens. Zunächst definiert man das Summenzeichen für die (einzige) nullelementige Teilmenge von I durch $\sum_{i\in\varnothing} a_i = 0$. Sei nun $n \in \mathbb{N}_0$, und nehmen wir an, dass das Summenzeichen für n-elementige Teilmengen von I bereits definiert wurde. Ist nun $|J| = n+1$ und $j \in J$, dann ist $J' = J \setminus \{j\}$ eine n-elementige Menge von I, und wir können $\sum_{i\in J} a_i = \sum_{i\in J'} a_i + a_j$ definieren.

Es muss allerdings noch überprüft werden, dass diese Definition unabhängig von der Auswahl des Elements j aus J ist; dieselbe Vorgehensweise mit einem anderen $j' \in J$ muss für $\sum_{i\in J} a_i$ denselben Wert liefern. Aus Zeitgründen verzichten wir auf die detaillierte Ausführung, zumal wir in erster Linie an der praktischen Verwendung des Summenzeichens interessiert sind.

Ist $(a_i)_{i\in I}$ besipielsweise eine Familie natürlicher Zahlen über einer zweielementigen Menge $I = \{k, \ell\}$, dann gilt $\sum_{i\in I} a_i = a_k + a_\ell$. Für eine dreielementige Menge $I = \{k, \ell, m\}$ erhält man entsprechend $\sum_{i\in I} a_i = a_k + a_\ell + a_m$, und so weiter. Besonders häufig verwenden als Indexmengen Teilmengen von $\mathbb{Z}$ der Form $M_n = \{1, 2, ..., n\}$ mit $n \in \mathbb{N}$ oder $M_{r,s} = \{r, r+1, ..., s-1, s\}$ mit $r, s \in \mathbb{Z}$ und $r \le s$. Außerdem bietet es sich an, $M_0 = \varnothing$ und $M_{r,s} = \varnothing$ für $r > s$ zu setzen. Ist nun $(a_m)_{m\in\mathbb{Z}}$ eine Familie natürlicher Zahlen mit $\mathbb{Z}$ als Indexmenge, dann setzt man

$$\sum_{i=1}^{n} a_i \quad = \quad \sum_{i\in M_n} a_i \qquad \text{und} \qquad \sum_{i=r}^{s} a_i \quad = \quad \sum_{i\in M_{r,s}} a_i.$$

Beispielsweise ist $\sum_{i=3}^{7} a_i = a_3 + a_4 + a_5 + a_6 + a_7$ und $\sum_{i=6}^{4} a_i = 0$. Die Formel für die Summe der ersten n natürlichen Zahlen aus Kapitel 1 kann mit dem Summenzeichen in der Form $\sum_{i=1}^{n} i = \frac{1}{2}n(n+1)$ geschrieben werden.

(4.2) Proposition (i) Seien $(a_i)_{i\in I}$, $(b_i)_{i\in I}$ zwei Familien natürlicher Zahlen mit endlicher Indexmenge I, und sei $c \in \mathbb{N}_0$. Dann sind $(a_i + b_i)_{i\in I}$ und $(ca_i)_{i\in I}$ ebenfalls Familien natürlicher Zahlen über I, und es gilt

$$\sum_{i\in I}(a_i + b_i) \quad = \quad \sum_{i\in I} a_i + \sum_{i\in I} b_i \qquad \text{und} \qquad \sum_{i\in I} ca_i \quad = \quad c\sum_{i\in I} a_i.$$

(ii) Sei J eine weitere endliche Menge und $\varphi : J \to I$ eine Bijektion. Dann ist $(c_j)_{j\in J}$ mit $c_j = a_{\varphi(j)}$ eine Familie natürlicher Zahlen mit Indexmenge J, und es gilt

$$\sum_{j\in J} c_j \quad = \quad \sum_{j\in J} a_{\varphi(j)} \quad = \quad \sum_{i\in I} a_i.$$

Auch hier verzichten wir auf die Angabe eines Beweises. Die letzte Gleichung bezeichnet man als ***Umparametrisierung*** einer Summe. Für Anwendungen ist vor allem der folgende Spezialfall interessant: Für $m, n, r \in \mathbb{Z}$ mit $m \le n$ ist durch $\varphi : M_{m+r,n+r} \to M_{m,n}$, $j \mapsto j - r$ eine Bijektion definiert, deren Umkehrabbildung offenbar durch $M_{m,n} \to M_{m+r,n+r}$, $i \mapsto i + r$ gegeben ist. Wir erhalten dann für eine beliebige Familie $(a_i)_{i\in\mathbb{Z}}$ die Gleichung

$$\sum_{i=m}^{n} a_i \quad = \quad \sum_{i\in M_{m,n}} a_i \quad = \quad \sum_{j\in M_{m+r,n+r}} a_{\varphi(j)} \quad = \quad \sum_{j=m+r}^{n+r} a_{j-r}.$$

Ist beispielsweise $(a_m)_{m\in\mathbb{Z}}$ gegeben durch $a_m = m^2$ für alle $m \in \mathbb{Z}$, dann gilt

$$\sum_{i=3}^{5} a_i \quad = \quad 3^2+4^2+5^2 \quad = \quad (6-3)^2+(7-3)^2+(8-3)^2 \quad = \quad \sum_{j=6}^{8} a_{j-3}.$$

Neben dem Summen- wird häufig auch das ***Produktzeichen*** verwendet. Hier ist das Ziel die Bildung von Produkten über beliebige endliche Mengen natürlicher Zahlen. Wie in Satz (4.1) kann gezeigt werden, dass für jede Familie $(a_i)_{i\in I}$ natürlicher Zahlen über einer Indexmenge I eine eindeutig bestimmte Zuordnung

$$\mathscr{P}_{\text{fin}}(I) \longrightarrow \mathbb{N}_0 \quad , \quad J \mapsto \prod_{i\in J} a_i$$

mit $\prod_{i\in\varnothing} a_i = 1$, $\prod_{i\in\{k\}} a_i = a_k$ und $\prod_{i\in J\cup K} a_i = \left(\prod_{i\in J} a_j\right)\left(\prod_{i\in K} a_k\right)$ existiert, falls $J, K \subseteq I$ endlich und disjunkte sind. Es gelten zu Satz (4.2) analoge Rechenregeln $\prod_{i\in I}(a_i b_i) = \left(\prod_{i\in I} a_i\right)\left(\prod_{i\in I} b_i\right)$ und $\prod_{j\in J} a_{\varphi(j)} = \prod_{i\in I} a_i$, falls $\varphi : J \to I$ eine Bijektion bezeichnet. Lediglich die Gleichung $\sum_{i\in I} ca_i = c\sum_{i\in I} a_i$ muss modifiziert werden: Beim Produktzeichen gilt $\prod_{i\in I}(ca_i) = c^n \prod_{i\in I} a_i$, falls $n = |I|$ ist.

Sowohl die Fakultätsfunktion als auch die Binomialkoeffizienten können mit Hilfe des Produktzeichens ausgedrückt werden: Es gilt

$$n! \quad = \quad \prod_{k=1}^{n} k \qquad \text{und} \qquad \binom{n}{k} \quad = \quad \frac{1}{k!}\frac{n!}{(n-k)!} \quad = \quad \frac{1}{k!}\prod_{\ell=n-k+1}^{n} \ell$$

für alle $k, n \in \mathbb{N}_0$ mit $k \leq n$.

Im weiteren Verlauf beschäftigen wir nun mit *unendlichen* Summen.

(4.3) Definition Sei $(a_n)_{n\in\mathbb{N}}$ eine Folge in $\mathbb{K}$. Dann bezeichnet man die Folge $(s_n)_{n\in\mathbb{N}}$ gegeben durch

$$s_n = \sum_{k=1}^{n} a_k \qquad \text{für} \quad n \in \mathbb{N}$$

als ***Reihe*** über $(a_n)_{n\in\mathbb{N}}$. Das Folgenglied s_n nennt man die ***n-te Partialsumme*** der Reihe.

Beginnt die Folge nicht bei 1, sondern an einem beliebigen Startwert $n_0 \in \mathbb{Z}$, dann beginnt auch die Folge der Partialsummen bei n_0. Die Reihe über $(a_n)_{n\geq n_0}$ ist dann gegeben durch

$$s_n \quad = \quad \sum_{k=n_0}^{n} a_k \qquad \text{für alle} \quad n \in \mathbb{Z} \text{ mit } n \geq n_0.$$

Die Reihe über $(a_n)_{n\in\mathbb{N}}$ bzw. $(a_n)_{n\geq n_0}$ wird mit $\sum_{n=1}^{\infty} a_n$ bzw. $\sum_{n=n_0}^{\infty} a_n$ bezeichnet. Wir betrachten einige konkrete Beispiele für Reihen.

(i) Die Reihe $\sum_{n=1}^{\infty} \frac{1}{n}$ besteht aus den Folgengliedern

$$s_1 = 1, \quad s_2 = 1+\tfrac{1}{2} = \tfrac{3}{2}, \quad s_3 = 1+\tfrac{1}{2}+\tfrac{1}{3} = \tfrac{11}{6}, \quad s_4 = 1+\tfrac{1}{2}+\tfrac{1}{3}+\tfrac{1}{4} = \tfrac{25}{12}, \quad \ldots$$

(ii) Die Reihe $\sum_{n=1}^{\infty}(-1)^n$ besteht aus den Folgengliedern

$$s_1 = -1\ ,\ s_2 = (-1)+1 = 0\ ,\ s_3 = (-1)+1+(-1) = -1\ ,\ s_4 = (-1)+1+(-1)+1 = 0\ ,\ \ldots$$

(4.4) Definition Sei $\sum_{n=1}^{\infty} a_n$ eine Reihe über $\mathbb{K}$ und $(s_n)_{n\in\mathbb{N}}$ die Folge ihrer Partialsummen. Man sagt, die Reihe ***konvergiert*** gegen einen Grenzwert $a \in \mathbb{R}$ (oder $\mathbb{C}$), wenn die Folge $(s_n)_{n\in\mathbb{N}}$ gegen a konvergiert. Existiert kein solcher Grenzwert, dann spricht man von einer ***divergenten*** Reihe.

Ebenso ordnet man der Reihe den Grenzwert $+\infty$ oder $-\infty$ zu, wenn die Folge $(s_n)_{n\in\mathbb{N}}$ uneigentlich gegen $+\infty$ oder $-\infty$ konvergiert. Konvergiert die Reihe $\sum_{n=1}^{\infty} a_n$, dann verwendet man die Notation $\sum_{n=1}^{\infty} a_n$ auch für den Grenzwert, man setzt also

$$\sum_{n=1}^{\infty} a_n \quad = \quad \lim_{n\to\infty} \sum_{k=1}^{n} a_k.$$

Die Bezeichnung $\sum_{n=1}^{\infty} a_n$ wird also für zwei verschiedene Objekte verwendet: einerseits für die Folge $(s_n)_{n\in\mathbb{N}}$ der Partialsummen, andererseits auch für den (evtl. uneigentlichen) Grenzwert dieser Folge, sofern er existiert. Welche Bedeutung gemeint ist, muss dem jeweiligen Kontext entnommen werden.

In einigen Fällen lässt sich der Grenzwert einer Reihe direkt angeben.

(4.5) Proposition Es gilt $\sum_{n=1}^{\infty} \frac{1}{n(n+1)} = 1.$

Beweis: Wir beweisen zunächst durch vollständige Induktion über $n \in \mathbb{N}$ die Summenformel $\sum_{k=1}^{n} \frac{1}{k(k+1)} = \frac{n}{n+1}$. Die Gleichung ist für $n = 1$ richtig, denn es gilt $\sum_{k=1}^{1} \frac{1}{k+1} = \frac{1}{2} = \frac{1}{1+1}$. Den Induktionsschluss von n auf $n+1$ erhält man durch

$$\sum_{k=1}^{n+1} \frac{1}{k(k+1)} \quad = \quad \sum_{k=1}^{n} \frac{1}{k(k+1)} + \frac{1}{(n+1)(n+2)} \quad \overset{\text{IV}}{=} \quad \frac{n}{n+1} + \frac{1}{(n+1)(n+2)} \quad =$$

$$\frac{n(n+2)+1}{(n+1)(n+2)} \quad = \quad \frac{n^2+2n+1}{(n+1)(n+2)} \quad = \quad \frac{(n+1)^2}{(n+1)(n+2)} \quad = \quad \frac{n+1}{n+2} \quad ,$$

wobei wir beim zweiten „=“ die Induktionsvoraussetzung verwendet haben. Die Folge $(s_n)_{n\in\mathbb{N}}$ ist also gegeben durch $s_n = \frac{n}{n+1}$, und es folgt $\sum_{n=1}^{\infty} \frac{1}{n(n+1)} = \lim_n \frac{n}{n+1} = 1$. □

(4.6) Proposition *(geometrische Reihe)*

Für alle $x \in \mathbb{C}$ mit $|x| < 1$ gilt $\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}$.

Beweis: Wir beweisen durch vollständige Induktion über $n \in \mathbb{N}_0$ die Summenformel

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

Für $n = 0$ gilt $\sum_{k=0}^{0} x^0 = 1 = \frac{1-x}{1-x}$. Der Schluss von n auf $n+1$ folgt aus der Rechnung

$$\sum_{k=0}^{n+1} x^k \quad = \quad \sum_{k=0}^{n} x^k + x^{n+1} \overset{\text{IV}}{=} \frac{1-x^{n+1}}{1-x} + x^{n+1} \quad = \quad \frac{1-x^{n+1}}{1-x} + \frac{x^{n+1}-x^{n+2}}{1-x} \quad = \quad \frac{1-x^{n+2}}{1-x}.$$

Die Folge $(s_n)_{n\in\mathbb{N}}$ der Partialsummen ist also gegeben durch $s_n = \frac{1-x^{n+1}}{1-x}$. Wegen $|x| < 1$ gilt $\lim\limits_{n\to\infty} x^n = 0$. Unter Verwendung der Grenzwertsätze erhalten wir also

$$\sum_{n=0}^{\infty} x^n \quad = \quad \lim_{n\to\infty} \frac{1-x^{n+1}}{1-x} \quad = \quad \frac{1}{1-x}. \qquad \square$$

(4.7) Satz Seien $\sum_{n=1}^{\infty} a_n$ und $\sum_{n=1}^{\infty} b_n$ konvergente Reihen über $\mathbb{K}$ und $\lambda \in \mathbb{K}$. Dann sind auch die Reihen $\sum_{n=1}^{\infty}(a_n + b_n)$, $\sum_{n=1}^{\infty}(a_n - b_n)$ und $\sum_{n=1}^{\infty} \lambda a_n$ konvergent, und es gilt

$$\sum_{n=1}^{\infty}(a_n + b_n) = \sum_{n=1}^{\infty} a_n + \sum_{n=1}^{\infty} b_n \ , \ \sum_{n=1}^{\infty}(a_n - b_n) = \sum_{n=1}^{\infty} a_n - \sum_{n=1}^{\infty} b_n \text{ und } \sum_{n=1}^{\infty} \lambda a_n = \lambda \sum_{n=1}^{\infty} a_n.$$

Beweis: Diese Aussagen ergeben sich ziemlich unmittelbar aus den Grenzwertsätzen für Folgen. Wir beschränken uns deshalb auf die Reihe $\sum_{n=1}^{\infty}(a_n + b_n)$. Seien $(s_n)_{n\in\mathbb{N}}$ und $(t_n)_{n\in\mathbb{N}}$ die Folgen der Partialsummen von $\sum_{n=1}^{\infty} a_n$ bzw. $\sum_{n=1}^{\infty} b_n$. Durch einen leichten Induktionsbeweis zeigt man zunächst, dass $s_n + t_n$ die n-te Partialsumme von $\sum_{n=1}^{\infty}(a_n + b_n)$ ist. Auf Grund des Grenzwertsatzes für Summen gilt nun

$$\sum_{n=1}^{\infty}(a_n + b_n) \quad = \quad \lim_{n\to\infty}(s_n + t_n) \quad = \quad \lim_{n\to\infty} s_n + \lim_{n\to\infty} t_n \quad = \quad \sum_{n=1}^{\infty} a_n + \sum_{n=1}^{\infty} b_n.$$

Der Beweis für $\sum_{n=1}^{\infty}(a_n - b_n)$ und $\sum_{n=1}^{\infty} \lambda a_n$ verläuft völlig analog. $\square$

(4.8) Proposition Seien $\sum_{n=1}^{\infty} a_n$ und $\sum_{n=1}^{\infty} b_n$ konvergente Reihen über $\mathbb{R}$. Gilt $a_n \le b_n$ für alle $n \in \mathbb{N}$, dann folgt $\sum_{n=1}^{\infty} a_n \le \sum_{n=1}^{\infty} b_n$.

Beweis: Sei $(s_n)_{n\in\mathbb{N}}$ die Folge der Partialsummen von $\sum_{n=1}^{\infty} a_n$ und $(t_n)_{n\in\mathbb{N}}$ die Folge der Partialsummen von $\sum_{n=1}^{\infty} b_n$. Durch Anwendung der Voraussetzung $a_n \le b_n$ und vollständige Induktion beweist man, dass $s_n \le t_n$ für alle $n \in \mathbb{N}$ gilt. Die entsprechende Aussage für Folgen, Satz (2.11), liefert nun die Behauptung. $\square$

Wir werden nun einige Kriterien für die Konvergenz von Reihen herleiten.

(4.9) Satz *(Cauchy-Kriterium)*
Eine Reihe $\sum_{n=1}^{\infty} a_n = (s_n)_{n\in\mathbb{N}}$ in $\mathbb{K}$ konvergiert genau dann, wenn für jedes $\varepsilon \in \mathbb{R}^+$ ein $N \in \mathbb{N}$ existiert, so dass $|\sum_{k=m+1}^{n} a_k| = |s_n - s_m| < \varepsilon$ für alle $m, n \in \mathbb{N}$ mit $n \ge m \ge N$ gilt.

Beweis: Das angegebene Kriterium ist genau dann erfüllt, wenn die Folge $(s_n)_{n\in\mathbb{N}}$ der Partialsummen eine Cauchyfolge bildet. Nach Satz (3.14), der, wie Anschluss bemerkt wurde, sowohl für $\mathbb{R}$ als auch für $\mathbb{C}$ gültig ist, ist die Cauchyfolgen-Eigenschaft äquivalent zur Konvergenz der Folge. □

(4.10) Folgerung Sei $\sum_{n=1}^{\infty} a_n$ eine konvergente Reihe in $\mathbb{R}$. Dann gilt $\lim_{n\to\infty} a_n = 0$.

Beweis: Sei $\varepsilon \in \mathbb{R}^+$ vorgegeben. Nach dem Cauchy-Kriterium gibt es ein $N \in \mathbb{N}$ mit $|\sum_{k=m+1}^{n} a_k| < \varepsilon$ für alle $m, n \geq N$. Wählen wir $n = m+1$, dann erhalten wir insbesondere $|a_{m+1}| < \varepsilon$ für alle $m \geq N$, also $|a_m| < \varepsilon$ für alle $m \geq N+1$. □

(4.11) Satz *(Monotoniekriterium)*
Sei $(a_n)_{n\in\mathbb{N}}$ eine Folge nichtnegativer reeller Zahlen. Genau dann konvergiert die Reihe $\sum_{n=1}^{\infty} a_n$, wenn die Folge $(s_n)_{n\in\mathbb{N}}$ der Partialsummen beschränkt ist.

Beweis: Setzen wir voraus, dass die Folge $(s_n)_{n\in\mathbb{N}}$ beschränkt ist. Wegen $a_n \geq 0$ für alle $n \in \mathbb{N}$ ist sie außerdem monoton wachsend. Also folgt die Konvergenz aus Satz (3.6). Umgekehrt folgt aus der Konvergenz von $(s_n)_{n\in\mathbb{N}}$ die Beschränktheit dieser Folge. □

(4.12) Satz Die ***harmonische Reihe*** $\sum_{n=1}^{\infty} \frac{1}{n}$ divergiert.

Beweis: Wir zeigen, dass die Folge $(s_n)_{n\in\mathbb{N}}$ der Partialsummen unbeschränkt ist. Dabei besteht die wesentliche Idee darin, die natürlichen Zahlen in Blöcke $\{2\}$, $\{3,4\}$, $\{5,6,7,8\}$ usw. einzuteilen. Allgemein hat ein solcher Block die Form $A_k = \{n \in \mathbb{N} \mid 2^{k-1}+1 \leq n \leq 2^k\}$. Für $p \in \mathbb{N}$ gilt somit

$$s_{2^p} = \sum_{n=1}^{2^p} \tfrac{1}{n} = 1 + \sum_{k=1}^{p} \left(\sum_{n=2^{k-1}+1}^{2^k} \tfrac{1}{n} \right) \geq 1 + \sum_{k=1}^{p} \left(\sum_{n=2^{k-1}+1}^{2^k} \frac{1}{2^k} \right)$$

$$= 1 + \sum_{k=1}^{p} 2^{k-1} \cdot \frac{1}{2^k} = 1 + \sum_{k=1}^{p} \tfrac{1}{2} = 1 + \tfrac{1}{2}p = \tfrac{1}{2}(p+2).$$

Dabei haben wir im zweiten Schritt die Summe lediglich in einzelne Blöcke aufgeteilt, d.h. wir haben Teilsummen von $1 + \frac{1}{2} + \frac{1}{3} + \ldots$ in der Form

$$1 + \tfrac{1}{2} + (\tfrac{1}{3} + \tfrac{1}{4}) + (\tfrac{1}{5} + \tfrac{1}{6} + \tfrac{1}{7} + \tfrac{1}{8}) + \ldots$$

zusammengefasst. Beim dritten Schritt in der Ungleichungskette oben beachte man, dass für alle $n \in A_k$ die Abschätzung $\frac{1}{n} \geq \frac{1}{2^k}$ gilt, und dass A_k insgesamt 2^{k-1} Elemente enthält. Wegen $\lim_p \frac{1}{2}(p+2) = +\infty$ zeigt die Ungleichungskette, dass die Partialsummen s_{2^p} tatsächlich unbeschränkt sind und die Reihe somit nach Satz (4.11) divergiert. □

(4.13) Satz *(Leibniz-Kriterium)*
Sei $(a_n)_{n\in\mathbb{N}}$ eine monoton fallende Folge nichtnegativer reeller Zahlen mit der Eigenschaft $\lim_n a_n = 0$. Dann konvergiert die Reihe $\sum_{n=0}^{\infty}(-1)^n a_n$.

Beweis: Die Teilfolge $(s_{2n})_{n\geq 0}$ der Folge $(s_n)_{n\in\mathbb{N}}$ der Partialsummen ist monoton fallend, denn es gilt

$$s_{2n+2} - s_{2n} \quad = \quad a_{2n+2} - a_{2n+1} \quad \leq \quad 0 \qquad \text{für alle} \quad n \in \mathbb{N}_0.$$

Die Rechnung $s_{2n+3} - s_{2n+1} = -a_{2n+3} + a_{2n+2} \geq 0$ zeigt, dass die Teilfolge $(s_{2n+1})_{n\geq 0}$ monoton wächst. Außerdem sind beide Folgen beschränkt, denn es gilt $s_{2n} = s_{2n-1} + a_{2n} \geq s_{2n-1} \geq s_1$ für alle $n \in \mathbb{N}$ und $s_{2n+1} = s_{2n} - a_{2n+1} \leq s_{2n} \leq s_0$ für alle $n \in \mathbb{N}_0$. Also existieren nach Satz (3.6) die beiden Grenzwerte

$$s = \lim_{n\to\infty} s_{2n} \qquad \text{und} \qquad s' = \lim_{n\to\infty} s_{2n+1}.$$

Wegen $s - s' = \lim_{n\to\infty}(s_{2n} - s_{2n+1}) = \lim_{n\to\infty} a_{2n+1} = 0$ stimmen s und s' überein.

Sei nun $\varepsilon \in \mathbb{R}^+$ vorgegeben. Wegen der Konvergenz der beiden Teilfolgen gibt es nun $N_1, N_2 \in \mathbb{N}$ mit $|s_{2n} - s| < \varepsilon$ für alle $n \geq N_1$ und $|s_{2n+1} - s| < \varepsilon$ für alle $n \geq N_2$. Sei $N = \max\{2N_1, 2N_2 + 1\}$ und $n \in \mathbb{N}$ mit $n \geq N$. Ist n gerade, $n = 2k$, dann gilt $2k \geq N \geq 2N_1$ und somit $k \geq N_1$, also $|s_n - s| = |s_{2k} - s| < \varepsilon$. Ist n dagegen ungerade, $n = 2k+1$, dann folgt $k \geq N_2$ und $|s_n - s| = |s_{2k+1} - s| < \varepsilon$. □

Nach dem Leibniz-Kriterium sind beispielsweise die Reihen

$$\sum_{n=1}^{\infty}(-1)^{n+1}\frac{1}{n} \qquad \text{und} \qquad \sum_{n=0}^{\infty}(-1)^n\frac{1}{2n+1} \qquad \text{konvergent.}$$

Die erste dieser beiden Reihen wird ***alternierende harmonische Reihe*** genannt. Bevor wir weitere Konvergenzkriterien aufstellen, führen wir einen stärkeren Konvergenzbegriff ein.

(4.14) Definition Eine Reihe $\sum_{n=1}^{\infty} a_n$ in $\mathbb{K}$ wird ***absolut konvergent*** genannt, wenn die Reihe $\sum_{n=1}^{\infty} |a_n|$ der Absolutbeträge konvergiert.

Zunächst überprüfen wir, dass diese Bedingung tatsächlich stärker ist als die herkömmliche Konvergenz.

(4.15) Proposition Jede absolut konvergente Reihe konvergiert im Sinne von Def. (4.4).

Beweis: Durch vollständige Induktion über n verallgemeinert man die Dreiecksungleichung $|x + y| \leq |x| + |y|$ zunächst von zwei auf n Summanden, d.h. man beweist

$$\left|\sum_{k=1}^{n} a_k\right| \quad \leq \quad \sum_{k=1}^{n} |a_k|$$

für beliebiges $n \in \mathbb{N}$ und $a_1, ..., a_n \in \mathbb{R}$. Sei nun $\sum_{n=1}^{\infty} a_n$ eine absolut konvergente Reihe und $\varepsilon \in \mathbb{R}^+$ vorgegeben. Da $\sum_{n=1}^{\infty} |a_n|$ konvergiert, existiert nach dem Cauchy-Kriterium (Satz (4.9)) ein $N \in \mathbb{N}$, so dass für alle $m, n \geq N$ die Abschätzung $\sum_{k=m+1}^{n} |a_k| < \varepsilon$ erfüllt ist. Es folgt

$$\left| \sum_{k=m+1}^{n} a_k \right| \quad \leq \quad \sum_{k=m+1}^{n} |a_k| \quad < \quad \varepsilon.$$

Die erneute Anwendung des Cauchy-Kriteriums liefert die gewöhnliche Konvergenz der Reihe. □

(4.16) Satz *(Majorantenkriterium)*
Sei $\sum_{n=1}^{\infty} c_n$ eine konvergente Reihe in $\mathbb{R}_+$ und $\sum_{n=1}^{\infty} a_n$ eine Reihe in $\mathbb{K}$ mit $|a_n| \leq c_n$ für fast alle (d.h. alle bis auf endlich viele) $n \in \mathbb{N}$. Dann konvergiert $\sum_{n=1}^{\infty} a_n$ absolut.

Beweis: Zunächst bemerken wir, dass (wie bei den Folgen) die Modifikation von endlich vielen Gliedern einer Reihe nichts an deren Konvergenzverhalten ändert. Deshalb können wir die Ungleichung $|a_n| \leq c_n$ für *alle* $n \in \mathbb{N}$ voraussetzen. Wir zeigen nun, dass eine Reihe unter der angegebenen Bedingung das Cauchy-Kriterium erfüllt.

Sei $\varepsilon \in \mathbb{R}^+$ vorgegeben. Auf Grund der Konvergenz von $\sum_{n=1}^{\infty} c_n$ gibt es ein $N \in \mathbb{N}$ mit

$$\sum_{k=m+1}^{n} c_k \quad = \quad \left| \sum_{k=m+1}^{n} c_k \right| \quad < \quad \varepsilon \qquad \text{für} \quad n, m \geq N.$$

Wegen $|a_k| \leq c_k$ für alle $k \in \mathbb{N}$ folgt $\sum_{k=m+1}^{n} |a_k| \leq \sum_{k=m+1}^{n} c_k < \varepsilon$. Damit sind die Voraussetzungen des Cauchy-Kriteriums nachgewiesen. □

(4.17) Folgerung *(Minorantenkriterium)*
Sei $\sum_{n=1}^{\infty} c_n$ eine divergente Reihe in $\mathbb{R}_+$ und $\sum_{n=1}^{\infty} a_n$ eine Reihe in $\mathbb{R}$ mit $a_n \geq c_n$ für fast alle $n \in \mathbb{N}$. Dann ist auch $\sum_{n=1}^{\infty} a_n$ divergent.

Beweis: Wie im vorherigen Beweis können wir die Ungleichung $a_n \geq c_n$ für alle $n \in \mathbb{N}$ voraussetzen. Wäre $\sum_{n=1}^{\infty} a_n$ konvergent, dann könnten wir das Majorantenkriterium anwenden und erhielten damit auch die Konvergenz der Reihe $\sum_{n=1}^{\infty} c_n$, im Widerspruch zur Voraussetzung. □

Wir illustrieren die Anwendung von Majoranten- und Minorantenkriterium an einer Reihe von Beispielen.

Beispiel 1: Für alle $k \geq 2$ ist die Reihe $\sum_{n=1}^{\infty} \frac{1}{n^k}$ konvergent.

Beweis: Durch direktes Nachrechnen haben wir bereits gezeigt, dass

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

eine konvergente Reihe ist. Da eine konvergente Reihe auch nach Multiplikation mit einem skalaren Vielfachen noch konvergent ist (vgl. Satz (4.7)) ist auch die Reihe über $(\frac{2}{n(n+1)})_{n\in\mathbb{N}}$ konvergent. Nun gelten für alle $n \in \mathbb{N}$ die Ungleichungen

$$n \geq 1 \quad \Leftrightarrow \quad n^2 \geq n \quad \Leftrightarrow \quad 2n^2 \geq n^2 + n \quad \Leftrightarrow \quad n^2 \geq \tfrac{1}{2}n(n+1).$$

also $\frac{1}{n^k} \leq \frac{1}{n^2} \leq \frac{2}{n(n+1)}$ für alle $n \in \mathbb{N}$. Das Majorantenkriterium beweist somit die Konvergenz der Reihe. □

Beispiel 2: Die Reihe $\displaystyle\sum_{n=1}^{\infty} \frac{1}{2n+7}$ ist divergent.

Beweis: Für alle $n \in \mathbb{N}$ mit $n \geq 7$ gilt

$$2n + 7 \leq 3n \quad \Leftrightarrow \quad \frac{1}{3n} \leq \frac{1}{2n+7.}$$

Wäre $\sum_{n=1}^{\infty} \frac{1}{3n}$ konvergent, dann nach Satz (4.7) auch $\sum_{n=1}^{\infty} \frac{1}{n}$, was wir bereits widerlegt haben. Also haben wir mit $\sum_{n=1}^{\infty} \frac{1}{3n}$ eine divergente Minorante für unsere Reihe gefunden. □

Beispiel 3: Die Reihe $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n^2-5}$ ist konvergent. *Beweis:* Für alle $n \in \mathbb{N}$ mit $n \geq 4$ gilt $n^2 \geq 10 \Leftrightarrow \frac{1}{2}n^2 \geq 5 \Leftrightarrow n^2 - 5 \geq \frac{1}{2}n^2$, also $\frac{1}{n^2-5} \leq \frac{2}{n^2}$. Somit ist $\sum_{n=1}^{\infty} \frac{2}{n^2}$ eine konvergente Majorante für $\sum_{n=1}^{\infty} \frac{1}{n^2-5}$. □

(4.18) Satz *(Quotientenkriterium)*
Sei $(a_n)_{n\in\mathbb{N}}$ eine Folge in $\mathbb{K}$ mit $a_n \neq 0$ für fast alle $n \in \mathbb{N}$. Ferner existiere ein $\theta \in \mathbb{R}$, $0 < \theta < 1$, so dass die Ungleichung

$$\left|\frac{a_{n+1}}{a_n}\right| \quad \leq \quad \theta \qquad \text{für fast alle} \quad n \in \mathbb{N}$$

erfüllt ist. Dann ist $\displaystyle\sum_{n=1}^{\infty} a_n$ absolut konvergent.

Beweis: Wir setzen voraus, dass die Bedingungen $a_n \neq 0$ und $|\frac{a_{n+1}}{a_n}| \leq \theta$ für alle $n \in \mathbb{N}$ erfüllt sind. Dann gilt $|a_2| \leq \theta|a_1|$, und durch vollständige Induktion beweist man $|a_{n+1}| \leq \theta^n|a_1|$ für alle $n \in \mathbb{N}_0$. Die Reihe $\sum_{n=0}^{\infty} \theta^n|a_1|$ stimmt bis auf den konstanten Faktor $|a_1|$ mit der geometrischen Reihe überein, konvergiert also gegen den Wert $|a_1|/(1-\theta)$. Somit ist $\sum_{n=0}^{\infty} \theta^n|a_1|$ eine konvergente Majorante der Reihe $\sum_{n=1}^{\infty} a_n$. □

Mit Hilfe des Quotientenkriteriums kann beispielsweise gezeigt werden, dass die Reihe

$$\sum_{n=1}^{\infty} \frac{n}{2^n}$$

konvergiert. Setzen wir nämlich $a_n = \frac{n}{2^n}$, dann gilt

$$\left|\frac{a_{n+1}}{a_n}\right| \quad = \quad \frac{a_{n+1}}{a_n} \quad = \quad \frac{\frac{n+1}{2^{n+1}}}{\frac{n}{2^n}} \quad = \quad \frac{n+1}{2^{n+1}} \cdot \frac{2^n}{n} \quad = \quad \frac{n+1}{2n} \quad = \quad \tfrac{1}{2}\left(1 + \frac{1}{n}\right)$$

und für $n \geq 2$ gilt

$$\tfrac{1}{n} \leq \tfrac{1}{2} \quad \Leftrightarrow \quad 1 + \tfrac{1}{n} \leq \tfrac{3}{2} \quad \Leftrightarrow \quad |a_{n+1}|/|a_n| = \tfrac{1}{2}(1 + \tfrac{1}{n}) \leq \tfrac{3}{4}.$$

Die Voraussetzungen des Quotientenkriteriums sind also mit der Konstante $\theta = \frac{3}{4}$ erfüllt.

Wir wenden das Quotientenkriterium nochmals auf eine etwas kompliziertere Situation an, indem wir die Konvergenz der Reihe

$$\sum_{n=1}^{\infty} \frac{n^3 + 5}{4^n}$$

beweisen. Setzen wir $a_n = \frac{n^3+5}{4^n}$ für $n \in \mathbb{N}$, dann gilt für $n \geq 2$ die Ungleichung $n + 1 \leq \frac{3}{2}n$, und somit

$$\begin{aligned} \left|\frac{a_{n+1}}{a_n}\right| &= \frac{(n+1)^3 + 5}{4^{n+1}} \cdot \frac{4^n}{n^3 + 5} = \tfrac{1}{4} \cdot \frac{(n+1)^3 + 5}{n^3 + 5} \leq \tfrac{1}{4} \cdot \frac{(\frac{3}{2}n)^3 + 5}{n^3 + 5} \\ &\leq \tfrac{1}{4} \cdot \frac{\frac{27}{8}n^3 + 5 \cdot \frac{27}{8}}{n^3 + 5} = \tfrac{1}{4} \cdot \frac{\frac{27}{8}(n^3 + 5)}{n^3 + 5} = \tfrac{1}{4} \cdot \tfrac{27}{8} = \tfrac{27}{32} < 1 \quad , \end{aligned}$$

wobei im vierten Schritt die Abschätzung $\frac{27}{8} \leq 5 \cdot \frac{27}{8}$ verwendetet wurde. Die Voraussetzungen des Quotientenkriteriums sind also mit der Konstanten $\theta = \frac{27}{32}$ erfüllt.

Man beachte, dass es für die Anwendung des Quotientenkriteriums ***nicht*** genügt, lediglich die Ungleichung $|\frac{a_{n+1}}{a_n}| < 1$ für fast alle $n \in \mathbb{N}$ zu beweisen. Betrachten wir zum Beispiel die Reihe $\sum_{n=1}^{\infty} a_n$ mit $a_n = \frac{1}{n}$. Es gilt

$$\left|\frac{a_{n+1}}{a_n}\right| = \frac{n}{n+1} < 1 \qquad \text{für alle} \quad n \in \mathbb{N} \quad ,$$

aber die Reihe ist, wie wir gezeigt haben, divergent.

Mit Hilfe der Quotienten $|\frac{a_{n+1}}{a_n}|$ kann auch die ***Divergenz*** von Folgen nachgewiesen werden. Sei $(a_n)_{n\in\mathbb{N}}$ eine Folge mit $a_n \neq 0$ und $|\frac{a_{n+1}}{a_n}| \geq 1$ für fast alle $n \in \mathbb{N}$. Dann divergiert die Reihe $\sum_{n=1}^{\infty} a_n$. Ist nämlich $n_0 \in \mathbb{N}$ so gewählt, dass die Bedingungen für alle $n \geq n_0$ erfüllt sind, dann erhält man die Ungleichungen

$$|a_{n_0}| \leq |a_{n_0+1}| \leq |a_{n_0+2}| \leq \ldots \quad ,$$

so dass die Folge a_n also keine Nullfolge sein kann. Deshalb ist die Reihe $\sum_{n=1}^{\infty} a_n$ divergent.

(4.19) Folgerung Sei $(a_n)_{n\in\mathbb{N}}$ eine Folge in $\mathbb{K}$ und $n_0 \in \mathbb{N}$, so dass $a_n \neq 0$ für $n \geq n_0$ gilt und die Folge $(|\frac{a_{n+1}}{a_n}|)_{n\geq n_0}$ konvergiert, mit $\alpha = \lim\limits_{n\to\infty} |\frac{a_{n+1}}{a_n}|$.

(i) Gilt $\alpha < 1$, dann konvergiert die Reihe $\displaystyle\sum_{n=1}^{\infty} a_n$ absolut.

(ii) Gilt $\alpha > 1$, dann divergiert sie.

Im Fall $\alpha = 1$ ist keine Aussage möglich, d.h. es gibt Reihen mit dieser Eigenschaft, die konvergieren, und andere, die divergieren.

Den Beweis stellen wir als Übungsaufgabe. Das folgende Kriterium ist dem Quotientenkriterium ähnlich, basiert aber auf der Wurzelfunktion.

(4.20) Satz *(Wurzelkriterium)*
Sei $(a_n)_{n\in\mathbb{N}}$ eine Folge in $\mathbb{K}$ mit der Eigenschaft, dass ein $\theta \in \mathbb{R}$, $0 < \theta < 1$ existiert, so dass $\sqrt[n]{|a_n|} \leq \theta$ für fast alle $n \in \mathbb{N}$ erfüllt ist. Dann konvergiert die Reihe $\sum_{n=1}^{\infty} a_n$ absolut.

Beweis: Sei die Bedingung $\sqrt[n]{|a_n|} \leq \theta$ für alle $n \in \mathbb{N}$ erfüllt. Dann gilt $|a_n| \leq \theta^n$ für alle $n \in \mathbb{N}$, also ist wiederum die geometrische Reihe $\sum_{n=0}^{\infty} \theta^n$ eine konvergente Majorante. □

Wie beim Quotientenkriterium gilt, dass im Falle von $\sqrt[n]{|a_n|} \geq 1$ für unendlich viele $n \in \mathbb{N}$ die Reihe $\sum_{n=1}^{\infty} a_n$ divergiert. Denn dann gilt unendlich oft $|a_n| \geq 1$, die Folge $(a_n)_{n\in\mathbb{N}}$ kann also keine Nullfolge sein. Genau wie in Folgerung (4.19) beweist man auch

(4.21) Folgerung Sei $(a_n)_{n\in\mathbb{N}}$ eine Folge mit der Eigenschaft, dass die Folge $(\sqrt[n]{|a_n|})_{n\in\mathbb{N}}$ konvergiert, mit $\alpha = \lim\limits_{n\to\infty} \sqrt[n]{|a_n|}$.

(i) Gilt $\alpha < 1$, dann konvergiert die Reihe $\displaystyle\sum_{n=1}^{\infty} a_n$ absolut.

(ii) Gilt $\alpha > 1$, dann divergiert sie.

Wiederum ist im Fall $\alpha = 1$ ist keine Aussage möglich.

Weder Quotienten- noch Wurzelkriterium sind stark genug, um die Divergenz der harmonischen Reihe oder die Konvergenz der Reihe $\sum_{n=1}^{\infty} \frac{1}{n^2}$ nachzuweisen, denn es gilt

$$\lim_{n\to\infty} \frac{\frac{1}{n+1}}{\frac{1}{n}} = \lim_{n\to\infty} \frac{n}{n+1} = 1 \qquad \text{und ebenso} \qquad \lim_{n\to\infty} \frac{(\frac{1}{n+1})^2}{(\frac{1}{n})^2} = \lim_{n\to\infty} \left(\frac{n}{n+1}\right)^2 = 1.$$

Wir formulieren nun ein Kriterium, dass für diese Reihen geeignet ist. Der Beweis ähnelt dem Verfahren, mit dem wir die Divergenz der harmonischen Reihe $\sum_{n=1}^{\infty} \frac{1}{n}$ nachgewiesen haben.

(4.22) Satz *(Verdichtungskriterium)*

Sei $(a_n)_{n\in\mathbb{N}}$ eine monoton fallende Folge nichtnegativer reeller Zahlen, und sei die Folge $(b_k)_{k\in\mathbb{N}_0}$ definiert durch $b_k = 2^k a_{2^k}$ für alle $k \in \mathbb{N}_0$. Dann gilt die Äquivalenz

$$\sum_{n=1}^{\infty} a_n \text{ ist konvergent} \quad \Leftrightarrow \quad \sum_{k=0}^{\infty} b_k \text{ ist konvergent.}$$

Beweis: „⇐“ Nach Voraussetzung ist die Reihe $\sum_{k=0}^{\infty} b_k$ konvergent. Zu zeigen ist, dass die Reihe $\sum_{n=1}^{\infty} a_n$ konvergiert. Für alle $k \in \mathbb{N}_0$ und $m \in \mathbb{N}_0$ mit $0 \le m < 2^k$ gilt auf Grund der Monotonie der Folge $(a_n)_{n\in\mathbb{N}}$ die Abschätzung $a_{2^k} \ge a_{2^k+m}$. Setzen wir also $c_{2^k+m} = a_{2^k}$ für $k \in \mathbb{N}_0$ und $0 \le m < 2^k$, dann gilt $c_n \ge a_n$ für alle $n \in \mathbb{N}$. Wenn wir zeigen können, dass die Reihe $\sum_{n=1}^{\infty} c_n$ konvergiert, dann folgt die (absolute) Konvergenz von $\sum_{n=1}^{\infty} a_n$ aus dem Majorantenkriterium (Satz (4.16)). Dafür wiederum genügt es zu zeigen, dass die Folge der Partialsummen $\sum_{n=1}^{p} c_n$ beschränkt ist. Sei $p \in \mathbb{N}$ vorgegeben und $r \in \mathbb{N}$ so gewählt, dass $2^{r-1} \le p < 2^r$ erfüllt ist. Dann gilt die Abschätzung

$$\begin{aligned}\sum_{n=1}^{p} c_n \quad &\le \quad \sum_{n=1}^{2^r-1} c_n \quad = \quad \sum_{k=0}^{r-1}\sum_{m=0}^{2^k-1} c_{2^k+m} \quad = \quad \sum_{k=0}^{r-1}\sum_{m=0}^{2^k-1} a_{2^k} \quad = \quad \sum_{k=0}^{r-1} 2^k a_{2^k} \\ &= \quad \sum_{k=0}^{r-1} b_k \quad \le \quad \sum_{k=0}^{\infty} b_k \quad < \quad +\infty.\end{aligned}$$

Sämtliche Partialsummen $\sum_{n=1}^{p} c_n$ sind also durch den endlichen Wert $\sum_{k=0}^{\infty} b_k$ nach oben beschränkt.

„⇒“ Diese Richtung beweisen wir durch Kontraposition: Wir setzen voraus, dass die Reihe $\sum_{k=0}^{\infty} b_k$ divergiert und leiten daraus die Divergenz von $\sum_{n=1}^{\infty} a_n$ ab. Für alle $k \in \mathbb{N}_0$ und $m \in \mathbb{N}_0$ mit $0 \le m < 2^k$ definieren wir $c_{2^k+m} = a_{2^{k+1}}$. Auf Grund der Monotonie der Folge $(a_n)_{n\in\mathbb{N}}$ gilt dann $c_n \le a_n$ für alle $n \in \mathbb{N}$. Wenn wir zeigen können, dass $\sum_{n=1}^{\infty} c_n$ divergiert, dann folgt die Divergenz von $\sum_{n=1}^{\infty} a_n$ aus dem Minorantenkriterium (Satz (4.17)). Sei $p \in \mathbb{N}$ mit $p \ge 2$ vorgegeben und $r \in \mathbb{N}$ so gewählt, dass $2^r \le p < 2^{r+1}$ erfüllt ist. Dann erhalten wir für die Partialsumme $\sum_{n=1}^{p} c_n$ die untere Abschätzung

$$\begin{aligned}\sum_{n=1}^{p} c_n \quad &\ge \quad \sum_{n=1}^{2^r-1} c_n \quad = \quad \sum_{k=0}^{r-1}\sum_{m=0}^{2^k-1} c_{2^k+m} \quad = \quad \sum_{k=0}^{r-1}\sum_{m=0}^{2^k-1} a_{2^{k+1}} \quad = \quad \sum_{k=0}^{r-1} 2^k a_{2^{k+1}} \\ &= \quad \tfrac{1}{2}\sum_{k=0}^{r-1} 2^{k+1} a_{2^{k+1}} \quad = \quad \tfrac{1}{2}\sum_{k=0}^{r-1} b_{k+1} \quad = \quad \tfrac{1}{2}\left(\sum_{k=0}^{r} b_k - a_1\right).\end{aligned}$$

Weil die Summe $\sum_{k=0}^{r} b_k$ auf Grund der Voraussetzung für $r \to \infty$ gegen $+\infty$ läuft, gilt auf Grund der Abschätzung dasselbe für die Partialsumme $\sum_{n=1}^{p} c_n$ für $p \to \infty$. Also ist $\sum_{n=1}^{\infty} c_n$ und damit auch $\sum_{n=1}^{\infty} a_n$ divergent. □

Als Anwendung des Verdichtungskriteriums beweisen wir, dass die Reihe $\sum_{n=1}^{\infty} n^{-\alpha}$ für jedes $\alpha \in \mathbb{Q}$ mit $\alpha > 1$ konvergiert. Setzen wir $a_n = n^{-\alpha}$ für alle $n \in \mathbb{N}$ und $\theta = 2^{1-\alpha}$, so gilt $0 \le \theta < 1$ und

$$\sum_{k=0}^{\infty} 2^k a_{2^k} \quad = \quad \sum_{k=0}^{\infty} 2^k \frac{1}{2^{k\alpha}} \quad = \quad \sum_{k=0}^{\infty} \left(2^{1-\alpha}\right)^k \quad = \quad \sum_{k=0}^{\infty} \theta^k.$$

Die Rechung zeigt, dass $\sum_{k=0}^{\infty} 2^k a_{2^k}$ mit der geometrischen Reihe zum Wert θ übereinstimmt und folglich konvergiert. Nach dem Verdichtungskriterium ist damit auch $\sum_{n=1}^{\infty} a_n$ konvergent. (Aussage und Beweis sind auch für irrationale $\alpha \in \mathbb{R}$ mit $\alpha > 1$ gültig, Allerdings haben wir $n^{-\alpha}$ für irrationales α noch nicht definiert, weil dafür der Stetigkeitsbegriff benötigt wird.)

Gelegentlich wird bei Rechnungen mit einer Reihe $\sum_{n=1}^{\infty} a_n$ die Reihenfolge der Glieder a_n abgeändert, und es soll sichergestellt werden, dass diese Operation keine Auswirkung auf das Konvergenzverhalten und den Grenzwert der Reihe hat. Wenn bei der ursprünglichen Reihe *absolute* Konvergenz vorliegt, so ist dies tatsächlich der Fall.

(4.23) Satz *(Umordnungssatz)*
Sei $\sum_{n=1}^{\infty} a_n$ eine absolut konvergente Reihe in $\mathbb{K}$. Ist $\tau : \mathbb{N} \to \mathbb{N}$ eine bijektive Abbildung, dann ist auch die Reihe $\sum_{n=1}^{\infty} a_{\tau(n)}$ absolut konvergent, und es gilt

$$\sum_{n=1}^{\infty} a_n \quad = \quad \sum_{n=1}^{\infty} a_{\tau(n)}.$$

Beweis: Für jedes $n \in \mathbb{N}$ sei $A_n = \{1, ..., n\}$, $B_n = \tau(A_n)$ und $b_n = \max(B_n)$. Zunächst beweisen wir die absolute Konvergenz von $\sum_{n=1}^{\infty} a_{\tau(n)}$. Es gilt

$$\sum_{k=1}^{n} \left|a_{\tau(k)}\right| \quad = \quad \sum_{k \in B_n} |a_k| \quad \leq \quad \sum_{k=1}^{b_n} |a_k| \quad \leq \quad \sum_{n=1}^{\infty} |a_n| \quad < \quad +\infty.$$

Dies zeigt, dass die Folge $\sum_{k=1}^{n} |a_{\tau(k)}|$ der Partialsummen von $\sum_{n=1}^{\infty} a_{\tau(n)}$ beschränkt ist und die Reihe somit nach Satz (4.11) tatsächlich absolut konvergiert.

Nun muss noch gezeigt werden, dass beide Reihen gegen denselben Wert konvergieren. Für jedes $n \in \mathbb{N}$ sei $s_n = \sum_{k=1}^{n} a_k$ und $s'_n = \sum_{k=1}^{n} a_{\tau(k)}$. Nach den Grenzwertsätzen genügt es dazu, die Gleichung $\lim_n (s_n - s'_n) = 0$ nachzuweisen. Sei also $\varepsilon \in \mathbb{R}^+$ vorgegeben. Weil die Reihe $\sum_{n=1}^{\infty} a_n$ absolut konvergiert, gibt es ein $N \in \mathbb{N}$ mit

$$\sum_{k=N+1}^{\infty} |a_k| \quad < \quad \tfrac{1}{2}\varepsilon.$$

Weil die Abbildung τ surjektiv ist, existiert ein $N' \geq N$ mit $B_{N'} = \tau(A_{N'}) \supseteq A_N$. Für alle $n \geq N'$ gilt dann ebenfalls $B_n \supseteq A_N$, was zu $A_N = B_n \cap A_N$ äquivalent ist. Für solche n erhalten wir

$$s_n - s'_n \quad = \quad \sum_{k \in A_n} a_k - \sum_{k \in B_n} a_k \quad = \quad \sum_{k \in A_n} a_k - \sum_{k \in B_n \cap A_N} a_k - \sum_{k \in B_n \setminus A_N} a_k \quad = \quad \sum_{k \in A_n \setminus A_N} a_k - \sum_{k \in B_n \setminus A_N} a_k.$$

Es folgt

$$|s_n - s'_n| \quad \leq \quad \sum_{k \in A_n \setminus A_N} |a_k| + \sum_{k \in B_n \setminus A_N} |a_k| \quad \leq \quad \sum_{k=N+1}^{\infty} |a_k| + \sum_{k=N+1}^{\infty} |a_k| \quad < \quad \tfrac{1}{2}\varepsilon + \tfrac{1}{2}\varepsilon \quad = \quad \varepsilon.$$

Weil $\varepsilon \in \mathbb{R}^+$ beliebig vorgegeben war, folgt daraus $\lim_n (s_n - s'_n) = 0$ wie gewünscht. □

Wir illustrieren an einem Beispiel, dass sich der Wert einer Reihe bei Umordnung ändern kann, wenn keine absolute, sondern nur „gewöhnliche“ Konvergenz vorliegt, die man auch als ***bedingte*** Konvergenz bezeichnet. Wie wir mit Hilfe des Leibniz-Kriteriums gezeigt haben, konvergiert die alternierende harmonische Reihe

$$\sum_{n=1}^{\infty} (-1)^{n+1} \frac{1}{n} \quad = \quad \sum_{k=1}^{\infty} \left(\frac{1}{2k-1} - \frac{1}{2k} \right).$$

Sie konvergiert aber nicht absolut, denn die harmonische Reihe $\sum_{n=1}^{\infty} |(-1)^{n+1} \frac{1}{n}| = \sum_{n=1}^{\infty} \frac{1}{n}$ ist divergent. Summieren wir nun die alternierende harmonische Reihe in der umgeordneten Form

$$1 - \tfrac{1}{2} - \tfrac{1}{4} \quad + \tfrac{1}{3} - \tfrac{1}{6} - \tfrac{1}{8} \quad + \tfrac{1}{5} - \tfrac{1}{10} - \tfrac{1}{12} + \quad \tfrac{1}{7} - \tfrac{1}{14} - \tfrac{1}{16} \quad + \tfrac{1}{9} - \tfrac{1}{18} - \tfrac{1}{20} \quad + \tfrac{1}{11} - \tfrac{1}{22} - \tfrac{1}{24} \quad + \ldots$$

Anstatt auf einen ungerade (positiven) Term nur einen geraden (negativen) Term folgen zu lassen, ordnen wir ihm jeweils zwei gerade Terme zu. Insgesamt wird die Summe also in „Dreierblöcken“ organisiert, die in der Form

$$\frac{1}{2k-1}-\frac{1}{2(2k-1)}-\frac{1}{4k} \quad = \quad \frac{2-1}{2(2k-1)}-\frac{1}{2\cdot 2k} \quad = \quad \tfrac{1}{2}\left(\frac{1}{2k-1}-\frac{1}{2k}\right)$$

geschrieben werden können. Als Summe über diese Terme erhalten wir

$$\sum_{k=1}^{\infty}\left(\frac{1}{2k-1}-\frac{1}{2(2k-1)}-\frac{1}{4k}\right) \quad = \quad \sum_{k=1}^{\infty}\tfrac{1}{2}\left(\frac{1}{2k-1}-\frac{1}{2k}\right) \quad = \quad \tfrac{1}{2}\sum_{n=1}^{\infty}(-1)^{n+1}\frac{1}{n}.$$

Gegenüber der herkömmlichen Reihenfolge hat sich der Grenzwert also halbiert.

Für eine Reihe von Anwendungen, zum Beispiel die Definition der Exponentialfunktion, die wir in einem der nächsten Kapitel behandeln werden, sind auch ***Produkte*** von Reihen von Interesse. Es ist allerdings nicht zu erwarten, dass eine so einfache Regel wie für Summen auch für Produkte gültig ist; schließlich ist die Gleichung $(\sum_{k=0}^{n} a_k)(\sum_{k=0}^{n} b_k) = \sum_{k=0}^{n} a_k b_k$ noch nicht einmal für endliche Summen richtig. Statt dessen erhält man durch Ausmultiplizieren

$$\left(\sum_{k=0}^{n} a_k\right)\left(\sum_{\ell=0}^{n} b_\ell\right) \quad = \quad \sum_{k=0}^{n}\sum_{\ell=0}^{n} a_k b_\ell \quad = \quad \sum_{r=0}^{2n} c_r \qquad \text{mit} \qquad c_r = \sum_{k=0}^{r} a_{r-k} b_k$$

wenn wir $a_k = b_\ell = 0$ für $k, \ell > n$ setzen. Diese Gleichung für endliche Produkte motiviert die folgende Definition.

(4.24) Definition Das ***Cauchy-Produkt*** zweier Reihen $\sum_{n=0}^{\infty} a_n$ und $\sum_{n=0}^{\infty} b_n$ im Körper $\mathbb{K}$ ist die Reihe $\sum_{n=0}^{\infty} c_n$ mit den Termen c_n gegeben durch

$$c_n \quad = \quad \sum_{k=0}^{n} a_{n-k} b_k \qquad \text{für alle } n \in \mathbb{N}_0.$$

Wir untersuchen nun, unter welchen Voraussetzungen die Reihe $\sum_{n=0}^{\infty} c_n$ das Produkt der beiden Ausgangsreihen liefert.

(4.25) Satz Seien $\sum_{n=0}^{\infty} a_n$ und $\sum_{n=0}^{\infty} b_n$ zwei konvergente Reihen in $\mathbb{K}$ mit Grenzwerten a und b, wobei wir bei mindestens einer der beiden Reihen absolute Konvergenz voraussetzen. Sei $\sum_{n=0}^{\infty} c_n$ das Cauchy-Produkt der beiden Reihen. Dann gilt

$$\sum_{n=0}^{\infty} c_n \quad = \quad ab \quad = \quad \left(\sum_{n=0}^{\infty} a_n\right)\left(\sum_{n=0}^{\infty} b_n\right) \quad ,$$

das Cauchy-Produkt konvergiert also gegen den Wert ab. Konvergieren beide Reihen absolut, dann gilt dasselbe auch für das Cauchy-Produkt.

Beweis: Ohne Beschränkung der Allgemeinheit können wir annehmen, dass $\sum_{n=0}^{\infty} a_n$ die absolut konvergente Reihe ist. Wir setzen $\nu = 1 + \sum_{n=0}^{\infty} |a_n|$. Für jedes $n \in \mathbb{N}_0$ definieren wir die Partialsummen $s_n = \sum_{k=0}^{n} a_k$, $t_n = \sum_{k=0}^{n} b_k$ und $u_n = \sum_{k=0}^{n} c_k$. Für jedes $n \in \mathbb{N}_0$ sei $U_n = \{(k,\ell) \in \mathbb{N}_0^2 \,|\, k+\ell \leq n\}$. Nach Definition des Cauchy-Produkts gilt $c_n = \sum_{k=0}^{n} a_{n-k} b_k$ und somit

$$u_n = \sum_{j=0}^{n} c_j = \sum_{j=0}^{n} \sum_{k=0}^{j} a_{j-k} b_k = \sum_{(j,k)\in U_n} a_j b_k = \sum_{j=0}^{n} \sum_{k=0}^{j} a_{n-j} b_k = \sum_{j=0}^{n} a_{n-j} \left(\sum_{k=0}^{j} b_k \right) = \sum_{j=0}^{n} a_{n-j} t_j.$$

Für die Differenz $u_n - ab$ ergibt sich damit

$$u_n - ab = \sum_{j=0}^{n} a_{n-j} t_j - ab = \sum_{j=0}^{n} a_{n-j}(t_j - b) + \sum_{j=0}^{n} a_{n-j} b - ab = \sum_{j=0}^{n} a_{n-j}(t_j - b) + (s_n - a)b.$$

Sei nun $\varepsilon \in \mathbb{R}^+$ vorgegeben. Wegen $\lim_n (s_n - a) = 0$ gibt es ein $N \in \mathbb{N}$ mit $|(s_n - a)b| < \frac{1}{3}\varepsilon$ für alle $n \geq N$. Für diese n betrachten wir nun die Summenzerlegung

$$\sum_{j=0}^{n} a_{n-j}(t_j - b) = \sum_{j=0}^{N-1} a_{n-j}(t_j - b) + \sum_{j=N}^{n} a_{n-j}(t_j - b).$$

Wegen $\lim_j (t_j - b) = 0$ können wir nach eventueller Vergrößerung von N davon ausgehen, dass $|t_j - b| \leq \frac{\varepsilon}{3\nu}$ für alle $j \geq N$ gilt. Der Betrag der zweiten Summe kann dann durch

$$\frac{\varepsilon}{3\nu} \sum_{j=N}^{n} |a_{n-j}| \leq \frac{\varepsilon}{3\nu} \sum_{n=0}^{\infty} |a_n| < \frac{\varepsilon}{3\nu} \cdot \nu = \tfrac{1}{3}\varepsilon$$

abgeschätzt werden. Auf Grund der Konvergenz der Reihe $\sum_{n=0}^{\infty} a_n$ gilt außerdem $\lim_n a_{n-j} = 0$ für $0 \leq j \leq N-1$. Deshalb existiert ein $M \in \mathbb{N}$ mit $|\sum_{j=0}^{N-1} a_{n-j}(t_j - b)| < \frac{1}{3}\varepsilon$ für alle $n \geq M$. Für alle $n \geq \max\{M, N\}$ gilt somit $|u_n - ab| < \frac{1}{3}\varepsilon + \frac{1}{3}\varepsilon + \frac{1}{3}\varepsilon = \varepsilon$. Damit ist die Gleichung $\lim_n u_n = ab$ bewiesen.

Sind die Reihen $\sum_{n=0}^{\infty} a_n$ und $\sum_{n=0}^{\infty} b_n$ beide absolut konvergent, dann konvergieren die Reihen $\sum_{n=0}^{\infty} |a_n|$ und $\sum_{n=0}^{\infty} |b_n|$. Auf Grund der soeben bewiesenen Aussage ist dann auch $\sum_{n=0}^{\infty} c'_n$ mit $c'_n = \sum_{k=0}^{n} |a_{n-k}||b_k|$ konvergent. Aus der Dreiecksungleichung folgt $|c_n| \leq c'_n$ für alle $n \in \mathbb{N}_0$. Aus dem Majorantenkriterium, Satz (4.16), folgt damit die absolute Konvergenz von $\sum_{n=0}^{\infty} c_n$. □

Eine wichtige, konkrete Anwendung der Theorie der konvergenten Reihen liegt in der Darstellung reeller Zahlen durch endliche oder unendliche Dezimalbrüche. Weil es keinen zusätzlichen Aufwand bedeutet, beschränken wir uns nicht auf die gewohnten Darstellungen zur Basis 10, sondern betrachten als Basis eine beliebige natürliche Zahl b mit $b > 1$.

(4.26) Definition Sei $b \in \mathbb{N}$ mit $b > 1$. Ein ***b*-adischer Bruch*** ist eine Reihe der Form

$$\sum_{n=-r}^{\infty} a_n b^{-n}$$

wobei $r \in \mathbb{Z}$ sowie $0 \leq a_n < b$ und $a_n \in \mathbb{Z}$ für alle $n \in \mathbb{Z}$ mit $n \geq -r$ gilt. Dabei bezeichnet man a_n als die ***n*-te Stelle** des b-adischen Bruchs.

Symbolisch stellt man b-adische Brüche in der Form $a_{-r}...a_{-1}a_0, a_1a_2...$ dar. Für $b = 10$ erhält man so die bekannte Dezimaldarstellung reeller Zahlen. Zum Beispiel gilt

$$0,33333... \quad = \quad \sum_{n=1}^{\infty} 3 \cdot 10^{-n} \quad = \quad \sum_{n=0}^{\infty} 3 \cdot 10^{-(n+1)} \quad = \quad \tfrac{3}{10} \cdot \sum_{n=0}^{\infty} 10^{-n} \quad = \quad \tfrac{3}{10} \cdot \frac{1}{1-\frac{1}{10}} \quad = \quad \tfrac{3}{10} \cdot \tfrac{10}{9} \quad = \quad \tfrac{1}{3}.$$

Dabei haben wir im zweiten Schritt die Formel für den Wert der geometrische Reihe $\sum_{n=0}^{\infty} q^n = \frac{1}{1-q}$ verwendet. Gibt es ein N mit $a_n = 0$ für $n \geq N$, dann können diese Stellen bei der symbolischen Darstellung auch weggelassen werden, d.h. $0,497$ und $0,49700000...$ bezeichnen denselben b-adischen Bruch. Sich periodisch wiederholende Stellen kennzeichnet man durch einen Überstrich: $0,\overline{3} = 0,33333...$ Wir betrachten einige weitere Beispiele.

(i) Die 10-adischen Brüche $1,\overline{0}$ und $0,\overline{9}$ stellen beide die reelle Zahl 1 dar. Im zweiten Fall folgt dies aus der Rechnung

$$0,\overline{9} \quad = \quad \sum_{n=1}^{\infty} 9 \cdot 10^{-n} \quad = \quad \tfrac{9}{10} \cdot \sum_{n=0}^{\infty} \left(\tfrac{1}{10}\right)^n \quad = \quad \tfrac{9}{10} \cdot \frac{1}{1-\frac{1}{10}} \quad = \quad \tfrac{9}{10} \cdot \frac{1}{\frac{9}{10}} \quad = \quad \tfrac{9}{10} \cdot \tfrac{10}{9} \quad = \quad 1.$$

Dies zeigt, dass die Darstellung einer reellen Zahl als b-adischer Bruch jedenfalls nicht völlig eindeutig ist.

(ii) Die b-adischen Darstellungen der reellen Zahl $\frac{1}{7}$ lautet $0,\overline{001}$ zur Basis $b = 2$. Zur Basis 10 erhalten wir $0,\overline{142857}$ für $b = 10$, und $0,1$ zur Basis $b = 7$.

Beweis: Wir rechnen lediglich nach, dass $0,\overline{001}$ die 2-adische Darstellung von $\frac{1}{7}$ ist, und überlassen den Rest dem Leser als Übungsaufgabe. Sei $(a_n)_{n\in\mathbb{N}}$ die Folge der Stellen des 2-adischen Bruchs. Dann gilt $a_{3k-2} = a_{3k-1} = 0$ und $a_{3k} = 1$ für alle $k \in \mathbb{N}$. Für jedes $n \in \mathbb{N}$ sei s_n die n-te Partialsumme der Reihe $\sum_{n=1}^{\infty} a_n 2^{-n}$. Wir erhalten

$$0,\overline{001} \quad = \quad \sum_{n=1}^{\infty} a_n 2^{-n} \quad = \quad \lim_{n\to\infty} s_n \quad = \quad \lim_{n\to\infty} s_{3n} \quad = \quad \lim_{n\to\infty} \sum_{k=1}^{n} \left(\sum_{\ell=0}^{2} a_{3k-\ell} 2^{\ell-3k} \right)$$

$$= \quad \sum_{k=1}^{\infty} \left(\sum_{\ell=0}^{2} a_{3k-\ell} 2^{\ell-3k} \right) \quad = \quad \sum_{k=1}^{\infty} (a_{3k-2} 2^{2-3k} + a_{3k-1} 2^{1-3k} + a_{3k} 2^{-3k}) \quad =$$

$$\sum_{k=1}^{\infty} (4a_{3k-2} + 2a_{3k-1} + a_{3k}) 2^{-3k} \quad = \quad \sum_{k=1}^{\infty} (4 \cdot 0 + 2 \cdot 0 + 1) 2^{-3k} \quad = \quad \sum_{k=1}^{\infty} 8^{-k} \quad =$$

$$\tfrac{1}{8} \sum_{k=0}^{\infty} 8^{-k} \quad = \quad \tfrac{1}{8} \frac{1}{1-\frac{1}{8}} \quad = \quad \tfrac{1}{8} \cdot \tfrac{8}{7} \quad = \quad \tfrac{1}{7}.$$

Dabei wurde im dritten Schritt Folgerung (3.11) angewendet: Ist die Folge $(s_n)_{n\in\mathbb{N}}$ konvergent, dann ist mit $(s_n)_{n\in\mathbb{N}}$ auch (s_{3n}) beschränkt, und kann wie $(s_n)_{n\in\mathbb{N}}$ nur einen Häufungspunkt besitzen, nämlich denselben wie $(s_n)_{n\in\mathbb{N}}$. Daraus folgt, dass die Grenzwerte beider Folgen, sofern sie existieren, übereinstimmen müssen. Dass sie existieren, ergibt sich unabhängig von der Rechnung auch aus dem nachfolgenden Satz. □

(4.27) Satz Jeder b-adische Bruch konvergiert gegen eine reelle Zahl.

Beweis: Sei $(a_n)_{n\geq -r}$ die Folge der Stellen und $(s_n)_{n\geq -r}$ die Folge der Partialsummen der Reihe $\sum_{n=-r}^{\infty} a_n b^{-n}$. Wir beweisen die Konvergenz der Reihe mit Hilfe des Cauchy-Kriteriums, Satz (4.9). Ist $\varepsilon \in \mathbb{R}^+$ vorgegeben, dann existiert ein $N \in \mathbb{N}$ mit $b^{-N} < \varepsilon$. Seien nun $n, m \in \mathbb{N}$ mit $m, n \geq N$ gegeben, wobei wir $n \geq m$ voraussetzen. Dann gilt

$$\begin{aligned} |s_n - s_m| &= \sum_{k=m+1}^{n} a_k b^{-k} \leq \sum_{k=m+1}^{n} (b-1)b^{-k} \leq \sum_{k=0}^{n-m-1} (b-1)b^{-k-m-1} \\ &\leq (b-1)b^{-m-1} \sum_{k=0}^{n-m-1} b^{-k} \leq (b-1)b^{-m-1} \sum_{k=0}^{\infty} (\tfrac{1}{b})^k \leq \\ (b-1)b^{-m-1} \frac{1}{1-\frac{1}{b}} &= (b-1)b^{-m-1} \frac{b}{b-1} = b^{-m} < \varepsilon. \end{aligned}$$

□

(4.28) Satz Jedes $a \in \mathbb{R}$ lässt sich in der Form $\pm s$ darstellen, wobei s den Wert eines b-adischen Bruchs $\sum_{n=-r}^{\infty} a_n b^{-n}$ bezeichnet.

Beweis: Wir können uns auf den Fall beschränken, dass $a \geq 0$ ist. Nach Satz (1.10) existiert ein $r \in \mathbb{N}$ mit $b^{r+1} > a$. Wir beweisen nun durch vollständige Induktion über $n \geq -r$, dass es jeweils Zahlen a_k in $\{0, ..., b-1\}$ gibt, so dass für $s_n = \sum_{k=-r}^{n} a_k b^{-k}$ jeweils die Abschätzungen

$$s_n \leq a < s_n + b^{-n} \qquad \text{erfüllt sind.}$$

Wir starten mit dem Induktionsanfang $n = -r$. Nach Voraussetzung gilt $0 \leq a < b^{r+1}$. Um zunächst eine geeignete Zahl a_{-r} zu bestimmen, betrachten wir im Intervall $\left[0, b^{r+1}\right[$ die Unterteilungspunkte $c\,b^r$ mit $c \in \{0, ..., b-1\}$. Da a in diesem Intervall enthalten ist, gibt es ein eindeutig bestimmtes $a_{-r} \in \{0, ..., b-1\}$ mit $a_{-r} b^r \leq a < (a_{-r}+1)b^r$. Für $s_{-r} = a_{-r} b^r$ ist die Bedingung $s_{-r} \leq a < s_{-r} + b^r$ dann offenbar erfüllt.

Gehen wir nun davon aus, dass $a_{-r}, ..., a_n$ bereits konstruiert sind. Nach Induktionsvoraussetzung gilt $s_n \leq a < s_n + b^{-n}$ für $s_n = \sum_{k=-r}^{n} a_k b^{-k}$. Wir betrachten nun im Intervall $[s_n, s_n + b^{-n}[$ die Unterteilungspunkte $s_n + c\,b^{-(n+1)}$ mit $c \in \{0, ..., b-1\}$. Wieder finden wir ein eindeutig bestimmtes $a_{n+1} \in \{0, ..., b-1\}$ mit der Eigenschaft, dass a im Teilintervall

$$\left[s_n + a_{n+1} b^{-(n+1)}, s_n + (a_{n+1}+1)b^{-(n+1)}\right[$$

liegt. Für $s_{n+1} = s_n + a_{n+1} b^{-(n+1)}$ gilt dann $s_{n+1} \leq a < s_{n+1} + b^{-(n+1)}$. Damit ist der Induktionsbeweis abgeschlossen. Für die konstruierte Folge $(s_n)_{n\geq -r}$ gilt nun $\lim_n s_n = a$. Wählen wir nämlich für ein vorgegebenes $\varepsilon \in \mathbb{R}^+$ ein $N \in \mathbb{N}$ mit $b^{-N} < \varepsilon$, dann gilt für alle $n \geq N$ wegen $s_n \leq a < s_n + b^{-n}$ die Abschätzung

$$|a - s_n| = a - s_n < (s_n + b^{-n}) - s_n = b^{-n} < \varepsilon.$$

□

(4.29) Lemma Sei $(a_n)_{n\in\mathbb{N}}$ eine Folge mit $a_n \in \mathbb{Z}$, $a_n \leq b-1$ für alle $n \in \mathbb{N}$. Dann gilt im Falle der Konvergenz die Abschätzung $\sum_{n=1}^{\infty} a_n b^{-n} \leq 1$, und Gleichheit genau dann, wenn $a_n = b-1$ für alle $n \in \mathbb{N}$ ist.

Beweis: Durch Proposition (4.8) erhalten wir zunächst die Abschätzung

$$\sum_{n=1}^{\infty} a_n b^{-n} \quad \leq \quad \sum_{n=1}^{\infty} (b-1)b^{-n} \quad = \quad (b-1)\sum_{n=1}^{\infty} b^{-n} \quad = \quad \frac{b-1}{b}\sum_{n=0}^{\infty} b^{-n} \quad =$$

$$\frac{b-1}{b}\frac{1}{1-\frac{1}{b}} \quad = \quad \frac{b-1}{b}\cdot\frac{b}{b-1} \quad = \quad 1.$$

Die Rechnung zeigt, dass im Fall $a_n = b-1$ für alle $n \in \mathbb{N}$ die Gleichung $\sum_{n=1}^{\infty} a_n b^{-n} = 1$ erfüllt ist. Setzen wir nun umgekehrt die Gleichung voraus und nehmen an, dass $a_{n_0} < b-1$ für ein $n_0 \in \mathbb{N}$ gilt. Dann folgt

$$\sum_{n=1}^{\infty} a_n b^{-n} \quad = \quad \lim_{n\to\infty} \sum_{k=1}^{n} a_k b^{-k} \quad \leq \quad \lim_{n\to\infty} \left(\sum_{k=1}^{n_0-1} (b-1)b^{-k} + a_{n_0} b^{-n_0} + \sum_{k=n_0+1}^{n} (b-1)b^{-k} \right)$$

$$= \quad \lim_{n\to\infty} \left(\sum_{k=1}^{n_0-1} (b-1)b^{-k} + (b-1)b^{-n_0} + \sum_{k=n_0+1}^{n} (b-1)b^{-k} + (a_{n_0} - b + 1)b^{-n_0} \right)$$

$$= \quad \lim_{n\to\infty} \left(\sum_{k=1}^{n} (b-1)b^{-k} \right) + (a_{n_0} - b + 1)b^{-n_0} \quad = \quad 1 + (a_{n_0} - b + 1)b^{-n_0} \quad < \quad 1$$

im Widerspruch zur angenommenen Gleichung. □

Der folgende Satz zeigt, inwieweit die Darstellung von reellen Zahlen als b-adische Brüche eindeutig ist.

(4.30) Satz Seien $c, d \in \mathbb{R}$, $c, d \geq 0$ und $c = \sum_{n=-r}^{\infty} c_n b^{-n}$, $d = \sum_{n=-r}^{\infty} d_n b^{-n}$ Darstellungen dieser Zahlen als b-adische Brüche. Genau dann gilt $c = d$, wenn einer der folgenden drei Bedingungen für die b-adischen Brüche zutrifft:

(i) Es gilt $c_n = d_n$ für alle $n \in \mathbb{Z}$, $n \geq -r$

(ii) Es gibt ein $N \in \mathbb{Z}$, $N \geq -r$, so dass $c_n = d_n$ für alle $n < N$, $d_N = c_N + 1$ und $c_n = b-1$, $d_n = 0$ für alle $n > N$.

(iii) Es gibt ein $N \in \mathbb{Z}$, $N \geq -r$, so dass $c_n = d_n$ für alle $n < N$, $c_N = d_N + 1$ und $c_n = 0$, $d_n = b-1$ für alle $n > N$.

Beweis: „⇐“ Im Fall (i) ist die Identität $c = d$ klar, denn gleiche Reihen konvergieren gegen denselben Wert. Im zweiten Fall erhalten wir die Identität durch die Rechnung

$$d - c \quad = \quad \sum_{n=N}^{\infty} (d_n - c_n)b^{-n} \quad = \quad b^{-N} + \sum_{n=N+1}^{\infty} (1-b)b^{-n} \quad =$$

$$b^{-N} + b^{-(N+1)}(1-b)\sum_{n=0}^{\infty} b^{-n} \quad = \quad b^{-N} + \frac{1-b}{b^{N+1}}\cdot\frac{b}{b-1} \quad = \quad b^{-N} - b^{-N} = 0.$$

Die Rechnung im Fall (iii) läuft vollkommen analog.

„⇒“ Sei $c = d$ vorausgesetzt. Angenommen, die b-adischen Brüche $\sum_{n=-r}^{\infty} c_n b^{-n}$ und $\sum_{n=-r}^{\infty} d_n b^{-n}$ sind nicht identisch, und $N \in \mathbb{Z}$ mit $N \geq -r$ ist minimal mit der Eigenschaft $c_N \neq d_N$. Betrachten wir zunächst den Fall, dass $c_N < d_N$ ist. Durch Umformungen erhalten wir

$$c = d \quad \Leftrightarrow \quad \sum_{n=-r}^{\infty} (d_n - c_n) b^{-n} = 0 \quad \Leftrightarrow \quad (d_N - c_N) b^{-N} + \sum_{n=N+1}^{\infty} (d_n - c_n) b^{-n} = 0$$

$$\Leftrightarrow \quad \sum_{n=1}^{\infty} (c_{N+n} - d_{N+n}) b^{-n} = d_N - c_N$$

Wegen $0 \leq c_{N+n}, d_{N+n} < b$ gilt für die Differenzen $c_{N+n} - d_{N+n} \leq b - 1$. Nach Lemma (4.29) folgt $d_N - c_N \leq 1$, also $d_N - c_N = 1$ auf Grund der Annahme $c_N < d_N$. Nochmalige Anwendung von Lemma (4.29) liefert $c_{N+n} = b - 1$ und $d_{N+n} = 0$ für alle $n \in \mathbb{N}$, d.h. die Bedingung (ii) ist erfüllt. Genauso führt die Annahme $c_N > d_N$ auf Bedingung (iii). □

Für jede Basis b findet man Zahlen $x \in \mathbb{R}$ mit zwei verschiedenen Darstellungen als b-adische Brüche. Andererseits gibt es immer auch Zahlen mit nur einer b-adischen Darstellung. Ist zum Beispiel $b = 10$, dann sind durch

$$1,5 \quad = \quad 1,5\overline{0} \qquad \text{und} \qquad 1,4\overline{9}$$

zwei Darstellungen von $x = \frac{3}{2}$ als 10-adischer Bruch gegeben. gegeben. Dagegen ist die Darstellung von $x = \frac{2}{3}$ zur Basis 10 eindeutig bestimmt, nämlich gleich $0,\overline{6}$.

Für $b = 2$ erhalten wir durch $1,\overline{0}$ und $0,\overline{1}$ zwei verschiedene Darstellungen von $x = 1$. Die Darstellung von $x = \frac{1}{5}$ gegeben durch $0,\overline{0011}$ ist wiederum eindeutig bestimmt. Man kann sich mit Hilfe von Satz (4.30) leicht überlegen, dass das Problem der Mehrdeutigkeit grundsätzlich nur bei rationalen Zahlen auftreten kann: Eine reelle Zahl mit einer ***abbrechenden*** Dezimalbruchentwicklung, für die also ein $N \in \mathbb{N}$ mit $N \geq -r$ und $c_n = 0$ für alle $n \geq N$ existiert, ist stets eine rationale Zahl.

§ 5. *Funktionsgrenzwerte und Stetigkeit*

Inhaltsübersicht

Nachdem wir uns in den letzten Kapiteln ausführlich mit Grenzwerten von Folgen beschäftigt haben, wird dieses Konzept nun hier verwendet, um *Grenzwerte von Funktionen* zu definieren. Intuitiv bedeutet die Gleichung $\lim_{x\to a} f(x) = c$ für eine Funktion f, dass sich $f(x)$ dem Wert c nähert, wenn x dem Wert a angenähert wird. Präzise definiert man dies mit Hilfe der Folgenkonvergenz.

Die Funktion f wird *stetig* im Punkt a genannt, wenn a zum Definitionsbereich von f gehört und die Gleichung $\lim_{x\to a} f(x) = f(a)$ erfüllt ist. In der Schulmathematik wird dies meist durch die Formulierung umschrieben, dass die Funktion f an der Stelle a nicht „springt“.

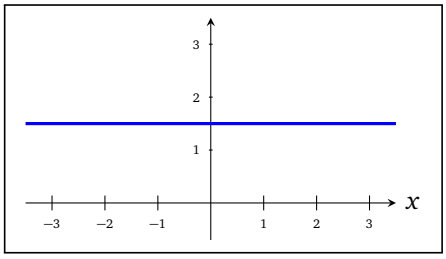
Zunächst betrachten wir eine Reihe von Beispielen für Grenzwerte und Stetigkeit und beschäftigen wir uns mit der Frage, wie Grenzwerte von Funktionen berechnet bzw. wie sich die Stetigkeit einer Funktion nachweisen lässt. Anschließend leiten wir mit zwei zentrale Sätze über stetige Funktionen her: den *Zwischenwertsatz* und das *Maximumsprinzip*.

Wichtige Begriffe und Sätze

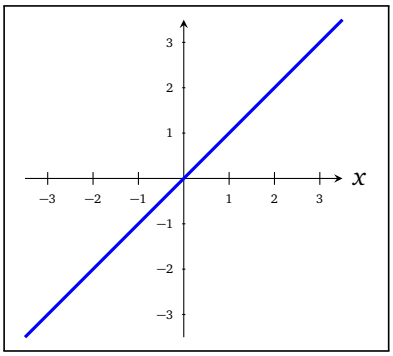
- Polynomfunktionen, rationale Funktionen
- Häufungspunkt, isolierter Punkt, Berührpunkt einer Teilmenge $D \subseteq \mathbb{R}$
- Definition des Grenzwerts $\lim\limits_{x\to a} f(x)$ einer Funktion $f : D \to \mathbb{R}$, wobei a Berührpunkt von D ist
- Summe, Produkt und Quotient von Funktionen
- Stetigkeit einer Funktion in einem Punkt ihres Definitionsbereichs
- Zwischenwertsatz
- Maximumsprinzip
- Stetigkeit der Umkehrfunktion einer stetigen, streng monoton wachsenden oder fallenden Funktion

Sei $D \subseteq \mathbb{R}$. Wir erinnern daran, dass eine (reellwertige) ***Funktion*** auf D nichts weiter als eine Abbildung $f : D \to \mathbb{R}$ ist. Dabei bezeichnet man D als den ***Definitionsbereich*** von f. Folgende Beispiele für Funktionen werden wir im Laufe der Vorlesung immer wieder aufgreifen.

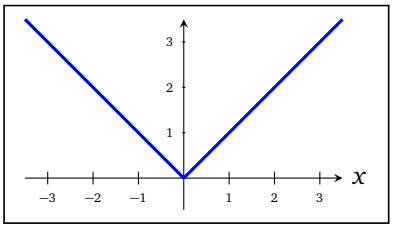
(i) Sei $c \in \mathbb{R}$. Dann ist die Abbildung $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto c$ die ***konstante*** Funktion mit Wert c.



(ii) Die identische Abbildung $\mathrm{id} : \mathbb{R} \to \mathbb{R}$, $x \mapsto x$ ist eine Funktion.



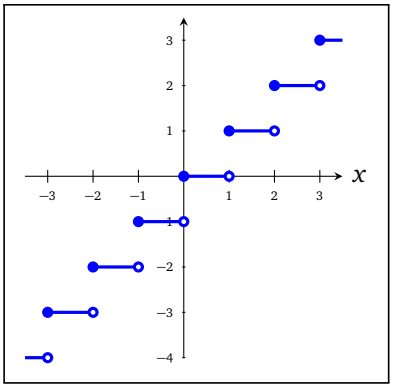
(iii) Die ***Betragsfunktion*** $\mathrm{abs} : \mathbb{R} \to \mathbb{R}$ ist gegeben durch $x \mapsto |x|$.



(iv) Für jedes $x \in \mathbb{R}$ definieren wir die ***untere Gaußklammer*** durch

$$\lfloor x \rfloor \quad = \quad \max \{r \in \mathbb{Z} \mid r \leq x\}.$$

Dann ist durch die Zuordnung $x \mapsto \lfloor x \rfloor$ ebenfalls eine Funktion auf ganz $\mathbb{R}$ definiert.



(v) Die Zuordnungen $\mathbb{R} \to \mathbb{R}$ gegeben durch

$$x \mapsto 2x^2 + 1 \qquad \text{und} \qquad x \mapsto x^3 - 3x^2 + 5x + 1$$

sind reellwertige Funktionen. Allgemeiner bezeichnet man eine Abbildung der Form

$$x \mapsto \sum_{k=0}^{m} a_k x^k \qquad \text{mit} \qquad m \in \mathbb{N}_0, \quad a_0, \ldots, a_m \in \mathbb{R}$$

als (reellwertige) ***Polynomfunktion***.

(vi) Seien $f, g : \mathbb{R} \to \mathbb{R}$ Polynomfunktionen und $D = \{x \in \mathbb{R} \mid g(x) \neq 0\}$. Eine Abbildung der Form

$$h : D \longrightarrow \mathbb{R} \quad , \qquad x \mapsto \frac{f(x)}{g(x)}$$

wird (reelle) ***rationale Funktion*** genannt. Konkrete Beispiele für rationale Funktionen sind

$$\mathbb{R} \longrightarrow \mathbb{R}, \quad x \mapsto \frac{1}{x^2+1} \qquad \text{und} \qquad \mathbb{R} \setminus \{-1, 1\} \longrightarrow \mathbb{R}, \quad x \mapsto \frac{3x+5}{x^2-1}.$$

Sei $D \subseteq \mathbb{R}$ eine beliebige Teilmenge. Eine Folge $(x_n)_{n\in\mathbb{N}}$ reeller Zahlen bezeichnen wir als ***Folge in*** D, wenn $x_n \in D$ für alle $n \in \mathbb{N}$ gilt.

(5.1) Definition Sei $D \subseteq \mathbb{R}$ und $a \in \bar{\mathbb{R}}$. Man nennt a einen ***Häufungspunkt*** von D, wenn eine Folge in $D \setminus \{a\}$ existiert, die (eventuell uneigentlich) gegen a konvergiert. Einen Punkt $a \in D$, der kein Häufungspunkt von D ist, nennt man einen ***isolierten*** Punkt von D.

Ist beispielsweise $D =]0, 1] \cup \{2, 3\}$, dann ist $[0, 1]$ die Menge der Häufungspunkte von D, während 2 und 3 isolierte Punkte sind. Das Beispiel zeigt, dass die Häufungspunkte einer Menge D nicht unbedingt in D enthalten sein müssen. Ein Häufungspunkt a von D mit $a \notin D$ wird ***Berührpunkt*** der Menge D genannt.

(5.2) Proposition Sei $D \subseteq \mathbb{R}$, $a \in \bar{\mathbb{R}}$ und $D_a = D \setminus \{a\}$.

(i) Im Fall $a \in \mathbb{R}$ ist a genau dann ein Häufungspunkt von D, wenn für jedes $\varepsilon \in \mathbb{R}^+$ ein $x \in D_a$ mit $|x - a| < \varepsilon$ existiert.

(ii) Der Punkt $a = +\infty$ ist genau dann Berührpunkt von D, wenn für jedes $\kappa \in \mathbb{R}^+$ ein $x \in D$ mit $x > \kappa$ existiert.

(iii) Der Punkt $a = -\infty$ ist genau dann Berührpunkt von D, wenn für jedes $\kappa \in \mathbb{R}^+$ ein $x \in D$ mit $x < -\kappa$ existiert.

Beweis: zu (i) „⇒“ Ist a ein Häufungspunkt von D, so gibt es eine Folge $(x_n)_{n\in\mathbb{N}}$ in D_a mit $\lim_n x_n = a$. Für vorgegebenes $\varepsilon \in \mathbb{R}^+$ existiert insbesondere ein $n \in \mathbb{N}$ mit $|x_n - a| < \varepsilon$, und x_n liegt in D_a. „⇐“ Erfüllt der Punkt a die angegebene Voraussetzung, so gibt es insbesondere für jedes $n \in \mathbb{N}$ ein $x_n \in D_a$ mit $|x_n - a| < \frac{1}{n}$. Somit ist $(x_n)_{n\in\mathbb{N}}$ eine Folge in D_a. Wir zeigen, dass $\lim_n x_n = a$ gilt. Ist $\varepsilon \in \mathbb{R}^+$ vorgegeben, dann gibt es ein $N \in \mathbb{N}$ mit $\frac{1}{N} < \varepsilon$. Es folgt $|x_n - a| < \frac{1}{n} \leq \frac{1}{N} < \varepsilon$ für alle $n \geq N$ und damit die Behauptung.

zu (ii) „⇒“ Ist $+\infty$ ein Berührpunkt von D, dann gibt es eine Folge $(x_n)_{n\in\mathbb{N}}$ in D mit der Eigenschaft $\lim_n x_n = +\infty$. Ist $\kappa \in \mathbb{R}^+$ vorgegeben, dann gibt es insbesondere ein $n \in \mathbb{N}$ mit $x_n > \kappa$, und x_n liegt in D. „⇐“ Erfüllt D die angegebene Voraussetzung, so exstiert insbesondere für jedes $n \in \mathbb{N}$ ein $x_n \in D$ mit $x_n > n$. Durch $(x_n)_{n\in\mathbb{N}}$ ist dann also eine Folge in D gegegeben. Ist $\kappa \in \mathbb{R}^+$, dann gibt es ein $N \in \mathbb{N}$ mit $N > \kappa$, und es folgt $x_n > n \geq N > \kappa$ für alle $n \geq N$. Damit ist $\lim_n x_n = +\infty$ bewiesen. Der Beweis von (iii) verläuft weitgehend analog. □

(5.3) Definition Seien $D \subseteq \mathbb{R}$, $f : D \to \mathbb{R}$ eine Funktion und $a, c \in \bar{\mathbb{R}}$. Ist a ein Berührpunkt von D, und gilt für ***jede*** Folge $(x_n)_{n\in\mathbb{N}}$ in D die Implikation

$$\lim_{n\to\infty} x_n = a \quad \Rightarrow \quad \lim_{n\to\infty} f(x_n) = c \quad ,$$

so bezeichnet man c als ***Grenzwert*** von f im Punkt a.

Falls a oder c in $\{-\infty, +\infty\}$ liegen, spricht man genauer von einem ***uneigentlichen***, sonst von einem ***eigentlichen*** Grenzwert der Funktion f. Grundsätzlich kann $f : D \to \mathbb{R}$ nur einen (eigentlichen oder uneigentlichen) Grenzwert im Punkt a besitzen. Sind nämlich $c, d \in \bar{\mathbb{R}}$ beides Grenzwerte von f im Punkt a, und ist $(x_n)_n$ eine Folge in D mit $\lim_n x_n = a$, so muss $c = \lim_n f(x_n) = d$ gelten. Deshalb ist es gerechtfertigt, im Fall der Existenz für den Grenzwert von f im Punkt a die Bezeichnung

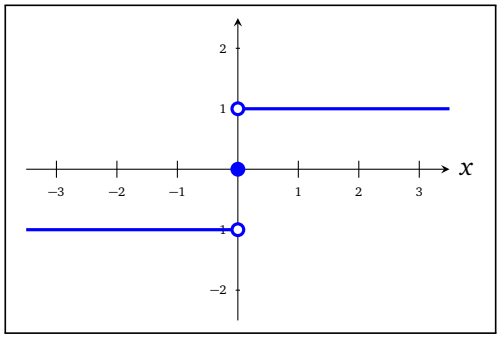
$$\lim_{x\to a} f(x)$$

einzuführen. Es ist aber durchaus möglich, dass $f : D \to \mathbb{R}$ in einem Punkt $a \in \bar{\mathbb{R}}$ keinen eigentlichen oder uneigentlichen Grenzwert besitzt, selbst wenn es sich bei a um einen Berührpunkt von D handelt.

Für das Verständnis der folgenden Anwendungen ist es wichtig, den soeben eingeführten Grenzwertbegriff *für Funktionen* unbedingt vom Grenzwertbegriff *für Folgen*, der bereits in § 8 definiert wurde, zu unterscheiden! Ein Ausdruck wie z.B. $\lim_{n\to 2} x_n$ ist für eine Folge $(x_n)_{n\in\mathbb{N}}$ nicht definiert, und für eine Funktion f hat der Ausdruck $\lim_n f(x_n)$ eine andere Bedeutung als $\lim_{x\to+\infty} f(x)$. Um sich den Unterschied klarzumachen, hilft es, die Definition von $\lim_{x\to+\infty} f(x)$ noch einmal selbstständig auszuformulieren. Das es sich um verschiedene Begriffe handelt, erkennt man schon daran, dass die Definition der Funktionsgrenzwerte auf der Definition der Folgengrenzwerte aufbaut.

Als erstes konkretes Beispiel für einen Funktionsgrenzwert betrachten wir eine Einschränkung der ***Signumsfunktion*** $\operatorname{sgn} : \mathbb{R} \to \mathbb{R}$, die durch

$$\operatorname{sgn}(x) \quad = \quad \begin{cases} -1 & \text{für } x < 0 \\ 0 & \text{für } x = 0 \\ 1 & \text{für } x > 0. \end{cases}$$



gegeben ist. Sei $f = \operatorname{sgn}|_D$ die Einschränkung von sgn auf die Menge $D = \mathbb{R} \setminus \{0\}$. Nehmen wir an, dass $c \in \bar{\mathbb{R}}$ der Grenzwert von f im Punkt $a = 0$ ist. Die beiden Folgen $(x_n)_n$ und $(y_n)_n$ gegeben durch $x_n = \frac{1}{n}$ und $y_n = -\frac{1}{n}$ für alle $n \in \mathbb{N}$ liegen beide in D und konvergieren gegen Null. Nach Definition des Grenzwerts müsste also sowohl

$$c \quad = \quad \lim_{n\to\infty} f(x_n) \quad = \quad \lim_{n\to\infty} f(\tfrac{1}{n}) \quad = \quad \lim_{n\to\infty} 1 \quad = \quad 1$$

als auch

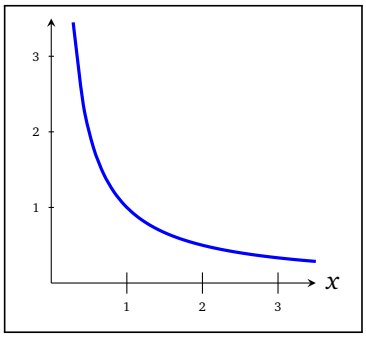
$$c \quad = \quad \lim_{n\to\infty} f(y_n) \quad = \quad \lim_{n\to\infty} f(-\tfrac{1}{n}) \quad = \quad \lim_{n\to\infty} (-1) \quad = \quad -1$$

gelten. Der resultierende Widerspruch $1 = c = -1$ lässt nur den Schluss zu, dass der Grenzwert $\lim\limits_{x\to 0} f(x)$ nicht existiert.

Wir studieren den Grenzwertbegriff für Funktionen anhand von einigen weiteren Beispielen.

Beispiel 1: Sei die Funktion $f : \mathbb{R}^+ \to \mathbb{R}$ gegeben durch $x \mapsto \frac{1}{x}$. Dann gilt

$$\lim_{x\to 0} f(x) = +\infty \quad \text{und} \quad \lim_{x\to+\infty} f(x) = 0.$$



Beweis: Für den Nachweis der ersten Gleichung bemerken wir zunächst, dass 0 ein Berührpunkt von $D = \mathbb{R}^+$ ist. Ist nämlich $(a_n)_{n\in\mathbb{N}}$ gegeben durch $a_n = \frac{1}{n}$ für alle $n \in \mathbb{N}$, dann liegen alle Folgenglieder in D, und es gilt $\lim_n a_n = 0$; außerdem ist $0 \notin D$.

Sei nun $(x_n)_{n\in\mathbb{N}}$ eine *beliebige* Folge in D mit $\lim_n x_n = 0$. Zu zeigen ist $\lim_n f(x_n) = +\infty$. Sei dazu $\kappa \in \mathbb{R}^+$ vorgegeben. Wegen $\lim_n x_n = 0$ gibt es ein $N \in \mathbb{N}$ mit $|x_n| < \kappa^{-1}$ für alle $n \geq N$. Weil die Folge $(x_n)_{n\in\mathbb{N}}$ in D liegt, gilt $x_n > 0$ und somit $|x_n| = x_n$ für alle $n \in \mathbb{N}$. Somit erhalten wir für alle $n \geq N$ die Abschätzung

$$f(x_n) \quad = \quad x_n^{-1} \quad = \quad |x_n|^{-1} \quad > \quad (\kappa^{-1})^{-1} \quad = \quad \kappa.$$

Weil $\kappa \in \mathbb{R}^+$ beliebig vorgegeben war, ist damit $\lim_n f(x_n) = +\infty$ bewiesen.

Überprüfen wir nun den zweiten Grenzwert. Zunächst stellen wir fest, dass $+\infty$ ein Berührpunkt von D ist, denn die Folge $(a_n)_{n\in\mathbb{N}}$ gegeben durch $a_n = n$ für alle $n \in \mathbb{N}$ liegt in D und konvergiert gegen $+\infty$. Sei nun $(x_n)_{n\in\mathbb{N}}$ eine *beliebige* Folge in D mit $\lim_n x_n = +\infty$. Zu zeigen ist $\lim_n f(x_n) = 0$. Sei dazu $\varepsilon \in \mathbb{R}^+$ vorgegeben. Wegen $\lim_n x_n = +\infty$ gibt es ein $N \in \mathbb{N}$ mit $x_n > \varepsilon^{-1}$ für alle $n \geq N$. Wegen $x_n > 0$ gilt auch $f(x_n) = x_n^{-1} > 0$ und somit $|f(x_n)| = f(x_n)$ für alle $n \in \mathbb{N}$. Insgesamt erhalten wir $|f(x_n)| = f(x_n) = x_n^{-1} < \varepsilon$ für alle $n \geq N$. □

Beispiel 2: Sei $D \subseteq \mathbb{R}$ und a ein Berührpunkt von D. Für jedes $c \in \mathbb{R}$ sei $f_c : D \to \mathbb{R}$ die konstante Funktion gegeben durch $f_c(x) = c$ für alle $x \in D$. Dann gilt

$$\lim_{x\to a} f_c(x) = c \qquad \text{und} \qquad \lim_{x\to a} \mathrm{id}_D(x) = a.$$

Beweis: Sei $(x_n)_{n\in\mathbb{N}}$ eine beliebige Folge in D mit $\lim_n x_n = a$. Dann gilt

$$\lim_{n\to\infty} f_c(x_n) = \lim_{n\to\infty} c = c \qquad \text{und} \qquad \lim_{n\to\infty} \mathrm{id}_D(x_n) = \lim_{n\to\infty} x_n = a. \qquad \square$$

(5.4) Proposition Sei f eine Polynomfunktion gegeben durch

$$f : \mathbb{R} \longrightarrow \mathbb{R} \quad , \quad x \mapsto x^m + a_{m-1}x^{m-1} + \ldots + a_1 x + a_0$$

mit $m \in \mathbb{N}$ und $a_0, \ldots, a_{m-1} \in \mathbb{R}$.

(i) Ist m gerade, dann gilt $\lim\limits_{x\to+\infty} f(x) = \lim\limits_{x\to-\infty} f(x) = +\infty$.

(ii) Für ungerades m gilt $\lim\limits_{x\to+\infty} f(x) = +\infty$ und $\lim\limits_{x\to-\infty} f(x) = -\infty$.

Beweis: Wir beschränken uns auf Teil (i) und die Bestimmung des ersten Grenzwertes. Zunächst schätzen wir die Funktionswerte $f(x)$ für $x > 1$ nach unten ab. Für solche x können wir die Funktion f in der Form $f(x) = x^m(1 + g(x))$ schreiben, mit

$$g(x) \quad = \quad \frac{a_{m-1}}{x} + \ldots + \frac{a_0}{x^m} \quad = \quad \sum_{k=0}^{m-1} \frac{a_k}{x^{m-k}}.$$

Sei nun $c = \max\{\, 2m|a_k| \mid 0 \le k < m\}$. Ist $x \in \mathbb{R}$ mit $x > c$ und $x > 1$, so erhalten wir $|x| > 2m|a_k|$ für $0 \le k < m$. Für alle k mit $a_k \neq 0$ gilt

$$\frac{|a_k|}{|x|} \quad < \quad \frac{|a_k|}{2m|a_k|} \quad = \quad \frac{1}{2m} \quad ,$$

und offensichtlich ist $\frac{|a_k|}{|x|} \le \frac{1}{2m}$ auch im Fall $a_k = 0$ erfüllt. Wir erhalten somit

$$|g(x)| \quad = \quad \left|\sum_{k=0}^{m-1} \frac{a_k}{x^{m-k}}\right| \quad \le \quad \sum_{k=0}^{m-1} \left|\frac{a_k}{x^{m-k}}\right| \quad \le \quad \sum_{k=0}^{m-1} \frac{|a_k|}{|x|} \quad < \quad \sum_{k=0}^{m-1} \frac{1}{2m} \quad = \quad m \cdot \frac{1}{2m} \quad = \quad \tfrac{1}{2} \quad ,$$

insbesondere $g(x) > -\frac{1}{2}$. Dies wiederum bedeutet $f(x) = x^m(1+g(x)) > x^m(1+(-\frac{1}{2})) = \frac{1}{2}x^m$ für alle $x > \max\{1, c\}$.

Für den Nachweis des Grenzwerts bemerken wir zunächst, dass $+\infty$ ein Berührpunkt von $D = \mathbb{R}$ ist, denn $(a_n)_{n\in\mathbb{N}}$ gegeben durch $a_n = n$ für alle $n \in \mathbb{N}$ ist eine Folge in D mit $\lim_n a_n = +\infty$. Sei nun $(x_n)_{n\in\mathbb{N}}$ eine *beliebige* Folge in D, die gegen $+\infty$ konvergiert. Zu zeigen ist $\lim_n f(x_n) = +\infty$. Sei dazu $\kappa \in \mathbb{R}^+$ vorgegeben. Dann gibt es ein $N \in \mathbb{N}$ mit $x_n > \max\{1, c, 2\kappa\}$ für alle $n \ge N$. Auf Grund unserer Überlegungen aus dem ersten Teil des Beweises folgt

$$f(x_n) \quad > \quad \tfrac{1}{2}x_n^m \quad \ge \quad \tfrac{1}{2}x_n \quad \ge \quad \tfrac{1}{2}(2\kappa) \quad = \quad \kappa$$

für alle $n \in \mathbb{N}$ mit $n \ge N$. □

(5.5) Definition Seien $D \subseteq \mathbb{R}$ und $f, g : D \to \mathbb{R}$ Funktionen auf D. Dann bezeichnet man die Funktionen $f + g$ und $f\,g$ auf D gegeben durch

$$(f+g)(x) = f(x) + g(x) \quad \text{und} \quad (f\,g)(x) = f(x)g(x) \qquad \text{für } x \in D$$

als ***Summe*** bzw. ***Produkt*** von f und g. Gilt zusätzlich $g(x) \neq 0$ für alle $x \in D$, dann definiert man

$$\left(\frac{f}{g}\right)(x) = \frac{f(x)}{g(x)} \qquad \text{für alle} \quad x \in D.$$

Die Funktion $\frac{f}{g}$ wird der ***Quotient*** von f und g genannt.

(5.6) Proposition Seien $D \subseteq \mathbb{R}$, $f, g : D \to \mathbb{R}$, und sei $a \in \bar{\mathbb{R}}$ ein Berührpunkt von D. Wenn die Grenzwerte

$$\lim_{x\to a} f(x) \qquad \text{und} \qquad \lim_{x\to a} g(x)$$

beide existieren und in $\mathbb{R}$ liegen, also endlich sind, dann gilt

$$\lim_{x\to a}(f+g)(x) = \lim_{x\to a} f(x) + \lim_{x\to a} g(x) \qquad \text{und} \qquad \lim_{x\to a}(fg)(x) = \Big(\lim_{x\to a} f(x)\Big)\Big(\lim_{x\to a} g(x)\Big) \quad ,$$

insbesondere existieren diese Grenzwerte. Ist der Quotient $\frac{f}{g}$ auf D definiert, und ist $\lim\limits_{x\to a} g(x) \neq 0$, so gilt auch

$$\lim_{x\to a}\left(\frac{f}{g}\right)(x) \quad = \quad \frac{\lim\limits_{x\to a} f(x)}{\lim\limits_{x\to a} g(x)}.$$

Beweis: Alle Aussagen lassen sich aus den Grenzwertsätzen ableiten. Wir demonstrieren dies am Beispiel der Funktion $f+g$. Sei $(x_n)_{n\in\mathbb{N}}$ eine Folge in D mit $\lim_n x_n = a$. Durch Anwendung von Satz (2.7) erhalten wir

$$\lim_{n\to\infty}(f+g)(x_n) \quad = \quad \lim_{n\to\infty}(f(x_n)+g(x_n)) \quad = \quad \lim_{n\to\infty} f(x_n) + \lim_{n\to\infty} g(x_n) \quad = \quad \lim_{x\to a} f(x) + \lim_{x\to a} g(x).$$

Damit ist die Gleichung $\lim\limits_{x\to a}(f+g)(x) = \lim\limits_{x\to a} f(x) + \lim\limits_{x\to a} g(x)$ bewiesen. □

(5.7) Definition Seien $D \subseteq \mathbb{R}$, $f : D \to \mathbb{R}$ und $a \in D$. Man sagt, die Funktion f ist ***stetig*** im Punkt a, wenn für jede Folge $(x_n)_n$ in D die Implikation

$$\lim_{n\to\infty} x_n = a \quad \Rightarrow \quad \lim_{n\to\infty} f(x_n) = f(a) \qquad \text{gilt.}$$

Die Funktion f wird stetig genannt, wenn sie in jedem Punkt von D stetig ist.

(5.8) Proposition Seien $D \subseteq \mathbb{R}$, $a \in D$, $D_a = D \setminus \{a\}$ und $\tilde{f} = f|_{D_a}$. Dann sind die folgenden beiden Aussagen äquivalent.

(i) Die Funktion f ist stetig in a.
(ii) Es gilt $\lim\limits_{x\to a} \tilde{f}(x) = f(a)$, oder a ist in D ein isolierter Punkt.

Beweis: „(i) ⇒ (ii)“ Ist a ein isolierter Punkt von D, dann brauchen wir nichts zu zeigen. Nehmen wir nun an, dass a ein Häufungspunkt von D ist. Dann ist a ein Berührpunkt der Menge D_a. Sei $(x_n)_{n\in\mathbb{N}}$ eine Folge in D_a mit $\lim_n x_n = a$. Dann ist $(x_n)_{n\in\mathbb{N}}$ erst recht eine Folge in D, und nach Definition der Stetigkeit gilt

$$\lim_{n\to\infty} \tilde{f}(x_n) \quad = \quad \lim_{n\to\infty} f(x_n) \quad = \quad f(a).$$

Damit ist $\lim\limits_{x\to a} \tilde{f}(x) = f(a)$ bewiesen.

„(ii) ⇒ (i)“ Betrachten wir zunächst den Fall, dass a ein isolierter Punkt von D ist. Sei $(x_n)_{n\in\mathbb{N}}$ eine Folge in D mit $\lim_n x_n = a$. Nach Proposition (5.2) gibt es ein $\varepsilon \in \mathbb{R}^+$, so dass $|x-a| < \varepsilon$ für kein $x \in D_a$ erfüllt ist. Sei nun $N \in \mathbb{N}$

so gewählt, dass $|x_n - a| < \varepsilon$ für alle $n \geq N$ gilt. Für jedes $n \in \mathbb{N}$ mit $n \geq N$ gilt dann $x_n \in D \setminus D_a$, also $x_n = a$. Es folgt $f(x_n) = f(a)$ für alle $n \geq N$ und damit insbesondere $\lim_n f(x_n) = f(a)$.

Sei nun a ein Häufungspunkt von D und wiederum $(x_n)_{n\in\mathbb{N}}$ eine Folge in D mit $\lim_n x_n = a$. Auf Grund der Häufungspunkt-Eigenschaft gibt es nach Proposition (5.2) für jedes $n \in \mathbb{N}$ ein $y_n \in D_a$ mit $|y_n - a| < \frac{1}{n}$. Wir definieren eine neue Folge (x'_n) in D_a durch

$$x'_n \quad = \quad \begin{cases} x_n & \text{falls} \quad x_n \in D_a \\ y_n & \text{falls} \quad x_n = a \end{cases}$$

und überprüfen, dass auch diese Folge gegen a konvergiert. Sei $\varepsilon \in \mathbb{R}^+$ vorgegeben. Wegen $\lim_n x_n = a$ gibt es ein $N_1 \in \mathbb{N}$ mit $|x_n - a| < \varepsilon$ für alle $n \geq N_1$. Sei $N_2 \in \mathbb{N}$ so gewählt, dass $\frac{1}{N_2} < \varepsilon$ ist, und $N = \max\{N_1, N_2\}$. Ist nun $n \in \mathbb{N}$ mit $n \geq N$, dann gilt $|x'_n - a| = |x_n - a| < \varepsilon$ im Fall $x_n \in D_a$ und $|x'_n - a| = |y_n - a| < \frac{1}{n} \leq \frac{1}{N_2} < \varepsilon$ im Fall $x_n = a$. Damit ist $\lim_n x'_n = a$ bewiesen, und auf Grund der Voraussetzung gilt

$$\lim_{n\to\infty} f(x'_n) \quad = \quad \lim_{n\to\infty} \tilde{f}(x'_n) \quad = \quad f(a).$$

Nun zeigen wir, dass daraus $\lim_n f(x_n) = f(a)$ folgt. Sei $\varepsilon \in \mathbb{R}^+$ vorgegeben. Wegen $\lim_n f(x'_n) = f(a)$ gibt es ein $N \in \mathbb{N}$ mit $|f(x'_n) - f(a)| < \varepsilon$ für alle $n \geq N$. Für jedes $n \geq N$ gilt entweder $x_n = x'_n$ oder $x_n = a$. Im ersten Fall ist $|f(x_n) - f(a)| = |f(x'_n) - f(a)| < \varepsilon$; im zweiten Fall ist diese Abschätzung wegen $|f(x_n) - f(a)| = |f(a) - f(a)| = 0$ ebenfalls erfüllt. Damit haben wir $\lim_n f(x_n) = f(a)$ nachgewiesen. □

Wir betrachten einige Beispiele für stetige Funktionen.

(i) Jede konstante Funktion auf einer beliebigen Teilmenge $D \subseteq \mathbb{R}$ ist stetig.

Beweis: Sei $a \in D$ beliebig vorgegeben und $(x_n)_{n\in\mathbb{N}}$ eine Folge in D mit $\lim_n x_n = a$. Dann folgt $\lim_n f(x_n) = \lim_n c = c = f(a)$. Weil die Folge beliebig vorgegeben war, folgt daraus die Stetigkeit von f im Punkt a. □

(ii) Für jede Teilmenge $D \subseteq \mathbb{R}$ ist die identische Abbildung id_D stetig.

Beweis: Sei $a \in D$ und $(x_n)_{n\in\mathbb{N}}$ eine Folge in D mit $\lim_n x_n = a$. Dann gilt $\lim_n \mathrm{id}_D(x_n) = \lim_n x_n = a = \mathrm{id}_D(a)$. Also ist id_D im Punkt a stetig. □

(iii) Die Funktion $\mathrm{abs} : \mathbb{R} \to \mathbb{R}$, $x \mapsto |x|$ ist stetig.

Beweis: Sei $a \in \mathbb{R}$ und $(x_n)_{n\in\mathbb{N}}$ eine Folge in $\mathbb{R}$ mit $\lim_n x_n = a$. In den Übungen haben wir gezeigt, dass für beliebige $x, y \in \mathbb{R}$ jeweils $||x| - |y|| \leq |x - y|$ gilt. Damit zeigen wir, dass $\lim_n \mathrm{abs}(x_n) = \mathrm{abs}(a)$ erfüllt ist. Sei $\varepsilon \in \mathbb{R}^+$ vorgegeben. Wegen $\lim_n x_n = a$ gibt es ein $N \in \mathbb{N}$ mit $|x_n - a| < \varepsilon$ für alle $n \geq N$. Es folgt

$$|\mathrm{abs}(x_n) - \mathrm{abs}(a)| \quad = \quad ||x_n| - |a|| \quad \leq \quad |x_n - a| \quad < \quad \varepsilon$$

für alle $n \geq N$. Weil $\varepsilon \in \mathbb{R}^+$ beliebig vorgegeben war, ist damit $\lim_n \mathrm{abs}(x_n) = \mathrm{abs}(a)$ bewiesen. □

(iv) Die Signumsfunktion $\operatorname{sgn} : \mathbb{R} \to \mathbb{R}$ ist im Punkt 0 unstetig, in allen übrigen Punkten stetig.

Beweis: Zunächst beweisen wir die Unstetigkeit im Punkt 0 und betrachten dazu die Folge $(x_n)_{n\in\mathbb{N}}$ gegeben durch $x_n = \frac{1}{n}$ für alle $n \in \mathbb{N}$. Es gilt $\lim_n x_n = \lim_n \frac{1}{n} = 0$. Wegen $x_n > 0$ gilt $\operatorname{sgn}(x_n) = 1$ für alle $n \in \mathbb{N}$. Daraus folgt

$$\lim_{n\to\infty} \operatorname{sgn}(x_n) \quad = \quad \lim_{n\to\infty} 1 \quad = \quad 1 \quad \neq \quad 0 \quad = \quad \operatorname{sgn}(0).$$

Also ist die Stetigkeitsbedingung im Punkt 0 nicht erfüllt. Sei nun $a \in \mathbb{R}^+$. Wir müssen zeigen, dass die Abbildung sgn in a stetig ist und betrachten dazu eine beliebige reelle Folge $(x_n)_{n\in\mathbb{N}}$ mit $\lim_n x_n = a$. Für $\varepsilon = a \in \mathbb{R}^+$ existiert auf Grund der Konvergenz ein $N \in \mathbb{N}$ mit $|x_n - a| < \varepsilon$ für alle $n \geq N$. Insbesondere gilt $x_n > a - \varepsilon = a - a = 0$ für alle $n \geq N$. Aus $x_n > 0$ folgt $\operatorname{sgn}(x_n) = 1$ für alle $n \geq N_1$ und somit $\lim_n \operatorname{sgn}(x_n) = 1 = \operatorname{sgn}(a)$.

Sei nun $a \in \mathbb{R}$ negativ und $(x_n)_{n\in\mathbb{N}}$ eine reelle Folge, die gegen a konvergiert. Für $\varepsilon = -a \in \mathbb{R}^+$ gibt es ein $N \in \mathbb{N}$ mit $|x_n - a| < \varepsilon$ für alle $n \geq N$. Dies bedeutet $x_n < a + \varepsilon = a + (-a) = 0$ und somit $\operatorname{sgn}(x_n) = -1$ für alle $n \geq N$. Es folgt $\lim_n \operatorname{sgn}(x_n) = -1 = \operatorname{sgn}(a)$. □

Ist $D \subseteq \mathbb{R}$ und $a \in D$ ein isolierter Punkt, so ist *jede* Funktion $f : D \to \mathbb{R}$ in a stetig. Denn in diesem Fall gibt es keine Folge $(x_n)_{n\in\mathbb{N}}$ in D_a, die gegen a konvergiert. Die Implikation $\lim_n x_n = a \Rightarrow \lim_n f(x_n) = f(a)$ ist damit für jede Folge in D erfüllt.

(5.9) Proposition Seien $D \subseteq \mathbb{R}$, $f, g : D \to \mathbb{R}$, und sei $a \in D$. Wenn die Funktionen f und g im Punkt a stetig sind, dann gilt dasselbe für $f + g$ und fg. Ist der Quotient $\frac{f}{g}$ auf D definiert, so ist auch $\frac{f}{g}$ in a stetig.

Beweis: Wir beschränken uns auf den Nachweis für die Funktion $f + g$. Ist $a \in D$ ein isolierter Punkt von D, dann ist $f + g$ in a auf jeden Fall stetig. Andernfalls ist a ein Berührpunkt der Menge $D_a = D \setminus \{a\}$. Setzen wir $\tilde{f} = f|_{D_a}$ und $\tilde{g} = g|_{D_a}$, dann folgt aus Proposition (5.8) und der Stetigkeit von f und g im Punkt a, dass

$$\lim_{x\to a} \tilde{f}(x) \quad = \quad f(a) \qquad \text{und} \qquad \lim_{x\to a} \tilde{g}(x) \quad = \quad g(a)$$

gilt. Durch Anwendung von Proposition (5.6) erhalten wir

$$\lim_{x\to a} (\tilde{f} + \tilde{g})(x) \quad = \quad \lim_{x\to a} \tilde{f}(x) + \lim_{x\to a} \tilde{g}(x) \quad = \quad f(a) + g(a) \quad = \quad (f + g)(a).$$

Wiederum auf Grund von Proposition (5.8) bedeutet dies, dass $f + g$ im Punkt a stetig ist. □

(5.10) Folgerung Alle Polynomfunktionen und rationalen Funktionen sind jeweils auf ihrem gesamten Definitionsbereich stetig.

Beweis: Dies folgt aus Proposition (5.9) und der Tatsache, dass alle rationalen Funktionen sich durch Addition, Multiplikation und Division aus der Identität id und den konstanten Funktionen zusammensetzen lassen. □

Reellwertige Funktionen lassen sich auch durch *Komposition* bereits gegebener Funktionen definieren. Die Komposition zweier Abbildungen hatten wir in §4 definiert. Sind $D, E \subseteq \mathbb{R}$ und $f : D \to \mathbb{R}$, $g : E \to \mathbb{R}$ Funktionen mit $f(D) \subseteq E$, dann ist auch $g \circ f$ eine reellwertige Funktion auf D. Ist zum Beispiel $f = \text{abs}$ und $g : \mathbb{R} \to \mathbb{R}$ gegeben durch $g(x) = x^3 + 3x + 5$ für alle $x \in \mathbb{R}$, dann sind die Kompositionen $g \circ \text{abs}$ und $\text{abs} \circ g$ gegeben durch

$$(g \circ \text{abs})(x) = g(|x|) = |x|^3 + 3|x| + 5 \qquad \text{und} \qquad (\text{abs} \circ g)(x) = |g(x)| = |x^3 + 3x + 5| \qquad \text{für alle } x \in \mathbb{R}.$$

(5.11) Proposition Seien $f : D \to \mathbb{R}$ und $g : E \to \mathbb{R}$ stetige Funktionen mit $f(D) \subseteq E$. Dann ist auch die Funktion $(g \circ f) : D \to \mathbb{R}$ stetig.

Beweis: Sei $a \in D$ und $(x_n)_{n\in\mathbb{N}}$ eine Folge in D mit $\lim_n x_n = a$. Wegen der Stetigkeit von f gilt dann $\lim_n f(x_n) = f(a)$. Nun ist durch $y_n = f(x_n)$ eine Folge in E mit $\lim_n y_n = f(a)$ gegeben. Die Stetigkeit von g im Punkt $f(a)$ liefert dann

$$\lim_{n\to\infty} (g \circ f)(x_n) \quad = \quad \lim_{n\to\infty} g(y_n) \quad = \quad g(f(a)) \quad = \quad (g \circ f)(a). \qquad \square$$

Also sind $x \mapsto |x|^3 + 3|x| + 5$ und $x \mapsto |x^3 + 3x + 5|$ beides stetige Funktionen.

Häufig verwendet man auch das folgende äquivalente Kriterium, um die Stetigkeit einer Funktion zu untersuchen.

(5.12) Satz *(ε-δ-Kriterium)*

Sei $f : D \to \mathbb{R}$ eine Funktion und $a \in D$. Genau dann ist f stetig in a, wenn für jedes $\varepsilon \in \mathbb{R}^+$ ein $\delta \in \mathbb{R}^+$ existiert, so dass die Implikation

$$|x - a| < \delta \qquad \Rightarrow \qquad |f(x) - f(a)| < \varepsilon \qquad \text{für alle} \quad x \in D \quad \text{erfüllt ist.}$$

Beweis: „⇐“ Nehmen wir an, dass ε-δ-Kriterium ist erfüllt. Sei $(x_n)_{n\in\mathbb{N}}$ eine Folge in D mit $\lim_n x_n = a$. Wir müssen zeigen, dass $\lim_n f(x_n) = f(a)$ erfüllt ist. Sei dazu $\varepsilon > 0$ vorgegeben. Dann gibt es auf Grund unseres Kriteriums ein $\delta > 0$, so dass aus $|x - a| < \delta$ stets $|f(x) - f(a)| < \varepsilon$ folgt, für alle $x \in D$. Wegen $\lim_n x_n = a$ gibt es ein $N \in \mathbb{N}$, so dass $|x_n - a| < \delta$ für alle $n \geq N$ erfüllt ist. Dann gilt aber auch $|f(x_n) - f(a)| < \varepsilon$ für alle $n \in \mathbb{N}$ mit $n \geq N$. Damit haben wir $\lim_n f(x_n) = f(a)$ bewiesen.

„⇒“ Angenommen, die Funktion f ist in a stetig, aber das ε-δ-Kriterium ist nicht erfüllt. Dann gibt es ein $\varepsilon > 0$ mit der Eigenschaft, dass für jedes $\delta > 0$ ein $x \in D$ mit $|x - a| < \delta$ und $|f(x) - f(a)| \geq \varepsilon$ existiert. Speziell finden wir für jedes $n \in \mathbb{N}$ ein $x_n \in D$ mit

$$|x_n - a| < \tfrac{1}{n} \qquad \text{und} \qquad |f(x_n) - f(a)| \geq \varepsilon.$$

Für die Folge $(x_n)_{n\in\mathbb{N}}$ gilt dann offenbar $\lim_n x_n = a$, wegen $|f(x_n) - f(a)| \geq \varepsilon$ für alle $n \in \mathbb{N}$ konvergiert die Folge $f(x_n)$ aber nicht gegen $f(a)$. Dies widerspricht unserer Annahme, dass f im Punkt a stetig ist. $\square$

Ein ähnliches Kriterium lässt sich auch für Funktionsgrenzwerte formulieren.

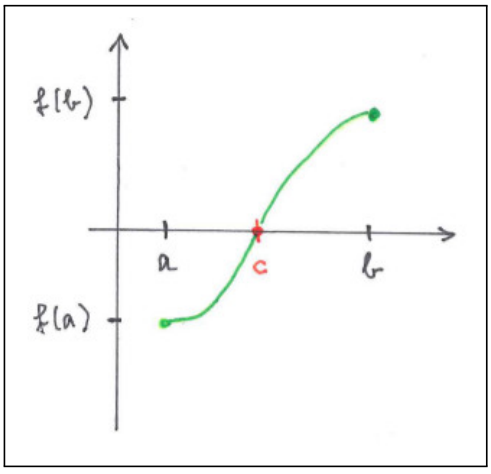
(5.13) Satz Seien $D \subseteq \mathbb{R}$, $f : D \to \mathbb{R}$ eine Funktion und $a, c \in \bar{\mathbb{R}}$, wobei a ein Berührpunkt von D sei. Unter diesen Voraussetzungen gilt $\lim\limits_{x \to a} f(x) = c$ genau dann, wenn für jede Umgebung V von c eine Umgebung U von a mit $f(U \cap D) \subseteq V$ existiert.

Beweis: „⇐“ Sei $(x_n)_{n \in \mathbb{N}}$ eine Folge in D mit $\lim_n x_n = a$ und V eine Umgebung von c. Wir müssen zeigen, dass ein $N \in \mathbb{N}$ mit $f(x_n) \in V$ für alle $n \geq N$ existiert, denn dann ist $\lim_n f(x_n) = c$ erfüllt. Nun gibt es nach Voraussetzung eine Umgebung U von a mit $f(U \cap D) \subseteq V$. Wegen $\lim_n x_n = a$ existiert ein $N \in \mathbb{N}$ mit $x_n \in U$ für alle $n \geq N$. Daraus folgt $f(x_n) \in f(U \cap D) \subseteq V$ für alle $n \geq N$.

„⇒“ Nehmen wir an, das $\lim_{x \to a} f(x) = c$ erfüllt ist, aber andererseits eine Umgebung V von c existiert mit der Eigenschaft, dass $f(U \cap D) \not\subseteq V$ für alle Umgebungen U von a gilt. Im Fall $a \in \mathbb{R}$ gibt es dann für jedes $n \in \mathbb{N}$ ein x_n in der Umgebung $U_n = \left]a - \frac{1}{n}, a + \frac{1}{n}\right[$ von a mit $f(x_n) \notin V$. Ist $a = +\infty$, dann existiert für jedes $n \in \mathbb{N}$ ein $x_n \in]n, +\infty[$ mit $f(x_n) \notin V$. Im Fall $a = -\infty$ finden für jedes $n \in \mathbb{N}$ ein $x_n \in]-\infty, -n[$ mit $f(x_n) \notin V$. In jedem der drei Fälle gilt dann $\lim_n x_n = a$, aber es existiert kein $N \in \mathbb{N}$ mit $f(x_n) \in V$ für alle $n \geq N$. Folglich konvergiert $(f(x_n))_{n \in \mathbb{N}}$ weder eigentlich noch uneigentlich gegen c. Dies steht zur Voraussetzung $\lim_{x \to a} f(x) = c$ im Widerspruch. □

Wir beweisen nun eine Reihe wichtiger Sätze über stetige Funktionen.

(5.14) Satz *(Zwischenwertsatz)*
Seien $a, b \in \mathbb{R}$ mit $a < b$ und $I = [a, b]$. Weiter sei $f : I \to \mathbb{R}$ eine stetige Funktion mit $f(a) < 0$ und $f(b) > 0$. Dann existiert ein $c \in I$ mit $f(c) = 0$. Dasselbe gilt, wenn $f(a) > 0$ und $f(b) < 0$ ist.



Beweis: Wir beschränken und auf den Fall $f(a) < 0$, $f(b) > 0$, da sich der Beweis im anderen Fall analog führen lässt. Um die Existenz einer Nullstelle c zu beweisen, konstruieren wir rekursiv eine Folge von Paaren (a_n, b_n), so dass folgende drei Bedingungen für alle $n \in \mathbb{N}_0$ erfüllt sind:

(i) Es gilt $[a_{n+1}, b_{n+1}] \subseteq [a_n, b_n]$.
(ii) Für die Länge des Intervalls $[a_n, b_n]$ gilt jeweils $b_n - a_n = 2^{-n}(b-a)$.
(iii) Es gilt jeweils $f(a_n) \leq 0$ und $f(b_n) \geq 0$.

Für $n = 0$ setzen wir $a_0 = a$ und $b_0 = b$. Nehmen wir nun an, das Paar (a_n, b_n) ist bereits konstruiert, und $m = \frac{1}{2}(a_n + b_n)$ ist der Mittelpunkt des Intervalls $[a_n, b_n]$. Ist $f(m) \geq 0$, dann setzen wir $a_{n+1} = a_n$ und $b_{n+1} = m$. Gilt dagegen $f(m) < 0$, dann sei $a_{n+1} = m$ und $b_{n+1} = b_n$. Man rechnet unmittelbar nach, dass das Paar (a_{n+1}, b_{n+1}) in beiden Fällen wieder die Bedingungen (i) bis (iii) erfüllt.

Die Folge $(a_n)_{n\in\mathbb{N}_0}$ ist nach Konstruktion monoton wachsend und nach oben beschränkt (mit b als oberer Schranke), also nach Satz (3.6) konvergent. Die Folge $(b_n)_{n\in\mathbb{N}_0}$ ist monoton fallend und nach unten beschränkt, deshalb existiert auch hier ein Grenzwert. Wegen

$$\lim_{n\to\infty} b_n - \lim_{n\to\infty} a_n \quad = \quad \lim_{n\to\infty} (b_n - a_n) \quad = \quad \lim_{n\to\infty} 2^{-n}(b-a) \quad = \quad 0$$

stimmen die beiden Grenzwerte überein, wir können also $c = \lim_n a_n = \lim_n b_n$ setzen. Da f stetig ist und $f(a_n) \leq 0$ für alle $n \in \mathbb{N}$ gilt, erhalten wir die Abschätzung

$$f(c) \quad = \quad \lim_{n\to\infty} f(a_n) \quad \leq \quad 0$$

nach Satz (2.11). Entsprechend liefert die Stetigkeit von f und die Ungleichung $f(b_n) \geq 0$ für alle $n \in \mathbb{N}$ die Abschätzung $f(c) \geq 0$. Insgesamt erhalten wir $f(c) = 0$, wir haben also eine Nullstelle c im Intervall I gefunden. □

(5.15) Folgerung Sei I wie oben und $f : I \to \mathbb{R}$ stetig und $d \in \mathbb{R}$ mit $f(a) < d < f(b)$. Dann gibt es ein $c \in I$ mit $f(c) = d$. Dasselbe gilt auch im Fall $f(a) > d > f(b)$.

Beweis: Wir wenden den Zwischenwertsatz auf die Funktion $g : I \to \mathbb{R}$ mit $g(x) = f(x) - d$ an und erhalten ein $c \in I$ mit $g(c) = 0$. Dieses c erfüllt dann die Gleichung $f(c) = d$. □

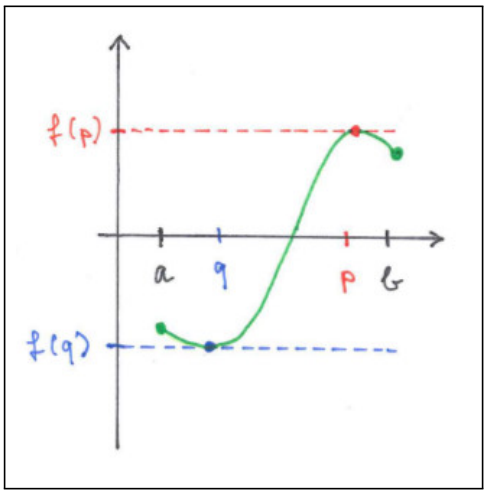
(5.16) Folgerung Ist $I \subseteq \mathbb{R}$ ein beliebiges Intervall, und ist $f : I \to \mathbb{R}$ eine stetige Funktion, dann ist auch $f(I)$ ein Intervall in $\mathbb{R}$.

Beweis: Seien $c, d \in f(I)$ mit $c < d$ und $q \in \mathbb{R}$ mit $c < q < d$ vorgegeben. Zu zeigen ist, dass q in $f(I)$ liegt; dies ist gleichbedeutend damit, dass ein $p \in I$ mit $f(p) = q$ existiert. Wegen $c, d \in f(I)$ gibt es $a, b \in I$ mit $f(a) = c$ und $f(b) = d$. Nach eventueller Vertauschung von a und b können wir davon ausgehen, dass $a < b$ gilt. Nach Folgerung (5.15) finden wir ein $p \in [a, b]$ mit $f(p) = q$. Weil es sich bei I um ein Intervall handelt, ist p in I enthalten. □

(5.17) Satz *(Maximumsprinzip)*
Seien $a, b \in \mathbb{R}$ mit $a < b$ und $I = [a, b]$. Sei $f : I \to \mathbb{R}$ stetig. Dann ist die Funktion f auf I ***beschränkt***, was bedeutet, dass die Wertemenge

$$M \quad = \quad \{f(x) \mid x \in I\} \quad \subseteq \quad \mathbb{R}$$

von f beschränkt ist. Außerdem nimmt die Funktion auf I ihr Maximum und ihr Minimum an, d.h. es gibt $p, q \in I$ mit $f(p) = \max M$ und $f(q) = \min M$.



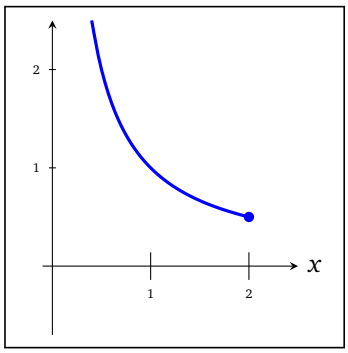
Maximumsprinzip: Der Wertebereich von f ist beschränkt.
Das Maximum wird im Punkt p, das Minimum in q angenommen.

Beweis: Wir beweisen nur, dass M nach oben beschränkt ist und das Maximum von f angenommen wird. Die entsprechenden Aussagen über das Minimum folgen dann durch Anwendung der bereits bewiesenen Aussagen auf die Funktion $g(x) = -f(x)$. Sei $s = \sup M$, wobei s in dem Fall, dass M nach oben unbeschränkt ist, den Wert $+\infty$ annimmt. Sowohl im Fall $s \in \mathbb{R}$ als auch im Fall $s = +\infty$ existiert eine Folge $(x_n)_{n\in\mathbb{N}}$ in I mit $\lim_n f(x_n) = s$. Ist nämlich $s \in \mathbb{R}$, dann existiert für jedes $n \in \mathbb{N}$ ein $x_n \in I$ mit $s - \frac{1}{n} < f(x_n) \leq s$; gäbe es ein solches x_n nicht, dann wäre bereits $s - \frac{1}{n}$ obere Schranke von M, im Widerspruch zur Definition des Supremums. Wegen $s - \frac{1}{n} < f(x_n) \leq s$, $\lim_n(s - \frac{1}{n}) = s$ und auf Grund des Sandwich-Lemmas, Satz (2.12), konvergiert dann $(f(x_n))_{n\in\mathbb{N}}$ gegen s.

Gilt nun statt dessen $s = +\infty$, dann existiert für jedes $n \in \mathbb{N}$ ein $x_n \in I$ mit $f(x_n) > n$, denn ansonsten wäre n eine obere Schranke von M. Es gilt dann $\lim_n f(x_n) = +\infty$. Ist nämlich $\kappa \in \mathbb{R}^+$ beliebig vorgegeben und $N \in \mathbb{N}$ mit $N \geq \kappa$, dann gilt für alle $n \geq N$ jeweils $f(x_n) > n \geq N \geq \kappa$.

Als Folge in I ist $(x_n)_{n\in\mathbb{N}}$ beschränkt. Auf Grund des Satzes (3.10) von Bolzano-Weierstrass besitzt sie einen Häufungspunkt. Es existiert somit eine konvergente Teilfolge $(x_{n_k})_{k\in\mathbb{N}}$, deren Grenwert c wegen $a \leq x_{n_k} \leq b$ für alle $k \in \mathbb{N}$ wiederum im Intervall I liegt. Auf Grund der Stetigkeit von f gilt $\lim_k f(x_{n_k}) = f(c)$, andererseits gilt wegen $\lim_n f(x_n) = s$ auch $\lim_k f(x_{n_k}) = s$. Wegen $f(c) \in \mathbb{R}$ ist der Fall $s = +\infty$ also ausgeschlossen. Dies zeigt, dass M nach oben beschränkt ist. Das Supremum s von M wird von f im Punkt c angenommen. Es gilt also $s = f(c) \in M$, und damit ist s auch das Maximum der Menge M. □

Für *offene* oder *halboffene* Intervalle ist die Aussage falsch, wie das Beispiel $f :]0,2] \to \mathbb{R}$, $x \mapsto \frac{1}{x}$ zeigt. Offenbar ist diese Funktion nach oben unbeschränkt.



(5.18) Definition Eine Funktion $f : D \to \mathbb{R}$ wird ***monoton wachsend*** genannt, wenn aus $a \leq b$ stets $f(a) \leq f(b)$ folgt. Von einer ***streng*** monoton wachsenden Funktion spricht man, wenn im Fall $a < b$ immer $f(a) < f(b)$ erfüllt ist. Entsprechend sind die Begriffe ***monoton fallend*** bzw. ***streng*** monoton fallend definiert.

Sei $f : D \to \mathbb{R}$ eine Funktion mit der Eigenschaft, dass die Menge D bijektiv auf $E = f(D)$ abgebildet wird. Dann nennen wir die Umkehrabbildung $g = f^{-1} : E \to D$ auch die ***Umkehrfunktion*** von f. Beispielsweise ist für ungerades $k \in \mathbb{N}$ die ***k-te Wurzelfunktion*** $r_k : \mathbb{R} \to \mathbb{R}$, $x \mapsto \sqrt[k]{x}$, gegeben durch die in § 1 definierte k-te Wurzel, die Umkehrfunktion von $p_k : \mathbb{R} \to \mathbb{R}$, $x \mapsto x^k$. Dies überprüft man anhand der Gleichungen

$$(r_k \circ p_k)(x) = r_k(x^k) = \sqrt[k]{x^k} = x \quad \text{und} \quad (p_k \circ r_k)(x) = p_k(\sqrt[k]{x}) = (\sqrt[k]{x})^k = x \quad \text{für alle} \quad x \in \mathbb{R}.$$

Eine entsprechende Aussage gilt auch für gerades $k \in \mathbb{N}$, nur dass hier die Definitionsbereiche $D = E = \mathbb{R}^+$ gewählt werden müssen.

(5.19) Satz *(Stetigkeit der Umkehrfunktion)*

Sei $I \subseteq \mathbb{R}$ ein Intervall, und sei $f : I \to \mathbb{R}$ stetig und streng monoton wachsend (bzw. fallend). Dann ist f eine Bijektion auf ihr Bild $f(I)$, und die Umkehrfunktion $f^{-1} : f(I) \to I$ ist ebenfalls stetig und streng monoton wachsend (bzw. fallend).

Beweis: Wir beschränken uns beim Beweis auf die streng monoton wachsenden Funktionen, weil der Beweis für die streng monoton fallenden vollkommen analog verläuft. Die Abbildung f ist injektiv, denn sind $a, b \in I$ mit $a \neq b$, dann gilt $f(a) < f(b)$ im Fall $a < b$ und $f(a) > f(b)$ im Fall $a > b$, auf jeden Fall also $f(a) \neq f(b)$. Als Abbildung $I \to f(I)$ ist f außerdem surjektiv, insgesamt also eine Bijektion zwischen I und $f(I)$. Sei $f^{-1} : f(I) \to I$ die Umkehrabbildung. Mit f ist auch f^{-1} streng monoton wachsend. Andernfalls gäbe es $c, d \in f(I)$ mit $c < d$ und $f^{-1}(c) \geq f^{-1}(d)$. Im Fall $f^{-1}(c) = f^{-1}(d)$ würde $c = f(f^{-1}(c)) = f(f^{-1}(d)) = d$ folgen, im Fall $f^{-1}(c) > f^{-1}(d)$ würden wir $c > d$ erhalten, in beiden Fällen also also $c \geq d$ im Widerspruch zur vorherigen Feststellung. Nach Folgerung (5.16) ist mit I auch $J = f(I)$ ein Intervall.

Beweisen wir nun die Stetigkeit von f^{-1}. Sei $q \in J$ und $(y_n)_{n\in\mathbb{N}}$ eine Folge in J mit $\lim_n y_n = q$. Weiter sei $p = f^{-1}(q)$ und $x_n = f^{-1}(y_n)$ für jedes $n \in \mathbb{N}$. Wir müssen zeigen, dass $\lim_n x_n = p$ gilt. Angenommen, dies ist nicht der Fall. Dann gibt es ein $\varepsilon \in \mathbb{R}^+$, so dass $|x_n - p| \geq \varepsilon$ für unendlich viele $n \in \mathbb{N}$ gilt. Betrachten wir die Mengen

$$M_+ = \{n \in \mathbb{N} \mid x_n \geq p + \varepsilon\} \qquad \text{und} \qquad M_- = \{n \in \mathbb{N} \mid x_n \leq p - \varepsilon\} \quad ,$$

dann ist M_+ oder M_- unendlich. Gehen wir davon aus, dass M_+ unendlich ist; im anderen Fall läuft der Beweis analog. Auf Grund der strengen Monotonie von f gilt für alle $n \in M_+$ jeweils

$$y_n \quad = \quad f(x_n) \quad \geq \quad f(p+\varepsilon) \quad > \quad f(p) \quad = \quad q.$$

Setzen wir $\varepsilon_1 = |f(p+\varepsilon) - q|$, dann gilt also $|y_n - q| > \varepsilon_1$ für unendlich viele $n \in \mathbb{N}$. Aber dies widerspricht der Voraussetzung $\lim_n y_n = q$. □

§ 6. *Potenzreihen, Exponentialfunktion, Sinus und Kosinus*

Inhaltsübersicht

Genau wie reelle oder komplexe Zahlen lassen sich auch Funktionen zu Folgen $(f_n)_{n\in\mathbb{N}}$ zusammenfassen und deren Konvergenz untersuchen. Für Funktionen existieren allerdings zwei unterschiedlich starke Konvergenzbegriffe. Für uns ist vor allem die ***gleichmäßige*** Konvergenz interessant, denn sie gewährleistet, dass bei stetigen Funktionen f_n auch die Grenzfunktion f stetig ist. Intuitiv bedeutet gleichmäßige Konvergenz, dass für jeden Punkt x des gemeinsamen Definitionsbereichs D der Funktionen f_n die Folgen $(f_n(x))_{n\in\mathbb{N}}$ „ungefähr gleich schnell" gegen ihren jeweiligen Grenzwert konvergieren.

Wir wenden das Konzept der gleichmäßigen Konvergenz dann auf einen bestimmten Typ von Funktionenfolgen an, die ***Potenzreihen***; hierunter kann man sich „unendliche Polynomausdrücke" vorstellen. Mit Hilfe dieser Potenzreihen definieren wir die Exponentialfunktion sowie die trigonometrischen Funktionen (Sinus und Kosinus). Wir leiten eine Reihe wichtiger Eigenschaften dieser Funktionen her. Außerdem befassen wir uns mit den Umkehrfunktionen dieser Funktionen, dem Logarithmus sowie den Arkusfunktionen (Arcus sinus, Arcus cosinus und Arcus tangens). Wie man darauf kommt, diese zum Teil bereits aus der Schule bekannten Funktionen als Potenzreihen darzustellen, werden wir erst im nächsten Kapitel sehen.

Wichtige Begriffe und Sätze

- punktweise und gleichmäßige Konvergenz von Funktionenfolgen
- Potenzreihen als Beispiele von Funktionenfolgen
- Die Grenzfunktion einer Folge stetiger Funktionen ist bei gleichmäßiger Konvergenz auf dem Definitionsbereich der Funktionenfolge wieder stetig. (Bei nur punktweiser Konvergenz ist dies im Allgemeinen falsch.)
- Definition des Konvergenzradius einer Potenzreihe, Berechnung durch den Limes superior
- Exponential-, Sinus- und Kosinusreihe; Grenzfunktionen sind die jeweils auf ganz $\mathbb{R}$ definierte Exponential-, Sinus- und Kosinusfunktion
- Exponentialfunktion und Logarithmus zu einer beliebigen Basis $b > 1$
- Tangensfunktion, Definition der Arkusfunktion als Umkehrfunktionen der trigonometrischen Funktionen

In diesem Abschnitt werden wir zunächst wieder mit der Bezeichnung $\mathbb{K}$ arbeiten, die für einen der beiden Körper $\mathbb{R}$ oder $\mathbb{C}$ steht.

(6.1) Definition Sei $D \subseteq \mathbb{K}$ eine beliebige Teilmenge und $(f_n)_{n\in\mathbb{N}}$ eine Folge $\mathbb{K}$-wertiger Funktionen $f_n : D \to \mathbb{K}$ auf D. Sei $f : D \to \mathbb{K}$ eine weitere Funktion. Man sagt, die Folge $(f_n)_{n\in\mathbb{N}}$ konvergiert

(i) ***punktweise*** gegen f, wenn für jedes $x \in D$ die Folge $(f_n(x))_{n\in\mathbb{N}}$ in $\mathbb{K}$ gegen den Wert $f(x)$ konvergiert

(ii) ***gleichmäßig*** gegen f, wenn für jedes $\varepsilon \in \mathbb{R}^+$ ein $N \in \mathbb{N}$ existiert, so dass $|f_n(x) - f(x)| < \varepsilon$ für alle $n \geq N$ und $x \in D$ erfüllt ist.

Offenbar folgt aus der gleichmäßigen Konvergenz auf D auch die punktweise Konvergenz auf D. In beiden Fällen nennt man f auch die ***Grenzfunktion*** der Funktionenfolge $(f_n)_{n\in\mathbb{N}}$.

(6.2) Definition Sei $(f_n)_{n\in\mathbb{N}}$ eine Folge $\mathbb{K}$-wertiger Funktionen auf einer Teilmenge $D \subseteq \mathbb{K}$. Dann bezeichnet man die Teilmenge

$$D_{\text{conv}} \quad = \quad \{x \in D \mid (f_n(x))_{n\in\mathbb{N}} \text{ konvergiert } \}$$

als ***Konvergenzbereich*** der Funktionenfolge.

Um den Unterschied zwischen beiden Konvergenzarten zu verdeutlichen, betrachten wir die Folge $(f_n)_{n\in\mathbb{N}}$ von Funktionen auf $[0,1]$ gegeben durch $f_n(x) = x^n$ für alle $x \in [0,1]$ und $n \in \mathbb{N}$. Sei $f : [0,1] \to \mathbb{R}$ gegeben durch

$$f(x) \quad = \quad \begin{cases} 0 & \text{falls } 0 \le x < 1 \\ 1 & \text{falls } x = 1. \end{cases}$$

Dann konvergiert die Folge $(f_n)_{n\in\mathbb{N}}$ zwar punktweise gegen f, denn nach Proposition (2.4) (i) konvergiert die Folge $(x^n)_{n\in\mathbb{N}}$ für $0 \le x < 1$ gegen 0, und die konstante Folge $(1^n)_{n\in\mathbb{N}}$ konvergiert gegen 1. Es handelt sich aber um keine gleichmäßige Konvergenz. Sei beispielsweise $\varepsilon = \frac{1}{2}$, und $x_n = \frac{1}{\sqrt[n]{2}}$ für jedes $n \in \mathbb{N}$. Dann gilt jeweils $|f_n(x_n) - f(x_n)| = |\frac{1}{2} - 0| = \varepsilon$. Es existiert also kein $N \in \mathbb{N}$, so dass $|f_n(x) - f(x)| < \varepsilon$ für alle $x \in [0,1]$ und $n \ge N$ erfüllt ist.

Für uns ist die gleichmäßige Konvergenz von Funktionenfolgen vor allem auf Grund der folgenden Eigenschaft von Interesse.

(6.3) Satz Sei $(f_n)_{n\in\mathbb{N}}$ eine Folge stetiger reellwertiger Funktionen mit Definitionsbereich $D \subseteq \mathbb{R}$, die auf D gleichmäßig gegen eine weitere Funktion $f : D \to \mathbb{R}$ konvergiert. Dann ist auch f eine stetige Funktion auf D.

Beweis: Seien $a \in D$ und $\varepsilon \in \mathbb{R}^+$ beliebig vorgegeben. Nach dem ε-δ-Kriterium, Satz (5.12), genügt es zu zeigen, dass ein $\delta \in \mathbb{R}^+$ existiert, so dass die Implikation

$$|x - a| < \delta \quad \Rightarrow \quad |f(x) - f(a)| < \varepsilon \qquad \text{für alle} \quad x \in D \quad \text{erfüllt ist.}$$

Auf Grund der gleichmäßigen Konvergenz existiert ein $N \in \mathbb{N}$, so dass $|f_n(x) - f(x)| < \frac{1}{3}\varepsilon$ für alle $x \in D$ und $n \ge N$ gilt. Außerdem gibt es auf Grund der Stetigkeit von f_N ein $\delta \in \mathbb{R}^+$, so dass die Implikation $|x - a| < \delta \Rightarrow |f_N(x) - f_N(a)| < \frac{1}{3}\varepsilon$ für alle $x \in D$ gültig ist. Ist nun $x \in D$ ein Punkt mit $|x - a| < \delta$, dann erhalten wir

$$\begin{aligned} |f(x) - f(a)| \quad &\le \quad |f(x) - f_N(x)| + |f_N(x) - f_N(a)| + |f_N(a) - f(a)| \\ &< \quad \tfrac{1}{3}\varepsilon + \tfrac{1}{3}\varepsilon + \tfrac{1}{3}\varepsilon \quad = \quad \varepsilon \qquad \text{wie gewünscht.} \end{aligned}$$

□

Der Satz bleibt gültig, wenn man für D eine Teilmenge von $\mathbb{K}$ und statt reellwertigen $\mathbb{K}$-wertige Funktionen betrachtet. Auch beim Beweis ergeben sich keinerlei Änderungen, wenn man unter $|\cdot|$ den Absolutbetrag auf $\mathbb{K}$ (also entweder den reellen oder komplexen Absolutbetrag) versteht. Allerdings haben wir in dieser Vorlesung die Stetigkeit nur für reellwertige Funktionen auf Teilmengen von $\mathbb{R}$ definiert.

Im weitere Verlauf der Vorlesung wird der folgende spezieller Typ von Funktionenfolgen für uns eine besondere Rolle spielen.

(6.4) Definition Sei $(a_n)_{n\in\mathbb{N}_0}$ eine Folge in $\mathbb{K}$ und $a \in \mathbb{K}$. Dann nennt man die Funktionenfolge $(f_n)_{n\in\mathbb{N}_0}$ gegeben durch

$$f_n(x) \quad = \quad \sum_{k=0}^{n} a_k (x-a)^k$$

für alle $x \in \mathbb{R}$ die ***Potenzreihe in Entwicklungspunkt*** a mit der Koeffizientenfolge $(a_n)_{n\in\mathbb{N}_0}$.

Ähnlich wie bei den Reihen in $\mathbb{K}$ bezeichnet der Ausdruck

$$\sum_{n=0}^{\infty} a_n (x-a)^n$$

sowohl die in Definition (6.4) angegebene Funktionenfolge als auch die Grenzfunktion $f : D_{\text{conv}} \to \mathbb{K}$. Bei der Auswertung der Potenzreihen ist darauf zu achten, dass $(x-a)^0$ immer gleich 1 gesetzt wird, unabhängig davon, ob $x = a$ oder $x \neq a$ gilt.

Ein einfaches Beispiel für eine Potenzreihe im Entwicklungspunkt 0 ist durch $\sum_{n=0}^{\infty} x^n$ gegeben. Weitere wichtige Beispiele für Potenzreihen mit dem gleichen Entwicklungspunkt sind die ***Exponentialreihe***

$$\exp(x) \quad = \quad \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

sowie die ***Sinus-*** und ***Kosinusreihe***

$$\sin(x) \quad = \quad \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!} \qquad \text{und} \qquad \cos(x) \quad = \quad \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!}.$$

Im folgenden werden wir nun den Konvergenzbereich einer Potenzreihe genauer untersuchen.

(6.5) Definition Sei $f(x) = \sum_{n=0}^{\infty} a_n (x-a)^n$ eine Potenzreihe und $D_{\text{conv}} \subseteq \mathbb{K}$ ihr Konvergenzbereich. Dann nennt man

$$\rho(f) \quad = \quad \sup\{|x-a| \mid x \in D_{\text{conv}}\}$$

den ***Konvergenzradius*** von $f(x)$. Nach Definition ist $\rho(f) \in \mathbb{R}_+$ oder $\rho(f) = +\infty$ möglich.

(6.6) Proposition Sei $f(x) = \sum_{n=0}^{\infty} a_n(x-a)^n$ Potenzreihe, $\rho = \rho(f)$ ihr Konvergenzradius und $x \in \mathbb{K}$.

(i) Ist $|x-a| < \rho$, dann konvergiert f in x absolut.
(ii) Ist $|x-a| > \rho$, dann divergiert die Potenzreihe.

Für Punkte $x \in K$ mit $|x-a| = \rho$ ist keine allgemeine Aussage möglich. Im Fall $\rho = 0$ gilt $D_{\text{conv}} = \{a\}$, im Fall $\rho = +\infty$ ist $D_{\text{conv}} = \mathbb{K}$.

Beweis: zu (i) Nach Definition des Konvergenzradius gibt es ein $y \in \mathbb{K}$, so dass $|x-a| < |y-a| < \rho$ gilt und die Reihe $\sum_{n=0}^{\infty} a_n(y-a)^n$ konvergiert. Nach Folgerung (4.10) gilt dann $\lim_n a_n(y-a)^n = 0$. Setzen wir $\theta = \frac{|x-a|}{|y-a|} < 1$, dann folgt $|a_n(x-a)^n| = |a_n(y-a)|^n\theta^n \leq \theta^n$ für alle bis auf endlich viele $n \in \mathbb{N}_0$. Weil die geometrische Reihe $\sum_{n=0}^{\infty} \theta^n$ nach Proposition (4.6) konvergiert, folgt die absolute Konvergenz der Reihe $\sum_{n=0}^{\infty} a_n(x-a)^n$ aus dem Majoranktenkriterium, Satz (4.16).

zu (ii) Nach Definition des Konvergenzradius ist ρ insbesondere eine obere Schranke der Werte $|x-a|$ für alle $x \in \mathbb{K}$ mit der Eigenschaft, dass $\sum_{n=0}^{\infty} a_n(x-a)^n$ konvergiert. Gilt also $|x-a| > \rho$, dann ist $\sum_{n=0}^{\infty} a_n(x-a)^n$ eine divergente Reihe. □

Ist der Konvergenzradius ρ einer Potenzreihe f endlich und positiv, dann gilt im Fall $\mathbb{K} = \mathbb{R}$ für den Konvergenzbereich D_{conv} von f also $]a-\rho, a+\rho[\subseteq D_{\text{conv}} \subseteq [a-\rho, a+\rho]$. Im Fall $K = \mathbb{C}$ enthält D_{conv} die offene Kreisscheibe vom Radius ρ um den Punkt a, und ist in der abgeschlossenen Kreisscheibe um a vom selben Radius enthalten. Man kann für den Konvergenzradius auch eine explizite Formel angeben.

(6.7) Satz *(Satz von Cauchy-Hadamard)*
Der Konvergenzradius ρ einer beliebigen Potenzreihe $f(x) = \sum_{n=0}^{\infty} a_n(x-a)^n$ über $\mathbb{K}$ ist gegeben durch

$$\rho \quad = \quad \left(\limsup_n \sqrt[n]{|a_n|}\right)^{-1}.$$

Dies soll im Fall $\limsup_n \sqrt[n]{|a_n|} = 0$ bedeuten, dass $\rho = +\infty$, und im Fall $\limsup_n \sqrt[n]{|a_n|} = +\infty$, dass $\rho = 0$ gilt.

Beweis: Sei $s = \limsup_n \sqrt[n]{|a_n|}$, und betrachten wir zunächst den Fall $s \in \mathbb{R}^+$. Zu zeigen ist $\rho = s^{-1}$. Sei dazu $\varepsilon \in \mathbb{R}^+$ mit $\varepsilon < s$ vorgegeben. Wir zeigen: Ist $x \in \mathbb{K}$ mit $|x-a| < (s+\varepsilon)^{-1}$, dann konvergiert die Potenzreihe in x (sogar absolut). Gilt dagegen $|x+a| > (s-\varepsilon)^{-1}$, dann divergiert die Potenzreihe in x. Aus diesen beiden Aussagen folgt dann die Gleichung $\rho = s^{-1}$. Denn die erste Aussage in Verbindung mit Proposition (6.6) (ii) liefert $\rho \geq (s+\varepsilon)^{-1}$, und weil $\varepsilon \in \mathbb{R}^+$ beliebig klein gewählt werden kann, folgt $\rho \geq s^{-1}$. Ebenso kann aus der zweiten Aussage in Verbindung mit Teil (i) von Proposition (6.6) abgeleitet werden, dass $\rho \leq s^{-1}$ gilt. Insgesamt erhalten wir $\rho = s^{-1}$, wie gewünscht.

Für den Nachweis der ersten Aussage sei nun $x \in \mathbb{K}$ mit $|x-a| < (s+\varepsilon)^{-1}$. Nach Definition von s als Limes superior gilt $\sqrt[n]{|a_n|} < s+\varepsilon$ für alle bis auf endlich viele $n \in \mathbb{N}$; denn gäbe es in $(\sqrt[n]{|a_n|})_{n\in\mathbb{N}_0}$ eine (beschränkte) Teilfolge, deren Häufungspunkte alle größer gleich $s+\varepsilon$ wären. Definieren wir $\theta = |x-a|(s+\varepsilon)$, dann ist $\theta < 1$ und $\sqrt[n]{|a_n(x-a)^n|} = |x-a|\sqrt[n]{|a_n|} \leq \theta$ für alle bis auf endlich viele $n \in \mathbb{N}_0$. Nach dem Wurzelkriterium, Satz (4.20), ist f in x also absolut konvergent.

Für den Nachweis der zweiten Aussage sei $x \in \mathbb{K}$ mit $|x-a| > (s-\varepsilon)^{-1}$. Nehmen wir an, dass die Potenzreihe in x konvergiert. Dann muss nach Folgerung (4.10) für die Glieder der Reihe $\lim_n |a_n(x-a)^n| = 0$ gelten. Weil s aber Häufungspunkt der Folge $(\sqrt[n]{|a_n|})_{n\in\mathbb{N}_0}$ ist, gibt es unendlich viele $n \in \mathbb{N}_0$ mit $\sqrt[n]{|a_n|} > s-\varepsilon$, und daraus folgt $\sqrt[n]{|a_n(x-a)^n|} = |x-a|\sqrt[n]{|a_n|} > 1$ und somit auch $|a_n(x-a)^n| > 1$ für unendlich viele $n \in \mathbb{N}_0$. Also divergiert die Potenzreihe in x.

Zum Schluss betrachten wir die Fälle $s = 0$ und $s = +\infty$. Im Fall $s = 0$ ist die erste Aussage von oben für *jedes* $s' \in \mathbb{R}^+$ an Stelle von s und für jedes $\varepsilon \in \mathbb{R}^+$ erfüllt. Daraus folgt $\rho \geq (s')^{-1}$ für jedes $s' \in \mathbb{R}^+$ und somit $\rho = +\infty$. Im Fall $s = +\infty$ ist die zweite Aussage von oben für jedes $s' \in \mathbb{R}^+$ an Stelle von s und für jedes $\varepsilon \in \mathbb{R}^+$ mit $\varepsilon < s'$ erfüllt. Daraus folgt $\rho \leq (s')^{-1}$ für jedes $s' \in \mathbb{R}^+$, und entsprechend $\rho = 0$. □

(6.8) Folgerung Seien die Bezeichnungen wie in Satz (6.6) und $r \in \mathbb{R}^+$ mit $r < \rho$, wobei wir uns aber auf den Fall $\mathbb{K} = \mathbb{R}$ beschränken. Dann konvergiert die Potenzreihe auf $[a-r, a+r]$ gleichmäßig gegen ihren Grenzwert und definiert dort eine stetige Funktion.

Beweis: Sei $\varepsilon \in \mathbb{R}^+$ vorgegeben. Nach Satz (6.6) konvergiert die Reihe $\sum_{n=0}^{\infty} a_n(x-a)^n$ im Punkt $x = a+r$ absolut, somit besitzt die Reihe $\sum_{n=0}^{\infty} |a_n| r^n$ einen endlichen Grenzwert. Es gibt also ein $N \in \mathbb{N}$ mit $\sum_{k=n}^{\infty} |a_k| r^k < \varepsilon$ für alle $n \geq N$. Ist nun $x \in [a-r, a+r]$ beliebig vorgegeben, dann folgt $|x-a| \leq r$, und wir erhalten

$$\left|\sum_{k=n}^{\infty} a_k(x-a)^k\right| \quad < \quad \sum_{k=n}^{\infty} |a_k||x-a|^k \quad \leq \quad \sum_{k=n}^{\infty} |a_k| r^k \quad < \quad \varepsilon$$

für alle $n \geq N$. Damit ist die gleichmäßige Konvergenz von $\sum_{n=0}^{\infty} a_n(x-a)^n$ auf $D = [a-r, a+r]$ bewiesen, und die Stetigkeit der Funktion $f : D \to \mathbb{R}$, $x \mapsto \sum_{n=0}^{\infty} a_n(x-a)^n$ folgt aus Satz (6.3). □

Wir wenden die bisherigen Ergebnisse auf eine Reihe von Beispielen an.

(i) Die Potenzreihe $\sum_{n=0}^{\infty} x^n$ zum Entwicklungspunkt $a = 0$ und mit der konstanten Koeffizientenfolge $(a_n)_{n\in\mathbb{N}}$ gegeben durch $a_n = 1$ für alle $n \in \mathbb{N}$ hat wegen $\limsup_n \sqrt[n]{|1|} = 1$ den Konvergenzradius $\rho = 1$. Für jedes $x \in \mathbb{R}$ mit $|x| < 1$ gilt

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

nach Proposition (4.6) über den Wert der geometrischen Reihe. Nach Folgerung (6.8) definiert die Reihe für jedes $r \in \mathbb{R}^+$ mit $r < 1$ auf $[-r, r]$ eine stetige Funktion, insgesamt also auf dem offenen Intervall $]-1, 1[$. (Natürlich wissen wir auch so bereits, dass $x \mapsto \frac{1}{1-x}$ als rationale Funktion auf dem Intervall $]-1, 1[$ stetig ist.)

(ii) Die Exponentialreihe $\exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!}$ hat den Konvergenzradius $\rho = +\infty$. Für beliebiges $x \in \mathbb{R}$, $x \neq 0$ und jedes $n \in \mathbb{N}$ gilt nämlich

$$\left|\frac{x^{n+1}/(n+1)!}{x^n/n!}\right| \quad = \quad \left|\frac{x^{n+1}}{(n+1)!}\frac{n!}{x^n}\right| \quad = \quad \left|\frac{x}{n+1}\right| \quad ,$$

und dieser Quotient ist $\leq \frac{1}{2}$ für hinreichend großes n. Also konvergiert die Reihe in x nach dem Quotientenkriterium, Satz (4.18). Nach Satz (6.6) kann der Konvergenzradius ρ keinen endlichen Wert haben, weil sonst die Reihe für ein $x \in \mathbb{R}$ divergieren müsste. Nach Folgerung (6.8) definiert $\exp(x)$ für jedes $r \in \mathbb{R}^+$ auf $[-r, r]$ eine stetige Funktion. Insgesamt erhalten wir eine stetige Funktion auf ganz $\mathbb{R}$ die wir als ***Exponentialfunktion*** bezeichnen.

(iii) Durch eine ähnliche Rechnung zeigt man, dass auch Sinus- und Kosinusreihe einen unendlichen Konvergenzradius haben und somit auf ganz $\mathbb{R}$ stetige Funktionen definieren. Die entsprechenden Funktionen werden als ***Sinus-*** bzw. ***Kosinusfunktion*** bezeichnet.

(6.9) Proposition Es gilt $\exp(1) = e$, wobei e die Eulersche Zahl bezeichnet, die in Proposition (3.15) definiert wurde.

Beweis: Seien die Folgen $(a_n)_{n\in\mathbb{N}}$ und $(s_n)_{n\in\mathbb{N}}$ gegeben durch $a_n = (1+\frac{1}{n})^n$ und $s_n = \sum_{k=0}^{n} \frac{x^n}{n!}$. Dann gilt $\lim_n a_n = e$ und $\lim_n s_n = \exp(1)$. Für alle $n \in \mathbb{N}$ gilt auf Grund des Binomischen Lehrsatzes die Abschätzung

$$a_n \quad = \quad (1+\tfrac{1}{n})^n \quad = \quad \sum_{k=0}^{n} \binom{n}{k} 1^{n-k} (\tfrac{1}{n})^k \quad =$$

$$\sum_{k=0}^{n} \frac{n!}{k!(n-k)!} \cdot \frac{1}{n^k} \quad = \quad \sum_{k=0}^{n} \frac{1}{k!} \frac{n\cdot(n-1)\cdot\ldots\cdot(n-k+1)}{n^k} \quad \leq \quad \sum_{k=0}^{n} \frac{1}{k!} \quad = \quad s_n.$$

Daraus folgt $e = \lim_n a_n \leq \lim_n s_n = \exp(1)$. Für die umgekehrte Abschätzung bemerken wir, dass für $m, n \in \mathbb{N}$ mit $m > n \geq 2$ gilt

$$a_m \quad = \quad (1+\tfrac{1}{m})^m \quad = \quad \sum_{k=0}^{m} \frac{1}{k!} \frac{m\cdot(m-1)\cdot\ldots\cdot(m-k+1)}{m^k} \quad =$$

$$1+1+\sum_{k=2}^{m} \frac{1}{k!} 1\cdot\frac{m-1}{m}\cdot\frac{m-2}{m}\cdot\ldots\cdot\frac{m-k+1}{m} \quad \geq \quad 1+1+\sum_{k=2}^{n} \frac{1}{k!} 1\cdot\frac{m-1}{m}\cdot\frac{m-2}{m}\cdot\ldots\cdot\frac{m-k+1}{m}.$$

Für jedes $n \geq 2$ gilt also

$$e \quad = \quad \lim_{m\to\infty} a_m \quad \geq \quad 1+1+\lim_{m\to\infty}\left(\sum_{k=2}^{n} \frac{1}{k!} 1\cdot\frac{m-1}{m}\cdot\frac{m-2}{m}\cdot\ldots\cdot\frac{m-k+1}{m}\right)$$

$$= \quad 1+1+\sum_{k=2}^{n} \frac{1}{k!}\left(\lim_{m\to\infty} 1\cdot\frac{m-1}{m}\cdot\frac{m-2}{m}\cdot\ldots\cdot\frac{m-k+1}{m}\right) \quad = \quad 1+1+\sum_{k=2}^{n} \frac{1}{k!} \quad = \quad s_n.$$

Aus $e \geq s_n$ für alle $n \geq 2$ wiederum folgt $e \geq \lim_n s_n = \exp(1)$, insgesamt also $e = \exp(1)$. □

Wir werden im folgenden Kapitel mit Hilfe der Differentiation noch einen etwas kürzeren Beweis für diese Identität angeben. Übrigens kann man sich mit dem Taschenrechner leicht davon überzeugen, dass die Reihe $\sum_{n=0}^{\infty} \frac{x^n}{n!}$ deutlich schneller gegen e konvergiert als die Folge $(a_n)_{n\in\mathbb{N}}$ gegeben durch $a_n = (1+\frac{1}{n})^n$.

(6.10) Satz Die Exponentialfunktion besitzt die folgenden Eigenschaften.

(i) Es gilt $\exp(0) = 1$, $\exp(x+y) = \exp(x)\exp(y)$, $\exp(-x) = \exp(x)^{-1}$ und $\exp(x) > 0$ für alle $x, y \in \mathbb{R}$.

(ii) $\lim\limits_{x\to-\infty} \exp(x) = 0$ und $\lim\limits_{x\to+\infty} \exp(x) = +\infty$

(iii) Die Exponentialfunktion ist streng monoton wachsend und definiert eine bijektive Abbildung zwischen $\mathbb{R}$ und $\mathbb{R}^+$.

Beweis: zu (i) Die erste Gleichung erhält man durch einfaches Einsetzen von 0 in die Exponentialreihe $\sum_{n=0}^{\infty} \frac{x^n}{n!}$. Alle Terme mit Ausnahme des ersten verschwinden, also ist $\exp(0) = 1$. Seien nun $x, y \in \mathbb{R}$ vorgegeben. Für den Beweis der Gleichung $\exp(x+y) = \exp(x)\exp(y)$ wenden wir Satz (4.25) über Cauchyprodukte an; weil die Reihen $\exp(x)$ und $\exp(y)$ absolut konvergieren, sind die Voraussetzungen des Satzes erfüllt. Seien die Folgen $(a_n)_{n\in\mathbb{N}_0}$ und $(b_n)_{n\in\mathbb{N}_0}$ so definiert, dass die Gleichungen

$$\exp(x) = \sum_{n=0}^{\infty} a_n \quad \text{und} \quad \exp(y) = \sum_{n=0}^{\infty} b_n$$

gelten, nämlich durch $a_n = \frac{x^n}{n!}$ und $b_n = \frac{y^n}{n!}$. Sei $\sum_{n=0}^{\infty} c_n$ das Cauchyprodukt der Reihen $\sum_{n=0}^{\infty} a_n$ und $\sum_{n=0}^{\infty} b_n$. Auf Grund des Binomischen Lehrsatzes $(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k$, den wir in den Übungen behandelt haben, erhalten wir

$$c_n = \sum_{k=1}^{n} a_{n-k} b_k = \sum_{k=0}^{n} \frac{x^{n-k}}{(n-k)!} \cdot \frac{y^k}{k!} = \sum_{k=0}^{n} \frac{1}{n!} \binom{n}{k} x^{n-k} y^k = \frac{1}{n!} \sum_{k=0}^{n} \binom{n}{k} x^{n-k} y^k = \frac{(x+y)^n}{n!} ,$$

und mit dem Satz über Cauchyprodukte erhalten wir

$$\exp(x)\exp(y) = \left(\sum_{n=0}^{\infty} a_n\right)\left(\sum_{n=0}^{\infty} b_n\right) = \sum_{n=0}^{\infty} c_n = \sum_{k=0}^{n} \frac{(x+y)^n}{n!} = \exp(x+y) ,$$

wie gewünscht. Die Gleichung $\exp(-x) = \exp(x)^{-1}$ für alle $x \in \mathbb{R}$ ergibt sich nun unmittelbar aus der Rechnung $\exp(x)\exp(-x) = \exp(x+(-x)) = \exp(0) = 1$. Schließlich gilt für jedes $x \in \mathbb{R}$ auch $\exp(x) > 0$. Im Fall $x > 0$ sieht man dies direkt, weil alle Terme der Exponentialreihe positiv sind. Für $x < 0$ erhält man die Positivität durch die Gleichung $\exp(-x) = \exp(x)^{-1}$, denn der Kehrwert einer positiven Zahl ist positiv; schließlich ist auch $\exp(0) = 1$ positiv.

zu (ii) Zum Nachweis von $\lim_{x\to+\infty}\exp(x)=+\infty$ sei $(x_n)_{n\in\mathbb{N}}$ eine Folge in $\mathbb{R}$ mit $\lim_n x_n=+\infty$. Zu zeigen ist $\lim_n\exp(x_n)=+\infty$. Sei dazu $\kappa\in\mathbb{R}^+$ vorgeben. Definieren wir $r_2(x)=\sum_{n=2}^{\infty}\frac{x^n}{n!}$ für $x\in\mathbb{R}$, dann gilt jeweils $\exp(x)=1+x+r_2(x)$, und $r_2(x)\geq 0$ falls $x>0$. Wegen $\lim_n x_n=+\infty$ gibt es nun ein $N\in\mathbb{N}$ mit $x_n>\kappa$ für alle $n\geq N$. Es folgt $\exp(x_n)=1+x_n+r_2(x_n)\geq 1+x_n>1+\kappa>\kappa$ für alle $n\geq N$. Damit ist der Beweis abgeschlossen.

Für den Nachweis des Grenzwerts $\lim_{x\to-\infty}\exp(x)=-\infty$ sei $(x_n)_{n\in\mathbb{N}}$ eine Folge in $\mathbb{R}$ mit $\lim_n x_n=-\infty$. Dann gilt $\lim_n(-x_n)=+\infty$. Ist nämlich $\kappa\in\mathbb{R}^+$ vorgegeben, dann gibt es ein $N\in\mathbb{N}$ mit $x_n<-\kappa$ für alle $n\geq N$, und für diese n gilt dann auch $-x_n>\kappa$. Wegen $\lim_{x\to+\infty}\exp(x)=+\infty$ folgt $\lim_n\exp(-x_n)=+\infty$. Auf Grund der Rechenregeln aus § 8 für uneigentliche Grenzwerte folgt daraus $\lim_n\exp(-x_n)^{-1}=0$, und wegen $\exp(-x_n)^{-1}=\exp(-(-x_n))=\exp(x_n)$ gilt auch $\lim_n\exp(x_n)=0$.

zu (iii) Um zu zeigen, dass exp streng monoton wachsend ist, bemerken wir zunächst, dass $\exp(x)=1+x+r_2(x)\geq 1+x>1$ für alle $x\in\mathbb{R}^+$ gilt. Seien nun $x,y\in\mathbb{R}$ mit $x<y$ vorgegeben. Aus $y-x>0$ folgt dann $\exp(y-x)>1$, und daraus folgt $\exp(x)<\exp(x)\exp(y-x)=\exp(x+(y-x))=\exp(y)$.

Als streng monoton wachsende Funktion ist exp jedenfalls injektiv. Zum Nachweis der Surjektivität sei $y\in\mathbb{R}^+$ vorgegeben. Zu zeigen ist, dass ein $x\in\mathbb{R}$ mit $\exp(x)=y$ existiert. Dazu sei $(x_n)_{n\in\mathbb{N}}$ eine Folge mit $\lim_n x_n=+\infty$. Wegen $\lim_n -x_n=-\infty$ und $\lim_{x\to-\infty}\exp(x)=0$ gilt $\lim_n\exp(-x_n)=0$. Insbesondere gibt es ein $m\in\mathbb{N}$ mit $\exp(-x_m)<y$. Wegen $\lim_{x\to+\infty}\exp(x)=+\infty$ gilt $\lim_n\exp(x_n)=+\infty$, und deshalb gibt es insbesondere ein $n\in\mathbb{N}$ mit $\exp(x_n)>y$, wobei wir $x_n>-x_m$ annehmen können. Auf Grund des Zwischenwertsatzes (5.15), angewendet auf die Funktion $\exp|_{[-x_m,y_n]}$, existiert ein $x\in\,]x_m,y_n[$ mit $\exp(x)=y$. Damit ist die Surjektivität nachgewiesen. □

Aus dem letzten Punkt es soeben bewiesenen Satzes folgt

(6.11) Satz Die Exponentialfunktion besitzt eine Umkehrfunktion $\ln:\mathbb{R}^+\to\mathbb{R}$, die als ***natürlicher Logarithmus*** bezeichnet wird. Sie hat die folgenden Eigenschaften.

(i) Sie ist stetig, streng monoton wachsend und definiert eine bijektive Abbildung zwischen $\mathbb{R}^+$ und $\mathbb{R}$.

(ii) Es gilt $\ln(\exp(x))=x$ für alle $x\in\mathbb{R}$ und $\exp(\ln(y))=y$ für alle $y\in\mathbb{R}^+$.

(iii) Weiter gilt $\ln(1)=0$, $\ln(e)=1$, $\ln(xy)=\ln(x)+\ln(y)$ und $\ln(x^{-1})=-\ln(x)$ für alle $x,y\in\mathbb{R}^+$.

(iv) $\lim\limits_{x\to0}\ln(x)=-\infty$ und $\lim\limits_{x\to+\infty}\ln(x)=+\infty$.

Beweis: zu (i) Als Umkehrfunktion einer bijektiven, stetigen und streng monoton wachsenden Funktion ist ln nach Satz (5.19) ebenfalls stetig und streng monoton wachsend, und natürlich ist die Umkehrabbildung einer bijektiven Abbildung selbst bijektiv.

zu (ii) Durch diese Gleichungen kommt lediglich zum Ausdruck, dass ln die Umkehrfunktion von exp ist.

zu (iii) Wegen $\exp(0) = 1$ und $\exp(1) = e$ gilt $\ln(1) = \ln(\exp(0)) = 0$ und $\ln(e) = \ln(\exp(1)) = 1$. Seien nun $x, y \in \mathbb{R}^+$ vorgegeben. Dann gilt $\ln(xy) = \ln(\exp(\ln(x))\exp(\ln(y))) = \ln(\exp(\ln(x) + \ln(y))) = \ln(x) + \ln(y)$ und $\ln(x^{-1}) = \ln(\exp(\ln(x))^{-1}) = \ln(\exp(-\ln(x))) = -\ln(x)$.

zu (iv) Sei $(x_n)_{n\in\mathbb{N}}$ eine Folge in $\mathbb{R}^+$ mit $\lim_n x_n = 0$. Zu zeigen ist $\lim_n \ln(x_n) = -\infty$. Sei dazu $\kappa \in \mathbb{R}^+$ vorgegeben und $\alpha = \exp(-\kappa)$. Dann gibt es ein $N \in \mathbb{N}$ mit $x_n < \alpha$ für alle $n \geq N$. Weil ln streng monoton wachsend ist, folgt daraus $\ln(x_n) < \ln(\alpha) = -\kappa$ für alle $n \geq N$. Damit ist der Beweis von $\lim_{x\to 0} \ln(x) = -\infty$ abgeschlossen. Der Beweis des zweiten Grenzwerts läuft nach demselben Schema. Sei $(x_n)_{n\in\mathbb{N}}$ eine Folge in $\mathbb{R}^+$ mit $\lim_n x_n = +\infty$. Dann muss $\lim_n \ln(x_n) = +\infty$ gezeigt werden. Sei dazu wieder ein $\kappa \in \mathbb{R}^+$ vorgegeben, und sei $\alpha = \exp(\kappa)$. Dann existiert ein $N \in \mathbb{N}$ mit $x_n > \alpha$ für alle $n \geq N$, und weil ln streng monoton wachsend ist, gilt für diese n auch $\ln(x_n) > \ln(\alpha) = \kappa$. □

Für Anwendungen im technisch-naturwissenschaftlichen Bereich ist es hilfreich, Exponentialfunktion und Logarithmus zu einer beliebigen Basis $b \in \mathbb{R}$ mit $b > 1$ zu definieren.

(6.12) Definition Für jedes $b \in \mathbb{R}$ mit $b > 1$ sind die Exponentialfunktion und der Logarithmus zur Basis b definiert durch

$$\exp_b : \mathbb{R} \to \mathbb{R}^+ \quad , \quad x \mapsto \exp(x\ln(b)) \qquad \text{und} \qquad \log_b : \mathbb{R}^+ \to \mathbb{R} \quad , \quad x \mapsto \frac{\ln(x)}{\ln(b)}.$$

Die Rechenregeln für die Exponentialfunktion und den natürlichen Logarithmus übertragen sich unmittelbar auf die entsprechenden Funktionen zur Basis b.

(6.13) Proposition Sei $b \in \mathbb{R}$ mit $b > 1$.

(i) Es gilt $\exp_b(0) = 1$ und $\exp_b(1) = b$, außerdem $\exp_b(x + y) = \exp_b(x)\exp_b(y)$, $\exp_b(-x) = \exp_b(x)^{-1}$ und $\exp_b(x) > 0$ für alle $x, y \in \mathbb{R}$.

(ii) Weiter gilt $\log_b(1) = 0$, $\log_b(b) = 1$, $\log_b(xy) = \log_b(x) + \log_b(y)$ und $\log_b(x^{-1}) = -\log_b(x)$ für alle $x, y \in \mathbb{R}^+$.

Beweis: zu (i) Seien $x, y \in \mathbb{R}$. Nach Definition von $\exp_b$ gilt $\exp_b(0) = \exp(0\cdot\ln(b)) = \exp(0) = 1$, $\exp_b(1) = \exp(1\cdot\ln(b)) = b$, $\exp_b(x+y) = \exp((x+y)\ln(b)) = \exp(x\ln(b)+y\ln(b)) = \exp(x\ln(b))\exp(y\ln(b)) = \exp_b(x)\exp_b(y)$ und $\exp_b(-x) = \exp((-x)\ln(b)) = \exp(-x\ln(b)) = \exp(x\ln(b))^{-1} = \exp_b(x)^{-1}$.

zu (ii) Für $x, y \in \mathbb{R}^+$ gilt $\log_b(1) = \frac{\ln(1)}{\ln(b)} = \frac{0}{\ln(b)} = 0$, $\log_b(b) = \frac{\ln(b)}{\ln(b)} = 1$, $\log_b(xy) = \frac{\ln(xy)}{\ln(b)} = \frac{\ln(x)}{\ln(b)} + \frac{\ln(y)}{\ln(b)} = \log_b(x) + \log_b(y)$ und $\log_b(x^{-1}) = \frac{\ln(x^{-1})}{\ln(b)} = -\frac{\ln(x)}{\ln(b)} = -\log_b(x)$. □

(6.14) Proposition Für alle $b \in \mathbb{R}$ mit $b > 1$ und alle $r \in \mathbb{Q}$ gilt $b^r = \exp_b(r)$, wobei der Ausdruck b^r wie in § 1 durch Potenzen und die Wurzelfunktionen definiert ist.

Beweis: Zunächst beweisen wir die Gleichung $b^n = \exp_b(n)$ für alle $n \in \mathbb{N}_0$ durch vollständige Induktion. Es ist $b^0 = 1 = \exp(0) = \exp(0 \cdot \ln(b)) = \exp_b(0)$. Ist $n \in \mathbb{N}_0$ und setzen wir $b^n = \exp_b(n)$ voraus, dann folgt $b^{n+1} = b^n \cdot b = \exp_b(n)\exp_b(1) = \exp_b(n+1)$. Für alle $n \in \mathbb{N}$ gilt auch $b^{-n} = (b^n)^{-1} = \exp_b(n)^{-1} = \exp_b(-n)$. Also ist $b^n = \exp_b(n)$ für alle $n \in \mathbb{Z}$ erfüllt. Sei nun $r \in \mathbb{Q}$, $r = \frac{m}{n}$ mit $m \in \mathbb{Z}$ und $n \in \mathbb{N}$. Zunächst bemerken wir, dass $\exp_b(r)^n = \exp_b(nr)$ gilt, wie man mit der Gleichung $\exp_b(x+y) = \exp_b(x)\exp_b(y)$ durch vollständige Induktion über n leicht zeigt. Es folgt nun $\exp_b(r)^n = \exp_b(nr) = b^{nr} = b^m$. Weil $\sqrt[n]{b^m}$ die eindeutig bestimmte Zahl in $\mathbb{R}_+$ ist, deren n-te Potenz gleich b^m ist, folgt $\exp_b(r) = \sqrt[n]{b^m} = b^{m/n} = b^r$. □

Wegen Prop. (6.13) ist es gerechtfertigt, b^r für alle $b \in \mathbb{R}^+$ und $r \in \mathbb{R}$ wie folgt zu definieren: Ist $b > 1$, dann setzt man $b^r = \exp_b(r)$. Für $r \in \mathbb{Q}$ stimmt die neue Definition von b^r mit der alten Definition aus § 6 überein. Gilt $0 < b < 1$, dann setzt man $b^r = \exp_{b^{-1}}(r)^{-1} = \exp_{b^{-1}}(-r)$. Schließlich definiert man noch $1^r = 1$ für alle $r \in \mathbb{R}$. Man beachte, dass man in allen drei Fällen genauso gut auch $b^r = \exp(r\ln(b))$ schreiben kann. Im Fall $b > 0$ folgt dies direkt aus der Definition. Im Fall $0 < b < 1$ ist $b^r = \exp_{b^{-1}}(-r) = \exp(-r\ln(b^{-1})) = \exp((-r)(-\ln(b))) = \exp(r\ln(b))$, und für $b = 1$ gilt ebenfalls $b^r = 1 = \exp(r \cdot 0) = \exp(r \cdot \ln(b))$.

(6.15) Folgerung Für alle $a, b \in \mathbb{R}^+$ und $x, y \in \mathbb{R}$ gelten die Potenzgesetze

$$(ab)^x = a^x b^x \quad , \quad a^{x+y} = a^x a^y \quad \text{und} \quad (a^x)^y = a^{xy}.$$

Beweis: Alle drei Regeln ergeben sich direkt aus der Gleichung $b^r = \exp(r\ln(b))$. Die erste Gleichung erhält man durch die Rechnung

$$\begin{aligned}(ab)^x &= \exp(x\ln(ab)) = \exp(x(\ln(a)+\ln(b))) = \exp(x\ln(a)+x\ln(b)) \\ &= \exp(x\ln(a))\exp(x\ln(b)) = a^x b^x\end{aligned}$$

die zweite durch $a^{x+y} = \exp((x+y)\ln(a)) = \exp(x\ln(a)+y\ln(a)) = \exp(x\ln(a))\exp(y\ln(a)) = a^x a^y$, und die dritte schließlich durch

$$(a^x)^y = \exp(y\ln(a^x)) = \exp(xy\ln(a)) = a^{xy}.$$ □

Zur Vereinfachung der Notation werden wir von nun an statt $\exp(x)$ meistens den Ausdruck e^x verwenden.

Wenden wir uns nun den trigonometrischen Funktionen, also Sinus und Kosinus, zu.

(6.16) Satz Sinus- und Kosinusreihe konvergieren für alle $x \in \mathbb{R}$ absolut. Sie definieren auf ganz $\mathbb{R}$ stetige Funktionen, die als ***Sinus-*** bzw. ***Kosinusfunktion*** bezeichnet werden. Für alle $x \in \mathbb{R}$ gilt $\cos(-x) = \cos(x)$ und $\sin(-x) = -\sin(x)$.

Beweis: Die Begründungen für die absolute Konvergenz und die Stetigkeit sind wortwörtlich dieselben wie bei der Exponentialfunktion. Die beiden Gleichungen ergeben sich direkt durch Einsetzen in die jeweilige Reihe. Beispielsweise gilt $\cos(-x) = \sum_{n=0}^{\infty}(-1)^n \frac{(-x)^{2n}}{(2n)!} = \sum_{n=0}^{\infty}(-1)^n \frac{x^{2n}}{(2n)!} = \cos(x)$ für alle $x \in \mathbb{R}$. □

Das folgende Resultat ermöglicht es, einen Teil unserer Ergebnisse zur Exponentialfunktion auf diese Funktionen zu übertragen. Weil der Konvergenzradius der Exponentialfunktion, wie wir oben gesehen haben, gleich $+\infty$ ist, dürfen nicht nur beliebige reelle Zahlen, sondern auch beliebige komplexe Zahlen in die Exponentialreihe eingesetzt werden. Wie bei den reellen Zahlen bezeichnen wir den Wert, der sich durch Einsetzen von $z \in \mathbb{C}$ in die Exponentialreihe ergibt, mit e^z. Es gilt nun die

(6.17) Satz *(Eulersche Gleichung)*
Für alle $x \in \mathbb{R}$ gilt $e^{ix} = \cos(x) + i\sin(x)$.

Beweis: Sei $x \in \mathbb{R}$ beliebig vorgegeben. Weil die Kosinus- und die Sinusreihe in x absolut konvergieren, dürfen wir die Grenzwertsätze (4.7) für Reihen anwenden und erhalten

$$\cos(x) + i\sin(x) = \sum_{n=0}^{\infty}(-1)^n \frac{x^{2n}}{(2n)!} + i\sum_{n=0}^{\infty}(-1)^n \frac{x^{2n+1}}{(2n+1)!} = \sum_{n=0}^{\infty}(-1)^n \left(\frac{x^{2n}}{(2n)!} + \frac{ix^{2n+1}}{(2n+1)!}\right)$$

Setzen wir $a_n = (-1)^n \left(\frac{x^{2n}}{(2n)!} + \frac{ix^{2n+1}}{(2n+1)!}\right)$ für alle $n \in \mathbb{N}_0$, dann gilt jeweils

$$\begin{aligned} a_{2n} + a_{2n+1} &= (-1)^{2n}\left(\frac{x^{4n}}{(4n)!} + \frac{ix^{4n+1}}{(4n+1)!}\right) + (-1)^{2n+1}\left(\frac{x^{4n+2}}{(4n+2)!} + \frac{x^{4n+3}}{(4n+3)!}\right) \\ &= \frac{x^{4n}}{(4n)!} + i\frac{x^{4n+1}}{(4n+1)!} - \frac{x^{4n+2}}{(4n+2)!} - i\frac{ix^{4n+3}}{(4n+3)!} \\ = \frac{(ix)^{4n}}{(4n)!} + \frac{(ix)^{4n+1}}{(4n+1)!} + \frac{(ix)^{4n+2}}{(4n+2)!} + \frac{(ix)^{4n+3}}{(4n+3)!} &= b_{4n} + b_{4n+1} + b_{4n+2} + b_{4n+3} \end{aligned}$$

wobei b_n für jedes $n \in \mathbb{N}_0$ den Term $\frac{(ix)^n}{n!}$ der Reihe $\exp(ix)$ bezeichnet. Damit erhalten wir

$$\begin{aligned} \cos(x) + i\sin(x) &= \sum_{n=0}^{\infty}(a_{2n} + a_{2n+1}) = \sum_{n=0}^{\infty}\big(b_{4n} + b_{4n+1} + b_{4n+2} + b_{4n+3}\big) \\ &= \lim_{n\to\infty}\sum_{k=0}^{4n+3} b_k = \lim_{n\to\infty}\sum_{k=0}^{n} b_k = \sum_{n=0}^{\infty} b_n = \exp(ix). \end{aligned}$$

Im vierten Schritt wurde verwendet, dass für eine konvergente Folge mit einem Grenzwert $a \in \mathbb{C}$ jede Teilfolge gegen denselben Grenzwert konvergiert, was im hier vorliegenden Fall auf den Grenzwert $a = \exp(ix)$ und die Teilfolge $(s_{4n+3})_{n\in\mathbb{N}_0}$ der Folge $(s_n)_{n\in\mathbb{N}}$ der Partialsummen von $\sum_{n=0}^{\infty}\frac{(ix)^n}{n!}$ angewendet wird. □

(6.18) Lemma Die ***Restglieder*** von Kosinus- und Sinusfunktion gegeben durch

$$\cos(x) = \sum_{k=0}^{n} (-1)^k \frac{x^{2k}}{(2k)!} + r_{n+1}(x) \qquad \text{und} \qquad \sin(x) = \sum_{k=0}^{n} (-1)^k \frac{x^{2k+1}}{(2k+1)!} + r'_{n+1}(x)$$

für alle $n \in \mathbb{N}_0$ erfüllen die Abschätzungen

$$|r_{n+1}(x)| \leq \frac{|x|^{2n+2}}{(2n+2)!} \quad \text{für alle} \quad x \in \mathbb{R} \text{ mit } |x| \leq 2n+3$$

und

$$|r'_{n+1}(x)| \leq \frac{|x|^{2n+3}}{(2n+3)!} \quad \text{für alle} \quad x \in \mathbb{R} \text{ mit } |x| \leq 2n+4.$$

Beweis: Für jedes $x \in \mathbb{R}$ und jedes $n \in \mathbb{N}_0$ gilt nach Definition des Restglieds

$$r_{n+1}(x) \quad = \quad \sum_{k=n+1}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!} \quad = \quad \sum_{k=0}^{\infty} (-1)^{n+1+k} \frac{x^{2n+2k+2}}{(2n+2k+2)!} \quad = \quad (-1)^{n+1} \frac{x^{2n+2}}{(2n+2)!} \sum_{k=0}^{\infty} (-1)^k a_k$$

mit

$$a_k = \frac{x^{2k}}{(2n+3)(2n+4) \cdot \ldots \cdot (2n+2k+2)} \qquad \text{für alle} \quad k \in \mathbb{N}_0.$$

Nach Definition ist dann $a_0 = 1$ und

$$a_{k+1} \quad = \quad a_k \frac{x^2}{(2n+2k+3)(2n+2k+4)} \quad \text{für alle} \quad k \in \mathbb{N}_0.$$

Auf Grund der Konvergenz der Reihe gilt $\lim_k a_k = 0$; unter der Voraussetzung $|x| \leq 2n+3$ gilt außerdem $a_{k+1}/a_k < 1$ und somit $a_{k+1} < a_k$ für alle $k \in \mathbb{N}_0$. Sei nun $s_n = \sum_{k=0}^{n} (-1)^k a_k$ die n-te Partialsumme der Reihe $\sum_{n=0}^{\infty} (-1)^n a_n$. Nach dem Leibniz-Kriterium, Satz (4.13), konvergiert diese Reihe. Wie im Beweis dieses Kriteriums überprüft man unmittelbar, dass die Teilfolge $(s_{2n})_{n\in\mathbb{N}}$ monoton fallend, die Teilfolge $(s_{2n+1})_{n\in\mathbb{N}}$ monoton wachsend ist. Wegen $a_1 < 1 = a_0$ ist $s_1 = a_0 - a_1 > 0$, und insgesamt gelten für alle $n \in \mathbb{N}_0$ die Ungleichungen

$$1 \quad = \quad s_0 \geq s_{2n} > s_{2n+1} \geq s_1 \quad > \quad 0 \quad ,$$

also insbesondere $s_n \in [0,1]$ für alle $n \in \mathbb{N}_0$. Damit liegt auch der Grenzwert der Reihe in $[0,1]$, und wir erhalten die gewünschte Abschätzung $|r_{n+1}(x)| \leq \frac{|x|^{2n+2}}{(2n+2)!}$. Für die Restgliedabschätzung der Sinusfunktion müssen die Formeln nur geringfügig modifiziert werden. Hier ist

$$r'_{n+1}(x) \quad = \quad \sum_{k=n+1}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!} \quad = \quad \sum_{k=0}^{\infty} (-1)^{n+1+k} \frac{x^{2n+2k+3}}{(2n+2k+3)!} \quad = \quad (-1)^{n+1} \frac{x^{2n+3}}{(2n+3)!} \sum_{k=0}^{\infty} (-1)^k b_k$$

mit

$$b_k \quad = \quad \frac{x^{2k}}{(2n+4)(2n+5) \cdot \ldots \cdot (2n+2k+3)} \qquad \text{für alle} \quad k \in \mathbb{N}_0.$$

woraus sich $b_0 = 1$ und die Rekursionsformel $b_{k+1} = b_k \frac{x^2}{(2n+2k+4)(2n+2k+5)}$ ergibt. Wie bei der Kosinusfunktion erhält man für $|x| \leq 2n+4$ die Ungleichungen $0 \leq \sum_{k=0}^{\infty} (-1)^k b_k \leq 1$ und damit die gewünschte Abschätzung für $r'_{n+1}(x)$. □

(6.19) Lemma Sinus- und Kosinusfunktion erfüllen für alle $x, y \in \mathbb{R}$ die Gleichungen

(i) $\sin(x+y) = \sin(x)\cos(y) + \cos(x)\sin(y)$
(ii) $\cos(x+y) = \cos(x)\cos(y) - \sin(x)\sin(y)$
(iii) $\sin(x) - \sin(y) = 2\cos(\frac{x+y}{2})\sin(\frac{x-y}{2})$
(iv) $\cos(x) - \cos(y) = -2\sin(\frac{x+y}{2})\sin(\frac{x-y}{2})$ und
(v) $\sin(x)^2 + \cos(x)^2 = 1$.

Die Gleichungen unter (i) und (ii) bezeichnet man als ***Additionstheoreme*** der Sinus- und der Kosinusfunktion.

Beweis: Die beiden Additionstheoreme ergeben sich aus der Eulerschen Gleichung (6.17) und der Tatsache, dass die Gleichung $e^{z+w} = e^z e^w$ auch für $z, w \in \mathbb{C}$ gültig ist; der Nachweis dafür funktioniert genauso wie in Satz (6.10). Für alle $x, y \in \mathbb{R}$ gilt nun

$$\begin{aligned} \cos(x+y) + i\sin(x+y) &= e^{i(x+y)} = e^{ix+iy} = e^{ix}e^{iy} \\ &= (\cos(x) + i\sin(x))(\cos(y) + i\sin(y)) \\ &= (\cos(x)\cos(y) - \sin(x)\sin(y)) + i(\sin(x)\cos(y) + \cos(x)\sin(y)) \quad , \end{aligned}$$

und durch Vergleich von Real- und Imaginärteil auf beiden Seiten erhält man (i) und (ii). Die übrigen Gleichungen können nun ihrerseits aus den Additionstheoremen abgeleitet werden. Setzen wir $u = \frac{x+y}{2}$ und $v = \frac{x-y}{2}$, dann gilt

$$\begin{aligned} \sin(x) - \sin(y) &= \sin(u+v) - \sin(u-v) = \\ \sin(u)\cos(v) + \cos(u)\sin(v) - \sin(u)\cos(-v) - \cos(u)\sin(-v) &= \\ \sin(u)\cos(v) + \cos(u)\sin(v) - \sin(u)\cos(v) + \cos(u)\sin(v) &= \\ 2\cos(u)\sin(v) &= 2\cos\left(\tfrac{x+y}{2}\right)\sin\left(\tfrac{x-y}{2}\right). \end{aligned}$$

und ebenso

$$\begin{aligned} \cos(x) - \cos(y) &= \cos(u+v) - \cos(u-v) = \\ \cos(u)\cos(v) - \sin(u)\sin(v) - \cos(u)\cos(-v) + \sin(u)\sin(-v) &= \\ \cos(u)\cos(v) - \sin(u)\sin(v) - \cos(u)\cos(v) - \sin(u)\sin(v) &= \\ -2\sin(u)\sin(v) &= -2\sin\left(\tfrac{x+y}{2}\right)\sin\left(\tfrac{x-y}{2}\right). \end{aligned}$$

Gleichung (v) folgt aus dem Additionstheorem (ii) durch die Rechnung $\cos(x)^2 + \sin(x)^2 = \cos(x)\cos(-x) - \sin(x)\sin(-x) = \cos(x + (-x)) = \cos(0) = 1$, wobei im ersten Schritt die Gleichungen aus Satz (6.16) verwendet wurden. □

(6.20) Lemma Die Kosinusfunktion besitzt im Intervall $[0, 2]$ genau eine Nullstelle ξ. Das Doppelte 2ξ dieser Nullstelle wird die ***Kreiszahl*** π genannt. Die Werte von Sinus- und Kosinusfunktion an den Stellen $\frac{k}{2}\pi$ mit $0 \le k \le 4$ sind durch folgende Tabelle gegeben.

x	0	$\frac{1}{2}\pi$	π	$\frac{3}{2}\pi$	2π
$\sin(x)$	0	1	0	-1	0
$\cos(x)$	1	0	-1	0	1

Beweis: Zunächst zeigen wir mit Hilfe der Restgliedabschätzungen aus Lemma (6.18), dass

$$\sin(x) > 0 \text{ für alle } x \in]0,2] \qquad \text{und} \qquad \cos(2) \le -\tfrac{1}{3} \quad \text{gilt.}$$

Nach Lemma (6.18) ist $\sin(x) = x + r_1'(x)$ mit $|r_1'(x)| \le \frac{1}{6}|x|^3$ für alle $x \in \mathbb{R}$ mit $|x| \le 4$. Für $x \in]0,2]$ folgt also $|\frac{r_1'(x)}{x}| \le \frac{1}{6}|x|^2 \le \frac{4}{6} = \frac{2}{3}$ und somit $\sin(x) = x + r_1'(x) = x(1 + \frac{r_1'(x)}{x}) \ge x(1 - \frac{2}{3}) = \frac{1}{3}x > 0$. Außerdem ist $\cos(x) = 1 - \frac{1}{2}x^2 + r_2(x)$ mit $|r_2(x)| \le \frac{1}{24}|x|^4$ für alle $x \in \mathbb{R}$ mit $|x| \le 5$. Insbesondere gilt $r_2(2) \le \frac{16}{24} = \frac{2}{3}$, also $\cos(2) \le 1 - \frac{1}{2} \cdot 2^2 + \frac{2}{3} = 1 - 2 + \frac{2}{3} = -\frac{1}{3}$.

Mit Hilfe der Gleichungen aus Lemma (6.19) können wir nun zeigen, dass die Kosinusfunktion im Intervall $[0,2]$ streng monoton fallend ist. Seien $x, y \in \mathbb{R}$ mit $0 \le x < y \le 2$ vorgegeben. Dann gilt $0 < \frac{1}{2}(x+y) \le 2$ und $0 < \frac{1}{2}(y-x) \le 1 < 2$. Daraus folgt $\sin(\frac{x+y}{2}) > 0$ und $\sin(\frac{y-x}{2}) > 0$, also $\cos(y) - \cos(x) = -2\sin(\frac{x+y}{2})\sin(\frac{y-x}{2}) < 0$ und somit $\cos(x) > \cos(y)$.

Durch Einsetzen in die Kosinusreihe findet man den Wert $\cos(0) = 1$. Auf Grund der Stetigkeit der Kosinusfunktion und wegen $\cos(0) = 1$ und $\cos(2) = -\frac{1}{3} < 0$ können wir den Zwischenwertsatz (5.14) anwenden. Dieser liefert uns eine Nullstelle ξ im Intervall $[0,2]$. Weil die Kosinusfunktion im Intervall $[0,2]$ streng monoton fallend ist, kann es in diesem Intervall keine weitere Nullstelle geben. Durch Einsetzen in die Sinusreihe findet man auch den Wert $\sin(0) = 0$. Außerdem gilt $\sin(\frac{1}{2}\pi)^2 + \cos(\frac{1}{2}\pi)^2 = 1$, wegen $\cos(\frac{1}{2}\pi)^2 = 0$ also $\sin(\frac{1}{2}\pi)^2 \in \{\pm 1\}$. Weil die Sinusfunktion im Intervall $]0,2]$ nur positive Werte annimmt und $\frac{1}{2}\pi$ in diesem Intervall enthalten ist, folgt $\sin(\frac{1}{2}\pi) = 1$.

Damit sind die ersten beiden Spalten der Tabelle verifiziert. Die übrigen Werte erhält man nun durch Anwendung der Additionstheoreme von Sinus und Kosinus. Es gilt $\sin(\pi) = \sin(\frac{1}{2}\pi + \frac{1}{2}\pi) = \sin(\frac{1}{2}\pi)\cos(\frac{1}{2}\pi) + \cos(\frac{1}{2}\pi)\sin(\frac{1}{2}\pi) = 1 \cdot 0 + 0 \cdot 1 = 0$ und $\cos(\pi) = \cos(\frac{1}{2}\pi)^2 - \sin(\frac{1}{2}\pi)^2 = 0^2 - 1^2 = -1$. Nach demselben Schema berechnet man

$$\begin{aligned}
\sin(\tfrac{3}{2}\pi) &= \sin(\pi + \tfrac{1}{2}\pi) = \sin(\pi)\cos(\tfrac{1}{2}\pi) + \cos(\pi)\sin(\tfrac{1}{2}\pi) = 0 \cdot 0 + (-1) \cdot 1 = -1 \\
\cos(\tfrac{3}{2}\pi) &= \cos(\pi + \tfrac{1}{2}\pi) = \cos(\pi)\cos(\tfrac{1}{2}\pi) - \sin(\pi)\sin(\tfrac{1}{2}\pi) = (-1) \cdot 0 - 0 \cdot 1 = 0 \\
\sin(2\pi) &= \sin(\pi)\cos(\pi) + \cos(\pi)\sin(\pi) = 0 \cdot (-1) + (-1) \cdot 0 = 0 \\
\cos(2\pi) &= \cos(\pi)\cos(\pi) - \sin(\pi)\sin(\pi) = (-1)(-1) - 0 \cdot 0 = 1
\end{aligned}$$

□

(6.21) Satz

(i) Die Sinusfunktion hat die Nullstellenmenge $\{k\pi \mid k \in \mathbb{Z}\}$. Sie ist im Bereich $[-\frac{1}{2}\pi, \frac{1}{2}\pi]$ streng monoton wachsend und im Bereich $[\frac{1}{2}\pi, \frac{3}{2}\pi]$ streng monoton fallend.

(ii) Die Kosinusfunktion hat die Nullstellenmenge $\{(k+\frac{1}{2})\pi \mid k \in \mathbb{Z}\}$. Sie ist im Bereich $[0, \pi]$ streng monoton fallend umd im Bereich $[\pi, 2\pi]$ streng monoton wachsend.

(iii) Sinus- und Kosinusfunktion erfüllen die Gleichungen $\sin(x + \pi) = -\sin(x)$, $\cos(x + \pi) = -\cos(x)$ und $\cos(x) = \sin(x + \frac{1}{2}\pi)$ für alle $x \in \mathbb{R}$.

Aus (iv) folgt insbesondere $\sin(x + 2\pi) = \sin(x)$ und $\cos(x + 2\pi) = \cos(x)$ für alle $x \in \mathbb{R}$. Man sagt, Sinus- und Kosinusfunktion sind ***periodische Funktionen*** mit Periodenläge 2π.

Beweis: Zunächst bemerken wir, dass sich die unter (iii) angegebenen Gleichungen direkt aus den Additionstheoremen und der Tabelle aus ((6.20)) ergeben. Für alle $x \in \mathbb{R}$ gilt $\sin(x+\pi) = \sin(x)\cos(\pi) + \sin(\pi)\cos(x) = \sin(x)\cdot(-1)+0\cdot\cos(x) = -\sin(x)$ und $\cos(x+\pi) = \cos(x)\cos(\pi) - \sin(x)\sin(\pi) = \cos(x)\cdot(-1) - \sin(x)\cdot 0 = -\cos(x)$. Außerdem ist $\sin(x+\frac{1}{2}\pi) = \sin(x)\cos(\frac{1}{2}\pi) + \cos(x)\sin(\frac{1}{2}\pi) = \sin(x)\cdot 0 + \cos(x)\cdot 1 = \cos(x)$.

Wir wissen bereits, dass $\cos(x)$ im Intervall $[0,2]$, also insbesondere auf $[0,\frac{1}{2}\pi]$ streng monoton fallend ist. Um zu zeigen, dass $\cos(x)$ sogar auf $[0,\pi]$ streng monoton fallend ist, verwenden wir die Gleichung $\cos(\pi - x) = -\cos(-x) = -\cos(x)$. Seien $x, y \in [0,\pi]$ mit $x < y$ vorgegeben. Liegen x und y beide im Intervall $[0,\frac{1}{2}\pi]$, dann ist die Ungleichung $\cos(x) < \cos(y)$ bereits bekannt. Gilt $\frac{1}{2}\pi \le x < y \le \pi$, dann definieren wir $x' = \pi - x$ und $y' = \pi - y$. Es gilt dann $x', y' \in [0,\frac{1}{2}\pi]$ und $y' < x'$. Wir erhalten $\cos(y') > \cos(x')$, also $-\cos(y) < -\cos(x)$ und damit auch hier $\cos(x) > \cos(y)$. Ist schließlich $x \in \left]0,\frac{1}{2}\pi\right[$ und $y \in \left]\frac{1}{2}\pi,\pi\right[$, dann gilt $1 = \cos(0) > \cos(x) > \cos(\frac{1}{2}\pi) = 0$ und $0 = \cos(\frac{1}{2}\pi) > \cos(y) > \cos(\pi) = -1$. Also ist $\cos(x) > \cos(y)$ auch hier erfüllt.

Um zu zeigen, dass die Kosinusfunktion auf $[\pi, 2\pi]$ streng monoton wachsend ist, seien $x, y \in [\pi, 2\pi]$ mit $x < y$ vorgegeben. Dann gilt $x - \pi, y - \pi \in [0,\pi]$ und $x - \pi < y - \pi$. Weil die Kosinusfunktion auf $[0,\pi]$ streng motonoton fallend ist, gilt $\cos(x-\pi) > \cos(y-\pi)$. Daraus folgt $-\cos(x) > -\cos(y)$ und schließlich $\cos(x) < \cos(y)$. Die Untersuchung des Monotonieverhaltens der Sinusfunktion kann entsprechend mit Hilfe der Gleichung $\cos(x) = \sin(x+\frac{1}{2}\pi)$ auf die Kosinusfunktion zurückgespielt werden. Wir beschränken uns auf den Nachweis für das Intervall $[\frac{1}{2}\pi, \frac{3}{2}\pi]$. Seien x, y in diesem Intervall mit $x < y$ vorgegeben. Dann gilt auch $x - \frac{1}{2}\pi < y - \frac{1}{2}\pi$, außerdem $x - \frac{1}{2}\pi, y - \frac{1}{2}\pi \in [0,\pi]$ und $\cos(x - \frac{1}{2}\pi) > \cos(y - \frac{1}{2}\pi)$, weil die Kosinusfunktion auf diesem Intervall streng monoton fallend ist. Es folgt $\sin(x) > \sin(y)$.

Weil die Kosinusfunktion auf $[0,\pi]$ streng monoton fallend ist und $\cos(0) = 1$ und $\cos(\pi) = -1$ gilt, ist $\frac{1}{2}\pi$ in diesem Intervall die einzige Nullstelle. Wegen $\cos(x) = -\cos(x-\pi)$ für alle $x \in [\pi, 2\pi]$ gibt es in diesem Intervall nur die Nullstelle $\frac{3}{2}\pi$. Also sind $\frac{1}{2}\pi, \frac{3}{2}\pi$ in $[0, 2\pi[$ die einzigen Nullstellen der Kosinusfunktion. Wegen $\cos(x+2\pi) = \cos(x)$ für alle $x \in \mathbb{R}$ ist die Nullstellenmenge insgesamt also durch

$$\bigcup_{k\in\mathbb{Z}} \{\tfrac{1}{2}\pi + 2\pi k, \tfrac{3}{2}\pi + 2\pi k\} \quad = \quad \{\tfrac{1}{2}\pi + (2k)\pi, \tfrac{1}{2}\pi + (2k+1)\pi\} \quad = \quad \{\tfrac{1}{2}\pi + k\pi \mid k \in \mathbb{Z}\}$$

gegeben. Die Nullstellenmenge der Sinusfunktion erhält man wiederum mit Hilfe der Gleichung $\cos(x) = \sin(x+\frac{1}{2}\pi)$. □

Schränkt man Sinus- und Kosinusfunktion auf bestimmte bestimmte Intervalle ein, so werden sie zu streng monoton wachsenden oder fallenden Funktionen, die eine Umkehrfunktion besitzen. Auch diese Funktionen spielen in Anwendungen eine wichtige Rolle.

(6.22) Satz

(i) Die Sinusfunktion bildet das Intervall $[-\frac{1}{2}\pi, \frac{1}{2}\pi]$ bijektiv auf das Intervall $[-1,1]$ ab. Die Umkehrfunktion $\arcsin : [-1,1] \to \mathbb{R}$ ist stetig und streng monoton wachsend.

(ii) Die Kosinusfunktion bilet das Intervall $[0,\pi]$ bijektiv auf $[-1,1]$ ab. Die Umkehrfunktion $\arccos : [-1,1] \to \mathbb{R}$ ist stetig und streng monoton fallend.

Man bezeichnet diese Umkehrfunktionen als ***Arcus sinus*** bzw. ***Arcus cosinus***.

Beweis: Wegen $\sin(-\frac{1}{2}\pi) = -1$ und $\sin(\frac{1}{2}\pi) = 1$ folgt aus dem Zwischenwertsatz, dass die Sinusfunktion $[-\frac{1}{2}\pi, \frac{1}{2}\pi]$ surjektiv auf $[-1, 1]$ abbildet, und auf Grund der strengen Monotonie ist die auf $[-\frac{1}{2}\pi, \frac{1}{2}\pi]$ eingeschränkte Abbildung auch injektiv. Entsprechend begründet man die Bijektivität der Kosinusfunktion auf $[0, \pi]$. Alle Aussagen ergeben sich nun direkt aus dem Satz (5.19) über die Stetigkeit von Umkehrfunktionen. □

(6.23) Satz Die ***Tangensfunktion*** ist auf $D = \mathbb{R} \setminus \{\frac{1}{2}\pi + k\pi \mid k \in \mathbb{Z}\}$ durch $\tan(x) = \sin(x)/\cos(x)$ definiert und dort stetig. Ihre Einschränkung auf $\left]-\frac{1}{2}\pi, \frac{1}{2}\pi\right[$ ist streng monoton wachsend und bildet dieses Intervall bijektiv auf $\mathbb{R}$ ab. Die Umkehrfunktion $\arctan : \mathbb{R} \to \mathbb{R}$ ist stetig und ebenfalls streng monoton wachsend. Sie wird als ***Arcus tangens*** bezeichnet.

Beweis: Die Sinusfunktion ist auf $\left]-\frac{1}{2}\pi, \frac{1}{2}\pi\right[$ streng monoton wachsend, die Kosinusfunktion ist auf demselben Intervall streng monoton fallend und positiv. Daraus folgt, dass $x \mapsto \cos(x)^{-1}$ auf dem Intervall streng monoton wachsend ist, und dass dasselbe auch für $\tan(x) = \sin(x)\cos(x)^{-1}$ gilt (denn für beliebige $a, b, c, d \in \mathbb{R}^+$ folgt aus $a < b$ und $c > d$ jeweils, dass $c^{-1} < d^{-1}$ und $ac^{-1} < bd^{-1}$ gilt).

Um zu zeigen, dass

$$\lim_{x \to -\frac{1}{2}\pi} \tan(x) = -\infty \qquad \text{und} \qquad \lim_{x \to +\frac{1}{2}\pi} \tan(x) = +\infty$$

gilt, seien $(x_n)_{n\in\mathbb{N}}$ und (y_n) Folgen in $\left]-\frac{1}{2}\pi, \frac{1}{2}\pi\right[$ mit $\lim_n x_n = -\frac{1}{2}\pi$ und $\lim_n y_n = \frac{1}{2}\pi$. Auf Grund der Stetigkeit von Sinus- und Kosinusfunktion gilt dann $\lim_n \sin(x_n) = -1$, $\lim_n \cos(x_n) = 0$, $\lim_n \sin(y_n) = 1$ und $\lim_n \cos(y_n) = 0$. Außerdem handelt es sich bei $(\cos(y_n))_{n\in\mathbb{N}}$ um Folgen positiver Zahlen. Ähnlich wie bei den Rechenregeln für uneigentliche Grenzwerte aus § 9 folgert man daraus $\lim_n \tan(x_n) = \lim_n \sin(x_n)/\cos(y_n) = -\infty$ und $\lim_n \tan(y_n) = \sin(y_n)/\cos(y_n) = +\infty$, was aber aus Zeitgründen hier nicht ausgeführt wird.

Weil die Tangensfunktion auf $\left]-\frac{1}{2}\pi, \frac{1}{2}\pi\right[$ streng monoton wachsend ist, ist die Einschränkung der Funktion auf dieses Intervall injektiv. Für den Nachweis der Surjektivität verwenden wir erneut den Zwischenwertsatz. Ist $a \in \mathbb{R}^+$ vorgegeben, dann existiert wegen $\lim_{x\to+\frac{1}{2}\pi} \tan(x) = +\infty$ ein $x_1 \in \left]0, \frac{1}{2}\pi\right[$ mit $\tan(x_1) > a$. Die Anwendung des Zwischenwertsatzes auf die Tangensfunktion und das Intervall $[0, x_1]$ liefert wegen $\tan(0) = 0 < \tan(x_1)$ ein $x \in]0, x_1[$ mit $\tan(x) = a$. Also werden alle $a \in \mathbb{R}^+$ von $\tan|_{]-\frac{1}{2}\pi, \frac{1}{2}\pi[}$ getroffen. Im Fall $-a < 0$ wählt man entsprechend ein $x_0 \in \left]-\frac{1}{2}\pi, 0\right[$ mit $\tan(x_0) < a$. Insgesamt ist dann wegen $\tan(x) = 0$ die Surjektivität nachgewiesen. Das streng monotone Wachstum und die Stetigkeit des Arcus tangens ergeben sich dann erneut aus Satz (5.19). □

§ 7. *Differenzierbarkeit und Ableitungsregeln*

Inhaltsübersicht

Eine Funktion $f : D \to \mathbb{R}$ wird ***differenzierbar*** in einem Punkt a des Definitionsbereichs D genannt, wenn a ein innerer Punkt von D ist und der Grenzwert von $\frac{f(x)-f(a)}{x-a}$ für $x \to a$ existiert. Diesen Grenzwert nennt man dann die ***Ableitung*** $f'(a)$ von f an der Stelle a. Aus der Schulmathematik ist bekannt, dass der Quotient $\frac{f(x)-f(a)}{x-a}$ als Steigung der Sekante durch die Punkte $(a, f(a))$ und $(x, f(x))$ des Funktionsgraphen interpretiert werden kann. Für $x \to a$ geht die Sekante über in die Tangente des Funktiongraphen im Punkt $(a, f(a))$. Somit kann die Ableitung $f'(a)$ als *Tangentensteigung* interpretiert werden.

Zunächst betrachten wir einige einfache Beispiele für differenzierbare Funktionen, ebenso die Betragsfunktion als prominentes Gegenbeispiel. Anschließend leiten wir die aus der Schulmathematik bekannten ***Ableitungsregeln*** (Summen-, Produkt- und Kettenregel) her. Als neue Regel hinzu kommt die Umkehrregel zur Ableitung von Umkehrfunktionen. Mit ihr kann zum Beispiel die Ableitung des Logarithmus oder des Arkus tangens bestimmt werden.

Wichtige Begriffe und Sätze

- innerer Punkt einer Teilmenge $D \subseteq \mathbb{R}$
- Differenzierbarkeit einer Funktion f in einem Punkt a
- Definition der Ableitung $f'(a)$
- Stetigkeit differezierbarer Funktionen
- Summen- und Produktregel
- Kettenregel
- Umkehrregel

Sei $D \subseteq \mathbb{R}$ eine beliebige Teilmenge. Man nennt $a \in D$ einen ***inneren Punkt*** von D, wenn ein $\varepsilon \in \mathbb{R}^+$ existiert, so dass $]a-\varepsilon, a+\varepsilon[\subseteq D$ erfüllt ist. Offenbar ist ein innerer Punkt immer auch ein Häufungspunkt von D: Wählt man nämlich N hinreichend groß, so dass $\frac{1}{N} < \varepsilon$ erfüllt ist, dann erhält man zum Beispiel durch $(a_n)_{n \geq N}$ mit $a_n = a + \frac{1}{n}$ eine Folge in D, die gegen a konvergiert.

(7.1) Definition Sei $D \subseteq \mathbb{R}$ und $a \in D$ ein innerer Punkt. Eine Funktion $f : D \to \mathbb{R}$ wird ***differenzierbar*** an der Stelle a genannt, wenn

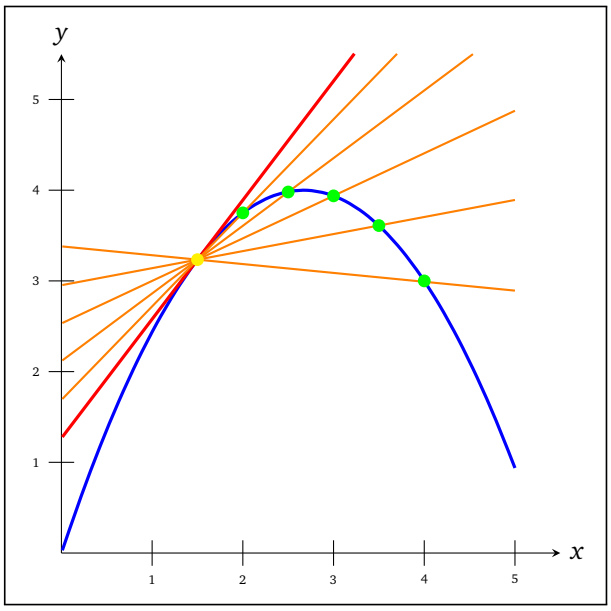
$$f'(a) \quad = \quad \lim_{x \to a} \frac{f(x)-f(a)}{x-a}$$

als eigentlicher Grenzwert existiert. Man nennt $f'(a)$ in diesem Fall die ***Ableitung*** von f an der Stelle a. Wir nennen die Funktion f *differenzierbar*, wenn sie an jeder Stelle ihres Definitionsbereichs differenzierbar ist.

Als innerer Punkt von D ist a zugleich ein Berührpunkt der Menge $D_a = D \setminus \{a\}$. Die Funktion

$$g(x) \quad = \quad \frac{f(x)-f(a)}{x-a}$$

ist auf D_a definiert, somit ist gewährleistet, dass der Grenzwert $\lim\limits_{x\to a} g(x)$ tatsächlich gebildet werden kann. Geometrisch lässt sich $g(x)$ als Steigung der ***Sekante*** des Funktionsgraphen von f durch die Punkte $(a, f(a))$ und $(x, f(x))$ interpretieren. Durch den Grenzübergang $x \to a$ erhält man die Steigung der ***Tangente*** an den Funktionsgraphen durch den Punkt $(a, f(a))$.



Veranschaulichung der Differenzierbarkeit:
Die Sekantensteigungen (orange) nähern sich für $x \to a$ der Tangentensteigung (rot) an.

(7.2) Proposition Sei $f : D \to \mathbb{R}$ eine Funktion und $a \in D$ ein innerer Punkt. Dann sind folgende Bedingungen äquivalent.

(i) Die Funktion f ist in a differenzierbar.

(ii) Es existiert der Grenzwert $\lim\limits_{h\to 0} \dfrac{f(a+h)-f(a)}{h}$.

(iii) Es gibt eine reelle Zahl $c \in \mathbb{R}$ und eine Funktion $\psi : D \to \mathbb{R}$ mit

$$f(a+h) \quad = \quad f(a)+ch+\psi(a+h) \qquad \text{und} \qquad \lim_{h\to 0} \frac{\psi(a+h)}{h} \quad = \quad 0.$$

Sind diese äquivalenten Bedingungen erfüllt, dann stimmt der Grenzwert in (ii) und die reelle Zahl c in (iii) mit dem Wert $f'(a)$ der Ableitung überein.

Beweis: „(i) ⇒ (ii)“ Die Funktion $g(h) = \frac{1}{h}(f(a+h)-f(a))$ besitzt die Menge

$$D' \quad = \quad \{h \in \mathbb{R} \mid a+h \in D\} \setminus \{0\}$$

als Definitionsbereich. Damit der Grenzwert definiert ist, müssen wir zeigen, dass 0 ein Berührpunkt von D' ist. Nach Voraussetzung gibt es eine Folge $(x_n)_{n\in\mathbb{N}}$ in D_a mit $\lim_n x_n = a$. Setzen wir $h_n = x_n - a$ für alle $n \in \mathbb{N}$, so erhalten wir wegen $a + h_n = x_n \in D_a$ eine Folge in D' mit $\lim_n h_n = 0$. Also ist 0 tatsächlich ein Berührpunkt. Ist nun $(h_n)_{n\in\mathbb{N}}$ eine *beliebige* Folge in D' mit $\lim_n h_n = 0$, dann ist durch $x_n = a + h_n$ eine Folge in D_a mit Grenzwert a gegeben, und aus der Differenzierbarkeit von f in a folgt

$$\lim_{n\to\infty} \frac{f(a+h_n)-f(a)}{h_n} \quad = \quad \lim_{n\to\infty} \frac{f(x_n)-f(a)}{x_n-a} \quad = \quad \lim_{x\to a} \frac{f(x)-f(a)}{x-a} \quad = \quad f'(a).$$

„(ii) ⇒ (iii)“ Wir definieren auf der Menge D' von oben die Funktion ϕ durch $\phi(h) = f(a+h) - f(a) - ch$, wobei c den Grenzwert unter (ii) bezeichnet, und setzen $\psi(x) = \phi(x-a)$ für alle $x \in D_a$. Dann gilt offenbar

$$f(a+h) \quad = \quad f(a) + ch + \phi(h) \quad = \quad f(a) + ch + \psi(a+h) \qquad \text{für alle} \quad h \in D'$$

und außerdem

$$\lim_{h\to 0} \frac{\psi(a+h)}{h} \quad = \quad \lim_{h\to 0} \frac{\phi(h)}{h} \quad = \quad \lim_{h\to 0} \frac{f(a+h)-f(a)}{h} - \lim_{h\to 0} \frac{ch}{h} \quad = \quad c - c \quad = \quad 0.$$

„(iii) ⇒ (i)“ Sei $c \in \mathbb{R}$ und $\psi : D_a \to \mathbb{R}$ eine Funktion mit der angegebenen Eigenschaft. Weiter sei $(x_n)_{n\in\mathbb{N}}$ eine Folge in D_a mit $\lim_n x_n = a$. Dann konvergiert die Folge gegeben durch $h_n = x_n - a$ gegen 0, und wir erhalten

$$\lim_{n\to\infty} \frac{f(x_n)-f(a)}{x_n-a} \quad = \quad \lim_{n\to\infty} \frac{f(a+h_n)-f(a)}{h_n} \quad = \quad \lim_{n\to\infty} \frac{ch_n}{h_n} + \lim_{n\to\infty} \frac{\psi(a+h_n)}{h_n} \quad = \quad c + 0 \quad = \quad c.$$

Dies zeigt, dass der Grenzwert $\lim_{x\to a} \frac{f(x)-f(a)}{x-a}$ existiert und mit c übereinstimmt. Also ist f in a differenzierbar, und es gilt $f'(a) = c$. □

Wir betrachten einige Beispiele für differenzierbare und nicht differenzierbare Funktionen.

(i) Sei $c \in \mathbb{R}$ und die Funktion f gegeben durch $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto c$. Dann gilt $f'(a) = 0$ für alle $a \in \mathbb{R}$, denn es gilt jeweils

$$\lim_{h\to 0} \frac{f(a+h)-f(a)}{h} \quad = \quad \lim_{h\to 0} \frac{c-c}{h} \quad = \quad \lim_{h\to 0} 0 \quad = \quad 0.$$

Die Ableitung von $\mathrm{id} : \mathbb{R} \to \mathbb{R}$ ist gegeben durch $\mathrm{id}'(x) = 1$ für alle $x \in \mathbb{R}$.

(ii) Die Funktion $\mathrm{id} : \mathbb{R} \to \mathbb{R}$, $x \mapsto x$ ist überall differenzierbar, mit Ableitung 1. Dies folgt für alle $a \in \mathbb{R}$ aus der Rechung

$$\lim_{h\to 0} \frac{\mathrm{id}(a+h)-\mathrm{id}(a)}{h} \quad = \quad \lim_{h\to 0} \frac{(a+h)-a}{h} \quad = \quad \lim_{h\to 0} \frac{h}{h} \quad = \quad \lim_{h\to 0} 1 \quad = \quad 1.$$

(iii) Die Funktion $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto x^2$ ist überall differenzierbar mit der Ableitung $f'(x) = 2x$ für alle $x \in \mathbb{R}$. Denn für jedes reelle a gilt

$$\lim_{h\to 0} \frac{f(a+h)-f(a)}{h} \quad = \quad \lim_{h\to 0} \frac{(a+h)^2-a^2}{h} \quad = \quad \lim_{h\to 0} \frac{2ah+h^2}{h} \quad = \quad 2a + \lim_{h\to 0} h$$

$$= \quad 2a + 0 \quad = \quad 2a.$$

(iv) Die Funktion $f : D \to \mathbb{R}$, $x \mapsto \frac{1}{x}$ ist auf ihrem gesamten Definitionsbereich $D = \mathbb{R} \setminus \{0\}$ differenzierbar, mit der Ableitung $f'(x) = -\frac{1}{x^2}$ für jedes $x \in D$. Zum Beweis bemerken wir zunächst, dass jedes $a \in D$ ein innerer Punkt des Definitionsbereichs ist. Setzen wir nämlich $\varepsilon = |a|$, dann gilt $\varepsilon \in \mathbb{R}^+$. Ist $a < 0$ und $x \in]a-\varepsilon, a+\varepsilon[$, dann gilt $x < a + \varepsilon = a + (-a) = 0$ und somit $x \in D$. Im Fall $a > 0$ und $x \in]a-\varepsilon, a+\varepsilon[$ gilt $x > a - \varepsilon = a - a = 0$, also wiederum $x \in D$. In beiden Fälle ist also $]a-\varepsilon, a+\varepsilon[\subseteq D$ erfüllt.
Bestimmen wir nun den Wert der Ableitung an einer beliebigen Stelle $a \in D$. Es gilt

$$\lim_{h\to 0} \frac{f(a+h)-f(a)}{h} = \lim_{h\to 0} \frac{\frac{1}{a+h} - \frac{1}{a}}{h} = \lim_{h\to 0} \frac{\frac{a-(a+h)}{a(a+h)}}{h} =$$
$$\lim_{h\to 0} \frac{\frac{(-h)}{a(a+h)}}{h} = \lim_{h\to 0} -\frac{1}{a(a+h)} = -\frac{1}{a^2}$$

(v) Die Funktion $\mathrm{abs} : \mathbb{R} \to \mathbb{R}$ ist an jedem Punkt $a \in \mathbb{R}^\times$ differenzierbar, aber im Punkt 0 nicht differenzierbar. Sei zunächst $a < 0$ und $(h_n)_{n\in\mathbb{N}}$ eine beliebige Folge mit $\lim_n h_n = 0$. Setzen wir $\varepsilon = |a|$, so gibt es ein $N \in \mathbb{N}$ mit $|h_n| < \varepsilon$ für alle $n \geq N$. Für diese n gilt dann $a + h_n < a + \varepsilon = a + (-a) = 0$ und somit $|a + h_n| = -(a + h_n)$. Weil sich endlich viele Folgenglieder auf den Grenzwert nicht auswirken, erhalten wir

$$\lim_{n\to\infty} \frac{\mathrm{abs}(a+h_n)-\mathrm{abs}(a)}{h_n} = \lim_{n\to\infty} \frac{|a+h_n|-|a|}{h_n} =$$
$$\lim_{n\to\infty} \frac{-(a+h_n)+a}{h_n} = \lim_{n\to\infty} \frac{-h_n}{h_n} = -1$$

Damit ist $\mathrm{abs}'(a) = -1$ bewiesen. Betrachten wir nun den Fall $a > 0$. Sei wiederum $(h_n)_{n\in\mathbb{N}}$ eine Folge mit $\lim_n h_n = 0$. Wie im vorherigen Fall zeigt man, dass $a + h_n > 0$ und somit $|a + h_n| = a + h_n$ für alle bis auf endlich viele $n \in \mathbb{N}$ gilt. Man erhält dann

$$\lim_{n\to\infty} \frac{\mathrm{abs}(a+h_n)-\mathrm{abs}(a)}{h_n} = \lim_{n\to\infty} \frac{a+h_n-a}{h_n} = \lim_{n\to\infty} \frac{h_n}{h_n} = 1$$

und somit $\mathrm{abs}'(a) = 1$. Nehmen wir nun an, die Funktion abs ist auch im Nullpunkt differenzierbar. Dies würde bedeuten, dass der Grenzwert

$$\lim_{h\to 0} \frac{\mathrm{abs}(h)-\mathrm{abs}(0)}{h} = \lim_{h\to 0} \frac{\mathrm{abs}(h)}{h}$$

existiert. Um ihn zu auszurechnen, betrachten wir die Folge $(\frac{1}{n})_{n\in\mathbb{N}}$ mit $\lim_n \frac{1}{n} = 0$. Wir erhalten

$$\lim_{h\to 0} \frac{\mathrm{abs}(h)}{h} = \lim_{n\to\infty} \frac{\mathrm{abs}(\frac{1}{n})}{\frac{1}{n}} = \lim_{n\to\infty} \frac{\frac{1}{n}}{\frac{1}{n}} = \lim_{n\to\infty} 1 = 1.$$

Wenn der Grenzwert also existiert, dann muss er mit 1 übereinstimmen. Betrachten wir aber andererseits die Folge $(-\frac{1}{n})_{n\in\mathbb{N}}$, die ebenfalls gegen Null konvergiert, so erhalten wir

$$\lim_{h\to 0} \frac{\mathrm{abs}(h)}{h} = \lim_{n\to\infty} \frac{\mathrm{abs}(-\frac{1}{n})}{(-\frac{1}{n})} = \lim_{n\to\infty} \frac{\frac{1}{n}}{(-\frac{1}{n})} = \lim_{n\to\infty} -1 = -1.$$

Der Widerspruch $1 = \lim_{h\to 0} \frac{\mathrm{abs}(h)}{h} = -1$ zeigt, dass der Grenzwert nicht existiert.

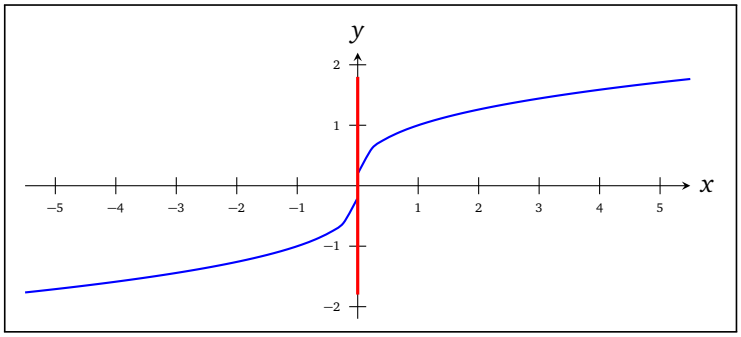
(vi) Die Funktion $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto \sqrt[3]{x}$ ist im Punkt $a = 0$ nicht differenzierbar. Zum Nachweis genügt es, eine Folge $(h_n)_{n\in\mathbb{N}}$ mit $\lim_n h_n = 0$ anzugeben, für die der Grenzwert

$$\lim_{n\to\infty} \tfrac{1}{h_n}(\sqrt[3]{h_n} - \sqrt[3]{0})$$

nicht existiert. Sei $h_n = \frac{1}{n^3}$ für alle $n \in \mathbb{N}$. Dann gilt $\lim\limits_{n\to\infty} h_n = 0$ und $\sqrt[3]{h_n} = \frac{1}{n}$ für alle $n \in \mathbb{N}$. Wir erhalten

$$\frac{\sqrt[3]{h_n} - \sqrt[3]{0}}{h_n} \quad = \quad \frac{\frac{1}{n}}{\frac{1}{n^3}} \quad = \quad \frac{1}{n} \cdot n^3 \quad = \quad n^2.$$

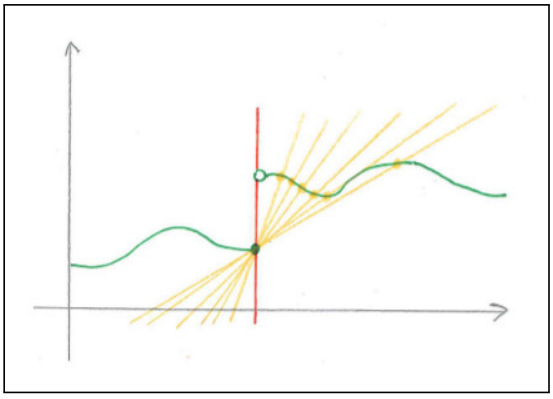
Die Folge ist unbeschränkt und somit nicht konvergent.



Die Kubikwurzel-Funktion $x \mapsto \sqrt[3]{x}$ ist im Nullpunkt nicht differenzierbar.
Der Funktionsgraph hat dort eine senkrechte Tangente (rot), die Steigung wird unendlich groß.

In jedem einzelnen Beispiel hätte man nach Proposition (7.2) an Stelle des Grenzwerts $\lim\limits_{h\to 0}$ auch den entsprechenden Grenzwert $\lim\limits_{x\to a}$ betrachten können, dies hätte an der Rechnung jeweils wenig geändert. Zum Abschluss bemerken wir noch

(7.3) Satz Sei $f : D \to \mathbb{R}$ eine Funktion und $a \in D$. Ist f im Punkt a differenzierbar, dann ist sie dort auch stetig.



An einer Unstetigkeitstelle a kann eine Funktion nicht differenzierbar sein.
Die Steigungen der Sekanten (gelb) laufen für $x \to a$ gegen einen unendlichen Wert.

Beweis: Wir zeigen, dass unter der angegebenen Voraussetzung $\lim_{x\to a} f(x) = f(a)$ erfüllt ist. Als innerer Punkt von D ist a auch ein Häufungspunkt von D und ein Berührpunkt von $D_a = D \setminus \{a\}$. Sei nun $(x_n)_{n\in\mathbb{N}}$ eine Folge in D_a mit $\lim_n x_n = a$. Die Differenzierbarkeit von f an der Stelle a liefert nun

$$\lim_{x\to a} f(x) \quad = \quad \lim_{x\to a}\left(f(a) + \frac{f(x)-f(a)}{x-a}(x-a)\right) \quad = \quad f(a) + \left(\lim_{x\to a}\frac{f(x)-f(a)}{x-a}\right)\left(\lim_{x\to a}(x-a)\right)$$

$$= \quad f(a) + f'(a)\cdot 0 \quad = \quad f(a). \qquad \square$$

Ist eine Funktion $f : D \to \mathbb{R}$ an einer Stelle a nicht stetig, so ist sie also dort auch nicht differenzierbar.

Um auch kompliziertere Funktionen ableiten zu können, beweisen wir nun eine Reihe von Ableitungsregeln, die zum Teil bereits aus der Schulmathematik bekannt sind.

(7.4) Satz *(Summenregel)*

Sei $D \subseteq \mathbb{R}$ eine beliebige Teilmenge, $f, g : D \to \mathbb{R}$ Funktionen und $a \in D$ ein Punkt, in dem f und g beide differenzierbar sind. Dann ist auch $f + g$ in a differenzierbar, und es gilt

$$(f+g)'(a) \quad = \quad f'(a) + g'(a).$$

Beweis: Nach Voraussetzung ist a ein innerer Punkt von D. Weil f und g in a differenzierbar sind, existieren die Grenzwerte

$$\lim_{x\to a}\frac{f(x)-f(a)}{x-a} = f'(a) \qquad \text{und} \qquad \lim_{x\to a}\frac{g(x)-g(a)}{x-a} = g'(a).$$

Auf Grund der Proposition (5.6) über Funktionsgrenzwerte erhalten wir

$$\lim_{x\to a}\frac{(f+g)(x)-(f+g)(a)}{x-a} \quad = \quad \lim_{x\to a}\frac{f(x)+g(x)-f(a)-g(a)}{x-a} \quad =$$

$$\lim_{x\to a}\left(\frac{f(x)-f(a)}{x-a} + \frac{g(x)-g(a)}{x-a}\right) \quad =$$

$$\lim_{x\to a}\frac{f(x)-f(a)}{x-a} + \lim_{x\to a}\frac{g(x)-g(a)}{x-a} \quad = \quad f'(a) + g'(a). \qquad \square$$

(7.5) Satz *(Produktregel)*

Sei $D \subseteq \mathbb{R}$ eine beliebige Teilmenge, $f, g : D \to \mathbb{R}$ Funktionen und $a \in D$ ein Punkt, in dem f und g beide differenzierbar sind. Dann ist auch $f\,g$ in a differenzierbar, und es gilt

$$(f\,g)'(a) \quad = \quad f'(a)g(a) + f(a)g'(a).$$

Beweis: Auch hier ist nach Voraussetzung ist a ein innerer Punkt von D, und weil f und g in a differenzierbar sind, existieren die Grenzwerte

$$\lim_{x\to a}\frac{f(x)-f(a)}{x-a}=f'(a)\qquad\text{und}\qquad\lim_{x\to a}\frac{g(x)-g(a)}{x-a}=g'(a).$$

Ferner ist nach s(7.3) eine differenzierbare Funktion immer auch stetig. Wenden wir dies auf die Funktion g an, so erhalten wir $\lim\limits_{x\to a}g(x)=g(a)$, vgl. Proposition (5.8).

Wiederum durch Anwendung der Rechenregeln für Funktionsgrenzwerte erhalten wir

$$\begin{aligned}\lim_{x\to a}\frac{(fg)(x)-(fg)(a)}{x-a} &= \lim_{x\to a}\frac{f(x)g(x)-f(a)g(a)}{x-a} =\\ \lim_{x\to a}\frac{f(x)g(x)-f(a)g(x)+f(a)g(x)-f(a)g(a)}{x-a} &= \lim_{x\to a}\frac{(f(x)-f(a))g(x)}{x-a}+\lim_{x\to a}\frac{f(a)(g(x)-g(a))}{x-a} =\\ \left(\lim_{x\to a}\frac{f(x)-f(a)}{x-a}\right)\left(\lim_{x\to a}g(x)\right)+f(a)\left(\lim_{x\to a}\frac{g(x)-g(a)}{x-a}\right) &= f'(a)g(a)+f(a)g'(a).\end{aligned}$$

□

(7.6) Folgerung Für jedes $n\in\mathbb{N}$ sei $f_n:\mathbb{R}\to\mathbb{R}$ gegeben durch $f_n(x)=x^n$ für alle $x\in\mathbb{R}$. Dann gilt $f_n'(a)=na^{n-1}$ für alle $a\in\mathbb{R}$ und $n\in\mathbb{N}$, insbesondere ist die Funktion f_n auf ihrem gesamten Definitionsbereich differenzierbar.

Beweis: Sei $a\in\mathbb{R}$. Wir beweisen die Gleichung $f_n'(a)=na^{n-1}$ durch vollständige Induktion über n. Für $n=1$ gilt $f_1'(a)=\mathrm{id}'(a)=1=1\cdot a^0$. Sei nun $n\in\mathbb{N}$, und setzen wir die Gleichung für n voraus. Dann folgt

$$\begin{aligned}f_{n+1}'(a) &= (f_nf_1)'(a) = f_n'(a)f_1(a)+f_n(a)f_1'(a) \overset{\text{Ind.–V.}}{=} na^{n-1}\cdot a+a^n\cdot 1\\ &= na^n+a^n = (n+1)a^n.\end{aligned}$$

□

Zusammen mit der Summenregel ergibt sich aus Folgerung (7.6) unmittelbar, dass alle Polynomfunktionen auf ganz $\mathbb{R}$ differenzierbar sind. Ohne Beweis merken wir an, dass Summen- und Produktregel auch auf Potenzreihen angewendet werden dürfen. (Der Beweis wird später im Rahmen der Funktionentheorie nachgeliefert.) Dies bedeutet genauer: Ist $f(x)=\sum_{n=0}^{\infty}a_n(x-a)^n$ eine Potenzreihe mit Entwicklungspunkt a und Konvergenzradius $\rho>0$, dann ist die Funktion f im Fall $\rho\in\mathbb{R}^+$ auf $]a-\rho,a+\rho[$, im Fall $\rho=+\infty$ auf ganz $\mathbb{R}$ differenzierbar. Die Ableitung erhält man in diesem Bereich durch

$$f'(x) = \sum_{n=1}^{\infty}na_n(x-a)^{n-1}.$$

Der Ausdruck auf der rechten Seite der Gleichung wird auch die ***formale Ableitung*** der Potenzreihe genannt, weil er dadurch zu Stande kommt, dass man Summen- und Produktregel auf die Summanden $a_n(x-a)^n$ einzeln anwendet.

Leitet man zum Beispiel die Terme $x^n/n!$ der Exponentialreihe ab, so erhält man

$$\exp'(x) = \sum_{n=1}^{\infty}\frac{nx^{n-1}}{n!} = \sum_{n=1}^{\infty}\frac{x^{n-1}}{(n-1)!} = \sum_{n=0}^{\infty}\frac{x^n}{n!} = \exp(x).$$

Die Exponentialfunktion stimmt also mit ihrer eigenen Ableitung überein!

(7.7) Satz *(Quotientenregel)*
Sei $D \subseteq \mathbb{R}$ eine beliebige Teilmenge, seien $f, g : D \to \mathbb{R}$ Funktionen mit $g(x) \neq 0$ für alle $x \in D$. Sei $a \in D$ ein Punkt mit der Eigenschaft, dass f und g in a differenzierbar sind. Dann ist die Funktion $\frac{f}{g}$ in a differenzierbar, und es gilt

$$\left(\frac{f}{g}\right)'(a) \quad = \quad \frac{f'(a)g(a) - f(a)g'(a)}{g(a)^2}.$$

Beweis: Auf Grund der Voraussetzungen existiert der Grenzwert

$$\lim_{x \to a} \frac{g(x) - g(a)}{x - a} \quad = \quad g'(a).$$

Nach Satz (7.3) ist g an der Stelle a auch stetig, es gilt also $\lim\limits_{x \to a} g(x) = g(a)$. Wir zeigen nun zunächst, dass die Funktion $\frac{1}{g}$ an der Stelle a differenzierbar ist und die Gleichung

$$\left(\frac{1}{g}\right)'(a) \quad = \quad -\frac{g'(a)}{g(a)^2}$$

erfüllt ist. Auf Grund der Rechenregeln für Funktionsgrenzwerte und wegen $g(a) \neq 0$ gilt

$$\begin{aligned} \lim_{x \to a} \frac{\frac{1}{g(x)} - \frac{1}{g(a)}}{x - a} \quad &= \quad \lim_{x \to a} \frac{g(a) - g(x)}{x - a} \frac{1}{g(x)g(a)} \quad = \quad \left(\lim_{x \to a} \frac{g(a) - g(x)}{x - a}\right)\left(\lim_{x \to a} \frac{1}{g(x)g(a)}\right) \\ &= \quad (-g'(a)) \cdot \frac{1}{g(a)g(a)} \quad = \quad -\frac{g'(a)}{g(a)^2}. \end{aligned}$$

Mit Hilfe der Produktregel erhalten wir nun

$$\begin{aligned} \left(\frac{f}{g}\right)'(a) \quad &= \quad \left(f \cdot \frac{1}{g}\right)'(a) \quad = \quad f'(a)\left(\frac{1}{g}\right)(a) + f(a)\left(\frac{1}{g}\right)'(a) \quad = \\ \frac{f'(a)}{g(a)} + f(a)\left(-\frac{g'(a)}{g(a)^2}\right) \quad &= \quad \frac{f'(a)g(a)}{g(a)^2} + \left(\frac{f(a)(-g'(a))}{g(a)^2}\right) \quad = \quad \frac{f'(a)g(a) - f(a)g'(a)}{g(a)^2}. \end{aligned}$$

□

Wir haben oben gezeigt, dass für jedes $n \in \mathbb{N}$ die Ableitung von $f_n(x) = x^n$ durch $f_n'(x) = nx^{n-1}$ gegeben ist. Mit der Quotientenregel kann man überprüfen, dass diese Formel auch im Fall $n \in \mathbb{Z}$ mit $n < 0$ die richtige Ableitung liefert. Es gilt dann nämlich $m = -n \in \mathbb{N}$, außerdem $f_m = \frac{1}{f_n}$, und auf Grund der bereits bewiesenen Formel erhalten wir für alle $x \in \mathbb{R}^\times$ jeweils

$$\begin{aligned} f_n'(x) \quad &= \quad \left(\frac{1}{f_m}\right)'(x) \quad = \quad -\frac{f_m'(x)}{f_m(x)^2} \quad = \quad -\frac{mx^{m-1}}{x^{2m}} \quad = \quad -\frac{(-n)x^{-n-1}}{x^{-2n}} \\ &= \quad nx^{-n-1} \cdot x^{2n} \quad = \quad nx^{n-1}. \end{aligned}$$

Der einzige Unterschied zum Fall $n \in \mathbb{N}$ besteht darin, dass der Definitionsbereich von f_n und f_n' nicht durch $\mathbb{R}$, sondern durch $\mathbb{R} = \mathbb{R} \setminus \{0\}$ gegeben ist.

(7.8) Satz *(Kettenregel)*
Seien D und E Teilmengen von $\mathbb{R}$ und $f : D \to \mathbb{R}$, $g : E \to \mathbb{R}$ Funktionen mit $f(D) \subseteq E$. Sei $a \in D$ ein Punkt mit der Eigenschaft, dass f in a differenzierbar ist und dass g in $f(a)$ differenzierbar ist. Dann ist auch $g \circ f : D \to \mathbb{R}$ im Punkt a differenzierbar, und es gilt

$$(g \circ f)'(a) \quad = \quad g'(f(a))f'(a).$$

Beweis: Sei $b = f(a)$ und die Funktion $\tilde{g} : E \to \mathbb{R}$ definiert durch

$$\tilde{g}(y) \quad = \quad \begin{cases} \dfrac{g(y)-g(b)}{y-b} & \text{für } y \neq b \\ g'(b) & \text{für } y = b. \end{cases}$$

Weil die Funktion g in b differenzierbar ist, gilt

$$\lim_{y\to b} \tilde{g}(y) \quad = \quad \lim_{y\to b} \frac{g(y)-g(b)}{y-b} \quad = \quad g'(b) \quad = \quad \tilde{g}(b) \quad ,$$

mit anderen Worten, die Funktion $\tilde{g}$ ist im Punkt b stetig. Weil f im Punkt a und $\tilde{g}$ im Punkt $b = f(a)$ stetig ist, handelt es sich bei $\tilde{g} \circ f$ um eine im Punkt a stetige Funktion. Wir betrachten nun den Differentialquotienten der Funktion $g \circ f$. Ist $x \in D$ ein Punkt mit $f(x) \neq f(a)$, dann gilt

$$\frac{(g \circ f)(x)-(g \circ f)(a)}{x-a} \quad = \quad \frac{(g \circ f)(x)-(g \circ f)(a)}{f(x)-f(a)} \cdot \frac{f(x)-f(a)}{x-a} \quad = \quad (\tilde{g} \circ f)(x) \cdot \frac{f(x)-f(a)}{x-a}.$$

Ist $x \neq a$ und $f(x) = f(a)$, dann ist die Gleichung

$$\frac{(g \circ f)(x)-(g \circ f)(a)}{x-a} \quad = \quad (\tilde{g} \circ f)(x) \cdot \frac{f(x)-f(a)}{x-a}$$

offenbar auch erfüllt, denn dann steht auf beiden Seiten eine Null. Insgesamt erhalten wir somit

$$\lim_{x\to a} \frac{(g \circ f)(x)-(g \circ f)(a)}{x-a} \quad = \quad \Big(\lim_{x\to a} (\tilde{g} \circ f)(x)\Big)\Big(\lim_{x\to a} \frac{f(x)-f(a)}{x-a}\Big) \quad =$$

$$(\tilde{g} \circ f)(a)f'(a) \quad = \quad g'(b)f'(a) \quad = \quad g'(f(a))f'(a) \qquad \square$$

Wir bestimmen die Ableitung der Funktion $h : \mathbb{R} \to \mathbb{R}$, $x \mapsto (x^2+5)^7$ mit Hilfe der Kettenregel. Seien $f, g : \mathbb{R} \to \mathbb{R}$ gegeben durch $f(x) = x^2 + 5$ und $g(x) = x^7$ für alle $x \in \mathbb{R}$. Dann ist $h = g \circ f$, und die Ableitungen von f und g sind gegeben durch $f'(x) = 2x$ und $g'(x) = 7x^6$ für $x \in \mathbb{R}$. Die Kettenregel liefert also

$$h'(x) \quad = \quad (g \circ f)'(x) \quad = \quad f'(x)g'(f(x)) \quad = \quad 2x \cdot 7(x^2+5)^6 \quad = \quad 14x(x^2+5)^6.$$

Ebenso kann mit der Kettenregel die Ableitung der Exponentialfunktion $\exp_b : \mathbb{R} \to \mathbb{R}$, $x \mapsto b^x$ zu einer beliebigen Basis $b > 1$ bestimmt werden. Weil nach Definition $b^x = \exp(x \ln(b))$ gilt, die Exponentialfunktion ihre eigene Ableitung ist und die innere Funktion $x \mapsto x \ln(b)$ die Ableitung $x \mapsto \ln(b)$ hat, erhalten wir mir mit der Kettenregel $\exp_b'(x) = \ln(b)\exp'(x \ln(b)) = \ln(b)\exp_b(x)$ für alle $x \in \mathbb{R}$.

(7.9) Satz *(Umkehrregel)*
Sei $I \subseteq \mathbb{R}$ ein (endliches oder unendliches) offenes Intervall, $f : I \to \mathbb{R}$ stetig und streng monoton wachsend. Sei $J = f(I)$ und $g = f^{-1}$ die Umkehrfunktion von f. Ist f an einer Stelle $a \in I$ differenzierbar und gilt $f'(a) \neq 0$, dann ist auch g an der Stelle $b = f(a)$ differenzierbar, und es gilt

$$g'(b) \quad = \quad \frac{1}{f'(a)} \quad = \quad \frac{1}{f'(g(b))}.$$

Beweis: Sei $(y_n)_{n\in\mathbb{N}}$ eine Folge in $J \setminus \{b\}$, die gegen b konvergiert. Definieren wir $x_n = g(y_n)$ für alle $n \in \mathbb{N}$, so ist $(x_n)_{n\in\mathbb{N}}$ eine Folge in $I \setminus \{a\}$. Nach Satz (5.19) ist mit der Funktion f auch die Funktion g stetig, also gilt $\lim_n x_n = \lim_n g(y_n) = g(b) = f^{-1}(f(a)) = a$. Auf Grund der Differenzierbarkeit von f im Punkt a und wegen $f'(a) \neq 0$ gilt

$$\lim_{n\to\infty} \frac{g(y_n)-g(b)}{y_n-b} \quad = \quad \lim_{n\to\infty} \frac{x_n-a}{f(x_n)-f(a)} \quad = \quad \frac{1}{\lim\limits_{n\to\infty} \frac{f(x_n)-f(a)}{x_n-a}}$$

$$= \quad \frac{1}{f'(a)} \quad = \quad \frac{1}{f'(g(b))}.$$

Damit ist $g'(b) = \lim\limits_{y\to b} \dfrac{g(y)-g(b)}{y-b} = \dfrac{1}{f'(g(b))}$ bewiesen. □

(7.10) Folgerung Sei $k \in \mathbb{N}$ und $g_k : \mathbb{R}^+ \to \mathbb{R}^+$, $x \mapsto \sqrt[k]{x}$. Dann ist g_k auf $\mathbb{R}^+$ differenzierbar, und es gilt $g_k'(x) = \frac{1}{k}(\sqrt[k]{x})^{1-k}$ für alle $x \in \mathbb{R}^+$.

Beweis: Wie wir bereits in § 7 festgestellt haben, ist g_k die Umkehrfunktion der Funktion $f_k : \mathbb{R}^+ \to \mathbb{R}^+$ gegeben durch $x \mapsto x^k$. Für alle $x \in \mathbb{R}^+$ gilt $f_k'(x) = kx^{k-1}$, die Umkehrregel liefert also

$$g_k'(x) \quad = \quad \frac{1}{f_k'(g_k(x))} \quad = \quad \frac{1}{f_k'(\sqrt[k]{x})} \quad = \quad \frac{1}{k(\sqrt[k]{x})^{k-1}} \quad = \quad \tfrac{1}{k}(\sqrt[k]{x})^{1-k}.$$ □

(7.11) Folgerung Die Ableitung der Logarithmusfunktion $\ln : \mathbb{R}^+ \to \mathbb{R}$ ist gegeben durch $\ln'(x) = \frac{1}{x}$ für alle $x \in \mathbb{R}^+$.

Beweis: Die Umkehrfunktion von $f(x) = \exp(x)$ ist die Logarithmusfunktion $g(x) = \ln(x)$. Durch Anwendung der Umkehrregel erhalten wir für alle $x \in \mathbb{R}^+$ jeweils

$$g'(x) \quad = \quad \frac{1}{f'(g(x))} \quad = \quad \frac{1}{\exp'(\ln(x))} \quad = \quad \frac{1}{\exp(\ln(x))} \quad = \quad \frac{1}{x}$$ □

§ 8. *Mittelwertsatz und Extremwertbestimmung*

Inhaltsübersicht

In diesem Abschnitt leiten wir die aus der Schulmathematik bekannten Regeln zur Bestimmung des Monotonieverhaltens einer Funktion und zur Bestimmung lokaler Extremstellen (Minima und Maxima) her, was dort häufig als Bestandteil der sog. „Kurvendiskussion“ behandelt wird. Grundlage für ersteres ist der ***Mittelwertsatz***, welcher besagt, dass die Ableitung einer differenzierbaren Funktion auf einem endlichen offenen Intervall mindestens in einem Punkt den Wert der durchschnittlichen Steigung animmt. Während der Zwischwertsatz aus § 12 also eine Aussage über die Werte der Funktion f selbst macht, liefert der Mittelwertsatz eine Aussage über die Werte der Ableitung f'. Ein notwendiges Kriterium für Extrema ergibt sich, wie wir zu Anfang des Kapitels sehen werden, direkt aus der Definition der Ableitung. Um auch ein hinreichendes Kriterium zu erhalten, betrachten wir im letzten Teil dieses Kapitels die höheren Ableitungen einer Funktion.

Wichtige Begriffe und Sätze

- lokales und globales Extremum (Minimum und Maximum) einer Funktion
- höhere Ableitungen, Taylorpolynome
- notwendiges Kriterium und hinreichendes Kriterium für Extrema
- Satz von Rolle und Mittelwertsatz der Differentialrechnung
- l'Hospitalsche Regel zur Bestimmung von Grenzwerten

(8.1) Definition Sei $D \subseteq \mathbb{R}$, $f : D \to \mathbb{R}$ eine Funktion und $a \in D$ ein beliebiger Punkt. Man nennt a ein

(i) ***lokales Maximum*** von f, wenn ein $\varepsilon \in \mathbb{R}^+$ existiert, so dass $f(a) \geq f(x)$ für alle $x \in D$ mit $|x - a| < \varepsilon$ gilt, und ein

(ii) ***globales Maximum*** von f, wenn $f(a) \geq f(x)$ für alle $x \in D$ erfüllt ist.

Entsprechend sind die Begriffe *lokales* bzw. *globales Minimum* definiert.

Das Wort ***Extremum*** ist der gemeinsame Oberbegriff für Minimum und Maximum. Man kann somit auch von lokalen oder globalen Extremstellen sprechen. Offenbar ist jedes globale Extremum auch ein lokales, während die Umkehrung im Allgemeinen falsch ist.

(8.2) Proposition *(notwendiges Kriterium für Extrema)*

Sei $D \subseteq \mathbb{R}$, $f : D \to \mathbb{R}$ eine Funktion und $a \in D$ ein lokales Extremum von f. Ist die Funktion f in a differenzierbar, dann gilt $f'(a) = 0$.

Beweis: Nach Voraussetzung ist a ein innerer Punkt von D, es gibt also ein $\varepsilon \in \mathbb{R}^+$ mit der Eigenschaft, dass $]a-\varepsilon, a+\varepsilon[$ in D enthalten ist. Wir gehen davon aus, dass es sich bei a um ein lokales Maximum von f handelt; im Fall eines Minimums läuft der Beweis völlig analog. Nach eventueller Verkleinerung von ε ist dann $f(a) \geq f(x)$ für alle $x \in D$ mit $|x-a| < \varepsilon$ erfüllt. Sei nun $N \in \mathbb{N}$ so groß gewählt, dass $\frac{1}{N} < \varepsilon$ gilt. Wir definieren Folgen $(x_n)_{n\geq N}$ und $(y_n)_{n\geq N}$ durch $x_n = a - \frac{1}{n}$ und $y_n = a + \frac{1}{n}$ für alle $n \in \mathbb{N}$. Wegen $f(x_n) - f(a) \leq 0$ und $x_n - a = -\frac{1}{n} < 0$ für alle $n \geq N$ gilt

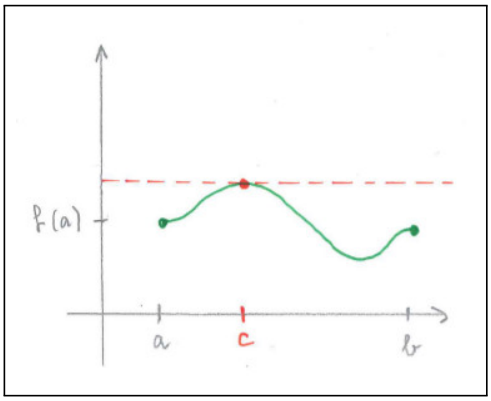
$$f'(a) \quad = \quad \lim_{n\to\infty} \frac{f(x_n)-f(a)}{x_n - a} \quad \geq \quad 0.$$

Andererseits gilt $f(y_n) - f(a) \leq 0$ sowie $y_n - a = \frac{1}{n} > 0$ für alle $n \geq N$ und somit

$$f'(a) \quad = \quad \lim_{n\to\infty} \frac{f(y_n)-f(a)}{y_n - a} \quad \leq \quad 0.$$

Insgesamt ist also nur $f'(a) = 0$ möglich. □

(8.3) Satz *(Satz von Rolle)*
Seien $a, b \in \mathbb{R}$ mit $a < b$. Ferner sei $f : [a,b] \to \mathbb{R}$ eine stetige Funktion mit $f(a) = f(b)$, die darüber hinaus auf $]a,b[$ differenzierbar ist. Dann gibt es ein $c \in]a,b[$ mit $f'(c) = 0$.

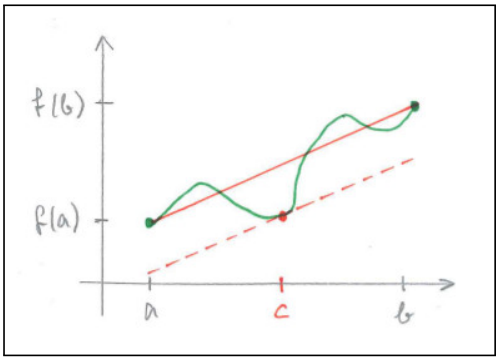


Beweis: Wenn f auf $[a,b]$ konstant ist, dann ist die Aussage trivial, denn in diesem Fall gilt $f'(x) = 0$ für alle $x \in]a,b[$. Nehmen wir nun an, dass ein $x \in]a,b[$ mit $f(x) > f(a)$ existiert. Nach dem Maximumsprinzip, Satz (5.17) nimmt f sein Maximum in einem Punkt $c \in [a,b]$ an, wobei $c \in \{a,b\}$ allerdings ausgeschlossen ist, denn dort ist der Funktionswert jedenfalls nicht maximal. Dieses c ist dann nach Definition auch ein lokales Maximum, und das notwendige Kriterium für Extrema, Proposition (8.2), liefert $f'(c) = 0$.

Wenn f auf $[a,b]$ nicht konstant ist und auch kein $x \in]a,b[$ mit $f(x) > f(a)$ existiert, dann gibt es ein $x \in]a,b[$ mit $f(x) < f(a)$. Der Beweis verläuft dann analog zum vorherigen Fall, an Stelle des globalen Maximums betrachtet man das globale Minimum. □

(8.4) Satz *(Mittelwertsatz der Differentialrechnung)*
Seien $a, b \in \mathbb{R}$ mit $a < b$. Ferner sei $f : [a,b] \to \mathbb{R}$ stetig und auf $]a,b[$ differenzierbar. Dann gibt es ein $c \in]a,b[$ mit

$$f'(c) \quad = \quad \frac{f(b)-f(a)}{b-a}.$$



Beweis: Wir wenden den Satz von Rolle auf die Funktion $g(x) = f(x) - \frac{f(b)-f(a)}{b-a}(x-a)$ an. Es gilt $g(a) = f(a) = g(b)$, außerdem ist die Funktion g stetig auf $[a,b]$ und differenzierbar auf $]a,b[$. Auf Grund des Satzes von Rolle finden wir ein $c \in]a,b[$ mit $g'(c) = 0$, und es folgt

$$f'(c) \quad = \quad g'(c) + \frac{f(b)-f(a)}{b-a} \quad = \quad \frac{f(b)-f(a)}{b-a}.$$ □

(8.5) Folgerung Seien $a, b \in \mathbb{R}$ mit $a < b$. Ferner sei $f : [a,b] \to \mathbb{R}$ stetig und es gelte $f'(x) = 0$ für alle $x \in]a,b[$. Dann ist f auf $[a,b]$ konstant.

Beweis: Angenommen, es gibt ein $c \in]a,b]$ mit $f(c) \neq f(a)$. Nach dem Mittelwertsatz finden wir dann ein $x \in]a,c[$ mit

$$f'(c) \quad = \quad \frac{f(c)-f(a)}{c-a} \quad \neq \quad 0$$

im Widerspruch zur Voraussetzung. □

Wir haben oben bereits ein notwendiges Kriterium für das Auftreten lokaler Extrema formuliert, nämlich das Verschwinden der ersten Ableitung. Wie bereits aus dem Schulunterricht bekannt, ist dieses Kriterium aber nicht hinreichend. Für die Funktion $f(x) = x^3$ gilt beispielsweise $f'(0) = 0$, ohne dass die Funktion an dieser Stelle ein lokales Extremum besitzt.

(8.6) Satz *(Ableitung und Monotonieverhalten)*
Seien $a, b \in \mathbb{R}$ mit $a < b$. Sei $f : [a, b] \to \mathbb{R}$ eine stetige, auf $]a, b[$ differenzierbare Funktion.

(i) Gilt $f'(x) \geq 0$ für alle $x \in]a, b[$, dann ist f auf $[a, b]$ monoton wachsend.

(ii) Gilt $f'(x) > 0$ für alle $x \in]a, b[$, dann ist f auf $[a, b]$ streng monoton wachsend.

(iii) Gilt $f'(x) \leq 0$ für alle $x \in]a, b[$, dann ist f auf $[a, b]$ monoton fallend.

(iv) Gilt $f'(x) < 0$ für alle $x \in]a, b[$, dann ist f auf $[a, b]$ streng monoton fallend.

Beweis: zu (i) Angenommen, es gilt $f'(x) \geq 0$ für alle $x \in]a, b[$, aber f ist auf I nicht monoton wachsend. Dann gibt es $x, y \in [a, b]$ mit $x < y$ und $f(x) > f(y)$. Nach dem Mittelwertsatz, Satz (8.4), gibt es ein $c \in]x, y[\subseteq \tilde{I}$ mit

$$f'(c) \quad = \quad \frac{f(y) - f(x)}{y - x} \quad < \quad 0$$

im Widerspruch zur Voraussetzung.

zu (ii) Nehmen wir an, es gilt $f'(x) > 0$ für alle $x \in]a, b[$, aber f ist auf $[a, b]$ nicht streng monoton wachsend. Dann finden wir $x, y \in [a, b]$ mit $x < y$ und $f(x) \geq f(y)$. Auf Grund des Mittelwertsatzes existiert ein $c \in]x, y[$ mit

$$f'(c) \quad = \quad \frac{f(y) - f(x)}{y - x} \quad \leq \quad 0 \quad ,$$

was erneut der Voraussetzung widerspricht. Der Beweis der Aussagen (iii) und (iv) verläuft analog. □

Als nächstes beweisen wir zwei wichtige Regeln zur Bestimmung von Funktionsgrenzwerten, bei denen die Ableitung eine Rolle spielt. Zur Vorbereitung benötigen wir eine verallgemeinerte Version des Mittelwertsatzes.

(8.7) Satz Seien $a, b \in \mathbb{R}$ mit $a < b$ und $f, g : [a, b] \to \mathbb{R}$ stetige, auf $]a, b[$ differenzierbare Funktionen, wobei wir $g'(x) \neq 0$ für alle $x \in]a, b[$ voraussetzen. Dann gibt es ein $c \in]a, b[$ mit

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

Beweis: Auf Grund unserer Voraussetzungen muss $g(a) \neq g(b)$ gelten, weil ansonsten nach dem Satz von Rolle ein $x \in]a, b[$ mit $g'(x) = 0$ existieren würde. Wir betrachten nun auf $[a, b]$ die Funktion

$$h(x) \quad = \quad f(x) - g(x)\frac{f(b) - f(a)}{g(b) - g(a)}.$$

Diese erfüllt die Voraussetzungen des Satzes von Rolle, denn es gilt

$$\begin{aligned} h(a) &= f(a)-g(a)\frac{f(b)-f(a)}{g(b)-g(a)} = \frac{f(a)g(b)-f(a)g(a)-f(b)g(a)+f(a)g(a)}{g(b)-g(a)} \\ &= \frac{f(a)g(b)-f(b)g(a)}{g(b)-g(a)} = \frac{f(b)g(b)-f(b)g(a)-f(b)g(b)+f(a)g(b)}{g(b)-g(a)} \\ &= f(b)-g(b)\frac{f(b)-f(a)}{g(b)-g(a)} = h(b). \end{aligned}$$

Es gibt also ein $x \in]a,b[$ mit

$$h'(x)=0 \quad\Leftrightarrow\quad f'(x)-g'(x)\frac{f(b)-f(a)}{g(b)-g(a)}=0 \quad\Leftrightarrow\quad \frac{f'(x)}{g'(x)}=\frac{f(b)-f(a)}{g(b)-g(a)} \qquad \square$$

(8.8) Satz *(l'Hospital'sche Regeln)*
Sei $I \subseteq \mathbb{R}$ ein Intervall, und sei $a \in \bar{\mathbb{R}}$ ein Berührpunkt von I. Seien $f, g : I \to \mathbb{R}$ differenzierbare Funktionen mit den Eigenschaften $g'(x) \neq 0$ für alle $x \in I$ und

$$\text{entweder} \quad \lim_{x\to a} f(x) = \lim_{x\to a} g(x) = 0 \quad \text{oder} \quad \lim_{x\to a} f(x), \lim_{x\to a} g(x) \in \{\pm\infty\}.$$

Dann gilt

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

Dies soll bedeuten: Existiert der Grenzwert $c \in \mathbb{R}$ auf der rechten Seite der Gleichung, dann existiert auch der Grenzwert auf der linken Seite und hat denselben Wert c.

Beweis: Sei $\varepsilon \in \mathbb{R}^+$ vorgegeben und c der Grenzwert von $f'(x)/g'(x)$ für $x \to a$. Nach Satz (5.13) genügt es zu zeigen, dass eine Umgebung U von a existiert, so dass $|f(x)/g(x) - c| < \varepsilon$ für alle $x \in I \cap U$ gilt. Auf Grund der Voraussetzung gibt es eine Umgebung V von a mit $|f'(x)/g'(x) - c| < \frac{1}{2}\varepsilon$ für alle $x \in I \cap V$, wobei dieses V als Intervall gewählt werden kann. Nach Ersetzung von I durch $I \cap V$ dürfen wir davon ausgehen, dass die Ungleichung für alle $x \in I$ erfüllt ist.

Betrachten wir nun zunächst den Fall, dass die Grenzwerte von $f(x)$ und $g(x)$ für $x \to a$ beide gleich 0 sind. Die Funktion g besitzt im Intervall I höchstens eine Nullstelle. Wären nämlich $x, y \in I$ zwei verschiedene Nullstellen, dann gäbe es nach dem Satz (8.3) von Rolle eine Stelle z in $]x,y[\subseteq I$ mit $g'(z) = 0$, im Widerspruch zur Voraussetzung. Nach weiterer Verkleinerung von I können wir o.B.d.A. $g(x) \neq 0$ für alle $x \in I$ annehmen. Sei nun $x \in I$ beliebig gewählt. Für alle $y \in I$ mit $y \neq x$ gilt, ebenfalls auf Grund des Satzes von Rolle, jeweils $g(x) - g(y) \neq 0$. Wegen $\lim_{y\to a} f(y) = \lim_{y\to a} g(y) = 0$ ist

$$\lim_{y\to a} \frac{f(x)-f(y)}{g(x)-g(y)} = \frac{f(x)}{g(x)}.$$

Wählen wir also $y \in I$ hinreichend nahe bei a, dann gilt

$$\left|\frac{f(x)}{g(x)} - \frac{f(x)-f(y)}{g(x)-g(y)}\right| < \tfrac{1}{2}\varepsilon.$$

Nach dem verallgemeinerten Mittelwertsatz, Satz (8.7), existiert ein z in $]x,y[$ oder $]y,x[$ mit

$$\frac{f'(z)}{g'(z)} \quad = \quad \frac{f(x)-f(y)}{g(x)-g(y)}.$$

Insgesamt erhalten wir damit

$$\left|\frac{f(x)}{g(x)}-c\right| \quad = \quad \left|\frac{f(x)}{g(x)}-\frac{f(x)-f(y)}{g(x)-g(y)}+\frac{f'(z)}{g'(z)}-c\right| \quad \leq \quad \left|\frac{f(x)}{g(x)}-\frac{f(x)-f(y)}{g(x)-g(y)}\right|+\left|\frac{f'(z)}{g'(z)}-c\right| \quad < \quad \tfrac{1}{2}\varepsilon+\tfrac{1}{2}\varepsilon=\varepsilon.$$

Die gewünschte Ungleichung ist also für alle $x \in I$ erfüllt. Setzen wir $U = V$, dann hat die Umgebung V hat also die gewünschte Eigenschaft.

Nun betrachten wir den Fall, dass die Grenzwerte von $f(x)$ und $g(x)$ für $x \to a$ beide in $\{\pm\infty\}$ liegen. Der wesentliche Unterschied zum vorherigen Fall besteht darin, dass wir die Differenz der Ausdrücke $\frac{f(x)-f(y)}{g(x)-g(y)}$ und $\frac{f(x)}{g(x)}$ nicht direkt betrachten, sondern zunächst deren Verhältnis. Nach Verkleinerung von I können wir $f(x), f(y) \neq 0$ für alle $x \in I$ annehmen, auf Grund des Grenzwerts von f. Für alle $x, y \in I$ mit $x \neq y$ gilt (wieder auf Grund des Satzes von Rolle) $g(x)-g(y) \neq 0$. Für jedes $y \in I$ definieren wir eine Hilfsfunktion h_y auf $I \setminus \{y\}$ durch

$$\frac{\frac{f(x)-f(y)}{g(x)-g(y)}}{\frac{f(x)}{g(x)}} \quad = \quad \frac{f(x)-f(y)}{g(x)-g(y)} \cdot \frac{f(x)^{-1}}{g(x)^{-1}} \quad = \quad \frac{1-\frac{f(y)}{f(x)}}{1-\frac{g(y)}{g(x)}} \quad = \quad h_y(x)$$

Weil die Grenzwerte von $f(x)$ und $g(x)$ für $x \to a$ in $\{\pm\infty\}$ liegen, ist der Grenzwert der Funktion $h_y(x)$ für $x \to a$ gleich 1. Sei nun $y \in I$ beliebig gewählt. Nach dem verallgemeinerten Mittelwertsatz existiert für jedes $x \in I$ mit $x \neq y$ ein z zwischen x und y mit $f'(z)/g'(z) = (f(x)-f(y))/(g(x)-g(y))$. Wegen $f'(z)/g'(z) \in \left]c-\frac{1}{2}\varepsilon, c+\frac{1}{2}\varepsilon\right[$ ist der Betrag von $(f(x)-f(y))/(g(x)-g(y))$ für alle $x \in I$ durch eine Konstante κ beschränkt. Wegen $\lim_{x\to a} h_y(x)^{-1} = 1$ gibt es eine Umgebung $U \subseteq V$ von a mit $|h_y(x)^{-1}-1| < \frac{\varepsilon}{2\kappa}$ für alle $x \in I \cap U$. Für diese x folgt

$$\left|\frac{f(x)}{g(x)}-\frac{f(x)-f(y)}{g(x)-g(y)}\right| \quad = \quad \left|\frac{f(x)-f(y)}{g(x)-g(y)}\right| \cdot |h_y(x)^{-1}-1| \quad < \quad \kappa \cdot \frac{\varepsilon}{2\kappa} \quad = \quad \tfrac{1}{2}\varepsilon.$$

Wie im vorherigen Fall kann nun mit Hilfe des verallgemeinerten Mittelwertsatzes gezeigt werden, dass für alle $x \in I \cap U$ jeweils $|f(x)/g(x)-c| < \varepsilon$ gilt. □

Als Anwendungsbeispiel für die l'Hospitalsche Regel berechnen wir den Grenzwert

$$\lim_{x\to 0} \frac{x}{1-e^x}$$

durch Betrachtung der Funktionen $f(x) = x$ und $g(x) = 1-e^x$ auf dem Intervall $I =]0,1[$ mit dem Berührpunkt 0. Beide Funktionen sind differenzierbar, und es gilt $g'(x) = -e^x \neq 0$ für alle $x \in D$. Auf der Stetigkeit von f und g im Nullpunkt gilt

$$\lim_{x\to 0} f(x) = f(0) = 0 \quad \text{und} \quad \lim_{x\to 0} g(x) = g(0) = 1-e^0 = 1-1 = 0.$$

Mit der Ableitung $f'(x) = 1$ und der l'Hospital'schen Regel erhalten wir somit

$$\lim_{x\to 0} \frac{x}{1-e^x} \quad = \quad \lim_{x\to 0} \frac{1}{-e^x} \quad = \quad \frac{1}{-1} \quad = \quad -1.$$

Als weitere Anwendung der l'Hospitalschen Regeln beweisen wir

(8.9) Satz Für jede Polynomfunktion p gilt

$$\lim_{x\to+\infty} \frac{p(x)}{e^x} \quad = \quad 0.$$

Beweis: Auf Grund der Rechenregeln für Grenzwerte in Proposition (5.6) genügt es zu zeigen, dass der Grenzwert von $x^n/p(x)$ für $x \to +\infty$ jeweils gleich null ist, für jedes $n \in \mathbb{N}_0$. Wir beweisen die Aussage durch vollständige Induktion über $n \in \mathbb{N}_0$ Vorweg bemerken wir, dass für jede Funktion f der Grenzwert von $f(x)$ für $x \to -\infty$ (sofern er existiert) übereinstimmt mit dem Grenzwert für $f(-x)$ für $x \to +\infty$, wie man durch Betrachtung einer $(x_n)_{n\in\mathbb{N}}$ mit $\lim_n x_n = +\infty$ unmittelbar überprüft. Für $n = 0$ folgt die Aussage nun direkt aus Teil (ii) von Satz (6.10), denn es gilt

$$\lim_{x\to+\infty} \frac{x^0}{e^x} \quad = \quad \lim_{x\to+\infty} \frac{1}{e^x} \quad = \quad \lim_{x\to+\infty} e^{-x} \quad = \quad \lim_{x\to-\infty} e^x \quad = \quad 0.$$

Setzen wir nun die Aussage für n voraus. Auf Grund von Proposition (5.4) und Satz (6.10) gilt

$$\lim_{x\to+\infty} x^{n+1} = +\infty \qquad \text{und} \qquad \lim_{x\to+\infty} e^x = +\infty.$$

Außerdem ist $g'(x) = e^x \neq 0$ für alle $x \in \mathbb{R}$. Zusammen mit der Induktionsvoraussetzung zeigt dies, dass die Voraussetzungen der l'Hospital'schen Regel erfüllt sind, und es folgt

$$\lim_{x\to+\infty} \frac{x^{n+1}}{e^x} \quad = \quad \lim_{x\to+\infty} \frac{nx^n}{e^x} \quad = \quad n \lim_{x\to+\infty} \frac{x^n}{e^x} \quad = \quad n \cdot 0 \quad = \quad 0. \qquad \square$$

Um die Untersuchung der Extrema von Funktionen weiterzuentwickeln, benötigen wir höhere Ableitungen. Sei $D \subseteq \mathbb{R}$ und $f : D \to \mathbb{R}$ eine Funktion. Definieren wir die Menge $D_1 \subseteq \mathbb{R}$ durch

$$D_1 \quad = \quad \{x \in D \mid f \text{ ist in } x \text{ differenzierbar } \} \quad ,$$

so erhalten wir durch die Ableitung f' eine Funktion $f' : D_1 \to \mathbb{R}$. Wir bezeichnen f als ***zweimal differenzierbar*** an der Stelle $x \in D$, wenn x in D_1 liegt und die Funktion f' in x differenzierbar ist. Man bezeichnet dann $f''(x) = (f')'(x)$ als die ***zweite Ableitung*** von f an der Stelle x.

Mit demselben Ansatz lassen sich auch höhere Ableitungen definieren. Sei $D_0 = D$ und $f^{(0)} = f$. Für jedes $n \in \mathbb{N}_0$ definieren wir rekursiv

$$D_{n+1} \quad = \quad \{x \in D_n \mid f^{(n)} \text{ ist in } x \text{ differenzierbar } \}$$

und $f^{(n+1)}(x) = (f^{(n)})'(x)$ für alle $x \in D_{n+1}$. Man sagt, die Funktion f ist an der Stelle x ***n-mal differenzierbar***, wenn x in D_n liegt. Gilt $D_n = D$, dann bezeichnet man f insgesamt als n-mal differenzierbare Funktion. Nach Definition gilt $D = D_0 \supseteq D_1 \supseteq \ldots \supseteq D_n$ für alle $n \in \mathbb{N}$.

Liegt x in D_n und ist $f^{(n)}$ in x stetig, so bezeichnet man f an der Stelle x als ***n-mal stetig differenzierbar***. Ist $D_n = D$ und $f^{(n)}$ auf ganz D stetig, so ist f insgesamt eine n-mal stetig differenzierbare Funktion auf D. Die Menge dieser Funktionen D wird mit $\mathscr{C}^n(D)$ bezeichnet. Insbesondere steht $\mathscr{C}^0(D)$ für die Menge der stetigen Funktionen auf D.

Gilt $D_n = D$ für alle $n \in \mathbb{N}$, so ist f beliebig oft differenzierbar und damit auch n-mal stetig differenzierbar für jedes $n \in \mathbb{N}$, denn jede differenzierbare Funktion ist nach Satz (7.3) stetig. Die Menge der Funktionen mit dieser Eigenschaft wird durch $\mathscr{C}^\infty(D)$ bezeichnet.

Wir haben in Proposition (7.2) die folgende äquivalente Charakterisierung differenzierbarer Funktionen angegeben: Eine Funktion $f : D \to \mathbb{R}$ ist genau dann differenzierbar an einer Stelle $a \in D$, wenn ein $c \in \mathbb{R}$ und eine Funktion $\psi : D \to \mathbb{R}$ mit

$$f(x) = f(a) + c(x-a) + \psi(x) \text{ für alle } x \in D \qquad \text{und} \qquad \lim_{x \to a} \frac{\psi(x)}{x-a} = 0$$

existiert. In diesem Fall gilt $c = f'(a)$. Die Funktion f wird an der Stelle a also durch eine Polynomfunktion ersten Grades approximiert, und zwar durch

$$\tau_1(f,a)(x) \quad = \quad f(a) + f'(a)(x-a).$$

(8.10) Definition Ist $f : D \to \mathbb{R}$ eine differenzierbare Funktion und $a \in D$, dann nennt man

$$\tau_1(f,a) \quad = \quad f(a) + f'(a)(x-a)$$

das ***Taylorpolynom erster Ordnung*** von f an der Stelle a.

Das Taylorpolynom ist dadurch gekennzeichnet, dass

$$\tau_1(f,a)(a) = f(a) \qquad \text{und} \qquad \tau_1(f,a)'(a) = f'(a) \qquad \text{erfüllt ist.}$$

Unser Ziel besteht nun darin, die zweimal differenzierbaren Funktionen auf ähnliche Weise zu charakterisieren. Dazu suchen wir eine Polynomfunktion zweiten Grades $p : \mathbb{R} \to \mathbb{R}$ mit $p^{(k)}(a) = f^{(k)}(a)$ für $k = 0, 1, 2$. Setzen wir p in der Form $p(x) = u(x-a)^2 + v(x-a) + w$ mit $u, v, w \in \mathbb{R}$ an, so gilt $p'(x) = 2u(x-a) + v$ und $p''(x) = 2u$, also $p''(a) = 2u$, $p'(a) = v$ und $p(a) = w$. Der Ansatz, dass die Ableitungen von p und f bis zum Grad 2 übereinstimmen, liefert $u = \frac{1}{2}f''(a)$, $v = f'(a)$ und $w = f(a)$. Dies führt uns nur folgenden Definition.

(8.11) Definition Ist $f : D \to \mathbb{R}$ eine zweimal differenzierbare Funktion und $a \in D$, dann nennt man

$$\tau_2(f,a) \quad = \quad f(a) + f'(a)(x-a) + \tfrac{1}{2}f''(a)(x-a)^2$$

das ***Taylorpolynom zweiter Ordnung*** von f an der Stelle a.

Wir werden uns nun anschauen, wie dieses Polynom mit der zweifachen Differenzierbarkeit einer Funktion zusammenhängt. Zur Vorbereitung beweisen wir

(8.12) Lemma Seien $a, b, c \in \mathbb{R}$ mit $a < c < b$ und $f :]a, b[\setminus \{c\} \to \mathbb{R}$ eine beliebige Funktion. Wenn die beiden Grenzwerte $\lim_{x \to c}(f|_{]a,c[})(x)$ und $\lim_{x \to c}(f|_{]c,b[})(x)$ existieren und übereinstimmen, dann existiert auch der Grenzwert $\lim\limits_{x \to c} f(x)$, und es gilt $\lim\limits_{x \to c} f(x) = \lim\limits_{x \to c}(f|_{]a,c[})(x) = \lim\limits_{x \to c}(f|_{]c,b[})(x)$.

Beweis: Sei u der gemeinsame Grenzwert der Funktionen $g = f|_{]a,c[}$ und $h = f|_{]c,b[}$ für $x \to c$. Erweitern wir den Definitionsbereich von g durch $g(c) = u$ auf $]a,c]$ und den von h auf $[c,b[$ durch $h(c) = u$, dann sind die Funktionen g und h beide im Punkt c stetig. Sei nun $(x_n)_{n\in\mathbb{N}}$ eine Folge in $]a,b[\setminus \{c\}$, die gegen c konvergiert. Zu zeigen ist $\lim_n f(x_n) = u$; sei dazu $\varepsilon \in \mathbb{R}^+$ vorgegeben. Nach dem ε-δ-Kriterium und auf Grund der Stetigkeit von g in c gibt es ein $\delta_1 \in \mathbb{R}^+$ mit der Eigenschaft, dass die Implikation $|x-c| < \delta_1 \Rightarrow |g(x)-u| < \varepsilon$ für alle $x \in]a,c[$ erfüllt ist. Ebenso finden wir ein $\delta_2 \in \mathbb{R}^+$ mit $|x-c| < \delta_2 \Rightarrow |h(x)-u| < \varepsilon$ für alle $x \in]c,b[$. Sei nun $\delta = \min\{\delta_1, \delta_2\}$ und $N \in \mathbb{N}$ so groß gewählt, dass $|x_n - c| < \delta$ für alle $n \geq N$ erfüllt ist. Für diese n gilt im Fall $x_n < c$ die Abschätzung

$$|f(x_n)-u| \quad = \quad |g(x_n)-u| \quad < \quad \varepsilon$$

im Fall $x_n > c$ ebenso $|f(x_n)-u| = |h(x_n)-u| < \varepsilon$. Damit ist $\lim_n f(x_n) = u$ nachgewiesen. □

Man bezeichnet $\lim\limits_{x\to c}(f|_{]a,c[})(x)$ als den ***linksseitigen*** und $\lim\limits_{x\to c}(f|_{]c,b[})(x)$ als den ***rechtsseitigen*** Grenzwert der Funktion f, während $\lim\limits_{x\to c} f(x)$ der ***beidseitige*** Grenzwert genannt wird. Aus dem Lemma folgt, dass die l'Hospitalsche Regel (8.8) auch für beidseitige Grenzwerte gültig ist.

(8.13) Satz Sei $f : D \to \mathbb{R}$ eine Funktion und $a \in D$ ein Punkt, in dem f zweimal differenzierbar ist. Dann gibt es eine Funktion $\psi : D \to \mathbb{R}$ mit

$$f(x) \quad = \quad \tau_2(f,a)(x) + \psi(x) \quad \text{für alle } x \in D \qquad \text{und} \qquad \lim_{x\to a} \frac{\psi(x)}{(x-a)^2} = 0.$$

Beweis: Für alle $x \in D$ definieren wir $\psi(x) = f(x) - \tau_2(f,a)(x)$. Mit f und $\tau_2(f,a)$ ist auch ψ zweimal differenzierbar. Mit Hilfe der beidseitigen l'Hospitalschen Regel (s.o.) und der beiden Gleichungen $\psi'(a) = f'(a) - \tau_2(f,a)'(a) = 0$ und $\psi''(a) = f''(a) - \tau_2(f,a)''(a) = 0$ erhalten wir

$$\lim_{x\to a} \frac{\psi(x)}{(x-a)^2} \quad = \quad \lim_{x\to a} \frac{\psi'(x)}{2(x-a)} \quad = \quad \tfrac{1}{2} \lim_{x\to a} \frac{\psi'(x)-\psi'(a)}{x-a} \quad = \quad \tfrac{1}{2}\psi''(a) \quad = \quad 0$$

□

(8.14) Satz *(hinreichendes Kriterium für lokale Maxima)*
Sei $f : D \to \mathbb{R}$ eine differenzierbare Funktion und $c \in D$ ein Punkt, in dem f sogar zweimal differenzierbar ist. Gilt nun

$$f'(c) = 0 \qquad \text{und} \qquad f''(c) < 0 \quad ,$$

dann besitzt f in c ein lokales Maximum.

Beweis: O.B.d.A. können wir annehmen, dass $D =]a,b[$ ist, mit $a, b \in \mathbb{R}$ und $a < c < b$. Nach Satz (8.13) gibt es eine Funktion $\psi : D \to \mathbb{R}$ mit $f(x) = f(c) + \frac{1}{2}f''(c)(x-c)^2 + \psi(x)$ für alle $x \in D$ und

$$\lim_{x\to c} \frac{\psi(x)}{(x-c)^2} \quad = \quad 0.$$

Nach Satz (5.13) existiert somit für $\varepsilon = -\frac{1}{2}f''(c) \in \mathbb{R}^+$ ein $\delta \in \mathbb{R}^+$, so dass $|\psi(x)/(x-c)^2| < \varepsilon$ für alle $x \in D$ mit $|x-c| < \delta$ erfüllt ist. Insbesondere gilt für diese x also $\psi(x) < \varepsilon(x-c)^2$, und wir erhalten

$$\begin{aligned} f(x) &= f(c) + \tfrac{1}{2}f''(c)(x-c)^2 + \psi(x) < f(c) + \tfrac{1}{2}f''(c)(x-c)^2 + \varepsilon(x-c)^2 \\ &= f(c) + (\tfrac{1}{2}f''(c) + \varepsilon)(x-c)^2 = f(c) + 0 \cdot (x-c)^2 = f(c). \end{aligned}$$

Also liegt an der Stelle c tatsächlich ein lokales Maximum vor. □

Nach demselben Schema beweist man

(8.15) Satz *(hinreichendes Kriterium für lokale Minima)*
Sei $f : D \to \mathbb{R}$ eine differenzierbare Funktion und $c \in D$ ein Punkt, in dem f sogar zweimal differenzierbar ist. Gilt nun

$$f'(c) = 0 \qquad \text{und} \qquad f''(c) > 0 \quad ,$$

dann besitzt f in c ein lokales Minimum.

Die gerade angegebenen Kriterien sind jeweils hinreichend, aber nicht notwendig für die Existenz einer Extremstelle. So besitzt zum Beispiel die Funktion $f(x) = x^4$ in $c = 0$ ein lokales Minimum, aber statt $f''(c) > 0$ gilt $f''(c) = 0$.

Ist $f : D \to \mathbb{R}$ eine mindestens n-mal differenzierbare Funktion, so kann man ihr für jeden inneren Punkt $a \in D$ das ***n-te Taylorpolynom***

$$\tau_n(f,a)(x) = \sum_{k=0}^{n} \frac{1}{k!} f^{(k)}(a)(x-a)^k.$$

Diese Polynomfunktion ist dadurch ausgezeichnet, dass dessen Ableitungen mit denen von f bis zur Ordnung n übereinstimmt, also $f^{(k)}(a) = \tau_n(f,a)^{(k)}(a)$ für $0 \le k \le n$ erfüllt ist.

Ist f beliebig oft differenzierbar, dann kann sogar die sogenante ***Taylorreihe***

$$\tau(f,a)(x) = \sum_{n=0}^{\infty} \frac{1}{n!} f^{(n)}(a)(x-a)^n$$

gebildet werden. Wir werden in der Funktionentheorie eine große Klasse von Funktionen kennenlernen, die sich in der Umgebung jedes Punktes ihres Definitionsbereich durch ihre (absolut konvergente) Taylorreihe beschreiben lassen, die sog. ***holomorphen*** Funktionen. Zu ihnen zählen alle Polynomfunktionen und rationalen Funktionen, die Exponentialfunktion, Sinus-, Kosinus und Tangensfunktion und die zugehörigen Umkehrfunktionen, der Logarithmus und die Arkusfunktionen. Man überprüft zum Beispiel leicht, dass die Exponentialfunktion exp mit ihrer eigenen Taylorreihe $\tau(\exp, 0)$ übereinstimmt, und dass dasselbe auch für Sinus und Kosinus gilt.

§ 9. *Das Riemann-Integral*

Inhaltsübersicht

In der Schule definiert man das Integral einer nicht-negativen Funktion durch die Fläche unter dem Funktionsgraphen. Diese Definition können wir hier nicht unmittelbar verwenden, weil wir den Begriff des Flächeninhalts für krummlinig begrenzte Flächen (noch) nicht zur Verfügung haben. Statt dessen werden wir das Integral zunächst approximieren, in dem wir die Fläche unter dem Funktionsgraphen durch Rechtecke ausschöpfen bzw. überdecken. Man spricht in diesem Zusammenhang von ***Unter-*** bzw. ***Obersummen***. Wenn sich diese Unter- und Obersummen beliebig nahe kommen, nennt man die Funktion ***Riemann-integrierbar***. Das Supremum der Untersummen, das in diesem Fall mit dem Infimum der Obersummen zusammenfällt, wird das ***Riemann-Integral*** der Funktion genannt.

Aus der Schulmathematik ist auch bereits bekannt, dass ein Integral $\int_a^b f(x)\,dx$ durch ***Stammfunktionen*** berechnet werden kann: Es stimmt mit der Differenz $F(b)-F(a)$ überein, wobei F eine differenzierbare Funktion mit $F'=f$ bezeichnet. Dieser Zusammenhang ist als ***Hauptsatz der Differential- und Integralrechnung*** bekannt und wird im zweiten Teil dieses Kapitels bewiesen. Direkt aus dem Hauptsatz ergeben sich auch zwei wichtige, aus der Schule bekannte Integrationsregeln, nämlich die ***Substitutionsregel*** und die ***partielle Integration***. Die Anwendung dieser Regeln werden wir anhand einiger Beispiele nachvollziehen.

Wichtige Begriffe und Sätze

- Zerlegung eines Intervalls $[a,b]$
- Unter- und Obersumme einer beschränkten Funktion $f:[a,b]\to\mathbb{R}$ bezüglich eines Intervalls
- Ober- und Unterintegral von f
- Riemann-Integrierbarkeit einer Funktion, Riemann-Integral
- Hauptsatz der Differential- und Integralrechnung
- Substitutionsregel und partielle Integration

Im gesamten Abschnitt bezeichnet $[a,b]\subseteq\mathbb{R}$ ein endliches, abgeschlossenes Intervall positiver Länge, mit $a,b\in\mathbb{R}$ und $a<b$. Eine endliche (eventuell auch leere) Teilmenge $\mathscr{Z}=\{x_1,\dots,x_{n-1}\}\subseteq\,]a,b[$ bezeichnen wir als ***Zerlegung*** (oder Unterteilung) von $[a,b]$. Dabei gehen wir davon stets davon aus, dass die Zahlen x_k so angeordnet sind, dass $x_k<x_{k+1}$ für $1\le k<n-1$ erfüllt ist. Außerdem setzen wir stets $x_0=a$ und $x_n=b$. Für die Menge *aller* Zerlegungen von $[a,b]$ verwenden wir die Bezeichung $\mathbf{Z}(a,b)$.

(9.1) Definition Sei $f:[a,b]\to\mathbb{R}$ eine beschränkte Funktion, und sei $\mathscr{Z}=\{x_1,\dots,x_{n-1}\}$ eine Zerlegung von $[a,b]$. Sei $x_0=a$ und $x_n=b$. Für jedes k mit $0\le k<n$ definieren wir $c_k=\inf f([x_k,x_{k+1}])$ und $d_k=\sup f([x_k,x_{k+1}])$. Dann bezeichnet man

$$\mathscr{S}_f^-(\mathscr{Z})=\sum_{k=0}^{n-1}c_k(x_{k+1}-x_k)\qquad\text{bzw.}\qquad\mathscr{S}_f^+(\mathscr{Z})=\sum_{k=0}^{n-1}d_k(x_{k+1}-x_k)$$

als ***Unter-*** bzw. ***Obersumme*** von f bezüglich der Zerlegung $\mathscr{Z}$.

Wir illustrieren den Begriff der Unter- und Obersumme am Beispiel der Funktion $f : [0,1] \to \mathbb{R}$ gegeben durch $x \mapsto x^2$. Für jedes $n \in \mathbb{N}$ betrachten wir die Zerlegung

$$\mathscr{Z}_n \quad = \quad \{\tfrac{1}{n}, \tfrac{2}{n}, ..., \tfrac{n-1}{n}\}.$$

Zunächst berechnen wir die Untersumme. Weil f streng monoton wachsend ist, gilt $\inf f([\tfrac{k}{n}, \tfrac{k+1}{n}]) = f(\tfrac{k}{n}) = \left(\tfrac{k}{n}\right)^2$. Verwenden wir die Summenformel für die ersten n Quadratzahlen

$$\sum_{k=1}^{n} k^2 \quad = \quad \tfrac{1}{6} n(n+1)(2n+1) \quad ,$$

so erhalten wir für die Untersumme den Wert

$$\mathscr{S}_f^-(\mathscr{Z}_n) \quad = \quad \sum_{k=0}^{n-1} f(\tfrac{k}{n}) \quad = \quad \sum_{k=0}^{n-1} \left(\tfrac{k}{n}\right)^2 \tfrac{1}{n} \quad = \quad \tfrac{1}{n^3} \sum_{k=0}^{n-1} k^2 \quad = \quad \tfrac{1}{n^3} \cdot \tfrac{1}{6}(n-1)n(2n-1) \quad = \quad \tfrac{1}{6}(1-\tfrac{1}{n})(2-\tfrac{1}{n}).$$

Nun bestimmen wir noch die Obersumme. Auf Grund des streng monotonen Wachstums der Funktion f gilt hier $\sup f([\tfrac{k}{n}, \tfrac{k+1}{n}]) = f(\tfrac{k+1}{n}) = \left(\tfrac{k+1}{n}\right)^2$ und somit

$$\mathscr{S}_f^+(\mathscr{Z}_n) \quad = \quad \sum_{k=0}^{n-1} f(\tfrac{k+1}{n}) \quad = \quad \sum_{k=1}^{n} \left(\tfrac{k}{n}\right)^2 \tfrac{1}{n} \quad = \quad \tfrac{1}{n^3} \sum_{k=1}^{n} k^2 \quad = \quad \tfrac{1}{n^3} \cdot \tfrac{1}{6} n(n+1)(2n+1) \quad = \quad \tfrac{1}{6}(1+\tfrac{1}{n})(2+\tfrac{1}{n}).$$

Wir bemerken, dass sich für $n \to \infty$ der Wert der Untersumme von unten, der Wert der Obersumme von oben gegen den exakten Wert $\tfrac{1}{6} \cdot 1 \cdot 2 = \tfrac{1}{3}$ des Integrals konvergiert.

(9.2) Definition Sei $f : [a,b] \to \mathbb{R}$ eine beschränkte Funktion. Dann bezeichnet man

$$\int_{a\bigstar}^{b} f(x)\,dx \quad = \quad \sup\left\{\mathscr{S}_f^-(\mathscr{Z}) \;\middle|\; \mathscr{Z} \in \mathbf{Z}(a,b)\right\} \quad \text{bzw.}$$

$$\int_{a}^{b\bigstar} f(x)\,dx \quad = \quad \inf\left\{\mathscr{S}_f^+(\mathscr{Z}) \;\middle|\; \mathscr{Z} \in \mathbf{Z}(a,b)\right\}$$

als ***Unter-*** bzw. ***Oberintegral*** der Funktion f. Stimmen Unter- und Oberintegral von f überein, dann bezeichnet man f als ***Riemann-integrierbar*** und nennt

$$\int_{a}^{b} f(x)\,dx \quad = \quad \int_{a\bigstar}^{b} f(x)\,dx \quad = \quad \int_{a}^{b\bigstar} f(x)\,dx$$

das ***Riemann-Integral*** der Funktion f.

Man beachte, dass Unter- und Oberintegral bei einer beschränkten Funktion f stets endliche reelle Zahlen sind. Sind nämlich $c, d \in \mathbb{R}$ Werte mit $c \le f(x) \le d$ für alle $x \in [a,b]$, und ist $\mathscr{Z} = \{x_1, ..., x_{n-1}\}$ eine Zerlegung von $[a,b]$, dann können wir die Unter- und Obersummen gegeben durch $\mathscr{S}_f^-(\mathscr{Z}) = \sum_{k=1}^{n} c_k(x_k - x_{k-1})$ und $\mathscr{S}_f^+(\mathscr{Z}) = \sum_{k=1}^{n} d_k(x_k - x_{k-1})$ wegen $c \le c_k \le d_k \le d$ durch $c(b-a)$ nach unten und durch $d(b-a)$ nach oben abschätzen.

Ist $[a,b]$ ein endliches, abgeschlossenes Intervall, $\mathscr{Z} = \{x_1, ..., x_{n-1}\}$ eine Zerlegung, und setzen wir $x_0 = a$, $x_n = b$, dann bezeichnet man die Zahl

$$\delta(\mathscr{Z}) \quad = \quad \max\{\, x_{k+1} - x_k \mid 0 \le k < n\} \qquad \text{als } \textbf{\textit{Feinheit}} \text{ der Zerlegung.}$$

Eine Zerlegung $\mathscr{Z}'$ heißt ***Verfeinerung*** von $\mathscr{Z}$, wenn $\mathscr{Z}' \supseteq \mathscr{Z}$ gilt.

(9.3) Lemma Sei $f : [a,b] \to \mathbb{R}$ eine beschränkte Funktion, und seien $\mathscr{Z}$, $\mathscr{Z}'$ Zerlegungen von $[a,b]$, wobei $\mathscr{Z}'$ eine Verfeinerung von $\mathscr{Z}$ ist. Dann gilt

$$\mathscr{S}_f^-(\mathscr{Z}') \quad \ge \quad \mathscr{S}_f^-(\mathscr{Z}) \qquad \text{und} \qquad \mathscr{S}_f^+(\mathscr{Z}') \quad \le \quad \mathscr{S}_f^+(\mathscr{Z}).$$

Beweis: Sei $\mathscr{Z} = \{x_1, ..., x_{m-1}\}$ und $\mathscr{Z}' = \{y_1, ..., y_{n-1}\}$. Wegen $\mathscr{Z} \subseteq \mathscr{Z}'$ gibt es für jedes k mit $1 \le k < m$ ein $\ell_k \in \{1, ..., n-1\}$ mit $x_k = y_{\ell_k}$. Außerdem definieren wir $x_0 = y_0 = a$, $x_m = y_n = b$ sowie $\ell_0 = 0$ und $\ell_m = n$. Für $0 \le k < m$ sei $c_k = \sup f([x_k, x_{k+1}])$, und für $0 \le \ell < n$ setzen wir $d_\ell = \sup f([y_\ell, y_{\ell+1}])$. Für jedes ℓ mit $\ell_k \le \ell < \ell_{k+1}$ gilt dann $[y_\ell, y_{\ell+1}] \subseteq [x_k, x_{k+1}]$ und folglich $d_\ell \le c_k$. Wir erhalten

$$\mathscr{S}_f^+(\mathscr{Z}') \quad = \quad \sum_{\ell=0}^{n-1} d_\ell (y_{\ell+1} - y_\ell) \quad = \quad \sum_{k=0}^{m-1} \sum_{\ell=\ell_k}^{\ell_{k+1}-1} d_\ell (y_{\ell+1} - y_\ell) \quad \le$$

$$\sum_{k=0}^{m-1} \sum_{\ell=\ell_k}^{\ell_{k+1}-1} c_k (y_{\ell+1} - y_\ell) \quad = \quad \sum_{k=0}^{m-1} c_k (x_{k+1} - x_k) \quad = \quad \mathscr{S}_f^+(\mathscr{Z}).$$

Der Beweis der anderen Ungleichung läuft völlig analog. □

(9.4) Folgerung Für jede beschränkte Funktion $f : [a,b] \to \mathbb{R}$ gilt

$$\int_{a\bigstar}^{b} f(x)\,dx \quad \le \quad \int_{a}^{b\bigstar} f(x)\,dx.$$

Beweis: Sei $\varepsilon \in \mathbb{R}^+$ vorgegeben. Nach Definition von Unter- und Oberintegral gibt es Zerlegungen $\mathscr{Z}$ und $\mathscr{Z}'$ mit

$$\mathscr{S}_f^-(\mathscr{Z}) \;>\; \int_{a\bigstar}^{b} f(x)\,dx - \varepsilon \qquad \text{und} \qquad \mathscr{S}_f^+(\mathscr{Z}') \;<\; \int_{a}^{b\bigstar} f(x)\,dx + \varepsilon.$$

Setzen wir $\mathscr{Z}'' = \mathscr{Z} \cup \mathscr{Z}'$, dann erhalten wir mit Lemma (9.3) die Ungleichungen $\mathscr{S}_f^-(\mathscr{Z}'') \ge \mathscr{S}_f^-(\mathscr{Z})$ und $\mathscr{S}_f^+(\mathscr{Z}'') \le \mathscr{S}_f^+(\mathscr{Z}')$. Da außerdem $\mathscr{S}_f^-(\mathscr{Z}'') \le \mathscr{S}_f^+(\mathscr{Z}'')$ gilt, erhalten wir insgesamt die Abschätzung

$$\int_{a}^{b\bigstar} f(x)\,dx - \int_{a\bigstar}^{b} f(x)\,dx \quad > \quad (\mathscr{S}_f^+(\mathscr{Z}') - \varepsilon) - (\mathscr{S}_f^-(\mathscr{Z}) + \varepsilon) \quad \ge$$

$$(\mathscr{S}_f^+(\mathscr{Z}'') - \varepsilon) - (\mathscr{S}_f^-(\mathscr{Z}'') + \varepsilon) \quad \ge \quad -2\varepsilon.$$

Da $\varepsilon \in \mathbb{R}^+$ beliebig klein gewählt werden kann, folgt daraus die Behauptung. □

Aus Folgerung (9.4) ergibt sich somit, dass für jede beschränkte Funktion $f : [a,b] \to \mathbb{R}$ und jede Zerlegung $\mathscr{Z}$ von $[a,b]$ jeweils die Ungleichungen

$$\mathscr{S}_f^-(\mathscr{Z}) \quad \leq \quad \int_{a\bigstar}^{b} f(x)\,dx \quad \leq \quad \int_{a}^{b\bigstar} f(x)\,dx \quad \leq \quad \mathscr{S}_f^+(\mathscr{Z}) \qquad \text{erfüllt sind.}$$

(9.5) Satz Sei $f : [a,b] \to \mathbb{R}$ beschränkt. Dann sind folgenden Aussagen äquivalent.

(i) Die Funktion f ist Riemann-integrierbar.

(ii) Für jedes $\varepsilon \in \mathbb{R}^+$ gibt es eine Zerlegung $\mathscr{Z}$ mit $\mathscr{S}_f^+(\mathscr{Z}) - \mathscr{S}_f^-(\mathscr{Z}) < \varepsilon$.

Beweis: „(i) ⇒ (ii)“ Sei $\varepsilon \in \mathbb{R}^+$ vorgegeben. Nach Definition des Unterintegrals gibt es eine Zerlegung $\mathscr{Z}$ mit

$$\mathscr{S}_f^-(\mathscr{Z}) \quad > \quad \int_{a\bigstar}^{b} f(x)\,dx - \tfrac{1}{2}\varepsilon.$$

Ebenso gibt es nach Definition des Oberintegrals eine Zerlegung $\mathscr{Z}'$ mit

$$\mathscr{S}_f^+(\mathscr{Z}') \quad < \quad \int_{a}^{b\bigstar} f(x)\,dx + \tfrac{1}{2}\varepsilon.$$

Weil f Riemann-integrierbar ist, stimmen Unter- und Oberintegral mit dem Riemann-Integral überein. Ersetzen wir $\mathscr{Z}$ und $\mathscr{Z}'$ durch die feinere Zerlegung $\mathscr{Z}'' = \mathscr{Z} \cup \mathscr{Z}'$, dann erhalten wir mit Lemma (9.3) insgesamt $\mathscr{S}_f^-(\mathscr{Z}'') > \int_a^b f(x)\,dx - \frac{1}{2}\varepsilon$ und $\mathscr{S}_f^+(\mathscr{Z}'') < \int_a^b f(x)\,dx + \frac{1}{2}\varepsilon$. Es folgt

$$\mathscr{S}_f^+(\mathscr{Z}'') - \mathscr{S}_f^-(\mathscr{Z}'') \quad < \quad \left(\int_a^b f(x)\,dx + \tfrac{1}{2}\varepsilon\right) - \left(\int_a^b f(x)\,dx - \tfrac{1}{2}\varepsilon\right) \quad = \quad \varepsilon.$$

„(ii) ⇒ (i)“ Sei $\varepsilon \in \mathbb{R}^+$ vorgegeben, und sei $\mathscr{Z}$ eine Zerlegung mit der unter (ii) angegebenen Eigenschaft. Nach Definition von Ober- und Unterintegral gilt

$$\mathscr{S}_f^-(\mathscr{Z}) \quad \leq \quad \int_{a\bigstar}^{b} f(x)\,dx \quad \leq \quad \int_{a}^{b\bigstar} f(x)\,dx \quad \leq \quad \mathscr{S}_f^+(\mathscr{Z}) \quad \leq \quad \mathscr{S}_f^-(\mathscr{Z}) + \varepsilon.$$

Weil $\varepsilon \in \mathbb{R}^+$ beliebig klein gewählt werden kann, müssen Unter- und Oberintegral übereinstimmen. Also ist f Riemann-integrierbar. □

(9.6) Satz Seien $f, g : [a,b] \to \mathbb{R}$ Riemann-integrierbare Funktionen und $\lambda \in \mathbb{R}$. Dann sind auch $f+g$ und λf Riemann-integrierbar. Außerdem gilt

(i) $\displaystyle\int_a^b (f+g)(x)\,dx = \int_a^b f(x)\,dx + \int_a^b g(x)\,dx$

(ii) $\displaystyle\int_a^b (\lambda f)(x)\,dx = \lambda \int_a^b f(x)\,dx$

(iii) Aus $f \leq g$ folgt $\displaystyle\int_a^b f(x)\,dx \leq \int_a^b g(x)\,dx.$

Beweis: zu (i) Zur Vorbereitung bemerken wir, dass für jede Zerlegung $\mathscr{Z} = \{x_1, ..., x_{n-1}\}$ von $[a,b]$ jeweils $\mathscr{S}^-_{f+g}(\mathscr{Z}) \geq \mathscr{S}^-_f(\mathscr{Z}) + \mathscr{S}^-_g(\mathscr{Z})$ und $\mathscr{S}^+_{f+g}(\mathscr{Z}) \leq \mathscr{S}^+_f(\mathscr{Z}) + \mathscr{S}^+_g(\mathscr{Z})$ gilt. Für jedes k mit $0 \leq k < n$ ist nämlich $\inf f([x_k, x_{k+1}]) + \inf g([x_k, x_{k+1}])$ eine untere Schranke von $(f+g)([x_k, x_{k+1}])$, denn für jedes $x \in [x_k, x_{k+1}]$ gilt

$$(f+g)(x) \quad = \quad f(x) + g(x) \quad \geq \quad \inf f([x_k, x_{k+1}]) + \inf g([x_k, x_{k+1}]).$$

Daraus folgt $\inf (f+g)([x_k, x_{k+1}]) \geq \inf f([x_k, x_{k+1}]) + \inf g([x_k, x_{k+1}])$, und Summation über alle k ergibt $\mathscr{S}^-_{f+g}(\mathscr{Z}) \geq \mathscr{S}^-_f(\mathscr{Z}) + \mathscr{S}^-_g(\mathscr{Z})$. Der Beweis der zweiten Ungleichung läuft analog.

Sei nun $\varepsilon \in \mathbb{R}^+$ vorgegeben. Weil f und g nach Voraussetzung Riemann-integrierbar Funktionen sind, gibt es nach Satz (9.5) (ii) Zerlegungen $\mathscr{Z}'$ und $\mathscr{Z}''$ von $[a,b]$ mit $\mathscr{S}^+_f(\mathscr{Z}') - \mathscr{S}^-_f(\mathscr{Z}') < \frac{1}{2}\varepsilon$ und $\mathscr{S}^+_g(\mathscr{Z}'') - \mathscr{S}^-_g(\mathscr{Z}'') < \frac{1}{2}\varepsilon$. Setzen wir $\mathscr{Z} = \mathscr{Z}' \cup \mathscr{Z}''$, dann gelten wegen Lemma (9.3) diese Ungleichungen auch für $\mathscr{Z}$ an Stelle von $\mathscr{Z}'$ bzw. $\mathscr{Z}''$. Es folgt

$$\begin{aligned} \mathscr{S}^+_{f+g}(\mathscr{Z}) - \mathscr{S}^-_{f+g}(\mathscr{Z}) \quad &\leq \quad \Big(\mathscr{S}^+_f(\mathscr{Z}) + \mathscr{S}^+_g(\mathscr{Z})\Big) - \Big(\mathscr{S}^-_f(\mathscr{Z}) + \mathscr{S}^-_g(\mathscr{Z})\Big) \quad = \\ \Big(\mathscr{S}^+_f(\mathscr{Z}) - \mathscr{S}^-_f(\mathscr{Z})\Big) + \Big(\mathscr{S}^+_g(\mathscr{Z}) - \mathscr{S}^-_g(\mathscr{Z})\Big) \quad &< \quad \tfrac{1}{2}\varepsilon + \tfrac{1}{2}\varepsilon \quad = \quad \varepsilon. \end{aligned}$$

Wegen Satz (9.5) (ii) ist $f+g$ also Riemann-integrierbar. Außerdem gilt sowohl

$$\mathscr{S}^-_f(\mathscr{Z}) + \mathscr{S}^-_g(\mathscr{Z}) \quad \leq \quad \mathscr{S}^-_{f+g}(\mathscr{Z}) \quad \leq \quad \int_a^b (f+g)(x)\,dx \quad \leq \quad \mathscr{S}^+_{f+g}(\mathscr{Z}) \quad \leq \quad \mathscr{S}^+_f(\mathscr{Z}) + \mathscr{S}^+_g(\mathscr{Z})$$

als auch

$$\mathscr{S}^-_f(\mathscr{Z}) + \mathscr{S}^-_g(\mathscr{Z}) \quad \leq \quad \int_a^b f(x)\,dx + \int_a^b g(x)\,dx \quad \leq \quad \mathscr{S}^+_f(\mathscr{Z}) + \mathscr{S}^+_g(\mathscr{Z}).$$

also

$$\left|\left(\int_a^b f(x)\,dx + \int_a^b g(x)\,dx\right) - \int_a^b (f+g)(x)\,dx\right| \leq \Big(\mathscr{S}^+_f(\mathscr{Z}) + \mathscr{S}^+_g(\mathscr{Z})\Big) - \Big(\mathscr{S}^-_f(\mathscr{Z}) + \mathscr{S}^-_g(\mathscr{Z})\Big) < \varepsilon.$$

Da $\varepsilon \in \mathbb{R}^+$ beliebig klein gewählt werden kann, stimmen die beiden Integrale überein.

zu (ii) Hier unterscheiden wir drei Fälle. Ist $\lambda = 0$, dann ist die Funktion λf konstant null, und es ist leicht zu sehen, dass $\mathscr{S}^-_{\lambda f}(\emptyset) = \mathscr{S}^+_{\lambda f}(\emptyset) = 0$ ist. Daraus folgt, dass Unter- und Oberintegral von λf ebenfalls null sind, und dass f somit Riemann-integrierbar ist, mit Integral null. Die Gleichung $\int_a^b (\lambda f)(x)\,dx = \lambda \int_a^b f(x)\,dx$ ist also für $\lambda = 0$ ebenfalls erfüllt.

Betrachten wir nun den Fall $\lambda > 0$. Zum Nachweis der Riemann-Integrierbarkeit von λf sei $\varepsilon \in \mathbb{R}^+$ vorgegeben. Auf Grund der Integrierbarkeit von f existiert eine Zerlegung $\mathscr{Z}$ mit $\mathscr{S}^+_f(\mathscr{Z}) - \mathscr{S}^-_f(\mathscr{Z}) < \frac{\varepsilon}{\lambda}$. Sei $\mathscr{Z} = \{x_1, ..., x_{n-1}\}$. Wir setzen als bekannt voraus, dass für jede nach oben beschränkte Teilmenge $A \subseteq \mathbb{R}$ und für $\lambda A = \{\lambda a \mid a \in \mathbb{R}\}$ jeweils die Beziehung $\sup(\lambda A) = \lambda \cdot \sup A$ gilt, und ebenso $\inf(\lambda A) = \lambda \cdot \inf A$ für jede nach unten beschränkte Teilmenge $A \subseteq \mathbb{R}$. Daraus folgt

$$\begin{aligned}\mathscr{S}^+_{\lambda f}(\mathscr{Z})-\mathscr{S}^-_{\lambda f}(\mathscr{Z}) &= \sum_{k=0}^{n-1}\sup(\lambda f)([x_k,x_{k+1}])(x_{k+1}-x_k)-\sum_{k=0}^{n-1}\inf(\lambda f)([x_k,x_{k+1}])(x_{k+1}-x_k)\\ &= \lambda\sum_{k=0}^{n-1}\sup f([x_k,x_{k+1}])(x_{k+1}-x_k)-\lambda\sum_{k=0}^{n-1}\inf f([x_k,x_{k+1}])(x_{k+1}-x_k)\\ &= \lambda\Big(\mathscr{S}^+_f(\mathscr{Z})-\mathscr{S}^-_f(\mathscr{Z})\Big) < \lambda\cdot\frac{\varepsilon}{\lambda} = \varepsilon\ ,\end{aligned}$$

wodurch die Riemann-Integrierbarkeit von λf nachgewiesen ist. Aus den Ungleichungen $\mathscr{S}^-_f(\mathscr{Z}) \le \int_a^b f(x)\,dx \le \mathscr{S}^+_f(\mathscr{Z})$ folgt sowohl

$$\lambda\mathscr{S}^-_f(\mathscr{Z}) \le \lambda\int_a^b f(x)\,dx \le \lambda\mathscr{S}^+_f(\mathscr{Z})$$

als auch

$$\lambda\mathscr{S}^-_f(\mathscr{Z}) = \mathscr{S}^-_{\lambda f}(\mathscr{Z}) \le \int_a^b(\lambda f)(x)\,dx \le \mathscr{S}^+_{\lambda f}(\mathscr{Z}) = \lambda\mathscr{S}^+_f(\mathscr{Z}).$$

Da die Differenz $\lambda\mathscr{S}^+_f(\mathscr{Z})-\lambda\mathscr{S}^-_f(\mathscr{Z})$ kleiner als ε ist, muss auch die Differenz von $\int_a^b(\lambda f)(x)\,dx$ und $\lambda\int_a^b f(x)\,dx$ betragsmäßig kleiner als ε sein. Weil $\varepsilon\in\mathbb{R}^+$ beliebig klein gewählt werden kann, stimmen die beiden Werte überein.

Im Fall $\lambda<0$ läuft die Argumentation weitgehend analog; es ist lediglich darauf zu achten, dass sich durch die Multiplikation mit einer negativen Zahl die Ungleichungen umdrehen. Diesmal basiert der Beweis darauf, dass für eine nach oben beschränkte Teilmenge A die Gleichung $\inf(\lambda A)=\lambda\cdot\sup A$ und für eine nach unten beschränkte Teilmenge jeweils $\sup(\lambda A)=\lambda\cdot\inf A$ gilt. Wieder sei $\varepsilon\in\mathbb{R}^+$ vorgegeben. Diesmal wählen wir die Zerlegung $\mathscr{Z}=\{x_1,\ldots,x_{n-1}\}$ so, dass $\mathscr{S}^+_f(\mathscr{Z})-\mathscr{S}^-_f(\mathscr{Z})<\frac{\varepsilon}{(-\lambda)}$ erfüllt ist. Weil die Obersummen durch Suprema und die Untersummen durch Infima definiert sind, liefern uns die Beziehungen von oben die Gleichungen $\mathscr{S}^+_{\lambda f}(\mathscr{Z})=\lambda\mathscr{S}^-_f(\mathscr{Z})$ und $\mathscr{S}^-_{\lambda f}(\mathscr{Z})=\lambda\mathscr{S}^+_f(\mathscr{Z})$. Wir erhalten

$$\mathscr{S}^+_{\lambda f}(\mathscr{Z})-\mathscr{S}^-_{\lambda f}(\mathscr{Z}) = \lambda\Big(\mathscr{S}^-_f(\mathscr{Z})-\mathscr{S}^+_f(\mathscr{Z})\Big) = (-\lambda)\Big(\mathscr{S}^+_f(\mathscr{Z})-\mathscr{S}^-_f(\mathscr{Z})\Big) < \lambda\cdot\frac{\varepsilon}{(-\lambda)} = \varepsilon\ ,$$

woraus sich wiederum die Riemann-Integrierbarkeit von λf ergibt. Aus den Ungleichungen $\mathscr{S}^-_f(\mathscr{Z}) \le \int_a^b f(x)\,dx \le \mathscr{S}^+_f(\mathscr{Z})$ folgt diesmal sowohl

$$\lambda\mathscr{S}^+_f(\mathscr{Z}) \le \lambda\int_a^b f(x)\,dx \le \lambda\mathscr{S}^-_f(\mathscr{Z})$$

als auch

$$\lambda\mathscr{S}^+_f(\mathscr{Z}) = \mathscr{S}^-_{\lambda f}(\mathscr{Z}) \le \int_a^b(\lambda f)(x)\,dx \le \mathscr{S}^+_{\lambda f}(\mathscr{Z}) \le \lambda\mathscr{S}^-_f(\mathscr{Z}).$$

Da die Differenz $\lambda\mathscr{S}^-_f(\mathscr{Z})-\lambda\mathscr{S}^+_f(\mathscr{Z})$ kleiner als ε ist, ergibt sich daraus wie im vorherigen Fall die Übereinstimmung von $\int_a^b(\lambda f)(x)\,dx$ und $\lambda\int_a^b f(x)\,dx$.

zu (iii) Sei $\varepsilon \in \mathbb{R}^+$ beliebig vorgegeben. Weil f und g Riemann-integrierbar sind, gibt es Zerlegungen $\mathscr{Z}'$ und $\mathscr{Z}''$ von $[a,b]$ mit $\mathscr{S}_f^+(\mathscr{Z}') - \mathscr{S}_f^-(\mathscr{Z}') < \varepsilon$ und $\mathscr{S}_g^+(\mathscr{Z}'') - \mathscr{S}_g^-(\mathscr{Z}'') < \varepsilon$. Setzen wir $\mathscr{Z} = \mathscr{Z}' \cup \mathscr{Z}''$, dann bleiben diese Ungleichungen wegen Lemma (9.3) erhalten. Wegen $f \le g$ gilt $\mathscr{S}_f^-(\mathscr{Z}) \le \mathscr{S}_f^+(\mathscr{Z}) \le \mathscr{S}_g^+(\mathscr{Z})$ und somit

$$\int_a^b g(x)\,dx - \int_a^b f(x)\,dx \quad \ge \quad \mathscr{S}_g^-(\mathscr{Z}) - \mathscr{S}_f^+(\mathscr{Z}) \quad \ge \quad \mathscr{S}_g^+(\mathscr{Z}) - \mathscr{S}_f^-(\mathscr{Z}) - 2\varepsilon \quad \ge \quad -2\varepsilon.$$

Weil $\varepsilon \in \mathbb{R}^+$ beliebig klein gewählt werden kann, folgt daraus die gewünschte Ungleichung. □

(9.7) Satz Sei $f : [a,b] \to \mathbb{R}$ beschränkt und $c \in]a,b[$ ein beliebig gewählter Punkt. Genau dann ist f Riemann-integrierbar, wenn die Funktionen $f|_{[a,c]}$ und $f|_{[c,b]}$ Riemann-integrierbar sind, und in diesem Fall gilt

$$\int_a^b f(x)\,dx \quad = \quad \int_a^c f(x)\,dx + \int_c^b f(x)\,dx.$$

Beweis: Zunächst beweisen wir die „genau dann"-Aussage zur Integrierbarkeit. Auch hierzu verwenden wir Satz (9.5). Sei $g = f|_{[a,c]}$ und $h = f|_{[c,b]}$. Zur Vorbereitung bemerken wir: Ist $\mathscr{Z}'$ eine Zerlegung von $[a,c]$ und $\mathscr{Z}''$ eine Zerlegung von $[c,b]$, dann ist durch $\mathscr{Z} = \mathscr{Z}' \cup \{c\} \cup \mathscr{Z}''$ eine Zerlegung von $[a,b]$ gegeben. Außerdem gilt

$$\mathscr{S}_f^-(\mathscr{Z}) = \mathscr{S}_g^-(\mathscr{Z}') + \mathscr{S}_h^-(\mathscr{Z}'') \qquad \text{und} \qquad \mathscr{S}_f^+(\mathscr{Z}) = \mathscr{S}_g^+(\mathscr{Z}') + \mathscr{S}_h^+(\mathscr{Z}''). \tag{9.8}$$

Wir beschränken und auf den Beweis der ersten Gleichung. Wir können die Elemente von $\mathscr{Z}'$ und $\mathscr{Z}''$ so durchnummerieren, dass $\mathscr{Z}' = \{x_1, \dots, x_{m-1}\}$ und $\mathscr{Z}'' = \{x_{m+1}, \dots, x_{n-1}\}$ gilt, mit $m, n \in \mathbb{N}_0$ und $m+1 < n$. Setzen wir außerdem $x_m = c$ und definieren $c_k = \inf f([x_k, x_{k+1}])$ für $0 \le k < n$, dann erhalten wir

$$\mathscr{S}_f^-(\mathscr{Z}) \quad = \quad \sum_{k=0}^{n-1} c_k(x_{k+1} - x_k) \quad = \quad \sum_{k=0}^{m-1} c_k(x_{k+1} - x_k) + \sum_{k=m+1}^{n-1} c_k(x_{k+1} - x_k) \quad = \quad \mathscr{S}_g^-(\mathscr{Z}') + \mathscr{S}_h^-(\mathscr{Z}'').$$

„⇒" Sei $\varepsilon \in \mathbb{R}^+$ vorgegeben. Weil f nach Voraussetzung integrierbar ist, existiert eine Zerlegung $\mathscr{Z}$ von $[a,b]$ mit $\mathscr{S}_f^+(\mathscr{Z}) - \mathscr{S}_f^-(\mathscr{Z}) < \varepsilon$. Dabei können wir annehmen, dass c in $\mathscr{Z}$ liegt, denn ersetzen wir $\mathscr{Z}$ durch $\mathscr{Z}_1 = \mathscr{Z} \cup \{c\}$, dann gilt $\mathscr{S}_f^-(\mathscr{Z}_1) \ge \mathscr{S}_f^-(\mathscr{Z})$ und $\mathscr{S}_f^+(\mathscr{Z}_1) \le \mathscr{S}_f^+(\mathscr{Z}_1)$ nach Lemma (9.3) und somit auch $\mathscr{S}_f^+(\mathscr{Z}_1) - \mathscr{S}_f^-(\mathscr{Z}_1) < \varepsilon$.

Setzen wir also $c \in \mathscr{Z}$ voraus, und definieren wir $\mathscr{Z}' =]a,c[\cap \mathscr{Z}$ sowie $\mathscr{Z}'' =]c,b[\cap \mathscr{Z}$. Dann ist $\mathscr{Z}'$ eine Zerlegung von $[a,c]$, die Menge $\mathscr{Z}''$ ist eine Zerlegung von $[c,b]$, und es gilt $\mathscr{Z} = \mathscr{Z}' \cup \{c\} \cup \mathscr{Z}''$. Wie oben gezeigt wurde, folgt daraus (9.8). Wegen $\mathscr{S}_g^+(\mathscr{Z}') - \mathscr{S}_g^-(\mathscr{Z}') \ge 0$, $\mathscr{S}_h^+(\mathscr{Z}'') - \mathscr{S}_h^-(\mathscr{Z}'') \ge 0$ und

$$\begin{aligned}\Big(\mathscr{S}_g^+(\mathscr{Z}') - \mathscr{S}_g^-(\mathscr{Z}')\Big) + \Big(\mathscr{S}_h^+(\mathscr{Z}'') - \mathscr{S}_h^-(\mathscr{Z}'')\Big) \quad &= \quad \Big(\mathscr{S}_g^+(\mathscr{Z}') + \mathscr{S}_h^+(\mathscr{Z}'')\Big) - \Big(\mathscr{S}_g^-(\mathscr{Z}') + \mathscr{S}_h^-(\mathscr{Z}'')\Big) \\ &= \quad \mathscr{S}_f^+(\mathscr{Z}) - \mathscr{S}_f^-(\mathscr{Z}) \quad < \quad \varepsilon\end{aligned}$$

gilt $\mathscr{S}_g^+(\mathscr{Z}') - \mathscr{S}_g^-(\mathscr{Z}') < \varepsilon$ und $\mathscr{S}_h^+(\mathscr{Z}'') - \mathscr{S}_h^-(\mathscr{Z}'') < \varepsilon$. Dies zeigt, dass g und h Riemann-integrierbar sind.

„⇐“ Wieder sei $\varepsilon \in \mathbb{R}^+$ vorgegeben. Weil g und h nach Voraussetzung integrierbar sind, gibt es Zerlegungen $\mathscr{Z}'$ von $[a,c]$ und $\mathscr{Z}''$ von $[c,b]$ mit $\mathscr{S}_g^+(\mathscr{Z}') - \mathscr{S}_g^-(\mathscr{Z}') < \frac{1}{2}\varepsilon$ und $\mathscr{S}_h^+(\mathscr{Z}'') - \mathscr{S}_h^-(\mathscr{Z}'') < \frac{1}{2}\varepsilon$. Definieren wir $\mathscr{Z} = \mathscr{Z}' \cup \{c\} \cup \mathscr{Z}''$, dann gelten wiederum die Gleichungen (9.8). Wir erhalten

$$\begin{aligned} \mathscr{S}_f^+(\mathscr{Z}) - \mathscr{S}_f^-(\mathscr{Z}) &= \Big(\mathscr{S}_g^+(\mathscr{Z}') + \mathscr{S}_h^+(\mathscr{Z}'')\Big) - \Big(\mathscr{S}_g^-(\mathscr{Z}') + \mathscr{S}_h^-(\mathscr{Z}'')\Big) = \\ \Big(\mathscr{S}_g^+(\mathscr{Z}') - \mathscr{S}_g^-(\mathscr{Z}')\Big) + \Big(\mathscr{S}_h^+(\mathscr{Z}'') - \mathscr{S}_h^-(\mathscr{Z}'')\Big) &< \tfrac{1}{2}\varepsilon + \tfrac{1}{2}\varepsilon = \varepsilon. \end{aligned}$$

Dies zeigt, dass f Riemann-integrierbar ist. Zum Schluss beweisen wir noch die angegebene Gleichung. Es gilt sowohl

$$\mathscr{S}_f^-(\mathscr{Z}) \le \int_a^b f(x)\,dx \le \mathscr{S}_f^+(\mathscr{Z})$$

als auch

$$\mathscr{S}_f^-(\mathscr{Z}) = \mathscr{S}_g^-(\mathscr{Z}') + \mathscr{S}_h^-(\mathscr{Z}'') \le \int_a^c g(x)\,dx + \int_c^b h(x)\,dx \le \mathscr{S}_g^+(\mathscr{Z}') + \mathscr{S}_h^+(\mathscr{Z}'') \le \mathscr{S}_f^+(\mathscr{Z}).$$

Daraus folgt

$$\begin{aligned} \left| \int_a^c f(x)\,dx + \int_c^b f(x)\,dx - \int_a^b f(x)\,dx \right| &= \left| \int_a^c g(x)\,dx + \int_c^b h(x)\,dx - \int_a^b f(x)\,dx \right| \\ &\le \mathscr{S}_f^+(\mathscr{Z}) - \mathscr{S}_f^-(\mathscr{Z}) < \varepsilon. \end{aligned}$$

Da $\varepsilon \in \mathbb{R}^+$ beliebig klein gewählt werden kann, erhalten wir die angegebene Gleichung. □

Unser nächstes Ziel ist der Nachweis, dass jede stetige Funktion f auf einem abgeschlossenen Intervall Riemann-integrierbar ist. Hierzu benötigen wir ein wenig Vorbereitung.

(9.9) Definition Eine Funktion $f : D \to \mathbb{R}$ wird auf einer Teilmenge $D \subseteq \mathbb{R}$ ***gleichmäßig stetig*** genannt, wenn für jedes $\varepsilon \in \mathbb{R}^+$ ein $\delta \in \mathbb{R}^+$ existiert, so dass die Implikation

$$|x-y| < \delta \quad \Rightarrow \quad |f(x)-f(y)| < \varepsilon$$

für alle $x, y \in D$ erfüllt ist.

Eine gleichmäßig stetige Funktion $f : D \to \mathbb{R}$ ist nach dem ε-δ-Kriterium offenbar in jedem Punkt $a \in D$ stetig. Umgekehrt braucht eine stetige Funktion aber nicht gleichmäßig stetig zu sein. Als Beispiel betrachten wir die Funktion

$$f : \mathbb{R}^+ \longrightarrow \mathbb{R} \quad , \qquad x \mapsto \frac{1}{x} \quad ,$$

die als rationale Funktion in jedem Punkt ihres Definitionsbereichs stetig ist. Sie ist aber nicht gleichmäßig stetig. Wäre sie es, dann gäbe es für $\varepsilon = 1$ ein $\delta > 0$, so dass

$$|x-y| < \delta \qquad \Rightarrow \qquad |f(x)-f(y)| < 1 \qquad \text{für alle} \quad x, y \in \,]0,1[$$

erfüllt ist. Wir wählen nun $n \in \mathbb{N}$ so groß, so dass $\frac{1}{2n} < \delta$ erfüllt ist. Für $x = \frac{1}{n}$ und $y = \frac{1}{2n}$ gilt dann zwar $|x-y| = |\frac{1}{n} - \frac{1}{2n}| = \frac{1}{2n} < \delta$, aber $|f(x)-f(y)| = |n-2n| = n \ge 1$, im Widerspruch zur Annahme.

(9.10) Satz Jede auf einem endlichen, abgeschlossenen Intervall $[a,b] \subseteq \mathbb{R}$ stetige Funktion ist dort auch gleichmäßig stetig.

Beweis: Sei $f : [a,b] \to \mathbb{R}$ stetig und nehmen wir an, f wäre nicht gleichmäßig stetig. Dann gäbe es ein $\varepsilon \in \mathbb{R}^+$ mit der Eigenschaft, dass für jedes $\delta \in \mathbb{R}^+$ ein Paar $x, y \in [a,b]$ mit $|x-y| < \delta$, aber $|f(x)-(y)| \geq \varepsilon$ existiert. Insbesondere gäbe es dann für jedes $n \in \mathbb{N}$ Punkte $x_n, y_n \in [a,b]$ mit $|x_n - y_n| < \frac{1}{n}$, aber $|f(x_n)-f(y_n)| \geq \varepsilon$.

Nach dem Satz (3.10) von Bolzano-Weierstraß besitzt $(x_n)_{n\in\mathbb{N}}$ einen Häufungspunkt p. Wir finden also eine streng monoton wachsende Folge $(n_k)_{k\in\mathbb{N}}$ natürlicher Zahlen mit $|x_{n_k}-p| < \frac{1}{k}$ für alle $k \in \mathbb{N}$. Die Folge $(x_{n_k})_{k\in\mathbb{N}}$ konvergiert offenbar gegen p. Wir behaupten, dass auch $\lim_k y_{n_k} = p$ gilt. Um dies zu beweisen, wählen wir für vorgegebenes $\varepsilon' \in \mathbb{R}^+$ ein so großes $K \in \mathbb{N}$, dass $\frac{1}{n_K} < \frac{1}{2}\varepsilon'$ und $|x_{n_k} - p| < \frac{1}{2}\varepsilon'$ für alle $k \geq K$ erfüllt ist. Es gilt dann

$$|y_{n_k} - p| \quad \leq \quad |y_{n_k} - x_{n_k}| + |x_{n_k} - p| \quad < \quad \frac{1}{n_k} + \tfrac{1}{2}\varepsilon' \quad < \quad \varepsilon' \qquad \text{für alle} \quad k \geq K.$$

Damit ist die Behauptung bewiesen. Auf Grund der Stetigkeit von f gilt nun $\lim_k (f(x_{n_k}) - f(y_{n_k})) = f(p) - f(p) = 0$. Dies widerspricht der Annahme $|f(x_n)-f(y_n)| \geq \varepsilon$ für alle $n \in \mathbb{N}$. □

Nach diesen Vorbereitungen können wir nun zeigen

(9.11) Satz Jede stetige Funktion auf einem endlichen, abgeschlossenen Intervall ist Riemann-integrierbar.

Beweis: Sei $f : [a,b] \to \mathbb{R}$ eine stetige Funktion und $\varepsilon \in \mathbb{R}^+$ vorgegeben. Nach Satz (9.10) ist f als stetige Funktion auf einem abgeschlossenen Intervall auch gleichmäßig stetig. Seien $\varepsilon' = \frac{\varepsilon}{b-a}$ und die Zahl $\delta \in \mathbb{R}^+$ so gewählt, dass

$$|x-y| < \delta \quad \Rightarrow \quad |f(x)-f(y)| < \varepsilon' \qquad \text{für alle} \quad x, y \in [a,b]$$

erfüllt ist. Sei außerdem $n \in \mathbb{N}$ so groß, dass $\frac{b-a}{n} < \delta$ gilt. Wir definieren nun $x_k = a + \frac{k}{n}(b-a)$ für $0 \leq k \leq n$ und setzen $\mathcal{Z} = \{x_1, \ldots, x_{n-1}\}$. Nach dem Maximumsprinzip, Satz (5.17), gibt es jeweils Punkte $y_k, z_k \in [x_k, x_{k+1}]$, so dass $c_k = f(y_k)$ das Minimum und $d_k = f(z_k)$ das Maximum von f auf $[x_k, x_{k+1}]$ ist. Daraus folgt, dass Unter- und Obersumme der Zerlegung $\mathcal{Z}$ durch

$$\mathcal{S}_f^-(\mathcal{Z}) = \sum_{k=0}^{n-1} c_k (x_{k+1} - x_k) \qquad \text{und} \qquad \mathcal{S}_f^+(\mathcal{Z}) = \sum_{k=0}^{n-1} d_k (x_{k+1} - x_k)$$

gegeben sind. Wegen $|y_k - z_k| \leq \frac{b-a}{n} < \delta$ gilt jeweils $d_k - c_k = f(z_k) - f(y_k) < \varepsilon'$. Damit erhalten wir

$$\mathcal{S}_f^+(\mathcal{Z}) - \mathcal{S}_f^-(\mathcal{Z}) \quad = \quad \sum_{k=0}^{n-1} (d_k - c_k)(x_{k+1} - x_k) \quad \leq$$

$$\sum_{k=0}^{n-1} \varepsilon'(x_{k+1} - x_k) \quad = \quad \varepsilon'(x_n - x_0) \quad = \quad \varepsilon'(b-a) \quad = \quad \varepsilon.$$

Da $\varepsilon \in \mathbb{R}^+$ beliebig klein gewählt werden kann, zeigt dies nach Satz (9.5) (ii), dass f Riemann-integrierbar ist. □

Der Satz lässt sich auf eine noch größere Klasse von Funktionen verallgemeinern. Eine Funktion $f : [a,b] \to \mathbb{R}$ wird ***stückweise stetig*** genannt, wenn eine Zerlegung $\mathscr{Z} = \{x_1, \ldots, x_{n-1}\}$ von $[a,b]$ existiert, so dass die Funktionen $g_k = f|_{]x_k, x_{k+1}[}$ stetig sind und

$$\lim_{x \to x_k} g_k \quad \text{und} \qquad \lim_{x \to x_{k+1}} g_k$$

als eigentliche Grenzwerte existieren, für $0 \leq k < n$. Dabei wird wie immer $x_0 = a$ und $x_n = b$ gesetzt. Aus den Grenzwertbedingungen folgt nämlich, dass wir g_k als stetige Funktion auf $[x_k, x_{k+1}]$ betrachten können, die außer in den Randpunkten x_k, x_{k+1} mit f übereinstimmen. Nach Satz (9.11) ist g_k und damit $f|_{[x_k, x_{k+1}]}$ integrierbar, und mit Satz (9.7) folgt daraus die Integrierbarkeit von f auf $[a,b]$.

Der folgende Satz wird im folgenden Abschnitt ein zentrales Hilfsmittel beim Beweis des Hauptsatzes der Differential- und Integralrechnung sein.

(9.12) Satz *(Mittelwertsatz der Integralrechnung)*
Seien $f, g : [a,b] \to \mathbb{R}$ stetige Funktionen, wobei wir $g \geq 0$ voraussetzen. Dann gibt es ein $c \in [a,b]$ mit

$$\int_a^b f(x)g(x)\,dx \quad = \quad f(c) \int_a^b g(x)\,dx.$$

Der Spezialfall $g = 1$ liefert die Gleichung $\int_a^b f(x)\,dx = f(c)(b-a)$ für ein $c \in [a,b]$.

Beweis: Wir definieren $u = \min f([a,b])$ und $v = \max f([a,b])$. Nach dem Maximumsprinzip sind diese Werte als reelle Zahlen definiert, und es existieren Punkte $p, q \in [a,b]$ mit $f(p) = u$ und $f(q) = v$. Wir können $p < q$ voraussetzen, denn im Fall $q > p$ läuft der Beweis völlig analog, und im Fall $p = q$ ist die Funktion f konstant und die Aussage daher trivial. Nach Voraussetzung gilt $ug \leq fg \leq vg$ auf dem gesamten Intervall $[a,b]$. Auf Grund von Satz (9.6) (iii) gelten deshalb die Ungleichungen

$$u \int_a^b g(x)\,dx \quad \leq \quad \int_a^b f(x)g(x)\,dx \quad \leq \quad v \int_a^b g(x)\,dx.$$

Insbesondere gibt es ein $\mu \in [u,v]$ mit $\int_a^b f(x)g(x)\,dx = \mu \int_a^b g(x)\,dx$. Da f stetig ist, liefert uns der Zwischenwertsatz (5.14) einen Punkt $c \in [p,q]$ mit $f(c) = \mu$. Damit ist die Aussage bewiesen. □

Für reelle Zahlen $a, b \in \mathbb{R}$ mit $a < b$ und integrierbare Funktionen f auf $[a,b]$ definieren wir

$$\int_a^a f(x)\,dx = 0 \qquad \text{und} \qquad \int_b^a f(x)\,dx = -\int_a^b f(x)\,dx.$$

Man beachte, dass der Mittelwertsatz der Integralrechnung für $a \geq b$ gültig bleibt. Ebenso leicht überzeugt man sich davon, dass die letzte Gleichung in Satz (9.7) für beliebige reelle Zahlen a, b, c gültig bleibt, unabhängig von deren Anordnung.

Im weiteren Verlauf untersuchen wir nun den Zusammenhang zwischen dem Riemann-Integral und der Ableitung von Funktionen. Als Ergebnis werden wir das aus der Schulmathematik bekannte Resultat herleiten, dass sich das Integral einer Funktion f mit Hilfe von Stammfunktionen berechnen lässt.

(9.13) Definition Sei $D \subseteq \mathbb{R}$ und $f : D \to \mathbb{R}$ eine beliebige Funktion. Eine Funktion $F : D \to \mathbb{R}$ wird ***Stammfunktion*** von f genannt, wenn F auf ganz D differenzierbar ist und $F' = f$ gilt.

Als erstes zeigen wir mit Hilfe des Mittelwertsatzes der Integralrechnung, dass das Riemann-Integral zur Definition von Stammfunktionen verwendet werden kann.

(9.14) Proposition Sei $f :]a,b[\to \mathbb{R}$ stetig und $c \in]a,b[$. Die Funktion $F :]a,b[\to \mathbb{R}$ sei definiert durch

$$F(x) \quad = \quad \int_c^x f(t)\,dt.$$

Dann ist F eine Stammfunktion von f.

Beweis: Sei $x \in]a,b[$ und $(x_n)_{n\in\mathbb{N}}$ eine Folge in $]a,b[\setminus \{x\}$ mit $\lim_n x_n = x$. Für den Differentialquotienten von F gilt

$$\frac{F(x_n)-F(x)}{x_n-x} \quad = \quad \frac{1}{x_n-x}\left(\int_c^{x_n} f(t)\,dt - \int_c^x f(t)\,dt\right) \quad = \quad \frac{1}{x_n-x}\int_x^{x_n} f(t)\,dt.$$

Nach dem Mittelwertsatz der Integralrechnung, angwendet auf f und die konstante Funktion $g(x) = 1$, gibt es für jedes $n \in \mathbb{N}$ im Fall $x < x_n$ ein $y_n \in [x, x_n]$ und im Fall $x > x_n$ ein $y_n \in [x_n, x]$ mit

$$\int_x^{x_n} f(t)\,dt \quad = \quad (x_n - x) f(y_n).$$

Nach Definition gilt $|y_n - x| \le |x_n - x|$ für alle $n \in \mathbb{N}$. Mit $(x_n)_{n\in\mathbb{N}}$ konvergiert also auch die Folge $(y_n)_{n\in\mathbb{N}}$ gegen x, und aus der Stetigkeit von f folgt

$$\lim_{n\to\infty} \frac{F(x_n)-F(x)}{x_n-x} \quad = \quad \lim_{n\to\infty} f(y_n) \quad = \quad f(x).$$

Also gilt tatsächlich $F'(x) = f(x)$ für alle $x \in]a,b[$. □

(9.15) Proposition Sei $f :]a,b[\to \mathbb{R}$ eine Funktion und $F :\to \mathbb{R}$ eine Stammfunktion. Eine weitere Funktion G auf $]a,b[$ ist genau dann eine Stammfunktion von f, wenn die Differenz $G - F$ auf $]a,b[$ konstant ist.

Beweis: „⇐“ Ist $G-F$ konstant, dann gilt $(G-F)'(x) = 0$ und somit $G'(x) = F'(x) = f(x)$ für alle $x \in]a,b[$. „⇒“ Seien F, G Stammfunktionen von f. Nach Voraussetzung gilt $(G-F)' = G' - F' = f - f = 0$, und nach Folgerung (8.5) ist $G - F$ damit auf $]a,b[$ konstant. □

(9.16) Satz *(Hauptsatz der Differential- und Integralrechnung)*
Sei $f :]a,b[\to \mathbb{R}$ stetig und F eine Stammfunktion von f. Für alle $c, d \in]a,b[$ gilt dann

$$\int_c^d f(x)dx \quad = \quad F(d) - F(c).$$

Beweis: Nach Proposition (9.14) ist durch $\tilde{F}(x) = \int_c^x f(t)\,dt$ jedenfalls *eine* Stammfunktion von f gegeben. Für diese gilt $\tilde{F}(c) = 0$ und $\tilde{F}(d) = \int_c^d f(t)\,dt$. Da F eine weitere Stammfunktion ist, existiert nach Proposition (9.15) ein $u \in \mathbb{R}$ mit $F(x) = \tilde{F}(x) + u$ für alle $x \in]a,b[$. Es gilt also

$$F(d) - F(c) \quad = \quad \tilde{F}(d) - \tilde{F}(c) \quad = \quad \int_c^d f(x)\,dx. \qquad \square$$

Um Schreibarbeit zu sparen, verwendet man im Zusammenhang mit einer Stammfunktion F häufig die Notation

$$\Big[F(x)\Big]_c^d \quad \text{für die Differenz} \quad F(d) - F(c).$$

Die Gleichung aus dem Hauptsatz kann dann in der Form $\int_c^d f(x)\,dx = \big[F(x)\big]_c^d$ geschrieben werden. Man kann die Schreibweise noch weiter vereinfachen, indem man einfach

$$\int f(x)\,dx \quad = \quad F(x)$$

notiert. Ausgeschrieben soll dies bedeuten, dass die Gleichung $\int_c^d f(x)\,dx = \big[F(x)\big]_c^d$ für beliebige c, d im Definitionsbereich von f erfüllt ist. Den Ausdruck auf der linken Seite bezeichnet man als ***unbestimmtes Integral***. Man beachte, dass die Funktion F auf der rechten Seite nur bis auf eine Konstante eindeutig bestimmt ist. Dort kann eine beliebige Stammfunktion von f erscheinen. So wäre zum Beispiel auch die Gleichung

$$\int x^3\,dx \quad = \quad \tfrac{1}{4}x^4 + 1$$

korrekt. Im allgemeinen bemüht man sich bei der Angabe von unbestimmten Integralen aber, eine möglichst „einfache“ Stammfunktion auszuwählen.

Für $n \in \mathbb{R}$ mit $n \neq -1$ gilt

$$\int x^n\,dx \quad = \quad \frac{x^{n+1}}{n+1} \quad , \quad \text{unter der Bedingung } x \neq 0 \text{ falls } n < 0.$$

Die Bedingung $x \neq 0$ bedeutet, dass die Null nicht im Integrationsintervall $[a,b]$ liegen darf, d.h. die ausführlich geschriebene Gleichung

$$\int_a^b x^n\,dx \quad = \quad \left[\frac{x^{n+1}}{n+1}\right]_a^b$$

ist nur gültig, wenn $a, b \in \mathbb{R}$ beide positiv oder beide negativ sind. Denn ansonsten wäre die Funktion x^n nicht an jeder Stelle des Intervalls $[a,b]$ definiert und daher auch $\int_a^b x^n\,dx$ undefiniert. Selbst wenn wir die Funktion an der Stelle 0 willkürlich fortsetzen würden, etwa in der Form

$$x \quad \mapsto \quad \begin{cases} x^n & \text{für} \quad x \neq 0 \\ 0 & \text{für} \quad x = 0 \end{cases}$$

hätte das Integralzeichen keine Bedeutung, da wir schon allein für die Definition des Ober- oder Unterintegrals eine *beschränkte* Funktion benötigen.

Als weiteres Beispiel beweisen wir die Gleichung

$$\int \frac{dx}{x} \quad = \quad \ln|x| \qquad \text{für} \quad x \neq 0.$$

Sei nämlich $[a,b]$ ein Integrationsintervall. Nach Voraussetzung gilt $0 \notin [a,b]$, d.h. die Grenzen $a,b \in \mathbb{R}$ sind entweder beide positiv oder beide negativ. Für $x > 0$ ist $\ln(x)$ eine Stammfunktion von $x \mapsto \frac{1}{x}$, also gilt

$$\int_a^b \frac{dx}{x} \quad = \quad \Big[\ln(x)\Big]_a^b \qquad \text{für} \quad a,b > 0$$

nach dem Hauptsatz der Differential- und Integralrechnung. Im Fall $x < 0$ ist $F(x) = \ln(-x)$ eine Stammfunktion für $x \mapsto \frac{1}{x}$, denn nach der Kettenregel gilt $F'(x) = (-1)\frac{1}{-x} = \frac{1}{x}$. Wir erhalten

$$\int_a^b \frac{dx}{x} \quad = \quad \Big[\ln(-x)\Big]_a^b \qquad \text{für} \quad a,b < 0.$$

Sowohl für $a,b > 0$ als auch für $a,b < 0$ gilt also die Gleichung $\displaystyle\int_a^b \frac{dx}{x} = \Big[\ln|x|\Big]_a^b$.

Aus dem Hauptsatz der Differential- und Integralrechnung ergibt sich wegen $\exp'(x) = \exp(x)$ auch unmittelbar die Integralformel

$$\int \exp(x)\,dx \quad = \quad \exp(x).$$

(9.17) Satz *(Substitutionsregel)*
Seien $I, J \subseteq \mathbb{R}$ offene Intervalle, $f : I \to \mathbb{R}$ stetig und $u : J \to \mathbb{R}$ stetig differenzierbar mit $u(J) \subseteq I$. Dann gilt

$$\int_a^b f(u(t))u'(t)\,dt \quad = \quad \int_{u(a)}^{u(b)} f(x)\,dx. \qquad \text{für alle} \quad a,b \in J.$$

Beweis: Sei $F : I \to \mathbb{R}$ eine Stammfunktion von f und $G = F \circ u$. Die Anwendung der Kettenregel liefert

$$(F \circ u)'(t)\,dt \quad = \quad F'(u(t))u'(t) \quad = \quad f(u(t))u'(t).$$

Diese Funktion ist auf Grund unserer Voraussetzungen stetig. Folglich erhalten wir nach dem Hauptsatz der Differential- und Integralrechnung

$$\int_a^b f(u(t))u'(t)\,dt \quad = \quad \int_a^b (F \circ u)'(t)\,dt \quad = \quad \Big[(F \circ u)(t)\Big]_a^b \quad = \quad \Big[F(x)\Big]_{u(a)}^{u(b)} \quad = \quad \int_{u(a)}^{u(b)} f(x)\,dx. \qquad \square$$

Als erstes Anwendungsbeispiel für die Substitutionsregel beweisen wir die Formel

$$\int \frac{dx}{x+c} \quad = \quad \ln|x+c| \qquad \text{für} \quad x \neq -c.$$

Seien $a, b \in \mathbb{R}$. Liegt $-c$ nicht im Intervall $[a, b]$, dann ist 0 nicht in $[a+c, b+c]$ enthalten. Anwendung der Substitutionsregel auf $u(t) = t + c$ mit Ableitung $u'(t) = 1$ liefert

$$\int_a^b \frac{dt}{t+c} \quad = \quad \int_a^b \frac{u'(t)\,dx}{u(t)} \quad = \quad \int_{a+c}^{b+c} \frac{dx}{x} \quad = \quad \Big[\, \ln|x| \,\Big]_{a+c}^{b+c} \quad = \quad \Big[\, \ln|x+c| \,\Big]_a^b$$

Als weiteres Beispiel überprüfen wir die Gleichung

$$\int x \exp(x^2)\,dx \quad = \quad \tfrac{1}{2}\exp(x^2).$$

Hier verwenden wir $u(t) = t^2$ mit der Ableitung $u'(t) = 2t$.

$$\int_a^b t e^{t^2}\,dt \quad = \quad \tfrac{1}{2}\int_a^b (2t) e^{t^2}\,dt \quad = \quad \tfrac{1}{2}\int_{a^2}^{b^2} e^x\,dx \quad = \quad \Big[\tfrac{1}{2}e^{x^2}\Big]_a^b$$

(9.18) Satz *(partielle Integration)*
Sei $I \subseteq \mathbb{R}$ ein offenes Intervall, und seien $f, g : I \to \mathbb{R}$ zwei stetig differenzierbare Funktionen. Dann gilt

$$\int_a^b f(x)g'(x)\,dx \quad = \quad \Big[f(x)g(x)\Big]_a^b - \int_a^b f'(x)g(x)\,dx \qquad \text{für alle} \quad a, b \in I.$$

Beweis: Wir wenden die Produktregel auf die Funktion $F = f\,g$ an und erhalten $F'(x) = f'(x)g(x) + f(x)g'(x)$. Der Hauptsatz der Differential- und Integralrechnung liefert

$$\int_a^b f'(x)g(x)\,dx + \int_a^b f(x)g'(x)\,dx \quad = \quad \int_a^b F'(x)\,dx \quad = \quad \Big[F(x)\Big]_a^b \quad = \quad \Big[f(x)g(x)\Big]_a^b.$$

Durch Umstellen erhalten wir die gewünschte Gleichung. □

Als Anwendung der partiellen Integration beweisen wir

$$\int x \exp(x) \quad = \quad (x-1)\exp(x).$$

Dazu definieren wir die Funktionen $f(x) = x$ und $g(x) = \exp(x)$ mit den Ableitungen $f'(x) = 1$ und $g'(x) = \exp(x)$. Für beliebige $a, b \in \mathbb{R}$ gilt dann

$$\int_a^b x \exp(x)\,dx \quad = \quad \int_a^b f(x)g'(x)\,dx \quad = \quad \Big[f(x)g(x)\Big]_a^b - \int_a^b f'(x)g(x)\,dx \quad =$$

$$\Big[x\exp(x)\Big]_a^b - \int_a^b 1 \cdot \exp(x)\,dx \quad = \quad \Big[x\exp(x)\Big]_a^b - \Big[\exp(x)\Big]_a^b \quad = \quad \Big[(x-1)\exp(x)\Big]_a^b.$$

Durch mehrfache Anwendung der partiellen Integration lassen sich kompliziertere Ausdrücke integrieren, beispielsweise

$$\int (x^2+5)\exp(x)\,dx \quad = \quad (x^2-2x+7)\exp(x).$$

Für beliebige $a, b \in \mathbb{R}$ gilt nämlich

$$\int_a^b (x^2+5)\exp(x)\,dx \quad = \quad \Big[(x^2+5)\exp(x)\Big]_a^b - \int_a^b 2x\exp(x)\,dx \quad =$$

$$\Big[(x^2+5)\exp(x)\Big]_a^b - \Big[2x\exp(x)\Big]_a^b + \int_a^b 2\exp(x)\,dx \quad = \quad \Big[(x^2+5)\exp(x)\Big]_a^b - \Big[2x\exp(x)\Big]_a^b + \Big[2\exp(x)\Big]_a^b$$

$$= \quad \Big[(x^2-2x+7)\exp(x)\Big]_a^b.$$

Wir geben eine Reihe weiterer wichtiger Beispiele für die Berechnung von Integralen mittels Substitutionsregel und partieller Integration an.

(i) $\displaystyle\int \frac{dx}{(x+c)^n} \quad = \quad \frac{1}{1-n}(x+c)^{1-n} \quad$ für $n \in \mathbb{N}$, $n \geq 2$.

Zur Notation sei angemerkt, dass allgemein $\int \frac{dx}{f(x)}$ nur eine vereinfachte Schreibweise für $\int \frac{1}{f(x)}\,dx$ ist. Wiederum verwenden wir die Substitution $u(t) = x + c$ mit $u'(t) = 1$ an und erhalten

$$\int_a^b \frac{dt}{(t+c)^n} \quad = \quad \int_{a+c}^{b+c} x^{-n}\,dx \quad = \quad \left[\frac{1}{1-n}x^{1-n}\right]_{a+c}^{b+c} \quad = \quad \left[\frac{1}{1-n}(x+c)^{1-n}\right]_a^b \quad .$$

(ii) $\displaystyle\int \ln(x)\,dx \quad = \quad x\ln(x) - x \quad$ für $x > 0$

Hier besteht der Trick darin, die partielle Integration auf das „triviale“ Produkt $\ln(x) = 1 \cdot \ln(x)$ anzuwenden.

$$\int_a^b \ln(x)\,dx \quad = \quad \int_a^b 1\cdot\ln(x)\,dx \quad = \quad \Big[x\ln(x)\Big]_a^b - \int_a^b x\cdot\frac{1}{x}\,dx$$

$$= \quad \Big[x\ln(x)\Big]_a^b - \int_a^b 1\,dx \quad = \quad \Big[x\ln(x) - x\Big]_a^b.$$

Allgemeiner kann mit partieller Integration und vollständiger Induktion gezeigt werden, dass

$$\int x^n \ln(x)\,dx \quad = \quad \frac{x^{n+1}}{(n+1)^2}\left((n+1)\ln(x) - 1\right) \qquad \text{für alle } n \in \mathbb{N}_0 \text{ gilt.}$$

(iii) $\displaystyle\int \frac{dx}{\sqrt{1-x^2}} \quad = \quad \arcsin(x) \quad$ für $|x| < 1$

Hier ist der entscheidende Trick die Substitution $x = \sin(t)$. Zu beachten ist, dass der Sinus das Intervall $[-\frac{1}{2}\pi, \frac{1}{2}\pi]$ bijektiv auf das Intervall $[-1, 1]$ abbildet. Für $-1 < a, b < 1$ liegen somit $\arcsin(a), \arcsin(b)$ im Intervall $[-\frac{1}{2}\pi, \frac{1}{2}\pi]$, und der Kosinus von Winkeln in diesem Bereich ist positiv. Daraus folgt $\sqrt{1-\sin(t)^2} = \cos(t)$, und nicht etwa

$\sqrt{1-\sin(t)^2} = -\cos(t)$, was ja mit der allgemeinen Gleichung $\sin(t)^2 + \cos(t)^2 = 1$ für beliebige $t \in \mathbb{R}$ ebenfalls vereinbar wäre.

$$\int_a^b \frac{dx}{\sqrt{1-x^2}} = \int_{\arcsin(a)}^{\arcsin(b)} \frac{\sin'(t)\,dt}{\sqrt{1-\sin(t)^2}} = \int_{\arcsin(a)}^{\arcsin(b)} \frac{\cos(t)\,dt}{\cos(t)} = \int_{\arcsin(a)}^{\arcsin(b)} 1\,dt$$

$$= [t]_{\arcsin(a)}^{\arcsin(b)} = \arcsin(b) - \arcsin(a) = [\arcsin(x)]_a^b.$$

(iv) $\displaystyle\int \frac{dx}{\sqrt{x^2-1}} = \operatorname{arcosh}(x)$ für $x > 1$

In den Übungen wurden die Funktionen *Sinus hyperbolicus* $\sinh(x) = \frac{1}{2}e^x - \frac{1}{2}e^{-x}$ und *Kosinus hyperbolicus* $\cosh(x) = \frac{1}{2}e^x + e^{-x}$ eingeführt. Der Sinus hyperbolicus ist streng monoton wachsend und bildet $\mathbb{R}$ bijektiv auf $\mathbb{R}$ ab. Wie man leicht nachrechnet, ist seine Umkehrfunktion gegeben durch $\operatorname{arsinh}(x) = \ln(x + \sqrt{x^2+1})$. Der Kosinus hyperbolicus ist auf $]-\infty, 0]$ streng monoton fallend und auf $[0, +\infty[$ streng monoton wachsend. Dabei werden die nichtnegativen reellen Zahlen bijektiv auf $[1, +\infty[$ abgebildet, und die Umkehrfunktion auf diesem Bereich ist gegeben durch $\operatorname{arcosh}(x) = \ln(x + \sqrt{x^2-1})$.

Die beiden Funktionen erfüllen die Funktionalgleichung $\cosh(x)^2 - \sinh(x)^2 = 1$, und ihre Ableitungen stehen durch $\sinh'(x) = \cosh(x)$ und $\cosh'(x) = \sinh(x)$ zueinander in Beziehung. Dies kann genutzt werden, und mittels Substitution die angebene Stammfunktion zu ermitteln.

$$\int_a^b \frac{dx}{\sqrt{x^2-1}} = \int_{\operatorname{arcosh}(a)}^{\operatorname{arcosh}(b)} \frac{\cosh'(t)\,dt}{\sqrt{\cosh(t)^2-1}} = \int_{\operatorname{arcosh}(a)}^{\operatorname{arcosh}(b)} \frac{\sinh(t)\,dt}{\sinh(t)} = \int_{\operatorname{arcosh}(a)}^{\operatorname{arcosh}(b)} 1\,dt$$

$$= [t]_{\operatorname{arcosh}(a)}^{\operatorname{arcosh}(b)} = [\operatorname{arcosh}(x)]_a^b.$$

Hier ist im dritten Schritt wieder zu beachten, dass die Integration über eine Teilmenge von $\mathbb{R}_+$ abläuft, wo der $\sinh(t)$ positive Werte annimmt. Aus diesem Grund ist $\sqrt{\cosh(t)^2-1} = \sinh(t)$ und nicht etwa $\sqrt{\cosh(t)^2-1} = -\sinh(t)$.

(v) $\displaystyle\int \frac{dx}{x^2+1} = \arctan(x)$

Wie wir aus § 7 wissen, bildet die Tangensfunktion $\tan(x)$ das Intervall $\left]-\frac{1}{2}\pi, \frac{1}{2}\pi\right[$ bijektiv auf $\mathbb{R}$ ab. Die Umkehrabbildung $\arctan : \mathbb{R} \to \left]-\frac{1}{2}\pi, \frac{1}{2}\pi\right[$ wird als *Arkus tangens* bezeichnet. Die Ableitung der Tangensfunktion erhält man durch die Quotientenregel; es gilt

$$\tan'(x) = \frac{\cos(x)\cos(x) - \sin(x)(-\sin(x))}{\cos(x)^2} = \frac{\cos(x)^2 + \sin(x)^2}{\cos(x)^2} = 1 + \tan(x)^2.$$

Diese Beziehung nutzt man zur Bestimmung der Stammfunktion durch Substitution.

$$\int_a^b \frac{dx}{x^2+1} \quad = \quad \int_{\arctan(a)}^{\arctan(b)} \frac{\tan'(t)\,dt}{\tan(t)^2+1} \quad = \quad \int_{\arctan(a)}^{\arctan(b)} \frac{1+\tan(t)^2}{\tan(t)^2+1}\,dt \quad = \quad \int_{\arctan(a)}^{\arctan(b)} 1\,dt$$

$$= \quad [t]_{\arctan(a)}^{\arctan(b)} \quad = \quad [\arctan(x)]_a^b .$$

Als weitere Anwendung für die partielle Integration beweisen wir eine Rekursionsformel für die Stammfunktion von $\sin(x)^n$ für $n \in \mathbb{N}$.

(vi) $$\int \sin(x)^n\,dx \quad = \quad \frac{n-1}{n}\int \sin(x)^{n-2}\,dx - \frac{1}{n}\cdot\cos(x)\sin(x)^{n-1} \qquad \text{für } n \geq 2$$

Durch partielle Integration erhält man

$$\int_a^b \sin(x)^n\,dx \quad = \quad \int_a^b \sin(x)\cdot\sin(x)^{n-1}\,dx \quad = \quad \Big[(-\cos(x))\sin(x)^{n-1}\Big]_a^b - \int_a^b (-\cos(x))\cos(x)(n-1)\sin(x)^{n-2}\,dx$$

$$= \quad -\Big[\cos(x)\sin(x)^{n-1}\Big]_a^b + (n-1)\int_a^b \cos(x)^2\sin(x)^{n-2}\,dx \quad = \quad -\Big[\cos(x)\sin(x)^{n-1}\Big]_a^b + (n-1)\int_a^b (1-\sin(x)^2)\sin(x)^{n-2}\,dx$$

$$= \quad -\Big[\cos(x)\sin(x)^{n-1}\Big]_a^b + (n-1)\int_a^b \sin(x)^{n-2}\,dx + (n-1)\int_a^b \sin(x)^n\,dx.$$

Diese Gleichung kann umgestellt werden zu

$$n\int_a^b \sin(x)^n\,dx \quad = \quad (n-1)\int_a^b \sin(x)^{n-2}\,dx - \Big[\cos(x)\sin(x)^{n-1}\Big]_a^b$$

und durch Multiplikation mit $\frac{1}{n}$ auf beiden Seiten erhält man die angegebene Gleichung.

Unser nächstes Ziel besteht darin, die Integration auch über nicht abgeschlossene oder unendlich ausgedehnte Intervalle zu ermöglichen. Dies geschieht durch eine geeignete Grenzwertbildung.

(9.19) Definition Seien $a \in \mathbb{R}\cup\{-\infty\}$ und $b \in \mathbb{R}$, mit $a < b$. Sei $f :]a,b] \to \mathbb{R}$ eine Funktion mit der Eigenschaft, dass f über jedes Teilintervall der Form $[c,b]$ mit $a < c < b$ integrierbar ist. Dann nennt man den Grenzwert

$$\int_a^b f(x)\,dx \quad = \quad \lim_{c\to a}\int_c^b f(x)\,dx\,, \quad \text{sofern er existiert,}$$

das ***uneigentliche Riemann-Integral*** von f über das halboffene Intervall $]a,b]$.

Dieselbe Definition ist auch für die rechte Intervallgrenze möglich.

(9.20) Definition Seien $a \in \mathbb{R}$ und $b \in \mathbb{R} \cup \{+\infty\}$, mit $a < b$. Sei $f : [a, b[\to \mathbb{R}$ eine Funktion mit der Eigenschaft, dass f über jedes Teilintervall der Form $[a, c]$ mit $a < c < b$ integrierbar ist. Dann nennt man den Grenzwert

$$\int_a^b f(x)\,dx \quad = \quad \lim_{c \to b} \int_a^c f(x)\,dx \qquad \text{im Falle der Existenz}$$

das ***uneigentliche Riemann-Integral*** von f über das halboffene Intervall $[a, b[$.

Beispiel 1: Berechnung von $\displaystyle\int_0^1 \frac{1}{\sqrt{x}}\,dx$

$$\int_0^1 \frac{1}{\sqrt{x}}\,dx \quad = \quad \lim_{c \to 0} \int_c^1 \frac{1}{\sqrt{x}}\,dx \quad = \quad \lim_{c \to 0} \Big[\, 2\sqrt{x} \,\Big]_c^1 \quad = \quad \lim_{c \to 0} \big(2 - 2\sqrt{c}\big) \quad = \quad 2.$$

Beispiel 2: Berechnung von $\displaystyle\int_{-\infty}^0 e^x\,dx$

$$\int_{-\infty}^0 e^x\,dx \quad = \quad \lim_{c \to -\infty} \int_c^0 e^x\,dx \quad = \quad \lim_{c \to -\infty} [\, e^x \,]_c^0 \quad =$$

$$\lim_{c \to -\infty} \left(e^0 - e^c\right) \quad = \quad \lim_{c \to -\infty} (1 - e^c) \quad = \quad 1 - 0 \quad = \quad 1.$$

Beispiel 3: Berechnung von $\displaystyle\int_1^{+\infty} 2xe^{-x^2}\,dx$

$$\int_1^{+\infty} 2xe^{-x^2}\,dx \quad = \quad \lim_{c \to +\infty} \int_1^c 2xe^{-x^2}\,dx \quad = \quad \lim_{c \to +\infty} -\int_1^c (-2x)e^{-x^2}\,dx \quad = \quad \lim_{c \to +\infty} -\int_{-1}^{-c^2} e^x\,dx \quad =$$

$$\lim_{c \to +\infty} \int_{-c^2}^{-1} e^x\,dx \quad = \quad \lim_{c \to +\infty} [\, e^x \,]_{-c^2}^{-1} \quad = \quad \lim_{c \to +\infty} \left(e^{-1} - e^{-c^2}\right) \quad = \quad e^{-1} - 0 \quad = \quad e^{-1}.$$

Schließlich ist es noch möglich, dass die Funktion an ***beiden*** Intervallenden undefiniert ist.

(9.21) Definition Seien $a \in \mathbb{R} \cup \{-\infty\}$ und $b \in \mathbb{R} \cup \{+\infty\}$, mit $a < b$. Sei $f :]a, b[\to \mathbb{R}$ eine Funktion mit der Eigenschaft, dass f über jedes Teilintervall der Form $[c, d]$ mit $a < c < d < b$ integrierbar ist. Sei nun $m \in]a, b[$ beliebig gewählt. Dann bezeichnet man die Summe der Grenzwerte

$$\int_a^b f(x)\,dx \quad = \quad \lim_{c \to a} \int_c^m f(x)\,dx \; + \; \lim_{d \to b} \int_m^d f(x)\,dx \quad ,$$

sofern sie beide existieren, als das ***uneigentliche Riemann-Integral*** von f über das offene Intervall $]a, b[$.

Es ist leicht zu sehen, dass sowohl die Existenz dieses uneigentlichen Integrals als auch dessen Wert unabhängig von der Wahl des Unterteilungspunktes m ist. Sind nämlich $m, m' \in \mathbb{R}$ mit $a < m < m' < b$, dann gilt für alle $c, d \in \mathbb{R}$ mit $a < c < m$ und $m' < d < b$ jeweils

$$\int_c^{m'} f(x)\,dx = \int_c^{m} f(x)\,dx + \int_m^{m'} f(x)\,dx \quad \text{und} \quad \int_m^{d} f(x)\,dx = \int_m^{m'} f(x)\,dx + \int_{m'}^{d} f(x)\,dx.$$

Die erste Gleichung zeigt, dass der Grenzwert

$$\lim_{c\to a}\int_c^{m} f(x)\,dx \quad \text{genau dann existiert, wenn der Grenzwert} \quad \lim_{c\to a}\int_c^{m'} f(x)\,dx \quad \text{existiert,}$$

und die zweite Gleichung liefert einen entsprechenden Zusammenhang zwischen der Existenz von

$$\lim_{d\to b}\int_{m'}^{d} f(x)\,dx \qquad \text{und} \qquad \lim_{d\to b}\int_m^{d} f(x)\,dx.$$

Desweiteren folgt aus der Gleichung

$$\int_c^{m} f(x)\,dx + \int_m^{d} f(x)\,dx \;=\; \int_c^{d} f(x)\,dx \;=\; \int_c^{m'} f(x)\,dx + \int_{m'}^{d} f(x)\,dx \;,$$

durch Grenzübergang $c \to a$ auf beiden Seiten zunächst die Gleichung

$$\lim_{c\to a}\int_c^{m} f(x)\,dx + \int_m^{d} f(x)\,dx \;=\; \lim_{c\to a}\int_c^{m'} f(x)\,dx + \int_{m'}^{d} f(x)\,dx \;,$$

und durch anschließenden Grenzübergang $d \to b$ auf beiden Seiten

$$\lim_{c\to a}\int_c^{m} f(x)\,dx + \lim_{d\to b}\int_m^{d} f(x)\,dx \;=\; \lim_{c\to a}\int_c^{m'} f(x)\,dx + \lim_{d\to b}\int_{m'}^{d} f(x)\,dx.$$

Beispiel 4: Berechnung von $\displaystyle\int_0^1 \left(\tfrac{1}{\sqrt{x}} + \tfrac{1}{\sqrt{1-x}}\right) dx$

Wir wählen als Unterteilungspunkt den Intervallmittelpunkt $m = \frac{1}{2}$. Es gilt

$$\int_0^{1/2} \left(\tfrac{1}{\sqrt{x}} + \tfrac{1}{\sqrt{1-x}}\right) dx \;=\; \lim_{c\to 0}\int_c^{1/2} \left(\tfrac{1}{\sqrt{x}} + \tfrac{1}{\sqrt{1-x}}\right) \;=\; \lim_{c\to 0}\Big[\, 2\sqrt{x} - 2\sqrt{1-x}\,\Big]_c^{1/2} \;=$$

$$\lim_{c\to 0}\left(2\sqrt{\tfrac{1}{2}} - 2\sqrt{\tfrac{1}{2}} - 2\sqrt{c} + 2\sqrt{1-c}\right) \;=$$

$$\lim_{c\to 0}\left(-2\sqrt{c} + 2\sqrt{1-c}\right) \;=\; -2\sqrt{0} + 2\sqrt{1} \;=\; 2.$$

Ebenso erhalten wir

$$\int_{1/2}^1 \left(\tfrac{1}{\sqrt{x}} + \tfrac{1}{\sqrt{1-x}}\right) dx \;=\; \lim_{c\to 1}\int_{1/2}^c \left(\tfrac{1}{\sqrt{x}} + \tfrac{1}{\sqrt{1-x}}\right) \;=\; \lim_{c\to 1}\Big[\, 2\sqrt{x} - 2\sqrt{1-x}\,\Big]_{1/2}^c \;=$$

$$\lim_{c\to 1}\left(2\sqrt{c} - 2\sqrt{1-c} - 2\sqrt{\tfrac{1}{2}} + 2\sqrt{\tfrac{1}{2}}\right) \;=$$

$$\lim_{c\to 1}\left(2\sqrt{c} - 2\sqrt{1-c}\right) \;=\; 2\sqrt{1} - 2\sqrt{0} \;=\; 2.$$

Insgesamt kommen wir damit zu dem Ergebnis

$$\int_0^1 \left(\tfrac{1}{\sqrt{x}} - \tfrac{1}{\sqrt{1-x}}\right) \quad = \quad \int_0^{1/2} \left(\tfrac{1}{\sqrt{x}} - \tfrac{1}{\sqrt{1-x}}\right) + \int_{1/2}^1 \left(\tfrac{1}{\sqrt{x}} - \tfrac{1}{\sqrt{1-x}}\right) \quad = \quad 2+2 \quad = \quad 4.$$

Zum Abschluss des Kapitels geben wir noch ein Verfahren an, mit dem sich Integrale beliebiger ***rationaler Funktionen*** berechnen lassen. Zur Erinnerung: Eine rationale Funktion ist eine Funktion f der Form $f(x) = g(x)/h(x)$, wobei g und h Polynomfunktionen bezeichnen und h ungleich null ist. Der Definitionsbereich einer rationalen Funktion ist $D = \mathbb{R} \setminus N$, wobei $N \subseteq \mathbb{R}$ die Nullstellenmenge des Polynoms h sei. Das Verfahren beruht auf dem folgenden Satz.

(9.22) Satz *(Satz von der Partialbruchzerlegung)*

Sei $f : D \to \mathbb{R}$ eine rationale Funktion wie oben beschrieben. Es seien $x_1, ..., x_m \in \mathbb{R}$ die verschiedenen rellen Nullstellen von h und $r_i \in \mathbb{N}$ jeweils die Vielfachheit der Nullstelle x_i. Weiter seien $z_1, \bar{z}_1, ..., z_n, \bar{z}_n$ die verschiedenen Paare nicht-reeller, konjugiert-komplexer Nullstellen, $p_i = 2\mathrm{Re}(z_i)$, $q_i = z_i\bar{z}_i = |z_i|^2$ und $s_i \in \mathbb{N}$ die gemeinsame Vielfachheit der Nullstellen $z_i, \bar{z}_i$. Dann gibt es Koeffizienten $a_{ij}, b_{ij} \in \mathbb{R}$ für $1 \le i \le m$, $1 \le j \le r_i$ und $c_{ij} \in \mathbb{R}$ für $1 \le i \le n$, $1 \le j \le s_i$ und eine Polynomfunktion q, so dass f in der Form

$$f(x) \quad = \quad q(x) + \sum_{i=1}^{m} \sum_{j=1}^{r_i} \frac{a_{ij}}{(x-x_i)^j} + \sum_{i=1}^{n} \sum_{j=1}^{s_i} \frac{b_{ij}x + c_{ij}}{(x^2 + p_i x + q_i)^j}$$

dargestellt werden kann.

Dieser Satz ist Bestandteil der Algebra und wird deshalb hier nicht bewiesen. Wir demonstrieren aber anhand zweier Beispiele, wie die im Satz angegebene Partialbruchzerlegung konkret berechnet werden kann.

1. Beispiel: Wir betrachten die rationale Funktion

$$f(x) \quad = \quad \frac{2x+3}{(x^2+2)(x+1)^2}.$$

Mit den Bezeichungen des Satzes ist $m = 1$ die Anzahl reeller Nullstellen, $n = 1$ die Anzahl der Paare konjugiert-komplexer Nullstellen, es ist $r_1 = 2$ die Vielfachheit der einzigen rellen Nullstelle und $s_1 = 1$ die gemeinsame Vielfachheit des einzigen Paares konjugiert-komplexer Nullstellen. Die einzige reelle Nullstelle ist $x_1 = -1$, und das einzige konjugiert-komplexe Paar ist $(z_1, \bar{z}_i) = (i\sqrt{2}, -i\sqrt{2})$. Die Koeffizienten a_{11}, a_{12}, b_{11} und c_{11} bestimmen wir durch den Ansatz

$$\frac{2x+3}{(x^2+2)(x+1)^2} \quad = \quad \frac{a_{11}}{x+1} + \frac{a_{12}}{(x+1)^2} + \frac{b_{11}x + c_{11}}{x^2+2}.$$

Die Multiplikation mit dem Hauptnenner liefert die Gleichung

$$2x+3 \quad = \quad a_{11}(x+1)(x^2+2) + a_{12}(x^2+2) + (b_{11}x + c_{11})(x+1)^2.$$

Diese kann umgeformt werden zu

$$2x+3 \quad = \quad (a_{11}+b_{11})x^3 + (a_{11}+a_{12}+2b_{11}+c_{11})x^2 + (2a_{11}+b_{11}+2c_{11})x + (2a_{11}+2a_{12}+c_{11}).$$

Durch Koeffizientenvergleich erhalten wir das linearen Gleichungssystem bestehend aus den Gleichungen

$$a_{11}+b_{11}=0 \quad , \quad a_{11}+a_{12}+2b_{11}+c_{11}=0 \quad , \quad 2a_{11}+b_{11}+2c_{11}=2 \quad , \quad 2a_{11}+2a_{12}+c_{11}=3$$

mit der eindeutigen Lösung $a_{11}=\frac{8}{9}$, $a_{12}=\frac{1}{3}$, $b_{11}=-\frac{8}{9}$, $c_{11}=\frac{5}{9}$. Die Funktion f hat somit die Form

$$f(x) \quad = \quad \frac{\frac{8}{9}}{x+1} + \frac{\frac{1}{3}}{(x+1)^2} + \frac{(-\frac{8}{9})x+\frac{5}{9}}{x^2+2}.$$

2. Beispiel: Diesmal sei f gegeben durch

$$f(x) \quad = \quad \frac{7x^2-3x+4}{(x^2+2)^2}.$$

Hier ist $m=0$, $n=1$, $s_1=2$ und $(z_1,\bar{z}_1)=(i\sqrt{2},-i\sqrt{2})$. Der Satz von der Partialbruchzerlegung liefert den Ansatz

$$\frac{7x^2-3x+4}{(x^2+2)^2} \quad = \quad \frac{b_{11}x+c_{11}}{x^2+2} + \frac{b_{12}x+c_{12}}{(x^2+2)^2}$$

mit $b_{11},c_{11},b_{12},c_{12}\in\mathbb{R}$. Diesselbe Vorgehensweise wie beim ersten Beispiel (Multiplikation der Gleichung mit dem Hauptnenner $(x^2+2)^2$ und Sortierung der rechten Seite der Gleichung, um die Koeffizienten der Monome $1,x,x^2,x^3$ zu bestimmen) liefert diesmal das lineare Gleichungssystem

$$2c_{11}+c_{12}=4 \quad , \quad 2b_{11}+b_{12}=-3 \quad , \quad c_{11}=7 \quad , \quad b_{11}=0$$

mit der eindeutigen Lösung $b_{11}=0$, $b_{12}=-3$, $c_{11}=7$ und $c_{12}=-10$. Wir erhalten somit

$$f(x) \quad = \quad \frac{7}{x^2+2} + \frac{(-3)x+(-10)}{(x^2+2)^2}.$$

Durch den Satz von der Partialbruchzerlegung ist die Integration beliebiger rationaler Funktionen auf die Integration der folgenden vier Typen von Funktionen zurückgeführt.

(i) $\displaystyle\int_a^b \frac{dx}{x-c}$ mit $c\notin[a,b]$

(ii) $\displaystyle\int_a^b \frac{dx}{(x-c)^n}$ mit $c\notin[a,b]$ und $n\in\mathbb{N}$, $n\geq 2$

(iii) $\displaystyle\int_a^b \frac{cx+d}{x^2+px+q}\,dx$ mit $c,d,p,q\in\mathbb{R}$, $p^2-4q<0$

(iv) $\displaystyle\int_a^b \frac{cx+d}{(x^2+px+q)^n}\,dx$ mit $c,d,p,q\in\mathbb{R}$, $p^2-4q<0$ und $n\in\mathbb{N}$, $n\geq 2$

Die Bedingung $p^2-4q<0$ beim Typ (iii) und (iv) besagt jeweils, dass die Diskriminante des Polynoms x^2+px+q negativ ist und das Polynom somit keine reellen Nullstellen, sondern ein Paar konjugiert-komplexer Nullstellen besitzt. Für den Typ (i) und (ii) haben wir das Integral weiter oben bereits ausgerechnet: Unter den angegebenen Bedingungen gilt

$$\int_a^b \frac{dx}{x-c} \quad = \quad [\,\ln|x-c|\,]_a^b$$

und

$$\int_a^b \frac{dx}{(x-c)^n} \quad = \quad \left[\ \frac{1}{(1-n)(x-c)^{n-1}}\ \right]_a^b.$$

Im Fall (iii) verwenden wir die bereits bekannt Tatsache, dass die Stammfunktion von $x \mapsto \frac{1}{x^2+1}$ der Arkus tangens ist. Für beliebige $c \in \mathbb{R}^+$ erhalten wir mit der Substitutionsregel

$$\int_a^b \frac{dx}{x^2+c} \quad = \quad \frac{1}{c}\int_a^b \frac{dx}{(\frac{x}{\sqrt{c}})^2+1} \quad = \quad \frac{1}{\sqrt{c}}\int_a^b \frac{\frac{1}{\sqrt{c}}\,dx}{(\frac{x}{\sqrt{c}})^2+1} \quad = \quad \frac{1}{\sqrt{c}}\int_{a/\sqrt{c}}^{b/\sqrt{c}} \frac{dt}{t^2+1} \quad =$$

$$\frac{1}{\sqrt{c}}\left[\ \arctan(t)\ \right]_{a/\sqrt{c}}^{b/\sqrt{c}} \quad = \quad \left[\ \frac{1}{\sqrt{c}}\arctan\left(\frac{x}{\sqrt{c}}\right)\ \right]_a^b.$$

Der allgemeine Fall, dass im Nenner ein Polynom der Form x^2+px+q mit $p^2-4q<0$ steht, führen wir durch erneute Anwendung der Substitutionsregel auf die soeben berechnete Form zurück. Entscheidend hierbei ist die Beobachtung, dass das Polynom x^2+px+q durch die Substitution $t = x+\frac{1}{2}p$ in das Polynom t^2+c mit $c = \frac{1}{4}(4q-p^2) > 0$ überführt wird: Es gilt

$$\begin{aligned} x^2+px+q \quad &= \quad (t-\tfrac{1}{2}p)^2+p(t-\tfrac{1}{2}p)+q \quad = \quad t^2-pt+\tfrac{1}{4}p^2+pt-\tfrac{1}{2}p^2+q \\ &= \quad t^2-\tfrac{1}{4}p^2+q \quad = \quad t^2+\tfrac{1}{4}(4q-p^2) \quad = \quad t^2+c. \end{aligned}$$

Durch die Substitution $t = x+\frac{1}{2}p$ erhalten wir somit

$$\int_a^b \frac{dx}{x^2+px+q} \quad = \quad \int_{a+\frac{1}{2}p}^{b+\frac{1}{2}p} \frac{dt}{t^2+c} \quad = \quad \left[\ \frac{1}{\sqrt{c}}\arctan\left(\frac{x}{\sqrt{c}}\right)\ \right]_{a+\frac{1}{2}p}^{b+\frac{1}{2}p} \quad =$$

$$\left[\ \frac{1}{\frac{1}{2}\sqrt{4q-p^2}}\arctan\left(\frac{x}{\frac{1}{2}\sqrt{4q-p^2}}\right)\ \right]_{a+\frac{1}{2}p}^{b+\frac{1}{2}p} \quad = \quad \left[\ \frac{2}{\sqrt{4q-p^2}}\arctan\left(\frac{2x}{\sqrt{4q-p^2}}\right)\ \right]_{a+\frac{1}{2}p}^{b+\frac{1}{2}p} \quad =$$

$$\left[\ \frac{2}{\sqrt{4q-p^2}}\arctan\left(\frac{2x+p}{\sqrt{4q-p^2}}\right)\ \right]_a^b.$$

Das Integral vom Typ (iii) wiederum kann durch Anwendung der Substitution $t = \varphi(x) = x^2+px+q$ auf diese Form zurückgeführt werden.

$$\int_a^b \frac{cx+d}{x^2+px+q}\,dx \quad = \quad \tfrac{1}{2}c\int_a^b \frac{2x+2\frac{d}{c}}{x^2+px+q}\,dx \quad = \quad \tfrac{1}{2}c\int_a^b \frac{2x+p+2\frac{d}{c}-p}{x^2+px+q}\,dx \quad =$$

$$\tfrac{1}{2}c\int_a^b \frac{2x+p}{x^2+px+q}\,dx \ + \ \tfrac{1}{2}c(2\tfrac{d}{c}-p)\int_a^b \frac{dx}{x^2+px+q} \quad = \quad \tfrac{1}{2}c\int_{\varphi(a)}^{\varphi(b)} \frac{dt}{t} \ + \ (d-\tfrac{1}{2}pc)\int_a^b \frac{dx}{x^2+px+q} \quad =$$

$$\left[\ \tfrac{1}{2}c\ln(t)\ \right]_{\varphi(a)}^{\varphi(b)} \ + \ \tfrac{1}{2}(d-\tfrac{1}{2}pc)\int_a^b \frac{dx}{x^2+px+q} \quad = \quad \left[\ \tfrac{1}{2}c\cdot\ln(x^2+px+q)\ \right]_a^b \ + \ (d-\tfrac{1}{2}pc)\int_a^b \frac{dx}{x^2+px+q}.$$

Mit Hilfe der Formeln für die Integrale vom Typ (i), (ii) und (iii) können wir nun die Integrale der beiden Funktionen angeben, die wir zuvor mittels Partialbruchzerlegung umgeformt haben.

$$\int \frac{2x+3}{(x^2+2)(x+1)^2}\,dx \quad = \quad \tfrac{8}{9}\cdot\ln|x+1| \;-\; \tfrac{4}{9}\cdot\ln(x^2+2) \;-\; \frac{1}{3x+3} \;+\; \tfrac{5}{18}\sqrt{2}\arctan\left(\frac{x}{\sqrt{2}}\right)$$

$$\int \frac{7x^2-3x+4}{(x^2+2)^2}\,dx \quad = \quad \frac{(-\frac{5}{2})x \;+\; \frac{3}{2}}{x^2+2} \;+\; \tfrac{9}{4}\sqrt{2}\cdot\arctan\left(\frac{x}{\sqrt{2}}\right).$$

Um auch die Integrale vom Typ (iv) berechnen zu können, bestimmen wir zunächst eine Rekursionsformel für die uneigentlichen Integral

$$I_n \quad = \quad \int \frac{dx}{(x^2+1)^n} \quad , \quad n \in \mathbb{N}.$$

Wir haben bereits gezeigt, dass $I_1 = \arctan(x)$ gilt. Um eine Formel für I_n in Abhängigkeit von I_{n-1} zu erhalten, beweisen wir zuerst für alle $a, b \in \mathbb{R}$ die Gleichung

$$\int_a^b \frac{2x}{(x^2+1)^n}\,dx \quad = \quad \left[\frac{1}{(1-n)}(x^2+1)^{n-1}\right]_a^b.$$

Tatsächlich erhält man durch Anwendung der Substitutionsregel

$$\int_a^b \frac{2x}{(x^2+1)^n}\,dx \quad = \quad \int\limits_{a^2+1} b^2+1t^{-n}\,dt \quad = \quad \left[\frac{1}{1-n}t^{1-n}\right]_{a^2+1}^{b^2+1}$$

$$= \quad \left[\frac{1}{(1-n)t^{n-1}}\right]_{a^2+1}^{b^2+1} \quad = \quad \left[\frac{1}{(1-n)(x^2+1)^{n-1}}\right]_a^b.$$

Die soeben bewiesene Gleichung in Verbindung mit partieller Integration liefert

$$[\; I_{n-1} \;]_a^b \quad = \quad \int_a^b \frac{dx}{(x^2+1)^{n-1}} \quad = \quad \int_a^b \frac{x^2+1}{(x^2+1)^n}\,dx \quad =$$

$$\int_a^b \tfrac{1}{2}x\cdot\frac{2x}{(x^2+1)^n}\,dx \;+\; \int_a^b \frac{dx}{(x^2+1)^n} \quad = \quad \int_a^b \tfrac{1}{2}x\cdot\frac{2x}{(x^2+1)^n}\,dx \;+\; [\; I_n \;]_a^b \quad =$$

$$\left[\tfrac{1}{2}x\cdot\frac{1}{(1-n)(x^2+1)^{n-1}}\right]_a^b - \int_a^b \tfrac{1}{2}\cdot\frac{dx}{(1-n)(x^2+1)^{n-1}} \;+\; [\; I_n \;]_a^b$$

$$= \quad \left[\frac{x}{2(1-n)(x^2+1)^{n-1}}\right]_a^b - \frac{1}{2(1-n)}\cdot[\; I_{n-1} \;]_a^b \;+\; [\; I_n \;]_a^b.$$

Durch Umstellen der Gleichung erhalten wir

$$[\; I_n \;]_a^b \quad = \quad \frac{2n-3}{2n-2}\cdot[\; I_{n-1} \;]_a^b \;+\; \left[\frac{x}{(x^2+1)^{n-1}}\right]_a^b.$$

Für das unbestimmte Integral gilt also die Gleichung

$$I_n \quad = \quad \frac{2n-3}{2n-2}\cdot I_{n-1} \;+\; \frac{1}{2n-2}\cdot\frac{x}{(x^2+1)^{n-1}}.$$

Damit erhalten wir beispielsweise

$$I_2 \quad = \quad \tfrac{1}{2}\arctan(x) \,+\, \frac{x}{2x^2+2} \quad \text{und} \quad I_3 \quad = \quad \tfrac{1}{8}\cdot\frac{3x^3+5x}{(x^2+1)^2} \,+\, \tfrac{3}{8}\arctan(x).$$

Wir führen nun die Integrale vom Typ (iv) auf die unbestimmten Integrale I_n zurück. Für alle $c \in \mathbb{R}^+$ und alle $a, b \in \mathbb{R}$ gilt jeweils

$$\int_a^b \frac{dx}{(x^2+c)^n} \quad = \quad \frac{1}{c^n}\int_a^b \frac{dx}{\left(\left(\frac{x}{\sqrt{c}}\right)^2+1\right)^n} \quad =$$

$$c^{\frac{1}{2}-n}\int_a^b \frac{\frac{1}{\sqrt{c}}\,dx}{\left(\left(\frac{x}{\sqrt{c}}\right)^2+1\right)} \quad = \quad c^{\frac{1}{2}-n}\int_{a/\sqrt{c}}^{b/\sqrt{c}} \frac{dt}{(t^2+1)^n} \quad = \quad c^{\frac{1}{2}-n}\left[\, I_n\left(\frac{x}{\sqrt{c}}\right)\right]_a^b.$$

Ähnlich wie bei den Integralen vom Typ (iii) gilt für die Integrale vom Typ (iv) eine Beziehung der Form

$$\int_a^b \frac{cx+d}{(x^2+px+q)^n}\,dx \quad = \quad \left[\frac{1}{2c(1-n)(x^2+px+q)^{n-1}}\right]_a^b \,+\, (d-\tfrac{1}{2}pc)\cdot\int_a^b \frac{dx}{(x^2+px+q)^n}.$$

Da der Rechenweg im Vergleich zu den Integralen vom Typ (iii) nahezu identisch ist, verzichten wir auf die Ausführung. Das Integral auf der rechten der Gleichung kann nun durch das unbestimmte Integral I_n ausgedrückt werden, wie die folgende Rechnung zeigt, bei der $c = q - \frac{1}{4}p^2$ gesetzt wird.

$$\int_a^b \frac{dx}{(x^2+px+q)^n}\,dx \quad = \quad \int_{a+\frac{1}{2}p}^{b+\frac{1}{2}p} \frac{dt}{(t^2+c)^n} \quad = \quad \left[c^{\frac{1}{2}-n}\cdot I_n\left(\frac{x}{\sqrt{c}}\right)\right]_{a+\frac{1}{2}p}^{b+\frac{1}{2}p} \quad =$$

$$\left[(q-\tfrac{1}{4}p^2)^{\frac{1}{2}-n}\cdot I_n\left(\frac{2x}{\sqrt{4q-p^2}}\right)\right]_{a+\frac{1}{2}p}^{b+\frac{1}{2}p} \quad = \quad \left[(q-\tfrac{1}{4}p^2)^{\frac{1}{2}-n}\cdot I_n\left(\frac{2x+p}{\sqrt{4q-p^2}}\right)\right]_a^b$$

Eine Anwendung dieser Formel liefert beispielsweise das unbestimmte Integral

$$\int \frac{dx}{(x^2+7x+17)^3} \quad = \quad (\tfrac{4}{19})^{5/2}\cdot I_3\left(\frac{2x+7}{\sqrt{19}}\right) \quad = \quad \frac{12x^3+126x^2+536x+847}{722\cdot(x^2+7x+17)^2} \,+\, \tfrac{12}{6859}\sqrt{19}\cdot\arctan\left(\frac{2x+7}{\sqrt{19}}\right).$$

Literaturverzeichnis

[Ap] J. Appell, *Analysis in Beispielen und Gegenbeispielen*. Springer-Verlag Heidelberg 2009.

[Fo] O. Forster, *Analysis 1*. Vieweg studium Grundkurs Mathematik, Braunschweig 1983.

[He] H. Heuser, *Lehrbuch der Analysis, Teil 1*. Teubner-Verlag Stuttgart 1991.

[Hi] S. Hildebrandt, *Analysis 1*. Springer-Verlag Heidelberg 2002.

[Ko] K. Königsberger, *Analysis 1*. Springer-Verlag Berlin 2001.